



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΒΙΟΛΟΓΙΑΣ, ΤΟΜΕΑΣ ΒΟΤΑΝΙΚΗΣ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΔΙΠΛΩΜΑ ΕΙΔΙΚΕΥΣΗΣ
«ΜΙΚΡΟΒΙΑΚΗ ΒΙΟΤΕΧΝΟΛΟΓΙΑ»

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Βιοσύνθεση και απομόνωση επιλεγμένων
κυκλικών πεπτιδίων σε καλλιέργειες
βακτηριακών κυττάρων**

Αλέξανδρος Καρυόλαιμος
Χημικός

ΥΠΕΥΘΥΝΟΣ
Δημήτρης Γ. Χατζηνικολάου

Αθήνα 2014



NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

FACULTY OF BIOLOGY, DEPARTMENT OF BOTANY

MSc DIPLOMA

«MICROBIAL BIOTECHNOLOGY»

MSc DIPLOMA THESIS

**Biosynthesis and isolation of cyclic peptides from
bacterial cell cultures**

Alexander Karyolaimos

Chemist

SUPERVISOR

Dimitris G. Hatzinikolaou

Athens 2014

*Στη Χριστίνα Καλπακίωτη και
Αγλαΐα Καρυολαίμου*

Η παρούσα Μεταπτυχιακή Εργασία εκπονήθηκε στα πλαίσια του μεταπτυχιακού προγράμματος Μικροβιακής Βιοτεχνολογίας του Τομέα Βοτανικής, του Τμήματος Βιολογίας, του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών.

ΤΡΙΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

1. Γεώργιος Διαλλινάς, Καθηγητής

Τμήμα Βιολογίας Τομέας Βοτανικής
Πανεπιστήμιο Αθηνών

2. Κωνσταντίνος Ε. Βοργιάς Καθηγητής

Τμήμα Βιολογίας, Τομέας Βιοχημείας και Μοριακής Βιολογίας
Πανεπιστήμιο Αθηνών

3. Δημήτριος Γ. Χατζηνικολάου, Επίκουρος Καθηγητής (Επιβλέπων)

Τμήμα Βιολογίας, Τομέας Βοτανικής
Πανεπιστήμιο Αθηνών

Αντί Προλόγου

Η παρούσα διπλωματική εργασία πραγματοποιήθηκε στο εργαστήριο της Μικροβιακής Βιοτεχνολογίας του τομέα Βοτανικής του τμήματος Βιολογίας, υπό την εποπτεία του επίκουρου καθηγητή Δ.Γ.Χατζηνικολάου. Ολοκληρώνοντας την μεταπτυχιακή μου διατριβή, θα ήθελα να ευχαριστήσω όλους τους καθηγητές και φίλους που με βοήθησαν και με ενθάρρυναν στην προσπάθειά μου αυτή.

Πρώτα από όλα θα ήθελα να ευχαριστήσω τους καθηγητές Α.Καραγκούνη, Γ.Διαλλινά και Δ.Γ.Χατζηνικολάου που μου έδωσαν τη δυνατότητα να συμμετάσχω στο Μ.Δ.Ε. «Μικροβιακή Βιοτεχνολογία», στο οποίο κατάφερα να αναπτύξω κριτική σκέψη ως προς ποικίλα επιστημονικά θέματα. Θα ήθελα να ευχαριστήσω από καρδιάς, τον επιβλέποντα της διατριβής μου Καθηγητή Δ.Γ.Χατζηνικολάου, όχι μόνο για τη εμπιστοσύνη που μου έδειξε αναθέτοντας μου μια τόσο ενδιαφέρουσα μελέτη, αλλά και για την βοήθειά του στο σχεδιασμό των πειραμάτων και την ερμηνεία των αποτελεσμάτων. Οι συμβουλές του και οι συζητήσεις μας, πάνω από όλα, μου δίδαξαν ένα νέο για εμένα τρόπο ερευνητικής σκέψης, κάτι που αποτελεί πολύτιμο εφόδιο στην ζωή μου.

Ένα μεγάλο ευχαριστώ οφείλω, στο κ. Γ.Διαλλινά για την πολύτιμη γνώση που μου πρόσφερε μέσω των διαλέξεων του Μεταπτυχιακού αλλά και προσωπικών συζητήσεων, που χωρίς αυτή δεν θα μπορούσα να τελειοποιήσω την διπλωματική μου εργασία επιτυχώς. Επιπλέον οι προσωπικές συζητήσεις μας, έχουν αποτελέσει μεγάλο κίνητρο για τις μετέπειτα μου επιστημονικές βλέψεις. Θα ήθελα να ευχαριστήσω επίσης και τον κ. Σ.Κατσίφα για τη βοήθεια που μου πρόσφερε σε διάφορα ζητήματα. Ευχαριστώ ιδιαίτερα τους φίλους και συνεργάτες που μοιραστήκαμε τον ίδιο εργαστηριακό χώρο Σ.Δελέντα, Τ.Γεωργούλη, Δ. Γεωργιάδου και Γ.Καπετανάκη για τη τρομερή στήριξη τους στις δύσκολες στιγμές μέσα στο εργαστήριο, όπως και τους συμφοιτητές μου Α.Παπαγεωργίου, Ν.Μάνο, Α.Δούκα, Γ.Παπαναστασίου και Μ.Βαλασοπούλου για τις ευχάριστες στιγμές κατά διάρκεια του Μεταπτυχιακού.

Ευχαριστώ θερμά την οικογένειά μου, τους γονείς Γίαννη και Φωτεινή, την αδερφή μου Χριστίνα, όχι μόνο για την ενίσχυση που μου προσέφεραν σε όλους τους τομείς, αλλά και για την αμέριστη κατανόηση και στήριξη σε όλες τις αποφάσεις που έχω πάρει. Τέλος θα ήθελα να ευχαριστήσω ιδιαίτερα τη σύντροφό μου Ειρήνη, γιατί είναι πάντα εκεί για εμένα, όποτε την έχω χρειαστεί.

ΠΕΡΙΕΧΟΜΕΝΑ

1. Εισαγωγή	1
1.1 Ιντεΐνες	3
1.1.1 Η τεράστια φυλογένεση των ιντεϊνών	6
1.1.2 Η σποραδική εξάπλωση των ιντεϊνών σε συνδυασμό με την εξελικτική ιστορία των ξενιστών τους	8
1.1.3 Η διάδοση των ιντεϊνών είναι κατευθυνόμενη προς εξειδικευμένες πρωτεΐνες.....	13
1.1.4 Οι ιντεΐνες βρίσκονται συνήθως στην λειτουργική περιοχή των πρωτεϊνών	14
1.1.5 Η εμφάνιση των διχασμένων ιντεϊνών.....	16
1.2 Πρωτεϊνικό μάτισμα: ο μηχανισμός απεμπλοκής της ιντεΐνης από την πρόδρομη πρωτεΐνη.....	18
1.2.1 Μηχανισμός πρωτεϊνικού ματίσματος κατηγορίας 1 ιντεϊνών:1 ^ο βήμα	22
1.2.2 Μηχανισμός πρωτεϊνικού ματίσματος κατηγορίας 1 ιντεϊνών:2 ^ο βήμα	23
1.2.3 Μηχανισμός πρωτεϊνικού ματίσματος κατηγορίας 1 ιντεϊνών:3 ^ο βήμα	23
1.2.4 Μηχανισμός πρωτεϊνικού ματίσματος κατηγορίας 1 ιντεϊνών:4 ^ο βήμα	25
1.2.5 Άλλοι μηχανισμοί πρωτεϊνικού ματίσματος.....	25
1.2.6 Πρωτεϊνικό μάτισμα υπό προϋποθέσεις	27
1.3 Δομικά και δυναμικά χαρακτηριστικά της ιντεΐνης.....	29
1.3.1 Αναδίπλωση ιντεϊνών	30
1.3.2 Η δομική βάση του πρωτεϊνικού ματίσματος	31
1.4 Εφαρμογές των ιντεϊνών στη Βιοτεχνολογία	39
1.4.1 Καθαρισμός πρωτεϊνών	39
1.4.2 Τροποποίηση πρωτεϊνών	41
1.4.3 Ιντεϊνο-βασίζόμενοι βιοαισθητήρες.....	43
1.4.4 Έλεγχος γονιδίων σε Διαγονιδιακούς οργανισμούς	46
1.5 SICLOPPS.....	47
1.5.1 Γενετική επιλογή στο SICLOPPS.....	48
1.5.2 Στόχευση σε μεταγραφικούς παράγοντες	52
1.5.3 Πέρα από <i>E. coli</i>	53

1.6	Στόχος.....	54
2.	Πειραματικό μέρος.....	57
2.1	Μέθοδοι και Υλικά.....	58
2.1.1	Κυτταρικές Καλλιέργειες.....	58
2.1.2	Αναλυτικές μέθοδοι	61
2.1.3	Συσκευές και μηχανήματα	68
2.2	Μελέτη συνθηκών υπερέκφρασης του πρωτεϊνικού συμπλέγματος πεπτιδίου-ιντεΐνης	69
2.2.1	Συγκέντρωση επαγωγή (L-Αραβινόζη).....	69
2.2.2	Χρόνος επαγωγής.....	69
2.2.3	Θερμοκρασία	69
2.3	Μελέτη βιοσύνθεσης κυκλικού πεπτιδίου 21	70
2.3.1	<i>In vivo</i> παραγωγή και απομόνωση με στήλη μοριακής διήθησης	70
2.3.2	<i>In vitro</i> παραγωγή και απομόνωση με στήλη συγγένειας.....	70
3.	Αποτελέσματα	72
3.1	Μελέτη συνθηκών υπερέκφρασης του πρωτεϊνικού συμπλέγματος πεπτιδίου-ιντεΐνης	73
3.1.1	Συγκέντρωση επαγωγή (L-Αραβινόζη).....	74
3.1.2	Χρόνος επαγωγής.....	76
3.1.3	Θερμοκρασία	78
3.1.4	Πρωτεϊνικό Μάτισμα	81
3.2	Μελέτη βιοσύνθεσης κυκλικού πεπτιδίου 21	82
3.2.1	<i>In vivo</i> παραγωγή	83
3.2.2	<i>In vitro</i> παραγωγή	84
4.	Συζήτηση-Συμπεράσματα.....	90
4.1	Μελέτη συνθηκών υπερέκφρασης του πρωτεϊνικού συμπλέγματος πεπτιδίου-ιντεΐνης	92
4.2	Μελέτη βιοσύνθεσης κυκλικού πεπτιδίου 21	98
5.	Περίληψη	100
6.	Βιβλιογραφία	103

1. Εισαγωγή

Τα τελευταία χρόνια έχει παρατηρηθεί ένα αυξημένο ενδιαφέρον στην παραγωγή κυκλικών πεπτιδίων τόσο ως θεραπευτικές ή διαγνωστικές ουσίες όσο και ως ανιχνευτές για την μελέτη Βιολογικών συστημάτων (Driggers et al., 2008; Marsault and Peterson, 2011). Το ενδιαφέρον αυτό μπορεί να οφείλεται στη βιοϊατρική αξία όπως και στην πληθώρα βιολογικών δράσεων ήδη γνωστών κυκλικών πεπτιδίων που έχουν απομονωθεί από τη φύση. Η ανακάλυψη της γραμισιδίνης S το 1944 από τους Gause και Brazhnikova αποτελεί κομβικό σημείο στην ιστορία κυκλικών πεπτιδίων και το πρώτο παράδειγμα χρήσης για ιατρικούς λόγους (Gause and Brazhnikova, 1944). Άλλα σημαντικά παραδείγματα αποτελούν η κυκλοσπορίνη (ανοσοκατασταλτικό), η κασποφουγκίνη (αντιμυκητιακό) και η πολυμυκίνη (αντιβιοτικό) (Nolan and Walsh, 2009). Πιο συγκεκριμένα, ο περιορισμός της διαμόρφωσης των κυκλικών πεπτιδίων μπορεί να έχει ως αποτέλεσμα την ενισχυμένη συγγένεια σύνδεσης με πρωτεΐνες (Tang et al., 1999), αυξημένη εκλεκτικότητα (Henchey et al., 2010), μεγαλύτερη διαπερατότητα από τις μεμβράνες (Rezai et al., 2006; Walensky et al., 2004) αλλά και μικρότερο κίνδυνο πρωτεόλυσης από έξω- και ένδο-πρωτεάσες σε σχέση με τα αντίστοιχα γραμμικά πεπτίδια (Fairlie et al., 2000; Wang et al., 2005), χαρακτηριστικά που είναι επιθυμητά στην περίπτωση *in vivo* εφαρμογών και σχεδιασμού θεραπευτικών ουσιών.

Υπάρχουν αρκετές χημικές και βιολογικές μέθοδοι για την παρασκευή βιβλιοθηκών κυκλικών πεπτιδίων (Chen et al., 2013; Diderich and Heinis, 2013; Morimoto et al., 2012; White and Yudin, 2011). Ωστόσο, οι χημικά κατασκευασμένες βιβλιοθήκες αντιμετωπίζουν προκλήσεις στο επίπεδο του προσδιορισμού, οι οποίες έχουν εν μέρει ξεπεραστεί με τη χρήση μεθοδολογιών DNA-κωδικοποίησης (Kleiner et al., 2011). Το μέγεθος όμως της βιβλιοθήκης μπορεί να περιοριστεί σε τέτοιες προσεγγίσεις λόγω της μειωμένης λειτουργικότητας των συνθετικών σταδίων για την αντιμετώπιση της αυξημένης πολυπλοκότητας.

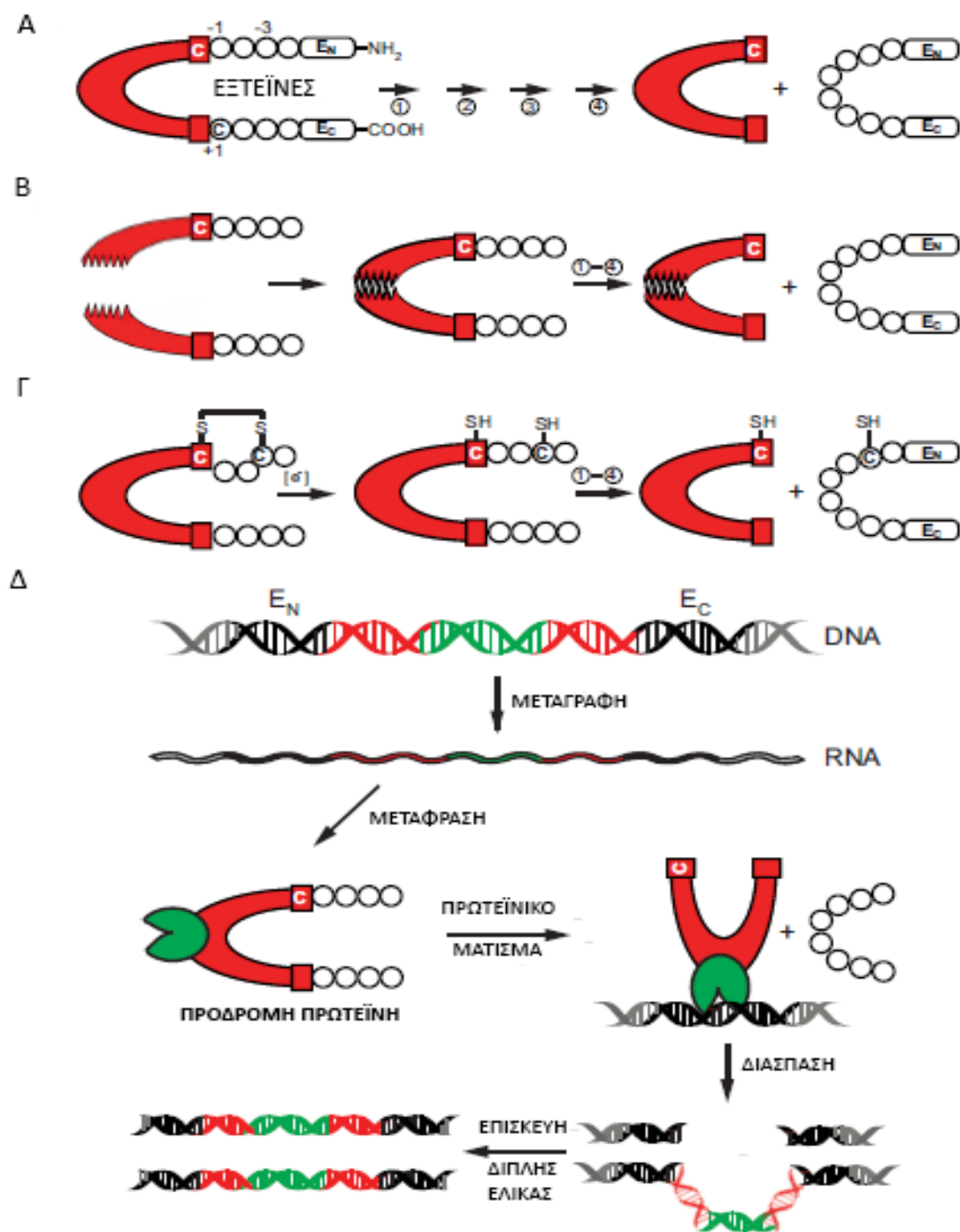
Ο σχεδιασμός γενετικά κωδικοποιημένων βιβλιοθηκών για τη κυκλική σύνθεση πεπτιδίων επιτρέπει την κατασκευή πολύ μεγάλων βιβλιοθηκών (10^6 με 10^{12} προϊόντα), ενώ υπερβαίνει τις προκλήσεις προσδιορισμού μέσω της αλληλουχίσης του DNA. Αρκετές προσεγγίσεις έχουν αναφερθεί που χρησιμοποιούν ριβοσωμική και μη-ριβοσωμική σύνθεση πεπτιδίου για την κατασκευή των γενετικά κωδικοποιημένων βιβλιοθηκών κυκλικών πεπτιδίων. Η χρήση μη-ριβοσωμικών συνθετάσεων πεπτιδίων είναι μια πολλά υποσχόμενη διεργασία που απαιτεί

περισσότερη ανάπτυξη (Bowers, 2012), ενώ αντιθέτως η χρήση ριβοσωμικής σύνθεσης πεπτιδίων είναι μια προσέγγιση που έχει χρησιμοποιηθεί ευρέως για την κατασκευή βιβλιοθηκών.

Είναι λογικό ότι υπάρχουν ποικίλες μέθοδοι για την κατασκευή γενετικά κωδικοποιημένων βιβλιοθηκών. Μια ενδιαφέρουσα προσέγγιση χρησιμοποιεί την ενσωμάτωση μη-φυσικών αμινοξέων μέσω της καταστολής του στοπ κωδικονίου και την GyrA ιντεΐνη για τη κυκλοποίηση υβρίδια όργανό-συντεθειμένων πεπτιδίων (Satyanarayana et al., 2012). Μία άλλη μέθοδος για την παραγωγή βιβλιοθηκών κυκλικών πεπτιδίων χρησιμοποιεί μια ανομοιογενή ριβοζύμη ακυλίωσης tRNA ικανή να φορτώσει μη-φυσικά αμινοξέα επάνω σε tRNAs, τα οποία χρησιμοποιούνται για την παραγωγή πολυπεπτιδίων σε ένα σύστημα *in vitro* μεταγραφής / μετάφρασης (Murakami et al., 2006). Επίσης έχει πρόσφατα αναφερθεί μέθοδος για την παραγωγή βιβλιοθηκών με τη χρήση της τεχνικής παρουσίασης μοριακών σχημάτων σε βακτηριακούς φάγους (phage display) (Chen et al., 2013; Diderich and Heinis, 2013). Πέρα από τα παραπάνω παραδείγματα το πιο συχνά χρησιμοποιημένο σύστημα είναι το SICLOPPS (split-intein circular ligation of peptides and proteins), το οποίο χρησιμοποιείται για τον προσδιορισμό κυκλικών πεπτιδίων που αλληλοεπιδρούν με μια συγκεκριμένη πρωτεΐνη στόχο. Το σύστημα αυτό, το οποίο αναφέρθηκε για πρώτη φορά το 1999 από το Benkovic χρησιμοποιεί μια συγκεκριμένη ιντεΐνη (intein, internal protein), τη Ssp DnaE διχασμένη ιντεΐνη για την αντίδραση κυκλοποίησης (Scott et al., 1999). Παρακάτω θα γίνει μια ανασκόπηση για το τι είναι ιντεΐνες και ποια είναι η βιολογική και βιοτεχνολογική τους σημασία.

1.1 Ιντεΐνες

Οι ιντεΐνες (intein, Internal Protein) είναι κινητά γενετικά στοιχεία που έχουν την δυνατότητα μετα-μεταφραστικού αυτοματίσματος (πρωτεϊνικό μάτισμα). Εμφανίζονται και στις τρεις επικράτειες της ζωής συμπεριλαμβανομένων των ιών και των βακτηριοφάγων, όπου παρατηρείται μια σποραδική κατανομή ακόμη και μεταξύ πολύ συγγενικών ειδών. Το πρωτεϊνικό μάτισμα (protein splicing) είναι μια φυσική βιοχημική διαδικασία που οδηγεί στην μετα-μεταφραστική τροποποίηση ενός προδρόμου πολυπεπτιδίου προς μία ώριμη και λειτουργική πρωτεΐνη, μέσω της αφαίρεσης ενός εσωτερικού στοιχείου της πρωτεΐνης που ονομάζεται ιντεΐνη (Εικ. 1.1Α). Η διαδικασία αυτή είναι ανάλογη με το μάτισμα του ιντρονίου σε επίπεδο RNA. Ο μηχανισμός του πρωτεϊνικού ματίσματος περιλαμβάνει μια σειρά από αυτόκαταλυούμενες αναδιατάξεις



Εικόνα 1.1.

Πρωτεϊνικό μάτισμα και οι διάφοροι τύποι ιντεϊνών. **Α)** σχηματική αναπαράσταση του μηχανισμού πρωτεϊνικού ματίσματος. Η ιντεΐνη που φαίνεται με κόκκινο είναι ενωμένη με τα δυο τμήματα εξτεϊνών (E_N και E_C) **Β)** Πρωτεϊνικό μάτισμα in trans από διχασμένη ιντεΐνη. **Γ)** Πρωτεϊνικό μάτισμα υπό προϋποθέσεις. Ένας δισουλφιδικός δεσμός μεταξύ της -3 Cys της εξτεϊνης και το πρώτο κατάλοιπο (1 Cys) της ιντεΐνης

παρεμποδίζει το πρωτεϊνικό μάτισμα. Παρουσία αναγωγικού παράγοντα ο δισουλφιδικός σπάει και το μάτισμα ενεργοποιείται επιτυχώς. Δ) Μηχανισμός διάδοσης των ιντεϊνών λόγω της ύπαρξης της ενδονουκλεάσης παλινόστησης (HEN) (Novikova et al., 2014).

πεπτιδικού δεσμού, όπου η ιντεΐνη κόβεται από το πρόδρομο πολυπεπτίδιο με ταυτόχρονη σύνδεση των πλευρικών αλληλουχιών, που ονομάζονται εξτεΐνες (N ή C-εξτεΐνη σε σχέση με τη θέση του δίπλα από την ιντεΐνη) (Mills et al., 2014). Ως αποτέλεσμα αυτής της διαδικασίας, παράγονται δύο πρωτεΐνες από ένα μόνο πολυπεπτιδικό προϊόν. Ο όρος ιντεΐνη αναφέρεται τόσο στο γενετικό στοιχείο του DNA ή RNA όσο και στην πρωτεϊνική οντότητα του ματίσματος.

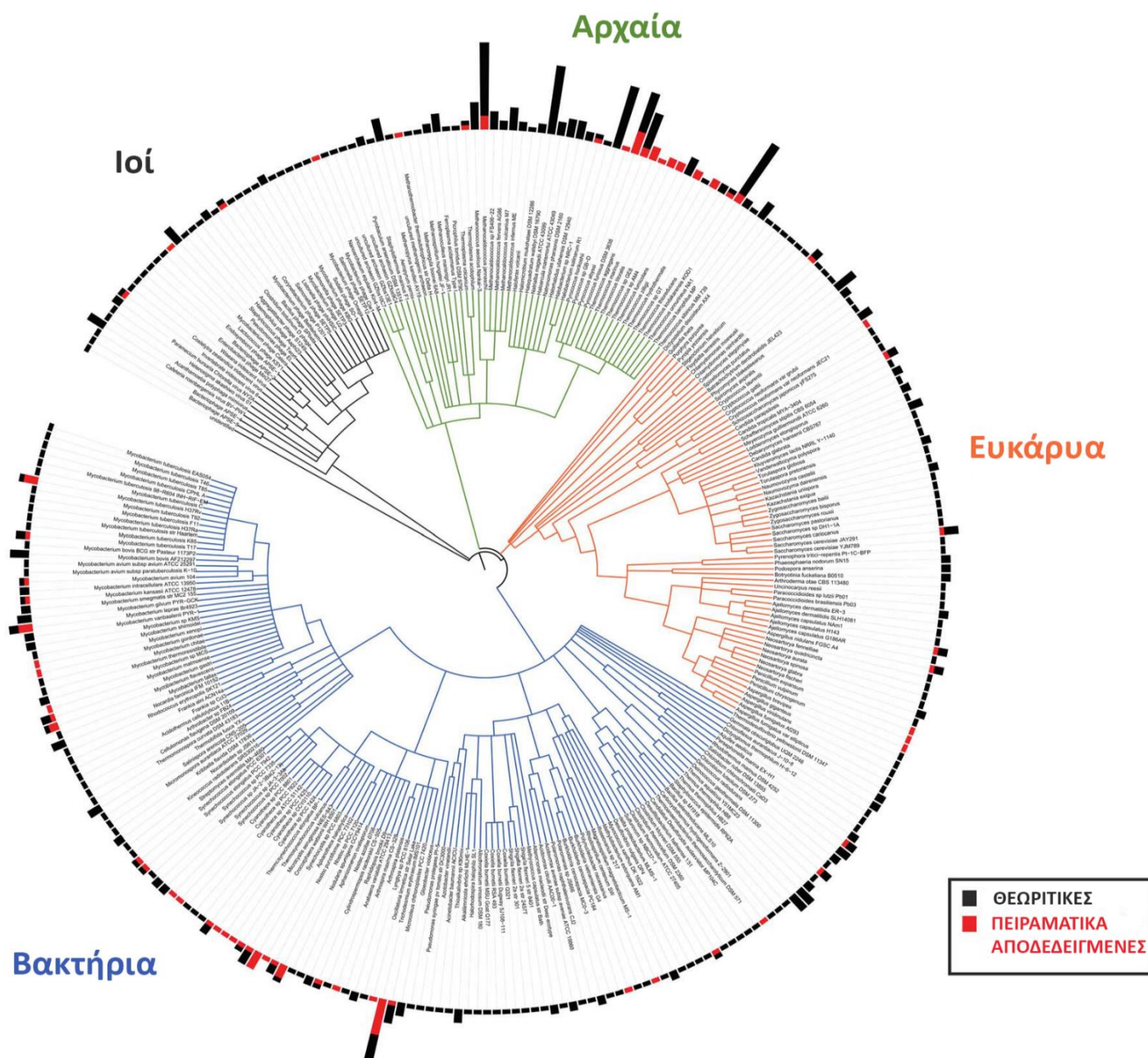
Οι περισσότερες ιντεΐνες εκφράζονται εντός μιας μόνο πολυπεπτιδικής αλυσίδας (cis-ματίσμα), αλλά έχουν παρατηρηθεί μερικές που χωρίζονται σε δύο πολυπεπτιδικά θραύσματα όπου το καθένα περιέχει μια εξτεΐνη και ένα κομμάτι ιντεΐνης (trans-μάτισμα) (Xu and Evans, 2005). Στην περίπτωση των διχασμένων ιντεϊνών, η επανασύνδεση των θραυσμάτων πραγματοποιείται με μια διαδικασία διασύνδεσης τύπου-φερμουάρ (Sorci et al., 2013), που προηγείται του πρωτεϊνικού ματίσματος (Εικ 1B). Τόσο οι ιντεΐνες που δρουν με το μηχανισμό του cis-ματίσματος όσο αυτές με του trans-ματίσματος χρησιμοποιούνται συχνά σε διάφορες βιοτεχνολογικές εφαρμογές, όπως για παράδειγμα καθαρισμός πρωτεϊνών, τροποποίηση, επισήμανση, και μετα-μεταφραστικό έλεγχο των εκφρασμένων πρωτεϊνών (Shah and Muir, 2014). Οι cis-ιντεΐνες (cis-μάτισμα) συχνά περιέχουν μια διακριτή περιοχή μιας ενδονουκλεάσης παλινόστησης (HEN, homing endonuclease) (Derbyshire et al., 1997).

Οι ιντεΐνες που περιέχουν την HEN ενδονουκλεάση χαρακτηρίζονται ως φυσικά κινητά γενετικά στοιχεία. Η παρουσία της HEN ενδονουκλεάσης παρέχει στις εκάστοτε ιντεΐνες την ικανότητα να μεταφέρουν το γονίδιο τους σε ομόλογες αλληλόμορφες περιοχές που δεν έχουν την αλληλουχία της ιντεΐνης (Εικ 1Δ). Αυτή η διαδικασία αυτοκατευθυνόμενης μεσολάβησης της HEN μπορεί να οδηγήσει σε οριζόντια μεταφορά γονιδίων (HGT) των ιντεϊνών, με την εισβολή διαφόρων ειδών, που ακολουθείται από κάθετη μετάδοση ιντεϊνών (Gogarten et al., 2002). Επιπλέον, οι ιντεΐνες που περιέχουν την HEN εμπλέκονται σε ένα λεγόμενο «κύκλο παλιννόστησης» (homing cycle), που περιλαμβάνει δύο αντίθετες διαδικασίες, την ακριβή απώλεια ιντεΐνης και την επανεισβολή μιας νεοσύστατης κενή θέσης παλιννόστησης. Πιστεύεται ότι ο «κύκλος παλιννόστησης» βοηθάει την HEN να αποφευχθεί τη στερέωση και τη λειτουργική αποσύνθεση σε ένα τόπο (Burt and Kouforanou, 2004).

1.1.1 Η τεράστια φυλογένεση των ιντεϊνών

Το 1990 βρέθηκε η πρώτη περιοχή πρωτεϊνικού ματίσματος ενσωματωμένη στο γονίδιο της πρωτόνιο-μεταθετικής ΑΤΡασης στο ζυμομύκητα *Saccharomyces cerevisiae* (Hirata et al., 1990). Από τότε, με την εκμετάλλευση πολλών εξαιρετικά συντηρημένων μοτίβων αλληλουχιών ιντεϊνης, έχουν εντοπιστεί με βιοπληροφορικές μεθόδους τουλάχιστον 600 πιθανά γονίδια ιντεϊνης στα γονιδιώματα μονοκύτταρων οργανισμών από όλες τις επικράτειες της ζωής (αρχαία, βακτήρια και ευκαρυώτες), καθώς επίσης και σε διάφορα ιικά γονιδιώματα (Εικ 1.2) (Perler, 2002). Ενδιαφέρον προκαλεί το γεγονός ότι αυτή η ευρεία κατανομή των ιντεϊνών δεν αφορά μόνο τη συστοιχία των οργανισμών στους οποίους βρίσκονται, αλλά επίσης και τη μεγάλη ποικιλία των γονιδίων ξενιστή, στον οποία είναι ενσωματωμένες. Μέχρι σήμερα, υπάρχουν τουλάχιστον 70 διαφορετικά αλληλόμορφα ιντεϊνων, που διακρίνονται όχι μόνο από τον τύπο του γονιδίου υποδοχής στο οποίο η ιντεϊνή είναι ενσωματωμένη, αλλά και από το σημείο ολοκλήρωσης εντός του γονιδίου υποδοχής (Pietrokovski, 2001). Επιπλέον, σε ορισμένες πρωτεΐνες έχουν βρεθεί να περιέχονται έως και τέσσερις ιντεΐνες ενσωματωμένες σε διάφορα σημεία ολοκλήρωσης μέσα στο γονίδιο τους, ενώ αρκετοί οργανισμοί έχουν παρατηρηθεί με περισσότερες από μία ιντεΐνη (έως και 19 ιντεΐνες) στο γονιδίωμα τους (Εικ 1.2) (Perler, 2002).

Οι ιντεΐνες είναι συνήθως ενσωματωμένες εντός ουσιωδών πρωτεϊνών που εμπλέκονται στην αντιγραφή, τη μεταγραφή και τη συντήρηση του DNA (π.χ. υπομονάδες DNA ή RNA πολυμερασών, DNA ελικάσες, DNA γυράσες, και ριβονουκλεοτιδική αναγωγή) ή σε άλλα house-keeping γονίδια βασικών πρωτεασών και μεταβολικών ενζύμων (Pietrokovski, 2001). Σε αυτές τις πρωτεΐνες, τα πεδία επεξεργασίας της περιοχής υποδοχής των ιντεϊνών παρεμβάλλουν συνήθως συντηρημένες περιοχές, που είναι ζωτικής σημασίας για τη λειτουργία της πρωτεΐνης ξενιστή (π.χ. θέσεις πρόσδεσης υποστρώματος ή ενεργό κέντρο ενζύμου) (Dalgaard et al., 1997). Με δεδομένο αυτό, πιστεύεται σε μεγάλο βαθμό ότι το πρωτεϊνικό μάτισμα της ιντεΐνης απαιτείται για την υποχρεωτική λειτουργία των πρωτεϊνών, ωστόσο δεν έχει αποδειχθεί αν οι ιντεΐνες έχουν ένα σαφές κανονιστικό ρόλο στην πρωτεΐνη ξενιστή ή παρέχουν ένα πλεονέκτημα υγείας στον οργανισμό (Frischkorn et al., 1998). Είναι ενδιαφέρον όμως ότι έχει αποδειχθεί, πως υπάρχει μια ελεγχόμενη οξειδοαναγωγική διαδικασία ματίσματος για ορισμένες ιντεΐνες. Έχει προταθεί ότι οι ιντεΐνες μπορούν να αναστείλουν τη συναρμολόγηση και τη λειτουργία της ενδογενούς πρωτεΐνης υποδοχής τους υπό «υπεροξικές συνθήκες», αν και αυτό δεν έχει πιστοποιηθεί στο φυσικό οργανισμό υποδοχής (Callahan et al., 2011). Η έλλειψη άμεσων αποδείξεων για τη βιολογική



Εικόνα 1.2.

Φυλογενετική κατανομή των οργανισμών που έχουν ιντεΐνες (περίπου 300). Η μικρότερη μπάρα υποδηλώνει μια ιντεΐνη ενώ η μεγαλύτερη δεκαεννέα. Το κόκκινο χρώμα στις μπάρες υποδεικνύει ιντεΐνες που έχουν προσδιοριστεί πειραματικά η ικανότητα τους για πρωτεϊνικό μάτισμα και το μαύρο χρώμα αυτές που έχουν προσδιοριστεί με βάση την αλληλουχία του γονιδίου χωρίς να έχει βρεθεί δράση ματίσματος. Τα δεδομένα

προέρχονται από τη βάση δεδομένων InBase της NEB που έχει ενημερωθεί τελευταία το 2010 (Shah and Muir, 2014).

σημασία της λειτουργίας των ιντεϊνών εγείρει το ερώτημα «Γιατί υπάρχουν οι ιντεΐνες;». Μέχρι στιγμής, η επικρατούσα άποψη είναι ότι οι ιντεΐνες αποτελούνται από εγωιστικά γονίδια χωρίς προφανή βιολογικό ρόλο. Ανεξάρτητα από αυτό, η παρουσία και η επιμονή τους στα μικροβιακά γονιδιώματα είναι αναμφίβολα συναρπαστική.

1.1.2 Η σποραδική εξάπλωση των ιντεϊνών σε συνδυασμό με την εξελικτική ιστορία των ξενιστών τους

Η ευρεία φυλογενετική κατανομή των ιντεϊνών υποδηλώνει ότι τα μόρια αυτά έχουν αρχαία προέλευση. Παρ' όλα αυτά, για τους λόγους που περιγράφονται κάτω, είναι σαφές ότι η επικράτηση τους δεν πρέπει να οφείλεται μόνο σε κάθετη, αλλά επίσης οριζόντια μεταφορά του γονιδίου: (1) Οι ιντεΐνες ενσωματώνονται σε μια ευρεία ποικιλία των πρωτεϊνών του ξενιστή. (2) Η ύπαρξη μιας ιντεΐνης σε ένα γονίδιο ενός μικροβιακού ξενιστή, δε σημαίνει ότι θα βρεθεί σε ένα ομόλογο γονίδιο ενός στενά σχετιζόμενου οργανισμού. (3) Οι ιντεΐνες απουσιάζουν εντελώς στους πολυκύτταρους οργανισμούς (αν και μερικοί πολυκύτταροι οργανισμοί παρουσιάζουν τις ομόλογες περιοχές με αντίστοιχες βιοχημικές λειτουργίες). (4) Αλληλόμορφα γονίδια ιντεϊνών έχουν συνήθως υψηλότερη ομολογία και διαφορετική χρήση κωδικονίων από τα γονίδια που τους φιλοξενούν. Αυτά τα στοιχεία υποδηλώνουν ότι υπάρχουν μηχανισμοί για τη διάδοση των ιντεϊνών, από ένα γονίδιο υποδοχής ή οργανισμό ξενιστή σε άλλο, όπως αντίστοιχα και μηχανισμοί για την απομάκρυνσή τους.

Όπως αναφέρθηκε πιο πάνω πολλές ιντεΐνες μέσα στην λειτουργική περιοχή τους περιλαμβάνουν, μια άλλη λειτουργική μονάδα, μια ενδονουκλεάση παλινόστησης (HEN, homing endonuclease). Οι HEN σπάζουν δίκλωνο DNA σε περιοχές που υπάρχουν αλληλουχίες αναγνώρισης οι οποίες κωδικοποιούνται εντός των γονιδίων του ξενιστή και είναι ελεύθερα αλληλόμορφα των ιντεϊνών-HEN. Αυτά τα σπασίματα αποτελούν έναρξη μιας διαδικασίας ανασυνδυασμού που οδηγεί στην δημιουργία ενός γονιδίου ιντεΐνης- HEN (Εικ. 1.2 και 1.3).

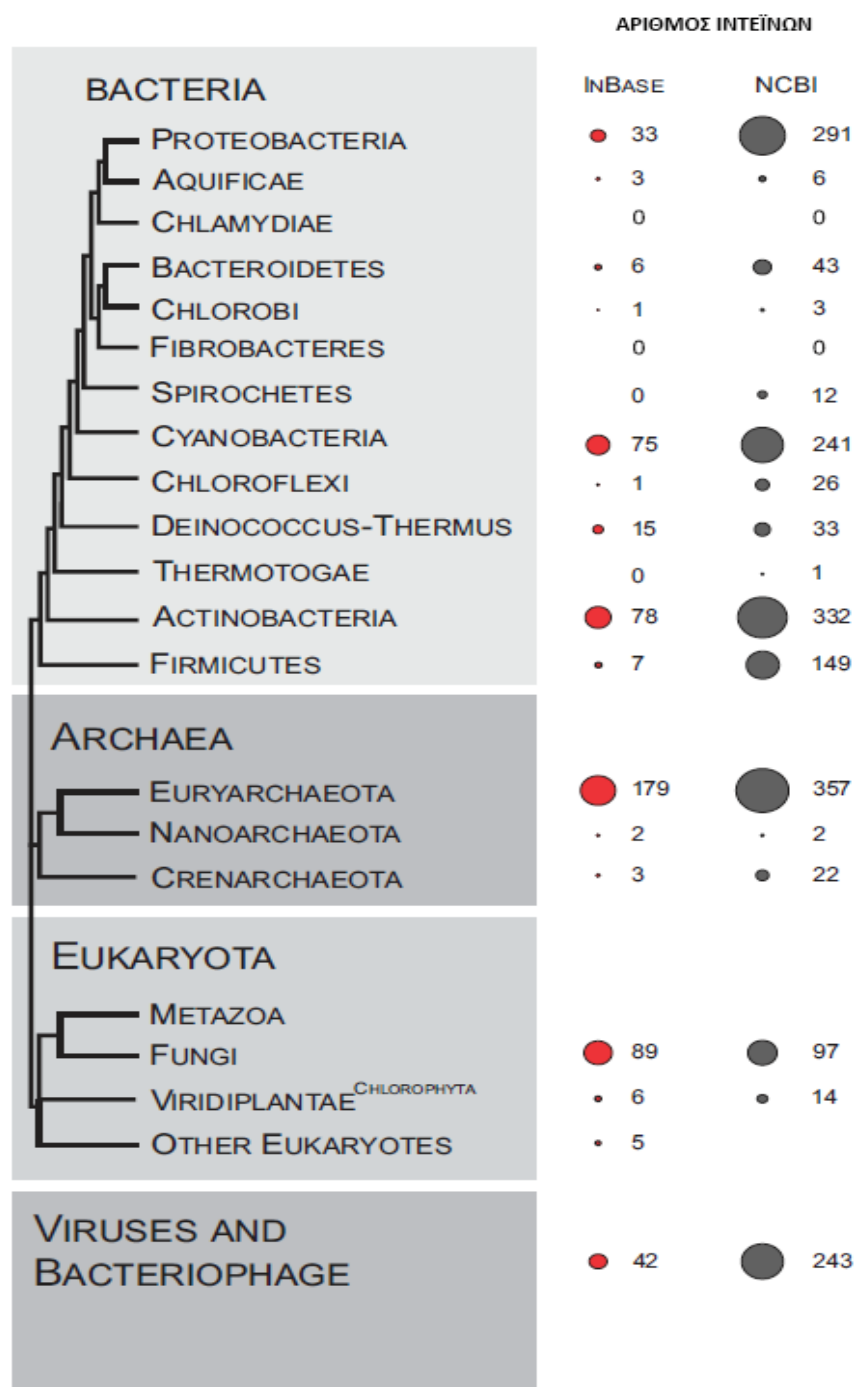
Ο μηχανισμός αυτός μπορεί να εξηγήσει τόσο την εντός όπως και την μεταξύ των ειδών εξάπλωση των ιντεϊνών, που αφορούν το ίδιο γονίδιο υποδοχής και την περιοχή προσθήκης (Pietrokovski, 2001). Ωστόσο, δεδομένης της ιδιαιτερότητας της αλληλουχίας των HEN, είναι απίθανο ότι ο μηχανισμός αυτός επιτρέπει την εξάπλωση ιντεϊνών σε διαφορετικές περιοχές προσθήκης εντός

του ίδιου γονιδίου ή διαφορετικών. Δεν είναι σαφές όμως πώς παρουσιάζεται αυτή η διαδικασία, ωστόσο, είναι αξιοσημείωτο ότι οι μη αλληλόμορφες ιντεΐνες έχουν εξαιρετικά χαμηλή ομολογία αλληλουχίας όταν συγκρίνονται με αλληλόμορφες (Pietrokonvski, 2001). Αυτό υποδηλώνει ότι η εξάπλωση των ιντεϊνών σε μια ποικιλία γονιδίων ξενιστών ήταν μια αρχαία διαδικασία και ότι από την εμφάνιση του αρχικού φαινομένου οι μη αλληλόμορφες έχουν αποκλίνει. Η πρόσφατη ανακάλυψη των ιντεϊνών στο ιικό γονιδίωμα μπορεί να δώσει εξηγήσεις για την γρήγορη διάδοση των ιντεϊνών (Raoult et al., 2004).

Η απώλεια μιας ιντεϊνης από ένα γονίδιο υποδοχής μπορεί να πραγματοποιηθεί με αρνητική επιλογή, άμα είναι επιζήμια στην υγεία του ξενιστή. Ωστόσο, η απώλεια μέσω διαγραφής του γονιδίου θα πρέπει να είναι δύσκολη, καθώς απαιτεί ακριβή αφαίρεση του γονιδίου της ιντεϊνης μέσα από ένα απαραίτητο γονίδιο. Προφανώς αυτή η διαδικασία είναι πιο πιθανό να εμφανίζεται σε διπλοειδείς ή πολυπλοειδείς οργανισμούς, όπου ένα αντίγραφο ενός γονιδίου που περιέχει την ιντεϊνη είναι αναλώσιμο σε σχέση με ένα γονίδιο χωρίς. Επιπλέον, η ικανότητα για ενδοχρωμοσωμικό ανασυνδυασμό παρέχει μια άλλη διαδρομή για την αποβολή ιντεϊνών, κάτι που μπορεί να εξηγήσει την έλλειψη ιντεϊνών σε πολυκύτταρους οργανισμούς.

Αξίζει να σημειωθεί ότι η εξάπλωση των ιντεϊνών σε διαφορετικά φύλα αλλά και μεταξύ στενά συγγενικών ειδών είναι σποραδική (Perler, 2002; Perler et al., 1997). Για να εξετάσει κανείς την εξάπλωση των ιντεϊνών, είναι σημαντικό να ερευνήσει τη βάση δεδομένων των ιντεϊνών inbase (Perler, 2002), για ιντεΐνες σε ολόκληρο το δέντρο της ζωής, συμπεριλαμβανομένων των ιών και των βακτηριοφάγων. Για εξονυχιστική μελέτη της εμφάνισης των ιντεϊνών προτιμάται η χρήση του National Center for Biotechnology Information (NCBI) και ειδικότερα η βάση δεδομένων των γονιδίων (www.ncbi.nlm.nih.gov/gene).

Υπάρχουν ελλείψεις στο περιεχόμενο της βάσης δεδομένων Inbase σε σύγκριση με τη βάση δεδομένων NCBI Gene, κάτι που οφείλεται στο γεγονός ότι η Inbase δημιουργήθηκε από παρατηρήσεις μεμονωμένων ερευνητών. Υπάρχουν φύλα που είτε είναι απόντα είτε δεν εκπροσωπούνται σε μεγάλο βαθμό στη inbase (π.χ. Spirochaet και Firmicutes), ενώ υπάρχουν περισσότερες υποθετικές ιντεΐνες στη βάση δεδομένων NCBI Gene σε σχέση με την Inbase (Εικ 1.3). Ωστόσο, υπάρχουν πολλές ομοιότητες μεταξύ των δύο βάσεων δεδομένων, όπως για φύση της εμφάνισης του ιντεϊνών υποδηλώνει μεροληπτική απόκτηση, συντήρηση ή και παράδειγμα ότι εμφανίζονται και στις δυο σποραδικά μοτίβα διάδοσης των ιντεϊνών ή άνιση απώλεια ιντεϊνών.

**Εικόνα 1.3.**

Διαφορές των αριθμών κατανομής των ιντεϊνών στις δυο βάσεις δεδομένων InBase και NCBI. (Novikova et al., 2014; Perler, 2002)

Το γεγονός ότι υπάρχει μια δραματική ποικιλότητα της εμφάνισης των ιντεϊνών, ακόμη και μεταξύ συγγενικών ειδών, που οφείλεται στη συνεχή διάδοση τους στα διάφορα γονιδιώματα υποδοχής, εξακολουθεί να αποτελεί πρόκληση στην εξελικτική βιολογία. Υπάρχουν πολλές δυνατότητες αλληλένδετων παραγόντων που συμβάλλουν στην παρατηρούμενη διάδοση των ιντεϊνών, που θα έπρεπε να εξεταστούν. Αναμφίβολα, η ποικιλότητα στην εξελικτική ιστορία των γονιδιωμάτων υποδοχής συνέβαλε στην ετερογένεια των ιντεϊνών σε όλες τις κλίμακες. Συνήθως, υποτίθεται ότι ιντεΐνες δεν είναι υπό ισχυρή επιλεκτική πίεση, είτε αρνητική είτε θετική, διότι προκαλούν αποτελεσματικό μάτισμα, παράγοντας πλήρως λειτουργικές πρωτεΐνες. Η μόνο εξελικτική πίεση που συνήθως αποδίδεται ότι δρα στις ιντεΐνες είναι η επιλογή για το αποτελεσματικό μάτισμα πρωτεϊνών (Gogarten et al., 2002). Ωστόσο, ένας μεγάλος αριθμός διεργασιών αναμόρφωσης του γονιδιώματος θα μπορούσε να επηρεάσει άμεσα τη διάδοση των ιντεϊνών.

Η απλούστευση του γονιδιώματος, δηλαδή η επιλεκτική εξάλειψη των περιττών κομματιών DNA για να μειωθεί η μεταβολική επιβάρυνση για την αντιγραφή του DNA, προτάθηκε ως μηχανισμός για τη μείωση του γονιδιώματος σε είδη που ζουν ελεύθερα με πολύ μεγάλο αριθμό πληθυσμού (Hessen et al., 2010) (Hessen et al., 2010). Σε αντίθεση, οι παρασιτικοί και συμβιωτικοί οργανισμοί υπόκεινται μια σημαντική μείωση του μεγέθους του γονιδιώματος λόγω χαλάρωσης από τη θετική επιλογή, μια τάση που ευνοεί τις διαγραφές σε σχέση με την αντικατάσταση και την ψευδο-ομογενοποίηση (Moran, 2002; Toft and Andersson, 2010). Τόσο ο εξορθολογισμός του γονιδιώματος όσο και η προσαρμογή σε ένα συμβιωτικό ή παρασιτικό τρόπο ζωής θα μπορούσε να οδηγήσει στην αποβολή ιντεϊνών. Η οριζόντια μεταφορά (που δεν συνδέεται με τη δραστηριότητα HEN και τη κινητικότητα των ιντεϊνών) από την άλλη θα μπορούσε να βοηθήσει στη διάδοση των ιντεϊνών μεταξύ μακρινών συγγενικών ειδών ή ακόμη και μεταξύ όλων των επικρατειών της ζωής. Η οριζόντια μεταφορά είναι κοινή μεταξύ Αρχαίων και Βακτηρίων, ενώ πιστεύεται ότι είναι ο ουσιαστικός μηχανισμός για την επίτευξη της γενετικής ποικιλότητας (Toft and Andersson, 2010). Ανάμεσα στα πιο γνωστά παραδείγματα μαζικών και επαναλαμβανόμενων οριζοντίων μεταφορών είναι η διάδοση των γονιδίων ανθεκτικότητας στα αντιβιοτικά μεταξύ των παθογόνων βακτηρίων.

Πίνακας 1.1

Κάποιες από τις πιο γνωστές πρωτεΐνες που περιέχουν ιντεΐνες (Gogarten et al., 2002; Novikova et al., 2014; Shah and Muir, 2014)

<i>Πρωτεΐνη (P-λούπα)</i>	<i>Λειτουργία</i>	<i>Κατανομή Ιντεϊνών</i>
<i>Ελικάση (DnaB)</i>	Αντιγραφική DNA ελικάση/ Αντιγραφή DNA	Βακτήρια
<i>(MCM)/ (Cdc21)</i>	Αντιγραφική DNA ελικάση/ Αντιγραφή DNA	Αρχαία
<i>Ρεκομπινάση (RecA/RadA)</i>	Αντιγραφή DNA	Βακτήρια
<i>SF2-Ελικάση (SWI/SNF2/Rad54)</i>	DNA ελικάση/ Επιδιόρθωση DNA	Βακτήρια
<i>Καταλυτική υπομονάδα DNA Πολυμεράσης III (DnaE)</i>	DNA πολυμεράση/Αντιγραφή DNA	Βακτήρια
<i>Πολυμεράση I (PolB)</i>	DNA πολυμεράση/Αντιγραφή DNA	Αρχαία
<i>Πολυμεράση II (PolC)</i>	DNA πολυμεράση/Αντιγραφή DNA	Βακτήρια
<i>Τοποϊσομεράση II-Υπομονάδα B DNA Γυράσης</i>	Τοποϊσομεράση/ Αντιγραφή DNA	Βακτήρια, Αρχαία, Βακτηριοφάγοι
<i>Βακτηριακή Πολυμεράση I</i>	DNA πολυμεράση/Αντιγραφή DNA	Βακτήρια
<i>Βακτηριακή Πολυμεράση II</i>	DNA πολυμεράση/Αντιγραφή DNA	Βακτήρια
<i>Καταλυτική υπομονάδα A ATP συνθάσης (VMA)</i>	DNA πολυμεράση/Αντιγραφή DNA	Ευκάρια, Αρχαία

Η οριζόντια μεταφορά είναι επίσης ένας σημαντικός μηχανισμός στην εξέλιξη του ευκαρυωτικού γονιδιώματος, ιδιαίτερα στους μονοκύτταρους οργανισμούς (Huang, 2013). Όλο και περισσότεροι υποψήφιοι για οριζόντια μεταφορά γονιδίων εντοπίζονται συνεχώς: από Βακτήρια σε Ευκαρυωτικά από Βακτήρια σε Αρχαία και μεταξύ των Ευκαρυωτικών. Οι ιντεΐνες μπορεί να μεταφέρονται μέσω της οριζόντιας μεταφοράς ως θραύσματα από γονίδια, συστάδες γονιδίων, θραύσματα γονιδιώματος, ή ακόμα και ολόκληρα χρωμοσώματα. Ιοί και βακτηριοφάγοι αποτελούν ιδανικούς φορείς οριζόντιας μεταφοράς για τη διάδοση ιντεϊνών (Filée et al., 2007). Μπορούν να μεταφέρουν αλληλουχίες ιντεϊνών πέρα από τα όρια των κυττάρων ως μέρος του δικού τους γονιδιώματος ή να μεταφέρουν κυτταρικά γονίδια που απέκτησαν ως αποτέλεσμα του ανεξέλεγκτου ανασυνδυασμού (Filée et al., 2007). Υπάρχουν πολλά παραδείγματα γονιδιωμάτων και βακτηρίων που φέρουν ιντεΐνες. Μεταξύ αυτών είναι ένας γιγαντιαίος ιός ενός σημαντικού θαλάσσιου μαστιγοφόρου *Cafeteria roenbergensis*, ο Mimivirus (Ogata et al., 2005) ένας ιός που

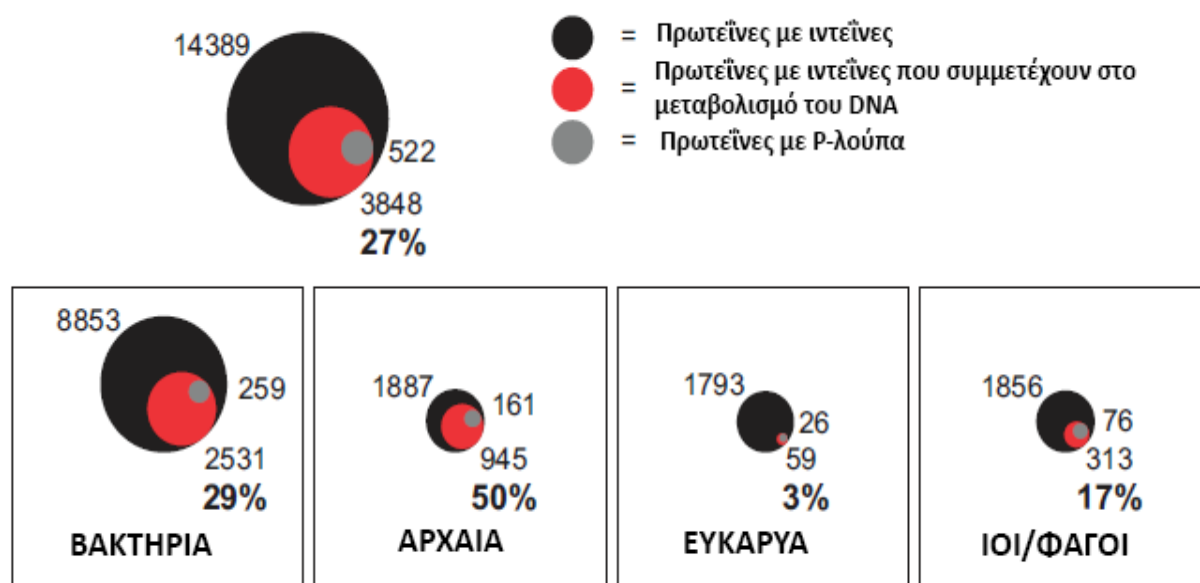
μολύνει τα φύκη *Heterosigma akashiwo* (Nagasaki et al., 2005) και *Chrysochromulina ericina* (Larsen et al., 2008). Εμφανίζονται επίσης ιντεΐνες σε ιούς Iridoviridae που μολύνουν ασπόνδυλα (Bigot et al., 2013), στον ιό Chlorella NY-2A (Fitzgerald et al., 2007), και στο Coccolithoviruse που μολύνει το *Emiliania huxleyi* (Allen et al., 2011). Βακτηριοφάγοι που φέρουν ιντεΐνες είναι επίσης πολύ διαδεδομένοι και έχουν αναφερθεί μεταξύ άλλων για τα *Bacillus* (Lazarevic et al., 1998), *Mycobacteria* (Hatfull, 2010) και *Caulobacter* (Gill et al., 2012).

1.1.3 Η διάδοση των ιντεϊνών είναι κατευθυνόμενη προς εξειδικευμένες πρωτεΐνες

Αν και οι ιντεΐνες εμφανίζονται σε πρωτεΐνες με διαφορετικές λειτουργίες, υπάρχει μια κατευθυνόμενη στόχευση των ιντεϊνών να εισάγονται σε πρωτεΐνες που εμπλέκονται στον μεταβολισμό του DNA, όπως πολυμεράσες (Pols), ελικάσες, τοποϊσομεράσες (TOPOs), και ριβονουκλεοτιδικές αναγωγάσες (RNRs) (ΠΙΝ.1.1, Εικ. 1.4) (Perler, 2002). Από τις 545 inteins που έχουν αναφερθεί στην Inbase, 266 εμφανίζονται είτε σε πολυμεράσες, όπως τις PolA, PolB, PolC και DnaE, , ή σε ελικάσες όπως και τις αντιγραφικές ελικάσες, την DnaB (βακτηριακή) και την MCM / Cdc21 (αρχαίων) (ΠΙΝ. 1.1). Μία έρευνα που έγινε βάση της βάσης δεδομένων του NCBI για τις πρωτεΐνες, αποκάλυψε ότι το 27% των πρωτεϊνών με ιντεΐνες σε όλες τις ομάδες αντιστοιχούσαν σε πρωτεΐνες που έχουν να κάνουν με το μεταβολισμό του DNA με το υψηλότερο ποσοστό στα Αρχαία (50% του συνόλου των υποθετικών πρωτεϊνών περιέχουν ιντεΐνες) και το χαμηλότερο στους Ευκαρυώτες (μόνο 3% Εικ. 1.4) (Novikova et al., 2014). Παρόλου που υπάρχουν διαφορές μεταξύ των δυο βάσεων δεδομένων (Inbase και NCBI), που αποδίδεται σε διαφορετικά κριτήρια υποβολής των δεδομένων, η υπεροχή των ιντεϊνών σε πρωτεΐνες του μεταβολισμού του DNA είναι εντυπωσιακή.

Αρκετές υποθέσεις έχουν προταθεί για την εισαγωγή των κατευθυνόμενων ιντεϊνών εντός αυτών των πρωτεϊνών. Πρώτον, επειδή η HEN -ιντεΐνη παράγεται ταυτόχρονα με την πρωτεΐνη υποδοχής της, ένα πιθανό πλεονέκτημα είναι να εξασφαλίσει μια ιντεΐνη την παρουσία της κατά την αντιγραφή και την επιδιόρθωση του DNA. Επειδή η παλινόστηση των ιντεϊνών απαιτεί τους μηχανισμούς αντιγραφής και επιδιόρθωσης του ξενιστή, είναι πιο αποτελεσματικό για τις ιντεΐνες να παράγονται μαζί με τις αντιγραφικές και επιδιορθωτικές πρωτεΐνες. Μια δεύτερη υπόθεση προτείνει ότι η κινητικότητα κατά τη διάρκεια της αντιγραφής και επιδιόρθωσης μειώνει την επιλογή προς τις ιντεΐνες. Πράγματι, ως αποτέλεσμα της μη ειδικής λειτουργίας της HEN -ιντεΐνης

η επισκευή μπορεί να γίνει αποτελεσματικά κατά τη διάρκεια αυτής της περιόδου, μειώνοντας τον κίνδυνο πολλαπλασιασμού ιντεϊνών. Τέλος, οι Pols, οι ελικάσες και οι TOPOs απαντώνται συχνά σε ιικά γονιδιώματα (περίπου 18% του συνολικού ιικού πρωτεώματος) (Filée et al., 2007). Οι ιοί είναι αποτελεσματικοί φορείς για την οριζόντια μεταφορά γονιδίων και ως εκ τούτου η εξάπλωση των ιντεϊνών πέρα από τα όρια των ειδών διευκολύνεται (Filée et al., 2007).



Εικόνα 1.4.

Διάδοση των ιντεϊνών στο πρωτέωμα . Παρουσιάζεται η συγκέντρωση των ιντεϊνών στις πρωτεΐνες που συμμετέχουν στο μεταβολισμό του DNA (Novikova et al., 2014).

1.1.4 Οι ιντεΐνες βρίσκονται συνήθως στην λειτουργική περιοχή των πρωτεϊνών

Οι ιντεΐνες εμφανίζονται συχνά σε κρίσιμες θέσεις για την λειτουργία των πρωτεϊνών, όπως στα καταλυτικά κέντρα και στις θέσεις πρόσδεσης υποκαταστατών. Ενδιαφέρον προκαλεί το γεγονός ότι ένα συντηρημένο Walker A μοτίβο που περιέχει (Walker et al., 1982) μια λούπα σύνδεσης φωσφορικών (P-loop) αποτελεί σημείο εισβολής από ιντεΐνες σε ορισμένες από αυτές τις πρωτεΐνες. Για παράδειγμα, από τα 33 είδη / στελέχη ιντεϊνών που βρεθήκαν για τις ρεκομπινάσες

RadA / RecA, 12 είχαν ενθέσεις στη P-loop, που έχουν ταυτοποιηθεί αντίστοιχα σε πολλές ATP- και GTP-δεσμευτικές πρωτεΐνες (Saraste et al., 1990) (Εικ. 1.3B). Επιπροσθέτως, οι ενθέσεις στη P-loop στις DnaB και MCM / Cdc21 ελικάσες βρέθηκαν σε 47 είδη και 51 στελέχη (Εικ 1.4). Αν και άλλες ελικάσες είχαν επίσης ενθέσεις ιντεϊνών στη P-loop, αυτό δεν παρατηρήθηκε στην περίπτωση των Pols (Πίνακας) (Perler, 2002).

Η λογική πίσω από τον εντοπισμό των ιντεϊνών στις πιο κρίσιμες περιοχές των πρωτεϊνών είναι ακόμα ένα θέμα συζήτησης, και θα μπορούσε να αντανakλά τη διαφορική στόχευση, τη συντήρηση, ή την απώλεια. Πρώτον, η στόχευση των συντηρημένων θέσεων θα μπορούσε να εξηγηθεί από την εξειδίκευση των HEN-ιντεϊνών. Συντηρημένα κατάλοιπα αμινοξέων στα σημαντικά λειτουργικά μοτίβα μιας πρωτεΐνης περιορίζουν το εύρος των νουκλεοτιδικών υποκαταστάσεων που μπορούν να είναι ανεκτές σε τέτοιες θέσεις. Έτσι, η πιθανότητα να χάσει τη θέση που αναγνωρίζεται από την HEN είναι χαμηλότερη σε ενεργά κέντρα σε σύγκριση με άλλες περιοχές της πρωτεΐνης. Δεύτερον, η διαδικασία καθαρισμού, η οποία διαδραματίζει σημαντικό ρόλο στη διατήρηση της μακροχρόνιας σταθερότητας των βιολογικών συστημάτων αφαιρώντας τις επικίνδυνες μεταλλάξεις, διατηρεί τα σημαντικά λειτουργικά μοτίβα, διευκολύνοντας την εισβολή των ιντεϊνών στον εν λόγω τόπο με οριζόντια μεταφορά σε ένα ευρύ φάσμα ειδών. Τρίτον, οι ιντεΐνες μπορούν να εισαχθούν σε διαφορετικές τοποθεσίες σε όλο το γονιδίωμα, αλλά μόνο σε συγκεκριμένες περιοχές οι παρεμβολές μπορούν να γίνουν σταθερές στον πληθυσμό. Η διατήρηση των ιντεϊνών σε συντηρημένα μοτίβα πρωτεΐνης μπορεί να οφείλεται στο χαμηλό ποσοστό της απώλειας μέσω της εκτομής. Η απομάκρυνση της ιντεΐνης θα διαταράξει τη λειτουργία της πρωτεΐνης και ως αποτέλεσμα θα γίνει μια συνολική διαγραφή (Gogarten et al., 2002). Εάν η ιντεΐνη αποκοπεί από το γονιδίωμα, είναι πιθανό ότι αυτή η νέα βιώσιμη παραλλαγή χωρίς την ιντεΐνη θα επαναπροσβληθεί από την HEN που περιέχει την ιντεΐνη, ως αποτέλεσμα του «κύκλου παλιννόστησης» (Gogarten et al., 2002). Ένα παρόμοιο επιχείρημα έχει προταθεί για τη διατήρηση των ιντρονίων αυτοματίσματος σε σημαντικά λειτουργικά μοτίβα (Goddard and Burt, 1999).

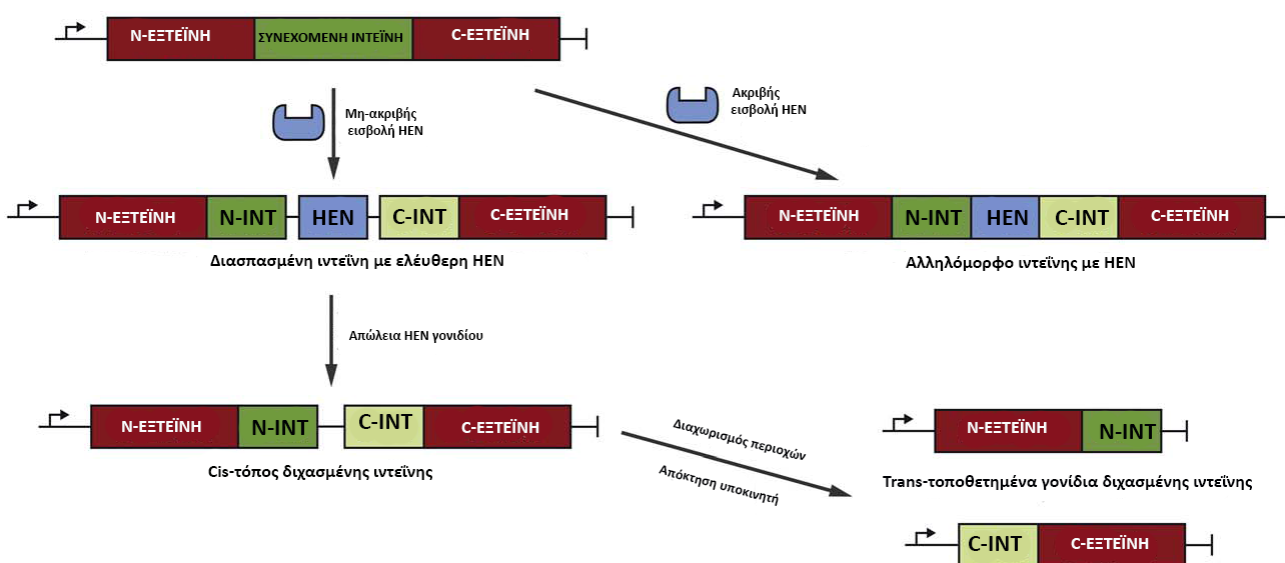
Τέλος, η παρουσία των ιντεϊνών σε ειδικότερα συντηρημένα μοτίβα θα μπορούσε να εξηγηθεί από ένα προσαρμοστικό ρόλο. Τα κινητά στοιχεία σε γενικές γραμμές φαίνεται να εξελίσσουν διαφορετικούς ρόλους (Faulkner and Carninci, 2009). Η τεράστια γνώση που έχει συσσωρευτεί για την οργάνωση του γονιδιώματος και της γονιδιακής έκφρασης κατά τη διάρκεια των δύο τελευταίων δεκαετιών έχει οδηγήσει σε μια φιλοσοφία ότι τα κινητά στοιχεία πρέπει θεωρούνται

μόνο ως παρασιτικές οντότητες για την καλύτερη κατανόηση του δυναμικού ρόλου τους στην εξέλιξη των ειδών (Faulkner and Carninci, 2009). Παρά τη σημασία τους στον τομέα της βιοτεχνολογίας, οι ιντεΐνες παραμένουν ένα από τα λιγότερο μελετημένα κινητά στοιχεία όσον αφορά την πιθανή λειτουργία τους. Παρ'όλα αυτά, τα διαθέσιμα δεδομένα, ειδικά όσον αναφορά στο πρωτεϊνικό μάτισμα, υποδηλώνουν ότι οι ιντεΐνες θα μπορούσαν να εμπλέκονται στην ρύθμιση της λειτουργίας της πρωτεΐνης του ξενιστή.

1.1.5 Η εμφάνιση των διχασμένων ιντεϊνών

Ένα μικρό κλάσμα (λιγότερο από 5%) των γονιδίων που κωδικοποιούν ιντεΐνες είναι γονίδια κωδικοποίησης διχασμένων ιντεϊνών (split inteins) (Perler, 2002). Σε αντίθεση με τις πιο κοινές συνεχόμενες ιντεΐνες, αυτές μεταγράφονται και μεταφράζονται ως δύο ξεχωριστά πολυπεπτίδια, οι N-ιντεΐνη και C-ιντεΐνη, το καθένα ενωμένο με μία εξεΐνη. Μετά τη μετάφραση, τα αυθόρμητα θραύσματα της ιντεΐνης συναρμολογούνται μη ομοιοπολικά προς την κανονική δομή της ιντεΐνης για να πραγματοποιήσουν *in trans* πρωτεϊνικό ματίσμα (trans protein splicing) (Εικ 1.1B).

Παρά το γεγονός ότι πολλές από τις καταγωγές της διαίρεσης αυτών των ιντεϊνών προέκυψαν ανεξάρτητα κατά τη διάρκεια της εξέλιξης, όπως αποδεικνύεται από τις αποκλίνουσες ακολουθίες τους στην ενσωμάτωσή τους σε τουλάχιστον πέντε διαφορετικές πρωτεΐνες ξενιστές (Caspi et al., 2003a; Dassa et al., 2009), ο ακριβής μηχανισμός τους δεν είναι σαφής. Είναι ενδιαφέρον ότι, η περιοχή της διάσπασης στις περισσότερες αποτελεί την περιοχή εισαγωγής της ενδονουκλεάσης σε πολλές συνεχόμενες ιντεΐνες. Στην πραγματικότητα, ορισμένα γονίδια διχασμένων ιντεϊνών χωρίζονται από ένα εκτός-πλαισίου, αυτοτελές HEN γονίδιο, γεγονός που υποδηλώνει ότι η παρεκκλίνουσα εισαγωγή ενός HEN σε ένα γονίδιο ιντεΐνης θα μπορούσε να το σπάσει (Dassa et al., 2009). Παραδόξως, η ευρύτερα διαδεδομένη οικογένεια διχασμένων ιντεϊνών, που βρέθηκε μέσα στα γονίδια της DnaE σε τουλάχιστον 20 είδη κυανοβακτηρίων, έχει μη συντηρημένη γονιδιωματική αρχιτεκτονική (Caspi et al., 2003a). Στα διάφορα είδη κυανοβακτηρίων, οι ιντεΐνες αυτές βρίσκονται σε δραματικά διαφορετικές περιοχές του βακτηριακού χρωμοσώματος και σε ορισμένες περιπτώσεις, τα θραύσματα κωδικοποιούνται σε αντίθετους κλώνους. Αυτό υποδηλώνει ότι μετά την πρωταρχική διάσπαση του γονιδίου της ιντεΐνης σε πρώιμο κυανοβακτήριο, ο προκύπτων γενετικός τόπος ήταν ασταθής και ανασυνδυάστηκε περαιτέρω μέχρι ο οργανισμός να κατηγοριοποιηθεί σε άλλα είδη (Εικ 1.5.).

**Εικόνα 1.5.**

Προτεινόμενος μηχανισμός διάσπασης ιντεϊνών, λόγω της εισβολής μιας ενδονουκλεάσης παλιννόστησης (HEN), προς τη δημιουργία διχασμένης ιντεϊνης (Shah and Muir, 2014).

Οι ιντεΐνες αυτές παρουσιάζουν μοναδικές εξελικτικές προκλήσεις και ευκαιρίες για τους οργανισμούς ξενιστές. Από τη μία πλευρά, η διάσπαση του γονιδίου της ιντεϊνης πρέπει να είναι αποτελεσματικά μη αναστρέψιμη, και τα προκύπτοντα θραύσματα γονιδίου, αν εξακολουθούν να μεταγράφονται και μεταφράζονται, πρέπει να πραγματοποιήσουν το μάτισμα της εξτεϊνης *in trans*. Οι περιορισμοί αυτοί παρέχουν ένα ισχυρό φίλτρο επιλογής που είτε θα οδηγήσουν στην παύση της εν λόγω καταγωγής ή στη σημαντική βελτιστοποίηση της πρόσφατα διχασμένης ιντεϊνης. Από την άλλη πλευρά, αυτά τα χωρισμένα γονίδια θα μπορούσαν να προάγουν με έναν τρόπο την ρύθμιση της δραστηριότητας του γονιδίου ξενιστή. Επιπλέον, έχει διατυπωθεί η άποψη ότι εάν ένας οργανισμός περιέχει πολλαπλές διασταυρούμενες διχασμένες ιντεΐνες, θα μπορούσε να χρησιμεύσει ως πλατφόρμα για την εξέλιξη της πρωτεϊνης μέσω ανακατέματος δομικών μονάδων (Perler and Adam, 2000). Ενώ αυτή η πρόταση είναι ενδιαφέρουσα, μέχρι στιγμής δεν υπάρχει οργανισμός που έχει αναγνωριστεί ότι περιέχει περισσότερες από μία διχασμένες ιντεΐνες στο γονιδιώμά του. Από την άλλη πλευρά, μια πρόσφατη βιοχημική και δομική μελέτη έδειξε ότι

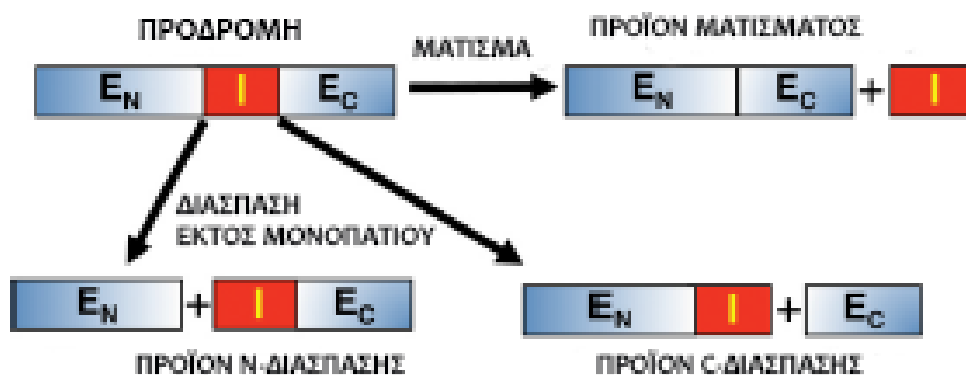
τα θραύσματα από τις διχασμένες ιντεΐνες μπορούν να κάνουν ανταλλαγή περιοχών με τις συνεχόμενες (Aranko et al., 2013a). Τα προκύπτοντα σύμπλοκα είναι ενεργά και μπορούν να αποδώσουν προϊόντα εναλλακτικού ματίσματος, διαφορετικά από εκείνα που κωδικοποιούνται σε ένα ενιαίο συντηγμένο γονίδιο ιντεΐνης-εξεΐνης. Αν και δεν είναι σαφές το φυσικό πλαίσιο του φαινομένου αυτού, έχει παρατηρηθεί ότι αρκετοί οργανισμοί περιέχουν μια μόνο διχασμένη ιντεΐνη και μία ή περισσότερες συνεχόμενες ιντεΐνες, κάτι που μπορεί να οδηγήσει στη συνεχή εναλλαγή πρωτεϊνικών περιοχών μεταξύ διχασμένων και συνεχόμενων που μπορεί να παρέχει τους οργανισμούς αυτούς με εξελικτικές ευκαιρίες μέσω του σχηματισμού ποικίλων πρωτεϊνικών προϊόντων.

Όπως αναφέρθηκε πιο πάνω οι ιντεΐνες εμφανίζονται σποραδικά μεταξύ συγγενικών ειδών και εμφανίζονται κατά βάση σε πρωτεΐνες που εμπλέκονται στο μεταβολισμό του DNA και ιδιαίτερα στα ενεργά κέντρα τους. Εξέταση της έμφυτης ικανότητας εισβολής των ιντεϊνών, η εξελικτική ιστορία των ειδών που τις περιέχουν, καθώς και τη φύση των πρωτεϊνών ξενιστών είναι απαραίτητη για μια βαθύτερη κατανόηση της δυναμικής και της διάδοσης των ιντεϊνών. Παρόλο που η απώλεια των ιντεϊνών είναι πιθανή λόγω της μείωσης του μεγέθους του γονιδιώματος σε πολλά είδη, η περιοδική εισβολή των ιντεϊνών είναι δυνατή σε μεγάλο βαθμό λόγω των εγγενών χαρακτηριστικών τους ως κινητά στοιχεία και την ικανότητά τους να εξαπλωθούν κάθετα και οριζόντια μεταξύ των ειδών. Η κατανόηση της εξέλιξης και της διάδοσης των ιντεϊνών θα δώσει φως σε θέματα που αφορούν το μηχανισμό λειτουργίας των ιντεϊνών.

1.2 Πρωτεϊνικό μάτισμα: ο μηχανισμός απεμπλοκής της ιντεΐνης από την πρόδρομη πρωτεΐνη

Το μάτισμα απαιτεί διαδοχικές αντιδράσεις πυρηνόφιλης προσβολής που καταλύονται από παρόμοιους μηχανισμούς πρωτεασών και λυασών ασπαραγίνης. Οι ιντεΐνες απαιτούν ακριβή συντονισμό της αντίδρασης και όχι γρήγορη κινητική ή εκλεκτική πρόσδεση του υποστρώματος, επειδή είναι τα μοναδικά ένζυμα αναδίπλωσης που είναι ομοιοπολικά συνδεδεμένα με τα υποστρώματά τους. Αυτό δίνει τη δυνατότητα στις ιντεΐνες να βρουν εναλλακτικούς μηχανισμούς με διαφορετικά βήματα ή να χρησιμοποιήσουν διαφορετικές διεργασίες για την ενεργοποίηση και το συντονισμό των βημάτων. Σημαντικά ζητήματα περιλαμβάνουν την κατανόηση των βασικών στοιχείων της κατάλυσης και πώς ελέγχονται τα βήματα του ματίσματος. Οι ιντεΐνες θεωρούνται

ένζυμα με κινητική, επειδή χρησιμοποιούν τις ίδιες στρατηγικές όπως τα κλασικά ένζυμα για να πραγματοποιηθεί η κατάλυση (Paulus, 2001). Το μάτισμα συμβαίνει απουσία οποιοδήποτε γνωστού συμπαράγοντα, σαπερόνης ή πηγής ενέργειας. Το μόνο που απαιτείται είναι η σωστή αναδίπλωση της ιντεΐνης στο πρόδρομο μόριο, για τη σωστή ευθυγράμμιση των πυρηνόφιλων καταλοίπων και των καταλοίπων που βοηθούν κατάλυση (Εικ. 1.7), με αποτέλεσμα ορισμένες ιντεΐνες να καλούνται καλλιτέχνες διαφυγής της φύσης (Evans and Xu, 1999). Επειδή ιντεΐνες χρησιμοποιούν όμοιες ομάδες πυρηνόφιλων, ανεπαίσθητες μεταβολές στη δραστικότητα αυτών των αμινοξέων απαιτούν χρήση διαφόρων βοηθητικών αμινοξικών καταλοίπων. Ως αποτέλεσμα, μερικά κατάλοιπα που διευκολύνουν τη κατάλυση να είναι ακόμα άγνωστα.



Εικόνα 1.6.

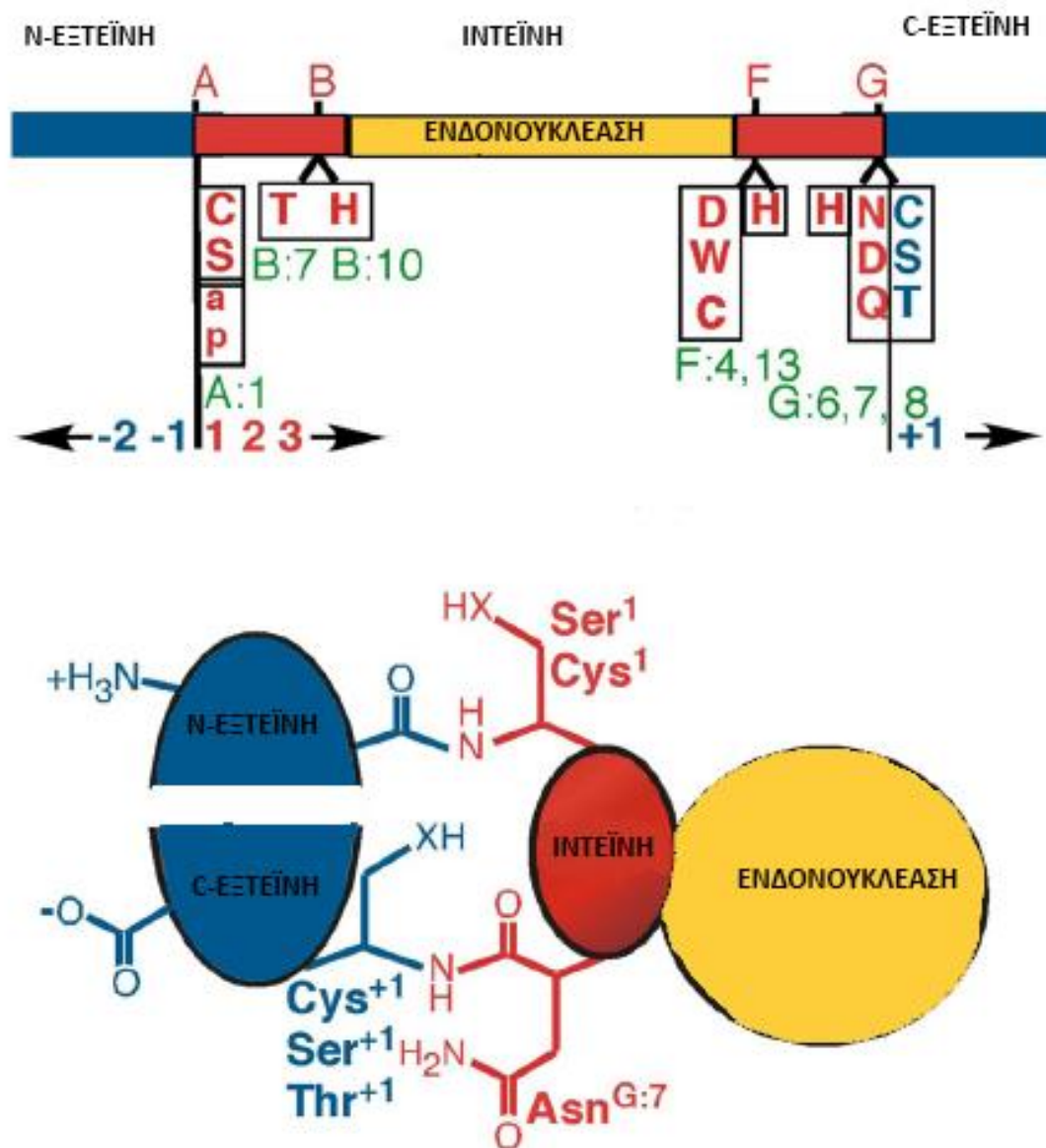
Οι πιθανές αντιδράσεις της ιντεΐνης. Εκτός από το πρωτεϊνικό μάτισμα υπάρχει η πιθανότητα της διάσπασης της ιντεΐνης από το ένα τμήμα εξτεΐνης.

Υπάρχουν τρεις κατηγορίες ιντεϊνών βάση των αλληλουχιών υπογραφής και των μηχανισμών ματίσματος (Tori et al., 2010). Ο μηχανισμός ματίσματος της πλέον πιο κλασικής κατηγορίας 1 ιντεϊνών (Εικ. 1.8) αποτελείται, όπως αναφέρθηκε και πριν, από τέσσερα στάδια 1) μια άκυλο-μετατόπιση για να μετατρέψει το N-τελικό πεπτιδικό δεσμό από ένα αμίδιο σε ένα θειοεστέρα, 2)

μία μετεστεροποίηση για να σχηματιστεί ένα διακλαδισμένο ενδιάμεσο, 3) κυκλοποίηση της Asn με διάσπαση της θέσης ματίσματος του C-τελικού και 4) μία δεύτερη άκυλο-μετατόπιση για το σχηματισμό ενός αμιδικού δεσμού μεταξύ των τμημάτων συντεθιμένης εξτεΐνης (Shao et al., 1996; Xu et al., 1994). Υπάρχει και η πιθανότητα διάσπασης, που συμβαίνει όταν ο συντονισμός των βημάτων διαταράσσεται με μετάλλαξη ή με έκφραση μεταξύ ξένων εξτεϊνών (Εικ. 1.6). Αυτό μπορεί να προκύψει από την αύξηση του ρυθμού διάσπασης σε αυτή τη θέση ματίσματος, τη μείωση του ρυθμού της αντίδρασης ενός άλλου σταδίου ή την παρεμβολή από μια διαμορφωτική αλλαγή λόγω μηχανισμού που απαιτείται για την προώθηση ενός καταβολικού βήματος.

Καθώς ο αριθμός των αλληλουχημένων ιντεϊνών έχει αυξηθεί, ευθυγραμμίσαις αλληλουχιών έχουν αποκαλύψει τέσσερα μοτίβα ματίσματος που ονομάζεται περιοχές A, B, F, και G (Εικ. 1.7) (Perler, 2002; Pietrokovski, 2001). Αν και οι περιοχές δεν είναι συντηρημένες στο σύνολό τους, αρκετές θέσεις σε κάθε μοτίβο περιέχουν υψηλά συντηρημένες ομάδες όμοιων αμινοξέων. Τα πυρηνόφιλα για κάθε βήμα είναι: Cys 1 ή Ser1 στο βήμα 1, Cys+1 ή Ser+1 ή Thr +1 στα βήματα 2 και 4 και η Asn στην περιοχή G: 7 του C-τελικού της ιντεΐνης στο βήμα 3 (Εικ. 1.7.). Γνωστά βοηθητικά κατάλοιπα περιλαμβάνουν τις θέσεις 7 και 10 στην περιοχή B (ThrB: 7 και HisB: 10), το προτελευταίο αμινοξύ της ιντεΐνης HisG: 6 και οι λιγότερο συντηρημένες θέσεις 4 και 13 στην περιοχή F (Εικ.1.7.). Η θέση F: 4 είναι συνήθως ένα Asp, ακολουθούμενο από μια Cys και μια Trp, ενώ η F:13 είναι συνήθως μια His (Mills et al., 2014; Perler, 2002; Tori et al., 2010).

Όπως έχει ήδη αναφερθεί υπάρχει μια μεγάλη ποικιλία ιντεϊνών. Οι περισσότερες είναι μεγάλες χίμαιρες που περιέχουν την περιοχή ματίσματος όπως και μια ενδονουκλεάση που ελέγχει την κινητικότητα ιντρονίων (Εικ.1.7) (Novikova et al., 2014; Pietrokovski, 2001). Άλλες είναι τόσο μικρές (mini-inteins) και έχουν μέχρι και 134 κατάλοιπα, ενώ στερούνται την ενδονουκλεάση. Μελέτες των φυσικών και ανασυνδυασμένων μικρών ιντεϊνών (mini-inteins) βοήθησαν στο να καθοριστεί η περιοχή της ιντεΐνης που πραγματοποιείται το μάτισμα (Mills et al., 2014; Pietrokovski, 2001). Τα γονίδια της ιντεΐνης μπορούν να χωριστούν μεταξύ των μοτίβων B και F. Ωστόσο τα εκφρασμένα θραύσματα της πρόδρομης πρωτεΐνης συναρμολογούνται ταχέως για να ακολουθηθεί μάτισμα *in trans* με τους ίδιους μηχανισμούς που χρησιμοποιούνται *in cis* (Iwai et al., 2006; Martin et al., 2001). Τόσο η εφαρμογή των φυσικών όσο και των ανασυνδυασμένων διχασμένων ιντεϊνών είναι σημαντική στη βιοτεχνολογία (Wood and Camarero, 2014).



Εικόνα 1.7.

Πρωτεϊνικές περιοχές στην πρόδρομη πρωτεΐνη και τα συντηρημένα μοτίβα. (Mills et al., 2014a)

Οι περιοχές ματίσματος μπορεί να έχουν προέλθει από αρχαία ένζυμα, επειδή είναι μικρές και στενά συνδεδεμένες δομικά, με συντηρημένα μοτίβα, όπως και να έχουν όμοιες ενζυματικές δραστηριότητες με τις Hedgehog πρωτεϊνικές περιοχές, οι οποίες ενεργοποιούν βασικές πρωτεΐνες σηματοδότησης για την ανάπτυξη μεταζώων (Hall et al., 1997). Υιοθέτηση ενός γενικού μηχανισμού για όλες τις ιντεΐνες σε ατομικό επίπεδο, είναι μάλλον απίθανο, διότι οι μέθοδοι που χρησιμοποιούνται για την προώθηση κάθε αναδιάταξης βοηθούνται από μια μεγάλη ποικιλία αμινοξικών καταλοίπων. Παρακάτω θα γίνει μια εκτενής ανάλυση των βημάτων του ματίσματος για την τάξη 1 ιντεϊνών και αντίστοιχα θα αναφερθούν περιπτώσεις εναλλακτικών μηχανισμών ματίσματος.

1.2.1 Μηχανισμός πρωτεϊνικού ματίσματος κατηγορίας 1 ιντεϊνών: 1^ο βήμα

Η μετατροπή των αμιδικών δεσμών σε θειοεστέρες ή εστέρες συνδεδεμένους με το ένζυμο, όπως παρατηρείται στο 1^ο βήμα του πρωτεϊνικού ματίσματος, είναι μια κοινή μέθοδος που χρησιμοποιείται στη κατάλυση από πρωτεάσες και αυτοεπεξεργαζόμενα ένζυμα, όπως και από γλυκοζυλοτρανσφεράσες (Evans and Xu, 1999; Paulus, 2001). Αν και το πρωτεϊνικό μάτισμα προκαλεί μια σειρά ομολόγων αναδιατάξεων και όχι τη διάσπαση του δεσμού που διευκολύνεται από πρωτεάσες, οι ιντεΐνες χρησιμοποιούν παρόμοιες στρατηγικές για να αποσταθεροποιήσουν το ομόλογο πεπτίδιο για να ευνοήσουν το σχηματισμό του θειοεστερικού δεσμού, με κλασική κατάλυση οξέος / βάσεως και μια οξυ-ανιονική οπή για τη σταθεροποίηση του τετραεδρικού ενδιάμεσου.

Τα περισσότερα κατάλοιπα που εμπλέκονται στην καταλυτική συναρμολόγηση κοντά στο κέντρο του δισκοειδούς HINT (Hedgehog-intein) αναδιπλώνονται κατά το μάτισμα (Hall et al., 1997). Για παράδειγμα, τα κατάλοιπα κοντά στο N-τελικό δεσμό σχάσης, περιλαμβάνουν τη ThrB: 7 και τη HisB: 10 σε μια τυπου Ι β-στροφή και τη Asp F: 4 σε μια β-πτυχή. Πειραματικά δεδομένα δείχνουν ότι η Thr B: 7, η HisB: 10 και τα κατάλοιπα F: 4 επηρεάζουν τη N-τελική αντίδραση ματίσματος, όπως και πλευρικά κατάλοιπα της εξτεΐνης (Perler, 2002; Volkmann and Mootz, 2013). Η HisB: 10 είναι το πιο συντηρημένο κατάλοιπο ιντεΐνης. Από τις δύο inteins που δεν έχουν τη HisB: 10, το ένα είναι ένα διασπασμένο ψευδογονιδίο, και το άλλο (η ιντεΐνη TKO CDC21-1 του *Thermococcus kodakaraensis*) χρησιμοποιεί τη Lys58 για να ενεργοποιήσει τη N-τελική θέση ματίσματος με πιθανή σταθεροποίηση του αρχικού τετραεδρικού ενδιάμεσου της N σε S άκυλο-μετατόπισης (Tori et al., 2012). Η Lys58 βρίσκεται έξω από τα συντηρημένα μοτίβα ιντεϊνών (Tori

et al., 2012). Τα κατάλοιπα της N-εξεΐνης φαίνεται να επηρεάζουν την ισορροπία μεταξύ αμιδίου και εστέρα στο N-τερματικό άκρο της Sce VMA ιντεΐνης (Chong et al., 1998a), ενώ οι αντιδράσεις στο N-τελικό άκρο της ιντεΐνης Pho RADA του *Pyrococcus horikoshii* επηρεάζονται από van der Waals αλληλεπιδράσεις από κατάλοιπα της αλληλουχίας της (Oeemig et al., 2012). Ορισμένες ιντεΐνες παραμορφώνουν το N-τελικό δεσμό σχάσης δημιουργώντας καταλυτική τόνωση ώστε να επιτευχθεί το 1^ο βήμα. Γειτονικά κατάλοιπα εξεΐνης και κατάλοιπα της περιοχής B βοηθούν στο σχηματισμό αυτής της τεταμένης τοπικής διάπλασης, όπως αποδεικνύεται από δομικές και βιοχημικές μελέτες. Η κρυσταλλική δομή της SCE VMA εμφανίζει ομολόγες παραμορφώσεις κοντά στο N-τελικό δεσμό ματίσματος (Poland et al., 2000).

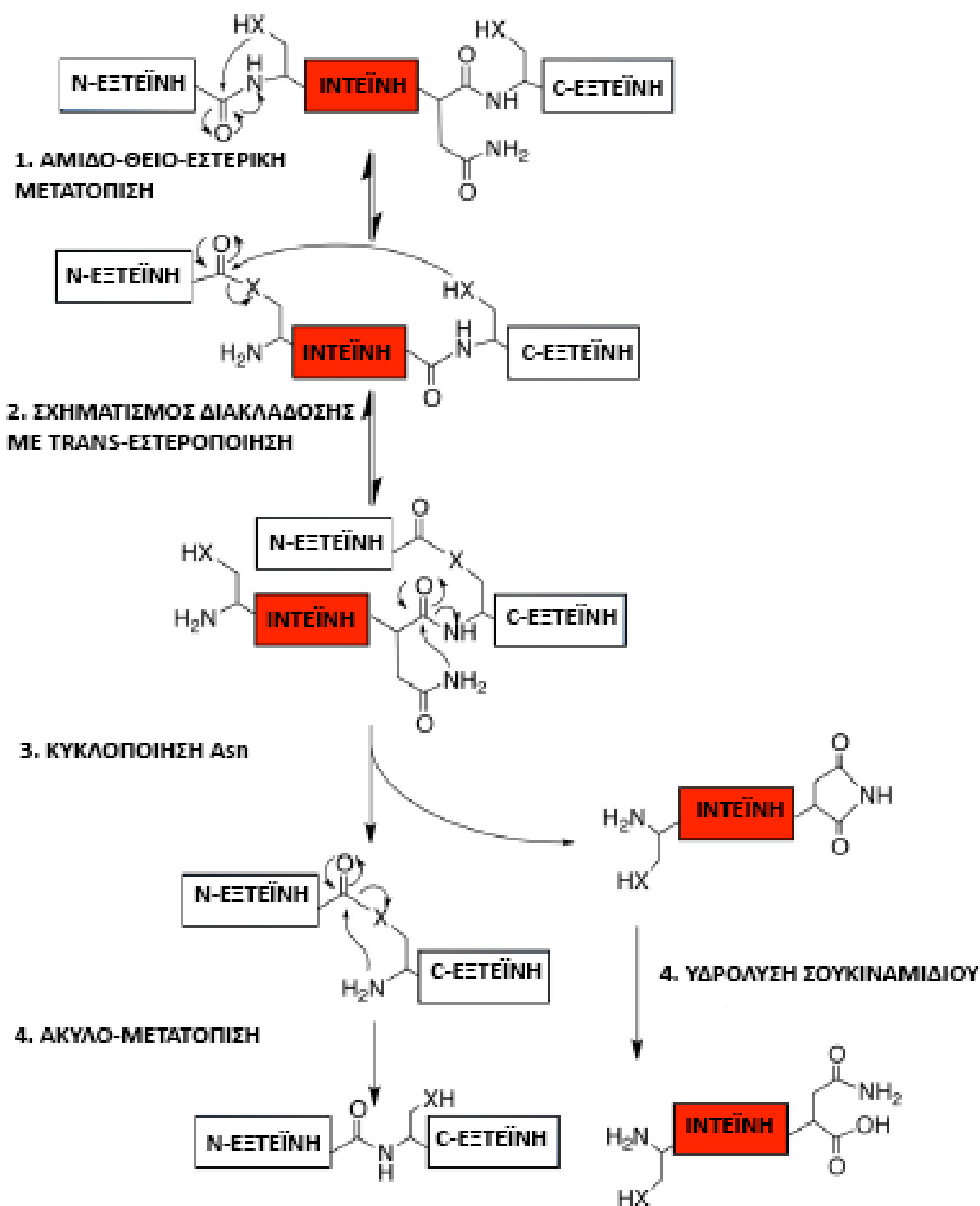
Η κρυσταλλική δομή της MXE GyrA ιντεΐνης στο *Mycobacterium xenopi* υποδεικνύει ένα cis-πεπτιδικό δεσμό που συνδέει τη N-εξεΐνη και την ιντεΐνη, ενώ δεδομένα από NMR υποδεικνύουν έλλειψη αμιδικού δεσμού όταν η His B: 10 μεταλλάσσεται (Romanelli et al., 2004). Η His B: 10 είναι σε απόσταση ενός δεσμού υδρογόνου με το άζωτο αμιδίου του N-τελικού δεσμού σχάσης σε αρκετές ιντεΐνες, γεγονός που υποδηλώνει ότι παίζει ρόλο στο συντονισμό της δεσμού σχάσης (Poland et al., 2000). Παρόμοιος ρόλος παρατηρήθηκε για τη ThrB: 7 από τη δομή της Ssp DnaE του *Synechocystis* (Dearden et al., 2013).

1.2.2 Μηχανισμός πρωτεϊνικού ματίσματος κατηγορίας 1 ιντεϊνών: 2^ο βήμα

Το δεύτερο βήμα έχει αποδειχθεί ότι είναι το πιο δύσκολο να μελετηθεί, καθώς είναι δύσκολο να απομονωθούν τα διακλαδισμένα ενδιάμεσα. Μεταλλάξεις που κανονικά πρέπει να οδηγήσουν σε συσσώρευση των διακλαδισμένων ενδιάμεσων συχνά οδηγούν σε αποσύνθεση σε προϊόντα N-τελικής διάσπασης, ειδικά όταν είναι παρόν ένας θειοεστερικός δεσμός. Ο σχηματισμός εστερο-συνδεδεμένων διακλαδισμένων ενδιάμεσων είναι αναστρέψιμος, πράγμα που μπορεί να οδηγήσει σε συσσώρευση του προδρόμου αντί του ενδιάμεσου (Mills et al., 2014).

1.2.3 Μηχανισμός πρωτεϊνικού ματίσματος κατηγορίας 1 ιντεϊνών: 3^ο βήμα

Αποδεικτικά στοιχεία για το 3^ο βήμα ματίσματος περιλαμβάνουν την απώλεια της C-τελικής θέσης διάσπασης του ματίσματος μετά τη μετάλλαξη ιντεΐνης στη C-τελική AsnG: 7 και την ανίχνευση αποκομμένων ιντεϊνών με C-τελικό σουκιναμίδιο (Shao et al., 1996; Xu et al., 1994). Η ιντεΐνη πρέπει καταλύσει την κυκλοποίηση της Asn, προκαλώντας απαμίδωση της πλευρικής αλυσίδας σε αντίθεση με άλλα συστήματα που γίνεται διάσπαση πεπτιδικού δεσμού (Brennan and Clarke, 1995).



Εικόνα 1.8.

Μηχανισμός πρωτεϊνικού ματίσματος ιντεϊνών της κατηγορίας I. (Volkman and Mootz, 2013)

Αρκετές στρατηγικές έχουν προταθεί για την ενζυμική ενεργοποίηση του 3^{ου} βήματος συμπεριλαμβανομένων τριών συζευγμένων τρόπων κατάλυσης: 1) Η His F: 13 αυξάνει την πυρηνοφιλία της C-τελικής AsnG: 7 με αποπρωτονίωση, 2) Το τετραεδρικό ενδιάμεσο σταθεροποιείται από τα φορτισμένα κατάλοιπα His F: 13 και His G: 6 και 3) Η ηλεκτροφιλικότητα του αμιδίου σκελετού μπορεί να αυξηθεί από τη His G: 6 (Mujika et al., 2012)

Εναλλακτικά, δεδομένου ότι η C-τελική αποκοπή ευνοείται σε χαμηλό pH, η πρωτονίωση του αζώτου του αμιδίου του σχάσιμου πεπτιδικού δεσμού μπορεί να έχει προτεραιότητα σε σχέση με την αποπρωτονίωση του αμιδίου της πλευρικής αλυσίδας της Asn (Mills et al., 2005). Ξεχωριστές μελέτες δείχνουν δύο άλλους τρόπους κατάλυσης: αλλαγή στο τοπικό περιβάλλον κοντά στο σχάσιμο δεσμό που εξαρτάται από το σχηματισμό διακλαδισμένου εστέρα και την αποσταθεροποίηση της σχιζόμενου πολωμένου ομολόγου κατάλοιπου C-εξτεΐνης (Shemella et al., 2011).

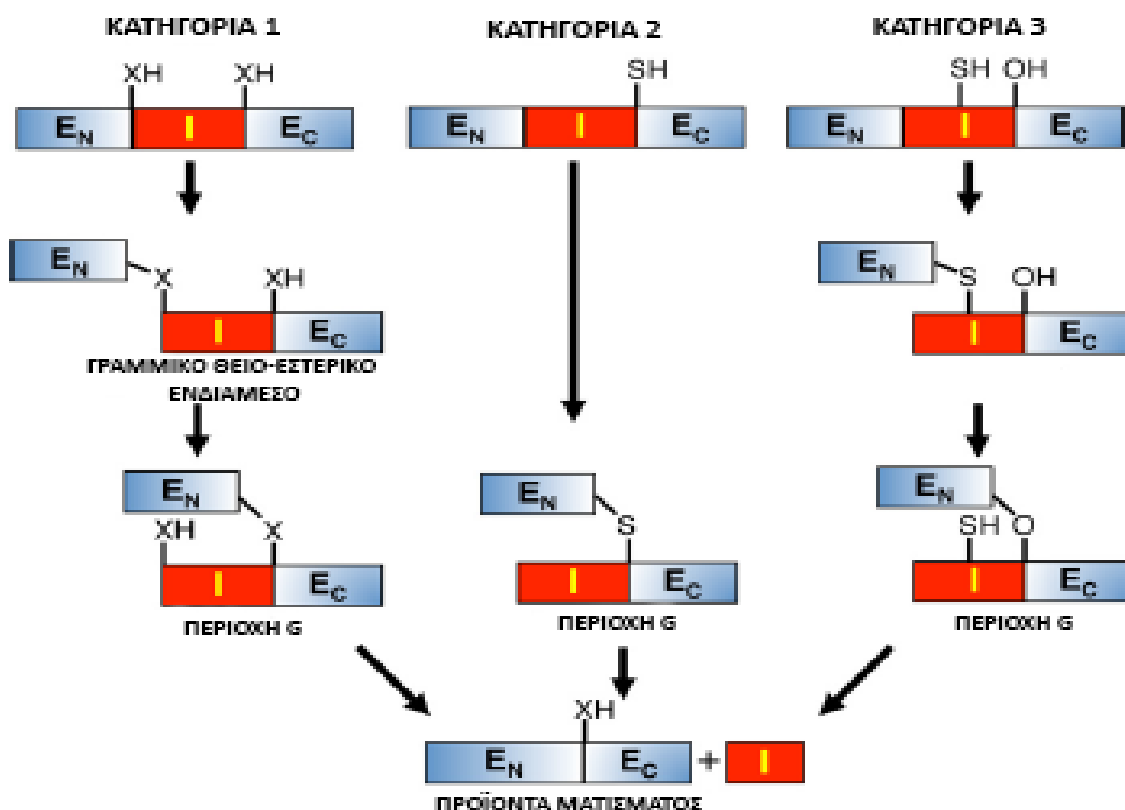
1.2.4 Μηχανισμός πρωτεϊνικού ματίσματος κατηγορίας 1 ιντεϊνών: 4^ο βήμα

Το 4^ο βήμα αποτελείται από δύο στάδια φινιρίσματος, σε κανένα από τα οποία η ιντεΐνη καταλύει τις αντιδράσεις. Η αμινοσουκινομιδική ιντεΐνη υδρολύεται βραδέως προς Asn ή ισο-Asn (Shao et al., 1996) και ο θειοεστέρας που συνδέει τα τμήματα της εξτεΐνης επανέρχεται στην αμιδική μορφή. Πειράματα με πεπτιδικά μοντέλα αποδεικνύουν ότι ο ρυθμός μετατροπής του θειοεστέρα σε αμίδιο είναι μεγαλύτερος από το συνολικό ρυθμό του ματίσματος (Shao and Paulus, 1997). Αυτή η τελική μετατόπιση ακυλίου ευνοείται θερμοδυναμικά και δεν επηρεάζεται από την παρουσία της ιντεΐνης (Frutos et al., 2010).

1.2.5 Άλλοι μηχανισμοί πρωτεϊνικού ματίσματος

Η στιβαρότητα του μηχανισμού πρωτεϊνικού ματίσματος των ιντεϊνών απεικονίζεται από τη συστοιχία των αποδεκτών τροποποιήσεις στον συνήθη μηχανισμό τεσσάρων βημάτων. Οι BIL στερούνται το +1 νουκλεόφιλο κατάλοιπο της C-εξτεΐνης και επομένως δεν είναι σε θέση να σχηματίσουν το διακλαδισμένο ενδιάμεσο (Aranko et al., 2013b). Οι δύο άλλες κατηγορίες ιντεϊνών μπορούν να προκαλέσουν πρωτεϊνικό μάτισμα, με μόνη διαφορά ότι στερούνται το πυρηνόφιλο κατάλοιπο Ser1 ή Cys 1 και συνεπώς δε μπορούν να παράγουν το γραμμικό ενδιάμεσο θειοεστέρα (Εικ. 1.9.)

Παρά το γεγονός ότι τα βασικά βήματα του πρωτεϊνικού ματίσματος έχουν διευκρινιστεί από τη δεκαετία του 90, υπάρχουν ακόμα ένα ερωτηματικό όσον αναφορά το συντανισμό. Οι δύο βασικές διαδικασίες προτείνονται: 1) οι διαμορφωτικές αλλαγές ενεργοποιούνται ως αποτέλεσμα ενός προηγούμενου σταδίου για το σχηματισμό ενός ισχυρού ενεργού κέντρου για το επόμενο βήμα και 2) οι διαφορές στις κινητικές τιμές για κάθε βαθμίδα εξασφαλίζουν τη σωστή σειρά της αντίδρασης. Οι διαμορφωτικές αλλαγές μπορεί να είναι απλές, για παράδειγμα καθορισμός διαφορετικών θέσεων ροτεμερούς, ή μπορεί να συνεπάγονται μεγαλύτερες κινήσεις. Αποδεικτικά στοιχεία για το διαμορφωτικό έλεγχο μπορούν να συλλεχθούν από την απόλυτη σύζευξη του N-τελικού και του C-τελικού, αντιδράσεις που παρατηρούνται σε ορισμένες ιντεΐνες όπου η διάσπαση του C-τελικού λαμβάνει χώρα μόνο εάν τα προηγούμενα στάδια έχουν ολοκληρωθεί (Frutos et al., 2010).



Εικόνα 1.9.

Παραλλαγές του μηχανισμού ματίσματος. Δείχνονται οι μηχανισμοί και των τριών κατηγοριών ιντεϊνών. Διαφορές φαίνονται στο δεύτερο βήμα του μηχανισμού πρωτεϊνικού ματίσματος (Mills et al., 2014).

Ο πιο εύκολος τρόπος για την εύρεση επιχειρημάτων για μεγαλύτερης κλίμακας μετακινήσεις ενεργών κέντρων των ιντεϊνών είναι μέσω της ανάλυσης των δομών τους. Μόνο οι δομές των Sce VMA (Mizutani et al., 2004) και Pho RADA (Oeemig et al., 2012) έχουν αποστάσεις μεταξύ του νουκλεόφιλου C-εξεΐνης και N-τελικού σχάσιμου δεσμού που είναι άμεσα συμβατές με τη κατάλυση (3.8 Å). Αυτή η απόσταση είναι πολύ μεγαλύτερη (8 Å) σε όλες τις άλλες δομές που υπάρχουν μέχρι σήμερα και απαιτούν μια διαμορφωτική αλλαγή για τη κατάλυση (Eryilmaz et al., 2014). Μένει να προσδιοριστεί γιατί οι ιντεΐνες εμφανίζουν ένα τόσο ανοικτό ενεργό κέντρο και αν αυτό αντιπροσωπεύει μια πραγματική διαμόρφωση ή ένα αποτέλεσμα των πειραματικών συνθηκών που εμποδίζουν το μάτισμα, συμπεριλαμβανομένων των μεταλλάξεων στα κατάλοιπα του ενεργού κέντρου και διαφορές στις αλληλουχίες ή το μέγεθος των εξεΐνων.

Κίνηση των πλευρικών αλυσίδων κατά τη διάρκεια του ματίσματος μπορεί να συντονίσει την αντίδραση με δημιουργία ή απώλεια δεσμών υδρογόνου και αλλαγές στις αλληλεπιδράσεις πακεταρίσματος van der Waals για την ευθυγράμμιση των καταλυτικών κατάλοιπων. Για παράδειγμα, οι δομές της Pho Rada δείχνουν ότι οι δεσμοί υδρογόνου του Asp F: 4 στη Asn G: 7, εμποδίζουν τη κυκλοποίηση της Asn μέχρι ο σχηματισμός του διακλαδισμένου ενδιάμεσου προκαλέσει αναπροσανατολισμό στη πλευρική αλυσίδα της Asp F: 4 (Oeemig et al., 2012). Ένα άλλο παράδειγμα προτείνει τη συζευγμένη διάσπαση του N-και C-τελικού στην Ssp DnaE, η οποία οφείλεται στη Tyr -1 που εμποδίζει το κατάλληλο προσανατολισμό της Arg B:11 έως το σχηματισμό του γραμμικού ή και διακλαδισμένου ενδιάμεσου θειοεστέρα που καταλήγει στη μετακίνηση της πλευρικής αλυσίδας της Tyr -1, επιτρέποντας στη πλευρική αλυσίδα της Arg B:11 να αναπροσανατολιστεί και να βοηθήσει στη κυκλοποίηση της Asn (Sun et al., 2005).

1.2.6 Πρωτεϊνικό μάτισμα υπό προϋποθέσεις

Τόσο οι φυσικές όσο και σχεδιασμένες ιντεΐνες μπορούν να πραγματοποιήσουν υπό όρους το πρωτεϊνικό μάτισμα (Conditional Protein Splicing CPS). Οι προϋποθέσεις εξαρτώνται από την παρουσία ενός συγκεκριμένου παράγοντα, όπως είναι η αλλαγή στην οξειδοαναγωγική κατάσταση, η θερμοκρασία, μικρά μόρια, ή το φως όπως έχει πρόσφατα αναφερθεί (Shah and Muir, 2014). Η ύπαρξη των παραγόντο-εξαρτώμενων ιντεϊνών υποδηλώνει την πιθανότητα ότι μερικές μπορούν να προσαρμόζονται στην ενδοκυτταρική τους θέση με το να γίνονται μετα-

μεταφραστικά ρυθμιστικά στοιχεία που ρυθμίζουν το πρωτεϊνικό μάτισμα σε συμφωνία με τις περιβαλλοντικές συνθήκες.

Ένα παράδειγμα ρυθμιζόμενου ματίσματος παρέχεται με σχηματισμό ενός δισουλφιδικού δεσμού από κατάλοιπα κυστεΐνης που εμπλέκονται στον μηχανισμό ματίσματος, παγιδεύοντας την ιντεΐνη στην πρόδρομη πρωτεΐνη ανάλογα με την οξειδοαναγωγική κατάσταση. Έχουν σχεδιαστεί αρκετές ιντεΐνες που σχηματίζουν δισουλφιδικό δεσμό για να ελέγχεται πρόωρα η διάσπαση ή η αντίδραση ματίσματος (Callahan et al., 2011, 2013) (Εικ.1.1Γ). Μια από αυτές ενέπνευσε μια έρευνα για φυσιολογική ρύθμιση του ματίσματος με ένα δισουλφιδικό δεσμό. Μετά τον καθορισμό από ανάλυση μεταλλάξεων αποδείχθηκε ότι η κυστεΐνη στη θέση +3 της εξτεΐνης μπορεί να παγιδεύσει τη κυστεΐνη (Cys 1) του καταλυτικού κέντρου της Ssp DnaE ιντεΐνης (Callahan et al., 2011, 2013), κάτι που έχει βρεθεί και σε άλλες ιντεΐνες όπως την MoaA από το θερμόφιλο αρχαίο *Pyrococcus abyssi*.

Εντυπωσιακά, έχει παρατηρηθεί ότι ορισμένες ιντεΐνες στις οποίες προβλέπεται η οξειδοαναγωγική ρύθμιση, οι εξτεϊνικές περιοχές εμφανίζουν ένα συντηρημένο μοτίβο CysXXXCysXXCys (όπου το X είναι οποιοδήποτε αμινοξύ). Σε όλες τις περιπτώσεις, η τελευταία Cys του μοτίβου CysXXX- CysXXCys είναι Cys 1 της ιντεΐνης, ενώ η Cys στη μέση αντιστοιχεί στην Cys +3 η οποία σχηματίζει ένα το δισουλφιδικό δεσμό με τη Cys 1 του MoaA. Χαρακτηρισμός της ιντεΐνης MoaA στη *E.coli* αποκάλυψε ότι ο δισουλφιδικός δεσμός Cys 3-προς-Cys 1 μπορεί να ελέγξει τη δραστηριότητα του ματίσματος, ανάλογα με την οξειδοαναγωγική κατάσταση του οργανισμού ξενιστή (Callahan et al., 2011).

Υπάρχουν και άλλα φυσικά παραδείγματα οξειδοαναγωγο-εξαρτώμενων ιντεϊνών. Η ιντεΐνη PolII του *P. abyssi*, μπορεί να σχηματίσει έναν δισουλφιδικό δεσμό μεταξύ Cys 1 και της Cys -1 της εξτεΐνης που αποτρέπει το ματίσμα (Chen et al., 2012). Επίσης η MMA PolII από το μεθανογόνο αρχαίο *Methanoculleus marisnigri*, μπορεί να σχηματίσει έναν εσωτερικό δισουλφιδικό δεσμό που διαμορφώνει τη δραστηριότητα ματίσματος της ιντεΐνης με την κατάσταση οξειδοαναγωγής της *E.coli* ή τον εντοπισμό της στο περίπλασμα ή το κυτταρόπλασμα (Nicastri et al., 2013). Το πρωτεϊνικό μάτισμα μπορεί επίσης να επηρεάζεται από τη θερμοκρασία. Αυτή η ρύθμιση μπορεί να έχει φυσιολογική σημασία, καθώς αποδείχθηκε ότι η δραστηριότητα των διαφόρων ιντεϊνών από ακραία θερμόφιλα εξαρτάται από τη θερμοκρασία (Mills et al., 2005; Xu and Perler, 1996). Επιπλέον, ορισμένες έχουν μια μέτρια εξάρτηση από το pH με μια προτίμηση για χαμηλό pH, το οποίο μπορεί να ενισχυθεί με μετάλλαξη (Wood et al., 1999).

Η ύπαρξη τόσο φυσικών όσο και ανασυνδιασμένων ιντεϊνών με CPS αναμένει περισσότερα παραδείγματα *in vivo* ρυθμιζόμενων ματίσμάτων με την εμφάνιση μιας λειτουργικής πρωτεΐνης που εξαρτάται από ένα συγκεκριμένο περιβαλλοντικό ερέθισμα. Πιστεύεται ότι οι φυσικές ιντεΐνες έχουν εξελιχθεί από επιδρομικά παράσιτα σε επίμονους μπουτουαλιστές που παρέχουν προσαρμοστική μετα-μεταφραστικά ρυθμιστικά πλεονεκτήματα για την επιβίωση του ξενιστή υπό συγκεκριμένες κυτταρικές και περιβαλλοντικές συνθήκες.

Μια μηχανιστική πρόκληση περιλαμβάνει την αποκρυπτογράφηση του πώς οι αντιδράσεις είναι συντονισμένες αλλά και η ανάλυση των ποικίλων τρόπων που οι ιντεΐνες προωθούν την κατάλυση. Είναι σημαντικό να αναφερθεί ότι πρέπει να γίνουν λεπτομερείς μελέτες των καταλυτικών μηχανισμών της ιντεϊνικής κινητικής και των δομών στο πλαίσιο των φυσικών εξτεϊνικών πρωτεϊνών, που μπορούν να τονίσουν τη διάκριση μεταξύ σημαντικών φυσιολογικών παρατηρήσεων και εκείνων που μπορεί να είναι αντικείμενα μελέτης των ετερόλογων συστημάτων. Επιπλέον, οι λεπτομερείς μελέτες πολλαπλών ιντεϊνών θα καθορίσουν κατά πόσο οι καταλυτικές στρατηγικές είναι καθολικές ή ειδικές για ένα υποσύνολο των ιντεϊνών.

1.3 Δομικά και δυναμικά χαρακτηριστικά της ιντεΐνης

Ενώ αρχικά πιστευόταν ότι είναι μια ανωμαλία που βρέθηκε σε λίγους μόνο οργανισμούς, το πρωτεϊνικό μάτισμα από τις ιντεΐνες έχει παρατηρηθεί έκτοτε σε μικροοργανισμούς από όλες τις επικράτειες της ζωής. Παρά την ευρεία φυλογενετική κατανομή, όλες οι ιντεΐνες μοιράζονται κοινά δομικά χαρακτηριστικά, όπως ψεύδο-συμμετρικές δομές και τύπου πέταλο, αρκετά κανονικά μοτίβα αλληλουχίας και παρόμοιους μηχανισμούς ματίσματος. Με περίεργο τρόπο, η βελτίωση της αποτελεσματικότητας του ματίσματος και η εξειδίκευση του υποστρώματος των διαφόρων ιντεϊνών διαφέρουν σημαντικά, άμα παρατηρήσουμε τις αλλαγές στο χημικό μηχανισμό ματίσματος, που συνδέεται με την τοπική δομή και τη δυναμική τους. Είναι ευρέως γνωστό πως η χημεία των ιντεϊνών παίζει σημαντικό ρόλο στις εφαρμογές της πρωτεϊνικής χημείας, όπως αντίστοιχα η κατανόηση των δομικών και δυναμικών πτυχών των ιντεϊνών είναι ζωτικής σημασίας για την μηχανική ιντεϊνών και της βελτίωσης των τεχνολογιών που βασίζονται σε αυτές. Παρά το γεγονός ότι όλες οι ιντεΐνες παρατηρούνται με παρόμοιο δίπλωμα και έχουν εντόνως συντηρημένα μοτίβα αλληλουχίας στις δραστικές θέσεις τους, οι ιντεΐνες έχουν εκπληκτικά διαφορετική αποτελεσματικότητα ματίσματος όπως και αντίστοιχα οι προτιμήσεις στις αλληλουχίες εξτεϊνης διαφέρουν δραματικά.

1.3.1 Αναδίπλωση ιντεϊνών

Η συντηρημένη δομή τύπου πετάλου (Εικ. 1.10Α) έχει παρατηρηθεί σε όλες τις δομές ιντεϊνών που έχουν λυθεί με NMR και κρυσταλλογραφία ακτίνων Χ (Aranko et al., 2013a, 2013b). Η αναδίπλωση περιλαμβάνει κυρίως β-πτυχές, λούπες και δύο μικρές έλικες, ενώ παρατηρείται και από ψευδό-συμμετρία (Εικ. 1.10Β). Δεδομένης αυτής της συμμετρίας, έχει προταθεί ότι αυτή η αναδίπλωση που προέκυψε μπορεί να οφείλεται σε μια εκδήλωση διπλασιασμού γονιδίου κάποιας μητρικής πρωτεΐνης (Hall et al., 1997). Η αναδίπλωση της έχει τρία αξιοσημείωτα χαρακτηριστικά: 1) η τοπολογία είναι πολύπλοκη, που συνεπάγεται πολλαπλά περάσματα της πολυπεπτιδικής αλυσίδας πίσω και εμπρός μεταξύ των συμμετρικών κομματιών (Εικ. 1.10Β), 2) τα άκρα της ιντεΐνης που φέρουν την εξτεΐνη έρχονται σε κοντινή απόσταση μεταξύ τους (10 Å) για μάτισμα και 3) οι πολλαπλές δραστικές θέσεις που μοιάζουν με ανάλογα πρωτεασών αποτελούνται από συντηρημένα μοτίβα αλληλουχίας χτισμένα γύρω από αυτά τα άκρα για τη διεξαγωγή όλων τα χημικών βημάτων που εμπλέκονται στο πρωτεϊνικό μάτισμα. Η αναδίπλωση της ιντεΐνης ακολουθεί την αντίδραση. Η αρχική δομή διευκολύνει το πρώτο βήμα του ματίσματος και στη συνέχεια κάθε βήμα του ματίσματος προκαλεί τοπική αναδιαμόρφωση που επηρεάζει την αναδίπλωση του καταλυτικού κέντρου και ως εκ τούτου επηρεάζοντας τον συντονισμό της αντίδρασης και μετατοπίζοντας της θέση ισορροπίας.

Τα μοτίβα του ματίσματος βρίσκονται σε δύο περιοχές ματίσματος, στη Ν-τελική περιοχή ματίσματος (N-intein) και στη C-τελική περιοχή ματίσματος (C-intein) (Εικ. 1.11.). Στις πρότυπες συνεχόμενες ιντεΐνες (cis-μάτισμα), οι δύο αυτές περιοχές ματίσματος χωρίζονται από μια ενδονουκλεάση ή από μια συνδετική αλληλουχία, ενώ στις διχασμένες ιντεΐνες (trans-μάτισμα), οι περιοχές ματίσματος μεταφράζονται ξεχωριστά. Σε μια λειτουργική αναδίπλωση ιντεΐνης, οι δύο περιοχές ματίσματος, είτε στη συνεχόμενη είτε στη διχασμένη, περιπλέκονται για να φέρουν όλα τα καταλυτικά κατάλοιπα και τα άκρα μαζί, το οποίο είναι ένα μοναδικό πρόβλημα αναδίπλωσης ή μοριακής αναγνώρισης που προηγείται του πρωτεϊνικού ματίσματος. Παρά το γεγονός ότι το πρωτεϊνικό μάτισμα ξεκινάει με αυτό το δίπλωμα, λίγες μελέτες έχουν ασχοληθεί με την αναδίπλωση των ιντεϊνών (Shah et al., 2013a; Sun et al., 2005).

Η αλληλένδετη αναδίπλωση των θραυσμάτων απαιτεί κάποια διαταραχή στα επιμέρους κομμάτια. Στη NPU DnaE διχασμένη ιντεΐνη, η C-ιντεΐνη (NpuC) είναι εντελώς διαταραγμένη, ενώ η Ν-ιντεΐνη είναι μερικώς διαταραγμένη στο C-τελικό άκρο (NpuNC) (Sun et al., 2005)(Εικ. 1.10Γ).

Η αναδίπλωση εκκινεί με ηλεκτροστατική συμπληρωματικότητα φορτίου μεταξύ της διαταραγμένης C-ιντεΐνης και το διαταραγμένο C-τελικό άκρο της N-ιντεΐνης (σύλληψη), στη συνέχεια το ενδιάμεσο σταθεροποιείται περαιτέρω με υδρόφοβες αλληλεπιδράσεις μεταξύ της C-ιντεΐνης και του N-τελικού άκρου της N-ιντεΐνης (κατάπτωση) (Shah et al., 2013a). Ο μηχανισμός "σύλληψης και κατάπτωσης" αποκαλύπτει ένα γενικό τρόπο αναδίπλωσης για όλες τις ιντεΐνες συμπεριλαμβανομένων και των συνεχόμενων ιντεϊνών. Η έναρξη ηλεκτροστατικών αλληλεπιδράσεων είναι μοναδική στις διχασμένες ιντεΐνες (Shah et al., 2011), αλλά είναι πιθανό ότι η «περιοχή διάσπασης» στις διχασμένες ιντεΐνες να ήταν ένα σημείο ένωσης στις συνεχόμενες. Στις συνεχόμενες με περιοχή ενδονουκλεάσης (HEN), η αναδίπλωση της περιοχής της ενδονουκλεάσης θα μπορούσε προκαλεί την έναρξη της αναδίπλωσης της ιντεΐνης. Η εγγενής διαταραχή που παρατηρείται στη NPU DnaE ιντεΐνη παρατηρείται επίσης και στη συγγενική Ssp DnaE ιντεΐνη, υποστηρίζοντας περαιτέρω την άποψη ότι η διεργασία αναδίπλωσης είναι συντηρημένη διαδικασία (Zheng et al., 2012). Οι N- και C-τελικές περιοχές ματίσματος διαφορετικών μορίων ιντεϊνών μπορούν να αλληλοεπιδράσουν για να σχηματίσουν ένα διμερές από αποτέλεσμα ανταλλαγής (Εικ. 1.10Δ). Δεδομένου ότι αυτό το είδος της μη ομοιοπολικής αλληλεπίδρασης μεταξύ των διαφόρων διχασμένων ιντεϊνών δημιουργεί υβριδικά ενεργές θέσεις και συνδυασμούς μικτών εξτεϊνών, έχει προταθεί ότι τέτοιες δομές μπορούν να δώσουν εναλλακτικό προϊόν ματίσματος, προϊόντα που θα μπορούσαν να επισπεύσουν τη διαδικασία εξέλιξης με την δημιουργία αναπληρωματικών φαινοτύπων υπό σταθερό γενετικό υπόβαθρο (Aranko et al., 2013a).

1.3.2 Η δομική βάση του πρωτεϊνικού ματίσματος

Το πρωτεϊνικό μάτισμα της κατηγορίας 1 ιντεϊνών είναι μια διαδικασία πολλών σταδίων (Εικ. 1.11). Όπως αναφέρεται πιο πάνω το πρώτο βήμα είναι N σε S / O ακυλό-μετατόπιση που ξεκινάει από το N-τελικό κατάλοιπο ιντεΐνης, μια κυστεΐνη ή σερίνη (Perler et al., 1997), που ενεργεί στο -1 κατάλοιπο του καρβονυλικού άνθρακα (Perler, 1998), με αποτέλεσμα τον σχηματισμό ενός θειοεστερικού γραμμικού ενδιαμέσου. Το επόμενο βήμα είναι η trans- θειοεστεροποίηση που προκαλείται από μια πυρηνόφιλη προσβολή από το κατάλοιπο +1 της C-εξτεΐνης το οποίο μπορεί να είναι μια κυστεΐνη ή μια σερίνη ή μια θρεονίνη. Το προκύπτον διακλαδισμένο ενδιάμεσο θειοεστέρα τροποποιείται με την κυκλοποίηση της συντηρημένης C-τελικής ασπαραγίνης της

ιντεΐνης. Μόλις σχηματιστεί το σουκιναμίδιο η ιντεΐνη αποκόπτεται από τις εξτεΐνες. Τέλος, η σύνδεση μεταξύ των εξτεϊνών διαμορφώνεται από ένα θειοεστέρα σε ένα αμίδιο ανεξάρτητα από την ιντεΐνη και το C-τελικό σουκιναμίδιο της ιντεΐνης υδρολύεται βραδέως (Frutos et al., 2010). Όλα αυτά τα διαδοχικά στάδια απαιτούν πολλαπλές δραστικές θέσεις με διακριτά κατάλοιπα όλα σε κοντινή απόσταση από τα άκρα της ιντεΐνης. Αυτά τα κατάλοιπα στο ενεργό κέντρο είναι συντηρημένα μεταξύ των περισσότερων ιντεϊνών, όμως υπάρχουν αρκετές εξαιρέσεις. Παρακάτω θα γίνει μια λεπτομερής περιγραφή του κάθε βήματος του πρωτεϊνικού ματίσματος της κατηγορίας 1 από μια δομική σκοπιά.

N-σε-O/S άκυλο-μετατόπιση

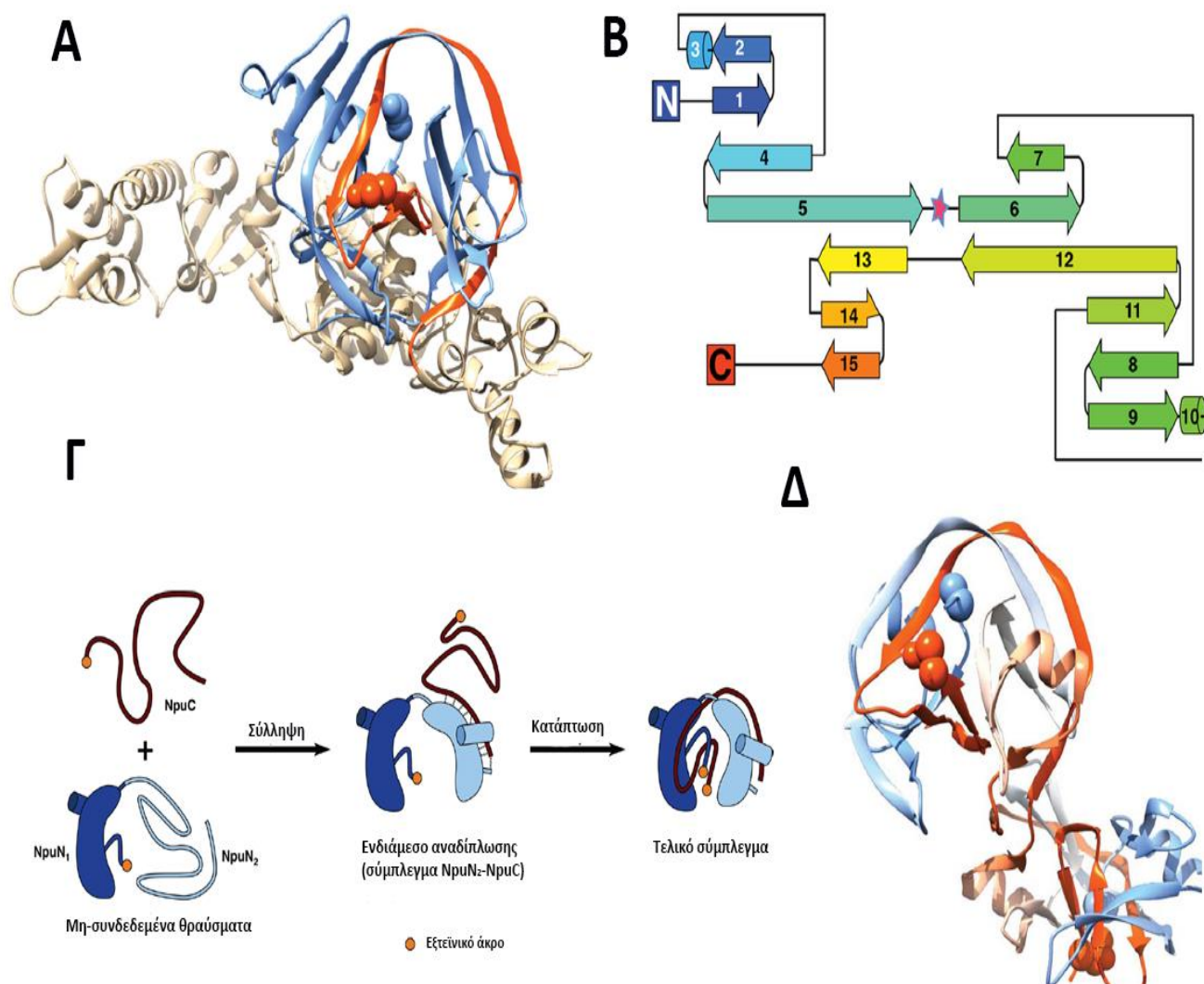
Αυτό το στάδιο διεξάγεται με την N-τελική κυστεΐνη ή σερίνη της ιντεΐνης, που προσβάλλει το C-τελικό κατάλοιπο της εξτεΐνης στον άνθρακα του καρβονυλίου της. Η προσβολή αυτή διευκολύνεται από το την θρεονίνη και ιστιδίνη της περιοχής B (TXXH) και το ασπαρτικό της περιοχής F (Εικ. 1.11). Το ασπαρτικό της περιοχής F χρησιμεύει ως δότης για την δημιουργία δεσμού υδρογόνου ενώ συγχρόνως σταθεροποιεί τη C1 θειολική ένωση, όπως και είναι ένα αυστηρό προαπαιτούμενο για την αντίδραση της N σε O / S άκυλο-μετατόπιση (Van Roey et al., 2007). Η ιστιδίνη της περιοχής B απουσιάζει μόνο στη TKO CDC21-1 ιντεΐνη του *Thermococcus kodakaraensis* intein. Η TKO CDC21-1 παρακάμπτει την έλλειψη της ιστιδίνης χρησιμοποιώντας έναν εναλλακτικό μηχανισμό που διευκολύνεται από Lys58, το οποίο είναι ένα από τα συντηρημένα μοτίβα (Tori et al., 2012). Αυτό το πρώτο βήμα του ματίσματος, της N σε S / O άκυλο-μετατόπισης, απαιτεί μια σημαντική πρόκληση για τη διάσπαση του πεπτιδικού δεσμού που σε γενικές γραμμές αποτελεί ένα μεγάλο ενεργειακό φράγμα. Είναι πιθανό ότι η ενέργεια που αποκτάται από την αναδίπλωση σε μια πολύ σταθερή διαμόρφωση να χρησιμοποιείται άμεσα για την υπέρβαση του ενεργειακού φράγματος για το πρώτο βήμα του ματίσματος. Ένας τέτοιος μηχανισμός παρατηρείται στην πρωτεϊνικές περιοχές του σπέρματος SEA (Sperm protein in Enterokinase and Agrin) της ανθρώπινης MUC 1, όπου η αυτοπρωτεόλυση καταλύεται από διαμορφωτικό στρες (Macao et al., 2006). Στη δομή της Ssp DnaE ιντεΐνης, οι εξτεΐνες είναι σφικτά πακεταρισμένες (Sun et al., 2005), ώστε μετά τη συναρμολόγηση της ιντεΐνης να μπορέσουν να δημιουργήσουν μια στερεοχημική σύγκρουση στο ενεργό κέντρο, έτσι ώστε η

αντίδραση ματίσματος να είναι ο πιο κινητικά προσβάσιμος μηχανισμός για τη διαφυγή από αυτή τη σύγκρουση.

Επίσης, από τη δομή της Ssp DnaE ιντεΐνης που πιάστηκε σε παγίδα οξειδοαναγωγής (Callahan et al., 2011), μαζί με συμπληρωματικές υπολογιστικές και κρυσταλλογραφικές αποδείξεις (Dearden et al., 2013), φαίνεται ότι οι αρχικές αναδιατάξεις των κατάλοιπων και το πρώτο βήμα της N σε O / S ακυλο-μετατόπισης, επιταχύνθηκε από εντοπισμένο διαρθρωτικό στέλεχος στο -1 κατάλοιπο της N-εξεΐνης, που δημιουργήσε η συντηρημένη θρεονίνη της περιοχής B, το οποίο καθιστά τις αλληλεπιδράσεις του δεσμού με C1 και του -2 κατάλοιπου N-εξεΐνης. Επιπλέον, ορισμένες ιντεΐνες στερούνται ένα νουκλεόφιλο στην περιοχή A, δηλαδή αντί αυτού έχουν μια αλανίνη ή μια προλίνη. Αρκετές μελέτες έχουν δείξει ένα στριμμένο ή αποσταθεροποιημένο σχάσιμο πεπτιδικό δεσμό στο N-τελικό άκρο της πρόδρομης πρωτεΐνης (Dearden et al., 2013). Αυτές οι ιντεΐνες άμεσα σχηματίζουν ένα τυπικό διακλαδισμένο ενδιάμεσο μετά την ενεργοποίηση του N-τελικού (Southworth et al., 2000), ή σε ορισμένες περιπτώσεις, χρησιμοποιούν μια κυστεΐνη της περιοχής F για να σχηματίσουν ένα διαφορετικό διακλαδισμένο ενδιάμεσο θειοεστέρα προτού το τυπικό διακλαδισμένο ενδιάμεσο (Tori et al., 2010). Η Mja KlbA ιντεΐνη, με μια αλανίνη στο N-τελικό άκρο, έχει ένα πεπτιδικό δεσμό σε cis-διαμόρφωση πιθανώς καθιστώντας τον ασταθή ή αυξάνοντας την ευαισθησία του στη πυρηνόφιλη προσβολή (Johnson et al., 2007).

trans-Θειοεστεροποίηση

Αυτό το βήμα είναι το λιγότερο κατανοητό βήμα στην πρωτεϊνικό μάτισμα, διότι είναι το πιο δύσκολο βήμα για να ερευνηθεί κανείς πειραματικά (το γραμμικό ενδιάμεσο θειοεστέρα δεν μπορεί να απομονωθεί). Η trans-θειοεστεροποίηση απαιτεί αποπρωτονίωση της +1 Cys /Ser / Thr στην πλευρική αλυσίδα και πυρηνόφιλη προσβολή στο δεσμο του γραμμικού θειοεστέρα στο Cys / Ser1, τα οποία είναι αρκετά μακριά ακόμη και σε δομές υψηλής ανάλυσης (9 -10 Å) (Ding et al., 2003; Sun et al., 2005)(Εικ. 1.11). Όλες όμως οι μελέτες υποδεικνύουν ότι τα δύο πρώτα στάδια του πρωτεϊνικού ματίσματος επιτυγχάνονται με ένα πολύπλοκο δίκτυο δεσμών υδρογόνου από τρία συντηρημένα κατάλοιπα, N-τελική κυστεΐνη / σερίνη, ιστιδίνη περιοχής B, και ασπαρτικό περιοχής F (Pereira et al., 2011).



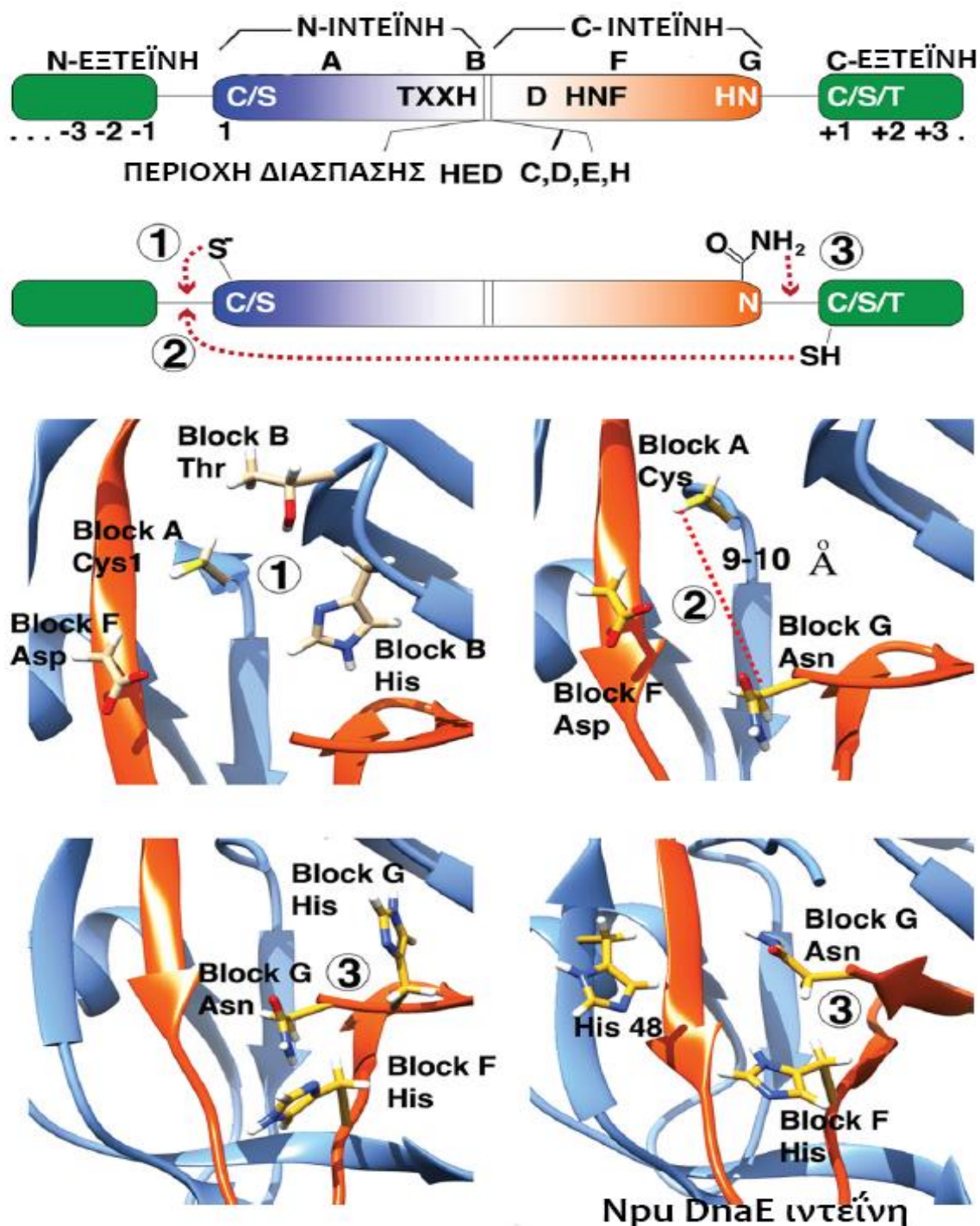
Εικόνα 1.10.

Αναδίπλωση ιντεΐνης. **A)** Pol-2 ιντεΐνη του *T. kodakaraensis* (Protein Data Bank (PDB) 2CW7) (Matsumura et al., 2006). Φαίνονται οι N-ιντεΐνη (μπλε) και C-ιντεΐνη (κόκκινη). **B)** Τοπολογικός χάρτης της δευτεροταγής δομής της πρωτεϊνικής περιοχής Hog της *Drosophila melanogaster*, που είναι συγγενής των ιντεϊνών (PDB 1AT0) (Hall et al., 1997). **Γ)** Ο μηχανισμός «σύλληψης και κατάπτωσης» της Npu DnaE ιντεΐνης και **Δ)** η τρισδιάστατη της δομή (Shah et al., 2013a).

Διακλαδισμένο ενδιάμεσο

Το διακλαδισμένο ενδιάμεσο είναι μη αντιστρεπτό και επιτυγχάνει την διάσπαση του πεπτιδικού δεσμού μεταξύ της ιντεΐνης και της C-εξτεΐνης. Το άζωτο της πλευρικής αλυσίδας της C-τελικής ασπαραγίνης εκτελεί μια πυρηνόφιλη προσβολή στον άνθρακα του καρβονυλίου του πεπτιδικού δεσμού για να σχηματίσουν ένα σουκιναμίδιο με αποτέλεσμα τη διάσπαση της ιντεΐνης από το διακλαδισμένο ενδιάμεσο (Xu et al., 1994). Πράγματι, η πυρηνόφιλη πλευρική αλυσίδα της Asn είναι βιοχημικά σπάνια (βρίσκεται επίσης στη N-γλυκοζυλίωση) (Schwarz and Aebi, 2011) και ο σχηματισμός του σουκιναμιδίου σε πρωτεΐνες είναι σύνηθες να συμβαίνει με πυρηνόφιλη προσβολή από ένα άζωτο αμιδίου του σκελετού στο καρβονυλικό άνθρακα της πλευρικής αλυσίδας της Asn, οδηγώντας σε απαμίδωση της πλευρικής αλυσίδας (Stephenson and Clarke, 1989). Τόσο οι δομικές όσο και οι υπολογιστικές μελέτες αποκάλυψαν ότι οι ιστιδίνες των περιοχών F και G (προτελευταία) είναι σημαντικές στο σχηματισμό του διακλαδισμένου ενδιάμεσου, ενεργοποιώντας την ασπαραγίνη και πρωτονιώνοντας το σχηματισμό αμίνης (Frutos et al., 2010) (Εικ. 1.11.). Είναι ενδιαφέρον ότι ορισμένες ιντεΐνες στερούνται την προτελευταία ιστιδίνη (Chen et al., 2000). Στην DnaE ιντεΐνη, αυτή η ιστιδίνη είναι είτε σερίνη ή μια αλανίνη (Caspi et al., 2003b). Δεν είναι σαφές ακόμη πώς αυτές οι ιντεΐνες μπορεί να παρακάμψουν την έλλειψη των δύο ιστιδινών και να διεξαγάγουν την αντίδραση ματίσματος. Ωστόσο, υπάρχουν και άλλα κοντινά κατάλοιπα τα οποία μπορούν να χρησιμεύσουν ως γενικά οξέα ή βάσεις κατά τη διάρκεια του ματίσματος. Είναι ενδιαφέρον ότι, αν και η His48 στη DnaE ιντεΐνη εμπλέκεται στο μάτισμα, είναι κοντά στο ενεργό κέντρο και μπορεί να υποκαταστήσει τη His περιοχής G.

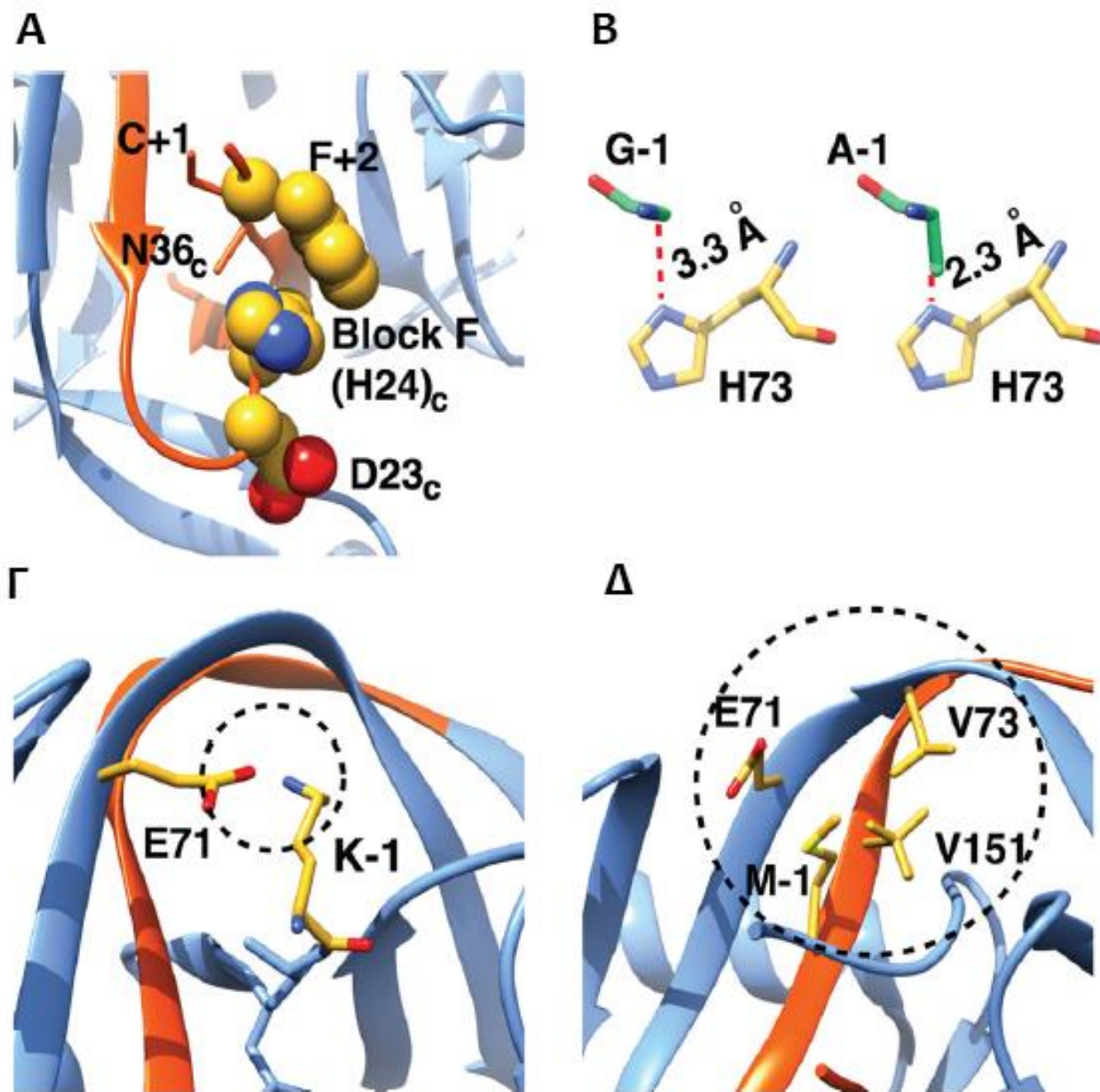
Αν και δεν είναι εμφανής σε κρυσταλλικές δομές, υποθέτετε ότι κάθε βήμα του ματίσματος προκαλεί τοπικές διαμορφωτικές μεταβολές, επηρεάζοντας το συντονισμό της αντίδρασης και μετατοπίζοντας τη θέση ισορροπίας. Συνεπώς με αυτή την έννοια, είναι ο σχηματισμός του διακλαδισμένου ενδιάμεσου τόσο στη MXE GyrA (Frutos et al., 2010) όσο και στη NPU DnaE που ενισχύει (Shah et al., 2013b) δραματικά τη κινητική κυκλοποίησης της Asn, υποδεικνύοντας ότι η διακλαδισμένη αρχιτεκτονική αλλάζει τη δομή και τη δυναμική του ενεργού κέντρου. Πράγματι, στοιχεία από NMR για την πρόδρομη GyrA ιντεΐνη και ένα συνθετικό παγιδευμένο διακλαδισμένο ενδιάμεσο δείχνουν ότι το χημικό περιβάλλον και η δυναμική χαλάρωση γύρω από τη σχάσιμο πεπτιδικό δεσμό αλλάζουν κατά τον σχηματισμό διακλαδώσεων.



Εικόνα 1.11.

Δομική απεικόνιση του μηχανισμού πρωτεϊνικού ματίσματος, καθώς και των μοτίβων που παίζουν ρόλο στο μάτισμα. Η δομή στα βήματα 1-3 προέρχεται από τη Mtu RecA ιντεΐνη PDB 2. Ένας Εξάρτηση του πρωτεϊνικού ματίσματος από την εξετίνική αλληλουχία εναλλακτικός μηχανισμός δείχνεται στο βήμα 3 δεξιά κάτω, από τη Npu DnaE ιντεΐνη στην οποία ελείπεται η προτελευταία His (Eryilmaz et al., 2014).

Είναι ενδιαφέρον το γεγονός ότι εκατοντάδες διαφορετικές ιντεΐνες μπορούν να πραγματοποιήσουν τον ίδιο χημικό μηχανισμό, αλλά με διαφορετική ειδικότητα σε εξτεΐνη. Διαφορετικές ιντεΐνες έχουν εξελιχθεί για να συνδέουν διαφορετικές πρωτεΐνες στη φύση και ως εκ τούτου οι προτιμήσεις των εξτεϊνών είναι σε συνάφεια με τα πλαίσια της φύσης. Απόκλιση από αυτές τις προτιμήσεις συχνά οδηγεί σε μια βαθιά επίδραση της κινητικής και της απόδοσης του μηχανισμού του ματίσματος, γεγονός που υποδηλώνει ότι οι εξτεΐνες μπορούν να συμβάλουν είτε χημικά ή δομικά στο ενεργό κέντρο της ιντεΐνης. Για παράδειγμα, στη NPU DnaE ιντεΐνη, το μάτισμα επηρεάζεται ελάχιστα από τις τοπική αλληλουχία N-εξτεΐνης, αλλά η διακύμανση της τοπικής αλληλουχίας C-εξτεΐνης επηρεάζει δραματικά την αποδοτικότητα του ματίσματος (Cheriyani et al., 2013; Shah et al., 2012) διαταράσσοντας το σχηματισμό του διακλαδισμένου ενδιάμεσου (Lockless and Muir, 2009). Στη δομή της ορθόλογης Ssp DnaE ιντεΐνης, η πλευρική αλυσίδα της ενδογενούς Phe +2 πακετάρεται κατά της σταθεροποίησης της His της περιοχής F (Εικ. 1.12). Αυτή η συντηρημένη His βρίσκεται επάνω σε μια εύκαμπτη λούπα και έχει χαρακτηριστεί ως γενική βάση ή οξύ στο μηχανισμό σχηματισμού του διακλαδισμένου ενδιάμεσου σε πολλές ιντεΐνες (Frutos et al., 2010; Sun et al., 2005). Η εύκαμπτη λούπα που περιέχει την ιστιδίνη στην περιοχή F διαμορφώνεται με βάση την αλληλουχία της C-εξτεΐνης και έχει χρησιμοποιηθεί για να κατασκευάσει μια πιο ανεκτική σε εξτεΐνες NPU ιντεΐνη. Η μετάλλαξη, D124Y, σε αυτό τη λούπα σταθεροποιεί και προσανατολίζει τη His125 σε μια καταλυτικά ευνοϊκή στάση, μειώνοντας έτσι την εξάρτηση από Phe +2 της εξτεΐνης στο NPU (36). Σε μια υπολογιστική μελέτη για την Mtu RecA ιντεΐνη, όπου χρησιμοποιήθηκαν μέθοδοι μοριακής μηχανικής και κβαντομηχανικής και η ηλεκτρονική δομή της τοπικής αλληλουχίας +1 μεταλλάξεις των διαφόρων κατάλοιπων που έχουν διερευνηθεί, διαπιστώθηκε ότι οι μεταλλάξεις έχουν διαφορετική συγγένεια ηλεκτρονίων και δυναμικά ιονισμού επιβάλλοντας διάφορα φράγματα ενέργειας (Shemella et al., 2011). Στη Ssp DnaB ιντεΐνη, η ανάλυση της κρυσταλλικής δομής εξηγεί γιατί η θέση -1 μπορεί να καταληφθεί μόνο από Gly. Η μοντελοποίηση της δομής προτείνει ότι οποιαδήποτε μεγαλύτερο κατάλοιπο θα καταργήσει το μάτισμα προκαλώντας στερεοχημικούς συγκρούσεις (Schwarzer et al., 2012) (Εικ. 1.12B). Επιπρόσθετα, στην Pho RadA ιντεΐνη, δείχθηκε ότι το μέγεθος και το φορτίο της πλευρικής αλυσίδας έχει αντίκτυπο στη θέση -1, λόγω ηλεκτροστατικών αλληλεπιδράσεων και επιδράσεις πακεταρίσματος στο ενεργό κέντρο. Αρνητικά φορτισμένα αμινοξέα, β-διακλαδισμένα αμινοξέα, η προλίνη, και αμινοξέα με μικρές πλευρικές αλυσίδες έχουν χαμηλότερη αποδοτικότητα ματίσματος (Oeemig et al., 2012) (Εικ. 1.12 Γ και Δ).



Εικόνα 1.12.

Επίδραση της εξτείνης στην **A)** Ssp DnaE ιντεΐνη **B)** Ssp DnaB ιντεΐνη και **Γ)** , **Δ)** Pho RadA ιντεΐνη

Μέχρι στιγμής, οι παραδοσιακές δομικές προσεγγίσεις βιολογίας και βιοχημείας έχουν διευκρινίσει πολλές λεπτομέρειες για τη δομική βάση του πρωτεϊνικού ματίσματος. Ωστόσο, μια λεπτομερή κατανόηση του πώς συντονίζει η αναδίπλωση των ιντεϊνών τη χημεία των πολλαπλών ενεργών κέντρων εξακολουθεί να είναι ανεπαρκής, όπως αντίστοιχα είναι και η κατανόηση των γενικών κανόνων για τις προτιμήσεις των εξτεϊνών. Το γεγονός ότι η ιντεϊνική δραστηριότητα μπορεί να διαμορφώνεται από μεταλλάξεις μακριά από το ενεργό κέντρο υποδηλώνει ότι τα αλλοστερικά δίκτυα μπορεί να παίζουν έναν μεγαλύτερο ρόλο στον καθορισμό της δραστικότητας των ιντεϊνών. Τελικά, μια καλύτερη κατανόηση της δομής και της δυναμικής των ιντεϊνών είναι ζωτικής σημασίας, καθώς με τη χρήση αυτών των πληροφοριών υπάρχει η δυνατότητα κατασκευής ιντεϊνών ως καλύτερα εργαλεία για την χρήση στη βιοτεχνολογική παραγωγή. Μερικά από τα αντιδρώντα ενδιάμεσα στο μηχανισμό του ματίσματος δεν είναι πειραματικά άμεσα παρατηρήσιμα με τη χρήση παραδοσιακών πειραματικών τεχνικών. Ωστόσο, με τον συνδυασμό συνθετικής χημείας, κατευθυνόμενης εξέλιξης και υπολογιστικής μοντελοποίησης είναι πιθανό να παρέχουν γνώσεις που θα δώσουν μια καλύτερη ιδέα για το μηχανισμό του πρωτεϊνικού ματίσματος

1.4 Εφαρμογές των ιντεϊνών στη Βιοτεχνολογία

Η ανακάλυψη των ιντεϊνών στις αρχές της δεκαετίας του 90 ήταν έναυσμα για τη δημιουργία ποικίλων νέων τεχνολογιών. Οι πρώτες ανασυνδιασμένες ιντεΐνες που προέρχονταν από διάφορους οργανισμούς, επέτρεψαν την ανάπτυξη μεθόδων όπως απομόνωση με συγγένεια με τη διάσπαση ετικέτας συγγένειας και σύνδεση τμημάτων πρωτεϊνών μέσω πρωτεϊνικού ματίσματος. Ορισμένες εφαρμογές έχουν αναπτυχθεί βάση φυσικών και ανασυνδιασμένων διχασμένων ιντεϊνών, που edώσαν τη δυνατότητα στη ξεχωριστή έκφραση τμημάτων μια πρωτεΐνης και την ένωση τους *in vitro*. Αυτές οι ιντεΐνες έχουν οδηγήσει ένα μεγάλο αριθμό εφαρμογών στη μεταβολική μηχανική, στη σήμανση πρωτεϊνών, στη κατασκευή βιοϋλικών, στη κυκλοποίηση πρωτεϊνών και στο καθαρισμό πρωτεϊνών. Η ικανότητα των ιντεϊνών να σχηματίζουν και να διασπούν ειδικούς πεπτιδικούς δεσμούς σε μια ποικιλία από περιβάλλοντα επέτρεψε την ανάπτυξη ισχυρών νέων εργαλείων στη μοριακή βιολογία.

1.4.1 Καθαρισμός πρωτεϊνών

Μία από τις πιο σημαντικές εφαρμογές των ιντεϊνών ήταν η ανάπτυξη ενός συστήματος αυτο-διάσπασης της ετικέτας συγγένειας, για την ανάκτηση της πρωτεΐνης χωρίς το κομμάτι αυτό κατά

την διάρκεια του καθαρισμού της, μετά από την έκφραση σε ανασυνδυασμένα συστήματα (Chong et al., 1997, 1998b). Σε αυτές τις εφαρμογές, μια τροποποιημένη ιντεΐνη εκφράζεται σε σύντηξη με μία ετικέτα συγγένειας και τη πρωτεΐνη στόχο, ώστε μετά των καθαρισμό να επαχθεί από την ίδια την ιντεΐνη, η διάσπαση της πρωτεΐνης στόχο από την ιντεΐνη και την ετικέτα. Το πρώτο εμπορικό σύστημα κυκλοφόρησε από την New England Biolabs (NEB) το 1997, που χρησιμοποιεί μια τροποποιημένη Sce VMA1 ιντεΐνη που ενεργοποιεί τη διάσπαση στο N-τελικό άκρο (σύστημα IMPACT) της ή και τα δύο N-και C-άκρα (IMPACT-CN σύστημα) με την προσθήκη ενώσεων θειόλης (Wood and Camarero, 2014). Λίγο αργότερα δημοσιεύθηκαν ιντεΐνες με ταχεία C-τελική δραστηριότητα διάσπασης. Αυτές οι ιντεΐνες εμφανίζουν καταστολή στη δραστηριότητα διάσπασης σε pH 8.5, που επιτρέπει το καθαρισμό του επισημασμένου στόχου, αφού στη συνέχεια διεγείρονται για να γίνει η διάσπαση με μία μετατόπιση σε pH 6.5. Η δραστηριότητα της διάσπασης αυτών επηρεάζεται επίσης έντονα από τη θερμοκρασία, όπου η πιο ταχεία διάσπαση παρατηρείται στους 37 ° C (Mathys et al., 1999). Η ανάπτυξη αυτών των ιντεϊνών οδήγησε σε μια ποικιλία νέων συστημάτων ετικετών και διαμορφώσεις για το καθαρισμό πρωτεϊνών, αρχικά με την χρήση των περιοχών πρόσδεσης της χιτίνης (CBD) (Chong et al., 1997) και της μαλτόζης (MBD) (Wood et al., 2000). Πρόσφατα, έχουν χρησιμοποιηθεί οι SUMO ετικέτες και ετικέτες ουβικιτίνης για την αύξηση της αποτελεσματικής έκφρασης και τον απλοποιημένο καθαρισμό (Wang et al., 2012; Wood et al., 2000).

Έλεγχος διάσπασης έναντι της επανασυναρμολόγησης των θραύσματος - Μια βασική απαίτηση για όλα τα συστήματα καθαρισμού που χρησιμοποιούν ιντεΐνες είναι να έχουν την ικανότητα της ελάχιστης διάσπασης κατά την διάρκεια της έκφρασης της πρωτεΐνης, αλλά να ενεργοποιούνται γρήγορα κατά το καθαρισμό. Πολλά παλιά συστήματα απαιτούν μεγάλους χρόνους επώασης για πλήρης C-τελική διάσπαση, ή υψηλές συγκεντρώσεις της θειολικών ενώσεων για την αποτελεσματική N-τελική διάσπαση. Είναι όμως γνωστό ότι οι ιντεΐνες με C-διάσπαση μαστίζονται από πρόωρη διάσπαση κατά τη διάρκεια της έκφρασης, ενώ υπάρχει δυσκολία χρήσης των ιντεϊνών N-διάσπασης με πρωτεΐνες-στόχους που περιέχουν δισουλφιδικούς δεσμούς. Για το λόγο αυτό, πρόσφατες μελέτες έχουν επικεντρωθεί στον έλεγχο της λειτουργίας μέσω της επανασυναρμολόγησης των ιντεϊνών trans-διάσπασης, καθώς και την εισαγωγή δισουλφιδικών δεσμών σε υπάρχοντα συστήματα ιντεϊνών (Callahan et al., 2013). Παρά το γεγονός ότι αυτή η στρατηγική ήταν αποτελεσματική στην καταστολή της πρόωρης διάσπασης κατά τη διάρκεια της έκφρασης, η επανασυναρμολόγηση των κομματιών ήταν συχνά αναποτελεσματική (Shi et al.,

2013). Η ανακάλυψη των φυσικώς διχασμένων ιντεϊνών είχε σημαντικό αντίκτυπο στον τομέα αυτό, αφού τα τμήματα αυτών είναι ικανά για αποτελεσματική έκφραση και επανασυναρμολόγηση τόσο *in vivo* όσο και *in vitro* (Evans et al., 2000; Martin et al., 2001). Πρόσφατα έχουν ανέφερθεί παραδείγματα συστημάτων με διάσπαση *in trans* που βασίζονται στη γενετικά ανασυνδυασμένη Ssp DnaB ιντεΐνη και τη φυσική NPU DnaE που προκαλεί trans-μάτισμα. Η Ssp DnaB έχει κατασκευαστεί για να καταστείλει την πρόωρη διάσπαση μέσω της διαγραφής 11 αμινοξέων στο N-τελικό άκρο, διαγραφή που καταργεί τη δραστηριότητα διάσπασης στην υπόλοιπη ιντεΐνη επιτρέποντας τον καθαρισμό μιας συντηγμένης πρωτεΐνης-στόχο στο C-τελικό άκρο. Η διάσπαση διεγείρεται με την προσθήκη του πεπτιδικού καταλοίπου 11. Από την άλλη το σύστημα καθαρισμού που βασίζεται στη NPU DnaE, η οποία παρατηρείται από φυσικά εξαιρετικά γρήγορο trans-μάτισμα, έχει κατασκευαστεί για να εμφανίζει πιο γρήγορη διάσπαση στο C-τελικό άκρο κατά την επανασυναρμολόγηση. Σε αυτό το σύστημα, το N-τελικό τμήμα της ιντεΐνης συντήκεται σε μία ετικέτα συγγένειας και ακινητοποιείται, ενώ το C-τελικό τμήμα συντήκεται με την πρωτεΐνη-στόχο. Αξίζει να σημειωθεί ότι, αυτές οι ιντεΐνες μπορούν να υποβάλουν πλήρης διάσπαση σε λιγότερο από 30 λεπτά σε θερμοκρασία δωματίου (Ramirez et al., 2013), και σε λίγες μόνο ώρες στους 6 ° C.

1.4.2 Τροποποίηση πρωτεϊνών

Οι ιντεΐνες χρησιμοποιούνται επίσης για την τροποποίηση πρωτεϊνών μέσω της αναδιάταξης των πεπτιδικών δεσμών. Οι εφαρμογές περιλαμβάνουν την παραγωγή των πρωτεϊνών με τις διαρθρωτικές αλλαγές, συμπεριλαμβανομένων σε αυτές τη κυκλοποίηση πεπτιδίων, τη τόπο-ειδική επισήμανση και την πρωτεόλυση. Σε κάθε μία από αυτές τις περιπτώσεις, μια μεταλλαγμένη ιντεΐνη μετακινείται σε ένα μη-φυσικό πλαίσιο, ώστε η τροποποιημένη αντίδραση ματίσματος που προκύπτει να χρησιμοποιηθεί για να παράγει ομοιοπολικές τροποποιήσεις στην πρωτεΐνη-στόχο.

Κυκλοποίηση πεπτιδίων. Η κυκλοποίηση ανασυνδυασμένων πολυπεπτιδίων μπορεί να επιτευχθεί *in vitro* ή *in vivo* είτε με έκφραση πρωτεϊνικής πρόσδεσης (expressed protein ligation EPL) είτε με πρωτεϊνικό trans-μάτισμα (PTS) (Aboye and Camarero, 2012). Η κυκλοποίηση στη EPL επιτυγχάνεται με τη σύντηξη του πολυπεπτιδίου στόχου στο N-τελικό άκρο με μια αλληλουχία οδηγού που τελειώνει με ένα κατάλοιπο Cys, ενώ το C-τελικό άκρο είναι συγχωνευμένο με μια ανασυνδυασμένη ιντεΐνη. Η αλληλουχία N-τερματικού μπορεί να διασπασθεί *in vivo* ή *in vitro* με ένα πρωτεολυτικό ή αυτο-πρωτεολυτικό συμβάν, αφήνοντας ένα N-τελικό κατάλοιπο Cys στο

πεπτίδιο στόχο. Το κατάλοιπο Cys μπορεί στη συνέχεια να αντιδράσει ενδομοριακά με ένα α-θειοεστέρα που παράγεται από την ιντεΐνη, παρέχοντας έτσι μια κυκλοποιημένο πεπτίδιο. Η χρήση της έκφραση πρωτεϊνικής πρόσδεσης για κυκλοποίηση αναφέρθηκε πρώτα από Camarero και το Muir, το 1999, χρησιμοποιώντας τη N-τελική SH3 περιοχή της πρωτεΐνης c-Crk ως πρότυπο σύστημα (Muir and Camarero). Η προκύπτουσα κυκλική περιοχή πρωτεΐνης διπλώθηκε ταχύτερα και ήταν πιο σταθερή από το γραμμικό ομόλογό (Camarero et al., 2001).

Μία άλλη προσέγγιση για την ανασυνδυασμένη παραγωγή κυκλικών πεπτιδίων *in vitro*, αναφέρθηκε πρώτα από το Benkovic χρησιμοποιώντας τη φυσική διχασμένη Ssp DnaE ιντεΐνη, για τη δημιουργία του PTS (Scott et al., 1999). Το πρωτεϊνικό trans-μάτισμα βασίζεται σε μια περιοχή ιντεϊνικού ματίσματος που χωρίζεται σε δύο τμήματα (I_N και I_C). Τα θραύσματα της διχασμένης ιντεΐνης δεν δραστηριοποιούνται μεμονωμένα, αλλά μπορεί να συνδεθούν με το άλλο με υψηλή εξειδίκευση υπό κατάλληλες συνθήκες για να σχηματίσει ένα ενεργό κέντρο ματίσματος *in trans*. Η αντιστροφή τη σειράς των θραυσμάτων και η σύντηξη με το πεπτίδιο που πρόκειται να κυκλοποιηθεί, έχει ως αποτέλεσμα το trans-μάτισμα για σχηματισμό ενός κυκλοποιημένου πεπτιδίου στη ραχοκοκαλιά του. Το πρωτεϊνικό trans-μάτισμα έχει χρησιμοποιηθεί για την παραγωγή αρκετών φυσικώς απαντώμενων κυκλικών πεπτιδίων, καθώς επίσης και για την κατασκευή μεγάλων βιβλιοθηκών μικρών κυκλικών ολιγοπεπτιδίων (Tavassoli and Benkovic, 2007). Μολονότι το πρωτεϊνικό trans-μάτισμα προσφέρει μια καλή εναλλακτική λύση έναντι της έκφρασης πρωτεϊνικής πρόσδεσης για την γρήγορη παραγωγή κυκλοποιημένων πεπτιδίων, θα πρέπει να σημειωθεί ότι τα περισσότερα τμήματα απαιτούν ειδικά κατάλοιπα αμινοξέων στις ενώσεις ιντεΐνης-εξεΐνης για ένα πιο αποδοτικό μάτισμα (Tavassoli and Benkovic, 2007).

Τοπο-ειδική επισήμανση πρωτεϊνών – Οι μέθοδοι που αναφερθήκαν (EPL και PTS) μπορούν επίσης να χρησιμοποιηθούν για την εξειδικευμένη εισαγωγή τροποποιήσεων σε πρωτεΐνες, όπως είναι η γλυκοζυλίωση, βιοτινυλίωση, ουβικιτινοποίησης, φωσφορυλίωση και τμηματική ισοτοπική επισήμανση (Berrade and Camarero, 2009; Volkmann and Iwai, 2010). Αυτές οι μέθοδοι μπορούν επίσης να εφαρμοστούν τη N-και C-τελική επισήμανση σε πρωτεΐνες σχετιζόμενες με τη μεμβρανική περιοχή. Ένα παράδειγμα αυτής της διεργασίας αποτελεί η επισήμανση του C-τελικού άκρου του ανθρώπινου υποδοχέα τρανσφερίνης, που σχετίζεται με τη μεμβράνη, με φλουορεσκεΐνη και βιοτίνη στην επιφάνεια ζωντανών θηλαστικών κυττάρων χρησιμοποιώντας τη διχασμένη SSP GyrB ιντεΐνη (Volkmann and Liu, 2009). Χρησιμοποιώντας μια παρόμοια

προσέγγιση, το N-άκρο του μονομερούς της κόκκινης φθορίζουσας πρωτεΐνης (mRFP) επισημάνθηκε με βιοτίνη στην επιφανείας κυττάρων CHO (Ando et al., 2007). Το πρωτεϊνικό trans-μάτισμα έχει επίσης χρησιμοποιηθεί για τοπο-ειδική επισήμανση των πρωτεϊνών μέσα σε ζωντανά κύτταρα. Τα τμήματα που χρησιμοποιούνται για την εκπλήρωση αυτού του καθήκοντος συνήθως σχηματίζεται είτε από μικρά θραύσματα I_N ή I_C , διευκολύνοντας έτσι την χημική σύνθεση με χημικούς ανιχνευτές. Η ιντεϊνο-κυτταρική σήμανση πρωτεΐνης έχει πολλά πλεονεκτήματα έναντι άλλων χημειοεπιλεκτών τεχνικών. Ειδικότερα, η επισήμανση με πρωτεϊνικό trans-μάτισμα παρέχει μοναδική τοποειδική ομοιοπολική τροποποίηση της πρωτεΐνης-στόχου και πιο σημαντικά η αντίδραση ματίσματος στηρίζεται σε ένα συγκεκριμένο γεγονός, στην αναγνώριση μεταξύ των αντίστοιχων I_N και I_C θραυσμάτων. Δηλαδή με τη χρήση διχασμένων ιντεϊνών που δεν αλληλοεπιδρούν μεταξύ τους, μπορεί να οδηγήσει στην ποικιλόμορφη σήμανση πρωτεϊνών εντός των κυττάρων (Borra et al., 2012).

Τόπο-ειδική πρωτεόλυση – Οι διχασμένες ιντεΐνες μπορούν επίσης να παρέχουν ένα μέσο για την αποτελεσματική τόπο-ειδική πρωτεόλυση σε ζωντανά κύτταρα. Για παράδειγμα, η μη κανονική Ssp DnaB S1 ιντεΐνη έχει πρόσφατα χρησιμοποιηθεί ως μια αποτελεσματική και ιδιαίτερα τόπο-ειδική κυτταρική πρωτεάση (Volkmann et al., 2012). Στην μελέτη αυτή, η αλληλουχία αναγνώρισης της πρωτεάσης σχηματίζεται από το 11^ο κατάλοιπο στο I_N θραύσμα και πέντε φυσικά κατάλοιπα στη N-εξτεΐνη. Αυτή η αλληλουχία αναγνωρίζεται από το αντίστοιχο θραύσμα I_C , που αποκαλείται ιντεϊνο-προερχόμενη πρωτεάση (Intein-derived Protease IP). Όταν τα δύο θραύσματα συνεκφράζονται, ενώνονται μεταξύ τους αναστήνοντας ένα πλήρως ενεργό τομέα ματίσματος που διασπά την πρωτεΐνη του ενδιαφέροντος μέσω N-τελικής διάσπασης. Αυτή η προσέγγιση, η οποία είναι ανάλογη με την αντίδραση διάσπασης N-τελικής ετικέτας, έχειδειχθεί να λειτουργεί αποτελεσματικά σε βακτηριακά και ευκαρυωτικά κύτταρα. Σε αντίθεση με άλλες λιγότερο ειδικές πρωτεάσες, IP πρωτεάσες που είναι πολύ ειδικές, παρουσιάζοντας εξαιρετικά χαμηλή δραστηριότητα προς κυτταρικές πρωτεΐνες που δεν περιέχουν την αλληλουχία αναγνώρισης I_N .

1.4.3 Ιντεϊνο-βασιζόμενοι βιοαισθητήρες

Οι ιντεϊνο-βασιζόμενοι βιοαισθητήρες συνήθως κατασκευάζονται από ανεξάρτητες πρωτεϊνικές περιοχές που μπορεί να εκφράζονται μέσα σε ζωντανά κύτταρα. Το σήμα αναγνωρίζεται από μια μονάδα ανίχνευσης, η οποία στη συνέχεια προκαλεί μια αλλαγή στην δραστηριότητα του

ματίσματος μιας συντηγμένης μονάδας ιντεΐνης, που έχει ως αποτέλεσμα την αλλαγή στη δραστηριότητα μιας μονάδας ανταποκριτή. Οι ιντεΐνο-βασιζόμενοι βιοαισθητήρες πρωτεϊνικό μάτισμα υπό προϋποθέσεις ενεργοποιούνται από ένα εξωτερικό ερέθισμα, όπως η σύνδεση ενός μικρού μορίου, το φως, μια αλλαγή στη θερμοκρασία, το pH ή η οξειδοαναγωγική κατάσταση (πρωτεϊνικό μάτισμα υπό προϋποθέσεις, CPS) (Mootz, 2009). Η αντίδραση ματίσματος αποκαθιστά την λειτουργία ενός τομέα ανταποκριτή, το οποίο στη συνέχεια παράγει ένα εύκολα ανιχνεύσιμο σήμα. Η ιντεϊνική περιοχή ενεργοποιεί μια ρυθμιστική δομή αυτών των αισθητήρων, που επιτρέπει την εύκολη ανταλλαγή περιοχών ανίχνευσης και ανταπόκρισης. Αυτό το χαρακτηριστικό έχει διευκολύνει το σχεδιασμό των αισθητήρων για την ανίχνευση αλληλεπιδράσεων μεταξύ πρωτεϊνών, μικρών μορίων, οξειδοαναγωγικών καταστάσεων, δραστηριότητας πρωτεασών, μεθυλίωσης του DNA και άλλων τύπων ερεθισμάτων.

Ανίχνευση αλληλεπιδράσεων μεταξύ πρωτεϊνών – Γίνεται χρήση βιοαισθητήρων για την ανίχνευση αλληλεπιδράσεων μεταξύ πρωτεϊνών με τεχνητές ή ανασυνδυασμένες διχασμένες ιντεΐνες με μειωμένη συγγένεια μεταξύ των I_N και I_C θραυσμάτων. Οι πρωτεΐνες ή πρωτεϊνικές περιοχές που αλληλοεπιδρούν συντήκονται στα C -και N -τελικά άκρα των ζευγών I_N και I_C θραυσμάτων αντίστοιχα, τα οποία βρίσκονται ενωμένα με μονάδα ανταποκριτή. Η αλληλεπίδραση μεταξύ των δυο πρωτεϊνών φέρνει σε κοντινή απόσταση τα δυο θραύσματα, οδηγώντας στην ενεργοποίηση της ιντεΐνης και μάτισμα της πρωτεΐνης ανταποκριτή. Ο Umezawa και οι συνεργάτες του χρησιμοποίησαν αυτήν την ιδέα για το σχεδιασμό αρκετών βιοαισθητήρων αλληλεπίδρασης μεταξύ πρωτεϊνών χρησιμοποιώντας GFP και λουσιφεράση ως ανταποκριτές σε μονοκύτταρους οργανισμούς και διαγονιδιακά ζώα (Kanno et al., 2009a, 2011; Ozawa et al., 2001). Αυτές οι προσεγγίσεις έχουν επίσης χρησιμοποιηθεί για την ενδοκυτταρική ανίχνευση των αλληλεπιδράσεων μεταξύ του υποστρώματος του φωσφορυλιωμένου υποδοχέα ινσουλίνης 1 και της N -τελικής περιοχής SH2 της PI3K κινάσης (Paulmurugan et al., 2002), για την αλληλεπίδραση μεταξύ των MyoD και Id πρωτεϊνών (Huang et al., 2012) όπως και για την επαγόμενη αλληλεπίδραση μεταξύ ενός ογκογόνου μεταλλάξεως της Ras και Raf-1 στον παράγοντα επιδερμικής ανάπτυξης (EGF) (Kanno et al., 2009b).

Ανίχνευση αλληλεπιδράσεων μεταξύ πρωτεϊνών και DNA – Ένας παρόμοιος βιοαισθητήρας πρωτεϊνικού trans-ματίσματος έχει επίσης σχεδιαστεί για την ανίχνευση ειδικών αλλαγών στη μεθυλίωση του DNA. Αυτός βιοαισθητήρας σχεδιάστηκε βάση μιας έκδοσης της See VMA

ιντεϊνης, όπου κάθε τμήμα της ιντεϊνης συγχωνεύεται σε μια περιοχή πενταδακτύλου ψευδαργύρου (pentadactyl zinc finger) και μια διάσπαση της περιοχής της λουσιφεράσης. Πρόσδεση των δακτύλων ψευδαργύρου σε συγκεκριμένες αλληλουχίες DNA στόχο επάγει ιντεϊνο-διαμεσολαβούμενη ανασύσταση του ανταποκριτή λουσιφεράσης (Huang et al., 2012). Αυτός ο βιοαισθητήρας χρησιμοποιήθηκε για να ανιχνεύσει την απώλεια της επιγενετικής σήγησης και την αυξημένη προσβασιμότητα μιας αλληλουχίας DNA κοντά στην περιοχή του υποκινητή του L1PA2 παρουσία ενός φάρμακού απομεθυλώσεως.

Ανίχνευση μικρών μορίων – Ο Wood και οι συνεργάτες του έχουν σχεδιάσει μια αλλοστερική ιντεϊνη ως βιοαισθητήρα για την ανίχνευση υποκατάστατων του ανθρώπινου πυρηνικού υποδοχέα ορμόνης (Skretas and Wood, 2005a, 2005b). Μια μικρή ιντεϊνη που δεν έχει τη δυνατότητα να πραγματοποιήσει μάτισμα δρα ως αλλοστερικός μετατροπέας σήματος μεταξύ του υποκατάστατη του υποδοχέα (LBD) και του ενζύμου ανταποκριτή. Ο βιοαισθητήρας βασίζεται στη σύντηξη της ιντεϊνικής περιοχής ματίσματος της Mtu RecA με το ανθρώπινο πυρηνικό υποδοχέα ορμονών LBD. Το κομμάτι αυτό στη συνέχεια συντήκεται στο καρβοξυτελικό άκρο της πρωτεΐνης σύνδεσης της μαλτόζης της *E.coli* (MBP), και στο N-τελικό άκρο της T4 θυμιδυλική συνθάσης (TS). Στην πράξη, η δραστηριότητα του ανταποκριτή TS διαμορφώνεται με τη σύνδεση ορμονών στο LBD, επιτρέποντας την παραγωγή του *E. coli* κυτταρικών σειρών που παρουσιάζουν την παρουσία υποκατάστατων τύπου ορμόνης ως ανάπτυξη φαινοτύπου (Skretas and Wood, 2005a). Αρχικές μελέτες χρησιμοποίησαν τους υποδοχείς των ανθρώπινων οιστρογόνων (ERα) και της θυρεοειδείς ορμόνης (TRβ-1), δείχνοντας ότι αυτός ο αισθητήρας μπορεί να διακρίνει μεταξύ συναγωνιστικούς και ανταγωνιστικούς υποδοχείς (Skretas et al., 2007).

Ανίχνευση οξειδοαναγωγικής κατάστασης - Η Ssp DnaE ιντεϊνη έχει πρόσφατα χρησιμοποιηθεί για την ανάπτυξη ενός βακτηριακού αισθητήρα οξειδοαναγωγής. Αυτός ο αισθητήρας σχεδιάστηκε με την εισαγωγή ενός νέου δισουλφιδικού δεσμού που περιλαμβάνει το καταλυτικό N-τελικό κατάλοιπο της ιντεϊνης Cys (Callahan et al., 2011, 2013). Η δραστηριότητα της ιντεϊνης έχει αναφερθεί από ένα ζεύγος μεταφοράς ενέργειας συντονισμού φθορισμού (FRET) που σχηματίζεται από κυανή και κίτρινη φθορίζουσα πρωτεΐνη. Σε περίπτωση της δεσμευμένης του δισουλφιδικού δεσμού, η ιντεϊνη είναι ανενεργή, παρέχοντας υψηλή FRET. Όταν ο δισουλφιδικός δεσμός ανάγεται, η ιντεϊνη γίνεται ενεργή, προκαλώντας τη N-τελική διάσπαση της κυανής φθορίζουσας πρωτεΐνης με μια συνακόλουθη μείωση του σήματος FRET. Αυτός ο βιοαισθητήρας

έχει χρησιμοποιηθεί επιτυχώς για την επιλογή υπεροξικών μεταλλαγμένων στελεχών *E. coli* (Callahan et al., 2011, 2013).

Ανίχνευση δραστικότητας πρωτεάσης - Η δραστικότητα πρωτεάσης μέσα στο κύτταρο έχει επίσης ανιχνευθεί χρησιμοποιώντας πρωτεϊνικό trans-μάτισμα (Kanno et al., 2007). Σε αυτή τη μελέτη, ο πρωτεϊνικός ανταποκριτής της λουσιφεράσης κυκλοποιήθηκε μέσω του πρωτεϊνικού trans-ματίσματος χρησιμοποιώντας τη Ssp DnaE ιντεΐνη μαζί με ένα συνδέτη υπόστρωματος κασπάσης-3. Στην κυκλοποιημένη κατάσταση, η δραστικότητα της λουσιφεράσης μειώνεται σημαντικά λόγω στερεοχημικής παρεμπόδισης, αλλά αποκαθίσταται πλήρως μετά τη διάσπαση από την κασπάση-3. Ο αισθητήρας αυτός επέτρεψε την ανίχνευση σε πραγματικό χρόνο της δραστικότητας κασπάσης σε ζωντανά ποντίκια (Kanno et al., 2007).

1.4.4 Έλεγχος γονιδίων σε Διαγονιδιακούς οργανισμούς

Διάσπαση της ιντεΐνης σε συνδυασμό με το πρωτεϊνικό trans-μάτισμα (PTS) μπορεί επίσης να χρησιμοποιηθεί για τον έλεγχο της παροχής ετερόλογων γονιδίων σε διαγονιδιακούς οργανισμούς (Evans et al., 2005). Αυτή η προσέγγιση βασίζεται στην διάσπαση της πρωτεΐνης-στόχου σε δύο τμήματα, η οποία μπορεί ανασυσταθεί μετα-μεταφραστικά *in vivo* από το πρωτεϊνικό trans-μάτισμα. Αυτή η προσέγγιση ελαχιστοποιεί τον κίνδυνο για τη μεταφορά συγκεκριμένων γονιδίων τα οποία προσδίδουν επιθυμητά χαρακτηριστικά σε ανεπιθύμητους ξενιστές, όπως στην περίπτωση των γονιδίων ανθεκτικότητας ζιζανιοκτόνου. Για παράδειγμα, το μεταβολικό ένζυμο ακετολακτική συνθάση (ALS) έχει χωριστεί και ανασυσταθεί με επιτυχία από το πρωτεϊνικό trans-μάτισμα σε *E.coli* (Sun et al., 2001). Με έναν παρόμοιο τρόπο, η βακτηριακή συνθάση της 5-φωσφο-3-φωσφορικής (EPSPS) είχε επίσης ανασυσταθεί από την PTS στους χλωροπλάστες του *Nicotiana tabacum* για να παράγει διαγονιδιακές καλλιέργειες ανθεκτικές στα ζιζανιοκτόνα (Chin et al., 2003; Dun et al., 2007). Με τη διαίρεση του γονιδίου μεταξύ των πυρηνικών χρωμοσώματων και των χλωροπλάστων, η πιθανότητα για ανασυνδυασμό και μεταφορά σε ένα ανεπιθύμητο ξενιστή αναστέλλεται έντονα. Παρόμοιες στρατηγικές με τη χρήση του PTS έχουν επίσης χρησιμοποιηθεί για την πλήρως ανασύσταση του ενεργού ετερόλογου β-γλουκορονιδάσης (GUS) στο φυτό *Arabidopsis thaliana* (Yang et al., 2003), καθώς και στα φύλλα του κριθαριού, σόγιας, αραβοσίτου και μπιζελίου (Yang et al., 2006).

Από την ανακάλυψή των ιντεΐνων έχει δοθεί μεγαλύτερη σημασία στις εφαρμογές τους στην Βιοτεχνολογία παρά στη φυσική τους ύπαρξη. Πάρα το γεγονός αυτό, η βασική μελέτη των

συστημάτων αυτών μπορεί να δώσει περεταίρω λύσεις όχι μόνο σε ερευνητές αλλά και στη Βιομηχανία. Τα παραδείγματα που αναφερθήκαν παραπάνω αποδεικνύουν τη χρήση των ιντεϊνών σε τομείς όπως της ιατρικής, φαρμακευτικής και των τροφίμων, στον αγροτικό τομέα ακόμα και στο τομέα ενέργειας.

Παρακάτω θα γίνει μια λεπτομερής ανάλυση μιας εφαρμογής των ιντεϊνών, του συστήματος SICLOPPS για την παραγωγή βιβλιοθηκών κυκλικών πεπτιδίων. Το σύστημα αυτό χρησιμοποιήθηκε ως συμπληρωματικό κομμάτι για τον προσδιορισμό των κυκλικών ολιγοπεπτιδίων που μελετήθηκαν σε αυτή την διατριβή.

1.5 SICLOPPS

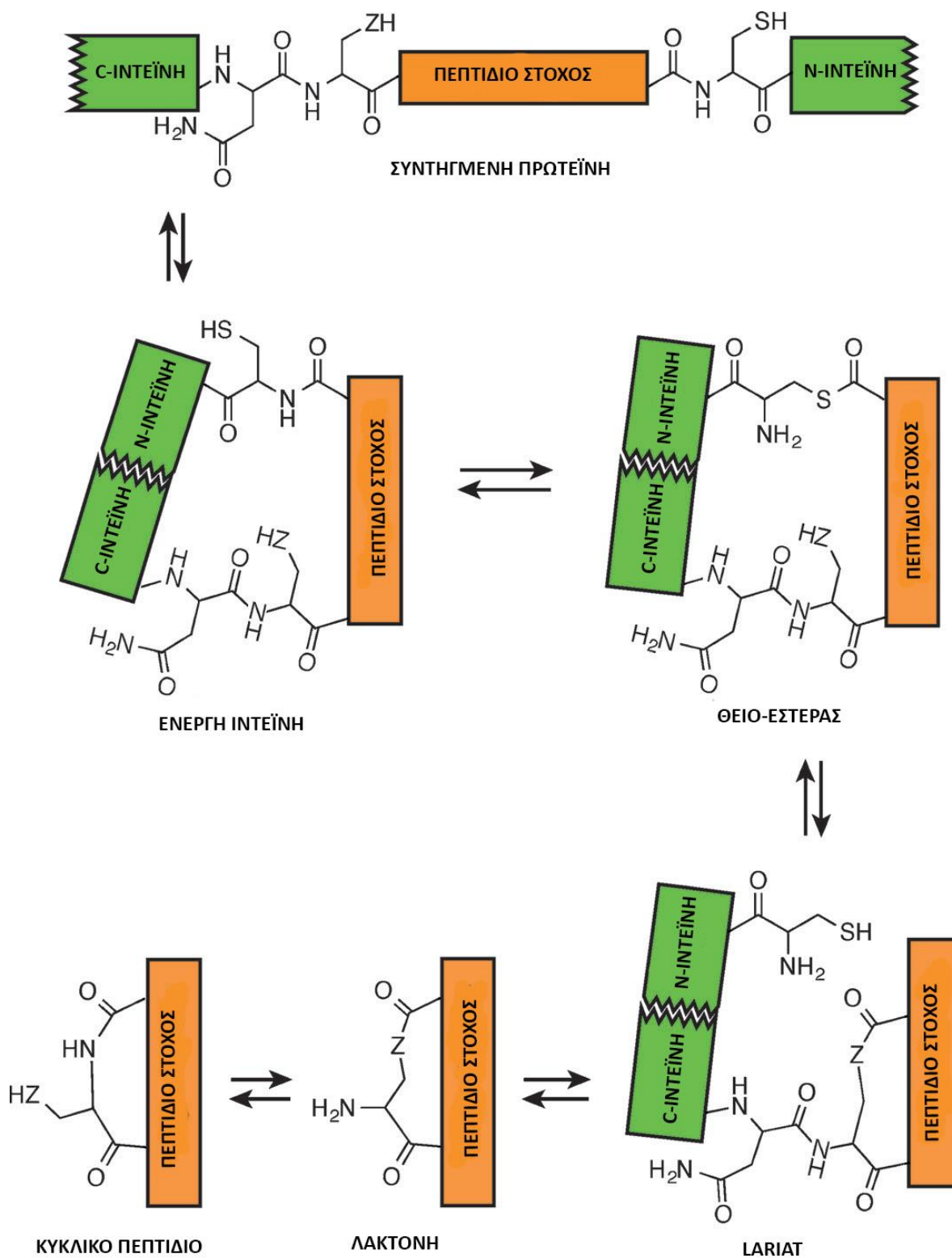
Το SICLOPPS χρησιμοποιεί την DnaE ιντεΐνη που είχε προσδιοριστεί στο γονίδιο που κωδικοποιεί την καταλυτική υπομονάδα της DNA πολυμεράσης III του *Synechocystis sp. PCC6803* (Martin et al., 2001). Το N-τελικό μισό της DnaE η οποία αποτελείται από ένα 123 αμινοξέα στην αλληλουχία της N-ιντεΐνης, κωδικοποιείται από ένα ανοικτό πλαίσιο ανάγνωσης περισσότερο από 745kb μακριά από εκείνο το οποίο κωδικοποιεί το C-τελικό μισό της πρωτεΐνης (προηγείται από μια αλληλουχία 36 αμινοξέων της C-ιντεΐνης). Στο SICLOPPS, η σειρά αυτών των ιντεϊνών είναι σχεδιασμένη έτσι ώστε η C-τελική περιοχή τομέα είναι πριν από τη N-τελική περιοχή. Έτσι μια ενεργή cis-ιντεΐνη μπορεί αποδώσει με μάτισμα που οδηγεί σε κυκλοποίηση της αλληλουχίας του πολυπεπτιδίου (ή ολιγοπεπτιδίου) μεταξύ των δυο ιντεϊνικών περιοχών (Εικ. 1.15) (Scott et al., 1999).

Οι διχασμένες ιντεΐνες έχουν χρησιμοποιηθεί για την παραγωγή βιβλιοθηκών κυκλικών πεπτιδίων μέσω ενός πλασμιδίου βιβλιοθήκη που κατασκευάστηκε χρησιμοποιώντας εκφυλισμένους ολιγονουκλεοτιδικούς εκκινητές και αλυσιδωτή αντίδραση πολυμεράσης (PCR) (Scott et al., 2001; Tavassoli and Benkovic, 2007). Το εκφυλισμένο ολιγονουκλεοτίδιο στους εκκινητές κωδικοποιεί τη μεταβλητή ως τμήμα επαναλήψεις του NNS, όπου το N αντιπροσωπεύει οποιαδήποτε από τις τέσσερις βάσεις του DNA (A, C, G ή T) και S αντιπροσωπεύει τη C ή τη G. Η αλληλουχία NNS καλύπτει τα 32 κωδικόνια και των 20 αμινοξέων, ενώ έχουν εξαλειφθεί τα κωδικόνια τερματισμού (UAA, UGA) από τη βιβλιοθήκη. Η χημεία της ιντεΐνης απαιτεί το πρώτο αμινοξύ να είναι μία πυρηνόφιλη κυστεΐνη ή σερίνη, ενώ για την έναρξη της αντίδρασης απαιτείται αλληλεπίδραση των δυο διακριτών κομματιών ιντεΐνης με δευτερεύοντες δεσμούς στην προτεινόμενη περιοχή ύπαρξης της ενδονουκλεάσης (βλέπε 1.1.5, Εικ.1.5). Δεν υπάρχουν όρια

στον αριθμό των αμινοξέων στο πεπτίδιο στόχου, επιτρέποντας τη δημιουργία κυκλικών πεπτιδίων διαφόρων μεγεθών, ανάλογα με τον αριθμό των επαναλήψεων NNS που ενσωματώνεται σε επίπεδο DNA. Τυπικά, πέντε ή έξι μεταβλητά αμινοξέα εισάγονται, ως ο συνολικός αριθμός των κυκλικών μελών της βιβλιοθήκης πεπτιδίου ($3,2 \times 10^6$ και $6,4 \times 10^7$ αντίστοιχα) να είναι χαμηλότερος από το μέγιστο αριθμό των μετασχηματισμένων *E.coli* (συνήθως $10^8 - 10^9$).

1.5.1 Γενετική επιλογή στο SICLOPPS

Παρά το γεγονός ότι το 1999 ήταν η πρώτη φορά που αναφέρθηκε αυτό το σύστημα (Scott et al., 1999), η δυναμική του SICLOPPS συνειδητοποιήθηκε πέντε χρόνια αργότερα, όταν μια προσέγγιση βασισμένη σε πλασμίδιο για την παραγωγή βιβλιοθηκών κυκλικών πεπτιδίων συνδυάστηκε με ένα βακτηριακό αντιστρεπτό υβριδικό σύστημα (reverse two-hybrid system RTHS) (Akoachere et al., 2007; Horswill et al., 2004; Tavassoli and Benkovic, 2005). Το RTHS που χρησιμοποιήθηκε ήταν βασισμένο από το ρυθμιστικό σύστημα ενός βακτηριοφάγου, που συνδέει την διάσπαση ενός ομο- ή ετερο-διμερούς με την έκφραση τριών γονιδίων ανταποκριτών. Τα HIS3 (φωσφορική δεϋδρατάση της ιμιδαζολο-γλυκερόλης) και KanR (3'-φωσφοτρανσφεράση αμινογλουκοσιδάσης για ανθεκτικότητα σε καναμυκίνη) είναι δύο χημικά συντονίσιμα γονίδια, που υπό επιλεκτικούς όρους απαιτούνται για την επιβίωση σε επιλεγμένα θρεπτικά μέσα. Το τρίτο γονίδιο ανταποκριτής, το LacZ (β-γαλακτοσιδάση) χρησιμοποιείται για την ποσοτικοποίηση της αλληλεπίδρασης πρωτεΐνης-πρωτεΐνης μέσω δοκιμασίες β-γαλακτοσιδάσης. Οι συνδυασμένες μεθοδολογίες δημιούργησαν μια ισχυρή μέθοδο για την ταυτοποίηση των αναστολέων της αλληλεπίδρασης πρωτεΐνης-πρωτεΐνης από τις γενετικώς κωδικοποιημένες βιβλιοθήκες κυκλικών πεπτιδίων. Το πρώτο παράδειγμα διαλογής βιβλιοθήκης SCIOPPS (κυκλο-CXXXXXX) ήταν για αναστολείς της αλληλεπίδρασης της ετεροδιμερικής αναγωγάσης του ριβονουκλεοτιδίου. (Horswill et al., 2004). Ενώ οι περισσότερες μελέτες έχουν επικεντρωθεί στην ταυτοποίηση αναστολέων της αλληλεπίδρασης πρωτεΐνης-πρωτεΐνης, η διαλογή μπορεί να γίνει για κάθε στόχο ή μονοπάτι χρησιμοποιώντας μια βιβλιοθήκη SICLOPPS με μόνο προαπαιτούμενο ένα εύκολα μετρήσιμο φαινότυπο για ενδοκυτταρική παρακολούθηση.

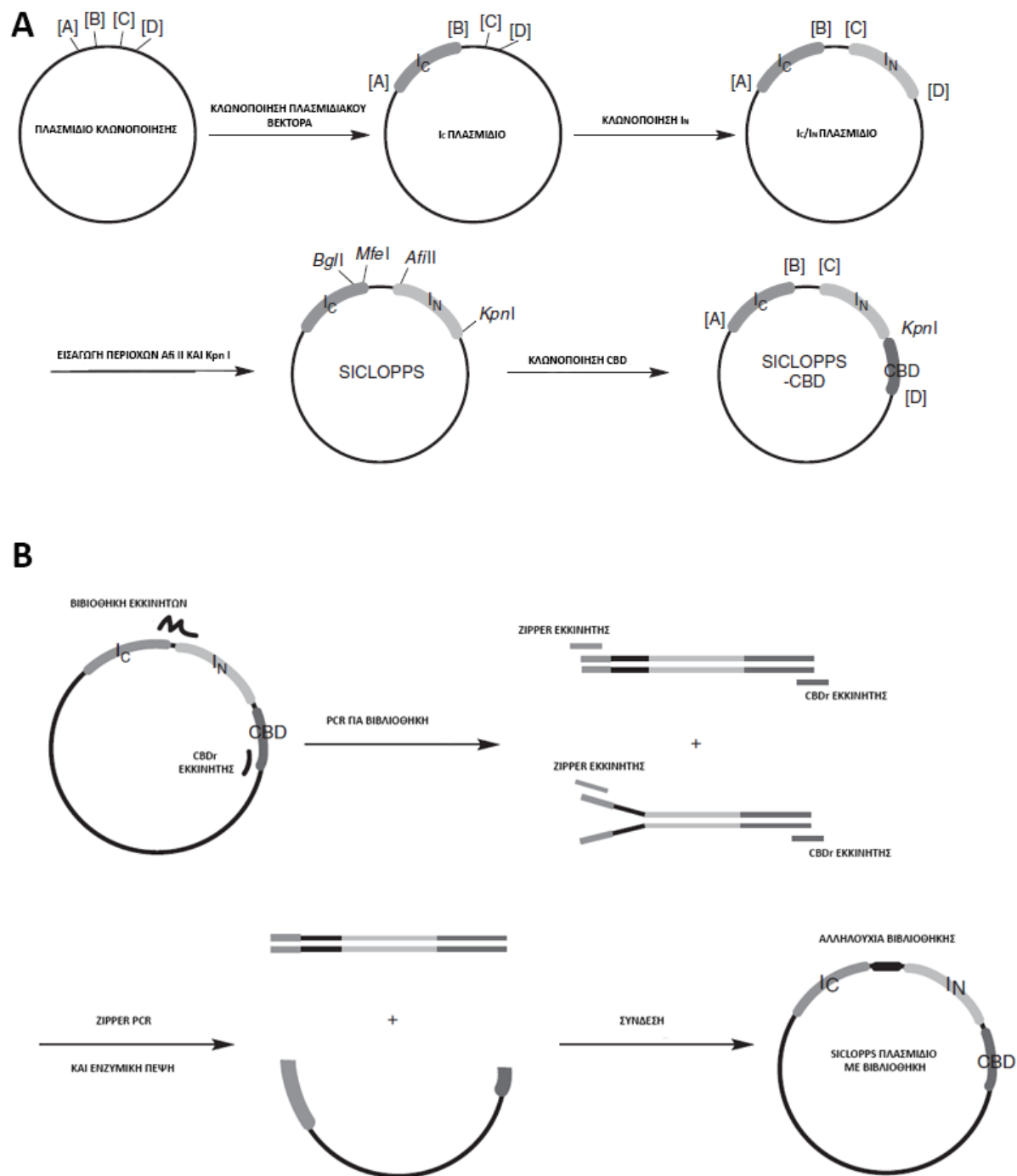


Εικόνα 1.13.

Μηχανισμός αντίδρασης του SICLOPPS.

Ένα παράδειγμα αυτού είναι ο προσδιορισμός ενός κυκλικού πεπτιδίου αναστολέας της πρωτεάσης ClpXP. Η παρακολούθηση έγινε χρησιμοποιώντας μια πρωτεΐνη ανταποκριτή που αποτελείται από μια πράσινη φθορίζουσα πρωτεΐνη (GFP) συντηγμένη σε μια ετικέτα αποικοδόμησης της ClpXP. Μόρια από τη βιβλιοθήκη αναστέλλοντας την πρωτεάση ClpXP εμποδίζουν επίσης την αποικοδόμηση της GFP με αποτέλεσμα τον πράσινο φθορισμό. Βιβλιοθήκες SICLOPPS έχουν επίσης κατασκευαστεί για την ταυτοποίηση ενός κυκλικού πεπτιδίου αναστολέα της DAM μεθυλτρανσφεράσης χρησιμοποιώντας ένα σύστημα παρακολούθησης βασισμένο σε μεταθετάση (Naumann et al., 2008). Μια ενδιαφέρουσα τροποποίηση μίας βιβλιοθήκης SICLOPPS έχει πρόσφατα αναφερθεί, όπου μια συν-μεταφραστική αλληλουχία σήματος μετατόπισης (πεπτίδιο σήμα) συντήχθηκε στο C-τελικό άκρο της ιντεΐνης, εκκρίνοντας εξωκυτταρικά τα προϊόντα της βιβλιοθήκης (Deschuyteneer et al., 2010). Είναι σημαντικό όμως να αναφερθεί, ότι η έκκριση των προϊόντων της βιβλιοθήκης SICLOPPS αφαιρεί το πλεονέκτημα εύκολης αποσυνέλιξης. Κατά συνέπεια, είναι αναγκαία η περαιτέρω διαμερισματοποίηση του κάθε κυττάρου ξενιστή πριν από την χρήση του για τη διαλογή κυκλικών πεπτιδίων, που προσδένουν σε εξωκυτταρικούς στόχους.

Το SICLOPPS μέχρι τώρα περιοριζόταν στην χρήση των 20 φυσικώς απαντώμενων αμινοξέων, ως δομικά στοιχεία για τους δυνητικούς αναστολείς. Αυτός ο περιορισμός έχει πρόσφατα αποφεχθεί με το συνδυασμό του SICLOPPS με ένα ζεύγος ορθογωνικής αμινοακυλο-tRNA συνθετάσης / tRNA_{CUA}, επιτρέποντας την παραγωγή βιβλιοθηκών με κυκλικά πεπτίδια που περιέχουν μη-κανονικά αμινοξέα (Young et al., 2011). Μία βιβλιοθήκη που ενσωματώνει το 21ο αμινοξύ, τη p-βενζοϋλοφαινυλαλανίνη, προβλήθηκε σε *E. coli* έναντι της πρωτεάσης του HIV. Η δοκιμασία χρησιμοποίησε μια παραλλαγή του γονιδίου αντίστασης σε τετρακυκλίνη που περιέχει μια αλληλουχία αναγνώρισης πρωτεάσης του HIV. Προϊόντα της βιβλιοθήκης SICLOPPS που αναστέλλουν την πρωτεάση HIV επίσης εμποδίζουν τη διάσπαση του γονιδίου αντίστασης σε τετρακυκλίνη, και ως εκ τούτου, τα κύτταρα-ξενιστές επιβίωσαν και αυξήθηκαν σε θρεπτικό μέσο που περιείχε τετρακυκλίνη. Όλες αυτές οι μελέτες αποδεικνύουν την ικανότητα και την προσαρμοστικότητα του συστήματος SICLOPPS στην παραγωγή ικανών βιβλιοθηκών, όταν συνδυάζεται με την παρακολούθηση στα κύτταρα, για να πιστοποιηθούν αναστολείς μιας ευρείας ποικιλίας αλληλεπιδράσεων πρωτεΐνης-πρωτεΐνης και ένζυμων.

**Εικόνα 1.14**

A) Κατασκευή πλασμιδίου SICLOPPS. Η μονάδα IC κλωνοποιείται μεταξύ των περιοχών περιορισμού [A] και [B]. Η IN κλωνοποιείται εντός του πλασμιδίου IC μεταξύ των περιοχών περιορισμού [C] και [D]. Εισαγωγή μιας AfiII περιοχής περιορισμού στο 5' άκρο του γονιδίου IN επιτρέπει την κλωνοποίηση των πεπτιδικών αλληλουχιών που στοχεύονται για κυκλοποίηση μεταξύ των περιοχών BglI / MfeI και AfiII.

Εισαγωγή μιας KpnI περιοχής στο 3' άκρο του IN επιτρέπει την εισαγωγή της ετικέτας συγγένειας. Η ετικέτα συγγένειας (CBD) κλωνοποιείται εντός του πλασμιδίου SICLOPPS μεταξύ των περιοχών περιορισμού KpnI και [D]. **B)** Κατασκευή γενετικής βιβλιοθήκης SICLOPPS. Ένας εκκινητής της βιβλιοθήκης που κωδικοποιεί το 5' άκρο του IC, μια εκφυλισμένη ακολουθία και το 3' τέλος του IN χρησιμοποιείται για να ενισχύσει ένα θραύσμα IN-CBD. Το προϊόν αντίδρασης της PCR περιέχει αταίριαστα ζευγη, οπότε εφαρμόζεται ένας δεύτερος γύρος ενίσχυσης με ένα Zipper εκκινητή. Τα ενισχυμένα θραύσματα και το πλασμίδιο SICLOPPS υποβάλλονται σε ενζυμική πέψη με τα αντίστοιχα ένζυμα περιορισμού. Τέλος γίνεται μια σύνδεση του πλασμιδίου με το ένθετο βιβλιοθήκης, με αποτέλεσμα την παραγωγή πλασμιδίων SICLOPPS με βιβλιοθήκη (Scott et al., 1999; Tavassoli and Benkovic, 2007).

1.5.2 Στόχευση σε μεταγραφικούς παράγοντες

Οι μεταγραφικοί παράγοντες θεωρούνται ότι είναι οι πιο χημικώς ανθεκτικοί βιολογικοί στόχοι, αλλά η διαμόρφωση αυτών των αλληλεπιδράσεων μεταξύ πρωτεϊνών προσφέρει τη δυνατότητα ρύθμισης των σηματοδοτικών μονοπατιών μόνο στο αρχικό στάδιο. Αυτή η σημαντική πρόταση έχει οδηγήσει σε συνεχιζόμενες προσπάθειες, αλλά σήμερα υπάρχουν μόνο λίγοι των αναστολές μεταγραφικών παραγόντων (Berg, 2008).

Πρόσφατα αναφέρθηκε η χρήση του SICLOPPS για το σκοπό αυτό. Σε ένα πρώτο παράδειγμα, μία βιβλιοθήκη μεγέθους 6.4×10^7 , που κωδικοποιεί το κυκλο-SGWXXXXXX, υποβλήθηκε σε διαλογή για αναστολές του ομοδιμερισμού της CtBP πρωτεΐνης. Η CtBP διμερίζεται με την παρουσία αυξημένων επιπέδων NADH και σχηματίζουν τον πυρήνα ενός ανασταλτικού συμπλόκου μεταγραφής που αναστέλλει τη μεταγραφή ενός δικτύου των ογκοκατασταλτικών γονιδίων. Προσδιορίστηκε το κυκλο-SGWTVVRMY ως ένας πρώτος αναφερόμενος αναστολέας του διμερισμού του CtBP, με την δραστηριότητα αυτού του αναστολέα να επικυρώνεται *in vitro* αλλά και στα κύτταρα (Birts et al., 2013). Το ταυτοποιημένο μόριο χρησιμοποιήθηκε και για την απόκτηση ενός νέου βιολογικού προσδιορισμού στο ρόλο του NADH-εξαρτώμενου διμερισμού της CtBP σε καρκινικά κύτταρα μαστού. Εκτός από αυτό το παράδειγμα, έχει αναφερθεί πρόσφατα η χρήση μιας βιβλιοθήκης SICLOPPS με 3.2×10^6 μέλη που κωδικοποιεί το κυκλο-CXXXXX για την ταυτοποίηση αναστολέων του επαγωγίμου παράγοντα HIF-1. Ο HIF-1 μεταγραφικός παράγοντας, που αποτελείται από μια οξυγόνο-ρυθμιζόμενη α-υπομονάδα και μια συνεχώς εκφραζόμενη β-υπομονάδα, είναι ένας στόχος υψηλής αξίας για την καταπολέμηση της καρκινικής ανάπτυξης, ρυθμίζει άμεσα την έκφραση εκατοντάδων γονιδίων σε απόκριση όταν υπάρχουν υποξικές συνθήκες. Ο HIF-1 είναι γνωστό ότι υπερεκφράζεται σημαντικά σε μια

ποικιλία καρκίνων και επιτρέπει στα καρκινικά κύτταρα να επιβιώσουν σε ένα περιβάλλον χαμηλού οξυγόνου. Από SICLOPPS διαλογή προσδιόρισε το κύκλο-CLLFVY ως ειδικός αναστολέα της αλληλεπίδρασης HIF-1a / HIF-1b (Miranda et al., 2013). Αυτά τα παραδείγματα αναδεικνύουν τη δύναμη και την ευελιξία των SICLOPPS βιβλιοθηκών για τον εντοπισμό λειτουργικών συνδετικών μορίων από μια ιδιαίτερα προκλητική κατηγορία πρωτεϊνών-στόχων, τονίζοντας επίσης την αξία των χημικών εργαλείων στην κυτταρική βιολογία.

1.5.3 Πέρα από *E. coli*

Το σύστημα SICLOPPS έχει αποδειχθεί να είναι λειτουργικό σε πολλαπλούς οργανισμούς και έχει χρησιμοποιηθεί σε φαινοτυπικές διαλογές για την πιστοποίηση κυκλικών πεπτιδίων που έχουν ανασταλτική δράση. Ένα χαρακτηριστικό παράδειγμα αυτής της προσέγγισης ήταν διαλογή από μία βιβλιοθήκη κυκλικών πεπτιδίων για την εύρεση αναστολέων της σηματοδότησης της ιντερλευκίνης σε ανθρώπινα B κύτταρα (Kinsella et al., 2002). Οι συγγραφείς απέδειξαν ότι η ιντεΐνη του SICLOPPS λειτουργεί σωστά στο κυτταρόπλασμα των κυττάρων θηλαστικού παράγοντας κυκλικά πεπτίδια. Σε ένα πιο πρόσφατο παράδειγμα, η ιντεΐνη του SICLOPPS χρησιμοποιήθηκε και λειτούργησε σε *Saccharomyces cerevisiae*, για τον προσδιορισμό κυκλικού πεπτιδίου που μειώνει τοξικότητα ασυνουκλεΐνης (Kritzer et al., 2009). Όπως και στο προηγούμενο παράδειγμα η λειτουργικότητα της ιντεΐνης προς το σχηματισμό κυκλικών πεπτιδίων αποδείχθηκε, αλλά το πιο ενδιαφέρον στοιχείο αυτής της μελέτης ήταν ότι τα χαμηλά επίπεδα μετασχηματισμού σε συνδυασμό με τη βιβλιοθήκη που κατασκευάστηκε για 7 κατάλοιπα υπέδειξαν ότι το μέγεθος της βιβλιοθήκης (1.28×10^9) ξεπέρασε τον αριθμό των επιλεγμένων μετασχηματισμένων κυττάρων (5×10^6). Επίσης έχει αναφερθεί ο σχεδιασμός συστήματος SICLOPPS (μετάλλαξη N36A στη C-ιντεΐνη) που μπορεί να χρησιμοποιεί σύστημα δυο υβριδίων (two hybrid system THS) σε ζύμες, σε αντίθεση με την ιντεΐνη του συμβατικού SICLOPPS που δεν μπορεί να χρησιμοποιηθεί με THS ζύμης λόγω των διαφορών στη μεταγραφική ρύθμιση προκαρυωτικών και ευκαρυωτικών οργανισμών (Barreto et al., 2009). Αυτές οι μελέτες υποδεικνύουν την πιθανή χρήση του SICLOPPS σε συνεργασία με ευκαρυωτικούς οργανισμούς για τον προσδιορισμό βιοενεργών κυκλικών πεπτιδίων.

Ο συνδυασμός των βιβλιοθηκών SICLOPPS με μεθοδολογίες βασισμένες στη κυτταρική επιλογή έχει ως αποτέλεσμα τη δημιουργία μιας ισχυρής γενετικά κωδικοποιημένης πλατφόρμας διαλογής (Passioura and Suga, 2013). Ενώ πολλές έχουν αναφερθεί έκτοτε άλλες προσεγγίσεις για την παραγωγή βιβλιοθηκών κυκλικών πεπτιδίων, το SICLOPPS παραμένει ένας ισχυρός και εύκολα

προσβάσιμος τρόπος για την παραγωγή αυτών των βιβλιοθηκών. Ένα μειονέκτημα του SICLOPPS είναι περιορισμός του στα 20 φυσικά αμινοξέα που ως αποτέλεσμα περιορίζει τη χημική ποικιλομορφία των μελών της βιβλιοθήκης, αν και αυτό έχει αντιμετωπιστεί εν μέρει με την προσθήκη ενός 21ου μη-φυσικού αμινοξέος. Οι βιβλιοθήκες SICLOPPS θα πρέπει επομένως να θεωρηθούν ως μία ταχεία και εύκολη μέθοδος για την ταυτοποίηση συνδεδετικών μορίων, μια διαδικασία που έχει αποδειχθεί ιδιαίτερα καρποφόρα για την ταυτοποίηση των αναστολέων της αλληλεπίδρασης μεταξύ πρωτεϊνών.

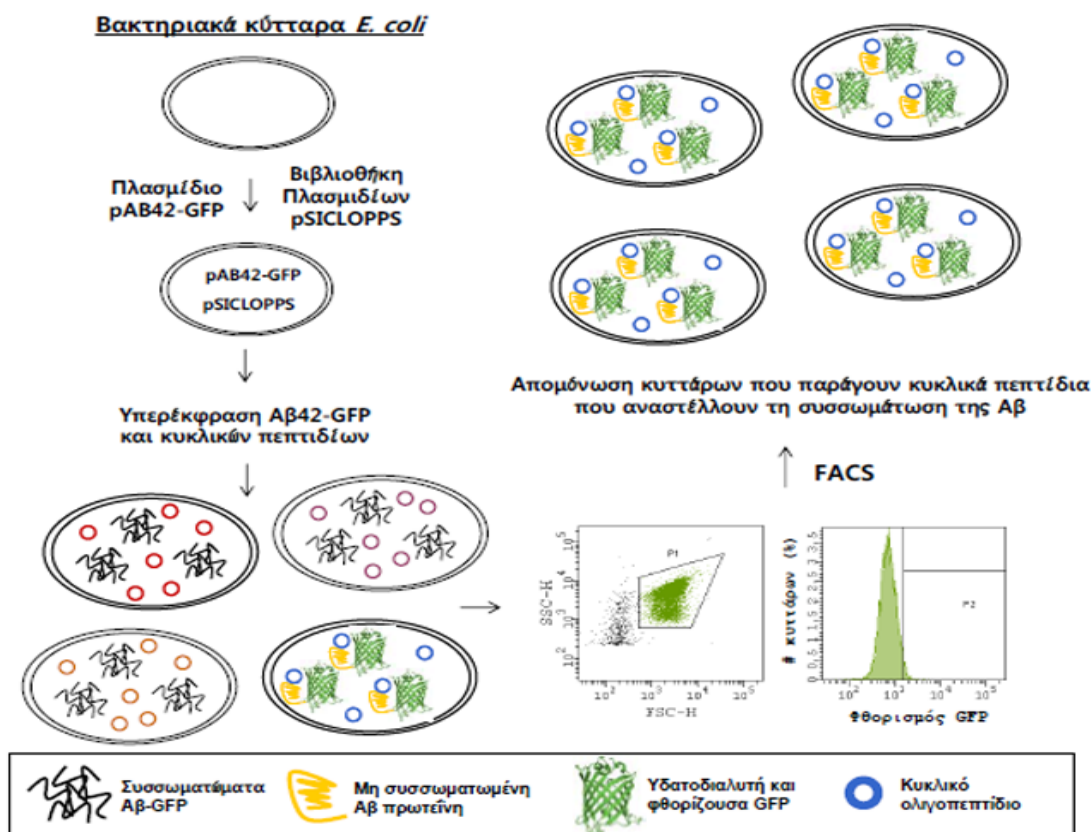
1.6 Στόχος

Ο γενικότερος στόχος αυτής της εργασίας ήταν η παραγωγή κυκλικών ολιγοπεπτιδίων σε κύτταρα *E.coli* με τη χρήση της DnaE ιντεΐνης του κυανοβακτηρίου *Synechocystis sp. PCC6803*. Προτού όμως αναλυθεί ο εξειδικευμένος στόχος αυτής της μελέτης, είναι σημαντικό αναφερθεί ο τρόπος διαλογής των πεπτιδικών αλληλουχιών, ώστε να κατανοηθεί και ο λόγος επιλογής των συγκεκριμένων πεπτιδίων που μελετήθηκαν.

Η χρήση του συστήματος SICLOPPS αποτέλεσε το μέσο διαλογής των πεπτιδίων, ως επακόλουθο μιας έρευνας για την καταπολέμηση του Alzheimer. Πιο συγκεκριμένα στόχος ήταν να βρεθούν κυκλικά ολιγοπεπτίδια (5-και 6-μερη) για την αναστολή της συσσωμάτωσης της Αβ πρωτεΐνης. Η διαλογή έγινε με τη χρήση ενός πλασμιδίου SICLOPPS (Tavassoli and Benkovic, 2007) και ενός πλασμιδίου που περιείχε μια πρωτεϊνική χίμαιρα μεταξύ της Αβ πρωτεΐνης και της πράσινης φθορίζουσας πρωτεΐνης (GFP).

Λόγω της προβληματικής αναδίπλωσης της Αβ, παρά το γεγονός ότι η χίμαιρα παράγεται σε υψηλές ποσότητες στο κυτταρόπλασμα δεν μπορεί να παραμείνει υδατοδιαλυτή, με αποτέλεσμα να συσσωρεύεται υπό τη μορφή κυτταροπλασματικών συσσωματωμάτων (inclusion bodies) στα οποία η GFP παραμένει ανενεργή και μη φθορίζουσα (Wurth et al., 2002). Ωστόσο, έχει βρεθεί πρόσφατα ότι παράγοντες, όπως μεταλλάξεις στο γονίδιο της Αβ ή χημικές ενώσεις με την ιδιότητα να αποτρέπουν τη συσσωμάτωση της Αβ, οδηγούν στην παραγωγή υδατοδιαλυτής χίμαιρας Αβ-GFP που περιλαμβάνει GFP με ενεργό φθοροφόρο, με αποτέλεσμα τα βακτηριακά αυτά κύτταρα να επιδεικνύουν φθορισμό (Baine et al., 2009). Εφαρμόζοντας το απλό αυτό γενετικό σύστημα μετασχηματιστήκαν κύτταρα *E.coli* με τα δυο διαφορετικά πλασμίδια και τα στελέχη που παρήγαγαν κυκλικά πεπτίδια ικανά να αναστείλουν την συσσωμάτωση της Αβ

πρωτεΐνης επιδείξαν φθορισμό. Η επιλογή των κατάλληλων στελεχών έγινε με τη τεχνική του φθορισμό-ενεργοποιούμενου διαχωρισμού κυττάρων (fluorescence-activated cell sorting, FACS). Από αυτή τη διαδικασία διαλογής κυκλικών πεπτιδίων επιλέχθηκαν τέσσερις αλληλουχίες πεπτιδίων που παρατηρήθηκαν με αυξημένο φθορισμό, αυτό πιθανώς να σημαίνει επιτυχημένη αναστολή της συσσωμάτωσης, για περαιτέρω μελέτη της παραγωγής τους. Αυτό το κομμάτι που αποτέλεσε και δουλεία αυτής της διπλωματικής εργασίας, είχε ως απώτερο σκοπό τον προσδιορισμό των βέλτιστων συνθηκών βιοσύνθεσης των τεσσάρων κυκλικών πεπτιδίων.



Εικόνα 1.15.

Γενετική επιλογή και εντοπισμός κυκλικών ολιγοπεπτιδίων με την ικανότητα να αναστέλλουν τη συσσωμάτωση της Aβ πρωτεΐνης. Βακτηριακά κύτταρα *E. coli* μετασχηματίστηκαν με τα πλασμίδια pAB42-GFP και pSICLOPPS που υπερεκφράζουν την πρωτεϊνική χίμαιρα Aβ42-GFP και συνδυαστικές βιβλιοθήκες τυχαίων κυκλικών ολιγοπεπτιδίων, αντίστοιχα. Τα βακτηριακά κύτταρα που παρήγαγαν κυκλικά ολιγοπεπίδια

με την ικανότητα να αναστέλλουν τη συσσωμάτωση της Αβ, οδηγούν στο σχηματισμό της υδατοδιαλυτής και φθορίζουσας Αβ42-GFP πρωτεΐνης με αποτέλεσμα την απόκτηση φθορίζοντα φαινότυπου. Τα σπάνια βακτηριακά στελέχη που παρήγαγαν τα λειτουργικά πεπτίδια απομονώθηκαν με FACS και στη συνέχεια τα λειτουργικά αυτά πεπτίδια ταυτοποιήθηκαν με αλληλούχιση του πλασμιδιακού DNA.

Για τη βιοσύνθεση χρησιμοποιήθηκε η ιντεΐνη του συστήματος SICLOPPS (Ssp DnaE) συντηγμένη μια CBD για επιτυχής διαχωρισμό σε στήλη συγγένειας (ρητίνη Χητίνης). Η μελέτη χωρίστηκε σε δυο μέρη, όπου στο πρώτο μέρος προσδιορίστηκαν οι βέλτιστες συνθήκες υπερέκφρασης του συνολικού συμπλέγματος (πεπτιδίου-ιντεΐνης-CBD) σε κύτταρα E.coli BL21 DE3 ενώ στο δεύτερο μέρος μελετήθηκε η κυκλοποίηση των πεπτιδίων τόσο *in vitro* σε στήλη συγγένειας όσο και *in vivo* μέσα στα κύτταρα. Όπως είναι αντιληπτό ο αρχικός στόχος της μελέτης αυτής φαίνεται να είναι η βιοσύνθεση κυκλικών πεπτιδίων. Ωστόσο, λόγω της έλλειψης πολλών πληροφοριών για τις ιντεΐνες, που οφείλεται στη πρόσφατη τους ανακάλυψη, τόσο από την άποψη της βιολογικής τους αξίας όσο και από τη βιοτεχνολογική τους εφαρμογή, είναι λογικό ότι τα αποτελέσματα αυτής της μελέτης μπορούν να προσφέρουν μια καλύτερη κατανόηση για την λειτουργία αλλά και ύπαρξη αυτών των βιολογικών παραγόντων

2. Πειραματικό μέρος

2.1 Μέθοδοι και Υλικά

Τα πλασμίδια με τις επιλεγμένες αλληλουχίες πεπτιδίων προμηθεύτηκαν από συνεργάτες μέσα σε βακτήρια της κυτταρικής σειράς *E.coli Mc1061*, γνωστά για τη χρήση τους στη κλωνοποίηση πλασμιδίων (Casadaban and Cohen, 1980), η εν λόγω σειρά χρησιμοποιήθηκε επίσης για την επιπλέον παραγωγή πλασμιδίων.

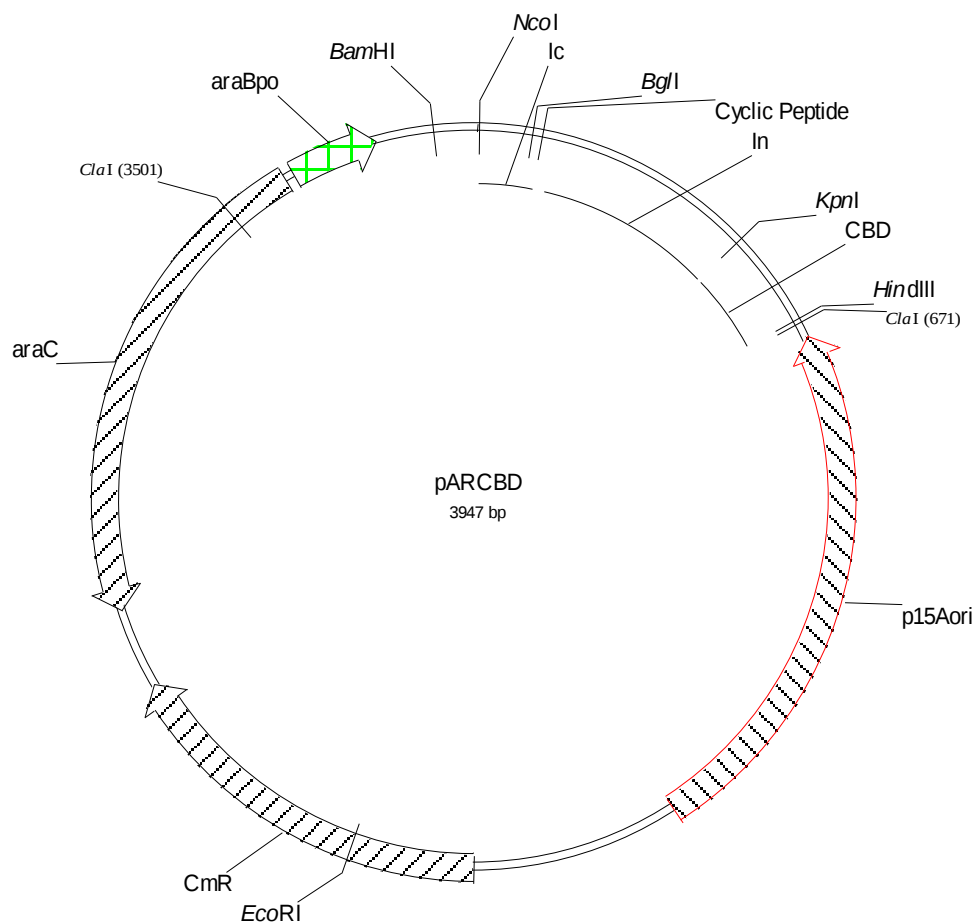
Η έκφραση του πρωτεϊνικού συμπλέγματος πεπτιδίου-ιντεΐνης-CBD για τη βιοσύνθεση των κυκλικών πεπτιδίων έγινε με τη κυτταρική σειρά *E.coli BL21 DE3*, που είναι γνωστή για τα υψηλά επίπεδα έκφρασης χάρη στην ύπαρξη του γονιδίου της T7 RNA πολυμεράσης (Studier and Moffatt, 1986).

Το πλασμίδιο που χρησιμοποιήθηκε για τις μελέτες και περιείχε τις διάφορες αλληλουχίες πεπτιδίων είναι το pARCB-p. Το συγκεκριμένο πλασμίδιο έχει ένα υποκινητή Αραβινόζης (araBp) υψηλής παραγωγικότητας (Schleif, 2000), άρα για την επαγωγή του συστήματος χρησιμοποιήθηκε διάλυμα L-Αραβινόζης σε διάφορες συγκεντρώσεις ανάλογα με τη μελέτη που έγινε σε κάθε περίπτωση. Στο πλασμίδιο υπάρχει επίσης ανθεκτικότητα στο αντιβιοτικό Χλωραμφενικόλη.

Για την εξαγωγή του πλασμιδιακού DNA χρησιμοποιήθηκε κιτ της εταιρίας QUIAGEN (Plasmid DNA extraction and purification). Μετά την απομόνωση ακολούθησε διαγνωστική πέψη με τη χρήση των περιοριστικών ενζύμων Bam HI και HindIII (Takara) για τη πιστοποίηση της ύπαρξης του πλασμιδίου. Ο μετασχηματισμός των βακτηριακών κυττάρων έγινε βάση του πρωτοκόλλου για μετασχηματισμό δεκτικών κυττάρων *E.coli* του βιβλίου Molecular Cloning: A Laboratory Manual, ενώ πρώτα πραγματοποιήθηκε διαδικασία για την μεταβολή κυττάρων στη δεκτική μορφή με πρωτόκολλο από το ίδιο βιβλίο (Sambrook and Russell, 2001).

2.1.1 Κυτταρικές Καλλιέργειες

Για τις υγρές κυτταρικές καλλιέργειες χρησιμοποιήθηκε το θρεπτικό Luria Bertani (LB) (1% Tryptone, 1% NaCl και 0.5% yeast extract), ενώ για στερεές καλλιέργειες σε τρυβλίο χρησιμοποιήθηκε το ίδιο θρεπτικό μαζί με 1.5% Άγαρ. Η θερμοκρασία επώασης ήταν ρυθμισμένη στους 37 °C, εκτός από την περίπτωση που έγινε μελέτη σε διάφορες τιμές θερμοκρασίας (βλέπε 2.2.3). Σε κάθε πείραμα έγιναν προκαλλιέργειες των 4-5 ml για 16-18 h στους 37 °C από μονήρη



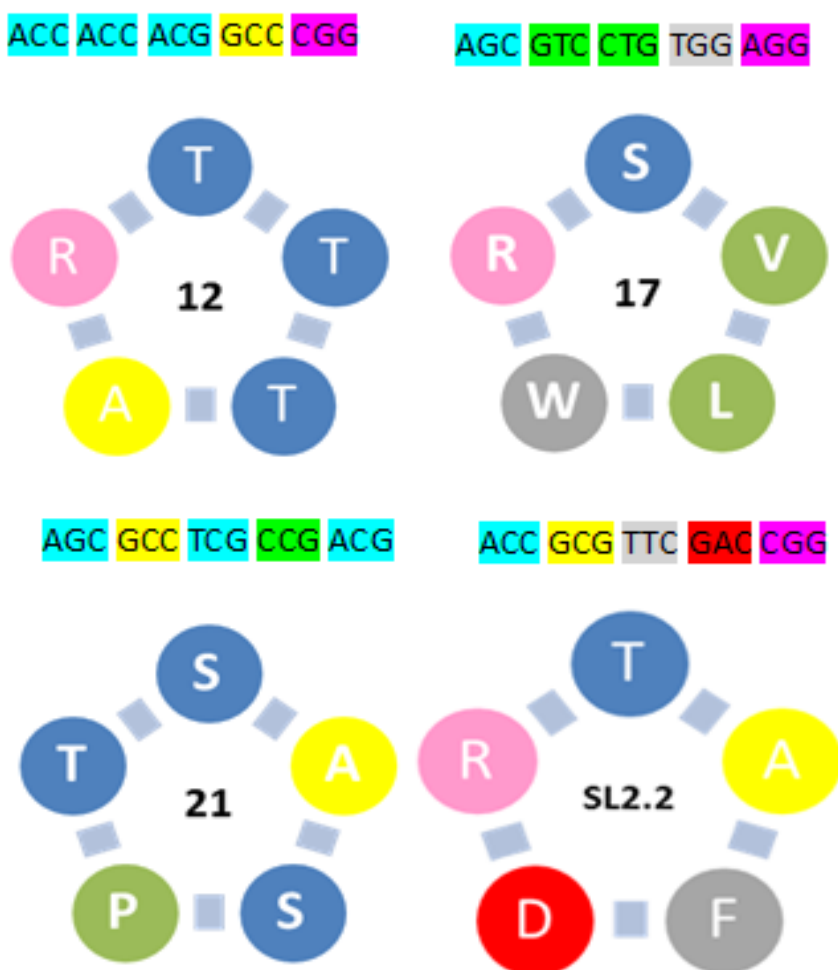
Εικόνα 2.1.

Χάρτης πλασμιδίου pARCB-p

αποικία, που είχε απομονωθεί από τριβλίο, για να γίνει η ανακαλλιέργεια την επόμενη μέρα σε μεγαλύτερο όγκο θρεπτικού αναλόγως με τα προαπαιτούμενα της κάθε μελέτης. Για κάθε πείραμα έγινε ασυνεχής υγρή καλλιέργεια (batch), με ανάδευση στις 150-200 rpm, όγκου 500-1000 ml, ανάλογα με τη μελέτη που διατελέστηκε. Η επαγωγή της έκφρασης του πρωτεϊνικού συμπλέγματος έγινε, για απορρόφηση των κυττάρων στα 600 nm $OD_{600} \approx 0.600$, με L-Αραβινόζη σε συγκεντρώσεις που κυμαίνονταν από 0.01% έως 0.5%. Για τη προστασία των καλλιεργειών από ανταγωνιστικούς μικροβιακούς παράγοντες αλλά και για τη διατήρηση του πλασμιδίου στα κύτταρα χρησιμοποιήθηκε το αντιβιοτικό χλωραμφενικόλη σε συγκέντρωση 30μg/ml για κάθε καλλιέργεια. Τα θρεπτικά που χρησιμοποιήθηκαν στις καλλιέργειες αποστειρώθηκαν σε αυτόκαυστο σε θερμοκρασία 120 °C και πίεση 2 bar, ενώ τα διαλύματα του αντιβιοτικού και της L-Αραβινόζη αποστειρώθηκαν με φίλτρα πάχους 0,2μm.

Μετά το τέλος της κάθε καλλιέργειας τα κύτταρα διαχωρίζονταν από το υγρό θρεπτικό με φυγοκέντρηση στις 6000 rpm για 25 min. Η λύση και ομογενοποίηση των κυττάρων έγινε με τη χρήση υπερήχων σε διαλύμα Tris-Hcl 20mM, NaCl 500mM pH=7.0 ή HEPES 20Mm NaCl 500mM pH=8.0 και φυγοκέντρηση στις 14.000 rpm για 30 min.

Για τα πειράματα χρησιμοποιήθηκαν τέσσερα διαφορετικά στελέχη κυττάρων *E.coli* ,που το καθένα περιείχε ένα από τα τέσσερα πλασμίδια με τις διαφορετικές αλληλουχίες πεπτιδίων που επιλέχθηκαν από συνεργάτες (βλέπε 1.6 Σκοπός) pARCB-12, pARCB-17, pARCB-21 και pARCB-SL2.2. Οι αλληλουχίες των πεπτιδίων φαίνονται στην εικόνα 2.2. και στοιχεία για τα πεπτίδια στον πίνακα 2.1.



Εικόνα 2.2.

Αλληλουχίες πεπτιδίων στόχων για κυκλοποίηση με πρωτεϊνικό μάτισμα.

Πίνακας 2.1.

Στοιχεία πεπτιδίων. Μοριακό βάρος (**MW**), σύσταση αλληλουχίας, ισοηλεκτρικό σημείο (**pI**) και φορτίο.

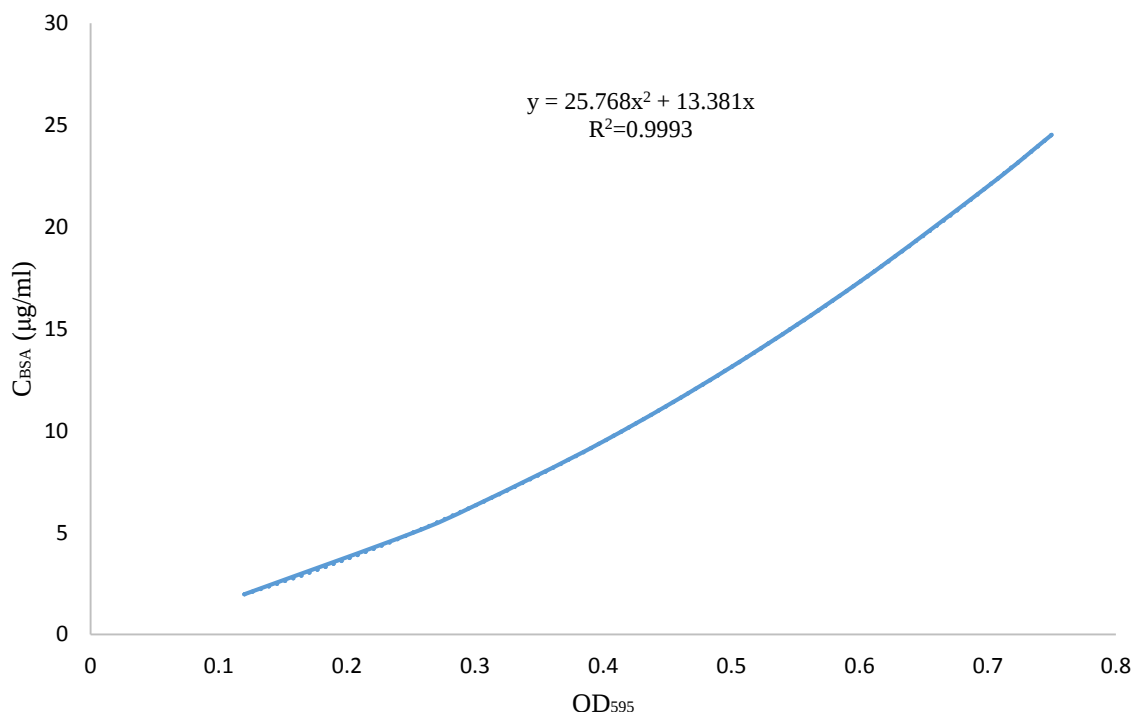
	MW	ΣΥΣΤΑΣΗ	PI	ΦΟΡΤΙΟ
12	548,61 g/mol	20% Βασικό, 20% Υδροφοβικό	11,18	+1
17	659.81 g/mol	20% Βασικό, 20% Ουδέτερο, 60% Υδροφοβικό	11,18	+1
21	461.48 g/mol	80% Ουδέτερο, 20% Υδροφοβικό	6,09	0
SL2.2	608.66 g/mol	20% Όξινο, 20% Ουδέτερο, 40% Υδροφοβικό, 20% Βασικό	6,95	0

2.1.2 Αναλυτικές μέθοδοι

Η εμφάνιση του πλασμιδίου μετά την εξαγωγή του από την κυτταρική σειρά *E.coli Mc1061* αλλά και των θραυσμάτων από τη διαγνωστική πέψη έγινε με ηλεκτροφόρηση σε πηκτή αγαρόζης, ενώ η καθαρότητα και η συγκέντρωση του πλασμιδιακού DNA προσδιορίστηκε με φωτομέτρηση σε μήκη κύματος 230nm, 260nm και 280nm. Η καθαρότητα του δείγματος ορίζεται από τους λόγους 260/280 και 260/230, όπου οι λόγοι αυτοί πρέπει να είναι ίσοι με περίπου 1.7. Στη περίπτωση που οι λόγοι είναι χαμηλότεροι από 1.7, υποδεικνύει στην περίπτωση του πρώτου μόλυνση από πρωτεΐνη ενώ στην περίπτωση του δεύτερου μόλυνση από οργανικά μόρια. Στη περίπτωση που οι λόγοι αποδεικνύουν τη καθαρότητα του δείγματος, για $OD_{260}=1$ η συγκέντρωση του DNA στο δείγμα είναι 50μg/ml (το DNA απορροφά στα 260nm).

Η συγκέντρωση της ολικής πρωτεΐνης στο κυτταροδιάλυμα προσδιορίστηκε με το Bradford Assay (Bradford, 1976). Για κάθε δείγμα χρησιμοποιήθηκαν 2ml αντιδραστηρίου Bradford και 20-200μl

δείγματος. Η απορρόφηση μετρήθηκε στα 595nm και οι συγκεντρώσεις των δειγμάτων προσδιορίστηκαν από την πρότυπη καμπύλη της πρωτεΐνης BSA (Εικ.2.3.).



Εικόνα 2.3.

Καμπύλη αναφοράς BSA.

2.1.2.1 Ηλεκτροφόρηση Πρωτεϊνών

Η πρωτεϊνική ανάλυση των δειγμάτων κυτταροδιαλύματος από τις καλλιέργειες, έγινε με ηλεκτροφόρηση σε πήκτωμα πολυακρυλαμίδης (Polyacrylamide gel electrophoresis, PAGE). Για τον ποιοτικό έλεγχο τόσο της έκφρασης του συμπλέγματος πεπτιδίου-ιντεΐνης όσο και της αντίδρασης ματίσματος, χρησιμοποιήθηκαν δυο μεθοδολογίες PAGE, SDS-PAGE και Native-PAGE (ΠΙΝ. 2.2.). Στην SDS-PAGE χρησιμοποιείται συγχρόνως ο αποδιατακτικός παράγοντας SDS για το σπάσιμο δευτερευόντων δεσμών (π.χ. δισουλφιδικοί δεσμοί) που δημιουργούνται λόγω των αλληλεπιδράσεων μεταξύ πρωτεϊνών, με σκοπό το καλύτερο διαχωρισμό σε πολύπλοκα δείγματα.

Πίνακας 2.2.*Αντιδραστήρια που χρησιμοποιήθηκαν για SDS-PAGE και Native-PAGE.*

ΔΙΑΛΥΜΑΤΑ	SDS-PAGE	NATIVE-PAGE
ΠΗΚΤΩΜΑ		
ΔΙΑΧΩΡΙΣΜΟΥ		
ΑΚΡΥΛΑΜΙΔΗ (30%)	+	+
1.5M TRIS-HCL PH=8.8	+	-
10% SDS	+	+
DH ₂ O	+	+
10% APS	+	+
TEMED	+	+
ΠΗΚΤΩΜΑ		
ΣΥΜΠΥΚΝΩΣΗΣ		
ΑΚΡΥΛΑΜΙΔΗ (30%)	+	+
1.5M TRIS-HCL PH=6.8	+	+
10% SDS	+	-
DH ₂ O	+	+
10% APS	+	+
TEMED	+	+
ΡΥΘΜΙΣΤΙΚΟ		
ΔΙΑΛΥΜΑ ΦΟΡΤΩΣΗΣ		
1M TRIS/HCL PH=6,8	+	+
20% SDS	+	-
ΓΛΥΚΕΡΟΛΗ	+	+
ΜΠΛΕ ΤΗΣ	+	+
ΒΡΩΜΟΦΑΙΝΟΛΗΣ		
ΜΕΡΚΑΠΤΟΑΙΘΑΝΟΛΗ	+	-
ΡΘΜΙΣΙΤΙΚΟ ΔΙΑΛΥΜΑ		
ΗΛΕΚΤΡΟΦΟΡΗΣΗΣ		

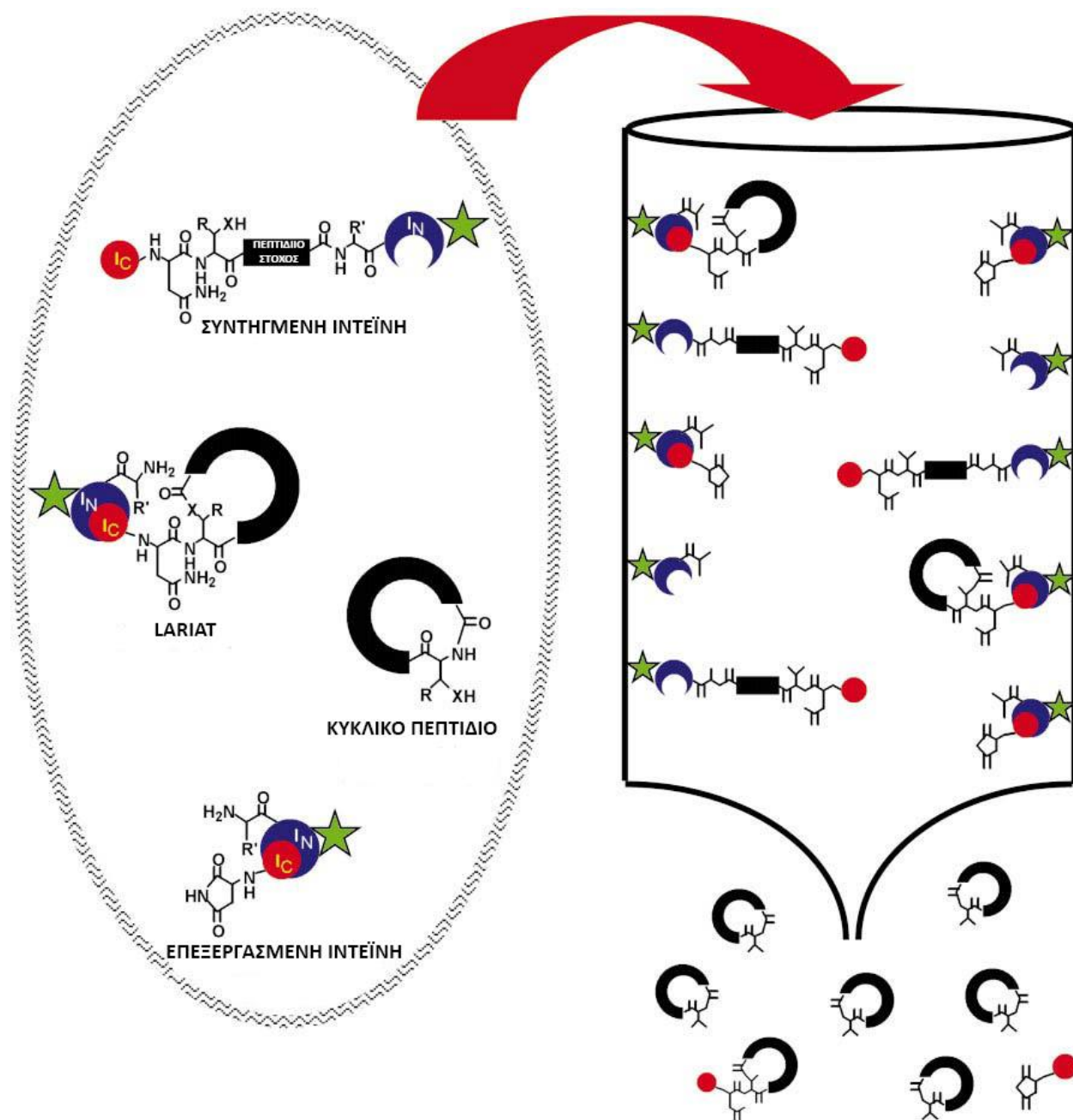
TRIS	+	+
ΓΛΥΚΙΝΗ	+	+
SDS	+	-

Συνέχεια ΠΙΝΑΚΑ 2.2.

Στην Native-PAGE, στην οποία δεν χρησιμοποιούνται αποδιακτικοί παράγοντες σε αντίθεση με την SDS-PAGE, σκοπός είναι ο διαχωρισμός των πρωτεϊνών ενώ βρίσκονται στη φυσική τους κατάσταση, κάτι που μπορεί να βοηθήσει στο να παρατηρηθούν οι αλληλεπιδράσεις μεταξύ πρωτεϊνικών δομικών μονάδων. Για την ανάλυση των δειγμάτων χρησιμοποιήθηκε πήκτωμα διαχωρισμού με 1.5% (σύσταση ακρυλαμίδης). Η εμφάνιση των πρωτεϊνών στα πηκτώματα έγινε με τη διαδικασία Silver Staining σύμφωνα με πρωτόκολλο του Wray και των συνεργατών του (Wray et al., 1981). Λόγω της μεγάλης διακριτικότητας αυτής της διαδικασίας (1ng ανά ζώνη) η συνολική μάζα πρωτεΐνης που φορτωνόταν στα πηκτώματα κυμαινόταν σε τιμές 1-5μg. Το πρόγραμμα ImageJ, για επεξεργασία εικόνας, χρησιμοποιήθηκε για τον ημι-ποσοτικό έλεγχο της έκφρασης και της αντίδρασης ματίσματος. Για την ανάλυση των αποτελεσμάτων έγινε σύγκριση μεταξύ ζωνών σε κάθε ηλεκτροφορητικό πήκτωμα ξεχωριστά, η οποία έγινε σε ποσοστό επί της 100. Τέλος έγινε στατιστική μελέτη, με τη χρήση των διαφόρων τιμών που προσδιορίστηκαν, ανάλογα με το ζητούμενο της κάθε μελέτης.

2.1.2.2 Διαχωρισμός με στήλη συγγένειας

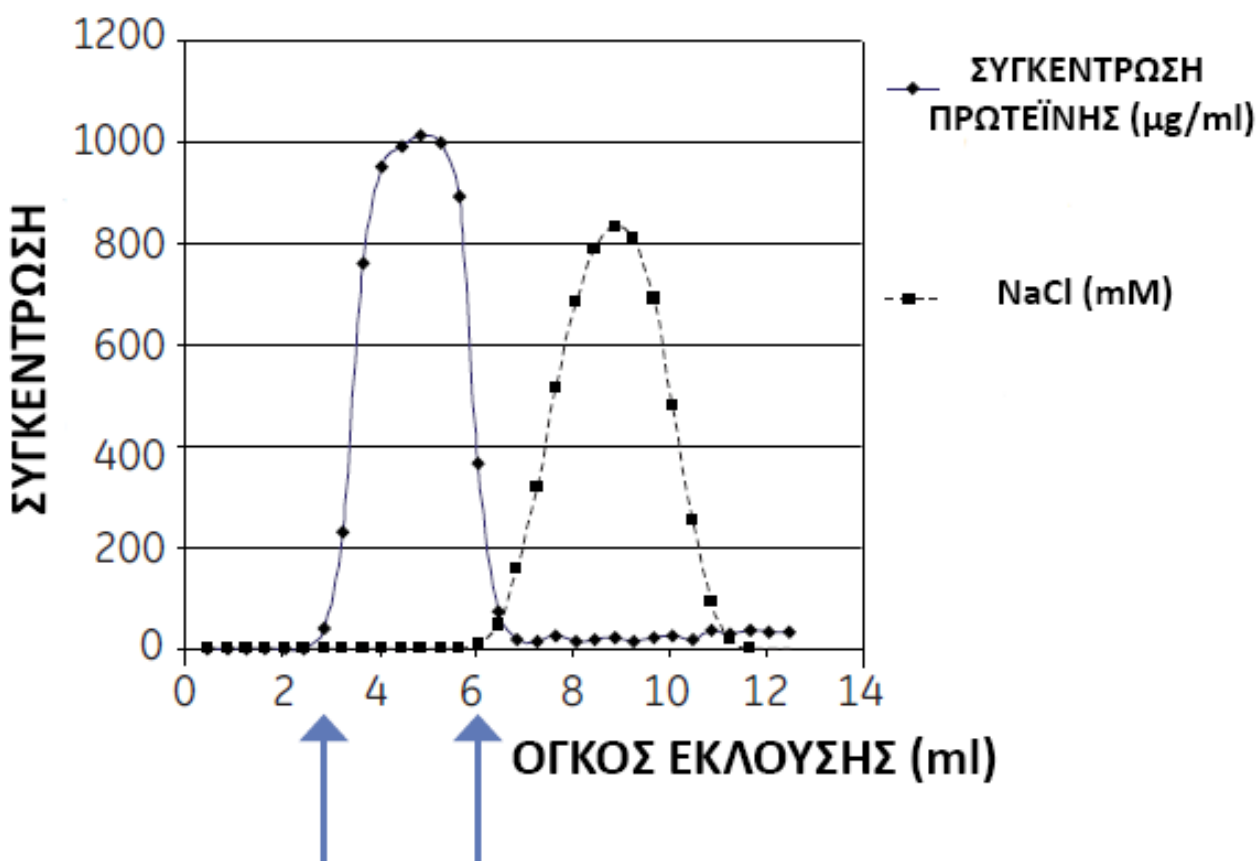
Για τον εκλεκτικό διαχωρισμό του συμπλέγματος πεπτιδίου-ιντεΐνης-CBD από το πολύπλοκο δείγμα κυτταροδιαλύματος, έγινε χρήση στήλης συγγένειας με ρητίνη χητίνης (NEB Chitin Beads). Ο διαχωρισμός γίνεται βάση της εκλεκτικής σύνδεσης της CBD μονάδας με τη χητίνη (Εικ. 2.3.) (Chong et al., 1997). Πριν τη χρήση της στήλης για το διαχωρισμό, η στήλη έπρεπε να περάσει από ένα στάδιο εξισορρόπησης με το διάλυμα έκλουσης, το οποίο διάλυμα ποικίλει αναλόγως με τη μελέτη που έγινε (βλέπε 2.5.2). Αφού κάθε δείγμα περάστηκε από τη στήλη έγινε ξέπλυμα με το αντίστοιχο διάλυμα έκλουσης (περίπου 10 όγκοι) και η στήλη αφέθηκε για επώαση με το αντίστοιχο διάλυμα επώασης σε θερμοκρασία δωματίου για 24-48 ώρες, ώστε να πραγματοποιηθεί η αντίδραση ματίσματος. Ο καθαρισμός της στήλης έγινε με διαδοχικά ξεπλύματα μεγάλων όγκων dH₂O και NaCl 3M, σε αυτή τη σειρά.

**Εικόνα 2.4.**

Διαχωρισμός του συμπλέγματος πεπτιδίου-ιντεΐνης-CBD με στήλη Χητίνης. Προϊόντα συντηγμένα με την ετικέτα συγγένειας (CBD) μπορούν να καθαριστούν από τις υπόλοιπες πρωτεΐνες που δεν έχουν συντηγμένη ετικέτα μετά την λύση των κυττάρων. Η απομόνωση των παραγόμενων *in vitro* κυκλικών πεπτιδίων όπως και των ενδιάμεσων της αντίδρασης γίνεται με μια δεύτερη έκπλυση μετά από επώαση για την πραγματοποίηση του πρωτεϊνικού ματίσματος (Scott et al., 2001).

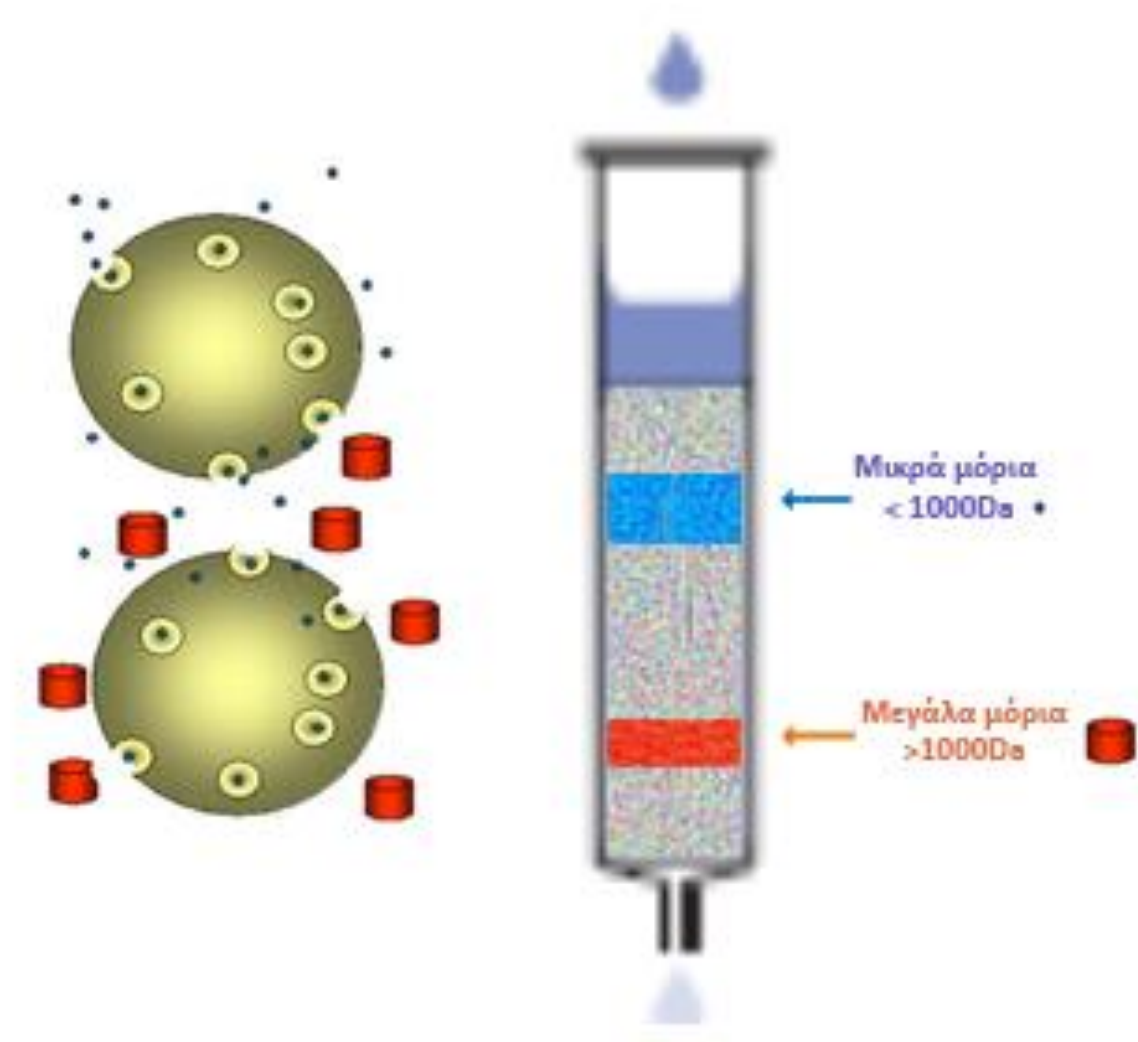
2.1.2.3 Διαχωρισμός με μοριακή διήθηση

Για το διαχωρισμό του *in vivo* παραγόμενου κυκλικού πεπτιδίου (βλέπε 2.3 *in vivo* παραγωγή) από το κυτταροδιάλυμα χρησιμοποιήθηκε η στήλη μοριακής διήθησης PD-10 (GE Healthcare) που περιέχει το υλικό Sephadex G-25. Ο διαχωρισμός πραγματοποιείται σε δυο βήματα έκλουσης βάση μοριακού βάρους με τη χρήση νερού ως διαλύτη. Το όριο διαχωρισμού μοριακού βάρους είναι 1000Da. Για κάθε 2.5 ml δείγματος συλλέγονται 3.5 ml μεγάλων μορίων (>1000Da) στη πρώτη έκλουση, ενώ 6 ml μικρών μορίων (<1000Da) στη δεύτερη (Εικ. 2.5.).



Εικόνα 2.5.

Στο σχήμα αυτό παρουσιάζεται ο πρότυπος διαχωρισμός της πρωτεΐνης BSA από το NaCl. Όπως φαίνεται η πρωτεΐνη εκλύεται σε κλάσματα μεταξύ 2.5-6.0 ml (>1000Da), ενώ το NaCl σε κλάσματα μεταξύ 6.0-12.0 ml (<1000Da).

**Εικόνα 2.6**

Διαχωρισμός πολύπλοκων δειγμάτων με στήλη μοριακής διήθησης PD-10. Τα μικρά μόρια εισέρχονται στις εγκολπώσεις του πορώδους υλικού(Sephadex g-25) και συγκρατούνται για μεγαλύτερο όγκο διαλύτη στη στήλη, κάνοντας δυνατό το διαχωρισμό τους από τα μεγάλα μόρια που διαπερνούν τις εσοχές λόγω μεγέθους και εκλούνται νωρίτερα.

2.1.2.4 Ανάλυση και καθαρισμός με Υγρή Χρωματογραφία Υψηλής Απόδοσης Αντίστροφης Φάσης (RP-HPLC)

Για την ανάλυση του παραγόμενου κυκλικού πεπτιδίου αλλά και το καθαρισμό του από βρωμιές και άλλες ενώσεις χρησιμοποιήθηκε η μέθοδος της RP-HPLC. Με τη μέθοδο αυτή γίνεται διαχωρισμός μορίων βάση της πολικότητας τους, δηλαδή οι πιο υδρόφιλες ενώσεις εκλούνται πριν από τις λιγότερο υδρόφιλες. Για τη διαδικασία αυτή χρησιμοποιήθηκαν δυο στήλες, μια

αναλυτική για τον έλεγχο και μια προπαρασκευαστική για το καθαρισμό (βλέπε 2.1.3 συσκευές). Χρησιμοποιήθηκαν δυο διαλύματα έκλουσης:

A. 2% Ακετονιτρίλιο 98% miliQ H₂O και 0.005% Τριφθοροξικό οξύ (TFA)

B. 95% Ακετονιτρίλιο 5% miliQ H₂O και 0.005% Τριφθοροξικό οξύ (TFA)

Η ανίχνευση έγινε σε μήκος κύματος 220nm και η ροή της αναλυτικής στήλης ήταν ρυθμισμένη στα 1ml/min ενώ της προπαρασκευαστικής στα 1.5-3.0ml/min αναλόγως με το διαχωρισμό.

2.1.3 Συσκευές και μηχανήματα

Για τις υγρές καλλιέργειες χρησιμοποιήθηκαν οι επωαστήρες New Brunswick G25 (περιβάλλον+5°C), Zhincheng ZHWY 100B (περιβάλλον+5°C) και ο ψυχόμενος Innova 4230 (από 5°C και πάνω), ενώ για τις στερεές καλλιέργειες σε τριβλία χρησιμοποιήθηκε ο επωαστήρας Termaks. Χρησιμοποιήθηκαν τρεις φυγόκεντροι στην πειραματική διαδικασία ανάλογα με τον όγκο των δειγμάτων, Hettich MIKPO 2000 για Eppendorf, Hettich Universal 32R με κεφαλή για σωλήνες Falcon 50 ml (swing-out) και Sigma 4KIJ με δυο κεφαλές, μια για σωλήνες όγκου 40 ml (12172-H) και μια για σωλήνες όγκου 300 ml (12256). Η διαδικασία λύσης των κυττάρων έγινε σε μηχανήμα υπερήχων της Sonic&Materials Inc (Vibra-Cell) και χρήση δυο κεφαλών (διάμετρος μύτης 3mm και 13mm) ανάλογα με τον όγκο του δείγματος κυττάρων. Για τη συμπύκνωση δειγμάτων έγινε χρήση του λυοφιλιωτή Christ Alpha 1-4. Το pH το διαλυμάτων ρυθμίστηκε με τη χρήση των pH-μέτρων YSI Enviromental για τιμές pH 4-7 και Comsort p800 για τιμές pH 7-10. Για τις φωτομετρικές αναλύσεις χρησιμοποιήθηκε το φωτόμετρο OPTIZEN. Οι ηλεκτροφορήσεις πρωτεϊνών έγιναν με συσκευή ηλεκτροφόρησης της BioRad Mini-PROTEAN® Tetra Cell. Για τη διαδικασία της υγρής χρωματογραφίας χρησιμοποιήθηκε HPLC συστοιχία, αντλία MERCK-HITACHI L7100, σύστημα εισαγωγής, απαερωτής και αναλυτής UV της HEWLETT-PACKARD Series 1050. Επίσης χρησιμοποιήθηκαν δυο στήλες για RP-HPLC μια αναλυτική Discovery® C18 HPLC Column 5 µm particle size, L × I.D. 25 cm × 4.6 mm και μια προπαρασκευαστική Discovery® C18 HPLC Column 5 µm particle size, L × I.D. 25 cm × 21.2 mm.

2.2 Μελέτη συνθηκών υπερέκφρασης του πρωτεϊνικού συμπλέγματος πεπτιδίου-ιντεΐνης

Έγινε μελέτη της υπερέκφρασης του συνολικού πρωτεϊνικού συμπλέγματος με βάση τρεις παράγοντες, τη συγκέντρωση του επαγωγέα L-Αραβινόζης, το χρόνο επαγωγής και τη θερμοκρασία επαγωγής στις 4 αλληλουχίες πεπτιδίων. Επίσης στα πλαίσια της ίδιας μελέτης έγινε και προσδιορισμός των βέλτιστων συνθηκών *in vivo* ματίσματος, δηλαδή πραγματοποίηση της αντίδρασης κυκλοποίησης στο εσωτερικό των κυττάρων.

2.2.1 Συγκέντρωση επαγωγέα (L-Αραβινόζη)

Η μελέτη αυτή είχε σκοπό την εύρεση της βέλτιστης συγκέντρωσης για τη μέγιστη έκφραση του πρωτεϊνικού συμπλέγματος. Χρησιμοποιήθηκαν τρεις διαφορετικές συγκεντρώσεις L-Αραβινόζης, 0.02%, 0.2% και 0.5%, σε συνδυασμό με τους διάφορους χρόνους επαγωγής. Μετά την επαγωγή των κυττάρων για OD₆₀₀ ίσο με 0,6 (βλέπε 2.1.1.) συλλέχθηκαν δείγματα κάθε ώρα για 5 ώρες και ένα δείγμα στις 24 ώρες. Τα δείγματα αυτά ομογενοποιήθηκαν σύμφωνα με την υποενότητα 2.1.1 και αναλύθηκαν με ηλεκτροφόρηση σε πήκτωμα πολυακρυλαμίδης (βλέπε 2.1.2.1) Η διαδικασία αυτή επαναλήφθηκε για κάθε στέλεχος. Τέλος προσδιορίστηκε η ιδανικότερη τιμή συγκέντρωσης από τα πηκτώματα πολυακρυλαμίδης με τη βοήθεια του ImageJ με σκοπό τη χρήση αυτών των δεδομένων για τη συνέχιση των υπολοίπων πειραμάτων.

2.2.2 Χρόνος επαγωγής

Με βάση τη προηγούμενη μελέτη χρησιμοποιήθηκε η καλύτερη τιμή συγκέντρωσης επαγωγέα για να προσδιοριστεί ο βέλτιστος χρόνος, τόσο για την υπερέκφραση του πρωτεϊνικού προϊόντος όσο και για το ιδανικότερο μάτισμα. Ύστερα από επαγωγή με L-Αραβινόζη σε συγκέντρωση 0.2%, συλλέχθηκαν δείγματα κάθε μια ώρα για 5 ώρες μετά τη επαγωγή. Όπως και στη παραπάνω μελέτη τα δείγματα ομογενοποιήθηκαν και αναλύθηκαν με SDS- και Native-PAGE, ενώ ο προσδιορισμός έγινε ημί-ποσοτικά βάση του προγράμματος ImageJ.

2.2.3 Θερμοκρασία

Σε αυτή τη μελέτη σκοπός ήταν να προσδιοριστεί η βέλτιστη τιμή θερμοκρασίας για τη μέγιστη παραγωγή του πρωτεϊνικού συμπλέγματος πεπτιδίου-ιντεΐνης, αλλά και ως ιδανικότερη συνθήκη για το *in vivo* μάτισμα. Ύστερα από επαγωγή με συγκέντρωση 0.5% L-Αραβινόζης για 3 ώρες σε θερμοκρασία 37 °C, χωρίστηκαν δείγματα σε πέντε διαφορετικές θερμοκρασίες, 5°C, 20°C, 25°C,

30°C, 37°C και επώαστηκαν για 24 ώρες. Η ανάλυση και ο προσδιορισμός έγινε όπως και στις προηγούμενες μελέτες.

2.3 Μελέτη βιοσύνθεσης κυκλικού πεπτιδίου 21

Η μελέτη για την παραγωγή του κυκλικού πεπτιδίου 21 χωρίστηκε σε δυο τμήματα, στη *in vivo* και τη *in vitro* παραγωγή του κυκλικού πεπτιδίου 21. Ο λόγος που έγινε η επιλογή του συγκεκριμένου πεπτιδίου, ήταν η ύπαρξη ποσότητας χημικά συντεθειμένου που χρησιμοποιήθηκε ως δείκτης για το προσδιορισμό με τη μέθοδο RP-HPLC του παραγόμενου από διεργασίες βιοσύνθεσης.

2.3.1 *In vivo* παραγωγή και απομόνωση με στήλη μοριακής διήθησης

Έγινε διαχωρισμός σε εκχύλισμα κυττάρων *E.Coli* BL21 pARCBD-21 με στήλη μοριακής διήθησης PD-10. Η καλλιέργεια των βακτηριακών κυττάρων επώαστηκαν πρώτα για 3 ώρες στους 37°C μετά την επαγωγή με L-Αραβινόζη σε συγκέντρωση 0.5 % και στη συνέχεια επώαστηκαν για άλλες 21 ώρες σε θερμοκρασία 25°C (συνολικός χρόνος επώασης 24 ώρες). Σκοπός ήταν ο διαχωρισμός του παραγόμενου κυκλικού πεπτιδίου από τις μεγάλες πρωτεΐνες (>1000Da). Η διαδικασία διαχωρισμού έγινε βάση της υποεπνότητας 2.1.2.3 Λόγω της ύπαρξης και άλλων μικρών μορίων στο δείγμα (<1000Da) έγινε πιστοποίηση με τη χρήση του δείκτη και καθαρισμός του πεπτιδίου με RP-HPLC. Πριν τη διαδικασία χρωματογραφίας το δείγμα συμπυκνώθηκε σε λυοφιλιωτή.

2.3.2 *In vitro* παραγωγή και απομόνωση με στήλη συγγένειας

Η *in vitro* παραγωγή έγινε με τη χρήση στήλης ρητίνης χητίνης (βλέπε 2.1.2.2) χάρη στην ικανότητα διαχωρισμού με βάση τη συγγένεια λόγω της ύπαρξης της CBD μονάδας στο πρωτεϊνικό σύμπλεγμα. Πρώτα έγινε μια μελέτη για την εύρεση του καταλληλότερου διαλύτη επώασης, που βοηθούσε για την πραγματοποίηση ικανοποιητικότερου πρωτεϊνικού ματίσματος. Οι διαλύτες που μελετηθήκαν ήταν:

- **Διαλύτης επώασης:** HEPES 20mM, NaCl 500mM, Triton, 0.1% EDTA 0.1mM, DTT 30mM pH=8.0
- **Διαλύτης έκλουσης:** HEPES 20mM, NaCl 500mM, 0.1% EDTA 0.1mM pH=8.0

- **Διαλύτης επώασης και έκλουσης:** HEPES 20mM, NaCl 500mM, 0.1% EDTA 0.1mM pH=8.0
- **Διαλύτης επώασης και έκλουσης:** Tris-HCl 20mM, NaCl 500mM pH=7.0

Η ανάλυση έγινε με SDS-PAGE και το πρόγραμμα ImageJ βάση των αποτελεσμάτων της ηλεκτροφόρησης. Για την εκχύλιση μετά την επώαση των συγγενικά κολλημένων τμημάτων του ιντεϊνικού συμπλέγματος έγινε χρήση του διαλύματος έκλουσης που περιείχε 10% SDS. Αφού προσδιορίστηκε ο καταλληλότερος διαλύτης επαναλήφθηκε η διαδικασία με αυτόν και το δείγμα αφού συμπυκνώθηκε σε λυοφιλιωτή αναλύθηκε με RP-HPLC.

3. Αποτελέσματα

3.1 Μελέτη συνθηκών υπερέκφρασης του πρωτεϊνικού συμπλέγματος πεπτιδίου-ιντεΐνης

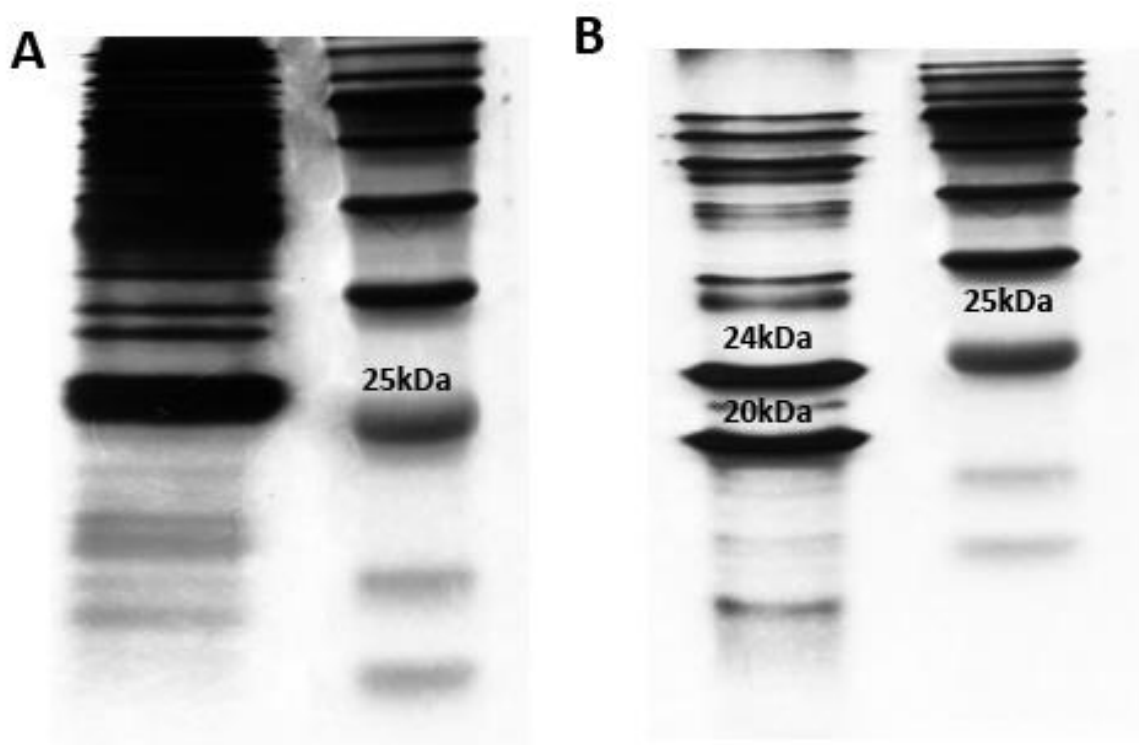
Όπως αναφέρεται στην υποενότητα 2.2 η μελέτη αυτή είχε ως σκοπό την εύρεση των βέλτιστων συνθηκών έκφρασης του συνολικού πρωτεϊνικού συμπλέγματος αλλά και πραγματοποίησης του πρωτεϊνικού ματίσματος. Για την ημι-ποσοτική έκφραση των δεδομένων έγινε χρήση του προγράμματος ImageJ (βλέπε 2.1.2.1). Ωστόσο, για να πραγματοποιηθεί η παραπάνω διαδικασία ημι-ποσοτικοποίησης πρέπει να γίνει μια αξιολόγηση του αποτελέσματος.

Η εφαρμογή των μεθόδων Native-PAGE και SDS-PAGE για τα ίδια δείγματα παρουσίασε διαφορές στη κατάσταση της πρωτεΐνης. Η εμφάνιση μιας ζώνης στη Native-PAGE έναντι των δυο στη SDS-PAGE, υποδεικνύει την πιθανότητα επιτυχούς πρωτεϊνικού ματίσματος (Εικ. 3.1). Για την έναρξη του μηχανισμού ματίσματος (βλέπε 1.5 Εικ.1.13) απαιτείται η αλληλεπίδραση των δυο διακριτών τμημάτων ιντεΐνης με δευτερεύοντες δεσμούς. Τα δυο τμήματα παραμένουν ενωμένα καθ' όλη τη διάρκεια της αντίδρασης ματίσματος με δευτερεύοντες αλληλεπιδράσεις, αλλά και ομοιοπολικά μέσω της αλληλουχίας του πεπτιδίου στόχου. Αφού ολοκληρωθεί η αντίδραση κυκλοποίησης, τα δυο τμήματα παραμένουν συνδεδεμένα μόνο με δευτερεύοντες δεσμούς, διατηρώντας μια εύθραυστη μορφή (Ic-I_N-CBD)στην οποία οι δεσμοί μπορούν να σπάσουν με τη χρήση ενός αποδιατακτικού παράγοντα όπως είναι το SDS. Συγκεκριμένα από την εικόνα της SDS-PAGE η ζώνη στα 24 kDa αποτελεί το μη-ματισμένο σύμπλεγμα και η ζώνη στα 20 kDa αποτελεί το θραύσμα I_N-CBD που είναι αποτέλεσμα της αντίδρασης ματίσματος (ΠΙΝ. 3.1), σε αντίθεση με τη περίπτωση της Native-PAGE όπου η ζώνη στα 25 kDa αποτελεί το άθροισμα του συνολικού μη-ματισμένου συμπλέγματος και της μορφής Ic-I_N-CBD.

Πίνακας 3.1

Μοριακό βάρος των τμημάτων του πρωτεϊνικού συμπλέγματος.

ΤΜΗΜΑΤΑ	MW
Ιντεΐνη –C	4kDa
Κυκλικό Πεπτιδιο	460-660Da
Ιντεΐνη-N	14kDa
CBD	6kDa

**Εικόνα 3.1.**

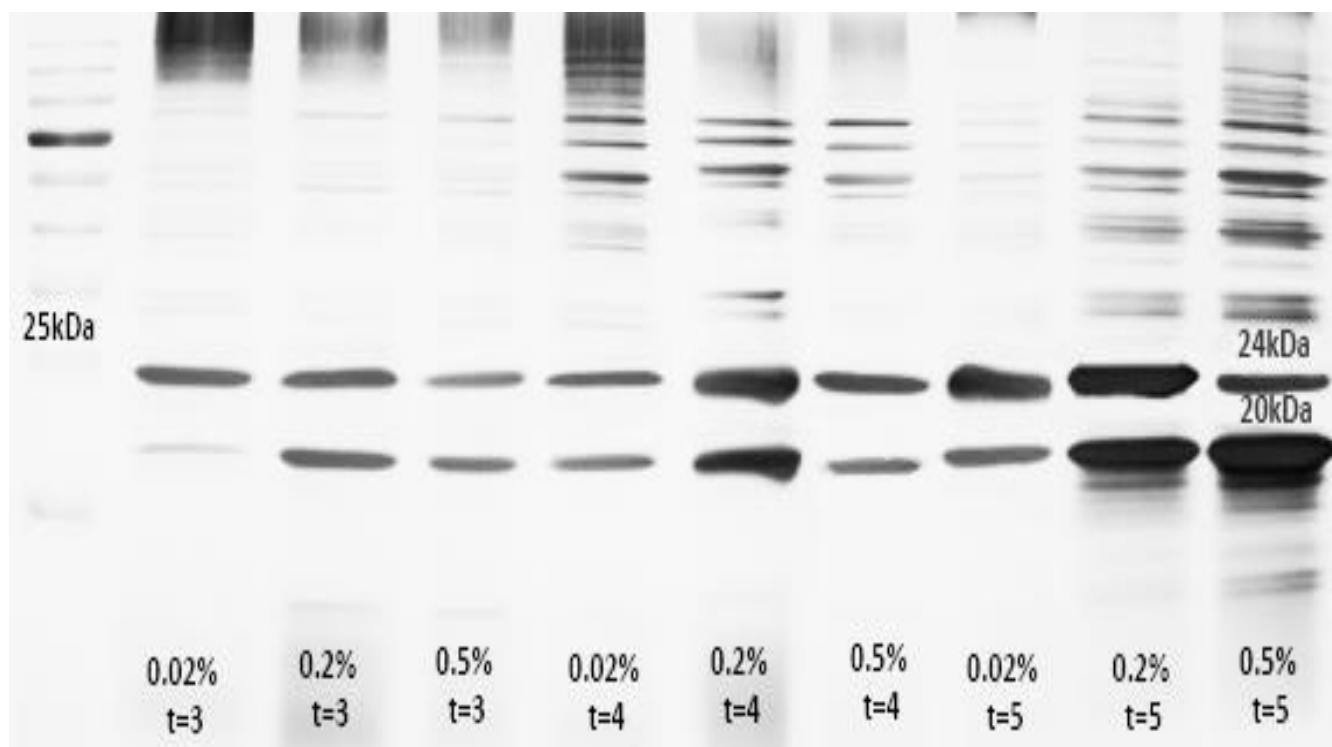
Ηλεκτροφόρηση κυτταροδιαλύματος A)Native-PAGE και B)SDS-PAGE

3.1.1 Συγκέντρωση επαγωγέα (L-Αραβινόζη)

Στη μελέτη αυτή συγκριθήκαν τα αποτελέσματα πειραμάτων έκφρασης με διαφορετικές συγκεντρώσεις L-Αραβινόζης, 0.02%, 0.2% και 0.5%, σε διάφορους χρόνους που αποτυπώθηκαν με πηκτώματα ηλεκτοφόρισης πρωτεϊνών (Εικ. 3.2). Στο γράφημα στην εικόνα 3.3 παρουσιάζονται τα ποσοστά έκφρασης του συνολικού συμπλέγματος (συμπεριλαμβανομένης της ποσότητας που ματίζεται) ως αποτέλεσμα της χρήσης των τριών διαφορετικών συγκεντρώσεων για χρόνους $t=3, 4, 5$ και 24 h μετά την επαγωγή (Εικ. 3.3). Όπως φαίνεται, για τη συγκέντρωση 0.2% η έκφραση ήταν πιο παραγωγική έναντι των άλλων δυο για τις πρώτες πέντε ώρες, με μέγιστη διαφορά (25-30%) για το χρόνο $t=4$ h. Ωστόσο σε χρόνο $t=5$ h παρατηρείται αύξηση της

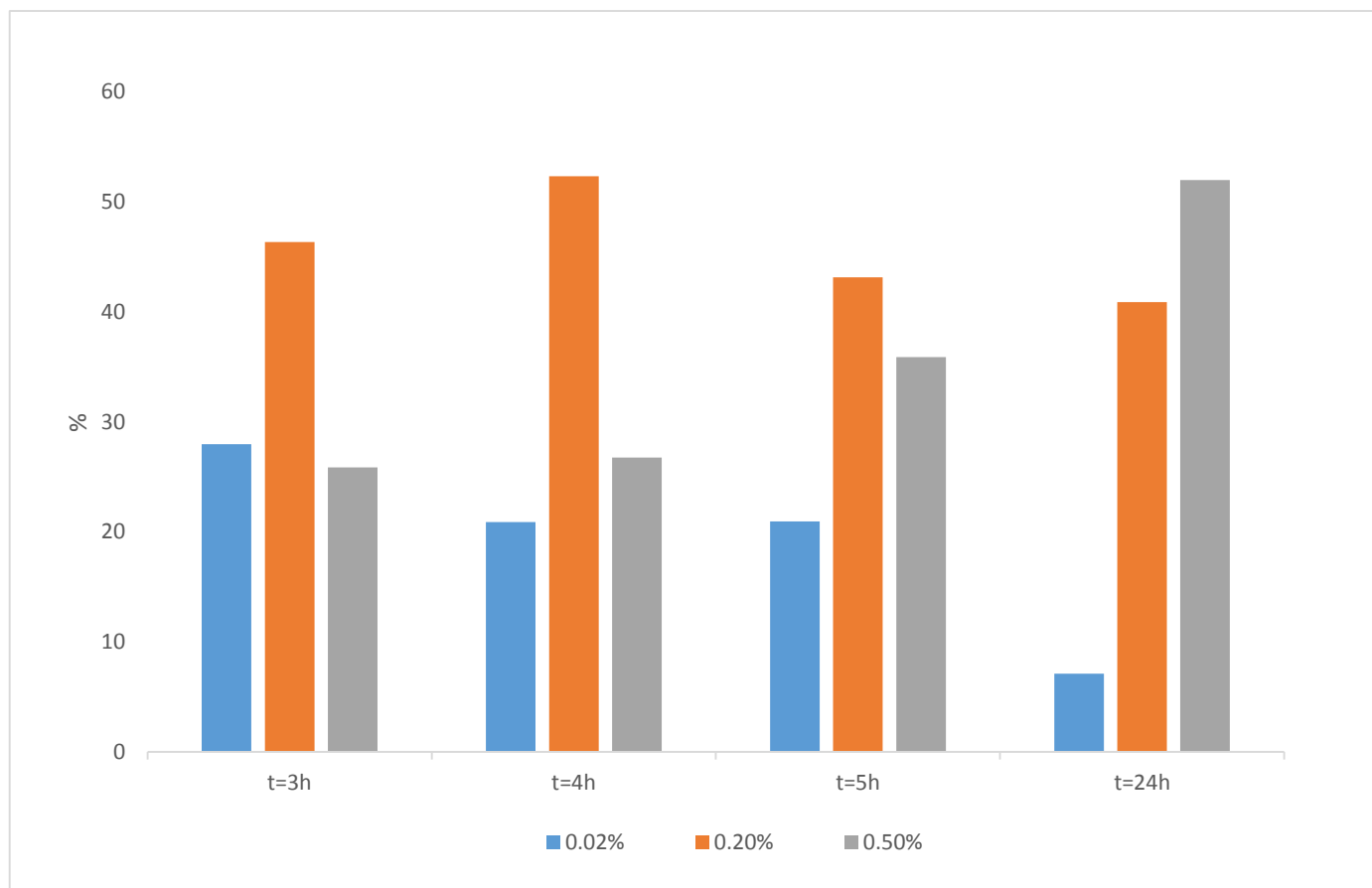
έκφρασης με συγκέντρωση 0.5%, με ποσοστό πολύ κοντινό στη συγκέντρωση 0.2% (διαφορά μικρότερη από 10%), ενώ στις 24 h μετά την επαγωγή εμφανίζεται αύξηση της έκφρασης του συμπλέγματος για συγκέντρωση 0.5% σε σχέση με τη συγκέντρωση 0.2% με διαφορά που ξεπερνάει τα 10 %.

Τα αποτελέσματα αυτής της μελέτης, όπως και των μετέπειτα, αποτέλεσαν δεδομένα για τις επόμενες. Συγκεκριμένα από τα εξαγόμενα στοιχεία αυτής έγινε αντιληπτό, ότι για τη γρήγορη παραγωγή της πρωτεΐνης (βλέπε 2.2.2 και 2.3.2), δηλαδή επαγωγή κάτω από 5 ώρες, έπρεπε να χρησιμοποιηθεί συγκέντρωση επαγωγέα 0.2%. Ωστόσο στην περίπτωση μέγιστης παραγωγής (βλέπε 2.2.3 και 2.3.1), στην οποία ο χρόνος επώασης έφτανε τις 24 ώρες μετά την επαγωγή, έγινε χρήση της τιμής 0.5%.



Εικόνα 3.2.

SDS-PAGE δειγμάτων από μελέτη προσδιορισμού βέλτιστης τιμής συγκέντρωσης επαγωγέα.



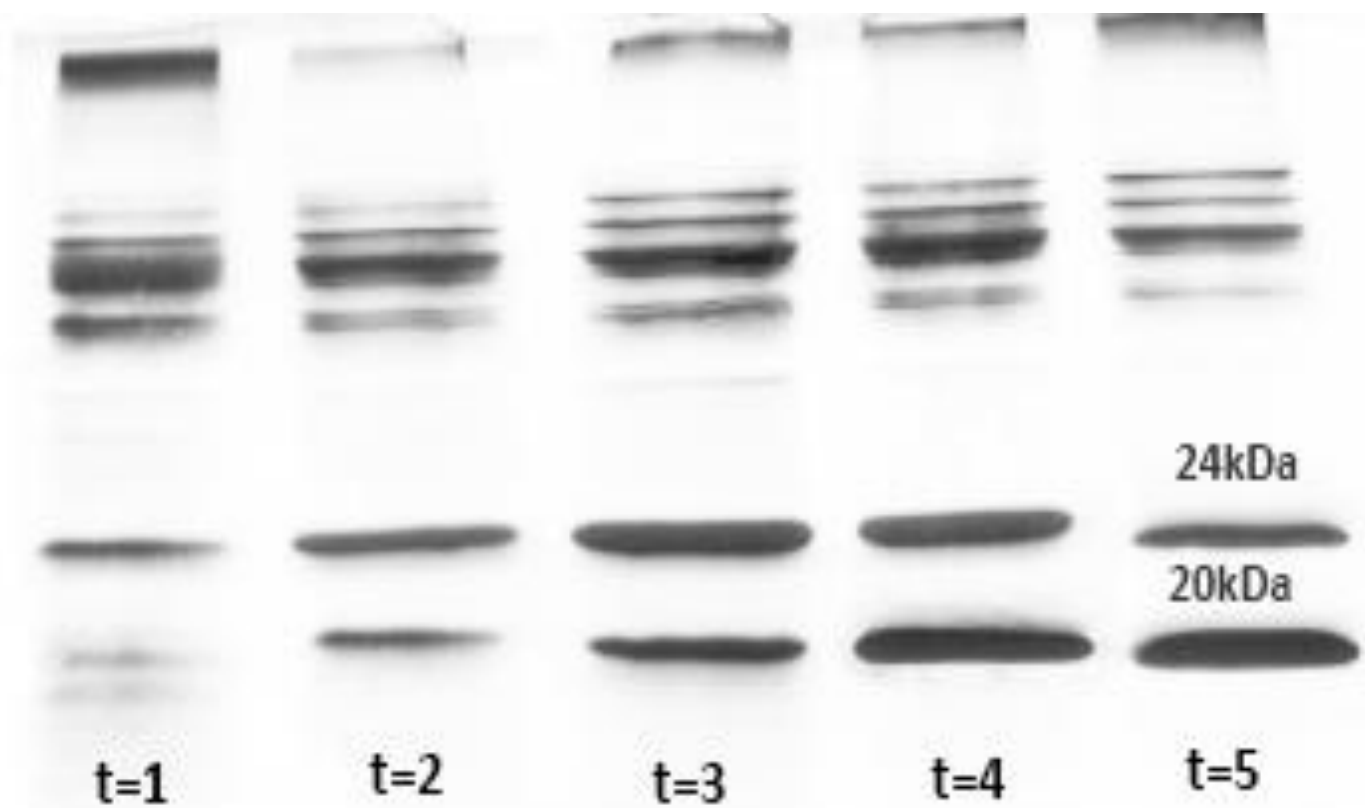
Εικόνα 3.3.

Στο ραβδόγραμμα συγκρίνεται το επίπεδο έκφρασης του συμπλέγματος πεπτιδίου-ιντεΐνης βάση των τριών συγκεντρώσεων L-Αραβινόζης (0.02, 0.2 και 0.5%) σε πέντε διαφορετικούς χρόνους μετά την επαγωγή.

3.1.2 Χρόνος επαγωγής

Βάση των αποτελεσμάτων της προηγούμενης μελέτης, χρησιμοποιήθηκε μια συγκέντρωση επαγωγέα 0.2%, για την άμεση παραγωγή του συμπλέγματος. Πιο συγκεκριμένα σκοπός ήταν η εύρεση της τιμής χρόνου που είχε την πιο αποδοτική έκφραση τόσο από άποψη συνολικού προϊόντος όσο και από άποψη μη-ματισμένης πρωτεΐνης. Όπως και στην προηγούμενη μελέτη έγινε στατιστική σύγκριση αποτελεσμάτων από ηλεκτροφόριση πρωτεϊνών (Εικ.3.4). Από το ραβδόγραμμα (Εικ.3.5.) είναι προφανής η βέλτιστη έκφραση για δυο χρόνους, t=3 και 4h, με

διαφορά πάνω από 5% σε σχέση με την επόμενη μεγαλύτερη τιμή. Οι δυο αυτές τιμές ενώ δεν παρουσιάζουν μεγάλη διαφορά μεταξύ τους (περίπου 1%) στην συνολική παραγωγή του συμπλέγματος, εμφανίζουν διαφορές στην πραγματοποίηση πρωτεϊνικού ματίσματος. Για χρόνο $t=3$ h η ματισμένη μορφή I_N-CBD αποτελεί το 41 % του συνολικού προϊόντος, ενώ στην περίπτωση για $t=4$ h η αντίδραση πρωτεϊνικού ματίσματος έχει πραγματοποιηθεί για το 53% του συνολικού παραγόμενου ιντεϊνικού συμπλέγματος.

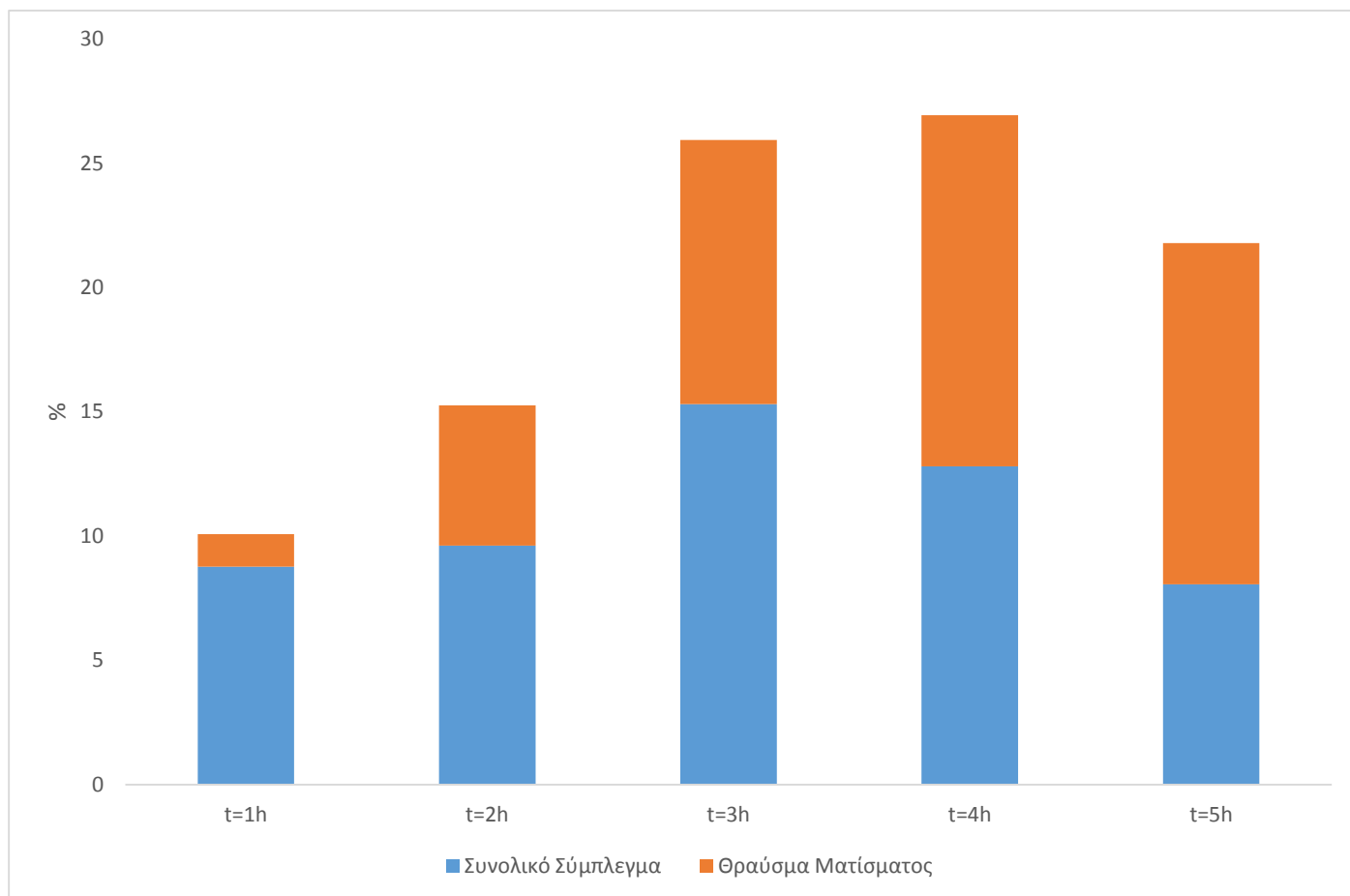


Εικόνα 3.4.

SDS-PAGE δειγμάτων από μελέτη προσδιορισμού βέλτιστης τιμής χρόνου επαγωγής για αποδοτική έκφραση.

Για τη μελέτη αυτή, όπως αναφέρεται και παραπάνω, ήταν σημαντικό να βρεθεί ο χρόνος επώασης που προσφέρει τη μεγαλύτερη έκφραση του πρωτεϊνικού συμπλέγματος, για χρόνους επαγωγής κάτω από 5 ώρες. Ως αποτέλεσμα αυτής της μελέτης, ήταν η δυνατότητα να χρησιμοποιηθεί αυτή

η τιμή στη *in vitro* παραγωγή του κυκλικού πεπτιδίου (βλέπε 2.3.2). Οπότε, είναι λογικό ότι κριτήριο επιλογής ανάμεσα στις δυο αυτές τιμές, είναι η εμφάνιση του μικρότερου ποσοστού ματισμένου προϊόντος, κάτι που παρατηρείται για χρόνο $t=3$ h.



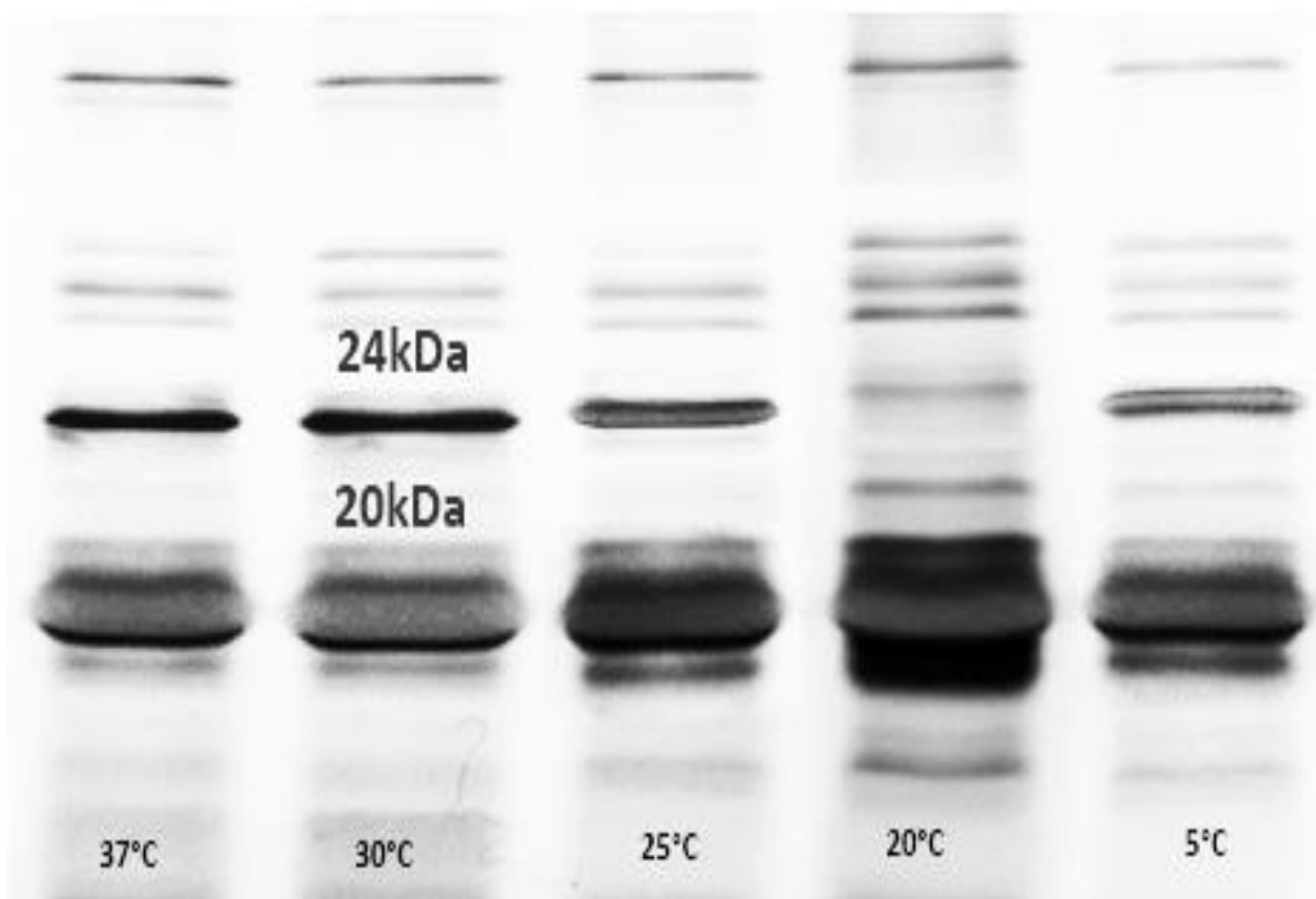
Εικόνα 3.5.

Στο ραβδόγραμμα συγκρίνεται το επίπεδο έκφρασης του συμπλέγματος πεπτιδίου-ιντεΐνης βάση του χρόνου επαγωγής. Παρουσιάζεται επίσης το ποσοστό επί της 100 ματισμένου και μη-ματισμένου συμπλέγματος για κάθε τιμή χρόνου.

3.1.3 Θερμοκρασία

Από τις προηγούμενες μελέτες προσδιορίστηκαν η βέλτιστη τιμή συγκέντρωσης επαγωγέα και του χρόνου για μια αποδοτική παραγωγή της πρωτεΐνης, αλλά επίσης κάτι που δεν αναφέρθηκε πριν είναι το γεγονός της μέγιστης παραγωγής συνολικού προϊόντος (ματισμένο και μη-) σε χρόνο

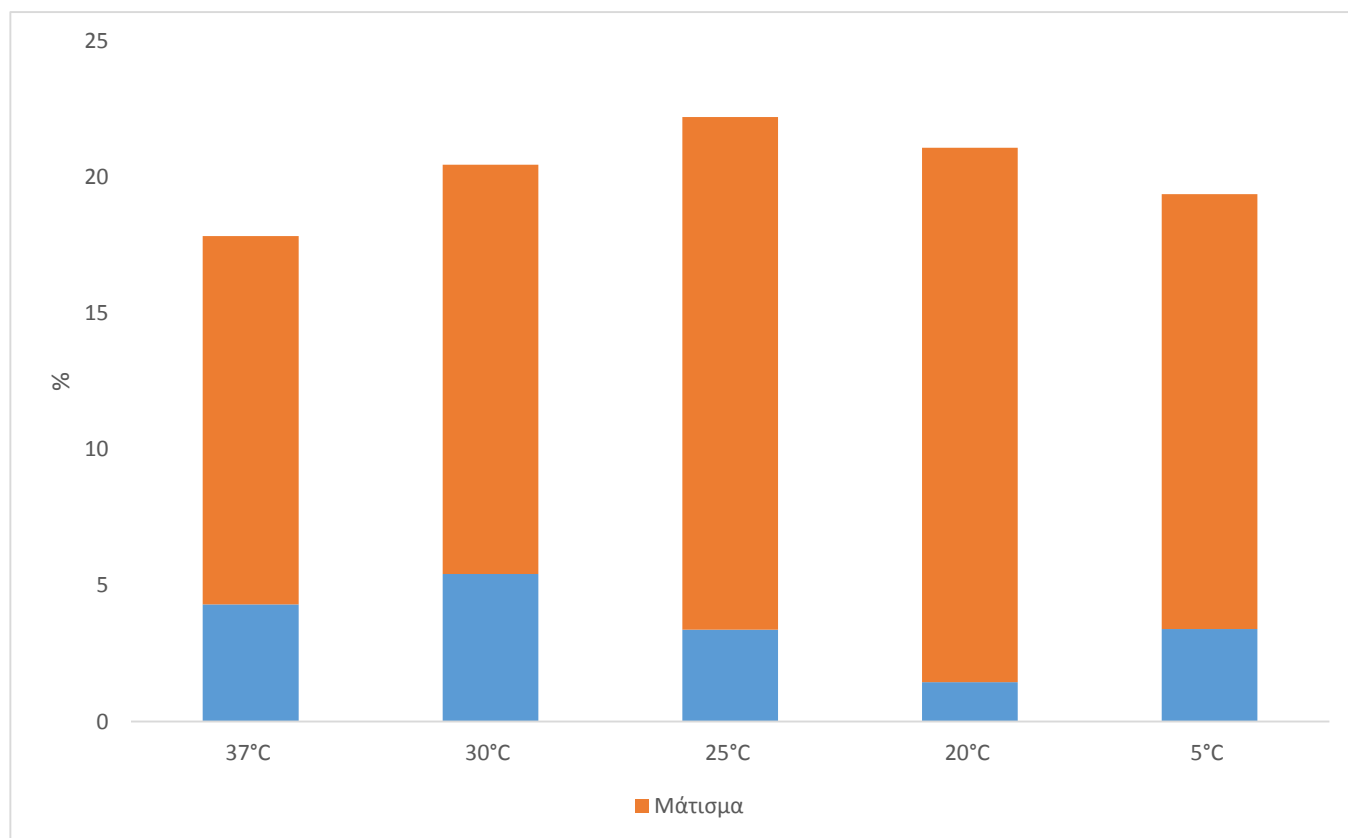
24 h από την επαγωγή. Με αυτό ως δεδομένο η μελέτη αυτή αποσκοπούσε στον προσδιορισμό της ικανοποιητικότερης τιμής θερμοκρασίας για τη παραγωγή του συμπλέγματος σε συνδυασμό με την ικανότητα εφαρμογής της αντίδρασης ματίσματος. Σε ορισμένα πηκτώματα που χρησιμοποιήθηκαν για αυτή τη μελέτη παρατηρήθηκε ότι για ορισμένες θερμοκρασίες το ποσοστό του ματίσματος ξεπερνούσε το 99% (Εικ.3.6.). Όπως ήταν αναμενόμενο σε θερμοκρασία 37 °C



Εικόνα 3.6.

SDS-PAGE δειγμάτων από μελέτη προσδιορισμού βέλτιστης τιμής θερμοκρασίας για μέγιστη παραγωγή.

όπως και 5° C παρατηρήθηκε η μικρότερη παραγωγή σε σχέση με όλες τις υπόλοιπες τιμές θερμοκρασίας (Εικ.3.7.). Η μέγιστη έκφραση όπως φαίνεται στο ραβδόγραμμα επιτυγχάνθηκε σε τιμή 25° C, στην οποία το *in vivo* μάτισμα πραγματοποιήθηκε σε ποσοστό 85%. . Όμως, το υψηλότερο ποσοστό ματίσματος παρατηρήθηκε σε θερμοκρασία 20° C, τιμή που παρουσίασε τη δεύτερη μεγαλύτερη παραγωγή του συνολικού προϊόντος.

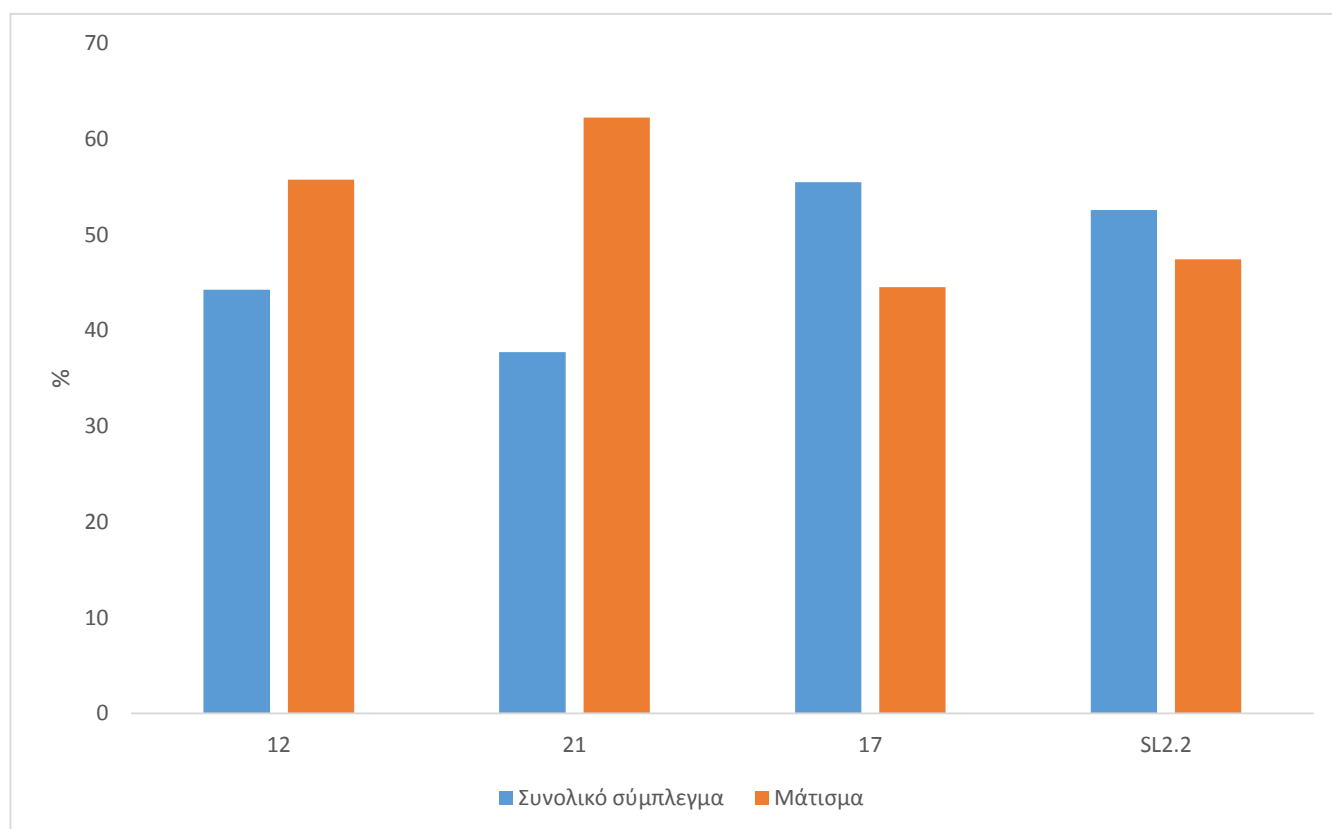


Εικόνα 3.7.

Στο ραβδόγραμμα συγκρίνεται το επίπεδο έκφρασης του συμπλέγματος πεπτιδίου-ιντεΐνης βάση της θερμοκρασίας επώασης. Παρουσιάζεται επίσης το ποσοστό επί της 100 ματισμένου και μη-ματισμένου συμπλέγματος για κάθε τιμή θερμοκρασίας.

3.1.4 Πρωτεϊνικό Μάτισμα

Όπως παρατηρήθηκε στα αποτελέσματα των προαναφερθεισών μελετών, σε κάθε περίπτωση έκφρασης του συμπλέγματος πεπτιδίου-ιντεΐνης εμφανίζεται ένα ποσοστό της μορφής I_C-I_N-CBD, κάτι που υποδυνάμει την πραγματοποίηση της αντίδρασης πρωτεϊνικού ματίσματος. Στο ραβδόγραμμα στην εικόνα 3.8., παρουσιάζονται τα ποσοστά ματισμένης και μη-ματισμένης μορφής και για τις τέσσερις μελετημένες αλληλουχίες πεπτιδίου στόχου. Όπως παρατηρείται για την αλληλουχία 21 παρουσιάζει την μεγαλύτερη άνεση στην πραγματοποίηση της αντίδρασης με ποσοστό ματισμένης μορφής να ξεπερνάει το 60% του συνολικού προϊόντος. Για την αλληλουχία του πεπτιδίου 17, που εμφανίζει το μικρότερο ποσοστό ματισμένης μορφής σε σχέση με τις άλλες αλληλουχίες, παρατηρείται η ύπαρξη μεγαλύτερου μέρους μη-ματισμένης μορφής, φαινόμενο που είναι εμφανές και στην αλληλουχία SL2.2.

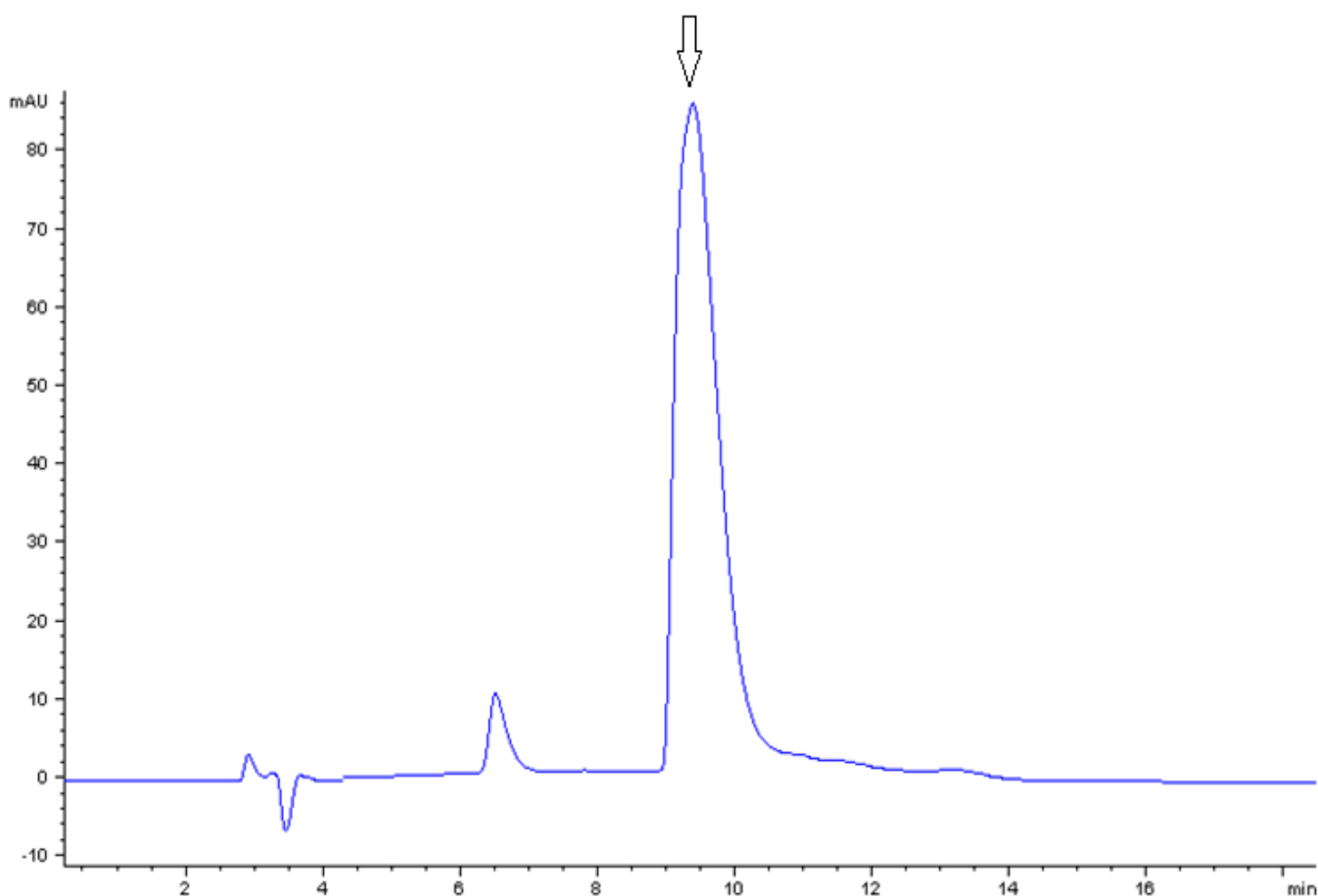


Εικόνα 3.8.

Στο ραβδόγραμμα παρουσιάζεται το ποσοστό επί της 100 ματισμένου και μη-ματισμένου συμπλέγματος για κάθε μελετημένη αλληλουχία πεπτιδίου στόχου. Τα δεδομένα έχουν ληφθεί από πειράματα έκφρασης στις εξής συνθήκες: 0.2% L-Αραβινόζη, 4h επώαση μετά την επαγωγή στους 37 °C

3.2 Μελέτη βιοσύνθεσης κυκλικού πεπτιδίου 21

Ένα σημαντικό στοιχείο που μπορεί να παρατηρηθεί από αυτό το σύστημα είναι η δυνατότητα του να πραγματοποιήσει την αντίδραση κυκλοποίησης *in vivo* αλλά και *in vitro*. Η επιλογή να μελετηθεί η βιοσύνθεση της πεπτιδικής αλληλουχίας 21 πρώτα, έγινε λόγω της δυναμικότητας που εμφάνισε αυτό το πεπτίδιο στη διαδικασία διαλογής. Το πεπτίδιο 21, με βάση τη μελέτη διαλογής, παρουσίασε τη μεγαλύτερη απορρόφηση (λόγω GFP) σε σχέση με τις υπόλοιπες επιλεγμένες αλληλουχίες, που ως επακόλουθο σημαίνει ότι προκαλεί μεγαλύτερη αναστολή της συσσώματωσης της πρωτεΐνης Αβ (βλέπε 1.6). Επίσης, μπορεί παρατηρηθεί από τις προηγούμενες μελέτες ότι το πεπτίδιο 21 εμφάνισε το μεγαλύτερο ποσοστό ματισμένου προϊόντος σε σχέση με



Εικόνα 3.9.

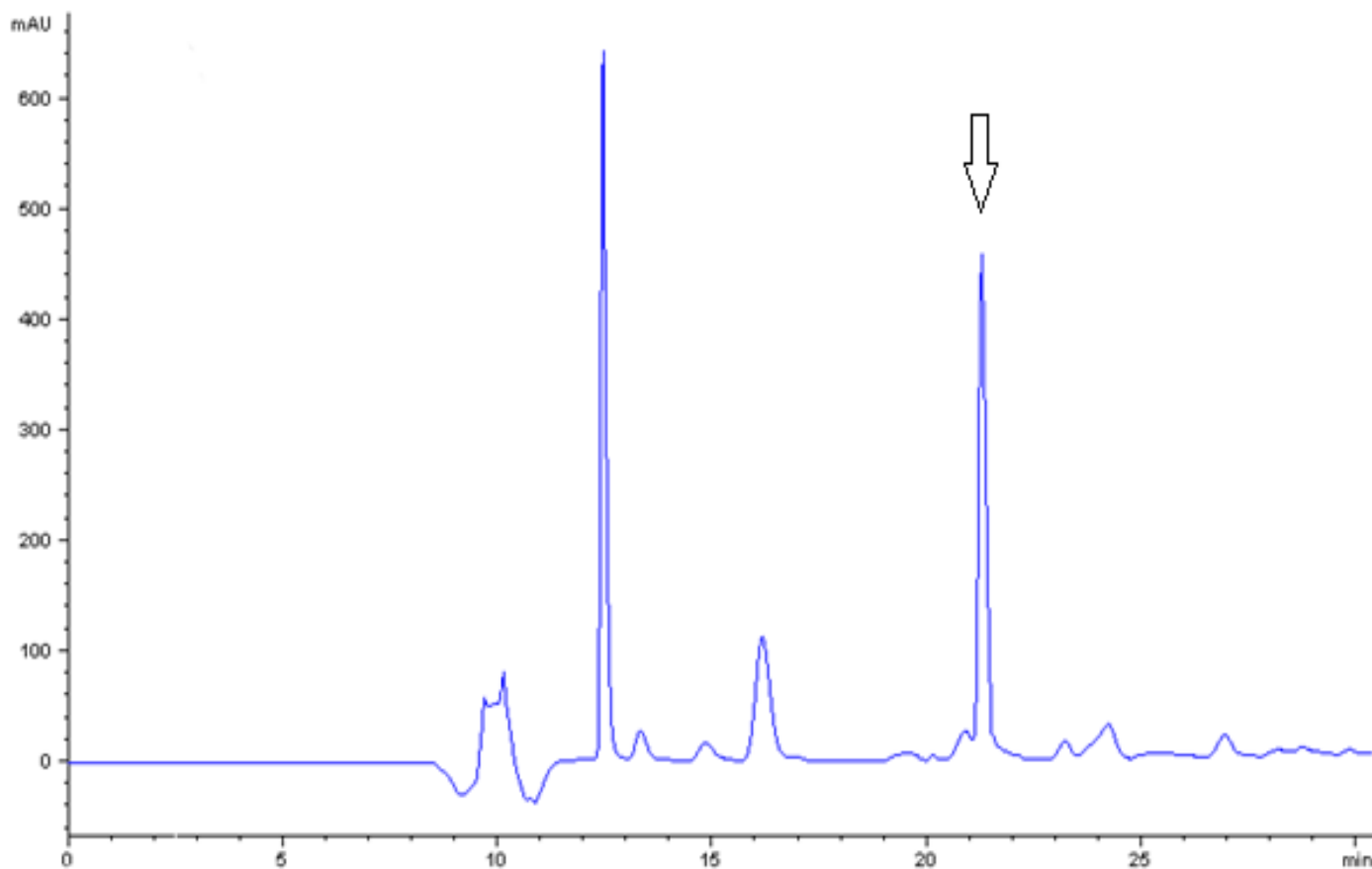
Χρωματογράφημα κυκλικού πεπτιδίου δείκτη. Η κορυφή του πεπτιδίου δείχνεται με το βέλος.

τις άλλες τέσσερις αλληλουχίες, κάτι που μπορεί να οφείλεται στην ύπαρξη ή μη ορισμένων αμινοξέων στην αλληλουχία, δίνοντας ένα προβάδισμα στην επιλογή αυτού του πεπτιδίου. Παρακάτω θα αναφερθούν αποτελέσματα από τις μελέτες για την βιοσύνθεση του πεπτιδίου μέσα στα βακτηριακά κύτταρα και *in vitro* σε στήλη συγγένειας.

3.2.1 *In vivo* παραγωγή

Όπως αναφέρεται στο τμήμα 2.3.1 ο διαχωρισμός του παραγόμενου κυκλικού πεπτιδίου έγινε με στήλη μοριακής διήθησης στο εκχύλισμα της βακτηριακής καλλιέργειας. Σκοπός αυτής της μελέτης ήταν η *in vivo* παραγωγή του κυκλικού πεπτιδίου, που σημαίνει ότι έπρεπε να χρησιμοποιηθούν οι βέλτιστες συνθήκες για τη πραγματοποίηση αντίδραση ματίσματος στο εσωτερικό των κυττάρων. Οπότε για τη μελέτη αυτή, επιλέχθηκαν οι 20° C ως θερμοκρασία επώασης για τη καλλιέργεια των κυττάρων, τιμή που παρουσίασε τη μεγαλύτερη ικανότητα ματίσματος σύμφωνα με τη μελέτη για την εύρεση της βέλτιστης θερμοκρασίας (βλέπε 3.1.3) έναντι της θερμοκρασία 25° C που παρουσίασε το μεγαλύτερο ποσό συνολικού προϊόντος και το δεύτερο μεγαλύτερο ποσοστό σε ματισμένο προϊόν.

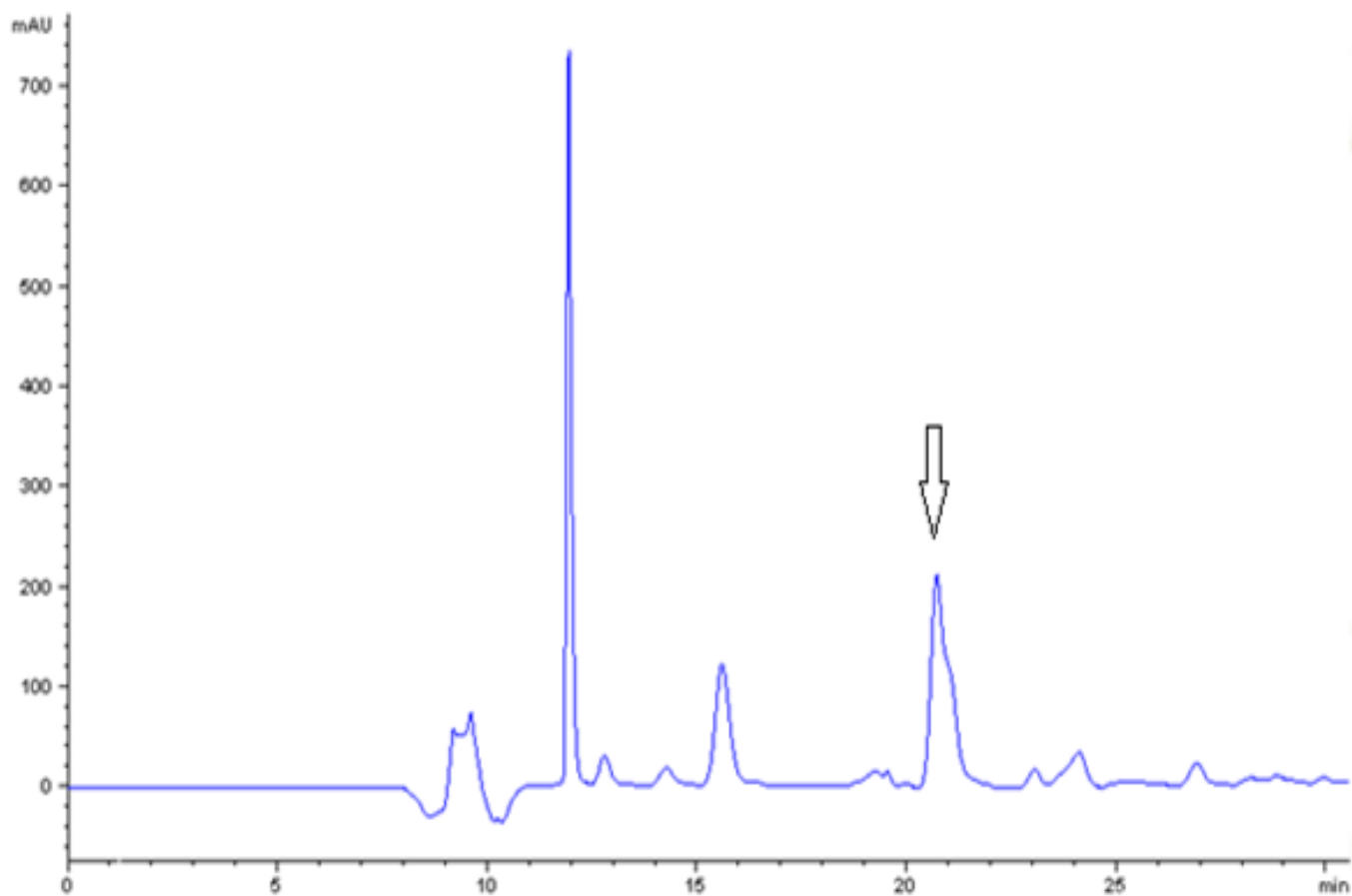
Τα δείγματα που απομονωθήκαν από την μοριακή διήθηση αναλυθήκαν περεταίρω με RP-HPLC. Για την πιστοποίηση της ύπαρξης του κυκλικού πεπτιδίου χρησιμοποιήθηκε ένα οργανικά συντεθειμένο κυκλικό πεπτίδιο για δείκτης. Το χρωματογράφημα του πεπτιδίου δείκτη φαίνεται στην εικόνα 3.9. Η εικόνα 3.10. αποτελεί χρωματογράφημα του εκχυλίσματος από τη μοριακή διήθηση. Όπως ήταν αναμενόμενο στο εκχύλισμα εμφανίζονται επίσης άλλες ενώσεις, ως αποτέλεσμα του μη εξειδικευμένου διαχωρισμού που γίνεται βάση μοριακού βάρους. Η πιστοποίηση της κορυφής του κυκλικού πεπτιδίου 21 έγινε με την πρόσθεση του δείκτη στο εκχυλισμένο δείγμα (spike ανάλυση) που είχε ως αποτέλεσμα την αύξηση του εμβαδού της συγκεκριμένης κορυφής (Εικ.3.11).

**Εικόνα 3.10.**

Χρωματογράφημα εκχυλίσματος από μοριακή διήθηση. Το κυκλικό πεπτίδιο δείχνεται με βέλος.

3.2.2 *In vitro* παραγωγή

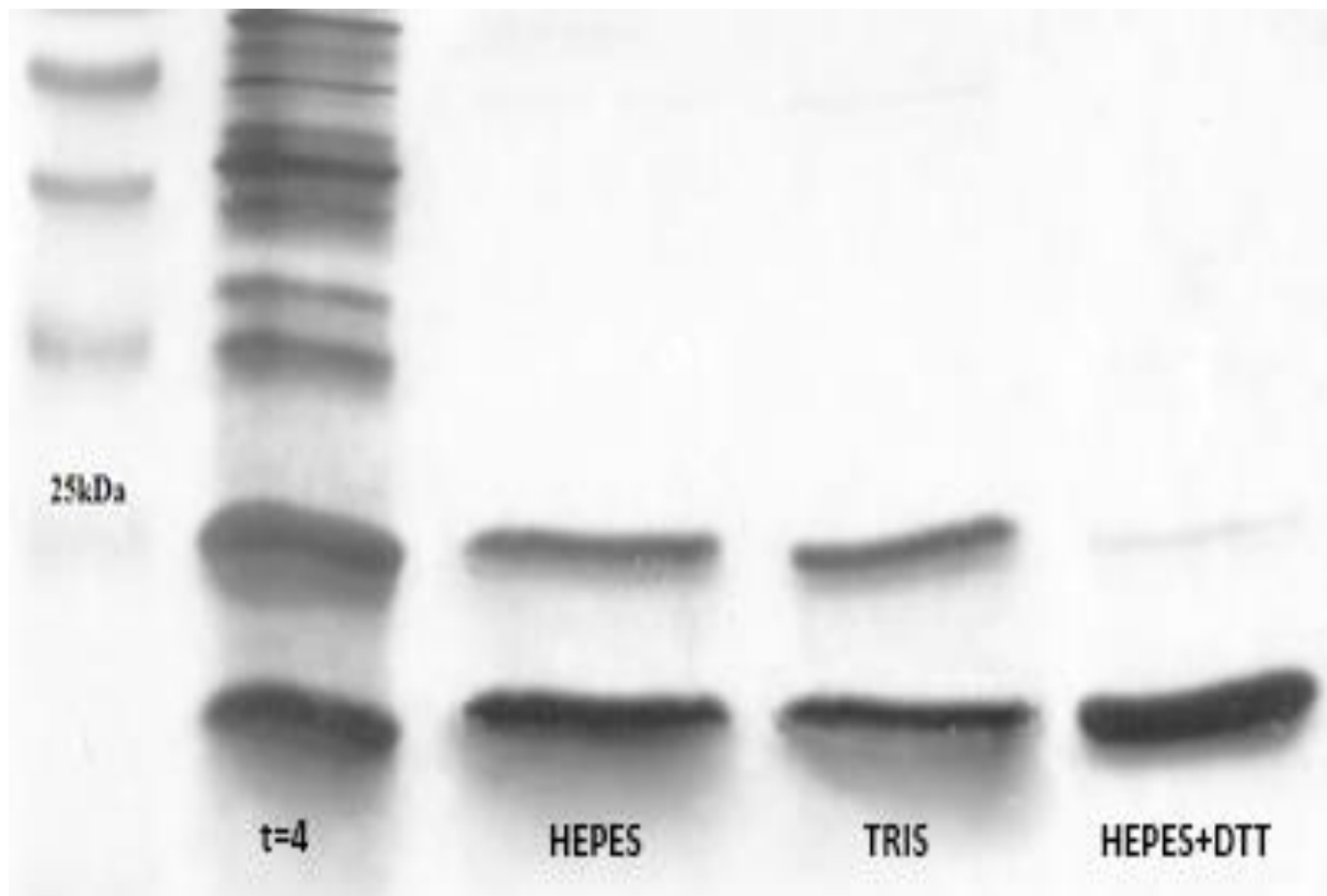
Η μελέτη αυτή είχε ως σκοπό την παρατήρηση της δυνατότητας βιοσύνθεσης του κυκλικού πεπτιδίου εκτός των κυττάρων, αφού απομονωθεί το μη-ματισμένο πρωτεϊνικό σύμπλεγμα με στήλη συγγένειας. Για τον διαχωρισμό του ιντεϊνικού συμπλέγματος, όπως αναφέρεται στο τμήμα 2.3.2, έγινε χρήση στήλης με ρητίνη χητίνης. Για να επιτευχθεί η μέγιστη παραγωγή κυκλικού προϊόντος στη στήλη, έγινε μια μελέτη για τον προσδιορισμό διαλύτη επώασης που προκαλεί το μέγιστο μάτισμα. Στη μελέτη αυτή έγινε χρήση τριών διαλυμάτων, εκ των οποίων το ένα περιείχε



Εικόνα 3.11

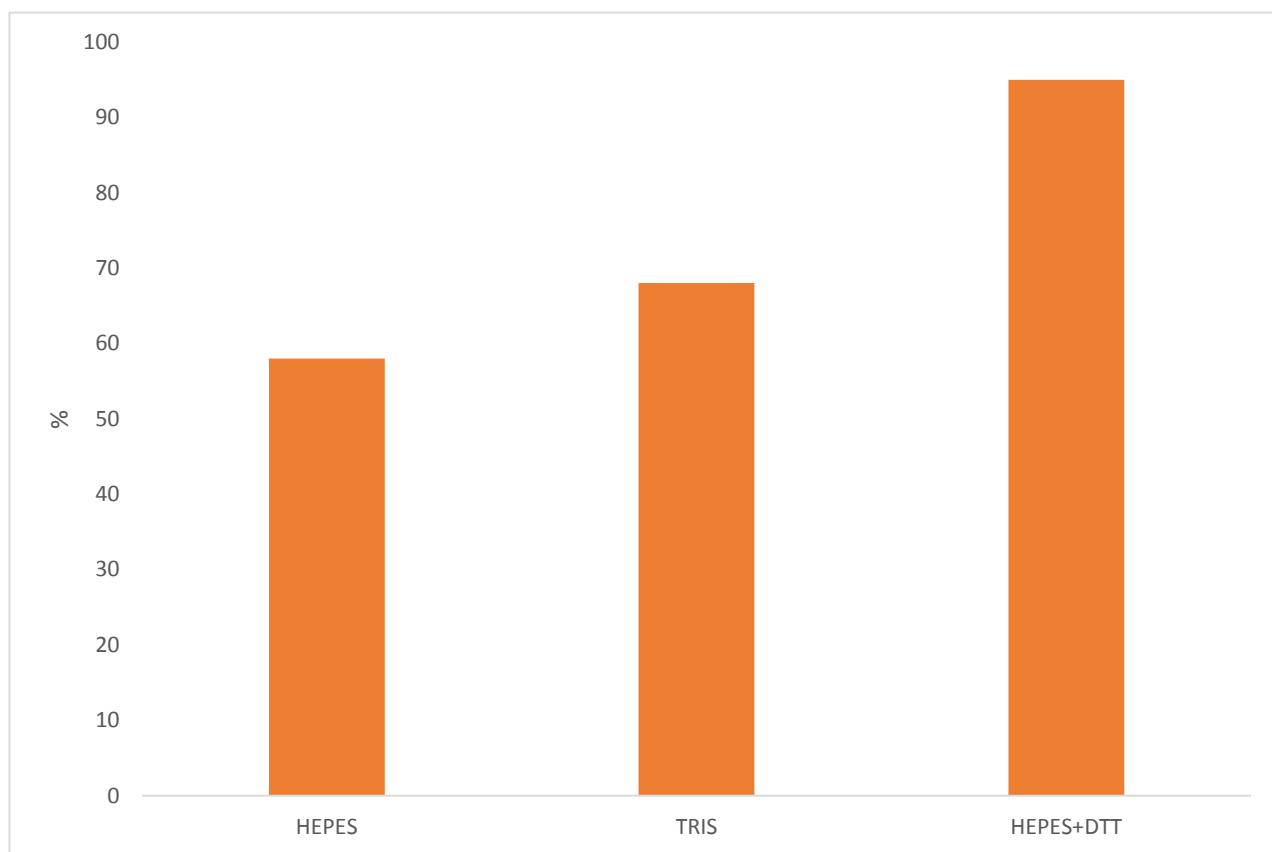
Χρωματογράφημα εκχυλίσματος από μοριακή διήθηση μετά από πρόσθεση του πεπτιδίου δείκτη. Το κυκλικό πεπτίδιο δείχνεται με βέλος.

και μικροποσότητα του αναγωγικού μέσου DTT, για την ενίσχυση του πρώτου βήματος της αντίδρασης ματίσματος (βλέπε Εικ. 1.8). Από πήκτωμα SDS-PAGE των εκχυλισμάτων της στήλης (διάλυμα έκλουσης και 10% SDS), παρατηρήθηκε μεγάλη διαφορά του ματισμένου προϊόντος, μεταξύ του δείγματος στο οποίο έγινε χρήση DTT και των υπόλοιπων (Εικ. 3.12).

**Εικόνα 3.12.**

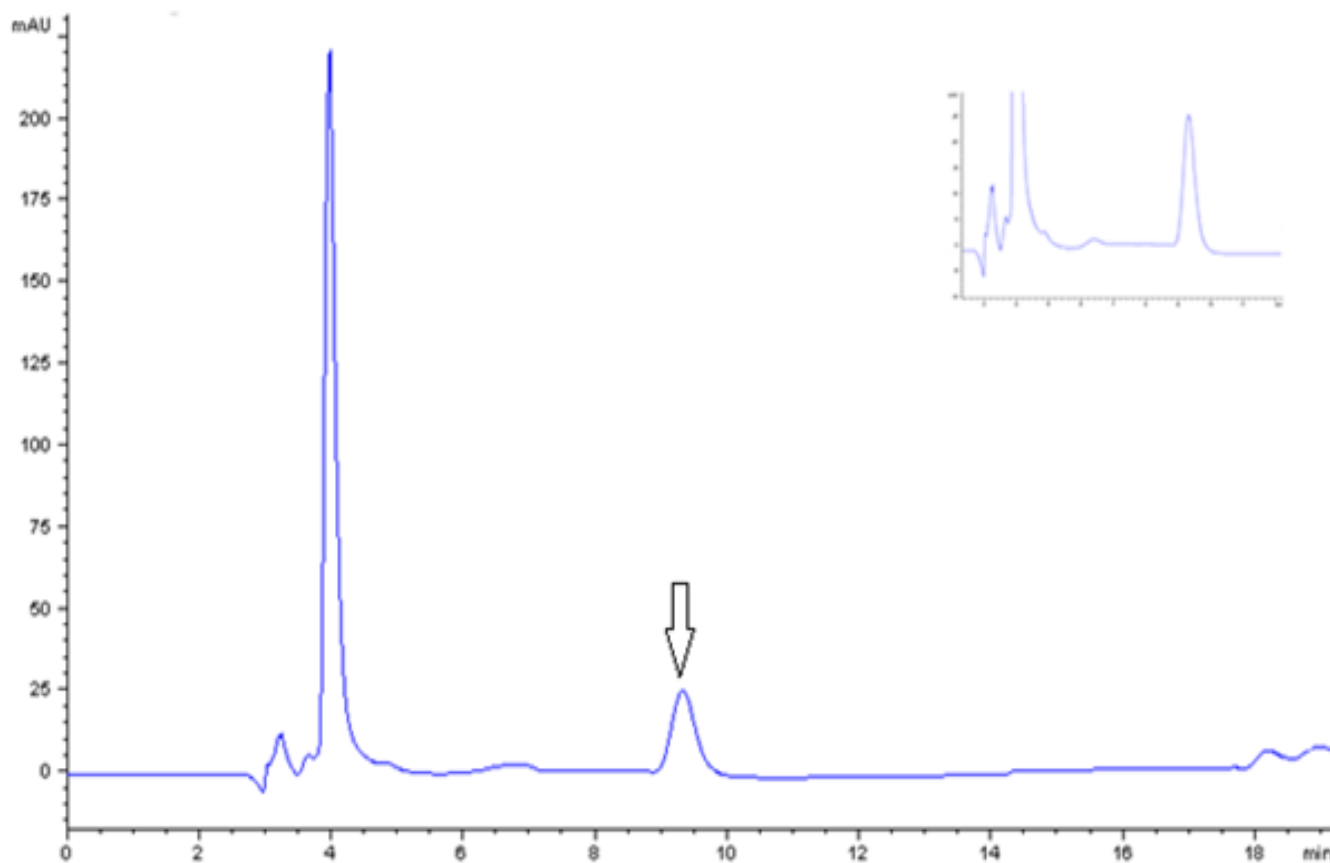
SDS-PAGE δειγμάτων από μελέτη προσδιορισμού βέλτιστου διαλύτη επώασης in vitro ματίσματος.

Η ανάλυση των αποτελεσμάτων με ImageJ έδωσαν μια πιο ξεκάθαρη εικόνα των αποτελεσμάτων. Στο ραβδόγραμμα (Εικ. 3.13) μπορεί εύκολα να παρατηρηθεί ότι η παρουσία του αναγωγικού παράγοντα DTT έχει θετική επίπτωση στην πραγματοποίηση του ματίσματος, εμφανίζοντας προϊόν ματίσματος σε ποσοστό πάνω από 90%, ενώ στη περίπτωση των άλλων δυο διαλυμάτων το ποσοστό του αντίστοιχου προϊόντος είναι κάτω από 70%. Αξιοσημείωτο είναι επίσης, το γεγονός ότι η χρήση του διαλύματος με Tris (pH=7.0) παρουσίασε μεγαλύτερο ποσοστό ματίσματος, διαφορά της τάξεως του 10%, από ότι η χρήση του εναλλακτικού διαλύματος άλατος HEPES ρυθμισμένο σε τιμή pH 8.0.



Εικόνα 3.13.

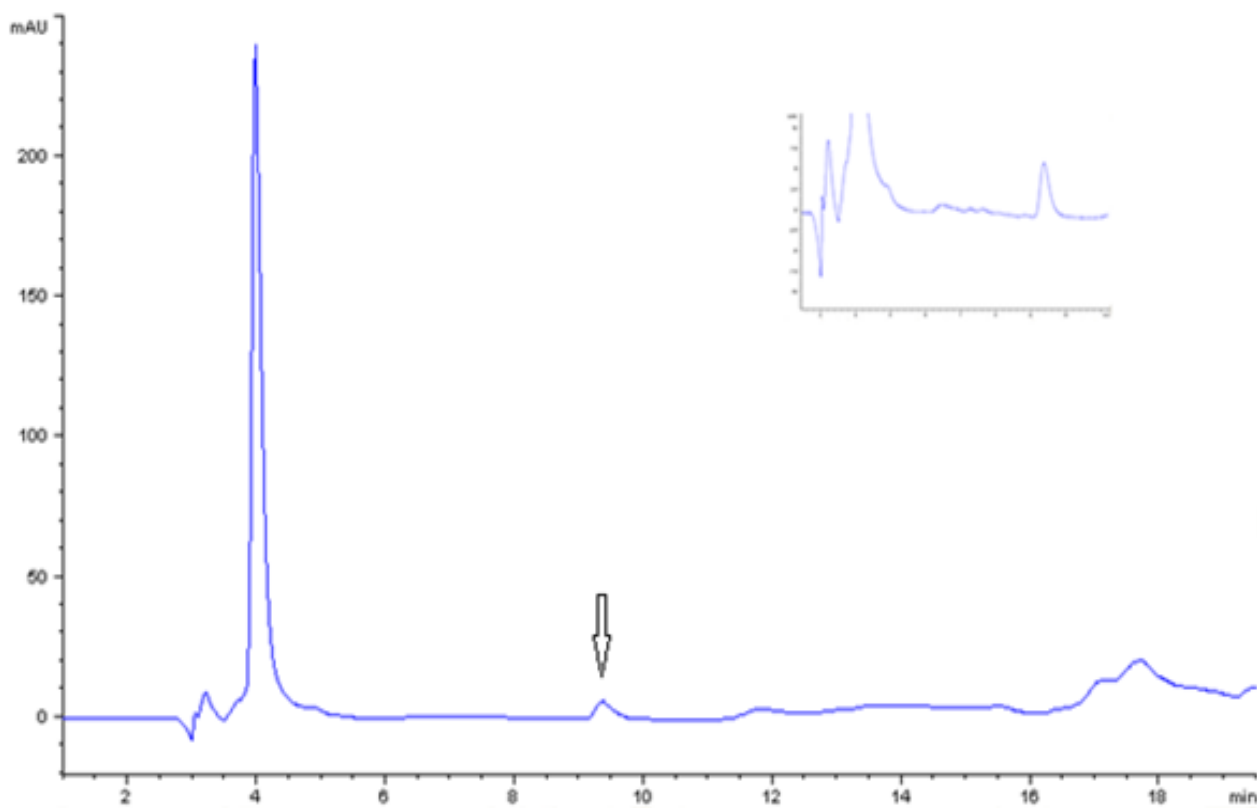
Στο ραβδόγραμμα παρουσιάζεται το ποσοστό επί της 100 ματισμένου προϊόντος μετά από επώαση σε στήλη χητίνης για κάθε ένα από τους τρεις διαλύτες.

**Εικόνα 3.14**

Χρωματογράφημα εκχύλισματος από στήλη συγγένειας . Το κυκλικό πεπτίδιο δείχνεται με βέλος.

Σε στήλη χητίνης που επώσθηκε δείγμα με διάλυμα HEPES 20mM, NaCl 500mM, Triton, 0.1% EDTA 0.1mM, και DTT 30mM pH=8.0, έγινε εκχύλιση με το αντίστοιχο διάλυμα έκλουσης για να απομονωθεί το παραγόμενο κυκλικό πεπτίδιο. Το εκχυλισμένο δείγμα αφού πρώτα συμπυκνώθηκε σε λυοφιλωτή, διαχωρίστηκε με διαδικασία RP-HPLC για την πιστοποίηση της ύπαρξης του κυκλικού πεπτιδίου. Όπως φαίνεται στη εικόνα 3.14, που αποτελεί χρωματογράφημα, με τη χρήση αναλυτικής στήλης, εμφανίζεται μια κορυφή λίγο πριν τα δέκα λεπτά η οποία πιθανότατα να αντιπροσωπεύει το κυκλικό πεπτίδιο 21. Για την σίγουρη πιστοποίηση, διαχωρίστηκε δείγμα στην αναλυτική στήλη, το οποίο περιείχε και ποσότητα του πεπτιδίου δείκτη. Από το χρωματογράφημα στην εικόνα 3.15, παρατηρείται η ύπαρξη μιας κορυφής στον ίδιο χρόνο έκλουσης με το προηγούμενο χρωματογράφημα. Από εδώ υποδεικνύεται

ότι η βιοσύνθεση του κυκλικού πεπτιδίου 21 ήταν επιτυχής *in vitro* όπως αντίστοιχα ήταν και *in vivo*.



Εικόνα 3.15

Χρωματογράφημα εκχυλίσματος από στήλη συγγένειας μετά από πρόσθεση του πεπτιδίου δείκτη. Το κυκλικό πεπτίδιο δείχνεται με βέλος.

4. Συζήτηση-Συμπεράσματα

Ως μια πιο γενική προσέγγιση, πρέπει να γίνει κατανοητό ότι η μελέτη αυτού του θέματος αποτελεί μια προοπτική της συνολικής αλλαγής στην παραγωγική διαδικασία βιοδραστικών ουσιών. Πιο συγκεκριμένα, υπάρχει μεγάλη ανάγκη, στη Βιομηχανία Τροφίμων και ειδικότερα Φαρμάκων, διεργασιών που μπορούν να διατελέσουν την παραγωγή βιομορίων οικονομικότερα και με λιγότερο αντίκτυπο στο περιβάλλον. Τα τελευταία χρόνια υπάρχει ένα ενδιαφέρον στη χρήση μικροοργανισμών ή ενζύμων ως εργαλεία για αυτή τη διαδικασία, με κρίσιμα αποτελέσματα στην παραγωγή σημαντικών για την ανθρωπότητα προϊόντων.

Το σύστημα που μελετήθηκε σε αυτή διατριβή αποτελεί μια καινοτόμα τεχνική που βασίζεται στη βιολογική δράση μιας ομάδας ενζύμων που πρόσφατα ανακαλύφθηκε και έχει προκαλέσει τεράστιο ενδιαφέρον σε ένα μεγάλο κομμάτι της ερευνητικής κοινότητας. Το γεγονός ότι οι ιντεΐνες μπορούν να δημιουργήσουν πεπτιδικό δεσμό, έχει δώσει πολλές βλέψεις για τη χειραγώγηση αυτών των μορίων. Μια από τις πιο σημαντικές διεργασίες για την οποία οι ιντεΐνες έχουν ανασχεδιαστεί, είναι η παραγωγή κυκλικών πεπτιδίων. Ωστόσο, παρά την απεριόριστη γνώση που υπάρχει αλλά και το γεγονός των πολλών παραδειγμάτων χειρισμού και ανασυνδυασμού ιντεϊνών, υπάρχουν πολλά αναπάντητα ερωτήματα. Σε αυτή την ενότητα, σκοπός είναι να απαντηθούν ορισμένα από αυτά τα ερωτήματα, μέσα από τα αποτελέσματα των διαφόρων πειραμάτων που έγιναν για αυτή τη διπλωματική εργασία σε συνδυασμό με την υπάρχουσα βιβλιογραφία. Επίσης θα αναφερθούν και προβληματισμοί, οι οποίοι έχουν να κάνουν τόσο με το βιολογικό μηχανισμό των ιντεϊνών όσο και με τη τεχνολογία που είναι βασισμένη σε αυτά τα βιομόρια και ήδη χρησιμοποιείται.

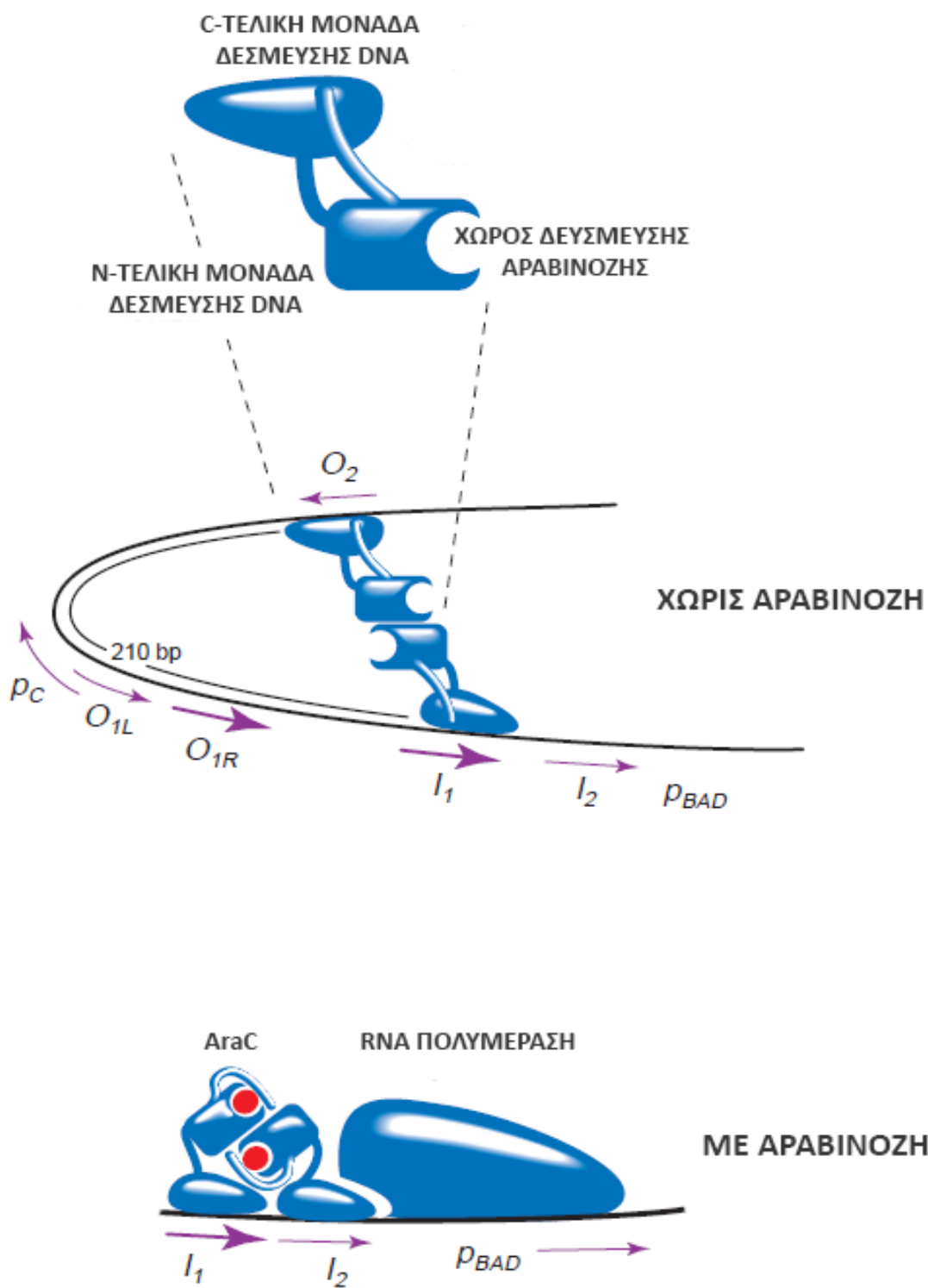
Όπως έχει ήδη παρατηρηθεί, οι μελέτες σε αυτή την πειραματική εργασία έχουν χωριστεί σε δυο μεγαλύτερα τμήματα, την παραγωγή του συνολικού συστήματος και την παραγωγή του τελικού προϊόντος, δηλαδή το κυκλικό ολιγοπεπτίδιο. Είναι προφανές όμως ότι αποτελέσματα από ορισμένες μελέτες αποτελούν δεδομένα για την εξαγωγή πειραμάτων για επόμενες. Για αυτό το λόγο οι μελέτες έχουν περιγραφεί με συγκεκριμένη σειρά, ώστε να εμφανίζεται μια ιεραρχία στην απόδοση των αποτελεσμάτων.

Η ανάλυση και συζήτηση των αποτελεσμάτων ακολουθεί την ίδια σειρά που παρατηρείται στην ενότητα 3 (βλέπε Εν. 3 Αποτελέσματα), πρώτα θα περιγραφούν τα αποτελέσματα των διαφόρων μελετών που είχαν να κάνουν με την έκφραση του συνολικού συστήματος ιντεϊνης-πεπτιδίου, ενώ στη συνέχεια θα αναλυθούν τα αποτελέσματα βιοσύνθεσης του κυκλικού πεπτιδίου 21.

4.1 Μελέτη συνθηκών υπερέκφρασης του πρωτεϊνικού συμπλέγματος πεπτιδίου-ιντεΐνης

Για τη διαδικασία αυτή, επιδιώχθηκε ο προσδιορισμός των ιδανικότερων συνθηκών παραγωγής του ιντεϊνικού συμπλέγματος από τα βακτηριακά κύτταρα. Ο διαχωρισμός με τη μετέπειτα μελέτη (βιοσύνθεση κυκλικού ολιγοπεπτιδίου) έγινε για τον ευκολότερο προσδιορισμό των παραγόντων που παίζουν ρόλο στην παραγωγή του πρόδρομου βιομορίου (πεπτίδιο-ιντεΐνη). Οπότε είναι λογικό ότι ο πρωταρχικός σκοπός σε αυτό το βήμα ήταν η μέγιστη παραγωγή του ιντεϊνικού συμπλέγματος, θεωρώντας ότι η ύπαρξη μεγάλης ποσότητας του συμπλέγματος θα έχει ως κατάληξη τη βιοσύνθεση αρκετά μεγάλης ποσότητας κυκλικού πεπτιδίου. Η μελέτη αυτή χωρίστηκε σε τρία τμήματα, όπου μελετηθήκαν τρεις διαφορετικοί παράγοντες που επηρεάζουν την υπερέκφραση της πρωτεΐνης, τη συγκέντρωση του επαγωγέα, το χρόνο και τη θερμοκρασία επώασης. Ως μια πρώτη ανάλυση θα συζητηθούν τα αποτελέσματα του προσδιορισμού της βέλτιστης συγκέντρωσης επαγωγέα σε συνδυασμό με το χρόνο επώασης.

Το σύστημα έκφρασης που είχε επιλεγεί σε αυτή την πειραματική διαδικασία ήταν ο υποκινητής *ara* από το οπερόνιο της L-Αραβινόζης (Schleif, 2000). Προτού όμως αναλυθούν τα αποτελέσματα, είναι σημαντικό να γίνει μια μικρή περιγραφή για το λόγο που επιλέχθηκε αυτό το σύστημα. Είναι κατανοητό ότι αυτό το σύστημα έχει τις αντίστοιχες δυνατότητες με το πιο συχνά μεταχειρισμένο σύστημα για ετερόλογη έκφραση πρωτεϊνών σε βακτηριακά κύτταρα *E.coli*, που χρησιμοποιεί τον υποκινητή *lac* του οπερονίου της λακτόζης. Η επιλογή του συστήματος αυτού δεν έχει να κάνει μόνο με τα επίπεδα υπερέκφρασης μετά από επαγωγή με L-Αραβινόζη, αλλά περισσότερο με το σκοπό της χρήσεως του συνολικού συστήματος. Στο επίπεδο της βιομηχανικής παραγωγής προτιμάται η χρήση μη τοξικών ουσιών αλλά και πιο σημαντικά οικονομικών, λόγοι για τους οποίους η επιλογή της L-Αραβινόζης προτείνεται έναντι του IPTG, επαγωγέα του υποκινητή *lac*. Ο υποκινητής *ara* σε συνδυασμό με το στέλεχος *E.coli* BL21 DE3 που παρουσιάζει τη δυνατότητα μέγιστης έκφρασης, που οφείλεται στην ύπαρξη του γονιδίου της T7 RNA πολυμεράσης (Studier and Moffatt, 1986), αποτελεί ένα αποδοτικό μικροβιακό εργοστάσιο.



Εικόνα 4.1

Ρύθμιση της έκφρασης από τη L-Αραβινόζη στο οπερόνιο αραβινόζης (Schleif, 2000).

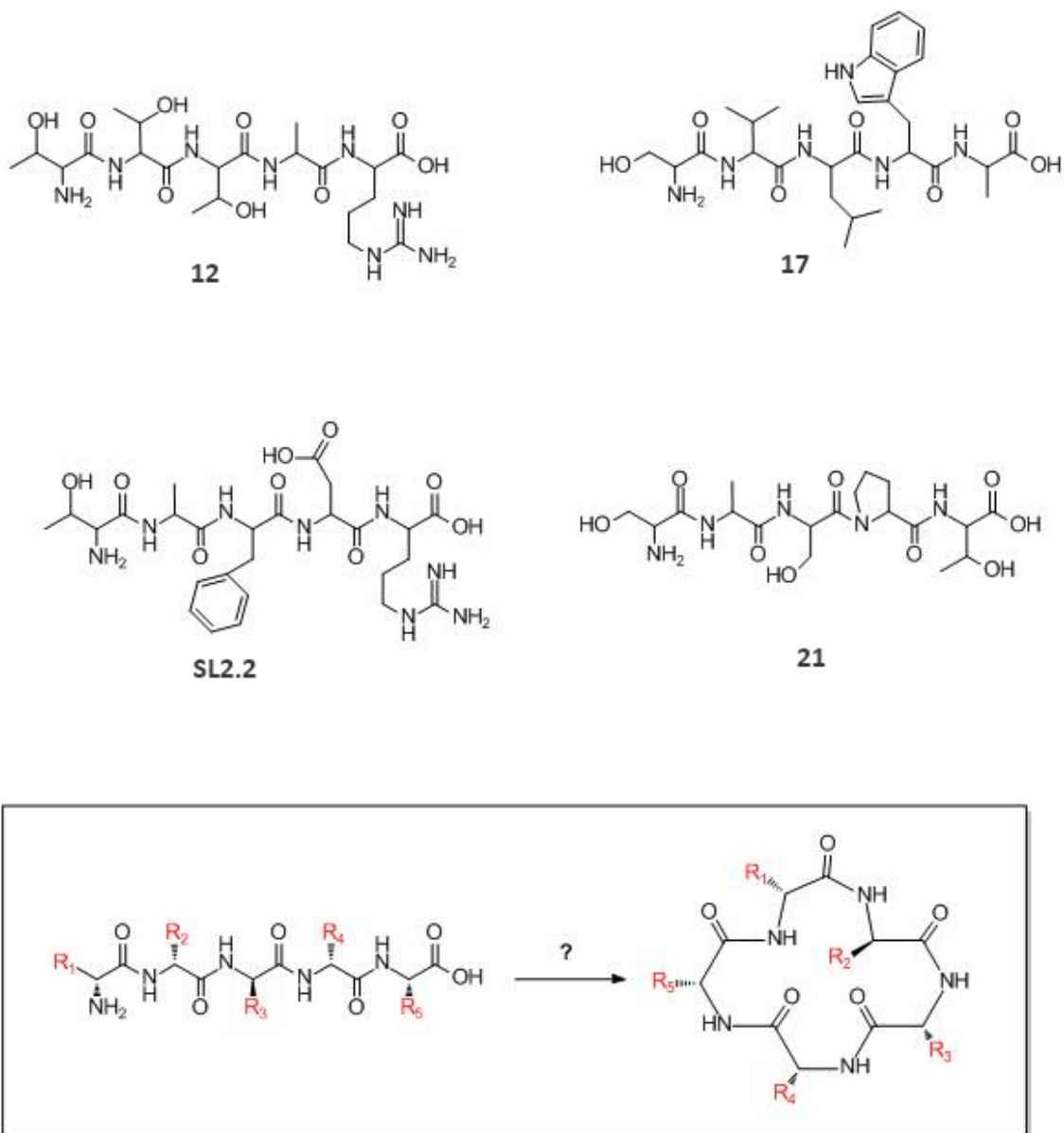
Με την πρόσθεση L-Αραβινόζης η παραμποδιστική δράση του AraC αναστέλεται και επιτρέπει στην RNA πολυμεράση να προσδεθεί στο τμήμα DNA (σύστημα θετικής ρύθμισης) (Εικ. 4.1). Οπότε με τη χρήση του επαγωγέα εμφανίζεται άμεση έκφραση της πρωτεΐνης και λογικά όσο μεγαλύτερη η συγκέντρωση τόσο μεγαλύτερη και η έκφραση. Παρά το γεγονός ότι σε χρόνο 24 ωρών παρατηρήθηκε η μέγιστη παραγωγή πρωτεΐνης με συγκέντρωση 0.5%, για χρόνους 3 και 4 ώρες παρατηρείται μεγαλύτερη έκφραση για συγκέντρωση 0.2% (βλέπε 3.1.1 Εικ.3.3). Αυτό μπορεί να οφείλεται στο κορεσμό του συστήματος κατά τη μεταφορά της L-Αραβινόζης (στη περίπτωση συγκέντρωσης 0.5%) στη διμερική πρωτεΐνη AraC. Η διαφορά αυτή στο χρόνο παραγωγής στις δυο συγκεντρώσεις, αποτέλεσε σημαντικό στοιχείο για την εξαγωγή των επόμενων μελετών. Πιο συγκεκριμένα για την μέγιστη παραγωγή πρωτεϊνικού προϊόντος έγινε χρήση της συγκέντρωσης 0.5% για χρόνο επώασης 24 h, ενώ για την άμεση παραγωγή του συμπλέγματος σε χρόνους κάτω από 5 h χρησιμοποιήθηκε η συγκέντρωση 0.2%

Είναι σημαντικό να αναφερθεί, ότι η συνολική αυτή μελέτη είχε ως γενικότερο σκοπό την εύρεση των συνθηκών για την αποδοτικότερη παραγωγή. Όπως αναφέρεται παραπάνω τα δεδομένα αυτής της μελέτης αποτελούν στοιχεία για την εξαγωγή επόμενων πειραμάτων, για παράδειγμα για την επίτευξη της *in vitro* βιοσύνθεσης έπρεπε να γίνει η μέγιστη έκφραση του ιντεϊνικού συμπλέγματος σε συνθήκες που να προκαλείται όσο το δυνατό λιγότερο πρωτεϊνικό μάτισμα. Ο πιο σημαντικός παράγοντας για το ζητούμενο αυτό ήταν ο χρόνος. Με δεδομένο από τη προηγούμενη μελέτη ότι η συγκέντρωση επαγωγέα 0.2% προκαλεί τη μέγιστη έκφραση σε χρόνους επώασης κάτω από 5 ώρες, συγκρίθηκε το ποσοστό ματισμένου προϊόντος για χρόνους 1-5 ώρες μετά την επαγωγή (βλέπε 3.1.2). Με ζητούμενα τη μικρή εμφάνιση αντίδρασης ματίσματος αλλά και τη μέγιστη έκφραση, παρατηρηθήκαν δυο τιμές που εμφανίζουν ικανοποιητικά αποτελέσματα. Οι διαφορές στους χρόνους επώασης 3 και 4 ωρών, τόσο στο ποσοστό μη ματισμένου προϊόντος όσο και συνολικής έκφρασης, είναι κάτω από 3% κάνοντας δύσκολη την επιλογή μεταξύ των δυο τιμών (Εικ.3.5). Ωστόσο το γεγονός ότι η διαδικασία μπορεί να γίνει πιο αποδοτικά για χρόνο επώασης 3 ωρών κρίθηκε στην ύπαρξη μεγαλύτερου ποσοστού μη-ματισμένης μορφής, αλλά και στη δυνατότητα να εκπληρωθεί της διαδικασίας σε μικρότερο χρόνο.

Αντίστοιχα, όπως έγινε μελέτη για την αποδοτικότερη παραγωγή, μελετήθηκε και η δυνατότητα μέγιστης παραγωγής για χρόνο 24 h μετά την επαγωγή με L-Αραβινόζη σε συγκέντρωση 0.5%.

Ως πρώτο συμπέρασμα για την επώαση στις 24 ώρες, είναι το γεγονός της εμφάνισης έντονων ζωνών, στα πηκτώματα SDS-PAGE, της ματισμένης μορφής. Με δεδομένο αυτό και ζητούμενο τη μέγιστη έκφραση και εμφάνιση του προϊόντος ματίσματος, έγινε σύγκριση διαφορετικών θερμοκρασιών επώασης (βλέπε 3.1.3). Τα αποτελέσματα αυτά αποτέλεσαν και στοιχεία για την εξαγωγή του πειράματος *in vivo* βοισύνθεσης. Η επιλογή των θερμοκρασιών, έγινε βάση των βιβλιογραφικών τιμών λειτουργίας του πρωτεϊνικού ματίσματος στη Ssp DnaE ιντεΐνη αλλά και της θερμοκρασίας καλλιέργειας του οργανισμού προέλευσης. Έχει αναφερθεί ότι η αντίδραση ματίσματος λειτουργεί βέλτιστα σε θερμοκρασίες 4-25 °C (Xu and Evans, 2001), ενώ η ιδανική θερμοκρασία ανάπτυξης του *Synechocystis* sp. PCC 6803 έχει παρατηρηθεί στους 30 °C (Yu et al., 2013). Με βάση αυτά τα δεδομένα επιλέχθηκαν τρεις θερμοκρασίες (5, 20 και 25 °C) στο εύρος τιμών του ιδανικού πρωτεϊνικού ματίσματος και η αναφερόμενη θερμοκρασία ανάπτυξης του *Synechocystis* sp. PCC 6803, ενώ μελετήθηκε επίσης η δυνατότητα πρωτεϊνικού ματίσματος στη βέλτιστη θερμοκρασία ανάπτυξης του βακτηρίου *E.coli* (37 °C).

Όπως παρατηρήθηκε στη μελέτη στους 20 και 25 °C παρατηρήθηκε η μέγιστη έκφραση του συμπλέγματος αλλά και παραγωγή ματισμένου προϊόντος, όπως αντίστοιχα είχε αναφερθεί βιβλιογραφικά (βλέπε Εικ. 3.7). Το γεγονός ότι στους 37 °C παρατηρείται η μικρότερη έκφραση, πιθανώς οφείλεται στην γρήγορη ανάπτυξη των κυττάρων, που οδηγεί σε πιο γρήγορη εμφάνιση της φάσης θανάτου όπου εμφανίζονται οι αυτό-πρωτεολυτικοί μηχανισμοί. Στους 30 °C παρατηρήθηκε το μικρότερο ποσοστό εμφάνισης του πρωτεϊνικού ματίσματος, κάτι που δείχνει ότι ο μηχανισμός δεν εξαρτάται από τη θερμοκρασία ανάπτυξης του οργανισμού προέλευσης. Το φαινόμενο αυτό μπορεί να εξηγηθεί από πολλές σκοπιές, όπως για παράδειγμα ότι η θερμοκρασία αλλά και η οξειδοαναγωγική κατάσταση μπορεί να μην επηρεάζουν το ουσιαστικό κομμάτι του χημικού μηχανισμού του ματίσματος αλλά την αναδίπλωση της ιντεΐνης, αποτρέποντας την έναρξη της αντίδρασης (Eryilmaz et al., 2014). Αυτή απορία θα μπορούσε να απαντηθεί από ένα βασικό ερώτημα που πηγάζει από την εξελικτική βιολογία «Ποιος είναι ο βιολογικός ρόλος των ιντεϊνών στους οργανισμούς;». Το πρωτεϊνικό μάτισμα από τις ιντεΐνες πιθανότατα αποτελεί βιολογικό μηχανισμό ως απόκριση σε στρες, που σημαίνει ότι δεν θα έπρεπε να λειτουργούν στις ιδανικές συνθήκες του οργανισμού. Επειδή όμως, η δραστηριότητα ελάχιστων ιντεϊνών έχει μελετηθεί στους ενδογενείς τους οργανισμούς, είναι δύσκολο να αντιληφθούμε τους φυσικούς παράγοντες που επηρεάζουν τη δράση τους.



Εικόνα 4.2

Χημικές δομές πεπτιδικών αλληλουχιών.

Ένας από τους σημαντικότερους παράγοντες επίδρασης της αντίδρασης πρωτεϊνικού ματίσματος των ιντεϊνών, είναι η αλληλουχία της εξτεΐνης. Έχουν αναφερθεί πολλές περιπτώσεις μη επιτυχημένου ματίσματος τόσο στη Ssp DnaE ιντεΐνη όσο και σε άλλες ιντεΐνες της κατηγορίας 1 που μπορεί να οφείλεται στην αλληλουχία της εξτεΐνης (βλέπε 1.3.2). Η ύπαρξη αμινοξέων με ογκώδης πλευρικές ομάδες μπορεί να λειτουργήσει ανασταλτικά στη δράση της ιντεΐνης, με αποτέλεσμα το πρωτεϊνικό μάτισμα να πραγματοποιείται με πιο αργούς ρυθμούς ή ακόμα και αναστέλλεται ο συνολικός μηχανισμός. Ένα παράδειγμα αυτής του μπλοκαρίσματος του μηχανισμού, είναι η δημιουργία δισουλφιδικού δεσμού μεταξύ της Cys του ενεργού κέντρου της ιντεΐνης και μιας Cys στην εξτεϊνική αλληλουχία (Callahan et al., 2011).

Από τα αποτελέσματα των πειραμάτων για τη μελέτη των συνθηκών, έγινε ανάλυση για την εξαγωγή των ποσοστών ματισμένου προϊόντος για κάθε μια από τις τέσσερις αλληλουχίες πεπτιδίων που μελετήθηκαν (βλέπε 3.1.4). Σκοπός ήταν να συγκριθεί η δυνατότητα ματίσματος *in vivo* για κάθε περίπτωση διαφορετικής εξτεϊνικής αλληλουχίας σε χρόνο 4 ωρών. Η μικρότερη εμφάνιση ματισμένου προϊόντος εμφανίστηκε στις αλληλουχίες 17 και SL2.2, με ποσοστά κάτω από 50%. Αντίθετα η πεπτιδική αλληλουχία 21 παρουσίασε το μέγιστο μάτισμα από όλα τα πεπτίδια με ποσοστό που ξεπερνούσε το 60%. Τέλος στο πεπτίδιο 12 παρατηρήθηκε η δεύτερη μεγαλύτερη τιμή της ματισμένης μορφής που σε αντίθεση με τα 17 και SL2.2 ξεπέρασε το 50%. Για να κατανοηθούν καλύτερα αυτά τα αποτελέσματα πρέπει να παρατηρηθούν τα αμινοξέα που συστήνουν αυτές τις πεπτιδικές αλληλουχίες (Εικ. 4.2). Τα πεπτίδια 17 και SL2.2 παρατηρούνται από την ύπαρξη ενός αμινοξέος με μια ογκώδη πλευρική ομάδα, Trp και Phe αντίστοιχα, που μπορεί να προκαλέσει δυσκολίες στην αναδίπλωση της ιντεΐνης (βλέπε Εικ. 1.12A). Πιο ανησυχητική είναι όμως η παρουσία της θετικά φορτισμένης Arg στην -1 θέση των πεπτιδίων 17, SL2.2 και 12, η οποία μπορεί να προκαλεί πρόβλημα εμποδίζοντας την αναδίπλωση της ιντεΐνης ή παρεμποδίζοντας το πρώτο βήμα του μηχανισμού (N-σε-S άκυλο-μετατόπιση) αλληλεπιδρώντας με κάποιο από το αμινοξέα της περιοχής B ή F που παίζουν ρόλο στο μηχανισμό (βλέπε 3.1.4 Εικ.1.11). Ωστόσο, μπορεί να παρατηρηθεί ότι η ύπαρξη αυτών των αμινοξέων στις εξτεϊνικές αλληλουχίες δεν προκαλεί την πλήρη αναστολή του πρωτεϊνικού ματίσματος, αλλά την επιβράδυνση του. Από την άλλη το πεπτίδιο 21 δεν εμφανίζει αμινοξέα με ογκώδεις πλευρικές ομάδες στην αλληλουχία του, αλλά το άκρως αντίθετο αμινοξέα μικρού μεγέθους και μη φορτισμένα. Το πιο σημαντικό σε αυτή την πεπτιδική αλληλουχία είναι η ύπαρξη μιας Pro στη θέση -2, η οποία μπορεί να ενισχύσει τη διαδικασία κυκλοποίησης χάρη στη ελαστικότητα της.

4.2 Μελέτη βιοσύνθεσης κυκλικού πεπτιδίου 21

Όλη η προηγούμενη διαδικασία πειραματικών μελετών αποτέλεσε δίοδο για τη καλύτερη κατανόηση του συστήματος ώστε να επιτευχθεί ο ουσιαστικός στόχος αυτής της διπλωματικής εργασίας, η βιοσύνθεση κυκλικού ολιγοπεπτιδίου από βακτηριακά κύτταρα. Για αυτό το κομμάτι επιλέχθηκε η πεπτιδική αλληλουχία 21, όπως αναφέρθηκε ήδη προηγουμένως (βλέπε 3.2) λόγω της διαλογής βάση μελετών με την Αβ πρωτεΐνη (βλέπε 1.6 Στόχος) και της εμφάνισης ικανοποιητικού ματίσματος, για αναπαραγωγή της διεργασίας κυκλοποίησης μέσω πρωτεϊνικού ματίσματος. Η μελέτη αυτή χωρίστηκε σε δυο τμήματα, στη *in vivo* και *in vitro* παραγωγή, λόγω των διαφορετικών ζητούμενων σε κάθε περίπτωση.

Γνωρίζοντας ότι η ιντεΐνη μπορεί να πραγματοποιήσει την αντίδραση ματίσματος εκτός κυττάρου (*in vitro*), επιδιώχτηκε ο εκλεκτικός διαχωρισμός του συνολικού συμπλέγματος με στήλη συγγένειας, ώστε να πραγματοποιηθεί η κυκλοποίηση μέσα στη στήλη και μετά από εκχύλιση να συλλεχθεί μόνο το κυκλικό πεπτίδιο. Αυτή η διαδικασία παρουσίασε εκλεκτικό διαχωρισμό του τελικού κυκλικού προϊόντος, όπως μπορεί να παρατηρηθεί από το χρωματογράφημα της RP-HPLC στην οποία εμφανίζεται μόνο μια κορυφή (βλέπε 3.14). Η ρητίνη που χρησιμοποιήθηκε για το πακετάρισμα της στήλης αποτελούταν από πολύ μικρά σφαιρίδια χητίνης. Η συγγένεια της CBD μονάδας με τα μόρια χητίνης, αποτελεί τη βάση αυτής της τεχνικής διαχωρισμού. Παρά όμως τη δυνατότητα του εκλεκτικού διαχωρισμού, εμφανιστήκαν διαφορά προβλήματα στη διαδικασία, όπως για παράδειγμα το γεγονός ότι η ρητίνη έχει δυνατότητα χωρητικότητας 2.0 mg πρωτεΐνης ανά ml χητίνης. Αυτό αποτέλεσε κύριο πρόβλημα στη παραγωγική διαδικασία, αφού το κόστος για την απομόνωση ιντεϊνικού συμπλέγματος από μια μεγάλη καλλιέργεια ήταν μεγάλο. Επίσης είναι πιθανή η ύπαρξη μονάδων στο κυτταρικό εκχύλισμα που έχουν συγγένεια με τη χητίνη, μειώνοντας έτσι την εκλεκτικότητα του διαχωρισμού αυτής της μεθόδου.

Ήταν απαραίτητο για αυτή την εργασία να βρεθεί μεθοδολογία για τη μαζική παραγωγή χωρίς το κόστος της διαδικασίας να είναι ασύμφορο. Για να αντιμετωπιστούν τα προβλήματα της προηγούμενης μεθοδολογίας, προτάθηκε να πραγματοποιηθεί η κυκλοποίηση στο εσωτερικό των κυττάρων και στη συνέχεια να διαχωριστούν τα μικρά μόρια με τα μεγάλα σε στήλη μοριακής διήθησης. Σκοπός ήταν η αύξηση της συγκέντρωσης του κυκλικού πεπτιδίου στο δείγμα, ώστε στη συνέχεια να προσδιοριστεί και να διαχωριστεί σε προπαρασκευαστική στήλη RP-HPLC. Από αυτή τη μελέτη προσδιορίστηκε μεγάλη ποσότητα του κυκλικού πεπτιδίου, με μόνο μειονέκτημα

την ύπαρξη και άλλων μικρών μορίων. Η διαδικασία αυτή έναντι του εκλεκτικού διαχωρισμού παρουσιάζει μεγαλύτερη αποδοτικότητα, τόσο σε επίπεδο ποσότητας παραγόμενου προϊόντος όσο και σε επίπεδο κόστους, αν θεωρηθεί ότι η χρωματογραφία (RP-HPLC) για την απομόνωση καθαρού προϊόντος είναι υποχρεωτική και στις δυο περιπτώσεις.

Η αντικατάσταση της διαδικασίας διαχωρισμού που χρησιμοποιεί τη συγγένεια σε χητίνη με μια παρόμοια μεθοδολογία που χρησιμοποιεί την συγγένεια μεταξύ His και νικελίου θα μπορούσε να δώσει καλύτερα αποτελέσματα στην *in vitro* παραγωγή κυκλικών ολιγοπεπτιδίων. Ωστόσο, τα αποτελέσματα από τη *in vivo* διαδικασία είναι δύσκολα ανταγωνίσιμα στο επίπεδο μαζικής παραγωγής, όμως έχει σημασία να αναφερθεί ότι σε βιομηχανική κλίμακα θα μπορούσαν να εφαρμοστούν και οι δυο διεργασίες συγχρόνως για τη μείωση του χρόνου παραγωγής. Από αυτή την εργασία μπορεί ευκολά να γίνει κατανοητό ότι η βιομηχανική παραγωγή κυκλικών πεπτιδίων, σύντομα θα πάψει να είναι αποτέλεσμα της Οργανικής Σύνθεσης αλλά νέων μεθοδολογιών όπως είναι το σύστημα SCICCLOPS. Όμως για να επιτευχθεί αυτή η αντικατάσταση πρέπει να διαλευκανθούν ζητήματα, όπως ο συνολικός μηχανισμός πρωτεϊνικού ματίσματος και επίδραση της εξτεΐνης στη δράση των ιντεϊνών.

5. Περίληψη

Η φαρμακευτική χρήση κυκλικών ολιγοπεπτιδίων έχει μεγάλη σημασία στην Ιατρική, καθώς θεωρείται ότι λόγω της δομής τους δεν είναι πιθανή η αποικοδόμηση τους από πρωτεάσες. Ένα παράδειγμα χρήσης τους, αποτελούν μελέτες σχετικά με την ανασταλτική δράση διαφόρων αλληλουχιών κυκλικών πεπτιδίων προς την συσσωμάτωση της Αβ πρωτεΐνης στη νόσο του Alzheimer. Επιπλέον, η παραγωγή μικρών πεπτιδίων και ειδικά κυκλικών ολιγοπεπτιδίων, για πολλά χρόνια γινόταν με την Οργανική Σύνθεση καθιστώντας το κόστος παραγωγής πολύ υψηλό. Στόχος αυτής της εργασίας ήταν η μικροβιακή παραγωγή κυκλικών ολιγοπεπτιδίων και ειδικότερα στο πρότυπο βακτηριακό σύστημα *E. Coli*, μια διαδικασία με μικρό κόστος σε σχέση με τις συμβατικές. Η παραγωγή των κυκλικών πεπτιδίων στην προτεινόμενη εργασία θα πραγματοποιηθεί μέσω μιας καινούριας τεχνολογίας *in vivo* κυκλοποίησης πρωτεϊνών και πεπτιδίων, η οποία ονομάζεται SICLOPPS (split intein circular ligation of peptides and proteins)). Η διεργασία αυτή χρησιμοποιεί ειδικά ανασχεδιασμένες ιντεΐνες (inteins, intervening protein sequences), όπως η διχασμένη DnaE trans-ιντεΐνη (split intein) από τον οργανισμό *Synechocystis* sp. PCC6803, οι οποίες έχουν την ιδιότητα να λειτουργούν ως πρωτεϊνικές λιγάσες και να συνδέουν τα N- και C- τελικά άκρα μιας πρωτεΐνης ή ενός πεπτιδίου (εξτεΐνη – extein, external protein sequence) με σταθερό πεπτιδικό δεσμό μέσω μιας αυτοκαταλυόμενης αντίδρασης που ονομάζεται μάτισμα πρωτεϊνών (protein splicing). Το σύστημα SICLOPPS μπορεί να παράγει κυκλικά προϊόντα το μέγεθος των οποίων ποικίλει από μεγάλες πρωτεΐνες μέχρι πολύ μικρά ολιγοπεπτίδια (μέχρι και τετραπεπτίδια). Η μοναδική προϋπόθεση για την πραγματοποίηση της αντίδρασης κυκλοποίησης είναι η ύπαρξη ενός πυρηνόφιλου αμινοξέως (κυστεΐνη (C), σερίνη (S), θρεονίνη (T)) στο N- τελικό άκρο του προς κυκλοποίηση πεπτιδίου.

The Medical use of cyclic oligopeptides is of great significance, as it has been considered that due to their structure degradation from proteases will not occur. For example there have been many studies on the inhibitory activity of cyclic peptides towards the aggregation of A β peptides in the Alzheimer disease. Furthermore the manufacturing of small peptides and especially cyclic oligopeptides, for many years has been a matter of Organic Synthesis making the production costs high. In this project we propose the production of a variety of cyclic oligopeptides from microorganisms, more specific from the model system *E.coli*, a cost-effective and easy downstream process. The production was carried out through a new technology of *in vivo* and *in vitro* cyclization of proteins and peptides, called SICLOPPS (split intein circular ligation of peptides and proteins). This process uses specially redesigned inteins (inteins, intervening protein sequences), which have the capacity to function as protein ligases and connecting the N- and C-termini of a protein or peptide (extein, external protein sequence) with a stable peptide bond through an autocatalytic reaction called protein splicing. The SICLOPPS system can produce cyclic products whose size varies from large proteins to small oligopeptides. The sole condition for performing the cyclization reaction is the presence of a nucleophilic amino acid (cysteine (C), serine (S), threonine (T)) at the N-terminus of the peptide towards cyclisation.

6. Βιβλιογραφία

- Aboye, T.L., and Camarero, J.A. (2012). Biological Synthesis of Circular Polypeptides. *J. Biol. Chem.* 287, 27026–27032.
- Akoachere, M., Squires, R.C., Nour, A.M., Angelov, L., Brojatsch, J., and Abel-Santos, E. (2007). Identification of an in Vivo Inhibitor of *Bacillus anthracis* Spore Germination. *J. Biol. Chem.* 282, 12112–12118.
- Allen, M.J., Lanzén, A., and Bratbak, G. (2011). Characterisation of the coccolithovirus intein. *Mar. Genomics* 4, 1–7.
- Ando, T., Tsukiji, S., Tanaka, T., and Nagamune, T. (2007). Construction of a small-molecule-integrated semisynthetic split intein for in vivo protein ligation. *Chem. Commun. Camb. Engl.* 4995–4997.
- Aranko, A.S., Oeemig, J.S., Kajander, T., and Iwai, H. (2013a). Intermolecular domain swapping induces intein-mediated protein alternative splicing. *Nat. Chem. Biol.* 9, 616–622.
- Aranko, A.S., Oeemig, J.S., and Iwai, H. (2013b). Structural basis for protein trans-splicing by a bacterial intein-like domain--protein ligation without nucleophilic side chains. *FEBS J.* 280, 3256–3269.
- Baine, M., Georgie, D.S., Shiferraw, E.Z., Nguyen, T.P.T., Nogaj, L.A., and Moffet, D.A. (2009). Inhibition of Aβ₄₂ aggregation using peptides selected from combinatorial libraries. *J. Pept. Sci. Off. Publ. Eur. Pept. Soc.* 15, 499–503.
- Barreto, K., Bharathikumar, V.M., Ricardo, A., DeCoteau, J.F., Luo, Y., and Geyer, C.R. (2009). A Genetic Screen for Isolating “Lariat” Peptide Inhibitors of Protein Function. *Chem. Biol.* 16, 1148–1157.
- Berg, T. (2008). Inhibition of transcription factors with small organic molecules. *Curr. Opin. Chem. Biol.* 12, 464–471.
- Berrade, L., and Camarero, J.A. (2009). Expressed protein ligation: a resourceful tool to study protein structure and function. *Cell. Mol. Life Sci. CMLS* 66, 3909–3922.
- Bigot, Y., Piégu, B., Casteret, S., Gavory, F., Bideshi, D.K., and Federici, B.A. (2013). Characteristics of inteins in invertebrate iridoviruses and factors controlling insertion in their viral hosts. *Mol. Phylogenet. Evol.* 67, 246–254.
- Birts, C.N., Nijjar, S.K., Mardle, C.A., Hoakwie, F., Duriez, P.J., Blaydes, J.P., and Tavassoli, A. (2013). A cyclic peptide inhibitor of C-terminal binding protein dimerization links metabolism with mitotic fidelity in breast cancer cells. *Chem. Sci.* 4, 3046–3057.
- Borra, R., Dong, D., Elnagar, A.Y., Woldemariam, G.A., and Camarero, J.A. (2012). In-Cell Fluorescence Activation and Labeling of Proteins Mediated by FRET-Quenched Split Inteins. *J. Am. Chem. Soc.* 134, 6344–6353.

- Bowers, A.A. (2012). Biochemical and biosynthetic preparation of natural product-like cyclic peptide libraries. *MedChemComm* 3, 905–915.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254.
- Brennan, T.V., and Clarke, S. (1995). Effect of adjacent histidine and cysteine residues on the spontaneous degradation of asparaginyl- and aspartyl-containing peptides. *Int. J. Pept. Protein Res.* 45, 547–553.
- Burt, A., and Koufopanou, V. (2004). Homing endonuclease genes: the rise and fall and rise again of a selfish element. *Curr. Opin. Genet. Dev.* 14, 609–615.
- Callahan, B.P., Topilina, N.I., Stanger, M.J., Van Roey, P., and Belfort, M. (2011). Structure of catalytically competent intein caught in a redox trap with functional and evolutionary implications. *Nat. Struct. Mol. Biol.* 18, 630–633.
- Callahan, B.P., Stanger, M., and Belfort, M. (2013). A redox trap to augment the intein toolbox. *Biotechnol. Bioeng.* 110, 1565–1573.
- Camarero, J.A., Fushman, D., Sato, S., Giriat, I., Cowburn, D., Raleigh, D.P., and Muir, T.W. (2001). Rescuing a destabilized protein fold through backbone cyclization. *J. Mol. Biol.* 308, 1045–1062.
- Casadaban, M.J., and Cohen, S.N. (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* 138, 179–207.
- Caspi, J., Amitai, G., Belenkiy, O., and Pietrokovski, S. (2003a). Distribution of split DnaE inteins in cyanobacteria. *Mol. Microbiol.* 50, 1569–1577.
- Caspi, J., Amitai, G., Belenkiy, O., and Pietrokovski, S. (2003b). Distribution of split DnaE inteins in cyanobacteria. *Mol. Microbiol.* 50, 1569–1577.
- Chen, L., Benner, J., and Perler, F.B. (2000). Protein Splicing in the Absence of an Intein Penultimate Histidine. *J. Biol. Chem.* 275, 20431–20435.
- Chen, S., Rentero Rebollo, I., Buth, S.A., Morales-Sanfrutos, J., Touati, J., Leiman, P.G., and Heinis, C. (2013). Bicyclic Peptide Ligands Pulled out of Cysteine-Rich Peptide Libraries. *J. Am. Chem. Soc.* 135, 6562–6569.
- Chen, W., Li, L., Du, Z., Liu, J., Reitter, J.N., Mills, K.V., Linhardt, R.J., and Wang, C. (2012). Intramolecular Disulfide Bond between Catalytic Cysteines in an Intein Precursor. *J. Am. Chem. Soc.* 134, 2500–2503.
- Cheriyian, M., Pedamallu, C.S., Tori, K., and Perler, F. (2013). Faster protein splicing with the *Nostoc punctiforme* DnaE intein using non-native extein residues. *J. Biol. Chem.* 288, 6202–6211.

Chin, H.G., Kim, G.-D., Marin, I., Mersha, F., Evans, T.C., Chen, L., Xu, M.-Q., and Pradhan, S. (2003). Protein trans-splicing in transgenic plant chloroplast: reconstruction of herbicide resistance from split genes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4510–4515.

Chong, S., Mersha, F.B., Comb, D.G., Scott, M.E., Landry, D., Vence, L.M., Perler, F.B., Benner, J., Kucera, R.B., Hirvonen, C.A., et al. (1997). Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* 192, 271–281.

Chong, S., Williams, K.S., Wotkowicz, C., and Xu, M.-Q. (1998a). Modulation of Protein Splicing of the *Saccharomyces cerevisiae* Vacuolar Membrane ATPase Intein. *J. Biol. Chem.* 273, 10567–10577.

Chong, S., Montello, G.E., Zhang, A., Cantor, E.J., Liao, W., Xu, M.Q., and Benner, J. (1998b). Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic Acids Res.* 26, 5109–5115.

Dalgaard, J.Z., Klar, A.J., Moser, M.J., Holley, W.R., Chatterjee, A., and Mian, I.S. (1997). Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family. *Nucleic Acids Res.* 25, 4626–4638.

Dassa, B., London, N., Stoddard, B.L., Schueler-Furman, O., and Pietrokovski, S. (2009). Fractured genes: a novel genomic arrangement involving new split inteins and a new homing endonuclease family. *Nucleic Acids Res.* gkp095.

Dearden, A.K., Callahan, B., Roey, P.V., Li, Z., Kumar, U., Belfort, M., and Nayak, S.K. (2013). A conserved threonine spring-loads precursor for intein splicing. *Protein Sci. Publ. Protein Soc.* 22, 557–563.

Derbyshire, V., Wood, D.W., Wu, W., Dansereau, J.T., Dalgaard, J.Z., and Belfort, M. (1997). Genetic definition of a protein-splicing domain: functional mini-inteins support structure predictions and a model for intein evolution. *Proc. Natl. Acad. Sci. U. S. A.* 94, 11466–11471.

Deschuyteneer, G., Garcia, S., Michiels, B., Baudoux, B., Degand, H., Morsomme, P., and Soumillion, P. (2010). Intein-Mediated Cyclization of Randomized Peptides in the Periplasm of *Escherichia coli* and Their Extracellular Secretion. *ACS Chem. Biol.* 5, 691–700.

Diderich, P., and Heinis, C. (2013). Directed Evolution of Bicyclic Peptides for Therapeutic Application. *Chim. Int. J. Chem.* 67, 910–915.

Ding, Y., Xu, M.-Q., Ghosh, I., Chen, X., Ferrandon, S., Lesage, G., and Rao, Z. (2003). Crystal structure of a mini-intein reveals a conserved catalytic module involved in side chain cyclization of asparagine during protein splicing. *J. Biol. Chem.* 278, 39133–39142.

Driggers, E.M., Hale, S.P., Lee, J., and Terrett, N.K. (2008). The exploration of macrocycles for drug discovery--an underexploited structural class. *Nat. Rev. Drug Discov.* 7, 608–624.

- Dun, B.-Q., Wang, X.-J., Lu, W., Zhao, Z.-L., Hou, S.-N., Zhang, B.-M., Li, G.-Y., Evans, T.C., Xu, M.-Q., and Lin, M. (2007). Reconstitution of Glyphosate Resistance from a Split 5-Enolpyruvyl Shikimate-3-Phosphate Synthase Gene in *Escherichia coli* and Transgenic Tobacco. *Appl. Environ. Microbiol.* 73, 7997–8000.
- Eryilmaz, E., Shah, N.H., Muir, T.W., and Cowburn, D. (2014). Structural and Dynamical Features of Inteins and Implications on Protein Splicing. *J. Biol. Chem.* 289, 14506–14511.
- Evans, T.C., and Xu, M.Q. (1999). Intein-mediated protein ligation: harnessing nature's escape artists. *Biopolymers* 51, 333–342.
- Evans, T.C., Xu, M.-Q., and Pradhan, S. (2005). Protein splicing elements and plants: from transgene containment to protein purification. *Annu. Rev. Plant Biol.* 56, 375–392.
- Evans, T.C., Jr, Martin, D., Kolly, R., Panne, D., Sun, L., Ghosh, I., Chen, L., Benner, J., Liu, X.Q., and Xu, M.Q. (2000). Protein trans-splicing and cyclization by a naturally split intein from the *dnaE* gene of *Synechocystis* species PCC6803. *J. Biol. Chem.* 275, 9091–9094.
- Fairlie, D.P., Tyndall, J.D., Reid, R.C., Wong, A.K., Abbenante, G., Scanlon, M.J., March, D.R., Bergman, D.A., Chai, C.L., and Burkett, B.A. (2000). Conformational selection of inhibitors and substrates by proteolytic enzymes: implications for drug design and polypeptide processing. *J. Med. Chem.* 43, 1271–1281.
- Faulkner, G.J., and Carninci, P. (2009). Altruistic functions for selfish DNA. *Cell Cycle Georget. Tex* 8, 2895–2900.
- Filée, J., Siguier, P., and Chandler, M. (2007). I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses. *Trends Genet. TIG* 23, 10–15.
- Fitzgerald, L.A., Graves, M.V., Li, X., Feldblyum, T., Nierman, W.C., and Van Etten, J.L. (2007). Sequence and annotation of the 369-kb NY-2A and the 345-kb AR158 viruses that infect *Chlorella* NC64A. *Virology* 358, 472–484.
- Frischkorn, K., Sander, P., Scholz, M., Teschner, K., Prammananan, T., and Böttger, E.C. (1998). Investigation of mycobacterial *recA* function: protein introns in the *RecA* of pathogenic mycobacteria do not affect competency for homologous recombination. *Mol. Microbiol.* 29, 1203–1214.
- Frutos, S., Goger, M., Giovani, B., Cowburn, D., and Muir, T.W. (2010). Branched intermediate formation stimulates peptide bond cleavage in protein splicing. *Nat. Chem. Biol.* 6, 527–533.
- Gause, G.F., and Brazhnikova, M.G. (1944). Gramicidin S and its use in the Treatment of Infected Wounds. *Nature* 154, 703–703.
- Gill, J.J., Berry, J.D., Russell, W.K., Lessor, L., Escobar-Garcia, D.A., Hernandez, D., Kane, A., Keene, J., Maddox, M., Martin, R., et al. (2012). The *Caulobacter crescentus* phage phiCbK: genomics of a canonical phage. *BMC Genomics* 13, 542.

- Goddard, M.R., and Burt, A. (1999). Recurrent invasion and extinction of a selfish gene. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13880–13885.
- Gogarten, J.P., Senejani, A.G., Zhaxybayeva, O., Olendzenski, L., and Hilario, E. (2002). Inteins: structure, function, and evolution. *Annu. Rev. Microbiol.* 56, 263–287.
- Hall, T.M., Porter, J.A., Young, K.E., Koonin, E.V., Beachy, P.A., and Leahy, D.J. (1997). Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins. *Cell* 91, 85–97.
- Hatfull, G.F. (2010). Mycobacteriophages: genes and genomes. *Annu. Rev. Microbiol.* 64, 331–356.
- Henchey, L.K., Porter, J.R., Ghosh, I., and Arora, P.S. (2010). High specificity in protein recognition by hydrogen-bond-surrogate α -helices: selective inhibition of the p53/MDM2 complex. *Chembiochem Eur. J. Chem. Biol.* 11, 2104–2107.
- Hessen, D.O., Jeyasingh, P.D., Neiman, M., and Weider, L.J. (2010). Genome streamlining and the elemental costs of growth. *Trends Ecol. Evol.* 25, 75–80.
- Hirata, R., Ohsumk, Y., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. (1990). Molecular structure of a gene, VMA1, encoding the catalytic subunit of H(+)-translocating adenosine triphosphatase from vacuolar membranes of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 265, 6726–6733.
- Horswill, A.R., Savinov, S.N., and Benkovic, S.J. (2004). A systematic method for identifying small-molecule modulators of protein–protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* 101, 15591–15596.
- Huang, J. (2013). Horizontal gene transfer in eukaryotes: The weak-link model. *BioEssays* 35, 868–875.
- Huang, X., Narayanaswamy, R., Fenn, K., Szpakowski, S., Sasaki, C., Costa, J., Blancafort, P., and Lizardi, P.M. (2012). Sequence-specific biosensors report drug-induced changes in epigenetic silencing in living cells. *DNA Cell Biol.* 31 Suppl 1, S2–S10.
- Iwai, H., Züger, S., Jin, J., and Tam, P.-H. (2006). Highly efficient protein trans-splicing by a naturally split DnaE intein from *Nostoc punctiforme*. *FEBS Lett.* 580, 1853–1858.
- Johnson, M.A., Southworth, M.W., Herrmann, T., Brace, L., Perler, F.B., and Wüthrich, K. (2007). NMR structure of a KlbA intein precursor from *Methanococcus jannaschii*. *Protein Sci. Publ. Protein Soc.* 16, 1316–1328.
- Kanno, A., Yamanaka, Y., Hirano, H., Umezawa, Y., and Ozawa, T. (2007). Cyclic luciferase for real-time sensing of caspase-3 activities in living mammals. *Angew. Chem. Int. Ed Engl.* 46, 7595–7599.

- Kanno, A., Umezawa, Y., and Ozawa, T. (2009a). Detection of apoptosis using cyclic luciferase in living mammals. *Methods Mol. Biol. Clifton NJ* 574, 105–114.
- Kanno, A., Ozawa, T., and Umezawa, Y. (2009b). Bioluminescent imaging of MAPK function with intein-mediated reporter gene assay. *Methods Mol. Biol. Clifton NJ* 574, 185–192.
- Kanno, A., Ozawa, T., and Umezawa, Y. (2011). Detection of protein-protein interactions in bacteria by GFP-fragment reconstitution. *Methods Mol. Biol. Clifton NJ* 705, 251–258.
- Kinsella, T.M., Ohashi, C.T., Harder, A.G., Yam, G.C., Li, W., Peelle, B., Pali, E.S., Bennett, M.K., Molineaux, S.M., Anderson, D.A., et al. (2002). Retrovirally Delivered Random Cyclic Peptide Libraries Yield Inhibitors of Interleukin-4 Signaling in Human B Cells. *J. Biol. Chem.* 277, 37512–37518.
- Kleiner, R.E., Dumelin, C.E., and Liu, D.R. (2011). Small-molecule discovery from DNA-encoded chemical libraries. *Chem. Soc. Rev.* 40, 5707–5717.
- Kritzer, J.A., Hamamichi, S., McCaffery, J.M., Santagata, S., Naumann, T.A., Caldwell, K.A., Caldwell, G.A., and Lindquist, S. (2009). Rapid selection of cyclic peptides that reduce α -synuclein toxicity in yeast and animal models. *Nat. Chem. Biol.* 5, 655–663.
- Larsen, J.B., Larsen, A., Bratbak, G., and Sandaa, R.-A. (2008). Phylogenetic Analysis of Members of the Phycodnaviridae Virus Family, Using Amplified Fragments of the Major Capsid Protein Gene. *Appl. Environ. Microbiol.* 74, 3048–3057.
- Lazarevic, V., Soldo, B., Dusterhoft, A., Hilbert, H., Mauel, C., and Karamata, D. (1998). Introns and intein coding sequence in the ribonucleotide reductase genes of *Bacillus subtilis* temperate bacteriophage SP? *Proc. Natl. Acad. Sci. U. S. A.* 95, 1692–1697.
- Lockless, S.W., and Muir, T.W. (2009). Traceless protein splicing utilizing evolved split inteins. *Proc. Natl. Acad. Sci.* 106, 10999–11004.
- Macao, B., Johansson, D.G.A., Hansson, G.C., and Härd, T. (2006). Autoproteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. *Nat. Struct. Mol. Biol.* 13, 71–76.
- Marsault, E., and Peterson, M.L. (2011). Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *J. Med. Chem.* 54, 1961–2004.
- Martin, D.D., Xu, M.Q., and Evans, T.C. (2001). Characterization of a naturally occurring trans-splicing intein from *Synechocystis* sp. PCC6803. *Biochemistry (Mosc.)* 40, 1393–1402.
- Mathys, S., Evans Jr., T.C., Chute, I.C., Wu, H., Chong, S., Benner, J., Liu, X.-Q., and Xu, M.-Q. (1999). Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene* 231, 1–13.

- Matsumura, H., Takahashi, H., Inoue, T., Yamamoto, T., Hashimoto, H., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T., and Kai, Y. (2006). Crystal structure of intein homing endonuclease II encoded in DNA polymerase gene from hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1. *Proteins* 63, 711–715.
- Mills, K.V., Dorval, D.M., and Lewandowski, K.T. (2005). Kinetic Analysis of the Individual Steps of Protein Splicing for the *Pyrococcus abyssi* PolII Intein. *J. Biol. Chem.* 280, 2714–2720.
- Mills, K.V., Johnson, M.A., and Perler, F.B. (2014). Protein Splicing: How Inteins Escape from Precursor Proteins. *J. Biol. Chem.* 289, 14498–14505.
- Miranda, E., Nordgren, I.K., Male, A.L., Lawrence, C.E., Hoakwie, F., Cuda, F., Court, W., Fox, K.R., Townsend, P.A., Packham, G.K., et al. (2013). A Cyclic Peptide Inhibitor of HIF-1 Heterodimerization That Inhibits Hypoxia Signaling in Cancer Cells. *J. Am. Chem. Soc.* 135, 10418–10425.
- Mizutani, R., Anraku, Y., and Satow, Y. (2004). Protein splicing of yeast VMA1-derived endonuclease via thiazolidine intermediates. *J. Synchrotron Radiat.* 11, 109–112.
- Mootz, H.D. (2009). Split Inteins as Versatile Tools for Protein Semisynthesis. *ChemBioChem* 10, 2579–2589.
- Moran, N.A. (2002). Microbial Minimalism: Genome Reduction in Bacterial Pathogens. *Cell* 108, 583–586.
- Morimoto, J., Hayashi, Y., and Suga, H. (2012). Discovery of Macrocyclic Peptides Armed with a Mechanism-Based Warhead: Isoform-Selective Inhibition of Human Deacetylase SIRT2. *Angew. Chem.* 124, 3479–3483.
- Muir, T.W., and Camarero, J.A. Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity. *J Am Chem Soc* 121, 5597–5598.
- Mujika, J.I., Lopez, X., and Mulholland, A.J. (2012). Mechanism of C-terminal intein cleavage in protein splicing from QM/MM molecular dynamics simulations. *Org. Biomol. Chem.* 10, 1207–1218.
- Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006). A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* 3, 357–359.
- Nagasaki, K., Shirai, Y., Tomaru, Y., Nishida, K., and Pietrokovski, S. (2005). Algal viruses with distinct intraspecies host specificities include identical intein elements. *Appl. Environ. Microbiol.* 71, 3599–3607.
- Naumann, T.A., Tavassoli, A., and Benkovic, S.J. (2008). Genetic Selection of Cyclic Peptide Dam Methyltransferase Inhibitors. *ChemBioChem* 9, 194–197.

- Nicastri, M.C., Xega, K., Li, L., Xie, J., Wang, C., Linhardt, R.J., Reitter, J.N., and Mills, K.V. (2013). Internal Disulfide Bond Acts as a Switch for Intein Activity. *Biochemistry (Mosc.)* 52, 5920–5927.
- Nolan, E.M., and Walsh, C.T. (2009). How Nature Morphs Peptide Scaffolds into Antibiotics. *ChemBioChem* 10, 34–53.
- Novikova, O., Topilina, N., and Belfort, M. (2014). Enigmatic distribution, evolution, and function of inteins. *J. Biol. Chem.* 289, 14490–14497.
- Oeemig, J.S., Zhou, D., Kajander, T., Wlodawer, A., and Iwai, H. (2012). NMR and crystal structures of the *Pyrococcus horikoshii* RadA intein guide a strategy for engineering a highly efficient and promiscuous intein. *J. Mol. Biol.* 421, 85–99.
- Ogata, H., Raoult, D., and Claverie, J.-M. (2005). A new example of viral intein in Mimivirus. *Virol. J.* 2, 8.
- Ozawa, T., Kaihara, A., Sato, M., Tachihara, K., and Umezawa, Y. (2001). Split Luciferase as an Optical Probe for Detecting Protein–Protein Interactions in Mammalian Cells Based on Protein Splicing. *Anal. Chem.* 73, 2516–2521.
- Passioura, T., and Suga, H. (2013). Flexizyme-Mediated Genetic Reprogramming As a Tool for Noncanonical Peptide Synthesis and Drug Discovery. *Chem. – Eur. J.* 19, 6530–6536.
- Paulmurugan, R., Umezawa, Y., and Gambhir, S.S. (2002). Noninvasive imaging of protein-protein interactions in living subjects by using reporter protein complementation and reconstitution strategies. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15608–15613.
- Paulus, H. (2001). Inteins as Enzyme. *Bioorganic Chem.* 29, 119–129.
- Pereira, B., Shemella, P.T., Amitai, G., Belfort, G., Nayak, S.K., and Belfort, M. (2011). Spontaneous proton transfer to a conserved intein residue determines on-pathway protein splicing. *J. Mol. Biol.* 406, 430–442.
- Perler, F.B. (1998). Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. *Cell* 92, 1–4.
- Perler, F.B. (2002). InBase: the Intein Database. *Nucleic Acids Res.* 30, 383–384.
- Perler, F.B., and Adam, E. (2000). Protein splicing and its applications. *Curr. Opin. Biotechnol.* 11, 377–383.
- Perler, F.B., Olsen, G.J., and Adam, E. (1997). Compilation and analysis of intein sequences. *Nucleic Acids Res.* 25, 1087–1093.
- Petrokovski, S. (2001). Intein spread and extinction in evolution. *Trends Genet.* 17, 465–472.

- Poland, B.W., Xu, M.-Q., and Quijcho, F.A. (2000). Structural Insights into the Protein Splicing Mechanism of PI-SceI. *J. Biol. Chem.* 275, 16408–16413.
- Ramirez, M., Valdes, N., Guan, D., and Chen, Z. (2013). Engineering split intein DnaE from *Nostoc punctiforme* for rapid protein purification. *Protein Eng. Des. Sel. PEDS* 26, 215–223.
- Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., Scola, B.L., Suzan, M., and Claverie, J.-M. (2004). The 1.2-Megabase Genome Sequence of Mimivirus. *Science* 306, 1344–1350.
- Rezai, T., Bock, J.E., Zhou, M.V., Kalyanaraman, C., Lokey, R.S., and Jacobson, M.P. (2006). Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: successful in silico prediction of the relative permeabilities of cyclic peptides. *J. Am. Chem. Soc.* 128, 14073–14080.
- Van Roey, P., Pereira, B., Li, Z., Hiraga, K., Belfort, M., and Derbyshire, V. (2007). Crystallographic and mutational studies of *Mycobacterium tuberculosis* recA mini-inteins suggest a pivotal role for a highly conserved aspartate residue. *J. Mol. Biol.* 367, 162–173.
- Romanelli, A., Shekhtman, A., Cowburn, D., and Muir, T.W. (2004). Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein–intein junction. *Proc. Natl. Acad. Sci. U. S. A.* 101, 6397–6402.
- Sambrook, J., and Russell, D.W. (2001). *Molecular Cloning: A Laboratory Manual* (CSHL Press).
- Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop — a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* 15, 430–434.
- Satyanarayana, M., Vitali, F., Frost, J.R., and Fasan, R. (2012). Diverse organo-peptide macrocycles via a fast and catalyst-free oxime/intein-mediated dual ligation. *Chem. Commun.* 48, 1461–1463.
- Schleif, R. (2000). Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet. TIG* 16, 559–565.
- Schwarz, F., and Aebi, M. (2011). Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.* 21, 576–582.
- Schwarzer, D., Ludwig, C., Thiel, I.V., and Mootz, H.D. (2012). Probing Intein-Catalyzed Thioester Formation by Unnatural Amino Acid Substitutions in the Active Site. *Biochemistry (Mosc.)* 51, 233–242.
- Scott, C.P., Abel-Santos, E., Wall, M., Wahnou, D.C., and Benkovic, S.J. (1999). Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13638–13643.
- Scott, C.P., Abel-Santos, E., Jones, A.D., and Benkovic, S.J. (2001). Structural requirements for the biosynthesis of backbone cyclic peptide libraries. *Chem. Biol.* 8, 801–815.

- Shah, N.H., and Muir, T.W. (2014). Inteins: nature's gift to protein chemists. *Chem. Sci.* 5, 446–461.
- Shah, N.H., Vila-Perelló, M., and Muir, T.W. (2011). Kinetic Control of One-Pot Trans-Splicing Reactions by Using a Wild-Type and Designed Split Intein. *Angew. Chem. Int. Ed.* 50, 6511–6515.
- Shah, N.H., Dann, G.P., Vila-Perelló, M., Liu, Z., and Muir, T.W. (2012). Ultrafast Protein Splicing is Common among Cyanobacterial Split Inteins: Implications for Protein Engineering. *J. Am. Chem. Soc.* 134, 11338–11341.
- Shah, N.H., Eryilmaz, E., Cowburn, D., and Muir, T.W. (2013a). Naturally Split Inteins Assemble through a “Capture and Collapse” Mechanism. *J. Am. Chem. Soc.* 135, 18673–18681.
- Shah, N.H., Eryilmaz, E., Cowburn, D., and Muir, T.W. (2013b). Extein Residues Play an Intimate Role in the Rate-Limiting Step of Protein Trans-Splicing. *J. Am. Chem. Soc.* 135, 5839–5847.
- Shao, Y., and Paulus, H. (1997). Protein splicing: estimation of the rate of O-N and S-N acyl rearrangements, the last step of the splicing process. *J. Pept. Res. Off. J. Am. Pept. Soc.* 50, 193–198.
- Shao, Y., Xu, M.Q., and Paulus, H. (1996). Protein splicing: evidence for an N-O acyl rearrangement as the initial step in the splicing process. *Biochemistry (Mosc.)* 35, 3810–3815.
- Shemella, P.T., Topilina, N.I., Soga, I., Pereira, B., Belfort, G., Belfort, M., and Nayak, S.K. (2011). Electronic structure of neighboring extein residue modulates intein C-terminal cleavage activity. *Biophys. J.* 100, 2217–2225.
- Shi, C., Meng, Q., and Wood, D.W. (2013). A dual ELP-tagged split intein system for non-chromatographic recombinant protein purification. *Appl. Microbiol. Biotechnol.* 97, 829–835.
- Skretas, G., and Wood, D.W. (2005a). Regulation of protein activity with small-molecule-controlled inteins. *Protein Sci. Publ. Protein Soc.* 14, 523–532.
- Skretas, G., and Wood, D.W. (2005b). A bacterial biosensor of endocrine modulators. *J. Mol. Biol.* 349, 464–474.
- Skretas, G., Meligova, A.K., Villalonga-Barber, C., Mitsiou, D.J., Alexis, M.N., Micha-Screttas, M., Steele, B.R., Screttas, C.G., and Wood, D.W. (2007). Engineered Chimeric Enzymes as Tools for Drug Discovery: Generating Reliable Bacterial Screens for the Detection, Discovery, and Assessment of Estrogen Receptor Modulators. *J. Am. Chem. Soc.* 129, 8443–8457.
- Sorci, M., Dassa, B., Liu, H., Anand, G., Dutta, A.K., Pietrokovski, S., Belfort, M., and Belfort, G. (2013). Oriented Covalent Immobilization of Antibodies for Measurement of Intermolecular Binding Forces between Zipper-Like Contact Surfaces of Split Inteins. *Anal. Chem.* 85, 6080–6088.

- Southworth, M.W., Benner, J., and Perler, F.B. (2000). An alternative protein splicing mechanism for inteins lacking an N-terminal nucleophile. *EMBO J.* 19, 5019–5026.
- Stephenson, R.C., and Clarke, S. (1989). Succinimide formation from aspartyl and asparaginyll peptides as a model for the spontaneous degradation of proteins. *J. Biol. Chem.* 264, 6164–6170.
- Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130.
- Sun, L., Ghosh, I., Paulus, H., and Xu, M.-Q. (2001). Protein trans-Splicing To Produce Herbicide-Resistant Acetolactate Synthase. *Appl. Environ. Microbiol.* 67, 1025–1029.
- Sun, P., Ye, S., Ferrandon, S., Evans, T.C., Xu, M.-Q., and Rao, Z. (2005). Crystal structures of an intein from the split dnaE gene of *Synechocystis* sp. PCC6803 reveal the catalytic model without the penultimate histidine and the mechanism of zinc ion inhibition of protein splicing. *J. Mol. Biol.* 353, 1093–1105.
- Tang, Y.Q., Yuan, J., Osapay, G., Osapay, K., Tran, D., Miller, C.J., Ouellette, A.J., and Selsted, M.E. (1999). A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated alpha-defensins. *Science* 286, 498–502.
- Tavassoli, A., and Benkovic, S.J. (2005). Genetically Selected Cyclic-Peptide Inhibitors of AICAR Transformylase Homodimerization. *Angew. Chem. Int. Ed.* 44, 2760–2763.
- Tavassoli, A., and Benkovic, S.J. (2007). Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in *E. coli*. *Nat. Protoc.* 2, 1126–1133.
- Toft, C., and Andersson, S.G.E. (2010). Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat. Rev. Genet.* 11, 465–475.
- Tori, K., Dassa, B., Johnson, M.A., Southworth, M.W., Brace, L.E., Ishino, Y., Pietrokovski, S., and Perler, F.B. (2010). Splicing of the Mycobacteriophage Bethlehem DnaB Intein IDENTIFICATION OF A NEW MECHANISTIC CLASS OF INTEINS THAT CONTAIN AN OBLIGATE BLOCK F NUCLEOPHILE. *J. Biol. Chem.* 285, 2515–2526.
- Tori, K., Cheriyan, M., Pedamallu, C.S., Contreras, M.A., and Perler, F.B. (2012). The *Thermococcus kodakaraensis* Tko CDC21-1 intein activates its N-terminal splice junction in the absence of a conserved histidine by a compensatory mechanism. *Biochemistry (Mosc.)* 51, 2496–2505.
- Volkman, G., and Iwai, H. (2010). Protein trans-splicing and its use in structural biology: opportunities and limitations. *Mol. Biosyst.* 6, 2110–2121.
- Volkman, G., and Liu, X.-Q. (2009). Protein C-terminal labeling and biotinylation using synthetic peptide and split-intein. *PloS One* 4, e8381.
- Volkman, G., and Mootz, H.D. (2013). Recent progress in intein research: from mechanism to directed evolution and applications. *Cell. Mol. Life Sci. CMLS* 70, 1185–1206.

Volkman, G., Volkman, V., and Liu, X.-Q. (2012). Site-specific protein cleavage in vivo by an intein-derived protease. *FEBS Lett.* **586**, 79–84.

Walensky, L.D., Kung, A.L., Escher, I., Malia, T.J., Barbuto, S., Wright, R.D., Wagner, G., Verdine, G.L., and Korsmeyer, S.J. (2004). Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science* **305**, 1466–1470.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951.

Wang, D., Liao, W., and Arora, P.S. (2005). Enhanced metabolic stability and protein-binding properties of artificial alpha helices derived from a hydrogen-bond surrogate: application to Bcl-xL. *Angew. Chem. Int. Ed Engl.* **44**, 6525–6529.

Wang, Z., Li, N., Wang, Y., Wu, Y., Mu, T., Zheng, Y., Huang, L., and Fang, X. (2012). Ubiquitin-intein and SUMO2-intein fusion systems for enhanced protein production and purification. *Protein Expr. Purif.* **82**, 174–178.

White, C.J., and Yudin, A.K. (2011). Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **3**, 509–524.

Wood, D.W., and Camarero, J.A. (2014). Intein Applications: From Protein Purification and Labeling to Metabolic Control Methods. *J. Biol. Chem.* **289**, 14512–14519.

Wood, D.W., Wu, W., Belfort, G., Derbyshire, V., and Belfort, M. (1999). A genetic system yields self-cleaving inteins for bioseparations. *Nat. Biotechnol.* **17**, 889–892.

Wood, D.W., Derbyshire, V., Wu, W., Chartrain, M., Belfort, M., and Belfort, G. (2000). Optimized single-step affinity purification with a self-cleaving intein applied to human acidic fibroblast growth factor. *Biotechnol. Prog.* **16**, 1055–1063.

Wray, W., Boulikas, T., Wray, V.P., and Hancock, R. (1981). Silver staining of proteins in polyacrylamide gels. *Anal. Biochem.* **118**, 197–203.

Wurth, C., Guimard, N.K., and Hecht, M.H. (2002). Mutations that reduce aggregation of the Alzheimer's Abeta42 peptide: an unbiased search for the sequence determinants of Abeta amyloidogenesis. *J. Mol. Biol.* **319**, 1279–1290.

Xu, M.-Q., and Evans, T.C. (2005). Recent advances in protein splicing: manipulating proteins in vitro and in vivo. *Curr. Opin. Biotechnol.* **16**, 440–446.

Xu, M.Q., and Evans, T.C., Jr (2001). Intein-mediated ligation and cyclization of expressed proteins. *Methods San Diego Calif* **24**, 257–277.

Xu, M.Q., and Perler, F.B. (1996). The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **15**, 5146–5153.

Xu, M.Q., Comb, D.G., Paulus, H., Noren, C.J., Shao, Y., and Perler, F.B. (1994). Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. *EMBO J.* *13*, 5517–5522.

Yang, J., Fox, G.C., and Henry-Smith, T.V. (2003). Intein-mediated assembly of a functional β -glucuronidase in transgenic plants. *Proc. Natl. Acad. Sci.* *100*, 3513–3518.

Yang, J., Henry-Smith, T.V., and Qi, M. (2006). Functional analysis of the split *Synechocystis* DnaE intein in plant tissues by biolistic particle bombardment. *Transgenic Res.* *15*, 583–593.

Young, T.S., Young, D.D., Ahmad, I., Louis, J.M., Benkovic, S.J., and Schultz, P.G. (2011). Evolution of cyclic peptide protease inhibitors. *Proc. Natl. Acad. Sci.* *108*, 11052–11056.

Yu, Y., You, L., Liu, D., Hollinshead, W., Tang, Y.J., and Zhang, F. (2013). Development of *Synechocystis* sp. PCC 6803 as a Phototrophic Cell Factory. *Mar. Drugs* *11*, 2894–2916.

Zheng, Y., Wu, Q., Wang, C., Xu, M.-Q., and Liu, Y. (2012). Mutual synergistic protein folding in split intein. *Biosci. Rep.* *32*, 433–442.

