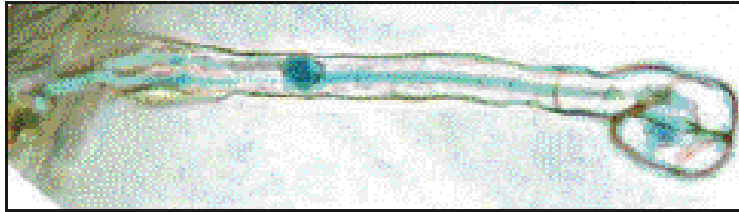


Initielle studier av Mll3104 og Mll3105 fra *Mesorhizobium loti*

- Proteiner antatt involvert i katabolisme av trehalose

Et mastergradsstudium i anvendt molekylærgenetikk og bioinformatikk



Av Linda J. Havdalen



Institutt for Biologi

Det Matematisk-Naturvitenskapelige Fakultet

Universitetet i Tromsø

2007

Forsidefoto: Osei Y. Ampomah, 2007

Innhold

Forord	6
Sammendrag	7
Forkortelser:	8
Introduksjon	10
Nitrogenfiksering	10
Rhizobiaceae	10
Leguminosae	10
Legum-rhizobia symbioser	10
Noduldannelsens forskjellige steg	11
Gjenkjennelse og adhesjon	11
Invasjon i rothår og dannelse av infeksjonstråd	12
Bakteroider, symbiosomer og nitrogenase	13
Overlevelse og kolonisering av rhizobia i rhizosfæren	14
Trehalose	15
Forekomst	15
Funksjoner	15
Biosyntese	16
Transport av trehalose	18
Katabolisme av trehalose	18
Trehalose hos rhizobia	18
Identifisering og karakterisering av trehaloserelaterte gener hos <i>Rhizobia sp.</i>	19
Hovedmål for studiet	21
Materiale	22
Metoder	26
Primer design	26
PCR- Polymerase Chain Reaction	27
Rensing av DNA fra agarosegel ved bruk av QIAquick Gel Extraction Kit	29

TOPO TA Cloning	32
Transformasjon av One Shot TOP10 elektrokompetente <i>E. coli</i>	33
Plasmidrensing ved bruk av QIAprep Miniprep kit.....	34
Ligering av <i>mll3104</i> og <i>mll3105</i> PCR-produkter i pET101/D-TOPO	35
Transformasjon av pET101/D-TOPO i <i>E. coli</i> BL21 Star	35
DNA-sekvensering	36
Pilotekspresjon av Mll3104- og Mll3105-proteiner med IPTG som transkripsjonsaktivator	38
Storskala IPTG-indusert uttrykk av Mll3104- og Mll3105-proteiner	39
Proteinekstrahering ved bruk av B-PER® Bacterial Protein Extraction Reagent	40
SDS-PAGE (Polyakrylamid Gel-elektroforese)	40
Rensing av polyhistidin-merkede Mll3104- og Mll3105 proteiner.....	43
Bio-Rad Protein Assay	45
Q-TOF MS/MS.....	46
Resultater.....	48
Primerdesign.....	48
Sekvenslikhet mellom <i>thuB</i> og <i>mll3104</i> og mellom <i>thuA</i> og <i>mll3105</i>	48
<i>thuB</i> og <i>mll3104</i>	48
<i>thuA</i> og <i>mll3105</i>	49
Genotypisk identifisering av trehalose assosierte gener; <i>mll3104</i> og <i>mll3105</i> fra <i>M. loti</i>	
MAFF303099 templat	49
Kvantitativ amplifisering av målgener <i>mll3104</i> og <i>mll3105</i>	50
Ekstrahering av amplifisert genomisk DNA fra <i>M. loti</i> etter agarosegel-elektroforese.....	50
Ligeringsreaksjon med ekstraherte <i>mll3104</i> og <i>mll3105</i> PCR-produkter i vektor pCR2.1 TOPO og transformasjon i <i>E. coli</i> TOP10 celler	51
Amplifisering av <i>mll3104</i> og <i>mll3105</i> med Pfx50 polymerase fra vektor pCR2.1-TOPO.....	53
Ligering av <i>mll3104</i>- og <i>mll3105</i> fragmenter i vektor pET101/D-TOPO.....	53
Transformasjon av pET1017D-TOPO i BL21 Star <i>E. coli</i> celler	55
DNA sekvensering av målsekvenser (<i>mll3104</i> og <i>mll3105</i>) etter kloning i pET101/D-TOPO	56

Sekvensering av <i>mll3104</i> -innsett i pETmll3104	57
Sekvensering av <i>mll3105</i> -innsett i pETmll3105	57
SDS-PAGE av proteinfraksjoner vasket med ulike imidazole konsentrasjoner	58
Mll3104-pilotekspresjons forsøk	58
Eluering av Mll3104-fraksjoner med ulike konsentrasjoner av imidazole.....	59
Mll3105-pilotekspresjons forsøk	60
Eluering av Mll3104-fraksjoner med ulike konsentrasjoner av imidazole.....	61
Q-TOF MS analyser	63
Sammenlikning av aminosyrer mellom eget innsett og opprinnelig sekvens av <i>mll3105</i>	65
Diskusjon	66
Ampicillin som seleksjonsmarkør	66
Amplifiserings resultater av genomisk <i>M. loti</i> DNA.....	67
Aminosyreendringer i <i>mll3105</i>-sekvens etter innsett i vektor, og mulige konsekvenser for proteinfunksjon	68
Rensing av Mll3104 og Mll3105 ved bruk av affinitetskolonne	70
DNA-sekvensering	73
Konklusjon.....	74
Referanser	75
Appendiks I	79
Appendiks II	80
Appendiks III	81
Appendiks IV	82
Appendiks V	83
Appendiks VI	84
Appendiks VII	85
Appendiks VIII	88

Forord

At jeg endte opp med et biologistudium er et resultat av forkjærlighet til naturvitenskap, samt utallige diskusjoner med familien. Gro, veien mot en universitetsgrad ble opprinnelig inspirert av våre sene kvelder der miljø møtte arv.

Dette studiet ble utført ved Institutt for biologi, avdeling for mikrobiell økologi ved Universitetet i Tromsø. Arbeidet begynte vinteren 2006, og ble avsluttet vinteren 2007.

Gjennom året som student hos førsteamanuensis John Beck Jensen har jeg fått utfordringer hele veien (det samme kan med høy sannsynlighet sies for hans del). Som veileder har John velvillig delt av sin kunnskap og forlangt progresjon fra min side, og som medmenneske har han vist empati og tålmodighet i perioder der det var vanskelig å holde et faglig fokus. Takk for at "du sto han av", John.

Til alle som utgjør det positive miljøet av Gjennomgangsbygget, jeg hadde ikke klart meg uten dere. Spesiell takk til jentene på laben: Coby (min nederlandske tulipan), Eli, Anne-Grete og Janne. Uten hjelpsomheten deres hadde jeg fortsatt stått på laben og klødd meg i hodet.

Hanne, min solstråle og humørspreder av en samboer i brakke nord, ord blir fattige. Du har vært min klagemur for alle sorger og bekymringer da jeg bare ville heise det kjente hvite flagget og kapitulere. I tillegg har du tidvis fungert som hobbypsykolog på et generelt grunnlag, DJ, kaffe-direktør, vekkeklokke og motivasjonskonsulent. Tusen takk, Hannemor!

Junita, Vero og Anders: Det har vært gjennomført >2 tildelinger av spark bak underveis i studiet, noe jeg er svært takknemlig for i dag. Samtidig har det alltid vært rom for diskusjoner av faglig og mye mindre faglig karakter på kollokvierommet i Storgata 48. Jeg føler meg privilegert som nyter deres bekjentskap, og dere utgjør en stor del av gode minner jeg forbinder studietiden med.

Takk til mamma og Erling for at dere er der med meg. Uten familiens bidrag hadde mang en eksamen gått rett vest. Og takk til alle andre som har bidratt med oppmuntring og faglige innspill underveis.

Robin og Sindre, jeg håper at jeg i fremtiden kan gjøre opp for tiden jeg ikke var tilstede. Dere er drivkraften og har vært i tankene mine hele veien.

Linda J. Havdalen
Tromsø, desember 2007

Sammendrag

Rhizobiacea utgjør en bakterieslekt der de fleste medlemmene er i stand til å inngå symbioser med diverse legumer. Slekten innbefatter *Rhizobium*, *Sinorhizobium*, *Mesorhizobium*, *Bradyrhizobium* og *Azorhizobium*, og kollektivt omtales disse for rhizobia. Partnerskapet med legumer er av mutualistisk karakter, der mikrosymbionten bidrar med fiksert nitrogen i form av NH_3 , og planteverten sørger for en kontinuerlig forsyning av karbon til sin bakterielle partner. Grunnlaget for nitrogenfikseringen er enzymet nitrogenase, som er unikt for nitrogenfikserende bakterier. Legum-rhizobia symbioser er spesifikke, og en type rhizobie kan vanligvis kun inngå symbioser med et fåtall legumarter. Mengden nitrogen som blir gjort tilgjengelig for planter ved denne typen symbiose er av stor betydning både for kommersielt landbruk og terrestrisk økologi i sin helhet.

I jordsamfunn der miljøendringer brått oppstår, er det essensielt at mikroorganismene kan omstille metabolismen for å takle ulike typer stress. Trehalose er et velkjent molekyl med mange stressrelaterte funksjoner. Disakkaridet er funnet i bakterier, sopp, insekter og planter, samt noen invertebrater. I tillegg til funksjoner knyttet til miljøstress, brukes trehalose også som energikilde, og enkelte rhizobier er i stand til å utnytte sakkariidet som eneste karbonkilde. Tidligere arbeid med *Sinorhizobium meliloti* har ført til identifisering og karakterisering av gener involvert i utnyttelse og transport av trehalose, *thuA* og *thuB*.

Ved fullsekvensering av genomet til *Mesorhizobium loti* ble det funnet gener med homologi til *thuA* og *thuB*; *mll3105* og *mll3104*. Eksperimenter har vist at *M. loti*- og *S. meliloti*-stammer med mutasjoner i *thuA/mll3105* eller *thuB/mll3104* ga redusert vekst på ulike medier, i forhold til villtypen. Funnene indikerer at *mll3104* og *mll3105* delvis har de samme funksjonene i *M. loti* som *thuAB* i *S. meliloti*, men lokalisasjonen for operonet er endog noe ulik mellom de to artene. I *M. loti* er operonet lokalisert proksimalt til gener involvert i aminosyremetabolisme, men hittil er ikke *mll3104* og *mll3105* karakterisert i den grad at en deltagende rolle under denne typen metabolisme kan bekreftes. Det er hittil ikke funnet system for katabolisme og transport av trehalose hos *M. loti*, og eksperimenter utført under dette studiet, tar sikte på å frembringe mer kunnskap om *M. loti*'s *thuAB* ortologer.

Forkortelser:

ABC	ATP-bindings cassette (transport system)
ADP	Adenin difosfat
BLAST	Basic Local Alignment Search Tool
BNF	Biologisk Nitrogenfiksering
bp	Basepar
BSA	Bovin Serum Albumin
cDNA	komplementær DNA
ddNTP	Dideoksynukleotid trifosfat
dH ₂ O	Destillert vann
DNA	Deoksyribonukleinsyre
dNTP	Deoksynukleotid trifosfat
dsDNA	Dobbeltrådet DNA
DTT	Dithiothreitol
EtBr	Ethidium bromid
IPTG	Isopropyl-thio-2-D-galaktopyranosid
Kb	Kilo base
kDa	Kilo Dalton
LB	Luria Bertani
MOTS	Maltooligosyl trehalose syntase
NCBI	National Center for Biotechnology Information
NH ₃	Ammonium
nm	Nanometer
OD	Optical Density (optisk tetthet)
PBS	Fosfat buffer saline
PCR	Polymerase Chain Reaction (polymerase kjedereaksjon)
PGPB	Plant Growth Promoting Bacteria (plantevekst fremmende bakterier)
PTS	Fosfoenolpyruvat fosfotransferase
Q-TOF	Quadrupole Time of Flight Mass Spectrometry

S.O.C.-medium	Super Optimal Broth
SDS-PAGE	Sodium Dodecyl Sulfat PolyAkrylamid Gel-elektroforese
TAE	Tris-acetat/EDTA
TBE	Tris-borat/EDTA
TEMED	Tetrametylendiamin
T _m	Smeltetemperatur
TPP	Trehalose 6-fosfat fosfatase
TreH	Trehalase
TreP	Trehalose fosforylase
TreT	Trehalose glykosyloverførende syntase
TreY	Maltooligosyl trehalose syntase
TreZ	Maltooligosyl trehalose trehalohydrolase
TS	Trehalose syntase
TSP	Trehalose 6-fosfat syntase
UDP	Uridin difosfoglukose
UV	Ultra fiolett
X-Gal	5-bromo-4-kloro-3-indolyl-β-D-galaktopyranosid

Introduksjon

Nitrogenfiksering

Planters behov for nitrogen er stort, og tilgjengeligheten er begrenset. Atmosfærisk N₂ er ikke anvendbart for planter, men noen plantegrupper tilegner seg likevel mengder nitrogen indirekte gjennom symbioser med prokaryoter. Nitrogenfiksering er en egenskap som kun finnes hos prokaryotene, med representanter både fra *Eubacteria* og *Archaeobacteria* (Fischer 1994). Nitrogenfikserende bakterier (diazotrofer) finnes som frittlevende celler og som symbionter med landlevende planter (Engelhard 2004). Nøkkelen til nitrogenfiksering er nitrogenase, et enzymkompleks unikt for diazotrofe bakterier. På grunn av fordelene samarbeidet gir symbiontene, blir nitrogenfikserende bakterier ofte omtalt som PGPB -Plant Growth Promoting Bacteria (Waisel 2002). Bruk av PGPB kalles biologisk nitrogenfiksering (BNF), da ekstern tilsetning av nitrogengjødsel og annen manipulering av nitrogentilsetning i jorda kan reduseres. Dette har resultert i bærekraftige økonomiske og miljømessige forhold (Gurevitch, Scheiner et al. 2006).

Rhizobiaceae

Rhizobier er Gram negative α -proteobakterier i familien *Rhizobiaceae*, og kjente PGPB som *Rhizobium*, *Sinorhizobium*, *Mesorhizobium*, *Azorhizobium* og *Bradyrhizobium* inkluderes (Glenn 1980; Chandra, Choure et al. 2007). Rhizobier inngår symbioser med legumer, men lever også som saprofytter i rhizosfæren, jorda som omgir planterøtter.

Leguminosae

Legumer tilhører familien *Fabaceae*, også kalt *Leguminosae*, og deles videre inn i tre underfamilier: *Caesalpinioideae*, *Mimososideae* og *Papilionideae*. Familien består av rundt 750 genera og 20.000 arter, der de fleste kan inngå nitrogenfikserende symbioser med rhizobier. Bidraget nitrogen som blir gjort tilgjengelig ved legum-rhizobia symbioser utgjør 30 % av årlig BNF, og er den vanligste gruppen nitrogenfikserende plantesymbioser (Sprent 2001)

Legum-rhizobia symbioser

Legum-rhizobia symbioser var de første til å bli beskrevet, og fungerer i dag som modellsystem for forskning og undervisning rundt fenomenet BNF. Endosymbiosen er mutualistisk, der bakterier bidrar med NH₃, og planteverten ofrer verdifulle karbonkilder. Symbiosene er svært spesifikke, og som regel vil en type rhizobia kunne inngå symbiose med

et fåtall legumarter (Townsend 2006). Det er rapportert at legumer kan skille ut substanser som kompatible rhizobier metaboliserer bedre enn eksudater fra annen kompetitiv flora. Eksempelvis vil betain frigjort fra spirende alfalfa-frø fungere både som næringskilde og osmoprotektant for *Sinorhizobium meliloti*, noe som kan indikere at tilstedeværelse av en passende vertsplante øker jordpopulasjoner av kompatible rhizobier (Sprent 2001).

Noduldannelsens forskjellige steg

Rotnoduler ble først observert på 1500-tallet, og ble lenge ansett som sykdomstegn hos planter. Allerede i 1888 demonstrerte Hellriegel og Wilfarth at rotnoduler kunne assosieres med opptak av nitrogen, men det er i løpet av de siste 35 årene at man har gjort de store framskrittene når det kommer til forståelsen av de fysiologiske, biokjemiske og molekylære egenskapene som inngår i BNF, inkludert dannelse av rotnoduler (Waisel 2002; Ampomah 2007).

Generelt inkluderes disse stegene under noduldannelse:

- 1) Gjenkjennelse av den korrekte partner for både plante og bakterie, samt adhesjon til rothår.
- 2) Ekskresjon av Nod-faktorer fra bakteriene.
- 3) Invasjon i rothår og dannelse av infeksjonstråd.
- 4) Dannelse av bakteroider i plantecellers indre, og dannelse av rotnodul.
- 5) Fortsatt bakteriell og plantecelle-deling, og modning av rotnodulen.

(Madigan, Martinko et al. 2002)

Gjenkjennelse og adhesjon

Før rhizobier kan infisere legumer må planten og dens kompatible bakterielle partner gjenkjenne hverandre, deretter må rhizobier kolonisere rotoverflaten og adhere til rothåret. Det er flere komponenter i gjenkjennelsesprosessen, der partene resiprokalt signaliserer hverandre (Gage 2004). Vanlige attraktanter for bakteriell kjemotaksis mot planterøtter er flavenoider, isoflavoner, betainer og enkle fenoler. Plantens spekter av disse stoffene avgjør dens kompatibilitet med visse bakterier (Broughton, Jabbouri et al. 2000; Brencic and Winans 2005). Sekrerte plantesubstanser bindes til bakterielle transkripsjonsregulerende NodD-proteiner med påfølgende induksjon av *nod*-gener (*nod*-, *nol*- og *noe*- gener) lokalisert i *nod*-bokser. Planter utsondrer kontinuerlig basale nivå av flavenoider, men konsentrasjonen er vist

å øke i rhizosfæren ved tilstedeværelse av kompatible rhizobier (Broughton, Jabbouri et al. 2000).

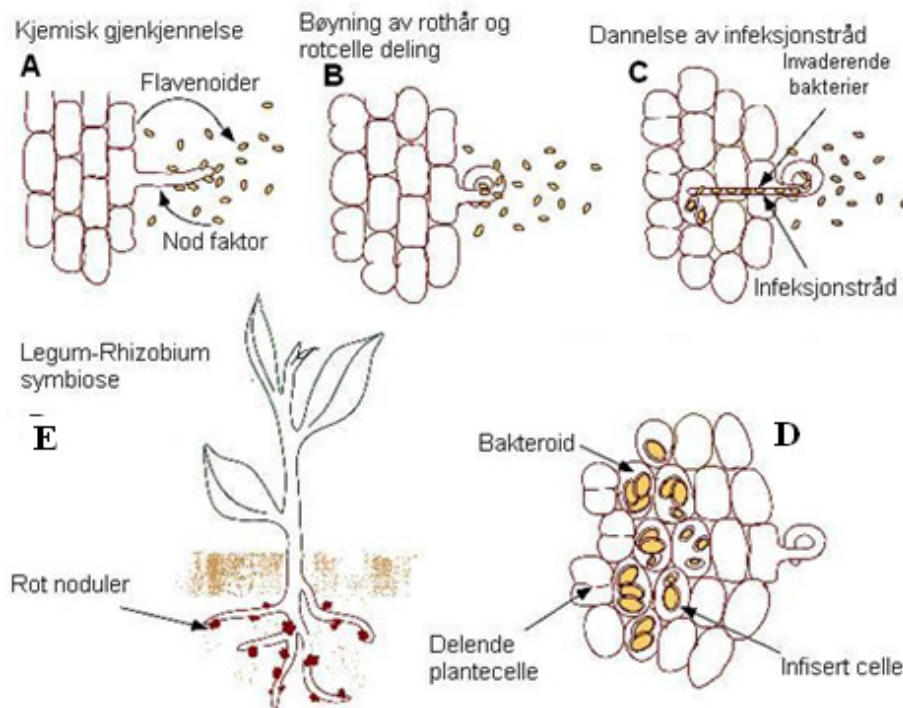
nod-gener koder for enzymer deltagende i biosyntese av vertsspesifikke nodulerings signaler (lipochitooligosakkarider) kalt Nod-faktor (van Rhijn and Vanderleyden 1995). Signalmolekylene sekreseres ut og diffunderer til rothåret, med påfølgende systemisk respons i planten, noe som induserer prosessen for noduldannelse. Både epidermale, kortikale og pericycle celler endres etter stimuli fra Nod-faktor (Sæki and Kouchi 2000).

Den grunnleggende strukturen til Nod-faktorer er ganske lik mellom forskjellige rhizobier, og består generelt av et β -1,4-koblet-N-acyl-D-glukosamin skjelett, med 4 eller 5 muligheter for tilleggsgrupper. Mange rhizobier produserer mer enn en type Nod-faktor, og genene som koder for syntese av Nod-faktor finnes hos *Rhizobium spp.* på såkalte symbiotiske (Sym) øyer i kromosomet (Whitehead NA 2001; Jones, Kobayashi et al. 2007).

Negativ regulering av Nod-faktor skjer via feedbackmekanismer i bakterier og fra etylen som sannsynligvis produseres i pericycle planteceller (Guerts R 2002). Planterot-hydrolaser vil på sin side modulere konsentrasjonen av Nod-faktor i rhizosfæren, og undertrykker dermed forsvarsresponsen som ville blitt stimulert pga de høye nivåene av bakterielle signalstoffer (Somers, Vanderleyden et al. 2004)

Invasjon i rothår og dannelse av infeksjonstråd

Det er kun rothår i vekstfasen som er mottakelig for invasjon av rhizobia, og initiell penetrering av bakterier skjer via tuppen av disse (Bhuvaneswari, Turgeon et al. 1980). Den vanligste måten for innpass er via invaginering av plantens plasmamembran, men skader i plantens epidermis er også årsak til infeksjoner (Sprent 2001). Kort tid etter stimuli fra Nod-faktor sveller rothåret og videre vekst bøyes mot én side. Det bøyde rothåret fikk navnet "shepherd's crook" av tidlige mikroskopister, og mikrokolonier med rhizobier entrer rothåret fra denne lommen. Videre migrasjon mot rotens sentrum skjer intercellulært gjennom en planteindusert tubulær struktur kalt infeksjonstråd (Sprent 2001; Madigan, Martinko et al. 2002) som ekskluderer bakterieceller fra plantecytoplasma. Simultant vil Nod-faktor stimulere kortikale celler til mitose (figur 1) (Guerts R 2002; Jones, Kobayashi et al. 2007).



Figur 1: Prosess mot dannelse av rotnoduler. **A)** Sekrerte flavenoider fra planten senses av nærliggende rhizobier. Respons på flavenoider fra korrekt plantepartner fører til induksjon av Nod-faktor fra kompatible bakterier som da syntetiserer Nod-faktor. **B)** Som respons på Nod-faktor vil rothår deformeres, og danne en "shepard's crook" der mikrokolonier fanges. **C)** Invaginasjon av planteplastamembranen med bakterieceller i lumen, kalles infeksjonstråd, og sørger for at bakteriene flyttes intercellulært mot rotas sentrum gjennom elongering av tråden. **D)** Modne rotnoduler der jevnverdig utveksling av nitrogen og karbon kontinuerlig skjer, vesikkelen med bakteroid omtales ofte som et symbiosom. **E)** Endocytosen av bakterieceller i plantecellene fører til morfologiske endringer av bakteriene, med terminal omdannelse til nitrogenfikserende bakteroider. Rotknoller kan observeres med det blotte øye, og kan vare ut plantens livstid. (Long and Staskawitz 1993).

Bakteroider, symbiosomer og nitrogenase

Det endelige målet for infeksjonstråden er de indre kortikale cellene. Herfra endocytteres enkeltbakterier og forblir i cytoplasma i plantederiverte vesikler der videre celledeling skjer (Gonzalez and Marketon 2003). Bakteriene endres morfologisk til bakteroider, med induksjon av *nif*- og *fix*- gener, hvis genprodukt gir opphav til enzymkomplekset nitrogenase (Oke and Long 1999). Enheten bakteroid og vertscellemembran refereres ofte til som et symbiosom, og enkeltceller kan huse tusenvis av symbiosomer (se fig 1 D). Simultant med induksjon av *nif*- og *fix*-genene, endres membranen som omgir bakteriene på en slik måte at en balansert utveksling av metabolitter skjer (Fischer 1994).

Overlevelse og kolonisering av rhizobia i rhizosfæren

Som saprofyttiske bakterier er rhizobia en del av et dynamisk og komplekst jordsamfunn, der den kvalitative og kvantitative sammensetningen av mikrobiell flora raskt endres. Tabell 1 viser et eksempel på relative antall (m^2/g) jordorganismer. Prøvene er tatt 18 cm under overflaten i vanlig markjords pløyselag, og gir et bilde av et habitat der ulike mikroorganismer lever i umiddelbar nærhet til hverandre og konkurransen om næring er stor.

Tabell 1: Relativt antall mikroorganismer i jord pr m^2/g (Brady 1984).

	Gjennomsnittlig antall i overflatejord	
	Per m^2	Per g
Bakterier	$10^{13}-10^{14}$	10^8-10^9
Sopp	$10^{11}-10^{11}$	10^5-10^6
Actinomyceter	$10^{12}-10^{13}$	10^7-10^8

Rhizobias kolonisering og overlevelse utenfor etablerte symbioser refereres til som saprofyttisk kompetanse, og evnen varierer mellom ulike stammer. Som frittlevende bakterier opptrer familien *Rhizobiales* hovedsaklig på rhizoplanet (rotoverflaten), i rhizosfæren, og til en viss grad i non-rhizosfæren (jord som ikke er kolonisert av planterøtter). Økning i bakterielle populasjoner rundt røtter der anvendbart roteksudat er tilstede, kalles rotkolonisering, og evnen til å katabolisere et bredt spekter av tilgjengelige energikilder er øker den saprofyttiske kompetansen.

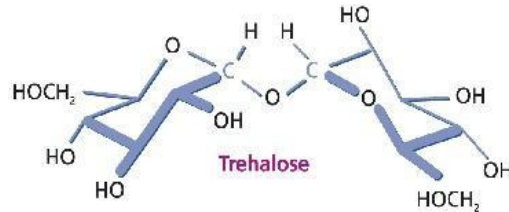
Faktorer som influerer sammensetningen av jordmikrober kan være type tilgjengelig karbonkilde, temperatur og klima, vegetasjon, jord-pH, salinitet, samt fuktighet og tørke. Tilstedeværende organismers næringskrav, vekstrate, formering, enzymproduksjon (cellulaser og kitinaser etc.), antibiotikaproduksjon og annen gjensidig påvirkning endres stadig, og sammenfattet utgjør dette ustabile habitater der inhabitantene er utsatt for intens næringskonkurranse i tillegg til ulike typer miljøstress (Jensen 1999)

For plutselige miljøendringer har bakterier og andre organismer evolvert verktøy i form av organiske molekyler som gjør dem i stand til å håndtere stress. Et velkjent stress-relatert molekyl som benyttes gjennom store deler av den biologiske verden, er disakkaridet trehalose.

Trehalose

Forekomst

Trehalose er et disakkarid som forekommer i tre isomerer, der glukosemolekylene er bundet av en 1,1-glykosid binding (fig 2).



Figur 2; Struktur av den naturlig forekommende isomeren av trehalose, $\alpha,\alpha,1,1$ -trehalose (Elbein, Pan et al. 2003).

Trehalose er vanlig i gjær og sopp, der sakkaridet opptre i sporer, fruktlegemer og vegetative celler. I dyreriket finnes trehalose i insektsblod hos larver, på puppestadiet og i voksne insekter. Disakkaridet er i tillegg identifisert i rundormegg, i larver og voksen rundorm, i reker, samt visse invertebrater (Shen 2006). Trehalose opptre også hos vaskulære planter, og finnes akkumulert i symbiotiske organer (Farias-Rodriguez, Mellor et al. 1998). I sistnevnte sammenheng er trehalose tenkt derivert fra mikrosymbionten, eksempelvis i mycorrhizarøtter, actinomycorrhizale noder eller nitrogenfikserende rotnoder (Engelhard 2004). Heller ikke hos prokaryoter er trehalose et ukjent molekyl, familier som *Streptomyces*, *Mycobacteria*, *Corynebacteria* og *Rhizobiaceae* er i stand til å syntetisere disakkaridet (Elbein, Pan et al. 2003).

Funksjoner

Man har tidligere ansett trehalose som en lagringsform av karbon, men nyere forskning viser at disakkaridet har langt flere funksjoner enn som energikilde. Trehalose inngår i transport, signalisering, regulering og som celleveggkomponent. I flere organismer stabiliserer og beskytter trehalose membraner og proteiner, og omtales som et osmotisk aktivt stoff mot stressfaktorer som varme, kulde, salinitet, dehydrering-rehydreringsskader, samt anoksi (Elbein, Pan et al. 2003; Avonce, Mendoza-Vargas et al. 2006). Mange organismer produserer og lagrer trehalose i mengder opptil 20 % av egen tørrvekt (Shen 2006). Flygende insekter bruker trehalose som energi for toraksmuskulatur under flyvning, mens planter anvender

sakkaridet som stressprotektant. I *Arabidopsis thaliana* har trehalose vist seg å være essensielt for sukkersignalisering og embryoutvikling (Avonce, Mendoza-Vargas et al. 2006).

Industrielt har trehalose gjort suksess på grunn av sine unike fysiske egenskaper. Molekylet har høy hydrofilitet, kjemisk stabilitet og fravær av interne hydrogenbindinger, noe som gjør det egnet til stabilisering av proteiner (Engelhard 2004). Kommersielt brukes derfor trehalose som tilsetningsstoff, stabilisator og søtningsmiddel i mat-, kosmetikk- og farmasøytisk industri (Shen 2006).

Forskning rundt trehalose konsentreres i stor grad rundt det faktum at trehalose bedrer planters viabilitet og produksjon under stressforhold. Det er foreslått at genetisk manipulert trehaloseakkumulering i tobakksplanter kan bedre toleransen for salinitet og dehydrasjon, og det er demonstrert at overproduksjon av trehalose kan brukes for utvikling av nye riskulturer med økt abiotisk stresstoleranse og avlingsproduktivitet (López, Herrera-Cervera et al. 2006).

Biosyntese

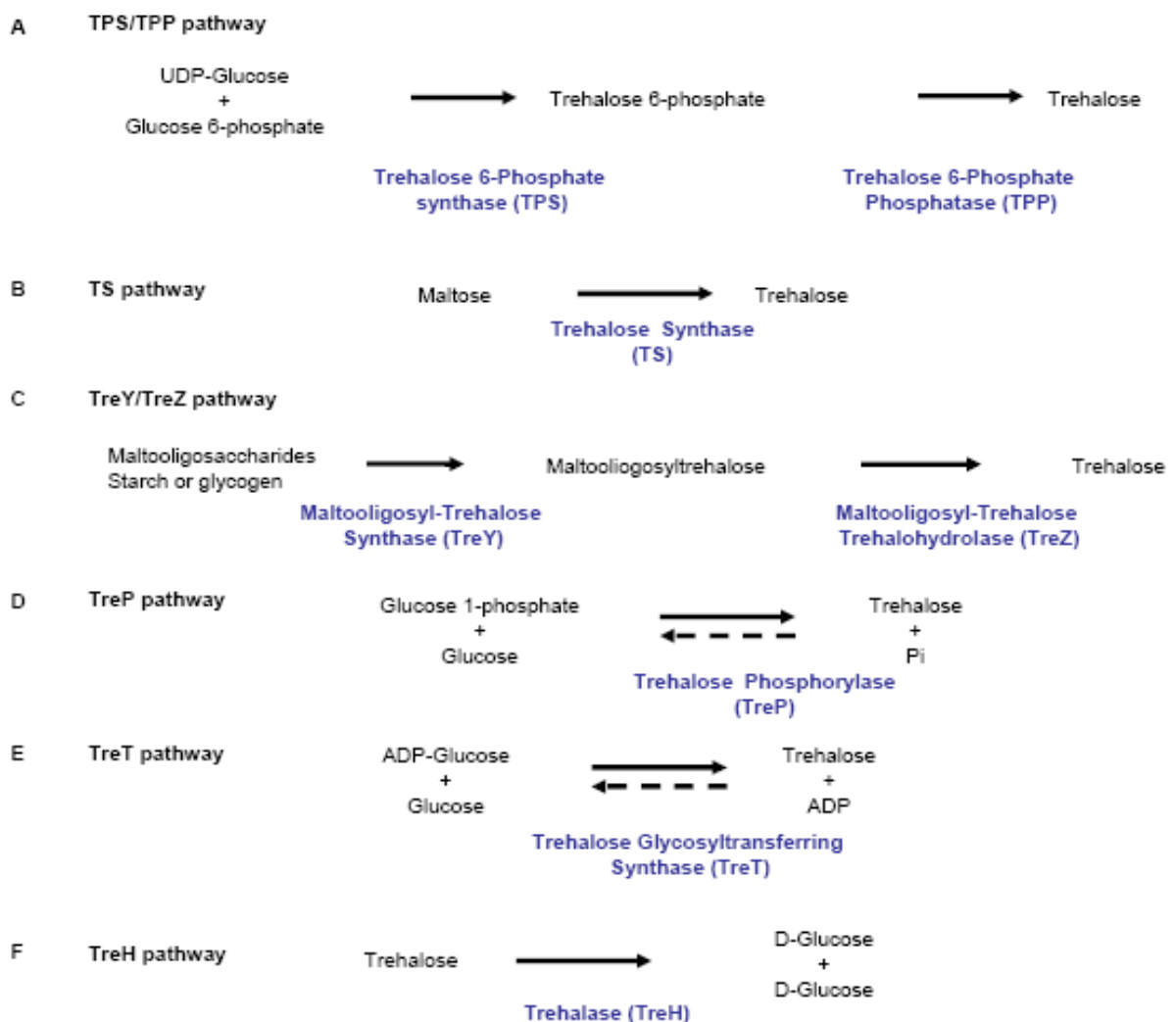
Flere biosynteseveier for trehalose er beskrevet. Den mest utbredte er Ots/AB (figur 3A) som danner trehalose-6-fosfat fra UDP-glukose og glukose-6-fosfat i en totrinns reaksjon som katalyseres av trehalose-6-fosfat syntase (TPS) og trehalose-fosfatase (TPP). Utbyttet er fritt trehalose, og syntesen er oftest beskrevet i sammenheng med beskyttelse mot stress (De Smet, Weston et al. 2000). Ots/B-biosyntese finnes i Eubacteria, archaea, sopp, insekter og planter (Elbein, Pan et al. 2003).

En andre syntesevei består av en enkel transglykolyserings reaksjon som gir trehalose fra maltose, og kalles TreS- eller TS-pathway. Enzymet trehalose syntase (TS) som katalyserer reaksjonen, ble først rapportert hos *Pimelobacter sp.* Ortologer av proteinet er senere funnet i andre eubakterier (figur 3B).

En tredje biosyntese av trehalose involverer konvertering av oligo/polymaltodextrin/glykogen til trehalose i en totrinns reaksjon, og refereres til som MOTS-pathway (figur 3C) (Maruta 1995). MOTS-biosyntese er rapportert hos den termofile archea-familien *Sulfolobus*. Organismene syntetiserer trehalose i to enzymatiske trinn som katalyseres av maltooligosyl trehalose syntase (TreY) og maltooligosyl trehalose trehalohydrolase (TreZ).

I tillegg finnes en biosyntese av trehalose som hittil kun er funnet hos visse sopparter, der reversibel hydrolysering av trehalose skjer i nærvær av uorganisk fosfat (figur 3D) (Avonce, Mendoza-Vargas et al. 2006).

En siste syntesevei er beskrevet hos den hypertermofile archaeae-bakterien *Thermococcus litoralis* (figur 3E og F). Trehalose glykosyloverførende syntase (TreT) katalyserer en reversibel dannelsen av trehalose fra ADP-glukose og glukose (Ampomah 2007).



Figur 3: Biosynteseveier for trehalose. Se tekst for forklaring og forkortelser. Alle synteseveiene bortsett fra D er funnet hos bakterier (Avonce, Mendoza-Vargas et al. 2006).

Transport av trehalose

Det er dokumentert tre transporttyper for trehalose hos bakterier, henholdsvis ATP-bindings cassette (ABC), fosfoenolpyruvat fosfotransferase (PTS) og galaktose permease transportsystemer (Boos, Ehmann et al. 1990; Ampomah 2007).

Katabolisme av trehalose

Per dags dato er det funnet tre kataboliseringssystemer for trehalose hos mikroorganismer;

1. Trehalose hydrolyseres til to glukoseenheter av enzymet trehalase, et system funnet i mange prokaryotiske- og eukaryotiske systemer.
2. Trehalose transport gjennom membraner ved hjelp av permeaser eller et PTS-system, der trehalose entrer cella som T6P. Videre kløyvning involverer trehalase, T6P-hydrolase eller andre enzymer.
3. Trehalose opptak via PTS-systemet som T6P, med videre kløyvning av enzymet trehalose-6-phosphate-phosphorylase. Den enzymatiske kløyvningen etterlater et fosforylert trehaloseprodukt som er vist benyttet av *Lactococcus lactis* (Jensen, Ampomah et al. 2005).

Trehalose hos rhizobia

Syntese og akkumulasjon av trehalose er detektert i forskjellige rhizobia som respons på miljømessig og fysiologisk stress, osmotisk press, mikroaerobe forhold og ved inngangen til stasjonær fase (Engelhard 2004).

To synteseveier for sakkaridet er kjent hos rhizobier:

1. MOTS-pathway
2. Ots/AB-pathway

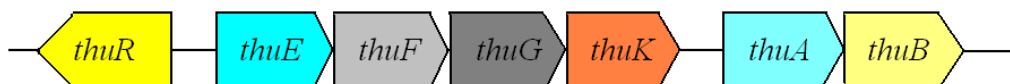
I legumrotnoduler opptrer trehalose som en av de vanligste karbohydratene, og er tenkt syntetisert av bakteroider. Mulig induksjonsfaktor kan være det lave O₂-nivået internt i rotnodulene. Akkumulert trehalose i rotnoduler øker resistensen mot tørke i nodulerte erteplanter, og det er derfor foreslått at symbiotisk trehaloseproduksjon kan være fordelaktig for vertsplanten ved å øke resistens mot tørke (Farias-Rodriguez, Mellor et al. 1998).

Identifisering og karakterisering av trehaloserelaterte gener hos *Rhizobia* sp.

Tidligere arbeid med *Sinorhizobium meliloti* har ført til identifisering og karakterisering av trehalose-induserbare gener hos arten, *thuA* og *thuB*. *S. meliloti* er en rhizobie med evne til å bruke trehalose som eneste karbonkilde for vekst, og arten inngår symbioser med legumen *Medicago sativa*. Genproduktene fra *thuAB* gir opphav til henholdsvis ThuA og ThuB-proteiner, som er vist å være involvert i katabolisme og transport av trehalose gjennom et ATP- bindings "cassette" transport system (ABC-transport system) (Jensen, Bhuvaneshwari et al. 1999). *thuAB*-ortologer er funnet i flere medlemmer av familien *Rhizobiaceae*, blant annet i *Rhizobium leguminosarum*, *Mesorhizobium loti*, *R. etli*, og den plantepatogene *Agrobacterium tumefaciens* (Ampomah 2007).

thuA og *thuB*'s genprodukter er ikke essensielle for infeksjonsprosessen for *S. meliloti*, men det er indikasjoner på at proteinene spiller en viktig rolle under kolonisering av røtter, samt konkurranse om noder. Det er påvist induksjon av genene under rotkolonisering, i bøyde rothår og i infeksjonstråder, noe som gir invaderende celler tilgjengelig trehalose tenkt benyttet for å håndtere høy temperatur, sult, tørke, osmotisk-, oksidativt- og etylenrelatert stress. Eksperimenter har vist at trehalose kan syntetiseres og akkumuleres hos *S. meliloti* under stress. I tillegg kan trehalose under slike forhold bedre proliferasjon av cellene (Jensen, Ampomah et al. 2005).

thuA og *thuB* inngår i *thuEFGKAB*-operonet som illustreres i figur 4. Operonet koder for et ABC-transport system for trehalose, maltose og sukrose. Mutasjoner i *thuA* eller *thuB* gjorde *S. meliloti* (Sm7023 og Sm7024) ute av stand til vekst med trehalose, maltitol, palatinose og leukrose som hovedkarbonkilde, og trehalose ble funnet til å akkumuleres i cellene (Ampomah 2007). Mutantstammene hadde dårligere rotkolonisering enn villtypestammen, men hadde like god eller bedre infeksjons- og noduleringsrate enn foreldrestammen.



Figur 4: Viser *thuEFGK*-operonet karakterisert i *S. meliloti* på megaplasmidet *pSymB*. Operonet inneholder gener med funksjoner involvert i utnyttelse og transport av trehalose. ThuR er den regulatoriske enheten for trehalose transportsystem, ThuE et trehalose/maltose bindingsprotein, ThuF et trehalose transporter permease protein, ThuG et trehalose transporter ATP-bindings protein, ThuA og ThuB er trehalose utnyttelsesproteiner. Pil angir transkripsjonsretning av individuelle gener (Jensen 1999).

Ved fullsekvensering av genomet til den nært beslektede *Mesorhizobium loti* MAFF303099, ble det funnet åpne leserammer med delvis homologi til *thuA/thuB*; *mll3105* og *mll3104*. Genene til *M. loti*'s *thuA* og *thuB* homologer er henholdsvis 71 % og 54 % identiske med *S. meliloti*'s *thuA* og *thuB* gener (Ampomah 2007). Selv om funksjonene til *mll3104* og *mll3105* ikke er bekreftet, har man indikasjoner på at de deltar i karbonmetabolisme hos *M. loti*. Ved mutasjon i *mll3105* hos arten, viste mutantstammen redusert vekst med sukrose, maltose og trehalose som karbonkilder, i forhold til villtypen. Ved et annet tilfelle ble *mll3104*- og *mll3105*-gener satt inn i en *S. meliloti* med mutasjon i *thuA*, og mutanten gjenvant da evnen til vekst med trehalose, palatinose og leukrose som karbonkilder. Resultatet indikerer en komplementering som tilsier at *mll3104* og *mll3105* er involvert i trehalosemetabolisme i *M. loti*. Selv om genene har vist seg å være deltagende i karbohydratmetabolisme, er lokalisasjonen av operonet noe annerledes enn i *S. meliloti*. *mll3104* og *mll3105* i *M. loti* er funnet lokalisert proksimalt til gener for aminosyremetabolisme, så det kan ikke utelukkes at genproduktene har roller i denne typen metabolisme.

Da dette studiet ble påbegynt antok man at *mll3104* og *mll3105* er proteinkodende DNA, med funksjoner i transport og katabolisme av karbohydrater. Det er hittil ikke identifisert egne systemer for transport og utnyttelse av trehalose hos *M. loti*, så studiets arbeid vil fokusere på å frembringe mer kunnskap om *thuAB* ortologene i *M. loti*.

Hovedmål for studiet

Overordnet mål for dette studiet er forsøk på å utrede funksjoner til *M. loti*'s antatte *mll3104* og *mll3105* proteiner. Det foreligger indikasjoner på at genene kan være ortologer av *S. meliloti*'s ThuA og ThuB proteiner, selv om genlokalisasjon på genomene er noe ulik mellom de to artene. Per dags dato har ingen ekspresjonsstudier vært gjort for genene, og studiet tar sikte på å prøve ut velkjente metoder som fungerer for uttrykk og isolering av *mll3104*'s og *mll3105*'s respektive genprodukter. Eksperimenter gjennomføres ved å bruke forskjellige molekylærgenetiske metoder som verktøy for:

- Identifisering og isolering av *mll3104* og *mll3105* fra *M. loti*.
- Ligering av *mll3104* og *mll3105* i passende vektor for transformasjon inn i ekspresjonsstamme.
- Storskala ekspresjon og opprensing av hypotetiske Mll3104 og Mll3105 proteiner.
- Mulig karakterisering av funksjonen til uttrykte proteiner.

Materiale

Bakterielle stammer

I dette studiet er *Mesorhizobium loti* MAFF303099 benyttet som referansestamme da denne er positiv for *mll3104* og *mll3105*. Arten danner determinante noder, blant annet modellegumen *Lotus japonicus* (Saeki and Kouchi 2000; Uchiumi 2004; Townsend 2006). Stammen ble fullsekvensert i 2000, og består av et enkelt kromosom på 7.036.071 bp, samt to plasmider på henholdsvis 351.911 bp og 208.315 bp (Hattori, Omori et al. 2002).

All data om referansestammen er hentet fra

<http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genome&Cmd=ShowDetailView&TermToSearch=173> og (Kaneko, Nakamura et al. 2000). Andre anvendte stammer er del av kommersielle kit fra Invitrogen life technologies.

Tabell 2: Bakterielle stammer og plasmider brukt i studiet.

Bakterielle stammer og plasmider	Relevante karakteristikk	Kilde/referanse
Bakterielle stammer		
<i>Mesorhizobium loti</i> MAFF3030900	Positiv for <i>mll3104</i> og <i>mll3105</i>	Hiroshi Taniguchi, Nat. Institute of Agrobiological Sciences, MAFF GENE BANK, Japan
One Shot <i>Escherichia coli</i> TOP10	Elektrokompetente celler. Brukt for generell kloning og β -galaktosidase seleksjon uten IPTG.	Invitrogen life technologies
<i>Escherichia coli</i> BL21 Star TM	Kjemokompetent ekspresjonsstamme.	Invitrogen life technologies
Plasmider		
pCR2.1 TOPO	Multiple kloningsseter, T7 promoter priming sete. Kanamycin og ampicillin resistens. Seleksjon med X-gal. IPTG induert uttrykk	Invitrogen life technologies
pUC19	Positiv kontroll plasmid	Invitrogen life technologies
pET101/D-TOPO	T7lac promoter for IPTG induert masseuttrykk av innsatt gen. C-terminalt plassert 6xHis markør. Ampicillin resistens for seleksjon. pBR322 opprinnelse for lav-kopi replikasjon og opprettholdelse i <i>E. coli</i>	Invitrogen life technologies

Det er noe ulik lagringsprosedyre for de ulike stammene. *M. loti* ble lagret i LB-medium med 15 % glyserol ved -20 °C. *E. coli* stammer fulgte med kit, og ble lagret ved -80 °C. Plasmider ble lagret ved -20 °C. Ved behov for ferske bakterier, ble det foretatt utstryk på agarplater (se tabell 3) med inkubasjon over natt ved 37 °C. Bortsett fra S.O.C-mediet, er alle dyrknings- og lagringsmedia selvprodusert.

Tabell 3: Vekstmedia brukt i dette studiet.

Medium	Formål	Innhold
TY-medium	Dyrkning av <i>M. loti</i>	5 g Trypton, 3 g gjærekstrakt, 0.87 g CaCl ₂ x2H ₂ O, 17 g Bactoagar, 1000 ml dH ₂ O
Luria Bertani agar (LB)	Dyrkning av <i>E. coli</i> (TOP10 og BL21 Star)	10 g Trypton, 1000 ml dH ₂ O, 5 g gjærekstrakt, 10 g NaCl. Dersom fast vekstmedium er ønskelig tilsettes 7 g Bactoagar.
Frysemedium for <i>E. coli</i>-stammer brukt i studiet	Lagring av stammer (stock)	85 % LB-medium uten agarose, 15 % glyserol
S.O.C-medium	Transformasjon av TOP10 elektrokompetente celler	2 % trypton, 0.5 % gjærekstrakt, 10 mM NaCl, 2.5 mM KCl, 10 mM MgSO ₄ , 20 mM glukose.

Tabell 4 viser forskjellige reagenser og kjemikalier brukt under studiet. Dersom en reagens er en del av et kommersielt kit, vil kitet refereres til i sin helhet under tabell 5 som lister opp alle kits og enzymer som er anvendt i eksperimenter under dette studiet.

Tabell 4: Kjemikalier og reagenser brukt i dette studiet.

Formål	Kjemikalie/reagens	Innhold/Produsent
Seleksjon ved dyrkning	Fosfomycin 50 µg/ml	Sigma P5396
	Kanamycin 50 µg/ml	Sigma-Aldrich K4000
	Ampicillin 50 µg/ml	Sigma A9518
	X-Gal	Invitrogen life technologies Kat.nr 15520-034
Elektroforese av DNA	1xTBE	10.8 g/L Tris Base, 5.5 g/L Borat, 0.93 g/L EDTA, dH ₂ O, pH 8
	Agarose LE, Analytical Grade	Promega
	Gel loading buffer x6	MBI-Fermentas R0611
	Ethidium Bromid (0.5 µg/ml)	EuroClone
	GeneRuler™ DNA Ladder Mix 0.5 µg/µl	MBI-Fermentas

Ekspresjon av <i>mll31054/mll3105</i> i BL21 Star <i>E. coli</i> celler	IPTG (Isopropyl-thio-2-D-galactopyranoside)	Invitrogen life technologies Kat. nr. 15529-019
Indusert ekspresjon av BL21 Star celler	Imidazole 1202	Aldrich
Protein ekstrahering	B-PER® Bacterial Protein Extraction Reagent	PIERCE Nr 78243
SDS-PAGE	Benchmark, protein marker 5x Tris-Glycin	Invitrogen life technologies
	Prestained SDS-PAGE Standard, Broad range	Bio-Rad
Rensing av His-merkede proteiner	Fosfatbuffer	KH ₂ PO ₄ 50 mM NaCl 0.5 M
	ProBond™ Resin	Invitrogen, 46-0019

Tabell 5: Kit og enzymer brukt under dette studiet

Metode	Reagenser, kits og enzymer	Produsent
PCR	Dynazym Polymerase	Finnzymes
	Dynazym Buffer 10x	Finnzymes
	Pfx50 Polymerase	Invitrogen life technologies kat nr 12355-012
	Pfx50 Buffer 10x	Invitrogen life technologies kat nr 12355-012
	dNTP-miks	Invitrogen life technologies
DNA ekstrahering fra gel	QIAquick Gel extraction Kit (50)	QIAGEN
	Kit	Kat.nr 28704
Transformasjon av TOP10 <i>E. coli</i> celler	One Shot	Invitrogen life technologies
	Kit	Kat nr C4040-03
Plasmidrensing	QIAprep Miniprep kit (250)	QIAGEN
	Kit	Kat.nr 27106
TOPO Directional Kloning	Champion™ pET Directional TOPO® Expression Kit	Invitrogen life technologies
	Kit	K101-01
DNA sekvensering	BigDye Terminator v3.1 Cycle Sequencing Kit	Applied Biosystems Kat.nr 4337455

Tabell 6 viser hvilke primere som er brukt for DNA amplifisering under studiet. Forward og revers primere for *mll3104* og *mll3105* er designet selv. Andre primere tilhørende kit.

Tabell 6: Oligonukleotid primere brukt under dette studiet.

Primer	Sekvens	Produsent
<i>mll3104-F (100 mM)</i>	5'-CACCATGCATCGCCTGCTCTTGCTC-3'	EUROGENTEC S.A
<i>mll3104-R (100 mM)</i>	5'-TGCGCCCATCGAGATCGGCAGCCTG-3'	EUROGENTEC S.A
<i>mll3105-F (100 mM)</i>	5'-CACCATGCCGATAAAACGTGTTGTC-3'	EUROGENTEC S.A
<i>mll3105-R (100 mM)</i>	5'-TGCACTCAACCCCTTGTCGCCATCG-3'	EUROGENTEC S.A
T7-R (0.1 µg/ml)	5'-TAGTTATTGCTCAGCGGTGG-3'	Invitrogen life technologies
T7-F (0.1 µg/ml)	5'-TTCCTTCGACGCTAACCTG-3'	Invitrogen life technologies

Tabell 7 viser annet relevant eksperimentelt utstyr anvendt under dette studiet.

Tabell 7: Eksperimentelt utstyr brukt under studiet.

Utstyr		Produsent
PCR-maskin	DNA Engine Gradient Cycler PTC-200	MJ Research
Elektroporator	Pulse Controller Gene Pulser	Bio-Rad
Foto system	Gel Doc 2000	Bio-Rad
Spektrofotometer for måling av protein konsentrasjon	SpectraMAX 250, SoftMaxPro 2.41 software	
Primer software	Oligo primer Analysis Software, Version 6.71	Molecular Biology insights, Inc.
Måling av DNA konsentrasjon	Nano Drop 1000	
SDS-PAGE system	Power Pac 300	Bio-Rad
Sekvenserings software	Sequence Skanner Software v1.0	Applied Biosystems Genetic Analyzers
DNA sekvensering	3130xl Genetic Analyzers	Applied Biosystems
Fraksjonssamler	Model 201 Fraction controller	GILSON

Metoder

Metodene som er anvendt under studiet er som regel hentet fra leverandørens manual. Det angis referanse for hver metode dersom prosedyren er fra et kommersielt kit.

Primer design

Primere er oligonukleotidsekvenser som er designet for hybridisering med regioner av DNA som flankerer målsekvensen. For at en primer skal binde effektivt til templatet er det visse kriterier som bør oppfylles. Den optimale primerlengden er mellom 18 og 25 nukleotider, og GC innholdet bør være mellom 50 % og 60 %. For å unngå dannelse av "hairpin" (hårnål) strukturer bør nukleotidrepetisjoner ikke forekomme, og designet må være av en slik karakter at dannelse av primer-dimerer unngås. Det er en fordel om primerne har samme smeltetemperatur (T_m). Dersom T_m variasjonen er for stor (>5 °C), vil oligonukleotidene kunne bindes til templatet ved ulike temperaturer, noe som kan føre til for få eller manglende produkter, eller uønskede amplifiserte segmenter. T_m for egne primere er oppgitt i tabell 8.

Tabell 8: Viser smeltetemperatur (T_m) for primere brukt under studiet.

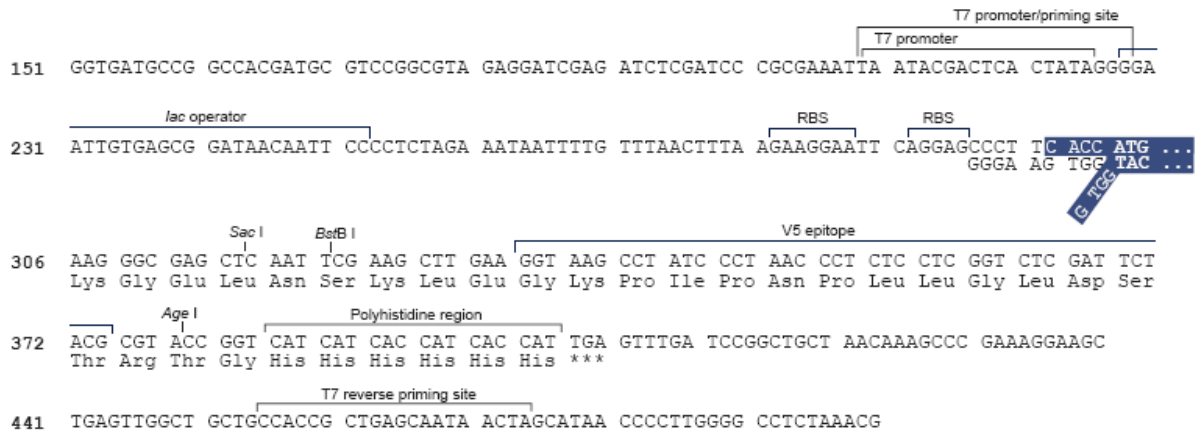
Primer	Temperatur (°C)
<i>mll3104-F</i>	2AT+4 GC: 80 °C og % GC: 58,1 °C.
<i>mll3104-R</i>	2AT+4 GC: 84 °C og % GC: 60, 8 °C
<i>mll3105-F</i>	2AT+4 GC: 74 °C og % GC: 52, 6 °C
<i>mll3105-R</i>	2AT+4 GC: 80 °C og % GC: 57, 5 °C

Målnukleotid-sekvensene (opprinnelig *M. loti* gener *mll3104* og *mll3105*) som er utgangspunkt for primerdesignet hentes fra søkefunksjoner i NCBI. Systemet er tilgjengelig ved: <http://www.ncbi.nlm.nih.gov/Genbank/index.html>.

Originale målsekvenser limes inn i Oligo Primer Analysis programvare. Programvaren angir mulige primere, intern stabilitet, kalkulerer T_m , og varsler om potensielle hårnålsformasjoner og dannelse av primer-dimerer. Sekvensene av primerne brukt under studiet er angitt i tabell 6.

Primerne brukt under dette studiet ble designet med tanke på amplifisering av målgener både fra *M. loti* genomsk templat, og for amplifisering av kloningsprodukter etter innsett i vektor. Ved bruk av kommersielt kit fra Invitrogen, designes primeren på en slik måte at genproduktet liggeres inn i vektor pET101/D-TOPO med et 5'-CACC overheng etter vektor

leverandørens anvisning, og med startkodon (ATG) for formylmethionin direkte etter dette (se figur 5). Målsekvensenes termineringskodon (TGA), ble endret ved at GC erstattet TG for å unngå prematur terminering av produktet. Nye baser gir aminosyrene alanin eller tyrosin.



Figur 5: Viser kloningssettet for pET101/D-TOPO. Start av innsett av egne fragmenter er markert i sort ved basepar 292. Ved induksjon fra IPTG vil transkripsjon starte ved T7 promoter sete. Den markerte polyhistidin regionen skal senere anvendes som markør for rensing av det transkriberte proteinet. Figur hentet fra (Invitrogen-life-technologies 2004).

PCR- Polymerase Chain Reaction

PCR tar utgangspunkt i vanlig DNA-replikasjon med forlengelse av DNAet, og anvendes på flere molekylærgenetiske områder der kloning og annen DNA-manipulering er nødvendig.

PCR kan kort beskrives som repetitiv kopiering av seksjoner med dobbeltrådet DNA (dsDNA), og metodens viktigste gevinst er muligheten for replikasjon av store mengder spesifikke DNA-sekvenser på kort tid som senere kan klones, sekvenseres og analyseres etc. (Reece 2004).

I dette studiet er det brukt to forskjellige leverandører for PCR buffer og polymerase, der det ene settet (Dynazym) er uten nøyaktig korrekturlesing (proofreading), se tabell 5 for detaljer. Erfaringsmessig har det vist seg vanskelig å amplifisere genomisk templat fra *M. loti* direkte ved bruk av Pfx50 polymerase og buffer, med proofreading aktivitet. Innledningsvis vil derfor all amplifisering av *M. loti* templat foregå ved bruk av Dynazym buffer og polymerase, med hensikt å ligere PCR-produktene inn i vektor pCR2.1-TOPO.

Etter innsett i vektoren vil genfragmentene amplifiseres videre ved bruk av Pfx50, som har en basekorrigerende 3'→5'-eksonuklease aktivitet. Neste mål for amplifisering av *mll3104/mll3105*-fragmenter vil være etter ligering i ekspresjonsvektor pET101/D-TOPO. Ved dette stadiet er det essensielt at basesammensetningen i innsettet er som originalekvensen, da mulige feil kan gjøre at senere uttrykk av genprodukter blir ufunksjonelle eller fraværende.

Prosedyre:

Standardoppsett for amplifisering av *mll3104* og *mll3105* er følgende:

1. *mll3104*-templat med tilhørende primere
2. Negativ kontroll for *mll3104* med dH₂O og primere for *mll3104*
3. *mll3105*-templat med tilhørende primere
4. Negativ kontroll for *mll3105* med dH₂O og primere for *mll3105*

Oppsett for standard mastermiksgjengis i tabell 9, og PCR-program i tabell 10.

Tabell 9: Viser standard mastermiksgjengis for PCR med bruk av Dynazym enzymer. Volum ganges opp med antall prøver. Mastermiksgjengis lages og holdes på is til bruk.

Reagens	Volum
Dynazym PCR buffer 10x	5 µl
Dynazym dNTP	1,5 µl
Dynazym DNA polymerase	1 µl
MgSO ₄ (10 mM)	0,5 µl
5 ng/µl templat (<i>mll3104</i> eller <i>mll3105</i>)	10 µl
Primere (<i>mll3104/mll3105</i> forward eller revers)	0,75 µl
dH ₂ O	Til 50 µl

Tabell 10: Viser PCR-syklus for reaksjoner med Dynazym enzymer. Trinn 2-4 gjentas 30 ganger.

Reaksjon	Temperatur	Tid
1: Denaturering	95 °C	7 minutter
2: Denaturering	95 °C	1 minutt
3: Annealing	60 °C	1 minutt
4: Extension	72 °C	2 minutter
5: Extension	72 °C	5 minutter
6: Nedkjøling	4 °C	∞

Standard mastermiks for reaksjoner med Pfx50 enzymer oppgis i tabell 11, og PCR program i tabell 12.

Tabell 11: Standard mastermiks for PCR med Pfx50 (volum ganges opp med antall prøver).

Reagens	Volum
Pfx50 PCR buffer 10x	5 µl
dNTP	1,5 µl
Pfx50 DNA polymerase	1 µl
MgSO ₄ (50 mM)	1 µl
5 ng/µl templat (<i>mll3104</i> eller <i>mll3105</i>)	10 µl
Primere (<i>mll3104/mll3105</i> forward eller revers)	0,75 µl
dH ₂ O	Til 50 µl

Tabell 12: Viser PCR-program der Pfx50 enzymer brukes. Trinn 2-4 gjentas 30 ganger.

Reaksjon	Temperatur	Tid
1: Denaturering	95 °C	7 minutter
2: Denaturering	95 °C	1 minutt
3: Annealing	60 °C	1 minutt
4: Extension	68 °C	2 minutter
5: Extension	68 °C	5 minutter
6: Nedkjøling	4 °C	∞

Gel-elektroforese:

PCR produkter (5 µl) separeres på 0.8 % agarosegel ved 80 V i ca 1.5 timer. Ved elektroforese med hensikt å rense DNA fra gel, tilsettes 25 µl i brønnene.

Agarosegeler tilsettes 1 µl/ml EtBr (0.5 µg/ml). Visualisering skjer ved bruk av Quantity One Gel Doc 2000 (Bio-Rad). For alle geler blir GeneRuler™ DNA Ladder Mix (MBI-Fermentas) brukt for å bestemme størrelse på produkter.

Rensing av DNA fra agarosegel ved bruk av QIAquick Gel Extraction Kit

Metoden er designet for rask DNA opprensing, og ekstraherer DNA-fragmenter fra 70 bp til 10 kb fra agarosegeler i TAE- eller TBE-buffer, samt DNA renses fra enzymatiske reaksjoner. QIAquick systemet benytter seg av spinkolonne-teknologi sammen med en silikagel membran

der DNA absorberes vha høy saltkonsentrasjon og riktig pH i bufferen, mens kontaminanter faller gjennom kolonnen. Absorpsjonen av nukleinsyrer til silikagelmembraner skjer kun ved tilstedeværelse av salter som endrer vannets struktur. DNA blir ytterligere renses ved bruk av vaskebufferen som medfølger kitet. Prosedyren avsluttes ved at renses DNA elueres med egen buffer eller dH₂O. All sentrifugering gjøres ved 17 900 x g.

Prosedyre:

1. Skjær DNA fragmentet fra agarosegelen med en ren skalpell, og minimer gelbiten så mye som mulig for å fjerne overflødig agarose.
2. Vei gelbiten i eppendorfrør og tilsett 3 volum QG buffer til 1 volum av gel (100mg ~ 100µl).
For >2 % agarose, tilsett 6 volum av QG buffer. Maksimal gel mengde per spinnkolonne er 400 mg. For gelbiter større enn dette, bruk mer enn en QIAquick kolonne.
3. Inkuber i 10 min ved 50 °C eller til gelen er fullstendig oppløst. For å bidra til hurtig oppløsning av gelen, er det mulig å vortexe røret hvert 2-3 minutt under inkubasjonen. Viktig: Løs opp agarosen fullstendig.
4. Etter at gelen er oppløst, sjekk at fargen er gul. Dette indikerer en optimal pH (7.5). Fargen skal være lik QG buffer uten oppløst agarose. Dersom fargen ikke tilsvarer ovennevnte, tilsett 10 µl 3 M natrium acetat, pH 5.0, og miks.
DNA absorpsjon til QIAquick membranen er kun effektiv ved pH ≤7.5. QG buffer inneholder en pH-indikator som er gul ved pH 7,5 og oransje eller fiolett ved høyere pH, noe som gir enkel pH bestemmelse for DNA-binding.
5. Plasser en QIAquick spinnkolonne i et 2 ml oppsamlingsrør.
6. Tilsett DNA-prøven i spinnkolonnen og sentrifuger 1 min. Dette for at DNA skal bindes. Maksimal kapasitet til spinnkolonnen er 800 µl, for prøver større enn dette, sett på mer prøve og spinn flere ganger.
7. Kast væsken som har gått gjennom kolonnen, og plasser spinnkolonnen tilbake i det samme røret.
8. Valgfritt: Tilsett 500 µl av QG buffer til kolonnen og sentrifuger 1 min. Dette trinnet fjerner alle spor av agarosen, men er imidlertid kun nødvendig dersom prøvematerialet skal brukes for direkte sekvensering, *in vitro* transkripsjon eller mikroinjeksjon.
9. For vasking: Tilsett 750 µl PE buffer til spinnkolonnen og sentrifuger 1 min.

NB: Dersom DNA prøven skal brukes i saltsensitive reaksjoner som for eksempel blunt-end liggering eller direkte sekvensering, la spinnkolonnen stå i 2-5 min etter tilsetning av PE bufferen før sentrifugering.

10. Kast væsken som har passert spinnkolonnen og sentrifuger 1 min.

Viktig: Rester av etanol fra PE bufferen vil ikke fjernes fullstendig med mindre væsken fjernes fra røret før den ekstra sentrifugeringen.

11. Overfør prøven til et 1.5 ml mikrosentrifugerør.

12. For å eluere DNA-prøven, tilsett 50 µl av EB buffer (10 mM Tris-Cl, pH 8.5) eller H₂O til sentrum av QIAquick membranen og sentrifuger kolonnen i 1 minutt.

Alternativt, for høyere DNA konsentrasjon, tilsett 30 µl elueringsbuffer til midten av QIAquick membranen, la kolonnen stå i 1 minutt, og sentrifuger så 1 minutt.

Viktig: Vær sikker på at elueringsbufferen settes direkte på QIAquick membranen for fullstendig eluering av bundet DNA. Gjennomsnittlig elueringsvolum er 48 µl fra 50 µl buffer volum, og 28 µl fra 30 µl.

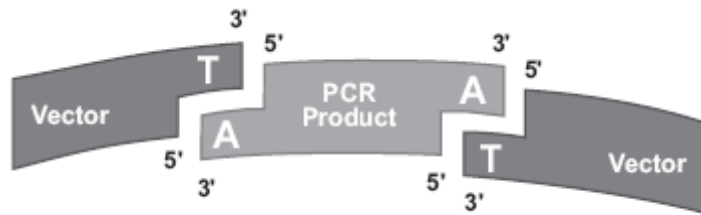
Elueringeffektiviteten er avhengig av pH. Maksimal effektivitet oppnås mellom pH 7.0 og 8.5. Ved bruk av vann, vær sikker på at pH-verdien er innenfor denne skalaen, og lagre DNA ved -20 °C ettersom DNA kan degradere ved fravær av et bufrende virkestoff. Opprensede DNA kan også elueres i TE (10 mM Tris-Cl, 1 mM EDTA, pH 8.0), men EDTA kan inhibere enzymatiske reaksjoner i ettertid.

(QIAGEN 2002)

TOPO TA Cloning

TOPO TA Cloning er designet for innsett av *Taq*-polymerase-amplifiserte PCR-produkter i en plasmidvektor som er supplert med:

- Et enkelt 3'-tymidin overheng for TA-kloningen
- En Topoisomerase 1 (top I) kovalent bundet til vektoren



Figur 6: Viser konseptet bak TA-kloning av PCR-produkter inn i vektoren (Invitrogen-life-technologies 2004).

Taq polymerase har en templat-uavhengig terminal transferaseaktivitet som tilsetter en enkel deoksyadenosine (A) til 3'enden av PCR-produktene. Den lineære vektoren har en enkel, overhengende 3'deoksytymidin (T) enhet som gjør at PCR innsettet ligeres effektivt med vektoren (figur 6). Top I er opprinnelig hentet fra *Vaccinia* virus, og binder til dupleks DNA ved spesifikke seter, og kløyver fosfodiester-delen etter 5'-CCCTT i den ene tråden. Salt (50 mM) tilsettes reaksjonen, dette forhindrer top I fra å danne potensielle kutt i DNA etter ligering.

Etter ligering av *mll3104* og *mll3105* i vektor pCR2.1-TOPO, vil plasmidet transformeres i elektrokompetente One Shot TOP10 *E. coli* celler som medfølger kitet. Et korrekt innsett vil ligeres inn den α -peptid kodende regionen (amino-terminal del) av enzymet β -galaktosidase i et *lacZ*-gen på vektor. β -galaktosidase er det enzymet som spalter X-gal, og gir galaktose og et 5-bromo-4-kloro-indoksyderivat. Sistnevnte vil spontant dimerisere og oksidere, og sluttproduktet er en uløselig blå farge som kommer til uttrykk i form av blå kolonier. Et innsett som sitter på riktig sted, vil gjøre β -galaktosidasen ufunksjonell ved at α -peptidet ikke inkluderes i proteinet. Resultatet er hvite kolonier som lett kan identifiseres (Reece 2004). Som positiv kontroll brukes pUC19, som selekteres for etter transformasjon med ampicillin (50 μ g/ml).

Tabell 13: Viser reaksjonsoppsett for ligering av egne genfragmenter i vektor pCR2.1-TOPO med bruk av elektrokompetente celler som medfølger kit.

Reagens	Volum
PCR produkt	0.5 µl til 4 µl
Saltløsning	1µl
Fortynnet saltløsning	--
dH ₂ O	til totalvolum 6 µl
TOPO vektor	1 µl
Totalvolum	6 µl

Prosedyre ligeringsreaksjon:

1. Reagenser (tabell 13) mikses forsiktig og inkuberes 5 minutter i romtemperatur (22-23 °C). For de fleste reaksjoner vil 5 minutter gi mange kolonier for analyse, men inkubasjonstiden kan varieres fra 30 sekunder til 30 minutter. For store innsett (>1kb) vil økt inkubasjonstid gi flere kolonier.
2. Plasser miksen på is for videre transformasjon. Dersom ikke transformasjon skal skje samme dag, kan kloningsreaksjonen fryses ned over natt ved -20 °C.

Transformasjon av One Shot TOP10 elektrokompetente *E. coli*

Prosedyre:

1. Tilsett 2 µl av kloningsreaksjonen i et rør med One Shot celler og miks forsiktig ved å knipse på røret med fingeren. IKKE pipettér opp og ned, da genomisk DNA tilstede i prøven kan kontaminere prøven.
2. Overfør løsningen forsiktig til en 0.1 cm kyvette, unngå bobledannelse.
3. Elektroporer prøvene ved bruk av egen protokoll for elektroporator (i dette studiet brukes Gene Pulser (Bio-Rad) og elektroporerer skjer ved 2.5 V, 200 OHMS og 25 µ FD)
4. Tilsett umiddelbart 250 µl romtemperert S.O.C. medium.
5. Overfør løsningen til 1.5 ml eppendorfrør, og plasser rørene på risting 1 time ved 37 °C. Dette for uttrykk av antibiotika resistensgener.

6. Spre ut 30 µl celler fra hver transformasjon på en forvarmede agarplater (LB plater med 40 mg/ml X-gal og 50 µg/ml kanamycin) med inkubasjon over natt ved 37 °C.
7. En effektiv kloningsreaksjon kan gi flere hundre kolonier. Plukk hvite kolonier for analyse.

(Invitrogen-life-technologies 2006)

Ut fra dyrkningsresultater på selektive plater velges hvite enkeltkolonier for videre dyrkning i 3 ml LB-medium tilsatt 50 µg/ml kanamycin (200 rpm ved 37 °C). Overnattkultur spinnes ned og fordeles i 1.5 ml eppendorfrør ved 4000 rpm i 1 minutt. Supernatanten fjernes før plasmidrensing. Pelletter kan fryses ved -20 °C.

Plasmidrensing ved bruk av QIAprep Miniprep kit

Etter liggering i pCR2.1-TOPO skal *mll3104* og *mll3105*-fragmenter amplifiseres fra vektor ved bruk av en polymerase med korrekturaktivitet. Før PCR reaksjoner settes opp, skal plasmid-DNA isoleres fra cellelysat. For enkelhets skyld vil vektor med ligert innsett refereres til etter type innsett; pCRmll3104 og pCRmll3105. All sentrifugering skjer ved 13.000 rpm.

Prosedyre:

1. Resuspender pellet med bakterieceller i 250 µl P1 buffer, og overfør til et mikrosentrifugerør.
2. Tilsett 250 µl P2 buffer og miks røret ved å vende det 4-6 ganger.
3. Tilsett 350 µl N3 buffer og miks umiddelbart ved å vende røret opptil 10 ganger. Løsningen skal bli blakket.
4. Sentrifuger 10 minutter.
5. Overfør forsiktig supernatanten fra trinn 4 til en QIAquick spinnkolonne.
6. Sentrifuger 60 sekunder, kast væske som har strømmet gjennom.
7. Vask spinnkolonnen ved å tilsette 750 µl PE buffer og sentrifuger 60 sekunder. Kast væske som har gått gjennom kolonnen, og sentrifuger ytterligere 60 sekunder.
8. Overfør spinnkolonnen til et rent 1.5 ml eppendorfrør. For eluering: Tilsett 50 µl dH₂O til sentrum av kolonnen, og vent 1 minutt før sentrifugering 1 minutt som over. Renset plasmid-DNA fryses ved -20 °C

(QIAGEN 2005)

For å undersøke om plasmidrensingen er vellykket utføres en gel-elektroforese. Avhengig av hvilken tilstand plasmidet er i, forventes bånd av ulik størrelse. PCR med Pfx50 enzymer og påfølgende DNA-rensing av PCR-produkter fra agarosegel utføres som ved tidligere angitt prosedyre.

Ligering av *mll3104* og *mll3105* PCR-produkter i pET101/D-TOPO

PCR-produktene amplifisert med Pfx50 polymerase, skal ligeres i ny vektor, pET101/D-TOPO. Vektoren er konstruert for høyt ekspresjonsnivå av ligerte fragmenter, og hensikten med ligeringsreaksjonen er å uttrykke *mll3104* og *mll3105* etter transformasjon i *E. coli* ekspresjonsstamme BL21 StarTM. Etter ligering med henholdsvis *mll3104/mll3105*-sekvenser, vil vektor navngis etter innsett: pETmll3104 og pETmll3105.

Vektor og stamme inkluderes i samme kit. Oppsettet for ligeringsreaksjonen er lik som for forrige ligering (TOPO TA Cloning).

Transformasjon av pET101/D-TOPO i *E. coli* BL21 Star

BL21 Star er en kjemokompetent *E. coli* stamme som medfølger ekspresjonskitet (tabell 2). Enkeltrør med celler (et for hver transformasjon) tines på is. Det settes opp 4 reaksjoner som blandes på is; pETmll3104, pETmll3105, negativ kontroll (dH₂O) og positiv kontroll (pUC19). Reaksjoner inkuberes 5 minutter i romtemperatur før reaksjonen igjen settes på is (volum og type reagenser er som for tabell 8):

Prosedyre:

1. Tilsett 3 µl av kloningsreaksjonen i rør med kjemokompetente celler, og miks forsiktig ved å knipse på røret, IKKE pipettér opp og ned, cellene er sensitive!
2. Inkubér rør på is i 10 minutter.
3. Overfør rør til vannbad, 42 °C i 30 sekunder, og sett på is igjen.
4. Tilsett 250 µl S.O.C.-medium i hvert rør (nytt totalvolum 300 µl)
5. Inkubér rør 1 time ved 37 °C og 200 rpm.
6. Plat ut 30 µl av kultur fra hver transformasjon på LB-plater tilsatt 50 µg/ml ampicillin, og dyrk over natt ved 37 °C. Enkeltkolonier med positive transformanter dyrkes videre i 10 ml LB tilsatt antibiotika over natt, og spinnes ned. Pellet fryses ved -20 °C.

DNA-sekvensering

Sekvenseringsreaksjonen utføres for å bestemme den nøyaktige sekvensen av baser som utgjør *mll3104*- og *mll3105*-fragmentene amplifisert fra ekspresjonsvektoren. Reaksjonen baseres på at forhåndsfargede dideoksynukleotid trifosfater (ddNTP) molekyler, sammen med ordinære deoksynukleotid trifosfater (dNTP) inkorporeres i voksende DNA-kjeder som ved vanlig DNA replikasjon. 3'- ddNTP-molekylene mangler en 3'-OH-gruppe ved polymeriseringssetet, og vil derfor terminere DNA-syntesen. Fargestrukturene som er festet til ddNTP-molekylene inneholder en donor- og en akseptorfarge, og kalles BigDye™ terminatorer. Under DNA-syntese vil fragmenter av varierende størrelse produseres, som hver korresponderer til punktet der ddNTP ble innsatt. En argon ione-laser vil under reaksjonen eksitere donorfargen, og fargen vil effektivt videreføre energien som utløses under eksitasjonen til en av de fire akseptorfargene, som har hver sitt distinktive emisjons spektrum. Hvert av de 4 ulike ddNTP-molekylene er merket med egen akseptorfarge, så DNA-fragmenter med forskjellig ddNTP vil fluoresere ved forskjellige bølgelengder. Sekvenseringsreaksjonen utføres i ett enkelt rør, og produktene separeres på et kapillærrør inneholdende en gelmatriks (kapillær elektroforese). Intensiteten og bølgelengden av den fluorescente emisjonen måles ettersom DNA-fragmentene beveges forbi laseren og en fluorescent detektor, og informasjonen overføres direkte til en computer med passende software. Her vil de fluorescente mønstrene konverteres til sekvenser med DNA-baser. Før reaksjonen utføres vil DNA prøvene felles ut og denatureres før de settes til sekvenseringsapparatet, 3130xl Genetic Analyzers (Applied Biosystems) som separerer, detekterer og analyserer de fluorescent-merkede DNA-fragmentene ved kapillær elektroforese. Molekyler fra prøvene injiseres elektroforetisk til tynne silika-fuserte kapillærer, og DNA-fragmentene migrerer mot kapillærens motsatte ende. Som ved annen elektroforese vil korte fragmenter vandre raskest og dermed lengst (Reece 2004).

Prosedyre:

Ved dette studiet blir ABI PRISM™ BigDye™ Terminator v3.1 Cycle Sequencing Reaction Kit brukt for sekvenseringen, sekvenseringsreaksjonen og utfelling utføres ved Universitetssykehuset i Nord-Norge. Reagenser og volum oppgis i tabell 14 og PCR-program i tabell 15.

Det settes opp 4 reaksjoner på is:

- a) Renset plasmid med *mll3104*-fragmenter og T7 forward primer
- b) Renset plasmid med *mll3104*-fragmenter og T7 revers primer
- c) Renset plasmid med *mll3105*-fragmenter og T7 forward primer
- d) Renset plasmid med *mll305*-fragmenter og T7 revers primer

Tabell 14: Viser oppsett for standard volum av reagenser for DNA-sekvenserings reaksjon ved bruk av Big Dye enzymer. Volum ganges opp med antall prøver.

Reagens	Volum
Big Dye 3.1	4 µl
5x Big Dye Sequencing Buffer	2 µl
Primer (T7 revers/forward)	3.2 pmol
Templat	200-300 ng ¹
dH₂O	til totalvolum 20 µl

1: Innledningsvis ble det brukt 200 ng templat, ved optimalisering økte konsentrasjon av templat til 300 ng.

Tabell 15: Viser PCR-program ved bruk av Big Dye enzymer. Punkt 2-4 gjentas 25 ganger (ABI-PRISM 2002)

Reaksjon	Temperatur	Tid
1: Denaturering	96 °C	1 minutt
2: Denaturering	96 °C	10 sekunder
3: Annealing	50 °C	5 sekunder
4: Extension	60 °C	4 minutter
5: Nedkjøling	4 °C	∞

Pilotekspresjon av Mll3104- og Mll3105-proteiner med IPTG som transkripsjonsaktivator

Dette eksperimentet utføres for uttrykk av Mll3104- og Mll3105-proteiner. Forut for ekspresjonseksperimentet er pETmll3104 og pETmll3105 transformert inn i separate BL21 Star *E. coli* celler konstruert for ekspresjon av T7 promoter regulerte gener. Vektoren bruker elementer fra bakteriofag T7 for å kontrollere ekspresjon av heterologe gener i *E. coli*. I pET101-TOPO vektorer vil uttrykk av det innsatte genet kontrolleres av en sterk bakteriofag T7 promoter som inneholder en *lac*-operator sekvens plassert nedstrøms for T7 promoter, og T7 RNA polymerase gjenkjenner spesifikt denne promoteren. Ved bruk av BL21 Star *E. coli* celler vil T7 RNA polymerase følge med stammen og sørge for regulert uttrykk av klonede gener. IPTG (isopropyl β -D-1-thiogalaktopyranosid) er en laktoseanalog som induserer transkripsjon av gener bak *lac*-operatoren, vanligvis uttrykk av β -galaktosidase. Ved induksjon med IPTG vil egne innsett transkriberes, da lokalisasjonen er 75 bp bak T7 promoteren. Det forventes at mengde uttrykt protein vil øke over tid, det tas derfor ut prøver hver 2. time for å finne optimalt ekspresjonsnivå.

Prosedyre:

1. Tin opp rør med pETmll3104- og pETmll3105-transformerte BL21 Star *E. coli* celler. Foreta fortynningsutstryk på agarplater med LB-medium tilsatt 50 μ g/ml ampicillin for transformantene.
2. Inkubér i varmeskap over natt ved 37 °C. Velg ut enkeltkolonier fra agarplatene for henholdsvis *mll3104*- og *mll3105*-transformanter. Enkeltkoloniene løses i Erlend-Meyer kolber inneholdende 10 ml LB-medium (ampicillin som over). I tilfelle dårlig vekst kan det være fordelaktig å inokulere flere kolber for hver av transformantene.
3. Inokuleres over natt i inkubator ved 200 rpm og 37 °C.
4. Fra overnattkulturene overføres 0.5 ml til nye Erlend-Meyer kolber med 10 ml (ampicillin som over), og kolbene inkuberes ved 200 rpm og 37 °C.
5. Mål OD₆₀₀ med jevne mellomrom, der første måling gjøres etter 1 time. Splitt kulturene ved OD₆₀₀ på 0.5 for både *mll3104*- og *mll3105*-transformeringskulturer.
6. Overfør 5 ml kultur for hver transformant over i 2 nye kolber der en kolbe tilsettes IPTG, og den andre dyrkes videre uten. Tilsett 1 μ l IPTG per ml LB. Fortsett inkubasjon som over.

7. Overfør 0.5 ml kultur fra hver av transformantene u/IPTG til eppendorfrør. Spinn ned rørene ved 13.000 rpm i 3 minutter. Fjern supernatanten. Rørene vil fungere som 0. time uttak av kulturene, der IPTG foreløpig ikke ville ha induisert ekspresjon fra *lac*-operator. Røret med pellet kan fryses ved -20 °C.
8. For hver 2. time, opp til 8. time, tas det ut nye 0.5 ml kultur som i forrige trinn.
9. Spinn ned rør og fjern supernatanten som ved forrige trinn. Frys ned rør ved -20 °C.

Storskala IPTG-indusert uttrykk av Mll3104- og Mll3105-proteiner

Etter eksperiment med pilotekspresjon, er tidspunktet for høyest ekspresjonsnivå etter IPTG-induksjon bestemt. For senere rensing av uttrykt protein fra kultur, vil mengde inokulum for henholdsvis *mll3104*- og *mll3105*-transformanter oppskaleres til 200 ml. Tidspunktet for optimalt uttrykk av proteiner ble funnet til å være etter 6 timers inokulasjon, derfor vil kolbene dyrkes uavbrutt i 6 timer før høsting.

Prosedyre:

1. Inokulér separate kolber med 10 ml LB-medium tilsatt 50 µg/ml ampicillin med BL21 Star *mll3104*- og *mll3105*-transformanter. Dyrk over natt ved 37 °C og 200 rpm.
2. Inokuler 200 ml LB tilsatt ampicillin som over med 1 ml av overnattkulturen.
3. Dyrk kulturen ved 37 °C og 200 rpm til $OD_{600} = \sim 0.5$, cellene bør da være i mid-log fase.
4. Tilsett 200 µl IPTG til kolbene, og dyrk kulturen 6 timer for optimalt uttrykk.
5. Fordel kulturen i 4 falconrør (50 ml) og sentrifuger 15 minutter ved 6000 rpm.
6. Fjern supernatanten og vask cellene med 40 ml PBS.
7. Spinn celler som i punkt 5.
8. Fjern supernatanten og tilsett 40 ml PBS.
9. Spinn rørene som i punkt 5.
10. Fjern supernatanten.
11. Frys rør med pellet ved -80 °C.

Proteinekstrahering ved bruk av B-PER® Bacterial Protein Extraction Reagent

B-PER Bacterial Protein Extraction Reagent muliggjør enkel ekstrahering av proteiner fra lyserte bakterier uten mekanisk påvirkning. Avhengig av type anvendelse kan tilleggskomponenter (lysozymer, proteasehemmere, salter, reduserende stoff etc.) tilsettes. Reagensen brukes både for løselig proteinekstrahering og rensing av inklusjonslegemer fra bakterielle cellelysater. I dette studiet ble reagensen tilført frosne pelletter av *E. coli* ekspresjonsstamme BL21 Star.

Prosedyre:

Maxi-Scale Bacterial Protein Extraction (200 ml bakteriekultur, $A_{600} = 1.5-3.0$)

1. Sentrifuger bakteriecellene ved $5.000 \times g$ i 10 minutter.
2. Tilsett 15 ml B-PER reagens og $50 \mu\text{l}$ proteasehemmer til pelleten. Pipetter opp og ned til løsningen er homogen. Rist løsningen forsiktig i 10 minutter.
3. Overfør løsningen til et sentrifugerør av passende størrelse. Sentrifuger ved $14.000 \times g$ i 15 minutter.
4. Overfør supernatanten til et nytt rør. Pellet kan eventuelt fryses ned ved -20°C , dersom ekstraheringen er utilstrekkelig og skal gjentas.

(PIERCE)

SDS-PAGE (Polyakrylamid Gel-elektroforese)

Denne typen gel-elektroforese er en av de mest brukte analytiske metodene for å separere og bestemme vekt på proteinfragmenter. Vanligvis brukes den anioniske detergenten Sodium Dodecyl Sulfat (SDS) for å denaturere proteiner i kombinasjon med et reduserende virkestoff, og varme for denaturering, før proteinprøvene settes på gelen. Denaturerte polypeptider bindes sterkt til SDS og blir negativt ladet. Fordi mengde SDS bundet nesten alltid er proporsjonal med den molekylære vekten av polypeptidet uavhengig av sekvens, vil SDS-polypeptidkomplekser migrere mot anoden gjennom polyakrylamidgeler etter størrelsen. Ved å bruke markører av kjent størrelse er det mulig å bestemme polypeptidkjedenes molekylære vekt ved å estimere relativ mobilitet av fragmenter gjennom gelen.

Polyakrylamidgeler er sammensatt av kjeder med akrylamid som kryssbindes med N,N'-metylen-bis-akrylamid. Effektiviteten av separeringen på gelen vil avhenge av konsentrasjonen av akrylamid brukt under støping av gelen, samt mengde kryssbindinger.

Kryssbindingene vil sørge for rigiditet og elastisitet i gelen, og sørger at det dannes porer der SDS-polyakrylamidkomplekser må vandre. Gelens egenskaper bestemmes av porenes størrelse, som vil være en funksjon av akrylamid konsentrasjon og mengde bisakrylamid brukt under støping av gelen. TEMED fungerer som en initiator for polymeriseringsprosessen, og mengden tilsatt vil innvirke på hvor raskt gelen polymeriseres (Sambrook and Russel 2001).

Prosedyre:

Tillaging av SDS-polyakrylamid Geler

1. Sett sammen glassplater etter leverandørens instruksjoner.
2. Bestem volum gelmengde (denne informasjonen er vanligvis gitt av leverandøren). Tilbered det passende volum av løsning med ønsket konsentrasjon av akrylamid for separerings gel. Komponentene mikses i rekkefølgen vist i tabell 16, og blandes i Erlendmeyer kolbe eller et engangs plastrør. Polymerisering vil starte så snart TEMED er tilsatt. Vend lett på miksturen, og fortsett til neste trinn.

NB: Konsentrasjonen av ammonium persulfat er høyere enn det som vanligvis brukes. Dette eliminerer nødvendigheten med å fjerne oppløst oksygen (som hemmer polymeriseringsprosessen) ved avgassing.

3. Hell akrylamidløsningen i åpningen mellom glassplatene. Etterlat nok rom for stacking gelen (lengden av tenner på kammen + 1 cm). Bruk en Pasteurpipette til å forsiktig legge lag med 0,1 % SDS. Plasser gelen i vertikal posisjon ved romtemperatur.

Topplaget med 0,1 % SDS vil hindre oksygendiffusjon inn i gelen, noe som reduserer polymeriseringen. Isobutanol løser opp plasten som noen av minigel-apparatene er laget av.

4. Etter polymerisering er fullstendig (ca 30 minutter), hell av laget med 0.1 % SDS fra gelen. Drenér bort så mye væske som mulig fra toppen av gelen, og fjern det resterende vannet ved hjelp av en papirkant som holdes inntil åpningen mellom glassplatene.
5. Tilbered running gel som følgende: I et engangs Falconrør tilberedes passende mengde løsning som inneholder ønsket konsentrasjon av akrylamid. Miks komponentene i den rekkefølgen som er vist i tabell 16. Polymerisering vil begynne så snart TEMED er tilsatt. Uten forsinkelser vendes miksturen lett.

6. Tøm stacking gel løsningen direkte på overflaten av den polymeriserte separerings gelen. Umiddelbart etter settes en plastkam (evt. teflon kam) i stacking gel løsningen, unngå at det setter seg luftbobler mellom tennene. Tilsett mer gel for å fylle opp rommet til kammen fullstendig. Plassér gelen i vertikal posisjon ved romtemperatur. Kammer bør rengjøres i H₂O, og skylles med etanol like før bruk.

Tabell 16: Viser reagenser og volum for SDS-PAGE geler benyttet under studiet.

Running gel (12%) x 2 volum	Mengde (ml)	Separerings gel (12%) x2 volum	Mengde (ml)
dH ₂ O	3.29	dH ₂ O	2.92
40 % Akrylamid	2.55	40 % Akrylamid	0.5
Tris 1.5 M	2.5	Tris 1.5 M	0.5
SDS 10 %	0.08	SDS 10 %	0.04
10 % Ammonium persulfat	0.08	10 % Ammonium persulfat	0.04
TEMED	0.004	TEMED	0.006
Totalvolum	8.504	Totalvolum	4.006

Tillaging av proteinprøver og kjøring av gel:

7. Mens stacking gel polymeriseres, tilberedes 2 X SDS gel-loading buffer som oppgitt i tabell 17.

Tabell 17: Viser reagenser og volum for 2x SDS gel-loading buffer.

Reagenser	Volum/vekt
0.5 M Tris-HCl, pH 6.8 (lysis og denaturering av prøver)	2.5 ml
Glyserol 100 % (tyngde til prøver)	2 ml
Bromofenolblå (visualisering av prøver)	0,02 g
SDS (lysis og denaturering av prøver)	0.4 g
dH ₂ O tilsettes til totalvolum	10 ml

Fortynn adekvat mengde 2x SDS samplebuffer til 1x SDS.

Hvert prøverør tilsettes:

40 µl 1 X SDS samplebuffer

16 µl DTT

24 µl dH₂O

Totalvolum 80 µl.

Mengden ganges opp med antall proteinprøver. Tilsett 80 µl samplebuffer til hver proteinprøve.

(Invitrogen-life-technologies 2004)

Prøvene med loadingbuffer og fraksjoner kokes i 5 minutter for proteindenaturering.

8. Etter polymeriseringen er fullstendig (ca 30 minutter) fjernes plastkammen forsiktig. Plassér gelen i elektroforeseapparatet. Tilsett Tris-Glycin elektroforese buffer til toppen av glassplatene. Viktig: Ikke sett elektrisitet til gelen før prøvene er tilsatt, dette vil ødelegge buffersystemene.
9. Tilsett 10 µl av hver prøve til bunnen av brønnene. Dette gjøres med vanlig gul pipettespiss (kan alternativt utføres ved bruk av Hamilton microliter syringe).
10. Fest elektroforeseapparatet til en elektrisk tilkobling kjør gelen på 90 V til det bromofenolblå når bunnen av gelen.
11. Fjern glassplatene fra elektroforese apparatet og bruk en ekstra gelspacer for å skille dem fra hverandre. Gelens orientering kan merkes ved å kutte et hjørne av gelen.
12. Gelen farges og fikseres med Coomassie Brilliant Blue fargeløsning.

(Sambrook and Russel 2001)

Rensing av polyhistidin-merkede Mll3104- og Mll3105 proteiner

Ved transkripsjon av ligerte sekvenser fra vektor pET101/D-TOPO i BL21 Star *E. coli* celler, vil priming fortsette forbi punkt for innsett. Dette fordi primere for *mll3104/mll3105* ble designet til å fjerne det opprinnelige stopkodonet i de respektive genene. T7 primer vil gå forbi den innsatte målsekvensen, og prime nedstrøms for innsettet. 72 bp etter innsettet i pET101/D-TOPO finnes en 6xHis region som blir transkribert sammen med målgenene, og stopkodon er plassert rett nedstrøms for polyhistidin regionen. Resultatet er translasjon av ligerte sekvenser inneholdende en polyhistidin markør.

ProBond™ er nikkel-ladete resiner som brukes for affinitetsrensing av uttrykte proteiner med en polyhistidin (6xHis) sekvens. Bindingskapasiteten til nye resiner er 1-5 mg protein per ml resin. Proteiner med affinitet til resin, forblir bundet til resinet i affinitetskolonnen, mens proteiner uten affinitet for resinets nikkel-ioner vaskes ut med fosfatbuffer. Lysat av BL21 Star celler med IPTG-indusert uttrykk av målsekvenser, påføres kolonnen. Cellelysering skjer ved tidspunkt for maksimalt uttrykk av proteiner. For opprensing av bundne proteinfraksjoner i etterkant, vaskes kolonnen med imidazole i fosfatbuffer. Imidazole ($C_3H_4N_2$) er et aromatisk alkaloid med en histidin-sidegruppe. Ved påføring av buffer med imidazole i kolonnen vil His-merket protein bundet til resin frigjøres, da imidazole har større affinitet til resinet. I dette studiet er 50, 100, 150, 200 og 300 mM konsentrasjoner av imidazole prøvd ut for optimal opprensing. Etter påføring av lysat med påfølgende vask med imidazole ble kolonnen klargjort for ny fraksjonsoppsamling ved vask med 50 ml fosfatbuffer og 50 ml 500 mM imidazole. Proteinkonsentrasjon bestemmes ved protein assay og SDS-PAGE separerer proteinfragmenter og angir størrelse på fraksjoner (Hengen 1995).

Prosedyre:

Eksperimentet utføres ved 4 °C, som er oppbevaringstemperatur for resinet. Resuspender flasken ProBond™ resin før bruk ved å vende flasken noen ganger. Tilbered en kolonne (ca 15 ml volum) med filter i bunnen (Whatman filter er brukt ved dette studiet). Tilsett resin sakte, og la resinet bunnfelle (5-10 minutter) før ytterligere tilsetning. Fyll kolonnen helt opp med resin, for å fortrenge O_2 . Resiner mister funksjonen ved tørke.

Følg instrumentleverandørens anvisning for programmering til oppsamling av 60 dråper i totalt 60 rør (fraksjoner). Etter hver vask med imidazol, gjennomføres et protein assay for bestemmelse av proteinkonsentrasjon i fraksjoner. Dersom konsentrasjonene er tilstrekkelig kjøres SDS-PAGE for relevante fraksjoner.

1. Rålysat tilføres kolonnen direkte fra oppbevaringsrør. Etter gjennomstrømming i kolonne kan lysatet samles opp for senere SDS-PAGE.
2. Vask med 50 ml fosfatbuffer (0.5 M K_2HPO_4 , 50 mM NaCl)
3. Vask med 50 ml 50 mM imidazol, samle opp fraksjoner maskinelt i glassrør.
4. Vask med 50 ml fosfatbuffer.
5. Evt: Vask med 50 ml 150 mM imidazole, samle opp fraksjoner maskinelt i glassrør.
6. Vask med 50 ml fosfatbuffer.

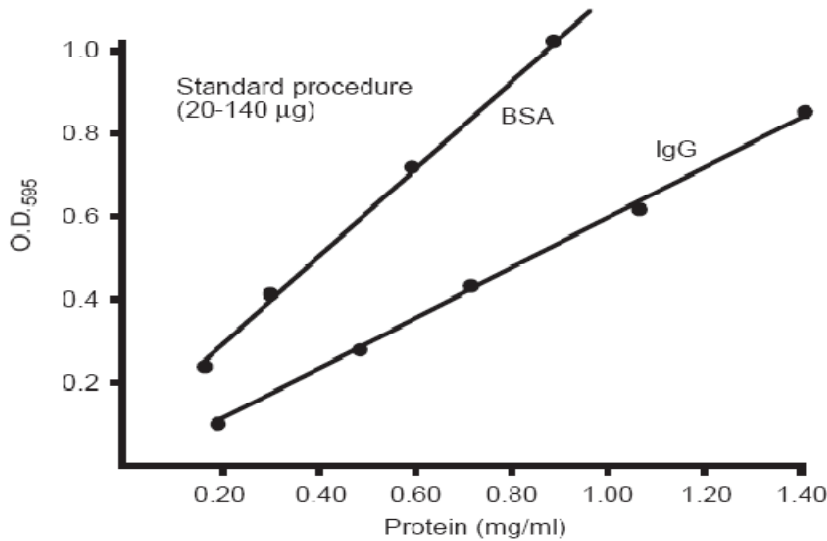
7. Evt: Vask kolonnen med 50 ml 200 mM imidazole, samle opp fraksjoner maskinelt i glassrør.
8. Vask kolonne med 50 ml fosfatbuffer.
9. Evt: Vask kolonnen med 50 ml 300 mM imidazole.

Bio-Rad Protein Assay

Bio-Rad Protein Assay er et fargebindings assay, der differensielle fargeendringer i proteinløsninger skjer som respons til variasjoner i proteinkonsentrasjoner. Sammenlikning med standardkurve (bovin serum albumin, BSA) sikrer en relativ måling av proteinkonsentrasjonen (figur 7). All målinger skjer spektrofotometrisk ved 595 nm (SpectraMax 250).

Prosedyre:

1. Preparer fargereagens ved å fortynne 1 del Dye Reagent konsentrat med 4 deler dH₂O. Den fortynnete fargeløsningen har en holdbarhet på 2 uker oppbevart i romtemperatur.
2. Den lineære rekkevidden av BSA er 0.2 til 0.9 mg/ml.
Preparer 5 fortynninger av standard proteinløsning representativt for proteinløsningen som skal måles.
0.0 mg/ml dH₂O
0.1 mg/ml BSA
0.2 mg/ml BSA
0.4 mg/ml BSA
0.8 mg/ml BSA
Standardkonsentrasjoner lages i eppendorfrør og bevares ved -20 °C.
Mål absorbans ved OD₅₉₅.



Figur 7. Vanlig standardkurve for bovin gamma globulin (IgG) og bovin serum albumin (BSA). Hentet fra Bio-Rad Protein Assay manual.

Prosedyre for mikrotiterplate assay:

3. Pipetter 10 µl av hver standard proteinløsning (BSA) i separate mikrotiterplate-brønner.
4. Tilsett 200 µl av fortynnet fargereagens til hver brønn. Miks prøvene forsiktig ved å pipettere opp og ned, unngå skumdannelse. Bytt pipettespiss for hver prøve.
5. Tilsett 10 µl av hver ukjente proteinfraksjon i separate mikrotiterplate-brønner, som ved forrige punkt.
6. Mål absorbans ved 595 nm.

(Bio-Rad Protein Assay Kit instruksjonshefte)

Q-TOF MS/MS

Moderne massespektrometri (MS) har utviklet seg til et av de viktigste verktøyene for studier av proteiner. Med MS kan molekylvekt (MW) av intakt protein måles med stor nøyaktighet, etter enzymatisk hydrolyse kan MW av de lett frigjorte peptidene måles, og hel eller delvis aminosyresekvens av peptider kan lett oppklares ved hjelp av MS. Flere ulike massespektrometriske teknikker kan benyttes. Den teknikken som gir mest informasjon om peptider, er Nano UPLC (Nano Ultra Performance Liquid Chromatography) og MS med Q-TOF konfigurasjon. NanoUPLC er en miniaturisert HPLC-teknikk, der analyser utføres ved å bruke kolonner med en indre diameter mellom 10-150 µm. Navnet refererer til flow-raten til

prøvematerialet, og oppgis som nanoliter per minutt. Fordelen ved bruk av kolonner med så liten indre diameter at mindre prøvemateriale trengs for analyse (Jensen 2007).

For mulig proteinidentifisering vil analysert peptidmasse søkes gjennom offentlige databaser for kjente hypotetiske proteiner. I dette studiet ble proteinsekvenser søkt opp i Mascot, et proteinidentifiserings-system fra Matrix Science. Søkefunksjonen tillater identifisering av hypotetiske proteiner ut fra peptidmasse og aminosyresekvenser.

Institutt for farmasi har laboratorier med ovennevnte fasiliteter, og utfører rutinemessig Q-TOF MS/MS. Etter SDS-PAGE av ukjente proteinfraksjoner fra ekspresjonseksperimenter, ble relevante bånd skåret ut fra gelen og sendt til analyse ved denne laben. Da vi kun fikk uttrykt *mll3105*-proteiner fra BL21 Star cellene, er det resultater for denne analysen som presenteres i appendiks VII.

Resultater

I dette kapitlet vil alle resultater som er oppnådd under arbeidet bli presentert. Dette arbeidet involverer genotypisk identifikasjon av målgener og primerdesign spesifikk for disse, amplifisering av genene, DNA-rensing, kloning og transformasjon til to ulike plasmider, og i to ulike *E. coli* stammer. I tillegg vil resultater fra DNA-sekvensering etter innsett av målgener i ekspresjonsvektor pET101 presenteres. Resultater fra ekspresjonseksperimenter og SDS-PAGE geler med uttrykte målproteiner gjennomgås, og resultater fra Nano UPLC med Q-TOF MS/MS vil også bli presentert.

Primerdesign

For å bekrefte at referansestammen, *M. loti* MAFF303099 var positiv for genene *mll3104* og *mll3105*, måtte spesifikke primere designes. All informasjon om *M. loti*'s *mll3104* og *mll3105* gener er hentet fra NCBI gjennom Genbank søkesystem. Systemet er tilgjengelig ved: <http://www.ncbi.nlm.nih.gov/Genbank/index.html>.

For amplifisering av målgensekvenser er det nødvendig med forward og revers primer. Anbefalt lengde på primere bør være mellom 18 og 25 baser, og ved dette studiet ble det designet 25 basers primere. Sekvenser er oppført i tabell 6, og kriterier for designet er beskrevet under metodedel for primerdesign. For å undersøke om egne primere fungerte og utviste spesifisitet for *mll3104* og *mll3105* ble det satt opp PCR-reaksjoner med påfølgende agarosegel-elektroforese. Figur 8 viser et eksempel på en gel-elektroforese der egne primere ble anvendt. Resultatet er tydelige enkeltbånd av omtrentlig størrelse for både *mll3104* og *mll3105*. Båndene bekrefter av baseparsammensetning i primersekvensene var funksjonelle og spesifikke for de respektive gener. Det ble derfor ikke gjort endringer for optimalisering, og primerne er brukt videre i alle eksperimenter der formålet er å amplifisere *mll3104*- og *mll3105*-fragmenter.

Sekvenslikhet mellom *thuB* og *mll3104* og mellom *thuA* og *mll3105*

thuB og *mll3104*

Appendiks II viser en sekvensoppstilling mellom genene utført i softwareprogram BLASTN (NCBI). Homologi mellom *thuB* (*Sinorhizobium meliloti*) og *mll3104* (*Mesorhizobium loti*) strekker seg over en enkelt region. Den totale prosentvise likheten mellom de to potensielle ortologene ble funnet å være 76 %. Det er kun 27 baser forskjell i total lengde mellom de to genene. Homologien mellom sekvensene og den totale størrelsen på gensekvensene indikerer

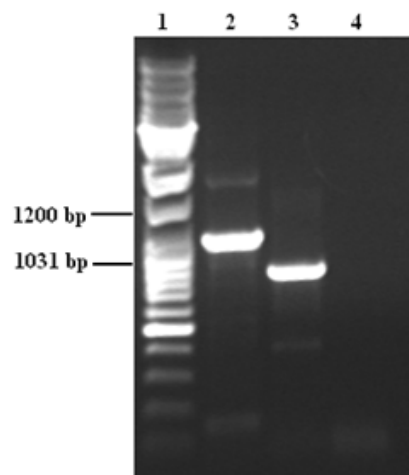
at det kan dreie seg om et hypotetisk protein med like funksjoner mellom de to artene, eller at sekvensene er deler av konserverte regioner.

thuA og *mll3105*

Appendiks I viser sekvensoppstilling mellom *thuA* (*S. meliloti*) og *mll3105* (*M. loti*). Søket er utført med samme programvare som over, og viser at det finnes 3 regioner med homologi for genene. Totalt er det 72, 79 og 81 % likhet mellom de 3 regionene, og det er 48 basepar forskjell i total basepar lengde mellom genene. Artene er nært beslektet, det kan derfor tenkes at *M. loti*'s *mll3105* har samme funksjon som *thuA* i *S. meliloti*, eller at sekvensene er deler av konserverte regioner.

Genotypisk identifisering av trehalose assosierte gener; *mll3104* og *mll3105* fra *M. loti* MAFF303099 templat

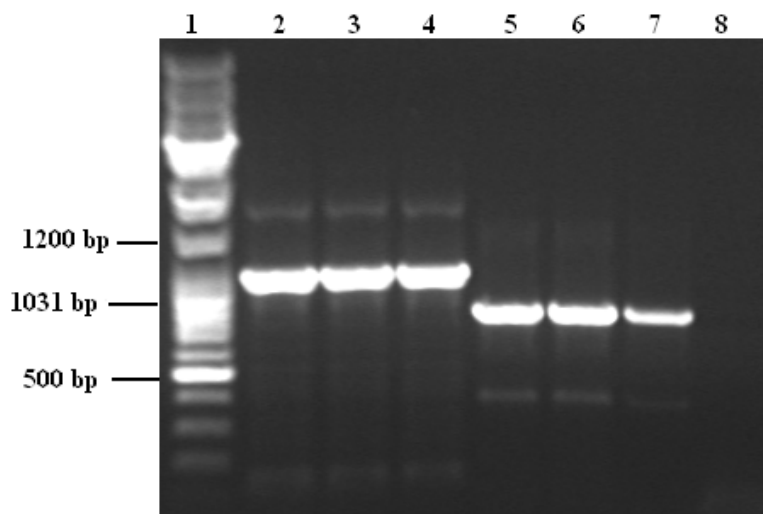
Eksperimentet viste at egne designede primere fungerte, i tillegg til amplifisering av målgener fra *M. loti* genomisk DNA. Figur 8 viser amplifiserte produkter av henholdsvis *mll3104* på ~1050 basepars lengde i brønn 2, og *mll3105* på ~786 bp lengde i brønn 3. Størrelsen på produktene er som forventet i forhold til størrelser på *mll3104/mll3105* i GenBank. Parallelt med amplifisering av målsekvensene ble det utført en negativ kontroll der dH₂O erstattet templat, og med primerbruk som for *mll3105* amplifisering. Den negative kontrollen er blank, og uten kontaminasjon.



Figur 8: 0.8 % agarosegel-elektroforese etter amplifisering av *mll3104* og *mll3105* fra *M. loti* genomisk DNA som templat. **Brønn 1:** GeneRuler DNA Ladder Mix (0.5 µg/µl). **Brønn 2:** *mll3104*-produkt, av ~1050 bp størrelse. **Brønn 3:** *mll3105*-produkt av ~780 bp størrelse. **Brønn 4:** Negativ kontroll (dH₂O og primer for *mll3105*)

Kvantitativ amplifisering av målgener *mll3104* og *mll3105*

Det ble utført repetitive eksperimenter med vellykket amplifisering av *mll3104* og *mll3105* ved bruk av spesifikke primere, der formålet var å kutte ut DNA fra agarosegelen. Figur 9 viser en gel der amplifiseringen ga forventet størrelse på båndene. Det ble kjørt 3 prøver for hvert gen, med hensikt å arbeide videre med høyest mulig DNA-konsentrasjon etter opprensing fra gelen. PCR-produktene ble skåret ut under UV-belysning, og rensset ved bruk av kommersielt gel-ekstraherings kit (QIAquick Gel extraction Kit, QIAGEN). Brønn 8 viser negativ kontroll med primerbruk som for forrige figur. Denne er uten kontaminasjon, noe som bekrefter at amplifiseringsproduktene er rene og kan benyttes i videre arbeid.



Figur 9: Agarosegel-elektroforese etter amplifisering av *mll3104* og *mll3105*. **Brønn 1:** GeneRuler DNA Ladder Mix. **Brønn 2-4:** amplifisert *mll3104*. **Brønn 5-7:** amplifisert *mll3105*. **Brønn 8:** Negativ kontroll med dH₂O og primer for *mll3105*. Bånd ble skåret ut og brukt til DNA-ekstrahering.

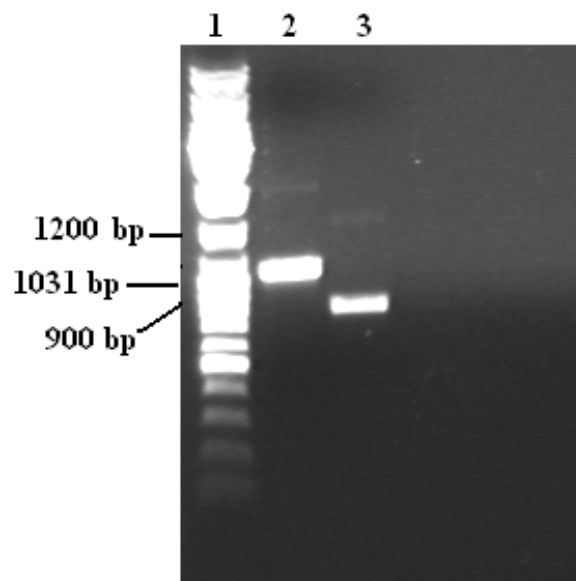
Ekstrahering av amplifisert genomisk DNA fra *M. loti* etter agarosegel-elektroforese

DNA-ekstrahering fra gel ble utført ved bruk av QIAquick Gel extraction kit, og det ble rutinemessig foretatt gel-elektroforese med formål å undersøke om ekstraheringen var vellykket. Figur 10 viser amplifiserte produkter fra *mll3104* og *mll3105*- *M. loti* genomisk templat rensset fra gel. Produktstørrelsen er som ved tidligere forsøk, noe som indikerer at DNA-ekstraheringen var vellykket. Det ble derfor besluttet å benytte DNAet for senere ligeringsreaksjoner. Ved ligeringsreaksjoner er det vanlig å benytte DNA-konsentrasjoner som samsvarer med konsentrasjon av vektor for best mulig kloningsresultat.

Konsentrasjonen på ekstrahert DNA som ble bevart for ligering ble målt på NanoDrop 1000 software:

-*mll3104*: 56,8 ng/μl

-*mll3105*: 38,2 ng/μl



Figur 10: Agarosegel-elektroforese etter DNA-ekstrahering fra *mll3104* og *mll3105* PCR-produkter. **Brønn 1:** GeneRuler DNA Ladder Mix 0.5 μg/μl. **Brønn 2:** *mll3104* DNA ekstrakt. **Brønn 3:** *mll3105* DNA ekstrakt. Båndene har omtrentlig størrelse som forventes etter amplifikasjon av målgenene.

Ligeringsreaksjon med ekstraherte *mll3104* og *mll3105* PCR-produkter i vektor pCR2.1 TOPO og transformasjon i *E. coli* TOP10 celler

Ligering av amplifiserte templat i vektor pCR 2.1-TOPO ble gjennomført etter leverandørens protokoll (Invitrogen) som oppgis under metodedel. Positivt innsett ble undersøkt ved å dyrke 30 μl av transformasjonsmikts over natt på agarplater med LB-medium tilsatt selektivt antibiotika (50 μg/ml kanamycin) og X-gal (40 mg/ml). Positiv kontroll (pUC19) har ampicillinresistens som seleksjonsmarkør, og ble dyrket på agarplater med 50 μg/ml ampicillin. I negativ kontroll er templat er erstattet med dH₂O (tabell 9). Et korrekt innsett skal gi lysblå eller hvite kolonier, dette på grunn av en ikke-funksjonell β-laktamase. Etter

gjentatte dårlige resultater som følge av feil templat:plasmid ratio, ble det utprøvd fortynnede løsninger av kloningsreaksjonene, noe som ga bedre resultater. I forsøkene ble det benyttet 4 plater til dyrkning etter hver transformasjon, bortsett fra positiv kontroll (2 plater). Tabell 18 viser et eksempel på resultat etter elektroporetisk transformasjon av pCR2.1 D-TOPO i *E. coli* TOP10 celler inkubert over natt på agarplater (LB-medium) med antibiotika ved 37 °C. Overskuddsplater inneholder all overskuddsmengde S.O.C-medium med transformerte celler, 160 µl. Erfaringsmessig har det vist seg at 30 µl kan gi dårlig vekst. Ved å dyrke overskuddsplater får man utnyttet alle cellene fra transformasjonsmiksen, noe som kan gi bedre vekst av transformanter.

Tabell 18: Viser et eksempel på resultat etter transformasjon av pCR2.1 i *E. coli* TOP10 celler dyrket på agarplater tilsatt antibiotika og X-gal. Plate 1-3 ble tilsatt 30 µl transformasjonsmiksen, bortsett fra overskuddsplaten (160 µl).

	Plate 1	Plate 2	Plate 3	Overskuddsplate
<i>mll3104</i> (56,8 ng/µl)	-	-	1 blå koloni	15 blå kolonier
<i>mll3104</i> fortynnet 1:4 (14,2 ng/µl)	>700 blå kolonier	Teppevekst	4 hvite og 2 blå kolonier	21 hvite og 8 blå kolonier
<i>mll3105</i> (38,2 ng/µl)	2 blå kolonier	3 hvite og 5 blå kolonier	6 hvite og 4 blå kolonier	7 hvite og 4 blå kolonier
<i>mll3105</i> fortynnet 1:3 (12,6 ng/µl)	74 hvite kolonier	2 blå	5 hvite og 14 blå kolonier	Hvit teppevekst og 4 blå kolonier
Negativ kontroll	1 blå	-	3 blå	2 blå
Positiv kontroll	Teppevekst	Teppevekst		

Positive transformanter ble dyrket videre i 3 ml LB tilsatt antibiotika, og plasmid ble rensset ut fra cellekulturere.

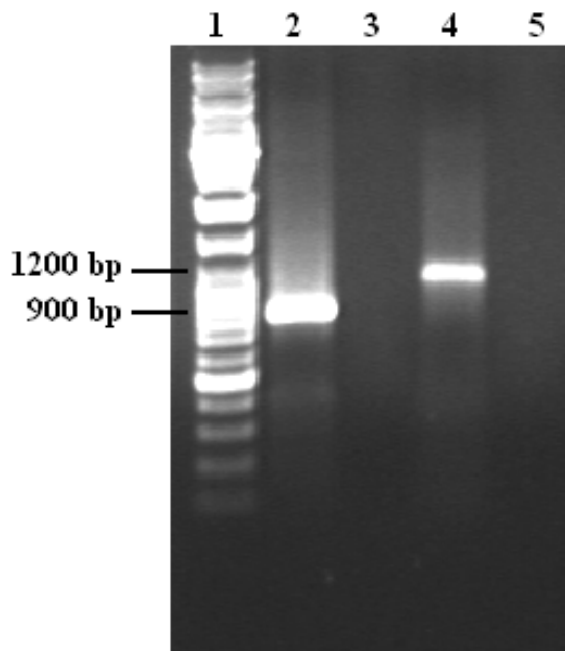
Konsentrasjonene av plasmid-DNA ble målt med NanoDrop 1000:

-*mll3104* plasmid-DNA: 31,2 ng/µl

-*mll3105* plasmid-DNA: 23,1 ng/µl

Amplifisering av *mll3104* og *mll3105* med Pfx50 polymerase fra vektor pCR2.1-TOPO

Amplifisering av *mll3104* og *mll3105* med Pfx50 polymerase fra genomisk DNA isolert fra *M. loti* ga dårlige resultater, det ble derfor besluttet å amplifisere gensekvensene med denne polymerasen etter innsett av fragmentene i vektor pCR2.1-TOPO. Det ble foretatt PCR som oppgitt under metodedel, og agarosegel-elektroforese ble utført for å undersøke om amplifiseringen var vellykket. Figur 11 viser en agarosegel der bånd i brønn 2 ser ut til å være produkt fra amplifisering av *mll3105*, og brønn 4 viser produkt fra *mll3104*, begge amplifisert ved bruk av Pfx50 polymerase og tilhørende buffer. Brønn 3 og 5 utgjør negative kontroller for de respektive gener, der dH₂O ble benyttet i stedet for templat. Da produktene hadde forventet størrelse sammenliknet med originale størrelser på målgenene, ble båndene kuttet ut fra gel ved bruk av QIAGEN Gel Extraction kit (QIAGEN), for videre ligering i ekspresjonsvektor pET101/D-TOPO.



Figur 11: Viser produkter etter amplifisering av målgener fra vektor pCR2.1-TOPO ved bruk av Pfx50 polymerase og tilhørende buffer. **Brønn 1:** GeneRuler DNA Ladder Mix. **Brønn 2:** *mll3105* amplifiseringsprodukt. **Brønn 4:** *mll3104*-amplifiseringsprodukt. **Brønn 2 og 5:** Negative kontroller for målgener.

Ligering av *mll3104*- og *mll3105* fragmenter i vektor pET101/D-TOPO

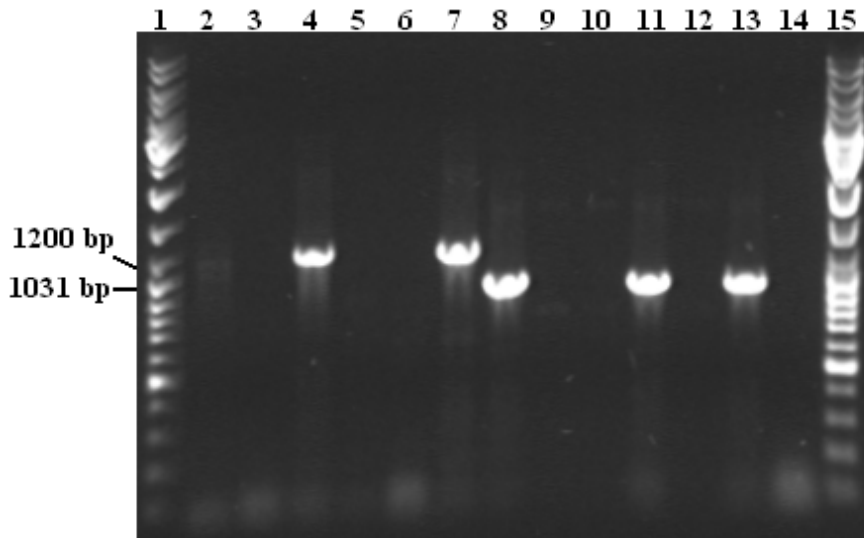
Ligeringsreaksjoner med ekspresjonsvektor pET101/D-TOPO og *mll3104/mll3105*- PCR-produkter ble utført. For enkelhets skyld vil respektive innsett av eget *mll3104*-fragment i

pET101/D-TOPO heretter refereres til som vektor pETmll3104, og innsett av *mll3105* i samme vektor refereres til som pETmll3105. For å undersøke om innsettene sto i riktig leseretning, ble det satt opp PCR med ulike primerkombinasjoner.

T7 forward og revers primer inkluderes i ekspresjonskitet fra Invitrogen, og i tillegg brukes egne primere (tabell 6). Dersom ligering har gitt innsett i riktig leseretning, skal amplifisering gi produkter ved bruk av T7 revers primer og *mll3104/mll3105* forward primere. Figur 12 viser et PCR-oppsett som nevnt over, med ulike primerkombinasjoner. Resultatet viser at enkelte innsett plassert i riktig leseretning, der bruk av T7 revers primer og *mll3104/mll3105* forward primer kan ha gitt bånd av omtrentlig forventet størrelse. Av tre ulike kulturer for *mll3104*, er det kun kultur C som viser et tilsynelatende korrekt innsett (brønn 7). For *mll3105*-ligering ser det ut som riktig innsett i alle 3 kulturer (brønn 8, 11 og 13). Brønn 4 viste seg å være en feil da forsøket ble repetert, og ekskluderes.

For å avgjøre leseretningen til et innsett er det nødvendig å kombinere en vektor-primer (forward eller revers), og en av egne primere (forward eller revers). Dersom en bare benytter seg av T7-primere, vil det resultere i et produkt der det ikke kan avgjøres om innsettet er korrekt. Ved utelukkende bruk av egne primere, vil man også få et produkt som ikke gir informasjon om hvilken retning innsettet sitter i vektoren.

Ved kloning i ekspresjonsvektor er det essensielt at innsettene liggeres i riktig leseretning dersom IPTG-indusert transkripsjon av målsekvensene skal translateres til funksjonelle proteiner. Tilsetning av IPTG i vekstmediet, fører til at T7 promoter binder til T7-promotersete og aktiverer transkripsjon. *mll3104*- og *mll3105*-målsekvenser er ligert inn nedstrøms for promotersetet, og transkriberes ved IPTG-indusert aktivering av *lac*-operonet.



Figur 12: Agarosegel-elektroforese etter amplifisering av *mll3104* og *mll3105* fra innsett i pET101/D-TOPO[®] rensset fra 3 forskjellige flytekulturer (A, B og C) av TOP10 *E. coli*-celler. I alle brønner med prøvemateriale og brønn for negativ kontroll det brukt primer T7-R i tillegg til en egendesignet primer for *mll3104/mll3105*-fragmenter. **Brønn 1 og 15:** GeneRuler DNA Ladder Mix. **Brønn 2:** Kultur *mll3104*-A; forward primer for *mll3104*. **Brønn 3:** Kultur *mll3104*-A; Primer *mll3104*-Rev. **Brønn 4:** Kultur *mll3104*-B; Primer *mll3104*-R. **Brønn 5:** Kultur *mll3104*-B; Primer for *mll3104*-Forw. **Brønn 6:** Kultur *mll3104*-C; Primer *mll3104*-Rev. **Brønn 7:** Kultur *mll3104*-C; Primer *mll3104*-Forw. **Brønn 8:** Kultur *mll3105*-A; Primer *mll3105*-Forw. **Brønn 9:** Kultur *mll3105*-A; Primer *mll3105*-Rev. **Brønn 10:** Kultur *mll3105*-B; Primer *mll3105*-Rev. **Brønn 11:** Kultur *mll3105*-B; Primer *mll3105*-Forw. **Brønn 12:** Kultur *mll3105*-C; Primer *mll3105*-Rev. **Brønn 13:** Kultur *mll3105*-C; Primer *mll3105*-Forw. **Brønn 14:** dH2O og primer *mll3105*-Forw.

Plasmid-DNA ble rensset ut som oppgitt i metodedel (QIAGEN), og konsentrasjonen ble målt på NanoDrop 1000:

-*mll3104*-C plasmid-DNA: 37 ng/ μ l

-*mll3105*-A plasmid-DNA: 44 ng/ μ l

-*mll3105*-B plasmid-DNA: 26 ng/ μ l

-*mll3105*-C plasmid-DNA: 21 ng/ μ l

Transformasjon av pET1017D-TOPO i BL21 Star *E. coli* celler

Oppsettet er som for tidligere transformasjoner, og 30 μ l transformeringsreaksjoner med BL21 Star *E. coli* elektrokompetente celler ble sådd ut på LB-plater tilsatt 50 μ g/ml ampisillin. Det ble sådd ut transformasjonsmikts på 4 plater (3x 30 μ l og en overskuddsplate)

for henholdsvis pETmll3104 og pETmll3105. Veksten var lik på alle platene, resultatet var over 900 kolonier per plate. Positiv kontroll hadde 5 kolonier, mens negativ kontroll var blank.

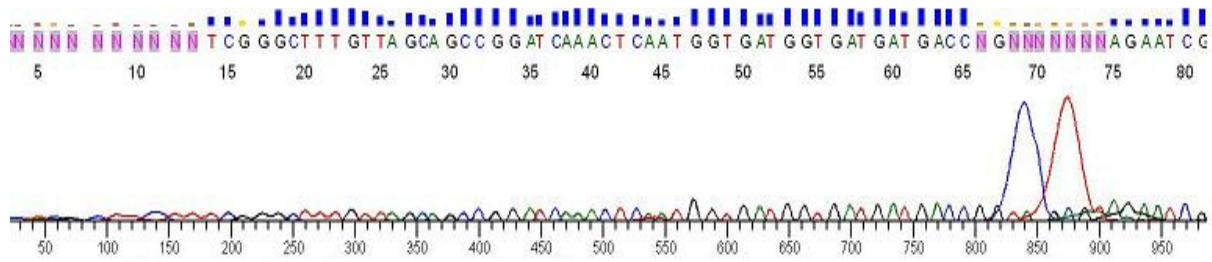
Videre ble det valgt ut 4 enkeltkolonier for pETmll3104- og pETmll3105-transformanter. Disse ble dyrket i kolber med 10 ml LB (m/ ampicillin), og det ble laget stock-løsning av kulturene (200 µl i eppendorfrør). Før ekspresjon av innsatte fragmenter fra vektor i BL21 Star cellene, var det nødvendig med DNA-sekvensering for å avgjøre om baserekkefølgen i fragmentene var korrekte. Oppsettet er angitt under metodedel for PCR med Big Dye enzymer.

DNA sekvensering av målsekvenser (*mll3104* og *mll3105*) etter kloning i pET101/D-TOPO

Sekvenseringsreaksjonen ble utført ved Universitetssykehuset i Nord-Norge. For å undersøke sekvensene, ble softwareprogrammene Sequence Scanner v1.0 og BLAST ("Blast two sequences"-funksjon) benyttet. Prøvene ble initialt vurdert i Sequence Scanner for å bestemme produktlengde og mengde ukjente baser. Dersom mengde ukjente baser var stor eller totallengden på produktet for kort, ble prøvene forkastet.

For å bestemme antall ukorrekte base-innsett, ble sekvensen sammenliknet med den opprinnelige sekvensen presentert for det enkelte gen i Genebank eller NCBI. BLAST har en funksjon der to sekvenser kan stilles opp mot hverandre for å bestemme prosentvis homologi, og for å finne mulige enkeltbaser som er feil ved sammenlikning med originalsekvensen.

Totalt ble det utført 5 sekvenseringsreaksjoner for begge genene der man brukte både T7 forward og revers primer enkeltvis i reaksjoner. Et gjennomgående problem ved resultatene var rester av overskudds templat-DNA/primer-DNA (figur 13). Mengden av overskudds-DNA ved felling gjorde at baser plassert 65-75 bp etter primer-start kom i bakgrunnen, og dermed ikke kunne identifiseres med sikkerhet.



Figur 13: Viser eksempel på utdrag av en DNA-sekvensering, der rester av templat-DNA og primer-DNA kommer til uttrykk rundt 65-75 basepar etter sekvenseringsstart. DNA-toppene var et gjennomgående problem for innsendt materiale, og gjorde det umulig å fastslå basesammensetningen rundt ligeringspunktet for egne innsett.

Det var av den grunn vanskelig å avgjøre om starten for innsett av egne fragmenter var korrekt. Tidligere studier av målgene har vært gjennomført, der man etter DNA-sekvensering fant korrekte innsett. For kontinuitet av studiets individuelle eksperimenter, ble det antatt at innsettene var korrekte.

Det ble senere byttet kit for DNA-presipitering ved sekvenseringslaben, og endringer for å oppnå bedre resultater ble foreslått. For å fastslå om resultatet ville endre seg ved optimalisering av reaksjonen, ble det sendt inn *mll3105*-sekvenser. Prøvematerialet var amplifisert fra pETmll3105 ved bruk av T7 forward og revers primer enkeltvis. Resultatet var vesentlig bedre enn ved tidligere sekvenseringsreaksjoner, og toppene med overskudds-DNA var fraværende i en enkelt prøve. Appendiks VI viser sekvensen etter at nytt fellings-kit for DNA i sekvenseringsreaksjoner ble innført.

Sekvensering av *mll3104*-innsett i pETmll3104

Det var kun to av sekvenseringsreaksjonene for innsettet som ble beholdt, andre ble forkastet på grunn av høy mengde ukjente baser. Appendiks III viser et sekvenseringsresultat der T7 forward primer ble benyttet for amplifisering av *mll3104*-sekvensen fra vektor pETmll3104. Det er ikke funnet ukorrekte baseinnsett i sekvenseringsresultatene, men ukjente baser gjør at det ikke er mulig å fastslå om innsettene i sin helhet er korrekte. Det andre resultatet er tilsvarende, og angis ikke i appendiks.

Sekvensering av *mll3105*-innsett i pETmll3105

For *mll3105* var resultatene bedre. Det viser seg at 4 enkeltbaser er feil innsatt ved sammenlikning av innsatt *mll3105*-sekvens og det opprinnelige *mll3105*-genet fra Genbank. Resultatet går igjen i alle sekvenseringsreaksjoner for *mll3105*-innsettene, både med bruk av forward og revers T7 primer (appendiks, IV, V og VI).

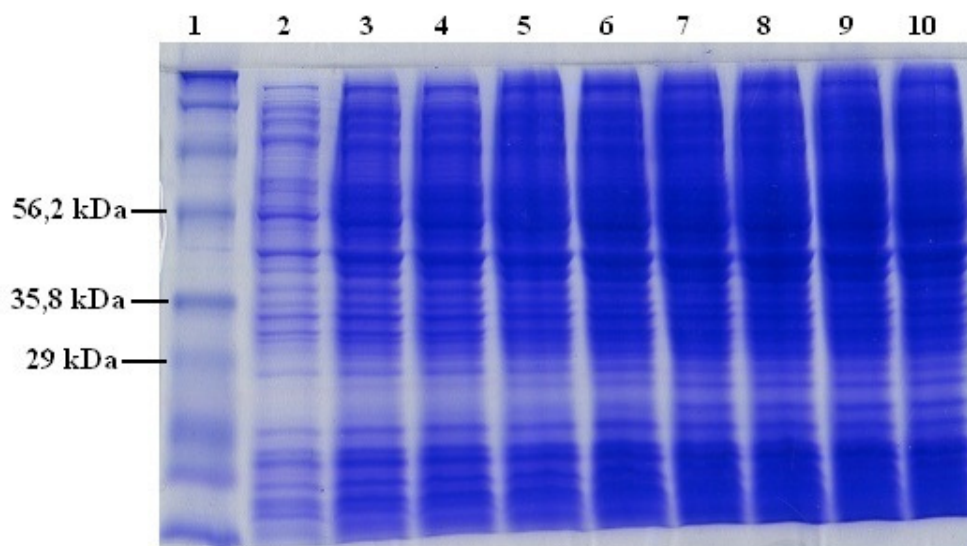
SDS-PAGE av proteinfraksjoner vasket med ulike imidazole konsentrasjoner

Mll3104-pilotekspresjons forsøk

Et eksperiment med pilotekspresjon av Mll3104-proteiner ble først gjennomført ved at IPTG ble tilsatt en kultur med *E. coli* ekspresjonsstamme BL21 Star med innsatt plasmid i LB-medium. Ved gitte tidspunkt (hver 2. time) ble det tatt ut 0.5 ml prøver. En kontrollkultur uten tilsatt IPTG ble gitt de samme vekstbetingelsene. Prøver fra begge kulturer ble fryst ned, og resultatet foreligger som bånd i en SDS-PAGE gel vist i figur 14. Lysatet inneholder samtlige proteiner representert i cellene ved lyseringstidspunkt, og man forventet derfor mange bånd. Mll3104-proteiner har en molekylvekt på ~42.5 kDa, og ved IPTG-indusert uttrykk skal proteinene synes i brønn 3, 5 og 7. Ingen av brønnene har synlig induksjon, noe som imidlertid ikke betyr at IPTG ikke aktiverte transkripsjon fra *mll3105*-innsett. Det kan tenkes at relevante proteiner ligger skjult i bånd med proteiner av samme størrelse. Mll3104-produktstørrelse estimeres ut fra følgende: 1 kb DNA = 37 kDa. Data fra GeneBank angir lengde på *mll3104*-gen til 1050 bp. Etter innsett i pETmll3104 vil tilleggsbaser fra kloningssetet til termineringskodon (99 bp) inkluderes i det endelige produktet:

mll3104-produkt: 1050 bp + 99 bp som gir 1149 bp.

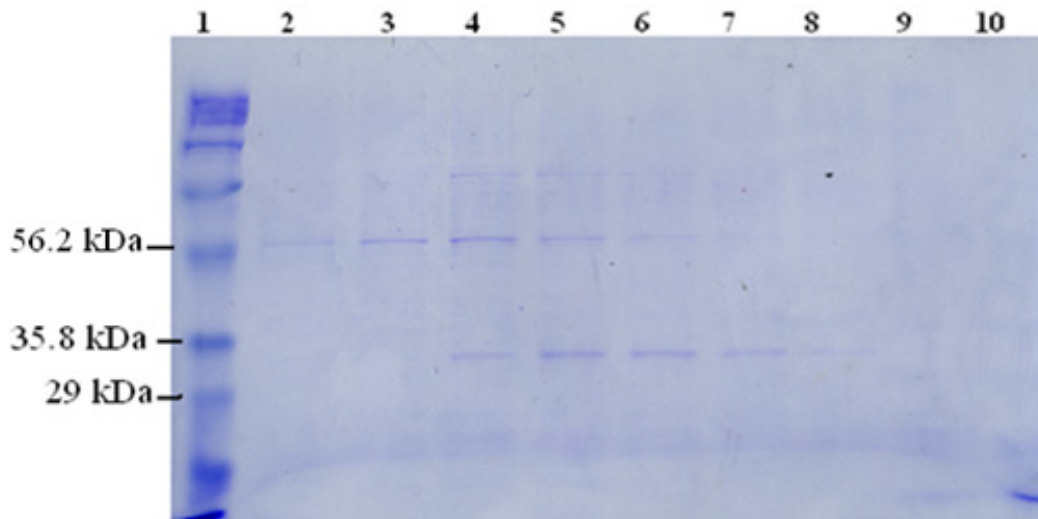
$1149 \times 0.037 \text{ kDa} = \text{endelig produkt på } \sim 42.5 \text{ kDa}$



Figur 14: SDS-PAGE med pilotekspresjon av antatt *mll3104* ved bruk av IPTG. **Brønn 1:** Prestained Protein Standard Broad Range, Invitrogen. **Brønn 2:** 0. time uten IPTG. **Brønn 3:** 2. time med IPTG. **Brønn 4:** 2. time uten IPTG. **Brønn 5:** 4. time med IPTG. **Brønn 6:** 4. time uten IPTG. **Brønn 7:** 6. time med IPTG. **Brønn 8:** 6. time uten IPTG. **Brønn 9:** 8. time med IPTG. **Brønn 10:** 8. time uten IPTG.

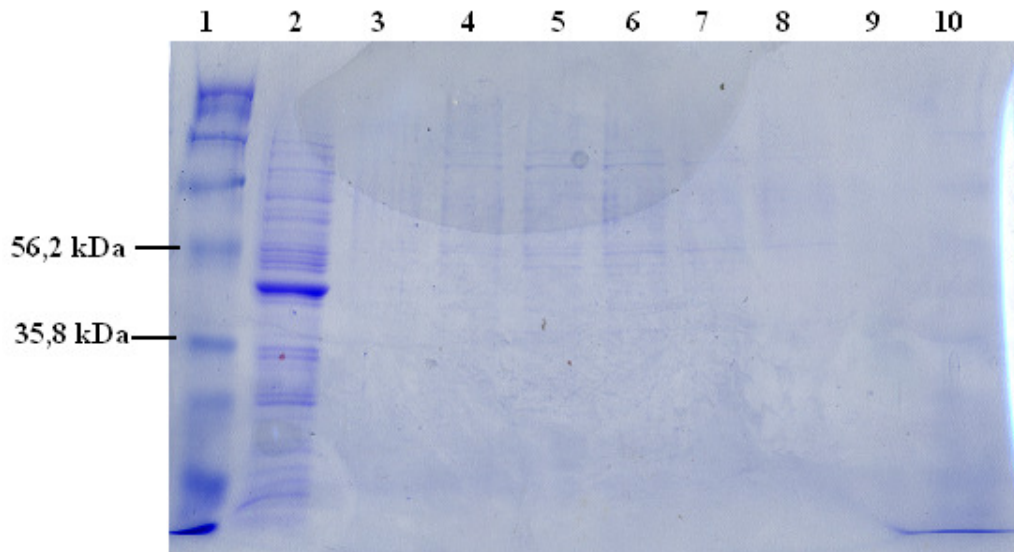
Eluering av MII3104-fraksjoner med ulike konsentrasjoner av imidazole

Proteinekstrahering av MII3104-proteiner ga jevnt over dårlig resultat, og figur 15 viser en SDS-PAGE gel der ukjente proteinfraksjoner er vasket ut fra affinitetskolonne med 200mM imidazol. Det ble utført vasking av kolonnen med henholdsvis 50, 150, og 200 mM imidazol. Samtlige vaskeprosesser for rensing av proteinet ga dårlig resultat, uten noen bånd av forventet størrelse. Proteiner av interesse burde ha ligget midt mellom 56.2 og 35.8 kDa, men ingen bånd av slik størrelse kan synes på gelen. Da resultatet er likt for andre vaskeprosesser med imidazole, blir kun et enkelt bilde presentert her.



Figur 15: Viser eksempel på SDS-PAGE med MII3104-proteinekstrakt vasket med 200 mM imidazol. **Brønn 1:** Markør (Prestained SDS-PAGE Standard, Bio-Rad). **Brønn 2-10:** Ukjente proteinfraksjoner.

Et enkelt unntak for vask av MII3104-proteiner er en SDS-gel der rålysatet ble samlet opp og kjørt på gel sammen med ukjente fraksjoner. I dette eksperimentet ble rålysat fra ekstraheringsprosessen (figur 16, brønn 2) samlet opp etter påføring i affinitetskolonnen, og dette ble kjørt på SDS-PAGE sammen med ukjente fraksjonsprøver (brønn 3-10). Ved dette ene tilfellet ses et bånd i brønn 2 som skiller seg ut. Båndet er meget markant og indikerer en høy konsentrasjon av proteiner. Av størrelse ser det ut til at proteinet har en molekylvekt på ~ 42 kDa, noe man kan forvente av MII3104-proteiner. Det er ingen andre interessante bånd for noen av de ukjente fraksjonsprøvene, noe som er gjennomgående for alle SDS-PAGE kjørt med MII3104-proteiner.



Figur 16: Viser en SDS-PAGE med antatte Mll3104 proteinfraksjoner og råekstrakt etter vask med 200mM imidazol. **Brønn 1:** Prestained SDS PAGE Standard, Invitrogen. **Brønn 2:** Lysat oppsamlet etter påføring i kolonne. Omtrentlig størrelse på 42,5 kDa. **Brønn: 3-10:** Ukjente fraksjonsprøver.

Mll3105-pilotekspresjons forsøk

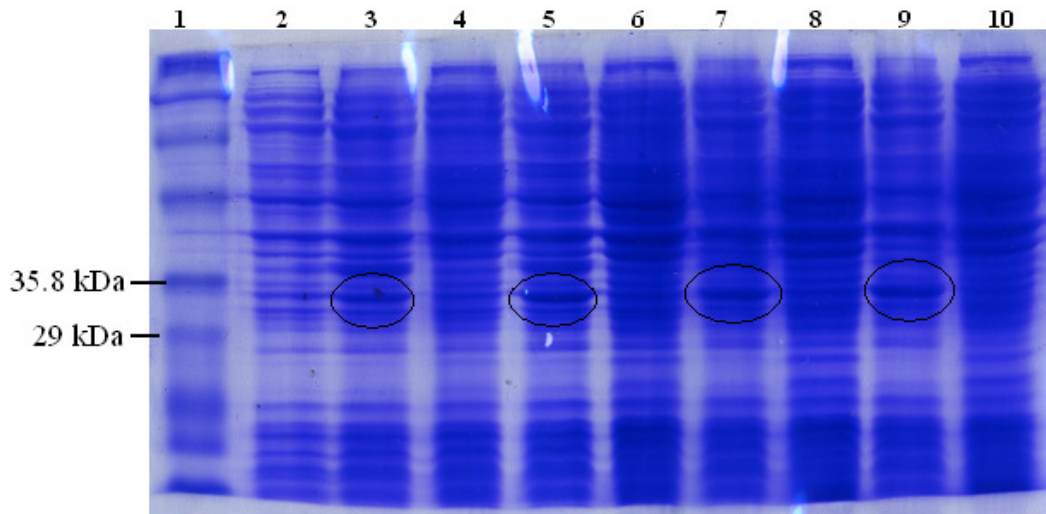
Pilotekspresjonsforsøk ble utført etter prosedyre angitt i metodedel. Forventet resultat er som for pilotekspresjonsforsøket med MLL3104-proteiner. SDS-gelen på figur 17 viser brønner tilsatt 10 μ l rålysate fra BL21 Star med pETmll3105, etter induksjon med IPTG (brønner 3, 5, 7 og 9) sammen med kontrollkultur uten IPTG induksjon (brønner 2, 4, 6, 8 og 10). Som forventet er det uttrykt store mengder proteiner i cellene. Formålet med dette eksperimentet var å identifisere Mll3105-proteiner, og resultatene tyder på at et mulig relevant protein er tilstede.

Produktstørrelse til et potensielt Mll3105-protein estimeres ut fra følgende: 1 kb DNA = 37 kDa. Formodet lengde på *mll3105*-gen er 786 bp (data fra GenBank). Etter innsett i pETmll3105 vil tilleggsbaser fra ligeringspunkt til termineringskodon (99 bp) inkluderes i det endelige produktet:

mll3105-produkt: 786 bp + 99 bp som gir endelig produkt på 885 bp.

$885 \times 0.037 \text{ kDa} = \text{endelig produktstørrelse på } \sim 32.7 \text{ kDa}$

Synlige bånd i brønnnummer 3, 5 og 7 antyder at tilsatt IPTG har induisert spesifikke sekvenser, noe som har resultert i bånd tilsynelatende unike for disse brønnene. Båndene av interesse er markert med sorte sirkler, og er $\sim 32 \text{ kDa}$.

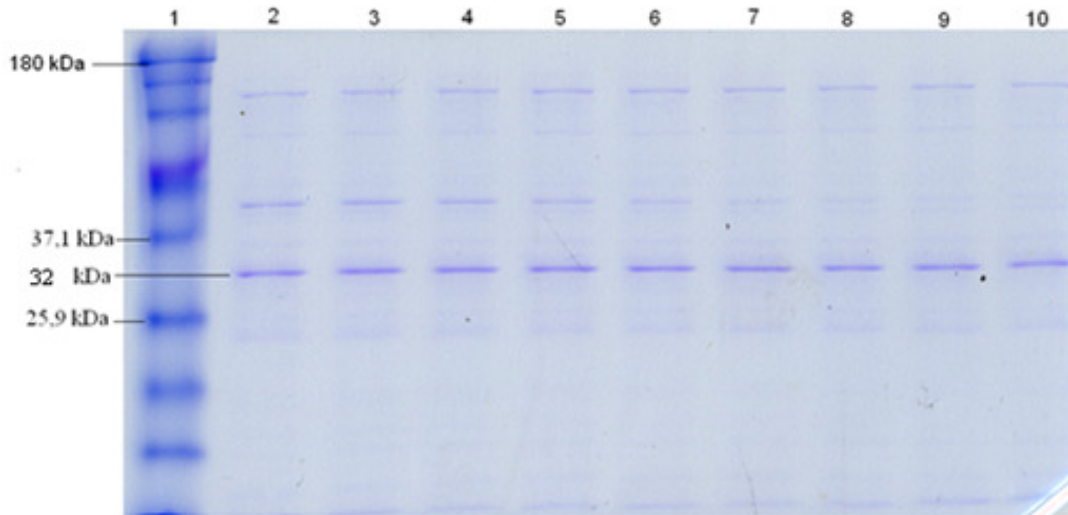


Figur 17: Viser en SDS-PAGE etter pilotekspresjonsekspertiment med IPTG-indusert uttrykk av Mll3105-proteiner fra rålysat av BL21 Star celler. **Brønn 1:** Markør (Prestained SDS-PAGE Standard, Invitrogen). **Brønn 2-10:** Ukjente proteinfraksjoner. Brønn 3 (2 timer), 5 (4 timer), 7 (6 timer) og 9 (8 timer) er tilsatt IPTG. Sorte sirkler i brønner med IPTG-indusert uttrykk indikerer potensielle Mll3105 proteiner av størrelse ~32 kDa.

Eluering av Mll3104-fraksjoner med ulike konsentrasjoner av imidazole

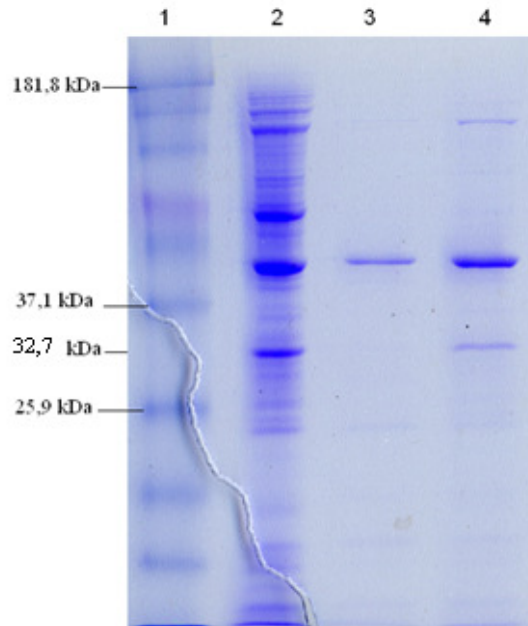
Etter bruk av ulike konsentrasjoner av imidazole under vaskeprosessen i affinitetskolonnen, ble det utført protein assay. Det ble utført forsøk med de samme imidazolkonsentrasjonene som for eksperimenter med Mll3104-protein uttrykk, med variable resultater.

Imidlertid ble det besluttet at 200 mM imidazol under vaskeprosessen, ga høyest protein innhold ved protein assay. Figur 18 viser en SDS-PAGE gel utført med 200 mM imidazole. Båndene av interesse er tilstede i brønn 2-10, og størrelsen er ~32 kDa. Det ble utført repetitive forsøk med denne konsentrasjonen av imidazole under vaskeprosess, og bånd ble skåret ut og sendt til Q-TOF MS ved Institutt for farmasi (UIT).



Figur 18: SDS-PAGE med antatte Mll3105-proteiner preparert fra proteinekstrahering BL21 Star *E. coli* celler. **Brønn 1:** Prestained Protein Standard markør (Benchmark, Invitrogen). **Brønn 2-9:** Ukjente fraksjons prøver. Båndstørrelse på ukjente prøver er ca 32 kDa.

Da resultatene var så ulike for fraksjonssamlinger i affinitetskolonnen, ble løsninger fra ulike trinn samlet og påført i kolonnen (figur 19). Det ble tatt prøve av supernatanten fra rålysat etter B-PER proteinekstrahering for å undersøke proteininnhold i løsningen, og brønn 2 viser supernatanten fra proteinekstraherings prosessen av BL21 Star celler med vektor pETmll3105. Som ventet er det flere bånd som er fremtredende i brønn 2, båndet som antas å inneholde Mll3105-proteiner har størrelse ~32 kDa. Brønn 3 inneholder supernatant samlet opp etter påføring i affinitetskolonnen. I brønnen er det aktuelle båndet fraværende, noe som indikerer at proteinene fra brønn 2 har hatt affinitet til kolonnematerialet. Det er imidlertid flere bånd tilstede i brønn 2, som etter påføring i kolonnen ikke finnes i brønn 3. Brønn 4 viser en prøve der man samlet opp proteinekstraktet etter vask med 50 mM imidazole. Det nederste båndet har om lag samme størrelse som båndet av interesse i brønn 2, og det kan indikere at det er samme protein. Det er flere ulike proteiner som har blitt vasket ut, men de andre har feil størrelse ut fra estimert produktstørrelse, og disse blir derfor ikke tatt med i videre betraktning.



Figur 19: SDS-PAGE av antatte MII3105 proteinprøver preparert fra ekstrahering av BL21 Star. Proteinet av interesse kan sees som nederste bånd i brønn 2 og 4, og har en omtrentlig størrelse på 32 kDa. **Brønn 1:** Prestained protein markør (Benchmark, Invitrogen). **Brønn 2:** Supernatant fra protein ekstraheringsprosess. **Brønn 3:** Supernatant samlet opp etter påføring i kolonne. **Brønn 4:** Løsning samlet opp etter vask med 50 ml 50mM imidazol.

Q-TOF MS analyser

Etter analyser av innsendte MII3105-proteinprøver, ble molekylære masse estimert for hver prøve, og søkt opp gjennom programvare (MASCOT) som sammenlikner kjente proteins molekylmasse mot ukjente analyseresultater. Tabell 19 viser en oversikt over de treff som ble gjort da innsendt prøve MII3105 a, b og c ble analysert.

Tabell 19: Viser treff etter Mascot (Matrix Science) søkefunksjon for proteinidentifikasjon. Tre ulike prøver (MII3105 a, b og c) er presentert i tabellen, og det angis enzymnavn og artsopprinnelse for enzymet. Proteintreff står oppført med synkende sannsynlighet for treff. Alle proteintreff viser enzymer av bakteriell opprinnelse, bortsett fra trypsinforløper, som er funnet i gris.

Proteintreff i Mascot database (NCBI)	MII3105 a	MII3105 b	MII3105 c
1	β-laktamase <i>Escherichia coli</i> (gil208244)	TEM-26B β-laktamase <i>Klebsiella oxytoca</i> (gil149169)	β-laktamase <i>E. coli</i> (gil208244)
2	ES-β-laktamase <i>E. coli</i> (gil2765071)	ES-β-lactamase <i>E. coli</i> (gil2765071)	ES-β-lactamase <i>E. coli</i> (gil2765071)
3	β-laktamase <i>Erwinia amylovora</i> (gil27228440)	β-laktamase <i>E. amylovora</i> (gil27228440)	Trypsin forløper <i>Sus scrofa</i> (eukaryotisk opprinnelse) (gil136429)
4	Glykosyltransferase <i>Ralstonia eutropha</i> <i>H16</i> (gil113868842)	Glykosyltransferase <i>R. eutropha H16</i> (gil113868842)	
5		Trypsin forløper <i>Sus scrofa</i> (eukaryotisk opprinnelse) (gil136429)	

Proteinene som ble funnet er ikke de man forventet ved analyse av MII3105-produkter. Et produkt fra *mII3105*-genet kan i teorien utgjøre et protein med funksjon for utnyttelse eller transport av trehalose, dersom det antas å ha samme funksjoner som *thuAB* i *S. meliloti*. De fleste treff indikerer at innsendt materiale er en β-laktamase, et enzym som har sitt opphav fra *bla*-gener i *E. coli*. Enzymet er uttrykt fra resistensgener mot penicilliner, som for eksempel ampicillin og brukes i vektorer. pET101/D-TOPO er designet med resistens for ampicillin, og antibiotikumet brukes som markør for transformanter. Det er derfor naturlig at det finnes β-laktamaser i lysat fra cellene. Årsaken til at β-laktamase har hatt affinitet til kolonnematerialet

er usikkert, og vil bli diskutert under neste seksjon. Imidlertid er en mulighet at enkelt-histidiner kan ha bundet seg til kolonnematerialet.

Sammenlikning av aminosyrer mellom eget innsett og opprinnelig sekvens av *mll3105*

For sammenlikning av aminosyrer ble den opprinnelige aminosyresekvensen for *mll3105* (*M. loti*) anvendt sammen med egen innsatt *mll3105*-sekvens fra den optimaliserte sekvenseringsreaksjonen (figur 20).

Sammenlikningen ble utført ved å importere innsatt *M. loti mll3105*-sekvens i programvaren Web Map References (http://pga.mgh.harvard.edu/web_apps/web_map/start), for deretter å bestemme hvilke aminosyrer som er ulike mellom de to sekvensene. Appendiks VIII viser resultatet av egen sekvens importert i Web Map References. Sekvensanalysen viser i tillegg til aminosyresekvensen at det finnes et enkelt stopkodon tilstede i sekvensen, og dette er lokalisert på rett sted. Totalt fremkommer 4 aminosyresubstitusjoner i eget *mll3105*-fragment, disse markeres i gult. Aminosyrene som er byttet ut, kan være utslagsgivende for proteinets funksjon:

1. Lysin i opprinnelig sekvens er erstattet med prolin i eget innsett.
2. Glutamat i opprinnelig sekvens er erstattet med glysin i eget innsett.
3. Tyrosin i opprinnelig sekvens er erstattet med cystein i eget innsett.
4. Aspargin i opprinnelig sekvens er erstattet med aspartat i eget innsett.

MPIKAVVWGE NVHEQTNAAV RDLYPLTMHG TIAAALNQDK GIEATTATLQ EPEHGLSEKR
MPIKAVVWGE NVHEQTNAAV RDLYPLTMHG TIAAALNQDK GIEATTATLQ EPEHGLSEPR
LAATDVLLWW GHAAHGEVKD EIVERVQKRV WEGMGLIVLH SGHYSKIFKR LMGTPCSLKW
LAATDVLLWW GHAAHGEVKD EIVERVQKRV WEGMGLIVLH SGHYSKIFKR LMGTPCSLKW
REAGERERVW AINRGHPIAQ GIGECLEIGE TEMYGEPFAV PEPMETV FVS WYEGGEVFRS
REAGERERVW AINRGHPIAQ GIGECLEIGE TEMYGEPFAV PEPMETV FVS WCEGGEVFRS
GLTYQRGAGR IFYFSPGHET YPIYHNEGVQ QVLRNAVHWA LNPAPAWSGI TNAPNVPTDK
GLTYQRGAGR IFYFSPGHET YPIYHNEGVQ QVLRDAVHWA LNPAPAWSGI TNAPNVPTDK
AKEKIVQKGL RLHADGDKGL S
AKEKIVQKGL RLHADGDKGL S

Figur 20: Viser aminosyresekvenser av henholdsvis det opprinnelige *mll3105* genet fra *M. loti* (øverste sekvens) hentet fra Genbank, og egen sekvens av *mll3105* etter en optimalisert sekvenseringsreaksjon (nederste sekvens). Aminosyrer som er byttet ut er markert i gult.

Diskusjon

Ampicillin som seleksjonsmarkør

En sentral faktor for vellykkede ekspresjons-eksperimenter, er seleksjon etter transformasjon ved bruk av antibiotika. Ampicillin er et semisyntetisk penicillin-derivat som ofte anvendes som seleksjonsmarkør i genkloning og ekspresjonseksperimenter av protein i *E. coli* og andre bakterier (Madigan, Martinko et al. 2002). Det er imidlertid visse begrensninger for bruk av markøren, og dette kan gi problemer under seleksjonsprosessen.

Grunnlaget for ampicillinseleksjon er hydrolyse av antibiotikumet ved katalytisk β -laktamase, uttrykt fra det plasmidbårne *bla*-genet. Ved tilstedeværelse av ampicillin vil transkripsjonen av *bla*-gener induseres, med påfølgende syntese av β -laktamase. Resultatet blir en ekstracellulær oppbygning av β -laktamase i kulturmediet som raskt kan inaktivere tilsatt ampicillin. Dette vil kunne fjerne seleksjonspresset, og gi forhold der celler uten plasmid kan få vekstbetingelser lik transformantenes. Resultatet kan da bli en kultur der mange av cellene mangler plasmidet, og med påfølgende dårlig opprensing av eventuelle plasmidpreparasjoner eller senere proteinekstrakt.

Et alternativ for ampicillin dersom man opplever problemer rundt seleksjonen, kan være å bruke karbencillin som seleksjonsmarkør. Denne typen antibiotika vil inaktiveres på samme måte som ampicillin, men over en lengre tidsperiode. Karbencillin er ikke prøvd ut i eksperimenter under dette studiet da man ikke kunne forutse eventuelle problemer ved bruk av ampicillin, i tillegg er karbencillin et mer kostbart alternativ.

En mulig årsak til at det kunne oppstått non-selektivt forhold i kulturen kan være bruk av gammel antibiotika, som kunne vært delvis degradert og inaktivt. Denne muligheten kan imidlertid ekskluderes da det ble laget ferske ampicillinpreparater for dette studiet.

Et mulig problem ved gjennomførelsen av mine eksperimenter kan være for høy celledetthet i flytekulturer, det anbefales OD_{600} under 3.0 (i LB-medium). Det ble ikke foretatt måling av OD_{600} etter vekst over natt for ekspresjonsstammen. Derfor er det ikke mulig å utrede om dette kan være en aktuell feilkilde for manglende uttrykk av innsatte gener, men muligheten kan ikke utelukkes. Dersom cellene har entret lag-fase, vil uttrykk av proteiner være på et betydelig lavere nivå.

I eksperimentene ble det brukt standard antibiotikakonsentrasjoner på 50 $\mu\text{g/ml}$, etter anbefaling fra kit-manual. En høyere konsentrasjon av ampicillin i mediet ville forlenget tiden

for inaktivering av ampicillinet, noe som ville forbedret seleksjonen. Dette ble ikke prøvd for eksperimentene, så det er ikke mulig å avgjøre om for lav antibiotikakonsentrasjon kunne vært årsak til et non-selektivt forhold i kulturene, med påfølgende plasmidtap og dårlig proteinuttrykk (Bennett and Chopra 1993; Invitrogen-life-technologies 2004).

Amplifiserings resultater av genomisk *M. loti* DNA

Innledningsvis under amplifisering fra templat ble det benyttet en polymerase uten korreksjonslesing (Dynazym polymerase). Det kan ikke utelukkes at ukorrekte baser ble satt inn under dette trinnet. Dersom dette er tilfelle ville alle forsøk videre bevart sekvensen med substituerte baser. Senere bruk av polymerase med eksonuklease korrekturaktivitet (Pfx50 polymerase) ville derfor ikke kunne rette opp de ukorrekte basene. For fremtidige eksperimenter bør det vurderes om man kun bør benytte den sistnevnte polymerasen, selv om det innebærer problemer med amplifiseringen fra genomisk *M. loti*-templat.

DNA-molekylene utgjør ikke stabile strukturer, og ved UV-bestråling er det essensielt å minimalisere strålingstiden for å unngå skader på DNA'et. Dannelse av pyrimidin-dimerer, tap av baser, brutte sukker-fosfat bindinger eller intertråd DNA-DNA kryssbindinger er vanlige UV-relaterte skader, og man må i tillegg ta med i betraktning de baseutbyttene som kan forekomme under senere replikasjon, der DNA-reparerings system kan sette inn ukorrekte baser for å unngå terminering av sekvenser. En forskningsrapport viser at den korte tiden (20-40 sekunder) man eksponerer DNA-båndene for UV-lys for å kutte bånd fra en agarosegel, er nok til å gjøre ekstensiv skade på DNA-molekylene. Forfatterne foreslår tilsetning av 1 mmol/L guanosin eller cytidin i buffer, både ved laging av agarosegel og ved kjøring av gelen. Guanosin har en struktur som vil absorbere UV-strålingen, og dermed gjøre DNA-fragmenter i gelen mindre utsatt for strålingsskader. Det er uvisst under hvilket trinn av studiet man fikk introdusert baseutbyttene på, men det er velkjent at DNA-skader lett kan oppstå under UV-stråling (Grundemann and Schømig 1996). Dersom tilsetning av guanosin eller cytidin i agarosegeler kan redusere problemet, og bør dette vurderes benyttet i eksperimenter som utsetter DNA-molekylene for denne typen skadelige fotoreaksjoner.

Aminosyreendringer i *mll3105*-sekvens etter innsett i vektor, og mulige konsekvenser for proteinfunksjon

Etter DNA-sekvensering av *mll3105*-fragmentet, viste det seg at 4 baser var substituert. De samme substitusjonene går igjen for alle sekvenseringsresultater, så man kan anta at dette ikke er et enkelttilfelle. Basesubstitusjoner trenger i utgangspunktet ikke å ha betydning for aminosyresekvensen, da flere ulike base-tripletter kan kode for samme aminosyre. En forskyvning av leserammen, ved addisjon eller delesjon av enkeltbaser under replikasjon, ville fått større konsekvenser.

Etter at *mll3105*-sekvensen ble eksportert inn i Web Map References, fant programvaren startkodon, stopkodon, samt aminosyrerekkefølge ved translasjon. Funnene viste at leserammen ikke var endret. Startkodon var ved -CACC-overhenget designet for ligering i ekspresjonsvektor pET101/D-TOPO. Stopkodon ble funnet direkte etter 6xHis-sekvensen. Histidin-sekvensen er vektor-DNA derivert, og utgjør delen av proteinet med affinitet til nikkelladede resiner ved rensing av proteinfraksjoner. Videre viste analysen at totalt 4 aminosyrer var endret:

1. Opprinnelig lysin ble erstattet med prolin.
2. Opprinnelig glutamat ble erstattet med glysin.
3. Opprinnelig tyrosin ble erstattet med cystein.
4. Opprinnelig asparagin ble erstattet med aspartat.

Dette har ført til en endring i aminosyrerekkefølgen (primærstrukturen) i proteinet. I tillegg vil sekundærstruktur en påvirkes, dersom de nye aminosyrene har andre egenskaper enn de opprinnelige. Proteinets sekundærstruktur sørger for stabile arrangementer av aminosyrene, og deles blant annet inn i "α-helikser", "β-sheets" og "turns". Videre kan proteinfoldingen (tertiærstruktur) påvirkes, terciærstrukturen blir spesielt affisert dersom de nye aminosyrene har egenskaper ulikt fra de opprinnelige. Utfallet kan i ytterste konsekvens være at proteinet taper funksjonen, eller ikke kan interagere med andre proteinsubenheter under kompleksdannelse. Kort sagt bestemmes peptideegenskapene av de fysiokjemiske egenskapene til enkelte aminosyrer og bindinger.

Det potensielle Mll3105-proteinet er ikke karakterisert tidligere. Man vet derfor lite om det både når det gjelder funksjon, stabilitet, bindinger og tredimensjonal struktur. Aminosyrenes struktur er i utgangspunktet ganske lik, alle innehar en karboksylgruppe og en aminogruppe

bundet til samme karbonatom. Av 20 standard aminosyrer har 19 sidegruppene sine bundet opp til α -karbonet, og forekommer i kjedeform. Unntaket er prolin, som har en syklisk struktur. Generelt måles de enkelte aminosyrers egenskaper etter den funksjonelle gruppen som varierer i størrelse, struktur og elektrisk ladning. I tillegg vil den funksjonelle gruppen avgjøre om aminosyren er løseligheten i vann. For å undersøke mulige konsekvenser av aminosyresubstitusjonene, ble de enkelte aminosyrers egenskaper vurdert.

Lysin → prolin substitusjon

Lysin er en positivt ladd hydrofil aminosyre, mens prolin er en ikke-polar hydrofob syre. En aminosyres løselighet forteller noe om hvordan molekylet står vendt i senere folding. En hydrofob aminosyre vil vende seg innover i molekylet, mens løselige aminosyrer vendes ut. Prolin har i tillegg en syklisk struktur, som gjør at polypeptider inneholdende prolin får en lite fleksibel struktur. Den rigide konfirmasjonen gjør at peptider med prolin er stabile strukturer. Lysin og prolin er aminosyrer med helt forskjellige egenskaper. Det kan derfor tenkes at substitusjonen er avgjørende for α -heliks struktur, og dermed foldingen, og endelig funksjonen av proteinet.

Glutamat → glysin substitusjon

Glutamat er en hydrofob aminosyre med negativt ladd sidegruppe. Glysin er ikke-polar og med en alifatisk sidegruppe, og av hydrofob natur. Glysin er den minste og enkleste aminosyren av de 20 standard aminosyrene, og skaper fleksibilitet i peptidkjedet på grunn av et H-atom som sidegruppe. Glysin vil imidlertid kunne indusere en knekk i α -heliks, og dermed endre foldingen. Glutamat er et vesentlig større molekyl, på grunn av en ekstra karboksylgruppe. De to syrene har svært forskjellige egenskaper, og substitusjonene kan tenkes å endre proteinfolding og funksjon.

Tyrosin → cystein substitusjon

Tyrosin har en aromatisk sidegruppe, og er relativt hydrofob. En hydroksylgruppe i tyrosin utgjør en viktig egenskap, hydroksylgruppen kan danne hydrogenbindinger, en viktig egenskap i noen enzymer. Cystein er hydrofil, med uladde sidegrupper. Cysteinets sulhydrylgruppe gjør at aminosyren oksiderer lett, og danner en kovalent bundet dimerisk aminosyre kalt cystin, der to cystiner bindes gjennom en disulfidbro. Disulfidbindinger er sterkt hydrofobisk, og spiller ofte en rolle i mange proteiner ved å danne kovalente bindinger

mellom deler av enkeltprotein og mellom ulike proteinkjeder. Substitusjonen kan tenkes å endre foldingen, spesielt dersom det finnes andre cysteiner i kjeden.

Aspargin → aspartat substitusjon

Disse to aminosyrene har relativt lik størrelse, og begge er polare. Aspartat er egentlig aspargin hydrolysert av en syre eller base, med en negativt ladd sidegruppe. Den ervervede OH-gruppen kan ha sentrale funksjoner for proteinet. Aspargin har god løselighet i vann, på grunn av sidegruppen som kan danne hydrogenbindinger med vann. Aspartat er også løselig, og er sammen med glutamat en av to sure aminosyrer. Selv om disse aminosyre er relativt like i oppbygning, størrelse og løselighet, kan man ikke utelukke at den totale funksjonen av det foldete proteinet endres (Nelson 2000).

Totalt sett kan det ikke utelukkes at aminosyresubstitusjonene har hatt konsekvenser for proteinfunksjonen. Det er svært vanskelig å forutsi endringer da man ikke vet hvor på proteinet endringene vil komme til uttrykk. Er det endringer i kritiske punkter, for eksempel i en α -heliks, er det ikke usannsynlig at foldingen blir annerledes eller at proteinet taper funksjonen. En annen faktor som må tas hensyn til, er den endelige foldingen av et potensielt funksjonelt protein. Dersom 6xhistidin-markøren vender innover i den hydrofobe delen av proteinet, vil dette kunne påvirke affiniteten til resinene.

Rensing av MII3104 og MII3105 ved bruk av affinitetskolonner

Fraksjonssamling av kolonneseparerte proteiner er en ofte anvendt metode som baserer seg på at proteinet av interesse tilføres en kolonne som mikstur fra et råekstrakt. Kolonnen er fylt med polymerer med en ligand spesifikk for proteinet som skal opprenses, og der proteinet bindes opp i kolonnematerialet. Uønskede proteiner uten affinitet til nikkelladde resiner, vaskes ut av kolonnen, mens det ønskede proteinet elueres ved bruk av en ligand-løsning med større affinitet til kolonnematerialet enn målprotein.

I dette studiet ble imidazol av forskjellige konsentrasjoner brukt for eluering av målprotein etter påføring i kolonnen. Imidazole er en ligand med sterkere affinitet til nikkelladde resiner enn målprotein, og vil fortrenge MII3104/MII3105 fra kolonnen ved adekvat konsentrasjon.

Det er flere trinn under prosessen som kan være avgjørende for om opprensingen blir vellykket. Dersom man antar at målprotein i utgangspunktet er uttrykt, men man ikke får resultater etter fraksjonssamling, er det naturlig å begynne med batchen av resiner. Denne er

ladd med Ni²⁺. Indikator for om resinene fungerer er en blå farge, og resinene som ble brukt under dette studiet var lys blå. Dette utelukker imidlertid ikke at produksjonen av resiner fungerer som den skal, da det ikke ble utført en positiv kontroll med proteiner man har god kunnskap om, og som er utprøvd ved tidligere anledninger. Av ukjent årsak tok hver fraksjonssamling lang tid (>24 t), og da studiet hadde en tidsbegrensning, valgte man å ikke prøve ut om systemet fungerte med en positiv kontroll. Det har tidligere vært gjennomført vellykket fraksjonssamling med homologer av målproteinene (ThuA og ThuB fra *S. meliloti*), og det ble derfor antatt at resinene fungerte. I ettertid er det klart at en positiv kontroll innledningsvis kunne avgjort om resinene fortsatt fungerte optimalt, og dette bør tas med i betraktning for fremtidige eksperimenter, uavhengig om det er samme batch eller en ny som anvendes. Det kan også være en fordel å sammenlikne fargen på den anvendte resin-batchen med en nyere batch. Eventuelt burde man antatt at resinet hadde redusert ladning, og ladet resinene på nytt med NiCl₂ forut for eksperimentene. Produsenten (Invitrogen life technologies) oppgir at resinet kan opplades på nytt opptil 3 ganger, noe som indikerer at det kan være fordelaktig med oppladning etter en tids bruk.

Før fraksjonssamling i kolonnen tok til, burde det vært utført et Western blot med anti-his antistoff for å avgjøre om målproteinene var uttrykt, og for å bestemme om histidin-markøren var tilstede og ikke kløyd av proteinet. Det finnes ulike kommersielle antistoffer, avhengig av om his-markøren er lokalisert ved N- eller C-terminal. Ved dette studiet har man ikke utført denne kontrollen, så det kan ikke utelukkes manglende uttrykk av målproteinene, eventuelt et uttrykk uten histidin-sekvensen.

Dersom man tar utgangspunkt i at proteinene uttrykkes med histidin-markøren intakt, er et mulig problem med de dårlige resultatene for dårlig denaturering. En løsning kan være å gjennomføre lyseringsprosedyren av ekspresjonsstammen i LB en ekstra gang ved å resuspendere pelletet, tilsette lyseringsmiksen, og sentrifugere en ytterligere en gang. Alternativt kan det vurderes om koketid av prøvene kan forlenges.

På samme måte som manglende degradering av proteiner kan utgjøre problemer, kan også proteiner degraderes i for stor grad. En løsning for dette kan være å øke mengde proteasehemmer i proteinekstraheringsmiksen, eller prøve en annen proteasehemmer i tilfelle at den anvendte ikke fungerer optimalt.

Et annet mulig problem kan være at tidspunktet for optimalisert uttrykk ikke er korrekt. Dersom man stopper kulturen for tidlig, vil ikke uttrykket være optimalt, og det samme gjelder for cellelysis av en kultur med for lang vekstperiode. Det siste er spesielt viktig dersom proteinet er ustabil og degraderes etter kort tid.

Det ble antatt at målproteiner ble optimalt uttrykt, og at his-markøren ikke var kløyvd. De fleste resultater etter fraksjonssamling var endog uten tilstedeværende proteiner. Dette kan tyde på at målproteiner ikke ble bundet opp initialt, eller at ønskede fraksjoner har for lav affinitet til kolonnematerialet. Resultatet vil da være at fraksjoner vaskes ut for tidlig, noe som kan indikeres i figur 16. I denne gelen er det satt på prøve etter påføring i kolonnen (brønn 2), og det er et spesielt intensivt bånd som skiller seg ut. Båndet har en omtrentlig størrelse av det man forventer av et estimert MII3104-protein, men det er vanskelig å avgjøre noe da man fikk gjennomgående dårlige resultater for alle fraksjonssamlinger med MII3104-proteiner.

For MII3105-proteiner kan figur 19 indikere at et protein har hatt affinitet til resinene i kolonnen. Brønn to viser en prøve av råekstraktet fra lyseringsprosessen. Det er et bånd av en størrelse man kan forvente av MII3105-proteiner som skiller seg ut. Brønn 3 viser oppsamlet løsning etter påføring på kolonnen, og båndet som skiller seg ut i brønn to er fraværende. Dette tyder på en affinitet for kolonnematerialet, selv om det ikke er ensbetydende med et MII3105-protein. Et bånd av samme størrelse er tilstede i brønn 3, der man har vasket kolonnen med 50 mM imidazole. Båndet er ikke så intenst, men det kan se ut som om det har riktig størrelse. Ved et annet tilfelle (figur 18) ble et protein av samme størrelse rensset ut fra kolonnen. Det er ikke mulig å fastslå at det er et MII3105-protein, men ut fra størrelsen er det en mulighet.

At man finner igjen proteinprøver i supernatant etter påføring på kolonnen, kan tyde på for mye tilsatt prøvemateriale. Dette kan være aktuelt for MII3104 (figur 19, brønn 2), og man kan da redusere mengde tilsatt prøve på kolonnen. For mye prøvemateriale er noe usannsynlig, da rensingen generelt ga dårlige resultater, men det kan likevel ikke utelukkes, siden man fant proteiner etter påføringen.

Da det viste seg at de antatte MII3105-målproteiner var en β -laktamase, må man anta at det finnes flere histidiner i aminosyresekvensen. Normalt skal ikke enkeltvis histidiner gi affinitet til kolonnematerialet, men i dette tilfellet har en β -laktamase klart hatt affinitet til resinene. En β -laktamase etter vaskeprosess kan tyde på at vaskeprosedyrene ikke er

stringente nok, en løsning for å gjøre resultatene bedre da er å senke pH i vaskebufferen, samt øke volum av buffer. Dette vil kunne fjerne uønskede proteiner fra kolonnen. Alternativt kan en eluere prøvematerialet en gang til på samme kolonne, eller anvende et annet kolonnesystem.

DNA-sekvensering

Alle sekvenseringresultater for *mll3105*, viste at 4 enkeltbaser var byttet ut, og man kan derfor anta at dette er et riktig resultat. Dette ville bare ha konsekvenser for senere proteinuttrykk dersom basesubstitusjonene førte til endringer i leserammen eller at triplettene kodet for nye aminosyrer. Ved bruk av en programvare som analyserer basesekvenser, kom det fram at baseutbyttene hadde hatt en konsekvens for proteinet ved at 4 aminosyrer var endret i forhold til originalsekvens.

Hovedproblemet med sekvenseringsresultatene var en fellingsprosedyre ved sekvenseringslaben, som ikke fungerte optimalt. Under studiets gang ble type kit for presipitering endret, og det nye skulle fjerne de DNA-toppene som kom etter ca 70-80 bp. Dette førte til at nyere innsendte prøver fikk et bedre sekvenseringsresultat. Hadde man fått et bra resultat ved de innledende sekvenseringer, ville man ikke gått videre til ekspresjonsforsøk, men heller amplifisert templatet til innsettet ble korrekt. Et av områdene det hersket usikkerhet rundt, da man fikk så dårlige resultater etter fraksjonssamlinger av antatt Mll315-protein, var om -CACC overhenget ikke var korrekt. Startkodonets baserekkefølge hadde gjennomgående vært vanskelig å lese, på grunn av overskudds-DNA fra sekvenseringsreaksjonen rundt 70-80 bp. Etter en optimalisering av sekvenseringsreaksjonene fikk vi bekreftet at -CACC ved innsettsstart var korrekt. Et manglende eller ufullstendig uttrykk av målproteiner, stammer da fra baseutbyttene.

For *mll3104*-sekvenser som ble sendt inn til sekvensering, var resultatene langt dårligere. Det ble ikke funnet enkeltbaser som var byttet ut, men det utelukker ikke muligheten for at det har skjedd en feil under amplifiseringen. Alle sekvenseringsresultatene for *mll3104* hadde for høyt antall ukjente baser til å kunne trekke konklusjoner. Etter at sekvenseringslaben endret fellingsprosedyren, ble det kun sendt inn prøver fra *mll3105*-amplifisering. En optimalisert sekvenseringsreaksjon burde vært gjennomført for *mll3104* også, spesielt da senere ekspresjons eksperimenter ga så dårlige resultater.

Konklusjon

Et av hovedmålene med dette studiet var om mulig å karakterisere hypotetiske Mll3104- og Mll3104-proteiner fra *Mesorhizobium loti*. Tidligere arbeid med en nært beslektet stamme, *Sinorhizobium meliloti*, førte til identifisering og karakterisering av genene *thuA* og *thuB*, involvert i transport og utnyttelse av trehalose.

På grunn av laboratoriske problemer i kombinasjon med tidsbegrensninger, kom man ikke så langt som til karakterisering av proteinene. Dette studiet var et pilotprosjekt på en sådan måte at tidligere arbeid med disse genene og deres produkter ikke har vært gjort. De samme metodene som fungerte under arbeid med *thuA* og *thuB*, viste seg å ikke fungere for ortologene *mll3104* og *mll3105*. Selv om de fleste metodene som ble anvendt ikke viste seg å fungere optimalt, har dette studiet bidratt til en innledende forståelse av hvilke prosedyrer som må modifiseres for at man skal frembringe resultater. Ved fremtidige eksperimenter er designede primere for amplifisering av *mll3104* og *mll3105*-sekvenser fullt anvendbare, selv om type polymerase som benyttes under amplifikasjon bør vurderes. I tillegg bør man ikke gå videre med kloningsdel før et riktig innsett av egne sekvenser er bekreftet av repetitive sekvenseringreaksjoner. Det bør vurderes om man bør gå over til et annet kit for lysering av ekspresjonskultur, da det ikke kan utelukkes at lyseringsprosedyrene ikke denaturerte proteiner tilstrekkelig før fraksjonssamling tok til. Resiner bør være nyladede, og det bør definitivt utføres en positiv kontroll med et allerede kjent fungerende protein parallelt med fraksjonssamlingen av ukjent prøvemateriale. Det kan vurderes om man bør øke mengde vaskebuffer etter påføring av prøvemateriale på kolonnen, eller evt å senke pH i kolonnen. Begge alternativer kan påvirke eluering av både ønskede og uønskede proteinfraksjoner.

Referanser

- ABI-PRISM (2002). BigDye@Terminator v3.1 Cycle Sequencing Kit instruction Manual.
- Ampomah, O. Y. (2007). Competition for nodule occupancy in *Medicago spp.* and *Lotus spp.* - Role of disaccharide utilization. Avdeling for biologi, Det Naturvitenskapelige Fakultet, Universitetet i Tromsø.
- Avonce, N., A. Mendoza-Vargas, et al. (2006). "Insights on the evolution of trehalose biosynthesis." BMC Evol Biol **6**: 109.
- Bennett, P. M. and I. Chopra (1993). "Molecular basis of beta-lactamase induction in bacteria." Antimicrob Agents Chemother **37**(2): 153-158.
- Bhuvaneswari, T., B. Turgeon, et al. (1980). "Early Events in the Infection of Soybean (*Glycine max* L. Merr) by *Rhizobium japonicum*
I. LOCALIZATION OF INFECTIBLE ROOT CELLS " Plant Physiol **66**(6): 1027–1031.
- Boos, W., U. Ehmann, et al. (1990). "Trehalose transport and metabolism in *Escherichia coli*." J. Bacteriol. **172**(6): 3450-3461.
- Brady, N. C. (1984). The Nature and Properties of Soil, Macmillan Publishing Company.
- Brencic, A. and S. C. Winans (2005). "Detection of and Response to Signals Involved in Host-Microbe Interactions by Plant-Associated Bacteria
" Microbiol. Mol. Biol. Rev **69**(1): 155-194.
- Broughton, W. J., S. Jabbouri, et al. (2000). "Keys to Symbiotic Harmony
" J. Bacteriol. **182**(20): 5641-5652.
- Chandra, S., K. Choure, et al. (2007). "Rhizosphere competent *Mesorhizobium loti* MP6 induces root hair curling, inhibits *Sclerotinia sclerotiorum* and enhances growth of Indian mustard (*Brassica campestris*)." Brazilian Journal of Microbiology **38**: 124-130.
- De Smet, K. A. L., A. Weston, et al. (2000). "Three pathways for trehalose biosynthesis in mycobacteria." Microbiology **146**(1): 199-208.
- Elbein, A. D., Y. T. Pan, et al. (2003). "New insights on trehalose: a multifunctional molecule." Glycobiology **13**(4): 17R-27R.
- Engelhard, M. (2004). Trehalose and the nitrogen fixing nodule symbiosis of legumes: studies on rhizobia deficient in the trehalose-6-phosphate synthase gene *otsA*. Philosophisch-Naturwissenschaftlichen Fakultät, Basel, Universität Basel. Doktors der Philosophie.
- Farias-Rodriguez, R., R. Mellor, et al. (1998). "The accumulation of trehalose in nodules of several cultivars of common bean (*Phaseolus vulgaris*) and its correlation with resistance to drought stress " physiol. Plant **102**: 353-359.

- Fischer, H. M. (1994). "Genetic regulation of nitrogen fixation in rhizobia." Microbiol Rev **58**(3): 352-86.
- Gage, D. J. (2004). "Infection and Invasion of Roots by Symbiotic, Nitrogen-Fixing Rhizobia during Nodulation of Temperate Legumes." Microbiol. Mol. Biol. Rev. **68**(2): 280-300.
- Glenn, A. R. (1980). "The uptake and hydrolysis of disaccharides by fast-and slow-growing species of Rhizobium." Archives of Microbiology **Volume 129**(3): 238-239.
- Gonzalez, J. E. and M. M. Marketon (2003). "Quorum Sensing in Nitrogen-Fixing Rhizobia" Microbiol. Mol. Biol. Rev **67**(4): 574-592.
- Grundemann, D. and E. Schømig (1996). "Protection of DNA During preparative Agarose Gel Electrophoresis Against Damage induced by Ultraviolet Light." Biotechniques **21**: 898-903.
- Guerts R, B. T. (2002). "Rhizobium Nod factor perception and signalling." The Plant Cell **14**(Supplement): S239–S249.
- Gurevitch, J., S. M. Scheiner, et al. (2006). The Ecology of Plants, Sinauer Associates Inc.
- Hattori, Y., H. Omori, et al. (2002). "Ordered cosmid library of the Mesorhizobium loti MAFF303099 genome for systematic gene disruption and complementation analysis." Plant Cell Physiol **43**(12): 1542-57.
- Hengen, P. (1995). "Purification of His-Tag fusion proteins from *Escherichia coli*." Trends Biochem Sci **20**(7): 285-6.
- Invitrogen-life-technologies (2004). Champion pET Directional TOPO Expression kit instruction manual.
- Invitrogen-life-technologies (2004). TA Cloning Kit Instruction manual
- Invitrogen-life-technologies (2006). TOPO TA Cloning, pCR2.1-TOPO Instruction manual.
- Jensen, Einar v/ Institutt for Farmasi, UIT (2007). Personlig kommunikasjon.
- Jensen, J. B. (1999). Transport and utilization of trehalose in *Sinorhizobium meliloti*. Avdeling for Biologi, Det Matematisk-Naturvitenskapelige Fakultet. Tromsø, Universitetet i Tromsø. **Doktorgrads avhandling**.
- Jensen, J. B., O. Y. Ampomah, et al. (2005). "Role of trehalose transport and utilization in Sinorhizobium meliloti-alfalfa interactions." Molecular Plant-Microbe Interactions **18**(7): 694-702.
- Jensen, J. B., T. V. Bhuvaneshwari, et al. (1999). "Identification of a new genetic locus in Sinorhizobium meliloti involved in trehalose utilization." Manuskript.

- Jones, K. M., H. Kobayashi, et al. (2007). "How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model." Nat Rev Micro **5**(8): 619-633.
- Kaneko, T., Y. Nakamura, et al. (2000). "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti (supplement)." DNA Res **7**(6): 381-406.
- Long, S. and B. J. Staskawitz (1993). "Prokaryotic plant parasites." Cell **73**(5): 921-935.
- López, M., J. A. Herrera-Cervera, et al. (2006). "Trehalose metabolism in root nodules of the model legume Lotus japonicus in response to salt stress." Physiologia Plantarum **128**(4): 701-709.
- Madigan, M. T., J. Martinko, et al. (2002). Brock Biology of Microorganisms, Prentice Hall.
- Maruta K, N. T., Kubota M, Chaen H, Sugimoto T, Kurimoto M, Tsujisaka Y (1995). "Formation of trehalose from maltooligosaccharides by a novel enzymatic system." Biosci Biotechnol Biochem **59**(10): 1829–1834.
- Nelson, D. L., Cox, Michael M. (2000). Lenhinger Principles of Biochemistry, Worth Publishers.
- Oke, V. and S. Long (1999). "Bacteroid formation in the Rhizobium–legume symbiosis." Curr Opin Microbiol. **2**(6): 641-646.
- PIERCE B-PER® Bacterial Protein Extraction kit Manual.
- QIAGEN (2002). QIAquick Spin handbook for QIAquick PCR Purification, Nucleotide Removal og Gel Extraction Kit.
- QIAGEN (2005). QIAprep Miniprep Handbook- for purification of molecular grade DNA.
- Reece, R. J. (2004). Analysis of Genes and Genomes, Wiley.
- Saeki, K. and H. Kouchi (2000). "The Lotus Symbiont, Mesorhizobium loti: Molecular Genetic Techniques and Application." Journal of Plant Research **113**(4): 457-465.
- Sambrook, J. and D. W. Russel (2001). Molecular Cloning: A Laboratory Manual. New York, Cold Spring Harbor Laboratory Press.
- Shen, W., Wu, J, Wang, J, Ma, Y, Duan, Z og Gong, Yifeng (2006). "Properties of recombinant Sulfolobus shibatae Maltioligosyltrehalose Synthase Expressed in HMS174." Nature and Science **4**(2): 52-57.
- Somers, E., J. Vanderleyden, et al. (2004). "Rhizosphere Bacterial Signalling: A Love Parade Beneath Our Feet." Critical Reviews in Microbiology **30**(4): 205-240.
- Sprent, J. (2001). Nodulation in Legumes, Royal Botanic Gardens, Kew Publishing.

Townsend, G. E., II, Forsberg, Lennart S., Keating, David H. (2006). "Mesorhizobium loti Produces nodPQ-Dependent Sulfated Cell-Surface Polysaccharides." J. Bacteriol. **188**(24): 8560–8572.

Uchiumi, T., Ohwada, Takuji, Itakura, Manabu, Mitsui, Hisayuki, Nukui, Noriyuki, Dawadi, Pramod, Kaneko, Takakazu, Tabata, Satoshi, Yokoyama, Tadashi, Tejima, Kouhei, Saeki, Kazuhiko, Omori, Hirofumi, Hayashi, Makoto, Maekawa, Takaki, Sriprang, Rutchadaporn, Murooka, Yoshikatsu, Tajima, Shigeyuki, Simomura, Kenshiro, Nomura, Mika, Suzuki, Akihiro, Shimoda, Yoshikazu, Sioya, Kouki, Abe, Mikiko, Minamisawa, Kiwamu (2004). "Expression Islands Clustered on the Symbiosis Island of the Mesorhizobium loti Genome." J. Bacteriol. **186**(8): 2439-2448.

van Rhijn, P. and J. Vanderleyden (1995). "The Rhizobium-plant symbiosis." Microbiol. Rev. **59**(1): 124-142.

Waisel, Y., Eshel Amram , Kafkafi Uzi (2002). Plant Roots: The Hidden Half - Third Edition, Marcel Dekker

Whitehead NA, B. A., Slater H, Simpson NJ, Salmond GP. (2001). "Quorum-sensing in Gram-negative bacteria." FEMS Microbiol Rev. **25**(4): 365-404.

Appendiks I

Resultat etter dobbel oppstilling av DNA-sekvenser av genene *thuA* (*Sinorhizobium meliloti*) og *mll3105* (*Mesorhizobium loti* MAFF303099). Region med homologi er illustrert i blått, og homologe sekvenser er angitt med "I" mellom basene. Totalt er regionen 76 % identisk (437 av 570 baser er like). Basepar angis etter posisjon i de komplette sekvensene: *thuA* base lokalisasjon oppgis i "Query", og *mll3105* base lokalisasjon i "Sbjct".

```

Sequence 1: thuA
Length = 813 (1.. 813)

Sequence 2: mll3105
Length = 786 (1.. 786)

-----
|-----|
-----

Score = 329 bits (171), Expect = 1e-86
Identities = 437/570 (76%), Gaps = 0/570 (0%)
Strand=Plus/Minus

Query 118  ACGATCGCAGCCGCACCTCAATTCCGACCCGGGATCGATGCGACGACGGCGACGCTGCAG 177
          |||||  ||||  ||||  ||  ||  ||||  ||||  ||  ||  ||  ||||  ||||
Sbjct 696  ACGATCGCGCCGCGCTCAACCAGGACAAGGGCATCGAGGCGACCACCGCCACCTTGCAG 637

Query 178  GAGCCGGAACACGGGCTGAGCGAGACACGGCTCGCCGCCCGGACGTGCTGCTGTGGTGG 237
          ||  ||  ||  |||||  |||||  ||  ||  |||||  |||||  |||||  |||||  |||||
Sbjct 636  GAACCCGAGCACGGCCTGAGCGAGAAGCGCCTGGCCGCCACCGACGTGCTTTTGTGGTGG 577

Query 238  GGCCACAAGGATCATGGCGCCGTCGATGACGCCATCGTAGAGCGCGTCGCCAAACGGCGTT 297
          |||||  |  ||  ||||  ||  ||  ||||  |||||  |||||  ||  ||  ||
Sbjct 576  GGCCACGCCGCACACGGCGAGGTCAAGGACGAGATCGTCGAGCGCGTGCAGAAGCGGGTC 517

Query 298  TGGGAAGGCATGGGTCTAATCGTTCTCCATTCCGGCCATTTCTCCAAGGTGTCAAGCGG 357
          |||||  |||||  ||  ||||  ||  |||||  ||||  |  ||  ||  |  |||||  ||||
Sbjct 516  TGGGAAGGCATGGGGCTGATCGTGTGCAATCCGGCCACTACTCGAAGATCTTCAAGCGG 457

Query 358  CTGATGGGCACGCCCTGCGCGCTTAAATGGCGCGAGGGCGGAGAGCGGAAACGGCTCTGG 417
          |||||  |||||  ||||  ||  |||||  |||||  |||||  |||||  ||  |||||
Sbjct 456  CTGATGGGCACGCCCTGCTCGCTAAAGTGGCGCGAGGGCGGCGAGCGCGAGCGGGTCTGG 397

Query 418  GTCGTCAATCCGCGCCACCCGATCGCCGAAGGACTCGGTGAGAACTTCGTATCGAGAAC 477
          |  ||||  |  |||||  |||||  ||||  ||||  ||  |  |  ||||  |
Sbjct 396  GCGATCAACCGCGGCCACCCGATCGCGCAAGGCATCGGCGAATGCCTGGAGATCGGGGAG 337

Query 478  GAGGAGATGTATGGCGAGCAATTCTCGGTGCCGGAGCCGCTCGAGACGGTATTCATCTCC 537
          ||  |||||  |||||  ||  |||||  |||||  ||  |||||  |||||  ||  |||||
Sbjct 336  ACGGAAATGTATGGCGAGCCCTTCGCGGTGCCGGAGCCGATGGAGACGGTGTTCGTCTCC 277

Query 538  TGGTTTGCCGGCGGCGAAGTCTTCCGGTCCGGCCCTCACCTGGCGGCGCGGCGGGCAAT 597
          ||||  |  |||||  ||  |||||  ||||  ||  ||||  |  |||||  |||||
Sbjct 276  TGGTACGAGGGCGGCGAGGTGTTCCGCTCGGGGCTGACCTACCAGCGCGGCGGGCCGC 217

Query 598  ATCTTTACTTCCGCCCGGGCATGAGACTTACCCGACCTATCACGACGCGACCGTACAC 657
          ||||  ||  ||  ||||  ||  ||||  ||  ||||  |||||  |  |  ||  ||
Sbjct 216  ATCTTCTATTCTCGCCCGGCCACGAGACCTATCCGATCTATCACAAATGAGGGTGTGCAG 157

Query 658  AAGGTGCTGCGGAACCGGTAAGTGGGCG 687
          |||||  ||  ||  ||  |  |||||
Sbjct 156  CAGGTGCTGCGCAATGCCGTGCACTGGGCG 127

```


Appendiks III

Eksempel på resultat etter dobbel oppstilling av referansesekvens av *M. lotis* gen *mll3104* fra Genebank, og opprensede *mll3104*-fragmenter etter kloning i ekspresjonsvektor pETmll3104. Totalt er sekvensene 98 % homologe, med 20 ukjente baser. Sort "." angir homologi, og "N" angir ukjente baser. "Query" er sekvens 1 base lokalisasjon, "Sbjct" er sekvens 2 baselokalisasjon.

```

Sequence 1: mll3104
Length = 1049 (1 .. 1050)

Sequence 2: mll3104 sekv 190407, T7-F
Length = 1129 (1 .. 1130)
Score = 1883 bits (979), Expect = 0.0
Identities = 999/1019 (98%), Gaps = 0/1019 (0%)
Strand=Plus/Minus

Query 24      TGCTCGCCGAACTCTCGAAGGCAGCGTCGATCAGCTTCTGCATGTGCGCGGCACGCCGGA 83
Sbjct 1096     ..N.....NHN.N.....N.N.....N.N.....N.N.N.N.....N.N..... 1037

Query 84      AGGACGGATCGCCATTCTGCCCGGCATCCAGCGCGTCGACAAAACGCCGGGCGTTGCGCT 143
Sbjct 1036     ..NHN.....N.....N..... 977

Query 144     TGACATCCGGCAGGGCCAGCTCTGCCAGAGCTGAACATCGACGCTCTCCCAGGCACG 203
Sbjct 976      ..... 917

Query 204     CCGACAGCCCGCGCTGCCCTTGCCATCGGTTTCGACCTTGAGCGCGCCCTTGGTGCCATGCA 263
Sbjct 916     ..... 857

Query 264     GCGCCAGCGACAGGTCATTGCCATGGCCGGTTGTGTAGCGGCTCGCCATGATCGTCGCCA 323
Sbjct 856     ..... 797

Query 324     GCGCGCCGGATGCCAGCCGTGCCGTCATGCCACACTGTCATTGGCGTCCAGCACATAGT 383
Sbjct 796     ..... 737

Query 384     CGCCGATCCGGTTGCCCTTCGGCCTTCGGGAACGTCACCAGATCGGCGTGCAGGCCGACAA 443
Sbjct 736     ..... 677

Query 444     TATCCTGGGCGGCCATAGTGGCGAAGTCCAGGATATGGATGCCGACATCGCCCAGAA 503
Sbjct 676     ..... 617

Query 504     CCCCAGTCCGACCCGTGCCGGGTCCGACAAATCGCCACAGCCATTTGCTCCTCGACGCGCCAGT 563
Sbjct 616     ..... 557

Query 564     CACCCAGGACTTGTGTGACCAGCCAGCTCTGCTTGTAGCTCGCCTCGACATGGCGCAATT 623
Sbjct 556     ..... 497

Query 624     CGCCGATCGCTCCCGCCAACACCAGACGCCGTGCTTCCCTGGATCGCCGGTGAATTGCGGT 683
Sbjct 496     ..... 437

Query 684     AGGTGAGATTGACCATGTTGACGAGGCCGGCCCTCGGCGGCTCGGTCATGGCAAGCG 743
Sbjct 436     ..... 377

Query 744     CATCGGCATGGTTCCGGTGCCAGCGGCTTTTCGCAAAAAGACATGTTTTCCCGCAGCCAGCA 803
Sbjct 376     ..... 317

Query 804     GCGCCAGCGTCGTCGCCTTGTGCACTCCGTCGGAGTGGCGTTGATGGCGCGCTCGAAAT 863
Sbjct 316     ..... 257

Query 864     CGCCCCAGGCGATCGCCGCTTCAAGGCTCTCGAAGGAAGCGCCGATCTCGTTGGCAGCGG 923
Sbjct 256     ..... 197

Query 924     CAAACGCCCGCGCGGCCAGGCAGTTGATCGACACAGGCGACAAATGCTGCAGCCAGGAA 983
Sbjct 196     ..... 137

Query 984     CCTTTGCGAACTCTTCGATATGATGTCCGGCGATCCAGCCGGTGGCGAGCAAGAGCAGG 1042
Sbjct 136     ..... 78
    
```

Appendiks IV

Resultat etter dobbel sekvensoppstilling mellom referansesekvens for *mll3105* (sekvens 1) fra Genbank og *mll3105*-fragmenter (sekvens 2), begge fra *M. loti* MAFF303099. Sekvensering etter kloning i ekspresjonsvektor pETmll3104. Totalt er sekvensene 99 % homologe, med 4 ukorrekte baser. Sort "." angir homologi, "A" og "C" angir ukorrekte baseinnsett under sekvenseringsreaksjonen. "Query" er sekvens 1 base lokalisasjon, "Sbjct" er sekvens 2 baselokalisasjon.

```

Sequence 1: putative mll3105
Length = 786 (1 .. 786)

Sequence 2: mll3105-T7F-110407
Length = 1181 (1 .. 1181)
Score = 1471 bits (765), Expect = 0.0
Identities = 773/777 (99%), Gaps = 0/777 (0%)
Strand=Plus/Minus

Query 4   ACTCAACCCCTTGTCGCCATCGGCATGTAGGCGCAGGCCCTTCTGAACGATCTTTTCCTT 63
Sbjct 849  .....

Query 64  GGCCTTGTCGGTCGGCAGGTCGGCGCATTGGTGATGCCCGACCATGCGGGCGCGGGATT 123
Sbjct 789  ......A.... 730

Query 124  GAGCGCCCAGTGCACGGCATTGCGCAGCACCTGCTGCACACCCTCATTGTGATAGATCGG 183
Sbjct 729  .....C..... 670

Query 184  ATAGGTCTCGTGGCCGGGCGAGAAATAGAAGATGCGGCCCGCGCCGCTGGTAGGTCAG 243
Sbjct 669  .....

Query 244  CCCCAGCGGAACACCTCGCCGCCCTCGTACCAGGAGACGAACACCGTCTCCATCGGCTC 303
Sbjct 609  .....C..... 550

Query 304  CGGCACCGCGAAGGGCTCGCCATACATTTCCGTCTCGCCGATCTCCAGGCATTGCGCGAT 363
Sbjct 549  .....C..... 490

Query 364  GCCTTGCGGATCGGGTGGCCGCGGTTGATCGCCCAGACCCGCTCGCGCTCGCCCGCCTC 423
Sbjct 489  .....

Query 424  GCGCCACTTTAGCGAGCAGGGCGTGCCCATCAGCCGCTTGAAGATCTTCGAGTAGTGCC 483
Sbjct 429  .....

Query 484  GGAATGCAGCACGATCAGCCCCATGCCTTCCCAGACCCGCTTCTGCACGCGCTCGACGAT 543
Sbjct 369  .....

Query 544  CTCGTCCTTGACCTCGCCGTGTGCGGCGTGGCCCCACCACAAAAGCACGTCGGTGGCGGC 603
Sbjct 309  .....

Query 604  CAGGCGCTTCTCGCTCAGGCCGTGCTCGGGTTCCTGCAAGGTGGCGGTGGTGCCTCGAT 663
Sbjct 249  .....

Query 664  GCCCTTGTCCTGGTTGAGCGCGCCGCGATCGTACCATGCATGGTCAAGGGATAGAGGTC 723
Sbjct 189  .....

Query 724  GCGGACGGCGCATTTGGTCTGTTTCATGGACATTCTCGCCCCAGACAACAGCTTTTAT 780
Sbjct 129  ..... 73
    
```

Appendiks V

Resultat etter dobbel sekvensoppstilling mellom referansesekvens for *mll3105* (sekvens 1) fra Genbank og *mll3105*-fragmenter (sekvens 2), begge fra *M. loti* MAF303099. Sekvensering (3/5-07) etter kloning i ekspresjonsvektor pETmll3104. T7-forward primer er anvendt. Totalt er sekvensene 99 % homologe, med 4 feile baser. Sort "." angir homologi, "A" og "C" angir ukorrekte baseinnsett under sekvenseringsreaksjonen. "Query" er sekvens 1 base lokalisasjon, "Sbjct" er sekvens 2 baselokalisasjon.

```

Sequence 1: putative mll3105
Length = 786 (1 .. 786)

Sequence 2: mll3105-T7F, 030507
Length = 1132 (1 .. 1132)
Score = 1469 bits (764), Expect = 0.0
Identities = 772/776 (99%), Gaps = 0/776 (0%)
Strand=Plus/Minus

Query 4   ACTCAACCCTTGTGCGCCATCGGCATGTAGGCGCAGGCCCTTCTGAACGATCTTTTCCTT 63
Sbjct 851  .....

Query 64  GGCCTTGTCCGTCGGCAGGTCGGCGCATTGGTGATGCCCGACCATGCGGGCGCGGGATT 123
Sbjct 791  .....A.... 732

Query 124 GAGCGCCCAGTGCACGGCATTGCGCAGCACCTGCTGCACACCCTCATTGTGATAGATCGG 183
Sbjct 731  .....C..... 672

Query 184 ATAGGTCTCGTGGCCGGGCGAGAAATAGAAAGATGCGGCCCGCGCCGCGCTGGTAGGTCAG 243
Sbjct 671  .....

Query 244 CCCCAGCGGAACACCTCGCCGCCCTCGTACCAGGAGACGAACACCGTCTCCATCGGCTC 303
Sbjct 611  .....C..... 552

Query 304 CGGCACCGCGAAGGGCTCGCCATACATTTCCGTCTCGCCGATCTCCAGGCATTGCGCGAT 363
Sbjct 551  .....C..... 492

Query 364 GCCTTGCGCGATCGGGTGGCCGCGGTTGATCGCCCAGACCCGCTCGCGCTCGCCCGCCTC 423
Sbjct 491  .....

Query 424 GCGCCACTTTAGCGAGCAGGGCGTGCCCATCAGCCGCTTGAAGATCTTCGAGTAGTGGCC 483
Sbjct 431  .....

Query 484 GGAATGCAGCACGATCAGCCCCATGCCTTCCCAGACCCGCTTCTGCACGCGCTCGACGAT 543
Sbjct 371  .....

Query 544 CTCGTCCTTGACCTCGCCGTGTGCGGCGTGCCCCACCACAAAAGCACGTCGGTGGCGGC 603
Sbjct 311  .....

Query 604 CAGGCGCTTCTCGCTCAGGCCGTGCTCGGGTTCTGCAAGGTGGCGGTGGTCGCCTCGAT 663
Sbjct 251  .....

Query 664 GCCCTTGTCTGTTGAGCGCGGCCGCGATCGTACCATGCATGGTCAAGGGATAGAGGTC 723
Sbjct 191  .....

Query 724 GCGGACGGCGGCATTGGTCTGTTCATGGACATTCTCGCCCCAGACAACAGCTTTTA 779
Sbjct 131  ..... 76

```

Appendiks VI

Resultat etter dobbel sekvensoppstilling mellom referansesekvens for *mll3105* (sekvens 1) fra Genbank og *mll3105*-fragmenter (sekvens 2), begge fra *M. loti* MAFF303099. Sekvensering (19/9-07) etter kloning i ekspressjonsvektor pETmll3105. Formål for sekvensering er optimalisering av reaksjonen, tilsatte 100 ng/μl DNA mer enn ved tidligere sekvenseringsreaksjoner. Totalt er sekvensene 99 % homologe, med 4 ukorrekte baser. T7-forward primer er brukt. Sort "." angir homologi, "A" og "C" angir feil baseinnsett under sekvenseringsreaksjonen. "Query" er sekvens 1 base lokalisasjon, "Sbjct" er sekvens 2 baselokalisasjon.

```

Sequence 1: putative mll3105
Length = 786 (1..786)

Sequence 2: mll3105-T7F optimalisering(300 ng DNA) 190907
Length = 1160 (1..1160)
Score = 1483 bits (771), Expect = 0.0
Identities = 779/783 (99%), Gaps = 0/783 (0%)
Strand=Plus/Minus

Query 4   ACTCAACCCCTTGTCGCCATCGGCATGTAGGCGCAGGCCCTTCTGAACGATCTTTTCCTT 63
Sbjct 848  .....

Query 64  GGCCTTGTCCGTCGGCACGTTCCGGCGCATTGGTGATGCCCGACCATGCGGGCGCGGGATT 123
Sbjct 788  .....A.....

Query 124  GAGCGCCCAGTGCACGGCATTGCGCAGCACCTGCTGCACACCCTCATTGTGATAGATCGG 183
Sbjct 728  .....C.....

Query 184  ATAGGTCTCGTGGCCGGCGAGAAATAGAAGATGCGGCCCGCGCCGCTGGTAGGTCAG 243
Sbjct 668  .....

Query 244  CCCCAGCGGAACACCTCGCCGCCCTCGTACCAGGAGACGAACACCGTCTCCATCGGCTC 303
Sbjct 608  .....C.....

Query 304  CGGCACCGCGAAGGGCTCGCCATACATTTCCGTCTCGCCGATCTCCAGGCATTGCGCGAT 363
Sbjct 548  .....C.....

Query 364  GCCTTGC CGATCGGGTGGCCGCGTTGATCGCCAGACCCGCTCGCGCTCGCCCGCCTC 423
Sbjct 488  .....

Query 424  GCGCCACTTTAGCGAGCAGGGCGTGCCCATCAGCCGCTTGAAGATCTTCGAGTAGTGGCC 483
Sbjct 428  .....

Query 484  GGAATGCAGCACGATCAGCCCCATGCCTTCCCAGACCCGCTTCTGCACGCGCTCGACGAT 543
Sbjct 368  .....

Query 544  CTCGTCCTTGACCTCGCCGTGTGCGGCGTGGCCCCACCACAAAAGCACGTCGGTGGCGGC 603
Sbjct 308  .....

Query 604  CAGGCGCTTCTCGCTCAGGCCGTGCTCGGGTTCTTGCAAGGTGGCGGTGGTGCCTCGAT 663
Sbjct 248  .....

Query 664  GCCCTTGTCTGTTGAGCGCGGCCGCGATCGTACCATGCATGGTCAAGGGATAGAGGTC 723
Sbjct 188  .....

Query 724  GCGGACGGCGGCATTGGTCTGTTCATGGACATTCTCGCCCCAGACAACAGCTTTTATCGG 783
Sbjct 128  .....

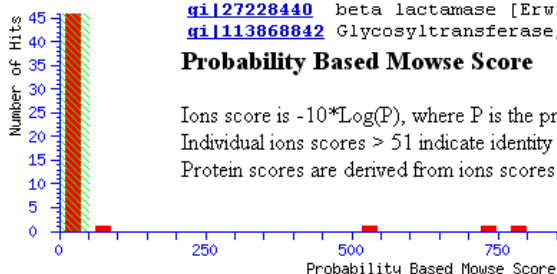
Query 784  CAT 786
Sbjct 68  ... 66
    
```

Appendiks VII

Eksempel på ett av tre resultater etter Q-TOF analyse for innsendte M113105-fraksjonsprøver. Innsendt materiale ble kuttet fra SDS-PAGE gel. "Protein hits" angir hvilke 4 treff man fikk etter søk i Mascot (etter proteiner med høyest mulig masselighet som i innsendt prøve "m113105 a"). Diagram "Probability Based Mowse Score" antyder sannsynlighet for proteinmatch, høyere verdi gir større sannsynlighet for eksakt treff under søk. Proteiner angis kronologisk med høyest sannsynlighet for eksakt treff under søk først, med synkende rekkefølge for sannsynlige match.

{MATRIX} Mascot Search Results

User : toril anne grønset
 Email : Toril-Anne.Gronset@farmasi.uit.no
 Search title : 3105 protein a
 MS data file : E:\071025 3105a.pkl
 Database : NCBIInr 20071019 (5565793 sequences; 1926068711 residues)
 Timestamp : 29 Oct 2007 at 09:10:34 GMT
 Protein hits : [gi|208244](#) b-lactamase
 [gi|2765071](#) ES-beta-lactamase [Escherichia coli]
 [gi|27228440](#) beta lactamase [Erwinia amylovora]
 [gi|113868842](#) Glycosyltransferase, probably involved in lipopolysaccharide biosynthesis



Probability Based Mowse Score

Ions score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event.
 Individual ions scores > 51 indicate identity or extensive homology ($p < 0.05$).
 Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

Peptide Summary Report

Format As: Peptide Summary [Help](#)

Significance threshold p < 0.05 Max. number of hits AUTO

Standard scoring MudPIT scoring Ions score or expect cut-off 0 Show sub-sets 0

Show pop-ups Suppress pop-ups Sort unassigned Decreasing Score Require bold red

Select All Select None Search Selected Error tolerant

1. [gi|208244](#) Mass: 31648 Score: 784 Queries matched: 14
 b-lactamase

Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank
<input checked="" type="checkbox"/> 2	494.7338	987.4530	987.4558	-2.78	0	35	2.3	1
<input checked="" type="checkbox"/> 4	536.7736	1071.5326	1071.5309	1.65	0	77	0.00018	1
<input checked="" type="checkbox"/> 5	544.3182	1086.6218	1086.6219	-0.09	0	61	0.0069	1
<input checked="" type="checkbox"/> 7	571.8390	1141.6634	1141.6706	-6.30	0	55	0.022	1
<input checked="" type="checkbox"/> 8	425.8880	1274.6422	1274.6619	-15.45	0	51	0.068	1
9	638.3394	1274.6642	1274.6619	1.87	0	(24)	31	3
<input checked="" type="checkbox"/> 10	638.3415	1274.6684	1274.6659	1.99	0	66	0.0023	1
<input checked="" type="checkbox"/> 11	643.8856	1285.7566	1285.7605	-3.02	0	85	1.8e-05	1
<input checked="" type="checkbox"/> 12	690.3396	1378.6646	1378.6585	4.49	0	58	0.012	1
<input checked="" type="checkbox"/> 13	856.4117	1710.8088	1710.7849	14.0	0	51	0.047	1
<input checked="" type="checkbox"/> 14	892.4520	1782.8894	1782.8822	4.06	0	138	1e-10	1
<input checked="" type="checkbox"/> 16	900.4526	1798.8906	1798.8771	7.52	0	(95)	2e-06	1
<input checked="" type="checkbox"/> 17	674.6646	2020.9720	2020.9742	-1.09	0	75	0.00024	1
<input checked="" type="checkbox"/> 19	888.1176	2661.3310	2661.3353	-1.64	0	33	1.9	1

Fortsettelse appendiks VII

Proteins matching the same set of peptides:

[qi|509142](#) Mass: 39814 Score: 784 Queries matched: 14
 beta-lactamase, human IgG3 hinge fusion [synthetic construct]
[qi|546975](#) Mass: 29031 Score: 784 Queries matched: 14
 TEM-2 beta-lactamase {Val-237 for Ala substitution} [Escherichia coli]
[qi|4557168](#) Mass: 31623 Score: 784 Queries matched: 14
 inhibitor resistant beta lactamase TEM-54 [Escherichia coli]
[qi|6176542](#) Mass: 31567 Score: 784 Queries matched: 14
 inhibitor-resistant beta-lactamase TEM-79 [Escherichia coli]
[qi|7766856](#) Mass: 28987 Score: 784 Queries matched: 14
 Chain A, S235a Mutant Of Tem1 Beta-Lactamase
[qi|8569629](#) Mass: 28999 Score: 784 Queries matched: 14
 Chain A, X-Ray Crystal Structure Of Tem-1 Beta Lactamase In Complex
[qi|8569630](#) Mass: 29003 Score: 784 Queries matched: 14
 Chain A, X-Ray Crystal Structure Of Tem-1 Beta Lactamase In Complex
[qi|10955180](#) Mass: 31666 Score: 784 Queries matched: 14
 beta-lactamase [Neisseria gonorrhoeae]
[qi|12018137](#) Mass: 31594 Score: 784 Queries matched: 14
 TEM-derived extended spectrum beta-lactamase [Escherichia coli]
[qi|13562035](#) Mass: 31608 Score: 784 Queries matched: 14
 TEM-90 beta-lactamase [Escherichia coli]
[qi|15149326](#) Mass: 31585 Score: 784 Queries matched: 14
 inhibitor resistant TEM beta-lactamase [Escherichia coli]
[qi|15149330](#) Mass: 31654 Score: 784 Queries matched: 14
 TEM beta-lactamase [Escherichia coli]
[qi|15624079](#) Mass: 31695 Score: 784 Queries matched: 14
 extended-spectrum beta-lactamase TEM-71 [Klebsiella pneumoniae]
[qi|16798376](#) Mass: 31667 Score: 784 Queries matched: 14
 inhibitor-resistant beta-lactamase TEM-84 [Escherichia coli]
[qi|20530259](#) Mass: 31608 Score: 784 Queries matched: 14
 beta-lactamase TEM-96 [Escherichia coli]
[qi|21465759](#) Mass: 29017 Score: 784 Queries matched: 14
 Chain A, Crystal Structure Of G238a Mutant Of Tem-1 Beta-Lactamase
[qi|24158842](#) Mass: 28934 Score: 784 Queries matched: 14
 Chain A, Crystal Structure Of Tem-30 Beta-Lactamase At 2.0 Angstrom
[qi|25900617](#) Mass: 31597 Score: 784 Queries matched: 14
 beta-lactamase [Escherichia coli]
[qi|27374443](#) Mass: 31670 Score: 784 Queries matched: 14
 beta-lactamase [Proteus mirabilis]
[qi|29650764](#) Mass: 31730 Score: 784 Queries matched: 14
 beta-lactamase TEM-120 [Klebsiella oxytoca]
[qi|37575400](#) Mass: 60561 Score: 784 Queries matched: 14
 Rdfusion [Ribosome display vector pRDV]
[qi|40737710](#) Mass: 31643 Score: 784 Queries matched: 14
 beta-lactamase TEM-127 [Escherichia coli]
[qi|40737712](#) Mass: 31680 Score: 784 Queries matched: 14
 beta-lactamase TEM-128 [Escherichia coli]
[qi|46370393](#) Mass: 31715 Score: 784 Queries matched: 14
 beta-lactamase [Escherichia coli]
[qi|46370393](#) Mass: 31715 Score: 784 Queries matched: 14
 beta-lactamase [Escherichia coli]
[qi|61967136](#) Mass: 31667 Score: 784 Queries matched: 14
 beta-lactamase TEM-141 [Enterobacter cloacae]
[qi|81365635](#) Mass: 31665 Score: 784 Queries matched: 14
 TEM-145 [Escherichia coli]
[qi|81365640](#) Mass: 31643 Score: 784 Queries matched: 14
 TEM-146 [Escherichia coli]
[qi|83031783](#) Mass: 31738 Score: 784 Queries matched: 14
 beta-lactamase TEM-55 [Escherichia coli]
[qi|89112911](#) Mass: 31648 Score: 784 Queries matched: 14
 TEM-1 beta lactamase gene [Neisseria gonorrhoeae]
[qi|112785108](#) Mass: 31636 Score: 784 Queries matched: 14
 TEM extended-spectrum beta-lactamase [Escherichia coli]
[qi|114881106](#) Mass: 31666 Score: 784 Queries matched: 14
 Beta Lactamase [Pseudomonas aeruginosa]

Fortsettelse appendiks VII

2. [qi|27265071](#) Mass: 31729 Score: 729 Queries matched: 14

ES-beta-lactamase [Escherichia coli]

 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
2	494.7338	987.4530	987.4558	-2.78	0	35	2.3	1	R.FPMMSTFK.V
4	536.7736	1071.5326	1071.5309	1.65	0	77	0.00018	1	R.VDAGQEQLGR.R
5	544.3182	1086.6218	1086.6219	-0.09	0	61	0.0069	1	K.VLLCGAVLSR.V
7	571.8390	1141.6634	1141.6706	-6.30	0	55	0.022	1	R.QIAEIGASLIK.H
8	425.8880	1274.6422	1274.6619	-15.45	0	51	0.068	1	K.ILESFRPEER.F
9	638.3394	1274.6642	1274.6619	1.87	0	(24)	31	3	K.ILESFRPEER.F
10	638.3415	1274.6684	1274.6659	1.99	0	66	0.0023	1	R.SALPAGWFIADK.S
11	643.8856	1285.7566	1285.7605	-3.02	0	85	1.8e-05	1	K.LLTGELLTLASR.Q
12	690.3396	1378.6646	1378.6585	4.49	0	58	0.012	1	R.DTTPAAMATTLR..
13	856.4117	1710.8088	1710.7849	14.0	0	51	0.047	1	R.WEPELNEAIPNDER
14	892.4520	1782.8894	1782.8644	14.0	0	83	3.1e-05	2	R.IVVIYMTGGQATMD
16	900.4526	1798.8906	1798.8594	17.4	0	(57)	0.015	2	R.IVVIYMTGGQATMD
17	674.6646	2020.9720	2020.9742	-1.09	0	75	0.00024	1	R.IHYSQNDLVEYSV'
19	888.1176	2661.3310	2661.3353	-1.64	0	33	1.9	1	R.ELCSAAITMSDNTA

3. [qi|27228440](#) Mass: 31263 Score: 555 Queries matched: 11

beta lactamase [Erwinia amylovora]

 Check to include this hit in error tolerant search

9	638.3394	1274.6642	1274.6619	1.87	0	(24)	31	3	K.ILESFRPEER.F
10	638.3415	1274.6684	1274.6659	1.99	0	66	0.0023	1	R.SALPAGWFIADK.S
11	643.8856	1285.7566	1285.7605	-3.02	0	85	1.8e-05	1	K.LLTGELLTLASR.Q
12	690.3396	1378.6646	1378.6585	4.48	0	23	40	2	R.DTGMFPVAMATTLR.K +
13	856.4117	1710.8088	1710.7849	14.0	0	51	0.047	1	R.WEPELNEAIPNDER.D
17	674.6646	2020.9720	2020.9742	-1.09	0	75	0.00024	1	R.IHYSQNDLVEYSVTER
19	888.1176	2661.3310	2661.3353	-1.64	0	33	1.9	1	R.ELCSAAITMSDNTAANI

4. [qi|113868842](#) Score: 53 Queries matched: 1

Glycosyltransferase, probably involved in lipopolysaccharide biosynthesis [Ralstonia eutropha]

 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
11	643.8856	1285.7566	1285.7904	-26.24	1	53	0.028	2	-.MRILLTTGLR.L

Peptide matches not assigned to protein hits: (no details means no match)

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 9	638.3394	1274.6642	1274.6765	-9.60	1	36	2	1	RMGATNVLIER + Oxide
<input checked="" type="checkbox"/> 3	525.2502	1048.4858	1048.4309	52.4	0	29	9.6	1	ESSEPDTER
<input checked="" type="checkbox"/> 15	897.9185	1793.8224	1793.8229	-0.26	0	26	16	1	DPSMFTIPCWIGNTR
<input checked="" type="checkbox"/> 6	563.3333	1124.6520	1124.6593	-6.49	1	24	23	1	FYNLKLSLK
<input checked="" type="checkbox"/> 1	412.7549	823.4952	823.4916	4.48	0	18	72	1	APAISLPR
<input checked="" type="checkbox"/> 18	726.6973	2177.0701	2177.0350	16.1	0	10	5.1e+02	1	TGSFAFYAPASSAIEMAEFLK
<input checked="" type="checkbox"/> 20	1024.8198	3071.4376	3071.2316	67.1	1	3	1.6e+03	1	SEGGSSMSEGGSSQERSGVGSF

Search Parameters

Type of search : MS/MS Ion Search
 Enzyme : Trypsin
 Fixed modifications : Carbamidomethyl (C)
 Variable modifications : Oxidation (M)
 Mass values : Monoisotopic
 Protein Mass : Unrestricted
 Peptide Mass Tolerance : ± 100 ppm
 Fragment Mass Tolerance : ± 0.1 Da
 Max Missed Cleavages : 1
 Instrument type : ESI-QUAD-TOF
 Number of queries : 20

Appendiks VIII

Viser basepar (1-1035 angitt til venstre), restriksjonssteder (markert over sekvens i rødt og blått) og aminosyrer for hver basetriplett (grønne bokstaver under sekvens, fra bp 41 til bp 923) av *mll3105*-sekvens innsatt i vektor pETmll3105. Sekvensen er hentet fra en optimalisert sekvenseringsreaksjon, og limt inn i analyseprogrammet (Web Map References). Programmet er tilgjengelig fra http://pga.mgh.harvard.edu/web_apps/web_map/start. Hele sekvensen ble funnet å tilhøre en enkelt ORF.

```

XmnI
EcoRI
SpDon
1 TTTTGTAACTTTAAGAAGGAATTCAGGAGCCCTTCACCATGCGGATAAAAAGCTGTTGCTGGGGCGAGAATGTCCATGAACAGACCAATGCCGCCGTC
-----
100 AAAACAAATTGAAATCTTCTTAAGTCTCGGGAAGTGTACGGCTATTTTCGACAAAGACCCCGCTTACAGGTACTTGTCTGGTTACGGCGGCAG
> M P I K A V V W G E N V H E Q T N A A V

NsiI
NotI
EagI
EfrBI
PvuI
101 CGCGACCTCTATCCCTTGACCATGCATGGTACGATCGCGGCCGCGCTCAACCAGGACAAGGGCATCGAGGCGACCCGCCACCTTGACAGAACCCGAGC
-----
200 GCGCTGGAGATAGGAACTGGTACGTACCATGCTAGCGCCGGCGGAGTTGGTCTGTTCCCGTAGCTCCGCTGGTGGCGGTGGAACGCTCCTTGGGCTCG
> R D L Y P L T H H G T I A A A L N Q D K G I E A T T A T L Q E P E H

BpuI0
BmgBI
AleI
DrdI
201 ACGGCCGTGAGCGAGAAGCGCTGGCCGCCACCGACGTGCTTTTGTGTGGGGCCACGCGCACACGGCGAGGTCAAGGACGAGATCGTCGAGCGCGTGA
-----
300 TGCCGGACTCGCTCTTCGCGGACCGCGGTGGCTGCACGAAAACACCACCCCGGTGCGCGGTGTGCGCTCCAGTTCCTGCTCTAGCAGCTCGCGCACGT
> G L S E K R L A A T D V L L W W G H A A H G E V K D E I V E R V Q

SpAcc
BsqI
BsmI
XmnI
Bgl2
BstAP
301 GAAGCGGGTCTGGGAAGGCATGGGGCTGATCGTGTGCATTCCGGCCACTACTCGAAGATCTTCAAGCGGCTGATGGGACCGCCCTGCTCGCTAAAGTGG
-----
400 CTTCCGCCAGACCCCTTCGGTACCCCGACTAGCAGACGTAAGGCCGGTGTAGGCTTCTAGAAGTTCGCCGACTACCCGTGGCGGACGAGCGGATTCACC
> K R V W E G M G L I V L H S G H Y S K I F K R L M G T P C S L K W

BsrBI
Sac2
PvuI
BsmI BsmBI
401 CGCGAGGGCGGGGAGCGGAGCGGTCTGGCGGATCAACCGCGGCCACCCGATCGCGCAAGGCATCGCGCAATGCCTGGGGATCGGCGAGACGAAATGT
-----
500 GCGCTCCGCCGCTCGGCTCGCCACAGCCCGTCTAGTTGGCGCGGTGGCTAGCGCGTTCGCTAGCCGCTTACGACCCCTAGCCGCTCGCTTTTACA
> R E A G E R E R V W A I N R G H P I A Q G I G E C L G I G E T E M Y

BsmBI
BsaXa
BsmBI
BsrBI
501 ATGGCGAGCCCTTCGCGGTGCGGAGCCGATGGAGACGGTGTCTGCTCCTGGTGGCGAGGGCGGCGAGGTGTTCGCTCGGGGCTGACCTACCAGCGCGG
-----
600 TACCGCTCGGAAGCGCCACGGCTCGGCTACCTCTGCCACAAGCAGAGGACCACGCTCCCGCGCTCCACAGGGCAGCCCGACTGGATGGTCCGCGCC
> G E P F A V P E P M E T V F V S W C E G G E V F R S G L T Y Q R G

BspMI
BstAP
BsaI
BfuAI
AarI
BstAP
AlwNI
ApaI
Btg2I
601 CGCGGGCCGCATCTTCTATTTCTCGCCCGGCCACGAGACCTATCCGATCTATCACAAATGAGGGTGTGCAGCAGGTGCTCGCGCATGCCGTGCCTGGGCG
-----
700 GCGCCCGCGTAGAAGATAAAGAGCGGGCGGTGCTCTGGATAGGCTAGATAGTGTACTCCACACGCTCGTCCACGACGGCTACGGCACGTGACCCGCA
> A G R I F Y F S P G H E T Y P I Y H N E G V Q Q V L R D A V H W A

BmrI
701 CTCATCTCGCGCCCGCATGGTCCGGCATCAACATGCGCCGAACGTCGCCGACGGACAAGGCCAAGGAAAAGATCGTTACAGAGGGCCTGCGCCTACATG
-----
800 GAGTTAGCGCGGGCGTACCAGCCGTAAGTGGTTACCGGCTTGCACGGCTGCCTGTTCCGGTTCCTTTTCTAGCAAGTCTTCCCGGACCGGGATGTAC
> L N L A P A W S G I T N A P N V P T D K A K E K I V Q K G L R L H A

Hind3
SacIBstBI
BseRI
BsaI
MluI
AgeI
801 CCGATGGCGACAAGGGGTTGAGTGCAAAAGGGGAGCTCAATTGCAAGCTTGAANGTAAGCCATCCCTAACCCCTCTCCTCGGTCTCGATTCTACCGGTAC
-----
900 GGCTACCGCTGTTCCCAACTCAGTTCCTCCGCTCGAGTTAAGCTTTCGAACCTTTCGGATAGGATAGGGATGGGAGAGGAGCCAGACTAAGATGGCGATG
> D G D K G L S A K G E L N S K L E X K P I P N P L L G L D S T R T

CGGTATCATCACCATCACCATTTGATTTGATCCGGCTGCTNACANNCCGAAAAGGAAGCTGANTNGGCTGTCGCCNNTGAGCANANTAGCANMANCC
901 GCCAGTAGTAGTGGTAGTGAACCTAACTAGGCCGACGANTGTNNGGGCTTTCCTTCGACTNANCCGACGCGGNNACTCGTNTNATCGTNNNTNG
-----
1000 > G H H H H H H *

CNNNGGNNNTAANGNNNNNNNGNNNTTGNINA
1001 ----- 1035
GNNNCCNNNATTCCNNNNNNNCCNNAAACNANT
    
```