

**ΕΘΝΙΚΟ & ΚΑΠΟΔΙΣΤΡΙΑΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**

Πρόγραμμα Μεταπτυχιακών Σπουδών στη Βιοπληροφορική
Τμήμα Βιολογίας



**Γεωμετρική Μοντελοποίηση και Εντοπισμός του Διαμεμβρανικού
Τμήματος σε Διαμεμβρανικές Πρωτεΐνες τύπου β-Βαρελίου**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Ιωάννη Κ. Βαλαβάνη
(Ηλεκτρολόγου Μηχ. και Μηχ. Υπολογιστών Ε.Μ.Π.)

Επίβλεψη: Ιωάννης Εμίρης
Αν. Καθηγητής Ε.Κ.Π.Α.

Αθήνα, Φεβρουάριος 2006

ΕΘΝΙΚΟ & ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΒΙΟΛΟΓΙΑΣ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ στη ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ

**Γεωμετρική Μοντελοποίηση και Εντοπισμός του Διαμεμβρανικού
Τμήματος σε Διαμεμβρανικές Πρωτεΐνες τύπου β-Βαρελίου**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Ιωάννη Κ. Βαλαβάνη
(Ηλεκτρολόγου Μηχ. και Μηχ. Υπολογιστών Ε.Μ.Π.)

Επίβλεψη: Ιωάννης Εμίρης
Αν. Καθηγητής Ε.Κ.Π.Α.

Τριμελής Επιτροπή:

Ι. ΕΜΙΡΗΣ
Αν. Καθηγητής Ε.Κ.Π.Α.

Σ. ΧΑΜΟΔΡΑΚΑΣ
Καθηγητής Ε.Κ.Π.Α.

Η. ΗΛΙΟΠΟΥΛΟΣ
Αν. Καθηγητής Γ.Π.Α

Αθήνα, Φεβρουάριος 2006

Πρόλογος

Η παρούσα διπλωματική εργασία εκπονήθηκε στα πλαίσια του Μεταπτυχιακού Προγράμματος στη Βιοπληροφορική που οργανώνει το τμήμα της Βιολογίας του Εθνικού & Καποδιστριακού Πανεπιστημίου Αθηνών. Η εκπόνηση της διπλωματικής εργασίας έγινε υπό την επίβλεψη του κ. Ιωάννη Εμίρη, αναπληρωτή καθηγητή του τμήματος Πληροφορικής και Τηλεπικοινωνιών του Ε.Κ.Π.Α., τον οποίο ευχαριστώ θερμά για την καθοδήγηση και τις συμβουλές του. Επίσης, θα ήθελα να ευχαριστήσω τα μέλη της τριμελούς επιτροπής, κ. Σ. Χαμόδρακα, καθηγητή του τμήματος Βιολογίας του Ε.Κ.Π.Α., και κ. Η. Ηλιόπουλο, αναπληρωτή καθηγητή του τμήματος Γεωπονικής Βιοτεχνολογίας του Γ.Π.Α.. Θα ήθελα, επίσης, να ευχαριστήσω θερμά το διδάκτορα Παντελή Μπάγκο για τις συμβουλές του κατά τη διάρκεια εκπόνησης της εργασίας. Ιδιαίτερη ευγνωμοσύνη εκφράζω, τέλος, προς την οικογένεια μου, η οποία με στήριξε σε όλη τη διάρκεια των σπουδών μου.

Θέμα της εργασίας αποτέλεσε η γεωμετρική μοντελοποίηση των διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού και ο εντοπισμός του διαμεμβρανικού τμήματος σε πρωτεΐνες αυτού του τύπου.

Στο 1^ο κεφάλαιο γίνεται μια εισαγωγή στις βιολογικές μεμβράνες, ενώ στο 2^ο κεφάλαιο περιγράφονται οι διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού. Στο 3^ο κεφάλαιο γίνεται μια βιβλιογραφική ανασκόπηση σχετικών ερευνητικών εργασιών και περιγράφεται σύντομα η παρούσα προσέγγιση. Στο 4^ο και 5^ο κεφάλαιο παρουσιάζονται οι αλγόριθμοι που αναπτύχθηκαν για τη γεωμετρική μοντελοποίηση και την εύρεση του διαμεμβρανικού τμήματος αντίστοιχα, ενώ στο 6^ο κεφάλαιο παρουσιάζονται και σχολιάζονται τα αποτελέσματα εφαρμογής των αλγορίθμων σε διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού. Η αντίστοιχη διαδικτυακή εφαρμογή που αναπτύχθηκε στα πλαίσια της εργασίας παρουσιάζεται στο 7^ο κεφάλαιο, ενώ στο 8^ο κεφάλαιο παρατίθενται τα συμπεράσματα της εργασίας. Ακολουθούν η Βιβλιογραφία και το Παράρτημα.

Περίληψη

Στόχος της παρούσης διπλωματικής εργασίας αποτελεί η γεωμετρική μοντελοποίηση και ο εντοπισμός του διαμεμβρανικού τμήματος σε διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού δεδομένης της τριτοταγούς δομής τους. Αρχικά, υλοποιήθηκε υβριδικός αλγόριθμος, που κάνοντας χρήση μεθόδου βελτιστοποίησης ελαχίστων τετραγώνων και γενετικού αλγόριθμου, υπολογίζει τον άξονα και την ακτίνα κυλίνδρου που μοντελοποιεί με τον καλύτερο δυνατό τρόπο το β-βαρέλι. Δεδομένου του άξονα και της ακτίνας του β-βαρελιού, υπολογίζεται η κλίση του άξονα προς του β-κλώνους καθώς και ο Shear Number, χαρακτηριστικά που χρησιμοποιούνται για τη γεωμετρική περιγραφή των β-βαρελιών. Σε επόμενο βήμα, αναπτύχθηκε αλγόριθμος για την εύρεση του διαμεμβρανικού κομματιού των πρωτεϊνών. Για το σκοπό αυτό, σύμφωνα και με εργασία του ερευνητή Wimley, χρησιμοποιήθηκε το προφίλ των εξωτερικών αμινοξέων του β-βαρελιού σε αλληλεπικαλυπτόμενα παράθυρα μήκους 5Å κατά μήκος του άξονα του β-βαρελιού. Η περισσότερο υδρόφοβη ζώνη, που χαρακτηρίζεται από έντονη παρουσία φορτισμένων αμινοξέων στο εξωτερικό της και από έντονη παρουσία αρωματικών αμινοξέων σε περιοχές γύρω από την αρχή και το τέλος της, θεωρήθηκε ότι αποτελεί τη διαμεμβρανική ζώνη της διαμεμβρανικής πρωτεΐνης. Τα αποτελέσματα εύρεσης του διαμεμβρανικού κομματιού συγκρίθηκαν με αυτά της ομάδας του ερευνητή Tusnadi και προέκυψαν αρκετά όμοια. Με βάση τα αποτελέσματα εύρεσης του διαμεμβρανικού κομματιού στο μη πλεοναστικό σύνολο των 22 πρωτεϊνών που χρησιμοποιήθηκε, έγινε μια συνολική μελέτη όσον αφορά την υδροφοβικότητα των αμινοξέων καθώς και την εμφάνιση συγκεκριμένων αμινοξέων ή ομάδων αμινοξέων (π.χ. αρωματικών ή φορτισμένων) κατά μήκος του άξονα των β-βαρελιών. Τα αποτελέσματα είναι συγκρίσιμα με αυτά παρόμοιας εργασίας του Wimley και δείχνουν αποτελέσματα που ήταν αναμενόμενα (π.χ. έντονη παρουσία αρωματικών αμινοξέων στις διεπιφάνειες της μεμβράνης και έντονη παρουσία υδρόφοβων αμινοξέων ως εξωτερικά του β-βαρελιού στο εσωτερικό της μεμβράνης). Τέλος, κατασκευάστηκε κατάλληλο διαδικτυακό εργαλείο, το TbB-Tool, που ενσωματώνει τους αλγόριθμους που αναπτύχθηκαν για τη γεωμετρική μοντελοποίηση και τον εντοπισμό του διαμεμβρανικού τμήματος σε διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού.

Περιεχόμενα

Πρόλογος	3
Περίληψη	4
Περιεχόμενα.....	5
1. Βιολογικές Μεμβράνες	6
1.1 Εισαγωγή.....	6
1.2 Μοριακή οργάνωση βιολογικών μεμβρανών	7
1.3 Δομή βιολογικών μεμβρανών	13
2. Διαμεμβρανικές Πρωτεΐνες τύπου β-Βαρελιού	17
2.1 Γενικά.....	17
2.2 Πού βρίσκονται οι διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού.....	18
2.3 Η δομή των διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού	29
3. Ανασκόπηση Σχετικών Εργασιών – Παρούσα Προσέγγιση	39
3.1 Ανασκόπηση Σχετικών Εργασιών	39
3.2. Παρούσα Προσέγγιση.....	41
4. Γεωμετρική Μοντελοποίηση του β-Βαρελιού.....	43
4.1 Αλγόριθμος για την εύρεση του άξονα και ακτίνας κυλίνδρου δεδομένου ενός αριθμού σημείων που ανήκουν σε αυτόν.....	43
4.2 Εύρεση του άξονα του διαμεμβρανικού β-βαρελιού με χρήση του αλγορίθμου εύρεσης άξονα και ακτίνας κυλίνδρου και Γενετικού Αλγορίθμου.	49
4.3 Υπολογισμός της μέσης κλίσης άξονα και β-κλώνων και του Shear Number του β-βαρελιού	57
5. Εύρεση του διαμεμβρανικού τμήματος διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού.....	59
5.1 Εισαγωγή.....	59
5.2 Προσέγγιση του Wimley	60
5.3 Αλγόριθμος για την εύρεση του διαμεμβρανικού τμήματος	65
6. Αποτελέσματα - Συζήτηση	73
6.1 Γεωμετρική Μοντελοποίηση	73
6.2 Εύρεση του διαμεμβρανικού τμήματος	74
6.3 Μέση υδροφοβικότητα και κατανομή των αμινοξέων κατά μήκος του άξονα του β-βαρελιού.....	89
7. Διαδικτυακή Εφαρμογή	98
8. Συμπεράσματα	102
Βιβλιογραφία	105
Παράρτημα	110
Α. Αλγόριθμοι που δεν περιγράφονται στο κύριο σώμα της εργασίας	110
Β. Εργαλεία που χρησιμοποιήθηκαν	117
Γ. Διαμεμβρανικά αμινοξέα των 22 πρωτεϊνών που μελετήθηκαν	117

1. Βιολογικές Μembrάνες

1.1 Εισαγωγή

«Η κυτταρική μεμβράνη διαιρεί το σύμπαν σε δυο μέρη, το «μέσα» και το «έξω». Έτσι, η μεμβράνη γίνεται το πιο σπουδαίο οργανίδιο του κυττάρου αφού αποτελεί το σημείο συνάντησης δύο κόσμων». Τα λόγια αυτά του A. Szent-Gyorgi δείχνουν τη σημασία που έχει η κυτταρική μεμβράνη και γενικότερα οι βιολογικές μεμβράνες [1] για την οργάνωση των κυττάρων και της ζωής. Η κυτταρική ή πλασματική μεμβράνη περιβάλλει το κύτταρο και το διαχωρίζει από το περιβάλλον του, αποτελώντας ταυτόχρονα ένα εκλεκτικό φίλτρο για την είσοδο και την έξοδο ιόντων και μορίων. Οι βιολογικές μεμβράνες περιλαμβάνουν και τις ενδοκυτταρικές μεμβράνες (π.χ. μεμβράνες μιτοχονδρίων ή χλωροπλαστών) οι οποίες δημιουργούν χωριστά διαμερίσματα μέσα στο κύτταρο, διαχωρίζοντας έτσι τις διάφορες βιοχημικές διεργασίες, ενώ μεμβράνες εμφανίζονται και σε προκαρυωτικούς οργανισμούς (π.χ. βακτήρια) για να τους ξεχωρίζουν από το εξωτερικό περιβάλλον.

Οι βιολογικές μεμβράνες που παρατηρούνται σε ευκαρυωτικούς και προκαρυωτικούς οργανισμούς έχουν κοινή δομή και αποτελούνται από λιπίδια και πρωτεΐνες που ενώνονται με μη ομοιοπολικές αλληλεπιδράσεις. Η ύπαρξη μιας λιπιδικής διπλοστιβάδας αποτελεί τη βασική δομή της μεμβράνης και ταυτόχρονα φράγμα για την είσοδο υδατοδιαλυτών μορίων. Τα πρωτεϊνικά μόρια βρίσκονται μέσα στο στρώμα των λιπιδίων και επιτελούν όλες τις λειτουργίες της μεμβράνης: άλλα μεταφέρουν εκλεκτικά μόρια μέσα ή έξω από το περιβάλλον που περικλείουν, άλλα καταλύουν ενζυμικές αντιδράσεις που συνδέονται με τη μεμβράνη και άλλα, τέλος, συνδέονται με στοιχεία του κυτταρικού σκελετού και της εξωκυττάριας ουσίας.

Όλες οι κυτταρικές μεμβράνες είναι δυναμικές, ρευστές δομές, μέσα στις οποίες τόσο τα λιπίδια όσο και οι πρωτεΐνες έχουν την ικανότητα μετακίνησης στο επίπεδο της μεμβράνης. Οι μεμβράνες έχουν διαφορετική σύσταση σε λιπίδια και πρωτεΐνες και αποτελούν ασύμμετρες δομές. Στις επόμενες παραγράφους περιγράφονται η μοριακή οργάνωση και η δομή των βιολογικών μεμβρανών.

1.2 Μοριακή οργάνωση βιολογικών μεμβρανών

1.2.1 Γενικά

Η χημική ανάλυση των μεμβρανών απαιτεί την απομόνωση τους από τα κύτταρα που περιβάλλουν. Κυριότερη πηγή για τη λήψη μεγάλων ποσοτήτων πλασματικών μεμβρανών αποτελούν τα ανθρώπινα ερυθροκύτταρα. Με την επίδραση υποτονικών διαλυμάτων τα κύτταρα αρχικά διογκώνονται και κατόπιν σπάζουν χάνοντας την αιμοσφαιρίνη που περιείχαν. Η μεμβράνη που απομένει ονομάζεται «φάντασμα ερυθροκυττάρου». Φαντάσματα ερυθροκυττάρων που προκύπτουν μετά από ήπια αιμόλυση, έχουν τη δυνατότητα να επαναλάβουν τις λειτουργίες της διαπερατότητας. Πλήρης απομάκρυνση της αιμοσφαιρίνης, μετά από δραστική αιμόλυση των ερυθροκυττάρων, δημιουργεί «λευκά φαντάσματα» τα οποία μπορούν να χρησιμοποιηθούν για βιοχημικές αλλά όχι για φυσιολογικές μελέτες. Πάντως ακόμη και στα λευκά φαντάσματα παρατηρούνται συστατικά π.χ. μικροϊνίδια, τα οποία δεν αποτελούν δομικά συστατικά της μεμβράνης των ερυθροκυττάρων.

Οι βιολογικές μεμβράνες αποτελούνται κυρίως από πρωτεΐνες, λιπίδια και μια μικρή αναλογία υδατανθράκων που συνδέονται με πρωτεΐνες (γλυκό-πρωτεΐνες) ή λιπίδια (γλυκολιπίδια). Η αναλογία λιπιδίων / πρωτεϊνών στις διάφορες μεμβράνες, διαφέρει ανάλογα με τον κυτταρικό τύπο. Π.χ. στη μυελίνη υπάρχουν, κατά βάρος, 90% λιπίδια και 10% πρωτεΐνες ενώ στην εσωτερική μεμβράνη του μιτοχονδρίου 25% λιπίδια και 75% πρωτεΐνες. Οι περισσότερες μεμβράνες έχουν αναλογία λιπιδίων / πρωτεϊνών που βρίσκεται ανάμεσα στις δύο αυτές ακραίες περιπτώσεις. Οι υδατάνθρακες αποτελούν το 2-10% του συνολικού βάρους της μεμβράνης και βρίσκονται τοποθετημένοι στην εξωτερική πλευρά της.

1.2.2 Μεμβρανικά Λιπίδια

Στα περισσότερα ζωικά κύτταρα, τα λιπίδια αποτελούν το 50% περίπου της μάζας της πλασματικής μεμβράνης τους. Οι κυριότεροι τύποι λιπιδίων που εμφανίζονται στις μεμβράνες είναι τα φωσφολιπίδια (που αποτελούν το πιο άφθονο συστατικό), η χοληστερόλη και τα γλυκολιπίδια. Όλα είναι αμφίτροπα μόρια, δηλ. έχουν ένα υδρόφιλο κι ένα υδρόφοβο άκρο.

Τα κυριότερα φωσφολιπίδια της πλασματικής μεμβράνης είναι τα φωσφογλυκερίδια. Τα φωσφογλυκερίδια έχουν δύο αλειφατικές αλυσίδες λιπαρών οξέων ενωμένες με τα δύο άτομα άνθρακα της γλυκερόλης ενώ στο τρίτο άτομο άνθρακα ενώνεται μια φωσφορική ομάδα. Στη φωσφορική ομάδα συνδέονται αλκοόλες, π.χ. χολίνη, αιθανολαμίνη, ινοσιτόλη κτλ. ή τα αμινοξέα σερίνη ή θρεονίνη. Το πιο σπουδαίο χαρακτηριστικό των φωσφογλυκεριδίων από την άποψη της μεμβρανικής δομής, είναι ο αμφίτροπος χαρακτήρας τους. Οι αλειφατικές αλυσίδες των λιπαρών οξέων είναι μη πολικές και υδρόφοβες, ενώ το φωσφορικό άκρο είναι πολικό και υδρόφιλο. Μια άλλη κατηγορία φωσφολιπιδίων που βρίσκονται συνήθως στην πλασματική μεμβράνη νευρικών και μυϊκών κυττάρων είναι τα πλασμαλογόνα ή αιθερολιπίδια. Τα αιθερολιπίδια έχουν την ίδια δομή με τα φωσφογλυκερίδια εκτός από το δεσμό της αλειφατικής αλυσίδας στον άνθρακα 1, όπου αντί για εστερικό δεσμό υπάρχει αιθερικός. Στη φωσφορική ομάδα συνήθως συνδέεται αιθανολαμίνη. Στα φωσφολιπίδια ανήκουν επίσης τα σφιγγολιπίδια, στα οποία βασικό στοιχείο είναι η αμινική αλκοόλη σφιγγοσίνη. Στη φωσφορική ομάδα των σφιγγολιπιδίων ενώνεται χολίνη, μονοσακχαρίτης ή πολυσακχαρίτης οπότε τα σφιγγολιπίδια ονομάζονται αντίστοιχα σφιγγομυελίνη, σερεβροζίδια και γαγγλιοζίδια.

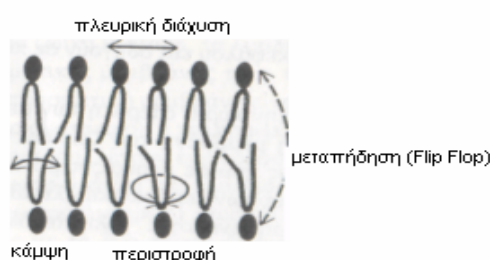
Σφιγγολιπίδια με υδατανθρακικές ομάδες αποτελούν τα πιο κοινά γλυκολιπίδια της πλασματικής μεμβράνης στα ζωικά και τα φυτικά κύτταρα. Στα κύτταρα θηλαστικών όλα σχεδόν τα γλυκολιπίδια είναι σφιγγολιπίδια και βρίσκονται σχεδόν αποκλειστικά στην πλασματική μεμβράνη. Τα γαγγλιοζίδια είναι τα πιο σύνθετα γλυκολιπίδια από δομική άποψη, περιέχουν μια ή περισσότερες ομάδες σιαλικού οξέος και είναι αρνητικά φορτισμένα. Τα γαγγλιοζίδια αποτελούν μεγάλο τμήμα της πλασματικής μεμβράνης των νευρώνων. Οι λειτουργίες που επιτελούν τα γαγγλιοζίδια δεν είναι γνωστές, αν και ορισμένα απ' αυτά αποτελούν υποδοχείς για ιούς και τοξίνες.

Η χοληστερόλη αποτελεί την κύρια στερόλη των μεμβρανών των ζωικών κυττάρων. Η πλασματική μεμβράνη των ζωικών κυττάρων περιέχει σημαντικά ποσά χοληστερόλης που σε μερικές περιπτώσεις φθάνουν και το 30% κατά βάρος, του συνόλου των μεμβρανικών λιπιδίων. Η πλασματική μεμβράνη περιέχει μεγαλύτερη αναλογία χοληστερόλης / φωσφολιπιδίων (0,3-1,2) σε σύγκριση με τις ενδοκυτταρικές μεμβράνες των μιτοχονδρίων (0,03-0,09) και του ενδοπλασματικού

δικτύου (0,03-0,08). Στις ενδοκυττάρειες μεμβράνες, επίσης, οι αλειφατικές αλυσίδες των φωσφολιπιδίων έχουν περισσότερους ακόρεστους δεσμούς σε σχέση με τα φωσφολιπίδια της πλασματικής μεμβράνης. Στα φυτά, από την άλλη πλευρά, υπάρχουν πολύ μικρά ποσά χοληστερόλης, ενώ στα βακτήρια και τα φύκη δεν υπάρχει σχεδόν καθόλου χοληστερόλη. Η αρχική θέση σύνθεσης των στερολών και των φωσφολιπιδίων είναι το ενδοπλασματικό δίκτυο, απ' όπου μεταφέρονται με κυστίδια στην πλασματική μεμβράνη. Ο διαχωρισμός των στερολών από τα φωσφολιπίδια γίνεται σε κάποιο αρχικό στάδιο της μεμβρανικής βιογένεσης.

Ο αμφίτροπος χαρακτήρας των λιπιδίων παίζει σπουδαίο ρόλο στη δημιουργία των βιολογικών μεμβρανών. Όταν αμφίτροπα μόρια βρεθούν σε υδατικό περιβάλλον, έχουν την τάση να σχηματίζουν συσσωματώματα «κρύβοντας» τις υδρόφοβες «ουρές» και εκθέτοντας στο H₂O τις υδρόφιλες ομάδες. Αυτό επιτυγχάνεται είτε με το σχηματισμό μυκηλίων είτε με τη δημιουργία διπλοστοιβάδας. Η λιπιδιακή διπλοστοιβάδα αποτελεί τη βασική δομή της μεμβράνης. Αυτό φαίνεται χαρακτηριστικά κατά τον ψυκτοτεμαχισμό όπου η θραύση της μεμβράνης γίνεται στο μέσο της διπλοστοιβάδας διαχωρίζοντας έτσι τις δύο λιπιδιακές μονοστοιβάδες.

Η διπλοστοιβάδα των λιπιδίων αποτελεί μια υγρή φάση στην οποία τα λιπιδιακά μόρια μπορούν να κινούνται ελεύθερα στη δική τους στοιβάδα με πλευρική διάχυση. Ένα λιπιδιακό μόριο μπορεί ακόμη να περιστραφεί γύρω από τον άξονα του, να καμφθεί, αλλά είναι πολύ δύσκολο να μεταπηδήσει από τη μια στοιβάδα στην άλλη (Εικόνα. 1.1).



Εικόνα 1.1 Διάφοροι τύποι κίνησης που μπορούν να κάνουν τα λιπίδια μέσα στη λιπιδιακή διπλοστοιβάδα

Έτσι, η λιπιδιακή διπλοστοιβάδα δείχνει δυναμικότητα και η ρευστότητα της μπορεί να μεταβληθεί από την επίδραση διαφόρων χημικών και φυσικών

παραγόντων. Οι χημικοί παράγοντες οι οποίοι επηρεάζουν τη ρευστότητα της λιπιδιακής διπλοστοιβάδας είναι το μήκος της αλειφατικής αλυσίδας, ο αριθμός των διπλών δεσμών, το ποσό της χοληστερόλης που υπάρχει στη διπλοστοιβάδα και το ποσό των μεμβρανικών πρωτεϊνών. Οι φυσικοί παράγοντες είναι η θερμοκρασία, η πίεση, το pH, και το μεμβρανικό δυναμικό και τα ιόντα ασβεστίου.

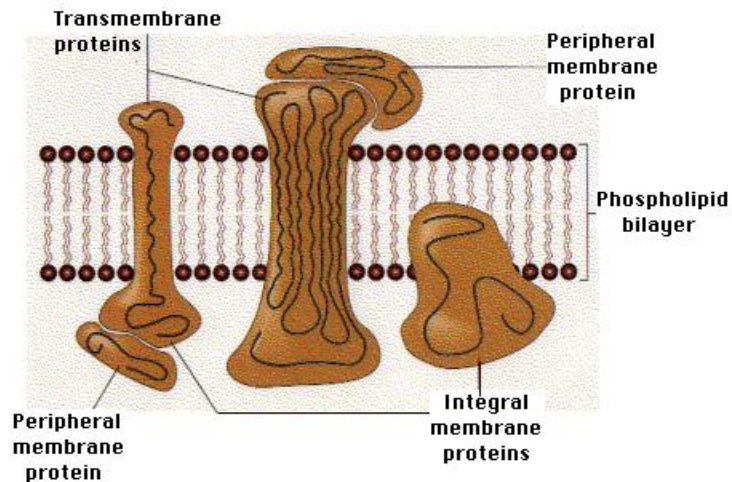
Αν και η ρευστότητα είναι χαρακτηριστικό γνώρισμα της διπλοστοιβάδας των λιπιδίων, υπάρχουν πειραματικά δεδομένα που δείχνουν ότι τα λιπίδια, όπως και οι πρωτεΐνες, είναι οργανωμένα κατά περιοχές. Η ετερογένεια αυτή στην κατανομή των λιπιδίων έχει λειτουργική και δομική σημασία για την μεμβράνη. Αν και υπάρχουν αντιρρήσεις, πιστεύεται ότι η οργάνωση αυτή επηρεάζεται άμεσα από τις μεμβρανικές πρωτεΐνες. Όλες οι βιολογικές μεμβράνες που έχουν εξεταστεί μέχρι σήμερα, δείχνουν ετερογένεια στη λιπιδιακή τους σύσταση. Η ετερογένεια αυτή δημιουργείται κατά τη βιογένεση των λιπιδίων, αφού η μετακίνηση λιπιδίων από τη μια στοιβάδα στην άλλη είναι σχεδόν ανύπαρκτη. Η ύπαρξη ασυμμετρίας σχετίζεται πιθανόν με το σωστό προσανατολισμό των μεμβρανικών πρωτεϊνών. Ενδεχομένως οι πρωτεΐνες για να μπορέσουν να δράσουν απαιτούν την ύπαρξη ή / και την αλληλεπίδραση συγκεκριμένων φωσφολιπιδίων. Αυτό ίσως αποτελεί μια λογική εξήγηση τόσο για την ασυμμετρία της λιπιδιακής διπλοστοιβάδας όσο και για τη μεγάλη ποικιλία των φωσφολιπιδίων που υπάρχουν σε αυτήν.

Έχει αναφερθεί από μερικούς ερευνητές ότι στη δομή της πλασματικής μεμβράνης συμμετέχουν και άλλες λιπιδιακές δομές εκτός από τη λιπιδιακή διπλοστοιβάδα. Οι διατάξεις αυτές των λιπιδίων είναι πολύ πιθανόν να διαδραματίζουν δυναμικό ρόλο σε διάφορες μεμβρανικές διεργασίες όπως π.χ. τη μεμβρανική σύντηξη και τη μετακίνηση λιπιδίων από τη μια στοιβάδα στην άλλη.

1.2.3 Μεμβρανικές Πρωτεΐνες

Οι πρωτεΐνες αντιπροσωπεύουν το κυριότερο συστατικό των περισσότερων βιολογικών μεμβρανών, όχι μόνο επειδή συμμετέχουν στη δομή τους αλλά και επειδή εκφράζουν τη βιολογική δράση των μεμβρανικών συστημάτων. Οι πρωτεΐνες είναι μόρια με υδρόφιλο σκελετό και πλευρικές αμινοξικές ακολουθίες. Τα πρωτεϊνικά μόρια είναι πολύ μεγαλύτερα σε σχέση με τα

λιπιδικά μόρια και γι' αυτό είναι αριθμητικά λιγότερα από τα λιπίδια. Οι μεμβρανικές πρωτεΐνες ανάλογα με τη θέση τους σε σχέση με τη λιπιδική διπλοστοιβάδα διακρίνονται σε περιφερειακές (peripheral) και εσωτερικές (integral). Οι περιφερειακές πρωτεΐνες αποχωρίζονται από τη λιπιδική διπλοστοιβάδα, αφήνοντας την άθικτη, υπό την επίδραση διαλυμάτων που σπάζουν ασθενείς ιοντικούς. Μπορούν να βρίσκονται, επίσης, τόσο στην κυτταροπλασματική όσο και στην εξωτερική πλευρά της μεμβράνης. Οι εσωτερικές πρωτεΐνες αποτελούν το 70% περίπου των μεμβρανικών πρωτεϊνών. Συνήθως βρίσκονται «βυθισμένες» μέσα στη διπλοστοιβάδα και για την απομόνωσή τους είναι απαραίτητη η χρήση απορρυπαντικών ή οργανικών διαλυτών που διαλύουν τα λιπίδια. Πολλές εσωτερικές πρωτεΐνες διασχίζουν ολόκληρη τη διπλοστοιβάδα και τα δύο άκρα τους βρίσκονται εκτεθειμένα στο υδατικό περιβάλλον. Οι πρωτεΐνες αυτές ονομάζονται διαμεμβρανικές (transmembrane). Άλλες εσωτερικές πρωτεΐνες εκθέτονται στο υδατικό περιβάλλον από τη μία μόνο πλευρά της λιπιδικής διπλοστοιβάδας. Η ύπαρξη διαμεμβρανικών πρωτεϊνών επιβεβαιώθηκε και με την τεχνική της ψυκτοεξάχνωσης, κατά την οποία οι πρωτεΐνες αυτές διακρίνονται ως μικρά εξογκώματα τα οποία ονομάζονται ενδομεμβρανικά σωματίδια ή κοκκία. Στις διαμεμβρανικές πρωτεΐνες και κυρίως στις διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού θα επανέλθουμε σε επόμενο κεφάλαιο, μια και οι τελευταίες αποτελούν αντικείμενο μελέτης της παρούσης εργασίας. Διάφορα είδη πρωτεϊνών ανάλογα με την θέση τους ως προς τη μεμβράνη φαίνονται στην Εικόνα 1.2.



Εικόνα 1. 2 Διάφοροι τύποι μεμβρανικών πρωτεϊνών ανάλογα με τη θέση τους ως προς τη μεμβράνη

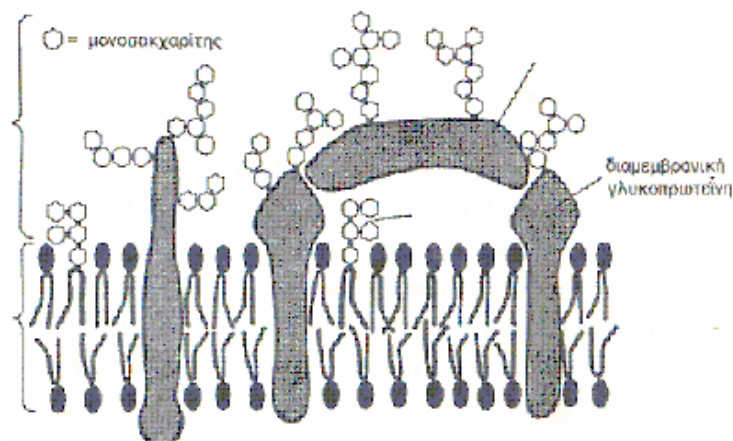
Οι πρωτεΐνες, όπως και τα λιπίδια, μπορούν να κινηθούν πλευρικά μέσα στην μεμβράνη. Η κίνηση αυτή όμως δεν είναι ελεύθερη αλλά περιορίζεται από διάφορους παράγοντες. Στο κορυφαίο τμήμα των επιθηλιακών κυττάρων π.χ. υπάρχουν διαφορετικές πρωτεΐνες απ' ότι στο πλευρικό ή το βασικό τμήμα. Οι πρωτεΐνες των διαφορετικών τμημάτων δεν αναμειγνύονται διότι η πλευρική διάχυση περιορίζεται από τους στενοσυνδέσμους που αποτελούν φραγμό διάχυσης. Τα κύτταρα συνεπώς δείχνουν πολικότητα, δηλ. διαφοροποίηση δομής και λειτουργίας ανάμεσα στα διάφορα τμήματα τους. Η ελεύθερη κίνηση των μεμβρανικών πρωτεϊνών μπορεί επίσης να περιοριστεί από εξωτερικές αλληλεπιδράσεις, όπως αυτό συμβαίνει κατά το σχηματισμό στοιχείων του συνδετικού συμπλέγματος. Ακόμη, η κίνηση και η κατανομή των μεμβρανικών πρωτεϊνών ελέγχεται και από τις αλληλεπιδράσεις κυτταροπλασματικών πρωτεϊνών, όπως π.χ. της ακτίνης.

1.2.4 Μεμβρανικοί Υδατάνθρακες

Οι υδατάνθρακες αποτελούν το τρίτο κύριο στοιχείο των μεμβρανών. Βρίσκονται πάντα στην εξωτερική πλευρά της μεμβράνης ενωμένοι είτε με πρωτεΐνες (γλυκοπρωτεΐνες) είτε με λιπίδια (γλυκολιπίδια). Ενώ οι περισσότερες μεμβρανικές πρωτεΐνες είναι ενωμένες με υδατάνθρακες, μόνο ένα στα δέκα λιπίδια ενώνεται με τα μόρια αυτά.

Αν και στη φύση υπάρχουν 100 περίπου μονοσακχαρίτες, μόνο εννέα απ' αυτούς αποτελούν συστατικά των γλυκοπρωτεϊνών και των γλυκολιπιδίων με κυριότερα τη γαλακτόζη, τη μαννόζη, τη γαλακτοζαμίνη, τη φουκόζη, τη γλυκόζη και το

σιαλικό οξύ. Οι μεμβρανοϊκοί υδατάνθρακες αποτελούν ένα κάλυμμα στο εξωτερικό τμήμα της πλασματικής μεμβράνης που ονομάζεται γλυκοκάλυκας (Εικόνα 1.3).



Εικόνα 1.3 Μεμβρανοϊκοί υδατάνθρακες στο εξωτερικό τμήμα μεμβράνης

Οι λειτουργίες που επιτελούν οι ολιγοσακχαριτικές ομάδες δεν είναι γνωστές με βεβαιότητα, αλλά πιστεύεται ότι βοηθούν τις διαμεμβρανικές πρωτεΐνες να παραμείνουν στη θέση τους, σταθεροποιώντας συγχρόνως και τη δομή τους. Οι μεμβρανοϊκές γλυκοπρωτεΐνες διαδραματίζουν σπουδαίο ρόλο στη ρύθμιση της κυτταρικής συμπεριφοράς. Έτσι θεωρείται ότι συμμετέχουν στην κυτταρική αύξηση, στη διαίρεση, τη διαφοροποίηση, στις κυτταρικές αλληλεπιδράσεις και την αναγνώριση διαφόρων μακρομορίων κτλ.

Συνοψίζοντας μπορούμε να πούμε σε αυτό το σημείο ότι όλες οι βιολογικές μεμβράνες έχουν αρκετά κοινά χαρακτηριστικά, όπως:

- α) Οι μεμβράνες διαχωρίζουν διαμερίσματα με διαφορετική σύσταση.
- β) Οι μεμβράνες αποτελούνται κυρίως από φωσφολιπίδια και πρωτεΐνες σε κατά βάρος αναλογία από 1: 4 μέχρι 4: 1.
- γ) Τα λιπίδια είναι αμφίτροπα μόρια και σε υδατικό περιβάλλον δημιουργούν αυτόματα λιπιδιακές διπλοστοιβάδες.
- δ) Οι μεμβράνες σχηματίζονται από μη ομοιοπολικές αλληλεπιδράσεις ανάμεσα στα δομικά συστατικά τους.
- ε) Οι μεμβράνες έχουν ρευστότητα και είναι ασύμμετρες.

1.3 Δομή βιολογικών μεμβρανών

Οι μελέτες του Overton, στο τέλος του περασμένου αιώνα, έδειξαν ότι μικρά μη πολικά μόρια μπορούν να μπουν μέσα στα κύτταρα πολύ πιο εύκολα απ' ότι πολικά μόρια ίδιου μεγέθους. Έτσι, το 1902 πρότεινε ότι η μεμβράνη που

καλύπτει εξωτερικά τα κύτταρα αποτελείται από ένα λεπτό στρώμα λιπιδίων. Το 1925 οι Gorter και Grendel βρήκαν ότι στη μεμβράνη των ερυθροκυττάρων υπάρχουν αρκετά λιπίδια τα οποία μπορούν να σχηματίσουν διπλοστοιβάδα. Με τις παρατηρήσεις αυτές άνοιξε ο δρόμος για τη μελέτη της μεμβρανικής δομής, ενώ αργότερα παρουσιάστηκαν διάφορα μοντέλα.

1.3.1 Το μοντέλο των Danielli-Davson

Το 1935 οι Danielli και Davson στηριζόμενοι στην εργασία των Gorter και Grendel αλλά και στα δικά τους πειραματικά δεδομένα παρουσίασαν το πρώτο μοντέλο για τη δομή της πλασματικής μεμβράνης. Σύμφωνα με αυτό, η μεμβράνη αποτελείται από ένα διμοριακό στρώμα λιπιδίων, τη λιπιδιακή διπλοστοιβάδα, στο οποίο τα πολικά τμήματα των λιπιδίων βρίσκονται προς την υδάτινη φάση, ενώ οι αλειφατικές αλυσίδες των λιπαρών οξέων συνδέονται στο κέντρο της διπλοστοιβάδας. Επειδή η επιφανειακή τάση στις πλασματικές μεμβράνες ζωντανών κυττάρων είναι αρκετά μικρότερη από τις τιμές επιφανειακής τάσης τεχνητών λιπιδιακών μεμβρανών, οι Danielli και Davson πρότειναν ότι η μεμβράνη καλύπτεται και απ' τις δύο πλευρές της από σφαιρικές πρωτεΐνες. Το μοντέλο αυτό τροποποιήθηκε αργότερα από τους Stein και Danielli, για να συμπεριλάβε και τα αποτελέσματα νεότερων ερευνών. Έτσι, το νέο αυτό μοντέλο που διατυπώθηκε το 1962 δείχνει τη διπλοστοιβάδα των λιπιδίων να διασχίζεται από πόρους μικρών διαστάσεων, μέσα από τους οποίους μπορούν να περάσουν πολικά μόρια. Αν και το μοντέλο του Danielli και Davson έχει τροποποιηθεί αρκετές φορές μέχρι σήμερα, η βασική αρχική ιδέα της λιπιδιακής διπλοστοιβάδας ισχύει ακόμη.

1.3.2 Μοντέλο Μονάδας Μεμβράνης

Στη δεκαετία του 1950 το ηλεκτρονικό μικροσκόπιο χρησιμοποιήθηκε για τη μελέτη της λεπτής δομής και των μεμβρανικών συστημάτων. Η χρησιμοποίηση του $KMnO_4$ ως μέσου μονιμοποίησης και των εποξειδικών ως μέσου έγκλεισης, επέτρεψε την πιο λεπτομερειακή παρατήρηση των μεμβρανών. Έτσι, στο ηλεκτρονικό μικροσκόπιο η πλασματική μεμβράνη φαίνεται ως μια τρίστρωμη δομή, σαν μια «σιδηροδρομική γραμμή» όπου δύο ακραίες σκοτεινές ζώνες πάχους περίπου 2,5 nm περιβάλλουν μια φωτεινή κεντρική ζώνη που έχει το ίδιο πάχος. Οι δύο σκοτεινές ζώνες αντιστοιχούν στα πρωτεϊνικά μόρια και η ανοιχτόχρωμη στη λιπιδιακή στοιβάδα. Τη δομή αυτή παρουσιάζουν όχι μόνο οι πλασματικές μεμβράνες αλλά και οι υπόλοιπες βιολογικές μεμβράνες. Έτσι, η δομή που παρουσιάζεται

στο ηλεκτρονικό μικροσκόπιο θεωρείται ότι αποτελεί τη μονάδα όλων των βιολογικών μεμβρανών και γι' αυτό αναφέρεται ως μονάδα μεμβράνης. Στο μοντέλο το οποίο προτάθηκε από τον Robertson οι πρωτεΐνες δεν θεωρούνται σφαιρικές όπως στο μοντέλο των Danielli και Davson αλλά πιστεύεται ότι σχηματίζουν επίπεδες επιφάνειες στα δύο άκρα της μεμβράνης. Το μοντέλο της «μονάδας μεμβράνης» στηρίζεται στο μοντέλο των Danielli και Davson αλλά προσθέτει δύο νέα στοιχεία: αποδεικνύει την «παγκοσμιότητα» της λιπιδιακής διπλοστοιβάδας και εισάγει για πρώτη φορά την έννοια της χημικής ασυμμετρίας των συστατικών της πλασματικής μεμβράνης.

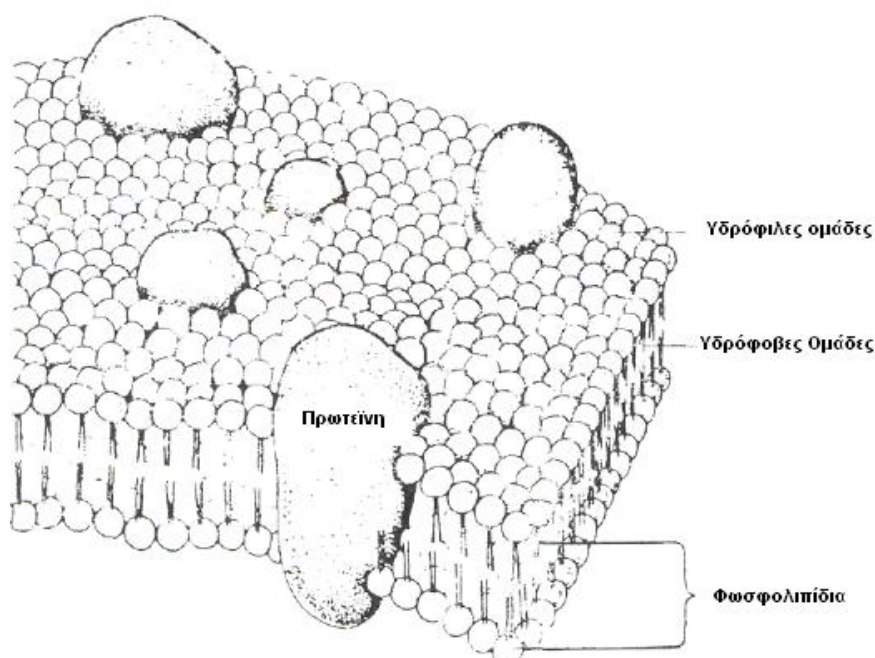
1.3.3 Το μοντέλο του Ρευστού Μωσαϊκού

Η εισαγωγή των τεχνικών του ψυκτοτεμαχισμού και της ψυκτοεξάχνωσης στο τέλος της δεκαετίας του '60 έδειξε ότι οι μεμβράνες σπάζουν στο κέντρο της λιπιδιακής διπλοστοιβάδας. Στις όψεις θράσης παρατηρούνται σωματίδια διαμέτρου 5-12 nm. Η ύπαρξη των ενδομεμβρανικών σωματιδίων και πειραματικά δεδομένα σχετικά με τη ρευστότητα και την τυχαία διανομή των μεμβρανικών πρωτεϊνών, οδήγησαν τους Singer και Nicolson να προτείνουν το 1972 ένα νέο μοντέλο για τη δομή των μεμβρανών, το μοντέλο του ρευστού μωσαϊκού. Τα βασικά σημεία του μοντέλου αυτού, το οποίο διατηρεί το κύριο στοιχείο της λιπιδιακής διπλοστοιβάδας είναι: α) οι πρωτεΐνες δεν βρίσκονται στην επιφάνεια της διπλοστοιβάδας αλλά σχηματίζουν ένα είδος μωσαϊκού καθώς παρεμβάλλονται ανάμεσα στα λιπίδια και β) οι βιολογικές μεμβράνες είναι ημίρευστες δομές στις οποίες τόσο τα λιπίδια όσο και οι πρωτεΐνες μπορούν να μετατοπίζονται πλευρικά μέσα στη λιπιδιακή διπλοστοιβάδα.

Με το μοντέλο του ρευστού μωσαϊκού εξηγείται η ύπαρξη των ενδομεμβρανικών κοκκίων που παρατηρούνται στις όψεις θραύσης με την τεχνική του ψυκτοτεμαχισμού και της ψυκτοεξάχνωσης. Η ρευστότητα των πλασματικών μεμβρανών υποστηρίζεται από αρκετά πειραματικά δεδομένα. Οι φασματοσκοπικές μελέτες και μελέτες πυρηνικού μαγνητικού συντονισμού έχουν δείξει ότι τόσο τα λιπίδια όσο και οι πρωτεΐνες δείχνουν πλευρική διάχυση στη στοιβάδα των λιπιδίων.

Σύμφωνα με το μοντέλο του ρευστού μωσαϊκού (Εικόνα 1.4), οι μεμβρανικές πρωτεΐνες αποτελούν «παγόβουνα» μέσα σε μια «θάλασσα» λιπιδίων. Η ύπαρξη πρωτεϊνικών μορίων ή συσσωματωμάτων μπορεί να ελαττώσει τη ρευστότητα των λιπιδίων που βρίσκονται σε στενή επαφή με τις πρωτεΐνες αυτές. Δημιουργείται έτσι

ένα «κάλυμμα» ακινητοποιημένων λιπιδίων που περιβάλλουν το πρωτεϊνικό μόριο. Λόγω της ειδικής θέσης των φορτισμένων ή μη ακραίων αμινοξέων της πρωτεΐνης, είναι απαραίτητη η αλληλεπίδραση τους σε συγκεκριμένα λιπίδια, έτσι ώστε η πρωτεΐνη να εκδηλώσει τη βιολογική της δράση.



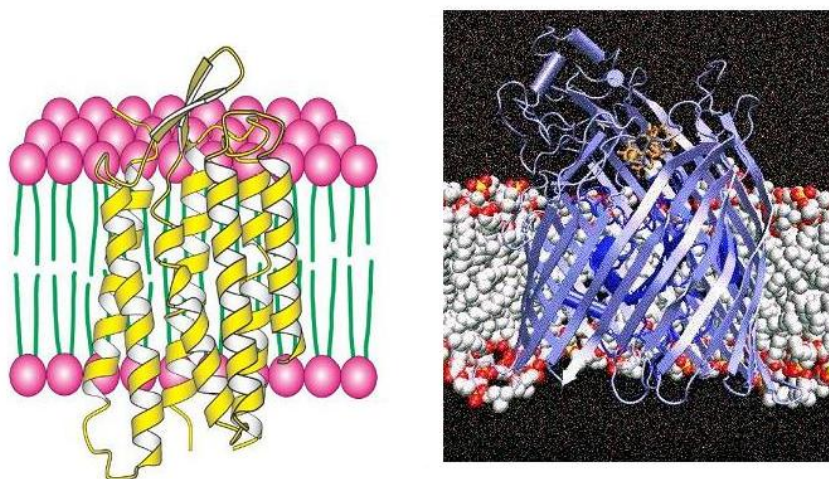
Εικόνα 1. 4 Μοντέλο ρευστού μωσαϊκού

Αν και υπάρχουν αρκετά ισχυρά δεδομένα για την αποδοχή του μοντέλου του ρευστού μωσαϊκού, διατυπώνονται μερικές αμφιβολίες και αντιρρήσεις ως προς την ορθότητα του. Πειραματικά δεδομένα δείχνουν ότι ανάμεσα στις μεμβρανικές πρωτεΐνες αναπτύσσονται μεγάλες αλληλεπιδράσεις κάτι που δεν θα έπρεπε να συμβαίνει αν οι πρωτεΐνες είναι απλώς «βυθισμένες» μέσα στη λιπιδιακή διπλοστοιβάδα. Έτσι, άλλα μοντέλα κάνουν σιγά την εμφάνιση τους, όπως το μοντέλο του υδρόφοβου φραγμού που διατυπώθηκε πρόσφατα από το Robertson. Στο μοντέλο αυτό περιλαμβάνονται όλα τα μέχρι τώρα δεδομένα, στην προσπάθεια διατύπωσης ενός παγκόσμια αποδεκτού μοντέλου για τη δομή των βιολογικών μεμβρανών.

2. Διαμεμβρανικές Πρωτεΐνες τύπου β -Βαρελίου

2.1 Γενικά

Όπως αναφέρθηκε στο προηγούμενο κεφάλαιο, οι διαμεμβρανικές πρωτεΐνες είναι εκείνες οι πρωτεΐνες που διαπερνούν τις βιολογικές μεμβράνες τους από το ένα άκρο τους στο άλλο. Αυτό μπορεί να γίνεται μια ή περισσότερες φορές και οι διαμεμβρανικές πρωτεΐνες μπορεί να είναι δύο τύπων: α) διαμεμβρανικές πρωτεΐνες τύπου α έλικας (το κομμάτι της πρωτεΐνης που διαπερνά τη μεμβράνη είναι δεμάτι α ελίκων, “ α helix bundle”, ή μία έλικα σε λιγότερες περιπτώσεις) β) διαμεμβρανικές πρωτεΐνες τύπου β -βαρελίου (το κομμάτι της πρωτεΐνης που διαπερνά τη μεμβράνη είναι μια κλειστή β -πτυχωτή επιφάνεια που σχηματίζει ένα βαρέλι). Οι τύποι των διαμεμβρανικών πρωτεϊνών φαίνονται σε αντιπροσωπευτικές δομές στην Εικόνα 2.1.



Εικόνα 2. 1 Διαμεμβρανική πρωτεΐνη τύπου α έλικας (αριστερά) και τύπου β -βαρελίου (δεξιά)

Σήμερα είναι γνωστές οι δομές 587 διαμεμβρανικών πρωτεϊνών, από τις οποίες οι 480 είναι του τύπου α έλικας και οι 100 είναι τύπου β -βαρελίου ([2], 02/02/06). Οι διαμεμβρανικές πρωτεΐνες τύπου β -βαρελίου απαντώνται στην εξωτερική μεμβράνη των αρνητικών κατά Gram βακτηρίων, ενώ πιστεύεται ότι υπάρχουν τέτοιες πρωτεΐνες στα μιτοχόνδρια και στους χλωροπλάστες [3,4]. Στο παρόν κεφάλαιο αναφερόμαστε στις διαμεμβρανικές πρωτεΐνες τύπου β -βαρελίου,

οι οποίες αποτελούν το αντικείμενο μελέτης της παρούσης εργασίας, όσον αφορά στους οργανισμούς ή τα ενδοκυτταρικά διαμερίσματα στα οποία απαντώνται αλλά και στη δομή τους.

2.2 Πού βρίσκονται οι διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού

2.2.1 Αρνητικά κατά Gram βακτήρια

2.2.1.1 Εισαγωγή

Οι διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού βρίσκονται κυρίως στην εξωτερική μεμβράνη των αρνητικών κατά Gram βακτηρίων. Για το λόγο αυτό στη παρούσα παράγραφο περιγράφονται τα βακτήρια, ενώ δίνεται έμφαση στα αρνητικά κατά Gram βακτήρια [5,6]

Τα βακτήρια μαζί με πρωτόζωα, τις άλγες, τους μύκητες και τους ιούς περιλαμβάνονται στα μικρόβια και έχουν μεγάλο ιατρικό ενδιαφέρον λόγω των λοιμώξεων που μπορούν να προκαλέσουν. Στα συνήθη βακτήρια που προκαλούν λοιμώξεις στον άνθρωπο περιλαμβάνονται οι σταφυλόκοκκοι, οι στρεπτόκοκκοι, οι γονιόκοκκοι, οι πνευμονιόκοκκοι, τα κολοβακτηρίδια, οι σαλμονέλλες και το τρεπόννημα το ωχρόν που προκαλεί τη σύφιλη. Είναι μια ομάδα με ποικίλα χαρακτηριστικά όσον αφορά την ανατομία τους, τη φυσιολογία και τη βιοχημεία τους. Στα βακτήρια υπάγονται, επίσης, και μικροοργανισμοί με ιδιαίτερα χαρακτηριστικά όπως τα χλαμύδια και οι ρικέτσιες (πολλαπλασιάζονται μόνο σε κύτταρα μεγαλοοργανισμών) και τα μυκοπλάσματα (στερούνται κυτταρικού τοιχώματος). Το ιατρικό ενδιαφέρον που παρουσιάζουν τα βακτηρία δικαιολογεί και το ενδιαφέρον που δείχνουν τα τελευταία χρόνια επιστήμονες ακόμα και στα πλαίσια της Βιοπληροφορικής, π.χ. για την εύρεση των διαμεμβρανικών κομματιών των πρωτεϊνών που διαπερνούν τις μεμβράνες τους.

Τα βακτήρια διαχωρίζονται με βάση τη μορφολογία τους (σχήμα και δομή), τους γενετικούς, φυσιολογικούς και βιοχημικούς τους χαρακτήρες και μερικές φορές από τη νόσο που προκαλούν. Ως προς το σχήμα συνήθως τα βακτήρια είναι σφαιρικά (κόκκοι), επιμήκη (βακτηρίδια) ή σπειροειδή (σπειροχαίτες), ενώ διατάσσονται μεμονωμένα, ανά δύο (διπόκοκκοι, διπλοβακτηρίδια), σε αλυσίδες (στρεπτόκοκκοι, στρεπτοβακτηρίδια), κατά ομάδες σαν σταφύλι

2.2.1.2 Κυτταρόπλασμα

Το κυτταρόπλασμα είναι το κύριο μικροβιακό σώμα και αποτελείται από πυκνόρρευστη ουσία, που περιέχει διάφορα οργανικά και ανόργανα συστατικά ενώ μέσα σ' αυτό βρίσκονται ο πυρήνας, τα ριβοσώματα, τα έγκλειστα, τα μεσοσώματα και μερικές φορές ο σπόρος.

Ο πυρήνας των βακτηρίων, που λέγεται και πυρηνική περιοχή ή πυρηνοειδής, αποτελείται από ένα απλό δίκλωνο μόριο DNA μαζί με RNA, πολυαμίνες και πρωτεΐνες. Ο πυρήνας αυτός δεν περιβάλλεται από πυρηνική μεμβράνη, δεν περιέχει ιστόνες και δεν διαιρείται με μίτωση. Το DNA του βακτηρίου, που αντιπροσωπεύει και το χρωμόσωμα του, έχει μήκος περίπου 110-1400 μm και βρίσκεται αναδιπλωμένο ή «υπερελικωμένο» για να περιληφθεί στο μικρό μέγεθος (0.5-5 μm) του βακτηρίου. Η υπερελίκωση του DNA ελέγχεται από τα ένζυμα DNA - γυράσες. Στο κυτταρόπλασμα του κυττάρου μπορεί να υπάρχουν και εξωχρωμοσωματικά μόρια DNA, τα λεγόμενα πλασμίδια. Σ' αυτά υπάρχουν γόνιμοι που κωδικοποιούν διάφορα χαρακτηριστικά των βακτηρίων, όπως την αντοχή σε αντιβιοτικά και την παραγωγή ορισμένων μεταβολικών ενζύμων και τοξινών.

Τα ριβοσώματα είναι οι θέσεις της πρωτεϊνοσύνθεσης. Είναι μικρά σφαιρικά σωμάτια μεγέθους περίπου 30 x 17 nm. Η χημική τους σύνθεση είναι κατά τα δύο τρίτα r-RNA και κατά το ένα τρίτο πρωτεΐνες. Σε ταχέως αναπτυσσόμενα βακτήρια αποτελούν περίπου το ένα τέταρτο της συνολικής μάζας του κυττάρου. Τα ριβοσώματα βρίσκονται διάσπαρτα μέσα στο κυτταρόπλασμα και στην εσωτερική επιφάνεια της κυτταροπλασματικής μεμβράνης. Συνήθως βρίσκονται κατά αθροίσματα, τα πολυριβοσώματα ή πολυσώματα, που ενώνονται με μόρια m-RNA για την πρωτεϊνοσύνθεση.

Τα έγκλειστα που βρίσκονται στο κυτταρόπλασμα μπορεί να είναι κενοτόπια αερίων, καρβοξυσώματα, αποθεματικά υλικά όπως πολυσακχαρίτες (άμυλο, γλυκογόνο), πολυφωσφορικά, πολυ-β-υδροξυβουτυρικό οξύ, αποθέματα θείου, κ.ά.

Οι σπόροι αποτελούν ανθεκτικές, και όχι αναπαραγωγικές, μορφές των βακτηρίων. Βρίσκονται μέσα στο κυτταρόπλασμα του βακτηριακού κυττάρου, γι' αυτό και λέγονται ενδοσπόρια. Το βακτηριακό κύτταρο που περιέχει σπόρο

λέγεται σποράγγειο. Σπορογόνα είναι κυρίως τα βακτήρια που ανήκουν στα γένη *Bacillus* και *Clostridium*. Οι σπόροι είναι πολύ ανθεκτικοί στη θερμότητα, στην ξηρασία, στην ακτινοβολία και σε τοξικές χημικές ουσίες.

2.2.1.3 Κυτταροπλασματική μεμβράνη

Το κυτταρόπλασμα των βακτηρίων περιβάλλεται από την κυτταροπλασματική μεμβράνη, που μαζί με το κυτταρικό τοίχωμα αποτελούν το κυτταρικό περίβλημα (cell envelope). Η κυτταροπλασματική μεμβράνη έχει εκλεκτική διαπερατότητα, καθώς επίσης αναπνευστικές και ενζυματικές λειτουργίες. Η κυτταροπλασματική μεμβράνη των βακτηρίων γενικά έχει απλούστερη κατασκευή από την κυτταρική μεμβράνη των ευκαρυωτικών κυττάρων. Αποτελείται κυρίως από λιποειδή και πρωτεΐνες. Τα κυριότερα λιποειδή της μεμβράνης είναι φωσφολιπίδια, που είναι οι εστέρες της γλυκερίνης με τα λιπαρά οξέα και ιονισμένα φωσφορικά. Η φωσφορική ομάδα ενώνεται με αζωτούχες βάσεις, όπως τη χολίνη, την αιθανολαμίνη, τη σερίνη. Το άκρο του φωσφολιπιδίου με τη φωσφορική ομάδα είναι πολικό και υδρόφιλο, ενώ το άκρο με το λιπαρό οξύ είναι μη πολικό και υδρόφοβο. Τα φωσφολιπίδια της κυτταροπλασματικής μεμβράνης διατάσσονται σε δύο στιβάδες με τις υδρόφοβες ομάδες εσωτερικά και τις υδρόφιλες εξωτερικά (προς το κυτταρικό τοίχωμα και προς το κυτταρόπλασμα). Οι πρωτεΐνες της κυτταροπλασματικής μεμβράνης κατανέμονται μεταξύ των μορίων των φωσφολιπιδίων. Μπορούν να αποτελούν και το 50% του ξηρού βάρους της μεμβράνης. Για να λειτουργεί σωστά η μεμβράνη, πρέπει τα λιπαρά οξέα να είναι σε ρευστή κατάσταση και αυτό εξαρτάται από τον κορεσμό και τις διακλαδώσεις τους, καθώς και από τη θερμοκρασία. Χαρακτηριστικά, στα ψυχρόφιλα βακτήρια υπάρχουν ακόρεστα, ενώ στα θερμοφιλα επικρατούν τα κορεσμένα διακλαδούμενα λιπαρά οξέα. Σε αντίθεση με τα ευκαρυωτικά κύτταρα, η μεμβράνη των προκαρυωτικών κυττάρων δεν περιέχει στερόλες, με μόνη εξαίρεση τα μυκοπλάσματα.

Οι κυριότερες λειτουργίες της κυτταροπλασματικής μεμβράνης είναι η μεταφορά διαφόρων ουσιών, η μεταφορά ηλεκτρονίων και η οξειδωτική φωσφορυλίωση, η έκκριση υδρολυτικών εξωενζύμων και η βιοσύνθεση διαφόρων μακρομορίων.

2.2.1.4 Περιπλασματικός χώρος

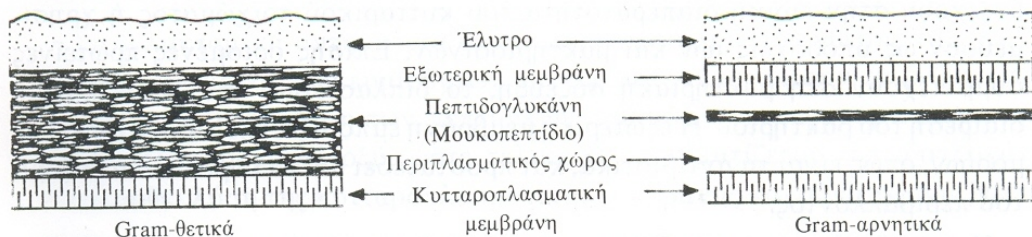
Περιπλασματικός λέγεται ο χώρος που βρίσκεται μεταξύ της κυτταροπλασματικής μεμβράνης και του κυτταρικού τοιχώματος. Είναι εμφανής στα Gram -αρνητικά βακτήρια (Εικόνες 2.3 και 2.4).

Στον περιπλασματικό χώρο υπάρχουν πρωτεΐνες και διάφορα ένζυμα, που διασπών τα αντιβιοτικά. Επίσης, εδώ βρίσκονται σε μεγάλες συγκεντρώσεις υδρολυτικά ένζυμα, όπως είναι οι φωσφατάσες, DNAάση, RNAάση, καθώς επίσης και ειδικές πρωτεΐνες, που συνδέονται με σάκχαρα, αμινοξέα και ανόργανα άλατα.

2.2.1.5 Κυτταρικό τοίχωμα

Γύρω από την κυτταροπλασματική μεμβράνη και τον περιπλασματικό χώρο βρίσκεται το κυτταρικό τοίχωμα του βακτηρίου, που καθορίζει τη μορφολογία του κυττάρου και το προστατεύει από αντίξοες συνθήκες του περιβάλλοντος. Η χημική σύνθεση του κυτταρικού τοιχώματος δεν είναι ίδια σ' όλα τα βακτήρια. Αυτή καθορίζει τη διαφορετική αντίδραση των βακτηρίων σε πολλούς φυσικούς και χημικούς παράγοντες, αλλά και το βασικό διαχωρισμό τους σε Gram - θετικά και Gram - αρνητικά βακτήρια (Εικόνα 2.4).

Το κυτταρικό τοίχωμα των Gram - θετικών βακτηρίων αποτελείται κυρίως (50%-90%) από πεπτιδογλυκάνη (μουκοπεπτιδιο). Σ' αυτό προστίθενται τα τειχοϊκά οξέα και οι πολυσακχαρίτες. Η πεπτιδογλυκάνη αποτελείται από τα αμινοσάκχαρα N-ακετυλογλυκοζαμίνη και N-ακετυλομουραμικό οξύ, που συνδέονται μεταξύ τους με 1,4 β-γλυκοσιδικό. Με την καρβοξυλική ομάδα του N-ακετυλομουραμικού οξέος ενώνονται τετραπεπτιδικές αλυσίδες που αποτελούνται από τα αμινοξέα D-γλουταμινικό οξύ, D- και L-αλανίνη, και διαμινοπιμελικό οξύ ή λυσίνη. Οι πεπτιδικές αλυσίδες του N-ακετυλομουραμικού οξέος ενώνονται μεταξύ τους με πενταπεπτιδική γέφυρα γλυκίνης. Τα τειχοϊκά οξέα είναι πολυμερή της γλυκερίνης και ριβιτόλης, που ενώνονται μεταξύ τους με φωσφορικές ομάδες. Οι υδροξυλομάδες της γλυκερίνης ή ριβιτόλης συνδέονται με αμινοξέα ή αμινοσάκχαρα.

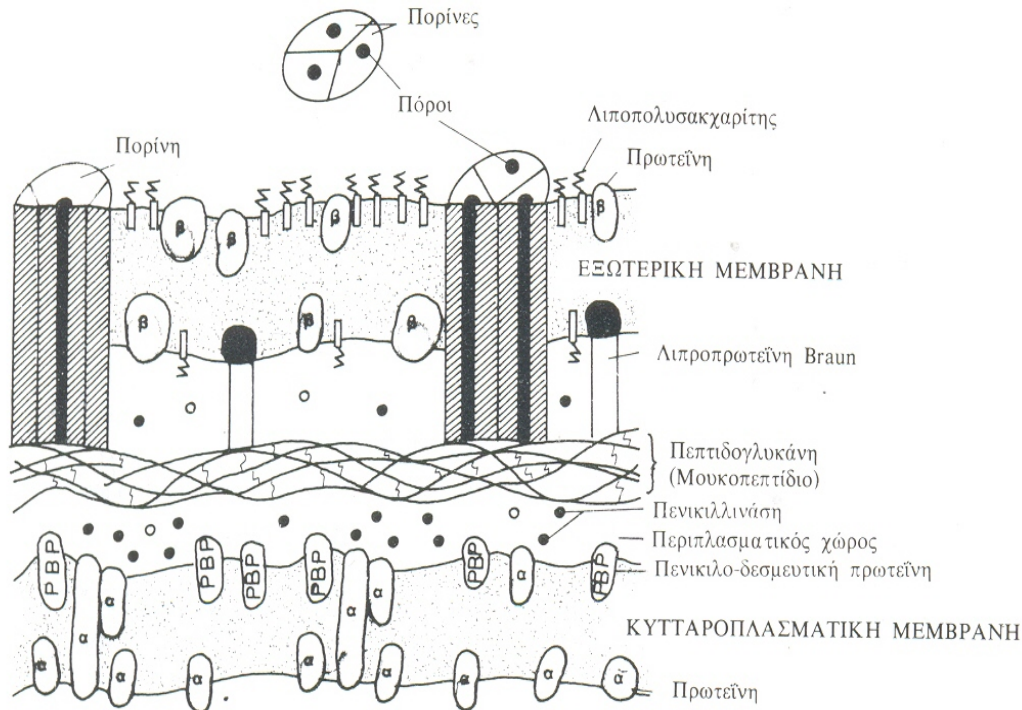


Εικόνα 2.3 Σύγκριση του κυτταρικού τοιχώματος των Gram- θετικών και Gram- αρνητικών βακτηρίων

Αντίθετα, το κυτταρικό τοίχωμα των Gram - αρνητικών βακτηρίων αποτελείται από περισσότερα στρώματα, αλλά είναι λεπτότερο. Εσωτερικά βρίσκεται λεπτό στρώμα (5%-10%) πεπτιδογλυκάνης, που περιβάλλεται από την εξωτερική μεμβράνη (Εικόνα 2.3 και 2.4). Η εξωτερική μεμβράνη συνδέεται με την πεπτιδογλυκάνη με λιποπρωτεΐνες και αποτελείται εσωτερικά από φωσφολιπίδια και εξωτερικά από λιποπολυσακχαρίτες. Μεταξύ των μορίων αυτών βρίσκονται διάφορες πρωτεΐνες. Ορισμένες πρωτεΐνες, που λέγονται και πρωτεΐνες σκελετού (πρόκειται για τις διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού που μελετώνται στη παρούσα εργασία) διαπερνούν και τις δύο πλευρές της εξωτερικής μεμβράνης και συνδέονται στενά με την πεπτιδογλυκάνη. Διατάσσονται δε σε τριάδες ώστε να σχηματίζουν πόρους (γι' αυτό λέγονται και πορίνες), από τους οποίους διέρχονται υδρόφιλα μόρια. Οι πρωτεΐνες σκελετού χρησιμεύουν και ως υποδοχείς ορισμένων ιών. Ορισμένες άλλες δευτερεύουσες πρωτεΐνες της εξωτερικής μεμβράνης συμμετέχουν στην ειδική διαπερατότητα του κυτταρικού τοιχώματος ή χρησιμεύουν ως υποδοχείς ιών και βακτηριοσινών. Επίσης ορισμένες πρωτεΐνες συμμετέχουν στη βακτηριακή σύζευξη, το διπλασιασμό του DNA και τη διαίρεση του βακτηρίου. Η εξωτερική μεμβράνη εμποδίζει την είσοδο μεγάλων μορίων, όπως είναι τα αντιβιοτικά, και προστατεύει απ' τα υδρολυτικά ένζυμα του περιβάλλοντος. Η πεπτιδογλυκάνη στα Gram - αρνητικά βακτήρια έχει, βασικά, την ίδια δομή με τη διαφορά ότι οι πεπτιδικές αλυσίδες του N-ακετυλομουραμικού οξέος ενώνονται μεταξύ τους με πεπτιδικό δεσμό μεταξύ της D-αλανίνης και του διαμινοπιμελικού οξέος,

Οι λιποπολυσακχαρίτες (LPS) του κυτταρικού τοιχώματος των Gram - αρνητικών βακτηρίων αποτελούνται από ένα σύνθετο λιποειδές, που ονομάζεται λιποειδές A, και που συνδέεται με ένα πολυσακχαρίτη αποτελούμενο από πυρήνα και ακραία αλυσίδα πολλών σακχάρων. Το λιποειδές A έχει ιδιότητες

τοξίνης και είναι ισχυρό πυρετογόνο, ενώ ο πολυσακχαρίτης έχει αντιγονικές ιδιότητες (αντιγόνο O).



Εικόνα 2. 4 Σχηματική αναπαράσταση του κυτταρικού τοιχώματος των Gram- αρνητικών βακτηρίων: α και β είναι οι πρωτεΐνες της κυτταροπλασματικής και εξωτερικής μεμβράνης αντίστοιχα, PBP είναι πενικιλλοδεσμευτικές πρωτεΐνες.

2.2.1.6 Έλυτρο

Σε αρκετά βακτήρια γύρω από το κυτταρικό τοίχωμα υπάρχουν μεγάλες ποσότητες εξωκυτταρίων πολυμερών, που διαμορφώνονται σε πυκνό βλενώδες στρώμα και σχηματίζουν το έλυτρο. Χημικά αποτελείται κυρίως από νερό και κατά μικρό μέρος (περίπου 2%) από πολυσακχαρίτες ή και από πολυπεπίδια. Η υφή του είναι πηκτωματώδης και τα μικρόβια που έχουν έλυτρο σχηματίζουν λείες ή βλενώδεις αποικίες, ενώ τα στελέχη των μικροβίων χωρίς έλυτρο μπορεί να σχηματίζουν ανώμαλες αποικίες. Η ύπαρξη του ελύτρου δεν είναι απαραίτητη για τη ζωή του μικροοργανισμού, αλλά σχετίζεται με τη λοιμογόνο δράση του. Τα στελέχη που φέρουν έλυτρο διαφεύγουν τη φαγοκυττάρωση και προσκολλώνται ευκολότερα στις διάφορες επιφάνειες. Το έλυτρο προστατεύει το κυτταρικό τοίχωμα από αντι-βακτηριακούς παράγοντες, όπως είναι οι βακτηριοφάγοι, οι κολισίνες, το συμπλήρωμα, η λυσοζύμη και άλλα λυτικά

ένζυμα. Ορισμένες φορές, τα εξωκυττάρια πολυμερή δεν διαμορφώνονται σε έλυτρο, αλλά σχηματίζουν ένα χαλαρό δίκτυο ινών γύρω από το κύτταρο (γλυκοκάλυκας), που παίζει σημαντικό ρόλο στην προσκόλληση των βακτηρίων στις διάφορες επιφάνειες.

2.2.1.7 Βλεφαρίδες και Ινίδια

Οι βλεφαρίδες ή μαστίγια είναι νηματοειδή εξαρτήματα και χρησιμεύουν ως όργανα κινήσεως. Ανάλογα με τη θέση των βλεφαρίδων τα βακτήρια διακρίνονται σε μονότριχα, λοφιοτρίχα ή περιτριχα. Οι βλεφαρίδες συνδέονται με την κυτταροπλασματική μεμβράνη, από την οποία αντλούν ενέργεια για την κίνηση τους. Οι βλεφαρίδες αποτελούνται από ειδικές υπομονάδες πρωτεϊνών, που λέγονται φλαγγελίνες. Οι πρωτεΐνες αυτές διατάσσονται κυλινδρικά και δεν είναι ίδιες σε όλα τα βακτήρια. Η κίνηση που προσδίδουν οι βλεφαρίδες στα βακτήρια είναι κυματοειδής και ταχεία (μέχρι 50 φορές το μήκος του βακτηρίου ανά δευτερόλεπτο).

Τα ινίδια ή φίμπριες είναι λεπτά και βραχέα εξαρτήματα που αποτελούνται από ειδικές πρωτεΐνες (πιλίνες) και ξεκινούν από την κυτταροπλασματική μεμβράνη. Διακρίνονται σε κανονικά ινίδια, που χρησιμεύουν ως όργανα προσκόλλησης των βακτηρίων στα κύτταρα του ξενιστού, και σε συζευκτικά ινίδια, που χρησιμεύουν για μεταφορά γενετικού υλικού κατά τη σύζευξη των βακτηρίων και ως υποδοχείς ορισμένων βακτηριοφάγων. Τα ινίδια δεν είναι όργανα κινήσεως.

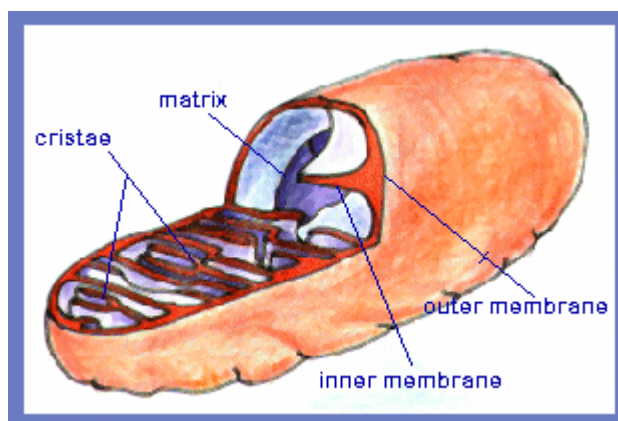
Στις σπειροχαίτες (τρεπτονήματα, λεπτοσπείρες, μπορέλιες), όργανα κινήσεως είναι τα αξονικά νημάτια, που έχουν την ίδια δομή με τις βλεφαρίδες, βρίσκονται στον περιπλασματικό χώρο και διατρέχουν αρκετό μέρος του βακτηριακού σώματος. Αυτά προσδίδουν την ελικοειδή εμφάνιση των σπειροχαιτών.

Πιστεύεται, επίσης, ότι οι διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού βρίσκονται και στην εξωτερική μεμβράνη των μιτοχονδρίων και χλωροπλαστών, γι' αυτό και παρατίθενται στο σημείο αυτό πληροφορίες σχετικές με αυτά τα οργανίδια.

2.2.2 Μιτοχόνδρια

Τα μιτοχόνδρια είναι οργανίδια που βρίσκονται σε όλα τα ευκαρυωτικά κύτταρα και επιτελούν σημαντικές λειτουργίες για την επιβίωση των κυττάρων [1]. Η πρώτη ίσως περιγραφή έγινε από τον Kolliker το 1857 ενώ μόλις το 1950 διατυπώθηκε από πολλές ερευνητικές ομάδες ότι τα μιτοχόνδρια δεν είναι μόνο θέσεις κυτταρικής αναπνοής αλλά αποτελούν και κέντρα παραγωγής ATP των αερόβιων κυττάρων. Σε πολλά κύτταρα τα μιτοχόνδρια είναι ομοιόμορφα κατανεμημένα στο κυτταρόπλασμα, ενώ σε μερικές περιπτώσεις συγκεντρώνονται σε συγκεκριμένες περιοχές, όπως στην περίπτωση των σπερματοζωαρίων όπου συγκεντρώνονται σε δακτυλίους γύρω από τις ουρές τους. Κατά τη μίτωση, τα μιτοχόνδρια συγκεντρώνονται κοντά στη μιτωτική άτρακτο και στο τέλος της διαίρεσης κατανέμονται σχεδόν εξίσου στα θυγατρικά κύτταρα. Το μέγεθος των μιτοχονδρίων συνήθως διαφέρει μεταξύ των κυτταρικών τύπων. Στα περισσότερα όμως κύτταρα έχουν πλάτος περίπου 0.5 μm και μήκος που συχνά μπορεί να φτάνει τα 7 μm, ενώ το μέγεθος τους μπορεί να αλλάζει κάτω από ορισμένες συνθήκες, φυσιολογικές ή μη. Το σχήμα, επίσης, των μιτοχονδρίων διαφέρει μεταξύ των κυτταρικών τύπων, αλλά συνήθως είναι ωοειδές ή νηματοειδές. Και ο αριθμός των μιτοχονδρίων εξαρτάται από το τύπο και το λειτουργικό στάδιο του κυττάρου (π.χ. σε ένα ηπατοκύτταρο βρίσκονται περίπου 1000-1600 μιτοχόνδρια, ενώ στα καρκινικά κύτταρα είναι πιο λίγα από τα αντίστοιχα μη καρκινικά κύτταρα λόγω της μείωσης της οξειδωτικής και αύξησης της γλυκολυτικής δράσης των καρκινικών κυττάρων).

Τα μιτοχόνδρια (Εικόνα 2.5) είναι διπλομεμβρανικά οργανίδια.



Εικόνα 2. 5 Δομή μιτοχονδρίου. cristae:μιτοχονδριακές αναδιπλώσεις, matrix:μιτοχονδριακή μήτρα, inner membrane: εσωτερική μεμβράνη, outer membrane: εξωτερική μεμβράνη [7]

Η εξωτερική μιτοχονδριακή μεμβράνη έχει πάχος περίπου 6 nm και διαχωρίζει το εσωτερικό του μιτοχονδρίου από το υπόλοιπο κυτταρόπλασμα. Η εσωτερική μεμβράνη έχει και αυτή πάχος 6 nm και σχηματίζει πολύπλοκες αναδιπλώσεις προς το εσωτερικό του μιτοχονδρίου. Οι αναδιπλώσεις αυτές ονομάζονται μιτοχονδριακές αναδιπλώσεις (cristae) και είναι συνήθως κάθετες προς τον επιμήκη άξονα του μιτοχονδρίου. Οι δύο μεμβράνες χωρίζουν το εσωτερικό του μιτοχονδρίου σε δύο διακριτές περιοχές:

i) το χώρο μεταξύ των μεμβρανών που ονομάζεται διαμεμβρανικός χώρος και
ii) το χώρο που περικλείεται από την εσωτερική μεμβράνη και ονομάζεται μιτοχονδριακή μήτρα (matrix). Η ηλεκτρονική πυκνότητα και η κοκκιώδης εμφάνιση της δείχνουν ότι η μιτοχονδριακή μήτρα πρέπει να βρίσκεται σε ημιστέρεη ή ιξώδη κατάσταση.

Στη μιτοχονδριακή μήτρα παρατηρούνται διάφορα νηματοειδή ή κοκκιώδη έγκλειστα. Τα πιο εμφανή από αυτά είναι ηλεκτρονικά πυκνά κόκκια διαμέτρου 25-30 nm που αποτελούν θέσεις δέσμευσης δισθενών ιόντων και ονομάζονται ενδομιτοχονδριακά κόκκια. Ένα ακόμη κοκκιώδες σωματίο διαμέτρου 12-15 nm παρατηρείται πολύ κοντά ή σε επαφή με τη μεμβράνη των μιτοχονδριακών μικρολοφιών, τα οποία έχει αποδειχθεί ότι είναι μιτοχονδριακά ριβισώματα. Τα ινόμορφα μιτοχονδριακά έγκλειστα είναι ινίδια DNA διαμέτρου 2-5 nm (είναι γνωστό πως τα μιτοχόνδρια έχουν δικό τους DNA). Συνήθως συνδέονται με την μεμβράνη σε ένα ή δύο σημεία. Το μιτοχονδριακό DNA είναι κυκλικό και έχει μήκος 4,5- 5,5 μm. Λιγότερα συχνά παρατηρούνται μέσα στην μιτοχονδριακή μήτρα μεγάλοι πρωτεϊνικοί κρύσταλλοι ή λεκιθικά κοκκία.

Όπως προαναφέρθηκε τα μιτοχόνδρια επιτελούν λειτουργίες σημαντικές για την επιβίωση των κυττάρων, όπως η οξειδωτική φωσφορυλίωση που παράγει την απαραίτητη ενέργεια με τη μορφή ATP, η πρόσληψη δισθενών κατιόντων, ο λιπιδιακός μεταβολισμός και η παραγωγή θερμότητας.

Όσον αφορά τη διαπερατότητα των μιτοχονδρίων, αναφέρουμε σε αυτό το σημείο ότι η εξωτερική μεμβράνη είναι διαπερατή στα περισσότερα ιόντα και οργανικά μόρια με μοριακό βάρος μέχρι και 10.000 daltons, ενώ η εσωτερική είναι αδιαπέραστη στις περισσότερες υδρόφιλες ουσίες εκτός από μια μικρή ομάδα μορίων που μεταφέρονται με τη βοήθεια μεμβρανικών πρωτεϊνών

μεταφοράς, όπως ο μεταφορέας του νουκλεοτιδίου της αδενίνης. Και οι δύο μεμβράνες είναι ελεύθερα διαπερατές στο H_2O , O_2 και το CO_2 .

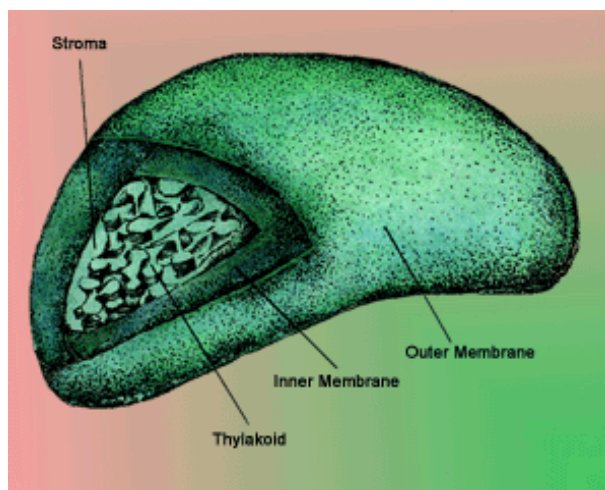
2.2.3 Χλωροπλάστες

Οι χλωροπλάστες είναι οργανίδια που βρίσκονται στα κύτταρα των φυτών και των φύκεων και συμμετέχουν ενεργά στην διαδικασία της φωτοσύνθεσης [8]. Λόγω του μεγάλου τους μεγέθους (1-5 μm πλάτος και 5-10 μm) ήταν από τα πρώτα οργανίδια που παρατηρήθηκαν: οι Antonie Van Leeuwenhoek και Nehemiah Grew τα περιέγραψαν για πρώτη φορά τον 17ο αιώνα όταν η μικροσκοπία ήταν στις πολύ αρχές της. Το κύτταρο ενός ώριμου φύλλου μπορεί να περιέχει 20-100 χλωροπλάστες ανάλογα με το είδος του φυτού και το περιβάλλον, ενώ το κύτταρο ενός φύκη περιέχει συνήθως έναν ή μερικούς χλωροπλάστες. Το σχήμα του είναι συνήθως πεπλατυσμένη σφαίρα στα φυτά ενώ στα πράσινα φύκη μπορεί να έχει και πιο ασυνήθιστο σχήμα, όπως τη μορφή ταινίας (λουρίδας) που εμφανίζεται στο φύκη *Spirogyra*. Οι χλωροπλάστες ανήκουν στα πλαστίδια και προέρχονται από τα προπλαστίδια. Τα προπλαστίδια, ανάλογα με το πού βρίσκονται και τις συνθήκες φωτισμού, πέρα από τους χλωροπλάστες μπορούν να εξελίσσονται σε αμυλοπλάστες, χρωμοπλάστες, πρωτεϊνοπλάστες κ.α.

Οι χλωροπλάστες (Εικόνα 2.6), όπως και τα μιτοχόνδρια, έχουν μια εξωτερική και μια εσωτερική μεμβράνη ανάμεσα στις οποίες υπάρχει ο διαμεμβρανικός χώρος. Η εσωτερική μεμβράνη εσωκλείει μια μήτρα από gel που ονομάζεται στρώμα και περιέχει ένζυμα που χρησιμεύουν στην αφομοίωση του άνθρακα, του αζώτου και του θείου. Η εξωτερική μεμβράνη περιέχει πορίνες, διαμεμβρανικές πρωτεΐνες που επιτρέπουν το πέρασμα ουσιών με μοριακό βάρος μέχρι και περίπου 5000, άρα και είναι διαπερατή στα περισσότερα μικρά οργανικά μόρια και ιόντα ενώ η εσωτερική μεμβράνη είναι διαπερατή σε λίγες ουσίες που τη διαπερνούν με τη βοήθεια πρωτεϊνών μεταφοράς. Και η εσωτερική μεμβράνη είναι όμως διαπερατή στο H_2O , O_2 και το CO_2 .

Οι χλωροπλάστες περιέχουν και ένα άλλο μεμβρανικό σύστημα, τα θυλακοειδή. Τα θυλακοειδή είναι επίπεδες δομές σε μορφή κυστών που βρίσκονται μέσα στο στρώμα και είναι οργανωμένα σε στοίβες που ονομάζονται γκράνα. Τα γκράνα που είναι σε στοίβες από νομίσματα είναι διασυνδεδεμένα σε ένα δίκτυο μέσω μακρών θυλακοειδών, που ονομάζονται θυλακοειδή

στρώματος. Πάνω ή μέσα στα θυλακοειδή βρίσκονται οι απαραίτητες ουσίες ή τα ένζυμα που είναι απαραίτητα για τη διαδικασία της φωτοσύνθεσης.



Εικόνα 2. 6 Δομή χλωροπλάστη. stroma: στρώμα, thylakoid: θυλακοειδές, inner membrane: εσωτερική μεμβράνη, outer membrane: εξωτερική μεμβράνη [9]

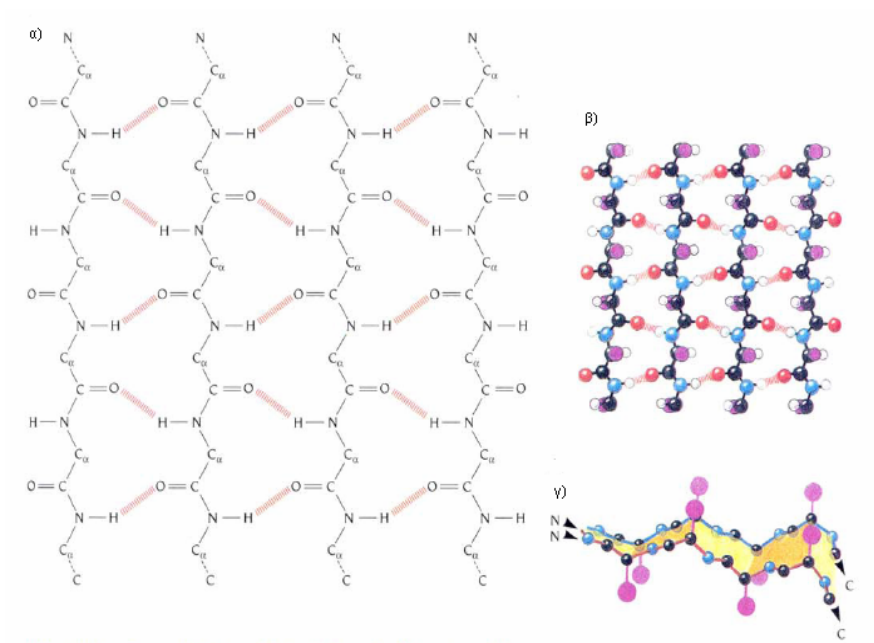
2.3 Η δομή των διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού

Τα β-βαρέλια αποτελούν κλειστές β-πτυχωτές επιφάνειες, γι' αυτό και οι τελευταίες περιγράφονται στην επόμενη παράγραφο.

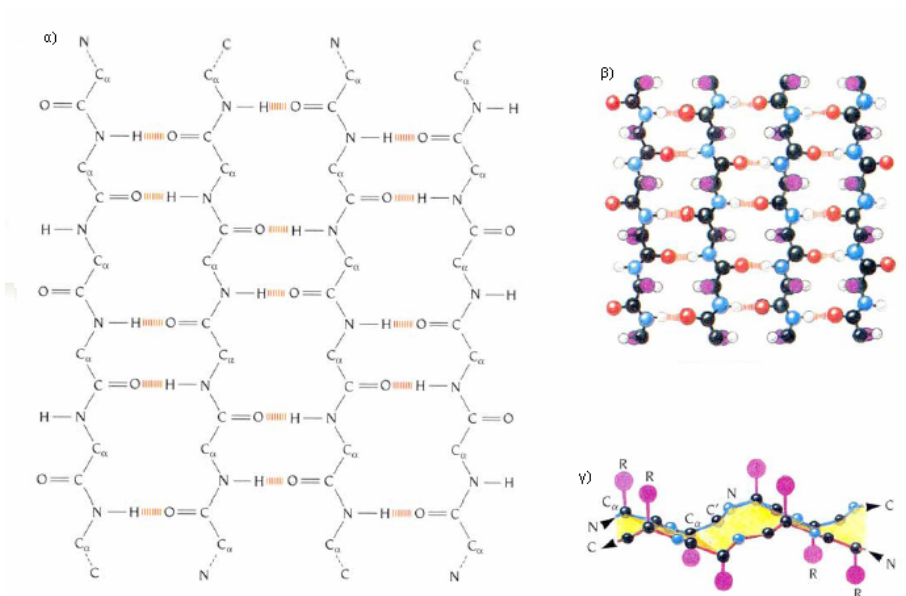
2.3.1 β-πτυχωτές επιφάνειες

Οι β-πτυχωτές επιφάνειες (β-sheets) [10] είναι δευτεροταγής δομή που προτάθηκε από τους Pauling και Corey (1951). Στις δομές αυτές σχηματίζονται υδρογονικοί δεσμοί μεταξύ καρβονυλικών ομάδων και αμινομάδων διαφορετικών πολυπεπτιδικών αλυσίδων ή μεταξύ διαφορετικών τμημάτων της ίδιας αλυσίδας, που ονομάζονται β-κλώνοι (β-strands). Οι αμινο- και καρβονυλο-ομάδες προσανατολίζονται περίπου κάθετα στον άξονα της αλυσίδας, που είναι σχεδόν τελείως εκτεταμένη, ώστε να είναι δυνατή η δημιουργία υδρογονικών δεσμών με αντίστοιχες ομάδες διπλανών αλυσίδων. Η ανάγκη σχηματισμού ισχυρών υδρογονικών δεσμών (περίπου ευθύγραμμων) έχει ως αποτέλεσμα οι β-κλώνοι να μην είναι τελείως εκτεταμένοι, οπότε η επιφάνεια που σχηματίζεται από αυτούς να παρουσιάζει μια ελαφρά κάμψη ή πτύχωση. Οι πλευρικές αλυσίδες τοποθετούνται κάθετα στον άξονα των β-κλώνων (και κάθετα στη β-πτυχωτή επιφάνεια) εναλλάξ πάνω και κάτω από την από την β-

πτυχωτή επιφάνεια. Είναι δυνατό, οι β-κλώνοι που αποτελούν μια β-πτυχωτή επιφάνεια να έχουν την ίδια διεύθυνση, οπότε η δομή χαρακτηρίζεται παράλληλη (Εικόνα 2.7), ή αντίθετες διευθύνσεις οπότε η δομή ονομάζεται αντιπαράλληλη (Εικόνα 2.8).



Εικόνα 2. 7 Παράλληλη β-πτυχωτή επιφάνεια

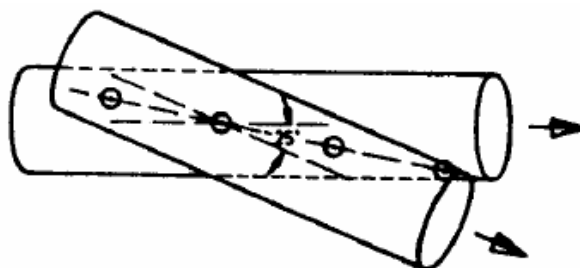


Εικόνα 2. 8 Αντιπαράλληλη β-πτυχωτή επιφάνεια

Οι περισσότερες β-πτυχωτές επιφάνειες περιέχουν λιγότερους από 6 β-κλώνους, ενώ περισσότερο εμφανίζονται επιφάνειες με περιττό αριθμό κλώνων.

Σπάνιες είναι οι παράλληλες επιφάνειες με λιγότερους από 4 β-κλώνους. Μικτές πτυχωτές επιφάνειες, δηλαδή με παράλληλους και αντιπαράλληλους κλώνους, εμφανίζονται μεν αλλά είναι λιγότερο συχνές.

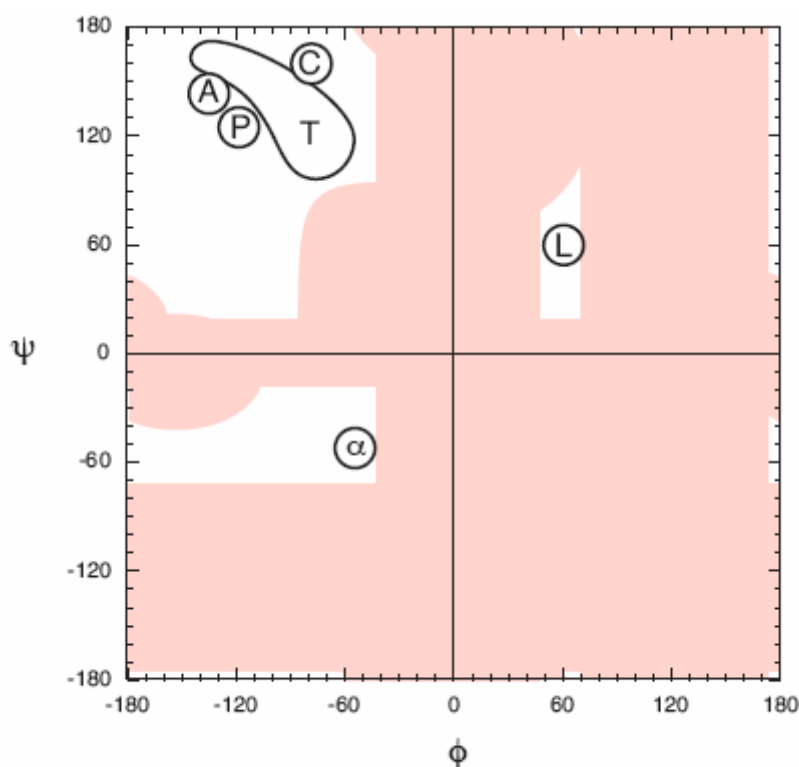
Οι β-πτυχωτές επιφάνειες, όπως προτάθηκαν από τους Pauling και Corey, αποτελούνται από επίπεδα παράλληλα μεταξύ τους ανά ζεύγη. Εν τούτοις, η μεγάλη πλειοψηφία των β-πτυχωτών επιφανειών που έχουν βρεθεί δεν έχουν αυτή την επίπεδη διάταξη: Αν παρατηρηθούν κατά μήκος του επιπέδου των επιφανειών και κάθετα στους κλώνους είναι αριστερόστροφες. Ισοδύναμα, αν παρατηρηθούν κατά μήκος των β-κλώνων φαίνονται δεξιόστροφες. Εμφανίζουν δηλ. μια χαρακτηριστική στροφή. Με αρκετή καλή προσέγγιση, ένας β-κλώνος μιας στραμμένης β-πτυχωτής επιφάνειας μπορεί να θεωρηθεί ως μια πολύ εκτεταμένη αριστερόστροφη έλικα (2.3 κατάλοιπα ανά στροφή, 3.3 Å απόσταση ανά κατάλοιπο παράλληλα με τον άξονα της έλικας και ακτίνα ίση με 1 Å). Η αριστερόστροφη αυτή έλικα αντιστοιχεί σε μία δεξιόστροφη περιστροφή των καρβονυλικών και αμιδικών ομάδων, περίπου 60° ανά 2 κατάλοιπα. Έτσι σχηματισμός ισχυρών υδρογονικών δεσμών μπορεί να γίνει μόνο αν οι κλώνοι σχηματίζουν γωνία περίπου 25° μεταξύ τους (Εικόνα 2.9). Έτσι παρατηρείται η στροφή της πτυχωτής επιφάνειας.



Εικόνα 2. 9 Σχηματισμός γωνίας 25° μεταξύ δύο κλώνων

Οι στραμμένες β-πτυχωτές επιφάνειες πιθανόν να παίζουν μεγάλο ρόλο στις αλληλεπιδράσεις πρωτεϊνών – DNA (ή RNA): η γεωμετρία μιας πλατείας, π.χ. με 6 β-κλώνους, ταιριάζει στην αυλάκωση της διπλής έλικας του DNA, ενώ μια στραμμένη αντιπαράλληλη επιφάνεια 2 κλώνων ταιριάζει στη μικρή αυλάκωση των DNA, RNA. Π.χ. προτάθηκε ότι στην χρωματίνη οι πρωταμίνες και οι ιστόνες θα πρέπει να περιέχουν στραμμένες β-πτυχωτές επιφάνειες, πράγμα που πρόσφατες ενδείξεις επιβεβαιώνουν.

Στα διαγράμματα Ramachandran (Εικόνα 2.10) όλα τα είδη πτυχωτών επιφανειών βρίσκονται σε αντίστοιχες επιτρεπόμενες περιοχές (αντιπαράλληλη (A) β-πτυχωτή επιφάνεια $\phi = -140^\circ$, $\psi = 135^\circ$ ανά κατάλοιπο, παράλληλη (P) β-πτυχωτή επιφάνεια $\phi = -120^\circ$, $\psi = 115^\circ$, στραμμένη β-πτυχωτή επιφάνεια $\phi = -120^\circ$, $\psi = 135^\circ$).



Εικόνα 2. 10 Διάγραμμα Ramachandran: Πάνω αριστερά υπάρχουν οι επιτρεπτές περιοχές για τις β-πτυχωτές επιφάνειες.

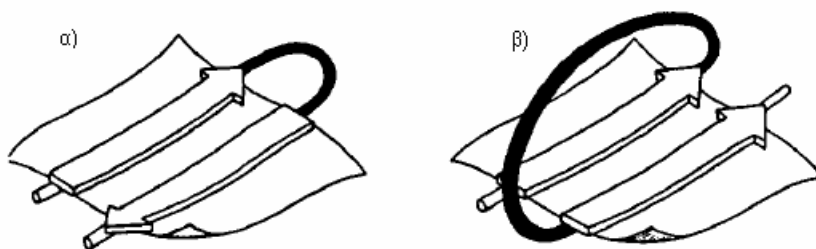
Κατάλοιπα που ευνοούν τη δημιουργία β-πτυχωτών επιφανειών είναι κυρίως κατάλοιπα με ογκώδεις υδροφοβικές αλυσίδες (Val, Ile, Leu, Trp, Cys, Tyr), οι οποίες διευθετούνται εύκολα στη στερεοδιάταξη της β-πτυχωτής επιφάνειας (το πακετάρισμα των πλευρικών αλυσίδων τους ευνοεί σημαντικά τις αλληλεπιδράσεις van der Waals).

Οι παράλληλες β-πτυχωτές επιφάνειες είναι συνήθως «θαμμένες» στο εσωτερικό των πρωτεϊνών και οι δύο επιφάνειες τους συνήθως «προστατεύονται» από άλλα τμήματα της κύριας αλυσίδας (συχνά α-έλικες), όντας ταυτόχρονα υδροφοβικές. Οι αντιπαράλληλες πτυχωτές επιφάνειες πολλές φορές έχουν την μια πλευρά τους εκτεθειμένη στο διαλύτη και την άλλη «θαμμένη», πράγμα που απεικονίζεται στην αμινοξική ακολουθία τους, που

εμφανίζει εναλλακτικά υδροφιλικά και υδροφοβικά κατάλοιπα. Οι β-πτυχωτές επιφάνειες γενικά, έχουν μεγαλύτερη υδροφοβικότητα στο κέντρο τους παρά στους ακραίους β-κλώνους.

Η συχνότητα εμφάνισης των β-πτυχωτών επιφανειών σε σφαιρικές υδατοδιαλυτές πρωτεΐνες είναι της τάξης του 20-25% και αυξάνει όσο αυξάνει ο αριθμός των κρυσταλλογραφικά λυμένων δομών. Απαντώνται σε πολλά είδη πρωτεϊνών, όπως σε πρωτεάσες, αφυδρογονάσες κ.λ.π. ή ακόμα σε ινώδεις (δομικές) πρωτεΐνες. Ακόμη εμφανίζονται και με τη μορφή β-βαρελιών όπως θα δούμε και παρακάτω.

Οι συνδέσεις των β-κλώνων (Εικόνα 2.11) ανήκουν σε 2 κατηγορίες: α) συνδέσεις τύπου φουρκέτας, στις οποίες η αλυσίδα επανέρχεται στη β-πτυχωτή επιφάνεια στο άκρο από το οποίο βγήκε και β) συνδέσεις τύπου διασταύρωσης, όπου η αλυσίδα επανέρχεται στη β-πτυχωτή επιφάνεια στο άλλο άκρο της.

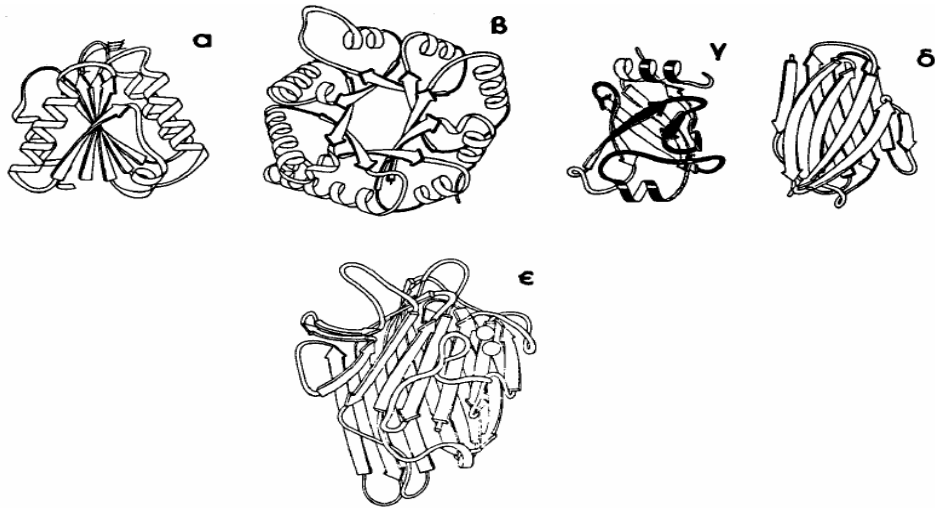


Εικόνα 2. 11 Συνδέσεις των β-κλώνων

2.3.2 β-βαρέλια

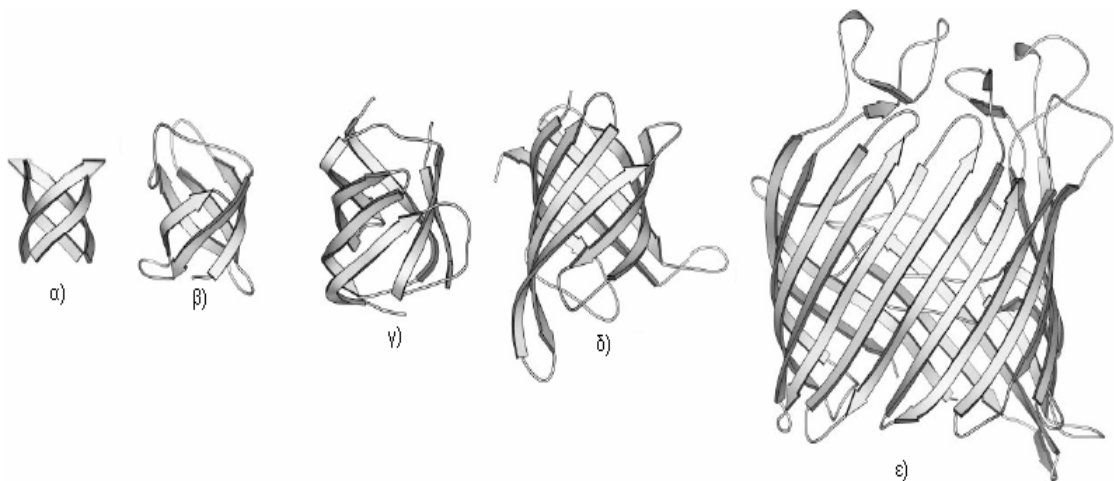
2.3.2.1 Γενικά

Τα β-βαρέλια αποτελούν είδος υπερδευτεροταγούς δομής των πρωτεϊνών (Εικόνα 2.12).



Εικόνα 2. 12 Υπερδευτεροταγείς δομές τύπου β-πτυχωτών επιφανειών

Οι παράλληλες β-πτυχωτές επιφάνειες συνήθως είναι μεγάλες στραμμένες επιφάνειες και μερικές διπλώνουν και συνιστούν κυλίνδρους με α-έλικες στο εσωτερικό τους. Οι μεγάλες αντιπαράλληλες β-πτυχωτές επιφάνειες, από την άλλη μεριά, διπλώνουν είτε μερικά είτε τελείως ώστε να δημιουργήσουν β-βαρέλια διαφόρων μεγεθών (Εικόνα 2.13).



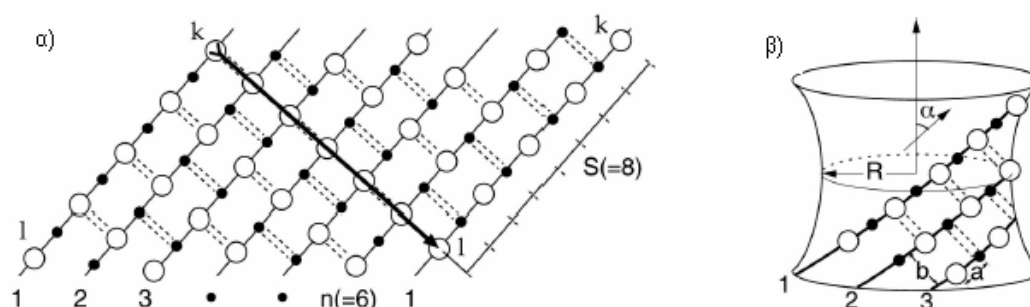
Εικόνα 2. 13. β-βαρέλια διαφόρων μεγεθών

Στο εσωτερικό των βαρελίων βρίσκονται καλά πακεταρισμένες υδρόφοβες πλευρικές αλυσίδες. Οι διατομές των β-βαρελίων παρουσιάζουν σημαντικές ομοιότητες, ανεξάρτητα από τον αριθμό των β-κλώνων. Η στροφή των β-κλώνων μεταβάλλεται μεταβαλλόμενου του αριθμού τους. Προφανώς τα β-βαρέλια [11-

13] αποτελούν διευθετήσεις ελάχιστης ενέργειας με συνεπακόλουθο αποτέλεσμα τη συχνή εμφάνιση τους.

2.3.2.2 Γεωμετρικά χαρακτηριστικά των β-βαρελιών

Όπως έχει ήδη αναφερθεί, το κύριο δομικό στοιχείο των β-βαρελιών είναι μια μεγάλη β-πτυχωτή επιφάνεια η οποία λόγω στρέψης της κλείνει και σχηματίζει μια κλειστή δομή (το β-βαρέλι) όπου ο τελευταίος κλώνος σχηματίζει υδρογονικούς δεσμούς με τον πρώτο. Πολλές μελέτες έχουν γίνει σχετικά με την απλή γεωμετρική δομή των β-βαρελιών [12-14]. Το πιο περιεκτικό και ευρέως αποδεκτό μοντέλο για τα β-βαρέλια είναι το μοντέλο “n,S”, το οποίο προτάθηκε αρχικά το 1979. Το μοντέλο “n,S” περιγράφει τα β-βαρέλια κάνοντας χρήση δύο παραμέτρων: τον αριθμό των κλώνων (n) και έναν αριθμό που δείχνει κατά πόσο είναι στραμμένη η β-πτυχωτή επιφάνεια που απαρτίζει το β-βαρέλι και ονομάζεται “shear number” (S)¹. Ο αριθμός S υπολογίζεται ως εξής (Εικόνα 2.14 α)): Θεωρούμε ως σημείο εκκίνησης το κατάλοιπο k του 1ου κλώνου και κινούμαστε σε κατεύθυνση κάθετη σε αυτή του 1ου κλώνου μέχρι να συναντήσουμε πάλι τον 1ο κλώνο. Τον κλώνο αυτό θα τον συναντήσουμε λόγω της στρέψης της πτυχωτής επιφάνειας στο κατάλοιπο l. Ως “shear number” (S) θεωρούμε τη διαφορά (l-k).



Εικόνα 2. 14 α) Εύρεση Shear Number του β-βαρελιού β) Ακτίνα R και κλίση α του β-βαρελιού

Ο αριθμός των κλώνων για τα βαρέλια των υδατοδιαλυτών πρωτεϊνών μπορεί να είναι και μέχρι 8, ενώ για τα βαρέλια των διαμεμβρανικών πρωτεϊνών το n μπορεί να πάρει τιμές από 8 μέχρι και 22. Το S έχει θετικές τιμές (μόνο σε περιπτώσεις β-ελίκων, αριστερόστροφων δηλαδή ελίκων, όπου και εκεί υπάρχει

¹ Μια ελληνική μετάφραση του “Shear Number” θα μπορούσε να είναι “Αριθμός Κάμψης”. Εν τούτοις στο παρόν κείμενο χρησιμοποιείται ο αγγλικός όρος.

αντίστοιχος όρος του “shear number”, το S παίρνει αρνητικές τιμές). Παίρνει συνήθως τιμές από n έως n+4 (το βαρέλι στην TolC για το οποίο S=n+8 αποτελεί εξαίρεση). Επίσης το S παίρνει άρτιες τιμές, λόγω του ότι η κατεύθυνση των υδρογονικών δεσμών εναλλάσσονται κατά μήκος ενός κλώνου οπότε για ένα πλήρη γύρο στο βαρέλι ο αριθμός των καταλοίπων ανάμεσα στο k και στο l κατάλοιπο είναι περιττός άρα το S=l-k είναι άρτιος αριθμός. Χρήσιμες πληροφορίες για το Shear Number μπορούν να βρεθούν στην [15]. Στον πίνακα 2.1 φαίνονται οι παράμετροι n, S για τα βαρέλια διάφορων πρωτεϊνών.

	n	S	R (Å)	α (deg)	Oligomeric state
Gramicidin A (native)	1	6	3.4	77	Head-to-head dimer traversing the membrane
Nanotube	1	-	5.6	90	Stack of about eight cyclic octapeptides through the membrane
OmpX	8	8	7.2	37	Monomer without channel
OmpA	8	10	7.9	43	Monomer without channel
OmpT	10	12	9.5	42	Monomer without channel
OmpLA ^c	12	12	10.6	37	Monomer without channel
TolC	12	20	13.6	51	Single β-barrel composed of a trimer, forms a channel
α-Haemolysin	14	14	12.3	37	Single β-barrel composed of a heptamer, forms a channel
Porin <i>Rhodobacter capsulatus</i>	16	20	15.5	43	Trimer of parallel β-barrels forming three channels
Porin OmpF (PhoE, OmpC)	16	20	15.5	43	Trimer of parallel β-barrels forming three channels
Porin <i>Rhodobacter blasticus</i>	16	20	15.5	43	Trimer of parallel β-barrels forming three channels
Porin <i>Paracoccus denitrificans</i>	16	20	15.5	43	Trimer of parallel β-barrels forming three channels
Porin Omp32	16	20	15.5	43	Trimer of parallel β-barrels forming three channels
Maltoporin (two species)	18	20	17.1	40	Trimer of parallel β-barrels forming three channels
Sucrose porin	18	20	17.1	40	Trimer of parallel β-barrels forming three channels
FhuA	22	24	19.9	39	Monomer clogged by a removable polypeptide domain
FepA	22	24	19.9	39	Monomer clogged by a removable polypeptide domain

Πίνακας 2. 1 Οι παράμετροι n,S,R,α για τα β-βαρέλια διάφορων πρωτεϊνών όπως παρουσιάζονται στη [14].

Η θεωρητική μελέτη του μοντέλου “n,S” έδειξε ότι οι παράμετροι n και S μπορούν να περιγράψουν πλήρως ένα ιδανικό β-βαρέλι.

Μάλιστα, κάνοντας τους εξής συμβολισμούς (Εικόνα 2.14 β):

α = μέση κλίση των κλώνων ως προς τον άξονα του βαρελιού

R = μέση ακτίνα του βαρελιού

a = η απόσταση μεταξύ των Ca ανθράκων κατά μήκος ενός κλώνου (a=3.3 Å για παράλληλες, αντιπαράλληλες και μεικτές πτυχωτές επιφάνειες)

b = η απόσταση μεταξύ δύο διαδοχικών κλώνων (b=4.4 Å για παράλληλες, αντιπαράλληλες και μεικτές πτυχωτές επιφάνειες)

ισχύει ότι:

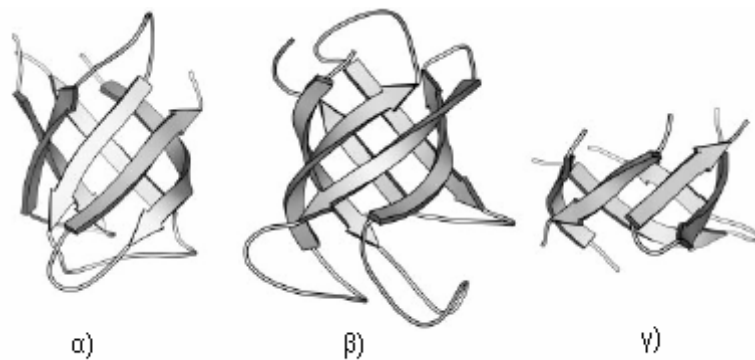
$$i) \tan\alpha = Sa / nb$$

$$ii) R = [(Sa)^2 + (nb)^2]^{0.5} / (2n \sin(\pi/n)) \cong [(Sa)^2 + (nb)^2]^{0.5} / 2\pi$$

iii) $R = nb / 2\pi \cos \alpha$ (circular cross section)

Επίσης, η ακτίνα του βαρελιού αυξάνει με την αύξηση του n αλλά και την αύξηση του R .

Στην Εικόνα 2.15 φαίνονται 3 διαφορετικές πρωτεΐνες με κοινό n ($n=6$) και διαφορετικό “shear number” ($S=8$ (α), $S=10$ (β) και $S=12$ (γ)). Οι αναμενόμενες τιμές για τα R , α , με χρήση των προαναφερόμενων σχέσεων, είναι αντίστοιχα 6.2 \AA και 45° , 7 \AA και 51° , και 7.9 \AA και 56° . Οι αντίστοιχες παρατηρηθείσες τιμές είναι 6.2 \AA και 43° , 7.1 \AA και 49° , και 8.0 \AA και 56° .



Εικόνα 2. 15. 3 διαφορετικές πρωτεΐνες με κοινό n ($n=6$) και διαφορετικό “shear number” ($S=8$ (α), $S=10$ (β) και $S=12$ (γ))

2.3.2.3 Δομικά χαρακτηριστικά των β -βαρελιών

Πέρα από τα χαρακτηριστικά της γεωμετρικής δομής των β -βαρελιών που περιγράφηκαν παραπάνω, υπάρχουν και άλλα χαρακτηριστικά που περιγράφουν τα διαμεμβρανικά β -βαρέλια [14,16]:

i) Όλοι οι β -κλώνοι είναι αντιπαράλληλοι και τοπικά συνδεδεμένοι με τους γειτονικούς τους κλώνους.

ii) Οι συνδέσεις των β -κλώνων στον εξωκυττάριο χώρο είναι μεγάλες λούπες (που συμβολίζονται ως L1, L2 κτλ) ενώ στον περιπλαστικό χώρο είναι στροφές μικρού μήκους (που συμβολίζονται ως T1, T2 κτλ.)

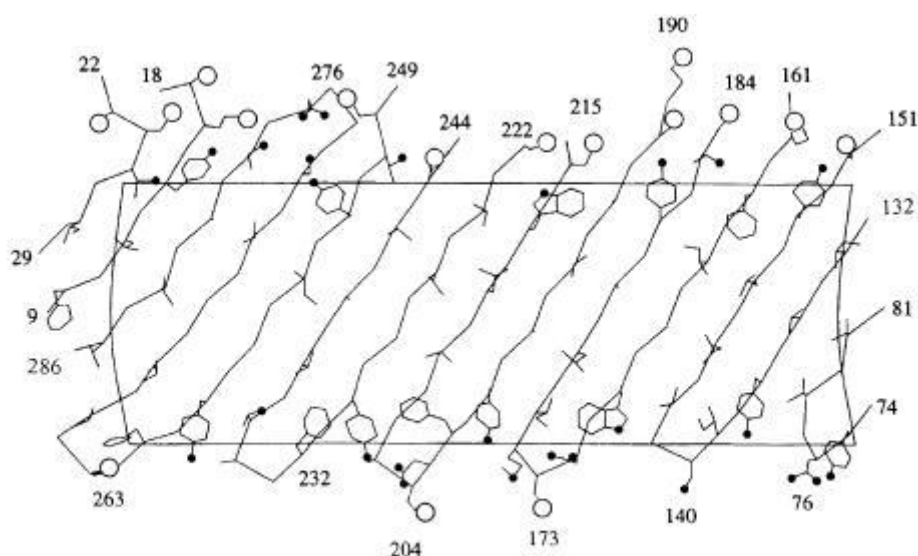
iii) Το αμινοτελικό καθώς και το καρβοτελικό άκρο (N-terminal και C-terminal) βρίσκονται στο περιπλαστικό μέρος, οπότε ο αριθμός των β -κλώνων που απαρτίζουν το β -βαρέλι είναι άρτιος.

iv) Στα τριμερή (πορίνες) στον άξονα του μορίου σχηματίζεται ένας υδρόφοβος πυρήνας, ώστε το κεντρικό κομμάτι ενός τριμερούς να θυμίζει μια υδατοδιαλυτή πρωτεΐνη.

v) Στις πορίνες το δέσιμο στο κέντρο του βαρελιού γίνεται μέσα από μια μεγάλη λούπα L3 που εισέρχεται στο βαρέλι.

vi) Στην αναπαράσταση του βαρελιού όπως φαίνεται στην Εικόνα 2.14 β) όπου ο περιπλαστικός χώρος βρίσκεται στο κάτω μέρος η αλυσίδα ξετυλίγεται από τα δεξιά προς τα αριστερά.

vii) Η επιφάνεια του β-βαρελιού που είναι σε επαφή με το μη πολικό εσωτερικό της μεμβράνης χαρακτηρίζεται από την ύπαρξη αλειφατικών μη πολικών αλυσίδων, ώστε να αποτελεί μια μη πολική ταινία. Τα όρια της ταινίας αυτής (πάνω και κάτω όριο) χαρακτηρίζονται από αρωματικές πλευρικές αλυσίδες (Εικόνα 2.16).



Εικόνα 2. 16 Το β-βαρέλι πορίνης στην *R. Blasticus* [17]. Η επιφάνεια του β-βαρελιού που είναι σε επαφή με το μη πολικό εσωτερικό της μεμβράνης χαρακτηρίζεται από την ύπαρξη αλειφατικών μη πολικών αλυσίδων ενώ τα όρια της ταινίας αυτής χαρακτηρίζονται από αρωματικές πλευρικές αλυσίδες

viii) Σύμφωνα με τον Wimley [16] έξω από τα όρια της μεμβράνης υπάρχουν περισσότερο φορτισμένα αμινοξέα από στους κλώνους του β-βαρελιού απ'ότι στο διαμεμβρανικό κομμάτι.

ix) Η ποικιλία στην αμινοξική ακολουθία των διαμεμβρανικών βαρελιών είναι μεγαλύτερη από αυτή στις υδατοδιαλυτές πρωτεΐνες και κυρίως στις λούπες του εξωκυττάρου χώρου.

3. Ανασκόπηση Σχετικών Εργασιών – Παρούσα Προσέγγιση

Στο παρόν κεφάλαιο γίνεται σύντομη παρουσίαση εργασιών που σχετίζονται με την παρούσα εργασία. Παράλληλα, παρουσιάζεται σύντομα η παρούσα προσέγγιση.

3.1 Ανασκόπηση Σχετικών Εργασιών

Τα τελευταία χρόνια αρκετή δουλειά έχει γίνει από ερευνητικές ομάδες που δραστηριοποιούνται στο χώρο της Βιοπληροφορικής και η οποία έχει να κάνει με την επεξεργασία ακολουθιών ή δομών διαμεμβρανικών πρωτεϊνών κάθε τύπου αλλά και συγκεκριμένα τύπου β-βαρελίου. Οι περισσότερες έχουν σκοπό την πρόγνωση της τοπολογίας του διαμεμβρανικού κομματιού από την ακολουθία ή τη δομή, ενώ γίνονται και προσπάθειες για τη γεωμετρική μοντελοποίηση των β-βαρελίων.

Επεξεργασία Ακολουθίας: Πολλές εργασίες έχουν δημοσιευθεί τα τελευταία χρόνια με σκοπό την πρόγνωση του διαμεμβρανικού κομματιού μιας διαμεμβρανικής πρωτεΐνης από την ακολουθία της [18]. Παλαιότερες εργασίες έχουν στηριχτεί στην υδροφοβική ανάλυση της ακολουθίας διαμεμβρανικών πρωτεϊνών τύπου β-βαρελίου κάνοντας χρήση αλληλοεπικαλυπτόμενων παραθύρων και της εναλλαγής της υδροφοβικότητας από υψηλή σε χαμηλή κατά μήκος των κλώνων [19,20].

Όσο νέα δεδομένα συλλέγονται και τα σύνολα εκπαίδευσης για τις μεθόδους πρόγνωσης μεγαλώνουν, νέες μέθοδοι αναπτύσσονται οι οποίες είναι λιγότερο απλοϊκές μια και χρησιμοποιούν νέες τεχνικές κυρίως από το χώρο της τεχνητής νοημοσύνης. Οι νέες μέθοδοι χρησιμοποιούν δεδομένα ακολουθίας μόνο ή δεδομένα ακολουθίας σε συνδυασμό με αναπτυξιακή (evolutionary) πληροφορία ενώ ως μηχανές πρόβλεψης έχουν χρησιμοποιηθεί τα τεχνητά νευρωνικά δίκτυα [21,22], τα μαρκοβιανά μοντέλα (Hidden Markov models: HMMs) [23-27] αλλά και οι μηχανές διανυσμάτων υποστήριξης (Support Vector Machines: SVMs) [28]. Σύμφωνα με την [18] οι μέθοδοι που χρησιμοποιούν HMMs (HMM-B2TMR,[23], ProfTMB [27], PRED-TMBB[26]) αποδίδουν καλύτερα από τις υπόλοιπες μεθόδους ενώ στην ίδια εργασία έγινε χρήση των προβλέψεων των

διαθέσιμων μεθόδων μαζί με δυναμικό προγραμματισμό με αποτέλεσμα την αύξηση της ακρίβειας πρόγνωσης.

Τέλος, στις εργασίες [19,29] περιγράφονται εργαλεία για την ταυτοποίηση πρωτεϊνών τύπου β-βαρελίου σε ολόκληρο το γονίδιο οργανισμών.

Επεξεργασία Δομής: Στην εργασία [30] παρουσιάζεται το πρόγραμμα Garlic το οποίο ταυτοποιεί τα διαμεμβρανικά κομμάτια μιας πρωτεΐνης κάνοντας χρήση της τριτοταγούς δομής της. Όσον αφορά τις διαμεμβρανικές πρωτεΐνες τύπου β-βαρελίου, θεωρεί ως πάχος της μεμβράνης σταθερό και ίσο με 21Å και κάνοντας χρήση εξαντλητικής αναζήτησης για τον άξονα του βαρελίου και περαιτέρω βελτίωση με χρήση των δακτυλίων αρωματικών αμινοξέων, ταυτοποιεί τον άξονα και ορίζει ως διαμεμβρανικό κομμάτι εκείνη τη ζώνη πάνω στον άξονα που έχει πάχος 21Å και έχει τη μέγιστη μέση υδροφοβικότητα. Ο χρήστης πρέπει να κατεβάσει το πρόγραμμα από αντίστοιχη ιστοσελίδα και να το εγκαταστήσει σε περιβάλλον Unix ή Linux.

Στην εργασία [31] παρουσιάζεται το εργαλείο (TM-DET) που έχει αναπτύξει η ομάδα του Tusnadi και των συνεργατών του για τη ταυτοποίηση του διαμεμβρανικού κομματιού διαμεμβρανικών πρωτεϊνών τύπου α-έλικας ή β-βαρελίου. Το εργαλείο αυτό αποτελεί σήμερα τη βασική μέθοδο επεξεργασίας δομής που χρησιμοποιείται σήμερα για την ταυτοποίηση του διαμεμβρανικού κομματιού μιας πρωτεΐνης. Παράλληλα, η ίδια ομάδα έχει δημιουργήσει βάση δεδομένων (PDB-TM) [32] που περιέχει όλες τις διαμεμβρανικές πρωτεΐνες με ορισμένο το διαμεμβρανικό τους κομμάτι, σύμφωνα με την έξοδο του εργαλείου. Το εργαλείο ταυτοποιεί μέσω εξαντλητικής αναζήτησης τη θέση της μεμβράνης κάνοντας χρήση μιας συνάρτησης ποιότητας για τις υποψήφιες θέσεις της μεμβράνης, που περιλαμβάνει παράγοντες υδροφοβικότητας και δομής των αμινοξέων που υπάρχουν σε ζώνη 15 Å για τις υποψήφιες θέσεις της μεμβράνης. Το πάχος της ζώνης αυξάνεται μέχρι το σημείο που δεν αλλάζει ο αριθμός των κομματιών της ακολουθίας που διέρχονται τη μεμβράνη (“crossing segments”). Στη συνέχεια, οι διαμεμβρανικές πρωτεΐνες κατηγοριοποιούνται στον τύπο α-έλικας ή β-βαρελίου. Το εργαλείο δέχεται ως είσοδο το .pdb αρχείο της πρωτεΐνης, ενώ την έξοδο απαρτίζουν 2 εικόνες που απεικονίζουν τη πρωτεΐνη με το διαμεμβρανικό της κομμάτι και ένα .xml αρχείο με τα αποτελέσματα επεξεργασίας της δομής (π.χ. τα διαμεμβρανικά αμινοξέα)

Στην [16] ο Wimley με σκοπό τη ταυτοποίηση των διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού στο γονιδίωμα οργανισμών, επεξεργάζεται τη δομή 15 λυμένων πρωτεϊνών αυτού του τύπου. Για τις δομές αυτές παρατηρεί το προφίλ της υδροφοβικότητας καθώς και παρουσίας φορτισμένων και αρωματικών αμινοξέων για αλληλοεπικαλυπτόμενα παράθυρα μήκους 5 Å κατά μήκος του άξονα του β-βαρελιού. Συμπεραίνει ότι το διαμεμβρανικό κομμάτι μιας διαμεμβρανικής πρωτεΐνης τύπου β-βαρελιού μπορεί να προσδιοριστεί με βάση αυτό το προφίλ.

Γεωμετρική μοντελοποίηση των β-βαρελιών: Στην εργασίες [12-14] οι συγγραφείς περιγράφουν γεωμετρικά τη δομή των β-βαρελιών καθώς και το μοντέλο “n,S” (n:αριθμός β-κλώνων, S:Shear Number) που χρησιμοποιείται για την περιγραφή τους. (στο 2^ο κεφάλαιο το μοντέλο αυτό παρουσιάζεται αναλυτικά).

3.2. Παρούσα Προσέγγιση

Στόχος της παρούσης εργασίας είναι η γεωμετρική μοντελοποίηση των β-βαρελιών των διαμεμβρανικών πρωτεϊνών αυτού του τύπου και η εύρεση του διαμεμβρανικού κομματιού τους κάνοντας χρήση της τριτοταγούς δομής τους. Οι αλγόριθμοι που υλοποιήθηκαν εφαρμόστηκαν σε ένα μη πλεοναστικό σύνολο 22 διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού (Πίνακας 3.1)

Η γεωμετρική μοντελοποίηση του β-βαρελιού περιλαμβάνει:

- Την εύρεση του άξονα και της ακτίνας του κυλίνδρου που προσεγγίζει με τον καλύτερο τρόπο το β-βαρέλι. Για το σκοπό αυτό, υλοποιήθηκε αλγόριθμος βελτιστοποίησης για την εύρεση του άξονα και της ακτίνας κυλίνδρου δεδομένου ενός αριθμού σημείων που ανήκουν σε αυτόν ο οποίος εμπλουτίστηκε κατάλληλα με Γενετικό Αλγόριθμο.
- Την εύρεση της μέσης κλίσης των β-κλώνων με τον άξονα του βαρελιού καθώς και τον θεωρητικό υπολογισμό του “Shear number” (S) σύμφωνα με το μοντέλο “n, S”.

Η εύρεση του διαμεμβρανικού κομματιού γίνεται κάνοντας χρήση του προφίλ υδροφοβικότητας καθώς και παρουσίας φορτισμένων και αρωματικών αμινοξέων για αλληλοεπικαλυπτόμενα παράθυρα μήκους 5 Å κατά μήκος του

άξονα του β-βαρελίου. Η εύρεση του διαμεμβρανικού κομματιού βασίζεται κυρίως στην εργασία του Wimley [16]. Τα αποτελέσματα σχετικά με την εύρεση του διαμεμβρανικού κομματιού συγκρίνονται με αυτά του Tusnadi. Τέλος, εξάγονται συμπεράσματα για την κατανομή των αμινοξέων και τη μέση υδροφοβικότητα κατά μήκος του άξονα του βαρελίου με βάση τις δομές που είναι στη διάθεση μας.

Όνομα Πρωτεΐνης	PDB ID	Αριθμός κλώνων	Οργανισμός	Αναφορά
NspA	1P4T	8	Neisseria Meningitidis	[33]
OmpX	1QJ8	8	Escherichia coli	[34]
Pagp	1MM4	8	Escherichia coli	[35]
OmpA	1QJP	8	Escherichia coli	[36]
OmpT	1I78	10	Escherichia coli	[37]
OpcA	1K24	10	Neisseria Meningitidis	[38]
Nalp	1UYN	12	Neisseria Meningitidis	[39]
OmpLA	1QD5	12	Escherichia coli	[40]
Tsx	1TLY	12	Escherichia coli	[41]
FadL	1TIL	14	Escherichia coli	[42]
Porin	2POR	16	Rhodobacter capsulatus	[43]
Porin	1PRN	16	Rhodopseudomonas blastica	[44]
OmpF	2OMF	16	Escherichia coli	[45]
Osmoporin	1OSM	16	Klebsiella pneumoniae	[46]
Omp32	1E54	16	Comamonas Acidovorans	[47]
Phosphoporin	1PHO	16	Escherichia coli	[48]
Sucrose porin	1A0S	18	Salmonella typhimurium	[49]
Maltoporin	2MPR	18	Salmonella typhimurium	[50]
FhuA	2FCP	22	Escherichia coli	[51]
FepA	1FEP	22	Escherichia coli	[52]
FecA	1KMO	22	Escherichia coli	[53]
BtuB	1NQE	22	Escherichia coli	[54]

Πίνακας 3.1 Το σύνολο των 22 πρωτεϊνών που χρησιμοποιήθηκε στη παρούσα μελέτη

4. Γεωμετρική Μοντελοποίηση του β-Βαρελιού

Στην παρόν κεφάλαιο παρουσιάζεται ο αλγόριθμος που αναπτύχθηκε, με σκοπό την γεωμετρική μοντελοποίηση του β-βαρελιού μιας διαμεμβρανικής πρωτεΐνης που είναι αυτού του τύπου. Για τις ανάγκες της γεωμετρικής μοντελοποίησης, το β-βαρέλι προσεγγίζεται από ένα κύλινδρο. Στόχος είναι η εύρεση του άξονα και της ακτίνας του κυλίνδρου που προσεγγίζει με βέλτιστο τρόπο το β-βαρέλι. Αναπτύχθηκε, λοιπόν, αλγόριθμος που υπολογίζει τον άξονα κυλίνδρου και την ακτίνα του δεδομένου ενός αριθμού σημείων που ανήκουν σε αυτόν, ο οποίος σε συνδυασμό με ένα Γενετικό Αλγόριθμο χρησιμοποιήθηκε για την εύρεση του άξονα και τις ακτίνας του β-βαρελιού. Η γεωμετρική μοντελοποίηση του β-βαρελιού ολοκληρώνεται με τον υπολογισμό της κλίσης των β-κλώνων (γωνία α) προς τον άξονα και του Shear Number (S) του βαρελιού, γεωμετρικών χαρακτηριστικών του β-βαρελιού που περιγράφηκαν σε προηγούμενο κεφάλαιο.

4.1 Αλγόριθμος για την εύρεση του άξονα και ακτίνας κυλίνδρου δεδομένου ενός αριθμού σημείων που ανήκουν σε αυτόν.

Στην παρούσα εργασία αναπτύχθηκε, αρχικά, αλγόριθμος για την εύρεση του άξονα κυλίνδρου και της ακτίνας του δεδομένου ενός αριθμού σημείων που ανήκουν σε αυτόν. Είσοδο στο σύστημα αποτελούν οι συντεταγμένες στον 3διάστατο χώρο των σημείων που ανήκουν στον κύλινδρο. Έστω, λοιπόν, N σημεία του 3διάστατου χώρου ($P_i=[x_i \ y_i \ z_i]$, $i=1..N$) τα οποία ανήκουν σε κύλινδρο με άξονα την ευθεία ε_0 και ακτίνα την R_0 . Αναζητούνται τα ε_0 και R_0 . Θεωρούμε ότι τα σημεία είναι ισοκατανεμημένα στην επιφάνεια του κυλίνδρου.

Θεωρώ το άθροισμα:

$$SSDs = \sum_{i=1}^N (dist(P_i, \varepsilon) - R)^2$$

, όπου $dist(P_i, \varepsilon)$ είναι η απόσταση του σημείου P_i από την ευθεία ε και R ένας θετικός πραγματικός αριθμός, ενώ ο συμβολισμός $SSDs$ αντιστοιχεί στο *Sum of Squared Distances*.

Επειδή τα σημεία $P_i=[x_i \ y_i \ z_i]$, $i=1..N$ ανήκουν στην επιφάνεια του κυλίνδρου με άξονα την ευθεία ε_0 και ακτίνα την R_0 , η ποσότητα $SSDs$ θα ελαχιστοποιείται

και θα γίνεται ίση με το μηδέν ($SSDs_{\min}=0$) όταν $\varepsilon = \varepsilon_0$ και $R = R_0$. Άρα η εύρεση του άξονα και την ακτίνας του κυλίνδρου ανάγεται σε ένα πρόβλημα ελαχιστοποίησης της εξής μορφής:

$$\min \left\langle \left(\sum_{i=1}^N (dist(P_i, \varepsilon) - R)^2 \right) \right\rangle$$

όπου τα ε και R μπορούν να μεταβάλλονται.

Οι περισσότερες μέθοδοι βελτιστοποίησης, εκτός από σύγχρονες μεθόδους όπως αυτή των γενετικών αλγορίθμων, εκτελούν τοπική βελτιστοποίηση και απαιτούν αρχικές εκτιμήσεις για τις μεταβλητές οι οποίες πρέπει να μεταβληθούν ώστε να επιτευχθεί η επιθυμητή βελτιστοποίηση. Στο παρόν πρόβλημα θα πρέπει να αναζητηθούν, λοιπόν, αρχικές εκτιμήσεις i) για την ευθεία ε (τον άξονα του κυλίνδρου) και ii) την ακτίνα R . Σημειώνουμε εδώ ότι η ακτίνα R θεωρείται ένας θετικός πραγματικός αριθμός, ενώ μια ευθεία ε θεωρούμε ότι ορίζεται από τις συντεταγμένες στον 3διάστατο χώρο δύο σημείων που ανήκουν σε αυτή (είναι γνωστό εξάλλου ότι 2 σημεία ορίζουν μοναδική ευθεία στο χώρο). Δηλαδή μια ευθεία συμβολίζεται ως $[x_A \ y_A \ z_A; x_B \ y_B \ z_B]$, όπου A και B είναι σημεία της ευθείας που χρησιμοποιούμε για να τον ορισμό της. Άρα το διάνυσμα που αποτελείται από τις τιμές $R, x_A, y_A, z_A, x_B, y_B, z_B$ είναι αυτό που πρέπει να βελτιστοποιηθεί κατά την ελαχιστοποίηση, ώστε να βρεθεί ο άξονας και η ακτίνα του κυλίνδρου

i) Αρχική εκτίμηση για την ακτίνα R κυλίνδρου($R_{αρχ}$): Παρατηρήθηκε ότι για N σημεία του 3διάστατου χώρου που είναι επαρκώς ομοιόμορφα κατανεμημένα στην επιφάνεια κυλίνδρου, η πιο πιθανά εμφανιζόμενη απόσταση των ανά 2 σημείων αποτελεί μια καλή προσέγγιση της διαμέτρου του κυλίνδρου. Αν υπολογίσουμε, λοιπόν, τις αποστάσεις των ανά 2 σημείων από τα N θα προκύψουν $0.5*N*(N-1)$ αποστάσεις. Οι αποστάσεις αυτές όταν απεικονιστούν σε ιστόγραμμα θα δώσουν ως στήλη με το μεγαλύτερο ύψος εκείνη τη στήλη που αντιστοιχεί στο διάστημα στο οποίο βρίσκεται η τιμή της διαμέτρου του κυλίνδρου ή μια πολύ καλή προσέγγιση της. Ο αριθμός των στηλών που έχει χρησιμοποιηθεί στην παρούσα εφαρμογή είναι 30. Παραδείγματα τέτοιων ιστογραμμάτων για κυλίνδρους με διαφορετικές ακτίνες φαίνονται παρακάτω,

όπου και παρατίθενται τα αποτελέσματα της μεθόδου για την εύρεση του άξονα κυλίνδρου τεχνητών δεδομένων.

Ως καλή εκτίμηση για τη διάμετρο ($\delta_{αρχ}$) μπορεί να θεωρηθεί το μέσο του διαστήματος που προκύπτει αν επεκτείνουμε το διάστημα που αντιστοιχεί στη στήλη με το μεγαλύτερο ύψος στο ιστόγραμμα των αποστάσεων κατά 25% δεξιά και αριστερά (έστω ότι είναι το διάστημα $[α β]$). Αυτό γίνεται για λόγους ασφαλείας στην εκτίμηση της διαμέτρου και χρησιμεύει στις περιπτώσεις που μια καλή εκτίμηση της διαμέτρου του κυλίνδρου βρίσκεται οριακά έξω από το διάστημα της στήλης με το μεγαλύτερο ύψος στο ιστόγραμμα. Τελικά, θεωρείται ως αρχική εκτίμηση για την ακτίνα του κυλίνδρου $R_{αρχ}$ το μισό της $\delta_{αρχ}$:

$$R_{αρχ} = 0.5 * \delta_{αρχ} = 0.25 * (α + β)$$

Ίδιας ποιότητας αποτελέσματα για την $R_{αρχ}$ θα παίρναμε αν θεωρούσαμε διαφορετικό αριθμό στηλών στο ιστόγραμμα των αποστάσεων και διαφορετική επέκταση δεξιά και αριστερά του διαστήματος που αντιστοιχεί στη στήλη με το μεγαλύτερο ύψος για λόγους ασφαλείας (διαφορετική από 25% δεξιά και αριστερά).

ii) Αρχική εκτίμηση για τον άξονα ε του κυλίνδρου ($\varepsilon_{αρχ}$): Ας θεωρήσω εκείνα τα ζεύγη σημείων από τα N σημεία του κυλίνδρου των οποίων οι αποστάσεις βρίσκονται στο διάστημα $[α β]$ και έστω ότι προκύπτουν $κ$ ζεύγη σημείων. Τα σημεία των ζευγών αυτών έχουν μεγάλη πιθανότητα να είναι αντιδιαμετρικά μεταξύ τους, οπότε ο γεωμετρικός τους μέσος (το μέσο του διανύσματος που τα ενώνει) έχει μεγάλη πιθανότητα να ανήκει στον άξονα του κυλίνδρου. Αν λάβω υπόψη τους $κ$ γεωμετρικούς μέσους των $κ$ θεωρημένων ζευγών σημείων και υπολογίσω την ευθεία που προκύπτει με γραμμική παρεμβολή τους, θα προκύψει μια καλή εκτίμηση για τον άξονα του κυλίνδρου, $\varepsilon_{αρχ}$. Για τη γραμμική παρεμβολή των $κ$ γεωμετρικών μέσων χρησιμοποιείται η μέθοδος της ορθογώνιας παλινδρόμησης η οποία περιγράφεται στο παράρτημα.

Στη συνέχεια, με χρήση κατάλληλου αλγορίθμου ελαχιστοποίησης πραγματοποιείται η ελαχιστοποίηση

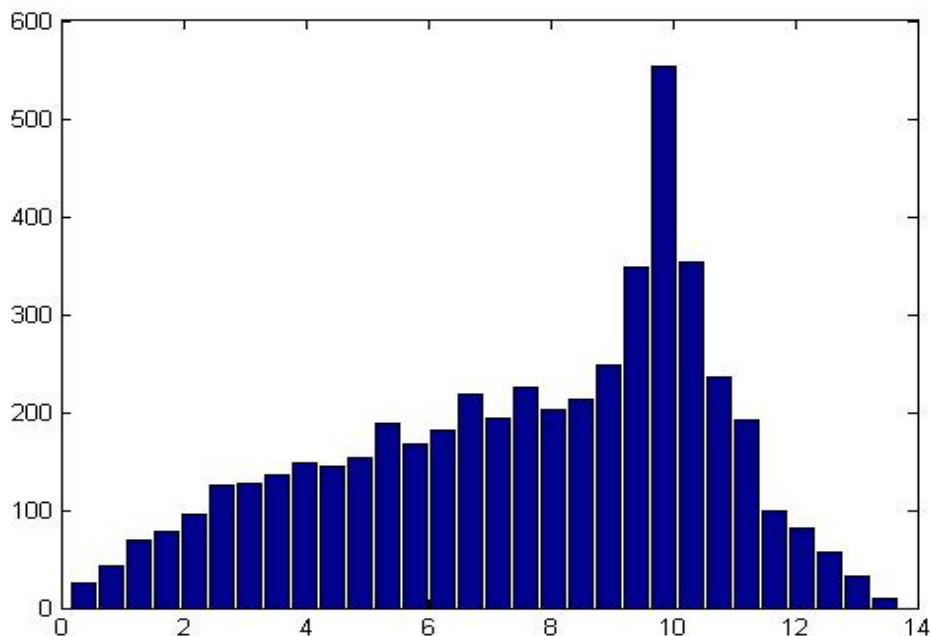
$$\min \left\langle \left(\sum_{i=1}^N (dist(P_i, \varepsilon) - R)^2 \right) \right\rangle$$

όπου η πρώτη επανάληψη του αλγορίθμου χρησιμοποιεί ως τιμές για την ευθεία ε και την ακτίνα R τα ήδη εκτιμημένα $\varepsilon_{αρχ}$ και $R_{αρχ}$. Ο αλγόριθμος ελαχιστοποίησης που χρησιμοποιείται αποτελεί αλγόριθμο μη γραμμικής βελτιστοποίησης ελαχίστων τετραγώνων και περιγράφεται σύντομα στο Παράρτημα. Ο συγκεκριμένος αλγόριθμος, με δεδομένες καλές αρχικές εκτιμήσεις για το $\varepsilon_{αρχ}$ και $R_{αρχ}$, μπορεί και βρίσκει την ευθεία ε_{min} και την ακτίνα R_{min} που αντιστοιχούν στο ολικό ελάχιστο του $SSDs$, δηλαδή μπορεί και βρίσκει τέλεια (ο αναγνώστης παραπέμπεται στα αποτελέσματα του αλγορίθμου σε τεχνητά δεδομένα παρακάτω) την ευθεία ε_o και την ακτίνα R_o που αντιστοιχούν στον άξονα και την ακτίνα του κυλίνδρου στον οποίο ανήκουν τα σημεία P_i ($i=1..N$).

Στο σημείο αυτό παρατίθενται τα αποτελέσματα του αλγορίθμου εύρεσης του άξονα και της ακτίνας κυλίνδρου δεδομένου ενός αριθμού σημείων που ανήκουν σε αυτόν σε 3 περιπτώσεις τεχνητών δεδομένων :

α) 100 τυχαία σημεία που ανήκουν σε κύλινδρο με άξονα την ευθεία $\varepsilon_o=xx'$, ο οποίος εκτείνεται από το $x=0$ έως το $x=10$ και έχει ακτίνα $R_o=5$

Το αντίστοιχο ιστόγραμμα αποστάσεων:



Σχήμα 4. 1 Ιστόγραμμα των αποστάσεων των ανά δύο σημείων που ανήκουν στον κύλινδρο της περίπτωσης α)

Με βάση το ιστόγραμμα αποστάσεων προέκυψαν οι εξής αρχικές εκτιμήσεις:

$$R_{αρχ} = 4.93 \text{ και}$$

$\varepsilon_{αρχ} = [9.801 \ 0.101 \ 0.216; -0.139 \ 0.381 \ -0.012]$ (η ευθεία με μια καλή προσέγγιση μπορεί να θεωρηθεί η ευθεία xx')

Με την ελαχιστοποίηση προέκυψαν:

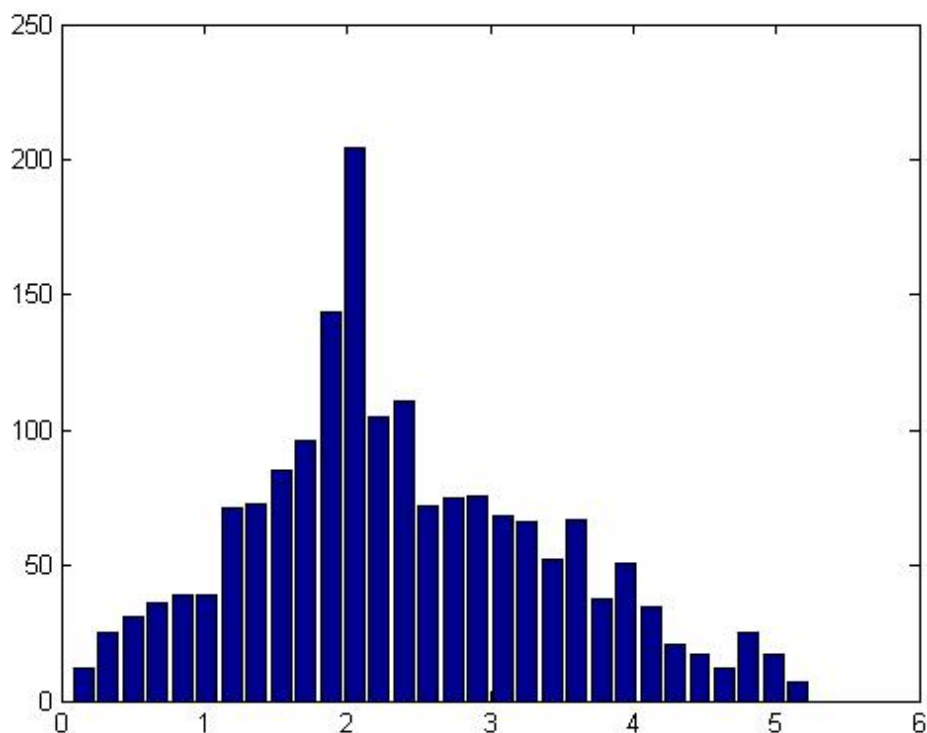
$$R_{min} = 5 \text{ και}$$

$\varepsilon_{min} = [9.688 \ 0 \ 0; 0.02 \ 0 \ 0]$ (η ευθεία αυτή αποτελεί τον άξονα xx')

Παρατηρούμε ότι με την ελαχιστοποίηση βρέθηκαν με μηδενικό σφάλμα ο άξονας και η ακτίνα του κυλίνδρου των σημείων που αποτέλεσαν την είσοδο του συστήματος

β) 60 τυχαία σημεία που ανήκουν σε κύλινδρο με άξονα την ευθεία $\varepsilon_0 = xx'$, ο οποίος εκτείνεται από το $x=0$ έως το $x=5$ και έχει ακτίνα $R_0=1$

Το αντίστοιχο ιστόγραμμα αποστάσεων:



Σχήμα 4. 2 Ιστόγραμμα των αποστάσεων των ανά δύο σημείων που ανήκουν στον κύλινδρο της περίπτωσης β)

Με βάση το ιστόγραμμα αποστάσεων προέκυψαν οι εξής αρχικές εκτιμήσεις:

$R_{αρχ}= 1.027$ και

$\epsilon_{αρχ}= [0.0143 \ 0.008 \ 0.0259; 4.868 \ 0.013 \ -0.0088]$ (η ευθεία με μια καλή προσέγγιση μπορεί να θεωρηθεί η ευθεία xx')

Με την ελαχιστοποίηση προέκυψαν:

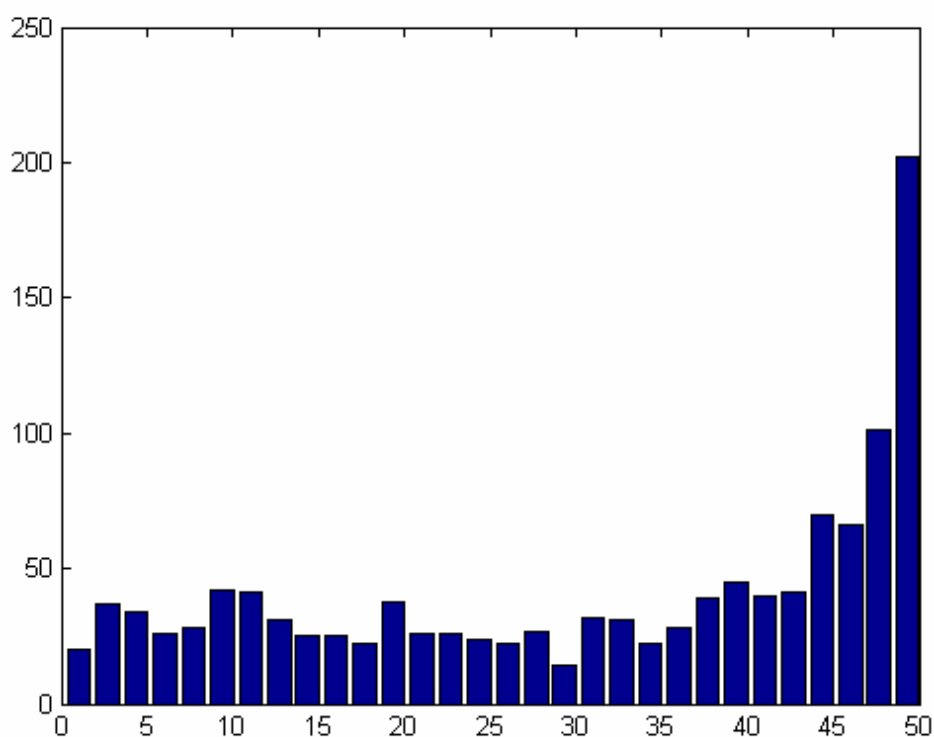
$R_{min}=1$ και

$\epsilon_{min}= [0.029 \ 0 \ 0; 4.88 \ 0 \ 0]$ (η ευθεία αυτή αποτελεί τον άξονα xx')

Παρατηρούμε, και εδώ, ότι με την ελαχιστοποίηση βρέθηκαν με μηδενικό σφάλμα ο άξονας και η ακτίνα του κυλίνδρου των σημείων που αποτέλεσαν την είσοδο του συστήματος

γ) 50 τυχαία σημεία που ανήκουν σε κύλινδρο με άξονα την ευθεία $\epsilon_0=xx'$, ο οποίος εκτείνεται από το $x=0$ έως το $x=5$ και έχει ακτίνα $R_0=25$

Το αντίστοιχο ιστόγραμμα αποστάσεων:



Σχήμα 4. 3 Ιστόγραμμα των αποστάσεων των ανά δύο σημείων που ανήκουν στον κύλινδρο της περίπτωσης γ)

Με βάση το ιστόγραμμα αποστάσεων προέκυψαν οι εξής αρχικές εκτιμήσεις

$R_{αρχ}= 24.66$

$\epsilon_{αρχ}= [3.19 \ -2.04 \ 24.87; 2.485 \ 1.299 \ -24.95]$

Μετά την ελαχιστοποίηση προέκυψαν:

$$R_{min}=25$$

$$\varepsilon_{min} = [4.906 \ 0 \ 0; 0.215 \ 0 \ 0] \text{ (η ευθεία αυτή αποτελεί τον άξονα } xx')$$

Παρατηρούμε, και εδώ, ότι με την ελαχιστοποίηση βρέθηκαν με μηδενικό σφάλμα ο άξονας και η ακτίνα του κυλίνδρου των σημείων που αποτέλεσαν την είσοδο του συστήματος. Αν και η αρχική εκτίμηση για τον άξονα του κυλίνδρου δεν ήταν καλή (λόγω του ότι ο κύλινδρος είναι αρκετά φαρδύς, $R \gg$ μήκος κυλίνδρου) ο αλγόριθμος ελαχιστοποίησης λόγω καλής αρχικής εκτίμησης για την ακτίνα μπόρεσε να βρει τη ζητούμενη ευθεία και τη ζητούμενη ακτίνα.

4.2 Εύρεση του άξονα του διαμεμβρανικού β-βαρελιού με χρήση του αλγόριθμου εύρεσης άξονα και ακτίνας κυλίνδρου και Γενετικού Αλγορίθμου.

Στην παρούσα παράγραφο παρουσιάζεται η μέθοδος που αναπτύχθηκε για την εύρεση άξονα και ακτίνας του διαμεμβρανικού β-βαρελιού.

Είσοδο στο σύστημα αποτελούν οι θέσεις των Ca ατόμων των αμινοξέων που ανήκουν στους β-κλώνους της β-πτυχωτής επιφάνειας που δημιουργεί το βαρέλι. Τα αμινοξέα που συνιστούν τους β-κλώνους του β-βαρελιού είναι αυτά που υποδεικνύει το πρόγραμμα DSSP, το οποίο εκτελείται πριν την εφαρμογή του αλγορίθμου για την εύρεση του άξονα και της ακτίνας του κυλίνδρου. Το πρόγραμμα DSSP [55] είναι ένα δημοφιλές πρόγραμμα που δέχεται ως είσοδο το .pdb αρχείο μιας πρωτεΐνης και μεταξύ άλλων υπολογίζει τη δευτεροταγή δομή των αμινοξέων της πρωτεΐνης. Το πρόγραμμα DSSP χαρακτηρίζει καθένα από τα αμινοξέα με ένα από τα εξής γράμματα ανάλογα με τη δευτεροταγή δομή στην οποία συμμετέχει:

- H = alpha helix
- B = residue in isolated beta-bridge
- E = extended strand, participates in beta ladder
- G = 3-helix (3/10 helix)
- I = 5 helix (pi helix)
- T = hydrogen bonded turn
- S = bend

Στην Εικόνα 4.1 φαίνεται ένα μέρος της εξόδου του προγράμματος DSSP για την πρωτεΐνη 1E54.pdb.

#	RESIDUE	AA	STRUCTURE	BP1	BP2	ACC	N-H-->O	O-->H-N	N-H-->O	O-->H-N	TCO	KAPPA	ALPHA	PHI	PSI	X-CA	Y-CA	Z-CA	
1	1	A	X			0	0, 0.0	2, -0.4	0, 0.0	46, -0.1	0.000	360.0	360.0	360.0	-142.0	54.7	14.6	13.5	
2	2	A	S	+	0	0	43, -0.2	2, -0.3	44, -0.1	43, -0.2	-0.944	360.0	145.1	-119.5	136.1	56.2	17.3	16.0	
3	3	A	S	E	-A	44	0A 0	41, -2.0	41, -2.2	2, -0.4	-0.988	34.1	-138.2	-163.8	160.1	54.5	20.3	17.5	
4	4	A	V	E	-Ab	43	336A 0	331, -0.5	333, -2.2	2, -0.4	-0.986	20.3	-164.7	-127.1	138.4	54.2	22.6	20.6	
5	5	A	T	E	-Ab	42	337A 0	37, -3.6	37, -2.7	2, -0.5	-0.982	15.0	-166.1	-131.8	135.4	51.0	23.9	22.0	
6	6	A	L	E	+A	41	0A 0	331, -2.7	333, -0.3	2, -0.4	-0.985	36.6	149.4	-115.2	120.6	50.1	26.6	24.5	
7	7	A	F	E	+A	40	0A 0	33, -2.7	33, -2.8	2, -0.5	-0.860	15.5	154.4	-145.9	178.9	46.5	26.2	25.6	
8	8	A	G	E	-A	39	0A 0	324, -2.2	324, -2.4	2, -0.3	-0.989	29.8	-117.1	175.1	-176.4	44.0	26.7	28.3	
9	9	A	I	E	-AC	38	331A 0	29, -2.5	29, -2.0	2, -0.3	-0.971	10.6	-161.7	-149.3	129.2	40.4	27.2	29.5	
10	10	A	V	E	+C	0	330A 0	320, -3.1	320, -3.0	2, -0.3	-0.961	18.9	169.7	-115.4	112.9	38.9	30.1	31.3	
11	11	A	D	E	+C	0	329A 0	-2, -0.6	21, -2.9	318, -0.3	2, -0.3	-0.953	6.1	162.3	-126.1	107.5	35.7	29.1	33.0
12	12	A	T	E	+DC	31	328A 0	316, -1.7	316, -2.2	2, -0.5	-0.947	8.4	171.0	-127.0	150.9	34.2	31.6	35.4	
13	13	A	N	E	-DC	30	327A 0	17, -2.6	17, -2.4	2, -0.3	-0.862	37.7	-108.0	-147.2	177.5	30.6	31.9	36.7	
14	14	A	V	E	-DC	29	326A 0	312, -2.7	312, -2.5	2, -0.3	-0.960	46.6	-171.1	-110.5	130.5	28.2	33.5	39.1	
15	15	A	A	E	-DC	28	325A 0	13, -2.8	13, -2.0	2, -0.5	-0.894	16.5	-173.3	-129.1	160.7	27.2	31.0	41.8	
16	16	A	Y	E	-DC	27	324A 0	308, -2.3	308, -2.9	2, -0.3	-0.981	6.8	-177.4	-153.9	135.8	24.7	30.6	44.6	
17	17	A	V	E	-DC	26	323A 0	9, -2.5	9, -3.0	2, -0.3	-0.997	26.5	-137.7	-132.0	133.3	24.1	28.1	47.4	
18	18	A	N	S	S+	0	0 0	304, -2.3	2, -0.3	2, -0.4	305, -0.1	0.630	81.5	13.3	-68.8	-14.2	21.1	28.5	49.7
19	19	A	K		-	0	0 0	303, -0.3	6, -0.2	7, -0.1	-1, -0.2	-0.954	55.7	-164.6	-165.0	143.4	23.1	27.6	52.9
20	20	A	D	>	-	0	0 0	4, -1.3	3, -1.8	2, -0.3	296, -0.4	-0.366	59.9	-69.7	-109.5	-167.2	26.6	27.1	54.3
21	21	A	A	T	S+	0	0 0	293, -0.3	294, -0.1	1, -0.3	293, -0.0	0.667	137.3	49.9	-61.4	-14.9	27.7	25.4	57.6
22	22	A	A	T	S+	0	0 0	2, -0.2	-1, -0.3	294, -0.1	3, -0.1	0.492	118.8	-109.9	-100.3	-8.0	26.2	28.4	59.4
23	23	A	G	S	<	0	0 0	-3, -1.8	2, -0.2	1, -0.3	-2, -0.1	0.434	77.2	125.9	93.7	-0.9	22.8	28.3	57.6
24	24	A	D	-	-	0	0 0	1, -0.1	-4, -1.3	-5, -0.1	-1, -0.3	-0.609	62.5	-114.0	-91.5	153.8	23.3	31.4	55.5
25	25	A	S	S	-	0	0 0	-6, -0.2	2, -0.4	-2, -0.2	-7, -0.2	-0.472	30.2	-136.2	-80.4	136.4	23.1	31.6	51.7
26	26	A	R	E	+D	17	0A 0	-9, -3.0	-9, -2.5	-2, -0.1	2, -0.3	-0.999	17.1	165.3	-139.3	134.8	26.2	32.4	49.8
27	27	A	Y	E	+D	16	0A 0	-2, -0.4	2, -0.3	-11, -0.2	-11, -0.2	-0.992	16.1	104.0	-147.8	153.6	26.7	33.9	46.3
28	28	A	G	E	+D	15	0A 0	-13, -2.0	-13, -2.8	-2, -0.3	2, -0.6	-0.983	67.4	-35.3	170.2	-159.8	29.4	35.4	44.2
29	29	A	L	E	+D	14	0A 0	-2, -0.3	-15, -0.3	-15, -0.2	2, -0.2	-0.840	62.1	167.0	-95.9	119.9	31.8	35.0	41.3
30	30	A	G	E	+D	13	0A 0	-17, -2.4	-17, -2.6	-2, -0.6	2, -0.3	-0.508	20.8	-125.7	-119.4	-168.5	33.2	31.4	41.1
31	31	A	T	E	+D	12	0A 0	277, -2.3	-19, -0.3	-19, -0.3	277, -0.3	-0.892	64.6	22.6	-140.5	167.7	35.1	29.3	38.6
32	32	A	S	+	0	0 0	-21, -2.9	79, -0.8	-2, -0.3	-1, -0.2	0.846	55.3	148.9	45.1	58.9	35.2	26.0	36.6	
33	33	A	G	S	S+	0	0 0	2, -0.1	80, -0.2	77, -0.1	77, -0.1	0.789	83.0	18.6	-86.9	-29.1	31.5	25.1	36.5
34	34	A	A	S	S+	0	0 0	1, -0.5	2, -0.3	294, -0.1	76, -0.1	0.748	140.1	4.8	-107.8	-40.9	31.6	23.3	33.1
35	35	A	S	S	S-	0	0 0	74, -0.1	-1, -0.5	3, -0.0	2, -0.3	-0.992	85.3	-98.7	-144.4	145.0	35.4	22.6	32.7
36	36	A	T	-	0	0 0	-2, -0.3	2, -0.1	72, -0.2	-3, -0.0	-0.486	43.3	-117.1	-66.8	127.9	38.3	23.2	35.1	
37	37	A	S	+	0	0 0	24, -0.3	24, -1.5	-26, -0.3	2, -0.3	-0.399	50.4	160.0	-64.0	140.8	40.0	26.4	34.2	

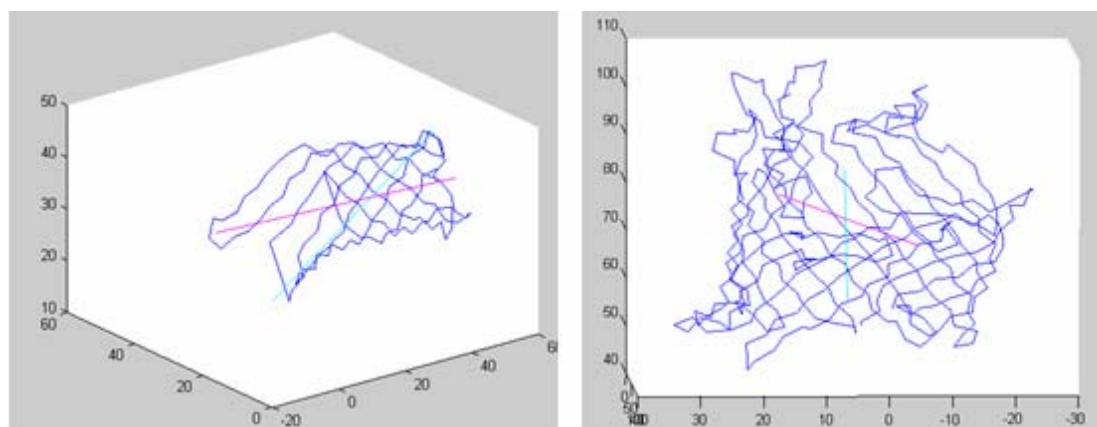
Εικόνα 4. 1 Μέρος της εξόδου του προγράμματος DSSP για την πρωτεΐνη 1E54.pdb

Με βάση τα αποτελέσματα της εξόδου του DSSP στην Εικόνα 4.1 συμπεραίνουμε ότι αμινοξικές ακολουθίες 3-17 και 26-31, που έχουν τη δευτεροταγή δομή E που αντιστοιχεί σε β-κλώνο, αποτελούν τους 2 πρώτους κλώνους του βαρελιού της πρωτεΐνης 1E54.pdb.

Ο άξονας και η ακτίνα του βαρελιού, που όπως προαναφέρθηκε προσεγγίζεται από κύλινδρο, γίνεται με χρήση του αλγόριθμου εύρεσης άξονα και ακτίνας κυλίνδρου, που περιγράφηκε αμέσως παραπάνω, και του Γενετικού Αλγορίθμου (ΓΑ). Παρακάτω, παρουσιάζονται οι λόγοι για τους οποίους είναι απαραίτητος ο συνδυασμός του αλγόριθμου για την εύρεση του άξονα και της ακτίνας κυλίνδρου και του ΓΑ καθώς και ο τρόπος με τον οποίο συνδυάζονται.

Ανάγκη συνδυασμού του αλγόριθμου για την εύρεση του άξονα και της ακτίνας κυλίνδρου και του ΓΑ: Ο ΓΑ αλγόριθμος χρησιμοποιήθηκε ως μέθοδος βελτιστοποίησης για την ανάδειξη εκείνων των σημείων (από όλα που αντιστοιχούν στις θέσεις των Ca ατόμων των αμινοξέων των β-κλώνων) που ως είσοδο στον αλγόριθμο εύρεσης άξονα και ακτίνας κυλίνδρου θα οδηγήσουν στη σωστή γεωμετρική μοντελοποίηση του βαρελιού. Αυτό που μας έκανε να χρησιμοποιήσουμε τον ΓΑ για μια τέτοιου είδους βελτιστοποίηση είναι το γεγονός ότι αρκετές φορές οι θέσεις των Ca ατόμων των αμινοξέων των β-κλώνων δεν αποτελούν σημεία που είναι ισοκαταναμημένα σε επιφάνεια κυλίνδρου, έστω και κατά προσέγγιση. Είναι πιθανό, για παράδειγμα, μερικοί κλώνοι να εκτείνονται περισσότερο από κάποιους άλλους, κυρίως προς το

εξωτερικό τοίχωμα της μεμβράνης και ακόμα πιο έξω από αυτό (θα μπορούσαμε κατά αντιστοιχία να φανταστούμε ένα κύλινδρο που του έχουν αφαιρεθεί σημεία από την μια πλευρά του και εκτείνεται κατά μήκος του άξονα μόνο από τη μια πλευρά του). Μια τέτοια προβληματική κατανομή των Ca ατόμων μπορεί να οδηγήσει σε μια λανθασμένη εκτίμηση για την ακτίνα του βαρελιού βασισμένη στο ιστόγραμμα των ανά δύο σημείων αποστάσεων και τελικά σε αποτυχημένη προσπάθεια εύρεσης του άξονα του βαρελιού. Στην Εικόνα 4.2, φαίνεται μια περίπτωση β-βαρελιού που η εφαρμογή του αλγορίθμου που αναπτύχθηκε για την εύρεση άξονα και ακτίνας κυλίνδρου θα ήταν αποτυχημένη χωρίς την προεπιλογή σημείων και μια περίπτωση που ο αλγόριθμος βρήκε επιτυχημένα τον άξονα χωρίς την προεπιλογή σημείων. Ο ΓΑ που αποτελεί μέθοδο βελτιστοποίησης που δεν απαιτεί την ύπαρξη καλών αρχικών συνθηκών και χρησιμοποιείται πολύ αποτελεσματικά για την επιλογή σε προβλήματα βελτιστοποίησης αποτέλεσε επιτυχημένα τη μέθοδο που επιλέχθηκε για την εύρεση εκείνων των Ca ατόμων που οι θέσεις τους ως είσοδο στον αλγόριθμο εύρεσης άξονα και ακτίνας κυλίνδρου θα οδηγήσουν στη σωστή γεωμετρική μοντελοποίηση του βαρελιού



Εικόνα 4. 2 Επιτυχημένη εύρεση του άξονα του βαρελιού (1QJP.pdb) χωρίς προ-επιλογή σημείων (αριστερά). Μη επιτυχημένη εύρεση του άξονα του βαρελιού (2OMF.pdb) χωρίς προ-επιλογή σημείων (δεξιά). Η πράσινη και κόκκινη ευθεία αντιστοιχούν στην αρχική εκτίμηση και το τελικό αποτέλεσμα αντίστοιχα κατά την εφαρμογή του αλγορίθμου εύρεσης άξονα και ακτίνας κυλίνδρου χωρίς προ-επιλογή σημείων.

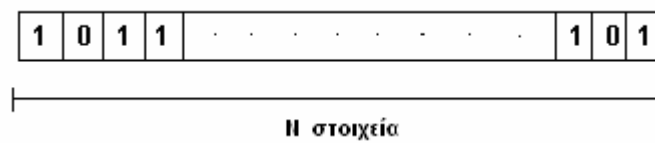
Σύντομη περιγραφή του ΓΑ: Η βασική ιδέα του ΓΑ, ο οποίος όπως προαναφέρθηκε αποτελεί μέθοδο βελτιστοποίησης που δε χρειάζεται αρχικές εκτιμήσεις για τις μεταβλητές προς βελτιστοποίηση, είναι η εξής:

Αρχικά δημιουργείται ένα τυχαίο σύνολο (γενιά) από υποψήφιες λύσεις. Καθεμία από τις υποψήφιες λύσεις ονομάζεται καταχρηστικά χρωμόσωμα. Ορίζεται μια συνάρτηση ποιότητας, η οποία χαρακτηρίζει κάθε χρωμόσωμα σε σχέση με το πόσο καλή λύση αποτελεί για την βελτιστοποίηση που επιθυμούμε. Για όλα τα χρωμοσώματα της αρχικής γενιάς υπολογίζεται η τιμή της συνάρτησης ποιότητας, ώστε να γίνει γνωστό ποιες από τις υποψήφιες λύσεις αποτελούν «καλές λύσεις» για το πρόβλημα βελτιστοποίησης. Στη συνέχεια, επιλέγονται εκείνα τα χρωμοσώματα της πρώτης γενιάς που αποτελούν «καλύτερες λύσεις», σύμφωνα με την τιμή της συνάρτησης ποιότητας. Τα επιλεγμένα χρωμοσώματα υπόκεινται στις πράξεις διασταύρωσης και μετάλλαξης (όπως περιγράφεται στο παράρτημα) με συγκεκριμένες πιθανότητες (π.χ. μπορούμε να θεωρήσουμε ότι το 80% και το 5% των επιλεγμένων χρωμοσωμάτων θα διασταυρωθούν και θα μεταλλαχθούν αντίστοιχα). Προκύπτει, έτσι, μια δεύτερη γενιά που στο σύνολο της περιέχει «καλύτερες λύσεις» για το πρόβλημα βελτιστοποίησης σε σχέση με την πρώτη. Η διαδικασία συνεχίζεται μέχρι την ικανοποίηση ενός κριτηρίου τερματισμού (όπως τη περαίωση ενός μέγιστου αριθμού γενεών ή την απόκτηση συγκεκριμένης τιμής στη συνάρτηση ποιότητας από ένα χρωμόσωμα). Στο τέλος της διαδικασίας, το χρωμόσωμα της τελευταίας γενιάς με την καλύτερη τιμή συνάρτησης ποιότητας είναι εκείνο που αποτελεί τη λύση του προβλήματος βελτιστοποίησης. Ο βασικός Γενετικός Αλγόριθμος (ΓΑ) παρουσιάζεται αναλυτικά στο παράρτημα.

Συνδυασμός του αλγόριθμου για την εύρεση του άξονα και της ακτίνας κυλίνδρου και του ΓΑ για την εύρεση του άξονα και της ακτίνας του β-βαρελίου: Ο αλγόριθμος για την εύρεση του άξονα κυλίνδρου και της ακτίνας του δεδομένου ενός αριθμού σημείων που ανήκουν σε αυτόν και ο ΓΑ συνδυάστηκαν για την εύρεση του άξονα και της ακτίνας του β-βαρελίου. Όπως, προαναφέρθηκε ο ΓΑ αναδεικνύει εκείνα τα Ca άτομα όλων των αμινοξέων της β-πτυχωτής επιφάνειας που οι θέσεις τους ως είσοδο στον αλγόριθμο εύρεση του άξονα και ακτίνας κυλίνδρου, θα οδηγήσουν στη σωστή εύρεση του άξονα και της ακτίνας του βαρελίου.

Έστω N ο αριθμός των Ca άτομα όλων των αμινοξέων της β-πτυχωτής επιφάνειας που δημιουργεί το βαρέλι. Σκοπός είναι να επιλεγθούν M από τα N σημεία που θα οδηγήσουν στη σωστή εύρεση του άξονα και της ακτίνας του

βαρελιού. Κάθε χρωμόσωμα του ΓΑ που αποτελεί υποψήφια λύση στο πρόβλημα ανάδειξης των M ατόμων από τα N είναι ένας μονοδιάστατος πίνακας διάστασης N . Κάθε $i^{\text{οστό}}$ στοιχείο του πίνακα είναι 1 ή 0, ανάλογα με τον αν η θέση του αντιστοίχου Ca ατόμου περιέχεται στη λύση ή όχι. Π.χ. στην υποψήφια λύση του Σχήματος 4.4 η θέση του $1^{\text{ου}}$ Ca ατόμου συμπεριλαμβάνεται ενώ το $2^{\text{ο}}$ δε συμπεριλαμβάνεται. Σημειώνεται, εδώ, ότι αναφερόμαστε σε 1° , 2° κ.τ.λ. Ca άτομο ανάλογα με τη σειρά με την οποία αναφέρονται τα αμινοξέα στη περιγραφή των β -κλώνων της β -πτυχωτής επιφάνειας στην έξοδο του προγράμματος του DSSP.



Σχήμα 4. 4 Ένα χρωμόσωμα υποψήφιας λύσης του ΓΑ για την ανάδειξη M από τα N σημεία που θα οδηγήσουν στη σωστή εύρεση του άξονα και της ακτίνας του βαρελιού. Η θέση του $1^{\text{ου}}$ Ca ατόμου συμπεριλαμβάνεται στην υποψήφια λύση του χρωμοσώματος, ενώ του $2^{\text{ου}}$ δε συμπεριλαμβάνεται.

Για κάθε χρωμόσωμα υπολογίζεται η συνάρτηση ποιότητας του ως εξής: Με βάση τις θέσεις των Ca ατόμων που υποδεικνύει το χρωμόσωμα, βρίσκονται οι αρχικές εκτιμήσεις για τον άξονα και την ακτίνα του βαρελιού, σύμφωνα με τον αλγόριθμο εύρεσης άξονα και ακτίνας ενός κυλίνδρου δεδομένου ενός αριθμού σημείων του. Χρησιμοποιώντας ως είσοδο την αρχική εκτίμηση του άξονα και της ακτίνας υπολογίζεται ο άξονας και η ακτίνα που υποδεικνύει η υποψήφια λύση του χρωμοσώματος, κάνοντας την ελαχιστοποίηση:

$$\min \left\langle \left(\sum_{i=1}^N (dist(P_i, \varepsilon) - R)^2 \right) \right\rangle, \text{ όπου } P_i \text{ η θέση του } i^{\text{οστού}} \text{ από τα } N \text{ } Ca \text{ άτομα}$$

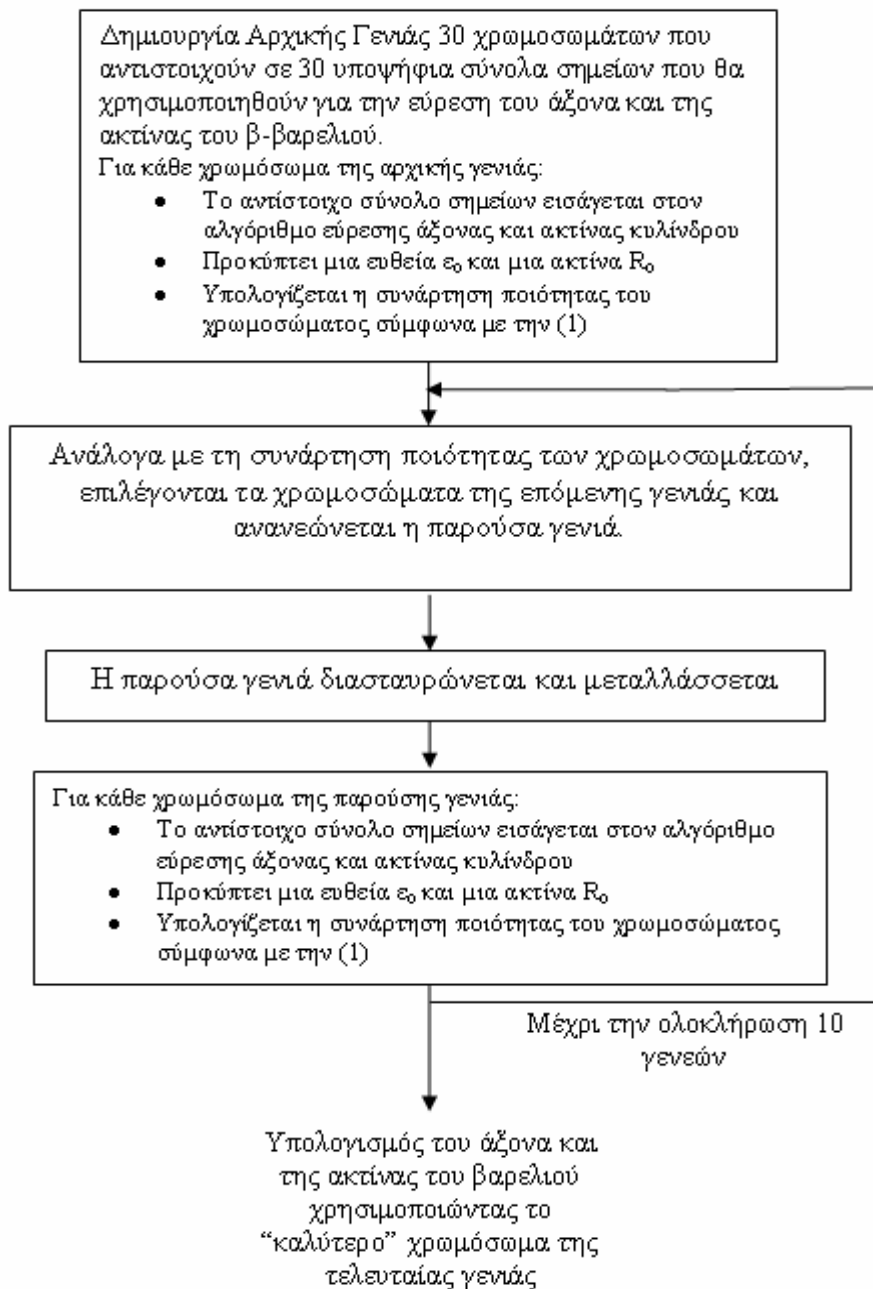
με χρήση της μεθόδου βελτιστοποίησης ελαχίστων τετραγώνων που περιγράφεται στο Παράρτημα. Σημειώνεται, εδώ, ότι στην ελαχιστοποίηση χρησιμοποιούνται οι θέσεις του 40% των N Ca ατόμων, που επιλέγονται τυχαία. Δε χρησιμοποιήθηκαν και τα N άτομα ώστε να μειωθεί το υπολογιστικό κόστος. Με βάση την ευθεία ε_0 και την ακτίνα R_0 που προέκυψαν από την παραπάνω ελαχιστοποίηση υπολογίζεται το άθροισμα

$$SSDs = \sum_{i=1}^N (dist(P_i, \varepsilon_o) - R_o)^2 \quad (1)$$

Το άθροισμα αυτό λαμβάνεται ως η τιμή της συνάρτηση ποιότητας του χρωμοσώματος. Όσο μικρότερη είναι η τιμή αυτή για ένα χρωμόσωμα, τόσο καλύτερο είναι το σύνολο των Ca ατόμων που υποδεικνύεται από το χρωμόσωμα στο πρόβλημα ανάδειξης του βέλτιστου συνόλου σημείων που θα χρησιμοποιηθούν για την εύρεση του άξονα και της ακτίνας του βαρελιού. Η τιμή αυτή μπορεί να πολύ διαφορετική από το μηδέν, όταν το χρωμόσωμα δεν υποδεικνύει καλό σύνολο σημείων για την εύρεση του άξονα και της ακτίνας του βαρελιού, ενώ αντίθετα είναι πολύ κοντά στο μηδέν όταν το χρωμόσωμα υποδεικνύει βέλτιστο σύνολο για την εύρεση του άξονα και της ακτίνας του βαρελιού.

Έχοντας ορίσει τη συνάρτηση ποιότητας των χρωμοσωμάτων του ΓΑ είναι δυνατό να προχωρήσουμε στην εφαρμογή του ΓΑ για την εύρεση του άξονα και της ακτίνας του βαρελιού: Δημιουργείται μια αρχική γενιά που περιέχει τυχαία χρωμοσώματα (έχει επιλεγθεί κάθε γενιά, άρα και η αρχική, να περιέχει 30 τυχαία χρωμοσώματα). Για καθένα από τα 30 χρωμοσώματα υπολογίζεται η τιμή της συνάρτησης ποιότητας, όπως περιγράφηκε παραπάνω. Ανάλογα με τις τιμές της συνάρτησης ποιότητας των χρωμοσωμάτων, επιλέγεται ένας υποπληθυσμός χρωμοσωμάτων που θα χρησιμοποιηθούν στη δημιουργία της επόμενης γενιάς. Τα χρωμοσώματα του υποπληθυσμού διασταυρώνονται μεταξύ τους και μεταλλάσσονται. Προκύπτει η δεύτερη γενιά, τα χρωμοσώματα της οποίας υποδεικνύουν καλύτερα σύνολα θέσεων Ca ατόμων ως εισόδους στον αλγόριθμο εύρεσης άξονα και ακτίνας κυλίνδρου. Η διαδικασία επαναλαμβάνεται μέχρι και τη δημιουργία 10 γενιών συνολικά. Το χρωμόσωμα της τελευταίας γενιάς με τη μικρότερη τιμή της συνάρτησης ποιότητας (αυτό που υποδεικνύει βέλτιστο σύνολο για την εύρεση του άξονα και της ακτίνας του βαρελιού) είναι και η έξοδος του ΓΑ. Το σύνολο σημείων που υποδεικνύει αυτό το χρωμόσωμα είναι αυτό που θα χρησιμοποιηθεί για την εύρεση του άξονα και της ακτίνας του βαρελιού.

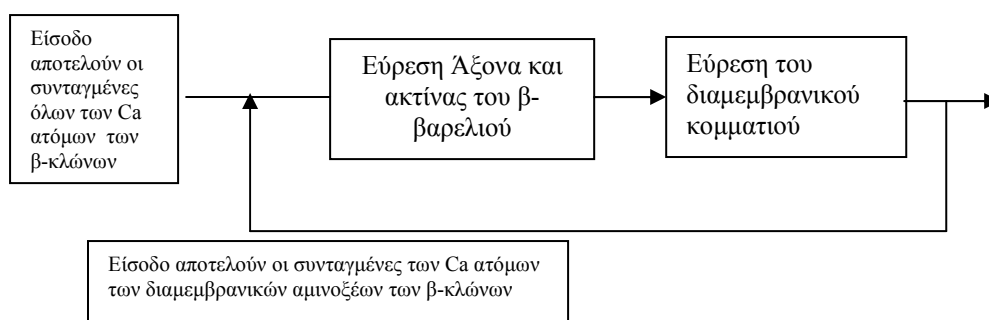
Το διάγραμμα του συνολικού αλγορίθμου για την εύρεση του άξονα και της ακτίνας του β-βαρελιού μια διαμεμβρανικής πρωτεΐνης φαίνεται στο Σχήμα 4.5.



Σχήμα 4. 5 Αλγόριθμος για την εύρεση του άξονα και της ακτίνας του β-βαρελιού

Όπως έχει ήδη αναφερθεί οι αλγόριθμοι που υλοποιήθηκαν στην παρούσα εργασία για τη γεωμετρική μοντελοποίηση των β-βαρελιών και την εύρεση του διαμεμβρανικού κομματιού των αντίστοιχων πρωτεϊνών εφαρμόστηκαν σε ένα μη πλεοναστικό σύνολο 22 πρωτεϊνών (Πίνακας 3.1). Πρέπει να αναφερθεί σε αυτό το σημείο ότι ο αλγόριθμος εύρεσης άξονα και ακτίνας όπως περιγράφηκε οδήγησε στις 18 από τις 22 πρωτεΐνες στη σωστή εύρεση άξονα, δηλαδή αυτοματοποιημένα προέκυψε άξονας που είναι σύμφωνος με αυτόν που φαντάζεται το ανθρώπινο μάτι. Στις υπόλοιπες τρεις πρωτεΐνες (1AOS.pdb,

1FEP.pdb, 1NQE.pdb) δεν προέκυψε με ακρίβεια ο αναμενόμενος άξονας, προέκυψε δηλαδή άξονας που έχει τη τάση να κλίνει περισσότερο προς τη μια πλευρά του βαρελιού και να μην ισαπέχει από όλα τους Ca άνθρακες των β-κλώνων του βαρελιού. Ωστόσο ήταν δυνατή μια περαιτέρω βελτίωση του άξονα του βαρελιού χρησιμοποιώντας τα αποτελέσματα από την εύρεση του διαμεμβρανικού κομματιού της αντίστοιχης πρωτεΐνης: χρησιμοποιώντας ως είσοδο στον αλγόριθμο εύρεσης άξονα και ακτίνας τις συντεταγμένες των Ca ατόμων των αμινοξέων που βρέθηκαν διαμεμβρανικά ήταν δυνατή η διόρθωση του άξονα. (και κατά συνέπεια και της ακτίνας). Οι θέσεις των Ca ατόμων των αμινοξέων που βρέθηκαν διαμεμβρανικά αποτέλεσαν σημεία που με μεγαλύτερη πιθανότητα στο σύνολο τους ανήκουν σε επιφάνεια κυλίνδρου σε σχέση με τις θέσεις των Ca ατόμων όλων των αμινοξέων των β-κλώνων. Η εύρεση του άξονα και της ακτίνας με αυτό τον αναδραστικό τρόπο που εφαρμόστηκε στις πρωτεΐνες 1AOS.pdb, 1FEP.pdb, 1NQE.pdb φαίνεται στο Σχήμα 4.6,

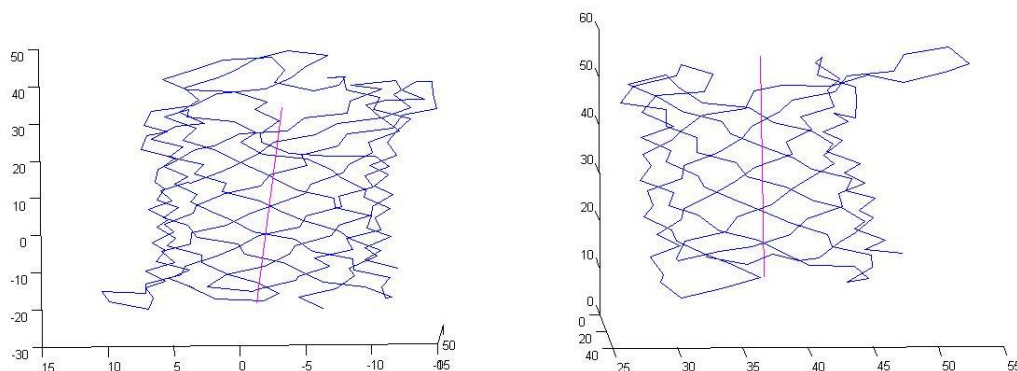


Σχήμα 4. 6 Αναδραστική βελτίωση του άξονα και της ακτίνας του β-βαρελιού στις πρωτεΐνες 1AOS.pdb, 1FEP.pdb, 1NQE.pdb

Προφανώς, στις 3 πρωτεΐνες που έγινε βελτίωση του άξονα και της ακτίνας του β-βαρελιού το διαμεμβρανικό κομμάτι επαναπροσδιορίστηκε μιας και όπως θα δούμε ο άξονας αποτελεί σημαντικό στοιχείο στον αλγόριθμο εύρεσης του διαμεμβρανικού κομματιού.

Στο σημείο αυτό δεν παρατίθενται αποτελέσματα όσον αφορά τον άξονα και την ακτίνα των πρωτεϊνών στις οποίες εφαρμόστηκε ο αλγόριθμος, μια και θα παρατεθούν συνολικά τα αποτελέσματα της παρούσης εργασίας σε επόμενο κεφάλαιο. Ενδεικτικά, παρουσιάζονται στην Εικόνα 4.3 οι άξονες όπως τους

υπολόγισε ο αλγόριθμος για τις πρωτεΐνες 1K24.pdb και 1P4T.pdb, οι οποίοι από οπτική επισκόπηση δείχνουν να έχουν υπολογιστεί σωστά,



Εικόνα 4.3 Οι άξονες των β-βαρελιών όπως τους υπολόγισε ο αλγόριθμος για τις πρωτεΐνες 1K24.pdb (αριστερά) και 1P4T.pdb (δεξιά)

4.3 Υπολογισμός της μέσης κλίσης άξονα και β-κλώνων και του Shear Number του β-βαρελιού

Η γεωμετρική μοντελοποίηση του β-βαρελιού των διαμεμβρανικών πρωτεϊνών του τύπου αυτού ολοκληρώνεται με τον υπολογισμό της μέσης κλίσης των β-κλώνων ως προς τον άξονα του β-βαρελιού και του Shear number του βαρελιού. Ήδη σε προηγούμενο κεφάλαιο έχουν περιγραφεί οι παράμετρος αυτές. Υπενθυμίζεται εδώ ότι:

αν κάνουμε τους εξής συμβολισμούς (Εικόνα 2.14 β)

S = Shear Number

a = μέση κλίση των κλώνων ως προς τον άξονα του βαρελιού

a = η απόσταση μεταξύ των $C\alpha$ ανθράκων κατά μήκος ενός κλώνου ($a=3.3 \text{ \AA}$ για παράλληλες, αντιπαράλληλες και μεικτές πτυχωτές επιφάνειες)

b = η απόσταση μεταξύ δύο διαδοχικών κλώνων ($b=4.4 \text{ \AA}$ για παράλληλες, αντιπαράλληλες και μεικτές πτυχωτές επιφάνειες)

n = ο αριθμός των β-κλώνων του βαρελιού

ισχύει ότι $\tan a = Sa / nb$ (2)

Στο τελευταίο στάδιο της γεωμετρικής μοντελοποίησης υπολογίζεται αρχικά η μέση κλίση των β-κλώνων με τον άξονα. Για κάθε κλώνο υπολογίζεται η ευθεία που προκύπτει αν ενώσουμε τους $C\alpha$ άνθρακες του πρώτου και του τελευταίου αμινοξέος του και υπολογίζεται η γωνία αυτής της ευθείας με τον άξονα (λαμβάνεται υπ' όψη βέβαια και η φορά του κλώνου). Έχοντας υπολογίζει

όλες τις κλίσεις των κλώνων ως προς τον άξονα υπολογίζεται η μέση κλίση (α). Στη συνέχεια υπολογίζεται ο Shear Number (S) από τη σχέση (2). Σε περίπτωση που το S δεν προκύψει άρτιος αριθμός στρογγυλοποιείται στον κοντινότερο άρτιο αριθμό. Σημειώνεται εδώ ότι με αυτό τον τρόπο υπολογίζεται θεωρητικά το S, και όχι διατρέχοντας το βαρέλι όπως περιγράφεται στην παράγραφο 2.2.2.2.

Αναλυτικά αποτελέσματα για τη μέση κλίση και το Shear Number των πρωτεϊνών που μελετήθηκαν παρουσιάζονται στο Κεφάλαιο 6.

5. Εύρεση του διαμεμβρανικού τμήματος διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού.

5.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται ο αλγόριθμος που αναπτύχθηκε για την εύρεση του διαμεμβρανικού τμήματος διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού. Η εύρεση του άξονα του β-βαρελιού όπως περιγράφηκε στο αμέσως προηγούμενο κεφάλαιο αποτελεί εργασία που πρέπει να προηγηθεί της εύρεσης του διαμεμβρανικού κομματιού μιας και η τελευταία στηρίζεται στην εξαγωγή συμπερασμάτων από τη μελέτη της κατανομής αμινοξέων κατά μήκος του άξονα του β-βαρελιού.

Ήδη έχουν περιγραφεί εργασίες στο 3^ο κεφάλαιο, οι οποίες σκοπό έχουν την εύρεση του διαμεμβρανικού κομματιού διαμεμβρανικών πρωτεϊνών δεδομένης της ακολουθίας των πρωτεϊνών ή της λυμένης δομής τους. Όλες οι εργασίες στηρίζονται στο γεγονός ότι η υδροφοβικότητα των αμινοξέων που ανήκουν στο διαμεμβρανικό κομμάτι είναι μεγαλύτερη από την υδροφοβικότητα των αμινοξέων που βρίσκονται έξω από αυτό. Αντίστοιχη λογική με προσθήκες που αφορούν στην ύπαρξη αρωματικών και φορτισμένων αμινοξέων σε συγκεκριμένες θέσεις κατά μήκος του άξονα του β-βαρελιού υιοθετείται στην παρούσα προσέγγιση.

Όπως έχει περιγραφεί στις εργασίες των Schulz [14] και Wimley [16] υπάρχουν κάποιοι κανόνες που περιγράφουν τα διαμεμβρανικά β-βαρέλια, οι οποίοι έχουν παρουσιαστεί στην παράγραφο 2.2.2.3. Από αυτούς, οι παρακάτω μπορούν να βοηθήσουν στην εύρεση του διαμεμβρανικού κομματιού:

- Η εξωτερική επιφάνεια του β-βαρελιού που αλληλεπιδρά με τη λιπιδική επιφάνεια της μεμβράνης είναι υδρόφοβη. Η εσωτερική επιφάνεια του β-βαρελιού είναι μέτριας υδροφοβικότητας. Ο διαχωρισμός σε εξωτερική και εσωτερική επιφάνεια του β-βαρελιού συνίσταται στην ύπαρξη εξωτερικών (external) και εσωτερικών (internal) αμινοξέων των β-κλώνων του β-βαρελιού. (Εύκολα μπορούν να βρεθούν ποια είναι τα εξωτερικά ή εσωτερικά αμινοξέα του β-βαρελιού

διατρέχοντας κάθε β-κλώνο και παρατηρώντας τη διαδοχική αύξηση και μείωση της απόστασης των αμινοξέων από τον άξονα του β-βαρελιού)

- Στις διεπιφάνειες εναλλαγής από διαμεμβρανικό σε μη διαμεμβρανικό κομμάτι των β-βαρελίων υπάρχουν δακτύλιοι εξωτερικών αρωματικών αμινοξέων (**Trp**: Τρυπτοφάνη, **Tyr**: Τυροσίνη, **Phe**: Φαινυλαλανίνη).
- Στις περιοχές που αντιστοιχούν σε μη διαμεμβρανική περιοχή της πρωτεΐνης υπάρχει πλεονασμός σε φορτισμένα αμινοξέα (**Lys**: Λυσίνη, **His**: Ιστιδίνη, **Glu**: Γλουταμινικό Οξύ, **Asp**: Ασπαρτικό Οξύ, **Arg**: Αργινίνη).

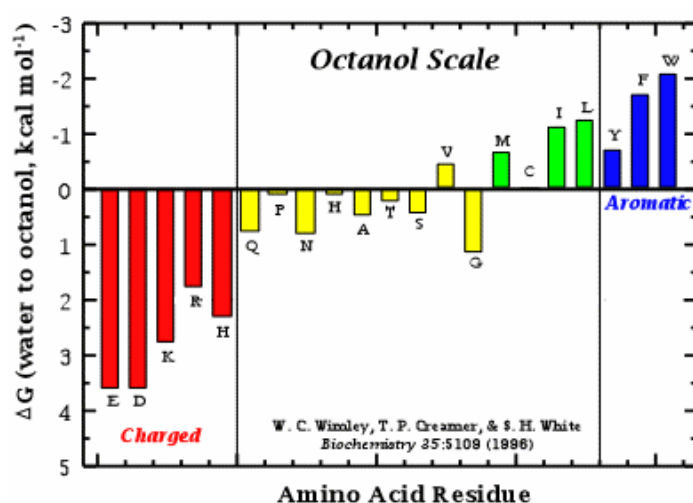
Όπως φαίνεται από τους παραπάνω κανόνες, 3 παρατηρήσεις στην κατανομή των εξωτερικών αμινοξέων του β-βαρελιού κατά μήκος του άξονα του βαρελιού μπορούν να δώσουν αρκετές πληροφορίες για τη θέση του διαμεμβρανικού κομματιού των διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού.

Η παρούσα προσέγγιση για την εύρεση του διαμεμβρανικού κομματιού βασίστηκε κυρίως στην εργασία των Wimley [16]. Άλλωστε ο σκοπός της εργασίας του Wimley ήταν η εξαγωγή συμπερασμάτων από ήδη λυμένες δομές σχετικών με τους 3 κανόνες που προαναφέρθηκαν τα οποία θα μπορούσαν να χρησιμοποιηθούν στην πρόβλεψη του διαμεμβρανικού κομματιού μόνο από την ακολουθία. Η εργασία αυτή περιγράφεται στην επόμενη παράγραφο. Στην παράγραφο 5.3, ο αναγνώστης θα έχει την ευκαιρία να δει τον αλγόριθμο εύρεσης του διαμεμβρανικού κομματιού που αναπτύχθηκε στην παρούσα εργασία.

5.2 Προσέγγιση του Wimley

Ο Wimley χρησιμοποίησε αλληλεπικαλυπτόμενα παράθυρα μήκους 5Å κατά μήκος του άξονα των β-βαρελίων στην εργασία του για να παρατηρήσει την κατανομή των εξωτερικών αμινοξέων των β-βαρελίων κατά μήκος του άξονα. Για την περιγραφή της υδροφοβικότητας των αμινοξέων έκανε χρήση μιας πειραματικής κλίμακας υδροφοβικότητας [56], η οποία χαρακτηρίζεται ως “whole residue hydrophobicity scale” μιας και λαμβάνει υπόψη τις συνεισφορές της πλευρικής αλυσίδας και του πεπτιδικού σκελετού στην υδροφοβικότητα του κάθε αμινοξέος. Για την εξαγωγή της κλίμακας έχουν μετρηθεί πειραματικά οι “whole residue” υδροφοβικότητες των αμινοξέων με μέτρηση των διαφορών της

ελεύθερης ενέργειας (ΔG) κάνοντας χρήση οκτανόλης. Σύμφωνα με την [16], αυτή η κλίμακα υδροφοβικότητας αποτελεί κλίμακα που παρέχει πιο χρήσιμα συμπεράσματα για τη μελέτη των διαμεμβρανικών πρωτεϊνών και την εύρεση του διαμεμβρανικού κομματιού σε αυτές. Παρακάτω παρατίθεται αυτή η κλίμακα υδροφοβικότητας (Εικόνα 5.1, Πίνακας 5.1). Αρνητική τιμή $\Sigma\Delta G$ (Kcal/mol) για το σύνολο των αμινοξέων που συνιστούν μια αμινοξική ακολουθία υποδεικνύουν προτίμηση της ακολουθίας να βρίσκεται σε υδρόφοβο περιβάλλον, όπως το λιπιδικό περιβάλλον της μεμβράνης. Ταυτόχρονα στον Πίνακα 5.1 παρατίθεται και η κλίμακα υδροφοβικότητας κατά Eisenberg [57] για λόγους σύγκρισης.



Εικόνα 5. 1 Κλίμακα “Whole residue” υδροφοβικότητας που εξάχθηκε με χρήση οκτανόλης.

Αμινοξύ	Υδροφοβικότητα κατά Wimley <i>et al</i> [56]	Υδροφοβικότητα κατά Eisenberg ² [57] (ΔG)
Ala	0.50	0.62
Arg	1.81	-2.53
Asn	0.85	-0.78
Asp	3.64	-0.9
Cys	-0.02	0.29
Gln	0.77	-0.85
Glu	3.63	-0.74
Gly	1.15	0.48
His	2.33	-0.4
Ile	-1.12	1.38
Leu	-1.25	1.06
Lys	2.80	-1.5
Met	-0.67	0.64
Phe	-1.71	1.19

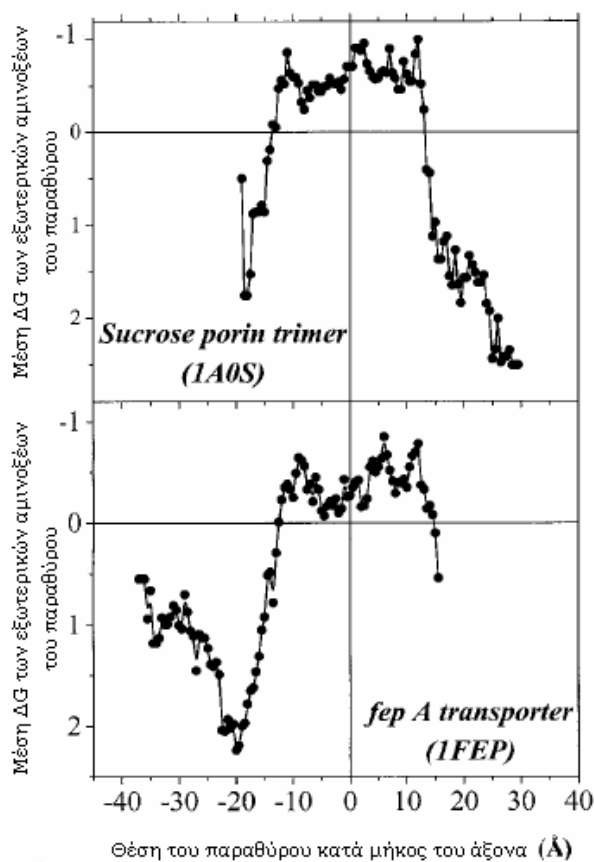
² Σε αντίθεση με τον την υδροφοβικότητα κατά Wimley *et al*, μεγαλύτερες τιμές υποδεικνύουν πιο υδρόφοβα αμινοξέα.

Pro	0.14	0.12
Ser	0.46	-0.18
Thr	0.25	-0.05
Trp	-2.09	0.81
Tyr	-0.71	0.26
Val	-0.46	1.08

Πίνακας 5. 1 Υδροφοβικότητα κατά Wimley και συνεργατών (2^η στήλη) και κατά Eisenberg (3^η στήλη)

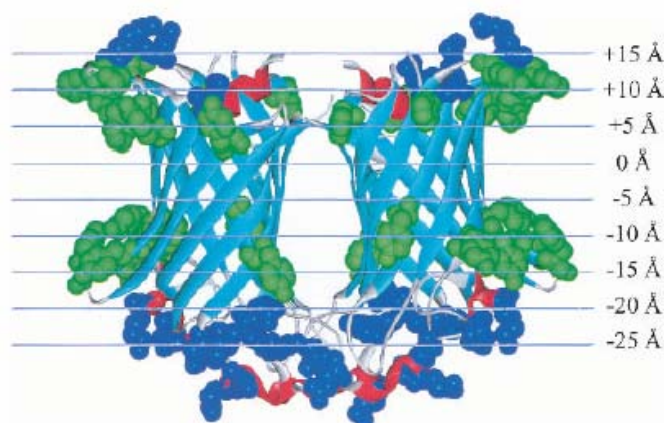
Ο Wimley παρατήρησε την κατανομή των εξωτερικών αμινοξέων κατά μήκος του άξονα του β-βαρελίου κάνοντας χρήση παραθύρων 5Å για ένα σύνολο 15 λυμένων διαμεμβρανικών πρωτεϊνών τύπου β-βαρελίου. Οι πρωτεΐνες που ανήκουν στο σύνολο αυτό παρουσιάζονται στην [16].

Στην Εικόνα 5.2 φαίνεται η μέση υδροφοβικότητα των εξωτερικών αμινοξέων στα αλληλεπικαλυπτόμενα παράθυρα κατά μήκος του άξονα του β-βαρελίου για τις πρωτεΐνες 1AOS και 1FEP. Ο κάθετος άξονας δείχνει τη μέση ΔG για το αντίστοιχο παράθυρο ενώ ο οριζόντιος άξονας τη θέση του παραθύρου κατά μήκος του άξονα του βαρελίου.



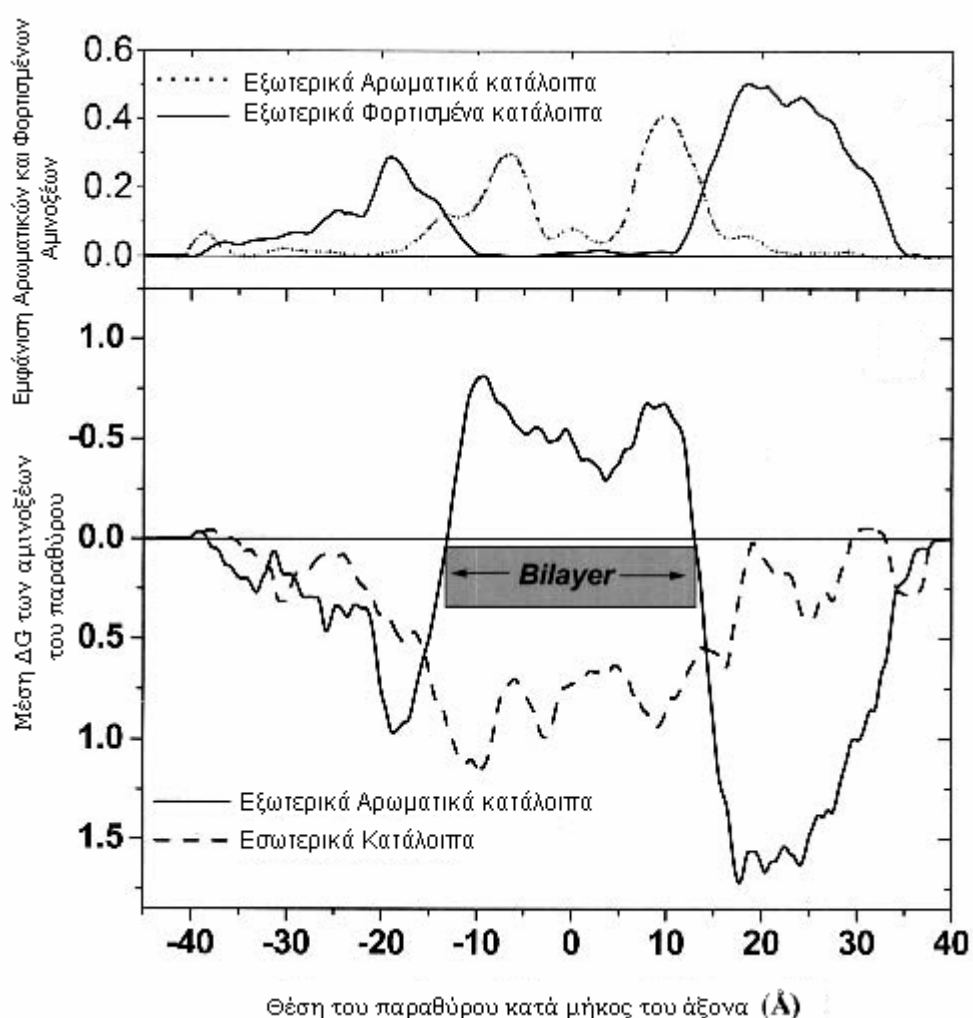
Εικόνα 5. 2 Μεταβολή της υδροφοβικότητας των εξωτερικών αμινοξέων στα αλληλεπικαλυπτόμενα παράθυρα μήκους 5Å κατά μήκος του άξονα των βαρελίων

Αρνητικές τιμές στον κάθετο άξονα υποδεικνύουν παράθυρα στα οποία τα εξωτερικά αμινοξέα είναι υδρόφοβα (κατά μέσο όρο βέβαια μιας και λαμβάνουμε υπόψη όλα τα εξωτερικά αμινοξέα που ανήκουν στο παράθυρο). Αρνητικές τιμές στον οριζόντιο άξονα υποδεικνύουν παράθυρα αμινοξέων κοντά στον περιπλαστικό χώρο (μιλάμε για διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού που απαντώνται στην εξωτερική μεμβράνη των Gram-αρνητικών βακτηρίων). Η ζώνη των παραθύρων με αρνητικές μέσες τιμές ΔG υποδεικνύουν τα αμινοξέα που ανήκουν στο διαμεμβρανικό κομμάτι, δηλαδή τα σημεία διασταύρωσης της καμπύλης της υδροφοβικότητας με τον οριζόντιο άξονα δείχνουν τα όρια του διαμεμβρανικού κομματιού [16]. Αντίστοιχο φαινόμενο έχει παρατηρηθεί για διαμεμβρανικές πρωτεΐνες τύπου α-έλικας [58]. Σημειώνεται εδώ ότι οι αποστάσεις των παραθύρων κατά μήκος του άξονα παρουσιάζονται στον οριζόντιο άξονα έτσι ώστε το μέσο της ζώνης των παραθύρων με αρνητικές μέσες ΔG να αποτελεί το σημείο με απόσταση 0 κατά μήκος του άξονα. Η θέση με απόσταση 0 κατά μήκος του άξονα ονομάζεται στην εργασία του Wimley και των συνεργατών του “bilayer midplane”. Στην Εικόνα 5.3 φαίνεται το διμερές OmpLA [40] όπου φαίνονται και οι αποστάσεις κατά μήκος του άξονα σε διάφορες θέσεις της πρωτεΐνης. Φαίνονται, επίσης, τα αρωματικά αμινοξέα (πράσινο) στις διεπιφάνειες εναλλαγής από διαμεμβρανικό σε μη διαμεμβρανικό κομμάτι καθώς και τα φορτισμένα αμινοξέα (μπλε) σε περιοχές που βρίσκονται έξω από το διαμεμβρανικό κομμάτι.



Εικόνα 5. 3 Απεικόνιση του διμερούς OmpLA [40] όπου φαίνονται και οι αποστάσεις κατά μήκος του άξονα σε διάφορες θέσεις της πρωτεΐνης.

Στην Εικόνα 5.4 φαίνονται τα αποτελέσματα της μελέτης του Wimley στις 15 λυμένες δομές. Φαίνεται η αρνητική μέση υδροφοβικότητα των εξωτερικών αμινοξέων στα αλληλεπικαλυπτόμενα παράθυρα (αρνητική ζώνη υδροφοβικότητας αντιστοιχεί στη περιοχή του μεμβρανικού λιπιδικού στρώματος), η αυξημένη παρουσία εξωτερικών αρωματικών αμινοξέων στις διεπιφάνειες της μεμβράνης και των εξωτερικών φορτισμένων αμινοξέων σε περιοχές εκτός του διαμεμβρανικού κομματιού.



Εικόνα 5. 3 Τα αποτελέσματα της μελέτης του Wimley στις 15 λυμένες δομές. Πάνω: Φαίνεται η αυξημένη παρουσία εξωτερικών αρωματικών αμινοξέων στις διεπιφάνειες της μεμβράνης και των εξωτερικών φορτισμένων αμινοξέων σε περιοχές εκτός του διαμεμβρανικού κομματιού. Κάτω: Φαίνεται η αρνητική μέση υδροφοβικότητα των εξωτερικών αμινοξέων στα αλληλεπικαλυπτόμενα παράθυρα (αρνητική ζώνη υδροφοβικότητας αντιστοιχεί στη περιοχή του μεμβρανικού λιπιδικού στρώματος) και των εσωτερικών αμινοξέων.

5.3 Αλγόριθμος για την εύρεση του διαμεμβρανικού τμήματος

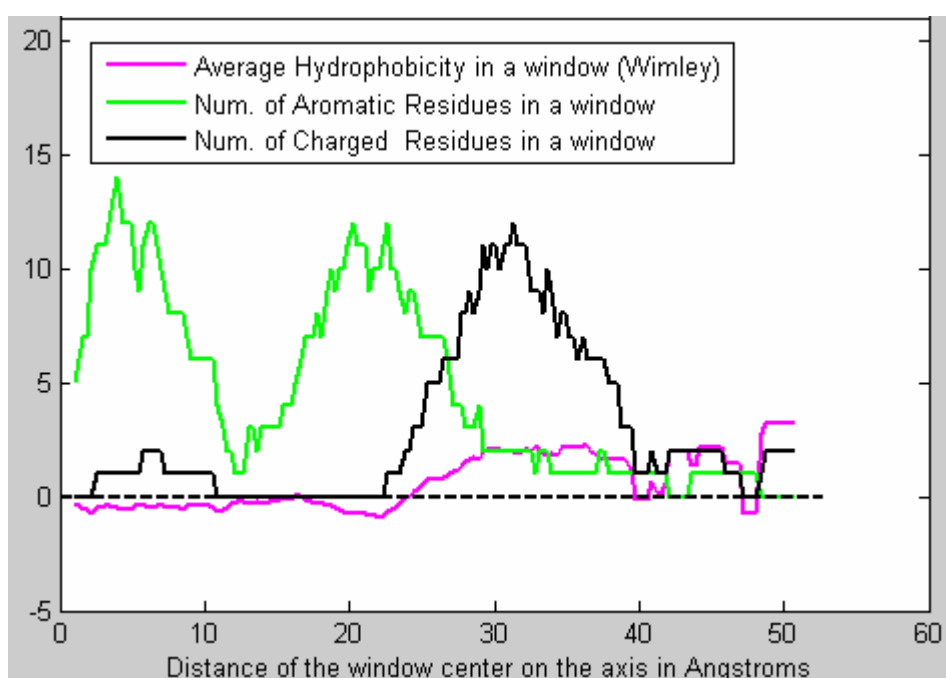
Στην παρούσα παράγραφο παρουσιάζεται ο αλγόριθμος εύρεσης του διαμεμβρανικού κομματιού που αναπτύχθηκε στην παρούσα εργασία. Ο αλγόριθμος αυτός βασίζεται στην προσέγγιση του Wimley, όπως περιγράφηκε στην προηγούμενη παράγραφο, ενώ σημειώνεται ότι η εύρεση του άξονα του β-βαρελιού (όπως περιγράφεται στο 4^ο Κεφάλαιο) πρέπει να προηγηθεί της εύρεσης του διαμεμβρανικού κομματιού.

Τα βήματα του αλγορίθμου περιγράφονται παρακάτω:

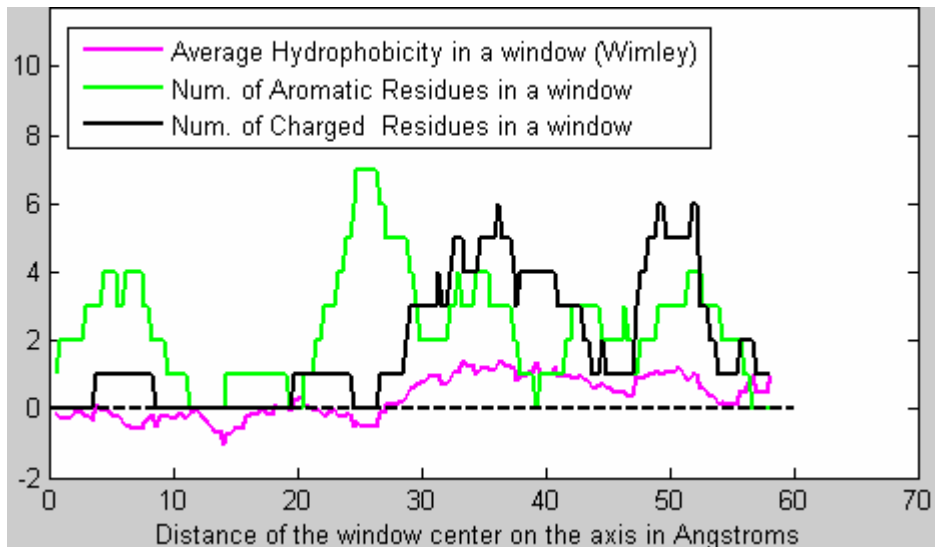
1. *Δημιουργία αλληλοεπικαλυπτόμενων παραθύρων εξωτερικών αμινοξέων των β-κλώνων κατά μήκος του άξονα του β-βαρελιού:* Κάνοντας χρήση του άξονα του βαρελιού, τα αμινοξέα των β-κλώνων του βαρελιού χαρακτηρίζονται ως εξωτερικά ή εσωτερικά. Μας ενδιαφέρουν μόνο τα εξωτερικά αμινοξέα, οι θέσεις των οποίων προβάλλονται στον άξονα. Κατά μήκος του άξονα του β-βαρελιού, δημιουργούνται αλληλοεπικαλυπτόμενα παράθυρα μήκους 5Å τα κέντρα των οποίων απέχουν απόσταση 0.30Å. Κάθε παράθυρο αντιστοιχεί σε μία λεπτή ζώνη του βαρελιού, η οποία περιέχει τα εξωτερικά αμινοξέα των οποίων οι προβολές πάνω στον άξονα συγκεντρώνονται σε ευθύγραμμο τμήμα μήκους 5Å πάνω σε αυτόν. Πρώτο παράθυρο θεωρείται το παράθυρο με κέντρο εκείνη την προβολή αμινοξέος που εμφανίζεται πρώτη πάνω στον άξονα. Μια και δεν υπάρχει προσανατολισμός του άξονα, ως “πρώτη” προβολή στον άξονα λαμβάνεται τυχαία μια από τις δύο προβολές που εμφανίζονται πιο απομακρυσμένες πάνω στον άξονα. Το πρώτο παράθυρο θεωρούμε ότι έχει τη θέση 0 πάνω στον άξονα, ενώ καθώς προχωρούμε προς το άλλο άκρο του άξονα η θέση των παραθύρων αυξάνεται κατά 0.30 Å. Για κάθε παράθυρο υπολογίζεται η μέση υδροφοβικότητα των εξωτερικών αμινοξέων κατά Wimley και συνεργατών (Πίνακας 5.1, 2^η στήλη), ο αριθμός των εξωτερικών αρωματικών αμινοξέων (**Trp**, **Tyr**, **Phe**) και εξωτερικών φορτισμένων αμινοξέων (**Lys**, **His**, **Glu**, **Asp**, **Arg**) που περιέχονται στο παράθυρο.

2. *Δημιουργία του προφίλ των αλληλοεπικαλυπτόμενων παραθύρων σχετικά με την υδροφοβικότητα των εξωτερικών αμινοξέων των β-κλώνων, την ύπαρξη εξωτερικών αρωματικών και φορτισμένων αμινοξέων κατά μήκος του άξονα:* Η υδροφοβικότητα των εξωτερικών αμινοξέων των β-κλώνων, ο αριθμός των εξωτερικών αρωματικών και φορτισμένων αμινοξέων για κάθε παράθυρο κατά

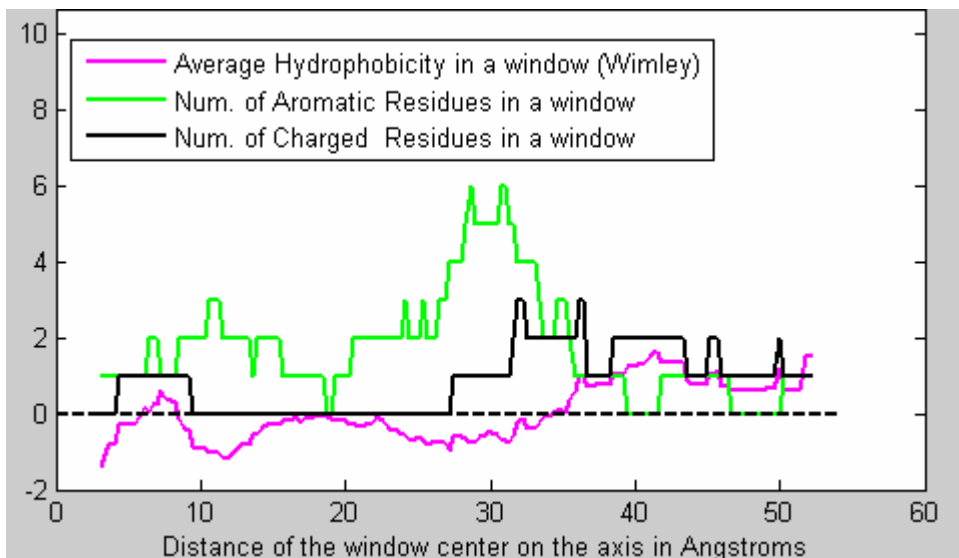
μήκος του άξονα παριστάνεται γραφικά. Με το γράφημα που προκύπτει, είναι ήδη φανερή μια αρνητική ζώνη υδροφοβικότητας, που φαίνεται να είναι το διαμεμβρανικό κομμάτι. Αυτό ενισχύεται περισσότερο από ζώνες αρωματικών αμινοξέων με κορυφές περίπου στα σημεία εναλλαγής του προσήμου της υδροφοβικότητας και από περιοχές με αυξημένη παρουσία φορτισμένων αμινοξέων έξω από την αρνητική ζώνη υδροφοβικότητας. Πρέπει να αναφερθεί εδώ ότι σε πολλές πρωτεΐνες στη ζώνη υδροφοβικότητας που μοιάζει να αντιστοιχεί στο διαμεμβρανικό κομμάτι υπάρχουν και παράθυρα με θετικές τιμές υδροφοβικότητας αλλά με χαμηλές απόλυτες τιμές. Ενδεικτικά παρουσιάζονται στις Εικόνες 5.5, 5.6, 5.7 τα γραφήματα που προέκυψαν για τις πρωτεΐνες 1NQE, 1I78 και 1K24 αντίστοιχα.



Εικόνα 5. 4 Το προφίλ της υδροφοβικότητας και της παρουσίας εξωτερικών αρωματικών και φορτισμένων αμινοξέων κατά μήκος του άξονα του βαρελιού για την πρωτεΐνη 1NQE



Εικόνα 5. 5 Το προφίλ της υδροφοβικότητας και της παρουσίας εξωτερικών αρωματικών και φορτισμένων αμινοξέων κατά μήκος του άξονα του βαρελιού για την πρωτεΐνη 1I78



Εικόνα 5. 6 Το προφίλ της υδροφοβικότητας και της παρουσίας εξωτερικών αρωματικών και φορτισμένων αμινοξέων κατά μήκος του άξονα του βαρελιού για την πρωτεΐνη 1K24

3. *Εύρεση του διαμεμβρανικού κομματιού με βάση το προφίλ των εξωτερικών αμινοξέων στα αλληλοεπικαλυπτόμενα παράθυρα κατά μήκος του άξονα:* Με βάση το προφίλ των αλληλοεπικαλυπτόμενων παραθύρων κατά μήκος του άξονα του βαρελιού ταυτοποιούνται όλα τα σημεία που μπορούν να αποτελέσουν σημεία αρχής ή τέλους του διαμεμβρανικού κομματιού (θέση αρχής θα αποτελεί σημείο αριστερά στον άξονα, ενώ θέση τέλους θα αποτελεί σημείο δεξιά στον άξονα). Ως πιθανό σημείο αρχής μπορεί να θεωρηθεί είτε το

κέντρο του πρώτου παραθύρου³ εφόσον έχει αρνητική υδροφοβικότητα είτε το κέντρο κάθε παραθύρου εφόσον αποτελεί σημείο μετάβασης από θετικές ή 0 τιμές υδροφοβικότητας σε αρνητικές τιμές υδροφοβικότητας. Ως πιθανό σημείο τέλους μπορεί να θεωρηθεί είτε το κέντρο του τελευταίου παραθύρου εφόσον έχει αρνητική υδροφοβικότητα είτε το κέντρο κάθε παραθύρου εφόσον αποτελεί σημείο μετάβασης από αρνητικές ή 0 τιμές υδροφοβικότητας σε θετικές τιμές υδροφοβικότητας. Οι ζώνες που προκύπτουν έχοντας ως αρχή οποιοδήποτε συνδυασμό αρχής και τέλους αποτελούν υποψήφιας λύσεις για τη θέση του διαμεμβρανικού κομματιού πάνω στον άξονα. Αποκλείονται εκείνες οι λύσεις που θα αντιστοιχούσαν σε υπερβολικά στενό ή πλατύ διαμεμβρανικό κομμάτι. Μάλιστα έχει ως τεθεί ελάχιστο πλάτος της μεμβράνης απόσταση ίση με 20Å, ενώ ως μέγιστο πλάτος απόσταση ίση με 32Å. Για καθεμιά από τις υποψήφιας ζώνες υπολογίζεται μια συνάρτηση ποιότητας, που χαρακτηρίζει ως πιο πιθανή λύση εκείνη τη ζώνη που είναι περισσότερο υδρόφοβη, έχει σε περιοχές γύρω από την αρχή και το τέλος της έντονη παρουσία αρωματικών αμινοξέων και έξω από αυτή είναι έντονη η παρουσία φορτισμένων αμινοξέων. Η συνάρτηση ποιότητας της κάθε ζώνης ορίζεται παρακάτω:

Έστω μια υποψήφια λύση με αρχή σημείο με απόσταση α στον άξονα και τέλος σημείο με απόσταση β στον άξονα. Η συνολική ποσότητα υδροφοβικότητας ($Sum_Hydrophobicity$) της ζώνης προκύπτει αν αθροίσουμε την υδροφοβικότητα όλων των παραθύρων (ονομάζω την υδροφοβικότητα του j -οστού παραθύρου $hydrophob_j$) που έχουν απόσταση (ονομάζω την απόσταση του j -οστού παραθύρου στον άξονα x_j) από α έως β στον άξονα:

$$Sum_Hydrophobicity_{\alpha,\beta} = \sum_{j,\alpha \leq x_j \leq \beta} hydrophob_j$$

Για κάθε j -οστό παράθυρο ονομάζω $aromatic_j$ το πηλίκο του αριθμού των αρωματικών αμινοξέων που περιέχονται στο j -οστό παράθυρο προς το συνολικό αριθμό αρωματικών αμινοξέων που περιέχονται σε όλα τα παράθυρα. Αντιστοίχως, ορίζω την ποσότητα $charged_j$. Για κάθε ζώνη ορίζεται ένας παράγοντας αρωματικών αμινοξέων, που αυξάνει καθώς αυξάνουν τα αμινοξέα που βρίσκονται στις περιοχές γύρω από την αρχή και το τέλος της ζώνης. Οι

³ Υπενθυμίζεται ότι τα σημεία του x άξονα (της απόστασης κατά μήκος του άξονα) αντιστοιχούν στα κέντρα διαδοχικών αλληλοεπικαλυπτόμενων παραθύρων

περιοχές αυτές ορίζονται ως περιοχές με μήκος $(\beta-\alpha)/2$ και με κέντρα α και β αντίστοιχα (Εικόνα 5.8). Ο παράγοντας αρωματικών αμινοξέων για ζώνη με αρχή α και β ορίζεται ο :

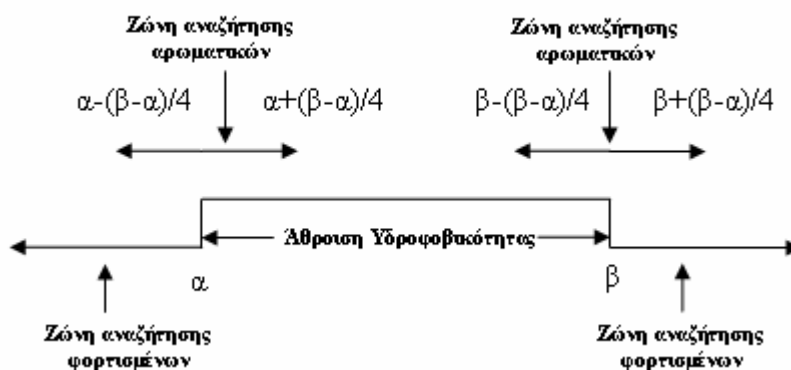
$$aromatic_factor_{\alpha,\beta} = \sum_j aromatic_j$$

$$\alpha - (\beta - \alpha) / 4 \leq x_j \leq \alpha + (\beta - \alpha) / 4 \quad \text{ή} \quad \beta - (\beta - \alpha) / 4 \leq x_j \leq \beta + (\beta - \alpha) / 4$$

Για κάθε ζώνη ορίζεται ένας παράγοντας φορτισμένων αμινοξέων, που αυξάνει καθώς αυξάνουν τα αμινοξέα που βρίσκονται στις περιοχές πριν από την αρχή και μετά το τέλος της ζώνης (Εικόνα 5.8). Ο παράγοντας φορτισμένων αμινοξέων για ζώνη με αρχή α και τέλος β ορίζεται ο :

$$charged_factor_{\alpha,\beta} = \sum_j charged_j$$

$$x_j < \alpha \quad \text{ή} \quad x_j > \beta$$



Εικόνα 5. 7 Άθροιση της υδροφοβικότητας και περιοχές αναζήτησης αρωματικών και φορτισμένων αμινοξέων για υποψήφια ζώνη με αρχή α και τέλος β .

Για καθεμιά από τις υποψήφιες ζώνες υπολογίζεται η συνάρτηση ποιότητας, η οποία περιλαμβάνει την υδροφοβικότητα της ζώνης καθώς και τους παράγοντες φορτισμένων αμινοξέων. Για ζώνη με αρχή α και τέλος β ορίζεται η εξής συνάρτηση ποιότητας:

$$Fitness_{\alpha,\beta} = (1 + 2 * charged_factor_{\alpha,\beta} + 2 * aromatic_factor_{\alpha,\beta}) * Sum_Hydrophobicity_{\alpha,\beta}$$

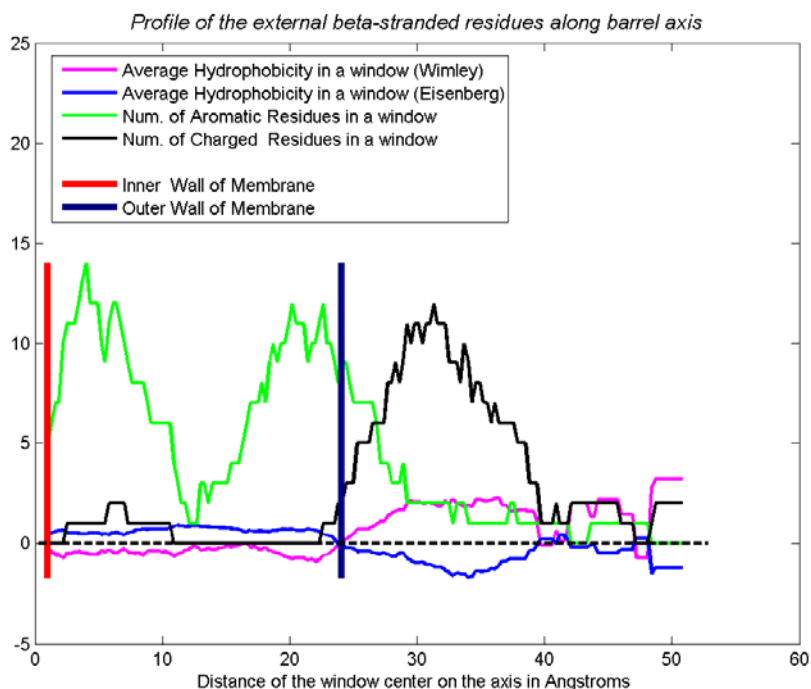
Όσο πιο αρνητική είναι η συνάρτηση ποιότητας για μια ζώνη τόσο πιο πιθανό είναι να αποτελεί τη διαμεμβρανική ζώνη του β-βαρελιού. Για παράδειγμα στην περίπτωση που έξω από την υποψήφια ζώνη υπάρχουν τα μισά φορτισμένα αμινοξέα από όσα υπάρχουν στο βαρέλι και στις περιοχές γύρω από την αρχή και το τέλος της ζώνης υπάρχουν τα μισά αρωματικά από όλα τα αρωματικά που υπάρχουν στο βαρέλι, η συνάρτηση ποιότητας είναι 3 φορές το (αρνητικό) άθροισμα της υδροφοβικότητας της υποψήφιας ζώνης.

Η ζώνη με την μικρότερη (πιο αρνητική) συνάρτηση ποιότητας αποτελεί τη διαμεμβρανική ζώνη. Η ζώνη αυτή στην ουσία αποτελεί την προβολή των διαμεμβρανικών αμινοξέων του β-βαρελιού στον άξονα. Τα δύο επίπεδα της μεμβράνης (εσωτερικό και εξωτερικό) είναι τα 2 επίπεδα που είναι κάθετα στον άξονα στα άκρα της διαμεμβρανικής ζώνης. Τα αμινοξέα που βρίσκονται στο χώρο που ορίζουν τα δύο επίπεδα είναι τα διαμεμβρανικά αμινοξέα. Είναι, λοιπόν, εύκολο να βρεθούν τα αμινοξέα που ανήκουν στο διαμεμβρανικό τμήμα (αυτά που ανήκουν στους β-κλώνους αλλά και άλλα, όπως αυτά που δημιουργούν έλικες στο εσωτερικό του β-βαρελιού). Μετά και την εύρεση των διαμεμβρανικών αμινοξέων είναι δυνατό να βρεθεί ποια από τις δύο θέσεις που αντιστοιχούν στα άκρα της διαμεμβρανικής ζώνης στον άξονα αντιστοιχεί στο εσωτερικό επίπεδο⁴ της μεμβράνης (π.χ. εκεί που τελειώνει ο περιπλασμικός χώρος για τα βαρέλια των Gram-αρνητικών βακτηρίων) ή στο εξωτερικό επίπεδο. Έτσι, μετρούνται τα αμινοξέα που βρίσκονται κάτω από το επίπεδο που αντιστοιχεί στην αρχή και εκείνα που βρίσκονται πάνω από το επίπεδο που αντιστοιχεί στο τέλος. Επειδή στο χώρο μέσα από την μεμβράνη οι κλώνοι ενώνονται με κοντές λούπες, ενώ στο χώρο έξω από αυτή ενώνονται με μακριές λούπες, θεωρείται ως εσωτερικό επίπεδο της μεμβράνης το επίπεδο εκείνο που αφήνει έξω από το διαμεμβρανικό κομμάτι λιγότερα αμινοξέα σε αριθμό. Το εναπομένον επίπεδο θεωρείται το εξωτερικό επίπεδο της μεμβράνης.

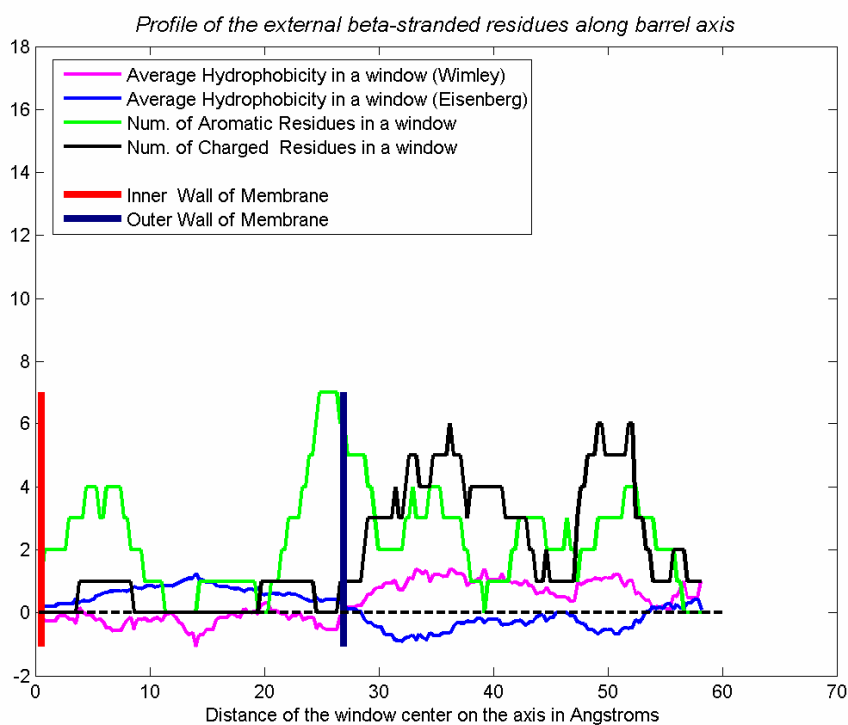
Στο σημείο αυτό παρατίθενται οι εικόνες με το προφίλ των β-βαρελιών σχετικά με την υδροφοβικότητα των εξωτερικών αμινοξέων των β-κλώνων, την ύπαρξη εξωτερικών αρωματικών και φορτισμένων αμινοξέων κατά μήκος του άξονα για τις πρωτεΐνες 1NQE, 1I78, 1K24 όπως έχουν ήδη παρουσιαστεί αλλά

⁴ Υπενθυμίζεται ότι τα δύο επίπεδα της μεμβράνης είναι τα 2 επίπεδα που είναι κάθετα στον άξονα στα άκρα της διαμεμβρανικής ζώνης

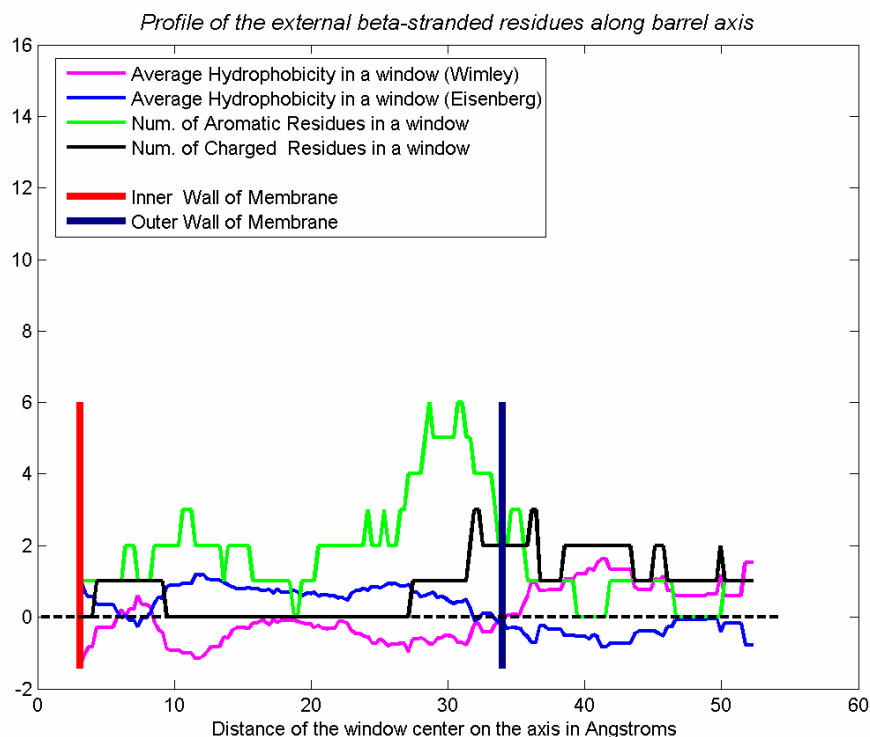
με φανερά τα δύο άκρα της μεμβράνης (Εικόνες 5.9, 5.10, 5.11 αντίστοιχα). Για λόγους πληρότητας φαίνεται και το προφίλ της υδροφοβικότητας κατά Eisenberg (Πίνακας 5.1, 3^η στήλη).



Εικόνα 5. 8 Προφίλ της 1NQE σχετικά με την υδροφοβικότητα των εξωτερικών αμινοξέων των β-κλώνων, την ύπαρξη εξωτερικών αρωματικών και φορτισμένων αμινοξέων κατά μήκος του άξονα. Είναι φανερές οι θέσεις του εσωτερικού και εξωτερικού τοιχώματος της μεμβράνης



Εικόνα 5. 9 Προφίλ της 1I78 σχετικά με την υδροφοβικότητα των εξωτερικών αμινοξέων των β-κλώνων, την ύπαρξη εξωτερικών αρωματικών και φορτισμένων αμινοξέων κατά μήκος του άξονα. Είναι φανερές οι θέσεις του εσωτερικού και εξωτερικού τοιχώματος της μεμβράνης



Εικόνα 5. 10 Προφίλ της 1K24 σχετικά με την υδροφοβικότητα των εξωτερικών αμινοξέων των β-κλώνων, την ύπαρξη εξωτερικών αρωματικών και φορτισμένων αμινοξέων κατά μήκος του άξονα. Είναι φανερές οι θέσεις του εσωτερικού και εξωτερικού τοιχώματος της μεμβράνης

Σε επόμενο κεφάλαιο παρουσιάζονται αναλυτικά για όλες τις πρωτεΐνες στις οποίες εφαρμόστηκε ο αλγόριθμος το προφίλ των πρωτεϊνών κατά μήκος του άξονα, ποιες βρέθηκαν να είναι οι θέσεις των τοιχωμάτων της μεμβράνης καθώς και ποια αμινοξέα βρέθηκαν να είναι διαμεμβρανικά.

6. Αποτελέσματα - Συζήτηση

Στην παρόν κεφάλαιο παρουσιάζονται i) τα αποτελέσματα εφαρμογής των αλγορίθμων για τη γεωμετρική μοντελοποίηση των διαμεμβρανικών πρωτεϊνών τύπου β-βαρελιού σε μη πλεοναστικό (non-redundant) σύνολο 22 πρωτεϊνών ii) τα αποτελέσματα εφαρμογής του αλγορίθμου για την εύρεση του διαμεμβρανικού κομματιού στο ίδιο σύνολο και iii) αποτελέσματα που προέκυψαν σχετικά με τη μέση υδροφοβικότητα και την κατανομή των αμινοξέων κατά μήκος του άξονα των β-βαρελιών μέσα από τη μελέτη του συνόλου των 22 πρωτεϊνών. Οι 22 πρωτεΐνες που χρησιμοποιήθηκαν στην παρούσα μελέτη παρουσιάζονται στον Πίνακα 3.1.

6.1 Γεωμετρική Μοντελοποίηση

Στο σημείο αυτό παρουσιάζονται τα αποτελέσματα εφαρμογής του αλγορίθμου γεωμετρικής μοντελοποίησης. Συγκεκριμένα στον Πίνακα 6.1 παρουσιάζονται για τις 22 πρωτεΐνες που μελετήθηκαν τα εξής στοιχεία:

- n (αριθμός κλώνων)
- ο άξονας του β-βαρελιού με τη μορφή 2 σημείων που ανήκουν σε αυτόν: το σημείο A αντιστοιχεί στην προβολή του καταλοΐπου που βρίσκεται περισσότερο κοντά στον χώρο μέσα από τη μεμβράνη (περιπλαστικό για τη περίπτωση των βακτηρίων) πάνω στον άξονα, ενώ το σημείο B ανήκει στην προβολή του καταλοΐπου που εμφανίζεται περισσότερο κοντά στην εξωτερική πλευρά της μεμβράνης πάνω στον άξονα. (ο αναγνώστης θα έχει την ευκαιρία να δει την οπτικοποίηση του άξονα σε σχέση με τη δομή τις πρωτεΐνης σε εικόνες της παραγράφου 6.2).
- Η ακτίνα του β-βαρελιού, $R_{\beta\epsilon\lambda.}$ όπως προέκυψε από τον αλγόριθμο βελτιστοποίησης (όπως περιγράφεται στην 4.2) αλλά και η ακτίνα του βαρελιού ($R_{\delta\iota\alpha\mu.}$) ως η μέση τιμή της απόστασης των διαμεμβρανικών αμινοξέων στους β-κλώνους από τον άξονα, όπως τα υπέδειξε ο αλγόριθμος που περιγράφεται στο κεφάλαιο 5.

- τη μέση κλίση (α) των β -κλώνων με τον άξονα, α , και το shear number, S^5 . Για τον υπολογισμό τους ισχύουν όσα αναφέρονται στην παράγραφο 4.3.

PDB ID	n	1° Σημείο Άξονα (A)			2° Σημείο Άξονα (B)			$R_{\beta\epsilon\lambda.}$ (Å)	$R_{\delta\iota\alpha\mu.}$ (Å)	α (μοίρες)	S
		x (Å)	y (Å)	z (Å)	x (Å)	y (Å)	z (Å)				
1P4T	8	37.07	14.19	58.53	37.19	19.67	7.52	8.431	8.266	44.2	10
1QJ8	8	71.77	16.2	74.42	42.56	12.01	22.86	7.438	7.314	38.2	8
1MM4	8	3.34	-23.91	2	-2.02	21.52	4.26	8.636	8.294	44.2	10
1QJP	8	-2.17	34.05	32.83	45.33	9.52	39.38	8.280	8.188	44.6	10
1I78	10	21.41	23.63	18.18	31.63	94.89	6.242	9.965	9.951	43.8	12
1K24	10	31.13	-1.02	-23.67	31.12	-3.80	46.82	9.626	9.905	44.5	14
1UYN	12	10.43	26.58	215.723	12.493	12.79	153.63	11.878	11.844	43.9	16
1QD5	12	25.15	10.96	42.89	46.80	50.315	65.56	12.189	12.305	46.7	16
1TLY	12	-18.09	116.02	6.04	-12.59	61.67	-6.02	12.179	11.935	47.2	18
1T1L	14	127.67	2.35	29.19	62.65	48.21	1.75	13.282	13.112	37.9	14
2POR	16	-1.907	-20.17	53.88	4.75	-22.33	-0.26	16.987	16.867	46.7	22
1PRN	16	-1.10	35.82	59.83	5.62	39.32	10.1	16.906	16.796	44.8	22
2OMF	16	5.84	46.10	6.67	-7.87	48.78	62.72	16.041	15.961	42.2	20
1OSM	16	-24.5	-19.77	49.75	3.84	-23.10	1.10	16.567	16.477	41.5	18
1E54	16	42.03	10.54	10.03	34.63	20.65	62.10	16.458	16.205	42.6	20
1PHO	16	3.50	45.66	6.09	-5.79	50.79	59.77	16.123	15.999	42.5	20
1A0S	18	-51.59	-17.87	-44.21	-49.10	-16.99	14.71	17.342	17.758	47.4	26
2MPR	18	-0.01	48.95	107.36	7.47	49.14	42.45	18.174	17.744	41.3	22
2FCP	22	94.47	59.24	68	59.40	114.64	34.50	19.971	20.641	40.8	26
1FEP	22	34.72	6.79	18.65	33.25	47.62	82.28	20.999	21.15	42.1	26
1KMO	22	24.61	9.85	-13.93	28.92	1.24	50.15	20.157	20.69	40.5	26
1NQE	22	-15.39	44.25	41.40	10.35	72.15	-8.96	21.12	21.121	43.5	28

Πίνακας 6.1 Αποτελέσματα της γεωμετρικής μοντελοποίησης των β -βαρελιών των πρωτεϊνών που μελετήθηκαν στη παρούσα εργασία

6.2 Εύρεση του διαμεμβρανικού τμήματος

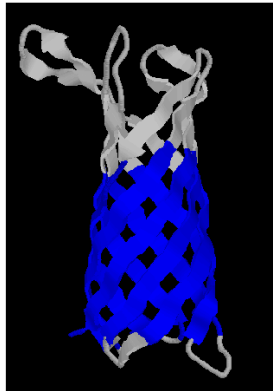
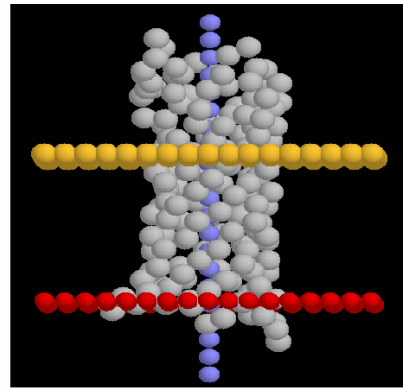
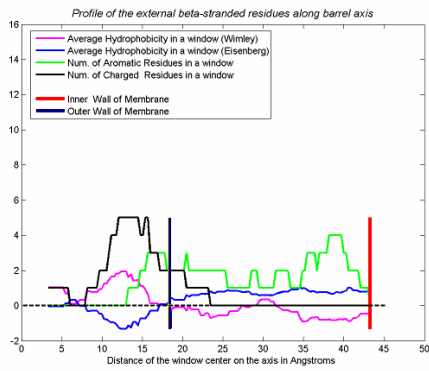
Στην παρούσα παράγραφο παρουσιάζονται τα αποτελέσματα της εφαρμογής του αλγορίθμου για την εύρεση του διαμεμβρανικού τμήματος των διαμεμβρανικών πρωτεϊνών τύπου β -βαρελιού στο σύνολο των 22 πρωτεϊνών που μελετήθηκαν στην παρούσα εργασία. Στις εικόνες 6.1-6.22, για καθεμία από τις πρωτεΐνες του Πίνακα 6.1 και με τη σειρά που εμφανίζονται στον πίνακα παρουσιάζονται:

⁵ Σημειώνεται εδώ ότι οι τιμές του S έχουν υπολογιστεί θεωρητικά όπως περιγράφεται στην παράγραφο 4.3 και όχι διατρέχοντας το βαρέλι όπως περιγράφεται στην παράγραφο 2.2.2.1. Μάλιστα για τον υπολογισμό του S έχει χρησιμοποιηθεί η τιμή $R_{\delta\iota\alpha\mu.}$ ως ακτίνα του βαρελιού.

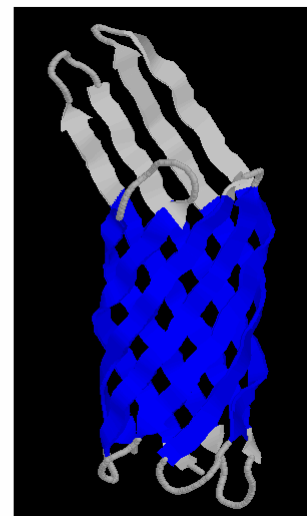
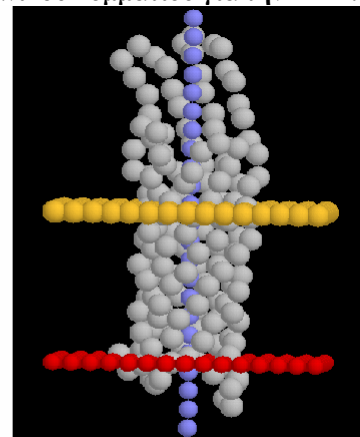
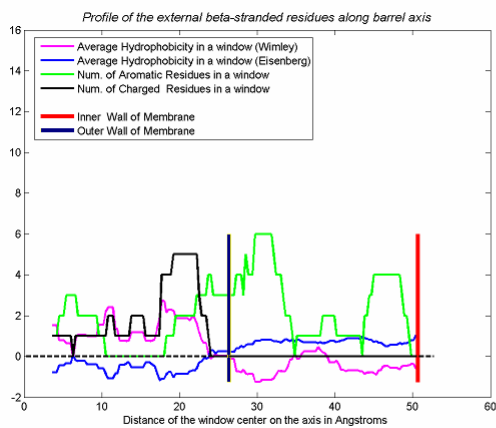
- το προφίλ των εξωτερικών αμινοξέων των β-κλώνων της πρωτεΐνης που βρίσκονται σε κυλιόμενα αλληλοεπικαλυπτόμενα παράθυρα μήκους 5Å κατά μήκος του άξονα. Για τις αποστάσεις κατά μήκος του άξονα ισχύουν όσα έχουν αναφερθεί στην παράγραφο 5.3. Το προφίλ αφορά στη μέση υδροφοβικότητα στα παράθυρα (η καμπύλη με χρώμα έντονο ροζ παρουσιάζει την υδροφιβικότητα κατά Wimley *et al* ενώ με μπλε γραμμή παρουσιάζεται για λόγους σύγκρισης η υδροφοβικότητα κατά Eisenberg) καθώς και στην ύπαρξη αρωματικών (ο αριθμός τους παρουσιάζεται με πράσινη καμπύλη) και φορτισμένων αμινοξέων (ο αριθμός τους παρουσιάζεται με μαύρη καμπύλη) κατά μήκος του άξονα. Το διαμεμβρανικό κομμάτι της πρωτεΐνης προσδιορίζεται από το προφίλ αυτό σύμφωνα με όσα έχουν αναφερθεί στο 5^ο κεφάλαιο. Σε κάθε εικόνα που παρουσιάζει το προφίλ των εξωτερικών αμινοξέων κατά μήκος του άξονα, παρουσιάζονται με κόκκινη και μπλε στήλη τα σημεία του άξονα όπου η εσωτερική και εξωτερική πλευρά της μεμβράνης τον τέμνουν αντίστοιχα (πάνω αριστερά).
- Οι Ca άνθρακες σε “space-fill” μοντέλο (γκρι σφαίρες), ο άξονας του βαρελιού (μπλε γραμμή), το εσωτερικό (κόκκινο στρώμα) και το εξωτερικό (κίτρινο στρώμα) της μεμβράνης (πάνω δεξιά).
- η πρωτεΐνη με τη μορφή “cartoons”, όπου με χρώμα μπλε έχουν χρωματιστεί τα διαμεμβρανικά αμινοξέα όπως τα υποδεικνύει ο αλγόριθμος που υλοποιήθηκε στην παρούσα εργασία (κάτω αριστερά)⁶.
- η πρωτεΐνη με τη μορφή “cartoons”, όπου με χρώμα μπλε έχουν χρωματιστεί τα διαμεμβρανικά αμινοξέα όπως τα υποδεικνύει το εργαλείο (TM-DET) του Tusnadi και συνεργατών [31] (κάτω δεξιά)⁷. Αυτή η παράθεση γίνεται για λόγους σύγκρισης .

⁶ Όταν η πρωτεΐνη είναι διμερές ή τριμερές, παρουσιάζονται όλες οι αλυσίδες. Ο αλγόριθμος έχει εφαρμοστεί σε μία αλυσίδα και στις υπόλοιπες αλυσίδες θεωρούμε ότι διαμεμβρανικά αμινοξέα είναι αυτά που υποδεικνύει ο αλγόριθμος όταν εφαρμόστηκε στην 1^η αλυσίδα (μια και οι αλυσίδες είναι ταυτόσημες στην πρωτοταγή τους δομή)

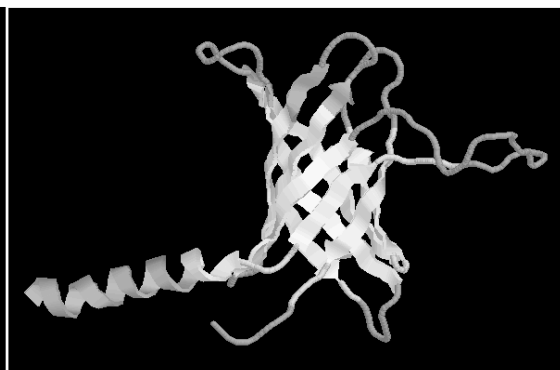
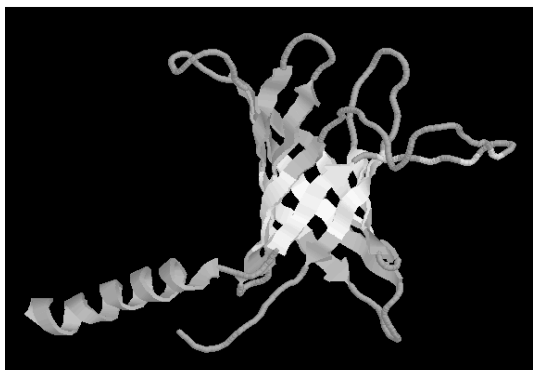
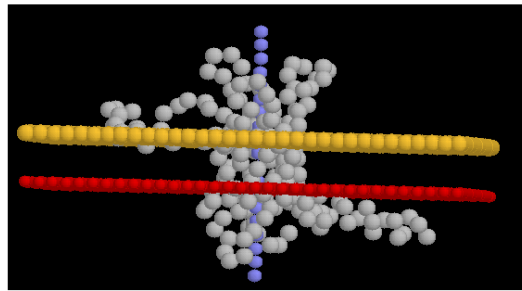
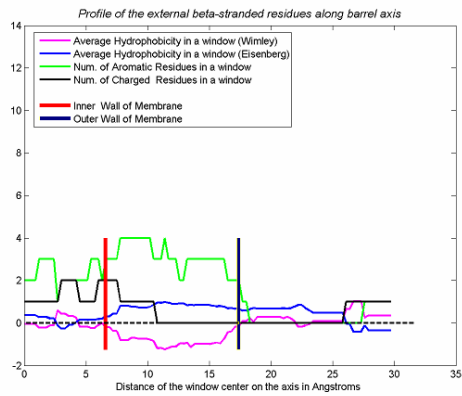
⁷ Το TM-DET, προφανώς από λάθος, δεν περιείχε στην έξοδο του 2 κλώνους όταν εφαρμόστηκε για την 1QD5. Οι δύο κλώνοι είναι αυτοί που δε φαίνονται επιλεγμένοι στην Εικόνα 6.8 (κάτω δεξιά), αλλά ο αναγνώστης μπορεί να καταλάβει το διαμεμβρανικό κομμάτι που προτείνει το TM-DET.



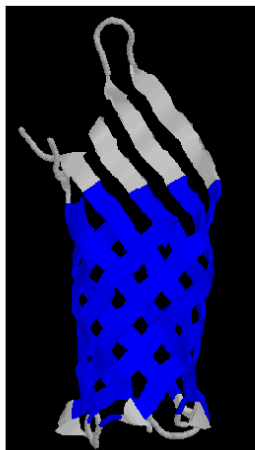
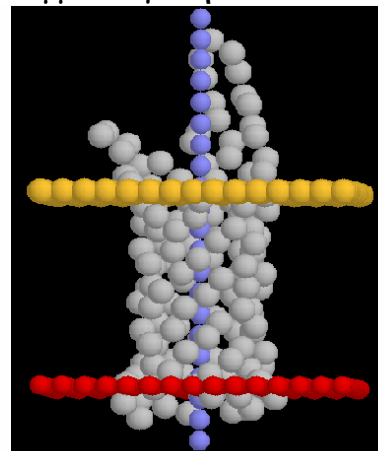
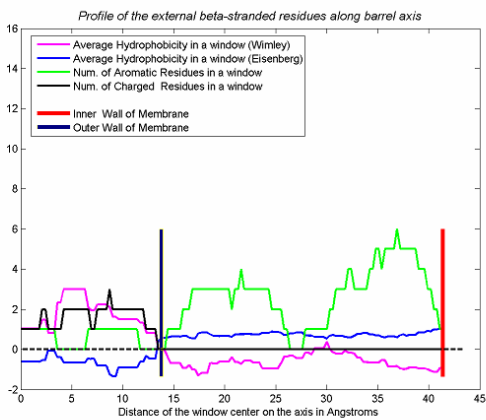
Εικόνα 6. 1 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1P4T.



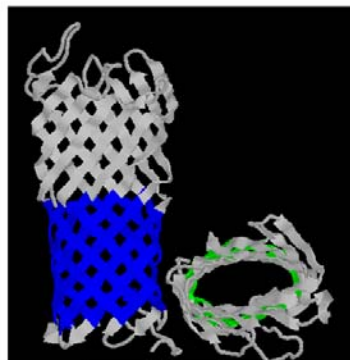
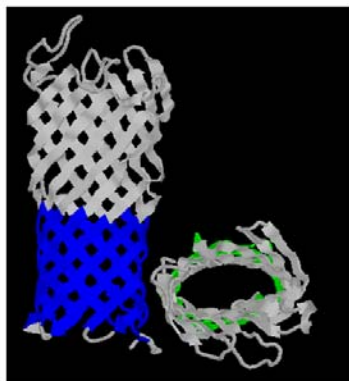
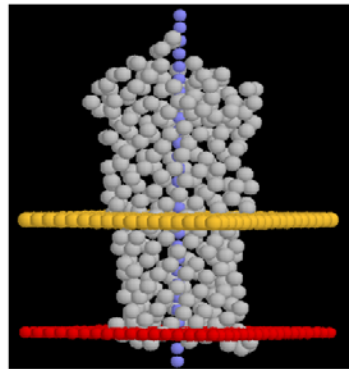
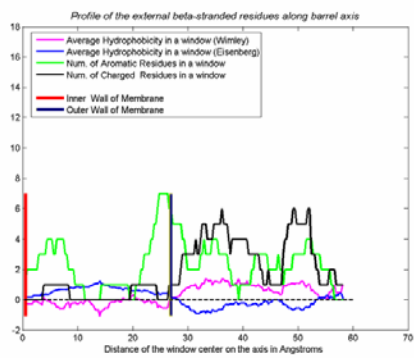
Εικόνα 6. 2 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1QJ8.



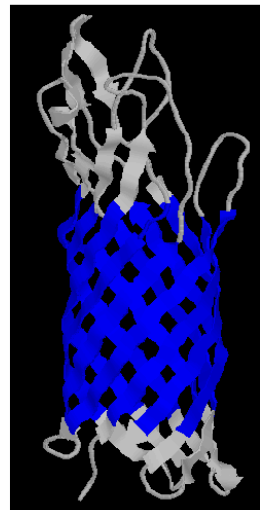
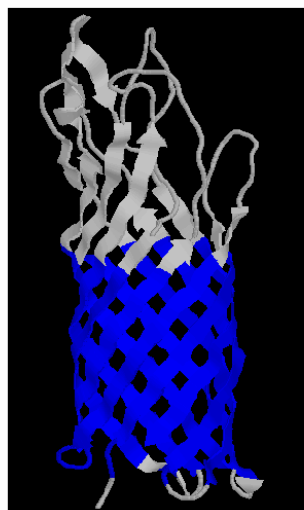
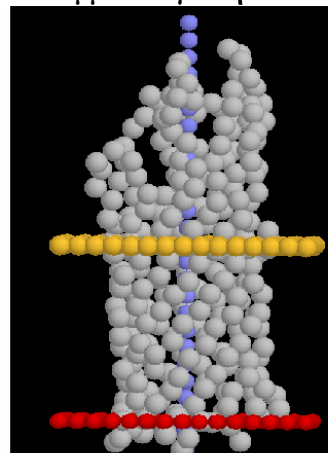
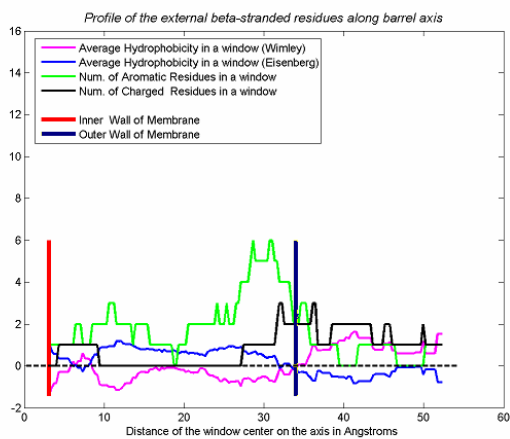
Εικόνα 6. 3 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1MM4.



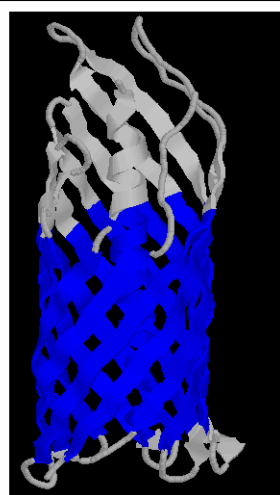
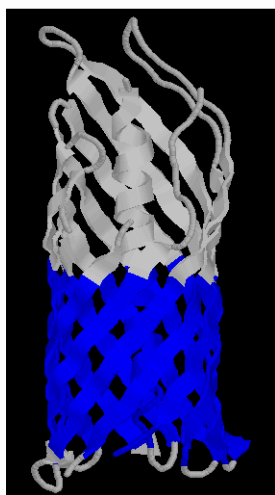
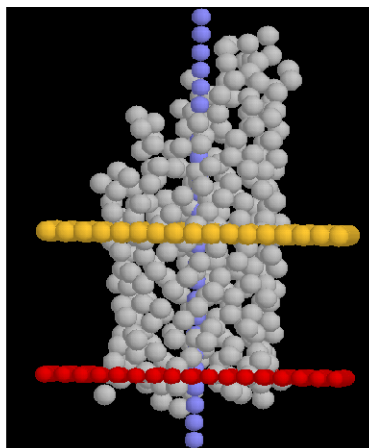
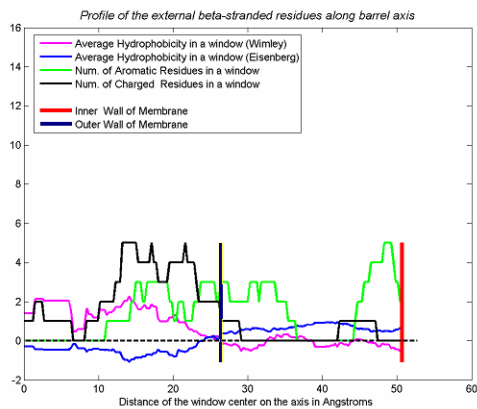
Εικόνα 6. 4 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1QJP.



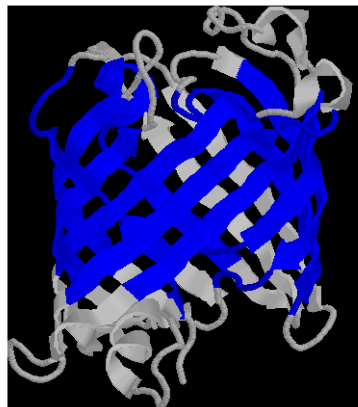
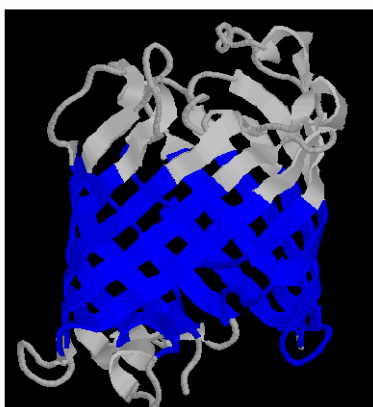
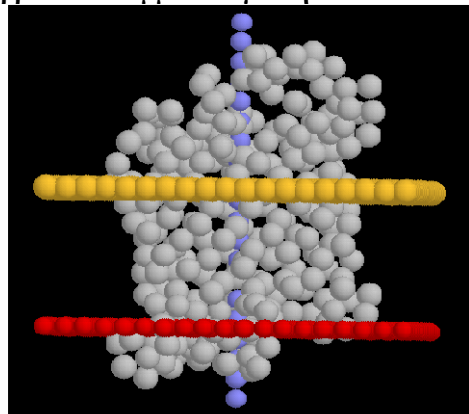
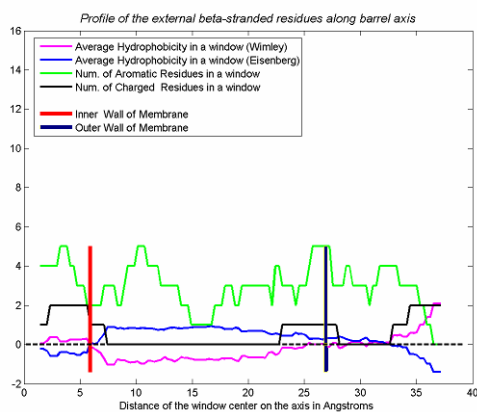
Εικόνα 6. 5 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1I78.



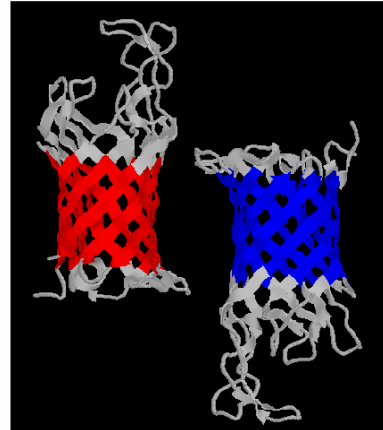
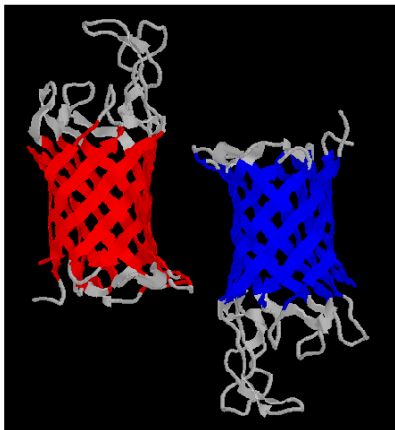
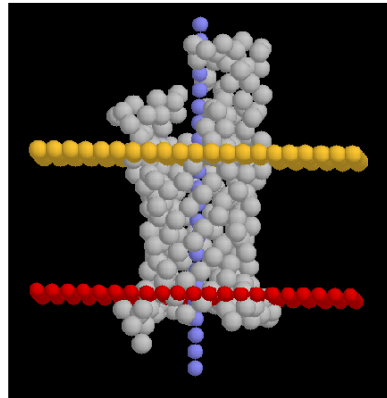
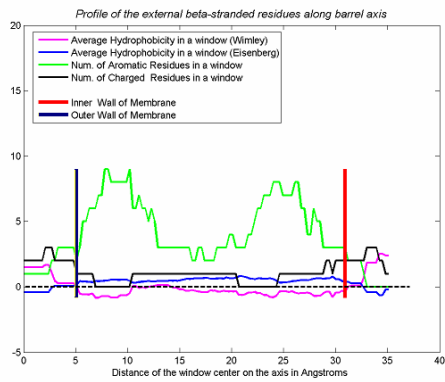
Εικόνα 6. 6 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1K24.



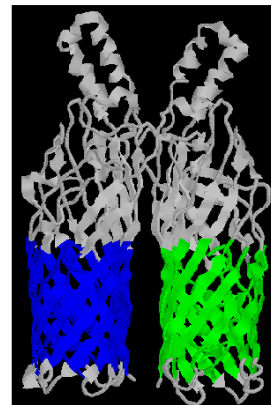
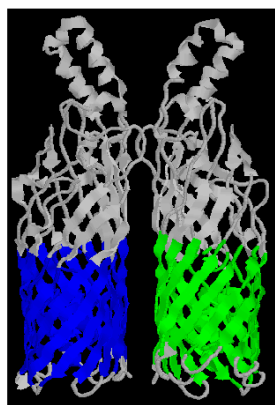
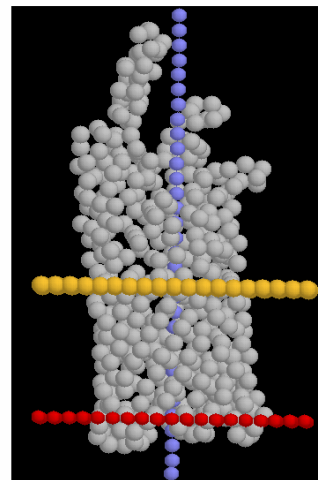
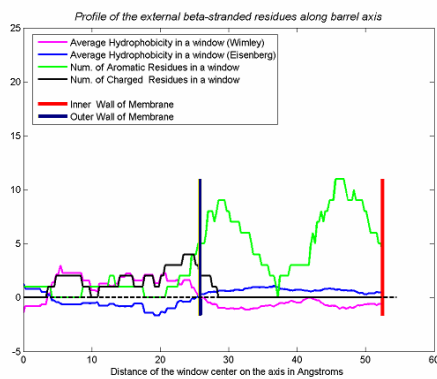
Εικόνα 6. 7 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1UYN.



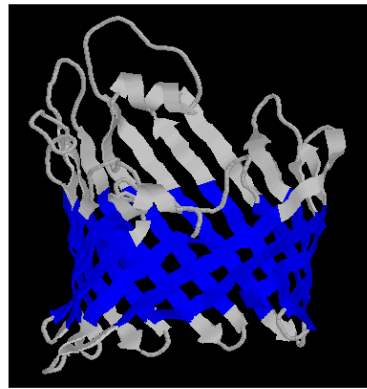
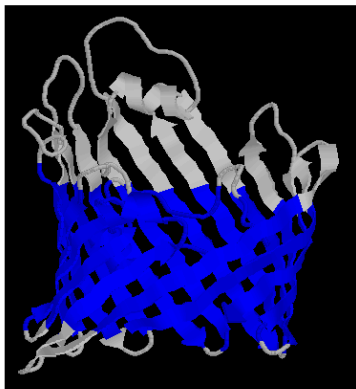
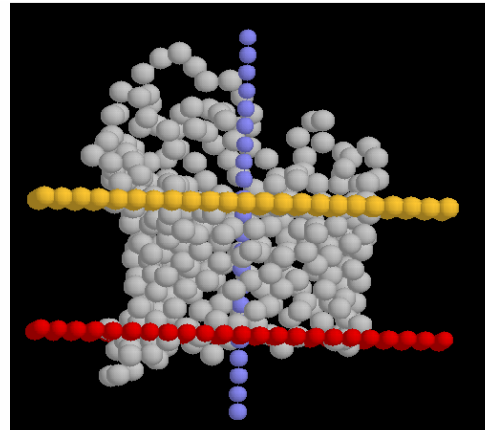
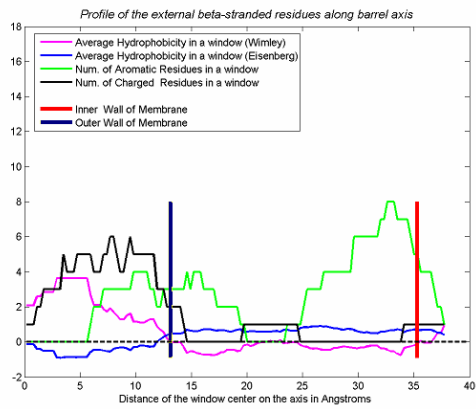
Εικόνα 6. 8 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1QD5.



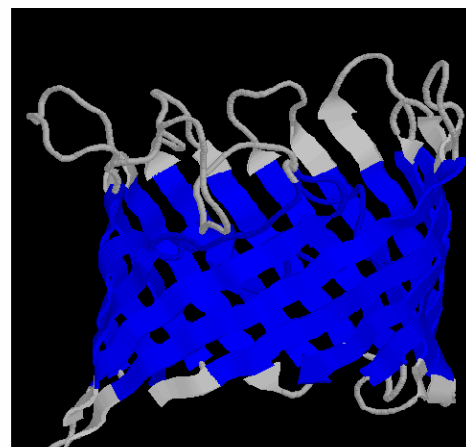
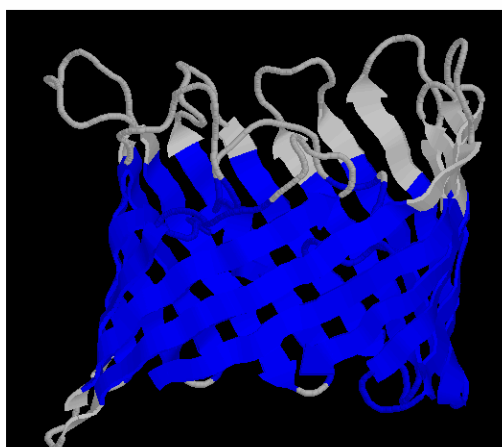
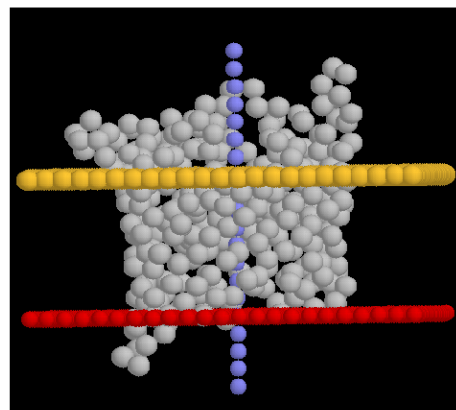
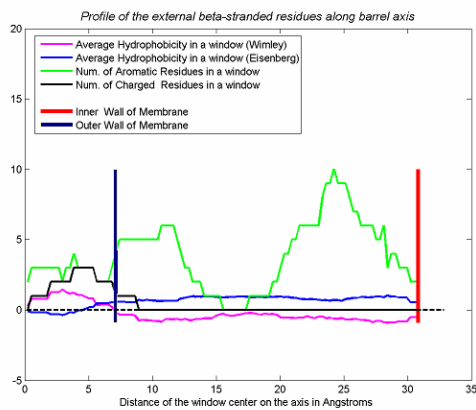
Εικόνα 6. 9 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1TLY.



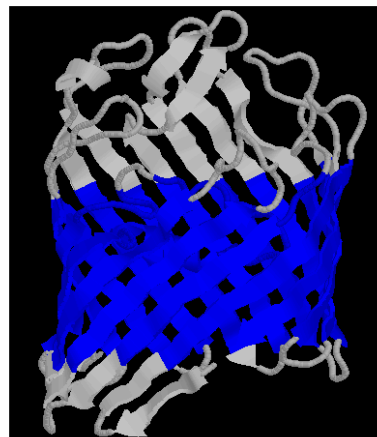
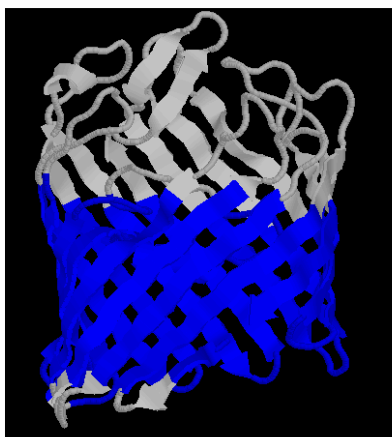
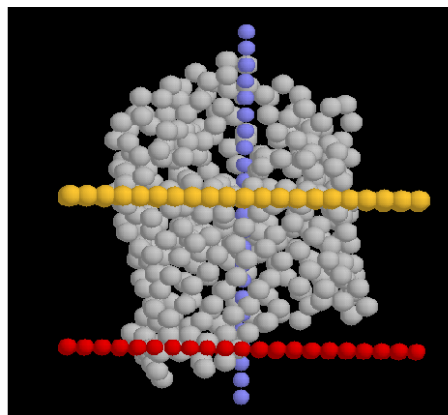
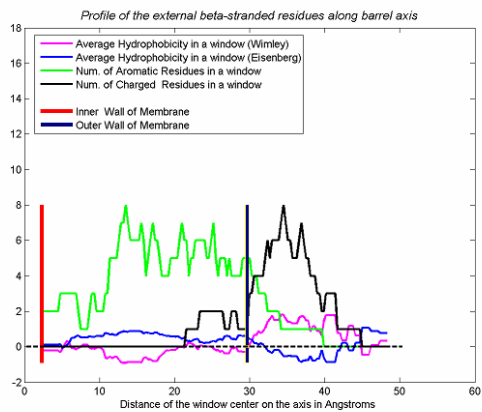
Εικόνα 6. 10 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1TIL.



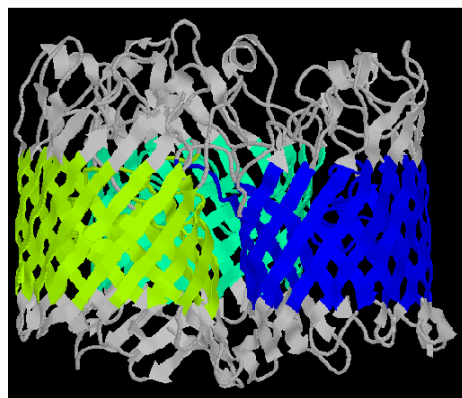
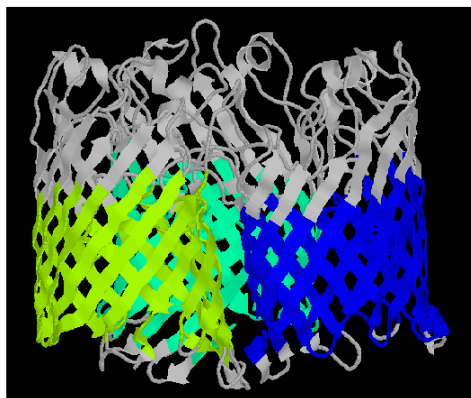
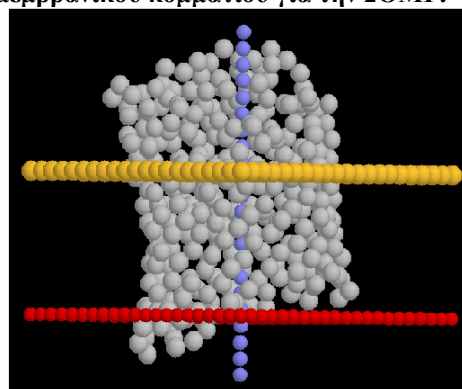
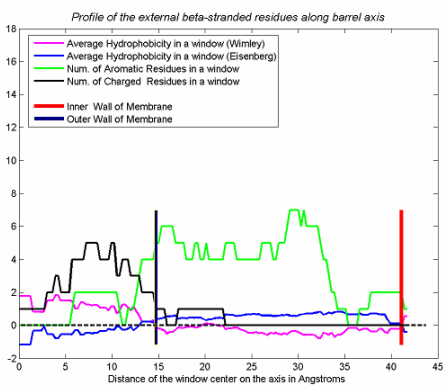
Εικόνα 6. 11 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 2POR.



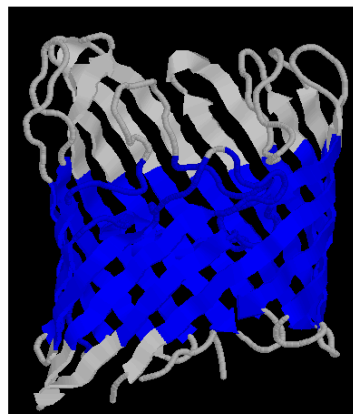
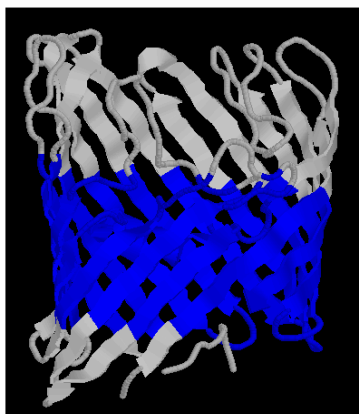
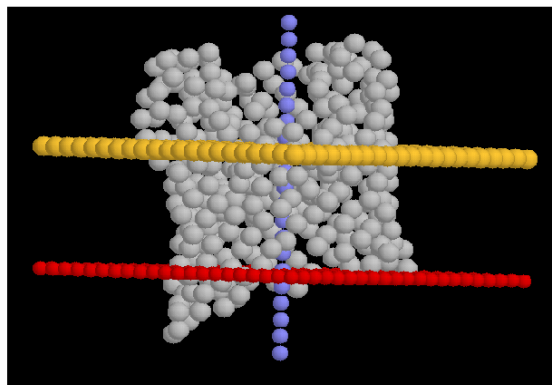
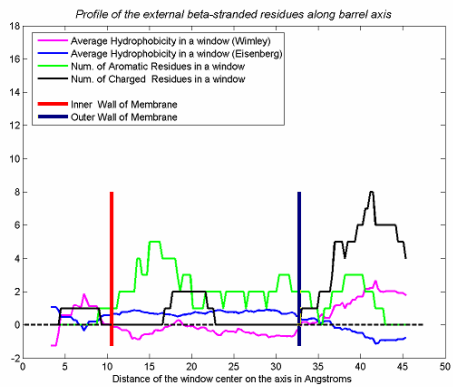
Εικόνα 6. 12 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1PRN.



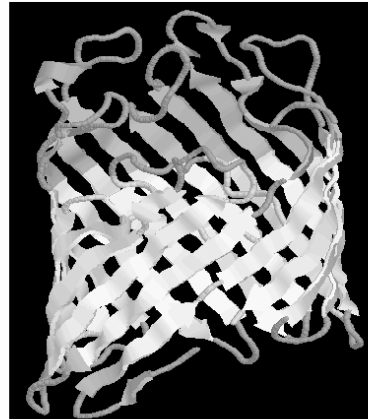
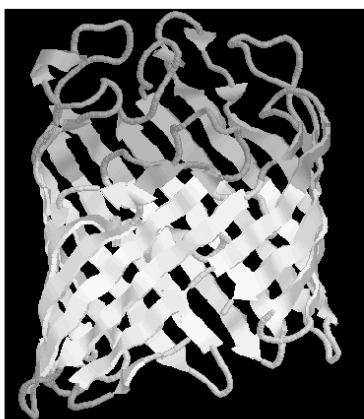
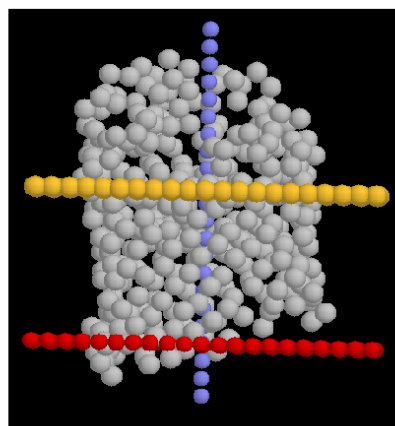
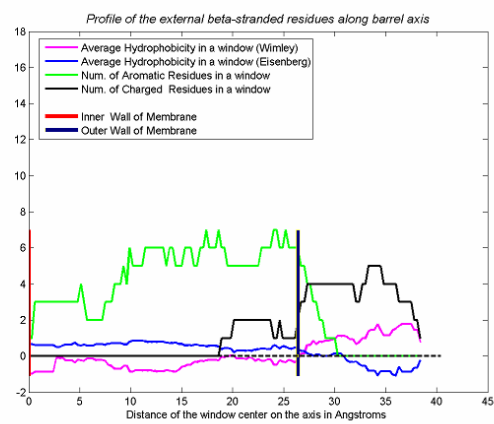
Εικόνα 6. 13 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 2OMF.



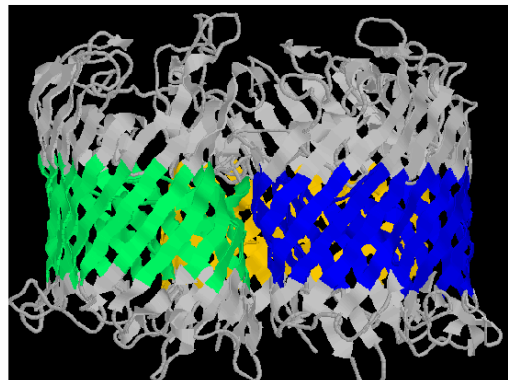
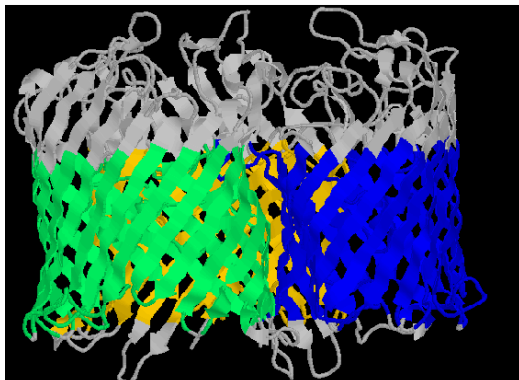
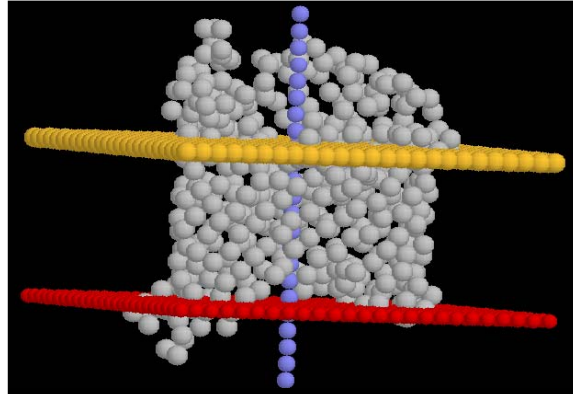
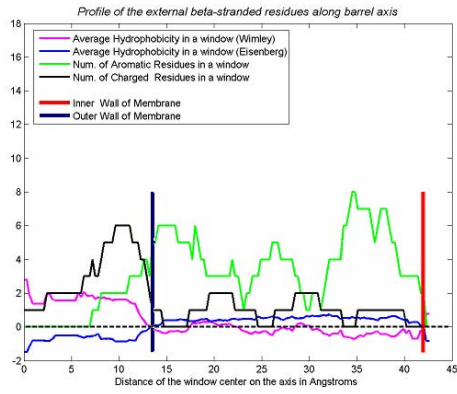
Εικόνα 6. 14 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1OSM.



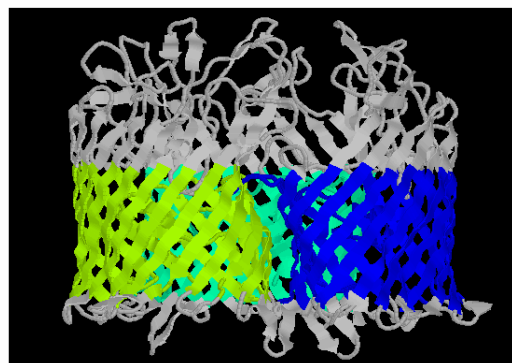
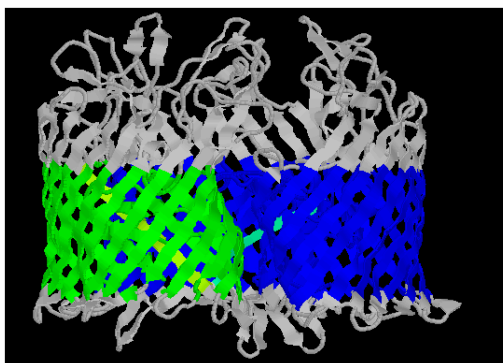
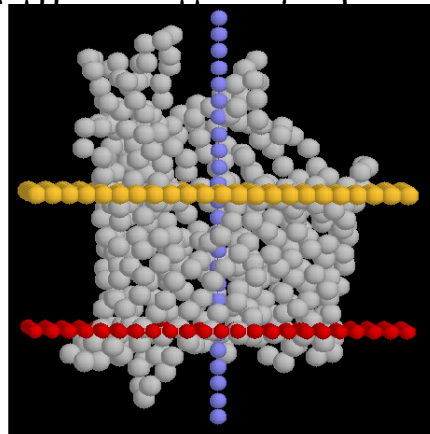
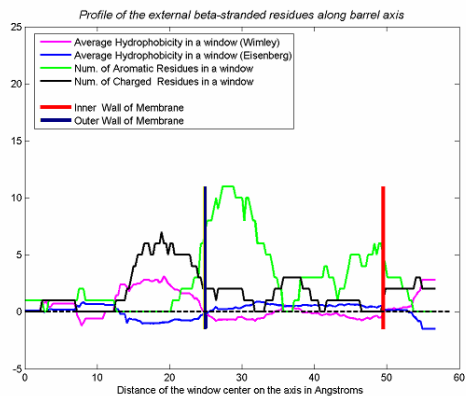
Εικόνα 6. 15 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1E54.



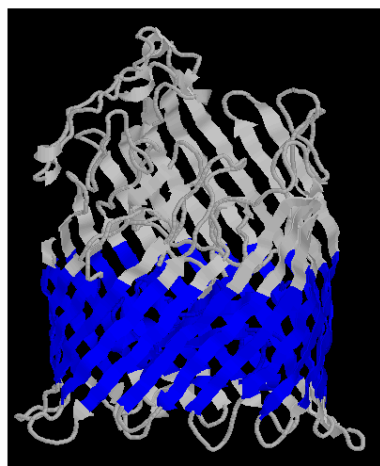
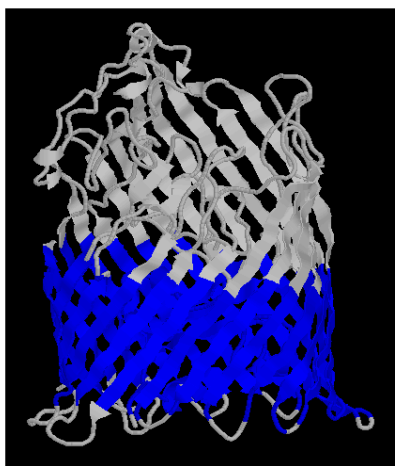
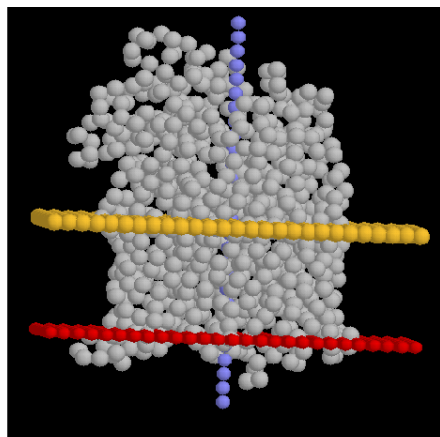
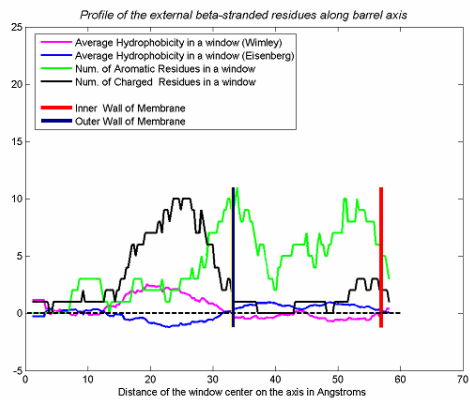
Εικόνα 6. 16 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1PHO.



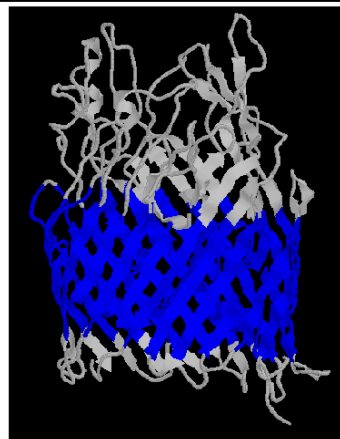
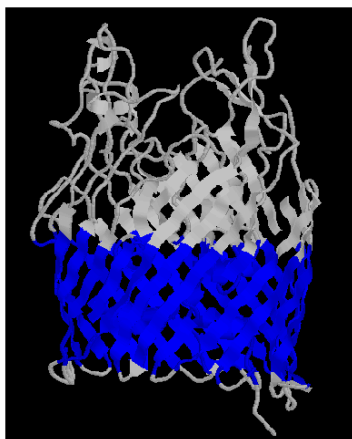
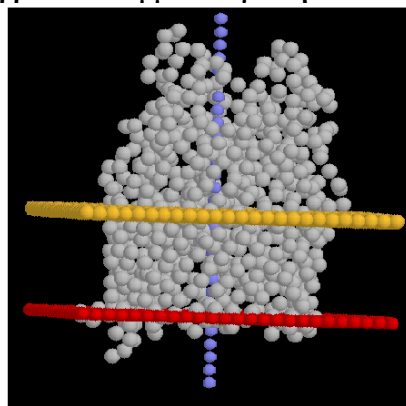
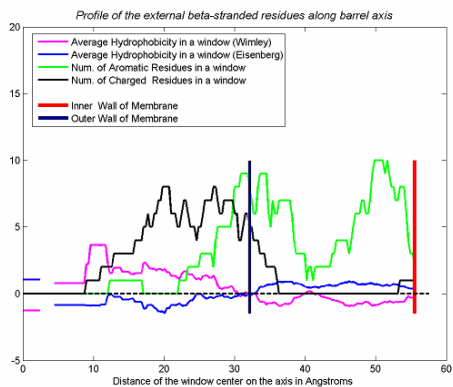
Εικόνα 6. 17 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1A0S.



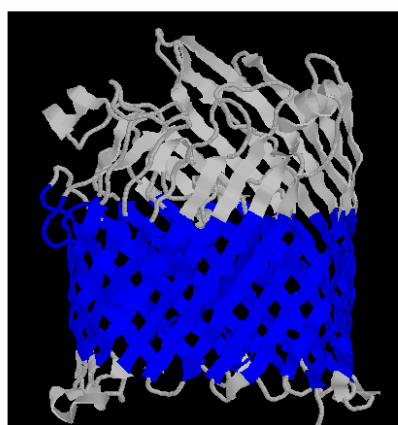
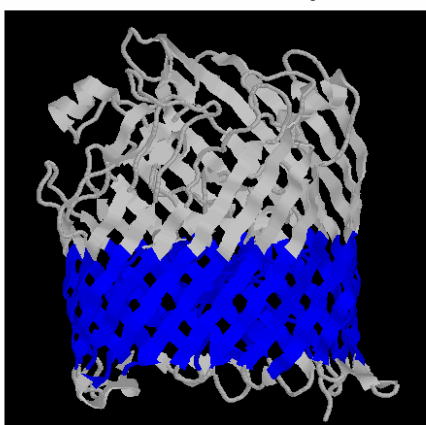
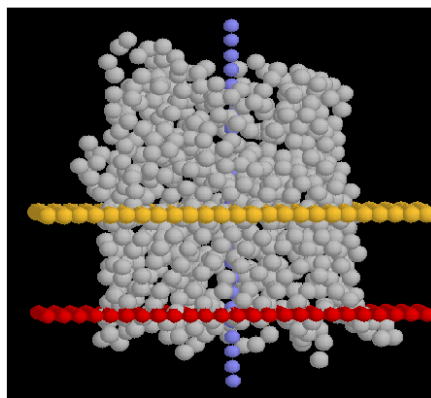
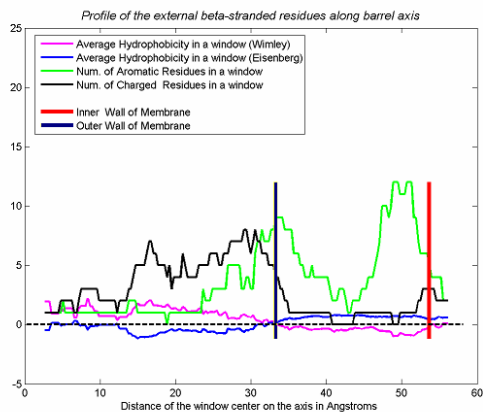
Εικόνα 6. 18 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 2MPR.



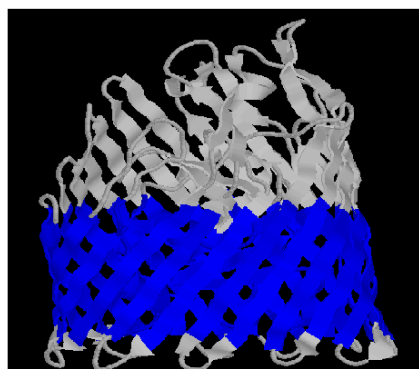
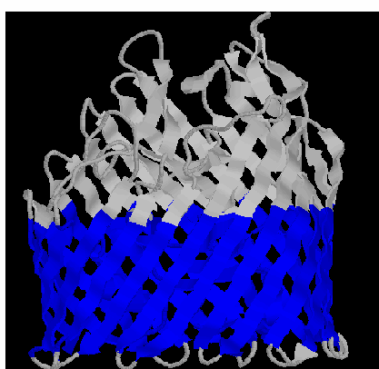
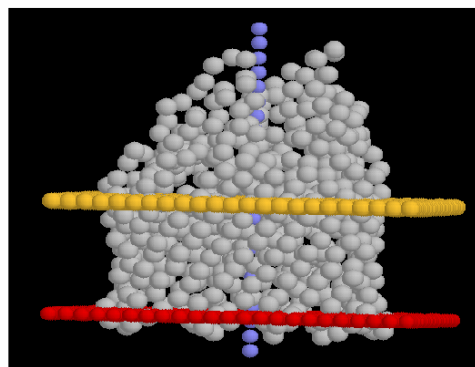
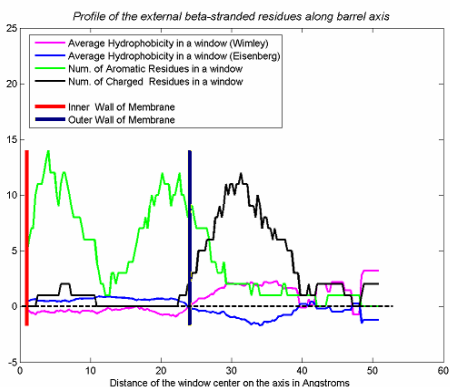
Εικόνα 6. 19 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 2FCP.



Εικόνα 6. 20 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1FEP.



Εικόνα 6. 21 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1KMO.



Εικόνα 6. 22 Αποτελέσματα για την εύρεση του διαμεμβρανικού κομματιού για την 1NQE.

Από τις παραπάνω εικόνες φαίνεται ότι οι άξονες των β-βαρελιών έχουν υπολογιστεί σωστά. Επίσης, παρατηρείται ότι σε όλες τις πρωτεΐνες το προφίλ των εξωτερικών αμινοξέων χαρακτηρίζεται από μια υδρόφοβη ζώνη αμινοξέων, που έχει έντονη παρουσία αρωματικών αμινοξέων γύρω από την αρχή και το τέλος της και περιβάλλεται από φορτισμένα αμινοξέα. Αυτή η παρατήρηση συμπίπτει με παρατήρηση του Wimley [16] και αυτή η ζώνη λαμβάνεται ως το διαμεμβρανικό κομμάτι, όπως έχει ήδη αναφερθεί.

Στον Πίνακα 6.2, παρουσιάζονται τα αποτελέσματα για το πάχος της μεμβράνης των πρωτεϊνών που μελετήθηκαν. Το πάχος της μεμβράνης προέκυψε ως το μήκος της προβολής του διαμεμβρανικού τμήματος πάνω στον άξονα. Στις εικόνες 6.1 έως 6.22 (πάνω αριστερά) αντιστοιχεί στην απόσταση πάνω στον άξονα της κόκκινης και από τη μπλε στήλη, οι οποίες αντιστοιχούν στα τοιχώματα της μεμβράνης.

PDB ID	πάχος μεμβράνης (Å)
1P4T	24.9
1QJ8	24.3
1MM4 ⁸	10.8
1QJP	27.6
1I78	26.4
1K24	30.9
1UYN	24.3
1QD5	21.0
1TLY	25.8
1T1L	26.7
2POR	22.2
1PRN	23.7
2OMF	27.3
1OSM	26.4
1E54	22.2
1PHO	26.4
1A0S	28.5
2MPR	24.6
2FCP	23.7
1FEP	23.4
1KMO	20.4
1NQE	23.1

Πίνακας 6. 2 Τα πάχη των πρωτεϊνών που μελετήθηκαν, σύμφωνα με τα αποτελέσματα εφαρμογής του αλγορίθμου εύρεσης του διαμεμβρανικού κομματιού

⁸ Για την 1MM4.pdb που είναι ένα πολύ κοντό β-βαρέλι χρησιμοποιήθηκαν πιο μικρά όρια για την ελάχιστη και μέγιστη τιμή του μήκους του διαμεμβρανικού κομματιού από 20Å και 32Å που χρησιμοποιήθηκαν για τις υπόλοιπες πρωτεΐνες

Θεωρώντας ως “βάση γνώσης” τα αποτελέσματα του εργαλείου του Tusnadi και συνεργατών (TM-DET), υπολογίστηκαν οι εξής δείκτες για κάθε πρωτεΐνη:

- TP_{TM-DET} : το ποσοστό των αμινοξέων που ο αλγόριθμος προέβλεψε σωστά ως διαμεμβρανικά σύμφωνα με το TM-DET.
- TN_{TM-DET} : το ποσοστό των αμινοξέων που ο αλγόριθμος προέβλεψε σωστά ως μη διαμεμβρανικά σύμφωνα με το TM-DET.
- Ποσοστό Ορθών Προβλέψεων ($\Pi.Ο.Π._{TM-DET}$): το ποσοστό των αμινοξέων που ο αλγόριθμος προέβλεψε σωστά ως διαμεμβρανικά ή μη διαμεμβρανικά σύμφωνα με το TM-DET.

Τα αποτελέσματα για τους προαναφερόμενους δείκτες παρουσιάζονται στον Πίνακα 6.3

PDB ID	TP_{TM-DET} (%)	TN_{TM-DET} (%)	$\Pi.Ο.Π._{TM-DET}$ (%)
1P4T	85.4 (70/82)	50.7 (37/73)	69
1QJ8	92.9 (65/70)	71.8 (56/78)	81.8
1MM4	95 (38/40)	61.5 (80/130)	69.4
1QJP	80.2 (73/91)	100 (46/46)	86.9
1I78	81.1 (86/106)	90.6 (173/191)	86.4
1K24	75.8 (94/124)	79.2 (99/125)	77.5
1UYN	91.6 (120/131)	71.6 (106/148)	81
1QD5 ⁹	-	-	-
1TLY	68.8 (97/141)	100 (110/110)	82.5
1TIL	93.2 (164/176)	93.5 (229/245)	93.3
2POR	72.9 (129/177)	99.2 (123/124)	83.7
1PRN	85.2 (155/182)	81.3 (87/107)	83.7
2OMF	80.3 (147/183)	85.4 (134/157)	82.6
1OSM	73.1 (136/186)	80.1 (125/156)	76.3
1E54	91.8 (157/171)	80.6 (129/160)	86.4
1PHO	85.6 (160/187)	86.7 (124/143)	86.1
1A0S	60.9 (154/253)	100 (160/160)	76
2MPR	96.8 (210/217)	96.6 (197/204)	96.7
2FCP	80 (236/295)	95.6 (392/400)	89.1
1FEP	85.2 (236/277)	81.1 (318/392)	82.8
1KMO	95.9 (233/243)	77.8 (325/418)	76.3
1NQE	90.7 (254/280)	95.5 (257/269)	93.1
<i>Μέση Τιμή</i>	83.9	84.7	82.9

Πίνακας 6. 3 Σύγκριση των αποτελεσμάτων του αλγορίθμου εύρεσης του διαμεμβρανικού κομματιού με τα αποτελέσματα του εργαλείου TM-DET. Με έντονη γραμματοσειρά φαίνονται οι ακραίες τιμές. Στην τελευταία γραμμή φαίνονται οι μέσες τιμές.

⁹ Το TM-DET, προφανώς από λάθος, δεν περιείχε στην έξοδο του 2 κλώνους όταν εφαρμόστηκε για την 1QD5. Για το λόγο αυτό δεν υπολογίστηκαν οι αντίστοιχοι δείκτες.

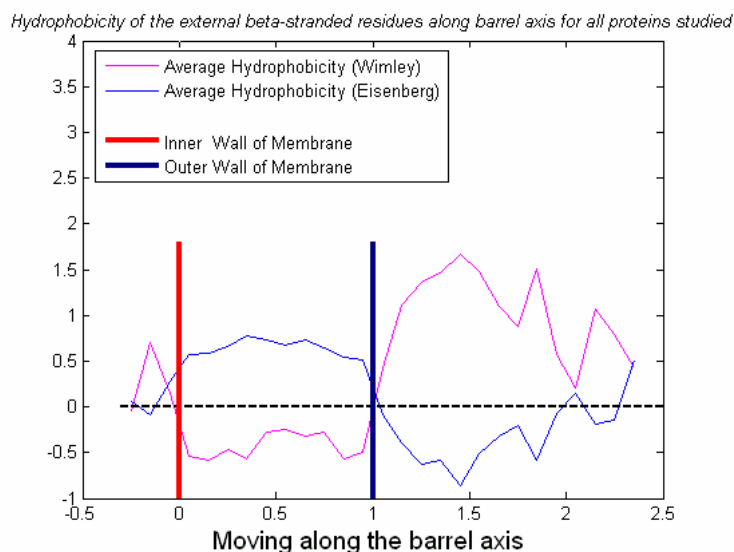
Από τις εικόνες 6.1-6.22 αλλά και από τους δείκτες του Πίνακα 6.3 γίνεται προφανές ότι ο αλγόριθμος που αναπτύχθηκε για την εύρεση του διαμεμβρανικού κομματιού είχε αρκετά παρόμοια αποτελέσματα με το εργαλείο TM-DET. Οι μέσες τιμές που προέκυψαν για τους δείκτες TP_{TM-DET} , TN_{TM-DET} , $Π.Ο.Π._{TM-DET}$ είναι 83.9%, 84.7% και 82.9%. Σημειώνεται, εδώ, ότι το εργαλείο TM-DET αποτελεί ένα ακόμη εργαλείο και δε μπορεί να θεωρηθεί ως πραγματική “βάση γνώσης” για το πρόβλημα εύρεσης των διαμεμβρανικών αμινοξέων μιας διαμεμβρανικής πρωτεΐνης. Στο παράρτημα, ο αναγνώστης μπορεί να δει αναλυτικά ποια αμινοξέα βρέθηκαν ως διαμεμβρανικά στις 22 πρωτεΐνες που μελετήθηκαν.

6.3 Μέση υδροφοβικότητα και κατανομή των αμινοξέων κατά μήκος του άξονα του β-βαρελιού

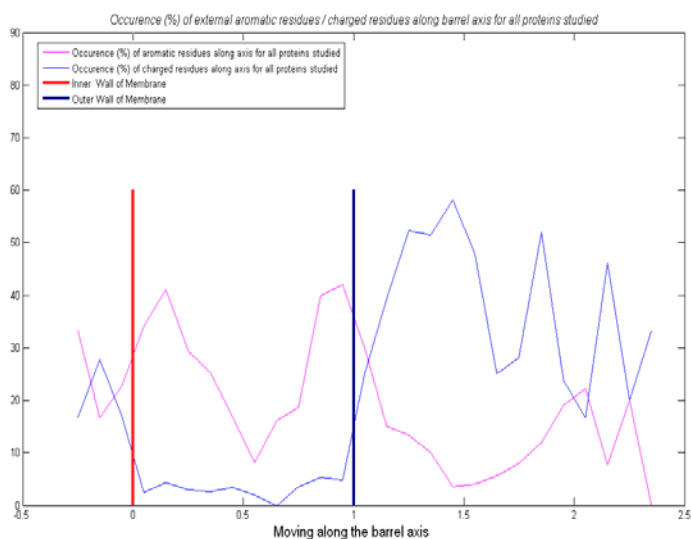
Έχοντας βρει τον άξονα των β-βαρελιών των 22 πρωτεϊνών που μελετήθηκαν στη παρούσα εργασία ήταν δυνατή η εξαγωγή συμπερασμάτων για τη μεταβολή της υδροφοβικότητας κατά μήκος του άξονα του β-βαρελιού στο σύνολο των 22 πρωτεϊνών, αλλά και για το ποια αμινοξέα εμφανίζονται περισσότερο (προτιμούνται) κατά μήκος του άξονα του βαρελιού στο σύνολο των 22 πρωτεϊνών. Η έννοια της απόστασης πάνω στον άξονα κανονικοποιήθηκε σε σχέση με τη θέση του διαμεμβρανικού κομματιού: το 0 αντιστοιχεί στο εσωτερικό τοίχωμα της μεμβράνης, ενώ το 1 στο εξωτερικό τοίχωμα της μεμβράνης. Παράλληλα, χρησιμοποιήθηκαν διαδοχικά παράθυρα μήκους 0.1 στην κανονικοποιημένη κλίμακα. Η μελέτη της μέσης υδροφοβικότητας και της κατανομής αμινοξέων κατά μήκος του άξονα των β-βαρελιών έγινε ξεχωριστά για τα εξωτερικά και εσωτερικά αμινοξέα των β-κλώνων των βαρελιών.

A) **Εξωτερικά αμινοξέα:** Στην Εικόνα 6.23 φαίνεται η μεταβολή της υδροφοβικότητας κατά μήκος του άξονα. Όπως και στην εργασία του Wimley [16], είναι φανερό ότι η μέση υδροφοβικότητα, κάνοντας χρήση της πειραματικής κλίμακας των Wimley *et al* (“octanol whole residue hydrophobicity scale”) [56] παίρνει αρνητικές τιμές για τα διαμεμβρανικά αμινοξέα. Στην Εικόνα 6.24 φαίνονται: i) η μεταβολή του λόγου των αρωματικών αμινοξέων (**Trp**, **Tyr**, **Phe**) προς το συνολικό αριθμό αμινοξέων που εμφανίζονται σε μία θέση κατά μήκος του άξονα (ροζ καμπύλη) ii) η μεταβολή του λόγου των φορτισμένων αμινοξέων (**Lys**, **His**, **Glu**, **Asp**, **Arg**)

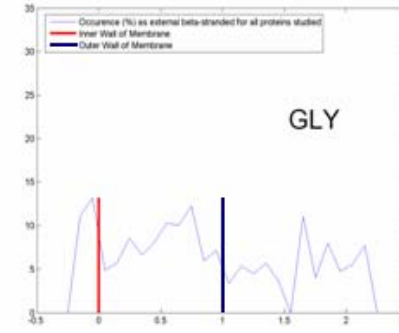
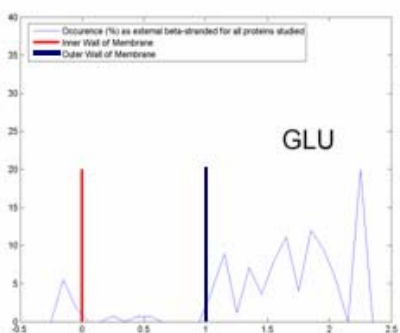
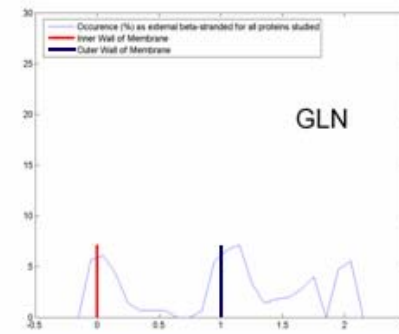
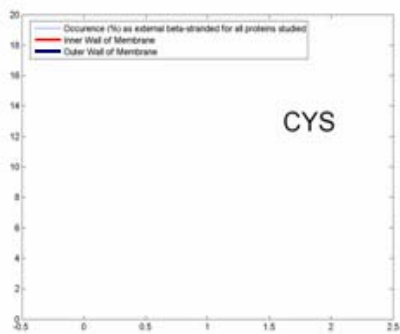
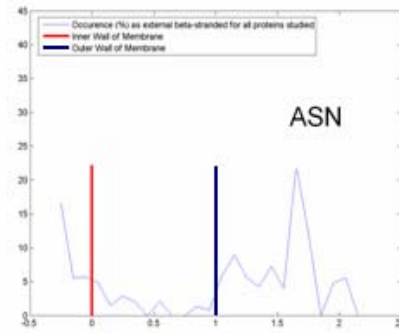
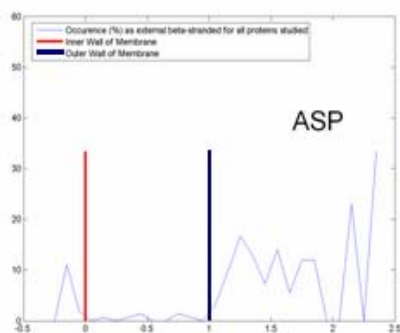
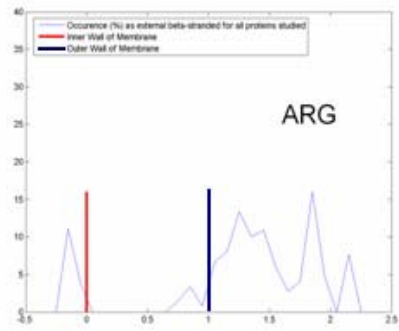
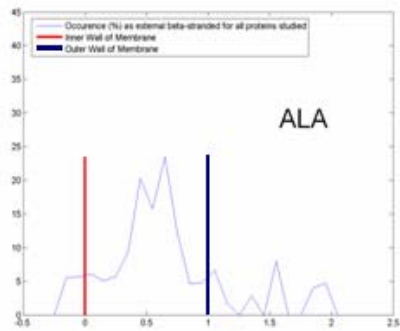
προς το συνολικό αριθμό αμινοξέων που εμφανίζονται σε μία θέση κατά μήκος του άξονα (μπλε καμπύλη). Είναι φανερή η υψηλή παρουσία αρωματικών αμινοξέων κοντά στα τοιχώματα της μεμβράνης καθώς και η αυξημένη παρουσία φορτισμένων αμινοξέων έξω από τα όρια της μεμβράνης, όπως ήδη παρατήρησε ο Wimley [16]. Στην Εικόνα 6. 25 φαίνεται η επί τοις εκατό (%) προτίμηση για καθένα από τα 20 αμινοξέα ως εξωτερικά των β-κλώνων κατά μήκος του άξονα για τις 22 πρωτεΐνες.



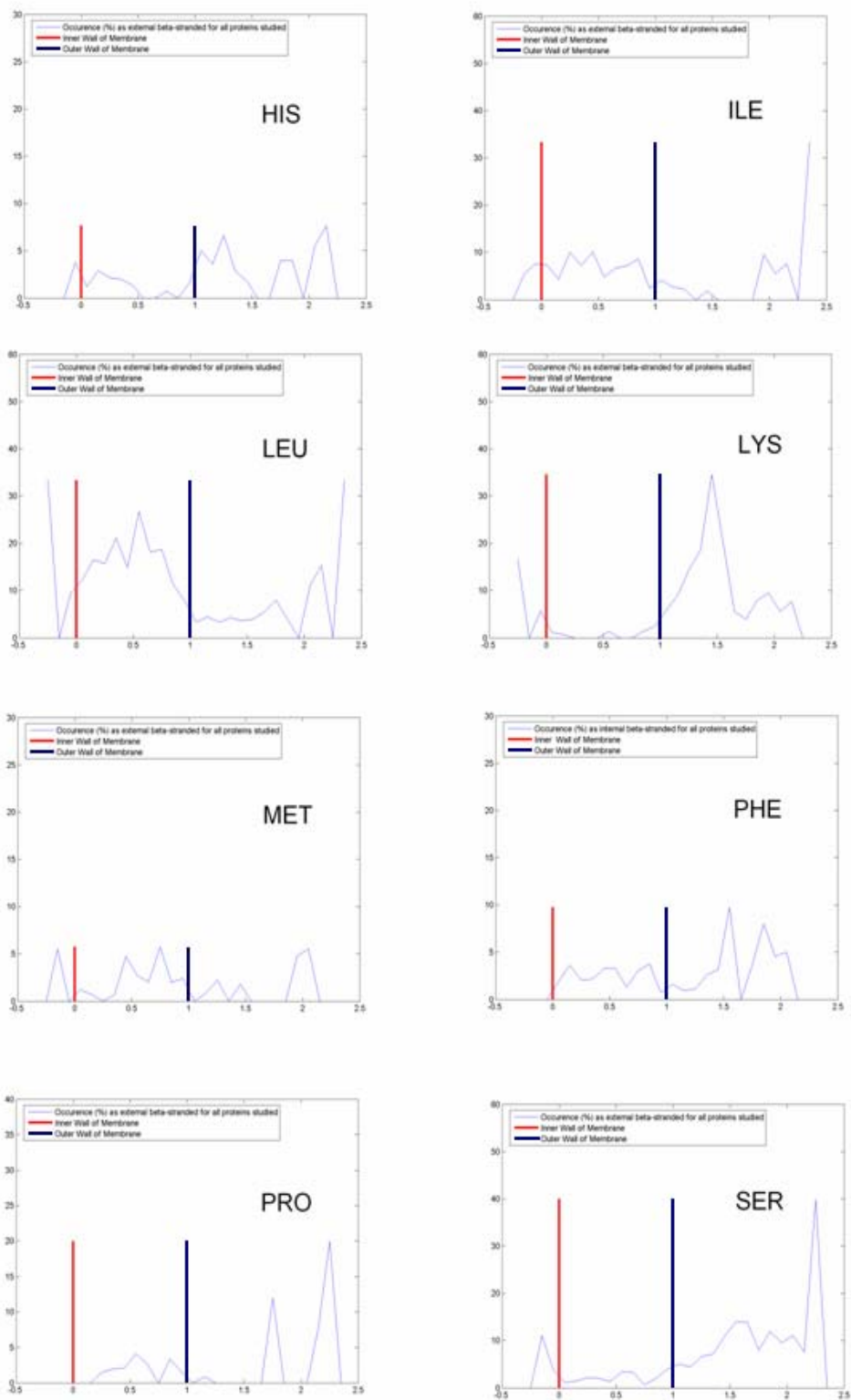
Εικόνα 6. 23 Μεταβολή της μέσης υδροφοβικότητας των εξωτερικών αμινοξέων κατά μήκος του άξονα για το σύνολο των 22 πρωτεϊνών που μελετήθηκαν



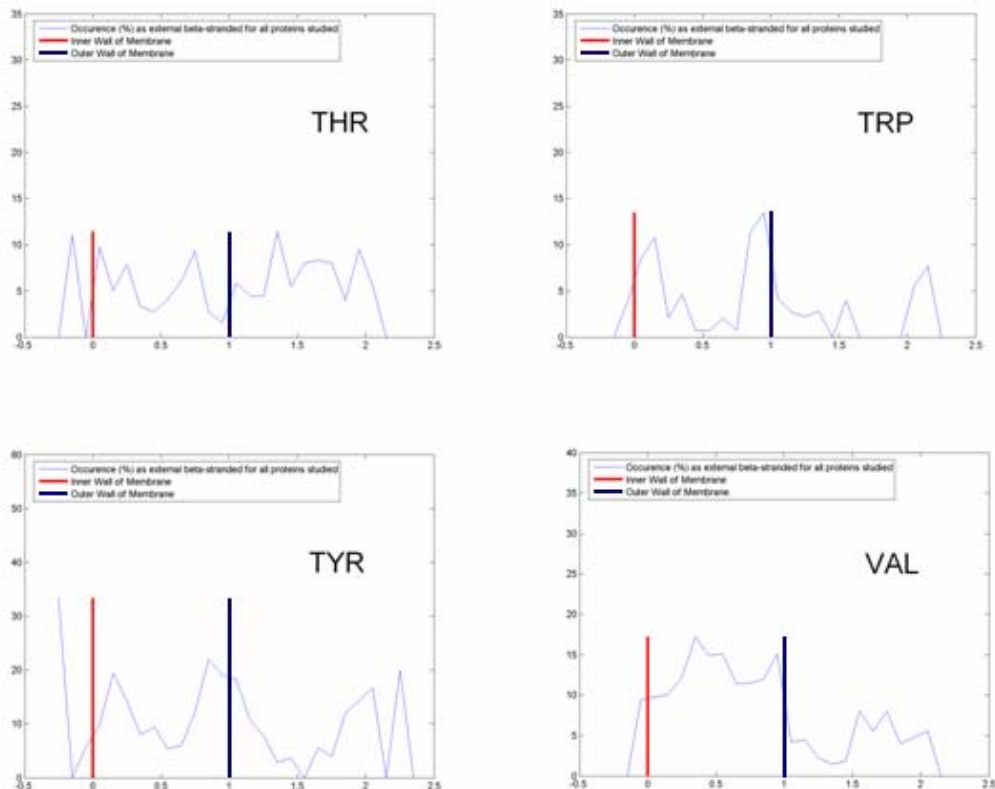
Εικόνα 6. 24 Ο λόγος του αριθμού των αρωματικών αμινοξέων προς το συνολικό αριθμό αμινοξέων που εμφανίζονται σε μία θέση κατά μήκος του άξονα (ροζ καμπύλη) και ο λόγος του αριθμού των φορτισμένων αμινοξέων προς το συνολικό αριθμό αμινοξέων που εμφανίζονται σε μία θέση κατά μήκος του άξονα (μπλε καμπύλη) για το σύνολο των εξωτερικών αμινοξέων των β-βαρελίων των 22 πρωτεϊνών που μελετήθηκαν.



Εικόνα 6. 25 Η επί τοις εκατό (%) προτίμηση για καθένα από τα 20 αμινοξέα ως εξωτερικά των β-κλώνων κατά μήκος του άξονα για τις 22 πρωτεΐνες που μελετήθηκαν. Τα 0 και 1 αντιστοιχούν στο εσωτερικό και εξωτερικό τοίχωμα της μεμβράνης.



Εικόνα 6. 25 Η επί τοις εκατό (%) προτίμηση για καθένα από τα 20 αμινοξέα ως εξωτερικά των β-κλώνων κατά μήκος του άξονα για τις 22 πρωτεΐνες που μελετήθηκαν. Τα 0 και 1 αντιστοιχούν στο εσωτερικό και εξωτερικό τοίχωμα της μεμβράνης (συνέχεια)

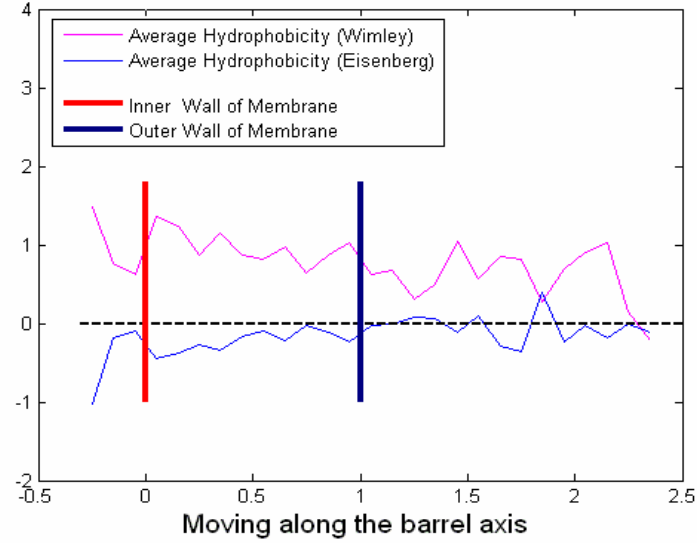


Εικόνα 6. 25 Η επί τοις εκατό (%) προτίμηση για καθένα από τα 20 αμινοξέα εξωτερικά των β-κλώνων κατά μήκος του άξονα για τις 22 πρωτεΐνες που μελετήθηκαν. Τα 0 και 1 αντιστοιχούν στο εσωτερικό και εξωτερικό τοίχωμα της μεμβράνης (συνέχεια).

Από την Εικόνα 6.25. παρατηρούμε αυξημένη παρουσία υδρόφοβων αμινοξέων στο εσωτερικό της μεμβράνης (κυρίως **Ala, Val, Leu** και λιγότερο **Pro, Met, Ile**), αυξημένη παρουσία αρωματικών (**Trp, Tyr, Phe**) αλλά και **Gln** κοντά στα τοιχώματα της μεμβράνης καθώς και αυξημένη παρουσία φορτισμένων (**Lys, His, Glu, Asp, Arg**) αλλά και **Ser** στην περιοχή εκτός της μεμβράνης. Τέλος, παρατηρείται η μηδενική παρουσία της κυστεΐνης (**Cys**). Οι παρατηρήσεις αυτές συμφωνούν με τις παρατηρήσεις του Wimley [16].

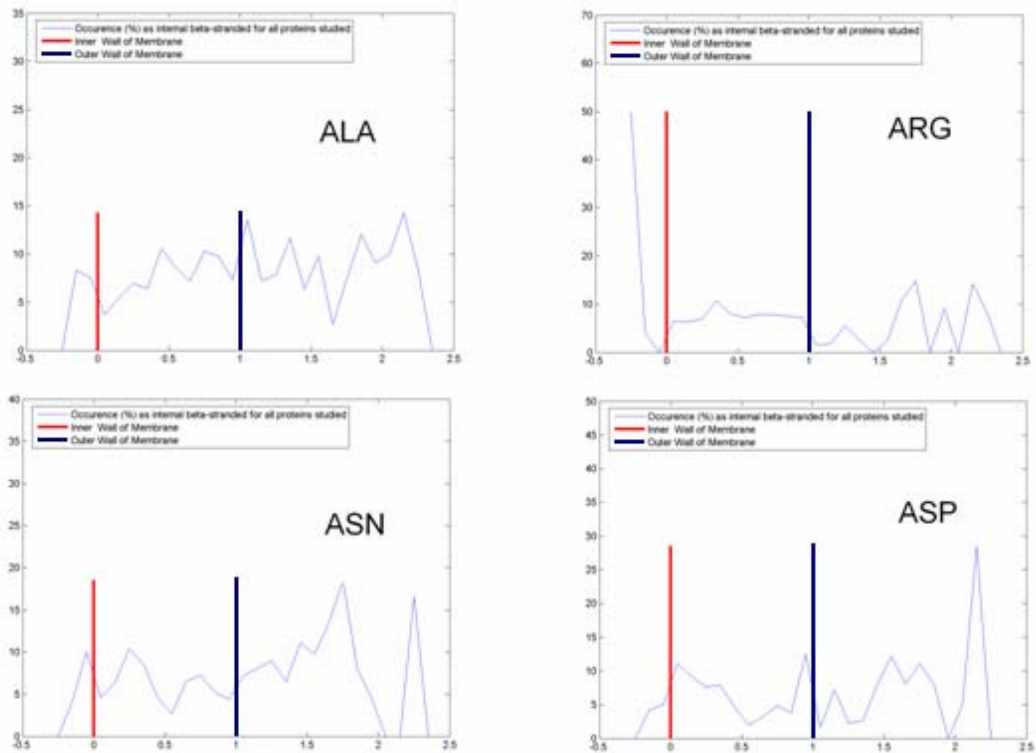
B) Εσωτερικά αμινοξέα: Στην Εικόνα 6.26 φαίνεται η μεταβολή της υδροφοβικότητας κατά μήκος του άξονα. Παρατηρείται, ότι στο εσωτερικό της μεμβράνης τα εσωτερικά αμινοξέα είναι υδρόφιλα (είναι φανερά λιγότερο υδρόφοβα από τα εξωτερικά αμινοξέα στο εσωτερικό αλλά και όχι μέγιστης υδροφιλικότητας). Κινούμενοι προς το εξωτερικό της μεμβράνης η υδροφιλικότητα των εσωτερικών αμινοξέων αυξάνεται

Hydrophobicity of the internal beta-stranded residues along barrel axis for all proteins studied

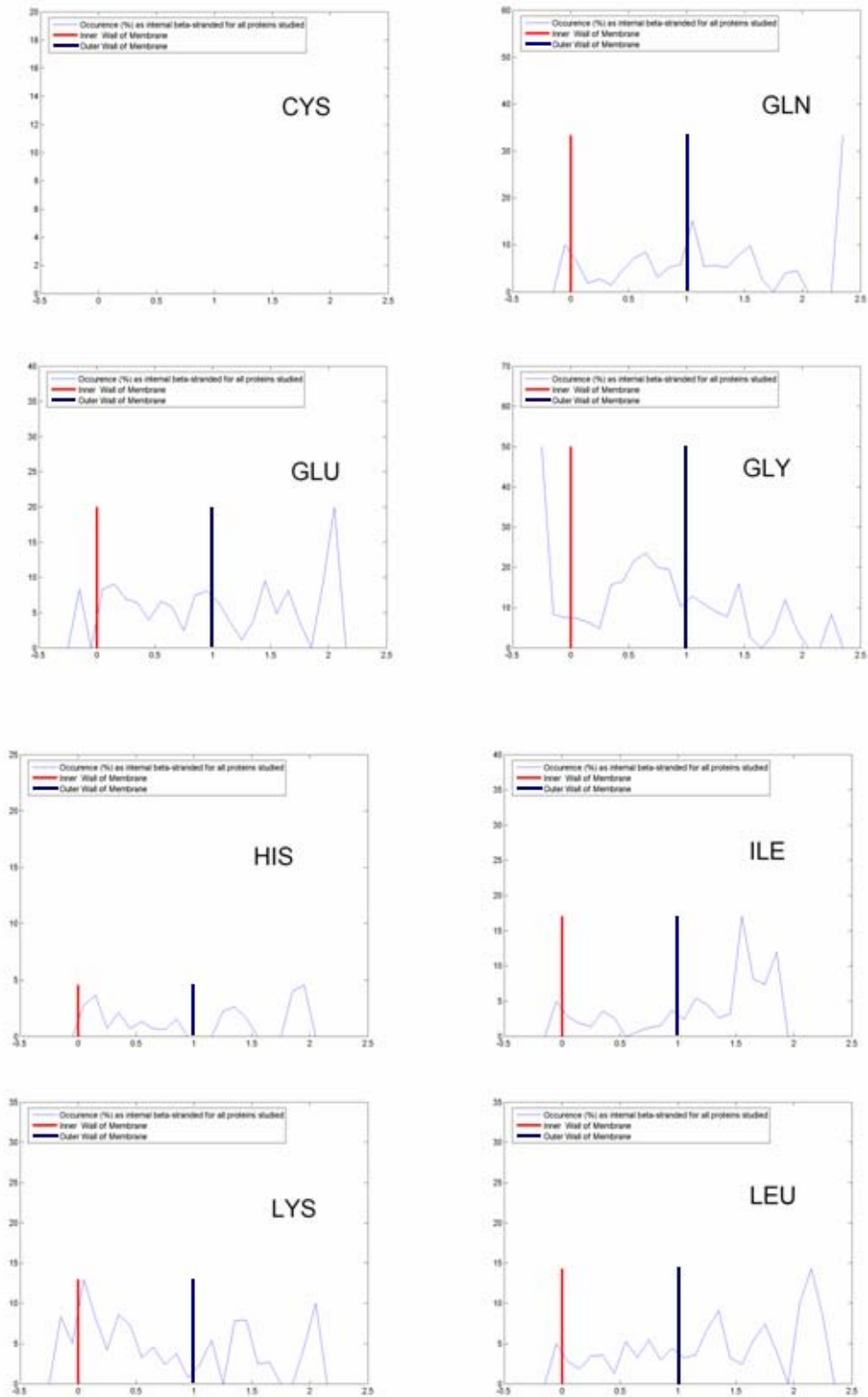


Εικόνα 6. 26 Μεταβολή της μέσης υδροφοβικότητας των εσωτερικών αμινοξέων κατά μήκος του άξονα για το σύνολο των 22 πρωτεϊνών που μελετήθηκαν

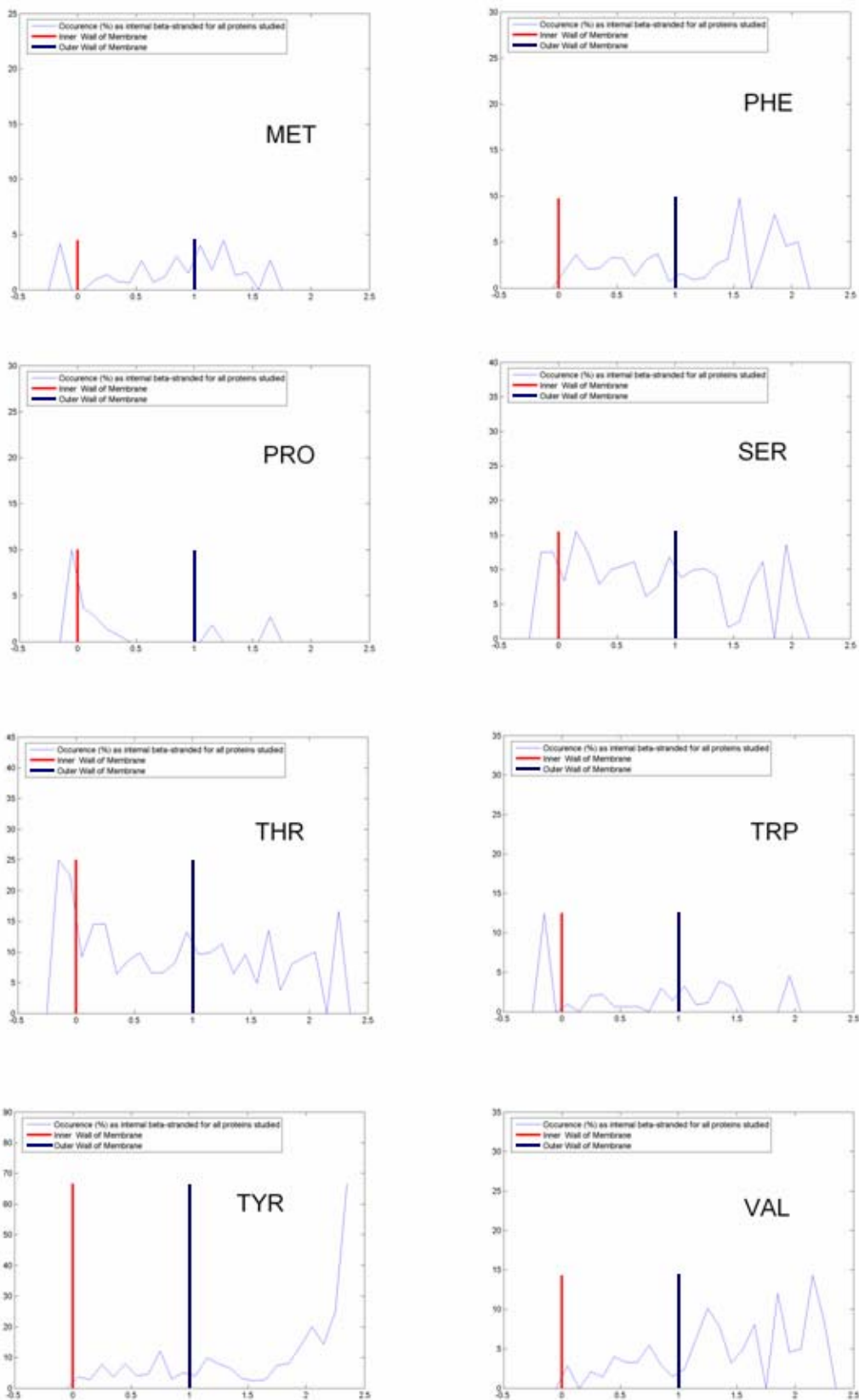
Στην Εικόνα 6. 27 φαίνεται η επί τοις εκατό (%) προτίμηση για καθένα από τα 20 αμινοξέα ως εσωτερικά των β-κλώνων κατά μήκος του άξονα για τις 22 πρωτεΐνες.



Εικόνα 6. 27 Η επί τοις εκατό (%) προτίμηση για καθένα από τα 20 αμινοξέα ως εσωτερικά των β-κλώνων κατά μήκος του άξονα για τις 22 πρωτεΐνες που μελετήθηκαν. Τα 0 και 1 αντιστοιχούν στο εσωτερικό και εξωτερικό τοίχωμα της μεμβράνης



Εικόνα 6. 27 Η επί τοις εκατό (%) προτίμηση για καθένα από τα 20 αμινοξέα ως εσωτερικά των β-κλώνων κατά μήκος του άξονα για τις 22 πρωτεΐνες που μελετήθηκαν. Τα 0 και 1 αντιστοιχούν στο εσωτερικό και εξωτερικό τοίχωμα της μεμβράνης (συνέχεια).



Εικόνα 6. 27 Η επί τοις εκατό (%) προτίμηση για καθένα από τα 20 αμινοξέα ως εσωτερικά των β-κλώνων κατά μήκος του άξονα για τις 22 πρωτεΐνες που μελετήθηκαν. Τα 0 και 1 αντιστοιχούν στο εσωτερικό και εξωτερικό τοίχωμα της μεμβράνης (συνέχεια).

Δεν παρατηρείται μια ξεκάθαρη αυξημένη παρουσία μερικών αμινοξέων ως εσωτερικά μέσα στη μεμβράνη σε σχέση με το χώρο έξω από τη μεμβράνη, παρά μόνο για τα αμινοξέα Gly και Ser. Πάντως στο διαμεμβρανικό χώρο της μεμβράνης προτιμούνται σε χαμηλό ποσοστό, ως εσωτερικά, τα υδρόφοβα αμινοξέα (π.χ. **Val, Trp, Met, Leu, Phe**). Τέλος, παρατηρείται η μηδενική παρουσία της κυστεΐνης (**Cys**).

7. Διαδικτυακή Εφαρμογή

Στα πλαίσια της διπλωματικής εργασίας δημιουργήθηκε κατάλληλο διαδικτυακό εργαλείο που δίνει τη δυνατότητα σε χρήστες του παγκόσμιου ιστού να εφαρμόζουν τον αλγόριθμο γεωμετρικής μοντελοποίησης και εύρεσης διαμεμβρανικού τμήματος σε διαμεμβρανικές πρωτεΐνες τύπου β-βαρελιού που εισάγουν στο σύστημα και να λαμβάνουν τα αποτελέσματα μέσω ηλεκτρονικού μηνύματος (email). Το εργαλείο ονομάστηκε **TbB-Tool** (Transmembrane beta Barrel Tool) και είναι διαθέσιμο στην ηλεκτρονική διεύθυνση www.di.uoa.gr/~ivalavan/TbB_Tool.htm. Μέρος της αρχικής σελίδας του εργαλείου φαίνεται στην Εικόνα 7.1

Welcome to TbB-Tool

The current tool processes a .pdb file containing a transmembrane β -barrel. It models geometrically the β -barrel protein and finds the transmembrane segments.

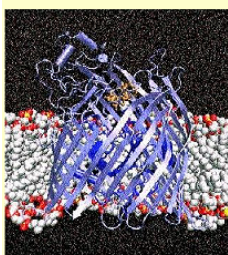


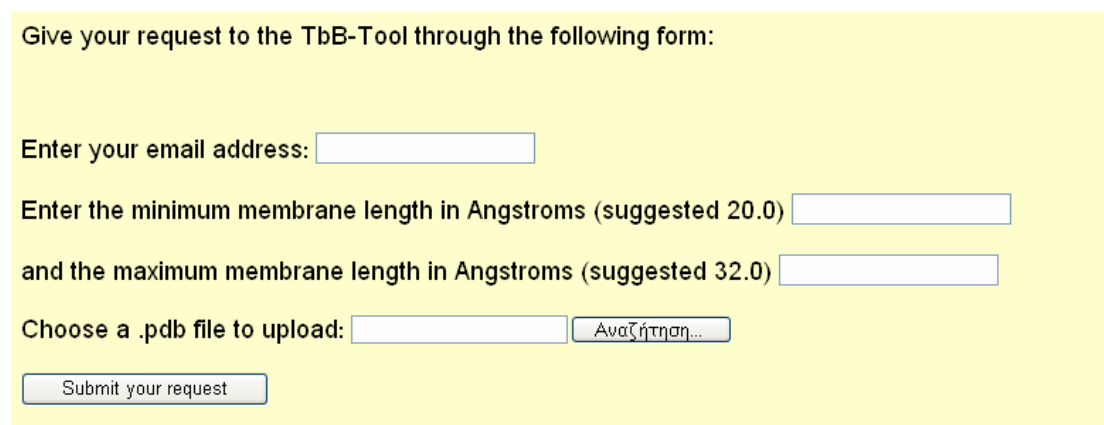
Fig. FhuA embedded in a lipid bilayer

Εικόνα 7. 1 Μέρος της αρχικής σελίδας του TbB-Tool

Το TbB-Tool μοντελοποιεί γεωμετρικά το β-βαρέλι (βρίσκει άξονα και ακτίνα), υπολογίζει τη μέση κλίση των β-κλώνων με τον άξονα του β-βαρελιού, βρίσκει θεωρητικά τη τιμή του Shear Number (σύμφωνα με όσα περιγράφονται στο κεφάλαιο 4) και βρίσκει τα διαμεμβρανικά κομμάτια με βάση το προφίλ των εξωτερικών αμινοξέων του βαρελιού (σύμφωνα με όσα περιγράφονται στο κεφάλαιο 5). Υπενθυμίζεται ότι σε 3 από τις 22 πρωτεΐνες που χρησιμοποιήθηκαν στη παρούσα μελέτη, ήταν απαραίτητη μια βελτίωση του άξονα κάνοντας χρήση μόνο των διαμεμβρανικών αμινοξέων που υπέδειξε ο αλγόριθμος καθώς και ο επαναπροσδιορισμός του διαμεμβρανικού τμήματος. Η

ίδια διαδικασία λαμβάνει χώρα και για κάθε πρωτεΐνη που εισάγεται στο TbB-Tool. Έτσι το TbB-Tool παράγει δύο σύνολα αποτελεσμάτων (ένα χωρίς επαναπροσδιορισμό του άξονα και ένα με επαναπροσδιορισμό). Για καθένα από τα δύο σύνολα, ο χρήστης μπορεί να δει στον τρισδιάστατο χώρο τον άξονα του βαρελιού σε σχέση με την πρωτεΐνη και να επιλέξει εκείνο το σύνολο των αποτελεσμάτων που αντιστοιχεί στον καλύτερο άξονα από τους δύο που βρέθηκαν¹⁰.

Ο χρήστης πρέπει να συμπληρώσει σε κατάλληλη φόρμα την ηλεκτρονική του διεύθυνση, ελάχιστη και μέγιστη τιμή για το πάχος της μεμβράνης (προτείνονται οι τιμές 20Å και 32Å αντίστοιχα) καθώς και να φορτώσει το .pdb αρχείο της πρωτεΐνης. (Εικόνα 7.2)



Give your request to the TbB-Tool through the following form:

Enter your email address:

Enter the minimum membrane length in Angstroms (suggested 20.0)

and the maximum membrane length in Angstroms (suggested 32.0)

Choose a .pdb file to upload: Αναζήτηση...

Εικόνα 7.2 Η φόρμα που καλείται να συμπληρώσει ο χρήστης του TbB-Tool.

Κατάλληλη ιστοσελίδα που εμφανίζεται αφού ο χρήστης φορτώσει το .pdb αρχείο της πρωτεΐνης τον ενημερώνει ότι το TbB-Tool ξεκινά τη διεργασία και ότι τα αποτελέσματα θα του σταλούν μέσω ηλεκτρονικού μηνύματος.

Μόλις ολοκληρωθεί η διεργασία, το TbB-Tool στέλνει αυτόματα τα αποτελέσματα στο χρήστη. Το κείμενο του αντίστοιχου ηλεκτρονικού μηνύματος ενημερώνει τον χρήστη ότι το TbB-Tool ολοκλήρωσε την εργασία καθώς και ότι συνημμένα στο ίδιο ηλεκτρονικό μήνυμα είναι τα δύο σύνολα αποτελεσμάτων.

¹⁰ Μπορεί οι δύο άξονες να μη διαφέρουν πολύ οπτικά, αλλά η διαφορά αυτή αντικατοπτρίζεται στα αποτελέσματα εύρεσης ακτίνας, κλίσης των β-κλώνων, Shear Number αλλά και διαμεμβρανικών αμινοξέων.

TbB-Tool

Your .pdb file: 1P4T.PDB has been uploaded and is going to be processed by the Tool.

If you would like to submit a new .pdb file or reload a .pdb file to TbB-Tool Tool click [here](#).

Εικόνα 7. 3 Ιστοσελίδα που εμφανίζεται όταν ο χρήστης φορτώσει το .pdb αρχείο.

Dear User,

Your .pdb file (1P4T.pdb) has been processed by TbB_Tool.

Two result-sets have been sent to you through this email:

One set regarding results with the axis found with no feedback (files named by "no_feedback_for_axis") and one set regarding results with the axis found with feedback (files named by "with_feedback_for_axis") (Please see the web-page of the tool for more details)

Check the two files which visualize the protein, the axis and the membrane walls (file names include: "protein_with_the_axis_and_membrane_walls"). Depending on which axis (with feedback or not) fits better to your eye, use this result-set and ignore the other one. (for finding the axis, the algorithm worked well with no feedback for 19 out of 22 beta-barrel proteins)

Thank you for using TbB_Tool.

Εικόνα 7. 4 Το σώμα κειμένου του ηλεκτρονικού μηνύματος που στέλνεται στο χρήστη TbB-Tool του όταν ολοκληρωθεί η διεργασία.

Για καθένα από τα σύνολα αποτελεσμάτων που στέλνονται στο χρήστη, στο ηλεκτρονικό μήνυμα είναι συνημμένα 5 αρχεία:

- ένα αρχείο κειμένου που περιέχει τα αποτελέσματα (αποτελέσματα γεωμετρικής μοντελοποίησης και εύρεσης του διαμεμβρανικού κομματιού) (Εικόνα 7.5).
- ένα αρχείο .png με το προφίλ των εξωτερικών αμινοξέων (π.χ. εικόνα 6.1, πάνω αριστερά)
- ένα αρχείο .html που απεικονίζει στον τρισδιάστατο χώρο την πρωτεΐνη, τον άξονα και τα τοιχώματα της μεμβράνης (π.χ. εικόνα 6.1, πάνω δεξιά).
- Ένα αρχείο συντεταγμένων .xyz που είναι απαραίτητο για τη προαναφερόμενη οπτικοποίηση.

- Ένα αρχείο .html που απεικονίζει στον τρισδιάστατο χώρο την πρωτεΐνη σε “cartoons” με επιλεγμένα (χρωματισμένα ανάλογα με την αλυσίδα στην οποία βρίσκονται) τα διαμεμβρανικά αμινοξέα (π.χ. εικόνα 6.1, κάτω αριστερά). Για την οπτικοποίηση αυτή είναι απαραίτητο το .pdb αρχείο, το οποίο, επίσης, στέλνεται στο χρήστη ως συνημμένο στο email.

Dear User,

Your .pdb file (1P4T.pdb) has been processed by the TbB-Tool and the results are the following:

The number of the beta-strands of the barrel is 8.

The Optimization procedure for finding the axis of the barrel yielded an axis in 3D space defined by the 2 following Points:

P1(x,y,z): 37.071 14.190 58.531 in Angstroms and

P2(x,y,z): 37.190 19.674 7.528 in Angstroms

The Optimization procedure for finding the axis of the barrel yielded a radius equal to: 8.431 Angstroms.

The mean radius, as calculated based on the distances of all transmembrane beta-stranded residues from the axis, is: 8.266.

The mean angle of the strands with the axis is 44.2 degrees and the theoretically calculated Shear Number(S) is 10.

The membrane thickness was found: 24.9000

The membrane edges (x1=18.37 and x2=43.27) can be observed in the picture of the profile of the external beta-stranded residues along the axis attached with this file.

The TRANSMEMBRANE residues found are the following:

```
Residue :      1 (E) (0)
Residue :      3 (A) (0)
Residues:   5 - 15 (GFYVQADAAHA) (1)
Residues:  28 - 29 (SP) (0)
Residues:  30 - 37 (RISAGYRI) (-1)
Residue :     39 (D) (0)
Residues:  40 - 47 (LRFVDVYT) (1)
Residues:  64 - 71 (IGASAIYD) (-1)
Residue :     72 (F) (0)
Residues:  78 - 87 (VEPYLGARLS) (1)
Residues: 108 - 117 (LGVLTVSYA) (-1)
Residues: 122 - 131 (VDLDAGYRYM) (1)
Residues: 145 - 154 (GELSAQVRVK) (-1)
Residue :    155 (F) (0)
```

{1 stands for a stranded segment with direction from inner to outer wall of the membrane, -1 for a stranded segment with opposite direction and 0 for a non-stranded segment}

You can see the profile of the external beta-stranded residues of the barrel along the axis in the .png file attached with this file.

Attached with this file, you can find a .html file with the beta-barrel protein and the transmembrane residues selected in 3D space.

Also you can see the protein, the membrane walls and the axis in a .html file attached with this file.

(For the visualization of the molecule using the attached .html files you will need the Chime Plugin in your Browser and also your .pdb file copied in the same directory)

Thank you for using the TbB-Tool.

Εικόνα 7. 5 Συνημμένο .txt αρχείο με τα αποτελέσματα του TbB-Tool που στέλνεται στον χρήστη

Τέλος, από την ιστοσελίδα του TbB-Tool ο χρήστης μέσω κατάλληλων συνδέσμων μπορεί να δει τα αποτελέσματα του εργαλείου σε 6 πρωτεΐνες (από το σύνολο των 22 που μελετήθηκαν) ή και τη σχετική βιβλιογραφία που χρησιμοποιήθηκε.

8. Συμπεράσματα

Αντικείμενο της παρούσης διπλωματικής εργασίας αποτέλεσαν οι διαμεμβρανικές πρωτεΐνες τύπου β-βαρελίου. Οι πρωτεΐνες αυτού του τύπου έχουν προσελκύσει το ενδιαφέρον αρκετών ερευνητών μια και μαζί με τις μεμβρανικές πρωτεΐνες τύπου α-έλικας συνιστούν το σύνολο των μεμβρανικών πρωτεϊνών, οι οποίες αποτελούν το 30% των προϊόντων κωδικοποίησης των γονιδίων. Οι διαμεμβρανικές πρωτεΐνες τύπου β-βαρελίου διαπερνούν τις βιολογικές μεμβράνες από το ένα άκρο τους στο άλλο και εμφανίζονται στις εξωτερικές μεμβράνες των αρνητικών κατά Gram βακτηρίων, ενώ πιστεύεται ότι εμφανίζονται και στις εξωτερικές μεμβράνες των μιτοχονδρίων και χλωροπλαστών.

Η παρούσα προσέγγιση εμπεριείχε την επεξεργασία δομών διαμεμβρανικών πρωτεϊνών τύπου β-βαρελίου δεδομένης της τριτοταγούς δομής τους (αμινοξική ακολουθία και συνταγμένες των Ca ατόμων των αμινοξέων) ενώ οι αλγόριθμοι που υλοποιήθηκαν εφαρμόστηκαν σε ένα μη πλεοναστικό σύνολο 22 πρωτεϊνών του τύπου αυτού.

Αρχικά, υλοποιήθηκε αλγόριθμος που, κάνοντας χρήση μεθόδου βελτιστοποίησης ελαχίστων τετραγώνων, υπολογίζει τον άξονα και την ακτίνα κυλίνδρου δεδομένου ενός αριθμού σημείων του. Ο αλγόριθμος αυτός συνδυάστηκε κατάλληλα με γενετικό αλγόριθμο ώστε να μοντελοποιεί γεωμετρικά τα β-βαρέλια, δηλαδή να υπολογίζει τον άξονα και την ακτίνα κυλίνδρου που μοντελοποιεί με τον καλύτερο τρόπο το β-βαρέλι. Μέσω οπτικοποίησης των αποτελεσμάτων, έγινε φανερό ότι ο υβριδικός αλγόριθμος υπολογίζει σωστά τον άξονα και την ακτίνα των β-βαρελίων. Δεδομένης της γεωμετρικής μοντελοποίησης των β-βαρελίων, υπολογίστηκαν η κλίση (α) του άξονα προς του β-κλώνους καθώς και ο Shear Number (S) για κάθε πρωτεΐνη, χαρακτηριστικά που έχουν ήδη προταθεί από ερευνητές για τη γεωμετρική περιγραφή των β-βαρελίων.

Επόμενο βήμα αποτέλεσε η εύρεση του διαμεμβρανικού τμήματος των πρωτεϊνών. Για το σκοπό αυτό, σύμφωνα και με εργασία του ερευνητή Wimley [16], χρησιμοποιήθηκε το προφίλ των εξωτερικών αμινοξέων του β-βαρελίου σε αλληλεπικαλυπτόμενα παράθυρα μήκους 5Å κατά μήκος του άξονα του β-

βαρελιού σχετικά με: i) την υδροφοβικότητα (χρησιμοποιήθηκε πειραματική κλίμακα των Wimley *et al*) ii) την ύπαρξη αρωματικών αμινοξέων και iii) την ύπαρξη φορτισμένων αμινοξέων. Η περισσότερο υδρόφοβη ζώνη, που χαρακτηρίζεται από έντονη παρουσία φορτισμένων αμινοξέων στο εξωτερικό της και από έντονη παρουσία αρωματικών αμινοξέων σε περιοχές γύρω από την αρχή και το τέλος της, θεωρήθηκε ότι αποτελεί τη διαμεμβρανική ζώνη της διαμεμβρανικής πρωτεΐνης. Κατάλληλη συνάρτηση ποιότητας υποδεικνύει τη ζώνη αυτή. Τα αποτελέσματα εύρεσης του διαμεμβρανικού κομματιού συγκρίθηκαν με αυτά της ομάδας του ερευνητή Tusnadi [31]. Τα αποτελέσματα των δύο μεθόδων δε διαφέρουν πολύ, πράγμα που προέκυψε με οπτική επισκόπηση των εξόδων των μεθόδων ή κάνοντας χρήση δεικτών ορθότητας χρησιμοποιώντας καταχρηστικά ως “βάση γνώσης” τα αποτελέσματα του Tusnadi *et al*. Συγκεκριμένα, θεωρώντας ως “βάση γνώσης” τα αποτελέσματα του Tusnadi, η μέθοδος που αναπτύχθηκε είχε ποσοστό ορθών προβλέψεων στη πρόβλεψη “διαμεμβρανικό ή μη” αμινοξύ ίσο με 82.9% στο σύνολο των πρωτεϊνών που χρησιμοποιήθηκαν στη παρούσα μελέτη.

Με βάση τα αποτελέσματα εύρεσης του διαμεμβρανικού τμήματος στο σύνολο των πρωτεϊνών που χρησιμοποιήθηκαν, έγινε μια συνολική μελέτη όσον αφορά την υδροφοβικότητα των αμινοξέων (εσωτερικών και εξωτερικών των β-βαρελίων) κατά μήκος του άξονα, μέσα και έξω από τη μεμβράνη, καθώς και την εμφάνιση συγκεκριμένων αμινοξέων ή ομάδων αμινοξέων (π.χ. αρωματικών ή φορτισμένων). Τα αποτελέσματα μοιάζουν με αυτά εργασίας του Wimley [16] και δείχνουν αποτελέσματα που ήταν αναμενόμενα. (π.χ. έντονη παρουσία αρωματικών αμινοξέων ως εξωτερικά του β-βαρελιού στις διεπιφάνειες της μεμβράνης, έντονη παρουσία φορτισμένων αμινοξέων ως εξωτερικά του β-βαρελιού στο εξωτερικό της μεμβράνης και έντονη παρουσία υδρόφοβων αμινοξέων ως εξωτερικά του β-βαρελιού στο εσωτερικό της μεμβράνης).

Τέλος, οι αλγόριθμοι που υλοποιήθηκαν στα πλαίσια της διπλωματικής εργασίας ενσωματώθηκαν σε κατάλληλο διαδικτυακό εργαλείο, το TbB-Tool. Το εργαλείο δέχεται ως είσοδο ελάχιστη και μέγιστη τιμή για το πάχος της μεμβράνης και το .pdb αρχείο της πρωτεΐνης, ενώ τα αποτελέσματα της γεωμετρικής μοντελοποίησης και εύρεσης του διαμεμβρανικού τμήματος

αποστέλλονται μέσω ηλεκτρονικού μηνύματος σε ηλεκτρονική διεύθυνση που καταχώρησε ο χρήστης στο εργαλείο.

Βιβλιογραφία

1. Θωμόπουλος (1990) Βιολογία του Κυττάρου, Εκδ. UNIVERSITY STUDIO PRESS, Θεσσαλονίκη
2. <http://pdbtm.enzim.hu/>
3. Tamm,L.K., Arora,A., and Kleinschmidt,J.H. (2001) Structure and assembly of beta-barrel membrane proteins J. Biol. Chem., 276, 32399-32402.
4. Tamm,L.K., Heedeok Hong, and Binyong Liang (2004) Folding and assembly of β -barrel membrane protein Biochim. Biophys. Acta, 1666, 250-263.
5. Γ.Ε. Χαλεβελάκης, Ν.Ι. Λεγάκης, Τ.Η. Περόγαμβρος (1997) Αντιβιοτικά και Συνήθεις Λοιμώξεις, Ιατρικές Εκδόσεις Π.Χ. Πασχαλίδης
6. Αντωνιάδης Α, Αντωνιάδης Γρ, Λεγάκης Ν., Τσελέντης Ι., (1998) Ιατρική Μικροβιολογία, Τόμος Ι., Ιατρικές Εκδόσεις Π.Χ. Πασχαλίδης
7. www.jdaros.mc.mail.co/mitochon.htm
8. Becker, Kleinsmith, and Hardin (2002) The world of the cell, 5th Edition, Benjamin-Cummings Publishing Company
9. <http://library.thinkquest.org/C004535/chloroplast.html>
10. Χαμόδρακας Σ.Ι. (1993) Θέματα Μοριακής Βιοφυσικής, Αθήνα, Εκδόσεις Συμμετρία
11. Wimley William (2003) The versatile β -barrel membrane protein Current opinion in Structural Biology, **13**, 404-411.
12. Murzin Alexey, Lesk Arthur, and Clothia Cyrus (1994) Principles Determining the Structure of B-Sheet Barrels in Proteins I. A Theoretical Analysis J. Mol. Biol., **236**, 1369-1381.
13. Murzin Alexey, Lesk Arthur, and Clothia Cyrus (1994) Principles Determining the Structure of B-Sheet Barrels in Proteins II. Observed Structures J. Mol. Biol., **236**, 1382-1400.
14. Schulz,G.E. (2002) The structure of bacterial outer membrane proteins Biochim. Biophys. Acta, **1565**, 308-317.
15. Liu,W.M. (1998) Shear numbers of protein beta-barrels: definition refinements and statistics J. Mol. Biol., **275**, 541-545.
16. Wimley,W.C. (2002) Toward genomic identification of beta-barrel membrane proteins: composition and architecture of known structures Protein Sci., **11**, 301-312.

17. Kreuzsch,A. and Schulz,G.E. (1994) Refined structure of the porin from *Rhodospseudomonas blastica*. Comparison with the porin from *Rhodobacter capsulatus* *J. Mol. Biol.*, **243**, 891-905.
18. Bagos,P.G., Liakopoulos,T.D., and Hamodrakas,S.J. (2005) Evaluation of methods for predicting the topology of beta-barrel outer membrane proteins and a consensus prediction method *BMC. Bioinformatics.*, **6**, 7.
19. Schirmer,T. and Cowan,S.W. (1993) Prediction of membrane-spanning beta-strands and its application to maltoporin *Protein Sci.*, **2**, 1361-1363.
20. Vogel,H. and Jahnig,F. (1986) Models for the structure of outer-membrane proteins of *Escherichia coli* derived from raman spectroscopy and prediction methods *J. Mol. Biol.*, **190**, 191-199.
21. Jacoboni,I., Martelli,P.L., Fariselli,P., De,P., V, and Casadio,R. (2001) Prediction of the transmembrane regions of beta-barrel membrane proteins with a neural network-based predictor *Protein Sci.*, **10**, 779-787.
22. Gromiha,M.M., Ahmad,S., and Suwa,M. (2004) Neural network-based prediction of transmembrane beta-strand segments in outer membrane proteins *J. Comput. Chem.*, **25**, 762-767.
23. Martelli,P.L., Fariselli,P., Krogh,A., and Casadio,R. (2002) A sequence-profile-based HMM for predicting and discriminating beta barrel membrane proteins *Bioinformatics.*, **18 Suppl 1**, S46-S53.
24. Liu,Q., Zhu,Y.S., Wang,B.H., and Li,Y.X. (2003) A HMM-based method to predict the transmembrane regions of beta-barrel membrane proteins *Comput. Biol. Chem.*, **27**, 69-76.
25. Bagos,P.G., Liakopoulos,T.D., Spyropoulos,I.C., and Hamodrakas,S.J. (2004) A Hidden Markov Model method, capable of predicting and discriminating beta-barrel outer membrane proteins *BMC. Bioinformatics.*, **5**, 29.
26. Bagos,P.G., Liakopoulos,T.D., Spyropoulos,I.C., and Hamodrakas,S.J. (2004) PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins *Nucleic Acids Res.*, **32**, W400-W404.
27. Bigelow,H.R., Petrey,D.S., Liu,J., Przybylski,D., and Rost,B. (2004) Predicting transmembrane beta-barrels in proteomes *Nucleic Acids Res.*, **32**, 2566-2577.
28. Natt,N.K., Kaur,H., and Raghava,G.P. (2004) Prediction of transmembrane regions of beta-barrel proteins using ANN- and SVM-based methods *Proteins*, **56**, 11-18.
29. Berven,F.S., Flikka,K., Jensen,H.B., and Eidhammer,I. (2004) BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria *Nucleic Acids Res.*, **32**, W394-W399.

30. Damir Zuzic and Davor Juretic (2004) Precise Annotation of Transmembrane Segments with Garlic - a Free Molecular Visualization Program *Croatica Chemica Acta*, **77**, 397-401.
31. Tusnady,G.E., Dosztanyi,Z., and Simon,I. (2005) TMDet: web server for detecting transmembrane regions of proteins by using their 3D coordinates *Bioinformatics.*, **21**, 1276-1277.
32. Tusnady,G.E., Dosztanyi,Z., and Simon,I. (2005) PDB_TM: selection and membrane localization of transmembrane proteins in the protein data bank *Nucleic Acids Res.*, **33**, D275-D278.
33. Vandeputte-Rutten,L., Bos,M.P., Tommassen,J., and Gros,P. (2003) Crystal structure of Neisserial surface protein A (NspA), a conserved outer membrane protein with vaccine potential *J. Biol. Chem.*, **278**, 24825-24830.
34. Vogt,J. and Schulz,G.E. (1999) The structure of the outer membrane protein OmpX from Escherichia coli reveals possible mechanisms of virulence *Structure.*, **7**, 1301-1309.
35. Hwang,P.M., Choy,W.Y., Lo,E.I., Chen,L., Forman-Kay,J.D., Raetz,C.R., Prive,G.G., Bishop,R.E., and Kay,L.E. (2002) Solution structure and dynamics of the outer membrane enzyme PagP by NMR *Proc. Natl. Acad. Sci. U. S. A*, **99**, 13560-13565.
36. Pautsch,A. and Schulz,G.E. (2000) High-resolution structure of the OmpA membrane domain *J. Mol. Biol.*, **298**, 273-282.
37. Vandeputte-Rutten,L., Kramer,R.A., Kroon,J., Dekker,N., Egmond,M.R., and Gros,P. (2001) Crystal structure of the outer membrane protease OmpT from Escherichia coli suggests a novel catalytic site *EMBO J.*, **20**, 5033-5039.
38. Prince,S.M., Achtman,M., and Derrick,J.P. (2002) Crystal structure of the OpcA integral membrane adhesin from Neisseria meningitidis *Proc. Natl. Acad. Sci. U. S. A*, **99**, 3417-3421.
39. Oomen,C.J., van Ulsen,P., van Gelder,P., Feijen,M., Tommassen,J., and Gros,P. (2004) Structure of the translocator domain of a bacterial autotransporter *EMBO J.*, **23**, 1257-1266.
40. Snijder,H.J., Ubarretxena-Belandia,I., Blaauw,M., Kalk,K.H., Verheij,H.M., Egmond,M.R., Dekker,N., and Dijkstra,B.W. (1999) Structural evidence for dimerization-regulated activation of an integral membrane phospholipase *Nature*, **401**, 717-721.
41. Ye J. and Van den Berg B. (2004) Crystal structure of the bacterial nucleoside transporter Tsx *EMBO J.*, **23**, 3187-3195.
42. Van den Berg B., Black,P.N., Clemons,W.M., Jr., and Rapoport,T.A. (2004) Crystal structure of the long-chain fatty acid transporter FadL *Science*, **304**, 1506-1509.

43. Weiss,M.S. and Schulz,G.E. (1992) Structure of porin refined at 1.8 A resolution J. Mol. Biol., **227**, 493-509.
44. Kreuzsch,A. and Schulz,G.E. (1994) Refined structure of the porin from *Rhodopseudomonas blastica*. Comparison with the porin from *Rhodobacter capsulatus* J. Mol. Biol., **243**, 891-905.
45. Cowan,S.W., Garavito,R.M., Jansonius,J.N., Jenkins,J.A., Karlsson,R., Konig,N., Pai,E.F., Paupit,R.A., Rizkallah,P.J., Rosenbusch,J.P., and . (1995) The structure of OmpF porin in a tetragonal crystal form Structure., **3**, 1041-1050.
46. Dutzler,R., Rummel,G., Alberti,S., Hernandez-Alles,S., Phale,P., Rosenbusch,J., Benedi,V., and Schirmer,T. (1999) Crystal structure and functional characterization of OmpK36, the osmoporin of *Klebsiella pneumoniae* Structure., **7**, 425-434.
47. Zeth,K., Diederichs,K., Welte,W., and Engelhardt,H. (2000) Crystal structure of Omp32, the anion-selective porin from *Comamonas acidovorans*, in complex with a periplasmic peptide at 2.1 A resolution Structure., **8**, 981-992.
48. Cowan,S.W., Schirmer,T., Rummel,G., Steiert,M., Ghosh,R., Paupit,R.A., Jansonius,J.N., and Rosenbusch,J.P. (1992) Crystal structures explain functional properties of two *E. coli* porins Nature, **358**, 727-733.
49. Forst,D., Welte,W., Wacker,T., and Diederichs,K. (1998) Structure of the sucrose-specific porin ScrY from *Salmonella typhimurium* and its complex with sucrose Nat. Struct. Biol., **5**, 37-46.
50. Meyer,J.E., Hofnung,M., and Schulz,G.E. (1997) Structure of maltoporin from *Salmonella typhimurium* ligated with a nitrophenyl-maltotrioxide J. Mol. Biol., **266**, 761-775.
51. Ferguson,A.D., Hofmann,E., Coulton,J.W., Diederichs,K., and Welte,W. (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide Science, **282**, 2215-2220.
52. Buchanan,S.K., Smith,B.S., Venkatramani,L., Xia,D., Esser,L., Palnitkar,M., Chakraborty,R., van der,H.D., and Deisenhofer,J. (1999) Crystal structure of the outer membrane active transporter FepA from *Escherichia coli* Nat. Struct. Biol., **6**, 56-63.
53. Ferguson,A.D., Chakraborty,R., Smith,B.S., Esser,L., van der,H.D., and Deisenhofer,J. (2002) Structural basis of gating by the outer membrane transporter FecA Science, **295**, 1715-1719.
54. Chimento,D.P., Mohanty,A.K., Kadner,R.J., and Wiener,M.C. (2003) Substrate-induced transmembrane signaling in the cobalamin transporter BtuB Nat. Struct. Biol., **10**, 394-401.
55. Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features Biopolymers, **22**, 2577-2637.

56. Wimley, W.C., Creamer, T.P., and White, S.H. (1996) Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides *Biochemistry*, **35**, 5109-5124.
57. Eisenberg D., Schwarz E., Komarony M., and Wall E. (1984) Amino acid scale: Normalized consensus hydrophobicity scale *J. Mol. Biol.*, **179**, 125-142.
58. Jayasinghe S., Hristova K., and White S. (2001) Energetics, stability and prediction of transmembrane helices *J. Mol. Biol.*, **312**, 927-934.
59. Eberly David (2001) Least Squares Fitting of Data, www.magic-software.com,
60. Τζαφέστας Σ. (2001) *Υπολογιστική Νοημοσύνη, Τομος Α: Μεθοδολογίες*
61. Siedlecki W. and Sklansky S. (1989) A Note on Genetic Algorithms for Large-Scale Feature Selection *Pattern Recognition Letters*, **10**, 335-347.
62. Coleman T. and Li Y. (1996) An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds *SIAM Journal on Optimization*, **6**, 418-445.
63. Coleman T. and Li Y. (1994) On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds *Mathematical Programming*, **67**, 189-224.

Παράρτημα

A. Αλγόριθμοι που δεν περιγράφονται στο κύριο σώμα της εργασίας

1) Εύρεση της “best-fit” ευθείας που περνά από σημεία του 3διάστατου χώρου με ορθογώνια παλινδρόμηση

Η γραμμική παρεμβολή, δηλαδή η εύρεση ευθείας που περνά από ένα αριθμό σημείων στο n -διάστατο χώρο, μπορεί να γίνει με τη μέθοδο της ορθογώνιας παλινδρόμησης, όπου η τεχνική των ελαχίστων τετραγώνων ελαχιστοποιεί το ορθογώνιο και όχι το τετραγωνικό σφάλμα [59].

Έστω ότι αναζητούμε την ευθεία που περνά από m σημεία του n -διάστατου χώρου $(\vec{X}_i, i=1..m)$, η οποία σε διανυσματική μορφή δίνεται από τη σχέση:

$$\vec{L}(t) = t \vec{D} + \vec{A} \quad (1)$$

, όπου \vec{D} μοναδιαίο διάνυσμα.

Καθένα από τα m σημεία μπορεί να γραφεί με τη μορφή: $\vec{X}_i = \vec{A} + d_i \vec{D} + p_i \vec{D}_i^\perp$,

όπου $d_i = \vec{D} \cdot (\vec{X}_i - \vec{A})$ και \vec{D}_i^\perp είναι κατάλληλο μοναδιαίο διάνυσμα κάθετο στο

\vec{D} και έχει κατάλληλο συντελεστή p_i .

Έστω, ακόμη, $\vec{Y}_i = \vec{X}_i - \vec{A}$. Το διάνυσμα από το σημείο \vec{X}_i στην προβολή του στην

ευθεία είναι το $\vec{Y}_i - d_i \vec{D} = p_i \vec{D}_i^\perp$ και το τετράγωνο του μήκους του είναι

$p_i^2 = (\vec{Y}_i - d_i \vec{D})^2$. Η συνάρτηση ενέργειας του σφάλματος που θα χρησιμοποιηθεί

στην ελαχιστοποίηση με χρήση των ελάχιστων τετραγώνων είναι

$E(\vec{A}, \vec{D}) = \sum_{i=1}^m p_i^2 \vec{D}_i^\perp$. Η συνάρτηση ενέργειας μπορεί να γραφεί με δύο τρόπους:

$$E(\vec{A}, \vec{D}) = \sum_{i=1}^m (\vec{Y}_i^t [I - \vec{D} \vec{D}^t] \vec{Y}_i) \quad \text{και}$$

$$E(\vec{A}, \vec{D}) = \vec{D}^t \left(\sum_{i=1}^m [(\vec{Y}_i \vec{Y}_i^t) I - \vec{Y}_i \vec{Y}_i^t] \right) \vec{D} = \vec{D}^t M(\vec{A}) \vec{D}$$

όπου ο πίνακας $M(A)$ εξαρτάται από τον πίνακα A .

Χρησιμοποιώντας την πρώτη μορφή της E , η παράγωγος της ως προς το A είναι η

$$\frac{\partial E}{\partial A} = -2[I - \vec{D} \vec{D}^t] \sum_{i=1}^m \vec{Y}_i$$

η οποία μηδενίζεται όταν $\sum_{i=1}^m \vec{Y}_i = 0$ οπότε και

$$\vec{A} = (1/m) \sum_{i=1}^m \vec{X}_i$$

Στη δεύτερη μορφή της E , η ποσότητα $\vec{D}^t M(A) \vec{D}$ είναι τετραγωνικής μορφής και η ελάχιστη τιμή της είναι η μικρότερη ιδιοτιμή του $M(A)$ (σύμφωνα με την άλγεβρα πινάκων και ιδιοτιμών) οπότε και σε αυτή την τιμή ελαχιστοποιείται η συνάρτηση ενέργειας του σφάλματος. Η ιδιοτιμή αυτή μπορεί να βρεθεί, αφού ο $M(A)$ εξαρτάται από το A , το οποίο υπολογίστηκε με χρήση της πρώτης μορφής της E . Το διάνυσμα \vec{D} που ελαχιστοποιεί τη συνάρτηση ενέργειας του σφάλματος, σύμφωνα με τη δεύτερη μορφή της E , είναι το ιδιοδιάνυσμα που αντιστοιχεί στην μικρότερη ιδιοτιμή του $M(A)$. Με τον υπολογισμό του \vec{D} ολοκληρώνεται ο υπολογισμός της ευθείας, η οποία δίνεται από την σχέση (1).

Σημείωση:

Για σημεία του 3-διάστατου χώρου, όπου $A = [a \ b \ c]$, ο πίνακας $M(A)$ δίνεται από τη σχέση:

$$M(A) = \delta \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \sum_{i=1}^m (x_i - a)^2 & \sum_{i=1}^m (x_i - a)(y_i - b) & \sum_{i=1}^m (x_i - a)(z_i - c) \\ \sum_{i=1}^m (x_i - a)(y_i - b) & \sum_{i=1}^m (y_i - b)^2 & \sum_{i=1}^m (y_i - b)(z_i - c) \\ \sum_{i=1}^m (x_i - a)(z_i - c) & \sum_{i=1}^m (y_i - b)(z_i - c) & \sum_{i=1}^m (z_i - c)^2 \end{bmatrix}$$

$$\text{όπου } \delta = \sum_{i=1}^m (x_i - a)^2 + \sum_{i=1}^m (y_i - b)^2 + \sum_{i=1}^m (z_i - c)^2$$

2) Γενετικός Αλγόριθμος

Στην παρούσα παράγραφο περιγράφεται ο Γενετικός Αλγόριθμος (ΓΑ) (Genetic Algorithm-GA), ο οποίος χρησιμοποιήθηκε στα πλαίσια της παρούσης διπλωματικής εργασίας κατά την γεωμετρική μοντελοποίηση του β-βαρελιού των διαμεμβρανικών πρωτεϊνών. Εδώ περιγράφεται ο βασικός ΓΑ, ενώ σε αντίστοιχη παράγραφο παρουσιάζονται οι λεπτομέρειες υλοποίησης του ΓΑ στον αλγόριθμο γεωμετρικής μοντελοποίησης.

Εισαγωγή στους Γενετικούς Αλγορίθμους

Οι ΓΑ [60-61] αναπτύχθηκαν από τον Holland αρχικά μεν για τη μελέτη του φαινομένου της «φυσικής προσαρμογής», μεταγενέστερα δε για την επίλυση πρακτικών προβλημάτων βελτιστοποίησης χωρίς τη χρήση των παραγώγων της αντικειμενικής συνάρτησης (κόστους ή συμπεριφοράς). Οι ΓΑ βασίζονται στην έννοια του πληθυσμού ατόμων, στα οποία εφαρμόζονται οι γενετικοί τελεστές της διασταύρωσης, μετάλλαξης και αντιστροφής. Οι ΓΑ λειτουργούν υπακούοντας στην αρχή της φυσικής επιλογής του Δαρβίνου για την οποία ο Άγγλος φιλόσοφος Herbert Spencer εισήγαγε τον όρο: «Επιβίωση του ισχυρότερου προσαρμοστή».

Ο ΓΑ δεν επιλύει το πρόβλημα με μαθηματικό τρόπο αλλά με βιολογικό. Έτσι έχει μεγάλη ενδογενή ευελιξία και ελευθερία να επιλέγει μια επιθυμητή βέλτιστη λύση σύμφωνα με τις προδιαγραφές σχεδίασης του προβλήματος. Οι ΓΑ έχουν την ίδια δυνατότητα προσδιορισμού της βέλτιστης (ή σχεδόν βέλτιστης) λύσης, ανεξάρτητα αν οι προδιαγραφές είναι μη γραμμικές, διακριτού χρόνου, πολλών ακρότατων, υποκείμενες σε ισοτικούς ή ανισοτικούς περιορισμούς, ή ακόμη μη πολυωνυμικά πλήρεις.

Μέσω των ΓΑ έννοιες από το χώρο της Βιολογίας μεταφέρονται στο χώρο επιστημονικών και τεχνολογικών προβλημάτων διατηρώντας μια αναλογία στις ιδιότητες τους. Έτσι, η έννοια του πληθυσμού αναφέρεται στο σύνολο των υποψήφιων λύσεων του προβλήματος βελτιστοποίησης. Το χρωμόσωμα, που για τη Βιολογία αντιστοιχεί σε μία αλυσίδα DNA, αναφέρεται σε μία υποψήφια λύση του προβλήματος, η οποία κωδικοποιείται με μία ακολουθία δυαδικών ψηφίων. Το γονίδιο αναφέρεται σε ένα τμήμα της ακολουθίας δυαδικών ψηφίων του χρωμοσώματος. Η διασταύρωση αναφέρεται στην ανταλλαγή τμήματος της αλυσίδας μεταξύ των υποψήφιων λύσεων και η μετάλλαξη συνίσταται στην αλλαγή του

δυναδικού ψηφίου μιας τυχαίως επιλεγμένης θέσης με ένα επίσης τυχαία επιλεγμένο δυαδικό ψηφίο. Ο γονότυπος αναφέρεται στη διάταξη των ψηφίων στο χρωμόσωμα που αντιστοιχεί στην υποψήφια λύση του προβλήματος ενώ ο φαινότυπος, όταν αυτός αναφέρεται, στην ίδια τη λύση. Επίσης, όπως στην Βιολογία η ποιότητα ενός οργανισμού (fitness) ορίζεται τυπικά ως η πιθανότητα που έχει ο οργανισμός να αναπαραχθεί, έτσι και στους ΓΑ η προσαρμογή ή ποιότητα μιας λύσης είναι η πιθανότητα της να είναι η νικητήρια σε ένα ανταγωνιστικό περιβάλλον. Από τα άτομα του πληθυσμού, τα οποία διασταυρώνονται μεταξύ τους αλλά και μεταλλάσσονται, προκύπτουν οι απόγονοι, οι οποίοι συμμετέχουν σε νέες διασταυρώσεις, υπόκεινται σε νέες μεταλλάξεις και δίνουν νέους απογόνους. Τελικά, μετά την ολοκλήρωση ενός αριθμού γενεών προκύπτει ο απόγονος που αντιστοιχεί στη βέλτιστη λύση του προβλήματος.

Βασικός Γενετικός Αλγόριθμος

Όπως αναφέρθηκε παραπάνω η ποιότητα ή προσαρμογή του χρωμοσώματος αντανακλά το πόσο «καλό» είναι το χρωμόσωμα, με το οποίο κωδικοποιείται η αντίστοιχη λύση για το πρόβλημα. Η ποιότητα ή προσαρμογή του χρωμοσώματος εκφράζεται από τη τιμή της συνάρτησης ποιότητας του χρωμοσώματος. Μέσω της γενετικής εξέλιξης, το «καλό» χρωμόσωμα τείνει να παράγει υψηλής ποιότητας απόγονο, που σημαίνει καλύτερη λύση στο πρόβλημα. Σε μια πρακτική εφαρμογή ΓΑ, δημιουργείται ένας πληθυσμός χρωμοσωμάτων, των οποίων η αρχική επιλογή είναι τυχαία, μπορεί να ορίζεται όμως και από τον χρήστη του λογισμικού. Το μέγεθος του πληθυσμού αυτού κυμαίνεται από πρόβλημα σε πρόβλημα.

Σε κάθε κύκλο της εξελικτικής διαδικασίας παράγεται μια επόμενη γενιά από τα χρωμοσώματα της παρούσας γενιάς. Αυτό μπορεί να επιτευχθεί μόνο αν μια ομάδα από τα χρωμοσώματα της παρούσας γενιάς επιλεχθούν με μια διαδικασία (ρουτίνα) επιλογής. Τα γονίδια των γονέων ανακατεύονται και ανασυνδυάζονται για την παραγωγή απογόνων στην επόμενη γενιά. Αναμένεται ότι από αυτήν την εξελικτική διαδικασία το «καλύτερο» χρωμόσωμα θα παραγάγει μεγαλύτερο αριθμό απογόνων και συνεπώς έχει μεγαλύτερη πιθανότητα επιβίωσης στις ακόλουθες γενιές, μιμούμενο το φυσικό μηχανισμό της επιβίωσης του ισχυρότερου, όπως είναι γνωστός από το χώρο της Βιολογίας.

Η απλούστερη μορφή ΓΑ περιλαμβάνει τρεις τελεστές: Επιλογή, διασταύρωση και μετάλλαξη. Οι τελεστές αυτοί περιγράφονται παρακάτω:

Επιλογή

Ο τελεστής αυτός επιλέγει χρωμοσώματα στον πληθυσμό προς αναπαραγωγή. Όσο μεγαλύτερη είναι η ποιότητα του χρωμοσώματος (fitter), τόσο αυξάνεται η πιθανότητα του να επιλεγεί περισσότερες φορές προς αναπαραγωγή.

Διασταύρωση

Ο τελεστής αυτός επιλέγει τυχαία μία θέση και ανταλλάσσει τις υποακολουθίες πριν και μετά από αυτή τη θέση μεταξύ δύο χρωμοσωμάτων για να παραγάγει δύο απογόνους. Για παράδειγμα, οι ακολουθίες 10000100 και 11111111 μπορούν να διασταυρωθούν μετά την τρίτη θέση κάθε μιας για να δώσουν τους δύο απογόνους: 10011111 και 11100100. Ο τελεστής διασταύρωσης μιμείται χονδρικά τη βιολογική διασταύρωση μεταξύ δύο απλοειδών (μονού χρωμοσώματος) οργανισμών.

Μετάλλαξη

Ο τελεστής αυτός, που εφαρμόζεται μετά τη διασταύρωση, αλλάζει τυχαία μερικά από τα ψηφία ενός χρωμοσώματος. Για παράδειγμα, η ακολουθία 0000100 θα μπορούσε να μεταλλαχθεί στη δεύτερη θέση, οπότε το αποτέλεσμα θα ήταν 0100100. Η μετάλλαξη μπορεί να λάβει χώρα σε κάθε θέση του χρωμοσώματος με κάποια πιθανότητα, που είναι συνήθως πολύ μικρή (π.χ. 0.001).

Ο εξελικτικός κύκλος επαναλαμβάνεται μέχρι να φθάσει σε κάποιο επιθυμητό κριτήριο τερματισμού. Το όριο αυτό μπορεί να είναι ο αριθμός των γενεών ή το πλήθος των μεταβολών των ατόμων ανάμεσα στις διάφορες γενιές ή μια προκαθορισμένη τιμή της συνάρτησης ποιότητας.

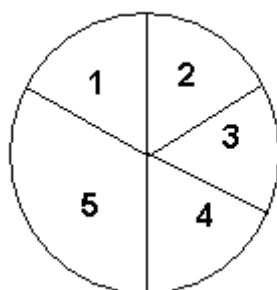
Ένα από τα συνηθέστερα σχήματα επιλογής γονέων είναι ο μηχανισμός «τροχού ρουλέτας», του οποίου τα βήματα περιγράφονται παρακάτω:

Επιλογή «τροχού ρουλέτας»

1. Αθροίζονται οι τιμές της συνάρτησης ποιότητας όλων των μελών του πληθυσμού για να βρεθεί η ολική τιμή της $f_{ολικό}$.
2. Επιλέγεται ένας τυχαίος αριθμός (n) μεταξύ 0 και $f_{ολικό}$.

3. Βρίσκεται το πρώτο μέλος του πληθυσμού του οποίου η τιμή της συνάρτησης ποιότητας, προστιθέμενη στις τιμές της συνάρτησης ποιότητας των προηγούμενων μελών του πληθυσμού, υπερβαίνει το n ή είναι ίση με αυτό.

Για παράδειγμα στο Σχήμα 1 η περιφέρεια του τροχού της ρουλέτας είναι ίση με το $f_{ολικό}$ των πέντε χρωμοσωμάτων. Το χρωμόσωμα 5 είναι το πιο ποιοτικό (fittest) και καταλαμβάνει το μεγαλύτερο τομέα κύκλου, ενώ το χρωμόσωμα 1 είναι το λιγότερο ποιοτικό χρωμόσωμα και καταλαμβάνει το μικρότερο τομέα στον τροχό. Για να επιλεγεί ένα χρωμόσωμα, παράγεται ένας τυχαίος αριθμός στο διάστημα $[0, f_{ολικό}]$ και επιλέγεται το χρωμόσωμα του οποίου ο τομέας προστιθέμενος στους τομείς των προηγούμενων χρωμοσωμάτων καλύπτει τον τυχαίο αυτό αριθμό.



Σχήμα 1. Επιλογή γονέων με το μηχανισμό της ρουλέτας

Η πιθανότητα της διασταύρωσης, P_{δ} συνήθως έχει τιμή μεταξύ 0.6 και 1.0. Αντίστοιχα, η πιθανότητα P_{μ} της μετάλλαξης είναι μικρότερη από 0.1. Η επιλογή των παραμέτρων P_{δ} και P_{μ} απαιτεί την επίλυση ενός πολύπλοκου προβλήματος βελτιστοποίησης.

Ενδεικτικές προτεινόμενες τιμές της πιθανότητας διασταύρωσης P_{δ} και της πιθανότητας μετάλλαξης P_{μ} είναι οι εξής :

Για μεγάλους πληθυσμούς: Ρυθμός διασταύρωσης: $P_{\delta} = 0.6$ και Ρυθμός μετάλλαξης: $P_{\mu} = 0.001$

Για μικρούς πληθυσμούς: Ρυθμός διασταύρωσης: $P_{\delta} = 0.9$ και Ρυθμός μετάλλαξης: $P_{\mu} = 0.01$

Τα βήματα του βασικού ΓΑ παρουσιάζονται παρακάτω:

Αρχικός Χρόνος : $0 \leftarrow t$

Βήμα 0: Επιλέγεται ένας τυχαίος αρχικός πληθυσμός $P(t)$ και υπολογίζεται η αρχική τιμή της συνάρτησης ποιότητας αυτού.

Κύκλος Γενετικού Αλγορίθμου

Βήμα 1: Μέχρι να ικανοποιηθεί το κριτήριο τερματισμού:

- Αυξάνεται ο χρόνος κατά 1
- Επιλέγεται ένας υποπληθυσμός για παραγωγή απογόνων

$$P' \leftarrow \text{επιλεγέντες γονείς } P(t)$$

- Διασταυρώνεται το $P'(t)$ με ρυθμό διασταύρωσης P_δ
- Μεταλλάσσεται το $P'(t)$ με ρυθμό μετάλλαξης P_μ
- Υπολογίζεται η νέα τιμή της συνάρτησης ποιότητας του $P'(t)$
- Επιλέγονται τα άτομα που θα επιβιώσουν από την τιμή της συνάρτησης

ποιότητας:

$$P \leftarrow P'(t)$$

Βήμα 2 : Όταν ικανοποιηθεί το κριτήριο τερματισμού σταματά η διαδικασία, αλλιώς γίνεται επαναφορά στο Βήμα 1.

3) Αλγόριθμος ελαχιστοποίησης για μη γραμμικά προβλήματα ελαχίστων τετραγώνων

Η ελαχιστοποίηση

$\min \left\langle \left(\sum_{i=1}^N (\text{dist}(P_i, \varepsilon) - R)^2 \right) \right\rangle$ η οποία πρέπει να γίνει για την εύρεση του άξονα

κυλίνδρου δεδομένου ενός αριθμού σημείων του σύμφωνα με όσα περιγράφονται στην παράγραφο 4.1, γίνεται με χρήση αλγορίθμου κατάλληλου για μη γραμμικά προβλήματα ελαχίστων τετραγώνων (least square non linear: **lsqnonlin**). Ο αλγόριθμος ανήκει στην κατηγορία των “subspace trust region” αλγορίθμων. Αυτό σημαίνει ότι σε κάθε βήμα ελαχιστοποίησης σε ένα σημείο x η συνάρτηση προς ελαχιστοποίηση (f) προσεγγίζεται από μία πιο απλή συνάρτηση (g) που έχει την ίδια συμπεριφορά σε μία μικρή περιοχή εμπιστοσύνης (“subspace trust region”) γύρω από το σημείο x . Η πιο απλή συνάρτηση g που χρησιμοποιείται ορίζεται χρησιμοποιώντας τη προσέγγιση των 2 πρώτων όρων της σειράς Taylor που προσεγγίζει την f στο σημείο x . Η περιοχή εμπιστοσύνης υπολογίζεται κάνοντας χρήση των αλγορίθμων PCG (Preconditioned conjugate gradient) και του πίνακα Hessian του προβλήματος ελαχιστοποίησης. Ο αλγόριθμος “lsqnonlin” παρουσιάζεται και στις [62-63].

B. Εργαλεία που χρησιμοποιήθηκαν

Το πρόγραμμα DSSP χρησιμοποιήθηκε για την εύρεση των αμινοξέων που απαρτίζουν τους β-κλώνους του β-βαρελιού των διαμεμβρανικών πρωτεϊνών. Το πρόγραμμα αυτό χρησιμοποιεί ως είσοδο το .pdb αρχείο της πρωτεΐνης, ενώ είναι ελεύθερα διαθέσιμο ως εκτελέσιμο αρχείο.

Για την υλοποίηση των μεθόδων βελτιστοποίησης που χρησιμοποιήθηκαν (αλγόριθμος “Isqnonlin” και γενετικοί αλγόριθμοι), έγινε χρήση των αντίστοιχων εργαλείων σε Matlab. Ήταν βέβαια δυνατή η περαιτέρω μεταγλώττιση του κώδικα σε αρχεία της C, η σύνδεση του κώδικα με τον υπόλοιπο κώδικα που υλοποιήθηκε σε Matlab και μεταγλωττίστηκε σε C και η δημιουργία τελικά ενός εκτελέσιμου αρχείου που υλοποιεί όλη την εφαρμογή. Το τελικό πρόγραμμα δέχεται ως είσοδο το αρχείο .pdb μιας πρωτεΐνης καθώς και τα όρια για το πάχος της μεμβράνης.

Για τη δημιουργία του διαδικτυακού εργαλείου έγινε χρήση της γλώσσας HTML και της γλώσσας PHP. Η γλώσσα PHP μπορεί και εκτελεί το εκτελέσιμο αρχείο που δημιουργήθηκε, όπως περιγράφεται παραπάνω, και μέσω κατάλληλης κλάσης να στέλνει τα αποτελέσματα μέσω ηλεκτρονικού μηνύματος στο χρήστη του εργαλείου.

Για την τρισδιάστατη οπτικοποίηση αποτελεσμάτων (αρχεία που στέλνονται στο χρήστη του διαδικτυακού εργαλείου και αρχεία που βρίσκονται στο συνοδευτικό ηλεκτρονικό υλικό) είναι απαραίτητο το πρόγραμμα Chime.

Γ. Διαμεμβρανικά αμινοξέα των 22 πρωτεϊνών που μελετήθηκαν

Στο σημείο αυτό παρατίθενται τα αποτελέσματα εύρεσης του διαμεμβρανικού κομματιού στο σύνολο των 22 πρωτεϊνών που χρησιμοποιήθηκε. Παρατίθενται τα διαμεμβρανικά αμινοξέα σε συνεχή τμήματα ακολουθίας τα οποία χαρακτηρίζονται με 1 εάν πρόκειται για τμήμα κλώνου με φορά από το εσωτερικό τοίχωμα προς τα έξω, -1 εάν πρόκειται για τμήμα κλώνου με αντίθετη φορά και 0 εάν τα αμινοξέα δεν ανήκουν σε β-κλώνο. Η αρίθμηση των αμινοξέων είναι ίδια με αυτή στα .pdb αρχεία.

1AOS

```
Residues: 73 - 85 (FEFHGYARSGVIM) (1)
Residues: 95 - 96 (GA) (0)
Residues: 107 - 116 (AIGRLGNQAD) (0)
Residues: 117 - 126 (TYVEMNLEHK) (-1)
Residues: 135 - 145 (TRFKVMVADGQ) (1)
```

Residues: 146 - 153 (TSYNDWTA) (0)
 Residues: 156 - 158 (SDL) (0)
 Residues: 159 - 168 (NVRQAFVELG) (-1)
 Residues: 176 - 177 (PF) (0)
 Residue : 180 (S) (0)
 Residues: 181 - 188 (TLWAGKRF) (1)
 Residues: 189 - 205 (DRDNFDIHWIDSDVVFL) (0)
 Residues: 206 - 217 (AGTGGGIYDVKW) (-1)
 Residue : 218 (N) (0)
 Residue : 220 (G) (0)
 Residues: 221 - 230 (LRSNFSLYGR) (1)
 Residues: 245 - 253 (YILTMNHFA) (-1)
 Residues: 254 - 255 (GP) (0)
 Residues: 256 - 263 (LQMMVSGL) (1)
 Residues: 288 - 298 (VHALLGLHNS) (-1)
 Residues: 299 - 303 (FYGLR) (0)
 Residues: 304 - 314 (DGSSKTALLYG) (1)
 Residues: 321 - 325 (VKGIG) (0)
 Residues: 337 - 346 (WRIASYGTTP) (-1)
 Residues: 347 - 348 (LS) (0)
 Residue : 350 (N) (0)
 Residues: 351 - 359 (WSVAPAMLA) (1)
 Residues: 375 - 384 (ATFNLRLIQA) (-1)
 Residues: 385 - 386 (IN) (0)
 Residue : 388 (N) (0)
 Residues: 389 - 397 (FALAYEGSY) (1)
 Residues: 418 - 427 (YKLTFAPTFK) (-1)
 Residues: 428 - 438 (VGSIGDFFSRP) (0)
 Residues: 439 - 446 (EIRFYTSW) (1)
 Residues: 474 - 482 (SFGVQMETW) (-1)
 Residue : 483 (F) (0)

1E54

Residues: 5 - 13 (TLFGIVDTN) (1)
 Residue : 31 (T) (-1)
 Residues: 32 - 37 (SGASTS) (0)
 Residues: 38 - 43 (RLGLRG) (-1)
 Residues: 54 - 60 (GFWLEGE) (1)
 Residues: 61 - 63 (IFG) (0)
 Residue : 66 (G) (0)
 Residues: 71 - 75 (FNFKR) (0)
 Residues: 76 - 81 (RSTVSL) (-1)
 Residues: 87 - 90 (EVRL) (1)
 Residues: 91 - 111 (GRDLVPTSQKLTSDYDLFSATG) (0)
 Residues: 114 - 115 (PF) (0)
 Residues: 132 - 136 (IRANN) (0)
 Residues: 137 - 141 (LISYY) (-1)
 Residue : 142 (T) (0)
 Residues: 148 - 155 (FNAGFGYA) (1)
 Residue : 156 (F) (0)
 Residues: 171 - 179 (YIGGYVAYD) (-1)
 Residue : 182 (P) (0)
 Residues: 183 - 190 (LSASLGLA) (1)
 Residues: 206 - 213 (ITLGASYN) (-1)
 Residues: 214 - 217 (FGVA) (0)
 Residues: 218 - 224 (KLSGLLQ) (1)
 Residues: 241 - 248 (YMLGASAP) (-1)
 Residues: 249 - 252 (VGGV) (0)
 Residues: 253 - 259 (GEVKLQY) (1)
 Residues: 275 - 281 (TLGYVHN) (-1)
 Residues: 282 - 285 (LSKR) (0)
 Residues: 286 - 291 (TALYGN) (1)
 Residues: 326 - 331 (VGIRHA) (-1)
 Residue : 332 (F) (0)

1FEP

Residues: 15 - 42 (VVTAAEQNLQAPGVSTITADEIRKNPVA) (0)
 Residues: 44 - 51 (DVSKIIRT) (0)
 Residues: 53 - 55 (PGV) (0)
 Residues: 84 - 90 (ILIDGKP) (0)
 Residue : 111 (T) (0)
 Residues: 113 - 117 (WVPPE) (0)

Residues: 119 - 134 (IERIEVLRGPAARYG) (0)
 Residues: 140 - 150 (GVVNIITKKGS) (0)
 Residues: 154 - 162 (HGSWDAYFN) (1)
 Residues: 173 - 181 (KRTNFSLTG) (-1)
 Residue : 183 (L) (0)
 Residues: 187 - 196 (FSFRLYGNLD) (1)
 Residues: 232 - 240 (DINGVVRWD) (-1)
 Residues: 245 - 253 (QSLELEAGY) (1)
 Residues: 288 - 296 (NYALTWNGG) (-1)
 Residue : 301 (V) (0)
 Residues: 302 - 309 (TTSNWWQY) (1)
 Residues: 347 - 355 (LHSEVNLPI) (-1)
 Residue : 357 (F) (0)
 Residues: 361 - 369 (QTLTLGTEW) (1)
 Residues: 411 - 419 (SLFAENN E) (-1)
 Residue : 420 (L) (0)
 Residue : 423 (S) (0)
 Residues: 424 - 431 (TIVTPGLR) (1)
 Residues: 444 - 451 (PALNISQG) (-1)
 Residue : 452 (L) (0)
 Residue : 455 (D) (0)
 Residues: 456 - 463 (FTLK GIA) (1)
 Residues: 510 - 517 (KEIGLEFK) (-1)
 Residue : 520 (G) (0)
 Residues: 521 - 528 (WLAGVTWF) (1)
 Residues: 566 - 574 (LEGLNVPV) (-1)
 Residues: 578 - 585 (V WTNNIT) (1)
 Residues: 607 - 614 (LNSTLSWQ) (-1)
 Residues: 619 - 627 (LS QTTFW) (1)
 Residues: 656 - 664 (IVGLSATWD) (-1)
 Residue : 665 (V) (0)
 Residue : 668 (N) (0)
 Residues: 669 - 676 (VSLTGGVD) (1)
 Residues: 677 - 678 (NL) (0)
 Residues: 715 - 723 (TWY SVNTH) (-1)
 Residue : 724 (F) (0)

1178

Residue : 2 (T) (0)
 Residues: 8 - 11 (TPDN) (0)
 Residues: 12 - 20 (INADISLGT) (1)
 Residues: 51 - 58 (IIKGAINW) (-1)
 Residues: 59 - 63 (DLMPQ) (0)
 Residues: 64 - 71 (ISIGAAGW) (1)
 Residues: 112 - 122 (FDLNIKWLLN) (-1)
 Residues: 126 - 135 (YRLGLMAGYQ) (1)
 Residues: 179 - 189 (PYIGLTGSYRY) (-1)
 Residues: 192 - 200 (FELGGTFKY) (1)
 Residues: 232 - 241 (SVAVNAGYYV) (-1)
 Residue : 244 (N) (0)
 Residues: 245 - 253 (AKVYVEGAW) (1)
 Residues: 288 - 296 (ITAGLKYT) (-1)
 Residue : 297 (F) (0)

1K24

Residues: 7 - 8 (AN) (0)
 Residues: 9 - 18 (EFTVHTDLSS) (1)
 Residues: 33 - 43 (AKHIGVRADIP) (-1)
 Residues: 44 - 47 (FDAN) (0)
 Residues: 48 - 57 (QGIRLEAGFG) (1)
 Residues: 90 - 100 (IDLYTGYTYTQ) (-1)
 Residue : 102 (L) (-1)
 Residues: 106 - 117 (LNFRVGAGLGFE) (1)
 Residues: 137 - 149 (WLAKVHADLSQL) (-1)
 Residues: 153 - 165 (WYINPWSEVKFDL) (1)
 Residues: 189 - 201 (WGFGLGANIGKKL) (-1)
 Residue : 204 (S) (0)
 Residues: 205 - 216 (ASIEAGPFYKQR) (1)
 Residues: 243 - 252 (REYGLRVGIK) (-1)
 Residue : 253 (F) (0)

1KMO

Residues: 85 - 118 (VGDWLGDARENDVFEHAGARDVIRREDFAKTGAT) (0)
 Residue : 120 (M) (0)
 Residues: 123 - 124 (VL) (0)

Residues: 126 - 128 (RIP) (0)
 Residues: 163 - 167 (LMDGI) (0)
 Residues: 184 - 203 (VSLGNMDAIDVVRGGGAVRY) (0)
 Residues: 211 - 220 (VVNFVTRAIP) (0)
 Residues: 225 - 231 (IEAGVEG) (1)
 Residues: 245 - 251 (THNLMVG) (-1)
 Residues: 259 - 266 (GTALLYSG) (1)
 Residues: 282 - 289 (DLMLKSKY) (-1)
 Residues: 296 - 303 (TFNSLLQY) (1)
 Residues: 339 - 346 (LASLGYQF) (-1)
 Residues: 353 - 359 (KFNIQGF) (1)
 Residues: 386 - 393 (GIEPRYSQ) (-1)
 Residues: 402 - 409 (HEVGVGYR) (1)
 Residues: 449 - 456 (WYLDDKID) (-1)
 Residues: 460 - 467 (WTITPGMR) (1)
 Residues: 493 - 501 (PALNVLYHL) (-1)
 Residues: 505 - 512 (WNLYTE) (1)
 Residues: 541 - 547 (ELGTRYD) (-1)
 Residues: 551 - 557 (LTAEMGL) (1)
 Residues: 587 - 593 (ETQARYD) (-1)
 Residue : 594 (L) (0)
 Residue : 601 (L) (0)
 Residue : 603 (N) (0)
 Residues: 604 - 610 (VSIYASY) (1)
 Residues: 637 - 642 (TLGVDY) (-1)
 Residue : 643 (K) (0)
 Residues: 647 - 653 (WTFNLNS) (1)
 Residues: 685 - 691 (GARVAYD) (-1)
 Residue : 699 (L) (0)
 Residues: 700 - 705 (NLAFGV) (1)
 Residues: 734 - 740 (YMQGSLK) (-1)

IMM4

Residues: 24 - 28 (YDLYI) (1)
 Residue : 30 (A) (0)
 Residue : 32 (T) (0)
 Residue : 34 (H) (0)
 Residue : 38 (A) (0)
 Residues: 56 - 59 (FGLS) (-1)
 Residues: 68 - 71 (HGLY) (1)
 Residues: 88 - 91 (YGWE) (-1)
 Residues: 106 - 109 (LGFT) (1)
 Residues: 128 - 132 (PLASV) (-1)
 Residues: 138 - 142 (TFQMT) (1)
 Residue : 156 (A) (0)
 Residues: 157 - 160 (WMRF) (-1)

INQE

Residues: 6 - 42 (DTLVVTANRFEQPRSTVLAPTTVVTRQDIDRWQSTSV) (0)
 Residues: 44 - 52 (DVLRRLLPGV) (0)
 Residues: 78 - 83 (VLIDGV) (0)
 Residues: 100 - 118 (PIALVQRVEYIRGPRSAVY) (0)
 Residues: 125 - 136 (GVVNIITRDEP) (0)
 Residues: 137 - 145 (GTEISAGWG) (1)
 Residues: 150 - 159 (QNYDVSTQQQ) (-1)
 Residue : 160 (L) (0)
 Residues: 164 - 173 (TRVTLLGDYA) (1)
 Residues: 201 - 209 (TLYGALEHN) (-1)
 Residue : 210 (F) (0)
 Residues: 214 - 222 (WSGFVRGYG) (1)
 Residues: 249 - 257 (SWDAGLRYN) (-1)
 Residues: 261 - 268 (IKSQLITS) (1)
 Residues: 297 - 305 (TVQWANNVI) (-1)
 Residue : 306 (V) (0)
 Residues: 309 - 316 (GSIGAGVD) (1)
 Residues: 340 - 347 (IYLTGLQQ) (-1)
 Residues: 351 - 358 (FTFEGAAR) (1)
 Residues: 371 - 379 (WQTSAGWEF) (-1)
 Residues: 383 - 390 (YRFIASYG) (1)
 Residues: 418 - 426 (WEGAFEGLT) (-1)
 Residues: 429 - 436 (VNWRISGY) (1)
 Residues: 464 - 472 (VEATANFDT) (-1)
 Residues: 475 - 482 (LTHTVSYD) (1)
 Residues: 503 - 510 (VKYQLDWQ) (-1)
 Residue : 514 (F) (0)
 Residues: 515 - 521 (DWGITYQ) (1)

Residues: 546 - 554 (LWDLAVAYP) (-1)
Residue : 555 (V) (0)
Residue : 558 (H) (0)
Residues: 559 - 566 (LTVRGKIA) (1)
Residues: 567 - 568 (NL) (0)
Residues: 586 - 593 (YTLSGSYT) (-1)
Residue : 594 (F) (0)

1OSM

Residue : 1 (A) (0)
Residues: 2 - 4 (EIY) (-1)
Residues: 10 - 20 (KLDLYGKIDGL) (1)
Residue : 39 (T) (0)
Residues: 40 - 49 (YMRLGVKGET) (-1)
Residues: 56 - 65 (TGYGQWEYNV) (1)
Residues: 80 - 89 (WTRLAFAGLK) (-1)
Residue : 90 (F) (0)
Residues: 93 - 94 (AG) (0)
Residues: 95 - 102 (SFDYGRNY) (1)
Residues: 103 - 120 (GVVYDVTSWTDVLPFEGG) (0)
Residues: 129 - 131 (LQS) (0)
Residues: 132 - 141 (RANGVATYRN) (-1)
Residues: 142 - 145 (SDFP) (0)
Residue : 150 (G) (0)
Residues: 151 - 158 (LNFALQYQ) (1)
Residue : 159 (G) (0)
Residues: 173 - 180 (GFGTSVTY) (-1)
Residues: 181 - 182 (IF) (-1)
Residues: 183 - 184 (DG) (0)
Residues: 185 - 193 (ISAGFAYAN) (1)
Residues: 214 - 222 (YTGGLKYDA) (-1)
Residues: 223 - 224 (NN) (0)
Residues: 225 - 232 (IYLATQYT) (1)
Residues: 257 - 263 (FEVAAQY) (-1)
Residues: 264 - 268 (QDFDG) (0)
Residues: 269 - 275 (LRPSVAY) (1)
Residues: 296 - 304 (DVGATYYFN) (-1)
Residues: 305 - 306 (KN) (0)
Residues: 307 - 312 (MSTYVD) (1)
Residues: 334 - 339 (LGLVYQ) (-1)
Residue : 340 (F) (0)

1P4T

Residue : 1 (E) (0)
Residue : 3 (A) (0)
Residues: 5 - 15 (GFYVQADAAHA) (1)
Residues: 28 - 29 (SP) (0)
Residues: 30 - 37 (RISAGYRI) (-1)
Residue : 39 (D) (0)
Residues: 40 - 47 (LRFVAVDT) (1)
Residues: 64 - 71 (IGASAIYD) (-1)
Residue : 72 (F) (0)
Residues: 78 - 87 (VKPYLGARLS) (1)
Residues: 108 - 117 (LGVLTVGSYA) (-1)
Residues: 122 - 131 (VDLDAGYRYN) (1)
Residues: 145 - 154 (GELSAGVRVK) (-1)
Residue : 155 (F) (0)

1PHO

Residues: 1 - 4 (AEIY) (0)
Residue : 9 (N) (0)
Residues: 10 - 14 (KLDVY) (1)
Residue : 15 (G) (0)
Residues: 16 - 20 (KVKAM) (1)
Residue : 39 (S) (0)
Residues: 40 - 49 (YIRFGFKGET) (-1)
Residues: 56 - 65 (TGYRWEEAF) (1)
Residues: 81 - 90 (TRLAFAGLKY) (-1)
Residue : 93 (L) (0)
Residues: 94 - 102 (GSFDYGRNL) (1)
Residues: 103 - 120 (GALYDVEAWTDMFPEFGG) (0)
Residues: 129 - 131 (MTK) (0)
Residues: 132 - 142 (RASGLATYRNT) (-1)
Residues: 143 - 145 (DFF) (0)
Residues: 148 - 150 (IDG) (0)
Residues: 151 - 158 (LNLTLQYQ) (1)

Residue : 159 (G) (0)
 Residues: 173 - 182 (GFGTSLTYDF) (-1)
 Residue : 183 (S) (0)
 Residues: 183 - 184 (SD) (0)
 Residues: 185 - 192 (FAISGAYT) (1)
 Residues: 214 - 222 (WATGLKYDA) (-1)
 Residues: 223 - 224 (NN) (0)
 Residues: 225 - 232 (IYLATFYS) (1)
 Residues: 257 - 263 (FEAVAQY) (-1)
 Residues: 264 - 268 (QFDFG) (0)
 Residues: 269 - 275 (LRPSLGY) (1)
 Residues: 296 - 302 (DVGATYY) (-1)
 Residues: 303 - 306 (FNKN) (0)
 Residues: 307 - 312 (MSAFVD) (1)
 Residues: 334 - 339 (VGMTYQ) (-1)
 Residue : 340 (F) (0)

IPRN

Residue : 1 (E) (0)
 Residues: 2 - 12 (ISLNGYGRFGL) (1)
 Residues: 28 - 39 (SSRLRINIVGTT) (-1)
 Residues: 46 - 56 (TFGAKLRMQWD) (1)
 Residues: 64 - 69 (TAGNAA) (0)
 Residues: 70 - 74 (QFWTS) (-1)
 Residue : 77 (G) (0)
 Residues: 78 - 83 (VTVSVG) (1)
 Residues: 84 - 106 (NVDTAFDVSVALTYDSEMGEASS) (0)
 Residues: 108 - 109 (GD) (0)
 Residues: 116 - 120 (AYNSK) (0)
 Residues: 131 - 139 (NGIAVTYSI) (-1)
 Residue : 141 (G) (0)
 Residues: 142 - 149 (VNLYLSYV) (1)
 Residues: 164 - 171 (FGIAADWS) (-1)
 Residue : 172 (N) (0)
 Residue : 174 (M) (0)
 Residues: 175 - 182 (ISLAAAYT) (1)
 Residues: 194 - 200 (AFVGAAY) (-1)
 Residues: 201 - 205 (KFNDA) (0)
 Residues: 206 - 212 (GTVGLNW) (1)
 Residues: 224 - 232 (VTLYGNYAF) (-1)
 Residues: 233 - 234 (GA) (0)
 Residues: 235 - 241 (TTVRAYV) (1)
 Residues: 254 - 262 (GIGADYQFA) (-1)
 Residues: 263 - 264 (EG) (0)
 Residues: 265 - 271 (VKVSGSV) (1)
 Residues: 282 - 288 (DVGVRFD) (-1)
 Residue : 289 (F) (0)

IQD5

Residues: 16 - 19 (GSII) (0)
 Residues: 30 - 32 (FTL) (0)
 Residues: 33 - 34 (YP) (1)
 Residues: 35 - 38 (YDTN) (0)
 Residues: 39 - 41 (YLI) (1)
 Residues: 69 - 78 (FQLSLAFPLW) (-1)
 Residues: 87 - 95 (VLGASYTQK) (1)
 Residues: 114 - 121 (YEPQLFLG) (-1)
 Residues: 135 - 142 (DVEMGYNH) (1)
 Residues: 157 - 165 (RLYTRLMAE) (-1)
 Residues: 169 - 173 (WLVEV) (1)
 Residues: 174 - 175 (KP) (0)
 Residues: 176 - 177 (WY) (1)
 Residues: 194 - 196 (YYQ) (0)
 Residues: 197 - 203 (LKIGYHL) (-1)
 Residue : 205 (D) (0)
 Residues: 206 - 213 (AVLSAKGQ) (1)
 Residues: 223 - 230 (AELGLSYP) (-1)
 Residues: 231 - 235 (ITKHV) (0)
 Residues: 236 - 241 (RLYTQV) (1)
 Residues: 261 - 264 (VGVM) (-1)
 Residues: 265 - 269 (LNDLF) (0)

IQJ8

Residues: 3 - 10 (STVTGGYA) (1)
 Residues: 24 - 31 (FNLKYRYE) (-1)
 Residues: 37 - 44 (LGVIGSFT) (1)

Residues: 64 - 72 (GITAGPAYR) (-1)
Residues: 77 - 85 (ASIYGVVGV) (1)
Residues: 108 - 115 (SYGAGLQF) (-1)
Residues: 116 - 117 (NP) (0)
Residues: 121 - 129 (VALDFSSEQ) (1)
Residues: 139 - 147 (TWIAGVGYR) (-1)

IQJP

Residue : 1 (A) (0)
Residues: 7 - 16 (WYTGAKLGWS) (1)
Residues: 35 - 45 (LGAGAFGGYQV) (-1)
Residue : 48 (Y) (0)
Residues: 49 - 58 (VGFEMGYDWL) (1)
Residues: 75 - 85 (QGVQLTAKLGY) (-1)
Residue : 87 (I) (0)
Residues: 91 - 102 (LDIYTRLGGMVW) (1)
Residues: 119 - 130 (VSPVFAGGVEYA) (-1)
Residues: 135 - 144 (IATRLEYQWT) (1)
Residue : 145 (N) (0)
Residue : 160 (G) (0)
Residues: 161 - 169 (MLSLGVSYR) (-1)
Residue : 170 (F) (0)

ITIL

Residue : 1 (A) (0)
Residues: 4 - 40 (QLNEFSSSGLGRAYSGEGAIADDAGNVSARNPALITMF) (0)
Residue : 43 (P) (0)
Residues: 44 - 52 (TFSAGAVYI) (1)
Residues: 77 - 87 (AWVPNMHFVAP) (-1)
Residues: 92 - 100 (FGWGASITS) (1)
Residue : 101 (N) (0)
Residues: 126 - 135 (MNLNLSGAYR) (-1)
Residue : 136 (L) (0)
Residue : 139 (A) (0)
Residues: 140 - 149 (WSFGLGFNAV) (1)
Residues: 207 - 216 (FGWNAGILYE) (-1)
Residues: 221 - 228 (NRYALTYR) (1)
Residue : 229 (S) (0)
Residues: 273 - 276 (EMWE) (0)
Residues: 277 - 281 (VSGYN) (-1)
Residue : 282 (R) (0)
Residues: 287 - 296 (WAIHYSLAYT) (1)
Residues: 325 - 333 (YRIALGTTY) (-1)
Residues: 339 - 340 (WT) (0)
Residues: 341 - 348 (FRTGIAFD) (1)
Residues: 367 - 375 (FWLSAGTTY) (-1)
Residues: 382 - 390 (SVDVGVSYM) (1)
Residues: 411 - 419 (WLFGTNFNY) (-1)

ITLY

Residues: 13 - 23 (WHQSVNVVGSY) (1)
Residues: 24 - 25 (HT) (0)
Residues: 33 - 34 (ND) (0)
Residues: 35 - 44 (TYLEYEAFK) (-1)
Residue : 47 (W) (0)
Residues: 48 - 57 (FDIFYGYADAP) (1)
Residue : 58 (V) (0)
Residues: 75 - 76 (PL) (0)
Residues: 77 - 85 (FMEIEPRFS) (-1)
Residues: 86 - 87 (ID) (0)
Residues: 89 - 90 (LT) (0)
Residue : 94 (L) (0)
Residues: 96 - 98 (FGP) (0)
Residue : 99 (F) (1)
Residues: 101 - 111 (EWYFANNYIYD) (1)
Residues: 121 - 131 (STWYMGLGTDI) (-1)
Residue : 133 (T) (0)
Residue : 137 (M) (0)
Residues: 138 - 147 (SLSMNVYAKY) (1)
Residues: 163 - 171 (YRFKIKYFV) (-1)
Residues: 181 - 190 (LSYIGFTNFD) (1)
Residues: 214 - 223 (IASSHILALN) (-1)
Residues: 227 - 236 (WHYSVVARYW) (1)
Residues: 262 - 272 (WGGYLVVGYNF) (-1)

IUYN

Residues: 802 - 815 (GRRLKAVSDGLDHN) (0)

Residues: 819 - 827 (LRVIAQTQQ) (1)
 Residues: 845 - 855 (TQTVGIAAKTG) (-1)
 Residue : 857 (N) (0)
 Residues: 858 - 866 (TTAAATLGM) (1)
 Residues: 884 - 893 (SLFAGIRHDA) (-1)
 Residue : 896 (I) (0)
 Residues: 897 - 904 (GYLKGLFS) (1)
 Residues: 933 - 939 (LGALGGV) (-1)
 Residues: 940 - 942 (NVP) (0)
 Residues: 948 - 956 (DLTVEGLLR) (1)
 Residues: 986 - 994 (GLAGLKLSQ) (-1)
 Residues: 1000 -1009 (AVLFATAGVE) (1)
 Residues: 1043 -1052 (LVAGLGADVE) (-1)
 Residues: 1057 -1066 (WNGLARYSYA) (1)
 Residues: 1074 -1083 (HSGRVGVGYR) (-1)

2FCP

Residues: 20 - 74 (SAWGPAATIAARQSATGKTDTPIQKVPQISVVTAEEMALHQPKSVKEALSYTP) (0)
 Residues: 105 - 110 (YLNGLK) (0)
 Residues: 121 - 133 (IDPYMLERAEIMR) (0)
 Residue : 136 (V) (0)
 Residues: 138 - 140 (VLY) (0)
 Residues: 148 - 157 (LLNMVSKRPT) (0)
 Residues: 161 - 169 (LKEVQFKAG) (1)
 Residue : 170 (T) (0)
 Residue : 173 (L) (0)
 Residues: 174 - 181 (FQTGFDFS) (-1)
 Residues: 191 - 200 (SYRLTGLARS) (1)
 Residues: 212 - 221 (RYAIAPAFW) (-1)
 Residues: 227 - 236 (TNFTFLSYFQ) (1)
 Residues: 280 - 289 (KMGYSFDHE) (-1)
 Residue : 293 (T) (0)
 Residues: 294 - 303 (FTVRQNLRF) (1)
 Residues: 356 - 366 (SVDTQLQSKFA) (-1)
 Residue : 367 (T) (0)
 Residue : 370 (I) (0)
 Residues: 371 - 380 (DHTLLTGVD) (1)
 Residues: 443 - 450 (VYVQDQAQ) (-1)
 Residues: 451 - 453 (WDK) (0)
 Residues: 454 - 461 (VLVTLGGR) (1)
 Residues: 487 - 493 (WRGGVNY) (-1)
 Residues: 494 - 495 (LF) (0)
 Residues: 497 - 498 (NG) (0)
 Residues: 499 - 506 (VTPYFYSY) (1)
 Residues: 531 - 536 (EVGVKY) (-1)
 Residues: 537 - 538 (VP) (0)
 Residue : 542 (P) (0)
 Residues: 543 - 549 (IVVTGAV) (1)
 Residues: 579 - 586 (VEIEAKAA) (-1)
 Residue : 587 (L) (0)
 Residue : 590 (S) (0)
 Residues: 591 - 597 (VNVVGSY) (1)
 Residues: 623 - 630 (ASLWADYT) (-1)
 Residues: 639 - 646 (LTLGTGGR) (1)
 Residues: 667 - 673 (VDALVRY) (-1)
 Residues: 674 - 675 (DL) (0)
 Residue : 683 (S) (0)
 Residues: 684 - 690 (NVALHVN) (1)
 Residue : 692 (L) (0)
 Residues: 714 - 722 (QVVATATFR) (-1)

2MPR

Residues: 3 - 13 (FHGYARSGIGW) (1)
 Residue : 20 (Q) (0)
 Residue : 22 (C) (0)
 Residues: 29 - 39 (QSKYRLGNECE) (0)
 Residues: 40 - 47 (TYAELKLG) (-1)
 Residues: 59 - 68 (YFDTNVAYSV) (1)
 Residues: 69 - 79 (NQQNDWESTDP) (0)
 Residues: 80 - 88 (AFREANVQG) (-1)
 Residues: 99 - 105 (TIWAGKR) (1)
 Residues: 106 - 123 (FYQRHDVHMIDFYWDIS) (0)
 Residues: 124 - 132 (GPGAGIENI) (-1)
 Residues: 138 - 146 (KLSLAATRS) (1)
 Residues: 170 - 177 (DVFDVRLA) (-1)
 Residue : 179 (L) (-1)

Residues: 186 - 194 (VLELGVDYD) (1)
 Residues: 213 - 221 (WMFTAHTQ) (-1)
 Residues: 227 - 235 (YNKFVVQYA) (1)
 Residues: 243 - 246 (GKGQ) (0)
 Residues: 280 - 288 (WRILDHGAI) (-1)
 Residues: 295 - 303 (DLMYVGMYY) (1)
 Residues: 316 - 324 (WTVGVRPMY) (-1)
 Residues: 330 - 339 (MSTLLEVGVD) (1)
 Residues: 353 - 362 (YKITLAQQWQ) (-1)
 Residues: 367 - 371 (IWSRP) (0)
 Residues: 372 - 380 (AIRIFATYA) (1)
 Residues: 417 - 426 (WTFGAQMEIW) (-1)
 Residue : 427 (W) (0)

2OMF

Residue : 1 (A) (0)
 Residues: 2 - 4 (EII) (-1)
 Residues: 9 - 19 (NKVDLYGKAVG) (1)
 Residues: 40 - 49 (YARLGFKGET) (-1)
 Residues: 56 - 65 (TGYGQWEYNF) (1)
 Residues: 81 - 90 (TRLAFAGLKY) (-1)
 Residues: 92 - 93 (DV) (0)
 Residues: 94 - 102 (GSFDYGRNY) (1)
 Residues: 103 - 120 (GVVYDALGYTDMLPEFGG) (0)
 Residue : 123 (A) (0)
 Residues: 129 - 131 (FVG) (0)
 Residues: 132 - 141 (RVGGVATYRN) (-1)
 Residues: 142 - 145 (SNFF) (0)
 Residues: 148 - 150 (VDG) (0)
 Residues: 151 - 158 (LNFAVQYL) (1)
 Residue : 159 (G) (0)
 Residues: 173 - 182 (GVGGSISYEY) (-1)
 Residues: 183 - 184 (EG) (0)
 Residues: 185 - 192 (FGIVGAYG) (1)
 Residues: 214 - 222 (WATGLKYDA) (-1)
 Residues: 223 - 224 (NN) (0)
 Residues: 225 - 232 (IYLAANYG) (1)
 Residues: 258 - 263 (LLVAQY) (-1)
 Residues: 264 - 268 (QDFDG) (0)
 Residues: 269 - 274 (LRPSIA) (1)
 Residues: 297 - 302 (VGATYY) (-1)
 Residues: 303 - 306 (FNKN) (0)
 Residues: 307 - 312 (MSTYVD) (1)
 Residues: 334 - 339 (VGIVYQ) (-1)
 Residue : 340 (F) (0)

2POR

Residues: 2 - 12 (VKLSGDARMGV) (1)
 Residues: 22 - 32 (SSRSRVLFMTS) (-1)
 Residues: 41 - 46 (FGASF) (1)
 Residues: 47 - 59 (AHESVGAETGEDG) (0)
 Residues: 60 - 64 (TVFLS) (-1)
 Residue : 67 (F) (0)
 Residues: 68 - 73 (GKIEMG) (1)
 Residues: 74 - 94 (DALGASEALFGDLYEVGYTDL) (0)
 Residues: 99 - 104 (GNDIPY) (0)
 Residues: 106 - 108 (TGD) (0)
 Residue : 117 (P) (0)
 Residues: 118 - 124 (VLLYTYS) (-1)
 Residues: 128 - 134 (FSVAASM) (1)
 Residues: 135 - 136 (SD) (0)
 Residues: 150 - 157 (MAVAAAYT) (-1)
 Residues: 161 - 168 (YTVGLGYE) (1)
 Residues: 184 - 192 (LELAAIAKF) (-1)
 Residue : 194 (A) (0)
 Residues: 195 - 202 (TNVKAYYA) (1)
 Residues: 232 - 240 (YGLSVDSTF) (-1)
 Residue : 242 (A) (0)
 Residues: 243 - 250 (TTVGGYVQ) (1)
 Residues: 263 - 271 (YGLGASYDL) (-1)
 Residue : 274 (G) (0)
 Residues: 275 - 281 (ASIVGGI) (1)
 Residues: 294 - 300 (DLGVKFK) (-1)
 Residue : 301 (F) (0)