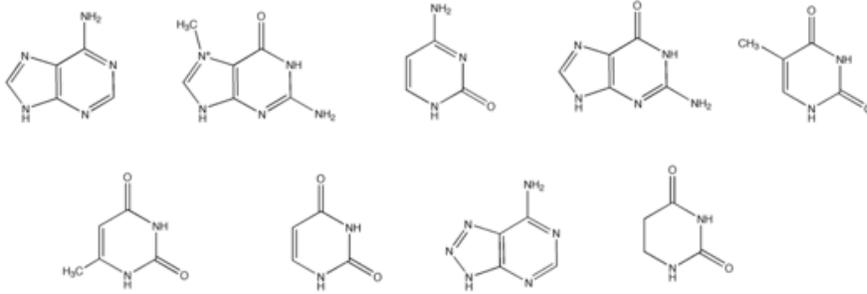


QUESTION 1 (20 points)

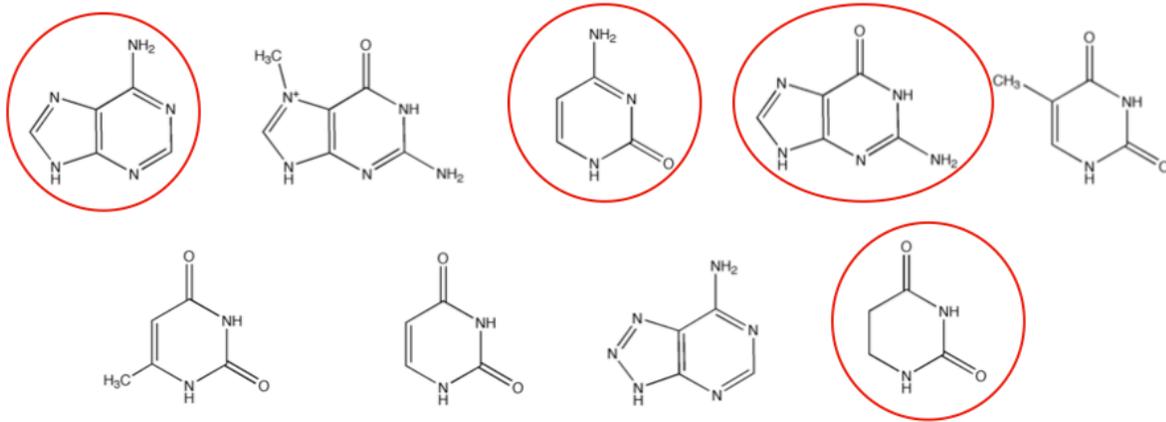
Q1A (9 points)

Shown below are nine nucleotide bases.

(i) Circle the four bases that are found in genetically-encoded RNA. (4 points)

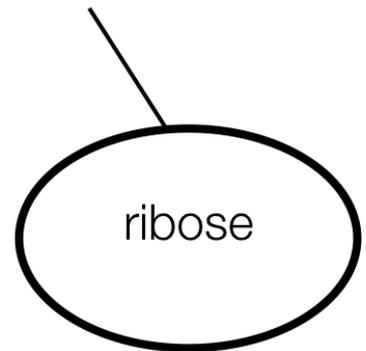
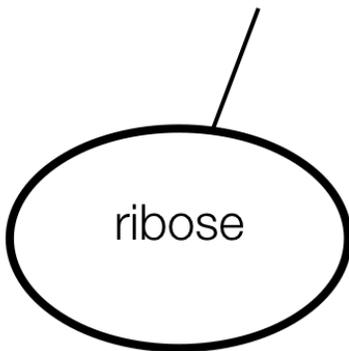


Answer:

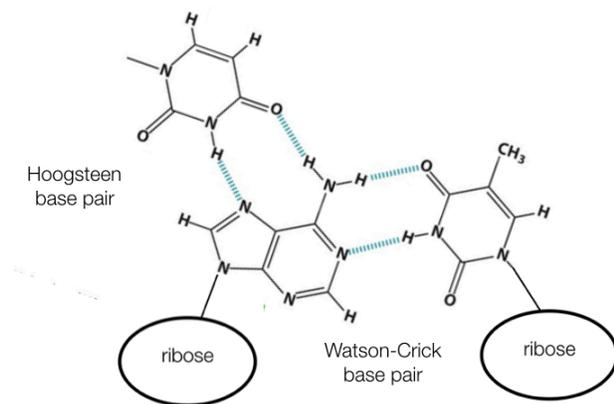


(ii) (3 points) In the diagram below, draw the chemical structure of the Watson-Crick base pair that is found in DNA, but not RNA. Indicate all the hydrogen bonds by dashed lines and the connection to the ribose sugar (the sugars and the glycosidic bonds have already been drawn for you, schematically, no need to add anything for them).

(iii) (2 points) A uracil makes a Hoogsteen base pair with the purine nucleotide in the major groove edge of the Watson-Crick base pair. Draw the structure of the uracil base and indicate the hydrogen bonds. No need to draw the glycosidic bond or sugar for the uracil. (2 points)

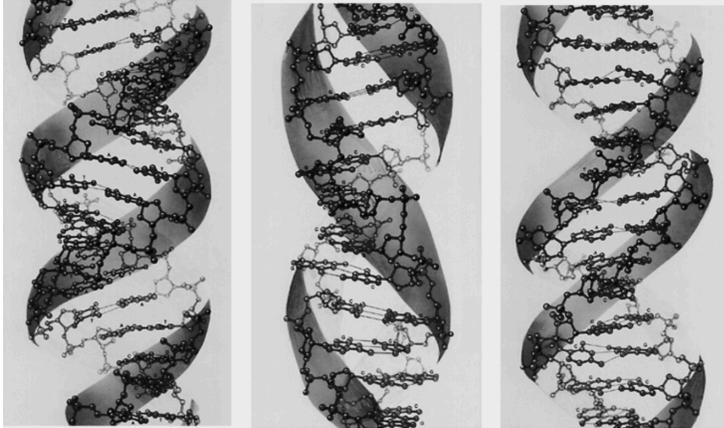


Answer:



Q1 (B). (7 points)

(i) Shown below are three double-helices. Identify each form by filling in the blanks below. (3 points)



(1) A FORM (2) Z FORM (3) B FORM

(ii) In the table shown below, put a “X” mark if the given form has that property. (4 points)

	Helical form 1	Helical form 2	Helical form 3
Left-handed helix		X	
RNA is able to adopt this form	X		
Results in negatively supercoiled DNA when the rest of the DNA is B-form		X	
Easiest access to the major groove for sequence-specific interactions			X

Question 1, contd.

Q1C (4 points) A protein involved in recombination, known as RecA, binds to single-stranded DNA and stretches the length of the strand. When the DNA is stretched by the protein, the bases are not evenly spaced. Instead, groups of nucleotides retain the B-form conformation, while the phosphate backbone between them is extended, as shown schematically in the diagram below.



This kind of staggered distortion uses less energy than forcing a uniform separation between all the base pairs. Describe the origin of the resistance of DNA to being uniformly stretched.

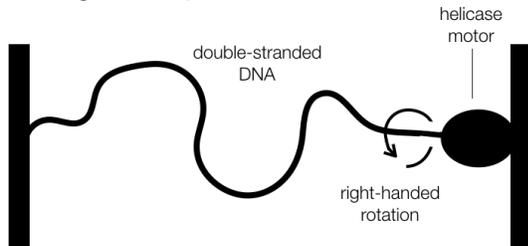
Answer:

Single-stranded DNA adopts the B-form conformation in order to preserve **base-stacking**, which provides the dominant stabilization energy of DNA. By separating the stretched base pairs into groups that retain the B-form conformation, base stacking is preserved to the extent possible.

QUESTION 2 (20 points)

Q2A (8 points)

A linear double-stranded DNA containing 1980 base pairs is fixed at one end, and attached to a rotating-motor protein at the other end, as shown schematically below.



Initially, the DNA is relaxed, with 11 base pairs per turn. The helicase motor then rotates the end of the DNA by 3 full turns in a right-handed sense (i.e., introduces additional turns into the DNA), and then keeps that end fixed. No additional base pairs are introduced in the process.

(i) What is the initial value of the twist, T , linking number, L , and writhe, W , before the helicase introduces any rotation? (2 points)

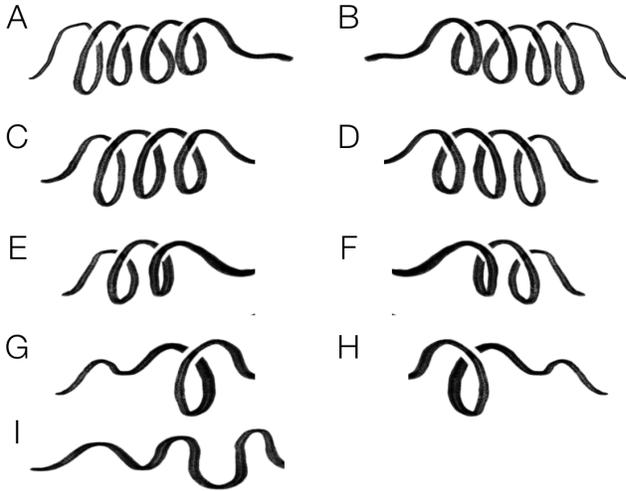
$$T = \underline{180} \quad L = \underline{180} \quad W = \underline{0}$$

(ii) What is the final value of the twist, T , linking number, L , and writhe, W , after the helicase has completed the rotations, and the DNA has achieved a lower energy state? Explain the logic of how you work out the answer. (2 points)

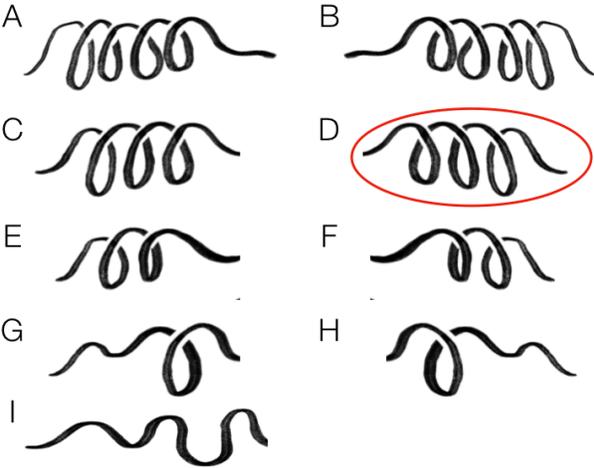
Answer: The motor increases the linking number to 183 by introducing 3 additional rotation. The DNA responds by reducing the twist back to the original value of 180, compensating with positive writhe ($W = +3$)

$$T = \underline{180} \quad L = \underline{183} \quad W = \underline{+3}$$

(iii) Shown below are **solenoidal** forms of the DNA after the process is completed. Circle the appropriate form of the DNA. (2 points)

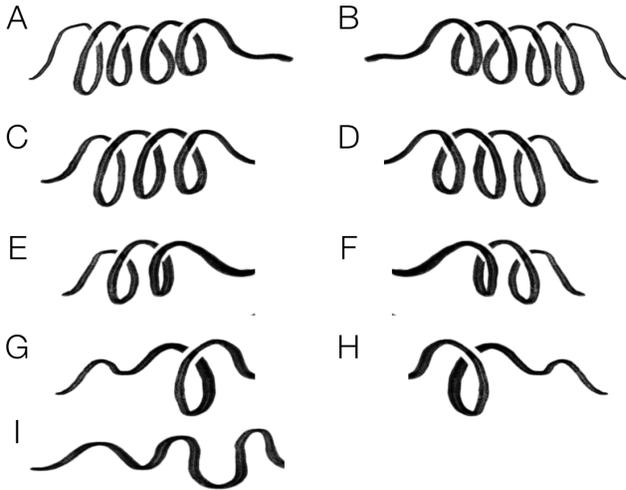


Answer:

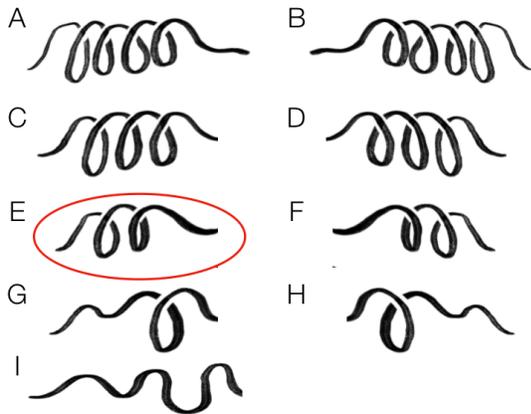


(iv) (2 points)

A nucleosome interacts with two superhelical turns of negatively supercoiled DNA. Circle the structure in the diagram below that corresponds to the DNA recognized by the nucleosome. (Note that this diagram is identical to the diagram on the previous page.)



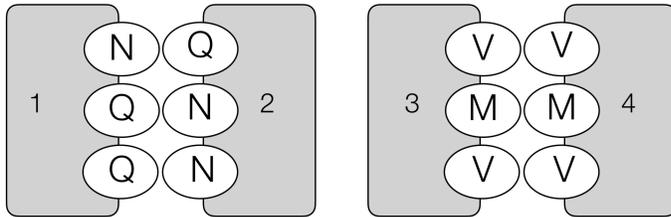
Answer:



Question 2, continued.

Q2B (8 points)

Shown below are schematic representations of four proteins, labeled 1, 2, 3 and 4. The proteins form dimers, as shown. The sidechains that interact to stabilize the dimer are shown by single-letter codes.



The strength of each dimer is measured in water and in octane.

- (i) When the 1-2 dimer is transferred from water to octane, does it get stronger or weaker? Explain. (2 points)

Answer:

It gets stronger, since there is no competition with water. Octane is a hydrophobic solvent, and cannot form hydrogen bonds.

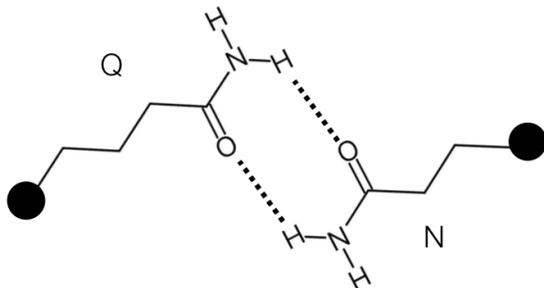
- (ii) When the 3-4 dimer is transferred from water to octane, does it get stronger or weaker? Explain. (2 points)

Answer:

It gets weaker, since the hydrophobic effect no longer is a driving force for association. Octane is a hydrophobic solvent.

- (iii) Draw the structure of the sidechains of one Q-N interacting pair, clearly indicating hydrogen bonds, if any. Indicate which is Q, and which is N, in your diagram. (2 points)

Answer: (other orientations are possible, but the hydrogen-bonding pattern must be maintained.)



(iv) The middle N residue in protein 2 is mutated to Q. The protein is then found to never form a dimer. Explain why. (2 points)

Answer:

Because the sidechains of N and Q are of different lengths, the hydrogen-bonding pattern that is set up for the original protein is disrupted by the mutant.

Q2C (4 points)

Consider a lipid bilayer made of lipids that are longer than normal, so that the hydrophobic position of the bilayer is 45 Å thick, instead of the normal ~35 Å. A protein contains **only one** α helix that spans this bilayer. Three different protein sequences are shown below:

Sequence 1: AVNSQLLKDFIALVGFMLFFLVILMSTLGTQALLAVALVLVAFMILVAAVLN

Sequence 2: AVTSQLLRRVIMAFFGGILVMALLAMFAMVLLAFGMVFWKRRDNSQPR

Sequence 3: AVLPGDRLLVLFMAVAVLLGLLGAMIFAAGLSQDLLATPNILVDKSRTKQ

(i) Which of the sequences is most consistent with this transmembrane helix? Briefly explain your choice. (2 points)

Answer: The rise per residue of an α helix is 1.5 Å. Hence, there will be approximately $45/1.5 = 30$ residues spanning the membrane. These should be hydrophobic. Sequence 2 best satisfies this condition.

(ii) In the diagram above, draw a box around the transmembrane section of the appropriate sequence. (2 points)

Answer:

Sequence 1: AVNSQLLKDFIALVGFMLFFLVILMSTLGTQALLAVALVLVAFMILVAAVLN

Sequence 2: AVTSQLLRRVIMAFFGGILVMALLAMFAMVLLAFGMVFWKRRDNSQPR

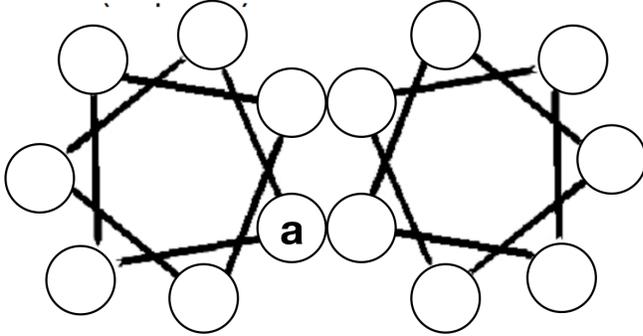
Sequence 3: AVLPGDRLLVLFMAVAVLLGLLGAMIFAAGLSQDLLATPNILVDKSRTKQ

BLANK SPACE (ANY WRITING HERE WILL NOT BE GRADED)

QUESTION 3 (20 points)

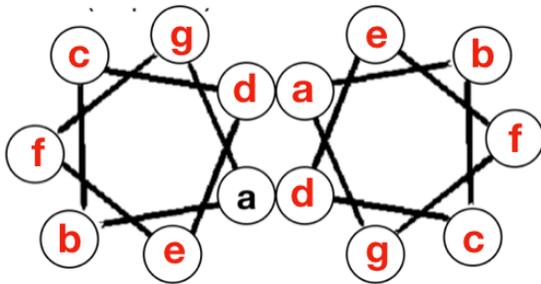
Q3A (10 points)

Shown below are schematic diagrams of two alpha helices forming a parallel coiled coil. The positions of residues within the coiled coil are indicated by *a*, *b*, *c*, *d*, *e*, *f* and *g*.



- (i) (2 points) Label each of the white circles with the appropriate position indicator (i.e., *a*, *b*, *c* ...). The *a* position of one residue is shown.

Answer:



- (ii) (2 points) Are the N-terminal ends of the helices above the plane of the page or below? Circle the appropriate choice below, and explain your answer.

- (a) N-terminus is below the plane (b) **N-terminus is above the plane.**

Explanation:

Since the helices are right-handed, the proper rotation of the helices puts the *a* residue above the *b* residue and so on.

(iii) (4 points) Consider two small proteins, A and B, of 14 residues each that contain only the following four amino acids: L, N, D and R.

Design sequences for A and B so that the proteins obey the following rules:

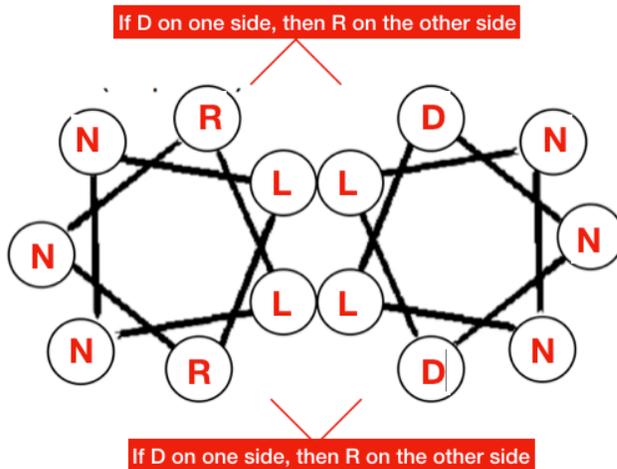
1. A and B form parallel coiled-coils. Ignore the possibility of forming antiparallel coiled-coils.
2. A-A and B-B coiled-coils do not form.
3. Each helix contains only two each of D and R **and four leucine residues**.

Write down your designed sequence in the table below. The parallel coils should be in register exactly as shown in the previous diagram. (There is blank space below, with the schematic diagram, to try out your answer.)

	a	b	c	d	e	f	g	a	b	c	d	e	f	g
A	L	N	N	L	D	N	D	L	N	N	L	R	N	R
B	L	N	N	L	R	N	R	L	N	N	L	D	N	R

(iii) (2 points) Explain briefly why your designed sequences prevent homo-dimerization (i.e., A-A or B-B dimers).

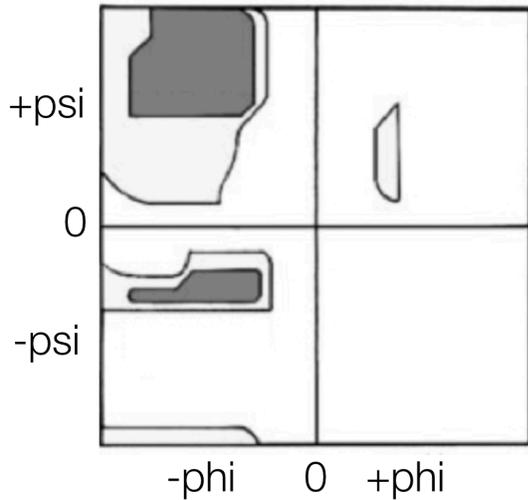
Answer: By putting only D in the first turn and R in the second turn of one helix, at the e and g position, and the opposite for the other helix, we ensure that these helices cannot form stable interactions with each other due to electrostatic repulsion.



 BLANK SPACE (ANY WRITING HERE WILL NOT BE GRADED)

Q3B (6 points)

Shown below is a Ramachandran Diagram.



(i) (2 points) Shown below are the values of ψ (psi) for a 25 residue protein. If the value of ψ for a particular amino acid residue is in the dark gray region in the lower left quadrant, the box corresponding to that residue is marked "-". Likewise, if the value of ψ for a particular residue is in the dark gray region in the upper left quadrant, the corresponding box is marked "+". In the empty boxes in the second row, write α or β , to indicate which conformation the residue is in.

ψ	+	+	+	+	+	+	+	-	+	+	-	+	-	+	+	+	+	+	+	+	+	+	+	+
α or β	α	β	β	α	β	α	β																	

(ii) (2 points) Is this protein likely to be an integral membrane protein? Explain your answer.

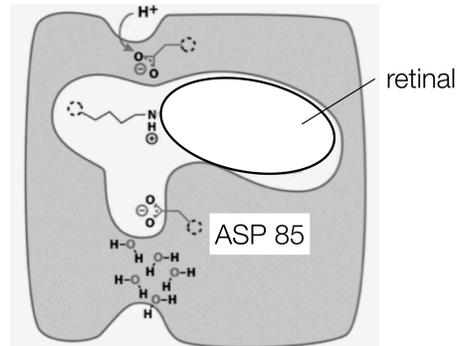
Answer: It is unlikely to be an integral membrane protein because the beta strands are interrupted by a loop (alpha and beta conformation mixed). The loop cannot be inside the membrane. Technically, the C-terminal 12 residues are in a beta conformation, but if it formed a beta-barrel the negatively charged C-terminus would barely extend out of the membrane.

(iii) (2 points) The one residue that does not obey the Ramachandran diagram is glycine. Explain why the Ramachandran diagram for glycine is different.

Answer: The Ramachandran diagram is based on van der Waals repulsion for an alanine sidechain. Since glycine has no sidechain (only a hydrogen), some regions that are disallowed for alanine will be allowed for glycine.

Q3C (4 points)

Shown below is a schematic diagram of bacteriorhodopsin.



(i) (2 points) A scientist modifies the retinal chromophore in bacteriorhodopsin by replacing all of the double bonds in retinal by single bonds, except for the one required for Schiff base formation. The modified chromophore binds to the protein, and forms a Schiff base, but it no longer acts as a proton pump. Explain why.

Answer: Such a molecule could no longer absorb light and undergo the cis-trans isomerization that is essential to drive proton transport in one direction.

(ii) (2 points) The scientist mutates the residue marked ASP 85 in bacteriorhodopsin to GLN. The mutated protein no longer works properly. Provide an explanation for this observation.

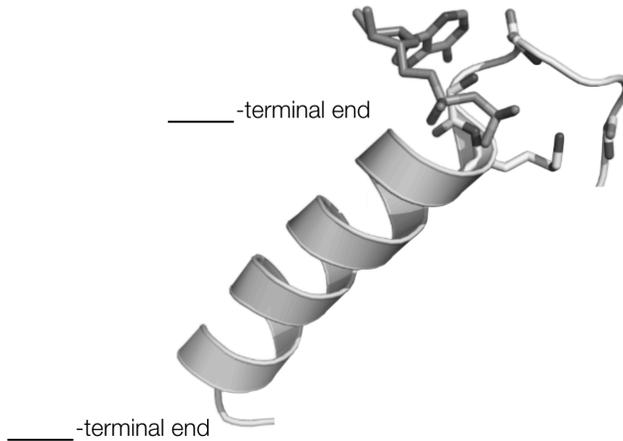
Answer: Asp 85 needs to accept and release protons in order for bacteriorhodopsin to function as a pump. Glutamine cannot titrate in this way, and so proton transport will be blocked.

QUESTION 4 (20 points)

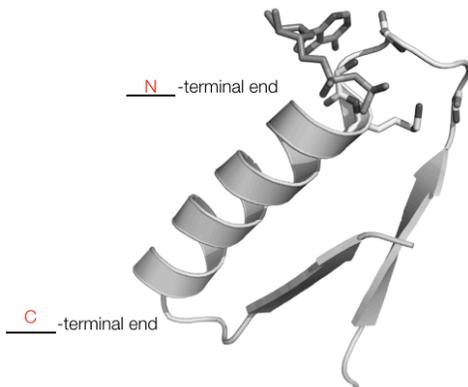
Q4A. (4 points)

Proteins bind to nucleotide phosphates by using an α -helix to stabilize the phosphate groups. In the diagram below, mark the N-terminal and C-terminal ends of the α -helix by writing N and C in the appropriate blanks. (2 points)

Explain the choice you made.



Answer:

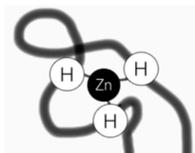


Explanation: (2 points)

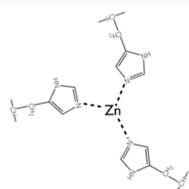
The helix has a dipole, with the positively charged end at the N-terminus of the helix. That is the end that will have favorable electrostatic interactions with the negatively charged phosphate groups.

Q4B (6 points)

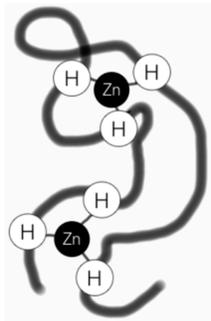
A protein contains 3 histidine residues, as shown schematically below. The details of how the histidine residues coordinate zinc are shown on the right.



Protein A



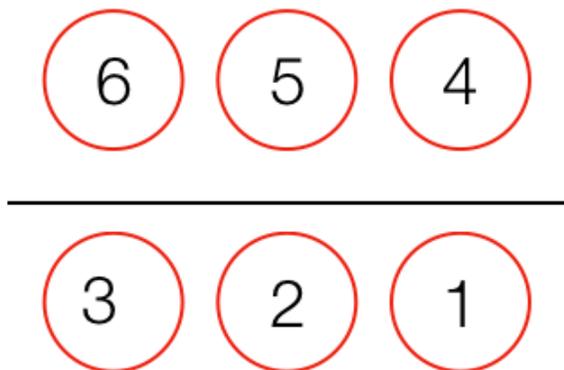
(i) When zinc is added in the presence of a denaturing agent, only one configuration of the protein is possible, corresponding to all three His coordinating the zinc. Consider, instead, a protein with two zincs bound, and six histidine residues, as shown below for one configuration:



Protein B

When zinc is added in the presence of a denaturing agent, how many distinct configurations are there of the protein that satisfy the 3-His-per-1-Zn rule? Assume that each histidine can coordinate either zinc atom effectively, and that both zinc atoms are coordinated by three histidine residues each.

Answer: There are 10 distinct configurations.



Ways of choosing 6 His = 6!

Correct for horizontal rearrangements: $\frac{6!}{3! \times 3!} = 20$

Correct for vertical rearrangements: $\frac{20}{2} = 10$

Q4C (10 points)

Shown below is a portion of the BLOSUM amino acid substitution matrix, containing the values of $S_{i,j}$, where S is the BLOSUM substitution score, and i and j are any two amino acids.

	P	Q	L	S	W
P	7	-1	-3	-1	-4
Q	-1	5	-2	0	-3
L	-3	-2	4	-2	-2
S	-1	0	-2	4	-3
W	-4	-3	-2	-3	11

- (i) (2 points) Explain why the scores for P-P and W-W substitutions are higher than for other self-substitutions.

Answer: Proline and Tryptophan are particularly special because no other amino acid has their properties. Hence, in related proteins, they tend to be conserved (not substituted by other amino acids).

- (ii) (2 points) The BLOSUM matrix shown above is calculated using the sequences of water-soluble proteins. Consider, instead, a BLOSUM matrix calculated using sequences of membrane proteins. In this matrix, will the value of $S_{i,j}$ for L-Q substitutions decrease (become more negative), stay the same, or increase (become more positive)? Explain your answer.

Answer:

The value of S_{LQ} will be smaller than for water soluble proteins. This is because leucine will be even more rarely replaced by glutamine (a polar residue) in membrane proteins.

Q4C, continued

iii) (6 points) Consider the three sequences shown below:

Sequence 1	Q	L	L	P
Sequence 2	Q	L	S	P
Sequence 3	Q	L	W	P

Based on the BLOSUM scores (also shown below), how much more likely is it that sequences 1 and 2 are evolutionarily related, compared to sequences 1 and 3?

That is, calculate the following ratio: $\frac{\text{likelihood that sequences 1 and 2 are related}}{\text{likelihood that sequences 1 and 3 are related}}$

	P	Q	L	S	W
P	7	-1	-3	-1	-4
Q	-1	5	-2	0	-3
L	-3	-2	4	-2	-2
S	-1	0	-2	4	-3
W	-4	-3	-2	-3	11

Answer:

There is only one position in which the sequences are different, so we need only consider the L,S and L,W substitutions. Note that these substitution scores are the same (-2), and so without doing any calculations we can conclude that the likelihood for the two relationships is the same. That is, the ratio = 1.0

QUESTION 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 following statements describe possible roles for hydrogen bonding in protein folding. Circle the *best* statement.
- (a) Hydrogen bonding drives the formation of α helices and β sheets.
 - (b) The formation of secondary structure elements satisfies backbone hydrogen bonding when the peptide chain is inaccessible to water**
 - (c) van der Waals attractions are stronger than hydrogen bonding in holding the structure together.
- (ii) The van der Waals contact distance between two nitrogen atoms is: (circle the best option)
- (a) 2.5 Å
 - (b) 3.5 Å**
 - (c) 4.5 Å
- (iii) Base pairs in RNA do not always obey the Watson-Crick rules because RNA has no requirement for the formation of a uniform double helix. **TRUE** / FALSE
- (iv) Carbon is less electronegative than hydrogen. TRUE / **FALSE**
- (v) The persistence length of DNA, which is the length over which the double helix is fairly rigid, is: (circle the best answer)
- (a) 50 base pairs
 - (b) 150 base pairs**
 - (c) 200 base pairs
 - (d) 250 base pairs
- (v) The lack of change in the globin fold over evolutionary time is due to which factor:
- (a) The maintenance of hydrophobic packing
 - (b) The necessity to function properly
 - (c) The limited ways in which helical ridges can pack into helical grooves**
 - (d) The high level of sequence identity between globin proteins
- (vi) Choose the amino acid substitution that results in the least change in hydrophobicity:
- (a) A \rightarrow N
 - (b) F \rightarrow Q
 - (c) M \rightarrow F**
 - (d) D \rightarrow H

(vii) A protein sequence with alternating hydrophilic and hydrophobic residues are likely to be found in what kind of secondary structure? Circle the best answer.

- (a) α helices
- (b) β strands**
- (c) Loops
- (d) left-handed helices

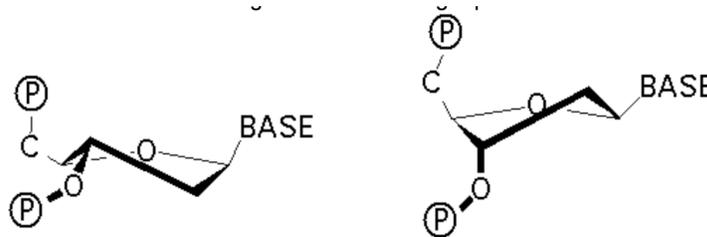
(viii) The major groove of DNA is the preferred site of interaction with proteins that recognize specific sequences because (circle the best answer):

- (a) α helices cannot enter the minor groove of DNA
- (b) Each base pair has a distinct pattern of interactions in the major groove**
- (c) Proteins can interact with the phosphate groups, which are more accessible in the major groove.

(ix) Mg^{2+} plays an important role in stabilizing DNA and RNA structure because (circle the best answer):

- (a) The hexacoordinate geometry of Mg^{2+} allows multiple direct or water-mediated interactions with phosphate groups
- (b) Since it has charge of 2+, it is effective at screening electrostatic repulsion.
- (c) All of the above.**

(x) Circle the correct representation of the 2'-endo sugar pucker in DNA.



Answer: The diagram on the right.

(xi) Which of the following statements is valid for a double helix in which one strand is RNA and one strand is DNA? Circle the best answer.

- (a) It will be in the A-form**
- (b) It will be in the B-form
- (c) It will be unstable

(xii) For proteins made of D-amino acids, which of the following is true? Circle the best answer.

- (a) The protein will form α helices, but not coiled-coils.
- (b) The protein will be unstable.
- (c) The protein will form left-handed coiled-coils.
- (d) The protein will form right-handed coiled-coils.**

(xiii) DNA is more suited than RNA to storing genetic information because the phosphate backbone is less likely to be cleaved. **TRUE** / FALSE