ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ

ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ ΚΑΙ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

«ΕΦΑΡΜΟΓΕΣ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ – ΜΟΡΙΑΚΗ ΓΕΝΕΤΙΚΗ – ΔΙΑΓΝΩΣΤΙΚΟΙ ΔΕΙΚΤΕΣ»

ΔΗΜΙΟΥΡΓΙΑ ΦΟΡΕΑ ΕΚΦΡΑΣΗΣ ΤΗΣ ΚΑΡΒΟΞΥΤΕΛΙΚΗΣ ΕΠΙΚΡΑΤΕΙΑΣ (CBM20) ΤΗΣ ΑΝΘΡΩΠΙΝΗΣ STBD1

ΑΘΑΝΑΣΙΟΣ ΤΟΥΛΙΟΣ

Λάρισα 2022

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

ΔΗΜΙΟΥΡΓΙΑ ΦΟΡΕΑ ΕΚΦΡΑΣΗΣ ΤΗΣ ΚΑΡΒΟΞΥΤΕΛΙΚΗΣ ΕΠΙΚΡΑΤΕΙΑΣ (CBM20) ΤΗΣ ΑΝΘΡΩΠΙΝΗΣ STBD1

CONSTRUCTION OF AN EXPRESSION VECTOR OF THE C-TERMINAL DOMAIN (CBM20) OF HUMAN STBD1

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Τριμελής Επιτροπή

Βασιλική Σκαμνάκη (επιβλέπουσα):

Επίκουρος καθηγήτρια Βιοχημείας – Μεταβολισμού, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Θεσσαλίας, Λάρισα, Ελλάδα

Δημήτριος Λεωνίδας:

Καθηγητής Βιοχημείας, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Θεσσαλίας, Λάρισα, Ελλάδα

Άννα-Μαρία Ψαρρά:

Αναπληρώτρια καθηγήτρια Βιοχημείας, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Θεσσαλίας, Λάρισα, Ελλάδα

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

ΕΥΧΑΡΗΣΤΙΕΣ

Η μεταπτυχιακή εργασία εκπονήθηκε στο εργαστήριο Δομικής και Λειτουργικής Βιοχημείας του Τμήματος Βιοχημείας και βιοτεχνολογίας, Σχολή Επιστημών Υγείας του Πανεπιστημίου Θεσσαλίας. Ξεκίνησα να συμμετέχω στις εργασίες του εργαστηρίου Δομικής και Λειτουργικής Βιοχημείας το 2020 εκπονώντας την πτυχιακή μου εργασία. Το εργαστήριο αυτό δεν αποτέλεσε αρχική μου επιλογή. Ωστόσο, η διερεύνηση των δυνατοτήτων που παρέχει η δομική βιοχημεία, άλλαξε ριζικά τον τρόπο σκέψης μου γύρω από τη σημασία των λειτουργιών των μορίων μέσα σε ένα κύτταρο. Η καθημερινή τριβή μου με το πεδίο αυτό μού προκάλεσε θαυμασμό παρατηρώντας τον καλά καθορισμένο συντονισμό τόσο των ενδομοριακών αλληλεπιδράσεων, που θα καθορίσουν τη διαμόρφωση των δομών και της λειτουργίας των βιολογικών μορίων, όσο των διαμοριακών αλληλεπιδράσεων που θα καθοριστούν από ολοκληρωμένες μορφές μέσα σε ένα υποσύνολο, ρυθμίζοντας και συντονίζοντας τη λειτουργία και την επιβίωση του ενδόκοσμου του κυττάρου. Η ομορφιά του μικρόκοσμου αυτού προκάλεσε το σεβασμό μου πάνω στη μελέτη του.

Θα ήθελα να ευχαριστήσω δεόντως την επιβλέπουσα της μεταπτυχιακής εργασίας μου, κα Επίκουρος Καθηγήτρια Βιοχημείας Βασιλική Σκαμνάκη, που μου επέτρεψε να ασχοληθώ με αυτό το θέμα. Την ευχαριστώ γιατί μου επέτρεψε να μελετήσω και να διερευνήσω καλύτερα και με μεγαλύτερη λεπτομέρεια το πεδίο αυτό.

Ένα τεράστιο ευχαριστώ στους συναδέλφους: Γεώργιο Στραβοδήμο, Χριστίνα Δράκου, Συμεών Κούλα, Σεραφείμ Αλεξόπουλο, Αχιλλέα Γεωργαντόπουλο, Ιωάννη Τσιάλτα, Όλγα Παπαϊωάννου που ήταν εκεί να με καθοδηγήσουν και να με βοηθήσουν σε όλα τα βήματα μέσα στο εργαστήριο.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

1 Πίνακας περιεχομένων

ΕΥΧΑΡΗΣΤΙΕΣ III
1 Πίνακας περιεχομένων IV
1 Εισαγωγή 1
1.1 Γλυκογόνο 1
1.1.1 Η δομή του γλυκογόνου 1
1.1.2 Η σύνθεση του γλυκογόνου 2
1.1.3 Αποικοδόμηση του γλυκογόνου 3
1.1.4 Η γλυκογονόλυση μέσω αυτοφαγίας 3
1.1.4.1 Αυτοφαγία 3
1.1.4.2 Γλυκοφαγία 4
1.2 Carbohydrate-Binding Modules 5
1.2.1 CBM20 7
1.2.1.1 Κατηγοριοποίηση 8
1.2.1.2 Συντηρημένα κατάλοιπα αμινοξέων 9
1.2.1.3 Αρχιτεκτονική CBM20 και δομικά χαρακτηριστικά 10
1.3 Starch Binding Domain-containing Protein 1 (STBD1) 12
2 ΣΚΟΠΟΣ 15
3 ΥΛΙΚΑ 16
3.1 Όργανα 16
3.2 Βακτηριακές κυτταρικές σειρές 16
3.3 Πλασμιδιακοί φορείς 16
3.4 Αντιδραστήρια 16
4 ΜΕΘΟΔΟΙ 18
4.1 Αρχή μεθόδων 18
4.1.1 Συστήματα Βιοπληροφορικής 18
4.1.1.1 ProtParam 18
4.1.1.2 DOMAINATION 19
4.1.1.3 INTERPRO 19
4.1.1.4 I-TASSER 20
4.1.1.5 PONDR 24
4.1.2 Παραγωγή Πρωτεϊνών στο Ετερόλογο Σύστημα *E.coli* 24
4.1.3 Τεχνολογία Ανασυνδυασμένου DNA 25
4.1.3.1 Κλωνοποίηση 25

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

4.1.3.2 Πλασμιδιακός φορέας 25
4.1.3.3 Ηλεκτροφόρηση σε πήκτωμα πολυακρυλαμίδης υπό αποδιατακτικές συνθήκες (SDS-PAGE) 27
4.1.3.4 Ανοσοαποτύπωση Πρωτεϊνών (Western Blot) 28
4.2 Βιοπληροφορική ανάλυση της πρωτεϊνικής επικράτειας CBM20 της πρωτεΐνης STBD1 28
4.3 Κατασκευή πλασμιδιακών φορέων 28
4.4 Χημικός μετασχηματισμός βακτηριακών κυττάρων 33
4.5 Δοκιμές έκφρασης και παραγωγής της πρωτεϊνικής επικράτειας CMB20 33
5 ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ 35
5.1 Βιοπληροφορική ανάλυση 35
5.1.1 Πρόβλεψη και ανάλυση δομής STBD1 35
5.1.1.1 Δευτεροταγής δομή και αστάθεια περιοχών 35
5.1.1.2 Τριτοταγής δομή 36
5.1.1.3 Καθορισμός της CBM20 επικράτειας της STBD1 38
5.1.2 Πρόβλεψη και ανάλυση της επικράτειας CBM20 της STBD1 38
5.1.2.1 Δευτεροταγής δομή και αστάθεια περιοχών 38
5.1.2.2 Τριτοταγής δομή 39
5.1.2.3 Πιθανά συντηρημένα κατάλοιπα στις θέσεις σύνδεσης των προσδετών της CBM20 41
5.1.2.4 Προσδέτες και θέσεις σύνδεσης 42
5.1.3 Ανασυνδυασμένα πλασμίδα 43
5.1.4 Δοκιμή έκφρασης 44
6 ΕΠΙΛΟΓΟΣ 48
7 Βιβλιογραφία i
8 Παράρτημα I
ΠΑΡΑΡΤΗΜΑ I: Αλληλουχίες από τη βάση δεδομένων *Protein* του NCBI I
ΠΑΡΑΡΤΗΜΑ II: ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΑΜΙΝΟΞΕΩΝ III
ΠΑΡΑΡΤΗΜΑ III: ΕΚΚΙΝΗΤΕΣ III
ΠΑΡΑΡΤΗΜΑ IV: Αποτελέσματα αλληλούχησης για τα constructs pETM-11(CBM20) και pGEX-6P-1(CBM20) III

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

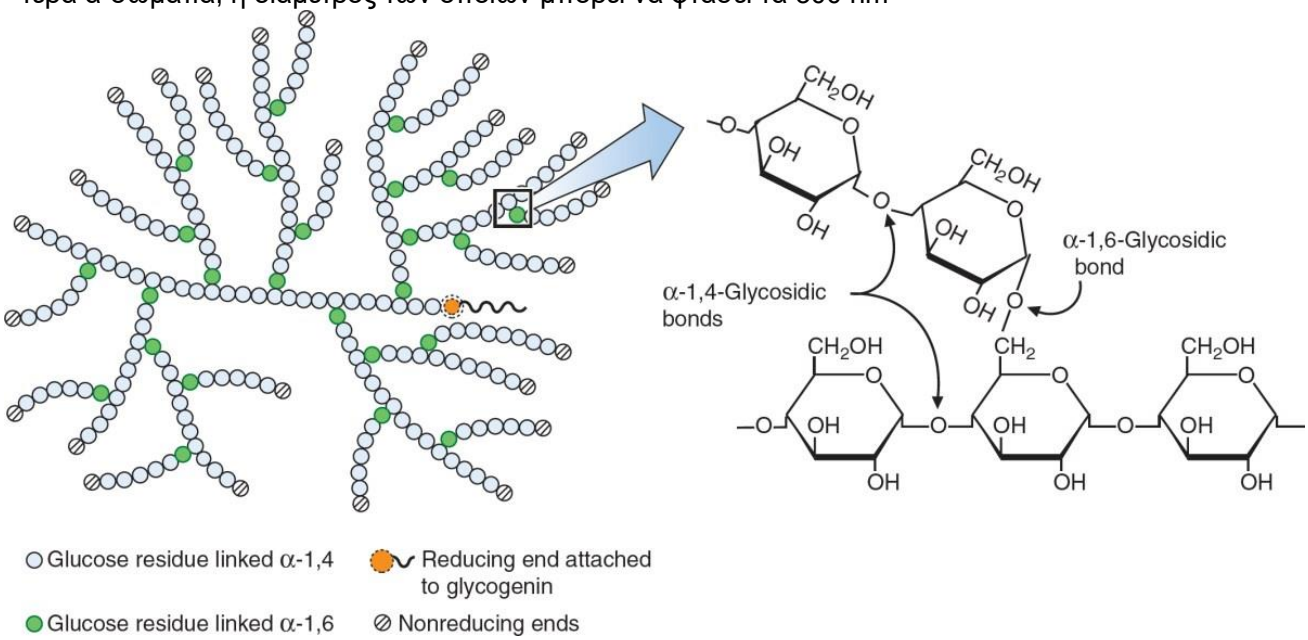
1 Εισαγωγή

1.1 Γλυκογόνο

1.1.1 Η δομή του γλυκογόνου

Το γλυκογόνο είναι ένα διακλαδισμένο πολυμερές καταλοίπων γλυκόζης που εξυπηρετεί ως ένα μέσο αποθήκευσης γλυκόζης για τη χρήση της και την παραγωγή ενέργειας σε περιπτώσεις που κρίνεται διατροφικά αναγκαίο από τον οργανισμό. Εντοπίζεται σε όλους τους οργανισμούς από τα βακτήρια μέχρι τα αρχαία και τον άνθρωπο. Τα φυτά συνθέτουν το άμυλο, ένα πολυμερές που συντίθεται από κατάλοιπα αμυλοπεκτίνης, έναν πολυσακχαρίτη χημικά όμοιο με το γλυκογόνο, και αμυλόζης, ένα μη διακλαδισμένο γραμμικό πολυμερές καταλοίπων γλυκόζης. Ο πολυμερισμός της γλυκόζης είναι ένας οικουμενικός μηχανισμός αποθήκευσης ενέργειας στη φύση[1].

Τα μόρια της γλυκόζης στο γλυκογόνο συνδέονται τόσο γραμμικά με α-1,4-γλυκοζιτικούς δεσμούς, αλλά και ανά 6 - 10 μόρια γλυκόζης σχηματίζοντας διακλαδώσεις από α-1,6-γλυκοζιτικούς δεσμούς [2]. Οι θέσεις των διακλαδώσεων δεν είναι γνωστές πάνω στο μόριο, με αποτέλεσμα μόρια γλυκογόνου παρόμοιου μοριακού βάρους να έχουν διαφορετική χημική δομή και κατ' επέκταση δεν υπάρχει ακριβής τριδιάστατη δομή του. Το γλυκογόνο οργανώνεται σε σωμάτια, στα οποία είναι συνδεδεμένο με πρωτεΐνες. Η δομή των σωματίων γλυκογόνου μπορεί να διαχωριστεί σε τρία επίπεδα: Σε μορφή ολιγομερούς μικρής αλυσίδας (συνήθως αναφέρεται ως γ-σωμάτιο), σε σφαιρικά β-σωμάτια με μέση διάμετρο 20 nm και σε μεγαλύτερα α-σωμάτια, η διάμετρος των οποίων μπορεί να φτάσει τα 300 nm [3]



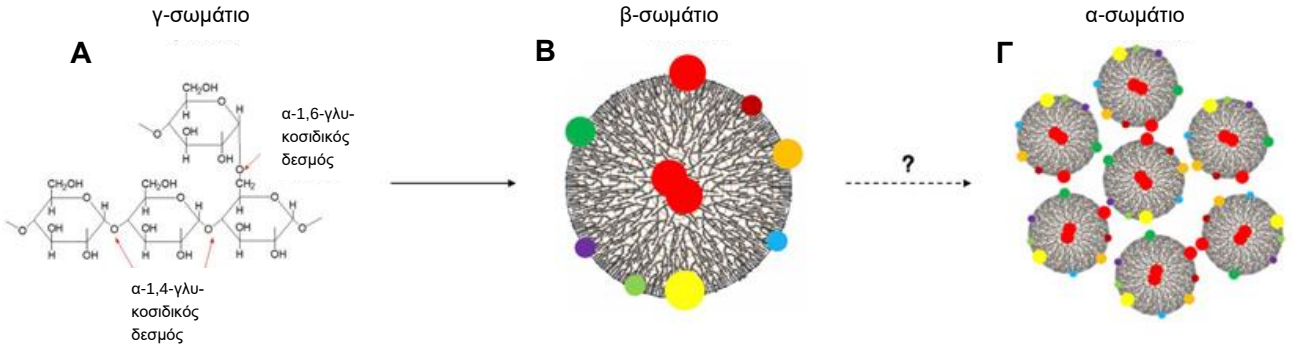
Εικόνα 1 Το γλυκογόνο αποτελείται από γλυκοσυλικές μονάδες που συνδέονται με α-1.4-γλυκοζυτικούς δεσμούς και α-1.6-γλυκοζυτικούς δεσμούς στις θέσεις διακλάδωσης [2]

Τα β-σωμάτια αποτελούνται από εξωτερικές Α-αλυσίδες που δεν φέρουν κάποια διακλάδωση και από εσωτερικές Β-αλυσίδες που διαθέτουν κατά μέσο όρο δύο διακλαδώσεις (Μοντέλο Whelan). Βάσει αυτού του μοντέλου, σε ένα τυπικό β σωμάτιο γλυκογόνου με διάμετρο 42 nm εντοπίζονται περίπου 55000 κατάλοιπα γλυκόζης[4]. Ωστόσο, το 2015, ο Besford, παρουσίασε με τη χρήση σκέδασης ακτινών Χ μικρής

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

γωνίας (small angle X-ray scattering) ότι στα β-σωμάτια γλυκογόνου των ποντικών ο πολυμερισμός είναι τυχαίος και όχι ιεραρχικός[5].

Τα α-σωμάτια θεωρούνται ως συσωματώματα β-σωματίων που συνδέονται μέσω δισουλφιδικών δεσμών ανάμεσα σε βασικές πρωτεΐνες που σχετίζονται άμεσα με το γλυκογόνο. Ο ακριβής μηχανισμός σχηματισμού των α-σωματίων δεν έχει χαρακτηριστεί καλά [1,6].



Εικόνα 2 **Σχηματική παρουσίαση των ευκαρυωτικών σωματίων γλυκογόνου:. (A) Μικρές αλυσίδες ολιγομερών. (B) β-σωμάτιο, (C) α-σωμάτιο. Το γλυκογόνο συνδέεται με πρωτεΐνες σχηματίζοντας το υποκυτταρικό πρωτέωμα του γλυκογόνου και παρουσιάζεται με χρωματιστές τελείες στο σχήμα. Το ερωτηματικό υποδεικνύει τον άγνωστο μηχανισμό αλληλεπίδρασης των β-σωματίων προς τον σχηματισμό των α-σωματίων[3].**

Οι διαφορετικοί τύποι του γλυκογόνου εντοπίζονται σε διαφορετικούς ιστούς. Οι αποθήκες γλυκογόνου στο ήπαρ αποτελούνται κυρίως από α-σωμάτια με τον λόγο της επιφάνειας προς τον όγκο του να είναι χαμηλός, χωρίς να είναι εύκολα προσβάσιμο από ένζυμα πολυμερισμού και με χαμηλή απελευθέρωση γλυκόζης από αυτό. Στο σκελετικό μυ, τα μικρότερα β-σωμάτια κυριαρχούν με τον λόγο της επιφάνειας προς τον όγκο του να είναι υψηλός, έχοντας τα ένζυμε πολυμερισμού μεγαλύτερη πρόσβαση ώστε να επάγεται μεγαλύτερη απελευθέρωση ελεύθερης γλυκόζης για την παραγωγή ενέργειας. Στον καρδιακό ιστό, εντοπίζονται και οι δύο τύποι σωματίων και σε περίπτωση μεταβολικού στρες μειώνεται η σύσταση των α-σωματίων από τα β-σωμάτια [6,7].

1.1.2 Η σύνθεση του γλυκογόνου

Η σύνθεση του γλυκογόνου απαιτεί τη συντονισμένη δράση διαφόρων ενζύμων. Η γλυκόζη εισέρχεται στα κύτταρα μέσω των μεταφορέων γλυκόζης (glucose transporters, GLUT) και φωσφορυλιώνεται από ισοένζυμα εξωκινασών (hexokinase) σε 6-γωσφορική γλυκόζη[8]. Μέσω του ενζύμου 1-φωσφογλυκομουτάση (phosphoglucomutase-1, PGM1), η 6-φωσφορική γλυκόζη ισομερίζεται σε 1-φωσφορική γλυκόζη[9,10]. Έπειτα το ένζυμο πυροφωσφατάση της 5'-διφωσφορικής ουριδίνης (UDP)-γλυκόζης (UDP-glucose pyrophosphorylase) καταλύει τη μετατροπή της 6-φωσφορικής γλυκόζης σε UDP-γλυκόζη, η οποία αποτελεί τον ενδιάμεσο δότη μορίων γλυκόζης για το σχηματισμό του γλυκογόνου[11].

Η γλυκογενίνη (glycogenin) ξεκινά τη σύνθεση του γλυκογόνου μέσω αυτοφωσφορυλίωσης, μεταφέροντας μόρια γλυκόζης από UDP-γλυκόζη και σχηματίζοντας μία μικρή αλυσίδα 10-20 καταλοίπων γλυκόζης[12]. Η επέκταση της αρχικής αυτής αλυσίδας καταλύεται από τη συνθάση του γλυκογόνου (glycogen synthase), η οποία μεταφέρει γλυκοσυλικά κατάλοιπα από UDP-γλυκόζη στην επεκτεινόμενη αλυσίδα, σχηματίζοντας α-1,4-γλυκοζυτικούς δεσμούς ανάμεσα στα κατάλοιπα γλυκόζης. Το ένζυμο διακλάδωσης του γλυκογόνου

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

(glycogen branching enzyme) εισάγει σημεία διακλάδωσης στο σωμάτιο του γλυκογόνου, σχηματίζοντας α-1,6-γλυκουτικούς δεσμούς[2].

1.1.3 Αποικοδόμηση του γλυκογόνου

Η αποικοδόμηση του γλυκογόνου λαμβάνει χώρα τόσο στο κυτταρόπλασμα όσο και στα λυσοσώματα. Στο κυτταρόπλασμα, ης σύνθεση του γλυκογόνου πραγματοποιείται μέσω της συντονισμένης δράσης δύο ενζύμων: της φωσφορυλάσης του γλυκογόνου (glycogen phosphorylase) και του ενζύμου αποδιακλάδωσης του γλυκογόνου (glycogen debranching enzyme). Η συνθάση του γλυκογόνου απελευθερώνει μόρια 1-φωσφορικής γλυκόζης από μία ευθεία αλυσίδα γλυκογόνου, διασπώντας τους α-1,4-γλυκοζυτικούς δεσμούς. Ωστόσο, η δράση το ενζύμου αυτού μπορεί να διακοπεί όταν βρεθεί τέσσερα μόρια γλυκόζης πριν από κάποιο σημείο διακλάδωσης. Στο σημείο αυτό, το ένζυμο αποδιακλάδωσης του γλυκογόνου λύει τα σημεία διακλάδωσης απελευθερώνοντας ελεύθερα μόρια γλυκόζης[2].

Στα λυσοσώματα, η αποικοδόμηση του γλυκογόνου (γνωστή και ως γλυκοφαγία [13]) πραγματοποιείται από τα ένζυμα όξινη α-γλυκοσιδάση (acid a-glucosidase, GAA) ή όξινη μαλτάση, που εντοπίζεται στα λυσοσώματα. Το γλυκογόνο μπορεί αν μεταφερθεί μέσα στα λυσοσώματα μέσω κυστιδίων αυτοφαγίας, τα οποία εγκλείουν ένα τμήμα του κυτταροπλάσματος, και συντήκονται με τα λυσοσώματα ούτως, ώστε να επεξεργαστεί το περιεχόμενο τους. Μέσα σε αυτά τα οργανίδια, το γλυκογόνο υδρολύεται από το ένζυμο όξινη α-γλυκοσιδάση [13].

1.1.4 Η γλυκογονόλυση μέσω αυτοφαγίας

1.1.4.1 Αυτοφαγία

Η αυτοφαγία είναι μία πολύ καλά συντηρημένη διαδικασία, κατά την οποία μη σωστά αναδιπλωμένες πρωτεΐνες του κυτταροπλάσματος, κατεστραμμένα οργανίδια, παθογόνοι μικροοργανισμοί ή ακόμα και πλεονάζον κυτοσόλιο μπορεί να αποικοδομηθεί στα λυσοσώματα [14]. Η διαδικασία της αυτοφαγίας λειτουργεί έτσι, ώστε να διατηρεί τα επίπεδα των πρωτεϊνών και των οργανιδίων σε χαμηλά επίπεδα. Επιπλέον, η αυτοφαγία σχετίζεται με ένα μεγάλο αριθμό άλλων λειτουργιών, όπως τη ρύθμιση του μεταβολισμού, την αύξηση και τη διαφοροποίηση και τη μεταφορά διαφόρων κυτταρικών αποθεμάτων ενέργειας, συμπεριλαμβανομένου υδατανθράκων, λιπιδίων και μετάλλων σε περιπτώσεις έλλειψης κυτταρικών συστατικών [15].

Μέχρι στιγμής, έχουν αναγνωριστεί τρεις κύριοι τύποι μη επιλεκτικής αυτοφαγίας στα θηλαστικά: η μακροαυτοφαγία, η μικροαυτοφαγία και η αυτοφαγία επαγόμενη από τις τσαπερονίνες. Η μακροαυτοφαγία χαρακτηρίζεται από το σχηματισμό των αυτοφαγοσωμάτων, για τη μεταφορά κυτταρικών συστατικών προς αποικοδόμηση τους στα λυσοσώματα. Τα κυτταρικά συστατικά επισημαίνονται από υποδοχείς αυτοφαγίας και εγκλείονται στα αυτοφαγοσώματα, κυστίδια διπλής μεμβράνης, μέσω της σύνδεση τους με μία πρωτεΐνη μέλος της οικογένειας πρωτεϊνών 8 που σχετίζονται με την αυτοφαγία [autophagy-related protein (Atg8) family protein], οι οποίες είναι αγκυροβολημένες στη μεμβράνη του αυτοφαγοσώματος. Έπειτα τα αυτοφαγοσώματα συντήκονται με τα λυσοσώματα, σχηματίζοντας τα αυτοφαγολυσοσώματα, ώστε να μεταφέρει σε αυτά τα κυτταρικά συστατικά που φέρει. Κατά την μικροαυτοφαγία, το ίδιο το λυσόσωμα εγκλείει

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

μικρά συστατικά του κυτταροπλάσματος μέσω της προς τα μέσα εγκόπλωσης της λυσοσωμικής μεμβράνης, με αποτέλεσμα της άμεσης αποικοδόμησης αυτών στον αυλό των λυσοσωμάτων. Κατά την αυτοφαγία επαγόμενης από τις τσαπερονίνες, πρωτεΐνες που φέρουν στην αμινική τους ακολουθία ένα πενταπεπτίδιο που σχετίζεται βιοχημικά με το μοτίβο KFERQ μπορούν να αναγνωριστούν από την κυτοσολική τσαπερονίνη Hsc70 και να μεταφερθούν στον λυσοσωμικό αυλό προς αποικοδόμηση τους [16].

1.1.4.2 Γλυκοφαγία

Η γλυκοφαγία είναι μία επιλεκτική διαδικασία αυτοφαγίας, κατά την οποία το κυτταρικό γλυκογόνο αποικοδομείται στα λυσοσώματα. Αν και ο επιλεκτικός μηχανισμός αυτός έχει ξεκινήσει να χαρακτηρίζεται πολύ πρόσφατα, η λυσοσωμική γλυκογονόλυση είναι γνωστή από τη δεκαετία του 1960 [17,18], συσχετίζοντας την με ασθένειες αποθηκών γλυκογόνου (glycogen storage diseases, GST). Η κληρονομική ανεπάρκεια του ενζύμου όξινη α-γλυκοσιδάση να αποικοδομήσει το λυσοσωμικό γλυκογόνο, γνωστή και ως ασθένεια Pompe, σχετίζεται με μια σημαντική συσσώρευση γλυκογόνου σε καρδιακούς και μυϊκούς ιστούς [8,19,20]. Έτσι, η αποικοδόμηση του γλυκογόνου μέσω της αυτοφαγίας δεν είναι απλά ένας εναλλακτικός τρόπος αποικοδόμησης του γλυκογόνου από τα κύτταρα, αλλά είναι μία καθοριστικής σημασίας διαδικασία για την διατήρηση της ακεραιότητας του κυττάρου και της ομοιόστασης του του γλυκογόνου. Ωστόσο, ο μηχανισμός της δεν είναι πλήρως γνωστός.

1.1.4.2.1 <u>Επισκόπηση της γλυκοφαγίας</u>

Ο καταβολισμός του γλυκογόνου μέσω της γλυκοφαγίας φαίνεται να συμβαίνει σε μεγάλο βαθμό στα νεογέννητα ζώα [13,21]. Κατά τη γέννηση, τα επίπεδα της γλυκόζης στο αίμα των νεογέννητων μειώνεται, λόγω της διακοπής παροχής θρεπτικών συστατικών μέσω του πλακούντα [22]. Ωστόσο, επειδή οι μηχανισμοί για παροχή γλυκόζης μέσω της γλυκονεογένεσης δεν είναι καλά εδραιωμένοι στα κύτταρα τους, το γλυκογόνο αποτελεί ένα αποθηκευτικό μέσο για την παροχή ενέργειας, παρέχοντας γλυκόζη στα κύτταρα. Κατά τη γλυκοφαγία, το γλυκογόνο στα λυσοσώματα αποικοδομείται από την όξινη α-γλυκοσιδάση (acid a-glucosidase, GAA), με αποτέλεσμα την απελευθέρωση α-γλυκόζης για παραγωγή ενέργειας στα κύτταρα. Η όξινα α-γλυκοσιδάση μπορεί να διασπάσει τόσο τους α-1,4-γλυκοζυτικούς δεσμούς, όσο και τους α-1,6-γλυκοζυτικούς δεσμούς, έχοντας, ωστόσο, μεγαλύτερη απόδοση διάσπασης των πρώτων [20,23,24]. Η υδρολυτική διάσπαση του γλυκογόνου μέσω των αυτοφαγικών κυστιδίων είναι μία σημαντική διεργασία για τα κύτταρα του ήπατος, της καρδιάς και των μυών.

Για την έναρξη της γλυκοφαγίας είναι απαραίτητος ο σχηματισμός της μεμβράνης του αυτοφαγοσώματος. Το αρχικό τμήμα της μεμβράνης, το φαγοφόρο (phagophore), φαίνεται να προέρχεται από θέσεις του ενδοπλασματικού δικτύου, όπως και από υποκυτταρικά διαμερίσματα που σχετίζονται με αυτό [25]. Κάποιες θέσεις από όπου προέρχεται αυτή η διπλή διπλοστοιβάδα είναι οι θέσεις εξόδου του ενδοπλασματικού δικτύου [26,27], η υπομονάδα του ενδοπλασματικού δικτύου, το ωμεγάσωμα (omegasome) [28-31], τα ενδιάμεσα διαμερίσματα του ενδοπλασματικού δικτύου-συσκευής Golgi [32] και οι μεμβράνες του ενδοπλασματικού δικτύου που σχετίζονται με τα μιτοχόνδρια (mitochondria-associated ER membranes, MAMs) [33]. Στις

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

θέσεις αυτές έχει εντοπιστεί μεγάλη συγκέντρωση γλυκογόνου στους καρδιακούς, ηπατικούς και μυϊκούς ιστούς.

Τα προς αποικοδόμηση γλυκογόνο πρέπει να σημανθεί από ειδικές πρωτεΐνες, τους υποδοχείς αυτοφαγίας. Οι πρωτεΐνες αυτές έχουν την ικανότητα να αλληλοεπιδρούν επιλεκτικά με συστατικά του κυτοσολίου, καθώς, επίσης, φέρουν περιοχές αλληλεπίδρασης με μέλη της οικογένειας πρωτεϊνών Atg8, μέσω του LIR/AIM μοτίβου (Atg8 family Interacting Motif). Η οικογένεια αυτή πρωτεϊνών περιλαμβάνει τις πρωτεΐνες LC3 (LC3a, LC3b, LC3c) και Gabarap (Gabarap, Gabarapl1, Gabarapl2) [34]. Οι πρωτεΐνες μοιάζουν με την ουβικουτίνη, είναι συνδεδεμένες με φωσφολιπίδια της επιφάνειας των σχηματισμένων αυτοφαγοσωμάτων, τυπικά με μόρια φωσφατιδυλεθανολαμίνης (phosphatidylethanolamine), και έχουν ρόλο στην μεταφορά της μεμβράνης και της αυτοφαγίας [35]. Με άλλα λόγια, η πρωτεΐνη-υποδοχέας σχηματίζει σύμπλοκο με το γλυκογόνο και ύστερα της αλληλεπίδρασής της με μέλη της οικογένειας πρωτεϊνών Atg8, συντονίζεται ο σχηματισμός των αυτοφαγοσωμάτων και επάγεται η διαδικασία της επιλεκτικής αυτοφαγίας του γλυκογόνου.

1.1.4.2.2 Ρύθμιση της γλυκοφαγίας

Η γλυκοφαγία επάγεται από το μονοπάτι της κυκλικής 5'-μονοφωσφωρικής αδενοσίνης (cyclic adenosine monophosphate, cAMP) /πρωτεϊνικής κινάσης A (Protein Kinase A, PKA). Το cAMP φαίνεται να αυξάνει το σχηματισμό των αυτοφαγοσωμάτων και την αποικοδόμηση του γλυκογόνου σε αυτά τα οργανίδια, ως απόκριση στην ανάγκη για παραγωγή γλυκόζης [36-38]. Δυο με τρεις ώρες μετά τη γέννα, τα νεογνά παθαίνουν υπογλυκαιμία. Η γλυκαγόνη, ύστερα από τη σύνδεσή της με τον β-αδρενεργικό υποδοχέα της, επάγει το σχηματισμό της 6-φωσφορικής γλυκόζης μέσω ενεργοποίησης του σηματοδοτικού μονοπατιού cAMP/PKA, το οποίο με τη σειρά του επάγει την ενεργοποίηση της φωσφορυλάσης του γλυκογόνου [39]. Ωστόσο, φαίνεται ότι το ίδιο αυτό μονοπάτι επάγει την αύξηση σχηματισμού των αυτοφαγωσωμάτων που περιέχουν μέσα τους γλυκογόνο και αυξημένη δράση της όξινης α-γλυκοσιδάσης.

Η γλυκοφαγία φαίνεται, επίσης, να ρυθμίζεται από το σηματοδοτικό μονοπάτι της κινάσης της 3-φωσφοϊοσιτόλης με το μεταγραφικό παράγοντα mTOR (PI3K-Akt/PKB-mTOR) [40]. Η σύνδεση της ινσουλίνης στον υποδοχέα της επιτρέπει την ενεργοποίηση της PI3K, η οποία με τη σειρά της ενεργοποιεί την Akt/PKB, επάγοντας το μεταγραφικό παράγοντα mTOR. Παράλληλα, η πρωτεϊνική φωσφατάση 2A (protein phosphatase 2A, PP2A), επιτρέπει τη σύνθεση της όξινης α-γλυκοσιδάσης [41-43]. Ο mTRO αναστέλλει τη δράση της PP2A, μειώνοντας τη έκφραση της όξινης α-γλυκοσιδάσης και, κατ' επέκταση, τα ρυθμό της γλυκοφαγίας στα κύτταρα.

1.2 Carbohydrate-Binding Modules

Πρωτεΐνες και ένζυμα που συνδέονται με υδατάνθρακες υπάρχουν σε πληθώρα σε όλα τα ζωντανά κύτταρα και οικογένειες οργανισμών και εμπλέκονται σε μυριάδες σημαντικές βιολογικές λειτουργίες, που περιλαμβάνουν τη σύνθεση, τη μεταφορά και την αποικοδόμηση διαφόρων μορφών υδατανθράκων [44,45]. Μικροβιακά εξωκυττάρια υδρολυτικά ένζυμα που καταλύουν την αποικοδόμηση κοκκιώδους μορφής αμύλου [46] ή φυτικών κυτταρικών τοιχωμάτων [47], στα οποία κυριαρχεί η κυτταρίνη, διαθέτουν μία ξεχωριστή αρθρωτή αρχιτεκτονική, ενώ άλλα φέρουν μη καταλυτικές βοηθητικές επικράτειες που αναφέρονται, με τον γενικό όρο, ως πρωτεϊνικές μονάδες που συνδέονται με υδατάνθρακες (carbohydrate-binding modules –

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

CBMs) [48]. Οι επικράτειες αυτές αποτελούν τόσο δομικά όσο και λειτουργικά ανεξάρτητες πρωτεϊνικές μονάδες και επιτρέπουν τη σύνδεση με ειδικές δομές ολιγο- ή πολυσακχαριτών και τη μεταφορά τους στο καταλυτικό κέντρο του ενζύμου [49].

Η πρώτη αναφορά τους έγινε το 1986, όταν στην κυτταρινάση cellobiohydrolase I (CBH I) του *Trichoderma reesei* βρέθηκε ότι μία επικράτεια της δεν είχε ικανότητα υδρόλυσης, αλλά σύνδεσης με κυτταρίνη, γνωστή έπειτα και ως cellulose-binding domains – CBDs [50,51]. Παράλληλα, παρόμοιες παρατηρήσεις έγιναν και για άλλα ένζυμα αποικοδόμησης πολυσακχαριτών, όπως για αμυλάσες που φέρουν επικράτειες σύνδεσης με το άμυλο (starch-binding domains – SBDs) [46,52]. Έτσι, το 1999, εισήχθη ο όρος των CBMs, οι οποίες είναι μία συνεχόμενη αμινοξική ακολουθία ενός ενζύμου ή μίας πρωτεΐνης που ενεργοποιείται από υδατάνθρακες (carbohydrate-active enzymes – CAZymes), αναδιπλώνεται στο χώρο ως ξεχωριστή επικράτεια και παρουσιάζει ικανότητα σύνδεσης με υδατάνθρακες [48].

Όλες οι πρωτεϊνικές αυτές μονάδες αποτελούν μέρος του συστήματος κατηγοριοποίησης που βασίζεται στην αλληλουχία όλων των ενζύμων που συνθέτουν, αποικοδομούν ή/ και τροποποιούν σάκχαρα – βάση δεδομένων CAZy (Carbohydrate-Active enzyme) [53-55]. Η βάση δεδομένων CAZy αποτελείται κυρίως από τις καταλυτικές επικράτειες των υδρολασών γλυκοσιδίων (glycoside hydrolases – GHs), των γλυκοσυλτρανσφερασών (glycosyl transferases – GTs), των λυασών πολυσακχαριδίων (polysaccharide lyases – PLs), των εστερασών υδατανθράκων (carbohydrate esterases – CEs) και των πρωτεϊνών στενής δραστικότητας (auxiliary activities – AAs). Βάσει της αλληλουχίας τους, της δομής και της λειτουργίας τους, οι CBMs ομαδοποιούνται σε 88 οικογένειες και έχουν καταγραφεί περισσότερα από 260000 μέλη (http://www.cazy.org/) [*]. Κάθε νέα πρωτεϊνική μονάδα που ανακαλύπτεται μπορεί να κατανεμηθεί σε μία οικογένεια CBM μόνο όταν έχει επιβεβαιωθεί η λειτουργία σύνδεσής της με τους υδατάνθρακες [49].

Παρά τη χαμηλή ομοιότητα των αλληλουχιών CBMs κατά τη στοίχισή τους, οι πρωτεϊνικές αυτές μονάδες φαίνεται να μοιράζονται κοινά χαρακτηριστικά στη δομή τους. Οι τρισδιάστατες δομές από 67 οικογένειες CMBs και 411 μέλη τους έχουν καθοριστεί μέσω μελετών κρυσταλλογραφίας ακτίνων X (X-ray crystallography) και φασματοσκοπίας πυρηνικού μαγνητικού συντονισμού [Nuclear Magnetic Resonance (NMR) spectrometry] (http://www.cazy.org/) [*]. Το 2004, ο Boraston et al. κατηγοριοποίησε 22 δομές CBM σε επτά δομικές οικογένειες βάσει της αναδίπλωσης των πρωτεϊνικών μονάδων στο χώρο: το μοτίβο β-σάντουιτς (β-sandwich) , το μοτίβο β-τριφύλλι (β-trefoil), το μοτίβο κόμβου κυστεΐνης (cysteine knot), το μοτίβο με μοναδική αναδίπλωση (unique fold), η αναδίπλωση με σύνδεση με ολιγονουκλεοτίδια/ ολιγοσακχαρίτες [oligonucleotide/ oligosaccharide-binding (OB) fold], η hevein αναδίπλωση, η αναδίπλωση παρόμοια με αυτή του τύπου hevein (hevein-like fold) [56]. Η δομή του β-sandwich είναι η κυρίαρχη δομή που συναντάται ανάμεσα στα μέλη των οικογενειών CBMs. Η δομή αυτή αποτελείται από δύο β-πτυχωτές επιφάνειες, με την κάθε μία να φέρει τρεις με έξι αντιπαράλληλες β-πτυχώσεις. Σχεδόν σε όλες τις δομές β-sandwich υπάρχει συνδεδεμένο ένα μεταλλικό ιόν, συνήθως ιόν ασβεστίου, το οποίο έχει δομικό ή/και λειτουργικό ρόλο, αποτελώντας γέφυρα στην αναγνώριση υποστρωμάτων από τις δομές CBMs. Πιο συγκεκριμένα, τα ιόντα ασβεστίου έχουν άμεση επίδραση, σχηματίζοντας πολικούς δεσμούς ανάμεσα με την πρωτεϊνική

[*] Οι αριθμοί των μελών και των οικογενειών CBM, αλλά και των δομών που έχουν καθοριστεί, ανανεώνονται τακτικά. Οι αριθμοί αυτοί βασίζονται στην τελευταία ενημέρωση της βάσης δεδομένων CAZy στις 18 Μαΐου 2021.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

μονάδα και το υπόστρωμα [57,58], είτε έμμεση, σταθεροποιώντας τις πλευρικές αλυσίδες των αρωματικών αμινοξέων στις θέσεις σύνδεσης των πρωτεϊνικών μονάδων [59,60]. Στην πλειοψηφία των CBMs με δομή β-sandwich, η θέση αναγνώρισης προσδετών εντοπίζεται στην ίδια επιφάνεια της δομής.

Βάσει της τοπολογίας της θέσης σύνδεσης των προσδετών, οι CBMs κατηγοριοποιούνται σε τρεις διαφορετικούς τύπους: τον τύπο A, τον τύπο B [ενδο-τύπος (endo-type) [61]] και τον τύπο C [εξω-τύπος (exo-type) [61]] [56]. Ο τύπος A έχει μία υδρόφοβη θέση σύνδεσης που αναγνωρίζει αδιάλυτους, κρυσταλλικούς πολυσακχαρίτες, όπως κυτταρίνης ή χιτίνης . Η θέση σύνδεσης των CBMs τύπου B μπορεί να εντοπιστεί σε μία ή/και δύο περιοχές της δομής τους [56]. Η ικανότητα πρόσδεσης των υδατανθράκων στις περιοχές αυτές καθορίζεται από τον βαθμό πολυμερισμού του (degree of polymerization – DP), καθώς παρουσιάζουν αυξημένη αγχιστεία με πολυσακχαρίτες έξι μονοσακχαριτών και αμελητέα αλληλεπίδραση με ολιγοσακχαρίτες [48,56]. Οι πρωτεϊνικές αυτές μονάδες μπορούν να αλληλοεπιδράσουν με διαλυτούς πολυσακχαρίτες παρά με κρυσταλλικούς. Ενώ οι CBMs τύπου B έχουν την ικανότητα να συνδέονται με το εσωτερικό τμήμα πληθώρας αλυσίδων γλυκανών (αναφορά ως ενδο-τύπος) [61], οι CBMs τύπου C έχουν την ικανότητα να συνδέονται με τα άκρα ολιγοσακχαριτών (μονό-, δι-, τρισακχαρίτες) και φέρουν παρόμοιες ιδιότητες με τις λεκτίνες [CBMs παρόμοιες με λεκτίνες (lectin-like CBMs) ή/και εξω-τύπος] [56].

Οι CBMs φαίνεται να έχουν τρεις βασικούς ρόλους στην λειτουργία των ενζύμων: τη δράση σε κοντινή απόσταση (proximity effect), την στόχευση του υποστρώματος (substrate targeting) και τη διατάραξη της δομής του υποστρώματος (substrate structure disruption) [48,56,62]. Οι δομές βρίσκονται σε στενή απόσταση από το υπόστρωμα, επιτρέποντας το ένζυμο να διατηρήσει στενή σχέση με τον προσδέτη του [63]. Αυτό έχει ως αποτέλεσμα την αποτελεσματική συγκέντρωση του ενζύμου στην επιφάνεια του υποστρώματος και την ενίσχυση της δράσης του ενζύμου, πχ της υδρόλυσης [63-66]. Επιπλέον, οι CBMs παρουσιάζουν επιλεκτική αγχιστεία προς διάφορες δομές διαλυτών και αδιάλυτων ή/και κρυσταλλικών πολυσακχαριτών [50,51,64,67,68]. Σημαντικό ρόλο έχουν η δομή των θέσεων σύνδεσης και ο προσανατολισμός των πλευρικών αλυσίδων ειδικά των αρωματικών καταλοίπων σε αυτές [56], αλλά και οι δεσμοί υδρογόνου που αναπτύσσονται από τα πολικά κατάλοιπα στις θέσεις αυτές [45]. Τέλος, η διάσπαση της δομής των πολυσακχαριτών απαιτεί πολυσθενείς αλληλεπιδράσεις. Οι CBMs έχουν την ικανότητα να ξεμπερδέψουν πολύπλοκες αλυσίδες πολυσακχαριτών στις επιφάνειες των υποστρωμάτων, επιτρέποντας τες να είναι αποτελεσματικότερα προσβάσιμες στο ένζυμο [48].

1.2.1 CBM20

Οι CBMs με αγχιστεία προς το άμυλο αναφέρονται ως SBDs (starch-binding domains) [48]. Για να ορισθεί μία CBMs ως SBD θα πρέπει να φέρει ορισμένα χαρακτηριστικά [49]. Πρώτα, θα πρέπει να είναι ένα συνεχές τμήμα μιας πολυπεπτιδικής αλυσίδας που σχηματίζει μία ξεχωριστή και λειτουργικά ανεξάρτητη μονάδα και να αποτελεί μέρος ενός ενζύμου ή, πιο σπάνια, να εντοπίζεται στα κύτταρα ως ξεχωριστή πρωτεΐνη. Επιπλέον, θα πρέπει να αποδεικνύεται μέσω βιοχημικών και μοριακών μελετών ότι συνδέεται με το άμυλο ή με κάποιο σχετικό ολιγο ή πολυσακχαρίτη α-γλυκανών, όπως το γλυκογόνο. Το άμυλο και το γλυκογόνο είναι παρόμοια χημικά μόρια, αλλά διαφέρουν σημαντικά στην τελική τους δομή, στις φυσικές τους ιδιότη-

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

τες και στην ευαισθησία της ενζυμικής τους αποικοδόμησης [48]. Όλα τα SBDs, για τα οποία έχει αποκαλυφθεί η τριδιάστατη δομή τους, παρουσιάζουν μία αναδίπλωση β-σάντουιτς με τοπολογία παρόμοια με αυτήν της ανοσοσφαιρίνης (Immunoglobulin-like fold) [69].

Οι πρώτες SBDs που ανακαλύφθηκαν τοποθετήθηκαν στην CBM20 οικογένεια, η οποία παραμένει η πιο καλά χαρακτηρισμένη οικογένεια SBD [48]. Η πρώτη αναγνωρίστηκε στις αρχές τις δεκαετίας του 1980, στο καρβοξυτελικό άκρο των γλυκοαμυλασών των *Aspergillus awamari* [70] και *Aspergillus niger* [71,72]. Έτσι, ενώ τα πρώτα CBM20s εντοπίστηκαν στο καρβοξυτελικό άκρο εξωκυττάριων αμυλασών που εκκρίνονται από μύκητες ή βακτήρια [72-74], ένας όλο και αυξανόμενος αριθμός CBM20s βρέθηκε να παρουσιάζει μία μεγάλη ποικιλία, όσον αφορά τη φυλογενετική τους προέλευση και τη λειτουργία των καταλυτικών μονάδων με τις οποίες συνδέονται στο ένζυμο [48]. Έτσι, οι CBM20s εντοπίζονται σε ένα μεγάλο εύρος εκκρινόμενων ή ενδοκυτταρικών, αμυλολυτικών και μη αμυλολυτικών ενζύμων των φυτών, των θηλαστικών, των βακτηρίων και των μυκήτων.

1.2.1.1 Κατηγοριοποίηση

Η οικογένεια CBM20 περιέχει περισσότερα από 2450 μέλη. Τα περισσότερα από αυτά ανήκουν σε βακτήρια (1987 μέλη) και λιγότερα σε ευκαρυώτες (477 μέλη), ενώ στα αρχαία έχουν αναγνωριστεί 14 μέλη (http://www.cazy.org/). Περισσότερα από 120 μέλη των CBM20s έχουν χαρακτηριστεί καλά και για 17 από αυτά έχει οριστεί η τριδιάστατη δομή τους. Τα ένζυμα, στα οποία οι CBM20s αποτελούν τμήμα της αλληλουχίας τους, είναι κυρίως α-αμυλάσες (GH13, οικογένεια 13 των υδρολασών γλυκοσιδίων) ή ένζυμα των οικογενειών GH70 και GH77 (οι τρεις οικογένειες GH13, GH70 και GH77 σχηματίζουν την τάξη GH-H [75]), αλλά μπορούν να ανήκουν στις β-αμυλάσες (GH14) και στις γλυκοαμυλάσες (GH15) [76]. Ενδιαφέρον παρουσιάζει το γεγονός ότι στα μέλη των τριών αυτών αμυλασών (GH13, GH14, GH15) τόσο στους προκαρυώτες όσο και στους μύκητες, οι CBM20s εντοπίζονται στο καρβοξυτελικό άκρο των ενζύμων και συνδέονται με την καταλυτική επικράτεια των ενζύμων είτε άμεσα [γλυκοσυλοτρανσφεράση (CGTase) του *Bacillus circulans str. 251*] είτε μέσω ενός πεπτιδικού συνδέτη πλούσιο σε Ser/Thr, τα οποία υφίστανται μεταμεταφραστικές τροποποιήσεις, όπως N- και O-γλυκοζυλίωση (γλυκοαμυλάση του *Aspergillus niger* και η γλυκοαμυλάση του *Hypocrea jecorina*) [77-85].

CBM20s έχουν αναγνωριστεί σε ένζυμα και πρωτεΐνες, των οποίων η καταλυτική τους επικράτεια δεν κατατάσσεται σε κάποια κατηγορία ενζύμων της βάσης δεδομένων CAZy. Τέτοια παραδείγματα είναι η φωσφογλυκάνη, δικινάση του νερού (phosphoglucan, water dikinase – GWD3) [86,87], η γλυκεροφωσφοδιεστερφωσφοδιεστεράση-5 (glycerophosphodiester phosphodiesterase-5 – GPDP5) [88], η λαφορίνη (laforin) [89] και η γενοθονίνη-1 [genethonin-1 – Starch-binding domain-containing protein 1 (STBD1)] [90,91]. Οι πρωτεϊνικές μονάδες CBM20 μπορεί να εντοπίζονται στο καρβοξυτελικό άκρο των ενζύμων, ωστόσο, στις πρωτεΐνες που δεν καταλύουν αντιδράσεις, οι CBM20s μπορεί να εντοπίζονται τόσο στο αμινοτελικό όσο και στο καρβοξυτελικό άκρο [92,93].

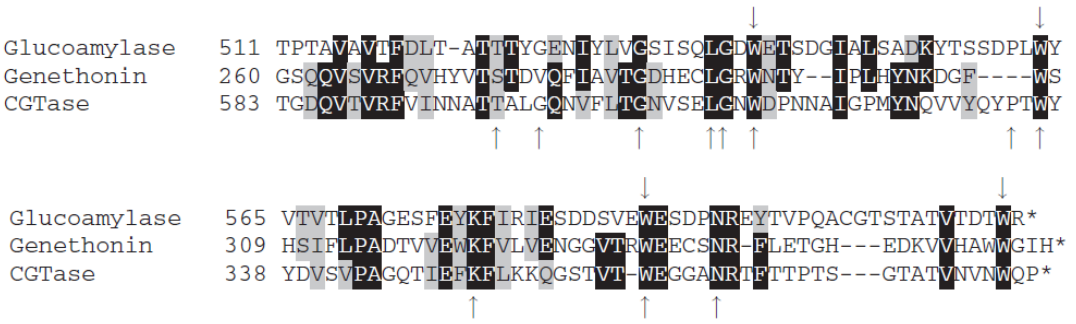
Τα μέλη της οικογένειας CBM20 εντοπίζονται, στην πλειοψηφία τους, ως μεμονωμένα αντίγραφα σε μία πρωτεΐνη και συνήθως δεν εντοπίζονται μαζί με άλλες πρωτεϊνικές μονάδες άλλων CBM οικογενειών στο ίδιο το πρωτεϊνικό μόριο [49]. Ωστόσο, έχουν παρατηρηθεί ορισμένες εξαιρέσεις, κατά τις οποίες οι CBM20

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

εντοπίζονται στο ίδιο πρωτεϊνικό μόριο με πρωτεϊνικές μονάδων των οικογενειών CBM25, CBM34 και CBM48 [94-98].

1.2.1.2 Συντηρημένα κατάλοιπα αμινοξέων

Οι CBM20s έχουν μήκος περίπου 90-130 αμινοξέα και αναλυτικές στοιχίσεις αλληλουχιών παρουσιάζουν ότι δεν υπάρχουν σταθερά, συνεχόμενα κατάλοιπα αμινοξέων στην οικογένεια αυτή. Παρόλα αυτά, η ικανότητα των CBM20s να συνδέονται με το άμυλο και με ανάλογα του φαίνεται να είναι αποτέλεσμα μεμονωμένων συντηρημένων αμινοξέων, που εντοπίζονται σε ένα μεγάλο ποσοστό των μελών αυτής της οικογένειας. Αρχικά, είχε προταθεί ότι βάσει της στοίχισης οκτώ διαφορετικών αλληλουχιών CBM20 από μύκητες και βακτήρια, υπάρχουν έντεκα συντηρημένα αμινοξέα. Με βάση την αρίθμηση των αμινοξέων της γλυκοαμυλάσης του *Aspergillus niger* αυτά τα συντηρημένα κατάλοιπα είναι: T525, G528, G535, L540, G541, W543, P561, W563, K578, W590 και N595 [99]. Ωστόσο, αναλύσεις μεγαλύτερων ομάδων CBM20s έδειξαν ότι ορισμένα από αυτά τα κατάλοιπα είναι πιο σημαντικά από άλλα, λόγω του μεγαλύτερου βαθμού συντήρησής τους ανάμεσα σε διαφορετικές ταξινομικές ομάδες και λειτουργίες [49,76,92,93,100]. Τα συντηρημένα κατάλοιπα τρυπτοφάνης πιθανώς να συμμετέχουν στην αλληλοεπίδραση με γλυκοσιδικά κατάλοιπα μέσω υδρόφοβων δεσμών υδρογόνου. Τα κατάλοιπα γλυκίνης και προλίνης, πιθανώς, είναι σημαντικά στη διαμόρφωση της πολυπεπτιδικής αλυσίδας, ενώ τα κατάλοιπα θρεονίνης, λυσίνης, ασπαραγίνης και λευκίνης μπορεί να έχουν δομικό ρόλο ή/και να αλληλοεπιδρούν με το άμυλο.



Εικόνα 3 Στοίχιση των αλληλουχιών SBD της γλυκοαμυλάσης του *Aspergillus niger*, της CGTase του *Bacillus sp. strain 1011* και της Genethonin (STBD1). Τα βέλη που βρίσκονται πάνω από τη στοίχηση δηλώνουν τη σημασία των καταλοίπων τρυπτοφάνης για τη δομή και τη λειτουργία της επικράτειας, ενώ τα βελή που βρίσκονται κάτω από τη στοίχιση δηλώνουν τα απαραίτητα κατάλοιπα για των δομών της επικράτειας σύμφωνα με τον Svensson.[98]

Τα έντεκα συντηρημένα κατάλοιπα αμινοξέων συνδέονται αρκετά με τις περιοχές σύνδεσης των υδατανθράκων. Η θέση σύνδεσης 1 περιέχει δύο συντηρημένα κατάλοιπα τρυπτοφάνης (W543 και W590, αρίθμηση βασισμένη στην γλυκοαμυλάση του *Aspergillus niger*) [101]. Το μοτίβο ^{540}LGxW^{543} είναι το πιο καλά διατηρημένο μοτίβο ανάμεσα στα μέλη της οικογένειας αυτής. Σε ορισμένες περιπτώσεις υπάρχει ένα από τα δύο συντηρημένα κατάλοιπα τρυπτοφάνης σε αυτήν την περιοχή, με μεγαλύτερη συχνότητα το πρώτο από το δεύτερο [92]. Η θέση σύνδεσης 2, φέρει μία τρυπτοφάνη στην γλυκοαμυλάση του *Aspergillus niger*, η οποία είναι συντηρημένη σε ποσοστό 96% ανάμεσα σε 103 αλληλουχίες CBM20s που έχουν αναλυθεί, και δύο κατάλοιπα τυροσίνης (Y527 και Y556, αρίθμηση βασισμένη στην γλυκοαμυλάση του *Aspergillus niger*), που δεν είναι καλά συντηρημένες και εμφανίζονται σε μικρότερο ποσοστό ανάμεσα σε αυτές τις 103 αλληλουχίες[92]. Στα έντεκα συντηρημένα κατάλοιπα προστίθεται και η φαινυλαλανίνη (F519, αρίθμηση

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

βασισμένη στην γλυκοαμυλάση του *Aspergillus niger*), που συντηρείται στο 87% των 103 αλληλουχιών, ενώ σε ορισμένες βακτηριακές β-αμυλάσες το κατάλοιπο αυτό αντικαθίσταται από λευκίνη, ισολευκίνη και βαλίνη [92]. Τέλος, το κατάλοιπο W615 της γλυκοαμυλάσης του *Aspergillus niger* έχει μάλλον δομικό ρόλο, σταθεροποιώντας την αρχιτεκτονική της πρωτεϊνικής μονάδας, και όχι άμεσα λειτουργικό στις θέσεις σύνδεσης [101].

1.2.1.3 Αρχιτεκτονική CBM20 και δομικά χαρακτηριστικά

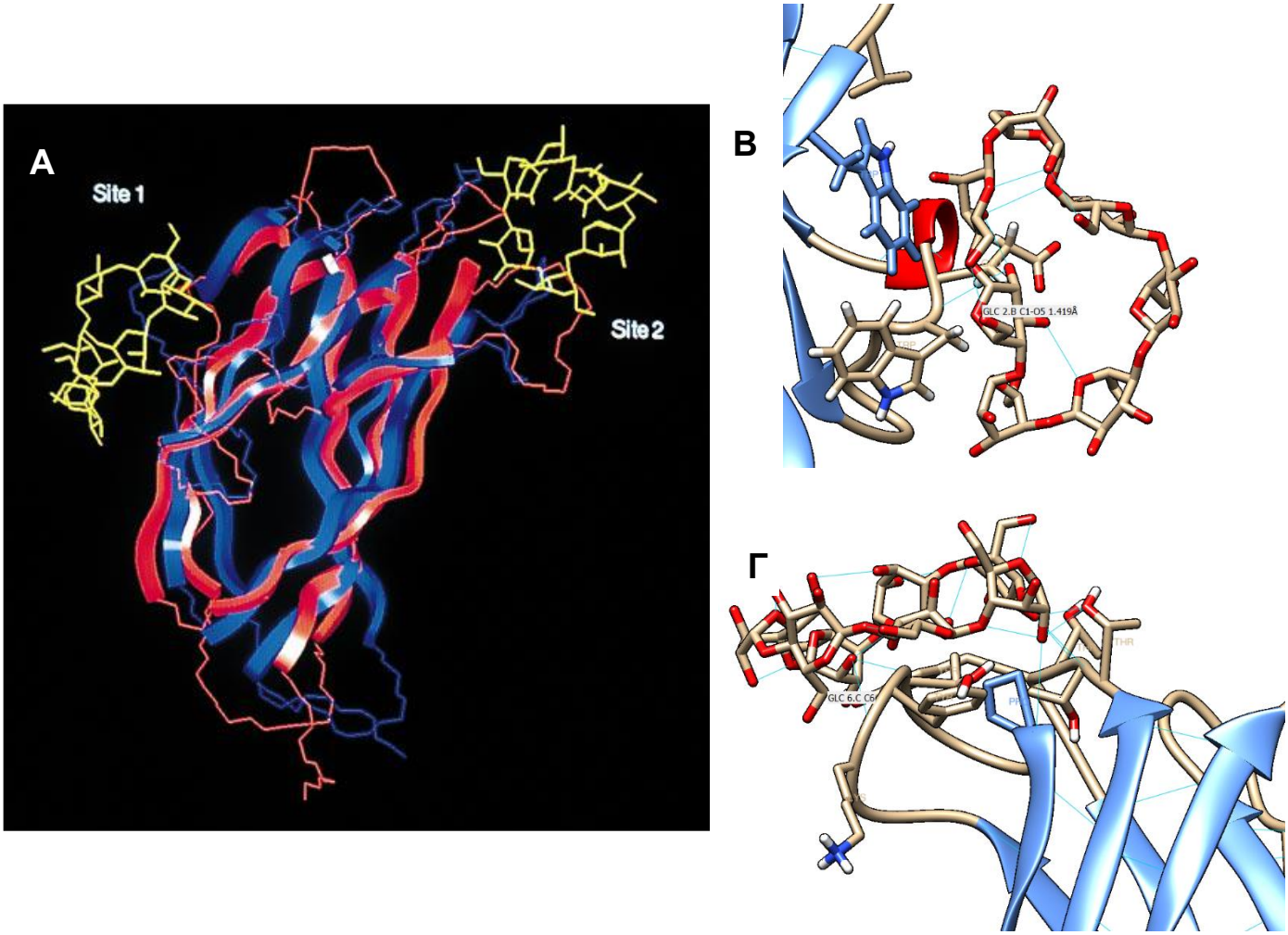
Έως και σήμερα, έχουν καθοριστεί οι τριδιάστατες δομές για 17 μέλη της οικογένειας CBM20. Κάποιες από αυτές τις δομές αναλύονται ως σύμπλοκα με διαφορετικά υποστρώματα, ενώ άλλα έχουν καθοριστεί χωρίς την παρουσία κάποιου υποστρώματος, όπως μίας α-γλυκάνης ή ενός αναστολέα (http://www.cazy.org/) [102]. Ορισμένες δομές έχουν αποσαφηνιστεί από απομονωμένα μόρια CBM20, ενώ άλλα ως μέρος ολόκληρου του ενζύμου στο οποίο ανήκουν. Όπως όλα τα SBDs, έτσι και οι CBM20s υιοθετούν μία αναδίπλωση β-σάντουιτς με τοπολογία παρόμοια με αυτήν της ανοσοσφαιρίνης (Immunoglobulin-like fold) [69], η οποία είναι χαρακτηριστική για πρωτεϊνικά μόρια που συνδέονται με υδατάνθρακες [56,103]. Από ανάλυση δομής διαφορετικών μελών της οικογένειας αυτής σε σύμπλοκο με υδατάνθρακες, οι CBM20s φέρουν μία ή δύο περιοχές σύνδεσης [46,80,82,104]. Οι δυο περιοχές σύνδεσης διαφέρουν μεταξύ τους τόσο στη δομή όσο και στη λειτουργία τους, αν και έχει προταθεί η σημασία και των δυο για την ορθή λειτουργία της πρωτεϊνικής μονάδας.

Οι CBM20s ανήκουν, βάσει της τοπολογίας τους, στον τύπο Β, σύμφωνα με τον οποίο οι υδατάνθρακες-υποστρώματα προσαρμόζονται σε μία ή σε δύο περιοχές σύνδεσης [92]. Η πρώτη περιοχή, που ονομάζεται θέση μεταβλητού βρόχου (variable loop site – VLS), εντοπίζεται στο τέλος της δομής και σε αυτήν μπορούν να εντοπιστούν πολλαπλοί βρόχοι. Η δεύτερη περιοχή, που ονομάζεται θέση κοίλης επιφάνειας (concave face site – CFS), εντοπίζεται στην κοίλη πλευρά της δομής. Στις περισσότερες περιπτώσεις, η θέση σύνδεσης εντοπίζεται στην περιοχή CFS, αλλά σε ορισμένα μέλη της οικογένειας CBM20 μπορεί να εντοπιστεί και η περιοχή VLS [45,61].

Η πιο καλά χαρακτηρισμένη CBM20 είναι αυτή της γλυκοαμυλάσης του *Aspergillus niger*, η οποία χρησιμοποιείται ως κύριος εκπρόσωπος των CBM20s. Η δομή της καθορίστηκε με NMR τόσο σε ελεύθερη όσο και σε δεσμευμένη κατάσταση παρουσία β-κυκλοδεξτρίνης και παρουσιάζει μία καλά καθορισμένη αναδίπλωση β-σάντουιτς με οκτώ β-πτυχώσεις χωρισμένες σε δύο β-πτυχωτές επιφάνειες (PDB: 1acz)[105]. Η μία β-πτυχωτή επιφάνεια έχει πέντε αντιπαράλληλες β-πτυχώσεις, ενώ η δεύτερη αποτελείται από ένα παράλληλο και ένα αντιπαράλληλο ζεύγος β-πτυχώσεων. Η αναδίπλωση αυτή σχηματίζει ένα παραμορφωμένο β-βαρέλι με ανοιχτές πλευρές που φέρει έξι βρόχους διαφορετικών μεγεθών, τέσσερις από τους οποίους είναι καλά καθορισμένοι. Η δομή της γλυκοαμυλάσης του *Aspergillus niger* έχει συνδεδεμένη την β-κυκλοδεξτρίνη, που μιμείται το άμυλο, σε δύο περιοχές σύνδεσης, υποδηλώνοντας τη δισθενή φύση αυτής της CBM20 [93]. Το αμινοτελικό και το καρβοξυτελικό άκρο της βρίσκονται στα αντίθετα άκρα του μεγαλύτερου άξονα του μορίου [52]. Ένας δισουλφιδικός δεσμός (C509-C604), ανάμεσα στην κυστεΐνη που εντοπίζεται στο αμινοτελικό άκρο και στο βρόχο που συνδέει τις β-πτυχώσεις 7 και 8, συμβάλλει στη δομική σταθερότητα της πρωτεϊνικής μονάδας [46,48,52].

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Η περιοχή σύνδεσης CFS αποτελείται από τα κατάλοιπα W543, K578, W590, E591 και N595, με τους ινδολικούς δακτυλίους των W543 και W590 να σχηματίζουν το κεντρικό τμήμα του επιπέδου σύνδεσης υδατανθράκων [46,48]. Τα κατάλοιπα αυτά σχηματίζουν μία συμπαγή, άκαμπτη, υδροφοβική επιφάνεια, με την απόσταση ανάμεσα στους δακτυλίους ινδολών να είναι κατάλληλη για την σύνδεσης μορίων γλυκόζης που συνδέονται με α-1,4-δεσμούς [46]. Αυτή η ρηχή και εκτεθειμένη στο διάλυμα περιοχή σύνδεσης υφίσταται μία μικρή αλλαγή της διαμόρφωσής της ύστερα από την πρόσδεσή της με το υπόστρωμα. Η περιοχή σύνδεσης VLS, που καθορίζεται από τα κατάλοιπα T526, Y527, G528, E529, N530, D554, Y526 και W563, είναι πιο εκτενής και υφίσταται μεγάλη αλλαγή στη διαμόρφωσή της ύστερα από την αλληλεπίδραση της με την β-κυκλοδεξτρίνη, γεγονός που υποδηλώνει ότι αυτή η περιοχή έχει μεγαλύτερη δομική ευελιξία από την περιοχή σύνδεσης CFS. Η κύρια διαφορά παρατηρείται σε έναν βρόχο κοντά στο καρβοξυτελικό άκρο, που είναι αρκετά ευέλικτος. Στο σύμπλοκο της CBM20 με την β-κυκλοδεξτρίνη, το κατάλοιπο Y556 πλησιάζει τα κατάλοιπα D554 και L555, επάγοντας μία σημαντική μεταβολή της θέσης D560, με αποτέλεσμα την μετατόπιση του C_a ατόμου του τελευταίου καταλοίπου κατά 13Å [46,48]. Ένα τρίτο καλά συντηρημένο κατάλοιπο τρυπτοφάνης, W563, θεωρείται ως ένα θαμμένο κατάλοιπο με αμελητέα προσβασιμότητα στην



Εικόνα 4 Α. Παρουσίαση της δομής NMR της επικράτειας SBD του ενζύμου γλυκοαμυλάση του *Aspergillus niger*. Στην εικόνα σημαίνεται με διαφορετικό χρώμα η διαμόρφωση της επικράτειας χωρίς (κόκκινο, PDB: 1ac0) και με (μπλε, PDB: 1acz) των προσδετών της, β-κυκλοδεξτρίνη, στις δύο διαφορετικές θέσεις σύνδεσης. Β. Παρουσιάζονται οι δεσμοί υδρογόνου (γαλάζιες γραμμές) που αναπτύσσονται ανάμεσα στην περιοχή σύνδεσης I και τον πρσδέτη β-κυκλοδεξτρίνη. Γ. Παρουσιάζονται οι δεσμοί υδρογόνου (γαλάζιες γραμμές) που αναπτύσσονται ανάμεσα στην περιοχή σύνδεσης II και το προσδέτη β-κυκλοδεξτρίνη [112]. Οι δομές παρουσιάζονται με τη βοήθεια του προγράμματος UCSF Chimera.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

επιφάνεια του μορίου [46]. Αν και είναι ανίκανο το κατάλοιπο αυτό να αναπτύξει δεσμούς με την β-κυκλοδεξτρίνη, είναι σε στενή επαφή με πολλά άλλα αμινοξέα γύρω από τη θέση σύνδεσης VLS, έχοντας σημαντικό ρόλο στη διαμόρφωση του βρόχου. Η δομική αυτή ευελιξία της θέσης σύνδεσης VLS φαίνεται να αποτελεί μία γενική ιδιότητα των CBM20s, καθώς σημαντικές αλλαγές στη διαμόρφωσή τους παρατηρήθηκαν και στο βρόχο της περιοχής 460-465 στη β-αμυλάση του *Bacillus cereus* και στο βρόχο της περιοχής 627-636 στην γλυκοσυλοτρανσφεράση (CGTase) του *Bacillus circulans str. 251*, όταν αυτές βρίσκονται σε σύμπλοκο με τη μαλτόζη [79,80].

Στις CBM20s, οι αλληλεπιδράσεις ανάμεσα στην πρωτεϊνική μονάδα και τον προσδέτη κυριαρχούνται από αρωματικά αμινοξέα μέσω υδρόφοβων και CH-π δεσμών και δεσμών van der Waals [45,48]. Οι δεσμοί υδρογόνου που σχηματίζονται, κυρίως, από πολικά κατάλοιπα καθορίζουν την επιλογή και τη συγγένεια των CBMs για τα υποστρώματα τους [45].

1.3 Starch Binding Domain-containing Protein 1 (STBD1)

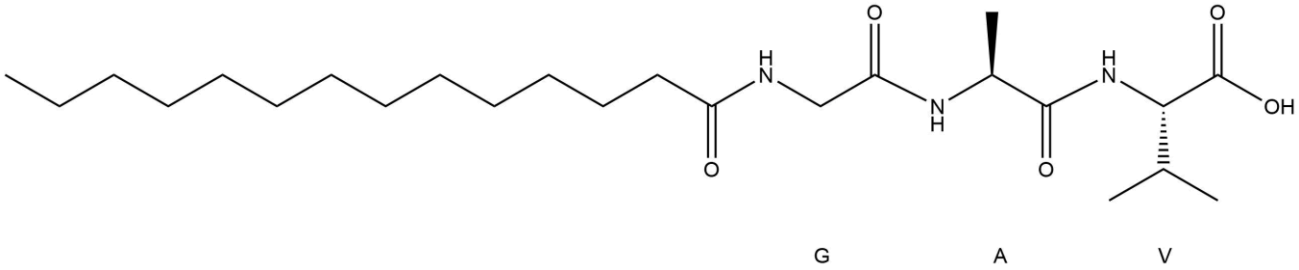
Η ανθρώπινη Starch Binding Domain-containing Protein 1 (STBD1) ή Genethonin-1, που αναφέρθηκε πρώτη φορά στη βιβλιογραφία το 1998, εκφράζεται από το γονίδιο GENX-3414 και κατηγοριοποιήθηκε ως μία πρωτεΐνη των εσωτερικών μεμβρανών [90]. Το γονίδιο αυτό μεταγράφεται σε ένα μόνο μετάγραφο. Το προ-mRNA φέρει δύο εξόνια, τα οποία χωρίζονται από ένα ιντρόνιο, με την κωδική περιοχή του να αποτελεί 1074 νουκλεοτίδια. Το mRNA *stbd1* εντοπίζεται, κυρίως, στα κύτταρα του σκελετικού μυός, του ήπατος, της καρδιάς και του πλακούντα. Η ανθρώπινη STBD1 απαρτίζεται από 358 αμινοξέα, με το θεωρητικό ισοηλεκτρικό της σημείο (pI) να είναι 5.46 και το πειραματικό του μοριακό βάρος να απαντάται σε δύο διαφορετικά μεγέθη,41 και 43 kDa [90]. Η έκφραση της STBD1 είναι υψηλότερη στο ήπαρ και στους σκελετικούς μύες, και λιγότερο στον σπλήνα, την καρδιά, το λιπώδη ιστό και τον πνευμονικό ιστό.

Η SBD1 χαρακτηρίστηκε ως πρωτεΐνη σύνδεσης με το γλυκογόνο μέσω ενός πρωτεωμικού ελέγχου σάρωσης απομονωμένου γλυκογόνου από ηπατικούς ιστούς ποντικιών και αρουραίων [106]. Η STBD1 φέρει μία υδρόφοβη περιοχή στο αμινοτελικό της άκρο (1-24 αα) και στο καρβοξυτελικό της άκρο μία επικράτεια σύνδεσης με υδρογονάνθρακες (Carbohydrate binding domain 20, CBM20) (258-357 αα) [90,91]. Επιπλέον, περιέχει μία περιοχή αλληλεπίδρασης με πρωτεΐνες της οικογένειας Atg8. Γενικά, η πεπτιδική ακολουθία της περιοχής αλληλεπίδρασης με πρωτεΐνες της οικογένειας Atg8 περιλαμβάνουν το μοτίβο [W/F/Y]-X-X-[I/L/V], γνωστό και ως AIM (Atg8 protein Interaction Motif) [107]. Τέτοια λειτουργικά μοτίβα εντοπίζονται συνήθως σε εγγενώς ασταθείς πρωτεΐνες, οι οποίες σχηματίζουν τεταρτοταγείς δομές ύστερα από την αλληλεπίδρασή τους με τις πρωτεΐνες Atg8 [108]. Στην περίπτωση της STBD1, η περιοχή αυτή είναι η ^{203}HEEWEV^{208}, η οποία φαίνεται να αλληλοεπιδρά με την GABARAPL1. Παράλληλα, η φωσφορυλίωση στο αμινοτελικό άκρο του AIM μοτίβου επιτρέπει τη σταθεροποίηση της αλληλεπίδρασης με τις Atg8 πρωτεΐνες. Ανηφορικά του μοτίβου στην STBD1 φαίνεται να φωσφορυλιώνεται το κατάλοιπο σερίνης S175 από την πρωτεϊνική κινάση που ενεργοποιείται από την 5'-μονοφωσφορική αδενοσίνη (AMP-activated protein kinase, AMPK), ενισχύοντας την παραπάνω υπόθεση αλληλεπίδρασης [109].

Δεδομένα από προγράμματα βιοπληροφορικής, σχετίζουν την υδροφοβικότητα του αμινοτελικού άκρου της πρωτεΐνης (*Kyle, Doolittle)* με τον υποκυτταρικό της εντοπισμό σε περιπυρηνικές δομές, και συγκεκρι-

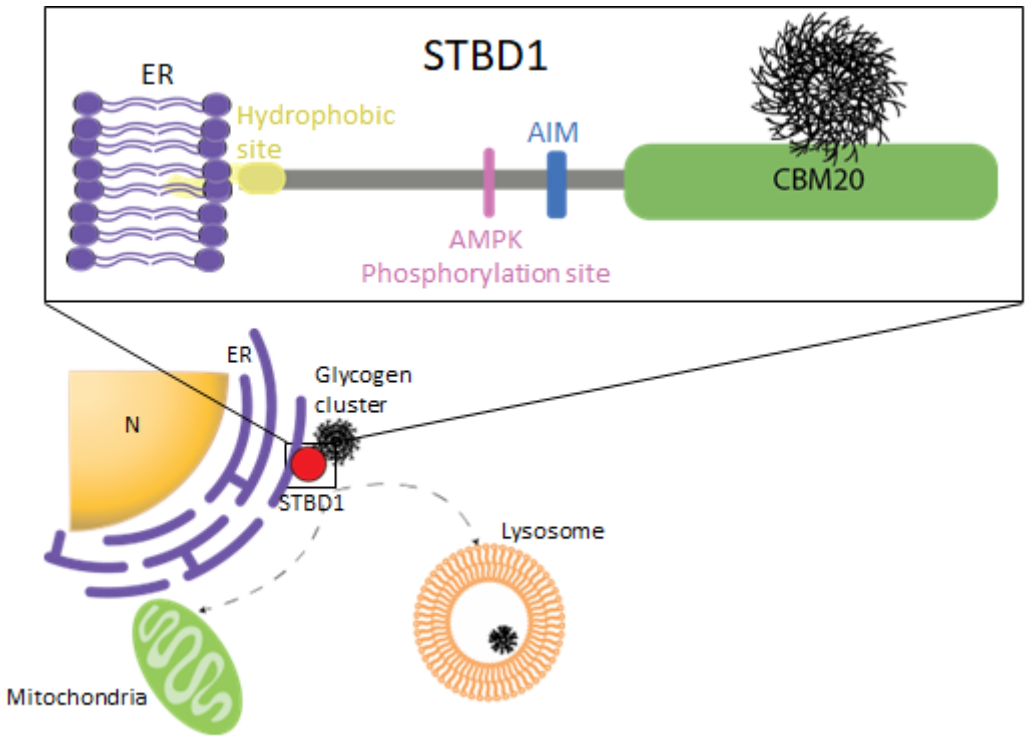
Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

μένα στις μεμβράνες του ενδοπλασματικού δικτύου, με την υδρόφοβη περιοχή να αποτελεί την διαμεμβρανική περιοχή της [110]. Ωστόσο, η STBD1 υφίσταται μία μετα-μεταφραστική τροποποίηση, τη N-μυριστοϋλίωση, η οποία της επιτρέπει να συνδέεται στο αδρό ενδοπλασματικό δίκτυο μέσω ενός λιπαρού οξέος, του μυριστοϊκού οξέος (ένα κορεσμένο λιπαρό οξύ 14 ατόμων άνθρακα). Το μυριστοϊκό οξύ συνδέεται στο πρώτο κατάλοιπο της γλυκίνης (G2) στο αμινοτελικό άκρο της STBD1 ύστερα από τη δράση του ενζύμου N-μυριστοϋλοτρανσφεράση (N-myristoyltransferase, NMT), ύστερα από την απομάκρυνση του αρχικού καταλοίπου μεθειονίνης [111].



Εικόνα 5 Το μυριστοϊκό οξύ συνδεδεμένο πάνω στην STDB1, ύστερα από τη σύνδεσηη της στο κατάλοιπο G2

Ενδοκυτταρικά, η STBD1 απαντάται σε δύο διαφορετικές μορφές, τη μυριστοϋλιωμένη και τη μη-μυριστοϋλιωμένη μορφή, οι οποίες εντοπίζονται με μεγάλη ειδικότητα σε διαφορετικά υποκυτταρικά διαμερίσματα, το αδρό ενδοπλασματικό δίκτυο (bulk endoplasmic reticulum) και τις περιοχές επαφής του ενδοπλασματικού δικτύου με τα μιτοχόνδρια (mitochondria associated membranes, MAMs), αντίστοιχα [111,112]. Η N-μυριστοϋλίωση επιτρέπει στη αμινοτελική περιοχή της STBD1 να «κλειδώσει», έχοντας ως αποτέλεσμα το σχηματισμό των οργανωμένων δομών του ενδοπλασματικού δικτύου, γνωστών ως δομές OSER (Organized Smooth Endoplasmic Reticulum). Αντίθετα, η έλλειψη της μετα-μεταφραστικής τροποποίησης



Εικόνα 6 Τα βασικά δομικά στοιχεία της STBD1. Η STBD1 συνδέεται στο ΕΔ μέσω μίας N-μυριστοϋλιωμένης ομάδας στο αμινοτελικό της άκρο. Φέρει το μοτίβο AIM και μπροστά από αυτό μία θέση φωσφορυλίωσης από την AMPK. Το καρβοξυτελικό της άκρο αποτελεί μία επικράτεια CBM20, η οποία έχει την ικανότητα σύνδεσης με το γλυκογόνο και με άλλες πρωτεΐνες του μεταβολισμού του γλυκογόνου.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

επιτρέπει τη διαμερισματοποίηση της πρωτεΐνης στις MAMs. Ο υποκυτταρικός αυτός εντοπισμός της πρωτεΐνης φαίνεται να σχετίζεται με τη μοναδική λιπιδική σύσταση των μεμβρανών αυτών, οι οποίες είναι πλούσιες σε χοληστερόλη και σφιγγολιπίδια. Στην τελευταία περίπτωση, το γλυκογόνο αποτελεί κύριο καθοριστή του υποκυτταρικού εντοπισμού του γλυκογόνου, καθόσον η μη-μυριστοϋλιωμένη μορφή της STBD1 δεν μπορεί να εντοπιστεί στις MAMs απουσία γλυκογόνου.

Η STBD1 δε φαίνεται να φέρει κάποια δομικά καταλυτική περιοχή, ωστόσο στο καρβοξυτελικό της άκρο περιέχεται μία επικράτεια σύνδεσης με υδατάνθρακες, την Carbohydrate Binding Domain 20 (CBM20) [91]. *In vitro* μελέτες έχουν δείξει ότι η STBD1 συνδέεται με πολυσακχαρίτες, όπως η αμυλοπεκτίνη και το γλυκογόνο, όπως επίσης αλληλοεπιδρά καλύτερα με λιγότερο διακλαδισμένους ή/και περισσότερο φωσφορυλιωμένους πολυσακχαρίτες [110]. Ενδοκυτταρικά, η STBD1 συνδέεται με σωμάτια γλυκογόνου σε περιπυρηνικές δομές, αλλά και σε δομές γύρω από τα μιτοχόνδρια. Η επικράτεια αυτή, πέρα από τη σύνδεση με το γλυκογόνο *in vivo*, έχει σημαντικό ρόλο στην αλληλοεπίδραση και με άλλες πρωτεΐνες *in vitro*, όπως η συνθάση του γλυκογόνου (Glycogen Synthase, GS) και το ένζυμο αποδιακλάδωσης του γλυκογόνου (Glycogen Debranching Enzyme, GDE), με το W293 κατάλοιπο να έχει ιδιαίτερη σημασία και στις δύο παραπάνω αλληλοεπιδράσεις [111-113].

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

2 ΣΚΟΠΟΣ

Η Starch binding domain-containing protein 1 (STBD1) είναι μία πρωτεΐνη σύνδεσης με υδατάνθρακες που εκφράζεται κυρίως στο ήπαρ, την καρδιά και τους μύες, και σχετίζεται με το μεταβολισμό και την κυτταρική μεταφορά του γλυκογόνου [90,110]. Λόγω της αλληλεπίδρασης της με την GABARAPL1 μέσω του μοτίβου Atg8-family interacting motif (AIM), προτάθηκε ως ένα υποδοοχέας της επιλεκτικής διαδικασίας της αυτοφαγίας του γλυκογόνου, γνωστής και ως γλυκοφαγία, επάγοντας τη μεταφορά του γλυκογόνου προς αποικοδόμηση στα λυσοσώματα στα κύτταρα του ήπατος [107,114]. Παρόλα αυτά, μέχρι και σήμερα δεν έχει καθοριστεί η δομή της STBD1.

Σε αυτήν την εργασία, με τη βοήθεια εργαλείων βιοπληροφορικής, γίνεται μία προσπάθεια ανάλυσής της. Επίσης, παρουσιάζεται η ετερόλογη έκφραση της επικράτειας CBM20 (258-358) της STBD1, κατασκευάζοντας, αρχικά, τον πλασμιδιακό φορέα pGEX-6P-1_CBM20 και δοκιμάζοντας, έπειτα, τις συνθήκες έκφρασής της.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

3 ΥΛΙΚΑ

3.1 Όργανα

	Εταιρεία
Φασματοφωτόμετρο ορατού/υπεριώδους UV 1600PC	VWR
Φασματοφωτόμετρο ορατού/υπεριώδους UV 1600PC	PerkinElmer
Συσκευή μέτρησης pH: 827 pH Lab	Metrohm
Φυγόκεντρος Eppendorf 5417R	Eppendorf
Φυγόκεντρος 50 mL, 500 mL	Hitachi
Συσκευή παραγωγής υπερήχων	Witeg
Heat block	Nippon Genetics
Αναλυτικός ζυγός	KERN
Παρασκευαστικός ζυγός	KERN
Ανακινούμενος ψυχόμενος επωαστήρας	LabTech
Συσκευή κάθετης ηλεκτροφόρησης: Mini-PROTEAN tetra cell	Bio-Rad
Σύστημα απεικόνισης πολυακρυλαμίδης: ChemiDox XR	Bio-Rad
Σύστημα οριζόντιας ηλεκτροφόρησης	Bio-Rad
Ξηρός κλίβανος	WTB
Κυκλομείκτες	Vortex

3.2 Βακτηριακές κυτταρικές σειρές

Βακτηριακό στέλεχος (DE3)	***Εταιρεία***
BL21Gold	Aligent technologies
Lemo21	Novagen
C41	Novagen
XL1-Blue	Novagen

3.3 Πλασμιδιακοί φορείς

Πλασμίδιο
pETM-11
pGEX-6P-1

3.4 Αντιδραστήρια

Αντιδραστήριο	***Εταιρεία***
Acetic Acid (CH_3COOH)	Sds
Acrylamide	Serva
Agarose	Serva

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Ammonium Persulfate	Serva
Ampicillin	Serva
BES	Sigma
Bis acrylamide	Serva
Bovine Serum Albumin (BSA)	New England Biolabs
Bradford protein assay 5x	Biorad
Bromophenole Blue	Research Organics
Calcium Chloride Dehydrate ($CaCl_2$)	Merck
Coomasie Brilliant Blue R-250	Fluka
Dimethyl sulfoxide (DMSO)	Panreac
Ethanol Absolute	Sigma
Ethylenediaminetetraacetic acid (EDTA)	Panreac
Glycerol	Applichem
Hepes	Serva
Hydrochloric acid 37% (HCl)	Applichem
Kanamycin	Serva
LB Agar	Sigma
LB Broth	Scharleu
MES	Apollo Scientific
Methanol	Scharleu
Phenylmethylsulfonyl fluoride (PMSF)	Applichem
Protease inhibitors	Roch
Sodium chloride	Scharleu
Sodium dodecyl sulfate (SDS)	Serva
Sodium hydroxide (NaOH)	Merck
TEMED	Sigma
Tris Ultra Pure	Invitrogen
Triton X-100	Applichem
a-D-Glucose	Sigma
β-Cyclodextrin (β-CD)	Serva
β-Mercaptoethanol (β-Merc)	Sigma

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

4 ΜΕΘΟΔΟΙ

4.1 Αρχή μεθόδων

4.1.1 Συστήματα Βιοπληροφορικής

4.1.1.1 ProtParam

Το ProtParam υπολογίζει διάφορες φυσικοχημικές ιδιότητες που μπορούν να συναχθούν από μια πρωτεϊνική ακολουθία. Δεν απαιτούνται πρόσθετες πληροφορίες σχετικά με την υπό εξέταση πρωτεΐνη. Δεν είναι δυνατό να προσδιοριστεί κάποια μετα-μεταφραστική τροποποίηση για την πρωτεΐνη-στόχο, το σχηματισμό διμερών ή πολυμερών. Οι παράμετροι που υπολογίζονται από το ProtParam περιλαμβάνουν το μοριακό βάρος (Mw), το θεωρητικό ισοηλεκτρικό σημείο (pI), τη σύνθεση των αμινοξέων, την ατομική σύνθεση, τον συντελεστή εξάλειψης, τον εκτιμώμενο χρόνο ημιζωής, τον δείκτη αστάθειας, τον αλειφατικό δείκτη και τον γενικό μέσο όρο της υδροπάθειας (Grand Average of Hydropathy, GRAVY) [115].

pI και Μοριακό βάρος

Αυτό το εργαλείο υπολογίζει το εκτιμώμενο pI και Mw μιας πρωτεϊνικής ακολουθίας. Το pI της πρωτεΐνης υπολογίζεται χρησιμοποιώντας τιμές pK των αμινοξέων, οι οποίες ορίστηκαν με εξέταση της μετατόπισης του πολυπεπτιδίου μεταξύ pH 4,5 έως 7,3 σε ακινητοποιημένο περιβάλλον πηκτής βαθμίδωσης pH με 9,2M και 9,8M ουρία στους 15°C ή 25°C. Η πρόβλεψη της πρωτεΐνης pI για πρωτεΐνες υψηλής βασικότητας δεν έχει ακόμη μελετηθεί και είναι πιθανό οι τρέχουσες προβλέψεις υπολογισμού pI/Mw να μην είναι επαρκείς για αυτόν τον σκοπό.

Το μοριακό βάρος της πρωτεΐνης υπολογίζεται με την προσθήκη των μέσων ισοτοπικών μαζών αμινοξέων στην πρωτεΐνη και της μέσης ισοτοπικής μάζας ενός μορίου νερού. Οι τιμές μοριακού βάρους δίνονται σε Dalton (Da). (Παράρτημα ΙΙ)

Συντελεστές απόσβεσης

Ο συντελεστής απόσβεσης δείχνει πόσο φως απορροφά μια πρωτεΐνη σε ένα συγκεκριμένο μήκος κύματος [116]. Έχει αποδειχθεί ότι είναι δυνατός ο υπολογισμός του συντελεστή γραμμομοριακής απόσβεσης μιας πρωτεΐνης από τη γνώση της σύστασης αμινοξέων της. Από τον μοριακό συντελεστή απόσβεσης της τυροσίνης, της τρυπτοφάνης και της κυστίνης, δηλαδή το αμινοξύ που σχηματίζεται όταν ένα ζεύγος μορίων κυστεΐνης ενώνονται με δισουλφιδικό δεσμό. (η κυστεΐνη δεν απορροφά σε μήκη κύματος >260 nm, ενώ η κυστίνη απορροφά) σε ένα δεδομένο μήκος κύματος, ο συντελεστής εξάλειψης της φυσικής πρωτεΐνης στο νερό μπορεί να υπολογιστεί χρησιμοποιώντας την ακόλουθη εξίσωση:

$$E_{PROT} = Numb_Y Ext_Y + Numb_W Ext_W + Numb_{Cystine} Ext_{Cystine},$$

όπου (για πρωτεΐνες σε νερό μετρημένες στα 280 nm): Ext_Y= 1490, Ext_W = 5500, $Ext_{Cystine}$= 125;

Η απορρόφηση (οπτική πυκνότητα) μπορεί να υπολογιστεί χρησιμοποιώντας τον ακόλουθο τύπο:

$$Absorb_{PROT} = {E_{PROT}}/{Mw}$$

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Αλειφατικός δείκτης

Ο αλειφατικός δείκτης μιας πρωτεΐνης ορίζεται ως ο σχετικός όγκος που καταλαμβάνεται από αλειφατικές πλευρικές αλυσίδες (αλανίνη, βαλίνη, ισολευκίνη και λευκίνη) [117]. Μπορεί να θεωρηθεί ως θετικός παράγοντας για την αύξηση της θερμοσταθερότητας των σφαιρικών πρωτεϊνών. Ο αλειφατικός δείκτης μιας πρωτεΐνης υπολογίζεται σύμφωνα με τον ακόλουθο τύπο:

$$Αλειφατικός\ δείκτης = X(A) + aX(V) + b\big(X(I) + X(L)\big),$$

όπου τα $X(A), X(V), X(I)\ και\ X(L)$ είναι τα ποσοστό των mole (100 X mole κλάσμα) των αμινοξέων αλανίνη, βαλίνη, ισολευκίνη και λευκίνη στο πρωτεϊνικό μόριο, αντίστοιχα. Οι συντελεστές a και b είναι ο σχετικός όγκος της πλευρικής αλυσίδας βαλίνης προς την πλευρική αλυσίδα της αλανίνης (a = 2,9) και των πλευρικών αλυσίδων Leu/Ile προς την πλευρική αλυσίδα της αλανίνης (b = 3,9).

GRAVY

Η τιμή GRAVY για ένα πεπτίδιο ή πρωτεΐνη υπολογίζεται ως το άθροισμα των τιμών υδροπάθειας όλων των αμινοξέων βάσει των Kyle-Doolittle, διαιρούμενο με τον αριθμό των καταλοίπων στην ακολουθία [118]. (Παράρτημα ΙΙ)

4.1.1.2 DOMAINATION

Το DOMAINATION θέτει τα όριο των επικρατειών, εφαρμόζοντας το PSI-BLAST με επαναλαμβανόμενο τρόπο. Η κατανομή των στοιχισμένων αμινοτελικών και καρβοξυτελικών άκρων από ευθυγραμμίσεις τοπικής ακολουθίας PSI-BLAST χρησιμοποιείται για τον εντοπισμό πιθανών ορίων επικρατειών. Το DOMAINATION ενσωματώνει μια νέα επαναληπτική στρατηγική για την κοπή και τη σύνδεση επικρατειών και τμημάτων επικρατειών σε μια προσπάθεια παρακολούθησης της «εξελικτικής οδού» μιας πρωτεΐνης από την απώλεια και την απόκτηση επικρατειών. Αυτό επιτρέπει την αναγνώριση τόσο των συνεχών όσο και των ασυνεχών επικρατειών. Για κάθε επικράτεια που προκύπτει από τις τοπικές ευθυγραμμίσεις PSI-BLAST, σχηματίζονται προφίλ μέσω φιλτραρίσματος περιττών ακολουθιών και επακόλουθων πολλαπλών στοιχίσεων ακολουθιών. Κάθε φιλτραρισμένο προφίλ χρησιμοποιείται στη συνέχεια σε περαιτέρω επαναληπτικές αναζητήσεις βάσης δεδομένων χρησιμοποιώντας PSI-BLAST. Όλα τα προφίλ απαιτείται να περιέχουν την αρχική ακολουθία επερώτησης σε κάθε επανάληψη του PSI-BLAST για την αποφυγή κάποιου λάθους προφίλ, αλλά οι παράμετροι ορίζονται για να διασφαλιστεί ότι τα προφίλ είναι αρκετά αποκλίνοντα, ώστε να καταγράφουν εξελικτικά απομακρυσμένα τμήματα ακολουθιών. Η όλη διαδικασία των επαναληπτικών αναζητήσεων PSI-BLAST επαναλαμβάνεται μέχρι να τεθεί η επικράτεια και να μην βρεθούν πλέον νέες ομόλογες ακολουθίες [119].

4.1.1.3 INTERPRO

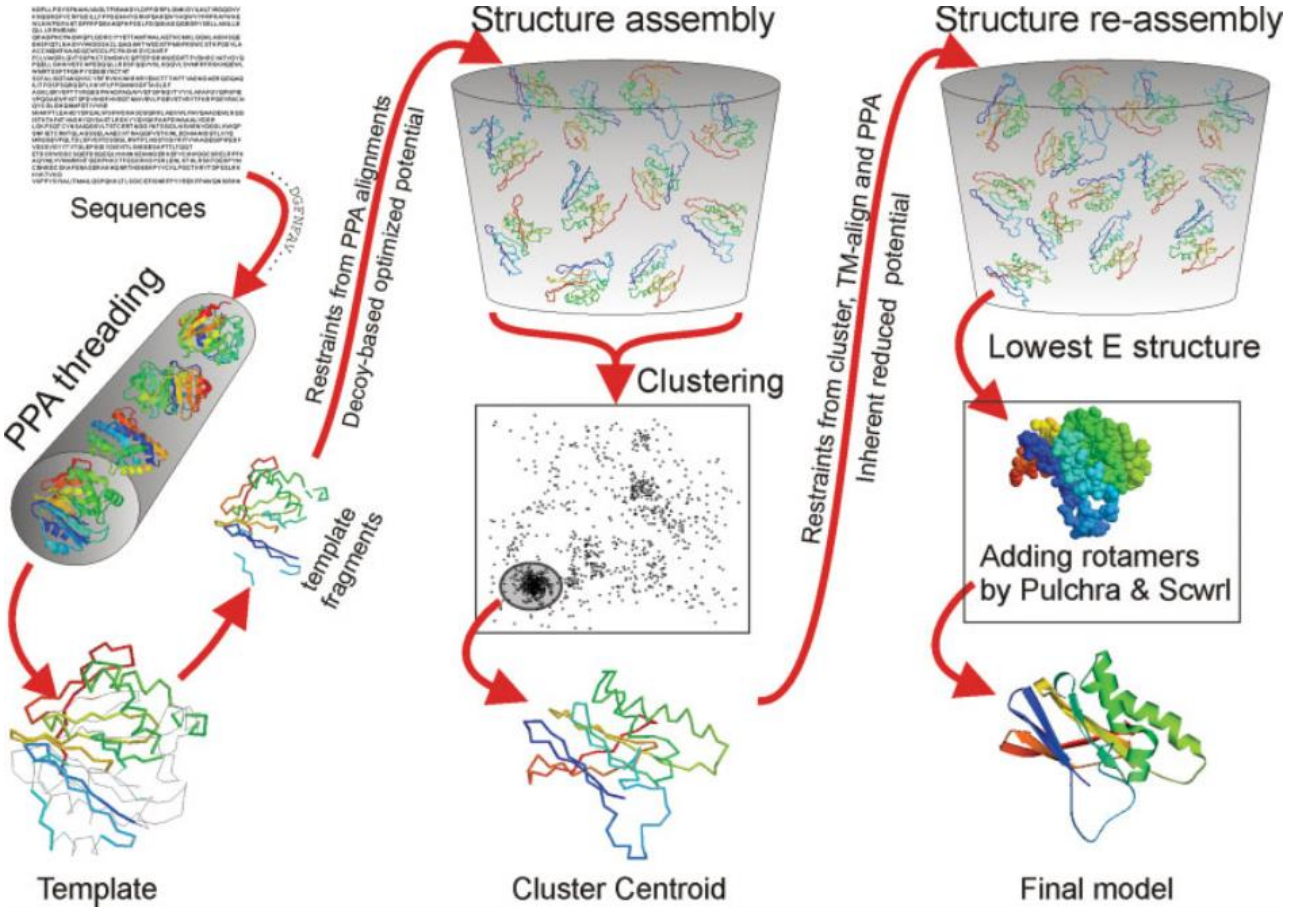
Ο διαδικτυακός διακομιστή InterPro χρησιμοποιείται για μία ευρεία σειρά διαφορετικών πεδίων, όπως είναι ο προσδιορισμός της οικογένειας πρωτεϊνών στην οποία ανήκει η επιθυμητή πρωτεΐνη, τα πρωτεϊνικές περιοχές, θέσεις ή/και άλλα χαρακτηριστικά αυτής, ο σχολιασμός της οικογένειας πρωτεϊνών που ανήκει. Η InterPro ενσωματώνει 13 διαφορετικές βάσεις δεδομένων πρωτεϊνών σε έναν κοινό πόρο. Ορισμένες από αυτές είναι η Cath-Gene3D, η Conserved Domains (CDD), η HAMAP, η PANTHER, η Pfam,η PIRSF,

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

η PROSITE, η SMART, βάση δεδομένων σύνδεσης Δομής και Λειτουργίας (SFLD), η SuperFamily και η TIGRFAM. Οι βάσεις δεδομένων παρέχουν συμπληρωματικά πεδία σχολιασμού των πρωτεϊνών και των πρωτεϊνικών οικογενειών, καθιστώντας το μια ολοκληρωμένη πηγή για την αναζήτηση οικογενειών, τομέων και λειτουργικών θέσεων των πρωτεϊνών [120].

4.1.1.4 I-TASSER

Η πρόβλεψη της δομής αναφέρεται στην προσπάθεια δημιουργίας τριδιάστατων μοντέλων από αμινοξικές αλληλουχίες χρησιμοποιώντας υπολογιστικούς αλγορίθμους. Το I-TASSER (Iterative Threading ASSEmbly Refinement) είναι μία διαδικτυακή πλατφόρμα, η οποία βάσει της ιεραρχικής προσέγγισης επιτρέπει την πρόβλεψη των πρωτεϊνικών δομών και της λειτουργίας του βάσει της δομής τους. Ο αλγόριθμος του I-TASSER αποτελείται από τρία διαδοχικά βήματα: την ύφανση των πρωτεϊνών (threading), τη συναρμολόγηση των τμημάτων και την επανάληψη [121-124].



Εικόνα 7 Το διάγραμμα της μεθοδολογίας I-TASSER

Κατά το πρώτο στάδιο, στοιχίζονται τα προφίλ των δευτεροταγών δομών (Profile-Profile Threading Alignment, PPA) ανάμεσα στην πρωτεΐνη-στόχο και στη βάση δεδομένων πρωτεϊνών (Protein DataBase), όπου εντοπίζονται γνωστές δομές γνωστών πρωτεϊνών (PPA threading). Ο βαθμός στοίχισης ανάμεσα στην ακολουθία επερώτησης (*i*) και την ακολουθία της δομής εκμαγείου (*j*) από τη βάση δεδομένων ορίζεται ως:

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

$$\mathrm{Score}(i,j) = \sum_{k=1}^{20} F_{query}(i,k) P_{template}(j,k) + c_1 \delta\left(s_{query}(i), s_{template}(j)\right) + c_2$$

όπου $F_{query}(i,k)$ είναι η συχνότητα ανίχνευσης *k* αριθμού αμινοξέων για μία θέση *i* της ακολουθίας επερώτησης που προκύπτει ύστερα από την στοίχιση πολλαπλών ακολουθιών στη βάση δεδομένων PDB βάσει των μεθόδων στοίχισης PSI-BLAST (Position-Specific Iterative Basic Local Alignment Search Tool) και HMM (Hidden Markov Model), $P_{template}(j,k)$ είναι οι αθροιστικές λογαριθμικές πιθανότητες ενός *k* αριθμού αμινοξέων για μία θέση *j* της πρωτεϊνικής ακολουθίας εκμαγείου που προκύπτει ύστερα από την στοίχιση πολλαπλών ακολουθιών στη βάση δεδομένων PDB βάσει των μεθόδων στοίχισης PSI-BLAST και HMM, $s_{query}(i)$ είναι ο συνδυασμός της προβλεπόμενης δευτεροταγούς δομή για τη θέση *i* της ακολουθίας επερώτησης με τη χρήση των προγραμμάτων PSIPRED και SAM, και $s_{template}(j)$ είναι η δευτεροταγή δομή που έχει ανατεθεί για τη θέση *j* της πρωτεΐνης εκμαγείου μέσω του αλγορίθμου DSSP (Define Secondary Structure of Proteins). Ο συνδυασμός των PSIPRED και SAM αθροίζει τις πιθανότητες για τη συγκεκριμένη θέση να αποτελεί μία από τις τρεις καταστάσεις έλικας/επιφάνειας/σπειράματος και έπειτα επιλέγεται η κατάσταση με τη μεγαλύτερη πιθανότητα, η οποία είναι και επακόλουθη της εξομάλυνσης όλων των μεμονωμένων δευτεροταγών δομών κατά μήκος της ακολουθίας επερώτησης. Ο δυναμικός προγραμματισμός των Needelman-Wunsch και Smith-Waterman χρησιμοποιείται για τον καθορισμό της καλύτερης αντιστοίχισης ανάμεσα στην ακολουθία επερώτησης και την ακολουθία εκμαγείου από τη βάση δεδομένων. Οι ακολουθίες στόχοι θα κατηγοριοποιηθούν, επομένως, ανάλογα της σημασίας και των αποτελεσμάτων των μεθόδων PPA: Ένας εύκολος στόχος θεωρείται όταν το *Z*-score είναι μεγαλύτερο του 8, ένας δύσκολος στόχος όταν το *Z*-score είναι μικρότερο του 7, όπου στην περίπτωση αυτή δε θεωρείται να υπάρχει κάποια αντιστοίχιση, και τέλος, για τις ενδιάμεσες τιμές *Z*-score οι στόχοι θεωρούνται μέτριοι.

Με την παραπάνω στοίχιση, ουσιαστικά οι ακολουθία στόχος χωρίζεται σε στοιχισμένες και σε μη στοιχισμένες περιοχές βάσει την δευτεροταγούς δομής. Τα τμήματα της ακολουθίας που έχουν στοιχηθεί εξέρχοναται αυτομάτων από την ακολουθία εκμαγείο και επιτρέπεται η περιστροφή τους σε ένα μη κυβικό σύστημα. Αντίθετα, τα μη στοιχισμένα τμήματα προσομοιώνται βάσει *ab initio* μοντελισμού σε ένα κυβικό σύστημα συντεταγμένων μεγέθους 0.87Å. Η τοπολογία ολόκληρου του μορίου είναι αποτέλεσμα του σχετικού επαναπροσανατολισμό των συνεχόμενων τμημάτων, ενώ τα μοντέλα που βρίσκονται σε πλέγμα (μη στοιχισμένα) λειτουργούν ως συνδέτες των κινήσεων στο χώρο των άκαμπτων τμημάτων (στοιχισμένων) του μορίου. Οι διαμορφώσεις της πρωτεΐνης παρουσιάζονται ως ίχνη των ατόμων C_a των αμινοξέων και της μάζας των πλευρικών αλυσίδων. Το πεδίο δύναμης είναι αποτέλεσμα μίας ποικιλίας γνωστών ενεργειακών όρων, που περιλαμβάνουν τις αλληλεπιδράσεις των πλευρικών αλυσίδων γειτονικών αμινοξέων, ο προσανατολισμός των κύριων προβλέψεων των δευτεροταγών δομών από τα προγράμματα PSIPRED και SAM, την προσβασιμότητα των διαλυμάτων βάσει το κατάλοιπο με χρήση νευρωνικών δικτύων, τους δεσμούς υδρογόνου του κορμού των δευτεροταγών δομών και των επαφών των πλευρικών αλυσίδων και της απόστασης των ατόμων C_a των αμινοξέων της κύριας ακολουθίας που προκύπτουν από τις πολλαπλές στοιχίσεις. Οι ισορροπίες ανάμεσα στους όρους της ενέργειας προσαρμόζονται ξεχωριστά για τις

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

τρεις κατηγορίες (Εύκολος, Μέτριος και Δύσκολος στόχος), λαμβάνοντας υπόψιν το μέγιστο βαθμό συσχέτισης της ολικής ενέργειας και του TM-score που βασίζεται στο σύνολο όλων των συνεχόμενων κατανεμημένων δομικών στόχων. Η διαδικασία συσχέτισης της δομής είναι το αποτέλεσμα ενός τροποποιημένου συστήματος προσομοίωσης Monte-Carlo, ενώ η ομαδοποίηση των δομών προκύπτει από το πρόγραμμα SPICKER. Τα κέντρα των ομάδων (cluster centroids) που σχηματίζονται καθορίζονται από τον υπολογισμό του μέσου όρου των συντεταγμένων όλων των δολωμάτων (πιθανές δομές) μίας ομάδας και ταξινομούνται βάσει της πυκνότητας της δομής τους.

Για τον αποκλεισμό των συγκρούσεων ανάμεσα στα κέντρα των δομών και για τη βελτιστοποίηση του μοντέλου, πραγματοποιείται ξανά η προσομοίωση της συναρμολόγησης των τμημάτων, με τη διαφορά ότι η προσομοίωση γίνεται πάνω στις ομάδες των δολωμάτων που προέκυψαν από την τελευταία προσομοίωση. Οι περιορισμοί στο χώρο εξάγονται από τα κέντρα των ομάδων και γίνεται αναζήτηση στη βάση δεδομένων PDB με το πρόγραμμα ευθυγράμμισης δομής TM-align. Τέλος, επιλέγονται οι δομές με τη χαμηλότερη ενέργεια σε κάθε τμήμα, με καθορισμένα άτομα C_a και πλευρικών αλυσίδων.

Στο I-TASSER ο βαθμός και η έκτασης της εγγενούς θερμικής κινητικότητας των καταλοίπων και τμημάτων της πρωτεΐνης επερώτησης καθορίζεται από τον B-Factor. Στο I-TASSER η τιμή αυτή συνάγεται από τις πρότυπες πρωτεΐνες threading από τη λίστα του PDB σε συνδυασμό με τα προφίλ αλληλουχίας που προέρχονται από τις βάσεις δεδομένων ακολουθιών. Στο παρακάτω σχήμα, οι αναφερόμενες τιμές αντιστοιχούν στον κανονικοποιημένο B-Factor (B-Factor profile, BFP) της πρωτεΐνης στόχου, που ορίζεται από τον τύπο:

$$B = {(B' - u)}/{s}$$

όπου B' είναι η τιμή του B-Factor και οι τιμές u και s είναι οι η μέση και η τυπική απόκλιση της τιμής B-Factor κατά μήκος της αλληλουχίας.

Για κάθε στόχο, οι προσομοιώσεις I-TASSER δημιουργούν ένα μεγάλο σύνολο δομικών διαμορφώσεων, που ονομάζονται δολώματα (decoys). Για να επιλέξει ο διακομιστής τα τελικά μοντέλα, το I-TASSER χρησιμοποιεί το πρόγραμμα SPICKER για να ομαδοποιήσει όλα τα δολώματα με βάση την ομοιότητα της δομής ανά ζεύγη και αναφέρει 5 μοντέλα που αντιστοιχούν στα 5 μεγαλύτερα σύμπλοκα δομών.

Το I-TASSER δημιουργεί μοντέλα πρωτεϊνών πλήρους μήκους, αποσπώντας συνεχή θραύσματα από στοιχίσεις αναγνώρισης δομών και στη συνέχεια τα επανασυναρμολογεί, χρησιμοποιώντας προσομοιώσεις Monte Carlo. Δολώματα χαμηλής θερμοκρασίας που δημιουργούνται κατά τη διάρκεια της προσομοίωσης συγκεντρώνονται από το SPICKER και οι κορυφαία πέντε ομαδοποιήσεις επιλέγονται για τη δημιουργία πλήρων ατομικών μοντέλων. Η πυκνότητα των ομάδων (cluster density) ορίζεται ως ο αριθμός των δομικών δολωμάτων όπως δίνεται από το SPICKER. Υψηλή cluster density σημαίνει ότι η δομή εμφανίζεται συχνότερα στην τροχιά προσομοίωσης και συνεπώς σημαίνει ότι είναι ένα μοντέλο καλύτερης ποιότητας. Ο βαθμός εμπιστοσύνης του κάθε μοντέλου υπολογίζεται ποσοτικά μέσω του C-score, μίας τιμής που υπολογίζεται βάσει της σημασίας των πρότυπων στοιχίσεων ανάγνωσης δομής και των παραμέτρων σύγκλισης των προσομοιώσεων συναρμολόγησης δομής, δηλαδή αποτελεί τον δείκτη της ποιότητας του πιθανού μοντέλου I-TASSER. Τυπικά το C-score έχει εύρος [-5,2], όπου το C-score υψηλότερης

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

τιμής σημαίνει ότι είναι ένα μοντέλο με μεγαλύτερο βαθμό εμπιστοσύνης και αντίστροφα. Ένα μοντέλο με σωστή ολικό τοπολογία έχει C-score >-1.5. Δεδομένου ότι τα top 5 μοντέλα ταξινομούνται βάσει του μεγέθους ομαδοποίησης, είναι πιθανό τα μοντέλα χαμηλότερης τάξης (2-5) να έχουν υψηλότερη τιμή C-score σε ορισμένες περιπτώσεις. Αν και το πρώτο μοντέλο έχει καλύτερη ποιότητα στις περισσότερες περιπτώσεις, είναι επίσης πιθανό τα μοντέλα χαμηλότερης τάξης να έχουν καλύτερη ποιότητα από τα μοντέλα υψηλότερης κατάταξης, όπως φαίνεται από τις δομές αναφοράς. Ένα μοντέλο με σωστή ολικό τοπολογία έχει C-score >-1.5.

Οι τιμές TM-score και RMSD είναι γνωστές σταθερές για τον υπολογισμό της δομικής ομοιότητας ανάμεσα σε δύο δομές που χρησιμοποιούνται, συνήθως, για τη μέτρηση της ακρίβειας της μοντελοποίησης των δομών, όταν είναι γνωστή η φυσική τους δομή. Στην περίπτωση που η φυσική τους δομή δεν είναι γνωστή, είναι απαραίτητο να προβλεφθεί η ποιότητας του προβλεπόμενου μοντέλου, δηλαδή ποια είναι η απόσταση ανάμεσα του προβλεπόμενου μοντέλου και των φυσικών δομών; Αυτές αναφέρονται μόνο για το μοντέλο 1, καθώς φαίνεται μέσα από πειράματα δοκιμών αναφοράς, πως η σχέση μεταξύ του C-score και της τιμής TM-score είναι μικρότερη σε μοντέλα μικρότερης κατάταξης. Για αυτό αναφέρεται μόνο η τιμή C-score.

Η τιμή TM-score είναι μια πρόσφατα προτεινόμενη κλίμακα για τη μέτρηση της δομικής ομοιότητας μεταξύ δύο δομών. Ο σκοπός της TM-score είναι να επιλυθεί το πρόβλημα του RMSD, που είναι ευαίσθητο στο τοπικό σφάλμα. Επειδή το RMSD είναι μια μέση απόσταση όλων των ζευγών των καταλοίπων σε δύο δομές, ένα τοπικό σφάλμα (π.χ. ένας αποπροσανατολισμός της ουράς) θα δημιουργήσει μια μεγάλη τιμή RMSD αν και η ολική τοπολογία είναι σωστή. Στην TM-score, ωστόσο, η μικρή απόσταση σταθμίζεται ισχυρότερη από τη μεγάλη απόσταση που καθιστά τη βαθμολογία μη ευαίσθητη στο τοπικό σφάλμα μοντελοποίησης. Μια βαθμολογία TM> 0,5 δείχνει ένα μοντέλο σωστής τοπολογίας και μια βαθμολογία TM <0,17 σημαίνει τυχαία ομοιότητα. Αυτά τα όρια δεν εξαρτώνται από το μήκος της πρωτεΐνης.

Η τοπική ακρίβεια ορίζεται ως η απόκλιση της απόστασης (σε Angstrom) μεταξύ των θέσεων των καταλοίπων στο μοντέλο και στης φυσικής δομής. Δεδομένου ότι η φυσική δομή είναι άγνωστη, τα σφάλματα απόστασης στα ακόλουθα γραφικά υπολογίζονται από το ResQ χρησιμοποιώντας μία παλινδρόμηση του φορέα υποστήριξης, που κάνει χρήση της κάλυψης της στοίχισης ανάγνωσης δομής, της απόκλισης των δολωμάτων προσομοίωσης I-TASSER και της δευτεροταγούς δομής και των προβλέψεων προσβασιμότητας διαλύτη βάσει των πρωτεϊνικών ακολουθιών. Όσο μικρότερο τόσο καλύτερα.

Το I-TASSER μπορεί και παραθέτει ορισμένα σχόλια για τη λειτουργία της πρωτεΐνης επερώτησης με τη χρήση των COFACTOR και COACH που βασίζονται στην πρόβλεψη της δομής που προκύπτει από το ίδιο το πρόγραμμα. Ενώ το COFACTOR συμπεραίνει τις λειτουργίες της πρωτεΐνης (ligand-binding sites, EC, GO) μέσω σύγκριση δομών και δίκτυα ανάμεσα σε πρωτεΐνες, το COACH είναι ένας διακομιστής μετασυστημάτων που συνδυάζει τα αποτελέσματα για τις παραπάνω λειτουργίες από το COFACTOR, TM-SITE και S-SITE.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Το COFACTOR είναι μία μέθοδος που βασίζεται στη δομή, την αλληλουχία και την αλληλεπίδραση ανάμεσα στις πρωτεΐνες για την πρόβλεψη της βιολογικής λειτουργίας των μορίων των πρωτεϊνών. Ξεκινώντας από το τρισδιάστατο μοντέλο της πρωτεΐνης, το COFACTOR κάνει treading της επερώτησης μέσω της βάσης δεδομένων λειτουργίας πρωτεϊνών BioLip από local και global αντιστοιχίσεις δομών για την αναγνώριση λειτουργικών περιοχών και ομολογιών. Οι προβλέψεις της λειτουργίας προκύπτουν από τα εκμαγεία που έχουν την καλύτερα λειτουργική ομολογία. Για το Gene Ontology (GO) επιπλέον προβλέψεις προκύπτουν από το UniProt-GOA μέσω της στοίχισης αλληλουχιών και αλληλουχιών-προφίλ και από το STRING. Το COACH είναι ένα διακομιστής μετασυστημάτων που χρησιμοποιείται για την πρόβλεψη των θέσεων σύνδεσης του προσδέτη στην πρωτεΐνη. Ξεκινώντας από τη δομή που δίνεται για την αλληλουχία επερώτησης, το COACH σχηματίζει συμπληρωματικές προβλέψεις θέσεων σύνδεσης του προσδέτη στην πρωτεΐνη χρησιμοποιώντας δύο διαφορετικές μεθόδους, το TM-SITE και το S-SITE, που αναγνωρίζουν από αναγνωρίζουν εκμαγεία σύνδεσης προσδετών από τη βάση δεδομένων λειτουργίας πρωτεϊνών BioLip μέσω σύγκρισης μικρότερων δομών σχετικά με τη σύνδεση των προσδετών και τα προφίλ αλληλουχιών. Αυτές οι προβλέψεις συνδυάζονται με τα αποτελέσματα από άλλες μεθόδους (COFACTOR, FINDSITE, ConCavity) για τη δημιουργία τελικών προβλέψεων θέσεων σύνδεσης προσδετών.

4.1.1.5 PONDR

Οι ασταθείς περιοχές (Disordered Regions, DRs) ορίζονται οι πρωτεΐνες ή τμήματα αυτών που στερούνται μία καθορισμένη τριτοταγή δομή. Με το διακομιστή PONDR (Predictor of Natural Disordered Regions) είναι εφικτή η πρόβλεψη τέτοιων περιοχών στην ακολουθία-στόχο. Ο καθορισμός της αστάθιας, των τμημάτων ή/και ολόκληρης της πρωτεΐνης, αφορά τον κορμό της πρωτεϊνικής ακολουθίας, και όχι τις πλευρικές αλυσίδες. Σε μία οικογένεια πρωτεϊνών, ο κορμός των δομικά καθορισμένων πρωτεϊνικών περιοχών έχουν τις ίδιες γωνίες Ramachandran, ενώ τα ασταθή τμήματα αυτών φέρουν διαφορετικές, συχνά δυναμικές, γωνίες Ramachandran. Αυτές οι περιοχές μπορεί να είναι είτε τυχαία σπειράματα είτε έχουν μία ορισμένη σύσταση δευτεροταγούς δομής, χωρίς μία καλά καθορισμένη τριτοταγή δομή (collapsed, πχ molten globun-like). Αυτές φαίνεται να εμπλέκονται σε μία ποικιλία λειτουργιών συμπεριλαμβανομένων: την αναγνώριση DAN/RNA/πρωτεϊνών, τη διαμόρφωση της συγγένειας/ειδικότητας της πρόσδεσης πρωτεϊνών, τη μοριακή αναγνώριση δομής.

Το PONDR προβλέπει ασταθή τμήματα σε μικρές αμινοξικές ακολουθίες μέσω εμπρόσθιων νευρωνικών δικτύων (feedforward neural networks) που χρησιμοποιούν χαρακτηριστικά της αλληλουχίας σε παράθυρα 9-21 αμινοξέων. Αυτά τα χαρακτηριστικά, όπως κλασματική σύσταση συγκεκριμένων αμινοξέων, υδροφοβικότητα, ή πολυπλοκότητα αλληλουχίας υπολογίζονται σε αυτό το εύρος παραθύρων. Τα αποτελέσματα των νευρωνικών δικτύων έχουν εύρος από 0 έως 1 και φιλτράρονται σε ένα παράθυρο 9 αμινοξέων. Εάν η τιμή ενός καταλοίπου είναι ίση ή μεγαλύτερη του 0.5 το κατάλοιπο ή η περιοχή θεωρείται ασταθές [125].

4.1.2 Παραγωγή Πρωτεϊνών στο Ετερόλογο Σύστημα *E.coli*

Η ανάγκη για την βιοχημική ανάλυση της δομής και της λειτουργίας των πρωτεϊνών, όπως και η μαζική παραγωγή τους για φαρμακευτικούς σκοπούς, επέτρεψε την ανάπτυξη μεθόδων που θα μπορέσουν να παράγουν μεγάλες ποσότητες των επιθυμητών πρωτεϊνών και σε υψηλό βαθμό καθαρότητας. Μια πλέον

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

διαδεδομένη βιοχημική τεχνική για την επίτευξη των παραπάνω αποτελεί η ετερόλογη έκφραση των πρωτεϊνών. Η ετερόλογη έκφραση των πρωτεϊνών είναι η εισαγωγή των γονιδίων ενός οργανισμού, που κωδικοποιούν για τις επιθυμητές πρωτεΐνες, σε ένα κύτταρο ενός άλλου οργανισμού (ξενιστές), τα οποία με τους δικούς τους μηχανισμούς θα καταφέρουν να τις εκφράσουν [126].

Η επιλογή του κυττάρου ξενική, του οποίου ο μηχανισμός παραγωγής των πρωτεϊνών θα χρησιμοποιηθεί, αποτελεί και καθοριστικό σημείο στη διαδικασία αυτή. Προκαρυωτικοί οργανισμοί που μπορούν να χρησιμοποιηθούν αποτελούν τα βακτήρια, ενώ ως ευκαρυωτικοί η ζύμη, οι νηματώδης μύκητες και τα μονοκύτταρα φύκη, τα έντομα, φυτά και ζώα, ακόμα και διάφοροι τύποι ιών. Από το 1977, όταν ο Itakura και οι συνεργάτες του κατάφεραν πρώτοι να εκφράσουν την πρώτη ανασυνδυασμένη πρωτεΐνη, το βακτηριακό σύστημα χρησιμοποιείται ευρύτατα για την ετερόλογη έκφραση των πρωτεϊνών [127].

Συγκεκριμένα, το Gram-αρνητικό βακτήριο *Escherichia coli* είναι το πλέον διαδεδομένο για το σκοπό αυτό, λόγω των πλεονεκτημάτων που φέρει. Το *E. coli* φέρει πολύ μεγάλη κινητική ταχείας ανάπτυξης, έχοντας χρόνο διπλασιασμού περίπου 20 λεπτά [128,129]. Οι καλλιέργειες του είναι υψηλής πυκνότητας, ενώ ενισχύεται με τη χρήση κατάλληλων άμεσα διαθέσιμων και, παράλληλα, φθηνών μέσων ανάπτυξής [130-132]. Τέλος, ο μετασχηματισμός του εξωγενούς γενετικού υλικού στο *E. coli* είναι εύκολος και γρήγορος [133].

Ωστόσο, ο μετασχηματισμός ενός ξένου γονιδίου στο *E. coli* επηρεάζει τις συνθήκες έκφρασης του, όπως την οσμωτικότητα, το δυναμικό οξειδοαναγωγής, η παρουσία συμπαραγόντων έκφραση, οι μηχανισμοί αναδίπλωσης και μετα- μεταφραστικές τροποποιήσεις [126]. Επιπλέον, σε υψηλής έκφρασης πρωτεϊνών, οι υδροφοβικές πολυπεπτιδικές ακολουθίες είναι διαθέσιμες για αλληλεπίδραση με παρόμοιες περιοχές, λόγω αυξημένης συγκέντρωσης τους. Όλοι οι παραπάνω παράγοντες οδηγούν σε αστάθεια της παραγόμενης πρωτεΐνης, καθώς δεν είναι διαλυτή και σωστά αναδιπλωμένη, με αποτέλεσμα η συσσώρευσή της να οδηγεί στο σχηματισμό συσσωματωμάτων, γνωστών και ως έγκλειστα σωμάτια (inclusion bodies) [134,135].

4.1.3 Τεχνολογία Ανασυνδυασμένου DNA

4.1.3.1 Κλωνοποίηση

Η μοριακή κλωνοποίηση παρέχει ένα μέσο για την εκμετάλλευση της ταχείας ανάπτυξης των βακτηριακών κυττάρων για την παραγωγή μεγάλων ποσοτήτων πανομοιότυπων θραυσμάτων DNA, τα οποία από μόνα τους δεν έχουν την ικανότητα να αναπαράγονται. Το θραύσμα του DNA που πρόκειται να ενισχυθούν εισάγεται πρώτα σε έναν φορέα κλωνοποίησης. Μετά την εισαγωγή του ξένου DNA στον φορέα, το κατασκευασμένο DNA μετασχηματίζεται σε ένα βακτηριακό κύτταρο, το πιο συχνά χρησιμοποιούμενο βακτήριο είναι το *E. coli*. Η αναπτυσσόμενη βακτηριακή καλλιέργεια αντιγράφει το ανασυνδυασμένο DNA σε εκατοντάδες αντίγραφα ανά κύτταρο. Αυτή η διαδικασία αποδίδει πολλαπλούς, πανομοιότυπους κλώνους του αρχικού ανασυνδυασμένου μορίου.

4.1.3.2 Πλασμιδιακός φορέας

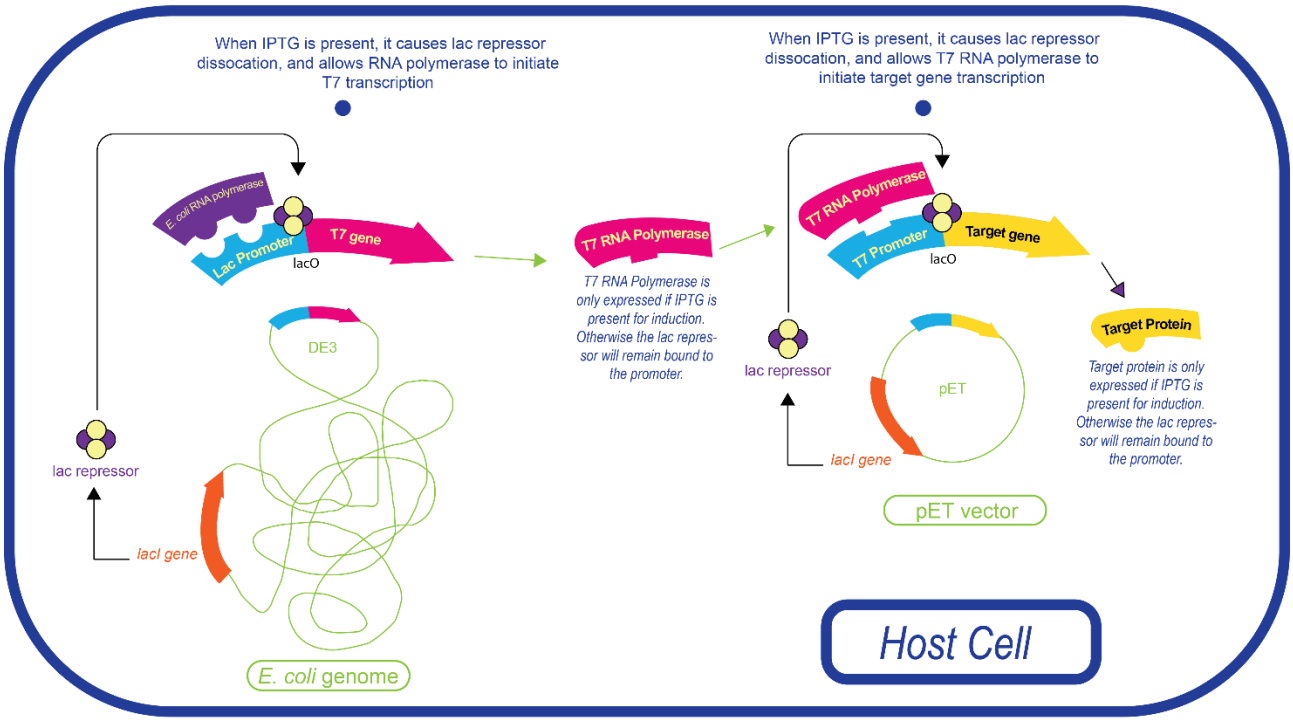
Τα πλασμίδια είναι οι πιο συχνά χρησιμοποιούμενοι φορείς βακτηριακής κλωνοποίησης, λόγω των ιδιοτήτων που έχουν. Τα πλασμίδια είναι αυτόνομα κυκλικά μόρια DNA που εντοπίζονται σε μία μεγάλη ποικιλία

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

μεγέθους και χαρακτηριστικών τόσο σε Gram- θετικά όσο και και σε Gram- αρνητικά βακτήρια. Τα πλασμίδια που χρησιμοποιούνται για κλωνοποίηση διακρίνονται από ορισμένα ειδικά χαρακτηριστικά, όπως είναι η περιοχή της αντιγραφής (ρεπλικόνιο), ο υποκινητής, η περιοχής κλωνοποίησης, το γονίδιο επιλογής του πλασμιδίου, ένας μοριακός δείκτης, η εισαγωγή ετικέτας στην παραγόμενη πρωτεΐνη και η θέση αφαίρεσης της ετικέτας. Ορισμένα πλασμίδια φέρουν γονίδια, τα οποία προσδίδουν ανθεκτικότητα σε ένα ή περισσότερα αντιβιοτικά στον ξενιστή στον οποίο βρίσκονται. Τέλος, στα πλασμίδια υπάρχει και η περιοχή του πολυσυνδέτη, μία περιοχή η οποία επιτρέπει την ένθεση θραυσμάτων DNA μέσω της χρήσης ενζύμων περιορισμού [126].

Ένα βασικό χαρακτηριστικό γνώρισμα των πλασμιδίων είναι ο φυσικός διαχωρισμός τους από το χρωμόσωμα του βακτηρίου και η σταθερή διατήρηση τους σε αυτήν την εξωχρωμοσωμική κατάσταση. Τα πλασμίδια μπορούν να αντιγραφούν αυτόνομα σε ένα βακτηριακό ξενιστή, ρυθμίζοντας μόνα τους τον τρόπο της αντιγραφής τους, ενώ παράλληλα χρησιμοποιούν τους τρόπους αντιγραφής του ίδιου του κυττάρου. Το χαρακτηριστικό τους αυτό είναι αποτέλεσμα παρουσίας μίας περιοχής έναρξης (origin, *ori*) στο πλασμίδιο που σχετίζεται με cis-δραστικά στοιχεία του πλασμιδίου. Η περιοχή αυτή καθορίζει και τον αριθμό αντιγράφων του πλασμιδίου μέσα στον ξενιστή [132,136].

Τα πλασμίδια φέρουν έναν υποκινητή μπροστά από την πρωτεΐνη στόχο ή και το γονίδιο ανθεκτικότητας σε αντιβιοτικά για τη συντονισμένη και ελεγχόμενη έκφραση ανασυνδυασμένων πρωτεϊνών. Οι πιο δημοφιλείς υποκινητές που χρησιμοποιούνται για τις ανασυνδυασμένες πρωτεΐνες είναι ο υποκινητής lac και ο υποκινητής του βακτηριοφάγου T7 [126]. Το οπερόνιο lac είναι απαραίτητο για τη μεταφορά και το μεταβολισμό της λακτόζης στα βακτήρια της *E. coli*. Αν και για το μεταβολισμό των βακτηρίων η γλυκόζη παρουσιάζεται ως η βασική πηγή ενέργειας, το οπερόνιο της λακτόζης επιτρέπει την αποτελεσματική πέψη της



Εικόνα 8 Το σύνολο ενός ετερόλογου συστήματος κάτω από την επαγωγή με IPTG, για την υπερέκφραση της ανασυνδυασμένης πρωτεΐνης-στόχου σε ένα βακτηριακό στέλεχος *E. coli* (DE3) [136].

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

λακτόζης όταν η γλυκόζη δεν είναι διαθέσιμη στο μέσο ανάπτυξης των βακτηρίων. Ο μηχανισμός μεταβολισμού της λακτόζης ρυθμίζεται από τον υποκινητή lac και την παρουσία του καταστολέα lacI στο χειριστή του οπερονίου [137]. Στο σύστημα του υποκινητή του βακτηριοφάγου T7, όταν το γονίδιο-στόχος βρίσκεται πίσω από τον υποκινητή που αναγνωρίζεται από την RNA πολυμεράση του φάγου T7 (T7 RNAP). Αυτή η υψηλά δραστική πολυμεράση πρέπει να παρέχεται από ένα άλλο πλασμίδιο ή εντοπίζεται στο βακτηριακό γένωμα κάτω από τον έλεγχο της έκφρασης του από έναν υποκινητή lacUV5 [138]. Έτσι, το σύστημα μπορεί να επάγεται είτε από τη λακτόζη είτε από το μη-υδρολυομένο ανάλογο της IPTG. Η έκφραση της ανασυνδυασμένης πρωτεΐνης ρυθμίζεται από τον αναστολέα $lacI^Q$ και τη συνέκφραση με τη λυσοζύμη T7 [139]. Η λυσοζύμη T7 συνδέεται με την T7 RNAP και αναστέλλει την μεταγραφή από τον υποκινητή T7. Η T7 λυσοζύμη παρέχεται από επιπλέον πλασμίδια στο βακτηριακό στέλεχος (pLysS ή pLysE). Ένα επιπλέον επίπεδο ελέγχου έγκειται στην εισαγωγή ενός χειριστή lacO μετά τον υποκινητή T7. Και οι τρεις αυτοί μηχανισμοί σε συνδυασμό καθιστούν το σύστημα του πλασμιδίου ιδανικό για την υπερέκφραση της επιθυμητής πρωτεΐνης [140].

Για τη βιοχημική μελέτη της ανασυνδυασμένης πρωτεΐνης, χρειάζεται η ανίχνευσή της ανάμεσα σε όλες τις άλλες παραγόμενες πρωτεΐνες, η παροχή σε αυτήν μέγιστης διαλυτότητας και ο εύκολος καθαρισμός της από το πρωτεΐνωμα του *E.coli* [141]. Για το σκοπό αυτό είναι απαραίτητος ο σχηματισμός μίας χιμαιρικής πρωτεΐνης, κατά την οποία θα φέρεται μία πεπτιδική ετικέτα ή μία πολυπεπτιδική αλυσίδα στο αμινο- ή καρβοξυτελικό άκρο της επιθυμητής πρωτεΐνης. Οι μικρές πεπτιδικές ετικέτες (πχ poly-Arg-, FLAG-, poly-His-, c-Myc-, S- και Strep II) είναι λιγότερο πιθανό να παρέμβουν στην διαλυτότητα της παραγόμενης πρωτεΐνης. Η επισημασμένη ανασυνδυασμένη πρωτεΐνη μπορεί να ανιχνευθεί εύκολα, όπως και μέσω της χρωματογραφίας συγγένειας ενός σταδίου καθίσταται εύκολος ο καθαρισμός τους. Αντίθετα, η προσθήκη ενός παράγοντα σύντηξης έχει το πλεονέκτημα ότι λειτουργεί ως ενισχυτής διαλυτότητας. Οι πιο δημοφιλείς ετικέτες σύντηξης είναι η πρωτεΐνη που δεσμεύει μαλτόζη (MBP), η θειορεδοξίνη (Trx), η S-τρανσφεράση της γλουταθειόνης (GST), η ουβικουιτίνη και η SUMO. Η ανίχνευση και ο καθαρισμός της χιμαιρικής πρωτεΐνης με χρωματογραφία συγγένειας είναι επίσης εφικτός. Οι ετικέτες μπορούν να αφαιρεθούν είτε ενζυμικά είτε χημικά για την βιοχημική ανάλυση της επιθυμητής πρωτεΐνης [126].

4.1.3.3 Ηλεκτροφόρηση σε πήκτωμα πολυακρυλαμίδης υπό αποδιατακτικές συνθήκες (SDS-PAGE)

Η ηλεκτροφόρηση είναι μια μέθοδος διαχωρισμού, οπτικοποίησης και απομόνωσης με εφαρμογές σε ποικίλα επιστημονικά πεδία. Η τεχνική αυτή είναι απλή, γρήγορη και χρησιμοποιείται, κυρίως, για τον προσδιορισμό του μεγέθους και της καθαρότητας ενός δείγματος, όπως επίσης, και για τον διαχωρισμό μιγμάτων μορίων, τα οποία δεν είναι δυνατόν να διαχωριστούν με άλλες τεχνικές. Δείγμα μπορεί να αποτελέσει οποιοδήποτε μόριο, το οποίο φέρει φορτίο. Η μέθοδος βασίζεται στην αρχή μετανάστευσης φορτισμένων μορίων κάτω από την επίδραση ενός εξωτερικά εφαρμοζόμενου ηλεκτρικού πεδίου. Η ηλεκτροστατική δύναμη που αναπτύσσεται κατευθύνει τα φορτισμένα μόρια προς το ηλεκτρόδιο του αντίθετου φορτίου. Εξαιτίας των διαφορετικών φορτίων και μαζών, τα διάφορα μόρια κινούνται με διαφορετικές ταχύτητες.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

4.1.3.4 Ανοσοαποτύπωση Πρωτεϊνών (Western Blot)

Το western blotting (γνωστό και ως ανοσοαποτύπωση πρωτεϊνών) χρησιμοποιείται για την ανίχνευση, τον ποσοτικό προσδιορισμό, καθώς και τον προσδιορισμό του μεγέθους των επιθυμητών πρωτεϊνών. Στα τέλη της δεκαετίας του 1970, ο Towbin και οι συνεργάτες του επέτρεψε τον διαχωρισμό των πρωτεϊνών μέσω ηλεκτροφόρησης σε πήκτωμα πολυακρυλαμίδης και ουρίας και τη μεταφορά τους σε μία μεμβράνη νιτροκυτταρίνης, και το 1981, ο Burnette εφάρμοσε την τεχνική αυτή χρησιμοποιώντας σύστημα ηλεκτροφόρησης υπό αποδιατακτικές συνθήκες (SDS-PAGE), καθιερώνοντας την τεχνική αυτή στο χώρο των μοριακών τεχνικών. Η ηλεκτροφόρηση σε πηκτή επιτρέπει το διαχωρισμό φυσικών ή μετουσιωμένων πρωτεϊνών βάσει μοριακού τους βάρους και, στη συνέχεια, μεταφέρονται σε μία μεμβράνη, συνήθως νιτροκυτταρίνης, και με τη χρήση ειδικών αντισωμάτων έναντι των πρωτεϊνών αυτών επιτρέπεται ο ποιοτικός και ποσοτικός προσδιορισμός της επιθυμητής πρωτεΐνης.

4.2 Βιοπληροφορική ανάλυση της πρωτεϊνικής επικράτειας CBM20 της πρωτεΐνης STBD1

Για τη βιοχημική μελέτη μίας πολυπετιδικής αλληλουχίας είναι απαραίτητη η βιοπληροφορική ανάλυση της και ο προσδιορισμος των χαρακτηριστικών της, όπως η θέση της στην πρωτεΐνη, ο χαρακτηρισμός των βιοχημικών της ιδιοτήτων (το θεωρητικό ισοηλεκτρικό σημείο και το μοριακό βάρος, την αμινική σύσταση, κλπ), την τριδιάστατη δομή και το βαθμό διαλυτότητας και παρουσίας ασταθών περιοχών σε αυτήν.

Ο καθορισμός της ακριβής θέσης της επικράτειας CBM20 στην πρωτεΐνη STBD1 ήταν αποτέλεσμα της ανάλυσης με τους διαδικτυακούς διακομιστές DOMAINATION και InterPro. Ο συνδυασμός και των δύο αυτών προγραμμάτων παρουσίασε ότι η θέση της πρωτεϊνικής αυτής επικράτειας εντοπίζεται στις θέσεις 258-358 της πρωτεϊνικής ακολουθίας της STBD1. Η επικράτεια αυτή χαρακτηρίστηκε για τις βιοχημικές της ιδιότητες με τη βοήθεια του προγράμματος ProtParam, δίνοντας στοιχεία για πρωτοταγή δομή της επικράτειας αυτής.

Σημαντικό σημείο για τη βιοχημική μελέτη της επικράτειας αποτελεί η πρόβλεψη της τριδιάστατης δομής της πρωτεϊνικής επικράτειας με τη βοήθεια του υπολογιστικού προγράμματος I-TASSER, το οποίο παρουσίασε στοιχεία τόσο για τη δευτεροταγή δομή, όσο και για την τριτοταγή δομή και την πιθανή λειτουργία της βάσει της δομής αυτής. Το πρόγραμμα αυτό επέτρεψε την παρουσία της σταθερότητας της επικράτειας στο χώρο, παρουσιάζοντας την συνολική ενέργεια και την κινητικότητα σε σχέση με τη δομή της. Το πρόγραμμα PONDR, σε συνδυασμό με το I-TASSER, επιτρέπει τον προσδιορισμό των ασταθών τμημάτων της επικράτειας, προσδιορίζοντας με τον τρόπο αυτό πιθανές θέσεις αλληλεπίδρασης με άλλα μόριο, πχ προσδέτες ή άλλα πρωτεϊνικά μόρια.

4.3 Κατασκευή πλασμιδιακών φορέων

Η πρωτεΐνη STBD1 εντοπίζεται κλωνοποιημένη σε πλασμιδιακό φορέα pGEX-6P-1, εκφράζοντας τη χιμαιρική πρωτεΐνη GST-STBD1. Η νουκλεοτιδική αλληλουχία της πρωτεΐνης σε αυτόν τον πλασμιδιακό φορέα είναι βελτιστοποιημένη, καθώς σε αυτήν τα σπάνια κωδικόνια έχουν αντικατασταθεί με τη άλλα πιο φιλικά για την *E. coli*. Η παραγόμενη ανασυνδυασμένη πρωτεΐνη δεν είναι διαλυτή in vivo, σχηματίζοντας έγκλειστα σωμάτια ύστερα από την ετερόλογη έκφρασή της σε σύστημα *E. coli*.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Προκειμένου να μελετηθεί η δομή και η λειτουργία της πρωτεϊνικής επικράτειας CBM20 καθίσταται αναγκαία η υποκλωνοποίηση της επικράτειας σε δύο διαφορετικούς πλασμιδιακούς φορείς, τον pETM-11 και τον pGEX-6P-1. Ο βασικός λόγος που επιλέχθηκαν αυτοί τα δύο πλασμίδια είναι ότι φέρουν διαφορετικές ετικέτες για την παραγόμενη ανασυνδυασμένη πρωτεΐνη, με τον pETM-11 να επιτρέπει το σχηματισμό χιμαιρικής πρωτεΐνης με μία ετικέτα έξι καταλοίπων ιστιδίνης και με τον pGEX-6P-1 να επιτρέπει το σχηματισμό με την ετικέτα GST. Για το πλασμίδιο pETM-11 επιλέγεται να γίνει κλωνοποίηση της επιθυμητής ακολουθίας ανάμεσα στις θέσεις αναγνώρισης από τα ένζυμα περιορισμού NcoI/XhoI, ενώ για το pGEX-6P-1 ανάμεσα στις θέσεις αναγνώρισης από τα ένζυμα περιορισμού BamHI/XhoI, σχηματίζοντας μονόκλωνα άκρα.

Τα ένζυμα περιορισμού επιλέχθηκαν έτσι, ώστε να μην μπορούν να εντοπίσουν κάποια θέση κοπής μέσα στην πρωτεΐνη STBD1, καθώς θα μπορούσε να εισαχθεί λάθος τμήμα του μέσα στον πλασμιδιακό φορέα. Επιπλέον, στο πλασμίδιο θα πρέπει να φέρει μόνο μία θέση κοπής και να εντοπίζεται στον πολυσυνδέτη του πλασμιδίου, ειδάλλως η κοπή σε δύο ή παραπάνω θέσεις θα είχε ως αποτέλεσμα την απώλεια των ιδιοτήτων του πλασμιδίου. Η επιλογή δύο ενζύμων περιορισμού επιτρέπει το σωστό προσανατολισμό του ενθέματος στο πλασμίδιο, ακολουθώντας το ανοιχτό πλαίσιο ανάγνωσης, που καθορίζεται από τον προσανατολισμό μεταγραφής βάσει του προαγωγέα που βρίσκεται μπροστά από την αλληλουχία και στον οποίο συνδέεται η RNA πολυμεράση και του πρώτου κωδικονίου που κωδικοποιεί για το αμινοξύ μεθειονίνη και βρίσκεται μπροστά από τη θέση του ενθέματος, επιτυγχάνοντας την έκφραση της ανασυνδυασμένης πρωτεΐνης. Κρίνεται αναγκαίο, επίσης, τα ένζυμα να μπορέσουννα λειτουργήσουν ως ζεύγη, ταυτόχρονα για την κοπή, τόσο του πλασμιδίου όσο και του ενθέματος, κάτω από κοινές συνθήκες θερμοκρασίας και ρυθμιστικού διαλύματος, αποδίδοντας τη μέγιστη αποτελεσματικότητα.

Το πλασμίδιο pETM-11 φέρει μία αλληλουχία που λειτουργεί ως ετικέτα, την MAD. Η ύπαρξη επιπλέον ετικέτας, πέραν της ετικέτας ιστιδίνης, μπορεί να φέρει προβλήματα τόσο στην σταθερότητα της ανασυνδυασμένης πρωτεΐνης όσο και στον μετέπειτα καθαρισμό αυτής. Έτσι, καθίσταται αναγκαία, η αφαίρεση της ετικέτας MAD από τον φορέα κλωνοποίησης pETM-11. Μπροστά από την αλληλουχία MAD εντοπίζεται μία θέση κοπής από το ένζυμο περιορισμού NcoI, το οποίο κόβει το πλασμίδιο μόνο μία φορά. Για το δεύτερο ένζυμο περιορισμού χρησιμοποιήθηκε το XhoI που βρίσκεται στην περιοχή του πολυσυνδέτη, μετά την αλληλουχία του γονιδίου MAD, καλύπτοντας τα παραπάνω κριτήρια. Για τον πλασμιδιακό φορέα pGEX-6P-1 χρησιμοποιήθηκαν δύο ένζυμα περιορισμού που καλύπτουν τα παραπάνω κριτήρια και οι θέσεις κοπείς τους εντοπίζονται στον πολυσυνδέτη του πλασμιδίου.

Η επικράτεια CBM20 της STBD1 ενισχύθηκε με τη χρήση της PCR, χρησιμοποιώντας εκκινητές οι οποίοι θα προσδώσουν στην τελική αλληλουχία προεξέχοντα άκρα. Οι προεκτάσεις φέρουν θέσεις κοπής από ένζυμα περιορισμού, προκειμένου να μπορέσει το ένθεμα να εισαχθεί στους πλασμιδιακούς φορείς με τη βοήθεια της λιγάσης T4. Για κάθε πλασμίδιο κατασκευάστηκαν διαφορετικοί ζεύγη εκκινητών, καθώς θα πρέπει να κοπούν τα άκρα τους από διαφορετικά ένζυμα περιορισμού για κάθε διαφορετικό πλασμίδιο.

Επιπλέον οι εκκινητές σχεδιάστηκαν έτσι, ώστε να πληρούν συγκεκριμένα κριτήρια:

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

- Θα έχουν μήκος 19-25 νουκλεοτιδίων. Το μήκος θα είναι τόσο μεγάλο ώστε να ικανοποιεί την ειδικότητα στο εκμαγείο και τόσο μικρό ώστε να μπορούν να συνδεθούν σε αυτό με σχετική ευκολία.
- Η θερμοκρασία αποδιάταξης (T_m) υποδηλώνει τη σταθερότητα του DNA. Εκκινητές με T_m εύρους 52-58ºC έχουν γενικά θετικά αποτελέσματα στη διαδικασία της PCR, ενώ όταν η τιμή ξεπερνά τους 65 ºC έχουν την τάση για σχηματισμό δευτεροταγών δομών. Για την πρόβλεψη της θερμακρασίας αποδιάταξης χρησιμοποιούνται εξισώσεις, οι οποίες βασίζονται στη θερμοδυναμική.
- Η θερμοκρασία υβριδοποίησης (T_a) επιτρέπει την υβριδοποίηση δύο συμπληρωματικών μονόκλωνων αλυσίδων. Η αποδιάταξη των κλώνων μπορεί να καταστρέψει όλους του δεσμούς υδρογόνου που συγκροτούν το δίκλωνο DNA, ενώ η θερμοκρασία υβριδοποίησης επιτρέπει τον επανασχηματισμό δεσμών υδρογόνου ανάμεσα σε συμπληρωματικές αλληλουχίες. Πολύ υψηλές θερμοκρασίες T_a μπορεί να οδηγήσει στη μη επιτυχημένη σύνδεση εκκινητή και εκμαγείου, όπως και να προκύψουν πολύ μικρές ποσότητες προϊόντων PCR. Αντίθετα, πολύ χαμηλές θερμοκρασίες T_a οδηγούν σε μη ειδική σύνδεση εκκινητή και εκμαγείου, που προκύπτει από κακό συνδυασμό ζευγών βάσεων σε υψηλό ποσοστό.
- Το περιεχόμενο των εκκινητών σε νουκλεοτίδια με γουανίνη και κυτοσίνη θα πρέπει να βρίσκεται σε εύρος 40-60%.
- Το ποσοστό των νουκλεοτιδιών με βάσεις γουανίνης και κυτοσίνης στις τελευταίες πέντε βάσεις στο 3’-άκρο των εκκινητών επηρεάζει την ειδική σύνδεση λόγω του μεγαλύτερου αριθμού δεσμών υδρογόνου που σχηματίζουν με τα συμπληρωματικά νουκλεοτίδια. Θεωρείται ότι δε θα πρέπει να υπάρχουν περισσότερα από τρία νουκλεοτίδια γουανίνης και κυτοσίνης.
- Η παρουσία των δευτεροταγών δομών στους εκκινητές είναι αποτέλεσμα διαμοριακών και ενδο μοριακών αλληλεπιδράσεων και οδηγούν σε χαμηλή ή /και σε καθόλου απόδοση προϊόντος. Μπορούν να επηρεάσουν δυσμενώς τη θερμοκρασία υβριδοποίησης των ακινήτων και κατ’ επέκταση την ενίσχυση το εκμαγείου, καθώς μειώνουν πολύ τη διαθεσιμότητα των εκκινητών στην αντίδραση. Κάποιες από αυτές τις δευτεροταγείς δομές των εκκινητών είναι:
    - Hairpins: Σχηματίζονται από ενδομοριακές αλληλεπιδράσεις μέσα στον εκκινητή και πρέπει να αποφεύγονται. Η σταθερότητα των hairpins συνήθως αναπαρίσταται από τις στιγμές ΔG, δηλαδή την ενέργεια που απαιτείται για να σπάσει η δευτεροταγής δομή. Πιο αρνητικές τιμές ΔG υποδηλώνει σταθερές, μη επιθυμητές hairpins. Ιδανικά, στο 3’-άκρο η hairpin θα μπορούσε να έχει ΔG=-2 kcal/mol, ενώ για ολόκληρο τον εκκινητή ΔG=-3 kcal/mol.
    - Self Dimer: ομο διμερή εκκινητών σχηματίζονται από διαμοριακές αλληλεπιδράσεις ανάμεσα σε δυο εκκινητές, όπου εκείνη της είναι ο μονόλογος στον εαυτό του. Γενικά, όταν ένας εκκινητής σχηματίζει διαμοριακά διμερή πιο συχνά απ’ ότι υβρίζεται με το DNA-στόχο ακόμα τότε μειώνεται η απόδοση της αντίδρασης. Ιδανικά, διμερή εκκιντών στο 3’-άκρο θα μπορούσε να έχουν ΔG=-5 kcal/mol, ενώ για εσωτερικά διμερή εκκινητών, το ΔG=-6 kcal/mol.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

- Cross Dimer: τα cross dimers των εκκινητών σχηματίζονται από την ίδια μοριακή αλληλεπίδραση ανάμεσα σε ζεύγη εκκινητών, όπου παρουσιάζουν ομολογία. Ιδανικά, διμερή εκκιντών στο 3'-άκρο θα μπορούσε να έχουν ΔG=-5 kcal/mol, ενώ για εσωτερικά διμερή εκκινητών, το ΔG=-6 kcal/mol.

- Οι επαναλήψεις δινουκλεοτιδίων που υπαρχουν πολλες φορες συνεχομενα στην αλληλουχία των εκκινητών θα πρέπει να αποφεύγεται κώμα καθώς θα μπορεί να ξεκινήσει λανθασμένα την αντίδραση PCR. Ο μέγιστος αριθμός επαναλήψεων δινουκλεοτιδίων σε ένα ολιγονουκλεοτίδιο δεν θα πρέπει να ξεπερνά τα 4 δινουκλεοτίδια.
- Μονές αλυσίδες αλληλουχιών νουκλεϊκών οξέων είναι αρκετά ασταθής και διπλώνονται σε διάφορες διαμορφώσεις τις δευτεροταγείς δομές. Η σταθερότητα αυτόν τον δεύτερο ταγών δομών εξαρτάται κυρίως στην ελεύθερη τους ενέργεια και στις τιμές της θερμοκρασίας αποδιάταξης. Ο σχεδιασμός των εκκινητών σε τέτοιες διαμορφώσεις των εκμαγείων δεν θα επιτρέψει τη σύνδεσή τους σε αυτό.
- Οι εκκινητές δεν θα πρέπει να ενισχύουν άλλα γονίδια στο δείγμα.
- Οι οι εκκινητές θα πρέπει να έχουν κοντινές θερμοκρασίες T_m.

Για το πλασμίδιο pETM-11, ο εμπρόσθιος εκκινητής φέρει θέση κοπής για το NcoI και ο οπίσθιος για το XhoI, ενώ για το πλασμίδιο pGEX-6P-1, ο εμπρόσθιος εκκινητής φέρει θέση κοπής για το BamHI και ο οπίσθιος για το XhoI. άλγη η αλληλουχίες των εκκινητών βρίσκονται στο παράρτημα III.

Αρχικά, πραγματοποιούνται δύο αντιδράσεις PCR για την ενίσχυση της επικρατείας CBM20 από το πλασμίδιο STBD1- pGEX-6P-1 για τα δύο ζεύγη εκκινητών ξεχωριστά. Τα συστατικά της αναγράφονται στον παρακάτω πίνακα:

Πίνακας 1 Συστατικά της αντίδρασης PCR ενίσχυσης του τμήματος CBM20 της STBD1 με διαφορετικά ζεύγη εκκινητών.

	Αντίδραση για κάθε ζεύγος αντίδρασης
Kapa High Fidelity HotStart	25.0 μL
Εμπρόσθιος εκκινητής (10 μM)	1.5 μL
Οπίσθιος εκκινητής (10 μM)	1.5 μL
STBD1- pGEX-6P-1	2.0 μL
ddH_2O	50.0 μL

Το θερμικό πρότυπο της αντίδρασης αναφέρεται στον επόμενο πίνακα:

Πίνακας 2 Το θερμικό πρότυπο της αντίδρασης PCR

	Θερμοκρασία	**Χρόνος**	
Αρχική Αποδιάταξη	95^oC	3 min	
Αποδιάταξη	95 ^oC	20 sec	
Υβριδοποίηση	60 ^oC	15 sec	30 κύκλους
Επέκταση	72 ^oC	1 min	

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Η ενίσχυση επιβεβαιώθηκε με ηλεκτροφόρηση σε πηκτή αγαρόζης. Ακολουθεί ο καθαρισμός των προϊόντων PCR και ο υπολογισμός της συγκέντρωσης και της καθαρότητάς τους. Για την ενισχυμένη επικράτεια CBM20 που φέρει στα άκρα της τις θέσεις κοπής των NcoI/XhoI, συγκέντρωσή της είναι 153.8 ng/μL με λόγο καθαρότητας (260/280) 1.78, ενώ την επικράτεια CBM20 που φέρει στα άκρα της τις θέσεις κοπής των BamHI/XhoI, με συγκέντρωση 77.3 ng/μL με βαθμό καθαρότητας 1,80.

Τα πλασμίδια pETM-11 και pGEX-6P-1 όπως και τα ενισχυμένα ενθέματα πέπτοναι με τα αντίστοιχα ένζυμα περιορισμού, όπως φαίνονται στους παρακάτω πίνακες, αντίστοιχα:

Πίνακας 3 Συστατικά της πέψης τόσο των δύο διαφορετικών πλασμιδίων όσο και των ενθεμάτων από τα ένζυμα περιορισμού.

	Πέψη με τα ένζυμα περιορισμού NcoI-HF και XhoI		**Πέψη με τα ένζυμα περιορισμού BamHI και XhoI**	
	pETM-11	**CBM20**	**pGEX-6P-1**	**CBM20**
DNA (2μgr)	170.0 μL	13.0 μL	10.0 μL	25.0 μL
rCutSmart Buffer (10X)	10.0 μL	10.0 μL	10.0 μL	10.0 μL
Ένζυμο περιορισμού NcoI-HF	2.0 μL	2.0 μL	2.0 μL	2.0 μL
Ένζυμο περιορισμού XhoI	2.0 μL	2.0 μL	2.0 μL	2.0 μL
ddH_2O	69.0 μL	73.0 μL	75.0 μL	61.0 μL

Η διαδικασία της πέψης πραγματοποιήθηκε στους 37 ^oC για δύο ώρες.

Τα προϊόντα της πέψης ηλεκτροφορούνται σε πηκτή αγαρόζης και τα επιθυμητά τμήματα (bands) απομονώνονται με τη βοήθεια ενός agarose gel band cutter. Τα δείγματα καθαρίζονται και υπολογίζεται η συγκέντρωσή τους και ο βαθμός καθαρότητάς τους.

Πίνακας 4 Οι συγκεντρώσεις και ο βαθμός καθαρισμού των προϊόντων της πέψης των πλασμιδίων και των ενθεμάτων από τα ένζυμα περιορισμού.

	pETM-11 NcoI/XhoI	**CBM20 NcoI/XhoI**	**pGEX-6P-1 BamHI/XhoI**	**CBM20 BamHI/XhoI**
Συγκέντρωση	42.0 ng/μL	32.0 ng/μL	31.9 ng/μL	35.07 ng/μL
260/280	1.90	1.90	1.81	1.74

Ύστερα από την πέψη, το πλασμίδιο pETM-11 έχει μέγεθος 5378 νουκλεοτιδικές βάσεις ενώ το ένθεμα CBM20 NcoI/XhoI 317 νουκλεοτιδικές βάσεις. Ενώ, το πλασμίδιο pGEX-6P-1 έχει μέγεθος 4960 νουκλεοτιδικές βάσεις ύστερα από την πέψη του, ενώ το ένθεμα CBM20 BamHI/XhoI έχει 321 νουκλεοτιδικές βάσεις. Για τον υπολογισμό της απαραίτητης ποσότητας ενθέματος που θα προστεθεί για κάθε 100 ng πλασμιδιακού φορέα για κάθε διαφορετικό μοριακό λόγο χρησιμοποιείται ο παρακάτω τύπος:

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

$$Μάζα\ ενθέματος\ (ng) = Μοριακός\ λόγος \times Μάζα\ φορέα\ (ng) \times \frac{Μήκος\ αλληλουχίας\ ενθέματος}{Μήκος\ αλληλουχίας\ πλασμιδίου}$$

Στον παρακάτω πίνακα παρουσιάζονται οι ποσότητες του ενθέματος και του πλασμιδίου που χρησιμοποιήθηκαν για την ένωσή τους από την λιγάση T4 για κάθε διαφορετικό μοριακό λόγο. Η T4 λιγάση επιτρέπει τη σύνδεση των τμημάτων στους 16 ^{o}C για 16 ώρες.

Πίνακας 5 Συστατικά της σύνδεση των πλασμιδίων με τα αντίστοιχα ενθέματα με τη βοήθεια της T4 DNA λιγάσης.

	pETM-11 και CBM20, πέψη με NcoI/XhoI			**pGEX-6P-1 και CBM20, πέψη με BamHI/XhoI**	
	2:1	**3:1**	**5:1**	**3:1**	**5:1**
Ρυθμιστικό διάλυμα λιγάσης (10X)	2.0 μL	2.0 μL	2.0 μL	2.0 μL	2.0 μL
Κομμένος φορέας	2.3 μL	2.3 μL	2.3 μL	3.0 μL	3.0 μL
Κομμένο ένθεμα	0.35 μL	0.52 μL	0.86 μL	0.6 μL	0.9 μL
T4 DNA λιγάση	1.0 μL	1.0 μL	1.0 μL	1.0 μL	1.0 μL
ddH_2O	14.35 μL	14.18 μL	13.84 μL	13.4 μL	13.1 μL

Η επιτυχία της κλωνοποίησης επιβεβαιώθηκε με ηλεκτροφόρηση σε πηκτή αγαρόζης και με αλληλούχιση (Παράρτημα IV).

4.4 Χημικός μετασχηματισμός βακτηριακών κυττάρων

Η χρήση της πολυαιθυλενικής γλυκόζης (PEG) και του χλωριούχου ασβεστίου ($CaCl_2$) αποτελούν τις πιο γνωστές μεθόδους χημικού μετασχηματισμού. Βακτηριακά κύτταρα τα οποία είχαν επεξεργαστεί με χλωριούχο ασβέστιο, ώστε να είναι δεκτικά σε εισαγωγή πλασμιδίου, επωάστηκαν με το επιθυμητό πλασμίδιο στον πάγο για περίπου 20 λεπτά, στη συνέχεια για 45 δευτερόλεπτα σε θερμοκρασία 42°C και έπειτα ξανά σε πάγο. Αυτή η απότομη διαδοχή στην αλλαγή της θερμοκρασίας οδηγεί σε εκπόλωση της βακτηριακής μεμβράνης των δεκτικών κυττάρων ώστε το πλασμίδιο να εισέλθει εντός του κυττάρου. Έπειτα, η μεταφορά των κυττάρων σε πάγο οδηγεί σε επαναφορά της κυτταρικής μεμβράνης στην αρχική της κατάσταση, ώστε το πλασμίδιο που εισήχθη να εγκλωβιστεί στο εσωτερικό του βακτηρίου. Ακολούθως, προστέθηκαν 900μL υγρού θρεπτικού υλικού και τα κύτταρα αναπτύχθηκαν σε θερμοκρασία 37 ^{o}C σε επωαστήρα υπό ανακίνηση (210 rpm) για μια ώρα. Τέλος, ποσότητα κυττάρων επιστρώθηκε σε τρυβλία Petri σε στερεό θρεπτικό υλικό παρουσία αντιβιοτικού για επιλογή των βακτηρίων στα οποία εισήχθη το πλασμίδιο. Αποικίες αναπτύχθηκαν έπειτα από 16 ώρες επώασης σε θερμοκρασία 37°C.

4.5 Δοκιμές έκφρασης και παραγωγής της πρωτεϊνικής επικράτειας CMB20

Για τον έλεγχο της έκφρασης της επικράτειας χρησιμοποιήθηκε το βακτηριακό στέλεχος *E. coli BL21Gold*. Τα γονίδια εισήχθησαν μέσω χημικού μετασχηματισμού και ποσότητα των κυττάρων επιστρώθηκε σε τρθβλία Petri με στερεό θρεπτικό υλικό LB-agar παρουσία 30 μg/mL καναμυκίνης για το πλασμίδιο pETM-11 και 50 μg/mL αμπικιλλίνης για το πλασμίδιο pGEX-6P-1.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Έπειτα ποσότητα κυττάρων μεταφέρθηκε σε υγρό θρεπτικό υλικό σε αναλογία 1:100 παρουσία των ίδιων συγκεντρώσεων αντιβιοτικών. Τα κύτταρα αναπτύχθηκαν σε θερμοκρασία 37°C σε επωαστήρα που ανακίνηση 210 rpm έως ότου η οπτική πυκνότητα της καλλιέργειας σε μήκος κύματος 600 nm έφτασε την τιμή περίπου 0,6 απορρόφησης. Η καλλιέργεια χωρίστηκε σε αποστειρωμένες υάλινες κωνικές φιάλες (10 mL σε κάθε κωνική φιάλη) και ακολούθησε προσθήκη IPTG σε συγκέντρωση 1 mM και επαγωγή της υπερέκφρασης σε θερμοκρασία 37°C 4 ώρες σε επωαστήρα υπό ανακίνηση (210 rpm). Υπερέκφραση έγινε επίσης σε θερμοκρασία 16 βαθμών κελσίου για 16 ώρες σε επωαστήρα υπό ανακίνηση (210 rpm). Σε αυτή την περίπτωση η καλλιέργεια τοποθετήθηκε πρώτα σε πάγο και έπειτα έγινε η προσθήκη του IPTG σε συγκέντρωση 0.1 mM. Σε κάθε περίπτωση, τα κύτταρα συλλέχθηκαν με φυγοκέντρηση μετά το πέρας της διάρκειας της επαγωγής της υπερέκφρασης (6000 rpm, 10 λεπτά, 4 °C). Το ίζημα από τα 10 γκολ καλλιέργειας διαλυτό ποιήθηκε σε ένα ml διαλύματος λύσης το οποίο περιείχε 20 mM Hepes, pH 7,4. Η λύση των κυττάρων έγινε με υπερήχους [(25% amplitude, 10 sec εφαρμογή/ 10 sec παύση) για δύο φορές, 4 ^{o}C] και το διαλυτό κλάσμα συλλέχθηκε με φυγοκέντρηση (13000 rpm, 30 λεπτά, 4 °C). Συλλέχθηκαν δείγματα πριν και μετά τη φυγοκέντρηση, δηλαδή το ολικό κύτταρο λήμμα αλλά και το διαλυτό κλάσμα. από τα δείγματα αυτά 30 μικρών γραμμάρια πρωτεϊνών αναλύθηκαν σε SDS-PAGE. Η παρουσία της πρωτεΐνης στο διαλυτό κλάσμα δηλώνει την παραγωγή διαλυτής πρωτεΐνης, ενώ η παρουσία της μόνο στο υλικό κυτταρόλυμα δηλώνει ότι η πρωτεΐνη εντοπίζεται στα έγκλειστα σωμάτια.

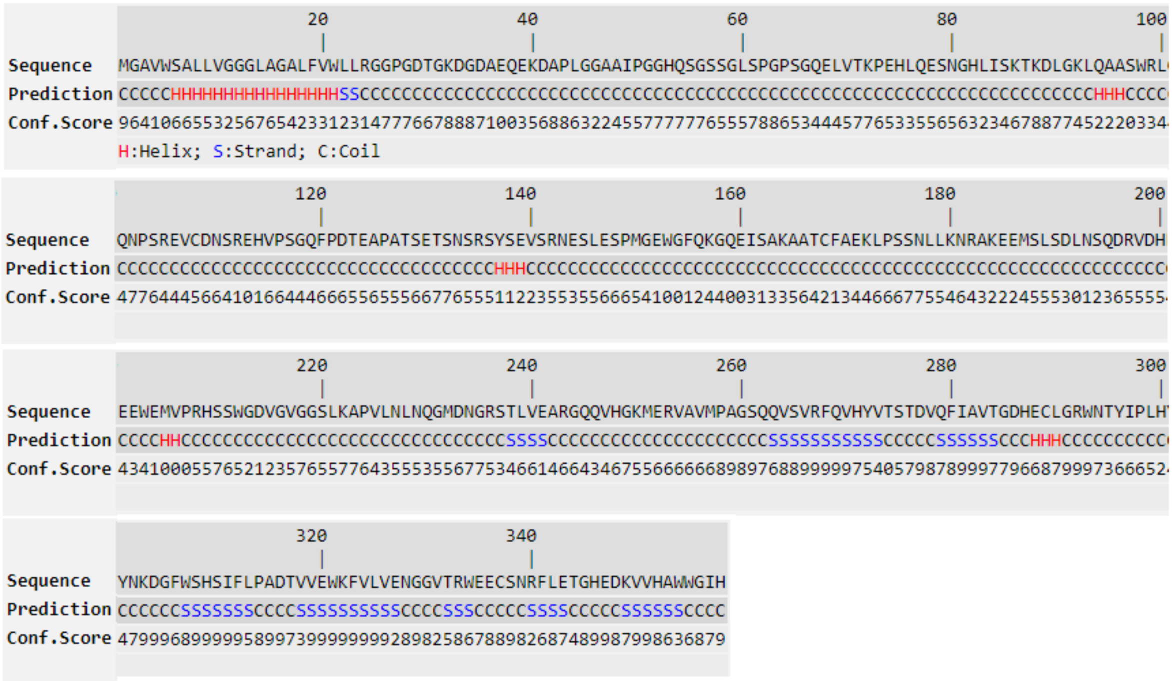
Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

5 ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ

5.1 <u>Βιοπληροφορική ανάλυση</u>

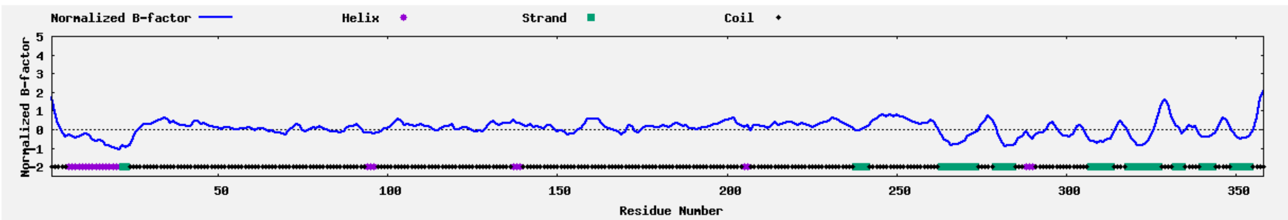
5.1.1 Πρόβλεψη και ανάλυση δομής STBD1

5.1.1.1 Δευτεροταγής δομή και αστάθεια περιοχών



Εικόνα 9 Πρόβλεψη της δευτεροταγούς δομής της STBD1 μέσω του προγράμματος I-TASSER

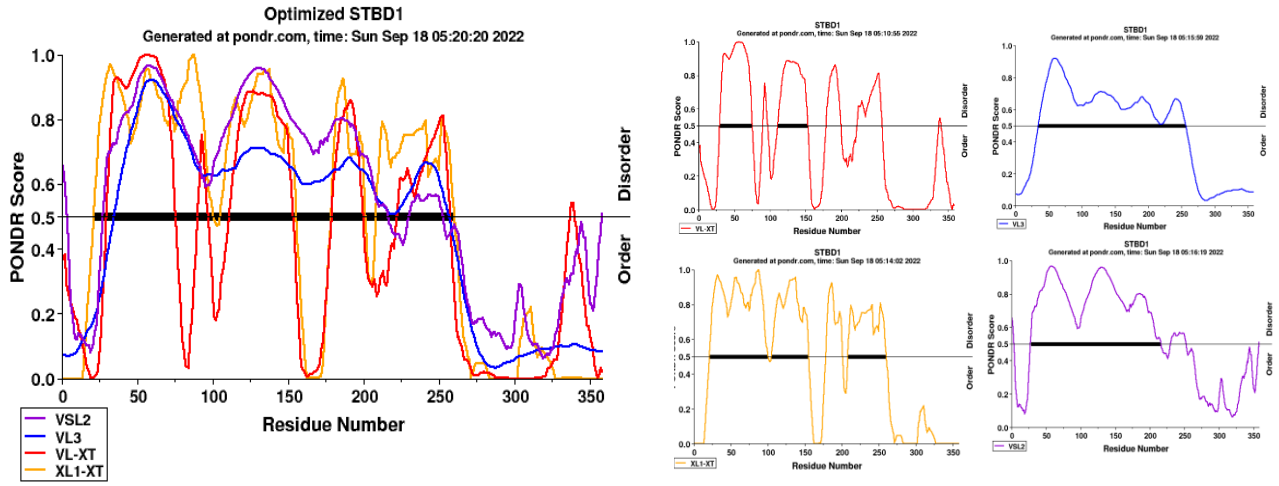
Η δευτεροταγής δομή της STBD1, σύμφωνα με το I-TASSER, φαίνεται να αποτελείται στο καρβοξυτελικό της άκρο από β-πτυχωτές επιφάνειες (263-355 αα), ενώ είναι πολύ καλά διακριτή μία α-έλικα στο αμινοτελικό της άκρο (6-21 αα). Ωστόσο, η υπόλοιπη διαμόρφωση της δεν έχει καλά καθορισμένη δομή και αναφέρεται ως σπειράματα. Επιπλέον, ο κανονικοποιημένος B-factor φαίνεται να έχει μεγάλη διακύμανση, με τιμές μεγαλύτερες του μηδενός σχεδόν καθόλο το μήκος της ακολουθίας. Αυτό αποδεικνύει και τη αστάθεια των περιοχών αυτών ή /και το μη εντοπισμό αντίστοιχων δομών από πειραματικές μελέτες άλλων συγγενών με την αλληλουχία πρωτεϊνών στόχων αναζήτησης. Παρόλα αυτά, το καρβοξυτελικό άκρα φαίνεται να φέρει μία καλά διατηρημένη δομή, με τις τιμές του B-factor που είναι μεγαλύτερες του μηδενός να αναφέρονται σε περιοχές αλληλεπίδρασης με άλλες πρωτεΐνες ή να συνδέονται με προσδέτες.



Εικόνα 10 Ο κανονικοποιημένος B-factor που προκύπτει από την ανάλυση της πρωτεΐνης STBD1 από το πρόγραμμα I-TASSER

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

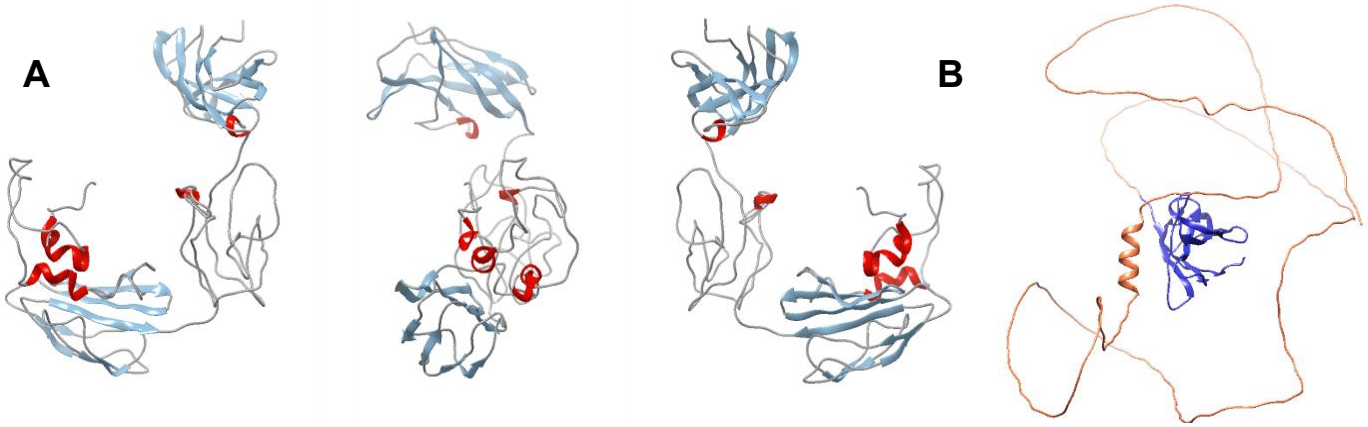
Το μήκος της αστάθειας της STBD1 είναι εμφανές και από το διακομιστή PONDR, όπου συγκρίνεται κυρίως ο βαθμός αστάθειας με αποτελέσματα κρυσταλλογραφίας και NMR. Στα διαγράμματα φαίνεται ότι σχεδόν όλη η επιφάνεια της πρωτεΐνης είναι ασταθής, εκτός από το καρβοξυτελικό της άκρο. Παρόλα αυτά διακρίνεται μία σταθερή διαμόρφωση στην περιοχή 150-175 αα, χωρίς όμως να πληροί όλα τα κριτήρια για να μπορεί να καθοριστεί ως σταθερή.



Εικόνα 11 Παρουσία των ασταθών περιοχών της πρωτεΐνης STBD1 από το πρόγραμμα PONDR. Σε αυτό παρουσιάζονται στοιχεία από τέσσερις διαφορετικούς λογαρίθμους, που λαμβάνουν υπόψιν τους κρυσταλλογραφικές δομές και δομές NMR.

5.1.1.2 Τριτοταγής δομή

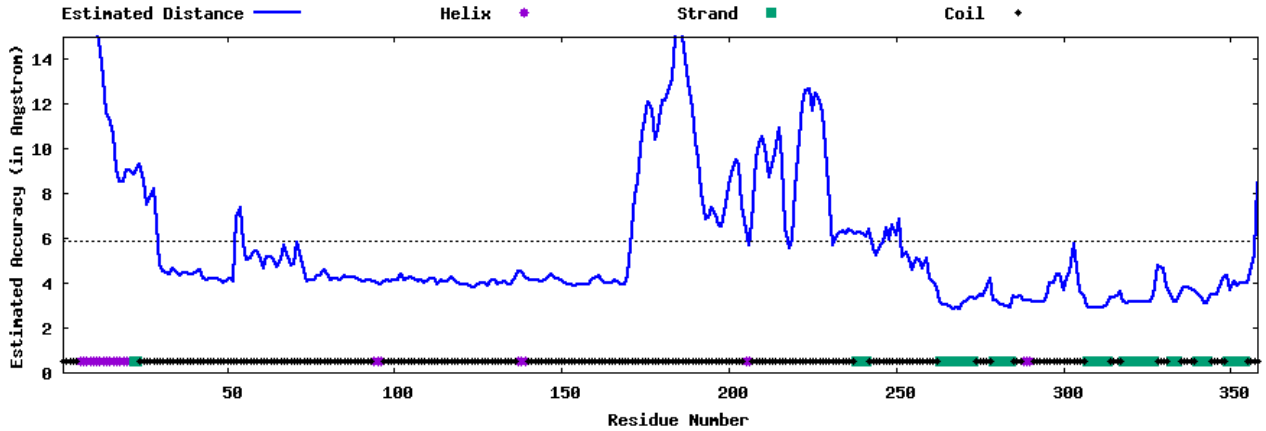
Από τα πέντε μοντέλα που προέκυψαν από το πρόγραμμα I-TASSER, το μοντέλο αυτό αποτελεί το επικρατέστερο της πρόβλεψης. Ο βαθμός εμπιστοσύνης (C-score) είναι -1.79, γεγονός που υποδηλώνει μια μη σωστή τοπολογία της εκτιμώμενης δομής της STBD1. Το TM-score είναι 0.5 (με βαθμό απόκλισης 0.15) και το RMSD είναι 10.7 Å (με βαθμό απόκλισης 4.6 Å), δηλαδή η τοπική απόκλιση είναι μεγάλη, ωστόσο έχει μεγάλο βαθμό αξιοπιστίας της ανάλυσης για τα τμήματα στα οποία έχουν δοθεί καθορισμένες δομές. Για το μοντέλο αυτό χρησιμοποιήθηκε κυρίως ως δομή εκμαγείο η γλυκοτρανσφεράση της κυκλοδεξτρίνης (CGTase) του *Geobacillus stearothermophilus* (PDB: 1CYG).



Εικόνα 12 A. Πρόβλεψη της τριτοταγούς δομής της STBD1 από το πρόγραμμα I-TASSER. Τα διαφορετικά χρώματα χαρακτηρίζουν τον τρόπο αναδίπλωσης των τμημάτων, β-πτυχώσεις (μπλε), α-έλικα (κόκκινο) B. Πρόβλεψη της τριτοταγούς δομής της STBD1 από το πρόγραμμα AlphaFold. Οι διαφορές στα χρώματα χαρακτηρίζουν το βαθμό εμπιστοσύνης ανά κατάλοιπο περιοχών. Με μπλε εμφανίζονται οι περιοχές, όπου pLDDT >90, ενώ με πορτοκαλί οι περιοχές όπου pLDDT <50. Οι δομές παρουσιάζονται με τη βοήθεια του προγράμματος UCSF Chimera.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Το N-αμινοτελικό άκρο της φαίνεται να φέρει δύο α-έλικες κοντά στο αμινοτελικό της άκρο (^{9}LVGGGLAGALFV^{20} και ^{50}PGGHQSGS^{57}) και επτά β-πτυχώσεις, που σχηματίζουν μία δομή βαρελιού, που αποτελείται από τρεις β-πτυχωτές επιφάνειες, με δύο ζεύγη αντιπαράλληλων β-πτυχώσεων να βρίσκονται απέναντι από τέσσερις αντιπαράλληλες β-πτυχώσεις. Πέραν από μίας περιοχής τεσσάρων αμινικών καταλοίπων (^{170}AEKL^{174}) που σχηματίζουν μία μη ολοκληρωμένη α-έλικα, το μεγαλύτερο τμήμα της πρωτεΐνης δε φαίνεται να έχει συγκεκριμένη δομή, εκτός από το C-τελικό της άκρο που αποτελεί και τη βασική δομική της επικράτεια (CBM20) και φαίνεται να είναι εξελικτικά καλά συντηρημένη. Το καρβοξυτελικό άκρο φαίνεται να είναι μία αναδίπλωση β-σάντουιτς με επτά β-πτυχώσεις χωρισμένες σε δύο β-πτυχωτές επιφάνειες. Η μία β-πτυχωτή επιφάνεια έχει τρεις αντιπαράλληλες β-πτυχώσεις και μία συνεχόμενη της άλλης, ενώ η δεύτερη αποτελείται από ένα παράλληλο και ένα αντιπαράλληλο ζεύγος β-πτυχώσεων. Η αναδίπλωση αυτή σχηματίζει ένα παραμορφωμένο β-βαρέλι με ανοιχτές πλευρές που φέρει έξι βρόχους διαφορετικών μεγεθών.



Εικόνα 13 Μέγεθος σφάλματος της προκύπτουσας ανάλυσης της δομής STBD1 από το πρόγραμμα I-TASSER.

Βάσει της αξιολόγησης του σφάλματος της μεθόδου για το προτεινόμενο από το I-TASSER μοντέλο φαίνεται να μην είναι αρκετά αξιόπιστη η δομή. Ωστόσο, μόνο οι δομές του C-τελικού άκρου φαίνεται να είναι καλά δομημένες σε σύγκριση με την υπόλοιπη πρωτεΐνη.

Σε αντίθεση με το I-TASSER, το οποίο βασίζεται στο threading και στην αναζήτηση εκμαγείων γνωστών πρωτεϊνικών δομών για την πρόβλεψη της τριτοταγούς δομής μίας πρωτεΐνης επερώτησης, το AlphaFold είναι ένα πρόγραμμα υπολογιστικής νοημοσύνης, που χρησιμοποιεί πολύπλοκα νευρωνικά-στοχαστικά συστήματα (deep learning). Το πρόγραμμα αυτό ξεκινά χρησιμοποιώντας στοιχίσεις πολλαπλών ακολουθιών (multiple sequence alignments, MSA) με κάθε περιοχή να έχει διαφορετικό βαθμό σημασίας. Έπειτα, χρησιμοποιώντας το Evformer εξάγει πληροφορίες σχετικά με τις αλληλεπιδράσεις μεταξύ των πρωτεϊνικών ακολουθιών και των πρότυπων δομών. Οι δομικές επικράτειες αντιμετωπίζονται στην πρωτεΐνη ως κατάλοιπα που μετακινούνται στο χώρο, σχηματίζοντας την τριδιάστατη δομή της πρωτεΐνης που ακολουθείται από τοπικές βελτιώσεις, παρέχοντας την τελική πρόβλεψη [142,143]. Ένας τρόπος αξιολόγησης των τμημάτων της προβλεπόμενης δομής είναι ο βαθμός εμπιστοσύνης ανά κατάλοιπο, που ονομάζεται ως pLDDT, με εύρος 0-100. Η ακρίβεια του τμήματος της δομής είναι μεγαλύτερη όσο ο αριθμός pLDDT είναι

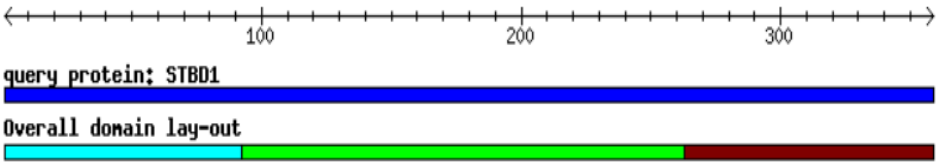
Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

υψηλότερος. Αντίθετα, σε περιπτώσεις χαμηλής ακρίβειας πρόβλεψης δομών τμημάτων της πρωτεΐνης (pLDDT <50), οι περιοχές έχουν μία απεικόνιση που μοιάζει με έλικα, ενώ μπορεί να αποτελεί ισχυρή ένδειξη ύπαρξης μίας ασταθούς περιοχής, η οποία μπορεί να μην έχει δομή σε φυσιολογικές συνθήκες ή να φέρει δομή ως μέρος ενός πρωτεϊνικού συμπλόκου.

Στην περίπτωση της STBD1, με το πρόγραμμα AlphaFold μόνο η επικράτεια CBM20 φέρει μία απεικόνιση της δομής της πρωτεΐνης, έχοντας και μεγάλο βαθμό αξιοπιστίας της δομής της επικράτειας (pLDDT >90). Αντίθετα, το υπόλοιπο τμήμα της δε φαίνεται να έχει κάποια διαμόρφωση στο χώρο, με τον αριθμό pLDDT να είναι μικρότερος του 50, υποδηλώνοντας την αστάθεια του τμήματος αυτού (Εικόνα 10Β).

5.1.1.3 Καθορισμός της CBM20 επικράτειας της STBD1

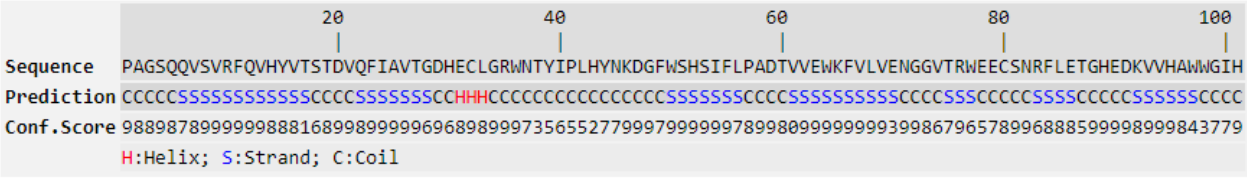
Η σταθερότητα της δομής και η σημασία του βιοχημικού ρόλου που έχει η επικράτεια CBM20 για το κύτταρο καθιστά απαραίτητη τη μελέτη της. Η θέση μέσα στην οποία βρίσκεται προκύπτει από την ανάλυση από τα εργαλεία DOMAINATION και InterPro. Βάσει του DOMAINATION, η επικράτεια αυτή φαίνεται να εντοπίζεται στις θέσεις 264-357. Ωστόσο, η ανάγνωση της αλληλουχίας από το InterPro από διαφορετικές βάσεις δεδομένων, παρουσιάζει ότι στη βάση δεδομένων της PROSITE, η επικράτεια εντοπίζεται στις θέσεις 258-357. Οι θέσεις 257-357 επιλέχθηκαν, έτσι, καθώς περιλαμβάνουν και τις δύο προβλέψεις των επικρατειών και από τους δύο διαφορετικούς διαδικτυακούς διακομιστές.



Εικόνα 14 Καθορισμος της επικράτειας CBM20 στην πρωεΐνη STBD1 με τη βοήθεια του διαδικτυακού διακομιστή DOMAINATION. Η επιθμητή επικράτεια συμβολίζεται με κόκκινο σκούρο χρώμα.

5.1.2 Πρόβλεψη και ανάλυση της επικράτειας CBM20 της STBD1

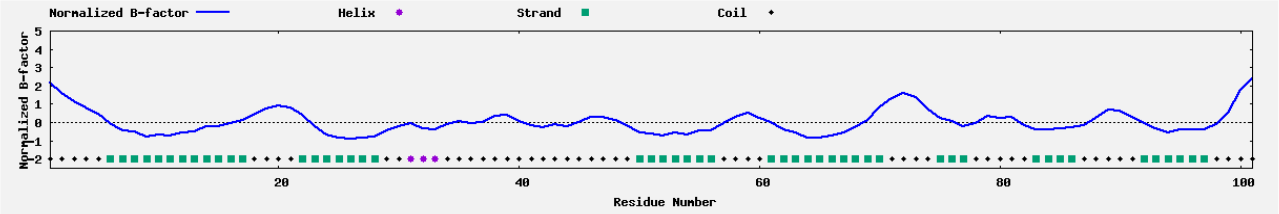
5.1.2.1 Δευτεροταγής δομή και αστάθεια περιοχών



Εικόνα 15 Πρόβλεψη της δευτεροταγούς δομής της CBM20 μέσω του προγράμματος I-TASSER

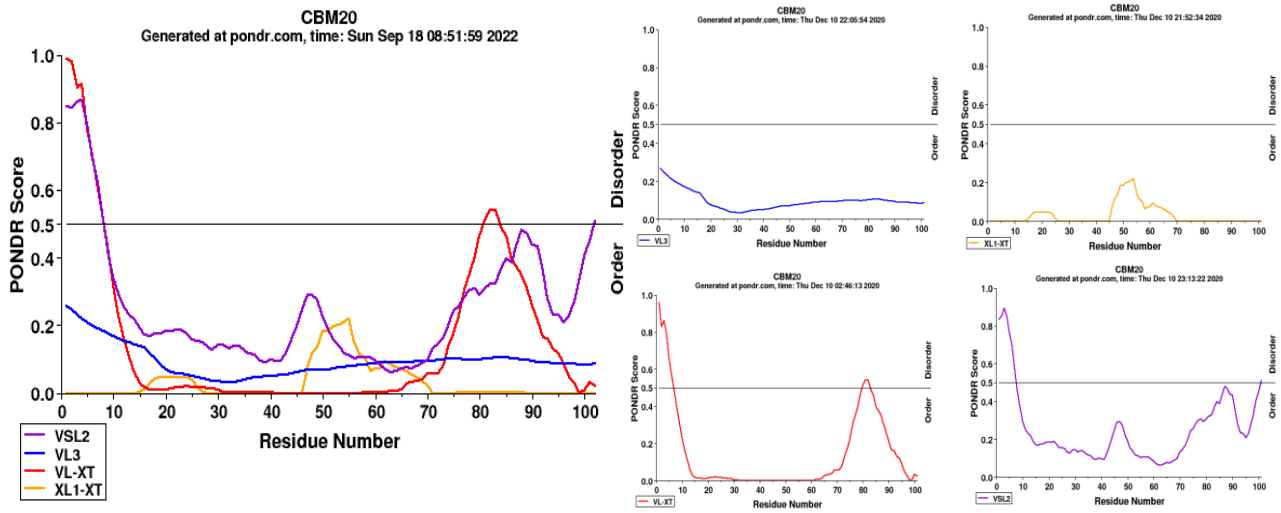
Βάσει της πρόβλεψης της δευτεροταγούς δομής της επικρατείας από το πρόγραμμα I-TASSER φαίνεται να αποτελείται από επτά β-πτυχώσεις και μία α-έλικα με μεγάλο βαθμό εμπιστοσύνης. Παράλληλα, παρατηρείται ότι και ο κανονικοποιημένος B-factor είναι σχετικά χαμηλός, με τις μη χαμηλές τιμές να μην υπερβαίνουν πολύ το μηδέν. Πιθανώς, στις θέσεις αυτές σχηματίζονται βρόγχοι που επιτρέπουν την ασταθή διαμόρφωσή τους, εικάζοντας ότι μπορεί να αποτελούν θέσεις αλληλεπίδρασης με άλλες πρωτεΐνες ή θέσεις σύνδεσης με άλλους προσδέτες.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84



Εικόνα 16 Ο κανονικοποιημένος B-factor που προκύπτει από την ανάλυση της επικράτειας CBM20 από το πρόγραμμα I-TASSER

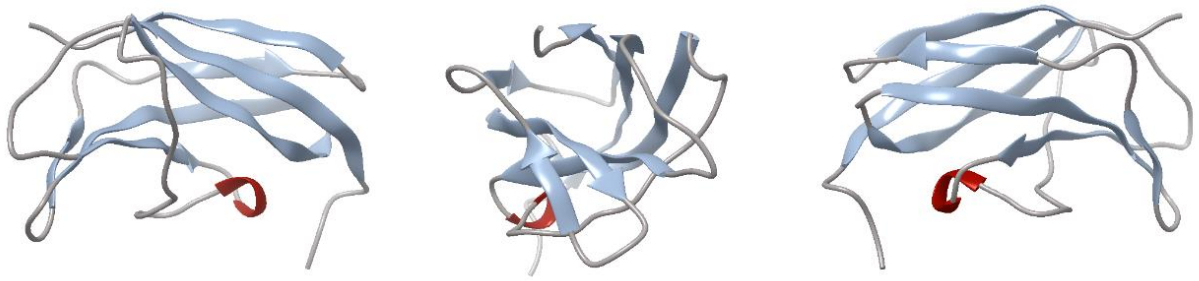
Τα αποτελέσματα αυτά επιβεβαιώνονται και από την ανάλυση με το εργαλείο PONDR. Σε αυτά τα διαγράμματα φαίνεται ότι η επιφάνεια της επικράτειας στο σύνολό της είναι αρκετά σταθερή, εκτός από ένα τμήμα κοντά στο καρβοξυτελικό της άκρο, το οποίο, όμως, και πάλι δεν ξεπερνά το κατώφλι για να θεωρηθεί ασταθής. Στην περίπτωση αυτή μπορεί, ακόμα μία φορά, να εννοηθεί ότι αποτελεί μία περιοχή αλληλεπίδρασης με κάποια πρωτεΐνη ή περιοχή σύνδεσης με κάποιον προσδέτη. Στην περιοχή αυτή, υπάρχει μία μικρή ταύτιση με το I-TASSER, όχι, όμως, σε κάποια άλλη.



Εικόνα 17 Παρουσία των ασταθών περιοχών της επικράτειας CBM20 από το πρόγραμμα PONDR. Σε αυτό παρουσιάζονται στοιχεία από τέσσερις διαφορετικούς λογαρίθμους, που λαμβάνουν υπόψιν τους κρυσταλλογραφικές δομές και δομές NMR.

5.1.2.2 Τριτοταγής δομή

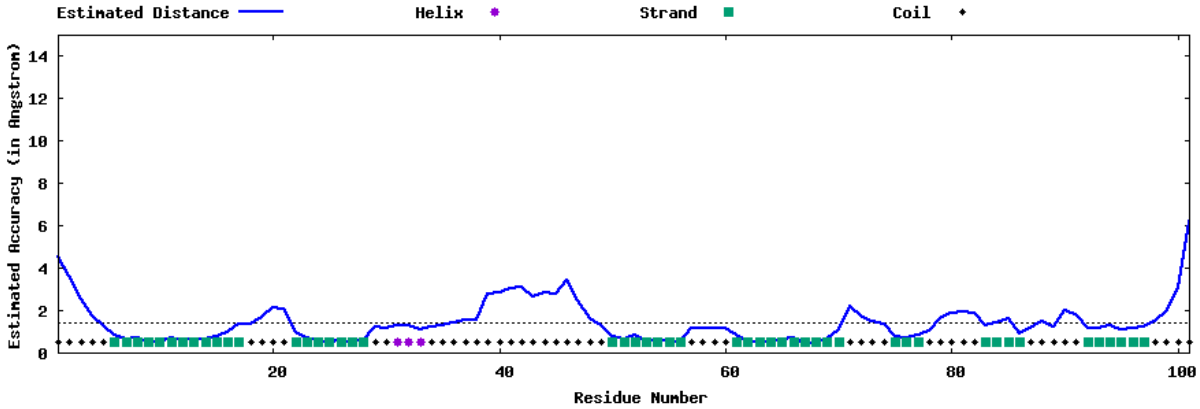
Για την επικράτεια CBM20 της STBD1, το I-TASSER δίνει μόνο μία διαμόρφωση. Ο βαθμός εμπιστοσύνης (C-score) είναι 0.65, αναγνωρίζοντας σε σημαντικό βαθμό τη σωστή τοπολογία της διαμόρφωσης που δίνεται. Το TM-score είναι 0.8 (με βαθμό απόκλισης 0.09) και το RMSD είναι 2.7 Å (με βαθμό απόκλισης



Εικόνα 18 Πρόβλεψη της τριτοταγούς δομής της CBM20 από το πρόγραμμα I-TASSER. Οι δομές παρουσιάζονται με τη βοήθεια του προγράμματος UCSF Chimera.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

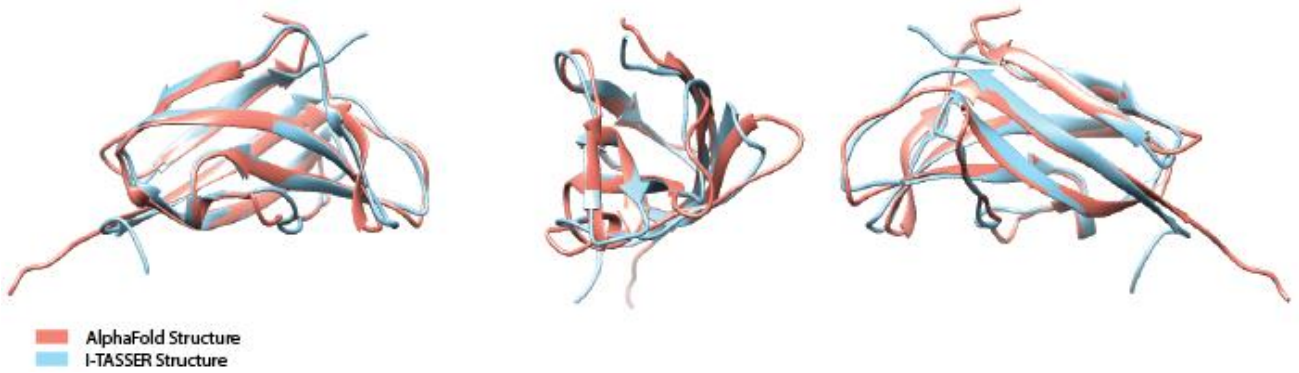
2.0 Å), υποδηλώνοντας με τον τρόπο αυτό την ακρίβεια της δομής με βάση ομόλογες άλλες δομές που έχουν ήδη αναγνωριστεί. Ο βαθμός σφάλματος επίσης είναι αμελητέος, εκτός από την περιοχή 38-45 αμινοξέων που φαίνεται να υπάρχει ένα μικρό σφάλμα στη διαμόρφωση και βάσει τη δευτεροταγή δομή φαίνεται να είναι ένα τυχαίο σπείραμα.



Εικόνα 19 Μέγεθος σφάλματος της προκύπτουσας ανάλυσης της δομής της επικράτειας CBM20 από το πρόγραμμα I-TASSER.

Όπως και στην πρόβλεψη της δομής της STBD1, η επικράτεια αυτή φαίνεται να είναι μία αναδίπλωση β-σάντουιτς με επτά β-πτυχώσεις χωρισμένες σε δύο β-πτυχωτές επιφάνειες. Η μία β-πτυχωτή επιφάνεια έχει τρεις αντιπαράλληλες β-πτυχώσεις και μία συνεχόμενη της άλλης, ενώ η δεύτερη αποτελείται από ένα παράλληλο και ένα αντιπαράλληλο ζεύγος β-πτυχώσεων. Η αναδίπλωση αυτή σχηματίζει ένα παραμορφωμένο β-βαρέλι με ανοιχτές πλευρές που φέρει έξι βρόχους διαφορετικών μεγεθών. Τέλος φαίνεται να σχηματίζεται και μια μικρή διαμόορφωση α-έλικας στις θέσεις ^{288}ECL^{290}.

Ενώ δεν υπάρχει σαφής δομή για το σύνολο της πρωτεΐνης STBD1, από το πρόγραμμα AlphaFold, καθίσταται σαφές ότι η δομή της CBM20 είναι καλά διαμορφωμένη. Εφαρμόζεται υπέρθεση των δύο προβλεπόμενων δομών από τα δύο διαφορετικά προγράμματα, I-TASSER και AlphaFold, προς σύγκριση τους. Με τη βοήθεια του προγράμματος UCSF Chimera, καθορίζονται οι παράμετροι της υπέρθεσης, δηλαδή ο αλγόριθμος στοίχισης είναι Needleman-Wunsch με μήτρα BLOSUM-62, στον τρόπο βαθμολόγησης της δομής συμπεριλαμβάνεται η δευτεροταγή δομή και σε ποσοστό 30% του βαθμού ομοιότητας των καταλοίπων, και τα ζεύγη των ατόμων των δύο υπερθετημένων δομών δεν θα υπερβαίνουν τα 2.0 Å. Από την παραπάνω υπέρθεση είναι σαφής η μεγάλη ομοιότητα στην τοπολογία των δ'υο προβλέψεων και η σε μεγαλύτερο βαθμό διαφορά τους στους βρόχους τους.



Εικόνα 20 Σύγκριση των δομών της CBM20 που προκύπτουν από τα προγράμματα I-TASSER (μπλε χρώμα) και AlphaFold (ροζ χρώμα) με το πρόγραμμα UCSF Chimera.

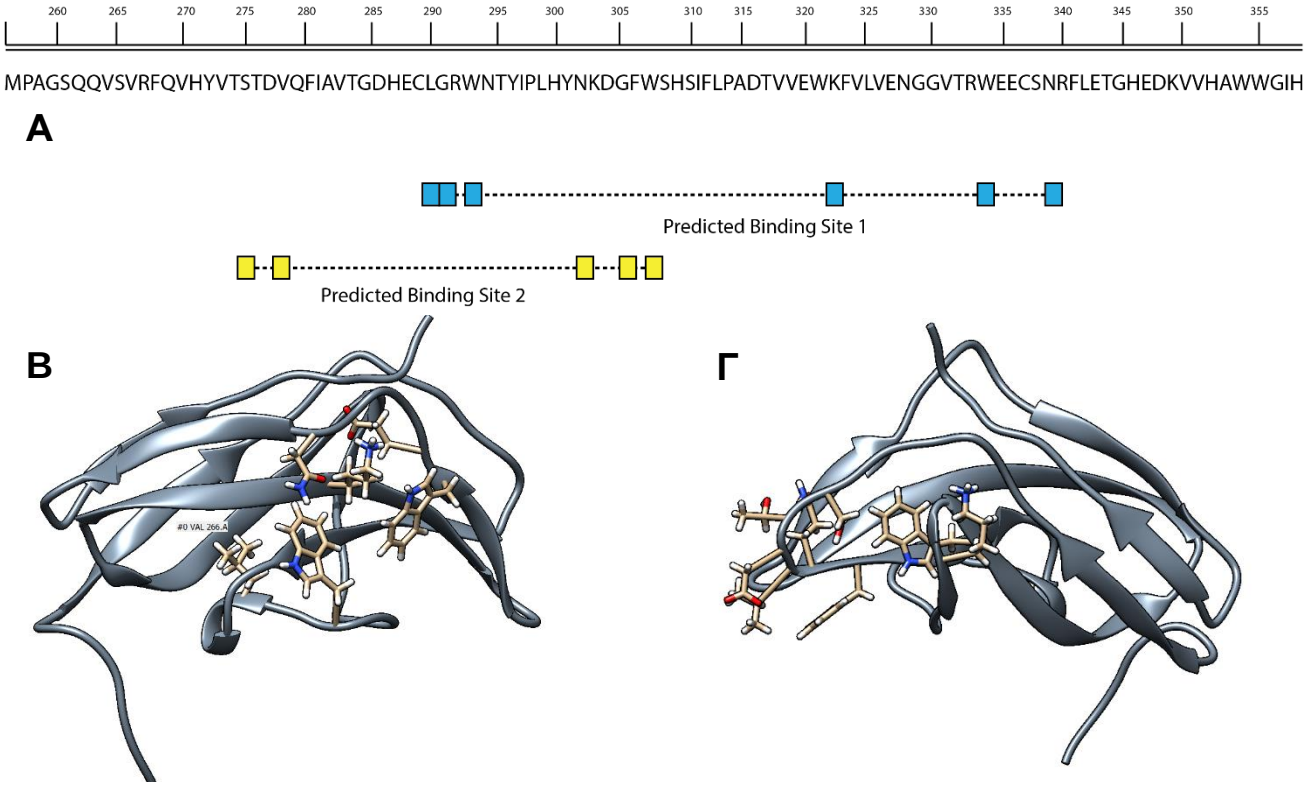
Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

5.1.2.3 Πιθανά συντηρημένα κατάλοιπα στις θέσεις σύνδεσης των προσδετών της CBM20

Μελετώντας προσεκτικά τη βιβλιογραφία είναι εφικτή η πρόβλεψη των καταλοίπων της CBM20 που είναι απαραίτητα για την αλληλεπίδρασή τους με τους προσδέτες. Όπως έχει αναφερθεί, σύμφωνα με τον Svensson, τα συντηρημένα κατάλοιπα οχτώ επικρατειών SBD είναι έντεκα, τα οποία είναι: T525, G528, G535, L540, G541, W543, P561, W563, K578, W590 και N595 (με βάση την αρίθμηση των αμινοξέων της γλυκοαμυλάσης του *Aspergillus niger*). Τα έντεκα αυτά συντηρημένα κατάλοιπα αμινοξέων συνδέονται αρκετά με τις περιοχές σύνδεσης των υδατανθράκων. Η θέση σύνδεσης 1 περιέχει τα δύο συντηρημένα κατάλοιπα τρυπτοφάνης, W543 και W590 [101]. Το μοτίβο ^{540}LGxW^{543} είναι το πιο καλά διατηρημένο μοτίβο ανάμεσα στα μέλη της οικογένειας αυτής. Η θέση σύνδεσης 2 φέρει έχει μεγάλο βαθμό συντήρησης της W563.. Στα έντεκα συντηρημένα κατάλοιπα προστίθεται και η F519 που συντηρείται στο μεγαλύτερο ποσοστό των CBM20 αλληλουχιών, ενώ σε ορισμένα ένζυμα το κατάλοιπο αυτό αντικαθίσταται από λευκίνη, ισολευκίνη και βαλίνη [92].

Στην περίπτωση της επικράτειας της CBM20, φαίνεται να έχουν συμβεί ορισμένες υποκαταστάσεις και ελλείψεις, δεδομένα τα οποία είτε προσθέτουν στην αρχική θεωρία της συντήρησης των αμινοξέων είτε την αναιρούν. Πιο συγκεκριμένα, το κατάλοιπο προλίνης (P561, με βάση την αρίθμηση των αμινοξέων της γλυκοαμυλάσης του *Aspergillus niger*) δεν παρουσιάζεται καθόλου στην πρωτεϊνική ακολουθία της STBD1, εμφανίζεται μία συντηρημένη υποκατάσταση, όπου το κατάλοιπο θρεονίνης, T525, αντισοιχεί για το κατάλοιπο S275 της STBD1 και μία μη συντηρημένη υποκατάσταση, όπου το κατάλοιπο γλυκίνης, G528, αντιστοιχεί για το κατάλοιπο V278 της STBD1.

Συμπερασματικά, από τα παραπάνω υποδεικνύεται, ότι η μία πιθανή θέση σύνδεσης των προσδετών της επικράτειας καθορίζεται από τα κατάλοιπα L290, G291, W293, K323, W334 και N339, ενώ η δεύτερη πιθανή θέση πρόσδεσης καθορίζεται από τα κατάλοιπα S275, V278, K303, G305 και W307. Η πρώτη



Εικόνα 21 Α. Σχεδιάγραμμα των πιθανών θέσεων σύνδεσης των προσδετών στην επικράτεια CBM20 πάνω στην αμινοξική ακολουθία της. Β. Πιθανή θέση σύνδεσης VLS, ως χημικές ενώσεις παρουσιάζονται τα πιθανά κατάλοιπα που καθορίζουν την περιοχή αυτή. Γ. Πιθανή θέση σύνδεσης CFS, ως χημικές ενώσεις παρουσιάζονται τα πιθανά κατάλοιπα που καθορίζουν την περιοχή αυτή. Οι δομές παρουσιάζονται με τη βοήθεια του προγράμματος UCSF Chimera.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

πιθανώς να αντιστοιχεί στην θέση μεταβλητού βρόχου (variable loop site, VLS), και η δεύτερη στην θέση κοίλης επιφάνειας (concave face site, CFS).

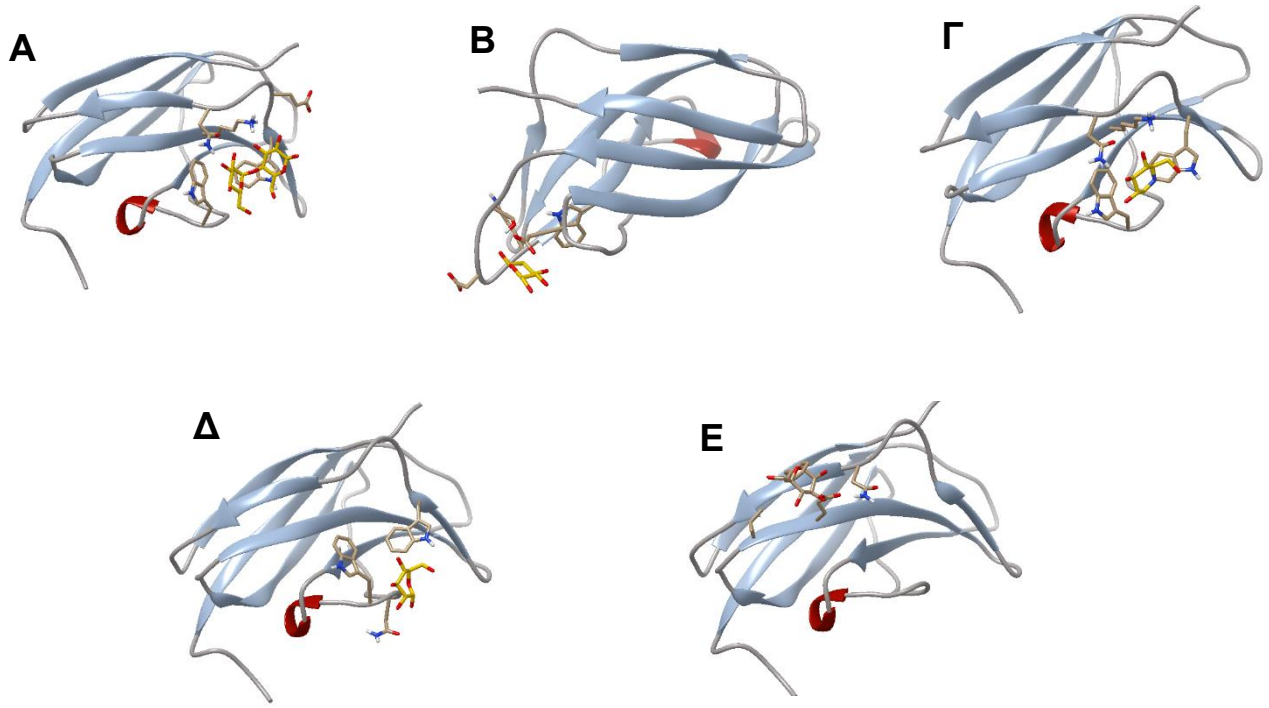
5.1.2.4 Προσδέτες και θέσεις σύνδεσης

Με το I-TASSER είναι δυνατή η πρόβλεψη της λειτουργίας της επικράτειας βάσει της δομής του, και αυτό επιτυγχάνεται με την αναγνώριση προσδετών πάνω σε αυτήν. Για την επικράτεια αυτή, έχουν αναγνωριστεί πέντε προσδέτες: η μαλτόζη, η β-D-γαλακτόζη, η NOJ (1-deoxynojirimycin), η α-D-γλυκόζη και η α-D-μαννόζη.

- Μαλτόζη (maltose, MAL): Ο βαθμός εμπιστοσύνης της σύνδεσης της μαλτόζης με την αλληλουχία είναι 0.69. Η μαλτόζη πιθανώς να αναπτύσσει δεσμούς να με τα αμινικά κατάλοιπα Trp36, Lys65, Trp77 και Asn82. Τα κατάλοιπα αυτά ανήκουν σε τέτοιες θέσεις της ακολουθίας που πιθανώς να σχηματίζεται Variable Loop Site (VLS), μία χαρακτηριστική θέση πρόσδεσης ενώσεων που είναι χαρακτηριστική για την ομάδα CBM20 πρωτεΐνες. Ο προσδέτης και οι θέσεις σύνδεσης του στο μοντέλο προκύπτουν ύστερα από αντιστοίχιση με τη δομή της γλυκοσυλοτρανσφεράσης της κυκλοδεξτρίνης (CGTase) του *Bacillus circulans str. 251* (PDB: 1CDG).
- β-D-γαλακτόζη (β-D-galactose, GAL): Ο βαθμός εμπιστοσύνης της σύνδεσης της β-D-γαλακτόζης με την αλληλουχία είναι 0.18. Η β-D-γαλακτόζη πιθανώς να αναπτύσσει δεσμούς να με τα αμινικά κατάλοιπα Ser18, Thr 19, Gln22, Asn45, Asp47, Trp50. Τα κατάλοιπα αυτά πιθανώς να ανήκουν στη δεύτερη ομάδα θέσεων σύνδεσης προσδετών της CBM20, την Concave Face Site (CFS), που εντοπίζεται στην κοίλη επιφάνεια της δομής και συνδέονται, όχι με επίπεδους πολυσακχαρίτες, αλλά με μία μεγάλη ποικιλία γλυκανών. Ο προσδέτης και οι θέσεις σύνδεσης του στο μοντέλο προκύπτουν ύστερα από αντιστοίχιση με τη δομή της μεταλλαγμένης (F283Y) γλυκοσυλοτρανσφεράσης της κυκλοδεξτρίνης (CGTase) του *Bacillus sp. 1011* (PDB: 1V3M).
- NOJ (1-deoxynojirimycin): Η 1-deoxynojirimycin είναι ένας αναστολέας των α-γλυκοσιδασών που έχει αντιϊική δράση. Ο βαθμός εμπιστοσύνης είναι 0.14. Η 1-deoxynojirimycin πιθανώς να συνδέεται με τα κατάλοιπα Trp36, Asn82. Τα κατάλοιπα αυτά ανήκουν προς το τέλος του μοτίβου CBM20, και πιθανώς αντιστοιχεί στην Variable Loop Site (VLS) των CBM20 πρωτεϊνών. Ο προσδέτης και οι θέσεις σύνδεσης του στο μοντέλο 1 προκύπτουν ύστερα από αντιστοίχιση με τη δομή της γλυκοσυλοτρανσφεράσης της κυκλοδεξτρίνης (CGTase) του *Bacillus sp. 1011* (PDB: 1I75).
- α-D-γλυκόζη (α-D-glucose, GLC): Ο βαθμός εμπιστοσύνης της σύνδεσης της α-D-γλυκόζη με την αλληλουχία είναι 0.05. Η α-D-γλυκόζη πιθανώς να αναπτύσσει δεσμούς να με τα αμινικά κατάλοιπα Trp36, Asn37, Trp77. Τα κατάλοιπα αυτά ανήκουν σε τέτοιες θέσεις της ακολουθίας που πιθανώς να σχηματίζεται Variable Loop Site (VLS), μία χαρακτηριστική θέση πρόσδεσης ενώσεων που είναι χαρακτηριστική για την ομάδα CBM20 πρωτεΐνες. Ο προσδέτης και οι θέσεις σύνδεσης του στο μοντέλο προκύπτουν ύστερα από αντιστοίχιση με τη δομή της μεταλλαγμένης (F283Y) γλυκοσυλοτρανσφεράσης της κυκλοδεξτρίνης (CGTase) του *Bacillus sp. 1011* (PDB: 1V3M).

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

- α-D-μαννόζη (α-D-mannose, MAN): Ο βαθμός εμπιστοσύνης της σύνδεσης της α-D-μαννόζη με την αλληλουχία είναι 0.02. Η α-D-μαννόζη πιθανώς να αναπτύσσει δεσμούς να με τα αμινικά κατάλοιπα Glu63, Asn82, Arg83, Phe84. Στην περίπτωση αυτή, η εικόνα δεν είναι αρκετά ικανοποιητική, καθώς η ένωση έχει πολλές επικαλύψεις με τις πλευρικές αλυσίδες των εκτεθειμένων αμινοξέων στην περιοχή αυτή.

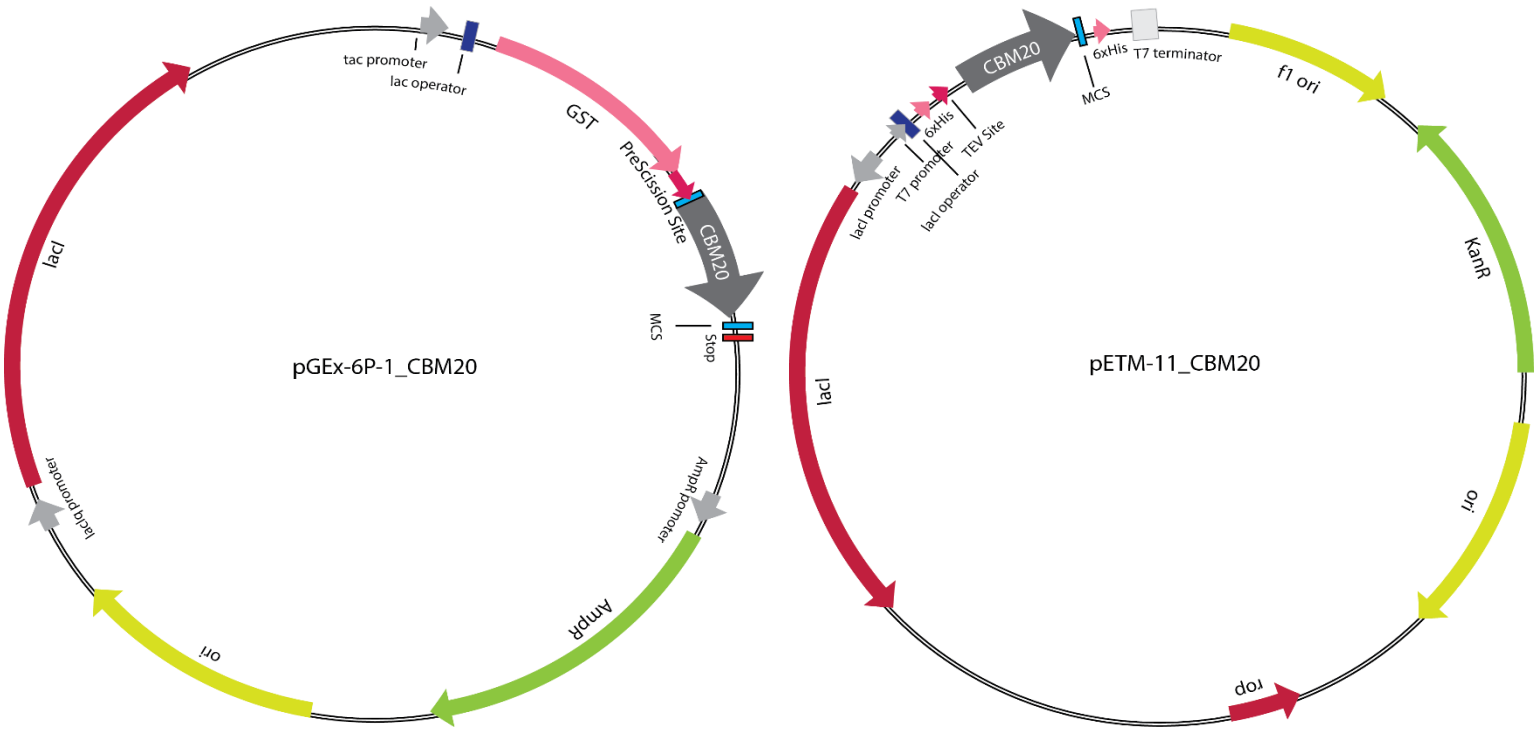


Εικόνα 22 Πρόβλεψη των πιθανών προσδετών στην επιφάνεια της επικράτειας CBM20 με τη βοήθεια του προγράμματος I-TASSER. Οι προσδέτες που προέκυψαν από την ανάλυση αυτή είναι: Α. Μαλτόζη, με βαθμό εμπιστοσύνης 0.69, Β. β-D-γαλακτόζη, με βαθμό εμπιστοσύνης 0.18, Γ. NOJ, με βαθμό εμπιστοσύνης 0.14, Δ. α-D-γλυκόζη, με βαθμό εμπιστοσύνης 0.05 και Ε. α-D-μαννόζη, με βαθμό εμπιστοσύνης 0.02

5.1.3 Ανασυνδυασμένα πλασμίδα

Με τη μεθοδολογία του ανασυνδυασμένου DNA που αναλύθηκε στην αντίστοιχη ενότητα, κατασκευάστηκαν τα πλασμίδια pGEX-6P-1_CBM20 και pETM-11_CBM20 (Εικόνα 23). Η κατασκευή των δύο αυτών διαφορετικών πλασμιδίων στοχεύει στην σύνδεση του ενθέματος με διαφορετικές ετικέτες στο αμινοτελικό άκρο του. Οι ανασυνδυασμένες πρωτεΐνες που θα παραχθούν θα επιτρέψουν, με αυτόν τον τρόπο, την απομόνωση και τον καθαρισμό του επιθυμητού γονιδίου. Τα συγκεκριμένα πλασμίδια επιλέχθηκαν καθώς είναι εύκολα στο χειρισμό και περιέχουν τις ετικέτες GST και ιστιδίνης, αντίστοιχα. Επιπλέον, οι συγκεκριμένες ετικέτες επιτρέπουν και την αύξηση της διαλυτότητας των αναδυνδυασμένων πρωτεϊνών *in vitro*.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

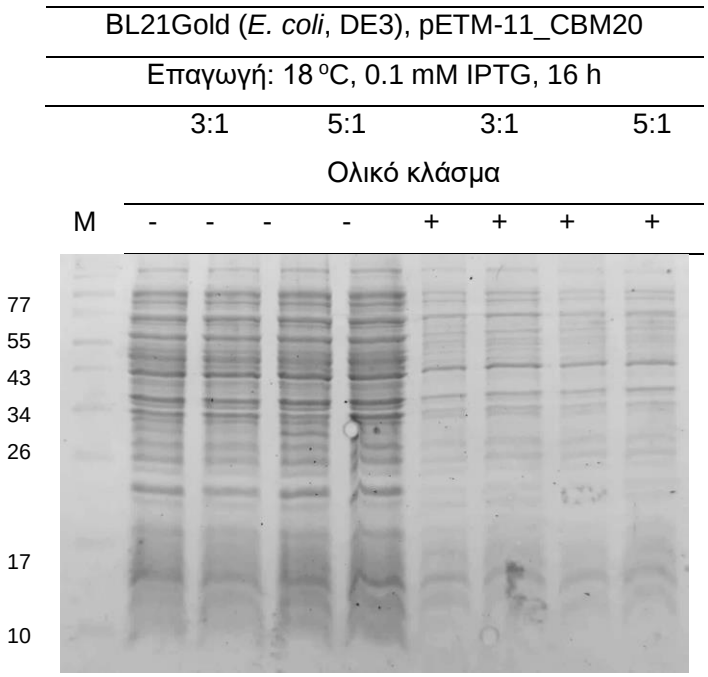


Εικόνα 23 Παρουσίαση των κατασκευασμένων πλασμιδίων. Οι συντμήσεις των παραπάνω τμημάτων των πλασμιδίων αναφέρονται ως: *MCS:* πολυ-συνδέτης, *tac promoter*: κατασκευασμένος DNA υποκινητής από το συνδυασμό των οπερονίων λακτόζης και τρυπτοφάνης, *lac operator*: χειριστής του οπερονίου της λακτόζης, *ori*: περιοχή έναρξης της αντιγραφής, *Stop*: κωδικόνια λήξης της μετάφρασης, *lacI*: το γονίδιο του αναστολέα του οπερονίου της λακτόζης,*AmpR*: γονίιδο ανθεκτικότητας στο αντιβιοτικό αμπικιλλίνη, *AmpR promoter*: ο υποκινητής του γονιδίου ανθεκτικότητας στο αντιβιοτικό αμπικιλλίνη, *lacI^q promoter*: ο μεταλλαγμένος υποκινητής του αναστολέα της λακτόζης, που επιτρέπει την αύξηση της έκφρασής του, *GST*: το γονίδιο της S-τρανσφεράσης της γλουταθειόνης, *PreScission Site*: θέση κοπής από την πρωτεάση PreScission, *KanR*: γονίδιο ανθεκτικότητας στο αντιβιοτικό καναμυκίνη , *T7 promoter*: ο υποκινητής που αναγνωρίζεται από την RNA πολυμεράση T7 , *T7 terminator*: αλληλουχίες λήξης της μεταγραφής από την RNA πολυμεράση T7 , *6xHis*: ετικέτα ιστιδίνης , *TEV Site*: θέση κοπής από την πρωτεάση του *Tobacco Etch Virus* (TEV).

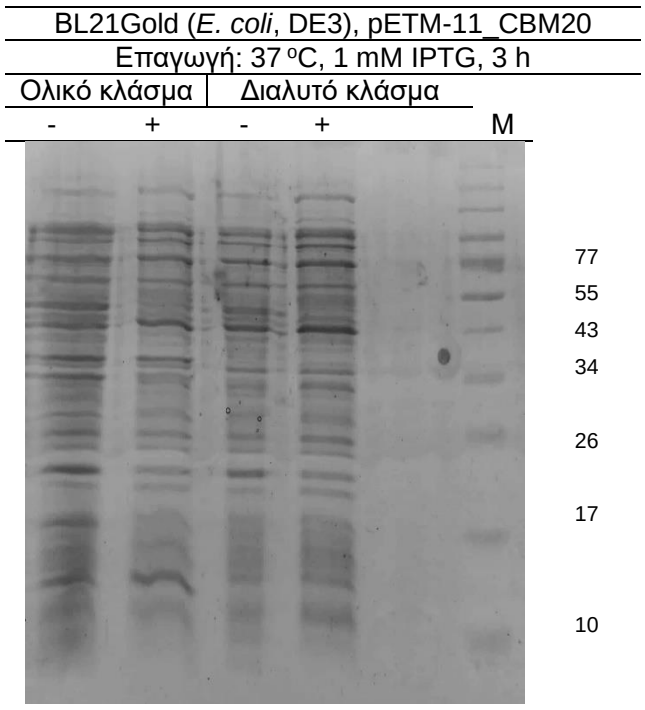
5.1.4 Δοκιμή έκφρασης

Εφόσον, η επικράτεια αυτή κλωνοποιήθηκε στα πλασμίδια, όπως αναφέρθηκε στη μεθοδολογία, έγιναν προσπάθειες της υπερέκφρασής της σε μικρή κλίμακα. Αρχικά, για την έκφραση της ανασυνδυασμένης επικράτειας CBM20 με τις ετικέτες ιστιδίνης και GST, κρατήσαμε ένα δείγμα από την καλλιέργεια πριν την προσθήκη του IPTG και επαγωγή της υπερέκφρασης έγινε σε θερμοκρασία 37 ^{o}C για 4 ώρες παρουσία συγκέντρωσης 1.0 mM IPTG. ακολουθήσαμε τη διαδικασία όπως περιγράφεται στην παράγραφο 4.5 και τα δείγματα αναλύθηκαν σε SDS-PAGE. Παρατηρήθηκε ότι τόσο στους 37 ^{o}C όσο και στους 18 ^{o}C η His-CBM20 δεν εκφράζεται ούτε στο ολικό ούτε στο διαλυτό κλάσμα, ενώ η GST-CMB20, στους 37 ^{o}C, φαίνεται να υπερεκφράζεται μόνο στο ολικό. Έτσι, δοκιμάστηκε και η υπερέκφραση της σε θερμοκρασία 18 ^{o}C για 16 ώρες παρουσία 0.1 mM IPTG, λαμβάνοντας θετικά αποτελέσματα διαλυτής έκφραση της επικράτειας. Τα αποτελέσματα της υπερέκφρασης επιβεβαιώθηκαν με Western blot.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

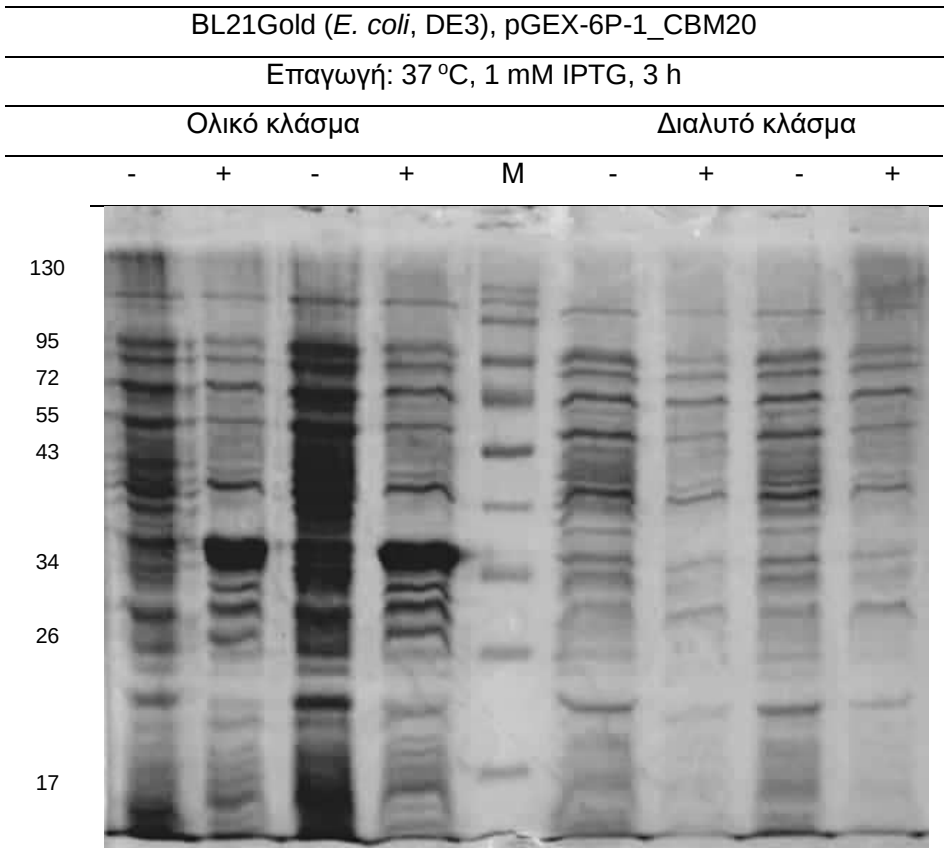


Εικόνα 24 Δοκιμή έκφρασης της His-CBM20 σε στέλεχος BL21Gold (DE3) για διαφορετικούς μοριακούς λόγους. Ως ολικό κλάσμα αναφέρουμε τα διαλυτοποιημένα κύτταρα στα οποία έχει επέλθει λύση με εφαρμογή υπερήχων. Ως το διαλυτό κλάσμα αναφέρουμε το δείγμα που προήλθε μετά από φυγοκέντρηση στο ολικό κλάσμα. Η έκφραση του γονιδίου επάγεται από την προσθήκη του IPTG. Επιπλέον, για τις διαφορετικές συνθήκες θερμοκρασία παρατηρείται ότι η πρωτεΐνη δεν εκφράζεται καθόλου (~11kDa)

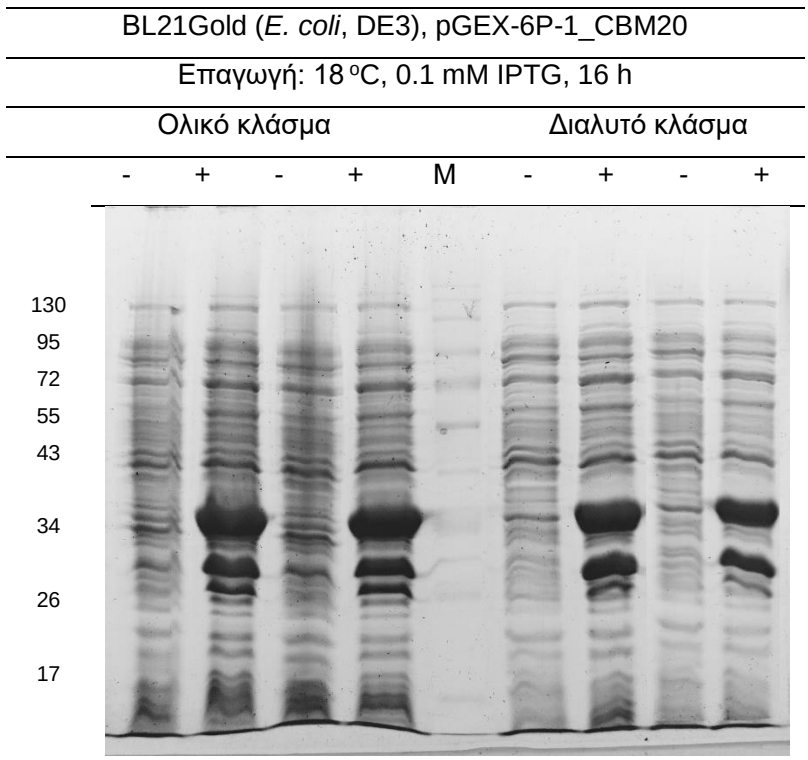


Εικόνα 25 Δοκιμή έκφρασης της His-CBM20 σε στέλεχος BL21Gold (DE3) για διαφορετικούς μοριακούς λόγους. Ως ολικό κλάσμα αναφέρουμε τα διαλυτοποιημένα κύτταρα στα οποία έχει επέλθει λύση με εφαρμογή υπερήχων. Ως το διαλυτό κλάσμα αναφέρουμε το δείγμα που προήλθε μετά από φυγοκέντρηση στο ολικό κλάσμα. Η έκφραση του γονιδίου επάγεται από την προσθήκη του IPTG. Επιπλέον, για τις διαφορετικές συνθήκες θερμοκρασία παρατηρείται ότι η πρωτεΐνη δεν εκφράζεται καθόλου (~11kDa)

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

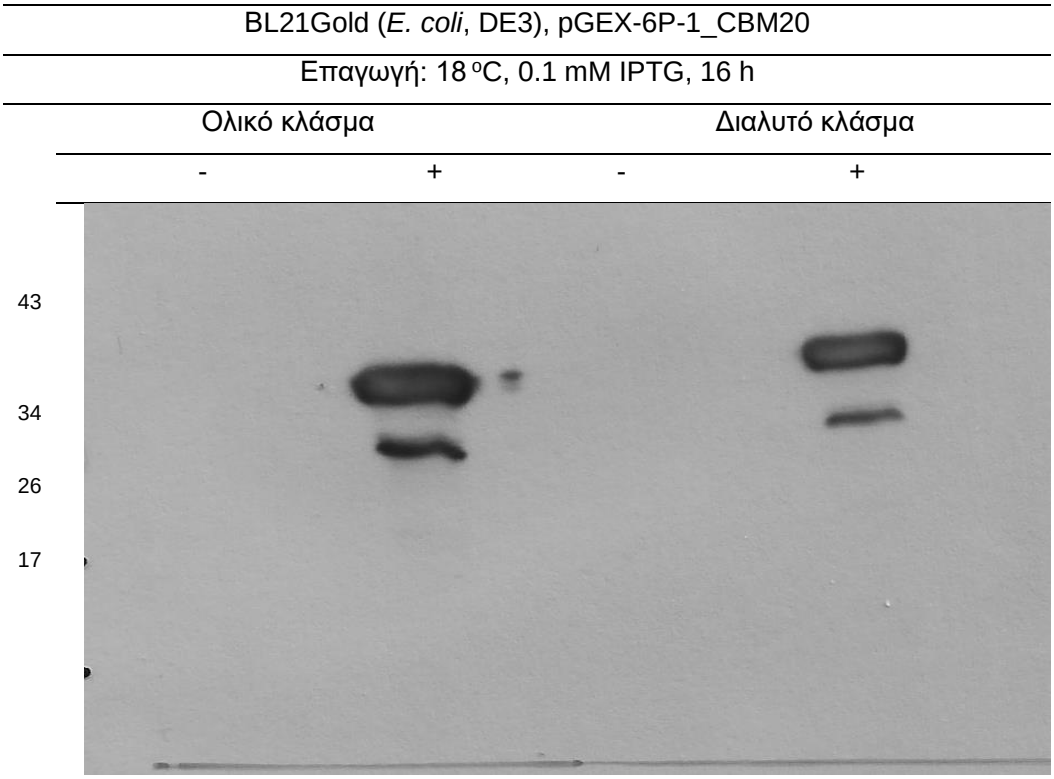


Εικόνα 26 Δοκιμή έκφρασης της GST-CBM20 σε στέλεχος BL21Gold (DE3). Ως ολικό κλάσμα αναφέρουμε τα διαλυτοποιημένα κύτταρα στα οποία έχει επέλθει λύση με εφαρμογή υπερήχων. Ως το διαλυτό κλάσμα αναφέρουμε το δείγμα που προήλθε μετά από φυγοκέντρηση στο ολικό κλάσμα. Η έκφραση του γονιδίου επάγεται από την προσθήκη του IPTG. Στους 37ºC φαίνεται πως η πρωτεΐνη εκφράζεται αλλά όχι στο διαλυτό κλάσμα (~36kDa)



Εικόνα 27 Δοκιμή έκφρασης της GST-CBM20 σε στέλεχος BL21Gold (DE3). Ως ολικό κλάσμα αναφέρουμε τα διαλυτοποιημένα κύτταρα στα οποία έχει επέλθει λύση με εφαρμογή υπερήχων. Ως το διαλυτό κλάσμα αναφέρουμε το δείγμα που προήλθε μετά από φυγοκέντρηση στο ολικό κλάσμα. Η έκφραση τυ γονιδίου επάγεται από την προσθήκη του IPTG. Στους 18C φαίνεται πως η πρωτεΐνη εκφράζεται στο διαλυτό κλάσμα (~36kDa).

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84



Εικόνα 28 Ανάλυση με Western blot δοκιμή έκφρασης της GST-CBM20 σε στέλεχος BL21Gold (DE3). Ως ολικό κλάσμα αναφέρουμε τα διαλυτοποιημένα κύτταρα στα οποία έχει επέλθει λύση με εφαρμογή υπερήχων. Ως το διαλυτό κλάσμα αναφέρουμε το δείγμα που προήλθε μετά από φυγοκέντρηση στο ολικό κλάσμα. Η έκφραση του γονιδίου επάγεται από την προσθήκη του IPTG. Στους 18C φαίνεται πως η πρωτεΐνη εκφράζεται στο διαλυτό κλάσμα (~36kDa). Πρώτο αντίσωμα anti-GST (1:5000) και δεύτερο αντίσωμα anti-mouse (1:500000). Διάρκεια έκθεσης στο διάλυμα Developer 5 sec. Η μπάντα μικρότερη των 34 kDa είναι η GST.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

6 ΕΠΙΛΟΓΟΣ

Στην παρούσα εργασία έγινε μια προσπάθεια ανάλυσης της προβλεπόμενης δομής της CBM20 της STBD1 από εργαλεία βιοπληροφορικής και έκφρασής της σε ετερόλογο σύστημα έκφρασης *E. coli*. Η ανάλυση της δομής με κρυσταλλογραφικές μεθόδους, θα επιτρέψει τον καλύτερο χαρακτηρισμό της δομής και θα επιτρέψει τον προσδιορισμό κατάλληλων προσδετών στην επιφάνεια της επικράτειας, επιτρέποντας περαιτέρω διερεύνηση και καθορισμό της λειτουργίας της STBD1 στο ίδιο το κύτταρο.

Η βιοπληροφορική ανάλυση, καθώς και προσπάθειες έκφρασης της STBD1, απέδειξαν την αστάθεια της πρωτεΐνης και τη δυσκολία έκφρασής της στο ετερόλογο σύστημα *E. coli*. Παρόλα αυτά, κατάφεραν να αναδείξουν τη σταθερή δομή που φέρει μέσα στην πρωτεϊνική της ακολουθία η επικράτεια CBM20. Τα εργαλεία βιοπληροφορικής επέτρεψαν τη διερεύνηση των φυσικοχημικών ιδιοτήτων που χαρακτηρίζουν την επικράτεια, και στην πρόβλεψη δομής της. Όλα τα παραπάνω συνέβαλαν ουσιαστικά στην πειραματική διαδικασία, όπως στον τρόπο με τον οποίο θα μπορέσουμε να κατασκευάσουμε τους πλασμιδιακούς φορείς (καθορισμός θέσεων της CBM20, έλεγχος διαλυτότητας, θέση ετικετών στη χιμαιρική πρωτεΐνη), καθώς και στον έλεγχο διαλυτότητας της στις δοκιμές έκφρασής της σε μικρή κλίμακα.

Κατά την πειραματική διαδικασία κατασκευάστηκαν δύο διαφορετικοί πλασμιδιακοί φορείς με σκοπό την έκφραση δύο χιμαιρικών πρωτεϊνών, που θα διαφέρουν στην επιλογή ετικέτας. Οι ετικέτες επιλογής αποτέλεσαν η ετικέτα ιστιδίνης και η GST. Ωστόσο, φαίνεται πως δεν επιτεύχθηκε η έκφραση και των δύο αυτών χιμαιρικών πρωτεϊνών, αλλά μόνο της επικράτειας CBM20 που είναι συνδεδεμένη στο αμινοτελικό της άκρο με την ετικέτα GST. Όπως είναι ήδη γνωστό από τη βιβλιογραφία, οι πολυπεπτιδικές ετικέτες επιτρέπουν την καλύτερη αναδίπλωση και διαλυτότητα της ανασυνδυασμένης πρωτεΐνης σε ετερόλογα συστήματα έκφρασης. Έτσι, σε αυτή την εργασία επιτεύχθηκε η έκφραση της GST-CBM20 στο βακτηριακό στέλεχος της *E. coli BL21Gold* στους 18°C για 16 ώρες και λύση των κυττάρων σε ρυθμιστικό διάλυμα 20 mM Hepes pH 7.4.

Τα παραπάνω στοιχεία αποτελούν ενθαρρυντικά στοιχεία για τον έλεγχο της έκφρασης μεγαλύτερης κλίμακας, την απομόνωση και τον καθαρισμό της επικράτειας, όπως και τον έλεγχο της μελέτης της δομής και τον καθορισμό προσδετών που θα συνδέονται με αυτήν.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

7 Βιβλιογραφία

1 Roach, P. J., Depaoli-Roach, A. A., Hurley, T. D. & Tagliabracci, V. S. Glycogen and its metabolism: some new developments and old themes. *Biochem J* **441**, 763-787, doi:10.1042/BJ20111416 (2012).

2 Berg, J. M., Tymoczko, J. L., Stryer, L. & Gatto, G. J. *Biochemistry*. International 7th edn, (W.H. Freeman/Palgrave Macmillan, 2012).

3 Liu, Q. H. *et al.* From Prokaryotes to Eukaryotes: Insights Into the Molecular Structure of Glycogen Particles. *Front Mol Biosci* **8**, 673315, doi:10.3389/fmolb.2021.673315 (2021).

4 Prats, C., Graham, T. E. & Shearer, J. The dynamic life of the glycogen granule. *J Biol Chem* **293**, 7089-7098, doi:10.1074/jbc.R117.802843 (2018).

5 Besford, Q. A., Zeng, X. Y., Ye, J. M. & Gray-Weale, A. Liver glycogen in type 2 diabetic mice is randomly branched as enlarged aggregates with blunted glucose release. *Glycoconj J* **33**, 41-51, doi:10.1007/s10719-015-9631-5 (2016).

6 Ryu, J. H. *et al.* Comparative structural analyses of purified glycogen particles from rat liver, human skeletal muscle and commercial preparations. *Int J Biol Macromol* **45**, 478-482, doi:10.1016/j.ijbiomac.2009.08.006 (2009).

7 Reichelt, M. E., Mellor, K. M., Curl, C. L., Stapleton, D. & Delbridge, L. M. Myocardial glycophagy - a specific glycogen handling response to metabolic stress is accentuated in the female heart. *J Mol Cell Cardiol* **65**, 67-75, doi:10.1016/j.yjmcc.2013.09.014 (2013).

8 Adeva-Andany, M. M., Gonzalez-Lucan, M., Donapetry-Garcia, C., Fernandez-Fernandez, C. & Ameneiros-Rodriguez, E. Glycogen metabolism in humans. *BBA Clin* **5**, 85-100, doi:10.1016/j.bbacli.2016.02.001 (2016).

9 Edwards, Y. H., Putt, W., Fox, M. & Ives, J. H. A novel human phosphoglucomutase (PGM5) maps to the centromeric region of chromosome 9. *Genomics* **30**, 350-353, doi:10.1006/geno.1995.9866 (1995).

10 Putt, W. *et al.* Phosphoglucomutase 1: a gene with two promoters and a duplicated first exon. *Biochem J* **296 (Pt 2)**, 417-422, doi:10.1042/bj2960417 (1993).

11 Yu, Q. & Zheng, X. The crystal structure of human UDP-glucose pyrophosphorylase reveals a latch effect that influences enzymatic activity. *Biochem J* **442**, 283-291, doi:10.1042/BJ20111598 (2012).

12 Gibbons, B. J., Roach, P. J. & Hurley, T. D. Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin. *J Mol Biol* **319**, 463-477, doi:10.1016/S0022-2836(02)00305-4 (2002).

13 Kotoulas, O. B., Kalamidas, S. A. & Kondomerkos, D. J. Glycogen autophagy. *Microsc Res Tech* **64**, 10-20, doi:10.1002/jemt.20046 (2004).

14 Tanida, I. Autophagy basics. *Microbiol Immunol* **55**, 1-11, doi:10.1111/j.1348-0421.2010.00271.x (2011).

15 Kaur, J. & Debnath, J. Autophagy at the crossroads of catabolism and anabolism. *Nat Rev Mol Cell Biol* **16**, 461-472, doi:10.1038/nrm4024 (2015).

16 Parzych, K. R. & Klionsky, D. J. An overview of autophagy: morphology, mechanism, and regulation. *Antioxid Redox Signal* **20**, 460-473, doi:10.1089/ars.2013.5371 (2014).

17 Nitowsky, H. M. & Grunfeld, A. Lysosomal alpha-glucosidase in type II glycogenosis; activity in leukocytes and cell cultures in relation to genotype. *J Lab Clin Med* **69**, 472-484 (1967).

18 Phillips, M. J., Unakar, N. J., Doornewaard, G. & Steiner, J. W. Glycogen depletion in the newborn rat liver: an electron microscopic and electron histochemical study. *J Ultrastruct Res* **18**, 142-165, doi:10.1016/s0022-5320(67)80236-3 (1967).

19 Chan, J. *et al.* The emerging phenotype of late-onset Pompe disease: A systematic literature review. *Mol Genet Metab* **120**, 163-172, doi:10.1016/j.ymgme.2016.12.004 (2017).

20 Palmer, T. N. The substrate specificity of acid -glucosidase from rabbit muscle. *Biochem J* **124**, 701-711, doi:10.1042/bj1240701 (1971).

21 Chandramouli, C. *et al.* Myocardial glycogen dynamics: new perspectives on disease mechanisms. *Clin Exp Pharmacol Physiol* **42**, 415-425, doi:10.1111/1440-1681.12370 (2015).

22 Kotoulas, O. B., Kalamidas, S. A. & Kondomerkos, D. J. Glycogen autophagy in glucose homeostasis. *Pathol Res Pract*

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

202, 631-638, doi:10.1016/j.prp.2006.04.001 (2006).

23 Roig-Zamboni, V. *et al.* Structure of human lysosomal acid alpha-glucosidase-a guide for the treatment of Pompe disease. *Nat Commun* **8**, 1111, doi:10.1038/s41467-017-01263-3 (2017).

24 Tashiro, K., Iwamasa, T., Kato, H., Ogata, S. & Anai, M. Purification and characterization of two components of acid alpha-glucosidase from pig liver. *J Biochem* **99**, 693-701, doi:10.1093/oxfordjournals.jbchem.a135528 (1986).

25 Mizushima, N. Physiological functions of autophagy. *Curr Top Microbiol Immunol* **335**, 71-84, doi:10.1007/978-3-642-00302-8_3 (2009).

26 Ge, L. *et al.* Remodeling of ER-exit sites initiates a membrane supply pathway for autophagosome biogenesis. *EMBO Rep* **18**, 1586-1603, doi:10.15252/embr.201744559 (2017).

27 Graef, M., Friedman, J. R., Graham, C., Babu, M. & Nunnari, J. ER exit sites are physical and functional core autophagosome biogenesis components. *Mol Biol Cell* **24**, 2918-2931, doi:10.1091/mbc.E13-07-0381 (2013).

28 Axe, E. L. *et al.* Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. *J Cell Biol* **182**, 685-701, doi:10.1083/jcb.200803137 (2008).

29 Hayashi-Nishino, M. *et al.* A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation. *Nat Cell Biol* **11**, 1433-1437, doi:10.1038/ncb1991 (2009).

30 Uemura, T. *et al.* A cluster of thin tubular structures mediates transformation of the endoplasmic reticulum to autophagic isolation membrane. *Mol Cell Biol* **34**, 1695-1706, doi:10.1128/MCB.01327-13 (2014).

31 Yla-Anttila, P., Vihinen, H., Jokitalo, E. & Eskelinen, E. L. 3D tomography reveals connections between the phagophore and endoplasmic reticulum. *Autophagy* **5**, 1180-1185, doi:10.4161/auto.5.8.10274 (2009).

32 Ge, L. & Schekman, R. The ER-Golgi intermediate compartment feeds the phagophore membrane. *Autophagy* **10**, 170-172, doi:10.4161/auto.26787 (2014).

33 Tang, Z. *et al.* TOM40 Targets Atg2 to Mitochondria-Associated ER Membranes for Phagophore Expansion. *Cell Rep* **28**, 1744-1757 e1745, doi:10.1016/j.celrep.2019.07.036 (2019).

34 Jatana, N., Ascher, D. B., Pires, D. E. V., Gokhale, R. S. & Thukral, L. Human LC3 and GABARAP subfamily members achieve functional specificity via specific structural modulations. *Autophagy* **16**, 239-255, doi:10.1080/15548627.2019.1606636 (2020).

35 Le Grand, J. N. *et al.* GABARAPL1 (GEC1): original or copycat? *Autophagy* **7**, 1098-1107, doi:10.4161/auto.7.10.15904 (2011).

36 Kalamidas, S. A., Kotoulas, O. B., Kotoulas, A. O. & Maintas, D. B. The breakdown of glycogen in the lysosomes of newborn rat hepatocytes: the effects of glucose, cyclic 3',5'-AMP and caffeine. *Histol Histopathol* **9**, 691-698 (1994).

37 Kondomerkos, D. J., Kalamidas, S. A., Kotoulas, O. B. & Hann, A. C. Glycogen autophagy in the liver and heart of newborn rats. The effects of glucagon, adrenalin or rapamycin. *Histol Histopathol* **20**, 689-696, doi:10.14670/HH-20.689 (2005).

38 Kotoulas, A. O., Kotoulas, O. B. & Kalamidas, S. An electron microscopic and biochemical study of the effects of cyclic 3', 5'-AMP, ergotamine or propranolol on the lysosomes of newborn rat hepatocytes. *Histol Histopathol* **6**, 421-426 (1991).

39 Yu, Q., Shuai, H., Ahooghalandari, P., Gylfe, E. & Tengholm, A. Glucose controls glucagon secretion by directly modulating cAMP in alpha cells. *Diabetologia* **62**, 1212-1224, doi:10.1007/s00125-019-4857-6 (2019).

40 Zoncu, R., Efeyan, A. & Sabatini, D. M. mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol* **12**, 21-35, doi:10.1038/nrm3025 (2011).

41 Petiot, A., Pattingre, S., Arico, S., Meley, D. & Codogno, P. Diversity of signaling controls of macroautophagy in mammalian cells. *Cell Struct Funct* **27**, 431-441, doi:10.1247/csf.27.431 (2002).

42 Yang, G., Wang, N., Seto, S. W., Chang, D. & Liang, H. Hydroxysafflor yellow a protects brain microvascular endothelial cells against oxygen glucose deprivation/reoxygenation injury: Involvement of inhibiting autophagy via class I PI3K/Akt/mTOR signaling pathway. *Brain Res Bull* **140**, 243-257, doi:10.1016/j.brainresbull.2018.05.011 (2018).

43 Zhang, Y. & Miao, J. M. Ginkgolide K promotes astrocyte proliferation and migration after oxygen-glucose deprivation via inducing protective autophagy through the AMPK/mTOR/ULK1 signaling pathway.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Eur J Pharmacol **832**, 96-103, doi:10.1016/j.ejphar.2018.05.029 (2018).
44 Quiocho, F. A. Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu Rev Biochem* **55**, 287-315, doi:10.1146/annurev.bi.55.070186.001443 (1986).
45 Armenta, S., Moreno-Mendieta, S., Sanchez-Cuapio, Z., Sanchez, S. & Rodriguez-Sanoja, R. Advances in molecular engineering of carbohydrate-binding modules. *Proteins* **85**, 1602-1617, doi:10.1002/prot.25327 (2017).
46 Sorimachi, K., Le Gal-Coeffet, M. F., Williamson, G., Archer, D. B. & Williamson, M. P. Solution structure of the granular starch binding domain of Aspergillus niger glucoamylase bound to beta-cyclodextrin. *Structure* **5**, 647-661, doi:10.1016/s0969-2126(97)00220-7 (1997).
47 Linder, M. T., T. T. . The roles and function of cellulose-binding domains. *Journal of Biotechnology* **57**, 14, doi:10.1016/S0168-1656(97)00087-4 (1997).
48 Christiansen, C. *et al.* The carbohydrate-binding module family 20--diversity, structure, and function. *FEBS J* **276**, 5006-5029, doi:10.1111/j.1742-4658.2009.07221.x (2009).
49 Janecek, S., Marecek, F., MacGregor, E. A. & Svensson, B. Starch-binding domains as CBM families-history, occurrence, structure, function and evolution. *Biotechnol Adv* **37**, 107451, doi:10.1016/j.biotechadv.2019.107451 (2019).
50 Tomme, P. *et al.* Studies of the cellulolytic system of Trichoderma reesei QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem* **170**, 575-581, doi:10.1111/j.1432-1033.1988.tb13736.x (1988).
51 Gilkes, N. R., Warren, R. A., Miller, R. C., Jr. & Kilburn, D. G. Precise excision of the cellulose binding domains from two Cellulomonas fimi cellulases by a homologous protease and the effect on catalysis. *J Biol Chem* **263**, 10401-10407 (1988).
52 Sorimachi, K. *et al.* Solution structure of the granular starch binding domain of glucoamylase from Aspergillus niger by nuclear magnetic resonance spectroscopy. *J Mol Biol* **259**, 970-987, doi:10.1006/jmbi.1996.0374 (1996).
53 Coutinho, P. M., Henrissat B. Carbohydrate-active enzymes : an integrated database approach. (1999).
54 Cantarel, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37**, D233-238, doi:10.1093/nar/gkn663 (2009).
55 Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490-495, doi:10.1093/nar/gkt1178 (2014).
56 Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382**, 769-781, doi:10.1042/BJ20040892 (2004).
57 Jamal-Talabani, S. *et al.* Ab initio structure determination and functional characterization of CBM36; a new family of calcium-dependent carbohydrate binding modules. *Structure* **12**, 1177-1187, doi:10.1016/j.str.2004.04.022 (2004).
58 Montanier, C. *et al.* Circular permutation provides an evolutionary link between two families of calcium-dependent carbohydrate binding modules. *J Biol Chem* **285**, 31742-31754, doi:10.1074/jbc.M110.142133 (2010).
59 Henshaw, J. *et al.* Family 6 carbohydrate binding modules in beta-agarases display exquisite selectivity for the non-reducing termini of agarose chains. *J Biol Chem* **281**, 17099-17107, doi:10.1074/jbc.M600702200 (2006).
60 Montanier, C. *et al.* Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. **106**, 3065-3070, doi:10.1073/pnas.0808972106 %J Proceedings of the National Academy of Sciences (2009).
61 Gilbert, H. J., Knox, J. P. & Boraston, A. B. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr Opin Struct Biol* **23**, 669-677, doi:10.1016/j.sbi.2013.05.005 (2013).
62 Ficko-Blean, E. & Boraston, A. B. Insights into the recognition of the human glycome by microbial carbohydrate-binding modules. *Curr Opin Struct Biol* **22**, 570-577, doi:10.1016/j.sbi.2012.07.009 (2012).
63 Bolam, D. N. *et al.* Pseudomonas cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. *Biochem J* **331 (Pt 3)**, 775-781, doi:10.1042/bj3310775 (1998).
64 Boraston, A. B., Kwan, E., Chiu, P., Warren, R. A. & Kilburn, D. G. Recognition and hydrolysis of noncrystalline cellulose. *J*

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Biol Chem **278**, 6120-6127, doi:10.1074/jbc.M209554200 (2003).

65 Takahashi, T., Kato, K., Ikegami, Y. & Irie, M. Different behavior towards raw starch of three forms of glucoamylase from a Rhizopus sp. *J Biochem* **98**, 663-671, doi:10.1093/oxfordjournals.jbchem.a135323 (1985).

66 Viksø-Nielsen, A., Andersen, C., Hoff, T. & Pedersen, S. Development of new α-amylases for raw starch hydrolysis. *Biocatalysis and Biotransformation* **24**, 7, doi:10.1080/10242420500519191 (2006).

67 Carrard, G., Koivula, A., Soderlund, H. & Beguin, P. Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc Natl Acad Sci U S A* **97**, 10342-10347, doi:10.1073/pnas.160216697 (2000).

68 Tomme, P., Creagh, A. L., Kilburn, D. G. & Haynes, C. A. Interaction of polysaccharides with the N-terminal cellulose-binding domain of Cellulomonas fimi CenC. 1. Binding specificity and calorimetric analysis. *Biochemistry* **35**, 13885-13894, doi:10.1021/bi961185i (1996).

69 Orengo, C. A. *et al.* The CATH Database provides insights into protein structure/function relationships. *Nucleic Acids Res* **27**, 275-279, doi:10.1093/nar/27.1.275 (1999).

70 Hayashida, S., Kunisaki, S.-i., Nakao, M. & Flor, P. Q. Evidence for Raw Starch-affinity Site on Aspergillus awamori Glucoamylase I. *Agricultural and Biological Chemistry* **46**, 83-89, doi:10.1080/00021369.1982.10865024 %J Agricultural and Biological Chemistry (1982).

71 Boel, E. *et al.* Glucoamylases G1 and G2 from Aspergillus niger are synthesized from two different but closely related mRNAs. *EMBO J.* **3**, 6 (1984).

72 Svensson, B., Pedersen, T. G., Svendsen, I., Sakai, T., & Ottesen, M. Characterization of two forms of glucoamylase from Aspergillus niger. *Carlsberg Res. Commun.* **47**, 15 (1982).

73 Flor, P. Q. & Hayashida, S. Production and Characteristics of Raw Starch-Digesting Glucoamylase O from a Protease-Negative, Glycosidase-Negative Aspergillus awamori var. kawachi Mutant. *Appl Environ Microbiol* **45**, 905-912, doi:10.1128/AEM.45.3.905-912.1983 (1983).

74 Svensson, B. Regional distant sequence homology between amylases, alpha-glucosidases and transglucanosylases. *FEBS Lett* **230**, 72-76, doi:10.1016/0014-5793(88)80644-6 (1988).

75 MacGregor, E. A., Janecek, S. & Svensson, B. Relationship of sequence and structure to specificity in the alpha-amylase family of enzymes. *Biochim Biophys Acta* **1546**, 1-20, doi:10.1016/s0167-4838(00)00302-2 (2001).

76 Janecek, S., Svensson, B. & MacGregor, E. A. Structural and evolutionary aspects of two families of non-catalytic domains present in starch and glycogen binding proteins from microbes, plants and animals. *Enzyme Microb Technol* **49**, 429-440, doi:10.1016/j.enzmictec.2011.07.002 (2011).

77 Bott, R. *et al.* Three-dimensional structure of an intact glycoside hydrolase family 15 glucoamylase from Hypocrea jecorina. *Biochemistry* **47**, 5746-5754, doi:10.1021/bi702413k (2008).

78 Klein, C. & Schulz, G. E. Structure of cyclodextrin glycosyltransferase refined at 2.0 A resolution. *J Mol Biol* **217**, 737-750, doi:10.1016/0022-2836(91)90530-j (1991).

79 Lawson, C. L. *et al.* Nucleotide sequence and X-ray structure of cyclodextrin glycosyltransferase from Bacillus circulans strain 251 in a maltose-dependent crystal form. *J Mol Biol* **236**, 590-600, doi:10.1006/jmbi.1994.1168 (1994).

80 Mikami, B. *et al.* Structure of raw starch-digesting Bacillus cereus beta-amylase complexed with maltose. *Biochemistry* **38**, 7050-7061, doi:10.1021/bi9829377 (1999).

81 Oyama, T., Kusunoki, M., Kishimoto, Y., Takasaki, Y. & Nitta, Y. Crystal structure of beta-amylase from Bacillus cereus var. mycoides at 2.2 A resolution. *J Biochem* **125**, 1120-1130, doi:10.1093/oxfordjournals.jbchem.a022394 (1999).

82 Penninga, D. *et al.* The raw starch binding domain of cyclodextrin glycosyltransferase from Bacillus circulans strain 251. *J Biol Chem* **271**, 32777-32784, doi:10.1074/jbc.271.51.32777 (1996).

83 Roth, C. *et al.* Structural insight into industrially relevant glucoamylases: flexible positions of starch-binding domains. *Acta Crystallogr D Struct Biol* **74**, 463-470, doi:10.1107/S2059798318004989 (2018).

84 Svensson, B., Larsen, K., Svendsen, I. & Boel, E. The complete amino acid sequence of the glycoprotein, glucoamylase G1, from Aspergillus niger. *Carlsberg Research Communications* **48**, doi:10.1007/BF02907555 (1983).

85 Williamson, G., Belshaw, N. J. & Williamson, M. P. O-glycosylation in

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Aspergillus glucoamylase. Conformation and role in binding. *Biochem J* **282 (Pt 2)**, 423-428, doi:10.1042/bj2820423 (1992).

86 Kotting, O. *et al.* Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. The phosphoglucan, water dikinase. *Plant Physiol* **137**, 242-252, doi:10.1104/pp.104.055954 (2005).

87 Orzechowski, S. *et al.* Analysis of the expression, subcellular and tissue localisation of phosphoglucan, water dikinase (PWD/GWD3) in Solanum tuberosum L.: a bioinformatics approach for the comparative analysis of two α-glucan, water dikinases (GWDs) from Solanum tuberosum L. *Acta Physiologiae Plantarum* **35**, 8, doi:10.1007/s11738-012-1091-y (2013).

88 Okazaki, Y. *et al.* A novel glycerophosphodiester phosphodiesterase, GDE5, controls skeletal muscle development via a non-enzymatic mechanism. *J Biol Chem* **285**, 27652-27663, doi:10.1074/jbc.M110.106708 (2010).

89 Minassian, B. A., Ianzano, L., Delgado-Escueta, A. V. & Scherer, S. W. Identification of new and common mutations in the EPM2A gene in Lafora disease. *Neurology* **54**, 488-490, doi:10.1212/wnl.54.2.488 (2000).

90 Bouju, S. *et al.* Molecular cloning and functional expression of a novel human gene encoding two 41-43 kDa skeletal muscle internal membrane proteins. *Biochem J* **335 (Pt 3)**, 549-556, doi:10.1042/bj3350549 (1998).

91 Janecek, S. A motif of a microbial starch-binding domain found in human genethonin. *Bioinformatics* **18**, 1534-1537, doi:10.1093/bioinformatics/18.11.1534 (2002).

92 Machovic, M., Svensson, B., MacGregor, E. A. & Janecek, S. A new clan of CBM families based on bioinformatics of starch-binding domains from families CBM20 and CBM21. *FEBS J* **272**, 5497-5513, doi:10.1111/j.1742-4658.2005.04942.x (2005).

93 Machovic, M. & Janecek, S. Starch-binding domains in the post-genome era. *Cell Mol Life Sci* **63**, 2710-2724, doi:10.1007/s00018-006-6246-9 (2006).

94 Kuchtova, A. & Janecek, S. In silico analysis of family GH77 with focus on amylomaltases from borreliae and disproportionating enzymes DPE2 from plants and bacteria. *Biochim Biophys Acta* **1854**, 1260-1268, doi:10.1016/j.bbapap.2015.05.009 (2015).

95 Lee, S. P., Morikawa, M., Takagi, M. & Imanaka, T. Cloning of the aapT gene and characterization of its product, alpha-amylase-pullulanase (AapT), from thermophilic and alkaliphilic Bacillus sp. strain XAL601. *Appl Environ Microbiol* **60**, 3764-3773, doi:10.1128/AEM.60.10.3764-3773.1994 (1994).

96 Oslancova, A. & Janecek, S. Oligo-1,6-glucosidase and neopullulanase enzyme subfamilies from the alpha-amylase family defined by the fifth conserved sequence region. *Cell Mol Life Sci* **59**, 1945-1959, doi:10.1007/pl00012517 (2002).

97 Saburi, W. *et al.* A thermophilic alkalophilic alpha-amylase from Bacillus sp. AAH-31 shows a novel domain organization among glycoside hydrolase family 13 enzymes. *Biosci Biotechnol Biochem* **77**, 1867-1873, doi:10.1271/bbb.130284 (2013).

98 Watanabe, H., Nishimoto, T., Kubota, M., Chaen, H. & Fukuda, S. Cloning, sequencing, and expression of the genes encoding an isocyclomaltooligosaccharide glucanotransferase and an alpha-amylase from a Bacillus circulans strain. *Biosci Biotechnol Biochem* **70**, 2690-2702, doi:10.1271/bbb.60294 (2006).

99 Svensson, B., Jespersen, H., Sierks, M. R. & MacGregor, E. A. Sequence homology between putative raw-starch binding domains from different starch-degrading enzymes. *Biochem J* **264**, 309-311, doi:10.1042/bj2640309 (1989).

100 Machovic, M. & Janecek, S. The evolution of putative starch-binding domains. *FEBS Lett* **580**, 6349-6356, doi:10.1016/j.febslet.2006.10.041 (2006).

101 Williamson, M. P. *et al.* Function of conserved tryptophans in the Aspergillus niger glucoamylase 1 starch binding domain. *Biochemistry* **36**, 7535-7539, doi:10.1021/bi9702896 (1997).

102 Burley, S. K. *et al.* RCSB Protein Data Bank: Sustaining a living digital data resource that enables breakthroughs in scientific research and biomedical education. *Protein Sci* **27**, 316-330, doi:10.1002/pro.3331 (2018).

103 Hashimoto, H. Recent structural studies of carbohydrate-binding modules. *Cell Mol Life Sci* **63**, 2954-2967, doi:10.1007/s00018-006-6195-3 (2006).

104 Raththagala, M. *et al.* Structural mechanism of laforin function in glycogen dephosphorylation and lafora disease. *Mol Cell* **57**, 261-272, doi:10.1016/j.molcel.2014.11.020 (2015).

105 Jacks, A. J. *et al.* 1H and 15N assignments and secondary structure of the starch-

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

binding domain of glucoamylase from Aspergillus niger. *Eur J Biochem* **233**, 568-578, doi:10.1111/j.1432-1033.1995.568_2.x (1995).
106 Stapleton, D. *et al.* Analysis of hepatic glycogen-associated proteins. *Proteomics* **10**, 2320-2329, doi:10.1002/pmic.200900628 (2010).
107 Jiang, S., Wells, C. D. & Roach, P. J. Starch-binding domain-containing protein 1 (Stbd1) and glycogen metabolism: Identification of the Atg8 family interacting motif (AIM) in Stbd1 required for interaction with GABARAPL1. *Biochem Biophys Res Commun* **413**, 420-425, doi:10.1016/j.bbrc.2011.08.106 (2011).
108 Popelka, H. & Klionsky, D. J. Analysis of the native conformation of the LIR/AIM motif in the Atg8/LC3/GABARAP-binding proteins. *Autophagy* **11**, 2153-2159, doi:10.1080/15548627.2015.1111503 (2015).
109 Ducommun, S. *et al.* Chemical genetic screen identifies Gapex-5/GAPVD1 and STBD1 as novel AMPK substrates. *Cell Signal* **57**, 45-57, doi:10.1016/j.cellsig.2019.02.001 (2019).
110 Jiang, S. *et al.* Starch binding domain-containing protein 1/genethonin 1 is a novel participant in glycogen metabolism. *J Biol Chem* **285**, 34960-34971, doi:10.1074/jbc.M110.150839 (2010).
111 Demetriadou, A. *et al.* Mouse Stbd1 is N-myristoylated and affects ER-mitochondria association and mitochondrial morphology. *J Cell Sci* **130**, 903-915, doi:10.1242/jcs.195263 (2017).
112 Lytridou, A. A. *et al.* Stbd1 promotes glycogen clustering during endoplasmic reticulum stress and supports survival of mouse myoblasts. *J Cell Sci*, doi:10.1242/jcs.244855 (2020).
113 Zhu, Y., Zhang, M., Kelly, A. R. & Cheng, A. The carbohydrate-binding domain of overexpressed STBD1 is important for its stability and protein-protein interactions. *Biosci Rep* **34**, doi:10.1042/BSR20140053 (2014).
114 Sun, T., Yi, H., Yang, C., Kishnani, P. S. & Sun, B. Starch Binding Domain-containing Protein 1 Plays a Dominant Role in Glycogen Transport to Lysosomes in Liver. *J Biol Chem* **291**, 16479-16484, doi:10.1074/jbc.C116.741397 (2016).
115 Wilkins, M. R. *et al.* Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **112**, 531-552, doi:10.1385/1-59259-584-7:531 (1999).
116 Gill, S. C. & von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **182**, 319-326, doi:10.1016/0003-2697(89)90602-7 (1989).
117 Ikai, A. Thermostability and aliphatic index of globular proteins. *J Biochem* **88**, 1895-1898 (1980).
118 Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132, doi:10.1016/0022-2836(82)90515-0 (1982).
119 George, R. A. & Heringa, J. Protein domain identification and improved sequence similarity searching using PSI-BLAST. *Proteins* **48**, 672-681, doi:10.1002/prot.10175 (2002).
120 Blum, M. *et al.* The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* **49**, D344-D354, doi:10.1093/nar/gkaa977 (2021).
121 Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **5**, 725-738, doi:10.1038/nprot.2010.5 (2010).
122 Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **12**, 7-8, doi:10.1038/nmeth.3213 (2015).
123 Zhang, Y. Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins* **69 Suppl 8**, 108-117, doi:10.1002/prot.21702 (2007).
124 Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40, doi:10.1186/1471-2105-9-40 (2008).
125 Garner, E., Romero, P., Dunker, A. K., Brown, C. & Obradovic, Z. Predicting Binding Regions within Disordered Proteins. *Genome Inform Ser Workshop Genome Inform* **10**, 41-50 (1999).
126 Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol* **5**, 172, doi:10.3389/fmicb.2014.00172 (2014).
127 Itakura, K. *et al.* Expression in Escherichia coli of a chemically synthesized gene for the hormone somatostatin. *Science* **198**, 1056-1063, doi:10.1126/science.412251 (1977).
128 Bentley, W. E., Mirjalili, N., Andersen, D. C., Davis, R. H. & Kompala, D. S. Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. *Biotechnol Bioeng* **35**, 668-681, doi:10.1002/bit.260350704 (1990).
129 Sezonov, G., Joseleau-Petit, D. & D'Ari, R. Escherichia coli physiology in Luria-Bertani

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

broth. *J Bacteriol* **189**, 8746-8749, doi:10.1128/JB.01368-07 (2007).
130 Lee, S. Y. High cell-density culture of Escherichia coli. *Trends Biotechnol* **14**, 98-105, doi:10.1016/0167-7799(96)80930-9 (1996).
131 Shiloach, J. & Fass, R. Growing E. coli to high cell density--a historical perspective on method development. *Biotechnol Adv* **23**, 345-357, doi:10.1016/j.biotechadv.2005.04.004 (2005).
132 Choi, J. H., Keum, K. C., Lee, S. Y. Production of recombinant proteins by high cell density culture of Escherichia coli. *Chemical Engineering Science* **61**, 876-885, doi:https://doi.org/10.1016/j.ces.2005.03.031 (2006).
133 Pope, B. & Kent, H. M. High efficiency 5 min transformation of Escherichia coli. *Nucleic Acids Res* **24**, 536-537, doi:10.1093/nar/24.3.536 (1996).
134 Carrio, M. M. & Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. *J Biotechnol* **96**, 3-12, doi:10.1016/s0168-1656(02)00032-9 (2002).
135 Hartley, D. L. & Kane, J. F. Properties of inclusion bodies from recombinant Escherichia coli. *Biochem Soc Trans* **16**, 101-102, doi:10.1042/bst0160101 (1988).
136 del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M. & Diaz-Orejas, R. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**, 434-464, doi:10.1128/MMBR.62.2.434-464.1998 (1998).
137 Müller-Hill, B. *The lac Operon : a short history of a genetic paradigm*. (Walter de Gruyter, 1996).
138 Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113-130, doi:10.1016/0022-2836(86)90385-2 (1986).
139 Moffatt, B. A. & Studier, F. W. T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49**, 221-227, doi:10.1016/0092-8674(87)90563-0 (1987).
140 Dubendorff, J. W. & Studier, F. W. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J Mol Biol* **219**, 45-59, doi:10.1016/0022-2836(91)90856-2 (1991).
141 Nilsson, J., Stahl, S., Lundeberg, J., Uhlen, M. & Nygren, P. A. Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. *Protein Expr Purif* **11**, 1-16, doi:10.1006/prep.1997.0767 (1997).
142 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
143 Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444, doi:10.1093/nar/gkab1061 (2022).

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

8 Παράρτημα

ΠΑΡΑΡΤΗΜΑ Ι: Αλληλουχίες από τη βάση δεδομένων *Protein* του NCBI

Παρακάτω παρουσιάζονται οι αμινοξικές αλληλουχίες των πρωτεϊνών που μελετήθηκαν στην παρούσα διπλωματική εργασία μαζί με τα χαρακτηριστικά τους, τα οποία προσδιορίστηκαν με το διαδικτυακό εργαλείο ProtParam στον διαδικτυακό διακομιστή ExPASy.

STBD1

```
        10         20         30         40         50         60
MGAVWSALLV GGGLAGALFV WLLRGGPGDT GKDGDAEQEK DAPLGGAAIP GGHQSGSSGL

        70         80         90        100        110        120
SPGPSGQELV TKPEHLQESN GHLISKTKDL GKLQAASWRL QNPSREVCDN SREHVPSGQF

       130        140        150        160        170        180
PDTEAPATSE TSNSRSYSEV SRNESLESPM GEWGFQKGQE ISAKAATCFA EKLPSSNLLK

       190        200        210        220        230        240
NRAKEEMSLS DLNSQDRVDH EEWEMVPRHS SWGDVGVGGS LKAPVLNLNQ GMDNGRSTLV

       250        260        270        280        290        300
EARGQQVHGK MERVAVMPAG SQQVSVRFQV HYVTSTDVQF IAVTGDHECL GRWNTYIPLH

       310        320        330        340        350
YNKDGFWSHS IFLPADTVVE WKFVLVENGG VTRWEECSNR FLETGHEDKV VHAWWGIH
```

Αριθμός αμινοξέων 358

Μοριακό βάρος: 39007,27

Θεωρητικό pI: 5,46

Ala (A)	24	6.7%	ΜΠ	Lys (K)	16	4.5%	Β
Arg (R)	16	4.5%	Β	Met (M)	7	2.0%	ΜΠ
Asn (N)	15	4.2%	Π	Phe (F)	10	2.8%	ΜΠ
Asp (D)	17	4.7%	Ο	Pro (P)	17	4.7%	ΜΠ
Cys (C)	4	1.1%	Π	Ser (S)	36	10.1%	Π
Gln (Q)	17	4.7%	Π	Thr (T)	15	4.2%	Π
Glu (E)	29	8.1%	Ο	Trp (W)	12	3.4%	ΜΠ

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

Gly (G)	40	11.2%	ΜΠ	Tyr (Y)	4	1.1%	Π
His (H)	14	3.9%	Β	Val (V)	29	8.1%	ΜΠ
Ile (I)	7	2.0%	ΜΠ	Pyl (O)	0	0.0%	
Leu (L)	29	8.1%	ΜΠ	Sec (U)	0	0.0%	

Ολικός αριθμός αρνητικά φορτισμένων επικρατειών (Asp + Glu): 46

Ολικός αριθμός θετικά φορτισμένων επικρατειών (Arg + Lys): 32

CBM20(258-358)

```
         10         20         30         40         50         60
MPAGSQQVSV RFQVHYVTST DVQFIAVTGD HECLGRWNTY IPLHYNKDGF WSHSIFLPAD

         70         80         90        100
TVVEWKFVLV ENGGVTRWEE CSNRFLETGH EDKVVHAWWG IH
```

Αριθμός αμινοξέων: 102

Θεωρητικό μοριακό βάρος: 11701.05 Da

Θεωρητικό pI: 5.7

Σύσταση από αμινοξέα

Ala (A)	4	4.0%	ΜΠ	Lys (K)	3	3.0%	Β
Arg (R)	4	4.0%	Β	Met (M)	0	0%	ΜΠ
Asn (N)	4	4.0%	Π	Phe (F)	6	5.9%	ΜΠ
Asp (D)	5	5.0%	Ο	Pro (P)	3	3.0%	ΜΠ
Cys (C)	2	2.0%	Π	Ser (S)	6	5.9%	Π
Gln (Q)	4	4.0%	Π	Thr (T)	7	6.9%	Π
Glu (E)	7	6.9%	Ο	Trp (W)	6	5.9%	ΜΠ
Gly (G)	8	7.9%	ΜΠ	Tyr (Y)	3	3.0%	Π
His (H)	7	6.9%	Β	Val (V)	13	12.9%	ΜΠ
Ile (I)	4	4.0%	ΜΠ	Pyl (O)	0	0.0%	
Leu (L)	5	5.0%	ΜΠ	Sec (U)	0	0.0%	

Ολικός αριθμός αρνητικά φορτισμένων επικρατειών (Asp + Glu): 12
Ολικός αριθμός θετικά φορτισμένων επικρατειών (Arg + Lys): 7

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

ΠΑΡΑΡΤΗΜΑ ΙΙ: ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΑΜΙΝΟΞΕΩΝ

Μέσος όρος ισοτοπικής μάζας των αμινοξέων

Αμινοξύ	*Μέσος Όρος Μάζας*
Αλανίνη (A)	71.0788
Αργινίνη (R)	156.1875
Ασπαραγίνη (N)	114.1038
Ασπαρτικό οξύ (D)	115.0886
Κυστεΐνη (C)	103.1388
Γλουταμινικό οξύ (E)	129.1155
Γλουταμίνη (Q)	128.1307
Γλυκίνη (G)	57.0519
Ιστιδίνη (H)	137.1411
Ισολευκίνη (I)	113.1594
Λευκίνη (L)	113.1594
Λυσίνη (K)	128.1741
Μεθειονίνη (M)	131.1926
Φαινυλαλανίνη (F)	147.1766
Προλίνη (P)	97.1167
Σερίνη (S)	87.0782
Θρεονίνη (T)	101.1051
Τρυπτοφάνη (W)	186.2132
Τυροσίνη (Y)	163.1760
Βαλίνη (V)	99.1326

Βαθμός Υδροφοβηκότητας βάσει των Kyle και Doolittle

Πλερική αλυσίδα	*Δείκτης Υδροφοβικότητας*
Ισολευκίνη (I)	4.5
Βαλίνη (V)	4.2
Λευκίνη (L)	3.8
Φαινυλαλανίνη (F)	2.8
Κυστεϊνη/Κυστίνη (C)	2.5
Μεθειονίνη (M)	1.9
Αλανίνη (A)	1.8
Γλυκίνη (G)	-0.4
Θρεονίνη(T)	-0.7
Τρυπτοφάνη(W)	-0.9
Σερίνη (S)	-0.8
Τυροσίνη (Y)	-1.3
Προλίνη (P)	-1.6
Ιστιδίνη(H)	-3.2
Γλουταμινικό οξύ (E)	-3.5
Γλουταμίνη (Q)	-3.5
Ασπαρτικό οξύ (D)	-3.5
Ασπαραγίνη (N)	-3.5
Λυσίνη (K)	-3.9
Αργινίνη (R)	-4.5

ΠΑΡΑΡΤΗΜΑ ΙΙΙ: ΕΚΚΙΝΗΤΕΣ

Εκκινητές κλωνοποίησης

Name	*Sequence*	*Length*	*Restriction Enzyme*
frw (BamHI) STBD1(CBM20)	CGCGGATCCGCGATGCCGGCGGGTAGCC	28 nts	BamHI
frw (NcoI) STBD1(CBM20)	CATGCCATGGCACCGGCGGGTAGCCAGCAA	30 nts	NcoI
rev (NotI) STBD1(CBM20)	ATTTGCGGCCGCTTAGTGAATACCCCACCACGC	33 nts	NotI
rev (XhoI) STBD1(CBM20)	CCGCTCGAGCGGTTAGTGAATACCCCACCACGC	33 nts	XhoI

ΠΑΡΑΡΤΗΜΑ IV: Αποτελέσματα αλληλούχησης για τα constructs pETM-11(CBM20) και pGEX-6P-1(CBM20)

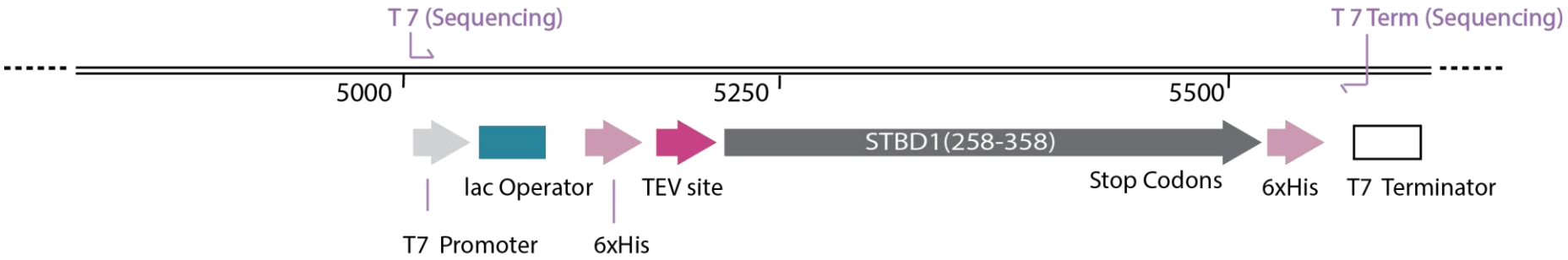
Παρακάτω παρουσιάζονται τα αποτελέσματα αλληλούχισης κατά Sanger, που χρησιμοποιήθηκε για την ακριβή ταυτοποίηση της ένθεσης της επικράτειας CBM20 στους πλασμιδιακούς φορείς pETM-11 και pGEX-6P-1 για τους διαφορετικούς μοριακούς λόγους. Για την αλληλούχιση αυτή χρησιμοποιήθηκαν οικουμενικοί εκκινητές που αντιστοιχούν σε

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

κάθε πλασμίδιο. Η οπτικοποίηση των αποτελεσμάτων πραγματοποιήθηκε με το πρόγραμμα SeqScanner2. Ακολουθείται από μία ανάλυση των ανοιχτών αναγνωστικών πεδίων για κάθε αλυσίδα του κάθε ανασυνδυασμένου πλασμιδίου με διαφορετικό μοριακό λόγο με τη βοήθεια του διαδικτυακού εργαλείου Translation Tool του διαδικτυακού διακομιστή ExPASy.

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

pETM11(CBM20)



Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

T7 Universal Primer

```
                    1
pETM-11(CBM20)      TTAATACG ACTCACTA TAGGGGAA TTGTGAGC GGATAACA ATTCCCCT CTAGAAAT AATTTTGA
Construct 3:1 (2)   -------- -------- -------- ---AGKTK RRMMGTAM ATTCCCCT CTAGAA-T AATTTTG-
Construct 3:1 (4)   -------- -------- -------- --GGKKKK RMSGGTTC ATT-CCCT CTAGAA-T AATTTTG-
Construct 5:1 (2)   -------- -------- -------- -AGKKGWM MGGGTACT TTTCCCCT CTAGAAAT AATTTTG-
Consensus100        NNNNNNNN NNNNNNNN NNNNNNNN NNNDKDDN VVVNDWHH WTTNCCCT CTAGAANT AATTTTGN
```

T7 Promoter | Lac Operator

```
                   65
pETM-11(CBM20)      TTTAACTT TAAGAAGG AGATATAC CATGAAAC ATCACCAT CACCATCA CCCCATGA GCGATTAC
Construct 3:1 (2)   TTTAACTT TAAGAAGG AGATATAC CATGAAAC ATCACCAT CMCYWTCA CCCCATGA GCGATTAC
Construct 3:1 (4)   TTTAACTT TAAGAAGG AGATATAC CATGAAAC ATCACCAT CACCATCA CCCCATGA GCGATTAC
Construct 5:1 (2)   TTTAACTT TAAGAAGG AGATATAC CATGAAAC ATCACCAT CACCATCA CCCCATGA GCGATTAC
Consensus100       TTTAACTT TAAGAAGG AGATATAC CATGAAAC ATCACCAT CMCYWTCA CCCCATGA GCGATTAC
```

His Tag

```
                  129
pETM-11(CBM20)      GACATCCC CACTACTG AGAATCTT TATTTTCA GGGCGCCA TGGCACCG GCGGGTAG CCAGCAAG
Construct 3:1 (2)   GACATCCC CACTACTG AGAATCTT TATTTTCA GGGCGCCA TGGCACCG GCGGGTAG CCAGCAAG
Construct 3:1 (4)   GACATCCC CACTACTG AGAATCTT TATTTTCA GGGCGCCA TGGCACCG GCGGGTAG CCAGCAAG
Construct 5:1 (2)   GACATCCC CACTACTG AGAATCTT TATTTTCA GGGCGCCA TGGCACCG GCGGGTAG CCAGCAAG
Consensus100        GACATCCC CACTACTG AGAATCTT TATTTTCA GGGCGCCA TGGCACCG GCGGGTAG CCAGCAAG
```

TEV Site | CBM20

```
                  193
pETM-11(CBM20)      TGAGCGTT CGTTTCCA GGTGCACT ATGTTACC AGCACCGA CGTGCAAT TTATTGCG GTTACCGG
Construct 3:1 (2)   TGAGCGTT CGTTTCCA GGTGCACT ATGTTACC AGCACCGA CGTGCAAT TTATTGCG GTTACCGG
Construct 3:1 (4)   TGAGCGTT CGTTTCCA GGTGCACT ATGTTACC AGCACCGA CGTGCAAT TTATTGCG GTTACCGG
Construct 5:1 (2)   TGAGCGTT CGTTTCCA GGTGCACT ATGTTACC AGCACCGA CGTGCAAT TTATTGCG GTTACCGG
Consensus100        TGAGCGTT CGTTTCCA GGTGCACT ATGTTACC AGCACCGA CGTGCAAT TTATTGCG GTTACCGG
```

CBM20

```
                  257
pETM-11(CBM20)      TGATCACG AATGCCTG GGCCGTTG GAACACCT ACATCCCG CTGCACTA TAACAAGG ACGGTTTC
Construct 3:1 (2)   TGATCACG AATGCCTG GGCCGTTG GAACACCT ACATCCCG CTGCACTA TAACAAGG ACGGTTTC
Construct 3:1 (4)   TGATCACG AATGCCTG GGCCGTTG GAACACCT ACATCCCG CTGCACTA TAACAAGG ACGGTTTC
Construct 5:1 (2)   TGATCACG AATGCCTG GGCCGTTG GAACACCT ACATCCCG CTGCACTA TAACAAGG ACGGTTTC
Consensus100        TGATCACG AATGCCTG GGCCGTTG GAACACCT ACATCCCG CTGCACTA TAACAAGG ACGGTTTC
```

CBM20

```
                  321
pETM-11(CBM20)      TGGAGCCA CAGCATTT TTCTGCCG GCGGATAC CGTGGTTG AGTGGAAA TTTGTGCT GGTTGAAA
Construct 3:1 (2)   TGGAGCCA CAGCATTT TTCTGCCG GCGGATAC CGTGGTTG AGTGGAAA TTTGTGCT GGTTGAAA
Construct 3:1 (4)   TGGAGCCA CAGCATTT TTCTGCCG GCGGATAC CGTGGTTG AGTGGAAA TTTGTGCT GGTTGAAA
Construct 5:1 (2)   TGGAGCCA CAGCATTT TTCTGCCG GCGGATAC CGTGGTTG AGTGGAAA TTTGTGCT GGTTGAAA
Consensus100        TGGAGCCA CAGCATTT TTCTGCCG GCGGATAC CGTGGTTG AGTGGAAA TTTGTGCT GGTTGAAA
```

CBM20

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
                    385
pETM-11(CBM20)      ACGGTGGC GTTACCCG TTGGGAGG AATGCAGC AACCGCTT TCTGGAGA CCGGTCAC GAAGATAA
Construct 3:1 (2)   ACGGTGGC GTTACCCG TTGGGAGG AATGCAGC AACCGCTT TCTGGAGA CCGGTCAC GAAGATAA
Construct 3:1 (4)   ACGGTGGC GTTACCCG TTGGGAGG AATGCAGC AACCGCTT TCTGGAGA CCGGTCAC GAAGATAA
Construct 5:1 (2)   ACGGTGGC GTTACCCG TTGGGAGG AATGCAGC AACCGCTT TCTGGAGA CCGGTCAC GAAGATAA
Consensus100        ACGGTGGC GTTACCCG TTGGGAGG AATGCAGC AACCGCTT TCTGGAGA CCGGTCAC GAAGATAA
```

CBM20

```
                    449
pETM-11(CBM20)      GGTGGTTC ATGCGTGG TGGGGTAT TCACTAAC CGCTCGAG CACCACCA CCACCACC ACTGAGAT
Construct 3:1 (2)   GGTGGTTC ATGCGTGG TGGGGTAT TCACTAAC CGCTCGAG CACCACCA CCACCACC ACTGAGAT
Construct 3:1 (4)   GGTGGTTC ATGCGTGG TGGGGTAT TCACTAAC CGCTCGAG CACCACCA CCACCACC ACTGAGAT
Construct 5:1 (2)   GGTGGTTC ATGCGTGG TGGGGTAT TCACTAAC CGCTCGAG CACCACCA CCACCACC ACTGAGAT
Consensus100        GGTGGTTC ATGCGTGG TGGGGTAT TCACTAAC CGCTCGAG CACCACCA CCACCACC ACTGAGAT
```

CBM20 His Tag

```
                    513
pETM-11(CBM20)      CCGGCTGC TAACAAAG CCCGAAAG GAAGCTGA GTTGGCTG CTGCCACC GCTGAGCA ATAACTAG
Construct 3:1 (2)   CCGGCTGC TAACAAAG CCCGAAAG GAAGCTGA GTTGGCTG CTGCCACC GCTGAGCA ATAACTAG
Construct 3:1 (4)   CCGGCTGC TAACAAAG CCCGAAAG GAAGCTGA GTTGGCTG CTGCCACC GCTGAGCA ATAACTAG
Construct 5:1 (2)   CCGGCTGC TAACAAAG CCCGAAAG GAAGCTGA GTTGGCTG CTGCCACC GCTGAGCA ATAACTAG
Consensus100        CCGGCTGC TAACAAAG CCCGAAAG GAAGCTGA GTTGGCTG CTGCCACC GCTGAGCA ATAACTAG
```

```
                    577
pETM-11(CBM20)      CATAACCC CTTGGGGC CTCTAAAC GGGTCTTG AGGGGTTT TTTGCTGA AAGGAGGA ACTATATC
Construct 3:1 (2)   CATAACCC CTTGGGGC CTCTAAAC GGGTCTTG AGGGGTTT TTTGCTGA AAGGAGGA ACTATATC
Construct 3:1 (4)   CATAACCC CTTGGGGC CTCTAAAC GGGTCTTG AGGGGTTT TTTGCTGA AAGGAGGA ACTATATC
Construct 5:1 (2)   CATAACCC CTTGGGGC CTCTAAAC GGGTCTTG AGGGGTTT TTTGCTGA AAGGAGGA ACTATATC
Consensus100        CATAACCC CTTGGGGC CTCTAAAC GGGTCTTG AGGGGTTT TTTGCTGA AAGGAGGA ACTATATC
```

T7 Terminator

```
                    641
pETM-11(CBM20)      CGGATTTG GCGAATGG GACGCGCC CTGTAGCG GCGCATTA AGCGCGGC GGGTGTGG TGGTTACG
Construct 3:1 (2)   CGGAT-TG GCGAATGG GACGCGCC CTGTAGCG GCGCATTA AGCGCGGC GGGTGTGG TGGTTACG
Construct 3:1 (4)   CGGAT-TG GCGAATGG GACGCGCC CTGTAGCG GCGCATTA AGCGCGGC GGGTGTGG TGGTTACG
Construct 5:1 (2)   CGGAT-TG GCGAATGG GACGCGCC CTGTAGCG GCGCATTA AGCGCGGC GGGTGTGG TGGTTACG
Consensus100        CGGATNTG GCGAATGG GACGCGCC CTGTAGCG GCGCATTA AGCGCGGC GGGTGTGG TGGTTACG
```

f1 ori

```
                    705
pETM-11(CBM20)      CGCAGCGT GACCGCTA CACTTGCC AGCGCCCT AGCGCCCG CTCCTTTC GCTTTCTT CCCTTCCT
Construct 3:1 (2)   CGCAGCGT GACCGCTA CACTTGCC AGCGCCCT AGCGCCCG CTCCTTTC GCTTTCTT CCCTTCCT
Construct 3:1 (4)   CGCAGCGT GACCGCTA CACTTGCC AGCGCCCT AGCGCCCG CTCCTTTC GCTTTCTT CCCTTCCT
Construct 5:1 (2)   CGCAGCGT GACCGCTA CACTTGCC AGCGCCCT AGCGCCCG CTCCTTTC GCTTTCTT CCCTTCCT
Consensus100        CGCAGCGT GACCGCTA CACTTGCC AGCGCCCT AGCGCCCG CTCCTTTC GCTTTCTT CCCTTCCT
```

f1 ori

```
                    769
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
pETM-11(CBM20)      TTCTCGCC ACGTTCGC CGGCTTTC CCCGTCAA GCTCTAAA TCGGGGGC TCCCTTTA GGGTTCCG
Construct 3:1 (2)   TTCTCGCC ACGTTCGC CGGCTTTC CCCGTCAA GCTCTAAA TCGGGGGC TCCCTTTA GGGTTCCG
Construct 3:1 (4)   TTCTCGCC ACGTTCGC CGGCTTTC CCCGTCAA GCTCTAAA TCGGGGGC TCCCTTTA GGGTTCCG
Construct 5:1 (2)   TTCTCGCC ACGTTCGC CGGCTTTC CCCGTCAA GCTCTAAA TCGGGGGC TCCCTTTA GGGTTCCG
Consensus100        TTCTCGCC ACGTTCGC CGGCTTTC CCCGTCAA GCTCTAAA TCGGGGGC TCCCTTTA GGGTTCCG
```

f1 ori

```
                  833
pETM-11(CBM20)      ATTTAGTG CTTTACGG CACCTCGA CCCCAAAA AACTTGAT TAGGGTGA TGGTTCAC GTAGTGGG
Construct 3:1 (2)   ATTTAGTG CTTTACGG CACCTCGA CCCCAAAA AACTTGAT TAGGGTGA TGGTTCAC GTAGTGGG
Construct 3:1 (4)   ATTTAGTG CTTTACGG CACCTCGA CCCCAAAA AACTTGAT TAGGGTGA TGGTTCAC GTAGTGGG
Construct 5:1 (2)   ATTTAGTG CTTTACGG CACCTCGA CCCCAAAA AACTTGAT TAGGGTGA TGGTTCAC GTAGTGGG
Consensus100        ATTTAGTG CTTTACGG CACCTCGA CCCCAAAA AACTTGAT TAGGGTGA TGGTTCAC GTAGTGGG
```

f1 ori

```
                  897
pETM-11(CBM20)      CCATCGCC CTGATAGA CGGTTTTT CGCCCTTT GACGTTGG AGTCCACG TTCTTTAA TAGTGGAC
Construct 3:1 (2)   CCATCGCC CTGATAGA CGGTTTTT CGCCCTTT GACGTTGG AGTCCACG TTCTTTAA TAGTGGAC
Construct 3:1 (4)   CCATCGCC CTGATAGA CGGTTTTT CGCCCTTT GACGTTGG AGTCCACG TTCTTTAA TAGTGGAC
Construct 5:1 (2)   CCATCGCC CTGATAGA CGGTTTTT CGCCCTTT GACGTTGG AGTCCACG TTCTTTAA TAGTGGAC
Consensus100        CCATCGCC CTGATAGA CGGTTTTT CGCCCTTT GACGTTGG AGTCCACG TTCTTTAA TAGTGGAC
```

f1 ori

```
                  961
pETM-11(CBM20)      TCTTGTTC CAAACTGG AACAACAC TCAACCCT ATCTCGGT CTATTCTT TTGATTTA TAAGGGAT
Construct 3:1 (2)   TCTTGTTC CAAACTGG AACAACAC TCAACCCT ATCTCGGT CTATTCTT T-GATTTA TAAGG-AT
Construct 3:1 (4)   TCTTGTTC CAAACTGG AAC-ACAC TCAACCCT ATCTCGGT CTATTCTT TTGATTTA TAGGG-AT
Construct 5:1 (2)   TCTTGTTC CAAACTGG AACAACAC TCAACCCT ATCTCGGT CTATTCTT TTGATTTA TAAGG-AT
Consensus100        TCTTGTTC CAAACTGG AACNACAC TCAACCCT ATCTCGGT CTATTCTT TNGATTTA TARGGNAT
```

f1 ori

```
                 1025
pETM-11(CBM20)      TTTGCCGA TTTCGGCC TATTGGTT AAAAAATG AGCTGATT TAACAAAA ATTTAACG CGAATTTT
Construct 3:1 (2)   T-TGCCGA TT-CGGC- TATTGGTT AAAAAATG AGCTGATT -A-CAAAA -TT-AACG CGA-TTT-
Construct 3:1 (4)   TTTGCCGA TTTCGGCC TATTGGTT AAAAAATG AGCTGATT -A-CAAAA -TT-A-CG CGAATTTT
Construct 5:1 (2)   TTTGCCGA TTTCG-CC TATTGGTT AAAAAATG AGCTGATT TAACAAAA -TTTAACG CGAATTT-
Consensus100        TNTGCCGA TTNCGNCN TATTGGTT AAAAAATG AGCTGATT NANCAAAA NTTNANCG CGANTTTN
```

f1 ori

```
                 1089
pETM-11(CBM20)      AACAAAAT ATTAACGT TTACAATT TCAGGTGG CACTTTTC GGGGAAAT G-TGCG-C GGAACCCC
Construct 3:1 (2)   A-CAAA-T AT-A-CGC TTACA-TT --AG-TG- CACTTT-C G--GAAAT G-TGS--C -GA-CCC-
Construct 3:1 (4)   A-CAAA-T AT-A-CGC T-ACA-TT --AGGTGG CACTTTTC --GGAA-T GGTGCG-C -GA-CCC-
Construct 5:1 (2)   A-CAAAAT AT-AACGC TTACAATT --AG-TGG CACTTTTC G-GGAA-T G-TGGCGC -GA-CCC-
Consensus100        ANCAAANT ATNANCGY TNACANTT NNAGNTGN CACTTTNC NNNGAANT GNTGSNNC NGANCCCN
```

f1 ori

```
                 1153
pETM-11(CBM20)      TATTTGTT TATTTTTC TAAATACA TTCAAATA TGTATCCG CTCATGAA TTAATTCT TAGAAAAA
Construct 3:1 (2)   TAT--GT- -ATTT--C TAA-TACA T-CAAAT- -GTATC-G CTCATGAA T-AAT-CT -AGAAAT-
Construct 3:1 (4)   TATTGGTT -ATTTT-C TAA-TACA TCCAWWT- TGTATCSG CTCATGA- T-A-TTCT --GAAA--
Construct 5:1 (2)   TATTGGTT -ATTTTTC TAA-TACA T-CAAWTW TGTWTCCG CTCATGA- T-AGT-CT T-GAAAA-
Consensus100        TATNNGTN NATTTNNC TAANTACA TNCAWWTN NGTWTCNG CTCATGAN TNANTNCT NNGAAANN
```

f1 ori

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
                 1217
pETM-11(CBM20)      CTCATCGA GCATCAAA TGAAACTG CAATTTAT TCATATCA GGATTATC AATACCAT ATT--
Construct 3:1 (2)   CT-ATCGA C-ATCA-A TGAAGCTG CAT-GTAT CCATTC-A GGA----- -------- ------
Construct 3:1 (4)   CT-ATCGR C-ATCA-A TGAAGCTG CAG-TTAT CAAATTCA GGATATTC A------- -----
Construct 5:1 (2)   CT-ATCGR C-ATCAA- TGA--CTG CAA-TTAT CCAAWTCA GGGATATC A------- -----
Consensus100        CTNATCGR SNATCANN TGANNCTG CADNKTAT YMAWWYNA GGRNNNNN NNNNNNNN NNNNNN
```

ExPASy translation tool

1. Construct 3:1 (2)

5'3' Frame 1

```
XXXVXSPLE-FCLTLRRRYTMKHHHXXHPMSDYDIPTTENLYFQGAMAPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLH
YNKDGFWSHSIFLPADTVVEWKFVLVENGGVTRWEECSNRFLETGHEDKVVHAWWGIH-PLEHHHHHH-DPAANKARKEAELAAATAEQ-L
A-PLGASKRVLRGFLLKGGTISGLANGTRPVAAH-ARRVWWLRAA-PLHLPAP-RPLLSLSSLPFSPRSPAFPVKL-IGGSL-GSDLVLYG
TSTPKNLIRVMVHVVGHRPDRRFFAL-RWSPRSLIVDSCSKLEQHSTLSRSIL-FIRICRFGYWLKNELITKLTRFTNIRLH-CTFGNVXT
LCISNTSNVSLMNNLEIYRHQ-SCMYPFR
```

2. Construct 3:1 (4)

5'3' Frame 1

```
XXXGSFPLE-FCLTLRRRYTMKHHHHHHPMSDYDIPTTENLYFQGAMAPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLH
YNKDGFWSHSIFLPADTVVEWKFVLVENGGVTRWEECSNRFLETGHEDKVVHAWWGIH-PLEHHHHHH-DPAANKARKEAELAAATAEQ-L
A-PLGASKRVLRGFLLKGGTISGLANGTRPVAAH-ARRVWWLRAA-PLHLPAP-RPLLSLSSLPFSPRSPAFPVKL-IGGSL-GSDLVLYG
TSTPKNLIRVMVHVVGHRPDRRFFAL-RWSPRSLIVDSCSKLEHTQPYLGLFF-FIGILPISAYWLKNELITKLREFYKYTLH-VALFGMV
RDPIGYFLIHPXCIXS-YSETIXINEAAVIKFRIF
```

3. Construct 5:1 (2)

5'3' Frame 1

```
XXXGTFPL-K-FCLTLRRRYTMKHHHHHHPMSDYDIPTTENLYFQGAMAPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPL
HYNKDGFWSHSIFLPADTVVEWKFVLVENGGVTRWEECSNRFLETGHEDKVVHAWWGIH-PLEHHHHHH-DPAANKARKEAELAAATAEQ-
LA-PLGASKRVLRGFLLKGGTISGLANGTRPVAAH-ARRVWWLRAA-PLHLPAP-RPLLSLSSLPFSPRSPAFPVKL-IGGSL-GSDLVLY
GTSTPKNLIRVMVHVVGHRPDRRFFAL-RWSPRSLIVDSCSKLEQHSTLSRSILLIYKDFADFAYWLKNELI-QNLTRIYKI-RLQLVALF
GNVARPYWLFF-YIXXXSAHDSLENYXHQ-LQLSXSGIS
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

T7 term Universal Primer

```
                    1
pETM-11(CBM20)      TTATGCTA GTTATTGC TCAGCGGT GGCAGCAG CCAACTCA GCTTCCTT TCGGGCTT TGTTAGCA
Construct 3:1 (2)   -------- -------- -------- ------GS SMAMMWMM RMTT-CTT TCGGGCTT TGTTAGCA
Construct 3:1 (4)   -------- -------- -------- ----TSSR MMCYWRAC TTCCTTTT TCGGGCTT TGTTAGCA
Construct 5:1 (2)   -------- -------- -------- ------CG SRMMAYYR SCTTYCTT TCGGGCTT TGTTAGCA
Consensus100        NNNNNNNN NNNNNNNN NNNNNNNN NNNNNNVV VVMHHNHV NHYYNYTT TCGGGCTT TGTTAGCA
                   65
pETM-11(CBM20)      GCCGGATC TCAGTGGT GGTGGTGG TGGTGCTC GAGCGGTT AGTGAATA CCCCACCA CGCATGAA
Construct 3:1 (2)   GCCGGATC TCAGTGGT GGTGGTGG TGGTGCTC GAGCGGTT AGTGAATA CCCCACCA CGCATGAA
Construct 3:1 (4)   GCCGGATC TCAGTGGT GGTGGTGG TGGTGCTC GAGCGGTT AGTGAATA CCCCACCA CGCATGAA
Construct 5:1 (2)   GCCGGATC TCAGTGGT GGTGGTGG TGGTGCTC GAGCGGTT AGTGAATA CCCCACCA CGCATGAA
Consensus100        GCCGGATC TCAGTGGT GGTGGTGG TGGTGCTC GAGCGGTT AGTGAATA CCCCACCA CGCATGAA
```

His Tag

CBM20

```
                  129
pETM-11(CBM20)      CCACCTTA TCTTCGTG ACCGGTCT CCAGAAAG CGGTTGCT GCATTCCT CCCAACGG GTAACGCC
Construct 3:1 (2)   CCACCTTA TCTTCGTG ACCGGTCT CCAGAAAG CGGTTGCT GCATTCCT CCCAACGG GTAACGCC
Construct 3:1 (4)   CCACCTTA TCTTCGTG ACCGGTCT CCAGAAAG CGGTTGCT GCATTCCT CCCAACGG GTAACGCC
Construct 5:1 (2)   CCACCTTA TCTTCGTG ACCGGTCT CCAGAAAG CGGTTGCT GCATTCCT CCCAACGG GTAACGCC
Consensus100        CCACCTTA TCTTCGTG ACCGGTCT CCAGAAAG CGGTTGCT GCATTCCT CCCAACGG GTAACGCC
```

CBM20

```
                  193
pETM-11(CBM20)      ACCGTTTT CAACCAGC ACAAATTT CCACTCAA CCACGGTA TCCGCCGG CAGAAAAA TGCTGTGG
Construct 3:1 (2)   ACCGTTTT CAACCAGC ACAAATTT CCACTCAA CCACGGTA TCCGCCGG CAGAAAAA TGCTGTGG
Construct 3:1 (4)   ACCGTTTT CAACCAGC ACAAATTT CCACTCAA CCACGGTA TCCGCCGG CAGAAAAA TGCTGTGG
Construct 5:1 (2)   ACCGTTTT CAACCAGC ACAAATTT CCACTCAA CCACGGTA TCCGCCGG CAGAAAAA TGCTGTGG
Consensus100        ACCGTTTT CAACCAGC ACAAATTT CCACTCAA CCACGGTA TCCGCCGG CAGAAAAA TGCTGTGG
```

CBM20

```
                  257
pETM-11(CBM20)      CTCCAGAA ACCGTCCT TGTTATAG TGCAGCGG GATGTAGG TGTTCCAA CGGCCCAG GCATTCGT
Construct 3:1 (2)   CTCCAGAA ACCGTCCT TGTTATAG TGCAGCGG GATGTAGG TGTTCCAA CGGCCCAG GCATTCGT
Construct 3:1 (4)   CTCCAGAA ACCGTCCT TGTTATAG TGCAGCGG GATGTAGG TGTTCCAA CGGCCCAG GCATTCGT
Construct 5:1 (2)   CTCCAGAA ACCGTCCT TGTTATAG TGCAGCGG GATGTAGG TGTTCCAA CGGCCCAG GCATTCGT
Consensus100        CTCCAGAA ACCGTCCT TGTTATAG TGCAGCGG GATGTAGG TGTTCCAA CGGCCCAG GCATTCGT
```

CBM20

```
                  321
pETM-11(CBM20)      GATCACCG GTAACCGC AATAAATT GCACGTCG GTGCTGGT AACATAGT GCACCTGG AAACGAAC
Construct 3:1 (2)   GATCACCG GTAACCGC AATAAATT GCACGTCG GTGCTGGT AACATAGT GCACCTGG AAACGAAC
Construct 3:1 (4)   GATCACCG GTAACCGC AATAAATT GCACGTCG GTGCTGGT AACATAGT GCACCTGG AAACGAAC
Construct 5:1 (2)   GATCACCG GTAACCGC AATAAATT GCACGTCG GTGCTGGT AACATAGT GCACCTGG AAACGAAC
Consensus100        GATCACCG GTAACCGC AATAAATT GCACGTCG GTGCTGGT AACATAGT GCACCTGG AAACGAAC
```

CBM20

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
                    385
pETM-11(CBM20)      GCTCACTT GCTGGCTA CCCGCCGG TGCCATGG CGCCCTGA AAATAAAG ATTCTCAG TAGTGGGG
Construct 3:1 (2)   GCTCACTT GCTGGCTA CCCGCCGG TGCCATGG CGCCCTGA AAATAAAG ATTCTCAG TAGTGGGG
Construct 3:1 (4)   GCTCACTT GCTGGCTA CCCGCCGG TGCCATGG CGCCCTGA AAATAAAG ATTCTCAG TAGTGGGG
Construct 5:1 (2)   GCTCACTT GCTGGCTA CCCGCCGG TGCCATGG CGCCCTGA AAATAAAG ATTCTCAG TAGTGGGG
Consensus100        GCTCACTT GCTGGCTA CCCGCCGG TGCCATGG CGCCCTGA AAATAAAG ATTCTCAG TAGTGGGG
```

```
                    449
pETM-11(CBM20)      ATGTCGTA ATCGCTCA TGGGGTGA TGGTGATG GTGATGTT TCATGGTA TATCTCCT TCTTAAAG
Construct 3:1 (2)   ATGTCGTA ATCGCTCA TGGGGTGA TGGTGATG GTGATGTT TCATGGTA TATCTCCT TCTTAAAG
Construct 3:1 (4)   ATGTCGTA ATCGCTCA TGGGGTGA TGGTGATG GTGATGTT TCATGGTA TATCTCCT TCTTAAAG
Construct 5:1 (2)   ATGTCGTA ATCGCTCA TGGGGTGA TGGTGATG GTGATGTT TCATGGTA TATCTCCT TCTTAAAG
Consensus100        ATGTCGTA ATCGCTCA TGGGGTGA TGGTGATG GTGATGTT TCATGGTA TATCTCCT TCTTAAAG
```

```
                    513
pETM-11(CBM20)      TTAAATCA AAATTATT TCTAGAGG GGAATTGT TATCCGCT CACAATTC CCCTATAG TGAGTCGT
Construct 3:1 (2)   TTAAA-CA AAATTATT TCTAGAGG GGAATTGT TATCCGCT CACAATTC CCCTATAG TGAGTCGT
Construct 3:1 (4)   TTAAA-CA AAATTATT TCTAGAGG GGAATTGT TATCCGCT CACAATTC CCCTATAG TGAGTCGT
Construct 5:1 (2)   TTAAA-CA AAATTATT TCTAGAGG GGAATTGT TATCCGCT CACAATTC CCCTATAG TGAGTCGT
Consensus100        TTAAANCA AAATTATT TCTAGAGG GGAATTGT TATCCGCT CACAATTC CCCTATAG TGAGTCGT
```

```
                    577
pETM-11(CBM20)      ATTAATTT CGCGGGAT CGAGATCT CGATCCTC TACGCCGG ACGCATCG TGGCCGGC ATCACCGG
Construct 3:1 (2)   ATTAATTT CGCGGGAT CGAGATCT CGATCCTC TACGCCGG ACGCATCG TGGCCGGC ATCACCGG
Construct 3:1 (4)   ATTAATTT CGCGGGAT CGAGATCT CGATCCTC TACGCCGG ACGCATCG TGGCCGGC ATCACCGG
Construct 5:1 (2)   ATTAATTT CGCGGGAT CGAGATCT CGATCCTC TACGCCGG ACGCATCG TGGCCGGC ATCACCGG
Consensus100        ATTAATTT CGCGGGAT CGAGATCT CGATCCTC TACGCCGG ACGCATCG TGGCCGGC ATCACCGG
```

```
                    641
pETM-11(CBM20)      CGCCACAG GTGCGGTT GCTGGCGC CTATATCG CCGACATC ACCGATGG GGAAGATC GGGCTCGC
Construct 3:1 (2)   CGCCACAG GTGCGGTT GCTGGCGC CTATATCG CCGACATC ACCGATGG GGAAGATC GGGCTCGC
Construct 3:1 (4)   CGCCACAG GTGCGGTT GCTGGCGC CTATATCG CCGACATC ACCGATGG GGAAGATC GGGCTCGC
Construct 5:1 (2)   CGCCACAG GTGCGGTT GCTGGCGC CTATATCG CCGACATC ACCGATGG GGAAGATC GGGCTCGC
Consensus100        CGCCACAG GTGCGGTT GCTGGCGC CTATATCG CCGACATC ACCGATGG GGAAGATC GGGCTCGC

                    705
pETM-11(CBM20)      CACTTCGG GCTCATGA GCGCTTGT TTCGGCGT GGGTATGG TGGCAGGC CCCGTGGC CGGGGGAC
Construct 3:1 (2)   CACTTCGG GCTCATGA GCGCTTGT TTCGGCGT GGGTATGG TGGCAGGC CCCGTGGC CGGGGGAC
Construct 3:1 (4)   CACTTCGG GCTCATGA GCGCTTGT TTCGGCGT GGGTATGG TGGCAGGC CCCGTGGC CGGGGGAC
Construct 5:1 (2)   CACTTCGG GCTCATGA GCGCTTGT TTCGGCGT GGGTATGG TGGCAGGC CCCGTGGC CGGGGGAC
Consensus100        CACTTCGG GCTCATGA GCGCTTGT TTCGGCGT GGGTATGG TGGCAGGC CCCGTGGC CGGGGGAC

                    769
pETM-11(CBM20)      TGTTGGGC GCCATCTC CTTGCATG CACCATTC CTTGCGGC GGCGGTGC TCAACGGC CTCAACCT
Construct 3:1 (2)   TGTTGGGC GCCATCTC CTTGCATG CACCATTC CTTGCGGC GGCGGTGC TCAACGGC CTCAACCT
Construct 3:1 (4)   TGTTGGGC GCCATCTC CTTGCATG CACCATTC CTTGCGGC GGCGGTGC TCAACGGC CTCAACCT
Construct 5:1 (2)   TGTTGGGC GCCATCTC CTTGCATG CACCATTC CTTGCGGC GGCGGTGC TCAACGGC CTCAACCT
Consensus100        TGTTGGGC GCCATCTC CTTGCATG CACCATTC CTTGCGGC GGCGGTGC TCAACGGC CTCAACCT
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
                   833
pETM-11(CBM20)       ACTACTGG GCTGCTTC CTAATGCA GGAGTCGC ATAAGGGA GAGCGTCG AGATCCCG GACACCAT
Construct 3:1 (2)    ACTACTGG GCTGCTTC CTAATGCA GGAGTCGC ATAAGGGA GAGCGTCG AGATCCCG GACACCAT
Construct 3:1 (4)    ACTACTGG GCTGCTTC CTAATGCA -GAGTCGC ATAAGGGA GAGCGTCG AGATCCCG GACACCAT
Construct 5:1 (2)    ACTACTGG GCTGCTTC CTAATGCA GGAGTCGC ATAAGGGA GAGCGTCG AGATCCCG GACACCAT
Consensus100         ACTACTGG GCTGCTTC CTAATGCA NGAGTCGC ATAAGGGA GAGCGTCG AGATCCCG GACACCAT

                   897
pETM-11(CBM20)       CGAATGGC GCAAAACC TTTCGCGG TATGGCAT GATAGCGC CCGGAAGA GAGTCAAT TCAGGGTG
Construct 3:1 (2)    CGAATGGC GCAAAACC TTTCGCGG TATGGCAT GATAGCGC CCGGAAGA GAGTCAAT TCAG-GTG
Construct 3:1 (4)    CGAATGGC GCAAAACC TTTCGCGG TATGGCAT GATAGCGC CCGGAAGA GAGTCAAT TCAG-GTG
Construct 5:1 (2)    CGAATGGC GCAAAACC TTTCGCGG TATGGCAT GATAGCGC CCGGAAGA GAGTCAAT TCAG-GTG
Consensus100         CGAATGGC GCAAAACC TTTCGCGG TATGGCAT GATAGCGC CCGGAAGA GAGTCAAT TCAGNGTG

                   961
pETM-11(CBM20)       GTGAATGT GAAACCAG TAACGTTA TACGATGT CGCAGAGT ATGCCGGT GTCTCTTA TCAGACCG
Construct 3:1 (2)    GTGAATGT GAAACCAG TAACGTTA TACGATGT CGCAGAGT ATGCCGGT GTCTCTTA TCAGACCG
Construct 3:1 (4)    GTGAATGT GAAACCAG TAACGTTA TACGATGT CGCAGAGT ATGCCGGT GTCTCTTA TCAGACCG
Construct 5:1 (2)    GTGAATGT GAAACCAG TAACGTTA TACGATGT CGCAGAGT ATGCCGGT GTCTCTTA TCAGACCG
Consensus100         GTGAATGT GAAACCAG TAACGTTA TACGATGT CGCAGAGT ATGCCGGT GTCTCTTA TCAGACCG

                  1025
pETM-11(CBM20)       TTTCCCGC GTGGTGAA CCAGG-CC AGCCACGT TTCTGCGA AAACGCGG GAAAAAGT GGAAGCGG
Construct 3:1 (2)    TTTCCCGC GTG-TGAA CCAGGGCC AGCCACGT TTCTGCGA AA-CGCGG GAAAAAGT GGAAGCGG
Construct 3:1 (4)    TTTCCCGC GTG-TGAA CCAGG-CC AGCCACGT TTCTGCGA AAACGCGG GAAAAAGT GGA-GCGG
Construct 5:1 (2)    TTTCCCGC GTG-TGA- CCAG--CC AGCCACGT T-CTGCGA AA-CGCGG GA-AAAGT GGA-GCGG
Consensus100         TTTCCCGC GTGNTGAN CCAGNNCC AGCCACGT TNCTGCGA AANCGCGG GANAAAGT GGANGCGG

                  1089
pETM-11(CBM20)       CGATGGCG GAGCTGAA TTACATTC CCAACCGC GTGGCACA ACAACTGG CGGGCAAA CAGTCGTT
Construct 3:1 (2)    CGATG-CG -AGCTGA- T-ACATTC C-AAC-GC GTG-CMCA C-AC-TGG CGG-CAAA CAGTCGT-
Construct 3:1 (4)    CGATGGCG -AGCTGA- TTACATTC CCA-CCGC GTG-CACA C-AC-TGG CGG-CAA- CAGTCGT-
Construct 5:1 (2)    CGATGGCG -AGCTGA- T-ACAT-C C-AACCGC GTGGCACA CCAC-TGG CGG-CAAA CAGTCGT-
Consensus100         CGATGNCG NAGCTGAN TNACATNC CNANCNGC GTGNCMCA MNAMNTGG CGGNCAAN CAGTCGTN

                  1153
pETM-11(CBM20)       GCTGATTG GCGTTGCC ACCTCCAG TCTGGCCC TGCACGCG CCGTCGCA AAT----- --------
Construct 3:1 (2)    GCTGAT-G -CGT-GC- AC-TC-AG TCTGAACC TGM-CG-- SCGTCGCA A-T-GTCG CGGCG-AT
Construct 3:1 (4)    GCTGAT-G -CGT-GC- ACCTC-AG TCTGA-CC TG--CACG SCGTCGCA A-TKGTC- CGGGGCAT
Construct 5:1 (2)    GCTGAT-G -CGT-GCC AC-TC-AG TCTGA-C- TGM-CG-- SCGTCGCA A-TKGTC- CGGS--AT
Consensus100         GCTGATNG NCGTNGCN ACNTCNAG TCTGRNCN TGNNCRNN SCGTCGCA ANTNNNNN NNNNNNNN

                  1217
pETM-11(CBM20)       -------- -------- -------- -------- -------- -------- -------
Construct 3:1 (2)    AT-CGGCG --ATCACT GTG-CA-C TAGTGK-T TC-GAATK GGT--GAA -CCCGA
Construct 3:1 (4)    AT-CCGCC CGATCACT GAKCCAGC TG-TGTGT TCC--AT- GGT--GAA -C----
Construct 5:1 (2)    ATTCCGCG --ATCACT GATGCA-C TGGTKGGT TCCG-AT- GGCT-GAA AC--GA-
Consensus100         NNNNNNNN NNNNNNNN NNNNNNNN NNNNNNNN NCNNNATN GGYNNGAA NCNNNNN
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

ExPASy Translation Tool

1. Construct 3:1 (2)

3'5' Frame 1

```
SGSXIRXTSAQ-SPISPRHCDXXRFRLSARISTTVCRQCXHALECISSHRRFHFFPRFAETWLALVHTRETV--ETPAYSATSYNVTGFTF
TT-IDSLPGAIMPYRERFCAIRWCPGSRRSPLCDSCIRKQPSSRLRPLSTAAARNGACKEMAPNSPPATGPATIPTPKQALMSPKWRARSS
PSVMSAI-APATAPVAPVMPATMRPA-RIEISIPRN-YDSL-GNCERITIPL-K-FCLTLRRRYTMKHHHHHHPMSDYDIPTTENLYFQGA
MAPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLHYNKDGFWSHSIFLPADTVVEWKFVLVENGGVTRWEECSNRFLETGH
EDKVVHAWWGIH-PLEHHHHHH-DPAANKARKXXXX
```

2. Construct 3:1 (4)

3'5' Frame 1

```
VHHGTHSWXSDRADMPRXIATXCRSD-GARISTTVAASVCTRWECNQLAIAAPLFPAFSQKRGWPGSHAGNGLIRDTGILCDIV-RYWFHI
HHLN-LSSGRYHAIPRKVLRHSMVSGISTLSLMRLCIRKQPSSRLRPLSTAAARNGACKEMAPNSPPATGPATIPTPKQALMSPKWRARSS
PSVMSAI-APATAPVAPVMPATMRPA-RIEISIPRN-YDSL-GNCERITIPL-K-FCLTLRRRYTMKHHHHHHPMSDYDIPTTENLYFQGA
MAPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLHYNKDGFWSHSIFLPADTVVEWKFVLVENGGVTRWEECSNRFLETGH
EDKVVHAWWGIH-PLEHHHHHH-DPAANKARKRKXXX
```

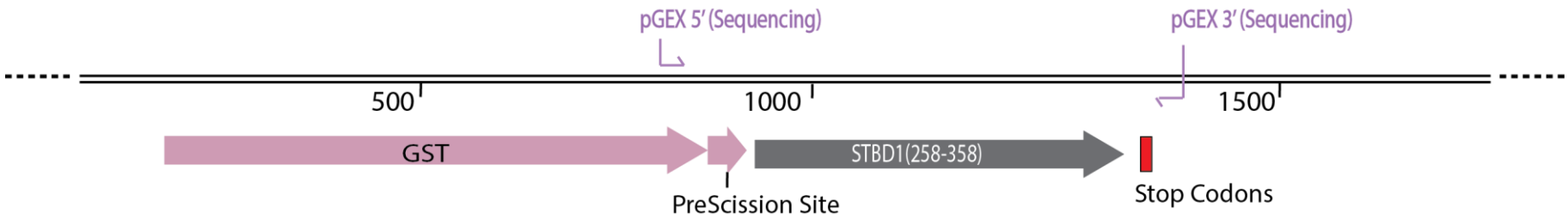
3. Construct 5:1 (2)

3'5' Frame 1

```
SFQPSEXTSASVIAEYXGXLRXXQSD-VARISTTVCRQWCATRLDVSARHRRSTFPAFRRTWLAGHTRETV--ETPAYSATSYNVTGFTFT
T-IDSLPGAIMPYRERFCAIRWCPGSRRSPLCDSCIRKQPSSRLRPLSTAAARNGACKEMAPNSPPATGPATIPTPKQALMSPKWRARSSP
SVMSAI-APATAPVAPVMPATMRPA-RIEISIPRN-YDSL-GNCERITIPL-K-FCLTLRRRYTMKHHHHHHPMSDYDIPTTENLYFQGAM
APAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLHYNKDGFWSHSIFLPADTVVEWKFVLVENGGVTRWEECSNRFLETGHE
DKVVHAWWGIH-PLEHHHHHH-DPAANKARKXXXXX
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

pGEX-6P-1(CBM20)



Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

pGEX 5′ Universal Primer

```
                    1
pGEX-6P-1(CBM20)    AGGGCTGGCA AGCCACGTTT GGTGGTGGCG ACCATCCTCC AAAATCGGAT CTGGAAGTTC
Construct 3:1 (3)   ---------- ---------- ---------- ----GACKWW YAATCGGGAT CTGGA-GTTC
Construct 3:1 (6)   ---------- ---------- ---------- ----AARKWA WCAATCGGAT CTGGA-GTTC
Consensus100        NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNDMVKHH HMAWYSGGAT CTGGANGTTC
```

GST Tag

PreScis-

```
                    61
pGEX-6P-1(CBM20)    TGTTCCAGGG GCCCCTGGGA TCCGCGATGC CGGCGGGTAG CCAGCAAGTG AGCGTTCGTT
Construct 3:1 (3)   TGTTCCAGGG GCCCCTGGGA TCCGCGATGC CGGCGGGTAG CCAGCAAGTG AGCGTTCGTT
Construct 3:1 (6)   TGTTCCAGGG GCCCCTGGGA TCCGCGATGC CGGCGGGTAG CCAGCAAGTG AGCGTTCGTT
Consensus100        TGTTCCAGGG GCCCCTGGGA TCCGCGATGC CGGCGGGTAG CCAGCAAGTG AGCGTTCGTT
```

PreScission site

CBM20

```
                    121
pGEX-6P-1(CBM20)    TCCAGGTGCA CTATGTTACC AGCACCGACG TGCAATTTAT TGCGGTTACC GGTGATCACG
Construct 3:1 (3)   TCCAGGTGCA CTATGTTACC AGCACCGACG TGCAATTTAT TGCGGTTACC GGTGATCACG
Construct 3:1 (6)   TCCAGGTGCA CTATGTTACC AGCACCGACG TGCAATTTAT TGCGGTTACC GGTGATCACG
Consensus100        TCCAGGTGCA CTATGTTACC AGCACCGACG TGCAATTTAT TGCGGTTACC GGTGATCACG
```

CBM20

```
                    181
pGEX-6P-1(CBM20)    AATGCCTGGG CCGTTGGAAC ACCTACATCC CGCTGCACTA TAACAAGGAC GGTTTCTGGA
Construct 3:1 (3)   AATGCCTGGG CCGTTGGAAC ACCTACATCC CGCTGCACTA TAACAAGGAC GGTTTCTGGA
Construct 3:1 (6)   AATGCCTGGG CCGTTGGAAC ACCTACATCC CGCTGCACTA TAACAAGGAC GGTTTCTGGA
Consensus100        AATGCCTGGG CCGTTGGAAC ACCTACATCC CGCTGCACTA TAACAAGGAC GGTTTCTGGA
```

CBM20

```
                    241
pGEX-6P-1(CBM20)    GCCACAGCAT TTTTCTGCCG GCGGATACCG TGGTTGAGTG GAAATTTGTG CTGGTTGAAA
Construct 3:1 (3)   GCCACAGCAT TTTTCTGCCG GCGGATACCG TGGTTGAGTG GAAATTTGTG CTGGTTGAAA
Construct 3:1 (6)   GCCACAGCAT TTTTCTGCCG GCGGATACCG TGGTTGAGTG GAAATTTGTG CTGGTTGAAA
Consensus100        GCCACAGCAT TTTTCTGCCG GCGGATACCG TGGTTGAGTG GAAATTTGTG CTGGTTGAAA
```

CBM20

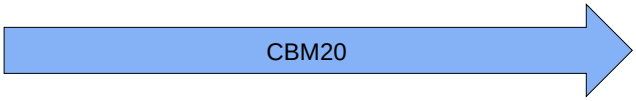
```
                    301
pGEX-6P-1(CBM20)    ACGGTGGCGT TACCCGTTGG GAGGAATGCA GCAACCGCTT TCTGGAGACC GGTCACGAAG
Construct 3:1 (3)   ACGGTGGCGT TACCCGTTGG GAGGAATGCA GCAACCGCTT TCTGGAGACC GGTCACGAAG
Construct 3:1 (6)   ACGGTGGCGT TACCCGTTGG GAGGAATGCA GCAACCGCTT TCTGGAGACC GGTCACGAAG
Consensus100        ACGGTGGCGT TACCCGTTGG GAGGAATGCA GCAACCGCTT TCTGGAGACC GGTCACGAAG
```

CBM20

```
                    361
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
pGEX-6P-1(CBM20)    ATAAGGTGGT TCATGCGTGG TGGGGTATTC ACTAACCGCT CGAGCGGCCG CATCGTGACT
Construct 3:1 (3)   ATAAGGTGGT TCATGCGTGG TGGGGTATTC ACTAACCGCT CGAGCGGCCG CATCGTGACT
Construct 3:1 (6)   ATAAGGTGGT TCATGCGTGG TGGGGTATTC ACTAACCGCT CGAGCGGCCG CATCGTGACT
Consensus100        ATAAGGTGGT TCATGCGTGG TGGGGTATTC ACTAACCGCT CGAGCGGCCG CATCGTGACT
```




```
                   421
pGEX-6P-1(CBM20)    GACTGACGAT CTGCCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC
Construct 3:1 (3)   GACTGACGAT CTGCCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC
Construct 3:1 (6)   GACTGACGAT CTGCCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC
Consensus100        GACTGACGAT CTGCCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC

                   481
pGEX-6P-1(CBM20)    TCCCGGAGAC GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG
Construct 3:1 (3)   TCCCGGAGAC GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG
Construct 3:1 (6)   TCCCGGAGAC GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG
Consensus100        TCCCGGAGAC GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG

                   541
pGEX-6P-1(CBM20)    GCGCGTCAGC GGGTGTTGGC GGGTGTCGGG GCGCAGCCAT GACCCAGTCA CGTAGCGATA
Construct 3:1 (3)   GCGCGTCAGC GGGTGTTGGC GGGTGTCGGG GCGCAGCCAT GACCCAGTCA CGTAGCGATA
Construct 3:1 (6)   GCGCGTCAGC GGGTGTTGGC GGGTGTCGGG GCGCAGCCAT GACCCAGTCA CGTAGCGATA
Consensus100        GCGCGTCAGC GGGTGTTGGC GGGTGTCGGG GCGCAGCCAT GACCCAGTCA CGTAGCGATA

                   601
pGEX-6P-1(CBM20)    GCGGAGTGTA TAATTCTTGA AGACGAAAGG GCCTCGTGAT ACGCCTATTT TTATAGGTTA
Construct 3:1 (3)   GCGGAGTGTA TAATTCTTGA AGACGAAAGG GCCTCGTGAT ACGCCTATTT TTATAGGTTA
Construct 3:1 (6)   GCGGAGTGTA TAATTCTTGA AGACGAAAGG GCCTCGTGAT ACGCCTATTT TTATAGGTTA
Consensus100        GCGGAGTGTA TAATTCTTGA AGACGAAAGG GCCTCGTGAT ACGCCTATTT TTATAGGTTA

                   661
pGEX-6P-1(CBM20)    ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG
Construct 3:1 (3)   ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG
Construct 3:1 (6)   ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG
Consensus100        ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG

                   721
pGEX-6P-1(CBM20)    GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
Construct 3:1 (3)   GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
Construct 3:1 (6)   GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
Consensus100        GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT

                   781
pGEX-6P-1(CBM20)    AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC
Construct 3:1 (3)   AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC
Construct 3:1 (6)   AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC
Consensus100        AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC

                   841
pGEX-6P-1(CBM20)    GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
Construct 3:1 (3)   GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
Construct 3:1 (6)   GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
Consensus100        GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
                   901
pGEX-6P-1(CBM20)   CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC
Construct 3:1 (3)  CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC
Construct 3:1 (6)  CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC
Consensus100       CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC

                   961
pGEX-6P-1(CBM20)   TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA
Construct 3:1 (3)  TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA
Construct 3:1 (6)  TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA
Consensus100       TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA

                  1021
pGEX-6P-1(CBM20)   TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTTGAC GCCGGGCAAG
Construct 3:1 (3)  TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGT-GAC GCCGG-CA-G
Construct 3:1 (6)  TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGT-GAC GCCGG-CAGG
Consensus100       TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTNGAC GCCGGNCANG

                  1081
pGEX-6P-1(CBM20)   AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC -TCACCAGTC
Construct 3:1 (3)  AGCAACTCG- TCGCCGCATA CACTATTCTC AGA-TGACTT G-T-GAGTAC -TCACCAGTC
Construct 3:1 (6)  AGCA-CTCGG TCGC-GCATA CACTATTCTC AGA-TGACTT GGTTGAGTAC CTCACCAGTC
Consensus100       AGCANCTCGN TCGCNGCATA CACTATTCTC AGANTGACTT GNTNGAGTAC NTCACCAGTC

                  1141
pGEX-6P-1(CBM20)   ACAGAAAAGC ATCTTACGGA -TGGCATGAC AGTAAGAGAA TTATGCAGTG -CTGCCATAA
Construct 3:1 (3)  CMAGAAAAGC ATCT-ACG-A ATG-CATGAC AGTA-GAGA- T-ATGCA-TG GCTGC-AT-A
Construct 3:1 (6)  MMAGAAAAGC ATCT-ACGGA ATGGCATGAC AGTA-GAGA- TTATGCA-TG CCTGC-ATTA
Consensus100       MMAGAAAAGC ATCTNACGNA NTGNCATGAC AGTANGAGAN TNATGCANTG NCTGCNATNA

                  1201
pGEX-6P-1(CBM20)   CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGA-CAA CGATCGGAGG A-CCG-AAGG
Construct 3:1 (3)  CCATGA-TGA TAMCTGCGC- -CCAACT-AC TTCTGA-CAG CGATCGGAAG AACCGGAAG-
Construct 3:1 (6)  ACATGA-TG- TAAAMCTGCG TMCA-CT-AC TTCTGAACAM -GATCGGGRG GACCT-AAGT
Consensus100       MCATGANTGN TAMMHSYGCN NMCANCTNAC TTCTGANCAV NGATCGGRRG RNCCKNAAGN

                  1261
pGEX-6P-1(CBM20)   AGCTAACCGC TTTTTTGCAC AACATGGGGG ATCATGTAA
Construct 3:1 (3)  ---------- ---------- ---------- -------
Construct 3:1 (6)  A-C------- ---------- ---------- ---------
Consensus100       NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

ExPASy Translation Tool

1. Construct 3:1 (3)

5'3' Frame 1

```
DXXSGSGVLFQGPLGSAMPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLHYNKDGFWSHSIFLPADTVVEWKFVLVENGG
VTRWEECSNRFLETGHEDKVVHAWWGIH-PLERPHRD-LTICLARFGDDGENL-HMQLPETVTACL-ADAGSRQARQGASAGVGGCRGAAM
TQSRSDSGVYNS-RRKGLVIRLFL-VNVMIIMVS-TSGGTFRGNVRGTPICLFF-IHSNMYPLMRQ-P--MLQ-Y-KRKSMSIQHFRVALI
PFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRVTPAEQLVAAYTILR-LVSTH
QSXKSIYECMTVEICMAAYHDDXCAQLLLTAIGRTG
```

2. Construct 3:1 (6)

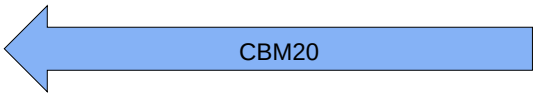
5'3' Frame 1

```
XXXIGSGVLFQGPLGSAMPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLHYNKDGFWSHSIFLPADTVVEWKFVLVENGG
VTRWEECSNRFLETGHEDKVVHAWWGIH-PLERPHRD-LTICLARFGDDGENL-HMQLPETVTACL-ADAGSRQARQGASAGVGGCRGAAM
TQSRSDSGVYNS-RRKGLVIRLFL-VNVMIIMVS-TSGGTFRGNVRGTPICLFF-IHSNMYPLMRQ-P--MLQ-Y-KRKSMSIQHFRVALI
PFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRVTPAGALGRAYTILR-LG-VP
HQXXKSIYGMA-Q-RLCMPALT-CKXAXTTSEXDRXDLS
```

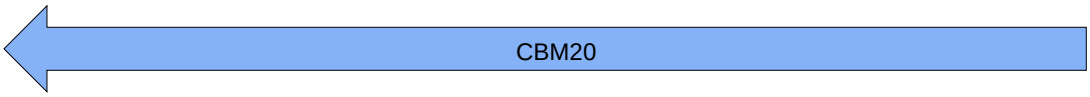

Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

pGEX 3′ Universal Primer

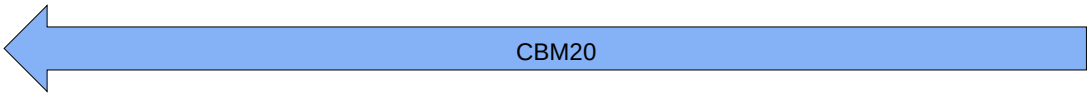
```
                    1
pGEX-6P-1(CBM20)    GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCAT-CACCG AAACGCGCGA GGCAGATCGT
Construct 3:1 (3)   ---------- ---------- --------GG RR-TMCACCG TAACGCGCGA GGCAGATCGT
Construct 3:1 (6)   ---------- ---------- --------GG RRATACAACG TAACGCGCGA GGCAGATCGT
Consensus100        NNNNNNNNNN NNNNNNNNNN NNNNNNNNSG DVNTNCAMCG WAACGCGCGA GGCAGATCGT

                   61
pGEX-6P-1(CBM20)    CAGTCAGTCA CGATGCGGCC GCTCGAGCGG TTAGTGAATA CCCCACCACG CATGAACCAC
Construct 3:1 (3)   CAGTCAGTCA CGATGCGGCC GCTCGAGCGG TTAGTGAATA CCCCMCCACG CATGAACCAC
Construct 3:1 (6)   CAGTCAGTCA CGATGCGGCC GCTCGAGCGG TTAGTGAATA CCCCACCACG CATGAACCAC
Consensus100        CAGTCAGTCA CGATGCGGCC GCTCGAGCGG TTAGTGAATA CCCCMCCACG CATGAACCAC
```



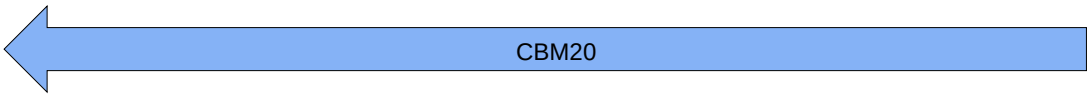

```
                  121
pGEX-6P-1(CBM20)    CTTATCTTCG TGACCGGTCT CCAGAAAGCG GTTGCTGCAT TCCTCCCAAC GGGTAACGCC
Construct 3:1 (3)   CTTATCTTCG TGACCGGTCT CCAGAAAGCG GTTGCTGCAT TCCTCCCAAC GGGTAACGCC
Construct 3:1 (6)   CTTATCTTCG TGACCGGTCT CCAGAAAGCG GTTGCTGCAT TCCTCCCAAC GGGTAACGCC
Consensus100        CTTATCTTCG TGACCGGTCT CCAGAAAGCG GTTGCTGCAT TCCTCCCAAC GGGTAACGCC
```



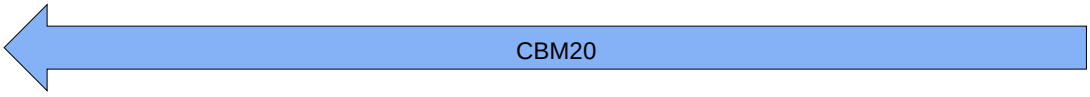

```
                  181
pGEX-6P-1(CBM20)    ACCGTTTTCA ACCAGCACAA ATTTCCACTC AACCACGGTA TCCGCCGGCA GAAAAATGCT
Construct 3:1 (3)   ACCGTTTTCA ACCAGCACAA ATTTCCACTC AACCACGGTA TCCGCCGGCA GAAAAATGCT
Construct 3:1 (6)   ACCGTTTTCA ACCAGCACAA ATTTCCACTC AACCACGGTA TCCGCCGGCA GAAAAATGCT
Consensus100        ACCGTTTTCA ACCAGCACAA ATTTCCACTC AACCACGGTA TCCGCCGGCA GAAAAATGCT
```




```
                  241
pGEX-6P-1(CBM20)    GTGGCTCCAG AAACCGTCCT TGTTATAGTG CAGCGGGATG TAGGTGTTCC AACGGCCCAG
Construct 3:1 (3)   GTGGCTCCAG AAACCGTCCT TGTTATAGTG CAGCGGGATG TAGGTGTTCC AACGGCCCAG
Construct 3:1 (6)   GTGGCTCCAG AAACCGTCCT TGTTATAGTG CAGCGGGATG TAGGTGTTCC AACGGCCCAG
Consensus100        GTGGCTCCAG AAACCGTCCT TGTTATAGTG CAGCGGGATG TAGGTGTTCC AACGGCCCAG
```




```
                  301
pGEX-6P-1(CBM20)    GCATTCGTGA TCACCGGTAA CCGCAATAAA TTGCACGTCG GTGCTGGTAA CATAGTGCAC
Construct 3:1 (3)   GCATTCGTGA TCACCGGTAA CCGCAATAAA TTGCACGTCG GTGCTGGTAA CATAGTGCAC
Construct 3:1 (6)   GCATTCGTGA TCACCGGTAA CCGCAATAAA TTGCACGTCG GTGCTGGTAA CATAGTGCAC
Consensus100        GCATTCGTGA TCACCGGTAA CCGCAATAAA TTGCACGTCG GTGCTGGTAA CATAGTGCAC
```




```
                   361
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
pGEX-6P-1(CBM20)    CTGGAAACGA ACGCTCACTT GCTGGCTACC CGCCGGCATC GCGGATCCCA GGGGCCCCTG
Construct 3:1 (3)   CTGGAAACGA ACGCTCACTT GCTGGCTACC CGCCGGCATC GCGGATCCCA GGGGCCCCTG
Construct 3:1 (6)   CTGGAAACGA ACGCTCACTT GCTGGCTACC CGCCGGCATC GCGGATCCCA GGGGCCCCTG
Consensus100        CTGGAAACGA ACGCTCACTT GCTGGCTACC CGCCGGCATC GCGGATCCCA GGGGCCCCTG
```

CBM20

PreScis-

```
                 421
pGEX-6P-1(CBM20)    GAACAGAACT TCCAGATCCG ATTTTGGAGG ATGGTCGCCA CCACCAAACG TGGCTTGCCA
Construct 3:1 (3)   GAACAGAACT TCCAGATCCG ATTTTGGAGG ATGGTCGCCA CCACCAAACG TGGCTTGCCA
Construct 3:1 (6)   GAACAGAACT TCCAGATCCG ATTTTGGAGG ATGGTCGCCA CCACCAAACG TGGCTTGCCA
Consensus100        GAACAGAACT TCCAGATCCG ATTTTGGAGG ATGGTCGCCA CCACCAAACG TGGCTTGCCA
```

PreScission site

GST Tag

```
                 481
pGEX-6P-1(CBM20)    GCCCTGCAAA GGCCATGCTA TATACTTGCT GGATTTCAAG TACTTATCAA TTTGTGGGAT
Construct 3:1 (3)   GCCCTGCAAA GGCCATGCTA TATACTTGCT GGATTTCAAG TACTTATCAA TTTGTGGGAT
Construct 3:1 (6)   GCCCTGCAAA GGCCATGCTA TATACTTGCT GGATTTCAAG TACTTATCAA TTTGTGGGAT
Consensus100        GCCCTGCAAA GGCCATGCTA TATACTTGCT GGATTTCAAG TACTTATCAA TTTGTGGGAT
```

GST Tag

```
                 541
pGEX-6P-1(CBM20)    AGCTTCAATA CGTTTTTTAA AACAAACTAA TTTTGGGAAC GCATCCAGGC ACATTGGGTC
Construct 3:1 (3)   AGCTTCAATA CGTTTTTTAA AACAAACTAA TTTTGGGAAC GCATCCAGGC ACATTGGGTC
Construct 3:1 (6)   AGCTTCAATA CGTTTTTTAA AACAAACTAA TTTTGGGAAC GCATCCAGGC ACATTGGGTC
Consensus100        AGCTTCAATA CGTTTTTTAA AACAAACTAA TTTTGGGAAC GCATCCAGGC ACATTGGGTC
```

GST Tag

```
                 601
pGEX-6P-1(CBM20)    CATGTATAAA ACAACATCAA GAGCGTCATA CAACATGAAG TCAGGATGGG TTACATGATC
Construct 3:1 (3)   CATGTATAAA ACAACATCAA GAGCGTCATA CAACATGAAG TCAGGATGGG TTACATGATC
Construct 3:1 (6)   CATGTATAAA ACAACATCAA GAGCGTCATA CAACATGAAG TCAGGATGGG TTACATGATC
Consensus100        CATGTATAAA ACAACATCAA GAGCGTCATA CAACATGAAG TCAGGATGGG TTACATGATC
```

GST Tag

```
                 661
pGEX-6P-1(CBM20)    ACCATTTAAA TATGTTTTAT GACATAAACG ATCTTCGAAC ATTTTCAGCA TTTCAGGTAG
Construct 3:1 (3)   ACCATTTAAA TATGTTTTAT GACATAAACG ATCTTCGAAC ATTTTCAGCA TTTCAGGTAG
Construct 3:1 (6)   ACCATTTAAA TATGTTTTAT GACATAAACG ATCTTCGAAC ATTTTCAGCA TTTCAGGTAG
Consensus100        ACCATTTAAA TATGTTTTAT GACATAAACG ATCTTCGAAC ATTTTCAGCA TTTCAGGTAG
```

GST Tag

```
                 721
pGEX-6P-1(CBM20)    CTTGCTAAGA AAATCAACTT TGAGAGTTTC AAAGTCTTTA CTATATGCAA TTCTCGAAAC
Construct 3:1 (3)   CTTGCTAAGA AAATCAACTT TGAGAGTTTC AAAGTCTTTA CTATATGCAA TTCTCGAAAC
Construct 3:1 (6)   CTTGCTAAGA AAATCAACTT TGAGAGTTTC AAAGTCTTTA CTATATGCAA TTCTCGAAAC
Consensus100        CTTGCTAAGA AAATCAACTT TGAGAGTTTC AAAGTCTTTA CTATATGCAA TTCTCGAAAC
```

GST Tag

```
                 781
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
pGEX-6P-1(CBM20)    ACCGTATCTA ATATCCAAAA CCGCTCCTTC AAGCATTGAA ATCTCTGCAC GCTCTTTTGG
Construct 3:1 (3)   ACCGTATCTA ATATCCAAAA CCGCTCCTTC AAGCATTGAA ATCTCTGCAC GCTCTTTTGG
Construct 3:1 (6)   ACCGTATCTA ATATCCAAAA CCGCTCCTTC AAGCATTGAA ATCTCTGCAC GCTCTTTTGG
Consensus100        ACCGTATCTA ATATCCAAAA CCGCTCCTTC AAGCATTGAA ATCTCTGCAC GCTCTTTTGG
```

GST Tag

```
                   841
pGEX-6P-1(CBM20)    ACAACCACCC AACATGTTGT GCTTGTCAGC TATATAACGT ATGATGGCCA TAGACTGTGT
Construct 3:1 (3)   ACAACCACCC AACATGTTGT GCTTGTCAGC TATATAACGT ATGATGGCCA TAGACTGTGT
Construct 3:1 (6)   ACAACCACCC AACATGTTGT GCTTGTCAGC TATATAACGT ATGATGGCCA TAGACTGTGT
Consensus100        ACAACCACCC AACATGTTGT GCTTGTCAGC TATATAACGT ATGATGGCCA TAGACTGTGT
```

GST Tag

```
                   901
pGEX-6P-1(CBM20)    TAATTTAACA TCACCATCAA TATAATAAGG AAGATTGGGA AACTCCAAAC CCAATTCAAA
Construct 3:1 (3)   TAATTTAACA TCACCATCAA TATAATAAGG AAGATTGGGA AACTCCAA-C CCAATTCA-A
Construct 3:1 (6)   TAATTTAACA TCACCATCAA TATAATAAGG AAGATTGGGA AACTCCAA-C CCAATTCAAA
Consensus100        TAATTTAACA TCACCATCAA TATAATAAGG AAGATTGGGA AACTCCAANC CCAATTCANA
```

GST Tag

```
                   961
pGEX-6P-1(CBM20)    CTTTTTGTTT CGCCATTTAT CACCTTCATC GCGCTCATAC AAATGCTCTT CATATTTTTC
Construct 3:1 (3)   CTTTTTGTTT CGCCATTTAT CACCTTCATC GCGCTCATAC AA-TGCTCTT CATATTTTTC
Construct 3:1 (6)   CTTTTTGTTT CGCCATTTAT CACCTTCATC GCGCTCATAC AA-TGCTCTT CATATTTTTC
Consensus100        CTTTTTGTTT CGCCATTTAT CACCTTCATC GCGCTCATAC AANTGCTCTT CATATTTTTC
```

GST Tag

```
                  1021
pGEX-6P-1(CBM20)    TTCAAGATAT TCCAAAAGAA GTCGAGTGGG TTGCACAAGG CCCTTAATTT TCCAATAACC
Construct 3:1 (3)   TTCA-GATAT TCCAAAAGAA GTCGAGTGGG TTGCACAAG- CC-TTTATTT -CCA-TAACC
Construct 3:1 (6)   TTCA-GATAT TCCAAAAGAA GTCGAGTGGG TTGCACA-GG CCCTTAATTT TCCAATAACC
Consensus100        TTCANGATAT TCCAAAAGAA GTCGAGTGGG TTGCACANGN CCNTTWATTT NCCANTAACC
```

GST Tag

```
                  1081
pGEX-6P-1(CBM20)    TAGTATAGGG GACATGAATA CTGTTTCCTG TG-TGAAATT GTTATCCGC- TCACAATTCC
Construct 3:1 (3)   TAGTATAGGG -AMATGAATA CTGTTCC-TG KGGTGAA-TT GTTATCCGCC TCACA-T-CC
Construct 3:1 (6)   TAGTATAGGG -AMATGAATA CTGTTTC-TG TG-TGAAATT GTTATC-GC- TCACA-T-C-
Consensus100        TAGTATAGGG NAMATGAATA CTGTTYCNTG KGNTGAANTT GTTATCNGCN TCACANTNCN
```

GST Tag

```
                  1141
pGEX-6P-1(CBM20)    ACACATTATA CGAGCCGATG ATTAATT-GT CAACAGCTCA TTTCAGAATA TTTGCCAGAA
Construct 3:1 (3)   MCACATTATA CGAGTCGAA- -TGAATATGT CA-CAGCTCA T-TC-GA-TA T-TGGCAGAT
Construct 3:1 (6)   ACACATTATA CGAGCG-ATG AT--ATGT-- CA-CAGCTCA TT-CAGA-TA C-TGGCAGAT
Consensus100        MCACATTATA CGAGYSNAWN NTNNATDNNN CANCAGCTCA TNNCNGANTA YNTGSCAGAW

                  1201
pGEX-6P-1(CBM20)    CCGTTATGAT GT-CGGCGCA AAAAACATTA TCCAGAACGG GAGTGC---- ----------
Construct 3:1 (3)   CGGT-ATGA- GTTCGGC-CA AAAA-CTT-A TCGAACGGRA GTGCCCTTGA GCGCACCGTA
Construct 3:1 (6)   CGGT-ATGA- GT-CGGC-CA ATAA-CT--A TCGACTG--- -GAGTGCCTT GAGCGAACCG
Consensus100        CSGTNATGAN GTNCGGCNCA AWAANCWNNA TCSAVHRNNN NDRBBSNNNN NNNNNNNNNN
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84

```
                 1261
pGEX-6P-1(CBM20)    ---
Construct 3:1 (3)   AT----
Construct 3:1 (6)   AAATWTGC
Consensus100        NNNNNNGC
```

ExPASy Translation Tool

1.Construct 3:1(3)

3'5' Frame 2

```
LRCAQGHXRSISFWPNSYRSANIE-AVTYSFDSYNVXM-GG-QFTXGTVFXSLY-VMEIKACATHSTSFGISEEKYEEHCMSAMKVINGET
KS-IGLEFPNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFED
RLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLEVLFQ
GPLGSAMPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLHYNKDGFWSHSIFLPADTVVEWKFVLVENGGVTRWEECSNRF
LETGHEDKVVHAWXGIH-PLERPHRD-LTICLARYGXX
```

2. Construct 3:1 (6)

3'5' Frame 1

```
AXFGSLKALQSIVIGRLIPICQYLNEL-HIIARIMCDVSDNNFTQKQYSXPYTRLLEN-GPVQPTRLLLEYLKKNMKSIV-AR-R--MAKQ
KV-IGLEFPNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFED
RLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLEVLFQ
GPLGSAMPAGSQQVSVRFQVHYVTSTDVQFIAVTGDHECLGRWNTYIPLHYNKDGFWSHSIFLPADTVVEWKFVLVENGGVTRWEECSNRF
LETGHEDKVVHAWWGIH-PLERPHRD-LTICLARYVVX
```


Institutional Repository - Library & Information Centre - University of Thessaly
17/02/2026 17:48:27 EET - 18.97.14.84