



ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΤΜΗΜΑ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ
ΕΡΓΑΣΤΗΡΙΟ ΓΕΝΕΤΙΚΗΣ**

**ΓΕΝΕΤΙΚΗ ΜΕΛΕΤΗ, ΕΞΕΛΙΚΤΙΚΕΣ ΣΧΕΣΕΙΣ ΚΑΙ
ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΑΠΑΚΕΤΥΛΑΣΩΝ**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΑΝΔΡΕΟΥ ΑΘΗΝΑ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: ΗΛΙΑΣ ΗΛΙΟΠΟΥΛΟΣ

ΑΘΗΝΑ 2019



AGRICULTURAL UNIVERSITY OF ATHENS

DEPARTMENT OF BIOTECHNOLOGY

LABORATORY OF GENETICS

**Genetic studies, Evolutionary relationships,
and Structural analysis of deacetylases**

PhD THESIS

Athena Andreou

Supervisor: Prof. Elias E. Eliopoulos

ATHENS 2019

“Γενετική μελέτη, Εξελικτικές σχέσεις και Δομική ανάλυση απακετυλασών ”
Διδακτορική Διατριβή
Υποβλήθηκε στο Τμήμα Βιοτεχνολογίας του Γεωπονικού Πανεπιστημίου Αθηνών
Αθήνα 2019
Αθηνά Ανδρέου

Τριμελής Συμβουλευτική Επιτροπή:
Ηλίας Ηλιόπουλος Καθηγητής, (Επιβλέπων)
Τμήμα Βιοτεχνολογίας, Γ.Π.Α

Ιωάννης Κλώνης Καθηγητής,
Τμήμα Βιοτεχνολογίας, Γ.Π.Α.

Κωνσταντίνος Μπεθάνης Επίκουρος Καθηγητής,
Τμήμα Βιοτεχνολογίας, Γ.Π.Α.

Εξεταστική Επιτροπή:
Ηλίας Ηλιόπουλος Καθηγητής,
Τμήμα Βιοτεχνολογίας, Γ.Π.Α
Επιβλέπων

Ιωάννης Κλώνης Καθηγητής,
Τμήμα Βιοτεχνολογίας, Γ.Π.Α
Μέλος τριμελούς συμβουλευτικής επιτροπής

Κωνσταντίνος Μπεθάνης Επίκουρος Καθηγητής,
Τμήμα Βιοτεχνολογίας, Γ.Π.Α.
Μέλος τριμελούς συμβουλευτικής επιτροπής

Γεώργιος Γουλιέλμος Καθηγητής,
Ιατρική Σχολή, Πανεπιστήμιο Κρήτης

Εμμανουήλ Ελευθέριος Σαρειδάκης Ερευνητής Β',
Ινστιτούτο Νανοεπιστήμης και Νανοτεχνολογίας, Ε.Κ.Ε.Φ.Ε. «Δημόκριτος»

Αθανάσιος Παπακυριακού Ερευνητής Β',
Ινστιτούτο Βιοεπιστημών και Εφαρμογών, Ε.Κ.Ε.Φ.Ε. «Δημόκριτος»

Αριάδνη Λουκία Χάγερ Επίκουρη Καθηγήτρια,
Τμήμα Επιστήμης Ζωικής Παραγωγής & Υδατοκαλλιεργειών Γ.Π.Α.

«Η έγκριση της παρούσης διδακτορικής διατριβής από το Τμήμα Βιοτεχνολογίας του Γεωπονικού Πανεπιστημίου Αθηνών δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.» [N.5343/1932, Άρθρο 202]

Στην Αθηνά της Αβησσυνίας

Περιεχόμενα

Ευχαριστίες	vii
Περίληψη	ix
Summary	xiii
Λέξεις κλειδιά	xvii
Συντμήσεις –Ξενόγλωσσοι όροι	xviii
1. Εισαγωγή	1
Α. Βιβλιογραφική ανασκόπηση	1
1.1 Διπλοποίηση γονιδίου (gene duplication)	1
1.1.1 Οικογένειες γονιδίων	1
1.1.2 Ο ρόλος της διπλοποίησης γονιδίων κατά την εξέλιξη	2
1.1.2.1 Η διπλοποίηση γονιδίου ως μηχανισμός γονιδιωματικής προσαρμογής σε ένα μεταβαλλόμενο περιβάλλον	3
1.1.3 Η εξέλιξη των πολυγονιδιακών οικογενειών	5
1.1.4 Λειτουργική διαφοροποίηση των διπλοποιημένων γονιδίων	7
1.1.4.1 Μηχανισμοί εξέλιξης	7
1.2 Η εξέλιξη των δομών των πρωτεϊνών και των δομικών συνόλων υπόκεινται σε λειτουργικούς περιορισμούς	8
1.3 Επίσταση	9
1.4. Οριζόντια μεταφορά του γονιδίου της ινωδονεκτίνης τύπου III	13
1.5 Τα βακτήρια <i>Bacillus anthracis</i> και <i>Bacillus cereus</i>	15
1.5.1 Η νόσος του άνθρακα	15
1.5.2 Μελέτη γονιδιωμάτων της ομάδας <i>B. cereus</i>	16
1.6 Λιποπρωτεΐνες (lipoproteins)	19
1.6.1 Λιποπρωτεΐνες των <i>Bacillus cereus</i> ATCC 14579 και <i>Bacillus anthracis</i> Ames	21
1.7 Κυτταρικό τοίχωμα κατά Gram-θετικών βακτηρίων	21
1.7.1 Η πεπτιδογλυκάνη	23
1.7.2 Συστατικά του κυτταρικού τοιχώματος των <i>Bacillus anthracis</i> και <i>B. cereus</i>	26
1.8 Τροποποίηση των αλυσίδων γλυκανών της πεπτιδογλυκάνης των βακτηρίων	31
1.8.1 Απακετυλίωση	32
1.9. Τα ένζυμα της βάσης δεδομένων CAZy	33

1.9.1 Η οικογένεια 4 υδατανθρακικών εστερασών της βάσης δεδομένων CAZy.....	33
1.9.2 Δομικές μελέτες των μελών της οικογένειας CE4	35
1.9.2.1 Δομικές μελέτες απακετυλασών πολυσακχαρίτη από την ομάδα του <i>B. cereus</i>	37
1.9.3 Μελέτες κυτταρικού εντοπισμού και χαρακτηρισμού απακετυλασών από την οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη του <i>B. anthracis</i> και του <i>B. cereus</i> .	39
1.10 Προοπτική	40
B. Σκοπός της διατριβής	43
2. Υλικά και Μέθοδοι	47
2.1. Υλικά.....	47
2.1.1. Εργαστηριακά όργανα	47
2.1.2. Αναλώσιμα	48
2.1.3. Αντιδραστήρια.....	49
2.1.4 Διαλύματα.....	50
2.1.5 Θρεπτικά υλικά.....	51
2.1.6. Εναρκτήρια μόρια (primers).....	51
2.1.7 Προσδέτες (ligands).....	52
2.1.8 Φορέας κλωνοποίησης (cloning vector) και βακτηριακό στέλεχος	52
2.1.9 Φορέας και στελέχη έκφρασης σε βακτηριακά συστήματα	54
2.2 Μέθοδοι.....	57
2.2.1 Γενετική Μηχανική	57
2.2.1.1 Απομόνωση πλασμιδιακού DNA.....	57
2.2.1.2 Καθαρισμός νουκλεϊκών οξέων (DNA purification).....	57
A. Με εκχύλιση με οργανικούς διαλύτες	57
B. Με συστηματοποιημένες χρωματογραφικές μεθόδους.....	58
2.2.1.3 Κατακρήμνιση νουκλεϊκών οξέων με αιθανόλη (DNA precipitation)	58
2.2.1.4 Ποσοτικός και ποιοτικός προσδιορισμός DNA.....	59
2.2.1.5 Ηλεκτροφόρηση DNA	59
2.2.1.6 Αντίδραση πέψης δίκλωνου DNA με περιοριστικές ενδονουκλεάσες.....	60
2.2.1.7 Αλυσιδωτή αντίδραση πολυμεράσης (PCR).....	60
2.2.1.8 Υποκλωνοποίηση δίκλωνου τμήματος DNA σε πλασμιδιακό φορέα	62
2.2.1.9 Αλληλούχιση DNA (DNA sequencing).....	63
2.2.2.Μέθοδοι Βιοχημείας.....	63
2.2.2.1 Έκφραση των πρωτεϊνών από μετασχηματισμένα βακτηριακά κύτταρα.....	63

A. Υγρές και στερεές καλλιέργειες βακτηριακών κυττάρων <i>Esherichia coli</i>	63
B. Μετασχηματισμός βακτηρίων με τη χημική μέθοδο του CaCl ₂	63
2.2.2.2 Καλλιέργειες έκφρασης σε βακτήρια	65
2.2.2.3 Λύση βακτηρίων και απομόνωση της ανασυνδυασμένης πρωτεΐνης.....	65
2.2.2.4 Χρωματογραφία συγγένειας (Affinity chromatography)	66
2.2.2.5 Χρωματογραφία μοριακής διήθησης (Size exclusion chromatography).....	67
2.2.2.6. Ποσοτική ανάλυση καθαρισμένων πρωτεϊνών.....	68
A. Φασματοσκοπικές μέθοδοι για την ποσοτικοποίηση πρωτεϊνών.....	68
B. Χρωματομετρική μέθοδος Bradford.....	68
2.2.2.7 Ηλεκτροφόρηση πρωτεϊνών σε πήκτωμα ακρυλαμιδίου υπό αποδιατακτικές συνθήκες	69
A. Χρώση με Coomassie:.....	69
2.2.2.8 Ανοσοαποτύπωση πρωτεϊνών	70
A. Ανοσοαποτύπωση πρωτεϊνών μετά από SDS PAGE (Western blotting):	70
B. Ανοσοαποτύπωση πρωτεϊνών υπό μη αποδιατακτικές συνθήκες	71
2.2.3 Κρυστάλλωση.....	71
2.2.3.1 Σάρωση συνθηκών κρυστάλλωσης.....	71
2.2.3.2 Βελτιστοποίηση αρχικών συνθηκών κρυστάλλωσης	72
2.2.3.3 Κρυστάλλωση συμπλόκων απακετυλάσης.....	73
Εμβάπτιση	73
Συγκρυστάλλωση	73
2.2.3.4 Συλλογή κρυστάλλων – Κρυοπροστασία – Αποθήκευση	74
2.3 Πρωτεϊνική Κρυσταλλογραφία	74
2.3.1 Συλλογή κρυσταλλογραφικών δεδομένων. Πειράματα περίθλασης ακτίνων-X.....	74
2.3.2 Επεξεργασία δεδομένων από περίθλαση ακτίνων-X.....	75
2.3.3 Επίλυση δομής-Βελτιστοποίηση	76
2.4 Βιοπληροφορική.....	77
2.4.1 Ανάκτηση ακολουθιών	77
2.4.2. Επιπρόσθετα προγράμματα ανάλυσης ακολουθιών	77
2.4.2.1 Στοιχισμός ακολουθιών.....	78
2.4.3 Φυλογενετική Ανάλυση - Μοριακή Εξέλιξη.....	78
2.5 Δομική Βιοπληροφορική.....	79
2.5.1 Πρόβλεψη και Μελέτη της αξιοπιστίας των τρισδιάστατων μοντέλων	80
2.5.2 Προσομοίωση της πρόσδεσης υποστρωμάτων.....	80

3. Αποτελέσματα.....	81
3.1. Έκφραση, απομόνωση και κρυστάλλωση της Ba3679	81
3.1.1. Έκφραση της Ba3679 από μετασχηματισμένα κύτταρα <i>E. coli</i>	81
3.1.2 Υποκλωνοποίηση της Ba3679 σε πλασμιδιακό φορέα έκφρασης	83
3.1.3 Καλλιέργειες μεσαίας κλίμακας και απομόνωση της πρωτεΐνης	87
3.1.4. Πειράματα κρυστάλλωσης	91
3.2 Μελέτη δομής της φερόμενης ως απακετυλάση πολυσακχαρίτη Ba0330 του <i>Bacillus anthracis</i>	92
3.2.1 Κρυστάλλωση – Βελτιστοποίηση κρυστάλλων	92
3.2.2 Πειράματα περίθλασης - Συλλογή κρυσταλλογραφικών δεδομένων.....	93
3.2.3 Επίλυση δομής – Βελτιστοποίηση.....	95
3.2.4 Περιγραφή δομής τριών διαστάσεων της Ba0330.....	97
3.3 Μελέτη δομής της φερόμενης ως απακετυλάση πολυσακχαρίτη Ba0331 του <i>Bacillus anthracis</i>	105
3.3.1 Κρυστάλλωση – Βελτιστοποίηση κρυστάλλων	105
3.3.2 Πειράματα περίθλασης - Συλλογή κρυσταλλογραφικών δεδομένων.....	106
3.3.3 Επίλυση δομής – Βελτιστοποίηση.....	107
3.3.4 Περιγραφή δομής τριών διαστάσεων της Ba0331.....	108
3.3.4.1 Ένα συντηρημένο καταλυτικό κέντρο	114
3.4 Μελέτη δομής της απακετυλάσης της πεπτιδογλυκάνης Bc1974 του <i>Bacillus cereus</i>	117
3.4.1 Κρυστάλλωση – Βελτιστοποίηση κρυστάλλων	117
3.4.1.1 Κρύσταλλοι συμπλόκων της Bc1974 με προσδέτες.....	117
3.4.2 Πειράματα περίθλασης - Συλλογή κρυσταλλογραφικών δεδομένων.....	118
3.4.3 Επίλυση δομής – Βελτιστοποίηση.....	119
3.4.4 Περιγραφή δομής τριών διαστάσεων της Bc1974.....	120
3.4.5 Περιγραφή δομής τριών διαστάσεων συμπλόκων της Bc1974.....	127
3.5 Φυλογενετική ανάλυση και ανάλυση αλληλουχίας στην οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη (PDA).....	135
3.5.1 Οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) στον <i>Bacillus anthracis</i> Ames.....	135
3.5.2 Οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) στον <i>Bacillus cereus</i> ATCC 14579.....	137
3.5.3 Φυλογενετική ανάλυση και ανάλυση νουκλεοτιδικής και αμινοξικής αλληλουχίας της δομικής ενότητας NodB	137

3.5.3.1 Φυλογενετική ανάλυση.....	139
3.5.3.2 Χαρακτηριστικά της αλληλουχίας.....	142
3.5.4 Φυλογενετική ανάλυση και ανάλυση αλληλουχίας της δομικής ενότητας Fn3.....	145
3.6 Τα γονίδια <i>ba0330</i> και <i>ba0331</i>	148
3.6.1 Η χρωμοσωμική οργάνωση των γονιδίων <i>ba0330</i> και <i>ba0331</i>	150
3.6.2 Σύγκριση νουκλεοτιδικής αλληλουχίας, αμινοξικής αλληλουχίας, τρισδιάστατης δομής	150
4. Συζήτηση.....	157
4.1. Χαρακτηριστικά της δομής και της αλληλουχίας CE4.....	157
4.1.1 Καθοριστικά δομικά μοτίβα. Μια συντηρημένη αμινοξική τριάδα στη θέση δέσμευσης του μετάλλου	157
4.1.2 Οργάνωση των δομικών ενοτήτων	159
4.1.3 Συγκρίσεις δομών PDA υψηλής διακριτικότητας	159
4.2 Εξελικτικές σχέσεις στην οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη του <i>Bacillus anthracis</i> και του <i>Bacillus cereus</i>	162
4.2.1 Φυλογενετική ανάλυση των απακετυλασών πολυσακχαρίτη με βάση τη νουκλεοτιδική και πρωτεϊνική αλληλουχία της καταλυτικής δομικής ενότητας NodB..	162
4.2.1.1 Ανίχνευση της διάδοσης της δομικής ενότητας ινωδονεκτίνης τύπου III.....	164
4.2.2 Εξέλιξη των θέσεων δέσμευσης των φερόμενων απακετυλασών πολυσακχαρίτη μετά από τα γεγονότα διπλοποίησης.	164
4.2.2.1 Όγκος της κοιλότητας δέσμευσης των απακετυλασών	166
4.2.2.2 Μεταλλάξεις στο ενεργό κέντρο καταστροφικές της καταλυτικής λειτουργίας	169
4.3 <i>ba0330</i> και <i>ba0331</i> δύο παράλογα γονίδια στην οικογένεια γονιδίων απακετυλασών πολυσακχαρίτη του <i>Bacillus anthracis</i>	170
A. Σύγκριση της τρισδιάστατης δομής. Διαφορές και ομοιότητες.	170
4.3.1 Η επιφάνεια διεπαφής των δομικών ενοτήτων NodB και Fn3	170
4.3.2 Μοριακή πλαστικότητα των δομών.....	172
4.3.2.1 Ανάλυση της κίνησης	173
4.3.3 Κρυσταλλική μοριακή διευθέτηση της δομικής ενότητας Fn3	175
4.3.4 Η NodB δομική ενότητα και η κοιλότητα δέσμευσης.....	177
4.3.4.1 Είναι δραματική η διαφορά της θέσης δέσμευσης μεταξύ της Ba0330 με την Ba0331.	180

B. Σύγκριση των πρωτοταγών αλληλουχιών των Ba0330, Ba0331 - Διαφορές και ομοιότητες	181
Γ. Η δομή φωτίζει την αλληλουχία.	182
4. 4 Ο ρόλος της δομικής ενότητας Fn3	183
4.5 Στοιχεία διαφοροποίησης λειτουργικών χαρακτηριστικών μέσω της δομικής ανάλυσης	185
4.6 Επίσταση	187
4.7 Γενετική μηχανική για την έκφραση των απακετυλασών	188
4.8 Προοπτική σχεδιασμού φαρμάκων με βάση τη δομή.	189
4.8.1 Αλληλεπιδράσεις υποστρώματος.	191
4.8.2.1 Δύο διακριτές διαμορφώσεις της καταλυτικής περιοχής MT4.....	197
4.8.3 Πρόταση φαρμακοφόρου.	200
4.9. Συμπερασματικά τα αποτελέσματα της παρούσης διατριβής.....	202
Βιβλιογραφία	205
Παράρτημα Α' -Πίνακας Ενώσεων –Προσδετών.....	233
Παράρτημα Β' –Ευρετήριο Εικόνων	237
Παράρτημα Γ' –Ευρετήριο Πινάκων.....	241
Παράρτημα Δ' -Δημοσιεύσεις / Ανακοινώσεις σε συνέδρια.....	243

Ευχαριστίες

Η παρούσα διδακτορική διατριβή πραγματοποιήθηκε στο Εργαστήριο Γενετικής του Γεωπονικού Πανεπιστημίου Αθηνών στα πλαίσια του προγράμματος διδακτορικών σπουδών του Τμήματος Βιοτεχνολογίας, υπό την επίβλεψη του Καθηγητή κ. Ηλία Ηλιόπουλου. Τον ευχαριστώ για την εμπιστοσύνη που μου έδειξε με την ανάθεση αυτής της διδακτορικής διατριβής, για τη συνεχή επιστημονική του παρουσία, τις καθοριστικές και εύστοχες παρεμβάσεις του στην κατεύθυνση της έρευνας, και ακόμη για τη φροντίδα του στην εκπαίδευση και κατάρτιση μου, και για τη χρηματοδότηση μέσω της οποίας έχω μπορέσει να συνεχίσω τις σπουδές μου αλλά και να διευρύνω τις επιστημονικές μου αναζητήσεις στην Ελλάδα αλλά και το εξωτερικό.

Θα ήθελα να ευχαριστήσω θερμά τα μέλη της τριμελούς συμβουλευτικής επιτροπής, τον Καθηγητή κ. Ιωάννη Κλώνη για τη στήριξη του και τον Επικ. Καθ. κ. Κωνσταντίνο Μπεθάνη γιατί ήταν πάντα εκεί με τη συμπαράστασή του και την άμεση ανταπόκρισή του να προσφέρει τη βοήθειά του.

Επίσης, ευχαριστώ θερμά όλα τα μέλη της επταμελούς εξεταστικής επιτροπής για την τιμή που μου έκαναν με την αποδοχή της πρότασης να συμμετάσχουν σε αυτήν, για το ενδιαφέρον τους και το χρόνο που μου αφιέρωσαν για την ανάγνωση της διατριβής. Τον Καθηγητή κ. Γεώργιο Γουλιέλμο θα ήθελα να τον ευχαριστήσω ξεχωριστά, για τις ευκαιρίες που μου έδωσε με τόσα ενδιαφέροντα ερευνητικά θέματα στη Γενετική που μου παρείχε και τη στήριξη και καθοδήγησή του. Την Επίκουρη Καθηγήτρια κα Αριάδνη Χάγερ για τις παρατηρήσεις της. Τον Κύριο Ερευνητή Δρ. Εμμανουήλ Σαρειδάκη για τις πολύτιμες επισημάνσεις του και τις συμβουλές του στην κρυστάλλωση πρωτεϊνών. Τον Κύριο Ερευνητή Δρ. Αθανάσιο Παπακυριακού γιατί απλόχερα μου έχει δώσει τις γνώσεις του, τη συνεργασία του αλλά και για τις καίριες παρατηρήσεις του επί του κειμένου της διατριβής.

Επιπλέον, θα ήθελα να ευχαριστήσω τον Ομ. Καθηγητή κ. Σωκράτη Τζάρτο για τη μοναδική ευκαιρία που μου έδωσε να διευρύνω τις γνώσεις μου με την επίσκεψή μου στο Εργαστήριο Μοριακής Νευροβιολογίας και Ανοσολογίας του Ελληνικού Ινστιτούτου Παστέρ καθώς και όλα τα μέλη της ομάδας του, για τη φιλική τους αντιμετώπιση, αλλά και γιατί μοιραστήκαμε πολλές όμορφες στιγμές. Θα ήθελα να ευχαριστήσω τον Δρ. Ελευθέριο Ζαρκάδα για τη βοήθειά του και για την εποικοδομητική, αλλά και ευχάριστη, συνεργασία μας και τους Δρ. Μάριο Ζουριδάκη, Δρ. Παρασκευή Ζησιμοπούλου, Δρ. Νικόλαο Τράκα και

Δρ. Νικόλαο Κουβάτσο όχι μόνο για τη φιλοξενία αλλά και για τις συζητήσεις επιστημονικού περιεχομένου και πειραματικών τεχνικών.

Το μεγαλύτερο και θερμότερο ευχαριστώ θα ήθελα να το εκφράσω στον Δρ. Πέτρο Γκιάστα, καθώς στάθηκε δίπλα μου σε όλη αυτήν την πορεία από την αρχή έως την ολοκλήρωση αυτής της διατριβής, και η συμβολή του ήταν καταλυτική. Του είμαι ευγνώμων. Η καθοδήγησή του, η ενεργή συμμετοχή του στα διάφορα στάδια της διατριβής, η ιδιαιτέρως σημαντική επιστημονική συμβολή του, οι γνώσεις που μου μετέδωσε, ήταν από τα στοιχεία εκείνα που οδήγησαν στην ολοκλήρωση της παρούσας διατριβής. Η ευκαιρία να συνεργαστώ μαζί του, ήταν καθοριστικής σημασίας για την κατάρτιση μου σε καινούργιους για μένα δρόμους στην δομική βιολογία, και τον ευχαριστώ για όλα όσα μου έμαθε σχετικά με την κρυσταλλογραφία μακρομορίων. Αλλά και γιατί συνάντησα έναν αληθινό φίλο.

Η παρούσα έρευνα κατά ένα μέρος της είναι μέρος ερευνητικού χρηματοδοτούμενου έργου «Θαλής» σε συνεργασία με το Εργαστήριο του Καθηγητή κ. Β. Μπουριώτη του Πανεπιστημίου Κρήτης, στο οποίο γινόταν και η μελέτη της ενζυμικής δράσης των απακετυλασών αλλά και η βιοχημική και βιολογική αποτίμηση των προτεινόμενων μέσω του δομικού προσδιορισμού, αναστολέων.

Τα κρυσταλλογραφικά δεδομένα συνελέγησαν με ακτινοβολία συγχρότρου σε αλληπάλληλες επισκέψεις στο σύγχροτρον Swiss Light Source (SLS) του Ινστιτούτου Paul Scherrer, στην Ελβετία στην τερματική γραμμή X06DA.

Ακόμη ευχαριστώ τον Δρ. Ηλία Χριστοφορίδη από το Εργαστήριο Φυσικής του Γ.Π.Α. για τη φιλία του, για τη συμπαράσταση και την πρόθυμη βοήθειά του.

Θα ήταν μεγάλη παράλειψη να μην ευχαριστήσω τον Ομ. Καθηγητή κ. Μιχαήλ Λουκά, ιδρυτή του τμήματος Βιοτεχνολογίας του Γ.Π.Α., για την εισαγωγή στον κόσμο της Γενετικής, η συμβολή του οποίου αποτελεί πάντοτε πηγή έμπνευσης και δημιουργικότητας.

Θα ήθελα να ευχαριστήσω την μητέρα μου, που με έχει στηρίξει και ανεχτεί στην παρουσία μου και στην απουσία μου. Τον πατέρα μου, που με μύησε από νήπιο στις θετικές επιστήμες και τον κόσμο της πειραματικής έρευνας.

Περίληψη

Οι απακετυλάσες πολυσακχαριτών (PDAs) είναι ένζυμα που ανήκουν στις υδατανθρακικές εστεράσες της οικογένειας 4 (CE4) και δρουν σε πολυσακχαρίτες όπως είναι η πεπτιδογλυκάνη, η χιτίνη, η χιτοζάνη, και η ακετυλοξυλάνη. Τα μέλη της CE4 είναι κυρίως μεταλλοεξαρτώμενες υδρολάσες, από ένα ιόν συνήθως Zn^{2+} ή Co^{2+} . Έχουν μια κοινή καταλυτική δομική ενότητα που ονομάζεται NodB και παρουσιάζουν τρισδιάστατη αναδίπλωση τύπου $(\alpha/\beta)_{6-8}$ με κοινό στοιχείο της πρωτεϊνικής τους αλληλουχίας τη συντήρηση σχηματισμού της θέσης δέσμευσης από πέντε μοτίβα (MT1-MT5). Σε αυτά περιλαμβάνονται τα αμινοξικά κατάλοιπα της τριάδας His–His–Asp τα οποία εμπλέκονται στην δέσμευση του μετάλλου. Ο προτεινόμενος μηχανισμός κατάλυσης θεωρείται τύπου οξέος/βάσης όπου τα καταλυτικά ασπαρτικό οξύ και ιστιδίνη του ενεργού κέντρου έχουν το ρόλο της βάσης και του οξέος, αντίστοιχα.

Η πεπτιδογλυκάνη του κυτταρικού τοιχώματος αναγνωρίζεται ως πρωταρχικός στόχος του έμφυτου ανοσοποιητικού συστήματος και συνήθως η αποσύνθεσή της οδηγεί σε βακτηριακή λύση. Η απακετυλίωση αποτελεί μια τροποποίηση μέσω της οποίας το βακτήριο διαφεύγει του ανοσοποιητικού συστήματος του ξενιστή, ο οποίος χρησιμοποιεί μία ποικιλία υδρολυτικών ενζύμων προκειμένου να διασπάσει τη στοιβάδα της πεπτιδογλυκάνης. Οι δομικές πληροφορίες για τη θέση δέσμευσης υποστρωμάτων που εξάγονται από τη μοριακή δομή των απακετυλασών καθώς και συμπλόκων τους με προσδέτες είναι πρωτίστης σημασίας, προκειμένου οι απακετυλάσες να χρησιμοποιηθούν ως μόρια στόχοι φαρμάκων.

Τα γονιδιώματα των *Bacillus cereus*, και *Bacillus anthracis*, δύο συγγενικών ειδών που ανήκουν στην ομάδα *B. cereus*, έχουν πολλαπλά γονίδια απακετυλάσης πολυσακχαρίτη. Στην παρούσα διατριβή, προσδιορίστηκαν με κρυσταλλογραφία ακτίνων-X και παρουσιάζονται οι ακόλουθες πρωτεϊνικές δομές: οι φερόμενες απακετυλάσες πολυσακχαρίτη, Ba0331 (διακριτικότητας 2,59Å) και Ba0330 (διακριτικότητας 1,48Å) από το βακτήριο *B. anthracis*, το οποίο είναι ο αιτιολογικός παράγοντας της ασθένειας του άνθρακα, καθώς και η απακετυλάση της N-ακετυλογλυκοζαμίνης της πεπτιδογλυκάνης, Bc1974 (διακριτικότητας 1,80Å), του *B. cereus*. Επίσης προσδιορίστηκαν και παρουσιάζονται οι κρυσταλλικές δομές συμπλόκων της Bc1974 με έξι προσδέτες, διακριτικότητας 1,45 Å, 2,44 Å, 2,45 Å, 2,73 Å, 2,80 Å και 3,05Å οι οποίες αποτελούν τις πρώτες δομές συμπλόκων απακετυλάσης της πεπτιδογλυκάνης με μικρά-μόρια προσδέτες.

Οι δομές των πρωτεϊνών Ba0331 και Ba0330 αποτελούνται από δυο ξεχωριστές δομικές ενότητες: μία στην καρβοξυτελική περιοχή που αναδιπλώνει ως μια τυπική υδατανθρακική εστεράση της οικογένειας 4 (CE4), επονομαζόμενη και ως καταλυτική NodB ομόλογη δομική ενότητα και μία κοντά στην αμινοτελική περιοχή, που συντίθεται από ένα μοτίβο (4 + 3) β-σάντουιτς, και εμφανίζει δομική ομοιότητα με δομικές ενότητες ινωδονεκτίνης τύπου III (fibronectin type III, Fn3). Οι δύο δομές διατηρούν στο σύνολό τους όμοια αρχιτεκτονική, παρά τις διαφοροποιήσεις α) σε συγκεκριμένες περιοχές της δομικής ενότητας Fn3, β) στις θέσεις δέσμευσης της δομικής τους ενότητας NodB όπου και παρουσιάζουν σημαντική διαφορά και γ) στη σχετική διευθέτηση των δύο διακριτών δομικών ενότητων, που όμως φανερώνει μια ευελιξία, κοινό χαρακτηριστικό των δύο πρωτεϊνών. Η Ba0331, εμφανίζει διακριτή θέση δέσμευσης, καθώς περιέχει ένα μοναδικό συνδυασμό πολικών και φορτισμένων καταλοίπων στην καταλυτική περιοχή NodB σχηματίζοντας ένα εκτεταμένο δίκτυο αλληλεπιδράσεων. Η παρατήρηση αυτή θέτει περαιτέρω ερωτήματα σχετικά με την απαιτούμενη εξειδίκευση του υποστρώματος, το οποίο ακόμα δεν έχει προσδιοριστεί, τόσο για την Ba0331, όσο και για την ομόλογή της Ba0330.

Η δομή της Bc1974 περιλαμβάνει μόνο την τυπική καταλυτική δομική ενότητα NodB. Αποκαλύπτονται δύο διακριτές διαμορφώσεις της περιοχής MT4 της Bc1974, υποδηλώνοντας εγγενή ευελιξία της συγκεκριμένης περιοχής, που μπορεί να παίζει ρόλο στην αναγνώριση του υποστρώματος, στην ικανότητα πρόσδεσής του και τελικά στην ενζυμική δραστηριότητα. Η δομική ανάλυση της Bc1974, χρησιμοποιήθηκε για να αποσαφηνιστεί ο ενζυμικός μηχανισμός απακετυλίωσης καθώς και τα προτεινόμενα φαρμακοφόρα χαρακτηριστικά θέτοντας τη βάση μιας προοπτικής ορθολογικού σχεδιασμού, με βάση τη δομή, νέων αντιβακτηριακών παραγόντων. Η ανάπτυξη εξειδικευμένων αναστολέων έναντι της Bc1974 θα έχει πιθανή εφαρμογή και στην ορθολογία της, απακετυλάση της πεπτιδογλυκάνης του *B. anthracis* Ba1977, η οποία είναι αποδεδειγμένα βάσει της βιβλιογραφίας λοιμογόνος παράγοντας του άνθρακα.

Η νουκλεοτιδική αλληλουχία του γονιδίου, η αμινοξική αλληλουχία της πρωτεΐνης και η τρισδιάστατη δομή της, καθώς και η βιολογική της λειτουργία είναι εγγενώς συνδεδεμένα μέσω της εξέλιξης και της γενετικής των πληθυσμών. Τα τελευταία χρόνια μελέτες έχουν περιγράψει τη λειτουργική διαφοροποίηση των γονιδίων και τη σημασία της στην εξέλιξη μέσω της σημαντικής διαδικασίας της διπλοποίησης γονιδίου.

Για τον καθορισμό φυλογενετικών σχέσεων στις οικογένειες γονιδίων απακετυλάσης πολυσακχαρίτη του *B. anthracis* Ames (12 ανοιχτά πλαίσια ανάγνωσης στο γονιδίωμά του) και του *B. cereus* ATCC14579 (11 ανοιχτά πλαίσια ανάγνωσης στο γονιδίωμά του)

χρησιμοποιήθηκαν οι αμινοξικές και οι νουκλεοτιδικές αλληλουχίες που απαρτίζουν χωριστά τις δομικές ενοότητες των PDAs. Η απεικόνιση των εξελικτικών γεγονότων στις δύο οικογένειες συνδυάστηκε και με την ανάλυση γονιδίων των PDAs από άλλα στελέχη του γένους *Bacillus* προκειμένου να ερευνηθούν εξελικτικές σχέσεις των γονιδίων PDAs. Διαδοχικά γεγονότα διπλοποίησης έχουν οδηγήσει σε ομάδες γονιδίων (ή υποοικογένειες) που εμφανίζουν μεταξύ τους μεγαλύτερη ομολογία σε σύγκριση με οποιοδήποτε άλλο γονίδιο εντός της κάθε οικογένειας.

Έγινε ανίχνευση της διάδοσης της δομικής ενότητας Fn3 μέσω της φυλογενετικής ανάλυσης, ενώ η δομική μελέτη οδήγησε σε επιπρόσθετες υποθέσεις για το ρόλο της, οι οποίες συνοψίζονται στα ακόλουθα: είτε στο να προσδώσει ιδιότητες προσκόλλησης/αλληλεπίδρασης βάση συγκεκριμένων διαφορών στην δομή των δύο πρωτεϊνών και σε συνδυασμό με το ιστορικό αυτής της δομικής ενότητας, είτε για να επεκτείνει την κοιλότητα δέσμευσης, ή/και για να εισάγει ευελιξία.

Αναλύθηκαν τα δύο παράλογα γονίδια *ba0330* και *ba0331* του *B. anthracis* μέσω συγκεκριμένων διαφορών στην αλληλουχία των γονιδίων, στην αλληλουχία των πρωτεϊνών, στην δομή των πρωτεϊνών, καθώς και μέσω της διάταξής τους στο γονιδίωμα, της ανίχνευσης της δράσης της επιλογής στην εξέλιξη της γονιδιακής διπλοποίησης, και του κυτταρικού εντοπισμού των αντίστοιχων πρωτεϊνών τους, (από τα διαθέσιμα πειράματα εντοπισμού στη βιβλιογραφία).

Οι τριςδιάστατες δομές απακετυλασών της οικογένειας των γονιδίων απακετυλάσης πολυσακχαρίτη χρησιμοποιήθηκαν για την εξελικτική μελέτη της θέσης δέσμευσης μετά από γεγονότα διπλοποίησης. Μέσα από την εργασία αυτή υπογραμμίζεται η σημασία της επίστασης στην εξέλιξη των πρωτεϊνών. Η δομική συγκριτική ανάλυση συνολικά επτά (τεσσάρων υπαρχουσών στη βιβλιογραφία και τριών από την παρούσα εργασία) προσδιορισμένων PDAs των *B. anthracis* και *B. cereus* υποστηρίζει τα συμπεράσματα της ανάλυσης αλληλουχίας και καταδεικνύει την ύπαρξη δομικών τροποποιήσεων στον σχηματισμό της θέσης δέσμευσης, παρά την υψηλή συντήρηση της πρωτεϊνικής αλληλουχίας της και του επίπεδου αναδίπλωσής της. Αυτό μπορεί να υποδεικνύει ότι τα τμήματα της αλληλουχίας που σχηματίζουν τη θέση δέσμευσης βρίσκονται υπό εξελικτική πίεση για να φιλοξενήσουν και να αναγνωρίσουν διακριτά κατάλοιπα πολυσακχαριτών ανάλογα με την κυτταρική θέση, τη χρήση ή το περιβάλλον.

Αν και η πρωτεϊνική δομή είναι πιο συντηρημένη από την αμινοξική ακολουθία, αυτή η μελέτη έδειξε ότι μικρές δομικές τροποποιήσεις στα πολλαπλά μοτίβα που συνθέτουν τη θέση δέσμευσης μπορεί να προκαλέσουν λειτουργική διαφοροποίηση. Η εξελικτική

λειτουργική διαφοροποίηση, και η ανάπτυξη της εξειδίκευσης της καταλυτικής δομικής ενότητας NodB με την έννοια των μικρών προσαρμογών της θέσης πρόσδεσης σε διαφορετικά αλλά ουσιαστικά παρόμοια υποστρώματα προφανώς είναι το αποτέλεσμα γεγονότων γονιδιακής διπλοποίησης.

Εν κατακλείδι, η σύγκριση των κοιλοτήτων δέσμησης του υποστρώματος μεταξύ των διαφόρων φερόμενων απακετυλασών πολυσακχαρίτη του *B. anthracis* και του *B. cereus* αποκάλυψε αξιοσημείωτη ποικιλότητα ως προς το μέγεθος και το σχήμα της κοιλότητας, καθώς επίσης και ως προς τον τύπο των αμινοξέων που τις συνθέτουν, εντοπίζοντας στοιχεία απόκλισης στα λειτουργικά χαρακτηριστικά με ταυτόχρονη διατήρηση των διακριτών τοπικών δομικών χαρακτηριστικών στην οικογένεια των PDAs. Παρότι απόψεις της λειτουργικής απόκλισης των απακετυλασών έχουν επισημανθεί σε αυτή τη διατριβή, η άμεση συσχέτιση των στην εξειδίκευση του υποστρώματος παραμένει ένα σημαντικό ερώτημα που μένει να απαντηθεί, ενώ και ο ρόλος των υπολοίπων φερόμενων απακετυλασών μένει να διελευκανθεί.

Summary

Polysaccharide deacetylases (PDAs), are members of the carbohydrate esterase family 4 (CE4), acting on polysaccharides, such as chitin and chitosans, peptidoglycans, acetylxylans, and β -1,6-GlcNAc polysaccharides. The CE4 family members share a conserved catalytic core termed the NodB homology domain which adopts an $(\alpha/\beta)_{6-8}$ three-dimensional fold, and they are mainly metal depended hydrolases, with Zn^{2+} and Co^{2+} as the most common metal cations. Common feature of their catalytic core is the conservation of five active site motifs in their protein sequence that include the His–His–Asp metal binding triad. The proposed catalytic mechanism is a type acid/base catalysis with the catalytic Asp and His residues as general base and general acid, respectively.

The cell wall peptidoglycan is recognized as a primary target of the innate immune system and usually its disintegration results in bacterial lysis. The N-deacetylation is an enzyme activated secondary modification of bacterial glycan, correlating with the bacterium resistance of the host lysozyme, and the effects on pathogen recognition via different host receptors. , by introducing additional positive charge into the cell wall that may increase the resistance of the bacterium to cationic antimicrobial peptides.

The structural information of the substrate binding site of deacetylases and their complexes with ligands is of prime importance for deacetylases to be used as drug target molecules.

Genomes of *Bacillus anthracis* and its relative *Bacillus cereus*, both belonging to the *B. cereus* group of species, contain multigenes of PDAs. In the present thesis, the 3D structures of the putative polysaccharide deacetylases Ba0331 (at 2.59Å resolution) and Ba0330 (at 1.48Å) from the *B. anthracis* bacterium, the causative agent of the anthrax disease, are presented. In addition the N- acetylglucosamine peptidoglycan deacetylase, Bc1974 (at 1.80Å) from *B. cereus* bacterium is presented. The structure of six Bc1974 complexes with distinct ligands have also been determined and presented at resolution of 1.45 Å, 2.44 Å, 2.45 Å, 2.73 Å, 2.80 Å and 3.05Å, which to the best of our knowledge, are the first crystal structures of a peptidoglycan GlcNAc deacetylase in complex with small-molecule ligands.

The structures of the Ba0331 and Ba0330 proteins revealed two distinct structural domains, a C-terminal domain with the typical fold of carbohydrate esterase family 4 (NodB catalytic domain) and an N-terminal domain composed of a two-layered (4 + 3) β -sandwich, with structural similarity to the fibronectin type III (Fn3) domain. The two structures retain a

similar architecture in general, despite the variations a) in specific regions of the Fn3 domain, b) in the binding sites of their NodB domains where they present a significant difference, and c) in the relative position of the two distinct domains, which reveals an interdomain structural flexibility that is a common feature of the two proteins. Ba0331, among the CE4 enzymes, exhibits a distinct binding site, as it contains a unique combination of polar and charged residues in the NodB catalytic domain, forming an extensive network of interactions.

The structure of Bc1974 retains the overall shape of the typical fold of the catalytic NodB domain. The structures of the protein-ligand complexes revealed two distinct conformations of the catalytic loop, known as motif 4, which suggests an intrinsic flexibility of this region that may play a role in the substrate recognition, binding and ultimately in enzymatic activity of Bc1974. Structural analysis was used to elucidate the enzyme deacetylation mechanism as well as to the pharmacophoric assertion, setting the basis of a rational structure based drug design of novel antibacterial agents. The development of specific inhibitors against Bc1974 will also be likely to be applied in its orthologous deacetylase of peptidoglycan of *B. anthracis* Ba1977, which is required for full virulence of *B. anthracis* as demonstrated in the literature.

The PDA gene nucleotide sequence, the amino acid sequence of the protein and its three-dimensional structure, as well as its biological function are inherently linked through evolution and population genetics. Recent studies have described the functional differentiation of genes and its significance in evolution through the important process of gene duplication.

Phylogenetic relationships in the polysaccharide deacetylase gene families of *B. anthracis* Ames (12 paralogous) and *B. cereus* ATCC14579 (11 paralogous) were investigated. Separate phylogenetic studies for the two distinct domains of the PDAs were performed for both amino acid and nucleotide sequences. In order to construct a composite phylogenetic tree these data have also been combined with the analysis of PDA genes from other strains of the genus *Bacillus*.

The spread of the Fn3 domain acquisition was traced by phylogenetic analysis. Also through the structural study additional assumptions about its role were proposed: a) to confer adhesion or recognition properties, based on specific differences in the domain structure between the two protein paralogues and based on the long history of this domain, b) to introduce length or/and c) to introduce mobility.

The two paralogous genes of *B. anthracis* *ba0330* and *ba0331* were analyzed via specific differences in the gene sequence, protein sequence, and 3D structure of the protein

molecules, as well as by their sequence arrangement in the genome, detection of selection in the evolution of gene duplication, and localization patterns available in the literature.

The three-dimensional structures of the deacetylases of the two polysaccharide deacetylase genes families were used for the evolutionary study of the binding site after gene duplication events. The structural comparative analysis of a total of seven (four existing in the literature and three of the present work) structurally determined PDAs of *B. anthracis* and *B. cereus*, supports the conclusions of the sequence analysis and demonstrates that, despite the high conservation on the protein sequence, there are distinct differences in the binding site formation, the folding level and the active site assembly. This may indicate that, the binding site forming sequence fragments are under functionally-driven evolutionary pressure in order to accommodate and recognize distinct polysaccharide residues according to cell location, use or environment. This is probably an indication of the distinct roles of PDAs at a cellular level, however it is still an open subject of biochemical verification.

Although protein structure is more conserved than the sequence, this study showed that small structural modifications in a multi motif assembled binding domain may cause functional diversification. Evolutionary functional diversification, as in the development of substrate specificity in the PDA NodB binding domain, in the form of small adaptations, apparently has occurred through gene duplication events. And the essential role of epistasis appears in protein evolution.

Overall, the comparison of the binding cavities among the putative polysaccharide deacetylases of *B. anthracis* and *B. cereus* revealed remarkable variation in cavity size and shape, as well as in the amino acid composition of the cavities, indicating elements of divergence in functional characteristics with concurrent conservation of distinct local structural features in the PDA family. Although aspects of the functional diversity have been highlighted in this thesis, direct correlation with substrate specificity of deacetylases remains an important question to be answered, while the role of the rest putative deacetylases remains to be elucidated.

Επιστημονική Περιοχή

Δομική Βιολογία Structural Biology

Εξέλιξη Evolution

Γενετική Genetics

Δομική εξέλιξη Structural Evolution

Λέξεις κλειδιά

Οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη, *Bacillus anthracis*, *Bacillus cereus*, διπλοποίηση γονιδίου, παράλογα γονίδια, λειτουργική απόκλιση, εξέλιξη πολυγονιδιακής οικογένειας, πρωτεϊνική δομή, γενετική πληθυσμών, Εξελικτική γενετική, Πίεση επιλογής, δομική εξέλιξη, πολυγονιδιακή οικογένεια, Επίσταση, CE4, Κρυσταλλική δομή ακτίνων-X, Κρυσταλλογραφία ακτίνων-X, Κρυσταλλική δομή Bc1974, Ba0330 κρυσταλλική δομή, Ba0331 κρυσταλλική δομή, φυλογένεια, απακετυλάση πεπτιδογλυκάνης, απακετυλάση πολυσακχαρίτη

Λέξεις κλειδιά στην αγγλική γλώσσα (Keywords)

Polysaccharide Deacetylase gene family; *Bacillus anthracis*; *Bacillus cereus*; gene duplication; paralogues; functional divergence; family evolution; protein structure; population genetic; evolutionary genetic; structural evolution; sselection pressure; multigene family; epistasis; CE4; X-ray crystal structure; X-ray crystallography; Bc1974 crystal structure; Ba0330 crystal structure; Ba0331 crystal structure; phylogeny; peptidoglycan deacetylase; polysaccharide deacetylase

Συντμήσεις – Ξενόγλωσσοι όροι

6xHIS: Hexahistidine tag region, Επίτοπος 6 ιστιδινών

ACT: acetate, μόριο οξικού

AM: asymmetric unit cell AUC, ασύμμετρη μονάδα

APS: Ammonium persulfate, Υπεροξθειϊκό αμμώνιο

Bis: Bis-acrylamide, Δις-ακρυλαμίδιο

C-τελικό άκρο: Carboxyl terminus, C- terminus, Καρβοξυτελικό άκρο

CAZy : Carbohydrate-Active enZymes Database

CDA: απακετυλάση χιτοολιγοσακχαριτών

CE4: Carbohydrate Esterase family 4, Η οικογένεια 4 των υδατανθρακικών εστερασών

CEs : Carbohydrate Esterases, Εστεράσες υδατανθράκων

CS: cleavage site, θέση αποκοπής

DMSO: dimethylsulfoxide, διμεθυλοσουλφοξείδιο

EDTA: EthyleneDiamineTetraAcetic acid, αιθυλενο-διαμινο-τετραοξικό οξύ

Fn3: Fibronectin type 3, ινοδωνεκτίνη τύπου III

GlcNAc: N-acetylglucosamine, N-ακετυλογλυκοζαμίνη

GlcNAc PDA: απακετυλάση Nακετυλογλυκοζαμίνης της πεπτιδογλυκάνης

HDAC8: histone deacetylase 8, απακετυλάση ιστόνης

Lpp: Lipoprotein, λιποπρωτεΐνη

MBP: Maltose Binding Protein, πρωτεΐνη δέσμευσης της μαλτόζης

MMP: matrix metalloproteinase

MR: Molecular Replacement, Μοριακή Αντικατάσταση

MT: motif, μοτίβο

MurNAc: N-acetylmuramic acid, N-ακετυλομουραμικό οξύ

MurNAc PDA: απακετυλάση N-ακετυλομουραμικού οξέος της πεπτιδογλυκάνης

N-τελικό άκρο: Amino terminus, N-terminus, Άμινο τελικό άκρο

OD600: Optical Density, οπτική πυκνότητα στα 600 nm

ORF: open reading frame, ανοιχτό πλαίσιο ανάγνωσης

PAGE: Polyacrylamide Gel Electrophoresis, Ηλεκτροφόρηση πηκτώματος πολυακρυλαμιδίου

PCR: PolymeraseChainReaction, Αλυσιδωτή Αντίδραση Πολυμεράσης

PDA: polysaccharide deacetylase, απακετυλάση πολυσακχαρίτη

PDB: Protein Data Bank, βάση δεδομένων Πρωτεϊνών

PDB ID: κωδικός καταχώρησης στην βάση δεδομένων πρωτεϊνών

PEG: PolyEthyleneGlycol, πολυαιθυλενογλυκόλη

PG: peptidoglycan, πεπτιδογλυκάνη

pI: Isoelectric Point, ισοηλεκτρικό σημείο

r.m.s.: root mean square, μέση τετραγωνική ρίζα

r.m.s.d.: r.m.s. deviation, απόκλιση r.m.s.

SDS: SodiumDodecylSulphate, δωδεκυλοθειϊκό νάτριο

TEMED: Tetramethylethylenediamine, Τετραμεθυλαιθυλενοδιαμίνη

Tris: Tris(hydroxymethyl)aminomethane, τρις(υδροξυμεθυλο)-αμινο-μεθάνιο

Συμβολισμός των αμινοξέων με τον κωδικό του ενός και των τριών γραμμάτων:

Αμινοξύ	AminoAcid	Κωδικός τριών Γραμμάτων	Κωδικός ενός Γράμματος
Αλανίνη	Alanine	Ala	A
Αργινίνη	Arginine	Arg	R
Ασπαραγίνη	Asparagine	Asn	N
Ασπार्टικό οξύ	Aspartic acid	Asp	D
Βαλίνη	Valine	Val	V
Γλουταμικό οξύ	Glutamicacid	Glu	E
Γλουταμίνη	Glutamine	Gln	Q
Γλυκίνη	Glycine	Gly	G
Θρεονίνη	Threonine	Thr	T
Ισολευκίνη	Isoleucine	Ile	I
Ιστιδίνη	Histidine	His	H
Κυστεΐνη	Cysteine	Cys	C
Λευκίνη	Leucine	Leu	L
Λυσίνη	Lysine	Lys	K
Μεθειονίνη	Methionine	Met	M
Προλίνη	Proline	Pro	P
Σερίνη	Serine	Ser	S
Θρυπτοφάνη	Tryptophan	Trp	W
Τυροσίνη	Tyrosine	Tyr	Y
Φαινυλαλανίνη	Phenylalanine	Phe	F

Ο Γενετικός Κώδικας που χρησιμοποιείται στη παρούσα διατριβή

TTT F Phe	TCT S Ser	TAT Y Tyr	TGT C Cys
TTC F Phe	TCC S Ser	TAC Y Tyr	TGC C Cys
TTA L Leu	TCA S Ser	TAA * Ter	TGA * Ter
TTG L Leu	TCG S Ser	TAG * Ter	TGG W Trp
CTT L Leu	CCT P Pro	CAT H His	CGT R Arg
CTC L Leu	CCC P Pro	CAC H His	CGC R Arg
CTA L Leu	CCA P Pro	CAA Q Gln	CGA R Arg
CTG L Leu	CCG P Pro	CAG Q Gln	CGG R Arg
ATT I Ile	ACT T Thr	AAT N Asn	AGT S Ser
ATC I Ile	ACC T Thr	AAC N Asn	AGC S Ser
ATA I Ile	ACA T Thr	AAA K Lys	AGA R Arg
ATG M Met	ACG T Thr	AAG K Lys	AGG R Arg
GTT V Val	GCT A Ala	GAT D Asp	GGT G Gly
GTC V Val	GCC A Ala	GAC D Asp	GGC G Gly
GTA V Val	GCA A Ala	GAA E Glu	GGA G Gly
GTG V Val	GCG A Ala	GAG E Glu	GGG G Gly

Ο γενετικός κώδικας που χρησιμοποιείται για τα βακτήρια. Ο εικονιζόμενος πίνακας ταυτίζεται με τον βασικό γενετικό κώδικα (standard code) (πηγή: *-The Bacterial, Archaeal and Plant Plastid Code, transl_ table=11-*, <https://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?chapter=tgencodes#SG1>)

1. Εισαγωγή

A. Βιβλιογραφική ανασκόπηση

1.1 Διπλοποίηση γονιδίου (gene duplication)

Το 1936, ο Calvin Bridges ανέφερε μία από τις πρώτες παρατηρήσεις διπλοποίησης* μιας χρωμοσωματικής περιοχής όταν μελέτησε την κυρίαρχη φυλοσύνδετη μεταλλαγή Bar που επηρεάζει τον αριθμό των ομματιδίων των συνθέτων οφθαλμών της δροσόφιλας *Drosophila melanogaster* (Bridges 1936). Κατόπιν ακολούθησαν εργασίες στις οποίες διατυπώθηκαν υποθέσεις για το δυνητικό ρόλο της διπλοποίησης γονιδίων στην εξέλιξη (Stephens 1951, Ohno 1967, Nei 1969) και το 1970 δημοσιεύθηκε η θεμελιώδης μονογραφία του Susumu Ohno «*Evolution by Gene Duplication*», «*Εξέλιξη με Γονιδιακή Διπλοποίηση*». Έκτοτε η διπλοποίηση γονιδίου έχει αναδειχθεί ως μια σημαντική διαδικασία για τη λειτουργική διαφοροποίηση των γονιδίων και εν γένει έχει πρωτεύοντα ρόλο στην εξέλιξη. Στα τέλη της δεκαετίας του 1990 όμως η σημασία της διπλοποίησης των γονιδίων αποδείχθηκε σαφώς, όταν άρχισαν να προσδιορίζονται και να αναλύονται πλήθος γονιδιωματικών ακολουθιών. Σήμερα βρισκόμαστε στην εποχή της πλήρους αλληλούχισης γονιδιωμάτων και σημειώνεται πρόοδος στην αποκάλυψη των μηχανισμών με τους οποίους τα διπλοποιημένα γονίδια αποκλίνουν σε λειτουργία αλλά και στην διατύπωση απόψεων για τη συμβολή τους στην εξέλιξη.

1.1.1 Οικογένειες γονιδίων

Η σύναψη των ομολόγων στην πρόφαση I είναι μια ακριβής διαδικασία (γονίδιο με γονίδιο), μερικές φορές (και ιδιαίτερα όταν υπάρχουν επαναλαμβανόμενες αλληλουχίες) η σύναψη είναι ασύμμετρη. Ο άνισος επιχιασμός (crossing over) έχει αποδειχθεί ότι αποτελεί το βασικό μηχανισμό διπλοποίησης γονιδίου και δημιουργίας οικογενειών γονιδίων (gene families) δηλαδή ομάδων συγγενικών γονιδίων τα προϊόντα των οποίων επιτελούν την ίδια ή παρόμοιες λειτουργίες και προέρχονται από ένα αρχέγονο γονίδιο. Επίσης ευθύνεται για τη δημιουργία εσωτερικών επαναλήψεων σε κωδικές αλληλουχίες γονιδίων (Λουκάς 2010).

*Παρόλο που σε πλήθος ελληνικών εκδόσεων-βιβλίων ο όρος αποδίδεται ως διπλασιασμός στο παρόν κείμενο θα ακολουθηθεί ο όρος διπλοποίηση όπως έχει αποδοθεί από τον Καθηγητή της Γενετικής κ. Μ. Λουκά.

1.1.2 Ο ρόλος της διπλοποίησης γονιδίων κατά την εξέλιξη

Ο Ohno (1970) κατέδειξε τη διπλοποίηση ως βασικό μηχανισμό δημιουργίας νέων γονιδίων κατά τη διάρκεια της εξέλιξης. Αν π.χ. τα προϊόντα μερικών γονιδίων είναι απαραίτητα για την επιβίωση ενός οργανισμού, τότε τα γονίδια αυτά δε συσσωρεύουν μεταλλάξεις οι οποίες θα μεταβάλλουν την πρωτογενή τους λειτουργία. Αν όμως ένα τέτοιο γονίδιο διπλοποιηθεί τότε το επιπλέον αντίγραφο μπορεί να υποστεί σημαντικές αλλαγές με τη συσσώρευση μεταλλαγών αφού η βασική λειτουργία του γονιδίου καλύπτεται από το άλλο χρωματοσωματικό αντίγραφο του. Έτσι είναι δυνατή μια διαφοροποίηση της λειτουργίας των διπλοποιημένων γονιδίων. Η νέα αυτή λειτουργία μπορεί να προσδώσει ένα εξελικτικό πλεονέκτημα στον οργανισμό με συνέπεια την αύξηση του συντελεστή προσαρμογής του. Οι διπλοποιήσεις μπορούν να συμβάλλουν όχι μόνο στην αύξηση της παραγωγής ενός προϊόντος από λειτουργικά πλεονάζοντα γονίδια, αλλά και στην αύξηση του γονιδιώματος και του πληροφοριακού του περιεχομένου (Λουκάς 2010). Το ίδιο το γεγονός διπλοποίησης μπορεί να αλλάξει ριζικά τον τρόπο με τον οποίο ρυθμίζεται η έκφραση του γονιδίου (Breen *et al.* 2012).

Ήδη από το 1976 ο Jensen υποστήριξε ότι τα ένζυμα που παρουσιάζουν ειδικότητα υποστρώματος μπορεί να έχουν προκύψει με γονιδιακή διπλοποίηση και απόκλιση προγονικών ενζύμων που παρουσίαζαν μικρότερη ειδικότητα υποστρώματος. Το 1983, ο Li αναφέρει ότι η διπλοποίηση γονιδίου μπορεί να δώσει πρωτεΐνες με καινούργια λειτουργία. Τα διπλοποιημένα γονίδια είναι δυνατόν να αποκλίνουν στην αλληλουχία και στην λειτουργία κάτω από την επίδραση της μεταλλαγής, της γενετικής παρέκκλισης και της φυσικής επιλογής (Futuyama 1991).

Διπλοποιημένα γονίδια εντός ενός είδους ονομάζονται παράλογα γονίδια (ή παράτροπα, *paralogous genes*) ενώ τα αντίστοιχα γονίδια σε διαφορετικά είδη ονομάζονται ορθόλογα (ή ορθότροπα, *orthologous*). Εξυπακούεται ότι η φυλογένεια των γονιδίων μπορεί να διαφέρει από τη φυλογένεια των ειδών (Page & Holmes 1998, Ροδάκης 2001).

1.1.2.1 Η διπλοποίηση γονιδίου ως μηχανισμός γονιδιωματικής προσαρμογής σε ένα μεταβαλλόμενο περιβάλλον

Η πληθώρα δεδομένων αλληλουχίας και θεωρητικές εκτιμήσεις οδηγούν στο συμπέρασμα ότι είναι πολύ ευκολότερο να δημιουργηθούν νέες λειτουργίες από προϋπάρχουσες (Li 1997, Hughes 1999, Lynch & Conery 2000, Prince & Pickett 2002, Conant & Wolfe 2008). Καθώς έγινε φανερό ότι γονιδιακά αντίγραφα ποικίλου βαθμού απόκλισης είναι παρόντα στα γονιδιώματα, το κύριο ερώτημα που προέκυψε ήταν πώς μπορούν να διατηρηθούν πλήρως πλεονάζοντα γονιδιακά αντίγραφα στο γονιδίωμα και να εξελιχθούν σε μια νέα λειτουργία χωρίς να εξαλειφθούν με συσσώρευση μεταλλάξεων. Σύντομα, δόθηκαν παραδείγματα προσαρμοστικότητας μέσω διπλοποίησης γονιδίου (Anderson & Roth 1977 και 1979, Koch *et al.* 1981, Stark & Wahl 1984, Sonti & Roth 1989, Tabashnik 1990, Taylor & Feyereisen 1996, Widholm *et al.* 2001, Francino 2005, Moore & Purugganan 2005, Craven & Neidle 2007, Hastings 2007). Ταυτόχρονα, δημοσιεύτηκαν δεδομένα από τον κλάδο της μικροβιολογίας που υποδήλωναν προσαρμοστικές επιπτώσεις των γονιδιακών διπλοποιήσεων υπό ορισμένες περιβαλλοντικές συνθήκες (Velkov 1982, Romero & Palacios 1997, Kondrashov *et al.* 2002)

Η διπλοποίηση γονιδίων ως τρόπος προσαρμογής σε διάφορες περιβαλλοντικές συνθήκες δεν είναι ούτε σπάνιος μηχανισμός, ούτε και αποκλειστικός, όπως έχει περιγραφεί. Ίσως να γίνονται αρκετές γενικεύσεις σχετικά με το πότε μπορεί να αναμένεται ότι μια προσαρμοστική διπλοποίηση θα διαδραματίσει ρόλο σε διάφορες προσαρμογές στη φύση (Kondrashov 2012). Πρώτον, σε περιπτώσεις όπου τα πρωτεϊνικά προϊόντα των γονιδίων αλληλεπιδρούν φυσικά με μόρια που σχετίζονται με ένα μεταβλητό περιβάλλον όπως οι τοξίνες ή τα θρεπτικά συστατικά. Δεύτερον, όταν οι πρωτεΐνες που κωδικοποιούνται από τα διπλοποιημένα γονίδια λειτουργούν στην ίδια οδό με εκείνες που φυσικά αλληλεπιδρούν με τέτοια μόρια. Τέλος, η προσαρμοστική διπλοποίηση φαίνεται να περιλαμβάνει εκείνα τα γονίδια, των οποίων το προϊόν πρέπει να παραχθεί γρήγορα ή συνεχώς σε υψηλό επίπεδο, όπως οι αντιψυκτικές ή οι αποθηκευτικές πρωτεΐνες. Η διπλοποίηση γονιδίων που εμπλέκονται στην ανθεκτικότητα των παθογόνων (Hanada *et al.* 2008, Cornman & Willis 2008, Gong *et al.* 2010, Schulenburg & Boehnisch 2008, Neiman *et al.* 2009) και στην παθογένεια (Jiang *et al.* 2010, Powell *et al.* 2008) υποδεικνύει ότι η διπλοποίηση γονιδίων μπορεί να έχει κάποιο ρόλο στην συνεξέλιξη του ξενιστή και του παθογόνου ή των συμβιωτών.

Η διπλοποίηση γονιδίων ή γονιδιωμάτων που σχετίζονται με την ανάπτυξη μπορεί να προσφέρει ένα προσαρμοστικό πλεονέκτημα σε ένα νέο περιβάλλον με την ανατομική τροποποίηση (Shigenobu *et al.* 2010). Σε φυτά η προσαρμογή στην καταπόνηση από την υψηλή περιεκτικότητα σε αλάτι έχει δειχθεί μέσω της πολυπλοειδίας (polyploidy) η οποία έχει συνδεθεί με την αντοχή στις υψηλές συγκεντρώσεις αλάτων (Ceccarelli *et al.* 2006, Gerstein *et al.* 2006, Saleh *et al.* 2008), γεγονός που υποδηλώνει ότι η πολυπλοειδία μπορεί να είναι μια γενική φυσιολογική προσαρμοστική απόκριση στο ωσμωτικό στρες (Dhar *et al.* 2011). Σε βακτήρια έχει παρατηρηθεί η διπλοποίηση ως προσαρμοστική απάντηση στη θεραπεία με αντιβιοτικά (Kondrashov & Kondrashov 2006, Craven & Neidle 2007, Sandegren & Andersson 2009). Ομοίως, αναγνωρίζεται ευρέως ότι διπλοποιήσεις εμφανίζονται σε καρκινικούς όγκους ως απόκριση σε διάφορες θεραπείες φαρμάκων (Kondrashov & Kondrashov 2006). Ενώ υπάρχουν δεδομένα σχετικά με διπλοποιήσεις γονιδίων ως απάντηση σε διάφορες θεραπείες της λείσμανιασης (Segovia 1994) και της ελονοσίας (Templeton 2009).

Αρχικώς μπορεί να φαίνεται ότι οι βραχυπρόθεσμες επιπτώσεις των γονιδιακών διπλοποιήσεων δεν είναι σημαντικές για τη μελλοντική εξέλιξη της γονιδιακής λειτουργίας. Ωστόσο, τουλάχιστον θεωρητικά, όταν ένα βραχυπρόθεσμο μειονέκτημα αντιβαίνει στα μακροπρόθεσμα οφέλη καθώς και όταν ένα βραχυπρόθεσμο πλεονέκτημα οδηγεί σε ενδεχόμενη εξαφάνιση, το εξελικτικό αποτέλεσμα συνήθως καθορίζεται από το βραχυπρόθεσμο αποτέλεσμα (Maynard Smith 1978, Webb 2003). Έτσι, οι βραχυπρόθεσμες επιδράσεις των γονιδιακών διπλοποιήσεων πρέπει να εξετάζονται με σκοπό την κατανόηση της μακροπρόθεσμης εξέλιξης των νέων γονιδιακών λειτουργιών. Υπάρχουν δύο σημαντικές υποθέσεις που σχετίζονται με τα αρχικά στάδια της ζωής μιας γονιδιακής διπλοποίησης. Η παραδοσιακή και ευρύτερα αποδεκτή άποψη υποστηρίζει ότι οι γονιδιακές διπλοποιήσεις είναι λειτουργικά περιττές και σταθεροποιούνται με γενετική παρέκκλιση (genetic drift). Η εναλλακτική υπόθεση υποστηρίζει ότι οι γονιδιακές διπλοποιήσεις αυξάνουν τη δοσολογία του γονιδίου και μπορούν να σταθεροποιηθούν με θετική επιλογή (Kondrashov 2012).

1.1.3 Η εξέλιξη των πολυγονιδιακών οικογενειών

Ο πιο προφανής τρόπος με τον οποίο μπορεί να αλλάξει ο αριθμός των γονιδίων μεταξύ των ειδών είναι μέσω της διπλοποίησης γονιδίου όπου δημιουργούνται απευθείας αντίγραφα των γονιδίων (ή μερών τους). Όπως ήδη αναφέρθηκε, αυτή αποτελεί μια σημαντική διαδικασία στην εξέλιξη των πολυγονιδιακών οικογενειών καθώς είναι το πρώτο βήμα για τη δημιουργία γονιδίων με νέες λειτουργίες και για την αύξηση του αριθμού αντιγράφων της οικογένειας (και του μεγέθους του γονιδιώματος), και έχει θεμελιώδη σημασία ως πηγή γενετικής ποικιλότητας.

Η διπλοποίηση γονιδίου μπορεί να συμβεί με διάφορους τρόπους. Ένας μηχανισμός όπως αναφέρθηκε είναι ο άνισος επιχiasμός. Αυτός είναι ιδιαίτερα σημαντικός στην αύξηση του αριθμού των αντιγράφων στις πολυγονιδιακές οικογένειες. Επίσης μπορεί να συμβεί αν ολόκληρο το γονιδίωμα διπλοποιηθεί όπως στον *Saccharomyces cerevisiae*. Ένας τρόπος που έχει συμβεί αυτό είναι με την πολυπλοειδία παρόλο που είναι κοινό στα φυτά παρά στα ζώα καθώς μπορούν να πολλαπλασιαστούν χωρίς την ανάγκη γαμετών, έχει παρουσιαστεί στα αμφίβια όπως έχει περιγραφεί στον *Xenopus laevis*. Η πολυπλοειδία είναι σημαντική στην ειδογένεση. Παράδειγμα η οικογένεια των σαλμονίδων που έχουν διπλάσια ποσότητα DNA από ότι συγγενικοί τους ιχθύες και που θεωρείται ότι προέρχεται μέσω πολυπλοειδίας. Άλλος τρόπος διπλοποίησης γονιδίου μπορεί να συμβεί μέσω μετάθεσης (transposition), η οποία λαμβάνει χώρα με έναν αριθμό διαφόρων μηχανισμών, με κύριο αντιπρόσωπο τα μεταθετά στοιχεία τα οποία αποτελούν κύριο συστατικό των ευκαρυωτικών γονιδιωμάτων (Page & Holmes 1998).

Όταν ένα γονίδιο δημιουργηθεί με διπλοποίηση ένας αριθμός συμβάντων θα μπορούσε να συμβεί σε αυτό. Θα μπορούσε να απαιτείται μόνο το ένα αντίγραφο έτσι το νέο αντίγραφο καθίσταται αυτομάτως περιττό. Σε αυτήν την περίπτωση το νέο διπλοποιημένο γονίδιο είναι ελεύθερο από περιορισμούς επιλογής και έτσι θα συσσωρεύσει μεταλλάξεις και ενδεχομένως θα γίνει ένα ψευδογονίδιο. Άλλη πιθανότητα είναι η φυσική επιλογή να στοχεύσει ένα γονίδιο ελαφρώς τροποποιημένης λειτουργίας η οποία θα μπορούσε να εκπληρωθεί από το νέο αντίγραφο όπως στην περίπτωση των γονιδίων της αιμοσφαιρίνης με τις διαφορετικές ικανότητες μεταφοράς οξυγόνου (Futuyama 1991, Page & Holmes 1998). Παρότι η διπλοποίηση είναι ο σημαντικότερος τρόπος με τον οποίο γονίδια με νέες λειτουργίες μπορούν να προκύψουν, σε κάποιες σπάνιες περιπτώσεις γονίδια έχουν αποκτήσει νέες λειτουργίες χωρίς να έχει μεσολαβήσει διπλοποίηση· παράδειγμα οι

κρυσταλλίνες. Φαίνεται ότι η αρχική λειτουργία της πρωτεΐνης ήταν εκείνη του ενζύμου αλλά κατόπιν στρατολογήθηκε σε ένα νέο δομικό ρόλο μέσω αλλαγών στην έκφραση του γονιδίου.

Κάποιες πολυγονιδιακές οικογένειες έχουν ένα μεγάλο αριθμό γονιδίων που επαναλαμβάνονται εν σειρά σχηματίζοντας συστοιχίες. Έτσι περιέχουν πολλαπλά αντίγραφα πανομοιότυπων γονιδίων με την ίδια λειτουργία. Προφανώς εξελίχθηκαν επειδή απαιτούνται μεγάλα ποσά της πρωτεΐνης που κωδικεύουν. Ένα καλό παράδειγμα αυτών είναι η rDNA συστοιχία η οποία κωδικεύει το ριβοσωμικό RNA, (rRNA). Λόγω του κρίσιμου ρόλου που παίζουν τα ριβοσώματα στην σύνθεση των πρωτεϊνών τα rRNAs βρίσκονται σε κάθε είδος και συνήθως απαντώνται σε τεράστιες ποσότητες. Παρά την ευρεία χρήση ως φυλογενετικό εργαλείο οι rDNA συστοιχίες, δεν εξελίσσονται πάντα με τον ίδιο τρόπο. Ειδικότερα όταν rDNA αλληλουχίες συγκριθούν εντός και μεταξύ ειδών ένα ενδιαφέρον πρότυπο αναδύεται: συχνά υπάρχει ένας μεγάλος βαθμός ομοιότητας της αλληλουχίας εντός του είδους αλλά μια μεγάλη απόκλιση μεταξύ των ειδών. Αυτό είναι μη σύνηθες διότι ένα χαμηλό επίπεδο της ποικιλότητας εντός των ειδών κανονικά συσχετίζεται με ένα χαμηλό επίπεδο της απόκλισης μεταξύ των ειδών. Αυτό το ασυνήθιστο εξελικτικό πρότυπο αναδύεται επειδή υπάρχουν γενετικοί μηχανισμοί, ιδιαίτερα της γονιδιακής μετατροπής (gene conversion) που σχετίζεται με το γενετικό ανασυνδυασμό και τη μεταφορά αλληλουχιών DNA μεταξύ των γονιδίων που έτσι εξελίσσονται μαζί. Αυτή η διεργασία είναι γνωστή ως εναρμονισμένη εξέλιξη (concerted evolution) και είναι μια από της σημαντικότερες που δρουν στις πολυγονιδιακές οικογένειες γιατί σημαίνει ότι μεταλλάξεις διαδίδονται σε όλα τα μέλη ακόμα και στην περίπτωση που αυτά εδρεύουν σε διαφορετικά χρωμοσώματα. Ενώ η εναρμονισμένη εξέλιξη επιτρέπει στα γονίδια να εξελίσσονται μαζί καθίσταται επίσης εξαιρετικά περίπλοκη η φυλογενετική ανάλυση των πολυγονιδιακών οικογενειών διότι γίνεται δύσκολη η διάκριση των ομολόγων γονιδίων. Επιπρόσθετα το σημείο της διαφοροποίησης της αλληλουχίας πλέον δεν αντιστοιχεί στη στιγμή του γεγονότος της διπλοποίησης του γονιδίου (Page & Holmes 1998, Ροδάκης 2001).

1.1.4 Λειτουργική διαφοροποίηση των διπλοποιημένων γονιδίων

Η εξέλιξη των οικογενειών παράλογων γονιδίων ενδιαφέρει διότι η διπλοποίηση είναι ένας σημαντικός μηχανισμός με τον οποίο εξελίσσονται τα νέα γονίδια. Έχουν αναπτυχθεί στατιστικές μέθοδοι για την ανίχνευση της λειτουργικής διαφοροποίησης στις οικογένειες γονιδίων (Casari *et al.* 1995, Gu *et al.* 2002, Gu *et al.* 2013) και διερευνώνται οι σχέσεις μεταξύ των προσαρμοστικών δυνάμεων που σταθεροποιούν τις διπλοποιήσεις και εκείνων που επιτρέπουν την εξέλιξη νέων λειτουργιών. Η εξέλιξη των διπλοποιημένων γονιδίων είναι ενδιαφέρουσα και σημαντική όχι μόνο διότι η διπλοποίηση είναι η κύρια πηγή νέων γονιδίων σε οποιοδήποτε γονιδίωμα, αλλά και διότι δίδει τη δυνατότητα για την εξέταση της λειτουργικής απόκλισης των γονιδίων υπό ίδιες συνθήκες. Παράλογα γονίδια υπάρχουν εντός του ίδιου γονιδιώματος και υπόκεινται π.χ. στις ίδιες περιβαλλοντικές πιέσεις και αυτό ενισχύει σημαντικά την ερμηνεία των διαφορών που έχουν συσσωρευτεί σε διπλοποιημένα γονίδια από τη στιγμή της προέλευσής τους. (Makino *et al.* 2006).

1.1.4.1 Μηχανισμοί εξέλιξης

Ψευδογονίδια: η πιο συνηθισμένη τύχη για ένα διπλοποιημένο γονίδιο προκύπτει από την τυχαία ουδέτερη συσσώρευση μεταλλάξεων, οι περισσότερες από τις οποίες είναι επιβλαβείς. Τελικά η διαδικασία οδηγεί στην εξάλειψη του ως λειτουργικό γονίδιο (Kondrashov *et al.* 2002). Υπολειτουργικοποίηση (Subfunctionalization): περιλαμβάνει ένα συνδυασμό ουδέτερων μηχανισμών και αρνητικής επιλογής που οδηγεί στην μείωση του πλεονασμού των διπλών αντιγράφων μέσω συμπληρωματικής απώλειας λειτουργικών χαρακτηριστικών τους. Γονίδια που έχουν περισσότερες ρυθμιστικές περιοχές, συμπεριλαμβανομένων αυτών που ρυθμίζουν την ανάπτυξη, θα είναι πιο επιρρεπή σε υπολειτουργικοποίηση. Νεολειτουργικοποίηση (neofunctionalization): συνεπάγεται την ανάπτυξη νέων λειτουργιών. Αυτό μπορεί να περιλαμβάνει τη δημιουργία εκ νέου παραγόντων μεταγραφής, την τροποποίηση των θέσεων δέσμευσης, την τροποποίηση των υφιστάμενων αλληλεπιδράσεων για την εξειδίκευση, τη συγγένεια ή την κινητική, την τροποποίηση των ρυθμιστικών στοιχείων και έναν αριθμό άλλων γεγονότων. Λόγω της πολυπλοκότητας των αλληλεπιδράσεων των μεταλλάξεων όχι μόνο εντός αλλά και μεταξύ των γονιδίων, οι ρυθμοί της νεολειτουργικοποίησης μπορεί να εμφανίζουν χρονική υστέρηση και είναι πιο πολύπλοκοι από εκείνους τους ρυθμούς που εμφανίζονται άμεσα σε περιπτώσεις ευεργετικών μεταλλάξεων μέσα σε έναν πληθυσμό. Επιπλέον, μια νέα λειτουργία σε μοριακό

επίπεδο δε συνεπάγεται απαραίτητα ένα επιλεκτικό πλεονέκτημα και θετική επιλογή, ορισμένες νέες μοριακές λειτουργίες θα είναι εξελικτικά ουδέτερες. Θα πρέπει να σημειωθεί όμως ότι η υπολειτουργικοποίηση θα μπορούσε να είναι και μια κατάσταση μετάβασης προς τη νεολειτουργικοποίηση με την έννοια ότι η υπολειτουργικοποίηση εξακολουθεί να είναι σημαντική για σύντομα χρονικά πλαίσια μετά το γεγονός της διπλοποίησης. Ωστόσο, με την πάροδο του χρόνου, ο ρόλος της φαίνεται να είναι η διατήρηση των διπλοποιημένων αντιγράφων για ενδεχόμενη νεολειτουργικοποίηση. Η σημασία της νεολειτουργικοποίησης ως τερματικής μοίρας γίνεται ακόμα πιο δραματική με την αύξηση του μεγέθους του πληθυσμού (Hughes 1994, Hughes 2005, Rastogi & Liberles 2005, Innan & Kondrashov 2010).

1.2 Η εξέλιξη των δομών των πρωτεϊνών και των δομικών συνόλων υπόκεινται σε λειτουργικούς περιορισμούς

Λειτουργικοί περιορισμοί και περιορισμοί αναδίπλωσης επιβάλλουν αλληλεξάρτηση μεταξύ αλληλεπιδρώντων θέσεων κατά μήκος της πρωτεϊνικής αλυσίδας όπως προβλέπεται μέσω της εξέλιξης της πρωτεϊνικής αλληλουχίας (Siltberg *et al.* 2011). Η αλληλουχία του γονιδίου, η αλληλουχία της πρωτεΐνης, καθώς και η δομή της και η λειτουργία της, συνδέονται εγγενώς με την εξέλιξη και τη γενετική του πληθυσμού. Επιπλέον, εξελικτικά, οι μεταβολές στις αλληλουχίες γονιδίων, όπως φιλτράρονται από τη δομή και τη λειτουργία των πρωτεϊνών, μπορούν να οδηγήσουν σε φαινοτυπική αλλαγή μέσω ουδέτερων και προσαρμοστικών μηχανισμών. Η επιλογή μπορεί να συμβεί τελικά στο επίπεδο της καταλληλότητας του μεμονωμένου οργανισμού, μέχρι το επίπεδο της πρωτεϊνικής λειτουργίας, της δομής και της αλληλουχίας. Τελικά, είναι απαραίτητη η κατανόηση της αλληλεπίδρασης της πρωτεϊνικής δομής τόσο με την εξέλιξη της αλληλουχίας όσο και με τη λειτουργική / φαινοτυπική εξέλιξη (Andreeva & Murzin 2006).

Η κατανόηση της εξέλιξης των πρωτεϊνικών δομών υπό λειτουργικό περιορισμό εξακολουθεί να είναι αντικείμενο υπό μελέτη και ένας κρίσιμος σύνδεσμος μεταξύ αλληλουχίας γονιδίων και μοριακής και βιολογικής λειτουργίας (Siltberg *et al.* 2011). Σε σφαιρικές πρωτεΐνες, οι συναρμολογημένες δομικές ενότητες χαρακτηρίζονται κατά ένα μεγάλο ποσοστό δευτεροταγούς δομής από έναν υδρόφοβο πυρήνα, μια κυρίως υδρόφιλη επιφάνεια και στις πρωτεΐνες με μια λειτουργία δέσμευσης, μια επιλεκτικά προσαρμοσμένη διεπαφή σύνδεσης εμπλεκόμενη στην αλληλεπίδραση πρωτεΐνης-προσδέτη (Shakhnovich *et al.* 2005). Κατά τη διάρκεια της εξέλιξης, αυτές οι περιοχές εμφανίζουν διαφορετικούς

ρυθμούς αντικατάστασης αμινοξέων με τον υδρόφοβο πυρήνα, συμπεριλαμβανομένων των καταλοίπων τα οποία μετέχουν στον προσδιορισμό της τοπολογίας της φυσικής αναδίπλωσης (Pereira de Araujo & Onuchic 2009), να εξελίσσονται πιο αργά από την υδρόφιλη επιφάνεια (Siltberg *et al.* 2011, Roth & Liberles 2006) Η διεπαφή της κοιλότητας δέσμευσης με τον προσδέτη βρίσκεται υπό λειτουργικό περιορισμό και εξελίσσεται πιο αργά, με διαφορές ρυθμού μεταξύ καταλοίπων που καθορίζουν τη συγγένεια και προσδιορίζουν την εξειδίκευση (Pechmann *et al.* 2009) σε σχέση με εκείνα στις περιφερειακές περιοχές (Pál *et al.* 2006). Εντός των οικογενειών πρωτεϊνών, οι αλλαγές των αμινοξέων του πυρήνα γίνονται πολύ αργά (Williams & Lovell 2009) διατηρώντας κυρίως τη χαρακτηριστική τοπολογία της αναδίπλωσης σε σχετικά μεγάλες εξελικτικές αποστάσεις. Η δομή υπαγορεύει τις αλληλεπιδράσεις που εμφανίζονται μεταξύ των καταλοίπων και η θερμοδυναμική ενδομοριακή σύζευξη αντικαταστάσεων είναι ανιχνεύσιμη από εξελικτικά δεδομένα (Lockless & Ranganathan 1999), οδηγώντας στη χρήση χαρτών επαφής και προβολής πρωτεϊνών σε ένα πλαίσιο δικτύου (Böde *et al.* 2007). Διαφορετικά δευτεροταγή δομικά στοιχεία παρουσιάζουν επίσης διαφορετικούς ρυθμούς εξέλιξης, με περιοχές βήτα-κλώνων να εξελίσσονται πιο αργά από τις περιοχές έλικας και περιοχές με στροφές και βρόχους να εξελίσσονται ταχύτερα (Schaefer *et al.* 2010).

1.3 Επίσταση

Έχουν περάσει περίπου 100 χρόνια από τότε που ο William Bateson (1909) εισήγαγε τον όρο "επίσταση" για να περιγράψει διαφορετικούς τύπους αλληλεπίδρασης γονιδίων, συσχετίζοντας το γονότυπο και τη φαινοτυπική αναλογία μιας διασταύρωσης. Η χρήση του όρου της επίστασης έχει επεκταθεί έκτοτε για την περιγραφή σχεδόν κάθε συνόλου αλληλεπιδράσεων μεταξύ γενετικών τόπων. Η επίσταση έχει χρησιμοποιηθεί για να περιγράψει μια σειρά φαινομένων, συμπεριλαμβανομένης της λειτουργικής αλληλεπίδρασης μεταξύ των γονιδίων, του γενετικού αποτελέσματος των μεταλλάξεων που δρουν εντός της ίδιας γενετικής οδού και της στατιστικής απόκλισης της προσθετικής δράσης των γονιδίων.

Η επίσταση ως πρωταρχικός παράγοντας στην μοριακή εξέλιξη συμβάλλει στη βαθύτερη κατανόηση της λειτουργικής βάσης των γονιδιακών αλληλεπιδράσεων δημιουργώντας μια συναρπαστική διασταύρωση ανάμεσα σε ένα ευρύ σύνολο γενετικών επιστημών, που κυμαίνονται από τη βιοχημεία πρωτεϊνών ως την εξελικτική γενετική. Ποτέ δεν υπήρξε καλύτερη δυνατότητα για την αξιολόγηση του ρόλου που διαδραματίζουν οι

γονιδιακές αλληλεπιδράσεις εντός των βιολογικών συστημάτων παρά τώρα με την τόσο διαθέσιμη πληροφορία από την πληθώρα δεδομένων πλήρους αλληλούχισης γονιδιωμάτων (Phillips 2008). Οι μοριακοί γενετιστές μελετούν τώρα πώς συγκεκριμένες αλληλεπιδράσεις αλληλομόρφων επιδρούν σε σύνθετα ρυθμιστικά δίκτυα, ενώ οι εξελικτικοί γενετιστές κινούνται από τις στατιστικές περιγραφές της γενετικής παραλλαγής έως την αναγνώριση των ειδικών αντικαταστάσεων νουκλεοτιδίων που είναι υπεύθυνες για την προσαρμοστική εξέλιξη. Αυτό που έγινε σαφές στον αιώνα που έχει περάσει από τότε που εισήχθη η έννοια της επίστασης είναι ότι τα περισσότερα από τα συστήματα που βασίζονται στην κυτταρική, αναπτυξιακή και φυσιολογική λειτουργία αποτελούνται από πολλά στοιχεία που αλληλεπιδρούν μεταξύ τους με συχνά πολύπλοκους τρόπους.

Οι χρήσεις του όρου "επίσταση" μπορούν να συμπτυχθούν σε τρεις μεγάλες κατηγορίες, καθεμία από τις οποίες έχει ονομαστεί με πολλούς διαφορετικούς τρόπους. 1) Λειτουργική επίσταση. Ο όρος αυτός αναφέρεται στις μοριακές αλληλεπιδράσεις που έχουν κυρίως οι πρωτεΐνες μεταξύ τους, είτε λειτουργούν εντός της ίδιας οδού, είτε αποτελούνται από πρωτεΐνες που συγκροτούν σύμπλοκα μεταξύ τους. Πρόκειται για μια αυστηρά λειτουργική περιγραφή, χωρίς άμεση γενετική σύνδεση, παρόλο που προφανώς προβλέπεται μια γενετική συνέπεια, εάν η λειτουργική σχέση διαταραχθεί. Ωστόσο αυτή η χρήση αποφεύγεται και περιορίζεται η χρήση μόνο για την περιγραφή των συνεπειών των πολυμορφισμών, όπως περιγράφεται παρακάτω. Υπάρχουν άλλοι περιγραφικοί όροι, όπως η αλληλεπίδραση πρωτεΐνης-πρωτεΐνης, που μπορούν να χρησιμοποιηθούν στο πλαίσιο της λειτουργικής επίστασης. 2) Συνθετική επίσταση. Έτσι περιγράφεται η παραδοσιακή χρήση της επίστασης ως παρεμπόδιση του αποτελέσματος ενός αλληλόμορφου από έναν αλληλόμορφο σε άλλο τόπο. Η συνθετική επίσταση περιγράφει τον τρόπο με τον οποίο ένας συγκεκριμένος γονότυπος υφίσταται και την επίδραση που έχει αυτό το συγκεκριμένο γενετικό υπόβαθρο στις επιδράσεις ενός δεδομένου συνόλου αλληλόμορφων. 3). Στατιστική επιστήμη. Έτσι περιγράφεται από τον Fisher (1918) όπου η μέση απόκλιση των συνδυασμών αλληλόμορφων σε διαφορετικούς τόπους υπολογίζεται σε όλους τους άλλους γονότυπους που υπάρχουν σε έναν πληθυσμό. Η προσέγγιση του Fisher επιλύει δύο προβλήματα. Πρώτον, δίνει τα κατάλληλα μέτρα για την περιγραφή της εξελικτικής αλλαγής. Δεύτερον, υπάρχουν πολύ περισσότεροι γενετικοί συνδυασμοί σε έναν πληθυσμό (ακόμη και σε βακτήρια) από ό, τι μπορεί ποτέ να πραγματοποιηθεί, οπότε δεν υπάρχει ένα πρότυπο γενετικό υπόβαθρο μέσα σε ένα φυσικό πληθυσμό, αλλά μόνο το σύνολο των υποβάθρων που είναι πιθανότερο να απαντηθούν σε ένα δεδομένο δείγμα πληθυσμού. Μια προσέγγιση συνθετικής επίστασης δεν μπορεί να εφαρμοστεί στους φυσικούς πληθυσμούς επειδή είναι αδύνατο να απαριθμηθούν

όλες οι πιθανές γενετικές αλληλεπιδράσεις για οποιοδήποτε πραγματικό πληθυσμό. Η στατιστική επίσταση μπορεί να θεωρηθεί ως η απόκλιση που προκαλείται από την ταυτόχρονη αντικατάσταση δύο αλληλόμορφων σε διαφορετικούς τόπους σε ένα τυχαία επιλεγμένο άτομο εντός του πληθυσμού, αφού ληφθεί υπόψη το τι θα αναμενόταν να είναι το αποτέλεσμα της αντικατάστασης κάθε αλληλόμορφου χωριστά. Η συνθετική επίσταση μετρά τα αποτελέσματα της αντικατάστασης αλληλόμορφων έναντι ενός συγκεκριμένου σταθερού γενετικού πλαισίου, ενώ η στατιστική επίσταση μετρά το μέσο αποτέλεσμα της αντικατάστασης αλληλόμορφων έναντι του μέσου γενετικού υποβάθρου του πληθυσμού (Phillips 2008).

Ένα από τα ωραιότερα παραδείγματα επίστασης εντός ενός τόπου προέρχεται από την έρευνα των Ortlund *et al.* (2007) για την εξέλιξη της νέας λειτουργίας σε υποδοχείς στεροειδών σπονδυλωτών. Αποδεικνύεται ότι η πρωτεΐνη των προγόνων είναι αρκετά μη εξειδικευμένη και αλληλεπιδρά με μια ποικιλία στεροειδών προσδετών, ακόμη και με προσδέτες που δεν υπάρχουν στον προγονικό οργανισμό εκείνη την εποχή. Επομένως, η εξειδίκευση πραγματοποιήθηκε μέσω της εξέλιξης ενός ειδικού για γλυκοκορτικοειδή υποδοχέα από ένα γενικότερο πρόγονο για κορτικοειδή, που επιτεύχθηκε μέσω αλλαγών σε δύο αλληλεπιδρούσες θέσεις. Μαζί, αυτές οι δομικές αλληλεπιδράσεις δημιουργούν μια συγκεκριμένη τάξη στην οποία πρέπει να εμφανιστούν οι εξελικτικές αντικαταστάσεις. Υπάρχουν διάφορα πιθανά μονοπάτια για αυτές τις αλλαγές, αλλά μόνο λίγα είναι λειτουργικώς βιώσιμα επειδή η αποκαλούμενη "στερεοδιατακτική επίσταση" που παράγεται από τη αποτυχία της δομικής αναδίπλωσης της πρωτεΐνης περιορίζει τις εξελικτικές επιλογές. Σε αυτό το παράδειγμα υπάρχει μια άμεση σύνδεση μεταξύ ειδικών αμινοξικών αλλαγών, αλληλεπιδράσεων επίστασης που παράγονται από την επίδρασή τους στην δομή της πρωτεΐνης και την επίδραση που αυτές οι αλληλεπιδράσεις έχουν στην μεταγενέστερη εξελικτική αλλαγή.

Η επίσταση με την έννοια των αλληλεπιδράσεων μεταξύ των γονιδίων εντός του γονιδιώματος μπορεί να είναι άμεση εάν τα γονίδια κωδικεύουν πρωτεΐνες που είναι, για παράδειγμα, ξεχωριστά συστατικά μιας πρωτεΐνης πολλών συστατικών όπως το ριβόσωμα, εάν αναστέλλουν τη δράση το ένα του άλλου ή εάν η πρωτεΐνη που κωδικεύεται από ένα γονίδιο τροποποιεί μια άλλη. Εναλλακτικά η αλληλεπίδραση μπορεί να είναι έμμεση, όπου τα γονίδια κωδικεύουν συστατικά μιας μεταβολικής οδού, μιας αναπτυξιακής οδού, ενός μονοπατιού σηματοδότησης ή ενός δικτύου (Phillips 2008).

Όπως αναφέρθηκε (§1.2) οι πρωτεΐνες διατηρούν την τριτοταγή δομή τους μέσω ενός καταναμημένου, εσωτερικού δικτύου αλληλεπιδράσεων (υδρόφοβες, πολικές, Van der Waals

και ηλεκτροστατικές -υδρογονικοί δεσμοί και αλληλεπιδράσεις φορτίου-). Οι επιστατικές αλληλεπιδράσεις συμβαίνουν κάθε φορά που μία μετάλλαξη μεταβάλλει το τοπικό περιβάλλον ενός άλλου αμινοξικού καταλοίπου (είτε με άμεση επαφή με αυτό είτε με επαγωγή αλλαγών στην δομή της πρωτεΐνης) (Avery & Wasserman 1992).

Για παράδειγμα, σε μια δισουλφιδική γέφυρα, μία μόνη κυστεΐνη δεν έχει καμία επίδραση στη σταθερότητα της πρωτεΐνης μέχρις ότου υπάρξει μια δεύτερη στη σωστή θέση, οπότε οι δύο κυστεΐνες σχηματίζουν ένα χημικό δεσμό που ενισχύει τη σταθερότητα της πρωτεΐνης (Harms & Thornton 2013, Witt 2008). Αυτό παρατηρήθηκε ως θετική επίσταση όπου η παραλλαγή με τη διπλή κυστεΐνη είχε πολύ μεγαλύτερη σταθερότητα από οποιαδήποτε από τις παραλλαγές μιας μόνο κυστεΐνης. Αντίθετα, όταν εισάγονται δηλητηριώδεις μεταλλάξεις, οι πρωτεΐνες συχνά εμφανίζουν μεταλλακτική ευρωστία, όπου καταστρέφονται ως σταθεροποιητικές αλληλεπιδράσεις, η πρωτεΐνη εξακολουθεί να λειτουργεί έως ότου φθάσει σε κάποιο όριο σταθερότητας, σημείο στο οποίο περαιτέρω αποσταθεροποιητικές μεταλλάξεις έχουν μεγάλες, επιζήμιες επιδράσεις καθώς η πρωτεΐνη δεν μπορεί πλέον να αναδιπλώνεται. Αυτό οδηγεί σε αρνητική επίσταση, όπου οι μεταλλάξεις που έχουν μικρή μόνο δράση ξεχωριστά έχουν μεγάλη, επιβλαβή επίδραση μαζί (Bershtein *et al.* 2006, Steinberg & Ostermeier 2016).

Στην τρισδιάστατη δομή των ενζύμων υπάρχουν ορισμένα αμινοξέα προσανατολισμένα σε ακριβείς γεωμετρίες, για να σχηματίσουν ένα ενεργό κέντρο ώστε να εκτελέσουν μια χημική αντίδραση (Halabi *et al.* 2009). Η μετάλλαξη οποιουδήποτε από αυτά τα αμινοξέα υποβαθμίζει τη δραστηριότητα και έτσι η μετάλλαξη ενός δευτέρου έχει σχετικά μικρή επίδραση στο ήδη αδρανοποιημένο ένζυμο. Μπορεί, η αντικατάσταση κάποιου μέλους της καταλυτικής τριάδας αμινοξέων ενός τέτοιου ενζύμου, να μειώσει τη δραστηριότητά του ακόμα και ως το επίπεδο της μη βιωσιμότητας του οργανισμού (Neet & Koshland 1966, Beveridge 1996, Sigal *et al.* 1982).

Η επίσταση ή οι αλληλεπιδράσεις μεταξύ των γονιδίων έχουν αναγνωριστεί από καιρό ως θεμελιωδώς σημαντικές για την κατανόηση τόσο της δομής και της λειτουργίας των γενετικών οδών όσο και της εξελικτικής δυναμικής των πολύπλοκων γενετικών συστημάτων. Με την εμφάνιση της λειτουργικής γονιδιωματικής υψηλής απόδοσης, την εμφάνιση των προσεγγίσεων συστημάτων στη βιολογία, καθώς και την ικανότητα παρακολούθησης της γενετικής βάσης της εξέλιξης μέσω συγκεκριμένων μοριακών μεταβολών, υπάρχει μια ανανεωμένη εκτίμηση, τόσο για τη σπουδαιότητα της μελέτης αλληλεπιδράσεων γονιδίων, όσο και για την αντιμετώπιση αυτών των ζητημάτων με ενοποιημένο και ποσοτικό τρόπο (Phillips 2008).

1.4. Οριζόντια μεταφορά του γονιδίου της ινωδονεκτίνης τύπου III

Η οριζόντια μεταφορά γονιδίων έχειδειχθεί ότι κατέχει αναπόσπαστο ρόλο στην εξέλιξη των βακτηριακών γονιδιωμάτων (Bork & Doolittle 1992, Leahy *et al.* 1992). Η αλληλουχία αμινοξέων της ινωδονεκτίνης (fibronectin) αποκάλυψε ότι είναι μια πρωτεΐνη αποτελούμενη από ομόλογες επαναλήψεις τριών πρωτότυπων τύπων δομικών ενοτήτων γνωστών ως τύποι I, II και III (Petersen *et al.* 1983, Kornbliht *et al.* 1985, Potts & Campbell 1994). Η δομική ενότητα ινωδονεκτίνης τύπου III (fibronectin type 3, Fn3 ή FnIII), είναι πολύ κοινή σε ζωικές πρωτεΐνες (Bork & Doolittle 1992). Παρουσιάζει ένα ξεχωριστό μοτίβο (Patthy 1990, Bazan 1990) που αποτελείται από 80-100 αμινοξικά κατάλοιπα ενός β-σάντουιτς με δύο β-πτυχωτές επιφάνειες τριών και τεσσάρων αντιπαράλληλων β-κλώνων, που δεν είναι διαφορετική από την αναδίπλωση που παρατηρείται στις ανοσοσφαιρίνες, χωρίς παρουσία δισουλφιδικών δεσμών (Baron *et al.* 1992, deVos *et al.* 1992, Leahy *et al.* 1992).

Αρχικώς η ανακάλυψη των πρωτεϊνών που περιέχουν ακολουθίες Fn3 στα βακτήρια αποκάλυψαν ότι, ανήκαν σε μία συγκεκριμένη κατηγορία πρωτεϊνών: εξωκυτταρικές γλυκοϋδρολάσες (Gilkes *et al.* 1991, Bork & Doolittle 1992, Hansen 1992). Δεδομένου ότι η κατανομή των Fn3 εμφανίζονταν σποραδικά σε βακτηριακά φυλογενετικά δένδρα και οι βακτηριακές αλληλουχίες ήταν περισσότερο όμοιες με κάποιες αλληλουχίες Fn3 από ζωικές πηγές παρά οι ζωικές αλληλουχίες Fn3 μεταξύ τους, η παρουσία τους στα βακτήρια θεωρείται ως το πιο πειστικό παράδειγμα οριζόντιας γονιδιακής μεταφοράς από ζωική πηγή σε προκαρυώτη (Bork & Doolittle 1992).

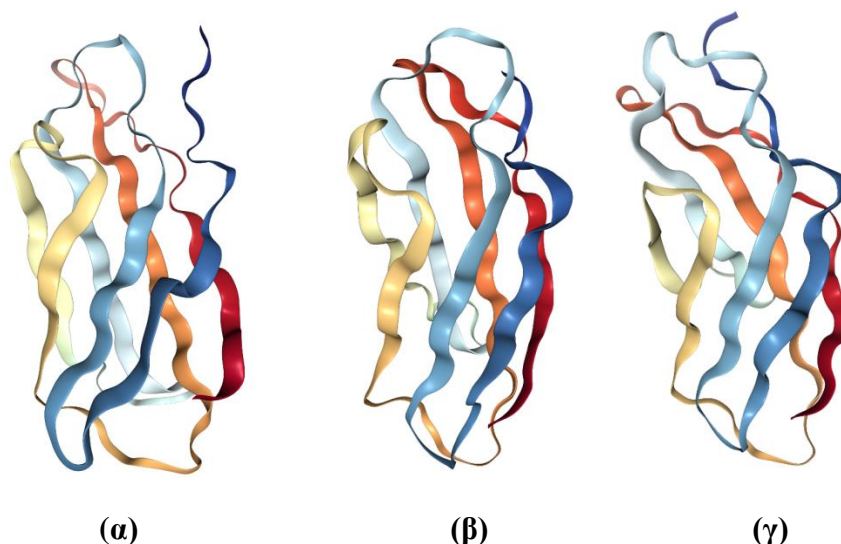
Στην ίδια εργασία οι Bork και Doolittle αναφέρουν ότι η πιο όμοια ζωική αλληλουχία με τις αλληλουχίες του βακτηριακού συνόλου που εξετάστηκε ήταν μια αλληλουχία του ανθρώπινου λευκοκυτταρικού αντιγόνου. Θεωρείται ότι τα τμήματα DNA που κωδικεύουν τη δομική ενότητα διακινούνται στα γονιδιώματα των βακτηρίων με τέτοιο τρόπο που υποδηλώνει ανακατάταξη των λειτουργικών μονάδων (Meinke *et al.* 1991, Gilkes *et al.* 1991, Bork & Doolittle 1992). Με βάση αυτήν την υπόθεση οι αναδιοργανώσεις αυτές είναι εξελικτικά πρόσφατες, και παραδείγματα του μηχανισμού της ανακατάταξης εξωνίων (exon shuffling) (Little *et al.* 1994). Σύμφωνα με το μηχανισμό αυτό νέα γονίδια έχουν προκύψει με αποτέλεσμα την εξέλιξη πρωτεϊνών με πολυλειτουργικές περιοχές, από τη σύνδεση δηλαδή διαφορετικών εξωνίων σε νέους συνδυασμούς (Doolittle 1985, Patthy 1985).

Οι ινωδονεκτίνες έχουν πολλές περιοχές που προσδίδουν την ικανότητα αλληλεπίδρασης με διάφορα εξωκυττάρια μόρια π.χ. κολλαγόνο, ινώδες (ινίνη, fibrin), ηπαρίνη, και υποδοχείς στην επιφάνεια των κυττάρων, μεμβρανικούς υποδοχείς ενώ έχει

αποδειχθεί ότι εμπλέκεται στην κυτταρική προσκόλληση (cell adhesion protein, Hynes 1990). Η περιοχή RGD (αργινίνη-γλυκίνη-ασπαρτικό οξύ) είναι χαρακτηριστική, αναγνωρίζεται από τις ιντεγκρίνες (integrins) και αρχικά ταυτοποιήθηκε στην ινωδονεκτίνη το 1983 (Pierschbacher & Ruoslahti 1984). Οι αλληλουχίες RGD που είναι δραστικές σε διαμοριακές αλληλεπιδράσεις έχουν βρεθεί σε πολυάριθμες πρωτεΐνες και μεσολαβούν σε μια ευρεία ποικιλία συμβάντων προσκόλλησης κυττάρων (Ruoslahti & Pierschbacher 1987) και εμπλέκονται σε ποικιλία βιολογικών διεργασιών όπως η θρόμβωση, η φλεγμονή, η μετάσταση όγκων και η διαμεσολάβηση μεταγωγής σήματος (Papadopoulos *et al.* 1988, Oberhauser *et al.* 1998). Εκτός από την τριάδα RGD στην αλληλεπίδραση μετέχει συνήθως και ένας γειτονικός επιφανειακά εκτεθειμένος βρόχος, η επονομαζόμενη συνεργιστική περιοχή (Main *et al.* 1992, Bateman *et al.* 1996, Leahy *et al.* 1996, Plaxco *et al.* 1996, Sharma *et al.* 1999).

Η πρώτη τρισδιάστατη δομή από μια βακτηριακή πρωτεΐνη προσδιορίστηκε με φασματοσκοπία NMR και ήταν η Fn3 της χιτινάσης A1 από τον *Bacillus circulans* WL-12 (Jee *et al.* 2002). Η δομή της δείχνει σημαντική ομοιότητα με Fn3 από ζωικές πρωτεΐνες θηλαστικών (**Εικόνα 1.1**).

Οι εντυπωσιακές ομοιότητες στις τριτοταγείς δομές των βακτηριακών Fn3 και των αντίστοιχων θηλαστικών τους μπορεί να υποστηρίξουν την υπόθεση της οριζόντιας μεταφοράς της δομικής ενότητας Fn3. Οι συγκρίσεις τρισδιάστατων δομών Fn3 βακτηρίων δείχνουν ότι τα κατάλοιπα σχηματισμού του πυρήνα είναι εξαιρετικά συντηρημένα και επομένως βρίσκονται υπό ισχυρή εξελικτική πίεση. Η έλλειψη επιφανειακά εκτεθειμένων αρωματικών καταλοίπων υποδεικνύει ότι ο ρόλος της Fn3 είναι διαφορετικός από εκείνους άλλων βακτηριακών δομικών ενοτήτων β-σάντουιτς, οι οποίες λειτουργούν ως ενότητες πρόσδεσης υδατανθράκων (Carbohydrate-Binding Modules, CBMs, Jee *et al.* 2002).



Εικόνα 1.1 (α) Δομική ενότητα Fn3 ανθρώπινης ιωδονεκτίνης που προσδιορίστηκε με φασματοσκοπία πυρηνικού μαγνητικού συντονισμού (NMR) (PDB ID: 1TTF, Main *et al.* 1992) (β) από την πρωτεΐνη του ανθρώπου τενασκίνη (γλυκοπρωτεΐνη) που προσδιορίστηκε με κρυσταλλογραφία ακτίνων-X (PDB ID: 1TEN, Leahy *et al.* 1992) (γ) της χιτινάσης A1 του *Bacillus circulans* WL-12 προσδιορίστηκε με φασματοσκοπία NMR και ήταν η πρώτη τρισδιάστατη δομή Fn3 από μια βακτηριακή πρωτεΐνη (PDB ID:1K85, Jee *et al.* 2002).

1.5 Τα βακτήρια *Bacillus anthracis* και *Bacillus cereus*

Ο *Bacillus cereus* είναι ένα ευρέως διαδεδομένο βακτήριο του εδάφους και ένα ευκαιριακό ανθρώπινο παθογόνο που προκαλεί προβλήματα στη βιομηχανία γάλακτος και γαλακτοκομικών προϊόντων και μπορεί να είναι η αιτία γαστρίτιδας και εντερίτιδας δύο συνδρόμων τροφικής δηλητηρίασης (Drobniowski 1993). Ο *B. anthracis* είναι υποχρεωτικό παθογόνο των θηλαστικών που προκαλεί την οξεία και συχνά θανατηφόρο νόσο του άνθρακα και έχει χρησιμοποιηθεί ως βιοτρομοκρατικός παράγοντας (Jernigan *et al.* 2001, Hughes & Gerberding 2002).

1.5.1 Η νόσος του άνθρακα

Η νόσος του άνθρακα είναι από τις παλαιότερες καταγεγραμμένες επιζωοτίες. Μνημονεύεται στην Παλαιά Διαθήκη (Εξοδος Θ' 9) και αναφέρεται από κλασσικούς

συγγραφείς όπως ο Ιπποκράτης, ο Γαληνός και ο Πλίνιος. Ο άνθρακας ήταν η πρώτη νόσος που ο αιτιολογικός της παράγοντας συνδέθηκε με μικρόβιο το 1863 από τον Casimir J. Davaine ενώ το απέδειξε ο Robert Koch το 1876 όταν το απομόνωσε.

Έχει ονομασθεί παγκοσμίως από την ελληνική λέξη *άνθραξ* διότι προκαλεί στο δέρμα μελανές νεκρωτικές εξελκώσεις που μοιάζουν με κάρβουνο. Δυνητικά όλα τα ζώα μπορούν να προσβληθούν συχνότερα όμως προσβάλλονται τα φυτοφάγα, βοοειδή, αιγοπρόβατα, ημίονοι και άλογα. Ενώ στα σαρκοφάγα και στον άνθρωπο μεταδίδεται από επαφή με πάσχοντα ζώα.

Τα τελευταία 140 περίπου χρόνια, ο άνθρακας αποτελούσε μόνο πρόβλημα της αγροτικής οικονομίας των αναπτυσσόμενων κυρίως χωρών ως σημαντική ασθένεια των εκτρεφόμενων ζώων. Στις μέρες μας παρουσιάζεται αναγκαίο να εστιάσουμε ξανά στον άνθρακα αλλά αυτήν τη φορά ως βιολογικό τρομοκρατικό παράγοντα. Ειδικότερα αφότου επιστολές στάλθηκαν με το στέλεχος *B. anthracis* Ames, κατά το διάστημα από την 11^η Σεπτεμβρίου 2001 έως την 29^η Ιανουαρίου 2002 (Spencer 2003, Matsumoto 2003). Έχουν υπάρξει αναφορές για τη χρήση του ως βιολογικό όπλο στο παρελθόν λόγω των εξαιρετικά ανθεκτικών του σπόρων (Meselson *et al.* 1994, Inglesby *et al.* 1999).

Κλινικά ο άνθρακας εμφανίζεται με τρεις μορφές. Η πιο κοινή είναι η δερματική, η οποία αντιπροσωπεύει πάνω από το 90% όλων των ανθρώπινων περιπτώσεων. Οι άλλες δύο μορφές είναι η γαστρεντερική και η πνευμονική. Η τελευταία φέρεται ως η πιο πιθανή μορφή στο πλαίσιο της βιολογικής τρομοκρατίας. Η πνευμονική μορφή του άνθρακα παρουσιάζει τη μεγαλύτερη θνητότητα ακόμη και με χορήγηση θεραπείας. Η αγωγή που ακολουθείται περιλαμβάνει χορήγηση τουλάχιστον διπλού σχήματος αντιβιοτικών. Ως πρώτο χορηγείται είτε σιπροφλοξασίνη ή δοξυκυκλίνη, ενώ επιλέγεται και συνδυασμός του τρόπου χορήγησης, ανάλογα με την περίπτωση. Τέλος θα πρέπει να αναφερθεί η εξαιρετικά παρατεταμένη χορήγηση των αντιβιοτικών που μπορεί να διαρκέσει για διάστημα ίσως και άνω των 60 ημερών (όταν γενικά, όπως είναι γνωστό, τα σχήματα των συνήθων περιπτώσεων έχουν διάρκεια 7-10 ημέρες) (Spencer 2003).

1.5. 2 Μελέτη γονιδιωμάτων της ομάδας *B. cereus*.

Το γένος *Bacillus* είναι ένα από τα καλύτερα χαρακτηρισμένα βακτηριακά γένη θετικών κατά Gram ραβδόμορφων βακτηρίων, που μπορεί να είναι αερόβια ή προαιρετικά αναερόβια και παράγουν εξαιρετικά ανθεκτικά ενδοσπόρια (Ravel & Fraser 2005). Η ομάδα

Bacillus cereus (*Bacillus cereus* group αναφερόμενη ως *B. cereus sensu lato*), αποτελεί μια ομάδα στενά συγγενικών βακτηρίων που περιλαμβάνει έξι είδη *B. anthracis*, *B. cereus*, *Bacillus thuringiensis*, *Bacillus mycoides*, *Bacillus pseudomycoides* and *Bacillus weihenstephanensis* (Helgason *et al.* 2000, Økstad & Kolstø 2011, Zwick *et al.* 2012). Το είδος *B. cytotoxicus* (ή *B. cytotoxicus*) έχει προταθεί ως μέλος της ομάδας (Lapidus *et al.* 2008) καθώς και τα *B. toyonensis*, *B. gaemokensis*, *B. manliponensis*, και *B. bingmayongensis* (βάση δεδομένων Tourasse & Kolstø—<http://mlstoslo.uio.no/>). Στην ομάδα υπάρχουν είδη με ιατρική και οικονομική σημασία (Tourasse *et al.* 2006). Η συστηματική ταξινόμηση, για παράδειγμα του στελέχους Ames του *B. anthracis*, έχει ως εξής: Bacteria, Firmicutes, Bacilli, Bacillales, Bacillaceae, *Bacillus*, *Bacillus cereus* group, *Bacillus anthracis*, *Bacillus anthracis* str. Ames.

Τα πρώτα γονιδιώματα που έγινε πλήρης αλληλούχιση από την ομάδα *B. cereus* και από τότε χρησιμεύουν ως γονιδιώματα αναφοράς, ήταν εκείνα του *B. anthracis* Ames (Read *et al.* 2003) και *B. cereus* ATCC 14579 (Ivanova *et al.* 2003).

Η ομάδα του *B. cereus* έχει το μεγαλύτερο αριθμό πλήρους αλληλούχισης, ολόκληρων γονιδιωμάτων των μελών της (Økstad & Kolstø 2011). Έχουν θεωρηθεί χωριστά είδη, που σε μεγάλο βαθμό αυτό έχει καθοριστεί από φαινοτυπικά χαρακτηριστικά και ιδιαίτερα από τα ειδικά πρότυπα παθογένειας. Ωστόσο, τα δεδομένα αλληλούχισης γονιδίου και γονιδιώματος επιβεβαίωσαν ότι πρόκειται για κοντινά συσχετιζόμενα είδη, όπως εξηγείται από το πολύ υψηλό επίπεδο ομοιότητας και σύνθεσης των χρωμοσωμάτων τους (Helgason *et al.* 2000, Ivanova *et al.* 2003, Read *et al.* 2003, Rasko *et al.* 2005). Παρουσιάζουν υψηλό βαθμό συντηρημένης γονιδιακής τάξης (Kolsto *et al.* 2009). Τα μέλη της ομάδας *B. cereus* είναι φαινοτυπικά και γενετικά ετερογενή. Ωστόσο, τα πολλά στελέχη τους μπορεί να εμφανιστούν φυλογενετικά αναμεμιγμένα στο χρωμοσωμικό επίπεδο (Økstad & Kolstø 2011).

Η ύπαρξη πλασμιδίων στην ομάδα *B. cereus* αποτελεί σημαντικό παράγοντα στον χαρακτηρισμό των μελών της (Tourasse *et al.* 2006). Η παθογένεια του *B. anthracis* προκαλείται κυρίως από δύο μεγάλα πλασμίδια (Mock & Fouet 2001, Bourgogne *et al.* 2003). Το pXO1 (182kb) (Okinaka *et al.* 1999) φέρει τα δομικά γονίδια που κωδικεύουν τα συστατικά της τοξίνης του άνθρακα, το προστατευτικό αντιγόνο (PA, *pag*), το θανατηφόρο παράγοντα (LF, *lef*) και τον παράγοντα οιδήματος (EF, *cya*) συναρμολογώντας τη θανατηφόρο τοξίνη (LT) και την τοξίνη του οιδήματος (ET) ως εξής: LT (PA+LF) και ET (PA+EF), και το pXO2 (95kb), που κωδικεύει την κάψουλα που περιβάλλει το βακτήριο, σε ένα οπερόνιο πέντε γονιδίων (*capBCADE*), η οποία αποτελείται από πολυ-γ-D-γλουταμικό

οξύ (Økstad & Kolstø 2011), σε αντίθεση με τις περισσότερες βακτηριακές κάψουλες πολυσακχαρίτη (Candela & Fouet 2006).

Η νόσος προκαλείται μέσω ενός συνδυασμού βακτηριακής λοίμωξης και τοξιναιμίας. Οι εξωτοξίνες που εκκρίνονται από τον *B. anthracis* παίζουν βασικούς ρόλους κατά τη διάρκεια της ασθένειας. Στα αρχικά στάδια της μόλυνσης, οι LT και ET συντονισμένα δρουν για να βλάψουν την έμφυτη ανοσοαπόκριση του ξενιστή, επιτρέποντας στον παθογόνο παράγοντα την επιτυχή μόλυνση. Στα τελευταία στάδια, όταν επιτυγχάνονται υψηλές συγκεντρώσεις τους, η LT και η ET προκαλούν απευθείας το θάνατο του ξενιστή μέσω στόχευσης ξεχωριστής ζωτικής σημασίας συστημάτων, ιδίως το καρδιαγγειακό σύστημα και το ήπαρ (Liu *et al.* 2014).

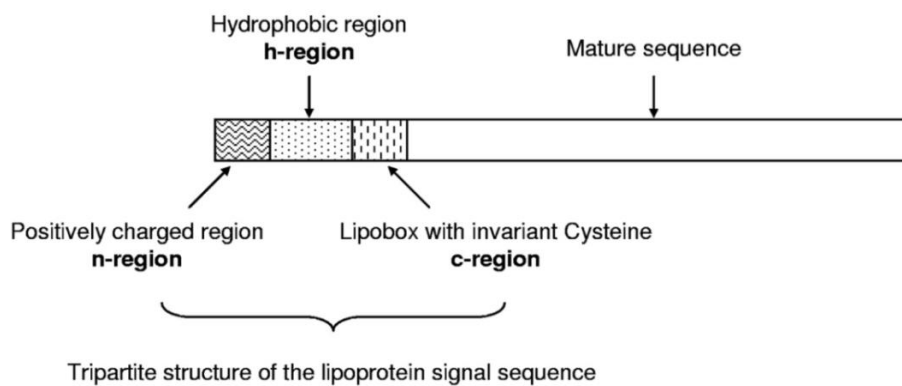
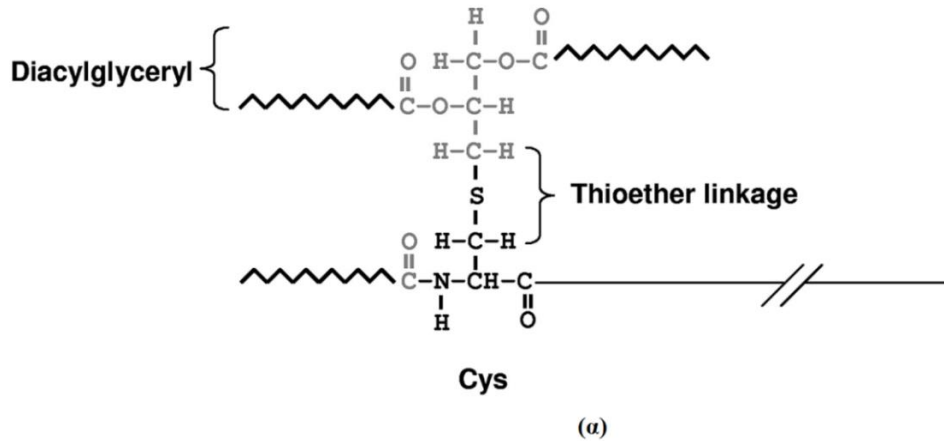
Σύμφωνα με τους Tourasse *et al.* (2006) η ομάδα *B. cereus* αποτελεί ένα συνεκτικό αλλά και δυναμικό πληθυσμό. Συνεκτικό γιατί όπως αναφέρθηκε τα μέλη της είναι ενοποιημένα ως προς την παρουσία πανταχού παρόντων συγκεκριμένων γενετικών στοιχείων, οι θέσεις και οι ακολουθίες των οποίων στο γονιδίωμα δεν επιτρέπουν καμία διάκριση μεταξύ τους. Αυτό φαίνεται από την εύρεση πολυάριθμων επαναλαμβανόμενων ακολουθιών που είναι κοινές σε όλα τα γονιδιώματα της ομάδας *B. cereus* που έχουν αλληλουχηθεί. Η ομάδα του *B. cereus* αποτελεί όμως και ένα δυναμικό πληθυσμό. Παρά το γεγονός ότι μοιράζονται ένα κοινό γενετικό υπόβαθρο, οι γενεαλογίες της ομάδας μπορούν να εξελιχθούν για να σχηματίσουν συμπλέγματα κλώνων, ενδεχομένως προσαρμοσμένα σε συγκεκριμένους ξενιστές ή/και περιβάλλοντα, και τα οποία τελικά παρουσιάζουν διαφορετικούς φαινοτύπους. Έχει συναχθεί προηγουμένως, ότι ο *B. anthracis* έχει αποκλίνει πρόσφατα από ένα σύμπλεγμα στελεχών των *B. cereus* και *B. thuringiensis* (Cherif *et al.* 2003, Helgason *et al.* 2004, Ko *et al.* 2004, Priest *et al.* 2004). Επίσης στελέχη της ομάδας *B. cereus* έχουν πρόσφατα εμφανιστεί ως μολυσματικοί κλώνοι. Ένας μηχανισμός που θα μπορούσε να συνεισφέρει σε αυτήν τη διαδικασία είναι η απόκτηση πλασμιδίων που φέρουν συγκεκριμένους παράγοντες δίδοντας ένα επιλεκτικό πλεονέκτημα. Η δυναμική συμπεριφορά του πληθυσμού της ομάδας *B. cereus* είναι επίσης εμφανής στο επίπεδο του γονιδιώματος, όπως υποδεικνύεται από το μη συντηρημένο εντοπισμό διαφόρων κινητών στοιχείων. Τέλος θα πρέπει να αναφερθεί ότι η άποψη που αποκομίζουμε για τον πληθυσμό της ομάδας *B. cereus* εξαρτάται σε μεγάλο βαθμό από τα στελέχη που χρησιμοποιούνται στις μελέτες (Tourasse *et al.* 2006).

1.6 Λιποπρωτεΐνες (lipoproteins)

Η ομοιοπολική τροποποίηση μιας πρωτεΐνης με λιπίδιο αναγνωρίστηκε για πρώτη φορά το 1969 σε πρωτεΐνη της εξωτερικής μεμβράνης του *E.coli* (Braun & Rehn 1969), ενώ η ακριβής χημική δομή διευκρινίστηκε το 1973. Η τροποποίηση της σουλφυδριλικής ομάδας μιας αμινοτελικής κυστεΐνης από διακυλογλυκερυλο ομάδα (μέσω σχηματισμού θειοαιθερικού δεσμού) καθώς και η ακυλίωση της κυστεϊνικής αμινομάδας με λιπαρό οξύ αποτέλεσαν μέρος της αρχικής ανακάλυψης από τον Braun το 1973 και γι' αυτό ονομάστηκε λιποπρωτεΐνη (Lpp) του Braun (Braun's lipoprotein, Hantke & Braun 1973). Στη συνέχεια εντοπίστηκαν περισσότερες από 400 πρωτεΐνες με την ίδια τροποποίηση σε όλα τα γνωστά βακτήρια. Αυτές οι λιποπρωτεΐνες είναι είτε δομικές πρωτεΐνες είτε ένζυμα είτε υποδοχείς ή μεταφορείς που εκτελούν βασικές λειτουργίες στη μεμβρανική-υδατική διασύνδεση (Braun & Wu 1993).

Οι βακτηριακές λιποπρωτεΐνες ανήκουν στην κατηγορία πρωτεϊνών που έχουν αγκυρωθεί με λιπίδια (lipid-anchored proteins) και που στα κατά Gram-αρνητικά βακτήρια είναι αγκυροβολημένες τόσο στις κυτταροπλασματικές όσο και στις εξωτερικές μεμβράνες, ενώ στα κατά Gram-θετικά βακτήρια είναι αγκυροβολημένα μόνο στην κυτταροπλασματική μεμβράνη (Nguyen & Götz 2016).

Οι βακτηριακές λιποπρωτεΐνες συντίθενται όπως οποιαδήποτε άλλη φυσιολογική μεμβρανική πρωτεΐνη και φέρουν αλληλουχία σήματος (Inouye *et al.* 1977). Η ανάλυση των αλληλουχιών αυτών σε διαφορετικές λιποπρωτεΐνες αποκάλυψε κοινά χαρακτηριστικά. Το πιο σημαντικό από όλα είναι η παρουσία μιας διακριτής αλληλουχίας στο καρβοξυτελικό άκρο του πεπτιδίου σήματος, που αναφέρεται ως λιποκυτίο (lipobox, Sankaran & Wu 1993, Braun & Wu 1993) (**Εικόνα 1.2**). Αυτό περιέχει μια αμετάβλητη κυστεΐνη Cys που είναι τροποποιημένη με λιπίδια και η -3 θέση από αυτήν καταλαμβάνεται από υδρόφοβο αμινοξύ, συνήθως Leu. Η αλληλουχία είναι [LVI][ASTVI][GAS][C].



Εικόνα 1.2 Η λιπιδική τροποποίηση στις λιποπρωτεΐνες. **(α)** Η σουλφυδρυλική ομάδα της N-τελικής κυστεΐνης (Cys) είναι τροποποιημένη με μία ομάδα διακυλγλυκερυλίου (diacylglyceryl) προσαρτημένη μέσω ενός θειοαιθερικού δεσμού (thioether linkage) και η αμινομάδα είναι ακυλιωμένη με ένα λιπαρό οξύ. **(β)** Η τριμερής κατασκευή της αλληλουχίας σήματος των λιποπρωτεΐνων. Η n -περιοχή (n-region) αποτελείται από πέντε έως επτά αμινοξικά κατάλοιπα και έχει τουλάχιστον δύο θετικά φορτισμένα κατάλοιπα (positively charged region). Η h -περιοχή (h-region) ή υδρόφοβη περιοχή (hydrophobic region), αποτελείται από 7 έως 22 κατά κύριο λόγο υδρόφοβα και μη φορτισμένα κατάλοιπα. Η c-περιοχή (c-region), (στην εικόνα αναγράφεται: lipobox with invariant Cysteine) παρουσιάζει την ακολουθία [LVI][ASTVI][GAS] ακολουθούμενη από μία αυστηρά συντηρημένη κυστεΐνη, που είναι το τροποποιημένο με λιπίδιο κατάλοιπο σε όλες τις βακτηριακές λιποπρωτεΐνες. Η τελευταία φέρει και τη χαρακτηριστική ονομασία λιποκυτίο (lipobox). (πηγή: <https://www.mrc-lmb.cam.ac.uk/genomes/dolop/biosyn.shtml>_ Babu *et al.* 2006)

Ο σημαντικός ρόλος της λιποπρωτεΐνης στα θετικά κατά Gram βακτήρια, ιδιαίτερα στο *phylum* Firmicutes, έχει αναδειχθεί τα τελευταία χρόνια ως βασικός παράγοντας στην ανοσοαπόκριση και την παθογένεια. Φυσιολογικά, η λιποπρωτεΐνη διαδραματίζει σημαντικό

ρόλο στην πρόσληψη θρεπτικών συστατικών και ιόντων, επιτρέποντας στα ιδιαίτερα παθογόνα είδη να επιβιώσουν καλύτερα στον ξενιστή. Επηρεάζουν επίσης λειτουργίες που σχετίζονται με τη σηματοδότηση, τη ρύθμιση της ανοσοαπόκρισης, την εμφάνιση φλεγμονής και την εν γένει παθογένεια του βακτηρίου (Nguyen & Götz 2016). Σε ομάδες κατά Gram-θετικών βακτηρίων που σχηματίζουν ενδοσπόρια, η λιποπρωτεΐνες αποδείχθηκαν ζωτικής σημασίας για τη βλάστηση των σπορίων. Συγκεκριμένα, όσον αφορά στο σχηματισμό σπορίων του *B. anthracis*, η βιογένεση της λιποπρωτεΐνης Igt αναδείχθηκε σημαντική για τη σπορίωση και επομένως επίσης για τη μετάδοση αυτού του παθογόνου (Okugawa *et al.* 2012).

1.6.1 Λιποπρωτεΐνες των *Bacillus cereus* ATCC 14579 και *Bacillus anthracis* Ames

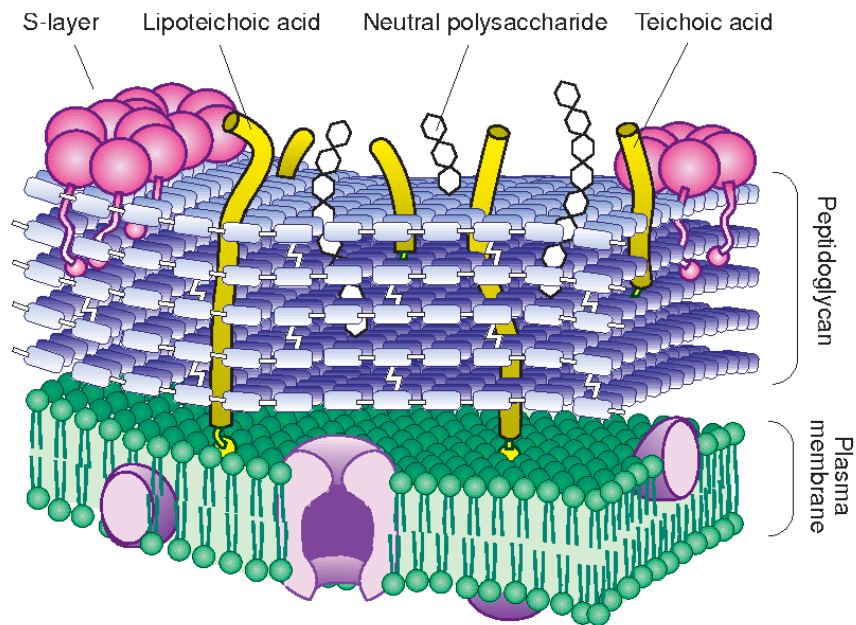
Στο στέλεχος *Bacillus cereus* ATCC 14579 ο αριθμός των προβλεπόμενων λιποπρωτεϊνών σύμφωνα με τη βάση DOLOP (Babu *et al.* 2006) είναι 105 σε σύνολο 5234 προβλεπόμενων ανοιχτών πλαισίων ανάγνωσης (ORFs) ήτοι ποσοστό 2% του γονιδιώματός του.

Στο στέλεχος *Bacillus anthracis* Ames ο αριθμός των προβλεπόμενων λιποπρωτεϊνών σύμφωνα με την ίδια βάση είναι 110 σε σύνολο 5311 προβλεπόμενων ORFs ήτοι ποσοστό 2,07% του γονιδιώματός του.

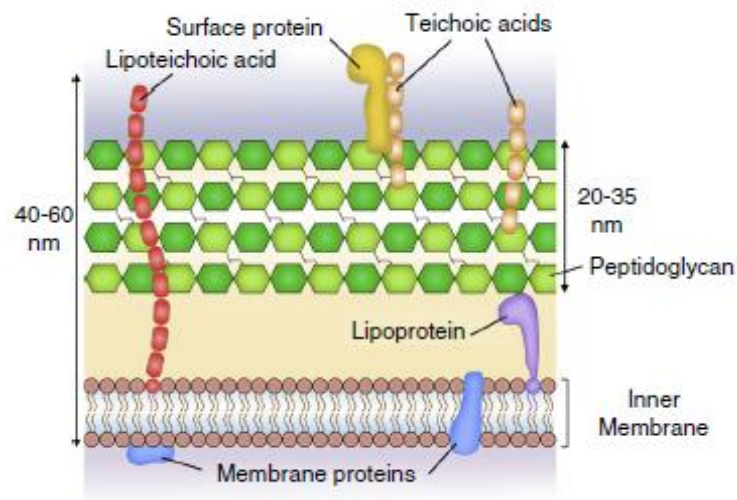
1.7 Κυτταρικό τοίχωμα κατά Gram-θετικών βακτηρίων

Γενικά, τα κυτταρικά τοιχώματα των θετικών κατά Gram βακτηρίων (Εικόνες 1.3 και 1.4) αποτελούνται από ένα παχύ στρώμα πεπτιδογλυκάνης, τειχοϊκά οξέα, λιποτειχοϊκά οξέα, κάψουλα από πολυσακχαρίτες και πρωτεΐνες της επιφάνειας του κυττάρου γνωστές ως S-layer πρωτεΐνες που είναι συχνά γλυκοζυλιωμένες (Messner 1977). Ωστόσο, το κυτταρικό τοίχωμα του *B. anthracis* διαφέρει από πολλές απόψεις από αυτήν τη γενικευμένη περιγραφή. Τα κύτταρα του *B. anthracis* περιβάλλονται από μια κάψουλα πολυ-D-γ-γλουταμικού και όχι μια πολυσακχαριτική κάψουλα. Δεύτερον, τα κυτταρικά τοιχώματα τους δεν περιέχουν τειχοϊκό οξύ (Molnar & Pragai 1971), και τέλος, οι πρωτεΐνες S-στρώματός τους δεν είναι

γλυκοζυλιωμένες (Messner 1977, Mesnage *et al.* 2000). Ωστόσο, συγκρίσεις για τη σύνθεση των γλυκοζυλικών ομάδων των κυτταρικών τοιχωμάτων του *B. anthracis*, *B. cereus* και *Bacillus thuringiensis* δείχνουν ότι περιέχουν γλυκοζυλικά κατάλοιπα και ότι διαφέρουν μεταξύ τους στη γλυκοζυλική τους σύνθεση (Fox *et al.* 2003).



Εικόνα 1.3 Σχέδιο που απεικονίζει το κυτταρικό τοίχωμα ενός κατά Gram-θετικού βακτηρίου. Η διλιπιδική μεμβράνη με ενσωματωμένες πρωτεΐνες καλύπτεται από πολλαπλές στρώσεις πεπτιδογλυκάνης (peptidoglycan), διακοσμημένη από ουδέτερους πολυσακχαρίτες (neutral polysaccharides), λιποτειχοϊκά οξέα (lipoteichoic acids) και τειχοϊκά οξέα (teichoic acids), και περιβάλλεται από το εξωτερικό περίβλημα των πρωτεϊνών της εξώτατης στοιβάδας (S-layer proteins), (για λόγους σαφήνειας, οι πρωτεΐνες που σχετίζονται με το κυτταρικό τοίχωμα δεν απεικονίζονται, πηγή: Delcour *et al.* 1999).



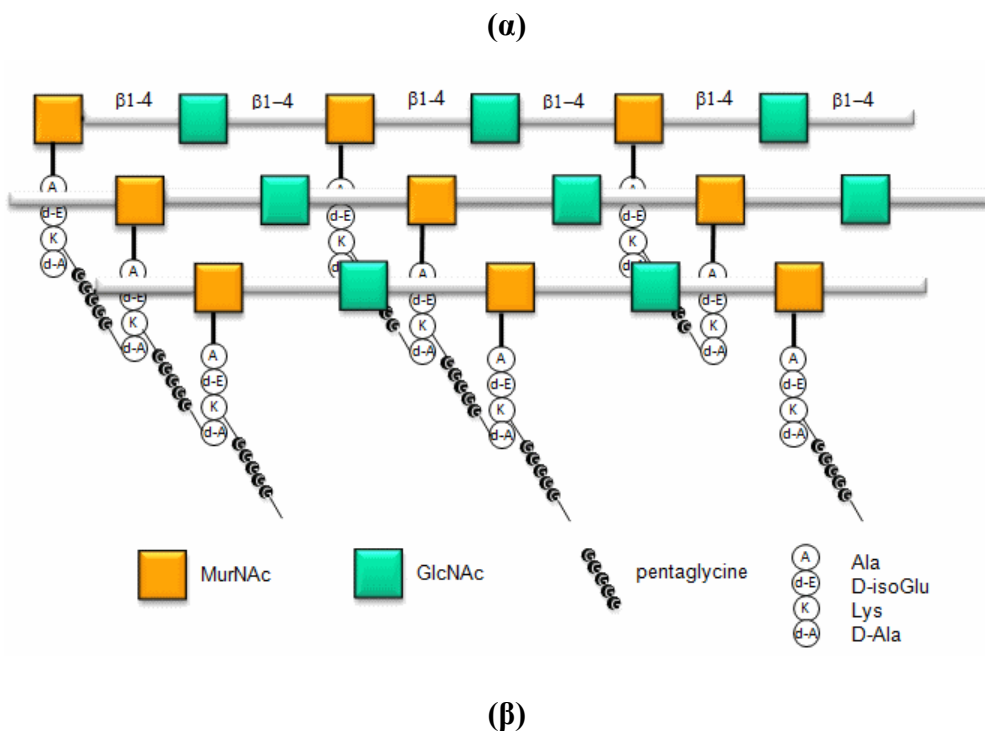
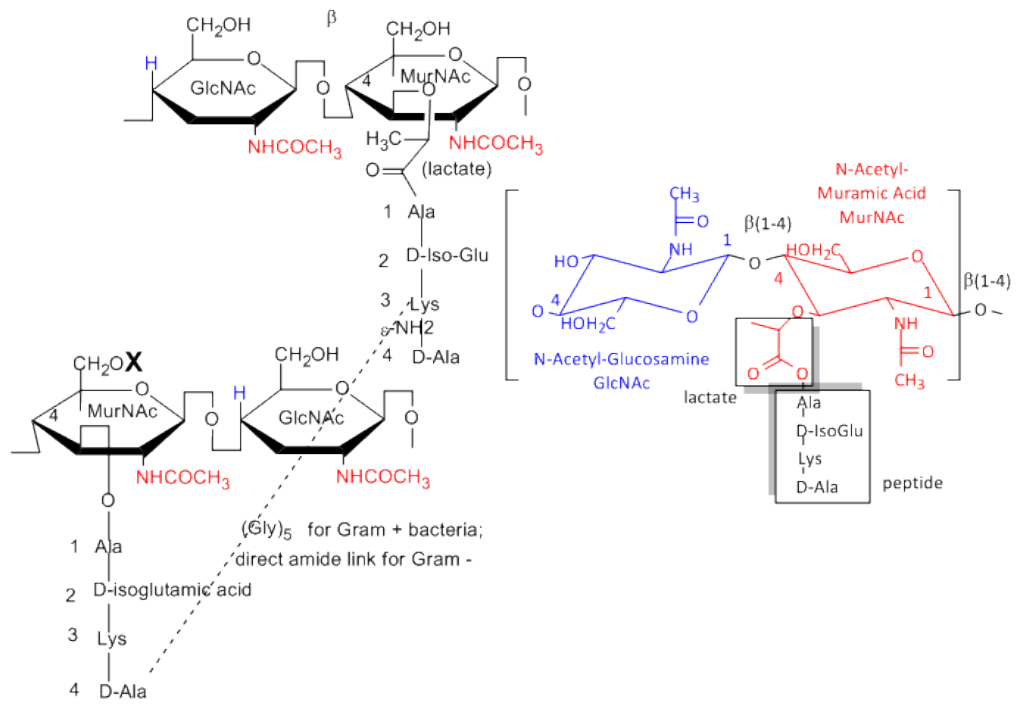
Εικόνα 1.4 Οργάνωση του κυτταρικού τοιχώματος των κατά Gram- θετικών. Η στοιβάδα PG περιλαμβάνει μακρά πολυμερή του επαναλαμβανόμενου δισακχαρίτη *N*-ακετυλογλυκοζαμίνη-*N*-ακετυλομουραμικό οξύ (σκούρο και ανοιχτό πράσινο αντίστοιχα) που συνδέονται μέσω πεπτιδικών γεφυρών. Η λιποπρωτεΐνη του Braun (μωβ) είναι προσαρτημένη στην εξωτερική μεμβράνη και διασυνδέεται ομοιοπολικά με την PG (πηγή: Alcorlo *et al.* 2017).

1.7.1 Η πεπτιδογλυκάνη

Η πεπτιδογλυκάνη (PG) επίσης γνωστή ως μουρεΐνη, βρίσκεται στο κυτταρικό τοίχωμα και περιβάλλει την κυτταροπλασματική μεμβράνη σχεδόν όλων των βακτηρίων. Η πεπτιδογλυκάνη είναι ετερογενές πολυμερές του βακτηριακού κυτταρικού τοιχώματος, η σύνθεση της χημικά περιλαμβάνει μεγάλες αλυσίδες γλυκάνης διασταυρωμένες με μικρά πεπτίδια, δημιουργώντας ένα μεγάλο μακρομόριο. Οι μακρές αλυσίδες γλυκανών της PG αποτελούνται από εναλλασσόμενες υπομονάδες *N*-ακετυλογλυκοζαμίνης (GlcNAc) και *N*-ακετυλομουραμικού οξέος (MurNAc) (van Heijenoort 1998, 2001), που συνδέονται με $\beta 1 \rightarrow 4$ γλυκοζιτικούς δεσμούς (**Εικόνα 1.5**). Το πεπτίδιο δημιουργίας σταυροδεσμών παρουσιάζει μερικά ενδιαφέροντα χαρακτηριστικά, όπως η παρουσία μη πρωτεϊνογενετικών αμινοξέων (π.χ. διαμινοπιμελικό οξύ, λανθιονίνη), εναλλασσόμενα *D*- και *L*-ισομερή και α -συνδεδεμένα μόρια *D*-γλουταμικού οξέος. Το πεπτίδιο του στελέχους συνήθως αποτελείται από *L*-Ala, *D*-Glu, meso-DAP, *D*-Ala (ενώ εμφανίζεται και το πενταπεπτίδιο *L*-Ala, *D*-Glu, meso-DAP, *D*-Ala, *D*-Ala). Τα πεπτίδια διασυνδέονται με απευθείας σύνδεση μεταξύ του meso-DAP και

του D-Ala στη θέση 4. Πολλές φορές αντί για μεσο-διαμινοπιμελικό οξύ (meso-DAP, *meso*-diaminopimelic acid), απαντάται L-lys.

Η πεπτιδογλυκάνη γενικά είναι ένα θεμελιώδες και ειδικό δομικό στοιχείο που βοηθά στην διατήρηση του σχήματος των κυττάρων και τα προστατεύει από την εσωτερική ωσμωτική πίεση. Ως αποτέλεσμα, οποιαδήποτε αναστολή της βιοσύνθεσής της ή η αποικοδόμησή της κατά τη διάρκεια της κυτταρικής ανάπτυξης οδηγεί σε λύση του κυττάρου. Επίσης χρησιμεύει ως ικρίωμα για τη σύνδεση ειδικών πρωτεϊνών και άλλων συστατικών του κυτταρικού τοιχώματος (Johannsen 1993, Vollmer *et al.* 2008, Mengin-Lecreulx *et al.* 2005, Wheeler *et al.* 2014). Η πεπτιδογλυκάνη είναι ένα δυναμικό μακρομόριο σε όλη τη διάρκεια ζωής του κυττάρου. Οι υδρολάσες της πεπτιδογλυκάνης έχουν κρίσιμες λειτουργίες στην ωρίμανση της, στην αύξηση του κυττάρου πριν την κυτταρική διαίρεση, στην αναδιαμόρφωση, στην επιμήκυνση, στον διαχωρισμό, στην ανακύκλωση, καθώς και στην κυτταρική αυτόλυση (Van Heijenoort 2011).

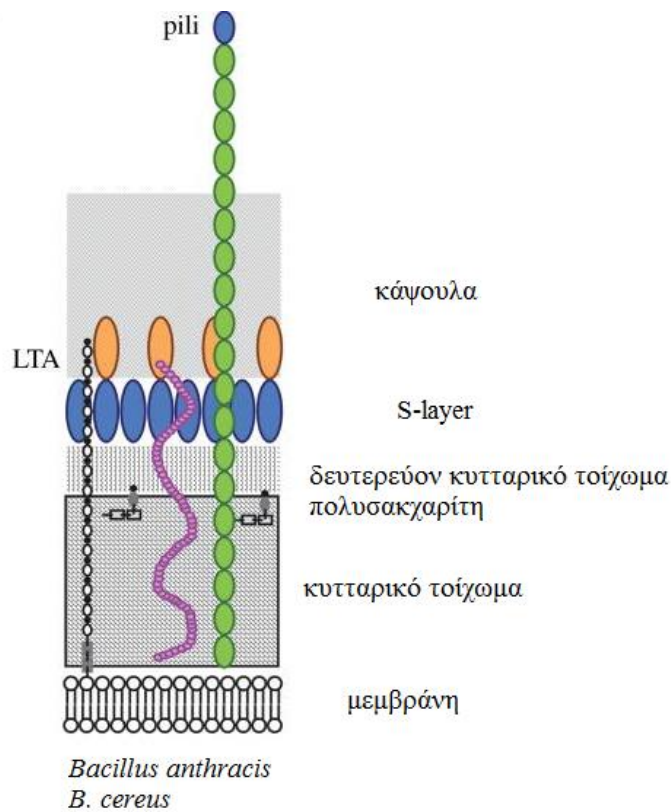


Εικόνα 1.5 (α) Η επαναλαμβανόμενη μονάδα σακχάρων, το τετραπεπτίδιο, και η θέση σύνδεσης των πεπτιδίων. **(β)** Η συγκρότηση της PG, GlcNAc (πράσινο) MurNAc (πορτοκαλί) Ένα τετραπεπτίδιο (Ala-D-isoGlu-Lys-D-Ala) προσαρτάται με αμιδικό δεσμό στην καρβοξυλική ομάδα του γαλακτικού στο MurNAc. Οι αλυσίδες GlcNAc (β 1-4) MurNAc συνδέονται ομοιοπολικά με γέφυρα πενταγλυκίνης διαμέσου της έψιλον αμινομάδας της Lys του τετραπεπτιδίου της μιας αλυσίδας και του D-Ala ενός τετραπεπτιδίου της άλλης αλυσίδας. (πηγή: College of St. Benedict/St. John's University https://bio.libretexts.org/Bookshelves/Biochemistry/The_LibreTexts_libraries/Prof._Henry_Jakubowski)

1.7.2 Συστατικά του κυτταρικού τοιχώματος των *Bacillus anthracis* και *B. cereus*

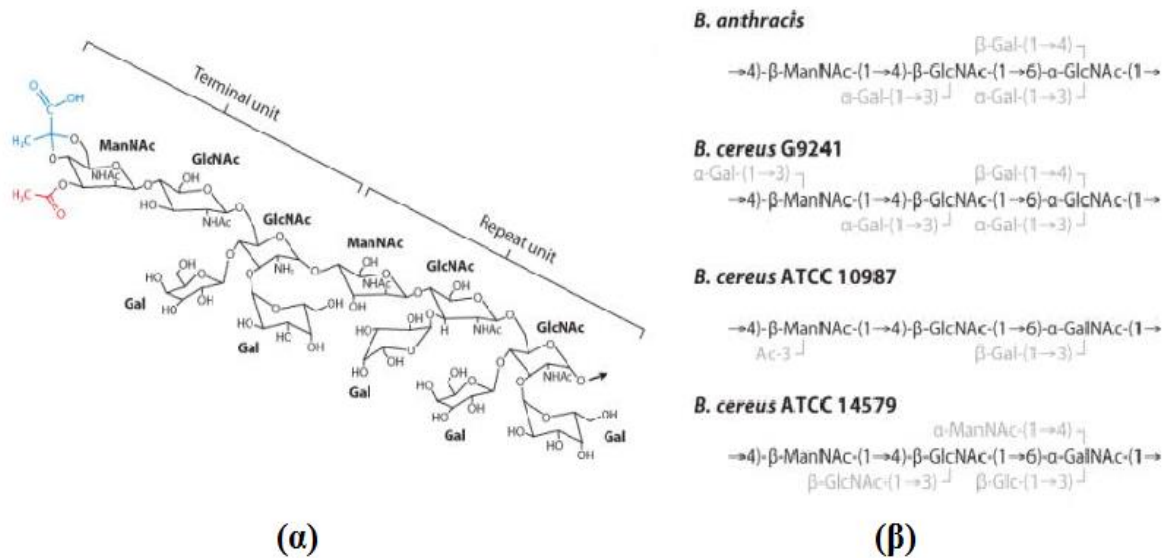
Το κυτταρικό τοίχωμα του *B.anthraxis* αποτελείται από την κυτταροπλασματική μεμβράνη και ένα παχύ στρώμα πεπτιδογλυκάνης, ένα στρώμα S (S-layer, a bi-dimensional crystalline layer,) και την πεπτιδική κάψουλα (Fouet & Mesnage 2002, Zipperle *et al.* 1984, Fouet 2009, Missiakas & Schneewind 2017) (**Εικόνα 1.6**). Το πεπτίδιο του στελέχους αποτελείται από L-Ala, D-Glu, meso-DAP, D-Ala. Οι δευτερογενείς πολυσακχαρίτες του κυτταρικού τοιχώματος (SCWPs Secondary cell wall polysaccharides) συνδέονται με την πεπτιδογλυκάνη και χρησιμεύουν ως συνδέτες για τις δομικές ενότητες SLH (surface layer homology) των πρωτεϊνών του στρώματος S. Ο ουδέτερος πολυσακχαρίτης αποτελείται από γαλακτόζη (Gal), GlcNAc και N-ακετυλομαννοζαμίνη (ManNac). Επομένως, ο σκελετός αμινο-σακχάρου έχει τον επαναλαμβανόμενο τρισακχαρίτη $[\rightarrow 4)\text{-}\beta\text{-ManNAc}\text{-}(1\rightarrow 4)\text{-}\beta\text{-GlcNAc}\text{-}(O3\text{-}\alpha\text{-Gal})\text{-}(1\rightarrow 6)\text{-}\alpha\text{-GlcNAc}\text{-}(O3\text{-}\alpha\text{-Gal}, O4\text{-}\beta\text{-Gal})\text{-}(1\rightarrow)6\text{-}12$ (Zipperle *et al.* 1984, Choudhury *et al.* 2006, Fouet 2009).

Δεδομένου ότι τα κυτταρικά περιβλήματα των περισσότερων από τα θετικά κατά Gram βακτήρια έχουν γλυκοπολυμερή που μπορούν να ποικίλουν σε μεγάλο βαθμό στη σύνθεση και τη δομή μεταξύ των ειδών και ότι διαφορετικά είδη εντός του ίδιου γένους μπορούν να έχουν διαφορετικές ομάδες πολυσακχαριτών (Zipperle *et al.* 1984, Choudhury *et al.* 2006), είναι πολύ πιθανό, στενά συνδεδεμένα στελέχη bacilli να παρουσιάζουν μία ευρεία ποικιλία διαφορετικών συστατικών ολιγοσακχαριτών που απαιτούν διαφορετικά PDAs.

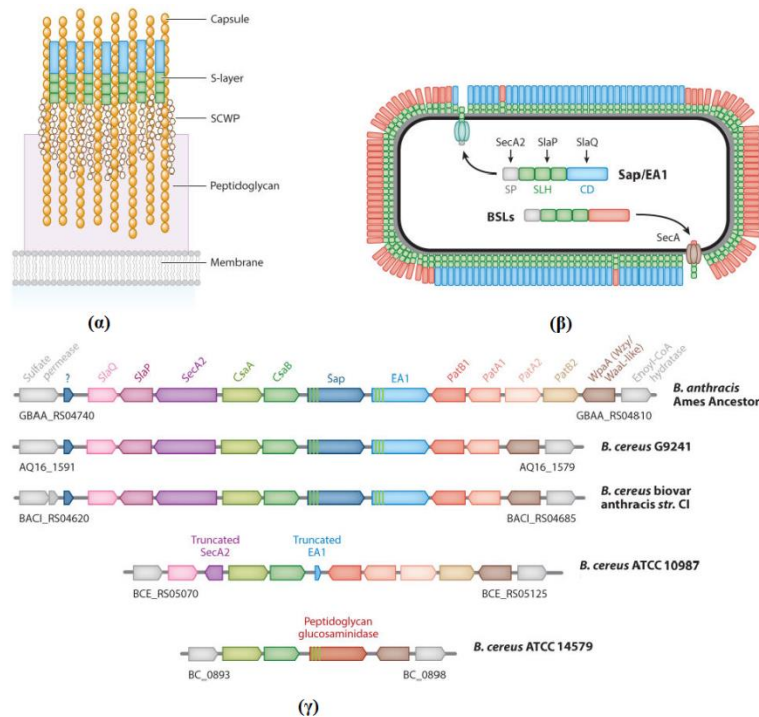


Εικόνα 1.6 Το περίβλημα των κατά Gram-θετικών *Bacillus anthracis* καθώς και του *Bacillus cereus* με την κυτταροπλασματική μεμβράνη και ένα παχύ στρώμα πεπτιδογλυκάνης. Οι πολυσακχαρίτες του δευτερογενούς κυτταρικού τοιχώματος (SCWPs) συνδέονται με την πεπτιδογλυκάνη και χρησιμεύουν ως συνδέτες για τις περιοχές SLH (surface layer homology) των πρωτεϊνών S-layer στρώματος. Εκτός από τις αγκυρωμένες πρωτεΐνες του κυτταρικού τοιχώματος, η πεπτιδογλυκάνη των βακίλων λειτουργεί επίσης ως σημείο αγκυροβόλησης για την κάψουλα και τα πρωτεϊνικά νημάτια σε μικροβιακές επιφάνειες, (που ονομάζονται pili ή κροσσάρια), παρέχουν συγκολλητικές λειτουργίες και παίζουν σημαντικούς ρόλους κατά τη μολυσματικότητα (Pichichero *et al.* 1982, Hultgren *et al.* 1993, Thelin & Taylor 1996). (πηγή: Schneewind & Missiakas 2012)

Υπάρχουν σημαντικές διαφορές μεταξύ των στελεχών του *B. cereus sensu lato* στη σύνθεση των συστατικών του δευτερογενούς κυτταρικού τοιχώματος (Εικόνες 1.7 και 1.8).

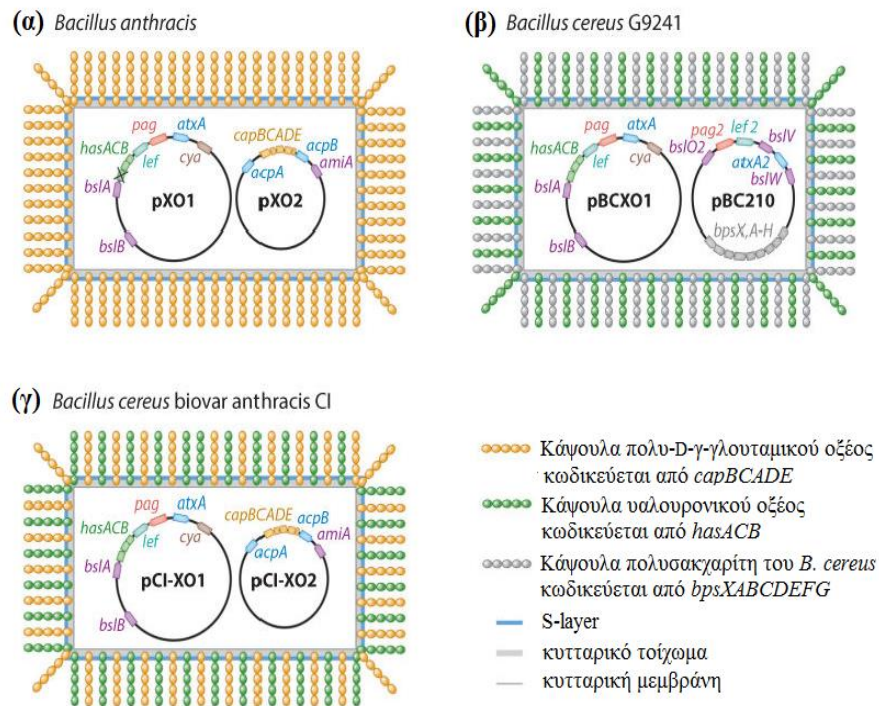


Εικόνα 1.7 Δομή του πολυσακχαρίτη του δευτερογενούς κυτταρικού τοιχώματος (SCWP) στελεχών *B. cereus sensu lato*. (α) Δομή της τελικής μονάδας και της μονάδας που επαναλαμβάνεται του SCWP του *B. anthracis*. Η τροποποίηση της πυροβουλ-κεταλικής ομάδας του ManNAc στο μη αναγωγικό απομακρυσμένο άκρο του SCWP παρουσιάζεται με χρώμα κυανό. Η ομάδα ακετυλίου του προτελευταίου κατάλοιπου GlcNAc στην τελική μονάδα χρωματίζεται κόκκινη. Το βέλος υποδεικνύει το γλυκοζιδικό δεσμό στην επόμενη επαναλαμβανόμενη μονάδα. (β) Δομή επαναλαμβανόμενης μονάδας του SCWP από *B. anthracis*, *B. cereus* G9241, *B. cereus* ATCC 10987 και *B. cereus* ATCC 14579. (πηγή: Missiakas & Schneewind 2017).



Εικόνα 1.8 (α) Μοντέλο που απεικονίζει τα διαφορετικά στρώματα, συμπεριλαμβανομένης της μεμβράνης, της πεπτιδογλυκάνης, του πολυσακχαρίτη του δευτερογενούς κυτταρικού τοιχώματος (SCWP, secondary cell wall polysaccharide) και της S-στοιβάδας (S-layer). Τα στρώματα SCWP και αυτό της κάψουλας του πολυ-D-γ-γλουταμικού οξέος συνδέονται με την πεπτιδογλυκάνη του *B. anthracis*. **(β)** Έκκριση και συναρμολόγηση των πρωτεϊνών της S-στοιβάδας (SLPs, S-layer proteins), των Sap και EA1 και των πρωτεϊνών που συνδέονται με την S-στοιβάδα (S-layer) του *B. anthracis* (BSLs). Κάθε πρωτεΐνη συντίθεται ως πρόδρομος με ένα αμινοτελικό πεπτίδιο σήματος (SP), τρεις περιοχές ομολογίας S-layer (SLH, S-layer homology) και μία καρβοξυτελική περιοχή. Το SecA2 και το SlaP συμβάλλουν στην έκκριση των Sap και EA1, ενώ το SlaQ συμβάλλει στη συναρμολόγηση των Sap και EA1 του S-layer. **(γ)** συναρμολόγηση των γονιδίων της S-στοιβάδας (*slc*) στον *B. anthracis* και άλλα *B. cereus sensu lato* στελέχη που έχουν απομονωθεί και μελετηθεί. Τα γονίδια *csaA* και *csaB* (κεταλική πυροϋβλομεταφοράση) και *wpaA* [Wzy- / WaaL (O-αντιγονική λιγάση)-like] είναι απολύτως συντηρημένα σε όλα τα είδη και απαιτούνται για τροποποιήσεις του SCWP που επιτρέπουν τη συναρμολόγηση SLP ή BSL. Τα στελέχη που προκαλούν τη νόσο του άνθρακα (*B. anthracis*) ή νόσο που μοιάζει με άνθρακα (*B. cereus* G9241 και *B. cereus* biovar anthracis CI) φέρουν γονίδια SLPs (Sap και EA1 με τρεις περιοχές SLH), για έκκριση και συναρμολόγηση SLP (slaQ, slaP και secA2) και ακετυλίωση SCWP (patA / patB). Το *slc* του *B. cereus* ATCC 10987 φιλοξενεί κολοβά γονίδια *eag* και *secA2* και στερείται των *sap* και *slaP*. Το στέλεχος *B. cereus* ATCC 14579 στερείται γονιδίων για SLPs, έκκριση και συναρμολόγηση SLP και ακετυλίωση SCWP αλλά κωδικεύει την υδρολάση της μουρείνης με τρεις περιοχές SLH στη θέση Sap και EA1. Τα γονίδια που συναρμολογούν το *slc* χρωματίζονται τυχαία με εξαίρεση τα γονίδια που το πλαισιώνουν αλλά δεν είναι μέρος του *slc* και τα οποία εμφανίζονται με γκρι χρώμα. (πηγή: Missiakas & Schneewind 2017)

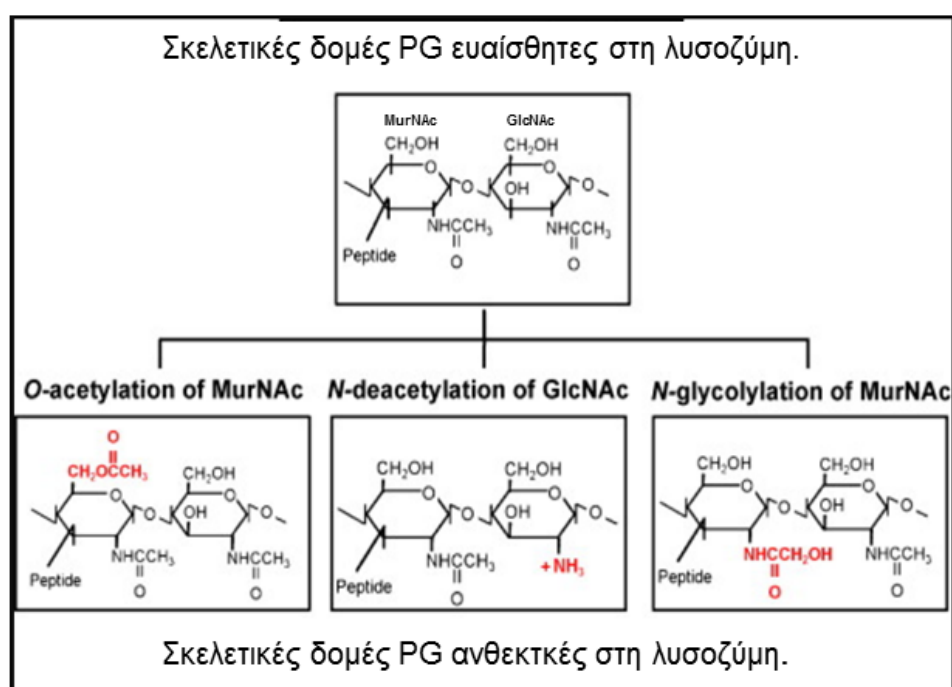
Στα στελέχη *B.cereus sensu lato* που προκαλούν τη νόσο του άνθρακα ή νόσους που μοιάζουν με τη νόσο του άνθρακα παρατηρούνται διαφορές και στην σύνθεση της κάψουλας όπως για παράδειγμα απεικονίζεται στην **εικόνα 1.9**.



Εικόνα 1.9 Απεικόνιση στελεχών *B.cereus sensu lato* που προκαλούν τη νόσο του άνθρακα ή νόσους που μοιάζουν με τη νόσο του άνθρακα: **(α)** *B. anthracis* με τα δύο πλασμίδια μολυσματικότητας, pXO1 και pXO2. **(β)** *B. cereus* G9241, που φέρει τα πλασμίδια pBCXO1 και pBC210, και **(γ)** *B. cereus* biovar anthracis CI με τα πλασμίδια pCI-XO1 και pCI-XO2. Τα στελέχη *B. cereus* G9241 και το biovar anthracis CI έχουν ένα τρίτο μικρό πλασμίδιο που δεν παρουσιάζεται εδώ. Επιπροσθέτως των γονιδίων της τοξίνης (*pag*, *lef* και *cya*), τα πλασμίδια *B. anthracis* pXO1 και pXO2 φέρουν γονίδια για τη σύνθεση της κάψουλας πολυ-D-γ-γλουταμικού οξέος (PDGA) (*capBCADE*) και τη σύνθεση της κάψουλας ναλουρονικού οξέος (HA) (*hasACB*). Ωστόσο, το γονίδιο *hasA*, φέρει μια μετάλλαξη απενεργοποίησης. Τα πλασμίδια φέρουν επίσης τρία γονίδια για πρωτεΐνες σχετιζόμενες με το μη συνεχές έλυτρο της εξώτατης στοιβάδας (S-layer) του *B. anthracis* (BSL): *bslA*, *bslB* και *amiA*. Η γονιδιακή έκφραση της μολυσματικότητας ρυθμίζεται από το προϊόν του *atxA*, το οποίο ενεργοποιεί την έκφραση των γονιδίων της τοξίνης, *bslA* και *capBCADE* ως απόκριση σε ένα σήμα από τον ξενιστή (αυξημένη συγκέντρωση CO₂). Το οπερόνιο του *capBCADE* ρυθμίζεται και από τα δύο *acpA* και *acpB*. Το pBC210 φέρει γονίδια τριών BSLs (*bslO2*, *bslV* και *bslW*), της τοξίνης Certhrax (*pag2*, *lef2*), ενός ρυθμιστή τύπου *atxA* (*atxA2*) και του εξω-πολυσακχαρίτη BPS (από το γονίδιο *bpsXABCDEFGHI*). Η κάψουλα του *B. cereus* biovar anthracis αποτελείται τόσο από PDGA όσο και από HA. (πηγή: Missiakas & Schneewind 2017)

1.8 Τροποποίηση των αλυσίδων γλυκανών της πεπτιδογλυκάνης των βακτηρίων

Ο κυρίαρχος μηχανισμός για να αποφύγει το βακτήριο το ανοσοποιητικό σύστημα του ξενιστή που χρησιμοποιείται από μεγάλο αριθμό θετικών κατά Gram και αρνητικών κατά Gram βακτηρίων είναι η μοριακή απόκρυψη: η τροποποίηση της αλυσίδας της PG (Psylinakis *et al.* 2005, Moynihan *et al.* 2014). Τρεις τροποποιήσεις έχουν περιγραφεί εκτενώς και περιλαμβάνουν *N*-γλυκολυλίωση, *O*-ακετυλίωση και *N*-απακετυλίωση (Vollmer 2008) (Εικόνα 1.10).



Εικόνα 1.10 Σκελετικές δομές PG ευαίσθητες και ανθεκτικές στη λυσοζύμη. Παθογόνα βακτήρια τροποποιούν την PG τους με *O*-ακετυλίωση (*O*-acetylation) του MurNAc στη θέση C-6, *N*-απακετυλίωση (*N*-deacetylation) του GlcNAc στη θέση C-2 και *N*-γλυκολυλίωση (*N*-glycolylation) του μουραμικού οξέως στη θέση C-2. Οι τροποποιήσεις προσδίδουν αντοχή κατά της λυσοζύμης που εκκρίνεται από τον ξενιστή κατά τη διάρκεια της μόλυνσης. (πηγή: Sukhithasri *et al.* 2013)

1.8.1 Απακετυλίωση

Η ενζυμική απακετυλίωση της πεπτιδογλυκάνης είναι γνωστή από το 1971 όταν ανακαλύφθηκε στα κυτταρικά τοιχώματα του *B. cereus* (Araki *et al.* 1971).

Η μετασυνθετική τροποποίηση της PG αντιπροσωπεύει μια σημαντική στρατηγική για τα παθογόνα βακτήρια με σκοπό να αποφύγουν την έμφυτη ανοσία και να ελέγξουν την αυτόλυση (Boneca *et al.* 2007). Η *N*-απακετυλίωση των καταλοίπων GlcNAc και MurNAc της PG συσχετίζεται με την ανθεκτικότητα του βακτηρίου στην λυσοζύμη του ξενιστή, τις επιδράσεις στην αναγνώριση των παθογόνων μέσω διαφόρων υποδοχέων του ξενιστή και την εισαγωγή επιπλέον θετικού φορτίου στο κυτταρικό τοίχωμα το οποίο μπορεί να αυξήσει την αντίσταση του βακτηρίου σε κατιονικά αντιμικροβιακά πεπτίδια (Vollmer 2008).

Η *N*-απακετυλίωση αυτών των σακχαριτών, η οποία επιτρέπει την αποικοδόμηση και την αναδιαμόρφωση, είναι μία από τις αμυντικές στρατηγικές των βακτηρίων έναντι των ενζύμων των ξενιστών αφού η τροποποιημένη πεπτιδογλυκάνη δεν αναγνωρίζεται πλέον (Boneca 2005, Popowska *et al.* 2009, Meyrand *et al.* 2007, Hebert *et al.* 2007, Kaoukab-Raji *et al.* 2012) Έτσι, η αναστολή των απακετυλασών πεπτιδογλυκάνης μπορεί να είναι ένας τρόπος πρόληψης της μολυσματικής διάδοσης.

Οι τροποποιήσεις της ραχοκοκαλιάς των αλυσίδων γλυκανών περιορίζονται γενικά στις C-6-υδροξυλ- και C-2 αμινομάδες, με τις συνηθέστερες να είναι η ακετυλίωση και η απακετυλίωση (Clarke 2014). Οι απακετυλάσες της πεπτιδογλυκάνης κάνουν *N*-απακετυλίωση σε κατάλοιπα GlcNAc ή MurNAc της πεπτιδογλυκάνης των βακτηρίων (Fukushima *et al.* 2002, Vollmer & Tomasz 2000) Συγκεκριμένα, οι απακετυλάσες της πεπτιδογλυκάνης υδρολύουν τον αμιδικό δεσμό και αφαιρούν την ακετυλομάδα από το άζωτο στη θέση 2 του άνθρακα των καταλοίπων GlcNAc ή MurNAc (Boneca 2005, Vollmer & Tomasz 2000, Vollmer & Tomasz 2002).

Η *N*-απακετυλίωση της στιβάδας PG παρατηρήθηκε στο ανθεκτικό στη λυσοζύμη *Streptococcus pneumoniae* όπου και ταυτοποιήθηκε το πρώτο γονίδιο απακετυλάσης GlcNAc που ονομάστηκε *PgdA* (Vollmer & Tomasz 2000). Οι Fukushima *et al.* ανέφεραν ότι η απακετυλάση PdaA του *B. subtilis* (*BsPdaA*) μπορεί να κάνει απακετυλίωση *in vitro* καταλοίπων MurNAc καθαρισμένων αλυσίδων γλυκανών της πεπτιδογλυκάνης του *B. subtilis* (Fukushima *et al.* 2005).

1.9. Τα ένζυμα της βάσης δεδομένων CAZy

Τα ένζυμα που διασπούν, τροποποιούν ή σχηματίζουν γλυκοζιδικούς δεσμούς έχουν κατηγοριοποιηθεί σε οικογένειες με βάση ομοιότητες στην αμινοξική τους αλληλουχία, σε μία διαρκώς αναπτυσσόμενη δομημένη βάση δεδομένων που περιλαμβάνει και πληροφορίες δομής και λειτουργίας, την CAZy (Carbohydrate-Active enzymes database <http://www.cazy.org/>, Lombard *et al.* 2014). Η CAZy σήμερα (Δεκέμβριος 2018) περιέχει 162 οικογένειες γλυκοζιδικών υδρολασών (Glycoside Hydrolases, GHs), 106 οικογένειες γλυκοζυλ-μεταφορασών (Glycosyl Transferases, GTs), 36 λυασών πολυσακχαριτών (Polysaccharide Lyases, PLs), 16 οικογένειες υδατανθρακικών εστερασών (Carbohydrate Esterases, CEs), και 16 οικογένειες που κατηγοριοποιούν ένζυμα οξειδοαναγωγής που δρουν σε συνδυασμό με άλλα ένζυμα της βάσης ως ένζυμα βοηθητικών δραστηριοτήτων (Auxiliary Activities, AA). Ακόμη στην ταξινόμηση περιλαμβάνονται και 84 οικογένειες στις οποίες εντάσσονται ενότητες πρόσδεσης υδατανθράκων (Carbohydrate-Binding Modules, CBMs), οι οποίες δεν παρουσιάζουν καταλυτική δράση.

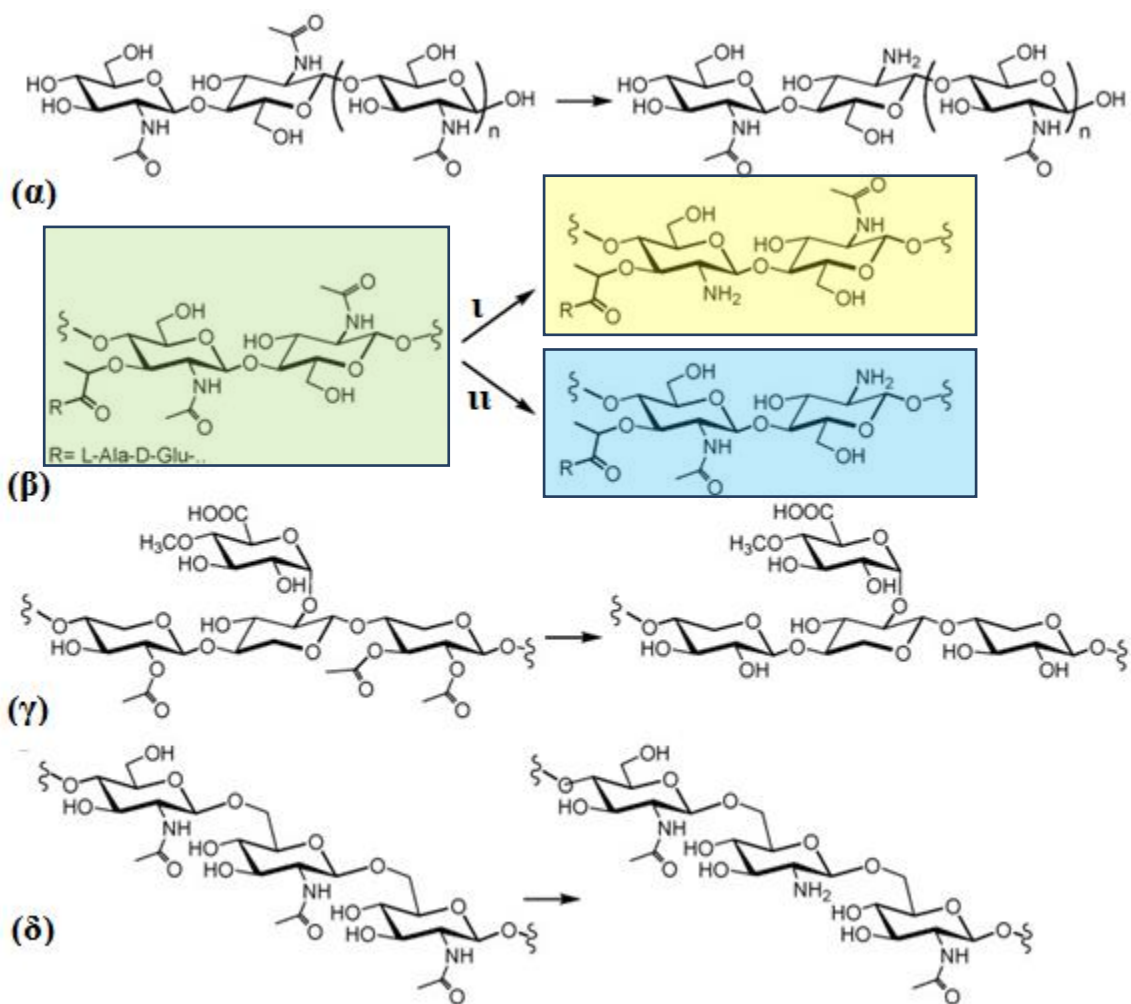
1.9.1 Η οικογένεια 4 υδατανθρακικών εστερασών της βάσης δεδομένων CAZy.

Οι υδατανθρακικές εστεράσες (CEs), όπως αναφέρθηκε, ταξινομούνται επί του παρόντος σε 16 οικογένειες οι οποίες αριθμούνται από την CE1 έως CE16. Ωστόσο, επειδή τα περισσότερα από τα μέλη της CE10, φαίνεται να είναι εστεράσες δραστικές έναντι μη υδατανθρακικών υποστρωμάτων, ο συνολικός αριθμός των οικογενειών CE περιορίζεται σε 15.

Η οικογένεια 4 των υδατανθρακικών εστερασών (Carbohydrate Esterase family 4, CE4) είναι μακράν η μεγαλύτερη οικογένεια CE που σήμερα απαριθμεί 23.081 ακολουθίες. Σχεδόν όλες οι ακολουθίες αυτής της οικογένειας (22.485) προέρχονται από βακτήρια. Τα υπόλοιπα μέλη προέρχονται από Ευκαρυωτικά (537), Αρχαία (17) ενώ περιλαμβάνονται και μη ταξινομημένα ένζυμα (41).

Στην CE4 ανήκουν οι απακετυλάσες πολυσακχαριτών που περιλαμβάνουν απακετυλάσες χιτίνης (chitin deacetylases, EC 3.5.1.41), απακετυλάσες χιτοολιγοσακχαριτών (chitooligosaccharide deacetylases, EC 3.5.1.-), εστεράσες ακετυλοξυλάνης (acetylxyylan esterases, EC 3.1.1.72), απακετυλάσες N-ακετυλογλυκοζαμίνης της πεπτιδογλυκάνης

(peptidoglycan N-acetylglucosamine deacetylases, EC 3.5.1.128), και απακετυλάσες N-ακετυλομουραμικού οξέος της πεπτιδογλυκάνης (peptidoglycan N-acetylmuramic acid deacetylases EC 3.5.1.-). Οι CEs είναι ένζυμα που καταλύουν την απακετυλίωση *O*- ή *N*-ακετυλιωμένων ομάδων γλυκανών (Vollmer & Tomasz 2000, Psylinakis *et al.* 2005, Caufrier *et al.* 2003, Tsigos *et al.* 2000, Kafetzopoulos *et al.* 1993 Moynihan *et al.* 2014) Στην **εικόνα 1.11** δίδονται παραδείγματα υποστρωμάτων των CE4 ενζύμων και αντιπροσωπευτικών απακετυλιωμένων προϊόντων.



Εικόνα 1.11 Σκελετικές δομές υποστρωμάτων των CE4 ενζύμων και αντιπροσωπευτικών απακετυλιωμένων προϊόντων. (α) ολιγοσακχαρίτης χιτίνης υπόστρωμα απακετυλασών της χιτίνης, (β) θραύσμα πεπτιδογλυκάνης υπόστρωμα (i) απακετυλάσης του MurNAc της πεπτιδογλυκάνης (ii) απακετυλάσης του GlcNAc της πεπτιδογλυκάνης. (γ) Acetyl-d-glucurono-d-xylan υπόστρωμα εστερασών ακετυλοξυλάνης. (δ) β-1,6-γλυκάνη υπόστρωμα απακετυλασών πολυ-β-1,6-N-ακετυλογλυκοζαμίνης (πηγή: Aragunde *et al.* 2018).

Μόνο το 0,3% περίπου των ενζύμων που ανήκουν στην οικογένεια CE4 έχουν χαρακτηριστεί βιοχημικά (60 ένζυμα), ενώ η τρισδιάστατη δομή έχει προσδιορισθεί μόλις για ποσοστό περίπου 0,1% επί του συνόλου, ήτοι σύμφωνα με την CAZy 25 καταχωρημένες δομές υπάρχουν στην βάση δομικών δεδομένων πρωτεϊνών (PDB, Protein Data Bank), στις οποίες συγκαταλέγονται και οι δομές που προσδιορίστηκαν κατά την παρούσα έρευνα.

1.9.2 Δομικές μελέτες των μελών της οικογένειας CE4

Η πρώτη τρισδιάστατη δομή μέλους της CE4 που προσδιορίστηκε κρυσταλλογραφικά ήταν η δομή της *BsPdaA* απακετυλάσης του MurNAc της πεπτιδογλυκάνης του *Bacillus subtilis* [με κωδικό καταχώρησης στην PDB (PDB ID) 1W17, Blair & Van Aalten 2004] (Εικόνα 1.12). Ακολούθησε η *SpPgdA* απακετυλάση της GlcNAc της πεπτιδογλυκάνης του *Streptococcus pneumonia* (PDB ID: 2C1G, Blair *et al.* 2005). Η δομή της *SpPgdA* επιλύθηκε σε διακριτικότητα 1,75Å και ήταν αυτή που αποκάλυψε την εξάρτηση της ενζυμικής δραστηρότητάς της από μεταλλικό ιόν, γεγονός που επαληθεύτηκε για την πλειονότητα των δομών των μελών της CE4 που ακολούθησαν, χαρακτηρίζοντας τα ως μεταλλοένζυμα, τα οποία εμφανίζουν ένα χαρακτηριστικό συντηρημένο καταλυτικό κέντρο αποτελούμενο από μια τριάδα συναρμογής του μετάλλου, His-His-Asp, και ο καταλυτικός μηχανισμός τους θεωρείται τύπου οξέος/βάσης (Εικόνα 1.13). Παρόμοια με τις υδρολάσες ψευδαργύρου, τα ένζυμα της CE4 περιέχουν ένα μόνο δισθενές μεταλλικό ιόν, συνήθως ψευδαργύρου (Zn^{2+}) ή κοβαλτίου (Co^{2+}), στο ενεργό κέντρο (Hernick & Fierke 2005). Τα κατάλοιπα που εμπλέκονται στη θέση πρόσδεσης των ενζύμων των μελών της CE4 ανήκουν σε πέντε διατηρημένα μοτίβα (MT1-MT5) αλληλουχίας (Blair *et al.* 2005). Αυτά τα μοτίβα παίζουν επίσης σημαντικούς ρόλους στην αναγνώριση του υποστρώματος.

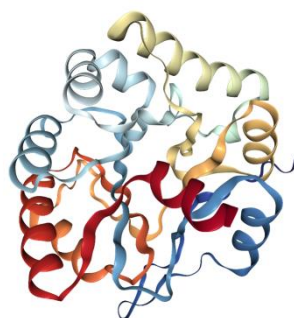
Η στοίχιση αλληλουχιών μελών της οικογένειας CE4 έδειξε ότι ο καταλυτικός πυρήνας εμφανίζεται συντηρημένος και φέρει την ονομασία NodB ομόλογη δομική ενότητα (NodB homology domain, Caufrier *et al.* 2003, Kafetzopoulos *et al.* 1993), από το πρώτο ένζυμο CE4 που χαρακτηρίστηκε, την NodB απακετυλάση χιτοολιγοσακχαρίτη, από το *Rhizobium meliloti* (John *et al.* 1993). Η καταλυτική δομική ενότητα NodB αναδιπλώνεται σε μία δομή α/β βαρελιού που αποτελείται από 6, 7 ή 8 παράλληλους β-κλώνους. Η καταλυτική

δομική ενότητα NodB εμφανίζεται στην βάση δεδομένων Pfam, [https://pfam.xfam.org_με_κωδικό PF01522](https://pfam.xfam.org_με_κωδικό_PF01522).

Έκτοτε οι δημοσιευμένες δομές περιλαμβάνουν: την απακετυλάση της GlcNAc της πεπτιδογλυκάνης SmPgdA του *Streptococcus mutans* (PDB ID: 2W3Z, Deng *et al.* 2009). Απακετυλάση της poly-β-1,6-*N*-acetyl-D-glucosamine όπως για παράδειγμα οι IcaB *Ammonifex degensii* (PDB ID: 4WCJ, Little *et al.* 2014), BpsB (PDB ID: 5BU6) από το *Bordetella bronchiseptica* (Little *et al.* 2015) και PgaB (PDB IDs: 3VUS, 4F9D και 4F9J) από το *Escherichia coli* (Little *et al.* 2012, Nishiyama *et al.* 2013).

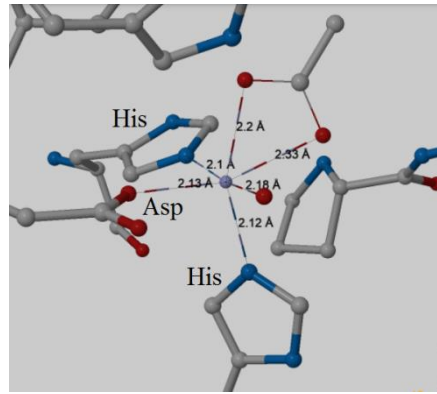
Η χιτίνη, η χιτοζάνη και οι χιτοολιγосακχαρίτες είναι πολυσακχαρίτες από μονάδες GlcNAc, οι οποίες μπορούν να απακετυλιωθούν από CE4 απακετυλάσες της χιτίνης (CDAs - EC 3.5.1.41). Δομές αυτών των ενζύμων είναι διαθέσιμες όπως η CICDA από το ευκαρυωτικό *Colletotrichum lindemuthianum* (PDB ID: 2IW0, Blair *et al.* 2006) και η VcCDA από το *Vibrio cholera* (PDB ID: 4NY2, Andrés *et al.* 2014). Ακόμη η δομή ECU11_0510 της φερόμενης απακετυλάσης πολυσακχαρίτη από το *Encephalitozoon cuniculi* (PDB ID: 2VYO, Urch *et al.* 2009).

Τέλος, οι διαθέσιμες δομές CE4 περιλαμβάνουν επίσης δύο AXEs με κάποιες δομικές διαφορές (EC3.1.1.72) του *Clostridium thermocellum* CtCE4 και του *Streptomyces lividans* SICE4 (PDB IDs: 2C71 και 2CC0, Taylor *et al.* 2006).



BsPdaA (1W17, 2005)

Εικόνα 1.12 Η πρώτη τρισδιάστατη δομή μέλους της CE4 που προσδιορίστηκε ήταν η δομή της BsPdaA του *B. subtilis* (Blair & Van Aalten 2004). Στην παρένθεση δίδεται ο κωδικός καταχώρησης και το έτος διάθεσης της δομής από την PDB.

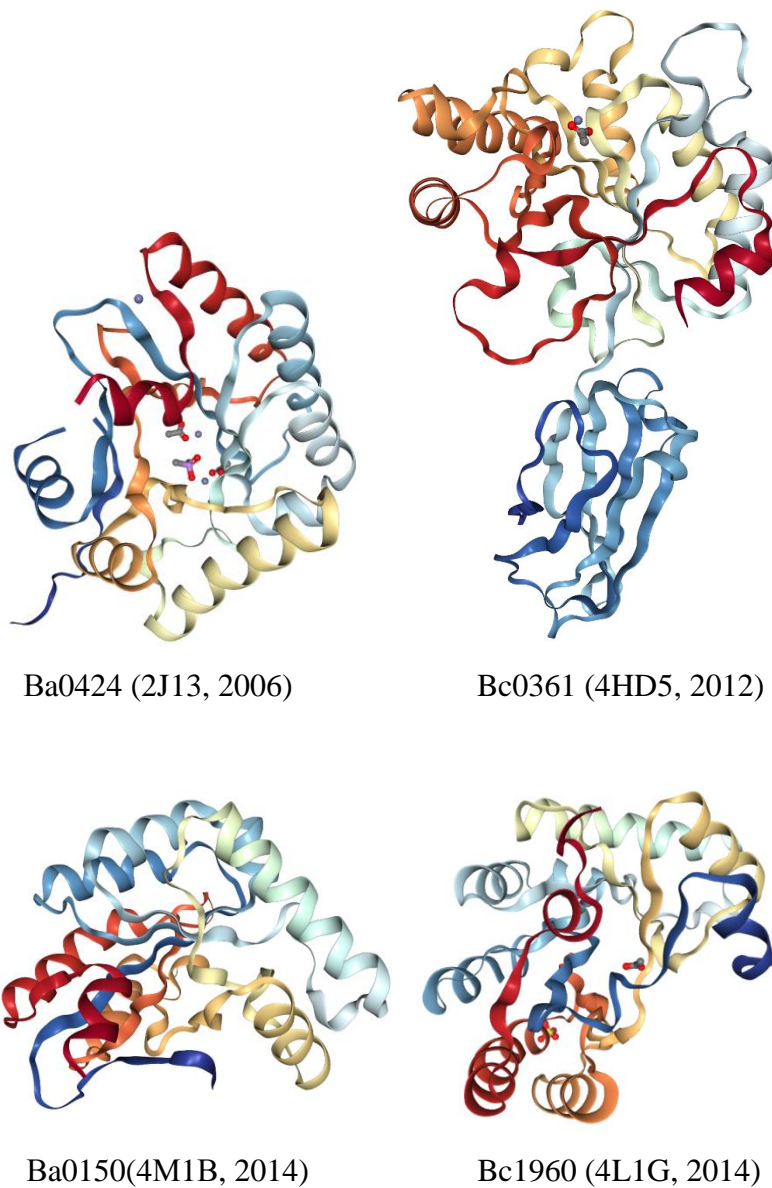


Εικόνα 1.13 Απεικόνιση της γεωμετρίας συναρμογής του μεταλλοϊόντος στο ενεργό κέντρο. Τα περισσότερα από τα μέλη CE4 έχουν χαρακτηριστεί ως μεταλλοένζυμα. Στην εικόνα παρουσιάζεται η συντηρημένη καταλυτική αμινοξική τριάδα Asp-His-His.

1.9.2.1 Δομικές μελέτες απακετυλασών πολυσακχαρίτη από την ομάδα του *B. cereus*

Οι δομές πρωτεϊνών που έχουν προσδιοριστεί με κρυσταλλογραφία ακτίνων $-X$ από τον *Bacillus anthracis* (**Εικόνα 1.14**) είναι αυτή της Ba0150 (με κωδικό καταχώρησης στην PDB: 4M1B, Strunk *et al.* 2014). Η κρυσταλλική δομή δεν περιέχει καταλυτικό μεταλλικό ιόν και η πρωτεΐνη δεν περιέχει τη συντηρημένη τριάδα δέσμευσης μετάλλου Asp-His-His που βρίσκεται στα περισσότερα μέλη της CE4, υποδεικνύοντας ότι δεν είναι μια λειτουργική απακετυλάση πολυσακχαρίτη, όμως δεδομένου ότι περιέχει τα συντηρημένα καταλυτικά Asp και His κατάλοιπα, η Ba0150 θεωρείται ότι μπορεί να έχει κάποια άλλη υδρολυτική δραστηριότητα (Strunk *et al.* 2014). Ακόμη αυτή της Ba0424 (με κωδικό καταχώρησης στην PDB: 2J13, Oberbarnscheidt *et al.* 2007) η οποία παρουσιάζει ομοιότητα με την BsPdaA απακετυλάση του MurNAc της πεπτιδογλυκάνης του *Bacillus subtilis*, υποδηλώνοντας έτσι ένα παρόμοιο υπόστρωμα εξειδίκευσης με τα κατάλοιπα MurNAc της πολυμερούς πεπτιδογλυκάνης (Fukushima *et al.* 2005).

Οι δομές που έχουν επιλυθεί με κρυσταλλογραφία ακτίνων $-X$ από τον *Bacillus cereus* (**Εικόνα 1.14**) είναι αυτή της Bc1960 (με κωδικό καταχώρησης στην PDB: 4L1G, Fadouloglou *et al.* 2017), που θεωρείται ως απακετυλάση GlcNAc της πεπτιδογλυκάνης (Balomenou *et al.* 2013). Ωστόσο, η κρυσταλλική δομή δεν περιέχει μεταλλοϊόν στο ενεργό κέντρο. Επίσης της φερόμενης απακετυλάσης πολυσακχαρίτη Bc0361 (με κωδικό καταχώρησης στην PDB:4HD5, Fadouloglou *et al.* 2013).



Εικόνα 1.14 Οι τρισδιάστατες δομές απακετυλάσεων πολυσακχαρίτη, του *B. cereus* (Bc) και του *B. anthracis* (Ba) Ba0424 (Oberbarnscheidt *et al.* 2007), Bc0361 (Fadoulglou *et al.* 2013), Ba0150 (Strunk *et al.* 2014), Bc1960 (Fadoulglou *et al.* 2017) έχουν προσδιοριστεί με κρυσταλλογραφία ακτίνων-X. Στην παρένθεση δίδεται ο κωδικός καταχώρησης και το έτος διάθεσης από την PDB (Berman *et al.* 2000) της κάθε δομής.

1.9.3 Μελέτες κυτταρικού εντοπισμού και χαρακτηρισμού απακετυλασών από την οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη του *B. anthracis* και του *B. cereus*

Αρχικά, ο βιοχημικός χαρακτηρισμός δύο γονιδιακών προϊόντων απακετυλάσης από τον *B. cereus*, της Bc1960 και της Bc3618, έδειξε τη δραστηριότητά τους ως απακετυλάσης της GlcNAc της πεπτιδογλυκάνης (Psylinakis *et al.* 2005). Στο *B. anthracis*, η Ba1977 έχει δειχθεί ότι είναι επίσης μία απακετυλάση της GlcNAc της πεπτιδογλυκάνης που εμπλέκεται στην αντοχή του βακίλλου στη λυσοζύμη του ξενιστή και απαιτείται για την πλήρη του μολυσματικότητα. Οι Ba1961 και Ba3679 συμμετέχουν στη βιογένεση της πεπτιδογλυκάνης κατά τη διάρκεια της επιμήκυνσης και της κυτταρικής διαίρεσης (Balomenou *et al.* 2013) και εμπλέκονται στην αγκυροβόληση της κάψουλας του πολυ- γ -D-γλουταμικού οξέος στην επιφάνεια του *B. anthracis* (Candela *et al.* 2014).

Πειράματα εντοπισμού μετά από σύντηξη του γονιδίου της εκάστοτε απακετυλάσης με αυτό της GFP (πράσινη φθορίζουσα πρωτεΐνη green fluorescent protein) έδειξαν ότι η Gfp-Ba0330 εντοπίζεται στην περιφέρεια του κυττάρου και ενισχύεται στα διαφράγματα (septa). Ενώ η Gfp-Ba0331 εντοπίζεται σε ξεχωριστά σημεία στην περιφέρεια και παρουσιάζει χαμηλό φθορισμό στα διαφράγματα (Arnaouteli *et al.* 2015).

Η Ba1977 εντοπίζεται στη κυτταρική μεμβράνη και η Ba2944 βρίσκεται έξω από τον κυτταρικό περίβλημα (Balomenou *et al.* 2015). Η Ba5436 είναι σημαντική για την προσάρτηση στην πεπτιδογλυκάνη του ουδέτερου πολυσακχαρίτη, στον οποίο αγκυροβολούν οι S-layer πρωτεΐνες και η Ba2944 είναι σημαντική για την τροποποίηση του ουδέτερου πολυσακχαρίτη (Balomenou *et al.* 2013). Τα πειράματα κυτταρικού εντοπισμού της Ba5436 έδειξαν ότι βρίσκεται κοντά στο κυτταρικό περίβλημα, ενώ εκείνα για τις Ba1961 και Ba3679 έδειξαν ότι βρίσκονται στο κυτταρικό τοίχωμα, σε ένα ελικοειδές πρότυπο και γύρω από το περίβλημα του κυττάρου και τα διαφράγματα, αντίστοιχα (Balomenou *et al.* 2015).

1.10 Προοπτική

Το 1928 ο Αλέξανδρος Φλέμινγκ ανακάλυψε την πενικιλίνη. Σχεδόν δυο δεκαετίες μετά 1945-1946 η πενικιλίνη έγινε ευρέως διαθέσιμη. Το 1945 η Dorothy Hodgkin χρησιμοποιώντας κρυσταλλογραφία ακτίνων-X προσδιόρισε τη δομή της και άνοιξε το δρόμο για τη χημική σύνθεση του φαρμάκου. Το 1952 αναπτύχθηκε μια μορφή του φαρμάκου κατάλληλη για χορήγηση από το στόμα, η πενικιλίνη V, η οποία το 1957 συντέθηκε χημικά και προχώρησε η ανάπτυξη και η σύνθεση αντιβιοτικών βασισμένων στην πενικιλίνη. Η πενικιλίνη δρα διακόπτοντας μία ουσιαστική διαδικασία του βακτηριακού κυτταρικού κύκλου, τη σύνθεση του κυτταρικού του τοιχώματος. Γύρω στο 1969 είχε γίνει πιστευτό ότι ο πόλεμος εναντίον των λοιμώξεων έχει τελειώσει. Τις προηγούμενες δύο δεκαετίες, τρεις μολυσματικές ασθένειες που αποτελούσαν μάστιγα της ανθρωπότητας για αιώνες - πολιομυελίτιδα, ευλογιά και φυματίωση - είχαν ουσιαστικά εξαλειφθεί σε όλο τον κόσμο. Αντιβιοτικά όπως η πενικιλίνη, η τετρακυκλίνη, η αμπικιλίνη και η χλωραμφενικόλη, έχουν σώσει αμέτρητες ζωές από βακτηριακές λοιμώξεις. Ωστόσο, η υπερβολική χρήση και η κατάχρηση των αντιβιοτικών, ώθησε την εξέλιξη ανθεκτικών βακτηριακών στελεχών. Τα βακτήρια ανέπτυξαν μηχανισμούς αντίστασης που τους επιτρέπουν να πολλαπλασιάζονται ακόμη και με την παρουσία των νεότερων αντιβιοτικών. Για παράδειγμα, κάθε χρόνο, περισσότεροι από 2 εκατομμύρια νοσοκομειακοί ασθενείς στις Ηνωμένες Πολιτείες αποκτούν ενδονοσοκομειακή λοίμωξη ανθεκτική στα αντιβιοτικά και ως αποτέλεσμα 90.000 πεθαίνουν (Griffiths *et al.* 2015).

Περισσότερο από το ήμισυ όλων των αντιβιοτικών που χρησιμοποιούνται σήμερα στοχεύουν στο βακτηριακό ριβόσωμα, όπου γίνεται η πρωτεϊνική σύνθεση στους προκαρυωτικούς οργανισμούς. Δομικές μελέτες έχουν εξετάσει τη μεγάλη υπομονάδα του ριβοσώματος σε σύμπλοκο με αρκετά διαφορετικά αντιβιοτικά. Παρόλο που τα ριβοσώματα των προκαρυωτικών και των ευκαρυωτικών είναι πολύ παρόμοια, εξακολουθούν να υπάρχουν λεπτές διαφορές. Λόγω αυτών των διαφορών, τα αντιβιοτικά είναι ικανά να στοχεύουν τα βακτηριακά ριβοσώματα αλλά να αφήνουν άθικτα τα ευκαρυωτικά ριβοσώματα. Από δομικές μελέτες, με κρυσταλλογραφία ακτίνων-X έχει δειχθεί ότι μεταλλάξεις του βακτηριακού rRNA και / ή των ριβοσωματικών πρωτεϊνών είναι υπεύθυνες για την ανοχή στα αντιβιοτικά. Αυτές οι μελέτες ταυτοποίησαν τα σημεία επαφής μεταξύ του αντιβιοτικού και του ριβοσώματος και, με αυτόν τον τρόπο, παρείχαν μια εξήγηση για το γιατί ορισμένα αντιβιοτικά απενεργοποιούν μόνο βακτηριακά ριβοσώματα. Για παράδειγμα, οι μακρολίδες είναι μια οικογένεια δομικά παρόμοιων ενώσεων που

περιλαμβάνει τα δημοφιλή αντιβιοτικά ερυθρομυκίνη και Zithromax. Αυτά τα αντιβιοτικά αναστέλλουν τη σύνθεση των πρωτεϊνών με την παύση του ριβοσώματος στο mRNA. Το κάνουν με σύνδεση με μια συγκεκριμένη περιοχή στο 23S rRNA στην μεγάλη υπομονάδα του ριβοσώματος εμποδίζοντας τη σήραγγα διέλευσης όπου το αναδυόμενο πολυπεπτίδιο εξέρχεται από τη μεγάλη υπομονάδα. Λόγω των μικρών διαφορών στην αλληλουχία μεταξύ των rRNAs των προκαρυωτικών και των ευκαρυωτικών, τα μακρολίδια αναστέλλουν μόνο τη βακτηριακή μετάφραση. Τα παθογόνα βακτήρια που έχουν αναπτύξει αντίσταση σε ορισμένα από αυτά τα αντιβιοτικά φαίνεται να έχουν μεταλλάξεις στο ριβόσωμα που κάνουν τη σήραγγα που οδηγεί στην έξοδο μεγαλύτερη. Έτσι, όπως περιγράφεται στο βιβλίο Γενετικής Ανάλυσης των Griffiths *et al.* (2015) αυτό που κάνουν οι γενετιστές σήμερα μέσα από τις δομικές μελέτες και την κατανόηση του ακριβούς τρόπου δράσης των αντιβιοτικών είναι να απαντήσουν σε ερωτήματα για τον τρόπο με τον οποίο τα αντιβιοτικά προσδένονται στα ριβοσώματα, για την κατανόηση της λειτουργίας, για το σχεδιασμό νέων αντιβιοτικών που να μπορούν να είναι δραστικά ενάντια στις μεταλλάξεις που έχουν δημιουργήσει ανθεκτικότητες. Η χρήση δομικών πληροφοριών σχετικά με την κυτταρική μηχανή με στόχο την ανάπτυξη νέων αντιβιοτικών και άλλων φαρμάκων αναφέρεται ως σχεδιασμός φαρμάκων που βασίζεται στην δομή (structure-based drug design).

Σήμερα λιγότερο από 100 χρόνια από την ανακάλυψη της πενικιλίνης από τον Φλέμινγκ η ανθρωπότητα βρίσκεται και πάλι μπροστά στην επιτακτική ανάγκη εύρεσης νέων αντιβακτηριακών παραγόντων και νέων βακτηριακών στόχων (Ventola 2015).

Οι απακετυλάσες της πεπτιδογλυκάνης που μελετήθηκαν μέχρι τώρα έχει αποδειχθεί ότι είναι μεταλλοένζυμα (Blair *et al.* 2005). Η στόχευση των βακτηριακών μεταλλοενζύμων είναι μια ελκυστική προσέγγιση για την ανάπτυξη νέων αντιβακτηριακών φαρμάκων, καθώς έχει αποδειχθεί ότι οι αναστολείς μεταλλοενζύμου μπορούν να είναι αρκετά επιλεκτικοί για τους στόχους τους και μπορούν να χρησιμοποιηθούν είτε μόνοι είτε σε συνδυασμό με γνωστά αντιβιοτικά για τη αντιμετώπιση των ανθεκτικών σε φάρμακα μηχανισμών (Blair *et al.* 2005, Fukushima *et al.* 2005). Οι απακετυλάσες της GlcNAc της πεπτιδογλυκάνης παρουσιάζουν ευνοϊκά χαρακτηριστικά ως αντιβακτηριακοί στόχοι, αφού εξαιτίας του εντοπισμού τους εξαλείφεται η ανάγκη για διαπερατά από τη μεμβράνη φάρμακα, και είναι προσιτοί σε αναστολείς μικρών μορίων. Τα μεταλλαγμένα στελέχη των απακετυλασών της GlcNAc της πεπτιδογλυκάνης από διάφορα παθογόνα, όπως και το Δba1977 μεταλλαγμένο στέλεχος του *B. anthracis* (Balomenou *et al.* 2013), εμφανίζουν μειωμένη λοιμογόνο δράση και ως εκ τούτου μπορούν να θεωρηθούν ως επικυρωμένοι αντιβιοτικοί στόχοι (Blair *et al.* 2005).

B. Σκοπός της διατριβής

Το 1970 ο Susumu Ohno διατύπωσε την ιδέα ότι η διπλοποίηση γονιδίου (gene duplication) είναι η κύρια δύναμη που δρα στην εξέλιξη των οργανισμών, διότι δημιουργεί το γενετικό υπόβαθρο, το οποίο στη συνέχεια θα τροποποιηθεί μέσω της φυσικής επιλογής. Αυτή η άποψη έχει επιβεβαιωθεί σε μεγάλο βαθμό μέσα από την αλληλούχιση γονιδιωμάτων.

Οι πολυγονιδιακές οικογένειες (multigene families) αποτελούν έναν απλό τρόπο και υψηλά προτιμητέο για τη δημιουργία γενετικής και λειτουργικής ευελιξίας. Όμως η σύνθετη διαδικασία που διαμορφώνει την εξέλιξή τους, μεταξύ άλλων μπορεί να σημαίνει ότι αυτές συχνά αποτελούνται από μωσαϊκά αλληλουχιών, καθένα από τα οποία έχει μια διαφορετική φυλογενετική ιστορία, παρά από αυστηρώς ομόλογα γονίδια που έχουν σταδιακά διαφοροποιηθεί κατά τη διάρκεια του χρόνου (Page & Holmes 1998).

Η εξέλιξη των διπλοποιημένων γονιδίων παρουσιάζει ενδιαφέρον και είναι σημαντική πέρα από το ότι είναι κύρια πηγή νέων γονιδίων στο κάθε γονιδίωμα, διότι τα διπλοποιημένα γονίδια δίνουν μια μοναδική και ισχυρή δυνατότητα για τη μελέτη της λειτουργικής διαφοροποίησής τους εντός ενός γονιδιώματος και κατ' επέκταση εντός κοινών συνθηκών π.χ. ίδιες περιβαλλοντικές πιέσεις (Makino *et al.* 2006). Μελέτες έχουν περιγράψει τη λειτουργική διαφοροποίηση των γονιδίων και τη σημασία της στην εξέλιξη μέσω της σημαντικής διαδικασίας της διπλοποίησης γονιδίου (Hughes 1994, Rastogi & Liberles 2005, Innan & Kondrashov 2010). Η μελέτη παράλογων γονιδίων (paralogous genes), ενισχύει την ερμηνεία των διαφορών που έχουν συσσωρευτεί στα διπλοποιημένα γονίδια από τη στιγμή της προέλευσής τους ως ταυτόσημα αντίγραφα.

Ο *Bacillus anthracis*, είναι ένα κατά Gram-θετικό, αερόβιο, ραβδόμορφο βακτήριο, που σχηματίζει σπόρια και είναι ο αιτιολογικός παράγοντας για την οξεία λοιμώδη νόσο του άνθρακα. Ανήκει στο γένος *Bacillus* της οικογένειας Bacillaceae. Η ομάδα *Bacillus cereus* (αναφερόμενη ως *B. cereus sensu lato*), μία υποομάδα συναφών Bacilli, που περιλαμβάνει τα *B. anthracis*, *B. cereus*, *B. thuringiensis*, *B. mycoides*, *B. pseudomycoides* και *B. weihenstephanensis* (Helgason *et al.* 2000, Økstad & Kolstø 2011). Τα πρώτα γονιδιώματα της ομάδας *B. cereus* των οποίων έγινε αλληλούχιση, και από τότε χρησιμεύουν ως γονιδιώματα αναφοράς, ήταν εκείνα των στελεχών *B. anthracis* Ames (Read *et al.* 2003) και *B. cereus* ATCC 14579 (Ivanova *et al.* 2003). Είναι αξιοσημείωτο ότι τα γονιδιώματα αυτά περιέχουν πολλαπλά γονίδια (ORFs) απακετυλάσων πολυσακχαρίτη (PDAs). Για τη διερεύνηση των εξελικτικών σχέσεων των γονιδίων τόσο εντός του είδους όσο και μεταξύ

των ειδών κατασκευάστηκαν φυλογενετικά δένδρα χρησιμοποιώντας τη νουκλεοτιδική αλληλουχία των γονιδίων αλλά και την αμινοξική αλληλουχία των αντίστοιχων πρωτεϊνών.

Οι απακετυλάσες πολυσακχαριτών μέλη της οικογένειας 4 των υδατανθρακικών εστερασών, εκτελούν κυρίως απακετυλίωση εξαρτώμενη από μέταλλα σε *O*- ή *N*-ακετυλιωμένους πολυσακχαρίτες όπως πεπτιδογλυκάνη, χιτίνη και ακετυλοξυλάνη μέσω ενός συντηρημένου καταλυτικού πυρήνα που ονομάζεται NodB ομόλογη δομική ενότητα (Kafetzopoulos *et al.* 1993, Tsigos *et al.* 2000, Caufrier *et al.* 2003).

Ο ρόλος της πεπτιδογλυκάνης (PG) είναι πολύ σημαντικός για την επιβίωση των βακτηρίων, καθώς είναι το κύριο συστατικό του βακτηριακού κυτταρικού τοιχώματος. Το ανοσοποιητικό σύστημα των θηλαστικών χρησιμοποιεί μία ποικιλία υδρολυτικών ενζύμων προκειμένου να διασπάσει τη στοιβάδα της πεπτιδογλυκάνης. Ωστόσο, τα βακτήρια έχουν αναπτύξει στρατηγικές συγκάλυψης, ώστε να αποφεύγουν την αναγνώριση από τα ένζυμα αυτά και να αυξάνουν τις πιθανότητες επιβίωσής τους. Δευτερογενείς τροποποιήσεις βακτηριακών αλυσίδων των γλυκανών περιλαμβάνουν *N*-απακετυλίωση, *N*-γλυκολυλίωση και *O*-ακετυλίωση. Η *N*-απακετυλίωση των καταλοίπων GlcNAc (*N*-ακετυλογλυκοζαμίνης) και MurNAc (*N*-ακετυλομουραμικού οξέος) της PG συσχετίζεται με την ανθεκτικότητα των βακτηρίων στη λυσοζύμη του ξενιστή, τις επιδράσεις στην αναγνώριση των παθογόνων μέσω διαφόρων υποδοχέων των ξενιστών ενώ εισάγει επιπλέον θετικό φορτίο στο κυτταρικό τοίχωμα που μπορεί να αυξήσει την αντίσταση του βακτηρίου στα κατιονικά αντιμικροβιακά πεπτίδια του ξενιστή (Vollmer 2008).

Η λεπτομερής γνώση της δομής των απακετυλασών σε ατομικό επίπεδο και η αποσαφήνιση του μηχανισμού απακετυλίωσης των ενζύμων αυτών και των δομικών τους ενοτήτων που συμμετέχουν στη μοριακή αναγνώριση, αναμένεται αφενός να βοηθήσει στην εξελικτική μελέτη της πολυγονιδιακής οικογένειας και αφετέρου αποτελεί απαραίτητη προϋπόθεση ώστε να γίνει δυνατός ο λελογισμένος σχεδιασμός βάσει δομής ειδικών και αποτελεσματικών αναστολέων των συγκεκριμένων ενζύμων.

Προκειμένου να καταστεί δυνατή η δομική μελέτη το πρώτο βήμα είναι η εξεύρεση κατάλληλου σχήματος ετερόλογης έκφρασης. Στο πλαίσιο αυτό, επιδιώχθηκε κατάλληλη κατασκευή μέσω γενετικής μηχανικής για την έκφραση ολόκληρης της Ba3679 του *B. anthracis* για την απομόνωση σταθερού και ομοιογενούς πρωτεϊνικού υλικού όπως απαιτούν οι δομικές μελέτες.

Η υψηλή ατομική διακριτικότητα των δομών μπορεί να επιτρέψει την άντληση συμπερασμάτων για τη συσχέτιση δομής και τον προσδιορισμό των στοιχείων λειτουργικής διαφοροποίησης της καταλυτικής δομικής ενότητας των συγκεκριμένων μορίων. Η μελέτη

δομής απακετυλασών της οικογενείας των γονιδίων απακετυλάσης πολυσακχαρίτη του *B. anthracis* και του *B. cereus* επιλέχθηκε να γίνει μέσω κρυσταλλογραφίας ακτίνων-Χ και η συλλογή των κρυσταλλογραφικών δεδομένων με ακτινοβολία συγχρότρου, προκειμένου η διακριτικότητα να είναι όσο το δυνατό υψηλότερη και να γίνει λεπτομερέστερη ανάλυση των δομικών χαρακτηριστικών τους. Για τη διερεύνηση αυτή, επιλέχθηκαν οι φερόμενες απακετυλάσες του βακτηρίου *B. anthracis* Ba0330, η οποία συμβάλλει στην προσαρμογή του βακτηρίου κάτω από ακραίες συνθήκες υψηλής συγκέντρωσης άλατος και της Ba0331 που συμβάλλει στην διατήρηση του κυτταρικού σχήματος (Arnaouteli *et al.* 2015), καθώς και η απακετυλάση της πεπτιδογλυκάνης της *N*-ακετυλογλυκοζαμίνης (GlcNAc) Bc1974 του βακτηρίου *B. cereus*. Επίσης επιχειρήθηκε και δομική ανάλυση συμπλόκων τους με προσδέτες για να αποσαφηνιστεί ο μηχανισμός απακετυλίωσης ώστε να οδηγηθούμε στην πρόταση φαρμακοφόρου θέτοντας τη βάση του ορθολογικού σχεδιασμού (με βάση τη δομή) νέων αντιβακτηριακών παραγόντων. Η ανάπτυξη εξειδικευμένων αναστολέων έναντι της Bc1974 θα έχει πιθανή εφαρμογή και στην ορθόλογή της, απακετυλάση της πεπτιδογλυκάνης του *B. anthracis* Ba1977, η οποία είναι αποδεδειγμένα λοιμογόνος παράγοντας του άνθρακα (Balomenou *et al.* 2013). Οι τρισδιάστατες δομές (3Δ) των Ba0330 και Ba0331 χρησιμοποιήθηκαν στην μελέτη των παράλογων γονιδίων *ba0330* και *ba0331*. Οι τρισδιάστατες δομές απακετυλασών που προέκυψαν από την παρούσα διατριβή σε συνδυασμό με τις ήδη υπάρχουσες 3Δ δομές χρησιμοποιήθηκαν για την εξελικτική μελέτη της θέσης δέσμευσης μετά από γεγονότα διπλοποίησης.

2. Υλικά και Μέθοδοι

2.1. Υλικά

2.1.1. Εργαστηριακά όργανα

Θερμικός κυκλοποιητής (Applied Biosystems, verity 96 well thermos cycler)

Αυτόματο ρομποτικό σύστημα κρυστάλλωσης πρωτεϊνών (Oryx4, Douglas Instruments Ltd)

Σύστημα παραγωγής υπερκάθαρου ύδατος (MilliQ Direct 8, Millipore)

Αυτόκαυστο (Rayra)

Επωαστήρας για καλλιέργειες βακτηρίων υπό ανάδευση (incubator shaker I26 New Brunswick Scientific)

Ψυχόμενη Φυγόκεντρος (Kubota 7780)

Επιτραπέζιες ψυχόμενες φυγόκεντροι (Hettich Mikro 200R και Rotina 420R, Hettich)

Επιτραπέζια φυγόκεντρος (Eppendor centrifuge 5415C)

Φασματοφωτόμετρο ορατού/υπεριώδους (Nanodrop 2000c ThermoScientific, UV/VIS spectrometer Lambda Bio Perkin-Elmer)

Υδατόλουτρο (BioBlock scientific polystat 86633)

Σύστημα μικροδιήθησης διαλυμάτων (Millipore)

Πιπέτες ρυθμιζόμενου όγκου (Gilson)

Συσκευές οριζόντιας ηλεκτροφόρησης πηκτώματος αγαρόζης (submarine gel system Thermo Scientific, FG006915 Consort)

Συσκευή κάθετης ηλεκτροφόρησης πρωτεϊνών (Mini-Protean tetra cell BIORAD)

Τροφοδοτικό ρεύματος (BIORAD 100/500 , CONSORT E834,)

Συσκευή μεταφοράς πρωτεϊνών σε νιτροκυτταρίνη (Mini-Trans-Blot, BIORAD)

Συσκευή προσρόφησης υπό κενό πρωτεϊνικών διαλυμάτων σε μεμβράνη νιτροκυτταρίνης (Bio-Dot, BIORAD)

Σύστημα υγρής χρωματογραφίας (GE Pharmacia ÄKTA purifier)

Αυτόματος συλλέκτης κλασμάτων (fraction collector, Pharmacia FRAC-100)

Συσκευή απεικόνισης πηκτωμάτων (Gel Documentation System, BIORAD)

Ψηφιακή φωτογραφική μηχανή (Power Shot A520 Canon)

Στερεοσκόπιο (Leica MZ 125)

Όργανο μέτρησης pH (pH meter HI2110, HANNA)

Ηλεκτρονικός ζυγός ακριβείας (Mettler)

Ηλεκτρονικός ζυγός (Mettler)

Αναδευτήρες: (περιστροφικός Vortex LABINCO B.V. 524, οριζόντιος θερμαινόμενος ή/και μαγνητικός αναδευτήρας FALC Instruments Srl, BioLine Scientific SMT150)

Φούρνος μικροκυμάτων (MATSUI Mat SMW17E)

Λουτρό υπερήχων (ElmasonicS30, Elma Schmidbauer GmbH)

Συσκευή υπερήχων (Sonicator-3000, Qsonica)

2.1.2. Αναλώσιμα

Πλαστικοί σωλήνες μιας χρήσης 1,5 mL (Greiner Bio-One)

Πλαστικοί σωλήνες μιας χρήσης 0,25 mL (Greiner Bio-One)

Πλαστικά ακρορύγχια και πλαστικές πιπέτες 1, 2, 5, 10 και 25 mL (Greiner Bio-One)

Τρυβλία petri για στερεές καλλιέργειες βακτηρίων (Greiner Bio-One)

Κωνικές φιάλες 500 mL, 1 L, 2 L, (SIMAX, OMNILAB)

Πλαστικοί σωλήνες φυγοκέντρησης 250 mL, 500 mL (Corning)

Δοκιμαστικοί σωλήνες πολυπροπυλενίου 15 και 50 mL (Greiner Bio-One)

Πλαστικές κυψελίδες μιας χρήσης (KARTELL SPA)

Κυψελίδες χαλαζία (quartzcells) οπτικής διαδρομής 10 mm (HELLMA)

Φίλτρα με διάμετρο πόρου 0,22 και 0,45 μm (NALGENE και Millipore)

Μεμβράνη νιτροκυτταρίνης (HybondP, Amersham Biosciences)

Χαρτί διήθησης 3 MM (Whatmann)

Στήλες μοριακής διήθησης Superose 12, Superdex 75, Superdex 200 (GE healthcare)

Σωλήνες κρυογονικής 1,8 mL αποστειρωμένοι (cryotubes, GreinerBio-One)

Θηλιά στερέωσης κρυστάλλων 0,05mm – 1,0 mm (Mounted Litholoops Molecular Dimensions, Mounted CryoLoop Hampton Research Inc.)

Δίσκοι κρυστάλλωσης 96 φρεατιδίων (MRC 2 Well Crystallization Plate Swissci, Molecular Dimensions)

Δοκιμαστικοί σωλήνες φυγοκέντρησης με φίλτρο υπερδιήθησης (Amicon)

Σιλικόνη (Hampton Research Inc.)

Καλυπτρίδες (Cover slips, Greiner Bio-One)

Μικρογέφυρες (Crystal Bridge, Greiner Bio-One)

Γυάλινες στήλες χρωματογραφίας (BIORAD)
Σφαιρίδια (Seed Bead, Hampton Research Inc.)
Δίσκοι κρυστάλλωσης 24 φρεατιδίων (VDX plates, Hampton Research Inc.)

2.1.3. Αντιδραστήρια

Χρησιμοποιήθηκαν χημικά αντιδραστήρια καθαρότητας αναλυτικού βαθμού από Applichem, Sigma, Merck, και Fluka.

Ταq DNA πολυμεράση και ρυθμιστικό διάλυμα αντίδρασης (Promega)

KAPA HiFi Hot Start DNA πολυμεράση και ρυθμιστικό διάλυμα αντίδρασης (Kapa Biosystems)

dNTPs (Kapa Biosystems και New England Biolabs)

Εναρκτήρια μόρια (Eurofins Genomics AT GmbH)

Περιοριστικές ενδονουκλεάσες με το αντίστοιχο ρυθμιστικό διάλυμα αντίδρασης (New England Biolabs)

T4 DNA λιγάση με το αντίστοιχο ρυθμιστικό διάλυμα αντίδρασης (New England Biolabs)

HRV 3C πρωτεάση (Pierce HRV 3C Protease, Thermo Scientific)

Μάρτυρες μοριακών μεγεθών DNA (New England Biolabs και Fermentas Life Sciences)

Ni²⁺-NTA-αγαρόζη, QIAquick Gel Extraction Kit, QIAquick PCR Purification kit (Qiagen)

Bacto-tryptone, Bacto-peptone, Bacto yeast extract και Bacto agar (Difco Laboratories)

Αμπικιλίνη (Bristol Mayers Squib)

Καναμυκίνη (Sigma)

Χλωραμφαινικόλη (Sigma)

Αγαρόζη (BioRad)

RNaseA

TEMED, APS, DAB, DTT, Ακρυλαμίδιο, SDS, β-μερκαπταιθανόλη, DMSO, NiCl₂ και μιδαζόλιο (Sigma)

Μίγμα αναστολέων πρωτεασών (PMSF, Thermo Scientific)

Αλβουμίνη από ορό βοός (BSA, Applichem)

Έγχρωμοι μάρτυρες μοριακών μεγεθών πρωτεϊνών (προ-χρωματισμένοι) Pink Prestained Protein Marker, 10,5-175kDa (Nippon Genetics), Page Ruler, 10-180 kDa, (Thermo Fisher Scientific)

Φαινόλη, ισο-προπανόλη, χλωροφόρμιο και ισο-αμυλική αλκοόλη (Sigma)
 Λυσοζύμη (Applichem)
 Χρωστική πρωτεϊνικών κρυστάλλων (Izit Crystal Dye, Hampton Research)
 Πολυκλωνικό αντίσωμα αντι-His (Santa Cruz)
 Πολυκλωνικό αντίσωμα goat anti-mouse-HRP (συζευγμένο με το ένζυμο υπεροξειδάση της ραπανίδος, Darko).
 Βιβλιοθήκες σάρωσης συνθηκών κρυστάλλωσης (crystallization screening kits, Molecular Dimensions (Αγγλία), Qiagen και Jena Bioscience (Γερμανία), Hampton Research (Αμερική).
 Συλλογές επιπρόσθετων χημικών ουσιών δοκιμασιών κρυστάλλωσης πρωτεϊνών (Additive Screen HT, Hampton Research)

2.1.4 Διαλύματα

Στην παρασκευή όλων των διαλυμάτων το ύδωρ που χρησιμοποιήθηκε ήταν δις απεσταγμένο και απιονισμένο (ddH₂O), μετά από διήθηση σε ηθμό με διάμετρο πόρων 0,22 μm και έκθεση του σε ακτινοβολία UV.

TBE	89 mM Tris-HCl, 90mM βορικού οξέος, 2 mM EDTA, pH 8,0
TAE	40 mM Tris-CH ₃ COOH, 5 mM EDTA, pH 8,0
TE (10:1)	10 mM Tris-HCl, 1 mM EDTA, pH 8,0
PBS	60 mM ρυθμιστικό διάλυμα φωσφορικών νατρίου, 137 mM (0,8% κ.β.) NaCl, 2,7 mM (0,02% κ.β.) KCl, pH 7,4
PBST	PBS + 0,1% κ.ο. Tween-20
GTE	50 mM γλυκόζη, 25 mM Tris-HCl, 10 mM EDTA, pH 8,0
TBS	20 mM Tris-HCl, 150 mM NaCl, 0,5% κ.β. NaN ₃ , pH 7,5

Χρωματογραφίας συγγένειας

Διάλυμα 1 ^{ης} έκπλυσης	50 mM Tris-HCl, 300 mM NaCl, 0,5% NaN ₃ , 10 mM, pH 8,0
Διάλυμα 2 ^{ης} έκπλυσης	50 mM Tris-HCl, 300 mM NaCl, 0,5% NaN ₃ , 30 mM μιδαζόλιο pH 8,0

Έκλουσης	50 mM Tris-HCl, 300 mM NaCl, 0,5% NaN ₃ , 200 mM ιμιδαζόλιο, pH 8,0
Χρωματογραφίας μοριακού αποκλεισμού Ηλεκτροφόρησης	20 mM Tris-HCl, 150 mM NaCl, 0,5% κ.β. NaN ₃ , pH 7,5 25 mM Tris-base, 200 mM γλυκίνη, 1% κ.β. SDS, pH 8,5
Ηλεκτρομεταφοράς (transfer buffer)	25 mM Tris-base, 200 mM γλυκίνη, 20% κ.β. μεθανόλη, pH 8,5
Χρώσης πηκτής πολυακρυλαμιδίου	0,1% κ.β. Coomassie R-250, 40% κ.ο. μεθανόλη, 10% κ.ο. CH ₃ COOH
Αποχρωματισμού πηκτής πολυακρυλαμιδίου	10% κ.ο. CH ₃ COOH, 40% κ.ο. μεθανόλη
Διάλυμα εμφάνισης με υπόστρωμα DAB	0,03% κ.β. DAB, 0,03% κ.β. NiCl ₂ , 0,009% κ.ο. H ₂ O ₂ σε PBS
Διάλυμα Φόρτωσης δείγματος πρωτεϊνών	45 mM Tris-HCl, pH 6,8, 10% κ.ο. γλυκερόλη, 1% κ.β. SDS, 0,01% κ.β. κυανού της βρωμοφαινόλης, 5% κ.ο. β-μερκαπτοαιθανόλη

2.1.5 Θρεπτικά υλικά

Τα θρεπτικά υλικά που χρησιμοποιήθηκαν για τις καλλιέργειες κυττάρων *E. coli* αποστειρώθηκαν σε κλίβανο υγρής αποστείρωσης (20 min, 121 °C, 2 Atm)

LB (Luria Bertani)
Υγρό

1% κ.β. NaCl, 1% κ.β. πεπτόνη, 0,5% κ.β. εκχύλισμα ζύμης, pH 7,0 (± κατάλληλο αντιβιοτικό)

Στερεό

Προσθήκη στο υγρό LB 1,5% κ.β. άγαρ

2.1.6. Εναρκτήρια μόρια (primers)

Οι νουκλεοτιδικές αλληλουχίες των εναρκτήριων μορίων (ή εκκινήτων) που χρησιμοποιήθηκαν ήταν οι εξής:

Pci-Ba3679-F: 5'–AATAACATGTTGCTCAGAAAAGAGCTAGAACC–3'

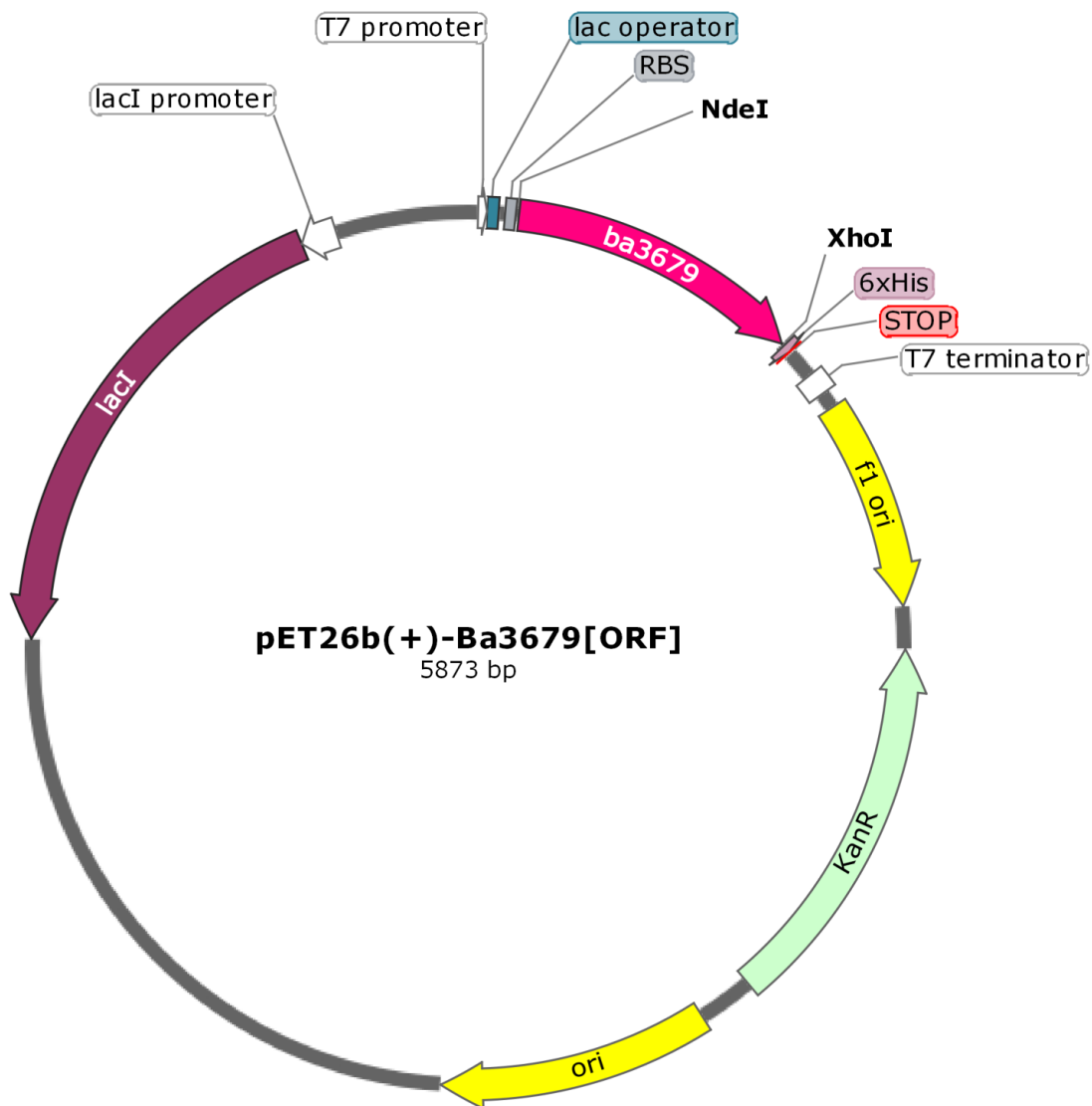
Sal-Ba3679-R: 5'–AATGTCGACTCAATGTTTATAACGTAATAATTCCG–3'

2.1.7 Προσδέτες (ligands)

Οι προσδέτες δίδονται στο παράρτημα Α' (Πίνακας 1) και πληροφορίες αντλούνται από τη βάση PubChem, (Kim *et al.* 2016). Οι προσδέτες ήταν της Sigma-Aldrich εκτός αν αναφέρεται διαφορετικά και η ελάχιστη καθαρότητά τους ήταν 95%, ενώ χρησιμοποιήθηκαν χωρίς περαιτέρω επεξεργασία. Οι ενώσεις 1 έως 6 και η 15 ήταν ευγενική χορηγία από τον καθηγητή S. M. Ulrich (Ithaca College, New York, USA).

2.1.8 Φορέας κλωνοποίησης (cloning vector) και βακτηριακό στέλεχος

Το γονίδιο της Ba3679 ήταν υποκλωνοποιημένο στον φορέα έκφρασης pET26b(+) μεταξύ των θέσεων περιορισμού NdeI και XhoI κατασκευάστηκε στο εργαστήριο του καθηγητή κ. Βασίλειου Μπουριώτη στο Πανεπιστήμιο Κρήτης. Για την παραγωγή αυτού του πλασμιδίου σε μεγάλη κλίμακα χρησιμοποιήθηκε το βακτηριακό στέλεχος TOP10F' του *E. coli* (Invitrogen). Το πλασμίδιο pET26b(+) διαθέτει γονίδιο ανθεκτικότητας στην καναμυκίνη, τον T7 υποκινητή (που αναγνωρίζει η T7 RNA πολυμεράση), επίσης διαθέτει γονίδιο για την έκφραση και περιοχή πρόσδεσης του μεταγραφικού παράγοντα lacI. Ο lacI παρεμποδίζει τη μεταγραφή του γονιδίου της πρωτεΐνης – στόχου αλλά αναστέλλεται από την παρουσία λακτόζης ή IPTG στο θρεπτικό μέσο. Με τον τρόπο αυτό ελαχιστοποιείται η έκφραση της πρωτεΐνης – στόχου πριν από την επαγωγή έκφρασης με προσθήκη IPTG στο θρεπτικό μέσο. Το κωδικόνιο λήξης της μετάφρασης βρίσκεται αμέσως μετά την ετικέτα των έξι καταλοίπων ιστιδίνης. Σχηματική αναπαράσταση του φορέα κλωνοποίησης παρουσιάζεται στην **Εικόνα 2.1**.

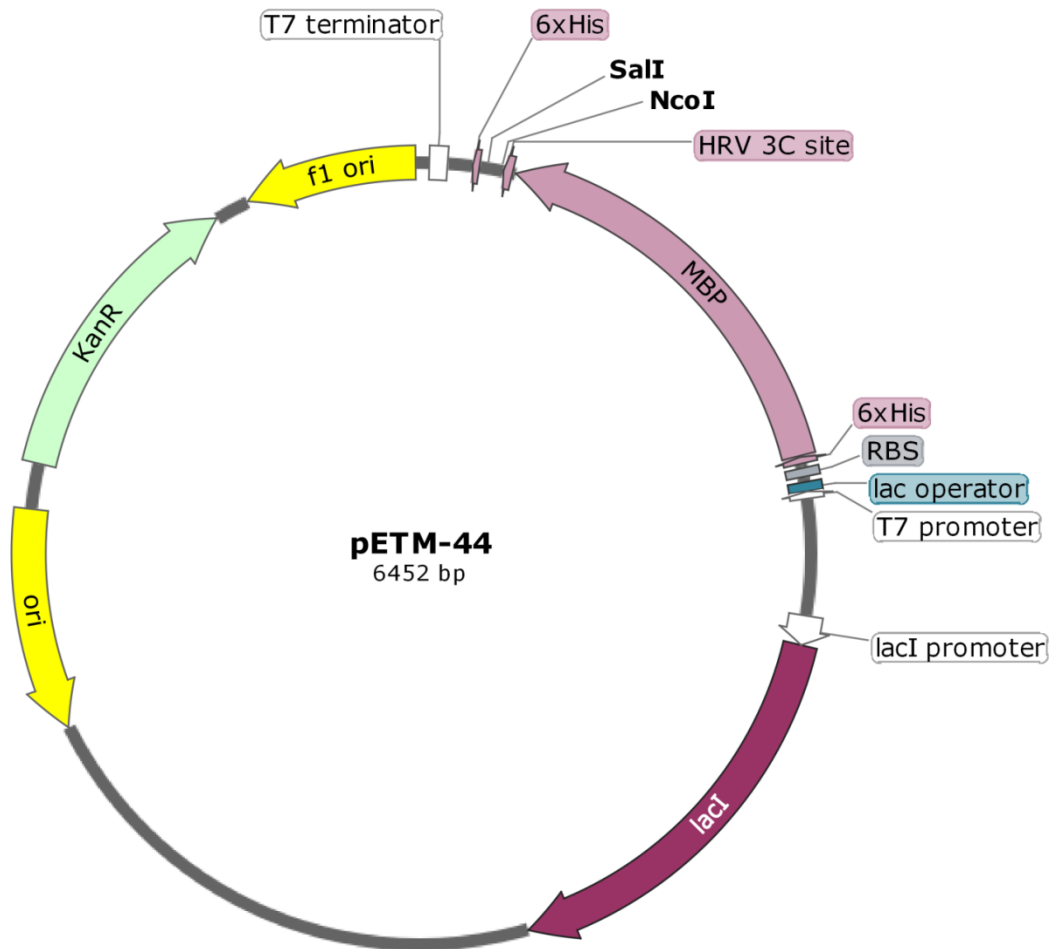


Εικόνα 2.1 Πλασμιδιακός χάρτης του φορέα κλωνοποίησης pET26b(+) με ένθεμα Ba3679 (Ba3679[ORF]) με προσαρτημένη καρβοξυτελική ετικέτα 6 καταλοίπων ιστιδίνης (6xHis). *T7 promoter* & *T7 terminator*: περιοχές έναρξης και λήξης της μεταγραφής, αντίστοιχα, *Ba3679[ORF]*: το ανοιχτό πλαίσιο ανάγνωσης της Ba3679, *STOP*: κωδικόνιο λήξης της μετάφρασης, *lac operator*: περιοχή πρόσδεσης του μεταγραφικού παράγοντα *lacI*, *RBS*: περιοχή πρόσδεσης του ριβοσώματος, *NdeI* & *XhoI*: θέσεις αναγνώρισης περιοριστικών ενδονουκλεασών, *KanR*: γονίδιο ανθεκτικότητας στην καναμυκίνη, *lacI* & *lacI promoter*: ο μεταγραφικός παράγοντας *lacI* και ο υποκινητής του, *ori* & *f1 ori*: περιοχές έναρξης της αντιγραφής του πλασμιδίου, bp: ζεύγη βάσεων.

2.1.9 Φορέας και στελέχη έκφρασης σε βακτηριακά συστήματα

Για την ετερόλογη έκφραση της Ba3679 χρησιμοποιήθηκαν βακτηριακά συστήματα. Αρχικά επιλέχθηκε το BL21(DE3) του *E. coli*. Ακολούθως τα στελέχη που δοκιμάστηκαν ήταν τα C41(DE3), C43(DE3) και Rosetta 2(DE3) (Novagen) του *Escherichia coli*, τα στέλεχη αυτά είναι θυγατρικά του στελέχους BL21. Επίσης, φέρουν ένα χρωμοσωμικό αντίγραφο του γονιδίου της T7 RNA πολυμεράσης υπό τον έλεγχο του *lacUV5* υποκινητή. Συνεπώς, όταν το γονίδιο της πρωτεΐνης – στόχου έχει υποκλωνοποιηθεί σε πλασμιδιακό φορέα έκφρασης της σειράς pET και έχει ακολουθήσει μετασχηματισμός σε DE3, η επαγωγή έκφρασης είναι εφικτή μετά από προσθήκη IPTG στο θρεπτικό μέσο.

Ως φορέας έκφρασης, χρησιμοποιήθηκε το πλασμίδιο pETM-44, που διαθέτει γονίδιο ανθεκτικότητας στην καναμυκίνη, τον T7 υποκινητή (που αναγνωρίζει η T7 RNA πολυμεράση), την πρωτεΐνη MBP, (Maltose Binding Protein, πρωτεΐνη δέσμευσης της μαλτόζης), δε διαθέτει ακολουθία που κωδικεύει βακτηριακό πεπτίδιο-οδηγό, και έτσι η MBP παραμένει στο κυτταρόπλασμα, επίσης διαθέτει την πρωτεολυτική θέση HRV 3C (**Εικόνα 2.2**).



Εικόνα 2.2. Πλασμιδιακός χάρτης του φορέα έκφρασης pETM-44. 6xHis: ετικέτα 6 καταλοίπων ιστοιδίνης, MBP: πρωτεΐνη δέσμευσης της μαλτόζης (maltose binding protein), HRV 3Csite:θέση αναγνώρισης της HRV (human rhinovirus) 3C πρωτεάσης, *T7 promoter* & *T7 terminator*: περιοχές έναρξης και λήξης της μεταγραφής, αντίστοιχα, *lac operator*: περιοχή πρόσδεσης του μεταγραφικού παράγοντα *lacI*, RBS: περιοχή πρόσδεσης του ριβοσώματος, *NcoI* & *SalI*: θέσεις αναγνώρισης περιοριστικών ενδονουκλεασών, *KanR*: γονίδιο ανθεκτικότητας στην καναμυκίνη, *lacI* & *lacI promoter*: ο μεταγραφικός παράγοντας *lacI* και ο υποκινητής του, *ori* & *f1 ori*: περιοχές έναρξης της αντιγραφής του πλασμιδίου, bp: ζεύγη βάσεων.

2.2 Μέθοδοι

2.2.1 Γενετική Μηχανική

2.2.1.1 Απομόνωση πλασμιδιακού DNA

Για την απομόνωση πλασμιδιακού DNA από τα μετασχηματισμένα βακτηριακά κύτταρα του στελέχους TOP10F' του *E. coli* των υγρών καλλιεργειών, χρησιμοποιήθηκε εμπορικό προϊόν συστηματοποιημένης μεθόδου (kit) σε μικρή κλίμακα εφαρμογής (mini-prep) και σε μεγάλη κλίμακα εφαρμογής (maxi-prep), σύμφωνα με τις οδηγίες της παρασκευάστριας εταιρείας (QIAGEN). Η αρχή της μεθόδου βασίζεται στην αλκαλική λύση κυττάρων (Birnboim & Doly 1979) και για την απομόνωση του πλασμιδιακού DNA χρησιμοποιούνται ειδικές χρωματογραφικές στήλες. Κατά την απομόνωση σε μικρή κλίμακα απομονώνεται μικρή ποσότητα καθαρισμένου πλασμιδίου που χρησιμοποιήθηκε στον έλεγχο της επιτυχίας της υποκλωνοποίησης των κατασκευών έκφρασης στους πλασμιδιακούς φορείς έκφρασης. Κατά την απομόνωση σε μεγάλη κλίμακα απομονώνεται μεγάλη ποσότητα (της τάξεως των mg) πλασμιδιακού DNA σε υψηλή καθαρότητα. Η μέθοδος αυτή χρησιμοποιήθηκε για την απομόνωση πλασμιδίων από ένα θετικό κλώνο σε ικανή ποσότητα, όπου αυτή απαιτείται π.χ. για την απομόνωση τμημάτων DNA, την υποκλωνοποίηση, το μετασχηματισμό βακτηρίων με τους ανασυνδυασμένους φορείς έκφρασης και τη παρακαταθήκη ενός μέρους.

2.2.1.2 Καθαρισμός νουκλεϊκών οξέων (DNA purification)

A. Με εκχύλιση με οργανικούς διαλύτες

Για τον καθαρισμό των νουκλεϊκών οξέων, χρησιμοποιήθηκε η μέθοδος εκχύλισης νουκλεϊκών οξέων, με οργανικούς διαλύτες. Κατά την εκχύλιση με βασικό διάλυμα εξισοροπημένης φαινόλης οι πρωτεΐνες απομακρύνονται στην οργανική φάση, ενώ το καθαρό DNA παραλαμβάνεται στην υδατική φάση. Το υδατικό διάλυμα DNA εκχυλίζεται με προσθήκη διαδοχικά ίσων όγκων διαλύματος φαινόλης, φαινόλης-χλωροφορμίου/ισοαμυλικής αλκοόλης 1:1 και χλωροφορμίου/ισοαμυλικής αλκοόλης. Κάθε φορά οι δύο φάσεις αναμιγνύονται μετά από έντονη ανάδευση και διαχωρίζονται μετά από σύντομη φυγοκέντρηση σε επιτραπέζια μικροφυγόκεντρο (13.000 rpm, 5 min, θερμοκρασία

δωματίου). Η υδατική φάση που απομονώνεται κάθε φορά χρησιμοποιείται στην διαδοχική και τελικά παραλαμβάνεται το καθαρό υδατικό διάλυμα νουκλεϊκού οξέος.

B. Με συστηματοποιημένες χρωματογραφικές μεθόδους

Για την απομόνωση τμημάτων DNA χρησιμοποιήθηκαν εμπορικά προϊόντα συστηματοποιημένης μεθόδου (kit), που διαχωρίζουν το DNA με χρωματογραφικές μεθόδους σύμφωνα με τις οδηγίες της παρασκευάστριας εταιρείας. Για την εξαγωγή μιας ζώνης DNA από πήκτωμα αγαρόζης μετά από αντίδραση πέψης ή PCR και ηλεκτροφόρηση σε πήκτωμα αγαρόζης χρησιμοποιήθηκε το QIAquick Gel Extraction Kit (Qiagen). Για τον καθαρισμό DNA προϊόντων απ' ευθείας από το μίγμα της αντίδρασης (π.χ. PCR ή αντίδραση πέψης) το QIAquick PCR purification Kit (Qiagen). Η έκλυση του DNA από τη στήλη έγινε σε 40 μ L ρυθμιστικού διαλύματος 10 mM Tris-HCl, pH 8,0.

2.2.1.3 Κατακρήμιση νουκλεϊκών οξέων με αιθανόλη (DNA precipitation)

Η κατακρήμιση των νουκλεϊκών οξέων εφαρμόζεται στη συμπύκνωση ή στην απομάκρυνση προσμίξεων από το υδατικό διάλυμα των νουκλεϊκών οξέων. Στο υδατικό διάλυμα του DNA που απομονώθηκε προστίθενται διπλάσιος όγκος απόλυτης αιθανόλης (η οποία έχει προψυχθεί) και 1:10 του όγκου διαλύματος 3 M οξικού νατρίου pH 4,8. Ακολουθούν ανάδευση, φυγοκέντρηση (13.000 rpm, 5 min, 4°C) και απόχυση του υπερκειμένου κατά το δυνατόν πλήρης. Κατόπιν μία πλύση του ίζηματος με αιθανόλη 70% κ.ο. και επανάληψη της φυγοκέντρησης και πλήρους απόχυσης. Το ίζημα αφήνεται να στεγνώσει με επώαση στους 37°C. Έπειτα από την πλήρη εξάτμιση της αιθανόλης, το ίζημα αναδιαλύεται στον επιθυμητό όγκο ddH₂O και ακολουθεί φωτομετρικός προσδιορισμός της συγκέντρωσης του δείγματος.

2.2.1.4 Ποσοτικός και ποιοτικός προσδιορισμός DNA

Η συγκέντρωση του νουκλεϊκού οξέος στο υδατικό διάλυμα προσδιορίστηκε φωτομετρικά με μέτρηση σε φασματοφωτόμετρο υπεριώδους ορατού (UV/VIS spectrometer Lambda Bio, Perkin-Elmer). Η ποσοτικοποίηση γίνεται από την τιμή της απορρόφησης στα 260 nm (A260), όπου απορροφούν οι αζωτούχες βάσεις (Kozutsumi *et al.* 1989). Ακολούθησε μέτρηση της απορρόφησης στα 280nm και 230 nm. Πληροφορίες για την καθαρότητα του DNA (Sambrook *et al.* 1989) παρέχουν ο λόγος A260/280 (για τιμή ~1,8 θεωρείται γενικώς αποδεκτή ενός καθαρού παρασκευάσματος DNA) -αν ο λόγος είναι αισθητά χαμηλότερος μπορεί να υποδεικνύει την παρουσία πρωτεΐνης, φαινόλης ή άλλων ουσιών που απορροφούν ισχυρά στα 280 nm ή κοντά στα 280 nm), και δευτερευόντως ο λόγος A260/230 (με τιμή συνήθως εντός του εύρους 1,8-2,2).

2.2.1.5 Ηλεκτροφόρηση DNA

Η μέθοδος της ηλεκτροφόρησης δείγματος DNA σε πήκτωμα αγαρόζης χρησιμοποιείται για την ταυτοποίηση, το διαχωρισμό ή την απομόνωση μορίων DNA με βάση το μέγεθός τους. Μαζί με τα προς εξέταση δείγματα, αναλύεται ενδεδειγμένος δείκτης γνωστών μοριακών μεγεθών προκειμένου να γίνει εφικτός ο προσδιορισμός του μοριακού μεγέθους των προς εξέταση δειγμάτων. Το μοριακό μέγεθος προσδιορίζεται μετά από έκθεση του πηκτώματος σε υπεριώδη ακτινοβολία (260-360 nm), λόγω της ενσωμάτωσης της φθορίζουσας χρωστικής βρωμιούχου αιθιδίου ανάμεσα στις αζωτούχες βάσεις της διπλής έλικας του DNA (Mc Donnell *et al.* 1977, Southern 1979). Ρόλο στην θέση της μετανάστευσης εκτός από το μοριακό μέγεθος έχουν κ άλλοι παράγοντες όπως η διαμόρφωση του μορίου του DNA, η συγκέντρωση της αγαρόζης, η ένταση του ηλεκτρικού πεδίου, το ρυθμιστικό διάλυμα ηλεκτροφόρησης.

Η παρασκευή του πηκτώματος γίνεται όταν κατάλληλη ποσότητα αγαρόζης (συνήθως τελική συγκέντρωση 0,8-1,2 % κ.β. που χρησιμοποιήθηκε και στην παρούσα διατριβή) διαλυθεί πλήρως σε ρυθμιστικό διάλυμα TBE υπό θέρμανση και μετά την προσθήκη βρωμιούχου αιθιδίου (0,5 μg/mL) αφηθεί να πήξει σε θερμοκρασία δωματίου μέσα στο φρεάτιο της συσκευής ηλεκτροφόρησης μαζί με κατάλληλο καλούπι για τη δημιουργία των φρεατιδίων στο πήκτωμα. Στα δείγματα DNA προστίθεται διάλυμα φόρτωσης και

τοποθετούνται στα φρεατίδια του οριζόντιου πήκτωματος. Ακολουθεί ηλεκτροφόρηση σε ρυθμιστικό διάλυμα TBE με εφαρμογή ρεύματος σταθερής τάσης 100 V σε θερμοκρασία δωματίου. Μετά την ολοκλήρωση της ηλεκτροφόρησης, το πήκτωμα εκτίθεται σε υπεριώδη ακτινοβολία, οπότε διακρίνονται φθορίζουσες ζώνες σε συγκεκριμένες θέσεις. Για την απομόνωση κάποιου από τα τμήματα DNA, μπορεί να αποκοπεί το τμήμα του πηκτώματος που περιέχει την επιθυμητή ζώνη και να ακολουθήσει ο καθαρισμός του με συστηματοποιημένη χρωματογραφική μέθοδο (§2.2.1.2β).

2.2.1.6 Αντίδραση πέψης δίκλωνου DNA με περιοριστικές ενδονουκλεάσες

Πραγματοποιήθηκαν αντιδράσεις πέψης του DNA με τις επιλεγμένες κάθε φορά περιοριστικές ενδονουκλεάσες για τη γραμμικοποίηση των πλασμιδίων, την παρασκευή και την απομόνωση τμημάτων δίκλωνου DNA με τα επιθυμητά μονόκλωνα άκρα (από τους φορείς κλωνοποίησης ή τα προϊόντα της PCR) και μετά από υποκλωνοποίηση των τμημάτων αυτών στους φορείς έκφρασης για τον έλεγχο των πλασμιδιακών κατασκευών. Οι συνθήκες των αντιδράσεων όπως ποσότητα ενζύμου, ρυθμιστικό διάλυμα, συγκέντρωση άλατος, θερμοκρασία και διάρκεια επώασης, ήταν εκείνες που αναγράφονται στα συνοδευτικά φυλλάδια των ενζύμων σύμφωνα με τις οδηγίες της παρασκευάστριας εταιρείας (New England Biolabs). Γενικά η σχέση που χρησιμοποιήθηκε ήταν 2 μονάδες δραστηριότητας περιοριστικού ενζύμου για κάθε 1 μg DNA σε 20 μL τελικού όγκου αντίδρασης. Μετά το πέρας της επώασης διάρκειας 1-6 ώρες, ακολούθησε ηλεκτροφόρηση του δείγματος σε πήκτωμα αгарόζης, είτε για τον αναλυτικό έλεγχο των προϊόντων των αντιδράσεων πέψης, είτε για την απομόνωση του επιθυμητού τμήματος DNA.

2.2.1.7 Αλυσιδωτή αντίδραση πολυμεράσης (PCR)

Η αλυσιδωτή αντίδραση πολυμεράσης (Polymerase chain reaction, PCR) είναι ίσως η ευρύτερα χρησιμοποιούμενη μέθοδος της μοριακής γενετικής, με αναρίθμητες εφαρμογές τόσο σε ερευνητικό όσο και σε διαγνωστικό επίπεδο. Ανακαλύφθηκε το 1983 από τον βιοχημικό Karry Mullis ο οποίος τιμήθηκε με το βραβείο Nobel χημείας για αυτό. Είναι μια ενζυμική μέθοδος ενίσχυσης συγκεκριμένων τμημάτων γενετικού υλικού *in vitro* (Mullis & Faloona 1987).

Οι αντιδράσεις πραγματοποιήθηκαν με τη χρήση συνθετικών ολιγονουκλεοτιδίων, ως εκκινητές (primers) για τον πολυμερισμό συμπληρωματικών προς τη μήτρα DNA αλυσίδων (DNA template), του πλασμιδίου. Το κύριο προϊόν των αντιδράσεων είναι ένα γραμμικό τμήμα δίκλωνου DNA, του οποίου τα άκρα ορίζονται από τα 5'-άκρα των δύο εκκινητών. Σε μια κλαστική PCR χρησιμοποιούνται δύο εκκινητές. Ο ένας είναι συμπληρωματικός με την περιοχή που βρίσκεται ακριβώς ανοδικά της αλληλουχίας-στόχου στον ένα κλώνο του DNA και συχνά αναφέρεται ως ανοδικός (upstream) ή εμπρόσθιος (forward) εκκινητής. Ο άλλος είναι συμπληρωματικός με την περιοχή που βρίσκεται ακριβώς καθοδικά της αλληλουχίας-στόχου στον άλλο κλώνο του DNA και συχνά αναφέρεται ως καθοδικός (downstream) ή ανάστροφος (reverse) εκκινητής. Στην παρούσα διατριβή οι εκκινητές που χρησιμοποιήθηκαν, είχαν σχεδιαστεί ώστε να διαθέτουν και αλληλουχίες αναγνώρισης από κατάλληλα περιοριστικά ένζυμα προκειμένου να καθίσταται εφικτή η υποκλωνοποίηση στον πλασμιδιακό φορέα έκφρασης. Η αντίδραση λαμβάνει χώρα παρουσία μοριακής περίσσειας των εκκινητών και των τεσσάρων τριφωσφορικών δεσοξυριβονουκλεοτιδίων (dATP, dTTP, dGTP και dCTP), τα οποία αναφέρονται ως dNTPs. Το ένζυμο που χρησιμοποιήθηκε για την ενίσχυση αλληλουχιών DNA είναι η υψηλής πιστότητας KAPA HiFi HotStart DNA πολυμεράση (KAPA Biosystems), που εκτός από την ενεργότητα 5'-3' DNA πολυμεράσης, έχει και ενεργότητα 3'-5' εξωνουκλεάσης ώστε να επιδιορθώνει τυχόν λάθος κατά την ενσωμάτωση των νουκλεοτιδίων, και που λόγω της τελευταίας παράγει γραμμικά μόρια DNA με μη κολλώδη δίκλιωνα άκρα (blunt end). Η προετοιμασία των αντιδράσεων γίνεται με προσεκτικό χειρισμό για την εξασφάλιση της καθαρότητας των αντιδράσεων υπό συνθήκες που απέτρεπαν τη μόλυνση τους από εξωγενές DNA. Όλα τα αντιδραστήρια της αλυσιδωτής αντίδρασης της πολυμεράσης φυλάσσονταν στους -20°C και ο χειρισμός τους κατά την προετοιμασία των αντιδράσεων γινόταν στον πάγο. Ως μάρτυρας συμπεριλήφθηκε και μια αντίδραση χωρίς γονιδιωματικό DNA, προκειμένου να αποκλειστεί η πιθανότητα μολύνσεων.

Για την ενίσχυση DNA αλληλουχιών με PCR, χρησιμοποιήθηκαν γενικά οι ακόλουθες συνθήκες: 1 ng μήτρας πλασμιδιακού DNA, 0,3 μM έκαστου εκκινητή, 0,3 mM έκαστου από τα dNTPs, μισή μονάδα ενεργότητας KAPA HiFi Hot Start DNA πολυμεράσης και εξειδικευμένο συνοδευτικό ρυθμιστικό διάλυμα της εταιρείας το οποίο μεταξύ άλλων περιέχει τελική συγκέντρωση 2,0mM MgCl₂ ανά αντίδραση που συμπληρώθηκε με υπερκάθαρο ύδωρ μέχρι τελικού όγκου των 25 μL.

Οι αντιδράσεις πραγματοποιήθηκαν σε θερμικό κυκλοποιητή. Η αρχική θερμοκρασία ήταν 95°C για 3 λεπτά, προκειμένου να αποδιαταχθεί πλήρως το δίκλινο μόριο του πλασμιδιακού DNA. Ακολούθησαν 25 κύκλοι ενίσχυσης με τις ακόλουθες συνθήκες:

Αποδιάταξης	98°C	για 20 s
Υβριδισμού	56-72°C	για 15 s
Πολυμερισμού	72°C	για 60 s/kbp που πρόκειται να ενισχυθεί

Μετά την ολοκλήρωση των κύκλων ενίσχυσης ακολούθησε διατήρηση της θερμοκρασίας στους 72°C για 5 λεπτά, ώστε να ολοκληρωθεί η επιμήκυνση σε προϊόντα που η διαδικασία πολυμερισμού ήταν ημιτελής. Με την ολοκλήρωση του προγράμματος η θερμοκρασία ρυθμίστηκε στους 4°C ώστε να σταματήσει η αντίδραση και να διατηρηθούν τα δείγματα. Ακολούθως, τα προϊόντα της PCR αναλύθηκαν με ηλεκτροφόρηση σε πήκτωμα αγαρόζης 1%κ.β. για να επιβεβαιωθεί ότι η διαδικασία πολλαπλασιασμού του DNA στόχου είχε το αναμενόμενο αποτέλεσμα είτε στη συνέχεια χρησιμοποιήθηκαν σε αντιδράσεις περιορισμού και σύνδεσης για υποκλωνοποίηση στον φορέα έκφρασης μετά από καθαρισμό.

2.2.1.8 Υποκλωνοποίηση δίκλωνου τμήματος DNA σε πλασμιδιακό φορέα

Για την υποκλωνοποίηση δίκλωνου τμήματος DNA σε πλασμιδιακό φορέα, αφού πρώτα έγιναν αντιδράσεις πέψης με τις κατάλληλες κάθε φορά περιοριστικές ενδονουκλεάσες, ακολούθησε απομόνωση του ανοιχτού γραμμικοποιημένου φορέα έκφρασης, και του ενθέματος το οποίο είχε απομονωθεί από τους φορείς κλωνοποίησης και τα προϊόντα της PCR. Στην παρούσα διατριβή προχώρησε η διαδικασία στην αντίδραση σύνδεσης (ligation), χωρίς να προηγηθεί αποφωσφορυλίωση του γραμμικού πλασμιδιακού DNA. Η αντίδραση σύνδεσης ανοιχτού, γραμμικού πλασμιδιακού DNA με το προς υποκλωνοποίηση τμήμα DNA πραγματοποιείται με τη χρήση ποσότητας 20 ng του πλασμιδίου και τριπλάσιας μοριακής αναλογίας του ενθέματος. Τα αντιδρώντα μόρια DNA προστίθενται μαζί με 2 μ L (1U/ μ L) του ενζύμου T4 DNA λιγάσης σε τελικό όγκο 20 μ L ρυθμιστικού διαλύματος 50 mM Tris-HCl, pH 7,5, 10 mM MgCl₂, 1 mM ATP, 1 mM DTT, 5% κ.β. PEG 8.000. Ακολουθεί επώαση του μίγματος σε θερμοκρασία δωματίου για 1 h. Τα προϊόντα των αντιδράσεων σύνδεσης χρησιμοποιούνται άμεσα για το μετασχηματισμό επιδεκτικών βακτηριακών κυττάρων *E. coli* του στελέχους TOP10F' (§2.2.2.1). Τα βήματα εν συντομία που ακολουθούν είναι καλλιέργειες του βακτηρίου, απομόνωση από έναν κλώνο πλασμιδιακού DNA, έλεγχος των κλώνων για την επιλογή αυτών που φέρουν το ορθό ανασυνδυασμένο πλασμίδιο, με χαρακτηριστικές πέψεις από περιοριστικές ενδονουκλεάσες και ανάλυση των προϊόντων των αντιδράσεων πέψης με ηλεκτροφόρηση σε πήκτωμα

αγαρόζης (0,8-1,2% κ.β.) και τέλος το πλασμιδιακό DNA που απομονώνεται από ένα θετικό κλώνο ελέγχεται στην συνέχεια με αλληλούχιση για να επιβεβαιωθεί η ορθότητά του.

2.2.1.9 Αλληλούχιση DNA (DNA sequencing)

Ο προσδιορισμός της νουκλεοτιδικής αλληλουχίας DNA των ανασυνδυασμένων φορέων έκφρασης, ώστε να επιβεβαιωθεί η ορθή κατασκευή τους, έγινε από εταιρείες που δραστηριοποιούνται στο χώρο της μοριακής γενετικής με τη χρήση κατάλληλων επιλεγμένων εκκινητών ευρείας χρήσης (universal primers). Η ανάλυση των αποτελεσμάτων της αλληλούχισης γίνεται ώστε να επιβεβαιωθεί ότι η κατασκευασμένη κασέτα έκφρασης είναι η θεωρητικά αναμενόμενη και ότι δε φέρει ανεπιθύμητες μεταλλάξεις.

2.2.2. Μέθοδοι Βιοχημείας

2.2.2.1 Έκφραση των πρωτεϊνών από μετασχηματισμένα βακτηριακά κύτταρα

A. Υγρές και στερεές καλλιέργειες βακτηριακών κυττάρων *E. coli*.

Χρησιμοποιήθηκαν στελέχη του *E. coli*, TOP10F' για την παρασκευή των πλασμιδίων, και τα BL21(DE3), C41(DE3), C43(DE3) και Rosetta 2(DE3) για την έκφραση, σε καλλιέργειες υγρού ή στερεού θρεπτικού μέσου LB παρουσία ή απουσία αντιβιοτικού κατά περίπτωση. Η ανάπτυξή τους πραγματοποιήθηκε στους 37°C, σε επωαστικούς θαλάμους. Για τις υγρές καλλιέργειες, η επώαση πραγματοποιήθηκε υπό ανάδευση (200 rpm) σε πλαστικούς αποστειρωμένους σωλήνες μιας χρήσης των 15 mL ή 50 mL ή σε υάλινες αποστειρωμένες κωνικές φιάλες των 0,5L ή 1L ή 2 L, για 12-16 h. Όλοι οι χειρισμοί έγιναν υπό στείρες συνθήκες. Τα βακτήρια διατηρούνται για μεγάλα χρονικά διαστήματα με αποθήκευση στους -80°C παρουσία 20% γλυκερόλης στο κατάλληλο θρεπτικό μέσο.

B. Μετασχηματισμός βακτηρίων με τη χημική μέθοδο του CaCl₂

Τα πλασμίδια χρησιμοποιούνται για το μετασχηματισμό βακτηρίων, ώστε να ακολουθήσει η παραγωγή και απομόνωσή τους. Για το μετασχηματισμό βακτηρίων με

πλασμιδιακό DNA χρησιμοποιήθηκε η χημική μέθοδος του χλωριούχου ασβεστίου (Sambrook *et al.* 1989). Με τη μέθοδο αυτή η ικανότητα μετασχηματισμού είναι περίπου $0,5-2,0 \times 10^7$ βακτήρια/μg υπερελικωμένου πλασμιδιακού DNA, που ενώ είναι μικρότερη από την τιμή της ικανότητας μετασχηματισμού που προκύπτει με τη μέθοδο της ηλεκτροδιάτρησης, παραμένει για τις συνήθεις διαδικασίες κλωνοποίησης ικανοποιητική. Ακολουθεί η περιγραφή της πορείας μετασχηματισμού των βακτηρίων με τη μέθοδο η οποία περιλαμβάνει δύο στάδια.

- Παρασκευή επιδεκτικών κυττάρων (competent cells) βακτηρίων με χρήση χλωριούχου ασβεστίου (CaCl₂)

Από μια καλλιέργεια σε στερεό θρεπτικό μέσο βακτηρίων του στελέχους TOP10F', γίνεται εμβολιασμός μίας αποικίας σε 5 mL LB και ακολουθεί επώαση υπό ανάδευση, στους 37°C για για 12-16 h. Στη συνέχεια με 2 mL από αυτή την καλλιέργεια εμβολιάζονται 100 mL LB, και ακολουθεί επώαση υπό ανάδευση στους 37°C, μέχρις ότου $O.D._{600} = 0,6-0,9$ (~2 h). Ακολουθεί μεταφορά της υάλινης φιάλης, μέσα στην οποία βρίσκεται η καλλιέργεια, σε πάγο (0°C), όπου παραμένει για 30 min. Έπειτα γίνεται μεταφορά της καλλιέργειας σε δύο στείρους προψυγμένους πλαστικούς σωλήνες των 50 mL και φυγοκέντρηση (4.000 rpm, 10 min, 4°C). Το ίζημα των βακτηρίων αναδιαλύεται σε 25 mL στείρου παγωμένου διαλύματος 50 mM CaCl₂. Ακολουθεί φυγοκέντρηση του εναιωρήματος στις 4.000 rpm για 10 min, στους 4°C και το ίζημα των βακτηρίων αναδιαλύεται σε 5 mL στείρου προψυγμένου διαλύματος 50 mM CaCl₂. Ακολουθεί τοποθέτηση του εναιωρήματος στον πάγο για 2-3 h. Τέλος είτε ακολουθεί άμεσα ο μετασχηματισμός των επιδεκτικών πλέον βακτηριακών κυττάρων ή αυτά αποθηκεύονται ανά 100 μL στους -80°C παρουσία 20% γλυκερόλης για μελλοντική χρήση.

- Μετασχηματισμός επιδεκτικών κυττάρων βακτηρίων

Το εναιώρημα των επιδεκτικών βακτηριακών κυττάρων αποψύχεται εν ηρεμία εντός πάγου (για ~45 min), στην περίπτωση χρήσης αποθηκευμένων επιδεκτικών κυττάρων. Στη συνέχεια γίνεται ανάμειξη 2μL κατάλληλα αραιωμένου με ddH₂O πλασμιδιακού DNA (~5 ng/μL) ή με το προϊόν της αντίδρασης σύνδεσης (§2.2.1.8), με 100 μL εναιωρήματος επιδεκτικών βακτηρίων. Ακολουθεί επώαση του μείγματος στον πάγο (0°C) για 45 min και θερμικό σοκ με επώαση σε 42°C (σε υδατόλουτρο) αυστηρά για 90 s. Το μείγμα τοποθετείται άμεσα σε πάγο και επωάζεται για 2-5 min. Γίνεται εν συνεχεία, προσθήκη στο μείγμα 1 mL LB και επώαση, υπό ανάδευση στους 37°C για 1 h. Τέλος, τα βακτήρια επιστρώνονται σε

κατάλληλο στερεό θρεπτικό μέσο επιλογής (π.χ. εμπλουτισμένο με αντιβιοτικό επιλογής) σε τρυβλίο Petri και επωάζονται για ~16 h στους 37°C, με αποτέλεσμα την εμφάνιση αποικιών επιτυχώς μετασχηματισμένων βακτηρίων.

2.2.2.2 Καλλιέργειες έκφρασης σε βακτήρια

Για την παρασκευή επιδεκτικών κυττάρων BL21(DE3), C41(DE3), C43(DE3) και Rosetta 2(DE3) και το μετασχηματισμό τους με τους φορείς έκφρασης, ακολουθήθηκε η ίδια πορεία που περιγράφηκε για το στέλεχος TOP10F', με τη διαφορά του εμπλουτισμού ή μη του θρεπτικού υλικού LB με το κατάλληλο αντιβιοτικό. Η επιλογή των μετασχηματισμένων κλώνων έγινε σε θρεπτικό μέσο εμπλουτισμένο με 50 µg/mL καναμυκίνη, λόγω του γονιδίου ανθεκτικότητας που φέρει ο φορέας έκφρασης που χρησιμοποιήθηκε.

Από ένα τρυβλίο Petri όπου αναπτύσσονται βακτήρια του εκάστοτε στελέχους μετασχηματισμένα με τον ανασυνδυασμένο φορέα έκφρασης pET26b(+) ή pETM-44, έγινε εμβολιασμός μίας και μόνο αποικίας σε 10 mL LB εμπλουτισμένο με 50 µg/mL καναμυκίνη και ακολούθησε επώαση, στους 37°C για 12-16 h υπό ανάδευση. Ακολούθως, η καλλιέργεια αυτή αραιώθηκε σε 1L LB εμπλουτισμένο με 50 µg/mL καναμυκίνη και επωάστηκε υπό ανάδευση στους 37°C, μέχρις ότου η οπτική πυκνότητα στα 600 nm γίνει 0,6 (~2 h). Ακολούθησαν οι διαδικασίες επαγωγής έκφρασης της ανασυνδυασμένης πρωτεΐνης με προσθήκη IPTG σε δοκιμές της συγκέντρωσής του, καθώς και του χρόνου και της θερμοκρασίας επώασης. Μετά το πέρας της επαγωγής έκφρασης, έγινε συλλογή των βακτηρίων με φυγοκέντρηση (5000 rpm, 15 min, 4°C) και είτε ακολούθησε άμεσα λύση των κυττάρων και απομόνωση της ανασυνδυασμένης πρωτεΐνης είτε αποθήκευση στους -80°C για μελλοντική χρήση.

2.2.2.3 Λύση βακτηρίων και απομόνωση της ανασυνδυασμένης πρωτεΐνης

Μετά την επαγωγή έκφρασης και τη συλλογή των κυττάρων, οι διαδικασίες διεξήχθησαν στους 4°C. Το ίζημα των βακτηριακών κυττάρων αναδιαλύεται σε 25 ml ρυθμιστικό διάλυμα λύσης (50 mM HEPES pH 7,5, 300 mM NaCl, 5% γλυκερόλη και μίγμα αναστολέων πρωτεασών) για κάθε λίτρο καλλιέργειας. Η λύση των κυττάρων έγινε με χρήση συσκευής υπερήχων με μικρο-ακρορύγχιο (microtip sonicator), με 10 κύκλους παλμών

υπερήχων έντασης 6-9 Watt για 10 s με μεσοδιαστήματα ηρεμίας 30 s. Στην συνέχεια, παραλαμβάνεται το υπερκείμενο μετά από μία σύντομη φυγοκέντρηση (2500 rpm, 5 min). Ακολούθησε επώαση του μίγματος για 2 h με ήπια ανάδευση υπό περιστροφή. Ακολούθησε φυγοκέντρηση και στο υπερκείμενο προστέθηκαν 3 ml σφαιριδίων Ni-NTA (εξισορροπημένων σε διάλυμα λύσης). Το μίγμα επώαστηκε για 2 h υπό ήπια ανάδευση με συνεχή περιστροφή του δοκιμαστικού σωλήνα. Στην συνέχεια, έγινε πακετάρισμα των σφαιριδίων Ni-NTA σε γυάλινη κυλινδρική στήλη χρωματογραφίας (BioRad). Η στήλη των σφαιριδίων Ni-NTA πλύθηκε αρχικά με 10 όγκους στήλης διαλύματος λύσης και έπειτα με 10 όγκους στήλης διαλύματος πλύσης 50 mM Tris-HCl pH 8, 300 mM NaCl, 5% γλυκερόλη και 10 mM ιμιδαζολίου. Η ανασυνδυασμένη πρωτεΐνη παραλήφθηκε σε 3 όγκους στήλης διαλύματος έκλουσης (50 mM Tris-HCl pH 8, 300 mM NaCl, 5% γλυκερόλη και 200 mM ιμιδαζολίου). Το πρωτεϊνικό διάλυμα συμπυκνώθηκε με διαδοχικές φυγοκεντρήσεις (2500 rpm, 5 min) με χρήση δοκιμαστικών σωλήνων που φέρουν φίλτρα υπερδιήθησης με διάμετρο πόρων 100 kDa (Amicon), σε τελικό όγκο ώστε η οπτική πυκνότητα στα 280 nm να είναι 1 με 2. Το συμπυκνωμένο πρωτεϊνικό διάλυμα υποβλήθηκε σε χρωματογραφία μοριακού αποκλεισμού σε ρυθμιστικό διάλυμα έκλουσης (χωρίς ιμιδαζόλιο), με χρήση της στήλης Superdex75 ή Superdex200 (GE Healthcare).

2.2.2.4 Χρωματογραφία συγγένειας (Affinity chromatography)

Το πρώτο στάδιο καθαρισμού των ανασυνδυασμένων πρωτεϊνών είναι βασισμένο στην ύπαρξη έξι καταλοίπων ιστιδίνης εν σειρά (6x HisTag) στο καρβοξυτελικό ή το αμινοτελικό άκρο τους. Σε ρυθμιστικό διάλυμα με pH 8, αυτές οι ανασυνδυασμένες πρωτεΐνες προσδένονται με υψηλή συγγένεια σε στήλη σφαιριδίων αγαρόζης-νικελίου συζευγμένης (Ni-NTA), μέσω της ετικέτας ιστιδινών (Hochuli *et al.* 1987). Η έκλουση των προσδεμένων πρωτεϊνών από τη στήλη Ni-NTA πραγματοποιείται με ρυθμιστικό διάλυμα αυξανόμενης συγκέντρωσης ιμιδαζολίου, το οποίο ανταγωνίζεται τους ιμιδαζολικούς δακτυλίους των ιστιδινών για την πρόσδεση στα ιόντα νικελίου της στήλης.

Για τον καθαρισμό των ανασυνδυασμένων πρωτεϊνών που παρήχθησαν με χρωματογραφία συγγένειας, χρησιμοποιήθηκαν σφαιρίδια Ni-NTA (Qiagen) σε εναιώρημα που περιέχει 50% κ.ο. σφαιρίδια και 20% αιθανόλη. Πριν από τη χρήση τους, έγιναν δύο πλύσεις των σφαιριδίων με υπερκάθαρο ύδωρ και δύο φορές με διάλυμα εξισορρόπησης (όπως π.χ. 50 mM Tris-HCl pH 8, 300 mM NaCl και 10 mM ιμιδαζολίου) ενώ κάθε φορά

ακολουθούσε φυγοκέντρηση (600 rcf, 5 min, 4°C) και πλήρης απομάκρυνση του υπερκειμένου.

Η εξισορροπημένη στήλη επωάζεται με το πρωτεϊνικό διάλυμα που περιέχει τις ανασυνδυασμένες πρωτεΐνες, υπό ανάδευση, στους 4°C και για κατάλληλο χρονικό διάστημα (2-12 h), ώστε αυτές να προσδεθούν στην στήλη. Τα σφαιρίδια τοποθετούνται σε γυάλινη στήλη χρωματογραφίας και το πρωτεϊνικό διάλυμα αφήνεται να περάσει μέσα από τη στήλη. Η στήλη πλένεται με ένα ή περισσότερα διαλύματα πλύσης, ίδιας σύστασης με το διάλυμα εξισορρόπησης αλλά με αυξανόμενη συγκέντρωση ιμιδαζολίου, ώστε να απομακρυνθούν οι πρωτεΐνες που έχουν δέσει μη-ειδικά στη στήλη. Η συγκέντρωση ιμιδαζολίου στο διάλυμα πλύσης είναι γενικά χαμηλή (10-50 mM). Η πρωτεΐνη – στόχος παραλαμβάνεται στο διάλυμα έκλουσης, που τυπικά περιέχει ιμιδαζόλιο σε υψηλή συγκέντρωση (200-500 mM).

2.2.2.5 Χρωματογραφία μοριακής διήθησης (Size exclusion chromatography)

Τη χρωματογραφία συγγένειας, ακολούθησε ένα δεύτερο στάδιο καθαρισμού, με χρωματογραφία μοριακού αποκλεισμού (size-exclusion chromatography, SEC, Barth *et al.* 1996), σε σύστημα υγρής χρωματογραφίας (FPLC). Στο στάδιο αυτό επιχειρείται ο διαχωρισμός της ανασυνδυασμένης πρωτεΐνης από προσμίξεις που (ενδεχομένως) βρίσκονται στα δείγματα έκλουσης από τη στήλη Ni-NTA αλλά και η απομόνωση ενός ομοιογενούς πληθυσμού της πρωτεΐνης – στόχου. Ο διαχωρισμός αυτός επιτυγχάνεται βάσει του διαφορετικού μεγέθους και σχήματός τους κατά τη διέλευσή του πρωτεϊνικού διαλύματος από στήλη μοριακού αποκλεισμού, το υλικό της οποίας φέρει συγκεκριμένη διάμετρο πόρων, ώστε να διαχωρίζει τα μόρια σε ένα συγκεκριμένο εύρος μοριακών μεγεθών.

Οι στήλες που χρησιμοποιήθηκαν, είχαν προηγουμένως εξισορροπηθεί στο κατάλληλο κάθε φορά ρυθμιστικό διάλυμα φιλτραρισμένο μέσω φίλτρων 0,2 μm και απαερωμένο με χρήση λουτρού υπερήχων. Τα δείγματα είχαν συμπυκνωθεί στην επιθυμητή συγκέντρωση και φιλτραριστεί μέσω φίλτρου διαμέτρου πόρων 0,2 μm και στην συνέχεια ενέθηκαν (100-400 μL κάθε φορά) στο σύστημα FPLC (με χρήση θηλιάς ένεσης του 1 mL). Ακολούθησε χρωματογραφία με συνεχή ροή διαλύματος ίση προς 0,5 mL/min και συλλογή κλασμάτων όγκου 0,5 mL. Μεταξύ της εξόδου της στήλης και του συλλέκτη κλασμάτων, παρεμβλήθηκε φασματοφωτόμετρο υπεριώδους για την ανίχνευση της εκλούμενης πρωτεΐνης στα κλάσματα.

2.2.2.6. Ποσοτική ανάλυση καθαρισμένων πρωτεϊνών

A. Φασματοσκοπικές μέθοδοι για την ποσοτικοποίηση πρωτεϊνών

Η συγκέντρωση της πρωτεΐνης στο διάλυμα προσδιορίζεται σε φασματοφωτόμετρο ορατού – υπεριώδους και με τη χρήση ακίδας (NanoDrop), όπου μετράται η απορρόφηση του πρωτεϊνικού δείγματος στα 280 nm (A280), όπου απορροφούν κυρίως τα αρωματικά αμινοξικά κατάλοιπα των πρωτεϊνών. Η μέθοδος αυτή έχει τα πλεονεκτήματα ότι είναι σύντομη, μη καταστρεπτική για την πρωτεΐνη και απαιτεί μικρή ποσότητα δείγματος (2 μL). Στη συνέχεια, η συγκέντρωση της πρωτεΐνης στο δείγμα υπολογίζεται σύμφωνα με τον εμπειρικό τύπο: $A_{280} (1 \text{ mg} / \text{mL}) = (5500 \cdot n_{\text{Trp}} + 1490 \cdot n_{\text{Tyr}} + 62,5 \cdot n_{\text{Cys}}) \cdot M^{-1}$ (Pace *et al.* 1995), όπου n_{Trp} , n_{Tyr} και n_{Cys} , ο αριθμός των αμινοξικών καταλοίπων τρυπτοφάνης, τυροσίνης και κυστεΐνης, αντίστοιχα, στην πρωτεΐνη που εξετάζεται, και M η μοριακή συγκέντρωση του πρωτεϊνικού δείγματος.

B. Χρωματομετρική μέθοδος Bradford

Η χρωματογραφική μέθοδος Bradford βασίζεται στην ιδιότητα της χρωστικής Coomassie Brilliant Blue G-250 να προσδένεται σε βασικά και αρωματικά αμινοξικά κατάλοιπα (κυρίως κατάλοιπα αργινίνης), που έχει αποτέλεσμα έγχρωμο προϊόν, έντασης, ανάλογης της συγκέντρωσης της πρωτεΐνης (Bradford 1976). Η τιμή της συγκέντρωσης της πρωτεΐνης ενός δείγματος προκύπτει από την τιμή απορρόφησης του έγχρωμου προϊόντος στα 595 nm (A595) και τη σύγκριση της μετρούμενης αυτής τιμής με πρότυπη καμπύλη αναφοράς, όπως έχει προκύψει από τιμές απορρόφησης A595 συμπλόκου χρωστικής με γνωστές συγκεντρώσεις πρότυπης πρωτεΐνης. Η πρότυπη καμπύλη με χρήση 6-8 διαφορετικών συγκεντρώσεων αλβουμίνης ορού βοός (BSA) κατασκευάζεται αρχικά. Διαλύματα 0,2-0,9 mg/mL BSA σε ddH₂O παρασκευάζονται και 20 μL από κάθε δείγμα προστίθενται σε 1 mL αντιδραστηρίου Bradford (BioRad protein assay, αραιωμένου 1:5 με ddH₂O) και μετά από επώαση 5 min σε θερμοκρασία δωματίου, μετράται η απορρόφηση των διαλυμάτων αυτών στα 595 nm. Ακολουθεί η κατασκευή πρότυπης καμπύλης που συσχετίζει τις μετρούμενες τιμές A595 με τις γνωστές συγκεντρώσεις των δειγμάτων BSA, με μία γραμμική σχέση. Για τον προσδιορισμό της συγκέντρωσης πρωτεϊνικού δείγματος, 20 μL του προς μέτρηση δείγματος (ή κατάλληλης αραιώσης) προστίθενται σε 1 mL αντιδραστηρίου

Bradford και γίνεται επώαση 5 min σε θερμοκρασία δωματίου. Έπειτα, μετράται η A595 και από την τιμή αυτή (και τον παράγοντα αραίωσης) υπολογίζεται η συγκέντρωση (mg/mL) του πρωτεϊνικού δείγματος, με βάση την πρότυπη καμπύλη αναφοράς. Κάθε φορά χρησιμοποιείται αρνητικός μάρτυρας (20 μ L του διαλύματος στο οποίο βρίσκεται κάθε φορά η προς μελέτη πρωτεΐνη) για τον προσδιορισμό της απορρόφησης υποβάθρου.

2.2.2.7 Ηλεκτροφόρηση πρωτεϊνών σε πήκτωμα ακρυλαμίδιου υπό αποδιατακτικές συνθήκες

Ο διαχωρισμός και η ταυτοποίηση των πρωτεϊνών μετά από ηλεκτροφόρηση σε κάθετο πήκτωμα πολυακρυλαμίδιου, παρουσία θειϊκού δωδεκυλικού νατρίου (SDS) γίνεται με βάση το μοριακό τους βάρος (Laemmli 1970, Brunelle & Green 2014). Το πήκτωμα συμπύκνωσης (stacking gel) περιέχει 0,125 M Tris-HCl, pH 6,8, 0,1% κ.β. SDS και 4% κ.β. ακρυλαμίδιο. Το πήκτωμα διαχωρισμού (separating gel), περιέχει 0,375 M Tris-HCl, pH 8,8, 0,1 % κ.β. SDS και 12 % κ.β. ακρυλαμίδιο, 1% κ.β. APS και 0,04% κ.ο. TEMED.

Τα δείγματα αναμειγνύονται με διάλυμα φόρτωσης πρωτεϊνών, το οποίο περιέχει SDS και β -μερκαπτοαιθανόλη και ακολουθεί επώαση για 5 min σε θερμοκρασία 100°C και οι πρωτεΐνες αποδιατάσσονται. Στην ηλεκτροφόρηση μαζί με τα προς εξέταση δείγματα, συμπεριλαμβάνεται και ένα δείγμα μείγματος πρωτεϊνών γνωστού μοριακού βάρους, ως μάρτυρας μοριακών μεγεθών. Η ηλεκτροφόρηση πραγματοποιείται υπό σταθερή τάση 120 V, σε ρυθμιστικό διάλυμα ηλεκτροφόρησης πρωτεϊνών, το οποίο περιέχει 0,1 % κ.β. SDS.

Μετά το πέρας της ηλεκτροφόρησης, είτε γίνεται άμεσα εμφάνιση των πρωτεϊνικών ζωνών με διάλυμα χρώσης του πηκτώματος πολυακρυλαμίδιου, είτε το πήκτωμα χρησιμοποιείται για μεταφορά των πρωτεϊνών σε μεμβράνη νιτροκυτταρίνης, ώστε να ακολουθήσει ανοσοαποτύπωση κατά Western (Burnette 1981).

A. Χρώση με Coomassie:

Οι ζώνες των πρωτεϊνών μονιμοποιούνται και εμφανίζονται με διάλυμα χρώσης του πηκτώματος πολυακρυλαμίδιου. Η χρώση του πηκτώματος γίνεται με ήπια ανάδευση σε διάλυμα χρώσης στους 37°C για 30 min. Η εμφάνιση των πρωτεϊνικών ζωνών γίνεται με συνεχείς πλύσεις του πηκτώματος με διάλυμα αποχρωματισμού σε RT, μέχρι την απομάκρυνση της περίσσειας χρωστικής και παραμονή μόνο όσης έχει δεσμευτεί στις πρωτεΐνες.

2.2.2.8 Ανοσοαποτύπωση πρωτεϊνών

A. Ανοσοαποτύπωση πρωτεϊνών μετά από SDS PAGE (Western blotting):

Για την εμφάνιση των πρωτεϊνικών ζωνών με χρήση εξειδικευμένου αντισώματος, οι πρωτεΐνες, κατόπιν του διαχωρισμού τους σε πήκτωμα πολυακρυλαμιδίου, μεταφέρονται σε σταθερό μεμβρανικό υπόστρωμα νιτροκυτταρίνης (Hybond-P, Amersham) με εφαρμογή ηλεκτρικού πεδίου (Burnette 1981). Για τη μετανάστευση των πρωτεϊνών από το πήκτωμα πολυακρυλαμιδίου στη μεμβράνη νιτροκυτταρίνης, χρησιμοποιείται κατάλληλο ρυθμιστικό διάλυμα ηλεκτρομεταφοράς (transfer buffer) και συσκευή ηλεκτρομεταφοράς (Mini-Trans-Blot, BioRad). Η μεταφορά των πρωτεϊνών πραγματοποιείται κατά τη διαβίβαση ρεύματος σταθερής έντασης 300 mA για 1 h στους 4°C.

Μετά την ολοκλήρωση της διαδικασίας της ηλεκτρομεταφοράς, ακολουθεί επώαση της μεμβράνης νιτροκυτταρίνης σε διάλυμα PBS με 3% κ.β. BSA, υπό ανάδευση σε θερμοκρασία δωματίου για 45 min όπου και δεσμεύονται οι ελεύθερες θέσεις της μεμβράνης νιτροκυτταρίνης από την BSA. Ακολουθεί επώαση υπό ανάδευση της μεμβράνης με το πρώτο αντίσωμα, το οποίο είναι κατάλληλα αραιωμένο σε διάλυμα PBS με 0,5 % κ.β. BSA, σε θερμοκρασία δωματίου για 2 h ή σε 4°C κατά τη διάρκεια της νύκτας. Το πρώτο αντίσωμα είναι ειδικό για την πρωτεΐνη – στόχο ή για την ετικέτα ιστιδινών (στην παρούσα εργασία anti-His Ab που έχουν παραχθεί σε ποντίκια) της πρωτεΐνης – στόχου. Μετά το πέρας της επώασης με το πρώτο αντίσωμα, πραγματοποιούνται τρεις διαδοχικές πλύσεις με PBS, διάρκειας 10 min, εκάστη. Ακολουθεί επώαση για 2 h της μεμβράνης, σε θερμοκρασία δωματίου υπό ανάδευση με το δεύτερο αντίσωμα που είναι ειδικό για ανοσοσφαιρίνες που έχουν παραχθεί σε ποντίκια και συζευγμένο με το ένζυμο υπεροξειδάση της ραπανίδος (anti-mouse IgG-HRP), αραιωμένο 1:2.000 σε PBS. Στη συνέχεια, πραγματοποιούνται τρεις διαδοχικές πλύσεις με PBST (βλ. §2.1.4), διάρκειας 10 min κάθε εκάστη. Ακολουθεί επώαση της μεμβράνης νιτροκυτταρίνης σε διάλυμα εμφάνισης με υπόστρωμα DAB. Το H₂O₂, χρησιμοποιείται ως υπόστρωμα από την υπεροξειδάση και έτσι παράγονται ελεύθερες ρίζες οξυγόνου. Παρουσία ελεύθερων ριζών οξυγόνου, η DAB οξειδώνεται προς ίζημα χρώματος καφέ σκούρου εμφανίζοντας τις πρωτεϊνικές ζώνες στις οποίες έχουν προσδεθεί τα δύο αντισώματα. Η αντίδραση πραγματοποιείται παρουσία ιόντων Ni²⁺ (NiCl₂), που αυξάνουν την ευαισθησία της, με αποτέλεσμα την εμφάνιση εντονότερου χρώματος. Οι ζώνες των πρωτεϊνών εμφανίζονται λίγα δευτερόλεπτα μετά την εμφάνιση της νιτροκυτταρίνης στο ρυθμιστικό διάλυμα εμφάνισης. Η αντίδραση σταματά με εμφάνιση της μεμβράνης σε ddH₂O, ώστε να αποφευχθεί η μη ειδική χρώση αυτής.

B. Ανοσοαποτύπωση πρωτεϊνών υπό μη αποδιατακτικές συνθήκες

Σε μία παραλλαγή της μεθόδου, η ανοσοαποτύπωση πραγματοποιείται σε δείγμα πρωτεϊνών προσροφημένων σε μεμβράνη νιτροκυτταρίνης υπό μη αποδιατακτικές συνθήκες (αποτύπωμα κηλίδας, dot blot). Στην περίπτωση αυτή, ένα διάλυμα πρωτεϊνικού δείγματος προσροφάται δια μέσω της μεμβράνης νιτροκυτταρίνης, υπό κενό. Η προσρόφηση πραγματοποιείται σε ειδική συσκευή προσρόφησης 96 θέσεων, όγκου 0,4 mL έκαστη. Η πορεία που ακολουθείται για την ανίχνευση των προσροφημένων πρωτεϊνών είναι ίδια όπως περιγράφεται παραπάνω, κατόπιν του σταδίου της ηλεκτρομεταφοράς. Η μέθοδος αυτή δεν απαιτεί διαχωρισμό των πρωτεϊνών του δείγματος και χρησιμοποιήθηκε για την επιλογή των μετασχηματισμένων κλώνων με την ισχυρότερη έκφραση από καλλιέργειες μικρής κλίμακας όπως υποδηλώνεται αναλογικά από την ένταση εμφάνισης του σήματος της πρωτεΐνης – στόχου.

2.2.3 Κρυστάλλωση

2.2.3.1 Σάρωση συνθηκών κρυστάλλωσης

Η εξεύρεση αρχικών συνθηκών κρυστάλλωσης έγινε με χρήση εμπορικά διαθέσιμων συλλογών συνθηκών καθορισμένης σύστασης (kit). Οι δοκιμές πραγματοποιήθηκαν με τη μέθοδο της διάχυσης ατμών και την τεχνική της επικαθήμενης σταγόνας, ενώ παράλληλα ελέγχθηκαν διαφορετικές θερμοκρασίες (4°C, 10°C, 16°C και 23°C) και διαφορετικές συγκεντρώσεις πρωτεΐνης (2-22 mg/mL). Το διάλυμα πρωτεΐνης αναμίχθηκε με το διάλυμα κρυστάλλωσης (μητρικό διάλυμα) σε αναλογία 2:1 ή 1:1 σχηματίζοντας σταγόνα τελικού όγκου 0,6 ή 0,8 μL αντίστοιχα, η οποία εξισορροπήθηκε έναντι 50 μL μητρικού διαλύματος σε κάθε φρεάτιο ενός δίσκου κρυσταλλώσεως πρωτεϊνών 96 φρεατιδίων, κλεισμένο αεροστεγώς. Για τα αρχικά πειράματα κρυστάλλωσης χρησιμοποιήθηκε αυτόματο ρομποτικό σύστημα (Oryx4, Douglas Instruments, Ltd). Ακολούθως, έγινε ανά τακτά χρονικά διαστήματα παρατήρηση της κάθε θέσης υπό φωτονικό στερεοσκόπιο, προκειμένου να διαπιστωθεί εάν σε κάποια από τις δοκιμαστικές σταγόνες έχουν αναπτυχθεί κρύσταλλοι.

2.2.3.2 Βελτιστοποίηση αρχικών συνθηκών κρυστάλλωσης

Μετά τον προσδιορισμό αρχικών ή υποσχόμενων συνθηκών που οδηγούσαν στο σχηματισμό πρωτεϊνικών κρυστάλλων ή ακόμα και ιζήματος κρυσταλλικής φύσεως, επιχειρήθηκε η βελτιστοποίηση των συνθηκών με σκοπό την ανάπτυξη κρυστάλλων κατάλληλων για συλλογή δεδομένων ακτίνων-X (McPherson 2016). Η διαδικασία αυτή συμπεριλάμβανε τροποποίηση μερικών εσωτερικών παραμέτρων του διαλύματος, σε ένα μικρό εύρος γύρω από την αρχική συνθήκη, όπως το pH και η ιοντική ισχύς, αλλά και εξωτερικών παραμέτρων, όπως η θερμοκρασία, ο όγκος της σταγόνας εξισορρόπησης, η συγκέντρωση της πρωτεΐνης και η τεχνική της κρυστάλλωσης. Επιπλέον, επιχειρήθηκε αντικατάσταση κάποιου εκ των συστατικών του διαλύματος κρυστάλλωσης από χημικά παραμφορείς ουσίες, όπως η χρήση αιθανόλης αντί ισοπροπανόλης και η χρήση PEG 4000 αντί PEG 3350. Δεδομένου ότι, τα μόρια των πρωτεϊνών, ειδικότερα των ανασυνδυασμένα εκπεφρασμένων και απομονωμένων, είναι σχετικά ασταθή, για την κρυσταλλική καταβύθισή τους αποφεύχθηκαν οι ακραίες τιμές pH και θερμοκρασιών. Επιπρόσθετη αιτιολογία για τη συγκεκριμένη επιλογή αποτελεί το γεγονός ότι η κατανόηση των βιολογικών και φυσιολογικών ιδιοτήτων ενός μακρομορίου μέσω της τρισδιάστατης δομής του, είναι πολύ ευκολότερη και λογικότερη, εάν οι συνθήκες ανάπτυξης των κρυστάλλων δεν απέχουν πολύ από το φυσιολογικό περιβάλλον στο οποίο το μόριο είναι δραστικό. Επίσης, επιχειρήθηκε ο έλεγχος της αύξησης του μεγέθους του κρυστάλλου στην συνθήκη παρουσία επιπρόσθετων χημικών ουσιών από εμπορικά διαθέσιμες συλλογές πρόσθετων προσμίξεων (στην παρούσα εργασία με χρήση του Additive Screen HT, Hampton Research), που θα μπορούσαν να καταστήσουν πιο ομοιόμορφη τη διευθέτηση των πρωτεϊνικών μορίων στον κρύσταλλο. Μία άλλη τεχνική που χρησιμοποιήθηκε, στην περίπτωση της Ba0331, ήταν αυτή της σποράς μικροκρυστάλλων (microseeding), σύμφωνα με την οποία κατόπιν της ανάπτυξης των πρώτων κρυστάλλων, ακολούθησε μηχανική θραύση τους για την παραγωγή αποθέματος μικροκρυστάλλων (Bergfors 2009). Η μηχανική θραύση των κρυστάλλων έγινε σε δοκιμαστικό πλαστικό σωλήνα μιας χρήσης 1,5 mL παρουσία ενός γυάλινου σφαιριδίου διαμέτρου 1mm με 5 κύκλους των 10 s έντονης ανάδευσης σε συσκευή περιστροφικής ανάδευσης (vortex) και μεσοδιαστήματα 20 s ηρεμίας με επώαση σε πάγο. Οι μικροκρύσταλλοι προστέθηκαν ως επιπλέον συστατικό της σταγόνας κρυστάλλωσης στην αρχική συνθήκη σε ποσοστό 5% v/v, χρησιμοποιώντας τις δυνατότητες που παρέχει το λογισμικό του ρομποτικού συστήματος Oryx4 και σύμφωνα με τις οδηγίες του κατασκευαστή. Οι μικρο- ή νάνο- κρύσταλλοι αυτοί αποτελούν θέσεις πυρήνωσης από τις

οποίες θα μπορούσαν να αναπτυχθούν μεγαλύτεροι κρύσταλλοι, παρακάμπτοντας το απαραίτητο πρωταρχικό στάδιο της δημιουργίας πυρήνων (Bergfors 2009). Επιπλέον, έγινε δοκιμή για την επιβράδυνση της ανάπτυξης των κρυστάλλων με προσθήκη γλυκερόλης 10% στο διάλυμα κρυστάλλωσης ή/και 10 μl μίγματος 1:1 παραφίνης-σιλικόνης από την εταιρεία Fluka (Buchs, Ελβετία) στο δοχείο εξισορρόπησης. Τέλος κατά τη βελτιστοποίηση των αρχικών συνθηκών προκειμένου να αναπτυχθούν καλύτεροι και μεγαλύτεροι κρύσταλλοι, πραγματοποιήθηκαν επίσης περιορισμένες δοκιμές κρυστάλλωσης χειρονακτικά σε δίσκους δοκιμασιών κρυστάλλωσης 24 φρεατιδίων (VDX plates, Hampton Research Inc.) με σταγόνες όγκου (2μL), με την τεχνική της κρεμάμενης ή/και της επικαθήμενης σταγόνας. Χρησιμοποιήθηκαν καλυπτρίδες (cover slips, Greiner Bio-One) και ως στεγανωτικό εφαρμόστηκε σιλικόνη (Hampton Research Inc.). Για την αξιολόγηση της ποιότητας των κρυστάλλων εξετάστηκε η ικανότητα περίθλασης, πολλών δειγμάτων δοκιμάζοντας ταυτόχρονα την επίδραση κρυοπροστατευτικών ενώσεων.

2.2.3.3 Κρυστάλλωση συμπλόκων απακετυλάσης

Η παραγωγή κρυστάλλων συμπλόκων των απακετυλάσεων Bc1974 και Ba0330 με τις ενώσεις που δίδονται στον πίνακα 1 (Παράρτημα Α') επιχειρήθηκε τόσο με τις τεχνικές της εμβάπτισης σε πυκνά διαλύματα ενώσεων, όσο και με προσπάθειες συγκρυστάλλωσης. Τα διαλύματα των ενώσεων παρασκευάστηκαν μετά από προσπάθειες διαλυτοποίησης τους σε ύδωρ, αιθανόλη, διαιθυλαιθέρα, ή DMSO.

Εμβάπτιση

Η τεχνική της εμβάπτισης (soaking) αναφέρεται στη μεταφορά του κρυστάλλου σε πυκνό διάλυμα της ένωσης ή στην προσθήκη πυκνού διαλύματος της ένωσης στον κρύσταλλο. Η αρχική αξιολόγηση της πορείας βασίστηκε σε χρωματικές αλλαγές του κρυστάλλου ή/και μακροσκοπικές αλλαγές στην επιφάνειά του όπως εμφάνιση ραγισμάτων, κυρτώσεων, κατακερματισμού ή ολοκληρωτικής διάλυσής του.

Συγκρυστάλλωση

Η τεχνική της συγκρυστάλλωσης (co-crystallisation) αναφέρεται στην εξαρχής προσπάθεια σχηματισμού κρυστάλλων με την παρουσία της ένωσης.

2.2.3.4 Συλλογή κρυστάλλων – Κρυοπροστασία – Αποθήκευση

Έγινε επιλογή και απομόνωση μονοκρυστάλλων οι οποίοι συλλέχθηκαν με θηλιά στερέωσης μικροκρυστάλλων. Η επιλογή του κρυοπροστατευτικού έγινε μετά από δοκιμή των ενώσεων μεθυλοπεντανοδιόλη (MPD), PEG 400, αιθυλενογλυκόλης (ethylene glycol) και γλυκερόλης σε συγκεντρώσεις 20-40% v/v σε μητρικό διάλυμα. Ακολούθησε εμβάπτιση του κρυστάλλου, για 5-15 δευτερόλεπτα, στο κρυοπροστατευτικό, και άμεση υαλοποίηση (vitrification) του περιεχόμενου της θηλιάς μέσω εμβάπτισης σε υγρό άζωτο. Εν συνεχεία, οι κρύσταλλοι αποθηκεύτηκαν εντός του υγρού αζώτου μέχρι τη συλλογή κρυσταλλογραφικών δεδομένων.

2.3 Πρωτεϊνική Κρυσταλλογραφία

Στην απόδοση ενός βιολογικού ρόλου στην πρωτεΐνη, παρότι δεν είναι πάντα εφικτή, μπορεί να συμβάλει ο προσδιορισμός της τρισδιάστατης δομής της (Chandonia & Brenner 2006). Μία από τις πλέον διαδεδομένες τεχνικές για τον προσδιορισμό της δομής των πρωτεϊνών σε υψηλή διακριτικότητα είναι η κρυσταλλογραφία ακτίνων Χ. Το 87% των καταχωρημένων δομών στη βάση δεδομένων (Protein Data Bank, PDB Berman *et al.* 2000), έχει προσδιοριστεί με τη μέθοδο αυτή. Σήμερα (1/10/2018) αριθμεί 145473 καταχωρημένες δομές βιολογικών μακρομορίων. Από αυτές ο αριθμός δομών που έχουν προσδιοριστεί συνολικά με κρυσταλλογραφία ακτίνων-Χ είναι 130191, με φασματοσκοπία πυρηνικού μαγνητικού συντονισμού (NMR) 12395 και ηλεκτρονική μικροσκοπία (EM) 2493, ενώ οι λιγιστές υπόλοιπες είναι αποτέλεσμα σκέδασης ακτίνων-Χ σε μικρές γωνίες (SAXS) ή συνδυασμού μεθόδων.

2.3.1 Συλλογή κρυσταλλογραφικών δεδομένων. Πειράματα περίθλασης ακτίνων-Χ

Τα κρυσταλλογραφικά δεδομένα συνελέγησαν με ακτινοβολία συγχρότρου σε αλληπάλληλες επισκέψεις στο σύγχροτρον Swiss Light Source (SLS) του Ινστιτούτου Paul Scherrer, στην Ελβετία (PSI, Villigen, Ελβετία) στην τερματική γραμμή X06DA. Χρησιμοποιήθηκε ακτινοβολία μήκους κύματος 1,00 Å. Η συλλογή κρυσταλλογραφικών δεδομένων έγινε υπό κρυογονικές συνθήκες αφότου κάθε μονοκρύσταλλος μεταφέρθηκε ταχύτατα σε ρεύμα αζώτου θερμοκρασίας 100 K. Η συλλογή έγινε σε ανιχνευτή δύο

διαστάσεων (pixel detector PILATUS 2M, DECTRIS) με τη μέθοδο περιστροφής του κρυστάλλου υπό σταθερό άξονα.

2.3.2 Επεξεργασία δεδομένων από περίθλαση ακτίνων-X

Από τη συλλογή των εικόνων περίθλασης έως τη δημιουργία μιας λίστας εντάσεων για όλες τις παρατηρηθείσες ανακλάσεις η επεξεργασία αρχικά πραγματοποιείται με δεικτοδότηση (indexing) των περιθλασιγραμμάτων, την απόδοση δεικτών Miller hkl στις κηλίδες και την εύρεση κατάλληλης μοναδιαίας κυψελίδας στο πλέγμα, και παράλληλη βελτιστοποίηση όλων των παραμέτρων του πειράματος (όπως διαστάσεις κυψελίδας, προσανατολισμός κρυστάλλου, απόσταση κρυστάλλου-ανιχνευτή). Σε κάθε παρατηρούμενη ανάκλαση αποδίδονται οι κατάλληλοι ακέραιοι δείκτες hkl, με βάση την επιλεγμένη μοναδιαία κυψελίδα, και από το προφίλ της έντασης της κηλίδας της, γίνεται η κλασματοποίηση της, χρησιμοποιώντας τα πακέτα XDS (Kabsch 2010), MOSFLM (Leslie & Powell 2007) και HKL2000 (Otwinowski & Minor 1997). Ακολούθησε ο προσδιορισμός της ομάδας συμμετρίας χώρου που έγινε με βάση αυτές τις παρατηρήσεις, με το πρόγραμμα POINTLESS (Evans 2006) της πλατφόρμας του CCP4 (Collaborative Computational Project Number 4 1994) (Winn *et al.* 2011). Μετά τον υπολογισμό των εντάσεων των συλλεχθέντων ανακλάσεων, μπορεί να καθοριστεί η συμμετρία της εικόνας περίθλασης (ομάδα Laue συν κάποιο πιθανό κέντρο συμμετρίας πλέγματος). Παράλληλα, είναι πιθανό να εντοπιστούν τυχόν άξονες συμμετρίας από τις συστηματικές απουσίες ανακλάσεων, που ιδανικά μπορεί να οδηγήσουν σε μια αδιαμφισβήτητη ένδειξη της ομάδας χώρου. Οι εντάσεις που υπολογίστηκαν με τα επιλεγμένα προγράμματα δεν είναι όλες στην ίδια κλίμακα, εξαιτίας ενός συνδυασμού φυσικών παραγόντων του πειράματος, κι επομένως ένας αριθμός διορθώσεων είναι αναγκαίο να εφαρμοστεί προκειμένου να επιτευχθεί η καλύτερη δυνατή εκτίμηση για τις πραγματικές έντασεις. Στην περίπτωση πολλαπλών μετρήσεων (εξαιτίας της στρατηγικής συλλογής δεδομένων και της Laue συμμετρίας), με την εύρεση των παραγόντων κλίμακας (scale factors), υπολογίζεται μια μέση ένταση για τις συμμετρικά ισοδύναμες ανακλάσεις που αποδίδεται σε κάθε μοναδική hkl ανάκλαση. Το στάδιο της συγχώνευσης ανακλάσεων (merging), παρέχει στατιστικούς δείκτες για την ποιότητα των δεδομένων, κυρίως εξετάζοντας τη συμφωνία μεταξύ των συμμετρικά-συσχετιζόμενων παρατηρήσεων. Η κανονικοποίηση και η συγχώνευση των δεδομένων έγιναν με το λογισμικό SCALA (CCP4) (Evans 2006)

2.3.3 Επίλυση δομής-Βελτιστοποίηση

Για την επίλυση της δομής χρησιμοποιούνται οι άμεσες μέθοδοι, η μέθοδος μοριακής αντικατάστασης (Molecular Replacement, MR) αλλά και πειραματικές μέθοδοι (Single Isomorphous Replacement SIR, Multiple Isomorphous Replacement MIR, κ.ά.) Η επίλυση της κρυσταλλικής δομής των Bc1974, Ba0330, Ba0331 και του καθενός συμπλόκου της Bc1974 έγινε με τη μέθοδο της μοριακής αντικατάστασης (Rossmann 1972) χρησιμοποιώντας το λογισμικό PHASER (McCoy *et al.* 2007). Η βελτιστοποίηση των δομών έγινε με το αυτοματοποιημένο λογισμικό PHENIX (Adams *et al.* 2010, Afonine *et al.* 2012), με εφαρμογή στερεοχημικών περιορισμών βάσει του μοντέλου των Engh & Huber 1991. Στα τελικά στάδια της βελτιστοποίησης χρησιμοποιήθηκε το μοντέλο TLS (Translation/Libration/Screw) για την εφαρμογή βελτιστοποίησης ανισοτροπικού τύπου (Painter & Merritt 2006) χωρίς την εισαγωγή καινούργιων παραμέτρων. Για την ανοικοδόμηση του τρισδιάστατου μοντέλου (model building) της δομής της πρωτεΐνης με βάση το χάρτη ηλεκτρονιακής πυκνότητας (electron density map) και για τα στάδια βελτιστοποίησης του πρωτεϊνικού μοντέλου στον πραγματικό χώρο χρησιμοποιήθηκε το λογισμικό COOT (Emsley & Cowtan 2004, Emsley *et al.* 2010). Το όριο διακριτικής ικανότητας των δεδομένων περίθλασης προσδιορίστηκε βάσει των κριτηρίων CC1/2 και $I/\sigma(I)$ με ουδό τιμών 50-70% και 1-1,5, αντίστοιχα (Evans & Murshudov 2013, Karplus & Diederichs 2012). Με χρήση του αυτοματοποιημένου πρωτόκολλου του COOT, προστέθηκαν τα μόρια ύδατος που σχημάτιζαν δίκτυα δεσμών υδρογόνου με την πρωτεΐνη και τα άλλα μόρια διαλύτη. Η αξιοπιστία του τελικού μοντέλου ελέγχθηκε με το πρόγραμμα MOLPROBITY (Chen *et al.* 2010) και το PROCHECK (Laskowski *et al.* 1993) καθώς και με ανάλυση του διαγράμματος Ramachandran (Ramachandran *et al.* 1963). Τέλος, το λογισμικό μοριακών γραφικών PyMOL (The PyMOL Molecular Graphics System, vers. 1.8 Schrödinger, LLC) χρησιμοποιήθηκε για την απεικόνιση των μοντέλων των δομών και για τη δημιουργία των σχημάτων όπου απεικονίζονται τρισδιάστατα μοντέλα πρωτεϊνικών δομών. Το τελικό στάδιο περιελάμβανε την κατάθεση των ατομικών συντεταγμένων και των παραγόντων δομής στην ενδεδειγμένη βάση δεδομένων Protein Data Bank (Berman *et al.* 2000).

2.4 Βιοπληροφορική

2.4.1 Ανάκτηση ακολουθιών

Οι πρωτεϊνικές και νουκλεοτιδικές αλληλουχίες της οικογένειας γονιδίων απακετυλάσης πολυσακχαρίτη του *B. anthracis* str. Ames και *B. cereus* str. ATCC14579 ανακτήθηκαν από τη βάση δεδομένων GENE BANK του NCBI (National Center for Biotechnology Information, Εθνικό Κέντρο Βιοτεχνολογικής Πληροφόρησης, Bethesda, MD, ΗΠΑ). Οι αναζητήσεις για τη σύγκριση της κάθε αλληλουχίας έγιναν μέσω του αλγόριθμου BLAST (Basic Local Alignment Search Tool, Βασικό Εργαλείο Ανίχνευσης Τοπικής Στοίχισης, Altschul *et al.* 1990, που υποστηρίζεται στο NCBI, και διατίθεται στην διεύθυνση <http://www.ncbi.nlm.nih.gov/BLAST/>) και πραγματοποιήθηκαν χρησιμοποιώντας το πρόγραμμα MegaBLAST με εφαρμογή στη βάση δεδομένων RefSeq (και επίσης στις PDB και SwissProt) χρησιμοποιώντας την παραλλαγή Blastp (πρωτεΐνη-πρωτεΐνη BLAST) με τις προεπιλεγμένες παραμέτρους, εκτός από τις μέγιστες ακολουθίες στόχων. Στη μελέτη των λιποπρωτεϊνών χρησιμοποιήθηκε η βάση δεδομένων για τις βακτηριακές λιποπρωτεΐνες DOLOP (Babu *et al.* 2006, www.mrc-lmb.cam.ac.uk/genomes/dolop). Επίσης πληροφορίες για την οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη εντός του κάθε στελέχους αντλήθηκαν και από τη βάση δεδομένων LocateP (Zhou *et al.* 2008, <http://www.cmbi.ru.nl/locatep-db/cgi-bin/locatepdb.py>) ενώ η ίδια βάση δεδομένων χρησιμοποιήθηκε για την πρόβλεψη του κυτταρικού εντοπισμού των πρωτεϊνών καθώς και για την αποκοπή ή μη της αλληλουχίας σήματος.

2.4.2. Επιπρόσθετα προγράμματα ανάλυσης ακολουθιών

Ο διαδικτυακός διακομιστής SignalP 4.1 (Petersen *et al.* 2011) χρησιμοποιήθηκε για την πρόβλεψη της παρουσίας και της θέσης αποκοπής του πεπτιδίου σήματος σε αλληλουχίες αμινοξέων. Για την απεικόνιση των νουκλεοτιδικών αλληλουχιών των πλασμιδίων χρησιμοποιήθηκε το πρόγραμμα SnapGene viewer 4.1.3 (SnapGene software της GSL Biotech, που διατίθεται στην διεύθυνση <http://www.snapgene.com>). Το πρόγραμμα GeneDoc (Nicholas *et al.* 1997) παρέχει ένα σύνολο εργαλείων για την απεικόνιση, επεξεργασία και ανάλυση πολλαπλών ευθυγραμμίσεων αλληλουχιών τόσο πρωτεϊνών όσο και νουκλεϊνικών οξέων σε ένα εξελικτικό πλαίσιο (χρησιμοποιήθηκε η έκδοση GeneDoc MFC Application vers. 2.6.0.2).

2.4.2.1 Στοιχίση ακολουθιών

Η προοδευτική στοιχίση για πολλαπλές ακολουθίες του Clustal Omega (Clustal-O, Sievers *et al.* 2011 διαθέσιμο στην διεύθυνση <https://www.ebi.ac.uk/Tools/msa/clustalo>) χρησιμοποιήθηκε για να εκτελέσει τις ευθυγραμμίσεις των νουκλεοτιδικών και αμινοξικών αλληλουχιών καθώς και η έκδοση με συντελεστές βαρύτητας ClustalW για ζευγάρια ακολουθιών. Εναλλακτικώς χρησιμοποιήθηκε ο εξυπηρετητής πολλαπλής στοιχίσης ακολουθιών T-Coffee (Notredame *et al.* 2000). Το πρόγραμμα ευθυγράμμισης Lalign (Huang & Miller 1991) χρησιμοποιήθηκε για τη σύγκριση ζευγαριών αλληλουχιών. Το ESPript χρησιμοποιήθηκε στην απεικόνιση της ευθυγράμμισης των αλληλουχιών και στην ενσωμάτωση πληροφοριών συντήρησης ή ποικιλομορφίας (Gouet *et al.* 2003).

2.4.3 Φυλογενετική Ανάλυση - Μοριακή Εξέλιξη

Τα δέντρα κατασκευάστηκαν χρησιμοποιώντας τη μέθοδο της μέγιστης πιθανοφάνειας όπως εφαρμόζεται στο MEGA7 (Kumar *et al.* 2016) με δοκιμή αξιοπιστίας (bootstrap, Felsenstein 1985). Για τις πρωτεϊνικές αλληλουχίες χρησιμοποιήθηκε η μέθοδος της μέγιστης πιθανοφάνειας με τη χρήση της μήτρας γενετικών αποστάσεων Jones-Taylor-Thornton (JTT, Jones *et al.* 1992), σε γ -κατανομή. Για το εξελικτικό δένδρο των νουκλεοτιδικών αλληλουχιών χρησιμοποιήθηκε η μέθοδος μέγιστης πιθανοφάνειας, με τη χρήση της GTR (General Time Reversible, Lanave *et al.* 1984) σε γ -κατανομή και διαχείριση των αμετάβλητων θέσεων (G+I μοντέλο) (Nei & Kumar 2000). Το μοντέλο GTR + G + I παράγαγε τις χαμηλότερες βαθμολογίες BIC (Bayesian Information Criterion, Schwarz 1978) και θεωρήθηκε ότι περιγράφει το πρότυπο αντικατάστασης για τις ακολουθίες γονιδίων που κωδικεύουν πρωτεΐνες. Σε όλες τις φυλογένειες εμφανίζεται το δέντρο με τη μέγιστη λογαριθμική πιθανοφάνεια. Τα αρχικά δέντρα για ευριστικές αναζητήσεις αποκτήθηκαν αυτόματα με την εφαρμογή των αλγορίθμων Neighbor-Join και BioNJ σε μια μήτρα ζευγών αποστάσεων, που εκτιμήθηκε χρησιμοποιώντας την προσέγγιση της Μέγιστης Συνθετικής Πιθανοφάνειας (MCL) και έπειτα επιλέγοντας την τοπολογία με τη μέγιστη λογαριθμική πιθανοφάνεια.

Το λογισμικό πακέτο φυλογενετικής ανάλυσης μέγιστης πιθανοφάνειας PAML (Phylogenetic Analysis by Maximum Likelihood, Yang 2007) χρησιμοποιήθηκε για τον υπολογισμό του ρυθμού συνώνυμων (dS) και μη συνώνυμων (dN) αντικαταστάσεων. Η δοκιμασία McDonald-Kreitman χρησιμοποιήθηκε για την ανίχνευση της επιλογής σε επίπεδο

DNA (McDonald & Kreitman 1991) μέσω του υπολογισμού του ρυθμού συνώνυμων και μη συνώνυμων πολυμορφισμών. Η σύγκριση πολυμορφισμού και απόκλισης είναι μία από τις πιο ισχυρές προσεγγίσεις για τη διερεύνηση της επιλογής σε επίπεδο DNA. Η δοκιμασία McDonald-Kreitman είναι σχεδόν απαραίτητο βήμα για να ακολουθήσει αυτή η προσέγγιση. Συγκρίνει το μέγεθος της μεταβολής εντός ενός είδους με την απόκλιση μεταξύ των ειδών σε δύο τύπους τοποθεσιών, εκ των οποίων ο ένας θεωρείται ουδέτερος και χρησιμοποιείται ως αναφορά για την ανίχνευση της επιλογής στον άλλο τύπο τοποθεσίας. Καθώς περιγράφηκε αρχικά (McDonald και Kreitman 1991), αυτές οι θέσεις ήταν συνώνυμες (υποθετικά ουδέτερες) και μη-συνώνυμες σε μία κωδική περιοχή. Ωστόσο, η δοκιμασία για την επιλογή μπορεί ενδεχομένως να επεκταθεί σε οποιουδήποτε δύο τύπους τοποθεσιών, υπό την προϋπόθεση ότι ένας από αυτούς υποτίθεται ότι εξελίσσεται ουδέτερα και ότι και οι δύο συνδέονται στο γονιδίωμα. Περαιτέρω, η λογική της δοκιμασίας μπορεί να χρησιμοποιηθεί για την ανάλυση πολλαπλών τόπων στο γονιδίωμα ταυτόχρονα, υπό τον όρο ότι εφαρμόζονται οι κατάλληλες στατιστικές δοκιμασίες πολλαπλών θέσεων.

2.5 Δομική Βιοπληροφορική

Όλες οι πειραματικώς προσδιορισμένες δομές με κρυσταλλογραφία ακτίνων X που συμπεριελήφθησαν στην παρούσα μελέτη ελήφθησαν από την πρωτεϊνική τράπεζα δεδομένων (PDB, Protein Data Bank). Οι δομές των *B. anthracis* και *B. cereus* που μελετήθηκαν στην παρούσα εργασία είναι οι εξής: Ba0424 (PDB ID: 2J13), Ba0150 (PDB ID: 4M1B), Bc0361 (PDB ID: 4HD5), Bc1960 (PDB ID: 4L1G), Ba0330 (PDB ID: 4V33), Bc1974 (PDB ID: 5N1J) και Ba0331 (PDB ID: 6G01). Οι πολλαπλές ευθυγραμμίσεις δομών των αντίστοιχων δομικών ενοτήτων (NodB, Fn3) πραγματοποιήθηκαν με τη χρήση του PyMOL (όπως αναφέρθηκε. §2.3.3) με βάση τη βελτιστοποίηση r.m.s.d. και την ευθυγράμμιση αλληλουχίας. Η αναζήτηση δομικών ομολόγων πραγματοποιήθηκε με τον εξυπηρετητή Dali (Holm *et al.* 2008), με βάση τις βαθμολογίες Dali Z και γραφήματα δομικής ομοιότητας (Holm & Laakso 2016). Το ENDscript (Robert *et al.* 2014) έχει χρησιμοποιηθεί για την απεικόνιση της συντήρησης αλληλουχίας στην δομή μεταξύ μελών της οικογένειας PDA, προκειμένου να εντοπιστούν μεταβλητές θέσεις στον λειτουργικό τομέα. Επίσης το ENDscript χρησιμοποιήθηκε για την απεικόνιση των πληροφοριών από την τρισδιάστατη δομή. Το πρόγραμμα LigPlot (Wallace *et al.* 1995) χρησιμοποιήθηκε για τη δημιουργία αυτόματων σχηματικών διαγραμμάτων αλληλεπιδράσεων πρωτεΐνης-προσδέτη από ένα δεδομένο αρχείο PDB. Οι αλληλεπιδράσεις που παρουσιάζονται είναι δεσμοί υδρογόνου και

υδρόφοβες αλληλεπιδράσεις. Εκτός από προγράμματα που ήδη αναφέρονται στην §2.3.3 ακόμη χρησιμοποιήθηκαν, ο διαδικτυακός διακομιστής 3V (Voss & Gerstein 2010) για την ανάλυση των εσωτερικών όγκων της κοιλότητας δέσμησης όλων των δομών. Στην μελέτη των επιφανειών που έρχονται σε επαφή με το διαλύτη αλλά και των διεπιφανειών των αλληλεπιδράσεων των δομικών ενοτήτων και της κρυσταλλικής μοριακής διεύθεσης χρησιμοποιήθηκε το πρόγραμμα PISA (Protein interfaces surfaces and assemblies service, Krissinel & Henrick 2007) της CCP4, διαθέσιμο και στο European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html). Όπου δεν αναφέρεται αλλιώς, οι παράμετροι των προγραμμάτων που χρησιμοποιήθηκαν ήταν οι προεπιλεγμένοι.

2.5.1 Πρόβλεψη και Μελέτη της αξιοπιστίας των τρισδιάστατων μοντέλων

Ο αλγόριθμος QUARK χρησιμοποιήθηκε για την πρόβλεψη τρισδιάστατης δομής ακολουθιών λίγων αμινοξέων με τη βοήθεια του ηλεκτρονικού διακομιστή σε απευθείας σύνδεση (Xu & Zhang 2011, 2012). Η μοντελοποίηση προτύπου ομολόγων χρησιμοποιήθηκε κυρίως για τον εντοπισμό ανεπαίσθητων αλλαγών στην κοιλότητα δέσμησης, κυρίως λόγω αντικαταστάσεων αμινοξέων στα μοτίβα που μελετήθηκαν (MT1-MT5). Η περιστροφική ανάλυση κίνησης των Ba0330 Bc0361 και Ba0331 ποσοτικοποιήθηκε με το DynDom (Hayward & Berendsen 1998).

2.5.2 Προσομοίωση της πρόσδεσης υποστρωμάτων

Κατά τον *in silico* σχεδιασμό υποστρωμάτων για την Bc1974 οι προσδέτες αρχικά αναπαραστάθηκαν από τα SMILES με τη χρήση του προγράμματος OMEGA (Open Eye Scientific Software, SantaFe, NM.) (Hawkins *et al.* 2010, Hawkins & Nicholls 2012). Οι υπολογισμοί ελλιμενισμού στην Bc1974 πραγματοποιήθηκαν με χρήση των AutoDock4.0 (Morris *et al.* 2009). Για τη σχεδίαση του φαρμακοφόρου εκτός από το πρόγραμμα PyMOL που έχει ήδη αναφερθεί (§2.3.3) χρησιμοποιήθηκε και το πρόγραμμα Maestro (Schrödinger LLC). Επιπροσθέτως το πρόγραμμα απεικόνισης VMD (visual molecular dynamics, Humphrey *et al.* 1996) χρησιμοποιήθηκε για την προβολή και την ανάλυση των προσδετών και των υποστρωμάτων. Τα αποτελέσματα των υπολογισμών ελλιμενισμού αξιολογήθηκαν μέσω οπτικής επιθεώρησης των προβλεπόμενων αλληλεπιδράσεων και της εκτιμώμενης ελεύθερης ενέργειας δέσμησης. (Huey *et al.* 2007)

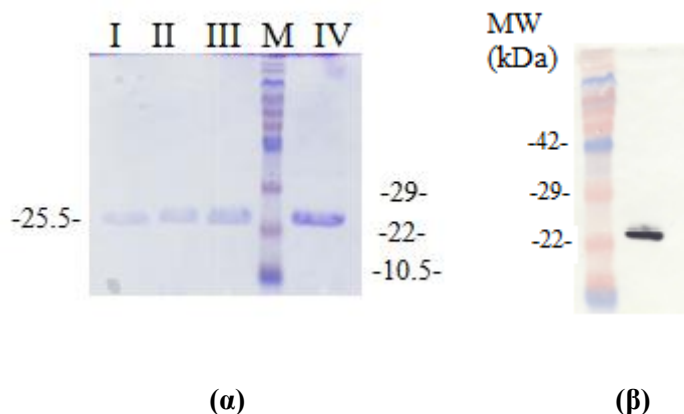
3. Αποτελέσματα

3.1. Έκφραση, απομόνωση και κρυστάλλωση της Ba3679

3.1.1. Έκφραση της Ba3679 από μετασχηματισμένα κύτταρα *E. coli*

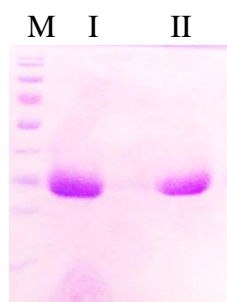
Στην κρυσταλλογραφία μακρομορίων, η επιλογή κατάλληλου πειραματικού σχήματος για την ετερόλογη έκφραση των πρωτεϊνών – στόχων είναι καθοριστικής σημασίας. Συνήθως, δοκιμάζονται περισσότερα από ένα προκειμένου να επιλεγθεί το καταλληλότερο. Λόγω του ότι πρόκειται για βακτηριακές πρωτεΐνες οι προσπάθειες έγιναν σε βακτηριακά στελέχη και επιλέχθηκαν στελέχη του *E. coli*, που παρουσιάζουν ένα σύστημα εύκολο στην χρήση και τους χειρισμούς, χαμηλού κόστους και υψηλής ταχύτητας.

Η Ba3679 όπως αναφέρθηκε προηγουμένως (βλ. §2.1.8) ολόκληρη ήταν υποκλωνοποιημένη στον φορέα έκφρασης pET26b(+), έτσι ώστε να βρίσκεται υπό τον έλεγχο του υποκινητή T7, και να εκφράζεται με μία καρβοξυτελική ετικέτα 6 καταλοίπων ιστιδίνης (6xHisTag) ενώ είχε αποκοπεί το πεπτίδιο οδηγός του φορέα έκφρασης. Ο ανασυνδυασμένος φορέας έκφρασης χρησιμοποιήθηκε για το μετασχηματισμό επιδεικτικών κυττάρων *E. coli* αρχικώς του στελεχούς BL21(DE3), και κατόπιν των στελεχών C41(DE3), C43(DE3) και Rosetta 2(DE3) (βλ. §2.1.9 και §2.2.2.2). Από τους μετασχηματισμένους κλώνους που αναπτύχθηκαν σε στερεό θρεπτικό μέσο με δείκτη επιλογής την καναμυκίνη έγινε τυχαία επιλογή ενός κλώνου, ο οποίος χρησιμοποιήθηκε για καλλιέργειες έκφρασης σε μεσαία κλίμακα. Η έκφραση επάγεται με προσθήκη IPTG στο υγρό θρεπτικό μέσο ανάπτυξης, που προκαλεί την ενεργοποίηση του υποκινητή T7. Η έλλειψη του πεπτιδίου οδηγού (pelb) του φορέα έκφρασης έχει αποτέλεσμα η πρωτεΐνη να παραμένει στο κυτταρόπλασμα. Το αναμενόμενο μοριακό βάρος είναι 25,5kDa. Έγιναν προσπάθειες επαγωγής με διάφορους συνδυασμούς συγκέντρωσης IPTG, θερμοκρασίας και διάρκειας έκφρασης. Συγκεκριμένα, χρησιμοποιήθηκε IPTG συγκέντρωσης: 0,2mM, 0,5mM, 1mM, και θερμοκρασίες με διάρκεια έκφρασης: 15°C για 7h, 20°C για 7h, 25°C για 6h, 30°C για 5h. Η έκφραση της ετερόλογης πρωτεΐνης ήταν υψηλότερη στα C41 (DE3) και ακολούθως C43 (DE3), ενώ σε όλες τις περιπτώσεις η απόδοση ήταν βέλτιστη με συγκέντρωση IPTG 0,2mM, σε θερμοκρασία 20°C για χρονικό διάστημα 7h (Εικόνα 3.1)



Εικόνα 3.1 (α) SDS-PAGE και χρώση με Coomassie brilliant blue κατόπιν έκλουσης από στήλη συγγενείας Ni-NTA. **I:** BL21(DE3), **II:** Rosetta 2(DE3), **III:** C43(DE3), **IV:** C41(DE3) (25,5 kDa) **M:** πρωτεϊνικός μάρτυρας. Παρουσιάζονται τα κλάσματα έκλουσης με ιμιδαζόλιο 200 mM, αφού πρώτα έγιναν εκπλύσεις με διάλυμα ιμιδαζολίου 20 και 30 mM. Σημειώνεται ότι σε κάθε φρέαρ του πηκτώματος φορτώθηκε ίδιος όγκος από το έκλουσμα ιμιδαζολίου 200mM, σταθερού όγκου συλλογής κλασμάτων. **(β)** SDS-PAGE και ανοσοαποτύπωση κατά Western με αντί-His αντίσωμα του IV.

Συνολικά αν και ανιχνεύεται στο θεωρητικά αναμενόμενο μοριακό βάρος, τα αποτελέσματα ήταν απογοητευτικά για πειράματα κρυστάλλωσης, ως προς την επάρκεια απαιτούμενης ποσότητας αλλά και ως προς τη σταθερότητα της πρωτεΐνης. Η πρωτεΐνη κατακρημνίζονταν σχετικά γρήγορα ακόμη και σε χαμηλές συγκεντρώσεις, τόσο στους 4°C όσο και σε RT, δείχνοντας ότι είναι ασταθής. Σημειώτεον μεγάλη ποσότητα της προσδεδεμένης πρωτεΐνης σε σφαιρίδια Ni-NTA κατακρημνίστηκε εντός της στήλης καθώς δεν εκλούστηκε ούτε μετά από επώαση των σφαιριδίων με διάλυμα ιμιδαζολίου 2M (**Εικόνα 3.2**) Επίσης η περισσότερη πρωτεΐνη βρισκόταν σε σωματίδια εγκλεισμού (inclusion bodies). Η απόφαση που λήφθηκε ήταν να γίνει μοριακή κατασκευή έτσι ώστε η εκπεφρασμένη πρωτεΐνη να μην περιέχει καρβοξυτελική ετικέτα ιστιδινών και να είναι συζευγμένη στο αμινοτελικό άκρο της με την MBP (Maltose Binding Protein), προκειμένου να βελτιωθεί η ποσότητα και ποιότητα της καθαρισμένης πρωτεΐνης, ενώ να υπάρχει και πρωτεολυτική θέση ώστε να μπορεί να αποκοπεί από την πρωτεΐνη που είναι συζευγμένη.



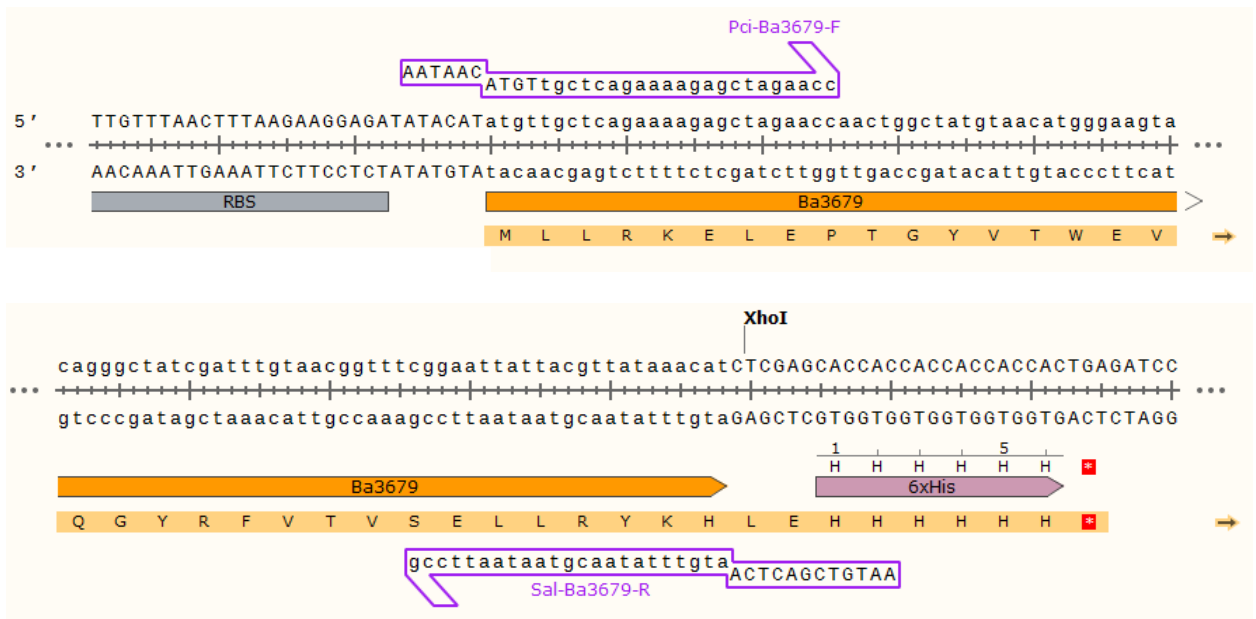
Εικόνα 3.2. SDS-PAGE και χρώση με Coomassie brilliant blue έχοντας φορτώσει ~10μL αναδιαλυμένων σφαιριδίων Ni-NTA σε H₂O. Η κατακρημνισμένη πρωτεΐνη (25,5kDa) δεν είχε εκλουστεί ούτε μετά την έκπλυση της στήλης με διάλυμα ιμιδαζολίου 2M. I: Ni-NTA σφαιρίδια C41, II: Ni-NTA σφαιρίδια C43 M: πρωτεϊνικός μάρτυρας.

3.1.2 Υποκλωνοποίηση της Ba3679 σε πλασμιδιακό φορέα έκφρασης

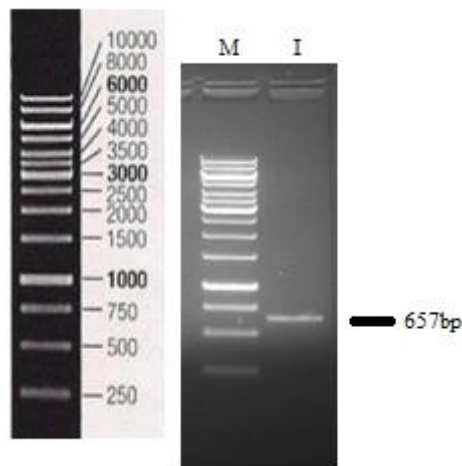
Το ανοιχτό πλαίσιο ανάγνωσης της Ba3679 για την έκφραση της αποκτήθηκε ως υποκλωνοποιημένο στον φορέα pET26b(+), μεταξύ των θέσεων περιορισμού NdeI και XhoI (pET26b(+)-Ba3679, βλ. εικόνα 2.1). Για τη μοριακή κατασκευή του πλαισίου ανάγνωσης της συζευγμένης Ba3679 με την MBP που φέρει αμινοτελική 6xHis ετικέτα και επίτοπο αναγνώρισης HRV 3C πρωτεάσης: NH₂-6xHis-MPB-HRV3C-Ba3679(1-213)-COOH χρησιμοποιήθηκε το πλασμίδιο pETM-44 (βλ. εικόνα 2.2). Το πρώτο βήμα ήταν μια αντίδραση PCR (**Εικόνα 3.3**), για την οποία χρησιμοποιήθηκε ως μήτρα DNA το πλασμίδιο pET26b(+)-Ba3679[ORF] και τα εναρκτήρια μόρια Pci-Ba3679-F και Sal-Ba3679-R (§2.1.6). Στην παρούσα διατριβή οι εκκινητές που χρησιμοποιήθηκαν, είχαν σχεδιαστεί ώστε να διαθέτουν και αλληλουχίες αναγνώρισης από κατάλληλα περιοριστικά ένζυμα προκειμένου να καθίσταται εφικτή η υποκλωνοποίηση στον πλασμιδιακό φορέα έκφρασης. Το εναρκτήριο μόριο Pci-Ba3679-F στο 3'-άκρο του είναι όμοιο με την αμινοτελική ακολουθία της Ba3679 για 26 νουκλεοτίδια, με σκοπό τον υβριδισμό του με την αλυσίδα της μήτρας. Το 5'-άκρο του δεν είναι συμπληρωματικό με τη μήτρα, αλλά φέρει την ακολουθία αναγνώρισης της PciI, η οποία εισάγεται στο προϊόν της αντίδρασης. Το εναρκτήριο μόριο Sal-Ba3679-R στο 3'-άκρο του είναι αντίστροφο και συμπληρωματικό με την καρβοξυτελική ακολουθία της Ba3679 για 23 νουκλεοτίδια, ενώ το 5'-άκρο του δεν είναι συμπληρωματικό με τη μήτρα, αλλά φέρει την ακολουθία αναγνώρισης της SalI, η οποία εισάγεται στο προϊόν της αντίδρασης. Έτσι, το

προϊόν της αντίδρασης PCR είναι ένα γραμμικό δίκλωνο τμήμα DNA που μεταξύ των θέσεων περιορισμού *PciI* και *SalI* φέρει την αλληλουχία που κωδικεύει την Ba3679 και ένα κωδικόνιο λήξης. Το προϊόν της PCR έχει αναμενόμενο μέγεθος 657 bp και ελέγχθηκε με ηλεκτροφόρηση (Εικόνα 3.4). Το δεύτερο βήμα ήταν η πέψη των προϊόντων των αντιδράσεων PCR με τα ένζυμα περιορισμού *PciI* και *SalI*, έτσι ώστε να προκύψουν μονόκλιωνα κολλώδη άκρα για τις δύο θέσεις. Αντίστοιχα, το πλασμίδιο pETM-44 υποβλήθηκε σε πέψη με *NcoI* και *SalI* και από τα προϊόντα της αντίδρασης, μετά από ηλεκτροφόρηση σε πήκτωμα αγαρόζης, απομονώθηκε και καθαρίστηκε ο ανοιχτός φορέας κλωνοποίησης. Το τμήμα DNA που απομονώθηκε από την πέψη των προϊόντων των αντιδράσεων PCR υποκλωνοποιήθηκε στον ανοιχτό φορέα έκφρασης που απομονώθηκε από την αντίδραση πέψης του pETM-44. Έτσι το τελικό προϊόν από την αντίδραση σύνδεσης είναι ο ανασυνδυασμένος φορέας έκφρασης pETM-44 που φέρει ένα τμήμα DNA που περιλαμβάνει την αλληλουχία που κωδικεύει το πλαίσιο ανάγνωσης μίας ετικέτας 6 καταλοίπων ιστιδίνης, την MBP πρωτεΐνη, τον επίτοπο αναγνώρισης της πρωτεάσης HRV 3C και την Ba3679 με ένα κωδικόνιο λήξης της μετάφρασης στο καρβοξυτελικό άκρο της pETM44-His-MBP-HRV3C-Ba3679[ORF] (Εικόνα 3.5) συνολικά ο αριθμός των αμινοξέων του προϊόντος είναι 600 με μοριακό βάρος 67106,50Da. Η μοριακή κατασκευή που έχει μέγεθος 7057bp, αρχικώς ελέγχθηκε με ηλεκτροφόρηση μετά από πέψη του πλασμιδίου με *ClaI* (Εικόνα 3.6) και ακολούθως επιβεβαιώθηκε με αλληλούχιση.

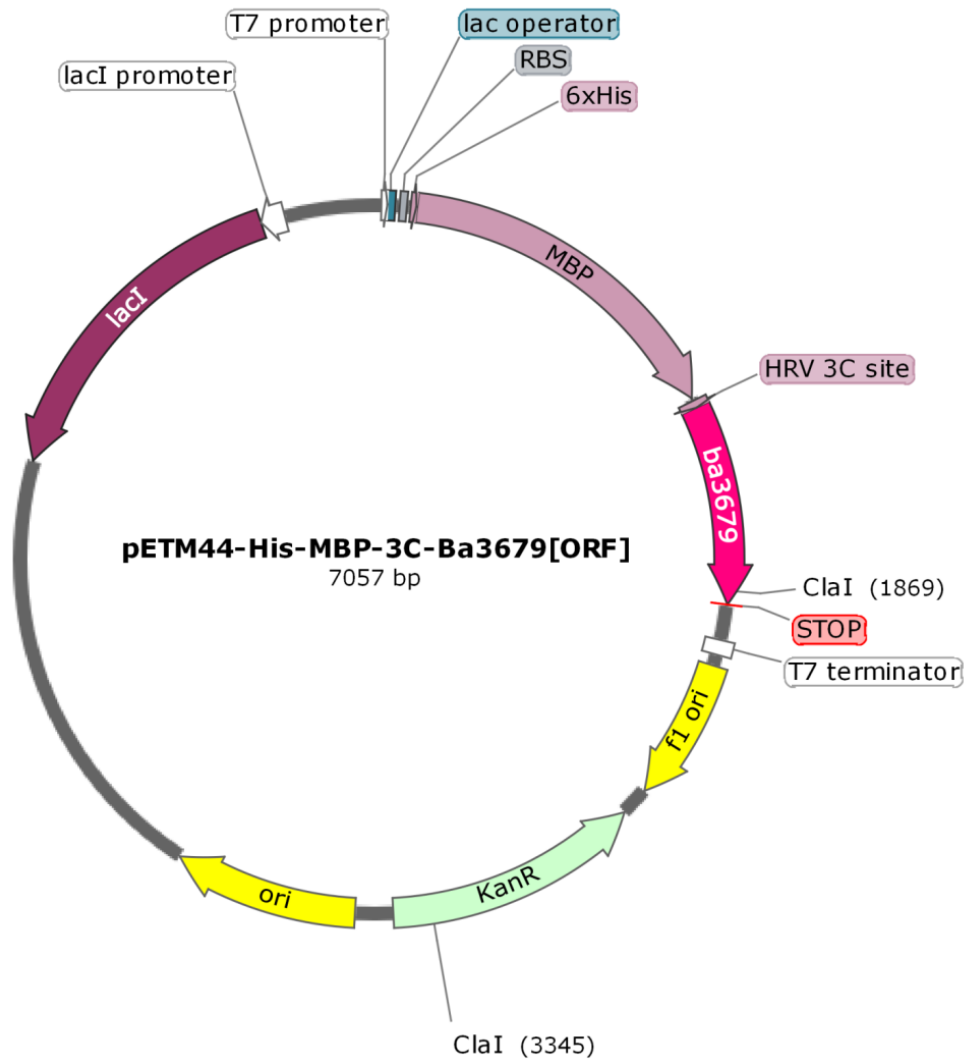
Η επίδραση της HRV 3C πρωτεάσης οδηγεί σε μόρια Ba3679 προσαυξημένα κατά δύο αμινοξικά κατάλοιπα (γλυκίνη και προλίνη) στο αμινοτελικό της άκρο ως αποτέλεσμα της πρωτεολυτικής αντίδρασης, αφού η HRV 3C πρωτεάση αναγνωρίζει την αλληλουχία Leu-Glu-Val-Leu-Phe-Gln-↓-Gly-Pro και κόβει ακριβώς μετά το κατάλοιπο γλουταμίνης.



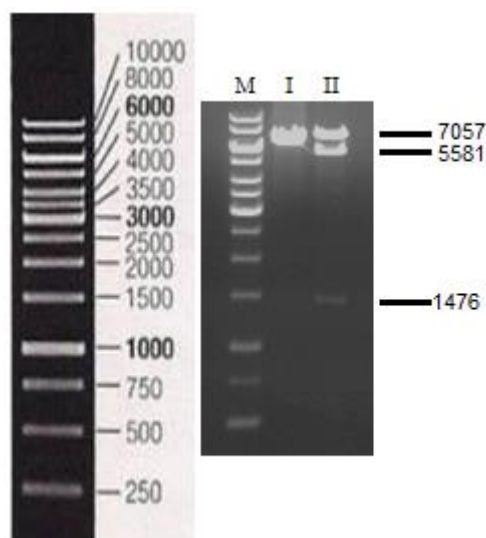
Εικόνα 3.3. Μέρος της αλληλουχίας του πλασμιδίου pET26b(+) μεταξύ των θέσεων περιορισμού *Nde*I και *Xho*I στις περιοχές υβριδισμού με τα εναρκτήρια μόρια Pci-Ba3679-F και Sal-Ba3679-R. Η αλληλουχία αυτή (βλ. Εικόνα 2.1) συμπεριλαμβάνει το ανοιχτό πλαίσιο ανάγνωσης της Ba3679 με προσαρτημένη ετικέτα ιστιδινών (6xHis). Δείχνονται μέρος του πλαισίου ανάγνωσης της Ba3679 [DNA (με πεζά) και αμινοξική ακολουθία], απεικονίζονται οι περιοχές υβριδισμού των εκκινητών και η αλληλουχία των εκκινητών (με κεφαλαία οι αλληλουχίες όπως σχεδιάστηκαν ώστε να συμπεριλαμβάνουν την αναγνώριση από τα περιοριστικά ένζυμα *Pci*I και *Sal*I).



Εικόνα 3.4. Ηλεκτροφόρηση των προϊόντων της PCR σε πήκτωμα αγαρόζης μετά την οπτικοποίηση μέσω έκθεσης σε υπεριώδη ακτινοβολία (Gel Documentation System, BIORAD) I: προϊόν PCR μεγέθους 657bp, M: δείκτης μοριακού μεγέθους (SM03111GeneRuler 1 kb DNA LadderFermentas) μπροστά δίδεται το πρότυπο των ζωνών του.



Εικόνα 3.5. Πλασμινδιακός χάρτης του φορέα έκφρασης pETM-44 με ένθεμα την Ba3679 συζευγμένη στο αμινοτελικό της άκρο με την MBP ακολουθούμενη από τον επίτοπο αναγνώρισης της HRV 3C πρωτεάσης. Η MBP φέρει ετικέτα 6 καταλοίπων ιστιδίνης (6xHis) προσαρτημένη στο αμινοτελικό της άκρο. T7 promoter & T7 terminator: περιοχές έναρξης και λήξης της μεταγραφής, αντίστοιχα, *ba3679*: το ανοιχτό πλαίσιο ανάγνωσης της Ba3679, STOP: κωδικόνιο λήξης της μετάφρασης, *ClaI*: θέσεις αναγνώρισης της περιοριστικής ενδονουκλεάσης *ClaI*.



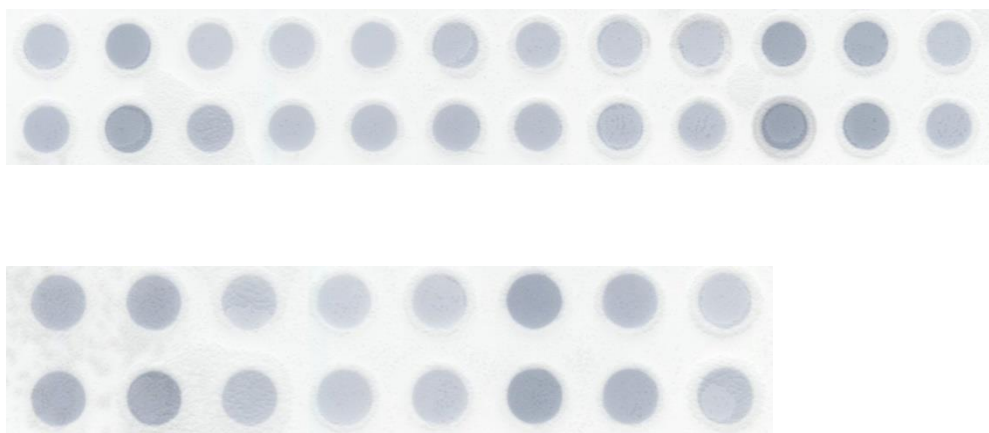
Εικόνα 3.6. Ηλεκτροφόρηση σε πήκτωμα αγαρόζης για τον έλεγχο της υποκλωνοποίησης στον φορέα έκφρασης pETM-44 μετά την οπτικοποίηση μέσω έκθεσης σε υπεριώδη ακτινοβολία (Gel Documentation System, BIORAD). I: η τελική πλασμιδιακή κατασκευή χωρίς να έχει υποστεί πέψη, II: προϊόντα της ημιτελούς πέψης του πλασμιδίου με την περιοριστική ενδονουκλεάση *ClaI* (βλ. εικόνα 3.5), M: δείκτης μοριακού μεγέθους (SM03111 Gene Ruler 1 kb DNA Ladder Fermentas) μπροστά δίδεται το πρότυπο των ζωνών του.

3.1.3 Καλλιέργειες μεσαίας κλίμακας και απομόνωση της πρωτεΐνης

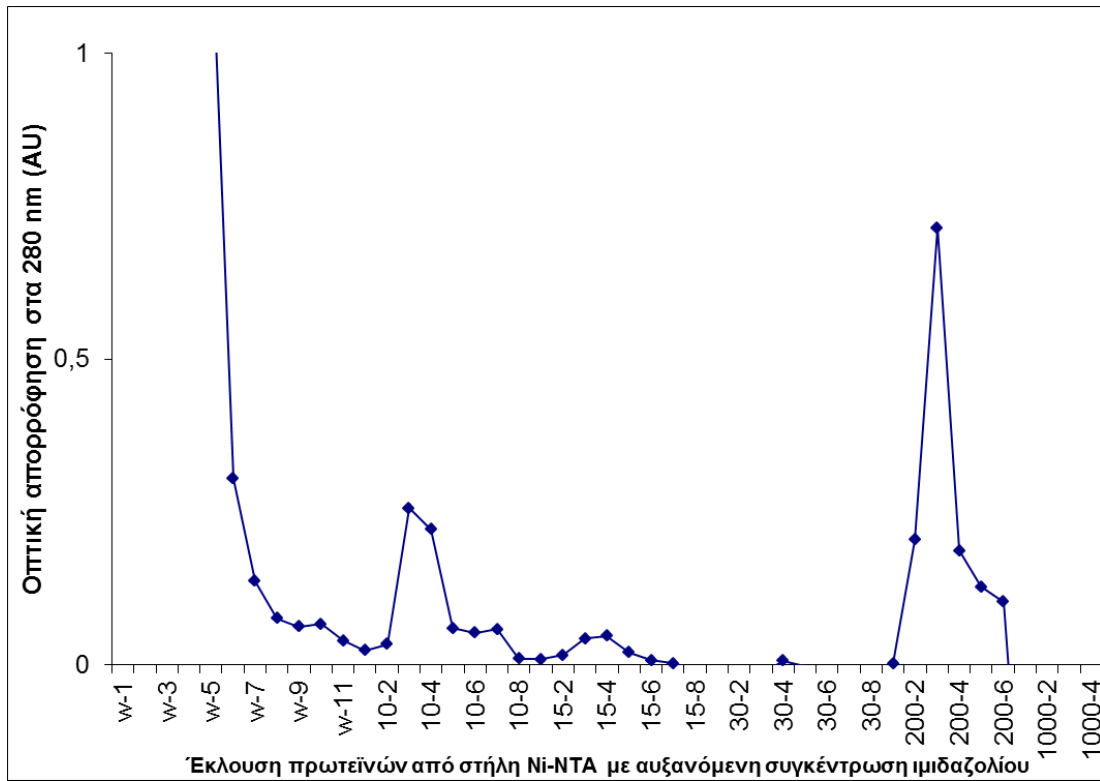
Όπως περιγράφηκε, η Ba3679 (ολόκληρη 213 αμινοξέα) υποκλωνοποιήθηκε στον φορέα έκφρασης pETM-44, έτσι ώστε να βρίσκεται υπό τον έλεγχο του υποκινητή T7, συζευγμένη στο αμινοτελικό της άκρο με την MBP (Maltose Binding Protein, πρωτεΐνη δέσμευσης της μαλτόζης), και εκφράζονται μαζί με μια ετικέτα 6 καταλοίπων ιστιδίνης που φέρει η MBP προσαρτημένη στο αμινοτελικό της άκρο. Ο ανασυνδυασμένος φορέας έκφρασης (Εικόνα 3.5) χρησιμοποιήθηκε για το μετασχηματισμό επιδεκτικών κυττάρων *E. coli* των στελεχών BL21(DE3), C41(DE3), C43(DE3) και Rosetta 2 (DE3). Από τους μετασχηματισμένους κλώνους του στελέχους C41(DE3) που αναπτύχθηκαν σε στερεό θρεπτικό μέσο με δείκτη επιλογής την καναμυκίνη έγινε επιλογή ενός κλώνου, ο οποίος χρησιμοποιήθηκε για καλλιέργειες έκφρασης. Η επιλογή του κλώνου βασίστηκε στην ένταση του σήματος των κλώνων κατά την εμφάνιση στην ανοσοαποτύπωση σε μεμβράνη νιτροκυτταρίνης (βλ. §2.2.2.8), καθώς η ισχυρότερη έκφραση υποδηλώνεται από το

ισχυρότερο σήμα (**Εικόνα 3.7**). Η έκφραση επάγεται με προσθήκη IPTG στο υγρό θρεπτικό μέσο ανάπτυξης, που προκαλεί την ενεργοποίηση του υποκινητή T7.

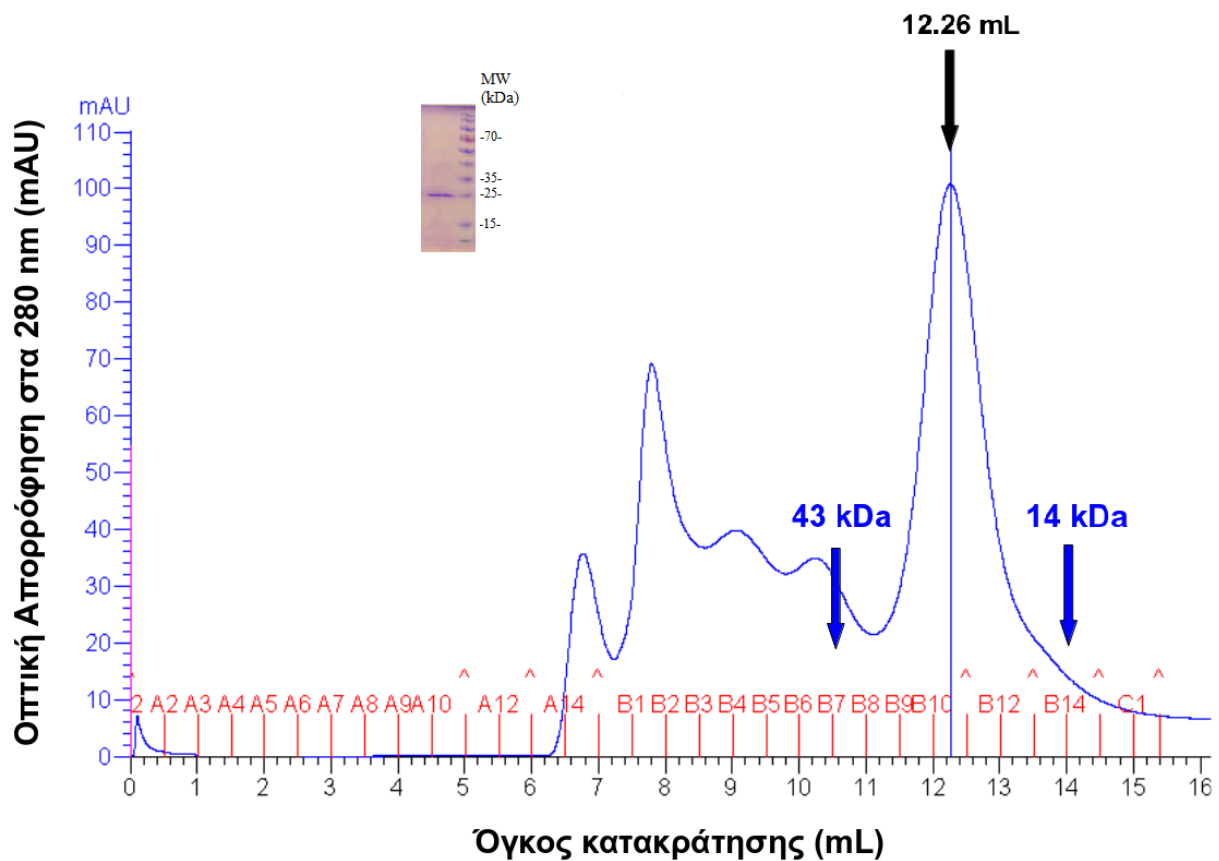
Η μεθοδολογία της απομόνωσης μέσω χρωματογραφίας συγγένειας και μοριακού αποκλεισμού περιγράφονται αναλυτικά στο Κεφάλαιο 2 (βλ. §2.2.2.4 και §2.2.2.5). Συνοπτικά, μετά το πέρας της επαγωγής έκφρασης έγινε συλλογή και λύση των κυττάρων, απομόνωση, το πρωτεϊνικό διάλυμα, επώαστηκε με 1 mL στήλης Ni-NTA και υποβλήθηκε σε χρωματογραφία συγγένειας, κατά την οποία έγινε απομόνωση των ανασυνδυασμένων πρωτεϊνών που φέρουν ετικέτα ιστιδινών (**Εικόνα 3.8**). Ακολούθησε η πρωτεολυτική πέψη με την HRV 3C πρωτεάση χρησιμοποιώντας μια μονάδα ενζύμου ανά 100 μg πρωτεΐνης στους 4°C επί 12 ώρες. Τέλος ακολούθησε χρωματογραφία μοριακού αποκλεισμού της Ba3679 (**Εικόνα 3.9**). Εκτός των μονομερών της πρωτεΐνης Ba3679, που αποτελούν τον κύριο υποπληθυσμό, εκλούνται και ολιγομερή, πιθανότατα λόγω μη ειδικών αλληλεπιδράσεων μεταξύ των μορίων Ba3679, αλλά και η ελάχιστη ποσότητα της πρωτεάσης HRV 3C η οποία όντας συζευγμένη με GST έχει αναμενόμενο όγκο έκλυσης ~10mL.



Εικόνα 3.7 Αποτύπωμα κηλίδας για την επιλογή του ισχυρότερου κλώνου ως προς την έκφραση της ανασυνδυασμένης πρωτεΐνης 6xHis-MPB-HRV3C-Ba3679. Η ανοσοαποτύπωση έγινε με πρωτογενές anti-His μονοκλωνικό αντίσωμα και HRP-συζευγμένου αντι-μυός δευτερογενές αντίσωμα.



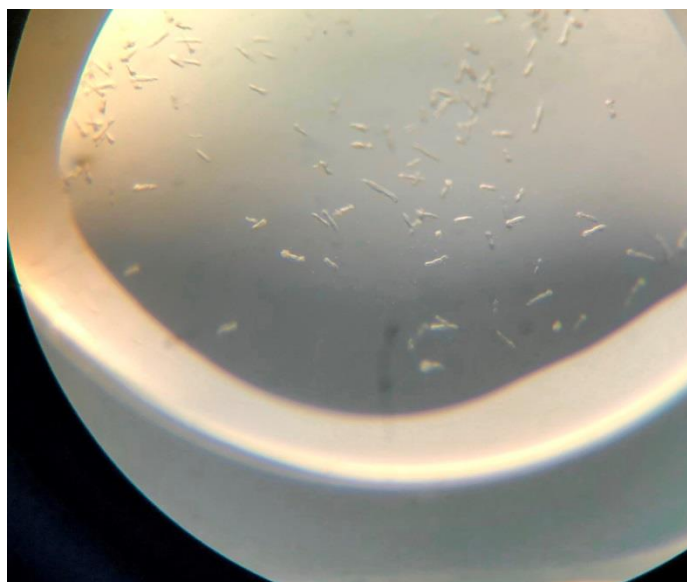
Εικόνα 3.8. Χρωματογραφία συγγενείας της 6xHis-MBP-HRV3C-Ba3679 προσδεδεμένης σε στήλη Ni-NTA. Γράφημα της έκλυσης των προσδεδεμένων πρωτεϊνών που πραγματοποιήθηκε με ρυθμιστικό διάλυμα αυξανόμενης συγκέντρωσης ιμιδαζολίου. Επιτελέστηκαν πλύσεις με 10, 15 και 30 mM ιμιδαζολίου και έκλυση της προσδεδεμένης πρωτεΐνης με 200 mM ιμιδαζολίου.



Εικόνα 3.9. Χρωματογραφία μοριακού αποκλεισμού της Ba3679. Χρωματογράφημα της Ba3679 που απομονώθηκε από χρωματογραφία συγγένειας μετά την επώαση με HRV 3C πρωτεάση. Τα μπλε βέλη δείχνουν τον όγκο της κορυφής έκλυσης πρότυπων πρωτεϊνών γνωστού μοριακού βάρους, ενώ το μαύρο δηλώνει τον όγκο έκλυσης της Ba3679 στις ίδιες συνθήκες [Στήλη Superdex75, ρυθμιστικό διάλυμα: 25 mM Tris HCl pH 7,5, 150 mM NaCl, 0,05% NaN₃, ροή: 0,5 ml/min]. Εκτός των μονομερών της πρωτεΐνης Ba3679, που αποτελούν τον κύριο υποπληθυσμό, εκλύονται και ολιγομερή πιθανότατα λόγω μη ειδικών αλληλεπιδράσεων μεταξύ των μορίων Ba3679. **Ένθετη εικόνα:** SDS-PAGE και χρώση με Coomassie brilliant blue από το κλάσμα των μονομερών.

3.1.4. Πειράματα κρυστάλλωσης

Η σάρωση συνθηκών κρυστάλλωσης της Ba3679 πραγματοποιήθηκε χρησιμοποιώντας αυτοματοποιημένη μηχανή κρυστάλλωσης Oryx4 (Douglas Instruments Ltd.) με τη μέθοδο διάχυσης των ατμών και την τεχνική της επικαθήμενης σταγόνας (όπως αναφέρθηκε στην §2.2.3.1). Τα πειράματα κρυστάλλωσης απέδωσαν αλλά οι καλύτεροι κρύσταλλοι (**Εικόνα 3.10**) δεν υπερέβησαν το μέγεθος των 10 μm, δεν είχαν σταθερή μορφολογία και πιθανότατα ήταν πολλαπλοί. Επακόλουθα πειράματα περίθλασης έδειξαν ότι περαιτέρω ανάλυση της δομής δεν ήταν εφικτή.

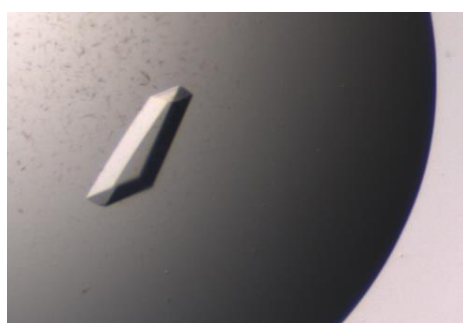


Εικόνα 3.10. Αντιπροσωπευτικοί κρύσταλλοι Ba3679.

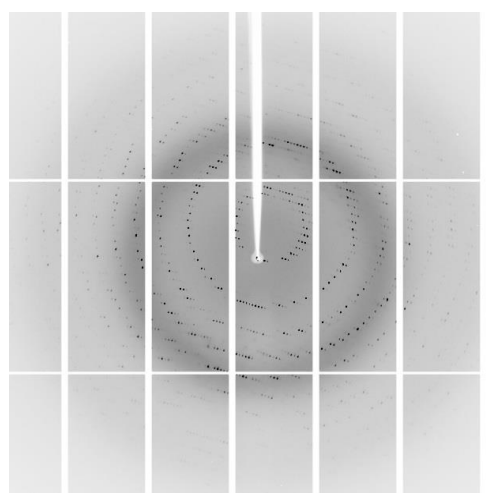
3.2 Μελέτη δομής της φερόμενης ως απακετυλάση πολυσακχαρίτη Ba0330 του *Bacillus anthracis*

3.2.1 Κρυστάλλωση – Βελτιστοποίηση κρυστάλλων

Η σάρωση για την εύρεση συνθηκών κρυστάλλωσης της Ba0330 πραγματοποιήθηκε χρησιμοποιώντας το αυτοματοποιημένο σύστημα κρυστάλλωσης Oryx4 (Douglas Instruments Ltd.) με τη μέθοδο διάχυσης των ατμών και την τεχνική της επικαθήμενης σταγόνας (βλ. § 2.2.3.1). Κρύσταλλοι πρωτεΐνης εμφανίστηκαν μόνο σε μία από τις 700 συνθήκες που δοκιμάστηκαν και στη συνέχεια ακολούθησε βελτιστοποίηση για την παράγωγη κρυστάλλων κατάλληλων για πειράματα περίθλασης. Τελικώς, οι βέλτιστες συνθήκες κρυστάλλωσης επιτεύχθηκαν μετά από ανάμιξη ίσων όγκων διαλύματος πρωτεΐνης συγκέντρωσης 20 mg mL^{-1} σε $50 \text{ mM Tris-HCl pH } 7,6$ 300 mM NaCl και μητρικού διαλύματος 100 mM διμεθυλαρσινικό νάτριο (κακοδυλικό νάτριο), $\text{pH } 6,5$ και $20\% \text{ w/v PEG } 3350$ ^a. Η εξισορρόπηση της σταγόνας έγινε με $50 \mu\text{L}$ του δεύτερου διαλύματος. Οι κρύσταλλοι (**Εικόνα 3.11α**) εμφανίστηκαν εντός 6 ημερών φθάνοντας το τελικό τους μέγεθος σε 2 εβδομάδες στους 16°C . Παραμονή των σταγόνων σε υψηλότερες θερμοκρασίες οδήγησε σε ταχύτατη ανάπτυξη πολλών μικρότερου μεγέθους κρυστάλλων, που περιθλούσαν σε σαφώς μικρότερη διακριτικότητα, ενώ μετά από πολυήμερη παραμονή των σταγόνων στους 4°C έδωσαν άμορφο ίζημα. Δοκιμασίες συγκρυστάλλωσης και εμβάπτισης της Ba0330 με τους προσδέτες (βλ. παράρτημα Α') δεν απέδωσαν κρυστάλλους συμπλόκων.



(α)



(β)

Εικόνα 3.11 (α) Κρύσταλλος της Ba0330. (β) Εικόνα περίθλασης ακτίνων-X από μονοκρύσταλλο της Ba0330.

^a Όλες οι τιμές συγκέντρωσης αναφέρονται στα αρχικά διαλύματα και όχι σε αυτά που προκύπτουν μετά την ανάμιξη τους.

3.2.2 Πειράματα περίθλασης - Συλλογή κρυσταλλογραφικών δεδομένων

Για τη συλλογή δεδομένων απομονώθηκε μονοκρύσταλλος, ο οποίος αφού εμβάπτιστηκε σε μητρικό διάλυμα το οποίο περιείχε 20% (v/v) αιθυλενογλυκόλη για περίπου 10 s, μεταφέρθηκε σε ρεύμα αζώτου θερμοκρασίας 100 K. Η επιλογή του κρυοπροστατευτικού έγινε μετά από δοκιμή των ενώσεων μεθυλοπεντανοδιόλη (MPD), PEG 400 και αιθυλενογλυκόλης σε συγκεντρώσεις 20-30% v/v σε μητρικό διάλυμα. Η τελική συγκέντρωση του κρυοπροστατευτικού και ο χρόνος εμβάπτισης του κρυστάλλου στο διάλυμα του, προέκυψε μετά από συμβιβασμό ανάμεσα στην επίτευξη της μέγιστης δυνατής πληρότητας δεδομένων και στην πρόκληση ελάχιστης μωσαϊκότητας του κρυστάλλου, αφού χαμηλότερες τιμές συγκέντρωσης κρυοπροστατευτικού οδηγούσαν στο σχηματισμό έντονων δακτυλίων πάγου, ενώ υψηλότερες τιμές επηρέαζαν την ακεραιότητα του κρυσταλλικού πλέγματος. Δεδομένα περίθλασης συλλέχθηκαν είτε αμέσως μετά την εμβάπτιση των κρυστάλλων στο κρυοπροστατευτικό, ή αφού είχε προηγηθεί τοποθέτηση των κρυστάλλων σε νάυλον βρόχο και εμβάπτισή τους σε υγρό άζωτο, ώστε να μεταφερθούν στην τερματική γραμμή συλλογής δεδομένων περίθλασης. Όλα τα πειράματα περίθλασης και τα κρυσταλλογραφικά δεδομένα συνελέγησαν με ακτινοβολία συγχρότρου στο κέντρο SLS του ινστιτούτου PSI (Villigen, Ελβετία). Η συλλογή των τελικών δεδομένων πραγματοποιήθηκε στην τερματική γραμμή ακτίνων-X X06DA με ακτινοβολία μήκους κύματος 1,000 Å. Η συλλογή έγινε σε ανιχνευτή δύο διαστάσεων (Dectris PILATUS 2M) με τη μέθοδο περιστροφής του κρυστάλλου (rotation method) σε μέγιστη διακριτικότητα 1,48 Å. Η μέγιστη διακριτικότητα δεν άλλαξε καθόλου κατά τη διάρκεια της συλλογής δεδομένων δείχνοντας ότι η φθορά λόγω ακτινοβολίας (radiation damage) ήταν ανεπαίσθητη, συνεπικουρώντας σε αυτό ο πολύ σύντομος χρόνος συλλογής που παρέχεται πλέον στις σύγχρονες εγκαταστάσεις. Συνελέγησαν 1080 εικόνες (**Εικόνα 3.11 β**) με απόσταση κρυστάλλου – ανιχνευτή 130 mm (υψηλής διακριτικότητας δεδομένα) με γωνιακή περιστροφή ($\Delta\phi$) μόλις 0,3°. Από την πρώτη εικόνα περίθλασης προσδιορίστηκε ότι ο κρύσταλλος ανήκε σε μονοκλινές κρυσταλλικό σύστημα με διαστάσεις μοναδιαίας κυψελίδας $a=108,9 \text{ \AA}$, $b=68,8 \text{ \AA}$, $c=132,0 \text{ \AA}$ και $\beta=96,86^\circ$. Η επεξεργασία, ολοκλήρωση και κανονικοποίηση των δεδομένων έγινε με το πρόγραμμα XDS (Kabsch 2010) και η ομάδα χώρου προσδιορίστηκε με το POINTLESS (Evans 2006) για το σύνολο των ασυγχώνευτων δεδομένων (unmerged data). (**Πίνακας 3.1**). Οι εντάσεις των ανακλάσεων μετατράπηκαν σε μέτρα παραγόντων δομής $|F_{hkl}|$ με το πρόγραμμα XDS. Ποσοστό 5% εξ αυτών ορίστηκε ως δεδομένα αναφοράς, τα οποία δε χρησιμοποιήθηκαν

κατά τη βελτιστοποίηση της δομής, αλλά χρησίμευσαν στην εξαγωγή του συντελεστή Rfree (Bruenger 1977). Τα κρυσταλλογραφικά δεδομένα συνοψίζονται στον **Πίνακα 3.1**.

Πίνακας 3.1. Συλλογή και επεξεργασία δεδομένων

Πηγή ακτίνων-X	X06DA, SLS
Μήκος κύματος (Å)	1,000
Θερμοκρασία (K)	100
Ανιχνευτής	Dectris PILATUS 2M
Εύρος περιστροφής ανά εικόνα (°)	0,3
Συνολικό εύρος (°)	324
Χρόνος έκθεσης ανά εικόνα (s)	0,3
Ομάδα χώρου	C2
Μοναδιαία κυψελίδα a, b, c (Å)	108,96, 68,76, 132,02
α, β, γ (°)	90,00, 96,86, 90,00
Μωσαϊκότητα (°)	0,4
Εύρος ανάλυσης (Å)	48,00-1,48 (1,53-1,48)
Συνολικός αριθμός ανακλάσεων	714399
Αριθμός μοναδικών ανακλάσεων	155203
Πολλαπλότητα	4,6 (4,2)
Πληρότητα δεδομένων (%)	96,3 (93,2)
$\langle I/\sigma(I) \rangle$	17,8 (2,0)
R _{meas}	0,062 (0,670)
CC _{1/2}	0,995 (0,710)
Overall B factor από διάγραμμα Wilson (Å ²)	15,19

Οι τιμές εντός των παρενθέσεων αντιστοιχούν στη στιβάδα της υψηλότερης διακριτικότητας.

Πίνακας 3.2. Επίλυση και βελτιστοποίηση δομής

Ανακλάσεις που χρησιμοποιήθηκαν για τη βελτιστοποίηση	147443
Ανακλάσεις που χρησιμοποιήθηκαν για το δείκτη αξιοπιστίας R-free	7760
R-work	0,157 (0,279)
R-free	0,175 (0,292)
Αριθμός μη-υδρογονικών ατόμων	
Πρωτεΐνης	5140
Προσδετών	34
Ύδατος	943

Αμινοξέα πρωτεΐνης	636
r.m.s.(μήκη δεσμών) (Å)	0,010
r.m.s.(γωνιών) (°)	1,024
Ramachandran περιοχές	
ευνοϊκές (%)	97,5
Μη επιτρεπτές (%)	0,5
MolProbity Clashscore	2,79
Μέσος παράγοντας B (Å ²)	
Πρωτεΐνης	23,7
Προσδέτες	26,6
Διάλυμα	33,3
Αριθμός TLS groups	5

Οι τιμές εντός των παρενθέσεων αντιστοιχούν στη στιβάδα της υψηλότερης διακριτικότητας

$$R_{meas} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_{j=1}^n |I_{hkl,j} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

$$R_{cryst} = \frac{\sum_{hkl} |F_{hkl}^{obs} - F_{hkl}^{calc}|}{\sum_{hkl} F_{hkl}^{obs}}$$

3.2.3 Επίλυση δομής – Βελτιστοποίηση

Η δομή επιλύθηκε με τη μέθοδο της μοριακής αντικατάστασης (Rossmann 1972) χρησιμοποιώντας ως αρχικό μοντέλο τη δομή της ομόλογης απακετυλάσης του βακτηρίου *Bacillus cereus* Bc0361 (κωδικός PDB: 4HD5, Fadoulglou *et al.* 2013). Από το αρχικό μοντέλο του μονομερούς της δομής της Bc0361 αφαιρέθηκαν τα ετεροάτομα (με τον όρο αυτόν υποδηλώνονται όλα τα άτομα πλην της πρωτεϊνικής αλυσίδας, δηλαδή μεταλλικά ιόντα, μόρια ύδατος, κρυοπροστατευτικά κλπ.) και οι πλευρικές αλυσίδες για να αποφευχθεί η εισαγωγή πιθανών σφαλμάτων από το προηγούμενο μοντέλο. Για την επίλυση χρησιμοποιήθηκε το πρόγραμμα PHASER (Mc Coy *et al.* 2007). Το πρόγραμμα αυτό είναι πλήρως αυτοματοποιημένο, καθώς ξεκινώντας από τα κρυσταλλογραφικά δεδομένα και το πρωτεϊνικό μοντέλο που θα χρησιμοποιηθεί για την επίλυση, επιλέγει όλες τις απαιτούμενες παραμέτρους και τελικά επιλέγει λύση περιστροφής κατάλληλα προσανατολισμένη, την οποία εν συνεχεία τοποθετεί στη μοναδιαία κυψελίδα (μεταφορά). Οι παράγοντες δομής

υπολογίζονται για κάθε προσανατολισμό / θέση του μοντέλου μέσω ταχέως μετασχηματισμού Fourier της ηλεκτρονιακής πυκνότητας. Η βέλτιστη λύση που παρήχθη ήταν κατά πολύ καλύτερη από τις αμέσως επόμενες, έχοντας τιμές αξιοπιστίας για τις συναρτήσεις περιστροφής και μεταφοράς σχεδόν διπλάσιες αυτών των υπολοίπων λύσεων. Η επίλυση της δομής, όπως αρχικά υπέδειξε ο υπολογισμός του συντελεστή Matthews (Matthews 1968) έδειξε την παρουσία δύο μορίων πρωτεΐνης ανά ασύμμετρη μονάδα. Η επιθεώρηση του διαφορικού χάρτη Fo-Fc ηλεκτρονιακής πυκνότητας, μετά από βελτιστοποίηση άκαμπτου σώματος (rigid body refinement), οδήγησε στη βέλτιστη τοποθέτηση των πλευρικών αλυσίδων των αμινοξέων και ακολούθως τοποθέτηση των μορίων ύδατος και κρυοπροστατευτικού στην πυκνότητα που αντιστοιχούσε. Οι βρόχοι χτίστηκαν στην ηλεκτρονιακή πυκνότητα στον πραγματικό χώρο και βελτιστοποιήθηκαν τοπικά βάσει επιτρεπτών τιμών για τα μήκη των χημικών δεσμών, γωνιών και διεδρων γωνιών. Το ιόν μετάλλου συμπεριλήφθηκε χωρίς επιπρόσθετους περιορισμούς λόγω της υψηλής διακριτικότητας.

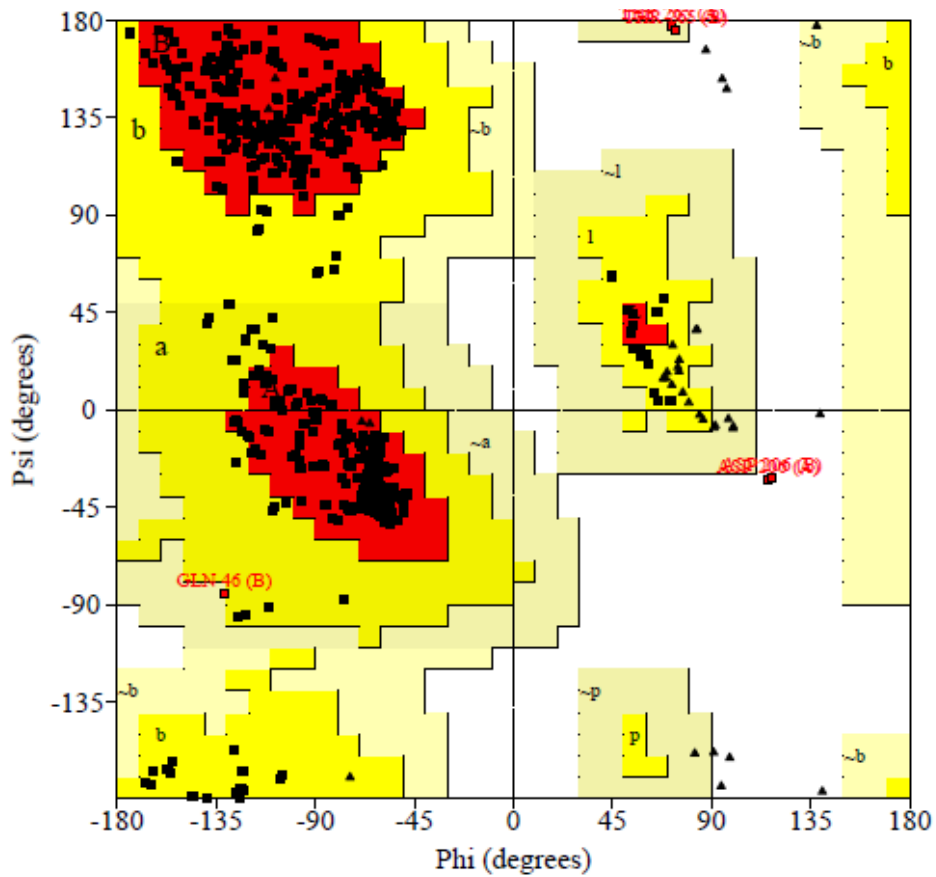
Η βελτιστοποίηση της δομής πραγματοποιήθηκε αρχικώς με τη μέθοδο της μεγιστοποιημένης πιθανότητας μέσω του προγράμματος PHENIX (Adams *et al.* 2010) εφαρμόζοντας διαδοχικούς κύκλους βελτιστοποίησης με ιστροπικούς θερμικούς παράγοντες, διακοπτόμενους από τακτικές επιθεωρήσεις των χαρτών ηλεκτρονιακής πυκνότητας 2Fo-Fc και Fo-Fc. Όπου κρίθηκε απαραίτητο ακολούθησε διόρθωση του μοντέλου με το πρόγραμμα COOT (Emsley *et al.* 2010).

Για την πολυπεπτιδική αλυσίδα εφαρμόστηκαν όλοι οι προτεινόμενοι από το πρόγραμμα περιορισμοί. Τα μόρια του διαλύτη τοποθετήθηκαν με αυτόματη επιλογή του προγράμματος και όπου κρίθηκε αναγκαίο με επιθεώρηση του διαφορικού χάρτη Fourier. Από την παραπάνω διαδικασία προέκυψε μοντέλο του οποίου ο συντελεστής R ήταν ~16% και ο ελεύθερος συντελεστής R_{free} ~18%. Η τελική δομή συνέκλινε σε πολύ καλά στατιστικά (Πίνακας 3.2)

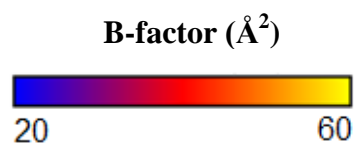
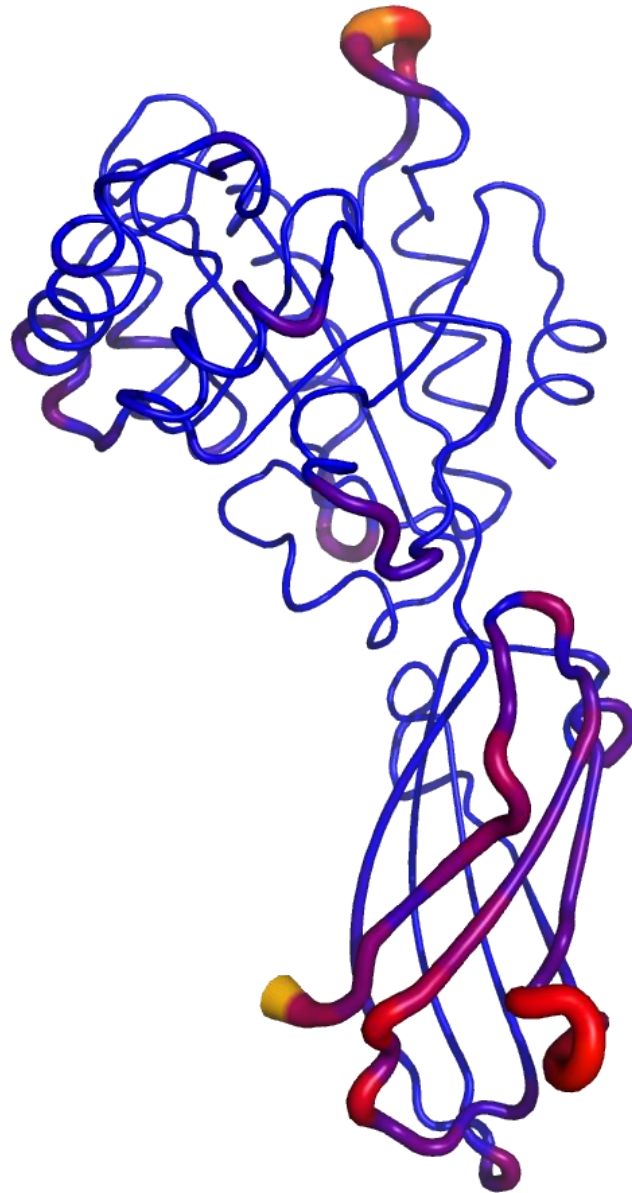
Μεταξύ των μορίων της ασύμμετρης μονάδας υπήρξε σχεδόν τέλεια υπέρθεση με μέση τετραγωνική ρίζα απόκλισης (r.m.s.d.) 0,234 Å για τα κύρια άτομα αλυσίδας. Πληροφορίες για τα αποτελέσματα της βελτιστοποίησης και το τελικό μοντέλο παρατίθενται στον πίνακα 3.2.

3.2.4 Περιγραφή δομής τριών διαστάσεων της Ba0330

Το γονίδιο της Ba0330, κωδικεύει πρωτεΐνη 360 αμινοξέων, εκ των οποίων στο προς κρυστάλλωση προϊόν βρίσκονταν τα 337 (από το κατάλοιπο 24 έως το 360). Στο τελικό μοντέλο που προέκυψε από την κρυσταλλική δομή, εμφανίστηκαν τα 316 κατάλοιπα καθώς η περιοχή 24-44 του αμινοτελικού άκρου δεν ανιχνεύθηκε στην ηλεκτρονιακή πυκνότητα. Ως εκ τούτου η αντίστοιχη περιοχή δε χτίστηκε στο μοντέλο. Το διάγραμμα Ramachadran υπολογισμένο από το πρόγραμμα PROCHECK (Laskowski *et al.* 1993) περιέχει 97,5% των καταλοίπων σε ευνοούμενες και επιτρεπόμενες περιοχές, ενώ 2 αμινοξέα λαμβάνουν σχεδόν μη επιτρεπτή διαμόρφωση (0,3%) (**Εικόνα 3.12**). Στην ανάλυση αυτή δεν περιλήφθηκαν τα κατάλοιπα γλυκίνης και προλίνης. Ουσιαστικά όλοι οι δεσμοί και οι γωνίες του πρωτεϊνικού μοντέλου, βρίσκονται εντός των ορίων που έχουν καθορισθεί από δομές υψηλής διακριτικότητας μακρομορίων (Engh & Huber 1991). Οι ατομικές συντεταγμένες και οι παράγοντες δομής της κρυσταλλικής δομής της Ba0330 κατατέθηκαν στην δημόσια βάση δεδομένων Protein Data Bank (PDB, Berman *et al.* 2000), από τους Giastas, P., Andreou, A., Bouriotis, V., και Eliopoulos, E. και έχουν καταχωρηθεί υπό τον κωδικό αριθμό καταχώρησης 4V33. Ο μέσος θερμικός παράγοντας (B factor) της δομής είναι $25,2 \text{ \AA}^2$. Στην **εικόνα 3.13** αναπαρίσταται ο θερμικός παράγοντας ανά αμινοξύ για τα άτομα της κύριας αλυσίδας όπου γίνεται φανερό ότι τα αμινοξέα είναι καλά εντοπισμένα. Οι βρόχοι β6-β7 και β9-β10 εμφανίζουν υψηλές τιμές για τους θερμικούς παράγοντες υποδηλώνοντας υψηλή κινητικότητα των περιοχών αυτών.

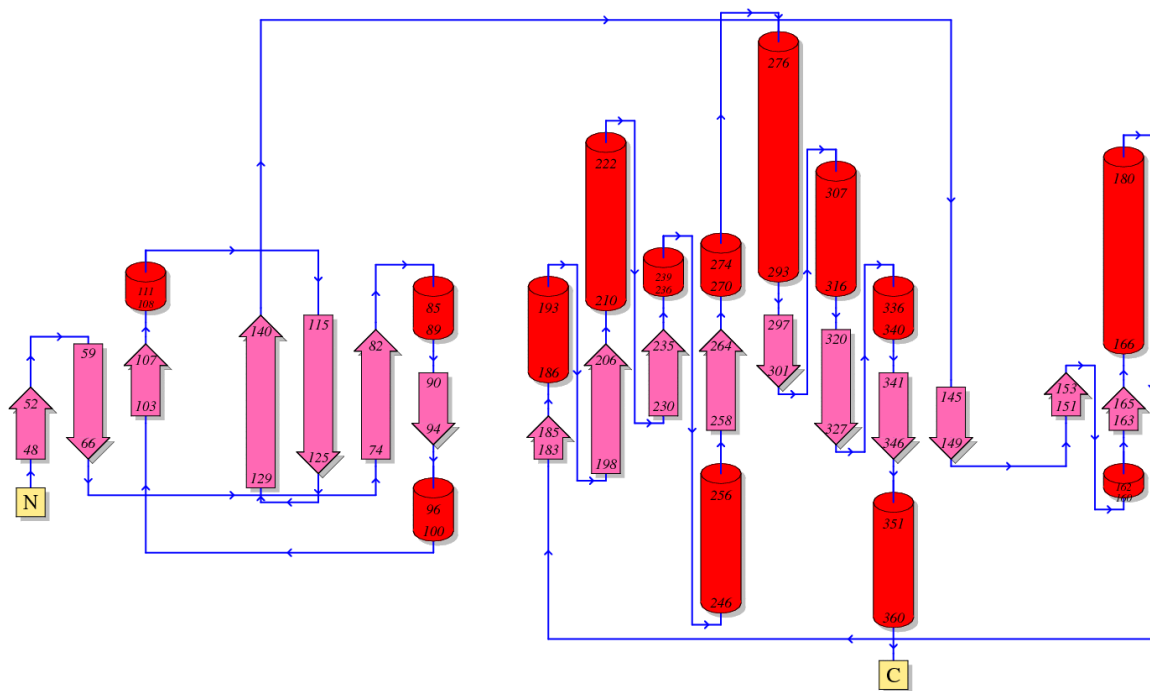


Εικόνα 3.12 Το διάγραμμα Ramachandran της δομής της Ba0330. Ο υπολογισμός του διαγράμματος έγινε με το πρόγραμμα PROCHECK. (ενοούμενες περιοχές - A, B, L, επιπρόσθετες επιτρεπόμενες περιοχές - a, b, l, p, γενικά επιτρεπόμενες περιοχές ~a,~b,~l,~p οι απαγορευμένες περιοχές είναι λευκές).

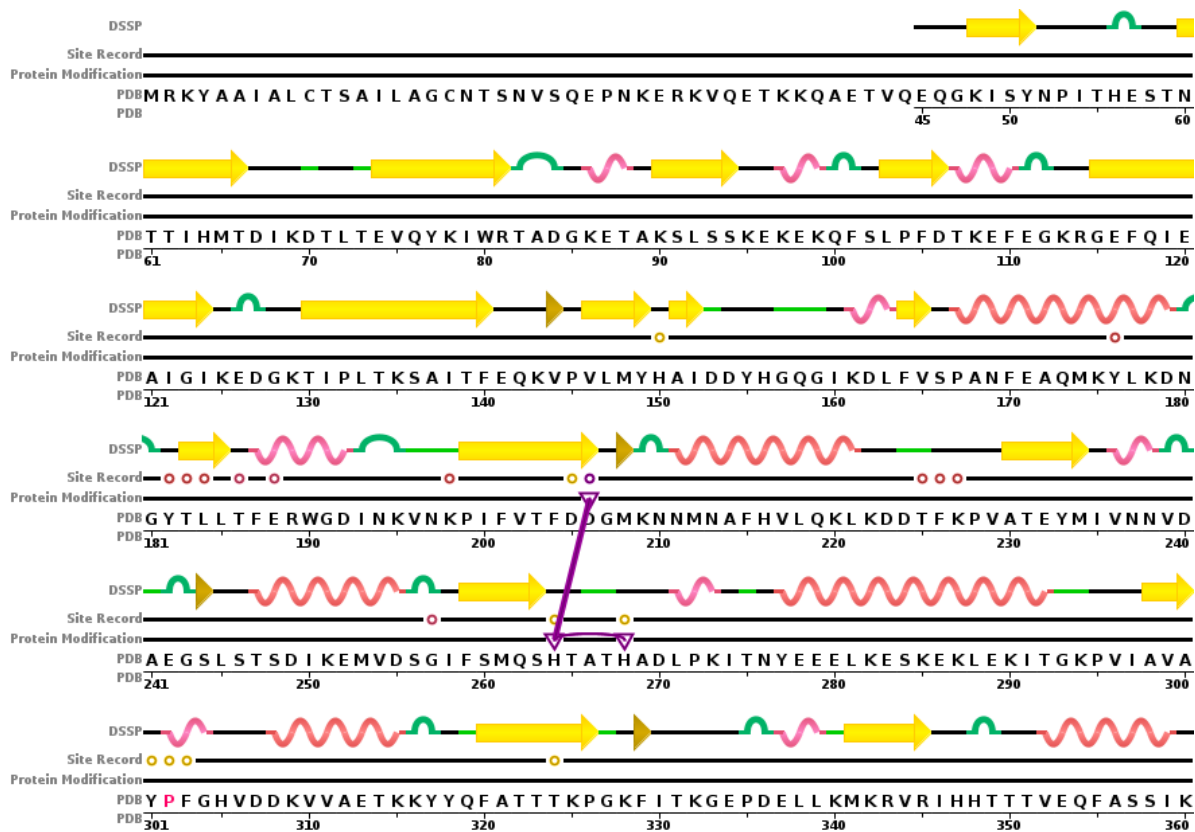


Εικόνα 3.13 Οι θερμικοί παράγοντες για τα άτομα της κύριας αλυσίδας της Ba0330.

Το διάγραμμα τοπολογίας και τα δευτεροταγή δομικά χαρακτηριστικά της δομής της Ba0330 δίδονται στις εικόνες 3.14 και 3.15. Η κρυσταλλική δομή της Ba0330 προσδιορίστηκε στα 1,48Å η οποία αποκάλυψε δύο δομικές ενότητες (Εικόνα 3.16). Επιπρόσθετα με αυτήν της εστεράσης τύπου CE4 (ή αλλιώς επονομαζόμενη NodB) (αμινοξικά κατάλοιπα 142-360), υπάρχει στο αμινοτελικό άκρο και μια δομική ενότητα παρόμοια με ινωδονεκτίνης τύπου III (fibronectin type III, Fn3) (αμινοξικά κατάλοιπα 45-141) που αποτελείται από δύο β-πτυχωτές επιφάνειες (4 + 3) αμινοξέων που σχηματίζουν το μοτίβο του β-σάντουιτς με τοπολογία ελληνικού κλειδιού (greek key). Η δομική ενότητα CE4 μοιάζει εκείνων από άλλες απακετυλάσες, που περιέχουν τον καλά σχηματισμένο αύλακα με το ιόν ψευδαργύρου να βρίσκεται στον πυθμένα. Η κλασική Asp-His-His διευθέτηση διατηρείται, υποδηλώνοντας τη δυνατότητα δέσμευσης ενός μεταλλικού κατιόντος στην κοιλότητα δέσμευσης.

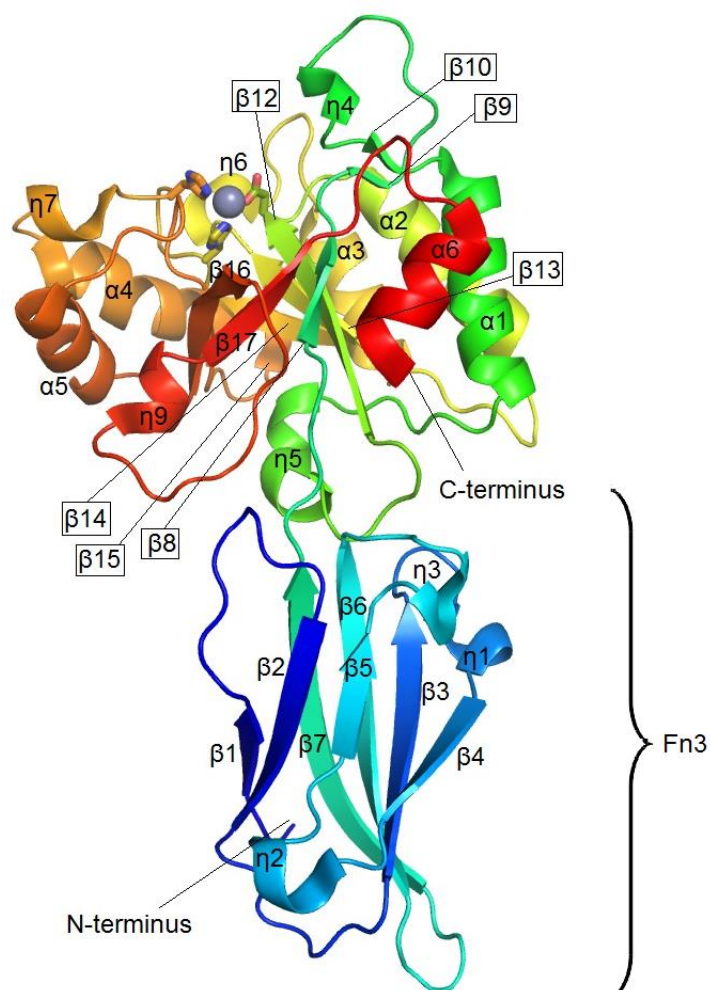


Εικόνα 3.14. Συγκρότηση δευτεροταγούς δομής της Ba0330.



- τόν ψευδαργύρου
- β-γέφυρα
- στροφή
- βρόχος
- β-κλώνος
- 3/10 έλικα
- α-έλικα

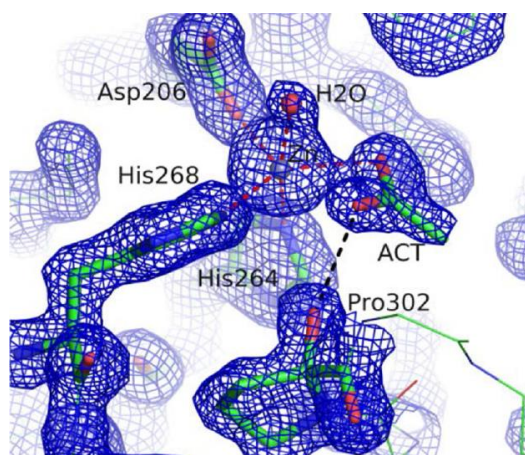
Εικόνα 3.15. Γραμμική απεικόνιση των δομικών χαρακτηριστικών της δευτεροταγούς δομής στην αμινοξική ακολουθία της Ba0330 (PDB ID: 4V33).



Εικόνα 3.16. Η δομή της Ba0330. Οι έλικες αντιπροσωπεύονται με κορδέλες, οι β-κλώνοι με βέλη και οι βρόχοι με χορδές. Τα στοιχεία της δευτεροταγούς δομής χρωματίζονται με χρωματική κλίμακα από μπλε (αμινοτελικό άκρο) έως κόκκινο (καρβοξυτελικό άκρο). Η δομική ενότητα NodB παρουσιάζεται στο πάνω μέρος της εικόνας και η δομική ενότητα Fn3 παρουσιάζεται στο κάτω μέρος. Τα αμινοξικά κατάλοιπα της συντηρημένης τριάδας συναρμογής του μεταλλοϊόντος που την αποτελούν ένα ασπαρτικό οξύ (Asp213) και δύο κατάλοιπα ιστιδίνης (His271 και His275) απεικονίζονται με ράβδους. Το ιόν Zn^{2+} εμφανίζεται ως σφαίρα χρώματος γκρι.

Η πυκνότητα ηλεκτρονίων στη θέση πρόσδεσης αποδόθηκε στο ιόν ψευδαργύρου (**Εικόνα 3.17**), το οποίο, εκτός των δύο ιστιδινών και του ασπαρτικού οξέος, αλληλεπιδρά με ένα μόριο ύδατος και ένα ιόν οξικού, σχηματίζοντας διάταξη τριγωνικής διπυραμίδας. Το ιοντισμένο μόριο οξικού οξέος πιθανώς υπάρχει ως υπολειμματική ουσία του διαλύματος της πολυαιθυλενογλυκόλης που χρησιμοποιείται στο ρυθμιστικό διάλυμα της κρυστάλλωσης (Blair *et al.* 2005). Το ιόν οξικού ήταν καλά διατεταγμένο (παρόν για άνω του 6σ στον Fo-Fc

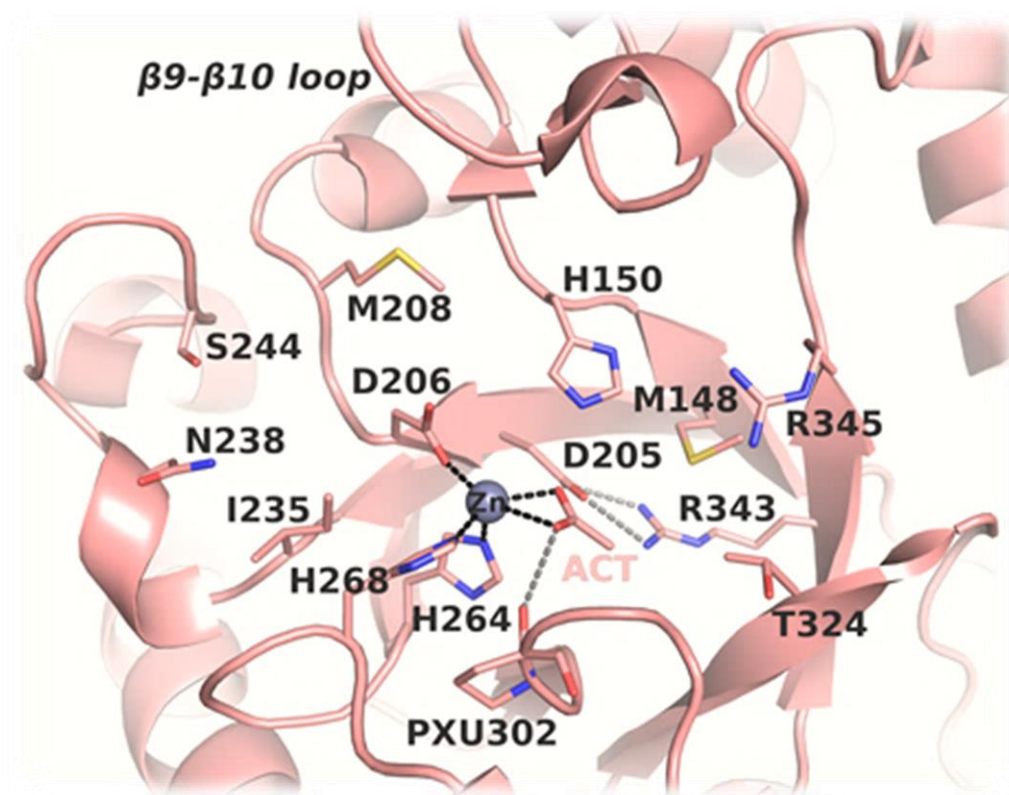
χάρτη), με μέσο συντελεστή θερμικών παραγόντων 27 \AA^2 , κοντά στην αντίστοιχη τιμή του ιόντος ψευδαργύρου. Το οξικό ανιόν αλληλεπιδρά εκτεταμένα με κατάλοιπα από τη θέση πρόσδεσης.



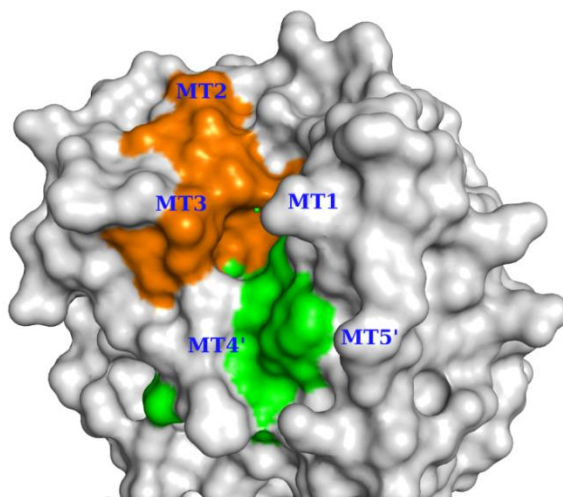
Εικόνα 3.17 Ηλεκτρονιακή πυκνότητα (χάρτης $2\text{Fo}-\text{Fc}$) στην περιοχή του ψευδαργύρου σε 2σ , υπολογισμένος με τις τελικές συντεταγμένες του μοντέλου. Η εικόνα σχεδιάστηκε με το πρόγραμμα PYMOL. Με χρώμα πράσινο παριστάνονται τα άτομα C, με κόκκινο τα άτομα O, με κυανό τα άτομα N. Ο ψευδάργυρος απεικονίζεται με μια γκρι σφαίρα και η κόκκινη σφαίρα δηλώνει ένα μόριο ύδατος. Το οξικό μόριο (ACT) απεικονίζεται με αναπαράσταση ράβδου. Η χαρακτηριστική καταλυτική αμινοξική τριάδα (His, His, και Asp) παρουσιάζεται στην Ba0330 στη θέση δέσμευσης του ψευδαργύρου στα πρότυπα των PDAs.

Ένα από τα οξυγόνα του μορίου οξικού αλληλεπιδρά με το υψηλά συντηρημένο Asp205, το οποίο επίσης αλληλεπιδρά ηλεκτροστατικά με την Arg343 (**Εικόνα 3.18**). Επί πλέον, το ίδιο άτομο οξυγόνου φαίνεται να αλληλεπιδρά με τη δομικά συντηρημένη His150. Το άλλο οξυγόνο του οξικού σχηματίζει έναν ισχυρό δεσμό υδρογόνου με την ομάδα υδροξυλίου της τροποποιημένης προλίνης (Pro302), η οποία χτίστηκε στο μοντέλο ως α -υδροξυ-L-προλίνη τοποθετημένη σε απόσταση $3,6 \text{ \AA}$ από το μεταλλικό ιόν. Παρόμοια παρατήρηση έγινε και κατά την ανάλυση της δομής της ορθόλογης Bc0361 από τον *B. cereus*, που στη θέση δέσμευσης περιέχει μια τροποποιημένη προλίνη στον Ca, η οποία στο τελικό μοντέλο βρίσκεται ως α -υδροξυ-L-προλίνη (Fadoulglou *et al.* 2017). Ο μηχανισμός τροποποίησης της προλίνης αυτής, καθώς και η λειτουργική σημασία της, εάν υφίσταται, δεν

είναι γνωστά. Η ομάδα μεθυλίου του οξικού καταλαμβάνει μια μικρή υδρόφοβη περιοχή που δημιουργείται από τη μεθυλομάδα της Thr324 και της Met148. Η κρυσταλλική δομή της Ba0330 στην δομική ενότητα της εστεράσης έχει μια καλά διαμορφωμένη κοιλότητα δέσμησης που είναι επενδεδυμένη με κάποια υδρόφοβα κατάλοιπα τα οποία προσανατολίζονται προς τα κατάλοιπα της θέσης δέσμησης (**Εικόνα 3.18**). Τα πέντε συντηρημένα μοτίβα MT1-MT5 που σχηματίζουν την κοιλότητα δέσμησης δίδονται στην **εικόνα 3.19**.



Εικόνα 3.18. Η φερόμενη κοιλότητα δέσμησης της Ba0330. Στο μέσο της εικόνας βρίσκεται το ιόν ψευδαργύρου (γκρι σφαίρα). Εμφανίζονται τα κατάλοιπα συναρμογής ψευδαργύρου (Asp206, His264 και His268) καθώς και κάποια από τα καταλυτικά κατάλοιπα. Το μόριο οξικού (ACT) εμφανίζεται με αναπαράσταση ράβδου. Η σφαίρα συναρμογής του ψευδαργύρου της Ba0330 παρουσιάζεται με μαύρες διακεκομμένες γραμμές και με γκρι διακεκομμένες γραμμές η ηλεκτροστατική αλληλεπίδραση της αργινίνης Arg343 με το ασπαρτικό οξύ (Asp205). Στο επάνω μέρος της εικόνας εμφανίζεται ο βρόχος β9-β10.



Εικόνα 3.19. Τα MT1-MT5' της κοιλότητας δέσμησης στην Ba0330.

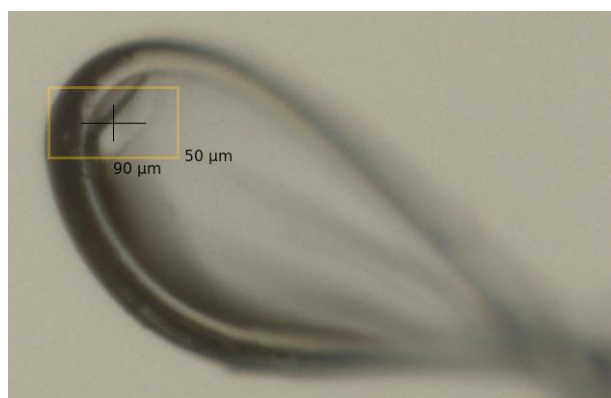
3.3 Μελέτη δομής της φερόμενης ως απακετυλάση πολυσακχαρίτη Ba0331 του *Bacillus anthracis*

3.3.1 Κρυστάλλωση – Βελτιστοποίηση κρυστάλλων

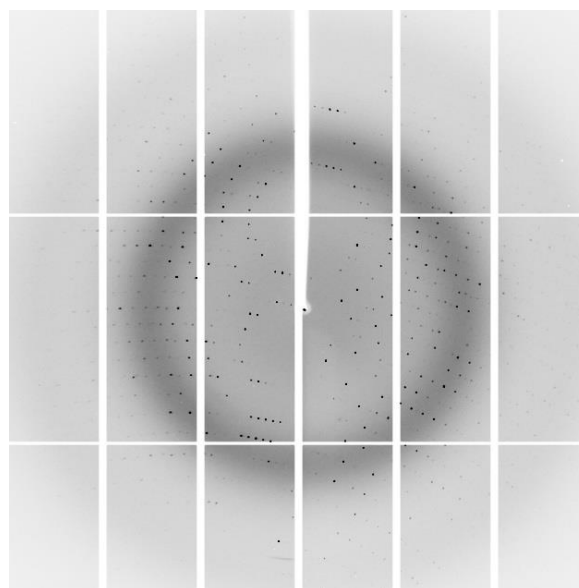
Η πρωτεΐνη Ba0331 από το βακτήριο *B. anthracis* κρυσταλλώθηκε στους 16° C με τη χρήση αυτόματης συσκευής κρυστάλλωσης Oryx4 της εταιρείας Douglas Instruments Ltd και τη μέθοδο της διάχυσης ατμών και την τεχνική της επικαθήμενης σταγόνας. Αρχικά, εμφανίστηκαν κρύσταλλοι σε μια από τις 500 συνθήκες που δοκιμάστηκαν σε διάφορες θερμοκρασίες (100 mM Tris HCl pH 7,5, 200 mM θειικό λίθιο και 5% w/v PEG 4000), που ήταν όμως κακής ποιότητας όσον αφορούσε στη μορφολογία αλλά και στην ικανότητα περίθλασης. Εκτενείς προσπάθειες βελτιστοποίησης των συνθηκών κρυστάλλωσης με χρήση σποράς μικροκρυστάλλων (microseeding) οδήγησε σε δημιουργία ραβδόμορφων κρυστάλλων κατάλληλων για πειράματα σκέδασης ακτίνων-X. Οι καλύτεροι κρύσταλλοι αναπτύχθηκαν με ανάμιξη ίσων όγκων διαλύματος πρωτεΐνης (συγκέντρωσης 25 mg ml⁻¹ σε 50mM Tris-HCl, pH 7,6, 200 mM NaCl) και μητρικού διαλύματος που περιείχε 100 mM Tris-HCl pH 8,0, 150 mM θειικού λιθίου, 5 mM σπερμιδίνης και 30% w/v PEG 4000. Οι κρύσταλλοι εμφανίστηκαν μέσα σε έξι ημέρες και απέκτησαν το τελικό τους μέγεθος σε δέκα ημέρες στους 16°C. Πριν την έκθεσή τους στην δέσμη των ακτίνων-X (**Εικόνα 3.20α**) εμβάπτιστηκαν για περίπου ~10 s σε διάλυμα κρυοπροστασίας που περιείχε το μητρικό διάλυμα παρουσία 25% v/v αιθυλικής γλυκόλης.

3.3.2 Πειράματα περίθλασης - Συλλογή κρυσταλλογραφικών δεδομένων

Το σύνολο δεδομένων συλλέχθηκε σε διακριτικότητα 2,6 Å υπό κρυογονικές συνθήκες που εξασφαλιζόνταν από ρεύμα συνεχούς παροχής αερίου αζώτου (100 K) στον τερματικό σταθμό εργασίας X06DA του συγχρότρου Swiss Light Source (Villigen, Switzerland), με χρήση μονοχρωματικής ακτινοβολίας μήκους κύματος 1,000 Å. Η δεικτοδότηση, ολοκλήρωση και κανονικοποίηση των εικόνων περίθλασης (**Εικόνα 3.20β**) έγινε με το XDS (Kabsch, 2010) σε ομάδα χώρου που υπέδειξε το POINTLESS (Evans 2006) από την ανάλυση μη συγχωνευμένων δεδομένων. Πληροφορίες που αφορούν στη συλλογή των κρυσταλλογραφικών δεδομένων δίδονται στον **Πίνακα 3.3**.



(α)



(β)

Εικόνα 3.20 (α) Μονοκρύσταλλος της Ba0331 σε θηλιά στερέωσης μικροκρυστάλλων αμέσως πριν τη συλλογή δεδομένων. **(β)** Εικόνα περίθλασης ακτίνων-X από μονοκρύσταλλο της Ba0331.

Πίνακας 3.3. Συλλογή και επεξεργασία δεδομένων

Πηγή ακτίνων-Χ	X06DA, SLS
Μήκος κύματος (Å)	1,000
Θερμοκρασία (K)	100
Ανιχνευτής	Dectris PILATUS 2M
Εύρος περιστροφής ανά εικόνα (°)	0,3
Συνολικό εύρος (°)	270
Χρόνος έκθεσης ανά εικόνα (s)	0,3
Ομάδα χώρου	$P2_1$
a, b, c (Å)	41,7, 211,1, 46,3
α, β, γ (°)	90,0, 91,3, 90,0
Μωσαϊκότητα (°)	0,4
Εύρος ανάλυσης (Å)	45,25 - 2,59 (2,69 - 2,59)
Συνολικός αριθμός ανακλάσεων	126746 (10953)
Αριθμός μοναδικών ανακλάσεων	24772 (2185)
Πολλαπλότητα	5,1 (5,0)
Πληρότητα δεδομένων (%)	98,62 (94,03)
$\langle I/\sigma(I) \rangle$	15,24 (1,36)
R_{meas}	0,09586 (0,5036)
$CC_{1/2}$	0,996 (0,841)
Μέσος όρος B factor διαγράμματος Wilson (Å ²)	49,34

Οι τιμές εντός των παρενθέσεων αντιστοιχούν στη στιβάδα της υψηλότερης διακριτικότητας.

3.3.3 Επίλυση δομής – Βελτιστοποίηση

Η δομή λύθηκε με μοριακή αντικατάσταση με τη χρήση του προγράμματος PHASER (McCoy *et al.*, 2007) και με μοντέλο τη δομή της Ba0330 (PDB ID: 4V33) και εν συνεχεία βελτιστοποιήθηκε με το PHENIX (Adams *et al.* 2010). Κατά τη διαδικασία αυτή, ο διαφορικός Fo-Fc και ο 2Fo-Fc χάρτης ηλεκτρονιακής πυκνότητας ερμηνεύονταν σε περιοδικά διαστήματα, προκειμένου να επιτευχθεί συμφωνία με το μοντέλο το οποίο αναπροσαρμόστηκε με το COOT (Emsley *et al.* 2010). Πληροφορίες για τα αποτελέσματα της βελτιστοποίησης και το τελικό μοντέλο παρατίθενται στον **πίνακα 3.4**.

Πίνακας 3.4. Επίλυση και βελτιστοποίηση δομής

Ανακλάσεις που χρησιμοποιήθηκαν για τη βελτιστοποίηση	23552 (2268)
Ανακλάσεις που χρησιμοποιήθηκαν για το δείκτη αξιοπιστίας R-free	1240 (113)
R-work	0,183
R-free	0,245
Αριθμός μη-υδρογονικών ατόμων	5294
Μακρομορίων	5182
Προσδετών	25
Διαλύτη	103
Αμινοξέα πρωτεΐνης	636
r.m.s.(δεσμών) (Å)	0,01
r.m.s.(γωνιών) (°)	1,24
Ramachandran περιοχές	
ευνοϊκές (%)	94,78
επιτρεπτές (%)	4,43
Μη-επιτρεπτές (%)	0,79
Μη-επιτρεπτά ροταμερή (%)	0,18
Clashscore	12,75
Μέσος παράγοντας B (Å ²)	63,00
Πρωτεϊνικό μόριο	63,16
Προσδέτες	72,57
Μόρια διαλύτη	53,56
Αριθμός TLS groups	17

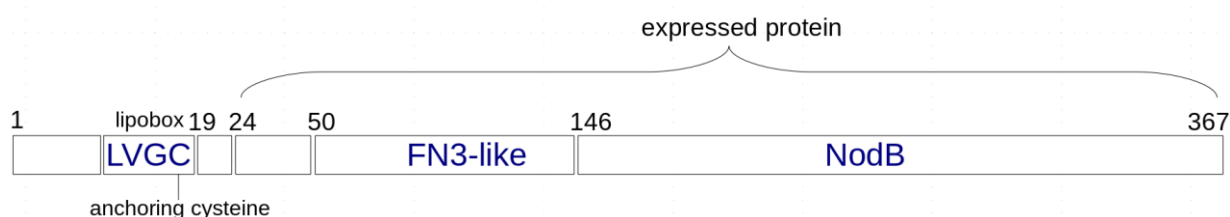
Οι τιμές εντός των παρενθέσεων αντιστοιχούν στη στιβάδα της υψηλότερης διακριτικότητας.

3.3.4 Περιγραφή δομής τριών διαστάσεων της Ba0331

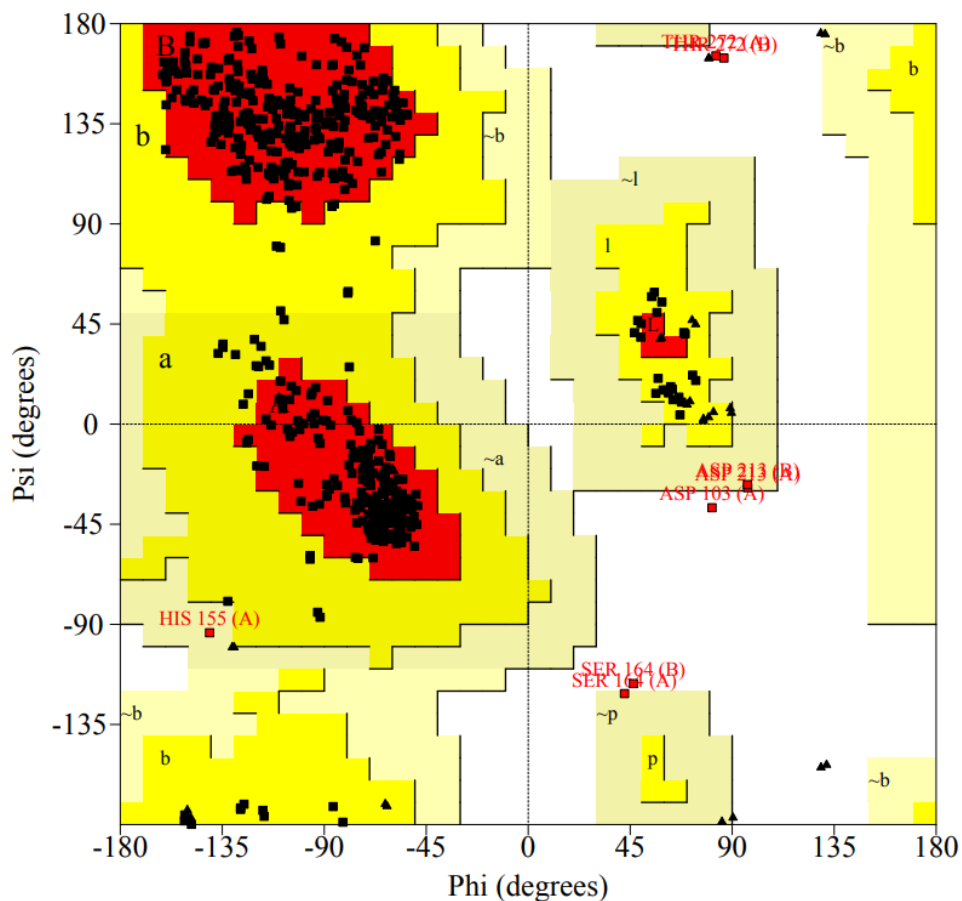
Το γονίδιο της Ba0331, κωδικεύει πρωτεΐνη 367 αμινοξέων, εκ των οποίων στο προς κρυστάλλωση προϊόν βρίσκονταν τα 344 (από το κατάλοιπο 24 έως το 367) (**Εικόνα 3.21**).

Η ασύμμετρη μονάδα περιέχει δυο μόρια Ba0331 τα οποία διαφέρουν ελάχιστα (με r.m.s.d. 0,306 Å για όλα τα άτομα της κύριας αλυσίδας). Το τελικό μοντέλο της Ba0331 περιλαμβάνει τα αμινοξέα 50 έως 367 της πρωτεϊνικής ακολουθίας τα οποία εντοπίστηκαν εντός της παρατηρούμενης ηλεκτρονιακής πυκνότητας και χτίστηκαν σε μια συνεχή αλυσίδα. Η δομή

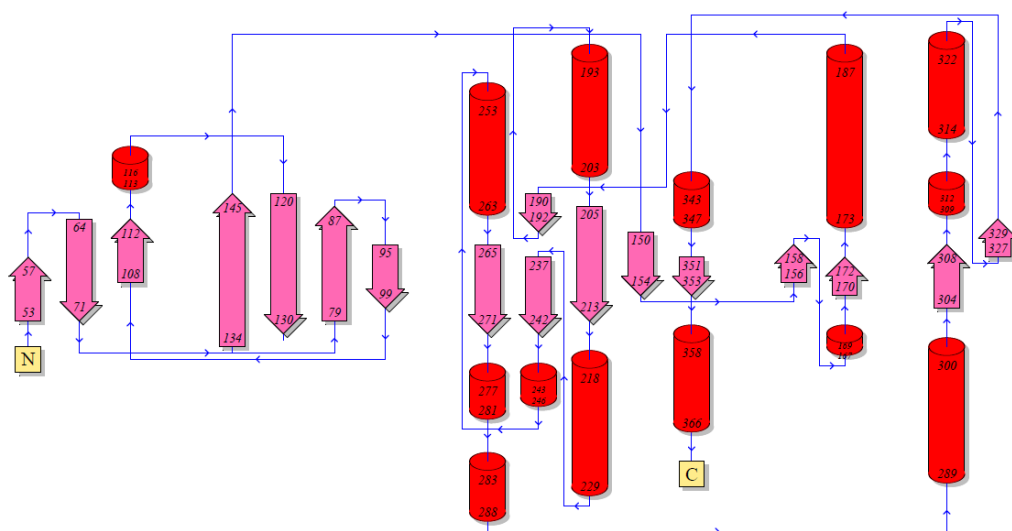
των αμινοξέων 24-49 δε συμπεριλήφθηκε στο τελικό μοντέλο καθώς οι θέσεις τους δεν προσδιορίστηκαν επαρκώς, πιθανότατα λόγω υψηλών παραγόντων θερμικής κίνησης. Το διάγραμμα Ramachandran έδειξε ότι 99,21% των αμινοξέων βρίσκονται ως προς τη στερεοδιάταξη της κύριας αλυσίδας σε ευνοούμενες ή επιτρεπτές περιοχές (**Εικόνα 3.22**). Η βελτιστοποίηση της δομής συνέκλινε στο τελικό μοντέλο με παράγοντα αξιοπιστίας (R factor) 0,1833 (Rfree = 0,2453) και καλή γεωμετρία με r.m.s.d. από τα ιδανικά μήκη δεσμών και γωνιών 0,010 Å και 1,20°, αντιστοίχως. Τα δευτεροταγή δομικά χαρακτηριστικά της δομής της Ba0331 και το διάγραμμα τοπολογίας δίδονται στις **εικόνες 3.23** και **3.24**.



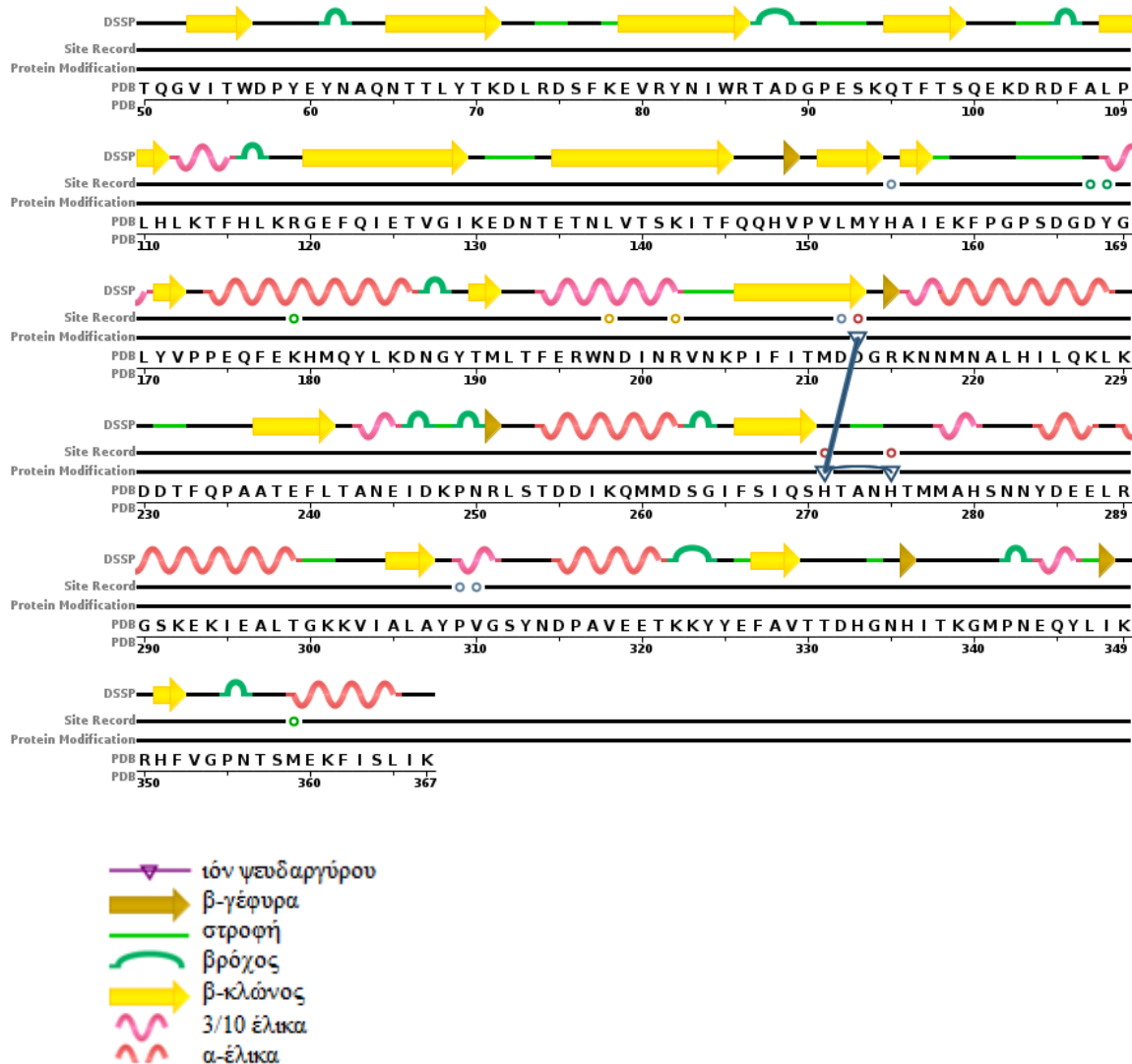
Εικόνα 3.21 Γραμμική αναπαράσταση της αμινοξικής αλληλουχίας της Ba0331 στην οποία επισημαίνονται βασικά χαρακτηριστικά της πρωτεϊνικής δομής. Στη θέση 18 βρίσκεται το αμινοξικό κατάλοιπο της κυστεΐνης που προσδένεται στη μεμβράνη, ενώ δίδεται αναλυτικά η ακολουθία του λιπόφιλου χαρακτηριστικού τμήματος της δομής που προηγείται της Cys. Ακολουθεί το τμήμα της πρωτεΐνης που εκφράστηκε χρησιμοποιώντας τεχνολογία γενετικής μηχανικής όπως δείχνει το άγκιστρο. Η ηλεκτρονική πυκνότητα αποκάλυψε τη κρυσταλλική δομή για την περιοχή 50-367, που περιλαμβάνει τις δομικές ενότητες Fn3 και NodB.



Εικόνα 3.22 Το διάγραμμα Ramachandran για τη δομή της Ba0331. Ο υπολογισμός του διαγράμματος έγινε με το πρόγραμμα PROCHECK. (ενοούμενες περιοχές - A, B, L, επιπρόσθετες επιτρεπόμενες περιοχές - a, b, l, p, γενικά επιτρεπόμενες περιοχές ~a,~b,~l,~p οι απαγορευμένες περιοχές είναι λευκές).



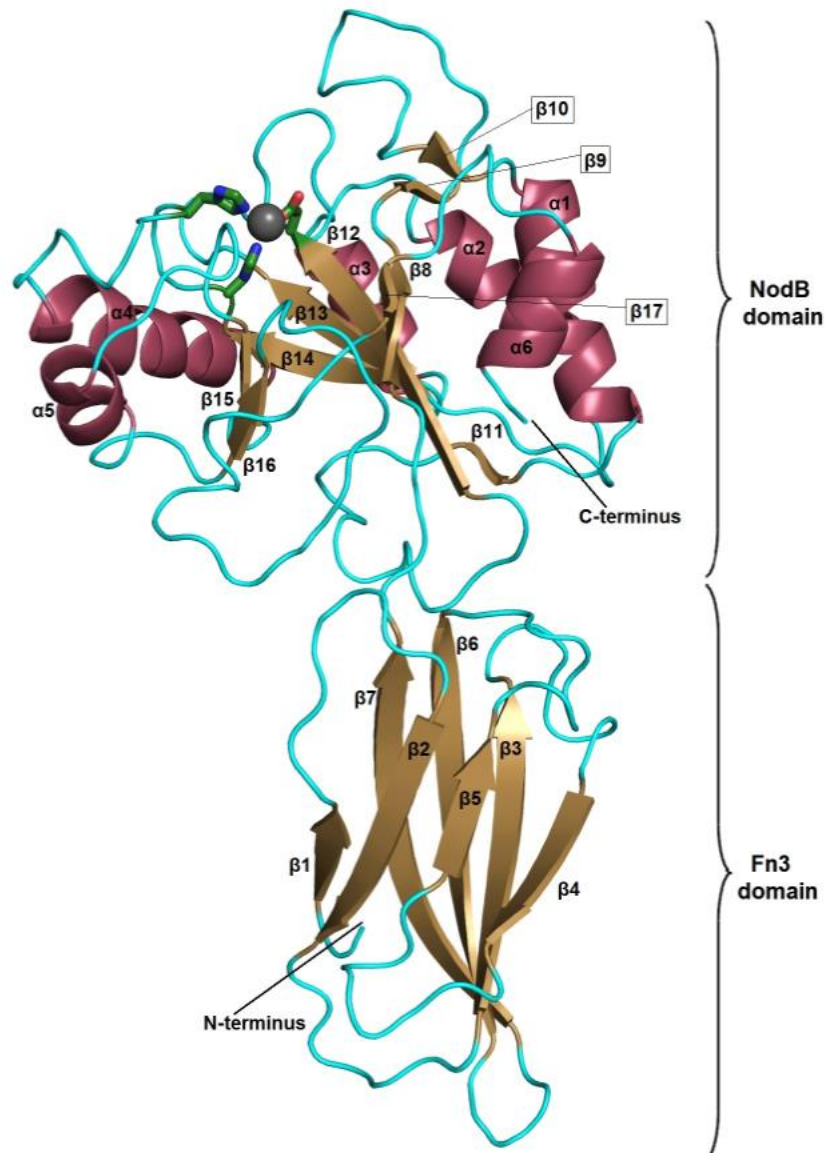
Εικόνα 3.23. Συγκρότηση δευτεροταγούς δομής της Ba0331.



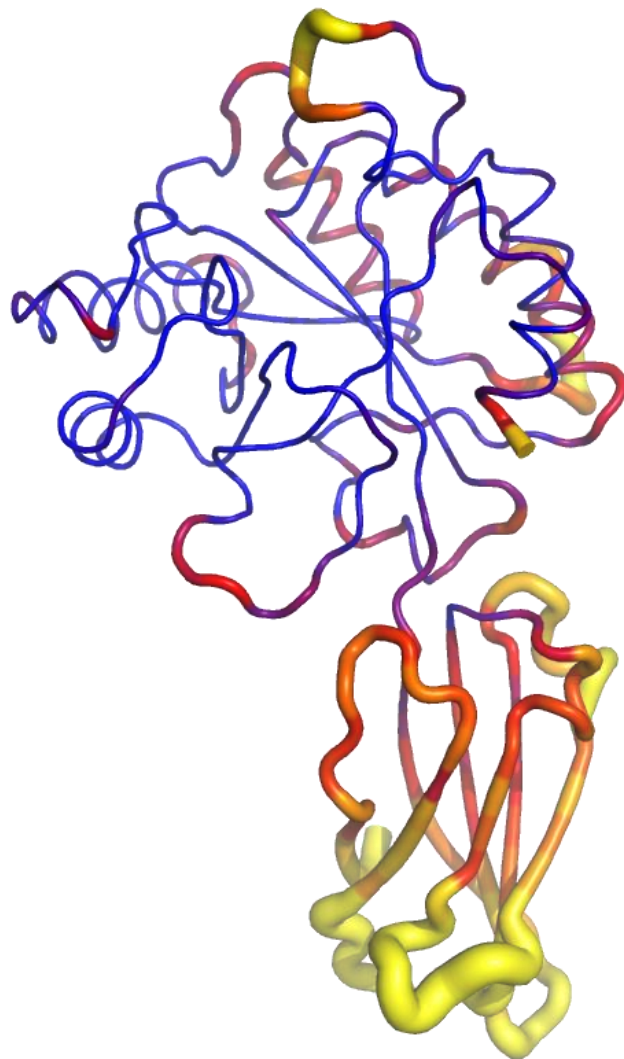
Εικόνα 3.24. Γραμμική απεικόνιση των δομικών χαρακτηριστικών της δευτεροταγούς δομής στην αμινοξική ακολουθία της Ba0331 (PDB ID: 6GO1).

Η κρυσταλλική δομή της Ba0331 (**Εικόνα 3.25**) αποτελείται από δυο διακριτές αλλά αλληλεπιδρούσες δομικές ενότητες (Andreou *et al.* 2019). Το αμινοτελικό άκρο είναι μια παρόμοια με ινωδονεκτίνη τύπου III (fibronectin type III, Fn3-like) δομική ενότητα (αμινοξέα 50-146) αποτελούμενη από δύο β-πτυχωτές επιφάνειες (4 + 3 β-φύλλων) που σχηματίζουν το μοτίβο του β-σάντουιτς με τοπολογία ελληνικού κλειδιού (greek key). Το καρβοξυτελικό άκρο της Ba0331 (αμινοξέα 147-367) υιοθετεί μια (α/β)₆ τοπολογία που αποτελεί χαρακτηριστική αναδίπλωση της NodB ομόλογης δομικής ενότητας, παρούσα στις CE4 εστεράσες. Οι ατομικές συντεταγμένες και οι παράγοντες δομής κατατέθηκαν στην βάση δεδομένων πρωτεϊνικών δομών (Protein Data Bank, <http://www.rcsb.org/pdb>, Berman *et al.*

2000) με κωδικό καταχώρησης 6GO1. Στην **εικόνα 3.26** παριστάνεται ο θερμικός παράγοντας ανά αμινοξύ για την κύρια αλυσίδα όπου γίνεται φανερό ότι τα αμινοξέα της δομικής ενότητας NodB είναι πολύ καλά εντοπισμένα, με αυτά του βρόχου β9-β10 να εμφανίζουν την υψηλότερη κινητικότητα. Αντίθετα, η δομική ενότητα Fn3 εμφανίζεται εν συνόλω πιο κινητική με τους βρόχους β3-β4 και β6-β7 να εμφανίζουν τις υψηλότερες τιμές για τους παράγοντες θερμικής κίνησης.



Εικόνα 3.25 Η δομή της Ba0331. Οι έλικες αντιπροσωπεύονται με κορδέλες χρώματος βυσσινί, οι β-κλώνοι με χρυσά βέλη και οι βρόχοι με κυανές χορδές. Η δομική ενότητα NodB παρουσιάζεται στο πάνω μέρος της εικόνας και η δομική ενότητα Fn3 παρουσιάζεται στο κάτω μέρος. Η συντηρημένη τριάδα των αμινοξικών καταλοίπων συναρμογής του ιόντος ψευδαργύρου (σφαίρα χρώματος γκρι) αποτελείται από ένα ασπαρτικό οξύ (Asp213) και δύο κατάλοιπα ιστιδίνης (His271 και His275) που απεικονίζονται με ράβδους.

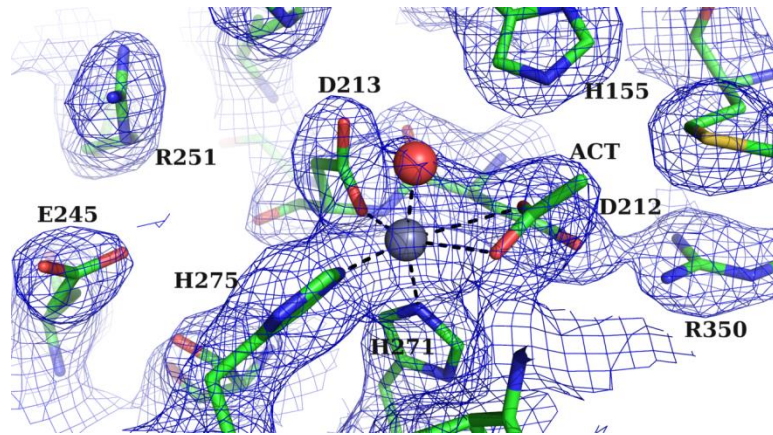


Εικόνα 3.26 Οι θερμοί παράγοντες (B-factors \AA^2) για τα άτομα της κύριας αλυσίδας.

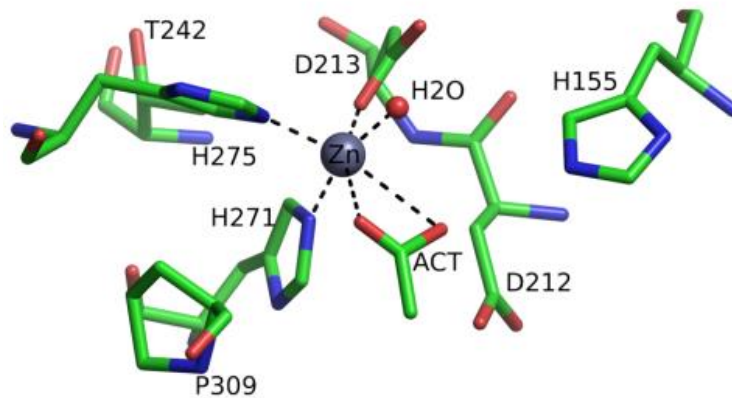
3.3.4.1 Ένα συντηρημένο καταλυτικό κέντρο

Η καρβοξυτελική NodB δομική ενότητα στην Ba0331, έχει έναν καταλυτικό αύλακα όπως και όλα τα μέλη της CE4 οικογένειας, που σχηματίζεται από πέντε συντηρημένα μοτίβα (MT1-MT5, Blair *et al.* 2005 και Arnaouteli *et al.* 2015) με μερικές τοπικές διαφοροποιήσεις που θα συζητηθούν σε επακόλουθη παράγραφο. Στον καταλυτικό αύλακα ένα ιόν ψευδαργύρου σχηματίζει ένωση συναρμογής με την υψηλά συντηρημένη, στην CE4 οικογένεια, τριάδα αμινοξικών καταλοίπων, Asp-His-His. Επίσης, συνδέεται με ένα οξικό ανιόν που εμφανίζεται καθαρά στον χάρτη ηλεκτρονιακής πυκνότητας (εμφανές σε 8σ στον διαφορικό Fo-Fc χάρτη) (**Εικόνα 3.27**) με μέση τιμή παράγοντα θερμικής κίνησης B-factor $\sim 62\text{\AA}^2$, η οποία είναι όμοια εκείνης των αμινοξέων συναρμογής του ψευδαργύρου. Το ανιόν οξικού οξέος είναι ένα από τα προϊόντα της αντίδρασης της απακετυλίωσης, αλλά υπάρχει στην κρυσταλλική δομή της Ba0331 πιθανώς ως υπολειμματική ουσία του διαλύματος της αιθυλενογλυκόλης που χρησιμοποιείται στο διάλυμα της κρυστάλλωσης (Blair *et al.* 2005). Ομοίως, στις δομές των ομολόγων Ba0330 και Bc1974, που προσδιορίστηκαν σε υψηλή διακριτικότητα το μόριο του οξικού βρέθηκε επίσης προσδεμένο στο ιόν ψευδαργύρου παρά την απουσία του από το αρχικό μίγμα συστατικών της κρυστάλλωσης. Η σχεδόν οκταεδρική γεωμετρία συναρμογής του ψευδαργύρου συμπληρώνεται από ένα μόριο ύδατος (**Εικόνες 3.28 και 3.29**).

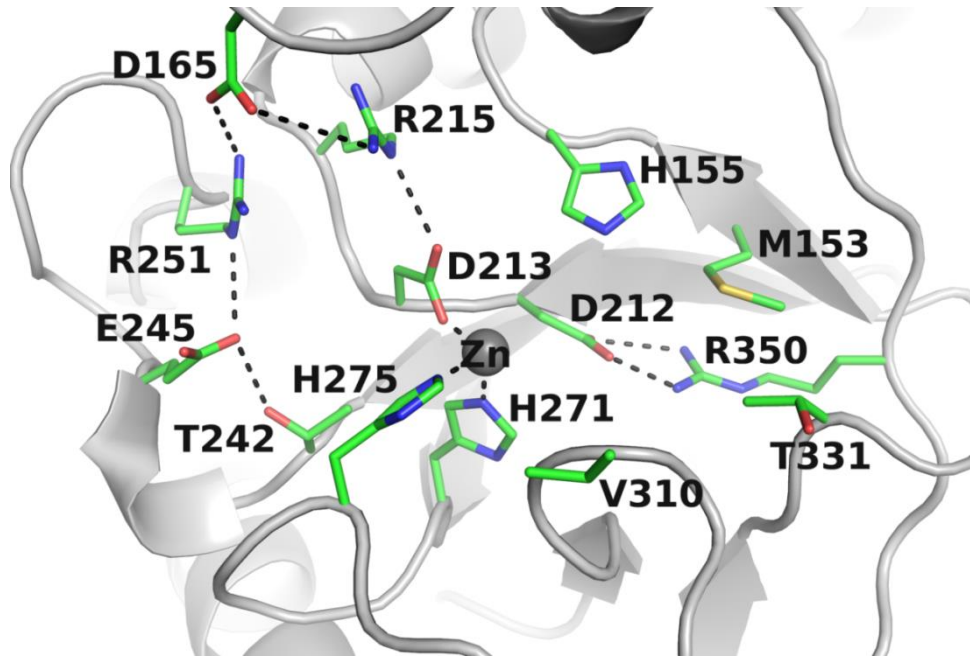
Η φερόμενη θέση πρόσδεσης του υποστρώματος στη NodB δομική ενότητα είναι ένας αύλακας με κατεύθυνση που σχηματίζει γωνία $\sim 45^\circ$ με το μεγάλο άξονα του μορίου της Ba0331 (**Εικόνα 3.30**).



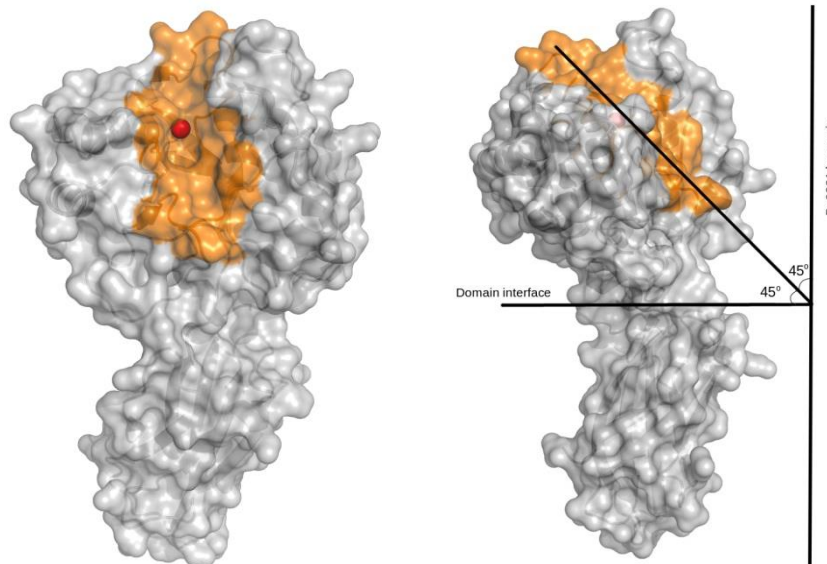
Εικόνα 3.27. Ηλεκτρονιακή πυκνότητα (χάρτης 2Fo-Fc) στην περιοχή του ψευδαργύρου σε επίπεδο 1,2σ, υπολογισμένη με τις τελικές συντεταγμένες του μοντέλου. Με χρώμα πράσινο παριστάνονται τα άτομα C, με κυανό τα άτομα N, με κόκκινο τα άτομα O, με κίτρινο τα άτομα S. Ο ψευδάργυρος απεικονίζεται με μια γκρι σφαίρα. Τα κατάλοιπα Asp213, His271, His275, ένα μόριο οξικού (ACT) και ένα μόριο ύδατος (κόκκινη σφαίρα) αποτελούν το περιβάλλον συναρμογής του ψευδαργύρου, με παραμορφωμένη οκταεδρική γεωμετρία. Επιπροσθέτως, απεικονίζονται τα πιθανά καταλυτικά κατάλοιπα Asp212, His155, Glu245, Arg251 και Arg350 που βρίσκονται σε άμεση γειτνίαση με το ιόν ψευδαργύρου.



Εικόνα 3.28. Παρουσίαση του σχήματος συναρμογής του ιόντος ψευδαργύρου (σφαίρα χρώματος γκρι). Η αμινοξική τριάδα Asp-His-His (απεικόνιση με ράβδους), το ιόν οξικού οξέος (ACT) και ένα μόριο ύδατος (εμφανίζεται ως κόκκινη σφαίρα). Μερικά από τα υποτιθέμενα καταλυτικά κατάλοιπα του δικτύου απεικονίζονται επίσης με ράβδους.



Εικόνα 3.29. Το φερόμενο ενεργό κέντρο της Ba0331 γύρω από το ιόν ψευδαργύρου. Περιέχει τα κατάλοιπα συναρμογής ψευδαργύρου (Asp213, His271 και His275) καθώς και κάποια από τα καταλυτικά κατάλοιπα. Το οξικό ιόν δεν παρουσιάζεται για λόγους σαφήνειας.



Εικόνα 3.30 Αναπαράσταση της επιφάνειας του μορίου της Ba0331. Η φερόμενη κοιλότητα δέσμευσης του υποστρώματος (που δείχνεται με πορτοκαλί χρώμα) της δομικής ενότητας NodB βρίσκεται να ολισθαίνει προς τα κάτω υπό γωνία $\sim 45^\circ$ σε σχέση με το μακρύ άξονα του μορίου. Ο ψευδάργυρος (κόκκινου χρώματος σφαίρα) εκτίθεται στον πυθμένα της κοιλότητας. Η δομική ενότητα NodB βρίσκεται στο πάνω μέρος του μορίου και η δομική ενότητα Fn3 βρίσκεται στο κάτω μέρος της εικόνας.

3.4 Μελέτη δομής της απακετυλάσης της πεπτιδογλυκάνης Bc1974 του *Bacillus cereus*.

3.4.1 Κρυστάλλωση – Βελτιστοποίηση κρυστάλλων

Η σάρωση για την εύρεση συνθηκών κρυστάλλωσης της απακετυλάσης της πεπτιδογλυκάνης του βακτηρίου *B. cereus*, Bc1974 πραγματοποιήθηκε με τη μέθοδο της διάχυσης ατμών με τη χρήση αυτόματης συσκευής κρυστάλλωσης Oryx4 (Douglas Instruments Ltd.) χρησιμοποιώντας τη τεχνική της επικαθήμενης σταγόνας. Παράλληλα δοκιμάστηκε η κρυστάλλωσή της και χειρονακτικά με τη τεχνική της κρεμάμενης σταγόνας και της επικαθήμενης σταγόνας. Πρωτεϊνικοί κρύσταλλοι σχηματίστηκαν σε μια μόνο από τις περίπου 1400 διαφορετικές συνθήκες που δοκιμάστηκαν, η οποία κατόπιν βελτιστοποιήθηκε προς παραγωγή κρυστάλλων κατάλληλων για πειράματα περίθλασης ακτίνων-X. Οι βέλτιστοι κρύσταλλοι αναπτύχθηκαν μετά από ανάμιξη ίσων όγκων πρωτεϊνικού διαλύματος συγκέντρωσης 22 mg mL^{-1} σε 25 mM Mes-NaOH, pH 6,5, 200 mM NaCl με μητρικό διάλυμα αποτελούμενο από: 100 mM κιτρικού νατρίου σε pH 5,6, 30% w/v PEG 4000, 10% αιθανόλης. Οι κρύσταλλοι εμφανίστηκαν σε διάστημα 7-10 ημερών και έλαβαν το τελικό τους μέγεθος σε 2 εβδομάδες μετά από επώαση στους 16°C (Εικόνα 3.31α).

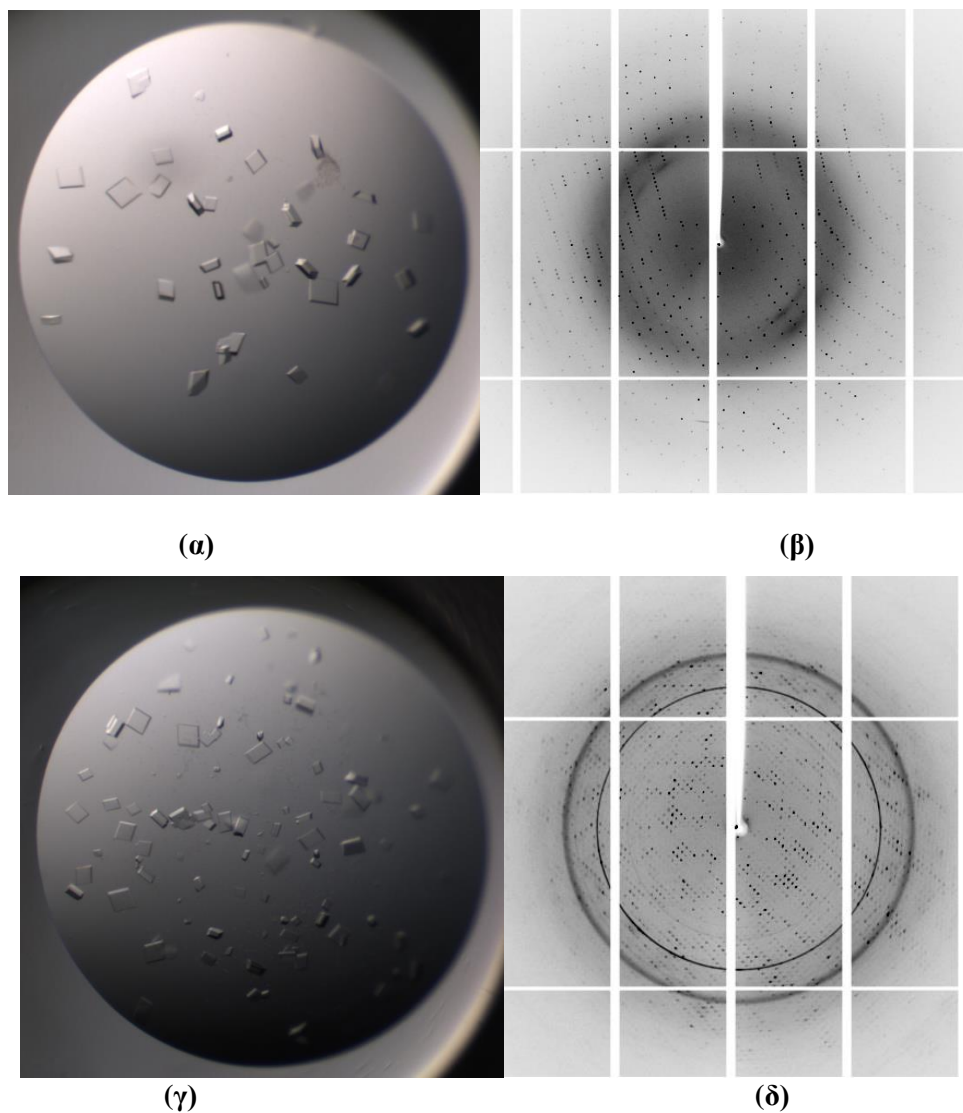
3.4.1.1 Κρύσταλλοι συμπλόκων της Bc1974 με προσδέτες

Οι κρύσταλλοι των συμπλόκων της Bc1974 με τους προσδέτες, σε όλες τις περιπτώσεις αναπτύχθηκαν κάτω από τις ίδιες συνθήκες κρυστάλλωσης σύμφωνα με το πρωτόκολλο που ακολουθεί. Οι προσδέτες διαλυτοποιήθηκαν σε διμέθυλο σουλφοξείδιο (DMSO) προς παραγωγή διαλύματος συγκέντρωσης 100 mM. Ακριβώς πριν την κρυστάλλωση, τα διαλύματα πρωτεΐνης και προσδέτη αναμίχθηκαν έτσι ώστε να προκύψει διάλυμα δεκαπλάσιας μοριακής συγκέντρωσης προσδέτη ως προς την πρωτεΐνη, το οποίο και επώαστηκε για 1-2 ώρες στους 4°C . Κρύσταλλοι σχηματίστηκαν όταν αναμίχθηκαν ίσοι όγκοι διαλύματος πρωτεΐνης-προσδέτη και μητρικού διαλύματος με τελική συγκέντρωση DMSO στη σταγόνα <4% κ.ο.

3.4.2 Πειράματα περίθλασης - Συλλογή κρυσταλλογραφικών δεδομένων

Οι κρύσταλλοι εμβαπτίστηκαν σε διάλυμα κρυσπροστασίας, που περιείχε πλέον του μητρικού διαλύματος 20% v/v αιθυλενική γλυκόλη, για 5-10 s πριν την έκθεσή τους σε υγρό άζωτο.

Οι μετρήσεις έγιναν στη γραμμή ακτινοβολήσης (beamline) X06DA του Swiss Light Source (Villigen, Ελβετίας), και τα δεδομένα περίθλασης συλλέχθηκαν σε διακριτικότητα 1,45–3,05 Å. Η επεξεργασία των εικόνων περίθλασης (**Εικόνα 3.31β**), (δεικτοδότηση, ολοκλήρωση και κανονικοποίηση) έγινε με το XDS (Kabsch 2010), στην ομάδα χώρου που υπέδειξε το POINTLESS (Evans 2006) για το σύνολο των ασυγχώνευτων δεδομένων (unmerged data).

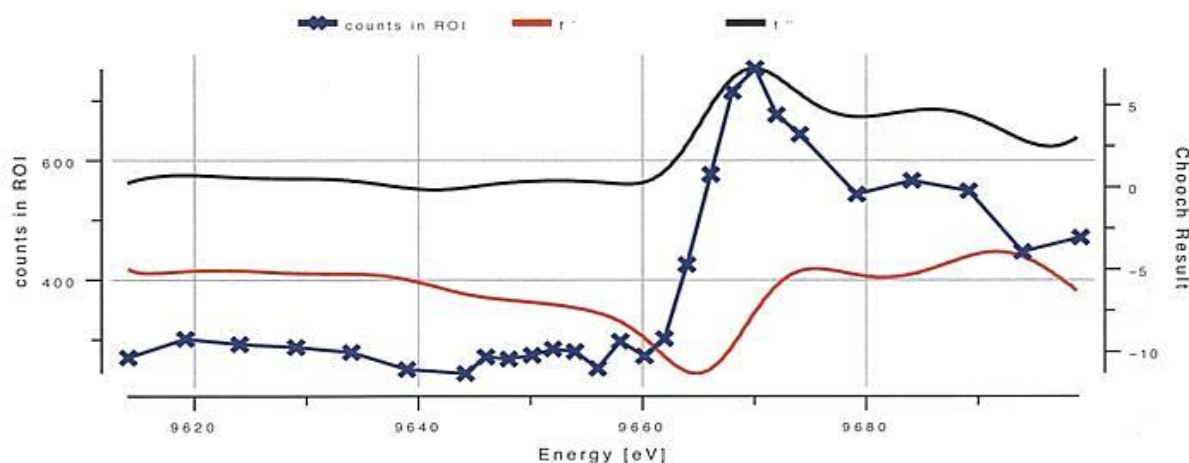


Εικόνα 3.31. α) κρύσταλλοι Bc1974 β) Περιθλασίγραμμα της Bc1974 γ) Bc1974 με τον προσδότη 1(1- Naphthohydroxamic acid) δ) Εικόνα περίθλασης ακτίνων-X από μονοκρύσταλλο της Bc1974 με τον προσδότη 1

3.4.3 Επίλυση δομής – Βελτιστοποίηση

Η δομή επιλύθηκε με τη μέθοδο της μοριακής αντικατάστασης (Rossmann 1972) χρησιμοποιώντας το λογισμικό PHASER (McCoy *et al.* 2007) με αρχικό μοντέλο τη δομή της ομόλογης απακετυλάσης του βακτηρίου *Streptococcus pneumoniae* SpPgdA (PDB ID: 2C1G, Blair *et al.* 2005). Από το αρχικό μοντέλο της SpPgdA αφαιρέθηκαν τα ετεροάτομα και οι πλευρικές αλυσίδες για να αποφευχθεί η εισαγωγή πιθανών σφαλμάτων από το προηγούμενο μοντέλο. Η επιθεώρηση του διαφορικού χάρτη Fo-Fc, μετά από τον αρχικό κύκλο βελτιστοποίησης άκαμπτου σώματος (rigid body refinement), αποκάλυψε τη θέση ενός βαρέως ατόμου εντός του ενεργού κέντρου της πρωτεΐνης, καθώς και τις διαμορφώσεις των περισσοτέρων πλευρικών αλυσίδων. Η βελτιστοποίηση της δομής πραγματοποιήθηκε αρχικώς με τη μέθοδο της μεγιστοποιημένης πιθανότητας μέσω του προγράμματος PHENIX (Adams *et al.* 2010) εφαρμόζοντας κύκλους βελτιστοποίησης διακοπτόμενους από τακτικές επιθεωρήσεις των χαρτών ηλεκτρονιακής πυκνότητας 2Fo-Fc και Fo-Fc. Στις περιοχές που κρίθηκε απαραίτητο ακολούθησε επανατοποθέτηση του μοριακού μοντέλου εντός της ηλεκτρονιακής πυκνότητας με το πρόγραμμα COOT (Emsley *et al.* 2010).

Τα στατιστικά βελτιστοποίησης για το τελικό μοντέλο δίδονται στο **Πίνακα 3.5**. Το είδος του βαρέως ατόμου προσδιορίστηκε με σάρωση φθορισμού ακτίνων-X στην περιοχή της Kedge του ψευδαργύρου (**Εικόνα 3.32**). Τα μόρια των προσδετών συμπεριλήφθηκαν στις περιπτώσεις που ήταν σαφώς διακριτά στους αρχικούς χάρτες Fo-Fc και 2Fo-Fc, και ο αριθμός κατάληψης τους προσδιορίστηκε με το πρόγραμμα PHENIX (Adams *et al.* 2010). Υπολογισμοί αφαιρετικών χαρτών (omit maps) ηλεκτρονιακής πυκνότητας χρησιμοποιήθηκαν στα τελικά στάδια βελτιστοποίησης προκειμένου να δειχθεί αμερόληπτα η παρουσία του προσδέτη. Οι τοπολογίες των προσδετών (μήκη δεσμών, γωνίες, διεδρες γωνίες μεταξύ των ατόμων τους) και οι συντεταγμένες τους δημιουργήθηκαν είτε με το πρόγραμμα PRODRG (Schuttelkopf & van Aalten 2004) είτε με το Jligand (Lebedev *et al.* 2012) και βελτιστοποιήθηκαν με το eLbow (Moriarty *et al.* 2009).



Εικόνα 3.32. Φάσματα φθορισμού ακτίνων-X περίξ της θεωρητικής Kedge του ψευδαργύρου στα 9659 eV. Η μπλε γραμμή δείχνει την ένταση του φθορισμού, ενώ η μαύρη και κόκκινη γραμμή δηλώνουν την κορυφή και σημείο καμπής της έντασης του φθορισμού, αντίστοιχα.

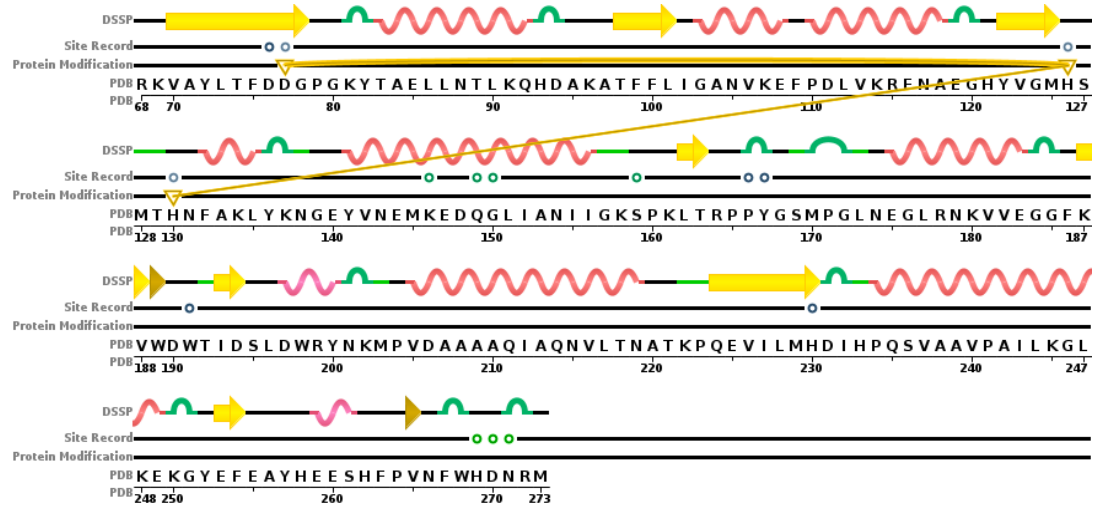
3.4.4 Περιγραφή δομής τριών διαστάσεων της Bc1974

Η ασύμμετρη μονάδα περιέχει τέσσερα μόρια της Bc1974, τα οποία διαφέρουν ελάχιστα (με r.m.s.d. 0.284 Å για όλα τα άτομα της κύριας αλυσίδας). Ολόκληρη η πρωτεΐνη Bc1974 έχει 273 αμινοξικά κατάλοιπα, ενώ το ανασυνδυασμένο πλασμίδιο δεν περιείχε τα πρώτα 25, τα οποία πιθανότατα ανήκουν σε περιοχή α-έλικας του αμινοτελικού άκρου, όπως έχει προσδιοριστεί με πρόβλεψη δευτεροταγούς δομής. Επιπλέον, το τελικό μοντέλο δεν περιλαμβάνει τα πρώτα 42 αμινοξέα του εκπεφρασμένου αμινοτελικού άκρου της πρωτεΐνης, δηλαδή οι κρυσταλλικές δομές αποκαλύπτουν τη δομή της περιοχής μεταξύ των αμινοξέων 68-273 (**Εικόνα 3.33 α**). Τα δευτεροταγή δομικά χαρακτηριστικά της δομής και το διάγραμμα τοπολογίας δίδονται στις **εικόνες 3.33** και **3.34**.

Το διάγραμμα Ramachadran υπολογισμένο από το πρόγραμμα PROCHECK (Laskowski *et al.* 1993) περιέχει το 97% των καταλοίπων στις ευνοούμενες και επιτρεπόμενες περιοχές, ενώ 2 αμινοξέα λαμβάνουν σχεδόν μη επιτρεπτή διαμόρφωση (0,6%) (**Εικόνα 3.35**).

MEKALKIKQIVVVLIAIAAVAIGYYMFQSITSPAKAVAKQENNVQLASEQPKVEMNKTPSRFNGKERKV
 AYLTFDDGPGKYTAELLNLTQHDATFFLLIGANVKEFPDLVKRENAEGHYVGMHSMTHNFAKLYKNGE
 YVNEMKEDQGLIANIIGKSPKLRPPYGSMPGLNEGLRNKVVVEGGFKVWDWTIDSLDWRYNKMPVDAAAA
 QIAQNVLTNATKPQEVILMHDIHPOQSVAAVPAILKGLKEKGYEFEAYHEESHFPVNFWDNRM

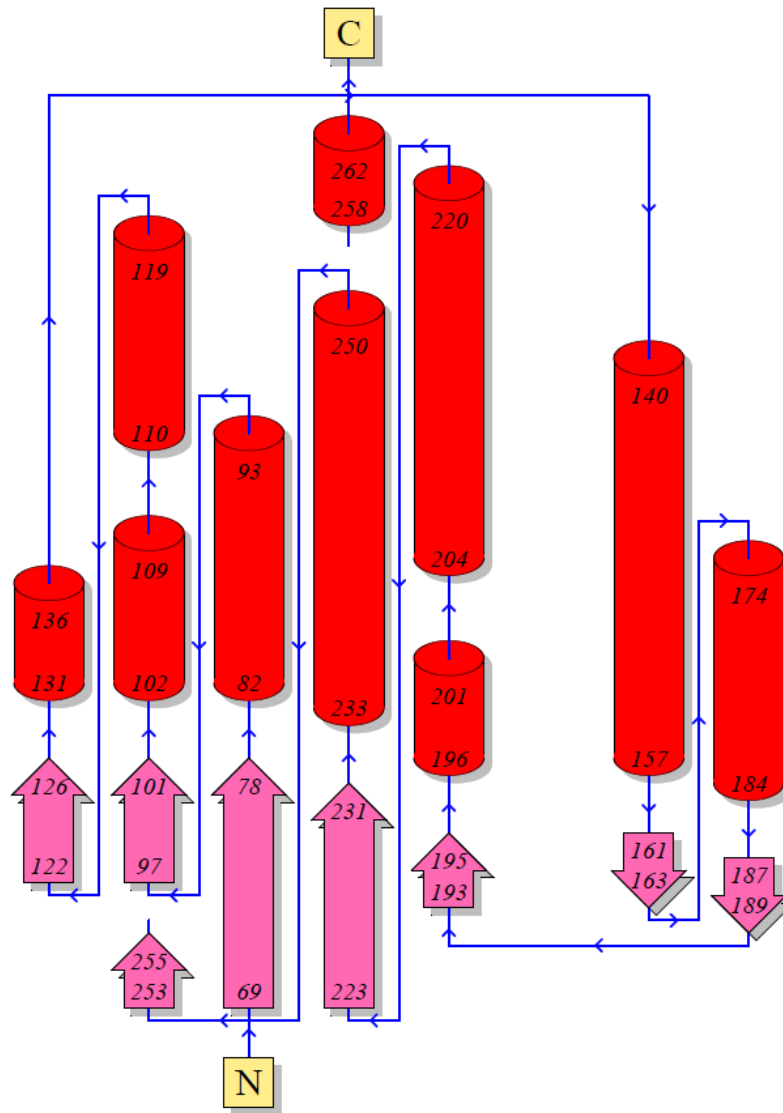
(α)



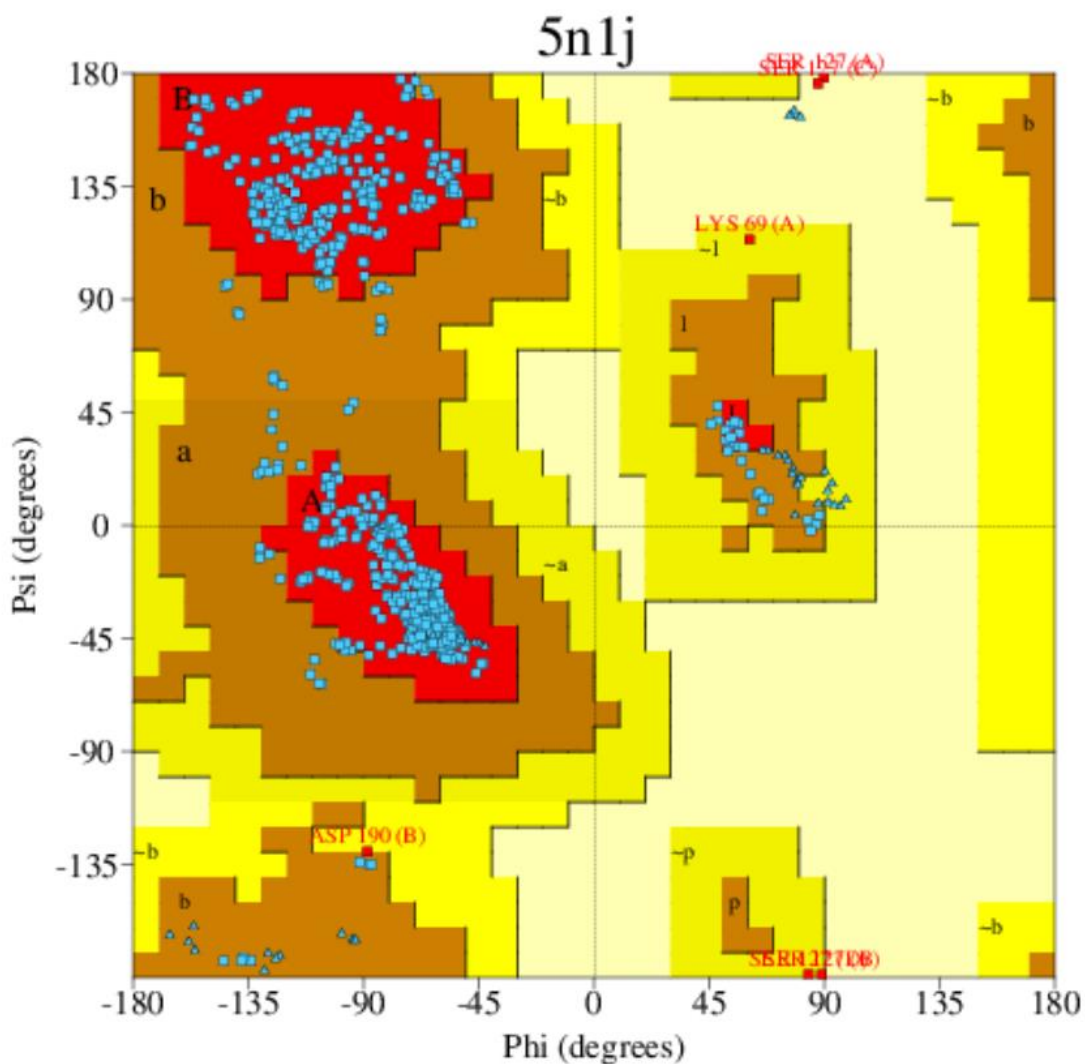
(β)

- τόν ψευδαργύρου
- β-γέφυρα
- στροφή
- βρόχος
- β-κλώνος
- 3/10 έλικα
- α-έλικα

Εικόνα 3.33. (α) Δίδεται η αλληλουχία της Bc1974. Με γκρι φόντο παρουσιάζονται τα αμινοξέα της κρυσταλλικής δομής. (β) Γραμμική απεικόνιση των δομικών χαρακτηριστικών της δευτεροταγούς δομής στην αμινοξική ακολουθία (PDB ID: 5N1J).



Εικόνα 3.34. Συγκρότηση της δευτεροταγούς δομής της Bc1974.



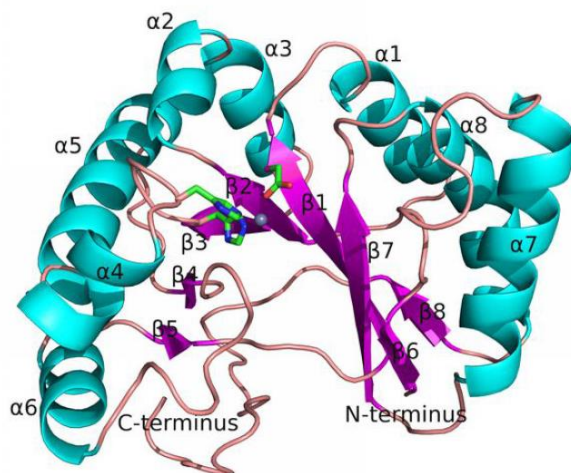
Εικόνα 3.35. Το διάγραμμα Ramachandran της δομής της Bc1974. (ενοούμενες περιοχές - A, B, L, επιπρόσθετες επιτρεπόμενες περιοχές - a, b, l, p, γενικά επιτρεπόμενες περιοχές ~a,~b,~l,~p οι απαγορευμένες περιοχές χρωματίζονται υποκίτρινες).

Η δομή της Bc1974 αποκαλύπτει μια μόνο δομική ενότητα (αμινοξέα 68–273), η οποία υιοθετεί την αναδίπλωση μιας $(\alpha/\beta)_8$ τοπολογίας (**Εικόνα 3.36**), χαρακτηριστική της NodB ομόλογης δομικής ενότητας που βρίσκεται σε όλες τις CE4 εστεράσες (Giastas *et al.* 2018). Οι ατομικές συντεταγμένες και οι παράγοντες δομής της κρυσταλλικής δομής της πρωτεΐνης χωρίς υπόστρωμα Bc1974 έχουν καταχωρηθεί στην δημόσια βάση δεδομένων Protein Data Bank (PDB, <http://wwpdb.org/>, Berman *et al.* 2000) υπό τον κωδικό αριθμό καταχώρησης 5N1J. Ο χάρτης ηλεκτρονιακής πυκνότητας έδειξε την αδιαμφισβήτητη

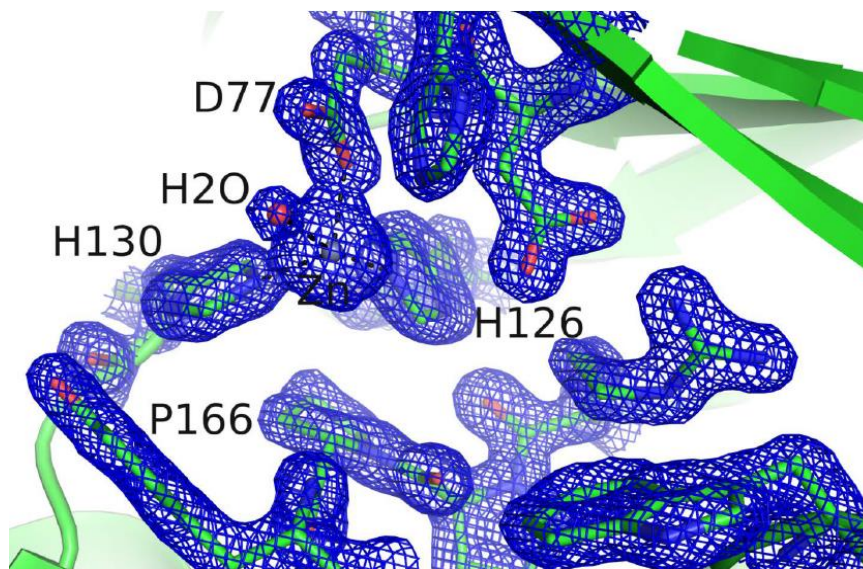
παρουσία ενός μεταλλικού ιόντος στον πυθμένα ενός καλά σχηματισμένου αύλακα [20-σ στον χάρτη διαφορικής ηλεκτρονιακής πυκνότητας, με θερμικό παράγοντα (B-factor) 25 Å²]. Το ιόν αυτό σταθεροποιείται αλληλεπιδρώντας με την υψηλά συντηρημένη τριάδα αμινοξικών καταλοίπων Asp-His-His (**Εικόνα 3.37**). Από την επίλυση της δομής βρέθηκαν τέσσερα μόρια στην ασύμμετρη μονάδα, όπως ήδη αναφέρθηκε ανωτέρω. Η βελτιστοποίηση της δομής έδειξε ότι η μέγιστη απόκλιση μεταξύ των τεσσάρων μορίων της ασύμμετρης μονάδας παρατηρείται στην περιοχή των αμινοξέων 190–205, τα οποία απαρτίζουν το μοτίβο 4 (MT4) και τη μετά-MT4 περιοχή. Σημειώνεται ότι ο βρόχος που σχηματίζουν τα αμινοξέα του MT4 φαίνεται να λαμβάνει δύο διακριτές στερεοδιατάξεις, μία όπως εμφανίζεται στα μόρια A και C και μία δεύτερη όπως εμφανίζεται στα μόρια B και D. Στην **εικόνα 3.38** παριστάνεται ο θερμικός παράγοντας ανά αμινοξύ για την κύρια αλυσίδα στα μόρια A και B όπου γίνεται φανερό ότι τα αμινοξέα είναι καλά εντοπισμένα. Η περιοχή MT4 είναι εκείνη που εμφανίζει τις υψηλότερες τιμές θερμικών παραγόντων, υποδηλώνοντας υψηλή κινητικότητα ιδιαίτερα στα δύο μόρια A και C.

Η οκταεδρική γεωμετρία συναρμογής του καταλυτικού ψευδαργύρου συμπληρώνεται από ένα μόριο ύδατος και ένα ιοντισμένο μόριο οξικού οξέος, που πιθανώς υπάρχει ως υπολειμματική ουσία του διαλύματος της πολυαιθυλενογλυκόλης που χρησιμοποιείται στο ρυθμιστικό διάλυμα της κρυστάλλωσης (Blair *et al.* 2005).

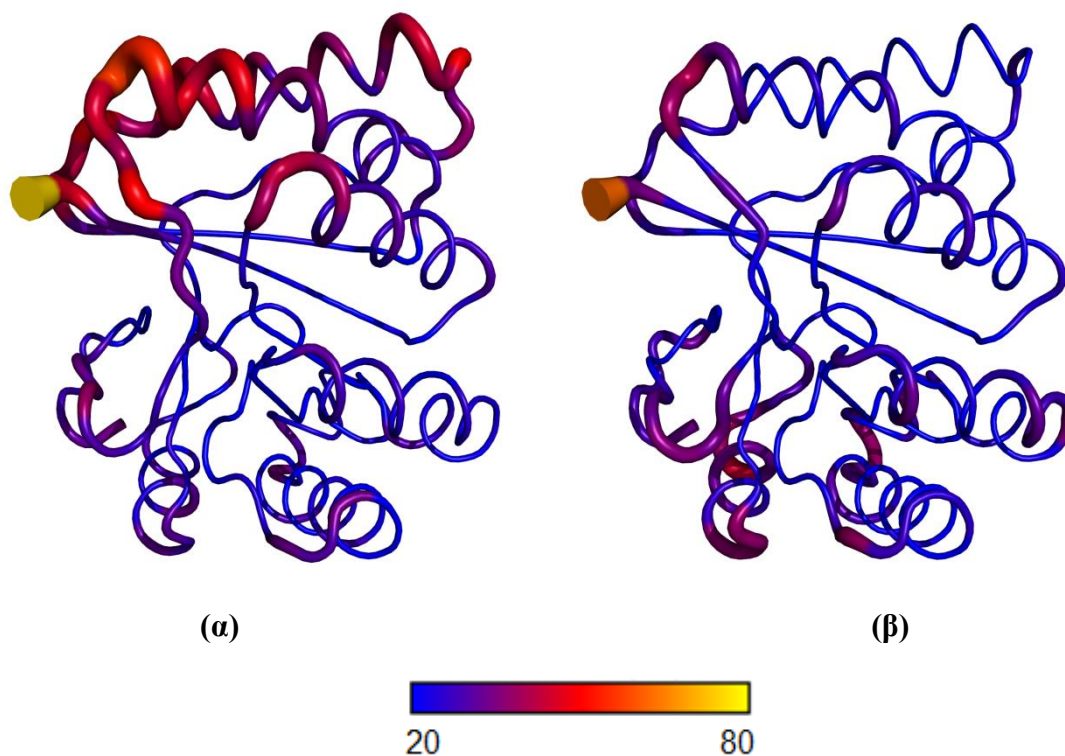
Το οξικό ανιόν βρέθηκε καλά εντοπισμένο (εμφανές σε χάρτη Fo-Fc σχεδιασμένο με κατώφλι άνω του 8-σ), με τιμή του θερμικού παράγοντα (B-factor) 29 Å² που είναι κοντά σε εκείνη του μεταλλικού ιόντος. Το οξικό ανιόν αλληλεπιδρά με αμινοξέα του ενεργού κέντρου με παρόμοιο τρόπο σε όλες τις αδέσμευτες δομές Bc1974 (**Εικόνα 3.39**). Ένα από τα άτομα οξυγόνου του αλληλεπιδρά με το αμινοξικό κατάλοιπο Asp76 που σχηματίζει γέφυρα άλατος με την Arg164, ένα καλά συντηρημένο αμινοξύ σε όλες τις απακετυλάσες εκτός από εκείνες που περιέχουν τις δομικές ενότητες Fn3 (Blair *et al.* 2005, Arnaouteli *et al.* 2015, Andreou *et al.* 2019). Το ίδιο οξυγόνο αλληλεπιδρά επίσης με τη δομικά συντηρημένη His230, που με τη σειρά της αλληλεπιδρά απευθείας με τη γειτονική Asp197 της περιοχής MT4. Η μεθυλομάδα του οξικού ανιόντος καταλαμβάνει μια μικρή υδρόφοβη περιοχή που σχηματίζεται από την ομάδα του ινδολίου της Trp191 και την πλευρική αλυσίδα της Leu228. Σε όλες τις δομές της Bc1974 που επιλύθηκαν, δεν υπήρχε ηλεκτρονιακή πυκνότητα γύρω από τον Ca της συντηρημένης Pro166 που να υποδεικνύει την ύπαρξη υδροξυλομάδας (**Εικόνα 3.37**), ένα χαρακτηριστικό που έχει παρατηρηθεί πρωτίτερα σε άλλες φερόμενες απακετυλάσες, όπως η Bd3279 (Lambert *et al.* 2016), Ba0330 (Arnaouteli *et al.* 2015) και Bc0361 (Fadoulglou *et al.* 2013).



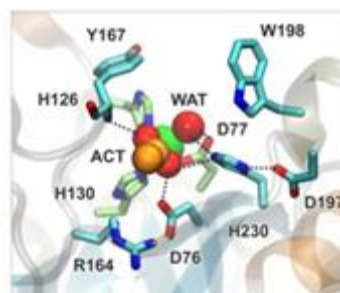
Εικόνα 3.36. Η δομή της Bc1974. Οι έλικες αντιπροσωπεύονται με κορδέλες χρώματος κυανού, οι β-κλώνοι με μωβ βέλη και οι βρόχοι με ροζ χορδές. Παρουσιάζεται η δομική ενότητα NodB. Το ιόν ψευδαργύρου στο ενεργό κέντρο απεικονίζεται με σφαίρα χρώματος γκρι. Παρουσιάζεται η συντηρημένη τριάδα των αμινοξικών καταλοίπων συναρμογής του μεταλλοϊόντος που αποτελείται από ένα ασπαρτικό οξύ (Asp77) και δύο κατάλοιπα ιστιδίνης (His126 και His130) τα οποία απεικονίζονται με ράβδους.



Εικόνα 3.37. Ηλεκτρονιακή πυκνότητα (χάρτης 2Fo-Fc) στην περιοχή του ψευδαργύρου σε 2σ , υπολογισμένος με τις τελικές συντεταγμένες του μοντέλου. Με χρώμα πράσινο παριστάνονται τα άτομα C, με κόκκινο τα άτομα O, με κυανό τα άτομα N. Το ιόν ψευδαργύρου απεικονίζεται με μια γκρι σφαίρα και η κόκκινη σφαίρα δηλώνει ένα μόριο ύδατος. Η χαρακτηριστική καταλυτική αμινοξική τριάδα (His, His, και Asp) παρουσιάζεται στην Bc1974 στη θέση δέσμευσης του ψευδαργύρου στα πρότυπα των PDAs. Φαίνεται η απουσία ηλεκτρονιακής πυκνότητας γύρω από τον Ca της συντηρημένης Pro166. Η εικόνα σχεδιάστηκε με το πρόγραμμα PYMOL.



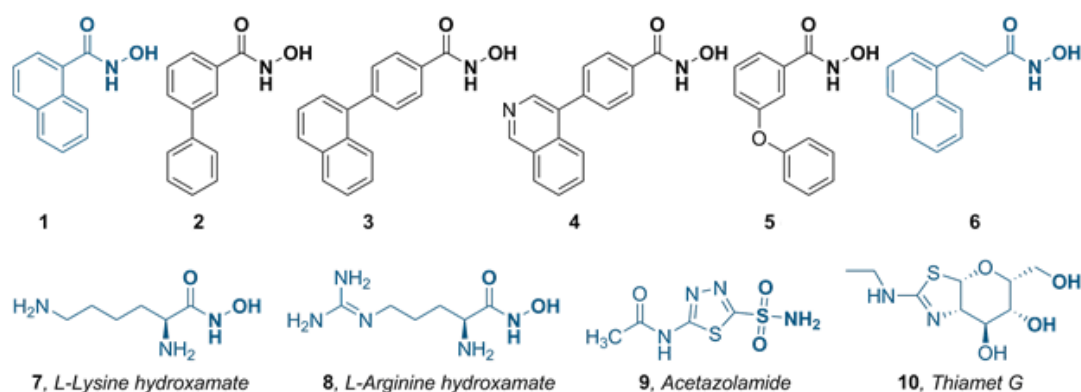
Εικόνα 3.38. Χρωματική απεικόνιση της δομής σύμφωνα με τους θερμικούς παράγοντες (B-factors \AA^2) για τα άτομα της κύριας αλυσίδας. Παρουσιάζονται οι αλυσίδες των μορίων A (**α**) και B (**β**). Στην περιοχή του MT4 παρατηρείται η μεγαλύτερη θερμική κίνηση.



Εικόνα 3.39 Το ενεργό κέντρο της Bc1974. Τα κατάλοιπα σύνδεσης του μετάλλου είναι χρωματισμένα με πράσινα άτομα C και τα αλληλεπιδρώντα καταλυτικά κατάλοιπα με γαλάζια άτομα C, όλα τα άλλα άτομα είναι χρωματισμένα με μπλε για το N και κόκκινο για το O. Τα μόρια οξικού (ACT) και ύδατος (WAT) παρουσιάζονται με σφαίρες, ενώ οι διακεκομμένες γραμμές δείχνουν αλληλεπιδράσεις δεσμού υδρογόνου.

3.4.5 Περιγραφή δομής τριών διαστάσεων συμπλόκων της Bc1974

Προκειμένου να δημιουργηθεί μια βάση πιθανών αρχικών δραστικών ενώσεων με απώτερο σκοπό τη μελέτη και εύρεση αναστολέων για τις Bc1974 και Ba1977 πραγματοποιήθηκε προκαταρτική σάρωση είκοσι πέντε μικρών μορίων με βάση τη στερεοδομή, με τις περισσότερες από αυτές να περιέχουν τμήμα υδροξαμικού οξέος ως ομάδα πρόσδεσης στον ψευδάργυρο (Πίνακας 1, Παράρτημα Α'). Οι επιλεγμένες ουσίες διαφέρουν σε μέγεθος (MB=120–558), υδροφοβικότητα (συντελεστής κατανομής οκτανόλης/ύδατος, cLogP: –4,4 έως 6,1), και αποτελούνται από ένα ποικίλο σύνολο αρυλ- και αμινοξικών υποκαταστατών υδροξαμικών οξέων (ενώσεις **1–8** στην **Εικόνα 3.40** και **11–15** στον πίνακα 1 παρ.Α'), μερικούς γνωστούς αναστολείς μεταλοπρωτεϊνών (MMP) (**18, 19, 21–23** στον πίνακα 1 παρ Α'), και αντιβιοτικούς ή αντιβακτηριακούς παράγοντες (**20, 24** στον πίνακα 1 παρ.Α'). Πιο συγκεκριμένα, τα υδροξαμικά οξέα **1–6** (Εικόνα 3.40, πίνακας 1 παρ. Α), **15** και Vorinostat (**16** ή SAHA στον πίνακα 1 παρ.Α') είναι γνωστοί αναστολείς απακετυλασών της ιστόνης (HDAC) (Krennhubec *et al.* 2007, Oehme *et al.* 2009), η ακεταζολαμίδη (**9**, Εικόνα 3.40), είναι ένας αναστολέας καρβονικής ανυδράσης με βάση το σουλφαμίδιο, Thiamet-G (**10**, Εικόνα 3.40) είναι ένας αναστολέας της ανθρώπινης O-GlcNAcase (Yuzwa *et al.* 2008), και PUGNAc (**25**, πίνακας 1 παρ.Α') είναι ένας αναστολέας της β-N-ακετυλογλυκοζαμινιδάσης (Horsch *et al.* 1991).



Εικόνα 3.40 Δομές αναστολέων που περιλαμβάνουν ομάδες υδροξαμικού οξέως **1-6** και οι προσδέτες που συγκρυσταλλώθηκαν με την Bc1974 **1, 6–10** (με μπλε).

Δοκιμές συγκρυστάλλωσης όλων των είκοσι πέντε ενώσεων έδωσαν κρυστάλλους της Bc1974 σε σύμπλοκο με έξι προσδέτες: Bc1974 με αναστολέα **1**, αναστολέα HDAC8 (**6**), δυο αμινοξικά υδροξαμικά οξέα (**7**, **8**) και δυο μη υδροξαμικά οξέα που είναι αναστολείς μεταλοενζύμων **9** και **10**. Έγινε συλλογή κρυσταλλογραφικών δεδομένων ακολουθώντας την ίδια διαδικασία που αναφέρεται για την πρωτεΐνη άνευ προσδέτη (βλ. §3.4.2). Μετά την επίλυση της κρυσταλλικής δομής των συμπλόκων και τη βελτιστοποίηση των δομών οι ατομικές συντεταγμένες και οι παράγοντες δομής κατατέθηκαν στην βάση δεδομένων πρωτεϊνικών δομών (Protein Data Bank, <http://www.rcsb.org/pdb>, Berman *et al.* 2000) και καταχωρήθηκαν υπό τους εξής κωδικούς καταχώρησης: 5N1P, 5NC6, 5NC9, 5NCD, 5NEK και 5NEL (**Πίνακας 3.5**).

Σε όλες τις περιπτώσεις, οι προσδέτες αρχικά εντοπίστηκαν στην ηλεκτρονιακή πυκνότητα του ενεργού κέντρου (Fo-Fc χάρτης) και ο βαθμός κατάληψης της θέσης τους βρέθηκε μετά από διαδοχικούς κύκλους βελτιστοποίησης της δομής. Η παρουσία πιστοποιήθηκε από αφαιρετικούς χάρτες ηλεκτρονιακής πυκνότητας, οι οποίοι έδειξαν ιδανικό ταίριασμα μεταξύ της πυκνότητας και των συντεταγμένων του τελικού μοντέλου (**Εικόνα 3.41 στήλη α**). Στις παρούσες δομές (**Πίνακας 3.5**), όπου το ιόν ψευδαργύρου δε βρίσκεται σε συναρμογή από κάποιο μόριο προσδέτη, ένα οξικό ανιόν διαφαίνεται ξεκάθαρα στους χάρτες ηλεκτρονιακής πυκνότητας. Όλα τα σύμπλοκα κρυστάλλωσαν στην ομάδα χώρου *P21* εκτός από το σύμπλοκο Bc1974–**1** που κρυστάλλωσε σε *C2*, και για το οποίο ελήφθησαν δεδομένα περίθλασης στην υψηλότερη διακριτικότητα (1,45 Å). Η πυκνότητα του προσδέτη φαίνεται καθαρά και στα 4 μόρια της ασύμμετρης μονάδας για την ένωση **1**, όπως και για την ένωση **7** (διακριτικότητα 2,44 Å) (**Πίνακας 3.5**). Αντίθετα, οι προσδέτες **6**, **9** και **10** (διακριτικότητα 2,80, 3,05 και 2,73 Å, αντίστοιχα) εντοπίστηκαν μόνο στα μόρια B και D, ενώ ο προσδέτης **8** (διακριτικότητα 2,45 Å) εντοπίστηκε μόνο στο μόριο B. Ενδιαφέρον παρουσιάζει το γεγονός πως ανεξάρτητα της χημική φύσης των έξι προσδετών στη συγκρυστάλλωση, η παρουσία τους μέσα στο ενεργό κέντρο δεν επηρεάζει την κύρια διαμόρφωση της Bc1974. Έτσι, η r.m.s.d. που παρατηρείται για τα άτομα της κύριας αλυσίδας στο σύμπλοκο και στην Bc1974 άνευ προσδέτη είναι 0,24–0,28 Å για το μόριο A, 0,25–0,27 Å για το μόριο B, 0,30–0,38 Å για το μόριο C, και 0,19–0,27 Å για το μόριο D.

Σε όλες τις δομές των συμπλόκων της Bc1974 η συναρμογή των ενώσεων με τον ψευδάργυρο γίνεται με χηλικό τρόπο (**Εικόνα 3.41 στήλη β**). Το υδροξαμικό τμήμα των **1** και **6–8** παρουσιάζει ένα κοινό μοτίβο αλληλεπίδρασης μέσω δεσμών- H με: α) η C=O ομάδα και η NH της Tyr167 στην κύρια αλυσίδα β) η NH και το ιμιδαζόλιο της His230, και γ) η OH ομάδα και η καρβοξυλομάδα της Asp76. Αξιοπρόσεκτο είναι το γεγονός πως οι 4 προσδέτες

υδροξαμικού οξέος είναι τοποθετημένοι σε ένα μη κανονικό οκτάεδρο που συμπληρώνεται με ένα μόριο διαλύτη (**Εικόνα 3.41 I, II, V, VI**). Αντίθετα, ο καταλυτικός ψευδάργυρος στο σύμπλοκο Bc1974–9 και Bc1974–10 σχηματίζει πεντάεδρο. Η ακεταζολαμίδα (**9**) δεσμεύεται στον ψευδάργυρο μέσω του N της αμίνης και ενός από τα άτομα οξυγόνου (O1) του σουλφαμιδίου (**Εικόνα 3.41 IV**). Το τμήμα του σουλφαμιδίου είναι σε ιδανική απόσταση από τις πλευρικές αλυσίδες της His230(S=O1), Asp76 (S=O2) για σχηματισμό δεσμού H, και του αμιδίου της κύριας αλυσίδας της Tyr167 (S–NH₂). Το Thiamet-G (**10**), είναι ένας πιθανός αναστολέας της ανθρώπινης O-GlcNAcase (Yuzwa *et al.* 2008), που βρίσκεται σε συναρμογή με τον ψευδάργυρο μέσω των O4 και O6 καρβοξυλομάδων του τμήματός του που προσομοιάζει σε σάκχαρο, σε μια διάταξη τριγωνικής διπυραμίδας (**Εικόνα 3.41 III**). Το υδροξύλιο O6 κάνει δεσμό H με το δακτύλιο του ιμιδαζολίου της His230 και του καρβοξυλίου της Asp76.

Ενώ οι ομάδες πρόσδεσης ψευδαργύρου των προσδετών εμφανίζουν πολικές αλληλεπιδράσεις εντός του ενεργού κέντρου της Bc1974, το υπόλοιπο τμήμα των προσδετών προσαρμόζεται μέσα σε μια κοιλότητα που αποτελείται από τη μη πολική Leu228 και τα αρωματικά αμινοξέα Tyr167, Trp191 και Trp198 (**Εικόνα 3**). Ως αποτέλεσμα, οι δακτύλιοι ναφθαλενίου των **1** και **6** εμφανίζουν αλληλεπιδράσεις π-π με αυτά τα τρία αρωματικά αμινοξέα, είτε με παράλληλο προσανατολισμό είτε με προσανατολισμό σχήματος T. Οι θετικά φορτισμένες των **7** και **8** εμφανίζουν αλληλεπιδράσεις κατιόντος-π με την Trp198 και τα αλειφατικά τους τμήματα εμπλέκονται σε επαφές van der Waals. Παρομοίως, οι ενώσεις **9** και **10** που περιλαμβάνουν ακόμη περισσότερες πολικές ομάδες εμπλέκονται μόνο σε αρωματικές αλληλεπιδράσεις και αλληλεπιδράσεις van der Waals (**Εικόνα 3.41 III, IV**). Αν και ο ακριβής προσανατολισμός των **6**, **8** και **9** ποικίλλει μεταξύ των μορίων της Bc1974 στην ασύμμετρη μονάδα, η μη δεσμευμένη ομάδα ψευδαργύρου των έξι υποκαταστατών παρουσιάζει παρόμοιες αλληλεπιδράσεις με τα αμινοξέα του ενεργού κέντρου και τα κατάλοιπα της θέσης δέσμευσης Leu228, Tyr167, Trp191 και Trp198 (**Εικόνα 3.41 II,IV,V στήλη γ**). Είναι αξιοσημείωτο, ότι οι μικρού μοριακού βάρους προσδέτες βρέθηκαν δεσμευμένοι στο ένζυμο Bc1974 κυρίως στις αλυσίδες B και D, όπου η MT4 υιοθετεί την κλειστή διαμόρφωσή της.

Στην ηλεκτρονιακή πυκνότητα για τον προσδέτη στο σύμπλοκο Bc1974–1 στα μονομερή A και C, σε αντίθεση με τα περισσότερα άλλα σύμπλοκα, που το οξικό αναγνωρίστηκε, παρατηρήθηκε ένα τριπλό αρωματικό π-επιστοίβαγμα αποτελούμενο από το ναφθαλένιο του **1**, την Trp198 του βρόχου MT4 και την His269 ενός γειτονικού μορίου της Bc1974 (B) (**Εικόνα 3.41 I στήλη γ**). Το ιμιδαζόλιο της His269 παρεμβαίνει με τις άλλες δύο

αρωματικές ομάδες ακινητοποιώντας το ναφθαλένιο στην καλά καθορισμένη του θέση και πιθανώς συνέβαλε στην μετατόπιση της MT4 στην μακρινή της θέση. Αυτό το τριπλό αρωματικό σύμπλεγμα βρέθηκε με την πιο ευνοϊκή, όσον αφορά την ενέργεια, παράλληλη εκτοπισμένη διάταξη. Μεταξύ των αλληλεπιδράσεων της ομάδας ναφθαλενίου με αμινοξέα της ίδιας αλυσίδας ήταν τα αρωματικά τύπου T με την Trp191 και την Tyr167, τα οποία υπάρχουν σε όλες τις αλυσίδες της ασύμμετρης μονάδας (**Εικόνα 3.41 στήλη β**). Εν τούτοις, στις αλυσίδες B και D όπου ο βρόχος MT4 εμφανίζει μια πιο συμπαγή διαμόρφωση, η Trp198 συνεισφέρει άμεσα στις αλληλεπιδράσεις πρωτεΐνης-προσδέτη, πλησιάζοντας το **1** στα 3,9 Å, με γωνία μεταξύ των δύο αρωματικών δακτυλίων ~ 40° μοιρών.

Το επιπρόσθετο τμήμα αιθυλενίου του **6** προσδίδει αυξημένη ευκαμψία στην ομάδα ναφθαλενίου, η οποία, σε αντίθεση με το **1**, προκαλεί δύο διαφορετικά σχήματα αλληλεπίδρασης εντός των δύο μονομερών του συμπλόκου Bc1974-6. Στο μονομερές B, η ομάδα ναφθαλενίου υιοθετεί μία ημι-παράλληλη διαμόρφωση με το τμήμα ινδολίου της Trp198 (γωνία 20° μοιρών) και σταθεροποιείται μέσω αλληλεπιδράσεων π-π επιστοιβάξης. Στην αλυσίδα D, το τμήμα του ναφθαλενίου υιοθετεί έναν κάθετο προσανατολισμό σε σχέση με τα αμινοξέα Trp191 και Trp198 μέσω αρωματικών αλληλεπιδράσεων σχήματος T και πιθανώς μέσω κατιονικών-π αλληλεπιδράσεων με άλλο αμινοξύ της MT4, Arg199 (**Εικόνα 3.41 I, II στήλη γ**)

Ο υδροξαμικός εστέρας L-λυσίνης (**7**) δεσμεύθηκε στον ψευδάργυρο με σχεδόν ταυτόσημο τρόπο μέσα σε όλα τα μονομερή της ασύμμετρης μονάδας, ενώ η πλευρική αλυσίδα ακολουθεί διαφορετικές τροχιές. Στις αλυσίδες B και D, όπου η θέση πρόσδεσης είναι ελεύθερη δομικών περιορισμών, η πυκνότητα ηλεκτρονίων του προσδέτη καλύπτει σαφώς το αλειφατικό τμήμα, σταθεροποιούμενη μέσω αλληλεπιδράσεων van der Waals με την Tyr167, ενώ το εύκαμπτο τμήμα, λόγω έλλειψης ισχυρών αλληλεπιδράσεων της πρωτοταγής αμίνης με τα γειτονικά αμινοξέα, εντοπίζεται οριακά στο χάρτη 2Fo-Fc (επίπεδο sigma 1,0), δείχνοντας προς την Trp198 (**Εικόνα 3.41 VI στήλη γ**). Εντούτοις, στις αλυσίδες A και C ο προσδέτης προσδιορίστηκε σαφώς σε όλη του την έκταση και η πρωτοταγής αμίνη του προσανατολίστηκε προς την MT5, τοποθετημένη σε σάντουιτς με την His230 της ίδιας αλυσίδας και την His269 μιας γειτονικής πολυπεπτιδικής αλυσίδας.

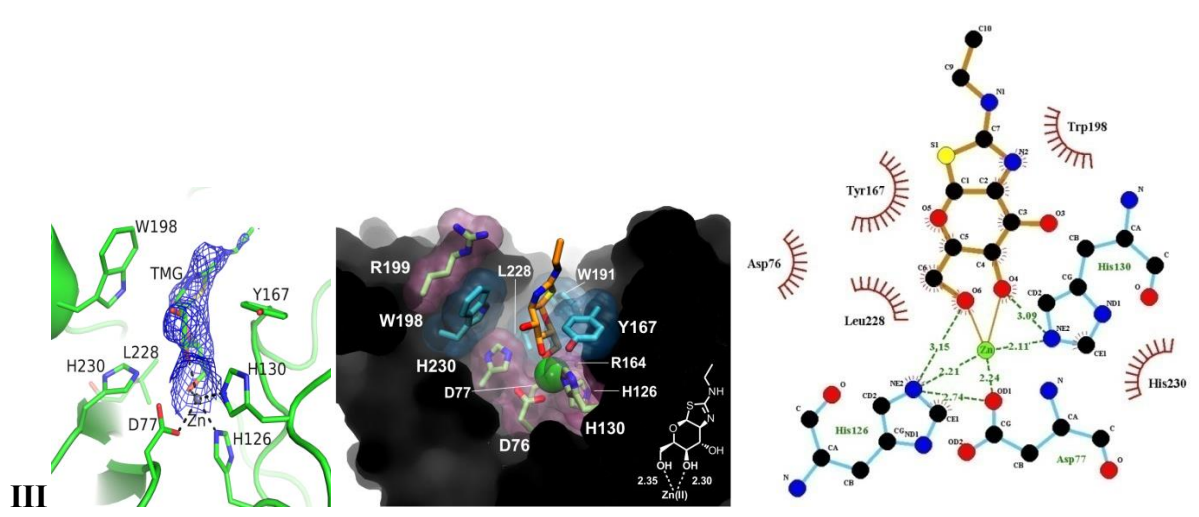
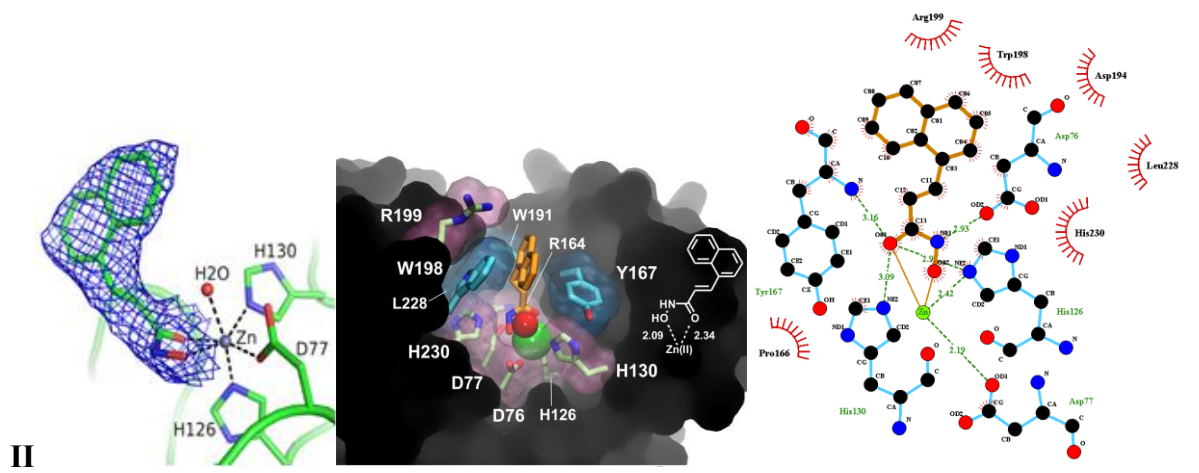
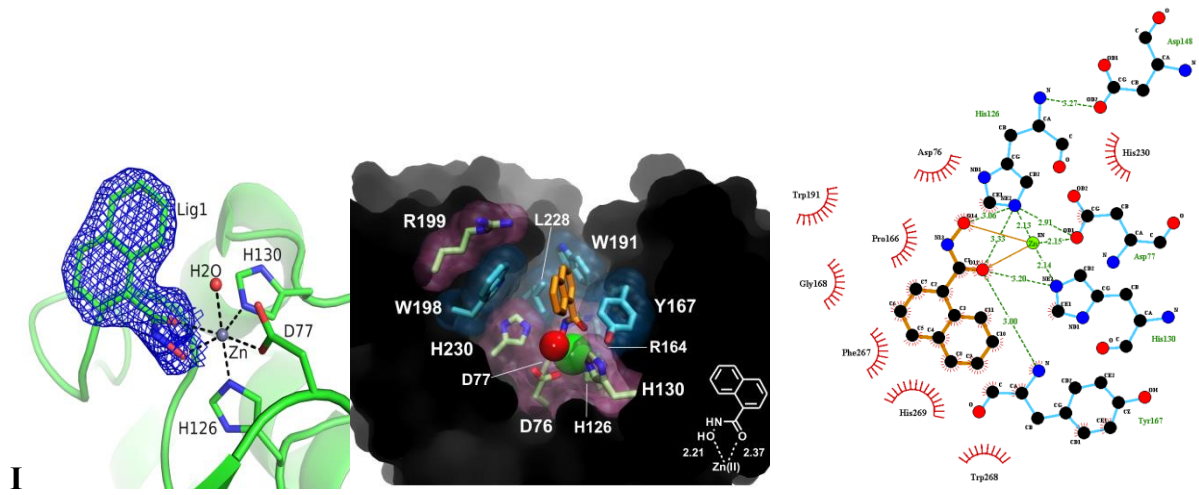
Η υδροξαμική L-αργινίνη (**8**) συνδέθηκε σαφώς με την αλυσίδα B, ενώ στην αλυσίδα D χτίστηκε ένα μόνο μόριο οξικού ανιόντος λόγω έλλειψης ισχυρής πυκνότητας ηλεκτρονίων που καλύπτει τη λειτουργική ομάδα γουανιδινίου. Η πλευρική αλυσίδα του υδροξαμικού αμινοξέος **8** υιοθετεί μια αποδεκτή ομάδα διεδρων γωνιών φέρνοντας την ομάδα γουανιδινίου

μεταξύ των αρωματικών ομάδων της Tyr167 και της Trp198. Αν και η τροχιά της πλευρικής αλυσίδας της διαφέρει εντελώς από εκείνη που εμφανίζεται στο σύμπλοκο με τον προσδέτη **7**, η ομάδα γουανιδινίου καταλαμβάνει χωρικά την ίδια περιοχή με τις πρωτοταγείς αμίνες του **7** στις αλυσίδες B και D και έτσι παρουσιάζει κατιονικές -π αλληλεπιδράσεις με το τμήμα ινδολίου της Trp198 (**Εικόνα 3.41V**).

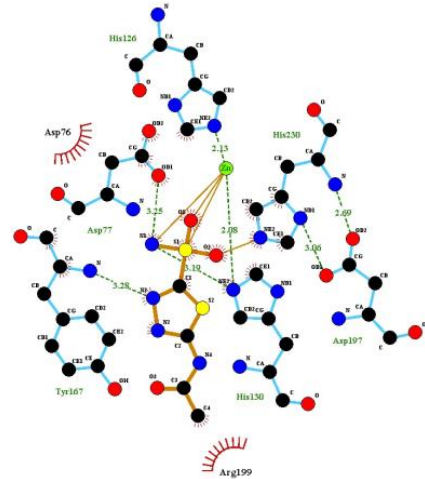
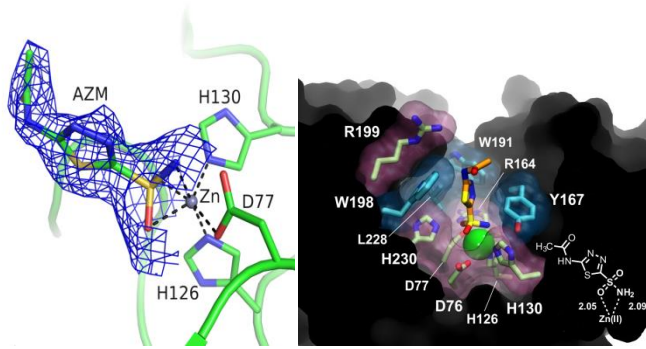
Πίνακας 3.5 Συλλογή δεδομένων, στατιστικά και βελτιστοποίηση.

Συλλογή δεδομένων	Bc1974	Σύμπλοκο Bc1974 -1	Σύμπλοκο Bc1974 - 6	Σύμπλοκο Bc1974 -7	Σύμπλοκο Bc1974 -8	Σύμπλοκο Bc1974 - 9	Σύμπλοκο Bc1974 -10
Ομάδα χώρου	$P2_1$	$C2$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
<i>Διαστάσεις μοναδιαίας κυψελίδας</i>							
a, b, c (Å)	50,28, 117,67, 99,77	196,10, 44,35, 99,13	49,36 118,01 98,62	49,77 117,83 99,06	50,65 117,44 99,49	49,39 118,21 98,27	49,44 117,97 98,20
α, β, γ (°)	90,00, 102,75, 90,00	90,00, 98,83, 90,00	90,00 102,28 90,00	90,00 102,38 90,00	90,00 102,92 90,00	90,00 102,09 90,00	90,00 102,10 90,00
Μήκος κύματος (Å)	1,000	1,000	1,000	1,000	1,000	1,000	1,000
Διακριτικότητα (Å)	48,65- 1,80 (1,87 -1,80)	48,4 - 1,45 (1,5 – 1,45)	47,35- 2,80 (2,92- 2,80)	48,61-2,44 (2,53-2,44)	48,48 -2,45 (2,53-2,45)	48,29- 3,05 (3,17-3,05)	48,34- 2,73 (2,83-2,73)
Μόρια πρωτεΐνης / AM	4	4	4	4	4	4	4
Μοναδικές ανακλάσεις	104462 (10419)	149678 (14328)	27111 (2632)	41046 (3720)	41648 (4057)	20589 (1717)	28074 (2256)
R_{meas}	0,082 (0,890)	0,069 (0,910)	0,093 (0,923)	0,097 (0,735)	0,117 (0,750)	0,074 (0,603)	0,103 (0,963)
$\langle I \rangle / \langle \sigma I \rangle$	13,1 (2,1)	15,2 (1,4)	6,6 (1,9)	9,0 (1,9)	7,2 (1,9)	11,1 (1,8)	9,6 (1,9)
Πληρότητα (%)	99,6 (99,7)	99,5 (96,0)	99,5 (95,7)	98,8 (90,8)	99,5 (97,3)	98,3 (83,5)	96,3 (85,4)
Πολλαπλότητα	6,7 (6,5)	5,1 (4,9)	3,9 (3,8)	4,5 (4,3)	4,5 (4,3)	3,5 (3,3)	2,9 (2,8)

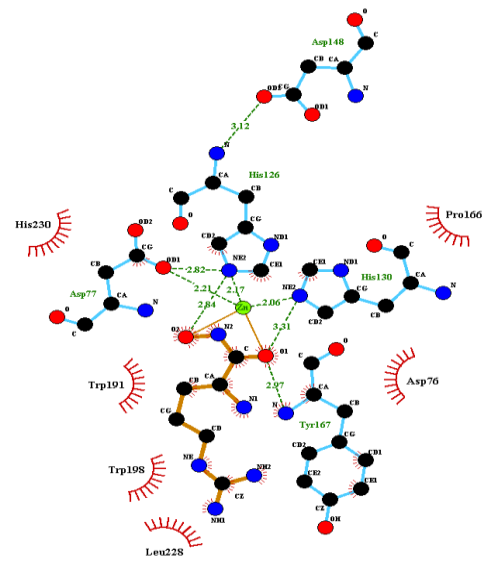
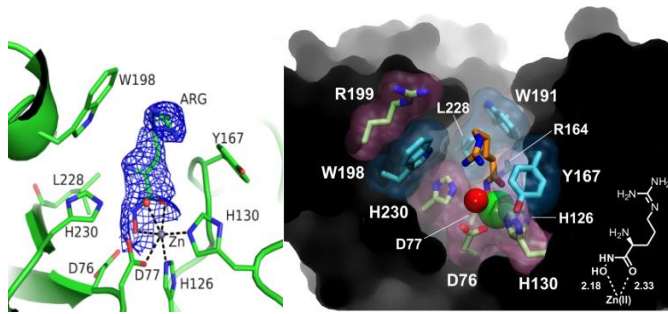
<i>Βελτιστοποίηση</i>							
Ανακλάσεις, σύνολο εργασίας / δείγμα ελέγχου	99165 / 5219	78443 / 3721	25796 / 1315	39170 / 1870	39625 / 2022	19594 / 996	26792 / 1280
$R_{work} /$ R_{free}	0,1733 / 0,2023	0,1613 / 0,1798	0,2157 / 0,2788	0,2220 / 0,2860	0,2256 / 0,2885	0,1792 / 0,2549	0,2051 / 0,2663
<i>Αριθμός ατόμων:</i>							
Πρωτεΐνης / προσδέτη / διαλύτη	6617 / 52 / 701	6628 / 76 / 622	6587 / 58 / 88	6591 / 65 / 91	6611 / 38 / 99	6542 / 48 / 19	6562 / 71 / 59
Παράγοντας Wilson B- factor (\AA^2)	24,84	22,30	49,21	45,71	41,84	60,66	54,91
<i>Μέσος θερμικός παράγοντας (Average B factor, \AA^2)</i>							
Πρωτεΐνης / προσδέτη / διαλύτη	31,20 / 37,40 / 37,00	30,20 / 46,30 / 41,40	57,11 / 72,00 / 37,60	61,70 / 65,90 / 51,30	56,10 / 73,30 / 47,00	73,50 / 83,40 / 49,60	69,60 / 82,00 / 51,40
Μήκη δεσμών r.m.s.ds (\AA)	0,007	0,006	0,012	0,009	0,008	0,010	0,009
Γωνίες δεσμών r.m.s.ds ($^\circ$)	1,024	1,100	1,491	1,220	1,080	1,360	1,190
<i>Περιοχές Ramachandran</i>							
Ευνοϊκές	97%	97%	97%	96%	95%	96%	95%
Μη επιτρεπτές	0,61%	0,49%	0,25%	0%	0,61%	0%	0,12%
PDB ID	5N1J	5N1P	5NC6	5NC9	5NCD	5NEK	5NEL



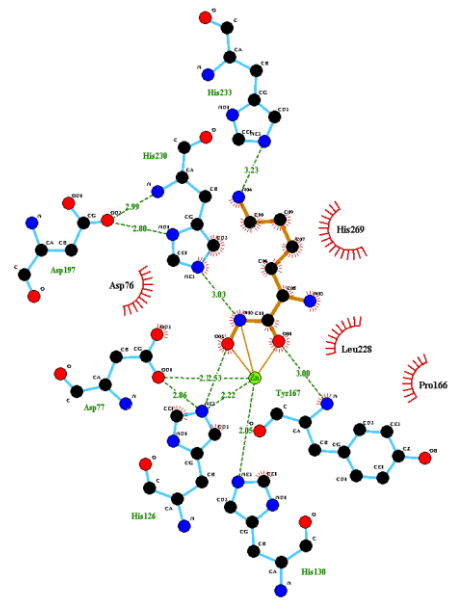
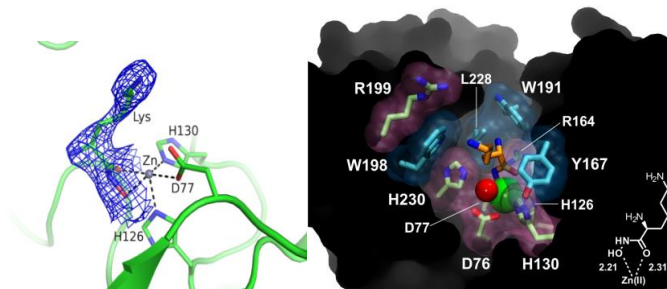
IV



V



VI



(α)

(β)

(γ)

Εικόνα 3.41. Τριάδες εικόνων για τον κάθε προσδέτη **I(1), II (6), III (10), IV (9), V (8), και VI (7),** (α) Ο χάρτης 2Fo-Fc σχεδιασμένος σε πυκνότητα > 3σ. (β) Αναπαράσταση του ενεργού κέντρου της Bc1974 από δομές που καθορίστηκαν με κρυσταλλογραφία ακτίνων-X στα σύμπλοκά της με προσδέτες (Σε ένθετο είναι οι σκελετικές δομές των προσδετών με τις αντίστοιχες αποστάσεις μεταξύ των ατόμων τους και του ψευδαργύρου). Ο χρωματικός κώδικας είναι ίδιος με της εικόνας 3.39. Σε όλες τις περιπτώσεις, απεικονίζεται το μόριο B της τετραμερούς ασύμμετρης μονάδας και ο προσανατολισμός του ενεργού κέντρου είναι όσο το δυνατό πανομοιότυπος προκειμένου να καταδειχθεί η γεωμετρία πρόσδεσης. Στα αρωματικά και μη πολικά αμινοξέα τα άτομα άνθρακα αποδίδονται με κυανό, στα πολικά με φυσικό, και στους προσδέτες με πορτοκαλί και το άτομο S είναι κίτρινο, ενώ όλα τα άλλα χρώματα όπως στην εικόνα 3.39 και (γ) το διάγραμμα LigPlot. Οι δεσμοί υδρογόνου υποδεικνύονται με διακεκομμένες γραμμές μεταξύ των εμπλεκόμενων ατόμων, ενώ οι υδρόφοβες αλληλεπιδράσεις αντιπροσωπεύονται από ένα τόξο με ακτίνες προς τα άτομα του προσδέτη που έρχονται σε επαφή.

3.5 Φυλογενετική ανάλυση και ανάλυση αλληλουχίας στην οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη (PDA)

3.5.1 Οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) στον *Bacillus anthracis* Ames

Στο γονιδίωμα του *B. anthracis* Ames (Read *et al.* 2003) συνολικά ανιχνεύθηκαν 12 ανοιχτά πλαίσια ανάγνωσης γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) χρησιμοποιώντας τη συλλογή αλληλουχιών αναφοράς (RefSeq) του NCBI και συγκεκριμένα τον κωδικό εγγραφής NC_003997.3. Λεπτομέρειες σχετικά με τις απακετυλάσες πολυσακχαρίτη του γονιδιώματος *B. anthracis* Ames που εξετάστηκαν στην παρούσα διατριβή αναφέρονται στον **Πίνακα 3.6**. Περιλαμβάνονται η ονομασία τους και ο κωδικός αριθμός από την NCBI, η υποτιθέμενη λειτουργία τους (Balomenou *et al.* 2015, Zhou *et al.* 2008), η πρόβλεψη της θέσης εντοπισμού τους και η παρουσία ή η απουσία θέσης αποκοπής σύμφωνα με τη βάση δεδομένων LocateP (<http://www.cmbi.ru.nl/locatep-db/cgi-bin/locatepdb.py>, Zhou *et al.* 2008), ο κωδικός στην τράπεζα δεδομένων πρωτεΐνης εάν υπάρχει κρυσταλλική δομή, και οι ομόλογες τους από τον *B. cereus* ATCC14579.

Πίνακας 3.6 Οι PDAs του *Bacillus anthracis* str. Ames που συμπεριλήφθησαν στη μελέτη και οι ομόλογες από τον *Bacillus cereus* str. ATCC 14579.

<i>Bacillus anthracis</i> Ames	NCBI RefSeq: NC_003997.3 Γονίδιο ID (CDS θέση)	NCBI RefSeq Πρωτεΐνη (aa)	Πιθανή λειτουργία	Πρόβλεψη κυτταρικής θέσης από τη Βάση δεδομένων LocateP (1) ταξινόμηση της SwissProt / (2) Μονοπάτι (CS)	Κωδικός PDBID της αντίστοιχης κρυσταλλικής δομής	<i>Bacillus cereus</i> ATCC 14579 ορθόλογο. NCBI Ref Seq: NC_004722.1 (1) Όνομα, (2) Κωδικός Πρωτεΐνης RefSeq (3) (aa)(4) % Ταυτότητα, (5) PDBID
Ba1961	1087084, (1847060, 1847887)	NP_844369 (275)	Απακετυλάση GlcNAc της πεπτιδογλυκάνης	Κυτταροπλασματική/ Ενδοκυτταρική (Χωρίς CS)	-	Bc1960, NP_831730 (275), 94,9 4L1G
Ba3679	1089257, (3383280, 3383921, complement)	NP_845942 (213)	Απακετυλάση GlcNAc της πεπτιδογλυκάνης	Κυτταροπλασματική/ Ενδοκυτταρική (Χωρίς CS)	-	Bc3618, NP_833348 (213), 97,7
Ba3480	1083768, (3198882, 3201665, complement)	NP_845761 (927)	Γλυκοζυλομεταφορά ση και απακετυλάση πολυσακχαρίτη	Μεμβρανική/πολλαλώς- διαμεμβρανική (Χωρίς CS)	-	-
Ba1977	1086062, (1861493, 1862314)	NP_844383 (273)	Απακετυλάση GlcNAc της πεπτιδογλυκάνης	Μεμβρανική/N-τελικώς αγκυροβολημένη (Χωρίς CS)	-	Bc1974, NP_831744 (273), 97,1 5N1J
Ba2944	1085595, (2708639, 2709466)	NP_845280 (275)	Απακετυλάση του ουδέτερου πολυσακχαρίτη	Εξωκυτταρική/Εκκρινόμενη(απελευθερώνεται) (με CS)	-	Bc2929, NP_832677 (275), 94,9
Ba5436	1085036, (4923836, 4924573)	NP_847604 (245)	Απακετυλάση GlcNAc της πεπτιδογλυκάνης	Εξωκυτταρική/Εκκρινόμενη (απελευθερώνεται) (με CS)	-	Bc5204, NP_834868 (245), 93,5
Ba0424	1087807, (444231, 445013)	NP_842967 (260)	Απακετυλάση MurNAc της πεπτιδογλυκάνης	Εξωκυτταρική/Εκκρινόμενη (απελευθερώνεται) (με CS)	2J13	Bc0467, NP_830306 (260), 98,5
Ba0150	1086778, (144110, 144874, complement)	NP_842717 (254)	Αδρανές ως ένζυμο απακετυλάσης πολυσακχαρίτη	Μεμβρανική/N-τελικώς αγκυροβολημένη (Χωρίς CS)	4M1B	Bc0171, NP_830050 (254), 95,3
Ba1836	1086408, (1722907, 1723611, complement)	NP_844255 (234)	Απακετυλάση πολυσακχαρίτη	Εξωκυτταρική/Εκκρινόμενη (απελευθερώνεται) (με CS)	-	Bc1768, NP_831543 (234), 91,9
Ba3943	1086849, (3618407, 3619306, complement)	NP_846187 (299)	Φερόμενη απακετυλάση πολυσακχαρίτη ;	Εξωκυτταρική/Εκκρινόμενη (απελευθερώνεται) (με CS)	-	Bc3804, NP_833526 (299), 95,7
Ba0330	1085388, (338279, 339361, complement)	NP_842877 (360)	Φερόμενη απακετυλάση πολυσακχαρίτη	Εξωκυτταρική/λιπιδιακά υροβολημένη (με CS)	4V33	Bc0361, NP_830200 (360), 90,6 4HD5
Ba0331	1085987, (339513, 340616, complement)	NP_842878 (367)	Φερόμενη απακετυλάση πολυσακχαρίτη	Εξωκυτταρική/λιπιδιακά αγκυροβολημένη (με CS)	6GO1	-

CDS: θέση στο γονιδίωμα του ανοικτού πλαισίου ανάγνωσης της ακολουθίας, aa: τα νούμερα στην παρενθέση αναφέρονται στον αριθμό των αμινοξέων της πρωτεΐνης, RefSeq: αλληλουχία αναφοράς στην NCBI, CS: θέση αποκοπής (cleavage site), η % ταυτότητα: υπολογίστηκε με χρήση του προγράμματος Lalign και αναφέρεται στο πλήρες μήκος της αμινοξικής αλληλουχίας (aa). PDB ID κωδικοί: 2J13 (Oberbarnscheidt *et al.* 2007), 4HD5 (Fadouloglou *et al.* 2013), 4M1B (Strunk *et al.* 2014), 4L1G (Fadouloglou *et al.* 2017), και 4V33, 5N1J, 6GO1 (παρούσα διατριβή).

3.5.2 Οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) στον *Bacillus cereus* ATCC 14579

Στο γονιδίωμα του *B. cereus* ATCC 14579 (Ivanova *et al.* 2003) ανιχνεύθηκαν 11 ανοικτά πλαίσια ανάγνωσης γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) χρησιμοποιώντας τη συλλογή αλληλουχιών αναφοράς (RefSeq) του NCBI και συγκεκριμένα τον κωδικό εγγραφής NC_004722.1. Τα 10 από αυτά αναφέρονται στον πίνακα 3.1 (βλ. §3.5.1). Επιπροσθέτως συμπεριλήφθηκε η πρωτεΐνη Bc3146 που δε βρέθηκε ορθόλογή της στον *Bacillus anthracis* Ames, η οποία έχει σύνολο αμινοξικών καταλοίπων 280. Στη συλλογή αλληλουχιών αναφοράς (RefSeq) του NCBI ο κωδικός εγγραφής της πρωτεΐνης είναι "NP_832887.1", ο κωδικός της εγγραφής του γονιδίου είναι "GeneID: 1205493" και η θέση της είναι 3116925... 3117767, (complement) στο γονιδίωμα του *B. cereus* ATCC 14579.

3.5.3 Φυλογενετική ανάλυση και ανάλυση νουκλεοτιδικής και αμινοξικής αλληλουχίας της δομικής ενότητας NodB

Στην ανάλυση της ακολουθίας της δομικής ενότητας NodB, της οικογένειας PDA του *B. anthracis* Ames, και αυτής του *B. cereus* ATCC14579 συμπεριλήφθηκαν 23 αλληλουχίες (βλ. §3.5.1 και §3.5.2). Ακολουθήθηκε τοπική στοίχιση (local alignment) που εφαρμόστηκε στο τμήμα της ακολουθίας της δομικής ενότητας NodB. Η ομολογία των αμινοξικών αλληλουχιών της δομικής ενότητας NodB της οικογένειας των γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) του *B. anthracis* Ames, παρουσιάζεται στον **πίνακα 3.7** και του *B. cereus* ATCC14579 παρουσιάζεται στον **πίνακα 3.8**. Η συντήρηση αμινοξέων κυμαίνονταν μεταξύ 21% και 43% (ταυτότητα) εντός της κάθε μιας από τις οικογένειες, με εξαίρεση τα ζεύγη Ba0331-Ba0330 (ταυτότητα 59,9%), Ba1977-Ba2944 (ταυτότητα 75,7%), Bc1974-Bc2929 (ταυτότητα 76,5%), καθώς και τα Bc1974-Bc3146 και Bc2929-Bc3146 (~50% ταυτότητα σε κάθε ζεύγος).

Πίνακας 3.7. Μήτρα της ομολογίας (ταυτότητα και ομοιότητα) επί τοις εκατό, μεταξύ των αμινοξικών αλληλουχιών της δομικής ενότητας NodB της οικογένειας των γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) του *B. anthracis* Ames, με τη χρήση του προγράμματος ευθυγράμμισης αλληλουχιών κατά ζεύγη Lalign.

NodB	Ba0331	Ba0330	Ba3679	Ba3943	Ba0424	Ba1961	Ba1836	Ba3480	Ba0150	Ba1977	Ba2944	Ba5436
Ba0331		59.9	29.2	28.3	27.9	27.8	25.6	24.3	24.0	24.6	23.1	21.0
Ba0330	82.4		25.4	25.9	29.5	23.4	30.6	23.9	30.3	30.6	27.9	23.4
Ba3679	54.0	57.7		28.4	33.0	37.4	32.8	38.9	36.7	35.0	32.3	29.9
Ba3943	59.6	54.5	67.5		31.1	31.2	27.1	31.0	30.2	28.9	27.4	28.4
Ba0424	61.2	62.0	64.5	69.4		33.8	27.7	33.2	30.8	29.9	27.7	30.1
Ba1961	58.3	63.9	71.3	67.3	65.7		32.5	39.6	29.0	40.2	33.0	28.1
Ba1836	58.3	64.9	68.2	64.1	62.4	63.4		30.3	25.4	31.8	29.1	28.1
Ba3480	59.1	61.5	73.2	65.0	65.3	71.1	63.6		32.8	34.7	34.0	33.8
Ba0150	55.8	69.7	68.8	70.4	68.2	68.2	65.8	67.2		33.3	31.5	27.4
Ba1977	51.6	57.7	65.5	60.2	61.3	68.8	60.6	60.2	64.1		75.7	40.9
Ba2944	56.2	59.5	62.6	58.5	57.7	63.8	60.7	60.0	59.4	91.7		39.8
Ba5436	54.3	52.1	59.7	58.9	56.6	65.9	65.9	62.6	61.1	72.6	69.7	

Ομοιότητα %

Ταυτότητα %

Πίνακας 3.8. Μήτρα της ομολογίας (ταυτότητα και ομοιότητα) επί τοις εκατό, μεταξύ των αμινοξικών αλληλουχιών της δομικής ενότητας NodB της οικογένειας των γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) του *B. cereus* ATCC14579 με τη χρήση του προγράμματος ευθυγράμμισης αλληλουχιών κατά ζεύγη Lalign.

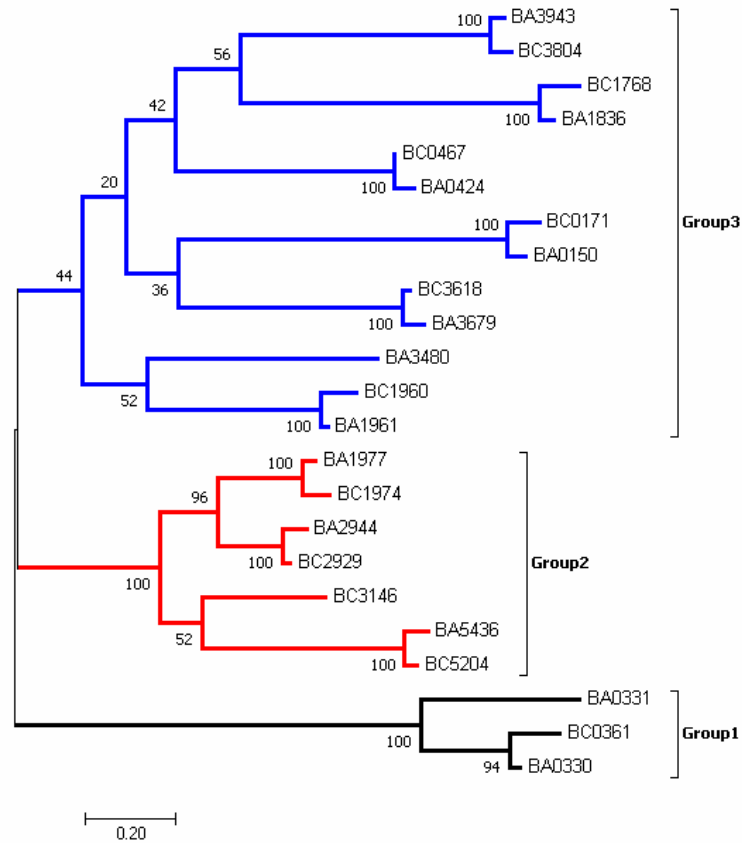
NodB	Bc0361	Bc3618	Bc3804	Bc0467	Bc1960	Bc1768	Bc3146	Bc0171	Bc1974	Bc2929	Bc5204
Bc0361		26.1	27.7	27.9	24.8	31.0	23.8	25.8	28.8	27.6	22.5
Bc3618	58.5		28.4	33.0	36.4	32.8	29.2	36.2	34.0	31.8	31.1
Bc3804	56.2	67.5		32.1	31.2	26.0	29.9	29.6	28.9	27.8	29.5
Bc0467	61.2	64.5	69.4		32.9	27.0	31.8	31.7	29.4	29.1	30.6
Bc1960	62.1	71.3	68.3	65.7		32.8	35.2	29.0	39.7	35.9	28.6
Bc1768	62.1	68.7	65.6	61.9	65.1		30.6	23.3	29.8	30.6	29.6
Bc3146	55.8	59.0	58.8	54.7	64.2	60.7		28.6	50.0	51.1	41.8
Bc0171	67.7	68.8	70.9	67.8	68.2	64.2	58.2		32.3	31.8	26.1
Bc1974	55.0	65.0	60.7	60.8	68.3	59.1	75.7	63.1		76.5	39.9
Bc2929	60.0	61.1	59.0	58.6	64.1	60.7	76.2	57.5	91.3		43.1
Bc5204	51.4	63.2	61.1	56.0	64.9	61.6	76.0	60.1	71.2	70.8	

Ομοιότητα %

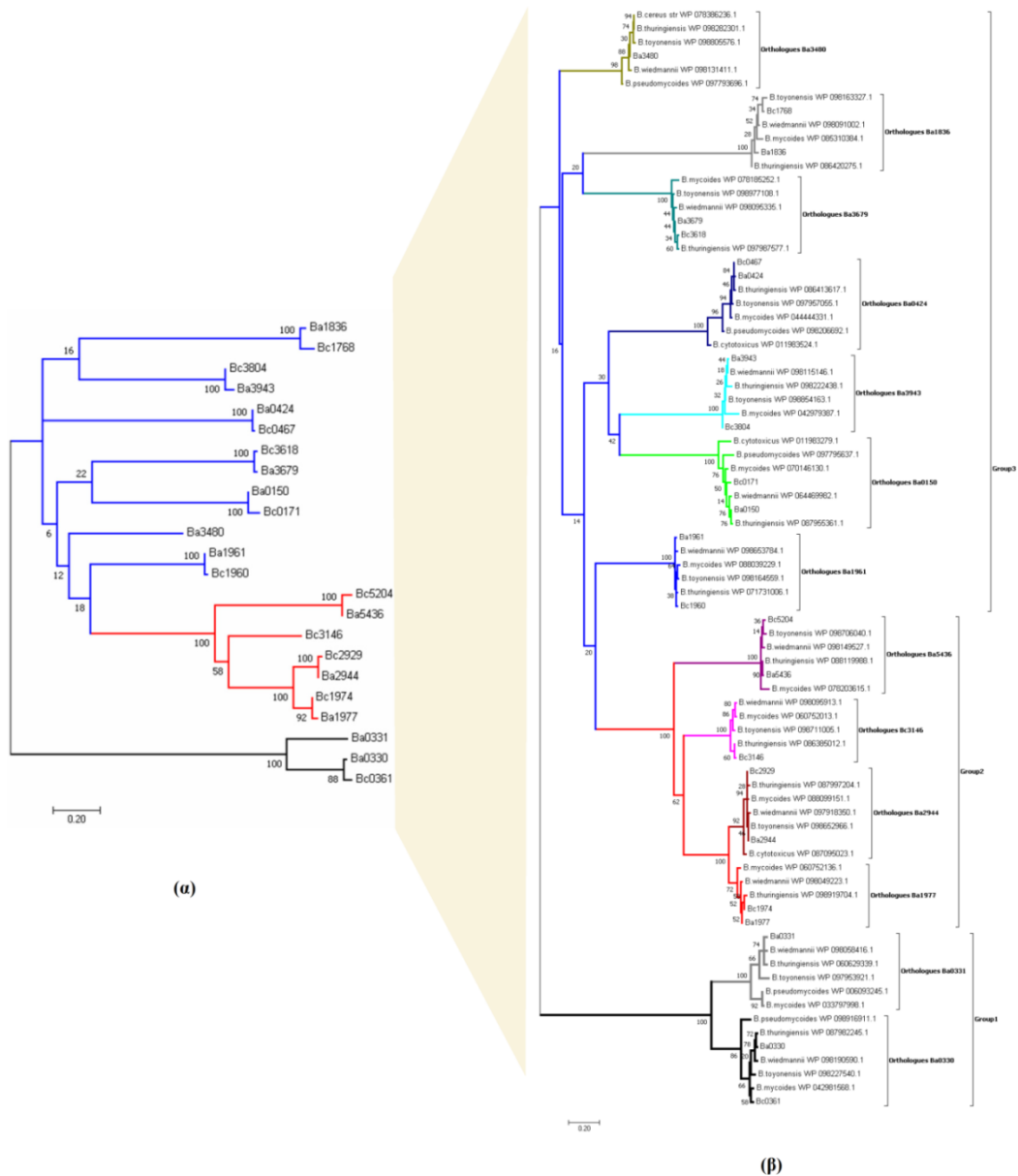
Ταυτότητα %

3.5.3.1 Φυλογενετική ανάλυση

Η κατασκευή δενδρογραμμμάτων έγινε τόσο με βάση τις νουκλεοτιδικές αλληλουχίες (Εικόνα 3.42) όσο και με βάση τις αμινοξικές αλληλουχίες (Εικόνα 3.43α). Η φυλογενετική ανάλυση της καταλυτικής περιοχής NodB αποκάλυψε την κατανομή των ακολουθιών σε τρεις ομάδες (1, 2 και 3 ή αλλιώς υποοικογένειες). Το πρόγραμμα BLAST, που βασίζεται σε ευρετικό αλγόριθμο αναζητώντας περιοχές τοπικής ομοιότητας, χρησιμοποιήθηκε για την αναζήτηση ομολόγων αλληλουχιών, προκειμένου να επεκταθεί η κατασκευή δενδρογράμματος ώστε να περιλαμβάνει και άλλα μέλη της ομάδας του *B. cereus*. Έτσι ένα σύνολο 80 αμινοξικών αλληλουχιών συμπεριλήφθηκαν στο δενδρόγραμμα που κατασκευάστηκε (Εικόνα 3.43 β). Η φυλογενετική ανάλυση της καταλυτικής περιοχής αποκάλυψε την κατανομή των ακολουθιών της NodB σε τρεις ομάδες (1, 2 και 3), οι οποίες διατηρούνται και στα δύο δενδρογράμματα. Επίσης η ίδια κατανομή στις τρεις ομάδες εμφανίζεται και στο δέντρο των νουκλεοτιδικών αλληλουχιών. Η συνολική τοπολογία του δένδρου των αμινοξικών αλληλουχιών συντηρείται στενά στο δένδρο των νουκλεοτιδικών αλληλουχιών. Η ομάδα 1 περιλαμβάνει μόνο τις Ba0330, Bc0361 και Ba0331. Στην ομάδα 2 συγκαταλέγεται η Bc1974 που είναι γνωστή PG GlcNAc απακετυλάση, και στην ομάδα 3 συγκαταλέγονται οι Ba0424, Bc1960 και Ba0150. Από τη φυλογενετική ανάλυση της διαπιστώνεται ότι σχηματίζονται ισχυρές μονοφυλετικές ομάδες εντός των ορθόλογων τους, που υποστηρίζεται ισχυρά από τις υψηλές τιμές ανάλυσης δοκιμασίας (bootstrap) κοντά στο 100%.



Εικόνα 3.42 Η μοριακή φυλογενετική ανάλυση της νουκλεοτιδικής αλληλουχίας της δομικής ενότητας NodB των *B. anthracis* Ames (BA) και *B. cereus* ATCC14579 (BC) των είκοσι τριών αλληλουχιών PDA που μελετώνται. Απεικονίζεται το φυλογενετικό δέντρο με τη μέγιστη λογαριθμική πιθανοφάνεια που υπολογίστηκε μετά την εφαρμογή της ανάλυσης δοκιμασίας bootstrap για 500 επαναλήψεις. Τα αρχικά δενδρόγραμμα χωρίς ρίζα προέκυψαν εφαρμόζοντας τους αλγόριθμους Neighbor-Join και BioNJ από τη μήτρα γενετικών αποστάσεων που υπολογίστηκε με τη μέθοδο Maximum Composite Likelihood (MCL). Η κατανομή των ακολουθιών σε τρεις ομάδες παρουσιάζονται με διαφορετικά χρώματα (Ομάδα 1 με μαύρο χρώμα, Ομάδα 2 με κόκκινο χρώμα και Ομάδα 3 με μπλε χρώμα). σε στελέχη της ομάδας του *B. cereus* Η φυλογενετική ανάλυση πραγματοποιήθηκε με το λογισμικό φυλογενετικής ανάλυσης MEGA (Kumar *et al.* 2016).



Εικόνα 3.43 Μοριακή φυλογενετική ανάλυση της περιοχής NodB PDA. **(α)** Το χωρίς ρίζα δέντρο με τη μέθοδο της Μέγιστης Πιθανοφάνειας των 23 πρωτεϊνικών αλληλουχιών της δομικής ενότητας PDA NodB από τα στελέχη *B. anthracis* Ames (Ba) και *B. cereus* ATCC14579 (Bc). Απεικονίζεται το φυλογενετικό δέντρο με τη μέγιστη λογαριθμική πιθανοφάνεια που υπολογίστηκε μετά την εφαρμογή της ανάλυσης δοκιμασίας bootstrap για 1000 επαναλήψεις. Οι ακολουθίες κατανέμονται σε τρεις ομάδες που παρουσιάζονται με διαφορετικά χρώματα (μαύρο, κόκκινο, μπλε). **(β)** Απεικονίζεται το χωρίς ρίζα φυλογενετικό δέντρο 80 αμινοξικών αλληλουχιών PDA ομολόγων των στελεχών *B. anthracis* Ames (Ba) και *B. cereus* ATCC14579 (Bc) με τη μέγιστη λογαριθμική πιθανοφάνεια που υπολογίστηκε μετά την εφαρμογή της ανάλυσης δοκιμασίας bootstrap για 1000 επαναλήψεις. Η ομαδοποίηση των ορθολόγων αναφορικά με τις 23 αλληλουχίες του *B. anthracis* / *B. cereus* παρουσιάζεται με διαφορετικά χρώματα. Η φυλογενετική ανάλυση πραγματοποιήθηκε με το λογισμικό φυλογενετικής ανάλυσης MEGA (Kumar *et al.* 2016).

3.5.3.2 Χαρακτηριστικά της αλληλουχίας

Οι PDAs του *B. anthracis* και *B. cereus* περιείχαν την καταλυτική περιοχή NodB στο καρβοξυτελικό άκρο τους, εκτός της Ba3480, όπου η περιοχή NodB ήταν τοποθετημένη μεταξύ των αμινοξικών καταλοίπων 295 και 491, σε σύνολο 927 αμινοξέων της πρωτεΐνης.

Τα κατάλοιπα της θέσης δέσμευσης συντηρούνται υψηλά στη καταλυτική περιοχή NodB (Εικόνα 3.44). Διατηρούν επίσης τη θέση τους στις τρισδιάστατες δομές. Λεπτομερώς, τρία από τα πέντε γνωστά μοτίβα (MT, Blair *et al.* 2005) στην αλληλουχία που σχετίζονται με την κατάλυση συντηρούνται σε όλες τις ομάδες κατά μήκος της αμινοξικής αλληλουχίας (εμφανίζονται με πορτοκαλί ράβδους στην εικόνα 3.44. Τα άλλα δύο ποικίλλουν ως προς το χαρακτήρα και τη θέση στις τρεις ομάδες (υποοικογένειες) των PDA (απεικονίζονται με ράβδους πλαισιωμένες με χρώμα πράσινο στην εικόνα 3.44. Στο μοτίβο MT1η ακολουθία TFDDG είναι εξαιρετικά συντηρημένη σε όλα τα μέλη (με M αντί για F στην Ba0331), εκτός από την περίπτωση των ορθόλογων Ba3943/Bc3804 που χάνεται πλήρως, των Ba0150/Bc0171 που συντηρείται η TFD και των Ba0424/Bc0171 που εδώ μία Asn βρίσκεται στη θέση του δεύτερου ασπαρτικού οξέος TFDNG (Oberbarnscheidt *et al.* 2007) όπως συμβαίνει και με την BsPdaA (Blair & van Aalten 2004).

Στο MT2 με [H][ST][AYWFM][NTSK][H], τρία αμινοξικά κατάλοιπα διαχωρίζουν τα δύο κατάλοιπα His που μετέχουν στην δέσμευση του μετάλλου, εκτός από τις δύο ορθόλογες Ba0150/Bc0171 που στερούνται αυτών των δύο συντηρημένων καταλοίπων His. Το μοτίβο MT3 είναι καλά συντηρημένο συνήθως εμφανίζεται ως gpPxG αναλυτικότερα [RA][PYV][PS][YGSRN][GE].

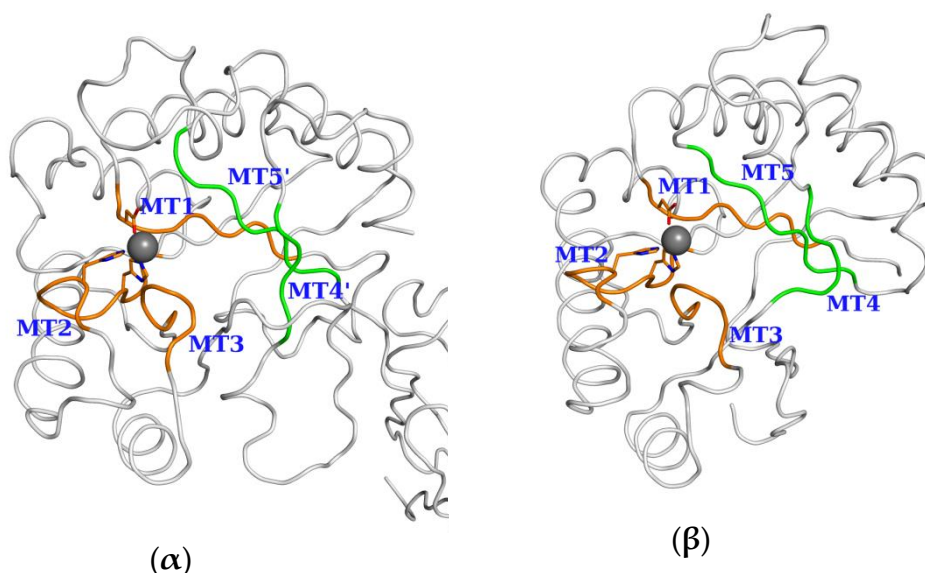
Στο MT4 εμφανίζεται [W]xx[DE]xx[D][W] για τις Ομάδες 2 και 3. Για την Ομάδα1, το MT4 βρίσκεται 15 κατάλοιπα καθοδικά της αλληλουχίας, μετονομάστηκε MT4' με αλληλουχία [R][HV][RF].

Το μοτίβο αλληλουχίας MT5 για τις Ομάδες 2 και 3 είναι [IV][IV][L][LMQ][H] ενώ για τις πρωτεΐνες της Ομάδας 1, εντοπίστηκε στο αμινοτελικό άκρο του NodB και μετονομάστηκε MT5' με αλληλουχία VLMYH (τα πρώτα τρία αμινοξικά κατάλοιπα του MT5' καθώς και το τελευταίο είναι σε συμφωνία με το MT5) (Εικόνα 3.44). Η τελική His στο MT5(5') συντηρείται αυστηρά.

Παρά τη μετατοπισμένη θέση των μοτίβων MT4' και MT5' στην αλληλουχία, οι θέσεις τους διατηρήθηκαν μεταξύ των τρισδιάστατων δομών των ενζύμων που

συγκρίθηκαν (**Εικόνα 3.45**). Τα MT1 και MT2 εμπλέκονται άμεσα με τη συναρμογή του μετάλλου στο ενεργό κέντρο. Το μοτίβο MT1 συνεισφέρει το ασπαρτικό οξύ και το MT2 συνεισφέρει τις δύο ιστιδίνες. Τα MT3 και MT4 ευρίσκονται αντικριστά μεταξύ τους οριοθετώντας τη θέση πρόσδεσης, ενώ το MT5 οριοθετεί το κατώτερο τμήμα της κοιλότητας δέσμευσης. (Arnaouteli *et al.* 2015, Giastas *et al.* 2018, Andreou *et al.* 2018, Andreou *et al.* 2019).

Εικόνα 3.44 Διάγραμμα στοίχισης αλληλουχίας της δομικής ενότητας NodB των είκοσι τριών *B. anthracis* Ames και *B. cereus* ATCC14579 PDAs. Τα υψηλά συντηρημένα κατάλοιπα μεταξύ των αλληλουχιών δίδονται με έντονα γράμματα. Οι πορτοκαλί και πράσινες ράβδοι κάτω από τις αλληλουχίες δείχνουν τα συντηρημένα μοτίβα (MTs). Οι πρωτεΐνες με κρυσταλλογραφικά γνωστές τριτοταγείς δομές έχουν το όνομα τους με χρώμα πορφυρό. Τα αμινοξέα που βρίσκονται στο σχηματισμό της θέσης πρόσδεσης πλαισιώνονται από μπλε, κυανό, φούξια, κίτρινο, πορτοκαλί ή κόκκινο χρώμα, αντίστοιχα. Η δευτεροταγής δομή των Ba0331 και Bc1974 παρουσιάζεται στην κορυφή και στο κάτω μέρος (α-έλικες ως σπείρες, β-κλώνοι ως βέλη, στροφές ως T). Η αρίθμηση δίδεται σύμφωνα με την Ba0331. Το διάγραμμα έγινε με χρήση του προγράμματος ESPRIPT (Gouet *et al.* 2003).

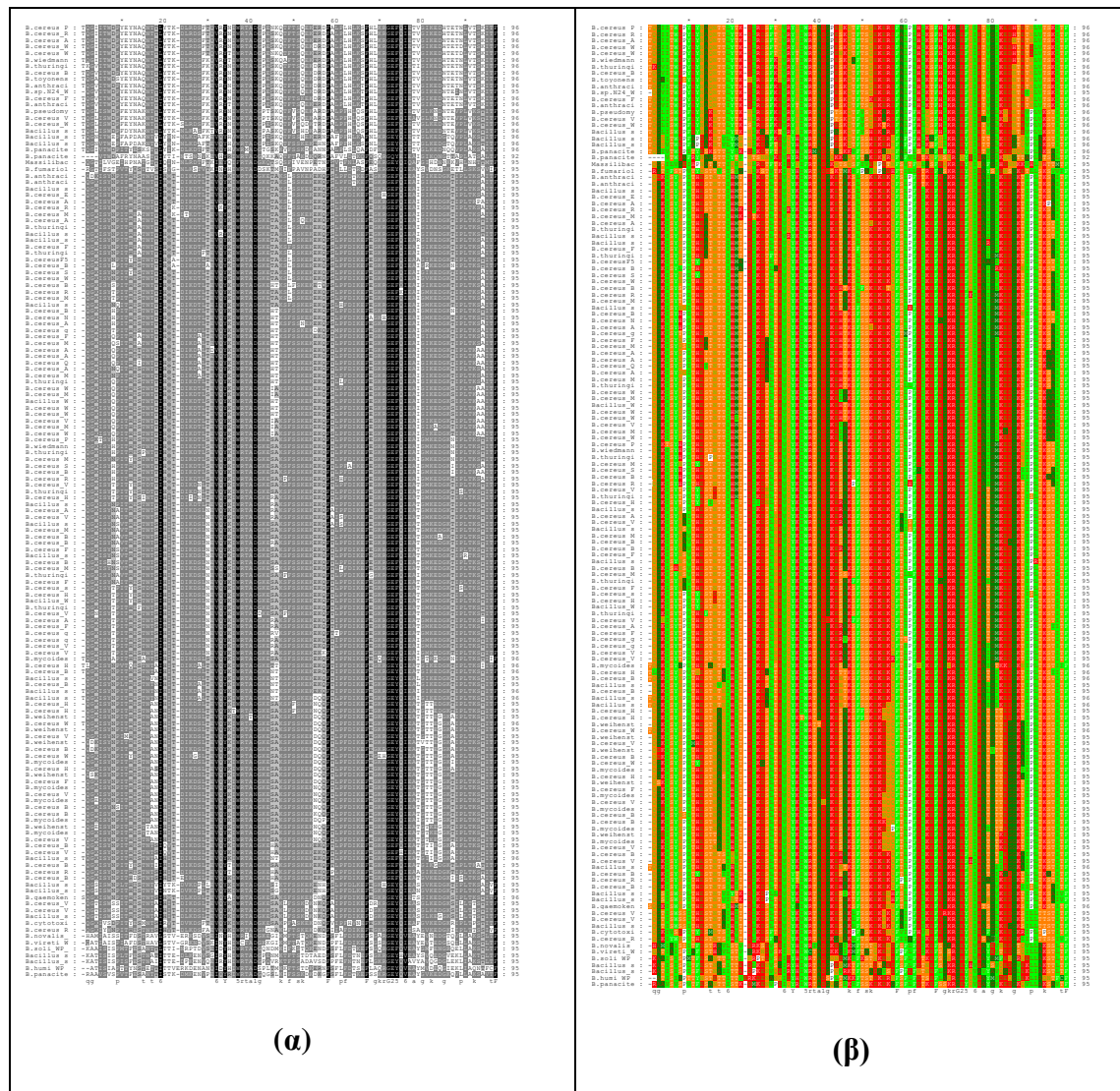


Εικόνα 3.45 Θέση των συντηρημένων μοτίβων MT1-MT5(') στην δομική ενότητα NodB των δομών Ba0331 και Bc1974. Τα στοιχεία που φαίνονται στην Εικόνα 3.44 της ευθυγράμμισης των αλληλουχιών με πορτοκαλί (MT1, MT2, MT3, MT4 και MT5) και πράσινες ράβδους (MT4' και MT5') παρουσιάζονται εδώ στην τρισδιάστατη αναπαράσταση των ατόμων Ca για τις δομές (α) Ba0331 και (β) Bc1974 που αντιπροσωπεύουν την Ομάδα 1 και την Ομάδα 2, αντίστοιχα.

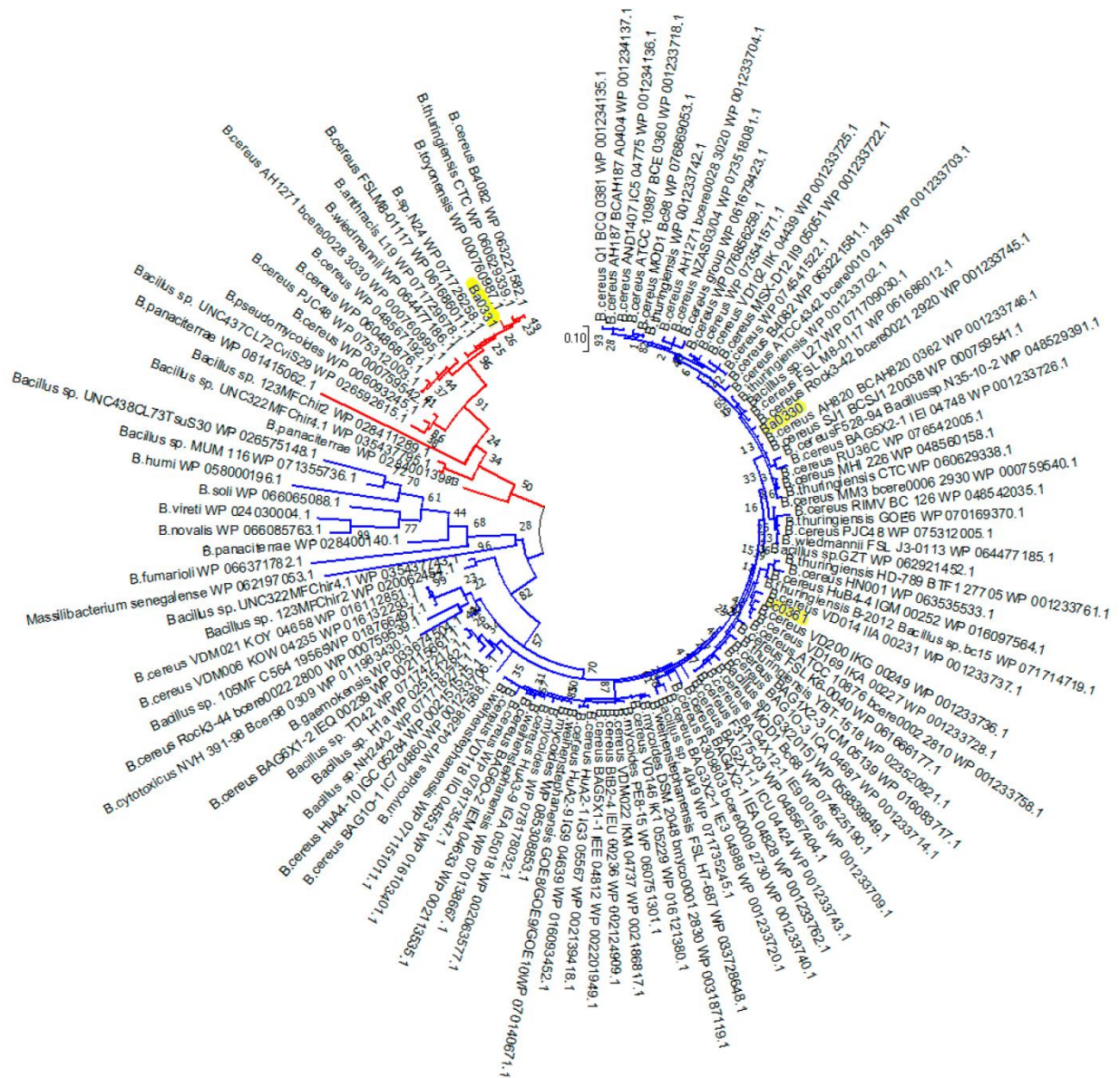
3.5.4 Φυλογενετική ανάλυση και ανάλυση αλληλουχίας της δομικής ενότητας Fn3

Οι πρωτεΐνες της Ομάδας 1 (υποοικογένεια 1) περιλαμβάνουν τη δομική ενότητα Fn3. Στην πολλαπλή ευθυγράμμιση αμινοξικών αλληλουχιών των Fn3 δομικών ενότητων από PDA βακίλων στην **εικόνα 3.46** ξεχωρίζουν η υψηλά συντηρημένη περιοχή RGEF_xI και η συνεργιστική της περιοχή YxxWRTAD. Στην τελική κατασκευή του δένδρογράμματος

συμπεριλήφθηκαν 117 αλληλουχίες από PDA Bacilli που βρέθηκαν να περιέχουν τη δομική ενότητα Fn3 μετά την απομάκρυνση των ισομορφών και των περιττών αλληλουχιών. Από αυτές, 105 αλληλουχίες ήταν από στελέχη της ομάδας *B. cereus* ενώ 12 ανήκαν στο γένος *Bacillus*. Οι αλληλουχίες της δομικής ενότητας Fn3 των Ba0330 και Ba0331 ταξινομήθηκαν σε δύο διακριτές ομάδες. Ο διαχωρισμός των ομολόγων Ba0330 (συμπεριλαμβανομένου της Bc0361) από την Ba0331 και των ομολόγων της φαίνεται στην (Εικόνα 3.47).



Εικόνα 3.46 Πολλαπλή ευθυγράμμιση αμινοξικών αλληλουχιών των δομικών ενότητων Fn3 από PDA βακίλων. **(α)** Η ένταση του χρώματος σε κλίμακα του γκρι δείχνει το βαθμό συντήρησης κατά μήκος της ευθυγράμμισης (μαύρο για τα υψηλά συντηρημένα κατάλοιπα και διαβαθμίσεων του γκρι για μερικώς συντηρημένα). **(β)** Τα κατάλοιπα είναι χρωματισμένα σύμφωνα με τις φυσικοχημικές τους ιδιότητες. Τα υδρόφοβα (Leu, Ile, Val, Phe, Trp και η Tyr) με ανοιχτό πράσινο, τα μικρά (Ala, Gly) με σκούρο πράσινο, τα θειούχα (Cys, Met) με κυπαρισσί, τα πολικά (Ser, Thr, Gln, Asn) με πορτοκαλί, τα φορτισμένα (Asp, Glu, Arg, Lys, His) με κόκκινο και η προλίνη με λευκό χρώμα. Οι εικόνες δημιουργήθηκαν με τη χρήση του προγράμματος GENEDOC (Nicholas *et al.* 1997).



Εικόνα 3.47 Μοριακή φυλογενετική ανάλυση της δομικής ενότητας Fn3 των PDAs. Το δέντρο χωρίς ρίζα 117 αμινοξικών αλληλουχιών κατασκευάστηκε με τη μέθοδο της Μέγιστης Πιθανοφάνειας. Απεικονίζεται το φυλογενετικό δέντρο με τη μέγιστη λογαριθμική πιθανοφάνεια που υπολογίστηκε μετά την εφαρμογή της ανάλυσης δοκιμασίας bootstrap για 1000 επαναλήψεις. Οι αλληλουχίες συγκεντρώνονται σε δύο διαφορετικές ομάδες (Ba0330-like, μπλε και Ba0331-like κόκκινο). Οι Ba0330, Bc0361 και Ba0331 επισημαίνονται με κίτρινο χρώμα. Οι αλληλουχίες είναι PDAs από είδη βακίλων που έχουν βρεθεί ότι περιέχουν την περιοχή Fn3. Η φυλογενετική ανάλυση πραγματοποιήθηκε με το λογισμικό φυλογενετικής ανάλυσης MEGA (Kumar *et al.* 2016).

3.6 Τα γονίδια *ba0330* και *ba0331*

Στις αλληλουχίες Ba0330 και Ba0331 ταυτοποιήθηκε ένα λιποκυτίο, σύμφωνα με τη βάση δεδομένων DOLOP (Babu *et al.* 2006), έτσι οι δύο πρωτεΐνες ανήκουν στην οικογένεια των λιποπρωτεϊνών (βλ. §1.6) (πίνακας 3.9).

Οι πρόδρομοι των λιποπρωτεϊνών (Lpp) περιείχαν ένα αμινοτελικό πεπτίδιο οδηγό μήκους 17 αμινοξέων. Το λιποκυτίο της Ba0330 και της ορθόλογης Bc0361 βρέθηκε LAGC και της Ba0331 LVGC. Η αλληλουχία του συνδέθηκε με τη δομική ενότητα Fn3 μέσω μιας α-έλικας επτά στροφών για την Ba0331 και πέντε στροφών για τις Ba0330 και Bc0361, (σύμφωνα με την πρόβλεψη του QUARK, Xu & Zhang 2012), που ίσως έχει το ρόλο άρθρωσης ευκινισίας ή ως μεσοδιαστηματικής περιοχής. Δύο στροφές της α-έλικας που περιείχε την ακολουθία QIQETTA υπήρχαν στην Ba0331 μετά το αμινοτελικό τμήμα πρόσδεσης. Εισαγωγή 15-νουκλεοτιδίων στο γονίδιο *ba0331* κωδικεύουν το φορτισμένο πενταπεπτίδιο EQKKA το οποίο βρίσκεται καθοδικά της μεσοδιαστηματικής περιοχής, και η πρόβλεψη έδειξε διαμόρφωση σχεδόν δύο στροφών α-έλικας. Συνολικά, η Ba0331 έχει επτά αμινοξέα περισσότερα από την Ba0330. Η αλληλουχία που σχηματίζει την α-έλικα QIQETTA βρίσκεται στην Ba0331 και επίσης βρίσκεται μοναδικά παρούσα στην διεπαφή σχηματισμού του ολιγομερούς διαύλου της δομικής ενότητας D2 του προστατευτικού αντιγόνου (PA) (Feld *et al.* 2010, Kintzer *et al.* 2009), ενός συστατικού της τοξίνης του άνθρακα υπεύθυνου μαζί με το LF και το EF για τη λοιμογόνο δράση του *B. anthracis* (βλ. 1.5.2)

Η γονιδιωματική αλληλουχία του *ba0331* σε στοίχιση με εκείνη από το PA (protective antigen)

```
ba0331: CAA ATA CAA GAA ACA ACG GCG
PA      : CAA ATT CAA GAA ACA ACT GCA
          Q   I   Q   E   T   T   A
```

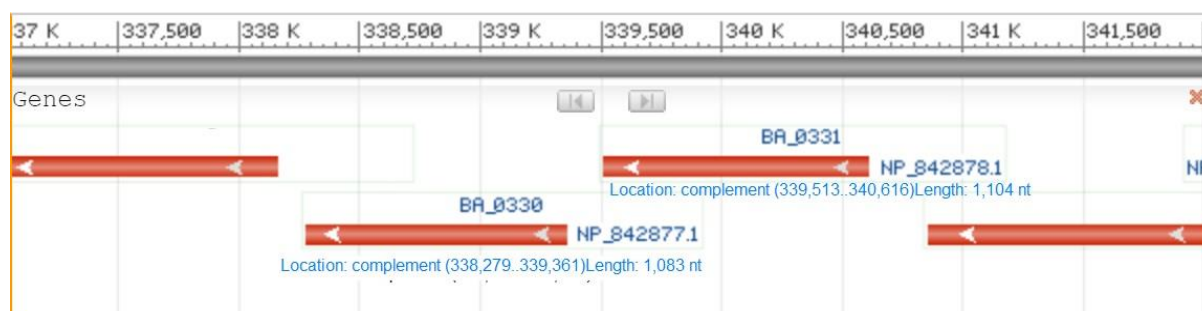
Τα στοιχεία του PA είναι από την εγγραφή πλήρους αλληλουχίας του πλασμιδίου PX01 του *B. anthracis*, αλληλουχία αναφοράς NCBI Reference Sequence: NC_001496.1, κωδικός εγγραφής γονιδίου "GeneID:3361714", κωδικός εγγραφής πρωτεΐνης NP_052806.1 προστατευτικό αντιγόνο (PA, protective antigen, pagA)

Πίνακας 3.9 Οι αλληλουχίες Ba0330, Ba0331, Bc0361 επισημαίνονται οι τρεις περιοχές της αλληλουχίας σήματος των λιποπρωτεϊνών.

<p>Ba0330 MRKYAAIALCTSAAILAGCNTSNVSQEPNKERKVKQETKKQAEVQEQGKISYNPIITHESTN TTIHMTDIKDTLTEVQYKIWRADGKETAKSLSSKEKEKQFSLPFDTKEFEGKRGEFQIE AIGIKEDGKTIPLTKSAITFEQKVPVLMYHAIDDDYHGQGIKDLFVSPANFEAQMKYLKDN GYTLLTFERWGDINKVNKPIFVTFDDGMKNMNAFHVLRQKLDKDDTFKPVATEYMIVNNVD AEGSLSTSDIKEMVDSGIFSMQSHATATHADLPKITNYEEELKESKEKLEKITGKPVIAVA YPFQGHVDDKVVAAETKQYQFATTTKPGKFITKGEPEDELLKMKRVRIHHTTTVEQFASSIK</p> <p>Μήκος αμινοξικής αλληλουχίας:360 κατάλοιπα MB: 40765,584 Da</p>	<p>n-region:MRK h-region: YAAIALCTSAI c-region:LAGC</p>
<p>Ba0331 MKKYTYIALLSSAILVGCNTSSASDKQIQETTAQAVEQKKAKAVELPLTQGVITWDPYE YNAQNTTLYTKDLRDSFKEVRYNIWRADGPESKQTFTSQEKDRDFALPLHLKTFHLKRG EFQIETVGIKEDNTETNLVTSKITFQQHVPVLMYHAIEKFPGPSDGDYGLYVPPEQFEKH MQYLKDNNGYTMLTFERWNDINRVNKPFIITMDDGRKNNMNAHLILQKLDKDDTFQPAATEF LTANEIDKPNRLSTDDIKQMMDSGIFSIQSHTANHTMMAHSNNYDEELRGSKEKIEALTG KKVIALAYPVGSYNDDPAVEETKQYEFVAVTTDHGNHITKGPNEQYLKRFHFGPNTSME KFISLIK</p> <p>Μήκος αμινοξικής αλληλουχίας:367 κατάλοιπα MB :42107,923 Da</p>	<p>n-region:MKK h-region: YTYIALLSSAI c-region:LVGC</p>
<p>Bc0361 MRKYAAIALCTSAAILAGCNTSNVSQEPKKEKKVQEVAIQKEALQEQGKISYTPITHESTN TSIHITDLKDSLNEVQYKIWRADGKERAKSFSSKEKEKQFTIPFDIKEFEGKRGEFQIE ATGMKEDGKTIPLTKSIIITFEQKVPVLMYHAIDDDYHGQGIKDLFVSPANFEAQMKHLKDN GYTLLTFERWGDINKVNKPIFVTFDDGMKNMNAFRVLRQKLDKDDTFKPAATEYMIVDNVD VEGALSTSEIKEMVDSGIFSVQSHTATHADLPKITNYEEELKGSKEKLEKITGKPVIAIA YPFQGHVDDKVVTTETKQYQFATTTKPGQFITKGEPEDELLKMKRVRIHHTTTVEQFASSIK</p> <p>Μήκος αμινοξικής αλληλουχίας :360 κατάλοιπα MB :40767,756 Da</p>	<p>n-region:MRK h-region: YAAIALCTSAI c-region:LAGC</p>

3.6.1 Η χρωμοσωμική οργάνωση των γονιδίων *ba0330* και *ba0331*

Η χρωμοσωμική οργάνωση των γονιδίων *ba0330* και *ba0331* παρουσιάζεται στην **Εικόνα 3.48**. Η διάταξη είναι εν σειρά και έχουν την ίδια κατεύθυνση.



Εικόνα 3.48 Γράφημα της διάταξης του γονιδιώματος του *Bacillus anthracis* str. Ames (NCBI Reference Sequence: NC_003997.3) στην περιοχή των γονιδίων *ba0330* και *ba0331* (στο γράφημα παρουσιάζονται με κόκκινα ορθογώνια πλαίσια οι αντίστοιχες πρωτεΐνες και οι κωδικοί αριθμοί τους).

3.6.2 Σύγκριση νουκλεοτιδικής αλληλουχίας, αμινοξικής αλληλουχίας, τρισδιάστατης δομής

Στον **πίνακα 3.10** παρουσιάζονται οι διαφορές στην αμινοξική αλληλουχία ανάμεσα στις Ba0330 και Ba0331.

Πίνακας 3.10 Πίνακας διαφορών αμινοξικής αλληλουχίας ανάμεσα στην Ba0330 και Ba0331

Περιοχή	Αριθμός αμινοξέων που διαφέρουν	Σύνολο αμινοξικής αλληλουχίας εξεταζόμενης περιοχής	Ομοιότητα %
signal peptide	6	18	67
hinge region	23	30	23,3
Fn3 domain	51	96	47,4
NodB domain	90	225	59,9

Ο υπολογισμός του ρυθμού συνώνυμων (d_S) και μη-συνώνυμων (d_N) αντικαταστάσεων της δομικής ενότητας NodB σύμφωνα με το πακέτο PAML (Goldman & Yang 1994) έχει ως εξής:

$$d_S = 63,7315 \quad (\text{Ρυθμός συνώνυμων αντικαταστάσεων})$$

$$d_N = 0,2917 \quad (\text{Ρυθμός μη-συνώνυμων αντικαταστάσεων})$$

$$d_N/d_S = 0,0046$$

Ο υπολογισμός του ρυθμού συνώνυμων (d_S) και μη-συνώνυμων d_N αντικαταστάσεων της δομικής ενότητας Fn3 σύμφωνα με το πακέτο PAML (Goldman & Yang 1994) έχει ως εξής:

$$d_S = 7,2222 \quad (\text{Ρυθμός συνώνυμων αντικαταστάσεων})$$

$$d_N = 0,4303 \quad (\text{Ρυθμός μη-συνώνυμων αντικαταστάσεων})$$

$$d_N/d_S = 0,0596$$

Η δομική ενότητα NodB παρουσιάζει πιο αργό ρυθμό εξέλιξης από τη δομική ενότητα Fn3 που φαίνεται να είναι δεκαπλάσιος

Ο υπολογισμός του ρυθμού συνώνυμων (P_S) και μη-συνώνυμων (P_N) πολυμορφισμών της δομικής ενότητας NodB και της τιμής NI με βάση τη δοκιμασία McDonald-Kreitman (McDonald & kreitman 1991) έχει ως εξής

	Πολυμορφισμοί	Απόκλιση	Σύνολο
	$P_S=106$	58,64	164,64
	$P_n=119$	10,13	129,13
Σύνολο	225	68,77	293,77

Η εκτίμηση του NI (Neutrality Index) είναι 6,498

Ο υπολογισμός του ρυθμού συνώνυμων (P_S) και μη-συνώνυμων (P_N) πολυμορφισμών της δομικής ενότητας Fn3 και της τιμής NI με βάση τη δοκιμασία McDonald-Kreitman (McDonald & kreitman 1991) έχει ως εξής

	Πολυμορφισμοί	Απόκλιση	Σύνολο
	$P_S=29$	12,79	41,79
	$P_n=67$	11,38	78,38
Σύνολο	96	24,17	120,17

Η εκτίμηση του NI (Neutrality Index) είναι 2,596

Η ευθυγράμμιση της αμινοξικής και της νουκλεοτιδικής ακολουθίας των Ba0330 και Ba0331 παρουσιάζεται στην **Εικόνα 3.49**. Η πρώτη περιοχή εισαγωγής κενών (gap) κατά την στοίχιση, αποτελεί ένδειξη ολίσθησης κλώνου (Strand slipping) όπου διακρίνονται οκτώ (A) αδενίνες σε σειρά στην αλληλουχία *ba0331*.

Στην **Εικόνα 3.50** παρουσιάζεται η υπέρθεση της NodB δομικής ενότητας στην περιοχή του ενεργού κέντρου. Η αναδίπλωση της καταλυτικής δομικής ενότητας των Ba0330 και Ba0331 παρουσιάζει την ίδια τοπολογία $(\alpha/\beta)_6$. Στην Ba0331, η περιοχή MT4 παρουσιάζει μοναδικά χαρακτηριστικά, που έχουν ένα μεικτό φορτισμένο-αρωματικό χαρακτήρα με τις πλευρικές ομάδες των κατάλοιπων Arg350 και Phe352 να δείχνουν προς τη θέση δέσμευσης (**Εικόνα 3.50**). Αυτό το κατάλοιπο φαινυλαλανίνης μαζί με την Met153 της περιοχής MT5' και την Leu 170 δημιούργησαν ένα υδρόφοβο τρίγωνο που οριοθετεί το κάτω τμήμα της κοιλότητας πρόσδεσης. Στην άλλη πλευρά της κοιλότητας πρόσδεσης, η Ba0331 παρουσιάζει και πάλι απόκλιση κατά την εξέταση της συγκεκριμένης περιοχής ενδιαφέροντος. Διακρινόμενη από τις Ba0330 και Bc0361 (η υπέρθεση των δύο δομών δείχνει ταύτιση), φορτισμένα και πολικά κατάλοιπα κατέλαβαν το άνω άνοιγμα της θέσης πρόσδεσης. Στις άλλες δύο φερόμενες απακετυλάσες πολυσακχαρίτη της ομάδας 1, αυτή η πλευρά καταλαμβάνεται από υδρόφοβα ή μικρά αμινοξέα, ενώ στην Ba0331 παρατηρήθηκε ένα εκτεταμένο δίκτυο γεφυρών άλατος και δεσμών υδρογόνου. Όλες αυτές οι διαφορές εμφανίστηκαν γειτονικά προς τα κατάλοιπα συναρμογής του ψευδαργύρου Asp213 και His275, σε μέγιστη απόσταση 6 Å. Συγκεκριμένα, στη θέση 215, η μεθειονίνη που βρέθηκε στα Ba0330 και Bc0361 αντικαταστάθηκε από μια αργινίνη, η οποία σταθεροποιήθηκε με αλληλεπιδράσεις με τα Asp213 και Asp165 του βρόχου $\beta 9$ - $\beta 10$ (Ile στις Ba0330 και Bc0361). Επιπλέον, το Asp165 σχημάτισε μια ισχυρή γέφυρα αλάτων με την Arg251 (Ser και Ala στις Ba0330 και Bc0361, αντίστοιχα) επιβάλλοντας μια μετατόπιση ολόκληρου του βρόχου $\beta 9$ - $\beta 10$ προς την κοιλότητα δέσμευσης, η οποία τοπικά μείωσε το διαθέσιμο όγκο της κοιλότητας πρόσδεσης για την Ba0331. Ο $\beta 9$ - $\beta 10$ βρόχος στην Ba0331 ήταν πιο εκτεταμένος σε σύγκριση με αυτόν της Ba0330 εξαιτίας της εισαγωγής δύο αμινοξέων (Ser164, Asp165) στην αλληλουχία (**Εικόνες 3.49 και 3.50**). Η ομάδα γουανιδινίου της Arg251 αλληλεπιδρά με το Glu245 (Asn στις Ba0330 και Bc0361) και ολόκληρο το δίκτυο τερματίζεται κυκλικά με αλληλεπίδραση μεταξύ του τελευταίου (γλουταμικό οξύ) και Thr242 (Ile στις Ba0330 και Bc0361).


```

Ba0331      A L H I L Q K L K D D T F Q P A A T E F
ba0331      GCTCTTCATATTTTACAAAATTTAAAGGATGATACATCCAGCCTGCAGCTACTGAATTC 720
ba0330      GCATTTTCATGTTTTACAAAAGCTAAAAGACGATACGTTTTAAACCGGTGGCAACAGAATAT 699
          ** ***** ***** ** * * * * * * * * * * * * * * * *
Ba0330      A F H V L Q K L K D D T F K P V A T E Y

Ba0331      L T A N E I D K P N R L S T D D I K Q M
ba0331      CTTACTGCAAATGAGATAGACAAACCAATCGTCTATCAACAGATGATATAAAGCAAATG 780
ba0330      ATGATTGTTAATAATGTTGATGCTGAAGGATCTCTTTCTACATCTGACATAAAGAATG 759
          * * * * * * * * * * * * * * * * * * * * * * * * * * * *
Ba0330      M I V N N V D A E G S L S T S D I K E M

Ba0331      M D S G I F S I Q S H T A N H T M M A H
ba0331      ATGGATTCAAGCATTTTTCTATTCATCAATCTCATACTGCTAATCATACGATGATGGCTCAC 840
ba0330      GTGGATTCCGGCATTTTTCTCTATGCAATCTCATACTGCAACACATGCCGACTTGCCGAAG 819
          ***** ***** ***** ***** * * * * * * * * * *
Ba0330      V D S G I F S M Q S H T A T H A D L P K

Ba0331      S N N Y D E E L R G S K E K I E A L T G
ba0331      TCCAATAATTATGATGAGGAATTACGTGGTTCAAAAGAGAAAATGAGAGCTCTTACTGGT 900
ba0330      ATTACAACACTATGAAGAAGAAATTAAGGAATCGAAAGAAAAGCTAGAAAAATAACAGGT 879
          * * * * * * * * * * * * * * * * * * * * * * * * * * * *
Ba0330      I T N Y E E E L K E S K E K L E K I T G

Ba0331      K K V I A L A Y P V G S Y N D P A V E E
ba0331      AAAAAGGTCATCGCTCTCGCCTATCCAGTAGGTTCTTATAATGATCCAGCTGTTGAAGAA 960
ba0330      AAACCCGTTATCGCAGTTGCTATCCATTCGGTCACGTAGATGATAAAGTCGTTGCAGAA 939
          *** ** ***** * ***** * * * * * * * * * * * * * * * *
Ba0330      K P V I A V A Y P F G H V D D K V V A E

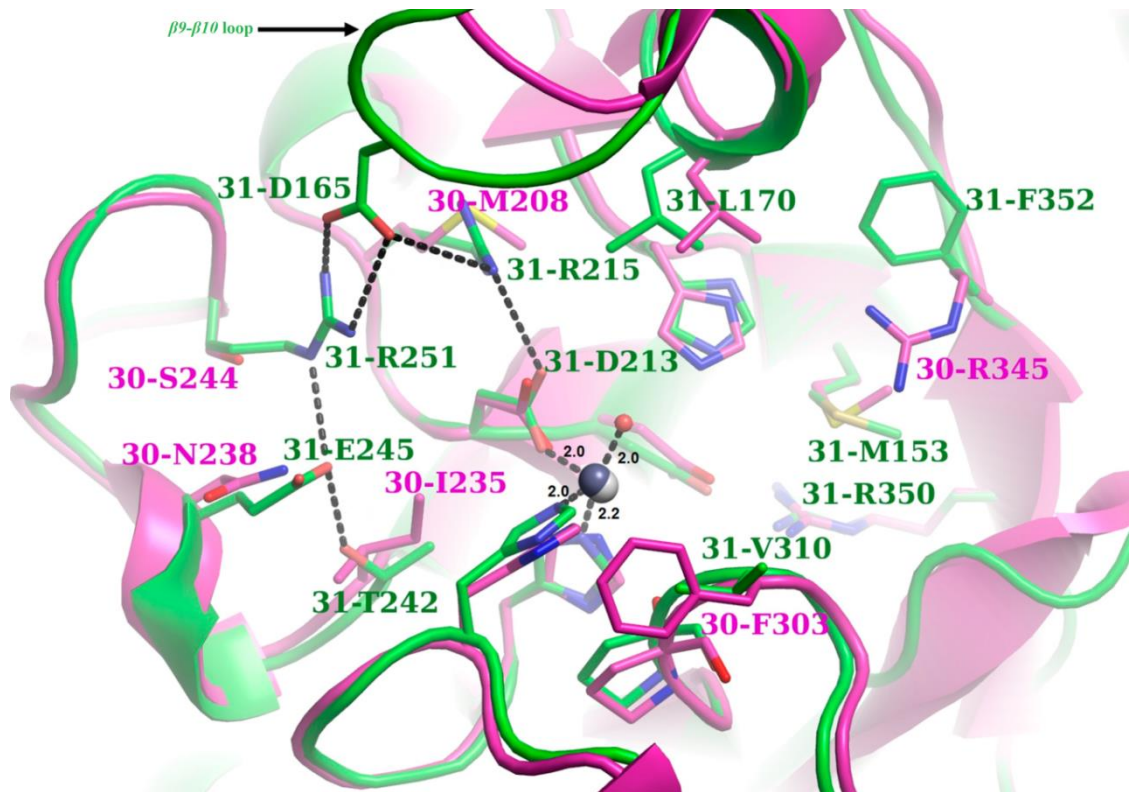
Ba0331      T K K Y Y E F A V T T D H G N H I T K G
ba0331      ACGAAAAAATATTTATGAGTTCGCAGTAACACTACTGATCACGGTAATCATATTTACAAAAGGG 1020
ba0330      ACGAAAAAATATTTATCAATTCGCAACAACAACAACCTGGAAAATTCATTACAAAAGGT 999
          ***** ***** * ***** * * * * * * * * * * * * * * * *
Ba0330      T K K Y Y Q F A T T T K P G K F I T K G

Ba0331      M P N E Q Y L I K R H F V G P N T S M E
ba0331      ATGCCGACGAACAGTACTTAATTAACGCCATTTTGTAGGGCCTAATACATCGATGGAA 1080
ba0330      GAACCCGATGAGTTGTTAAAATGAAGCGTGTTCGCATACACCATACGACGACCGTAGAA 1059
          ** * * * * * * * * * * * * * * * * * * * * * * * * * * * *
Ba0330      E P D E L L K M K R V R I H H T T T V E

Ba0331      K F I S L I K -
ba0331      AAATTCATATCTCTTATAAAGTAA 1104
ba0330      CAATTTGCTTCTTCGATTAATAA 1083
          **** * * * * * * * * * *
Ba0330      Q F A S S I K -

```

Εικόνα 3.49 Ευθυγράμμιση αμινοξικής και νουκλεοτιδικής ακολουθίας της Ba0330 και της Ba0331. Απεικονίζονται οι περιοχές η- (κίτρινο), h- (πράσινο) και c- (μπλε) της τριμερής κατασκευής της αλληλουχίας σήματος των λιποπρωτεϊνών (βλ. §1.6) σύμφωνα με τη βάση δεδομένων DOLOP. Επίσης απεικονίζονται η περιοχή της άρθρωσης ή μεσοδιαστηματική περιοχή της έλικας (γκρι), η δομική ενότητα Fn3 (φούξια) και η δομική ενότητα NodB (κυανό). Οι δύο ενθέσεις κωδικονίων (απεικονίζονται με τα δύο ορθογώνια πλαίσια) στην Ba0331 εντοπίζονται στην ελικοειδή άρθρωση και στον βρόχο β9-β10 της δομικής ενότητας NodB, αντίστοιχα. Οι νουκλεοτιδικές αλληλουχίες αναγράφονται με μπλε χρώμα και πρωτεϊνικές με κόκκινο χρώμα. Η ευθυγράμμιση πραγματοποιήθηκε χρησιμοποιώντας το πρόγραμμα ClustalW.



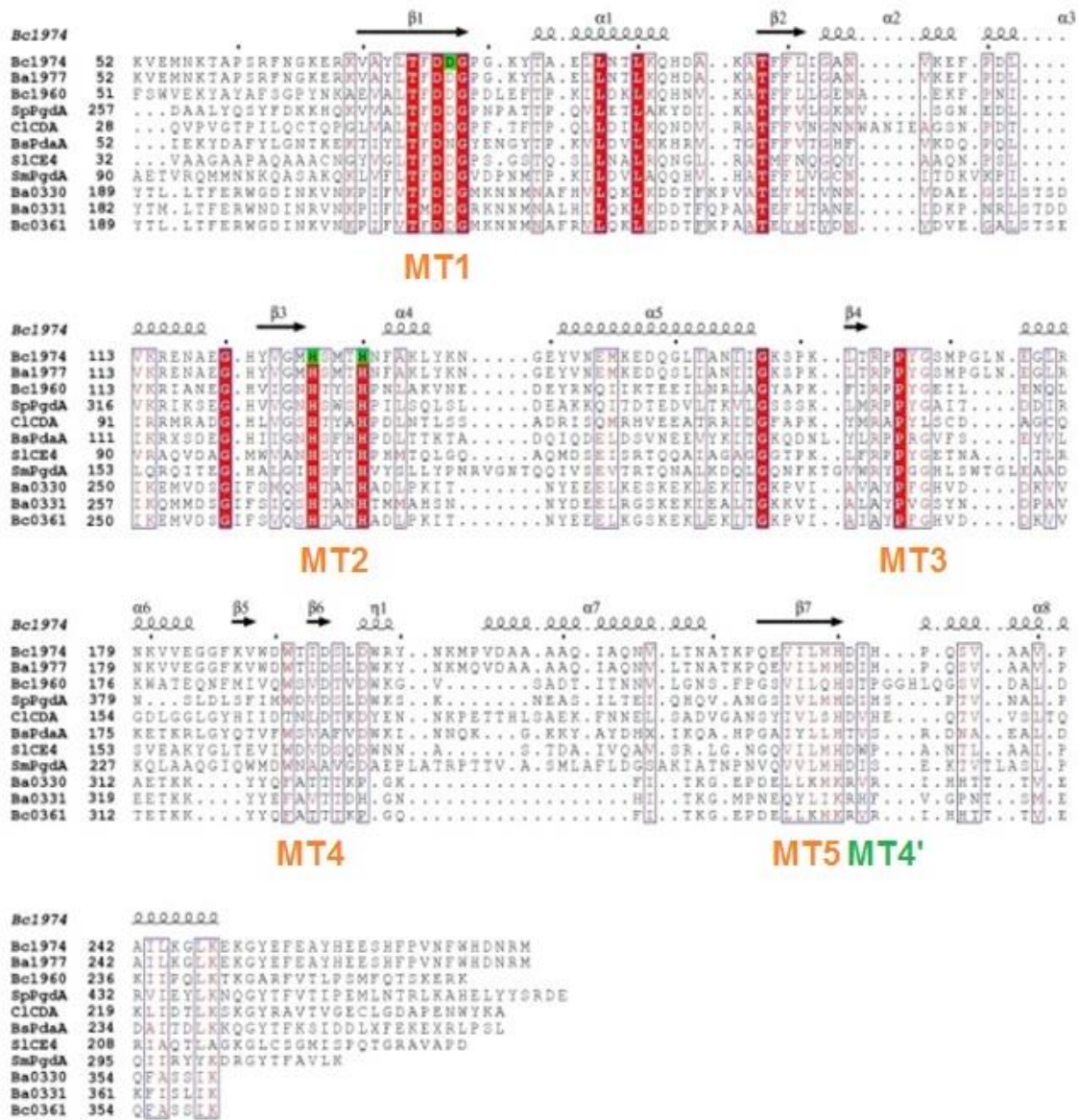
Εικόνα 3.50 Υπέρθηση μέρους της δομικής ενότητας NodB που περιέχει τις θέσεις πρόσδεσης των Ba0331 και Ba0330. Η Ba0331 εμφανίζεται με πράσινο χρώμα και η Ba0330 εμφανίζεται με φούξια. Η εικόνα έχει κέντρο το ιόν ψευδαργύρου και επομένως εμφανίζεται μόνο η κοιλότητα της θέσης δέσμευσης των δύο πρωτεϊνών. Τα κατάλοιπα που είτε ανήκουν άμεσα στην περιοχή ενός εκ των MT1-MT5 είτε βρίσκονται σε απόσταση 5 Å από ένα από αυτά τα μοτίβα παρουσιάζονται με απεικόνιση ράβδου. Τα οξικά μόρια απομακρύνθηκαν για λόγους ευκρίνειας. Με τη μαύρη σφαίρα αντιπροσωπεύεται το ιόν ψευδαργύρου στο ενεργό κέντρο της Ba0331 και η γκριζα σφαίρα δείχνει το ιόν ψευδαργύρου της Ba0330. Η σφαίρα συναρμογής του ψευδαργύρου της Ba0331 παρουσιάζεται με μαύρες διακεκομμένες γραμμές (οι αποστάσεις δίδονται σε Ångströms) καθώς και το μοναδικό δίκτυο αλληλεπιδράσεων της. Ένα ισχυρό δίκτυο αλληλεπιδράσεων (στην αριστερή πλευρά της εικόνας), το οποίο αντιστοιχεί στο άνω χείλος της κοιλότητας δέσμευσης, παρατηρήθηκε στην Ba0331. Σε αντίθεση με τις απακετυλάσες Ba0330 και Bc0361 όπου η συγκεκριμένη πλευρά καταλαμβάνεται από υδρόφοβα ή μικρά αμινοξέα, στην Ba0331 εντοπίστηκαν πολλά φορτισμένα ή υδρόφιλα αμινοξέα. Οι ισχυρές αλληλεπιδράσεις του D165 με θετικά φορτισμένα αμινοξέα σταθεροποιούν το βρόχο $\beta 9$ - $\beta 10$ σε μια κλειστή διαμόρφωση, μειώνοντας έτσι το διαθέσιμο όγκο της κοιλότητας δέσμευσης. Στην δεξιά πλευρά του σχήματος, που αντιστοιχεί στην κάτω πλευρά της κοιλότητας δέσμευσης, η παρουσία των M153, L170 και F352 χαρακτηρίζει την περιοχή ως υδρόφιλη.

4. Συζήτηση

4.1. Χαρακτηριστικά της δομής και της αλληλουχίας CE4.

4.1.1 Καθοριστικά δομικά μοτίβα. Μια συντηρημένη αμινοξική τριάδα στη θέση δέσμευσης του μετάλλου

Η στοίχιση της αλληλουχίας της Bc1974 με ομόλογα ένζυμα παρουσιάζεται στην **Εικόνα 4.1**. Τα αυστηρά συντηρημένα κατάλοιπα είναι λευκά σε κόκκινο φόντο ενώ παρόμοια κατάλοιπα είναι κόκκινα εντός πλαισίου. Εντοπίζονται τα αμινοξέα στην καταλυτική περιοχή και στην περιοχή πρόσδεσης του μετάλλου που είναι συντηρημένα και επισημαίνεται η κατανομή τους στα πέντε καταλυτικά μοτίβα (MT1–MT5) που χαρακτηρίζουν ενζυμικά και δομικά τις CE4 εστεράσες (Blair *et al.* 2005). Η εξέταση της οικογένειας των ομολόγων πρωτεϊνών από τις οποίες είναι γνωστές κάποιες τρισδιάστατες δομές κάνει σαφές ότι περιοχές και κατάλοιπα κρίσιμα για το μηχανισμό λειτουργίας της πρωτεΐνης έχουν συντηρηθεί σε μεγαλύτερο βαθμό από ότι άλλα κατάλοιπα. Τα MT1 και MT2 εμπλέκονται άμεσα με τη συναρμογή του μετάλλου στο ενεργό κέντρο. Το μοτίβο MT1 συνεισφέρει το ασπαρτικό οξύ και το MT2 συνεισφέρει τις δύο ιστιδίνες. Η ρύθμιση της ενζυμικής δράσης, η μεταβολή στην εξειδίκευση του υποστρώματος και η προτίμηση του είδους του μετάλλου στη θέση δέσμευσης των ενζύμων της CE4 μπορούν να πραγματοποιηθούν μέσω τροποποιήσεων στα MT1-MT5 (Arnaouteli *et al.* 2015).



Εικόνα 4.1 Ευθυγράμμιση αλληλουχιών της δομικής ενότητας NodB των απακετυλασών πεπτιδογλυκάνης Bc1974, Ba1977, των φερόμενων απακετυλασών πολυσακχαριτών Ba0330, Bc0361, Ba0331 με αντιπροσωπευτικά μέλη της οικογένειας CE4 περιλαμβανομένων των Bc1960, των απακετυλασών πεπτιδογλυκάνης SpPdgA, BsPdaA, της απακετυλάσης της χιτίνης C1CDA, της εστεράσης της ακετυλοξυλάνης S1CE4. Τα στοιχεία δευτεροταγούς δομής της Bc1974 εμφανίζονται στην κορυφή της ευθυγράμμισης, ενώ τα καθοριστικά μοτίβα (MT1 έως MT5) φαίνονται στο κάτω μέρος (Το MT5' των Ba0330-Bc0361 και Ba0331 που βρίσκεται στην αρχή δεν παρουσιάζεται). Τα αμινοξέα με πράσινο φόντο στην Bc1974 δείχνουν την αντίστοιχη στήλη που βρίσκεται η θέση καθενός από τα αμινοξέα της τριάδας συναρμογής του μετάλλου (Asp, His, His). Η ευθυγράμμιση πραγματοποιήθηκε με το T-coffee (Notredame *et al.* 2000) και σχεδιάστηκε με το ESPRIPT (Gouet *et al.* 2003).

4.1.2 Οργάνωση των δομικών ενοτήτων

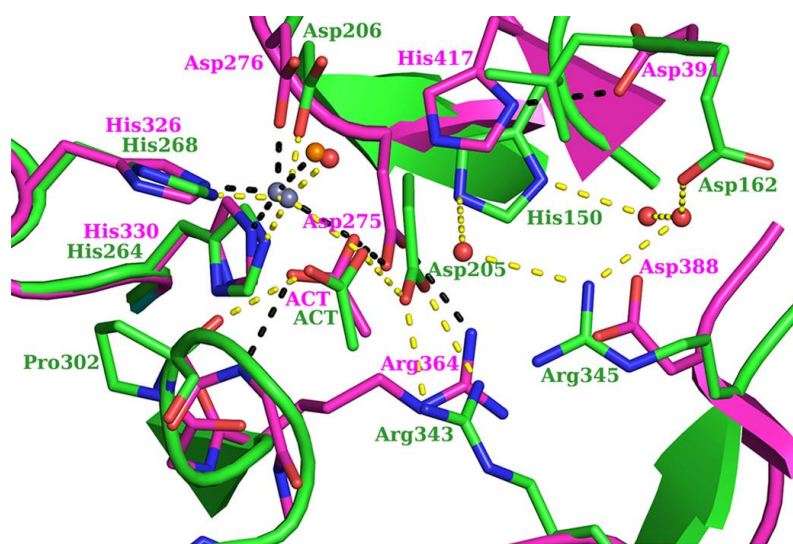
Εκτός από την κοινή καταλυτική δομική ενότητα NodB, τα μέλη CE4 ενδέχεται να έχουν πρόσθετες δομικές ενότητες με μοναδικά χαρακτηριστικά για την κάθε πρωτεΐνη, όπως α-έλικες (Deng *et al.* 2009), τοπολογίες α/β (Blair *et al.* 2005), (4+3) β -σάντουιτς (Fadoulglou *et al.* 2013), και CBM (carbohydrate-binding modules, ενότητες πρόσδεσης υδατανθράκων, Andrés *et al.* 2014). Η δομή των Ba0330 και Ba0331 έχει μια μοναδική αρχιτεκτονική τριών δομικών ενοτήτων. Μια αμινοτελική δομική ενότητα α -έλικας που ίσως προσδίδει μεσοδιάστημα και ευλυγισία στην πρωτεΐνη και ίσως διαδραματίζει κάποιο ρόλο στην αναγνώριση. Στη μέση υπάρχει μια Fn3 (Arnaouteli *et al.* 2015, Andreou *et al.* 2019) και η καρβοξυτελική καταλυτική δομική ενότητα NodB που υιοθετεί μια $(\alpha/\beta)_6$ τοπολογία. Η Bc1974 δεν έχει μεσαία δομική ενότητα και η καταλυτική NodB περιοχή στο καρβοξυτελικό άκρο υιοθετεί $(\alpha/\beta)_8$ τοπολογία.

4.1.3 Συγκρίσεις δομών PDA υψηλής διακριτικότητας

Στην υπέρθεση των θέσεων δέσμευσης της SpPgdA και της Ba0330 παρουσιάζει ενδιαφέρον το κατάλοιπο Arg345 στην Ba0330, καθώς η αργινίνη αυτή βρίσκεται στη θέση ενός συντηρημένου, μεταξύ των μελών της οικογένειας CE4, ασπαρτικού οξέος Asp388 στη SpPgdA και πιστεύεται ότι είναι υπεύθυνη για τη ρύθμιση της τιμής pKa της καταλυτικής ιστιδίνης (His150) (Εικόνα 4.2). Έχει αναφερθεί προηγουμένως ότι η μετάλλαξη D391N της SpPgdA απενεργοποίησε εντελώς το ένζυμο (Blair *et al.* 2005). Μια παρόμοια παρατήρηση έχει αναφερθεί για την PgaB από το *E. coli* όπου ένα μόριο ύδατος στη θέση του συντηρημένου ασπαρτικού οξέος (Little *et al.* 2012), ενδεχομένως εξηγεί τη μειωμένη απόδοση του ενζύμου. Από την υπέρθεση των ενεργών κέντρων των SpPgdA και Ba0330 (Εικόνα 4.2), είναι προφανές ότι η συντηρημένη ιστιδίνη με ρόλο τροποποιητή του pKa του καταλυτικού Asp391 στη SpPgdA αντικαθίσταται από μια Leu στην Ba0330. Ωστόσο, το ρόλο του ρυθμιστή της κατάστασης πρωτονίωσης στο καταλυτικό κέντρο της Ba0330 μπορεί να παίζει η κοντινή Arg345, η οποία αλληλεπιδρά μέσω δομικών μορίων διαλύτη με τα κατάλοιπα His150 και Asp162. Οι δομές έδειξαν την ύπαρξη ενός μορίου οξικού οξέος προσδεμένο στο ιόν ψευδαργύρου σε όλες τις πρωτεΐνες χωρίς υπόστρωμα. Το μόριο του οξικού οξέος αλληλεπιδρά εκτεταμένα με κατάλοιπα από τη θέση πρόσδεσης. Ένα από τα οξυγόνα του αλληλεπιδρά με το υψηλά συντηρημένο Asp205, το οποίο επίσης αλληλεπιδρά

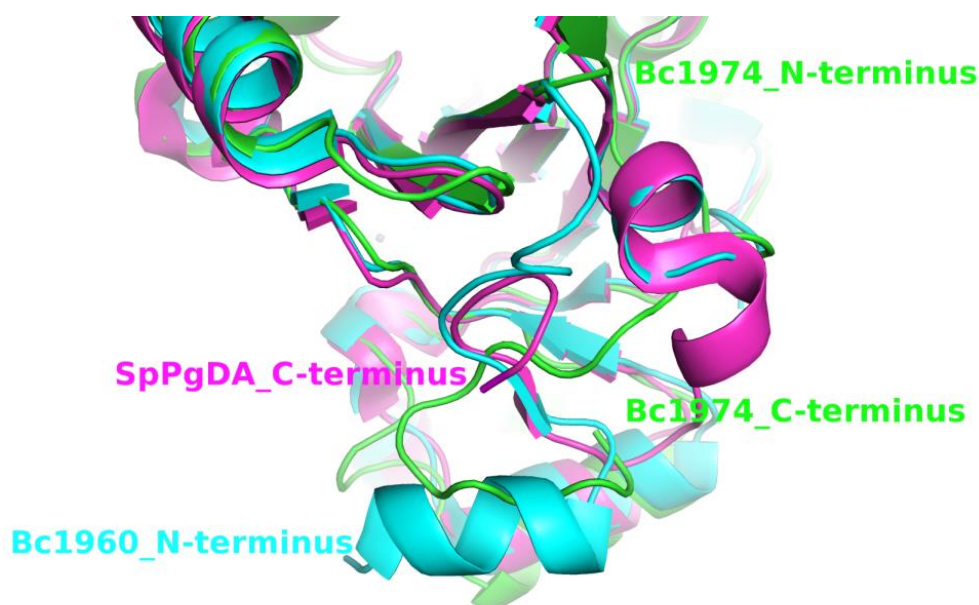
ηλεκτροστατικά με την Arg343 (που υπάρχει μόνο στις απακετυλάσες που περιέχουν τη δομική ενότητα Fn3), της οποίας η ομάδα γουανιδίνης επικαλύπτεται χωρικά με την αντίστοιχη πλευρική ομάδα της αργινίνης (Arg363) στην πρωτεΐνη SpPgdA (Εικόνα 4.2). Η αλληλεπίδραση του ασπαρτικού οξέος (Asp205) με την αργινίνη (Arg343) μάλλον είναι απαραίτητη για την κατάλυση. Επιπλέον, το ίδιο άτομο οξυγόνου φαίνεται να αλληλεπιδρά με τη δομικά συντηρημένη His150, που σε αντίθεση με άλλες απακετυλάσες, δεν αλληλεπιδρά άμεσα με κοντινά αμινοξικά κατάλοιπα.

Η θέση κατάλυσης διαφέρει από αυτή που απαντάται συνήθως στις PDA στα ακόλουθα σημεία: α) Η Arg που αλληλεπιδρά με το καταλυτικό Asp βρίσκεται συνήθως στην αρχή του μοτίβου 3 (MT3). Εδώ, η αντίστοιχη αργινίνη της Arg343, βρίσκεται στον τελευταίο κλώνο. β) Η καταλυτική ιστιδίνη (His150), η οποία είναι συνήθως εντός του μοτίβου 5 (MT5), στην Ba0330 βρίσκεται σε μη συντηρημένη θέση της αλληλουχίας. γ) Η Ba0330 στερείται του ασπαρτικού οξέος, το οποίο συνήθως παρέχεται από το μοτίβο 4, και πιστεύεται ότι είναι υπεύθυνο για τη ρύθμιση της τιμής pKa της καταλυτικής ιστιδίνης. Αντ' αυτού, η δομή της Ba0330 αποκαλύπτει ένα κατάλοιπο αργινίνης (Arg345) σε εγγύτητα με την His150 και πιθανή αλληλεπίδραση μεταξύ τους μέσω ενός μορίου ύδατος. Επιπλέον, το Asp162 μπορεί να αλληλεπιδράσει με την καταλυτική His150 μέσω γέφυρας δύο μορίων ύδατος.



Εικόνα 4.2 Υπέρθωση των θέσεων δέσμευσης της SpPgdA με φούξια (PDB ID: 2C1G) και της Ba0330 με πράσινο χρώμα. Τα κατάλοιπα που εμπλέκονται σε αλληλεπιδράσεις παρουσιάζονται με αναπαράσταση ράβδων. Τα μόρια ύδατος στην Ba0330 παρουσιάζονται με κόκκινες σφαίρες ενώ στην SpPgdA με πορτοκαλί σφαίρες. Οι αλληλεπιδράσεις δίδονται με κίτρινες και μαύρες διακεκομμένες γραμμές για τις δομές των Ba0330 και SpPgdA, αντιστοίχως. Τα κατάλοιπα της Ba0330 επιγράφονται με πράσινο χρώμα ενώ εκείνα της SpPgdA με φούξια.

Η υπέρθεση της Bc1974 με την ομόλογη Bc1960 (39% ταυτότητα) εμφανίζει r.m.s.d. 1,32 Å για τα άτομα τις κύριας αλυσίδας 179 καταλοίπων, ενώ τα δύο άκρα αμινοτελικό και καρβοξυτελικό παρουσιάζουν αξιοσημείωτη διαφοροποίηση (**Εικόνα 4.3**). Ο χώρος στον οποίο βρίσκεται η πρώτη κοντή έλικα της Bc1960, στην δομή της Bc1974 καταλαμβάνεται από ένα μακρύ καρβοξυτελικό βρόχο, μια μοναδική διαμόρφωση που εμφανίζεται σε αυτό το ένζυμο, η οποία σταθεροποιείται με πολυάριθμες ενδομοριακές αλληλεπιδράσεις με παρακείμενα κατάλοιπα. Η παρατήρηση μεγάλων αποκλίσεων μεταξύ αυτών των περιοχών ομολόγων ενζύμων παρά τη συνολικά διατηρημένη $(\alpha/\beta)_8$ τοπολογία δεν είναι ασυνήθιστη. Τελικά, παρά τη σημαντική διαφοροποίηση που παρατηρείται στις αμινοτελικές και καρβοξυτελικές περιοχές της Bc1974 σε σχέση με τις αντίστοιχες περιοχές των ομολόγων Bc1960 και SpPgdA, το ενεργό κέντρο της Bc1974 είναι χαρακτηριστικό της PDAs.



Εικόνα 4.3 Υπέρθωση της NodB ομόλογης δομικής ενότητας των Bc1974 (PDB ID: 5N1J, πράσινο), Bc1960 (PDB ID: 4L1G, κυανό) και SpPgdA (PDB ID: 2C1G, φούξια). Η εικόνα επικεντρώνεται στην καρβοξυτελική περιοχή της Bc1974, δείχνοντας τις σημαντικές δομικές διαφορές στην συγκεκριμένη περιοχή μεταξύ των ενζύμων και τη μοναδική τροχιά που ακολουθείται από τα δεκαοκτώ καρβοξυτελικά κατάλοιπα της Bc1974. Αξιοσημείωτα, η καρβοξυτελική α -έλικα της δομής της SpPgdA αν και ακολουθεί το ίχνος της αντίστοιχης περιοχής της Bc1960, στη συνέχεια σχηματίζει τον όγδοο β -κλώνο στην SpPgdA στην περιοχή όπου στην Bc1960 βρίσκεται ο αμινοτελικός βρόχος.

4.2 Εξελικτικές σχέσεις στην οικογένεια γονιδίων απακετυλάσης πολυσακχαρίτη του *Bacillus anthracis* και του *Bacillus cereus*

4.2.1 Φυλογενετική ανάλυση των απακετυλασών πολυσακχαρίτη με βάση τη νουκλεοτιδική και πρωτεϊνική αλληλουχία της καταλυτικής δομικής ενότητας NodB.

Για τη διερεύνηση φυλογενετικών σχέσεων στις οικογένειες γονιδίων απακετυλάσης πολυσακχαρίτη του *B. anthracis* Ames (12 παράλογα γονίδια) και του *B. cereus* ATCC14579 (11 παράλογα) χρησιμοποιήθηκαν οι αμινοξικές και νουκλεοτιδικές αλληλουχίες που απαρτίζουν χωριστά τις δομικές ενότητες των PDAs.

Η φυλογενετική ανάλυση μεταξύ των PDAs του *B. anthracis* και του *B. cereus* της καταλυτικής περιοχής NodB, αποκάλυψε την κατανομή σε τρεις ομάδες (υποοικογένειες) (1, 2 και 3), που παρουσιάζονται συντηρημένες μεταξύ των διαφόρων ειδών. Οι διαδοχικές διπλοποιήσεις οδήγησαν σε ομάδες γονιδίων (υποοικογένειες), τα οποία εμφανίζουν μεταξύ τους πολύ μεγαλύτερη ομολογία σε σύγκριση με οποιοδήποτε άλλο γονίδιο της ίδιας οικογένειας. Ερώτημα αποτελεί η διατήρηση της λειτουργικής ομοιότητας των νεοεμφανιζόμενων γονιδίων εντός των ομάδων. Τα τελευταία χρόνια μελέτες έχουν περιγράψει τη νεολειτουργικοποίηση, την υπολειτουργικοποίηση και την απολειτουργικοποίηση των πρωτεϊνών μέσω της σημαντικής διαδικασίας της διπλοποίησης γονιδίου (Hughes 1994, Rastogi & Liberles 2005, Konrad *et al.* 2011).

Σε ένα αποδεκτό εξελικτικό δενδρόγραμμα των γονιδίων των απακετυλασών πολυσακχαρίτη του *B. anthracis* οι κόμβοι δηλώνουν τα γεγονότα γονιδιακής διπλοποίησης. Όμως θα πρέπει να ληφθεί υπόψη ότι οι κόμβοι των δενδρογραμμμάτων παράλογων μάλλον δηλώνουν τα τελευταία γεγονότα γονιδιακής ανάμειξης, σύμφωνα με το φαινόμενο της γονιδιακής σύγκλισης (gene convergence), και όχι το ίδιο το γεγονός της διπλοποίησης του γονιδίου (Page & Holmes, 1998). Η γενετική απόκλιση των αλληλουχιών της Ομάδας 1 από τις άλλες γίνεται φανερή τόσο από τα φυλογενετικά δέντρα όσο και από τον προσδιορισμό των δομών. Αν και διατηρήθηκε ο ίδιος μηχανισμός ενεργοποίησης μετάλλου, η δομική ενότητα NodB μεταξύ των τριών ομάδων τροποποιήθηκε σημαντικά. Αυτό επιτεύχθηκε με τροποποιήσεις στα κατάλοιπα σχηματισμού της θέσης πρόσδεσης όπως αυτή διαμορφώνεται από συντηρημένα μοτίβα της αλληλουχίας (MT), προκαλώντας αφενώς την ομαδοποίηση των ομολόγων αλληλουχιών στο εξελικτικό δένδρο των δύο ειδών, υποδεικνύοντας αφετέρου τη λειτουργική διαφοροποίησή τους. Το αποτέλεσμα αυτό ισχυροποιείται επίσης από τη διατήρηση της ομαδοποίησης των ομολόγων στην ανάλυση που επεκτείνεται και σε άλλα είδη.

Οι πολυγονιδιακές οικογένειες (multigene families) αποτελούν έναν απλό τρόπο και υψηλά προτιμητέο για τη δημιουργία γενετικής και λειτουργικής ευελιξίας. Όμως η σύνθετη διαδικασία που διαμορφώνει την εξέλιξή τους μεταξύ άλλων μπορεί να σημαίνει ότι αυτές συχνά αποτελούνται από μωσαϊκά αλληλουχιών καθένα από τα οποία έχει μια διαφορετική φυλογενετική ιστορία, παρά από αυστηρώς ομόλογα γονίδια που έχουν σταδιακά διαφοροποιηθεί κατά τη διάρκεια του χρόνου (Page & Holmes, 1998).

Μόνο οι πρωτεΐνες της Ομάδας 1 και όχι των Ομάδων 2 και 3 φαίνεται να περιέχουν μία δομική ενότητα τύπου Fn3, εκτός από την καταλυτική περιοχή. Εντός της Ομάδας 1 ένα γεγονός διπλοποίησης γονιδίου είχε ως αποτέλεσμα την Ba0331. Εν τούτοις, το διπλοποιημένο παράλογο της Ba0331 απουσίαζε στα περισσότερα από τα γονιδιώματα των στελεχών της ομάδας του *B. cereus* που μελετήθηκαν. Η υψηλή ομολογία μεταξύ της Ba0330 και Ba0331, παραπέμπει μάλλον σε μικρό χρόνο απόκλισης. Η γνώση της δομής και της οργάνωσης στο χρωμόσωμα των γονιδίων των απακετυλασών πολυσακχαρίτη μπορεί να επιτρέψει τη σκιαγράφηση της εξελικτικής τους ιστορίας. Παρότι η παρουσία του αντιγράφου της Ba0331 σε περιορισμένα στελέχη καταδεικνύει ότι το γεγονός διπλοποίησης που τα δημιούργησε είναι μάλλον πρόσφατο, είναι σαφές ότι εκτίμηση του μοριακού ρολογιού δεν μπορεί να γίνει από την παρούσα μελέτη, αφού θα ήταν αναγκαίο να εξευρεθούν πολλαπλά σημεία χρονολόγησης. Έχει συναχθεί προηγουμένως από μελέτες, με τη μέθοδο τυποποίησης με βάση την αλληλουχία πολλαπλών γενετικών τόπων (MLST-Multilocus Sequence typing) και ανάλυσης διαγονιδιακής αλληλουχίας ριβοσωμικού DNA, ότι ο *B. anthracis* έχει αποκλίνει πρόσφατα από σχετικούς κλάδους στελεχών *B. cereus* και *B. thuringiensis* (Cherif *et al.* 2003, Helgason *et al.* 2004, Ko *et al.* 2004, Priest *et al.* 2004).

Η φυλογένεια της Ομάδας 2 αποκάλυψε τη στενή σχέση μεταξύ της Ba1977 και της Ba2944 στο *B. anthracis* και στην οικογένεια του *B. cereus* τα αντίστοιχα ορθόλογα (Bc1974 και Bc2929 Με βάση την ομολογία Ba1977-Ba2944 (ταυτότητα 75,7%), Bc1974-Bc2929 (ταυτότητα 76,5%) και την πρόβλεψη ότι οι Ba1977 και Bc1974 είναι πρωτεΐνες αγκυροβολημένες στη μεμβράνη χωρίς θέση αποκοπής (CS), ενώ οι Ba2944 και Bc2929 είναι εξωκυτταρικές πρωτεΐνες με μια θέση αποκοπής, μπορεί να συναχθεί ότι πρόκειται για πρωτεΐνες που έχουν την ίδια λειτουργία και εξειδίκευση υποστρώματος σε διαφορετικούς τόπους (Balomenou *et al.* 2015). Όμοια, όπως αναφέρθηκε για την παρουσία των ορθόλογων της Ba0331, με βάση ότι το διπλοποιημένο παράλογο της Ba2944 απουσίαζε στα περισσότερα από τα γονιδιώματα των στελεχών της ομάδας του *B. cereus* που εξετάστηκαν, μπορεί να υποθεθεί ότι ένα πρόσφατο γεγονός διπλοποίησης γονιδίου, είχε ως αποτέλεσμα την παρουσία των ορθόλογων της Ba2944.

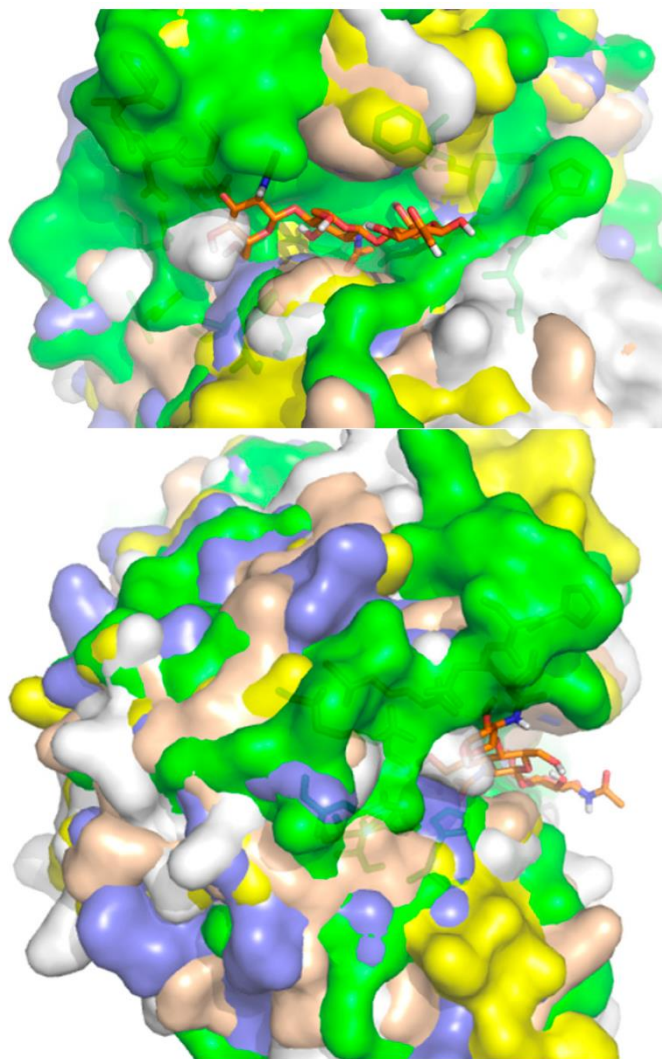
Συνολικά το φυλογενετικό δέντρο της καταλυτικής ενότητας NodB αντικατοπτρίζει τα γεγονότα διπλοποίησης (στους κόμβους που οδηγούν στα ορθόλογα) χωρίς να παραβλέπεται ο κίνδυνος λάθους συμπεράσματος στην περίπτωση που δεν εμπεριέχονται όλα τα πολλαπλά γονίδια στο δέντρο και περισσότερα του ενός είδους ώστε να αποκαλυφθούν (οι κόμβοι) τα γεγονότα διπλοποίησης.

4.2.1.1 Ανίχνευση της διάδοσης της δομικής ενότητας ινωδονεκτίνης τύπου III

Το φυλογενετικό δέντρο με τις αλληλουχίες που αποτελούν τις καταλυτικές μονάδες φανερώνει τον αρχικό εντοπισμό της δομικής ενότητας Fn3. Στις Ομάδες 2 και 3 ανήκουν οι ομόλογες πρωτεΐνες που στερούνται της δομικής ενότητας Fn3. Στα φυλογενετικά δέντρα που κατασκευάστηκαν από αυτά τα ομόλογα τμήματα ξεχωρίζει η Ομάδα 1 εντός της οποίας οι μονάδες Fn3 για πρώτη φορά εμφανίζονται ως μια ομάδα.

4.2.2 Εξέλιξη των θέσεων δέσμευσης των φερόμενων απακετυλασών πολυσακχαρίτη μετά από τα γεγονότα διπλοποίησης.

Οι υπάρχουσες κρυσταλλογραφικά προσδιορισμένες PDAs των *B. anthracis* και *B. cereus* χρησιμοποιήθηκαν για τη μελέτη της κοιλότητας πρόσδεσης. Αυτές είναι οι Ba0424 (Oberbarnscheidt *et al.* 2007), Bc0361 (Fadouloglou *et al.* 2013), Ba0150 (Strunk *et al.* 2014), Bc1960 (Fadouloglou *et al.* 2017) μαζί με τις τρεις από την παρούσα διατριβή Ba0330, Ba0331 και Bc1974. Στην **εικόνα 4.4** παρουσιάζεται η υπέρθεση τους, όπου κάθε δομή αναπαριστάται με διαφορετικό χρώμα, και αναδεικνύονται οι διαφορές της επιφάνειας μεταξύ των PDAs. Παρά όμως τις τοπικές διαφοροποιήσεις, υπάρχει διατήρηση της κοινής τοπολογίας αναδίπλωσης των γενικών δομικών χαρακτηριστικών, και γι' αυτό στην εικόνα της υπέρθεσης η θέση της κοιλότητας δέσμευσης διατηρείται, όπως φαίνεται από τη θέση ενός υποτιθέμενου υποστρώματος εντός της κοιλότητας.



Εικόνα 4.4. Αναπαράσταση της επιφάνειας της καταλυτικής δομικής ενότητας NodB με υπέρθεση των κρυσταλλογραφικά προσδιορισμένων PDAs των *B. anthracis* και *B. cereus* Bc1974 (κίτρινο), Ba0331 (μπλε), Ba0330 (πράσινο), Bc1960 (μπεζ), Ba0424 (λευκό), σε δύο κάθετες όψεις ώστε να επισημανθούν ομοιότητες του σχήματος και οι διαφορές μεταξύ των PDA. Το μοντέλο ενός τρισακχαρίτη (GlcNAc)₃ στην Bc1974 (Giastas *et al.* 2018) απεικονίζεται με αναπαράσταση ράβδων (τα Ca χρωματίζονται πορτοκαλί), και προσδιορίζει τη θέση δέσμευσης που σχηματίζεται (οριζόντια). Η Bc0361 αντιπροσωπεύεται από την Ba0330 καθώς τα σχήματά τους ταυτίζονται. Η Ba0150 δε συμπεριλήφθηκε σε αυτήν την αναπαράσταση διότι υπερκαλύπτει όλες τις άλλες, λόγω του πολύ κλειστού σχήματος της κοιλότητας της.

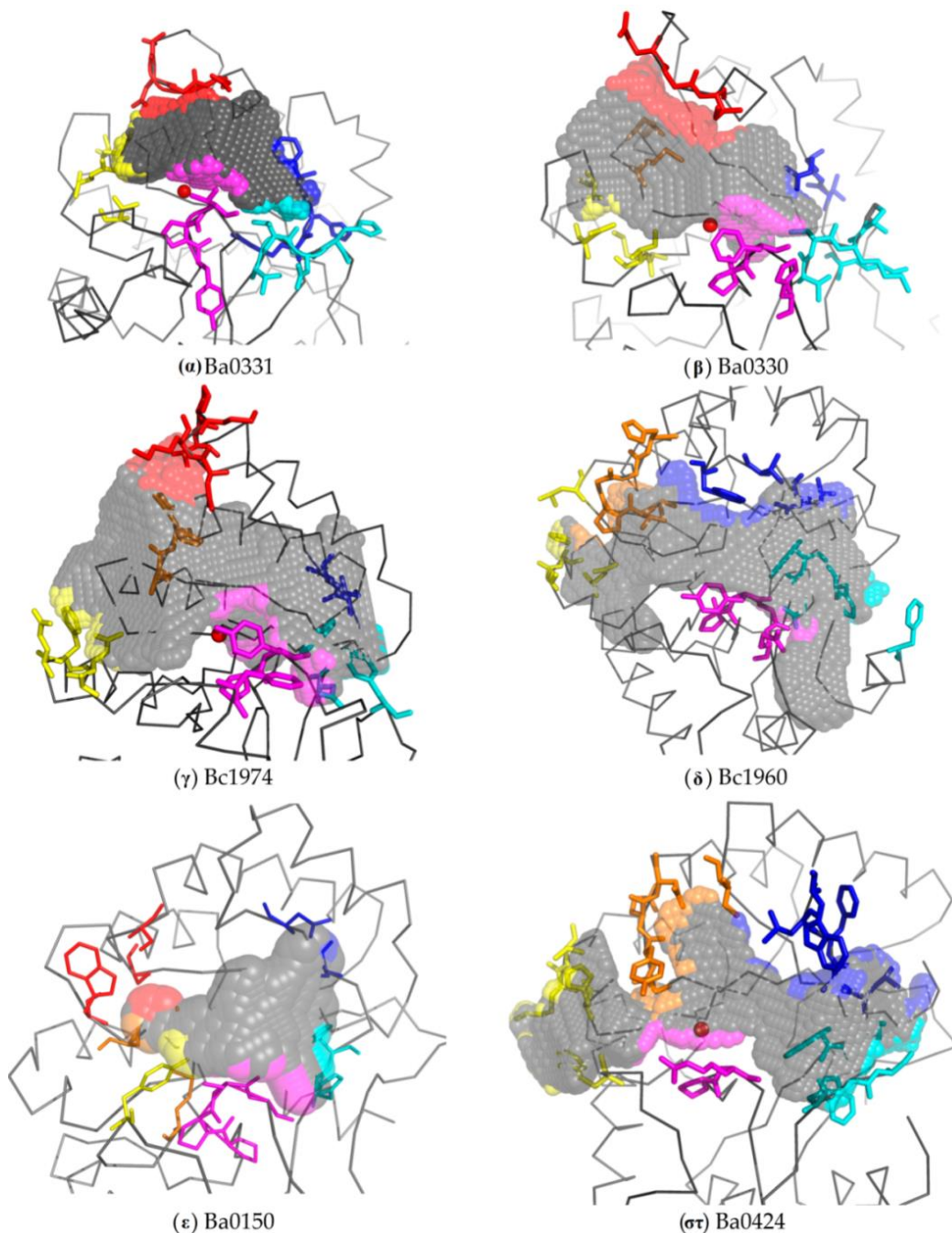
4.2.2.1 Όγκος της κοιλότητας δέσμευσης των απακετυλασών

Η εξέταση του όγκου της κοιλότητας δέσμευσης των απακετυλασών δεν έχει αναφερθεί στην βιβλιογραφία, και στην παρούσα διατριβή απεικονίζεται με μια μορφή χρωματικής κωδικοποίησης (**Εικόνα 4.5**). Σε αντίθεση με τα σχετικά συντηρημένα μοτίβα MT1 και MT2, τα οποία συμμετέχουν στη συναρμογή του μετάλλου, τα άλλα τρία μοτίβα MT3, MT4 και MT5 καθώς και άλλοι περιφερειακοί βρόχοι που πλαισιώνουν την κοιλότητα δέσμευσης και προσδίδουν στην εξειδίκευση του υποστρώματος, παρουσιάζουν σημαντικές διαφορές. Η θέση αυτών επί της πρωτοταγούς δομής των πρωτεϊνών που μελετήθηκαν στην παρούσα διατριβή δεν είναι υψηλά συντηρημένη και ακόμη μεγαλύτερες διαφορές προκύπτουν από τη σύγκριση των φυσικοχημικών ιδιοτήτων των αμινοξέων τους. Τελικά παρατηρήθηκε σημαντική ποικιλότητα του όγκου και του σχήματος των κοιλοτήτων πρόσδεσης (**Πίνακας 4.1, Εικόνα 4.5**). Η παρατηρούμενη ποικιλότητα δεν είναι ειδική εντός συγκεκριμένων ομάδων, καθώς είναι παρούσα και μεταξύ των PDA της Ομάδας 2 και 3, που περιείχαν μόνο την καταλυτική NodB δομική ενότητα αλλά και εντός της Ομάδας 1 στην οποία ανήκουν οι φερόμενες απακετυλάσες πολυσακχαρίτη που περιέχουν και τη δομική ενότητα τύπου Fn3.

Μια χωρική συντήρηση του μοτίβου MT3 (που παρουσιάζεται με φούξια) αποκαλύφθηκε κοντά στην καταλυτική θέση σε όλες τις διαθέσιμες δομές, παρά τις διαφορές στα κατάλοιπα που περιέχονται. Οι μικρότερες διαφορές παρατηρήθηκαν μεταξύ των χωρικών διευθετήσεων των μοτίβων MT4 (που παρουσιάζονται με μπλε χρώμα), παρά τις μεγάλες διακυμάνσεις στα αμινοξέα που τα αποτελούν. Μεγαλύτερες αποκλίσεις παρατηρήθηκαν στους βρόχους μετά-MT1 και προ-MT2 (που φαίνονται με κίτρινο χρώμα), οι οποίοι επίσης αποτελούνται από μεγάλες διαφοροποιήσεις στην αλληλουχία τους. Αυτοί οι βρόχοι έθεσαν τα άνω σύνορα της κοιλότητας δέσμευσης και επιπλέον διαμόρφωσαν τη χωρική διάταξη των βρόχων μετά-MT4 ή μετά-MT5 (που φαίνονται με κόκκινο) είτε αλληλεπιδρώντας μεταξύ τους όταν κυριαρχούσαν στην αλληλουχία τους τα πολικά κατάλοιπα (σχηματίζοντας ρητές πολικές αλληλεπιδράσεις), ή αγνοώντας ο ένας τον άλλον όταν επικρατούσαν κατάλοιπα διαφορετικών φυσικοχημικών ιδιοτήτων. Θα πρέπει να σημειωθεί ότι ο τελευταίος βρόχος υπήρχε μόνο στις λιποπρωτεΐνες της Ομάδας 1.

Το κατώτερο άνοιγμα της κοιλότητας δέσμευσης (χρωματίζεται με κυανό), το οποίο σχηματίζεται από αμινοξικά κατάλοιπα κοντά στο καρβοξυτελικό άκρο στις PDAs που περιέχουν μόνο τη δομική ενότητα NodB (Ομάδα 2 και Ομάδα 3) ενώ στις πρωτεΐνες που περιέχουν και τη δομική ενότητα Fn3 (της Ομάδας 1), αποτελείται από αμινοξικά κατάλοιπα

της προ-MT4 περιοχής. Ωστόσο, τα συμμετέχοντα αμινοξικά κατάλοιπα διαφοροποιήθηκαν σημαντικά μεταξύ όλων των πρωτεϊνών που μελετήθηκαν με εξαίρεση τις πρωτεΐνες της Ομάδας 2 (Εικόνα 3.44). Η περιοχή MT5 για τις πρωτεΐνες της Ομάδας 1 εντοπίστηκε κοντά στο αμινοτελικό άκρο της δομικής τους ενότητας NodB (ονομαζόμενη MT5'), ενώ η περιοχή MT4 μετατοπίστηκε καθοδικά, προς το καρβοξυτελικό τους άκρο, κατά περίπου 15 κατάλοιπα (Εικόνα 3.44). Επιπλέον, σε αυτές τις τρεις πρωτεΐνες, η περιοχή MT4 φάνηκε να μεταβάλλεται ηλεκτροστατικά, αφού φέρει θετικά φορτισμένα κατάλοιπα, αντίθετα με άλλα μέλη της οικογένειας όπου αυτή η θέση καταλαμβάνεται από την καρβοξυλική ομάδα ενός ασπαρτικού οξέος.



Εικόνα 4.5 Οι όγκοι της κοιλότητας δέσμευσης των πειραματικά προσδιορισμένων με κρυσταλλογραφία ακτίνων-Χ απακετυλασών πολυσακχαρίτη. Ο όγκος της κοιλότητας που εκτείνεται σε κάθε μια πλευρά του ιόντος ψευδαργύρου πληρούται με γκρίζες σφαίρες. Οι περιοχές που ορίζουν τα εξωτερικά όρια της κοιλότητας δέσμευσης χρωματίζονται με μπλέ, κυανό, φούξια, κίτρινο, πορτοκαλί ή κόκκινο σύμφωνα με την εικόνα 3.44 και τα αντίστοιχα αμινοξικά κατάλοιπα παρουσιάζονται με αναπαράσταση ράβδου. Σε όποια από τις δομές υπάρχει το ιόν ψευδαργύρου, παρουσιάζεται με μία κόκκινη σφαίρα. Παρά τη συνολική όμοια αναδίπλωση όλων των δομικών ενοτήτων NodB, η σχετική ανάλυση αποκάλυψε βαθιές παραλλαγές μεταξύ των Ba0330 (**α**) και Ba0331 (**β**) που περιέχουν και τη δομική ενότητα Fn3 καθώς και μεταξύ των υψηλά ομολόγων Bc1974 (**γ**) και Bc1960 (**δ**). Η Ba0150 (**ε**) που είναι μια εκ των οποίων δεν υπάρχει μέταλλο, έχει μια μάλλον μικρή κοιλότητα, ενώ η κοιλότητα της θεωρούμενης απακετυλάσης του N-ακετυλομουραμικού οξέος Ba0424 (**στ**) περιέχει ένα δακτύλιο περιορισμού στην μετά-MT1 περιοχή (στην εικόνα είναι χρωματισμένο πορτοκαλί).

Πίνακας 4.1 Μέτρηση του όγκου της κοιλότητας δέσμευσης.

NodB κοιλότητα δέσμευσης	Όγκος (Å³)
Ba0331	1577
Ba0330	2731
Bc1974	4528
Bc1960	4206
Ba0424	3728
Ba0150	1097

Συλλογικά, η σύγκριση των κοιλοτήτων δέσμευσης μεταξύ των διαφόρων φερόμενων απακετυλασών πολυσακχαρίτη του *B. anthracis* και του *B. cereus* αποκάλυψε αξιοσημείωτη ποικιλότητα ως προς το μέγεθος και το σχήμα της κοιλότητας καθώς επίσης και ως προς τον τύπο αμινοξέος που τις συνθέτουν. Γίνεται φανερό ότι προκύπτουν σοβαρά εμπόδια όσον αφορά στην πρόβλεψη εξειδίκευσης του υποστρώματος μέσω ομοιότητας ή τον προσδιορισμό πιθανών θέσεων αγκύρωσης του υποστρώματος ή ακόμη και των κρίσιμων καταλοίπων που εμπλέκονται σε παροδικές αλληλεπιδράσεις που προάγουν τη μοριακή αναγνώριση. Οι μεταλλάξεις στην δομή γύρω από το ενεργό κέντρο στη θέση δέσμευσης μπορούν να οδηγήσουν στην *βελτιστοποίησή* της με την έννοια των μικρών προσαρμογών της σε διαφορετικά αλλά ουσιαστικά παρόμοια υποστρώματα. Θα μπορούσε όλα αυτά τα

επονομαζόμενα PDA ένζυμα να λειτουργούν σε διαφορετικά επίπεδα των βιοσυνθετικών μονοπατιών του βακτηριακού κυτταρικού τοιχώματος. Αυτή η ανάλυση καταδεικνύει τους ξεχωριστούς ρόλους τους σε κυτταρικό επίπεδο.

4.2.2.2 Μεταλλάξεις στο ενεργό κέντρο καταστροφικές της καταλυτικής λειτουργίας

Μετά το γεγονός της διπλοποίησης είναι εξαιρετικά εύκολο να χαθεί η ενζυμική λειτουργία εξαιτίας μεταλλάξεων στα αμινοξικά κατάλοιπα του ενεργού κέντρου (Leslie 2013). Αυτό είναι κάτι που δεν μπορεί να αγνοηθεί στην περίπτωση πρωτεϊνών όπως η Ba0150 και η Ba3943 του *B. anthracis* και των αντίστοιχων ορθόλογων τους στον *B. cereus*. Οι ορθόλογες Ba0150-Bc0171, στερούνται των δύο ιστιδινών της περιοχής MT2 καθώς στερούνται και του μετάλλου. Στην Ba0150 είναι αξισημείωτα συγκριτικά μικρός σε όγκο ο χώρος της κοιλότητας δέσμευσής της ώστε να αμφισβητείται η ικανότητα δέσμευσης υποστρώματος, και κατά συνέπεια πρόκειται μάλλον για ένα ανενεργό ένζυμο με άγνωστο ρόλο. Έτσι, για κάποιες εκ των PDAs δυνητικά δεν μπορεί να εξαιρεθεί η πιθανότητα ότι έχουν ένα διαφορετικό ρόλο, για παράδειγμα ίσως δομικό ή στηρικτικό κλπ. Ο Leslie (2013) αναφερόμενος στα νεκρά ένζυμα, επισημαίνει ότι παρά την αβεβαιότητα για την κατάλυση μιας χημικής αντίδρασης, ενδέχεται να έχουν ένα ρόλο μιας άλλης κυτταρικής λειτουργίας, για παράδειγμα ίσως ένα ρυθμιστικό ρόλο στη βιοχημική αντίδραση του ενεργού συγγενή τους με βάση το ενδεχόμενο να εντοπίζονται στο ίδιο σημείο και την ίδια χρονική στιγμή με τους ενεργούς συγγενείς τους ώστε ενδεχομένως να επικουρούν στον έλεγχο της βιοχημικής αντίδρασης, ή να έχουν διατηρήσει την ικανότητα δέσμευσης για άλλους σκοπούς.

4.3 *ba0330* και *ba0331* δύο παράλογα γονίδια στην οικογένεια γονιδίων απακετυλασών πολυσακχαρίτη του *Bacillus anthracis*

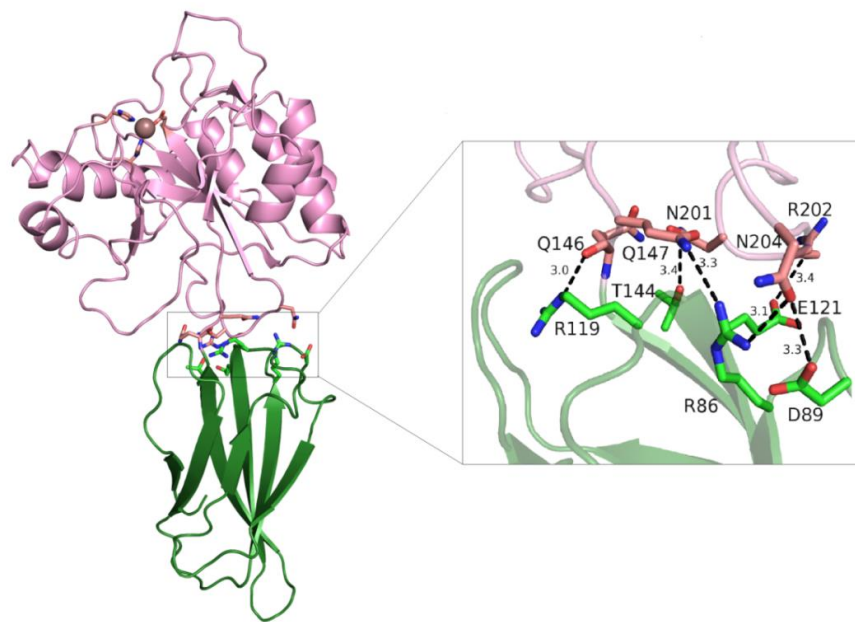
A. Σύγκριση της τρισδιάστατης δομής. Διαφορές και ομοιότητες.

Η Ba0331 και Ba0330 από τον *B. anthracis* έχουν μια παρόμοια αρχιτεκτονική που αποτελείται από μια καρβοξυτελική δομική ενότητα NodB, την οποία μοιράζονται όλα τα ένζυμα της CE4, και μια αμινοτελική δομική ενότητα Fn3 η οποία δεν είναι ένα κοινό χαρακτηριστικό. Οι παρατηρήσεις από τις κρυσταλλικές δομές τους αποκάλυψαν σημαντικές διαφοροποιήσεις και συμβάλλουν στον χαρακτηρισμό των δύο παράλογων γονιδίων.

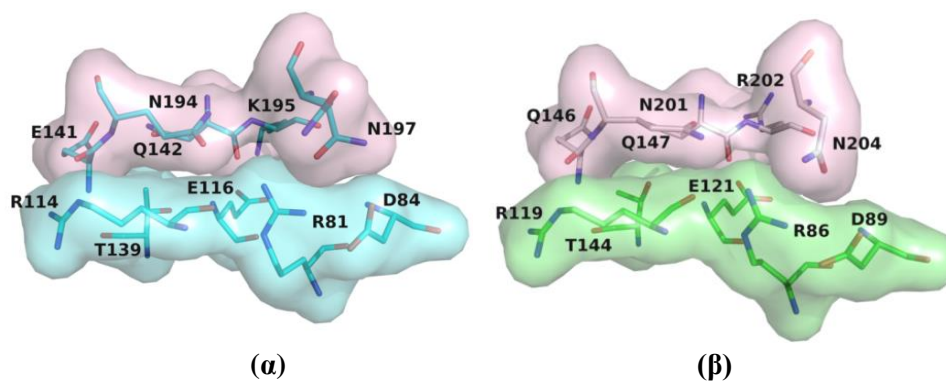
4.3.1 Η επιφάνεια διεπαφής των δομικών ενότητων NodB και Fn3

Η αλληλουχία μοριακής αναγνώρισης του τριπεπτιδίου RGD είναι ένας καλά χαρακτηρισμένος βρόχος διαμοριακής αλληλεπίδρασης (Ruoslahti & Pierschbacher 1987, Papadopoulos *et al.* 1998, Leahy *et al.* 1996, Wojcik *et al.* 2010). Στην περίπτωση των τριών φερόμενων PDAs που περιέχουν την Fn3-like δομική ενότητα, Ba0331, Ba0330 και Bc0361, εμφανίζεται στη γνωστή μετάλλαξη RGE (Xu & Mosher 2011). Η σύγκριση των τριών PDAs αποκαλύπτει ομοιότητες στη θέση και τον προσανατολισμό των δυο δομικών ενότητων (NodB και Fn3-like). Μια προσεκτική εξέταση της επιφάνειας διεπαφής των δομικών ενότητων στις Ba0331 και Ba0330 ή Bc0361, αποκαλύπτει σημαντικές ομοιότητες όσον αφορά στην χωρική τοποθέτηση των πολικών και φορτισμένων αμινοξέων (Εικόνες **4.6**, **4.7 α** και **β**). Οι αλληλεπιδράσεις των μοτίβων στις σχετικές δομικές ενότητες είναι βασικά πανομοιότυπες πηγάζοντας από το διατηρημένο συνολικό σχήμα των απακετυλασών που περιέχουν Fn3-like δομικές ενότητες.

Οι δυο δομικές ενότητες σχηματίζουν μια ενεργή επιφάνεια διεπαφής μεγέθους 448,0 Å² με τη συμμετοχή ενός συντηρημένου μοτίβου (βρόχος RGE, αμινοξέα 119-121) από την Fn3-like δομική ενότητα και ισχυρούς πολικούς δεσμούς με την NodB δομική ενότητα (Εικόνες **4.7 α** και **β**). Το χαρακτηριστικό τριπεπτίδιο μοριακής αλληλεπίδρασης της Fn3 καθώς και η συνεργιστική περιοχή (Leahy *et al.* 1996) εδώ WRTAD (αμινοξέα 85-89) χρησιμοποιούνται στην αλληλεπίδραση με την NodB δομική ενότητα και στην σταθεροποίηση της αρχιτεκτονικής κατασκευής της πρωτεΐνης.



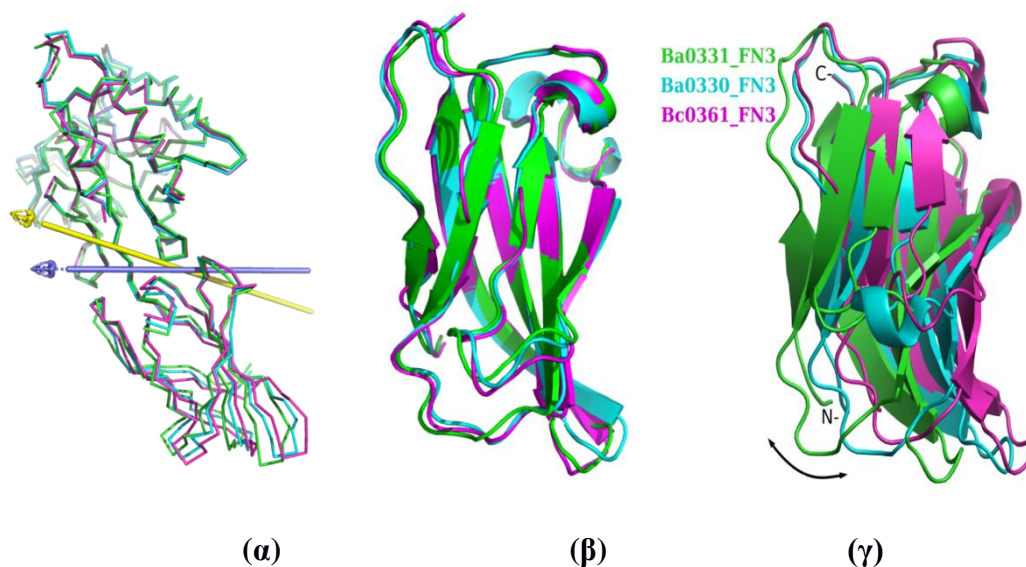
Εικόνα 4.6 Η αρχιτεκτονική της δομής της πρωτεΐνης Ba0331 (αριστερά), στο πλαίσιο (δεξιά) παρουσιάζεται σε μεγέθυνση η τοπολογία της περιοχής της διεπαφής μεταξύ NodB (χρωματίζεται με ροζ) και Fn3 (πράσινο). Εμφανίζονται μόνο κατάλοιπα που αλληλεπιδρούν είτε μέσω δεσμών υδρογόνου είτε μέσω ηλεκτροστατικών αλληλεπιδράσεων. Τα συντηρημένα R119 και E121 του χαρακτηριστικού τριπεπτιδίου RGE από την Fn3 συμμετέχουν σε ισχυρές αλληλεπιδράσεις της διεπαφής.



Εικόνα 4.7 (α) Αναπαράσταση της επιφάνειας της διεπαφής των NodB και Fn3 καθώς και των αμινοξέων (με ράβδους) της Ba0331 και **(β)** Η αντίστοιχη διεπαφή της Ba0330. Τα κατάλοιπα που εμπλέκονται στις πολικές αλληλεπιδράσεις μεταξύ των δύο δομικών ενοτήτων σημαίνονται ρητά. Το τμήμα της δομικής ενότητας NodB εμφανίζεται με τη χρωματισμένα ροζ επιφάνεια και το τμήμα της δομικής ενότητας Fn3-like εμφανίζεται με τη χρωματισμένα με πράσινο επιφάνεια για την Ba0331 ή με κυανό για την Ba0330. Τα κατάλοιπα της δομικής ενότητας Fn3-like είναι ταυτόσημα, ενώ οι διεπαφές των NodB παρουσιάζουν μικρές διαφοροποιήσεις και συγκεκριμένα Q(146) αντί E(141) και R(202) αντί K(195). Ωστόσο, η χωρική διάταξη των αντίστοιχων καταλοίπων είναι σχεδόν ταυτόσημη.

4.3.2 Μοριακή πλαστικότητα των δομών

Εξετάζοντας τη σχετική θέση των δυο δομικών ενοτήτων στις τρεις κρυσταλλικές δομές με υπέρθεση των NodB δομικών ενοτήτων τους αποκαλύπτεται μια διαφορετική πλευρική συστροφή των Fn3-like δομικών ενοτήτων (**Εικόνα 4.8**). Οι άξονες περιστροφής των ζευγών δε συμπίπτουν, αλλά διαφέρουν περίπου 19° φανερώνοντας πως οι Fn3-like δομικές ενότητες δεν περιστρέφονται κατά τον ίδιο τρόπο. Οι μεταξύ των δομικών ενοτήτων διαφορές διευθέτησης δεν προκύπτουν από δομικές παραλλαγές ούτε μεταξύ των NodB δομικών τους ενοτήτων ούτε μεταξύ των Fn3 δομικών τους ενοτήτων. Καίτοι το πρότυπο αλληλεπίδρασης μεταξύ των NodB και Fn3 ταυτίζεται στις Ba0331 και Ba0330 / Bc0361 (Εικόνες 4.7), η συστροφική κίνηση των αντίστοιχων δομικών τους ενοτήτων Fn3 αποκαλύπτει μια ενδογενή ευκινησία, για την οποία δεν μπορεί να παραβλεφθεί κάποιος δυνητικά φυσιολογικός ρόλος. Ωστόσο, πρέπει να σημειωθεί πως οι τρεις πρωτεΐνες έχουν κρυσταλλωθεί σε διαφορετικές ομάδες χώρου και οι δυο δομικές ενότητες τους συμμετέχουν σε διαφορετικές κρυσταλλικές επαφές (ακολουθεί περιγραφή στην §4.3.3). Οι αντίρροπες δυνάμεις εφαρμόζουν επιπρόσθετους περιορισμούς στην Fn3-like δομική ενότητα με αποτέλεσμα να παρατηρούνται διαφορετικές περιστροφικές κινήσεις στις τρεις πρωτεΐνες.



Εικόνα 4.8 (α) Η υπέρθεση των δομικών ενοτήτων NodB της Ba0331 (με πράσινο χρώμα), της Ba0330 (με κυανό) και της Bc0361 (με φούξια, PDB ID: 4HD5) αποκαλύπτει την περιστροφική μετατόπιση της αντίστοιχής δομικής ενότητας Fn3 τύπου. Οι άξονες περιστροφής παρουσιάζονται ως βέλη (κίτρινο για το ζεύγος Ba0331-Ba0330 και μωβ για το ζεύγος Ba0331-Bc0361) **(β)** υπέρθεση των Fn3. **(γ)** Σχετικός προσανατολισμός των δομικών ενοτήτων Fn3 τύπου μετά την υπέρθεση των δομικών ενοτήτων τους NodB. Η Fn3 δομική ενότητα της Bc0361 περιστρέφεται κατά 16° , ενώ εκείνη της Ba0330 κατά 11° σε σχέση με αυτήν της Ba0331.

4.3.2.1 Ανάλυση της κίνησης

Η ανάλυση της περιστροφής μεταξύ των ζευγών Ba0331-Ba0330 και Ba0331-Bc0361 ποσοτικοποιήθηκε με το πρόγραμμα DynDom (Hayward & Berendsen 1998). Η σχετική περιστροφική μετατόπιση των δομικών ενότητων NodB-Fn3 για κάθε ζεύγος πρωτεϊνών, δίδεται από τις γωνίες περιστροφής (RA) γύρω από τον προβλεπόμενο άξονα περιστροφής (Πίνακας 4.2). Η περιστροφή αυτή υποδεικνύει μια κίνηση της άρθρωσης με συμμετοχή των αμινοξέων που εμπλέκονται στην σύνδεση των δύο δομικών ενότητων NodB-Fn3 (Thr144-Val149), όπως φανερώνουν οι διαφορετικές τιμές των γωνιών συστροφής στις κύριες αλυσίδες (Πίνακας 4.3).

Πίνακας 4.2 Ανάλυση κίνησης με χρήση του προγράμματος DynDom

Ba0331- Ba0330								
	Ratio	r.m.s.d. (Å)		RA (°)	TR (Å)	a_CM (°)	d_CM (Å)	CL (%)
	1,022	D1	D2	11,03	-0,185	61,47	1,944	77,8
		0,998	1,223					
Ba0331- Bc0361								
	1,462	D1	D2	15,9	-0,346	73,63	4,148	92,055
			0,936	1,386				

Ratio: αναλογία ενδομοριακής προς διαμοριακής μετατόπισης, r.m.s.d.: για όλα τα άτομα της δομική ενότητας, D1: σταθερό τμήμα δομική ενότητας, D2: κινούμενο τμήμα δομική ενότητας, RA: γωνία περιστροφής, TR: μεταφορά, a_CM: γωνία μεταξύ άξονα περιστροφής και ευθείας που ενώνει τα κέντρα μάζας, d_CM: απόσταση μεταξύ του άξονα και ευθείας που ενώνει τα κέντρα μάζας, CL: απόσταση εγγύτητας

Η σύγκριση των δομών της Ba0330 και Bc0361 με την Ba0331 πραγματοποιήθηκε με άκαμπτου-σώματος μετακίνηση της Fn3 δομικής ενότητας (κινητή δομική ενότητα D2) σε σχέση με εκείνη της NodB (σταθερή δομική ενότητα D1). Οι προεπιλεγμένες παράμετροι του προγράμματος χρησιμοποιήθηκαν στην ανάλυση. Η παραμετρική εξίσωση του άξονα περιστροφής για το ζεύγος Ba0331- Ba0330 είναι $(x,y,z) = (0,340, -0,043, 1,036) + t(-0,051, 0,0085, -0,025)$ ενώ για το Ba0331-Bc0361 είναι $(x,y,z) = (0,366, -0,048, 0,953) + t(-0,032, 0,010, -0,037)$ με σύστημα αναφοράς το Ba0331 και εκπεφρασμένη σε κλασματικές μονάδες των διαστάσεων a, b, c της μοναδιαίας κυψελίδας.

Πίνακας 4.3. Γωνίες συστροφής των εμπλεκόμενων αμινοξέων στην κύρια αλυσίδα.

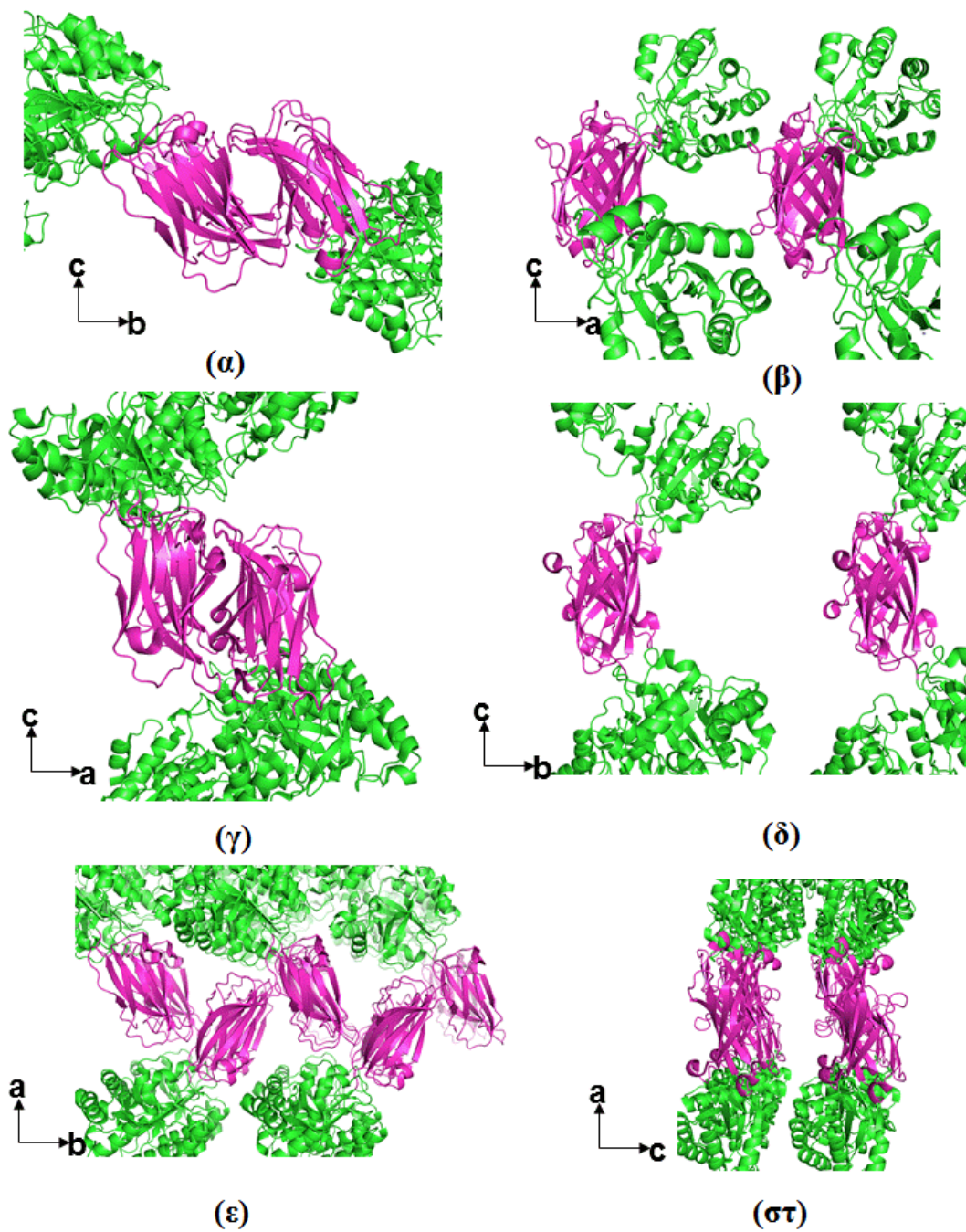
Ba0331						
	Thr144	Phe145	Gln146	Gln147	His148	Val149
ϕ (°)	-121,4	-112,3	-58,2	-110,9	-74,6	-129,2
ψ (°)	132,4	110,3	141,1	145,5	140,0	114,3
Ba0330						
	Thr139	Phe140	Glu141	Gln142	Lys143	Val144
ϕ (°)	-120,8	-96,50	-51,20	-91,60	-72,40	-132,50
ψ (°)	119,3	104,80	129,40	143,70	147,50	102,00
Bc0361						
	Thr139	Phe140	Glu141	Gln142	Lys143	Val144
ϕ (°)	-120,8	-96,50	-51,20	-91,60	-72,40	-132,50
ψ (°)	119,3	104,80	129,40	143,70	147,50	102,00
Ba0331 - Ba0330						
$\Delta\phi$ (°)	0,6	15,8	7,0	19,3	2,2	-3,3
$\Delta\psi$ (°)	-13,1	-5,5	-11,7	-1,8	7,5	-12,3
Ba0331 - Bc0361						
$\Delta\phi$ (°)	-4,6	22,7	4,4	8,0	1,8	-2,7
$\Delta\psi$ (°)	-20,2	-0,8	-8,2	6,9	7,5	-8,2

Οι (ϕ), και (ψ) γωνίες συστροφής για τις Ba0331, Ba0330, Bc0361 και $\Delta\phi$, $\Delta\psi$ των στοιχισμένων αμινοξέων σε υπέρθεση για τα ζεύγη Ba0331-Ba0330 και Ba0331-Bc0361. Η ανάλυση δείχνει πως η περιστροφή για τα δυο ζεύγη, όπως μαρτυρούν οι τιμές των $\Delta\phi$ και $\Delta\psi$ (με χαρακτηρισ χρώματος πράσινου), οφείλεται κυρίως στις αλλαγές των ϕ και ψ στα Thr144-Phe145 και Gln146-Gln147 (η αρίθμηση των αμινοξέων δίδεται για την Ba0331).

4.3.3 Κρυσταλλική μοριακή διευθέτηση της δομικής ενότητας Fn3

Οι τρεις πρωτεΐνες έχουν κρυσταλλωθεί σε διαφορετικές ομάδες χώρου και οι δυο δομικές τους ενότητες συμμετέχουν σε ξεχωριστές κρυσταλλικές επαφές, όπως ήδη αναφέρθηκε (βλ. §4.3.2). Η Ba0331 κρυστάλλωσε σε $P2_1$ έχοντας δυο μόρια στην ασύμμετρη μονάδα και η Fn3-like δομική ενότητα αλληλεπιδρά με την αντίστοιχη Fn3 ενός μόριου σε αντιπαράλληλη διευθέτηση κατά τον άξονα **b**, σχηματίζοντας μια διεπαφή $462,4 \text{ \AA}^2$ σύμφωνα με το πρόγραμμα PISA (Krissinel & Henrick 2007) (Εικόνα 4.9 α και β). Όμοια για την Ba0330, η οποία ωστόσο κρυστάλλωσε στο $C2$, η Fn3-like δομική ενότητα αλληλεπιδρά με τη συμμετρική της Fn3-like κατά τον άξονα **a** κατά αντιπαράλληλο πάλι τρόπο (Εικόνα 4.9 γ και δ). Η επιφάνεια διεπαφής ανάμεσα στις γειτονικές Fn3-like δομικές ενότητες στην δομή της Ba0330 υπολογίστηκε $564,8 \text{ \AA}^2$. Οι κρυσταλλικές επαφές στη Bc0361, που κρυστάλλωσε στην ομάδα χώρου $P2_1$, διαφέρουν σημαντικά. Κάθε Fn3-like δομική ενότητα αλληλεπιδρά με δυο συμμετρικές της προσανατολισμένες υπό γωνία και στις δυο πλευρές της στον άξονα **b** και επιπλέον με δύο NodB δομικές ενότητες κατά τον άξονα **a** (Εικόνα 4.9 ε και στ). Έτσι, η επιφάνεια διεπαφής για κάθε Fn3-like περιοχή γειτονικών μορίων είναι $310,9 \text{ \AA}^2$, και μεταξύ δυο NodB περιοχών $362,1 \text{ \AA}^2$ συνολικά. Από την ανάλυση της κρυσταλλικής μοριακής διευθέτησης συμπεραίνουμε πως οι αντισταθμιστικές δυνάμεις επιβάλλουν μια πρόσθετη τάση στην Fn3-like δομική ενότητα με αποτέλεσμα να την οδηγούν σε σχετική περιστροφική μετατόπιση μεταξύ των τριών πρωτεϊνών.

Συνολικά, η αρχική ιδέα περί δομικής σταθερότητας στις πρωτεΐνες της PDA οικογένειας που περιέχουν Fn3-like δομικές ενότητες φαίνεται να αμφισβητείται λόγω της σχετικής ευκαμψίας μεταξύ των δομικών ενότητων.



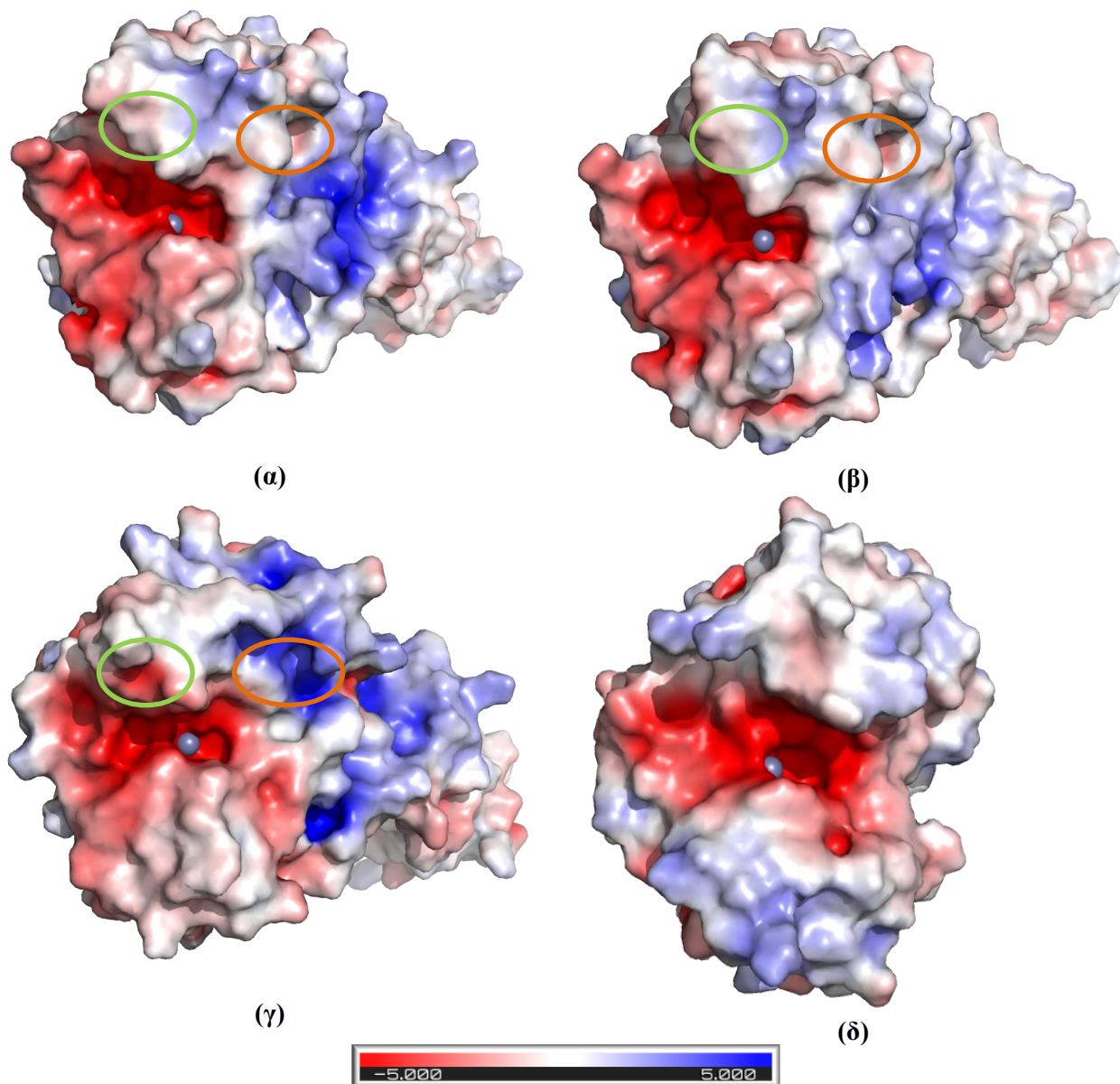
Εικόνα 4.9 Αλληλεπιδράσεις της κρυσταλλικής μοριακής διεύθετησης της Fn3-like δομικής ενότητας στην Ba0331 (α και β), στην Ba0330 (γ και δ) και στη Bc0361 (ε και στ).

4.3.4 Η NodB δομική ενότητα και η κοιλότητα δέσμευσης

Οι παράλογες PDAs Ba0331 και Ba0330 διαφέρουν ουσιαστικά όταν συγκρίνονται οι αύλακες που σχηματίζουν τα μοτίβα της καταλυτικής θέσης όσον αφορά τη θέση ή τις φυσικοχημικές τους ιδιότητες ή και τα δύο.

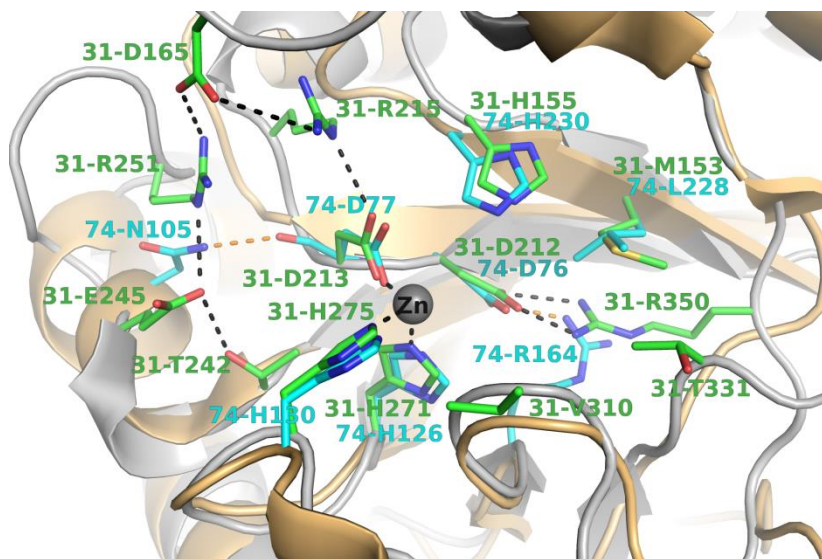
Η σύγκριση των επιφανειών ηλεκτροστατικού δυναμικού όπως υπολογίστηκε από το πρόγραμμα APBS (Baker *et al.* 2001), και παρουσιάζεται εδώ αποκαλύπτει πως η Ba0331, σε αντίθεση με τις δύο ορθόλογες Ba0330 και Bc0361, περιέχει μια εμφανώς αρνητικά φορτισμένη περιοχή εντοπισμένη στην κοίλη επιφάνεια του αύλακα και στα γειτονικά της εγκολλώματα της περιφέρειάς του (**Εικόνα 4.10**).

Η μη ύπαρξη πειραματικών δεδομένων για τις ικανότητες πρόσδεσης υποστρωμάτων των τριών απακετυλασών Ba0331, Ba0330 και Bc0361, καθιστά σημαντική τη σύγκριση της δομής της Ba0331 με την γνωστή ως απακετυλάση της GlcNAc της πεπτιδογλυκάνης Bc1974, αποκαλύπτοντας σημαντικές πληροφορίες όσον αφορά στην τροποποίηση της κοιλότητας δέσμευσης. Η Bc1974, όπως και η Ba0331, εντοπίζεται στον εξωκυτταρικό χώρο, και είναι μια μεμβρανική πρωτεΐνη που αγκυροβολεί με το αμινοτελικό της άκρο (Giastas *et al.* 2018) και η δράση απακετυλάσης στην *N*-ακετυλογλυκοζαμίνη της πεπτιδογλυκάνης έχει καθοριστεί (Balomenou *et al.* 2013). Υπέρθωση της NodB δομικής ενότητας της Ba0331 με την Bc1974 δείχνει σημαντική δομική απόκλιση στην καταλυτική κοιλότητα εκτός από τα συντηρημένα στοιχεία που συμμετέχουν στη συναρμογή του ψευδαργύρου (βλ. §4.1.1). Ενδιαφέρον παρουσιάζει το γεγονός πως τα άνω όρια του αύλακα της κοιλότητας στην Ba0331 είναι κατειλημμένα από φορτισμένα και πολικά αμινοξέα που σχηματίζουν ένα εκτεταμένο δίκτυο γεφυρών αλάτων και δεσμών υδρογόνου που ρυθμίζουν τη χωρική διεύθυνση των βρόχων που ξεκινούν από αυτές (**Εικόνα 4.11**). Ειδικότερα, όπως ήδη έχει περιγραφεί (§3.6.2) βασικά και όξινα αμινοξέα (Asp213, Arg215, Asp165, Arg251, Glu245) παρόντα με εναλλακτικό τρόπο σχηματίζουν ένα αλληλεπιδρών δίκτυο που τερματίζει στην πολική Thr242 και αναγκάζει το β9-β10 βρόχο (Ser164-Asp165-Gly166) να μετακινηθεί προς την κοιλότητα δέσμευσης μειώνοντας τοπικά το διαθέσιμο όγκο της κοιλότητας. Από την άλλη μεριά, οι αντίστοιχες περιοχές στην Bc1974 απαρτίζονται από μικρότερα και κυρίως υδρόφοβα αμινοξέα (Gly80, Gly103, Ala104, Val106) τα οποία, με εξαίρεση ένα δεσμό υδρογόνου μεταξύ της Asn105 και του καρβονυλικού οξυγόνου του Asp77, δεν προάγουν ισχυρές αλληλεπιδράσεις.



Εικόνα 4.10 Ηλεκτροστατικό επιφανειακό δυναμικό της **(α)** Ba0330, **(β)** Bc0361(PDB ID: 4HD5) **(γ)** Ba0331 και **(δ)** Bc1974. Στην καταλυτική κοιλότητά τους υπάρχει σε όλες ένα ιόν ψευδαργύρου σε συναρμογή με τη συντηρημένη τριάδα Asp-His-His. Οι θέσεις που παρουσιάζουν διαφορές στο φορτίο μεταξύ της Ba0331 σε σύγκριση με τις ομόλογες της Ba0330 και Bc0361 εμφανίζονται με τα οβάλ σχήματα χρώματος πράσινου και πορτοκαλί. Μια περιοχή φορτισμένη θετικά βρίσκεται κοντά στο μοτίβο MT1 στην Ba0331. Μια παρόμοια ηλεκτροστατική ανάλυση της Bc1974 δείχνει ότι η κοιλότητα δέσμευσης της είναι περισσότερο ηλεκτραρνητική και λιγότερο περιορισμένη. Το ηλεκτροστατικό επιφανειακό δυναμικό των μοντέλων υπολογίστηκε από το APBS (Adaptive Poisson-Boltzmann Solver) πρόσθετο του PyMOL, με τις προεπιλεγμένες ρυθμίσεις παραμέτρων. Η μη γραμμική εξίσωση Poisson-Boltzmann περιγράφεται από το $\pm 5kT / e$ ηλεκτροστατικό δυναμικό που απεικονίζει τις αρνητικά και θετικά φορτισμένες επιφάνειες χρωματισμένες με κόκκινο και μπλε, αντίστοιχα.

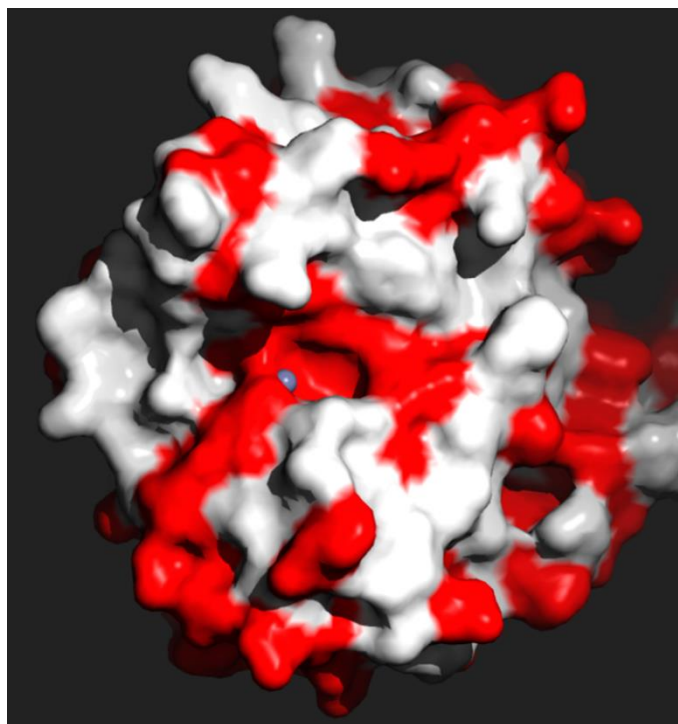
Η σύγκριση της κοιλότητας δέσμευσης της Ba0331 και της Bc1974 έχει αποκαλύψει ένα συντηρημένο δομικό χαρακτηριστικό, παρά τις διαφορές τους στην πρωτοταγή δομή. Στην Bc1974 το Asp76, ένα απόλυτα συντηρημένο αμινοξύ της θέσης πρόσδεσης (πλην της Ba3943/Bc3804), εμπλέκεται στο σχηματισμό του αρνητικά φορτισμένου περιβάλλοντος στην απομακρυσμένη πλευρά από το ιόν ψευδαργύρου. Αυτό ενώνεται με την Arg164, ένα υψηλά συντηρημένο αμινοξύ σε σχεδόν όλες τις απακετυλάσες του MT3 εκτός από εκείνες που φέρουν τη δομική ενότητα της ινωδονεκτίνης τύπου III (Fn3), προς σχηματισμό μιας ηλεκτροστατικής αλληλεπίδρασης. Στις Ba0331, Ba0330 και Bc0361 η τελευταία θέση έχει αντικατασταθεί από μια αλανίνη στην αλληλουχία, ενώ μια αργινίνη (Arg350 στην Ba0331, Arg343 στην Ba0330) από άλλο τμήμα της δομής και διαφορετική κατεύθυνση, από το MT4', διατηρεί μια ισοδύναμη, (το αντίστοιχο Asp76 στην Bc1974) ηλεκτροστατική αλληλεπίδραση με το ασπαρτικό οξύ (Asp212 στην Ba0331, Asp205 στην Ba0330) στην περιοχή (Εικόνα 4.11). Αυτό το γεγονός υποδεικνύει τη σημασία της τοποθεσίας του Asp στη θέση πρόσδεσης και της ηλεκτροστατικής αλληλεπίδρασης με την Arg διαμορφώνοντας τις αλληλεπιδράσεις του υποστρώματος, που ακόμα μένει να βρεθεί.



Εικόνα 4.11 Υπέρθηση των θέσεων πρόσδεσης της Bc1974 (με χρυσό) και Ba0331 (με ασημένιο). Η εικόνα είναι εστιασμένη γύρω από το ιόν ψευδαργύρου και επομένως εμφανίζονται μόνο οι δύο κοιλότητες της θέσης δέσμευσης. Τα κατάλοιπα που εμπλέκονται στις αλληλεπιδράσεις παρουσιάζονται με αναπαράσταση ράβδων και οι αλληλεπιδράσεις με μαύρες και πορτοκαλί διακεκομμένες γραμμές για τη δομή της Ba0331 και της Bc1974, αντίστοιχα. Τα κατάλοιπα της Ba0331 σχεδιάζονται με πράσινο χρώμα και της Bc1974 με κυανό.

4.3.4.1 Είναι δραματική η διαφορά της θέσης δέσμευσης μεταξύ της Ba0330 με την Ba0331.

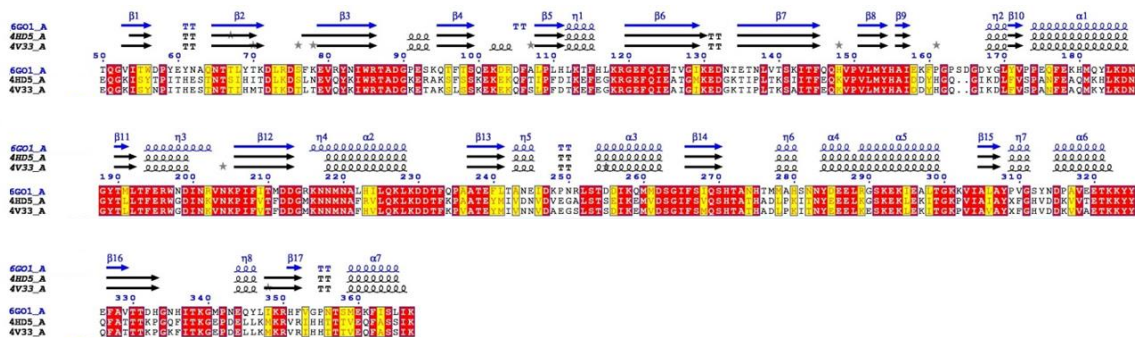
Στην εικόνα 4.12 παρουσιάζεται μια αναπαράσταση της επιφάνειας της καταλυτικής ενότητας NodB των Ba0331, Ba0330 και Bc0361, στην οποία δίδεται η συντήρηση. Διαφοροποίηση αναδεικνύεται και από το ηλεκτροστατικό επιφανειακό δυναμικό (βλ. εικόνα 4.10). Όπως παρουσιάστηκε στην §4.2.2.1, μια σημαντική διακύμανση στο μήκος, το πλάτος και το βάθος δημιουργείται στη θέση πρόσδεσης μεταξύ των δύο μορίων. Συνολικά παρά την ομοιότητα που παρουσιάζουν στην αναδίπλωση, η διαφορά στη θέση δέσμευσης μεταξύ της Ba0331 και των Ba0330/Bc0361 (βλ. Εικόνα 3.50) είναι εκπληκτική.



Εικόνα 4.12 Αναπαράσταση της επιφάνειας της καταλυτικής ενότητας NodB των Ba0331, Ba0330 και Bc0361. Τα απόλυτα συντηρημένα αμινοξέα χρωματίζονται με κόκκινο χρώμα. Η θέση του ψευδαργύρου υποδεικνύεται με γκρί σφαίρα.

B. Σύγκριση των πρωτοταγών αλληλουχιών των Ba0330, Ba0331 - Διαφορές και ομοιότητες.

Στην **Εικόνα 4.13**, παρουσιάζεται η συντήρηση που παρατηρείται εντός των αλληλουχιών της Ομάδας 1.



Εικόνα 4.13 Η συντήρηση εντός των αλληλουχιών της Ομάδας 1. Τα απολύτως συντηρημένα αμινοξικά κατάλοιπα δείχνονται με κόκκινο χρώμα φόντου και τα μερικώς συντηρημένα δείχνονται με κίτρινο ή λευκό φόντο. Τα δευτεροταγή δομικά στοιχεία για τις τρεις δομές δίνονται στο πάνω μέρος (α-έλικες ως σπείρες, β-κλώνοι ως βέλη, στροφές ως T). Η απεικόνιση πραγματοποιήθηκε με το ENDSCRIPT.

Σύμφωνα με τη φυλογενετική ανάλυση ο διαχωρισμός αφενός των ομολόγων με την καταλυτική δομική ενότητα και αφετέρου της δομικής ενότητας Fn3 και η σύγκριση τους αποκαλύπτει το πρώτο γεγονός της απόκτησης της δομικής ενότητας Fn3 στην οικογένεια των γονιδίων απακετυλάσης πολυσακχαρίτη. Με αποτέλεσμα να εμφανίζεται στις ορθόλογες της Ba0330. Κατόπιν η Ba0331 εμφανίζεται ως προϊόν ενός γεγονότος διπλοποίησης γονιδίου. Φαίνεται ότι πρόκειται για επαναλαμβανόμενη διπλοποίηση καθώς η διάταξη είναι εν σειρά και ίδια με την αρχική. Το αντίγραφο *ba0331* παραμένει δίπλα στο *ba0330*. Η σύγκριση τους αφορά τη γονιδιακή διπλοποίηση που τα δημιούργησε.

Τα αποτελέσματα της εξελικτικής ανάλυσης υποδεικνύουν ότι οι δύο διακριτές δομικές ενότητες NodB και Fn3, έχουν διαφορετικούς ρυθμούς εξέλιξης. Η μέτρηση του ρυθμού μη συνώνυμων προς συνώνυμων αντικαταστάσεων (με $dN/dS < 1$) παραπέμπει σε ένδειξη μιας νέας λειτουργίας. Επειδή όμως η σύγκριση αυτή γίνεται μεταξύ πολύ κοντινών συσχετιζόμενων γονιδιακών αντιγράφων καθιστά αναγκαία την υπόθεση μειωμένης πιθανότητας τα δύο αντίγραφα ήδη να παρουσιάζουν μια νέα λειτουργία. Στις περιπτώσεις

που δεν είναι δυνατό να μετρηθεί μια στατιστικώς σημαντική διαφορά από την ουδετερότητα η μέτρηση dN/dS δε θα μπορούσε να χρησιμοποιηθεί. Έτσι επιλέχθηκε και ο υπολογισμός του ρυθμού μη συνώνυμων προς συνώνυμων πολυμορφισμών από όπου προέκυψε με βάση τη δοκιμασία McDonald-Kreitman και πάλι αρνητική επιλογή ($NI > 1$) η οποία και εμποδίζει τη σταθεροποίηση επιβλαβών μεταλλάξεων.

Γ. Η δομή φωτίζει την αλληλουχία.

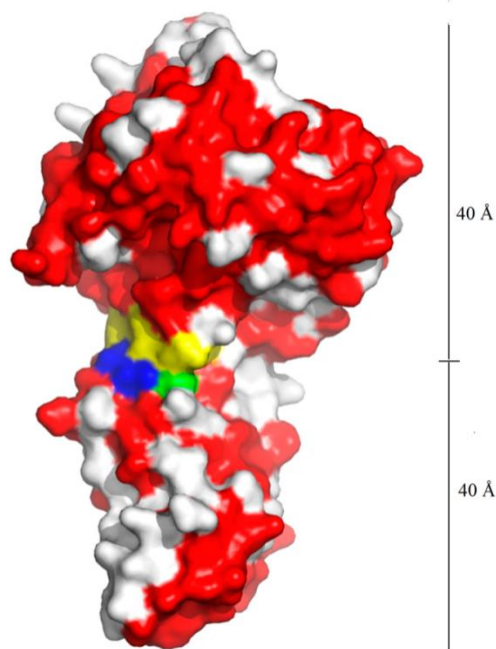
Η ύπαρξη σημαντικής ομοιότητας μεταξύ των αλληλουχιών των δύο μορίων Ba0330 και Ba0331 σε επίπεδο αλληλουχίας νουκλεοτιδίων και αμινοξέων είναι σαφές ότι εκδηλώνεται ως ομοιότητα στην τρισδιάστατη δομή και υποδηλώνει την ίδια εξελικτική προέλευση, και όμοιους μηχανισμούς λειτουργίας. Στην περίπτωση διπλοποίησης ενός γονιδίου δημιουργούνται δύο αντίγραφα του, που αρχικά έχουν την ίδια λειτουργία, στην πορεία όμως της εξέλιξης αποκλίνουν κατόπιν συσσώρευσης μεταλλάξεων. Η ανάλυση της αλληλουχίας τους παρέχει μερική μόνο απάντηση στην ένδειξη της απόκλισης, αντίθετα αναδεικνύεται σημαντικής σημασίας η ανάλυση της τρισδιάστατης δομής η οποία μπορεί να αποκαλύψει τα ακριβή στοιχεία της διαφοροποίησης αυτής, που δεν ανιχνεύονται εύκολα στο επίπεδο της αλληλουχίας και μόνο. Όπως π.χ. η ένθεση των δύο αμινοξέων στο β9-β10 βρόχο και η εκπληκτική διαφορά στη θέση δέσμευσης τους όπως αποκαλύφθηκε με αποτέλεσμα δύο μοναδικές κοιλότητες δέσμευσης (βλ. Εικόνες 3.49 και 3.50). Όπως αναφέρθηκε (βλ. §1.2) οι βρόχοι είναι περιοχές που εξελίσσονται πιο γρήγορα. Η σύγκριση της τρισδιάστατης δομής, όπως παρουσιάστηκε (βλ. §4.3.1-4.3.4) από τις διαφορές και τις ομοιότητες, οδηγεί στη πορεία της εξέλιξης δύο παραλόγων γονιδίων. Οι διαφορές στις λεπτομέρειες της τρισδιάστατης δομής τους αντικατοπτρίζει τη διαφορά που ενδεχομένως έχουν στο βιοχημικό τους ρόλο. Τα ευρήματα θα είναι συνεπώς χρήσιμα για τη διερεύνηση πιθανών αποκλινόντων ρόλων μεταξύ του *ba0330* και του διπλοποιημένου γονιδίου *ba0331*, στα υψηλά μολυσματικά στελέχη.

4. 4 Ο ρόλος της δομικής ενότητας Fn3

Τα Fn3 στοιχεία εμφανίζονται συχνά σε βακτηριακές εξωκυτταρικές καρβονικές ανυδράσες υδατανθράκων όπως χιτινάσες, αμυλάσες, κυτταρινάσες και πιστεύεται ότι συμμετέχουν στην προαγωγή της υδρόλυσης των υποστρωμάτων υδατανθράκων τροποποιώντας τις επιφάνειές τους, ενώ παράλληλα παίζουν σημαντικό ρόλο σχηματίζοντας αλληλεπιδράσεις πρωτεΐνης- πρωτεΐνης (Campbell & Spitzfaden 1994, Kataeva *et al.* 2002, Spitzfaden *et al.* 1997, Zhao *et al.* 2017). Η Fn3 δομική ενότητα είναι παρούσα σε διάφορα υδρολυτικά ένζυμα της πεπτιδογλυκάνης, όπως οι D,D-καρβοξυπεπτιδάσες PPB5 και PBP6 στο *E. coli* (Nelson *et al.* 2002, Chen *et al.* 2009). Έχει προταθεί πως οι διαφορές στις Fn3 δομικές ενότητες πιθανόν να επηρεάζουν τη θέση των πρωτεϊνών, την ενζυμική δραστηριότητα ή αλληλεπιδράσεις με άλλα στοιχεία του βιοσυνθετικού μηχανισμού της πεπτιδογλυκάνης (Chen *et al.* 2009).

Στην δομική ενότητα Fn3 μεταξύ των Ba0330 και Ba0331 (47,4% ομοιότητα) η περιοχή KRGEFQIE (αμινοξέα 114-121 στην Ba0330 και 119-126 στην Ba0331) που περιέχει την RGE και η συνεργιστική της περιοχή IWRTADG (αμινοξέα 79-85 στην Ba0330 και 84-90 στην Ba0331) (εικόνα 4.13) εμφανίζουν πλήρη ταύτιση τόσο στην αλληλουχία όσο και στη τριτοταγή δομή (οι δύο περιοχές έχουν συζητηθεί και στην §4.3.1). Στην υπόλοιπη όμως ακολουθία οι δύο πρωτεΐνες εμφανίζουν αξιοσημείωτη διαφοροποίηση (λευκές περιοχές στις εικόνες 4.12 και **4.14**).

Ειδικότερα για δύο περιοχές, η μια περιοχή εντοπίζεται στο βρόχο β6-β7, που αποτελείται από τα αμινοξέα Q100→D105, T66→K71, L73→F78, N74→K79, T134→V139, K135→T140 (γράφονται πρώτα τα αμινοξέα που απαντώνται στην Ba0330 και ακολουθούν αυτά της Ba0331, Ba0330→Ba0331), και η δεύτερη στο τέλος του κλώνου β2 και αποτελείται από τα Q76→R81, K95→Q100, GKTIP(128-132)→NTETN(133-137), οι διαφοροποιήσεις παρουσιάζονται σε αμινοξέα που βρίσκονται στο εξωτερικό μέρος, γεγονός που δεν αποκλείει την πιθανότητα οι Ba0330 και Ba0331 να εμπλέκονται σε διαφορετικές αλληλεπιδράσεις. Επιπροσθέτως η ευκαμψία που ήδη συζητήθηκε (βλ. §4.3.2) μεταξύ της δομικής ενότητας Fn3 και NodB αλλά και η αναμενόμενη μεταξύ της αμινοτελικής α-έλικας και της Fn3, δεν μπορεί να αγνοηθεί ότι πιθανόν διαδραματίζει ρόλο στη φυσιολογία της πρωτεΐνης.



Εικόνα 4.14 Αναπαράσταση της επιφάνειας των δομών των Ba0331, Ba0330 και Bc0361, όπου αναδεικνύεται η συντήρηση. Τα απόλυτα συντηρημένα αμινοξικά κατάλοιπα χρωματίζονται με κόκκινο χρώμα, ενώ τα υπόλοιπα είναι λευκά. Η επιφάνεια διεπαφής μεταξύ των δομικών ενοτήτων επισημαίνεται με πράσινο (περιοχή RGE), μπλε (συνεργιστική περιοχή WRTAD και κίτρινο (περιοχή της NodB). Αριστερά δίδεται το μήκος των δομικών ενοτήτων σε Ångströms.

Επιπροσθέτως είναι πιθανό ότι ο διαφορετικός κυτταρικός εντοπισμός των Ba0330 και Ba0331 (Arnaouteli *et al.* 2015) να σχετίζεται με διαφορετικές αλληλεπιδράσεις των δύο δομικών ενοτήτων Fn3. Παρόλο που παρουσιάζεται μεταβολή στα κατάλοιπα στο εξωτερικό της δομικής ενότητας Fn3 της πρωτεΐνης διατηρείται ο χαρακτήρας τους ως προς τις φυσικοχημικές τους ιδιότητες. Συνολικά, το μήκος του μορίου της πρωτεΐνης Ba0330 και Ba0331, λαμβάνοντας υπόψη το μήκος της αμινοτελικής α-έλικας με 5-7 στροφές (περίπου 25Å), την περιοχή Fn3 (περίπου 40 Å) και την περιοχή NodB (περίπου 40 Å) (Εικόνα 4.14) μπορεί να φθάσει σε μήκος άνω των 105 Å, το οποίο είναι πολύ μεγαλύτερο από τα άλλα ένζυμα PDA που έχουν αγκυρωθεί στη μεμβράνη όπως το Bc1974 και ούτω καθεξής. Συνοψίζοντας η παρουσία της δομικής ενότητας Fn3, μοναδική στην αρχιτεκτονική των Ba0330, Ba0331, και Bc0361 για τις δύο οικογένειες των *B. anthracis*, και *B.cereus* που εξετάζονται, προσδίδει κάποιο φυσικό πλεονέκτημα, είτε μέσω της εισαγωγής μήκους-επέκτασης (ως περιοχή μεσοδιαστήματος), ή/και της εισαγωγής ιδιοτήτων προσκόλλησης ή ιδιοτήτων αλληλεπίδρασης, ή/και της εισαγωγής ευλυγισίας.

4.5 Στοιχεία διαφοροποίησης λειτουργικών χαρακτηριστικών μέσω της δομικής ανάλυσης

Οι δομικές πληροφορίες που παρέχονται από τις κρυσταλλογραφικές μελέτες ακτίνων -X για τη δομική ενότητα NodB των εστερασών CE4 μπορούν να συμβάλλουν στην κατανόηση των βιολογικών λειτουργιών αυτών των πρωτεϊνών και του μηχανισμού που δρουν σε ολιγοσακχαρίτες και πολυσακχαρίτες. Παρόλο που αυτές οι δομικές ενότητες των CE4 μπορούν να ομαδοποιηθούν με βάση τη συντήρηση της αλληλουχίας ή της πρωτεϊνικής αναδίπλωσης, τέτοιες ομαδοποιήσεις δε δίνουν πρόγνωση για τη λειτουργία. Η σύγκριση της κοιλότητας δέσμευσης μεταξύ των διαφόρων PDA αποκάλυψε αξιοσημείωτη μεταβλητότητα από την άποψη του μεγέθους, του σχήματος και του χαρακτήρα της κοιλότητας καθώς και ως προς τη σύνθεση τύπου αμινοξέων, που αντικατοπτρίζουν τις βιολογικές λειτουργίες και την τοπική θέση αυτών των πρωτεϊνών, καθώς και των μηχανισμών με τους οποίους δρουν στο υπόστρωμα τους (ολιγοσακχαρίτες και πολυσακχαρίτες). Η προσπάθεια μηχανιστικής κατανόησης της μοριακής λειτουργικότητας των οικογενειών απακετυλάσης πολυσακχαριτών *B. anthracis* και *B. cereus* μέσω της δομικής ανάλυσης έδειξε ότι μικρές δομικές τροποποιήσεις στη συναρμολόγηση πολλαπλών μοτίβων μπορεί να προκαλέσουν λειτουργική διαφοροποίηση. Εδώ παρουσιάζονται στοιχεία απόκλισης στα λειτουργικά χαρακτηριστικά με ταυτόχρονη διατήρηση των διακριτών τοπικών δομικών χαρακτηριστικών στην οικογένεια των απακετυλασών πολυσακχαρίτη. Αυτό μπορεί να υποδεικνύει ότι τα τμήματα της αλληλουχίας που σχηματίζουν τη θέση δέσμευσης βρίσκονται υπό εξελικτική πίεση για να φιλοξενήσουν και να αναγνωρίσουν διακριτά κατάλοιπα πολυσακχαριτών ανάλογα με την κυτταρική θέση, τη χρήση ή το περιβάλλον. Παρότι αυτό αποτελεί ανοιχτό αντικείμενο βιοχημικής επαλήθευσης, υποδεικνύει πιθανότατα τους ξεχωριστούς ρόλους σε κυτταρικό επίπεδο για τις PDAs των υπό εξέταση οικογενειών.

Από τις μελέτες σχετικά με τη βιολογική λειτουργία των μελών της οικογένειας PDA, είναι κατανοητό ότι η θέση πρόσδεσης των εν λόγω ενζύμων μπορεί να φιλοξενήσει διαφορετικά υποστρώματα και να λειτουργεί σε διαφορετικά περιβάλλοντα (κυτταρική θέση, αλατότητα κλπ., Blair & van Aalten 2004, Balomenou *et al.* 2015). Αυτή η μελέτη, ακολουθώντας τις πρόσφατες μελέτες κρυσταλλογραφίας και μοντελοποίησης δέσμευσης προσδότη σε μέλη της οικογένειας (Arnaouteli *et al.* 2015, Fadoulglou *et al.* 2017, Giastas *et al.* 2018, Andreou *et al.* 2018, Andreou *et al.* 2019), υποδηλώνει ότι παρόλο που η αναδίπλωση της καταλυτικής περιοχής και το ενεργό κέντρο είναι καλά διατηρημένα, οι αντικαταστάσεις μπορεί να έχουν ως αποτέλεσμα τη διαφοροποίηση της λειτουργικότητας.

Ειδικότερα, για την καταλυτική δομική ενότητα NodB, από τα δομικά χαρακτηριστικά που σχηματίζουν τη θέση πρόσδεσης, αποδείχθηκε ότι τρία ειδικά μοτίβα (MT1, MT2 και MT3) που περιείχαν αμινοξικά κατάλοιπα που εμπλέκονται στο ενεργό κέντρο, παρέμειναν σχετικά καλά διατηρημένα στη σύνθεση και σταθερά στη θέση κατά μήκος της ακολουθίας. Δύο άλλα, MT4 και MT5, φαίνεται ότι έχουν αποκλίσει στην ακολουθία για να σχηματίσουν μια θέση πρόσδεσης κατάλληλη για το στοχευόμενο τροποποιημένο πολυσακχαρίτη ή/και το τοπικό περιβάλλον. Σε συσχετισμό με τα πρώιμα βιοχημικά ευρήματα (Balomenou *et al.* 2015), για την εξειδίκευση δέσμευσης ή το τοπικό περιβάλλον των PDAs στα είδη της ομάδας του *B. cereus*, συνήχθει το συμπέρασμα ότι εντός της οικογένειας PDA, τα εμπλεκόμενα γονίδια έχουν εξελιχθεί ώστε να συμβάλλουν στην διαφοροποιημένη λειτουργικότητα.

Επιπλέον, για ορισμένες PDAs των λοιμογόνων στελεχών, η εισαγωγή των N-τελικών ελίκων και των επιπρόσθετων δομικών ενότητων Fn3 που μπορεί να έχουν υιοθετηθεί μέσω γονιδιακής μεταφοράς και υποτίθεται ότι δρουν ως στοιχεία επέκτασης, ή που έχουν μια άγνωστη μέχρι σήμερα λειτουργία διαμεσολάβησης που υποδεικνύεται από το συντηρημένο μοτίβο στην έλικα ή το ιστορικό αλληλεπίδρασης μορίου-μορίου της δομικής ενότητας Fn3 θα μπορούσε να επηρεάσει την αναγνώριση των βακτηρίων από τους ανοσολογικούς παράγοντες του ξενιστή και να συμβάλλει στην αντίσταση των βακτηρίων από τους αμυντικούς παράγοντες του ξενιστή, αυξάνοντας έτσι την παθογένεια.

Σε μια πιο πρόσφατη υπολειτουργικοποίηση PDA, η οποία μπορεί να συνδέεται με τις τροποποιήσεις των πιο λοιμογόνων στελεχών, οι διαφορές αλληλουχίας στη θέση πρόσδεσης των στενά σχετιζόμενων PDAs Ba0331 και Ba0330-Bc0361, που όλες περιέχουν ένα λιποκυτίο και τις αμινοτελικές περιοχές α-έλικας και Fn3, δείχνουν ότι η ανάγκη για λειτουργική διαφοροποίηση έχει οδηγήσει σε περαιτέρω προσαρμογή της θέσης πρόσδεσης στα πρωτεϊνικά μόρια, ώστε ενδεχομένως να μπορούν να ασκήσουν εναρμονισμένους ή/και ανεξάρτητους ρόλους. Ακόμη και αν η ομολογία αλληλουχίας μεταξύ Ba0331 και Ba0330 είναι υψηλή, η ομαδοποίηση των ομολόγων στο δένδρο δείχνει εξελικτική απόσταση. Αυτό υποστηρίζεται και από τις σημαντικές διαφορές της κοιλότητας δέσμευσης του ενζύμου Ba0331, σε σύγκριση με αυτήν του Ba0330, στο σχήμα, τον όγκο, το μήκος και τις συγκεκριμένες θέσεις καταλοίπων που τροποποιούν την τοπογραφία της θέσης πρόσδεσης, με αποτέλεσμα ένα δραματικά διαφορετικό δίκτυο και χαρακτήρα δυναμικών αλληλεπιδράσεων με το υπόστρωμα.

Η εξελικτική λειτουργική διαφοροποίηση στην ανάπτυξη της εξειδίκευσης της δομικής ενότητας NodB της θέσης δέσμευσης προφανώς είναι αποτέλεσμα γεγονότων

διπλοποίησης γονιδίου. Βέβαια καίτοι απόψεις λειτουργικής ποικιλομορφίας έχουν επισημανθεί σε αυτήν την εργασία, η άμεση συσχέτιση με το υπόστρωμα παραμένει ένα σημαντικό ερώτημα που μένει να απαντηθεί.

4.6 Επίσταση

Μελέτες έχουν επισημάνει το σημαντικό αντίκτυπο της επίστασης στον τρόπο με τον οποίο εξελίσσονται οι πρωτεΐνες (Wanz & Pollock 2005, Lunzer *et al.* 2010, Breen *et al.* 2012, Pollock *et al.* 2012, Starr & Thornton 2016). Ένα από τα καλύτερα συστήματα για τη διερεύνηση των λειτουργικών και εξελικτικών συνεπειών της επίστασης είναι μέσω των αλληλεπιδράσεων εντός ενός τόπου οι οποίες χαρακτηρίζουν την αναδίπλωση και τη δραστικότητα των πρωτεϊνών (Phillips 2008). Στην παρούσα μελέτη η περιοχή πρόσδεσης και αναγνώρισης του υποστρώματος υποδεικνύει τη σημασία της επίστασης.

Οι διαφορές στην τρισδιάστατη τοπολογία της αλληλουχίας που διατηρεί τα μοτίβα που σχηματίζουν τη θέση δέσμευσης είναι καθοριστικός παράγοντας για την εξειδίκευση της δέσμευσης. Η πρόβλεψη της εξειδίκευσης του προσδέτη, που βασίζεται αποκλειστικά στην αναδίπλωση της δομικής ενότητας NodB, είναι ανεπαρκής. Μερικοί σημαντικοί παράγοντες φαίνεται να είναι η θέση των αρωματικών καταλοίπων και δομών βρόχου που διαμορφώνουν τις θέσεις σύνδεσης για να αντικατοπτρίζουν ενδεχομένως τη διαμόρφωση του προσδέτη (βλ. Εικόνα 3.44). Η αλληλεπίδραση πλευρικών αλυσίδων αρωματικών αμινοξέων με τον προσδέτη-υπόστρωμα είναι πανταχού παρούσα σε αναγνώριση υδατανθράκων στην NodB (βλ. §3.4.5 και §4.1.1). Οι πλευρικές αλυσίδες τρυπτοφάνης, τυροσίνης και, λιγότερο συχνά, φαινυλαανίνης σχηματίζουν υδρόφοβα μοτίβα των θέσεων δεσμεύσεως της δομικής ενότητας NodB, οι οποίες μπορεί να είναι μακρές, βραχείες ή βαθιές σε σχέση με το ενεργό κέντρο (Εικόνα 4.5). Επιπροσθέτως, το μέγεθος ή το σχήμα της θέσης δέσμευσης που διαφοροποιείται εξαιτίας εμφανίσεως λεπτών δομικών αλλαγών ίσως χρησιμοποιούνται από το βακτήριο για να ικανοποιήσει τις διαμορφώσεις των διαλυτών ολιγοσακχαριτικών υποστρωμάτων (βλ. §4.5). Η υδατοδιαλυτή φύση των υδατανθράκων, λόγω των υδροξυλομάδων, προσθέτει στην αναγνώρισή τους μέσω του σχηματισμού δεσμών υδρογόνου με πολικά ή φορτισμένα κατάλοιπα τοποθετημένα εντός ή στο χείλος των θέσεων δέσμευσης (βλ. §4.2.1).

Η εξειδίκευση υποστρώματος μπορεί να μεταβληθεί με πολύ λίγες αντικαταστάσεις αμινοξέων ή επανατοποθέτηση εντός και γύρω από την κρίσιμη κοιλότητα που είναι υπεύθυνη για την εξειδίκευση του υποστρώματος και τη σύνδεση. Αυτή η εκλεκτικότητα στην εξειδίκευση, λόγω μεταβολών στο μέγεθος και στο σχήμα του αύλακα που δεσμεύει το υπόστρωμα, επηρεάζει τη σύνδεση ή το ταίριασμα του στοχευόμενου υποστρώματος. Ωστόσο, επιπρόσθετες διαφοροποιήσεις στα αμινοξικά κατάλοιπα και μεταβολές στη συνολική δομή στο εξωτερικό μέρος της κοιλότητας δέσμευσης μπορεί επίσης να έχουν συμβάλει στις διαφορές εξειδίκευσης υποστρώματος μεταξύ των διαφορετικών ενζύμων. Αυτό υποδηλώνει ότι η εξειδίκευση του υποστρώματος διέπεται επίσης από την αλληλουχία αμινοξέων και τα δομικά χαρακτηριστικά περιφερειακά της θέσης δέσμευσης.

4.7 Γενετική μηχανική για την έκφραση των απακετυλασών

Η πρόταση για τη μοριακή κατασκευή της κασέτας έκφρασης των απακετυλασών έτσι ώστε το καρβοξυτελικό άκρο να είναι απελευθερωμένο όπως υφίσταται στην φυσική πρωτεΐνη, προέκυψε από τα προβλήματα που παρουσιάστηκαν στην αναδίπλωση κατά την απομόνωση (βλ. §3.1). Η προτεινόμενη μοριακή κατασκευή για την απομόνωση κατάλληλων προς δομικές μελέτες απακετυλασών φαίνεται ότι αποδίδει.

Τα δύο μεταλλαγμένα στελέχη του *E. coli* BL21 (DE3), C41 (DE3) και C43 (DE3) που περιγράφηκαν αρχικά από τους Miroux και Walker, συχνά χρησιμοποιούνται για να υπερνικήσουν την τοξικότητα που σχετίζεται με την υπερέκφραση των ανασυνδυασμένων πρωτεϊνών. Ακόμη και όταν η τοξικότητα των πλασμιδίων είναι τόσο υψηλή που εμποδίζει το μετασχηματισμό στο στέλεχος BL21 (DE3), οι πρωτεΐνες μπορούν συχνά να εκφράζονται με επιτυχία σε C41 (DE3) και /ή C43 (DE3) (Dümmler *et al.* 2005). Στην παρούσα εργασία το C41 (DE3) οδήγησε στην επιτυχημένη έκφραση.

Ακόμη θα πρέπει να αναφερθεί η περαιτέρω δυνατότητα μελέτης της οικογένειας των απακετυλασών με τη μέθοδο της κατευθυνόμενης εξέλιξης η οποία χρησιμοποιείται στην πρωτεϊνική μηχανική, και μιμείται τη διαδικασία της φυσικής επιλογής για να κατευθύνει τις πρωτεΐνες ή τα νουκλεϊνικά οξέα προς ένα στόχο καθορισμένο από το χρήστη με την εφαρμογή επαναληπτικών κύκλων μετάλλαξης (Lutz 2010, Cobb *et al.* 2013). Τελευταία έχει γίνει ένα κοινό εργαστηριακό εργαλείο για τη μεταβολή και τη βελτιστοποίηση της πρωτεϊνικής λειτουργίας (Arnold 2018) Σε αυτό το σημείο προτείνεται ως εργαλείο για την κατανόηση των εξελικτικών σχέσεων εντός των οικογενειών που εξετάστηκαν μέσω της

επιλεκτικής μεταλλαξιγένεσης καταλοίπων. Παρότι η επέκταση των διδαγμάτων από την κατευθυνόμενη εξέλιξη στη φυσική εξέλιξη απαιτεί προσοχή, επειδή αυτές οι διαδικασίες αναζήτησης λειτουργούν υπό διαφορετικές χρονικές κλίμακες, μεγέθη πληθυσμού, ποσοστά μετάλλαξης, δύναμη επιλογής κλπ. και παρότι η φυσική εξέλιξη λειτουργεί σε διαφορετικό φυσικό περιβάλλον, και οι διαφορές αντικατοπτρίζουν συνέπειες των αλληλεπιδράσεων μεταξύ της πρωτεΐνης και του κυτταρικού περιβάλλοντος που μπορεί να περιλαμβάνουν περιορισμούς σχετιζόμενους με τη ρύθμιση, τις μη ειδικές αλληλεπιδράσεις ή άλλους παράγοντες, η αποσύνδεση μιας πρωτεΐνης από την *in vivo* λειτουργία της είναι ένα πλεονέκτημα της κατευθυνόμενης εξέλιξης. διότι επιτρέπει τη διερεύνηση φυσικά επιτρεπτών πρωτεϊνών (Romero & Arnold 2009).

4.8 Προοπτική σχεδιασμού φαρμάκων με βάση τη δομή.

Ο βάκιλλος του άνθρακα *B. anthracis* είναι αιτιολογικός παράγοντας λοιμώδους ανθρωποζωονόσου ο αντίκτυπος της οποίας έχει περιγραφεί από την αρχαία βιβλιογραφία, που χρονολογείται πάνω από 2000 χρόνια, και εξακολουθεί να έχει τεράστιο ενδιαφέρον για την τρέχουσα επιστημονική έρευνα.

Αποτελεί μείζονα προτεραιότητα στην έρευνα για την αντιμετώπιση των απειλητικών για τη ζωή λοιμώξεων, η ανάπτυξη νέων αντιβακτηριακών παραγόντων για την καταπολέμηση της επιδείνωσης της ανθεκτικότητας στα αντιβιοτικά που παρουσιάζουν τα βακτήρια.

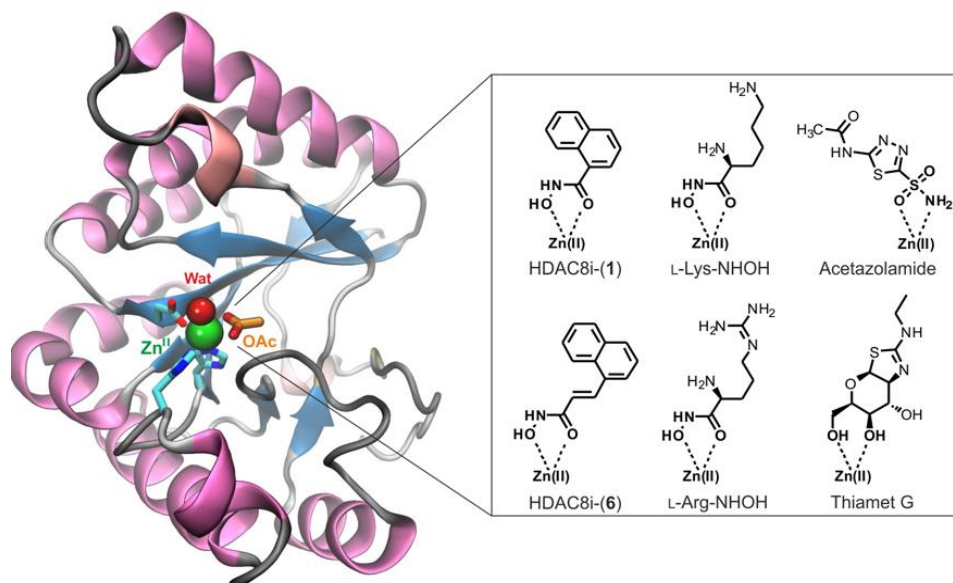
Η μοναδική φύση του κυτταρικού τοιχώματος των βακτηρίων, επιτρέπει να χρησιμοποιηθούν θεραπευτικά ειδικοί αναστολείς με λίγες μόνο ανεπιθύμητες παρενέργειες για τους ασθενείς. Ένας σημαντικός αριθμός αντιμικροβιακών που στοχεύουν στο βακτηριακό κυτταρικό τοίχωμα δρα μέσω των κυτταροπλασματικών προδρόμων των κυτταρικών τοιχωμάτων και του βιοσυνθετικού μηχανισμού τους (Nikolaidis *et al.* 2014). Ωστόσο, μια άλλη κατηγορία μορίων θα μπορούσε να δρα ως παράγοντες που στοχεύουν εξωκυτταρικές πρωτεΐνες που είναι προσδεμένες στη μεμβράνη, οι οποίες είναι υπεύθυνες για τροποποιήσεις βακτηριακών κυτταρικών τοιχωμάτων. Οι απακετυλάσες της πεπτιδογλυκάνης ως τα εργαλεία για την τροποποίησή της αποτελούν κρίσιμα ένζυμα των βακτηρίων και έχουν αποτελέσει αντικείμενο εντατικής έρευνας στα πεδία της βιοχημείας, της δομικής βιολογίας και της φαρμακευτικής χημείας. Η λεπτομερής γνώση της δομής τους αναμένεται να

βοηθήσει στην κατανόηση του φυσιολογικού τους ρόλου, του μηχανισμού λειτουργίας τους, αλλά και στον ορθολογικό σχεδιασμό νέων αποτελεσματικών και επιλεκτικών αναστολέων της απακετυλάσης της πεπτιδογλυκάνης για την επιτυχή θεραπεία των λοιμώξεων που προκαλούν τα κατά Gram-θετικά βακτήρια.

Η Bc1974 είναι μια εξαρτώμενη από μέταλλο PG GlcNAc απακετυλάση με μια μάλλον ασυνήθιστη εξειδίκευση υποστρώματος, καθώς είναι δραστική για ένα ευρύ φάσμα υποστρωμάτων (Psylinakis *et al.* 2005, Balomenou *et al.* 2013) και η οποία είναι ορθόλογη με την Ba1977 του *B. anthracis* (98,5% ταυτότητα για την περιοχή NodB). Έχει αποδειχθεί προηγουμένως ότι η Ba1977 σχετίζεται με την πλευρική τροποποίηση της πεπτιδογλυκάνης και ότι εμπλέκεται στην αντίσταση του *B. anthracis* στην λυσοζύμη του ξενιστή, είναι εντοπισμένη στην κυτταρική μεμβράνη και η έκφρασή της απαιτείται για την πλήρη μολυσματικότητα του *B. anthracis* (Balomenou *et al.* 2013).

Στην παρούσα έρευνα διεξήχθησαν δομικές μελέτες συμπλόκων της Bc1974 του *B. cereus*, με μια ποικιλία προσδετών (ligands). Παρουσιάστηκαν οι κρυσταλλικές δομές ακτίνων-X της Bc1974 από τον *B. cereus* σε μη προσδεδεμένη μορφή αλλά και σε σύμπλοκα με δύο αμινοξικά υδροξυαμικά οξέα, δύο αναστολείς της απακετυλάσης της ιστόνης (HDAC), έναν αναστολέα καρβονικής ανυδράσης και έναν αναστολέα της ανθρώπινης O-GlcNAcαση, όλα σε συναρμογή με το καταλυτικό ιόν ψευδαργύρου (**Εικόνα 4.15**). Αυτές αποτελούν τις πρώτες κρυσταλλικές δομές συμπλόκων μιας απακετυλάσης (GlcNAc) της πεπτιδογλυκάνης με προσδέτες μικρού μοριακού βάρους, ένας από τους οποίους παρουσιάζει και ανασταλτική δραστηριότητα του ενζύμου Bc1974. Είναι ενδιαφέρον ότι οι δομές της Bc1974 αποκαλύπτουν ότι ένας βρόχος του ενεργού κέντρου (μοτίβο-4 ή MT4) μπορεί να υιοθετήσει τουλάχιστον δύο διακριτές διαμορφώσεις, υποδεικνύοντας μια εγγενή ευκαμψία αυτής της περιοχής που προηγουμένως παρατηρήθηκε μόνο στις δομές που συνδέονται με το υπόστρωμα της απακετυλάσης χιτίνης (CDA) του *Vibrio cholerae* (Andrés *et al.* 2014, Giastas *et al.* 2018). Με χρήση υπολογιστικής πρόβλεψης ελλιμενισμού (molecular docking) έγινε αρχικά ο προσδιορισμός της δομής ενός μοντέλου υποστρώματος προσδεδεμένου στην καταλυτική περιοχή της Bc1974. Με βάση αυτό το μοντέλο, προτάθηκε ο καταλυτικός μηχανισμός της Bc1974 και έγινε μια υπόθεση για το ρόλο των δύο διακριτών διαμορφώσεων της περιοχής MT4 στην διευκόλυνση της δέσμευσης του υποστρώματος PG και στην κατάλυση της αντίδρασης απακετυλίωσης (Giastas *et al.* 2018). Και οι δύο υποθέσεις είναι σύμφωνες με αυτές που είχαν προταθεί για την CDA του *V. cholerae*, και οι οποίες βασίστηκαν σε πειραματικές κρυσταλλογραφικές δομές (Andrés *et al.* 2014). Τα αποτελέσματα που παρουσιάζονται εδώ παρέχουν σημαντικές δομικές πληροφορίες που θα μπορούσαν να

θέσουν τη βάση για την ανάπτυξη αντιβακτηριδιακών παραγόντων που στοχεύουν την Bc1974 και την ορθόλογη της Ba1977.



Εικόνα 4.15 Μοριακή απεικόνιση της δομής της Bc1974 και οι συντακτικοί τύποι των υποκαταστατών/προσδετών στους οποίους φαίνεται η ομάδα πρόσδεσης, που επιλύθηκαν στην παρούσα εργασία. Το καταλυτικό ιόν ψευδαργύρου απεικονίζεται ως πράσινη σφαίρα και τα αμινοξικά κατάλοιπα συναρμογής του μετάλλου φαίνονται με κυανά άτομα C. Ένα μόριο οξικού (OAc) που βρέθηκε σε συναρμογή με τον ψευδάργυρο στην πρωτεΐνη άνευ προσδέτη απεικονίζεται με πορτοκαλί άτομα C, ενώ όλα τα άλλα άτομα είναι χρωματισμένα με μπλε για το N και κόκκινο για το O. Ένα μόριο ύδατος (Wat) που επίσης βρέθηκε στην διάταξη συναρμογής του ιόντος Zn²⁺ απεικονίζεται ως κόκκινη σφαίρα.

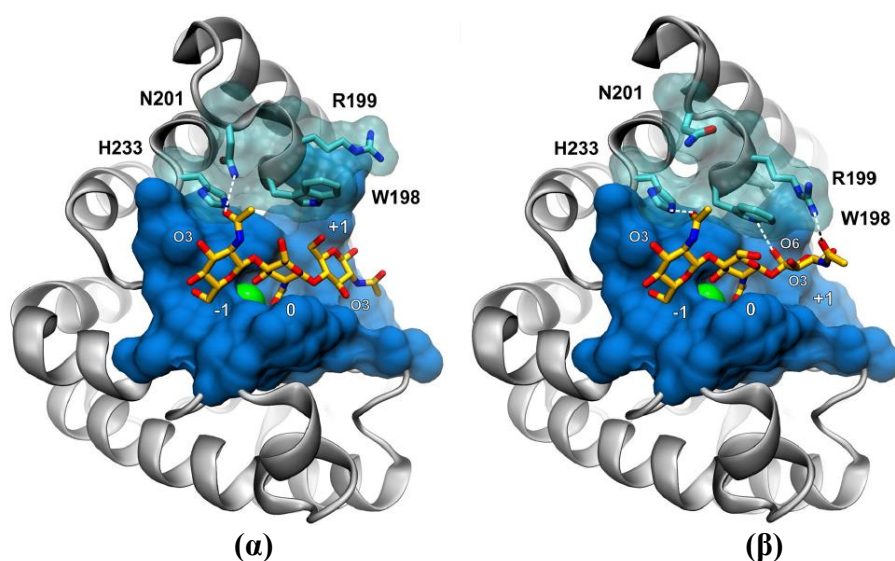
4.8.1 Αλληλεπιδράσεις υποστρώματος.

Με απώτερο σκοπό να προταθεί ένας καταλυτικός μηχανισμός για την Bc1974, υπολογίστηκε πρώτα μια πιθανή διαμόρφωση σύνδεσης ενός μοντέλου-υποστρώματος (GlcNAc)₃ στην καταλυτική κοιλότητα. Τόσο το μόριο A όσο και το B της ασύμμετρης κρυσταλλογραφικής μονάδας χρησιμοποιήθηκαν στους υπολογισμούς πρόσδεσης με το πρόγραμμα AutoDock 4 (Morris *et al.* 2009), προκειμένου να συμπεριληφθούν οι διαφορετικές διαμορφώσεις του καταλυτικού βρόχου MT4. Στις αντιπροσωπευτικές διαμορφώσεις του υποστρώματος που δεσμεύεται στα δυο μόρια της Bc1974 A και B

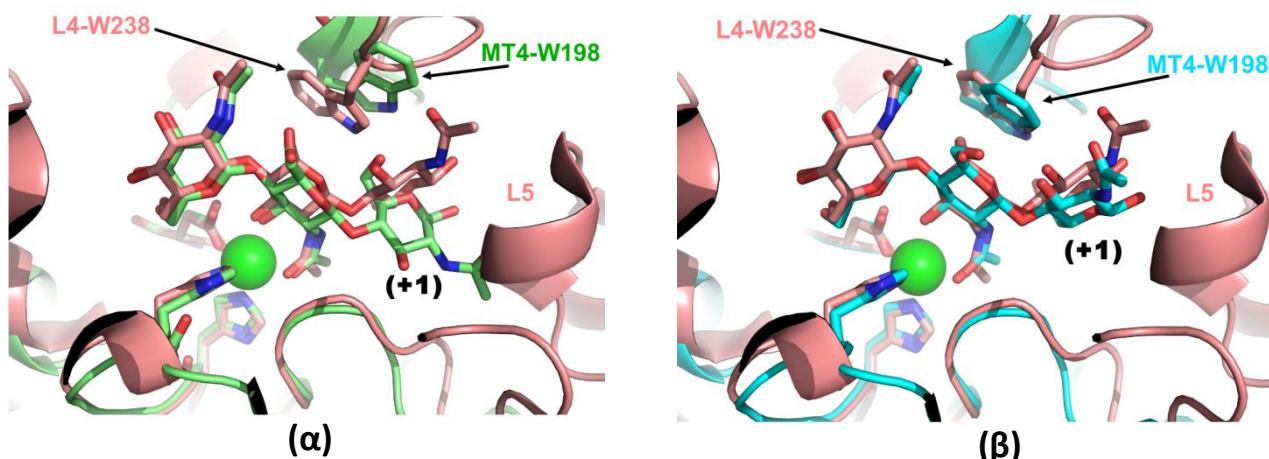
(**Εικόνα 4.16 α και β**), η *N*-ακετυλομάδα της κεντρικής γλυκάνης (σημειώνεται ως 0) προβλέπεται ότι συνδέεται με τον καταλυτικό ψευδάργυρο μέσω του καρβονυλίου της. Τα πλευρικά σάκχαρα (σημειώνονται ως -1 και +1) καταλαμβάνουν έναν αύλακα που διατρέχει όλη τη πρόσοψη της πρωτεΐνης και είναι τοποθετημένα ώστε τα άτομα O3 να δείχνουν προς το εξωτερικό του αύλακα, δίνοντας πρόσβαση στο μεγάλο μόριο της πεπτιδογλυκάνης. Οι διαμορφώσεις που προέβλεψε το πρόγραμμα για τα (0) και (-1) τμήματα «ταιριάζουν» αρκετά καλά και στα δυο μόρια A και B της Bc1974. Η ομάδα καρβονυλίου του (0) σταθεροποιείται μέσω ενός δεσμού-H με την αμινομάδα της Tyr167, ενώ το αμιδίο της *N*-ακετυλομάδας του βρίσκεται σε εγγύτητα με το δακτύλιο του ιμιδαζολίου της His230. Επιπλέον, η GlcNAc (0) εμφανίζει έναν αριθμό αρωματικών αλληλεπιδράσεων C^π-H με την ομάδα ινδολίου της Trp198. Η μόνη διαφορά στις αλληλεπιδράσεις του GlcNAc με τα δυο μόρια εντοπίζεται στη μονάδα (-1), όπου στο μόριο A η καρβονυλομάδα σχηματίζει δεσμό H με την πλευρική αλυσίδα του αμιδίου της Asn201 (**Εικόνα 4.16 α**), ενώ στο μόριο B με το ιμιδαζόλιο της His233 (**Εικόνα 4.16 β**). Αυτή η διαφορά είναι αποτέλεσμα της πιο κλειστής διαμόρφωσης του MT4 που παρατηρείται στο μόριο B, στο οποίο τα αμινοξέα Trp198 και Arg199 περικλείουν τη μονάδα (+1), ενώ το κατάλοιπο Asn201 μετακινείται μακριά από τον αύλακα πρόσδεσης του υποστρώματος. Ως αποτέλεσμα, οι διαμορφώσεις του συνδεδεμένου GlcNAc (+1) διαφέρουν σημαντικά μεταξύ των μορίων A και B της Bc1974, υιοθετώντας σχεδόν κάθετες διευθετήσεις (**Εικόνα 4.16 α και β**). Στο μόριο A, η μοναδική πολική αλληλεπίδραση προβλέπεται μεταξύ της καρβονυλικής ομάδας και του Asp194 (δεν απεικονίζεται), αλλά στο B, η μονάδα (+1) σταθεροποιείται από επιπλέον δυο υδρογονικούς δεσμούς μεταξύ του O6 και του ινδολικού NH της Trp198, καθώς και μεταξύ του καρβονυλίου της *N*-ακετυλομάδας και της γουανιδινομάδας της Arg199.

Αυτές οι πιθανές διαμορφώσεις του υποστρώματος (GlcNAc)₃ οδηγούν στην υπόθεση πως τα μόρια A (όπως ακριβώς και το C) της Bc1974 στην κρυσταλλική δομή είναι αντιπροσωπευτικές καταστάσεις υποδοχής που διευκολύνουν την αρχική πρόσδεση του μεγάλου PG υποστρώματος και επιτρέπουν την ορθή οργάνωση της κεντρικής μονάδας της γλυκάνης εντός του ενεργού κέντρου. Η πιο «κλειστή» διαμόρφωση της MT4 στην περιοχή πρόσδεσης (+1) που παρατηρήθηκε στα μόρια B και D, υποθέτουμε ότι θα σταθεροποιεί το υπόστρωμα περαιτέρω, προάγοντας την ενζυμική δραστηριότητα. Προκειμένου να αξιολογηθούν τα μοντέλα πρόσδεσης και οι υποθέσεις των λειτουργικών ρόλων των διακριτών διαμορφώσεων της MT4, πραγματοποιήθηκε δομική υπέρθεση του μοντέλου του συμπλόκου Bc1974 – (GlcNAc)₃ με τη δομή του συμπλόκου της VcCDA – τριακετυλχιτοτριόζη (VcCDA– DP3) (Andrés *et al.* 2014). Ενδιαφέρον παρουσιάζει η

παρατήρηση ότι τα ενεργά κέντρα των δυο ενζύμων έχουν εκπληκτικές ομοιότητες, ιδιαίτερα στο κατώτατο τμήμα του αύλακα πρόσδεσης, παρά τις σημαντικές διαφοροποιήσεις στους καταλυτικούς βρόχους (**Εικόνα 4.17 α και β**). Σύγκριση της MT4 διαμόρφωσης της Bc1974 με την DP3-VcCDA δομή έδειξε πως οι αρωματικοί δακτύλιοι ινδολίου στα κατάλοιπα Trp198 στην Bc1974 και Trp238 στη VcCDA συμπίπτουν χωρικά, λαμβάνουν παρόμοιες διευθετήσεις και συμβάλλουν στην σταθεροποίηση των σακχάρων που καταλαμβάνουν τη μονάδα του σακχάρου (+1) (**Εικόνα 4.17β**). Έτσι, η θέση του (GlcNAc)₃ που υπολογίστηκε στη Bc1974 συμπίπτει σχεδόν τέλεια με τη δεσμευμένη DP3 στη VcCDA στις θέσεις (-1) και (0), ενώ διαφέρει εντός εύρους από 1,2 έως 2,1 Å για τα αντίστοιχα άτομα του δακτυλίου των σακχάρων στη θέση (+1), καθώς ο προσανατολισμός των τμημάτων σακχάρου είναι ο ίδιος (**Εικόνα 4.17β**). Το τελευταίο χαρακτηριστικό δεν παρατηρείται στην ανοικτή διαμόρφωση της MT4 της Bc1974, όπου η Trp198 της MT4 παραμένει περιφερική, που οδηγεί σε μια σχεδόν κάθετη διευθέτηση της (+1) θέσης κατάληψης σακχάρου (**Εικόνα 4.17α**).



Εικόνα 4.16 Μοριακά μοντέλα του συμπλόκου Bc1974–(GlcNAc)₃ που απεικονίζει (**α**) την ανοικτή διαμόρφωση της κοιλότητας δέσμευσης του υποστρώματος στο μόριο A σε σχέση (**β**) με το B. Τα άτομα C του (GlcNAc)₃ δείχνονται με πορτοκαλί ενώ των καταλοίπων της Bc1974 δείχνονται με κυανό.

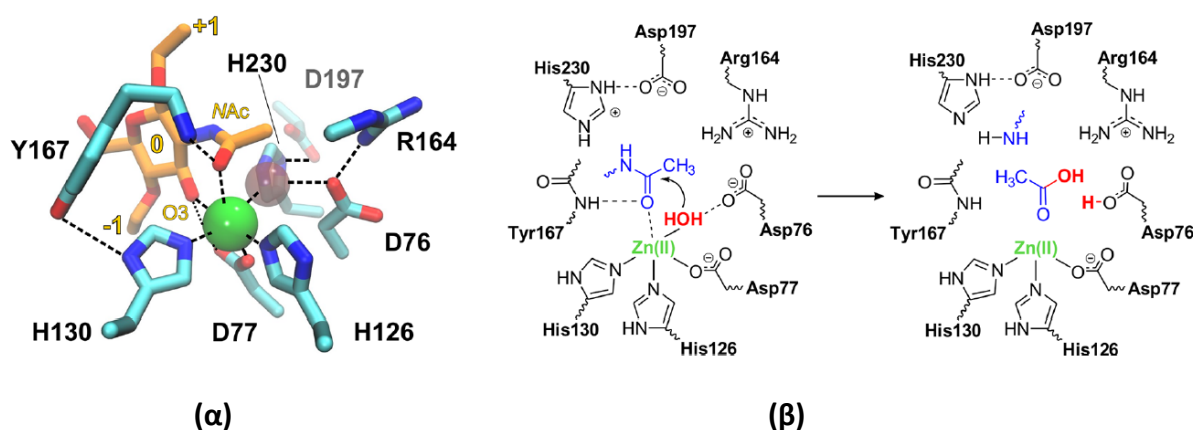


Εικόνα 4.17 Υπέρθηση της κρυσταλλικής δομής της VcCDA με σύμπλοκο με τριόζη (ροζ) (PDB ID: 4OUI) με τα μοριακά μοντέλα του (GlcNAc)₃ στην κοιλότητα δέσμευσης της Bc1974 (πράσινο ή κυανό) (PDB ID: 5N1J). **(α)** Η περιοχή MT4 της Bc1974 βρίσκεται στην ανοικτή διαμόρφωσή της όπως βρέθηκε στα μόρια A και C της κρυσταλλικής δομής, ενώ στην **(β)** βρίσκεται στην κλειστή διαμόρφωσή της που βρίσκεται στα μόρια B και D. Μόνο η W198 της περιοχής MT4 της Bc1974 και η W238 του βρόχου 4 της VcCDA παρουσιάζονται (σε αναπαράσταση ράβδων), ενώ μερικοί βρόχοι της VcCDA έχουν παραλειφθεί για λόγους σαφήνειας. Εμφανίζεται ο εκτεταμένος βρόχος 5 της VcCDA ο οποίος είναι κοντά στην υπο-θέση (+1) της κοιλότητας δέσμευσης. Στην κλειστή διαμόρφωση της κρυσταλλικής δομής Bc1974, η W198 συμπίπτει χωρικά με το ισοδύναμο αμινοξικό κατάλοιπο W238 της VcCDA, όταν το τελευταίο κρυσταλλώθηκε με το μόριο υποστρώματος χιτοτριόζης. Επιπρόσθετα, το αποτέλεσμα των υπολογισμών πρόσδεσης του μοντέλου-υποστρώματος (GlcNAc)₃ υπερτίθεται σχεδόν τέλεια με το πειραματικά προσδιορισμένο υπόστρωμα που δεσμεύεται στην VcCDA στις θέσεις (-1) και (0), και παρουσιάζει απόκλιση στη θέση (+1) εντός ενός εύρους 1,2–2,1 Å, αν και ο προσανατολισμός του τμήματος του σακχάρου είναι ο ίδιος. Αντίθετα, όταν συγκρίνονται η ανοικτή διαμόρφωση της Bc1974 (α) με τη συμπλοκοποιημένη δομή της VcCDA με τη χιτοτριόζη, οι αποκλίσεις είναι πιο έντονες υποδεικνύοντας ότι η κλειστή διαμόρφωση του MT4 πιθανώς αντιστοιχεί στην καταλυτική δραστική κατάσταση της Bc1974.

4.8.2 Προτεινόμενος καταλυτικός μηχανισμός.

Ο ευρέως αποδεκτός καταλυτικός μηχανισμός των εξαρτώμενων από μέταλλο (ψευδάργυρο) *N*-απακετυλασών ξεκινάει από την πυρηνόφιλη προσβολή ενός μορίου ύδατος στον αμιδικό C του μεσαίου σακχάρου (ο υπο-τόπος 0 στην **Εικόνα 4.18**). Το πολωμένο μόριο του διαλύτη ενεργοποιείται από ένα συντηρημένο όξινο κατάλοιπο, το οποίο είναι σωστά προσανατολισμένο κοντά στο ιόν του ψευδαργύρου και δρα ως βάση που σταθεροποιεί το τετραεδρικό ενδιάμεσο (Whittington *et al.* 2003). Στην περίπτωση της

Bc1974, το μόριο διαλύτη που είναι συνδεδεμένο στον ψευδάργυρο σχηματίζει ένα δεσμό υδρογόνου με το κατάλοιπο Asp76, το οποίο δρα ως βάση στον προτεινόμενο καταλυτικό μηχανισμό (Εικόνα 4.18α). Το Asp76 σταθεροποιείται μέσω αλληλεπίδρασης γεφυρών άλατος με την ομάδα γουανιδινίου της Arg164, η οποία με τη σειρά της σταθεροποιείται μέσω δεσμού υδρογόνου με την πλευρική αλυσίδα της Thr74. Κατ' αναλογία με τον καταλυτικό μηχανισμό της θερμολυσίνης (Matthews 1988) η καρβοξυλική ομάδα του Asp197 δρα ως δέκτης δεσμού υδρογόνου από τη θετικά φορτισμένη His230, μια αλληλεπίδραση που εξασφαλίζει το σωστό προσανατολισμό της πλευρικής αλυσίδας του ιμιδαζολίου έτσι ώστε να δρα ως οξύ και πιθανότατα ως δότης πρωτονίων στην αποχωρούσα αμινομάδα (Εικόνα 4.18β).



Εικόνα 4.18 (α) Το ενεργό κέντρο του μοντέλου Bc1974-(GlcNAc)₃. Οι αλληλεπιδράσεις μεταξύ του μεσαίου σακχάρου (0) του μοντέλου υποστρώματος, του δεσμευμένου με τον ψευδάργυρο μορίου ύδατος και των καταλοίπων του ενεργού κέντρου υποδεικνύονται με διακεκομμένες γραμμές. **(β)** Προτεινόμενος καταλυτικός μηχανισμός της Bc1974, που ξεκινά από την πυρηνόφιλη προσβολή ενός μορίου ύδατος (ή ανιόντος ⁻OH) στον καρβονυλικό άνθρακα της *N*-ακετυλομάδας.

Αυτός ο μηχανισμός είναι σε άριστη συμφωνία με τον προτεινόμενο μηχανισμό για την GlcNAc απακετυλάση της πεπτιδογλυκάνης *SpPgdA* (Blair *et al.* 2005) του *Streptococcus pneumoniae*. Σε εκείνη την εργασία οι ερευνητές έδειξαν μέσω πειραμάτων μεταλλαξιγένεσης ότι η ανάλογη αλληλεπίδραση μεταξύ της His417 (His230 στην Bc1974) και του Asp391 (Asp197 στην Bc1974) ρυθμίζει την pKa της ιστιδίνης που είναι απαραίτητη στον καταλυτικό μηχανισμό.

Τα δομικά ευρήματα της παρούσας μελέτης υποδεικνύουν ότι η περιοχική δέσμευση της απακετυλάσης δε μεταβλήθηκε με αλλαγή στην διαμόρφωση της κύριας αλυσίδας, ή με περιστροφές στις πλευρικές ομάδες κατά τη σύνδεση διαφόρων ενώσεων στην καταλυτική κοιλότητα ψευδαργύρου, υποδηλώνοντας έτσι την ακαμψία της θέσης πρόσδεσης. Εν

τούτοις, όλοι οι προσδέτες που συγκρυσταλλώθηκαν με την Bc1974 στην παρούσα μελέτη ήταν σχεδόν ίσου μεγέθους και στις περισσότερες περιπτώσεις σταθεροποιήθηκαν στις θέσεις τους μέσω αλληλεπιδράσεων με κατάλοιπα των περιοχών MT1-MT3 και σε μικρότερο βαθμό με κατάλοιπα των περιοχών MT4 και MT5. Συγκεκριμένα, η περιοχή MT4 είχε μόνο μικρή συμβολή στη σταθεροποίηση των ενώσεων που στερούνται συγκεκριμένων λειτουργικών ομάδων. Πράγματι, το κατάλοιπο Trp198 που φέρει έναν αρωματικό δακτύλιο, το οποίο αναδύεται από την περιοχή MT4 προς την κοιλότητα δέσμησης, συμμετείχε στην σύνδεση των ενώσεων **1** και **6** (στην αλυσίδα B) μέσω των αρωματικών αλληλεπιδράσεων π-π του δακτυλίου ινδολίου με του δακτυλίου ναφθαλενίου. Εδώ θα πρέπει να αναφερθεί ότι η ικανότητα αναστολής του προσδέτη **1** ήταν υψηλή ακόμη και σε συγκεντρώσεις 62 μM, ενώ στα 500 μM εξαλείφθηκε η ενζυμική δραστηριότητα όπως προσδιορίστηκε από πειράματα δοκιμασιών αναστολής του ενζύμου *in vitro* με αυτούς τους προσδέτες καθώς και πειράματα δέσμησης με θερμιδομετρία ισόθερμης τιτλοδότησης (Isothermal Titration Calorimetry, ITC) (Giastas *et al.* 2018). Επιπροσθέτως, η ίδια τρυπτοφάνη (Trp198) συνέβαλε στη σταθεροποίηση της ομάδας γουανιδίνης της ένωσης υδροξαμικής-L-αργινίνης μέσω αλληλεπιδράσεων κατιόντος-αρωματικού π-νέφους, υποβοηθούμενων από την αρωματική πλευρική αλυσίδα της Tyr167. Ωστόσο, η ομάδα ναφθαλενίου της ένωσης **6** που βρίσκεται δεσμευμένη στην αλυσίδα D, αντί να αλληλεπιδρά με την Trp198 μέσω παράλληλης στοιβαξης, σχημάτισε αρωματικές αλληλεπιδράσεις σχήματος T (κάθετης διευθέτησης) και πιθανώς αλληλεπιδράσεις κατιόντος-π με την Arg199 της περιοχής MT4. Η τελευταία δείχνει ότι είναι δυνατή και η αντίστροφη συμμετοχή ανάμεσα στο ζεύγος ενζύμου-προσδέτη όσον αφορά την αλληλεπίδραση κατιόντος-π νέφους. Η διαφορετική δέσμηση της ένωσης **6** έδειξε ότι η προσθήκη ακόμη και της μικρής ομάδας του αιθυλενίου μεταξύ της δραστικής ομάδας συναρμογής του ψευδαργύρου και της αρωματικής ομάδας του υποκαταστάτη προκάλεσε την έλλειψη εξειδίκευσης του προσδέτη για το ένζυμο. Τέλος, βρέθηκε μια άλλη αλληλεπίδραση κατιόντος-π νέφους μεταξύ της πρωτοταγούς αμίνης της υδροξαμικής-L-λυσίνης και της His230. Αν και αυτή η αλληλεπίδραση διέπεται από περιορισμούς λόγω της κρυσταλλικής μοριακής διευθέτησης, είναι πιθανό να συμβεί στο απομονωμένο ένζυμο, καθώς η κρυσταλλική διευθέτηση δεν επηρέασε τη διαμόρφωση των εμπλεκόμενων καταλοίπων. Επιπλέον, ίσως να έχει βιοχημική σημασία αφού η εμφάνισή της θα μπορούσε να μεταβάλει την ισορροπία προς όφελος της πρωτονιωμένης μορφής του καταλοίπου His230, ρυθμίζοντας έτσι τη δραστηριότητα του ενζύμου.

Για να προσδιοριστεί εάν υπάρχει προτιμητέος προσανατολισμός κατά τη σύνδεση ενώσεων στην κοιλότητα πρόσδεσης της Bc1974, οι δεσμευμένες με προσδέτη δομές

υπερτέθηκαν και υπολογίστηκε η γεωμετρία δέσμευσης του ψευδαργύρου. Σε όλες τις περιπτώσεις, το επίπεδο που ορίζεται από τη δισθενή ομάδα και τον ψευδάργυρο τέμνει το επίπεδο της τριάδας των καταλοίπων His-His-Asp υπό πολύ παρόμοιες γωνίες, καταδεικνύοντας μια βέλτιστη γεωμετρία δέσμευσης που μπορεί να χρησιμοποιηθεί ως περιορισμός σε μελλοντικές *in silico* μελέτες αναστολέων των απακετυλασών της πεπτιδογλυκάνης.

Σε προηγούμενες μελέτες της SpPgdA, χρησιμοποιώντας στοχευμένη υπολογιστική αξιολόγηση, ακολουθούμενη από βιοχημική αξιολόγηση, εντοπίστηκαν δύο υποσχόμενοι αναστολείς με τιμές IC₅₀ 580 μM και 130 μM (Bui *et al.* 2011). Και οι δύο ενώσεις περιείχαν αρωματικές ομάδες κοντά στην ομάδα συναρμογής του δισθενούς ψευδαργύρου και προτάθηκαν πιθανές θέσεις υποκατάστασης του βενζολικού δακτυλίου προς ενίσχυση της ικανότητας αναστολής. Τα πειραματικά δεδομένα που προέκυψαν ευθυγραμμίζονται με τα ανωτέρω αποτελέσματα και επιπλέον υποδεικνύουν ότι η ύπαρξη μακράς αλειφατικής αλυσίδας που προηγείται της αρωματικής ομάδας δεν ευνοεί τη δραστικότητα των αναστολέων, όπως παρατηρήθηκε στην περίπτωση της ένωσης **7** (SAHA). Επιπλέον, σε μια σχετικά πρόσφατη μελέτη των HDAC8, η οποία περιλάμβανε κρυσταλλογραφικές μελέτες και μελέτες *in silico*, χρησιμοποιήθηκε ένα υποσύνολο των υδροξαμικών οξέων που περιλαμβάνονται και στην παρούσα έρευνα (Tabackman *et al.* 2016). Η δομή της HDAC8 με την ένωση **3** (Εικόνα 3.40) που επιλύθηκε (ένωση **6** στην εργασία των Tabackman *et al.*) έδειξε με έμφαση τη σημασία ενός δι-αρυλικού συστήματος χωρίς πλευρική αλειφατική ομάδα στην δέσμευση της ιστόνης με υψηλή συγγένεια, καθώς αμφότερες οι ομάδες βενζολίου και ναφθαλενίου σταθεροποιήθηκαν μέσω ισχυρών αλληλεπιδράσεων με τα αρωματικά κατάλοιπα της HDAC8

4.8.2.1 Δύο διακριτές διαμορφώσεις της καταλυτικής περιοχής MT4

Όπως αναφέρθηκε στην §3.4.4 η Bc1974 έχει 4 μόρια στην ασύμμετρη μονάδα. Η μέγιστη απόκλιση μεταξύ των τεσσάρων μορίων της ασύμμετρης μονάδας παρατηρείται στην περιοχή των αμινοξέων 197–205, τα οποία απαρτίζουν το μοτίβο 4 (MT4). Ενώ τα πρωτεϊνικά μόρια A και C υιοθετούν μια πανομοιότυπη διαμόρφωση στην συγκεκριμένη περιοχή του MT4, διαφέρουν ουσιαστικά από τις αντίστοιχες περιοχές των μορίων B και D. Υπέρθωση των αλυσίδων A και B (ή C και D) οδηγούν σε μια απόκλιση r.m.s.d. = 0,28 Å των ατόμων της κύριας αλυσίδας εκτός της περιοχής MT4, ενώ τα C^a άτομα των αμινοξέων

Trp198 και Arg199 της MT4 διαφέρουν κατά 1,0 Å και 1,2 Å μεταξύ των δυο αλυσίδων, αντίστοιχα. Το μέγιστο της απόκλισης εντοπίζεται στην Lys202 φτάνοντας στην τιμή των 4,2 Å για τα C^α άτομα. Δεν είναι η πρώτη φορά που ένας καταλυτικός βρόχος ενός CE4 ενζύμου εμφανίζεται σε δυο διακριτές διαμορφώσεις. Όπως αναφέρθηκε, οι Andres *et al.* (2014) παρουσίασαν τις δομές των VcCDA δεσμευμένων σε υποστρώματα, οι οποίες έδειξαν μεγάλες μεταβολές στην διαμόρφωση του βρόχου 4 (παρόμοια με την περιοχή MT4 εδώ) σε σχέση με το ελεύθερο ένζυμο. Η παρόμοια δομική διαφοροποίηση μεταξύ των αλυσίδων A, C και B, D στην παρούσα μελέτη μπορεί να αποδοθεί στις κρυσταλλικές επαφές μεταξύ των μορίων της Bc1974. Ειδικότερα, η His269 των αλυσίδων B και D διεισδύει στον καταλυτικό αύλακα των αλυσίδων C και A αντίστοιχα, και σταθεροποιείται μέσω π-π αλληλεπιδράσεων με το ινδόλιο της Trp198. Ως εκ τούτου, όλη η MT4 περιοχή αποκλίνει από το ενεργό κέντρο, το οποίο επάγει μια σημαντική αύξηση του όγκου της καταλυτικής κοιλότητας κατά ~7% (Εικόνα 4.19). Δεδομένου ότι οι δυο διαφορετικές διαμορφώσεις της MT4 στις παρούσες δομές δεν ήταν σε συναρμογή με κάποιον προσδέτη, αποτελούν ενδείξεις μιας έμφυτης ευκινησίας της περιοχής MT4, η οποία μπορεί να παίζει ρόλο στην αναγνώριση του υποστρώματος, στην δέσμευση του και τελικά στην ενζυμική δράση της Bc1974.

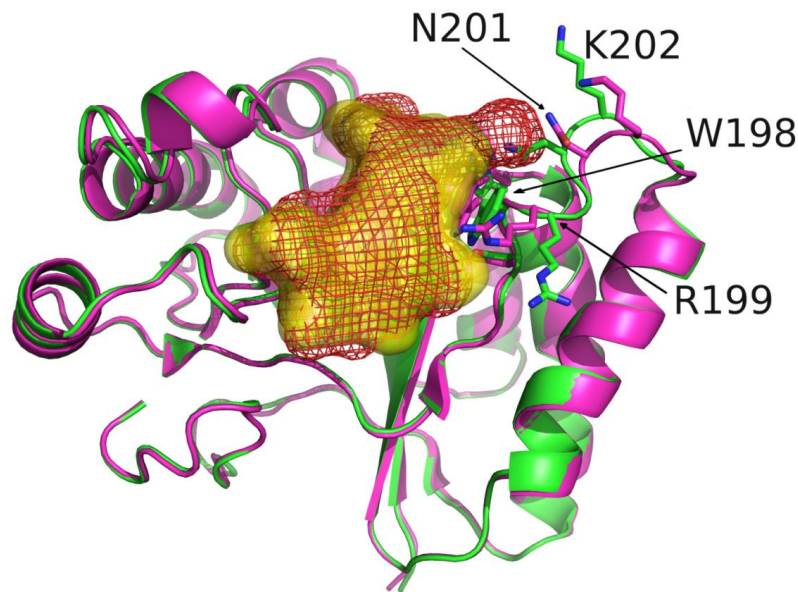
Καμμία από τις έξι ενώσεις που προσδιορίστηκαν κρυσταλλογραφικά δεν προκαλεί σημαντικές διαμορφωτικές αλλαγές των περιοχών που συμμετέχουν στην δέσμευση (μοτίβων MT1-MT5) μεμονωμένων καταλοίπων όταν συγκριθούν οι δομές των συμπλόκων με τη δομή της πρωτεΐνης χωρίς προσδέτη. Η οργάνωση του ενεργού κέντρου της Bc1974 (Εικόνες 4.17β και 4.18α) ελάχιστα επηρεάζεται από τις διαφορετικές διαμορφώσεις των αμινοξέων της MT4 και επομένως και οι δυο διακριτές διαμορφώσεις μπορούν να είναι παρούσες στην δομική ενότητα NodB μιας ενεργής Bc1974. Επομένως, ήταν εκπληκτική η ανακάλυψη μιας τέτοιας σημαντικής μεταβολής της διαμόρφωσης στην περιοχή MT4, που προηγουμένως είχε θεωρηθεί ως άκαμπτη περιοχή.

Συμπερασματικά, οι δυο καταστάσεις της περιοχής MT4 υποδεικνύουν μια έμφυτη ελαστικότητα του βρόχου, με πιθανές φυσιολογικές λειτουργίες, όπως συμπεραίνετε από τις ομοιότητες της Bc1974 με τις δομές της VcCDA δεσμευμένης σε υποστρώματα, και επομένως θα πρέπει να λαμβάνονται υπόψη σε προσπάθειες σχεδιασμού αναστολέων.

Επιπλέον, οι υπολογιστικές μελέτες μοριακής πρόσδεσης έδειξαν ότι αυτή η διαμορφωτική πλαστικότητα της περιοχής MT4 είναι σε συμφωνία με τον προτεινόμενο καταλυτικό μηχανισμό για την αντίδραση απακετυλίωσης, υποδεικνύοντας μια εύλογη ενδογενή κινητικότητα του MT4. Είναι λοιπόν λογικό να υποτεθεί ότι και οι δύο προσδιορισμένες διαμορφώσεις είναι παρούσες στο ένζυμο και ότι ένας περιορισμένος

αριθμός ατομικών αλληλεπιδράσεων μεταξύ ενός προσδέτη ή ενός υποστρώματος και της Bc1974 μπορεί να προωθήσει την κλειστή διαμόρφωση. Αν συνέβαινε αυτό, η ανοικτή διαμόρφωση είναι πιθανότατα υπεύθυνη για την αρχική σύλληψη του υποστρώματος, του οποίου η κατάλυση μπορεί κατόπιν να διευκολυνθεί από ένα περαιτέρω κλείσιμο και την ενίσχυση των αλληλεπιδράσεων. Σε αυτό το πλαίσιο, μπορεί να θεωρηθεί ότι η κλειστή διαμόρφωση της μόνης γνωστής κρυσταλλικής δομής των απακετυλασών της πεπτιδογλυκάνης με ένα μονοσακχαρίτη ως μόριο που μοιάζει με το υπόστρωμα, επαληθεύει την παραπάνω υπόθεση (Blair & van Aalten 2004). Αξιοσημείωτα, καθώς οι παρατηρούμενες αλληλεπιδράσεις κυρίως της Trp198 και δευτερευόντως της Arg199 με τις ομάδες των προσδετών που μελετήθηκαν στην παρούσα έρευνα ήταν εκείνες που ποικίλουν, θα μπορούσε να αποδοθεί ένας ενισχυμένος ρόλος της περιοχής MT4 στην εξειδίκευση του προσδέτη.

Είναι αξιοσημείωτο ότι οι μικρού μοριακού βάρους προσδέτες βρέθηκαν δεσμευμένοι στο ένζυμο Bc1974 κυρίως στις αλυσίδες B και D, όπου η MT4 υιοθετεί την κλειστή διαμόρφωσή της, γεγονός που υποστήριξε την πρόβλεψη σχετικά με τους ρόλους των διακριτών διαμορφώσεων της MT4 στη σύνδεση υποστρώματος,.

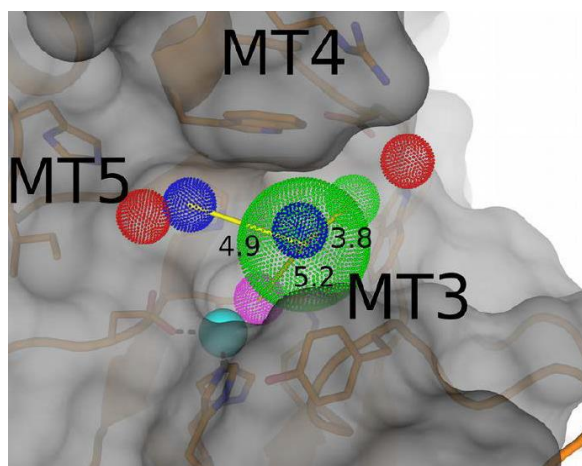


Εικόνα 4.19 Υπέρθωση των μονομερών A (φούξια) και B (πράσινο) της ασύμμετρης μονάδας της Bc1974 χωρίς προσδέτη. Η κίτρινη διαφανής επιφάνεια αντιπροσωπεύει την κοιλότητα δέσμευσης του υποστρώματος στο B, ενώ το κόκκινο πλέγμα απεικονίζει τον αυξημένο όγκο στο μονομερές A. Παρουσιάζονται επιλεγμένα κατάλοιπα της περιοχής MT4 με την Trp198 να είναι το καθοριστικό κατάλοιπο, για την παρατηρούμενη αλλαγή του όγκου της κοιλότητας.

Συνολικά, τα πειραματικά δεδομένα υποδηλώνουν ότι οι HDAC αναστολείς υδροξαμικού οξέος είναι επίσης ελπιδοφόροι αναστολείς της Bc1974 (Εικόνα 3.40). Παρότι θα πρέπει να επισημανθεί ότι τα υδροξαμικά είναι γνωστά για την τοξικότητά τους (Shena & Kozikowskia 2016). Ένας από τους ισχυρότερους αναστολείς της Bc1974, η ένωση 3, έχει πρόσφατα κρυσταλλωθεί σε σύμπλοκο με την HDAC8 (Tabackman *et al.* 2016), γεγονός που αποκάλυψε τον εξέχοντα ρόλο του ογκώδους αρυλικού τμήματος. Η δραστηριότητά της για την Bc1974 είναι πιθανώς ένα συνδυασμένο αποτέλεσμα της χηλικής ομάδας συναρμογής στον ψευδάργυρο και της άμεσης προσκόλλησής της στον αρωματικό υποκαταστάτη. Αυτή η παρατήρηση υποστηρίζεται περαιτέρω από την έλλειψη ανασταλτικής δραστηριότητας έναντι της Bc1974 που εμφανίζεται από τους πιθανούς αναστολείς της HDAC 6, 15 και 16, όπου όλοι φέρουν έναν αλειφατικό σύνδεσμο μεταξύ του υδροξαμικού οξέος και των αρυλικών τους τμημάτων (Παράρτημα Α').

4.8.3 Πρόταση φαρμακοφόρου.

Οι κρυσταλλογραφικές δομές των συμπλόκων της Bc1974 με τους έξι προσδέτες που παρουσιάστηκαν αποκαλύπτουν φερόμενα φαρμακοφορικά στοιχεία, τα οποία μπορούν να χρησιμοποιηθούν στην ανάπτυξη αντιβακτηριακών παραγόντων που στοχεύουν την Bc1974 ή/και την ορθόλογό της Ba1977 του *B. anthracis*. Οι δύο αυτές πρωτεΐνες παρουσιάζουν 98,5% ταυτότητα στην NodB ομόλογη δομική ενότητα και διαφέρουν μόνο σε δυο επιφανειακά αμινοξέα καθώς και σε ένα αμινοξύ στην περιοχή MT4 όπου η Arg199 της Bc1974, αντικαθίσταται από την όμοια σε μέγεθος και φυσικοχημικές ιδιότητες Lys199. Συλλογικά, φαίνεται να υπάρχει μια σαφής συσχέτιση μεταξύ της ανασταλτικής ισχύος και της συγγένειας του προσδέτη με τον αριθμό και τη θέση των αρωματικών καταλοίπων στους προσδέτες που μελετήθηκαν (Giastas *et al.* 2018), κάνοντας επιτρεπτή την πρόταση βασικών φαρμακοφόρων που παρουσιάζονται στην **Εικόνα 4.20**.



Εικόνα 4.20 Αναπαράσταση της επιφάνειας του ενεργού κέντρου της Bc1974 που απεικονίζει βασικά φαρμακοφορικά στοιχεία όπως αναγνωρίστηκαν από τις κρυσταλλικές δομές σε σύμπλοκα με προσδέτες. Τα κατάλοιπα της Bc1974 εμφανίζονται με ράβδους (άτομα C χρώματος πορτοκαλί), το καταλυτικό ιόν Zn^{2+} με σφαίρα χρώματος κυανού και τα φαρμακοφόρα δείχνονται με στικτές σφαίρες οι οποίες είναι χρωματισμένες ως εξής: η ομάδα δέσμευσης ψευδαργύρου παρουσιάζεται με φούξια και η μεγάλη πράσινη σφαίρα υποδεικνύει τη θέση μιας αρωματικής ομάδας που αλληλεπιδρά μέσω π-π αλληλεπιδράσεων με Tyr167, Trp191 και Tyr198. Θετικά φορτισμένες ομάδες που θα μπορούσαν να συμμετέχουν σε κατιονικές-π αλληλεπιδράσεις με κατάλοιπα των περιοχών MT4 και MT5 εμφανίζονται με μπλε χρώμα. Το τμήμα που επικαλύπτει την αρωματική περιοχή θα πρέπει να θεωρείται ως μια εναλλακτική επιλογή. Οι κόκκινες σφαίρες δείχνουν πιθανές θέσεις αρνητικά φορτισμένων ομάδων που θα μπορούσαν να αλληλεπιδράσουν με τα βασικά κατάλοιπα His230, His233, Arg199 και ακόμη και την Lys202 που εκτείνεται προς την κορυφή της καταλυτικής κοιλότητας. Οι κίτρινες γραμμές συνδέουν τα κύρια φαρμακοφόρα και οι αποστάσεις τους δίδονται σε Å.

Οι δομές της Bc1974 που συνδέονται με τις ενώσεις **1** και **6** (στην αλυσίδα B), υποδεικνύουν ότι για την επίτευξη υψηλής ικανότητας αναστολής απαιτούνται δύο φαρμακοφορικά συστατικά, εκτός από την ομάδα δέσμευσης ψευδαργύρου: α) ένας αρωματικός δακτύλιος, άμεσα συνδεδεμένος με την υδροξυαμική ομάδα, ο οποίος ευθύνεται για το μεγαλύτερο μέρος της ισχύος, πιθανώς λόγω της αλληλεπίδρασής της με την Trp198 (Εικόνα 4.19) και β) η παρουσία ενός άλλου αρωματικού δακτυλίου που εκτείνεται, είτε άμεσα είτε μετά από μια ομάδα που παρεμβάλλεται, έπειτα από την πρώτη αρωματική ομάδα. Η ενέργεια αλληλεπίδρασης μεταξύ των αρωματικών ομάδων μπορεί να ποικίλλει σημαντικά, ανάλογα με τον ακριβή προσανατολισμό και τη φύση των δακτυλίων που συμμετέχουν. Η συσσώρευση αρωματικών καταλοίπων στην καταλυτική κοιλότητα της Bc1974 σηματοδοτεί την έμφυτη τάση του ενζύμου να φιλοξενήσει προσδέτες που φέρουν αρωματικές ομάδες, η

χωρική διαμόρφωση των οποίων θα μπορούσε να διαδραματίσει αποφασιστικό ρόλο στην συγγένεια της δέσμευσης. Η δέσμευση της ένωσης **6** στην αλυσίδα D υποδηλώνει μια εύλογη χωρική διευθέτηση της δεύτερης αρωματικής ομάδας, η οποία προφανώς αυξάνει περαιτέρω την ανασταλτική ισχύ. Αυξημένη εκλεκτικότητα θα μπορούσε να επιτευχθεί με την προσθήκη φορτισμένων τμημάτων είτε θετικών για την προαγωγή των αλληλεπιδράσεων κατιόντων-π με τα αρωματικά κατάλοιπα Trp198, Trp191 και Tyr167, είτε αρνητικών για να επιτραπούν αλληλεπιδράσεις με τα θετικά φορτισμένα κατάλοιπα Arg199, His230, His233 ή ακόμη και Lys202, που εκτείνεται προς την κορυφή της κοιλότητας δέσμευσης.

Συμπερασματικά οι δομικές μελέτες της Bc1974 αποκάλυψαν νέα χαρακτηριστικά της κοιλότητας δέσμευσης (κινητικότητα της περιοχής MT4), που προηγουμένως δεν είχαν παρατηρηθεί μεταξύ των ομολόγων ενζύμων μελών της CE4, ενώ η συσχέτιση των δομών με προσδέτες οδήγησε στον προσδιορισμό δυνητικών φαρμακοφορικών στοιχείων. Περαιτέρω διερεύνηση, χρησιμοποιώντας βιοχημικές μελέτες, μεταλλαξιγένεση και δομικές μελέτες είναι απαραίτητη προκειμένου να εξακριβωθούν πρόσθετες απαιτήσεις πιθανών αναστολέων με αυξημένη ισχύ και εκλεκτικότητα.

4.9. Συμπερασματικά τα αποτελέσματα της παρούσης διατριβής

Στην παρούσα διατριβή, προσδιορίστηκαν με κρυσταλλογραφία ακτίνων-X και παρουσιάζονται οι ακόλουθες πρωτεϊνικές δομές: οι φερόμενες απακετυλάσες πολυσακχαρίτη, Ba0331 (διακριτικότητας 2,59Å) και Ba0330 (διακριτικότητας 1,48Å) από το βακτήριο *B. anthracis*, το οποίο είναι ο αιτιολογικός παράγοντας της ασθένειας του άνθρακα, καθώς και η απακετυλάση της *N*-ακετυλογλυκοζαμίνης της πεπτιδογλυκάνης, Bc1974 (διακριτικότητας 1,80Å), από το βακτήριο *B. cereus*. Τα δομικά δεδομένα έδειξαν ότι όλες ήταν μεταλλοπρωτεΐνες. Επίσης προσδιορίστηκαν και παρουσιάζονται οι κρυσταλλικές δομές συμπλόκων της Bc1974 με έξι προσδέτες, διακριτικότητας από 1,45 Å έως 3,05Å, οι οποίες αποτελούν τις πρώτες δομές συμπλόκων απακετυλάσης της πεπτιδογλυκάνης με μικρά-μόρια προσδέτες.

Τα γονιδιώματα των *B. cereus*, και *B. anthracis*, δύο συγγενικών ειδών που ανήκουν στην ομάδα *B. cereus*, έχουν πολλαπλά γονίδια απακετυλάσης πολυσακχαρίτη. Για τον καθορισμό φυλογενετικών σχέσεων στις οικογένειες γονιδίων απακετυλάσης πολυσακχαρίτη (PDAs) του *B. anthracis* Ames (12 ανοιχτά πλαίσια ανάγνωσης στο γονιδίωμά του) και του *B. cereus* ATCC14579 (11 ανοιχτά πλαίσια ανάγνωσης στο γονιδίωμά του)

χρησιμοποιήθηκαν οι αμινοξικές και οι νουκλεοτιδικές αλληλουχίες που απαρτίζουν χωριστά τις δομικές ενότητες των PDAs. Επίσης συνδυάστηκαν και με την ανάλυση γονιδίων των PDAs από άλλα στελέχη του γένους *Bacillus* προκειμένου να ερευνηθούν εξελικτικές σχέσεις των γονιδίων PDAs.

Διαδοχικά γεγονότα διπλοποίησης έχουν οδηγήσει σε ομάδες γονιδίων (ή υποοικογένειες) που εμφανίζουν μεταξύ τους μεγαλύτερη ομολογία σε σύγκριση με οποιοδήποτε άλλο γονίδιο εντός της κάθε οικογένειας. Έγινε ανίχνευση της διάδοσης της δομικής ενότητας Fn3 μέσω της φυλογενετικής ανάλυσης, ενώ η δομική μελέτη οδήγησε σε επιπρόσθετες υποθέσεις για το ρόλο της, οι οποίες συνοψίζονται στα ακόλουθα: είτε στο να προσδώσει ιδιότητες προσκόλλησης/αλληλεπίδρασης βάση συγκεκριμένων διαφορών στην δομή των δύο πρωτεϊνών και σε συνδυασμό με το ιστορικό αυτής της δομικής ενότητας, είτε για να επεκτείνει την κοιλότητα δέσμευσης, ή/και για να εισάγει ευελιξία. Αναλύθηκαν τα δύο παράλογα γονίδια *ba0330* και *ba0331* του *B. anthracis*. Το γονίδιο *ba0331* εμφανίστηκε ως προϊόν ενός γεγονότος διπλοποίησης του *ba0330*.

Οι τρισδιάστατες δομές απακετυλασών της οικογένειας των γονιδίων απακετυλάσης πολυσακχαρίτη χρησιμοποιήθηκαν για την εξελικτική μελέτη της θέσης δέσμευσης μετά από γεγονότα διπλοποίησης. Η δομική συγκριτική ανάλυση συνολικά επτά (τεσσάρων υπαρχουσών στη βιβλιογραφία και τριών από την παρούσα εργασία) προσδιορισμένων PDAs των *B. anthracis* και *B. cereus* υποστηρίζει τα συμπεράσματα της ανάλυσης αλληλουχίας και καταδεικνύει την ύπαρξη δομικών τροποποιήσεων στον σχηματισμό της θέσης δέσμευσης, παρά την υψηλή συντήρηση της πρωτεϊνικής αλληλουχίας της και τη διατήρηση της κοινής τοπολογίας αναδίπλωσης των γενικών δομικών χαρακτηριστικών της. Αυτό μπορεί να υποδεικνύει ότι τα τμήματα της αλληλουχίας που σχηματίζουν τη θέση δέσμευσης βρίσκονται υπό εξελικτική πίεση για να φιλοξενήσουν και να αναγνωρίσουν διακριτά κατάλοιπα πολυσακχαριτών ανάλογα με την κυτταρική θέση, τη χρήση ή το περιβάλλον. Αν και η πρωτεϊνική δομή είναι πιο συντηρημένη από την αμινοξική ακολουθία, αυτή η μελέτη έδειξε ότι μικρές δομικές τροποποιήσεις στα πολλαπλά μοτίβα που συνθέτουν τη θέση δέσμευσης μπορεί να προκαλέσουν λειτουργική διαφοροποίηση.

Η εξελικτική λειτουργική διαφοροποίηση, και η ανάπτυξη της εξειδίκευσης της καταλυτικής δομικής ενότητας NodB με την έννοια των μικρών προσαρμογών της θέσης πρόσδεσης σε διαφορετικά αλλά ουσιαστικά παρόμοια υποστρώματα προφανώς είναι το αποτέλεσμα γεγονότων γονιδιακής διπλοποίησης.

Εν κατακλείδι, η σύγκριση των κοιλοτήτων δέσμευσης του υποστρώματος μεταξύ των διαφόρων φερόμενων απακετυλασών πολυσακχαρίτη του *B. anthracis* και του *B. cereus*

αποκάλυψε αξιοσημείωτη ποικιλότητα ως προς το μέγεθος και το σχήμα της κοιλότητας, καθώς επίσης και ως προς τον τύπο των αμινοξέων που τις συνθέτουν, εντοπίζοντας στοιχεία απόκλισης στα λειτουργικά χαρακτηριστικά με ταυτόχρονη διατήρηση των διακριτών τοπικών δομικών χαρακτηριστικών στην οικογένεια των PDAs. Στοιχεία της λειτουργικής απόκλισης των απακετυλασών έχουν επισημανθεί σε αυτήν τη διατριβή.

Τέλος από τις προσδιορισθείσες κρυσταλλικές δομές αποδεικνύεται η δυνατότητα σχηματισμού συμπλόκου. Εξάγονται ακριβείς πληροφορίες για τη στοιχειομετρία, τη διάταξη των συμπλόκων, τις ενδομοριακές αλλά και διαμοριακές αλληλεπιδράσεις, τη διαμόρφωση της πρωτεΐνης-προσδέτη στα σύμπλοκα. Οι δομικές πληροφορίες για τη θέση δέσμευσης υποστρωμάτων που εξάγονται από τη μοριακή δομή των απακετυλασών καθώς και συμπλόκων τους με προσδέτες είναι πρωτίστης σημασίας, προκειμένου οι απακετυλάσες να χρησιμοποιηθούν ως μόρια στόχοι φαρμάκων και χρησιμοποιούνται στην πρόταση φαρμακοφόρου.

Βιβλιογραφία

Στην ελληνική γλώσσα

Λουκάς Μ. Γ. (2010). Εισαγωγή στη γενετική. Εκδόσεις Σταμούλη Α.Ε., ISBN 978-960-351-814-3.

Ροδάκης Γ. Κ. (2001). Εισαγωγή στην Εξελικτική Βιολογία. Ιατρικές Εκδόσεις Λίτσας.

Futuyama D. J. (1991). Εξελικτική Βιολογία. Πανεπιστημιακές εκδόσεις Κρήτης.

Στην αγγλική γλώσσα

Adams P. D., Afonine P. V., Bunkóczi G., Chen V. B., Davis I. W., Echols N., Headd J. J., Hung L.-W., Kapral G. J., Grosse-Kunstleve R. W., McCoy A. J., Moriarty N. W., Oeffner R., Read R. J., Richardson D. C., Richardson J. S., Terwilliger T. C., and Zwart P. H. (2010). *it* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D* **66**, 213–221. doi:10.1107/S0907444909052925

Afonine P. V., Grosse-Kunstleve R. W., Echols N., Headd J. J., Moriarty N. W., Mustyakimov M., Terwilliger T. C., Urzhumtsev A., Zwart P. H., and Adams P. D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367. doi:10.1107/S0907444912001308

Alcorlo M., Martínez-Caballero S., Molina R., and Hermoso J. A. (2017). Carbohydrate recognition and lysis by bacterial peptidoglycan hydrolases. *Carbohydr. Feast Struct. Glycobiol. • Seq. Topol. Comput. Stud. Protein-Protein Interact.* **44**, 87–100. doi:10.1016/j.sbi.2017.01.001

Altschul S. F., Gish W., Miller W., Myers E. W., and Lipman D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. doi:10.1016/S0022-2836(05)80360-2

Anderson R. P., and Roth J. R. (1977). Tandem genetic duplications in phage and bacteria. *Annu. Rev. Microbiol.* **31**, 473–505. doi:10.1146/annurev.mi.31.100177.002353

Anderson R. P., and Roth J. R. (1979). Gene duplication in bacteria: alteration of gene dosage by sister-chromosome exchanges. *Cold Spring Harb. Symp. Quant. Biol.* **43 Pt 2**, 1083–1087.

- Andreeva A., and Murzin A. G. (2006). Evolution of protein fold in the presence of functional constraints. *Curr. Opin. Struct. Biol.* **16**, 399–408. doi:10.1016/j.sbi.2006.04.003
- Andreou A., Giastas P., Arnauteli S., Tzanodaskalaki M., Tzartos S. J., Bethanis K., Bouriotis V., and Eliopoulos E. E. (2019). The putative polysaccharide deacetylase Ba0331: cloning, expression, crystallization and structure determination. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **75**, 312–320. doi:10.1107/S2053230X19001766
- Andreou A., Giastas P., Christoforides E., and Eliopoulos E. Elias (2018). Structural and Evolutionary Insights within the Polysaccharide Deacetylase Gene Family of *Bacillus anthracis* and *Bacillus cereus*. *Genes* **9**, doi:10.3390/genes9080386
- Andrés E., Albesa-Jové D., Biarnés X., Moerschbacher B. M., Guerin M. E., and Planas A. (2014). Structural Basis of Chitin Oligosaccharide Deacetylation. *Angew. Chem. Int. Ed.* **53**, 6882–6887. doi:10.1002/anie.201400220
- Aragunde H., Biarnés X., and Planas A. Substrate Recognition and Specificity of Chitin Deacetylases and Related Family 4 Carbohydrate Esterases. *Int. J. Mol. Sci.* **2018**, *19*, 412.
- Araki Y., Fukuoka S., Oba S., and Ito E. (1971). Enzymatic deacetylation of N-acetylglucosamine residues in peptidoglycan from *Bacillus cereus* cell walls. *Biochem Biophys Res Commun.* **45**, 751-758.
- Arnaouteli S., Giastas P., Andreou A., Tzanodaskalaki M., Aldridge C., Tzartos S. J., Vollmer W., Eliopoulos E., and Bouriotis V. (2015). Two Putative Polysaccharide Deacetylases Are Required for Osmotic Stability and Cell Shape Maintenance in *Bacillus anthracis*. *J. Biol. Chem.* **290**, 13465–13478. doi:10.1074/jbc.M115.640029
- Arnold F. H. (2018). Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem. Int. Ed Engl.* **57**, 4143–4148. doi:10.1002/anie.201708408
- Avery L., and Wasserman S. (1992). Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends Genet. TIG* **8**, 312–316. doi:10.1016/0168-9525(92)90263-4
- Babu M. M., Priya M. L., Selvan A. T., Madera M., Gough J., Aravind L., and Sankaran K. (2006). A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J. Bacteriol.* **188**, 2761–2773. doi:10.1128/JB.188.8.2761-2773.2006

- Baker N. A., Sept D., Joseph S., Holst M. J., and McCammon J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10037–10041. doi:10.1073/pnas.181342398
- Balomenou S., Arnaouteli S., Koutsoulis D., Fadouloglou V., and Bouriotis V. (2015). Polysaccharide Deacetylases: New Antibacterial Drug Targets. ‘Frontiers in Anti-Infective Drug Discovery’. 63. pp. 68–130. (Bentham Science)
- Balomenou S., Fouet A., Tzanodaskalaki M., Couture-Tosi E., Bouriotis V., and Boneca I. G. (2013). Distinct functions of polysaccharide deacetylases in cell shape, neutral polysaccharide synthesis and virulence of *Bacillus anthracis*. *Mol. Microbiol.* **87**, 867–883. doi:10.1111/mmi.12137
- Baron M., L. Main A., Driscoll P., J. Mardon H., Boyd J., and Campbell I. (1992). ‘Proton NMR assignment and secondary structure of the cell adhesion type III module of fibronectin.’ doi:10.1021/bi00122a025
- Barth H. G., Boyes B. E., and Jackson C. (1996). Size Exclusion Chromatography. *Anal. Chem.* **68**, 445–466. doi:10.1021/a19600193
- Bateman A., Jouet M., MacFarlane J., Du J. S., Kenwick S., and Chothia C. (1996). Outline structure of the human L1 cell adhesion molecule and the sites where mutations cause neurological disorders. *EMBO J.* **15**, 6050–6059.
- Bateson W., and Mendel G. (1909). ‘Mendel’s principles of heredity.’ (Cambridge :University Press,1909.) <https://doi.org/10.5962/bhl.title.44575>
- Bazan, J.F. (1990). Structural design and molecular evolution of a cytokine receptor superfamily. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 6934–6938.
- Bergfors, T. M., Protein Crystallization 2nd edition, Iul Biotechnology Series 2009 ISBN 978-0972077446, 97-108, 157-170.
- Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., and Bourne P. E. (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242.
- Bershtein S., Segal M., Bekerman R., Tokuriki N., and Tawfik D. S. (2006). Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature* **444**, 929–932. doi:10.1038/nature05385
- Beveridge A. J. (1996). A theoretical study of the active sites of papain and S195C rat trypsin: implications for the low reactivity of mutant serine proteinases. *Protein Sci. Publ. Protein Soc.* **5**, 1355–1365. doi:10.1002/pro.5560050714
- Birnboim H. C., and Doly J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513–1523.

- Blair D. E., and van Aalten D. M.F (2004). Structures of Bacillus subtilis PdaA, a family 4 carbohydrate esterase, and a complex with N-acetyl-glucosamine. *FEBS Lett.* **570**, 13–19. doi:10.1016/j.febslet.2004.06.013
- Blair D.E., Hekmat O., Schuttelkopf A.W., Shrestha B., Tokuyasu K., Withers S.G., Van Aalten D.M.F. (2006). Structure and Mechanism of Chitin Deacetylase from the Fungal Pathogen Colletotrichum Lindemuthianum. *Biochemistry* **45**, 9416
- Blair D. E., Schuttelkopf A. W., MacRae J. I., and van Aalten D. M. F. (2005). Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15429–15434. doi:10.1073/pnas.0504339102
- Böde C., Kovács I. A., Szalay M. S., Palotai R., Korcsmáros T., and Csermely P. (2007). Network analysis of protein dynamics. *FEBS Lett.* **581**, 2776–2782. doi:10.1016/j.febslet.2007.05.021
- Boneca I. G. (2005). The role of peptidoglycan in pathogenesis. *Curr. Opin. Microbiol.* **8**, 46–53. doi:10.1016/j.mib.2004.12.008
- Boneca I. G., Dussurget O., Cabanes D., Nahori M.-A., Sousa S., Lecuit M., Psylinakis E., Bouriotis V., Hugot J.-P., Giovannini M., Coyle A., Bertin J., Namane A., Rouselle J.-C., Cayet N., Prevost M.-C., Balloy V., Chignard M., Philpott D. J., Cossart P., and Girardin S. E. (2007). A critical role for peptidoglycan N-deacetylation in Listeria evasion from the host innate immune system. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 997–1002. doi:10.1073/pnas.0609672104
- Bork P., and Doolittle R. F. (1992). Proposed acquisition of an animal protein domain by bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8990–8994.
- Bourgogne A., Drysdale M., Hilsenbeck S. G., Peterson S. N., and Koehler T. M. (2003). Global effects of virulence gene regulators in a Bacillus anthracis strain with both virulence plasmids. *Infect. Immun.* **71**, 2736–2743.
- Bradford M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254. doi:10.1016/0003-2697(76)90527-3
- Braun V., and Wu. H. C. (1993). Lipoproteins, structure, function, biosynthesis and model for protein export, p. 319 - 342. In J-M Ghuyssen and R. Hakenback (ed), Comprehensive biochemistry, vol 27. Bacterial cell wall. Elsevier Science Publishers, B.V. Amsterdam

- Braun V., and Rehn K. (1969). Chemical characterization, spatial distribution and function of a lipoprotein (murein-lipoprotein) of the *E. coli* cell wall. The specific effect of trypsin on the membrane structure. *Eur. J. Biochem.* **10**, 426–438.
- Breen M. S., Kemena C., Vlasov P. K., Notredame C., and Kondrashov F. A. (2012). Epistasis as the primary factor in molecular evolution. *Nature* **490**, 535–538. doi:10.1038/nature11510
- Bridges C. B. (1936). The Bar ‘gene’ a duplication. *Science* **83**, 210–211. doi:10.1126/science.83.2148.210
- Bruenger A.T. (1977). Free R value: cross-validation in crystallography. *Methods Enzymol.* **277**, 366–396.
- Brunelle J.L., and Green R. (2014). Chapter Twelve - One-dimensional SDS-Polyacrylamide Gel Electrophoresis (1D SDS-PAGE). *Methods in Enzymology.* **541**, 151–159.
- Bui N. K., Turk S., Buckenmaier S., Stevenson-Jones F., Zeuch B., Gobec S., and Vollmer W. (2011). Development of screening assays and discovery of initial inhibitors of pneumococcal peptidoglycan deacetylase PgdA. *Biochem. Pharmacol.* **82**, 43–52. doi:10.1016/j.bcp.2011.03.028
- Burnette W. N. (1981). ‘Western blotting’: electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* **112**, 195–203.
- Campbell I. D., and Spitzfaden C. (1994). Building proteins with fibronectin type III modules. *Struct. Lond. Engl.* **1993** **2**, 333–337.
- Candela T., Balomenou S., Aucher W., Bouriotis V., Simore J.-P., Fouet A., and Boneca I. G. (2014). N-acetylglucosamine deacetylases modulate the anchoring of the gamma-glutamyl capsule to the cell wall of *Bacillus anthracis*. *Microb. Drug Resist. Larchmt.* **N 20**, 222–230. doi:10.1089/mdr.2014.0063
- Candela T., and Fouet A. (2006). Poly-gamma-glutamate in bacteria. *Mol. Microbiol.* **60**, 1091–1098. doi:10.1111/j.1365-2958.2006.05179.x
- Casari G., Sander C., and Valencia A. (1995). A method to predict functional residues in proteins. *Nat. Struct. Biol.* **2**, 171–178.
- Caufrier F., Martinou A., Dupont C., and Bouriotis V. (2003). Carbohydrate esterase family 4 enzymes: substrate specificity. *Carbohydr. Res.* **338**, 687–692.
- Ceccarelli M., Santantonio E., Marmottini F., Amzallag G. N., and Cionini P. G. (2006). Chromosome endoreduplication as a factor of salt adaptation in *Sorghum bicolor*. *Protoplasma* **227**, 113–118. doi:10.1007/s00709-005-0144-0

- Chandonia J.-M., and Brenner S. E. (2006). The impact of structural genomics: expectations and outcomes. *Science* **311**, 347–351. doi:10.1126/science.1121018
- Chen V. B., Arendall W. B. 3rd, Headd J. J., Keedy D. A., Immormino R. M., Kapral G. J., Murray L. W., Richardson J. S., and Richardson D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21. doi:10.1107/S0907444909042073
- Chen Y., Zhang W., Shi Q., Heseck D., Lee M., Mobashery S., and Shoichet B. K. (2009). Crystal structures of penicillin-binding protein 6 from *Escherichia coli*. *J. Am. Chem. Soc.* **131**, 14345–14354. doi:10.1021/ja903773f
- Cherif A., Borin S., Rizzi A., Ouzari H., Boudabous A. and Daffonchio D. (2003). *Bacillus anthracis* diverges from related clades of the *Bacillus cereus* group in 16S-23S ribosomal DNA intergenic transcribed spacers containing tRNA genes. *Appl Environ Microbiol* **69**, 33–40.
- Choudhury B., Leoff C., Saile E., Wilkins P., Quinn C. P., Kannenberg E. L., and Carlson R. W. (2006). The structure of the major cell wall polysaccharide of *Bacillus anthracis* is species-specific. *J. Biol. Chem.* **281**, 27932–27941. doi:10.1074/jbc.M605768200
- Clarke A.J. (2014). Chemical biology of peptidoglycan acetylation and deacetylation. *Bioorg. Chem.* **54**, 44–50.
- Cobb R. E., Chao R., and Zhao H. (2013). Directed Evolution: Past, Present and Future. *AIChE J. Am. Inst. Chem. Eng.* **59**, 1432–1440. doi:10.1002/aic.13995
- Conant G. C., and Wolfe K. H. (2008). Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet.* **9**, 938–950. doi:10.1038/nrg2482
- Cornman R. S., and Willis J. H. (2008). Extensive gene amplification and concerted evolution within the CPR family of cuticular proteins in mosquitoes. *Insect Biochem. Mol. Biol.* **38**, 661–676. doi:10.1016/j.ibmb.2008.04.001
- Craven S. H., and Neidle E. L. (2007). Double trouble: medical implications of genetic duplication and amplification in bacteria. *Future Microbiol.* **2**, 309–321. doi:10.2217/17460913.2.3.309
- Delcour J., Ferain T., Deghorain M., Palumbo E., and Hols P. (1999). The biosynthesis and functionality of the cell-wall of lactic acid bacteria. *Antonie Van Leeuwenhoek* **76**, 159–184.
- Deng D. M., Urch J. E., ten Cate J. M., Rao V. A., van Aalten D. M. F., and Crielaard W. (2009). *Streptococcus mutans* SMU.623c codes for a functional, metal-dependent

- polysaccharide deacetylase that modulates interactions with salivary agglutinin. *J. Bacteriol.* **191**, 394–402. doi:10.1128/JB.00838-08
- DHAR R., SÄGESSER R., WEIKERT C., YUAN J., and WAGNER A. (2011). Adaptation of *Saccharomyces cerevisiae* to saline stress through laboratory evolution. *J. Evol. Biol.* **24**, 1135–1153. doi:10.1111/j.1420-9101.2011.02249.x
- Doolittle R.F. (1985). The genealogy of some recently evolved vertebrate proteins. *Trends Biochem. Sciences* **10**, 233-237.
- Drobniewski F. A. (1993). *Bacillus cereus* and related species. *Clin. Microbiol. Rev.* **6**, 324–338.
- Dummler A., Lawrence A.-M., and de Marco A. (2005). Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb. Cell Factories* **4**, 34. doi:10.1186/1475-2859-4-34
- Emsley P., and Cowtan K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132. doi:10.1107/S0907444904019158
- Emsley P., Lohkamp B., Scott W. G., and Cowtan K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501. doi:10.1107/S0907444910007493
- Engh R. A., and Huber R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. Sect. A* **47**, 392–400. doi:10.1107/S0108767391001071
- Evans P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82. doi:10.1107/S0907444905036693
- Evans P. R., and Murshudov G. N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214. doi:10.1107/S0907444913000061
- Fadoulglou V. E., Balomenou S., Aivaliotis M., Kotsifaki D., Arnaouteli S., Tomatsidou A., Efstathiou G., Kountourakis N., Miliara S., Griniezaki M., Tsalafouta A., Pergantis S. A., Boneca I. G., Glykos N. M., Bouriotis V., and Kokkinidis M. (2017). Unusual alpha-Carbon Hydroxylation of Proline Promotes Active-Site Maturation. *J. Am. Chem. Soc.* **139**, 5330–5337. doi:10.1021/jacs.6b12209
- Fadoulglou V. E., Kapanidou M., Agiomirgianaki A., Arnaouteli S., Bouriotis V., Glykos N. M., and Kokkinidis M. (2013). Structure determination through homology modelling and torsion-angle simulated annealing: application to a polysaccharide deacetylase

- from *Bacillus cereus*. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 276–283. doi:10.1107/S0907444912045829
- Feld G. K., Thoren K. L., Kintzer A. F., Sterling H. J., Tang I. I., Greenberg S. G., Williams E. R., and Krantz B. A. (2010). Structural basis for the unfolding of anthrax lethal factor by protective antigen oligomers. *Nat. Struct. Mol. Biol.* **17**, 1383–1390. doi:10.1038/nsmb.1923
- Felsenstein Joseph (1985). CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution* **39**, 783–791. doi:10.1111/j.1558-5646.1985.tb00420.x
- Fisher R. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans R. Soc Edinb.* 399–433.
- Fouet A. (2009). The surface of *Bacillus anthracis*. *Mol. Aspects Med.* **30**, 374–385. doi:10.1016/j.mam.2009.07.001
- Fouet A., and Mesnage S. (2002). *Bacillus anthracis* cell envelope components. *Curr. Top. Microbiol. Immunol.* **271**, 87–113.
- Fox A., Stewart G. C., Waller L. N., Fox K. F., Harley W. M., and Price R. L. (2003). Carbohydrates and glycoproteins of *Bacillus anthracis* and related bacilli: targets for biodetection. *J. Microbiol. Methods* **54**, 143–152. doi:10.1016/S0167-7012(03)00095-2
- Francino M. P. (2005). An adaptive radiation model for the origin of new gene functions. *Nat. Genet.* **37**, 573–577. doi:10.1038/ng1579
- Fukushima T., Kitajima T., and Sekiguchi J. (2005). A Polysaccharide Deacetylase Homologue, PdaA, in *Bacillus subtilis* Acts as an N-Acetylmuramic Acid Deacetylase In Vitro. *J. Bacteriol.* **187**, 1287. doi:10.1128/JB.187.4.1287-1292.2005
- Fukushima T., Yamamoto H., Atrih A., Foster, S.J., Sekiguchi J, A. (2002). polysaccharide deacetylase gene (pdaa) is required for germination and for production of muramic delta-lactam residues in the spore cortex of *Bacillus subtilis* . *J. Bacteriol.* **184**, 6007–6015. doi: 10.1128/JB.184.21.6007-6015.2002
- Gerstein A. C., Chun H.-J. E., Grant A., and Otto S. P. (2006). Genomic Convergence toward Diploidy in *Saccharomyces cerevisiae*. *PLOS Genet.* **2**, e145. doi:10.1371/journal.pgen.0020145
- Giastas P., Andreou A., Papakyriakou A., Koutsioulis D., Balomenou S., Tzartos S. J., Bouriotis V., and Eliopoulos E. E. (2018). Structures of the Peptidoglycan N-

- Acetylglucosamine Deacetylase Bc1974 and Its Complexes with Zinc Metalloenzyme Inhibitors. *Biochemistry* **57**, 753–763. doi:10.1021/acs.biochem.7b00919
- Gilkes N. R., Henrissat B., Kilburn D. G., Miller R. C. Jr, and Warren R. A. (1991). Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. *Microbiol. Rev.* **55**, 303–315.
- Goldman N., and Yang Z. (1994). A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol Biol Evol.* **11**, 725-736.
- Gong D., Wilson P. W., Bain M. M., McDade K., Kalina J., Hervé-Grépinet V., Nys Y., and Dunn I. C. (2010). Gallin; an antimicrobial peptide member of a new avian defensin family, the ovodefensins, has been subject to recent gene duplication. *BMC Immunol.* **11**, 12. doi:10.1186/1471-2172-11-12
- Gouet P., Robert X., and Courcelle E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3323.
- Griffiths A.J.F., S. R. Wessler , S. B. Carroll, Doebley J. (2015). An Introduction to Genetic Analysis 11th Edition, Freeman W.H. Edition.
- Gu X., and Vander Velden K. (2002). DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. *Bioinforma. Oxf. Engl.* **18**, 500–501.
- Gu X., Zou Y., Su Z., Huang W., Zhou Z., Arendsee Z., and Zeng Y. (2013). An update of DIVERGE software for functional divergence analysis of protein family. *Mol. Biol. Evol.* **30**, 1713–1719. doi:10.1093/molbev/mst069
- Halabi N., Rivoire O., Leibler S., and Ranganathan R. (2009). Protein sectors: evolutionary units of three-dimensional structure. *Cell* **138**, 774–786. doi:10.1016/j.cell.2009.07.038
- Hanada K., Zou C., Lehti-Shiu M. D., Shinozaki K., and Shiu S.-H. (2008). Importance of Lineage-Specific Expansion of Plant Tandem Duplicates in the Adaptive Response to Environmental Stimuli. *Plant Physiol.* **148**, 993. doi:10.1104/pp.108.122457
- Hansen C. K. (1992). Fibronectin type III-like sequences and a new domain type in prokaryotic depolymerases with insoluble substrates. *FEBS Lett.* **305**, 91–96.
- Hantke K., and Braun V. (1973). Covalent binding of lipid to protein. Diglyceride and amide-linked fatty acid at the N-terminal end of the murein-lipoprotein of the Escherichia coli outer membrane. *Eur. J. Biochem.* **34**, 284–296.

- Harms M. J., and Thornton J. W. (2013). Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat. Rev. Genet.* **14**, 559–571. doi:10.1038/nrg3540
- Hastings P. J. (2007). Adaptive Amplification. *Crit. Rev. Biochem. Mol. Biol.* **42**, 271–283. doi:10.1080/10409230701507757
- Hawkins P. C. D., Skillman A. G., Warren G. L., Ellingson B. A., and Stahl M. T. (2010). Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **50**, 572–584. doi:10.1021/ci100031x
- Hawkins P.C.D., and Nicholls A. (2012). Conformer generation with OMEGA: learning from the data set and the analysis of failures. *J Chem Inf Model.* **52**, 2919-2936. doi: 10.1021/ci300314k.
- Hayward S., and Berendsen H. J. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30**, 144–154.
- Hebert L., Courtin P., Torelli R., Sanguinetti M., Chapot-Chartier M.-P., Auffray Y., and Benachour A. (2007). *Enterococcus faecalis* constitutes an unusual bacterial model in lysozyme resistance. *Infect. Immun.* **75**, 5390–5398. doi:10.1128/IAI.00571-07
- van Heijenoort J. (1998). Assembly of the monomer unit of bacterial peptidoglycan. *Cell. Mol. Life Sci. CMLS* **54**, 300–304. doi:10.1007/s000180050155
- van Heijenoort J. (2001). Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology* **11**, 25R-36R.
- Helgason E., Okstad O. A., Caugant D. A., Johansen H. A., Fouet A., Mock M., Hegna I., and Kolsto A. B. (2000). *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*--one species on the basis of genetic evidence. *Appl. Environ. Microbiol.* **66**, 2627–2630.
- Helgason E., Tourasse N.J., Meisal R., Caugant D.A. and Kolsto A.B. (2004). Multilocus sequence typing scheme for bacteria of the *Bacillus cereus* group. *Appl Environ Microbiol* **70**, 191–201.
- Hernick M., and A Fierke C. (2005). ‘Zinc hydrolases: The mechanisms of zinc-dependent deacetylases.’ doi:10.1016/j.abb.2004.08.006
- Hochuli E., Döbeli H., and Schacher A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr. A* **411**, 177–184. doi:10.1016/S0021-9673(00)93969-4

- Holm L., Kaariainen S., Rosenstrom P., and Schenkel A. (2008). Searching protein structure databases with DaliLite v.3. *Bioinforma. Oxf. Engl.* **24**, 2780–2781. doi:10.1093/bioinformatics/btn507
- Holm L., and Laakso L. M. (2016). Dali server update. *Nucleic Acids Res.* **44**, W351–W355. doi:10.1093/nar/gkw357
- Horsch M., Hoesch L., Vasella A., and Rast D. M. (1991). N-acetylglucosaminono-1,5-lactone oxime and the corresponding (phenylcarbamoyl)oxime. Novel and potent inhibitors of beta-N-acetylglucosaminidase. *Eur. J. Biochem.* **197**, 815–818.
- Huang X., and Miller W. (1991). A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math.* **12**, 337–357. doi:10.1016/0196-8858(91)90017-D
- Huey R., Morris G. M., Olson A. J., and Goodsell D. S. (2007). A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.* **28**, 1145–1152. doi:10.1002/jcc.20634
- Hughes A. L. (1994). The evolution of functionally novel proteins after gene duplication. *Proc. Biol. Sci.* **256**, 119–124. doi:10.1098/rspb.1994.0058
- Hughes A. L. (1999). ‘Adaptive Evolution of Genes and Genomes.’ (Oxford University Press)
- Hughes A. L. (2005). Gene duplication and the origin of novel proteins. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 8791. doi:10.1073/pnas.0503922102
- Hughes J. M., and Gerberding J. L. (2002). Anthrax bioterrorism: lessons learned and future directions. *Emerg. Infect. Dis.* **8**, 1013–1014. doi:10.3201/eid0810.020466
- Hultgren S.J., Abraham S., Caparon M., Falk P., St. Geme J.W., and Normark S. (1993). Pilus and nonpilus bacterial adhesins: assembly and function in cell recognition. *Cell* **73**, 887–901.
- Humphrey W., Dalke A., and Schulten K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–8, 27–28.
- Hynes, R.O. (1990). Structure of fibronectins. In *Fibronectins*; Rich, A., Ed.; Springer: New York, NY, USA, pp. 113–175. ISBN 978-1-4612-3264-3.
- Inglesby T. V., Henderson D. A., Bartlett J. G., Ascher M. S., Eitzen E., Friedlander A. M., Hauer J., McDade J., Osterholm M. T., O’Toole T., Parker G., Perl T. M., Russell P. K., and Tonat K. (1999). Anthrax as a biological weapon: medical and public health management. Working Group on Civilian Biodefense. *JAMA* **281**, 1735–1745.
- Innan H., and Kondrashov F. (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* **11**, 97.

- Inouye S., Wang S., Sekizawa J., Halegoua S., and Inouye M. (1977). Amino acid sequence for the peptide extension on the prolipoprotein of the Escherichia coli outer membrane. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 1004–1008.
- Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B., Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A., Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T., Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S. D., Overbeek R., and Kyrpides N. (2003). Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis. *Nature* **423**, 87–91. doi:10.1038/nature01582
- Jee J.-G., Ikegami T., Hashimoto M., Kawabata T., Ikeguchi M., Watanabe T., and Shirakawa M. (2002). Solution structure of the fibronectin type III domain from Bacillus circulans. *J. Biol. Chem.* **277**, 1388–1397. doi:10.1074/jbc.M109726200
- Jensen R. A. (1976). Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**, 409–425. doi:10.1146/annurev.mi.30.100176.002205
- Jernigan J. A., Stephens D. S., Ashford D. A., Omenaca C., Topiel M. S., Galbraith M., Tapper M., Fisk T. L., Zaki S., Popovic T., Meyer R. F., Quinn C. P., Harper S. A., Fridkin S. K., Sejvar J. J., Shepard C. W., McConnell M., Guarner J., Shieh W. J., Malecki J. M., Gerberding J. L., Hughes J. M., and Perkins B. A. (2001). Bioterrorism-related inhalational anthrax: the first 10 cases reported in the United States. *Emerg. Infect. Dis.* **7**, 933–944. doi:10.3201/eid0706.010604
- Jiang S.-Y., Ma Z., and Ramachandran S. (2010). Evolutionary history and stress regulation of the lectin superfamily in higher plants. *BMC Evol. Biol.* **10**, 79. doi:10.1186/1471-2148-10-79
- JOHANNSEN L. (1993). Biological properties of bacterial peptidoglycan. *APMIS* **101**, 337–344. doi:10.1111/j.1699-0463.1993.tb00119.x
- John M., Röhrig H., Schmidt J., Wieneke U., and Schell J. (1993). Rhizobium NodB protein involved in nodulation signal synthesis is a chitoooligosaccharide deacetylase. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 625–629.
- Jones D. T., Taylor W. R., and Thornton J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci. CABIOS* **8**, 275–282.
- Kabsch W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132. doi:10.1107/S0907444909047337

- Kafetzopoulos D., Martinou A., and Bouriotis V. (1993). Bioconversion of chitin to chitosan: purification and characterization of chitin deacetylase from *Mucor rouxii*. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 2564–2568.
- Kaoukab-Raji A., Biskri L., Bernardini M.-L., and Allaoui A. (2012). Characterization of SfPgdA, a *Shigella flexneri* peptidoglycan deacetylase required for bacterial persistence within polymorphonuclear neutrophils. *Microbes Infect.* **14**, 619–627. doi:10.1016/j.micinf.2012.01.009
- Karplus P. A., and Diederichs K. (2012). Linking crystallographic model and data quality. *Science* **336**, 1030–1033. doi:10.1126/science.1218231
- Kataeva I. A., Seidel R. D. 3rd, Shah A., West L. T., Li X.-L., and Ljungdahl L. G. (2002). The fibronectin type 3-like repeat from the *Clostridium thermocellum* cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying its surface. *Appl. Environ. Microbiol.* **68**, 4292–4300.
- Kim S., Thiessen P. A., Bolton E. E., Chen J., Fu G., Gindulyte A., Han L., He J., He S., Shoemaker B. A., Wang J., Yu B., Zhang J., and Bryant S. H. (2016). PubChem Substance and Compound databases. *Nucleic Acids Res.* **44**, D1202–1213. doi:10.1093/nar/gkv951
- Kintzer A. F., Thoren K. L., Sterling H. J., Dong K. C., Feld G. K., Tang I. I., Zhang T. T., Williams E. R., Berger J. M., and Krantz B. A. (2009). The Protective Antigen Component of Anthrax Toxin Forms Functional Octameric Complexes. *J. Mol. Biol.* **392**, 614–629. doi:10.1016/j.jmb.2009.07.037
- Ko M., Choi H. and Park C. (2002). Group I self-splicing intron in the recA gene of *Bacillus anthracis*. *J Bacteriol* **184**, 3917–3922.
- Koch A. L. (1981). Evolution of antibiotic resistance gene function. *Microbiol. Rev.* **45**, 355–378.
- Kolsto A.-B., Tourasse N. J., and Okstad O. A. (2009). What sets *Bacillus anthracis* apart from other *Bacillus* species? *Annu. Rev. Microbiol.* **63**, 451–476. doi:10.1146/annurev.micro.091208.073255
- Kondrashov F. A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc. Biol. Sci.* **279**, 5048–5057. doi:10.1098/rspb.2012.1108
- Kondrashov F. A., and Kondrashov A. S. (2006). Role of selection in fixation of gene duplications. *J. Theor. Biol.* **239**, 141–151. doi:10.1016/j.jtbi.2005.08.033

- Kondrashov F. A., and Koonin E. V. (2004). A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications. *Trends Genet. TIG* **20**, 287–290. doi:10.1016/j.tig.2004.05.001
- Kondrashov F. A., Rogozin I. B., Wolf Y. I., and Koonin E. V. (2002). Selection in the evolution of gene duplications. *Genome Biol.* **3**, research0008.1. doi:10.1186/gb-2002-3-2-research0008
- Konrad A., Teufel A. I., Grahnen J. A., and Liberles D. A. (2011). Toward a General Model for the Evolutionary Dynamics of Gene Duplicates. *Genome Biol. Evol.* **3**, 1197–1209. doi:10.1093/gbe/evr093
- Kornbliht A.R., Umezawa K., Vibe-Pedersen K., Baralle F.E. (1985). Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene *EMBO J.* **4**, 1755-1759.
- Kozutsumi Y., Normington K., Press E., Slaughter C., Sambrook J., and Gething M. J. (1989). Identification of immunoglobulin heavy chain binding protein as glucose-regulated protein 78 on the basis of amino acid sequence, immunological cross-reactivity, and functional activity. *J. Cell Sci. Suppl.* **11**, 115–137.
- Krennhrubec K., Marshall B. L., Hedglin M., Verdin E., and Ulrich S. M. (2007). Design and evaluation of ‘Linkerless’ hydroxamic acids as selective HDAC8 inhibitors. *Bioorg. Med. Chem. Lett.* **17**, 2874–2878. doi:10.1016/j.bmcl.2007.02.064
- Krissinel E., and Henrick K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797. doi:10.1016/j.jmb.2007.05.022
- Kumar S., Stecher G., and Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874. doi:10.1093/molbev/msw054
- Laemmli U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**, 680–685. doi:10.1038/227680a0
- Lambert C., Lerner T. R., Bui N. K., Somers H., Aizawa S.-I., Liddell S., Clark A., Vollmer W., Lovering A. L., and Sockett R. E. (2016). Interrupting peptidoglycan deacetylation during *Bdellovibrio* predator-prey interaction prevents ultimate destruction of prey wall, liberating bacterial-ghosts. *Sci. Rep.* **6**, 26010. doi:10.1038/srep26010
- Lanave C., Preparata G, Saccone C., Serio G. (1984). A new method for calculating evolutionary substitution rates *J. Mol. Evol.* **20**, 86–93.

- Lapidus A., Goltsman E., Auger S., Galleron N., Segurens B., Dossat C., Land M. L., Broussolle V., Brillard J., Guinebretiere M.-H., Sanchis V., Nguen-The C., Lereclus D., Richardson P., Wincker P., Weissenbach J., Ehrlich S. D., and Sorokin A. (2008). Extending the *Bacillus cereus* group genomics to putative food-borne pathogens of different toxicity. *Chem. Biol. Interact.* **171**, 236–249. doi:10.1016/j.cbi.2007.03.003
- Laskowski R., MacArthur M. W., Moss D. S., and Thornton J. (1993). 'PROCHECK: A program to check the stereochemical quality of protein structures.' doi:10.1107/S0021889892009944
- Leahy D. J., Aukhil I., and Erickson H. P. (1996). 2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. *Cell* **84**, 155–164.
- Leahy D. J., Hendrickson W. A., Aukhil I., and Erickson H. P. (1992). Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science* **258**, 987–991.
- Lebedev A. A., Young P., Isupov M. N., Moroz O. V., Vagin A. A., and Murshudov G. N. (2012). JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 431–440. doi:10.1107/S090744491200251X
- Leslie A. G. W., and Powell H. R. (2007). Processing diffraction data with mosflm. In Read RJ, and Sussman JL (eds) 'Evolving Methods for Macromolecular Crystallography', pp. 41–51. (Springer Netherlands)
- Leslie M. (2013). 'Dead' Enzymes Show Signs of Life. *Science* **340**, 25-27. doi: 10.1126/science.340.6128.25
- Li W. H. (1983). Evolution of duplicate genes and pseudogenes. 'Evolution of Genes and Proteins'. (Eds M Nei and M Koehn) pp. 14–37. Sinauer Associates: Sunderland, MA
- Li W. H. (1997). Molecular Evolution. Sinauer Associates Inc
- Little D. J., Bamford N. C., Pokrovskaya V., Robinson H., Nitz M., and Howell P. L. (2014). Structural basis for the De-N-acetylation of Poly-beta-1,6-N-acetyl-D-glucosamine in Gram-positive bacteria. *J. Biol. Chem.* **289**, 35907–35917. doi:10.1074/jbc.M114.611400
- Little E., Bork P., and Doolittle R. F. (1994). Tracing the spread of fibronectin type III domains in bacterial glycohydrolases. *J. Mol. Evol.* **39**, 631–643. doi:10.1007/BF00160409
- Little D. J., Milek S., Bamford N. C., Ganguly T., DiFrancesco B. R., Nitz M., Deora R., and Howell P. L. (2015). The protein BpsB is a poly-beta-1,6-N-acetyl-D-glucosamine

- deacetylase required for biofilm formation in *Bordetella bronchiseptica*. *J. Biol. Chem.* **290**, 22827–22840. doi:10.1074/jbc.M115.672469
- Little D. J., Poloczek J., Whitney J. C., Robinson H., Nitz M., and Howell P. L. (2012). The structure- and metal-dependent activity of *Escherichia coli* PgaB provides insight into the partial de-N-acetylation of poly-beta-1,6-N-acetyl-D-glucosamine. *J. Biol. Chem.* **287**, 31126–31137. doi:10.1074/jbc.M112.390005
- Liu S., Moayeri M., and Leppla S. H. (2014). Anthrax lethal and edema toxins in anthrax pathogenesis. *Trends Microbiol.* **22**, 317–325. doi:10.1016/j.tim.2014.02.012
- Lockless S. W., and Ranganathan R. (1999). Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* **286**, 295–299.
- Lombard V., Golaconda Ramulu H., Drula E., Coutinho P. M., and Henrissat B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–495. doi:10.1093/nar/gkt1178
- Lunzer M., Golding G. B., and Dean A. M. (2010). Correction: Pervasive Cryptic Epistasis in Molecular Evolution. *PLOS Genet.* **6**, 10.1371/annotation/d618ce28-5010-47df-a44e-148ecdc0fef6. doi:10.1371/annotation/d618ce28-5010-47df-a44e-148ecdc0fef6
- Lutz S. (2010). Beyond directed evolution--semi-rational protein engineering and design. *Curr. Opin. Biotechnol.* **21**, 734–743. doi:10.1016/j.copbio.2010.08.011
- Lynch M., and Conery J. S. (2000). The Evolutionary Fate and Consequences of Duplicate Genes. *Science* **290**, 1151. doi:10.1126/science.290.5494.1151
- Main A. L., Harvey T. S., Baron M., Boyd J., and Campbell I. D. (1992). The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions. *Cell* **71**, 671–678.
- Makino T., Suzuki Y., and Gojobori T. (2006). Differential evolutionary rates of duplicated genes in protein interaction network. *Gene* **385**, 57–63. doi:10.1016/j.gene.2006.06.028
- Matsumoto G. (2003). Bioterrorism. Anthrax powder: state of the art? *Science* **302**, 1492–1497. doi:10.1126/science.302.5650.1492
- Matthews B. W. (1988). Structural basis of the action of thermolysin and related zinc peptidases. *Acc. Chem. Res.* **21**, 333–340. doi:10.1021/ar00153a003
- Maynard Smith J. 1920- (1978). ‘The evolution of sex / John Maynard Smith.’ (Cambridge University Press: Cambridge [Eng.] ; New York)

- McCoy A. J., Grosse-Kunstleve R. W., Adams P. D., Winn M. D., Storoni L. C., and Read R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674. doi:10.1107/S0021889807021206
- McDonald J.H, and Kreitman M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*. **351**, 652-654.
- McDonell M. W., Simon M. N., and Studier F. W. (1977). Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. *J. Mol. Biol.* **110**, 119–146.
- McPherson, A., *Macromolecular Crystallography*, 2nd Edition, Wiley-Blackwell, 2016, ISBN 978-0470185902, 28.
- Meinke A., Gilkes N. R., Kilburn D. G., Miller R. C., and Warren R. A. (1991). Multiple domains in endoglucanase B (CenB) from *Cellulomonas fimi*: functions and relatedness to domains in other polypeptides. *J. Bacteriol.* **173**, 7126. doi:10.1128/jb.173.22.7126-7135.1991
- Mengin-Lecreulx D., and Lemaitre B. (2005). Structure and metabolism of peptidoglycan and molecular requirements allowing its detection by the *Drosophila* innate immune system. *J. Endotoxin Res.* **11**, 105–111. doi:10.1179/096805105X35233
- Meselson M., Guillemin J., Hugh-Jones M., Langmuir A., Popova I., Shelokov A., and Yampolskaya O. (1994). The Sverdlovsk anthrax outbreak of 1979. *Science* **266**, 1202–1208.
- Mesnager S., Fontaine T., Mignot T., Delepierre M., Mock M., and Fouet A. (2000). Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *EMBO J.* **19**, 4473–4484. doi:10.1093/emboj/19.17.4473
- Messner P. (1997). Bacterial glycoproteins. *Glycoconj. J.* **14**, 3–11.
- Meyrand M., Boughammoura A., Courtin P., Mezange C., Guillot A., and Chapot-Chartier M.-P. (2007). Peptidoglycan N-acetylglucosamine deacetylation decreases autolysis in *Lactococcus lactis*. *Microbiol. Read. Engl.* **153**, 3275–3285. doi:10.1099/mic.0.2007/005835-0
- Missiakas D., and Schneewind O. (2017). Assembly and Function of the *Bacillus anthracis* S-Layer. *Annu. Rev. Microbiol.* **71**, 79–98. doi:10.1146/annurev-micro-090816-093512
- Mock M., and Fouet A. (2001). Anthrax. *Annu. Rev. Microbiol.* **55**, 647–671. doi:10.1146/annurev.micro.55.1.647

- Molnar J., and Pragai B. (1971). Attempts to detect the presence of teichoic acid in *Bacillus anthracis*. *Acta Microbiol. Acad. Sci. Hung.* **18**, 105–108.
- Moore R. C., and Purugganan M. D. (2003). The early stages of duplicate gene evolution. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15682–15687. doi:10.1073/pnas.2535513100
- Moriarty N. W., Grosse-Kunstleve R. W., and Adams P. D. (2009). It electronic Ligand Builder and Optimization Workbench (it eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. Sect. D* **65**, 1074–1080. doi:10.1107/S0907444909029436
- Morris G. M., Huey R., Lindstrom W., Sanner M. F., Belew R. K., Goodsell D. S., and Olson A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791. doi:10.1002/jcc.21256
- Moynihan P. J., Sychantha D., and Clarke A. J. (2014). Chemical biology of peptidoglycan acetylation and deacetylation. *Bioorganic Chem.* **54**, 44–50. doi:10.1016/j.bioorg.2014.03.010
- Mullis K. B., and Faloona F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**, 335–350.
- Neet K. E., and Koshland D. E. Jr (1966). The conversion of serine at the active site of subtilisin to cysteine: a ‘chemical mutation’. *Proc. Natl. Acad. Sci. U. S. A.* **56**, 1606–1611. doi:10.1073/pnas.56.5.1606
- Nei M. (1969). Gene Duplication and Nucleotide Substitution in Evolution. *Nature* **221**, 40–42. doi:10.1038/221040a0
- Nei M. and Kumar S. (2000). *Molecular Evolution and Phylogenetics* Oxford University Press: Oxford, UK, pp. 17–49, ISBN 0 19 513584 9.
- Neiman M., Olson M. S., and Tiffin P. (2009). Selective histories of poplar protease inhibitors: elevated polymorphism, purifying selection, and positive selection driving divergence of recent duplicates. *New Phytol.* **183**, 740–750. doi:10.1111/j.1469-8137.2009.02936.x
- Nelson D. E., Ghosh A. S., Paulson A. L., and Young K. D. (2002). Contribution of Membrane-Binding and Enzymatic Domains of Penicillin Binding Protein 5 to Maintenance of Uniform Cellular Morphology of *Escherichia coli*. *J. Bacteriol.* **184**, 3630. doi:10.1128/JB.184.13.3630-3639.2002
- Nguyen M. T., and Gotz F. (2016). Lipoproteins of Gram-Positive Bacteria: Key Players in the Immune Response and Virulence. *Microbiol. Mol. Biol. Rev. MMBR* **80**, 891–903. doi:10.1128/MMBR.00028-16

- Nicholas K. B. (1997). GeneDoc: analysis and visualization of genetic variation. *Embnew News* **4**, 14.
- Nikolaidis I., Favini-Stabile S., and Dessen A. (2014). Resistance to antibiotics targeted to the bacterial cell wall. *Protein Sci. Publ. Protein Soc.* **23**, 243–259. doi:10.1002/pro.2414
- Nishiyama T., Noguchi H., Yoshida H., Yong Park S., and Tame J. (2013). ‘The structure of the deacetylase domain of Escherichia coli PgaB, an enzyme required for biofilm formation: A circularly permuted member of the carbohydrate esterase 4 family. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 44-51 doi:10.1107/S0907444912042059
- Notredame C., Higgins D. G., and Heringa J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217. doi:10.1006/jmbi.2000.4042
- Oberbarnscheidt L., Taylor E. J., Davies G. J., and Gloster T. M. (2007). Structure of a carbohydrate esterase from Bacillus anthracis. *Proteins* **66**, 250–252. doi:10.1002/prot.21217
- Oberhauser A. F., Marszalek P. E., Erickson H. P., and Fernandez J. M. (1998). The molecular elasticity of the extracellular matrix protein tenascin. *Nature* **393**, 181–185. doi:10.1038/30270
- Oehme I., Deubzer H. E., Wegener D., Pickert D., Linke J.-P., Hero B., Kopp-Schneider A., Westermann F., Ulrich S. M., von Deimling A., Fischer M., and Witt O. (2009). Histone deacetylase 8 in neuroblastoma tumorigenesis. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **15**, 91–99. doi:10.1158/1078-0432.CCR-08-0684
- Ohno S. (1967). ‘Sex Chromosomes and Sex-Linked Genes.’ (Springer, Berlin, Heidelberg)
- Ohno S. (1970). ‘Evolution by gene duplication.’ (London: George Allen & Unwin Ltd. Berlin, Heidelberg and New York: Springer-Verlag.)
- Okinaka R. T., Cloud K., Hampton O., Hoffmaster A. R., Hill K. K., Keim P., Koehler T. M., Lamke G., Kumano S., Mahillon J., Manter D., Martinez Y., Ricke D., Svensson R., and Jackson P. J. (1999). Sequence and organization of pXO1, the large Bacillus anthracis plasmid harboring the anthrax toxin genes. *J. Bacteriol.* **181**, 6509–6515.
- Økstad O. A., and Kolstø A.-B. (2011). Genomics of Bacillus Species. ‘Genomics of Foodborne Bacterial Pathogens’. (Eds M Wiedmann and W Zhang) pp. 29–53. (Springer New York: New York, NY) doi:10.1007/978-1-4419-7686-4_2
- Okugawa S., Moayeri M., Pomerantsev A. P., Sastalla I., Crown D., Gupta P. K., and Leppla S. H. (2012). Lipoprotein biosynthesis by prolipoprotein diacylglyceryl transferase is

- required for efficient spore germination and full virulence of *Bacillus anthracis*. *Mol. Microbiol.* **83**, 96–109. doi:10.1111/j.1365-2958.2011.07915.x
- Ortlund E. A., Bridgman J. T., Redinbo M. R., and Thornton J. W. (2007). Crystal Structure of an Ancient Protein: Evolution by Conformational Epistasis. *Science* **317**, 1544. doi:10.1126/science.1142819
- Otwinowski Z., and Minor W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
- Pace C. N., Vajdos F., Fee L., Grimsley G., and Gray T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci. Publ. Protein Soc.* **4**, 2411–2423. doi:10.1002/pro.5560041120
- Page R. D. M., and Holmes E. C. (1998). ‘Molecular Evolution: A Phylogenetic Approach.’ (Wiley-Blackwell)
- Painter J., and Merritt E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 439–450. doi:10.1107/S0907444906005270
- Pál C., Papp B., and Lercher M.J. (2006). An integrated view of protein evolution. *Nat Rev Genet.* **7**, 337-348. doi:10.1038/nrg1838
- Papadopoulos G. K., Ouzounis C., and Eliopoulos E. (1998). RGD sequences in several receptor proteins: novel cell adhesion function of receptors? *Int. J. Biol. Macromol.* **22**, 51–57.
- Patthy L. (1985). Evolution of the proteases of blood coagulation and fibrinolysis by assembly from modules. *Cell* **41**, 657–663. doi:10.1016/S0092-8674(85)80046-5
- Patthy L. (1990). Homology of a domain of the growth hormone/prolactin receptor family with type III modules of fibronectin. *Cell* **61**, 13-14.
- Pechmann S., Levy E. D., Tartaglia G. G., and Vendruscolo M. (2009). Physicochemical principles that regulate the competition between functional and dysfunctional association of proteins. *Proc. Natl. Acad. Sci.* **106**, 10159. doi:10.1073/pnas.0812414106
- Pereira de Araujo A. F., and Onuchic J. N. (2009). A sequence-compatible amount of native burial information is sufficient for determining the structure of small globular proteins. *Proc. Natl. Acad. Sci.* **106**, 19001. doi:10.1073/pnas.0910851106
- Petersen T. E., Thogersen H. C., Skorstengaard K., Vibe-Pedersen K., Sahl P., Sottrup-Jensen L., and Magnusson S. (1983). Partial primary structure of bovine plasma fibronectin:

- three types of internal homology. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 137–141. doi:10.1073/pnas.80.1.137
- Petersen T. N., Brunak S., von Heijne G., Nielsen H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **8**, 785–786. doi:10.1038/nmeth.1701.
- Phillips P. C. (2008). Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867. doi:10.1038/nrg2452
- Pichichero M.E., Loe M., Anderson A., and Smith D.H. (1982). Do pili play a role in pathogenicity of *Haemophilus influenzae* type B? *Lancet* 8305, 960–962.
- Pierschbacher M.D., and Ruoslahti E. (1984). Variants of the cell recognition site of fibronectin that retain attachment-promoting activity. *Proc Natl Acad Sci U S A.* **81**, 5985–5988.
- Plaxco K. W., Spitzfaden C., Campbell I. D., and Dobson C. M. (1996). Rapid refolding of a proline-rich all-beta-sheet fibronectin type III module. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10703–10706.
- Pollock D. D., Thiltgen G., and Goldstein R. A. (2012). Amino acid coevolution induces an evolutionary Stokes shift. *Proc. Natl. Acad. Sci.* **109**, E1352. doi:10.1073/pnas.1120084109
- Popowska M., Kusio M., Szymanska P., and Markiewicz Z. (2009). Inactivation of the wall-associated de-N-acetylase (PgdA) of *Listeria monocytogenes* results in greater susceptibility of the cells to induced autolysis. *J. Microbiol. Biotechnol.* **19**, 932–945.
- Potts J., and Campbell I. (1994). Fibronectin structure and assembly. *Curr. Opin. Cell Biol.* **6**, 648–655. doi:10.1016/0955-0674(94)90090-6
- Powell A. J., Conant G. C., Brown D. E., Carbone I., and Dean R. A. (2008). Altered patterns of gene duplication and differential gene gain and loss in fungal pathogens. *BMC Genomics* **9**, 147. doi:10.1186/1471-2164-9-147
- Priest F.G., Barker M., Baillie L.W., Holmes E.C. and Maiden M.C. (2004). Population structure and evolution of the *Bacillus cereus* group. *J Bacteriol* **186**, 7959–7970.
- Prince V. E., and Pickett F. B. (2002). Splitting pairs: the diverging fates of duplicated genes. *Nat. Rev. Genet.* **3**, 827–837. doi:10.1038/nrg928
- Psylinakis E., Boneca I. G., Mavromatis K., Deli A., Hayhurst E., Foster S. J., Varum K. M., and Bouriotis V. (2005). Peptidoglycan N-acetylglucosamine deacetylases from *Bacillus cereus*, highly conserved proteins in *Bacillus anthracis*. *J. Biol. Chem.* **280**, 30856–30863. doi:10.1074/jbc.M407426200

- RAMACHANDRAN G. N., RAMAKRISHNAN C., and SASISEKHARAN V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99.
- Rasko D. A., Altherr M. R., Han C. S., and Ravel J. (2005). Genomics of the *Bacillus cereus* group of organisms. *FEMS Microbiol. Rev.* **29**, 303–329. doi:10.1016/j.femsre.2004.12.005
- Rastogi S., and Liberles D. A. (2005). Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol. Biol.* **5**, 28. doi:10.1186/1471-2148-5-28
- Ravel J., and Fraser C. M. (2005). Genomics at the genus scale. *Trends Microbiol.* **13**, 95–97. doi:10.1016/j.tim.2005.01.004
- Read T. D., Peterson S. N., Tourasse N., Baillie L. W., Paulsen I. T., Nelson K. E., Tettelin H., Fouts D. E., Eisen J. A., Gill S. R., Holtzapple E. K., Okstad O. A., Helgason E., Rilstone J., Wu M., Kolonay J. F., Beanan M. J., Dodson R. J., Brinkac L. M., Gwinn M., DeBoy R. T., Madpu R., Daugherty S. C., Durkin A. S., Haft D. H., Nelson W. C., Peterson J. D., Pop M., Khouri H. M., Radune D., Benton J. L., Mahamoud Y., Jiang L., Hance I. R., Weidman J. F., Berry K. J., Plaut R. D., Wolf A. M., Watkins K. L., Nierman W. C., Hazen A., Cline R., Redmond C., Thwaite J. E., White O., Salzberg S. L., Thomason B., Friedlander A. M., Koehler T. M., Hanna P. C., Kolsto A.-B., and Fraser C. M. (2003). The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* **423**, 81–86. doi:10.1038/nature01586
- Robert X., and Gouet P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-324. doi:10.1093/nar/gku316
- Romero P. A., and Arnold F. H. (2009). Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **10**, 866–876. doi:10.1038/nrm2805
- Romero D., and Palacios R. (1997). Gene amplification and genomic plasticity in prokaryotes. *Annu. Rev. Genet.* **31**, 91–111. doi:10.1146/annurev.genet.31.1.91
- Rossmann M. G. (1972). *The Molecular Replacement Method*. New York: Gordon & Breach.
- Roth C., and Liberles D. A. (2006). A systematic search for positive selection in higher plants (Embryophytes). *BMC Plant Biol.* **6**, 12. doi:10.1186/1471-2229-6-12
- Ruoslahti E., and Pierschbacher M. D. (1987). New perspectives in cell adhesion: RGD and integrins. *Science* **238**, 491–497.
- Saleh B., Allario T., Dambier D., Ollitrault P., and Morillon R. (2008). Tetraploid citrus rootstocks are more tolerant to salt stress than diploid. *C. R. Biol.* **331**, 703–710. doi:10.1016/j.crv.2008.06.007

- Sambrook J., E.F. Fritsch and T. Maniatis, 1989, *Molecular cloning: A laboratory manual*. Second edition. Volumes 1, 2, and 3. 1,626 pp., soft cover Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N
- Sandegren L., and Andersson D. I. (2009). Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat. Rev. Microbiol.* **7**, 578–588. doi:10.1038/nrmicro2174
- Sankaran K., and Wu H. C. (1992). Bacterial Lipoproteins. ‘Lipid Modifications of Proteins’. pp. 163–208. (CRC Press)
- Schaefer C., Schlessinger A., and Rost B. (2010). Protein secondary structure appears to be robust under in silico evolution while protein disorder appears not to be. *Bioinformatics* **26**, 625–631. doi:10.1093/bioinformatics/btq012
- Schneewind O., and Missiakas D. M. (2012). Protein secretion and surface display in Gram-positive bacteria. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **367**, 1123–1139. doi:10.1098/rstb.2011.0210
- Schrödinger, LLC (2015). The PyMOL Molecular Graphics System, Version 1.8.
- Schulenburg H., and Boehnisch C. (2008). Diversification and adaptive sequence evolution of Caenorhabditis lysozymes (Nematoda: Rhabditidae). *BMC Evol. Biol.* **8**, 114. doi:10.1186/1471-2148-8-114
- Schüttelkopf A. W., and van Aalten D. M. F. (2004). *PRODRG*: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363. doi:10.1107/S0907444904011679
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464. doi:10.1214/aos/1176344136.
- Segovia M. (1994). Leishmania gene amplification: a mechanism of drug resistance. *Ann. Trop. Med. Parasitol.* **88**, 123–130. doi:10.1080/00034983.1994.11812849
- Shakhnovich B. E., Deeds E., Delisi C., and Shakhnovich E. (2005). Protein structure and evolutionary history determine sequence space topology. *Genome Res.* **15**, 385–392. doi:10.1101/gr.3133605
- Sharma A., Askari J. A., Humphries M. J., Jones E. Y., and Stuart D. I. (1999). Crystal structure of a heparin- and integrin-binding segment of human fibronectin. *EMBO J.* **18**, 1468–1479. doi:10.1093/emboj/18.6.1468
- Shen S., and Kozikowski A. P. (2016). Why Hydroxamates May Not Be the Best Histone Deacetylase Inhibitors--What Some May Have Forgotten or Would Rather Forget? *ChemMedChem* **11**, 15–21. doi:10.1002/cmdc.201500486

- Shigenobu S., Bickel R. D., Brisson J. A., Butts T., Chang C.-C., Christiaens O., Davis G. K., Duncan E. J., Ferrier D. E. K., Iga M., Janssen R., Lin G.-W., Lu H.-L., McGregor A. P., Miura T., Smaghe G., Smith J. M., van der Zee M., Velarde R. A., Wilson M. J., Dearden P. K., and Stern D. L. (2010). Comprehensive survey of developmental genes in the pea aphid, *Acyrtosiphon pisum*: frequent lineage-specific duplications and losses of developmental genes. *Insect Mol. Biol.* **19 Suppl 2**, 47–62. doi:10.1111/j.1365-2583.2009.00944.x
- Sievers F., Wilm A., Dineen D., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Soding J., Thompson J. D., and Higgins D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539. doi:10.1038/msb.2011.75
- Sigal I. S., Harwood B. G., and Arentzen R. (1982). Thiol-beta-lactamase: replacement of the active-site serine of RTEM beta-lactamase by a cysteine residue. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 7157–7160. doi:10.1073/pnas.79.23.7157
- Siltberg-Liberles J., Grahnen J. A., and Liberles D. A. (2011). The evolution of protein structures and structural ensembles under functional constraint. *Genes* **2**, 748–762. doi:10.3390/genes2040748
- Sonti R. V., and Roth J. R. (1989). Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources. *Genetics* **123**, 19–28.
- Southern E. (1979). Gel electrophoresis of restriction fragments. *Methods Enzymol.* **68**, 152–176.
- Spencer R. C. (2003). *Bacillus anthracis*. *J. Clin. Pathol.* **56**, 182–187. doi:10.1136/jcp.56.3.182
- Spitzfaden C., Grant R., Mardon H., and Campbell I. (1997). Module-module interactions in the cell binding region of fibronectin: stability, flexibility and specificity. *J. Mol. Biol.* **265**, 565–579. doi:10.1006/jmbi.1996.0736
- Stark G. R., and Wahl G. M. (1984). Gene amplification. *Annu. Rev. Biochem.* **53**, 447–491. doi:10.1146/annurev.bi.53.070184.002311
- Starr T. N., and Thornton J. W. (2016). Epistasis in protein evolution. *Protein Sci. Publ. Protein Soc.* **25**, 1204–1218. doi:10.1002/pro.2897
- Steinberg B., and Ostermeier M. (2016). Environmental changes bridge evolutionary valleys. *Sci. Adv.* **2**, e1500921. doi:10.1126/sciadv.1500921

- Stephens S. G. (1951). Possible Significance of Duplication in Evolution. ‘Advances in Genetics’. (Ed M Demerec) pp. 247–265. (Academic Press) doi:10.1016/S0065-2660(08)60237-0
- Strunk R. J., Piemonte K. M., Petersen N. M., Koutsioulis D., Bouriotis V., Perry K., and Cole K. E. (2014). Structure determination of BA0150, a putative polysaccharide deacetylase from *Bacillus anthracis*. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **70**, 156–159. doi:10.1107/S2053230X13034262
- Sukhithasri V., Nisha N., Biswas L., Anil Kumar V., and Biswas R. (2013). Innate immune recognition of microbial cell wall components and microbial strategies to evade such recognitions. *Microbiol Res* **168**, 396-406. doi: 10.1016/j.micres.2013.02.005
- Tabackman A. A., Frankson R., Marsan E. S., Perry K., and Cole K. E. (2016). Structure of ‘linkerless’ hydroxamic acid inhibitor-HDAC8 complex confirms the formation of an isoform-specific subpocket. *J. Struct. Biol.* **195**, 373–378. doi:10.1016/j.jsb.2016.06.023
- Tabashnik B. E. (1990). Implications of gene amplification for evolution and management of insecticide resistance. *J. Econ. Entomol.* **83**, 1170–1176. doi:10.1093/jee/83.4.1170
- Taylor M., and Feyereisen R. (1996). Molecular biology and evolution of resistance of toxicants. *Mol. Biol. Evol.* **13**, 719–734. doi:10.1093/oxfordjournals.molbev.a025633
- Taylor E. J., Gloster T. M., Turkenburg J. P., Vincent F., Brzozowski A. M., Dupont C., Shareck F., Centeno M. S. J., Prates J. A. M., Puchart V., Ferreira L. M. A., Fontes C. M. G. A., Biely P., and Davies G. J. (2006). Structure and activity of two metal ion-dependent acetylxyloxyesterases involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases. *J. Biol. Chem.* **281**, 10968–10975. doi:10.1074/jbc.M513066200
- Templeton T. J. (2009). The varieties of gene amplification, diversification and hypervariability in the human malaria parasite, *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **166**, 109–116. doi:10.1016/j.molbiopara.2009.04.003
- Thelin K. H., and Taylor R. K. (1996). Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by *Vibrio cholerae* O1 El Tor biotype and O139 strains. *Infect Immun* **64**, 2853–2856.
- Tourasse N. J., Helgason E., Okstad O. A., Hegna I. K., and Kolsto A.-B. (2006). The *Bacillus cereus* group: novel aspects of population structure and genome dynamics. *J. Appl. Microbiol.* **101**, 579–593. doi:10.1111/j.1365-2672.2006.03087.x

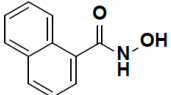
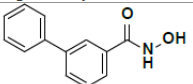
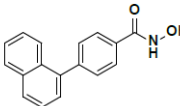
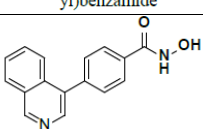
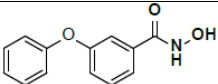
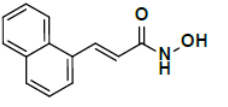
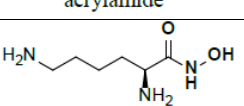
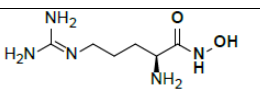
- Tsigos I., Martinou A., Kafetzopoulos D., and Bouriotis V. (2000). Chitin deacetylases: new, versatile tools in biotechnology. *Trends Biotechnol.* **18**, 305–312.
- Urch J. E., Hurtado-Guerrero R., Brosseau D., Liu Z., Eijssink V. G. H., Texier C., and van Aalten D. M. F. (2009). Structural and functional characterization of a putative polysaccharide deacetylase of the human parasite *Encephalitozoon cuniculi*. *Protein Sci. Publ. Protein Soc.* **18**, 1197–1209. doi:10.1002/pro.128
- Vel'kov V. V. (1982). [Gene amplification in prokaryote and eukaryote systems]. *Genetika* **18**, 529–543.
- Ventola C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *P T Peer-Rev. J. Formul. Manag.* **40**, 277–283.
- Vollmer W. (2008). Structural variation in the glycan strands of bacterial peptidoglycan. *FEMS Microbiol. Rev.* **32**, 287–306. doi:10.1111/j.1574-6976.2007.00088.x
- Vollmer W., Blanot D., and de Pedro M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **32**, 149–167. doi:10.1111/j.1574-6976.2007.00094.x
- Vollmer W., and Tomasz A. (2000). The *pgdA* gene encodes for a peptidoglycan N-acetylglucosamine deacetylase in *Streptococcus pneumoniae*. *J. Biol. Chem.* **275**, 20496–20501. doi:10.1074/jbc.M910189199
- Vollmer W., and Tomasz A. (2002). Peptidoglycan N-acetylglucosamine deacetylase, a putative virulence factor in *Streptococcus pneumoniae*. *Infect. Immun.* **70**, 7176–7178. doi:10.1128/IAI.70.12.7176-7178.2002
- de Vos A. M., Ultsch M., and Kossiakoff A. A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **255**, 306–312.
- Voss N. R., and Gerstein M. (2010). 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res.* **38**, W555–W562. doi:10.1093/nar/gkq395
- Wallace A. C., Laskowski R. A., and Thornton J. M. (1995). LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.*, **8**, 127-134.
- Wang Z. O., and Pollock D. D. (2005). Context dependence and coevolution among amino acid residues in proteins. *Methods Enzymol.* **395**, 779–790. doi:10.1016/S0076-6879(05)95040-4
- Webb C. (2003). A complete classification of Darwinian extinction in ecological interactions. *Am. Nat.* **161**, 181–205. doi:10.1086/345858

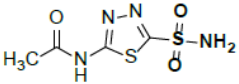
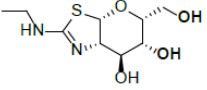
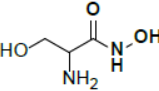
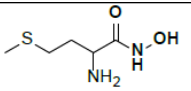
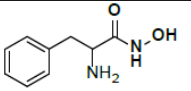
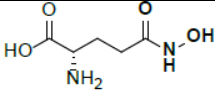
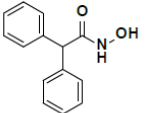
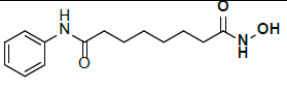
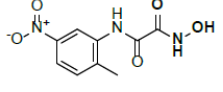
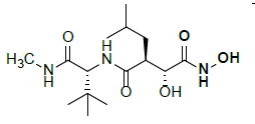
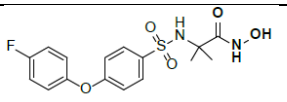
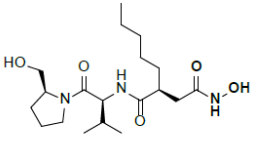
- Wheeler R., Chevalier G., Eberl G., and Gomperts Boneca I. (2014). The biology of bacterial peptidoglycans and their impact on host immunity and physiology. *Cell. Microbiol.* **16**, 1014–1023. doi:10.1111/cmi.12304
- Whittington D. A., Rusche K. M., Shin H., Fierke C. A., and Christianson D. W. (2003). Crystal structure of LpxC, a zinc-dependent deacetylase essential for endotoxin biosynthesis. *Proc. Natl. Acad. Sci.* **100**, 8146. doi:10.1073/pnas.1432990100
- Widholm J. M., Chinnala A. R., Ryu J.-H., Song H.-S., Eggett T., and Brotherton J. E. (2001). Glyphosate selection of gene amplification in suspension cultures of 3 plant species. *Physiol. Plant.* **112**, 540–545.
- Williams S. G., and Lovell S. C. (2009). The Effect of Sequence Evolution on Protein Structural Divergence. *Mol. Biol. Evol.* **26**, 1055–1065. doi:10.1093/molbev/msp020
- Winn M. D., Ballard C. C., Cowtan K. D., Dodson E. J., Emsley P., Evans P. R., Keegan R. M., Krissinel E. B., Leslie A. G. W., McCoy A., McNicholas S. J., Murshudov G. N., Pannu N. S., Potterton E. A., Powell H. R., Read R. J., Vagin A., and Wilson K. S. (2011). Overview of the it CCP4 suite and current developments. *Acta Crystallogr. Sect. D* **67**, 235–242. doi:10.1107/S0907444910045749
- Witt D. (2008). Recent Developments in Disulfide Bond Formation. *Synthesis* **2008**, 2491–2509. doi:10.1055/s-2008-1067188
- Wojcik J., Hantschel O., Grebien F., Kaupe I., Bennett K. L., Barkinge J., Jones R. B., Koide A., Superti-Furga G., and Koide S. (2010). A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. *Nat. Struct. Mol. Biol.* **17**, 519–527. doi:10.1038/nsmb.1793
- Wu S., and Zhang Y. (2007). LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Res.* **35**, 3375–3382. doi:10.1093/nar/gkm251
- Xu D., and Zhang Y. (2011). Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-Step Atomic-Level Energy Minimization. *Biophys. J.* **101**, 2525–2534. doi:10.1016/j.bpj.2011.10.024
- Xu D., and Zhang Y. (2012). Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* **80**, 1715–1735. doi:10.1002/prot.24065
- Xu J., and Mosher D. (2011). Chapter 2 Fibronectin and other adhesive glycoproteins . In Mecham R.P. ed., *The Extracellular Matrix: an Overview, Biology of Extracellular Matrix*, 41 Springer-Verlag Berlin

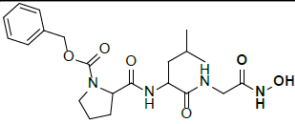
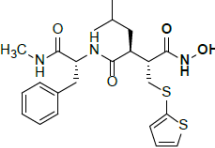
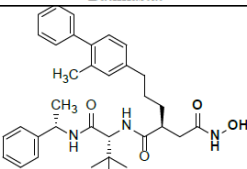
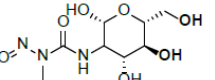
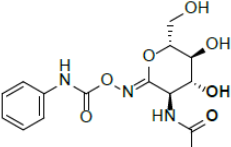
- Yang J., and Zhang Y. (2015). Protein Structure and Function Prediction Using I-TASSER. *Curr. Protoc. Bioinforma.* **52**, 5.8.1-5.8.15. doi:10.1002/0471250953.bi0508s52
- Yang Z. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood Molecular Biology and *Evolution* **24**, 1586–1591. doi:10.1093/molbev/msm088
- Yuzwa S. A., Macauley M. S., Heinonen J. E., Shan X., Dennis R. J., He Y., Whitworth G. E., Stubbs K. A., McEachern E. J., Davies G. J., and Vocadlo D. J. (2008). A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo. *Nat. Chem. Biol.* **4**, 483–490. doi:10.1038/nchembio.96
- Zhang Y., and Skolnick J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins* **57**, 702–710. doi:10.1002/prot.20264
- Zhao J., Ren J., Wang N., Cheng Z., Yang R., Lin G., Guo Y., Cai D., Xie Y., and Zhao X. (2017). Crystal structure of the second fibronectin type III (FN3) domain from human collagen alpha1 type XX. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **73**, 695–700. doi:10.1107/S2053230X1701648X
- Zhou M., Boekhorst J., Francke C., and Siezen R. J. (2008). LocateP: Genome-scale subcellular-location predictor for bacterial proteins. *BMC Bioinformatics* **9**, 173. doi:10.1186/1471-2105-9-173
- Zipperle G. F. J., Ezzell J. W. J., and Doyle R. J. (1984). Glucosamine substitution and muramidase susceptibility in *Bacillus anthracis*. *Can. J. Microbiol.* **30**, 553–559.
- Zwick M. E., Joseph S. J., Didelot X., Chen P. E., Bishop-Lilly K. A., Stewart A. C., Willner K., Nolan N., Lentz S., Thomason M. K., Sozhamannan S., Mateczun A. J., Du L., and Read T. D. (2012). Genomic characterization of the *Bacillus cereus* sensu lato species: backdrop to the evolution of *Bacillus anthracis*. *Genome Res.* **22**, 1512–1524. doi:10.1101/gr.134437.111

Παράρτημα Α' -Πίνακας Ενώσεων –Προσδετών

Πίνακας 1. Λίστα χημικών ενώσεων μικρού μοριακού βάρους που μελετήθηκαν στην παρούσα διατριβή ως προσδέτες (πιθανοί αναστολείς) της Bc1974. Παρατίθενται στον πίνακα, όπου είναι εφικτό, ο κωδικός καταχώρησης των ενώσεων στη βάση PubChem (CID) και εντός παρενθέσεως, ο κωδικός καταχώρησης στην PDB για το αντίστοιχο σύμπλοκο της Bc1974. Το μοριακό βάρος, ο συντελεστής κατανομής σε διάλυμα οκτανόλης/ύδατος (cLogP) και ο δείκτης τοπολογικής επιφάνειας (tPSA) δίδονται στον πίνακα σύμφωνα με την PubChem (Kim *et al.* 2016).

a/a	PubChem CID (PDB ID)	Ένωση Σκελετική δομή Κοινή ονομασία	MW	cLogP	tPSA (Å ²)
1	23382 (5N1P)	 1-Naphthohydroxamic acid	187	2,4	49
2	44369352	 3-Biphenyl hydroxamic acid	213	2,3	49
3	17748273	 N-hydroxy-4-(naphthalen-1-yl)benzamide	263	3,6	49
4	<i>n.a.</i>	 N-hydroxy-4-(isoquinolin-4-yl)benzamide	264	2,2	62
5	10192705	 N-hydroxy-3-phenoxybenzamide	229	3,2	59
6	24880249 (5NC6)	 N-hydroxy-3-(naphthalen-1-yl)acrylamide	213	2,4	49
7	168298 (5NC9)	 L-Lysine hydroxamate	161	-1,7	101
8	151275 (5NCD)	 L-Arginine hydroxamate	189	-2,7	140

9	1986 (5NEK)		222	-0,3	152
		Acetazolamide			
10	122172999 (5NEL)		248	-0,9	120
		Thiamet G			
11	101173		120	-2,5	96
		DL-Serine hydroxamate			
12	352864		164	-0,8	101
		DL-Methionine hydroxamate			
13	352861		180	0,0	75
		DL-Phenylalanine hydroxamate			
14	449178		162	-4,4	113
		N ⁵ -Hydroxy-L-glutamine			
15	239512		227	2,4	49
		N-hydroxy-2,2-diphenylacetamide			
16	5311		264	1,9	78
		Vorinostat (SAHA)			
17	4357555		239	1,1	124
		N'-hydroxy-N-(2-methyl-5-nitrophenyl) oxamide			
18	119031		331	0,5	128
		Marimastat			
19	9907286		368	2,0	113
		CP-471474			
20	10475188		385	1,5	119
		Actinonin			

21	561533		434	1,4	137
Z-Pro-Leu-Gly hydroxamate					
22	5362422		478	3,5	161
Batimastat					
23	10370504		558	6,1	108
UK-356618					
24	45357367		265	-1,4	152
Streptozotocin					
25	9576811		353	-0,3	150
(Z)-PUGNAc					

Παράρτημα Β' –Ευρετήριο Εικόνων

- Εικόνα 1.1. Δομικές ενότητες γλυκοπρωτεϊνών.
- Εικόνα 1.2. Η λιπιδική τροποποίηση στις λιποπρωτεΐνες.
- Εικόνα 1.3. Απεικόνιση του κυτταρικού τοιχώματος ενός κατά Gram-θετικού βακτηρίου.
- Εικόνα 1.4. Οργάνωση του κυτταρικού τοιχώματος των κατά Gram- θετικών.
- Εικόνα 1.5. Η πεπτιδογλυκάνη.
- Εικόνα 1.6. Το περίβλημα του *B.anthraxis* και του *Bacillus cereus*
- Εικόνα 1.7. Δομή του πολυσακχαρίτη του δευτερογενούς κυτταρικού τοιχώματος (SCWP) στελεχών *B. cereus sensu lato*.
- Εικόνα 1.8. (α) Περίβλημα του *B. anthracis* . (β) Έκκριση και συναρμολόγηση των πρωτεϊνών της S-στοιβάδας (γ) συναρμολόγηση των γονιδίων της S-στοιβάδας στον *B. anthracis* και άλλα *B. cereus sensu lato* στελέχη.
- Εικόνα 1.9. Απεικόνιση στελεχών *B.cereus sensu lato*.
- Εικόνα 1.10. Σκελετικές δομές PG ευαίσθητες και ανθεκτικές στη λυσοζύμη.
- Εικόνα 1.11. Σκελετικές δομές υποστρωμάτων των CE4 ενζύμων και αντιπροσωπευτικών απακετυλιωμένων προϊόντων.
- Εικόνα 1.12. Η πρώτη τρισδιάστατη δομή μέλους της CE4
- Εικόνα 1.13. Απεικόνιση της γεωμετρίας συναρμογής του μεταλλοϊόντος στο ενεργό κέντρο.
- Εικόνα 1.14. Οι τρισδιάστατες δομές απακετυλάσεων πολυσακχαρίτη, του *B. cereus* (Bc) και του *B. anthracis* (Ba).
- Εικόνα 2.1. Πλασμιδιακός χάρτης του φορέα κλωνοποίησης pET26b(+) με ένθεμα την Ba3679.
- Εικόνα 2.2. Πλασμιδιακός χάρτης του φορέα έκφρασης pETM-44.
- Εικόνα 3.1. α) SDS-PAGE και χρώση με Coomassie brilliant blue β) SDS-PAGE και ανοσοαποτύπωση κατά Western.
- Εικόνα 3.2. SDS-PAGE και χρώση με Coomassie brilliant blue.
- Εικόνα 3.3. Μέρος της αλληλουχίας του πλασμιδίου pET26b(+) μεταξύ των θέσεων περιορισμού NdeI και XhoI στις περιοχές υβριδισμού με τα εναρκτήρια μόρια Pci-Ba3679-F και Sal-Ba3679-R.
- Εικόνα 3.4. Ηλεκτροφόρηση των προϊόντων της PCR σε πήκτωμα αγαρόζης.
- Εικόνα 3.5. Πλασμιδιακός χάρτης του φορέα έκφρασης pETM-44 με ένθεμα την Ba3679.
- Εικόνα 3.6. Ηλεκτροφόρηση σε πήκτωμα αγαρόζης για τον έλεγχο της υποκλωνοποίησης στον φορέα έκφρασης pETM-44
- Εικόνα 3.7. Αποτύπωμα κηλίδας
- Εικόνα 3.8. Χρωματογραφία συγγενείας της 6xHis-MBP-HRV3C-Ba3679 προσδεδεμένης σε στήλη Ni-NTA. Γράφημα της έκλουσης των προσδεδεμένων πρωτεϊνών που πραγματοποιήθηκε με ρυθμιστικό διάλυμα αυξανόμενης συγκέντρωσης ιμιδαζολίου.
- Εικόνα 3.9. Χρωματογραφία μοριακού αποκλεισμού της Ba3679 Χρωματογράφημα της Ba3679 που απομονώθηκε από χρωματογραφία συγγενείας. Ένθετη εικόνα: SDS-PAGE και χρώση με Coomassie brilliant blue από το κλάσμα των μονομερών
- Εικόνα 3.10. Αντιπροσωπευτικοί κρύσταλλοι Ba3679.
- Εικόνα 3.11. (α) Κρύσταλλος της Ba0330. (β) Εικόνα περίθλασης ακτίνων-X από μονοκρύσταλλο της Ba0330.
- Εικόνα 3.12. Το διάγραμμα Ramachandran της δομής της Ba0330.
- Εικόνα 3.13. Οι θερμικοί παράγοντες για τα άτομα της κύριας αλυσίδας της Ba0330.
- Εικόνα 3.14. Συγκρότηση δευτεροταγούς δομής της Ba0330.
- Εικόνα 3.15. Διάγραμμα τοπολογίας της Ba0330.

- Εικόνα 3.16. Η δομή της Ba0330.
- Εικόνα 3.17. Ηλεκτρονική πυκνότητα (χάρτης 2Fo-Fc) στην περιοχή του ψευδαργύρου
- Εικόνα 3.18. Η κοιλότητα δέσμευσης της Ba0330
- Εικόνα 3.19. Τα MT1-MT5' της κοιλότητας δέσμευσης στην Ba0330
- Εικόνα 3.20. (α) Κρύσταλλος της Ba0331. (β) Εικόνα περίθλασης ακτίνων-X από μονοκρύσταλλο της Ba0331
- Εικόνα 3.21. Γραμμική αναπαράσταση της αλληλουχίας της Ba0331
- Εικόνα 3.22. Το διάγραμμα Ramachandran για τη δομή της Ba0331
- Εικόνα 3.23. Συγκρότηση δευτεροταγούς δομής της Ba0331
- Εικόνα 3.24. Διάγραμμα τοπολογίας της Ba0331
- Εικόνα 3.25. Η δομή της Ba0331.
- Εικόνα 3.26. Οι θερμοικοί παράγοντες (B-factors Å²) για τα άτομα της κύριας αλυσίδας.
- Εικόνα 3.27. Ηλεκτρονική πυκνότητα (χάρτης 2Fo-Fc) στην περιοχή του ψευδαργύρου
- Εικόνα 3.28. Παρουσίαση του σχήματος συναρμογής του ιόντος ψευδαργύρου
- Εικόνα 3.29. Το φερόμενο ενεργό κέντρο της Ba0331
- Εικόνα 3.30. Αναπαράσταση της επιφάνειας του μορίου της Ba0331.
- Εικόνα 3.31. α) κρύσταλλοι Bc1974 β) Εικόνα περίθλασης ακτίνων-X από μονοκρύσταλλο της Bc1974 γ) Bc1974 με τον προσδέτη 1(1- Naphthohydroxamic acid) δ) Εικόνα περίθλασης ακτίνων-X από μονοκρύσταλλο της Bc1974 με τον προσδέτη 1
- Εικόνα 3.32. Φάσματα φθορισμού ακτίνων-X
- Εικόνα 3.33. (α) Η αλληλουχία της Bc1974. (β) Διάγραμμα τοπολογίας της Bc1974
- Εικόνα 3.34. Συγκρότηση της δευτεροταγούς δομής της Bc1974.
- Εικόνα 3.35. Το διάγραμμα Ramachandran της δομής της Bc1974.
- Εικόνα 3.36. Η δομή της Bc1974.
- Εικόνα 3.37. Ηλεκτρονική πυκνότητα στην περιοχή του ψευδαργύρου
- Εικόνα 3.38. Οι θερμοικοί παράγοντες (B-factors Å²) για τα άτομα της κύριας αλυσίδας των μορίων A και B
- Εικόνα 3.39. Το ενεργό κέντρο της Bc1974.
- Εικόνα 3.40. Δομές αναστολέων που περιλαμβάνουν ομάδες υδροξαμικού οξέως και οι προσδέτες που συγκρυσταλλώθηκαν με την Bc1974 (με μπλε).
- Εικόνα 3.41. Τριάδες εικόνων για τον κάθε προσδέτη, (α) Ο χάρτης 2Fo-Fc σχεδιασμένος σε πυκνότητα > 3σ, (β) Αναπαράσταση του ενεργού κέντρου της Bc1974 και (γ) το διάγραμμα lig-plot.
- Εικόνα 3.42. Η μοριακή φυλογενετική ανάλυση της νουκλεοτιδικής αλληλουχίας της δομικής ενότητας NodB των PDA των *B. anthracis* Ames και *B. cereus* ATCC14579
- Εικόνα 3.43. Μοριακή φυλογενετική ανάλυση της περιοχής NodB PDA. (α) των 23 πρωτεϊνικών αλληλουχιών της δομικής ενότητας PDA NodB από τα στέλεχη *B. anthracis* Ames και *B. cereus* ATCC14579 (β) Απεικονίζεται το χωρίς ρίζα φυλογενετικό δένδρο 80 αμινοξικών αλληλουχιών PDA (ομόλογων των στελέχων *B. anthracis* στέλεχος Ames και *B. cereus* ATCC14579. Η ομαδοποίηση των ορθολόγων αναφορικά με τις 23 αλληλουχίες του *B. anthracis* / *B. cereus* παρουσιάζεται με διαφορετικά χρώματα.
- Εικόνα 3.44. Διάγραμμα στοίχισης αλληλουχίας της δομικής ενότητας NodB των είκοσι τριών *B. anthracis* Ames και *B. cereus* ATCC14579 PDAs.
- Εικόνα 3.45. Θέση των συντηρημένων μοτίβων MT1-MT5(') στην δομική ενότητα NodB για τις αντιπροσωπευτικές δομές (α) Ba0331 και (β) Bc1974 της Ομάδας 1 και της Ομάδας 2, αντίστοιχα.
- Εικόνα 3.46. Πολλαπλή ευθυγράμμιση αμινοξικών αλληλουχιών των δομικών ενότητων Fn3 από PDA βακίλλων.
- Εικόνα 3.47. Μοριακή Φυλογενετική ανάλυση της δομικής ενότητας Fn3 των PDAs.

- Εικόνα 3.48. Γράφημα της διάταξης του γονιδιώματος του *B. anthracis* str. Ames (NCBI Reference Sequence: NC_003997.3) στην περιοχή των γονιδίων *ba0330* και *ba0331*
- Εικόνα 3.49. Ευθυγράμμιση αμινοξικής και νουκλεοτιδικής ακολουθίας της Ba0330 και της Ba0331.
- Εικόνα 3.50. Υπέρθωση μέρους της δομικής ενότητας NodB που περιέχει τις θέσεις πρόσδεσης των Ba0331 και Ba0330.
- Εικόνα 4.1. Ευθυγράμμιση αλληλουχιών της δομικής ενότητας NodB των απακετυλασών πεπτιδογλυκάνης Bc1974, Ba1977, των φερόμενων απακετυλασών πολυσακχαριτών Ba0330, Bc0361, Ba0331 με αντιπροσωπευτικά μέλη της οικογένειας CE4
- Εικόνα 4.2. Υπέρθωση των θέσεων δέσμωσης της SpPgdA με φούξια (PDB ID: 2C1G) και της Ba0330 με πράσινο χρώμα.
- Εικόνα 4.3. Υπέρθωση της NodB ομόλογης δομικής ενότητας των Bc1974 (PDB ID: 5N1J, πράσινο), Bc1960 (PDB ID: 4L1G, κυανό) και SpPgdA (PDB ID: 2C1G, φούξια).
- Εικόνα 4.4. Αναπαράσταση της επιφάνειας της θέσης δέσμωσης της δομικής ενότητας NodB με υπέρθεση των κρυσταλλογραφικά προσδιορισμένων PDAs των *B. anthracis* και *B. cereus*
- Εικόνα 4.5. Οι όγκοι της κοιλότητας δέσμωσης των πειραματικά προσδιορισμένων με κρυσταλλογραφία ακτίνων-X απακετυλασών πολυσακχαρίτη των *B. anthracis* και *B. cereus*.
- Εικόνα 4.6. Αρχιτεκτονική της δομής της πρωτεΐνης Ba0331 παρουσιάζεται σε μεγέθυνση η τοπολογία της περιοχής της διεπαφής μεταξύ NodB (χρωματίζεται με ροζ) και Fn3 (πράσινο).
- Εικόνα 4.7. (α) Αναπαράσταση της επιφάνειας της διεπαφής των NodB και Fn3 καθώς και των αμινοξέων (με ράβδους) της Ba0331 και (β) Η αντίστοιχη διεπαφή της Ba0330.
- Εικόνα 4.8. (α) Η υπέρθεση των δομικών ενοτήτων NodB της Ba0331, της Ba0330 και της Bc0361 (PDB ID: 4HD5) αποκαλύπτει την περιστροφική μετατόπιση της αντίστοιχης δομικής ενότητας τύπου Fn3. (β) υπέρθεση των Fn3. (γ) Σχετικός προσανατολισμός των δομικών ενοτήτων τύπου Fn3 μετά την υπέρθεση των δομικών ενοτήτων τους NodB. Η Fn3 δομική ενότητα της Bc0361 περιστρέφεται κατά 16°, ενώ εκείνη της Ba0330 κατά 11° σε σχέση με αυτήν της Ba0331.
- Εικόνα 4.9. Αλληλεπιδράσεις της κρυσταλλικής μοριακής διευθέτησης της Fn3 τύπου δομικής ενότητας στη Ba0331 (α και β), στη Ba0330 (γ και δ) και στη Bc0361 (ε και στ).
- Εικόνα 4.10. Ηλεκτροστατικό επιφανειακό δυναμικό της (α) Ba0330, (β) Bc0361(PDB ID: 4HD5) (γ) Ba0331 και (δ) Bc1974.
- Εικόνα 4.11. Υπέρθωση των θέσεων πρόσδεσης της Bc1974 και Ba0331
- Εικόνα 4.12. Αναπαράσταση της επιφάνειας της καταλυτικής ενότητας NodB των Ba0331, Ba0330 και Bc0361 που παρουσιάζεται η συντήρηση.
- Εικόνα 4.13. Η συντήρηση των αλληλουχιών της Ομάδας 1.
- Εικόνα 4.14. Αναπαράσταση της επιφάνειας των δομών των Ba0331, Ba0330 και Bc0361
- Εικόνα 4.15. Οι προσδέτες των συμπλόκων με την Bc1974 της παρούσης διατριβής.
- Εικόνα 4.16. Μοριακά μοντέλα του συμπλόκου Bc1974-(GlcNAc)₃ που απεικονίζει (α) την ανοικτή διαμόρφωση της κοιλότητας δέσμωσης του υποστρώματος στο μόριο A σε σχέση (β) με το B.
- Εικόνα 4.17. Υπέρθωση της κρυσταλλικής δομής της VcCDA με σύμπλοκο με τριόζη (ροζ) (PDB ID: 4OUI) με τα μοριακά μοντέλα του (GlcNAc)₃ στην κοιλότητα δέσμωσης της Bc1974 (α) Η περιοχή MT4 της Bc1974 υιοθετεί την ανοικτή διαμόρφωσή της όπως βρέθηκε στα μόρια A και C της κρυσταλλικής δομής, ενώ στην (β) υιοθετεί την κλειστή διαμόρφωσή της που βρίσκεται στα μόρια B και D.

Εικόνα 4.18. (α) Κοντινή όψη στο ενεργό κέντρο του μοντέλου Bc1974-(GlcNAc)₃. (β) Προτεινόμενος καταλυτικός μηχανισμός της Bc1974.

Εικόνα 4.19. Υπέρθωση των μονομερών Α και Β της ασύμμετρης μονάδας της Bc1974 χωρίς προσδέτη.

Εικόνα 4.20. Ημι-διαφανής αναπαράσταση της επιφανείας του ενεργού κέντρου της Bc1974 που απεικονίζει βασικά φαρμακοφορικά στοιχεία όπως ταυτοποιήθηκαν από τις κρυσταλλικές δομές των συμπλόκων της.

Παράρτημα Γ' –Ευρετήριο Πινάκων

Πίνακας 3.1. Συλλογή και επεξεργασία δεδομένων της Ba0330

Πίνακας 3.2. Επίλυση και βελτιστοποίηση της δομής Ba0330

Πίνακας 3.3. Συλλογή και επεξεργασία δεδομένων της Ba0331

Πίνακας 3.4. Επίλυση και βελτιστοποίηση της δομής Ba0331

Πίνακας 3.5 Συλλογή δεδομένων, στατιστικά και βελτιστοποίηση της Bc1974 και συμπλόκων της με προσδέτες

Πίνακας 3.6 Οι PDAs του *Bacillus anthracis* Ames που συμπεριλήφθησαν στη μελέτη και οι ομόλογες από τον *Bacillus cereus* ATCC 14579.

Πίνακας 3.7. Μήτρα της ομολογίας (ταυτότητα και ομοιότητα) επί τοις εκατό, μεταξύ των αμινοξικών αλληλουχιών της δομικής ενότητας NodB της οικογένειας των γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) του *B. anthracis* Ames.

Πίνακας 3.8. Μήτρα της ομολογίας (ταυτότητα και ομοιότητα) επί τοις εκατό, μεταξύ των αμινοξικών αλληλουχιών της δομικής ενότητας NodB της οικογένειας των γονιδίων απακετυλάσης πολυσακχαρίτη (PDA) του *B. cereus* ATCC14579.

Πίνακας 3.9 Οι αλληλουχίες Ba0330, Ba0331, Bc0361

Πίνακας 3.10 Πίνακας διαφορών αμινοξικής αλληλουχίας ανάμεσα στην Ba03330 και Ba0331

Πίνακας 4.1 Μετρήση του όγκου της κοιλότητας δέσμησης.

Πίνακας 4.2 Ανάλυση κίνησης με χρήση του προγράμματος DynDom

Πίνακας 4.3. Γωνίες συστροφής των εμπλεκόμενων αμινοξέων.

Παράρτημα Δ' -Δημοσιεύσεις / Ανακοινώσεις σε συνέδρια

Μέρος των αποτελεσμάτων της διατριβής έχουν συμπεριληφθεί στις κάτωθι Δημοσιεύσεις σε ξενόγλωσσα περιοδικά με κριτές:

- **A. Andreou**, P. Giastas, S. Arnaouteli, M. Tzanodaskalaki, S. Tzartos, K. Bethanis, V. Bouriotis and E. E. Eliopoulos (2019). The putative polysaccharide deacetylase Ba0331: cloning, expression, crystallization and structure determination. *Acta Crystallogr Sect F Structural Biology Com.*, 75 (4): 312-20.
- **A. Andreou**, P. Giastas, E. Christoforides, E. E. Eliopoulos (2018). Structural and Evolutionary Insights within the Polysaccharide Deacetylase Gene Family of *Bacillus anthracis* and *Bacillus cereus*. *Genes*, 9 (8): 386.
- P. Giastas, **A. Andreou**, A. Papakyriakou, D. Koutsioulis, S. Balomenou, S. J. Tzartos, V. Bouriotis, and E. E. Eliopoulos (2018). Structures of the Peptidoglycan N-Acetylglucosamine Deacetylase Bc1974 and Its Complexes with Zinc Metalloenzyme Inhibitors *Biochemistry*, 57 (5): 753–63.
- S. Arnaouteli, P. Giastas, **A. Andreou**, M. Tzanodaskalaki, C. Aldridge, S. J. Tzartos, W. Vollmer, E. Eliopoulos, V. Bouriotis (2015). Two Putative Polysaccharide Deacetylases Are Required for Osmotic Stability and Cell Shape Maintenance in *Bacillus anthracis*. *J Biol Chem*. 21: 13465-78.

Δημοσιεύσεις κατά το διάστημα εκπόνησης της διατριβής σε διαφορετικό θέμα.

- Zervou M. I., Dimopoulou D. G., Eliopoulos E., Trachana M., Pratsidou-Gkertsis P., **Andreou A.**, Sidiropoulos P., Spandidos D. A., Garyfallos A., Goulielmos G.N. (2017). The genetics of juvenile idiopathic arthritis: Searching for new susceptibility loci. *Mol Med Rep*. 16:6 8793-8.

Παροχή επικουρικού εκπαιδευτικού έργου σε προπτυχιακούς φοιτητές κατά το διάστημα εκπόνησης της διατριβής.

Πρακτική Άσκηση και Πτυχιακές εργασίες προπτυχιακών φοιτητών στο Εργαστήριο Γενετικής του Γ.Π.Α. Αναφέρονται τα θέματα των πτυχιακών μελετών που εκπονήθηκαν:

- «ΣΥΣΧΕΤΙΣΗ ΓΕΝΕΤΙΚΟΥ ΠΟΛΥΜΟΡΦΙΣΜΟΥ ΤΟΥ ΓΟΝΙΔΙΟΥ C1Q ΜΕ ΤΑ ΑΥΤΟΑΝΟΣΑ ΝΟΣΗΜΑΤΑ ΤΟΥ ΣΑΚΧΑΡΩΔΗ ΔΙΑΒΗΤΗ ΤΥΠΟΥ Ι ΚΑΙ ΤΟΥ ΣΥΣΤΗΜΑΤΙΚΟΥ ΕΡΥΘΗΜΑΤΩΔΗ ΛΥΚΟΥ» παρουσιάστηκε Απρίλιο 2018.
- «ΣΥΣΧΕΤΙΣΗ ΤΟΥ ΓΕΝΕΤΙΚΟΥ ΠΟΛΥΜΟΡΦΙΣΜΟΥ ΣΤΟ ΙΝΤΡΟΝΙΟ 4 ΤΟΥ ΓΟΝΙΔΙΟΥ ΤΗΣ ΕΝΔΟΘΗΛΙΑΚΗΣ ΙΣΟΜΟΡΦΗΣ ΤΗΣ ΣΥΝΘΕΤΑΣΗΣ ΤΟΥ ΜΟΝΟΞΕΙΔΙΟΥ ΤΟΥ ΑΖΩΤΟΥ (eNOS) ΜΕ ΤΟ ΑΥΤΟΑΝΟΣΟ ΝΟΣΗΜΑ ΤΗΣ ΝΕΑΝΙΚΗΣ ΡΕΥΜΑΤΟΕΙΔΟΥΣ ΑΘΡΙΤΙΔΑΣ» παρουσιάστηκε Ιούνιο 2016.
- «ΣΥΣΧΕΤΙΣΗ ΤΟΥ rs11003125C/G SNP ΤΟΥ ΓΟΝΙΔΙΟΥ MBL2 ΜΕ ΤΑ ΑΥΤΟΑΝΟΣΑ ΝΟΣΗΜΑΤΑ ΤΟΥ ΣΑΚΧΑΡΩΔΗ ΔΙΑΒΗΤΗ ΤΥΠΟΥ Ι ΚΑΙ ΤΟΥ ΣΥΣΤΗΜΑΤΙΚΟΥ ΕΡΥΘΗΜΑΤΩΔΟΥΣ ΛΥΚΟΥ» παρουσιάστηκε Μάιο 2014.
- «ΚΡΥΣΤΑΛΛΟΓΡΑΦΙΚΕΣ ΜΕΛΕΤΕΣ ΤΗΣ ΠΡΩΤΕΪΝΗΣ ΤΟΥ ΓΑΛΑΚΤΟΣ β-ΛΑΚΤΟΣΦΑΙΡΙΝΗΣ» παρουσιάστηκε Σεπτέμβριο 2013.

**Επιλεγμένες ανακοινώσεις σε Συνέδρια
όπου μέρος των αποτελεσμάτων της διατριβής έχουν συμπεριληφθεί**

Συμμετοχή στις ανακοινώσεις:

- 9th International Conference of the Hellenic Crystallographic Association, Συνεδριακό και Πολιτιστικό Κέντρο Πανεπιστημίου Πατρών, 05-07 Οκτωβρίου 2018. Τίτλος ανακοίνωσης: «*Evolution through the Crystal Lens*».
- 12th Conference of the Hellenic Society for Computational Biology and Bioinformatics, Ελληνικό Ινστιτούτο Παστέρ, Αθήνα, 11-13 Οκτωβρίου 2017. Τίτλος ανακοίνωσης: «*Pharmacophore assertions: towards the design of potent inhibitors using structural analysis of liganded deacetylase complexes*».
- Joint Conference of the Hellenic Crystallographic Association and the Hellenic Society for Computational Biology and Bioinformatics, Γεωπονικό Πανεπιστήμιο Αθηνών, 07-09 Οκτωβρίου 2016. Τίτλος ανακοίνωσης: «*Structural studies on the peptidoglycan N-acetylglucosamine deacetylase Bc1974 from Bacillus cereus and reflections to the Ba1977 from Bacillus anthracis*».
- FP7-REGPOT Project SEE-DRUG Conference NMR Applications in Life Science and Exploring Peptides & Proteins, Συνεδριακό και Πολιτιστικό Κέντρο Πανεπιστημίου Πατρών, 18-20 Ιουνίου 2015. Τίτλος ανακοίνωσης: «*Structure based inhibitor design targeting the bacterial GlcNAc deacetylase Bc1974*».
- Current Trends in Structural Biology & 7th International Conference of the Hellenic Crystallographic Association, Ίδρυμα Τεχνολογίας και Έρευνας, Ηράκλειο Κρήτης, 19-21 Σεπτεμβρίου 2014. Τίτλος ανακοίνωσης: «*Diverse bacterial polysaccharide deacetylases contain protein-protein interface features*».

Επιλεγμένες ανακοινώσεις σε Συνέδρια κατά το διάστημα εκπόνησης της διατριβής σε διαφορετικό θέμα

Συμμετοχή στις ανακοινώσεις:

- Annual European Congress of Rheumatology, session: Genomics, genetic basis of disease and HLA/T cell recognition Madrid, Spain, 14-17 Ιουνίου 2017. Τίτλος ανακοίνωσης: «*Investigation of juvenile idiopathic arthritis in Greece: new susceptibility loci*».
- Joint Conference of the Hellenic Crystallographic Association and the Hellenic Society for Computational Biology and Bioinformatics, Γεωπονικό Πανεπιστήμιο Αθηνών, 07-09 Οκτωβρίου 2016. Τίτλος ανακοίνωσης: «*Employing a Structural Model and the Mechanism of Insect Chitin Deacetylase for the Design of Novel Control Agents*».
- 36^o Επιστημονικό Συνέδριο Ελληνικής Εταιρείας Βιολογικών Επιστημών, Ιωάννινα, 8-10 Μαΐου 2014. Τίτλος ανακοίνωσης: «*Εξελικτική μελέτη μιτοχονδριακού γονιδιώματος 51 ειδών Ημιπτέρων*».
- Current Trends in Structural Biology & 7th International Conference of the Hellenic Crystallographic Association, Ίδρυμα Τεχνολογίας και Έρευνας, Ηράκλειο Κρήτης, 19-21 Σεπτεμβρίου 2014. Τίτλος ανακοίνωσης: «*β-Lactoglobulin as a carrier of molecules with biological significance*».