



Πανεπιστήμιο Κρήτης
Σχολή Θετικών και Τεχνολογικών Επιστημών
Τμήμα Χημείας

Γενικό Μεταπτυχιακό Πρόγραμμα

Μεταπτυχιακό Δίπλωμα Ειδίκευσης

*Διαχωρισμός και χαρακτηρισμός του πρωτεόματος
του πράσινου θειούχου φωτοσυνθετικού βακτηρίου*

Chlorobium tepidum

Καλλιόπη Κουγιανού

Επιβλέπων Καθηγητής: Γεώργιος Τσιώτης

Ηράκλειο, 2009

University of Crete
Department of Chemistry
Biochemistry Division

Master in Science

*Prefractionation and characterization of the
proteome of the photosynthetic green sulphur
bacterium Chlorobium tepidum*

Kalliopi Kouyianou

Supervising Professor: Georgios Tsiotis

Heraklion, 2009

Μέλη τριμελούς εξεταστικής επιτροπής:

Γεώργιος Τσιώτης, Αναπληρωτής καθηγητής Τμήματος
Χημείας Πανεπιστημίου Κρήτης

Δημήτριος Γανωτάκης, Καθηγητής Τμήματος Χημείας
Πανεπιστημίου Κρήτης

Σπύρος Περγαντής, Αναπληρωτής καθηγητής Τμήματος
Χημείας Πανεπιστημίου Κρήτης

Ευχαριστίες

Η παρούσα μεταπτυχιακή εργασία εκπονήθηκε στον τομέα Βιοχημείας του Τμήματος Χημείας του Πανεπιστημίου Κρήτης υπό την επιστημονική επιμέλεια του αναπληρωτή καθηγητή ***Γεωργίου Τσιώτη***, τον οποίο θα ήθελα να ευχαριστήσω θερμά για τη βοήθεια, την ενθάρρυνση, την κατανόηση και τη συμπαράσταση του κατά τη διάρκεια πραγματοποίησης της μεταπτυχιακής μου εργασίας στο εργαστήριό του. Οι πάντα εποικοδομητικές συζητήσεις για τα προβλήματα που παρουσιάζονταν κατά τη διάρκεια των πειραμάτων, αποτέλεσαν το σημαντικότερο βοήθημα για τη πραγμάτωση της εργασίας μου.

Ευχαριστώ επίσης τους καθηγητές ***Δημήτριο Γανωτάκη*** και ***Σπύρο Περγαντή*** του τμήματος Χημείας του Πανεπιστημίου Κρήτης που δέχτηκαν να συμμετάσχουν στην τριμελή εξεταστική επιτροπή μου.

Ακόμη, ευχαριστώ πολύ όλα τα άτομα που κατά καιρούς εργάστηκαν στο εργαστήριο και ιδιαίτερα τον Μιχάλη Αϊβαλιώτη για τις γνώσεις και τη βοήθεια που μου προσέφερε το αρχικό διάστημα που βρέθηκα στο εργαστήριο. Επίσης ευχαριστώ πολύ για την ηθική συμπαράσταση, τη βοήθεια και τη φιλία τους τα χρόνια αυτά την Αναστασία Παπαδιώτη, την Αντιγόνη Νικολάκη και την Κατερίνα Αρβανίτη.

Επίσης θα ήθελα να ευχαριστήσω όλα τα άτομα του εργαστηρίου του καθηγητή. κ. Γανωτάκη του τμήματος Χημείας και συγκεκριμένα το Λευτέρη Τουλουπάκη, την Γιάννα Παναγάκου, την Ειρήνη Κωστομοίρη, την Αριάδνη Παπαδάκη και τον Δημήτρη Στεφανάκη για την ενθάρρυνση, την συνεργασία τους και την βοήθεια τους σε οποιοδήποτε πρόβλημα παρουσιαζόταν.

Ευχαριστώ τον καθηγητή ***Απόστολο Ρίζο*** του τμήματος Χημείας του Πανεπιστημίου Κρήτης, για τα πειράματα DLS των χλωροσωμάτων που πραγματοποίησε με την ερευνητική του ομάδα.

Θα ήθελα να ευχαριστήσω πολύ τον καθηγητή **Kris Gevaert** του τμήματος Βιοχημείας στο πανεπιστήμιο του Ghent, στο Βέλγιο για την N-τερματική COFRADIC ανάλυση και την nLC-MS/MS ανάλυση των πρωτεϊνικών κλασμάτων, και τον **P.J. De Brock** που επιμελήθηκε των αναλύσεων αυτών.

Ακόμη, ευχαριστώ τους ερευνητές του εργαστηρίου ηλεκτρονικής μικροσκοπίας Βασίλης Γαλανόπουλος του τμήματος Βιολογίας του Πανεπιστημίου Κρήτης για τις μελέτες με ηλεκτρονικό μικροσκόπιο.

Επίσης θα ήθελα να ευχαριστήσω πολύ όλους τους φίλους μου και τον Νίκο Βελόγλου για την ισορροπία και τη συμπαράσταση που μου προσέφεραν στο διάστημα εκπόνησης της μεταπτυχιακής μου εργασίας.

Κυρίως ωστόσο θα ήθελα να ευχαριστήσω τους γονείς μου **Σωκράτη** και **Άννα** και την αδερφή μου **Τίνα** για την αγάπη τους, για τη συνεχή τους ενθάρρυνση και την οικονομική και ηθική τους στήριξη σε όλα αυτά τα χρόνια των σπουδών μου.

Πίνακας Περιεχομένων

Ευχαριστίες	4
Περίληψη	8
Συνοπτομογραφίες.....	14
1. ΕΙΣΑΓΩΓΗ	16
1.1 Πράσινα θειοβακτήρια	16
1.2 Chlorobium tepidum	18
1.1.3 Φωτοσύνθεση.....	22
1.1.4 Μεταβολισμός θείου	27
1.2 Πρωτεομική.....	30
1.2.1 Gel-based και gel-free πρωτεομική	33
1.2.2 Το COFRADIC ως εργαλείο διαχωρισμού πεπτιδίων [37-40]	36
1.3 Σκοπός Μεταπτυχιακής Εργασίας	42
2. ΥΛΙΚΑ – ΜΕΘΟΔΟΙ.....	44
2.1.1 Οργανισμός.....	44
2.1.2 Ανάπτυξη και απομόνωση κυττάρων του Chlorobium tepidum	44
2.1.2 Διάρρηξη κυττάρων: απομόνωση περιπλασματικού - κυτταροπλασματικού πρωτεόματος	48
2.1.3 Διακριτή απομόνωση υδατοδιαλυτού και μεμβρανικού τμήματος του Chl. tepidum	49
2.1.4 Απομόνωση χλωροσωμικού κλάσματος	50
2.2 Αναλυτικές μέθοδοι.....	52
2.2.1 Φασματοσκοπία απορρόφησης ορατής ακτινοβολίας	52
2.2.2 Στατική και δυναμική σκέδαση φωτός	52
2.2.3 Ηλεκτρονική μικροσκοπία	53
2.3 Μέθοδοι Ανάλυσης και Χαρακτηρισμού Πρωτεϊνών	54
2.3.1 Tricine-SDS ηλεκτροφόρηση σε πηκτή πολυακρυλαμιδίου [47].....	54
2.3.2 Μέθοδος Bradford [49].....	58
2.3.3 Συμπύκνωση – καταβύθιση πρωτεϊνών	60

2.3.5 N-τερματική ανάλυση COFRADIC [38-40].....	60
2.3.4 Φασματομετρία μάζας και ταυτοποίηση πρωτεϊνών χλωροσωμικού κλάσματος	63
2.4 Επεξεργασία των αποτελεσμάτων της πρωτεομικής ανάλυσης με μεθόδους βιοπληροφορικής	66
3. ΑΠΟΤΕΛΕΣΜΑΤΑ	67
3.1 Αναλυτικές μέθοδοι.....	67
3.1.1 Φασματοσκοπία απορρόφησης ορατής ακτινοβολίας	67
3.1.2 Δυναμική και Στατική σκέδαση φωτός[45-46]	70
3.1.3 Ηλεκτρονική μικροσκοπία	72
3.2. Μέθοδοι Ανάλυσης και Χαρακτηρισμού Πρωτεϊνών	75
3.2.1 Tricine-SDS ηλεκτροφόρηση σε πηκτή πολυακρυλαμιδίου	75
3.2.2 N-τερματική COFRADIC ανάλυση υποπρωτεομάτων του <i>Chl. terpidum</i> [38-40].....	79
3.2.2.1 N-τερματική COFRADIC ανάλυση υδατοδιαλυτού κλάσματος	80
3.2.2.2 N-τερματική COFRADIC ανάλυση περιπλασματικού κλάσματος	86
3.2.2.3 N-τερματική COFRADIC ανάλυση κυτταροπλασματικού κλάσματος.....	92
3.2.2.4 N-τερματική COFRADIC ανάλυση ολικού μεμβρανικού κλάσματος	98
3.2.3 Φασματομετρία μάζας και ταυτοποίηση πρωτεϊνών χλωροσωμικού κλάσματος	104
4. ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ.....	109
4.1 Ανάλυση και χαρακτηρισμός χλωροσωμικού κλάσματος.....	109
4.2 N-τερματική COFRADIC ανάλυση	112
4.3 Σύνολο ταυτοποιημένων πρωτεϊνών <i>Chl. Terpidum</i>	118
5. ΣΥΜΠΕΡΑΣΜΑΤΑ ΜΕΛΛΟΝΤΙΚΟΙ ΣΤΟΧΟΙ	131
Βιβλιογραφία	134

Περίληψη

Στην παρούσα ερευνητική εργασία συνδυάζοντας πρωτεομικές μεθόδους κλασματοποίησης με gel-free μεθόδους πρωτεομικής ανάλυσης (N-terminal COFRADIC και nLC-MS/MS) μελετήθηκε κυρίως η έκφραση αλλά και η τοπολογία των πρωτεϊνών του *Chl. tepidum*. Επιπλέον εξετάστηκαν τα απομονωμένα χλωροσώματα του *Chl. tepidum*, ως προς το μέγεθος τους με χρήση DLS και ηλεκτρονική μικροσκοπία και ως προς την πρωτεϊνική τους σύσταση χρησιμοποιώντας nLC-MS/MS.

Το *Chlorobium tepidum* είναι ένα θερμοφιλο Gram-αρνητικό, πράσινο θειοβακτήριο (*Chlorobia*), είναι αποκλειστικά αναερόβιο φωτολιθοαυτότροφο και είναι διαδεδομένο σε υδατικά περιβάλλοντα, όπου ανοξικά στρώματα που περιέχουν ανηγμένες ενώσεις θείου εκτίθενται σε φως. Τα πράσινα θειοβακτήρια χρησιμοποιούν μοναδικά συστήματα για την συλλογή του φωτός, τα χλωροσώματα, στα οποία οι χρωστικές οργανώνονται σε σωληνοειδείς δομές, χωρίς να αλληλεπιδρούν με πρωτεΐνες.

Το βακτήριο αυτό αποτελεί οργανισμό μοντέλο για τη μελέτη των διεργασιών του μεταβολισμού θείουχων ενώσεων και της φωτοσύνθεσης. Προσφάτως, το γονιδίωμα του *Chl. tepidum* αναλύθηκε πλήρως και προβλέφθηκε ότι κωδικοποιεί 2288 γονίδια, αλλά η λειτουργία ενός μεγάλου αριθμού αυτών παραμένει άγνωστη. Πολλά γονίδια βρέθηκε ότι είναι σε μεγάλο βαθμό συντηρημένα μεταξύ των φωτοσυνθετικών οργανισμών με μη ξεκάθαρο ρόλο στο *Chl. tepidum*, αλλά αυτά τα γονίδια πιθανόν να έχουν συγκεκριμένους ρόλους στην φωτοσύνθεση και στην φωτοβιολογία. Επίσης, αν και μελέτες βασισμένες σε SDS-PAGE, ανοσοκαθήλωση και μοριακή βιολογία σε απομονωμένα χλωροσώματα του *Chl. tepidum* έχουν διαπιστώσει ότι τα χλωροσώματα περιέχουν 10 χλωροσωμικές πρωτεΐνες, δεν έχει διευκρινιστεί ακόμα η λειτουργία τους καθώς και η πρωτεϊνική σύνθεση των άθικτων χλωροσωμάτων.

Η ανάλυση και ο χαρακτηρισμός των γονιδιακών προϊόντων ενός κυττάρου, ιστού ή οργανισμού ορίζεται ως πρωτεομική ανάλυση (*proteomics*). Η πρωτεομική ανάλυση παρουσιάζει μεγάλο ερευνητικό ενδιαφέρον, παρά τους περιορισμούς που παρουσιάζει αναφορικά με την πολυπλοκότητα των δειγμάτων τα οποία καλείται να αναλύσει, των διαφορετικών φυσικοχημικών τους ιδιοτήτων και το μεγάλο εύρος των συγκεντρώσεων τους, καθότι οι πρωτεΐνες είναι υπεύθυνες για τον φαινότυπο των κυττάρων και οι λειτουργίες που επιτελούνται σε αυτά δεν είναι δυνατόν να διερευνηθούν μελετώντας αποκλειστικά το γονιδίωμα των κυττάρων.

Στην μελέτη αυτή, πραγματοποιήθηκε κλασματοποίηση του πρωτεόματος του *Chl. tepidum* και τα υποπρωτεόματα του κυτταροπλάσματος, του περιπλάσματος, των ολικών μεμβρανών καθώς και το κλάσμα του υδατοδιαλυτού τμήματος του βακτηρίου χαρακτηρίστηκαν με φασματοσκοπία απορρόφησης και Tricine SDS-PAGE και αναλύθηκαν με N-terminal COFRADIC, μια αναλυτική τεχνική που στοχεύει μέσω χημικών τροποποιήσεων και χρωματογραφικών μεθόδων στην επιλογή των N-τερματικών πεπτιδίων που προκύπτουν από τη θρυψινοποίηση του πρωτεϊνικού δείγματος πριν την ανάλυση τους με LC-MS/MS.

Η N-terminal COFRADIC ανάλυση οδήγησε στην ταυτοποίηση 805 πρωτεϊνών εκ των οποίων με βάση την προβλεπόμενη [51-52] υποκυτταρική τοπολογία τους το μεγαλύτερο μέρος είναι κυτταροπλασματικές πρωτεΐνες, ενώ 277 από τις ταυτοποιημένες πρωτεΐνες εμφανίστηκαν ως άγνωστης κυτταρικής τοπολογίας. Αναφορικά με την προβλεπόμενη λειτουργία τους [3], η πλειοψηφία των ταυτοποιημένων πρωτεϊνών δεν εμφανίζει γνωστή λειτουργία, ενώ το μεγαλύτερο μέρος των ταυτοποιημένων πρωτεϊνών γνωστής λειτουργίας εμπλέκεται στο μεταβολισμό της ενέργειας. Η ταυτοποίηση πολλών πρωτεϊνών σε επιμέρους πρωτεϊνικά κλάσματα και όχι στο σύνολο τους, ενισχύει την σημασία της κλασματοποίησης πολύπλοκων πρωτεομάτων καθώς και την σημασία της επιλογής μιας αντιπροσωπευτικής τάξης των προκύπτοντων πεπτιδίων όπως είναι τα N-τερματικά πεπτίδια προκειμένου να ταυτοποιηθεί όσο το δυνατόν μεγαλύτερο μέρος των πρωτεϊνών των πολύπλοκων αυτών δειγμάτων.

Επιπλέον, απομονώθηκε το χλωροσωμικό κλάσμα του *Chl. Trepidum* και χαρακτηρίστηκαν με φασματοσκοπία απορρόφησης, Tricine SDS-PAGE,

ηλεκτρονική μικροσκοπία και δυναμική σκέδαση φωτός (DLS). Στην Tricine SDS-PAGE εμφανίστηκαν πρωτεϊνικές ζώνες που αντιστοιχούν σε περισσότερες από 20 πρωτεΐνες με μοριακά βάρη μεταξύ 6 kDa και 70 kDa. Με βάση τα αποτελέσματα της ηλεκτρονικής μικροσκοπίας και της DLS, τα χλωροσώματα εμφανίζονται να έχουν μήκος 170 nm και 140 nm, αντίστοιχα. Επιπρόσθετα, η πρωτεϊνική σύσταση του απομονωμένου χλωροσωμικού κλάσματος μελετήθηκε με nLC-MS/MS, οδηγώντας στην ταυτοποίηση 270 πρωτεϊνών, στις οποίες συμπεριλαμβάνονται και οι 10 προβλεπόμενες από το γονιδίωμα ως χλωροσωμικές πρωτεΐνες, καθώς και κάποιες επιπλέον πρωτεΐνες που προτείνεται [8, 12, 14] ότι αλληλεπιδρούν με τα χλωροσώματα.

Συνολικά, η πρωτεομική ανάλυση που πραγματοποιήθηκε στο *Chl. tepidum* οδήγησε στην ταυτοποίηση 891 πρωτεϊνών, δηλαδή του 38.94% επί του θεωρητικού πρωτεόματος του *Chl. tepidum* [3], που αποτελεί την ταυτοποίηση του μεγαλύτερου ποσοστού πρωτεϊνών επί του θεωρητικού πρωτεόματος του βακτηρίου αυτού που έχει αναφερθεί στη βιβλιογραφία.

Abstract

In the present study, the expression and the topology of the proteins of *Chl. tepidum* were examined, by combining proteomics fractionation techniques with gel-free proteomics analysis (*N*-terminal COFRADIC and nLC-MS/MS). Additionally, isolated chlorosomes from *Chl. tepidum* were studied, regarding their size, using DLS and EM, and their protein composition, using nLC-MS/MS.

Chl. tepidum is a thermophile, Gram-negative green-sulfur bacterium. It is an obligatory anaerobic photolithoautotroph, found in aquatic environments where the highly abundant in reduced sulfur compounds anoxic layers are exposed to light. Green sulfur bacteria use a unique apparatus for light harvesting, the chlorosome, in which pigments are organised in aggregates, rather than associated with proteins.

Chl. tepidum is a model organism for the study of inorganic sulfur metabolism and photosynthesis. The genome of *chl. Tepidum* has recently been sequenced and it was found to comprise of 2288 assigned coding sequences. The function of a large number of encoded proteins remains unknown, however many of these proteins appear to be highly conserved among photosynthetic organisms. Therefore, it can be inferred that these proteins may be functionally involved in photosynthesis. In addition, although studies based on SDS-PAGE and immunodetection have established that chlorosomes contain 10 gene products, their function, as well as the protein composition of intact chlorosomes is yet to be determined.

The analysis and characterization of the gene products of a cell, tissue or organism is termed proteomics. Proteomic analysis is a very interesting research field, despite its limitations regarding the increased sample complexity, their diverse physical and chemical properties and the dynamic range of the protein concentrations in any given proteome. The reason why this field attracts so much academic interest is because it is

proteins which are responsible for the phenotype of the cells and because the cells' functions cannot be fully understood only by genomic studies.

In this study, the proteome of *Chl. tepidum* was fractionated, and the resulting cytoplasmic, periplasmic, total membrane and water-soluble sub-proteomes were characterized using absorption spectroscopy and Tricine SDS-PAGE, and they were subsequently analysed using *N*-terminal COFRADIC. *N*-terminal COFRADIC is an analytical technique, aiming through chemical modifications and chromatography techniques, to select the *N*-terminal peptides that emerge after trypsin digestion of the protein sample. After the selection of this class of peptides, the less complex peptide mixture is subjected to nLC-MS/MS analysis.

The *N*-terminal COFRADIC analysis of the fractionated proteome of *Chl. tepidum*, led to the identification of 805 proteins, most of which, based on their predicted subcellular topology [51-52] are cytoplasmic proteins, whereas 277 were of unknown topology. With regards to their annotated function [3], the majority of the identified proteins were of unknown function, whereas most of the remaining proteins are involved in energy metabolism. The identification of a large number of proteins in some protein fractions but not in all of them, stresses the importance of prefractionation techniques in proteomic analysis, as well as the importance of selecting a representative class of peptides, such as *N*-terminal peptides, in order to identify the largest possible number of proteins in such complex protein samples.

Furthermore, the chlorosome fraction of *Chl. tepidum* was isolated and characterized using absorption spectroscopy, Tricine SDSPAGE, electron microscopy and dynamic light scattering (DLS). Tricine SDS-PAGE showed the presence of at least twenty proteins with molecular weights ranging between 6 kDa and 70 kDa. Light scattering and electron microscopy revealed particles of 140 nm and 170 nm in length respectively. Additionally, the protein composition of the intact chlorosome was obtained by MS/MS-driven proteomics, leading to the identification of 270 proteins, including all 10 chlorosomal proteins, along with a number of proteins which are suggested to interact with the chlorosome [8, 12, 14].

All in all, the proteomic analysis of *Chlorobium tepidum* led to the identification of 891 proteins, accounting for 38.9% of the predicted proteome of *Chl. tepidum* [3]. Therefore, this proteomics study, has led to the identification of the largest number of proteins of *Chlorobium tepidum* to date, providing information about their state and topology.

Keywords: Chlorobium tepidum, photosynthesis, COFRADIC, light scattering, green sulfur bacteria, electron microscopy, mass spectrometry, Tricine-SDS PAGE

Συντομογραφίες

ATP : αδενόσινο-5-τριφωσφορικός εστέρας

BChl : Βακτηριοχλωροφύλλη

Chl : χλωροφύλλη

bp : ζεύγη βάσεων

CO₂ : διοξείδιο του άνθρακα

FMO : Fenna-Matthews-Olson πρωτεΐνη

EDTA : αιθυλενο-διαμινο-τετραοξικό οξύ

kDa : kilodalton

CM : κυτταροπλασματική μεμβράνη

OM : εξωτερική μεμβράνη

OMP : πρωτεΐνες εξωτερικής μεμβράνης

NADP : νικοτινάμιδο-αδενινο-δινουκλεοτίδιο φωσφορικού εστέρα

NADH : νικοτινάμιδο-αδενινο-δινουκλεοτίδιο

PAGE : ηλεκτροφόρηση σε πηκτή πολυακρυλαμιδίου

2-DE : δύο-διαστάσεων ηλεκτροφόρηση

IPG : πηκτή ακρυλαμιδίου με ακινητοποιημένους αμφολύτες

RC : κέντρο αντίδρασης

RTCA : αναγωγικός κύκλος τρικαρβοξυλικού οξέος

SDS : δωδεκυλοθειϊκό νάτριο

MS : Φασματοσκοπία μάζας

MALDI : ιονισμός με laser απόπτωση, υποβοηθούμενη με υπόστρωμα

TOF : ανιχνευτής χρόνου πτήσης

ESI : ιονισμός με ηλεκτροψεκασμό

LC : υγρή χρωματογραφία

Tris : τρις-υδροξυμεθυλ-αμινομεθάνιο

Tricine : N-τρις (υδροξυμέθυλο-) μέθυλο-γλυκίνη

DTT : διθειοθρεϊτόλη

TEMED : N, N, N', N'-τετραμεθυλενοδιαμίνη

TCA : τριγλωροοξικό οξύ

GRAVY : τιμή μέσης υδροφοβικότητας των πρωτεϊνών

COFRADIC : Combined fractional diagonal chromatography

DLS : Dynamic light scattering

1. ΕΙΣΑΓΩΓΗ

1.1 Πράσινα θειοβακτήρια

Το φύλο *Chlorobi* (που περικλείει τα πράσινα θειούχα βακτήρια) είναι μια μεταβολικά περιορισμένη, καλά καθορισμένη φυσιολογικά και γενετικά στενά συνδεδεμένη φυλογενετική ομάδα. Η ομάδα αυτή είναι ευρέως διαδεδομένη σε υδατικά περιβάλλοντα και αποτελείται από υποχρεωτικά αναερόβια, αποκλειστικά φωτότροφα βακτήρια [1].

Τα βακτήρια αυτά χωρίζονται σε δύο οικογένειες, τα *Chlorobiaceae* και τα *Chloroflexaceae*. Εκτός της φυλογενετικής διαφοροποίησης τα *Chlorobiaceae* και τα *Chloroflexaceae* διαφέρουν και ως προς άλλους παράγοντες. Τα *Chlorobiaceae* συνθέτουν βακτηριοχλωροφύλλες (BChls) και είναι υποχρεωτικά ή επιλεκτικά φωτότροφα, ενώ τα *Chloroflexaceae* δε συνθέτουν BChls και δεν είναι φωτότροφα [1].

Τα *Chlorobia* είναι υποχρεωτικά αναερόβια φωτοαυτότροφα τα οποία: α) οξειδώνουν θειούχες ενώσεις, H_2 ή Fe^{+2} , β) καθηλώνουν άνθρακα μέσω του αναγωγικού κύκλου τρικαρβοξυλικού οξέος, γ) συνθέτουν BChl c, d, ή e, καθώς και BChla και Chla και δ) διαθέτουν κέντρο αντίδρασης τύπου I, την Fenna-Matthews-Olson (FMO) BChl-a-binding πρωτεΐνη και χλωροσώματα [2].

Τα *Chlorobia* και τα Gram-θετικά *heliobacteria* είναι οι μοναδικοί οργανισμοί που διαθέτουν ομοδιμερή κέντρα αντίδρασης τύπου I, σε αντίθεση με τους λοιπούς φωτοσύνθετικούς οργανισμούς που διαθέτουν ετεροδιμερή κέντρα αντίδρασης. Σε αντίθεση με τα φυτά, τα πράσινα φύκη και τα κυανοβακτήρια, τα *Chlorobia* διεξάγουν αναερόβια φωτοσύνθεση. Επίσης, αντί να χρησιμοποιούν τον κύκλο του Calvin,

πραγματοποιούν καθήλωση CO₂ μέσω του αναγωγικού κύκλου του τρικαρβοξυλικού

οξέος (TCA), χρησιμοποιώντας ηλεκτρόνια προερχόμενα από H₂ ή από ανηγμένες θειούχες ενώσεις. Αυτός ο κύκλος, που ανακαλύφθηκε στα *Chlorobia*, συναντάται μόνο σε έναν περιορισμένο αριθμό οργανισμών. Λόγω των μοναδικών αυτών χαρακτηριστικών τους, οι μελέτες στελεχών των *Chlorobia* είναι σημαντικές για την κατανόηση της εξέλιξης και των μηχανισμών της φωτοσύνθεσης και του μεταβολισμού της ενέργειας [3].

Αν και τα πράσινα θειοβακτήρια περιγράφηκαν αρχικά πριν περισσότερο από έναν αιώνα, η έρευνα στα πράσινα θειοβακτήρια έχει καθυστερήσει σε σχέση με τους περισσότερους από τους άλλους φωτοσυνθετικούς οργανισμούς. Το γεγονός αυτό πιθανόν οφείλεται στη δύσκολη καλλιέργεια τους, στην υποχρεωτικά αναερόβια φύση τους, καθώς και στο ότι πολλές από τις πρωτεΐνες τους χάνουν την ενεργότητα τους μετά από έκθεση τους σε οξυγόνο καθιστώντας την απομόνωση και τη μελέτη τους δυνατά μόνο σε αναερόβιες συνθήκες [4].

Στα πλαίσια μίας προσπάθειας αύξησης της κατανόησης της βακτηριακής ποικιλότητας γενικότερα, και της γνώσης και του ενδιαφέροντος στα πράσινα θειοβακτήρια ειδικότερα, προσφάτως αποκωδικοποιήθηκε το γονιδίωμα του *Chlorobium tepidum* [3] (TIGR, Rockville, Maryland). Τα γονιδιώματα 10 στελεχών *Chlorobia* που έχουν αποκωδικοποιηθεί και των 2 που αναμένεται να αποκωδικοποιηθούν περιλαμβάνουν το μεγαλύτερο μέρος της έως τώρα γνωστής διαφοροποίησης αυτής της ομάδας. Με βάση τα γονιδιώματα αυτά, προκύπτει ότι τα στελέχη του *Chlorobi* κωδικοποιούν περίπου 1750-2800 γονίδια και μοιράζονται έναν κοινό πυρήνα από περίπου 1400-1500 γονίδια.

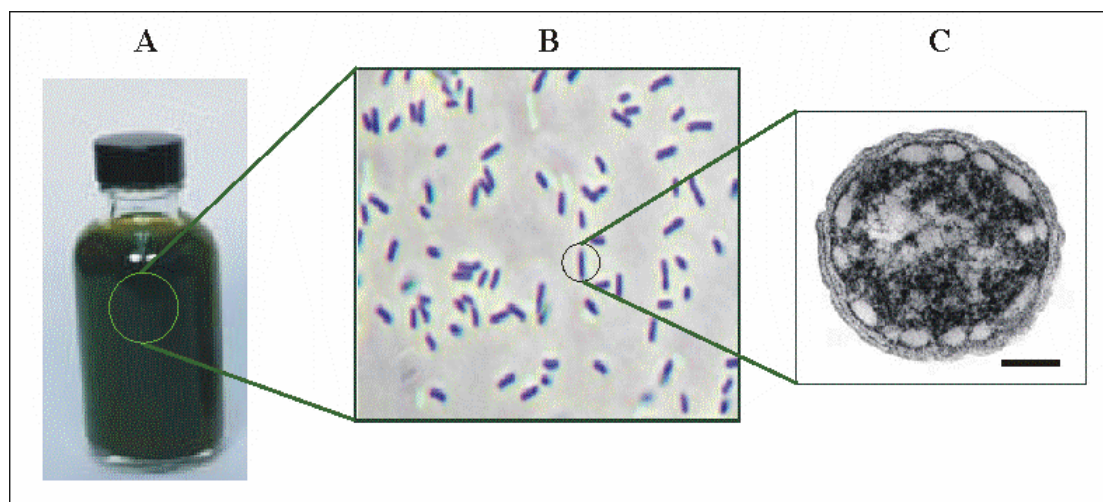
Τα γονιδιώματα των στελεχών των *Chlorobia* κωδικοποιούν μόνο μερικούς προβλεπόμενους μεταφορείς οργανικών μορίων, έχουν ένα μικρό αριθμό προβλεπόμενων ρυθμιστών μεταγραφής, ενώ παρουσιάζεται και έλλειψη σε μεγάλο βαθμό από ιστιδινοκινάσες και ρυθμιστές απόκρισης. Αυτές οι παρατηρήσεις οδηγούν στο συμπέρασμα ότι τα *Chlorobia* βρίσκονται σε ένα σχετικά σταθερό (και περιορισμένης ενέργειας) περιβάλλον και ότι πιθανόν διαθέτουν περιορισμένη ικανότητα να αντιδρούν σε αλλαγές στο

φυσικοχημικό τους περιβάλλον [5].

1.2 *Chlorobium tepidum*

Το *Chlorobium tepidum*, που πρόσφατα μετονομάστηκε σε *Chlorobaculum tepidum* [6], είναι ένα Gram-αρνητικό βακτήριο που ανήκει στην οικογένεια των πράσινων θειοβακτηρίων, *Chlorobia* (Εικόνα 1).

Απομονώθηκε αρχικά από θερμές, υψηλής περιεκτικότητας σε σουλφίδια, πηγές της Νέας Ζηλανδίας [7]. Το γένος *Chlorobium* περιλαμβάνει έξι είδη από τα οποία το *Chlorobium tepidum* είναι το μόνο θερμόφιλο, καθώς αναπτύσσεται στους 48°C. Το *Chl. tepidum* έχει αναδειχθεί ως ο οργανισμός μοντέλο των *Chlorobia*, κυρίως λόγω της εύκολης καλλιέργειας του, καθώς και επειδή ήταν το πρώτο βακτήριο της οικογένειας των πράσινων θειοβακτηρίων, του οποίου λύθηκε το γονιδίωμα [8].



Εικόνα 1: Εικόνα υγρής καλλιέργειας του *Chl. tepidum* (A), εικόνα κυττάρων *Chl. tepidum* από οπτικό μικροσκόπιο (B) και τομή του κυττάρου του *Chl. tepidum* από ηλεκτρονικό μικροσκόπιο (C). Η γραμμή στην εικόνα C αντιστοιχεί σε 100 nm.

Το πλήρες γονιδίωμα του *Chl. tepidum* αποκωδικοποιήθηκε από τον Eisen et al. (2002) [3] και βρέθηκε ότι είναι ένα απλό κυκλικό χρωμόσωμα 2,154,946 bp. Τα γενικά χαρακτηριστικά του γονιδιώματος του *Chl. tepidum* παρουσιάζονται στον Πίνακα 1. Σύγκριση με άλλα γονιδιώματα έδειξε την ύπαρξη γονιδίων τα οποία είναι συντηρημένα μεταξύ των φωτοσυνθετικών οργανισμών. Η λειτουργία πολλών από αυτά δεν είναι γνωστή και πιθανόν να παίζουν πρωτότυπους ρόλους στην φωτοσύνθεση ή στην φωτοβιολογία. Από φυλλογονιδιακή ανάλυση προέκυψαν πιθανοί διπλασιασμοί γονιδίων που σχετίζονται με βιοσυνθετικά μονοπάτια της φωτοσύνθεσης και του μεταβολισμού του θείου και του αζώτου. Επίσης, προέκυψαν σημαντικές ομοιότητες μεταξύ μεταβολικών διαδικασιών του *Chl. tepidum* και πολλών Αρχαίων. Φυλογενετικές αναλύσεις στα απαραίτητα (housekeeping) γονίδια του *Chl. tepidum*, έδειξαν μεγάλη ομοιότητα με τα Κυτοφάγα-Φλαβοβακτήρια-βακτηροΐδης (CFB), και συγκεκριμένα με το *Porphoryomonas gingivalis*. Επίσης, ένα μεγάλο ποσοστό (~12%) των πρωτεϊνών του *Chl. tepidum* μοιάζουν περισσότερο με πρωτεΐνες Αρχαίων.

Πλήρες μέγεθος του γονιδιώματος (bp)	2,154,946
Συνολικός αριθμός αλληλουχιών	36,670
Ποσοστό % G+C	56.5
Συνολικός αριθμός CDS	2,288
Μέσο μέγεθος CDS (bp)	837
Ποσοστό % που κωδικοποιείται	88.9
Αριθμός οπερονίων rRNA (16S-23S-5S)	2
Αριθμός γονιδίων tRNA	50
Αριθμός γονιδίων sRNA	2
Πρωτεΐνες με ομοιότητες με πρωτεΐνες γνωστής λειτουργίας και λειτουργικής κατηγορίας	1,217
Πρωτεΐνες με ομοιότητες με πρωτεΐνες γνωστής λειτουργίας αλλά άγνωστης λειτουργικής κατηγορίας	98
Συντηρητικές υποθετικές πρωτεΐνες	293
Υποθετικές πρωτεΐνες	680

Πίνακας 1: Γενικά χαρακτηριστικά του γονιδιώματος του *Chl. tepidum* [3].

Στην Εικόνα 2 συνοψίζονται όλες οι γνωστές και προβλεπόμενες διαδικασίες που πραγματοποιούνται στο *Chl. tepidum* καθώς και οι αντίστοιχες πρωτεΐνες που συμμετέχουν σε αυτές [3]. Οι προτεινόμενες διαδικασίες όπως και οι αντίστοιχες

πρωτεΐνες πρέπει να επιβεβαιωθούν πειραματικά για να γίνει εφικτός ο προσδιορισμός της λειτουργίας ενός μεγάλου αριθμού πρωτεϊνών που δεν έχουν χαρακτηριστεί ακόμη. Συγκριτικά, το μεγαλύτερο ερευνητικό ενδιαφέρον στη μελέτη του *Chl. tepidum* και των πράσινων θειούχων βακτηρίων γενικότερα, επικεντρώνεται στην μελέτη των διαδικασιών της φωτοσύνθεσης και του μεταβολισμού θειούχων ενώσεων.

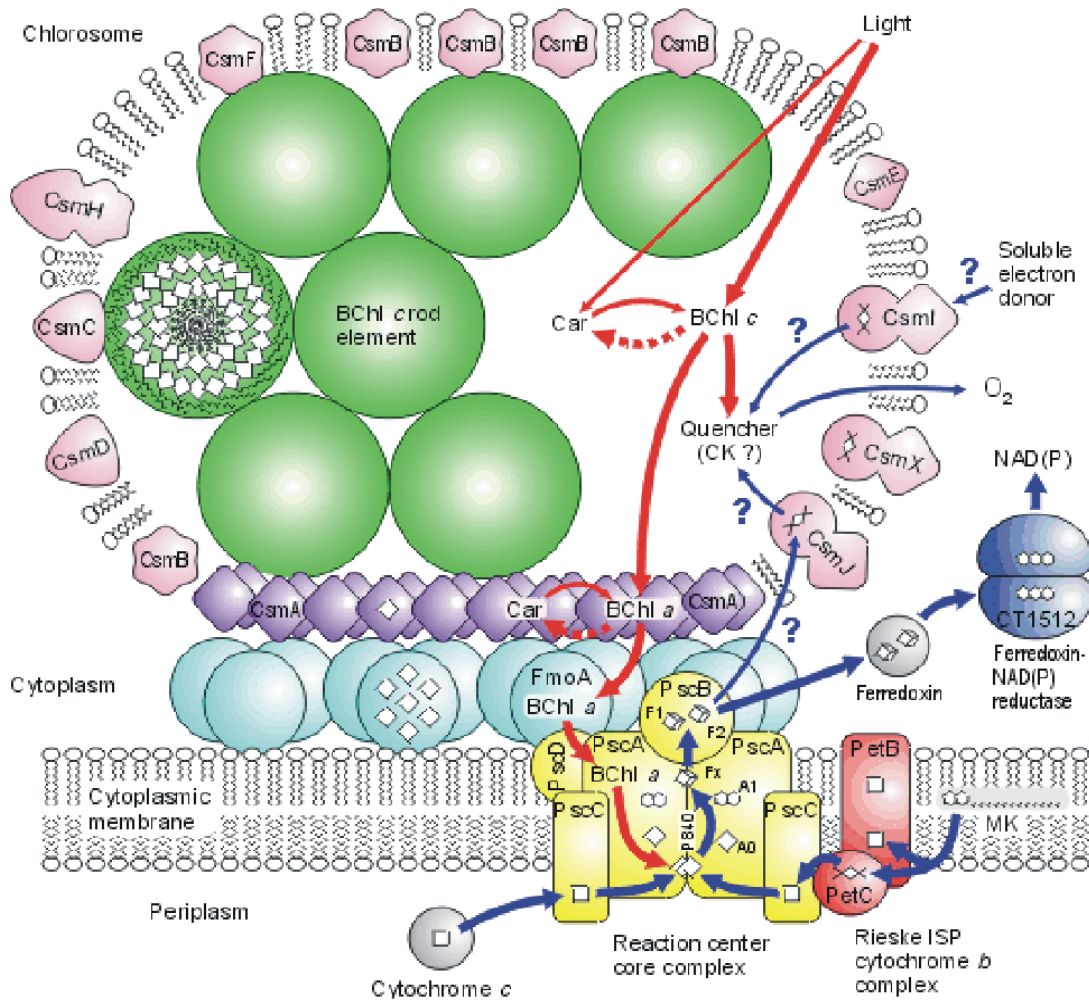
1.1.3 Φωτοσύνθεση

Για την συλλογή του φωτός, σε αντίθεση με όλους τους γνωστούς τύπους φωτοσυνθετικών κεραιών που είναι είτε μεμβρανικές είτε υδατοδιαλυτές πρωτεΐνες, το *Chlorobium tepidum* χρησιμοποιεί τα χλωροσώματα. Τα χλωροσώματα περιέχουν διαφορετικές BChls και καροτενοειδή που βρίσκονται διατεταγμένα μέσα σε λιπίδια δημιουργώντας κυστίδια προσδεμένα στην κυτταροπλασματική μεμβράνη (Εικόνα 3). Εξαιτίας, της μεγάλης αποτελεσματικότητας των χλωροσωμάτων στη συλλογής φωτός, το *Chl. tepidum* είναι ο πιο ενεργητικά αποτελεσματικός οργανισμός από όλους τους φωτότροφους οργανισμούς. Το γεγονός αυτό, επιτρέπει στο βακτήριο να επιβιώνει σε περιβάλλον όπου η ένταση του φωτός είναι μικρότερη από το 0,01% του τυπικού φωτός της ημέρας. Επιπλέον, ένας μεταβολισμός σε ένα χαμηλό οξειδοαναγωγικό δυναμικό μπορεί να πλεονεκτεί ενεργητικά: η χρήση της φερρεδοξίνης στην καθήλωση του CO₂, είναι ενεργητικά προτιμητέα από την χρήση NAD(P)H και αντίστροφη ροή ηλεκτρονίων για την αναγωγή του NAD(P)⁺ δεν είναι απαραίτητη όπως είναι στα μωβ φωτοσυνθετικά βακτήρια και σε άλλους οργανισμούς [9].

Τα χλωροσώματα αποτελούνται κυρίως από BChl τύπου *c* [10] που βρίσκονται διατεταγμένες σε σωληνοειδείς σωρούς εντός μια γαλακτολιπιδικής στοιβάδας που σταθεροποιείται από μεμβρανικές πρωτεΐνες. Είναι προσδεμένα στην εσωτερική πλευρά της κυτταροπλασματικής μεμβράνης στην οποία βρίσκονται τα κέντρα αντίδρασης (RCs). Τα χλωροσώματα έχουν την ικανότητα να δεσμεύουν αποτελεσματικά το φως με περίπου 200.000 μόρια BChl *c*, *d* ή *e* ανά χλωρόσωμα και αποτελούν την μεγαλύτερη, γνωστή, εξωτερική κεραία με μια φωτοσυνθετική μονάδα μερικών χιλιάδων μορίων χλωροφύλλης ανά RC (Εικόνα 3). Με 5000 χλωροφύλλες ανά RC θα πρέπει να υπάρχουν περίπου 40 RCs ανά χλωρόσωμα. Αν και τα χλωροσώματα δεν έχουν αυστηρά καθορισμένα μεγέθη, είναι περίπου 100-200 nm σε μήκος, 30-70 nm σε πλάτος και 30-40 nm σε ύψος.

Υψηλής καθαρότητας χλωροσώματα του βακτηρίου *Chl. tepidum* βρέθηκαν να περιέχουν 10 πολυπεπίδια. Πειράματα κυτταρικής τοπολογίας έχουν δείξει ότι όλες οι χλωροσωμικές πρωτεΐνες απομονώνονται σε μια σταθερή αναλογία με BChlc και ότι καμία από αυτές δεν σχετίζεται με άλλα υποκυτταρικά πρωτεϊνικά κλάσματα. Αν και διαφορετικά μοντέλα έχουν προταθεί για την οργάνωση των χλωροσωμικών πρωτεϊνών, δεν υπάρχουν ευρέως αποδεκτά δεδομένα για την πρωτεϊνική σύσταση των χλωροσωμάτων.

Είναι αποδεκτό ότι οι αλληλεπιδράσεις μεταξύ χρωστικών, και όχι μεταξύ χρωστικών και πρωτεϊνών, είναι αυτές που παίζουν τους κύριους ρόλους στον καθορισμό της δομής και των ιδιοτήτων των χλωροσωμάτων. Αρκετά μοντέλα έχουν προταθεί που περιγράφουν τα δομικά χαρακτηριστικά της χωροθέτησης των μορίων BChl στα χλωροσώματα. Πρέπει να σημειωθεί ωστόσο ότι οι ακριβείς τοπολογίες των πρωτεϊνών στον χλωροσωμικό φάκελο δεν είναι γνωστές, καθώς και ότι οι ρόλοι των πρωτεϊνών των χλωροσωμάτων είναι κατ' ουσία άγνωστοι.[11-14] Πιθανός ρόλος των χλωροσωμικών πρωτεϊνών θα μπορούσε να είναι η σταθεροποίηση της ασυνήθιστης μεμβρανικής δομής του χλωροσώματος στο κυτταρόπλασμα [10, 15].



Εικόνα 3: Το σύστημα συλλογής του φωτός από τα χλωροσώματα στα πράσινα θειοβακτήρια το οποίο συνδέεται μέσω της πρωτεΐνης FMO με το ομοδιμερές φωτοσυνθετικό κέντρο αντίδρασης τύπου I [4]

Εκτός από τα συσσωματώματα BChlc, τα χλωροσώματα περιέχουν BChla, καροτενοειδή, κινόνες, λιπίδια και πρωτεΐνες. Τα λιπίδια θεωρητικά σχηματίζουν τη μεμβρανική μονοστιβάδα και η σύζευξη με την κυτταροπλασματική μεμβράνη επιτυγχάνεται μέσω μιας βάσης από την CsmA πρωτεΐνη που προσδένει BChla [4]. Οι πρωτεΐνες δεν φαίνεται να αποτελούν στοιχεία σημαντικά για τη δομή και λειτουργία του χλωροσώματος και πιστεύεται ότι βρίσκονται στη χλωροσωμική βάση και στο χλωροσωμικό φάκελο [16]

(

Table 1: Interactions and Locations of Chlorosome Envelope Proteins of *C. tepidum*

protein	mass (kDa)	copies per chlorosome ^a	cross-linking properties ^b	location ^b
CsmA	6.2	~2700	dimers, trimers up to dodecamers; cross-linked to pre-CsmA, CsmB, CsmF, CsmD, FmoA	baseplate; binds BChl <i>a</i> and carotenoids
pre-CsmA	8.3	~200	cross-linked to CsmA, CsmB, CsmF	edge of baseplate
CsmB	7.5	~1000	cross-linked to CsmA, pre-CsmA, CsmC, CsmD, CsmE, CsmH, CsmI, CsmJ, CsmX	everywhere except the baseplate
CsmC	14.3	~300	dimers, trimers, and tetramers; cross-linked to CsmB	cytoplasmic surface
CsmD	11.1	~300	dimers and trimers; cross-linked to CsmA and CsmB	cytoplasmic surface
CsmE	7.5	~200	cross-linked to CsmB	no evidence
CsmF	7.7	~200	cross-linked to CsmA and pre-CsmA	edge of the baseplate
CsmH	21.8	~100	cross-linked to CsmB and probably CsmA	cytoplasmic surface
CsmI	25.9	~100	cross-linked to CsmJ and CsmB	cytoplasmic surface
CsmJ	23.9	~100	cross-linked to CsmI and CsmB	cytoplasmic surface
CsmX	24.0	~5	cross-linked to CsmB	cytoplasmic surface

^a Data taken from refs 2, 13, and 21. ^b This work.

Πίνακας 2).

Ανάμεσα τους, σε μεγαλύτερη συγκέντρωση βρίσκεται η χλωροσωμική πρωτεΐνη A (CsmA) η οποία προσδένει BChl*a* και καροτενοειδή και αποτελεί την επονομαζόμενη χλωροσωμική βάση, που μπορεί να παρατηρηθεί ως μία 2-D παρακρυσταλλική δομή με χρήση ηλεκτρονικής μικροσκοπίας [17].

Table 1: Interactions and Locations of Chlorosome Envelope Proteins of *C. tepidum*

protein	mass (kDa)	copies per chlorosome ^a	cross-linking properties ^b	location ^b
CsmA	6.2	~2700	dimers, trimers up to dodecamers; cross-linked to pre-CsmA, CsmB, CsmF, CsmD, FmoA	baseplate; binds BChl <i>a</i> and carotenoids
pre-CsmA	8.3	~200	cross-linked to CsmA, CsmB, CsmF	edge of baseplate
CsmB	7.5	~1000	cross-linked to CsmA, pre-CsmA, CsmC, CsmD, CsmE, CsmH, CsmI, CsmJ, CsmX	everywhere except the baseplate
CsmC	14.3	~300	dimers, trimers, and tetramers; cross-linked to CsmB	cytoplasmic surface
CsmD	11.1	~300	dimers and trimers; cross-linked to CsmA and CsmB	cytoplasmic surface
CsmE	7.5	~200	cross-linked to CsmB	no evidence
CsmF	7.7	~200	cross-linked to CsmA and pre-CsmA	edge of the baseplate
CsmH	21.8	~100	cross-linked to CsmB and probably CsmA	cytoplasmic surface
CsmI	25.9	~100	cross-linked to CsmJ and CsmB	cytoplasmic surface
CsmJ	23.9	~100	cross-linked to CsmI and CsmB	cytoplasmic surface
CsmX	24.0	~5	cross-linked to CsmB	cytoplasmic surface

^aData taken from refs 2, 13, and 21. ^bThis work.

Πίνακας 2: Αλληλεπιδράσεις και τοπολογία των χλωροσωμικών πρωτεϊνών του *Chl. tepidum* [12]

Μετά την συλλογή της ηλιακής ενέργειας από τα συσσωματώματα BChl_c, η ενέργεια διέγερσης μεταφέρεται στην παρακρυσταλλική βάση, και ακολούθως στην Fenna-Matthews-Olson (FMO) πρωτεΐνη, μια υδατοδιαλυτή πρωτεΐνη-αντένα που βρίσκεται ανάμεσα στο χλωρόσωμα και το κέντρο αντίδρασης στην κυτταροπλασματική μεμβράνη. Στη συνέχεια η ενέργεια διέγερσης μεταφέρεται στα κέντρα αντίδρασης (RCs). Τα RCs είναι ειδικά μεμβρανικά πρωτεϊνικά σύμπλοκα στην κυτταροπλασματική μεμβράνη που χρησιμοποιούν ηλιακή ενέργεια για την προώθηση αντιδράσεων μεταφοράς ηλεκτρονίων που οδηγούν στην παραγωγή της πρωτονιοκινητηρίου δύναμης για την σύνθεση ATP και/ή την παραγωγή αναγωγικής ισχύος.[18] Οι χλωροπλάστες των φυτών και τα κυανοβακτήρια χρησιμοποιούν τύπου I (δέκτες Fe-S, Φωτοσύστημα I) και τύπου II (δέκτες κινόνες, Φωτοσύστημα II) RCs για να οξειδώσουν νερό, παράγοντας οξυγόνο και πρωτόνια. Πολλά βακτήρια, συμπεριλαμβανομένων και των πράσινων μη θειούχων βακτηρίων και οι α, β και γ υποκατηγορίες των πρωτεοβακτηρίων, περιέχουν μόνο τύπου II RCs. Το RC των *Chlorobiaceae* και των Gram-θετικών ηλιοβακτηρίων (τύπου I) αποτελούνται από δυο όμοιες μεμβρανικές πρωτεΐνες (ομοδιμερή RCs) ενώ όλα τα άλλα

RCs είναι ετεροδιμερή.

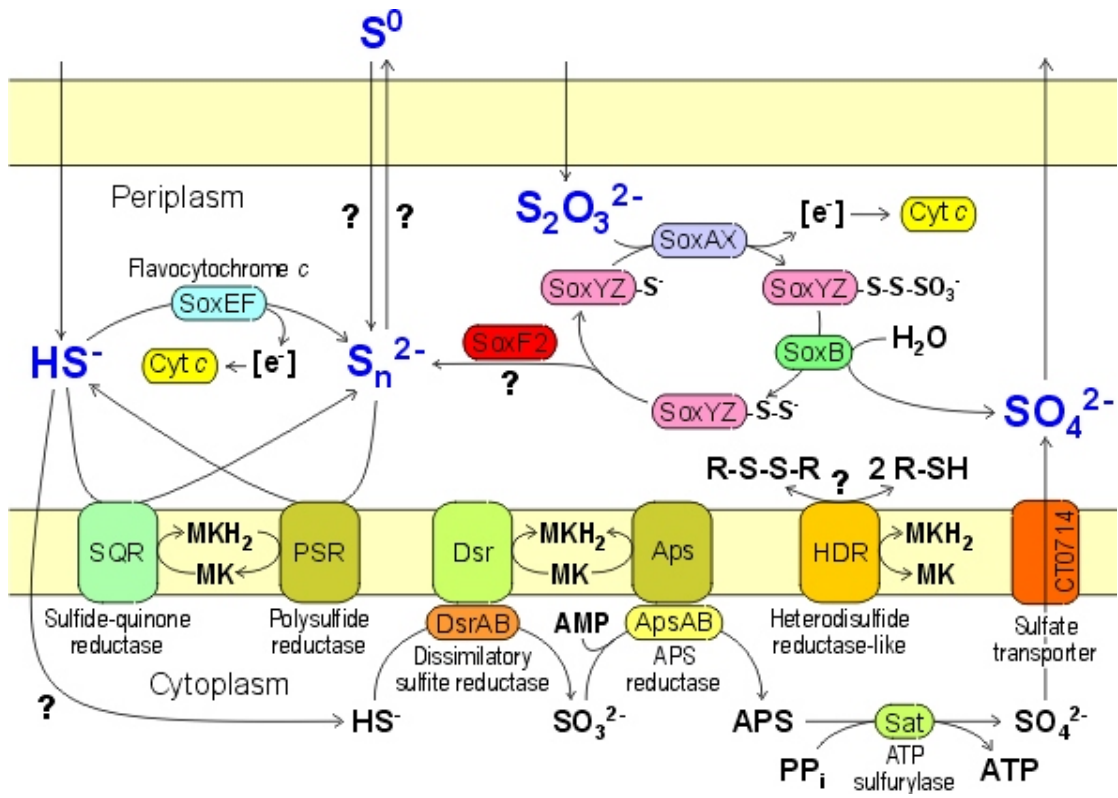
Η ενέργεια μεταφέρεται μέσω μιας μικρότερης δεξαμενής BChl *a*-795, στην βάση του χλωροσώματος, στην πρωτεΐνη FMO, με μέγιστο απορρόφησης στα 808 nm. Η FMO μεταφέρει την ενέργεια στον ομοδιμερή πυρήνα του RC, με ένα διμερές BChl*a* ως τον αρχικό δότη ηλεκτρονίων, P840, και Chl *a*-670 ως τον κύριο δέκτη ηλεκτρονίων A₀ (Εικόνα 3).

Ο πυρήνας του RC του *Chl. tepidum* αποτελείται από δύο αντίγραφα της μεγάλης διαμεμβρανικής πρωτεΐνης PscA και ένα αντίγραφο της περιφερειακής πρωτεΐνης PscB [18]. Η PscA προσδένει τον αρχικό δότη ηλεκτρονίων P840, τον αρχικό δέκτη ηλεκτρονίων A₀, 20 BChl*a* και το 4Fe-4S σύμπλεγμα Fx. Η PscB προσδένει τα δύο τελικά 4Fe-4S συμπλέγματα F_A και F_B, τα επονομαζόμενα κέντρα 1 και 2, και σχετίζεται με την βακτηριακή φερρεδοξίνη [19]. Το φωτοσυνθετικό κέντρο αντίδρασης του *Chl. tepidum* συμπληρώνουν δύο αντίγραφα της πρωτεΐνης PscC, ένα της πρωτεΐνης PscD και έξι της πρωτεΐνης FMO (Εικόνα 3) [18]. Η πρωτεΐνη PscC είναι ένα μονοαιμικό κυτόχρωμα c, το οποίο λειτουργεί ως ο βασικός δότης ηλεκτρονίων στο P840, ενώ ο ρόλος της πρωτεΐνης PscD δεν είναι γνωστός. Η πρωτεΐνη FMO σχηματίζει δύο τριμερή ανά RC και συνδέει το RC με το χλωρόσωμα.

Το κέντρο αντίδρασης όταν ανάγεται, παράγει χαμηλού δυναμικού αναγωγικά που ανάγουν την κυτταροπλασματική φερρεδοξίνη. Οι αρχικοί δότες ηλεκτρονίων στο κέντρο αντίδρασης είναι κυτοχρώματα c. Ο ενδιάμεσος δότης είναι ένα μεμβράνο-συνδετό κυτόχρωμα c (PscC), αλλά, έχουν απομονωθεί και υδατοδιαλυτά περιπλασμικά κυτοχρώματα c από τα πράσινα θειοβακτήρια [20-22]. Όταν οι ανόργανες ενώσεις θείου οξειδώνονται, τα ηλεκτρόνια μπορούν να εισαχθούν στην φωτοσυνθετική μεταφορά ηλεκτρονίων μέσω αυτών των κυτοχρωμάτων [23], που ανάγουν ξανά την πρωτεΐνη PscC.

1.1.4 Μεταβολισμός θείου

Οι ανηγμένες ανόργανες θειούχες ενώσεις αποτελούν τους πρωταρχικούς δότες ηλεκτρονίων για τη φωτοσυνθετική καθήλωση του άνθρακα στα πράσινα θειούχα βακτήρια [24]. Το *Chl. tepidum* χρησιμοποιεί υδρόθειο (H_2S), θειοθειικό ($S_2O_3^{2-}$), ή στοιχειακό θείο (S_0) ως δότες ηλεκτρονίων και συνεπώς αποτελεί ιδανικό μοντέλο για τη μελέτη του μεταβολισμού του ανόργανου θείου στα πράσινα θειούχα βακτήρια, αλλά και για τα αναερόβιων θειούχων φωτότροφων οργανισμών γενικότερα. Υπό φωτοτροφικές συνθήκες ανάπτυξης, τα περισσότερα *Chlorobiaceae* οξειδώνουν το υδρόθειο προς στοιχειακό θείο, το οποίο εξάγεται στον εξωκυττάριο χώρο [23]. Μετά την κατανάλωση του υδρόθειου, τα περισσότερα στελέχη οξειδώνουν το στοιχειακό θείο προς σχηματισμό θειϊκού. Επίσης, πολλά στελέχη των *Chlorobiaceae*, συμπεριλαμβανομένου και του *Chl. tepidum*, μπορούν να οξειδώσουν θειοθειικό προς θειϊκό [23]. Ωστόσο, ελάχιστα είναι γνωστά σχετικά με την δυνατότητα των *Chlorobiaceae* να ρυθμίζουν και να ενσωματώνουν τις αλληλοεξαρτώμενες διαδικασίες της οξείδωσης των δοτών ηλεκτρονίων, της συλλογής του φωτός και της καθήλωσης του CO_2 . Το πλήρως κωδικοποιημένο γονιδίωμα του *Chl. tepidum* έχει χρησιμοποιηθεί για να προταθούν μοντέλα για τα μονοπάτια οξείδωσης του θείου στο βακτήριο αυτό (Εικόνα 4) [3, 25].



Εικόνα 4: Προτεινόμενο μοντέλο για την οξείδωση θειούχων ενώσεων στο *Chlorobium tepidum*

Παρόλα αυτά, λίγες πληροφορίες είναι διαθέσιμες σχετικά με τις αρχικές οξειδωτικές αντιδράσεις, ενώ ακόμα λιγότερα είναι γνωστά σχετικά με τον τρόπο που το *Chl. tepidum* παράγει και στη συνέχεια καταναλώνει τα σφαιρίδια θείου που σχηματίζονται όταν υπάρχει περίσσεια θείου στο θρεπτικό υλικό. Τουλάχιστον δύο μονοπάτια φαίνεται να υπάρχουν για την οξείδωση του υδρόθειου στο περίπλασμα που περιλαμβάνουν εναλλακτικά την flavocytochrome c sulfide dehydrogenase (SoxEF/CT2080–2081) και την sulfide–quinone reductase (SQR) [26].

Επιπλέον, ένα θέμα που αφορά το σύστημα οξείδωσης του θειοθειικού και των σχετιζόμενων με αυτό μονοπατιών μεταφοράς ηλεκτρονίων, δεν έχει διασαφηνιστεί ακόμα στα πράσινα θειούχα βακτήρια. Ένα πολυενζυμικό Sox σύστημα, που εμπλέκεται στην οξείδωση του θειοθειικού, έχει βρεθεί να αποτελείται από τουλάχιστον 6 πρωτεΐνες (SoxAX, YZ, CD, και B) στο χημειολιθοτροφικό βακτήριο *Paracoccus pantotrophus* [27]. Αυτό το είδος ενζυμικού συστήματος φαίνεται να είναι ευρέως διαδεδομένο στα βακτήρια

που χρησιμοποιούν θειοθειικό [28]. Η γονιδιακή ανάλυση του *Chl. tepidum* έχει επίσης

αποκαλύψει την παρουσία μιας ομάδας από sox γονίδια, που υποθετικά εμπλέκονται στην οξείδωση του θειοθειικού [3, 29]. Ωστόσο, τα soxCD γονίδια, που φέρονται να αλληλεπιδρούν με την SoxY, απουσιάζουν από το γονιδίωμα του *Chl. tepidum* [1]. Έτσι, εναλλακτικοί μηχανισμοί έχουν προταθεί για την οξείδωση του θειοθειικού σε SO_4^{2-} , απουσία των SoxC/D [29], που πιθανόν να περιλαμβάνουν ένα ομόλογο της SoxF.

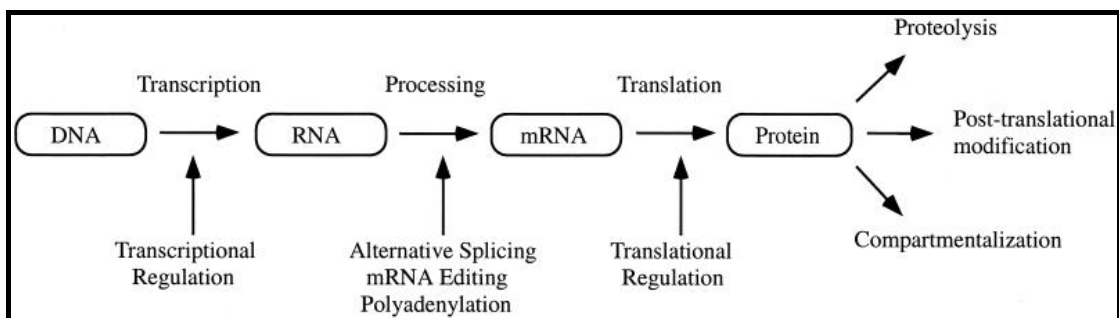
Αρκετά ένζυμα πιθανόν εμπλέκονται στο σχηματισμό των σφαιριδίων στοιχειακού θείου, λόγω της ύπαρξης γονιδίων που κωδικοποιούν πιθανές αναγωγάσες πολυσουλφιδίων, αναγωγάσες ετεροσουλφιδίων και θειοαναγωγάσες. Ωστόσο δεν έχει εξακριβωθεί κατά πόσο οι πρωτεΐνες αυτές εμπλέκονται στον μεταβολισμό του ανόργανου θείου. Σε κάθε περίπτωση περισσότερες πρωτεΐνες με άγνωστη λειτουργία θα πρέπει να εμπλέκονται στο μεταβολισμό των σφαιριδίων θείου [1].

1.2 Πρωτεομική

Ο όρος «πρωτεομική» χρησιμοποιήθηκε για πρώτη φορά το 1995 και ορίστηκε ως μεγάλης κλίμακας χαρακτηρισμός του συνόλου των πρωτεϊνών μιας κυτταρικής σειράς, ενός ιστού ή οργανισμού. Ο στόχος της πρωτεομικής είναι να αποκτήσουμε μια πιο καθολική και απαρτιωμένη θεώρηση της βιολογίας, μέσω της μελέτης όλων των πρωτεϊνών ενός κυττάρου, και όχι της καθεμίας ξεχωριστά [30]. Με βάση τον ορισμό της πρωτεομικής ανάλυσης πολλοί διαφορετικοί τομείς έρευνας εντάσσονται στο πλαίσιο της. Αυτοί περιλαμβάνουν μελέτες αλληλεπίδρασης, τροποποίησης, λειτουργίας και τοπολογίας πρωτεϊνών, καθώς και αρκετές άλλες. Σκοπός της πρωτεομικής είναι όχι απλά να ταυτοποιήσει όλες τις πρωτεΐνες σ' ένα κύτταρο, αλλά και να χαρακτηρίσει την τοπολογία κάθε πρωτεΐνης. Αυτοί οι φιλόδοξοι στόχοι απαιτούν την σύμπραξη πολλών

διαφορετικών επιστημονικών τομέων, όπως μοριακής βιολογίας, βιοχημείας, αναλυτικής χημείας και βιοπληροφορικής.

Στην προσπάθεια χαρακτηρισμού του ολικού πρωτεόματος ενός κυττάρου ή οργανισμού, πρέπει να λαμβάνεται υπόψη ότι το πρωτέομα είναι δυναμικό. Το πρωτέομα ενός κυττάρου αντανακλά το άμεσο περιβάλλον μέσα στο οποίο αυτό μελετάται. Αντιδρώντας σε εσωτερικά ή εξωτερικά ερεθίσματα, οι πρωτεΐνες μπορούν να τροποποιηθούν μετα-μεταφραστικά, να αλλάξουν τοπολογία στο κύτταρο, ή να αλλάξουν τα επίπεδα σύνθεσης και αποικοδόμησης τους [31]. Συνεπώς, εξετάζοντας το πρωτέομα ενός κυττάρου, λαμβάνουμε ένα στιγμιότυπο του πρωτεϊνικού περιβάλλοντος μία δεδομένη στιγμή. Λαμβάνοντας υπόψη όλες τις παραμέτρους άλλωστε, είναι πιθανό από κάθε γονιδίωμα να παραχθεί ένας άπειρος αριθμός πρωτεομάτων [32] (Εικόνα 5).



Εικόνα 5: Η δυναμική φύση του πρωτεόματος [33].

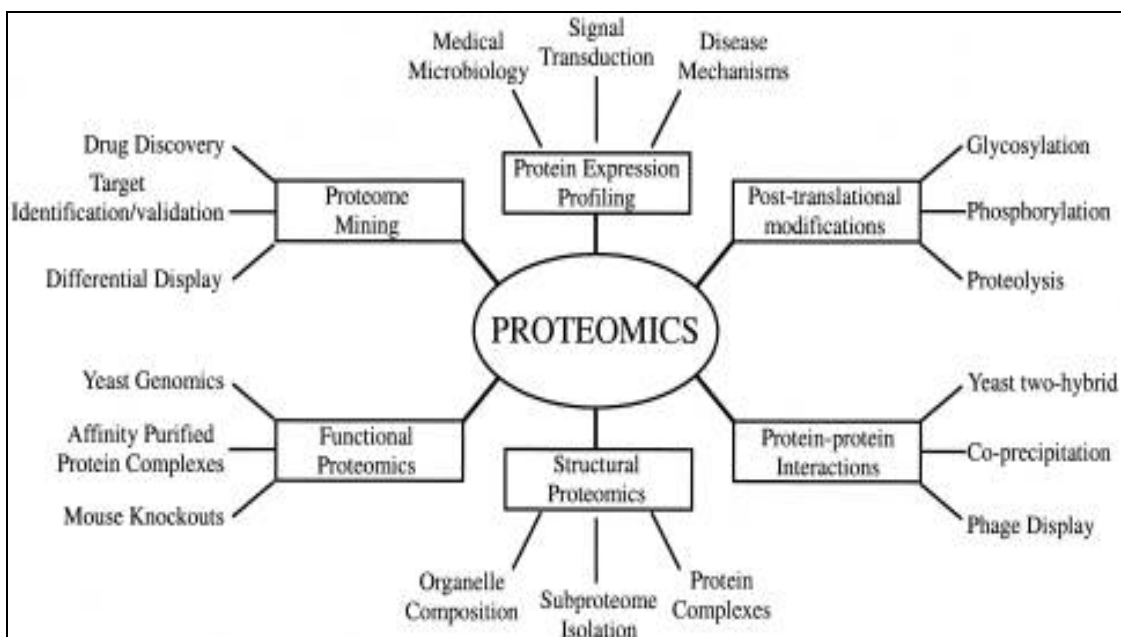
Τα πειράματα πρωτεομικής που περιλαμβάνουν χρήση φασματομετρίας μάζας, χωρίζονται σε 5 κατηγορίες: (1) ταυτοποίησης πρωτεϊνών, (2) ποσοτικοποίησης ή διαφορικής ανάλυσης, (3) πρωτεϊνικών αλληλεπιδράσεων, (4) μετά- μεταφραστικών τροποποιήσεων, και (5) δομικής πρωτεομικής ανάλυσης [34].

Η πρωτεομική θεωρείται μια έρευνα που ωθείται περισσότερο με βάση τα ευρήματα παρά με βάση την αρχική υπόθεση. Έτσι, στόχος της είναι να παράγει μεγάλες ποσότητες δεδομένων που μπορούν να χρησιμοποιηθούν για ένα τωρινό ή μελλοντικό επιστημονικό ερώτημα ή για να βρεθούν μοτίβα από τα οποία μπορούν να αντληθούν πληροφορίες, παρά να αποδειχθεί ή να καταρριφθεί μια στενά καθορισμένη υπόθεση [34].

Βασικός λόγος που η επιστημονική έρευνα εστιάζει στον τομέα τα πρωτεομικής είναι ότι

πολλές πληροφορίες δεν μπορούν να ληφθούν αποκλειστικά και μόνο από την μελέτη των γονιδίων. Για παράδειγμα, οι πρωτεΐνες, και όχι τα γονίδια, είναι υπεύθυνες για τον φαινότυπο των κυττάρων. Είναι αδύνατον να διασαφηνιστούν οι μηχανισμοί γήρανσης ή οι επιδράσεις του περιβάλλοντος αποκλειστικά μέσω της μελέτης του γονιδιώματος. Μόνο μέσω της μελέτης των πρωτεϊνών μπορούν να χαρακτηριστούν οι πρωτεϊνικές τροποποιήσεις και να αναγνωριστούν οι στόχοι νέων φαρμάκων. Γενικότερα, η πρωτεομική ανάλυση συμπληρώνει την γενετική πληροφορία και δίνει πληροφορίες για την έκφραση και την λειτουργία των πρωτεϊνών.

Η πρωτεομική βρίσκει πολλές εφαρμογές στην μελέτη της πρωτεϊνικής έκφρασης, μετά-μεταφραστικών τροποποιήσεων, πρωτεϊνικών αλληλεπιδράσεων, καθώς και σε δομικές, λειτουργικές και ποσοτικές μελέτες. (Εικόνα 6)



Εικόνα 6: Κατηγορίες πρωτεομικής και οι εφαρμογές τους στην βιολογία [33].

Ένας σημαντικός παράγοντας που βοήθησε σημαντικά στην ανάπτυξη της πρωτεομικής είναι η ραγδαία εξέλιξη των μεθόδων ιονισμού των πρωτεϊνών τα τελευταία χρόνια. Πριν από 35 περίπου χρόνια, όταν εμφανίστηκε η 2-DE, υπήρχαν πολύ λίγα εργαλεία για την πρωτεομική. Αν και μέθοδοι διαχωρισμού των πρωτεϊνών είναι γνωστές για δεκαετίες, η επανάσταση στην πρωτεομική έλαβε χώρα με την εισαγωγή 2 ριζοσπαστικών μεθόδων MS, που τιμήθηκαν με βραβείο Nobel, για την ανάλυση μεγάλων βιομορίων, του ESI [35] και του MALDI [36]. Οι μεθοδολογίες αυτές διέθεταν την απαραίτητη τεχνολογία για την ανάλυση πρωτεϊνών και πεπτιδίων με μεγάλη ευαισθησία, και αποτέλεσαν τον καταλυτικότερο παράγοντα για την μετέπειτα ραγδαία ανάπτυξη του τομέα της πρωτεομικής.

Έκτοτε, νέες τεχνολογίες έχουν εμφανιστεί και παλαιότερες έχουν βελτιωθεί σε τομείς που εκτείνονται από το διαχωρισμό των πρωτεϊνών μέχρι και την ταυτοποίηση τους. Ωστόσο, ακόμα πολλοί τύποι πειραμάτων πρωτεομικής δεν είναι δυνατόν να διεξαχθούν λόγω τεχνολογικών περιορισμών. Αυτά τα προβλήματα πρέπει να λυθούν και νέες τεχνολογίες πρέπει να αναπτυχθούν ώστε η πρωτεομική ανάλυση να φτάσει σε ιδανικά επίπεδα αποτελεσμάτων. Ένα τυπικό πείραμα πρωτεομικής μπορεί να χωριστεί στα παρακάτω βασικά βήματα: α) διαχωρισμός και απομόνωση των πρωτεϊνών ενός κυττάρου, ιστού ή οργανισμού, β) αναγνώριση και χαρακτηρισμός των πρωτεϊνών, και γ) χρήση βάσεων δεδομένων για περαιτέρω αξιολόγηση των ληφθέντων πληροφοριών.

1.2.1 Gel-based και gel-free πρωτεομική

Η πρωτεομική ανάλυση είναι στενά συνδεδεμένη με την δισδιάστατη ηλεκτροφόρηση σε πηκτή ακρυλαμίδιου, λόγω της μεγάλης διαχωριστικής της ικανότητας και της ποσοτικοποίησης που παρέχεται μέσω της χρώσης και της οπτικής ανάλυσης της προκύπτουσας πηκτής. Υπάρχουν ωστόσο αρκετοί περιορισμοί στην ανάλυση με 2-DE, όπως ο μη ικανοποιητικός διαχωρισμός στα ακραία μοριακά βάρη (<10 kDa, >150 kDa) ή pI (<3, >10), περιορισμοί στην ανίχνευση μέσω της χρώσης πρωτεϊνών που βρίσκονται σε

μικρή αφθονία (>1 ng), προβλήματα διαλυτότητας κάποιων μεμβρανικών πρωτεϊνών καθώς και το γεγονός ότι αυτή η μέθοδος είναι χρονοβόρα. Αν και γίνονται προσπάθειες για την αντιμετώπιση πολλών από αυτά τα προβλήματα, η κλασική δισδιάστατη ηλεκτροφόρηση δεν αποτελεί μια καθολική τεχνική διαχωρισμού για όλες τις τάξεις των πρωτεϊνών. Με την εισαγωγή των MALDI και ESI [35-36] κατέστη δυνατή η ανάλυση πρωτεϊνών και πεπτιδίων με αυξημένη ευαισθησία. Με βάση τις αναλυτικές προκλήσεις της πρωτεομικής, πρόσφατα παρατηρείται ραγδαία ανάπτυξη τεχνικών που δεν βασίζονται πλέον στην 2-DE για τη διερεύνηση των πολλών διαφορετικών ιδιοτήτων που παίζουν σημαντικό ρόλο στην πρωτεομική ανάλυση. Η ραγδαία βελτίωση στην οργανολογία του MS ώστε να επιτευχθεί μεγαλύτερη ευαισθησία και αυτοματοποίηση, έθεσε τις βάσεις για προσεγγίσεις μεγάλης εμβέλειας στη μελέτη των πολύπλοκων και ετερογενών ιδιοτήτων των πρωτεομάτων. Το γεγονός αυτό, σε συνδυασμό με τη συνεχώς αυξανόμενη διάθεση πλήρως αποκωδικοποιημένων γονιδιωμάτων, έχουν οδηγήσει στη δημιουργία ενός νέου ερευνητικού τομέα της πρωτεομικής ανάλυσης: gel-free, ή πεπτίδο-κεντρική πρωτεομική. Ο τομέας αυτός συνδυάζει gel-free διαχωρισμούς πρωτεϊνών και πεπτιδίων, νέες πρωτεϊνικές χημείες και μεθόδους εμπλουτισμού τους, με πολύ αυτοματοποιημένη οργανολογία MS, για να παρέχει νέα εργαλεία ανάδειξης των ιδιοτήτων των πρωτεομάτων με αυξημένη ευαισθησία και απόδοση. Στην περίπτωση αυτή, αντί να αναλύεται απευθείας η πρωτεΐνη, αναλύονται τα πεπτίδια της, τα οποία είναι σε γενικές γραμμές πιο υδατοδιαλυτά και μπορούν να αναλυθούν κατευθείαν με φασματομετρία μάζας.

Η παντελής κατάργηση του ηλεκτροφορητικού διαχωρισμού απαιτεί τη χρήση αποτελεσματικών βημάτων κλασματοποίησης του πρωτεϊνικού μίγματος ή πολύ υψηλής διακριτικής ικανότητας διαχωρισμό στην περίπτωση μίγματος πεπτιδίων. Για παράδειγμα, ένα τυπικό κύτταρο θηλαστικού μπορεί να εκφράσει μέχρι και 10.000 διαφορετικές πρωτεΐνες με τεράστιο εύρος στις συγκεντρώσεις τους. Μια μέση πρωτεΐνη 50 kDa αναμένεται να μετατραπεί σε περίπου 40 πεπτίδια μετά από θρυψινοποίηση της. Έτσι, η επεξεργασία με θρυψίνη ενός ολικού κυτταρικού εκχυλίσματος θα μπορούσε να αποδώσει μέχρι και 400.000 πεπτίδια με μεγάλο εύρος στις συγκεντρώσεις τους. Είναι συνεπώς προφανές ότι πολλαπλά βήματα διαχωρισμού είναι απαραίτητα να αποτελούν μέρος της ανάλυσης ενός τόσο πολύπλοκου μίγματος [37].

Οι τεχνικές κλασματοποίησης των πρωτεομάτων στοχεύουν στην μείωση της ετερογένειας και της πολυπλοκότητας των δειγμάτων, αυξάνοντας έτσι τον αριθμό και τη συγκέντρωση των διαφορετικών πρωτεϊνών στα υπό-πρωτεόματα που προκύπτουν. Πρωτεϊνικά δείγματα μπορούν να κλασματοποιηθούν με βάση τη σχετική τους διαλυτότητα σε μια σειρά από ρυθμιστικά διαλύματα, με βάση στις διαφορετικές τους φυσικοχημικές ιδιότητες (π.χ. φορτίο, μέγεθος, υδροφοβικός χαρακτήρας, ή συγγένεια). Η υποκυτταρική κλασματοποίηση αποτελεί ακόμη μια χρήσιμη προσέγγιση που στοχεύει τον εμπλουτισμό συσχετισμένων πρωτεϊνών με βάση την κοινή τους τοπολογία στο κύτταρο. Γενικά, οι διαδικασίες κλασματοποίησης θα πρέπει να παραμένουν όσο το δυνατόν απλούστερες, να στοχεύουν σε διαφορετικές μοριακές ιδιότητες και να χρησιμοποιούν τεχνικές και διαλύματα συμβατά με τις μεθόδους ταυτοποίησης των πρωτεϊνών, ώστε να υπάρχει η μικρότερη δυνατή απώλεια δείγματος, αποικοδόμηση πρωτεϊνών ή εισαγωγή *in vitro* τροποποιήσεων σε αυτές [37].

Στην gel-free πρωτεομική, μετά την απομόνωση ενός (υπό-) πρωτεόματος, λαμβάνει χώρα ενζυματική πέψη που μετατρέπει τις πρωτεΐνες σε πεπτίδια, κατά κανόνα με ένα ειδικό ένζυμο όπως η θρυψίνη, και στη συνέχεια το μίγμα των πεπτιδίων που προκύπτει αναλύεται με φασματομετρία μάζας.

Το προκύπτον μίγμα πεπτιδίων συχνά περιέχει δεκάδες χιλιάδες πεπτίδια και συνεπώς είναι πολύπλοκο για απευθείας MS/MS ανάλυση. Δύο γενικές προσεγγίσεις ακολουθούνται για την αντιμετώπιση αυτού του προβλήματος: είτε το μίγμα των πεπτιδίων διαχωρίζεται με χρήση διαφορετικών τεχνικών υγρής χρωματογραφίας πριν το MS/MS, ή συγκεκριμένες τάξεις πεπτιδίων επιλέγονται με βάση τη συγγένεια τους και χρησιμοποιούνται ως αντιπροσωπευτικά της αρχικής πρωτεΐνης από την οποία προήλθαν. Ωστόσο με την LC-MS/MS ανάλυση πολύπλοκων μιγμάτων πεπτιδίων, αν και υπάρχει διαχωρισμός τους με βάση κάποιες φυσικοχημικές τους ιδιότητες, ο συνολικός αριθμός των προς ανάλυση πεπτιδίων παραμένει ο ίδιος. Έτσι, και όταν ακόμα ο διαχωρισμός επιτρέπει την MS/MS ανάλυση των περισσότερων πεπτιδίων, η ανάλυση παραμένει μια πολύ χρονοβόρα διαδικασία.

Η επιλογή ενός μικρότερου αριθμού πεπτιδίων, αντιπροσωπευτικών της κάθε πρωτεΐνης

που αρχικά υπήρχε στο μίγμα, μειώνει τον όγκο του αρχικού προς ανάλυση δείγματος, μειώνοντας έτσι και το χρόνο ανάλυσης του και τον όγκο των δεδομένων που προκύπτει από αυτήν. Ωστόσο, όταν μειώνεται η πολυπλοκότητα ενός προς ανάλυση δείγματος, δύο σημαντικά θέματα πρέπει να λαμβάνονται υπ' όψιν. Πρώτον, το επιλεγμένο μίγμα των πεπτιδίων θα πρέπει να αντιπροσωπεύει το αρχικό μίγμα στο μεγαλύτερο δυνατό βαθμό, έχοντας παράλληλα και τη μέγιστη δυνατή μείωση της πολυπλοκότητας του. Δεύτερον, η διαδικασία απομόνωσης των αντιπροσωπευτικών πεπτιδίων θα πρέπει να είναι όσο το δυνατόν πιο ειδική και να έχει πολύ καλή απόδοση, ώστε να απομονώνονται όλα τα επιθυμητά πεπτίδια και να αποφεύγονται προσμίξεις [37].

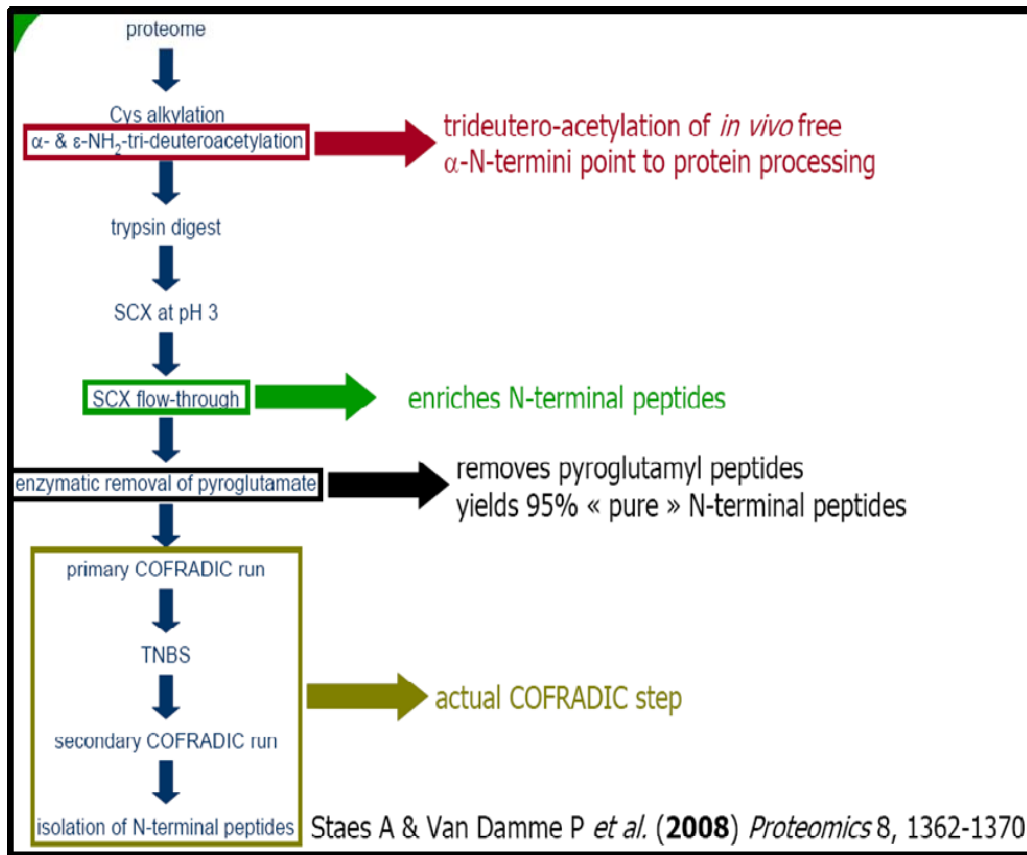
1.2.2 Το COFRADIC ως εργαλείο διαχωρισμού πεπτιδίων [37-40]

Το COFRADIC (COmbined FRActional DIagonal Chromatography) αποτελεί μια σειρά από τεχνικές που βασίζονται στην αρχή της διαγώνιας χρωματογραφίας για την επιλογή αντιπροσωπευτικών πεπτιδίων, ανάλογα με το εκάστοτε βιολογικό ενδιαφέρον, από πρωτεόματα μετά την κατεργασία τους με θρυψίνη. Η τεχνική αυτή βασίζεται σε έναν αρχικό διαχωρισμό του μίγματος των πεπτιδίων με RP-HPLC, μια χημική αντίδραση που τροποποιεί εκλεκτικά την δομή των προς επιλογή πεπτιδίων και μια επανάληψη του αρχικού διαχωρισμού του μίγματος. Τα πεπτίδια που παραμένουν τα ίδια, εκλούνται στην ίδια θέση και στις 2 χρωματογραφικές αναλύσεις, ενώ τα τροποποιημένα πεπτίδια εκλούνται νωρίτερα ή αργότερα, ανάλογα με την χημική αντίδραση που έχει προηγηθεί. Για να μειωθεί ο αριθμός των επαναλαμβανόμενων χρωματογραφικών αναλύσεων, αρκετά κλάσματα από την πρώτη ανάλυση RP-HPLC μπορούν να συνδυαστούν και να υποβληθούν στην χημική αντίδραση τροποποίησης.

Το COFRADIC συνεπώς είναι μια πολύ ευέλικτη και ευπροσάρμοστη τεχνική στα ζητούμενα του εκάστοτε βιολογικού ερωτήματος, καθώς οποιαδήποτε χημική ή

ενζυματική τροποποίηση που είναι πολύ εκλεκτική, ποσοτική και προκαλεί αρκετά μεγάλη μετατόπιση στις κορυφές των τροποποιημένων πεπτιδίων, μπορεί να χρησιμοποιηθεί για την επιλογή αντιπροσωπευτικών πεπτιδίων με διαφορετικές κάθε φορά ιδιότητες. Έτσι μπορούν να επιλεγθούν πεπτίδια με βάση τη χημική φύση των πλευρικών αλυσίδων των αμινοξέων τους, πχ κυστεΐνες, ή με βάση τις μετά-μεταφραστικές τροποποιήσεις που έχουν υποστεί, πχ φωσφορυλίωση [37, 41].

Ιδανικά, η μείωση της πολυπλοκότητας του δείγματος θα πρέπει να είναι όσο το δυνατόν μεγαλύτερη και όλες οι πρωτεΐνες που αρχικά υπήρχαν στο δείγμα θα πρέπει να αντιπροσωπεύονται από τουλάχιστον ένα πεπτίδιο. Η μεγαλύτερη δυνατή μείωση της πολυπλοκότητας ενός μίγματος πεπτιδίων μέσω της τεχνικής COFRADIC επιτυγχάνεται όταν επιλέγονται και τροποποιούνται τα N-τερματικά πεπτίδια που αντιστοιχούν στις πρωτεΐνες του αρχικού δείγματος (Εικόνα 7). Έτσι κάθε πρωτεΐνη θα αντιπροσωπεύεται από ένα πεπτίδιο και επιπλέον μπορούν να ανιχνευθούν μετά-μεταφραστικές τροποποιήσεις των πρωτεϊνών στο άκρο αυτό καθώς και ισομορφές πρωτεϊνών, οι οποίες διαφοροποιούνται κυρίως στα ακραία αμινοξέα τους [38-40].



Εικόνα 7: Σχηματική αναπαράσταση της N-τερματικής COFRADIC ανάλυσης [39].

Στην περίπτωση του N-terminal COFRADIC, το αρχικό πρωτεϊνικό δείγμα υφίσταται αναγωγή και αλκυλίωση όλων των κυστεϊνών του. Αυτό συμβαίνει γιατί το αντιδραστήριο της ενδιάμεσης χημικής αντίδρασης τροποποίησης (TNBS, 2,4,6-τρινιτροβενζενосуλφονικό οξύ, Εικόνα 8) μπορεί να αντιδράσει και με ελεύθερες θειόλες και να υπάρχει αλλοίωση των αποτελεσμάτων της χημικής τροποποίησης. Στη συνέχεια, όλες οι ελεύθερες αμίνες (α- και ε-αμίνες) προστατεύονται με ακετυλίωση, χρησιμοποιώντας sulfo-N-hydroxysuccinimide acetate. Μετά από ένα βήμα αφάλατωσης, το τροποποιημένο αυτό πρωτόμα θρυψινοποιείται, με την θρυψίνη όμως να πρωτεολύει ειδικά μόνο μετά από αργινίνη, αφού οι ακετυλιωμένες πλέον λυσίνες δεν αποτελούν πια υπόστρωμα της. Το προκύπτον πρωτεολυμένο δείγμα περιέχει 2 τύπους πεπτιδίων: N-

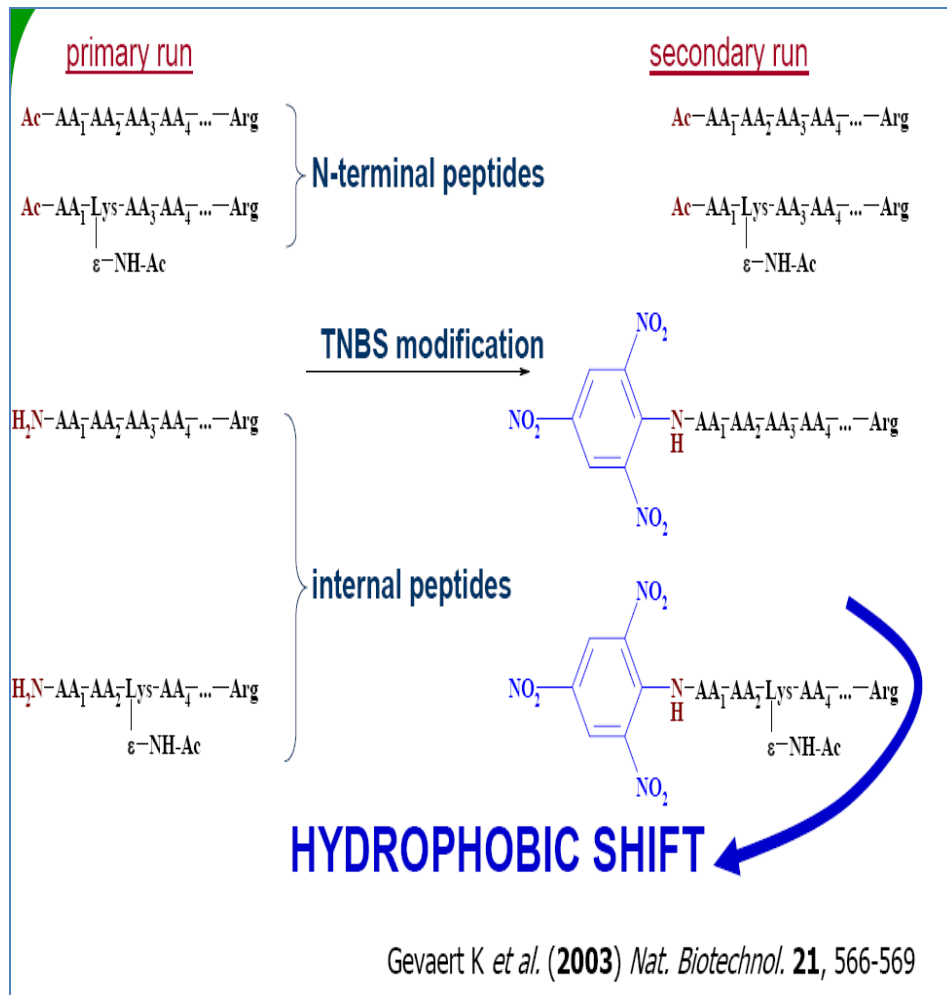
τερματικά και εσωτερικά. Τα N-τερματικά πεπτίδια διαθέτουν προστατευμένο α-αμινοτερματικό (είτε *in vivo* ή *in vitro* λόγω της αντίδρασης ακετυλίωσης), ενώ τα εσωτερικά πεπτίδια διαθέτουν ελεύθερο α-αμινο-τελικό άκρο.

Δεδομένης της φύσης των πεπτιδίων που προκύπτουν μετά την θρυψινοποίηση, διαφορετικά φορτισμένα πεπτίδια αναμένεται να προκύψουν σε pH 3. Το συνολικό φορτίο των πεπτιδίων αυτών θα καθορίζεται από τις πλευρικές ομάδες αργινίνης και ιστιδίνης, καθώς και από τις ελεύθερες α-αμινομάδες που είναι θετικά φορτισμένες σε pH 3. Εκτός από τις C-τερματικές καρβοξυλομάδες των πεπτιδίων που έχουν $pK_a < 3$, σε pH 3 οι πλευρικές ομάδες του ασπαρτικού οξέος και του γλουταμινικού οξέος είναι ουδέτερες. Η ε-αμίνη της λυσίνης είναι και αυτή ουδέτερη λόγω της ακετυλίωσης της. Έτσι, συνολικά, τα N-τερματικά πεπτίδια που είναι προστατευμένα -δηλαδή τα (τριδευτερο) ακετυλιωμένα πεπτίδια και αυτά που εκκινούν με pyroglutamate carboxylic acid ή κυκλοποιημένη S-καρβαμυλμεθυλκυστεΐνη- θα περιέχουν ένα λιγότερο θετικό φορτίο σε pH 3, σε σύγκριση με τα πεπτίδια που διαθέτουν ελεύθερη α-N-αμινομάδα ή που εκκινούν με προλίνη. Συνεπώς, με την εξαίρεση των προστατευμένων πεπτιδίων που περιέχουν ιστιδίνη ή που διαθέτουν εσωτερικά κατάλοιπα αργινίνης λόγω ανεπαρκούς θρυψινοποίησης, η πλειοψηφία των προστατευμένων πεπτιδίων περνάει από τη SCX στήλη, ενώ τα περισσότερα πεπτίδια που διαθέτουν ελεύθερη α-N-ομάδα συγκρατούνται. Έτσι η κατιοντική στήλη χρησιμοποιείται για να εμπλουτίσει σε πρώτο στάδιο τα N-τερματικά πεπτίδια. Για να αυξηθεί περαιτέρω ο τελικός αριθμός των α-αμινο-ακετυλιωμένων (N-

τερματικών) πεπτιδίων εισήχθη ένα επιπλέον ενζυμικό βήμα πριν την πρώτη RP-HPLC ανάλυση για την αφαίρεση του pyroglutamate carboxylic acid από το N-τερματικό άκρο των πεπτιδίων που είχαν επιλεγεί μέσω της SCX στήλης. Οι pGAPασες διασπούν αποτελεσματικά το δεσμό του pyroglutamate carboxylic acid και ως αποτέλεσμα τα pyroglutamyl πεπτίδια χάνουν το πρώτο τους αμινοξύ και διαθέτουν πλέον μια νέα ελεύθερη α-αμινομάδα. Αυτή η ομάδα μπορεί να αντιδράσει με το TNBS τα πεπτίδια αυτά μπορούν πλέον να διαχωριστούν ως N-τερματικά (Εικόνα 8). Χρησιμοποιείται παράλληλα και η Qκυκλάση για να υπάρξει πλήρης κυκλοποίηση των N-τερματικών γλουταμικών

οξέων πριν τη χρήση της pGAPασης. Το ενζυμικό αυτό βήμα λαμβάνει χώρα μετά την SCX χρωματογραφία γιατί παρουσία θρυψίνης θα υπήρχε απενεργοποίηση των ενζύμων αυτών και επειδή αυτά τα ένζυμα θα έχαναν μέρος της ενεργότητας τους στις συνθήκες που απαιτούνται στην κατιοντική χρωματογραφία.

Το μίγμα των πεπτιδίων που προκύπτει διαχωρίζεται αρχικά με RP-HPLC και τα κλάσματα που προκύπτουν, αφού ξηραθούν και επαναδιαλυτοποιηθούν σε ρυθμιστικό διάλυμα βασικού pH (9.5), υφίστανται κατεργασία με TNBS. Το αντιδραστήριο αυτό αντιδρά ποσοτικά με τις ελεύθερες αμίνες και κατ' αυτόν τον τρόπο τροποποιεί τα εσωτερικά πεπτίδια, εισάγοντας τους μια υδρόφοβη τρινιτροφαινυλομάδα (Εικόνα 8). Η χημική δομή των N-τερματικών πεπτιδίων ωστόσο δεν αλλάζει, δεδομένου ότι όλες οι αμίνες τους είναι ήδη προστατευμένες. Κατά τον επόμενο διαχωρισμό του μίγματος των πεπτιδίων με RP-HPLC, τα τροποποιημένα εσωτερικά πεπτίδια υφίσταται έντονη μετατόπιση λόγω του αυξημένου υδρόφοβου χαρακτήρα τους, ενώ τα N-τερματικά πεπτίδια εκλούνται εντός των ορίων της έκλουσης τους κατά την πρώτη RP-HPLC ανάλυση. Τα πεπτίδια που αρχίζουν με κυκλικά αμινοξέα, όπως προλίνη, εμφανίζονται ταυτόχρονα με τα N-τερματικά πεπτίδια, καθώς ούτε αυτά δεν αντιδρούν με το TNBS. Ωστόσο, ο αριθμός των πεπτιδίων αυτών είναι πολύ μικρός και κατά μέσο όρο 70-80% των πεπτιδίων που τελικά απομονώνονται είναι N-τερματικά [38-40].



Εικόνα 8: Ενδιάμεση χημική αντίδραση τροποποίησης N-τερματικής COFRADIC ανάλυσης με TNBS [37].

Συμπερασματικά, η τεχνική N-terminal COFRADIC παρουσιάζει εξαιρετικό ενδιαφέρον, διότι δεν μειώνει απλά στο μεγαλύτερο δυνατό βαθμό την πολυπλοκότητα του δείγματος, διατηρώντας παράλληλα αντιπροσωπευτικά πεπτιδία του, αλλά επίσης να εξετάσει τις τροποποιήσεις που υφίστανται τα N-τερματικά άκρα των πρωτεϊνών, μια διαδικασία που ενώ είναι πολύ σημαντική δεν υπάρχουν μέχρι στιγμής τεχνικές για τη πρωτεομική μελέτη της σε μεγάλη κλίμακα. Γενικότερα η COFRADIC ανάλυση παρουσιάζει το επιπλέον πλεονέκτημα της ευελιξίας, καθώς επιτρέπει την απομόνωση οποιασδήποτε τάξης πεπτιδίων που μπορούν να υποστούν ειδική τροποποίηση, τέτοια ώστε να αλλάξουν οι χρωματογραφικές τους ιδιότητες [38-40].

1.3 Σκοπός Μεταπτυχιακής Εργασίας

Σκοπός της παρούσας εργασίας ήταν η κλασματοποίηση του πρωτεόματος του πράσινου θειούχου βακτηρίου *Chlorobium tepidum* και ο χαρακτηρισμός του με gel-free μεθόδους πρωτεομικής ανάλυσης που εστιάζουν στη χρήση φασματοσκοπίας μάζας (nLC-MS/MS και N-terminal COFRADIC). Το πρωτέομα του *Chl. tepidum* διαχωρίστηκε στα υποπρωτέομα του κυτταροπλάσματος, του περιπλάσματος, των ολικών μεμβρανών (κυτταροπλασματική και εξωτερική μεμβράνη και χλωροσώματα), των χλωροσωμάτων, καθώς και στο υδατοδιαλυτό υποπρωτέομα που περιλαμβάνει το κυτταρόπλασμα και το περίπλασμα του βακτηρίου. Στόχο της εργασίας αυτής αποτελεί η πρωτεομική μελέτη του *Chl. tepidum*, δηλαδή ο η μελέτη κυρίως της έκφρασης, αλλά και της τοπολογίας-μέσω της κλασματοποίησης- των πρωτεϊνών που εκφράζονται στον οργανισμό αυτόν, στόχος που καθίσταται εφικτός λόγω του πρόσφατου προσδιορισμού της πλήρους αλληλουχίας των νουκλεϊκών οξέων του γονιδιώματος του *Chl. tepidum* [3] που έχει προσφέρει την δυνατότητα διερεύνησης σε ευρεία κλίμακα, της έκφρασης των γονιδίων του οργανισμού αυτού.

Επιπλέον, χαρακτηρίστηκαν τα απομονωμένα χλωροσώματα του *Chl. tepidum* με δυναμική και στατική σκέδαση φωτός (DLS) και ηλεκτρονική μικροσκοπία (EM) προκειμένου να ληφθούν πληροφορίες για τις διαστάσεις τους προτού αναλυθεί η πρωτεϊνική τους σύσταση με nLC-MS/MS.

Συμπερασματικά, η πρωτεομική μελέτη του πράσινου θειούχου βακτηρίου *Chlorobium tepidum* με κλασματοποίηση του πρωτεόματος του και με χαρακτηρισμό των επιμέρους υποπρωτεομάτων με N-terminal COFRADIC και nLC-MS/MS, στοχεύει στην ταυτοποίηση του μεγαλύτερου δυνατού ποσοστού πρωτεϊνών επί του θεωρητικού πρωτεόματος του βακτηρίου αυτού. Ο συνδυασμός της απλοποίησης του πρωτεϊνικού

δείγματος με την κλασματοποίηση του και gel-free τεχνικών που επικεντρώνονται στην ανάλυση συγκεκριμένων τάξεων πεπτιδίων, απλοποιώντας περαιτέρω το προς ανάλυση δείγμα, στοχεύει στην μεγαλύτερη δυνατή αποτελεσματικότητα της πρωτεομικής μελέτης ενός τόσο πολύπλοκου πρωτεϊνικού δείγματος όπως είναι το ολικό πρωτέομα του *Chl. tepidum*.

2. ΥΛΙΚΑ – ΜΕΘΟΔΟΙ

2.1.1 Οργανισμός

Το πράσινο θειούχο βακτήριο *Chl. tepidum* που μελετήθηκε είναι ευγενική προσφορά του Prof. G. Hauska, University of Regensburg, Germany.[42]

2.1.2 Ανάπτυξη και απομόνωση κυττάρων του *Chlorobium tepidum*

Παρασκευή καλλιέργειας Ht.

Για την παρασκευή της ποσότητας του θρεπτικού υλικού για την ανάπτυξη του *Chlorobium tepidum* χρησιμοποιήθηκαν οι εξής ποσότητες των παρακάτω χημικών ενώσεων, σύμφωνα με [42]

Αντιδραστήριο	Ποσότητα
$\text{Na}_2\text{S}_2\text{O}_3 \times 5\text{H}_2\text{O}$	1gr
KH_2PO_4	0.50gr
$\text{CH}_3\text{COONH}_4$	0.50gr
NH_4Cl	0.40gr
NaCl	0.40gr
$\text{MgSO}_4 \times 7\text{H}_2\text{O}$	0.20gr
$\text{CaCl}_2 \times 2\text{H}_2\text{O}$	0.050gr
EDTA	0.0125gr
NaHCO_3	2.0gr
Μείγμα Ιχνοστοιχείων	1ml
Βιταμίνη B12 (20mg/l)	2ml

Ιχνοστοιχεία:

Αντιδραστήριο	Ποσότητα
EDTA	0.67gr
FeCl ₃ x 6H ₂ O	0.2gr
CoCl ₂ x 6H ₂ O	19mg
Na ₂ MoO ₄ x 2H ₂ O	19mg
ZnSO ₄ x 7H ₂ O	15mg
MnCl ₂ x 4H ₂ O	10mg
VSO ₄ x 2H ₂ O	3mg
NiCl ₂ x 6H ₂ O	2.5mg
CuCl ₂ x 2H ₂ O	1.7mg
H ₃ BO ₃	0.6mg
NaWO ₄ x 2H ₂ O	0.2mg
Na ₂ HSeO ₃	0.2mg
NaOH	έως pH = 8

Οι παραπάνω ενώσεις διαλύονται σε 1 lt απεσταγμένου νερού και το προκύπτον διάλυμα υφίσταται ανάδευση έως ότου διαλυθούν όλα τα στερεά υπολείμματα. Έπειτα 800 ml του διαλύματος τοποθετούνται σε φιάλη 1 lt. Ακολουθεί απαέρωση στη φιάλη για 15 min με διαβίβαση αερίου αζώτου υπό χαμηλή πίεση, αφού προηγουμένως έχει κλειστεί το στόμιο της φιάλης, ώστε να μη μπαίνει ατμοσφαιρικό οξυγόνο. Στη συνέχεια προστίθεται στο θρεπτικό 5 ml υδατικού διαλύματος που περιέχει 0.6 gr Na_2S και η φιάλη κλείνεται με πώμα.

Ακολουθεί αποστείρωση για 20 min στους 120°C και έπειτα ρύθμιση του pH περιοχή 6.7-6.9 χρησιμοποιώντας αποστειρωμένο διάλυμα HCl 10% (περίπου 2,3 ml). Στη συνέχεια προστίθενται με σύριγγα 20 ml διαλύματος βακτηρίων υπό αποστειρωμένες συνθήκες και η φιάλη τοποθετείται σε υδατόλουτρο σε θερμοκρασία $40-45^\circ\text{C}$ και πλήρη φωτισμό για μία περίπου μέρα. Η πλήρης ανάπτυξη των βακτηρίων διαπιστώνεται από την αλλαγή του χρώματος της φιάλης από ανοικτό πράσινο σε σκούρο πράσινο και ταυτόχρονη αύξηση της απορρόφησης στα 750 nm, γεγονός το οποίο πιστοποιεί ότι η καλλιέργεια έχει φτάσει στη μέγιστη ανάπτυξη της. Τότε η φιάλη απομακρύνεται από το υδατόλουτρο και είτε φυλάσσεται στους 4°C , είτε συλλέγονται τα κύτταρα.

Συλλογή κυττάρων

Τα κύτταρα συλλέγονται με φυγοκέντρηση στα 6000 g για 20 min. Μετά το τέλος της φυγοκέντρησης παρατηρείται η δημιουργία πολύ σκούρου πράσινου ιζήματος όπου βρίσκονται τα κύτταρα του *Chlorobium tepidum*, καθώς και ο σχηματισμός υποκίτρινου ιζήματος το οποίο οφείλεται στην παρουσία στοιχειακού θείου. Κατά τη διάρκεια της συλλογής των κυττάρων ωστόσο, αποφεύγεται η λήψη του ιζήματος που οφείλεται στο στοιχειακό θείο. Τα κύτταρα τέλος φυλάσσονται για περαιτέρω χρήση.

2.1.2 Διάρρηξη κυττάρων: απομόνωση περιπλασματικού - κυτταροπλασματικού πρωτεόματος

Η ποσότητα των κυττάρων που έχει συλλεχθεί υπό μορφή ιζήματος, διαλύεται-σε αναλογία 10 ml διαλύματος ανά γραμμάριο ιζήματος-σε ένα υδατικό ρυθμιστικό διάλυμα όσμωσης το οποίο αποτελείται από: 50 mM Tris pH=8, 0,5 M Σακχαρόζη και 1mM EDTA. Ο ρόλος του ρυθμιστικού αυτού διαλύματος όσμωσης είναι με τη βοήθεια ανάδευσης, να εισέλθει διαμέσου της εξωτερικής μεμβράνης και με αξιοποίηση του φαινομένου της όσμωσης να προκαλέσει οπές σ' αυτή τη μεμβράνη, επιτρέποντας έτσι στο περίπλασμα να εξέλθει του κυττάρου, αφήνοντας την κυτταροπλασματική μεμβράνη άθικτη.

Στη συνέχεια τα κύτταρα ομογενοποιούνται για 20min, ενώ καθ' όλη την διαδικασία απομόνωσης του περιπλάσματος τα δείγματα διατηρούνται σε πάγο.

Στη συνέχεια το διάλυμα φυγοκεντρείται για 15min σε 10.000 g. Μετά τη φυγοκέντρηση, στο υπερκείμενο βρίσκεται το περίπλασμα των κυττάρων διαλυμένο στο ρυθμιστικό διάλυμα όσμωσης, ενώ στο ίζημα βρίσκονται τα άσπαστα κύτταρα καθώς και οι μεμβράνες και τα λοιπά κυτταρικά τμήματα των βακτηρίων, πλην του περιπλάσματος. Ακολούθως, το περίπλασμα συλλέγεται και φυλάσσεται στους -20°C.

Το ίζημα διαλύεται σε αναλογία 1:20 σε ρυθμιστικό διάλυμας TT και προσαρμόζονται σε αυτό υπέρηχοι. Ο κάθε υπέρηχος είναι έντασης 30 Watt, διαρκεί 15 sec και ακολουθεί αναμονή 45 sec. Συνολικά εκτελούνται 10 επαναλήψεις. Η παρουσία υπερήχων έχει ως στόχο τη διάσπαση του κυτταρικού τοιχώματος και τη διάρρηξη της κυτταρικής μεμβράνης με αποτέλεσμα την απελευθέρωση του ενδοκυτταρικού διαλύματος. Το διάστημα των 45 sec όπου πραγματοποιείται διακοπή της λειτουργίας των υπερήχων είναι απαραίτητο ώστε να επέλθει θερμική ισορροπία στο διάλυμα και για αυτό τον λόγο άλλωστε πρέπει καθ' όλη τη διάρκεια της πειραματικής διαδικασίας να

υπάρχει πάγος. Στη συνέχεια, το διάλυμα που προκύπτει τοποθετείται σε σωλήνα υπερφυγοκέντρωσης και φυγοκεντρείται για 1,5 h σε 200.000 g στους 4⁰C. Στο υπερκείμενο, το οποίο συλλέγεται και διατηρείται στους -20⁰C βρίσκεται το κυτταρόπλασμα του κυττάρου.

Ρυθμιστικό Διάλυμα TT

20 mM Tricine Base

Tris Base (έως pH = 7.5)

1 mM Benzamidine

1 mM 6–amino-caproic acid

2.1.3 Διακριτή απομόνωση υδατοδιαλυτού και μεμβρανικού τμήματος του *Chl. teridum*

Η ποσότητα των βακτηριακών κυττάρων που έχει συλλεχθεί κατά τη διαδικασία λήψης κυττάρων και βρίσκεται υπό μορφή ιζήματος διαλύεται κατ' αναλογία 1:20 σε ρυθμιστικό διάλυμα TTN το οποίο περιέχει αναστολείς πρωτεασών ώστε να αποφευχθεί πιθανή πρωτεόλυση των πρωτεϊνών. Το μίγμα ομογενοποιείται αρχικά με τη βοήθεια πιπέτας, ακολούθως με την χρήση ομογενοποιητή και στη συνέχεια με υπέρηχους. Η διαδικασία διάρρηξης με υπέρηχους στη μέγιστη ισχύ (20 επαναλήψεις, 15 sec με ενδιάμεσες παύσεις 45 sec) λαμβάνει χώρα σε πάγο. Η ομογενοποίηση πραγματοποιείται παρουσία ομογενοποιητή μέσα σε πάγο, για να προκύψει όσο το δυνατόν περισσότερο ομογενοποιημένο μίγμα και για να αποφευχθεί ο σχηματισμός συσσωματωμάτων.

Το προκύπτον διάλυμα φυγοκεντρείται για 20 min σε 10.000 g στους 6⁰C (Beckmann) ώστε να απομακρυνθούν υπό μορφή ιζήματος τα άθραυστα κύτταρα. Το υπερκείμενο της φυγοκέντρωσης υφίσταται υπερφυγοκέντρωση για 1,5 h σε 200.000 g. Μετά το πέρας της υπερφυγοκέντρωσης στο υπερκείμενο βρίσκονται το κυτταρόπλασμα και το

περίπλασμα του κυττάρου (υδατοδιαλυτό τμήμα κυττάρου), ενώ ως ίζημα έχουν καταβυθιστεί οι μεμβράνες του βακτηρίου.

Το υπερκείμενο, και το ίζημα μετά από διάλυση σε 1 ml διαλύματος TTN και ομογενοποίηση, συλλέγονται και διατηρούνται στους -40°C .

2.1.4 Απομόνωση χλωροσωμικού κλάσματος

Το μεμβρανικό κλάσμα του *Chl. tepidum* απομονώθηκε σύμφωνα με τη διαδικασία που περιγράφηκε από Gerola et al.[43]. Τα χλωροσώματα απομονώθηκαν χρησιμοποιώντας τροποποιημένη τη μεθοδο [44]. Αναλυτικά, 4 gr συλλεγμένων κυττάρων διαλυτοποιήθηκαν σε 80 ml διαλύματος απομόνωσης (20 mM Tricine-Tris pH 7.5, 2 M NaSCN, 1 mM βενζαμιδίνης 1 mM αμινο-καπροϊκού οξέος, 5 mM EDTA, 2 mM DTT και 50 $\mu\text{g}/\text{ml}$ λυσοζύμης) και ομογενοποιήθηκαν. Το προκύπτον διάλυμα στη συνέχεια τα κύτταρα διαρρήχθηκαν με υπερήχους (20 επαναλήψεις, 15 sec με ενδιάμεσες παύσεις 45 sec), διατηρώντας το δείγμα σε πάγο για να διατηρηθεί η θερμοκρασία του κάτω από 10°C . Το εναιώρημα αφέθηκε για 30 min σε θερμοκρασία δωματίου προκειμένου να δράσει η λυσοζύμη. Τα άθραυστα κύτταρα απομακρύνθηκαν με φυγοκέντρηση στα 10.000 g για 15 min στους 4°C . Το προκύπτον υπερκείμενο φυγοκεντρήθηκε στα 200,000 g για 1.5 h στους 4°C για να ληφθεί ως ίζημα το ολικό μεμβρανικό κλάσμα, που συλλέχθηκε και αποθηκεύτηκε στους -80°C για περαιτέρω χρήση.

Το ολικό μεμβρανικό κλάσμα επαναδιαλυτοποιήθηκε σε ρυθμιστικό διάλυμα απομόνωσης, απουσία λυσοζύμης μέχρι οπτικής πυκνότητας 3 στα 750 nm και στη συνέχεια προστέθηκαν 0,3 gr ιωδιούχου νατρίου ανά ml εναιωρήματος, οπότε και το δείγμα υπέστη κατεργασία με υπερήχους 2 φορές για 1 min με Broxon ultrasonic processor στη μέγιστη ισχύ. Έπειτα προστέθηκαν σακχαρόζη και το διπολικό ιονικό

απορρυπαντικό (Nlauryl- β-ιμινοδιπροπιονικό-Na άλας) σε συγκεντρώσεις 120 mg/ml και 0,08% (w/v), αντίστοιχα. Το εναιώρημα αφέθηκε για 30 min σε πάγο και στη συνέχεια φυγοκεντρήθηκε για 2 h σε 80.000 g (4°C), σχηματίζοντας ένα υπερκείμενο ίζημα, μια ενδιάμεση στιβάδα και ένα ίζημα. Το υπερκείμενο ίζημα, εμπλουτισμένο σε χλωροσώματα, αφαιρέθηκε, διαλυτοποιήθηκε σε 20 όγκους ρυθμιστικού διαλύματος απομόνωσης απουσία λυσοζύμης, έχοντας παράλληλα τελική συγκέντρωση Deriphat 0,16% (w/v) και διατηρήθηκε στους 4°C για 12- 18 h.

Στη συνέχεια 2 ml από το δείγμα αυτό τοποθετήθηκαν στην κορυφή μιας διαβάθμισης σακχαρόζης 3 βημάτων (2 ml 40% w/v, 3 ml 30% w/v και 3 ml 20% w/v) η οποία επίσης περιείχε 0.16% Deriphat και ακολούθησε φυγοκέντρωση για 18 h σε 200.000 g, 4°C. Το τελικό χλωροσωμικό κλάσμα εμφανίστηκε στο κλάσμα που περιείχε 20% ζάχαρη, στη συνέχεια συλλέχθηκε και αποθηκεύτηκε στους -80°C.

2.2 Αναλυτικές μέθοδοι

2.2.1 Φασματοσκοπία απορρόφησης ορατής ακτινοβολίας

Τα οπτικά φάσματα ελήφθησαν σε ένα φασματοφωτόμετρο Aminco dual wavelength DW 2000 uv-vis (SLM Aminco). Τα φάσματα απορρόφησης των δειγμάτων ελήφθησαν σε κυψελίδες διαμέτρου διέλευσης 1 cm σε θερμοκρασία δωματίου χρησιμοποιώντας φασματοφωτόμετρο Aminco DW-2000.

2.2.2 Στατική και δυναμική σκέδαση φωτός

Η διαδικασία σκέδασης του φωτός καθορίζει έναν κυματικό φορέα $q=(4\pi n/\lambda) \sin(\theta/2)$, όπου λ είναι το μήκος κύματος της προσκίπτουσας ακτινοβολίας υπό κενό, θ είναι η γωνία σκέδασης και n είναι ο δείκτης διάθλασης του μέσου. Στην μελέτη αυτή η πειραματική χρονική συνάρτηση αυτοσυσχέτισης full homodyne intensity $g(q,t)$ was μετρήθηκε σε θερμοκρασία δωματίου, χρησιμοποιώντας ένα εύρος γωγιών σκέδασης μεταξύ 15° και 150° (ALV-5000 multi-bit, multi-t full digital correlator που καλύπτει ένα ευρύ φάσμα περίπου 10 δεκάδων). Η δέσμη ενός solid state diode laser (Coherent Model 2020) που λειτουργούσε στα 532 nm με κάθετα πολωμένο φως εστιάστηκε στην κυψελίδα του δείγματος μέσω ενός θαλάμου ελεγχόμενης θερμοκρασίας (η θερμοκρασία

ελέχθηκε σε ένα εύρος $\pm 0.01^\circ\text{C}$) που περιείχε τολουένιο με ίδιο δείκτη διάθλασης. Οι προσπίπτουσες και σκεδαζόμενες δέσμες πολώθηκαν με Glan and Glan-Thompson πολωτές που έχουν μοριακούς συντελεστές απορρόφησης υψηλότερους από 10^{-6} και 10^{-7} αντίστοιχα. Τα διαλύματα των δειγμάτων διηθήθηκαν χρησιμοποιώντας φίλτρα $0.45\ \mu\text{m}$ (Millipore) και στη συνέχεια τοποθετήθηκαν απευθείας σε προ-καθαρισμένες υψηλής ποιότητας κυψελίδες σκέδασης φωτός [45-46].

2.2.3 Ηλεκτρονική μικροσκοπία

Για την αρνητική χρώση, μια σταγόνα περίπου 5 μl δείγματος προσροφήθηκε σε δισκίο χαλκού επικαλυπτόμενο με άνθρακα και χρυσό η αρνητική χρώση πραγματοποιήθηκε με 2% (w/v) οξικού ουρανυλίου. Τα δισκία αναλύθηκαν με CM 100 ηλεκτρονικό μικροσκόπιο διέλευσης (Philips, Eindhoven, The Netherlands) και με JEM 100 (JEOL, Japan). Η τάση στο μικροσκόπιο ήταν 80 kV.

2.3 Μέθοδοι Ανάλυσης και Χαρακτηρισμού Πρωτεϊνών

2.3.1 Tricine-SDS ηλεκτροφόρηση σε πηκτή πολυακρυλαμιδίου [47]

Η ηλεκτροφόρηση σε πηκτή SDS-Tricine πολυακρυλαμιδίου προσφέρει καλύτερη ανάλυση και διαχωριστική ικανότητα των πρωτεϊνών – ιδιαίτερα αυτών μικρού μοριακού βάρους (1-100 kDa) - σε σχέση με την πηκτή SDS πολυακρυλαμιδίου. Η πηκτή που χρησιμοποιήθηκε στην παρούσα πειραματική διαδικασία είχε συνολική συγκέντρωση ακρυλαμιδίου και δις-ακρυλαμιδίου (T) 10% και συγκέντρωση 3% σε cross-linker (C) δις-ακρυλαμίδιο, έτσι ώστε να επιτευχθεί καλύτερη ανάλυση στις πρωτεΐνες μοριακού βάρους 20-100 kDa.

Ρυθμιστικά και Διαλύματα

AB-mix (49.5%T, 3%C)

Acrylamide 48 gr/100 ml

Bisacrylamide 1.5 gr/100 ml

Gel Buffer (3x)

Tris 3M

SDS 0.3% (w/v)

pH=8.4 (ρύθμιση με HCl)

Sample Buffer

SDS 12% (w/v)

Glycerol 10% (w/v)

Mercaptoethanol 6%

Serva blue G 0.05%

Tris-HCl, pH=7.0 150 mM

Anode Buffer (10x)

Tris 1M

pH=8.9 (ρύθμιση με HCl)

Cathode Buffer (10x)

Tris 1M

Tricine 1M

SDS 1.0% (w/v)

pH=8.25

Προετοιμασία των Gel

Οι πηκτές SDS-Tricine πολυακρυλαμιδίου παρασκευάστηκαν σε πλάκες διαστάσεων 16 cm x 18 cm και πάχους 2 mm. Παρακάτω αναγράφεται η σύσταση των διαλυμάτων που χρησιμοποιήθηκαν για την παρασκευή των πηκτών:

Διάλυμα Πηκτής Διαχωρισμού (300ml)

60 ml AB-mix S

100ml	Gel buffer (3x)
30gr	Glycerol
H ₂ O (μέχρι τελικού όγκου 300ml)	
1.5ml	APS 10% (w/v)
150μl	TEMED

Διάλυμα Πηκτής Στοιβαξης (96ml)

8ml	AB-mix S
24ml	Gel buffer (3x)
H ₂ O (μέχρι τελικού όγκου 96 ml)	
720μl	APS 10% (w/v)
72μl	TEMED

Αρχικά τοποθετείται το διάλυμα πηκτής διαχωρισμού σε συσκευή η οποία μπορεί να παρασκευάσει 4 gel. Για να προκύψει ομαλό και ομοιόμορφο πακετάρισμα τοποθετείται πάνω από το διάλυμα πηκτής διαχωρισμού μικρή ποσότητα βουτανόλης, κορεσμένης με νερό. Μετά τον πολυμερισμό του ακρυλαμιδίου, απομακρύνεται η βουτανόλη και προστίθεται το διάλυμα πηκτής στοιβαξης. Αμέσως μετά προστίθενται χτενάκια και η πηκτή αφήνεται να πολυμεριστεί.

Πριν την ηλεκτροφόρηση τα δείγματα διαλυτοποιούνται σε ρυθμιστικό διάλυμα μετουσίωσης, σε αναλογία 1:1, και έπειτα θερμαίνονται στους 42⁰C για 20 min για να σπάσουν οι δεσμοί υδρογόνου και οι δισουλφιδικοί δεσμοί στις πρωτεΐνες ώστε να αποδιαταχθούν. Για να προσδιοριστούν τα μοριακά βάρη των πρωτεϊνών που

περιέχονται στο δείγμα ,τοποθετείται και μάρτυρας, ο οποίος περιέχει μίγμα πρωτεϊνών με γνωστά μοριακά βάρη.

Η ηλεκτροφόρηση πραγματοποιείται αρχικά υπό χαμηλή τάση (60V), έως ότου τα

δείγματα περάσουν από την πηκτική στοιβαξης στην πηκτική διαχωρισμού, και στη συνέχεια η τάση αυξάνεται στα 130 – 180 V για περίπου 8 h, μέχρι το μέτωπο του διαλύτη να φτάσει λίγο πριν το κατώτερο άκρο των πλακών ηλεκτροφόρησης. Η ηλεκτροφόρηση διαρκεί 8-11 h και πραγματοποιείται σε χαμηλές θερμοκρασίες.

Πρότυπες Πρωτεΐνες

a-lactalbumin from bovine serum	67 kDa
a-lactalbumin from chicken egg white	45 kDa
Glyceraldehydes-3-phosphate-dehydrogenase from rabbit muscle	36 kDa
Carbonic anhydrase from bovine erythrocytes	29 kDa
Trypsinogen from bovine pancreas	24 kDa
Trypsin inhibitor from soybean	20 kDa
a-lactalbumin from bovine milk	14.2 kDa

Χρόση Colloidal Coomassie [48]

Μετά το πέρας της ηλεκτροφόρησης, η πηκτική αφήνεται για περίπου 18 h σε διάλυμα:

Μεθανόλη	450 ml
Οξικό οξύ	10 ml
H ₂ O	540 ml

Έπειτα η πηκτή τοποθετείται στο παρακάτω διάλυμα για 12-18 h:

Θεικό Αμμώνιο	17% (w/v)
Μεθανόλη	34%
Οξικό οξύ	0.5%
Coomasie G250	0.1% (w/v)

Στη συνέχεια, η επιθυμητή χρωματική αντίθεση επιτυγχάνεται χρησιμοποιώντας απιονισμένο H₂O.

Όλα τα πρωτεϊνικά δείγματα υπέστησαν αραίωση 1 προς 3 σε ρυθμιστικό διάλυμα μετουσίωσης πριν την ηλεκτροφόρηση, εκτός από κάποια δείγματα περιπλάσματος, στα οποία πραγματοποιήθηκε εναλλακτικά κατεργασία με TCA (10%) και ακετόνη, για περαιτέρω συμπύκνωση. Αυτά τα δείγματα διαλυτοποιήθηκαν σε 100 μl ρυθμιστικού διαλύματος μετουσίωσης.

2.3.2 Μέθοδος Bradford [49]

Η μέθοδος Bradford χρησιμοποιείται για τον προσδιορισμό της συνολικής πρωτεϊνικής πυκνότητας σε ποικίλα δείγματα. Αρχικά κατασκευάζεται πρότυπη καμπύλη με γνωστές συγκεντρώσεις κάποιας δεδομένης πρωτεΐνης (BSA) ως προς την απορρόφηση στο ορατό (595 nm). Στη συνέχεια μετράται η απορρόφηση στα 595 nm του προς ανάλυση δείγματος και μέσω της καμπύλης υπολογίζεται σ' αυτό η συνολική

πρωτεϊνική πυκνότητα.

Κατασκευή Πρότυπης Καμπύλης

Αρχικά λαμβάνονται 0, 10, 20, 30, 40 και 50 μl από το πρότυπο διάλυμα αλβουμίνης από ορό βοδιού (BSA) και αραιώνονται με νερό μέχρι τελικού όγκου 50 μl. Έπειτα προστίθενται σε κάθε διάλυμα άλλα 50 μl ρυθμιστικού διαλύματος (π.χ. 20 mM Tris) και τέλος κάθε διάλυμα συμπληρώνεται με 900 μl αντιδραστηρίου Bradford. Ακολουθώς και μετά από ισχυρή ανάδευση, τα δείγματα αφήνονται σε ηρεμία για 5min και καταγράφεται η απορρόφηση τους στα 595 nm, στο μήκος κύματος δηλαδή όπου απορροφά το σύμπλοκο χρωστική – πρωτεΐνη, ενώ το πρώτο δείγμα χρησιμοποιείται ως τυφλό και η τιμή που λαμβάνεται από το φασματοφωτόμετρο για αυτό αφαιρείται από τις τιμές των υπόλοιπων δειγμάτων. Αφού κατασκευαστεί η καμπύλη αναφοράς

συναρτήσει της μάζας της BSA σε μg, παρασκευάζονται τα άγνωστα δείγματα με παρόμοιο τρόπο, μετράται η απορρόφηση τους στα 595 nm και ακολούθως υπολογίζεται η πρωτεϊνική τους πυκνότητα μέσω της ανωτέρω καμπύλης.

Για τα συγκεκριμένα πειράματα η πρότυπη καμπύλη ήταν της μορφής

$$Y = a X \text{ με } a = 0.0524$$

όπου Y η συγκέντρωση και X η απορρόφηση.

Κατά τη διαδικασία ποσοτικού προσδιορισμού πρωτεϊνών με τη μέθοδο Bradford πρέπει να δοθεί ιδιαίτερη προσοχή ώστε το διάλυμα Bradford να φυλάσσεται σε σκοτεινό μέρος, η μέτρηση της απορρόφησης να έχει πραγματοποιηθεί μέσα σε μία ώρα και όλα τα πρωτεϊνικά δείγματα να βρίσκονται σε χαμηλή θερμοκρασία.

Η συγκέντρωση του δείγματος υπολογιζόταν μέσω του νόμου Beer-Lambert.

Έτσι ισχύει:

$$A = \varepsilon \cdot b \cdot c$$

Όπου: A = απορρόφηση

ε = μοριακός συντελεστής απορρόφησης, (για το κυτόχρωμα c είναι 17)

C = η συγκέντρωση σε mM

b = η οπτική διαδρομή της κυψελίδας (1cm)

2.3.3 Συμπύκνωση - καταβύθιση πρωτεϊνών

Στο περιπλασματικό εκχύλισμα προστέθηκε υποδεκαπλάσιος όγκος υδατικού διαλύματος TCA 100% και μετά από ανάδευση σε vortex το μίγμα τοποθετήθηκε στους

4⁰C για 16 h. Στη συνέχεια το δείγμα φυγοκεντρήθηκε σε 10.000 g στους 4⁰C για 15 min. Το υπερκείμενο απομακρύνθηκε και προστέθηκε ίση ποσότητα παγωμένης ακετόνης με το αρχικό πρωτεϊνικό δείγμα για να ξεπλυθεί το ίζημα. Το μίγμα αναδεύτηκε σε vortex και έπειτα φυγοκεντρήθηκε όπως αναφέρθηκε ανωτέρω. Ακολούθως, το υπερκείμενο που περιείχε την ακετόνη απομακρύνθηκε και η διαδικασία έκπλυσης έλαβε χώρα συνολικά τρεις φορές. Τελικά, το ίζημα μετά την απομάκρυνση του υπερκείμενου, αφέθηκε σε πάγο για να στεγνώσει

2.3.5 N-τερματική ανάλυση COFRADIC [38-40]

Στερεή Gu.HCl προστέθηκε στα δείγματα τελικής συγκέντρωσης περίπου 4 M και οι κυστεΐνες ανήχθησαν και αλκυλιώθηκαν για 60 min στους 37 ⁰C με προσθήκη Tris (2

carboxyethyl)phosphine και ιωδοακεταμιδίου σε τελικές συγκεντρώσεις 3 και 6 M, αντίστοιχα. Η περίσσεια των αντιδραστηρίων αφαιρέθηκε με αφαλάτωση σε μια στήλη PD-10 (Amersham Biosciences, Uppsala, Sweden) και οι πρωτεΐνες εκλούσθηκαν σε 2.5 ml Gu.HCl (2 M) σε 50 mM φωσφορικού νατρίου (pH 8.0). Οι πρωτοταγείς ελεύθερες αμίνες τριδευτέρο-ακετυλιώθηκαν για 2 h στους 37 °C με προσθήκη *N*-hydroxysuccinimide trideutero-acetate σε τελική συγκέντρωση 10 mM. Η Ο-ακετυλίωση αντεστράφη προσθέτοντας υδροξυλαμίνη σε τελική συγκέντρωση 40 mM και με incubation για 10 min στους 37 °C. Το αντιδραστήριο της ακετυλίωσης απενεργοποιήθηκε περαιτέρω για 10 min στους 37 °C με γλυκίνη (τελική συγκέντρωση 20 mM). Η περίσσεια των αντιδραστηρίων αφαιρέθηκε με αφαλάτωση με στήλη PD-10 και έλαβε χώρα έκλυση των πρωτεϊνών με 2.5 ml διαλύματος 20 mM NH₄HCO₃ (pH 7.9).

Οι πρωτεϊνικές συγκεντρώσεις μετρήθηκαν με χρήση του BioRad's Protein Assay (BioRad Laboratories, Munich, Germany) σύμφωνα με τις οδηγίες του κατασκευαστή. Μετά απο θέρμανση των πρωτεϊνών για 5 min στους 95 °C, ακολούθησε τοποθέτηση τους σε πάγο για 5 min και στη συνέχεια ολονύχτια θρυψινοποίηση στους 37 °C (sequencing-grade, modified trypsin, Promega, Madison, WI, USA). Η αναλογία ενζύμου-υποστρώματος ήταν 1:50 w/w.

Ακολούθως, το προκύπτον μίγμα πεπτιδίων (600 mg) στεγνώθηκε σε speedVac και επαναδιαλυτοποιήθηκε σε 1 ml διαλύματος 0.08% TFA σε 50% HPLC-graded ACN (J.T. Baker, Deventer, The Netherlands), ώστε το τελικό διάλυμα να έχει pH 3. Στη συνέχεια το δείγμα φορτώθηκε σε μια κατιοντική στήλη AccuBONDII SCX SPE cartridge (Agilent Technologies, Waldbronn, Germany), ισορροπήθηκε με 10 mM φωσφορικού νατρίου σε ρυθμιστικό διάλυμα 50% ACN (pH 3) χρησιμοποιώντας Agilent vacuum manifold. Μετά την φόρτωση του δείγματος ακολούθησε ξέπλυμα της στήλης με 3 ml διαλύματος φωσφορικού νατρίου σε 50% ACN (pH 3). Τα πεπτίδια που ανακτήθηκαν στο κλάσμα που δεν προσδέθηκε, συλλέχθηκαν σε συνολικό όγκο 4 ml και ακολούθησε ξήρανση υπό κενό και επαναδιαλυτοποίηση σε 30 ml διαλύματος 0.5 M NaOH (με ρύθμιση τελικού pH 7.5). Στο δείγμα προστέθηκε μίγμα

pGAPάσης/Qκυκλάσης και η ενζυμική αντίδραση συνεχίστηκε για 1 h στους 37 °C.

Διάλυμα pGAPάσης (25 ml, 25 U/ml) (από το TAGZyme kit, Qiagen, Hilden, Germany) ενεργοποιήθηκε για 10 min στους 37 °C προσθέτοντας 10 ml διαλύματος 200 mM φωσφορικού νατρίου (pH 7.5), 10 ml διαλύματος 800 mM NaCl and 10 ml διαλύματος 50 mM κυστεαμίνης-HCl και έπειτα προστέθηκε διάλυμα (25 ml, 50 U/ ml) Qκυκλάσης (TAGZyme kit).

Το δείγμα που προέκυψε ήταν εμπλουτισμένο σε N-τερματικά πεπτιδία και είχε υποστεί κατεργασία για αφαίρεση των πυρολιδονικών καρβοξυλικών οξέων, Στη συνέχεια το δείγμα αυτό οξινοποιήθηκε με 5 ml οξικού οξέος. Προς αποφυγήν οξειδωσης των αμινοξέων μεθειονίνης μεταξύ των δύο RP-HPLC (που θα οδηγούσε σε μετατόπιση των πεπτιδίων αυτών λόγω σχηματισμού σουλφοξειδίου), οι μεθειονίνες οξειδώθηκαν πριν την πρώτη RP-HPLC ανάλυση για 30 min στους 37 °C προσθέτοντας 2 ml διαλύματος 30% w/v H₂O₂ (τελική συγκέντρωση 0.06%). Το μίγμα των πεπτιδίων εισήχθη στην RP-στήλη (Zorbax-300SB-C18 Narrowbore, 2.1 mm (id) 6150 mm μήκος, 5 mm σωματίδια, Agilent) για την πρώτη N-τερματική COFRADIC ανάλυση. Μετά απο διέλευση του διαλύτη A (10 mM οξικού αμμωνίου σε νερό/ACN (98:2 v/v), pH 5.5) για 10 min, εκκίνησε μια διαβάθμιση με 1% αύξηση διαλύτη B ανά λεπτό (ο διαλύτης B αποτελείται απο 10 mM οξικό αμμώνιο σε ACN/ H₂O (70:30 v/v), pH 5.5). Στη συνέχεια η στήλη έτρεξε με 100% διαλύτη B για 5 min, έπειτα με 100% διαλύτη A και ισορροπήθηκε για 20 min. Η ροή κρατήθηκε σταθερή στα 80 µl/min χρησιμοποιώντας μια capillary αντλία της Agilent's 1100 series με ρυθμιστή ροής 100 µL/min. Συλλέχθηκαν κλάσματα ανά 4 min μεταξύ 20 και 40 min μετά την εισαγωγή του δείγματος.

Τα κλάσματα αυτά ξηράθηκαν και επαναδιαλυτοποιήθηκαν σε 50 ml διαλύματος βορικού νατρίου 50 mM (pH 9.5). Διάλυμα TNBS (10 ml, 15 mM) (Fluka BioChemika, Buchs, Switzerland) σε 50 mM βορικού νατρίου (pH 9.5) προστέθηκε και αντέδρασε με τις ελεύθερες αμίνες για 1 h στους 37 °C. Το τελευταίο αυτό βήμα επαναλήφθηκε 4 φορές χωρίς ενδιάμεση ξήρανση των δειγμάτων, και στη συνέχεια τα προκύπτοντα τροποποιημένα κλάσματα οξινοποιήθηκαν με 5 ml συμπυκνωμένου οξικού οξέος.

Έπειτα, επαναφορτώθηκαν στην ίδια RP- κολώνα και διαχωρίστηκαν με την ίδια

διαβάθμιση όπως και στην πρώτη RP-HPLC ανάλυση. Τα προκύπτοντα κλάσματα, που ήταν αμπλουτισμένα σε N-τερματικά πεπτίδια συλλέχθηκαν στο ίδιο χρονικό πλαίσιο με την πρώτη ανάλυση, αλλά τώρα πραγματοποιήθηκαν 8 συλλογές κλασμάτων 30 sec έκαστη, με την επιπρόσθετη συλλογή 2 κλασμάτων, ενός 4 min πριν και ενός 4 min μετά το χρονικό πλαίσιο της πρώτης ανάλυσης λαμβάνοντας υπόψη φαινόμενα peak broadening κατά τη δεύτερη χρωματογραφική ανάλυση.

Τα συλλεγμένα κλάσματα ξηράθηκαν και επαναδιαλυτοποιήθηκαν σε 20 ml H₂O/CAN (98.2 v/v) και το μισό κάθε μίγματος αναλυτή αναλύθηκε με LC-MS/MS χρησιμοποιώντας μια microfluid διεπιφάνεια (Agilent's Chip Cube) σε έναν Agilent XCT-Ultra IT φασματογράφο μάζας.

Τα δημιουργηθέντα αρχεία Mascot αναλύθηκαν στην NCBI βάση δεδομένων (version 17 December 2007). Ως πρωτεολυτικό ένζυμο επελέγη η θρυψίνη με μέγιστο αριθμό αποτυχημένων πρωτεολύσεων 1. Τα αρχικά αρχεία του Mascot επεξεργάστηκαν περαιτέρω με λογισμικό και μόνο τα φάσματα MS/MS που υπερέβαιναν το όριο ταυτοποίησης σε 95% επίπεδο εμπιστοσύνης κρατήθηκαν. Τα φάσματα που έλαβαν χαμηλό σκόρ στο Mascot (10 ή λιγότερες μονάδες πάνω από το όριο) επαναθεωρήθηκαν και μόνο αυτά που διέθεταν ένα σημαντικό αριθμό από ιόντα πεπτιδίων έγιναν δεκτά ως ταυτοποιημένα.

2.3.4 Φασματομετρία μάζας και ταυτοποίηση πρωτεϊνών χλωροσωμικού κλάσματος

Το χλωροσωμικό κλάσμα του *Chlorobium tepidum* επαναδιαλυτοποιήθηκε σε 100 μl διαλύματος 4 M guanidinium hydrochloride (Gu.HCl) σε 50 mM NH₄HCO₃ (pH περίπου 8). Μετά την διαλυτοποίηση με 50 mM NH₄HCO₃ μέχρι τελικής συγκεντρώσεως Gu.HCl 500 mM, μετρήθηκε η πρωτεϊνική συγκέντρωση

χρησιμοποιώντας Bio-Rad's Protein Assay (Bio-Rad laboratories, Munich, Germany) σύμφωνα με τις οδηγίες του κατασκευαστή. Μετά από θέρμανση του δείγματος στους 95°C για 5 min, ακολουθούμενη από τοποθέτηση του σε πάγο για 10 min, οι πρωτεΐνες υπέστησαν ολονύκτια πρωτεόλυση με θρυψίνη (sequencing-grade, modified, Promega Corporation, Madison, WI, USA) στους 37°C. Η αναλογία ενζύμου-υποστρώματος ήταν 1:50 (w/w).

Το μίγμα των πεπτιδίων που αντιστοιχούσε σε 70 μg πρωτεολυμένων πρωτεϊνών, ξηράνθηκε, επαναδιαλυτοποιήθηκε σε 100 μl διαλύματος οξικού οξέος 1% και κλασματοποιήθηκε με RP-HPLC (η RP στήλη που χρησιμοποιήθηκε ήταν Zorbax-300SB-C18 Narrow-bore, 2.1 mm (I.D.) x 150 mm μήκος, 5 μm σωματίδια, Agilent). Μετά από 10 min ισοκρατικής άντλησης με διαλύτη A (10 mM οξικού αμμωνίου σε H₂O/ακετονιτρίλιο (98:2, v/v), pH 5.5), δημιουργήθηκε μια διαβάθμιση ξεκινώντας από 1% αύξηση διαλύτη B ανά λεπτό (ο διαλύτης B είναι 10 mM οξικού αμμωνίου σε ακετονιτρίλιο/H₂O (70/30, v/v), pH 5.5). Η κολώνα τελικά πλύθηκε με τον διαλύτη B για 5 min και στη συνέχεια ισορροπήθηκε για 20 min με τον διαλύτη A. Κατά την RP-HPLC ανάλυση, η ροή διατηρήθηκε σταθερή στα 80 μl/min χρησιμοποιώντας την αντλία της Agilent's 1100 με ελεγκτή ροής 100 μl/min. Τα λήφθηκαν είχαν εύρος 1 min και συλλέχθηκαν μεταξύ των 20 και 80 min μετά την εισαγωγή του δείγματος στη

στήλη. Τα προκύπτοντα κλάσματα που είχαν διαφορά έκλουσης 20 min, συλλέχθηκαν, ξηράνθηκαν και επαναδιαλυτοποιήθηκαν σε 20 μl H₂O/ακετονιτρίλιο (98:2, v/v). Το ήμισυ κάθε μίγμα αναλυτή αναλύθηκε με n-LC-MS/MS χρησιμοποιώντας Agilent XCT-Ultra ion trap MS[50]

Τα δημιουργηθέντα αρχεία Mascot αναλύθηκαν στην NCBI βάση δεδομένων (version 17 December 2007) χρησιμοποιώντας και τον DBToolKit αλγόριθμο. Οι τελευταίες αυτές βάσεις δεδομένων χρησιμοποιήθηκαν καθότι συνυπολογίζουν αποτελεσματικά την απομάκρυνση των N-τερματικών σηματοδοτικών αλληλουχιών και συνεπώς αυξάνουν την συνολική αποτελεσματικότητα της ταυτοποίησης στα MS/MS φάσματα. Τέθηκαν οι παρακάτω παράγοντες αναζήτησης στο Mascot. Οι ανοχές πεπτιδικής μάζας για το πρόδρομο ιόν και το fragment ion ρυθμίστηκαν σε ± 0.5 Da, το ESI-IT επιλέχθη για τους κανόνες θραυσματοποίησης των πεπτιδίων, το φορτίο των πεπτιδίων ρυθμίστηκε σε 1+,2+,3+, ενώ δεν επιλέχθηκε καμία fixed τροποποίηση. Ως variable τροποποιήσεις επελέχθησαν η οξείδωση της μεθειονίνης στο σουλφοξειδικό της ανάλογο, ο σχηματισμός πυρογλουταμικού (N-τερματική Gln) και η αποαμίνωση (στην περίπτωση των Gln και Asn). Η επιλογή του Mascot C13 ρυθμίστηκε στο 1. Ως πρωτεολυτικό ένζυμο επελέγη η θρυψίνη με μέγιστο αριθμό αποτυχημένων πρωτεολύσεων 1. Τα αρχικά αρχεία του Mascot επεξεργάστηκαν περαιτέρω με λογισμικό και μόνο τα φάσματα MS/MS που υπερέβαιναν το όριο ταυτοποίησης σε 95% επίπεδο εμπιστοσύνης κρατήθηκαν. Τα φάσματα που έλαβαν χαμηλό σκόρ στο Mascot (10 ή λιγότερες μονάδες πάνω από το όριο) επαναθεωρήθηκαν και μόνο αυτά που διέθεταν ένα σημαντικό αριθμό από ιόντα πεπτιδίων έγιναν δεκτά ως ταυτοποιημένα. Όλα τα ταυτοποιημένα πεπτίδια αποθηκεύτηκαν σε μία MySQL βάση δεδομένων (http://genesis.ugent.be/ms_lims/) στην οποία δημιουργήθηκαν σύνδεσμοι για τα MS/MS φάσματα και τις πρόδρομες πρωτεΐνες τους.

2.4 Επεξεργασία των αποτελεσμάτων της πρωτεομικής ανάλυσης με μεθόδους βιοπληροφορικής

Οι ταυτοποιημένες πρωτεΐνες που προέκυψαν από κάθε μέθοδο που χρησιμοποιήθηκε επεξεργάστηκαν με διάφορα προγράμματα που είναι διαθέσιμα στο δίκτυο (<http://expasy.org>), για τον προσδιορισμό διαφόρων ιδιοτήτων τους. Η θέση και λειτουργία των ταυτοποιημένων πρωτεϊνών στο βακτήριο προβλέφθηκε με τα προγράμματα PSORT [51-52] και SignalIP [53-54]. Το ισοηλεκτρικό σημείο, το μοριακό βάρος καθώς και η τιμή GRAVY κάθε πρωτεΐνης προσδιορίστηκαν με το πρόγραμμα ProtParam (<http://expasy.org>).

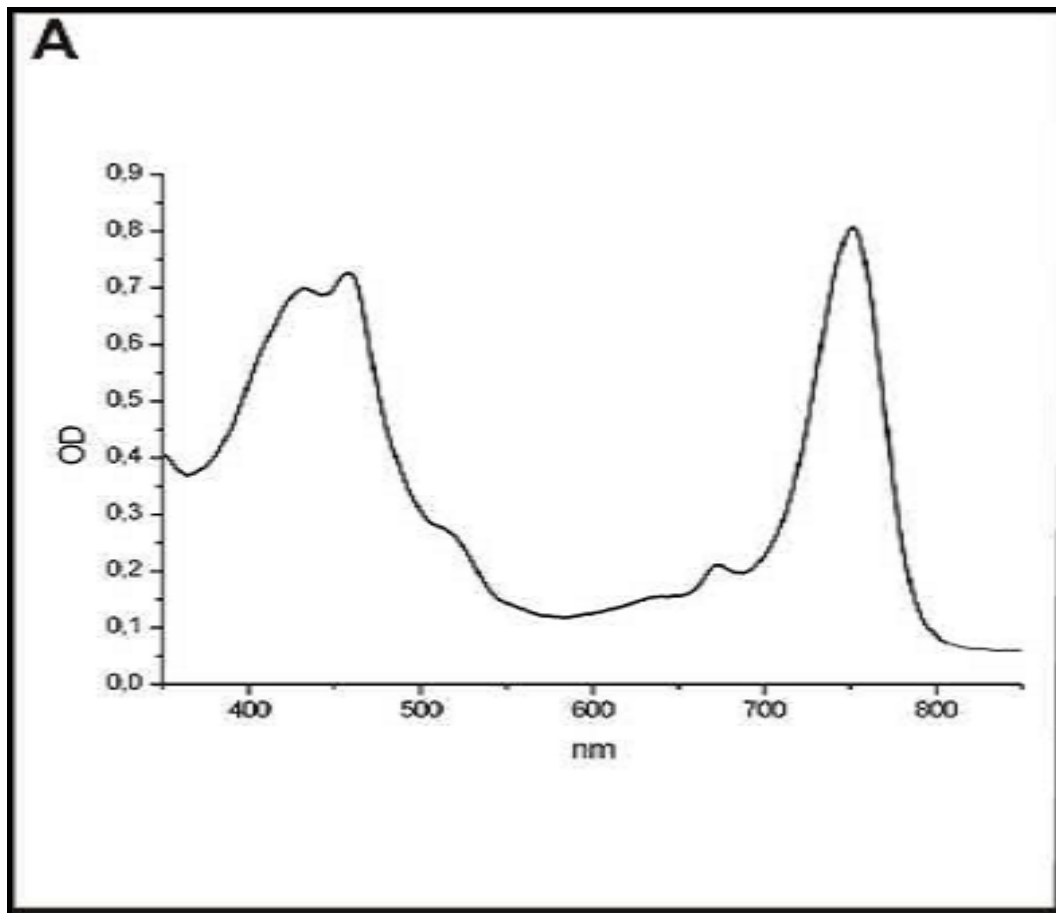
3. ΑΠΟΤΕΛΕΣΜΑΤΑ

3.1 Αναλυτικές μέθοδοι

3.1.1 Φασματοσκοπία απορρόφησης ορατής ακτινοβολίας

Το *Chl. tepidum*, όντας φωτοσυνθετικό βακτήριο, περιέχει μεγάλο αριθμό χρωστικών (BChls και καροτενοειδή) που είτε βρίσκονται ελεύθερες είτε προσδεμένες σε πρωτεΐνες. Το γεγονός αυτό καθιστά δυνατή την μελέτη του βακτηρίου αλλά και των επιμέρους συστατικών του (μεμβράνες, κυτταρόπλασμα, πρωτεΐνες κλπ) με φασματοσκοπία απορρόφησης ορατού.

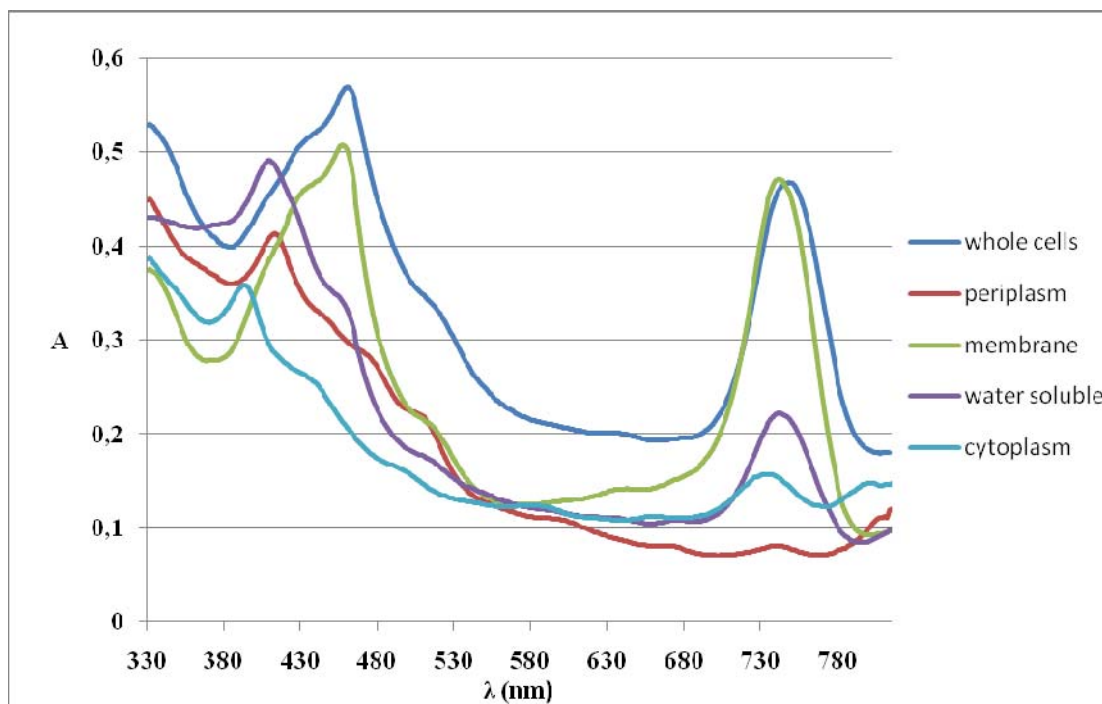
Στο Σχήμα 1 εμφανίζεται το φάσμα απορρόφησης των απομονωμένων χλωροσωμάτων του *Chl. tepidum*. Τα χλωροσώματα επιδεικνύουν μια Qy ταινία χαρακτηριστική για την συσσωματωμένη BChl *c*, με κορυφές μεταξύ 742 και 747 nm, και μια ταινία Soret στα 460 nm. Ο ώμος που εμφανίζεται περίπου στα 515 nm αποδίδεται σε καροτενοειδή.



Σχήμα 1: Φάσμα απορρόφησης των απομονωμένων χλωροσωμάτων.

Στο Σχήμα 2 εμφανίζονται τα φάσματα απορρόφησης των λοιπών κλασμάτων που απομονώθηκαν από το *Chl. tepidum* καθώς και το φάσμα των κυττάρων του βακτηρίου.

Το φάσμα των κυττάρων του *Chl. tepidum* εμφανίζει μια έντονη διπλή ταινία απορρόφησης στα 400-550 nm που αποδίδεται στην ύπαρξη καροτενοειδών και βακτηριοχλωροφυλλών a και c, καθώς και μια ταινία λίγο χαμηλότερης έντασης στα 720-780 nm που οφείλεται στη βακτηριοχλωροφύλλη c. Το ίδιο μοτίβο εμφανίζεται και στο φάσμα του μεμβρανικού κλάσματος.



Σχήμα 2: Φάσματα απορρόφησης των απομονωμένων κλασμάτων του *Chl. tepidum* (κύτταρα βακτηρίου, περίπλασμα, μεμβράνες, υδατοδιαλυτό κλάσμα και κυτταρόπλασμα).

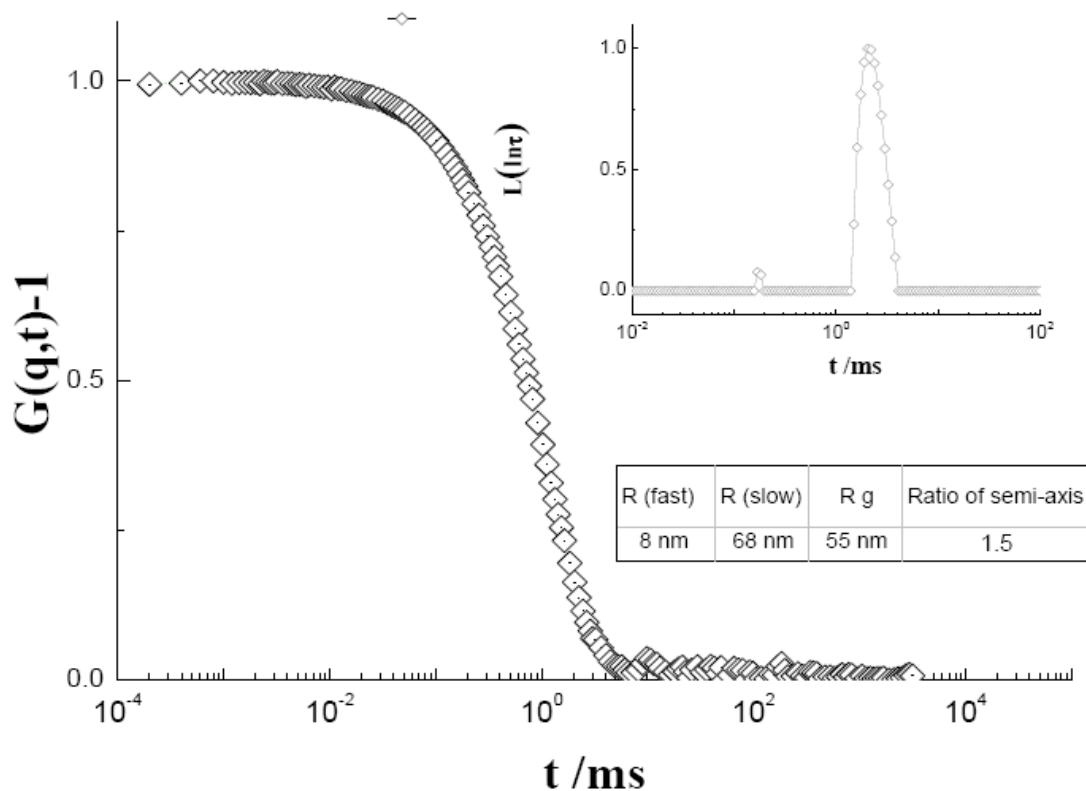
Αναφορικά με το φάσμα του υδατοδιαλυτού κλάσματος, παρατηρούνται δύο ταινίες απορρόφησης, μια στα 380-550 nm και μία στα 720-780 nm. Σε σύγκριση με τα δύο προαναφερόμενα κλάσματα παρατηρείται μείωση στην ένταση της μπάντας στην

περιοχή των 400-480 nm, οδηγώντας στην εμφάνιση ενός ώμου στα 380-400 nm. Παράλληλα, παρατηρείται σημαντική μείωση στην ένταση της κορυφής στα 720-780 nm. Το μοτίβο αυτό εμφανίζεται και στο φάσμα απορρόφησης του κυτταροπλασματικού κλάσματος, ενώ στο περιπλασματικό κλάσμα, δεν εμφανίζεται η ταινία απορρόφησης στα 720-780 nm που οφείλεται στην βακτηριοχλωροφύλλη c.

3.1.2 Δυναμική και Στατική σκέδαση φωτός[45-46]

Προκειμένου να ληφθούν πληροφορίες για την ομοιογένεια, το μέγεθος και τη μορφή των απομονωμένων οργανιδίων, το δείγμα που απομονώθηκε από τη διαβάθμιση σακχαρόζης μελετήθηκε σε υδατικό περιβάλλον. Η ομοιογένεια και ο ρόλος του ρυθμιστικού μελετήθηκε με δυναμική σκέδαση φωτός (DLS).

Στην Εικόνα 9 παρουσιάζονται η πειραματική χρονική συνάρτηση αυτοσυσχέτισης για τα απομονωμένα χλωροσώματα σε γωνία σκέδασης 45° καθώς και η κατανομή των πειραματικών συναρτήσεων αυτοσυσχέτισης χρησιμοποιώντας την ανάλυση REPES. Η ανάλυση αυτή εμφάνισε δύο δυναμικές διαδικασίες χαλάρωσης, μια γρήγορη και μία αργή. Η γρήγορη διαδικασία αποδίδεται σε πληθυσμούς μικρού μεγέθους και στην προκειμένη περίπτωση του ρυθμιστικού διαλύματος καθώς οι μικροί πληθυσμοί εμφανίζονται στο διάλυμα απουσία των χλωροσωμάτων. Η αργή διαδικασία αποδίδεται σε μεγαλύτερους πληθυσμούς, δηλαδή εν προκειμένω στα χλωροσώματα. Και οι δύο διαδικασίες σχετίζονται με φαινόμενα διάχυσης, καθώς οι ρυθμοί χαλάρωσης $\Gamma(\Gamma=1/\tau)$ είναι γραμμικά εξαρτημένοι από το τετράγωνο του διανύσματος σκέδασης (q^2). Με βάση τους χαρακτηριστικούς ρυθμούς χαλάρωσης, μπορεί να ληφθεί ο συντελεστής διάχυσης D ($D=\Gamma/q^2$), και συνεπώς η αντίστοιχη υδροδυναμική ακτίνα R_h χρησιμοποιώντας την σχέση Stokes–Einstein για τη σφαίρα.



Εικόνα 9: Πειραματικές χρονικές συναρτήσεις αυτοσυσχέτισης για τα απομονωμένα χλωροσώματα και οι κατανομές των πειραματικών συναρτήσεων αυτοσυσχέτισης, χρησιμοποιώντας ανάλυση REPES.

Τόσο η υδροδυναμική, όσο και η γυροσκοπική ακτίνα ενός μακρομορίου δίνουν πληροφορίες σε μια διάσταση και αποτελούν χρονική αποτύπωση της συμπεριφοράς του. Επομένως, τα μακρομόρια δεν είναι δυνατόν να περιγραφούν επαρκώς μόνο από τη γνώση μίας εκ των δύο ακτινών. Ο συνδυασμός τους όμως μπορεί να μας δώσει στοιχεία για τη διαμόρφωση του συσσωματώματος. Μέσω των μετρήσεων στατικής σκέδασης

υπολογίζεται η ένταση του σκεδαζόμενου φωτός. Το αντίστροφο της σκεδαζόμενης έντασης ως συνάρτηση του q^2 μας παρέχει τη δυνατότητα υπολογισμού της γυροσκοπικής ακτίνας και της μοριακής μάζας των υπό μελέτη σωματιδίων. Πιο συγκεκριμένα, ο προσδιορισμός της διαμόρφωσης κάθε μακρομορίου είναι εφικτός, αν υπολογισθεί το πηλίκο το οποίο συμβολίζεται με ρ και καλείται συντελεστής σχήματος ή ασύμμετρος συντελεστής. Οι τιμές του ρ είναι δυνατό να μας παράσχουν πληροφορίες σχετικά με την διαμόρφωση του συσσωματώματος, καθώς τιμές του ρ κοντά στο 0,78

αντιστοιχούν σε ομογενείς σφαίρες, ενώ τιμές μεγαλύτερες του 2 αντιστοιχούν σε πιο εκτεταμένες μορφές (επιμήκη ελλειψοειδή).

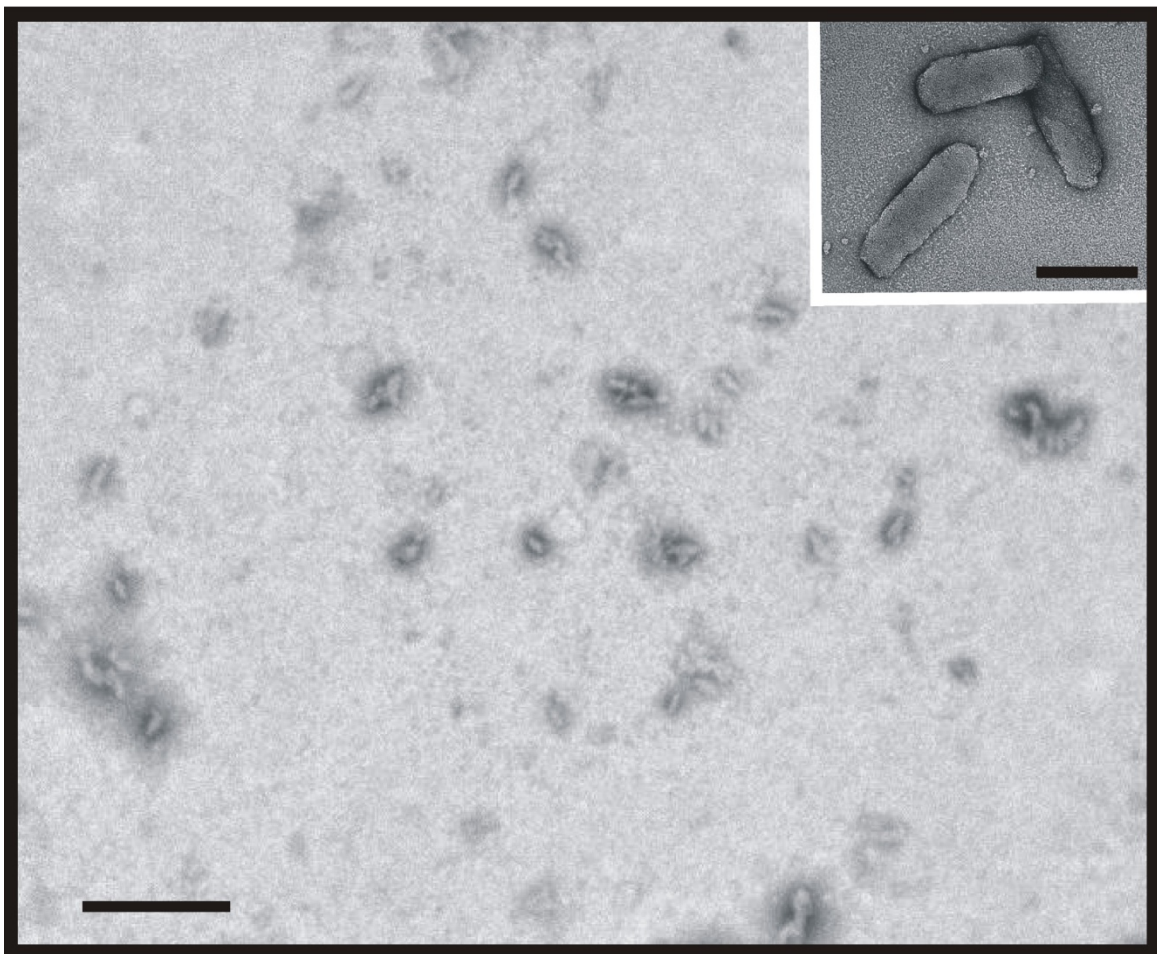
Η τιμή $R_h=68$ nm για την αργή διαδικασία χαλάρωσης, σε συνδυασμό με την τιμή της γυροσκοπικής ακτίνας R_g από τα πειράματα στατικής σκέδασης φωτός (static light scattering) οδηγεί στον υπολογισμό του λόγου R_g/R_h που είναι 0.81, γεγονός που υποδεικνύει ότι το σχήμα των χλωροσωμάτων είναι επιμηκυμένο και διαθέτει λόγο semi-axis 1.5. Με βάση την αναλογία αυτή, μπορούν να εκτιμηθούν διαστάσεις 136 x 90 nm για τα απομονωμένα χλωροσώματα.

3.1.3 Ηλεκτρονική μικροσκοπία

Το ηλεκτρονικό μικροσκόπιο είναι ένα επιστημονικό όργανο το οποίο χρησιμοποιεί μία υψηλής ενέργειας δέσμη ηλεκτρονίων για την εξέταση αντικειμένων σε μικροσκοπική κλίμακα. Η εξέταση αντικειμένων με ηλεκτρονικό μικροσκόπιο παρέχει πληροφορίες για την τοπογραφία, μορφολογία, σύσταση καθώς και κρυσταλλογραφικές πληροφορίες.

Προκειμένου να ληφθούν πληροφορίες υψηλής ευκρίνειας, τα χλωροσώματα μετά από αρνητική χρώση με οξικό ουρανύλιο απεικονίστηκαν με χρήση EM. Στην Εικόνα 10 φαίνονται χαρακτηριστικές απεικονίσεις των χλωροσωμάτων με ηλεκτρονικό

μικροσκόπιο, όπου εμφανίζονται ως ελλειψοειδείς δομές με ανομοιόμορφες άκρες, διαστάσεων 170 x 70 nm. Οι διαστάσεις αυτές που ελήφθησαν με EM εμφανίζουν τα χλωροσώματα αρκετά πιο επιμηκυμένα από τις αντίστοιχες μετρήσεις DLS.

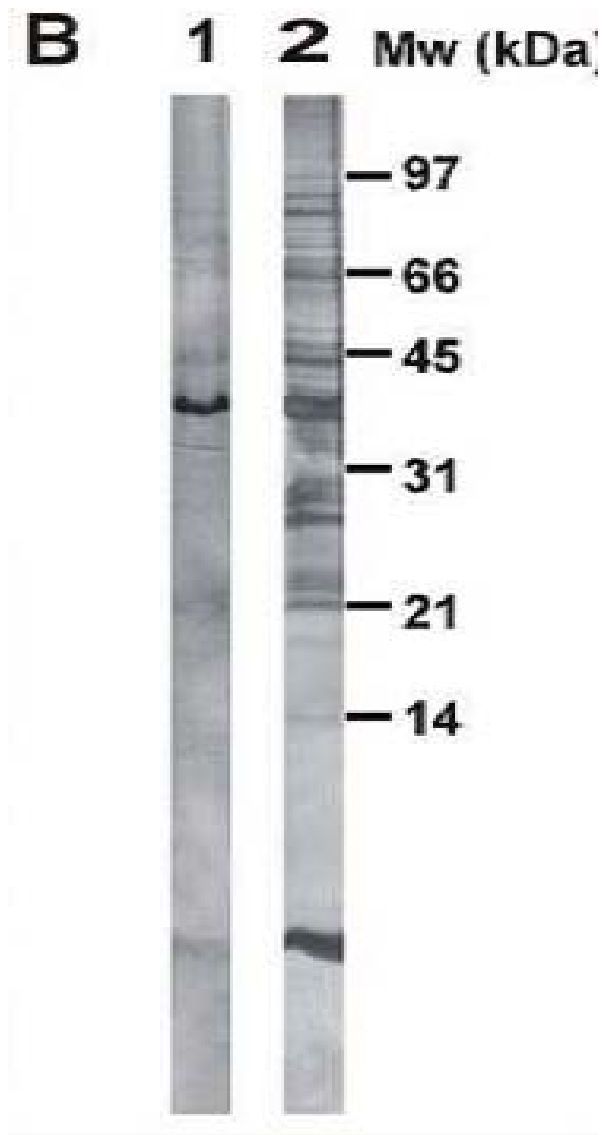


Εικόνα 10: Απεικόνιση με ηλεκτρονική μικροσκοπία των απομονωμένων χλωροσωμάτων του *Chl. tepidum*. Η μπάρα στην εικόνα αντιστοιχεί σε 1000 nm και στην ένθετη εικόνα σε 100 nm.

3.2. Μέθοδοι Ανάλυσης και Χαρακτηρισμού Πρωτεϊνών

3.2.1 Tricine-SDS ηλεκτροφόρηση σε πηκτή πολυακρυλαμιδίου

Στην Εικόνα 11 παρουσιάζεται το καταβυθισμένο με TCA κλάσμα των χλωροσωμάτων μετά από Tricine SDS-PAGE ανάλυση και χρώση με κολλοειδή Coomassie blue. Οι κύριες πρωτεΐνες των χλωροσωμάτων του *Chl. tepidum* που εμφανίζουν μοριακά βάρη μικρότερα των 30 kDa προέρχονται από τις πρωτεΐνες CsmA, CsmB/E/F, CmcD, CsmC, CmsH, CmsK, CmsI, CmsJ και CmsX σύμφωνα με βιβλιογραφικά δεδομένα [14-15]. Κάποιες από τις πρωτεΐνες αυτές παρατηρήθηκαν μόνο μετά από χρώση αργύρου που έχει μεγαλύτερη ευαισθησία. Επίσης, εμφανίζονται έντονες πρωτεϊνικές μπάντες στην περιοχή της πηκτής που αντιστοιχεί σε μοριακά βάρη 30–65 kDa.

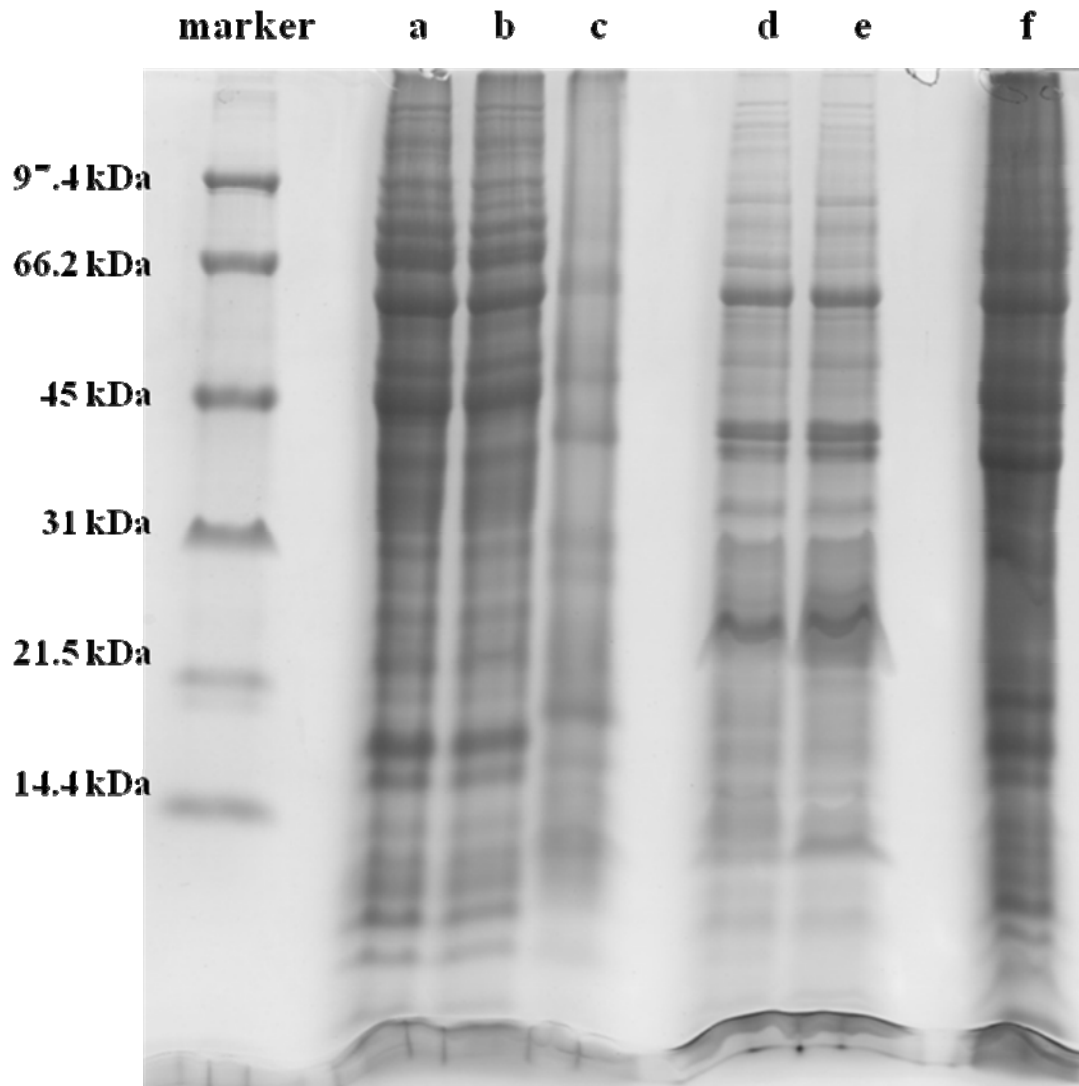


Εικόνα 11: Tricine SDS-PAGE απομονωμένων χλωροσωμάτων: 1. χρώση πρωτεϊνών με Coomassie blue, 2. χρώση αργύρου

Επιπρόσθετα πραγματοποιήθηκε ηλεκτροφόρηση σε πηκτή ακρυλαμιδίου SDS-Tricine του λοιπών πρωτεϊνικών κλασμάτων που απομονώθηκαν στην παρούσα μελέτη, καθώς και του δείγματος κυττάρων του *Chl. tepidum* (Εικόνα 12).

Στην περίπτωση του δείγματος των κυττάρων του βακτηρίου εμφανίζεται ένας πολύ μεγάλος αριθμός από πρωτεϊνικές μπάντες, κάνοντας τη διάκριση μεταξύ τους δύσκολη,

γεγονός αναμενόμενο δεδομένης της πολυπλοκότητας του δείγματος. Είναι εμφανής συνεπώς η αναγκαιότητα περαιτέρω κλασματοποίησης του πρωτεϊνικού δείγματος.



Εικόνα 12: Tricine-SDS-PAGE των πρωτεϊνικών κλασμάτων: α)υδατοδιαλυτό, β)κυτταροπλασματικό, γ) περιπλασματικό, δ) μεμβranικό μετά την απομάκρυνση του υδατοδιαλυτού, ε) μεμβranικό μετά την απομάκρυνση περιπλασματικού και κυτταροπλασματικού, ς)κύτταρα *Chl. tepidum*

Τα πρωτεϊνικά προφίλ του υδατοδιαλυτού και του κυτταροπλασματικού κλάσματος

εμφανίζουν μεγάλη ομοιότητα στην ανάλυση τους με Tricine-SDS-PAGE, με διαφοροποιήσεις κυρίως στην ένταση των πρωτεϊνικών μπαντών, γεγονός που επαληθεύει την πολυπλοκότητα των δειγμάτων αυτών, αλλά και ότι εμφανίζουν σημαντικές ομοιότητες στην πρωτεϊνική τους σύσταση.

Αναφορικά με το ηλεκτροφορητικό μοτίβο του περιπλασματικού κλάσματος, εμφανίζεται ένα πολύ πιο απλουστευμένο σε σχέση με το υδατοδιαλυτό κλάσμα πρωτεϊνικό προφίλ, γεγονός που επαληθεύει την επίτευξη της κλασματοποίησης. Επίσης παρατηρείται διαφοροποίηση κυρίως στην ένταση κάποιων από τις εμφανιζόμενες πρωτεϊνικές μπάντες σε σχέση με το υδατοδιαλυτό και το κυτταροπλασματικό κλάσμα.

Τέλος, τα ηλεκτροφορητικά μοτίβα των μεμβρανικών κλασμάτων που απομονώθηκαν από τις δύο διαδικασίες κλασματοποίησης του ολικού πρωτεόματος του *Chl. tepidum* εμφανίζονται όμοια, οδηγώντας στο συμπέρασμα ότι η διαδικασία απομόνωσης του ολικού μεμβρανικού κλάσματος του βακτηρίου δεν επηρεάζει την πρωτεϊνική του σύσταση. Συγκριτικά με τα λοιπά πρωτεϊνικά κλάσματα που παρουσιάζονται στην Εικόνα 12, τα μεμβρανικά κλάσματα εμφανίζουν μεγάλη διαφοροποίηση, γεγονός που αποτελεί ένδειξη της απουσίας σημαντικών προσμίξεων από το υδατοδιαλυτό και το κυτταροπλασματικό κλάσμα αντίστοιχα.

3.2.2 N-τερματική COFRADIC ανάλυση υποπρωτεομάτων του *Chl. tepidum* [38-40]

Στην παρούσα μελέτη πραγματοποιήθηκε κλασματοποίηση του πρωτεόματος του πράσινου θειούχου βακτηρίου *Chlorobium tepidum* με βάση την υποκυτταρική τοπολογία των πρωτεϊνών του.

Το *Chl. tepidum* εντάσσεται στα Gram-αρνητικά βακτήρια. Το κυτταρικό τοίχωμα των Gram-αρνητικών βακτηρίων έχει διαφορετική δομή από το αντίστοιχο κυτταρικό τοίχωμα των Gram-θετικών βακτηρίων. Υπάρχει μια επιπλέον κυτταρική μεμβράνη η εξωτερική μεμβράνη (OM) εκτός από την κυτταροπλασματική (CM). Η CM περιβάλλεται από μια ανοιχτή περιοχή η οποία ονομάζεται περίπλασματικός χώρος, στην οποία υπάρχει μια στοιβάδα πεπτιδογλυκάνης και εξωτερικά της πεπτιδογλυκάνης υπάρχει η OM.

Ο περιπλασματικός χώρος είναι ο χώρος μεταξύ της εσωτερικής μεμβράνης και της εξωτερικής μεμβράνης, όπου περιέχονται αρκετές διαφορετικές πρωτεΐνες. Ο ρόλος των πρωτεϊνών του περιπλάσματος είναι να ανιχνεύουν το περιβάλλον και να μεταφέρουν τα απαραίτητα θρεπτικά συστατικά εντός του κυττάρου. Μερικά παραδείγματα περιπλασματικών ενζύμων αποτελούν τα υδρολυτικά ένζυμα, όπως φωσφατάσες, πρωτεάσες και ενδονουκλεάσες, πρωτεΐνες πρόσδεσης, που αναγνωρίζουν συγκεκριμένες ενώσεις (σάκχαρα, αμινοξέα, ανόργανα ιόντα, βιταμίνες) και τις μεταφέρουν διαμέσου της μεμβράνης, ένζυμα αποτοξίνωσης, που τροποποιούν τοξικές για το κύτταρο ενώσεις πριν εισέλθουν στο κύτταρο, και χημειοποδοχείς που βοηθούν στην αναγνώριση της χημικής σύστασης του περιβάλλοντος από το κύτταρο.

Η πεπτιδογλυκάνη στα Gram-αρνητικά βακτήρια περιέχει λιγότερες διασυνδέσεις και καθόλου πεπτιδικές διασυνδέσεις συγκριτικά με την στοιβάδα της πεπτιδογλυκάνης των

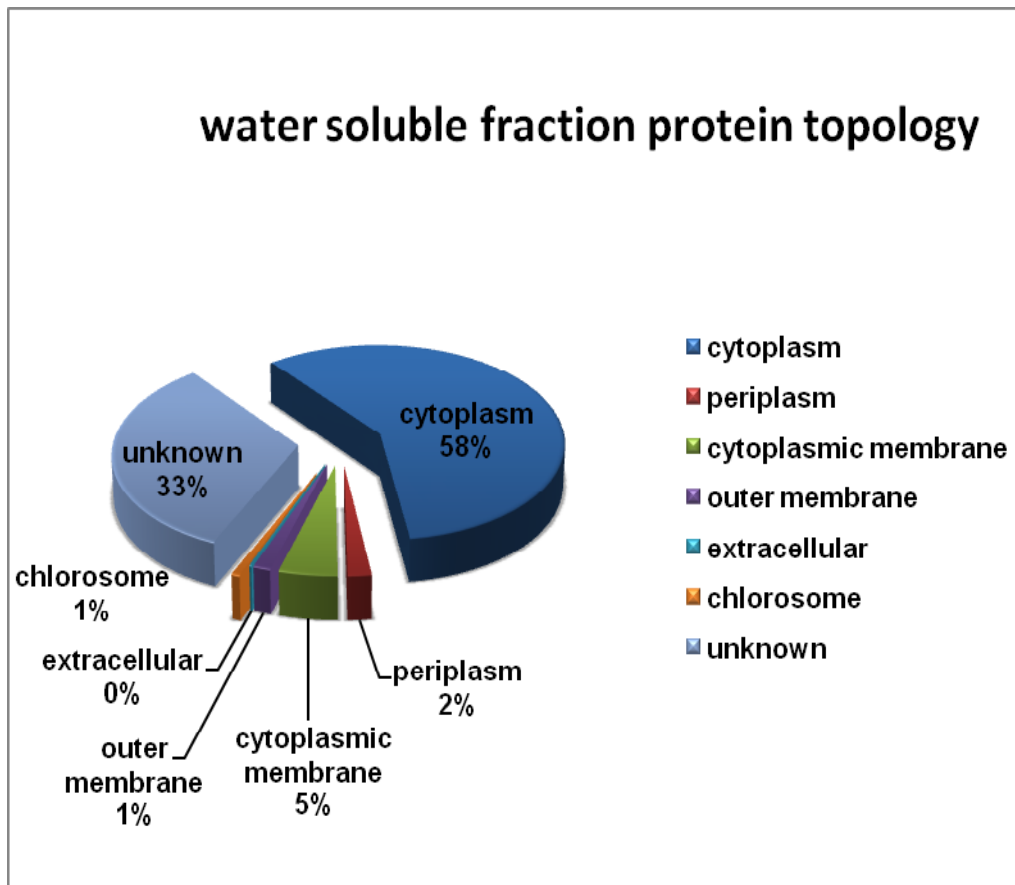
Gram-θετικών βακτηρίων. Ομοιοπολικά προσδεμένες στην πεπτιδογλυκάνη είναι λιποπρωτεΐνες οι οποίες έχουν υδρόφοβες περιοχές που προσδένονται στην ΟΜ. Με αυτόν τον τρόπο η πεπτιδογλυκάνη προσδέεται ισχυρά στην εξωτερική μεμβράνη.

Το κυτταρόπλασμα είναι το μέρος του κυττάρου που περιβάλλεται από την CM και αποτελείται κυρίως από νερό και πρωτεΐνες, αλλά και από βιταμίνες, ιόντα, νουκλεϊκά οξέα και τις πρόδρομες ενώσεις τους, αμινοξέα και τις πρόδρομες ενώσεις τους, υδατάνθρακες, υδρογονάνθρακες και τα παράγωγα τους καθώς και από λιπαρά οξέα και τα παράγωγα τους. Σε αυτό πραγματοποιούνται οι ζωτικές για το κύτταρο χημικές αντιδράσεις. Οι χημικές αντιδράσεις που πραγματοποιούνται στο κυτταρόπλασμα είναι πολύ σημαντικές για την ανάπτυξη του κυττάρου, τον μεταβολισμό και την αντιγραφή. Περιέχει επίσης το χρωμοσωμικό DNA του βακτηρίου το οποίο δε περιβάλλεται από πυρηνική μεμβράνη όπως στα ευκαρυωτικά κύτταρα και διακριτές περιοχές στις οποίες πραγματοποιούνται διάφορες βιοχημικές αντιδράσεις.

3.2.2.1 N-τερματική COFRADIC ανάλυση υδατοδιαλυτού κλάσματος

Ως υδατοδιαλυτό τμήμα του βακτηρίου ορίζεται το σύνολο των πρωτεϊνών που βρίσκονται στο κυτταρόπλασμα και το περίπλασμα του. Στην παρούσα εργασία απομονώθηκε το υδατοδιαλυτό κλάσμα του *Chlorobium tepidum* και αναλύθηκε με N-terminal COFRADIC, οδηγώντας στην ταυτοποίηση 579 πρωτεϊνών, η προβλεπόμενη [51-52] τοπολογία των οποίων εμφανίζεται στο Σχήμα 3. Παρατηρείται ότι η πλειοψηφία των ταυτοποιημένων πρωτεϊνών είναι κυτταροπλασματικές, ενώ 189 πρωτεΐνες είχαν άγνωστη τοπολογία. Επίσης βρέθηκαν 11 περιπλασματικές πρωτεΐνες, 27 κυτταροπλασματικής μεμβράνης και 8 που ανήκουν στην εξωτερική βακτηριακή μεμβράνη. Τέλος ταυτοποιήθηκε μια εκκρινόμενη πρωτεΐνη και οι 4 από τις 10 προβλεπόμενες από το γονιδίωμα χλωροσωμικές πρωτεΐνες. Σημειώνεται επίσης ότι στο

υδατοδιαλυτό κλάσμα βρέθηκαν 36 ριβοσωμικές πρωτεΐνες, γεγονός που υποδηλώνει την αφθονία τους στο κύτταρο καθώς και ενδεχομένως την εύκολη ταυτοποίηση τους λόγω εύκολου ιονισμού τους με ESI.



Σχήμα 3: Προβλεπόμενη κυτταρική τοπολογία των ταυτοποιημένων πρωτεϊνών στο υδατοδιαλυτό κλάσμα.

Συνολικά 212 πρωτεΐνες ταυτοποιήθηκαν με βάση το N-τερματικό τους πεπτίδιο, 118 με πεπτίδιο που εκκινούσε από το δεύτερο αμινοξύ της προβλεπόμενης από το γονιδίωμα πρωτεϊνικής αλληλουχίας ενώ σε συγκεκριμένες τάξεις πρωτεϊνών παρατηρήθηκε η εκτενής θραυσματοποίηση τους ανά ένα αμινοξύ στο N-τερματικό του πεπτιδίου που οδήγησε στην ταυτοποίηση τους. Οι πρωτεΐνες αυτές είτε σχετίζονται με την πρόσδεση σε DNA ή RNA ή με τη μεταφορά ηλεκτρονίων (οξειδοαναγωγάσες, κυτοχρώματα

φερρεδοξίνες, πρωτεΐνες της οικογένειας sox), είτε είναι σαπερονίνες ή πρωτεάσες. Το φαινόμενο αυτό μπορεί να ερμηνευθεί με βάση το γεγονός ότι το *Chl. tepidum* είναι αναερόβιο βακτήριο και η έκθεση του σε αερόβιες συνθήκες κατά την συλλογή και διάρρηξη των κυττάρων του, ενδεχομένως οδήγησε στην πρωτεόλυση των πρωτεϊνών αυτών κατά τα αρχικά στάδια απομόνωσης και κατεργασίας του δείγματος. Στην

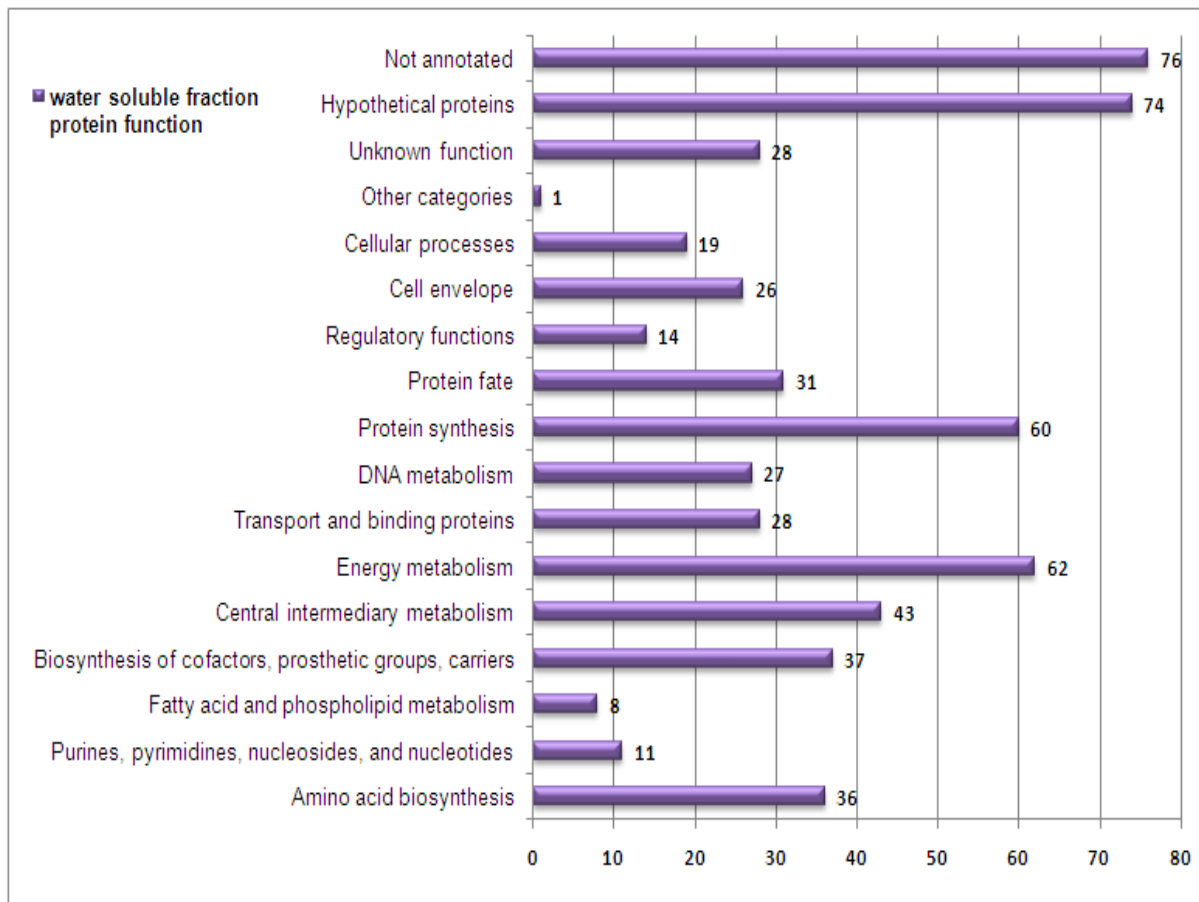
περίπτωση των πρωτεϊνών που σχετίζονται με το DNA, το RNA, των πρωτεασών και των σαπερονινών, η θραυσματοποίηση τους πιθανόν να οφείλεται στο γεγονός ότι η έκφραση τους στα κύτταρα είναι δυναμική και εξαρτάται από τις εκάστοτε συνθήκες που επικρατούν σε αυτά. Συνεπώς, θα πρέπει να εκφράζονται και να αποικοδομούνται με μια δυναμική, κατ' αντιστοιχία με τις ανάγκες του κυττάρου και τα εξωτερικά ερεθίσματα που αυτό δέχεται.

Οι χλωροσωμικές πρωτεΐνες που ταυτοποιήθηκαν είναι οι CsmA, CsmC, CsmI και CsmX. Η CsmA είναι η μεγαλύτερη σε αφθονία χλωροσωμική πρωτεΐνη, βρίσκεται στην παρακρυσταλλική επιφάνεια που συνδέει τα χλωροσώματα με την κυτταροπλασματική μεμβράνη και συνεπώς η εμφάνιση της στο υδατοδιαλυτό τμήμα είναι αναμενόμενη [55]. Σύμφωνα με προηγούμενες μελέτες, η CsmC (καθώς και η CsmD) είναι σε υψηλό ποσοστό προσβάσιμες σε πρωτεάσες και αποδεσμεύονται από τα χλωροσώματα πολύ πιο εύκολα από οποιεσδήποτε άλλες χλωροσωμικές πρωτεΐνες, γεγονός που δικαιολογεί την ταυτοποίηση της CsmC στο υδατοδιαλυτό κλάσμα. Η CsmX βρίσκεται σε μικρότερη αφθονία σε σχέση με τις υπόλοιπες χλωροσωμικές πρωτεΐνες στο χλωροσώμα, ωστόσο σημειώνεται ότι οι CsmA, CsmC, CsmI ταυτοποιήθηκαν με N-τερματικό πεπτίδιο, ενώ η CsmX, με βάση εσωτερικό της πεπτίδιο που ενδεχομένως υποδεικνύει πρωτεόλυση της πριν την ανάλυση.

Η εκκρινόμενη πρωτεΐνη που ταυτοποιήθηκε είναι η Rab family protein (CT1526) η οποία έχει άγνωστο λειτουργικό ρόλο ωστόσο φέρεται να προσδένει GTP και να εμπλέκεται στη μεταφορά σημάτων.

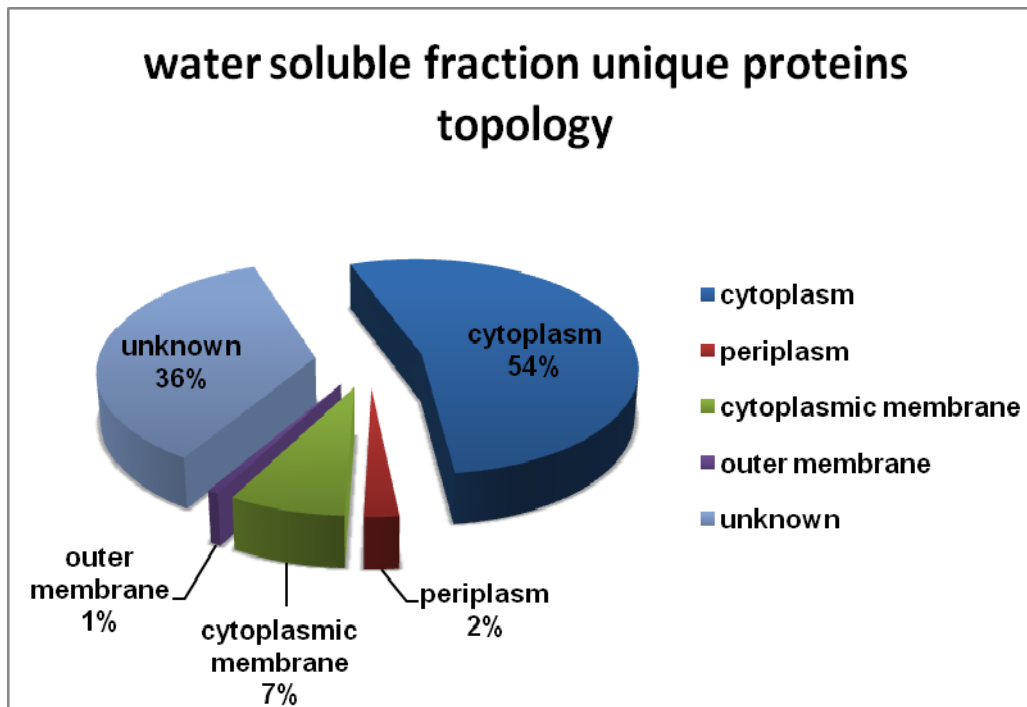
Οι πρωτεΐνες που ταυτοποιήθηκαν στο υδατοδιαλυτό κλάσμα κατηγοριοποιήθηκαν με

βάση τις λειτουργίες τους όπως φαίνεται στο Σχήμα 4. Οι περισσότερες εξ' αυτών δεν έχουν γνωστό λειτουργικό ρόλο (178 πρωτεΐνες), ενώ η πλειοψηφία γνωστής λειτουργίας εμπλέκονται στο μεταβολισμό της ενέργειας, στην πρωτεϊνοσύνθεση, στον κεντρικό μεταβολισμό του κυττάρου καθώς και στη βιοσύνθεση αμινοξέων, προσθετικών ομάδων και συμπαραγόντων.



Σχήμα 4: Προβλεπόμενη λειτουργία των ταυτοποιημένων πρωτεϊνών στο υδατοδιαλυτό κλάσμα.

Από τις συνολικά 579 πρωτεΐνες που ταυτοποιήθηκαν στο υδατοδιαλυτό κλάσμα, 135 βρέθηκαν μόνο σε αυτό το πρωτεϊνικό δείγμα. Το μεγαλύτερο ποσοστό αυτών με βάση την προβλεπόμενη τοπολογία τους [51-52] αντιστοιχεί σε κυτταροπλασματικές πρωτεΐνες, όπως φαίνεται στο Σχήμα 5. Σε σύγκριση με το ολικό υδατοδιαλυτό κλάσμα, παρατηρείται μια αύξηση στις πρωτεΐνες της κυτταροπλασματικής μεμβράνης και στις άγνωστης τοπολογίας πρωτεΐνες.



Σχήμα 5: Προβλεπόμενη κυτταρική τοπολογία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο υδατοδιαλυτό κλάσμα.

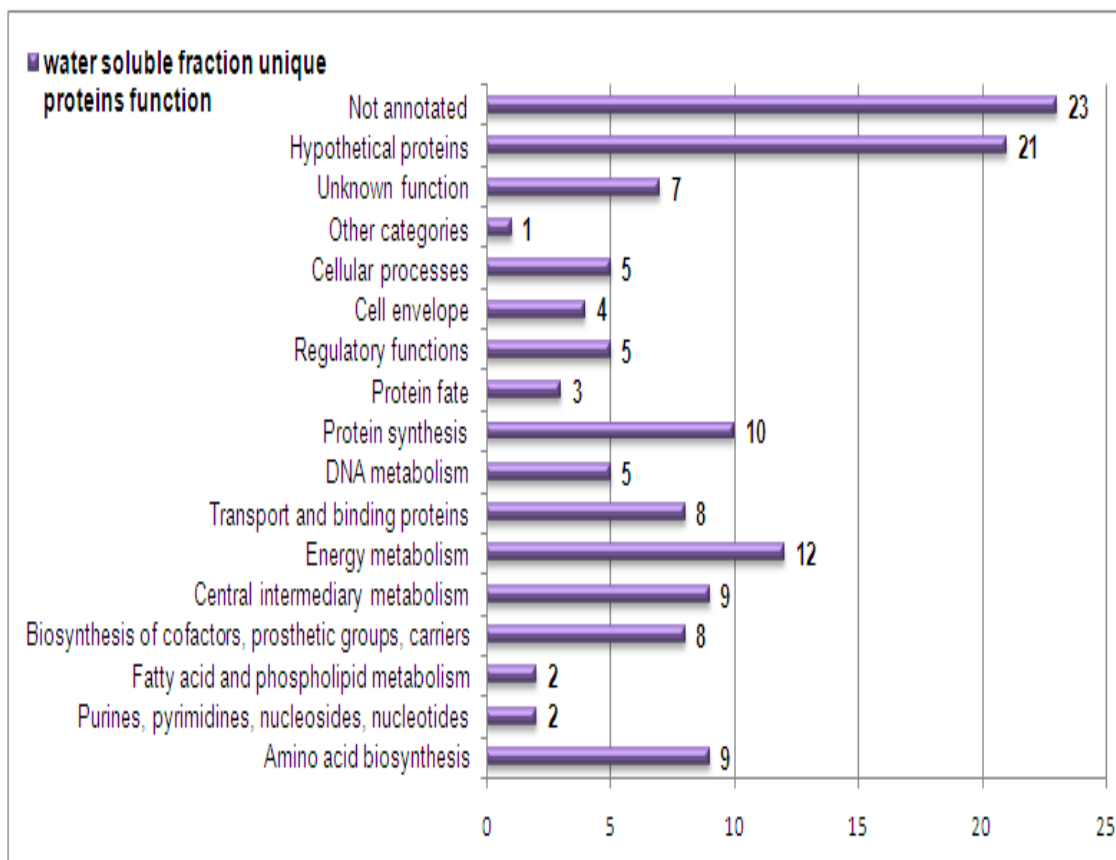
Από τις περιπλασματικές πρωτεΐνες που ταυτοποιήθηκαν η Putative uncharacterized protein (CT1867), η Adhesion protein, putative (CT2106) [Cell envelope: Surface structures] και η Thiol:disulfide interchange protein, thioredoxin family (CT1023) [μεταβολισμός θείου και μεταφορά ηλεκτρονίων] βρέθηκαν μόνο στο υδατοδιαλυτό κλάσμα.

Επίσης από τις προβλεπόμενες ως εξωτερικής μεμβράνης πρωτεΐνες η μόνη που ταυτοποιήθηκε αποκλειστικά στο υδατοδιαλυτό κλάσμα ήταν η Peptidoglycan-associated lipoprotein (CT0638) η οποία εμπλέκεται στη μεταφορά και την πρόσδεση πρωτεϊνών και αλληλεπιδρά με το στρώμα πεπτιδογλυκάνης του βακτηρίου.

Αναφορικά με τις προβλεπόμενες πρωτεΐνες της κυτταροπλασματικής μεμβράνης, από τις 27 που ταυτοποιήθηκαν στο υδατοδιαλυτό κλάσμα, 10 βρέθηκαν αποκλειστικά σε αυτό. Πιο συγκεκριμένα, βρέθηκαν οι Cytochrome d ubiquinol oxidase, subunit I (CT1818), Molybdenum ABC transporter, permease protein (CT0451), Sulfate

transporter family protein (CT0714), Putative uncharacterized protein (CT1149), Multidrug resistance protein A (CT0105), Penicillin-binding protein 1 (CT0176), sensor histidine kinase (CT0195), Cell division protein FtsH (CT0297), ExbD/TolR family protein (CT0634), και Putative uncharacterized protein (CT0748), που εμπλέκονται κατά κύριο λόγο στην μεταφορά και την πρόσδεση, σε ρυθμιστικές λειτουργίες του κυττάρου, στον κυτταρικό φάκελο αλλά και στο μεταβολισμό της ενέργειας.

Συνοπτικά, στο Σχήμα 6 παρουσιάζεται η λειτουργική κατηγοριοποίηση των πρωτεϊνών που ταυτοποιήθηκαν αποκλειστικά στο υδατοδιαλυτό κλάσμα του *Chlorobium tepidum*. Παρατηρείται ότι κατά κύριο λόγο ο λειτουργικός τους ρόλος δεν είναι γνωστός και από τις γνωστές λειτουργίες οι περισσότερες εμπλέκονται στις ίδιες κυτταρικές λειτουργίες που σημειώθηκαν και στο ολικό υδατοδιαλυτό κλάσμα (μεταβολισμός ενέργειας, πρωτεϊνοσύνθεση, κεντρικός μεταβολισμός, βιοσύνθεση αμινοξέων, προσθετικών ομάδων και συμπαραγόντων).

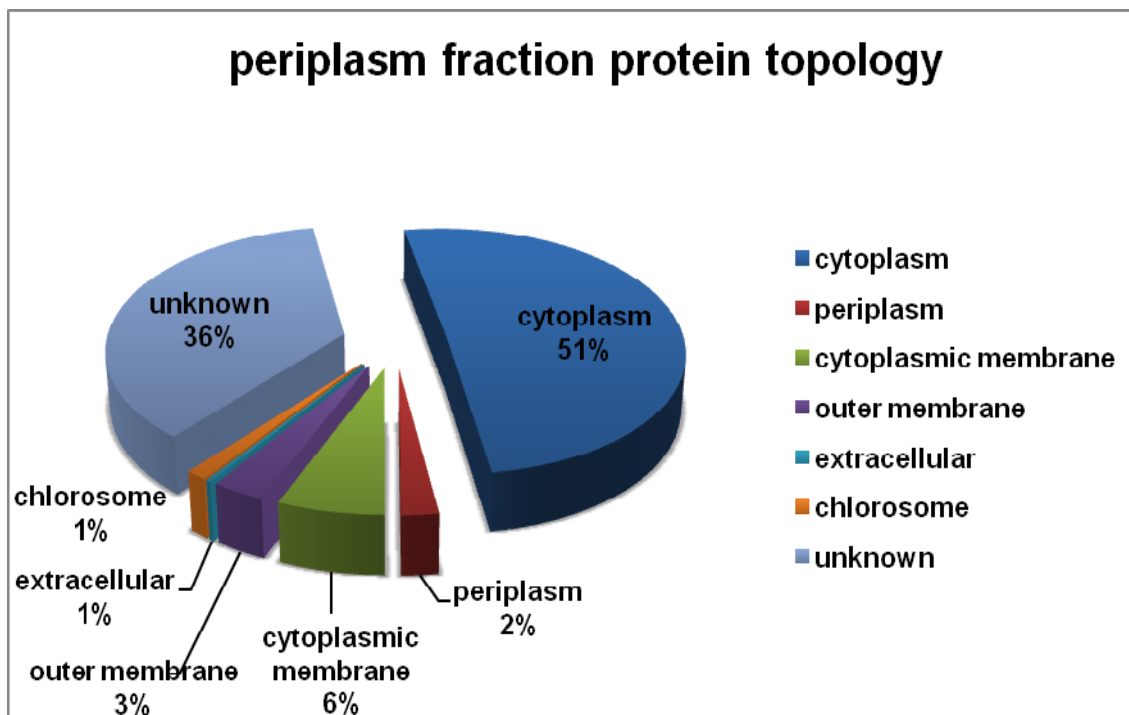


Σχήμα 6: Προβλεπόμενη λειτουργία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο υδατοδιαλυτό κλάσμα.

3.2.2.2 N-τερματική COFRADIC ανάλυση περιπλασματικού κλάσματος

Ως περιπλασματικό τμήμα του βακτηρίου ορίζεται το σύνολο των πρωτεϊνών που βρίσκονται στο περίπλασμα του. Στην παρούσα εργασία απομονώθηκε το περιπλασματικό κλάσμα του *Chlorobium tepidum*, και στη συνέχεια πραγματοποιήθηκε

N-τερματική ανάλυση του με COFRADIC, οδηγώντας στην ταυτοποίηση 317 πρωτεϊνών, η προβλεπόμενη [51-52]τοπολογία των οποίων εμφανίζεται στο Σχήμα 7. Παρατηρείται ότι η πλειοψηφία των πρωτεϊνών αυτών είναι κυτταροπλασματικές (160 πρωτεΐνες), ενώ 115 πρωτεΐνες είχαν άγνωστη τοπολογία. Επίσης βρέθηκαν 6 περιπλασματικές πρωτεΐνες, 20 κυτταροπλασματικής μεμβράνης και 10 που ανήκουν στην εξωτερική μεμβράνη του βακτηρίου. Τέλος ταυτοποιήθηκε μια εκκρινόμενη πρωτεΐνη-που είχε ταυτοποιηθεί και στο υδατοδιαλυτό κλάσμα- και 4 από τις 10 προβλεπόμενες από το γονιδίωμα χλωροσωμικές πρωτεΐνες. Σημειώνεται επίσης ότι στο υδατοδιαλυτό κλάσμα ταυτοποιήθηκαν 24 ριβοσωμικές πρωτεΐνες.



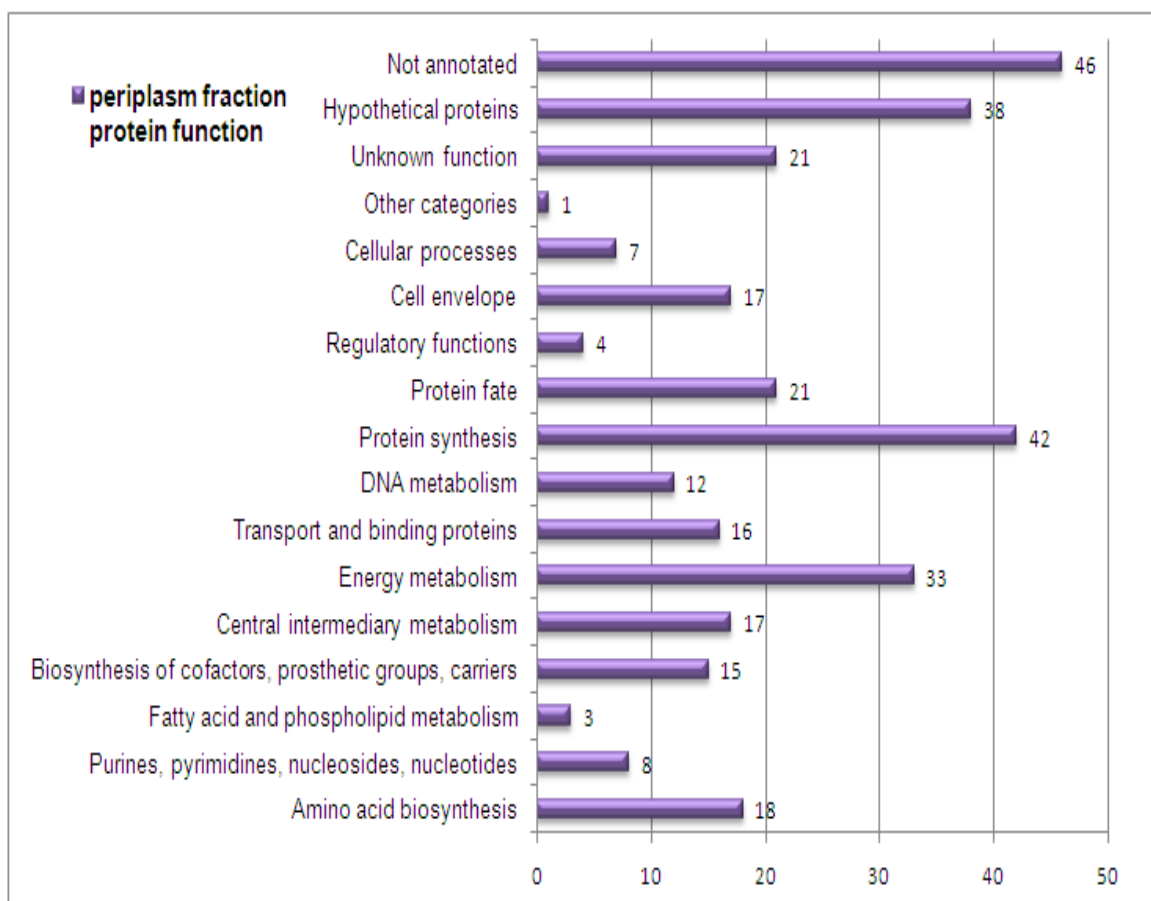
Σχήμα 7: Προβλεπόμενη κυτταρική τοπολογία των ταυτοποιημένων πρωτεϊνών στο περιπλασματικό κλάσμα.

Συνολικά 112 πρωτεΐνες ταυτοποιήθηκαν με βάση το N-τερματικό τους πεπτίδιο, 62 με βάση πεπτίδιο που εκκινούσε από το δεύτερο αμινοξύ της προβλεπόμενης με βάση το γονιδίωμα πρωτεϊνικής αλληλουχίας, ενώ σε συγκεκριμένες τάξεις πρωτεϊνών παρατηρήθηκε και πάλι η εκτενής θραυσματοποίηση τους ανά ένα αμινοξύ στο N-τερματικό άκρο του πεπτιδίου που οδήγησε στην ταυτοποίηση τους. Οι πρωτεΐνες αυτές είτε σχετίζονται με την πρόσδεση σε DNA ή με τη μεταφορά ηλεκτρονίων (οξειδοαναγωγάσες, κυτοχρώματα, πρωτεΐνες της οικογένειας sox), είτε είναι πρωτεάσες.

Οι χλωροσωμικές πρωτεΐνες που ταυτοποιήθηκαν είναι οι CsmA, CsmC, CsmI και CsmX, δηλαδή οι ίδιες που ταυτοποιήθηκαν και στο υδατοδιαλυτό κλάσμα.

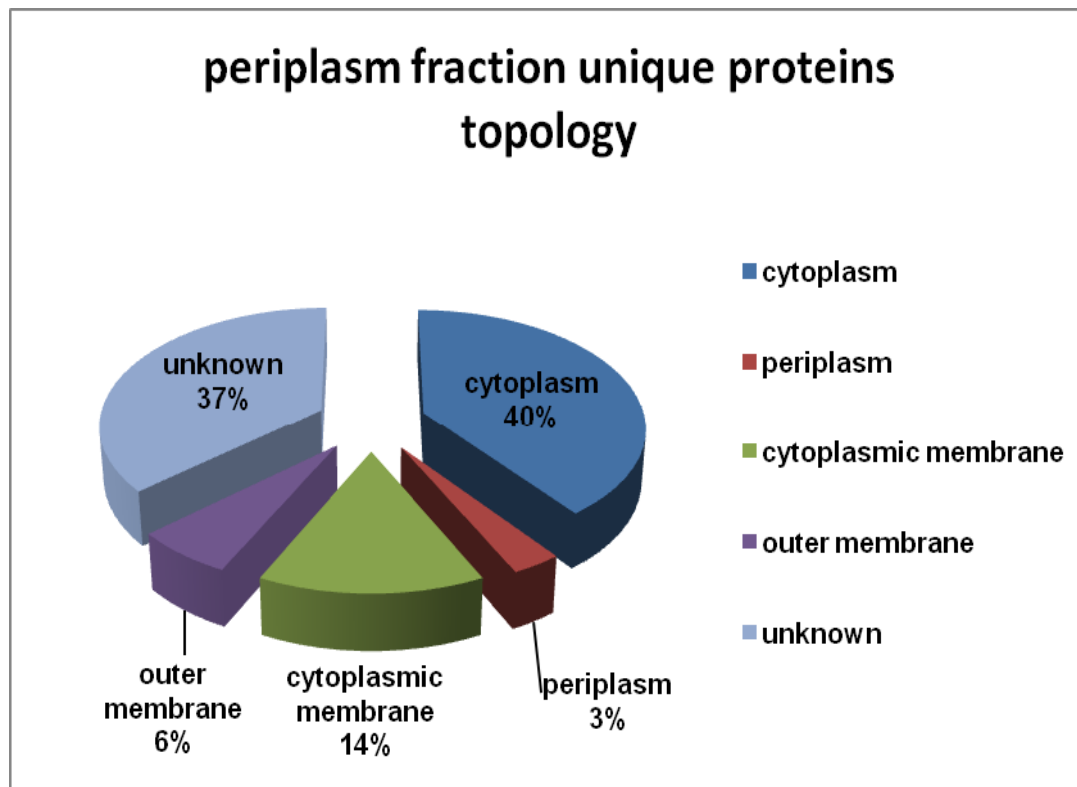
Οι περιπλασματικές πρωτεΐνες που ταυτοποιήθηκαν είναι η Thiol:disulfide interchange protein DsbE, putative (CT1072), η Thiol:disulfide interchange protein, thioredoxin family protein (CT2233), η Putative uncharacterized protein (CT0600), η Protein tolB [Precursor] (CT0636), η Serine protease (CT1447), και η Cytochrome subunit of sulfide dehydrogenase [Precursor] (CT2080).

Οι πρωτεΐνες που ταυτοποιήθηκαν στο περιπλασματικό κλάσμα κατηγοριοποιήθηκαν με βάση τις λειτουργίες τους όπως φαίνεται στο Σχήμα 8 . Οι περισσότερες εξ' αυτών δεν έχουν γνωστό λειτουργικό ρόλο (105 πρωτεΐνες), ενώ η πλειοψηφία γνωστής λειτουργίας εμπλέκονται στην πρωτεϊνοσύνθεση, στον μεταβολισμό της ενέργειας, στον κεντρικό μεταβολισμό του κυττάρου καθώς και στη βιοσύνθεση αμινοξέων, και στην πρωτεϊνική αναδίπλωση και στόχευση (protein fate).



Σχήμα 8: Προβλεπόμενη λειτουργία των ταυτοποιημένων πρωτεϊνών του περιπλασματικού κλάσματος.

Από τις συνολικά 317 πρωτεΐνες που ταυτοποιήθηκαν στο περιπλασματικό κλάσμα, 65 βρέθηκαν μόνο σε αυτό το πρωτεϊνικό δείγμα. Το μεγαλύτερο ποσοστό αυτών με βάση την προβλεπόμενη τοπολογία τους [51-52] αντιστοιχεί σε κυτταροπλασματικές πρωτεΐνες, όπως φαίνεται στο Σχήμα 9. Σε σύγκριση με το ολικό περιπλασματικό κλάσμα, παρατηρείται μια σημαντική μείωση στις κυτταροπλασματικές πρωτεΐνες, με ταυτόχρονη αύξηση στις μεμβρανικές πρωτεΐνες, και ιδιαίτερα στις πρωτεΐνες της κυτταροπλασματικής μεμβράνης.



Σχήμα 9: Προβλεπόμενη κυτταρική τοπολογία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο περιπλασματικό κλάσμα

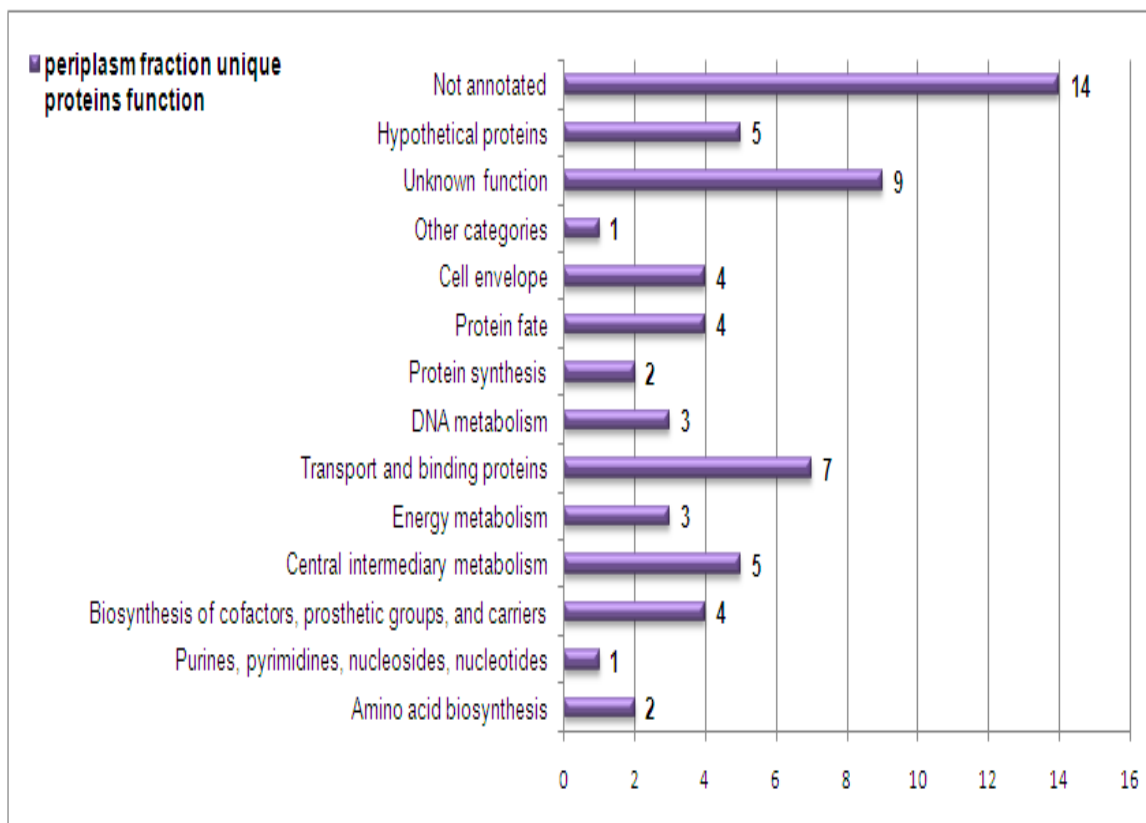
Από τις περιπλασματικές πρωτεΐνες που ταυτοποιήθηκαν η Serine protease (CT1547) και η Cytochrome subunit of sulfide dehydrogenase [Precursor] (CT2080) βρέθηκαν μόνο στο περιπλασματικό κλάσμα οι οποίες εμπλέκονται στην αποικοδόμηση πρωτεϊνών και την μεταφορά ηλεκτρονίων αντίστοιχα.

Επίσης από τις ταυτοποιημένες προβλεπόμενες ως εξωτερικής μεμβράνης πρωτεΐνες αποκλειστικά στο περιπλασματικό κλάσμα ταυτοποιήθηκαν: η Ferric siderophore receptor, putative, TonB receptor family (CT1953) και η Receptor, putative (CT1754) που εμπλέκονται στη μεταφορά και την πρόσδεση μορίων, η υποθετική πρωτεΐνη Putative uncharacterized protein (CT0893), καθώς και η Outer surface protein, putative (CT2144) που συμμετέχει στην διαδικασία βιοσύνθεσης και αποικοδόμησης πολυσακχαριτών και λιποπολυσακχαριτών.

Αναφορικά με τις προβλεπόμενες πρωτεΐνες της κυτταροπλασματικής μεμβράνης, από

τις 20 που ταυτοποιήθηκαν στο περιπλασματικό κλάσμα, 9 βρέθηκαν αποκλειστικά σε αυτό. Πιο συγκεκριμένα, βρέθηκαν οι Putative uncharacterized proteins (CT1247, CT2058 και CT0635), η Protease IV (CT1644), η Na⁺/H⁺ antiporter, putative (CT1144), η ABC transporter, ATP-binding protein (CT0375), η Iron(III) ABC transporter, permease protein, putative (CT0411), η Ribonucleotide reductase family protein (CT0578) και η NfeD protein (CT0795) που εμπλέκονται κατά κύριο λόγο στην μεταφορά και την πρόσδεση μορίων, αλλά και στην αντιγραφή του DNA, ή έχουν πρωτεολυτική δράση.

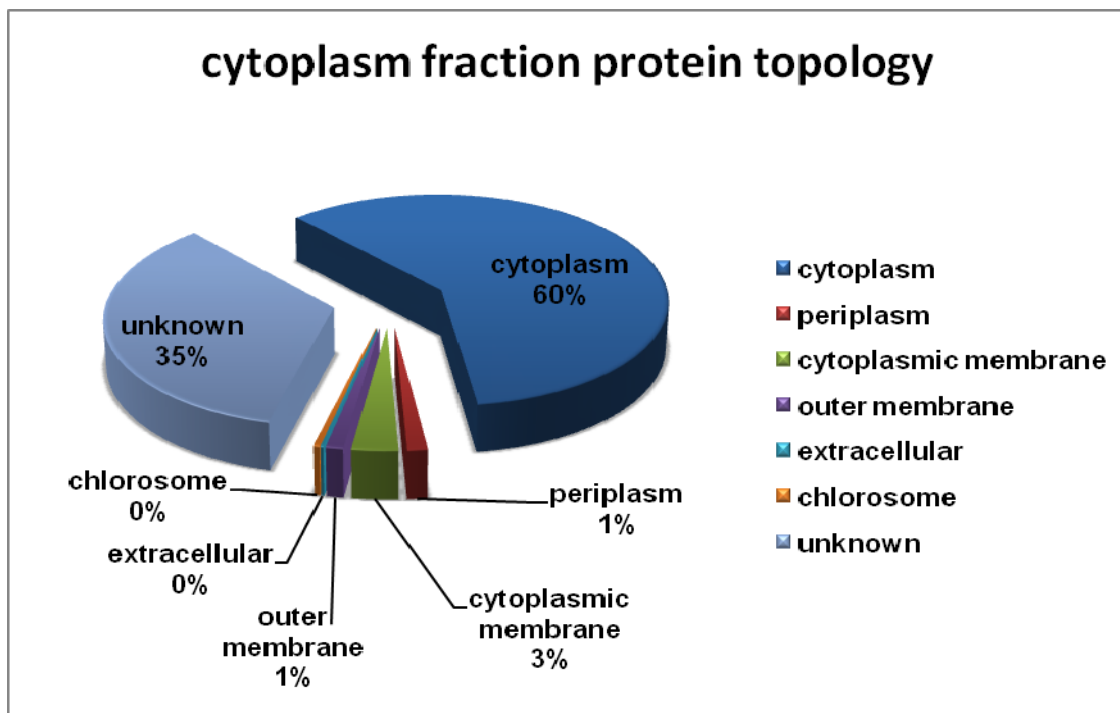
Συνοπτικά, στο Σχήμα 10 παρουσιάζεται η λειτουργική κατηγοριοποίηση των πρωτεϊνών που ταυτοποιήθηκαν αποκλειστικά στο περιπλασματικό κλάσμα του *Chlorobium tepidum*. Παρατηρείται ότι κατά κύριο λόγο ο λειτουργικός τους ρόλος δεν είναι γνωστός αλλά από τις γνωστές λειτουργίες πρωτεΐνες, οι περισσότερες εμπλέκονται κυρίως στην μεταφορά και την πρόσδεση μορίων και τον κεντρικό μεταβολισμό.



Σχήμα 10: Προβλεπόμενη λειτουργία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο περιπλασματικό κλάσμα.

3.2.2.3 N-τερματική COFRADIC ανάλυση κυτταροπλασματικού κλάσματος

Ως κυτταροπλασματικό τμήμα του βακτηρίου ορίζεται το σύνολο των πρωτεϊνών που βρίσκονται στον κυτταροπλασματικό του χώρο του. Στην παρούσα εργασία απομονώθηκε το κυτταρόπλασμα του *Chlorobium tepidum*, και αναλύθηκε με N-terminal COFRADIC. Ταυτοποιήθηκαν 540 πρωτεΐνες, η προβλεπόμενη τοπολογία των οποίων [51-52] εμφανίζεται στο Σχήμα 11. Παρατηρείται ότι η πλειοψηφία των πρωτεϊνών αυτών είναι κυτταροπλασματικές (322 πρωτεΐνες), ενώ 183 πρωτεΐνες είχαν άγνωστη τοπολογία. Επίσης βρέθηκαν 7 περιπλασματικές πρωτεΐνες, 16 κυτταροπλασματικής μεμβράνης και 6 που ανήκουν στην εξωτερική μεμβράνη του βακτηρίου. Τέλος ταυτοποιήθηκε μια εκκρινόμενη πρωτεΐνη-που είχε ταυτοποιηθεί και στο υδατοδιαλυτό και το περιπλασματικό κλάσμα- και 2 από τις 10 προβλεπόμενες από το γονιδίωμα χλωροσωμικές πρωτεΐνες. Σημειώνεται επίσης ότι στο περιπλασματικό κλάσμα βρέθηκαν 36 ριβοσωμικές πρωτεΐνες.



Σχήμα 11: Προβλεπόμενη κυτταρική τοπολογία των ταυτοποιημένων πρωτεϊνών του κυτταροπλασματικού κλάσματος.

Συνολικά 211 πρωτεΐνες ταυτοποιήθηκαν με βάση το N-τερματικό τους πεπτίδιο, 122 με βάση πεπτίδιο που εκκινούσε από το δεύτερο αμινοξύ της προβλεπόμενης με βάση το γονιδίωμα πρωτεϊνικής αλληλουχίας, ενώ σε συγκεκριμένες τάξεις πρωτεϊνών παρατηρήθηκε και πάλι η εκτενής θραυσματοποίηση τους ανά ένα αμινοξύ στο N-τερματικό άκρο του πεπτιδίου που οδήγησε στην ταυτοποίηση τους. Οι πρωτεΐνες αυτές είτε σχετίζονται με την πρόσδεση σε DNA ή RNA ή με τη μεταφορά ηλεκτρονίων (οξειδοαναγωγάσες, κυτοχρώματα, πρωτεΐνες της οικογένειας sox), είτε είναι πρωτεολυτικά ένζυμα ή σαπερονίνες.

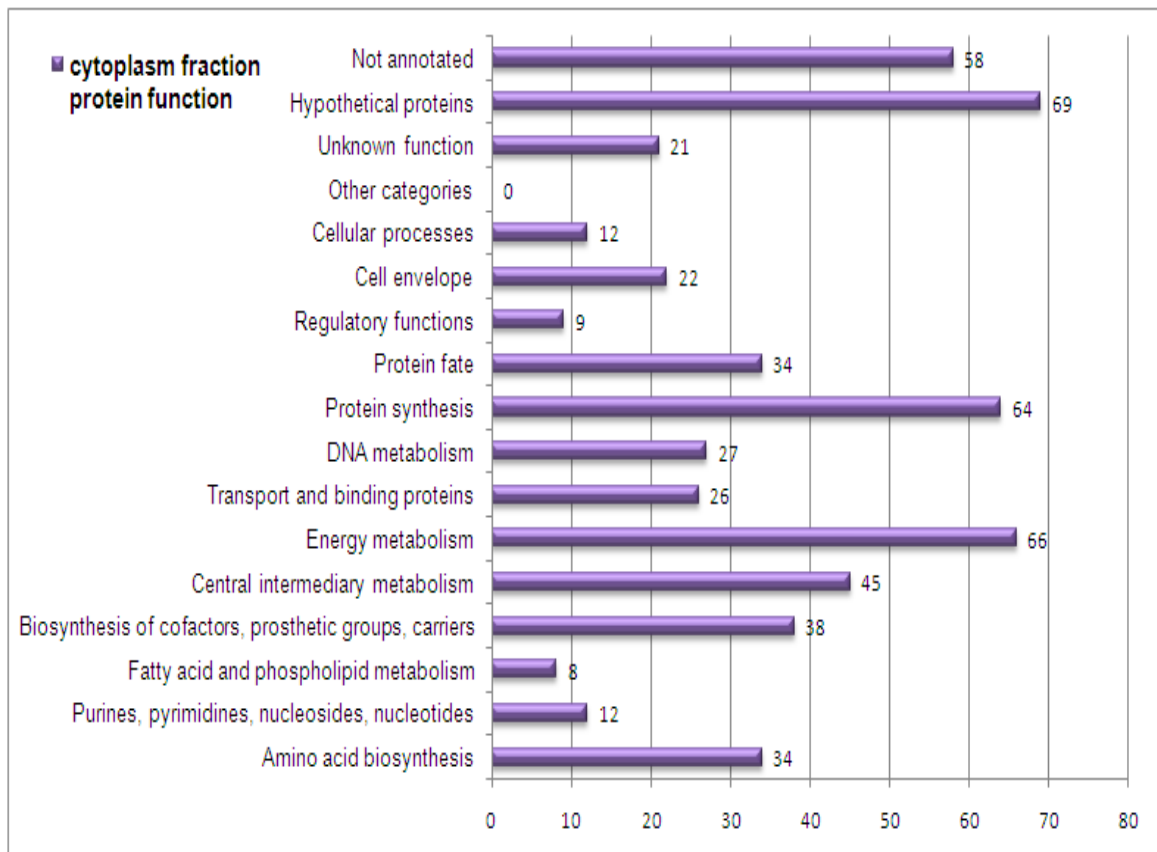
Οι χλωροσωμικές πρωτεΐνες που ταυτοποιήθηκαν είναι οι CsmA και CsmC, δηλαδή σε σύγκριση με το υδατοδιαλυτό και το περιπλασματικό κλάσμα απουσιάζουν η CsmI, που αποδεσμεύεται από χλωρόσωμα πιο δύσκολα από τις άλλες δύο και η CsmX, που

αναλογικά βρίσκεται σε μικρότερη αφθονία σε σύγκριση με τις υπόλοιπες 10 χλωροσωμικές πρωτεΐνες. Επιπρόσθετα, η απουσία των 2 χλωροσωμικών αυτών πρωτεϊνών από το κυτταροπλασματικό κλάσμα και η εμφάνισή τους και στο

υδατοδιαλυτό και στο περίπλασμα, αποτελεί μια ένδειξη την επιτυχία της κλασματοποίησης και τη συμπληρωματικότητα περιπλασματικού και κυτταροπλασματικού κλάσματος που συνδυαζόμενα σχηματίζουν το υδατοδιαλυτό κλάσμα.

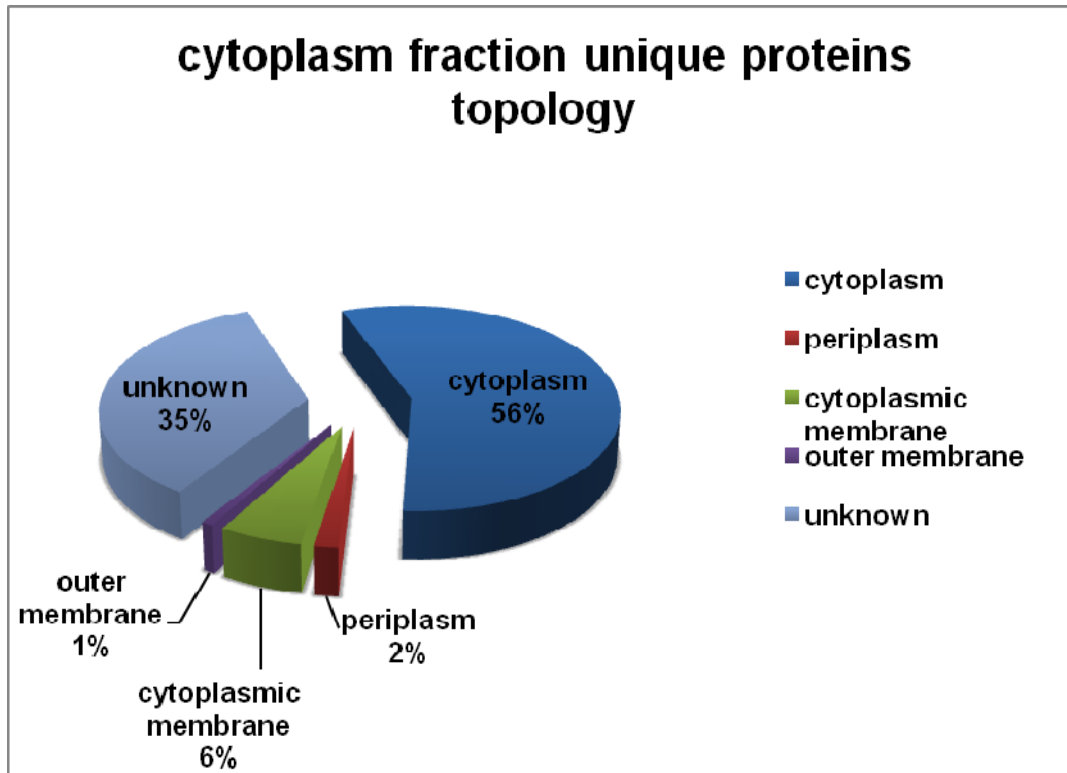
Σε σύγκριση με τα άλλα 2 κλάσματα που παρουσιάστηκαν παραπάνω, παρατηρείται μια αύξηση στις κυτταροπλασματικές πρωτεΐνες και μια μείωση στις πρωτεΐνες της κυτταροπλασματικής μεμβράνης. Συγκριτικά αποκλειστικά με το περιπλασματικό κλάσμα αυτή η τάση γίνεται πολύ εντονότερη, επισημαίνοντας την συμπληρωματικότητα των 2 αυτών κλασμάτων.

Οι πρωτεΐνες που ταυτοποιήθηκαν στο κυτταροπλασματικό κλάσμα κατηγοριοποιήθηκαν με βάση τις λειτουργίες τους όπως φαίνεται στο Σχήμα 12. Οι περισσότερες εξ' αυτών δεν έχουν γνωστό λειτουργικό ρόλο (148 πρωτεΐνες), ενώ η πλειοψηφία γνωστής λειτουργίας εμπλέκονται στον μεταβολισμό της ενέργειας, στην πρωτεϊνοσύνθεση, στον κεντρικό μεταβολισμό του κυττάρου καθώς και στη βιοσύνθεση αμινοξέων, προσθετικών ομάδων και συμπαραγόντων, και στην πρωτεϊνική αναδίπλωση και στόχευση (protein fate).



Σχήμα 12: Προβλεπόμενη λειτουργία των ταυτοποιημένων πρωτεϊνών του κυτταροπλασματικού κλάσματος.

Από τις συνολικά 540 πρωτεΐνες που ταυτοποιήθηκαν στο κυτταροπλασματικό κλάσμα, 122 βρέθηκαν μόνο σε αυτό το πρωτεϊνικό δείγμα. Το μεγαλύτερο ποσοστό αυτών (69 πρωτεΐνες) με βάση την προβλεπόμενη τοπολογία τους [51-52] αντιστοιχεί σε κυτταροπλασματικές πρωτεΐνες (Σχήμα 13). Σε σύγκριση με το ολικό κυτταροπλασματικό κλάσμα, παρατηρείται μια μικρή μείωση στις κυτταροπλασματικές πρωτεΐνες, με ταυτόχρονη αύξηση στις πρωτεΐνες της κυτταροπλασματικής μεμβράνης.



Σχήμα 13: Προβλεπόμενη κυτταρική τοπολογία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο κυτταροπλασματικό κλάσμα.

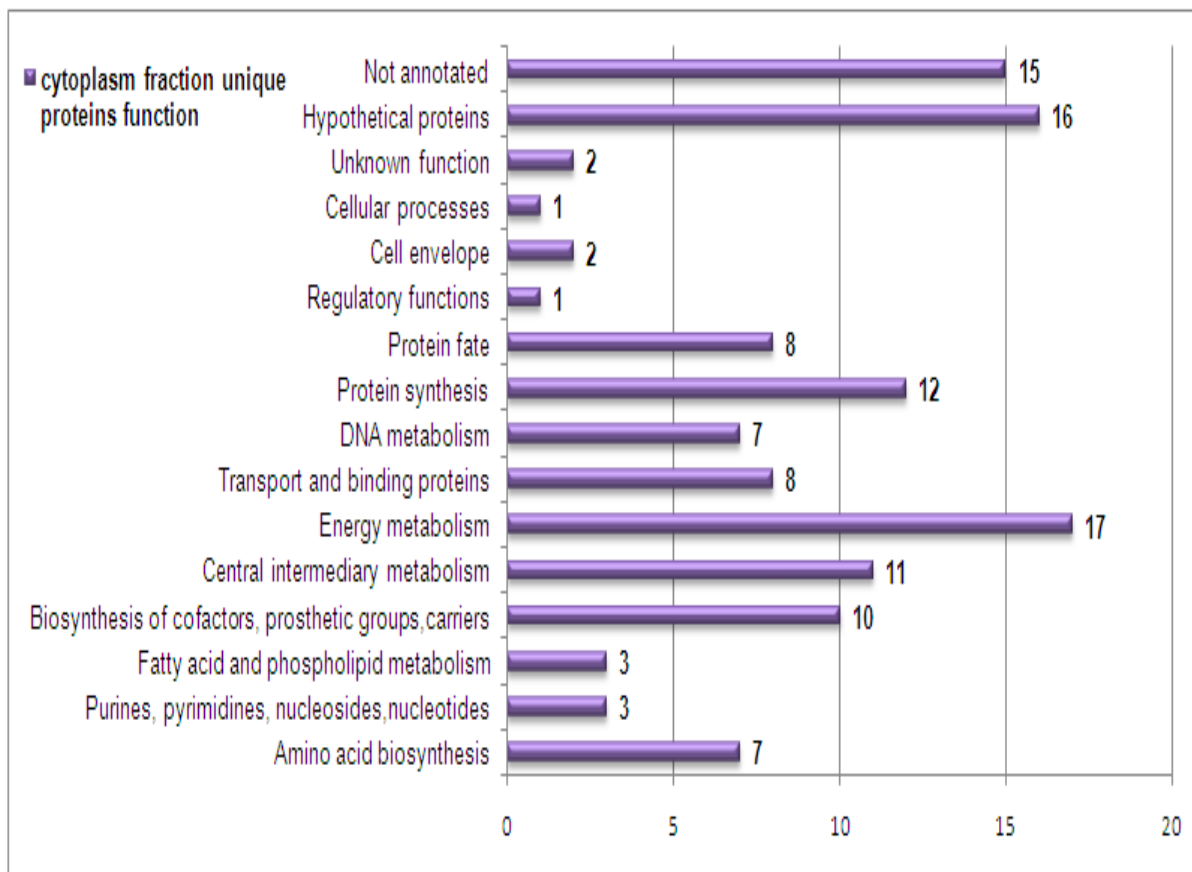
Από τις περιπλασματικές πρωτεΐνες που ταυτοποιήθηκαν η Prolyl oligopeptidase family protein (CT1301) και η Polysulfide reductase, subunit A, putative (CT0496) βρέθηκαν μόνο στο κυτταροπλασματικό κλάσμα οι οποίες εμπλέκονται στην πρωτεόλυση πρωτεϊνών και στον μεταβολισμό του θείου αντίστοιχα.

Επίσης από τις ταυτοποιημένες προβλεπόμενες ως εξωτερικής μεμβράνης πρωτεΐνες αποκλειστικά στο κυτταροπλασματικό κλάσμα ταυτοποιήθηκε η Outer membrane efflux protein, putative (CT0758) που εμπλέκεται στη μεταφορά και την πρόσδεση μορίων.

Αναφορικά με τις προβλεπόμενες πρωτεΐνες της κυτταροπλασματικής μεμβράνης, από τις 16 που ταυτοποιήθηκαν στο κυτταροπλασματικό κλάσμα, 7 βρέθηκαν αποκλειστικά σε αυτό. Πιο συγκεκριμένα, βρέθηκαν οι Putative uncharacterized proteins (CT0816 και CT2016), η NADH dehydrogenase I, 49 kDa subunit (CT0769), η ABC transporter, ATP-binding protein (CT1092), η Anthranilate synthase component I (CT1448), η

Precorrin-3B C17-methyltransferase/precorrin-8X methylmutase (CT0378), και η Phosphate ABC transporter, permease protein, putative (CT0897) που εμπλέκονται κατά κύριο λόγο στην μεταφορά και την πρόσδεση μορίων, αλλά και στην βιοσύνθεση αμινοξέων και προσθετικών ομάδων, συμπαραγόντων και φορέων και στην μεταφορά ηλεκτρονίων.

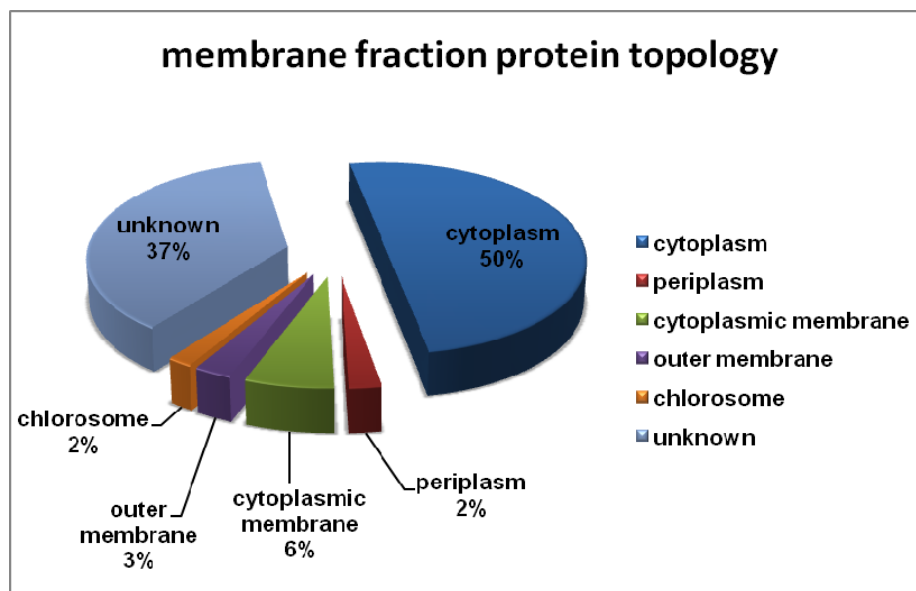
Συνοπτικά, στο Σχήμα 14 παρουσιάζεται η λειτουργική κατηγοριοποίηση των πρωτεϊνών που ταυτοποιήθηκαν αποκλειστικά στο κυτταροπλασματικό κλάσμα του *Chlorobium terpidum*. Παρατηρείται ότι κατά κύριο λόγο ο λειτουργικός τους ρόλος δεν είναι γνωστός αλλά από τις γνωστές λειτουργίας πρωτεΐνες, οι περισσότερες εμπλέκονται κυρίως στον μεταβολισμό της ενέργειας, στην πρωτεϊνοσύνθεση, τον κεντρικό μεταβολισμό του κυττάρου και στη βιοσύνθεση προσθετικών ομάδων, συμπαραγόντων και φορέων.



Σχήμα 14: Προβλεπόμενη λειτουργία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο κυτταροπλασματικό κλάσμα.

3.2.2.4 N-τερματική COFRADIC ανάλυση ολικού μεμβρανικού κλάσματος

Ως ολικό μεμβρανικό κλάσμα του βακτηρίου *Chlorobium tepidum* ορίζουμε το σύνολο των πρωτεϊνών που βρίσκονται στην εξωτερική και εσωτερική του μεμβράνη καθώς και στα χλωροσώματα. Στην παρούσα εργασία απομονώθηκαν οι ολικές μεμβράνες του *Chlorobium tepidum*. Από την N-τερματική ανάλυση του ολικού μεμβρανικού δείγματος με COFRADIC, ταυτοποιήθηκαν 220 πρωτεΐνες, η προβλεπόμενη τοπολογία [51-52] των οποίων εμφανίζεται στο Σχήμα 15. Παρατηρείται ότι οι μισές εξ αυτών είναι κυτταροπλασματικές (110 πρωτεΐνες), ενώ 81 πρωτεΐνες είχαν άγνωστη τοπολογία. Επίσης βρέθηκαν 5 περιπλασματικές πρωτεΐνες, 14 κυτταροπλασματικής μεμβράνης και 6 που ανήκουν στην εξωτερική μεμβράνη του βακτηρίου. Τέλος ταυτοποιήθηκαν και 4 από τις 10 προβλεπόμενες από το γονιδίωμα χλωροσωμικές πρωτεΐνες. Σημειώνεται επίσης ότι στο μεμβρανικό κλάσμα βρέθηκαν 18 ριβοσωμικές πρωτεΐνες.



Σχήμα 15: Προβλεπόμενη κυτταρική τοπολογία των ταυτοποιημένων πρωτεϊνών του μεμβρανικού κλάσματος.

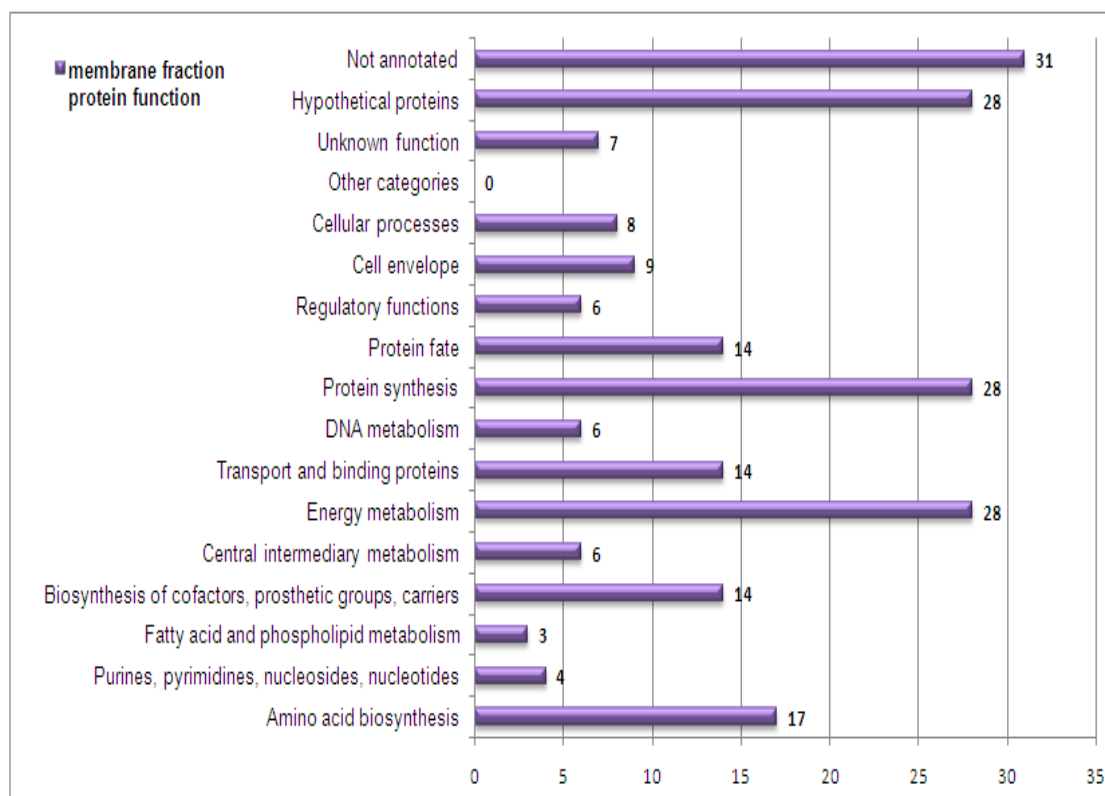
Συνολικά 101 πρωτεΐνες ταυτοποιήθηκαν με βάση το N-τερματικό τους πεπτιδίο, 55 με βάση πεπτιδίο που εκκινούσε από το δεύτερο αμινοξύ της προβλεπόμενης με βάση το γονιδίωμα πρωτεϊνικής αλληλουχίας, ενώ σε συγκεκριμένες τάξεις πρωτεϊνών παρατηρήθηκε και πάλι η εκτενής θραυσματοποίηση τους ανά ένα αμινοξύ στο N-τερματικό άκρο του πεπτιδίου που οδήγησε στην ταυτοποίησή τους. Οι πρωτεΐνες αυτές είτε σχετίζονται με την πρόσδεση σε DNA ή με τη μεταφορά ηλεκτρονίων (πρωτεΐνες της οικογένειας sox), είτε είναι πρωτεολυτικά ένζυμα ή σαπερονίνες.

Οι χλωροσωμικές πρωτεΐνες που ταυτοποιήθηκαν είναι οι CsmA, CsmC, CsmI και CsmJ, δηλαδή σε σύγκριση με το υδατοδιαλυτό και το περιπλασματικό κλάσμα απουσιάζει η CsmX, που αναλογικά βρίσκεται σε μικρότερη αφθονία σε σύγκριση με τις υπόλοιπες 10 χλωροσωμικές πρωτεΐνες και ενδεχομένως δεν κατέστη δυνατό να ταυτοποιηθεί λόγω της περιπλοκότητας και της φύσης του δείγματος. Επιπρόσθετα, ταυτοποιήθηκε η πρωτεΐνη CsmJ που δε βρέθηκε σε κανένα από τα υπόλοιπα ανελυμένα με N-τερματικό COFRADIC δείγματα που προαναφέρθηκαν. Σύμφωνα με τη βιβλιογραφία, η CsmJ φέρεται να σχηματίζει ετεροτετραμερή με την CsmI, τύπου CsmI₂-CsmJ₂.

Σε σύγκριση με τα άλλα 3 κλάσματα που παρουσιάστηκαν παραπάνω, αναφορικά με την τοπολογία των ταυτοποιημένων πρωτεϊνών, παρατηρείται μια μεγάλη ομοιότητα μεταξύ του μεμβρανικού κλάσματος και αυτού του περιπλάσματος. Το φαινόμενο αυτό, πιθανόν να μπορεί να ερμηνευτεί λαμβάνοντας υπόψη τις δυσκολίες που παρουσιάζονται κατά την ανάλυση μεμβρανικών πρωτεϊνών με μεθόδους φασματομετρίας μάζας [62]. Ως αποτέλεσμα, ταυτοποιούνται σε μεγάλο ποσοστό- που δεν αγγίζει ωστόσο αυτό του υδατοδιαλυτού και κυτταροπλασματικού κλάσματος- κυτταροπλασματικές πρωτεΐνες των οποίων η ανάλυση με αυτές τις μεθοδολογίες είναι πολύ ευκολότερη. Αξίζει να σημειωθεί επίσης ότι οι μεμβρανικές πρωτεΐνες συνήθως δεν βρίσκονται σε εξίσου μεγάλη αφθονία στα κύτταρα με τις κυτταροπλασματικές πρωτεΐνες .

Οι πρωτεΐνες που ταυτοποιήθηκαν στο μεμβρανικό κλάσμα κατηγοριοποιήθηκαν με βάση τις λειτουργίες τους όπως φαίνεται στο Σχήμα 16. Οι περισσότερες εξ' αυτών δεν

έχουν γνωστό λειτουργικό ρόλο (66 πρωτεΐνες), ενώ η πλειοψηφία γνωστής λειτουργίας εμπλέκονται στον μεταβολισμό της ενέργειας, στην πρωτεϊνοσύνθεση, στον κεντρικό μεταβολισμό του κυττάρου καθώς και στη βιοσύνθεση αμινοξέων, προσθετικών ομάδων και συμπαραγόντων, στην μεταφορά και την πρόσδεση μορίων και στην πρωτεϊνική αναδίπλωση και στόχευση (protein fate).



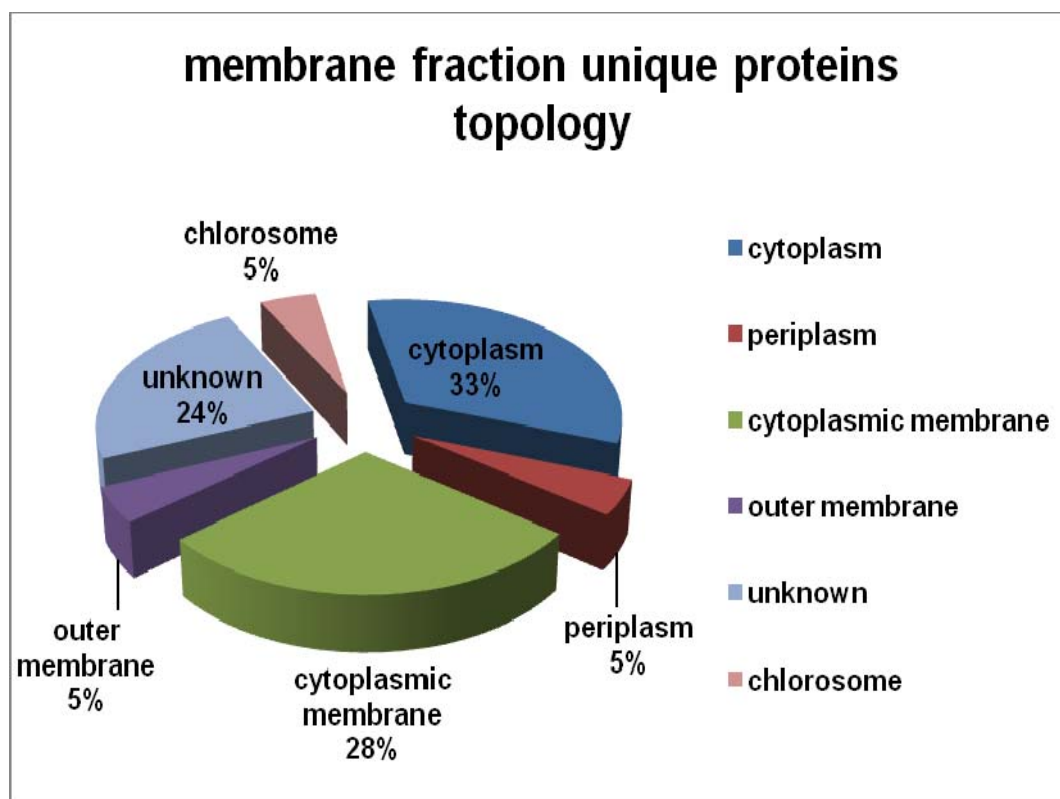
Σχήμα 16: Προβλεπόμενη λειτουργία των ταυτοποιημένων πρωτεϊνών του μεμβρανικού κλάσματος.

Από τις συνολικά 220 πρωτεΐνες που ταυτοποιήθηκαν στο μεμβρανικό κλάσμα, 21 βρέθηκαν μόνο σε αυτό το πρωτεϊνικό δείγμα. Το μεγαλύτερο ποσοστό αυτών με βάση την προβλεπόμενη τοπολογία τους [51-52] αντιστοιχεί σε κυτταροπλασματικές πρωτεΐνες (7 πρωτεΐνες), όπως φαίνεται στο Σχήμα 17. Σε σύγκριση με το ολικό μεμβρανικό κλάσμα ωστόσο, παρατηρείται μια σημαντική μείωση στις κυτταροπλασματικές πρωτεΐνες και στις πρωτεΐνες άγνωστης τοπολογίας, με ταυτόχρονη ανάλογη αύξηση στις μεμβρανικές πρωτεΐνες και κυρίως σε αυτές της κυτταροπλασματικής μεμβράνης. Πιο συγκεκριμένα αποκλειστικά στο μεμβρανικό κλάσμα ταυτοποιήθηκαν 6 πρωτεΐνες που προβλέπεται να ανήκουν στην

κυτταροπλασματική μεμβράνη, η Cell division ATP-binding protein FtsE (CT0590), η ABC transporter, ATP-binding protein CydC (CT1822), η Tryptophan synthase beta

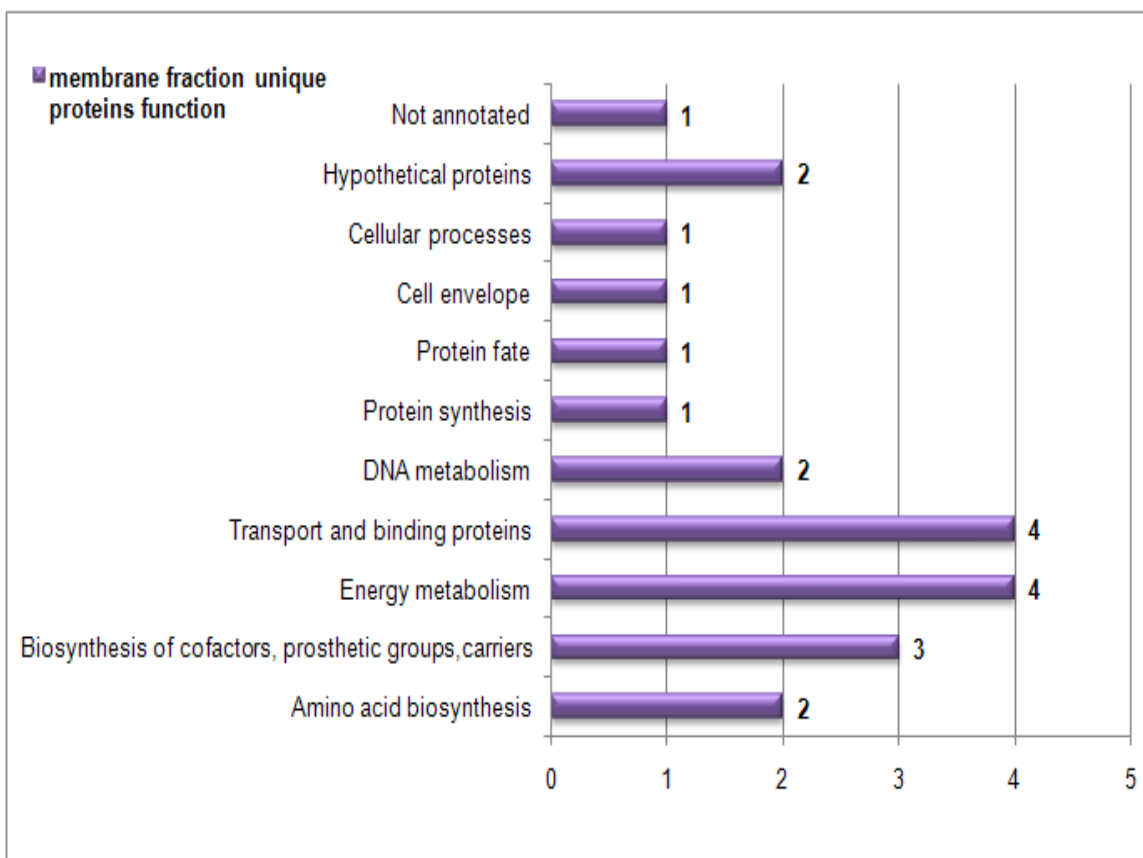
chain (CT0521), η Cobalamin synthase (CT0948), η Bacteriochlorophyll synthase, 34 kDa subunit (CT1270) και η Photosystem P840 reaction center, large subunit (CT2020). Οι πρωτεΐνες αυτές εμπλέκονται κατά κύριο λόγο στη βιοσύνθεση αμινοξέων, προσθετικών ομάδων και συμπαραγόντων, στην φωτοσύνθεση αλλά και στην κυτταρική διαίρεση. Συμπληρωματικά, ως προβλεπόμενη πρωτεΐνη εξωτερικής μεμβράνης, ταυτοποιήθηκε μόνο στο μεμβρανικό κλάσμα η ABC-type export system, outer membrane channel protein (CT0554).

Τις πρωτεΐνες που ταυτοποιήθηκαν μόνο από το μεμβρανικό δείγμα συμπληρώνουν η περιπλασματική πρωτεΐνη Peptide ABC transporter, periplasmic peptide-binding protein (CT1632) που εμπλέκεται στην πρόσδεση πεπτιδίων, η χλωροσωμική πρωτεΐνη CsmJ, καθώς και 7 κυτταροπλασματικές πρωτεΐνες. Οι πρωτεΐνες αυτές είναι η DNA deoxyribodipyrimidine photolyase, classII (CT0511), η ATP-dependent hsl protease ATP-binding subunit hslU (CT1191), η Cytosine deaminase (CT0588), η LAO/AO transport system kinase (CT0798), η DNA polymerase III, alpha subunit (CT0840), η Arginine biosynthesis bifunctional protein argJ beta chain (CT1110), και η Dihydroneopterin aldolase (CT1937) που συμμετέχουν κυρίως στον μεταβολισμό του DNA και την βιοσύνθεση συμπαραγόντων, προσθετικών ομάδων και φορέων. Τέλος ταυτοποιήθηκαν και 5 πρωτεΐνες άγνωστης τοπολογίας, οι Putative uncharacterized proteins (CT0118, CT1519, CT1064) που δεν παρουσιάζουν γνωστή λειτουργία, η TRNA methyltransferase SpoU (CT0879), και η ADP-L-glycero-D-mannoheptose-6-epimerase (CT1258) που σχετίζονται με την πρωτεϊνοσύνθεση και τον κυτταρικό φάκελο, αντίστοιχα.



Σχήμα 17: Προβλεπόμενη κυτταρική τοπολογία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο μεμβρανικό κλάσμα.

Συνοπτικά, στο Σχήμα 18 παρουσιάζεται η λειτουργική κατηγοριοποίηση των πρωτεϊνών που ταυτοποιήθηκαν αποκλειστικά στο μεμβρανικό κλάσμα του *Chlorobium tepidum*. Παρατηρείται ότι από τις γνωστές λειτουργίας πρωτεΐνες, οι περισσότερες εμπλέκονται κυρίως στον μεταβολισμό της ενέργειας, στην μεταφορά και την πρόσδεση μορίων και στη βιοσύνθεση προσθετικών ομάδων, συμπαραγόντων και φορέων.



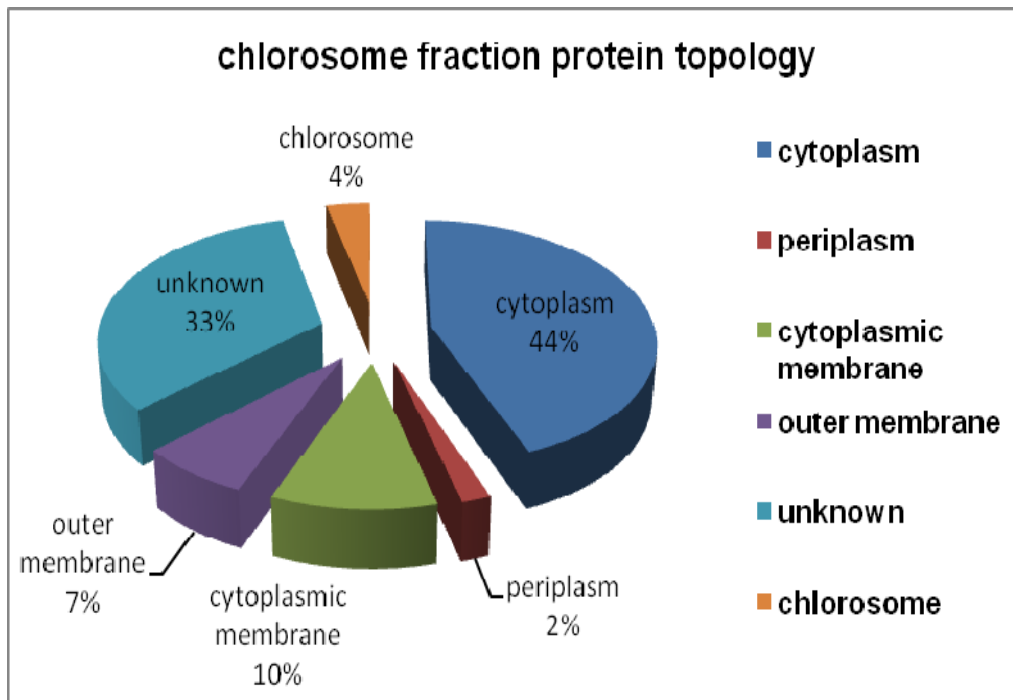
Σχήμα 18: Προβλεπόμενη λειτουργία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο μεμβρανικό κλάσμα.

3.2.3 Φασματομετρία μάζας και ταυτοποίηση πρωτεϊνών χλωροσωμικού κλάσματος

Μελετήθηκε η σύσταση των πρωτεϊνών των απομονωμένων χλωροσωμάτων μετά από πρωτεόλυση τους με θρυψίνη και ταυτοποίηση τους με φασματομετρία μάζας. Το χλωροσωμικό πρωτεϊνικό δείγμα υπέστη κατεργασία με θρυψίνη και τα πεπτίδια που προέκυψαν, διαχωρίστηκαν με RP-HPLC. Από τα 1514 φάσματα MS/MS, ταυτοποιήθηκαν 835 διαφορετικά πεπτίδια που αντιστοιχούν συνολικά σε 270 διαφορετικές πρωτεΐνες. Στις πρωτεΐνες αυτές βρέθηκαν οι 10 πρωτεΐνες που με βάση το

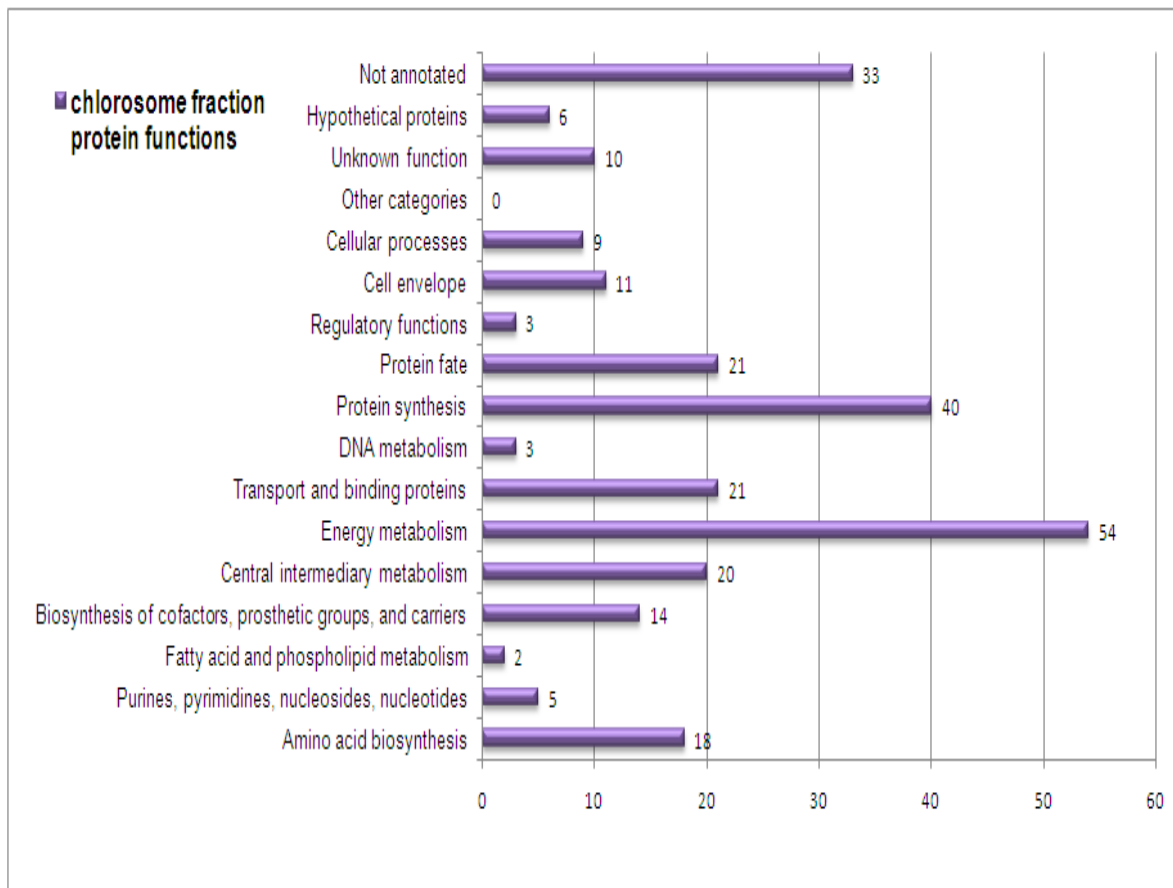
γονιδίωμα και πειραματικά ευρήματα αναφέρονται ως χλωροσωμικές πρωτεΐνες και σχετίζονται δομικά και λειτουργικά με τον χλωροσωμικό φάκελο.

Η προβλεπόμενη υποκυτταρική τοπολογία των ταυτοποιημένων πρωτεϊνών παρουσιάζεται στο Σχήμα 19 [51]. Πιο συγκεκριμένα, βρέθηκαν 120 κυτταροπλασματικές πρωτεΐνες, 5 περιπλασματικές, 28 κυτταροπλασματικής μεμβράνης, και 19 πρωτεΐνες εξωτερικής μεμβράνης. Αξίζει να σημειωθεί ότι αν και ταυτοποιήθηκαν οι χλωροσωμικές πρωτεΐνες στο σύνολο τους, αντιστοιχούν στο 4% επί του συνόλου των ταυτοποιημένων πρωτεϊνών, καθώς και ότι 90 πρωτεΐνες εμφανίστηκαν ως άγνωστης τοπολογίας [51].



Σχήμα 19: Προβλεπόμενη κυτταρική τοπολογία των ταυτοποιημένων πρωτεϊνών του χλωροσωμικού κλάσματος.

Από την λειτουργική κατηγοριοποίηση των πρωτεϊνών που ταυτοποιήθηκαν, προέκυψαν 15 κατηγορίες, που παρουσιάζονται στο Σχήμα 20. Συγκεκριμένα 3 πρωτεΐνες έχουν ρυθμιστικό ρόλο, 40 εμπλέκονται στην πρωτεϊνοσύνθεση, 21 στη μεταφορά και την πρόσδεση μορίων, 11 σε κυτταρικές διαδικασίες, 109 σε ποικίλα μεταβολικά και βιοσυνθετικά μονοπάτια (μεταβολισμός ενέργειας, σύνθεση αμινοξέων, συμπαραγόντων, πουρίνης και πυριμιδίνης, καθώς και μεταβολισμός DNA), 11 πρωτεΐνες είναι μεμβρανικές, και 21 πρωτεΐνες εμπλέκονται στην τροποποίηση των πρωτεϊνών. Αξίζει να σημειωθεί ότι 39 από τις ταυτοποιημένες πρωτεΐνες είναι υποθετικές και τέλος ότι 10 από τις πρωτεΐνες έχουν άγνωστο ρόλο.

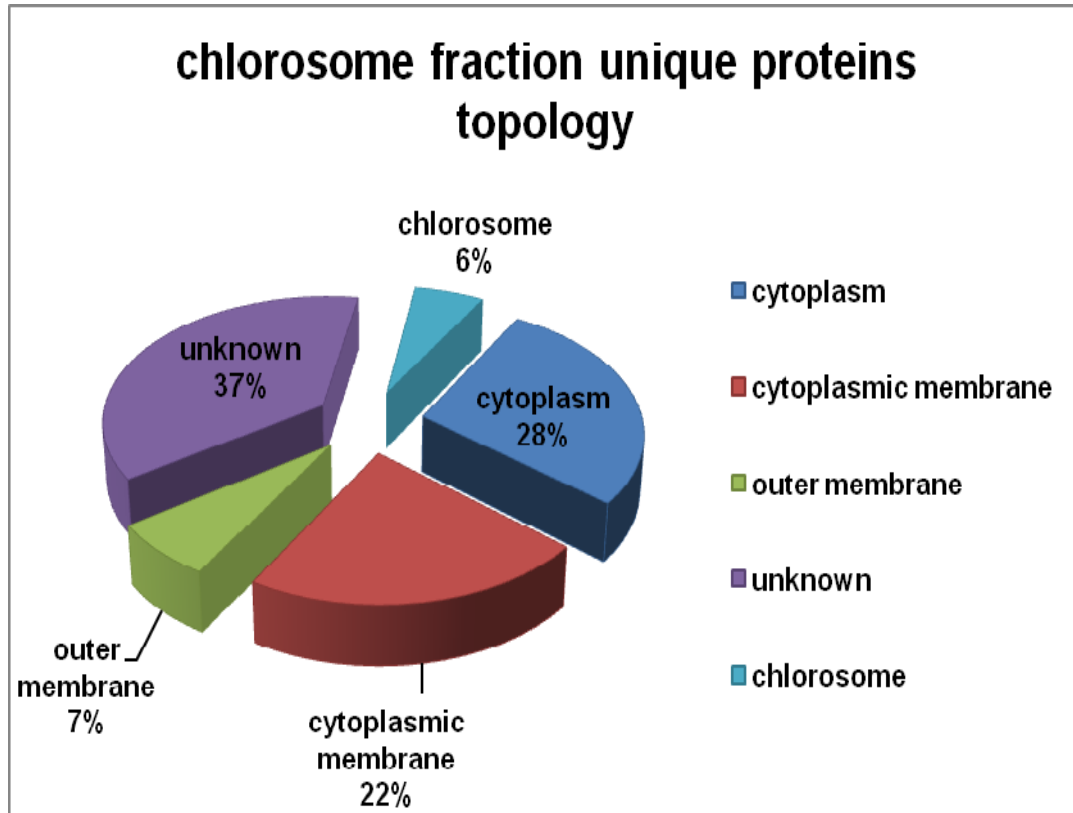


Σχήμα 20: Προβλεπόμενη λειτουργία των ταυτοποιημένων πρωτεϊνών στο χλωροσωμικό κλάσμα.

Ο αριθμός των διαφορετικών πεπτιδίων που χρησιμοποιήθηκαν για την ταυτοποίηση των πρωτεϊνών κυμαινόταν από 1 μέχρι και 38, με 3 πρωτεΐνες να έχουν ταυτοποιηθεί με περισσότερα από 20 διαφορετικά πεπτίδια (η conserved hypothetical protein CT0065 με 38 πεπτίδια, η CmsA χλωροσωμική πρωτεΐνη CT1942 με 21 και η groEL CT0530 με 19 διαφορετικά πεπτίδια).

Μετά απο σύγκριση με τα λοιπά πρωτεϊνικά κλάσματα που απομονώθηκαν απο το *Chl. teridum* στην παρούσα μελέτη, 86 πρωτεΐνες δεν ταυτοποιήθηκαν σε κανένα απο τα άλλα κλάσματα, η προβλεπόμενη τοπολογία των οποίων παρουσιάζεται στο Σχήμα 21 [51]. Το γεγονός αυτό ενδεχομένως να οφείλεται σε εμπλουτισμό του δείγματος λόγω μείωσης της πολυπλοκότητας του ή και στο ότι ακολουθήθηκαν διαφορετικές διαδικασίες ανάλυσης των πρωτεϊνών σε κάθε περίπτωση. Αξίζει να σημειωθεί ωστόσο ότι το μεγαλύτερο ποσοστό των πρωτεϊνών αυτών προβλέπεται να είναι άγνωστης

τοπολογίας, ενώ 25 πρωτεΐνες είναι μεμβρανικές (19 κυτταροπλασματικής μεμβράνης και 6 εξωτερικής μεμβράνης) και 24 πρωτεΐνες κατηγοριοποιούνται ως κυτταροπλασματικές. Επίσης 5 από τις 10 συνολικά χλωροσωμικές πρωτεΐνες ταυτοποιήθηκαν μόνο στο χλωροσωμικό κλάσμα.

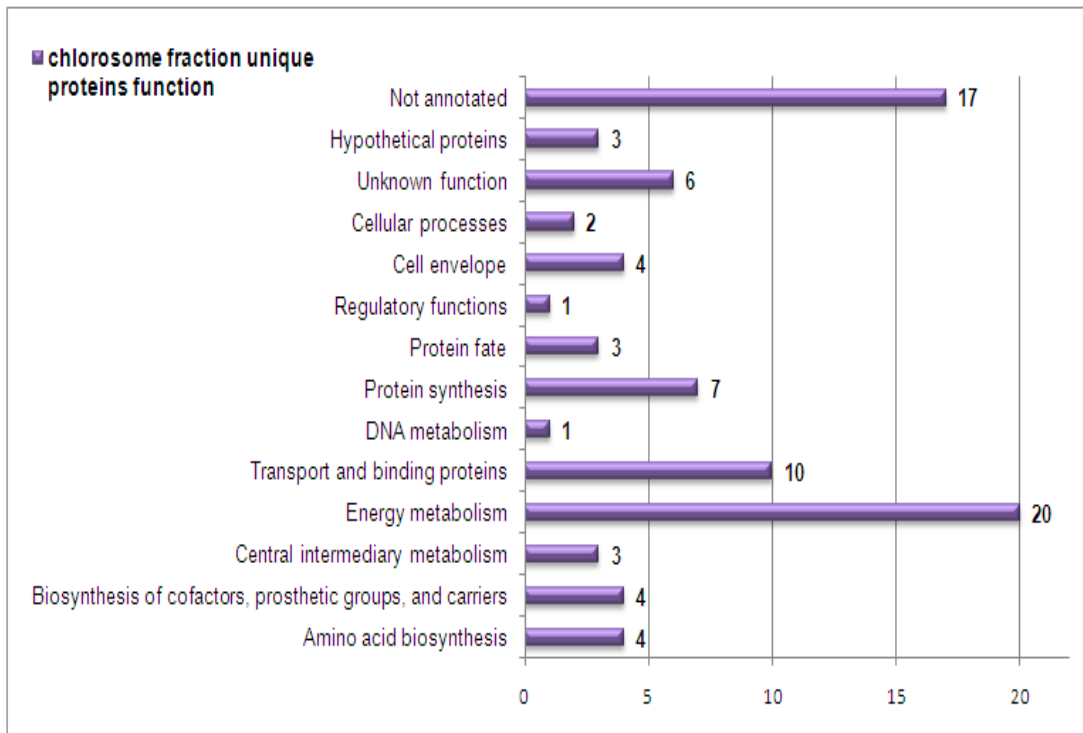


Σχήμα 21: Προβλεπόμενη κυτταρική τοπολογία των πρωτεϊνών που ταυτοποιήθηκαν αποκλειστικά στο χλωροσωμικό κλάσμα.

Αναφορικά με τη λειτουργική κατηγοριοποίηση των πρωτεϊνών που βρέθηκαν μόνο στο χλωροσωμικό κλάσμα (Σχήμα 22), το μεγαλύτερο ποσοστό αυτών εμπλέκονται στο μεταβολισμό της ενέργειας, ενώ ένας σημαντικός αριθμός τους είναι υποθετικές, γεγονός αναμενόμενο, καθώς το ένα τρίτο των πρωτεϊνών που προβλέπεται να κωδικοποιούνται από το γονιδίωμα του *Chl. tepidum* είναι υποθετικές πρωτεΐνες.

Σημειώνεται επίσης ότι ο μεταβολισμός του θείου και η φωτοσυνθετική διεργασία δεν

έχουν διασαφηνιστεί πλήρως ακόμα, άρα ενδεχομένως κάποιες από τις υποθετικές πρωτεΐνες που ταυτοποιήθηκαν στο χλωροσωμικό κλάσμα να εμπλέκονται στα μονοπάτια αυτά.



Σχήμα 22: Προβλεπόμενη λειτουργία των πρωτεϊνών που ταυτοποιήθηκαν μόνο στο χλωροσωμικό κλάσμα.

4. ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

4.1 Ανάλυση και χαρακτηρισμός χλωροσωμικού κλάσματος

Το μέγεθος των χλωροσωμάτων αποτελεί έναν σημαντικό τομέα της έρευνας των πράσινων θειούχων βακτηρίων. Διαφορετικές μέθοδοι έχουν εφαρμοστεί για την διερεύνηση του θέματος αυτού περιλαμβάνοντας DLS, TEM και AFM [56-58]. Με βάση τα αποτελέσματα της ηλεκτρονικής μικροσκοπίας διέλευσης, τα απομονωμένα χλωροσώματα εμφανίζονται να έχουν μήκος 110–180 nm και διάμετρο 40–60 nm, ενώ με βάση την AFM, τα απομονωμένα χλωροσώματα εμφανίζονται λίγο μεγαλύτερα, δηλαδή 170-260 nm σε μήκος, 90-160 nm σε πλάτος και 30-40 nm σε ύψος. Η ηλεκτρονική μικροσκοπία χρησιμοποιείται για την λήψη πληροφοριών για το σχήμα και το μέγεθος των χλωροσωμάτων, αλλά δεν παρουσιάζει μια εικόνα των χλωροσωμάτων στην εγγενή τους μορφή. Συνεπώς η DLS απεδείχθη συμπληρωματική της τεχνική και στάθηκε ικανή να προσφέρει πολύ πιο λεπτομερείς, ποσοτικές πληροφορίες σε σχέση με τη μορφολογία των χλωροσωμάτων. Με βάση τα αποτελέσματα των δύο μεθόδων, οι διαστάσεις που ελήφθησαν από το TEM ήταν μεγαλύτερες από αυτές που υπολογίστηκαν μέσω DLS. Η διαφοροποίηση αυτή ενδεχομένως να οφείλεται στην αλληλεπίδραση των χλωροσωμάτων με το στρώμα άνθρακα των δισκίων χαλκού που πιθανόν να προκαλεί κάποια παραμόρφωση των χλωροσωμάτων. Αξίζει ωστόσο να σημειωθεί ότι οι διαστάσεις που ελήφθησαν από το TEM είναι παρόμοιες με μετρήσεις

σε προηγούμενες μελέτες [59], ενώ οι μετρήσεις με DLS εμφανίζουν την διάμετρο των χλωροσωμάτων μεγαλύτερη από αυτήν που αναφέρεται στη βιβλιογραφία [58].

Στην προσπάθεια απόκτησης μιας πιο λεπτομερούς εικόνας για την πρωτεϊνική σύνθεση των χλωροσωμάτων του *Chlorobium tepidum*, πραγματοποιήθηκε nLC-ESI-MS/MS ανάλυση των απομονωμένων χλωροσωμικών δειγμάτων. Σύμφωνα με προηγούμενες μελέτες, τα χλωροσώματα του *Chl. tepidum* περιέχουν 10 πρωτεΐνες, εκ των οποίων η CsmA η οποία προσδένει ένα μόριο BChl a ανά μονομερές αντιστοιχεί σχεδόν στη μισή ολική πρωτεϊνική συγκέντρωση χλωροσώματα. Στο *Chl. tepidum*, εννέα ακόμα πρωτεΐνες παρίστανται σε μικρότερες συγκεντρώσεις (CsmB, CsmC, CsmD, CsmE, CsmF, CsmH, CsmI, CsmJ, CsmX), χωρίς οι ρόλοι τους όμως να είναι εξίσου καλά χαρακτηρισμένοι. Αν και με βάση τα δεδομένα της φασματομετρίας μάζας δεν είναι δυνατόν να υπολογιστεί η ακριβής στοιχειομετρία των πρωτεϊνών, συγκρίνοντας τον αριθμό των ταυτοποιημένων πεπτιδίων ανά πρωτεΐνη μπορεί να εξαχθεί το συμπέρασμα ότι η CmsA φαίνεται να είναι η πρωτεΐνη που βρίσκεται σε μεγαλύτερη αφθονία, σε σχέση πάντα με τις λοιπές χλωροσωμικές πρωτεΐνες.

Σε προηγούμενες μελέτες, κάποιες επιπλέον πρωτεΐνες απομονώθηκαν στο χλωροσωμικό κλάσμα, και πιο συγκεκριμένα αναφορές υπάρχουν για τις AtpF (CT0019), μια heat shock protein (Hsp20 family, CT1970) και για μια putative outer surface protein (CT2144) [14]. Πρόσφατα οι μεταφορικές πρωτεΐνες (CT0104 and CT0105) αναφέρθηκαν σε *Chl. tepidum bchK* mutants, που δεν είχαν την ικανότητα να παράγουν BChl_c. Αυτές οι πρωτεΐνες είναι ομόλογα των EmrB και EmrA ενός μεταφορικού συμπλόκου στο *E. coli* που είναι υπεύθυνο για την εξαγωγή υδρόφοβων, ξενοβιοτικών ενώσεων από το κυτταρόπλασμα [4], γεγονός το οποίο υποδεικνύει ότι τα χλωροσώματα πιθανόν περιέχουν επιπλέον πρωτεΐνες εκτός των 10 αναφερόμενων χλωροσωμικών πρωτεϊνών. Η ταυτοποίηση των CT1970 και CT2144 στην παρούσα πρωτεομική μελέτη ενισχύει την υπόθεση της ύπαρξης επιπλέον πρωτεϊνών στο χλωρόσωμα. Επίσης στο χλωροσωμικό κλάσμα, ταυτοποιήθηκαν και πρωτεΐνες που εμπλέκονται στη φωτοσύνθεση, όπως οι PscA και PscC υπομονάδες του κέντρου αντίδρασης του φωτοσυστήματος I, η FMO πρωτεΐνη, η Rubisco-like protein (CT1772) καθώς και κυτοχρώματα για τη μεταφορά ηλεκτρονίων (cyt c, putative (CT1704), cyt c-

555 membrane-bound (CT0073), cyt bc complex, cyt b subunit (CT0303) και cyt b6-f complex iron-sulfur subunit (CT0302)). Τέλος, ταυτοποιήθηκαν οι υπομονάδες άλφα, βήτα και γάμμα του υδρόφιλου καταλυτικού τμήματος F1 της πρωτονοκινητήριας ATP συνθάσης.

Τέλος στο χλωροσωμικό κλάσμα ταυτοποιήθηκε και ένας σημαντικός αριθμός πρωτεϊνών που εμπλέκονται στην οξειδωση ανόργανων θειούχων ενώσεων στο *Chlorobium tepidum*. Συγκεκριμένα, βρέθηκαν οι πρωτεΐνες SoxZ, SoxY, SoxB, SoxA, καθώς και η putative uncharacterized protein (CT1020) η οποία σύμφωνα με πρόσφατες μελέτες [26], φέρεται να σταθεροποιεί το σύμπλοκο SoxAX. Επίσης ταυτοποιήθηκε η heterodisulfide reductase, putative (CT0868), η οποία φαίνεται να εμπλέκεται στη μεταφορά του στοιχειακού θείου διαμέσω της κυτταροπλασματικής μεμβράνης [26]. Τέλος βρέθηκαν και 12 ακόμα πρωτεΐνες που συμμετέχουν στο μεταβολισμό των θειούχων ενώσεων (Adenylylsulfate reductase, alpha subunit (CT0865) και beta subunit (CT0864), Heterodisulfide reductase, subunit A (CT0866), Heterodisulfide reductase, subunit A/hydrogenase, delta subunit, putative (CT0867), Thioredoxin-2 (CT0841), Thioredoxin reductase (CT0842), Sulfite reductase, dissimilatory-type, gamma subunit (CT0851), alpha subunit (CT0852) and, beta subunit (CT0853), Iron-sulfur cluster-binding protein, gltD family (CT0854), DsrF protein (CT0856), και Sulfate adenylyltransferase (CT0862).

Σε προηγούμενες μελέτες έχουν αναφερθεί αλληλεπιδράσεις μεταξύ εξωτερικής και κυτταροπλασματικής μεμβράνης [60]. Πρόσφατα, μελέτες ηλεκτρονικής μικροσκοπίας εμφάνισαν ένα δίκτυο που περιλαμβάνει την εξωτερική και κυτταροπλασματική μεμβράνη, τα χλωροσώματα και στοιχεία που συνδέουν τις 2 κυτταρικές μεμβράνες [61]. Τα ευρήματα της παρούσας μελέτης υποστηρίζουν αυτήν την θεωρία, καθώς το χλωρόσωμα φαίνεται να αλληλεπιδρά με πρωτεΐνες που εμπλέκονται στην μεταφορά ηλεκτρονίων και τη φωτοσύνθεση, αλλά και με υποθετικές πρωτεΐνες της εξωτερικής μεμβράνης, με μεταφορικές πρωτεΐνες, καθώς και με πρωτεΐνες εμπλεκόμενες στο μεταβολισμό της ενέργειας και στην αναδίπλωση των πρωτεϊνών. Επιπρόσθετα εμφανίζονται και 27 ριβοσωμικές πρωτεΐνες, δεδομένου όμως της παρουσίας των περισσοτέρων εξ αυτών και στα λοιπά πρωτεϊνικά κλάσματα που αναλύθηκαν,

συμπεραίνεται ότι οι ριβοσωμικές πρωτεΐνες βρίσκονται σε μεγάλη αφθονία στο πρωτόμα του βακτηρίου και η απομάκρυνση τους αποτελεί πρόκληση.

Με βάση των αριθμών των φασμάτων που οδήγησαν στην ταυτοποίηση των πρωτεϊνών που χλωροσωμικού κλάσματος μπορεί να εξαχθεί το συμπέρασμα ότι κάποιες από αυτές τις πρωτεΐνες βρίσκονται σε μεγαλύτερη αφθονία στα χλωροσώματα, ωστόσο η φαινομενική αυτή αφθονία ίσως οφείλεται στη φύση των προκυπτόντων πεπτιδίων που ευνοεί την ταυτοποίησή τους. Παρόλα αυτά, περαιτέρω διερεύνηση απαιτείται προκειμένου να κατανοηθεί ο ρόλος των πρωτεϊνών αυτών και η πιθανή τους συσχέτιση με το χλωρόσωμα.

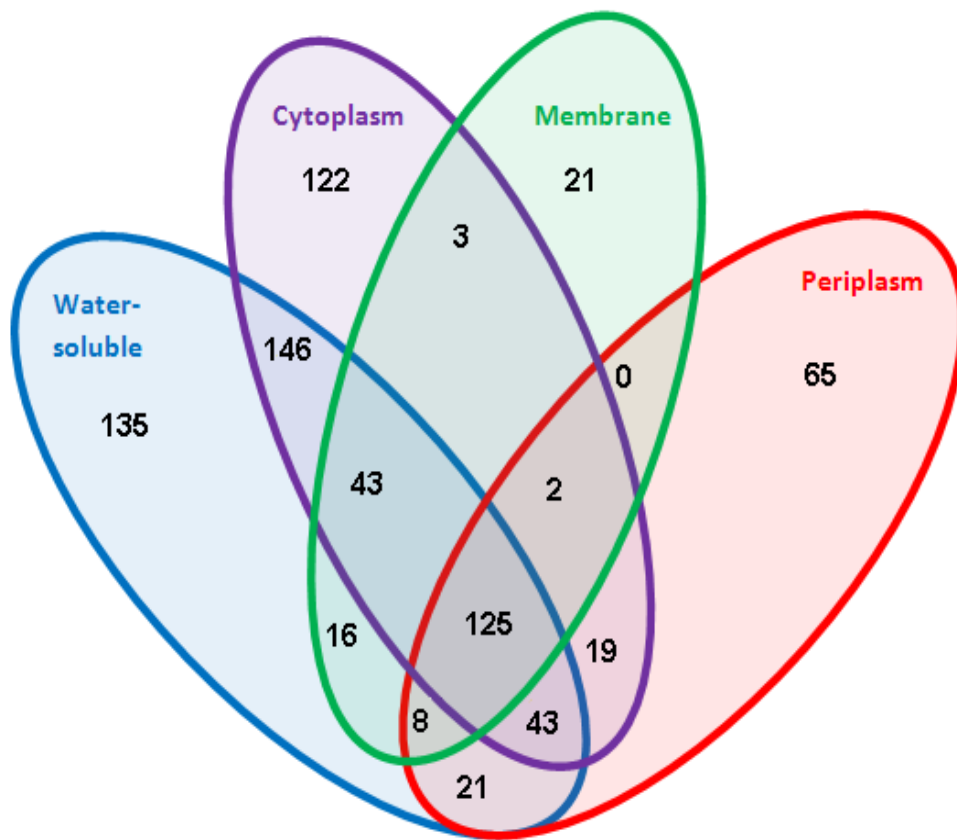
4.2 N-τερματική COFRADIC ανάλυση

Συνολικά μέσω της N-τερματικής COFRADIC ανάλυσης του συνόλου των πρωτεϊνικών κλάσμάτων που απομονώθηκαν από το *Chlorobium tepidum*, ταυτοποιήθηκαν 805 πρωτεΐνες.

	water-soluble	cytoplasm	periplasm	membrane
identifications	2700	2252	938	699
unique proteins	673	644	339	231
unique peptides	979	808	408	259
formyl - 1st AA	63 (2.3%)	24 (1.1%)	8 (0.9%)	16 (2.3%)
AcD3 - 1st AA	927 (34.3%)	868 (38.5%)	217 (23.1%)	242 (34.6%)
AcD3 - 2nd AA	455 (16.9 %)	497 (22.1%)	138 (14.7%)	105 (15%)
AcD3 internal	893 (33.1%)	614 (27.3%)	469 (50%)	268 (38.3%)
in vivo acetylated	6 (0.2%)	6 (0.3%)	19 (2%)	25 (3.6%)
formyl internal	6 (0.2%)	0	8 (0.9%)	7 (1%)
unmodified	246 (9.1%)	121 (5.4%)	66 (7%)	26 (3.7%)
pyrocarbamidomethyl	98 (3.6%)	105 (4.7%)	5 (0.5%)	9 (1.3%)
pyro - glu	6 (0.2%)	17 (0.8%)	8 (0.9%)	1 (0.1%)

Πίνακας 3: Συνοπτική παρουσίαση των ταυτοποιήσεων με N-terminal COFRADIC και των τροποποιήσεων τους.

Στον Πίνακα 3 παρουσιάζονται οι ταυτοποιήσεις με N-terminal COFRADIC των επιμέρους κλασμάτων που αναλύθηκαν στην παρούσα μελέτη και οι τροποποιήσεις των πεπτιδίων που οδήγησαν στην ταυτοποίηση τους. Οι τροποποιήσεις που εμφανίζονται στις 4 τελευταίες ομάδες οφείλονται στην μεθοδολογία της N-τερματικής ανάλυσης COFRADIC, ενώ εμφανίζεται ένας σημαντικός αριθμός ταυτοποιημένων εσωτερικών πεπτιδίων, που οφείλεται εναλλακτικά σε πρωτεόλυση των πρωτεϊνών από τις οποίες προήλθαν κατά την απομόνωση τους, ή σε ατελή διαχωρισμό των N-τερματικών από τα εσωτερικά πεπτίδια κατά την N-τερματική COFRADIC ανάλυση.



Εικόνα 13: Αριθμός ταυτοποιημένων πρωτεϊνών με N-terminal COFRADIC στα πρωτεϊνικά κλάσματα όπου βρέθηκαν.

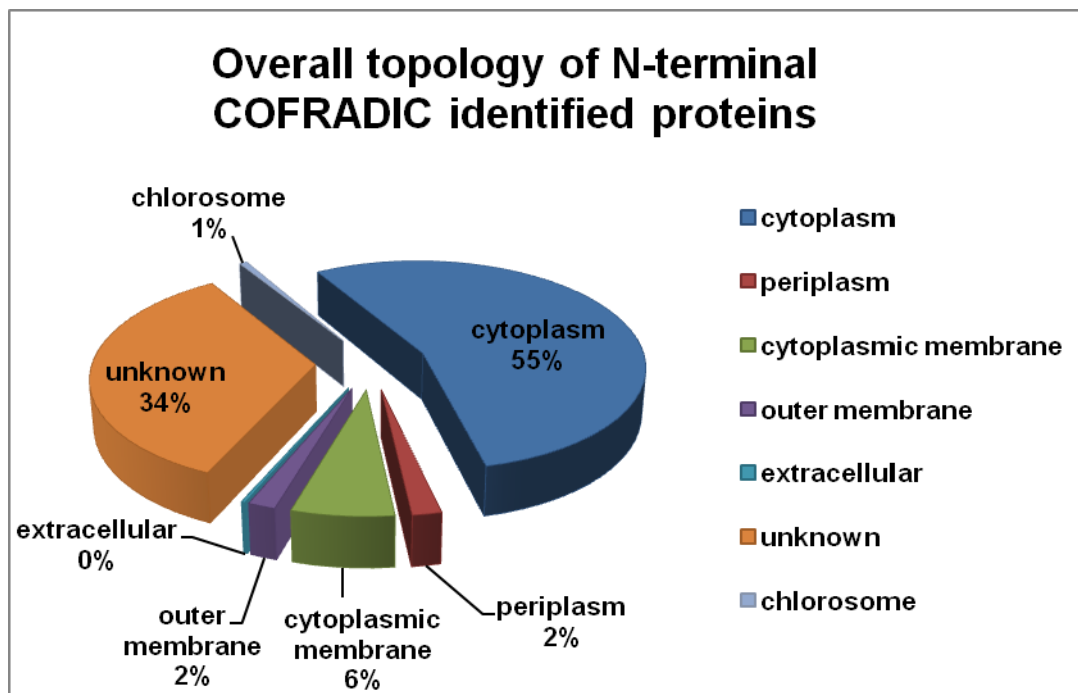
Οι 125 από τις πρωτεΐνες αυτές βρέθηκαν στο σύνολο των κλασμάτων που αναλύθηκαν, γεγονός που υποδεικνύει ότι η κλασματοποίηση που έλαβε χώρα δεν οδήγησε σε πλήρη διαχωρισμό. Ένα επιπλέον στοιχείο της ατελούς κλασματοποίησης είναι ότι 79 πρωτεΐνες ταυτοποιήθηκαν και στο υδατοδιαλυτό κλάσμα, αλλά και στα κλάσματα του κυτταροπλάσματος και του περιπλάσματος. Σημειώνεται ότι 147 πρωτεΐνες ταυτοποιήθηκαν από κοινού στο υδατοδιαλυτό και το κυτταροπλασματικό κλάσμα, ενώ κατ' αντιστοιχία 21 ήταν οι κοινές πρωτεΐνες μεταξύ υδατοδιαλυτού και περιπλασματικού κλάσματος. Τέλος, αξίζει να αναφερθεί ότι 19 πρωτεΐνες

ταυτοποιήθηκαν στα κλάσματα του κυτταροπλάσματος και του περιπλάσματος, αλλά

δεν βρέθηκαν στο υδατοδιαλυτό κλάσμα. Με βάση το φαινόμενο αυτό, και λαμβάνοντας υπόψη ότι ταυτοποιήθηκε ένας μεγάλος αριθμός πρωτεϊνών επιλεκτικά σε κάποια κλάσματα, ενισχύει ότι η επιμέρους κλασματοποίηση και συνεπώς απλοποίηση ενός πολύπλοκου πρωτεϊνικού δείγματος, όπως είναι το πρωτέομα ενός βακτηρίου, οδηγεί στην ταυτοποίηση περισσότερων πρωτεϊνών του, κάνοντας παράλληλα την διαδικασία της ανάλυσης του ευκολότερη. Συνυπολογίζοντας και την αρχή της τεχνικής του N-τερματικού COFRADIC, που βασίζεται στην όσο το δυνατόν απλούστευση του πεπτιδικού μίγματος που προκύπτει μετά την πρωτεόλυση του προς ανάλυση πρωτεϊνικού κλάσματος, επιλέγοντας-μέσω χημικών και χρωματογραφικών τεχνικών-για την ταυτοποίηση των πρωτεϊνών μόνο το N-τερματικό τους πεπτίδιο, μπορεί να προκύψει το συμπέρασμα ότι η απλούστευση του δείγματος με κλασματοποίηση του και με επιλογή πεπτιδίων αντιπροσωπευτικών των πρωτεϊνών στις οποίες ανήκουν, οδηγεί σε ταυτοποίηση μεγαλύτερου αριθμού πρωτεϊνών [62].

Παράλληλα, ο μεγαλύτερος αριθμός των ταυτοποιημένων πρωτεϊνών του μεμβρανικού κλάσματος βρέθηκε από κοινού και στα λοιπά κλάσματα, υποδηλώνοντας προσμίξεις από υδατοδιαλυτές, κυρίως κυτταροπλασματικές πρωτεΐνες στο κλάσμα των ολικών μεμβρανών που απομονώθηκε κατά την μελέτη αυτή. Το φαινόμενο αυτό μπορεί να αιτιολογηθεί καθώς κατά τη λύση του κυττάρου με διάρρηξη του, οι υδρόφοβες μεμβράνες δύνανται να λάβουν τις σταθερότερες θερμοδυναμικά διαμορφώσεις, σχηματίζοντας μικκύλια που περιβάλλουν υδατοδιαλυτό τμήμα του βακτηρίου.

Με βάση την προβλεπόμενη τοπολογία τους με μεθόδους βιοπληροφορικής [51], 440 είναι κυτταροπλασματικές, 15 περιπλασματικές, 52 κυτταροπλασματικής μεμβράνης, 14 εξωτερικής μεμβράνης, 2 εκκρινόμενες και 5 χλωροσωμικές πρωτεΐνες, ενώ 277 εμφανίστηκαν ως πρωτεΐνες άγνωστης κυτταρικής τοπολογίας (Σχήμα 23).



Σχήμα 23: Προβλεπόμενη τοπολογία των πρωτεϊνών που ταυτοποιήθηκαν με N-terminal COFRADIC.

Γενικά παρατηρείται ταυτοποίηση σε μεγαλύτερο ποσοστό υδατοδιαλυτών πρωτεϊνών, γεγονός που δικαιολογείται με βάση την δυσκολία που εμφανίζει η ανάλυση υδρόφωβων πρωτεϊνών με μεθόδους φασματομετρίας μάζας [62].

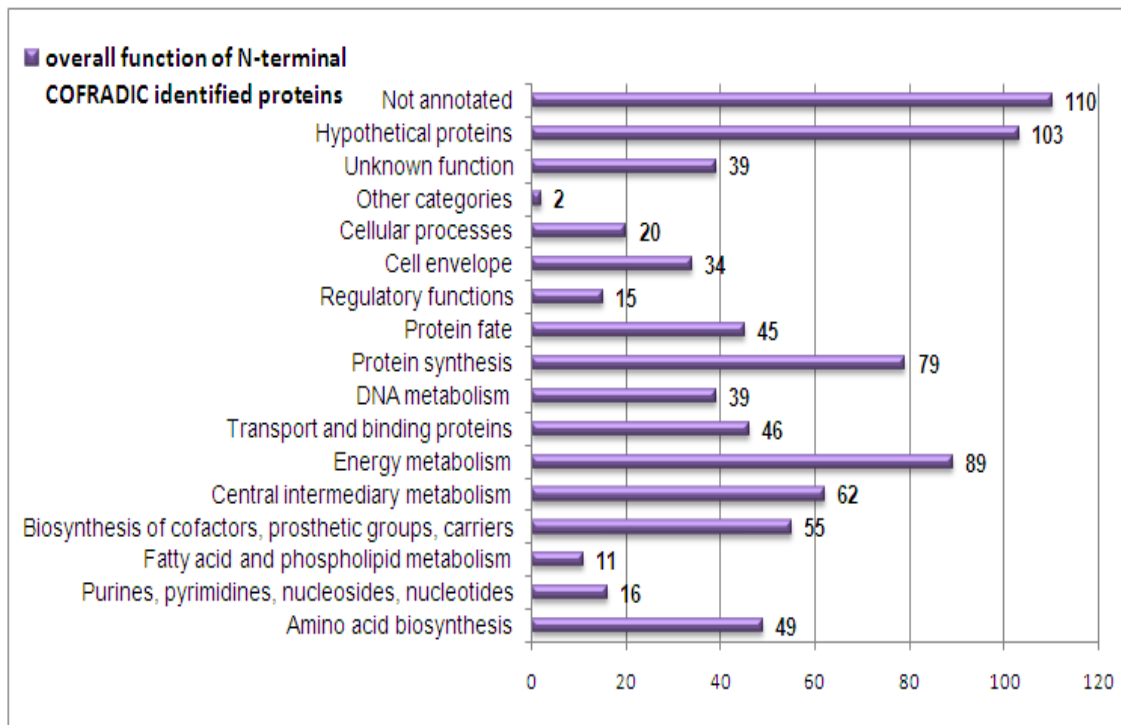
Αξιοσημείωτη ωστόσο είναι η φαινομενικά χαμηλή απόδοση σε ταυτοποίηση περιπλασματικών πρωτεϊνών που παρατηρείται, δεδομένης και της υδρόφιλης φύσης των πρωτεϊνών αυτών που καθιστά την ταυτοποίηση τους ευκολότερη. Σημειώνεται ωστόσο ότι με βάση το πρόγραμμα rsort, συνολικά 21 πρωτεΐνες προβλέπονται ως περιπλασματικές, με αποτέλεσμα οι 15 πρωτεΐνες που ταυτοποιήθηκαν στην παρούσα μελέτη και προβλέπεται ότι έχουν περιπλασματική τοπολογία να αποτελούν το 71% επί του συνόλου των προβλεπόμενων περιπλασματικών πρωτεϊνών του *Chl. tepidum*.

Αναφορικά με τις χλωροσωμικές πρωτεΐνες, συνολικά ταυτοποιήθηκαν οι 5 από τις 10 προβλεπόμενες από το γονιδίωμα ως χλωροσωμικές πρωτεΐνες, οι οποίες εμφανίστηκαν

κυρίως στο περιπλασματικό και το μεμβρανικό κλάσμα, ενώ μόνο οι 2 εμφανίστηκαν στο κυτταροπλασματικό κλάσμα, επιβεβαιώνοντας τον επιτυχή διαχωρισμό κυτταροπλάσματος-περιπλάσματος με την διαδικασία που έλαβε χώρα στην παρούσα μελέτη.

Όσον αφορά τις μεμβρανικές πρωτεΐνες που ταυτοποιήθηκαν στο σύνολο της N-τερματικής COFRADIC ανάλυσης του *Chlorobium tepidum*, παρατηρείται ότι ταυτοποιήθηκαν πολλές περισσότερες πρωτεΐνες κυτταροπλασματικής μεμβράνης σε σχέση με αυτές της εξωτερικής μεμβράνης. Ωστόσο σημειώνεται ότι με βάση το γονιδίωμα 27 συνολικά πρωτεΐνες εμφανίζουν τοπολογία εξωτερικής μεμβράνης ενώ 323 πρωτεΐνες εμφανίζονται ως κυτταροπλασματικής μεμβράνης. Συνεπώς, με την μεθοδολογία της N-τερματικής ανάλυσης COFRADIC ταυτοποιήθηκε το 51% των προβλεπόμενων πρωτεϊνών της εξωτερικής μεμβράνης και το 16% επί των προβλεπόμενων πρωτεϊνών κυτταροπλασματικής μεμβράνης του *Chl. tepidum*.

Συνοπτικά, στο Σχήμα 24 παρουσιάζεται η λειτουργική κατηγοριοποίηση των πρωτεϊνών που ταυτοποιήθηκαν μέσω N-τερματικής COFRADIC ανάλυσης συνολικά όλων των πρωτεϊνικών κλασμάτων του *Chlorobium tepidum*. Οι περισσότερες εξ' αυτών δεν έχουν γνωστό λειτουργικό ρόλο (252 πρωτεΐνες), ενώ η πλειοψηφία γνωστής λειτουργίας εμπλέκονται στον μεταβολισμό της ενέργειας, στην πρωτεϊνοσύνθεση, στον κεντρικό μεταβολισμό του κυττάρου καθώς και στη βιοσύνθεση αμινοξέων, προσθετικών ομάδων και συμπαραγόντων, στην μεταφορά και την πρόσδεση μορίων και στην πρωτεϊνική αναδίπλωση και στόχευση (protein fate).



Σχήμα 24: Προβλεπόμενη λειτουργία των ταυτοποιημένων πρωτεϊνών με N-terminal COFRADIC.

4.3 Σύνολο ταυτοποιημένων πρωτεϊνών *Chl. Terpidum*

Συνολικά στην παρούσα μελέτη, χρησιμοποιώντας μεθόδους κλασματοποίησης του πρωτεόματος του βακτηρίου *Chl. terpidum*, απομονώθηκαν τα υποπρωτεόματα του κυτταροπλάσματος, του περιπλάσματος, των ολικών μεμβρανών, των χλωροσωμάτων καθώς και το υδατοδιαλυτό πρωτεϊνικό κλάσμα. Τα κλάσματα αυτά στη συνέχεια αναλύθηκαν με N-τερματικό COFRADIC, με την εξαίρεση του χλωροσωμικού κλάσματος που αναλύθηκε με nLC-MS/MS. Από το σύνολο των διαδικασιών αυτών, ταυτοποιήθηκαν 891 πρωτεΐνες εκ των 2288 πρωτεϊνών που προβλέπεται ότι εκφράζονται στο *Chl. terpidum* με βάση το γονιδίωμα του. Το 38,9% επί του συνολικού προβλεπόμενου πρωτεόματος του *Chl. terpidum* που ταυτοποιήθηκε στην παρούσα μελέτη, αποτελεί το μεγαλύτερο ποσοστό ταυτοποιημένων πρωτεϊνών που έχει

καταγραφεί σε έως τώρα πρωτεομικές μελέτες του βακτηρίου αυτού.

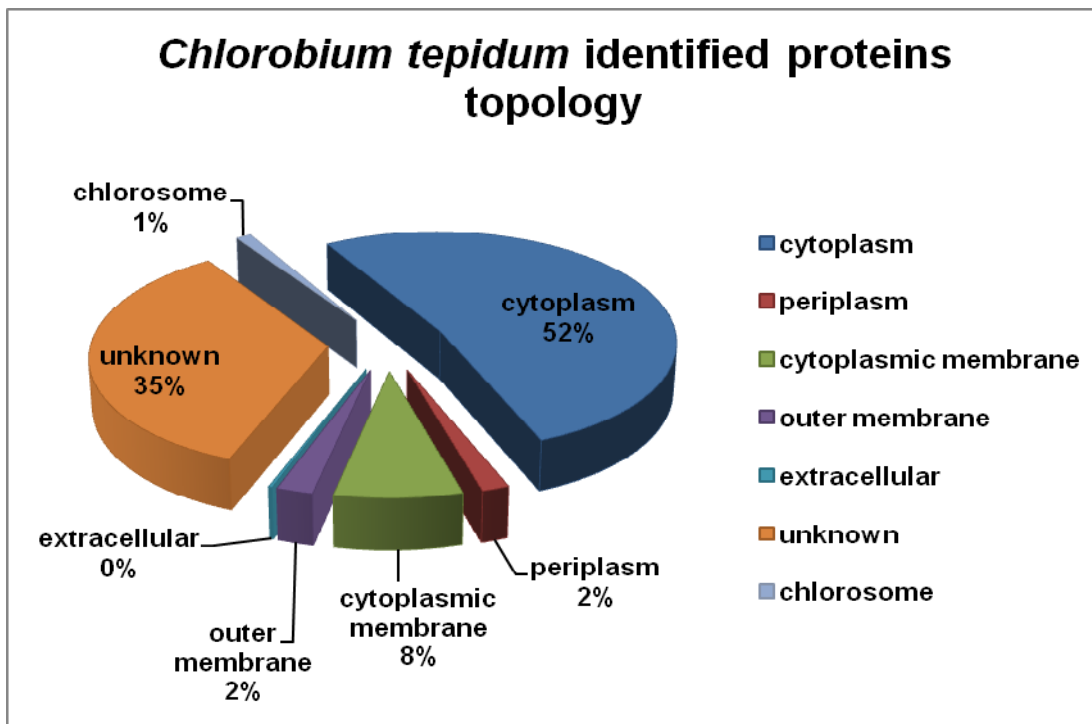
Αναφορικά με την τοπολογία των ταυτοποιημένων πρωτεϊνών, η οποία παρουσιάζεται στο Σχήμα 25, 465 πρωτεΐνες προβλέπεται ότι είναι κυτταροπλασματικές, 15 περιπλασματικές, 70 κυτταροπλασματικής μεμβράνης, 20 εξωτερικής μεμβράνης, 2 εκκρινόμενες, και 10 χλωροσωμικές πρωτεΐνες, ενώ 309 πρωτεΐνες είχαν άγνωστη τοπολογία [51-52]. Αναλογικά με την προβλεπόμενη τοπολογία του θεωρητικού πρωτεόματος του *Chl. tepidum*, ταυτοποιήθηκε το 54,3% των προβλεπόμενων κυτταροπλασματικών πρωτεϊνών, το 71,4% των προβλεπόμενων περιπλασματικών πρωτεϊνών, το 21,6% του θεωρητικού πρωτεόματος της κυτταροπλασματικής μεμβράνης, το 74% των προβλεπόμενων πρωτεϊνών εξωτερικής μεμβράνης, το 50% των εκκρινόμενων πρωτεϊνών, το 100% των προβλεπόμενων χλωροσωμικών πρωτεϊνών και το 31,2% των προβλεπόμενων ως άγνωστης τοπολογίας πρωτεϊνών [51-52].

Οι περισσότερες πρωτεΐνες γνωστής προβλεπόμενης τοπολογίας που ταυτοποιήθηκαν είναι υδρόφιλες, κυρίως κυτταροπλασματικές πρωτεΐνες, γεγονός που οφείλεται σε μεγάλο βαθμό στον εύκολο ιονισμό κατά την ανάλυση με φασματομετρία μάζας των πεπτιδίων που προκύπτουν μετά την θρυψινοποίηση των πρωτεϊνών αυτών.

Αναφορικά με τις προβλεπόμενες πρωτεΐνες εξωτερικής μεμβράνης, 7 πρωτεΐνες δεν προβλέπεται να διαθέτουν σηματοδοτική αλληλουχία και στην παρούσα μελέτη ταυτοποιήθηκαν με εσωτερικά πεπτίδια. Από τις υπόλοιπες 13 προβλεπόμενες πρωτεΐνες της εξωτερικής μεμβράνης που προβλέπεται ότι έχουν σηματοδοτική αλληλουχία, οι 5 ταυτοποιήθηκαν με βάση το N-τερματικό πεπτίδιο που προκύπτει μετά την απομάκρυνση του signal peptide. Επιπλέον 7 πρωτεΐνες, ενώ προβλέπεται ότι διαθέτουν σηματοδοτική αλληλουχία, στην παρούσα μελέτη ταυτοποιήθηκαν με εσωτερικά πεπτίδια. Τέλος, στην περίπτωση της πρωτεΐνης OmpA family protein (CT1353), ενώ προβλέπεται ότι η ακολουθία της ώριμης πρωτεΐνης μετά την αφαίρεση του signal peptide εκκινεί από το 31ο αμινοξύ της αρχικής πρωτεΐνης, κατά την N-τερματική COFRADIC ανάλυση της, η πρωτεΐνη ταυτοποιήθηκε με βάση πεπτίδιο που εκκινούσε στο 19ο αμινοξύ της.

Όσον αφορά τις 70 προβλεπόμενες πρωτεΐνες κυτταροπλασματικής μεμβράνης που

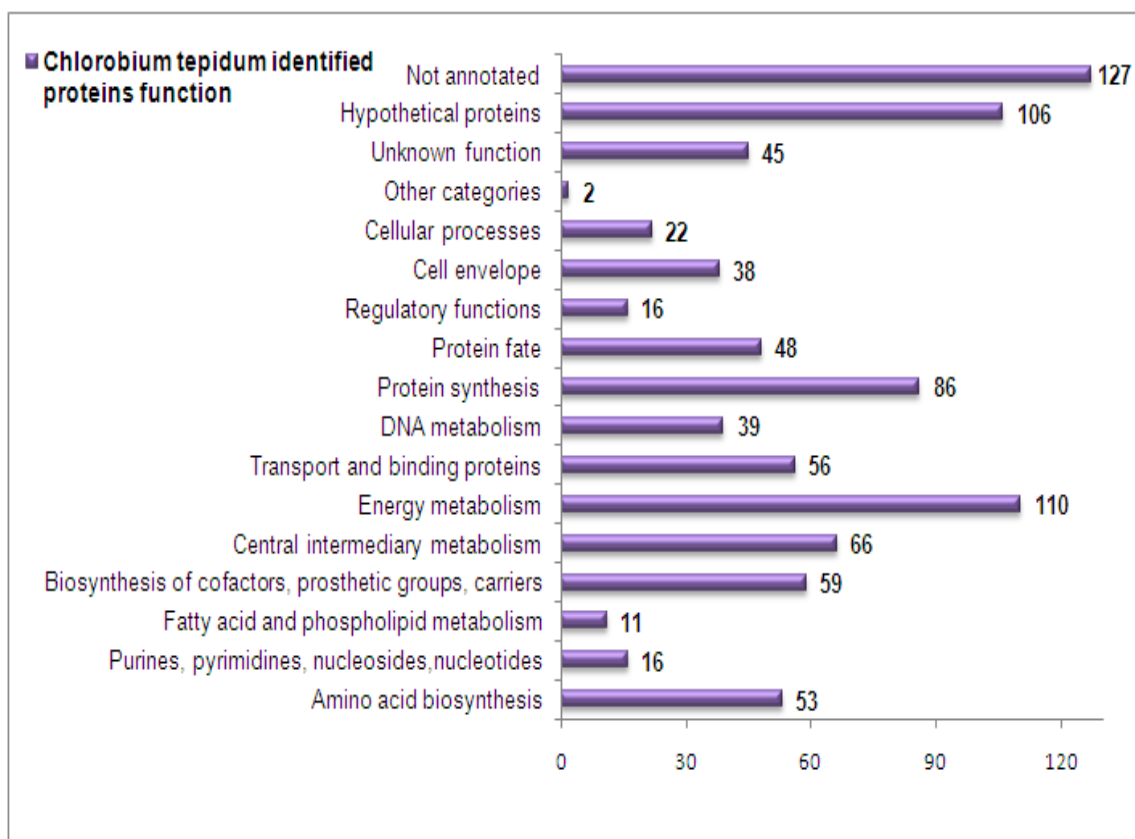
ταυτοποιήθηκαν στην παρούσα μελέτη, 10 προβλέπεται να διαθέτουν σηματοδοτική αλληλουχία, εκ των οποίων οι 8 ταυτοποιήθηκαν με εσωτερικά τους πεπτίδια. Μια πρωτεΐνη ταυτοποιήθηκε με βάση το N-τερματικό πεπτίδιο που προκύπτει μετά την απομάκρυνση της σηματοδοτικής της αλληλουχίας. Τέλος, η υποθετική πρωτεΐνη CT0917 ενώ προβλέπεται ότι η ακολουθία της ώριμης πρωτεΐνης μετά την αφαίρεση του signal peptide εκκινεί από το 39ο αμινοξύ της αρχικής πρωτεΐνης, κατά την N-τερματική COFRADIC ανάλυση της, η πρωτεΐνη ταυτοποιήθηκε με βάση το αρχικό της N-τερματικό πεπτίδιο.



Σχήμα 25: Προβλεπόμενη κυτταρική τοπολογία των ταυτοποιημένων πρωτεϊνών του *Chl. tepidum*.

Στο Σχήμα 26 παρουσιάζεται συνολικά η λειτουργική κατηγοριοποίηση των πρωτεϊνών του *Chlorobium tepidum* που ταυτοποιήθηκαν στην παρούσα μελέτη. Οι περισσότερες εξ' αυτών δεν έχουν γνωστό λειτουργικό ρόλο (278 πρωτεΐνες), ενώ η πλειοψηφία γνωστής λειτουργίας εμπλέκονται στον μεταβολισμό της ενέργειας, στην πρωτεϊνοσύνθεση, στον κεντρικό μεταβολισμό του κυττάρου καθώς και στη βιοσύνθεση προσθετικών ομάδων και συμπαραγόντων, και στην μεταφορά και την πρόσδεση

μορίων.



Σχήμα 26: Προβλεπόμενη λειτουργία των ταυτοποιημένων πρωτεϊνών του *Chl. tepidum*.

Πιο συγκεκριμένα, ταυτοποιήθηκαν συνολικά 110 πρωτεΐνες οι οποίες εμπλέκονται στο μεταβολισμό της ενέργειας, που εμπεριέχει τις διεργασίες της φωτοσύνθεσης και της μεταφοράς ηλεκτρονίων, ενώ ταυτοποιήθηκαν 66 πρωτεΐνες που συμμετέχουν στον κεντρικό μεταβολισμό του κυττάρου που εμπεριέχει την διαδικασία μεταβολισμού του θείου.

Στην παρούσα μελέτη ταυτοποιήθηκε το 47,2% των προβλεπόμενων πρωτεϊνών που σχετίζονται με τη μεταφορά ηλεκτρονίων στο *Chl. tepidum*. Στον Πίνακα 4 παρουσιάζονται οι πρωτεΐνες που προβλέπεται από το γονιδίωμα ότι εμπλέκονται στον μεταβολισμό της ενέργειας και συγκεκριμένα στην υποκατηγορία της μεταφοράς

ηλεκτρονίων στο *Chl. tepidum*.

Το κυτοχρωμικό σύμπλοκο bc του *Chl. tepidum* αποτελείται σε αντίθεση με τα υπόλοιπα φωτοσυνθετικά βακτήρια, από την πρωτεΐνη Rieske (cytochrome b6-f complex, iron-sulfur subunit, PetC, CT0302) και την υπομονάδα του κυτοχρώματος b (cytochrome b-c complex, cytochrome b subunit, PetB, CT0303) και φαίνεται ότι δεν περιέχει ένα αντίστοιχο κυτόχρωμα του κυτοχρώματος c1 που υπάρχει στους περισσότερους οργανισμούς [63-64]. Η υπομονάδα του συμπλόκου bc PetC (CT0302), ταυτοποιήθηκε σε όλα τα πρωτεϊνικά κλάσματα, ενώ η PetB (CT0303) ταυτοποιήθηκε, μόνο στο χλωροσωμικό κλάσμα, γεγονός που αποτελεί ένδειξη ότι το σύμπλοκο έχει διασπαστεί στις υπομονάδες του.

Το *Chl. tepidum* περιέχει ένα φαινομενικό οπερόνιο *ndhCHJKAIGFDB*, που κωδικοποιεί τις 11 υπομονάδες μίας τύπου I NADH αφυδρογονάσης (NDH-1). Στην μελέτη αυτή ταυτοποιήθηκαν οι υπομονάδες I (CT0771), J (CT0768) και K (CT0769) αποκλειστικά στο κυτταροπλασματικό κλάσμα, με την εξαίρεση της J (CT0768) που βρέθηκε και στο υδατοδιαλυτό κλάσμα.

Electron transport								
Gene Locus	Gene symbol	protein name	Gene Locus	Gene symbol	protein name	Gene Locus	Gene symbol	protein name
CT0073		cytochrome c-555, membrane-bound	CT1248		hydrogenase, iron-sulfur binding	CT1818	cydA	cytochrome d ubiquinol
CT0075		cytochrome c-555	CT1249	hydB-2	hydrogenase/sulfur reductase, beta	CT1819	cydB	cytochrome d ubiquinol
CT0117		sulfide-quinone reductase, putative	CT1250	hydG-2	hydrogenase/sulfur reductase,	CT1821	cydD	ABC transporter, ATP-binding
CT0168		ferredoxin, 4Fe-4S, putative	CT1260		ferredoxin, 4Fe-4S	CT1822	cydC	ABC transporter, ATP-binding
CT0188		c-type cytochrome,	CT1261		ferredoxin, 4Fe-4S	CT1891	hydB-1	hydrogenase/sulfur reductase,
CT0302	petC	cytochrome b6-f complex, iron-	CT1327		rubrerythrin	CT1892	hydG-1	hydrogenase/sulfur reductase,
CT0303	petB	cytochrome b-c complex,	CT1512		thioredoxin reductase,	CT1893	hydD	hydrogenase/sulfur reductase,
CT0322		iron-sulfur cluster-binding protein	CT1541		ferredoxin, 2Fe-2S	CT1894	hydA	hydrogenase/sulfur reductase,
CT0369	ndh	NADH dehydrogenase	CT1559	ycf5	cytochrome c biogenesis protein	CT2024	rbr-3	rubredoxin
CT0409		ferredoxin, 4Fe-4S	CT1628	nifJ	pyruvate flavodoxin/ferredoxin	CT2026		c-type cytochrome, sulfide
CT0484		NADH dehydrogenase I	CT1639	pscC	photosystem P840 reaction	CT2080	fccA	dehydrogenase, sulfide
CT0766	ndhC	NADH dehydrogenase I,	CT1655		ferredoxin, 2Fe-2S	CT2081	fccB-2	dehydrogenase, sulfide
CT0767	ndhH	NADH dehydrogenase I,	CT1701		iron-sulfur cluster-binding protein	CT2141	etfA	electron transfer flavoprotein,
CT0768	ndhJ	NADH dehydrogenase I,	CT1704		cytochrome c, putative	CT2142	etfB	electron transfer flavoprotein, beta
CT0769	ndhK	NADH dehydrogenase I,	CT1727		glutaredoxin family protein	CT2146		iron-sulfur cluster binding protein
CT0770	ndhA	NADH dehydrogenase I,	CT1734		cytochrome c, putative	CT2233		thiol:disulfide interchange
CT0771	ndhI	NADH dehydrogenase I,	CT1736		ferredoxin, 4Fe-4S	CT2242		cytochrome DsrJ
CT0772	ndhG	NADH dehydrogenase I,	CT1738	fld	flavodoxin	CT2247		iron-sulfur cluster binding protein,
CT0773	ndhE	NADH dehydrogenase I,	CT1023		thiol:disulfide interchange	CT2285		rubredoxin oxygen
CT0774	ndhF	NADH dehydrogenase I,	CT1087		sulfide-quinone reductase	CT0854		iron-sulfur cluster binding protein,
CT0775	ndhD	NADH dehydrogenase I,	CT1100	rbr-1	rubredoxin	CT0876		sulfide-quinone reductase,
CT0776	ndhB	NADH dehydrogenase I,	CT1101	rbr-2	rubredoxin	CT1013		cytochrome c peroxidase,
CT0777	hupL	uptake hydrogenase, large	CT1247	mvhD	hydrogenase, methyl-violgen-	CT1015	fccB-1	sulfide dehydrogenase,
CT0778		nickel-dependent hydrogenases b-	CT0842	trxB	thioredoxin reductase	CT0841	trx-2	thioredoxin
CT0785	trx-1	thioredoxin	CT1016	soxX	sulfur oxidation protein SoxX			

Πίνακας 4: Πρωτεΐνες που προβλέπεται με βάση το γονιδίωμα του *Chl. tepidum* ότι εμπλέκονται στην μεταφορά ηλεκτρονίων.

Στην παρούσα μελέτη ταυτοποιήθηκε το 47,7% των προβλεπόμενων πρωτεϊνών που σχετίζονται με την φωτοσύνθεση στο *Chl. tepidum*. Στον Πίνακα 5 παρουσιάζονται οι

πρωτεΐνες που προβλέπεται από το γονιδίωμα ότι εμπλέκονται στον μεταβολισμό της ενέργειας και συγκεκριμένα στην υποκατηγορία της φωτοσύνθεσης στο *Chl. tepidum*.

Ο μεταβολισμός της ενέργειας στο *Chl. tepidum* όσον αφορά την φωτοσύνθεση εκκινεί με τη συλλογή του φωτός από τα χλωροσώματα. Η πρωτεϊνική σύσταση των χλωροσωμάτων με βάση το γονιδίωμα του βακτηρίου περιλαμβάνει 10 πρωτεΐνες: CsmA, CsmB, CsmC, CsmD, CsmE, CsmF, CsmH, CsmI, CsmJ, και CsmX. Συνολικά στη μελέτη αυτήν ταυτοποιήθηκαν και οι 10 αυτές πρωτεΐνες, ενώ σημειώνεται ότι στο χλωροσωμικό κλάσμα βρέθηκαν όλες οι χλωροσωμικές πρωτεΐνες, ενώ στα λοιπά πρωτεϊνικά κλάσμα ταυτοποιήθηκαν 5 εξ' αυτών. Συγκεκριμένα, οι περισσότερες από αυτές βρέθηκαν στο περιπλασματικό και το μεμβρανικό κλάσμα, ενώ στο κυτταροπλασματικό κλάσμα ταυτοποιήθηκαν μόνο 2 χλωροσωμικές πρωτεΐνες, επιβεβαιώνοντας την επιτυχία της κλασματοποίησης.

Photosynthesis								
Gene Locus	Gene symbol	protein name	Gene Locus	Gene symbol	protein name	Gene Locus	Gene symbol	protein name
CT0063		magnesium chelatase, subunit I, putative	CT1776		bacteriochlorophyll c3(1) hydratase	CT1610	bchG	bacteriochlorophyll synthase, 34 kDa subunit
CT0420		magnesium-chelatase, subunit D/I family	CT1777	bchQ	bacteriochlorophyll c8 methyltransferase	CT1639	pscC	photosystem P840 reaction center cytochrome c-551
CT0421		magnesium-chelatase, subunit I family	CT1826	bchY	chlorophyllide reductase, BchY subunit	CT1641		magnesium-chelatase, subunit D/I family
CT0641	pscD	photosystem P840 reaction center protein PscD	CT1942	csmA	chlorosome envelope protein A	CT1772	rbcL	ribulose bisphosphate carboxylase, large subunit
CT0651	csmJ	chlorosome envelope protein J	CT1943	csmC	chlorosome envelope protein C	CT2151	bchB	protochlorophyllide reductase, ChlB subunit
CT0652	csmX	chlorosome envelope protein X	CT1957	bchH-1	magnesium-protoporphyrin methyltransferase	CT2152	bchN	protochlorophyllide reductase, ChlN subunit
CT1046	csmF	chlorosome envelope protein F	CT1958	bchM	magnesium-protoporphyrin IX monomethyl ester	CT2256	bchP	geranylgeranyl hydrogenase
CT1197		chlorobiumquinone synthase BchC related protein	CT1959	bchE	magnesium-protoporphyrin IX monomethyl ester	CT1422	bchC	2-desacetyl-2-hydroxyethyl bacteriochlorophyllide A
CT1232		geranylgeranyl hydrogenase BchP, putative	CT1992	bchK	bacteriochlorophyll c synthase	CT1423	bchX	chlorophyllide reductase, BchX subunit
CT1270	chlG	bacteriochlorophyll synthase, 34 kDa subunit	CT2014	bchJ	bacteriochlorophyll synthase, 23 kDa subunit	CT1441	pucC	pucC protein
CT1295	bchH-3	magnesium-protoporphyrin methyltransferase	CT2019	pscB	photosystem P840 reaction center iron-sulfur protein	CT1499	fmoA	bacteriochlorophyll A protein
CT1296	bchD	magnesium-chelatase, subunit D	CT2020	pscA	photosystem P840 reaction center, large subunit	CT2125	bchZ	chlorophyllide reductase, bchZ subunit
CT1297	bchI	magnesium-chelatase, subunit I	CT2054	csmB	chlorosome envelope protein B	CT2150	bchL	protochlorophyllide reductase, iron-sulfur ATP-
CT1382	csmI	chlorosome envelope protein I	CT2062	csmE	chlorosome envelope protein E	CT1421	bchF	2-vinyl bacteriochlorophyllide hydratase
CT1417	csmH	chlorosome envelope protein H	CT2064	csmD	chlorosome envelope protein D			

Πίνακας 5: Πρωτεΐνες που προβλέπεται με βάση το γονιδίωμα του *Chl. tepidum* ότι εμπλέκονται στην φωτοσύνθεση.

Μετά την συλλογή της ηλιακής ενέργειας μέσω των χλωροσωμάτων, η ενέργεια μεταφέρεται υπό μορφή ηλεκτρονίων στην παρακρυσταλλική βάση, που συνδέει τα χλωροσώματα με την κυτταροπλασματική μεμβράνη και ακολούθως στην Fenna-Matthews-Olson (FMO) πρωτεΐνη, μια υδατοδιαλυτή πρωτεΐνη-αντένα που βρίσκεται

ανάμεσα στο χλωρόσωμα και το κέντρο αντίδρασης στην κυτταροπλασματική μεμβράνη. Η FMO πρωτεΐνη (CT1499) ταυτοποιήθηκε σε όλα τα κλάσματα που αναλύθηκαν με N-τερματικό COFRADIC πλην του μεμβρανικού, υποδηλώνοντας την ασθενή της αλληλεπίδραση με την κυτταροπλασματική μεμβράνη του βακτηρίου. Σημειώνεται επίσης ότι η FMO πρωτεΐνη ταυτοποιήθηκε και στο χλωροσωμικό κλάσμα.

Στη συνέχεια η ενέργεια μεταφέρεται στα κέντρα αντίδρασης (RCs) με την υπομονάδα A των οποίων αλληλεπιδρά η FMO πρωτεΐνη. Ο πυρήνας του κέντρου αντίδρασης του *Chl. tepidum* είναι ένα ετεροτριμερές που αποτελείται από δύο αντίγραφα της μεγάλης διαμεμβρανικής πρωτεΐνης PscA (CT2020) και ένα αντίγραφο της περιφερικής πρωτεΐνης PscB (CT2019). Το φωτοσυνθετικό κέντρο αντίδρασης του *Chl. tepidum* συμπληρώνεται με τις πρωτεΐνες PscC (CT1639) και PscD (CT0641). Στην μελέτη αυτή ταυτοποιήθηκε η PscA (CT2020) στο μεμβρανικό και στο χλωροσωμικό κλάσμα. Η PscC (CT1639) ταυτοποιήθηκε σε όλα τα κλάσματα εκτός από το μεμβρανικό ενώ η PscD (CT0641) δεν βρέθηκε στο χλωροσωμικό πρωτεϊνικό κλάσμα. Τέλος η PscB δεν ταυτοποιήθηκε σε κανένα κλάσμα του *Chl. tepidum*.

Η ATP συνθάση είναι ένα χαρακτηριστικό μεμβρανικό σύμπλοκο που βρίσκεται στην εσωτερική μεμβράνη των μιτοχονδρίων, στην θυλακοειδή μεμβράνη των χλωροπλαστών και στην κυτταροπλασματική μεμβράνη των βακτηρίων [65]. Η ATP συνθάση συνθέτει ATP από ADP και φωσφορικό χρησιμοποιώντας μια διαβάθμιση πρωτονίων μεταξύ των δύο πλευρών της κυτταροπλασματικής μεμβράνης, η οποία σχηματίζεται από τη μεταφορά ηλεκτρονίων κατά τη διαδικασία της φωτοσύνθεσης. Η ATP συνθάση είναι ένα μεγάλο (500 kDa) μεμβρανικό σύμπλοκο που αποτελείται από ένα διαμεμβρανικό τμήμα F₀ που ενώνεται με ένα σφαιρικό καταλυτικό τμήμα F₁ που βρίσκεται εκτός της μεμβράνης. Το καταλυτικό τμήμα F₁ αποτελείται από τρεις άλφα- και τρεις βήτα-υπομονάδες και από μια γάμα, δέλτα και έψιλον υπομονάδα. Το μεμβρανικό τμήμα F₀ αποτελείται από δυο b υπομονάδες και 10-12 c υπομονάδες.

ATP-proton motive force interconversion		
Gene Locus	Gene symbol	protein name
CT0018	atpH	ATP synthase F1, delta subunit
CT0019	atpF	ATP synthase F0, B subunit
CT0020	atpE	ATP synthase F0, C subunit
CT0021	atpB-2	ATP synthase F0, A subunit
CT1029	atpB-1	ATP synthase F0, A subunit
CT1031		ATP synthase, putative
CT1032	atpC-1	ATP synthase F1, epsilon subunit
CT1033	atpD-1	ATP synthase F1, beta subunit
CT2032	atpG	ATP synthase F1, gamma subunit
CT2033	atpA	ATP synthase F1, alpha subunit
CT2234	atpD-2	ATP synthase F1, beta subunit
CT2235	atpC-2	ATP synthase F1, epsilon subunit

Πίνακας 6: Πρωτεΐνες που προβλέπεται με βάση το γονιδίωμα ότι αποτελούν την πρωτονοκινήτρια ATP συνθάση του του *Chl. tepidum*.

Στην μελέτη αυτή ταυτοποιήθηκαν οι υπομονάδες άλφα (CT2033), βήτα (CT2234) και γάμα (CT2032) από το καταλυτικό τμήμα F1, και η υπομονάδα b (CT0019) από το μεμβρανικό τμήμα. Αναφορικά με το καταλυτικό τμήμα της ATP συνθάσης, η υπομονάδα γάμα (CT2032) βρέθηκε σε όλα τα πρωτεϊνικά κλάσματα που απομονώθηκαν από το *Chl. tepidum*. Η υπομονάδα βήτα (CT2234) δε βρέθηκε στο μεμβρανικό κλάσμα, ενώ η υπομονάδα άλφα (CT2033) ταυτοποιήθηκε αποκλειστικά στο χλωροσωμικό κλάσμα. Η υπομονάδα b (CT0019) του διαμεμβρανικού τμήματος F0 ταυτοποιήθηκε στο κυτταροπλασματικό και στο υδατοδιαλυτό κλάσμα.

Στην παρούσα μελέτη ταυτοποιήθηκε το 64,5% των προβλεπόμενων πρωτεϊνών που σχετίζονται με τον μεταβολισμό του θείου στο *Chl. tepidum*. Στον Πίνακα 7 παρουσιάζονται οι πρωτεΐνες που προβλέπεται από το γονιδίωμα ότι εμπλέκονται στον μεταβολισμό του θείου στο *Chl. tepidum*.

Sulfur metabolism								
Gene Locus	Gene symbol	protein name	Gene Locus	Gene symbol	protein name	Gene Locus	Gene symbol	protein name
CT0117		sulfide-quinone reductase, putative	CT2241		polysulfide reductase, subunit B, putative	CT1023		thiol:disulfide interchange protein, thioredoxin family
CT0494		polysulfide reductase, subunit C, putative	CT2243	dsrK	dsrK protein	CT1025		sulfide dehydrogenase, flavoprotein subunit, putative
CT0495		polysulfide reductase, subunit B, putative	CT2244	dsrM	dsrM protein	CT1087		sulfide-quinone reductase
CT0496		polysulfide reductase, subunit A, putative	CT2247		iron-sulfur cluster-binding protein, qItD family	CT1245		heterodisulfide reductase, subunit C/succinate
CT0851	dsrC-1	sulfite reductase, dissimilatory-type, gamma	CT2248	dsrB-2	sulfite reductase, dissimilatory-type, beta subunit, authentic	CT1246	hdrA-2	heterodisulfide reductase, subunit A
CT0852	dsrA-1	sulfite reductase, dissimilatory-type, alpha	CT2249	dsrA-2	sulfite reductase, dissimilatory-type, alpha subunit	CT1247	mvhD	hydrogenase, methyl-violgen-reducing type, delta subunit
CT0853	dsrB-1	sulfite reductase, dissimilatory-type, beta	CT2250	dsrC-2	sulfite reductase, dissimilatory-type, gamma subunit	CT1248		hydrogenase, iron-sulfur binding protein, putative
CT0854		iron-sulfur cluster-binding protein, qItD family	CT2251	dsrN	dsrN protein	CT1249	hydB-2	hydrogenase/sulfur reductase, beta subunit
CT0855	dsrE	dsrE protein	CT0868		heterodisulfide reductase, putative	CT1250	hydG-2	hydrogenase/sulfur reductase, gamma subunit
CT0856	dsrF	dsrF protein	CT0876		sulfide-quinone reductase, putative	CT1891	hydB-1	hydrogenase/sulfur reductase, beta subunit
CT0857	dsrH	dsrH protein	CT1015	fccB-1	sulfide dehydrogenase, flavoprotein subunit	CT1892	hydG-1	hydrogenase/sulfur reductase, gamma subunit
CT0862	sat	sulfate adenylyltransferase	CT1016	soxX	sulfur oxidation protein SoxX	CT1893	hydD	hydrogenase/sulfur reductase, delta subunit
CT0864	aspB	adenylylsulfate reductase, beta subunit	CT1017	soxY	sulfur oxidation protein SoxY	CT1894	hydA	hydrogenase/sulfur reductase, alpha subunit
CT0865	apsA	adenylylsulfate reductase, alpha subunit	CT1018	soxZ	sulfur oxidation protein SoxZ	CT2080	fccA	sulfide dehydrogenase, cytochrome subunit
CT0866	hdrA-1	heterodisulfide reductase, subunit A	CT1019	soxA	sulfur oxidation protein SoxA	CT2081	fccB-2	sulfide dehydrogenase, flavoprotein subunit
CT0867		heterodisulfide reductase, subunit A/hydrogenase, delta	CT1021	soxB	sulfur oxidation protein SoxB	CT2240		polysulfide reductase, subunit C, putative

Πίνακας 7: Πρωτεΐνες που προβλέπεται με βάση το γονιδίωμα του *Chl. tepidum* ότι εμπλέκονται στον μεταβολισμό του θείου.

Το υδρόθειο (H_2S), το θειοθειϊκό ($S_2O_3^{2-}$), και στοιχειακό θείο (S_0) αποτελούν τους πρωταρχικούς δότες ηλεκτρονίων για τη φωτοσυνθετική καθήλωση του άνθρακα στο *Chl. tepidum* [24]. Τουλάχιστον δύο μονοπάτια φαίνεται να υπάρχουν για την οξείδωση του υδρόθειου στο περίπλασμα που περιλαμβάνουν εναλλακτικά την flavocytochrome c sulfide dehydrogenase (SoxEF/CT2080–2081) και την sulfide– quinone reductase (SQR) [66].

Συνολικά, το γονίδιο που κωδικοποιεί υπομονάδα φλαβοπρωτεΐνης του φλαβοκυτοχρώματος c (SoxF) εκφράζει 3 πρωτεΐνες στο *Chl. tepidum* (CT1015, CT1025, CT2081), με τα δύο επιπλέον ομόλογα της SoxF πιθανόν να συμμετέχουν σε άλλες αντιδράσεις μεταφοράς H_2S ή και στο Sox σύστημα για την οξείδωση του θειοθειϊκού. Ομοίως, τρεις πρωτεΐνες εκφράζονται και στην περίπτωση του γονιδίου sqf στο *Chl. tepidum* (CT0117, CT0876, and CT1087). Ενδιαφέρον παρουσιάζει το γεγονός ότι καμία από τις 3 αυτές κωδικοποιημένες πρωτεΐνες δεν περιέχει την περιοχή που θεωρείται συντηρημένη και απαραίτητη για την οξείδωση του θειώδους από την SDQ [66].

Στην πεπτιδοκεντρική πρωτεομική μελέτη του *Chl. tepidum*, ταυτοποιήθηκε η κυτοχρωμική υπομονάδα της sulfide dehydrogenase (CT2080) μόνο στο περιπλασματικό κλάσμα, ενώ δεν ταυτοποιήθηκε κανένα από τα παράγωγα του sqf γονιδίου. Αναφορικά με τις πρωτεΐνες που εκφράζονται από το γονίδιο της SoxF, οι Sulfide dehydrogenases, flavoprotein subunit, putative CT1025 και CT1015 ταυτοποιήθηκαν στο υδατοδιαλυτό και το κυτταροπλασματικό κλάσμα, ενώ η CT1015 βρέθηκε και στο περιπλασματικό δείγμα.

Αναφορικά με το σύστημα οξείδωσης του θειοθειϊκού και των σχετιζόμενων με αυτό μονοπατιών μεταφοράς ηλεκτρονίων, η γονιδιακή ανάλυση του *Chl. tepidum* έχει αποκαλύψει την παρουσία μιας ομάδας από sox γονίδια (SoxAX, YZ, CD, και B), που είναι σε μεγάλο βαθμό συντηρημένα και που υποθετικά εμπλέκονται στην οξείδωση του θειοθειϊκού [3, 29]. Ωστόσο, τα soxCD γονίδια, που φέρονται να αλληλεπιδρούν με την SoxY, απουσιάζουν από το γονιδίωμα του *Chl. tepidum* [1]. Έτσι, εναλλακτικοί μηχανισμοί έχουν προταθεί για την οξείδωση του θειοθειϊκού σε SO_4^{2-} , απουσία των SoxC/D [29], που πιθανόν να περιλαμβάνουν ένα ομόλογο της SoxF. Πρόσφατα πειραματικά ευρήματα έδειξαν ότι το κυτόχρωμα c-554 παίζει το ρόλο του δέκτη

ηλεκτρονίων κατά την οξείδωση του $S_2O_3^{2-}$ [67], αλλά και η πρωτεΐνη RuBisCO-like protein (RLP) του *Chl. tepidum* (CT1772) εμπλέκεται στο μεταβολισμό του θειοθειικού [68-69]. Μια πρόσφατη μελέτη έδειξε ότι η SoxK (CT1020), η οποία συμμετέχει σε ένα σύμπλοκο με την SoxAX (CT1019, CT1016), ήταν απαραίτητη για να επιτευχθεί ο μέγιστος ρυθμός αναγωγής του κυτοχρώματος *cyt c-554* κατά την οξείδωση του θειοθειικού [26].

Στην παρούσα πρωτεομική μελέτη του *Chl. tepidum*, ταυτοποιήθηκαν όλες οι πρωτεΐνες του Sox συστήματος (CT1016-CT1019 και CT1021) καθώς και η putative uncharacterized protein (CT1020) η οποία σύμφωνα με πρόσφατες μελέτες [26], φέρεται να σταθεροποιεί το σύμπλοκο SoxAX. Η SoxA (CT1019), και η υποθετική πρωτεΐνη (CT1020) ταυτοποιήθηκαν σε όλα τα πρωτεϊνικά δείγματα, ενώ η SoxX (CT1016), με την οποία σχηματίζεται το σύμπλοκο SoxAX, ταυτοποιήθηκε μόνο στο περιπλασματικό και το χλωροσωμικό κλάσμα. Κατά αντιστοιχία η SoxY (CT1017) δε βρέθηκε στο μεμβρανικό κλάσμα, ενώ η SoxZ (CT1018) με την οποία σχηματίζεται το σύμπλοκο SoxYZ δεν ταυτοποιήθηκε στο μεμβρανικό και το περιπλασματικό κλάσμα. Τέλος η SoxB (CT1021) δε βρέθηκε στο χλωροσωμικό κλάσμα. Επιπρόσθετα ταυτοποιήθηκε η πρωτεΐνη RuBisCO-like protein (RLP, CT1772) σε όλα τα πρωτεϊνικά κλάσματα.

Επίσης ταυτοποιήθηκε η heterodisulfide reductase, putative (CT0868) σε όλα τα κλάσματα πλην του κυτταροπλασματικού, η οποία φαίνεται να εμπλέκεται στη μεταφορά του στοιχειακού θείου διαμέσω της κυτταροπλασματικής μεμβράνης [26].

5. ΣΥΜΠΕΡΑΣΜΑΤΑ ΜΕΛΛΟΝΤΙΚΟΙ ΣΤΟΧΟΙ

Ο πρόσφατος προσδιορισμός της πλήρης αλληλουχίας νουκλεϊκών οξέων του γονιδιώματος του *Chl. tepidum* [3] έχει προσφέρει την δυνατότητα διερεύνησης σε ευρεία κλίμακα, της έκφρασης των γονιδίων του *Chl. tepidum*. Η πρωτεομική στοχεύει εν μέρει στον προσδιορισμό όλων των πρωτεϊνών που εκφράζονται σε έναν οργανισμό επιτρέποντας την ανάλυση της έκφρασης γονιδίων στο επίπεδο της πρωτεϊνικής σύνθεσης.

Στην παρούσα ερευνητική εργασία συνδυάζοντας gel-free μεθόδους πρωτεομικής ανάλυσης (N-terminal COFRADIC και nLC-MS/MS) μελετήσαμε κυρίως την έκφραση αλλά και την τοπολογία πρωτεϊνών του *Chl. tepidum*. Από την πρωτεομική ανάλυση που πραγματοποιήσαμε στο *Chl. tepidum* ταυτοποιήθηκαν συνολικά 891 πρωτεΐνες από τις συνολικά 2,288 πρωτεΐνες που προβλέπονται από το γονιδίωμα [3], δηλαδή το 38.9% επί του θεωρητικού πρωτεόματος του *Chl. tepidum*. Αναλογικά με την προβλεπόμενη τοπολογία του θεωρητικού πρωτεόματος του *Chl. tepidum*, ταυτοποιήθηκε το 54,3% των προβλεπόμενων κυτταροπλασματικών πρωτεϊνών (465 πρωτεΐνες), το 71,4% των προβλεπόμενων περιπλασματικών πρωτεϊνών (15 πρωτεΐνες), το 21,6% του θεωρητικού πρωτεόματος της κυτταροπλασματικής μεμβράνης (70 πρωτεΐνες), το 74% των προβλεπόμενων πρωτεϊνών εξωτερικής μεμβράνης (20 πρωτεΐνες), το 50% των εκκρινόμενων πρωτεϊνών, το 100% των προβλεπόμενων χλωροσωμικών πρωτεϊνών (10 πρωτεΐνες), και το 31,2% (309 πρωτεΐνες), των προβλεπόμενων ως άγνωστης τοπολογίας πρωτεϊνών [51-52]

Με βάση τον προβλεπόμενο από το γονιδίωμα κυτταρικό ρόλο των ταυτοποιημένων πρωτεϊνών [3] η πλειοψηφία των πρωτεϊνών προβλέφθηκε ότι συμμετέχει στον

μεταβολισμό και συγκεκριμένα οι περισσότερες εξ' αυτών στον μεταβολισμό της ενέργειας. Ένα μεγάλο ποσοστό πρωτεϊνών συμμετέχει στην πρωτεϊνική σύνθεση, ενώ

ταυτοποιήθηκαν πειραματικά 233 υποθετικές πρωτεΐνες.

Ένας επιπλέον σκοπός της παρούσας ερευνητικής εργασίας ήταν η μελέτη των χλωροσωμάτων αναφορικά με το μέγεθος και τη πρωτεϊνική τους σύσταση. Το μέγεθος των χλωροσωμάτων αποτελεί έναν σημαντικό τομέα της έρευνας των πράσινων θειούχων βακτηρίων και χρησιμοποιώντας DLS και ηλεκτρονική μικροσκοπία προέκυψαν διαστάσεις για τα χλωροσώματα 136x90 nm και 170x70 nm αντίστοιχα, που συμφωνούν σε μεγάλο βαθμό με τη βιβλιογραφία. Αναφορικά με την πρωτεϊνική σύσταση των χλωροσωμάτων, ταυτοποιήθηκαν και οι 10 προβλεπόμενες από το γονιδίωμα ως χλωροσωμικές πρωτεΐνες, καθώς και κάποιες επιπλέον πρωτεΐνες που προτείνεται [8, 12, 14] ότι αλληλεπιδρούν με τα χλωροσώματα, σε ένα σύνολο 270 ταυτοποιημένων πρωτεϊνών στο χλωροσωμικό κλάσμα. Το μεγαλύτερο μέρος αυτών ήταν με βάση την προβλεπόμενη τοπολογία τους [51-52] κυτταροπλασματικές (120 πρωτεΐνες), ενώ 90 πρωτεΐνες προβλέπονταν ως άγνωστης τοπολογίας. Αναφορικά με τις προβλεπόμενες με βάση το γονιδίωμα [3] λειτουργίες των ταυτοποιημένων πρωτεϊνών του χλωροσωμικού κλάσματος, η πλειοψηφία αυτών εμπλέκεται στον μεταβολισμό της ενέργειας, που εμπεριέχει και τις διαδικασίες της φωτοσύνθεσης και της μεταφοράς ηλεκτρονίων.

Συμπερασματικά, η πρωτεομική μελέτη του πράσινου θειούχου βακτηρίου *Chlorobium tepidum* με κλασματοποίηση του πρωτεόματος του και με χαρακτηρισμό των επιμέρους υποπρωτεομάτων με N-terminal COFRADIC και nLC-MS/MS, οδήγησε στην ταυτοποίηση του μεγαλύτερου ποσοστού πρωτεϊνών επί του θεωρητικού πρωτεόματος του βακτηρίου αυτού που έχει αναφερθεί μέχρι σήμερα. Ο συνδυασμός απλοποίησης του πρωτεϊνικού δείγματος με την κλασματοποίηση του και gel-free τεχνικών που επικεντρώνονται στην ανάλυση συγκεκριμένων τάξεων πεπτιδίων, απλοποιώντας περαιτέρω το προς ανάλυση δείγμα, συνεπώς αποδεικνύεται ο πλέον αποτελεσματικός για την πρωτεομική μελέτη πολύπλοκων πρωτεϊνικών δειγμάτων όπως είναι το ολικό πρωτέομα του *Chl. tepidum*.

Μελλοντικοί στόχοι της μελέτης αυτής αποτελούν ο επιπλέον καθαρισμός του χλωροσωμικού κλάσματος για να διευκρινιστεί κατά πόσο μπορούν να απομακρυνθούν τυχόν προσμίξεις από το δείγμα αυτό χωρίς να επηρεαστεί το χλωροσώμα στη

φυσιολογία του. Η μελέτη της πρωτεϊνικής σύστασης του μεμβρανικού κλάσματος του *Chl. tepidum*, το οποίο προκύπτει μετά την απομάκρυνση του χλωροσωμικού κλάσματος αποτελεί έναν επιπλέον μελλοντικό στόχο, καθώς η έως τώρα μελέτη του ολικού μεμβρανικού κλάσματος δυσχεραινόταν από την παρουσία των χλωροσωμάτων στο δείγμα. Τέλος, η μελέτη αυτή αποσκοπεί στην ανάπτυξη του βακτηρίου παρουσία διαφορετικών μορφών και συγκεντρώσεων θειούχων ενώσεων και την πρωτεομική μελέτη των κυττάρων αυτών, για να διερευνηθούν οι διαφοροποιήσεις στις ταυτοποιημένες πρωτεΐνες του *Chl. tepidum* που εμπλέκονται στο μεταβολισμό του θείου.

Βιβλιογραφία

1. Frigaard, N.U. and D.A. Bryant, *Seeing green bacteria in a new light: genomics-enabled studies of the photosynthetic apparatus in green sulfur bacteria and filamentous anoxygenic phototrophic bacteria*. Arch Microbiol, 2004. **182**(4): p. 265-76.
2. Hauska, G., et al., *The reaction center of green sulfur bacteria(1)*. Biochim Biophys Acta, 2001. **1507**(1-3): p. 260-77.
3. Eisen, J.A., et al., *The complete genome sequence of Chlorobium tepidum TLS, a photosynthetic, anaerobic, green-sulfur bacterium*. Proc Natl Acad Sci U S A, 2002. **99**(14): p. 9509-14.
4. Frigaard, N.U., et al., *Chlorobium tepidum: insights into the structure, physiology, and metabolism of a green sulfur bacterium derived from the complete genome sequence*. Photosynth Res, 2003. **78**(2): p. 93-117.
5. Bryant, D.A. and N.U. Frigaard, *Prokaryotic photosynthesis and phototrophy illuminated*. Trends Microbiol, 2006. **14**(11): p. 488-96.
6. Imhoff, J.F., *Phylogenetic taxonomy of the family Chlorobiaceae on the basis of 16S rRNA and fmo (Fenna-Matthews-Olson protein) gene sequences*. Int J Syst Evol Microbiol, 2003. **53**(Pt 4): p. 941-51.
7. Wahlund, T.M. and M.T. Madigan, *Genetic transfer by conjugation in the thermophilic green sulfur bacterium Chlorobium tepidum*. J Bacteriol, 1995. **177**(9): p. 2583-8.
8. Hohmann-Marriott, M.F. and R.E. Blankenship, *Hypothesis on chlorosome biogenesis in green photosynthetic bacteria*. FEBS Lett, 2007. **581**(5): p. 800-3.
9. Gupta, R.S., T. Mukhtar, and B. Singh, *Evolutionary relationships among photosynthetic prokaryotes (Heliobacterium chlorum, Chloroflexus aurantiacus, cyanobacteria, Chlorobium tepidum and proteobacteria): implications regarding the origin of photosynthesis*. Mol Microbiol, 1999. **32**(5): p. 893-906.
10. Olson, J.M., M. Miller, and J. D'Oliesslager, *The asymmetry of P+ in bacterial reaction centers revealed by circular dichroism spectroscopy*. Biochemistry, 1995. **34**(46): p. 15230-4.
11. Arellano, J.B., et al., *Hexanol-induced order-disorder transitions in lamellar self-assembling aggregates of bacteriochlorophyll c in Chlorobium tepidum chlorosomes*. Langmuir, 2008. **24**(5): p. 2035-41.

12. Li, H., N.U. Frigaard, and D.A. Bryant, *Molecular contacts for chlorosome envelope proteins revealed by cross-linking studies with chlorosomes from Chlorobium tepidum*. *Biochemistry*, 2006. **45**(30): p. 9095-103.
13. Buchanan, S.K., et al., *Crystal structure of the outer membrane active transporter FepA from Escherichia coli*. *Nat Struct Biol*, 1999. **6**(1): p. 56-63.
14. Vassilieva, E.V., et al., *Subcellular localization of chlorosome proteins in Chlorobium tepidum and characterization of three new chlorosome proteins: CsmF, CsmH, and CsmX*. *Biochemistry*, 2002. **41**(13): p. 4358-70.
15. Bryant, D.A., et al., *Selective protein extraction from Chlorobium tepidum chlorosomes using detergents. Evidence that CsmA forms multimers and binds bacteriochlorophyll a*. *Biochemistry*, 2002. **41**(48): p. 14403-11.
16. Psencik, J., et al., *Lamellar organization of pigments in chlorosomes, the light harvesting complexes of green photosynthetic bacteria*. *Biophys J*, 2004. **87**(2): p. 1165-72.
17. Pedersen, M.O., et al., *The three-dimensional structure of CsmA: a small antenna protein from the green sulfur bacterium Chlorobium tepidum*. *FEBS Lett*, 2008. **582**(19): p. 2869-74.
18. Remigy, H.W., et al., *The reaction center complex from the green sulfur bacterium Chlorobium tepidum: a structural analysis by scanning transmission electron microscopy*. *J Mol Biol*, 1999. **290**(4): p. 851-8.
19. Oh-oka, H., et al., *The protein responsible for center A/B in spinach photosystem I: isolation with iron-sulfur cluster(s) and complete sequence analysis*. *J Biochem*, 1988. **103**(6): p. 962-8.
20. Fulco, A.J., *P450BM-3 and other inducible bacterial P450 cytochromes: biochemistry and regulation*. *Annu Rev Pharmacol Toxicol*, 1991. **31**: p. 177-203.
21. Meyer, T.E., et al., *The cytochromes of Chlorobium thiosulfatophilum*. *Biochim Biophys Acta*, 1968. **153**(4): p. 854-61.
22. Kusai, A. and T. Yamanaka, *Cytochrome c (553, Chlorobium thiosulfatophilum) is a sulphide-cytochrome c reductase*. *FEBS Lett*, 1973. **34**(2): p. 235-7.
23. Brune, D.C., *Isolation and characterization of sulfur globule proteins from Chromatium vinosum and Thiocapsa roseopersicina*. *Arch Microbiol*, 1995. **163**(6): p. 391-9.
24. Holt, J.A., *The metabolism of sulphur in relation to the biochemistry of cystine and cysteine: its fundamental importance in biology. A cyclic interchange between their mono- and di-sulphides is the unique reaction creating life and intelligence*. *Med Hypotheses*, 2001. **56**(5): p. 658-76.
25. Chan, L.K., et al., *A genomic region required for phototrophic thiosulfate oxidation in the green sulfur bacterium Chlorobium tepidum (syn. Chlorobaculum tepidum)*. *Microbiology*, 2008. **154**(Pt 3): p. 818-29.
26. Ogawa, T., et al., *SoxAX binding protein, a novel component of the thiosulfate-oxidizing multienzyme system in the green sulfur bacterium Chlorobium tepidum*. *J Bacteriol*, 2008. **190**(18): p. 6097-110.
27. Rother, D., et al., *Novel genes of the sox gene cluster, mutagenesis of the flavoprotein SoxF, and evidence for a general sulfur-oxidizing system in Paracoccus pantotrophus GB17*. *J Bacteriol*, 2001. **183**(15): p. 4499-508.
28. Friedrich, C.G., et al., *Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism?* *Appl Environ Microbiol*, 2001. **67**(7): p. 2873-82.

29. Gomez Maqueo Chew, A., N.U. Frigaard, and D.A. Bryant, *Identification of the bchP gene, encoding geranylgeranyl reductase in Chlorobaculum tepidum*. J Bacteriol, 2008. **190**(2): p. 747-9.
30. Naaby-Hansen, S., M.D. Waterfield, and R. Cramer, *Proteomics--post-genomic cartography to understand gene function*. Trends Pharmacol Sci, 2001. **22**(7): p. 376-84.
31. de Hoog, C.L. and M. Mann, *Proteomics*. Annu Rev Genomics Hum Genet, 2004. **5**: p. 267-93.
32. Grant, S.G. and W.P. Blackstock, *Proteomics in neuroscience: from protein to network*. J Neurosci, 2001. **21**(21): p. 8315-8.
33. Graves, P.R. and T.A. Haystead, *Molecular biologist's guide to proteomics*. Microbiol Mol Biol Rev, 2002. **66**(1): p. 39-63; table of contents.
34. Stults, J.T. and D. Arnott, *Proteomics*. Methods Enzymol, 2005. **402**: p. 245-89.
35. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules*. Science, 1989. **246**(4926): p. 64-71.
36. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. Anal Chem, 1988. **60**(20): p. 2299-301.
37. Gevaert, K., et al., *A la carte proteomics with an emphasis on gel-free techniques*. Proteomics, 2007. **7**(16): p. 2698-718.
38. Sandra, K., et al., *Combination of COFRADIC and high temperature-extended column length conventional liquid chromatography: a very efficient way to tackle complex protein samples, such as serum*. J Sep Sci, 2007. **30**(5): p. 658-68.
39. Staes, A., et al., *Improved recovery of proteome-informative, protein N-terminal peptides by combined fractional diagonal chromatography (COFRADIC)*. Proteomics, 2008. **8**(7): p. 1362-70.
40. Van Damme, P., et al., *A review of COFRADIC techniques targeting protein N-terminal acetylation*. BMC Proc, 2009. **3 Suppl 6**: p. S6.
41. Gevaert, K., et al., *Diagonal reverse-phase chromatography applications in peptide-centric proteomics: ahead of catalogue-omics?* Anal Biochem, 2005. **345**(1): p. 18-29.
42. Wahlund, T.M., et al., *A thermophilic green sulfur bacterium from New Zealand hot springs, Chlorobium tepidum sp. nov.* Arch Microbiol, 1991. **156**: p. 81-90.
43. Gerola, P.D. and J.M. Olson, *A new bacteriochlorophyll a-protein complex associated with chlorosomes of green sulfur bacteria*. Biochim Biophys Acta, 1986. **848**(1): p. 69-76.
44. Feick, R.G., M. Fitzpatrick, and R.C. Fuller, *Isolation and characterization of cytoplasmic membranes and chlorosomes from the green bacterium Chloroflexus aurantiacus*. J Bacteriol, 1982. **150**(2): p. 905-15.
45. Aivaliotis, M., et al., *Molecular size determination of a membrane protein in surfactants by light scattering*. Biochim Biophys Acta, 2003. **1615**(1-2): p. 69-76.
46. Rizos, A.K., Tsatsakis, A. M., Shtilman, M. I. & Brown, W., *Dynamic light scattering study of 1-naphthylacetic acid polymeric derivative, a novel plant hormone*. Polymer, 1998(39): p. 4729-4734.
47. Schagger, H. and G. von Jagow, *Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa*. Anal Biochem, 1987. **166**(2): p. 368-79.
48. Celis, J.E., et al., *The human keratinocyte two-dimensional gel protein database: update 1993*. Electrophoresis, 1993. **14**(11): p. 1091-198.

49. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Anal Biochem, 1976. **72**: p. 248-54.
50. Staes, A., et al., *Assessing a novel microfluidic interface for shotgun proteome analyses*. J Sep Sci, 2007. **30**(10): p. 1468-76.
51. Nakai, K. and P. Horton, *PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization*. Trends Biochem Sci, 1999. **24**(1): p. 34-6.
52. Nakai, K., *Protein sorting signals and prediction of subcellular localization*. Adv Protein Chem, 2000. **54**: p. 277-344.
53. Nielsen, H., et al., *Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites*. Protein Eng, 1997. **10**(1): p. 1-6.
54. Nielsen, H., S. Brunak, and G. von Heijne, *Machine learning approaches for the prediction of signal peptides and other protein sorting signals*. Protein Eng, 1999. **12**(1): p. 3-9.
55. Pedersen, M.O., et al., *A reconstituted light-harvesting complex from the green sulfur bacterium Chlorobium tepidum containing CsmA and bacteriochlorophyll a*. Biochemistry, 2008. **47**(5): p. 1435-41.
56. Golecki, J.R.O., J., *Quantitative relationships between bacteriochlorophyll content, cytoplasmic membrane structure and chlorosome size in Chloroflexus aurantiacus*. Arch Microbiol, 1987(148): p. 236-241.
57. Martinez-Planells, A., et al., *Determination of the topography and biometry of chlorosomes by atomic force microscopy*. Photosynth Res, 2002. **71**(1-2): p. 83-90.
58. Wang, Z.-Y., Marx, G., Umetsu, M., Kobayashi, M., Mimuro, M. & Nozawa, T., *Morphology and spectroscopy of chlorosomes from Chlorobium tepidum by alcohol treatments*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1995(1232): p. 187-196.
59. Frigaard, N.U., et al., *Nine mutants of Chlorobium tepidum each unable to synthesize a different chlorosome protein still assemble functional chlorosomes*. J Bacteriol, 2004. **186**(3): p. 646-53.
60. Bayer, M.E., ed. *The fusion sites between outer membrane and cytoplasmic membrane of bacteria: their role in membrane assembly and virus infection*. Bacterial Outer Membranes, Biogenesis and Functions, ed. M. Inouye. 1979, John Wiley & Sons: New York. 167-202.
61. Hohmann-Marriott, M.F., R.E. Blankenship, and R.W. Roberson, *The ultrastructure of Chlorobium tepidum chlorosomes revealed by electron microscopy*. Photosynth Res, 2005. **86**(1-2): p. 145-54.
62. Santoni, V., M. Molloy, and T. Rabilloud, *Membrane proteins and proteomics: un amour impossible?* Electrophoresis, 2000. **21**(6): p. 1054-70.
63. Harnisch, U., H. Weiss, and W. Sebald, *The primary structure of the iron-sulfur subunit of ubiquinol-cytochrome c reductase from Neurospora, determined by cDNA and gene sequencing*. Eur J Biochem, 1985. **149**(1): p. 95-9.
64. Gabellini, N. and W. Sebald, *Nucleotide sequence and transcription of the fbc operon from Rhodospseudomonas sphaeroides. Evaluation of the deduced amino acid sequences of the FeS protein, cytochrome b and cytochrome c1*. Eur J Biochem, 1986. **154**(3): p. 569-79.
65. Yoshida, M., E. Muneyuki, and T. Hisabori, *ATP synthase--a marvellous rotary engine of the cell*. Nat Rev Mol Cell Biol, 2001. **2**(9): p. 669-77.

66. Azai, C., et al., *Sulfur oxidation in mutants of the photosynthetic green sulfur bacterium Chlorobium tepidum devoid of cytochrome c-554 and SoxB*. Photosynth Res, 2009. **100**(2): p. 57-65.
67. Tsukatani, Y., et al., *Parallel electron donation pathways to cytochrome c(z) in the type I homodimeric photosynthetic reaction center complex of Chlorobium tepidum*. Biochim Biophys Acta, 2008. **1777**(9): p. 1211-7.
68. Hanson, T.E. and F.R. Tabita, *A ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO)-like protein from Chlorobium tepidum that is involved with sulfur metabolism and the response to oxidative stress*. Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4397-402.
69. Hanson, T.E. and F.R. Tabita, *Insights into the stress response and sulfur metabolism revealed by proteome analysis of a Chlorobium tepidum mutant lacking the Rubisco-like protein*. Photosynth Res, 2003. **78**(3): p. 231-48.

IIAPAPTHMA

Protein Name	Score	Pepti	EXPASY	Gene	Theoretic	Subcellular	Cellular function from
chlorosome fraction							
50S ribosomal protein	61 (27)	7- 16	Q8KG17	CT01	19282.5	unknown	Protein synthesis: Ribosomal
	33 (24)	109-					
	35 (31)	138-					
	62 (26)	159-					
	59 (26)	159-					
ABC transporter, ATP-	40 (36)	207-	Q4W545	CT11	55270.6	cytoplasmic	
Acetyl-coenzyme A	51 (26)	560 -	Q8KBY0	CT16	73575.5	cytoplasmic	Energy metabolism:
AcrB/AcrD/AcrF family	51 (25)	161-	Q8KAV4	CT20	118868.6	cytoplasmic	Transport and binding
	61 (27)	171-					
	48 (26)	303-					
	70 (26)	318-					
	59 (25)	323-					
	48 (26)	664-					
	89 (25)	679-					
	30 (26)	744-					
	69 (26)	748-					
	39 (25)	760-					
33 (25)	813-						
Alcohol dehydrogenase,	56 (26)	349-	Q8KDA2	CT11	40657.2	cytoplasmic	Energy metabolism:
Aminotransferase, class V	34 (26)	5- 20	Q8KG98	CT00	41767.0	cytoplasmic	Central intermediary
ATP	65 (30)	37-	Q8KB10	CT19	32103.0	Cytoplasmic	Amino acid biosynthesis:
ATP synthase subunit	40 (26)	2- 15	Q8KAW8	CT20	56847.5	unknown	Energy metabolism: ATP-
	51 (26)	2- 16					
	55 (26)	93-					
	55 (26)	109-					
	47 (27)	134-					
55 (27)	306-						
Band 7 family protein	64 (25)	199-	Q8KE93		27873.5	unknown	Cell envelope: Other
Branched-chain amino	60 (25)	202-	Q8KC21	CT16	33945.5	unknown	Amino acid biosynthesis:
Cell division protein FtsH	31 (30)	78-	Q8KG41	CT01	78396.7	cytoplasmic	Cellular processes: Cell
	42 (26)	362-					
	58 (26)	383-					
	46 (25)	394-					
	131 (26)	409-					
	28 (25)	432-					
39 (26)	481-						
Chlorosome envelope	89 (26)	34-	Q46383	CT20	7608.7	chlorosome	Energy metabolism:
	51 (25)	48-					
	29 (25)	59-					
Chlorosome envelope	38 (25)	12-	Q46387	CT20	11102.5	chlorosome	Energy metabolism:
Chlorosome protein E	78 (30)	6- 17	Q46386	CT20	8954.0	chlorosome	Energy metabolism:
	61 (26)	44-					
Chlorosome envelope	33 (23)	35-	O68986	CT10	7742.7	chlorosome	Energy metabolism:
	69 (32)	57-					
	46 (26)	68-					
	56 (26)	68-					
	81 (26)	2- 15					

Chlorosome envelope	75 (26)	99-	Q68991	CT14	21779.5	chlorosome	Energy metabolism:
	44 (26)	187-					
Cytochrome bc complex	68 (26)	83-	Q9F721	CT03	47495.2	cytoplasmic	Energy metabolism: Electron
Cytochrome c, putative	54 (26)	124-	Q8KBS9	CT17	17037.5	unknown	Energy metabolism: Electron
Cytochrome c-555,	61 (24)	37-	Q8KG95		14512.7	unknown	Energy metabolism: Electron
	64 (24)	42-					
	59 (24)	109-					
	87 (26)	121-					
	47 (26)	137-					
DNA methylase, putative	75 (26)	655-	Q8KBQ5		127470.8	unknown	DNA metabolism:
DNA-directed RNA	42 (26)	25-	Q8KAJ8		36936.0	cytoplasmic	Transcription: DNA-
	65 (26)	33-					
	66 (25)	45-					
	58 (26)	260-					
	108 (26)	275-					
Ferrous iron transport	40 (26)	238-	Q8KGB2	CT00	78352.2	cytoplasmic	Transport and binding
	53 (30)	578-					
Glutamyl-tRNA reductase	58 (26)	414-	Q93ST4	CT14	48133.0	cytoplasm	Biosynthesis of cofactors,
Hemagglutinin-related	84 (27)	59-	Q8KGA0	CT00	22145.8	unknown	Unknown function: General
	62 (27)	59-					
	101 (26)	71-					
	42 (26)	90-					
	36 (26)	179-					
	65 (26)	179-					
Heterodisulfide	49 (26)	388-	Q8KE26	CT08	45094.9	unknown	Central intermediary
Histidinol dehydrogenase	80 (26)	41-	Q8KEY6	CT05	46263.5	cytoplasmic	Amino acid biosynthesis:
HIT family protein	30 (27)	155-	Q8KEC9	CT07	20189.2	cytoplasm	Unknown function: General
Putative uncharacterized	60 (26)	147-	Q8KGF1	CT00	44022.2	unknown	
	55 (26)	167-					
	46 (26)	178-					
	61 (26)	202-					
	41 (25)	259-					
	43 (26)	345-					
Putative uncharacterized	41 (26)	179-	Q8KG59	CT01	29644.2	unknown	
	32 (27)	196-					
	60 (25)	196-					
	90 (26)	206-					
	43 (26)	239-					
Putative uncharacterized	44 (26)	7- 19	Q8KFL6		7409.4	unknown	
	35 (24)	43-					
Putative uncharacterized	52 (27)	130-	Q8KF09	CT05	107606.9	outer membrane	
	71 (26)	161-					
	56 (27)	284-					
	87 (26)	354-					
	61 (26)	439-					
	71 (26)	524-					
Putative uncharacterized	100 (26)	43-	Q8KDR1	CT09	24884.7	outer membrane	
	102 (26)	177-					
Putative uncharacterized	29 (25)	193-	Q8KDL2	CT10	23487.7	cytoplasmic	
Putative uncharacterized	39 (31)	110-	Q8KD97	CT11	25393.9	unknown	
	56 (26)	180-					
Putative uncharacterized	32 (26)	79-	Q8KD76	CT11	31827.8	unknown	
	91 (26)	178-					
	30 (26)	221-					
Putative uncharacterized	55 (25)	161-	Q8KD43	CT12	33731.9	cytoplasmic	
	44 (25)	200-					
	41 (25)	247-					

Putative uncharacterized	38 (26)	270 -	Q8KC55	CT15	47489.5	cytoplasm	
	36 (30)	330-					
Putative uncharacterized	60 (26)	307-	Q8KBW3	CT16	83642.1	outer membrane	
Putative uncharacterized	58 (26)	137-	Q8KB32	CT19	18593.9	unknown	
Putative uncharacterized	32 (26)	73-	Q8KAY5		63227.1	unknown	
	28 (27)	96-					
	73 (26)	151-					
	32 (27)	189-					
	39 (25)	237-					
	55 (27)	306-					
	44 (26)	440-					
	89 (26)	455-					
Putative uncharacterized	98 (26)	75-	Q8KAV3		18127.8	unknown	
	46 (26)	89-					
Putative uncharacterized	83 (26)	52-	Q8KA97	CT22	21451.5	unknown	
Putative uncharacterized	36 (25)	24-	Q8KA88		12560.2	cytoplasmic	
	40 (26)	97-					
LemA family protein	82 (25)	98-	Q4W537		21817.0	unknown	Hypothetical proteins:
	46 (31)	130-					
	36 (26)	147-					
	44 (26)	149-					
Lipoprotein, putative	39 (26)	69-	Q8KEH8	CT07	49189.9	cytoplasmic	Hypothetical proteins:
	33 (26)	191-					
BchH2 or Magnesium-	37 (26)	69-	Q93SV6		145575.2	cytoplasmic	Biosynthesis of cofactors,
Magnesium-transporting	84 (26)	205-	Q8KEQ5		97911.5	cytoplasmic	Transport and binding
Malate dehydrogenase	64 (26)	27-	P80039	CT15	97911.5	cytoplasmic	Energy metabolism: TCA
	33 (25)	42-					
	64 (31)	139-					
Membrane fusion efflux	67 (27)	209-	Q8KCR5	CT13	46300.7	cytoplasmic	Transport and binding
	37 (27)	223-					
Membrane protein,	84 (26)	309-	Q8KEE9		50583.3	cytoplasmic	Hypothetical proteins:
Molybdopterin	58 (27)	65-	Q8KCT1	CT13	45154.0	unknown	Biosynthesis of cofactors,
Multidrug resistance	37 (25)	99-	Q8KCW9	CT12	41784.1	cytoplasmic	Transport and binding
	46 (26)	140-					
	37 (26)	282-					
	35 (25)	369-					
Multidrug resistance	76 (26)	249-	Q8KCX1		51903.5	outer membrane	Transport and binding
Outer membrane	46 (26)	142-	Q8KBG4	CT18	19591.5	unknown	Cell envelope: Other
Inner membrane protein	43 (26)	316-	Q8KGG2		65489.7	cytoplasmic	Unknown function: General
Pentapeptide repeat	83 (26)	53-	Q8KDK2	CT10	60492.3	unknown	Unknown function: General
	44 (26)	479-					
Peptide ABC transporter,	61 (27)	137-	Q8KCJ8		64707.0	cytoplasmic	Transport and binding
Phosphate ABC	72 (25)	116-	Q8KDZ6	CT08	49369.5	cytoplasmic	Transport and binding
Phosphoglucomutase/pho	89 (26)	461-	Q8KFR2		51378.7	cytoplasm	Energy metabolism: Sugars
	85 (26)	216-	Q8KDT4	CT09	26540.5	cytoplasm	Purines, pyrimidines,
PilB-related protein	58 (26)	61-	Q8KG65	CT01	20583.3	unknown	Unknown function: General
Polyribonucleotide	51 (31)	48-	Q8KBY3	CT16	79937.8	cytoplasm	Transcription: Degradation of
	45 (26)	430-					
	43 (26)	572-					
	38 (26)	631-					
	43 (26)	648-					
Light-independent	43 (25)	301-	Q9F715	CT21	58968.5	cytoplasm	Biosynthesis of cofactors,
	76 (27)	517-					
	26 (25)	10-					
	66 (26)	28-					
	37 (25)	138-					

Proton transporting	27 (25)	385-	Q8KBU9		96714.7	cytoplasmic	Transport and binding
	81 (26)	414-					
	88 (26)	432-					
	63 (27)	433-					
	74 (26)	491-					
	58 (26)	502-					
	75 (26)	553-					
	41 (26)	553-					
50S ribosomal protein	47 (30)	42-	Q8KAJ0	CT21	20083.3	unknown	Protein synthesis: Ribosomal
	34 (27)	63-					
	85 (27)	88-					
	42 (25)	151-					
50S ribosomal protein L21	37 (26)	158-	Q8KCB8	CT15	10840.7	unknown	Protein synthesis: Ribosomal
	69 (26)	27-					
	62 (26)	37-					
50S ribosomal protein L5	31 (26)	85-	Q8KAI4	CT21	21037.6	cytoplasm	Protein synthesis: Ribosomal
	32 (26)	50-					
	42 (27)	104-					
30S ribosomal protein	106 (26)	135-	Q8KAJ6	CT21	13666.7	unknown	Protein synthesis: Ribosomal
	86 (25)	56-					
30S ribosomal protein Rubrerythrin	62 (25)	97-	Q8KAI5	CT21	10226.0	unknown	Protein synthesis: Ribosomal
	70 (26)	29-					
Signal recognition	71 (26)	10 -	Q8KCT4	CT13	18165.5	cytoplasm	Energy metabolism: Electron
Succinate/fumarate	39 (27)	306-	Q8KD87		49603.0	cytoplasm	Protein fate:
Succinyl-CoA ligase	42 (26)	276-	Q8KAW0	CT20	39528.6	cytoplasmic	Energy metabolism:
	58 (26)	69-					
	77 (26)	220-					
Sulfate	61 (26)	364-	Q8KFE7	CT03	42217.4	cytoplasm	Energy metabolism: TCA
76 (31)	12-						
Superoxide dismutase	72 (26)	102-	Q8KE30	CT08	45036.7	cytoplasm	Central intermediary
Threonine synthase	71 (26)	299-	Q8KD44	CT12	22066.5	unknown	Cellular processes:
	44 (26)	432-					
Threonyl-tRNA	71 (26)	299-	Q8KAW6	CT20	48723.2	unknown	Amino acid biosynthesis:
Tia invasion determinant-	52 (26)	55-					
Toluene transport protein,	37 (25)	44-	Q8KAN0	CT21	74884.7	cytoplasm	Protein synthesis: tRNA
	27 (24)	36-					
	55 (27)	65-					
Probable transaldolase	57 (26)	232-	Q8KF15	CT05	23521.6	outer membrane	Unknown function: General
	43 (26)	129-					
Transcriptional regulator,	67 (30)	75-	Q8KAD3	CT22	49259.6	cytoplasmic	Transport and binding
Transketolase	84 (27)	23-					
Type II secretion system	68 (26)	591-	Q8KGF8	CT00	23766.5	cytoplasm	Energy metabolism: Pentose
	73 (24)	30-					
	62 (26)	66-					
	32 (26)	79-					
	57 (26)	127-					
	46 (25)	127-					
	89 (25)	284-					
	53 (27)	308-					
48 (24)	322-						
UDP-glucose/GDP-	37 (26)	415-	Q8KBB9	CT18	74372.5	cytoplasmic	Energy metabolism: Pentose
	85 (26)	120-					
	62 (26)	354-					
UDP-N-	61 (26)	421-	Q8KDZ1	CT09	27147.3	cytoplasm	Regulatory functions: DNA
	70 (26)	61-					
VpsC protein	89 (26)	219-	Q8KCE3	CT14	51627.4	unknown	Cell envelope: Biosynthesis
	60 (26)	255-					
			Q8KFK1	CT03	31895.7	unknown	Protein fate: Protein and

common in all fractions							
DNA polymerase III, beta	55 (27)	1- 9	Q8KGG7	CT00	42158.24	cytoplasm	DNA metabolism: DNA
hypothetical protein	69 (26)	1- 17	Q8KGE1	CT00	7554.60	unknown	
Phosphohistidine	68 (26)	1- 8	Q8KGG77	CT00	18596.24	cytoplasm	Regulatory functions: Protein
Phosphoribosylformylglyci	81 (27)	221-	Q8KGG74	CT00	35690.24	cytoplasm	Purines, pyrimidines,
RNA-binding protein	86 (26)	1- 18	Q8KGG36	CT01	10429.58	unknown	DNA metabolism: Other
P-II family protein	56 (26)	1- 12	Q8KGG34	CT01	12587.41	cytoplasm	Regulatory functions: Protein
Translation elongation	47 (27)	1- 10	Q8KGG26	CT01	76538.26	cytoplasm	Protein synthesis:
Elongation factor P	78 (25)	1- 9	Q8KGG11	CT01	21090.88	cytoplasm	Protein synthesis:
DNA-binding protein HU-	100 (26)	1- 23	Q8KG10	CT01	9594.95	unknown	DNA metabolism:
	36 (26)						
	33 (31)	2- 23					
Putative uncharacterized	63 (26)	1- 10	Q8KFY3	CT01	29972.17	cytoplasm	Hypothetical proteins:
Tryptophan synthase,	66 (26)	2- 17	Q8KFX8	CT01	50154.66	unknown	Amino acid biosynthesis:
Putative uncharacterized	39 (26)	2- 12	Q8KFS4	CT02	13673.14	cytoplasm	
Glutathione S-transferase,	73 (25)	1- 12	Q8KFS3	CT02	18404.71	cytoplasm	Cellular processes: Toxin
Isoprenyl synthetase	66 (26)	2- 15	Q8KFR6	CT02	36263.42	cytoplasm	Biosynthesis of cofactors,
Isopentenyl-diphosphate	61 (26)	1- 12	Q8KFR5	CT02	38265.32	cytoplasm	
Succinyl-CoA ligase	54 (26)	2- 10	Q8KFAQ3	CT02	31075.92	unknown	Energy metabolism: TCA
DNA-binding protein HU-	49 (26)	2- 17	Q8KFM7	CT02	11325.02	cytoplasm	DNA metabolism:
Cytochrome b6-f complex	40 (26)	2- 12	Q9F722	CT03	18863.90	cytoplasmic	Energy metabolism: Electron
Glucose-1-phosphate	105 (26)	1- 13	Q8KFM1	CT03	32521.00	cytoplasm	Cell envelope: Biosynthesis
2,3-bisphosphoglycerate-	68 (26)	1- 8	Q8KFC8	CT03	28434.34	unknown	Energy metabolism:
Imidazole glycerol	84 (26)	1- 13	Q8KF56	CT04	22259.31	unknown	Amino acid biosynthesis:
Putative uncharacterized	78 (27)	1- 12	Q8KF47	CT04	7113.19	unknown	
Putative uncharacterized	68 (26)	1- 9	Q8KF45	CT04	14754.87	cytoplasm	
P-II family protein	73 (27)	1- 9	Q8KF33	CT04	13594.30	cytoplasm	Regulatory functions: Protein
10 kDa chaperonin	68 (27)	1- 9	Q8KF03	CT05	10165.77	cytoplasm	Protein fate: Protein folding
or Protein Cpn10	31 (26)						
60 kDa chaperonin or	72 (26)	2- 13	Q8KF02	CT05	58079.56	cytoplasm	Protein fate: Protein folding
Probable pyruvoyl-	55 (26)	2- 16	Q8KEX0	CT05	19975.59	unknown	Hypothetical proteins:
2-isopropylmalate	87 (26)	1- 11	Q8KES0	CT06	58546.48	unknown	Amino acid biosynthesis:
	46 (26)	2- 11					
3-isopropylmalate	68 (26)	2- 12	Q8KER8	CT06	46671.06	unknown	Amino acid biosynthesis:
PTS system, IIA	78 (26)	1- 13	Q8KEQ6	CT06	29515.59	unknown	Transport and binding
P840 reaction center 17	51 (26)	1- 7	Q8KEP5	CT06	16607.18	cytoplasm	Energy metabolism:
Heat shock protein,	47 (26)	1- 11	Q8KEP2	CT06	15054.07	cytoplasm	Protein fate: Protein folding
Cysteine synthase	80 (27)	2- 13	Q8KEI9	CT07	32631.47	cytoplasm	Amino acid biosynthesis:
Epoxide hydrolase,	40 (26)	1- 13	Q8KEH2		36423.78	unknown	Cellular processes:
Methylenetetrahydrofolate	91 (26)	1- 14	Q8KEH0		33243.69	unknown	Purines, pyrimidines,
Glucose-1-phosphate	65 (26)	1- 13	Q8KEG2	CT07	35573.16	cytoplasm	Cell envelope: Biosynthesis
Thioredoxin-1	57 (26)	2- 10	Q8KEA4	CT07	11247.23	cytoplasm	Energy metabolism: Electron
Putative uncharacterized	44 (26)	1- 18	Q8KE70	CT08	12156.90	unknown	Hypothetical proteins:
Putative uncharacterized	90 (26)	1- 12	Q8KE66	CT08	19047.69	cytoplasm	Hypothetical proteins:
Lipoprotein, putative	55 (26)	1- 17	Q8KE20	CT08	30187.4	cytoplasm	Hypothetical proteins:
Phosphate ABC	69 (26)	261-	Q8KDZ9	CT08	28378.9	unknown	Transport and binding
	62 (26)						
Putative uncharacterized	76 (26)	2- 12	Q8KDT5	CT09	15258.0	cytoplasm	
Phospho-2-dehydro-3-	65 (26)	1- 8	Q8KDR4	CT09	39319.9	unknown	Amino acid biosynthesis:
Sulfur oxidation protein	67 (26)	115-	Q8KDM7	CT10	32133.7	unknown	Central intermediary
Dihydroorotase	59 (26)	2- 9	Q8KDK5	CT10	47039.0	unknown	Purines, pyrimidines,
Putative uncharacterized	72 (30)	1- 17	Q8KDH4	CT10	24933.6	cytoplasm	Hypothetical proteins:
Putative uncharacterized	86 (26)	1- 9	Q8KDG8	CT10	34575.4	cytoplasm	Hypothetical proteins:
Ornithine	47 (26)	2- 13	Q8KDE2	CT11	38105.7	cytoplasm	Amino acid biosynthesis:
Putative uncharacterized	81 (26)	53-	Q8KD96	CT11	8237.5	cytoplasm	
50S ribosomal protein L19	78 (27)	1- 14	Q8KD91	CT11	13347.5	unknown	Protein synthesis: Ribosomal
Putative uncharacterized	43 (26)	2- 15	Q8KD84	CT11	25006.5	Cytoplasm	Hypothetical proteins:

Putative uncharacterized	54 (26)	1- 10	Q8KD70	CT11	28671.8	unknown	
Aspartate	60 (27)	1- 9	Q8KD15	CT12	45530.2	unknown	Amino acid biosynthesis:
Peptidyl-tRNA hydrolase	69 (26)	1- 13	Q8KD05	CT12	20096.3	Cytoplasm	Protein synthesis: Other
50S ribosomal protein L25	76 (25)	72-	Q8KCQ1	CT13	21531.5	cytoplasm	Cellular processes:
	79 (26)	186-					
Methylenetetrahydrofolate	87 (26)	1- 10	Q8KCP5	CT13	32772.6	Cytoplasm	Amino acid biosynthesis:
Glutamine synthetase	29 (26)	2- 24	Q8KCK4		78551.8	Cytoplasm	Amino acid biosynthesis:
Porphobilinogen	59 (26)	1- 10	Q8KCJ4	CT14	34393.8	Cytoplasm	Biosynthesis of cofactors,
Uroporphyrinogen-III	70 (25)	1- 8	Q8KCJ3		26616.5	Cytoplasm	Biosynthesis of cofactors,
GTP-binding protein lepA	77 (27)	2- 12	Q8KCHO	CT14	67767.9	Cytoplasm	Unknown function: General
Probable GTPase engC	82 (31)	1- 11	Q8KC52	CT15	34460.0	Cytoplasm	Hypothetical proteins:
Glycogen phosphorylase	63 (26)	2- 13	Q8KC30	CT15	80667.8	unknown	Energy metabolism:
50S ribosomal protein L28	63 (26)	2- 11	Q93SU8	CT16	8403.9	unknown	Protein synthesis: Ribosomal
Glycine cleavage system	46 (26)	1- 8	Q8KC04	CT16	13543.2	unknown	Energy metabolism: Amino
Phosphoribosylamine--	79 (26)	1- 11	Q8KBV8		45064.4	Cytoplasm	Purines, pyrimidines,
Putative uncharacterized	59 (26)	1- 19	Q8KBP1		9959.5	unknown	Hypothetical proteins:
Molybdenum-pterin-	41 (25)	1- 6	Q8KBM1	CT17	14275.5	Cytoplasm	Transport and binding
Putative uncharacterized	55 (26)	2- 14	Q8KBL7	CT17	15176.5	unknown	
RuBisCO-like protein or	79 (31)	1- 13	Q8KBL4		48002.9	Cytoplasm	Energy metabolism:
Elongation factor Ts or	92 (26)	2- 13	Q8KBK7	CT17	31544.1	Cytoplasm	Protein synthesis:
30S ribosomal protein S2	43 (31)	2- 12	Q8KBK6	CT17	31544.1	Cytoplasm	Protein synthesis:
30S ribosomal protein S9	80 (26)	1- 10	Q8KBK5	CT17	14541.0	Cytoplasm	Protein synthesis: Ribosomal
ATP-binding protein,	32 (26)	2- 23	Q8KBK2	CT17	39209.0	unknown	Unknown function: General
Putative uncharacterized	56 (26)	365-	Q8KBI3		42810.9	outer membrane	
Putative uncharacterized	80 (26)	1- 12	Q8KBH3	CT18	25394.8	Cytoplasm	Hypothetical proteins:
Translation initiation	77 (26)	1- 13	Q8KBH1		38690.1	unknown	Protein synthesis:
3-deoxy-manno-	68 (26)	1- 10	Q8KBF7	CT18	27729.7	Cytoplasm	Cell envelope: Biosynthesis
Aspartyl/glutamyl-	91 (26)	2- 16	Q8KBF4	CT18	10669.0	unknown	Protein synthesis: tRNA
Dihydrodipicolinate	76 (26)	1- 10	Q8KBD8	CT18	27027.5	Cytoplasm	Amino acid biosynthesis:
5-methyltetrahydrofolate-	71 (26)	1- 12	Q8KBD1		134894.2	Cytoplasm	Amino acid biosynthesis:
Holo-[acyl-carrier-	90 (31)	1- 12	Q8KB56	CT19	12610.5	Cytoplasm	Fatty acid and phospholipid
Chlorosome protein A or	82 (26)	1- 16	P0A314 or Q46368	CT19	8289.4	chlorosome	Energy metabolism:
	68 (27)						
	96 (31)						
Chlorosome protein C or	50 (26)	2- 9	Q46367	CT19	14325.0	chlorosome	Energy metabolism:
CRISPR-associated	45 (26)	1- 10	Q8KB23	CT19	37908.7	Cytoplasm	Hypothetical proteins:
D,D-heptose 1,7-	60 (26)	1- 11	Q8KAY7	CT20	22675.3	Cytoplasm	Unknown function: General
ATP synthase gamma	59 (26)	2- 8	Q8KAW9	CT20	32053.1	cytoplasm	Energy metabolism: ATP-
Uroporphyrinogen	60 (26)	1- 9	Q8KAW2	CT20	39104.2	Cytoplasm	Biosynthesis of cofactors,
CBS domain protein	72 (26)	1- 8	Q8KAV7	CT20	16716.4	cytoplasm	Unknown function: General
Hydroxyacylglutathione	53 (27)	2- 9	Q8KAV6	CT20	23289.1	unknown	Energy metabolism: Other id
OrfM or Response	40 (26)	1- 13	Q46385	CT20	13052.4	cytoplasm	Regulatory functions: Other
Putative uncharacterized	87 (26)	1- 10	Q8KAQ5		12531.5	cytoplasm	Hypothetical proteins:
50S ribosomal protein L32	54 (26)	2- 12	Q8KAP4	CT21	6963.1	unknown	Protein synthesis: Ribosomal
3-oxoacyl-(Acyl-carrier-	122 (26)	1- 14	Q8KAP0	CT21	25865.7	cytoplasm	Fatty acid and phospholipid
Probable glycine	69 (26)	1- 11	Q8KAN3	CT21	53809.8	unknown	Energy metabolism: Amino
50S ribosomal protein	85 (26)	2- 13	Q8KAM7	CT21	13126.4	unknown	Protein synthesis: Ribosomal
Phenylalanyl-tRNA	47 (26)	1- 18	Q8KAM6	CT21	38499.2	Cytoplasm	Protein synthesis: tRNA
50S ribosomal protein L9	48 (26)	1- 6	Q8KAM4	CT21	16499.0	cytoplasm	Protein synthesis: Ribosomal
50S ribosomal protein L36	52 (26)	1- 10	Q8KAJ4	CT21	4585.5	unknown	Protein synthesis: Ribosomal
	Methionine	66 (29)	1- 8	Q8KAJ2	CT21	28190.0	cytoplasm
50S ribosomal protein L30	68 (26)	2- 12	Q8KAI9	CT21	6838.1	cytoplasm	Protein synthesis: Ribosomal
30S ribosomal protein S8	107 (25)	2- 13	P59031	CT21	14973.3	unknown	Protein synthesis: Ribosomal
30S ribosomal protein S3	70 (26)	2- 11	Q8KAH8	CT21	28453.0	cytoplasm	Protein synthesis: Ribosomal
50S ribosomal protein L22	54 (26)	1- 8	Q8KAH7	CT21	13408.7	unknown	Protein synthesis: Ribosomal
50S ribosomal protein L3	75 (26)	2- 13	Q8KAH2	CT21	22182.6	unknown	Protein synthesis: Ribosomal
Elongation factor Tu	41 (26)	2- 8	Q8KAH0	CT21	42899.2	Cytoplasm	Protein synthesis:

30S ribosomal protein	65 (26)	2- 9	Q8KAG8	CT21	14435.9	unknown	Protein synthesis: Ribosomal
Phosphocarrier protein	87 (27)	1- 9	Q8KAF1	CT22	9577.1	Cytoplasm	Transport and binding
Porphyrin biosynthesis	51 (26)	1- 14	Q8KAC4	CT22	17833.2	cytoplasm	Biosynthesis of cofactors,
Probable chaperone	45 (36)	181-	Q8KGG79	CT00	49654.71	cytoplasm	Protein fate: Degradation of
Outer membrane protein,	88 (32)	15-	Q8KGG61	CT01	47364.5	outer membrane	Protein fate: Protein and
Bacterial surface antigen	75 (31)	31-	Q8KFG6	CT02	93746.5	outer membrane	Cell envelope: Other
Putative uncharacterized	37 (33)	37-	Q8KET2		20803.1	periplasm	Hypothetical proteins:
	90 (32)	27-					
	58 (32)	33-					
Protein tolB [Precursor]	69 (32)	28-	Q8KEQ0		47275.42	Periplasm	Transport and binding
Putative uncharacterized	50 (31)	17-	Q8KDX0	CT09	10484.9	unknown	
Putative uncharacterized	38 (36)	23-	Q8KDM6		11684.4	unknown	
	59 (31)	24-					
	71 (36)	25-					
	98 (32)	28-					
	75 (32)	33-					
	43 (36)						
33 (32)	35-						
Sulfur oxidation protein	63 (32)	60-	Q8KDM5	CT10	68234.8	unknown	Central intermediary
	52 (33)						
Carboxyl-terminal	41 (31)	3- 24	Q8KDJ6	CT10	75483.3	unknown	Protein fate: Degradation of
	48 (31)	8- 24					
Thiol:disulfide interchange	56 (31)	12-	Q8KDH5		63811.7	cytoplasmic	Protein fate: Protein folding
	57 (37)	556-					
	60 (37)	557-					
Putative uncharacterized	64 (32)	24-	Q8KDH0		28666.0	unknown	
CRISPR-associated	82 (38)	476-	Q8KDC1		67092.9	unknown	
Putative uncharacterized	84 (32)	5- 16	Q8KD60		21796.1	cytoplasm	
Putative uncharacterized	64 (32)	38-	Q8KCC6		28291.7	unknown	
	64 (38)	211-					
Putative uncharacterized	54 (37)	199-	Q8KB89	CT19	43375.6	unknown	
TPR domain protein	87 (32)	8- 19	Q8KAT0		32241.4	unknown	Unknown function: General
Ribosomal protein S1	39 (30)	2- 17	Q8KFN4	CT02	61516.5	unknown	Protein synthesis: Ribosomal
unique cytoplasm							
Deoxyhyposine synthase-	96 (26)	129-	Q8KGF7	CT00	39045.5	cytoplasm	Regulatory functions
Hydrolase, haloacid	38 (27)	1- 25	Q8KGF4	CT00	25419.2	unknown	Central intermediary
4-hydroxybutyrate	47 (25)	5- 20	Q8KGD9	CT00	46983.5	cytoplasm	Energy metabolism:
Cell division protein FtsZ	29 (27)	222-	Q8KGD8	CT00	46081.2	cytoplasm	Cellular processes: Cell
DNA topoisomerase I	56 (25)	780-	Q8KGB3	CT00	89970.7	cytoplasm	DNA metabolism: DNA
Putative uncharacterized	29 (26)	1- 8	Q8KGG68	CT01	44549.1	cytoplasm	Hypothetical proteins:
Glucosamine--fructose-	62 (25)	2- 10	Q8KGG38	CT01	67576.9	cytoplasm	Central intermediary
Putative uncharacterized	50 (31)	35-	Q8KGG37	CT01	33752.7	unknown	Hypothetical proteins:
Alpha oxoglutarate	58 (26)	76-	Q8KGG08		37798.7	unknown	Energy metabolism: TCA
Alpha oxoglutarate	66 (26)	157-	Q8KGG07	CT01	67379.4	cytoplasm	Energy metabolism: TCA
Proline iminopeptidase,	43 (25)	2- 8	Q8KGFZ3	CT01	34254.3	unknown	Protein fate: Degradation of
Glycerol kinase	71 (25)	2- 20	Q8KGFY5	CT01	52226.5	unknown	Energy metabolism: Other id
Phosphoenolpyruvate-	34 (26)	81-	Q8KFW6	CT02	67005.9	cytoplasm	Transport and binding
Transcription termination	64 (26)	233-	Q8KFG8		48469.7	unknown	Transcription: Transcription
Putative uncharacterized	33 (27)	174-	Q8KFP7		30669.6	unknown	
Isoleucyl-tRNA	36 (26)	2- 22	Q8KFL5	CT03	124097.1	cytoplasm	Protein synthesis: tRNA
DNA mismatch binding	38 (27)	1- 14	Q8KFJ2	CT03	89896.9	cytoplasm	DNA metabolism: DNA
Putative uncharacterized	57 (26)	261-	Q8KFG9	CT03	30065.3	unknown	Hypothetical proteins:
Type II secretion system	41 (26)	2- 10	Q8KFG93	CT04	52852.7	cytoplasm	Protein fate: Protein and
Oxidoreductase, FAD-	77 (25)	251-	Q8KEW7		54843.4	cytoplasm	Central intermediary
O-acetylhomoserine	43 (25)	2- 8	Q8KES8	CT06	46676.7	cytoplasm	Amino acid biosynthesis:
Seryl-tRNA synthetase	64 (26)	1- 8	Q8KES6	CT06	47990.5	cytoplasm	Protein synthesis: tRNA
	70 (26)	268-					

Dihydroxy-acid	62 (26)	17-	Q8KER4		58915.6	cytoplasm	Amino acid biosynthesis:
N-(5'-	53 (26)	1- 11	Q8KEH1		23801.1	cytoplasm	Amino acid biosynthesis:
Adenosylhomocysteinase	54 (30)	237-	Q8KEG8		51949.4	cytoplasm	Energy metabolism: Amino
Putative uncharacterized	52 (31)	164-	Q8KEG4	CT07	62851.6	unknown	
ABC transporter, ATP-	28 (27)	148-	Q8KED6	CT07	30940.2	cytoplasm	Transport and binding
Outer membrane efflux	31 (25)	209-	Q8KED1	CT07	52717.1	outer membrane	Transport and binding
NADH dehydrogenase I,	40 (25)	309-	Q8KEC0	CT07	41764.3	cytoplasmic	Energy metabolism: Electron
NADH dehydrogenase I,	66 (26)	33-	Q8KEB8	CT07	23568.9	unknown	Energy metabolism: Electron
Transketolase, C-	69 (27)	2- 16	Q8KE86	CT08	35681.1	cytoplasm	Energy metabolism: Pentose
Putative uncharacterized	31 (26)	1- 7	Q8KE74	CT08	24742.9	cytoplasmic	
Putative uncharacterized	72 (26)	1- 9	Q8KE59	CT08	58510.2	cytoplasm	Hypothetical proteins:
Adenylylsulfate reductase,	89 (26)	97-	Q8KE27	CT08	73418.4	unknown	Central intermediary
	92 (26)	178-					
	29 (26)	613-					
	56 (26)	651-					
Cobyrinic acid a,c-	36 (26)	1- 9	Q8KDW5	CT09	50546.6	unknown	Biosynthesis of cofactors,
	62 (31)						
Transketolase, N-terminal	68 (26)	100-	Q8KDT1		33431.1	cytoplasm	Energy metabolism: Pentose
Aspartate	39 (25)	2- 8	Q8KDS8	CT09	44007.3	cytoplasm	Amino acid biosynthesis:
Putative uncharacterized	81 (26)	2- 17	Q8KDS6	CT09	37599.1	unknown	Hypothetical Proteins:
Methionyl-tRNA	64 (25)	216-	P59077	CT09	79553.4	cytoplasmic	Protein synthesis: tRNA
tRNA delta(2)-	28 (26)	1- 23	Q8KDS2	CT09	37067.1	cytoplasm	Protein synthesis: tRNA and
Putative uncharacterized	52 (26)	2- 18	Q8KDI7		38189.0	unknown	Hypothetical proteins:
	41 (26)						
ABC transporter, ATP-	46 (30)	2- 8	Q8KDF8	CT10	24297.9	cytoplasmic	Transport and binding
CRISPR-associated	55 (25)	1- 7	Q8KDC2	CT11	32883.1	unknown	Hypothetical proteins:
	48 (25)	2- 7					
	27 (26)	379-					
CRISPR-associated	27 (26)	379-	Q8KDB9	CT11	74497.1	cytoplasm	Transcription: Other
Carbohydrate kinase,	73 (26)	1- 11	Q8KD86	CT11	35123.4	cytoplasm	Energy metabolism: Sugars
Probable cytosol	59 (26)	70 -	Q8KD74		52994.8	Cytoplasm	Protein fate: Degradation of
DHH family protein	32 (26)	1- 8	Q8KD73	CT11	39711.8	unknown	Unknown function: General
Polyprenyl synthetase	59 (26)	1- 20	Q8KD49	CT12	36157.4	cytoplasm	Biosynthesis of cofactors,
Dihydrolipoyl	48 (27)	307-	Q8KCW2	CT12	48007.4	Cytoplasm	Energy metabolism:
Prolyl oligopeptidase	67 (26)	616-	Q8KCV9	CT13	77943.2	periplasm	Protein fate: Degradation of
Oxidoreductase, short-	56 (26)	1- 21	Q8KCV2		25132.9	cytoplasm	Central intermediary
	31 (26)	2- 21					
DNA polymerase III,	57 (25)	2- 8	Q8KCT6	CT13	69350.0	Cytoplasm	DNA metabolism: DNA
Putative uncharacterized	37 (31)	2- 17	Q8KCP1	CT13	3718.5	unknown	
Putative uncharacterized	50 (26)	163-	Q8KCL7	CT13	49410.2	cytoplasm	Hypothetical proteins:
Delta-aminolevulinic acid	31 (26)	310-	Q8KCJ0	CT14	36432.6	cytoplasm	Biosynthesis of cofactors,
Anthranilate synthase	28 (25)	4- 25	Q8KCH3	CT14	57182.5	cytoplasmic	Amino acid biosynthesis:
Putative uncharacterized	30 (26)	82-	Q8KCF9	CT14	20886.7	unknown	Hypothetical proteins:
Gamma-glutamyl	56 (27)	169-	Q8KCE9	CT14	45960.6	Cytoplasm	Amino acid
Chaperone protein dnaJ	54 (26)	14-	Q8KCD8	CT14	43250.8	Cytoplasm	Protein fate: Protein folding
	74 (26)	346-					
Putative addiction module	62 (26)	2- 14	Q8KCA3	CT15	8525.7	cytoplasm	
Acetyl-CoA carboxylase,	44 (26)	97-	Q8KC71		30719.5	unknown	Fatty acid and phospholipid
Putative uncharacterized	82 (26)	1- 10	Q8KC68	CT15	63911.9	unknown	Hypothetical proteins:
Putative uncharacterized	27 (26)	62-	Q8KC65	CT15	12373.3	unknown	Hypothetical proteins:
Aldehyde dehydrogenase	66 (30)	1- 29	Q8KC53	CT15	49794.1	Cytoplasm	Energy metabolism:
3,4-dihydroxy-2-butanone	32 (26)	292-	Q8KC35	CT15	47821.9	cytoplasm	Biosynthesis of cofactors,
Carbamoyl-phosphate	63 (26)	91-	Q8KC34	CT15	68479.1	cytoplasm	Purines, pyrimidines,
Phosphofructokinase	27 (25)	280-	Q8KC23	CT16	44092.3	cytoplasm	Energy metabolism:
	30 (27)	382-					
	74 (26)	392-					
Pyruvate	60 (26)	217-	Q8KC02	CT16	130174.5	unknown	Energy metabolism: Electron
	52 (31)	257-					

HPr kinase/phosphorylase	66 (26)	1- 11	Q8KBZ7		39123.0	cytoplasm	Transport and binding
Putative uncharacterized	68 (26)	2- 13	Q8KBZ3		27023.8	cytoplasm	Hypothetical proteins:
	42 (26)	14-					
Leucyl-tRNA synthetase	78 (26)	223-	Q8KBY2	CT16	92478.2	Cytoplasm	Protein synthesis: tRNA
Putative uncharacterized	35 (31)	83-	Q8KBR6	CT17	12928.1	unknown	
Hydrogenase	47 (26)	2- 13	Q8KBJ5		36593.0	unknown	Protein fate: Protein folding
Hydrogenase accessory	86 (27)	2- 16	Q8KBI9		31073.9	cytoplasm	Protein fate: Protein folding
Formyltetrahydrofolate	47 (31)	1- 9	Q8KBG2		32984.0	Cytoplasm	Purines, pyrimidines,
Putative uncharacterized	64 (26)	1- 12	Q8KBE7		19331.3	unknown	
Type I restriction system	43 (25)	1- 7	Q8KBA8	CT18	57556.7	unknown	DNA metabolism:
TIM-barrel protein, nifR3	27 (26)	162-	Q8KB61	CT19	38878.5	unknown	Unknown function: General
IscS protein	40 (26)	222-	Q8KB03	CT19	43001.0	cytoplasm	Biosynthesis of cofactors,
	66 (25)	363-					
Succinate/fumarate	43 (26)	251-	Q8KAV9	CT20	62545.9	unknown	Energy metabolism:
	96 (25)						
	64 (26)	284-					
UDP-N-	75 (26)	2- 14	Q8KAR1	CT20	37930.8	unknown	Cell envelope: Biosynthesis
Glutamate-1-	49 (26)	7- 27	Q8KAQ7	CT20	46183.1	Cytoplasm	Biosynthesis of cofactors,
30S ribosomal protein S4	73 (26)	43-	P59129	CT21	23318.9	unknown	Protein synthesis: Ribosomal
Geranylgeranyl	27 (26)	1- 23	Q8KAB0	CT22	43089.7	unknown	Biosynthesis of cofactors,
50S ribosomal protein L34	72 (33)	43-	Q8KGG5		6157.1	unknown	Protein synthesis: Ribosomal
C-20 methyltransferase	66 (37)	300-	Q8KGE0	CT00	37934.7	Cytoplasm	Biosynthesis of cofactors,
UDP-N-acetylmuramate-	46 (37)	161-	Q8KGD5		51464.5	Cytoplasm	Cell envelope: Biosynthesis
DNA gyrase subunit A	46 (41)	807-	Q8KG27	CT01	92693.1	cytoplasm	DNA metabolism: DNA
50S ribosomal protein	45 (36)	102-	Q8KG16	CT01	12995.9	unknown	Protein synthesis: Ribosomal
Glutamyl-tRNA	33 (31)	83-	Q9F724		57810.1	Cytoplasm	Protein synthesis: tRNA
	45 (37)	479-	Q8KFL3	CT03	55515.2	Cytoplasm	Purines, pyrimidines,
Pyridoxine 5'-phosphate	66 (37)	157-	Q8KFJ5	CT03	26450.3	Cytoplasm	
Enoyl-(Acyl-carrier-	52 (37)	98-	Q8KFH6	CT03	32444.8	unknown	Fatty acid and phospholipid
Precorrin-3B C17-	36 (30)	12-	Q8KFE0	CT03	51490.0	cytoplasmic	Biosynthesis of cofactors,
Putative uncharacterized	36 (32)	71-	Q8KFD1	CT03	18090.8	unknown	
Glutamate synthase, large	64 (26)	96-	Q8KFC6	CT04	169774.9	unknown	Amino acid biosynthesis:
	61 (26)	109-					
Putative uncharacterized	37 (32)	80-	Q8KF96	CT04	17646.2	unknown	
Hydrogenase, putative	47 (37)	122-	Q8KF58	CT04	30268.6	unknown	Central intermediary
Polysulfide reductase,	58 (37)	110-	Q8KF37	CT04	22206.4	cytoplasm	Central intermediary
Polysulfide reductase,	41 (37)	218-	Q8KF36	CT04	101997.9	periplasm	Central intermediary
Putative uncharacterized	36 (32)	30 -	Q8KEV9	CT05	9341.8	unknown	
Putative uncharacterized	55 (38)	113-	Q8KEK7	CT06	36968.3	cytoplasm	
Putative uncharacterized	36 (32)	41-	Q8KEJ9	CT06	6882.7	cytoplasm	
Thiamin biosynthesis	51 (35)	4- 12	Q8KEJ3		26703.6	unknown	Biosynthesis of cofactors,
Phosphate ABC	39 (37)	397-	Q8KDZ7	CT08	48166.0	cytoplasmic	Transport and binding
Phosphate import ATP-	73 (31)	4- 18	Q8KDZ5	CT08	31931.7	cytoplasm	Transport and binding
Exodeoxyribonuclease V,	33 (32)	85-	Q8KDI1	CT10	134287.2	cytoplasm	DNA metabolism: DNA
Putative uncharacterized	51 (32)	74-	Q8KD85	CT11	23964.6	unknown	Hypothetical proteins:
Heterodisulfide reductase,	43 (42)	172-	Q8KD11	CT12	73175.9	Cytoplasm	Central intermediary
Putative uncharacterized	35 (32)	68-	Q8KCX7	CT12	7766.9	unknown	
Putative uncharacterized	50 (33)	40 -	Q8KCH6	CT14	8728.9	cytoplasm	
	37 (36)	60-					
DNA mismatch repair	41 (37)	504-	Q8KCC0	CT15	97797.1	Cytoplasm	DNA metabolism: DNA
Putative uncharacterized	32 (31)	29-	Q8KC94		4561.1	unknown	
Valyl-tRNA synthetase	54 (38)	677-	Q8KC74	CT15	102951.3	Cytoplasm	Protein synthesis: tRNA
Pyruvate,orthophosphate	48 (37)	890-	Q8KBV0	CT16	100664.4	cytoplasm	Energy metabolism: TCA
Putative uncharacterized	61 (31)	10-	Q8KBB3	CT18	55167.0	cytoplasm	Hypothetical proteins:
Putative uncharacterized	59 (32)	12-	Q8KB19	CT19	32740.3	cytoplasm	Hypothetical proteins:
Putative uncharacterized	46 (38)	240-	Q8KAY4	CT20	33259.6	cytoplasmic	Transport and binding
Esterase/lipase, putative	34 (32)	50-	Q8KAR8	CT20	25309.8	unknown	Fatty acid and phospholipid

Putative uncharacterized	68 (38)	178-	Q8KAK9	CT21	42377.1	cytoplasm	Hypothetical proteins:
50S ribosomal protein L17	41 (38)	145-	Q8KAJ9	CT21	17400.0	cytoplasm	Protein synthesis: Ribosomal
Glycyl-tRNA synthetase	70 (37)	320-	Q8KAB1	CT22	53718.0	Cytoplasm	Protein synthesis: tRNA
S-adenosylmethionine	59 (37)	280-	Q8KEG7		44177.7	Cytoplasm	Central intermediary
	61 (25)	290-					
	43 (37)	390-					
unique periplasm							
Arginyl-tRNA synthetase	35 (27)	259-	Q8KGF3	CT00	61452.9	cytoplasm	Protein synthesis: tRNA
NLP/P60 family protein	54 (26)	83-	Q8KG71		23837.5	unknown	Unknown function: General
Aminotransferase, class	27 (26)	13-	Q8KG58	CT01	44012.5	cytoplasm	Central intermediary
Undecaprenyl	32 (24)	244-	Q8KFQ5	CT02	30540.0	cytoplasm	Biosynthesis of cofactors,
Exodeoxyribonuclease V,	27 (27)	783-	Q8KDI2	CT10	120524.3	cytoplasm	DNA metabolism: DNA
Na ⁺ /H ⁺ antiporter,	28 (26)	675-	Q8KDB0	CT11	87355.9	cytoplasmic	Transport and binding
NLP/P60 family protein	29 (26)	125-	Q8KDA9	CT11	22855.8	unknown	Unknown function: General
Putative uncharacterized	29 (27)	194-	Q8KCL4	CT14	32948.6	unknown	
Putative uncharacterized	29 (26)	434-	Q8KCF8		59931.7	unknown	
TPR domain protein	26 (25)	532-	Q8KC01		62157.4	unknown	Unknown function: General
Transcriptional regulator,	33 (27)	66-	Q8KBX8	CT16	10759.5	Cytoplasm	Protein synthesis: Ribosomal
	37 (26)	752-	Q8KAX1	CT20	87914.8	cytoplasm	Amino acid biosynthesis:
Pentapeptide repeat	31 (26)	427-	Q8KAT8	CT20	46534.1	unknown	Unknown function: General
Glycosyl transferase	27 (26)	53-	Q8KAF3		31388.2	unknown	Cell envelope: Biosynthesis
Putative uncharacterized	90 (37)	89-	Q8KGC7	CT00	17415.0	unknown	
	49 (38)	97-					
Segregation and	39 (38)	108-	Q8KG32	CT01	28183.7	cytoplasm	Hypothetical proteins:
	44 (38)	175-					
Oxidoreductase, short-	39 (33)	81-	Q8KG01	CT01	26808.9	unknown	Central intermediary
Glycosyl transferase	39 (37)	268-	Q8KFX4	CT01	36370.7	cytoplasm	Cell envelope: Biosynthesis
UPF0090 protein	35 (32)	24-	Q8KFT3	CT02	18871.5	cytoplasm	Hypothetical proteins:
Putative uncharacterized	51 (32)	35-	Q8KFR4	CT02	45473.8	cytoplasmic	Hypothetical proteins:
Putative uncharacterized	55 (32)	36-	Q8KFP6	CT02	18989.9	unknown	
	40 (37)	126-					
Hydroxyneurosporene	39 (37)	331-	Q9F723	CT03	44954.6	unknown	Biosynthesis of cofactors,
ABC transporter, ATP-	39 (38)	562-	Q8KFF2		74074.1	cytoplasmic	Transport and binding
Iron(III) ABC transporter,	33 (32)	64-	Q8KFB6	CT04	37802.3	cytoplasmic	Transport and binding
Putative uncharacterized	36 (32)	47-	Q8KF65		9994.3	cytoplasm	
Segregation and	36 (33)	41-	Q8KF54	CT04	23924.1	cytoplasm	Hypothetical proteins:
Putative uncharacterized	31 (30)	59-	Q8KF25		8120.7	unknown	
	33 (30)	60-					
	52 (31)	63-					
	83 (31)	65-					
	51 (36)	68-					
49 (33)	69-						
Tryptophan synthase	38 (37)	136-	Q8KEZ7	CT05	29259.9	cytoplasm	Amino acid biosynthesis:
Ribonucleotide reductase	40 (37)	344-	Q8KEV4		124981.3	cytoplasmic	Purines, pyrimidines,
Nudix/MutT family	48 (38)	154-	Q8KEQ9		19027.8	cytoplasm	Unknown function: Enzymes
Putative uncharacterized	92 (37)	213-	Q8KEQ1		36116.7	cytoplasmic	
Putative uncharacterized	47 (37)	223-	Q8KEP6	CT06	29147.9	unknown	Hypothetical proteins:
Putative uncharacterized	47 (32)	50-	Q8KEH7	CT07	7909.9	unknown	
Putative uncharacterized	38 (32)	88-	Q8KEF8	CT07	11979.7	cytoplasm	
Putative uncharacterized	37 (31)	83-	Q8KE99	CT07	13475.6	unknown	
NfeD protein	32 (32)	77-	Q8KE94	CT07	45657.5	cytoplasmic	Unknown function: General
Putative uncharacterized	39 (37)	242-	Q8KE01	CT08	42409.6	outer membrane	
Putative uncharacterized	51 (37)	114-	Q8KDV8		18104.9	unknown	
ArsA ATPase family	39 (33)	74-	Q8KDR5	CT09	44541.5	cytoplasm	Transport and binding
PhnA protein	44 (33)	34-	Q8KDP2		12346.0	cytoplasm	Transport and binding
Sulfur oxidation protein	43 (32)	88-	Q8KDN0	CT10	15334.6	unknown	Central intermediary

Peptidase, M20/M25/M40	39 (31)	18-	Q8KDJ8		43957.2	unknown	Protein fate: Degradation of
Exonuclease SbcC	39 (37)	568-	Q8KDA7	CT11	136898.0	extracellular	DNA metabolism: DNA
Thiamine-phosphate	58 (32)	60-	Q8KD79	CT11	21474.5	cytoplasm	Biosynthesis of cofactors,
Putative uncharacterized	37 (37)	147-	Q8KD20	CT12	21584.4	cytoplasmic	
Phytoene desaturase	52 (38)	180-	Q8KCM8	CT13	36031.2	unknown	Biosynthesis of cofactors,
Serine protease	50 (33)	83-	Q8KCH4	CT14	53528.1	periplasm	Protein fate: Degradation of
Rod shape-determining	67 (38)	190-	Q8KCF2		37973.8	Cytoplasm	Cell envelope: Biosynthesis
Sec-independent protein	46 (36)	17-	Q8KCA8	CT15	7435.6	unknown	Protein fate: Protein and
Putative uncharacterized	42 (38)	222-	Q8KC57	CT15	46413.0	cytoplasm	
Holliday junction ATP-	56 (38)	185-	Q8KC00	CT16	37370.9	cytoplasm	DNA metabolism: DNA
Protease IV	74 (38)	418-	Q8KBY8	CT16	64208.2	cytoplasmic	Protein fate: Degradation of
Es1 family protein	35 (32)	61-	Q8KBV2		22871.6	cytoplasm	Unknown function: General
Putative uncharacterized	39 (32)	9- 18	Q8KBU1	CT16	12454.4	cytoplasm	
Type III restriction	46 (37)	892-	Q8KBQ4	CT17	107460.9	cytoplasm	
MoaA/nifB/pqqE family	39 (37)	85-	Q8KBN6		42556.1	cytoplasm	Unknown function: General
Receptor, putative	39 (38)	157-	Q8KBN2	CT17	76404.9	outer membrane	Transport and binding
ComEA-related protein	33 (30)	76-	Q8KB66	CT19	18586.5	unknown	Unknown function: General
	67 (31)	84-					
Ferric siderophore	45 (37)	598-	Q8KB41	CT19	93821.8	outer membrane	Transport and binding
C-type cytochrome,	36 (31)	59-	Q8KAX5		16776.2	unknown	Central intermediary
	50 (32)	66-					
	43 (36)	92-					
	50 (36)	93-					
	85 (37)	97-					
	80 (38)	102-					
Flavocytochrome c	46 (32)	17-	Q8KAS5	CT20	11437.9	periplasm	Central intermediary
Outer surface protein,	50 (32)	77-	Q8KAL2	CT21	20056.5	outer membrane	Cell envelope: Biosynthesis
Transposase, putative	44 (43)	289-	Q8KAK4	CT21	35134.0	unknown	Other categories:
6-phosphogluconate	28 (25)	115-	Q8KBB5		31368.4	cytoplasm	Energy metabolism: Pentose
unique water soluble							
Putative Holliday junction	55 (25)	104-	Q8KG86	CT00	15704.96	cytoplasmic	Hypothetical proteins:
Exopolyphosphatase,	45 (26)	211-	Q8KG69	CT00	35799.32	cytoplasmic	Central intermediary
Multidrug resistance	58 (26)	2- 23	Q8KG63	CT01	35184.50	Cytoplasmic	Transport and binding
Acetyl-coenzyme A	45 (26)	2- 24	Q8KG09	CT01	37324.93	cytoplasmic	Fatty acid and phospholipid
Penicillin-binding protein 1	77 (26)	323-	Q8KFZ4	CT01	83628.42	cytoplasmic	Cell envelope: Biosynthesis
	36 (26)	414-					
sensor histidine kinase	52 (26)						
	28 (26)	275-	Q8KFX5	CT01	47741.70	Cytoplasmic	Regulatory functions: Other
Putative uncharacterized	30 (26)	731-	Q8KFW2	CT02	162849.19	unknown	
Ribosome-binding factor	47 (27)	27-	Q8KFT0	CT02	13976.07	cytoplasmic	Transcription: RNA
Riboflavin biosynthesis	43 (31)	215-	Q8KFS8	CT02	35069.39	cytoplasmic	Biosynthesis of cofactors,
30S ribosomal protein	74 (26)	55-	Q8KFS7	CT02	10326.11	cytoplasmic	Protein synthesis: Ribosomal
UDP-N-	41 (26)	2- 15	Q8KFS1	CT02	26985.92	cytoplasmic	Cell envelope: Biosynthesis
Putative uncharacterized	37 (26)	104-	Q8KFR9	CT02	15966.78	unknown	
Holliday junction ATP-	38 (26)	1- 6	Q8KFR0	CT02	21539.98	cytoplasmic	DNA metabolism: DNA
GTP-binding protein	38 (26)	344-	Q8KFN7	CT02	39339.26	cytoplasmic	Unknown function: General
Cell division protein FtsH	33 (26)	1- 12	Q8KFM5	CT02	73225.20	cytoplasmic	Cellular processes: Cell
Nucleoside-	43 (31)	216-	Q8KFJ6	CT03	24410.67	cytoplasmic	Unknown function: General
Death on curing protein	39 (25)	15-	Q8KFJ3	CT03	14553.81	cytoplasmic	Other categories: Prophage
Putative uncharacterized	67 (27)	1- 15	Q8KF49	CT04	32619.53	cytoplasm	Hypothetical proteins:
Putative uncharacterized	69 (25)	58-	Q8KF07	CT05	15066.52	cytoplasmic	Hypothetical proteins:
Cold shock-like protein	30 (27)	86-	Q8KES2	CT06	10577.87	cytoplasmic	Cellular processes:
3-isopropylmalate or Beta-IPM	33 (26)	135-	P59028		38169.44	cytoplasmic	Amino acid biosynthesis:
		67 (26)					
ExbD/TolR family protein	63 (26)	1- 8	Q8KEQ2	CT06	15571.43	cytoplasmic	Transport and binding

Peptidoglycan-associated	58 (26)	109-	Q8KEP8	CT06	17849.75	outer membrane	Transport and binding
RNA polymerase sigma-	31 (30)	181-	Q8KEN8		23492.39	cytoplasmic	Transcription: Transcription
Putative uncharacterized	46 (26)	1- 7	Q8KEL2		62565.51	cytoplasmic	Hypothetical proteins:
Thiazole biosynthesis	33 (26)	242-	Q8KEJ1	CT06	27070.33	cytoplasmic	Biosynthesis of cofactors,
Putative uncharacterized	32 (30)	2- 12	Q8KEE1	CT07	14421.19	cytoplasmic	Hypothetical proteins:
Fumarate hydratase,	28 (25)	207-	Q8KE57	CT08	54186.90	cytoplasmic	Energy metabolism: TCA
Putative uncharacterized	54 (27)	2- 11	Q8KE09	CT08	59411.9	unknown	
Phosphate transport	46 (26)	64-	Q8KDZ4	CT09	26232.2	cytoplasmic	Regulatory functions: DNA
Pentapeptide repeat	58 (26)	298-	Q8KDX6	CT09	40536.5	unknown	Unknown function: General
Putative uncharacterized	40 (31)	1- 20	Q8KDU5	CT09	14486.6	unknown	
	45 (26)						
tRNA pseudouridine	43 (26)	25-	Q8KDR7	CT09	27737.8	unknown	Protein synthesis: tRNA and
Ribose-5-phosphate	75 (26)	83-	Q8KDK0	CT10	15794.7	unknown	Energy metabolism: Pentose
tRNA (guanine-N(1)-)	41 (26)	79-	Q8KD90	CT11	25869.8	cytoplasmic	Protein synthesis: tRNA and
Putative uncharacterized	44 (27)	82-	Q8KD81	CT11	16814.4	unknown	Hypothetical proteins:
Acetyltransferase GNAT	63 (26)	154-	Q8KCY9	CT12	17881.4	cytoplasmic	Central intermediary
Alcohol dehydrogenase,	32 (26)	312-	Q8KCY2	CT12	36402.6	cytoplasmic	Energy metabolism:
Inosine-5'-	26 (24)	1- 33	Q8KCW4	CT12	53019.2	cytoplasmic	Purines, pyrimidines,
Thymidylate kinase or	47 (26)	1- 22	Q8KCU7	CT13	24146.6	cytoplasmic	Purines, pyrimidines,
Putative uncharacterized	66 (26)	184-	Q8KCU6	CT13	51009.0	cytoplasmic	Hypothetical proteins:
2-C-methyl-D-erythritol 4-	94 (26)	1- 15	Q8KCU3	CT13	27022.0	unknown	Biosynthesis of cofactors,
Putative uncharacterized	39 (26)	1- 10	Q8KCQ9	CT13	15960.1	unknown	
Myo-inositol-1(Or 4)-	70 (26)	232-	Q8KCJ1	CT14	28114.0	unknown	Fatty acid and phospholipid
	59 (26)	245-					
Putative uncharacterized	76 (26)	108-	Q8KCD3		29189.9	unknown (non	
Transcriptional regulator,	75 (31)	2- 7	Q8KCC3		15692.9	cytoplasmic	Regulatory functions: DNA
NifB protein	30 (26)	336-	Q8KC85	CT15	46793.8	cytoplasmic	Biosynthesis of cofactors,
50S ribosomal protein	73 (27)	65-	Q8KC50		8404.6	unknown	Protein synthesis: Ribosomal
DNA primase, putative	29 (25)	1- 6	Q8KC38		70220.7	cytoplasmic	DNA metabolism: DNA
	30 (26)	104-					
Periplasmic divalent	49 (26)	2- 19	Q8KC19		12609.2	cytoplasmic	Cellular processes:
Prephenate dehydratase	68 (26)	121-	Q8KBW6	CT16	31099.3	cytoplasmic	Amino acid biosynthesis:
Shikimate dehydrogenase	58 (30)	198-	Q8KBH8		31733.5	unknown	Amino acid biosynthesis:
Purine nucleoside	36 (26)	154-	Q8KBH2	CT18	29968.8	cytoplasmic	Amino acid biosynthesis:
	39 (27)	183-					
Cytochrome d ubiquinol	42 (26)	1- 8	Q8KBG9	CT18	49750.2	cytoplasmic	Energy metabolism: Electron
Thioesterase,	30 (26)	133-	Q8KBE3	CT18	30730.2	cytoplasmic	Biosynthesis of cofactors,
Putative uncharacterized	33 (26)	34-	Q8KBC7	CT18	64280.8	unknown (non	Hypothetical proteins:
Putative uncharacterized	48 (26)	115-	Q8KBC2	CT18	24297.1	periplasmic	Hypothetical proteins:
Collagenase	60 (27)	253-	Q8KB99	CT18	44765.0	cytoplasmic	Protein fate: Degradation of
Acetyltransferase,	70 (26)	143-	Q8KB90		17983.4	cytoplasmic	Central intermediary
3-phosphoshikimate 1-	91 (26)	286-	Q8KB71	CT19	46626.4	cytoplasmic	Amino acid biosynthesis:
Dihydroflavonol 4-	93 (26)	2- 18	Q8KB60	CT19	36515.7	cytoplasmic	Unknown function: General
2-amino-4-hydroxy-6-	68 (26)	2- 17	Q8KB53	CT19	19396.3	cytoplasmic	Biosynthesis of cofactors,
Putative uncharacterized	51 (26)	2- 8	Q8KB01		16018.5	unknown	Hypothetical proteins:
Succinate/fumarate	39 (26)	87-	Q8KAW1		31422.5	cytoplasmic	Energy metabolism:
Putative uncharacterized	58 (26)	1- 10	Q8KAT6		12111.7	unknown	
Oxidoreductase, short-	45 (25)	1- 23	Q8KAS0	CT20	24083.7	cytoplasmic	Central intermediary
Putative uncharacterized	36 (26)	269-	Q8KAR0	CT20	34058.3	unknown	
Adhesion protein, putative	28 (25)	206-	Q8KAQ0	CT21	37924.3	periplasmic	Cell envelope: Surface
30S ribosomal protein	64 (31)	62-	Q8KAM3	CT21	8671.1	unknown	Protein synthesis: Ribosomal
AP (Apurinic or	37 (26)	217-	Q8KAM0	CT21	28638.8	cytoplasmic	DNA metabolism: DNA
Putative uncharacterized	43 (26)	1- 19	Q8KAL8	CT21	19417.9	unknown	
Putative uncharacterized	47 (26)	1- 15	Q8KAL6	CT21	15839.2	unknown	Hypothetical proteins:
	71 (26)	64-					
Probable GTP-binding	71 (26)	32-	Q8KAK5	CT21	21171.2	unknown	Unknown function: General
50S ribosomal protein L2	62 (26)	157-	Q8KAH5	CT21	30373.2	unknown	Protein synthesis: Ribosomal

50S ribosomal protein L4	43 (31)	80-	Q8KAH3	CT21	22643.1	cytoplasm	Protein synthesis: Ribosomal
	39 (25)	96-					
Siroheme synthetase,	30 (26)	192-	Q8KAC5	CT22	51106.4	cytoplasmic	Biosynthesis of cofactors,
UPF0102 protein CT2262	54 (26)	71-	Q8KAA4	CT22	14386.5	cytoplasmic	Hypothetical proteins:
Putative uncharacterized	38 (32)	44-	Q8KGA3	CT00	12954.2	unknown	Hypothetical proteins:
Putative uncharacterized	49 (32)	58-	Q8KG24	CT01	12026.2	cytoplasmic	
50S ribosomal protein L11	62 (30)	78-	Q8KG19	CT01	14951.6	unknown	Protein synthesis: Ribosomal
	35 (32)	83-					
	70 (37)	118-					
tRNA pseudouridine	73 (31)	50-	Q8KFS9	CT02	26846.7	cytoplasmic	Protein synthesis: tRNA and
Alpha-amylase family	48 (37)	210-	Q8KFM2	CT03	132099.1	unknown	Energy metabolism:
ABC transporter, ATP-	50 (38)	235-	Q8KFK0		27007.3	cytoplasmic	Transport and binding
Molybdenum ABC	33 (32)	64-	Q8KF79	CT04	24222.8	cytoplasmic	Transport and binding
Glutamate synthase,	65 (37)	359-	Q8KFS9		51394.1	cytoplasmic	Amino acid biosynthesis:
Tyrosyl-tRNA synthetase	51 (38)	274-	Q8KEW9		45724.2	cytoplasmic	Protein synthesis: tRNA
Transcriptional regulator,	88 (32)	20-	Q8KEP0		14962.3	cytoplasmic	Regulatory functions: DNA
	84 (38)	107-					
	65 (38)	108-					
Trans-sulfuration enzyme	38 (32)	45-	Q8KEI8		41321.2	cytoplasmic	Amino acid biosynthesis:
Sulfate transporter family	78 (36)	451-	Q8KEH5	CT07	65728.8	cytoplasmic	Transport and binding
	63 (36)	455-					
Putative uncharacterized	36 (32)	3- 10	Q8KEG6	CT07	10145.8	unknown	
Putative uncharacterized	45 (32)	3- 18	Q8KEF7		10086.5	cytoplasmic	Hypothetical proteins:
Probable thiol	53 (37)	123-	Q8KED5	CT07	17714.3	unknown	Cellular processes:
Putative uncharacterized	37 (32)	69-	Q8KE98	CT07	15612.3	unknown	
Thioredoxin-2	69 (32)	58-	Q8KE49	CT08	11858.7	cytoplasmic	Energy metabolism: Electron
	62 (32)	59-					
	57 (33)	65-					
DsrF protein	84 (29)	68-	Q8KE36		14033.0	unknown	Central intermediary
DsrH protein	62 (32)	8- 20	Q8KE35	CT08	10824.3	unknown	Central intermediary
Phosphoglycolate	44 (37)	207-	Q8KE05		24570.2	unknown	Energy metabolism: Sugars
Nitroreductase family	36 (31)	31-	Q8KDV7	CT09	24132.6	cytoplasmic	Central intermediary
Cob(I)alamin	38 (36)	17-	Q8KDV3	CT09	19193.0	cytoplasmic	Biosynthesis of cofactors,
Putative uncharacterized	49 (32)	55-	Q8KDU2	CT09	17312.6	unknown (non	Hypothetical proteins:
Thiol:disulfide interchange	32 (28)	29-	Q8KDM3		19842.3	periplasmic	Central intermediary
Rubredoxin 1	35 (31)	64-	P58992	CT11	7821.8	cytoplasmic	Energy metabolism: Electron
Acetylglutamate kinase or	42 (35)	255-	P59297	CT11	30155.8	cytoplasmic	Amino acid biosynthesis:
CRISPR-associated	34 (31)	79-	Q8KDC4	CT11	38755.8	cytoplasmic	Hypothetical proteins:
Putative uncharacterized	36 (32)	54-	Q8KDA5	CT11	48934.9	cytoplasmic	Hypothetical proteins:
Putative uncharacterized	54 (30)	42-	Q8KD50	CT12	28401.8	unknown	Hypothetical proteins:
Putative uncharacterized	64 (32)	5- 18	Q8KD38	CT12	6837.5	cytoplasmic	
Putative uncharacterized	49 (37)	120-	Q8KD19	CT12	14487.6	unknown	
Protein translocase	64 (38)	947-	Q8KD18	CT12	118004.1	cytoplasmic	Protein fate: Protein and
Sensory box protein	38 (33)	13-	Q8KCN5		27003.8	cytoplasmic	Regulatory functions: Protein
Putative uncharacterized	41 (31)	11-	Q8KCK2		10430.7	unknown	
Putative uncharacterized	49 (36)	92-	Q8KC32	CT15	22316.0	unknown	
Putative uncharacterized	34 (32)	5- 14	Q8KBY9		18406.0	unknown	
Putative uncharacterized	35 (32)	44-	Q8KBX9		5747.5	unknown	
Bifunctional enzyme	41 (37)	300 -	Q8KBX0	CT16	51844.2	cytoplasmic	Cell envelope: Biosynthesis
Indole-3-glycerol	60 (37)	239-	Q8KBW1		28664.7	unknown	Amino acid biosynthesis:
Hydroxylamine reductase	56 (37)	122-	Q8KBT4		58654.1	cytoplasmic	Unknown function: General
Putative uncharacterized	45 (32)	6- 22	Q8KBS2	CT17	16902.1	unknown	
Ferredoxin-3 or FdIII	49 (32)	52-	Q8KBP9	CT17	6426.3	cytoplasmic	Energy metabolism: Electron
Putative uncharacterized	87 (32)	46-	Q8KBG0	CT18	41706.6	unknown	
UPF0235 protein	72 (36)	3- 16	Q8KBF5	CT18	10746.5	unknown	Hypothetical proteins:
Heavy-metal-associated	46 (29)	42-	Q8KBF2	CT18	7487.5	cytoplasmic	Transport and binding
ABC transporter, ATP-	60 (30)	5- 20	Q8KB88	CT19	27694.9	cytoplasmic	Transport and binding

Putative uncharacterized	45 (33)	26-	Q8KB75	CT19	51728.8	cytoplasmic	
Trigger factor	49 (32)	27-	Q8KB57	CT19	48608.6	cytoplasmic	Protein fate:Protein and
	73 (32)	28-					
WD-repeat family protein	59 (37)	94-	Q8KB12	CT19	35899.9	unknown	Unknown function: General
50S ribosomal protein	52 (36)	29-	Q8KAI3	CT21	8968.6	cytoplasmic	DNA metabolism: Ribosomal
Putative uncharacterized	52 (32)	61-	Q8KAD4	CT22	7922.0	unknown	Hypothetical proteins:
Cytochrome DsrJ	56 (32)	36-	Q8KAC1	CT22	13020.8	cytoplasmic	Energy metabolism: Electron
Ribonuclease G	61 (32)	76-	Q8KAA6	CT22	63233.0	Cytoplasmic	Transcription: RNA
	71 (32)	77-					
Putative uncharacterized	73 (37)	65-	Q8KA93	CT22	26706.3	unknown	Hypothetical proteins:
Peptidase, M23/M37	33 (26)	201-	Q4W535	CT19	52148.2	cytoplasmic	
Putative uncharacterized	52 (27)	159-	Q4W552	CT05	24381.5	unknown	
Nitroreductase family	30 (26)	9- 16	Q8KFI1	CT03		cytoplasmic	
unique membrane							
DNA	44 (30)	2- 10	Q8KF21	CT05	52540.75	Cytoplasmic	DNA metabolism: DNA
ABC-type export system,	60 (26)	172-	Q8KEX8	CT05	53358.5	outer membrane	Transport and binding
Cell division ATP-binding	33 (25)	1- 12	Q8KEU2	CT05	26496.0	cytoplasmic	Cellular processes: Cell
ATP-dependent hsl	31 (27)	475-	Q8KD63	CT11	55513.9	Cytoplasmic	Protein fate: Protein folding
Putative uncharacterized	52 (26)	164-	Q8KCA5		22386.6	unknown	Hypothetical proteins:
Peptide ABC transporter,	52 (25)	77-	Q8KBZ8		61554.1	periplasmic	Transport and binding
ABC transporter, ATP-	30 (25)	419-	Q8KBG5	CT18	62382.4	cytoplasmic	Transport and binding
Putative uncharacterized	33 (32)	37-	Q8KG50	CT01	6314.5	unknown	
Tryptophan synthase beta	60 (33)	49-	Q8KF11	CT05	43639.1	cytoplasmic	Amino acid biosynthesis:
Cytosine deaminase	32 (32)	35-	Q8KEU4	CT05	16103.2	cytoplasmic	Purines, pyrimidines,
Chlorosome protein J	94 (37)	183-	Q68983	CT06	23895.6	chlorosome	Energy metabolism:
LAO/AO transport	48 (37)	113-	Q8KE91	CT07	36488.5	cytoplasmic	Transport and binding
DNA polymerase III,	38 (38)	718-	Q8KE50	CT08	132765.0	cytoplasmic	DNA metabolism: DNA
TRNA methyltransferase	40 (37)	165-	Q8KE14		25863.7	unknown	Protein synthesis: tRNA and
Cobalamin synthase	79 (32)	14-	Q8KDU6	CT09	26726.8	cytoplasmic	Biosynthesis of cofactors,
	66 (34)						
Putative uncharacterized	48 (32)	30-	Q8KDI6		25294.6	unknown	Hypothetical proteins:
Arginine biosynthesis	42 (37)	315-	P59611	CT11	22481.4	cytoplasmic	Amino acid biosynthesis:
ADP-L-glycero-D-	34 (32)	85-	Q8KCZ9		37722.7	unknown	Cell envelope: Biosynthesis
Bacteriochlorophyll	81 (31)	38-	Q8KCY7	CT12	40301.3	cytoplasmic	Biosynthesis of cofactors,
Dihydroneopterin	33 (32)	78-	Q8KB54		14030.0	cytoplasmic	Biosynthesis of cofactors,
Photosystem P840	36 (32)	39-	Q8KAY0	CT20	81713.1	cytoplasmic	Energy metabolism:
cytoplasm-periplasm							
2-dehydro-3-	44 (26)	262-	Q8KG80		30318.0	cytoplasm	Cell envelope: Biosynthesis
ATP-dependent Clp	78 (26)	1- 8	Q8KFC3	CT04	48331.6	unknown	Protein fate: Degradation of
S-	74 (26)	1- 14	Q8KEY7	CT05	37546.9	cytoplasm	Protein synthesis: tRNA and
Chaperone protein htpG	71 (31)	2- 10	Q8KE61	CT08	71037.2	cytoplasm	Protein fate: Protein folding
Putative uncharacterized	37 (31)	2- 7	Q8KDF6	CT10	93345.3	unknown	
Lysyl-tRNA synthetase or	33 (25)	2- 26	Q8KCM7	CT13	58338.7	Cytoplasm	Protein synthesis: tRNA
NADPH-dependent 7-	28 (26)	1- 16	Q9F719	CT16	13564.3	Cytoplasm	Hypothetical proteins:
Pantothenate synthetase	39 (26)	1- 18	Q8KBY5	CT16	31470.9	Cytoplasm	Biosynthesis of cofactors,
Hydrolase, haloacid	30 (26)	88-	Q8KBS5	CT17	25867.3	Cytoplasm	Central intermediary
Oxidoreductase,	39 (27)	37-	Q8KBB8		29879.2	Cytoplasm	Central intermediary
Putative uncharacterized	56 (27)	2- 10	Q8KB74		79467.0	unknown	
Cysteine	49 (25)	1- 22	Q8KB68		49945.0	Cytoplasm	Amino acid biosynthesis:
Putative uncharacterized	63 (26)	1- 8	Q8KAU7	CT20	38826.9	cytoplasm	Hypothetical proteins:
Alpha-amylase family	41 (25)	1- 12	Q8KAR6		76607.8	cytoplasm	Energy metabolism:
Adenylosuccinate lyase	42 (33)	16-	Q8KDQ6	CT09	49191.5	cytoplasm	Purines, pyrimidines,
Acetate kinase	39 (37)	202-	Q8KC99		44575.8	cytoplasm	Energy metabolism:
Cytochrome c, putative	85 (32)	29-	Q8KBQ1	CT17	15103.4	unknown	Hypothetical proteins:
Putative uncharacterized	45 (32)	54-	Q8KBF8		27510.2	cytoplasmic	Hypothetical proteins:
Elongation factor G	53 (41)	435-	Q8KAG9	CT21	77951.8	Cytoplasm	Protein synthesis:

cytoplasm-water soluble							
UDP-N-acetylmuramoyl-Adenosylmethionine--8-	38 (24)	1- 5	Q8KGC9	CT00	54301.82	cytoplasm	Cell envelope: Biosynthesis
Glycerophosphoryl diester	58 (26)	2- 9	Q8KGC1	CT00	46901.5	cytoplasm	Biosynthesis of cofactors,
Probable inorganic	30 (27)	150-	Q8KG89	CT00	29973.08	cytoplasm	Fatty acid and phospholipid
Glycosyl hydrolase, family	95 (26)	1- 11	Q8KG83	CT00	30988.51	cytoplasm	Hypothetical proteins:
Aspartokinase	65 (31)	179-	Q8KG78	CT00	40437.47	unknown	Energy metabolism:
1-deoxy-D-xylulose 5-Transcription	58 (26) 99 (26)	1- 18 2- 18	Q8KG73	CT00	50416.77	cytoplasm	Amino acid biosynthesis:
50S ribosomal protein L1	47 (25)	1- 22	Q8KG43	CT01	41343.71	cytoplasm	Biosynthesis of cofactors,
ATP-dependent Clp	74 (27)	176-	Q8KG20	CT01	21580.10	cytoplasm	Transcription: Transcription
DNA polymerase,	46 (26)	2- 9	Q8KG18	CT01	24797.68	cytoplasm	Protein synthesis: Ribosomal
Cytidyltransferase family	36 (26)	1- 8	Q8KFX9	CT01	95863.05	cytoplasm	Protein fate: Degradation of
Aspartate 1-	62 (25)						
4-hydroxy-3-methylbut-2-	40 (26)	147-					
Putative uncharacterized	51 (26)	329-					
Adenine	33 (31)	1- 30	Q8KFW4	CT02	23223.05	cytoplasm	DNA metabolism: DNA
DTDP-4-	52 (26)	2- 8	Q8KFW1	CT02	17788.52	cytoplasm	Central intermediary
Phosphoribosylglycinamid	43 (30)	78-	Q8KFS0	CT02	14558.72	cytoplasm	Biosynthesis of cofactors,
Isocitrate dehydrogenase,	51 (26)	1- 7	Q8KFN9		61656.13	cytoplasm	Amino acid biosynthesis:
Acetylornithine	31 (25)	157-	Q8KFN1	CT02	18644.78	unknown	
Putative uncharacterized	34 (26)	9- 21	Q8KFM9	CT02	19319.45	cytoplasmic	Purines, pyrimidines,
Oxygen-independent	74 (26)	1- 9	Q8KFL9	CT03	32396.77	unknown	Cell envelope: Biosynthesis
Precorrin-6X	52 (25)	1- 7	Q8KFK7	CT03	22271.77	cytoplasm	Purines, pyrimidines,
Glutamate synthase,	56 (26)	382-	Q8KFH5		80838.13	unknown	Energy metabolism: TCA
DNA ligase	73 (26)	347-	P59316	CT03	43548.11	cytoplasm	Amino acid biosynthesis:
1-(5-phosphoribosyl)-5-	98 (26)	28-	Q8KFF7	CT03	18685.62	unknown	Hypothetical proteins:
Putative uncharacterized	28 (26)	2- 19	Q8KFF5	CT03	55089.39	cytoplasm	Biosynthesis of cofactors,
RNA polymerase sigma-Sensor protein	48 (26)	446-					
Rod shape-determining	61 (25)	1- 12	Q8KFE3	CT03	65757.77	cytoplasmic	Biosynthesis of cofactors,
Chaperonin, 10 kDa. 10	33 (26)	153-					
Quinolinate synthetase A	71 (25)	2- 11	Q8KFC5	CT04	54086.94	cytoplasm	Amino acid biosynthesis:
Metallo-beta-lactamase	43 (31)	1- 14	Q8KF74	CT04	74783.62	cytoplasm	DNA metabolism: DNA
DNA-binding response	33 (31)						
Nudix/MutT family protein	39 (26)	1- 16	Q8KF55	CT04	28711.16	cytoplasm	Amino acid biosynthesis:
Nucleotidyltransferase	74 (25)	1- 9	Q8KF39		21144.05	cytoplasm	Hypothetical proteins:
Serine cycle enzyme,	75 (26)	2- 12	Q8KF30	CT05	22402.48	unknown	Transcription: Transcription
Universal stress protein	74 (26)	2- 9	Q8KF01	CT05	79024.83	cytoplasmic	Regulatory functions: Other
Adenylylsulfate reductase,	55 (26)	2- 9	Q8KEY5	CT05	36893.88	cytoplasm	Cell envelope: Biosynthesis
Internalin-related protein	78 (26)	321-					
Sulfite reductase,	28 (26)	1- 6	Q8KEW3	CT05	14767.76	cytoplasm	Protein fate: Protein folding
Adenylylsulfate reductase,	34 (26)	122-					
Quinolinate synthetase A	54 (27)	297-	Q8KEW2	CT05	35718.38	cytoplasm	Biosynthesis of cofactors,
Metallo-beta-lactamase	99 (27)	1- 10	Q8KET4	CT05	53146.31	cytoplasm	Unknown function: Enzymes
DNA-binding response	81 (27)	2- 13	Q8KEN1	CT06	26796.81	cytoplasm	Regulatory functions: DNA
Nudix/MutT family protein	41 (26)	221-					
NADH-quinone	82 (26)	1- 13	Q8KEK1	CT06	16147.36	cytoplasm	DNA metabolism: DNA
Universal stress protein	50 (26)	1- 9	Q8KEJ4	CT06	11132.95	cytoplasm	Central intermediary
Oxaloacetate	42 (27)	130-	Q8KEG9		22071.71	unknown	Central intermediary
Internalin-related protein	37 (26)	124-	Q8KEG0	CT07	15352.74	cytoplasm	Unknown function: Enzymes
Sulfite reductase,	73 (26)	1- 25	Q8KEC1		19882.59	unknown	Energy metabolism: Electron
Adenylylsulfate reductase,	37 (25)	1- 22	Q8KE92	CT07	16483.40	unknown	Cellular processes:
Adenylylsulfate reductase,	73 (26)	45-					
Internalin-related protein	31 (24)	1- 5	Q8KE56	CT08	68429.57	cytoplasm	Energy metabolism:
Sulfite reductase,	33 (26)	176-	Q8KE54	CT08	45608.15	cytoplasm	Unknown function: General
Adenylylsulfate reductase,	38 (27)	408-	Q8KE5E9	CT08	46350.8	unknown	Central intermediary
Adenylylsulfate reductase,	89 (26)	97-	Q8KE27	CT08	73418.4	unknown	Central intermediary
Adenylylsulfate reductase,	92 (26)	178-					

	29 (26)	613-					
	56 (26)	651-					
Iron-sulfur cluster-binding	81 (25)	145-	Q8KE38	CT08	62177.9	cytoplasm	Central intermediary
Heterodisulfide reductase,	40 (26)	489-	Q8KE25	CT08	82601.4	unknown	Central intermediary
Methyltransferase,	34 (30)	201-	Q8KE04	CT08	23723.8	cytoplasm	Central intermediary
Phosphate import ATP-	38 (27)	1- 18	Q8KDZ5	CT08	31931.7	cytoplasm	Transport and binding
Putative uncharacterized	36 (26)	1- 17	Q8KDY7	CT09	27590.7	cytoplasm	
Type III restriction system	74 (26)	1- 12	Q8KDY6	CT09	74929.2	cytoplasm	DNA metabolism:
Putative uncharacterized	30 (25)	2- 10	Q8KDY5	CT09	45453.7	unknown	
Putative uncharacterized	50 (26)	1- 7	Q8KDV0	CT09	29573.5	cytoplasm	
Cobalamin biosynthesis	75 (27)	2- 12	Q8KDU9	CT09	19295.3	cytoplasm	Biosynthesis of cofactors,
Carbohydrate kinase,	72 (26)	2- 20	Q8KDR9	CT09	34008.4	cytoplasm	Energy metabolism: Sugars
	63 (26)	281-					
Oxidoreductase,	31 (26)	325-	Q8KDQ4	CT09	38144.8	cytoplasm	Central intermediary
Sulfide dehydrogenase,	34 (27)	2- 22	Q8KDM1	CT10	45492.6	unknown	Central intermediary
	28 (26)						
Polyphosphate kinase 2	72 (26)	2- 20	O68984	CT10	81891.4	cytoplasm	Central intermediary
	32 (26)						
	28 (26)	522-					
Putative uncharacterized	45 (25)	1- 6	Q8KDI4	CT10	23434.1	cytoplasm	Hypothetical proteins:
Putative uncharacterized	32 (27)	262-	Q8KDG6	CT10	32178.8	cytoplasm	Hypothetical proteins:
	64 (26)	281-					
Citrate lyase, subunit1	28 (26)	385-	Q8KDG1	CT10	43776.2	cytoplasm	Energy metabolism: TCA
	37 (26)						
Arginine repressor	42 (24)	1- 6	Q8KDE1	CT11	16224.6	cytoplasm	Regulatory functions: DNA
Nitroreductase family	82 (26)	1- 10	Q8KDA4	CT11	30792.6	cytoplasm	Central intermediary
Type III restriction system	65 (26)	1- 9	Q8KD95	CT11	24114.6	unknown	DNA metabolism:
Type III restriction system	47 (26)	2- 10	Q8KD94	CT11	59223.3	cytoplasm	DNA metabolism:
	33 (27)						
Florfenicol resistance	54 (26)	1- 10	Q8KD71	CT11	42197.8	cytoplasm	Cellular processes: Toxin
	71 (26)						
	31 (30)	304-					
Thioredoxin	85 (26)	203-	Q8KD40	CT12	30037.1	cytoplasm	Biosynthesis of cofactors,
Hydrogenase, iron-sulfur	32 (26)	2- 8	Q8KD09	CT12	30718.2	Cytoplasm	Central intermediary
Methylated-DNA--protein-	40 (26)	1- 16	Q8KCW8	CT12	17915.6	unknown	DNA metabolism: DNA
Magnesium-chelatase 38	37 (27)	92-	Q93SW1	CT12	42799.6	Cytoplasm	Biosynthesis of cofactors,
ComA2 protein	33 (31)	2- 28	Q8KCU2	CT13	16168.4	unknown	Unknown function: General
Putative uncharacterized	74 (27)	7- 15	Q8KCT3	CT13	18606.2	unknown	Hypothetical proteins:
UDP-3-O-[3-	92 (26)	1- 12	Q8KCQ3	CT13	36726.0	unknown	Cell envelope: Biosynthesis
Diaminopimelate	43 (25)	357-	Q8KCP2	CT13	45694.1	unknown	Amino acid biosynthesis:
	29 (25)						
Putative uncharacterized	49 (27)	175-	Q8KCN8	CT13	29043.7	cytoplasm	Hypothetical proteins:
	52 (26)	243-					
Queuine tRNA-	52 (27)	1- 10	Q8KCL8		42661.5	cytoplasm	Protein synthesis: tRNA and
Putative uncharacterized	50 (31)	178-	Q8KCK0		22274.4	unknown	Hypothetical proteins:
Prolyl-tRNA synthetase or	53 (26)	2- 8	Q8KCD1	CT14	54950.4	Cytoplasm	Protein synthesis: tRNA
	67 (25)	22-					
Carbon-nitrogen	48 (25)	1- 7	Q8KCC8	CT14	31773.1	unknown	Central intermediary
Carbon-nitrogen	52 (25)	1- 7	Q8KCB4	CT15	32582.6	cytoplasm	Central intermediary
Rab family protein	65 (27)	2- 8	Q8KC98		123777.4	extracellular	Unknown function: General
Thymidylate synthase,	55 (26)	82-	Q8KC72	CT15	25513.1	Cytoplasm	Purines, pyrimidines,
Ribosome recycling factor	65 (26)	30-	Q8KC49		21222.5	Cytoplasm	Protein synthesis:
Recombination/replication	56 (26)	1- 8	Q8KC12		62979.4	cytoplasm	DNA metabolism: DNA
Extragenic suppressor	48 (26)	1- 22	Q8KBX1	CT16	28514.3	Cytoplasm	Unknown function: General
Dihydropteroate	65 (26)	2- 11	Q8KBS7	CT17	32809.5	Cytoplasm	Biosynthesis of cofactors,
N utilization substance	32 (27)	1- 5	Q8KBS4	CT17	19214.2	Cytoplasm	Transcription: Transcription
Ferrous iron transport	82 (26)	1- 9	Q8KBP3	CT17	84466.4	cytoplasmic	Transport and binding

Putative uncharacterized	48 (31)	4- 13	Q8KBL8		14895.8	unknown	
Putative uncharacterized	28 (26)	94-	Q8KBL3	CT17	32521.6	unknown	Hypothetical proteins:
GTP-binding protein engA	81 (27)	1- 10	Q8KBK3	CT17	48977.5	cytoplasm	Unknown function: General
Peptidyl-prolyl cis-trans	83 (25)	2- 24	Q8KBH4		17897.1	periplasm	Protein fate: Protein folding
O-succinylbenzoic acid--	88 (25)	1- 22	Q8KBE0	CT18	49152.3	Cytoplasm	Biosynthesis of cofactors,
Type I restriction system	43 (26)	449-	Q8KBB1		123529.7	unknown	DNA metabolism:
	26 (25)	889-					
Hydrogenase/sulfur	68 (26)	1- 12	Q8KB98	CT18	41243.8	cytoplasm	Central intermediary
Putative uncharacterized	50 (26)	73-	Q8KB82	CT19	12045.5	unknown	
Putative uncharacterized	36 (25)	1- 6	Q8KB69		21086.0	cytoplasm	
Putative uncharacterized	87 (26)	1- 12	Q8KB63		9396.7	cytoplasm	Hypothetical proteins:
	57 (26)	2- 12					
Aspartate-semialdehyde	56 (24)	2- 8	Q8KB62	CT19	36840.0	Cytoplasm	Amino acid biosynthesis:
	55 (26)	255-					
CRISPR-associated	49 (26)	1- 11	Q8KB25		20114.7	unknown	
	28 (26)						
CRISPR-associated	63 (31)	1- 14	Q8KB20	CT19	12256.2	unknown	Hypothetical proteins:
IscU protein	52 (25)	124-	Q8KB04	CT19	22620.5	cytoplasm	Biosynthesis of cofactors,
Bacterioferritin	26 (24)	1- 29	Q8KAZ7	CT20	16353.8	unknown	Cellular processes:
Nucleoside diphosphate	91 (26)	86-	Q8KAZ6	CT20	15445.7	Cytoplasm	Purines, pyrimidines,
Putative uncharacterized	65 (25)	1- 17	Q8KAU5		43328.6	cytoplasm	Hypothetical proteins:
	124 (25)						
2-isopropylmalate	70 (26)	2- 20	Q8KAP9	CT21	62370.7	unknown	Amino acid biosynthesis:
Malonyl CoA-acyl carrier	55 (26)	1- 18	Q8KAP1	CT21	32451.1	unknown	Fatty acid and phospholipid
Translation initiation factor	43 (26)	1- 17	Q8KAM9		24810.5	Cytoplasm	Protein synthesis:
30S ribosomal protein	111 (25)	57-	Q8KAJ5	CT21	13834.0	Cytoplasm	Protein synthesis: Ribosomal
50S ribosomal protein L6	49 (27)	86-	Q8KAI6	CT21	19597.9	unknown	Protein synthesis: Ribosomal
50S ribosomal protein L29	56 (26)	1- 24	Q8KAI0	CT21	8083.4	unknown	Protein synthesis: Ribosomal
50S ribosomal protein L16	31 (26)	1- 6	Q8KAH9	CT21	15786.6	unknown	Protein synthesis: Ribosomal
Uncharacterized	72 (26)	2- 9	P59599	CT22	17468.3	Cytoplasm	Central intermediary
Exodeoxyribonuclease 7	42 (25)	87-	Q8KAE9	CT22	10642.9	Cytoplasm	DNA metabolism:
Putative uncharacterized	59 (26)	110-	Q8KAB7	CT22	13640.5	unknown	
DNA gyrase subunit B	59 (26)	604-	Q8KAA3	CT22	72546.5	Cytoplasm	DNA metabolism: DNA
ATP synthase B chain	74 (32)	80-	Q8KGE9	CT00	19436.6	Cytoplasm	Energy metabolism: ATP-
N-acetylmuramoyl-L-	61 (37)	524-	Q8KGB4	CT00	60430.3	unknown	Cell envelope: Biosynthesis
Putative uncharacterized	43 (26)	203-	Q8KFY1		30076.0	cytoplasm	Hypothetical proteins:
Putative uncharacterized	57 (37)	102-	Q8KFX7		12430.9	unknown	
Putative uncharacterized	52 (32)	83-	Q8KEZ5		29526.6	cytoplasm	
ATP-dependent RNA	94 (31)	14-	Q8KET5		73443.5	cytoplasm	Transcription: Other
	46 (32)	20-					
3-isopropylmalate	56 (37)	171-	Q8KER9		20658.4	cytoplasm	Amino acid biosynthesis:
Ketol-acid	88 (37)	249-	Q8KER7	CT06	35858.00	cytoplasm	Amino acid biosynthesis:
Chaperone protein dnaK	103 (37)	616-					
	48 (37)		Q8KEP3	CT06	68387.56	unknown	Protein fate: Protein folding
	62 (37)	617-					
Soluble inorganic	57 (36)	113-	Q8KE67	CT08	15607.9	Cytoplasm	Central intermediary
Sialic acid synthase	72 (37)	148-	Q8KE65	CT08	31218.5	Cytoplasm	Cell envelope: Biosynthesis
	60 (37)	172-					
Sulfite reductase,	62 (32)	73-	Q8KE39	CT08	40108.8	unknown	Central intermediary
Sulfur oxidation protein	40 (37)	89-	Q8KDM8	CT10	10640.2	unknown	Central intermediary
Lipoyl synthase or Lipoic	39 (37)	164-	Q8KDH2	CT10	32247.8	cytoplasm	Biosynthesis of cofactors,
Putative uncharacterized	39 (32)	31-	Q8KDD3	CT11	9782.2	unknown	
Oxidoreductase, FAD-	38 (37)	420-	Q8KDB5	CT11	129688.5	cytoplasm	Central intermediary
DegT/DnrJ/EryC1/StrS	56 (38)	376-	Q8KDA3		41879.0	cytoplasm	Unknown function: General
ATP-dependent protease	56 (33)	11-	Q8KD62	CT11	19668.6	Cytoplasm	Protein fate: Protein folding
Ferredoxin-2	34 (32)	53-	Q8KCZ7		6624.5	Cytoplasm	Energy metabolism: Electron
Ferredoxin-1	40 (36)	52-	Q8KCZ6		6330.1	unknown	Energy metabolism: Electron

Protein grpE or HSP-70	42 (29)	17-	Q8KCD7		21889.9	Cytoplasm	Protein fate: Protein folding
	55 (30)	19-					
	56 (30)	21-					
	49 (30)	22-					
	51 (36)	26-					
Putative uncharacterized	37 (32)	60-	Q8KBT0		13158.6	cytoplasm	
NifU protein, putative	48 (31)	65-	Q8KBK1	CT17	8995.5	cytoplasm	Central intermediary
Putative uncharacterized	96 (38)	481-	Q8KB51	CT19	75041.2	unknown	Hypothetical proteins:
	61 (36)	538-					
CRISPR-associated	40 (37)	494-	Q8KB27		94073.0	unknown	Hypothetical proteins:
Rubredoxin 3 or rbr-3 or	75 (31)	38-	P58025	CT20	5869.4	cytoplasm	Energy metabolism: Electron
LipD protein, putative	47 (38)	229-	Q8KAV2		50088.8	outer membrane	Protein fate: Protein and
ArsA ATPase family	85 (37)	304-	Q8KAT5	CT20	44671.4	cytoplasm	Transport and binding
50S ribosomal protein L35	34 (32)	51-	Q8KAM8	CT21	7527.9	unknown	DNA metabolism: Ribosomal
Single-stranded DNA-	55 (33)	73-	Q8KAM2	CT21	16991.7	cytoplasm	DNA metabolism: DNA
30S ribosomal protein S6	43 (36)	111-	Q8KAM1	CT21	14741.0	cytoplasm	Protein synthesis: Ribosomal
Iron-sulfur cluster-binding	42 (37)	561-	Q8KAB6	CT22	62206.9	cytoplasm	Central intermediary
	45 (37)	565-					
Sulfite reductase,	60 (37)	99-	Q8K5F4	CT08	12213.0	Cytoplasm	Central intermediary
Peptidyl-prolyl cis-trans	81 (33)	21-	Q8KAA2	CT22	48562.7	periplasm	Protein fate: Protein folding
cytoplasm-membrane							
Histidyl-tRNA synthetase	53 (26)	2- 11	Q8KFT6		48131.1	cytoplasm	Protein synthesis: tRNA
Hydrolase, alpha/beta	60 (26)	1- 9	Q8KDU7	CT09	33133.4	unknown	Central intermediary
Electron transfer	77 (27)	1- 10	Q8KAL5	CT21	31552.0	unknown	Energy metabolism: Electron
water soluble-periplasm							
Putative uncharacterized	63 (26)	115-	Q8KFH9	CT03	20213.57	unknown	
DNA helicase, putative	28 (26)	473-	Q8KC08	CT16	168045.9	Cytoplasm	DNA metabolism: DNA
Putative uncharacterized	56 (27)	1- 10	Q8KBU7	CT16	6757.1	unknown	Hypothetical proteins:
Hydrogenase maturation	35 (24)	70-	Q8KBJ0	CT17	83956.9	Cytoplasm	Protein fate: Protein folding
3-oxoacyl-[acyl-carrier-	69 (26)	1- 24	Q8KAP2		36140.2	Cytoplasm	Fatty acid and phospholipid
50S ribosomal protein L14	28 (26)	1- 20	Q8KAI2	CT21	13411.8	cytoplasm	Protein synthesis: Ribosomal
Glycosyl transferase,	29 (26)	281-	Q8KAF5	CT22	35998.3	cytoplasm	Cell envelope: Biosynthesis
Thiol:disulfide interchange	60 (26)	166-	Q8KAD0	CT22	18991.1	periplasm	Energy metabolism: Electron
Probable chaperone	54 (26)	313-	Q8KA87	CT22	48928.9	Cytoplasm	Protein fate: Degradation of
Outer membrane protein	54 (38)	100-	Q8KFR8	CT02	20449.69	unknown	Cell envelope: Other
Putative uncharacterized	66 (37)	245-	Q8KF98	CT04	29034.7	unknown	Hypothetical proteins:
Putative uncharacterized	43 (32)	50-	Q8KEX5		21303.5	unknown	
Chlorosome envelope	37 (26)	1- 8	Q8KEN5		23970.4	Chlorosome	Energy metabolism:
Putative uncharacterized	56 (32)	67-	Q8KDT7	CT09	39256.7	unknown	
Serine acetyltransferase	47 (32)	79-	Q8KCI5	CT14	30313.8	Cytoplasm	Amino acid biosynthesis:
Putative uncharacterized	76 (37)	309-	Q8KC16	CT16	39608.6	unknown	
SMC family protein	49 (37)	896-	Q8KBS6		134314.9	cytoplasm	Cellular processes: Cell
Putative uncharacterized	50 (29)	25-	Q8KBP0		29008.2	unknown	
Nudix/MutT family protein	38 (32)	17-	Q8KBI5		15546.6	cytoplasm	Unknown function: Enzymes
Putative uncharacterized	41 (37)	113-	Q8KBC6	CT18	22729.8		Hypothetical proteins:
	76 (37)	115-					
DNA mismatch repair	46 (33)	81-	Q8KAX3		69836.6	Cytoplasm	DNA metabolism: DNA
periplasm- membrane 0							
water soluble-							
Peptidyl-prolyl cis-trans	34 (25)	2- 8	Q8KFW7	CT20	77049.14	periplasmic	Protein fate: Protein folding
Membrane protein,	68 (26)	405-	Q8KF44	CT04	48751.03	cytolasmic	Cell envelope: Other
ABC-type export system,	41 (26)	223-	Q8KEX9	CT05	34024.15	unknown	Transport and binding
Putative uncharacterized	65 (26)	2- 14	Q8KEV7	CT05	32041.74	unknown	
Putative uncharacterized	61 (31)	15-	Q8KEP4	CT06	16141.73	cytoplasmic	

Riboflavin synthase, alpha	94 (27)	1- 17	Q8KED3	CT07	23062.14	unknown	Biosynthesis of cofactors,
Putative uncharacterized	53 (31)	1- 11	Q8KDX7	CT09	47232.5	cytoplasmic	
Dihydroorotate	29 (27)	2- 11	Q8KC29		32269.6	Cytoplasmic	Purines, pyrimidines,
6,7-dimethyl-8-	40 (26)	55-	Q8KAW4		16307.8	unknown	Biosynthesis of cofactors,
Putative uncharacterized	88 (26)	1- 13	Q8KAV5		11261.2	unknown	
Preprotein translocase	47 (25)	1- 8	Q8KAJ1		48295.0	cytoplasmic	Protein fate: Protein and
Putative uncharacterized	55 (37)	35-	Q8KG57	CT01	15626.19	cytoplasmic	Hypothetical proteins:
Peptide chain release	32 (31)	22-	Q8KG45	CT01	40600.78	cytoplasmic	Protein synthesis:
Putative uncharacterized	73 (37)	156-	Q8KFAQ7		22684.6	cytoplasmic	Hypothetical proteins:
	63 (37)	158-					
Dihydrodipicolinate	41 (31)	54-	Q8KC06		32386.1	Cytoplasmic	Amino acid biosynthesis:
Putative uncharacterized	46 (31)	19-	Q8KBH5		26485.3	unknown	
cytoplasm-periplasm-							
Probable	27 (26)	1- 11	Q8KCA0	CT15	90024.6	cytoplasmic	Hypothetical proteins:
Putative uncharacterized	74 (26)	1- 8	Q8KAU7	CT20	38826.9	cytoplasmic	Hypothetical proteins:
water soluble-							
Ferrous iron transport	78 (25)	1- 11	Q8KGB1	CT00	10779.59	cytoplasm	Transport and binding
Carbamoyl-phosphate	96 (26)	1- 17	Q8KGA2	CT00	40110.62	unknown	Amino acid biosynthesis:
DNA-directed RNA	30 (26)	753-	Q8KG15	CT01	145469.06	cytoplasm	Transcription: DNA-
DNA-directed RNA	39 (27)	85-	Q8KG14	CT01	166281.64	cytoplasm	Transcription: DNA-
N utilization substance	28 (26)	222-	Q8KFT2	CT02	60221.84	cytoplasm	Transcription: Transcription
Translation initiation factor	45 (26)	2- 9	Q8KFT1	CT02	100647.32	cytoplasm	Protein synthesis:
Putative uncharacterized	94 (26)	1- 13	Q8KFS6	CT02	33714.55	cytoplasm	Hypothetical proteins:
Guanylate kinase	67 (26)	2- 11	Q8KFS5	CT02	21772.73	cytoplasm	Purines, pyrimidines,
Mannose-6-phosphate	48 (26)	1- 15	Q8KFL7		51881.45	unknown	Cell envelope: Biosynthesis
Putative uncharacterized	51 (26)	1- 7	Q8KFC4	CT04	9981.23	cytoplasm	Hypothetical proteins:
Pseudouridine synthase	77 (26)	1- 11	Q8KF53	CT04	28030.46	cytoplasm	Protein synthesis: tRNA and
SsrA-binding protein	32 (31)	2- 17	Q8KEW8	CT05	18280.35	unknown	Unknown function: General
Putative uncharacterized	34 (26)	2- 6	Q8KES5	CT06	16765.41	cytoplasm	Hypothetical proteins:
Bacterioferritin	48 (26)	2- 26	Q8KEM5		17815.35	cytoplasm	Cellular processes:
Putative uncharacterized	37 (25)	29-	Q8KEI2	CT07	16458.79	cytoplasm	Hypothetical proteins:
GTP cyclohydrolase 1	44 (31)	1- 15	Q8KEA8	CT07	25368.00	cytoplasm	Biosynthesis of cofactors,
6-pyruvoyl	57 (27)	1- 9	Q8KEA7	CT07	17318.71	cytoplasm	Biosynthesis of cofactors,
Putative uncharacterized	64 (26)	1- 9	Q8KE76	CT08	27932.19	unknown	Hypothetical proteins:
Thioredoxin reductase	41 (25)	1- 6	Q8KE48	CT08	34048.91	cytoplasm	Energy metabolism: Electron
Adenylylsulfate reductase,	69 (26)	2- 18	Q8KE28	CT08	16104.6	unknown	Central intermediary
Sulfide dehydrogenase,	59 (26)	252-	Q8KDN1	CT10	46240.0	unknown	Central intermediary
CBS domain protein	44 (26)	1- 13	Q8KDK9	CT10	73274.4	cytoplasm	Unknown function: General
Putative uncharacterized	56 (26)	1- 12	O68985	CT10	22808.2	cytoplasm	Central intermediary
Fructose-bisphosphate	41 (26)	1- 16	Q8KDJ7	CT10	35124.1	cytoplasm	Energy metabolism:
Thiol:disulfide interchange	111 (25)	153-	Q8KDH8	CT10	18393.4	periplasm	Protein fate: Protein folding
30S ribosomal protein	36 (25)	1- 5	Q8KD88	CT11	15170.5	unknown	Protein synthesis: Ribosomal
	35 (26)	1- 8					
Peptide methionine	36 (31)	1- 12	Q8KCX9	CT12	32160.0	unknown	Protein fate: Protein
Aspartyl-tRNA synthetase	83 (26)	2- 14	Q8KCT7	CT13	68598.6	Cytoplasm	Protein synthesis: tRNA
ABC transporter, ATP-	29 (26)	1- 9	Q8KCR4		25808.2	Cytoplasmic	Transport and binding
Ribose-phosphate	76 (26)	1- 11	Q8KCQ2	CT13	35265.6	Cytoplasm	Purines, pyrimidines,
N-acetylmuramoyl-L-	32 (26)	1- 13	Q8KD30	CT12	29113.5	unknown	Cell envelope: Biosynthesis
50S ribosomal protein L33	38 (26)	2- 8	Q8KCP0	CT13	7039.2	Cytoplasm	Protein synthesis: Ribosomal
MoaA/nifB/pqqE family	41 (26)	1- 12	Q8KCI1		34524.8	Cytoplasm	Unknown function: General
	31 (25)	2- 12					
DegT/DnrJ/EryC1/StrS	74 (26)	1- 12	Q8KCE5	CT14	40255.1	Cytoplasm	Unknown function: General
Glyceraldehyde 3-	52 (26)	2- 13	Q8KCE2	CT14	35614.1	unknown	Energy metabolism:
Peptidase, M16 family	32 (26)	2- 25	Q8KC77		109748.8	Cytoplasm	Protein fate: Degradation of
Serine	70 (25)	1- 8	Q8KC36		47447.6	Cytoplasm	Amino acid biosynthesis:

Ferrous iron transport	82 (26)	1- 13	Q8KBP2	CT17	8423.9	unknown	Transport and binding
50S ribosomal protein L13	37 (26)	2- 19	Q8KBK4		16902.6	unknown	Protein synthesis: Ribosomal
Exodeoxyribonuclease 7	90 (26)	1- 13	Q8KBF9	CT18	44259.7	Cytoplasm	DNA metabolism:
Protein recA or	37 (25)	1- 18	O52393		37101.7	Cytoplasm	DNA metabolism: DNA
Dihydroorotate	89 (26)	1- 11	Q8KB48	CT19	27561.8	unknown	Purines, pyrimidines,
Heat shock protein,	37 (26)	2- 21	Q8KB28	CT19	16008.1	unknown	Protein fate: Protein folding
CRISPR-associated	73 (25)	1- 10	Q8KB24		23603.1	unknown	
CRISPR-associated	75 (26)	1- 14	Q8KB21	CT19	33099.4	cytoplasmic	Hypothetical proteins:
Rare lipoprotein A	53 (27)	100-	Q8KAS8	CT20	17736.4	unknown	Cell envelope: Other
Adenylosuccinate	55 (26)	1- 8	Q8KAK6	CT21	47773.4	Cytoplasm	Purines, pyrimidines,
Translation initiation factor	33 (26)	2- 23	Q8KAJ3	CT21	8269.6	Cytoplasm	Protein synthesis:
30S ribosomal protein S5	40 (24)	2- 7	Q8KAI8	CT21	18090.1	cytoplasm	Protein synthesis: Ribosomal
50S ribosomal protein L18	52 (26)	2- 9	Q8KAI7	CT21	12998.9	cytoplasm	Protein synthesis: Ribosomal
	31 (26)	11-					
50S ribosomal protein L23	63 (27)	1- 7	Q8KAH4	CT21	11766.7	unknown	Protein synthesis: Ribosomal
30S ribosomal protein S7	76 (26)	2- 11	P59060	CT21	17161.8	unknown	Protein synthesis: Ribosomal
Aspartyl/glutamyl-	35 (26)	145-	Q8KAE8	CT22	53091.5	cytoplasm	Protein synthesis: tRNA
Putative uncharacterized	28 (26)	2- 13	Q8KAE3	CT22	18274.8	unknown	Hypothetical proteins:
Phosphoglycerate kinase	61 (26)	1- 15	Q8KAE1	CT22	42289.5	Cytoplasm	Energy metabolism:
ATP synthase subunit	54 (25)	1- 34	Q8KAC9	CT22	50125.2	cytoplasm	Energy metabolism: ATP-
UDP-N-acetylmuramoyl-	42 (31)	3- 14	Q8KGD0		51348.2	Cytoplasm	Cell envelope: Biosynthesis
	40 (32)						
HIT family protein	40 (32)	63-	Q8KGB0	CT00	13922.93	unknown	Unknown function: general
	55 (36)	84-					
	54 (36)	86-					
	48 (36)	89-					
Cytochrome c-555	63 (37)	91-	Q8KG93	CT00	10891.6	unknown	Energy metabolism: Electron
	52 (38)						
	41 (37)	94-					
	94 (37)	95-					
Oxidoreductase, short-	42 (38)	112-	Q8KG22	CT01	28662.39	cytoplasm	Central intermediary
	41 (38)	168-					
Signal recognition	48 (37)	215-	Q8KFX0	CT02	34392.54	cytoplasm	Protein fate:Protein and
DnaK suppressor protein,	66 (31)	5- 23	Q8KFL4	CT03	17380.73	cytoplasm	DNA metabolism: DNA
Menaquinone	36 (31)	3- 18	Q8KF69		27306.8	unknown	Biosynthesis of cofactors,
Putative uncharacterized	42 (30)	28-	Q8KEK2	CT06	20559.5	unknown	
Membrane-bound lytic	45 (37)	478-	Q8KDR6	CT09	60082.3	cytoplasmic	Cell envelope: Biosynthesis
	50 (37)	97-					
Sulfur oxidation protein	40 (37)	98-	Q8KDM9	CT10	16077.7	unknown	Central intermediary
	68 (38)	99-					
Putative uncharacterized	44 (32)	32-	Q8KDI9		26286.1	unknown	Hypothetical proteins:
	77 (37)	92-					
UPF0251 protein CT1277	53 (37)	95-	Q8KCY0	CT12	21345.1	cytoplasm	Hypothetical proteins:
Multidrug resistance	47 (38)	256-	Q8KCX0		115534.3	cytoplasmic	Transport and binding
Putative uncharacterized	50 (32)	24-	Q8KCV1	CT13	19251.2	unknown	
	55 (32)	72-					
Putative uncharacterized	56 (32)	32-	Q8KCS8		17905.7	cytoplasmic	
	43 (30)	66-					
	57 (31)	67-					
OmpA family protein	90 (37)	150-	Q8KCR0	CT13	24334.4	outer membrane	Cell envelope: Other
	85 (38)	155-					
	47 (39)	224-					
Bacteriochlorophyll a	55 (37)	208-	Q46393	CT14	40294.5	unknown	Energy metabolism:
Cytochrome c or	52 (38)	144-	O07091	CT16	22714.7	cytoplasmic	Energy metabolism: Electron
	42 (37)	106-					
Phosphoenolpyruvate	47 (38)	225-	Q8KBZ2		103807.4	cytoplasm	Energy metabolism: TCA
	67 (32)	31-					

Putative uncharacterized	59 (33)	33-	Q8KBR2	CT17	12795.8	unknown	Hypothetical proteins:
	47 (36)	60-					
BchE or Magnesium-	62 (37)	508-	Q93SV2	CT19	63404.3	cytoplasm	Biosynthesis of cofactors,
Putative uncharacterized	63 (32)	4- 13	Q8KAP5	CT21	19996.4	Cytoplasm	Hypothetical proteins:
CysteinyI-tRNA	43 (38)	136-	Q8KAK2	CT21	54600.5	Cytoplasm	Protein synthesis: tRNA
water soluble-							
BirA bifunctional protein	90 (26)	1- 11	Q8KGB5	CT00	36097.49	cytoplasm	Biosynthesis of cofactors,
Magnesium chelatase,	95 (26)	1- 12	Q8KGA5	CT00	37467.29	cytoplasm	Biosynthesis of cofactors,
Enolase 2	68 (25)	379-	Q8KG25	CT01	46946.07	cytoplasm	Energy metabolism:
Acetyl-CoA carboxylase,	54 (26)	1- 10	Q8KG13	CT01	49180.13	cytoplasm	Fatty acid and phospholipid
ABC-type drug export	75 (27)	1- 14	Q8KFY6	CT01	29758.40	cytoplasm	Transport and binding
Thiamine-monophosphate	72 (27)	2- 17	Q8KFX1	CT01	38741.20	cytoplasm	Biosynthesis of cofactors,
Glutamate synthase,	32 (31)	52-	Q8KFP0		61656.13	cytoplasm	Amino acid biosynthesis:
Bifunctional purine	59 (27)	1- 8	Q8KFK6	CT03	56743.11	unknown	Purines, pyrimidines,
Leucyl/phenylalanyl-	74 (26)	1- 9	Q8KFI0	CT03	22135.71	unknown	Protein fate: Protein
Phosphoribosylaminoimid	46 (26)	50-	Q8KFF6	CT03	18407.37	unknown	Purines, pyrimidines,
Putative uncharacterized	39 (25)	1- 7	Q8KFC9	CT03	46427.93	cytoplasm	Hypothetical proteins:
	49 (25)						
Molybdenum transport	74 (26)	1- 12	Q8KF78	CT04	29773.27	cytoplasm	Transport and binding
Aconitate hydratase	54 (25)	2- 8	Q8KEY9	CT05	92162.84	cytoplasm	Energy metabolism: TCA
	63 (27)	450-					
	27 (26)	539-					
UDP-N-	56 (27)	1- 7	Q8KEX7	CT05	45759.83	cytoplasm	Cell envelope: Biosynthesis
Putative uncharacterized	32 (26)	34-	Q8KDY2	CT09	35048.7	unknown	
	71 (26)	151-					
	26 (25)	260-					
3-dehydroquinate	110 (27)	1- 17	Q8KD64	CT11	16053.4	unknown	Amino acid biosynthesis:
	121 (27)						
UDP-N-	32 (26)	2- 13	Q8KD27	CT12	58445.0	Cytoplasm	Cell envelope: Biosynthesis
YjeF family protein	67 (26)	1- 8	Q8KD16	CT12	55378.2	cytoplasm	Unknown function: General
Putative uncharacterized	84 (26)	1- 11	Q8KCW0	CT13	20990.0	Cytoplasm	Hypothetical proteins:
	55 (27)						
Putative metalloprotease	66 (27)	2- 11	Q8KCN1	CT13	15678.6	Cytoplasm	Hypothetical proteins:
Putative uncharacterized	43 (31)	2- 26	Q8KCE0		17102.9	unknown	
Thioredoxin peroxidase	57 (26)	127-	Q8KCD0	CT14	21764.8	cytoplasm	Cellular processes:
Putative uncharacterized	86 (26)	1- 13	Q8KC60		43894.2	unknown	Hypothetical proteins:
Glycine cleavage system	91 (27)	2- 11	Q8KC05	CT16	47739.3	unknown	Energy metabolism: Amino
Putative uncharacterized	62 (26)	1- 9	Q8KBW4	CT16	8827.5	unknown	
Putative uncharacterized	35 (27)	1- 13	Q8KBT1	CT17	17096.8	Cytoplasm	Hypothetical proteins:
Mg2+ transporter MgtE	63 (26)	1- 11	Q8KBD4	CT18	50986.4	cytoplasmic	Transport and binding
2-phosphoglycerate	125 (26)	1- 15	Q8KB35	CT19	46838.2	cytoplasm	Energy metabolism:
	107 (26)						
Response regulator	38 (26)	1- 17	Q8KAU3		13886.9	cytoplasm	Regulatory functions: Other
	27 (26)						
30S ribosomal protein	48 (26)	81-	Q8KAI1	CT21	10075.7	unknown	Protein synthesis: Ribosomal
30S ribosomal protein	50 (26)	2- 8	Q8KAH1	CT21	11820.9	unknown	Protein synthesis: Ribosomal
Putative uncharacterized	66 (27)	38-	Q8KAE7	CT22	20975.3	cytoplasmic	
Phosphoenolpyruvate	60 (26)	1- 13	Q8KAD1	CT22	69932.6	Cytoplasm	Energy metabolism:
Muramoyltetrapeptide	52 (27)	1- 10	Q8KAC7	CT22	35018.2	Cytoplasm	Cell envelope: Biosynthesis
Putative uncharacterized	77 (32)	6- 16	Q8KFW9		21521.9	unknown	Hypothetical proteins:
CobN protein, putative	38 (37)	1045-	Q8KFA9	CT04	141198.0	cytoplasm	Biosynthesis of cofactors,
	44 (38)	1046-					
Putative uncharacterized	41 (37)	157-	Q8KDP3		19134.2	cytoplasm	
Putative uncharacterized	84 (31)	5- 21	Q8KCA2	CT15	8232.4	Cytoplasm	
ATP-dependent Clp	39 (32)	26-	Q8KC73		24998.4	Cytoplasm	Protein fate: Degradation of
	43 (32)	27-					
	94 (32)						

	56 (32)	28-					
Peptidyl-prolyl cis-trans	106 (32)	15-	Q8KB93		15387.5	unknown	Protein fate: Protein folding
Putative uncharacterized	41 (32)	29-	Q8KAQ9	CT20	8973.9	unknown	Hypothetical proteins:
Putative uncharacterized	96 (26)	1- 31	Q8KDI3	CT10	6677.9	cytoplasm	Hypothetical proteins:
Probable peroxiredoxin	40 (26)	146-	Q8KBN8	CT17	24046.7	cytoplasm	Cellular processes:
water soluble-							
Heterodisulfide reductase,	58 (26)	2 -15	Q8KE24	CT08	43788.9	cytoplasmic	Central intermediary
Citrate lyase, subunit2	70 (27)	2- 10	Q8KDG2	CT10	65750.0	cytoplasmic	Energy metabolism: TCA
Chlorosome protein I	36 (26)	1- 23	O68988	CT13	25911.1	Chlorosome	Energy metabolism:
Partitioning protein, ParB	38 (26)	53-	Q8KBD7	CT18	32844.9	unknown	Cellular processes: Cell
Putative uncharacterized	77 (36)	31-	Q8KF52	CT04	74459.1	outer membrane	
TonB-dependent	86 (31)	48-	Q8KDV9	CT09	73909.4	outer membrane	Unknown function: General
Putative uncharacterized	44 (30)	28-	Q8KBI6	CT18	32338.7	unknown	
Polysulfide reductase,	66 (31)	32-	Q8KAC2	CT22	28963.3	unknown	Central intermediary