

# Protein Modeling Division C - Master Key

Northview Science Olympiad Invitational

*December 7<sup>th</sup>, 2019*



*Authored by:*

*<sup>1</sup>Erich Beregszaszy, University of Michigan-Dearborn, B.S. Chemistry and Physics 2014*

*<sup>2</sup>Sophia Velasco, University of Florida, B.S. Biochemistry 2023*

## **Instructions and Clarifications:**

- You have **50** minutes to finish this exam and the computer exploration of protein structure. The **Exam** portion is composed of Section **I** and **II** and is worth **30%** of the overall score. The **Computer Exploration** portion is composed of Section **III** and is worth **30%** of the overall score.
- Each **participant** may bring **one** 8.5" x 11" sheet of paper that contains information about this event without any annotations or labels affixed along with writing utensils for each participant.
- You **may** split the exam but you are responsible for placing them in the correct order afterwards.
- Write your team number in **every** page of the answer sheet.
- There are **5** tiebreakers in this exam. They are marked as **TB#**.
- Anything written on the exam will **not** be graded. Only the **answer sheet** will be graded.
- If you have any comments or questions about this exam, feel free to contact us at the emails below. **Happy testing!**

<sup>1</sup>*eberegszaszy@troy.k12.mi.us*

<sup>2</sup>*velasco.scienceolympiad@gmail.com*

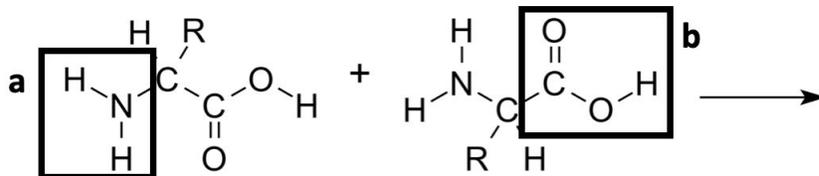
# I. General Biochemistry and CRISPR-Cas Systems<sup>1</sup>

**Multiple Choice:** Choose the most appropriate answer option for the following questions. Each question is worth **two** points. (18)

1. Proteins are polymers of...
  - a. Amino acids
  - b. Peptides
  - c. Saccharides
  - d. Fatty Acids
  - e. Nucleic Acids
  
2. DNA and RNA are polymers of...
  - a. Amino acids
  - b. Peptides
  - c. Saccharides
  - d. Fatty Acids
  - e. Nucleic Acids
  
3. Which of the following would you **not** see in globular proteins?
  - a. Alpha helices
  - b. Beta Sheets
  - c. Double helices
  - d. Beta turns
  - e. Random coils
  
4. What does CRISPR stand for?
  - a. Clustered Regularly Interspersed Short Polymorphic Repeats
  - b. Clustered Regularly Interspaced Small Palindromic Repeats
  - c. Clustered Regularly Interspaced Short Palindromic Repeats
  - d. Clustered Regularly Integral Short Palindromic Repeats
  - e. Clustered Reversibly Interspersed Short Palindromic Repeats
  
5. The CRISPR system functions as a component of the:
  - a. viral immune system
  - b. viral reproductive system
  - c. human immune system
  - d. bacterial reproductive system
  - e. bacterial immune system

6. Which of the following relationships in biology is/are always true?
- a. DNA is transcribed to RNA and RNA is translated into proteins
  - b. RNA is transcribed to DNA and DNA is translated into proteins
  - c. DNA and RNA can both be transcribed to proteins
  - d. A. and B.
  - e. B. and C.
7. Which of the following is **not** a type of non-covalent interaction that would stabilize protein structure?
- a. Hydrogen bonds
  - b. Hydrophobic interactions
  - c. Salt Bridges
  - d. Van Der Waals attractions
  - e. Pi-stacking interactions
8. Which of the following best describes palindromic DNA sequences?
- a. The nucleic acid order on a single strand of DNA repeats in a given pattern
  - b. The nucleic acid order on a single strand of DNA has the same pattern forwards and backwards
  - c. The nucleic acid order on both strands has the same pattern forwards and backwards
  - d. The nucleic acid order on one strand appears in the reverse on the other strand in the same location
  - e. None of the above
9. Which of the following best describes the idea of denaturing of a protein structure?
- a. Going from an unfolded state to a folded state
  - b. Going from a folded state to an unfolded state
  - c. Tightening of alpha helices and changing of the tertiary structure
  - d. Relaxing of the beta sheets and changing of the tertiary structure
  - e. More than one of the above.

10. (5 pts) Complete the reaction below for the joining of two amino acids. Label the peptide bond formed. Name the functional groups outlined with boxes.

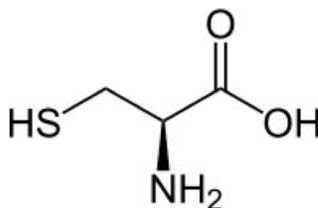


A. Amine (primary)

B. Carboxylic acid

One point per box (2 points max), one point per correct naming of functional group (2 points max), one point for combining the peptide bond correctly. (5 points total)

Questions 11-17 refer to the amino acid below. Question 18 is a tiebreaker and extra credit. Providing the wrong answer to question 18 will **not** impact your exam score. Each question is worth **one** point unless stated otherwise. (10)



11. The amino acid shown above is \_\_\_\_\_ cysteine

12. \_\_\_\_\_ (true/false) The side chain of this amino acid is nonpolar. false

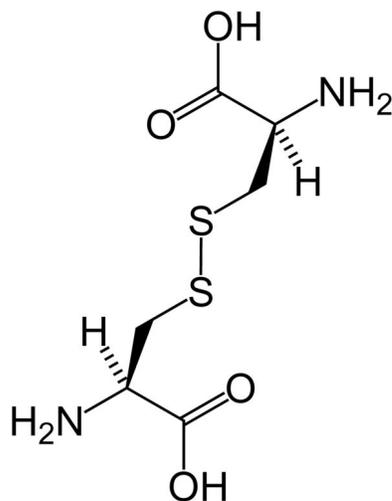
13. \_\_\_\_\_ (true/false) The side chain of this amino acid is acidic. true

14. \_\_\_\_\_ (true/false) The SH region is capable of hydrogen bonding between molecules. false

15. \_\_\_\_\_ (true/false) This amino acid contains a secondary amine. false

16. The SH functional group is called a \_\_\_\_\_. thiol

17. The molecule above can form disulfide bonds through a redox process. Please create this reaction in the space provided in the answer sheet. **Label** the disulfide bond explicitly. (4)



One point for each correctly drawn cysteine. Two points for correct peptide bond.

18. (TB#1/Extra Credit) Fun fact – how is this amino acid important to sheep? (3)  
Allows them to make wool.

## II. DNA, RNA, and CRISPR-Cas Systems<sup>2</sup>

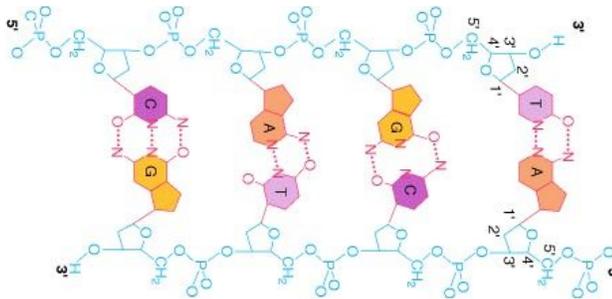
**Multiple Choice:** Choose the most appropriate answer option for the following questions. Each question is worth **two** points. **(26)**

- cas3*, *cas9* and *cas10* are the signature genes for what CRISPR type?
  - I**
  - II
  - III
  - B and C
  - All of the above
- What CRISPR-Cas system has primary processing that occurs in the context of Cas9, which is followed by the trimming of the 5'-end repeat and spacer sequences of the int-crRNA to create mat-crRNAs?
  - I
  - II**
  - III
  - B and C
  - All of the above
- According to genetic evidence, which CRISPR-Cas system (sub)type cleaves RNA molecules?
  - I
  - II
  - III-A
  - III-B**
  - More than one of the above
- According to genetic evidence, which CRISPR-Cas system (sub)type cleaves DNA sequences?
  - I
  - II
  - III-A
  - III-B**
  - More than one of the above

5. The PAM sequence of the Type I E system of *Escherichia coli* is composed of a three-nucleotide motif that is located immediately upstream of the protospacer and then recognized by the Cascade complex. Determine the three-nucleotide motif described. **TB#2**
- AUG
  - AAU
  - AUU
  - AAG**
  - All of the above are possible motifs.
6. In which step of the CRISPR-Cas immune system do crRNAs guide Cas nucleases for specific cleavage of homologous sequences?
- Adaptation
  - crRNA biogenesis
  - Targeting**
  - More than one of the above
  - None of the above
7. Which of the following descriptions below correctly describes PAM?
- It is flanked by a system-specific, highly conserved *CRISPR motif*.
  - PAMs are typically 10-13 nt, poorly conserved sequence motifs.
  - PAMs do not immediately flank one side of the protospacer.
  - PAMs are located within 1-4 nt of one extremity relative to the protospacer.
  - More than one of the above**
8. Which of the following proteins that are widely distributed across all the CRISPR-Cas systems catalyze the capture and integration of new spacers into CRISPR loci?
- Cas 1 and 2**
  - Cas 2 and 3
  - Cas 9 and 10
  - Cas 4, and 5
  - None of the above
9. In a DNA molecule, what structure forms from the base and the deoxyribose if the phosphate group is not present?
- Nucleotide
  - Nucleoside**
  - Hydroxide
  - Ribonucleic acid
  - None of the above

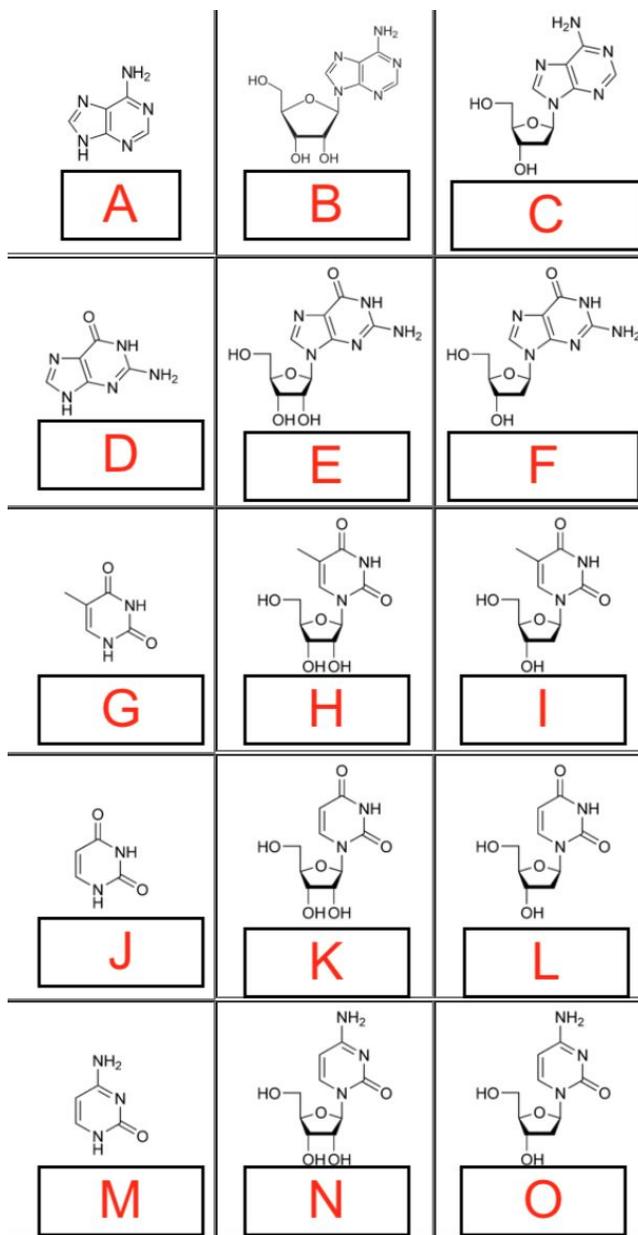
10. Which of the following statements about DNA nucleotides is correct?
- The amount of T equals the amount of C
  - The amount of A equals the amount of G.
  - The amount of T equals the amount of U.
  - The amount of A equals the amount of U.
  - The amount of T equals the amount of A.
11. Which of the following statements about hydrogen bonds in DNA molecules is correct?
- Each hydrogen atom in the  $\text{NH}_2$  group is  $\delta^-$  because the nitrogen atom is not attracted to the electrons of the N-H bond.
  - The hydrogen bond forms between one H and N.
  - Hydrogen bonds are really strong and prominent in DNA molecules, and its strength is 64% of the strength of a covalent bond.
  - The hydrogen bond is stronger if the atoms in the hydrogen bond are “pointing at each other” in ideal orientations.
  - More than one of the above

Use the image below to answer the following multiple choice questions. Each question is worth two points.



12. The hydrogen bonds in the image are indicated how?
- Hexagons
  - Pentagons
  - Solid lines
  - Dotted lines
  - None of the above
13. Describe the orientation of the two backbones.
- Antiparallel
  - Parallel
  - Sometimes A, Sometimes B
  - There is more than one backbone.
  - None of the above

**Identification:** Identify and state the **full name** of the structures below. Abbreviated names will **not** be awarded credit. Each structure is worth **two points**. **TB#3 (30)**



A. Adenine

B. Adenosine

C. Deoxyadenosine

D. Guanine

E. Guanosine

F. Deoxyguanosine

G. Thymine

H. 5-Methyluridine

I. Deoxythymidine

J. Uracil

K. Uridine

L. Deoxyuridine

M. Cytosine

N. Cytidine

O. Deoxycytidine

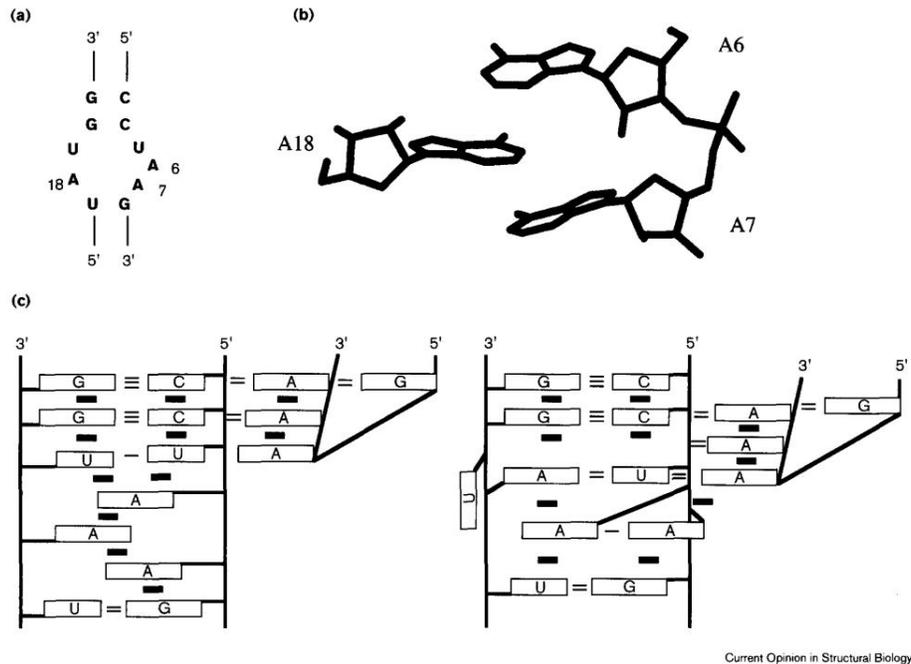
**Short Answer:** The point values for each question are addressed in the parentheses. Complete sentences are **not** required in this question, and coherent, specific and direct answers are encouraged. **(68)**

1. There are several types of CRISPR-Cas systems that each have unique characteristics that distinguish them from one another. Answer the following questions about these systems. **(7)**
  - a. Which CRISPR-Cas type require minimal Cas machinery for immunity? (1) **II**
  - b. Which CRISPR-Cas type uses Cas9, a single large multidomain protein, to cleavage DNA targets? (1) **II**
  - c. Which CRISPR-Cas Type III subtype cleaves RNA targets *in vitro*? (1) **Type III-B (do not accept Type III)**
  - d. Which CRISPR-Cas system further trims crRNAs at the 3' end by an unknown nuclease? (1) **III**
  - e. Which **two** CRISPR-Cas systems rely on the Cas6 endoribonuclease family to cleave repeat sequences of the crRNA precursor? There is no partial credit for this question. (2) **I and III**
  
2. CRISPR-Cas systems are currently utilized in genetic engineering in different domains. Answer the following questions about using CRISPR-Cas systems in genetic engineering. **(11)**
  - a. Do type I and III CRISPR-CAS systems provide a DNA- or RNA-guided nuclease activity? (1) **RNA-guided**
  - b. Support your answer in 2(a) by stating the structure that allows the nuclease activity. There is no partial credit for this question. (3) **crRNA-Cas ribonucleoprotein complex**
  - c. What is the name of the “enabling RNA” in Type II CRISPR-Cas systems? (2) **tracrRNA**
  - d. What is the name of the structure that loads the crRNA guide in Type II CRISPR-CAS systems? (2) **RNaseIII (provide 1 point for RNase only; do not provide credit for RNase I or RNase II)**
  - e. The structures described in 2(c) and 2(d) can be bypassed by using a chimeric nucleic acid guide. What is the name of this structure? (*Hint: This guide is formed by fusion the structures in 2(c) and 2(d).*) (3) **sgRNA (single-guide RNA)**

3. Imagine that you are a scientist utilizing a CRISPR-Cas9 system for genome engineering. The questions below will take you step-by-step in your experiment, and test your application skills regarding the complex. Answer the questions directed at each step of your experiment. **TB#4 (50)**
- a. First, you have to identify and select your target genomic site. There are two primary rules for identifying a target site for the SpCas9 system. **(8)**
    - i. What PAM sequence is required for SpCas9 targeting? (2) **NGG**
    - ii. Describe the **guide** sequence (length, location relative to PAM, and which end) that you will pick for your experiment. (6) **20 bp in length (2) upstream of the PAM (2) to its 5' end (2)**
  - b. Next, it's time to design oligos! **(10)**
    - i. What is the purpose of oligos? (4) **for inserting guide sequences (2) into backbone vectors (2)**
    - ii. What is the cloning vector typically used for wild type SpCas9? (2) **pX330**
    - iii. What is a common mistake that happens when designing oligos? (4) **including the "NGG" PAM sequence (2) in the guide sequence (2)**
  - c. Let's move on to screening multiple guides. The following questions below apply to most CRISPR-Cas 9 applications. **(8)**
    - i. What is the least number of guide sequences screened within the target region in order to find the most efficient ones? (2) **3**
    - ii. The CRISPR-Cas9 system is a very efficient system. Why should we even bother screening multiple guides? Support your answer. (4) **the actual cleavage efficiency could be affected guide-specific or cell-type-specific factors (2), including but not limited to the following: guide sequence, accessibility of local chromatin, activity of endogenous DNA repair pathways (provide 2 points for one of these factors or another answer similar to it that supports the competitor's answers)**
    - iii. **True or False:** A guide sequence that is verified in one cell type is not guaranteed to work to the same efficiency in another cell type or condition. (1) **True**
  - d. After more steps, it's time to analyze the SURVEYOR assays results! **(10)**
    - i. What formulae are used to calculate the genome cleavage efficiency of a tested target? (10) (*Hint: In one of the formulae, "a" and "b" represents relative quantities of cleaved bands, and "c" represents the relative quantity of the non-cut full-length PCR product.*)  
**Indel percentage (%) =  $(1 - \sqrt{1 - x}) \times 100$ , where  $x = (a + b)/(a + b + c)$**   
**5 points per correct equation. There is no partial credit for this question.**
  - e. Lastly, let's design and make a repair template for your HR experiment. **(13)**
    - i. What pathway should you use to introduce a precise genome modification into the genome? (2) **HDR pathway**

- ii. How can you maximize the efficiency of HR? (8) the cleavage site (2) of the guide (2) should be as close as possible (2) to the junction of the homology arm (2)
  - iii. What design can be used to minimize off-target cleavage? (3) double nickase design
- f. It's time to consider some choices of HR template. (8)
- i. What is the best HR template for transfection in order to introduce a single-point mutation? (2) single-stranded DNA (ssDNA) oligos
  - ii. What can be used to introduce a larger genomic modification? (2) plasmid DNA vector
  - iii. A "protospacer + PAM" sequence within the HR template can be degraded by Cas9. How can you use a silent mutation to prevent this degradation from even happening? (4) possible answers: mutate the PAM "NGG" (2) within the HR template (2) to "NGT" or "NGC" OR mutations in the PAM and in the spacer itself (4)
- g. Congratulations on finishing your CRISPR-Cas9 experiment! For 3 points extra credit, draw an x below (2fiii) in your answer sheet. Make it big! x

**Diagram-based Questions:** Answer the following questions based on the given image. The point values for each question are addressed in the parentheses. *Image credit: Graeme L Conn and David E Draper. (32)*



1. Portions of a macromolecule is shown in the given image. Identify the macromolecule. (2) **RNA**
2. The structure labeled **Part (a)** in the given image is in what structure level? (2) **secondary structure**
3. How many conserved nucleotides are present in **Part (a)** of the given image? (2) **11**
4. What motif is shown in **Part (b)** of the given image? (4) **adenosine (2) base-zipper motif (2)**
5. The horizontal lines in **Part (c)** of the given image represents what type of bond? (2) **Hydrogen bonds**
6. The solid rectangles in **Part (c)** of the given image represents what interactions? (2) **stacking interactions**
7. **Part (c)** of the given image shows what conversion? (6) **conversion of the free tetraloop receptor (3) to the bound form (3)**
8. What motif is present in the left part in **Part (c)** of the given image? (4) **adenosine (2) zipper motif (2)**
9. Tertiary folding events in large versions of this macromolecule can be examined at nucleotide resolution through the use of short bursts of synchrotron radiation. What type of radicals is produced in this radiation type? **TB#5** (4) **hydroxyl (OH) radicals**
10. The stability of the macromolecule in the given image is very sensitive to one specific ion. What is this ion? (4) **magnesium**

### III. Computer Exploration<sup>2</sup>

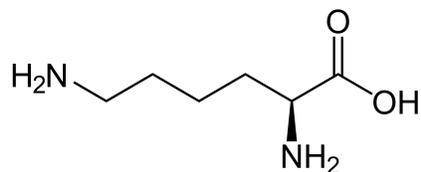
**Directions:** The following questions below is based on the crystal structure of MBP fused activation-induced cytidine deaminase (AID) in complex with cytidine (**5w1c.pdb**). Answer these questions in the most concise and specific way possible. The point values for each question are addressed in the parentheses. **(63)**

1. The following questions below refer to **residue 175. (8)**
  - a. Identify the amino acid. (1) **glycine (accept Gly)**
  - b. This amino acid has a significant role in the formation of what secondary structure? (2) **alpha helices**
  - c. **True or False:** This is the only achiral proteinogenic amino acid that exists. (1) **True**
  - d. Would this amino acid fit into an environment that is hydrophilic, hydrophobic, or both? (1) **both**
  - e. Explain your answer for question 1(d). (3) **The side chain is only one hydrogen atom.**
  
2. The following questions refer to **residue 113. (8)**
  - a. Identify the amino acid. (1) **alanine (accept Ala, A)**
  - b. What is the IUPAC systematic name of this amino acid? (2) **2-aminopropanoic acid**
  - c. Is this amino acid polar or nonpolar? (1) **nonpolar**
  - d. Based on your answer in part 2(c) and the Jmol environment, where do you expect to see this amino acid? (2) **interior/inside of the protein**
  - e. Which isomer of this amino acid is the only one incorporated into proteins? (2) **L-isomer**
  
3. The following questions refer to **residue 367. (10)**
  - a. Identify the amino acid. (1) **Threonine (Thr, T)**
  - b. This amino acid is a precursor of what other amino acid? (2) **glycine**
  - c. **True or False:** The side chains of this amino acid does not hydrogen bond well. (1) **False**
  - d. Theoretically, if this residue was near serine, what common small motif can form between these two amino acids? (3) **accept one of the following and award 3 points: ST turns, ST motifs, ST staples**
  - e. The sidechain of this amino acid can undergo what type of post translational modification? (3) **O-linked glycosylation**

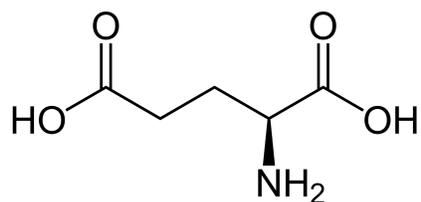
4. The following questions below refer to **residue 233. (9)**
- State the full name of this amino acid. (1) **Tryptophan**
  - State the one letter name of this amino acid. (1) **W**
  - What is the X, Y, Z coordinates of this amino acid? (3) **283.83, -5.566, 1252.864**  
(award one point for each coordinate; the order is X, Y, Z, respectively)
  - What is the chain letter of this amino acid? (2) **B**
  - In two letters, describe the atom type of this amino acid. (2) **CA**
5. The following questions below refer to the relative positions of amino acids to each other. Write as many decimal values as you can. Be mindful of your units. **(10)**
- What is the distance between residues 53 and 67 in nanometers? (2) **7.967**
  - What is the distance between residues 53 and 58 in nanometers? (2) **1.167**
  - What is the distance between residues 37 and 101 in angstroms? (2) **28.81**
  - What is the distance between residues 21 and 119 in angstroms? (2) **36.84**
  - What is the distance between residues 113 and 119 in picometers? (2) **1940.0**
6. How many hydrogen bonds are there in the entire protein structure? (2) **654**
7. How many hydrogen bonds are there in the entire nucleic structure? (2) **28**
8. How many struts can you add in the entire protein structure? (2) **159**
9. How many struts can you add in the entire nucleic structure? (2) **0/none**
10. The following questions below refer to **residues 120-399. (10)**
- How many total atoms are there in residues 120-399? (2) **3856**
  - How many struts can you add in residues 120-399? (2) **65**
  - How many aromatic atoms are there in residues 120-399? (2) **1728**
  - How many aliphatic atoms are there in residues 120-399? (2) **2466**
  - How many polar atoms are there in residues 120-399? (2) **4059**

11. Sketch the **full structure** of the following residues as **organic compounds** (do not display hydrogens bonded to carbon; do not display carbon). If you do not sketch the residues as organic compounds, you will **not** receive credit. Each sketch is worth **4** points. There is **no** partial credit granted for this question. However, the canonical or zwitterionic versions of some residues will be accepted. **(20)**

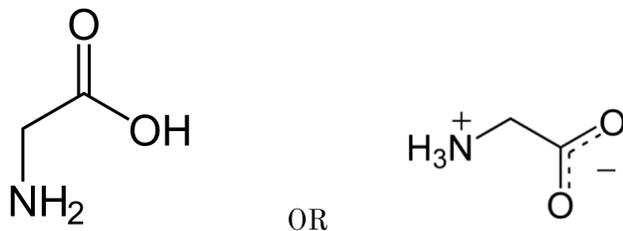
a. Residue **296**.



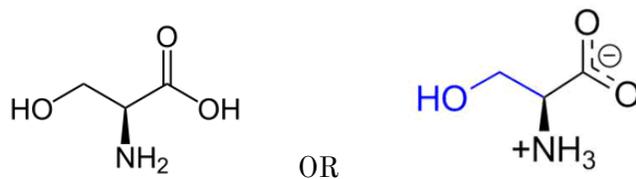
b. Residue **372**.



c. Residue **144**.



d. Residue **307**.



e. Residue **297**.

