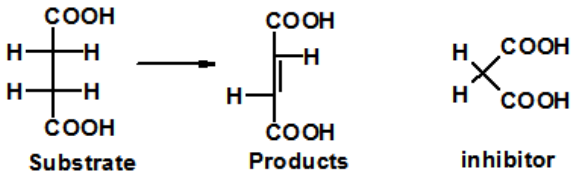
This exam contains 100 points on 6 pages. Allot 1 min/2 pts. Please use the space provide, or the back of the previous page. In questions with choices, all answers will be graded and you will receive the highest grade. Be sure to make it clear which choice you are answering.

Section A: Multiple Choice (8 pts)

1. In DNA sequencing which of the following nucleotides are "colored"?
   a) ddNMPs (dideoxynucleotide monophosphates)
   b) dNTPs (deoxynucleotide triphosphates)
   c) NTPs (nucleotide triphosphates)
   d) dNMP (deoxynucleotide monophosphates)
   e) None of these are correct. (ddNTPs are colored)

2. In DNA sequencing, the largest colored DNA fragment is detected first, followed by the other fragments from largest to smallest.
   a) True
   b) False, the smallest ones are detected first, giving the order of addition of the bases to the primer.

3. The substrate and product of an enzymatic reaction are shown on the left. An inhibitor is show on the right. This is a
   a) allosteric inhibitor
   b) competitive inhibitor



4. Which of the following activities **cannot** be accomplished by the HIV reverse transcriptase enzyme
   a) RNA polymerase, using RNA as the template.
   b) DNA polymerase, using DNA as the template.
   c) DNA polymerase, using RNA as the template.
   d) Ribonuclease activity, digestion of RNA.

Section B: Short Answer

1. (4 pts) Briefly describe the structure of a virus and explain why they are not considered to be living organisms.

Viruses contain some kind of genetic material, RNA or DNA.
A protein coat that surrounds the genetic material
Some viruses have an external lipid coat.
They are non-living because they cannot replicate without a host.

2. (9 pts) Provide an overview of the HIV life cycle, in your answer you want to address:
   i) Name the key enzymes that are involved in the life cycle
   ii) Briefly describe the function of the key enzymes.
   iii) Explain why they are particularly good drug targets.

Virus attached to cell and crosses membrane
Its viral RNA is converted to ds DNA by reverse transcriptase.
The dsDNA is inserted into the host chromosome by integrase
Viral mRNA and vRNA are produced from the DNA
The immature viral proteins are cut by HIV protease to generate mature proteins that assemble into the virus.

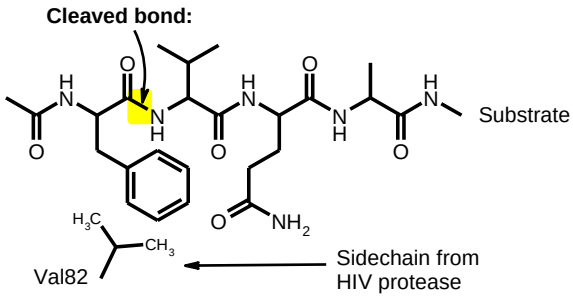
These are good drug targets because they are i) unique to the virus – not found in the host, ii) essential for viral replication.

3. (2 pts) The active site of HIV protease contains two aspartic acid residues and a valine at position 82. What is the function, or role, of the two aspartic acid residues in the enzymatic mechanism?

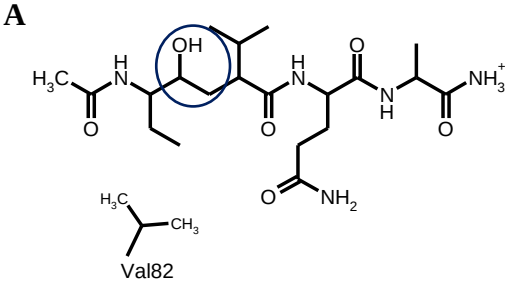
These are responsible for cleaving the peptide bond (valine 82 is responsible for substrate specificity).

4. (12 pts) Two inhibitors (drugs) (A and B) of HIV protease are shown below. The normal substrate for the enzyme is shown on the right. All diagrams show the sidechain of Val 82, which is part of the HIV protease enzyme.

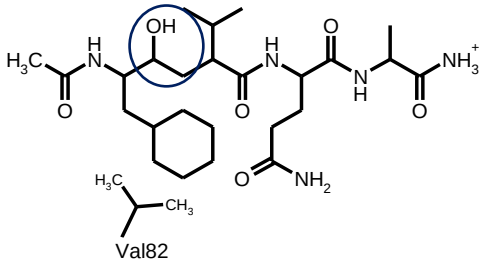
i) Explain why these are competitive inhibitors of the enzyme (6 pts).



A



B



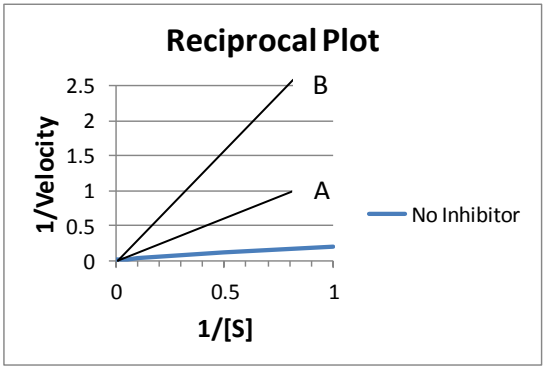
Both drugs look like the substrate.

They lack a peptide bond, as indicated by the circled region. Therefore they cannot be cleaved.

ii) Draw on the double reciprocal plot the line you would expect to observe for inhibitor A and inhibitor B. The plot already contains the line that would be obtained in the absence of inhibitor. **Justify** your answer with a brief discussion of interaction between the enzyme and the inhibitors. (6 pts).

Inhibitor B should bind better because it makes more extensive van der Waals contacts with valine 82. It has a larger non-polar group, so there would be a larger hydrophobic effect driving binding as well. On a double reciprocal plot this will give a steeper line.

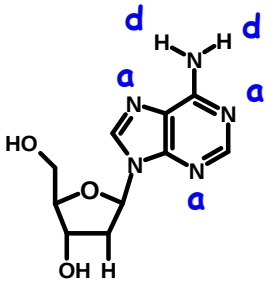
Inhibitor A has a smaller group, replacing the cyclohexane ring, leading to poorer van der Waals contacts, and a reduction in the hydrophobic effect. On a double reciprocal plot this will give a line with a smaller slope.



5. (6 pts)

i) Is the base a purine or a pyrimidine (circle correct answer)?

ii) Label **all** hydrogen bond donors (d) and acceptors (a) on the base.



6.(5 pts) Please do one of the following two choices.

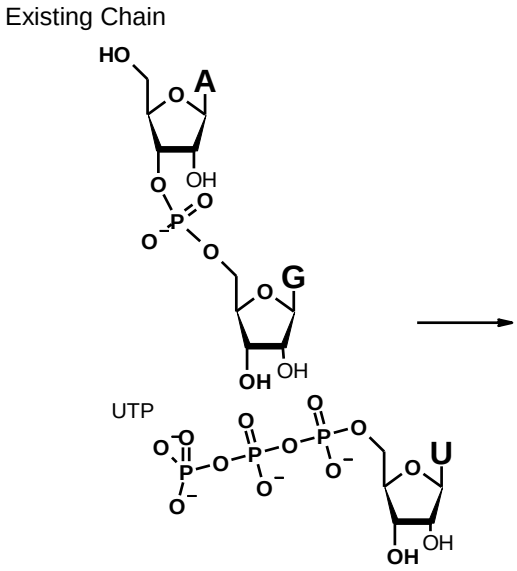
Choice A: What is the principal difference between the *chemical* structure of RNA and DNA? How does this difference affect the stability of RNA versus that of DNA.

Choice B: Briefly describe the major and minor grooves in DNA. Where do proteins typically bind, in the major or minor groove? Why?

Choice A: The sugar in RNA (ribose) has an –OH group on the 2' carbon. This group can attack the phosphate, breaking the phosphodiester bond between the bases, making RNA less stable than DNA. DNA has a hydrogen at this position.

Choice B: The major and minor grooves refer to a difference in phosphate spacing in the DNA double helix. The major groove has a wider spacing. Proteins typically bind in the major groove because it is larger.

7. (14 pts) The following diagram shows the addition of a UTP to an existing chain.



i) identify the phosphodiester bond in the existing chain.

It joins the two riboses.

ii) Indicate the 5' and 3' end of the existing chain.

5' end is at the top

iii) What type of enzyme is likely catalyzing this reaction?

A polymerase (RNA polymerase to be exact).

iv) How would the enzyme know to add a U to existing chain?

It has to use a template, there would be an A opposite the U.

v) Briefly indicate how the chemical bond forms between the existing strand and the new base would occur. You need not draw a chemical structure.

The 3' OH attacks the 1st phosphate, releasing pyrophosphate.

vi) Write the sequence of the product of the reaction.

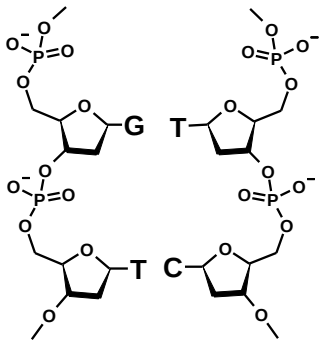
The sequence is written the 5' to 3' direction, AGU.

8. (4 pts) What are the two general classes of HIV reverse transcriptase inhibitors?

Both competitive and allosteric inhibitors.

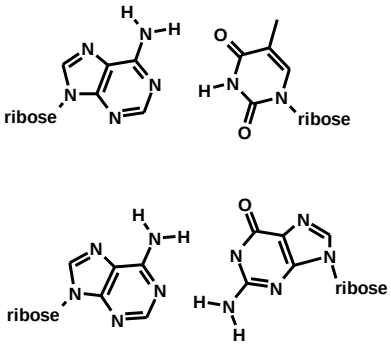
9. (4 pts) Please do choice A or choice B.

Choice A: The double stranded structure of DNA is shown on the right. What are the errors in the structure, if any?



The chains are parallel – they should be anti-parallel.
G pairs with C, T with A.

Choice B: Two possible basepairs are shown on the right, which of these two is most likely to be found in DNA, why?



The upper one, because it is a purine-pyrimidine match, which is the right size to fit into the helix.

10. (6 pts) Briefly explain the mechanism by which drug resistant HIV viruses develop.

The polymerase involved in copying the viral RNA to DNA lacks the 3′ to 5′ exonuclease activity that corrects errors. Therefore changes occur in the viral genetic information (genome), giving rise to altered proteins that the drugs bind poorly to.

11. (18 pts) You wish to amplify a gene from a drug resistant virus with the eventual goal of producing the protein in bacteria for further study. The nucleotide sequence of the region that contains this gene is shown below, with the sequence of the gene in uppercase letters. The amino acid sequence of the protein is also given.

```
          MetArgTyrValAlaValTyrArgArgSerIleLeuProGlyTyrSerAlaSerAspSerSTP
ggcctaggtATGCGGTACGTAGCGGTATATAGGCGATCGATTCTGCCCGGGTATAGCGCTAGCGATAGCTAAcgcgttagtagta
ccggatccaTACGCCATGCATCGCCATATATCCGCTAGCTAAGACGGGCCCATATCGCGATCGCTATCGATTgcgcaatcatcat
```

The plasmid that your plan to insert the PCR product into has a single BamH1 site (G^GATCC), so your PCR product should look like:

```
          MetArgTyrValAlaValTyrArgArgSerIleLeuProGlyTyrSerAlaSerAspSerSTP
GGATCCATGCGGTACGTAGCGGTATATAGGCGATCGATTCTGCCCGGGTATAGCGCTAGCGATAGCTAAGGATCC
CCTAGGTACGCCATGCATCGCCATATATCCGCTAGCTAAGACGGGCCCATATCGCGATCGCTATCGATTCCTAGG
```

i) Give 12 bases of the sequence of the right and left PCR primers that you would be need to produce this product (6 pts)

The left primer is just the first 12 bases on the top strand of the PCR product, in the 5′ to 3′ direction: GGATCCATGCGG

The right primer is just the first 12 bases of the bottom strand of the PCR product, in the the 5′ to 3′ direction: GGATCCTTAGCT

ii) Briefly describe the steps involved in a PCR reaction and why amplification of the target region occurs (6 pts).

The steps are:

i) denaturation at $T>T_M$
ii) annealing of primers, at $T<T_M$
iii) polymerization

Amplification occurs because the primers "point towards" each other and copy the DNA between the primers (and the primers themselves) each cycle, doubling the amount of PCR product per cycle.

iii) How will digesting your PCR product with BamH1, and the plasmid with BamH1, allow you to insert the PCR product into the plasmid? A well-labeled diagram is an acceptable answer (6 pts)

Digestion would produce the same sticky ends on both the PCR product and the plasmid.
The sticky ends would anneal by complementary Watson-Crick base pairing
DNA ligase would be used to restore the phosphodiester bond.

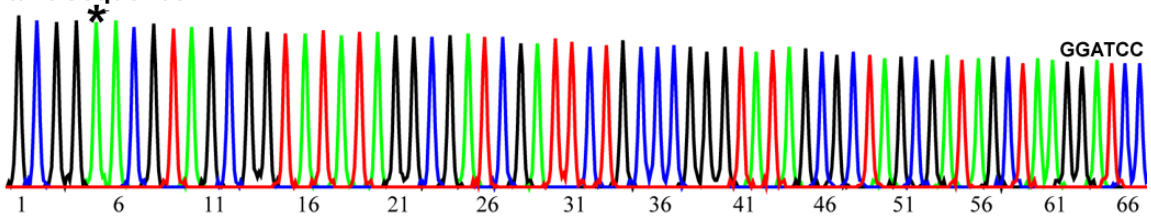
12. (8 pts) You sequence the wild-type and mutant versions of this gene to determine what mutation is responsible for drug resistance. The location of the mutation is indicated with a "*" and the peak associated with that change is green. The first forty bases of sequence from the wild-type are also given. Identify the altered amino acid in the mutant gene. Please indicate how you obtained your answer.

(The sequence of the PCR product and the protein sequence of the wild-type protein are repeated here for your convenience.)

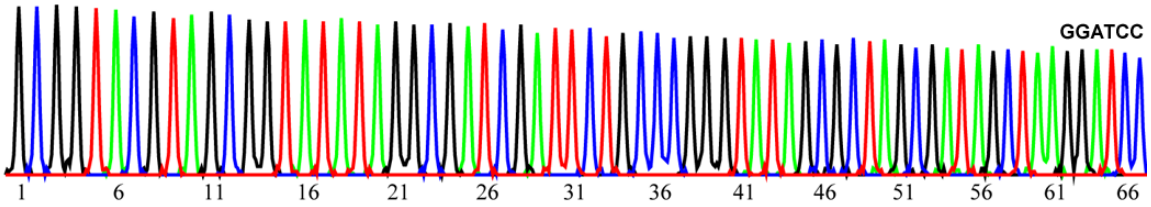
```
                MetArgTyrValAlaValTyrArgArgSerIleLeuProGlyTyrSerAlaSerAspSerSTP
GGATCCATGCGGTACGTAGCGGTATATAGGCGATCGATTCTGCCCGGGTATAGCGCTAGCGATAGCTAAGGATCC
CCTAGGTACGCCATGCATCGCCATATATCCGCTAGCTAAGACGGGCCCATATCGCGATCGCTATCGATTCCTAGG
```

G=black, C=blue, A=green, T=red.

Mutant Sequence



Wild-type Sequence



The first forty bases from data for the sequence of the wild-type:

```
1         10        20        30        40
GCGGTACGTAGCGGTATATAGGCGATCGATTCTGCCCGGG
GCGGAACG
```

The mutation is at position 5 – a T has been changed to an A.

The first G of the sequence is the last G of the first codon, so the correct reading frame for the mutant sequence is:

G CGG AAC G
MET ARG Asn VAL

The mutation changes a Tyr to an Asn.

14. Bonus (4 pts) What was the sequence of the primer that was used for sequencing?

It was GGATCCAT, since the G was the first base added.

Codon Table:

5' Base	Middle Base				3'
	T	C	A	G	
T	Phe	Ser	Tyr	Cys	T
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	**Term**	**Term**	A
	Leu	Ser	**Term**	Trp	G
C	Leu	Pro	His	Arg	T
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	T
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	T
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G