

# UNIVERSITETET I OSLO

## Det matematisk-naturvitenskapelige fakultet

**Eksamen i:**

**MBV4010 Arbeidsmetoder i molekylærbiologi og biokjemi I**

**Eksamensdag: Torsdag 25. september 2008**

**Tid for eksamen: Kl . 09.00 – 12.00**

**Oppgavesettet er på 8 sider.**

**Vedlegg: 3 (3 sider)**

**Tillatte hjelpemidler: Kalkulator**

*Kontroller at oppgavesettet er komplett  
før du begynner å besvare spørsmålene.*

*Les nøye gjennom hver enkelt oppgave før  
du begynner å besvare delspørsmålene.*

# UNIVERSITETET I OSLO

## Det matematisk-naturvitenskapelige fakultet

**Exam in:**

**MBV4010 Methods in molecular biology and biochemistry I**

**Day of exam: Thursday 25 September 2008**

**Exam hours: 09.00 – 12.00**

**This examination paper consists of 8 pages.**

**Appendices: 3 (3 pages)**

**Permitted materials: Calculator**

*Make sure that your copy of this examination paper is complete before answering.*

*Read thoroughly through the entire problem before starting to answer the sub-questions.*

## Oppgave 1

a) DNA-sekvensen vist under er del av sense-tråden til et protein-kodende gen. Translater sekvensen i de tre mulige leserammer (bruk den genetiske kode i vedlegg 1). Hvilke(n) av disse leserammene er åpne (uten stopp-kodoner)?

5' GGCAGGTAATGTGATTCA 3'

b) Noen DNA-polymeraser som brukes til PCR har såkalt "proofreading"-aktivitet. Hva betyr dette, og hvorfor kan dette være en nyttig egenskap?

c) Gel-elektroforese er mye brukt til å separere DNA-molekyler basert på størrelse. I hvilken retning vandrer et DNA-fragment i gelen; fra "minus"- til "pluss"-elektroden, eller fra "pluss" til "minus"? Hvorfor?

d) Hvilket (i, ii, eller iii) av forslagene for å lage 1 l av en 5 M løsning av NaCl (MW = 58,4) er korrekt? Forklar hvorfor.

i) Bland 292 g NaCl med 1 l H<sub>2</sub>O, og løs opp.

ii) Tilsett 292 g NaCl til en målekolbe (eller liknende), tilsett så H<sub>2</sub>O til et totalvolum på 1 l, og løs opp.

iii) Tilsett 292 g NaCl til en målekolbe (eller liknende), tilsett H<sub>2</sub>O til et totalvolum på ca. 0,9 l, og løs opp. Tilsett så H<sub>2</sub>O til et totalvolum på 1 l, og bland.

e) Noen restriksjonesenzymer med ulike gjenkjenningssekvenser gir fragmenter som lett kan liggeres sammen; pga. dannelsen av såkalte "kompatible kohesive ender". Nedenfor er det vist gjenkjenningsekvenser, kuttsetter (pil) og navn for seks par av restriksjonesenzymer som alle gjenkjenner klassiske palindromiske sekvenser. Indiker hvilke par som vil danne compatible kohesive ender, og indiker hvorfor.

i)	AclI	↓ AACGTT	and	NarI	↓ GGCGCC
ii)	MfeI	↓ CAATTG	and	AflII	↓ CTTAAG
iii)	NsiI	↓ ATGCAT	and	PstI	↓ CTGCAG
iv)	NarI	↓ GGCGCC	and	BbeI	↓ GGCGCC
v)	SalI	↓ GTCGAC	and	XhoI	↓ CTCGAG
vi)	StuI	↓ AGGCCT	and	SmaI	↓ CCCGGG

## Problem 1

a) The DNA sequence shown below is part of the sense strand of a protein-coding gene. Translate the sequence in the three possible reading frames (use the genetic code in appendix 1). Which of the reading frames are open (without stop codons)?

5' GGCAGGTAATGTGATTCA 3'

b) Some DNA polymerases used for PCR have a so-called "proof-reading" activity. What does this mean, and why can this property be useful?

c) Gel electrophoresis is commonly used to separate DNA molecules according to size. In what direction does a DNA fragment migrate in the gel; from the "minus" to the "plus" electrode, or from "plus" to "minus"? Why?

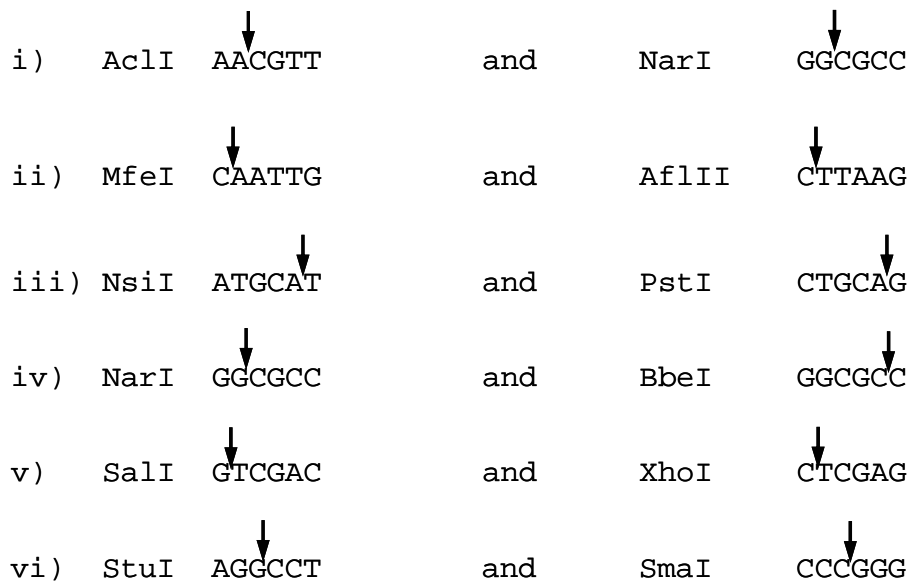
d) Which (i, ii, or iii) of the suggestions for preparing 1 l of a 5 M solution of NaCl (MW = 58.4) are correct? Explain why.

i) Mix 292 g NaCl with 1 l of H<sub>2</sub>O, and dissolve.

ii) Add 292 g NaCl to a volumetric flask (or similar), add H<sub>2</sub>O to a total volume of 1 l, and dissolve.

iii) Add 292 g NaCl to a volumetric flask (or similar), add H<sub>2</sub>O to a total volume of ca. 0.9 l water, and dissolve. Then add H<sub>2</sub>O to a total volume of 1 liter, and mix.

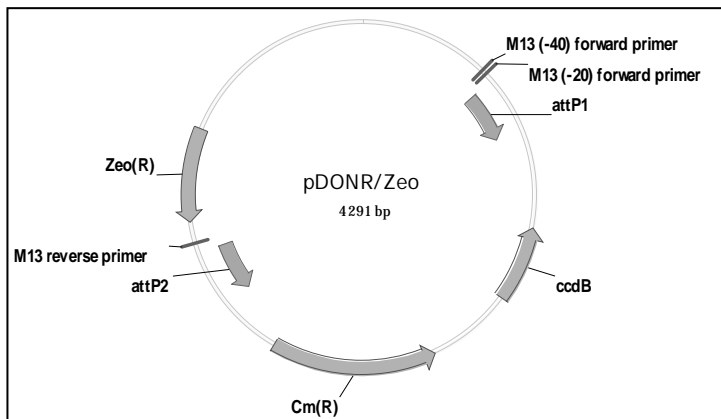
e) Some restriction enzymes that have different recognition sequences generate fragments that can be easily ligated together, due to the formation of so-called "compatible cohesive ends". Below are shown the recognition sequences, cleavage sites (arrow) and names of six pairs restriction enzymes that all recognize classic, palindromic sequences. Indicate which pairs that will form compatible cohesive ends, and indicate why.



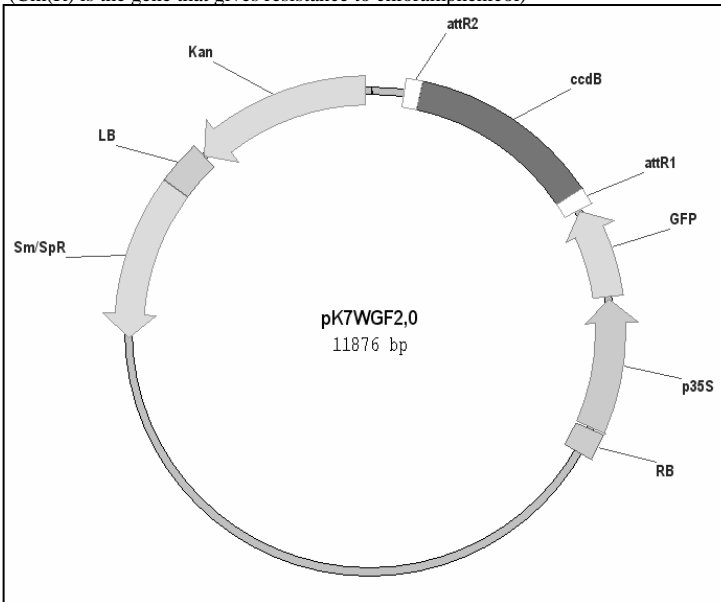
## Oppgave 2

a) Forklar kort prinsippene for Gateway-teknologien, og nevnt hvilke fordeler denne kloningsmetoden har sammenliknet med klassisk (restriksjonsenzym-basert) kloning.

b) Du skal studere den subcellulære lokaliseringen til SET-domene-proteinet SUVR4 fra *Arabidopsis* ved å bruke Grønt Fluoriserende Protein (GFP) som reporter. For dette formålet skal du klonere cDNAet som koder for proteinet inn i en destinasjonsvektor ved å bruke Gateway-teknologien. Lag en skjematisk tegning av kloningsstrategien du vil bruke, og beskriv hva som skjer i hvert trinn. Husk å ta med hvilken seleksjonsmarkør du trenger på LB-skålene du bruker, og indiker hvilke plasmider som vil gi opphav til overlevende kolonier på de respektive plater, og hvorfor. Du skal bruke komponentene vist under.



(Cm(R) is the gene that gives resistance to chloramphenicol)

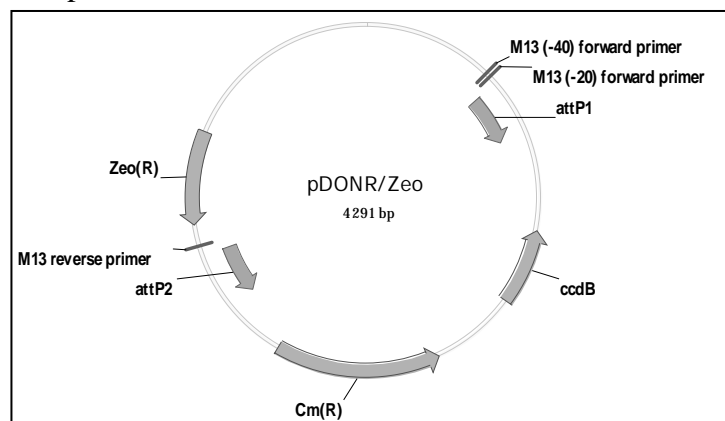


c) Du er interessert i å finne ut hvilken rolle SUVR4 har i utviklingen av *Arabidopsis*. Som en første tilnærming vil du, ved hjelp av RT-PCR, undersøke i hvilke vev dette genet er uttrykt. Gi en kort beskrivelse av hvordan du vil gå fram, og hvordan du vil designe primerne dersom genet har introns, slik at du kan skille mellom cDNA og kontaminering fra genomisk DNA.

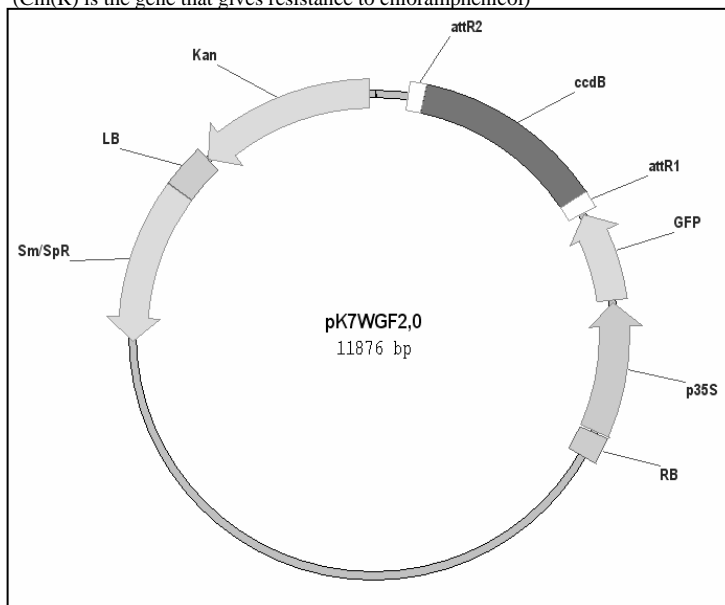
## Problem 2

a) Briefly explain the principles of the Gateway-technology, and mention which advantages this cloning method has compared to classic (restriction-enzyme based) cloning.

b) You are going to study the subcellular localization of the *Arabidopsis* SET domain protein SUVR4 by using the Green Fluorescent Protein (GFP) as a reporter. For this purpose you need to clone the cDNA encoding this protein into a destination vector using the Gateway technology. Make a schematic drawing of the cloning strategy you will use, and describe what is occurring in each step. Remember to include what kind of selection you need on your LB plates, and indicate which plasmids that will give rise to surviving clones on the respective plates, and why. You should use the components shown below.



(Cm(R) is the gene that gives resistance to chloramphenicol)



c) You are then interested in finding out what role SUVR4 has during development of *Arabidopsis*. As a first approach you decide to study, by RT-PCR, in which tissues this gene is expressed. Give a brief description of how you would proceed and how you would design your primers if the gene has introns, to be able to distinguish between cDNA and contamination from genomic DNA.

### Oppgave 3

I vedlegg 2 finner du DNA og proteinsekvensen til AlkB from *Agrobacterium tumefaciens* (AtAlkB). I vedlegg 3 finner du kart over vektoren pET-45b(+).

a) Design PCR primere for å klonere AtAlkB-genet mellom NcoI- (CCATGG) og XhoI- (CTCGAG) setene i pET-45b(+), for dermed å lage et plasmid som du kaller pET-45b-AtAlkB. Design forward primer slik at initiator ATG (koder for Met1) svarer til den ATG som finnes i NcoI-setet (CCATGG), and design reverse primer slik at en S-tag blir skjøtet på (i riktig leseramme) proteinets C-terminale ende. Primerne skal inneholde 20 nukleotider med sekvens som baseparrer med AtAlkB-genet. Siden noen restriksjonsenzymmer kutter dårlig ved enden av DNA-fragmenter, føy sekvensen TAAAAT til 5'-enden til primerne.

b) Under finner du deler av en sekvens-sammenstilling av AlkB-proteiner fra ulike bakterier (*Escherichia coli*, *Mycobacterium tuberculosis*, *Xanthomonas campestris*, and *Agrobacterium tumefaciens*). Du ønsker, ved å bruke QuikChange-metoden på pET-45b-AtAlkB, å mutere en konserverert His residu i AtAlkB til Trp. Denne residuen finnes som del av et konserverert HXD-motiv, og svarer til His131 i AlkB fra *Escherichia coli*. Du vil også introdusere en "stille" tilleggsmutasjon i nærheten, slik at et kuttesete for enzymet Eco47III (AGCGCT) blir introdusert. Du skal bruke PCR-primere som inneholder 12 nukleotider med ikke-mutert sekvens på hver side av de muterte nukleotidene. Skriv sekvensen på forward primeren du vil bruke, og indiker hvilke nukleotider som har blitt mutert. Plasmidet som er resultat av mutagenesen kaller du pET-45b-AtAlkB-mut.

```

Escherichia/1-216 : TMSVAMTNCGRLL---GWTTHRQGLYLSPIIDPQTNKPP-AMPQSFHNLCQRAATAAGYEDFQPD : 115
Mycobacterium/1-205 : QMYDRVVDVPRLL-----VSFHDLTIEDPPH---E---QLARMRRRLNDIYGGELGEE---ET : 107
Xanthomonas/1-201 : RMEGRVVDSPRLSSWIGDADASYRNISGTQFAPQ---EMLLEALQVVRTRIQD---ETGSE---EN : 99
Agrobacterium/1-195 : ---RIVQHFQ-----YRYDMLKVRVAVTEP---AMLGELPPWLGLEAERLVAD-GYCRTVPP : 91

Escherichia/1-216 : ACLINRYAEGAKLSLHODKDEEDL--RAPIVSVSLCLPAIEQEGGLKFNDDPLKRIIL-LEHGDVV : 176
Mycobacterium/1-205 : TAGLCYYRDSGSDSVAMHGDTIGRGSTEDTMVAIVSLCATRVFALRPRGRGESLR--EPLANGDLL : 170
Xanthomonas/1-201 : SVLVNRYRSGADAMGWHSDDEPEELGAQPVIASLSLCAARREAFRHR-HDAALKQTELELGHGDLL : 162
Agrobacterium/1-195 : QVTIANEYLLGQ-CISARVDC-VECF--DDTIVSISLSLACEMVFDL-RGEGIRSVLHPRSG--V : 149

```

c) Beregn størrelsen på fragmentene som blir resultatet av følgende kuttereaksjoner med restriksjonsenzymmer\*.

- pET-45b(+) kuttet med NcoI and XhoI.
- pET-45b-AtAlkB kuttet med NcoI and XhoI.
- pET-45b-AtAlkB kuttet med NcoI and Eco47III.
- pET-45b-AtAlkB-mut kuttet med NcoI and Eco47III.

\* Anta at det kun er ett enkelt kuttesete for NcoI and XhoI in pET-45b(+), og at det ikke er noen kutteseter for Eco47III, NcoI eller XhoI i AtAlkB-genet. Du kan utføre beregningene som om alle restriksjonsenzymmer kutter midt i sin gjenkjennelsessekvens.

### Problem 3

In appendix 2 you will find the DNA and protein sequence of AlkB from *Agrobacterium tumefaciens* (AtAlkB). In appendix 3 you will find a map of the vector pET-45b(+).

a) Design PCR primers for cloning the AtAlkB gene between the NcoI (CCATGG) and XhoI (CTCGAG) sites in pET-45b(+), to thereby obtain a plasmid you denote pET-45b-AtAlkB. Design the forward primer so that the initiator ATG (encoding Met1) corresponds to the ATG present in the NcoI site (CCATGG), and design the reverse primer so that an in-frame S-tag is fused to the C-terminus of the encoded protein. The primers should contain 20 nucleotides of sequence annealing to the AtAlkB gene, and, since some restriction enzymes tend to cleave poorly when present at end of DNA fragments, add the sequence TAAAAT to the 5' end of the primers.

b) Below you find part of an alignment of AlkB-proteins from various bacteria (*Escherichia coli*, *Mycobacterium tuberculosis*, *Xanthomonas campestris*, and *Agrobacterium tumefaciens*). You would like, by the using the QuikChange-method on pET-45b-AtAlkB, to mutate a conserved His residue in AtAlkB to Trp. This conserved His residue is found as part of a conserved HXD motif, and corresponds to His131 in AlkB from *Escherichia coli*. You will also introduce an additional, silent mutation in the close vicinity, leading to the introduction of recognition site for the restriction enzyme Eco47III (AGCGCT). You should use PCR-primers that contain 12 residues of non-mutated sequence flanking the mutated nucleotides. Write the sequence of the forward primer that you will use, indicating the mutated nucleotides. The plasmid resulting from the mutagenesis will be denoted pET-45b-AtAlkB-mut.

```

*           80           *           100           *           120           *
Escherichia/1-216 : TMSVAMTNCGRLL----GWTTHRQGLYSPIDPQTNKPPMP-AMEQSFHNLCQRAATAAGYEDFQPD : 115
Mycobacterium/1-205 : QMYDRVVDVPERL-----VSFHDLTIEDPPH---E---QLARMRRRNDIYGGELGEE---ET : 107
Xanthomonas/1-201 : RMEGRVVDSPRLSSWIGDADASYRISGTOFAEQ---PWLEALQVVRTRIQD----ETGSE---EN : 99
Agrobacterium/1-195 : ---RIVQHEG-----YRYDQKVRRAVTEP---AYLGEPLPWLGLFAERLVAD-CYCRTVPD : 91

          140           *           160           *           180           *
Escherichia/1-216 : ACLINRYAPGA-KLSLHODKDEEDL--RAPTVSVSLCLPATFQEGGLKENDPLKRLI-IERGDVV : 176
Mycobacterium/1-205 : TAGLCYYRDGSDSWAMHGDTIGRGSTEDTMVAIVSLCATRVEALFPRGRGPELRL--LPLANGDILL : 170
Xanthomonas/1-201 : SVLVNRYRSCADAMGWHSDD-EEELIGAQPVIASLSLCAARRFAFRHR-HDAALKQTEELGHCDDL : 162
Agrobacterium/1-195 : QVIANEYLLGQ-GISAHVDC-VECF--DDTIVSISLSLACEMVERDL-RGEGFRSVDLHPRSG--V : 149

```

c) Calculate the size of the DNA fragments resulting from the following restriction enzyme cleavage reactions\*.

- pET-45b(+) cut with NcoI and XhoI.
- pET-45b-AtAlkB cut with NcoI and XhoI.
- pET-45b-AtAlkB cut with NcoI and Eco47III.
- pET-45b-AtAlkB-mut cut with NcoI and Eco47III.

\* Assume that there is only a single site for NcoI and XhoI in pET-45b(+), and that there are no sites for Eco47III, NcoI nor XhoI in the AtAlkB gene. You may perform these calculations as if all restriction enzymes cleave in the middle of their recognition sequence.

# The genetic code

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

Amino acid	Abbreviation (three letters)	Abbreviation (one letter)
Alanine	Ala	A
Cysteine	Cys	C
Aspartate	Asp	D
Glutamate	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophane	Trp	W
Tyrosine	Tyr	Y

## Vedlegg 2 / Appendix 2

# AtAlkB

```
1 M D C V A N P T L L P H D I M Y F D G F
1 ATGGATTGCGTCGCTAATCCAACCTCTTCTACCGCACGACATCATGTATTTTGATGGCTTT
21 L S S E D E A F V A D R L D A G E W S T
61 CTCTCTTCAGAGGACGAGGCCTTCGTTGCGGACAGATTGGATGCTGGCGAGTGGAGCACG
41 E L K R R V Q H F G Y R Y D Y K V R A V
121 GAGCTGAAGCGTCGTGTGCAGCATTTCGGTTATCGTTATGACTATAAAGTGCGCGCTGTT
61 T P D A Y L G P L P P W L G L F A E R L
181 ACGCCAGATGCCTATCTCGGGCCTTTGCCGCCCTGGCTGGGCTTGTTCGCCGAACGATTA
81 V A D G Y C R T V P D Q V I A N E Y L L
241 GTGGCGGATGGCTACTGCCGGACAGTGCCGGATCAAGTTATTGCGAACGAGTATCTTCTG
101 G Q G I S A H V D C V P C F D D T I V S
301 GGGCAGGGGATCAGCGCGCATGTTCGATTGTGTCCCCTGTTTTGACGATACGATCGTATCG
121 I S L L S A C E M V F R D L R G P G I R
361 ATAAGCCTTTTATCCGCATGCGAGATGGTTTTCCGCGATTTGCGCGGACCGGGCATAACGC
141 S V L H P R S G V L L R G S S R Y D W T
421 AGCGTTCTGCATCCTCGTTCAGGCGTACTGCTCAGGGGCTCCAGCCGCTACGATTGGACT
161 H E I P A R K S D I V N G V K T A R S R
481 CATGAAATCCCGGCTCGCAAATCAGATATTGTTAACGGGGTAAAGACCGCCCGAAGCAGA
181 R I S L T F R K V L G L K R P -
541 AGGATTTCCCTCACGTTCCGCAAGGTCCTGGGTTTGAAACGGCCCTGA
```

