UNIVERSITETET I OSLO

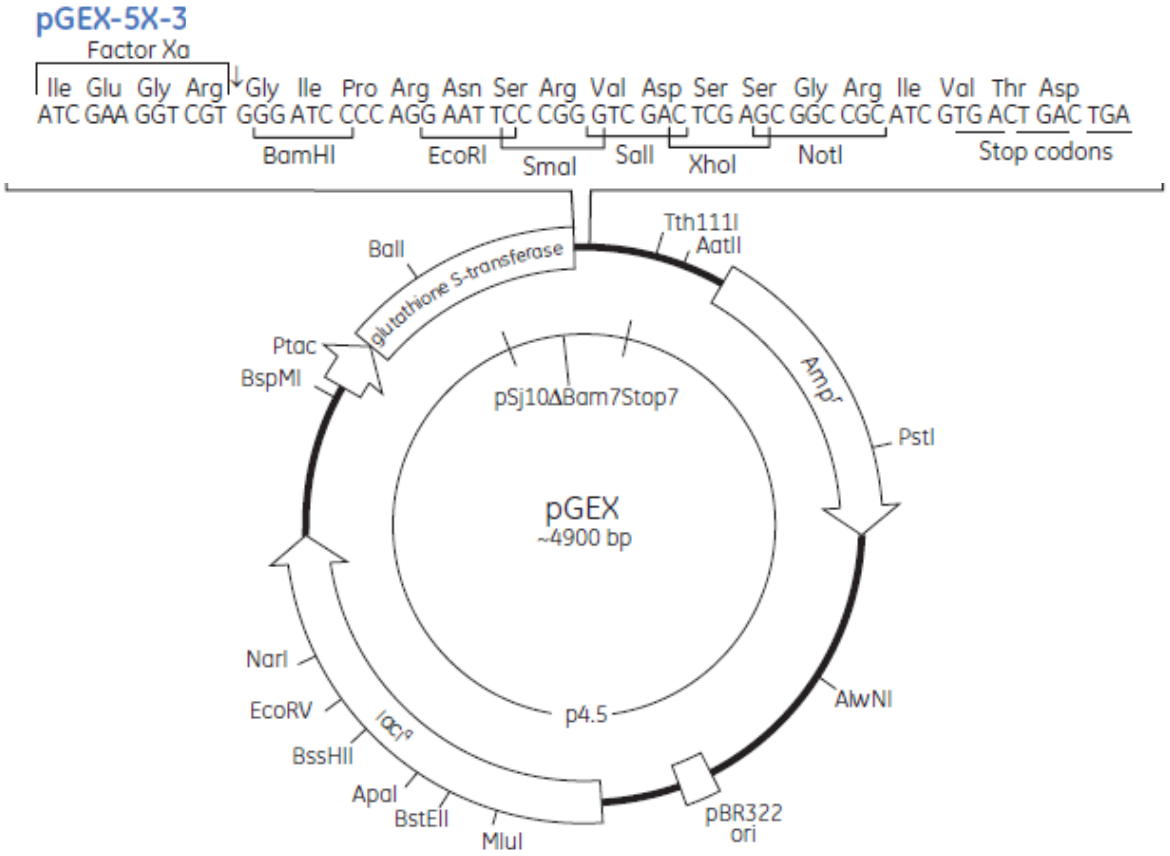
Det matematisk-naturvitenskapelige fakultet

Exam in:
MBV4010 Arbeidsmetoder i molekylærbiologi og biokjemi I
MBV4010 Methods in molecular biology and biochemistry I
Day of exam: Friday 25 September 2015
Exam hours: 10:00 – 13:00
This examination paper consists of 5 pages.
Appendices: 2 (2 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.

Problem 1

The Efm5 protein from the yeast *Saccharomyces cerevisiae* is a methyltransferase that specifically targets a lysine residue in a protein translation factor. You would like to clone the gene encoding Efm5 into the expression vector pGEX-5X-3 (illustrated below) for expression of the Efm5 protein as a fusion protein with glutathinone S-transferase (GST).

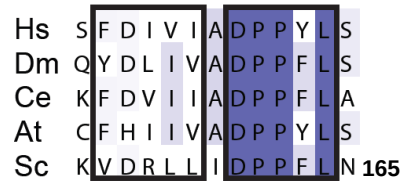


For this cloning you would like to use the enzyme NotI, and one of the other five enzymes (BamHI, EcoRI, SalI, SmaI, XhoI) found in the multiple cloning site region of pGEX-5X-3. EcoRI is the only one of these five enzymes that cuts inside the *EFM5* gene.

a)
Which two out of the five enzymes (BamHI, EcoRI, SalI, SmaI, XhoI) would you avoid using for this cloning? Why?

b)
In addition to NotI, you choose to use the enzyme BamHI for the cloning. In Appendix 1, you find the DNA and protein sequence of Efm5. Complying with the following criteria, design a pair of primers for cloning the Efm5 gene into pGEX-5X-3:
- Avoid adding extra amino acids to the GST-Efm5 fusion protein.
- To assure efficient restriction enzyme cleavage, add six extra nucleotides of DNA sequence 5' of the restriction enzyme recognition sites, indicated as NNNNNN.
- The length of the part of the primers that anneals to the EFM5 gene should be determined by using the Wallace rule so that the estimated melting point for the primer-template hybrid is ~54 ^{o}C The Wallace rule: Tm = 2(A+T) + 4(G+C)
Additional info: All the mentioned restriction enzymes recognize palindromic sequences.

c)
The *S. cerevisiae* (Sc) Efm5 protein and putative orthologues from other organisms such as nematode (Ce), fruitfly (Dm) and plant (At) share a conserved, functionally important DPP(F/Y)L motif, as indicated below.



Design pair of mutagenesis primers for mutating the second conserved proline residue of this motif into glycine, using the QuikChange method. The primers should be designed so that the introduction of the desired amino acid substitution is accompanied by the introduction of a SmaI recognition sequence, and the mutation(s) should be flanked by 12 nucleotides of non-mutated sequence on each side. In Appendix 2 you will find the genetic code and a list of amino acid abbreviations. If you find more than one solution to this problem, choose the one that gives the lowest number of nucleotide substitutions.

d)
What is the size of the small fragment obtained when cutting the plasmid resulting from the mutagenesis in c) with the enzymes SmaI and BamHI?

Problem 2

a)
The below DNA sequence encodes amino acids 24-28 of protein X. What is the actual amino acid sequence of this segment? (In Appendix 2 you will find the genetic code and a list of amino acid abbreviations)

CTACTGATTACTAGTCA

b)
Answer briefly (max 2-3 sentences and/or a simple drawing) to the following questions:

i)
What are the major differences between DNA polymerases and RNA polymerases?

ii)
What is PCR SOEing, and what are the primary applications of this technique?

iii)
Which important features make the budding yeast *Saccharomyces cerevisiae* a useful model organism?

iv)
What does it mean that a DNA polymerase has “proofreading” activity?

Problem 3

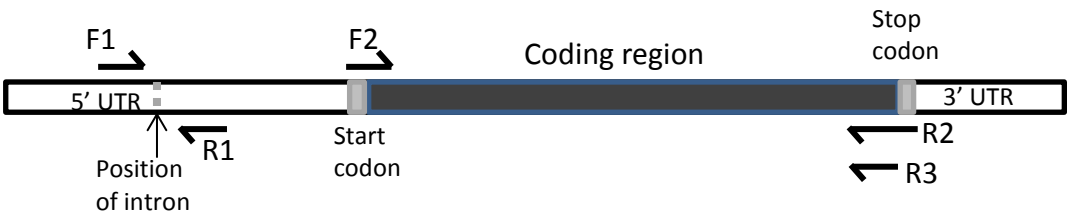
You want to investigate whether your favorite gene is expressed in different tissues in your model organism. Therefore you isolate RNA and make cDNA.

a) Your protocol says you should use RNAsin and DNAse. Explain their function and what the consequences might be if you do not use them.

b) You need to use reverse transcriptase for the cDNA synthesis. How does this enzyme work, and what does it require for function?

c) Explain what kind of controls you need to be sure that
- your mRNA isolation and cDNA synthesis has been successful
- your PCR set-up has been satisfactory

d) Your favorite gene encodes a protein and has one intron in the 5' UTR. Here is a schematic drawing of the full-length cDNA, showing 5' and 3' UTRs, coding region with start and stop codons, the position of the intron and the positions where two forward primers (F1 and F2) and three reverse primers (R1, R2 and R3) anneal.



Explain which primers you would use if you want to
- clone the cDNA for your gene into a vector designed for making a GFP-fusion protein with GFP in the C-terminal end
- clone the cDNA for your gene into a vector designed for making a GFP-fusion protein with GFP in the N-terminal end
- run PCR to check that your cDNA is not contaminated by genomic DNA
- run qPCR

e) You run qPCR on cDNA from three different tissues, having three biological and two technical replicates. Your lab has two different genes they often use for normalization.

In one experiment you got the Cp values seen in the table (next page).
Explain based on this table
- which of the two standards you would use
- whether the quality of each of your technical replicates is satisfactory
- whether it is likely that the expression level of your gene differs between the three tissues.

	Biological replilcates	Technical replicate 1	Technical replicate 2	Average Cp	Standard deviation
Favorite gene		Cp	Cp	Cp	
	Tissue 1				
	Repl1	30.03	30.10	30.07	0.05
	Repl2	29.82	29.78	29.8	0.03
	Repl3	29.75	32.40	31.075	1.87
	Average	29.87	30.76		
	Tissue 2				
	Repl1	35.66	35.24	35.45	0.30
	Repl2	33.60	33.40	33.5	0.14
	Repl3	33.20	32.99	33.095	0.15
	Average	34.15	33.88		
	Tissue 3				
	Repl1	30.07	29.8	29.935	0.19
	Repl2	29.81	30.19	30	0.27
	Repl3	29.55	29.56	29.555	0.01
	Average	29.81	29.85		

	Biological replilcates	Technical replicate 1	Technical replicate 2	Average Cp	Standard deviation
Standard A					
	Tissue 1				
	Repl1	30.52	30.34	30.43	0.13
	Repl2	29.67	30.05	29.86	0.27
	Repl3	30.75	34.47	32.61	2.63
	Average	30.31	31.62		
	Tissue 2				
	Repl1	22.49	22.62	22.56	0.09
	Repl2	22.32	22.49	22.405	0.12
	Repl3	22.21	22.50	22.355	0.21
	Average	22.34	22.54		
	Tissue 3				
	Repl1	21.94	22.00	21.97	0.04
	Repl2	22.68	22.58	22.63	0.07
	Repl3	22.07	22.18	22.125	0.08
	Average	22.23	22.25		

	Biological replilcates	Technical replicate 1	Technical replicate 2	Average Cp	Standard deviation
Standard B					
	Tissue 1				
	Repl1	21.94	22.00	21.97	0.04
	Repl2	22.68	22.58	22.63	0.07
	Repl3	22.07	24.13	23.1	1.46
	Average	22.23	22.90		
	Tissue 2				
	Repl1	22.49	22.62	22.56	0.09
	Repl2	22.32	22.49	22.405	0.12
	Repl3	22.21	22.50	22.355	0.21
	Average	22.34	22.54		
	Tissue 2				
	Repl1	21.94	22.00	21.97	0.04
	Repl2	22.68	22.58	22.63	0.07
	Repl3	22.07	22.18	22.125	0.08
	Average	22.23	22.25		

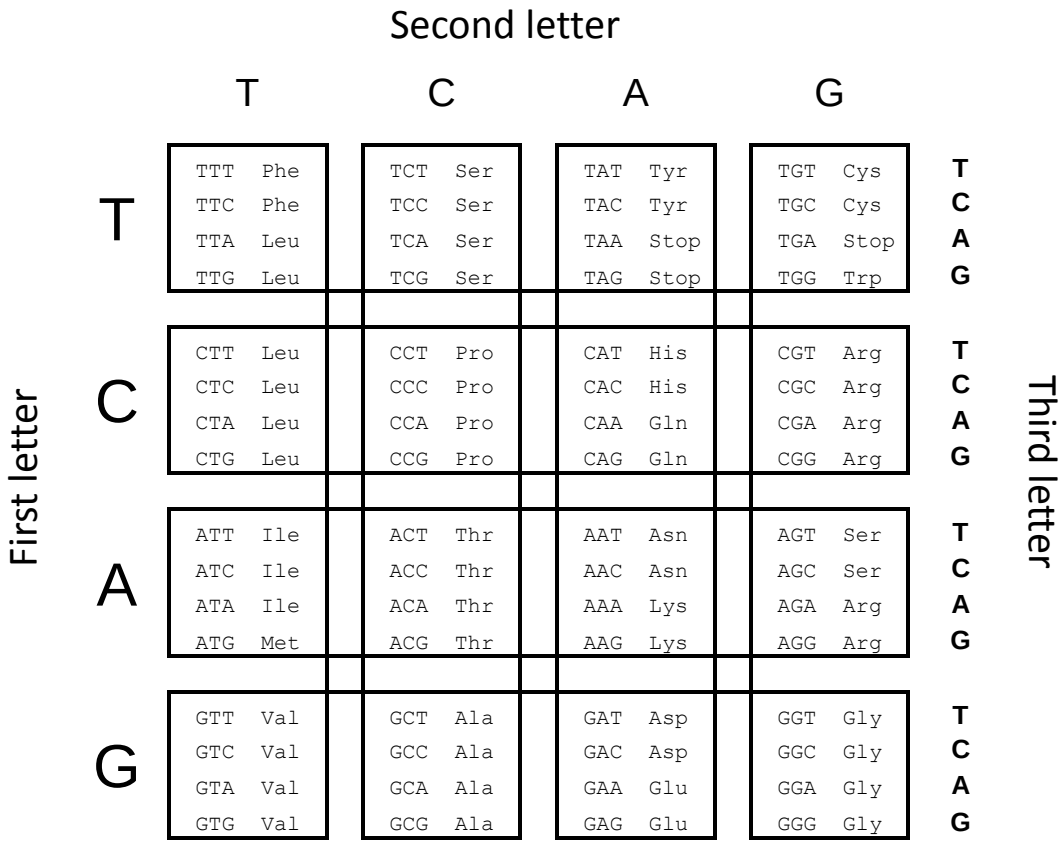
Appendix 1. DNA and protein sequence of *S. cerevisiae* Efm5

(248 amino acids, 747 nucleotides incl. stop codon)

```
  1   M  S  D  S  D  S  D  S  D  Y  E  L  T  L  S  A  N  A  L  A
  1  ATGTCCGACTCCGACTCCGACTCCGATTATGAATTGACACTTTCTGCTAATGCCCTCGCT

 21   A  L  E  E  F  K  R  E  E  Q  Q  H  Q  E  A  F  Q  K  L  Y
 61  GCCCTTGAAGAATTCAAAAGAGAGGAACAACAACATCAAGAAGCCTTTCAAAAGCTTTAC

 41   D  E  T  D  E  D  F  Q  K  K  K  K  E  E  G  M  K  L  F  K
121  GACGAAACGGATGAAGACTTCCAAAAGAAAAAAAAAAGAAGAAGGGATGAAGCTTTTCAAG

 61   E  D  W  Q  L  S  Q  F  W  Y  S  D  D  T  A  A  I  L  A  D
181  GAAGATTGGCAGCTTTCCCAGTTTTGGTACAGCGATGACACAGCCGCAATTTTAGCAGAT

 81   A  I  L  E  G  A  D  E  N  T  V  I  A  I  V  S  A  P  S  V
241  GCCATATTGGAAGGTGCGGACGAAAACACTGTAATTGCAATAGTTAGTGCGCCATCCGTC

101   Y  A  A  I  Q  K  K  P  T  N  E  I  P  T  E  H  I  Y  L  F
301  TATGCTGCCATTCAAAAGAAACCTACTAATGAAATTCCAACCGAACATATCTACTTGTTC

121   E  F  D  K  R  F  E  L  L  A  G  R  D  H  F  F  F  Y  D  Y
361  GAATTTGATAAGAGATTCGAGTTGTTAGCTGGAAGAGATCATTTCTTTTTTTATGATTAC

141   N  K  P  L  D  F  S  D  E  I  K  G  K  V  D  R  L  L  I  D
421  AATAAACCACTGGACTTCAGTGATGAAATTAAAGGAAAAGTTGATAGATTGTTAATTGAC

161   P  P  F  L  N  E  D  C  Q  T  K  S  S  I  T  A  K  C  L  L
481  CCACCTTTTTTAAATGAAGATTGTCAAACAAAGTCCTCTATCACTGCAAAGTGTTTATTA

181   A  P  N  D  N  S  K  T  K  K  G  V  F  K  H  R  L  I  S  C
541  GCGCCAAATGACAACTCTAAAACTAAGAAAGGAGTTTTTAAGCACCGCCTCATAAGTTGT

201   T  G  E  R  M  S  E  V  I  S  K  V  Y  S  D  T  R  I  T  T
601  ACTGGTGAAAGAATGTCTGAAGTCATATCCAAAGTTTATTCTGACACAAGAATAACAACT

221   F  L  P  E  H  S  N  G  L  S  N  E  F  R  C  Y  A  N  F  E
661  TTTCTTCCTGAACATTCTAATGGTTTGAGTAACGAATTTAGGTGCTACGCAAATTTTGAA

241   C  S  S  W  K  F  A  S  -
721  TGTAGTTCATGGAAATTTGCTTCTTGA
```

Appendix 2

The genetic code



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