Name_____ **Section**_____ **TA**_____

7.016 Problem Set 2- 2018

Question 1

Melatonin is a hormone that regulates sleep and is used to treat insomnia (inability to fall asleep). It acts by binding to the melatonin receptors (MT_R), which exists in different isoforms: MT_{R1}, encoded by *MT-1 gene* and MT_{R2}, encoded by *MT-2 gene.* The following is the sequence corresponding to base pairs (bp) 90-140 of *MT-1* and *MT-2* genes.

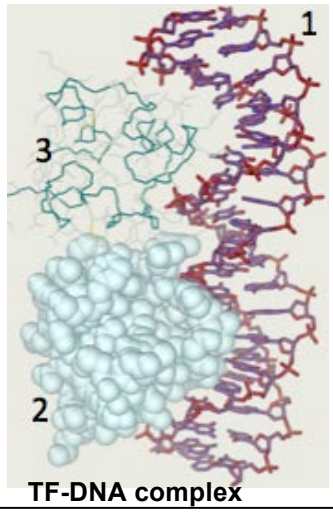
```
                90                                                  140
MT-1: 5'------GATATGCCCCCCCCGGCGCGCGATATGCCCCCCCCGGCGCGCGTGCGTGA-------3'
      3'------CTATACGGGGGGGGCCGCGCGCTATACGGGGGGGGCCGCGCGCACGCACT-------5'

                90                                                  140
MT-2: 5'------GATATGATATATATATATATAGATATGAAAAATTTTTATATAGTGCGTGA-------3'
      3'------CTATACTATATATATATATATCTATACTTTTTAAAAAATTATCACGCACT-------5'
```

a) Which of the above sequences will denature at a lower temperature and **why**?

b) In the sequences above, which end receives the incoming nucleotide: the **5' or the 3' end**?

c) Melatonin binds to MT_R and activates it. The activated MTR triggers a signaling pathway that activates two specific transcription factors (TF) that play a critical role in regulating sleep.



TF-DNA complex

i. On the schematic of the TF-DNA complex to the left, which macromolecule(s) are the likely transcription factors and **why**: **1/ 2/ 3**?

ii. You denature the TF-DNA complex. Which of these molecules may likely renature and **why**: **1/ 2/ 3**?

iii. Name a **region on a gene** to which the TFs can bind to regulate gene expression.

iv. Which of the following amino acids in the TFs can potentially form an ionic bond with the DNA and **why**: **Serine/ Alanine/ Glutamic acid/ Lysine**? ***Note:*** *A chart of amino acids was given on the last page of Problem set 1.*

Name_____ **Section**_____ **TA**_____

Question 1 continued

d) You identify three individuals with the following mutations in the *MT-1* gene. **Explain** how melatonin prescription would affect Individuals 1-3.

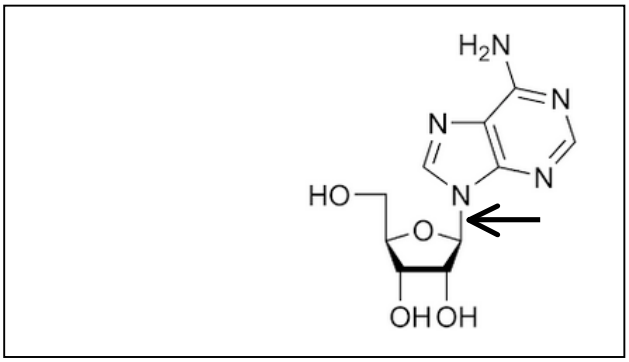
I. **Individual 1:** The Histone proteins bound to the *MT-1* gene in Individual 1 are constitutively (always) acetylated:

II. **Individual 2:** Activating transcription factors (TF) for the *MT-1* gene expression are absent.

III. **Individual 3:** Promoter region of *MT-1* gene does not have methylated bases:

Question 2

Nucleic acids are made of five bases –A, T, G, C and U. The following is the chemical structure of one of the five bases. For parts **(a) – (d)**, on the schematic…



a) Number the carbon atoms of the sugar.

b) Add three phosphates to the 5'C of the nucleoside (base + sugar) so that you now show it as a nucleotide (base + sugar + triphosphate).

c) Put an **" X"** next to the atoms/ groups of the nucleotide that participate in hydrogen bonding with the complementary base.

d) Circle the group that would participate in the formation of a phosphodiester bond if this nucleotide were **added to** the growing end of a nucleic acid chain.

e) Identify the base on the schematic as **purine or pyrimidine**: _____

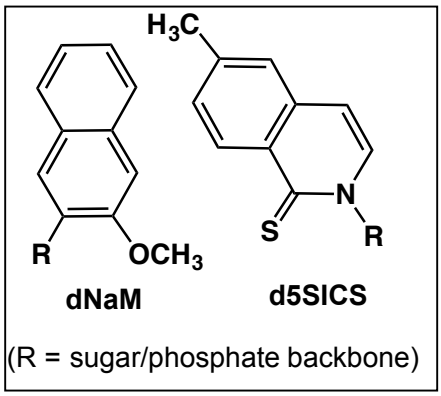
f) Name the **covalent bond** shown by an arrow in the schematic above: _____

g) Which nucleic acid is the nucleotide in the schematic above a building block for and **why**: DNA/ RNA?

Name_____ **Section**_____ **TA**_____

Question 2 continued

Organisms are defined by the information encoded in their genomes. The four bases in DNA (A, T, G, C) can be combined to form 64 different codons or "words". Recently ***Malyshev et al (Nature, 2014)*** have expanded the genetic code in bacteria to include two synthetic nucleotides, which are designed to form a new complementary base pair d5SICS–dNaM as shown below.



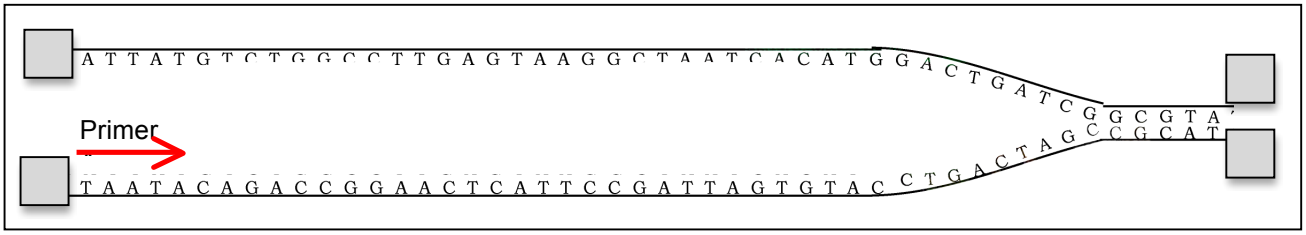
h) How many possible tri-nucleotide codons exist using the bases **A, T, G, C, d5SICS** and **dNaM**?

i) Circle the most likely interaction between the dNaM-d5SICS base pair: **covalent bond/ ionic bond/ hydrophobic interaction/ hydrogen bonding**.

j) Explain how the interaction that you circled in part (i) above differs from the type of interactions between A/T and G/C base pairs

Question 3

The *MT-1* gene is located on chromosome 11. The diagram below depicts a replication fork at one origin of replication site (ori) on Chromosome 11 <u>in a cell</u> undergoing DNA replication. **<u>Note:</u>** *The primer is shown by an arrow i.e. 5'→3'.*



a) On the schematic of the replicating DNA above, label the 5' and 3' ends of the template DNA strands by filling in the shaded boxes and show the direction of movement of the replication fork by an arrow.

b) Label the parental strands as the templates for **leading (continuous)** or **lagging (discontinuous)** strands synthesis.

c) Give the sequence of the five base long primer that is shown as an arrow in the diagram above:
5'_____**3'**

d) Give the sequence of the **<u>next four bases</u>** that would be added to the growing end of the primer:
5'_____**3'**

Name_____ **Section**_____ **TA**_____

Question 3 continued

e) You grow the normal, healthy *MT-1 gene* expressing cells in the presence of nutrients and the following compounds. ***Note:*** *For this question you may assume that these compounds are able to get into the cells.*

- **Plate 1** contains TA-65, a compound that prevents the shortening of chromosomes.
- **Plate 2** contains doxorubicin, a compound that promotes DNA supercoiling.
- **Plate 3** contains an RNAse inhibitor.

Explain how the compounds would affect DNA replication in the following plates.

i. **Plate 1:**

ii. **Plate 2:**

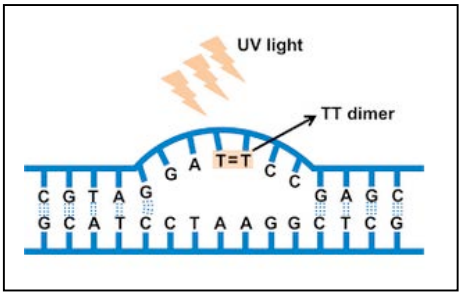
iii. **Plate 3:**

Question 4

a) Prokaryotes such as bacteria only have one origin of replication (ori site) in the bacterial genome. If DNA polymerase enzyme adds nucleotides at the rate of 10,000 base pairs/minute ***in one direction*** and the bacterial cells replicate every 20 minutes, what is the size (in terms of base pairs) of the bacterial genome? **Show your work.** ***Note:*** *In your calculation be sure to account for the fact that replication is bidirectional.*

b) Unlike the prokaryotes, eukaryotic genomes have multiple ori sites. Using the following information, calculate the total number of ori sites on a single human chromosome that is 252 million base pairs long: *DNA polymerase adds nucleotides at the rate of 3,000 base pairs/minute in one direction. In your calculation be sure to account for the fact that replication is bidirectional. The S phase (DNA replication phase) of cell cycle is 5 hours long.*

c) UV radiation from the sun or tanning salons can result in the formation of thymidine dimers (T-T) in the DNA of skin and hair follicle cells. These T-T dimers, if left unrepaired can result in rapid aging of skin, freckles and even melanoma (a form of skin cancer). Which process will repair the T-T dimers: **DNA Proofreading/ mismatch repair/ Nucleotide excision repair**? **Explain** why you selected this process over the others.



Name_____ **Section**_____ **TA**_____

Question 4 continued

d) Does the formation of T-T dimers impact the…

i. **Fidelity of Replication**? If so, how? **Note:** *Your explanations may vary.*

ii. **Process of Transcription**? If so, how? **Note:** *Your explanation may vary.*

Question 5

The *MT-1* gene is comprised of four exons (Exon 1, 2, 3, and 4) that are each 1.5 kilo base (kb) long and three introns (Introns 1, 2, and 3) that are each 3kb long.

a) Based on the information provided, draw the *MT-1* gene in the box below. Label its 5' and 3' ends, draw and label the promoter, the transcription start and stop sites, exons 1-4 (as blank boxes) and introns 1-3 (as shaded boxes). ***Note:*** *The arrow in the schematic shows the direction of transcription.*

b) Draw the **nascent (pre-spliced/ newly synthesized) mRNA** transcribed from the *MT-1* gene. Label its 5' and 3' ends, exons 1-4, introns 1-3 and any modifications to the mRNA. Label each splice donor site with a "D" and each splice acceptor site with an "A".

c) Assuming all introns are spliced out, give **all** the possible **mature (spliced) mRNA transcripts** resulting from the *MT-1* gene.

I. Label the 5' and 3' ends of each mature transcript.

II. Specify the exons that make the transcripts and give the size (in terms of kb) for each mature transcript.

III. List any modification(s) to the 5' and the 3' ends of the mature mRNA transcripts and **briefly describe** their significance.

Name_____ **Section**_____ **TA**_____

Question 6

The following is the DNA sequence for the transcription initiation region of the *MT-1* gene. ***Note:*** *Part of the* ***promoter region*** *is boxed and the dashes represent bases that are not shown. Transcription* ***begins at and includes*** *the bold and underlined* ***T/A*** *base pair.*

```
5′--TGGACTGCTATGATAGCAGTTCTGCTGAGACGATGGCCATACGGCCATGGTTCATAAAGT----3′ TOP
3′--ACCTGACGATACTATCGTCAAGACGACTCTGCTACCGGTATGCCGGTACCAAGTATTTCA----5′BOTTOM
```

a) Identify the template strand for transcription: **Top / Bottom**? _____

b) Write the **first 9 nucleotides** of the newly transcribed *MT-1* mRNA: **5'**_____**3'**

c) Using the codon chart on the last page of this problem set, write the **first 2 amino acids** of the newly synthesized MT_{R1} receptor: **N-**_____**-C**

d) Write the sequence of the **anticodon corresponding to the 2nd amino acid** of newly translated MT_{R1} protein: **5'**-_____-3'

e) The last (C-terminal) 5 amino acids (296-300), which are **critical** for the normal function of MT_{R1} receptor are: **N –Val^{296}-Ser^{297}-Asn^{298}-Ser^{299}-Met^{300}- C**

The DNA sequence encoding the C- terminus of the wild- type and mutant forms of the MT_{R1} protein is included within the sequence below. The **point mutations** in mutants 1 and 2 are bold & shaded and the stop codon (5'TAG3') is underlined. ***Note:*** *The codon table is provided on the last page of the problem set.*

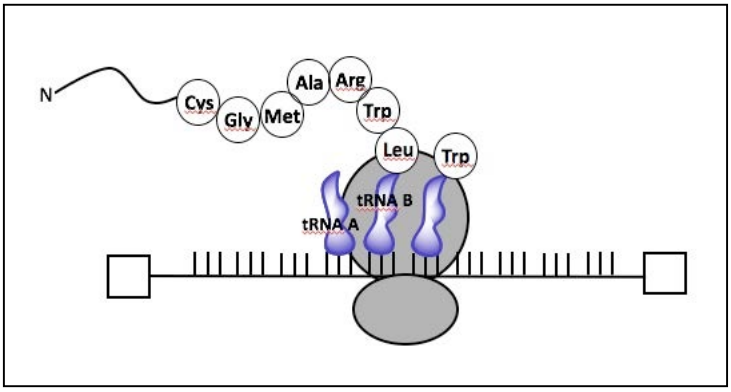
```
Wild- type: 5′-TCGTATCGAATTCCATGTAGC-3′
            3′-AGCATAGCTTAAGGTACATCG-5′

Mutant 1:   5′-TCGTATAGAATTCCATGTAGC-3′
            3′-AGCATATCTTAAGGTACATCG-5′

Mutant 2:   5′-TCGTATCGAACTCCATGTAGC-3′
            3′-AGCATAGCTTGAGGTACATCG-5′
```

Compared to the wild- type, which mutant allele will most likely not express the *MT-1* gene and **why**?

f) Below is a drawing of a ribosome actively translating *MT-1* mRNA. The horizontal line represents the mRNA, each vertical line represents a base (A, U, G, C) and two tRNA molecules (tRNA A and tRNA B) are labeled and drawn in purple on the posted version of the problem set.



i. Draw an arrow to indicate the movement of mRNA relative to the ribosome during translation.

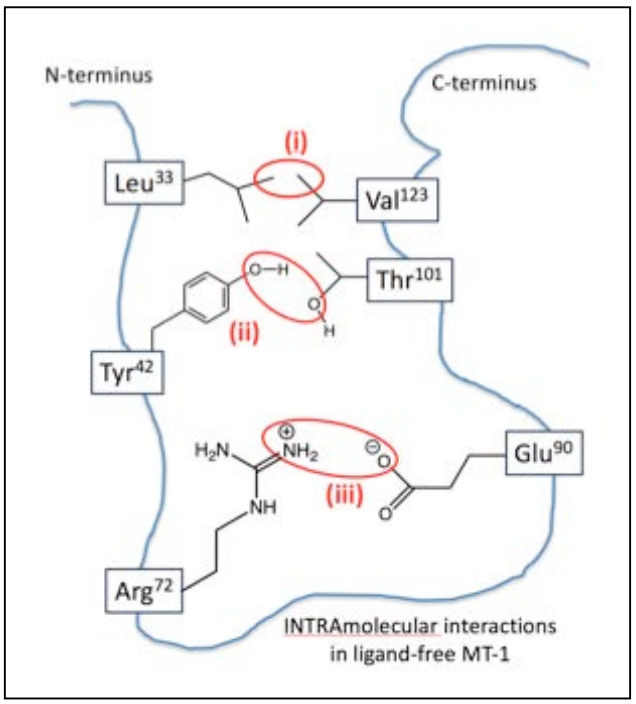
ii. Label the 5' and the 3' ends of the mRNA.

iii. Name the amino acid that was originally bound to **tRNA A**: _____

Name_____ **Section**_____ **TA**_____

Question 7

The diagram below shows the intra-molecular interaction between the side-chains of amino acids at positions (i)-(iii) in the MT_{R1} protein. Each of these interactions is critical for the correct 3D folding of MT_{R1} protein, which allows it to bind to melatonin.



a) What is the strongest non-covalent interaction at positions (i)–(iii): **Ionic interaction/ hydrophobic interaction/ hydrogen bonding?**

i. _____

ii. _____

iii. _____

b) You analyze both alleles of the *MT-1* gene in four individuals (1-4). Each of these individuals has a specific mutation in one or both alleles of the *MT-1* gene as specified in the table below. ***Note:*** *Please refer to the information at the beginning of Question 1 of this problem set.*

Which of the above individuals are most likely to have insomnia and **why**?

Individuals	Mutations
1	5'CUG3' (Leu) -> 5'CUA3' at (i)
2	5'UAU3' (Tyr) -> 5'UAA3' at (ii)
3	5'CGA3' (Arg) -> 5'CAA3' at (iii)
4	5'UAU3' (Tyr) -> 5'UAG3' at (ii)

CODON CHART

	U	C	A	G	
U	UUU **phe**	UCU **ser**	UAU **tyr**	UGU **cys**	U
	UUC **phe**	UCC **ser**	UAC **tyr**	UGC **cys**	C
	UUA **leu**	UCA **ser**	**UAA STOP**	**UGA STOP**	A
	UUG **leu**	UCG **ser**	**UAG STOP**	UGG **trp**	G
C	CUU **leu**	CCU **pro**	CAU **his**	CGU **arg**	U
	CUC **leu**	CCC **pro**	CAC **his**	CGC **arg**	C
	CUA **leu**	CCA **pro**	CAA **gln**	CGA **arg**	A
	CUG **leu**	CCG **pro**	CAG **gln**	CGG **arg**	G
A	AUU **ile**	ACU **thr**	AAU **asn**	AGU **ser**	U
	AUC **ile**	ACC **thr**	AAC **asn**	AGC **ser**	C
	AUA **ile**	ACA **thr**	AAA **lys**	AGA **arg**	A
	AUG met	ACG **thr**	AAG **lys**	AGG **arg**	G
G	GUU **val**	GCU **ala**	GAU **asp**	GGU **gly**	U
	GUC **val**	GCC **ala**	GAC **asp**	GGC **gly**	C
	GUA **val**	GCA **ala**	GAA **glu**	GGA **gly**	A
	GUG **val**	GCG **ala**	GAG **glu**	GGG **gly**	G

7.016 Introductory Biology
Fall 2018