UNIVERSITY OF CALIFORNIA, BERKELEY
CHEM C130/MCB C100A
MIDTERM EXAMINATION #1
SEPTEMBER 22, 2016

INSTRUCTORS: John Kuriyan and David Savage

THE TIME LIMIT FOR THIS EXAMINATION: 1 HOUR 50 MINUTES

SIGNATURE:

_____

Please **SIGN** your name on the line above in **INDELIBLE INK**.

YOUR NAME:
_____KEY_____

PLEASE PRINT your name (**IN INDELIBLE INK**) on the line above (& on the top right hand corner of every page).

PLEASE CIRCLE the name of your GSI:

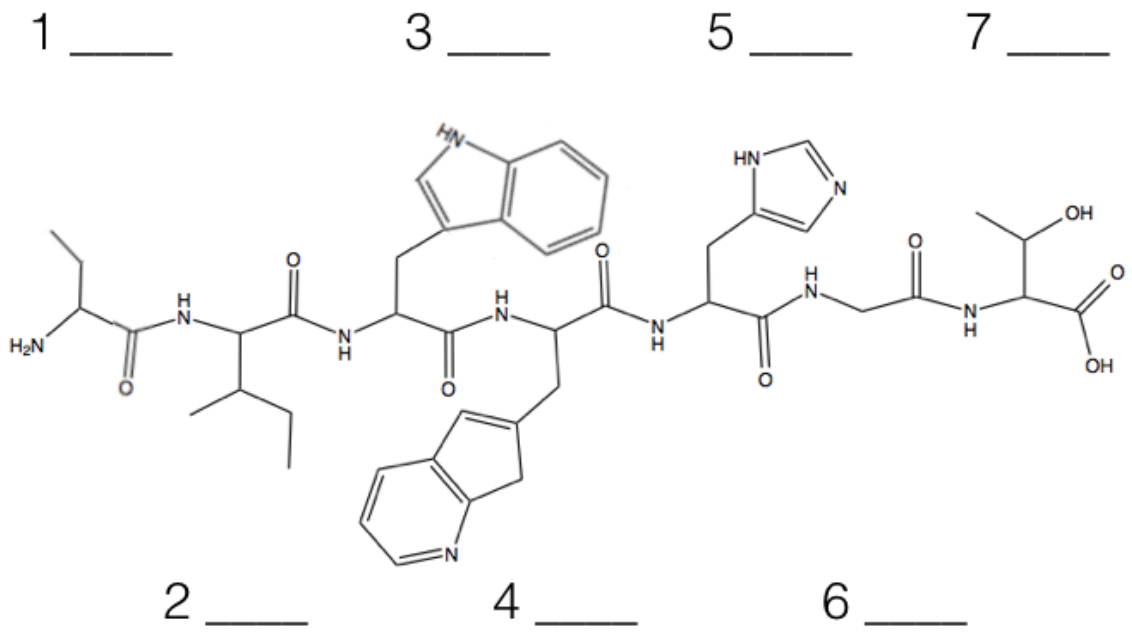
Alec Heckert	Thomas Laughlin
Pei Liu	Mark O'Dair
Kersh Thevasundaram	Carl Ward

PLEASE WRITE all of your answers **AS LEGIBLY AS POSSIBLE**.
Note that any exam submitted for a regrade should have been written in indelible ink.
SCORING. The exam consists of 5 questions totaling 100 points as broken down in this table:

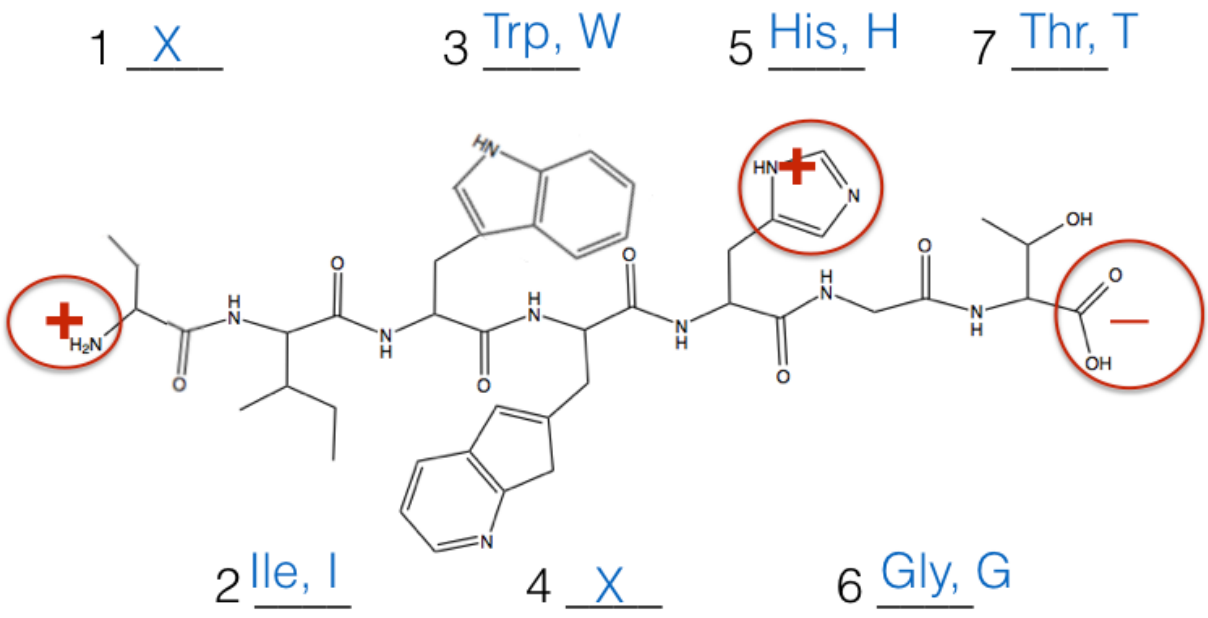
Question	Part A	Part B	Part C	Part D	Your Total	Max Score
1.	6	6	6	2		20
2.	5	5	4	6		20
3.	9	7	4			20
4.	6	8	6			20
5	----------	----------	----------	----------		20
				TOTAL		100

Q1 (20 points)

Q1A (6 points)



(i) (4 points) The diagram above shows the structure of a peptide with five natural amino acid residues and two unnatural ones. For the natural amino acids, write down the one-letter and three-letter codes in the appropriate blank spaces. For the unnatural amino acids, write X.

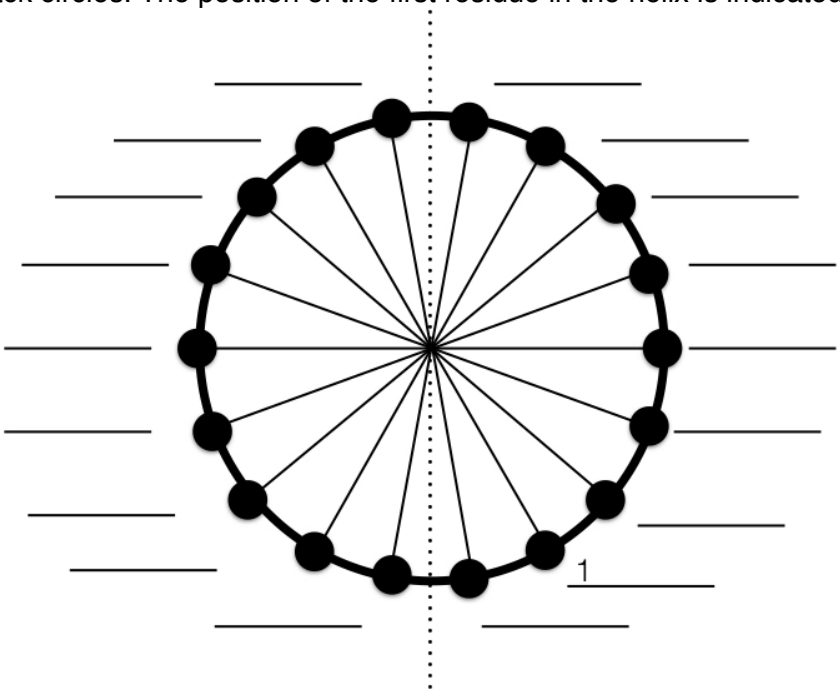


(ii) (1 point) The peptide shown above is in the completely neutral form. At pH 6.0, some of the chemical groups in the peptide will be substantially in a charged form. Circle the groups that will be in the charged form and indicate the charge next to the circle with either "+" or "–".

(iii) (1 point) The interfaces between helices in transmembrane proteins are usually composed of hydrophobic residues. Since these regions face other protein regions, it is possible to design membrane proteins so that polar groups form hydrogen bonds across such interfaces. Natural membrane proteins rarely do this however. What is the principal reason for this?

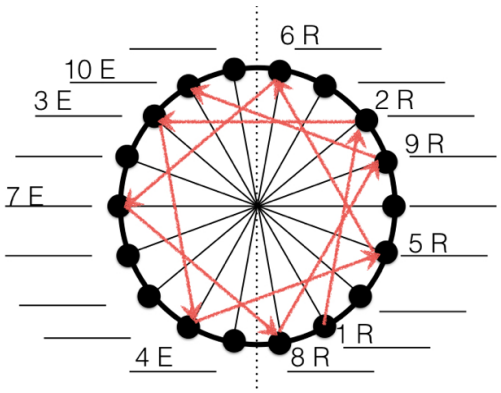
Answer: Such a design would be mutationally sensitive. Since hydrogen bonds have precise geometries, mutation of one partner to any other amino acid typically will break the hydrogen bond. The un-mutated partner will then no longer be able to form a hydrogen bond, which has a high desolvation penalty within the membrane.

Q1B (6 points) Shown below is a helical wheel. The circle is divided into 18 equal sectors. Assume that the helix has the N-terminal end below the plane of the page and the C-terminal end above the plane of the page. The Cα atoms of residues on the helical wheel are shown as black circles. The position of the first residue in the helix is indicated by the number 1.

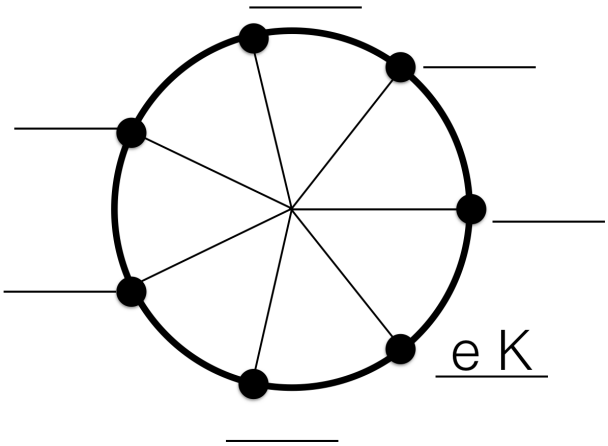


A ten residue peptide forms a monomeric α-helix that interacts with the negatively charged surface of a lipid bilayer. The peptide is made up of only two kinds of residues, glutamate and arginine. One face of the helix interacts with the surface of the bilayer, and the other side interacts with water.

In the helical wheel shown above, the face of the helix to the right of the dotted line interacts with the lipid headgroups, and that to the left interacts with water. Starting from the Cα atom labeled “1”, complete the helical wheel by writing down the one letter code for the reside at positions 1 to 10. Be cautious about the direction you go around the circle. Helical wheels for you to practice on are given two pages down.

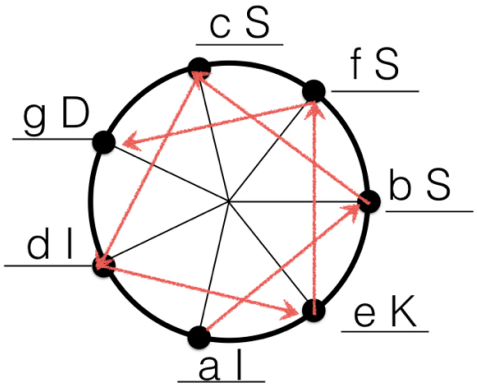


Q1C (6 points) Shown below is a helical wheel representation for a peptide that forms a coiled coil. As before, the Cα atoms of residues on the helical wheel are shown as black circles.
(i) (2 points) Why is this circle only divided into seven sectors, as opposed to the 18 sectors in the helical wheel in part B? Provide a convincing reason for this.



(ii) (4 points) A peptide containing only the residues serine, isoleucine, aspartate and lysine forms a dimeric coiled coil in water, but the coiled coil is not stable when salt is added.

Use the helical wheel shown above to draw the sequence of seven residues in the coiled coil (denoted *a* through *g*), such that no hydrophobic residues are solvent exposed when the dimer is formed. Assume that the direction of the helix is the same as in Q1B(ii), with the N-terminus below the page and the C-terminus above. The residue at the *e* position, and its identity, is indicated. Draw arrows connecting subsequent residues, and label each one with the position (*a* through *g*) and the single letter code for the residue type. There is a practice helix on the next page.

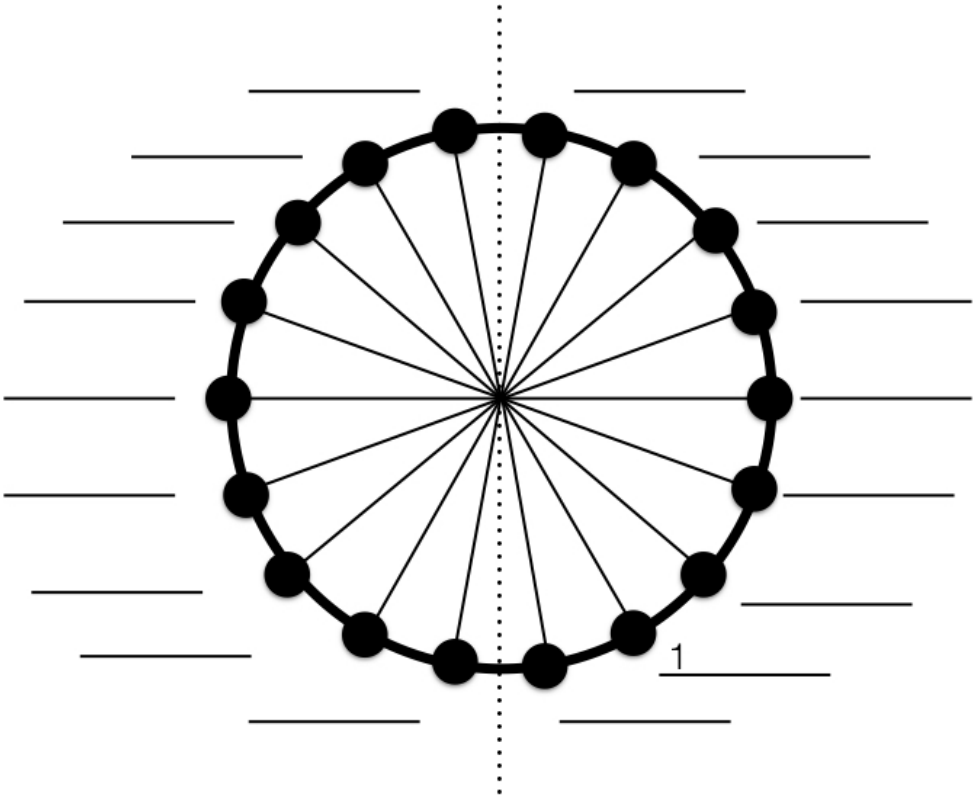


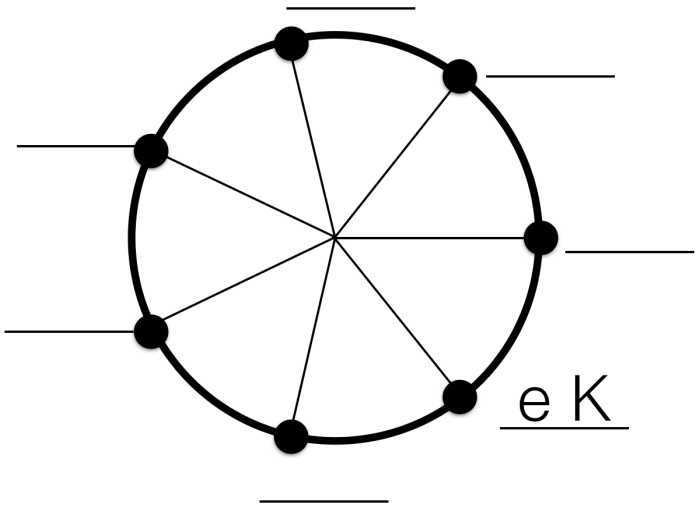
Note: It is acceptable to put D and K at the b, c and f positions as well.

Q1D (2 points)
In bacteriorhodopsin, there are several residues that are important for the proton pumping function of the protein. One of these residues is mutated to asparagine, and the protein stops working, even though the protein is folded properly and retinal is bound correctly. Explain why the mutant protein no longer functions as a pump.

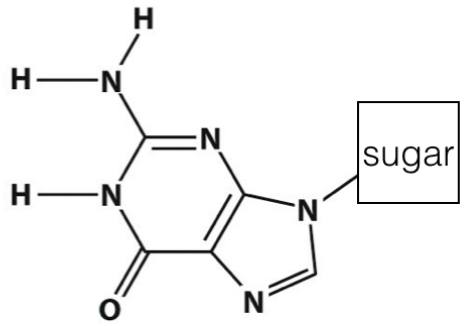
Answer: There are glutamate and aspartate sidechains that can protonate/deprotonate, and thereby help in shuttling protons up or down the pump. Asparagine cannot titrate easily, and so its presence would block the flow of protons.

Practice helical wheels. Provide your final answer to Q2, parts i and ii on the previous page. Any drawings or writings on this page will be ignored during grading.

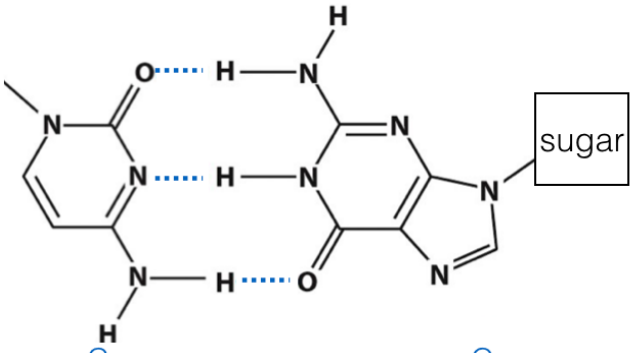




Q2. (20 points).
(A) (5 points) In the schematic below, a nucleotide forms a Watson-Crick basepair interaction with a nucleotide that is not shown. Draw the missing nucleotide of this pair, draw the proper hydrogen bonding pattern, and label them both with their correct single-letter code in the space provided.

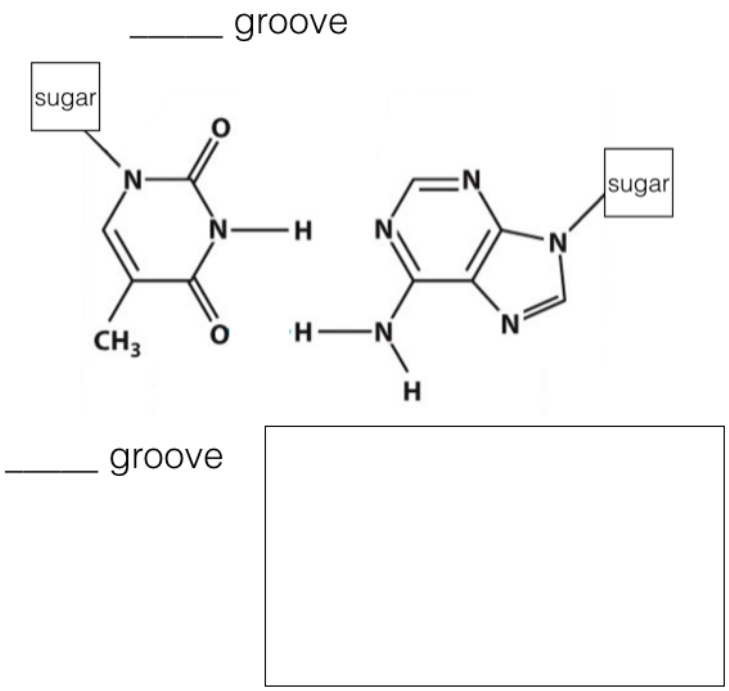


nucleotide:_____ nucleotide:_____



nucleotide: C nucleotide: G

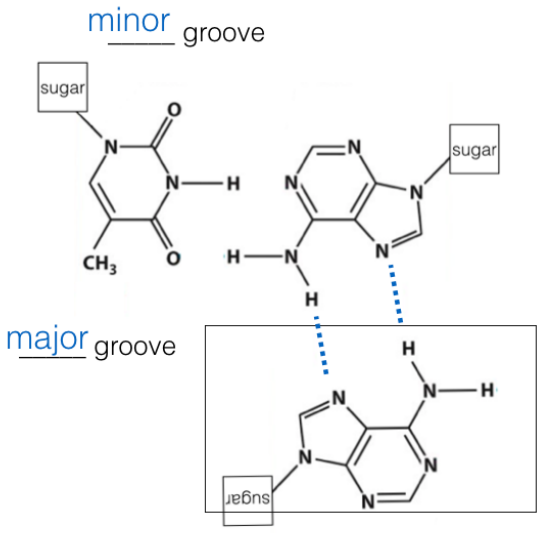
(B) (5 points) The schematic below shows another Watson-Crick base pair.

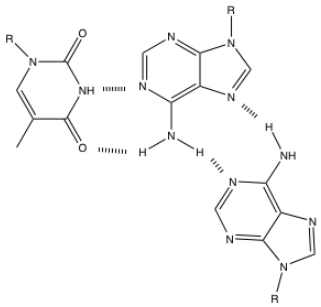


(i) (1 point) Draw the hydrogen bonds between the base pairs

(ii) (1 point) Identify the major and minor grooves by filling in the blanks

(iii) (3 points) In the space provided, an adenine base forms a Hoogsteen base pair with one of the Watson-Crick base pairs. Draw the adenine base, and indicate the hydrogen bonds that it forms.





The above Hoogsteen base pair is also acceptable.

Q2 (C) (4 points)

(i) (1 point) DNA in cells is typically *negatively* supercoiled. One advantage of this is that the DNA is slightly unwound, allowing easier access to enzymes that work on DNA. Discuss *another* major advantage for the cell that the negative supercoiling provides.

Answer: Negative supercoiling of DNA is used to package DNA around histones (nucleosomes), allowing substantial **compaction** of the genome.

(ii) (1 point) Base-stacking interactions are the principal source of stabilization of the DNA double helix. Explain why base-stacking interactions are so stable.

Answer: The **partial charges** on the nucleotides are aligned in base stacks so that the **electrostatic energy** is favorable.

(iii) (2 points) The major groove of the A-form double helix is less accessible for proteins than the major groove of B-form DNA. What is the principal disadvantage of the narrower major groove in terms of one of the functions of DNA?

Answer: Proteins, such as transcription factors, use the major groove for **sequence-specific recognition of DNA**. The **minor groove has less sequence-specificity.** The narrower major groove of A-form double helix makes it harder for proteins to achieve sequence-specific recognition of the nucleotides.

Q2D (6 points) Assume that in relaxed double-helical DNA the rotation between adjacent base pairs is 36°. Consider a closed circular DNA of 2000 base pairs in a relaxed and flat form.

(i) (2 points) What is the value of the linking number, L? Explain how you work out the answer.

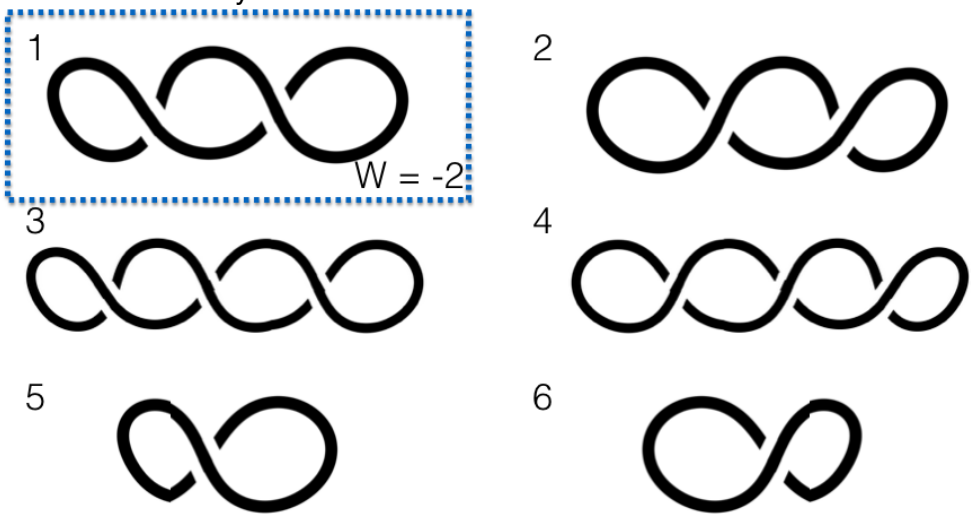
Answer: There are 10 residues per turn, so the DNA has 2000/10 = 200 turns. Hence the linking number is 200.

(ii) (2 points) Now assume that the closed circle is broken, the DNA in unwound by two full turns, and then the circle is resealed. What is the value of the twist, T, if the DNA is held flat?

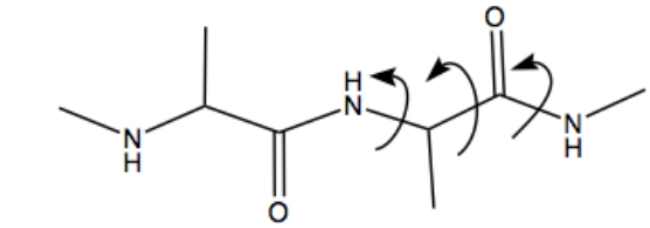
Answer: Since the DNA is **unwound** by 2 turns, the linking number will be reduced to 200 - 2 = 198.

To answer the next question, realize that the twist will initially also be 198. But the twist will return to 200 in order to maintain the base-stacking angle. Thus, there will be a +2 change in twist when the DNA relaxes in 3 dimensions. This will be compensated by a -2 change in writhe. Since the starting value of W is zero, the resulting value of writhe will be -2.

(iii) (2 points) The circular DNA is then allowed to relax in three dimensions. Circle the diagram below that correctly describes the writhe of the relaxed DNA.

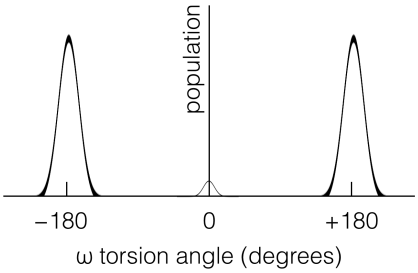


Q3 (20 points)
Q3A (9 points)
(i) (2 points) Shown below is a segment of a peptide chain, with three backbone torsion angles indicated. Write down the name of the angles in the space provided.



torsion angles: Φ ψ ω

(ii) (3 points) A scientist measures the value of the ω torsion angle for thousands of bonds in protein structures in the Protein Databank, and obtains the distribution of torsion angles shown below.



Why is the distribution is peaked at 0^0 and 180^0? (1 point)

Answer: Because of **partial delocalization** of the amide bond, the angle is constrained to be in the cis (0^0) or trans (180^0) conformations.

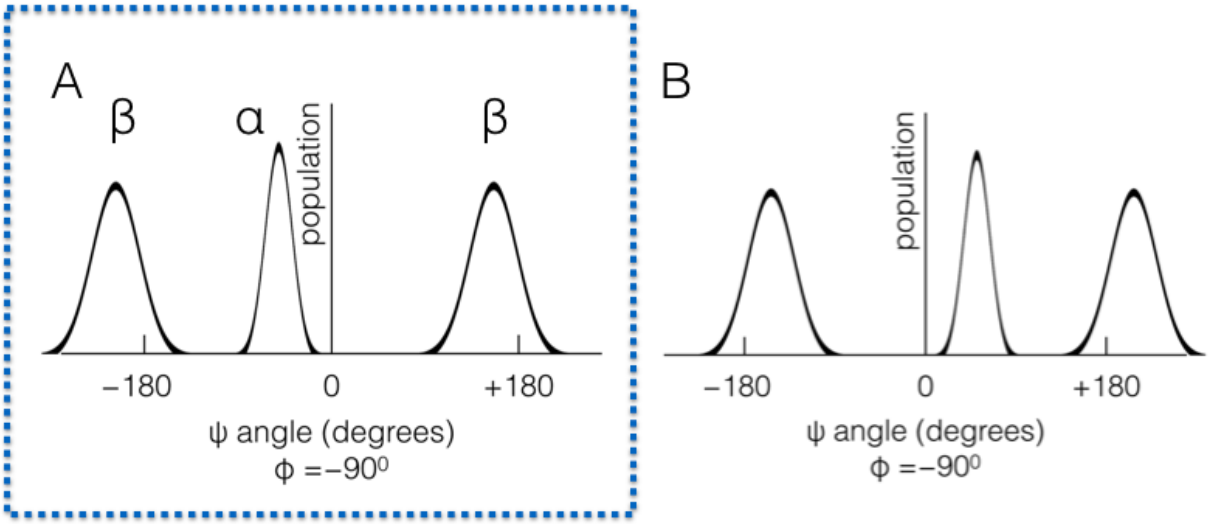
Why is the peak at 0^0 is so much smaller than the peak at 180^0? (1 point)

Answer: The **cis conformation** causes more steric clashes, so it is rare.

Which type of residue contributes the most to the peak at 0^0? Provide the three letter code. (1 point)

Answer: Pro, P (I the answer is given as Gly, G I would accept that as reasonable.)

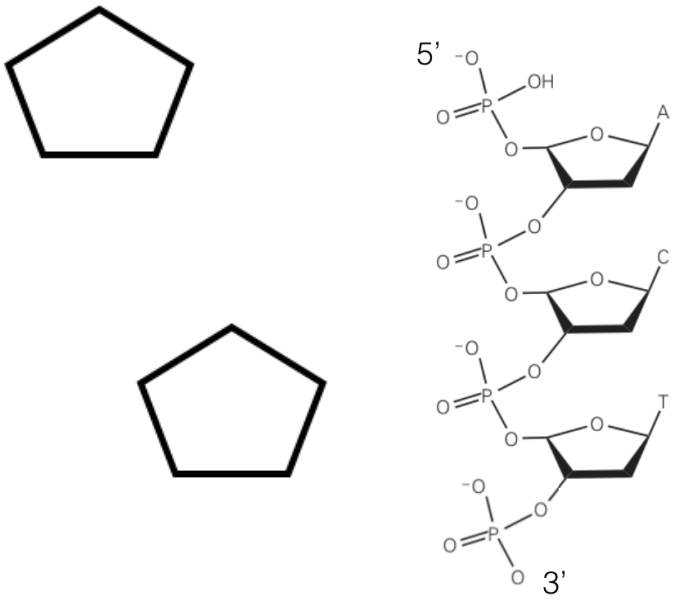
Q3A (iii) (4 points) The scientist also measures the distribution of the values of the ψ torsion angle for a fixed value of ϕ (–90⁰) for all the proteins in the Protein Databank.



Which of the two distributions shown above corresponds to the distribution expected from the Ramachandran Diagram? **Cross out** the incorrect distribution. (2 points)

For the correct distribution, identify the secondary structures associated with the peaks by writing α or β above the appropriate peaks. (2 points)

Q3B. (7 points)
(i) (3 points) Shown below are two deoxyribose ring skeletons from a single strand of DNA. Draw the missing atoms of the backbone for this DNA. Label the 3' and 5' ends of the chain, and show the phosphate groups at the 3' and 5' ends.



(ii) (2 points) Suppose this organism uses nucleotides such as the molecule 2'-deoxyadenosine 3'-triphosphate for its DNA synthesis. How would the mechanism of polymerization of DNA differ from that normally seen in cells?.

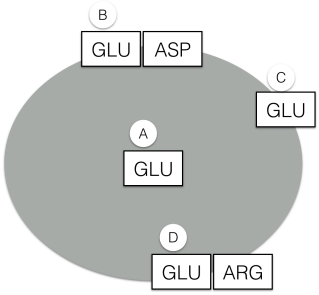
Answer: This nucleotide possesses the triphosphoryl group on the opposite C (3') as normal dATP (5'). As such, since polymerization results from a attack of the alcohol group on the pyrophosphoryl bond, the directionality of synthesis must be the opposite. Polymerization would go in a 3' to 5' direction.

(iii) (2 points). Chemists have made derivatives of DNA and RNA with an amide bond (e.g. like a peptide) instead of a phosphodiester linkage. Surprisingly, these nucleic acid analogs form double helices, which are significantly more stable than double helices formed by normal DNA and RNA. Explain why this is the case.

Answer: Base stacking interactions contribute the most stabilizing energy to the structure of nucleic acid helices. Base stacking in the modified oligonucleotides will also stabilize the formation of helices, provided that the linker group is able to adopt a geometry consistent with the helix. The negatively charged phosphate groups in oligonucleotides repel each other and destabilize the helical structure. In the modified oligonucleotides, there is no charge on the linker group, they do not repel each other, and the structure is stabilized with respect to normal oligonucleotides.

Q3C (4 points)

A protein contains four glutamate residues, which are labeled A, B, C and D in the diagram below. Glu A is buried in the hydrophobic core of the protein. Glu B is on the surface of the protein, located close to an aspartate sidechain. Glu C is on the surface of the protein and does not interact with any other residues. Glu D is on the surface, and next to an arginine sidechain.



The glutamate residues have pKa values of either 7.0, 4.5 or 3.0 (only 3 possible values).

Write down the expected pKa values of each Glu residue below:

Glu A: _____ 7.0 (easier to protonate, since removal from water will favor neutral form)

Glu B: _____ 7.0 (easier to protonate, since electrostatic repulsion will favor neutral form).

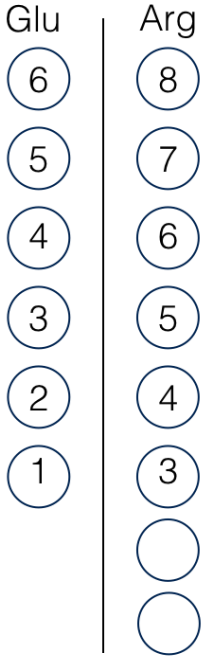
Glu C: _____ 4.5 (will behave like a normal isolated glutamate sidechain)

Glu D: _____ 3.0 (harder to protonate, since ion-pair formation with arginine favors the charged form)

Q4. (20 points)
Q4A. (6 points) (i) (2 points) In Anfinsen's experiments concerning the folding of ribonuclease, in one experiment he unfolded ribonuclease in urea and reducing agent, then removed the reducing agent and added oxygen in the presence of urea, then removed the urea. Briefly explain the key conclusion that Anfinsen reached from the results.

Answer: Because the recovered activity was so low, and consistent with random pairing of cysteines, he was able to conclude that the presence of urea and denaturing agent had completely unfolded the ribonuclease molecule. By doing the refolding the other way he was able to regain full activity, thereby proving that the conversion of the fully unfolded form to the folded form could happen in a test-tube, without any other information.

(ii) (4 points) A protein contains six glutamate residues and eight arginine residues.
In the unfolded protein, the possible conformations all have at least six ion pairs, between the six glutamate residues and six of the eight arginine residues. Assuming that all glutamates can interact with all arginines, how many ion-paired configurations are there in the unfolded protein?



Ways of choosing Glu residues: 6!

Ways of choosing Arg residues: 8!/2

Ways of rearranging the six ion pairs (vertical order does not count): 6!

So: total number of different ion pairs:

$$\frac{6! \times 8!}{6! \times 2} = \frac{8!}{2} = 20{,}160$$

Q4(B) (8 points) Shown below are short segments of three proteins, containing only the amino acids D, F, G, L and K.

```
Protein A: DFGLLK
Protein B: KFGFLK
Protein C: LFLDLK
```

(i) (3 points) Using the BLOSUM substitution score matrix (given below), explain whether protein A is more closely related to protein B or to protein C. Explain your reasoning completely.

	D	F	G	L	K
D	6	-3	-1	-4	-1
F	-3	6	-3	0	-3
G	-1	-3	6	-4	-2
L	-4	0	-4	4	-2
K	-1	-3	-2	-2	5

Answer:

$$S_{AB} = S_{DK} + S_{GG} + S_{LF} \left(+ \text{ constant, which can be ignored}\right)$$
$$= -1 + 6 + 0 = 5$$

$$S_{AC} = S_{DL} + S_{GL} + S_{LD} \left(+ \text{ constant, which can be ignored}\right)$$
$$= -4 + (-4) + (-4) = -12$$

Hence, A is more closely related to B, since the net BLOSUM score is higher for the A-B alignment.

(ii) (5 points) What is the relative likelihood that protein A is related to protein B versus protein C? That is, calculate the following ratio:

$$\frac{\text{likelihood that A and B have a common ancestor}}{\text{likelihood that A and C have a common ancestor}}$$

Answer:

$$\frac{L_{AB}}{L_{AC}} = \frac{2^{S_{AB}/2}}{2^{S_{AC}/2}} = 2^{\frac{1}{2}(S_{AB} - S_{AC})} = 2^{\frac{1}{2}(17)} = 2^{8.5} = 362$$

Q4C (6 points)

Q4C(i) (3 points) Shown below is a small sequence alignment block, in which only G, D, L and V residues are shown. The gray boxes represent other amino acids (not G, D, L or V).

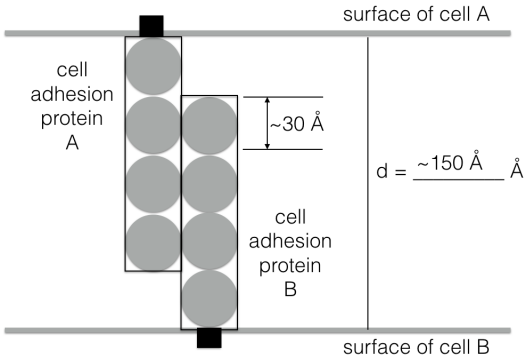
G		L	D
	L	V	
D			L
	V		
L		V	G

Based on this matrix, and considering how the BLOSUM scores are calculated, which would be the higher substitution score, S_{GD} or S_{LV}? Explain your answer without doing any calculations.

Answer: A reasonable guess is that S(G,D) will be larger than S(L,V). There are more L and V residues in the block, so the random probability of getting L,V pairs is higher, reducing the score. (There is one more LV pair in the block than GD pairs, so you will also be given credit for reasoning that the LV score is higher, although it turns out to be not so.)

Q4C(ii) (3 points)

Two cells make contact with each other through cell adhesion proteins. These adhesion proteins consist of four typically sized protein domains arranged in a row, as shown in the schematic diagram. Based on what you know about the sizes of typical protein domains, estimate the distance, d, between the surfaces of the cells. Write down your estimate in units of Å in the space provided above. Your answer must be within a factor of 3.0 of the correct answer to get full credit.



Q 5. (20 points) Multiple choice and True/False questions. Circle the ***best*** option (or TRUE or FALSE).
+2 points for each correct answer, -1 points for each wrong answer.
To get the maximum score you do not need to answer all the questions, so be careful not to answer questions incorrectly.
Maximum points: 20. Minimum points: 0.

(i) The helical structure of DNA must be right handed so that bases can hydrogen bond in a complementary fashion. TRUE / FALSE

(ii) The Ramachandran diagram is calculated using a hard-sphere contact potential for alanine residues, and yet it explains the structure of real proteins with all kinds of amino acids (except glycine). Circle the best reason for this:
(a) The Ramachandran diagram is consistent with secondary structure formation
(b) The alanine sidechain is smaller than other sidechains (except glycine)
(c) Alanine is a representative sidechain, so it generates a Ramachandran diagram that is generally applicable

(iii) Watson-Crick base pairing is the defining energetic contributor to stabilize DNA in the double helical form. TRUE / FALSE

(iv) RNA structures contain non-Watson-Crick base pairs such as G-U because the necessity to maintain the uniform dimensions of base pairs does not apply in RNA. TRUE / FALSE

(v) DNA is the repository of genetic information because the absence of a hydroxyl group at the 2' position makes it less susceptible to cleavage. TRUE / FALSE

(vi) Ionic interactions tend to be weaken because of the screening effect of water. TRUE / FALSE

(vii) Choose the amino acid substitution that results in the most conservative change in hydrophobicity:
(a) V S
(b) F N

(c) L F
(d) Q I

(viii) Membrane proteins can insert into the membrane with a β-α-β motif that is entirely within the membrane. TRUE / FALSE

(ix) Two α helices pack against each other with an inter-helix angle of ~50^0. Choose the best explanation for why this happens:
(d) The ridges and grooves formed by the i, i+4 residues pack against each other
(e) The ridges formed by the i, i+4 residues pack into grooves formed by the i, i+3 residues
(f) The ridges and grooves formed by the i, i+3 residues pack against each other

(x) Circle the substitution or mutation that would have the *least* effect on protein stability:
(a) L (in center of an α helix) to P
(b) G (in a turn or loop) to P
(c) I (in the hydrophobic core) to M
(d) L (in a transmembrane helix) to Q

(xi) Shown below is a list of the twenty amino acids, in decreasing level of hydrophobicity from the left to right.

W F L I R Y V C P H^0 T S A Q N G M K E D

Which two residues are clearly out of order? Circle the best answer.

(a) H^0 M
(b) R M
(c) R Y
(d) T Q

(xii) Single-stranded DNA and RNA can form helices without base pairing. The rise per turn of such single-stranded helices is usually close to that of double-stranded RNA or DNA. This is because (circle the best answer):
(a) the rise per turn is determined by the phosphate-phosphate distance
(b) the rise per turn is independent of whether the helix is formed by RNA or DNA
(c) van der Waals contact distances determine the spacings between stacked basepairs
(d) the size of the sugar group determines the rise per turn

(xiii) When comparing the sequences of two proteins, what is the level of sequence identity below which structural similarity becomes uncertain? Circle the best answer:
(a) 55%
(b) 45%
(c) 35%
(d) 25%
(e) 15%