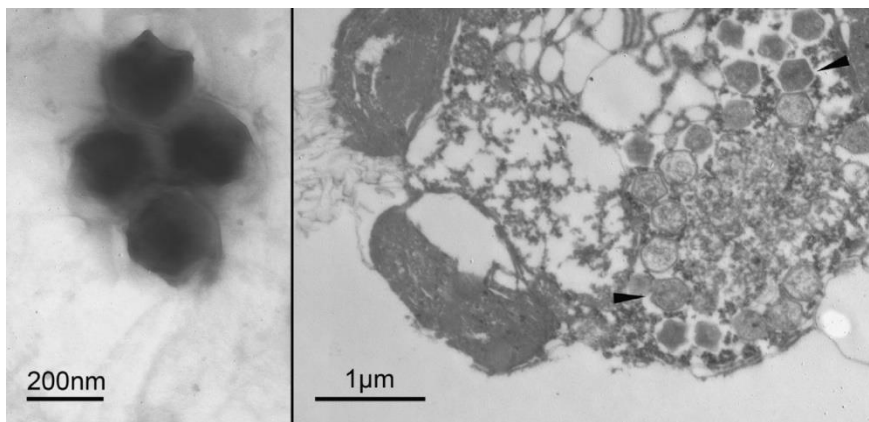


# Karakterisering av proteiner fra *Prymnesium* kappa virus (PkV RF01) et potensial for bioprospektering



Ida Marie Lervåg

Masteroppgave

Institutt for biovitenskap

Mikrobiologi

Vår 2019



UNIVERSITETET I BERGEN  
*Det matematisk-naturvitenskapelige fakultet*

## Sammendrag

Prymnesium kappa virus (PkV RF01) ble først isolert i Raunefjorden i tidsrommet 2009-2011 og er det største algeviruset så langt isolert (kapsid størrelse 310 nm). Viruset kan lysere flere stammer av *Haptolina ericina* og en stamme av *Prymnesium kappa*, noe som er uvanlig da kultiverte virus ofte er artsspesifikke. I dette viruset har det nylig blitt oppdaget flere gener som ikke tidligere har blitt beskrevet i virus. Dette inkluderer blant annet alle genene involvert i succinat dehydrogenase kompleks II, som er en del av trikarboksylsyresyklus og den oksidative fosforyleringen. I tillegg har det blitt oppdaget flere interessante auksiliære metabolske gener, hvor flere ikke er beskrevet i virus tidligere, og disse kan ha potensiale innen bioprospektering. PkV RF01 er dermed et virus av stor interesse for forskning. Målet for dette masterprosjektet er derfor å se nærmere på sju utvalgte gener for å undersøke om disse blir uttrykt i viruset, som er det første steget i å finne funksjonen deres. Arbeid med *sdhA* genet ble gjennomført for å finne ut om det blir transkribert under infeksjon, og eventuelt når i infeksjonssyklusen dette skjer. For å undersøke dette ble det isolert mRNA, som ble brukt som templat under RT-PCR, og deretter blir DNA amplifisert ved PCR for å undersøke om genet har blitt transkribert. Dette ble ikke vellykket da DNA ikke ble fjernet under DNase I behandling ved mRNA isoleringen.

Videre ble det forsøkt å klonere, uttrykke og rense fem auksiliære metabolske gener for å videre kunne karakterisere de og se på muligheten for økonomisk verdi. Dette ble gjort ved hjelp av pET ekspressjonssystem hvor *E. coli* ble brukt som vertsorganisme. DNA polymerase X ble vellykket renset, og sendt for videre karakterisering. De fire andre proteinene ble ikke vellykket renset, som kan tyde på at de ikke er løselige eller bare delvis løselig.

## Summary

*Prymnesium kappa* virus (PkV RF01) was first isolated in Raunefjorden during the years 2009-2011, being the largest algae virus yet to be isolated (capsid size 310 nm). This virus can lyse several strains of *Haptolina ericina* and one strain of *Prymnesium kappa*, which is rather unusual due to cultured viruses often being species specific. In this virus, several genes never previously described in viruses have recently been discovered. Including amongst others, all genes involved in succinate dehydrogenase complex II, which is part of TCA cycle and the oxidative phosphorylation. In addition, several interesting auxiliary metabolic genes that may have potential in bioprospecting have also been discovered. Thus, PkV RF101 is a virus of great interest when it comes to research. The aim of this master thesis is therefore to take a closer look at seven selected genes to investigate whether these genes are expressed in the virus, being the first step in deciding their function in the virus. The work with the *sdhA* gene was conducted to find out if the gene is transcribed during infection, and if possible when in the infection cycle it will be transcribed. mRNA was isolated and used as a template during RT-PCR, then DNA was amplified by PCR to find out if the gene was transcribed. This was not a success due to challenges removing DNA with DNase I treatment during mRNA isolation.

Five auxiliary metabolic genes were attempted cloned, expressed and purified, in order to characterize these and view further possibilities for economic value. This was done with the help of pET expression system where *E. coli* was used as the host organism. DNA polymerase X was successfully purified and shipped off for further characterization. The remaining four proteins were not successfully purified, which may indicate these proteins not being soluble or only partially soluble.

## Forord

Dette prosjektet har vært en bratt læringskurve med mye spennende laboratoriearbeid. Først vil jeg takke mine veiledere Ruth-Anne Sandaa og Ida Helene Steen for muligheten til å jobbe med en spennende oppgave som jeg har lært utrolig mye av. Tusen takk for gode tilbakemeldinger og all hjelp under arbeidet med oppgaven. Jeg vil også takke Hilde Marie Kristiansen Stabell og Hasan Arsin for den kjempegode opplæringen og all hjelp på laben. Takk skal også rettes til Elzbieta Petelencz-Kurdziel for opplæring på flowcytometer, og Katrine Sandnes Skaar for opplæring med ddPCR. Samtidig vil jeg takke Håkon Dahle for all hjelp med arbeid rundt fylogenetiske analysene.

Takk til familie og venner som har støttet meg under hele perioden, og spesielt en stor takk til samboeren min, som har vært en stor støtte disse to årene.

## Ordliste

<b>NCLDV</b>	Nukleocytoplasmiske store DNA-virus
<b>L-LDH</b>	L-laktat dehydrogenase
<b>PDH</b>	Prolin dehydrogenase
<b>GCDH</b>	Glutaryl-CoA dehydrogenase
<b>Pol X</b>	DNA polymerase X
<b>GDH</b>	Glutamat dehydrogenase
<b>SDHA</b>	Succinat dehydrogenase subenhet A
<b>SDHB</b>	Succinat dehydrogenase subenhet B
<b>SDHC</b>	Succinat dehydrogenase subenhet C
<b>SDHD</b>	Succinat dehydrogenase subenhet D
<b>MCP</b>	Stort kapsid protein
<b>PkV RF01</b>	Prynmesium kappa virus
<b>HeV RF02</b>	Haptolina ericina virus
<b>AaV</b>	Aureococcus anophagefferens
<b>CeV 01B</b>	Chrysochromulina ericina virus
<b>TetV1</b>	Tetraselmis virus
<b>PoV 01B</b>	Pyramimonas orientalis virus

# Innholdsfortegnelse

<b>1</b>	<b>INTRODUKSJON.....</b>	<b>1</b>
1.1	VIRUS I DET MARINE MILJØET.....	1
1.2	ALGEVIRUS .....	1
1.2.1	<i>Mimiviridae</i> .....	3
1.3	AUKSILIÆRE METABOLSKE GENER .....	5
1.4	VIRALE AUKSILIÆRE METABOLSKE GENER MED POTENSIALE FOR BIOPROSPEKTERING ...	6
1.4.1	<i>L-laktat dehydrogenase (L-LDH)</i> .....	7
1.4.2	<i>Prolin dehydrogenase (PDH)</i> .....	8
1.4.3	<i>Glutaryl-CoA dehydrogenase (GCDH)</i> .....	8
1.4.4	<i>DNA polymerase X (Pol X)</i> .....	8
1.4.5	<i>Glutamat dehydrogenase (GDH)</i> .....	9
1.4.6	<i>Succinat dehydrogenase kompleks II</i> .....	9
1.5	MARIN BIOPROSPEKTERING.....	12
1.6	METODER BENYTTET I PROTEINARBEID OG KVANTIFISERING AV GENPRODUKT.....	13
1.6.1	<i>Replikering av virus for gruppen NCLDV</i> .....	13
1.6.2	<i>Droplet digital PCR (ddPCR)</i> .....	15
1.7	MÅL FOR DENNE MASTEROPPGAVEN .....	16
<b>2</b>	<b>MATERIALER OG METODER .....</b>	<b>17</b>
2.1	DYRKING AV PRYMNESIUM KAPPA VIRUS (PkV RF01).....	18
2.1.1	<i>Dyrking av verten, haptofyten Haptolina ericina</i> .....	18
2.1.2	<i>Dyrking av algeviruset PkV RF01</i> .....	18
2.1.3	<i>Infeksjonsforsøk PkV RF01</i> .....	18
2.1.4	<i>Flowcytometri (FCM)</i> .....	19
2.2	DNA-ISOLERING .....	19
2.2.1	<i>PkV RF01</i> .....	19
2.2.2	<i>He028</i> .....	20
2.3	PCR-AMPLIFISERING AV SUCCINATE DEHYDROGENASE A (SDHA).....	20
2.4	DIGITAL DROPLET PCR (DDPCR) AV <i>SDHA</i> GENET.....	22
2.5	TRANSKRIPSJONSFORSØK – <i>SDHA</i> .....	24
2.5.1	<i>Vasking av celler og isolering av RNA</i> .....	24

2.5.2	<i>Revers transkripsjon (RT) PCR</i> .....	24
2.5.3	<i>Verifisering av transkripsjon og nedbrytingen av DNA</i> .....	25
2.6	<b>KLONING AV VIRUSGENER TIL PKV RF01</b> .....	26
2.6.1	<i>Bakteriestammer og plasmid</i> .....	26
2.6.2	<i>Primere</i> .....	27
2.6.3	<i>Amplifisering av målgener og vektor</i> .....	30
2.6.4	<i>Overføring av målgener til pET-21a vektor</i> .....	32
2.7	<b>TRANSFORMASJON NEB 5-ALPHA KOMPETENTE CELLER</b> .....	33
2.8	<b>KOLONI PCR</b> .....	33
2.9	<b>TRANSFORMASJON BL21-GOLD(DE3) CELLER</b> .....	34
2.10	<b>PROTEINRENSING - SMÅSKALA</b> .....	35
2.10.1	<i>Ekspresjon</i> .....	35
2.10.2	<i>SDS-PAGE av proteiner</i> .....	37
2.10.3	<i>Rensing av protein</i> .....	37
2.11	<b>RENSING AV PROTEIN - STORSKALA</b> .....	38
2.12	<b>BIOINFORMATISKE ANALYSER</b> .....	40
<b>3</b>	<b>RESULTAT</b> .....	<b>42</b>
3.1	<b>INFEKSJONSSYKLUS - PKV RF01</b> .....	42
3.2	<b>AMPLIFISERING AV SDHA GENET</b> .....	43
3.3	<b>TRANSKRIPSJONSFORSK</b> .....	44
3.4	<b>DDPCR - SDHA</b> .....	46
3.5	<b>KLONING AV MÅLGENER OG VEKTORER</b> .....	47
3.5.1	<i>Amplifisering av målgener og vektor</i> .....	47
3.5.2	<i>Kloning av målgener i pET-21a</i> .....	48
3.6	<b>EKSPRESJON AV MÅLGENER I E. COLI</b> .....	50
3.7	<b>RENSING AV PROTEINER</b> .....	53
3.7.1	<i>L-laktat dehydrogenase og glutaryl-CoA dehydrogenase</i> .....	53
3.7.2	<i>DNA polymerase X</i> .....	54
3.8	<b>BIOINFORMATISKE ANALYSER</b> .....	57
3.8.1	<i>Succinat dehydrogenase A</i> .....	57
3.8.2	<i>DNA Polymerase X</i> .....	58
3.8.3	<i>Domene figurer for de sju proteinene</i> .....	59
<b>4</b>	<b>DISKUSJON</b> .....	<b>60</b>

4.1	ER SDHA OG SDHB AKTIVE PROTEINER I PkV RF01?.....	60
4.1.1	<i>Transkripsjonsforsøk</i> .....	60
4.1.2	<i>Ekspresjon av SDHA og SDHB</i> .....	61
4.2	NÅR I INFEKSJONSSYKLUSEN TIL PkV RF01 BLIR SDHA EVENTUELT UTTRYKT?.....	62
4.3	HVA ER DEN EVOLUSJONISTISKE OPPRINNELSEN TIL SDHA I PkV RF01? .....	63
4.4	ER DET MULIG Å UTTRYKKE UTVALGTE PROTEINER FRA PkV RF01 I <i>E. COLI</i> ?.....	63
4.5	ER DE UTTRYKTE VIRUS PROTEINENE LØSELIGE?.....	67
4.6	HVA ER DEN EVOLUSJONISTISKE OPPRINNELSEN TIL DNA POLYMERASE X I PkV RF01? 70	
<b>5</b>	<b>KONKLUSJON OG VIDERE ARBEID</b> .....	<b>71</b>
<b>6</b>	<b>REFERANSER</b> .....	<b>72</b>
<b>7</b>	<b>VEDLEGG A: RESULTAT</b> .....	<b>87</b>
A.1	TRANSKRIPSJONSFORSØK .....	87
A.2	RENSING AV GLUTARYL-CoA DEHYDROGENASE.....	88
A.3	RENSING AV DNA POLYMERASE X – SMÅSKALA.....	89
<b>8</b>	<b>VEDLEGG B: LØSNINGER</b> .....	<b>90</b>
B.1	IMR/2 MEDIUM I 1L SJØVANN .....	90
B.2	SM BUFFER .....	91
B.3	0,5 M EDTA.....	91
B.4	LA MEDIUM MED AMPICILLIN (100 MG/ML) OG LB MEDIUM.....	91
<b>9</b>	<b>VEDLEGG C: PROTOKOLLER OG METODER</b> .....	<b>92</b>
C.1	INSTRUMENTINNSTILLINGER OG FORMEL – FLOWCYTOMETER .....	92
C.2	ZYMO DNA CLEAN AND CONCENTRATOR™-10 KIT .....	93
C.3	QUBIT 2.0 FLUOROMETER.....	93
C.4	VERIFISERING AV TRANSKRIPSJON OG NEDBRYTING AV DNA .....	94
C.5	BLOOD AND CELL CULTURE DNA MINI KIT (QIAGEN).....	94
C.6	MONARCH PLASMID MINIPREP KIT .....	95
<b>10</b>	<b>VEDLEGG D: SEKVENSER</b> .....	<b>96</b>
D.1	L-LAKTAT DEHYDROGENASE (L-LDH).....	96
D.2	PROLIN DEHYDROGENASE (PDH) .....	96
D.3	GLUTARYL-CoA DEHYDROGENASE (GCDH).....	96



D.4 GLUTAMAT DEHYDROGENASE (GDH) .....	97
D.5 SUCCINATE DEHYDROGENASE A (SDHA) .....	97
D.6 SUCCINATE DEHYDROGENASE B (SDHB).....	98

# 1 INTRODUKSJON

## 1.1 Virus i det marine miljøet

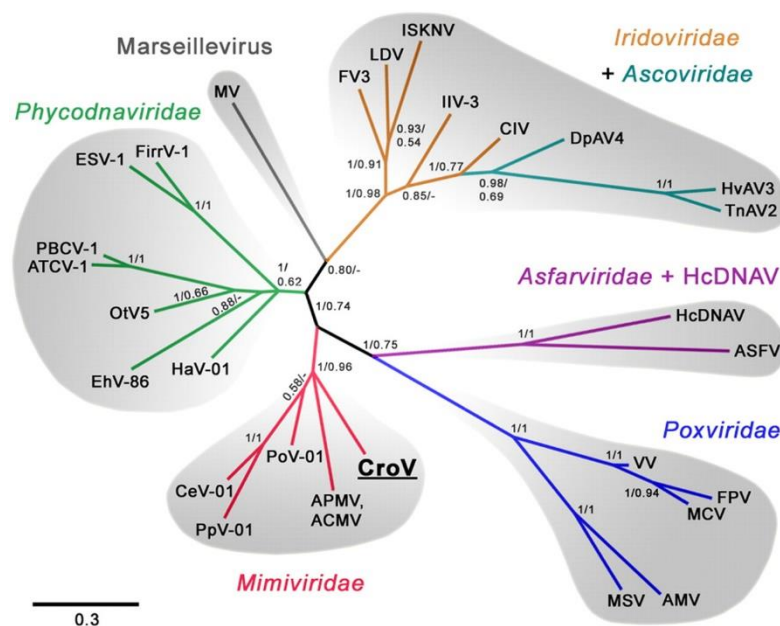
Den første bakteriofagen ble isolert fra det marine miljøet i 1955 (Spencer, 1955), men det var først 20 år senere at det ble oppdaget at virus er de mest tallrike biologiske enhetene i havet (ca.  $10^7$ - $10^8$  ml<sup>-1</sup>) (Bergh et al., 1989), og i dag vet man at virus er den gruppen som har størst mangfold (Sandaa, 2008). Virus har en stor innvirkning på økosystemer, og de påvirker alle mikrobielle prosesser gjennom lysis, inkludert reprogrammering av vertsmetabolismen, horisontal genoverføring og diversiteten av vertssamfunnet (Suttle, 2007). Som et eksempel, har virus utviklet mekanismer som gjør de i stand til å bruke vertens proteiner til deres eget forsvar, og som gjør de i stand til å bruke vertsmetabolismene til sin fordel (Roucourt and Lavigne, 2009). Overføring av gener mellom organismer er en viktig mekanisme innen evolusjon (Sullivan et al., 2006). Gjennom å gi verten nytt genetisk materiale kan virus påvirke evolusjonsraten til verten, og virus kan bytte ut vertsgener med virale homologer (Filée et al., 2003, 2002). I tillegg kan virussamfunnet bli påvirket ved at det tar opp DNA fra verten (Filée et al., 2003), og ved bytte av gener fra andre virus (Botstein, 1980). Ved å lysere nærmere 50 % av biomassen, påvirker virus alle de biokjemiske prosessene i havet når organiske og uorganiske næringsstoffer frigjøres fra vertsorganismer (Fuhrman, 1999). Det er estimert at  $10^{23}$  mikrober i havet blir infisert av en fag hvert sekund (Hendrix, 2003). Generelt så vil antall virus korrelere med produktiviteten i systemet og vil dermed være høyest der hvor bakterier og klorofyll er høyest (Suttle, 2005).

## 1.2 Algevirus

Virus som infiserer alger er sentrale i struktureringen av mikrobielle samfunn og biogeokjemiske sykluser på grunn av deres høye antall og mangfold i marine miljøer (Horas et al., 2018). Jo høyere og kortere generasjonstiden til virus er, jo mer vil de påvirke miljø- og vertssamfunnet (Edwards and Steward, 2018). Algevirus bidrar mest til disse prosessene når de infiserer algeoppblomstringer (Coy et al., 2018). Under algeoppblomstringer kan virus være ansvarlige for 25-100 % av algedødeligheten (Short, 2012). Så langt har man bare så vidt rørt overflaten av mangfoldet til virus som infiserer alger da det er kun ca. 60 algevirus i kultur (Coy et al., 2018), og man antar at alle de ca. 350 000 marine algeartene som finnes har ett eller flere spesifikke virus (Short, 2012). Femti-90 % av gen sekvensene til algevirus har ingen

signifikante treff i databaser (Kristensen et al., 2010). Fylogenetiske analyser viser at en stor del av genomet til disse virusene har blitt ervervet fra eukaryote verter ved horisontal genoverføring (Koonin and Yutin, 2010). Dette gjør at virusene er mer autonome fra vertscellen, som gir dem mulighet til å kontrollere ulike prosesser i eukaryote celler (Claverie and Abergel, 2010).

Haptofytter er en gruppe av encellede alger, som utgjør en viktig del av det autotrofe piko- og nano planktonsamfunnet (Thomsen et al., 1994). Haptofytter finnes overalt, og ofte i høyt antall i det marine miljøet (Liu et al., 2009), og de blir sett på som noen av de viktigste primærprodusentene i havet (Eikrem et al., 2016). Noen haptofytter som *Emiliania huxleyi* og *Phaeocystis* spp. danner omfattende oppblomstringer som kan være skadelige for annet liv (Eikrem et al., 2016; Larsen et al., 2004). Mens slektene som *Haptolina*, *Chrysochromulina* og *Prymnesium* forekommer i mindre antall og produserende ikke like omfattende oppblomstringer (Edvardsen and Paasche, 1998; Gjørseter et al., 2000), samt at enkelte arter innen disse slektene kan produsere toksiner som er skadelig for fisk og andre organismer (Dahl et al., 2005).



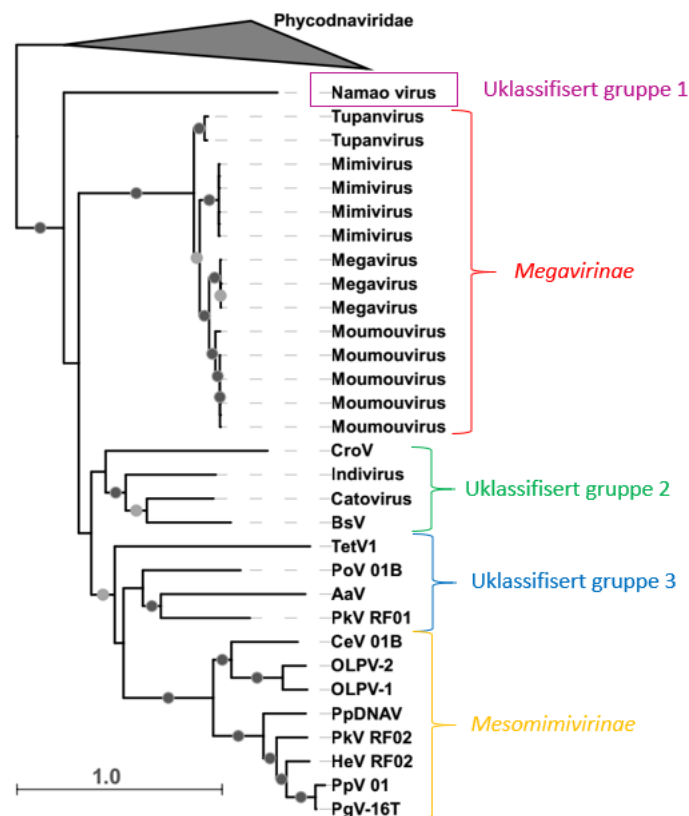
Figur 1.1 Fylogenetisk tre til NCLDV (Fischer et al., 2010). Viser de 7 familiene som hører til NCLDV, og noen av virusene til hver familie.

Alle isolerte virus som infiserer haptofytter har dsDNA, og tilhører familiene *Mimiviridae* og *Phycodnaviridae* (Claverie and Abergel, 2018; Wilson et al., 2009). Disse familiene, sammen med fem andre familier (*Poxviridae*, *Marseilleviridae*, *Asfarviridae*, *Asco-* og *Iridoviridae*) hører til gruppen nukleocytoplasmiske store DNA-virus (NCLDV, figur 1.1) (Iyer et al., 2006,

2001). *Mimiviridae* og *Phycodnaviridae* dominerer NCLDV populasjonene i marine metagenomer med en konsentrasjon opptil  $10^5$  genom per ml i den fotiske sonen (Hingamp et al., 2013), og etter bakteriofager er NCLDV den gruppen med høyest forekomst av virus (Kristensen et al., 2010). De store genomene til disse virusene antyder at infeksjonsprosessene i vertene er komplekse, og disse oppdagelsene har startet en diskusjon om opprinnelsen og evolusjonen til disse virusene (Claverie and Abergel, 2010; Koonin and Yutin, 2010).

### 1.2.1 *Mimiviridae*

Virus som tilhører familien *Mimiviridae* er de mest tallrike og diverse virusene som infiserer autotrofe og heterotrofe eukaryote verter (Claverie and Abergel, 2018; Hingamp et al., 2013; Larsen et al., 2008; Mihara et al., 2018). Familien består av veldig store virus med en genomstørrelse opp til 1,2 Mbp og mer enn 900 kodete proteiner (Raoult et al., 2004). De store genomene koder for flere gener enn hos mange bakterier og små eukaryote mikroorganismer (Abergel et al., 2015). Mangfoldet hos disse virusene tyder på at virusene har en dyp koevolusjonær historie med dets marine verter som går langt tilbake i tid (Koonin and Yutin, 2010).



Figur 1.2: Fylogenetisk tre for *Mimiviridae* familien, hvor de fem undergruppene er markert (Håkon Dahle, ikke publisert).

*Mimiviridae* familien kan deles inn i fem undergrupper basert på fylogenetiske analyser og gruppering i monofyletiske grupper; *Mesomimivirinae*, *Megavirinae* og tre uklassifiserte grupper (Figur 1.2) (Sandaa et al., innsendt til gjennomgang, Encyclopedia of Viruses, Elsevier forlag). Subfamilien *Mesomimivirinae* består av algevirus, hvor verten er *Prymnesiophyceae*. I denne gruppen finner man blant annet Haptolina ericina virus (HeV RF02), Prynesium kappa virus (PkV RF02), Phaeocystis pouchetii (PpV 01) og Chrysochromulina ericina virus (CeV 01B) (Jacobsen et al., 1996; Johannessen et al., 2015; Sandaa et al., 2001). *Megavirinae* består av 14 virus som infiserer *Amoebozoa*. Dette inkluderer Megavirus, Mimivirus, Tupanvirus og Moumouvirus (Abrahão et al., 2018; Legendre et al., 2012; Raoult et al., 2004; Yoosuf et al., 2012). De uklassifiserte virusene deler seg i tre grupper. Gruppe 1 består av Namao virus som infiserer *Actinopterygii* (Clouthier et al., 2013). Gruppe 2 består av fire virus (Bodo saltans virus (BsV), Catovirus, Indivirus og Cafeteria virus (CroV)) som infiserer ulike verter, som blant annet *Bikosea* og *Kinetoplastida* (Claverie and Abergel, 2018; Deeg et al., 2018; Fischer et al., 2014). Gruppe 3 består av de fire virusene Tetraselmis virus (TetV1), Pyramimonas orientalis virus (PoV 01B), Aureococcus anophagefferens (AaV) og Prynesium kappa virus (PkV RF01), som infiserer arter innen slektene av *Prasinophyceae*, *Pelgaophyceae* og *Prymnesiophyceae* (Johannessen et al., 2015; Moniruzzaman et al., 2014; Sandaa et al., 2001; Schvarcz and Steward, 2018).

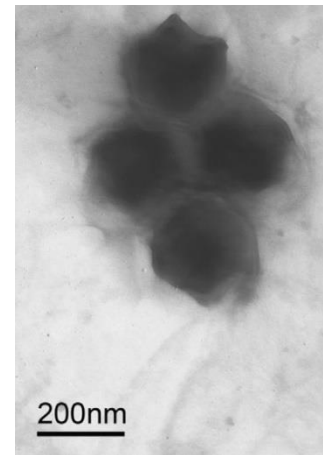
PpV 01, CeV 01B og PoV 01B var de første store virusene som ble isolert (Jacobsen et al., 1996; Sandaa et al., 2001). Da disse ble isolert på 1990-tallet ble de først beskrevet som *Phycodnaviridae*. Det er først i de senere år man har funnet ut at disse tilhører gruppen *Mimiviridae* (Claverie and Abergel, 2018). I dag vet man at virusene i *Mimiviridae* infiserer et høyt mangfold av eukaryote encellede organismer, som inkluderer seks store fyla; *Amoebozoa*, *Haptophyta*, *Chlorophyta*, *Exavata*, *Heterokonta* og *Opisthokonta* (Claverie and Abergel, 2018).

Alle isolerte algevirus i *Mimiviridae* familien har icosahedral kapsidstrukturer, dsDNA genomer og de har en latent periode fra 10 – 72 timer (Baudoux and Brussaard, 2005; Brown and Bidle, 2014; Jacobsen et al., 1996; Johannessen et al., 2015; Sandaa et al., 2001; Schvarcz and Steward, 2018). Virusene i familien er mest lik virusene som infiserer heterotrofe protister og mindre lik de som infiserer andre autotrofe protister i *Phycodnaviridae* familien (Sandaa et al., innsendt til gjennomgang, Encyclopedia of Viruses, Elsevier forlag). Samtidig finner man flere gener av *Phycodnaviridae* i genomer til *Mimiviridae* virus, noe som kan tyde på en felles opprinnelse (Moniruzzaman et al., 2014). For virusene som tilhører NCLDV har man oppdaget

177 ortologe gener som finnes i minst en av familiene (Koonin and Yutin, 2010), og av disse genene finner man syv i genomsekvenserte *Mimiviridae* virus. DNA polB, RNAPol, MCP, Asparagin syntese (AsnS), A32-lignende virionpakket ATPase, eukaryotisk translasjonsinitierende faktor 4E (eIF-4E) og MutS7 finnes i alle *Mimiviridae* virusene bortsett fra AaV som mangler AsnS og eIF-4E homologer (Sandaa et al., innsendt til gjennomgang, Encyclopedia of Viruses, Elsevier forlag).

#### 1.2.1.1 *Prymnesium kappa virus (PkV RF01)*

*Prymnesium kappa virus* (PkV RF01) ble først isolert i Raunefjorden i tidsrommet 2009-2011 og er det største algeviruset så langt isolert (kapsid størrelse 310 nm) (Johannessen et al., 2015). Viruset kan lysere flere stammer av *Haptolina ericina* og en stamme av *Prymnesium kappa*, noe som er uvanlig da kultiverte virus ofte er artsspesifikke. Tidligere er dette bare kjent fra CbV-PW1 som infiserer *Haptolina brevifila* og *Chrysochromulina strobilus* (Suttle and Chan, 1995). PkV RF01 har en icosahedral symetri bestående av tyveflatede kapsider og viruspartiklene (Figur 1.3) blir trolig produsert i cytoplasma (Johannessen et al., 2015).



Figur 1.3: PkV RF01 partikler i lysat (Johannessen et al., 2015)

Viruset har en latent periode på 24-32 timer og lyserer bare 90 % av vertscellene (Johannessen et al., 2015). Virusets nærmeste, men likevel fjerne slektning er AaV (Sandaa et al., innsendt til gjennomgang, Encyclopedia of Viruses, Elsevier forlag).

### 1.3 Auksiliære metabolske gener

Virus påvirker som nevnt tidligere det marine økosystemet ved blant annet reprogrammering av vertsmetabolismen (Suttle, 2007), og her spiller auksiliære metabolske gener (AMG) en viktig rolle. Tidligere trodde man at disse genene bare fantes i cellulært liv, men de har nå blitt identifisert i flere virale genomer (Breitbart et al., 2007; Lindell et al., 2005; Sullivan et al., 2006; Weigle et al., 2007). I gruppen NCLDV er det beskrevet et høyt antall av auksiliære metabolske gener (Koonin and Yutin, 2010), men man vet fortsatt ikke funksjonen til mange av disse. I andre virus har det blitt beskrevet AMGer hvor funksjonen påvirker verten (Breitbart et al., 2007; Monier et al., 2017). Auksiliære metabolske gener er gener som har blitt overført fra verten til virus, ved horisontal genoverføring (Breitbart, 2011; Breitbart et al., 2007).

Uttrykkelsen av disse genene er antatt å være kritisk for den økologiske suksessen til viruset (Breitbart, 2011). Disse genene kan utvide evnene til verten under infeksjonen, og de er ofte involvert i kritiske steg i vertsmetabolismen (Enav et al., 2014; Hurwitz et al., 2013). Det har blitt oppdaget AMGer som blant annet er involvert i fotosyntesen, svovelmetabolismen, nitrogenmetabolismen, karbon metabolismen og DNA/RNA behandling (Anantharaman et al., 2014; Arslan et al., 2011; Monier et al., 2017; Sharon et al., 2009; Thompson et al., 2011).

Stadig flere studier viser at genoverføring mellom virus og vert er en viktig del av den virale evolusjonen (Lindell et al., 2004; Monier et al., 2017, 2009). Men for flere AMGer vet man fremdeles ikke hvilken funksjon de har i viruset. Nylig har det blitt oppdaget et AMG som er involvert i nitrogenopptak i marine fytoplankton (Monier et al., 2017). N-transport genet (*vAmt*) ble funnet i ett algevirus (*OtV6*) som infiserer grønnalgen *Ostreococcus tauri*. Monier et al. 2017 har bevist at genet blir transkribert under infeksjon, og at det transporterer  $\text{NH}_4^+$  og en rekke andre nitrogen kilder. Ved lave konsentrasjoner av ammonium, viser det seg at den virale versjonen av genet har et høyere opptak av  $\text{NH}_4^+$  sammenlignet med *O.tauri* homologen. Tilgang på næringsstoff har vist seg å være viktig for fytoplanktonvirus, og lave konsentrasjoner av P og N kan påvirke latenttid og lysis av celler (Cheng et al., 2015; Mojica and Brussaard, 2014). Når virale proteiner da endrer vertens nitrogen metabolisme, vil dette være til fordel for viruset (Monier et al., 2017).

#### 1.4 Virale auxiliære metabolske gener med potensiale for bioprospektering

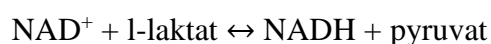
I denne oppgaven ble det arbeidet med sju gener fra PkV RF01 (Tabell 1.1, sekvenser er vist i vedlegg D.1-6). Disse proteinene er enzymer som katalyserer ulike prosesser i vertscellen, men man vet enda ikke hvilken funksjon de har for PkV RF01. Disse proteinene er interessante for grunnforskning, men å finne funksjonen til virus-genprodukter kan være vanskelig. For å undersøke om de predikerte funksjonene til de valgte ORFene stemmer og for å bekrefte at de har en spesifikk funksjon i virus, er det viktig å produsere og karakterisere proteinene. Seks av enzymene er dehydrogenaser, som er en gruppe enzymer som hører til oksidoreduktaser, og som oksiderer et substrat ved å redusere en elektronakseptor (Alberts et al., 2002a). L-laktat dehydrogenase (L-LDH), Glutaryl-CoA dehydrogenase (GCDH), glutamat dehydrogenase (GDH) og de to subenhetene (SDHA og SDHB) involvert i succinat dehydrogenase kompleks II har tidligere ikke blitt beskrevet i virus.

Tabell 1.1: Proteiner fra PkV RF01 som det ble arbeidet med i denne oppgaven, og informasjon om deres plassering i PkV RF01 genomet (ORF), størrelse (antall aminosyrer og basepar), og om proteinene har blitt beskrevet i virus tidligere. Informasjonen er hentet fra annoteringstabell av PkV RF01 genomet (Ruth-Anne Sandaa, UiB, 2019)

Navn	ORF	Antall aminosyrer	Antall basepar	Funnet i virus tidligere
<b>L-laktat dehydrogenase</b>	ORF79	394	1184	Nei
<b>Prolin dehydrogenase</b>	ORF188	277	833	Ostrecoccus tauri virus 1
<b>Glutaryl-CoA dehydrogenase</b>	ORF236	410	1232	Nei
<b>DNA polymerase X</b>	ORF630	404	1214	Phaeocystis globosa virus (og flere)
<b>Glutamat dehydrogenase</b>	ORF951	415	1247	Nei
<b>Succinate dehydrogenase A</b>	ORF893	602	1808	Nei
<b>Succinate dehydrogenase B</b>	ORF900	233	701	Nei

#### 1.4.1 L-laktat dehydrogenase (L-LDH)

L-laktat dehydrogenase (L-LDH) har tidligere ikke blitt oppdaget i virus, men ifølge protein databasen COG (Tatusov et al., 1997, <http://www.ncbi.nlm.nih.gov/COG>) har dette enzymet en predikert funksjon i lipidtransport og metabolisme, samt energiproduksjon. I andre organismer vet man at enzymet katalyserer reaksjonen:



Pyruvat er sluttproduktet i glykolysen, som er startfasen i alle typer cellerespirasjoner. Dette er en katabolisme i mange levende celler hvor glukose blir omdannet til pyruvat. Videre kan laktat dehydrogenase katalyserer reaksjonen hvor pyruvat blir omdannet til laktat. Samtidig produseres  $\text{NAD}^+$  fra  $\text{NADH}$ , noe som er nødvendig for å lage ATP gjennom glykolysen selv uten oksygen (Markert, 1984).

L-LDH er et viktig kommersielt enzym, da det ofte blir brukt til å lage melkeprodukter, samtidig som enzymet har en tydelig sur smak som utnyttes ved lagning av øl og vin (Ren, 2010).



#### 1.4.2 Prolin dehydrogenase (PDH)

Prolin dehydrogenase (PDH) er et universelt protein i levende organismer som inngår i prolin katabolismen. Det har blitt oppdaget i viruset *Ostrecoccus lucimarinus* virus 1 (Weynberg et al., 2011), men det er fortsatt ukjent hvilken funksjon dette genet har hos virus. I andre organismer er PDH et FAD-avhengig enzym som katalyserer det første steget i to-steps oksidasjonen av prolin (Servet et al., 2012). L-prolin blir omgjort til delta1-pyrrolin-5-karboksyilat (P5C), som videre blir omgjort til glutamat (Serrano and Blanchard, 2013).

I dag har PDH en viktig rolle i blant annet ulike biosensorer og kit brukt innen diagnostikk (Shahbaz- Mohammadi and Omidinia, 2011).

#### 1.4.3 Glutaryl-CoA dehydrogenase (GCDH)

Glutaryl-CoA dehydrogenase (GCDH) er tidligere ikke beskrevet i virus, men den predikerte funksjonen inngår i lipidtransport og lipidmetabolisme. I cellulære organismer er enzymet viktig i fettsyre- og aminosyre metabolisme (Shen et al., 2009). Enzymet katalyserer den oksidative dekarboksyleringen av glutaryl-CoA til krotonyl-CoA og CO<sub>2</sub> hvor enzymet bruker flavoprotein som elektronakseptor (Goodman et al., 1995; Lenich and Goodman, 1986). I pattedyr er glutaryl-CoA et mellomprodukt i oksideringen av aminosyrene lysin, hydroksylisin og tryptofan (Goodman et al., 1995).

#### 1.4.4 DNA polymerase X (Pol X)

DNA polymeraser er både viktige for cellulære organismer og virus, da de er avhengig av enzymet for replikering og vedlikehold av genomet, men det er også en viktig del av molekylærbiologien, da de brukes i teknikker som blant annet kloning og PCR (Gardner and Kelman, 2014).

Hovedfunksjonen til polymeraser i organismer, er å kopiere templat tråden for å lage en datter tråd (Coen, 1996). Dette vil si at den primære katalytiske aktiviteten til polymeraser er å overføre en nukleotiddel av et NTP, som er komplementær til templattråden, til 3' enden av en voksende dattertråd av RNA eller DNA (Alberts et al., 2002b). Virus bruker polymeraser ved flere steg i livssyklusen. Først må viruset konvertere genomet til mRNA som videre benyttes i proteinsyntese. Deretter må identiske kopier av genomet blir syntetisert for nye viruskapsider. Om viruset har behov for en polymerase for transkripsjon og/eller replikasjon avhenger av den

cellulære plasseringen til viruset (Coen, 1996). Hvis viruset tilbringer hele livssyklusen i cytoplasma har de ikke tilgang på vertens polymeraser, og må da kode for sine egne polymeraser. Mens andre virus (ofte små virus) tilbringer syklusen sin i nukleus og kan dermed bruke vertens maskineri for replikasjon og mRNA syntese (Choi, 2012). To metallioner;  $Mg^{2+}$  og  $Mn^{2+}$  er nødvendige for polymerase aktiviteten, samtidig som at polymerasen må ha et bindingssete for templatet og NTP.

Familie X DNA polymeraser er involvert i steget hvor mellomrom (gaps) fylles inn ved baseutkuttingsreparasjon (BER) mekanismen, som reparerer skader på de enkelte basene (Fernández-García et al., 2017). Disse polymerasene er små, monomere protein som katalyserer innsettingen av nukleotider og som mangler en egen korrekturlesningsaktivitet (Moon et al., 2007). I tillegg til bakterier, arkeer og eukaryote har DNA polymerase X blitt beskrevet i virus tilhørende noen av familiene i NCLDV (*Marseilleviridae*, *Mimiviridae* og *Poxviridae*). PgV-16T og CeV er de virusene hvor det først har blitt beskrevet at DNA ligase genet er festet til C-terminalen av DNA polymerase X (Gallot-Lavallée et al., 2015; Santini et al., 2013). Det viser seg at PgVPolX deltar i de tre siste stegene i BER mekanismen, da dette genet i tillegg til polymerase aktivitet, har en lyase og ligase aktivitet (Fernández-García et al., 2017).

#### 1.4.5 Glutamat dehydrogenase (GDH)

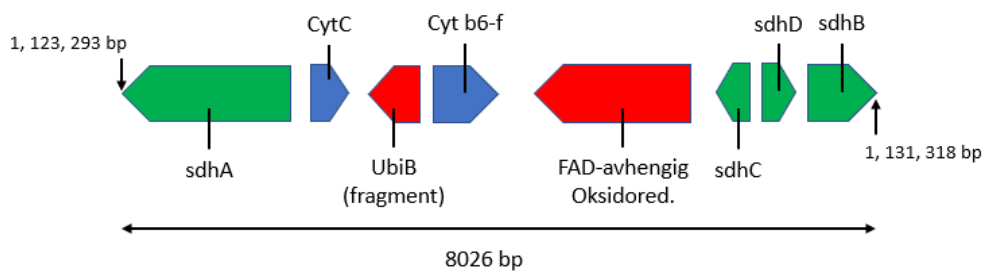
Katalyserer assimileringen av  $NH_3$ .

Glutamat dehydrogenase (GDH) har tidligere ikke blitt beskrevet i marine virus. Dette er et enzym som inngår i aminosyre transport og metabolisme, hvor det katalyserer den reversible omdannelsen av glutamat til alfa-ketoglutarat og ammonium samtidig som det reduserer  $NAD(P)^+$  til  $NAD(P)H$  (Plaitakis et al., 2017). I tillegg til å bidra til anaplerotiske reaksjoner og energiproduksjon i trikarboksylsyresyklus, er GDH-funksjonen knyttet til redoks homeostase og celledesignal prosesser (Durán et al., 2012; Jin et al., 2015). Studier viser at GDH er et av enzymene som spiller en viktig rolle i å opprettholde nitrogen og karbon balansen i celler (Mifflin and Habash, 2002), og har en rolle i assimileringen av  $NH_3$ .

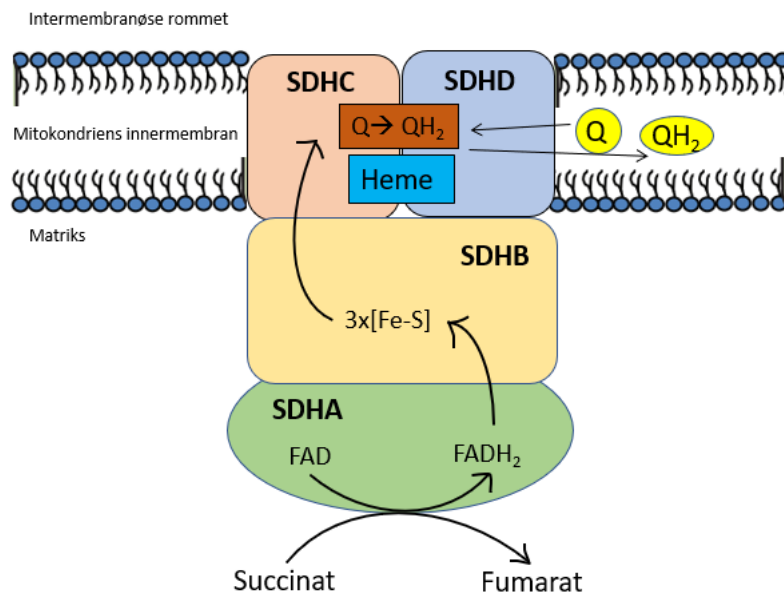
#### 1.4.6 Succinat dehydrogenase kompleks II

I genomet til PkV RF01 har man funnet alle genene involvert i succinat dehydrogenase kompleks II i den oksidative fosforyleringen og trikarboksylsyresyklus. Dette er gener som

tidligere ikke har blitt oppdaget i virus (Blanc-Mathieu et al., under utarbeiding). Alle genene er funnet i samme område på virus genomet (Figur 1.4), sammen med to andre gener (merket rødt på figur 1.4) som kan være involvert i andre prosesser. I de fleste eukaryote organismer finner man disse genene fordelt mellom mitokondriet og nukleus (Adams et al., 2001). Det er derfor blitt spekulert i om plasseringen av genene i viruset kan forklares med at viruset tilegnet seg disse genene veldig tidlig i evolusjonen, før fordelingen av genene mellom mitokondriet og nukleus i eukaryote organismer.

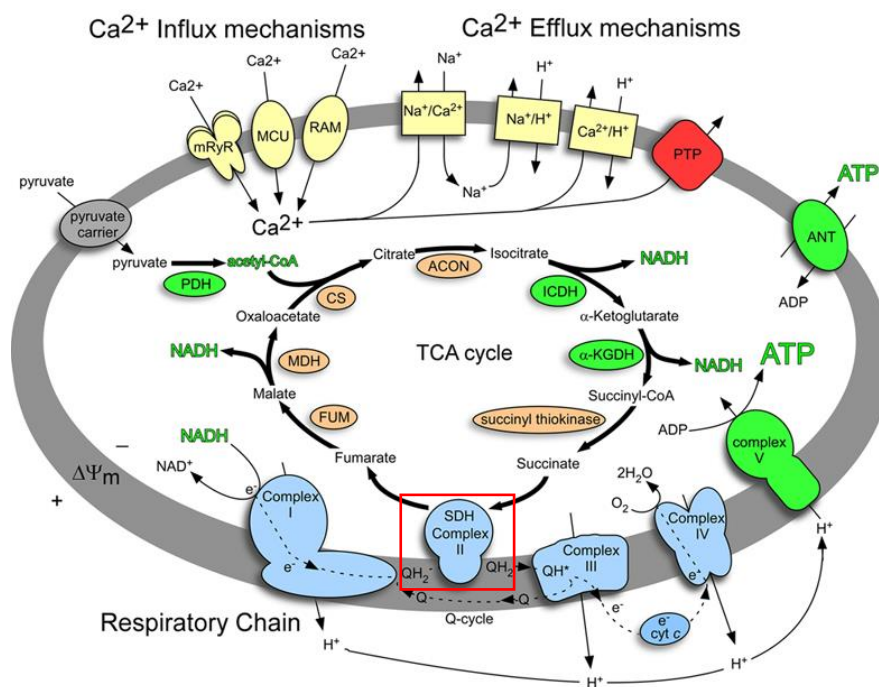


Figur 1.4: Gener funnet i PkV RF01 som er involvert i SDH kompleks II. Grønn viser de fire subenhetene (A-D), blå er cytokrom gener, mens rødt viser de to genene som kan være involvert i andre prosesser. Retning på pilen representerer retning til genet i genomet.



Figur 1.5: Plasseringen av succinat dehydrogenase kompleks II i eukaryote celler. Basert på figur fra <https://commons.wikimedia.org/w/index.php?curid=2287606>

Mitochondrie er en organell som er funnet i de fleste eukaryote celler (Sun et al., 2005). Mitokondrielle respirasjonssystemer består av fem protein komplekser (figur 1.6) og produserer det meste av energien i eukaryote celler (Saraste, 1999). Succinat dehydrogenase kompleks II er plassert i den indre membranen av mitokondriet. Komplekset oksiderer succinat til fumarat i trikarboksylsyresyklus og denne oksideringen er koblet til reduseringen av ubikinon til ubiquinol i den indre membranen som en del av elektrontransportkjeden (Sun et al., 2005). Komplekset består av fire subenheter (figur 1.5); et flavoprotein (SDHA) som har en FAD kofaktor, et jernsvovel protein (SDHB) som har tre Fe-S grupper, og to små membranproteiner (SDHC og SDHD) som binder heme for å danne cytokrom b (Sun et al., 2005). SDHA er den subenheten som oksiderer succinat til fumarat, samtidig som at FAD omdannes til FADH<sub>2</sub>. Elektronene fra FADH<sub>2</sub> vil så bli transportert til SDHB hvor de når Fe-S gruppene, og deretter videre til ubiquinon pool via SDHC og SDHD hvor ubiquinon blir redusert til ubiquinol (Huang and Millar, 2013).



Figur 1.6: Trikarboksylsyresyklus og den oksidative fosforyleringen i cellen, hvor man ser at SDH kompleks II (markert med rødt) har en viktig rolle i begge prosessene (Batty et al., 2009).

## 1.5 Marin bioprospektering

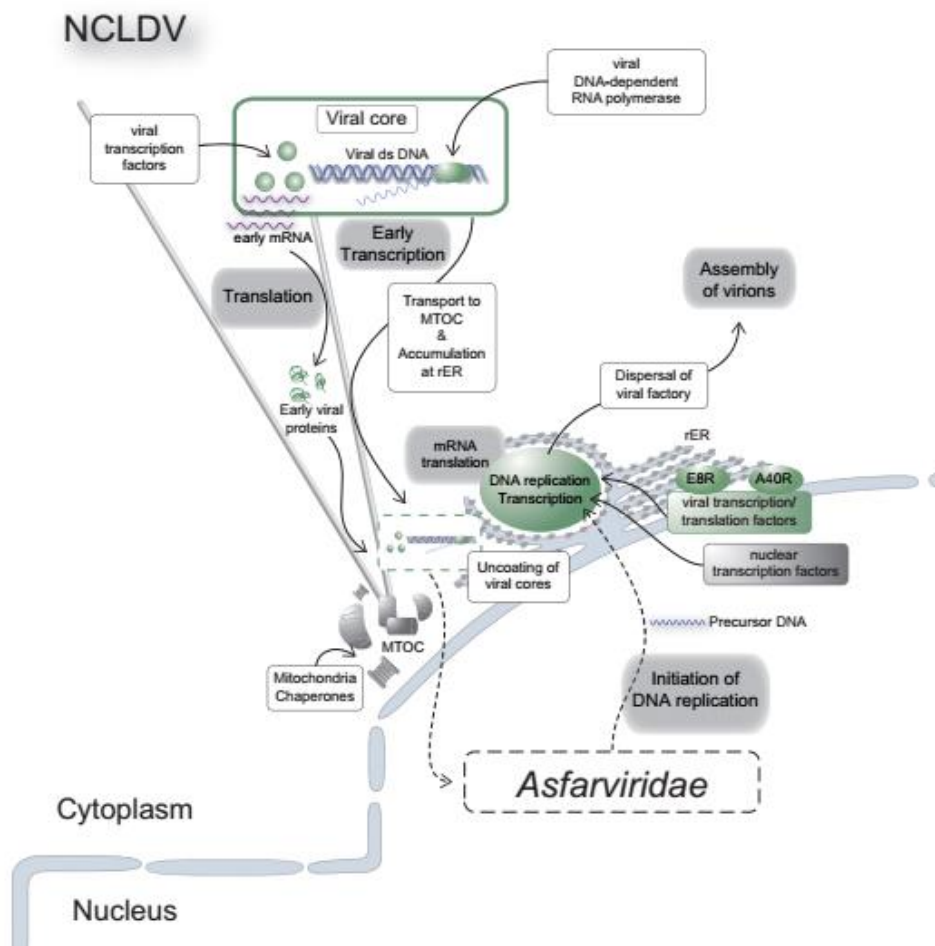
Marin bioprospektering er letingen etter nye verdifulle bioaktive molekyler i organismer med mål om å utnytte dette til å blant annet lage nye medisiner, i næringsmiddelindustri eller kjemisk industri (Bhatia and Chugh, 2015; de Pascale et al., 2012). I dagens utnyttelse av marine ressurser i forbindelse med bioprospektering er marine habitater en viktig del da de har ett høyt mikrobielt mangfold, hvor både bakterier, sopp og virus har genomer med mye genetisk informasjon som inngår i forskjellige metabolske prosesser (Lozada and M. Dionisi, 2015). Virussamfunn er i dag det største reservoaret for ukjent genetisk diversitet, og er av stor interesse når det kommer til utnyttelse av marine ressurser. På bakgrunn av den uvanlige biologien og livssyklusene til virus har de utviklet spesielle replikasjonsmaskinerier og har enzymer som er unike sammenlignet med verts enzymer (McGrath et al., 2004). DNA og RNA modifierende enzymer har en viktig rolle i molekylærbiologien, og dette inkluderer enzymer som blant annet polymeraser, ligaser og reverse transkriptaser. Disse og andre enzymer kan knyttes til replikasjonszyklusen til virus. Blant annet er T4 bakteriofagen kilden til flest nyttige virale enzymer, og har minst 85 gener som er involvert i DNA replikering, reperatur, nukleotidmetabolisme, transkripsjon og translasjon (Miller et al., 2003). Dette viruset og andre har vært viktige i utvikling av bioteknologien, på grunn av det store antallet av visse gen i genomet. Til sammenligning har et 2 Mb bakteriegenom bare et *polI* gen (koder for DNA polymerase I), men for virus kan det bli funne mellom 20-40 *pol* gener per 2 Mb i virale metagenomiske sekvenser (Schoenfeld et al., 2010) og flere av disse har vist seg å være nyttige.

Utnyttelsen av enzymer har store teknologiske og økonomiske fordeler. De er blant annet brukt i produksjonen av søtningsmiddel og vaskemidler, og de har en sentral rolle i ulike analytiske verktøy innen rettsmedisin. Frem til 1970-tallet var det vanlig å bruke enzymer hentet fra dyr og planter. Men fordelene med mikrobielle enzymer ble etter hvert oppdaget og man så store forskjeller ved å bruke mikrobielle enzymer sammenlignet med å hente de samme enzymene fra dyr og planter (Robinson, 2015). Blant annet er det mye enklere å produsere store kvanta av mikrobielle enzymer, og man kan gjøre dette på mye mindre fasiliteter, samtidig som at mange av de mikrobielle enzymene er mye enklere å ekstrahere og rense. I tillegg har mikrobielle enzymer ofte egenskaper som gjør de mer passende for kommersiell utnyttelse. Blant annet er de mer stabile, og de er mer mottagelig for modifisering ved å bruke enkle metoder som plasmidinnsetting (Robinson, 2015).

## 1.6 Metoder benyttet i proteinarbeid og kvantifisering av genprodukt

### 1.6.1 Replikering av virus for gruppen NCLDV

Alle virusene som tilhører gruppen NCLDV danner cytoplasmiske replikasjonskammer, ofte kalt virale fabrikker, som er involvert i replikasjon og sammenstillingen av nye virus i vertscellen. Disse kammerene skjerner de for forsvarsmekanismer til verten samtidig som viruset vil stoppe proteinsyntesen til verten (Figur 1.7) (Sánchez et al., 2013). Disse kammerene blir ofte dannet ved omfattende omorganisering av vertens cellemembranrom (Schmid et al., 2014). Noen virus (*Poxviridae*) replikerer bare i cytoplasma, mens andre (*Asfarviridae*) replikerer både i cytoplasma og kjernen (Iyer et al., 2006, 2001).



Figur 1.7: Oversikt over replikasjonskammer som dannes i alle virus som tilhører gruppen NCLDV (Schmid et al., 2014).

*Asfarviridae* blir replikert i virale fabrikker som fordeles i cytoplasmaet. Genomreplikasjonen til dette viruset starter i cellekjernen, hvor det dannes korte DNA fragmenter (García-Beato et al., 1992). Deretter blir disse DNA fragmenter transportert fra kjernen til de cytoplasmiske replikasjonskammerene, hvor de brukes som primere for genomreplikasjon (Rojo et al., 1999). Transkripsjonen av tidlige gener og post transkripsjonelle modifiseringer av mRNA blir initiert av transkripsjonsfaktorer og RNA polymerase i viruskjernen (Figur 1.7), og er uavhengig av cellulære enzymer (Kuznar et al., 1980; Salas et al., 1986). Dermed vil mRNA dannes som vil bli transportert til cytoplasma hvor de vil inngå i translasjonen. Deretter vil tidlige proteiner bli transportert til mikrotubuli organiserende senter (MTOC) hvor de vil akkumulere og vil bli frigjort når virus sammenstillingen starter. Samtidig er sent genuttrykk avhengig av den virale DNA-replikasjonen og virale tidlige proteiner. Mens virusgenomet blir replikert vil viruset gjøre store forandringer på vertens celle. Blant annet vil mitokondriet, chaperoner og ER membraner bli flyttet til replikasjonskammerene, og trans-Golgi nettverket vil bli ødelagt (McCrossan et al., 2001) .

#### *1.6.1.1 Første steg i bestemmelsen av genets funksjon*

Transkripsjon er første steget ved ekspresjon av gener hvor en RNA-kopi (mRNA) blir laget av en DNA sekvens. For et gen som koder for protein vil informasjonen i RNA kopien være nødvendig for å bygge proteinet. Dette vil si at hvis genet ikke blir transkribert, tyder dette på at genet ikke har en funksjon i cellen. Ved å isolere mRNA og bruke dette som templat under RT-PCR kan man undersøke om et gen blir transkribert eller ikke. RT-PCR står for revers transkripsjon PCR og er en metode som gjør det mulig å detektere og amplifisere mRNA. I prosessen kalt reverse transkripsjon bruker enzymet reverse transkriptase mRNA i prøven for å lage en komplementær enkeltrådet DNA tråd som kalles cDNA. Deretter vil DNA polymerase konvertere enkeltrådet cDNA til dobbeltrådet DNA, som kan bli brukt som templat i en PCR-reaksjon (Carter and Shieh, 2015). Om man klarer å amplifisere DNA i PCR-reaksjonen, har man bevist at genet blir transkribert i cellen, og første steg i prosessen med å finne funksjonen til genet er gjennomført.

#### *1.6.1.2 Kloning, uttrykkelse og rensing av proteiner*

For å kunne karakterisere et protein må proteinet først uttrykkes og renses slik at man får store kvanta av proteinet. Muligheten til å kunne uttrykke og rens rekombinante proteiner i store mengder har gjort at dette har blitt en mye brukt metode ved proteinkarakterisering, ved bruk i industrielle prosesser og i utviklingen av kommersielle varer (Rosano and Ceccarelli, 2014). I denne oppgaven ble et pET ekspresjonssystem brukt for å klonere og uttrykke fem proteiner.



Dette systemet under kontroll av T7 bakteriofag. Målggen blir først klonet ved å bruke en vert som ikke har T7 RNA polymerase genen, for å eliminere ustabilitet på grunn av produksjon av toksiske proteiner. Deretter blir plasmidet overført til en ekspresjonsvert som inneholder en kromosomkopi av T7 RNA polymerase genen som er under kontroll av *lacUV5*. Ekspresjonen av proteinet blir induisert ved tilsetning av IPTG, som vil gjøre at *lacUV5* ikke lenger er hemmet og ekspresjon av målggen vil forekomme da dette er plassert bak RNA polymerase genen (Novagen, 1999).

### 1.6.2 Droplet digital PCR (ddPCR)

Droplet digital PCR (ddPCR) er en nyere teknologi som blir brukt for å kvantifisere DNA i en prøve (Vogelstein and Kinzler, 1999). I denne teknikken blir PCR-reaksjonsmiksen fordelt på flere tusen små dråper (droplets), hvor hver enkelt dråpe inneholder en uavhengig PCR-reaksjon (Figur 1.9). Ved å måle fluorescens intensitet til hver dråpe blir mengden DNA så bestemt basert på antall dråper som er positive, dvs. antall dråper som har amplifisert DNAet av interesse (Shehata et al., 2017). Dette er en metode som blant annet brukes for å kvantifisere genuttrykk, og i denne oppgaven ble ddPCR brukt for å kvantifisere uttrykkelsen av *sdhA* genen ved forskjellige tidspunkter under infeksjonssyklusen.



Figur 1.8: Trinnvis fremstilling av hvordan DNA blir kvantifisert ved bruk av ddPCR metoden (Christof Winter, 2016).

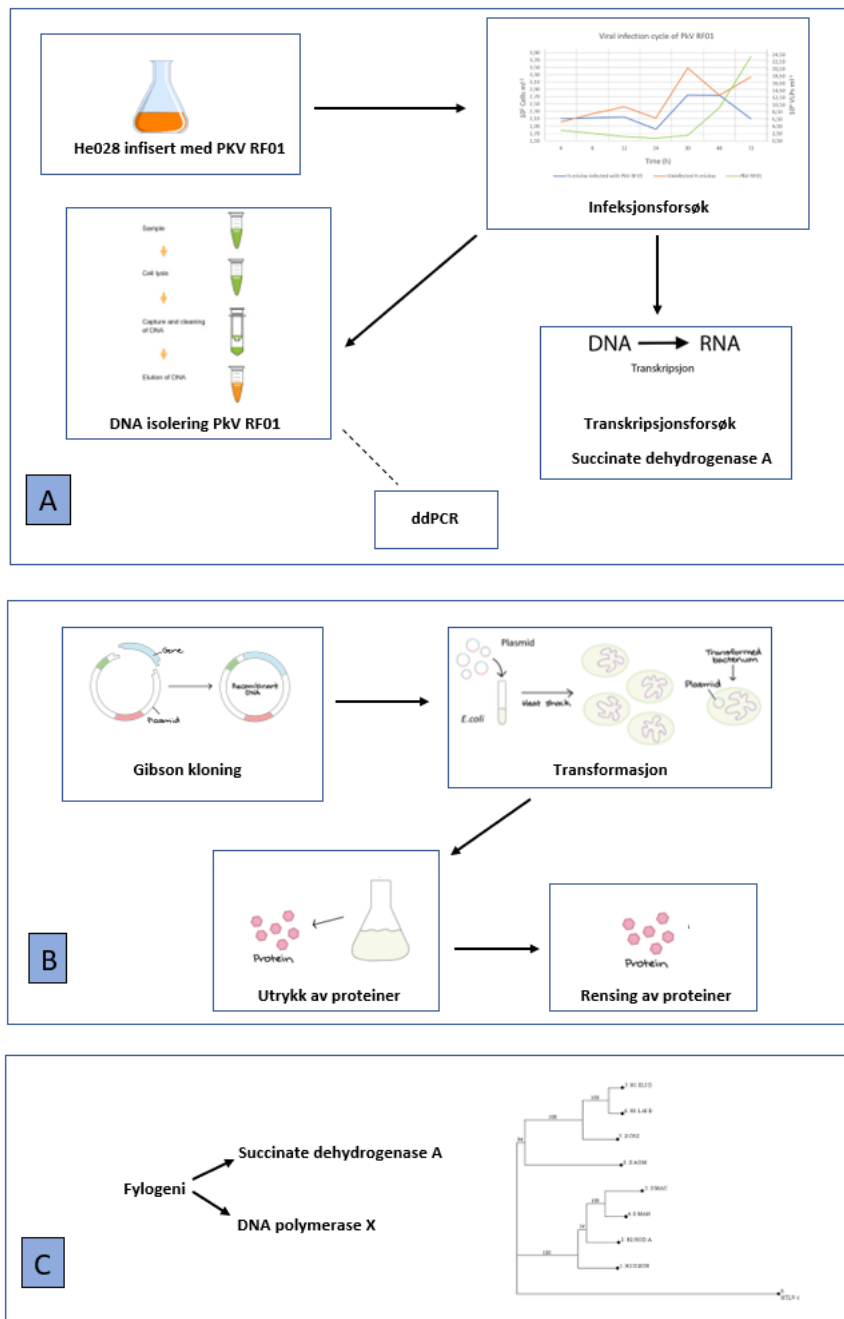


## 1.7 Mål for denne masteroppgaven

Målet med dette masterprosjektet var å undersøke ulike proteiner i PkV RF01 som kan ha både en økologisk og økonomisk verdi. Oppgaven ble utført i forbindelse med EU Horizon 2020 prosjektet Virus-X prosjekt (No 685778). Oppgaven ble begrenset til arbeid rundt algeviruset PkV RF01, da dette er det største marine viruset så langt isolert, og hvor mange av genene som finnes i genomet samt den økologiske rollen til viruset fortsatt ikke er beskrevet. Da det nylig har blitt oppdaget at viruset har hele succinat dehydrogenase kompleks II, som inngår i den oksidative fosforyleringen og trikarboksylsyresyklus, er det av stor interesse å finne ut hvilken funksjon genene i dette komplekset har for viruset. Første steg i denne oppgaven er dermed å finne ut om genet blir transkribert under virusinfeksjon, og videre om disse genene ble uttrykt og eventuelt når i infeksjonssyklusen. Det ble også undersøkt evolusjonistisk opprinnelse til dette genet ved hjelp av fylogenetiske metoder. Videre ble det arbeidet med fem auksiliære metabolske virusgener med mulig potensiale innen bioprospektering. Disse ble forsøkt klonet, uttrykt og rensset for å videre kunne karakterisere de og for å se mulighetene for økonomisk verdi. Basert på dette ble det definert to delprosjekter med følgende forskningsspørsmål:

1. Succinat dehydrogenase subenhet A og B
  - a. Er SDHA og SDHB aktive proteiner i PkV RF01?
  - b. Når i infeksjonssyklusen til PkV RF01 blir SDHA eventuelt uttrykt?
  - c. Hva er den evolusjonistiske opprinnelsen til SDHA i PkV RF01?
  
2. Auksiliære metabolske virusgener med mulig potensiale for bioprospektering
  - a. Er det mulig å uttrykke utvalgte proteiner fra PkV RF01 i *E. coli*?
  - b. Er de uttrykte PkV RF01 proteinene løselige?
  - c. Hva er den evolusjonistiske opprinnelsen til DNA polymerase X i PkV RF01?

## 2 MATERIALER OG METODER



Figur 2.1.1: En enkel oversikt over hvordan eksperimentet ble gjennomført i denne oppgaven. Eksperimentet har tre deler, hvor arbeidet i del A handler om å finne ut om proteinet Succinate dehydrogenase A (PkV RF01) er aktivt, mens del B omfatter kloning, uttrykking og rensing av 5 gener tilhørende PkV RF01. Del C omhandler fylogenetiske undersøkelser rundt Succinate dehydrogenase A og DNA polymerase X. ddPCR i del A ble ikke gjennomført slik som planlagt ved eksperimentets start, og er derfor merket med en stiptet linje. Bare optimaliseringen av ddPCR ble gjennomført.

## 2.1 Dyrking av Prymnesium kappa virus (PkV RF01)

### 2.1.1 Dyrking av verten, haptofytten *Haptolina ericina*

En kultur av haptofytten *Haptolina ericina*, isolat He028, ble hentet fra algekultursamlingen ved Institutt for biovitenskap ved Universitet i Bergen. Kulturen ble dyrket i ½ IMR medium med selen (vedlegg B.1) ved 16 °C i lys-mørkesykluser med fordeling 14:10 under ~50  $\mu\text{mol fotoner m}^{-2} \text{s}^{-1}$ . For å holde cellene i vekstfase ble kulturen overført ukentlig. Kulturen ble brukt til produksjon av Prymnesium kappa virus (PkV RF01), samt i infeksjonsforsøk med PkV RF01 og til DNA isolering av både virus og vert.

### 2.1.2 Dyrking av algeviruset PkV RF01

Viruset som ble brukt i denne oppgaven var Prymnesium kappa virus (PkV RF01), som ble isolert fra Raunefjorden (60°16.2'N, 5°12.5'E) i tidsrommet 2009-2011 (Johannessen et al., 2015). PkV RF01 ble hentet fra virussamlingen ved Institutt for biovitenskap ved Universitet i Bergen. En He028 kultur i eksponentiell vekst (ca  $10^5$  celler  $\text{ml}^{-1}$ ) ble infisert med PkV RF01 i forholdet 1:100 (virus:vert, ca. 10 PkV RF01 per 1 vertscelle). Den infiserte kulturen og en kontrollkultur (ikke tilsatt virus) ble inkubert ved 16 °C under vanlige vekstforhold for He028 frem til synlige tegn på lysis, sammenlignet med kontrollkulturen. Ved infeksjonsforsøk ble antall virus- og vertspartikler bestemt ved flowcytometri (FCM) (se 2.1.4).

Cellerester ble fjernet fra virus lysatet ved sentrifugering (Beckman Coulter Allegra 2IR Centrifuge, Indianapolis, USA) i 15 minutter ved  $4600 \times g$  (10 °C). Lysatet ble oppbevart ved 4 °C frem til bruk.

### 2.1.3 Infeksjonsforsøk PkV RF01

Det ble satt opp et infeksjonsforsøk hvor prøver ble tatt ved bestemte tidspunkter. Tidspunktene var: 0, 3, 15, 24, 27 og 39 timer etter tilsetning av virus. Ved hvert tidspunkt ble det tatt ut prøver til ddPCR (2.4), transkripsjonsforsøk (2.5) og telling av verts- og virusceller på flowcytometer (2.1.4). Til forsøket ble det laget til en 180 ml kultur av He028 ( $10^5$  celler) infisert med 20 ml PkV RF01. Som kontroll ble det brukt en tilsvarende kultur av He028 som ikke ble tilsatt PkV RF01.

#### 2.1.4 Flowcytometri (FCM)

Ved hjelp av instrumentet FACS Calibur BC Flow cytometer (Becton-Dickinson, Franklin Lakes, NJ, USA) ble prøvene analysert som beskrevet i (Marie et al., 1999). Prøvene ble fiksert med 20 µl 25 % glutaraldehydløsning og vendt flere ganger for å blande løsningen grundig med kulturprøvene. De ble deretter oppbevart 30 minutter i mørket ved 4 °C. Prøvene ble hurtigfrost i flytende nitrogen (-200 °C), og oppbevart der frem til analysering.

PkV RF01 ble målt ved 10x, 100x og 1000x fortyninger med 0,2 µm filtersterilisert TE (10 mM Tris, 1 mM EDTA, pH 8). Fortyningene ble tilsatt 10 µl SYBR Green I DNA fargestoff 100x fortynnet (Life Technologies, Canada). Prøvene ble inkubert 10 minutter i et 80 °C vannbad, før de ble oppbevart i mørket ved romtemperatur frem til telling (minst 10 minutter). Ved telling av algen (He028) ble 1 ml av prøvene brukt direkte uten fortykning eller tilsetning av DNA fargestoff, da algen ble telt basert på dens egen klorofyll fluorescens. Virus prøvene ble telt på høy strømningshastighet (ca. 100 µl/minutter) i 60 sekunder, mens algen ble telt på medium strømningshastighet (ca. 50 µl/minutter) i 60 sekunder. Instrumentinnstillinger og formel brukt til omregning fra antall telte partikler på FCM til konsentrasjon av partikler per ml er beskrevet i vedlegg C.1.

## 2.2 DNA-isolering

### 2.2.1 PkV RF01

1,5 L virus lysat ble sentrifugert (Beckman Coulter Allegra 2IR Centrifuge, Indianapolis, USA) ved  $4600 \times g$  i 30 minutter ved 10 °C. For å pelletere viruset ble supernatanten ultrasentrifugert (Beckman Optima L-90K ultrasentrifuge, Indianapolis, USA) ved 25 000 o/min i 1,5 time ved 4 °C. Supernatanten ble fjernet, og 100 µl natrium-magnesium (SM) buffer (Vedlegg B.2) ble tilsatt pellet. For at pelleten skulle løse seg opp ble rørene med pellet og SM buffer oppbevart ved 4 °C over natten.

DNA-isoleringsmetoden som ble brukt for å isolere virus DNA fra konsentratet er beskrevet i Pagarete et al, 2013. 500 µl viruskonsentrat ble overført til 1,5 ml eppendorf rør, og lysis av viruspartiklene ble utført ved å inkubere viruskonsentratet ved 90 °C for 1 minutt, deretter 1 minutt på is. Videre lysis ble utført ved å tilsette 20 µl 0,5 M EDTA (pH 8, vedlegg B.3). For å ødelegge virus kapsidene ble 5 µl proteinase K, fortynnet i sdH<sub>2</sub>O (10 mg/ml) tilsatt, før røret ble inkubert ved 55 °C i 10 minutter. For å løse opp lipider i viruskapsidene ble 25 µl 10 % Natrium Dodecyl Sulfat (SDS) tilsatt og prøvene ble videre inkubert ved 55 °C i 1 time.

Virus lysatet ble rensset ved å bruke ZYMO DNA clean and concentrator<sup>TM</sup>-10 kit (Zymo Research, Irvine, CA, USA), slik som beskrevet i protokollen til produsenten (vedlegg C.2). Kvantifisering av isolert DNA ble utført ved å bruke Qubit 2.0 fluorometer (Invitrogen, Singapore) som beskrevet i instruksjonsmanualen til instrumentet (vedlegg C.3). Prøvene ble oppbevart i  $-80\text{ }^{\circ}\text{C}$  frem til bruk i PCR-reaksjoner (se avsnitt 2.3, 2.4 og 2.6.3).

### 2.2.2 He028

Fire  $\times$  1 ml av en He028 kultur i eksponentiell vekst ble tatt ut for DNA isolering. Lysis av cellene ble utført ved å bruke Blood & Cell culture DNA mini kit slik som beskrevet i QIAGEN genomic DNA Handbook 2015 (QIAGEN, Germantown, USA). Algecellene ble pelletert (Sigma 1-14, Osterode am Harz, Tyskland) ved  $1500 \times g$  i 10 minutter, og supernatanten ble fjernet. Deretter ble cellene vasket to ganger med 1xPBS (fosfatbufret saltvann, pH 7,4) og resuspendert i 0,5 ml kald ( $4\text{ }^{\circ}\text{C}$ ) 1xPBS. For å lysere cellene ble 100  $\mu\text{l}$  av buffer C1 (lysis buffer) og 300  $\mu\text{l}$  MilliQ vann tilsatt, og prøvene ble inkubert på is i 10 minutter. For å pellete de lyserte cellene ble de sentrifugert ved  $1300 \times g$  i 15 minutter og supernatanten ble fjernet. 0,25 ml kald buffer C1 og 0,75 ml MilliQ vann ble tilsatt og prøvene ble sentrifugert ved  $1300 \times g$  i 15 minutter. Videre lysis og denaturering av proteiner ble gjort ved å tilsette 1 ml buffer G2. Løsningen ble blandet godt. 25  $\mu\text{l}$  QIAGEN Proteinase K ble tilsatt, og prøvene ble inkubert ved  $50\text{ }^{\circ}\text{C}$  i 1 time.

Videre ble DNAet rensset ved å følge protokollen til kitet (QIAGEN, Germantown, USA) (vedlegg C.4). Qubit 2.0 fluorometer (Invitrogen, Singapore) ble brukt for kvantifisering av det isolerte DNAet (vedlegg C.3). Prøvene ble oppbevart i  $-80\text{ }^{\circ}\text{C}$  frem til bruk i PCR-reaksjoner (se 2.3)

## 2.3 PCR-amplifisering av Succinate dehydrogenase A (SDHA)

DNA fra PkV RF01 og He028 ble brukt til å optimalisere en PCR-reaksjon med spesifikke primere for Succinate dehydrogenase A genet (*sdhA*);

<b>Primer</b>	<b>Sekvens</b>
primer F	5'- ATGTGCCGAGAAGCTCCTAA - 3'
primer R	5'-CTGCACAGGCTGTTCGATAA – 3'

Det isolerte DNAet til PkV RF01 ble brukt for å optimalisere PCR-reaksjonen, og det isolerte DNAet til He028 ble brukt som kontroll på at primerparet er spesifikt for det virale *sdhA* genet.

To fortyninger av templat ble brukt, 1:1 og 1:10, hvor 2 µl templat ble tilsatt til 48 µl mastermiks (Tabell 2.1). I tillegg ble det laget en negativ kontroll med 2 µl sdH<sub>2</sub>O som templat. For å finne optimal hybridiseringstemperatur til primerne ble prøvene amplifisert ved å bruke et gradientprogram (55-60 °C, tabell 2.2) på PCR-maskinen (Thermo Fisher Veriti Thermal Cycler, Massachusetts, USA). PCR-produktene ble kontrollert ved å analysere 3 µl av reaksjonsblandingen på en 1,5 % agarosegel.

Det ble laget en 1,5 % agarose gel ved å blande 0,8 g agarose (SeaKem LE Agarose, USA) med 60 ml 1xTAE buffer (40 mM tris, 20 mM eddiksyre, 2,5 mM EDTA). Blandingen ble varmet i en mikrobølgeovn til agarosen var løst opp. Deretter ble 2 µl 10 000X GelRed (Biotium, 41003, USA) farge tilsatt. Når løsningen hadde en temperatur ca. 50 °C ble den overført til et gelkar med 32 brønner, og inkubert ved romtemperatur i ca. 30 minutter. Gelen ble overført til et fryst elektroforesekar og karet ble fylt med ca. 250 ml 1x Tris-acetat-EDTA buffer (TAE buffer). 3 µl MassRuler DNA Ladder mix (Thermo Fisher Scientific, SM0403, Litauen) ble tilsatt til brønnene på hver side av gelen. På parafilm ble 3 µl PCR-produkt (PkV RF01 og He028) samt negativ kontroll godt blandet med 1 dråpe MassRuler LD (Thermo Fisher Scientific, SM0403, Litauen) før de ble tilsatt brønnene på gelen. Elektroforesen ble utført ved 200 V i ca. 30 minutter, og deretter ble gelen plassert i BIO RAD Molecular Imager (ChemiDOC XRS, Hertfordshire, England) og et bilde ble tatt ved å bruke Image Lab software.

*Tabell 2.1: PCR-reaksjonsmiks for amplifisering av sdhA genet*

<b>Innhold i reaksjonsmiks</b>	<b>Mengde (µl) per reaksjon</b>
<b>sdH<sub>2</sub>O</b>	31,75
<b>10x buffer</b>	5
<b>10 µM dNTP</b>	4
<b>10 µM forward primer</b>	2,5
<b>10 µM reverse primer</b>	2,5
<b>Templat (100 ng/ml)</b>	2
<b>BSA (2 %)</b>	2
<b>Taq</b>	0,25
<b>Totalt</b>	50

Tabell 2.2: PCR-program med gradient for amplifisering av *sdhA* genet

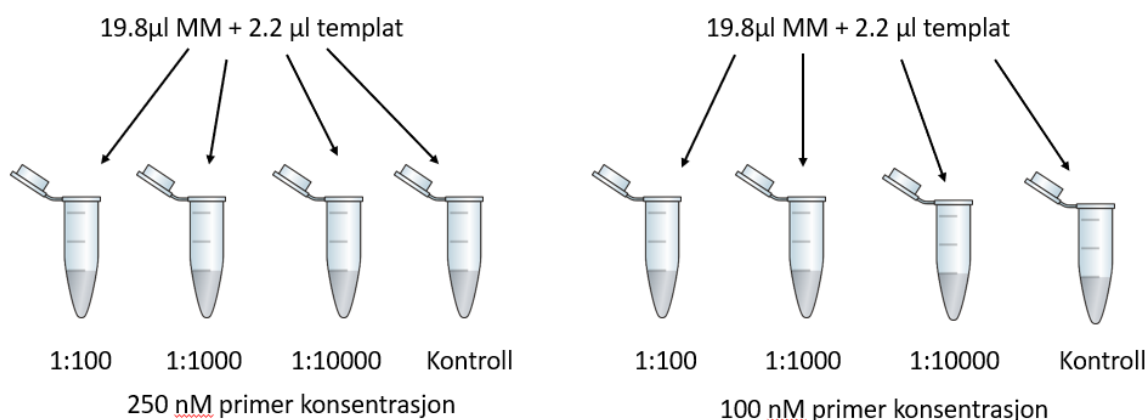
Steg i PCR	Temperatur (°C)	Tid	Antall sykluser
<b>Denaturering</b>	94	5 minutter	1
	94	30 sekunder	
<b>Hybridisering</b>	55-60	30 sekunder	35
<b>Polymerisering</b>	72	30 sekunder	
	72	12 minutter	1
<b>Hold</b>	4	∞	

#### 2.4 Digital droplet PCR (ddPCR) av *sdhA* genet

Spesifikke primere for *sdhA* (se avsnitt 2.3) ble brukt for å optimalisere ddPCR-reaksjonen for dette primersettet, og samtidig få en absolutt kvantifisering av antall *sdhA* kopier i prøven. Det ble testet to ulike primer konsentrasjoner; 250 nM og 100 nM. Samtidig ble det laget fortynningsrekke av templatet. 1:10 fortynningen ble laget ved å tilsette 2 µl virus DNA (100 ng/ml, se 2.2.1) til 18 µl TRIS buffer, for deretter å lage 1:100, 1:1000 og 1:10000 fortynning. For hver reaksjon ble 19,8 µl av mastermiks (Tabell 2.3) og 2,2 µl templat (fortynningsrekke) (Figur 2.2) overført til ett Eppendorfrør og blandet godt. Deretter ble rørene oppbevart ved romtemperatur frem til videre arbeid.

Tabell 2.3: PCR-reaksjonsmiks for ddPCR.

Primer konsentrasjon (nM)	Innhold i reaksjonsmiks	Mengde (µl)
<b>250</b>	2x QX200 ddPCR EvaGreen Supermix	11
	10 mM forward primer	0,55
	10 mM reverse primer	0,55
	sdH <sub>2</sub> O	7,7
	<b>Totalt</b>	<b>19,8</b>
<b>100</b>	2x QX200 ddPCR EvaGreen Supermix	11
	10 mM forward primer	0,22
	10 mM reverse primer	0,22
	sdH <sub>2</sub> O	8,36
	<b>Totalt</b>	<b>19,8</b>



Figur 2.2: Viser oppsettet for ddPCR. Det ble tilsatt 19,8 µl mastermix (MM) og 2,2 µl templat til hvert rør. Kontrollen har 2.2 µl H<sub>2</sub>O som templat.

20 µl av hver prøve ble overført til prøvebrønnene (midterste raden) på DG8 prøvebrettet. Deretter ble 70 µl av droplet generasjons (DG) oljen tilsatt i oljebrønnene (nederste raden). Pakningen ble festet over prøvebrettet og plassert i instrumentet (QX200 droplet generator, BIO-RAD, Hertfordshire, USA), hvor små dråper (droplets) ble generert. Etter 5 minutter ble prøvebrettet tatt ut og pakningen ble fjernet. 40 µl av dråpesuspensjonen (øverste rad) ble pipetert over i en 96-brønns PCR-plate. PCR-platen ble forseglet med folie (PX1 PCR-plate sealer, BIO-RAD, Hertfordshire, USA), og kjørt i PCR-maskinen (Bio-Rad C1000, USA) ved å bruke et program som er beskrevet i tabell 2.4. Kvantifiseringen av antall kopier i hver prøve ble utført ved å bruke QX200 Droplet reader (Bio-Rad, Hertfordshire, USA).

Tabell 2.4: ddPCR-program for amplifisering av droplets (SDHA)

Steg i PCR	Temperatur °C	Tid	Ramp rate	Antall sykluser
<b>Enzym aktivering</b>	95	5 minutter	2 °C/sekunder	1
<b>Denaturering</b>	95	30 sekunder		40
<b>Hybridisering</b>	60	1 minutter		40
<b>Signal stabilisering</b>	4	5 minutter		1
	90	5 minutter		1
<b>Hold</b>	4	∞	1	



## 2.5 Transkripsjonsforsøk – SDHA

Transkripsjon er første steg i uttrykkelsen av et gen, og målet med transkripsjon er å lage en RNA-kopi av DNA sekvensen. Ved å bruke Power SYBR Green Cells to- $C_T$  kit (Ambion, Texas, USA) kan prøver bli brukt direkte til isolering av mRNA. Ved å bruke dette kitet blir mRNA brukt som templat i en reaksjon hvor cDNA blir syntetisert av enzymet revers transkriptase. Isolering av mRNA baserer seg på lysis av cellene etterfulgt av nedbryting av DNA, og hvor man tilslutt står igjen med cDNA som er et mål på hvor mye mRNA man har klart å isolere. Ved å bruke PCR og gelelektroforese kan man undersøke om det aktuelle genet blir transkribert eller ikke.

### 2.5.1 Vasking av celler og isolering av RNA

Ved hvert tidspunkt i infeksjonsforsøket (se avsnitt 2.1.3) ble det tatt ut 0,5 ml av den infisert kultur og kontrollkulturen. Vasking og lysis av cellene ble utført ved å følge protokollen til kitet (Ambion, Texas, USA). For å vaske cellene ble de først sentrifugert ved  $3000 \times g$  i 10 minutter, deretter ble pelleten resuspendert i 50  $\mu\text{l}$  kald 1xPBS. De ble så sentrifugert ved  $3000 \times g$  i 10 minutter. Det ble laget to forskjellige konsentrasjoner av de vaskede cellene ved å tilsette 50  $\mu\text{l}$  ( $10^4$  celler  $\text{ml}^{-1}$ ) og 5  $\mu\text{l}$  ( $10^5$  celler  $\text{ml}^{-1}$ ) 1xPBS før de ble lagt på is.

For lysis av cellene og nedbrytning av DNA ble 49,5  $\mu\text{l}$  lysisløsning og 0,5  $\mu\text{l}$  DNase I tilsatt til 5  $\mu\text{l}$  cellesuspensjon. Ved bruk av en pipette ble løsningen forsiktig blandet, for deretter å bli inkubert. For å undersøke om temperatur og inkuberingstid kunne påvirke isoleringen av mRNA og nedbrytingen av DNA ble det valgt to ulike tider og temperatur. To prøver ble inkubert ved romtemperatur, hvor den ene ble inkubert i 5 minutter og den andre i 9 minutter. To andre prøver ble inkubert ved samme tider, men ved 25 °C. Stoppløsningen ble laget ved å tilsette 1  $\mu\text{l}$  kontroll mRNA til 5  $\mu\text{l}$  stoppløsning, for deretter å tilsette 6  $\mu\text{l}$  løsning til cellene. Løsningen og cellene ble blandet godt før de ble inkubert i 2 minutter ved romtemperatur.

Prøvene ble oppbevart ved  $-80$  °C frem til videre bruk.

### 2.5.2 Revers transkripsjon (RT) PCR

For syntesen av mRNA til cDNA ble det kjørt en RT PCR-reaksjon med lysatet (avsnitt 2.5.1) som templat. 10  $\mu\text{l}$  av lysatet ble overført til PCR-rør med 40  $\mu\text{l}$  RT reaksjonsmiks (Tabell 2.5). Negativ kontroll ble laget med 10  $\mu\text{l}$  H<sub>2</sub>O som templat. PCR ble utført i PCR-maskinen (Thermo Fisher Veriti Thermal Cycler, Massachusetts, USA) etter temperatursyklusen beskrevet i tabell 2.6. Deretter ble prøvene oppbevart ved  $-20$  °C frem til videre bruk.

Tabell 2.5: RT-PCR-reaksjonsmik

Amplifisert gen	Innhold i reaksjonsmik	Mengde (µl) per reaksjon
<i>sdhA</i>	2x SYBR RT buffer	25
	20x RT enzyme mix	2,5
	H <sub>2</sub> O	12,5
	Totalt	40

Tabell 2.6: RT-PCR-program

Steg	Temperatur (°C)	Tid	Sykluser
<b>Revers transkripsjon</b>	37	60 minutter	1
<b>RT inaktivering</b>	95	5 minutter	1
<b>Hold</b>	4	∞	

### 2.5.3 Verifisering av transkripsjon og nedbrytingen av DNA

Protokollen ble optimalisert ved hjelp av major capsid protein (MCP), ettersom man vet at viruset har *mcp* gen, og at dette blir uttrykt ved produksjon av viruspartikler. Følgende MCP spesifikke primere ble benyttet;

Primer	Sekvens
PkV-RF01-MCP F	5'- GATGAACCTTGCCCACAACACT- 3'
PkV-RF01-MCP R	5'- GTGCATGGTACGTTTTTCGTG- 3'

For å undersøke om transkripsjonen av *sdhA* gen var vellykket, ble genet forsøkt amplifisert med SDHA primere (primerne er vist i avsnitt 2.3). I den første PCR-reaksjonen ble lysatet som inneholdt mRNA benyttet som templat for å kontrollere at alt DNAet var brutt ned etter DNase I trinnet. Som kontroll på at genet var transkribert ble produktet fra RT-PCR benyttet som templat. For PCR-reaksjonen med MCP spesifikke primere ble HotStarTaq master mix (QIAGEN, Germantown, USA) benyttet (Vedlegg C.4, Tabell 9.3), mens ved PCR-reaksjonen med SDHA spesifikke primere ble reaksjonsmiksen laget som beskrevet i tabell 2.1. Genene ble amplifisert ved bruk av PCR-programmet beskrevet i vedlegg C.4, tabell 9.4. Det ble benyttet 1,5 % agarose gelelektroforese som beskrevet i avsnitt 2.3 til å evaluere størrelsen til de amplifiserte PCR-produktene.

## 2.6 Kloning av virusgener til PkV RF01

Kloning er prosessen der et spesifikt gen blir satt inn en kloningsvektor, og introdusert til en passende organisme hvor det klonede genet sammen med vektoren blir replikert (Griffiths et al., 1999). Når målet er å få ekspresjon av det klonede genet brukes ofte plasmid som kloningsvektor, og som vertscelle er det vanlig å bruke en stamme av *E. coli* (Rosano and Ceccarelli, 2014).

Ved å følge instruksjonsmanualen til Gibson Assembly kloning kit (New England BioLabs, England) ble de fem målgenene l-laktat dehydrogenase (L-LDH), prolin dehydrogenase (PDH), glutaryl-CoA dehydrogenase (GCDH), DNA polymerase X (pol X) og glutamat dehydrogenase (GDH) inkorporert i kloningsvektoren pET-21a. Deretter ble *E. coli* NEB 5-alfa kompetente celler brukt som vertscelle for det rekombinante plasmidet.

### 2.6.1 Bakteriestammer og plasmid

#### 2.6.1.1 *E. coli* NEB 5-alfa kompetente celler

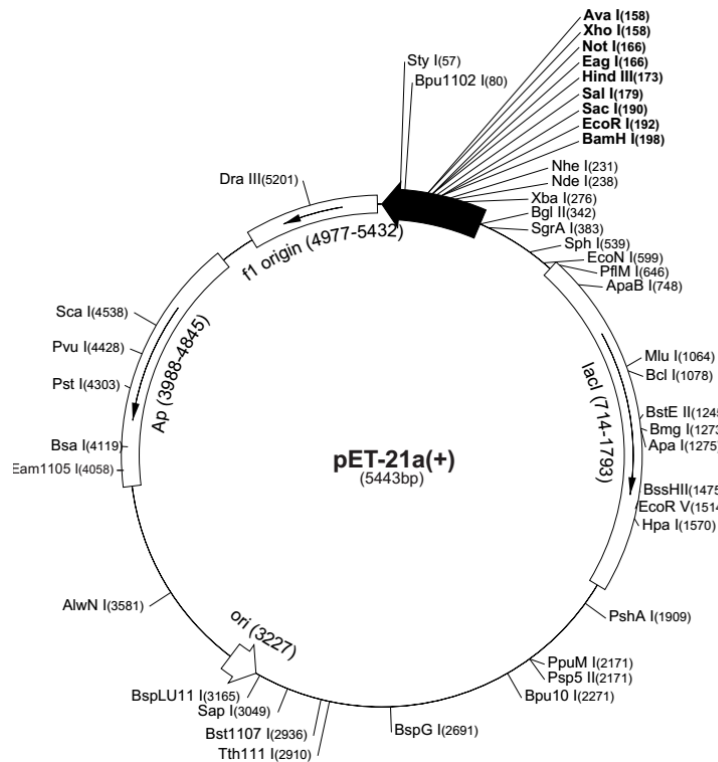
Genotype: fhuA2  $\Delta$ (argF-lacZ)U169 phoA glnV44  $\Phi$ 80 $\Delta$  (lacZ)M15 gyrA96 recA1 relA1 end. NEB 5-alfa kompetent *E. coli* er en K12 stamme og er ett derivat av DH5 $\alpha$  celler. Stammen har en endA1 mutasjon i genet som inaktiverer endonukleaser, for å unngå at plasmid DNA nedbrytes. Cellene har ikke genet for T7-RNA-polymerase, som bidrar til å øke stabiliteten av plasmidene i vertscellene. Denne stammen ble brukt til oppformering av rekombinant plasmid. Kjemisk kompetente NEB 5-alfa *E. coli* celler ble kjøpt fra New England Biolabs (Massachusetts, USA).

#### 2.6.1.2 *E. coli* BL21-GOLD (DE3) celler

Genotype: *E. coli* B F<sup>-</sup> ompT hsdS(rB – mB –) dcm<sup>+</sup> Tetr gal  $\lambda$ (DE3) endA Hte. BL21-GOLD (DE3) ble brukt til ekspresjon av de 5 målgenene. Stammen er et derivat fra *E. coli* B, og mangler både Lon protease og OmpT protease, som kan bryte ned proteiner under rensing. Stammen har genet for T7-RNA-polymerase, som er under kontroll av *lacUV5* promoter. For å inducere ekspresjon av T7-RNA-polymerase blir Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) tilsatt til kulturen. Kjemisk kompetente BL21-GOLD (DE3) celler ble kjøpt fra Agilent Technologies (California, USA).

### 2.6.1.3 pET-21a vektor

pET-21a vektoren er et plasmid på 5443 basepar og koder for ampisillin resistens. Vektoren har en bakteriofag T7-promotor for T7-RNA-polymerase som gir høy-ekspressjon av klonede gener. Vektoren ble brukt til kloning og ekspressjon av de 5 målgenene. Vektoren ble kjøpt fra Novagen.



Figur 2.3: Viser pET-21-a vektor.

### 2.6.2 Primere

Spesifikke primere for de fem målgenene og deres tilhørende pET-21a vektor ble designet som anbefalt i protokollen som fulgte med kitet. For at målgenet skal bli riktig plassert i vektoren er det anbefalt at det er et 15-25 nt sekvensoverlapp mellom gensekvens og vektorsekvens og en smeltetemperatur høyere enn 48 °C. Når sekvensoverlappen blir plassert på 5' enden av primeren til vektoren, er denne homolog til 5' enden av primeren til målgenet. Primerne ble designet med sekvensoverlapp ved å bruke Gibson Assembly Primer Design Tool (SGI-DNA, <https://tools.sgidna.com/gibson-assembly-primers.html>). Steg 1 er å sette inn vektorsekvensen i FASTA format. Videre bestemmer man hvor genet skal bli satt inn ved å manuelt finne HIS tag på vektor sekvensen, og velger at man vil sette inn gensekvensen foran denne. Deretter setter man inn gensekvensen, og man får ett forslag for primere og smeltetemperatur til hver primer. For å få en oversikt over genet sin plassering i vektoren ble programmet CLC genomic workbench 9.5.2 (Qiagen Bioinformatics, Århus, Danmark) brukt.

Smeltetemperaturen ( $T_m$ ) til hver primer ble sjekket ved å benytte en  $T_m$  kalkulator (ThermoFisher, <https://www.thermofisher.com/no/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html>). Denne kalkulatoren beregner  $T_m$  basert på primerpar sekvensen, primer konsentrasjon og hvilken DNA polymerase som skal benyttes. Det ble designet primere med en  $T_m$  fra 53-56 °C for samtlige gener slik at de kunne PCR-amplifiseres ved bruk av samme hybridiseringstemperatur for dermed å effektivere forsøkene.  $T_m$  for vektorene var den samme, da det bare er sekvensoverlapp som varierer avhengig av hvilket gen vektoren tilhører. Primerne ble bestilt fra Biomers.net (Donau, Tyskland) (Tabell 2.7 og 2.8). Hver primer ble tilsatt oppgitt mengde sdH<sub>2</sub>O som anbefalt av produsent og plassert i kjøleskap over natten (4 °C) slik at de skulle løse seg opp. Deretter ble de oppbevart i -20 °C frem til bruk.

Tabell 2.7: Oversikt over sekvens og  $T_m$  for primere benyttet til amplifisering av de fem målgenene. Små bokstaver representerer delen av primer-sekvens som hybridiserer, og store bokstaver representerer sekvensoverlapp. Sekvensene vises i retning 5'-3'.

Målgen		Primer sekvens	$T_m$ (°C)
<i>l-ldh</i>	Forward	atgtatcaaaaaatgattagacctg	55,8
	Reverse	CTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGT GGTGGTGattaccaataaatgacat	53,6
<i>pdh</i>	Forward	atgttgagatacattggtgg	54,5
	Reverse	CTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGT GGTGGTgaataaacatatatttactgtatctatgttt	53,9
<i>gcdh</i>	Forward	atgattagagcaattttaataaaaat	56,2
	Reverse	CTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGT GGTGGTGaaatgcactaatattagtaattgac	56,7
<i>pol X</i>	Forward	atgacaactgaaaaaaaaaatatcaat	54,5
	Reverse	CTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGT GGTGGTGattttgattattttaattttattcagtctca	54,6
<i>gdh</i>	Forward	atgtcaaaaataactgatttaattaaaaatc	55,2
	Reverse	CTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGT GGTGGTGaaatcctcttcttaagtatacttc	53,0

<b><i>sdhA</i></b>	Forward	atgaaaatTTTtacaagaaaatttca	53,9
	Reverse	CTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGT GGTGGTGgtaaactcttttgctaaag	53,5
<b><i>sdhB</i></b>	Forward	atgaatatttttagaagaactttatca	53,2
	Reverse	CTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGT GGTGGTGatgtaactcttgtaatttttaattt	52,8

Tabell 2.8: Oversikt over sekvens og  $T_m$  for primere benyttet til amplifisering av pET-21a vektor. Små bokstaver representerer delen av primer-sekvens som hybridiserer, og store bokstaver representerer sekvensoverlapp. Sekvensene vises i retning 5'-3'.

<b>Vektor tilhørende:</b>		<b>Primer sekvens</b>	<b><math>T_m</math> (°C)</b>
<b><i>l-ldh</i></b>	Forward	ATcaccaccaccaccactg	64,8
	Reverse	GTAAATAATTAATAACAGGTCTAATCATTTTTTG ATACATatgtatatctccttcttaaagtt	53,9
<b><i>pdh</i></b>	Forward	TTcaccaccaccaccactg	64,8
	Reverse	GAACTTGATTTTATATTTGATCCACCAATGTATCTC AACATatgtatatctccttcttaaagtt	53,9
<b><i>gcdh</i></b>	Forward	TTcaccaccaccaccactg	64,8
	Reverse	GATGATTAATTTTATTTTATTAATAAATTGCTCTA ATCATatgtatatctccttcttaaagtt	53,9
<b><i>pol X</i></b>	Forward	ATcaccaccaccaccactg	64,8
	Reverse	CTCTTGGTTTTTTTATTGATATTTTTTTTCAGTTG TCATatgtatatctccttcttaa	53,9
<b><i>gdh</i></b>	Forward	TTcaccaccaccaccactg	64,8
	Reverse	TATCAATTTGATTTTAAATTAATCAGTTATTTTT GACATatgtatatctccttcttaa	53,9
<b><i>sdhA</i></b>	Forward	caccaccaccaccactg	64,8
	Reverse	CTATAATTTTTTTGTGAAATTTTCTTGTAATAAATT TTCATatgtatatctccttcttaataatgat	53,7
<b><i>sdhB</i></b>	Forward	caccaccaccaccactg	64,8
	Reverse	TTTTTACAGTTGTTGATAAAGTTCTTCTAAAAATA TTCATatgtatatctccttcttaaagtttaa	53,5

### 2.6.3 Amplifisering av målgener og vektor

Primerne (Tabell 2.7 og 2.8) ble brukt for å amplifisere opp målgen og vektor. For de fem målgenene ble det laget en 50 µl Phusion reaksjonsmiks (New England Biolabs, England) (Tabell 2.9), hvor 2 µl virus DNA ble brukt som templat (35,4 ng/ml, se avsnitt 2.2.1). Som negativ kontroll ble det brukt 2 µl sdH<sub>2</sub>O. Vektorene ble amplifisert ved å bruke Q5 High-Fidelity mastermiks (New England Biolabs, England), hvor det ble tilsatt 2 µl pET-21a (500 ng/ml, fortynnet 1:50 i sdH<sub>2</sub>O) som templat til 25 µl reaksjonsmiks (Tabell 2.11). Som positiv kontroll ble plasmidet AF0809\_pET21a (2 µl) brukt, hvor et gen (AF0809) fra den sulfatreduserende arkéen *Archaeoglobus fulgidus* tidligere var blitt klonet og sekvensert av ingeniør Anita-Elin Fedøy ved Dyphavbiologigruppen ved BIO. Som negativ kontroll ble det brukt 2 µl sdH<sub>2</sub>O. PCR ble utført i PCR-maskiner (Thermo Fisher Veriti Thermal Cycler, Massachusetts, USA) etter temperatursyklusen beskrevet i tabell 2.10. SDHA og SDHB ble i tillegg forsøkt amplifisert ved hjelp av gradient PCR. For vektor til SDHA ble denne forsøkt amplifisert ved 58-64 °C, og målgenet *sdhB* ble forsøkt amplifisert ved 55-59 °C.

For å sjekke størrelsen på PCR-produktene, ble prøvene til målgene analysert ved hjelp av 1,5 % agarose gelelektroforese som forklart i avsnitt 2.3. PCR-amplifisert vektorer ble analysert på en 1 % agarose gel laget ved å blande 0,4 g agarose med 40 ml 1xTAE. Løsningen ble varmet opp i mikrobølgeovn frem til agarosen var løst opp. Deretter ble 4 µl 10 000x GelGreen (Biotium, USA) tilsatt og løsningen ble overført til et gelkar med brønner (BIO-RAD, Hertfordshire, USA) som sto i romtemperatur i ~30 minutter. Deretter ble karet fylt med 1xTAE buffer. 5 µl GeneRuler 100 bp Plus DNA ladder (Thermo Fisher Scientific, Litauen) ble tilsatt til brønnene på hver side av gelen. 5 µl PCR-produkt, samt positiv og negativ kontroll ble blandet med 1 µl MassRuler LD (Thermo Fisher Scientific, Litauen) før de ble tilsatt brønnene på gelen. Elektroforesen ble utført ved 50 V i 35 minutter (BIO-RAD powerpac basic, Hertfordshire, USA). Deretter ble gelen plassert i G-box gel doc (Syngene, Cambridge, England) og et bilde ble tatt ved å bruke Genesnap software (Syngene, Cambridge, England).

Tabell 2.9: PCR-reaksjonsmiks for amplifisering av gener og vektor.

Amplifisering av	Innhold i reaksjonsmiks	Mengde ( $\mu$ l)
<i>l-ldh</i> <i>pdh</i> <i>gcdh</i> <i>pol X</i> <i>gdh</i> <i>sdhA</i> <i>sdhB</i>	5x phusion HF	5
	10 $\mu$ M dNTP	0,5
	10 $\mu$ M forward primer	1,25
	10 $\mu$ M reverse primer	1,25
	34,5 ng/ml templat	2
	Phusion DNA polymerase	0,25
	DMSO	0,75
	sdH <sub>2</sub> O	Til volum
	<b>Totalt</b>	<b>50</b>
<b>pET-21-a</b>	Q5 master mix	12,5
	10 $\mu$ M forward primer	1,25
	10 $\mu$ M reverse primer	1,25
	500 ng/ml templat	2
	sdH <sub>2</sub> O	Til volum
	<b>Totalt</b>	<b>25</b>

Tabell 2.10: PCR-program for amplifisering av gen og pET-21a.

Amplifisering av	Tid	Temperatur ( $^{\circ}$ C)	Antall Sykluser
<i>l-ldh</i> <i>pdh</i> <i>gcdh</i> <i>pol X</i> <i>gdh</i> <i>sdhA</i> <i>sdhB</i>	30 sekunder	98	1x
	10 sekunder	98	
	30 sekunder	56	35x
	30 sekunder	72	
	10 minutter	72	1x
	$\infty$	4	
<b>pET-21-a</b>	30 sekunder	98	1x
	10 sekunder	98	
	30 sekunder	60	35x
	2,5 minutter	72	
	2 minutter	72	1x
	$\infty$	4	



#### 2.6.4 Overføring av målgener til pET-21a vektor

De amplifiserte produktene for målgenene og vektorene ble først behandlet med *DpnI* (New England Biolabs, England) som er et restriksjonsenzym som kutter ikke amplifisert templat. Enzymet kutter ved sekvensen GATC når A er metylert, og siden PCR-produkt ikke har metylerte A-er vil bare original templat bli ødelagt. Det ble laget en reaksjonsmiks for hvert av PCR-produktene (Tabell 2.11). Deretter ble rørene inkubert i PCR-maskin (Veriti Thermal Cycler, Thermo Fisher) ved 37 °C i 30 minutter for å kutte original templat og 80 °C i 15 minutter for å inaktivere *DpnI* (Tabell 2.12).

Tabell 2.11: Reaksjonsmiks og inkuberingstid for *DpnI* behandling av PCR-produktene

<b>Reaksjonsmiks</b>	5 µl PCR-produkt 1 µl CutSmart 1 µl <i>DpnI</i> 3 µl H <sub>2</sub> O
<b>Inkubering</b>	37 °C i 30 minutter 80 °C 15 minutter

De *DpnI* behandlede PCR-produktene ble videre brukt til overføringen av målgenene til vektorene. Dette ble gjort ved å følge protokollen beskrevet i Gibson Assembly Cloning kit Manual (New England BioLabs, England). Reaksjonsmiks og inkuberingstid for overføringen av målgen til pET-21a er beskrevet i tabell 2.12. Mastermiksen inneholder en eksonuklease, DNA-polymerase og DNA-ligase. Målgenet blir satt inn i vektoren ved at eksonukleasen tygger tilbake 5' enden på målgenet og vektor, deretter vil DNA polymerasen forlenge 3' enden og tilslutt vil DNA ligase lime sammen DNA tråden, og genet har blitt inkorporert i vektoren.

Tabell 2.12: Reaksjonsmiks og inkuberingstid for overføringen av målgen til pET-21a

<b>Reaksjonsmiks</b>	10 µl Gibson assembly master mix (2X) 7 µl H <sub>2</sub> O 2 µl <i>DpnI</i> behandlet målgenprodukt 2 µl <i>DpnI</i> behandlet vektorprodukt
<b>Inkubering</b>	50 °C i 40 minutter

## 2.7 Transformasjon NEB 5-alpha kompetente celler

Som beskrevet i manualen (New England BioLabs, England) ble genet transformert inn i *E. Coli* NEB 5-alfa kompetente celler ved at 2 µl av rekombinant plasmidet med ligert målgen ble overført til røret med de kompetente cellene. Cellene ble forsiktig blandet ved å pipettere opp og ned før røret ble plassert på is i 30 minutter. Deretter ble cellene utsatt for et varmesjokk ved 42 °C i 30 sekunder, før de ble plassert på is i 2 minutter. Når cellene blir utsatt for et kortvarig varmesjokk vil de ha mulighet til å ta opp det rekombinante plasmidet. For å øke transformasjonseffektiviteten ble 950 µl romtemperert super optimal buljong med glukose (SOC) vekstmedium (New England BioLabs, England) tilsatt før cellene ble inkubert ved 30 °C i 60 minutter. 100 µl av cellene ble deretter spredd på LA skåler (med ampicillin, vedlegg B.4), og inkubert ved 37 °C over natten.

To kolonier fra hver kloningsreaksjon ble analysert ved hjelp av PCR for å bekrefte at det korrekte rekombinante plasmidet var transformert inn i de kompetente cellene (se avsnitt 2.8).

## 2.8 Koloni PCR

Ved bruk av en pipettespiss ble små mengder av hver koloni på LA-skålene (se avsnitt 2.7) overført til PCR-rør ved å plassere spissen oppi røret og blande godt med 24,5 µl reaksjonsmiksblanding (Tabell 2.13). Flankerende pET-21a primere ble brukt for å undersøke om korrekt rekombinant plasmid var transformert:

<b>Primer</b>	<b>Sekvens</b>
pET-21a forward	5'-acgactcactataggggaatt- 3'
pET-21a reverse	5'-gctagttattgctcagcgg- 3'

PCR ble utført i PCR-maskinen (Thermo Fisher Veriti Thermal Cycler) etter programmet beskrevet i tabell 2.14. Deretter ble størrelsen på PCR-produktene analysert ved 1 % agarose gelelektroforese som beskrevet i avsnitt 2.6.3. Kolonier med korrekt størrelse av målgen og som dermed inneholdt korrekt rekombinant plasmid ble vurdert som positive og benyttet som inokulum for kulturer ved påfølgende plasmid-rensing.

Tabell 2.13: PCR-reaksjonsmiks for amplifisering av plasmidet med klonet gen

Amplifisering av	Innhold i reaksjonsmiks	Mengde ( $\mu$ l)
<b>Rekombinant plasmid etter transformasjon</b>	HotStarTaq mastermiks	12,5
	Primer pET-21-a forward	0,5
	Primer pET-21-a reverse	0,5
	H <sub>2</sub> O	11,5
	Totalt	24,5

Tabell 2.14: PCR-program for amplifisering av plasmid med klonet gen

Steg i PCR	Temperatur ( $^{\circ}$ C)	Tid	Antall sykluser
<b>Denaturering</b>	94	30 sekunder	1
	94	20 sekunder	
<b>Hybridisering</b>	60	20 sekunder	25
<b>Polymerisering</b>	68	45 sekunder	
	68	5 minutter	1
<b>Hold</b>	4	$\infty$	

## 2.9 Transformasjon BL21-GOLD(DE3) celler

Før de rekombinante plasmidene kunne bli transformert inn i ekspresjonsceller, ble plasmidet renset. Kulturer av positive kolonier verifisert med koloni-PCR, ble inokulert i Falconrør med 5 ml LB medium og 5  $\mu$ l ampisillin (100  $\mu$ g/ml) og inkubert ved 37  $^{\circ}$ C over natten. Deretter ble plasmidet renset ved å følge protokollen til Monarch Plasmid miniprep kit (New England BioLabs, vedlegg C.4). DNA-utbyttet ble målt med Quantus fluorometer (Promega, Wisconsin, USA).

Ved å følge protokollen som er basert på prinsippet presentert i (Pope and Kent, 1996) ble det rensede plasmidet transformert inn i BL21-GOLD(DE3) celler. 1  $\mu$ l renset plasmid ble overført til 20  $\mu$ l BL21-GOLD(DE3) celler og inkubert i 20 minutter på is, for deretter å bli spredd ut på LA skåler (med ampisillin) og inkubert ved 37  $^{\circ}$ C over natten. Deretter ble resultatet analysert ved koloni PCR som beskrevet i avsnitt 2.8.

## 2.10 Proteinrensing - småskala

### 2.10.1 Ekspresjon

Det ble laget pre-kulturer for celler hvor rekombinant plasmid med korrekt størrelse på målgen var bekreftet. Disse ble dyrket i Falconrør med 5 ml LB-medium (vedlegg B.4) og 5  $\mu$ l ampicillin (100 mg/ml) og inkubert ved 37 °C over natten. Neste dag ble pre-kulturen brukt til å lage 10 ml kulturer. 250  $\mu$ l pre-kultur ble overført til et Falconrør med 10 ml LB medium og 10  $\mu$ l ampicillin. Kulturene ble inkubert ved 37 °C og 250 rpm (Infors HT Incubator Shaker, Bottmingen, Sveits). Cellene ble inkubert til optisk tetthet ( $OD_{600nm}$ ) var over 0,4. Deretter ble temperaturen skrudd ned til 30 °C og *lac* promotor ble induert ved å tilsette 1 mM IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside). Dette inducerer T7 RNA polymerase, som igjen vil starte transkripsjonen av målgenet på klonet i pET-21a-plasmidet. Etter induksjon ble kulturene dyrket videre i 3 timer. Kulturene ble høstet ved sentrifugering i 12 minutter ved  $4750 \times g$  (Beckman Coulter Allegra centrifuge, Indianapolis, USA), før supernatanten ble fjernet og pelleten ble fryst ved -20 °C.

Isoelektriske punkt (pI) og størrelse i kDa til proteinene ble predikert ved bruk av ExPasy (<https://web.expasy.org/cgi-bin/protparam/protparam>). Basert på proteinenes isoelektriske punkt (pI, Tabell 2.15), ble alle løsningene til videre arbeid laget med en pH som var ca. 15-20 ganger surere enn pI til hvert protein. Dette var nødvendig for at proteinet skulle være ladet, og dermed frastøte hverandre og hindre aggregering. Lysis buffer ble laget med to ulike pH verdier da GDH har et lavere pI enn de andre proteinene (Tabell 2.16). Lysisbufferen ble tilsatt lysozym, som er et enzym som vil ødelegge celleveggen til *E. Coli*.

Tabell 2.15: Teoretisk pI og størrelse i kDa for de fem proteinene.

Navn	Teoretisk pI	kDa
L-laktat dehydrogenase	8,71	44,18
Prolin dehydrogenase	8,61	32,80
Glutaryl-CoA dehydrogenase	8,81	46,49
DNA polymerase X	9,19	46,92
Glutamat dehydrogenase	7,98	47,03

Tabell 2.16: Oversikt over bufferløsninger som ble brukt under småskala rensing av proteiner. Lysisbuffer med pH 6 ble brukt for GDH, og pH 7,5 ble benyttet for de fire andre proteinene.

Buffer	NaH <sub>2</sub> PO <sub>4</sub>	NaCl	Imidazol	Lysozym	pH	
<b>Lysis buffer</b>	50 mM	200 mM	n/a	0,25 mg/ml	6	7,5
<b>Buffer A (ekvilibrering og vask)</b>	50 mM	300 mM	5 mM	n/a	7,5	
<b>Buffer B (Eluering)</b>	50 mM	300 mM	500 mM	n/a	7,5	

For å lysere cellene ble pellet tint på is, og 1 ml lysis buffer med tilpasset pH (Tabell 2.16) ble tilsatt før løsningen ble inkubert i 30 minutter på is. For å være sikker på fullstendig lysis ble cellene utsatt for sonikering, som er en teknikk hvor man utsetter cellene for lydenergi høyere enn 20 kilohertz, slik at cellemembraner blir ødelagt. Ved bruk av denne teknikken er det viktig at cellene er plassert på is både før og etter sonikeringen. Samtidig er det viktig å velge en amplitude som ikke er for høy. Dette fordi energien som dannes vil avgi mye varme og blir prøvene utsatt for høy varme kan proteinet denatureres. Av den grunn ble cellene utsatt for lydenergi i 10 sekunder ved amplitude 20-30 % (Sonics Vibra cell, Connecticut, USA) for deretter å plasseres på is i 15 sekunder. Dette ble gjentatt fire ganger.

For å finne forholdene som ga høyest ekspresjon ble det kjørt flere runder med ekspresjon med ulike typer medium, IPTG konsentrasjoner og inkubasjonstider. IPTG inducere den virale T7-RNA-polymerasen, som vil gjør at genet blir uttrykt da genet er introdusert bak *lacI*-promotoren i genomet, og om man da øker konsentrasjoner og inkubasjonstider vil dette påvirke ekspresjonen av genet.

Først ble det gjennomført ekspresjon med LB-medium (tilsatt ampicillin 100 µg/ml) og med IPTG konsentrasjon på 1 mM. Når OD<sub>600nm</sub> til kulturen var 0,64 ble IPTG tilsatt og kulturen ble så indusert i 3 timer. Deretter ble prøvene analysert ved SDS-PAGE. For å se om induseringstiden påvirket ekspresjonen, ble det satt opp et nytt forsøk hvor bare induseringstiden ble endret, og kulturen ble indusert over natten (ca. 16 timer). Tilslutt ble det satt opp et nytt forsøk med en induseringstid på 4 timer. Her ble det brukt to medier; LB og 2YT. 2YT er et medium med mer gjærekstrakt enn LB medium, og dermed mer næring til bakterien. I tillegg ble det brukt to konsentrasjoner av IPTG; 1 mM og 4 mM, for å se om mengde induser påvirker ekspresjonen.

### 2.10.2 SDS-PAGE av proteiner

SDS-PAGE er en metode som separerer proteiner basert på størrelse. For at separasjon skal skje på grunnlag av størrelse må proteinene ha samme masse- og ladningsforhold. Dette oppnås ved å tilsette en prøvebuffer som inneholder SDS og  $\beta$ -mercaptoetanol. SDS er en detergent som vil denaturere proteinet, mens  $\beta$ -mercaptoetanol vil redusere disulfidbroer. SDS-PAGE ble benyttet for å sjekke ekspresjon av proteinene.

Lysatet (se avsnitt 2.10.1) består av både løselige og uløselige proteiner, og denne prøven blir dermed referert til som den totale prøven. Ved å ta ut 16  $\mu$ l av lysatet ble det undersøkt om man har det aktuelle proteinet. For å finne ut om proteinet var løselig, ble lysatet sentrifugert (Eppendorf centrifuge 5424, Oslo, Norge) ved  $18\,000 \times g$  i 2,5 minutter, og 16  $\mu$ l av supernatanten ble tatt ut. Denne prøven blir videre i oppgaven referert til som den løselige prøven. Uløselig protein vil foreligge i pelleten etter sentrifugering. SDS-PAGE prøvebuffer (GenScript, Leiden, Nederland) ble tilsatt prøvene i forholdet 1:4, og prøvene ble inkubert ved  $95\text{ }^{\circ}\text{C}$  i 10 minutter. 15  $\mu$ l av løsningene ble applisert på gelen (GenScript express plus PAGE gel, Leiden, Nederland). 7  $\mu$ l Precision plus protein Dual-color standards (BIO-RAD, Hertfordshire, USA) ble tilsatt brønnen på hver side av gelen. Elektroforesen ble utført ved 200 V i 25 minutter (BIO-Rad mini Protean Tetra system, Hertfordshire, USA). Etter elektroforesen ble gelen rensert med MilliQ vann før den ble farget med Coomassie Brilliant Blue. Gelen ble farget til man ikke lenger kunne se båndene til standarden på gelen. Dette var ca. 10 minutter, etter 5 minutter ble gelen plassert i mikrobølgeovn for ca. 15 sekunder for å redusere inkuberingstiden. For å fjerne farge som ikke var spesifikt bundet til proteiner, ble gelen avfarget med en løsning av 20 % etanol og 10 % eddiksyre i 1-2 timer. Deretter ble gelen plassert i G-box gel doc (Syngene, Cambridge, England) og et bilde ble tatt ved å bruke Genesnap Software (Syngene, Cambridge, England).

### 2.10.3 Rensing av protein

Affinitetskromatografi baserer seg på molekylets evne til å binde seg til en spesifikk ligand. Kloningsdesignet benyttet i dette prosjektet fører til at det rekombinante proteinet som produseres har et såkalt histidin-merke (HIS-TAG) bestående av 6 påfølgende aminosyrer av typen histidin knyttet til C terminalen. Denne påsettes for rensing ved bruk av immobiliserings metal affinitetskromatografi (IMAC)-kolonner. Prinsippet for IMAC-kolonner er basert på affiniteten histidin har for metallioner. I dette forsøket ble det brukt kolonner hvor metallioner blir immobilisert på en uløselig matriks ved å bruke chelateringsmiddelet iminodieddiksyre, og denne er forhåndslandet med  $\text{Ni}^{2+}$  av produsenten (BIO-RAD, Hertfordshire, USA). Siden

histidin-merkede protein har høy affinitet til dette metallionet vil de binde seg sterkt til IMAC-kolonnen. De øvrige proteinene vil binde seg svakt eller ikke binde seg. Imidazol er et organisk stoff som finnes i sidekjeden til histidin, og kan dermed konkurrere med histidin-merkede proteiner i å binde seg til metallionet. Ved å tilsette en lav konsentrasjon av imidazol til buffer A (appliserings- og vaskebuffer), kan man hindre andre proteiner i å binde seg til metallionene. Ved å tilsette en mye høyere konsentrasjon til buffer B (elueringsbuffer), vil imidazol utkonkurrere de bundne histidin-merkede proteinene og de blir da eluert dvs. de løsner fra og vaskes ut av kolonnen.

Proteinene ble rensset ved å følge instruksjonsmanualen til Bio-Scale mini IMAC kolonnen (BIO-RAD, Hertfordshire, USA). Kolonnen ble brukt sammen med lavtrykkskromatografi systemet Bio-Rad's BioLogic LP (BIO-RAD, Hertfordshire, USA). Bufferløsningene ble laget som beskrevet i tabell 2.16. For å fjerne pakningsvæsken (20 % EtOH) ble kolonnen vasket med to kolonnevolum (CV) MilliQ vann ved strømningshastighet på 2 ml/minutter.

For å fjerne partikler >20 µm ble supernatanten av lysatet filtrert gjennom et 20 µm filter. Deretter ble den filtrerte løsningen overført til en 5 ml sprøyte for applisering av prøven. Rensingen av proteiner ble gjennomført som beskrevet i tabell 2.17. Systemet var direkte knyttet til en PC hvor programmet LP data View Software (BIO-RAD, Hertfordshire, USA), ble benyttet til å følge renseforløpet. Eluatet ble samlet opp i fraksjoner à 1,5 ml som ble analysert ved SDS-PAGE som beskrevet i avsnitt 2.10.3.

*Tabell 2.17: Program for rensing av proteiner med 1ml IMAC kolonne*

<b>Steg</b>	<b>Kolonnevolum (CV)</b>	<b>Strømningshastighet</b>
<b>Ekvibrering (buffer A)</b>	5	2 ml/min
<b>Lysat</b>	2	2 ml/min
<b>Vask (Buffer A)</b>	6	2 ml/min
<b>Eluering (Buffer B)</b>	5	2 ml/min

### 2.11 Rensing av protein - storskala

På bakgrunn av resultatet i småskala forsøket, ble det satt opp et nytt forsøk hvor det ble laget større kulturer. Det ble laget både 500 ml og 1 l kulturer, og disse ble inkubert som beskrevet i avsnitt 2.10.1. Når OD<sub>600nm</sub> var over 0,4, ble ekspresjon induisert med 1 mM IPTG og kulturene inkubert videre i 4 timer. Deretter ble kulturene sentrifugert og pellet ble fryst (-20 °C).

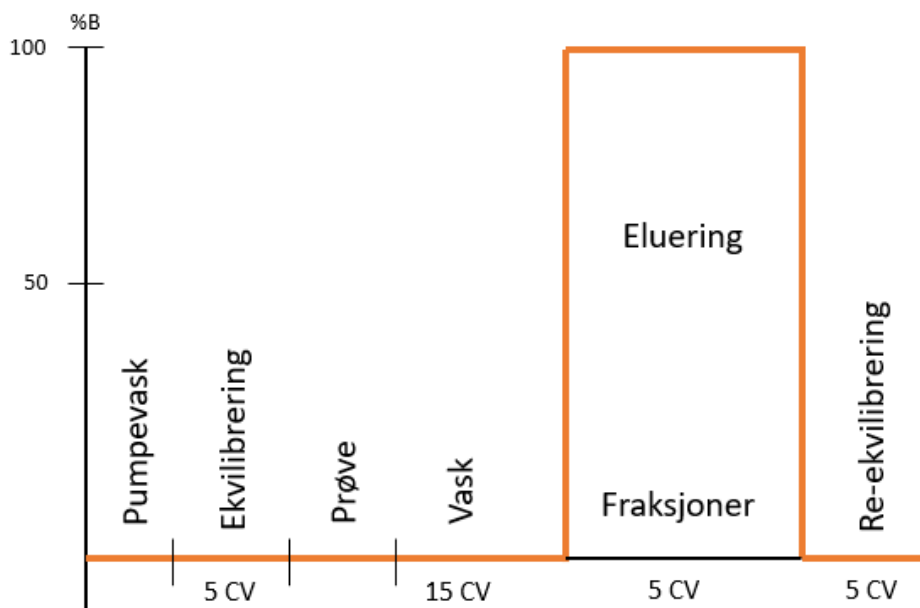
Det ble tatt ut ca. 1,5 ml pellet som ble tint og tilsatt 5 ml lysis buffer (Tabell 2.18). Etter 30 minutter ble lysatet sonikert og sentrifugert som beskrevet i avsnitt 2.10.1. For å rense proteinet ble en 5 ml HisTrap excel Affinity purification column satt opp på et ÄKTA protein rensesystem (Bionity, Berlin). Prinsippet for proteinrensing ved bruk av denne kolonnen er tilsvarende som for Bio-Scale mini IMAC kolonnen (se 2.10.3). Rensingen ble gjennomført i et kjølerom (4 °C) for å holde temperaturen så lav som mulig, og ikke utsette proteinet for temperaturer som kan forårsake denaturering. Proteinet ble rensert som vist i figur 2.4. Før rensing ble kolonnen klargjort ved bruk av 5 ml vaskebuffer og 5 ml ekvilibreringsbuffer (Tabell 2.18). Deretter ble 4,3 ml lysat applisert. Da rensingen var gjennomført ble aktuelle fraksjonene (bestemt fra kromatogrammet) tatt ut og plassert på is. For å vurdere om proteinet var rensert ble fraksjonene analysert ved hjelp av SDS-PAGE sammen med en prøve av lysatet som beskrevet i 2.10.2. Deretter ble fraksjonene konsentrert ved ultrafiltrering (Amicon Ultra-4 Centrifugal Filter Unit, 10 kDa), hvor filteret gjentatte ganger ble fylt på med fraksjoner med rensert protein og sentrifugert (Thermo Fisher heraeus multifuge 3SR+) ved  $2000 \times g$  i 5 minutter ved 4 °C. Dette ble utført til totalvolumet på konsentratet var ca. 500 µl. Den konsentrerte prøven ble så overført til et Eppendorf rør og 1 ml lagringsbuffer (Tabell 2.18) ble tilsatt før røret ble fryst (-20 °C).

Protein utbyttet fra fraksjonene og den oppkonsentrerte løsningen ble målt med Quantus fluorometer (Promega, Wisconsin, USA).

*Tabell 2.18: Oversikt over bufferløsningene brukt for å rense protein. Alle løsningene har pH 7,5*

	<b>HEPES</b>	<b>NaCl</b>	<b>Imidazol</b>	<b>Glyserol</b>	<b>Lysozym</b>	<b>TCEP</b>
<b>Lysisbuffer</b>	50 mM	300 mM	N/A	10 %	0,25 mg/ml	N/A
<b>Ekvilibreringsbuffer</b>	20 mM	0,5 M	N/A	N/A	N/A	N/A
<b>Vaskebuffer</b>	20 mM	500 mM	10 mM	N/A	N/A	N/A
<b>Elueringsbuffer</b>	20 mM	500 mM	500 mM	N/A	N/A	N/A
<b>Lagringsbuffer</b>	20 mM	300 mM	N/A	10 %	N/A	2 mM





Figur 2.4: Rensprogrammet brukt for å rense protein på ÄKTA systemet. CV står for kolonne volum og viser hvor mye av hver buffer som blir sendt gjennom systemet. Den oransje linjen illustrerer når elueringsbufferen (tabell 2.19) er i systemet, og man ser at dette bare er under eluering. På samme tid vil systemet samle fraksjoner med det rensende proteinet. Figur fra protokoll ÄKTA.

## 2.12 Bioinformatiske analyser

For å se på den evolusjonære sammenhengen og slektskapet til pol X og SDHA ble det laget fylogenetiske trær. Homologe sekvenser til pol X ble funnet ved å utføre aminosyresekvenssøk ved NCBI og UniProt. Ved NCBI ble det brukt programmet Blastp (Altschul et al., 1990) ved EBI (<https://www.ebi.ac.uk/Tools/sss/ncbiblast/>) og ved UniProt ble det gjennomført et søk i UniRef90 databasen (<https://www.uniprot.org/blast/>). Deretter ble disse to søkene sammenlignet, og homologe sekvenser ble valgt til videre arbeid. For SDHA ble en FASTA fil med homologe sekvenser laget av Håkon Dahle (Universitetet i Bergen) benyttet. Sekvensene til pol X og SDHA ble hentet fra annoteringstabell til PkV RF01 (fra Ruth-Anne Sandaa, 2019). Sekvenssammenstillinger og trær ble konstruert ved hjelp av MEGA X (<https://www.megasoftware.net/>). I dette programmet ble sekvensene sammenstilt ved å bruke MUSCLE, og metoden «maximum likelihood» med bootstrapping (50 runder) ble brukt for å konstruere de to trærne.

HMMR (Potter et al., 2018) ved EBI (<https://www.ebi.ac.uk/Tools/hmmer/>) ble brukt for å gjøre et domene søk for alle de sju proteinene. Proteiner har vanligvis ett eller flere funksjonelle

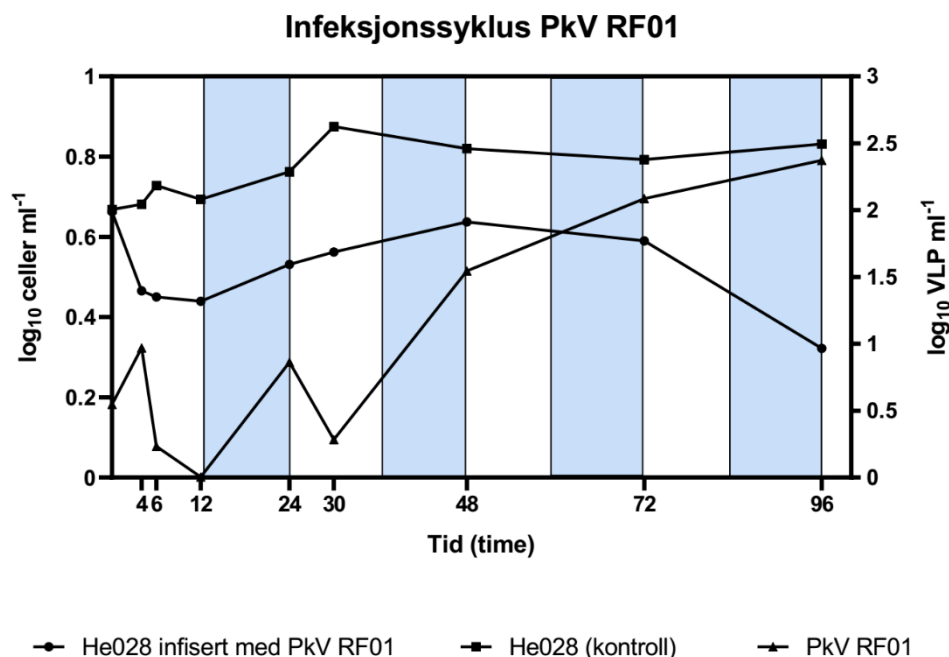
områder, noe som kalles domener. Dette er områder på proteinet som utgjør en strukturell enhet og som kan ha en funksjon uavhengig av resten av proteinkjeden. Forskjellige kombinasjoner av disse domenene gjør at man har en stor variasjon av proteiner i naturen. Identifiseringen av domener i hvert protein gjør at man får en forståelse for funksjonen til proteinet (Potter et al., 2018).

For å forsøke å predikere funksjonen til DNA polymerase X, ble sekvensen sammenlignet med proteinene WP\_001413640.1 og Q5UPZ0.1 ved å bruke Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

### 3 Resultat

#### 3.1 Infeksjonssyklus - PkV RF01

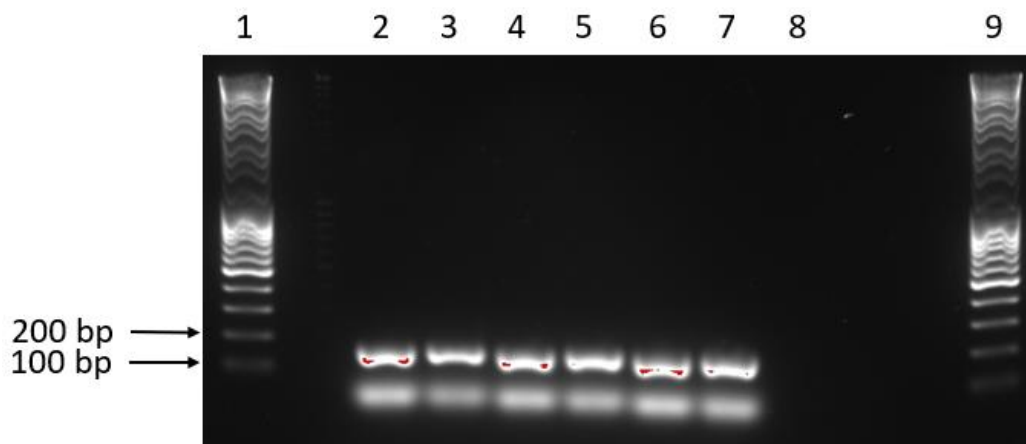
Det ble satt opp et infeksjonsforsøk hvor prøver ble tatt ut til ddPCR (se avsnitt 2.4) og transkripsjonsforsøket (se avsnitt 2.5), og antall verts- og virus celler ble telt ved flowcytometri (se avsnitt 2.1.4). Disse tellingene ble videre brukt til å lage en graf som viser antall vert- og virusceller gjennom infeksjonssyklusen til PkV RF01 (Figur 3.1). Både den infiserte kulturen og kontrollkulturen startet med  $4 \times 10^5$  algeceller. Av figuren ser man at begge kulturer vokser frem til 48 timer. Deretter ser man en nedgang i algeceller i den infiserte kulturen, mens antall celler i kontrollkulturen fortsetter å stige. Dette viser at PkV RF01 har en latent periode mellom 30-48 timer (Figur 3.1). Antall virus partikler øker jevnt etter 24 timer. Fullstendig lysis var nådd rundt 96 timer, som også ble bekreftet visuelt ved å sammenligne den infiserte kulturen med kontrollkulturen, og lysis ble bekreftet ved fargeforandring i den infiserte kulturen. Antall algeceller i kontroll kulturen økte jevnt gjennom hele perioden, og viste ingen tegn til lysis.



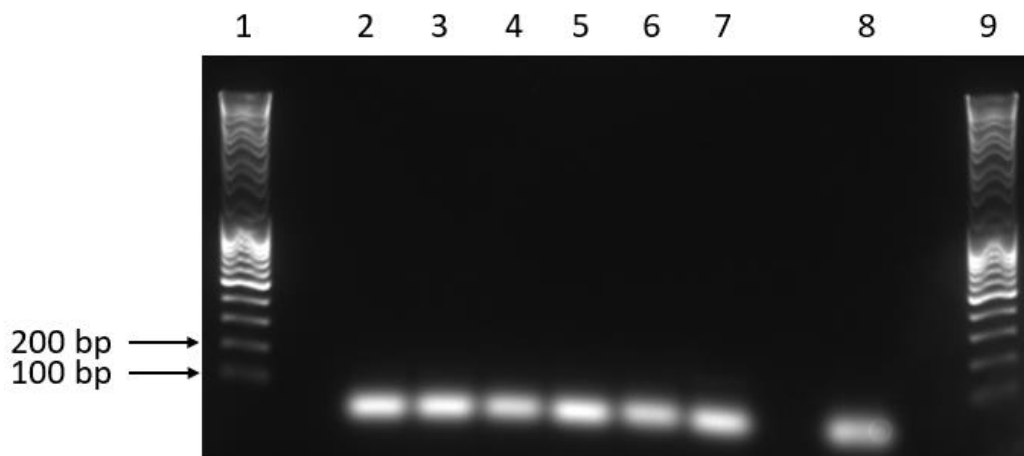
Figur 3.1: Infeksjonssyklus til PkV RF01 dyrket med *H.ericina* (He28). Grafen viser antall virus partikler per ml<sup>-1</sup> (trekant) og vertsceller (kontroll – firkant, infisert vert – sirkel) som har blitt telt på flowcytometer ved gitte tidspunkter i løpet av en syklus. Antall timer siden infeksjon er oppgitt på x-aksen. Venstre y-akse viser antall alge partikler per ml, og høyre akse viser antall virus partikler per ml. Y-aksene er vist som log<sub>10</sub>. De blå feltene viser når kulturen har stått i mørket.

### 3.2 Amplifisering av *sdhA* genet

Optimalisering av PCR-protokollen for *sdhA* genet fra P.kappa virus RF01 (154 bp) ble utført med DNA fra PkV RF01 som templat og DNA fra He028 som kontroll på at primerne var spesifikke for den virale versjonen av *sdhA*. Tydelige bånd med størrelse på ca. 150 bp (Figur 3.2) ble identifisert ved hjelp av agarose gelelektroforese, og dette viste at amplifiseringen av virus versjonen av *sdhA* genet var vellykket. Ingen bånd på PCR med He028 som templat (Figur 3.3) viste at de designede primerne var spesifikke for det virale *sdhA* genet. Det virale genet ble amplifisert i temperaturgradienten 55-60 °C, og basert på resultatet ble 59 °C benyttet videre i PCR-protokollen for *sdhA*.



Figur 3.2: Gelelektroforese (1,5 % agarose) av PCR-produktet til *sdhA*. Brønn 1 og 9: 1 kb DNA-standard. Brønn 2-7: PCR produkt av *sdhA* med PkV RF01-DNA som templat. Gradient: 55 °C (brønn 2) – 60 °C (brønn 8). Brønn 8: Negativ kontroll med *sdH<sub>2</sub>O* som templat.

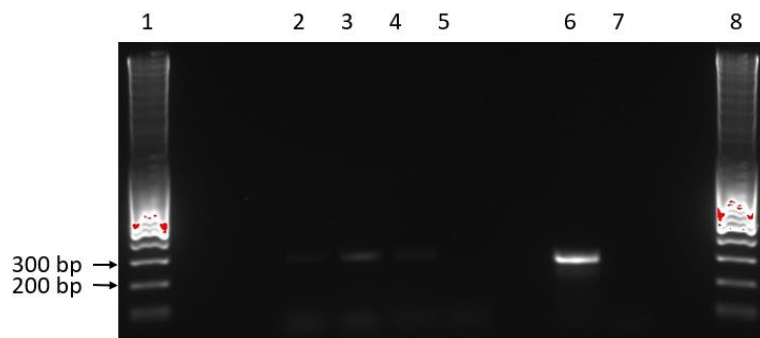


Figur 3.3: Gelelektroforese (1,5 % agarose) til kontroll (He028). Brønn 1 og 9: 1 kb DNA-standard. Brønn 2-7: He028-DNA som templat. Gradient: 55 °C (brønn 2) – 60 °C (brønn 8). Brønn 8: Negativ kontroll med *sdH<sub>2</sub>O* som templat.

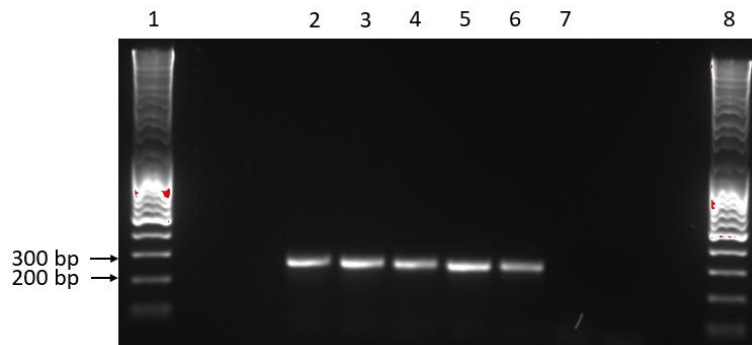
### 3.3 Transkripsjonsforsøk

For å optimalisere protokollen for lysis og RT-PCR ble *mcp* genet til PkV RF01 benyttet. *mcp* genet (246 bp) ble ikke amplifisert (Figur 3.4) når lysatet ble undersøkt ved hjelp av PCR og gelelektroforese noe som tyder på at DNase1 har fungert. Samtidig ble det identifisert tydelige bånd med størrelse ca. 300 bp (Figur 3.5) som tyder på at det har blitt dannet cDNA ved RT-PCR. Basert på dette resultatet ble det bestemt at under transkripsjonen av det virale *sdhA* genet skulle prøvene inkuberes i 5 minutter ved 25 °C etter tilsetning av DNase I

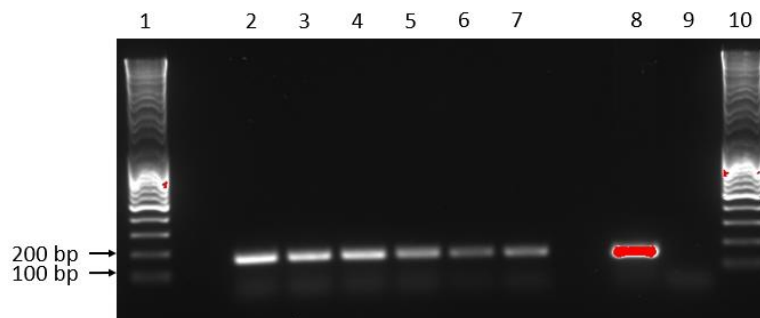
Optimaliserte forhold for lysis ble benyttet for å transkribere *sdhA* genet. Tydelige bånd med størrelse ca. 150 bp (Figur 3.6) ble identifisert når lysatet ble undersøkt ved hjelp av gelelektroforese, noe som tyder på at DNase1 ikke har brutt ned DNAet i prøvene. Samtidig identifiseres bånd på ca. 150 bp (Figur 3.7) som tyder på at *sdhA* genet har blitt amplifisert under RT-PCR. mRNA kontroll (Vedlegg A, Figur 5.1) er vellykket da man har bånd på ca. 150 bp. Det ble flere ganger forsøkt å transkribere *sdhA* genet ved å følge protokollen, men man klarte ikke å bevise at transkriberingen av *sdhA* genet ble vellykket.



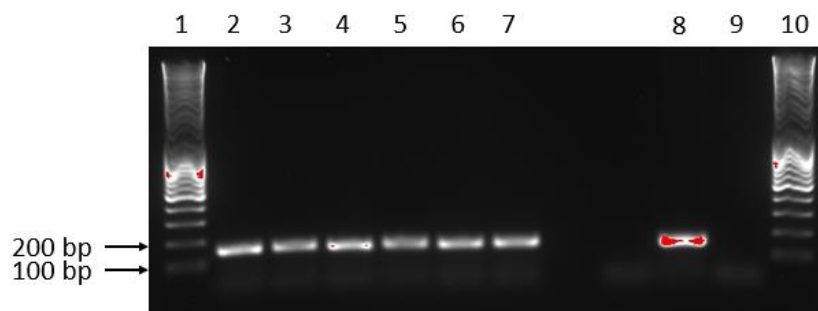
Figur 3.4: Gel elektroforese (1,5 % agarose): kontroll av DNaseI trinnet i lysistrinnet ved amplifisering av *mcp* genet. Brønn 1 og 8: 1 kb DNA-standard. Brønn 2-5: lysat som templat. Brønn 2: prøve tatt 24 timer etter infeksjon med PkV RF01 og med 5 minutt inkubering etter tilsetning av DNase I. Brønn 3: prøve tatt etter 48 timer og 5 minutter inkubering. Brønn 4: Prøve tatt etter 24 timer og 9 minutter inkubering. Brønn 5: Prøve tatt etter 48 timer og 9 minutter inkubering. Brønn 6: positiv kontroll med PkV RF01-DNA som templat. Brønn 7: Negativ kontroll med  $sdH_2O$  som templat.



Figur 3.5: Gelelektroforese (1,5 % agarose) av RT-PCR-produktet til *mcp* genet. Brønn 1 og 8: 1 kb DNA-standard. Brønn 2-5: RT-PCR produkt som templat. Brønn 2: prøve tatt 24 timer etter infeksjon med PkV RF01 og med 5 minutt inkubering etter tilsetning av DNase I. Brønn 3: prøve tatt etter 48 timer og 5 minutter inkubering. Brønn 4: Prøve tatt etter 24 timer og 9 minutter inkubering. Brønn 5: Prøve tatt etter 48 timer og 9 minutter inkubering. Brønn 6: positiv kontroll med PkV RF01-DNA som templat. Brønn 7: Negativ kontroll med  $sdH_2O$  som templat.



Figur 3.6: Gelelektroforese (1,5 % agarose): kontroll av DNaseI trinnet i lysistrinnet ved amplifisering av *sdhA* genet. Brønn 1 og 10: 1 kb DNA-standard. Brønn 2-7: lysat som templat. Brønn 2: prøve tatt 24 timer etter infeksjon med PkV RF01 ( $10^5$  celler). Brønn 3: prøve tatt etter 24 timer ( $5 \times 10^4$  celler). Brønn 4: Prøve tatt etter 24 timer ( $10^4$  celler). Brønn 5: prøve tatt 48 timer ( $10^5$  celler). Brønn 6: prøve tatt etter 48 timer ( $5 \times 10^4$  celler). Brønn 7: Prøve tatt etter 24 timer ( $10^4$  celler). Brønn 8: positiv kontroll med PkV RF01-DNA som templat. Brønn 9: Negativ kontroll med  $sdH_2O$  som templat.

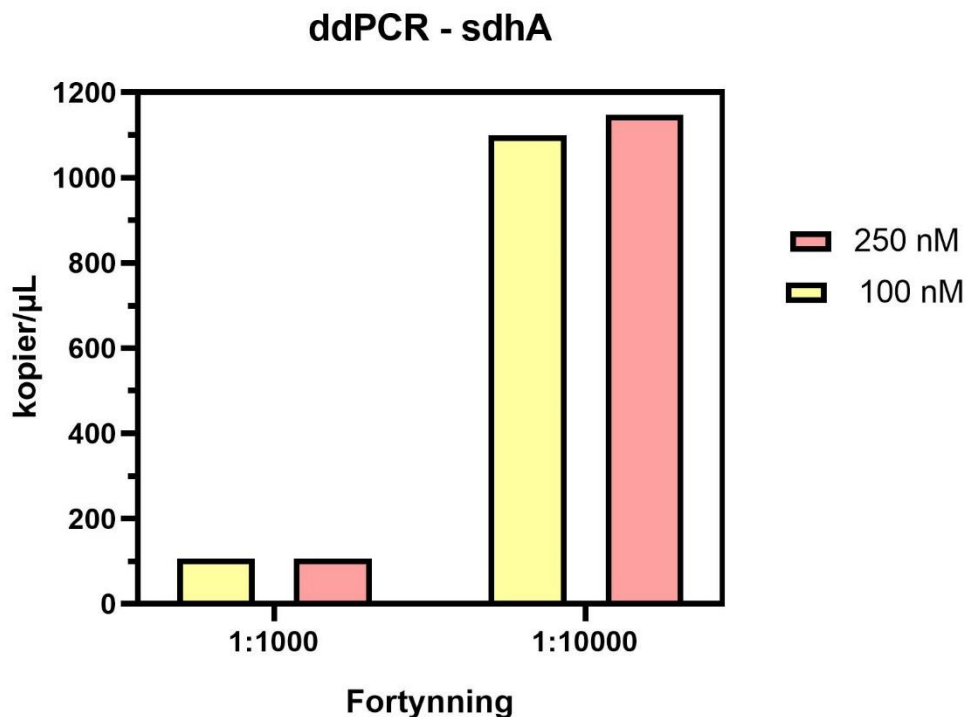


Figur 3.7: Gelelektroforese (1,5 % agarose) av RT-PCR-produktet til *sdhA* genet. Brønn 1 og 10: 1 kb DNA-standard Brønn 2-7: RT-PCR produkt som templat. Brønn 2: prøve tatt 24 timer etter infeksjon med PkV RF01 ( $10^5$  celler). Brønn 3: prøve tatt etter 24 timer ( $5 \times 10^4$  celler). Brønn 4: Prøve tatt etter 24 timer ( $10^4$  celler). Brønn 5: prøve tatt 48 timer ( $10^5$  celler). Brønn 6: prøve tatt etter 48 timer ( $5 \times 10^4$  celler). Brønn 7: Prøve tatt etter 24 timer ( $10^4$  celler). Brønn 8: positiv kontroll med PkV RF01-DNA som templat. Brønn 9: Negativ kontroll med  $sdH_2O$  som templat.

### 3.4 ddPCR - SDHA

ddPCR-protokollen for SDHA ble optimalisert med DNA fra PkV RF01 for å finne ut hvilke konsentrasjoner av primere og fortynninger av DNA som gav best resultat. Videre ble det undersøkt hvor mange kopier av *sdhA* genet som blir dannet i de to fortynningene. Fortynningen 1:100 er ikke vist på grafen da det ikke ble registrert noen kopier i denne fortynningen. Ved 1:1000 fortynning av DNA (Tabell 3.8) ble det dannet 106,5 og 107 kopier/ $\mu$ l. Mens med 1:10 000 fortynning ble det dannet 1148 kopier/ $\mu$ l ved 250 nM primer konsentrasjon, sammenlignet med 1099 kopier/ $\mu$ l ved 100 nM.

ddPCR skulle benyttes for å kvantifisere transkripsjonen av det virale *sdhA* genet gjennom infeksjonssyklusen til PkV RF01. Men da man ikke klarte å transkribere genet, ble ikke dette gjennomført.

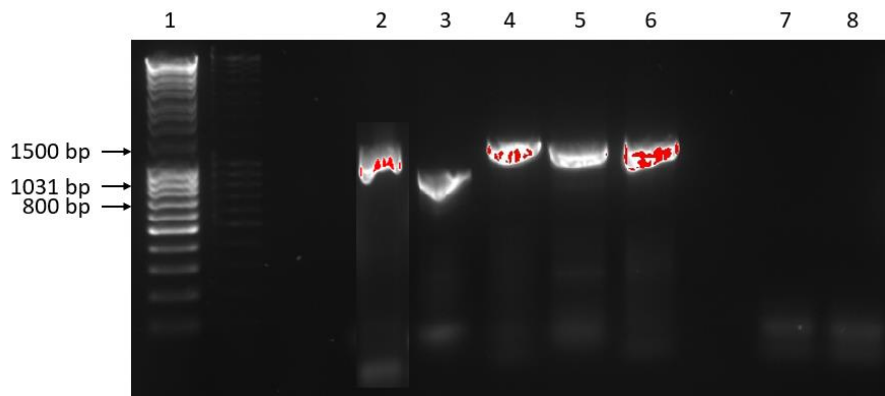


Figur 3.8: Viser antall *sdhA* kopier/ $\mu$ l etter amplifisering ved ddPCR med spesifikke primere for *sdhA*. 250 nM primer konsentrasjon vises i rosa og 100 nM primer konsentrasjon vises i gul. For hver primerkonsentrasjon har man to DNA fortynninger; 1:1000 og 1:10 000.

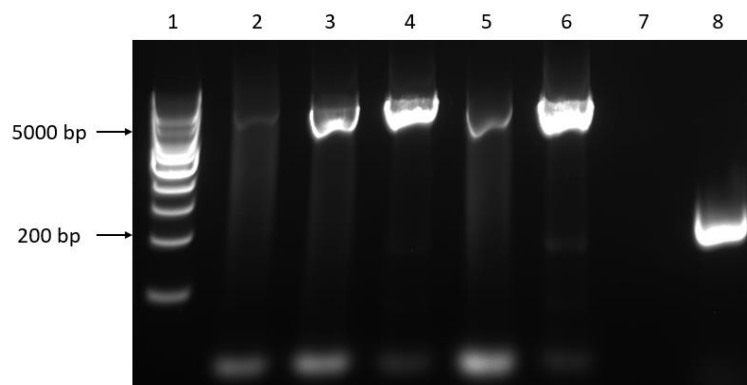
### 3.5 Kloning av målgener og vektorer

#### 3.5.1 Amplifisering av målgener og vektor

Korrekt størrelse på PCR-produkt av L-LDH (1184 bp), PDH (833 bp), GCDH (1232 bp), pol X (1212 bp), GDH (1247 bp), SDHA (1808 bp) og SDHB (701 bp) samt korresponderende pET-21a PCR-produkt for hvert målgen ble verifisert ved hjelp av agarose gelelektroforese (Figur 3.9 og 3.10). Et tydelig bånd med rett størrelse for hvert av målgenene og vektorene viste at amplifiseringen var vellykket for alle genene utenom *sdhA* og *sdhB*. For amplifiseringen av *sdhA* genet fikk man et produkt som var dobbelt så stort som det skulle, samtidig som at amplifiseringen av vektor-produktet til SDHA viste mye uspesifikke bindinger (Figur 3.11). For *sdhB* ble ikke amplifiseringen av genet vellykket, mens det tilhørende vektorproduktet amplifisert er for stort i forhold til hva som var forventet på ca. 5000 bp (Figur 3.11). Dermed ble SDHA og SDHB ikke forsøkt klonet.

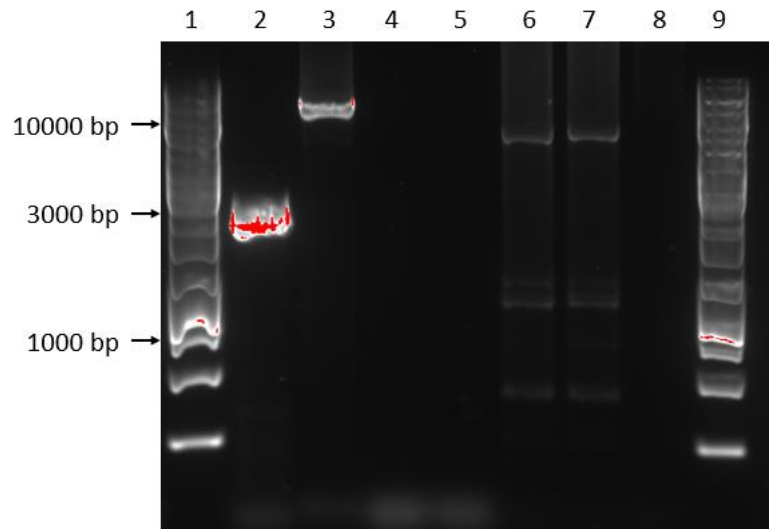


Figur 3.9: Gelelektroforese (1,5 % agarose) av PCR-produktet til målgenene. Brønn 1: 1 kb DNA standard. Brønn 2: PCR-produkt av *l*-ldh. Brønn 3: PCR-produkt av *pdh*. Brønn 4: PCR-produkt av *gcdh*. Brønn 5: PCR produkt av *pol X*. Brønn 6: PCR produkt av *gdh*. Brønn 7: Negativ kontroll med He028-DNA som templat. Brønn 8: Negativ kontroll med  $sdH_2O$  som templat.



Figur 3.10: Gelelektroforese (1 % agarose) av PCR-produktene til *pet-21a* vektor. Brønn 1: 1 kb DNA standard. Brønn 2: PCR produkt av vektor til *gcdh*. Brønn 3: PCR produkt av vektor til *l*-ldh. Brønn 4: PCR produkt av vektor til *pol X*. Brønn 5: PCR produkt av vektor til *gdh*. Brønn 6: PCR produkt av vektor til *pdh*. Brønn 7: negativ kontroll med  $sdH_2O$  som templat. Brønn 8: Positiv kontroll med AF0209 *pET-21a* som templat.

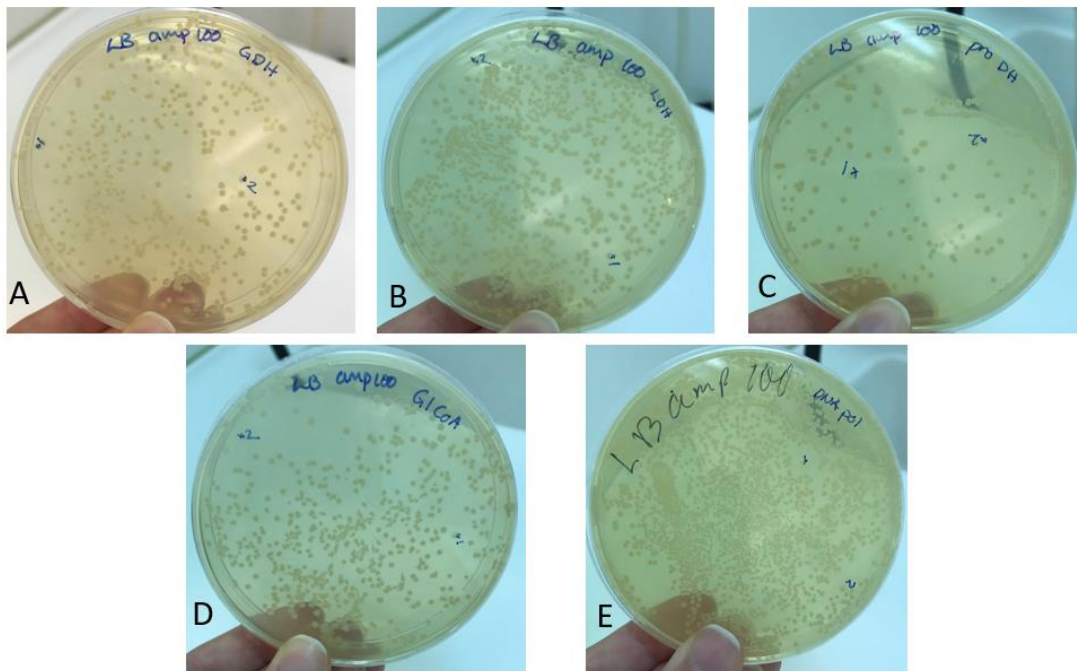




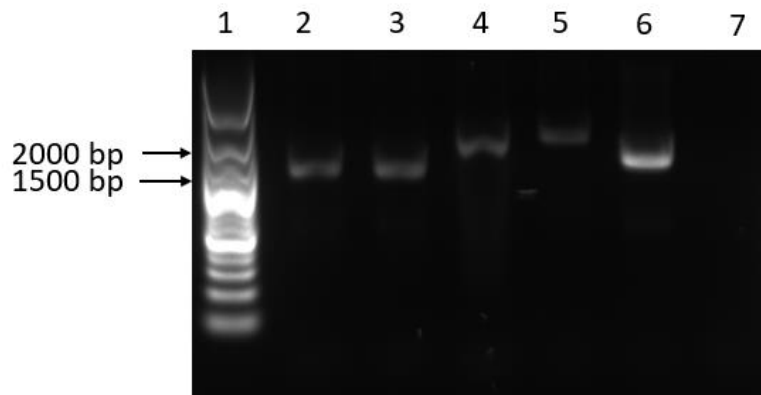
Figur 3.11 Gelelektroforese (1,5 % agarose) av gen-og vektor produkt til *sdhA* og *sdhB*. Brønn 1 og 9: 1 kb DNA standard. Brønn 2: PCR-produkt av *sdhA*. Brønn 3: PCR-produkt av vektor tilhørende *sdhB*. Brønn 4: PCR-produkt av *sdhB* (54 °C). Brønn 5: PCR produkt av *sdhB* (56 °C). Brønn 6: PCR-produkt av vektor tilhørende *sdhA* (58 °C). Brønn 7: PCR-produkt av vektor tilhørende *sdhA* (60 °C). Brønn 8: Negativ kontroll med H<sub>2</sub>O.

### 3.5.2 Kloning av målgener i pET-21a

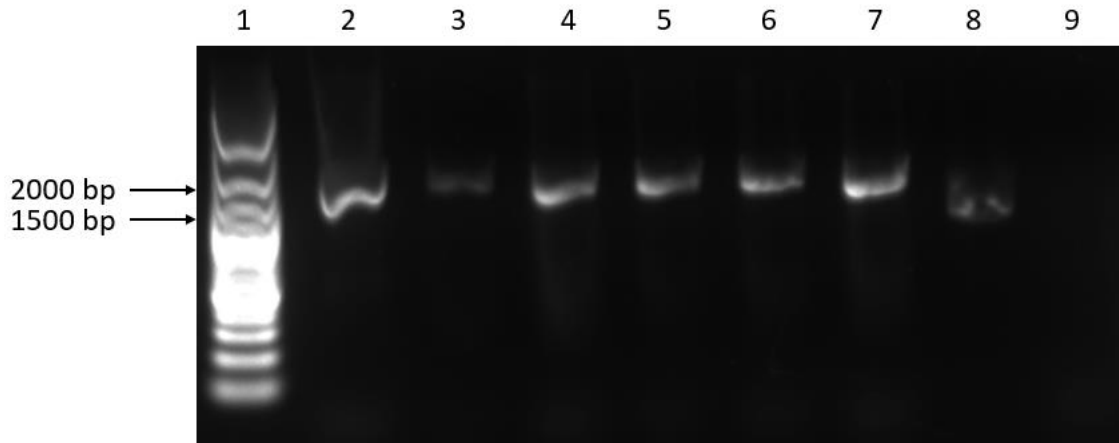
Det ble observert god vekst med god spredning av kolonier på skålene etter transformasjon i *E. coli* NEB 5-alfa kompetente celler av hver av kloningsreaksjonene av PCR-amplifisert målgen og korresponderende pET-21a plasmid (Figur 3.12). Videre ble korrekt størrelse på målgen funnet ved koloni-PCR ved bruk av flankerende pET-21a primere. Forventet størrelse på hvert målgen vises i tabell 2.15 (antall basepar), og resultatet av gelelektroforesen viser at *E. coli* koloniene har fått transformert den riktige klonen (Figur 3.13 og 3.14). Basert på dette ble koloni 1 for hvert av målgenene valgt til videre arbeid (markert på skålene Figur 3.12).



Figur 3.12: Viser LA-skåler etter transformasjon av det klonede produktet inn i *E. coli* NEB 5-alfa kompetente celler. Skål A: GDH. Skål B: L-LDH. Skål C: PDH. Skål D: GCDH. Skål E: Pol X. På hver skål er det markert hvilke to kolonier som ble brukt videre (1 og 2).



Figur 3.13: Gelelektroforese (1% agarose) av pET-21a-plasmid fra 2 av transformantene til PDH og Pol X. Brønn 1: DNA standard. Brønn 2 og 3: plasmid fra transformant til PDH. Brønn 4 og 5: plasmid fra transformant til Pol X. Brønn 6: positiv kontroll med AF0209 pET-21a som templat. Brønn 7: negativ kontroll med sdH<sub>2</sub>O som templat.



Figur 3.14: Gelelektroforese (1 % agarose) av pET-21a-plasmid fra 2 av transformantene til GCDH, L-LDH og GDH. Brønn 1: DNA standard. Brønn 2 og 3: plasmid fra transformant til GCDH. Brønn 4 og 5: plasmid fra transformant til L-LDH. Brønn 6 og 7: plasmid fra transformant til GDH. Brønn 8: positiv kontroll med AF0209 pET-21a som templat. Brønn 9: negativ kontroll med sdH<sub>2</sub>O som templat.

### 3.6 Ekspresjon av målgener i *E. coli*

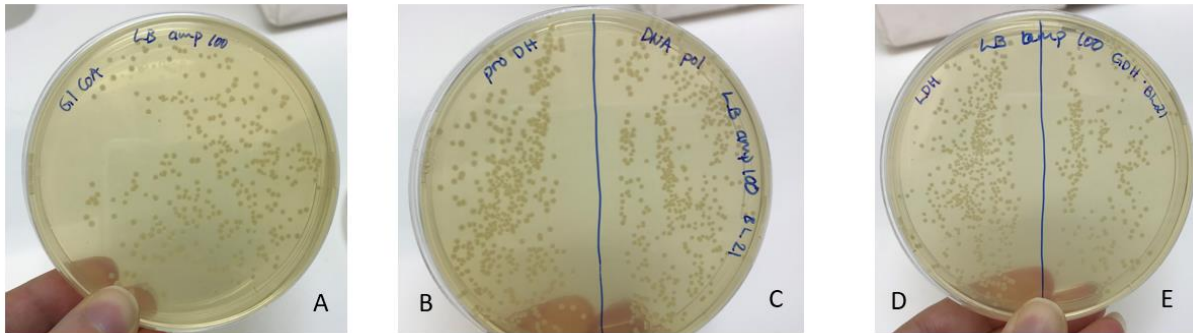
Plasmidisolering resulterte i følgende plasmid-DNA konsentrasjoner (24–55 ng/μl, Tabell 3.1) etter transformasjon inn i NEB 5-alfa kompetente celler. Dette er betydelig høyere enn anbefalt konsentrasjon fra produsenten (Agilent, < 0,1 ng/μl per 100 μl kompetente celler).

Tabell 3.1: Viser plasmid-DNA konsentrasjonen til hver av de 5 målgenene målt med Quantus fluorometer.

Gen	Plasmid DNA
<i>pol X</i>	28 ng/μl
<i>pdh</i>	36 ng/μl
<i>gcdh</i>	55 ng/μl
<i>l-ldh</i>	24 ng/μl
<i>gdh</i>	33 ng/μl

Det ble observert god vekst med god spredning av kolonier på skålene etter transformasjon i *E. coli* BL21-GOLD(DE3) av hver av det rensede plasmid-DNAet (Figur 3.15). Korrekt størrelse på målgen ble funnet ved koloni-PCR ved å bruke flankerende pET-21a primere. Det ble bekreftet at transformantene til fire av målgenene hadde riktig størrelse (se tabell 1.1 for størrelse). Kolonien som ble valgt for L-LDH hadde ingen synlig bånd på agarosegelen, som tyder på at plasmidet ikke var korrekt transformert. Dermed ble seks nye kolonier til L-LDH

valgt og undersøkt ved PCR og gelelektroforese. Fem av disse koloniene fikk tydelig bånd (resultat ikke vist), og dermed en bekreftelse på at riktig plasmid var transformert. For hvert av målgene ble det dermed valgt en koloni til videre arbeid.

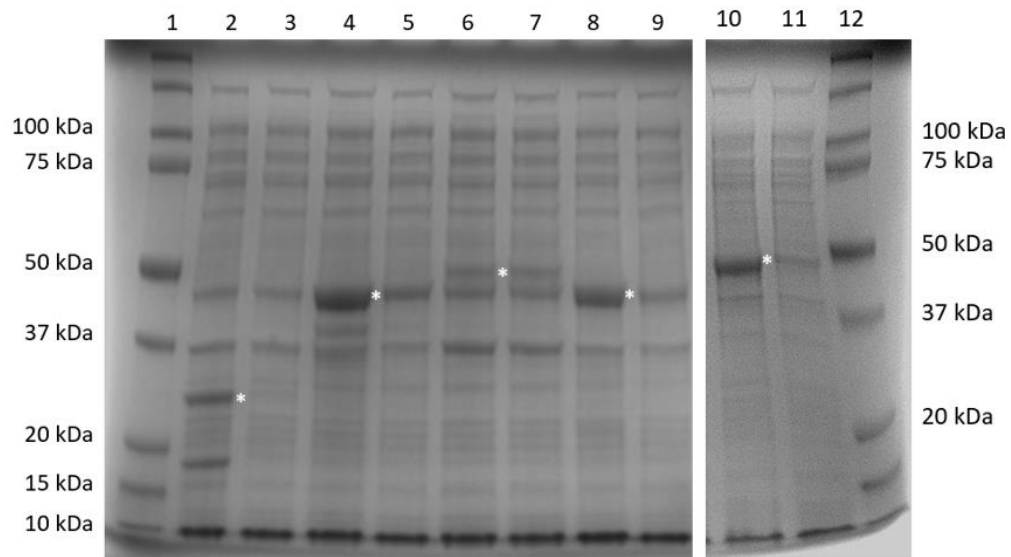


Figur 3.15: LA-skåler etter transformasjon inn i *E. coli* BL21-GOLD(DE3) celler. A: *gcdh*. B: *pdh*. C: *pol X*. D: *l-ldh*. E: *gdh*.

Det ble satt opp flere ekspresjonsforsøk for å finne forholdene som ga høyest ekspresjon (resultat ikke vist). Resultatet av SDS PAGE ved de forskjellige forsøkene viste at det ikke var betydelig høyere ekspresjon ved å øke induseringstiden eller konsentrasjonen av IPTG, og 2YT medium hadde heller ingen innvirkning på ekspresjonen. Dermed ble det bestemt at ved videre ekspresjonsforsøk skulle det brukes LB-medium, 1 mM IPTG og 4 timer indusering.

Ved å studere resultatet av SDS PAGE, ser man at flere av proteinene har bånd i både den totale og den løselige prøven (Figur 3.16), noe som kan indikere at de er løselige proteiner. PDH har en estimert molekylvekt på 32,80 kDa, og i den totale prøven (brønn 2) ser man et tydelig bånd på ca. 32 kDa. Dette båndet er svakt synlig i den løselige prøven (brønn 3). Men basert på SDS-PAGE fra samtlige ekspresjonsforsøk er båndet så svakt at det kan tyde på at dette proteinet er uløselig. GCDH har en estimert molekylvekt på 46,49 kDa, og man kan se et tydelig og sterkt bånd i den totale prøven (brønn 4) på ca. 46 kDa, som også er tilstede i den løselige prøven (brønn 5), og dette kan tyde på at proteinet er løselig. Pol X har en estimert molekylvekt på 46,92 kDa, og i den totale prøven (brønn 6) er det et tydelig bånd på ca. 47 kDa som også er tilstede i den løselige prøven (brønn 7). Ved å sammenligne med de andre brønnene ser man at dette ikke er et bånd som er i andre prøver og dermed indikerer dette at det ikke kan være fra *E. coli*. GDH har en estimert molekylvekt på 47,03 kDa, og har et veldig tydelig bånd i den totale prøven (brønn 8), som også er tilstede i den løselige prøven (brønn 9), som kan tyde på at proteinet er løselig. L-LDH har en estimert molekylvekt på 44,18 kDa, og man ser et tydelig bånd i den totale prøven (brønn 10) på ca. 45 kDa, og dette er tilstede i den løselige prøven

(brønn 11) som kan tyde på at proteinet er løselig. Basert på dette ble det bestemt at proteinene GCDH, L-LDH og pol X skulle renses.



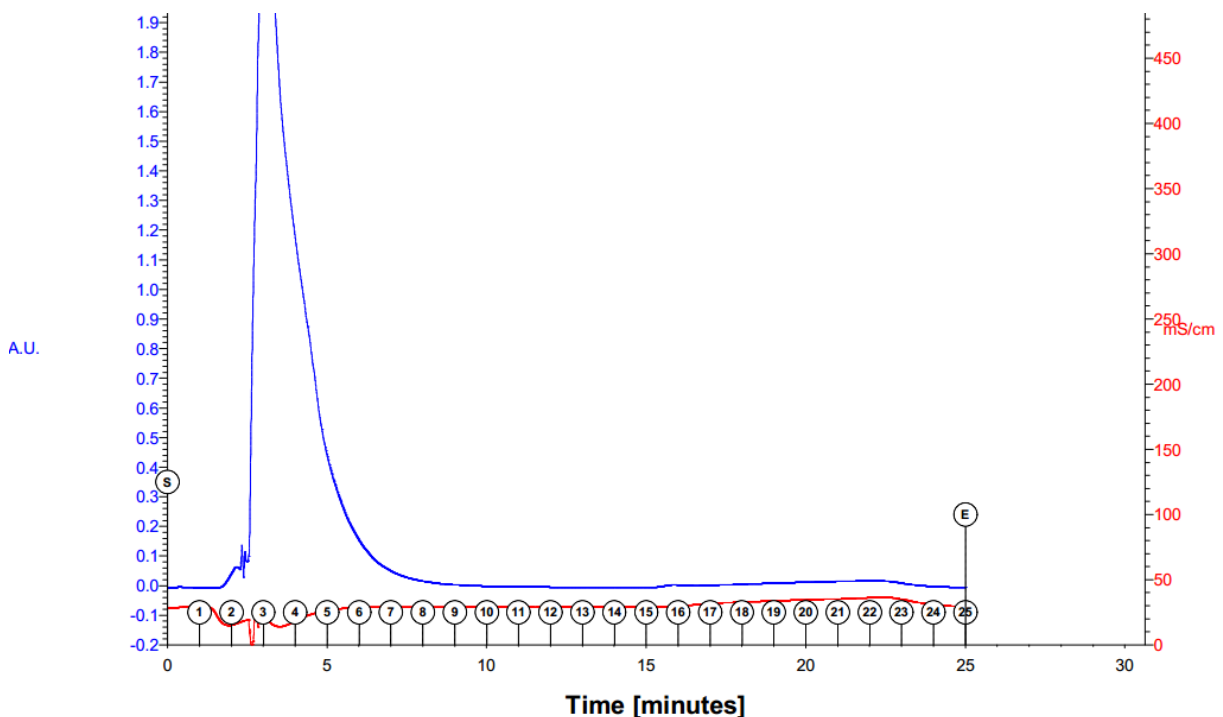
*Figur 3.16: SDS-PAGE analyse av ekspresjon av proteinene i BL21-GOLD(DE3) med ampicillin som seleksjonsantibiotika. Brønn 1 og 12: SDS-PAGE Precision plus protein standards, 10-250 kDa (BIO RAD). Brønn 2: PDH total. Brønn 3: PDH løselig. Brønn 4: GCDH total. Brønn 5: GCDH løselig. Brønn 6: pol X total. Brønn 7: pol X løselig. Brønn 8: GDH total. Brønn 9: GDH løselig. Brønn 10: L-LDH total. Brønn 11: L-LDH løselig. \* viser hvor det induerte båndet ligger for hvert protein.*

### 3.7 Rensing av proteiner

L-LDH, GCDH og pol X ble forsøkt rensed da resultatet fra SDS PAGE viste at alle tre proteinene hadde ett bånd med rett molekylvekt i både den totale og den løselige prøven, som tydet på at proteinet var løselig.

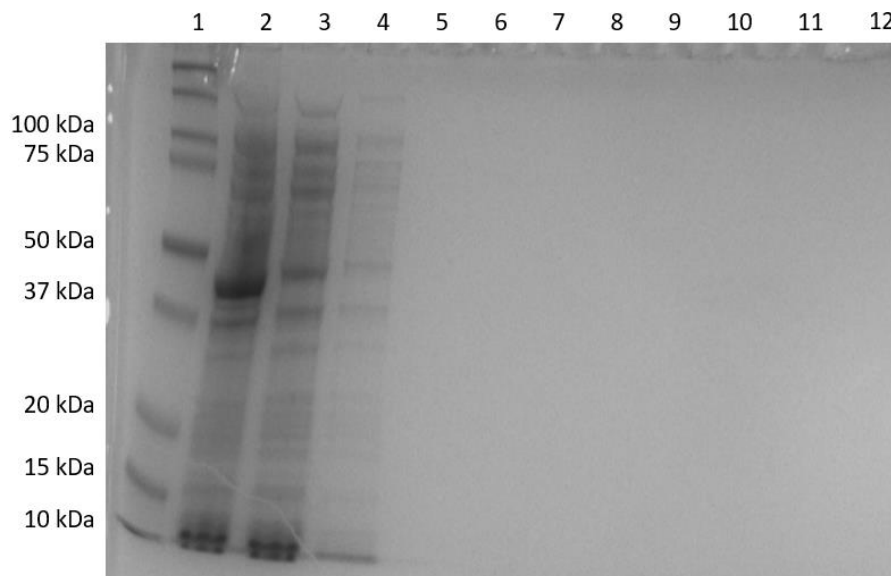
#### 3.7.1 L-laktat dehydrogenase og glutaryl-CoA dehydrogenase

Ved rensing av L-LDH ble det observert en sterk stigning i absorbans og ett fall i konduktivitet en stund etter at vaskebuffer var tilsatt til systemet (Figur 3.17). Dette tyder på at en stor mengde proteiner har passert gjennom kolonnen, og blitt detektert av UV monitoren. Dette er proteiner som ikke er histidin-merket og som dermed har blitt vasket rett gjennom kolonnen. Etter tilsetning av elueringsbuffer, som har en høy konsentrasjon av Imidazol, ser man fra kromatogrammet (Figur 3.17) en svak stigning i absorbans etter ca. 20 minutter, noe som kan tyde på at man har en liten mengde protein som har bundet seg til metallionene i kolonnen, og dermed har blitt eluert når imidazolmolekylene binder seg til metallionene. Basert på grafen ble det bestemt at fraksjon nr. 4, 11, 17, 18, 19, 20, 21, 22 og 23 skulle undersøkes ved SDS PAGE for å bekrefte at proteinet L-LDH var blitt rensed.



Figur 3.17: Diagrammet viser elueringsprofilen for rensing av L-LDH på IMAC kolonnen. X-aksen viser tiden i minutter. Tallene med rounding rundt (langs x-aksen) viser fraksjoner, og hver fraksjon var på 1,5 ml. Venstre y-akse (blå) viser UV absorbans (A.U.), og høyre y-akse (rød) viser konduktivitet (mS/cm).

SDS-PAGE-analysen viste at rensing av L-LDH ikke var vellykket (Figur 3.18). Som en bekreftelse på at proteinet ble rensset, skulle fraksjon 17-23 fått bånd på gelen. Dette båndet skulle ha vært rundt 44 kDa. Da det ikke er noen bånd på noen av brønnene, utenom fraksjon 4, har man ikke det aktuelle proteinet.



*Figur 3.18: SDS-PAGE-analyse av prøver ved rensingen av rekombinant l-ldh. Brønn 1: SDS-PAGE Precision plus protein standards, 10-250 kDa (BIO RAD). Brønn 2: lysat l-ldh (total). Brønn 3: lysat l-ldh filtrert (løselig). Brønn 4: Fraksjon 4. Brønn 5: Fraksjon 11. Brønn 6: Fraksjon 17. Brønn 7: Fraksjon 18. Brønn 8: Fraksjon 19. Brønn 9: Fraksjon 20. Brønn 10: Fraksjon 21. Brønn 11: Fraksjon 22. Brønn 12: Fraksjon 23.*

Da GCDH ble forsøkt rensset ved samme metode som L-LDH, ble heller ikke her rensingen vellykket. Elueringsprofilen (se Vedlegg A.2, Figur 7.3) er veldig lik som ved L-LDH, og fraksjoner ble valgt basert på denne. Den har en svak stigning etter tilsetning av elueringsbuffer, men ingen bånd på SDS-PAGE gelen viser at proteinet ikke har blitt rensset (se Vedlegg A.2, Figur 7.4).

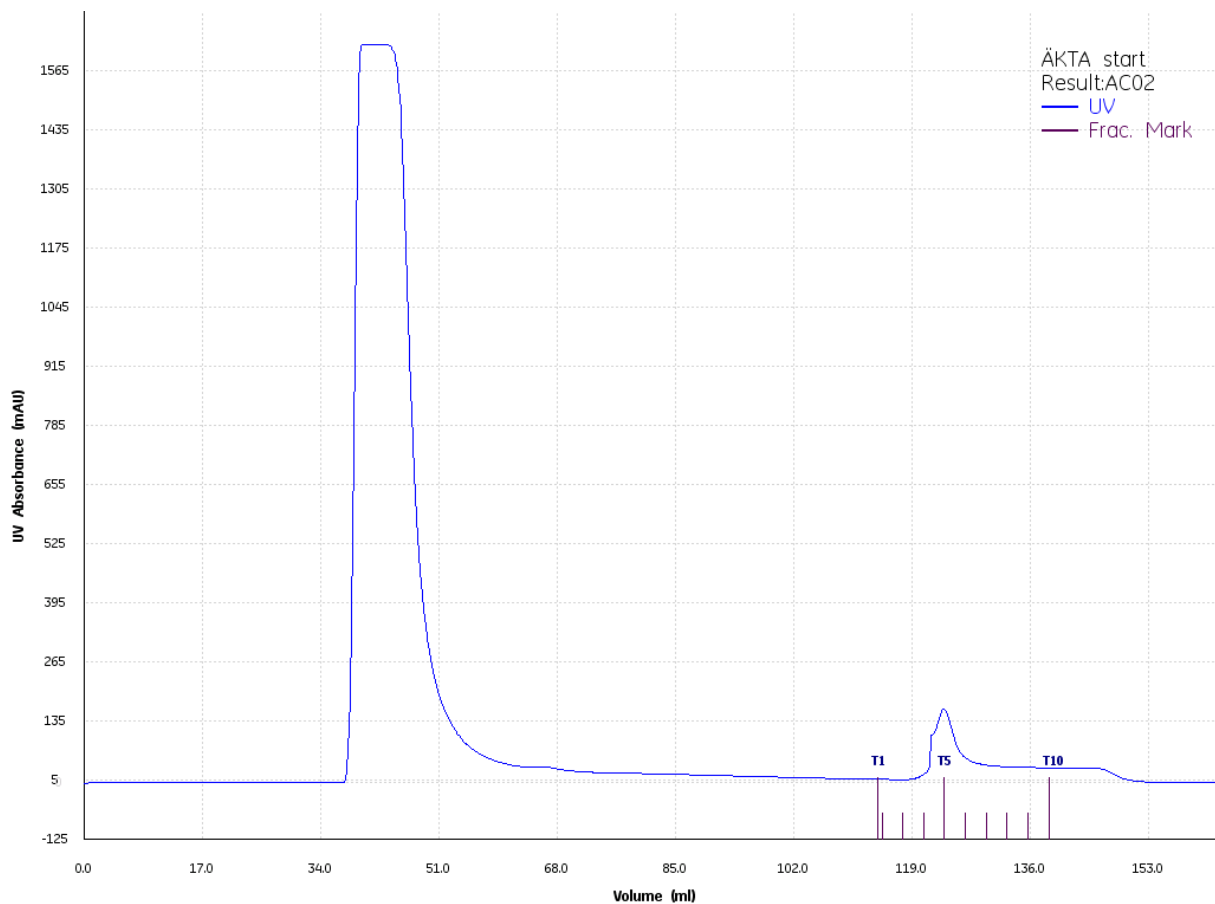
### 3.7.2 DNA polymerase X

DNA polymerase X (pol X) ble vellykket rensset i småskala forsøket (se Vedlegg A.3, figur 7.5 og 7.6). Det ble dannet et tydelig bånd fra fraksjon 9, som stemmer med molekylvekten til proteinet (46,92 kDa). Det ble bestemt at proteinet skulle renses i større skala, og bli sendt til Tromsø for karakterisering.

Elueringsprofilen til pol X (Figur 3.19) viser samme mønster som ved rensing av L-LDH og GCDH. Etter tilsetning av prøve og vaskebuffer får man en høy stigning i absorbansen grunnet proteinene som ikke fester seg til kolonnen. Deretter får man en mye mindre stigning når det



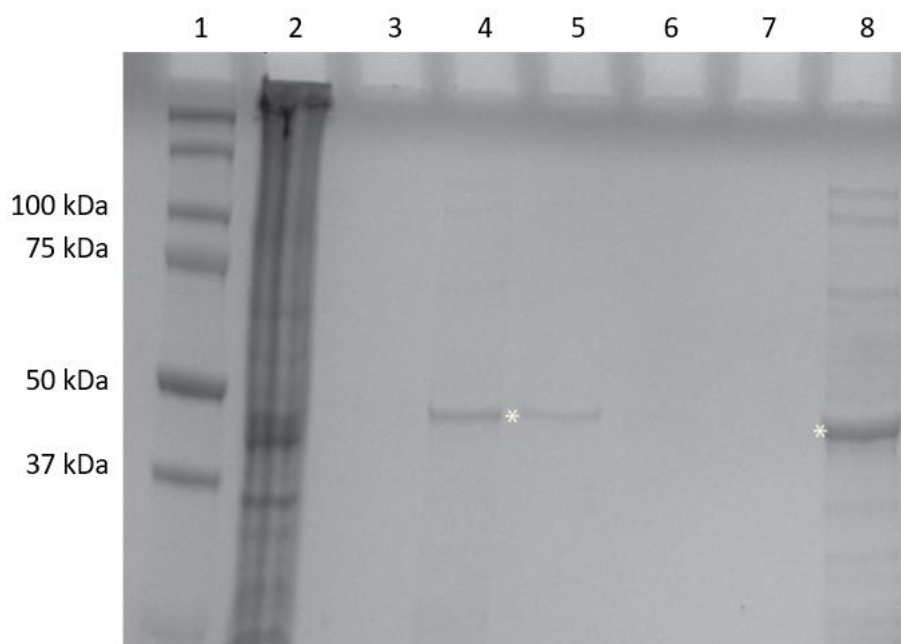
his-merkede proteinet blir eluert. Basert på diagrammet ble fraksjon 3, 4, 5, 6 og 7 undersøkt på SDS-PAGE, samt lysatet og det oppkonsentrerte eluatet.



Figur 3.19: Diagrammet viser elueringsprofilen for rensing av DNA polymerase X på HisTrap kolonnen. X-aksen viser volum (ml). T1-T10 står for fraksjoner. Y-aksen viser UV absorbans (mAU).

SDS-PAGE-analysen viste at rensingen av pol X var vellykket (Figur 3.20). Prøven i brønn 2 inneholder alle løselige proteiner, mens brønn 4 og 5 inneholder bare det rensede proteinet. Molekylvekten til pol X er 46,92 kDa, noe som stemmer med størrelsen til båndet i brønn 4 og 5, og proteinet har da blitt rensset. Dette båndet ser man også i brønn 8 hvor man har det oppkonsentrerte eluatet.





Figur 3.20: SDS-PAGE-analyse av prøver ved rensingen av rekombinant Pol X. Brønn 1: SDS-PAGE Precision plus protein standards, 10-250 kDa (BIO RAD). Brønn 2: lysat Pol X (løseligl). Brønn 3: Fraksjon 3. Brønn 4: Fraksjon 4. Brønn 5: Fraksjon 5. Brønn 6: Fraksjon 6. Brønn 7: Fraksjon 7. Brønn 8: Oppkonsentrert eluat. \* viser hvor båndet ligger for Pol X.

### 3.7.2.1 Protein konsentrasjoner etter rensing

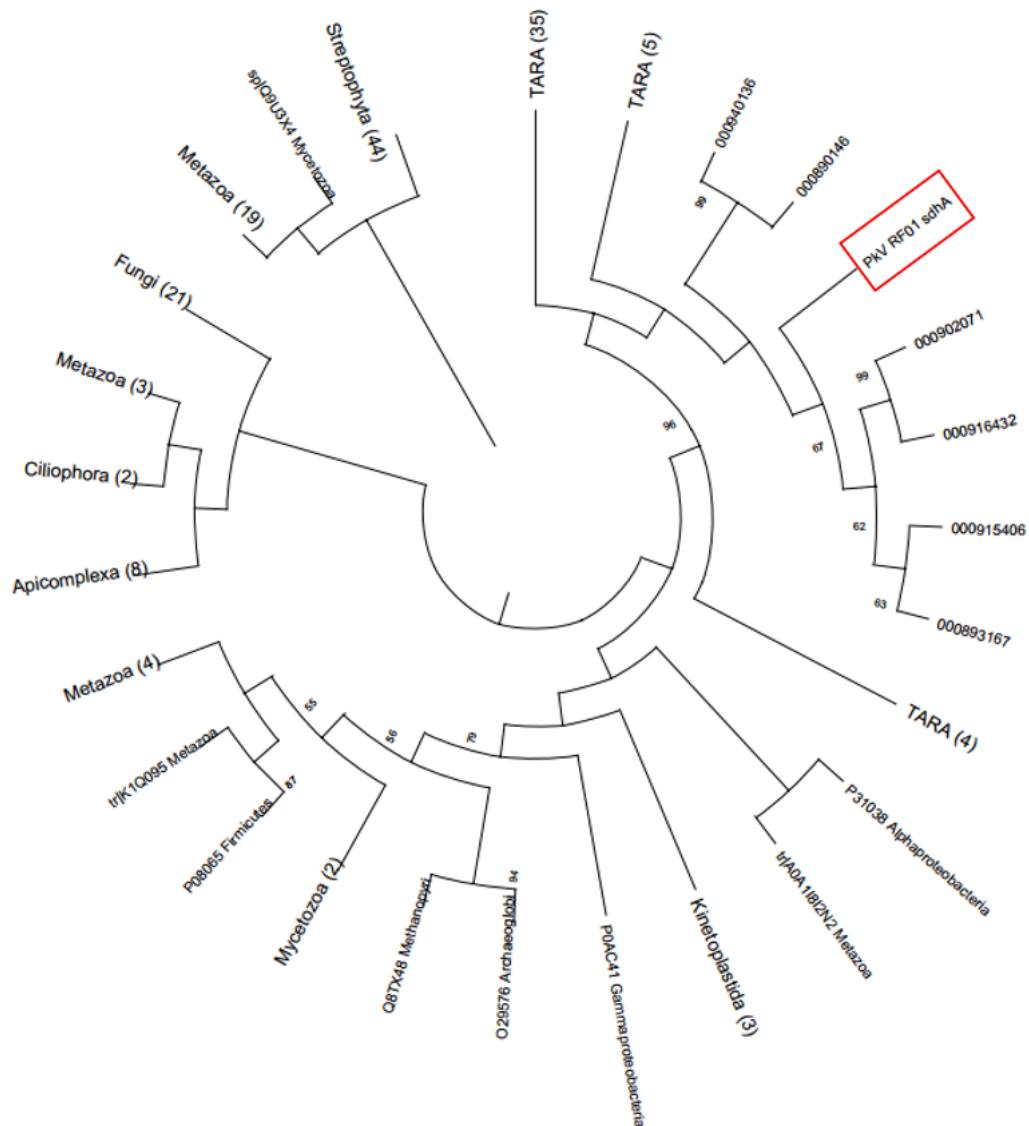
Etter at pol X ble rensert i stor skala, ble protein konsentrasjonen målt på Quantus fluorometer (Promega, Wisconsin, USA). Dette ble gjort med de fraksjonene som fikk bånd på SDS-PAGE gelen, fraksjon 4 og 5 samt det oppkonsentrerte eluatet. Det ble kjørt en standard, som fikk en konsentrasjon på 1958  $\mu\text{g/ml}$ . Det oppkonsentrerte eluatet (som består av fraksjon 4 og 5) fikk en konsentrasjon på 0,7 mg/ml (Tabell 3.2). På forhånd ble det gitt beskjed om at prøven som skulle sendes måtte ha en konsentrasjon på over 0,2 mg/ml for at karakteriseringen kunne gjennomføres. Pol X konsentrasjonen er godt over dette, og ble dermed tilsatt lagringsbuffer, og oppbevart i  $-20\text{ }^{\circ}\text{C}$  frem til sending.

Tabell 3.2: Protein konsentrasjon målt på Quantus fluorometer etter rensing av pol X.

Prøve	Protein konsentrasjon ( $\mu\text{g/ml}$ )
Standard	1958
Fraksjon 4	367
Fraksjon 5	167
Oppkonsentrert eluat	697

### 3.8 Bioinformatiske analyser

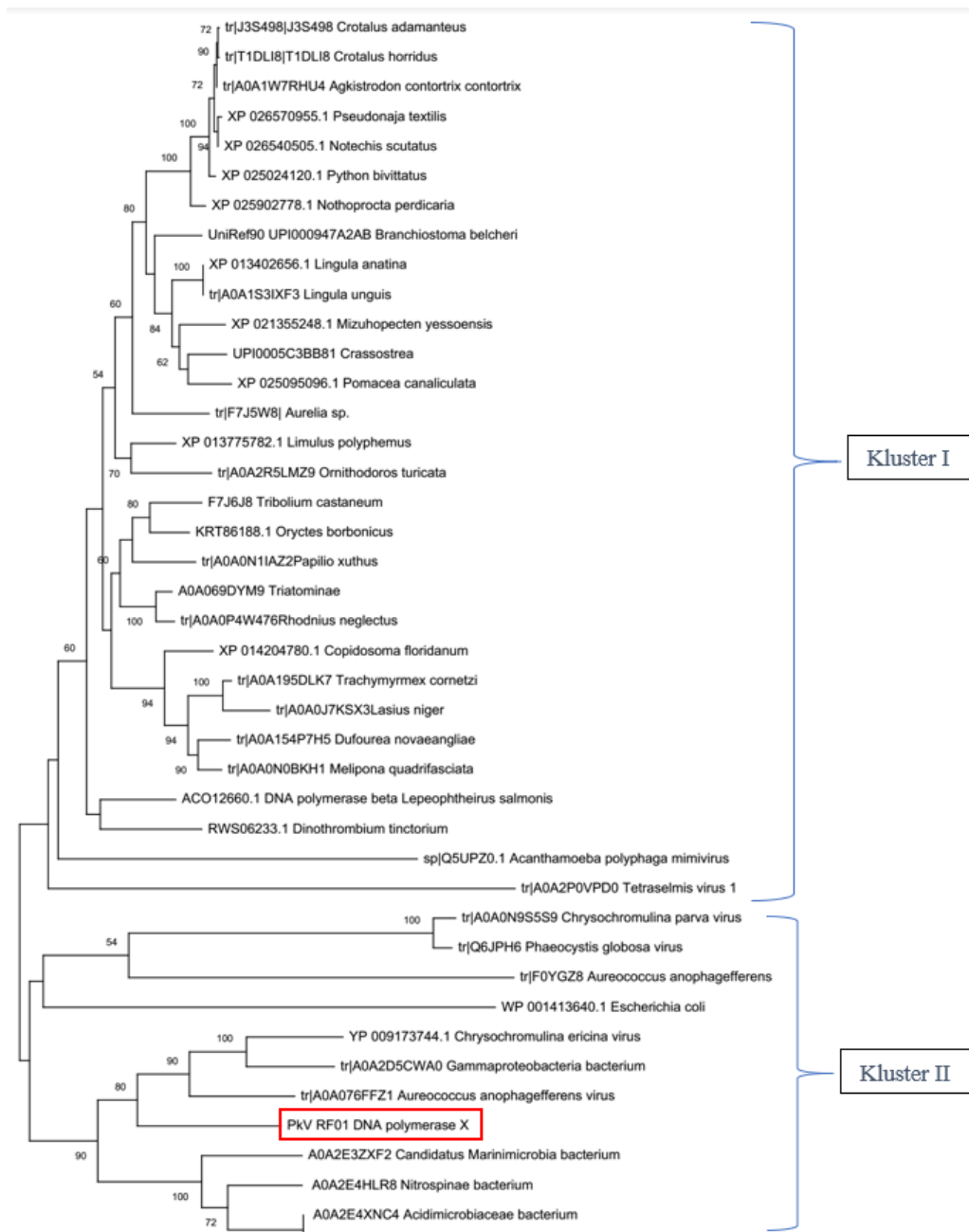
#### 3.8.1 Succinat dehydrogenase A



Figur 3.21

For å lage et fylogenetisk tre ble 164 homologe sekvenser til *sdhA* genet i PkV RF01 valgt, hvor 50 var TARA sekvenser, 50 var toptreff ved BLAST søk i UniProt og resten var andre homologe sekvenser for at flere taxa skulle bli representert. BLAST søk viser at toptreffene er eukaryote organismer. Av det fylogenetiske treet (Figur 3.22) ser man at *sdhA* til PkV RF01 danner en dyp forgreining i klusteret av *sdhA* gener sammen med TARA sekvenser.

### 3.8.2 DNA Polymerase X

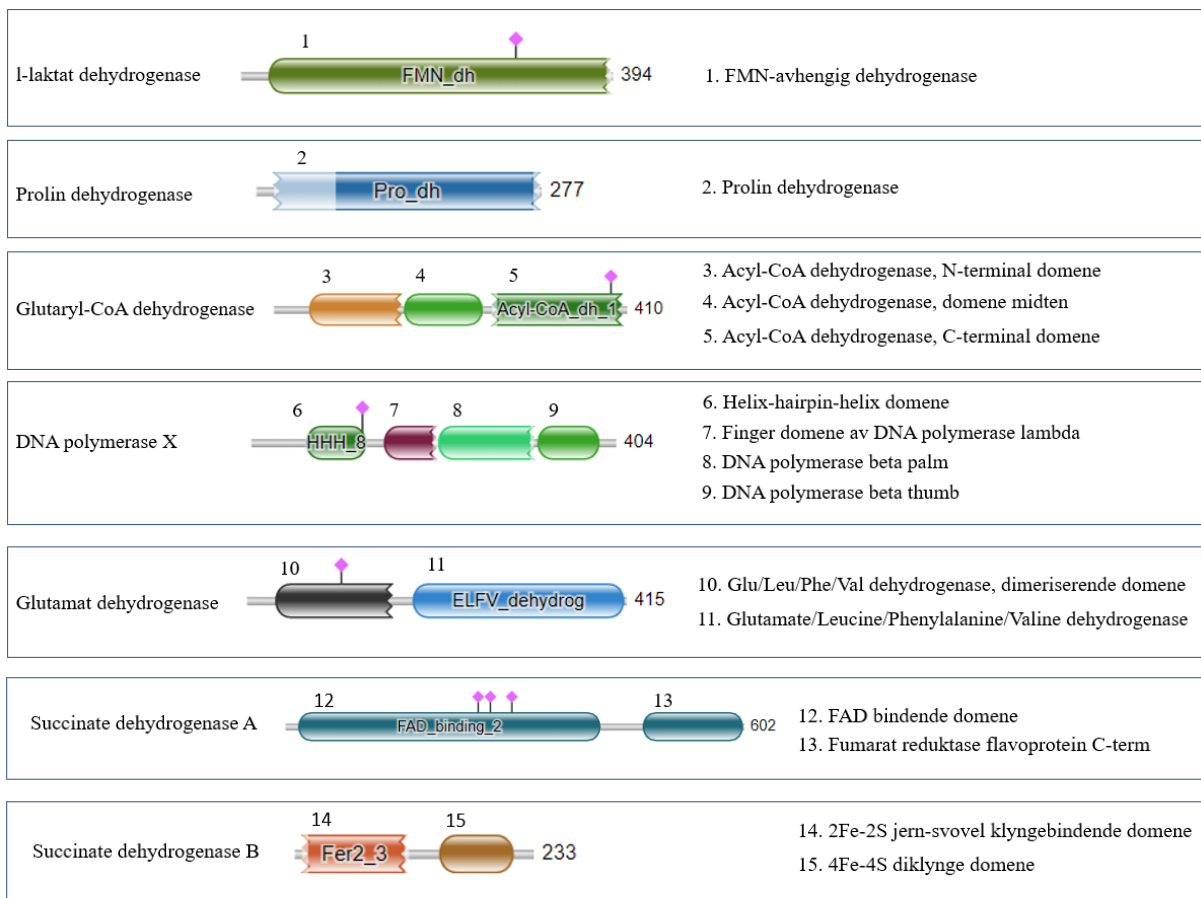


Figur 3.22: Fylogenetisk tre til DNA polymerase X i PkV RF01. Pol X fra PkV RF01 er markert med rødt.

For å lage det fylogenetiske treet ble de første 42 toptreffene fra BLAST søket brukt. Mange av disse treffene var eukaryote organismer, samt virus og bakterier. Det dannes to hovedkluster (Figur 3.21); kluster 1 mot eukaryote organismer og to virus, og kluster 2 mot virus, bakterier og alger. Pol X i PkV RF01 danner en gren ut fra de andre i kluster 2, hvor genet er nærmest beslektet med AaV.

### 3.8.3 Domene figurer for de sju proteinene

Analyse av domener i de syv proteinene i denne oppgaven ble utført ved å gjøre et HMMR søk ved EBI, og det ble laget en oversikt over domene til hvert protein (Figur 3.22). HMMR søket viser at L-LDH og PDH har et funksjonelt område (domene), hvor dette er et flavin mononukleotid (FMN)-avhengig dehydrogenase for L-LDH, og en prolin dehydrogenase for PDH. GCDH har tre domener, som er acyl-CoA dehydrogenaser. GDH, SDHA og SDHB har alle to funksjonelle områder, mens pol x har fire domener.



Figur 3.23: Viser domener til hvert protein etter et HMMR søk.

## 4 Diskusjon

### 4.1 Er SDHA og SDHB aktive proteiner i PkV RF01?

#### 4.1.1 Transkripsjonsforsøk

Som første steg i arbeidet med å finne ut om *sdhA* genet er funksjonelt i PkV RF01, ble det gjennomført et transkripsjonsforsøk. Det ble valgt å jobbe med subenheten SDHA da dette genet ligger i matriks og kan være enklere å jobbe med enn de membranbundne SDHC og SDHD. Når et gen er aktivt, vil det først bli transkribert og det blir dannet mRNA (Griffiths et al., 2000). Derfor må man isolere mRNA som er fritt for DNA for ikke å amplifisere DNA under PCR trinnet. For å undersøke om mRNA ble isolert, ble det utført RT-PCR og deretter en PCR-reaksjon. Under RT-PCR blir mRNA brukt som templat for å danne cDNA. Med spesifikke primere mot *sdhA* genet, ble RT-PCR produktet brukt som templat for å bevise at det var blitt isolert mRNA, og at genet dermed blir transkribert. For å kunne bevise transkripsjon av et gen må genomisk DNA bli fjernet under mRNA isoleringen ved hjelp av DNase I behandling, slik at det er cDNA som blir amplifisert, og ikke genomisk DNA.

Det ble flere ganger forsøkt å isolere mRNA for å bevise at *sdhA* genet blir transkribert under infeksjon uten å lykkes i å fjerne rester av genomisk DNA ved DNase I behandling av det isolerte mRNAet (Figur 3.6 og 3.7). Dette skjedde selv om mRNA for *sdhA* ble isolert med de parameterne som ble bestemt ved optimalisering av reaksjonen ved å benytte et gen som en vet blir transkribert (*mcp* genet) (Figur 3.4 og 3.5). Når det fortsatt er genomisk DNA i prøven, vet man ikke om det er DNA fra mRNA isoleringen eller cDNA fra RT-PCR trinnet som blir amplifisert under PCR. Dermed ble det vanskelig å få bekreftet om SDHA ble transkribert under infeksjon eller ikke.

Ifølge produsenten (Ambion) er kitet som ble benyttet i denne oppgaven ikke optimalisert for isolering av RNA fra virus. Triyatni et al. 2011 har brukt dette kitet da de isolerte RNA fra Hepatitis C virus (Triyatni et al., 2011). Samtidig fungerte metoden på kontrollgenet *mcp* hvor man klarte å isolere mRNA og fjerne rester av DNA ved DNase I behandling. *mcp* genet ble deretter transkribert ved RT-PCR, noe som ble bevist ved hjelp av PCR med spesifikke primere mot *mcp* genet. En mRNA kontroll ble inkludert for å verifisere at protokollen fungerte, og her fikk man gode resultater. Dette tyder da på at problemet ligger i genet av interesse, og ikke protokollen eller kitet.

Kitet ble kjøpt inn til bruk i denne oppgaven, og var dermed ikke tidligere åpent. Et alternativ til hvorfor det fortsatt var DNA i prøven er at det skjedde en DNA kontaminasjon ved bruk av reagensene. Ved samtlige forsøk ble de samme reagensene brukt, og dermed er faren for kontaminasjon tilstede. Men samme reagenser ble benyttet for optimaliseringen av protokollen ved hjelp av *mcp* genet, og i dette tilfellet var det ikke tegn på DNA i prøven, som betyr at man kan utelukke DNA kontaminasjon. RNA er veldig ustabil og kan brytes ned raskt om ikke forhåndsregler blir tatt. Grunner til at RNA brytes ned kan være aktivt RNase i prøven, og disse må bli inaktivert under prøvebehandlingen. I tillegg kan prøver bli kontaminert med ytterligere RNase fra omgivelser under prøvebehandlingen. Det viktigste tiltaket man gjør for å ikke kontaminere prøven med RNase er å jobbe i RNase frie områder, at pipetter, annet utstyr og reagenser ikke inneholder RNase. For å unngå dette ble det brukt RNaseZAP for å fjerne eventuelle RNase på benk, pipetter og annet utstyr som ble brukt.

Siden ikke alt genomisk DNA ble brutt ned ved DNase I behandling, er det vanskelig å si noe om hvorvidt genet blir transkribert under infeksjon eller ikke, da man ikke kan utelukke at det er genomisk DNA som blir amplifisert. Forsøket ble gjentatt flere ganger med ulike temperaturer og inkuberingsstider. Hovedproblemet ligger i kontaminering av DNA i RNA isoleringstrinnet, og hvordan man skal få fjernet DNAet uten å bryte ned RNA. En mulighet for dette er å benytte seg av et annet kit for isolering av RNA (Monier et al., 2017).

Det har i senere tid blitt gjennomført et nytt transkripsjonsforsøk for *sdhA* genet i PkV RF01 hvor det ble brukt et annet RNA-isoleringskitt i kombinasjon med DNase I behandling av det isolerte RNAet før RT-PCR trinnet (samme kitt som i Monier et al., 2017). Her ble alt genomisk DNA fjernet fra RNA isoleringstrinnet, og PCR med spesifikke *sdhA* primere viste at *sdhA* genet har blitt transkribert under infeksjon (samtale med Ruth-Anne Sandaa, UiB, 2019).

#### 4.1.2 Ekspresjon av SDHA og SDHB

I tillegg til transkripsjonsforsøket ble det bestemt at SDHA og SDHB skulle forsøkes å klones for å finne ut av om genene lar seg uttrykke i *E. coli* og dermed er funksjonelle i PkV RF01. Første steg var å designe primere med overheng, og bruke disse for å amplifisere genet og tilhørende vektor. Amplifiseringen av genene ble ikke vellykket verken for SDHA eller SDHB. Ved amplifiseringen av vektoren til SDHA fikk man flere uspesifikke bånd, og ved amplifiseringen av SDHB fikk man ingen bånd som tyder på at genet ikke har blitt amplifisert (Figur 3.11). En grunn til dette kan være at primerne ikke har blitt designet riktig slik at man kan få uspesifikke bindinger på genet. En annen grunn kan det høye A + T innholdet i PkV RF01, noe som har blitt bevist for flere algevirus som tilhører familien *Mimiviridae* (Gallot-

Lavallée et al., 2015). Blant annet har CeV 01B et G + C innhold på ca. 20 % (Gallot-Lavallée et al., 2015). Ved en typisk PCR-reaksjon har man ett polymeriseringssteg på 72 °C etter denatureringen av dobbeltråden og hybridiseringen av primerne. Ved denne temperaturen vil termostabile polymeraser replikere DNAet (Gurr, 1991). Men om man bruker flankerende primere, slik som med amplifisering av genene i denne oppgaven, er det ofte at disse har veldig høyt A + T innhold, ofte over 90 % (Su et al., 1996) og man vil da kunne få problemer med å amplifisere genet, da temperaturen kan være for høy og DNAet kan smelte som vil gjør at polymerasen ikke kan replikere DNAet (Su et al., 1996). Samtidig er det ofte slik at store DNA fragmenter (> 5 kb) også forventes å ha områder med > 90 % A + T innhold. I Su et al. 1996 viser de at ved å senke polymeriseringsstemperaturen til 60 °C klarer man likevel å amplifisere DNA med høyt A + T innhold, og at denne temperaturen er anbefalt å bruke om man har områder med veldig høyt A + T innhold (Su et al., 1996).

#### 4.2 Når i infeksjonssyklusen til PkV RF01 blir SDHA eventuelt uttrykt?

For å undersøke når i infeksjonssyklusen til PkV RF01 SDHA blir uttrykt, var det planlagt å utføre ddPCR for kvantifisering av PCR produktet. Dette ble ikke utført da det ikke kunne bevises at *sdhA* genet ble transkribert. Om transkripsjonsforsøket hadde vært vellykket, skulle ddPCR bli brukt for å bestemme når i infeksjonssyklusen genet ble transkribert og hvor mange kopier av genet man hadde ved ulike tidspunkt under infeksjonen. På denne måten kunne man sagt noe om når genet var aktivt, om det var mer aktivt tidlig i infeksjonssyklusen, eller om mørke/lys hadde en innvirkning på genets aktivitet. Da man vet at dette er et gen som inngår i trikarboksylsyresyklusen og den oksidative fosforyleringen, kan man anta at dette er et gen som vil bli uttrykt tidlig i infeksjonssyklusen. Disse prosessene er en viktig del av respirasjonen til eukaryote organismer, og er prosesser som cellene vil holde i gang kontinuerlig, men når virus infiserer verten er det vanlig at verten går i forsvarsmodus. Og dermed er det mulig at viruset benytter dette genet og resten av komplekset for å øke eller opprettholde metabolismen til verten gjennom infeksjonen (Blanc-Mathieu, ikke publisert).

Det ble gjennomført en vellykket optimalisering av ddPCR, med spesifikke SDHA primere, hvor det ble undersøkt hvilke primer konsentrasjoner og fortyninger av templat som ga best resultat. Grafen (Figur 3.8) viser at primer konsentrasjoner og fortyninger gir veldig likt resultat, men ved 1:1000 templatfortynning og 100 nM primer konsentrasjon får man flest kopier

(1148000 kopier/ $\mu$ l). Dette kan tyde på at det er disse parameterne som bør brukes når ddPCR skal benyttes for å undersøke når i infeksjonssyklusen SDHA blir uttrykt.

#### 4.3 Hva er den evolusjonistiske opprinnelsen til *sdhA* i PkV RF01?

Sdh komplekset finnes i alle levende organismer som bruker oksygen som en del av celleåndingen. Samtidig har studier blant annet vist at kompleks II har en viktig rolle i den anaerobe energimetabolismen til parasitter, hvor den fungerer som en quinol-fumarat reduktase (Kita et al., 2002). *sdhA* genet funnet i PkV RF01 vet man har blitt overført fra andre organismer, da dette er et gen som virus normalt ikke har. Samtidig kan man anta at dette er et gen som viruset har plukket opp fra verten, og at det dermed kan ha opprinnelse fra eukaryote organismer, noe som toppreffene på BLAST søket bekrefter. Av det fylogenetiske treet ser man at PkV RF01 danner en dyp forgreining i det eukaryote klusteret av *sdhA* gener, noe som tyder på at SDHA i PkV RF01 er nærmest beslektet eukaryote organismer, men at det ikke tilhører en kjent gruppe eukaryote organismer (Figur 3.21)

Tidlig i den eukaryote evolusjonen og raskt etter den endosymbiotiske organisering av mitokondriet, ble mange mitokondrie gener overført til nukleus (Adams et al., 2001). I eukaryote organismer som har blitt beskrevet er SDHA ikke tilstede i mitokondrie (Lang et al., 1999), noe som tyder på at overføring til nukleus fant sted tidlig evolusjonen hos eukaryote organismer. De andre subenhetene har i få tilfeller blitt funnet i mitokondriet (Gray et al., 1998), men for alle beskrevne dyr og sopp arter er de fire subenhetene plassert i nukleus (Boore, 1999; Lang et al., 1999). Da alle genene i SDH kompleks II har blitt funnet samlet i PkV RF01, kan dette tyde på at viruset har tatt opp dette genet før genoverføringen til nukleus, noe som kan tyde på at det er tatt opp tidlig i evolusjonen muligens før den endosymbiotiske organiseringen av mitokondriet.

#### 4.4 Er det mulig å uttrykke utvalgte proteiner fra PkV RF01 i *E. coli*?

Det finnes mange systemer tilgjengelig for protein ekspresjon, og *E. coli* er den verten som har blitt mest vanlig å bruke de siste årene, da den har mange fordeler (Sezonov et al., 2007). Ved å bruke *E. coli* som vert har man mulighet til å lage kulturer som vokser raskt og gir høyt utbytte i et medium som er rimelig. Samtidig som man har tilgang på et høyt antall kloningsvektorer og vertsstammer (Sivashanmugam et al., 2009). Tidligere har blant annet *mcp* genet fra *Emiliana huxleyi* blitt klonet og uttrykt ved bruk av *E. coli* stammer (Liu et al., 2013).



Ekspresjon og proteinstabilitet avhenger av valg av vektor, *E. coli* stamme og dyrkningsparametere (McCormick et al., 2014). I dette eksperimentet ble det valgt *E. coli* stammer som har vært mye brukt i proteinarbeid, hvor blant annet BL21(DE3) stammene er de mest brukte for ekspresjon av protein (Gottesman, 1996; Grodberg and Dunn, 1988). Samtidig ble det brukt et vektorsystem som benytter seg av T7 *lac* promotor og ampisillin resistens som gir en stabil ekspresjon av rekombinert protein (McCormick et al., 2014).

For at målgenet skal bli plassert korrekt i vektoren, ble det designet primere med overheng. Disse ble brukt til å amplifisere opp målgen og tilhørende vektor før kloning. Amplifiseringen av målgen og vektor ble vellykket, som da viser at primerne og overhenget på disse var korrekt designet (Figur 3.9 og 3.10). Som bekreftelse på at målgenet var plassert korrekt i vektoren, og at dette plasmidet var blitt korrekt overført til *E. coli* NEB-5 alfa kompetente cellene ble det utført en koloni-PCR på de utvalgte koloniene for hvert gen. Resultatet av koloni-PCR viste at de valgte *E. coli* koloniene hadde det rette plasmidet (Figur 3.13 og 3.14). Deretter ble det rensede plasmidet overført til *E. coli* BL21-GOLD (DE3) celler, og disse koloniene ble undersøkt ved koloni PCR. For fire av genene fikk man amplifisering ved de valgte koloniene, som tyder på at disse hadde det korrekte plasmidet. For den valgte kolonien til l-laktat dehydrogenase (L-LDH) fikk man ingen bånd på gelen, som tyder på at det ikke har blitt overført det korrekte plasmidet til denne kolonien. Ved å studere LB-skålene (figur 3.15) ser man at noen av koloniene ligger litt tett. Når man skal velge kolonier er det viktig å velge små, isolerte kolonier for å sikre at man plukker bare av den ene kolonien og ikke av en klase av kolonier. For å få bedre distribusjon av kolonier kan man velge en isolert koloni og stryke denne ut på nytt eller velge en kortere inkuberingsstid slik at man får mindre vekst på skålen (McCormick et al., 2014). I dette tilfellet med L-LDH er det stor sannsynlighet for at den kolonien som ble valgt manglet det rette plasmidet, da det ble positivt resultat når det ble plukket ut seks nye kolonier. Som en ekstra sikkerhet for at rett plasmid har blitt overført kunne man gjennomført Sanger sekvensering på koloniene (Sanger et al., 1977).

Da det ble bekreftet at alle de fem proteinene hadde kolonier med det korrekte plasmidet, ble det bestemt at ekspresjonsforsøk skulle utføres for alle fem proteinene. Dette ble gjennomført i småskala først for å finne ut hvilke proteiner som var løselige (diskuteres i avsnitt 4.5). Basert på resultatet fra SDS PAGE (Figur 3.16) ser det ut som alle fem proteinene har blitt uttrykt i *E. coli* da man har fått bånd med rett størrelse, og det ble valgt å jobbe videre med tre proteiner (pol X, L-LDH og GCDH) basert på løseligheten til proteinene (se avsnitt 4.5). Pol X ble vellykket rensset hvor man fikk et høyt utbytte (697 µg/ml ved 1 l kultur), noe som dermed

bekrefter at dette proteinet har blitt uttrykt i *E. coli*. L-LDH og GCDH ble ikke vellykket rensset, noe det kan være flere grunner til. Dårlig ekspresjon av et protein kan ofte være forårsaket av protein toksisitet, inaktivt protein, ukontrollert basal ekspresjon, kodon bias og inklusjonslegemer (Rosano and Ceccarelli, 2014). Samtidig som at faktorer som antibiotika og IPTG konsentrasjoner, induseringstid og temperatur kan påvirke ekspresjonen av proteinet (Graslund et al., 2008; Sivashanmugam et al., 2009).

Protein toksisitet er den viktigste grunnen til problemer med kloning og ekspresjon (Dumon-Seignovert et al., 2004; Miroux and Walker, 1996) da proteinet kan ha en unødvendig og skadelig funksjon som påvirker vertens normale proliferasjon og homeostasen, som igjen vil føre til lavere vekstrate og celletetthet (Doherty et al., 1993; Dong et al., 1995). For å kunne utelukke protein toksisitet burde man ha inkludert en kontrollvektor uten proteinet, slik at man kunne ha gjennomført en sammenligning av vekst og ekspresjon av de to kulturene (Saïda et al., 2016). Proteinet kan være skadelig for verten både før og etter indusering (Rosano and Ceccarelli, 2014). Før indusering vil man kunne oppdage toksisiteten ved at det blant annet ikke dannes kolonier på LB-skålene, man har få positive kloner eller at målgenet ikke er satt inn rett vei i vektoren. Samtidig får man en lav transformasjon effektivitet og vekstraten i kulturene er lave (Doherty et al., 1993). Protein toksisitet kan gi en lav eller ingen ekspresjon samtidig som ekspresjonen kan være veldig varierende (Dong et al., 1995). Siden koloni-PCR viser at kloning og transformasjon av disse to proteinene var vellykket, samtidig som man har godt vekst på LB-skålene, er det liten sannsynlighet for at det var problemer med toksisitet for verten på dette stadiet. Det er vanskelig å si om proteinene omtalt her er skadelig for verten etter indusering eller ikke uten å gjøre mer arbeid, som blant annet å inkludere en kontrollvektor og kontrollkultur. Men det kan tyde på at problemet med ekspresjonen av proteinene dukker opp etter induseringen med IPTG. Hvis det er protein toksisitet som påvirker ekspresjonen kan man bruke andre *E. coli* stammer som er tilpasset til skadelige proteiner, som for eksempel *E. coli* C41(DE3) eller C43(DE3) (Miroux and Walker, 1996). En annen løsning er å fjerne proteinet fra cellen, og flytte det til periplasmaet eller mediumet ved å bruke et signal peptid (de Marco, 2009; Mergulhão et al., 2005)

Proteiner kan ofte ha kodoner som ikke vanligvis blir brukt av verten, dette kan føre til at proteiner kan være vanskelig å uttrykke utenfor deres egne celler (Gustafsson et al., 2004). Dette kan føre til kodon bias, som forekommer når synonyme kodoner i det fremmede DNA er signifikant forskjellig fra vertens (Rosano and Ceccarelli, 2014). Når man har full syntese av det rekombinante proteinet, er det vanlig at det blir mangel på tRNA. Dette kan føre til

manglende inkorporering av aminosyrer og at polypeptider blir kuttet, som igjen vil påvirke ekspresjonen av proteinet (Gustafsson et al., 2004). For *E. coli* blir slike uvanlige kodoner definert som kodoner brukt av *E. coli* ved en frekvens  $< 1\%$  (Kane, 1995). For å unngå kodon bias slik at dette ikke påvirker ekspresjonen av proteinet, kan man enten foreta en kodon optimalisering eller øke tilgangen til tRNA (Sørensen and Mortensen, 2005). Målet med kodon optimalisering er å modifisere kodonene som er uvanlige i verten slik at disse kan bli mer lik de som verten bruker (Burgess-Brown et al., 2008), men det er viktig å ikke forandre aminosyre sekvensen i denne prosessen (Menzella, 2011). Når det blir produsert store mengder med protein, vil ikke tilgjengeligheten av tRNA være nok og dette kan ofte bli den begrensende faktor (Pedersen, 1984). Mangel på tRNA kan forårsake produksjon av uferdig produkt (Buchan and Stansfield, 2007). For å ha tilgang på nok tRNA finnes det flere plasmider som inneholder kopier av tRNA gener som kan bli brukt for å forhindre dette, som blant annet stammen BL21(DE3) codonPlus som har plasmidet pRIL.

Når ekspresjonen av proteinet er lav så kan utbytte økes ved å endre noen parametere under dyrkingen av *E. coli*. LB er det vanligste mediumet å bruke når man dyrker *E. coli* fordi det er enkelt å lage mediumet og det er rikt på næringsstoffer, samtidig som det gir optimal vekst (Rosano and Ceccarelli, 2014). Om man får problemer med celleveksten ved bruk av LB medium kan man øke mengden av pepton eller gjærekstrakt, og man kan supplere med kationer som  $MgSO_4$  som vil resultere i høyere cellevekst (Studier, 2005). Hvis LB medium ikke gir tilfredsstillende utbytte kan man vurdere å bruke medium som blant annet 2YT, TB eller SB, som inneholder større mengder med næringsstoffer som vil gi høyere celletetthet (Atlas, 2004; Studier, 2005). I Studier, 2005 presenteres konseptet autoinduseringsmedium, som består av glukose, laktose og glyserol (Studier, 2005). Her brukes glukose som karbonkilden og som brytes ned under vekst slik at laktosen ikke blir tatt opp før det er mangel på glukose. Når laktose og glyserol blir tatt opp, vil laktose fungere som en inducer for *lac* operon (Studier, 2005). På denne måten er det ikke nødvendig å overvåke kulturen og man kan unngå å tilsette IPTG eller andre inducere ved feil tidspunkt (Studier, 2005).

Da det kan se ut som at det ble dannet bånd med rett størrelse ved SDS PAGE analysen for alle proteinene som ble forsøkt uttrykt i *E. coli*, kan dette tyde på at alle proteinene kan uttrykkes i *E. coli*. For å være sikker på at dette stemmer, bør man klare å rense proteinet, og da dette bare ble vellykket for pol X kan man ikke være sikker for de fire andre proteinene. For å kunne si noe om ekspresjonen av disse proteinene bør man prøve å gjøre ekspresjonsforsøk hvor man tar hensyn til faktorene nevnt ovenfor.

#### 4.5 Er de uttrykte virus proteinene løselige?

For å kunne bestemme om proteinene var løselige, ble det laget til to forskjellige prøver som ble undersøkt ved SDS PAGE, en metode som separerer protein basert på molekylvekt. Den ene prøven inneholdt alle proteinene (tatt direkte fra lysatet), mens den andre ble sentrifugert og dermed inneholder supernatanten bare de løselige proteinene, mens pellet inneholder resten av cellene, membraner og inklusjonslegemer (Duong-Ly and Gabelli, 2014). Ved å få bånd med korrekt størrelse i begge prøvene kunne man bekrefte at man har protein i prøven, og om disse var løselige. Av resultatet ser man at for alle proteinene har det blitt dannet et bånd med rett størrelse basert på den predikerte molekylvekten. Dette kan bety at alle proteinene har blitt uttrykt. For noen av proteinene var det vanskelig å skille om båndet var fra det induerte proteinet eller fra et protein tilhørende *E. coli* da det ser ut som de kan ha lik størrelse rundt ca. 50 kDa. Dette gjelder tre proteiner; GCDH, GDH og L-LDH. Hvis man studerer båndene til proteinene som hører til *E. coli* ser man i prøven til PDH og pol X at det har blitt dannet et bånd med molekylvekt rundt 50 kDa, men disse båndene tilhører ikke PDH og pol X da de har en annen molekylvekt. Dette kan tyde på at et protein som tilhører *E. coli* har samme størrelse som noen av induerte proteinene. Dette gjør det vanskelig å vite hvilket protein båndet tilhører. Samtidig ser man at båndene i prøvene til de tre proteinene er mye større enn de som er i prøvene til PDH og pol X. Basert på denne størrelsesforskjellen, og dermed forskjell i ekspresjon kan dette tyde på at det proteinet som har blitt uttrykt i prøvene til GCDH, GDH og L-LDH ikke er et protein tilhørende *E. coli*, men de induerte proteinene da man forventer en høyere ekspresjon av disse. For disse tre proteinene har man et bånd med korrekt størrelse både i den totale og den løselige prøven. Båndene i den totale prøven er mye større enn for de i den løselige prøven, noe som kan tyde på at de bare delvis er løselige. For pol X blir det dannet bånd både i den totale prøven og den løselige prøven, og disse er ikke tilstede i noen av prøvene til de andre proteinene, som dermed tyder på at dette tilhører pol X fra PkV RF01. For PDH har man et bånd i den totale prøven, men som er veldig svakt synlig i den løselige prøven. Dette kan bety at dette proteinet ikke er løselig, eller at det ikke har blitt produsert store mengder av det løselige proteinet grunnet faktorer nevnt nedenfor. Basert på disse resultatene ble det bestemt at pol X, GCDH og L-LDH skulles forsøkes å renses da resultatet tydet på at disse var størst sjans for at var løselige. Rensingen av pol X ble vellykket, mens GCDH og L-LDH ikke ble vellykket, noe som tydet på at de ikke var løselige.

Løseligheten til et protein påvirkes av aminosyresammensetningen og hvilke polare og ikke-polare grupper som finnes i aminosyrene. Samtidig blir løseligheten til protein påvirket av

miljøfaktorer som pH, ionestyrke, buffer, temperatur og hvordan prøvene blir behandlet (Zayas, 1997). Disse faktorene kan føre til ustabile proteiner og aggregering (Hartley and Kane, 1988), som igjen kan føre til at proteinet ikke er løselig. Under ekspresjon er det flere faktorer som kan påvirke løseligheten til proteiner. Ved å senke produksjonshastigheten til proteiner, vil transkriberte proteiner få tid til å folde seg riktig og man får mindre aggregering av proteiner (Graslund et al., 2008). Dette kan gjøres ved å senke inkuberingstemperaturen og IPTG konsentrasjonen (Zayas, 1997). Hvis man har problemer med inklusjonslegemer viser flere studier at ved å senke inkuberingstemperaturen til mellom 15-25 °C vil dette gi en bedre produksjon av løselige proteiner (Liu et al., 2013; Vera et al., 2007). Det er vanlig å bruke IPTG konsentrasjoner mellom 0,05-2 mM, men studier viser at en lavere IPTG konsentrasjon (rundt 0,1 mM) vil gi en mer korrekt folding av proteinet (Picaud et al., 2007; Terpe, 2006). Når disse faktorene endres, er det viktig å huske på at celleveksten er lavere, og at man dermed trenger en lenger induseringstid. I Liu et al, 2013 beskriver de hvordan *mcp* genet til *E. huxleyi* virus blir uttrykt i *E. coli*. Her hadde de problemer med at mesteparten av proteiner var uløselige, men ved å optimalisere ekspresjonsforholdene klarte de å øke produksjonen av løselige proteiner. Ved å indusere proteinet med 0,3 mM IPTG ved 20 °C i 8 timer oppnådde de den høyeste produksjonen av løselig protein (Liu et al., 2013). Som nevnt tidligere ble IPTG konsentrasjon, induseringstid og medium endret for å se om dette ville påvirke ekspresjon og løselighet for proteinene uten å se en betydelig forskjell. Temperaturen som ble brukt var 37 °C frem til indusering, og temperaturen ble da senket til 30 °C. Basert på nevnte studier, bør man prøve å uttrykke proteinene ved lavere temperaturer for å se om dette vil utgjøre en forskjell.

En viktig faktor for at protein skal være løselig er pH til buffer. Graden av løseligheten baserer seg på elektrostatiske og hydrofobe interaksjoner mellom proteinmolekylene. Løseligheten øker hvis den elektrostatiske avstøtningen mellom molekylene er høyere enn de hydrofobe interaksjonene (Kinsella and Melachouris, 1976). Det er flere faktorer i bufferløsningen som kan ha en innvirkning på løseligheten til proteinene. For at proteiner skal være løselige bør ikke løsningene ha en pH som tilsvarer proteinets pI. Ved det isoelektriske punktet, vil proteiner ha en nettoladning lik null, og da vil molekyler ha en tendens til å ville binde seg sammen som resulterer i uløselige proteiner (Zayas, 1997). For å få høyest mulig ekspresjon av proteiner som er løselige bør pH være lavere enn pI. Da vil det bare være positive og nøytrale grupper som ikke vil kunne binde seg til hverandre. Siden proteinene er HIS-merkede for rensing, er dette også viktig å ta i betraktning når man velger pH. pKa (syrekonstanten) til histidin er 6,5, noe som gjør at pH til buffer bør være høyere enn 6,5 for å ikke ha en innvirkning på histidin

affiniteten under rensing (Bornhorst and Falke, 2000). Hvis det er mulig er det vanlig å velge en pH som ligger rundt 7,5, da dette er nærme de fysiologiske forholdene til cellen. Under ekspresjon og rensing av proteinene i denne oppgaven ble det valgt pH til buffer basert på pI til hvert protein. For alle proteinene utenom GDH, ble det valgt en pH på 7,5. For GDH er pI 7,98 og en buffer med pH 7,5 ville da være for nært pI til proteinet, og kunne ha påvirket løseligheten. Dermed ble det bestemt å lage en lysisbuffer med pH 6 for GDH. Det er viktig å tilsette NaCl til bufferen for å holde proteinet løselig og for å etterligne fysiologiske forhold i cellen. NaCl er med på å redusere uspesifikke hydrofobe protein interaksjoner (Bornhorst and Falke, 2000). Det er anbefalt å bruke en konsentrasjon mellom 300–500 mM (Graslund et al., 2008). I denne oppgaven ble det brukt en økende konsentrasjon av salt under rensingen for å forhindre uspesifikke bindinger mellom proteiner og affinitetskolonnen. I lysisbuffer og lagringsbuffer ble det tilsatt glyserol, da dette holder proteinet stabilt. Protein aggregering vil bli hemmet, samtidig som at proteinet vil få en mer kompakt tilstand ved tilsetning av glyserol (Vagenende et al., 2009).

For å oppnå den aktive tredimensjonale konformasjonen er dannelsen av disulfidbindinger avgjørende for mange rekombinante proteiner. Om disse ikke dannes riktig kan dette føre til at proteinet ikke folder seg rett eller at det dannes inklusjonslegemer. I *E. coli* blir disulfidbindingene dannet i periplasmaet, noe som påvirker produksjonen av rekombinante proteiner med disulfidbindinger, da disse ofte blir uttrykt i cytoplasma (Derman et al., 1993). For at disulfidbindinger skal dannes, kan man overføre proteinet til periplasmaet eller man kan bruke stammer av *E. coli* hvor disulfidbindinger dannes i cytoplasmaet (Bessette et al., 1999; Derman et al., 1993).

Det er vanskelig å si noe om hvorfor ikke alle proteinene i denne oppgaven kunne renses, men som nevnt tidligere tyder det på at proteinene har blitt uttrykt i *E. coli*, og at det dermed er løseligheten til proteinene som er problemet. Grafen som viser renseforløpet til L-LDH og GCDH detekterer en liten mengde HIS-merkede proteiner, men det dannes ikke bånd ved SDS PAGE analysen. En grunn til dette kan være lav ekspresjon av proteinet, og at det dermed ikke er nok løselig protein i prøven til å gi et positivt resultat. Grunnen til dette kan være at proteinene bare er delvis løselige, blant annet hvis det dannes inklusjonslegemer eller hvis noen av de andre faktorene nevnt har en innvirkning på ekspresjonen av de løselige proteinene. For å kunne si noe om hvorvidt disse har en innvirkning på proteinet må det gjøres mer arbeid på disse proteinene. Hvis noen av proteinene er membranbundne vil dette kunne påvirke ekspresjonen av proteinene da det kan være vanskelig å uttrykke slike proteiner. Noen av proteinene ble ikke

forsøkt rensset selv om det kunne tyde på at de var løselige, da man måtte gjøre et valg på grunn av tiden til rådighet. Pol X var det eneste proteinet som vellykket ble uttrykt og rensset, men selv om man har klart å uttrykke proteinet, betyr ikke dette at proteinet er aktivt. Proteinene kan ha en ufullstendig folding, hvor proteinet vil kunne være i en stabil løselig form, samtidig som det aktive setet fortsatt er uegnet for aktivitet. Derfor må proteinet karakteriseres og undersøkes nærmere.

#### 4.6 Hva er den evolusjonistiske opprinnelsen til DNA polymerase X i PkV RF01?

DNA polymerase X i PkV RF01 danner en gren med høy bootstrappingverdi i klusteret av virus og bakterier (Figur 3.22), som viser at man kan stole på plasseringen av genet i dette klusteret. Basert på plasseringen av genet i det fylogenetiske treet laget i denne oppgaven ser det ut til at pol X i PkV RF01 kan være nærmest beslektet til andre virus og bakterier, og nærmest til AaV. Man vet at mange virale gener har en opprinnelse fra eukaryote organismer, men samtidig finnes det virusgener som har en opprinnelse fra prokaryote organismer. Blant annet vet man at *MutS7* genet som har blitt beskrevet i *Mimiviridae*, sannsynligvis har en opprinnelse fra Epsilon proteobakterier (*Arcobacter* og *Sulfurospirillum*) (Claverie and Abergel, 2018). Basert på det fylogenetiske treet til pol X, og at kluster II består av blant annet bakterier er det en mulighet for at pol X i PkV RF01 kan ha en opprinnelse fra prokaryote organismer.

Basert på sammenligninger gjort med Clustal Omega, hvor sekvensen til pol X ble sammenlignet med DNA ligase sekvenser fra *E. coli* og mimivirus (WP\_001413640.1 og Q5UPZ0.1) ser det ut som genet kan ha en ligase funksjon i tillegg til polymerase funksjonen. Dette har tidligere vært beskrevet i Fernández-García et al., 2017, hvor de har beskrevet to gener som tilhører virus i NCLDV gruppen som har et DNA ligase gen festet til C-terminalen av det PolX kodede området. De fleste virusene i NCLDV er i liten grad avhengig av vertens replikasjon- og transkripsjonssystemer for å gjennomføre deres egen replikasjon (Van Etten, 2003). Dette er fordi disse virusene har flere konserverte gener, som DNA polymeraser, helikaser og primaser, som er essensielle for virus replikasjon. Disse genene har mest sannsynlig blitt ervervet fra forskjellige cellulære organismer (Yutin and Koonin, 2012).

## 5 Konklusjon og videre arbeid

I denne oppgaven var ett av målene å kunne bevise at *sdhA* genet i PkV RF01 blir transkribert under infeksjon av He028. Dette ble ikke vellykket, da man ikke klarte å fjerne alt av DNA fra det isolerte mRNAet ved DNase I behandlingen. Ved å bytte kit og legge til et ekstra DNase I trinn har man i senere tid klart å isolere DNA fritt mRNA, og bevise at *sdhA* genet blir transkribert under infeksjon. Da det ikke ble bevist i denne oppgaven at *sdhA* genet blir transkribert under infeksjon, ble det ikke gjennomført ddPCR for å kvantifisere PCR produkt. Fylogenetiske analyser viser at *sdhA* genet har en opprinnelse fra eukaryote organismer, men at det ikke tilhører en kjent gruppe av eukaryote organismer. Det tyder på at hele kompleks II har blitt tatt opp av PkV RF01 tidlig i evolusjonen av eukaryote organismer. Veien videre vil bli å undersøke når dette genet blir uttrykt i infeksjonssyklusen ved hjelp av ddPCR, og eventuelt hvilken funksjon dette har hos PkV RF01. Samtidig er det av stor interesse å gjøre arbeid på de andre subenhetene (SDHB, SDHC og SDHD) for å finne ut om hele SDH kompleks II blir transkribert under infeksjon, og hva funksjonen til kompleks II er i PkV RF01. Basert på hvilken funksjon dette komplekset har i eukaryote organismer, kan man anta at viruset eventuelt bruker det for å øke eller opprettholde vertsmetabolismen under infeksjon.

Det ble forsøkt å klonere, uttrykke og rense fem auksiliære metabolske virusproteiner som kan ha potensiale innen bioprospektering. Basert på resultatet etter ekspresjon, tyder det på at alle fem proteinene har blitt uttrykt i *E. coli*, da man ved SDS PAGE analyse fikk bånd med korrekt størrelse for alle de fem proteinene. Av disse ble proteinene L-LDH, GCDH og pol X forsøkt rensset, hvor L-LDH og GCDH ikke ble vellykket, mest sannsynlig fordi de bare er delvis løselige. DNA polymerase X ble vellykket rensset og sendt til Tromsø for karakterisering. Dette er et enzym som er viktig i forbindelse med DNA reparasjon i virus, og kan ha stort potensial innen bioprospektering da det er mulighet for at det kan benyttes innen molekylærbiologien. Bioinformatiske analyser viser at pol X i PkV RF01 kan ha en opprinnelse fra prokaryote organismer da nærmest beslektede er bakterier og virus. Veien videre bør være å forsøke på nytt å uttrykke og rense, deretter eventuelt karakterisere de fire proteinene som ikke ble vellykket. Da flere av de ikke har blitt beskrevet i virus tidligere, er det av stor interesse å finne ut hvilken funksjon disse har i PkV RF01.

Viruser i gruppen NCLDV har store genomer og dermed mye genetisk materiale som gjør de attraktive for marin bioprospektering. Det er fortsatt mye arbeid som gjenstår før man har full forståelse for funksjonene til alle genene i PkV RF01, da 70-80 % av genene fortsatt er definert som hypotetiske gener.



## 6 Referanser

- Abergel, C., Legendre, M., Claverie, J.-M., 2015. The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiol. Rev.* 39, 779–796. <https://doi.org/10.1093/femsre/fuv037>
- Abrahão, J., Silva, L., Silva, L.S., Khalil, J.Y.B., Rodrigues, R., Arantes, T., Assis, F., Boratto, P., Andrade, M., Kroon, E.G., Ribeiro, B., Bergier, I., Seligmann, H., Ghigo, E., Colson, P., Levasseur, A., Kroemer, G., Raoult, D., Scola, B.L., 2018. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* 9, 749. <https://doi.org/10.1038/s41467-018-03168-1>
- Adams, K.L., Rosenblueth, M., Qiu, Y.L., Palmer, J.D., 2001. Multiple losses and transfers to the nucleus of two mitochondrial succinate dehydrogenase genes during angiosperm evolution. *Genetics* 158, 1289–1300.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002a. Catalysis and the Use of Energy by Cells. *Mol. Biol. Cell* 4th Ed.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002b. DNA Replication Mechanisms, in: *Molecular Biology of the Cell*. Garland Science.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Anantharaman, K., Duhaime, M.B., Breier, J.A., Wendt, K.A., Toner, B.M., Dick, G.J., 2014. Sulfur Oxidation Genes in Diverse Deep-Sea Viruses. *Science* 344, 757–760. <https://doi.org/10.1126/science.1252229>
- Arslan, D., Legendre, M., Seltzer, V., Abergel, C., Claverie, J.-M., 2011. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl. Acad. Sci. U. S. A.* 108, 17486–17491. <https://doi.org/10.1073/pnas.1110889108>
- Atlas, R.M., 2004. *Handbook of Microbiological Media*, 3rd ed. CRC Press LLC, Boca Raton.
- Baudoux, A.-C., Brussaard, C.P.D., 2005. Characterization of different viruses infecting the marine harmful algal bloom species *Phaeocystis globosa*. *Virology* 341, 80–90. <https://doi.org/10.1016/j.virol.2005.07.002>
- Bergh, Ø., Børsheim, K.Y., Bratbak, G., Heldal, M., 1989. High abundance of viruses found in aquatic environments. *Nature* 340, 467. <https://doi.org/10.1038/340467a0>

- Bessette, P.H., Åslund, F., Beckwith, J., Georgiou, G., 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc. Natl. Acad. Sci.* 96, 13703–13708. <https://doi.org/10.1073/pnas.96.24.13703>
- Bhatia, P., Chugh, A., 2015. Role of marine bioprospecting contracts in developing access and benefit sharing mechanism for marine traditional knowledge holders in the pharmaceutical industry. *Glob. Ecol. Conserv.* 3, 176–187. <https://doi.org/10.1016/j.gecco.2014.11.015>
- Boore, J.L., 1999. Animal mitochondrial genomes. *Nucleic Acids Res.* 27, 1767–1780. <https://doi.org/10.1093/nar/27.8.1767>
- Bornhorst, J.A., Falke, J.J., 2000. [16] Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzymol.* 326, 245–254.
- Botstein, D., 1980. A Theory of Modular Evolution for Bacteriophages\*. *Ann. N. Y. Acad. Sci.* 354, 484–491. <https://doi.org/10.1111/j.1749-6632.1980.tb27987.x>
- Breitbart, M., 2011. Marine Viruses: Truth or Dare. *Annu. Rev. Mar. Sci.* 4, 425–448. <https://doi.org/10.1146/annurev-marine-120709-142805>
- Breitbart, M., Thompson, L., Suttle, C., Sullivan, M., 2007. Exploring the Vast Diversity of Marine Viruses. *Oceanography* 20, 135–139. <https://doi.org/10.5670/oceanog.2007.58>
- Brown, C.M., Bidle, K.D., 2014. Attenuation of virus production at high multiplicities of infection in *Aureococcus anophagefferens*. *Virology*, Special issue: Giant Viruses 466–467, 71–81. <https://doi.org/10.1016/j.virol.2014.07.023>
- Buchan, J.R., Stansfield, I., 2007. Halting a cellular production line: responses to ribosomal pausing during translation. *Biol. Cell* 99, 475–487. <https://doi.org/10.1042/BC20070037>
- Burgess-Brown, N.A., Sharma, S., Sobott, F., Loenarz, C., Oppermann, U., Gileadi, O., 2008. Codon optimization can improve expression of human genes in *Escherichia coli*: A multi-gene study. *Protein Expr. Purif.* 59, 94–102. <https://doi.org/10.1016/j.pep.2008.01.008>
- Carter, M., Shieh, J., 2015. Chapter 10 - Molecular Cloning and Recombinant DNA Technology, in: Carter, M., Shieh, J. (Eds.), *Guide to Research Techniques in Neuroscience (Second Edition)*. Academic Press, San Diego, pp. 219–237. <https://doi.org/10.1016/B978-0-12-800511-8.00010-1>
- Cheng, Y.-S., Labavitch, J., VanderGheynst, J.S., 2015. Organic and Inorganic Nitrogen Impact *Chlorella variabilis* Productivity and Host Quality for Viral Production and

- Cell Lysis. *Appl. Biochem. Biotechnol.* 176, 467–479. <https://doi.org/10.1007/s12010-015-1588-0>
- Choi, K.H., 2012. Viral Polymerases. *Adv. Exp. Med. Biol.* 726, 267–304.  
[https://doi.org/10.1007/978-1-4614-0980-9\\_12](https://doi.org/10.1007/978-1-4614-0980-9_12)
- Claverie, J.-M., Abergel, C., 2018. Mimiviridae: An Expanding Family of Highly Diverse Large dsDNA Viruses Infecting a Wide Phylogenetic Range of Aquatic Eukaryotes. *Viruses* 10, 506. <https://doi.org/10.3390/v10090506>
- Claverie, J.-M., Abergel, C., 2010. Mimivirus: the emerging paradox of quasi-autonomous viruses. *Trends Genet. TIG* 26, 431–437. <https://doi.org/10.1016/j.tig.2010.07.003>
- Clouthier, S., VanWalleghem, E., Copeland, S., Klassen, C., Hobbs, G., Nielsen, O., Anderson, E., 2013. A new species of nucleo-cytoplasmic large DNA virus (NCLDV) associated with mortalities in Manitoba lake sturgeon *Acipenser fulvescens*. *Dis. Aquat. Organ.* 102, 195–209. <https://doi.org/10.3354/dao02548>
- Coen, D.M., 1996. Viral DNA Polymerases.
- Coy, S.R., Gann, E.R., Pound, H.L., Short, S.M., Wilhelm, S.W., 2018. Viruses of Eukaryotic Algae: Diversity, Methods for Detection, and Future Directions. *Viruses* 10, 487.  
<https://doi.org/10.3390/v10090487>
- Dahl, E., Bagøien, E., Edvardsen, B., Stenseth, N., 2005. The dynamics of *Chrysochromulina* species in the Skagerrak in relation to environmental conditions - ScienceDirect [WWW Document]. URL <https://www.sciencedirect.com/science/article/abs/pii/S1385110105000079> (accessed 3.31.19).
- de Marco, A., 2009. Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb. Cell Factories* 8, 26.  
<https://doi.org/10.1186/1475-2859-8-26>
- de Pascale, D., De Santi, C., Fu, J., Landfald, B., 2012. The microbial diversity of Polar environments is a fertile ground for bioprospecting. *Mar. Genomics* 8, 15–22.  
<https://doi.org/10.1016/j.margen.2012.04.004>
- Deeg, C.M., Chow, C.-E.T., Suttle, C.A., 2018. The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea. *eLife* 7.  
<https://doi.org/10.7554/eLife.33014>
- Derman, A.I., Prinz, W.A., Belin, D., Beckwith, J., 1993. Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*. *Science* 262, 1744–1747.  
<https://doi.org/10.1126/science.8259521>

- Doherty, A.J., Connolly, B.A., Worrall, A.F., 1993. Overproduction of the toxic protein, bovine pancreatic DNaseI, in *Escherichia coli* using a tightly controlled T7-promoter-based vector. *Gene* 136, 337–340. [https://doi.org/10.1016/0378-1119\(93\)90491-K](https://doi.org/10.1016/0378-1119(93)90491-K)
- Dong, H., Nilsson, L., Kurland, C.G., 1995. Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J. Bacteriol.* 177, 1497–1504. <https://doi.org/10.1128/jb.177.6.1497-1504.1995>
- Dumon-Seignovert, L., Cariot, G., Vuillard, L., 2004. The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr. Purif.* 37, 203–206. <https://doi.org/10.1016/j.pep.2004.04.025>
- Duong-Ly, K.C., Gabelli, S.B., 2014. Chapter Eighteen - Explanatory Chapter: Troubleshooting Protein Expression: What to do When the Protein is not Soluble, in: Lorsch, J. (Ed.), *Methods in Enzymology, Laboratory Methods in Enzymology: Protein Part C*. Academic Press, pp. 231–247. <https://doi.org/10.1016/B978-0-12-420119-4.00018-5>
- Durán, R.V., Oppliger, W., Robitaille, A.M., Heiserich, L., Skendaj, R., Gottlieb, E., Hall, M.N., 2012. Glutaminolysis activates Rag-mTORC1 signaling. *Mol. Cell* 47, 349–358. <https://doi.org/10.1016/j.molcel.2012.05.043>
- Edwardsen, B., Paasche, E., 1998. Bloom dynamics and physiology of *Prymnesium* and *Chrysochromulina*. *Physiol. Ecol. Harmful Algal Blooms* 193–208.
- Edwards, K.F., Steward, G.F., 2018. Host Traits Drive Viral Life Histories across Phytoplankton Viruses. *Am. Nat.* 191, 566–581. <https://doi.org/10.1086/696849>
- Eikrem, W., Medlin, L., Henderiks, J., Rokitta, S., Rost, B., Probert, I., Throndsen, J., Edwardsen, B., 2016. Haptophyta. [https://doi.org/10.1007/978-3-319-32669-6\\_38-1](https://doi.org/10.1007/978-3-319-32669-6_38-1)
- Enav, H., Mandel-Gutfreund, Y., Béjà, O., 2014. Comparative metagenomic analyses reveal viral-induced shifts of host metabolism towards nucleotide biosynthesis. *Microbiome* 2, 9. <https://doi.org/10.1186/2049-2618-2-9>
- Fernández-García, J.L., Ory, A. de, Brussaard, C.P.D., Vega, M. de, 2017. *Phaeocystis globosa* Virus DNA Polymerase X: a “Swiss Army knife”, Multifunctional DNA polymerase-lyase-ligase for Base Excision Repair. *Sci. Rep.* 7, 6907. <https://doi.org/10.1038/s41598-017-07378-3>
- Filée, J., Forterre, P., Laurent, J., 2003. The role played by viruses in the evolution of their hosts: a view based on informational protein phylogenies. *Res. Microbiol.* 154, 237–243. [https://doi.org/10.1016/S0923-2508\(03\)00066-4](https://doi.org/10.1016/S0923-2508(03)00066-4)

- Filée, J., Forterre, P., Sen-Lin, T., Laurent, J., 2002. Evolution of DNA Polymerase Families: Evidences for Multiple Gene Exchange Between Cellular and Viral Proteins. *J. Mol. Evol.* 54, 763–773. <https://doi.org/10.1007/s00239-001-0078-x>
- Fischer, M.G., Kelly, I., Foster, L.J., Suttle, C.A., 2014. The virion of *Cafeteria roenbergensis* virus (CroV) contains a complex suite of proteins for transcription and DNA repair. *Virology*, Special issue: Giant Viruses 466–467, 82–94. <https://doi.org/10.1016/j.virol.2014.05.029>
- Fuhrman, J.A., 1999. Marine viruses and their biogeochemical and ecological effects. *Nat. Lond.* 399, 541–8. <http://dx.doi.org.pva.uib.no/10.1038/21119>
- Gallot-Lavallée, L., Pagarete, A., Legendre, M., Santini, S., Sandaa, R.-A., Himmelbauer, H., Ogata, H., Bratbak, G., Claverie, J.-M., 2015. The 474-Kilobase-Pair Complete Genome Sequence of CeV-01B, a Virus Infecting *Haptolina* (*Chrysochromulina*) *ericina* (*Prymnesiophyceae*). *Genome Announc.* 3. <https://doi.org/10.1128/genomeA.01413-15>
- García-Beato, R., Salas, M.L., Viñuela, E., Salas, J., 1992. Role of the host cell nucleus in the replication of African swine fever virus DNA. *Virology* 188, 637–649.
- Gardner, A.F., Kelman, Z., 2014. DNA polymerases in biotechnology. *Front. Microbiol.* 5. <https://doi.org/10.3389/fmicb.2014.00659>
- Gjørseter, J., Lekve, K., Stenseth, N.C., Petter Leinaas, H., Christie, H., Dahl, E., S. Danielssen, D., Edvardsen, B., Olsgard, F., Oug, E., Paasche, E., 2000. A long-term perspective on the *Chrysochromulina* bloom on the Norwegian Skagerrak coast 1988: A catastrophe or an innocent incident? *Mar. Ecol. Prog. Ser.* 207, 201–218. <https://doi.org/10.3354/meps207201>
- Goodman, S., Kratz, L.E., DiGiulio, K.A., Biery, B.J., Goodman, K.E., Isaya, G., Frerman, F.E., 1995. Cloning of glutaryl-CoA dehydrogenase cDNA, and expression of wild type and mutant enzymes in *Escherichia coli*. *Hum. Mol. Genet.* 4, 1493–1498. <https://doi.org/10.1093/hmg/4.9.1493>
- Gottesman, S., 1996. Proteases and their targets in *Escherichia coli*. *Annu. Rev. Genet.* 30, 465–506. <https://doi.org/10.1146/annurev.genet.30.1.465>
- Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B., 2008. Protein production and purification. *Nat. Methods* 5, 135–146. <https://doi.org/10.1038/nmeth.f.202>
- Gray, M.W., Lang, B.F., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T.G., Plante, I., Rioux, P., Saint-Louis, D.,

- Zhu, Y., Burger, G., 1998. Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res.* 26, 865–878.
- Griffiths, A.J., Gelbart, W.M., Miller, J.H., Lewontin, R.C., 1999. Cloning a Specific Gene. *Mod. Genet. Anal.*
- Griffiths, A.J., Miller, J.H., Suzuki, D.T., Lewontin, R.C., Gelbart, W.M., 2000. Transcription and RNA polymerase. *Introd. Genet. Anal.* 7th Ed.
- Grodberg, J., Dunn, J.J., 1988. ompT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification. *J. Bacteriol.* 170, 1245–1253.
- Gurr, S.J., 1991. PCR protocols-a guide to methods and applications: Edited by M A Innis, D H Gelfand, J J Sninsky and T J White. pp 482. Academic Press, London 1990. \$39.95 ISBN 0-12-372181-4. *Biochem. Educ.* 19, 45–45. [https://doi.org/10.1016/0307-4412\(91\)90165-5](https://doi.org/10.1016/0307-4412(91)90165-5)
- Gustafsson, C., Govindarajan, S., Minshull, J., 2004. Codon bias and heterologous protein expression. *Trends Biotechnol.* 22, 346–353. <https://doi.org/10.1016/j.tibtech.2004.04.006>
- Hartley, D.L., Kane, J.F., 1988. Properties of inclusion bodies from recombinant Escherichia coli. *Biochem. Soc. Trans.* 16, 101–102. <https://doi.org/10.1042/bst0160101>
- Hendrix, R.W., 2003. Bacteriophage genomics. *Curr. Opin. Microbiol.* 6, 506–511. <https://doi.org/10.1016/j.mib.2003.09.004>
- Hingamp, P., Grimsley, N., Acinas, S.G., Clerissi, C., Subirana, L., Poulain, J., Ferrera, I., Sarmiento, H., Villar, E., Lima-Mendez, G., Faust, K., Sunagawa, S., Claverie, J.-M., Moreau, H., Desdevises, Y., Bork, P., Raes, J., de Vargas, C., Karsenti, E., Kandels-Lewis, S., Jaillon, O., Not, F., Pesant, S., Wincker, P., Ogata, H., 2013. Exploring nucleo-cytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. *ISME J.* 7, 1678–1695. <https://doi.org/10.1038/ismej.2013.59>
- Horas, E.L., Theodosiou, L., Becks, L., 2018. Why Are Algal Viruses Not Always Successful? *Viruses* 10, 474. <https://doi.org/10.3390/v10090474>
- Huang, S., Millar, A.H., 2013. Succinate dehydrogenase: the complex roles of a simple enzyme. *Curr. Opin. Plant Biol., Physiology and metabolism* 16, 344–349. <https://doi.org/10.1016/j.pbi.2013.02.007>
- Hurwitz, B.L., Hallam, S.J., Sullivan, M.B., 2013. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol.* 14, R123. <https://doi.org/10.1186/gb-2013-14-11-r123>

- Iyer, L.M., Aravind, L., Koonin, E.V., 2001. Common Origin of Four Diverse Families of Large Eukaryotic DNA Viruses. *J. Virol.* 75, 11720–11734.  
<https://doi.org/10.1128/JVI.75.23.11720-11734.2001>
- Iyer, L.M., Balaji, S., Koonin, E.V., Aravind, L., 2006. Evolutionary genomics of nucleocytoplasmic large DNA viruses. *Virus Res.* 117, 156–184.  
<https://doi.org/10.1016/j.virusres.2006.01.009>
- Jacobsen, A., Bratbak, G., Heldal, M., 1996. Isolation and Characterization of a Virus Infecting *Phaeocystis Pouchetii* (prymnesiophyceae)1. *J. Phycol.* 32, 923–927.  
<https://doi.org/10.1111/j.0022-3646.1996.00923.x>
- Jin, L., Li, D., Alesi, G.N., Fan, J., Kang, H.-B., Lu, Z., Boggon, T.J., Jin, P., Yi, H., Wright, E.R., Duong, D., Seyfried, N.T., Egnatchik, R., DeBerardinis, R.J., Magliocca, K.R., He, C., Arellano, M.L., Khoury, H.J., Shin, D.M., Khuri, F.R., Kang, S., 2015. Glutamate dehydrogenase 1 signals through antioxidant glutathione peroxidase 1 to regulate redox homeostasis and tumor growth. *Cancer Cell* 27, 257–270.  
<https://doi.org/10.1016/j.ccell.2014.12.006>
- Johannessen, T.V., Bratbak, G., Larsen, A., Ogata, H., Egge, E.S., Edvardsen, B., Eikrem, W., Sandaa, R.-A., 2015. Characterisation of three novel giant viruses reveals huge diversity among viruses infecting Prymnesiales (Haptophyta). *Virology* 476, 180–188.  
<https://doi.org/10.1016/j.virol.2014.12.014>
- Kane, J.F., 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* 6, 494–500.  
[https://doi.org/10.1016/0958-1669\(95\)80082-4](https://doi.org/10.1016/0958-1669(95)80082-4)
- Kinsella, J.E., Melachouris, N., 1976. Functional properties of proteins in foods: A survey. *C R C Crit. Rev. Food Sci. Nutr.* 7, 219–280.  
<https://doi.org/10.1080/10408397609527208>
- Kita, K., Hirawake, H., Miyadera, H., Amino, H., Takeo, S., 2002. Role of complex II in anaerobic respiration of the parasite mitochondria from *Ascaris suum* and *Plasmodium falciparum*. *Biochim. Biophys. Acta BBA - Bioenerg.* 1553, 123–139.  
[https://doi.org/10.1016/S0005-2728\(01\)00237-7](https://doi.org/10.1016/S0005-2728(01)00237-7)
- Koonin, E.V., Yutin, N., 2010. Origin and Evolution of Eukaryotic Large Nucleocytoplasmic DNA Viruses. *Intervirology* 53, 284–292.  
<https://doi.org/10.1159/000312913>

- Kristensen, D.M., Mushegian, A.R., Dolja, V.V., Koonin, E.V., 2010. New dimensions of the virus world discovered through metagenomics. *Trends Microbiol.* 18, 11–19.  
<https://doi.org/10.1016/j.tim.2009.11.003>
- Kuznar, J., Salas, M.L., Viñuela, E., 1980. DNA-dependent RNA polymerase in African swine fever virus. *Virology* 101, 169–175. [https://doi.org/10.1016/0042-6822\(80\)90493-6](https://doi.org/10.1016/0042-6822(80)90493-6)
- Lang, B.F., Gray, M.W., Burger, G., 1999. Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.* 33, 351–397.  
<https://doi.org/10.1146/annurev.genet.33.1.351>
- Larsen, A., Flaten, G.A.F., Sandaa, R.-A., Castberg, T., Thyraug, R., Erga, S.R., Jacquet, S., Bratbak, G., 2004. Spring phytoplankton bloom dynamics in Norwegian coastal waters: Microbial community succession and diversity. *Limnol. Oceanogr.* 49, 180–190. <https://doi.org/10.4319/lo.2004.49.1.0180>
- Larsen, J.B., Larsen, A., Bratbak, G., Sandaa, R.-A., 2008. Phylogenetic Analysis of Members of the Phycodnaviridae Virus Family, Using Amplified Fragments of the Major Capsid Protein Gene. *Appl Env. Microbiol* 74, 3048–3057.  
<https://doi.org/10.1128/AEM.02548-07>
- Legendre, M., Arslan, D., Abergel, C., Claverie, J.-M., 2012. Genomics of Megavirus and the elusive fourth domain of Life. *Commun. Integr. Biol.* 5, 102–106.
- Lenich, A.C., Goodman, S.I., 1986. The purification and characterization of glutaryl-coenzyme A dehydrogenase from porcine and human liver. *J. Biol. Chem.* 261, 4090–4096.
- Lindell, D., Jaffe, J.D., Johnson, Z.I., Church, G.M., Chisholm, S.W., 2005. Photosynthesis genes in marine viruses yield proteins during host infection. *Nat. Lond.* 438, 86–9.
- Lindell, D., Sullivan, M.B., Johnson, Z.I., Tolonen, A.C., Rohwer, F., Chisholm, S.W., 2004. Transfer of photosynthesis genes to and from Prochlorococcus viruses. *Proc. Natl. Acad. Sci. U. S. A.* 101, 11013–11018. <https://doi.org/10.1073/pnas.0401526101>
- Liu, H., Probert, I., Uitz, J., Claustre, H., Aris-Brosou, S., Frada, M., Not, F., Vargas, C. de, 2009. Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *Proc. Natl. Acad. Sci.* 106, 12803–12808.  
<https://doi.org/10.1073/pnas.0905841106>
- Liu, J., Zhang, Z., Liu, X., Cai, Y., Cai, H., 2013. Cloning, expression of the major capsid protein gene from marine algae *Emiliana huxleyi* virus and the possible use in



- detection of virus infection. *Microbiol. Res.* 4, e5–e5.  
<https://doi.org/10.4081/mr.2013.e5>
- Lozada, M., M. Dionisi, H., 2015. Microbial Bioprospecting in Marine Environments, in: Springer Handbook of Marine Biotechnology. pp. 307–326.  
[https://doi.org/10.1007/978-3-642-53971-8\\_11](https://doi.org/10.1007/978-3-642-53971-8_11)
- Marie, D., Brussaard, C.P.D., Thyraug, R., Bratbak, G., Vaulot, D., 1999. Enumeration of Marine Viruses in Culture and Natural Samples by Flow Cytometry. *Appl. Environ. Microbiol.* 65, 45–52.
- Markert, C.L., 1984. Lactate dehydrogenase. Biochemistry and function of lactate dehydrogenase. *Cell Biochem. Funct.* 2, 131–134.  
<https://doi.org/10.1002/cbf.290020302>
- McCormick, A.M., Jarmusik, N.A., Endrizzi, E.J., Leipzig, N.D., 2014. Expression, Isolation, and Purification of Soluble and Insoluble Biotinylated Proteins for Nerve Tissue Regeneration. *J. Vis. Exp. JoVE.* <https://doi.org/10.3791/51295>
- McCrossan, M., Windsor, M., Ponnambalam, S., Armstrong, J., Wileman, T., 2001. The trans Golgi Network Is Lost from Cells Infected with African Swine Fever Virus. *J. Virol.* 75, 11755–11765. <https://doi.org/10.1128/JVI.75.23.11755-11765.2001>
- McGrath, S., Fitzgerald, G.F., van Sinderen, D., 2004. The impact of bacteriophage genomics. *Curr. Opin. Biotechnol.* 15, 94–99. <https://doi.org/10.1016/j.copbio.2004.01.007>
- Menzella, H.G., 2011. Comparison of two codon optimization strategies to enhance recombinant protein production in *Escherichia coli*. *Microb. Cell Factories* 10, 15.  
<https://doi.org/10.1186/1475-2859-10-15>
- Mergulhão, F.J.M., Summers, D.K., Monteiro, G.A., 2005. Recombinant protein secretion in *Escherichia coli*. *Biotechnol. Adv.* 23, 177–202.  
<https://doi.org/10.1016/j.biotechadv.2004.11.003>
- Mifflin, B., Habash, D., 2002. The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in nitrogen utilization of crops. <https://doi.org/10.1093/jexbot/53.370.979>
- Mihara, T., Koyano, H., Hingamp, P., Grimsley, N., Goto, S., Ogata, H., 2018. Taxon Richness of “Megaviridae” Exceeds those of Bacteria and Archaea in the Ocean. *Microbes Environ.* 33, 162–171. <https://doi.org/10.1264/jsme2.ME17203>
- Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., Rüger, W., 2003. Bacteriophage T4 Genome. *Microbiol. Mol. Biol. Rev.* 67, 86–156.  
<https://doi.org/10.1128/MMBR.67.1.86-156.2003>

- Miroux, B., Walker, J.E., 1996. Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *J. Mol. Biol.* 260, 289–298. <https://doi.org/10.1006/jmbi.1996.0399>
- Mojica, K.D.A., Brussaard, C.P.D., 2014. Factors affecting virus dynamics and microbial host–virus interactions in marine environments. *FEMS Microbiol. Ecol.* 89, 495–515. <https://doi.org/10.1111/1574-6941.12343>
- Monier, A., Chambouvet, A., Milner, D.S., Attah, V., Terrado, R., Lovejoy, C., Moreau, H., Santoro, A.E., Derelle, É., Richards, T.A., 2017. Host-derived viral transporter protein for nitrogen uptake in infected marine phytoplankton. *Proc. Natl. Acad. Sci. U. S. A.* 114, E7489–E7498. <https://doi.org/10.1073/pnas.1708097114>
- Monier, A., Pagarete, A., de Vargas, C., Allen, M.J., Read, B., Claverie, J.-M., Ogata, H., 2009. Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus. *Genome Res.* 19, 1441–1449. <https://doi.org/10.1101/gr.091686.109>
- Moniruzzaman, M., LeCleir, G.R., Brown, C.M., Gobler, C.J., Bidle, K.D., Wilson, W.H., Wilhelm, S.W., 2014. Genome of brown tide virus (AaV), the little giant of the Megaviridae, elucidates NCLDV genome expansion and host–virus coevolution. *Virology*, Special issue: Giant Viruses 466–467, 60–70. <https://doi.org/10.1016/j.virol.2014.06.031>
- Moon, A.F., Garcia-Diaz, M., Batra, V.K., Beard, W.A., Bebenek, K., Kunkel, T.A., Wilson, S.H., Pedersen, L.C., 2007. The X Family Portrait: Structural Insights into Biological Functions of X Family Polymerases. *DNA Repair* 6, 1709–1725. <https://doi.org/10.1016/j.dnarep.2007.05.009>
- Novagen, 1999. pET System Manual.
- Pedersen, S., 1984. *Escherichia coli* ribosomes translate in vivo with variable rate. *EMBO J.* 3, 2895–2898.
- Picaud, S., Olsson, M.E., Brodelius, P.E., 2007. Improved conditions for production of recombinant plant sesquiterpene synthases in *Escherichia coli*. *Protein Expr. Purif.* 51, 71–79. <https://doi.org/10.1016/j.pep.2006.06.025>
- Plaitakis, A., Kalef-Ezra, E., Kotzamani, D., Zaganas, I., Spanaki, C., 2017. The Glutamate Dehydrogenase Pathway and Its Roles in Cell and Tissue Biology in Health and Disease. *Biology* 6. <https://doi.org/10.3390/biology6010011>
- Pope, B., Kent, H.M., 1996. High efficiency 5 min transformation of *Escherichia coli*. *Nucleic Acids Res.* 24, 536–537.

- Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., Finn, R.D., 2018. HMMER web server: 2018 update. *Nucleic Acids Res.* 46, W200–W204.  
<https://doi.org/10.1093/nar/gky448>
- Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., Scola, B.L., Suzan, M., Claverie, J.-M., 2004. The 1.2-Megabase Genome Sequence of Mimivirus. *Science* 306, 1344–1350. <https://doi.org/10.1126/science.1101485>
- Ren, J., 2010. Lactic Acid, in: Ren, J. (Ed.), *Biodegradable Poly(Lactic Acid): Synthesis, Modification, Processing and Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 4–14. [https://doi.org/10.1007/978-3-642-17596-1\\_2](https://doi.org/10.1007/978-3-642-17596-1_2)
- Robinson, P.K., 2015. *Enzymes: principles and biotechnological applications*. *Essays Biochem.* 59, 1–41. <https://doi.org/10.1042/bse0590001>
- Rojo, G., García-Beato, R., Viñuela, E., Salas, M.L., Salas, J., 1999. Replication of African Swine Fever Virus DNA in Infected Cells. *Virology* 257, 524–536.  
<https://doi.org/10.1006/viro.1999.9704>
- Rosano, G.L., Ceccarelli, E.A., 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* 5.  
<https://doi.org/10.3389/fmicb.2014.00172>
- Roucourt, B., Lavigne, R., 2009. The role of interactions between phage and bacterial proteins within the infected cell: a diverse and puzzling interactome. *Environ. Microbiol.* 11, 2789–2805. <https://doi.org/10.1111/j.1462-2920.2009.02029.x>
- Saïda, F., Uzan, M., Odaert, B., Bontems, F., 2016. Expression of Highly Toxic Genes in *E. coli*: Special Strategies and Genetic Tools.
- Salas, M.L., Rey-Campos, J., Almendral, J.M., Talavera, A., Viñuela, E., 1986. Transcription and translation maps of african swine fever virus. *Virology* 152, 228–240.  
[https://doi.org/10.1016/0042-6822\(86\)90387-9](https://doi.org/10.1016/0042-6822(86)90387-9)
- Sánchez, E.G., Quintas, A., Nogal, M., Castelló, A., Revilla, Y., 2013. African swine fever virus controls the host transcription and cellular machinery of protein synthesis. *Virus Res.*, African swine fever virus 173, 58–75.  
<https://doi.org/10.1016/j.virusres.2012.10.025>
- Sandaa, R.-A., 2008. Burden or benefit? Virus - host interactions in the marine environment.
- Sandaa, R.-A., Heldal, M., Castberg, T., Thyraug, R., Bratbak, G., 2001. Isolation and Characterization of Two Viruses with Large Genome Size Infecting *Chrysochromulina ericina* (Prymnesiophyceae) and *Pyramimonas orientalis* (Prasinophyceae). *Virology* 290, 272–280. <https://doi.org/10.1006/viro.2001.1161>

- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 74, 5463–5467.
- Santini, S., Jeudy, S., Bartoli, J., Poirot, O., Lescot, M., Abergel, C., Barbe, V., Wommack, K.E., Noordeloos, A.A.M., Brussaard, C.P.D., Claverie, J.-M., 2013. Genome of *Phaeocystis globosa* virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 110, 10800–10805. <https://doi.org/10.1073/pnas.1303251110>
- Saraste, M., 1999. Oxidative Phosphorylation at the fin de siècle. *Science* 283, 1488–1493.
- Schmid, M., Speiseder, T., Dobner, T., Gonzalez, R.A., 2014. DNA Virus Replication Compartments. *J. Virol.* 88, 1404–1420. <https://doi.org/10.1128/JVI.02046-13>
- Schoenfeld, T., Liles, M., Wommack, K.E., Polson, S.W., Godiska, R., Mead, D., 2010. Functional viral metagenomics and the next generation of molecular tools. *Trends Microbiol.* 18, 20. <https://doi.org/10.1016/j.tim.2009.10.001>
- Schvarcz, C.R., Steward, G.F., 2018. A giant virus infecting green algae encodes key fermentation genes. *Virology* 518, 423–433. <https://doi.org/10.1016/j.virol.2018.03.010>
- Serrano, H., Blanchard, J.S., 2013. Kinetic and Isotopic Characterization of L-Proline Dehydrogenase from *Mycobacterium tuberculosis*. *Biochemistry* 52. <https://doi.org/10.1021/bi400338f>
- Servet, C., Ghelis, T., Richard, L., Zilberstein, A., Savoure, A., 2012. Proline dehydrogenase: a key enzyme in controlling cellular homeostasis.
- Sezonov, G., Joseleau-Petit, D., D’Ari, R., 2007. *Escherichia coli* Physiology in Luria-Bertani Broth. *J. Bacteriol.* 189, 8746–8749. <https://doi.org/10.1128/JB.01368-07>
- Shahbaz- Mohammadi, H., Omidinia, E., 2011. Screening and characterization of proline dehydrogenase flavoenzyme producing *Pseudomonas entomophila*. *Iran. J. Microbiol.* 3, 201–209.
- Sharon, I., Alperovitch, A., Rohwer, F., Haynes, M., Glaser, F., Atamna-Ismaeel, N., Pinter, R.Y., Partensky, F., Koonin, E.V., Wolf, Y.I., Nelson, N., Béjà, O., 2009. All-in-one: Photosystem-I gene cassettes in marine viruses. *Nature* 461, 258–262. <https://doi.org/10.1038/nature08284>
- Shehata, H.R., Li, J., Chen, S., Redda, H., Cheng, S., Tabujara, N., Li, H., Warriner, K., Hanner, R., 2017. Droplet digital polymerase chain reaction (ddPCR) assays integrated with an internal control for quantification of bovine, porcine, chicken and

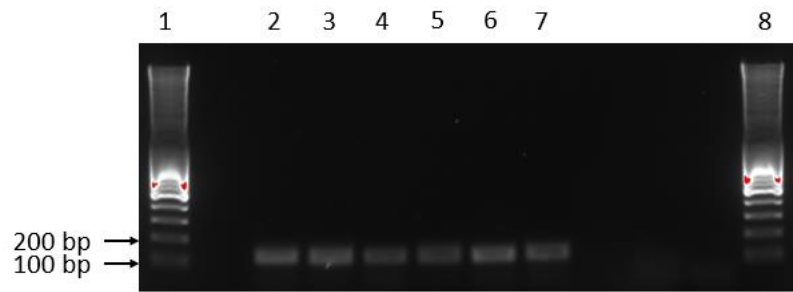
- turkey species in food and feed. PLoS ONE 12.  
<https://doi.org/10.1371/journal.pone.0182872>
- Shen, Y.-Q., Lang, B.F., Burger, G., 2009. Diversity and dispersal of a ubiquitous protein family: acyl-CoA dehydrogenases. *Nucleic Acids Res.* 37, 5619–5631.  
<https://doi.org/10.1093/nar/gkp566>
- Short, S.M., 2012. The ecology of viruses that infect eukaryotic algae. *Environ. Microbiol.* 14, 2253–2271. <https://doi.org/10.1111/j.1462-2920.2012.02706.x>
- Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., Li, Q., 2009. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci. Publ. Protein Soc.* 18, 936–948. <https://doi.org/10.1002/pro.102>
- Sørensen, H.P., Mortensen, K.K., 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J. Biotechnol.* 115, 113–128.  
<https://doi.org/10.1016/j.jbiotec.2004.08.004>
- Spencer, R., 1955. A Marine Bacteriophage | Nature [WWW Document]. URL <https://www.nature.com/articles/175690a0> (accessed 4.29.19).
- Studier, F.W., 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234. <https://doi.org/10.1016/j.pep.2005.01.016>
- Su, X., Wu, Y., Sifri, C.D., Welles, T.E., 1996. Reduced Extension Temperatures Required for PCR Amplification of Extremely A+T-rich DNA. *Nucleic Acids Res.* 24, 1574–1575. <https://doi.org/10.1093/nar/24.8.1574>
- Sullivan, M.B., Lindell, D., Lee, J.A., Thompson, L.R., Bielawski, J.P., Chisholm, S.W., 2006. Prevalence and Evolution of Core Photosystem II Genes in Marine Cyanobacterial Viruses and Their Hosts. *PLOS Biol.* 4, e234.  
<https://doi.org/10.1371/journal.pbio.0040234>
- Sun, F., Huo, X., Zhai, Y., Wang, A., Xu, J., Su, D., Bartlam, M., Rao, Z., 2005. Crystal Structure of Mitochondrial Respiratory Membrane Protein Complex II. *Cell* 121, 1043–1057. <https://doi.org/10.1016/j.cell.2005.05.025>
- Suttle, C., 2005. Viruses in the sea. <https://doi.org/10.1038/nature04160>
- Suttle, C., Chan, A., 1995. Viruses infecting the marine Prymnesiophyte *Chrysochromulina* spp.: isolation, preliminary characterization and natural abundance. *Mar. Ecol. Prog. Ser.* 118, 275–282. <https://doi.org/10.3354/meps118275>
- Suttle, C.A., 2007. Marine viruses — major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. <https://doi.org/10.1038/nrmicro1750>

- Terpe, K., 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 72, 211. <https://doi.org/10.1007/s00253-006-0465-8>
- Thompson, L.R., Zeng, Q., Kelly, L., Huang, K.H., Singer, A.U., Stubbe, J., Chisholm, S.W., 2011. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc. Natl. Acad. Sci.* 108, E757–E764. <https://doi.org/10.1073/pnas.1102164108>
- Thomsen, H.A., Buck, K.R., Chavez, F.P., 1994. Haptophytes as components of marine phytoplankton. *Haptophyte Algae The Systematics Association Special Volume Bo.* 51, 187–208.
- Triyatni, M., Berger, E.A., Saunier, B., 2011. A New Model to Produce Infectious Hepatitis C Virus without the Replication Requirement. *PLoS Pathog.* 7. <https://doi.org/10.1371/journal.ppat.1001333>
- Vagenende, V., Yap, M.G.S., Trout, B.L., 2009. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry* 48, 11084–11096. <https://doi.org/10.1021/bi900649t>
- Van Etten, J.L., 2003. Unusual life style of giant chlorella viruses. *Annu. Rev. Genet.* 37, 153–195. <https://doi.org/10.1146/annurev.genet.37.110801.143915>
- Vera, A., González-Montalbán, N., Arís, A., Villaverde, A., 2007. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnol. Bioeng.* 96, 1101–1106. <https://doi.org/10.1002/bit.21218>
- Vogelstein, B., Kinzler, K.W., 1999. Digital PCR. *Proc. Natl. Acad. Sci. U. S. A.* 96, 9236–9241.
- Weigele, P.R., Pope, W.H., Pedulla, M.L., Houtz, J.M., Smith, A.L., Conway, J.F., King, J., Hatfull, G.F., Lawrence, J.G., Hendrix, R.W., 2007. Genomic and structural analysis of Syn9, a cyanophage infecting marine *Prochlorococcus* and *Synechococcus*. *Environ. Microbiol.* 9, 1675–1695. <https://doi.org/10.1111/j.1462-2920.2007.01285.x>
- Weynberg, K.D., Allen, M.J., Gilg, I.C., Scanlan, D.J., Wilson, W.H., 2011. Genome Sequence of *Ostreococcus tauri* Virus OtV-2 Throws Light on the Role of Picoeukaryote Niche Separation in the Ocean ▽. *J. Virol.* 85, 4520–4529. <https://doi.org/10.1128/JVI.02131-10>
- Wilson, W.H., Van Etten, J.L., Allen, M.J., 2009. The Phycodnaviridae: The Story of How Tiny Giants Rule the World. *Curr. Top. Microbiol. Immunol.* 328, 1–42.

- Yoosuf, N., Yutin, N., Colson, P., Shabalina, S.A., Pagnier, I., Robert, C., Azza, S., Klose, T., Wong, J., Rossmann, M.G., La Scola, B., Raoult, D., Koonin, E.V., 2012. Related Giant Viruses in Distant Locations and Different Habitats: *Acanthamoeba polyphaga* mousmouvirus Represents a Third Lineage of the Mimiviridae That Is Close to the Megavirus Lineage. *Genome Biol. Evol.* 4, 1324–1330.  
<https://doi.org/10.1093/gbe/evs109>
- Yutin, N., Koonin, E.V., 2012. Hidden evolutionary complexity of Nucleo-Cytoplasmic Large DNA viruses of eukaryotes. *Viol. J.* 9, 161. <https://doi.org/10.1186/1743-422X-9-161>
- Zayas, J.F., 1997. Solubility of Proteins, in: Zayas, J.F. (Ed.), *Functionality of Proteins in Food*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 6–75.  
[https://doi.org/10.1007/978-3-642-59116-7\\_2](https://doi.org/10.1007/978-3-642-59116-7_2)

## 7 Vedlegg A: Resultat

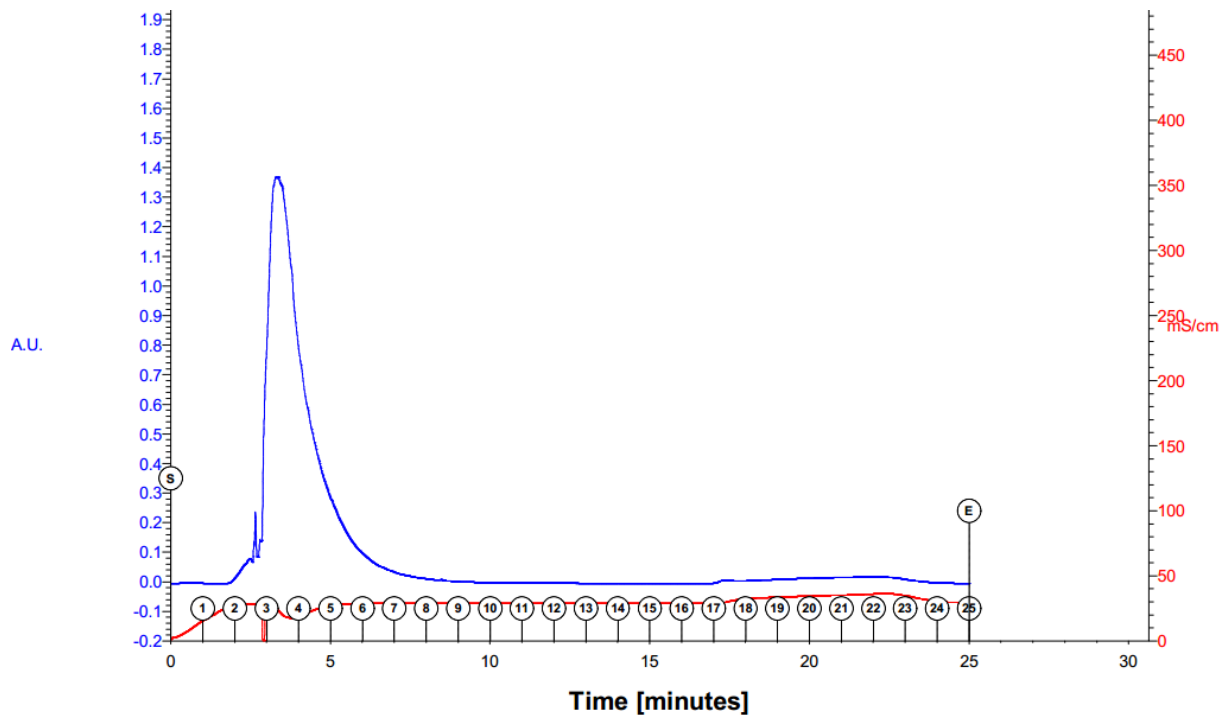
### A.1 Transkripsjonsforsøk



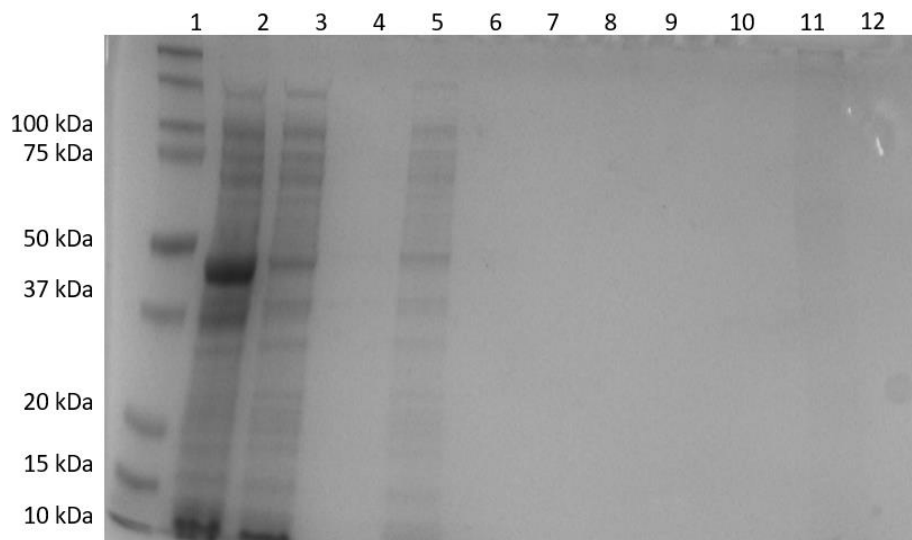
*Figur 7.1: mRNA kontroll*



## A.2 Rensing av glutaryl-CoA dehydrogenase

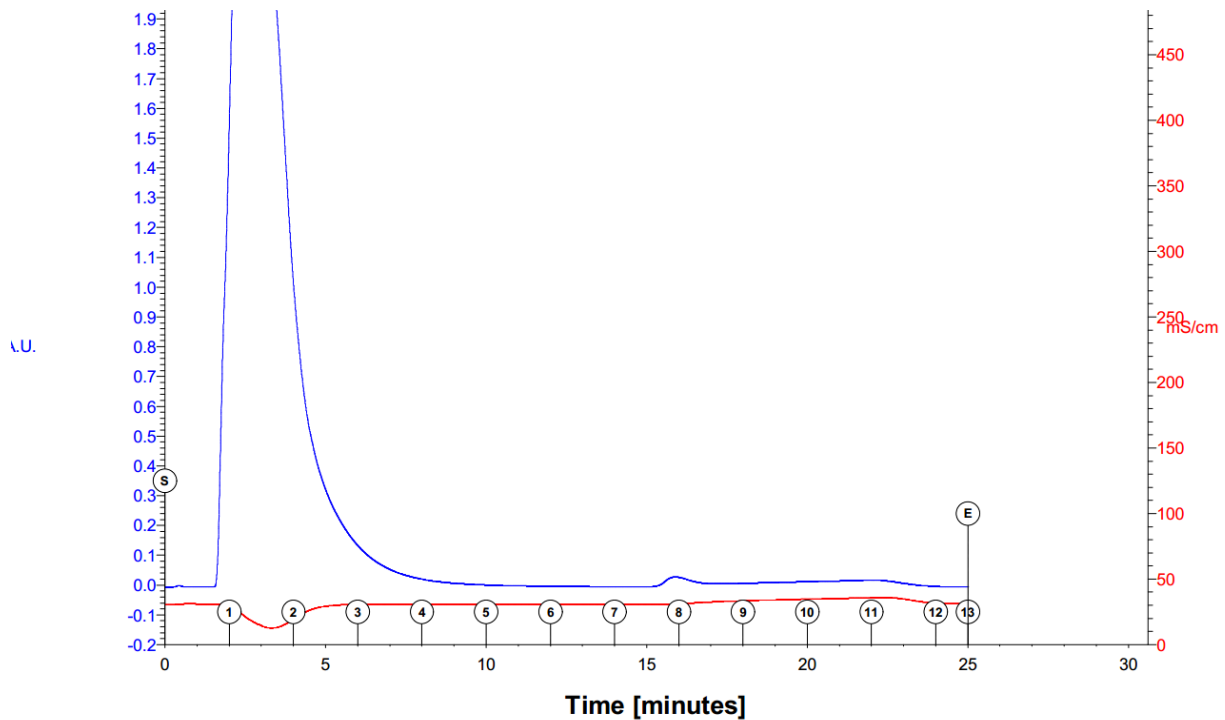


Figur 7.2: Diagrammet viser elueringsprofilen for rensing av Gl CoA på IMAC kolonnen. X-aksen viser tiden i minutter. Tallene med runding rundt (langs x-aksen) viser fraksjoner, og hver fraksjon var på 1,5 ml. Venstre y-akse (blå) viser UV absorbans (A.U.), og høyre y-akse (rød) viser konduktivitet (mS/cm).

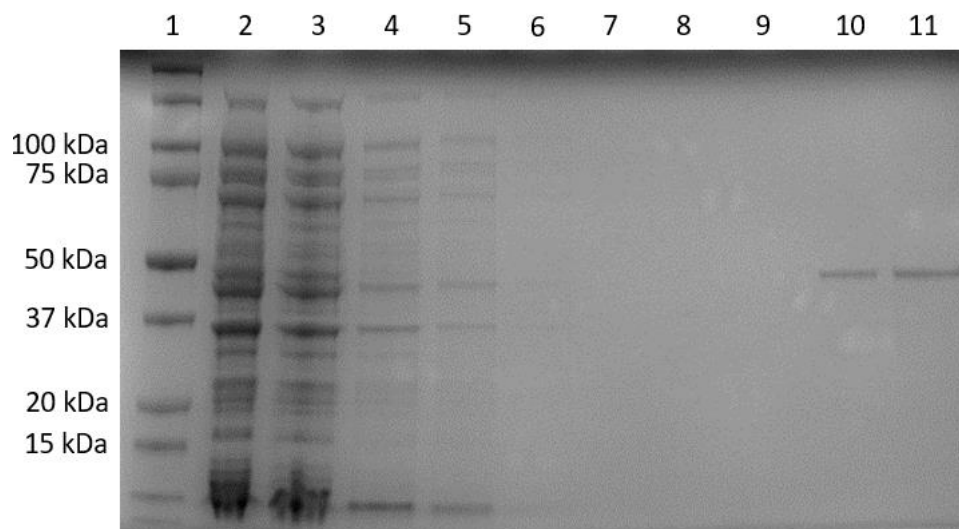


Figur 7.3: SDS-PAGE-analyse av prøver ved rensingen av rekombinant Gl CoA. Brønn 1: SDS-PAGE Precision plus protein standards, 10-250 kDa (BIO RAD). Brønn 2: lysat gcdh (total). Brønn 3: lysat gcdh filtrert (løselig). Brønn 4: Fraksjon 4. Brønn 5: Fraksjon 5. Brønn 6: Fraksjon 8. Brønn 7: Fraksjon 18. Brønn 8: Fraksjon 19. Brønn 9: Fraksjon 20. Brønn 10: Fraksjon 21. Brønn 11: Fraksjon 22. Brønn 12: Fraksjon 23.

### A.3 Rensing av DNA polymerase X – småskala



Figur 7.4: Diagrammet viser elueringsprofilen for rensing av DNA polymerase X på IMAC kolonnen. X-aksen viser tiden i minutter. Tallene med runding rundt (langs x-aksen) viser fraksjoner, og hver fraksjon var på 1,5 ml. Venstre y-akse (blå) viser UV absorbans (A.U.), og høyre y-akse (rød) viser konduktivitet (mS/cm).



Figur 7.5: SDS-PAGE-analyse av prøver ved rensingen av rekombinant DNA polymerase X. Brønn 1: SDS-PAGE Precision plus protein standards, 10-250 kDa (BIO RAD). Brønn 2: lysat pol x ufiltrert (uløselig). Brønn 3: lysat pol x filtrert (løselig). Brønn 4: Fraksjon 2. Brønn 5: Fraksjon 3. Brønn 6: Fraksjon 4. Brønn 7: Fraksjon 5. Brønn 8: Fraksjon 8-15  $\mu$ l tilsatt brønn. Brønn 9: Fraksjon 8-30  $\mu$ l tilsatt brønn. Brønn 10: Fraksjon 9-15  $\mu$ l tilsatt brønn. Brønn 11: Fraksjon 9-30  $\mu$ l tilsatt brønn.

## 8 Vedlegg B: Løsninger

### B.1 IMR/2 medium i 1L sjøvann

IMR/2-medium	Innhold	Stamløsning	Tilsetning til 1L 70 % SW
KH <sub>2</sub> PO <sub>4</sub>		0,68 g / 100 ml MQ	0,5 ml
KNO <sub>3</sub>		5 g / 100 ml MQ	0,5 ml
Spormetall løsning	Na <sub>2</sub> EDTA (3 g) NaFeEDTA (800 mg) MnSO <sub>4</sub> 7H <sub>2</sub> O (410 mg) ZnSO <sub>4</sub> 7H <sub>2</sub> O (125 mg) Na <sub>2</sub> MoO <sub>4</sub> 2H <sub>2</sub> O (65 mg) MilliQ-vann (500 ml) CuSO <sub>4</sub> /CoCl <sub>2</sub> -løsning (0,5 ml) <ul style="list-style-type: none"><li>- CuSO<sub>4</sub> 5H<sub>2</sub>O (0,4 g)</li><li>- CoCl<sub>2</sub> 6H<sub>2</sub>O (0,4 g)</li><li>- MilliQ-vann (100 ml)</li></ul>		0,5 ml
Vitamin løsning	Thiamin (50 mg) Biotin (0,5 mg) B12-løsning (0,5 ml) <ul style="list-style-type: none"><li>- 100 mg B12/100ml MilliQ</li></ul> MilliQ-vann (500 ml)		0,5 ml
Selen			0,25 ml

## B.2 SM buffer

0,1 M NaCl

8 mM MgSO<sub>4</sub>·7H<sub>2</sub>O

50 mM Tris-HCl (pH 8)

0,005 % Glyserin

Fordeles på 100 ml flasker og autoklaveres.

## B.3 0,5 M EDTA

186,1 g disodium EDTA (Na<sub>2</sub>EDTA)

800 ml dH<sub>2</sub>O

Juster pH til 8,0 med NaOH (~50 ml NaOH)

Tilsett sdH<sub>2</sub>O til 1 l og autoklaver

## B.4 LA medium med ampicillin (100 mg/ml) og LB medium

**LA medium: Tilsetning til 500 ml**    **LB medium: Tilsetning til 1 l**  
**MilliQ vann (til skåler)**                      **MilliQ vann**

<b>Trypton</b>	5 g	10 g
<b>NaCl</b>	5 g	10 g
<b>Gjærekstrakt</b>	2,5 g	5 g
<b>Agar</b>	10 g	N/A
<b>Ampicillin</b>	500 µl	N/A

Autoklaveres i 20 minutter ved 121 °C. LA medium overføres til skåler, som oppbevares ved 4 °C. LB medium oppbevares enten ved romtemperatur eller ved 4 °C.

## 9 Vedlegg C: protokoller og metoder

### C.1 Instrumentinnstillinger og formel – Flowcytometer

Tabell 9.1: Innstillinger for telling av virus og bakterier

<b>Virus/bakterier</b>	<b>FSC</b>	<b>SSC</b>	<b>FL1</b>	<b>FL2</b>	<b>FL3</b>	<b>Terskelverdi</b>	<b>Hovedplot</b>
<b>Spenning</b>	E02	620	530	500	700	20, FL1-H	FL1 vs SSC FL1 vs FL3 SSC histogram
<b>60 s @ 50 µl/min</b>							

Tabell 9.2: Innstillinger for telling av algen

<b>Alger</b>	<b>FSC</b>	<b>SSC</b>	<b>FL1</b>	<b>FL2</b>	<b>FL3</b>	<b>Terskelverdi</b>	<b>Hovedplot</b>
<b>Spenning</b>	E01	300	400	500	300	52, FL3-H	FL1 vs SSC FL2 vs FL1 FL1 histogram
<b>60 s @ 100 µl/min</b>							

Resultatet fra tellingen på FCM ble manuelt sortert i et regneark i Excel, deretter ble antall telte partikler omgjort til konsentrasjonen av partikler per ml:

$$\frac{\text{antall telte partikler} \times \text{fortynning} \times 1000}{\text{strømhastighet} \times \text{tid}}$$

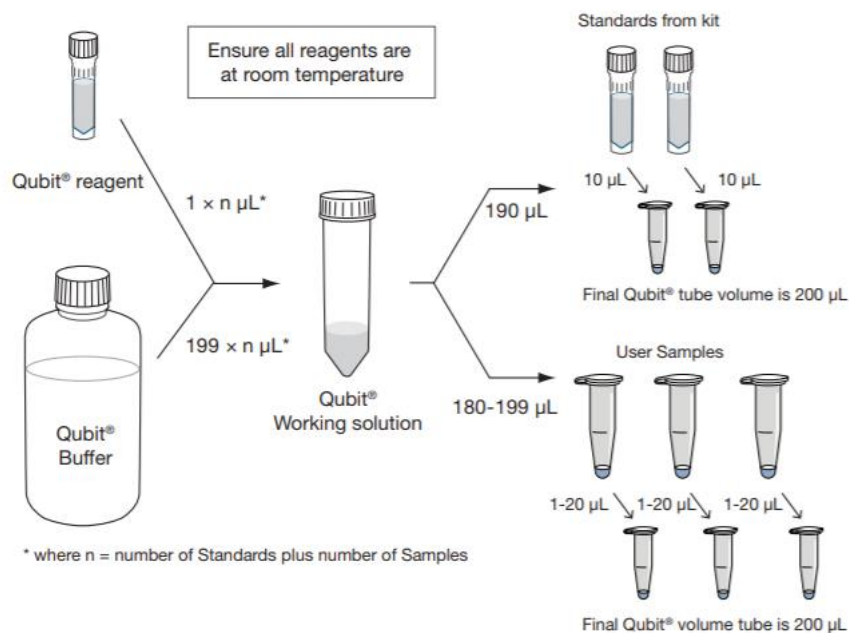
Disse innstillingene og formelen er en del av protokollen av Elzbieta Petelencz-Kurdziel for FCM ved Universitet i Bergen 2016.

## C.2 ZYMO DNA clean and concentrator™-10 kit

Protokoll med noen endringer gjort ved avdeling for mikrobiologi (UiB).

1. Tilsett 2-5 volum av ChIP DNA bindende buffer til hvert volum av DNA prøve (200  $\mu$ l DNA bindende buffer : 100  $\mu$ l DNA prøve).
2. Fordel løsningen på to Zymo-Spin™ IC-XL kolonne med ett oppsamlingsrør.
3. Sentrifuger ved 13000 g i 1 minutt, og kast avfallet.
4. Tilsett 200  $\mu$ l DNA vaske buffer til kolonnen. Sentrifuger ved 13000 g i 1 minutt.
5. Gjenta steg 4.
6. Sentrifuger prøven ved 13 000 g i 1 minutt.
7. Overfør kolonnen til et 1,5 ml Eppendorf rør.
8. Eluer med 10  $\mu$ l forvarmet Tris buffer (10mM), og inkuber i 1 minutt ved romtemperatur. Sentrifuger prøven ved 13 000 g i 30 sekund.
9. Gjenta steg 8.

## C.3 Qubit 2.0 Fluorometer



Figur 9.1: Viser hvordan man lager til standard og prøver for avlesning på Qubit 2.0 fluorometer.

Alle prøvene ble blandet godt og inkubert i 2 minutter ved romtemperatur. Instrumentet ble kalibrert ved å lese av standardene, og deretter ble prøvene lest av.

## C.4 Verifisering av transkripsjon og nedbryting av DNA

Tabell 9.3: PCR-reaksjonsmiks for amplifisering av SDHA

Amplifisert gen	Innhold i reaksjonsmiks	Mengde ( $\mu$ l) per reaksjon
<i>mcp</i>	HStaq master mix	10
	sdH <sub>2</sub> O	6
	10 $\mu$ M forward primer	1
	10 $\mu$ M reverse primer	1
	Templat	2
	Totalt	20

Tabell 9.4: PCR-program med gradient for amplifisering av SDHA

Steg i PCR	Temperatur ( $^{\circ}$ C)	Tid	Antall sykluser
<b>Denaturering</b>	94	5 minutter	1
	94	30 sekunder	
<b>Hybridisering</b>	59	30 sekunder	35
<b>Polymerisering</b>	72	30 sekunder	
	72	12 minutter	1
<b>Hold</b>	4	$\infty$	

## C.5 Blood and Cell culture DNA mini Kit (QIAGEN)

1. Ekvilibrer en QIAGEN genomic tip med 1 ml buffer QBT, og la bufferen renne gjennom
2. Bland prøven ved vortexing for 10 sekunder, og tilsett prøven til kolonnen og la det renne gjennom.
3. Vask kolonnen med 3 x 1 ml buffer QC.
4. Plassert kolonnen i et nytt rør og eluer med 2 x 1 ml buffer QC.
5. Fell ut DNAet ved å tilsette 1,4 ml romtemperert isopropanol til det eluerte DNAet.
6. Bland og sentrifuger ved  $5000 \times g$  i 15 minutter og fjern supernatanten.
7. Vask DNAet med 1 ml kald 70 % etanol, og sentrifuger ved  $5000 \times g$  i 10 minutter. Fjern supernatanten og resuspender med 2 ml TE buffer.

## C.6 Monarch Plasmid miniprep kit

1. Falconrøret med kultur ble sentrifugert i 3 minutter ved  $5000 \times g$  og supernatanten ble fjernet.
2. 200  $\mu\text{l}$  plasmid resuspension buffer (B1) ble tilsatt, og blandet godt slik at cellene er fullstendig resuspendert.
3. 200  $\mu\text{l}$  plasmid lysis buffer (B2) ble tilsatt, røret ble vendt 5-6 ganger før det ble inkubert ved romtemperatur i 1 minutt.
4. 400  $\mu\text{l}$  neutralization buffer (B3) ble tilsatt og røret ble vendt forsiktig frem til all væske ble gul (nøytralisert). Inkubering i 2 minutter.
5. Lysatet ble sentrifugert i 3 minutter ved  $5000 \times g$ .
6. Supernatanten ble overført til en kolonne og sentrifugert i 1 min ved  $5000 \times g$ . Kast avfallet.
7. 200  $\mu\text{l}$  plasmid wash buffer 1 ble tilsatt og røret ble sentrifugert i 1 minutt ved  $5000 \times g$ .
8. 400  $\mu\text{l}$  av wash buffer 2 ble tilsatt og røret ble sentrifugert i 1 minutt ved  $5000 \times g$ .
9. Kolonnen ble overført til et nytt 1 ml eppendorf rør, og tilsatt 30  $\mu\text{l}$  elueringsbuffer. Deretter ble røret inkubert i 1 minutt og sentrifugert i 1 minutt ved  $5000 \times g$ .



## 10 Vedlegg D: Sekvenser

### D.1 L-laktat dehydrogenase (L-LDH)

atgtatcaaaaaatgattagacctgtatttaattattacccaatcaaagattagctaagcttataatgtaatggatataagaaaatgtgctca  
acaaagacacataaaatggtggttgattatattgattctggagcagatgatgaattagctttacaatgggcaaatgacgcttataataaaa  
aatcttttagttataaagttttatctggagttctccaccaattgatttaataattctgtcggtaaacattaaatttgacttcctttttaattgtcc  
aactgctggaatagaatgtttcatagcgaaggagaattagcagtagctaaaacagctgaaaaatftaagactatgtatagtatatcttcatt  
atgtacaaccaaattagataaaaataaaacaagttcataatggtcctcaagattttcaactttacgtttgaaagatagacatttataaaagat  
gttatacaattagctaaagaaaacaattataatcttagctttaactgcagattttacctggatggtaatcgagaaagagataaaaaaat  
ggattcacaattcctccgcatattcaatgaatcaagtaattgaagcttataaaaagccagcttggacattagattttgatgaatcctaatt  
ataactatgctgtcttaaacaggatgctccagctgaaagtttagcacacttcgtaaatgtcaaatctgtcctgaatttaattgaaagatgc  
tgaatgggtactaggtgaatggaatcatggctcactgctttgaaaggtgctgttagaactgatgatgctaaaatggcaattgaaactggtt  
ttgatattttggatttcaacctgggtggcgtcaactggaaactcaatagctccattcgaggcattacagaaaataaacctcatattgg  
aaatacacaattatattagatggaggagttagaaggggaatagatatagtaaaagcaataagtttaggggcagatgctgctggaattgg  
aaaaccttacttatatggtttggctgctggaggacaaaatggtgttgaagagtatatgaaatftaatagatgagattcaagagctatgg  
gtttattgggagtttcatccatagctgaattgaaaagaatggctcatttttggtaaattag

### D.2 Prolin dehydrogenase (PDH)

atggtgagatacattggtgatcaaatataaatcaagttcttcaagaagtagaaatatattgaataatcgaagattccagttattaactatat  
aaaagaaaacagtgatattttaaataaaacagattataatagtgtgaacaagaatataatcaatcattccttactggataatcaatttaga  
gttgcctttaaattatcatctttaaatttaattatatttctgttcagacttaattcacaatttgacatcaaaaggaattcaattatgattgatca  
gaaaatagtcaaaataatcaaaagatcaagaaactaattggttaataaatgatttaataaagataaagtttaataatatttaaacttatca  
aatgtatcgtaaagatagctgttaatttagaaaatgatattaaatcaatcaagaacttggtaattatttaggtgttaattggttaagaggtg  
cttattggaatcaagaaaaagataatctagaattttacaataaatatgatactgatatgaattacaatcgagctatttattctttatgaaa  
aagaaataaaatcgatttagccactcataatcatgaatctattaatttaggattaattttaaataagataaaataaatgataattctaaacaattg  
aatttgctcacctattaggaatgaatgtaaataaatatgaatcaattaaagataaaataaagttaatgtctaccttccttatggtccatacaag  
catatgatgccttatttgatacgaagattatataaaaacatagatacagtaaaatataatggtttattaa

### D.3 Glutaryl-CoA dehydrogenase (GCDH)

atgattagagcaatttttaataaaaataaaatcaatccattttataaattatctacaacctagtaaatatattagaaaatttagtaattgga  
aagatccataaatctagagattgatcttaccgaagaagaaaaattagtaaaaaactactaaatcttttgtgaaaaagaattaatgcca  
gaattatatcagacacaagaaatgaatattttaaataaagaataatcaaaaaataggtgatttaggattttgaaactttaaagcagg

aaaattatgtttcatagcgattaatatgcagagagatagaacgaattgatagttcattgatctgcatttagcgtacaatcatcattggtaatt  
tatccaattaatctctatgggtccaataaccaacgagaaaaagtatcttaatgacttaataatggaactaaaattgggtgctttggttaacag  
aaccaaatcatgggtcggatccctcttcaatggagactaaagctaaattaaagatgattattatatttaaacggtagcaaatcttgattac  
taattctccgatagctgatatttttagttgggctaaagatgtgattcaataagattagaggattatattagacagagatttgaaggttta  
agcacacaaaaataaattgtaagtagctttaaagcttcagagacaggatatataattttagaagatgtaaagttccattaaagaattg  
tctcccaaatattgaaggattaaaaggtccttttagttggttaacaaagcaagatatggaatatcttggggagcattaggagctgcagaat  
tttgtttgaattagtaagagattatactttaagtagacaacaattggaaaaccattagctcaaatcaattaattcaaaaaaattagcagat  
atgcaaatc gatatagctttaggattaaattcatgtttaaagtaggtagttaataaaaaagggttagatgataattcaatgatatactatact  
aaaaaaaactcttgtaaaaatctataaatattgtcgtactctagagatattttaggtgctaattggtatctcagatgaatatcatgttattag  
acatttaataaattagaagcagttaacacatatgagggacaagtgatattcattcattaattattggcaagcaattactaatattagtgcatt  
tttaa

#### D.4 Glutamat dehydrogenase (GDH)

atgtcaaaaataactgatttaattaaatcaaatgataaagcatcacaattatccaaaattaatccatcggttatgaaatcttaattaagcc  
aaaacataaattagaataaatttccagtcataaataatgataaattgaagtatttactggttcagagttcaacataatgatatacttgg  
ccatttaaaggaggattaagattcatccctcagtatctatcagcaggtgtaattgtttatctcaatggatgactfacaatgtgctctcaag  
atttacccttggaggtggtaaaggaggattatctattgataatcaaaatatacacctgaacaaatagaagaaattagtagaaaatttggca  
aatcactatataattatattggaaatacaaaagatattccagctcctgatctaggaactaattctcaaatcatggattggatgatggacgaat  
ataatcgaatgagtggaagaagctataacatcaaatatgaaaagtatttactggaaaatcaactatttggaggtatagaggacg  
tgaagaagctactggaagaggtatagctataattattaaagagtgggtataaataataatattgattaaaagggtcaaaattatattgtcca  
aggttttgaaacgttgatctttactgctgaaatattatcttctatggaatgaattgataggggtggtgaccatactggttactattttt  
aatgaaggttttaattgttataagttaaaaaacatgttagtcaaaaaggagtttaaaagagtatttactggtgagaatatttctaaagaag  
aattctttaaaattgaatgtgatattataattccagcagcattagaattacaaatcaatgaattagaagctaaaaatataaaaactaaattaata  
gttgaagctgctaacggaccaatattttagaagctgaaaaattttattagaaaatggaattgaaattatactgatataattagctaattcgg  
gaggagtagtagtaagttattatgaatggttacaaaataaaagagatgagattatggaatataaagatgttattataatctagataaaaaa  
tgaaaaaagcattttataaagtgtcaaatattatcaataaatacaaatgtacacttagagaagccagttatatttctctaaagaaaataga  
agaagtatacttaagaagaggattt taa

#### D.5 Succinate dehydrogenase A (SDHA)

atgaaaattttacaagaaaaattcacaaaaattatagaacacacatatgatgcaatagtaataggtgctgggtgctggtttgagggc  
cactatgggttagctgaaaaaattataaagttgcttatttcaaaattatttccaactcggctcactactgtagctgctcaaggaggaatc  
aatgctgctctggggaatatgacacctgatgattggaatggcattttatgatactgttaagggaagtgattggttaggtgatcaagatgct

atcattatatgtccgagaagctcctaaagctgactgaattagaaaattatggttacccttttctagaactgaagatggaaaaatttate  
aaagagcttttgggtgagctcttggaaatggaaaaggaggacaagcttatcgaacagcctgtgcagcagatagaactggacacgc  
aatgctccatactttatggaaatcattaaaatgaaagctgattttatattgaatatttgccttagattattttaatcatgatagatcggct  
tgtattggagctttagcttataacatggaagatgggtgttatcatcaattacggctaaaaatacggttattgccactgggtggtatggaaagt  
gttattttcggcaaccagtgtcactactgtacaggtgatgtaatgctatggattaagacaaggtataaccactcaagattcagaattcgt  
tcaatttcaccaaccggagtttgggtgctggtgttttattgactgaaggatgctgaggtgaaggaggatttttagtcaatggacaaggag  
aaagatttatggaaagatatgcaccaagtgcaaaggatctagctagtcgagatgtagtatctagatctatgaccattgaaattaatgaagg  
aagaggtgctggcctaataaagatcatattttattacaattacatcattacccaagaaatctacaagatagacttccgggtatttctgaa  
acagctagtattttgccggtgctgatgtaactaaagaaccagttccagtaattccaacgggtcattataatgggaggtattccaactaatt  
ggaaagggcaagttattaatccatcttctaaagatgaacataatgttataatggactttggcagctggagaagctgctgttctcctgggt  
catggtgcaaatcgcttaggagctaaactcttattggatattgtagtattggtaaagcttgctgagaatattctgaaatggatcaacctg  
gtcagcatattgaatctattgataaaaatcaggtagataaattaattgatttatcaaaaactttttatcaaaaaggagatatacatattcag  
atthaagactagaatgcagaagataatgcaacaacatgctggagtatttcgaaataatgaatctttagaacaagggtgtttaaaaataaat  
caagtttatgatcaattaaacatgtcaaatagatgataaatcaaatattttaactgaattaccgaaatgctagaatttaagaacttatta  
gataatgctatagttacaattcatagtgccaattttagaaaagagtcgagaggagctcattctcatgatgattatcctgaaagagatgatca  
aaattggttagttcacactagcttatcttaaaatgatcgagtaaaatfatcaactgtagtgaatagaatactttaaatgatgaagttg  
attcaattccttagcaaaaagagttactaa

## D.6 Succinate dehydrogenase B (SDHB)

Atgaatatttttagaagaactttatcaacaactgtaaaaatataatcgcataatgctgttaagactaaaatggatcacattttcatttgataaaaa  
ggattgtggtcctatggtttagatgcttaatacatataaaaaataatttagataagactgtatcttttagaagatcttgcgggaaggtatct  
gtggatcatgtgctatgaatataaatggtaaaaatacattggcttgccttaaacctatggaagagaaaattcaatctatccattaccacata  
tgctgttattaagatttagttacagatctatcgaattttacaacaatataaagagattcaacctgggttaaaaaattcgtctacaggtcaa  
gaattaaacaatctcctcaagatcgtaaaaatagatggatgtatgaatgtattttatgtgctgtgttctacttcatgtccctcttattggg  
ggaatagtgataaatatttaggtccagctgttttaatgcaagcttatagatggattcaagattcaagagatcaggatactattaatcgtatga  
aatatgtagatgatgctatgaaattatcgtgtaaaacaattatgaattgtactaataactgtcctaaaaggattaaatcctggaaaagctat  
cggagagttaaaattaaaaatcaacaaggttacattaa