BCH3356 *Molecular Biology Laboratory*

This final exam is worth 40 % of the final mark for BCH3356

Name:_____
Student number:_____

Instructions

All questions should be answered within the space provided on THIS COPY EXAM.

You may use the endorsement of pages for your draft calculations.

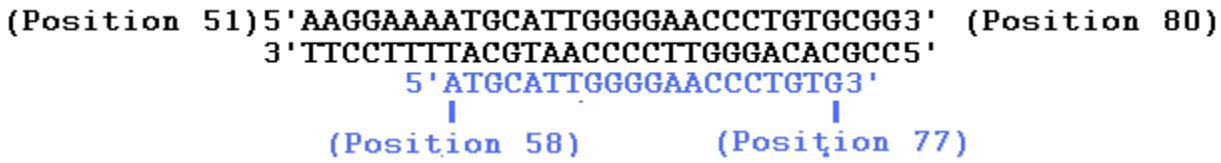
Students are allowed to use a calculator.

This is a closed book exam

Part I (/30 marks): *Basic Molecular Biology Techniques*

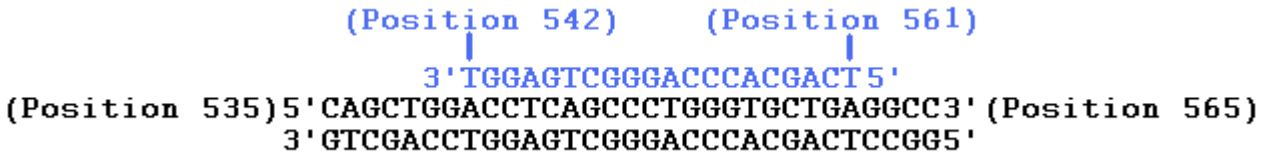
1. **5 MARKS: RT-PCR** Design a pair of PCR primers that could be used to specifically amplify the coding part of the human leptin mRNA (sequence displayed in Appendix II). The 5'UTR and 3'UTR should be excluded from your PCR product. Each of your two primers should contain 20 nucleotides.

   a. Forward primer: 5'_____3'
   Annealing position is from position _____ to position _____ along the mRNA template provided in Appendix II.



   1 mark for primer sequence, 0.5 mark for each alignment position number (58 and 77)

   b. Reverse primer: 5'_____3'
   Annealing position along is from position _____ to position _____ along the mRNA template provided in Appendix II..



   1 mark for primer sequence, 0.5 mark for each alignment position number (58 and 77)

   c. Which of the two primers should be used for the RT preparatory step? Why?
   The reverse primer (0.5 mark) as it is the only one that can anneal onto the mRNA (0.5 mark)

2. **5 MARKS: PROBE SYNTHESIS FOR NORTHERN AND SOUTHERN ANALYSES**
   a. Describe the expected result for the synthesis of a digoxygenin-labeled probe complementary to the fragment 251-550 of the cDNA sequence coding for the human leptin gene (see sequence in appendix II) if the two following primers were used:

      Primer 1: 5'ggctccaccccatcctgacc3'
      Primer 2: 5'caccgtcgacctggagtcgg3'

   Primer 1 is a 'functional' forward primer, although primer 2 is a 'dysfunctional' reverse primer as it cannot anneal onto mRNA (its sequence should be reversed to anneal onto mRNA) (1 mark).
   There will be no amplification (1 mark)

   b. For question b and c, you should assume that a double-stranded probe had been successfully prepared. Which of the two strands of the probe having been successfully synthesized would effectively hybridize onto the mRNA coding for human leptin if a Northern analysis were performed? Justify your answer.

The reverse primer (primer 2) is the one that can anneal with mRNA. The strand whose elongation is initiated with primer 2 is accordingly the only strand that could hybridize during Northern analysis (1.5 marks)

c. Which of the two strands of the probe having been successfully synthesized would effectively hybridize onto the genomic DNA coding for human leptin if a Southern analysis were performed? Justify your answer.
Genomic DNA is double stranded so that each of the two probe strands could anneal onto their complementary gDNA strand (1.5 marks).

3. **5 MARKS: MUTAGENESIS** During the semester you used DpnI to selectively digest the parental non mutated plasmid DNA. How is it that DpnI can selectively digest and eliminate the non mutated plasmid DNA without affecting the copies of the plasmid having been mutated? Explain your answer.

Non mutated, parental plasmid DNA was obtained by miniprep from a cell culture. Parental plasmid DNA was then DAM methylated in vivo (2.5 marks).
Mutated DNA was amplified by PCR in vitro and was accordingly NOT methylated (2.5 marks).
DpnI can only digest methylated DNA, i.e., the non mutated, parental plasmid DNA.

4. **5 MARKS** Genetic features of cloning vectors
   a. Describe the genetic feature of pBluescript regulating its replication.
   RNAII, a RNA molecule encoded by pBluescript, can anneal near the origin of replication and, upon partial digest by an RNase, initiate replicative elongation at its free 3' end. RNAI can form a dimer with RNAII and accordingly minimize its activity. Finally, the ROP protein can bind onto the RNAII/RNAI complex and then lessen replication. 1.5 marks

   b. How is it that all colonies containing cells transformed with the pET-15b vector were white whereas some blue and white colonies were noticed with the pBluescript? Your answer should emphasize the underlying molecular principles.

   The pET-15b vector lacks the colorimetric screening system (1 mark) that is found in pBluescript and which is based on the alpha complementation of galactosidase (1 mark). All transformant colonies containing the pET-15b vector are accordingly white.

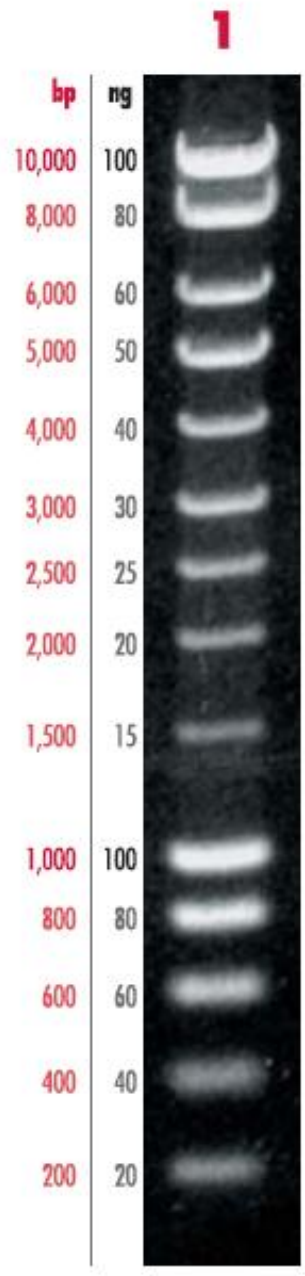
   c. Would it be possible to ligate an insert outside the multiple cloning site of pBluescript? If no, explain why. If yes, explain how the cloning strategy might be affected?
   This should be considered as an open question. A yes or no answer was considered as correct if justified properly. Yes, it would be possible to ligate an insert outside the MCS of pBluescript, but that would inactivate the blue/white colorimetric screening system and possibly alter other house keeping genes such as the one coding for resistance to ampicillin. In other words, it might be possible to ligate an insert outside the MCS, but it is not recommended.

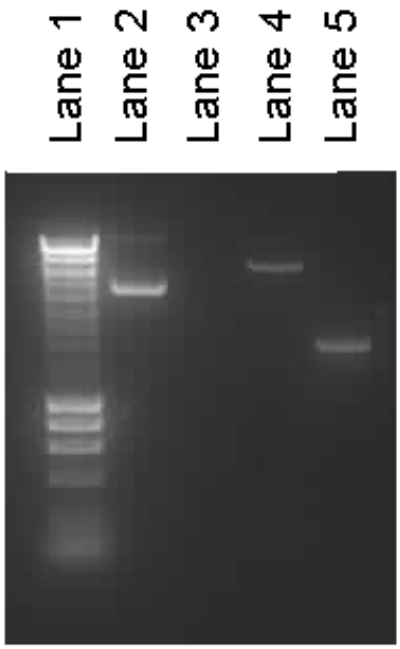
5. **5 MARKS** Define the following concepts: 1 mark per definition
    a. Denaturing agarose gel:
    Separation technique used to resolve RNA or DNA molecules on an agarose gel under denaturing conditions to minimize conformational effects.
    b. Electroporation:
    Genetic transformation procedure in which short electrical pulses are applied to a cell suspension to form transient pores into the cell membrane and facilitate the incorporation of plasmid DNA.
    c. Preproalbumin:
    The precursor of serum albumin (containing the signal peptide at the N terminus along with a short peptide that is cleaved in the Golgi apparatus to produce the secreted albumin).
    d. Primary antibody:
    The antibody used in Western analysis to specifically target the antigen or protein of interest.
    e. Base-calling:
    Computerized procedure used to analyze and convert the fluorescence profile recorded at the output of the capillary tube of a DNA sequencer into the corresponding sequence of nucleotide bases.

6. **5 MARKS** Explain the molecular principle for the alkaline lysis miniprep procedure originally designed by Dr. Birnboim, a researcher from the Biochemistry, Microbiology and Immunology department at the University of Ottawa. Use point form to explain what specifically happens at each of the different steps of the procedure.
    - Pellet of cells is suspended with the suspension buffer
    - Lysis buffer with SDS and NaOH is added to denature DNA and proteins (2 marks)
    - Neutralization buffer with KAc is added; plasmid DNA reanneals faster than genomic DNA that aggregates with proteins (2 marks)
    - Centrifugation to pellet genomic DNA and protein aggregate; plasmid DNA remains in the supernatant (1 mark)

Part II (/40 marks): *Analytical Skills*

1. **10 MARKS** The figure below displays the AlphaQuant 1 DNA ladder along with a gel picture obtained by one group during the semester. Assuming that 5ul had been loaded in each lane, explain how you would prepare the aliquots for (1) the PCR product and (2) the pET15-b vector for proceeding to a ligation reaction with a 7:1 insert:vector molar ratio. Include all your calculations and explain your reasoning assuming that the amount required for ligation is 20ng and that the insert and plasmid have a length of 1000 bps and 5000 bps, respectively. Use next page to answer this question.



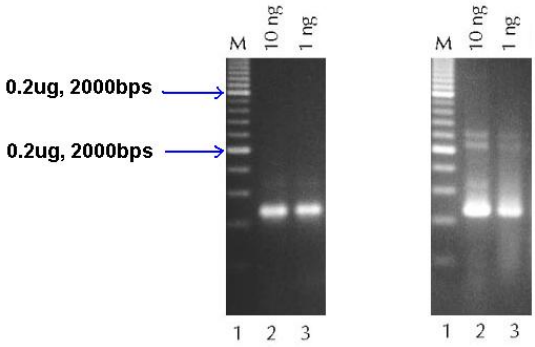


Lane 1: Alpha DNA ladder
Lane 2: Undigested control
Lane 3: Blank
Lane 4: Digested/purified pET15-b
Lane 5: Digested/purified PCR product

This blank sheet should be used to answer question 1 of Part II.

- Lane 4 from gel: about 30ng/5ul of digested and purified pET15-b vector5
- Lane 5 from gel: about 50ng/5ul of digested and purified insert
- The question specifies that 20ng (of vector) is required for ligation:
    - 20ng x (1nmol bps/660ng) x (1nmol vector/5000nmol bps) = $6x10^{-6}$nmol vector
    - 20ng/(30ng/5ul)=3.3ul of the vector solution loaded on the gel is required
- Cloning molar ratio of 7: $42x10^{-6}$nmol insert is necessary
    - Amount of insert required in ng
        - $42x10^{-6}$nmol ins. x (1000nmol bps/nmol ins.) x (660ng/nmol bps) = 28ng
    - Volume of insert required for ligation
        - V = q/C = 28ng/(50ng/5ul) = 2.8 ul of the insert solution loaded on the gel
- 5 marks for correct answer
- 5 marks for reasoning

- **10 MARKS** Two students use RT-PCR to amplify the same cDNA fragment coding for a protein, phtalalamin, with two different amounts of the same total RNA template, i.e 10ng and 1ng. The two gels of the RT-PCR products are shown below. Lanes 1, 2 and 3 are for DNA markers, and the RT-PCR products obtained using 10ng and 1 ng of total RNA, respectively.

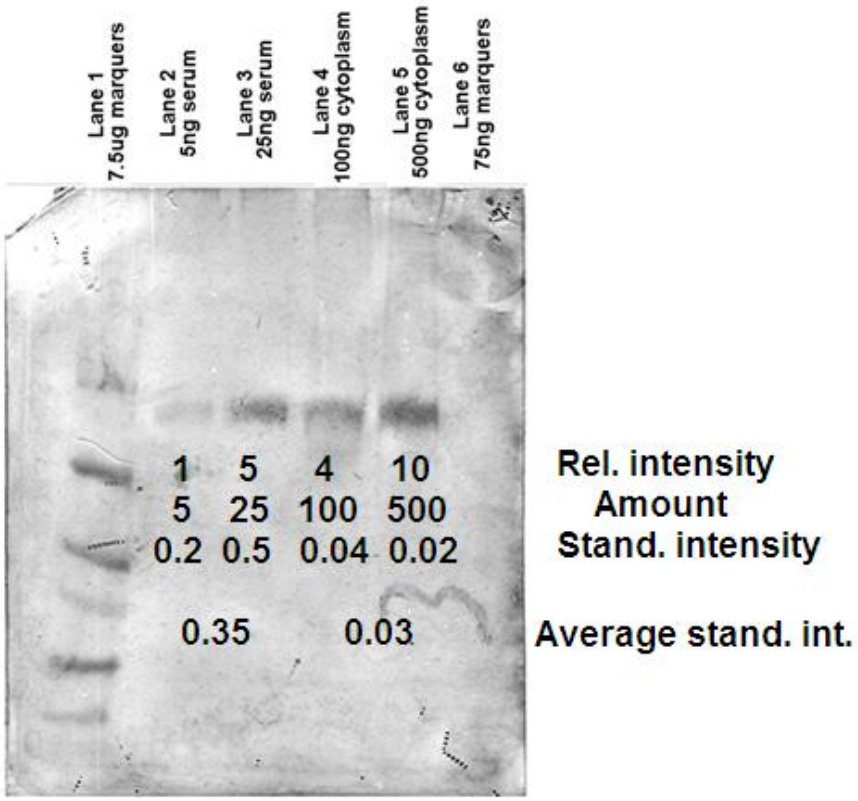


a. Assuming that all samples from both students were run under similar conditions, explain what might justify the non specific amplification that is noticed on the gel displayed at the right.
Was not graded due to redundancy between a and b.

b. Identify three PCR parameters that could be modified by the students whose results are displayed at the right to improve their amplification specificity. Also explain how each of those three parameters can affect the specificity of PCR amplification.
   - Better primer design: make primers longer or more specific, adjust GC content
   - Verify specificity of primers using BLAST
   - Raise annealing temperature up to 5C below the mean annealing temperature of the two primers
   - Keep PCR mixture on ice until addition of Taq then immediately transfer to the thermocycler to begin amplification
   - …
   - 2 marks per correct answer for a total of 6 marks

c. The amount of cDNA amplified is comparable for both amounts of the RNA template used, i.e., 10ng and 1ng. How is it that the amount of PCR product amplified is not linearly related with the initial amount of template added to the PCR mixture?
   - Amplification is exponential
     - # of amplified strands = $\#_0 \times 2^n$
     - Where $\#_0$ is the number of initial copies and n the number of amplification cycles
   - The number of strands amplified should be linear with the number of initial copies (or template amounts), but it is not the case. One can thus conclude that lack of linearity between amounts amplified and initial amounts of RNA is due to the fact that something else (dNTPs, enzyme availability, primers, …) was limiting during amplification (a plateau for PCR amplification was reached).
   - 2 marks for reference to linear amplification
   - 2 marks for reference to a saturation effect

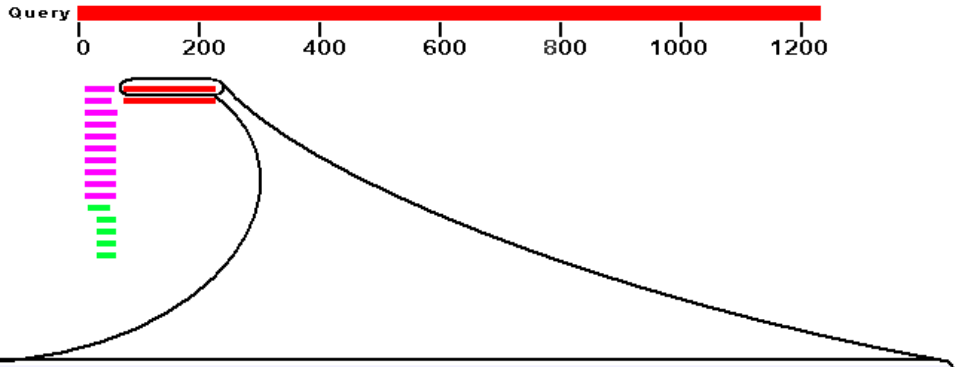
3. **10 MARKS** The Western analysis result obtained by one group during the semester is displayed below.



a. Refer to the results displayed on the figure just above to discuss the relative abundance of rat albumin between the two fractions that were investigated, i.e., serum and liver cytoplasm. Show your calculations and justify your reasoning.
Calculations displayed on gel picture. Conclusion is that rat serum albumin is roughly 12 times more abundant in the serum than the cytoplasmic fractions.
2 marks for relative intensities
2 marks for standardized intensities
2 marks for conclusion with quantitative appreciation of relative abundance

b. What can you say about the specificity of the primary antibody used for this Western analysis with regard to the information specifically obtained from the protein markers?
Primary antibody was raised against rat serum albumin. It should accordingly not detect any of the protein marker bands if it is specific. No band was detected in lane 6 for which 75 ng of the protein marquers had been loaded suggesting good specificity of the primary antibody. For lane 1, a larger amount of the protein markers, i.e., 7.5 ug, had been used to ensure that the protein marker bands can be seen even without any detection. The analysis of the protein markers thus suggests that the primary antibody was relatively specific for the conditions used for the Western analysis.
2 marks for reference to amounts of the markers
2 marks for proper conclusion

4. **10 MARKS** One sequencing result obtained during the semester was submitted to nucleotide BLAST and the NCBI server returned the information displayed below.




```
 Score =  289 bits (150),  Expect = 3e-74
 Identities = 150/150 (100%), Gaps = 0/150 (0%)
 Strand=Plus/Plus

Query  80    AGTGGGTCCTTGGTGGAAAGACGGCCATGTTTCTCTGCTCTGACAGTTGACGAGACATAT  139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1504  AGTGGGTCCTTGGTGGAAAGACGGCCATGTTTCTCTGCTCTGACAGTTGACGAGACATAT  1563

Query  140   GTCCCCAAAGAGTTTAAAGCTGAGACCTTCACCTTCCACTCTGATATCTGCACACTCCCA  199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1564  GTCCCCAAAGAGTTTAAAGCTGAGACCTTCACCTTCCACTCTGATATCTGCACACTCCCA  1623

Query  200   GACAAGGAGAAGCAGATAAAGAAGCAAACG  229
             |||||||||||||||||||||||||||||
Sbjct  1624  GACAAGGAGAAGCAGATAAAGAAGCAAACG  1653
```

Query sequence is sequencing result and Sbjct sequence corresponds to rat serum albumin(V01222)

a. How is it that many alignments are matching the very first part of the sequencing results, i.e., the roughly 40 first nucleotides?
The first 40 nucleotides having been sequenced correspond to the pBluescript fragment located between the T7 primer binding site and the insertion site. It appears that a few sequences reported in the mRNA rat database were cloned and sequenced from vectors that are homologous to the fragment of pBluescript having been sequenced.
2 marks for reference to the pBluescript fragment having been sequenced
2 marks for reference to many pBluescript variants

b. For the purpose of this question, you should assume that the first 1000 out of the nearly 1225 nucleotides that had been sequenced gave a reliable signal with well resolved peaks. What can you conclude with regard to the nature of the insert that had been sequenced knowing that the BLAST nucleotide alignment was performed against the mRNA rat database. Justify your reasoning.
From the 1000 nucleotides having been accurately sequenced, only the 80-229 stretch was aligned against the rat serum albumin mRNA. That means that the remaining sequence located between positions 230 and 1000 of the query or sequencing result could not be aligned with something else from the mRNA rat database. It is likely that the 230-1000 fragment of the sequencing result is an intron. The sequenced insert was thus likely a genomic fragment of the rat serum albumin gene with an exon at its 5' end and an intron at its 3'end.
3 marks for reference to exon
3 marks for conclusion: genomic DNA insert coding for a fragment of rat serum albumin gene

Part III (/30 marks): *Subcloning Project*

1. For the purpose of this question you should assume that the exact full length cDNA coding for the human leptin gene (Appendix 2) was directionally ligated at the SmaI and XbaI sites of the yT&A vector with the start and stop codons at the SmaI and XbaI sites, respectively. Refer to Appendix 3 to get the complete list of recognition sites within the leptin sequence (notice that there are two tables: one for the enzymes that cut only once and another with enzymes that cut twice).

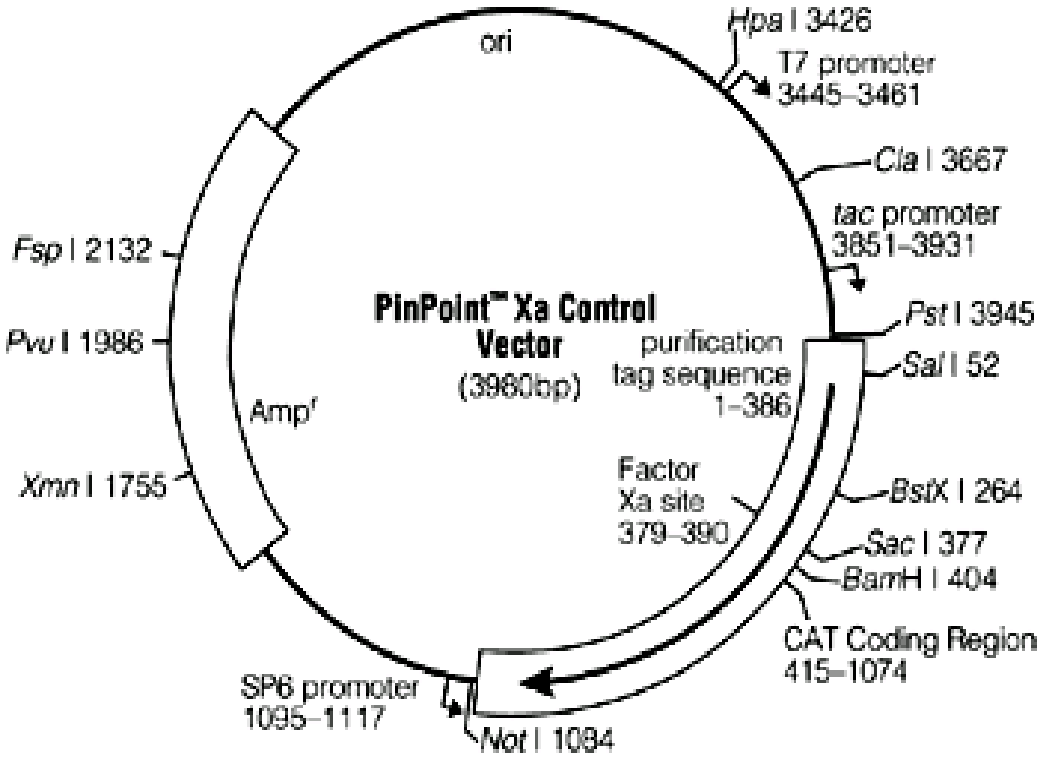
   a. **10 Marks** Design two primers that you could use to amplify the exact full length coding cDNA insert (5' and 3' UTRs to be excluded) from the recombinant yT&A plasmid in order to subclone it into the pinpoint Xa plasmid (see details in Appendix Ia and Ib). Justify your reasoning.
   This question is essentially a repeat of question 1 of part 1, except that recognition sites have to be introduced at the 5' end of each primer. The recognition sites added to the primers should (1) be found within the MCS of the receptor plasmid, i.e., pinpoint Xa plasmid and (2) not found within the insert. The recognition sequences of the following six enzymes are the only ones to respect those two criteria: Kpn1, EcoRV, BgIII, SmaI, Acc651 and BamHI.
   2.5 marks for each of the two annealing sequences of the primers
   2.5 marks for each of the two recognition sequences

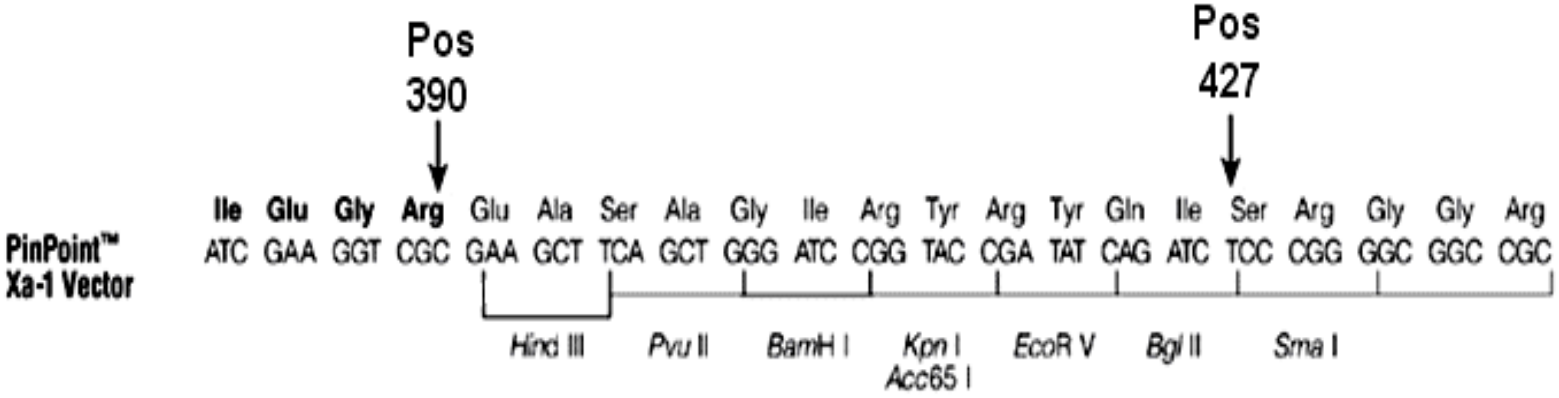
   b. **5 Marks** For questions b and c, you should assume that ligation of the PCR product amplified with the two primers listed in a had been successfully done into the PinPoint Xa Control vector. After transformation had been done, you noticed 15 white colonies for the negative ligation control performed with the pinpoint Xa plasmid pre-digested with your two enzymes specified in a. How could you explain the presence of those colonies for that negative ligation control? Should you systematically repeat the transformation treatment? Justify your reasoning.
   If white colonies are observed, that means that the 15 transformant colonies contain the PinPoint Xa vector. It might be that some of the plasmids were cut only once leaving complementary ends that could have been ligated. It is also possible that a few multimers of the plasmid were ligated.
   As for the second question, more meticulous individuals might prefer to repeat the whole transformation procedure whereas more opportunistic ones might prefer to screen a larger range of transformant colonies hoping that at least one has the expected insert. Reasoning should be provided in either option.
   2.5 marks for providing at least one explanation for the presence of background white colonies.
   2.5 marks for deciding and justifying whether to repeat or not the transformation procedure

15 Marks Explain a screening strategy based on restriction enzymes that you could use to digest the miniprep products of five transformant colonies observed after transformant cells had been plated onto an agar petri dish with ampicillin. You should specify the length of the expected fragment(s).
7.5 marks for reasoning
7.5 marks for correct lengths of restriction fragments for both empty and recombinant plasmids

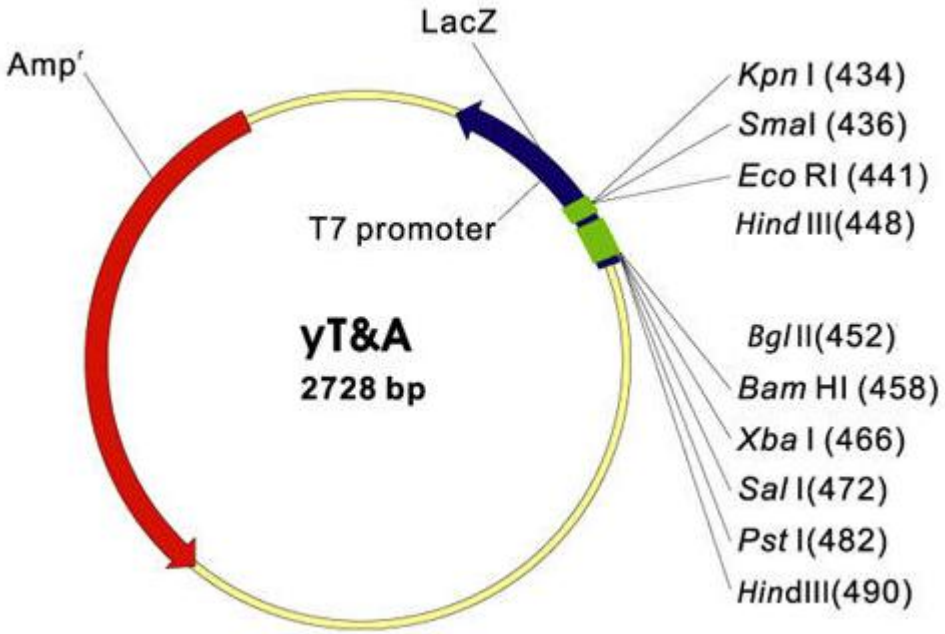
Appendix Ia: Information Relative to the Pinpoint Xa Cloning Vector

The PinPoint™ Xa Protein Purification System is designed for the production and purification of fusion proteins that are biotinylated in vivo. Biotinylated fusion proteins are produced in *E. coli* and are affinity-purified using the SoftLink™ Soft Release Avidin Resin. This proprietary resin allows elution of the fusion protein under nondenaturing conditions. The PinPoint™ Vectors feature the encoded endoproteinase Factor Xa (pronounced “ten a”) proteolytic site that provides a way to separate the purification tag from the native protein. The PinPoint™ Xa Control Vector contains the chloramphenicol acetyltransferase (CAT) gene and is provided as a means of monitoring protein expression, purification and processing conditions.





Appendix Ib: Informative Relative to the yT&A cloning vector



Appendix 2: Homo sapiens leptin (LEP), mRNA

```
REFERENCE   1  (bases 1 to 3444)
PRIMARY     REFSEQ_SPAN         PRIMARY_IDENTIFIER PRIMARY_SPAN        COMP
            1-309               DA762132.1         1-309
            310-1807            U43653.1           309-1806
            1808-3383           AC018635.10        34398-35973
            3384-3444           BU752306.1         1-61                c
FEATURES             Location/Qualifiers
     gene            1..1002
     CDS             58..561
                     /translation="MHWGTLCGFLWLWPYLFYVQAVPIQKVQDDTKTLIKTIVTRIND
                     ISHTQSVSSKQKVTGLDFIPGLHPILTLSKMDQTLAVYQQILTSMPSRNVIQISNDLE
                     NLRDLLHVLAFSKSCHLPWASGLETLDSLGGVLEASGYSTEVVALSRLQGSLQDMLWQ
                     LDLSPGC"
     sig_peptide     58..120
ORIGIN
        1 gtaggaatcg cagcgccagc ggttgcaagg cccaagaagc ccatcctggg aaggaaaatg
       61 cattggggaa ccctgtgcgg attcttgtgg ctttggccct atcttttcta tgtccaagct
      121 gtgcccatcc aaaaagtcca agatgacacc aaaaccctca tcaagacaat tgtcaccagg
      181 atcaatgaca tttcacacac gcagtcagtc tcctccaaac agaaagtcac cggtttggac
      241 ttcattcctg ggctccaccc catcctgacc ttatccaaga tggaccagac actggcagtc
      301 taccaacaga tcctcaccag tatgccttcc agaaacgtga tccaaatatc caacgacctg
      361 gagaacctcc gggatcttct tcacgtgctg gccttctcta agagctgcca cttgccctgg
      421 gccagtggcc tggagacctt ggacagcctg gggggtgtcc tggaagcttc aggctactcc
      481 acagaggtgg tggccctgag caggctgcag gggtctctgc aggacatgct gtggcagctg
      541 gacctcagcc ctgggtgctg aggccttgaa ggtcactctt cctgcaagga ctacgttaag
      601 ggaaggaact ctggcttcca ggtatctcca ggattgaaga gcattgcatg gacacccctt
      661 atccaggact ctgtcaattt ccctgactcc tctaagccac tcttccaaag gcataagacc
      721 ctaagcctcc ttttgcttga aaccaaagat atatacacag gatcctattc tcaccaggaa
      781 gggggtccac ccagcaaaga gtgggctgca tctgggattc ccaccaaggt cttcagccat
      841 caacaagagt tgtcttgtcc cctcttgacc catctccccc tcactgaatg cctcaatgtg
      901 accaggggtg atttcagaga gggcagaggg gtaggcagag cctttggatg accagaacaa
      961 ggttccctct gagaattcca aggagttcca tgaagaccac at
//
```

Appendix 3: Restriction Sites within the DNA sequence provided in appendix II

Single cutters

#	Enzyme	Specificity	Sites & flanks	Cut positions (blunt - 5' ext. - 3' ext.)
1	AccI	GT^MK_AC	list	299/301
2	AgeI	A^CCGG_T	list	*229/233
3	AleI	CACNN_^NNGTG	list	484
4	AlwNI	CAG_NNN^CTG	list	291/288
5	ApoI	R^AATT_Y	list	973/977
6	BaeGI	G_KGCM^C	list	125/121
7	BamHI	G^GATC_C	list	760/764
8	BanII	G_RGCY^C	list	254/250
9	BbsI	GAAGACNN^NNNN_	list	822/826
10	BsaAI	YAC_^GTR	list	*384
11	BsaWI	W^CCGG_W	list	229/233
12	BseYI	C^CCAG_C	list	790/794
13	BsmFI	GGGAC(N)$_{10}$^NNNN_	list	842/846
14	BsmI	GAATG_CN^	list	892/890
15	BspQI	GCTCTTCN^NNN_	list	631/634
16	BsrDI	GCAATG_NN^	list	641/639
17	BsrFI	R^CCGG_Y	list	*229/233
18	BstAPI	GCAN_NNN^NTGC	list	525/522
19	EcoNI	CCTNN^N_NNAGG	list	661/662
20	EcoRI	G^AATT_C	list	973/977
21	HaeII	R_GCGC^Y	list	*16/12
22	HhaI	G_CG^C	list	*15/13
23	HinP1I	G^CG_C	list	*13/15
24	HindIII	A^AGCT_T	list	464/468
25	MfeI	C^AATT_G	list	167/171
26	MmeI	TCCRAC(N)$_{18}$_NN^	list	*374/372
27	MseI	T^TA_A	list	596/598
28	NciI	CC^S_GG	list	*370/371
29	NsiI	A_TGCA^T	list	62/58
30	NspI	R_CATG^Y	list	528/524

31	PflFI	GACN▾N▴NGTC	list	670/671
32	PflMI	CCAN▴NNN▾NTGG	list	291/288
33	PmlI	CAC▾▴GTG	list	*384
34	PshAI	GACNN▾▴NNGTC	list	169
35	PvuII	CAG▾▴CTG	list	537
36	SapI	GCTCTTCN▾NNN▴	list	631/634
37	SfaNI	GCATC(N)$_5$▾NNNN▴	list	817/821
38	StuI	AGG▾▴CCT	list	563
39	Tth111I	GACN▾N▴NGTC	list	670/671
40	XcmI	CCANNNN▴N▾NNNNTGG	list	486/485

Two cutters

#	Enzyme	Specificity	Cut positions (blunt - 5' ext. - 3' ext.)
1	AciI	C▾CG▴C	*19/21, *77/79
2	AcuI	CTGAAG(N)$_{14}$▴NN▾	452/450, 816/814
3	BbvCI	CC▾TCA▴GC	544/547, 558/561
4	BsaI	GGTCTCN▾NNNN▴	427/431, 518/522
5	BseRI	GAGGAG(N)$_8$▴NN▾	201/199, 678/676
6	Bsp1286I	G▴DGCH▾C	125/121, 254/250
7	HpaII	C▾CG▴G	*230/232, *369/371
8	Hpy166II	GTN▾▴NAC	300, 787
9	Hpy188I	TC▴N▾GA	916/915, 970/969
10	MlyI	GAGTC(N)$_5$▾▴	661, 679
11	MslI	CAYNN▾▴NNRTG	319, 484
12	MspA1I	CMG▾▴CKG	*19, 537
13	MspI	C▾CG▴G	230/232, 369/371
14	PleI	GAGTCNNNN▾N▴	661/662, 679/680
15	PstI	C▴TGCA▾G	509/505, 521/517
16	SfcI	C▾TRYA▴G	505/509, 517/521