# Reference Solution: Problem Set #2

Structure and thermodynamics of protein-DNA complexes

## Question 1: Specific and Unspecific Protein-DNA Interactions

Specific and unspecific interactions between proteins and the DNA double helix are mediated between various types of amino acid side chains with the DNA double helix.

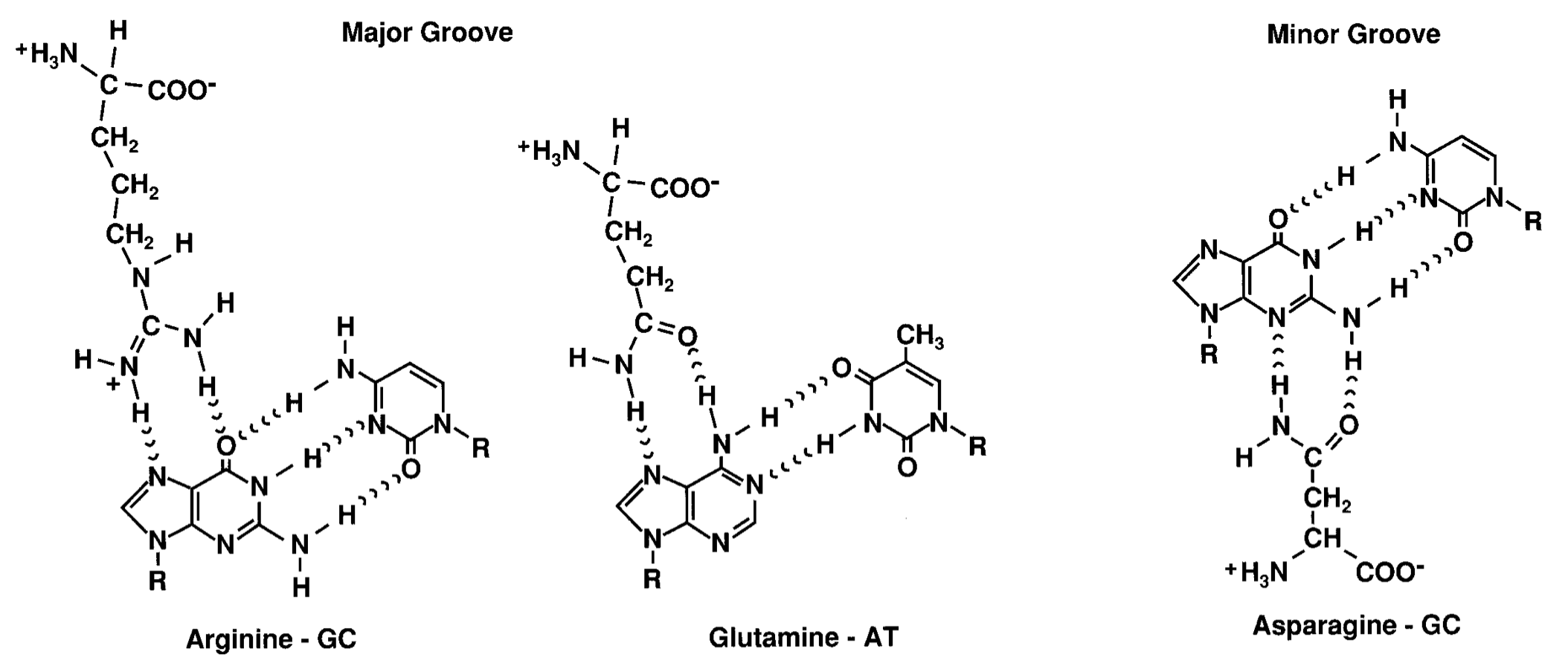
### a) Specific interactions between amino acid side chains and DNA bases

Give an example of specific interactions between an amino acid side chain and a guanine residue in the DNA and another example of a specific interaction with adenine. What type of protein secondary structure will facilitate these interactions, and where on the DNA helix do they occur?

**Guanine-specific interaction:**  
Arginine forms specific hydrogen bonds with guanine in the major groove. The guanidinium group of arginine side chain interacts with the O6 and N7 atoms of guanine through hydrogen bonding. This interaction is stabilized by the positive charge of arginine's side chain.

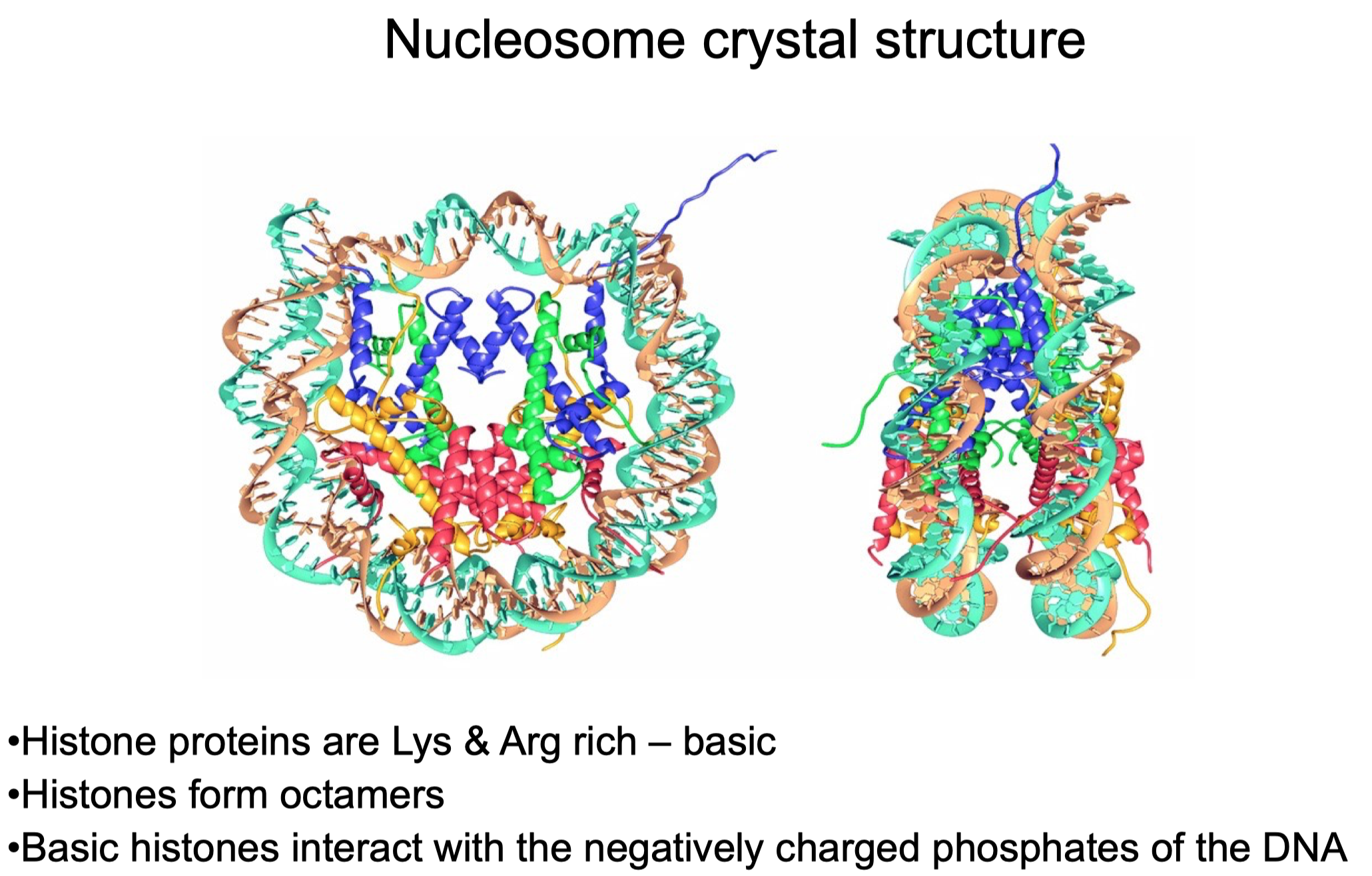
**Adenine-specific interaction:**  
The asparagine or glutamine side chains can form specific hydrogen bonds with adenine in the major groove. The amide group of asparagine/glutamine forms hydrogen bonds with the N6 and N7 atoms of adenine.

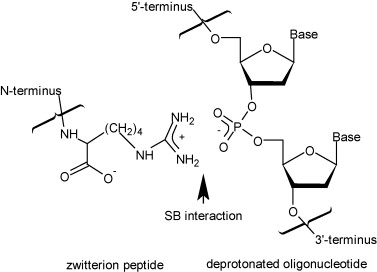
**Facilitating protein structure:**  
These interactions are primarily facilitated by α-helices, particularly in motifs like helix-turn-helix (HTH). The α-helix diameter (1.2 nm) matches the width of the DNA major groove (1.2 nm), enabling precise positioning of amino acid side chains for base-specific recognition. These interactions occur predominantly in the major groove where base-specific hydrogen bond donors and acceptors are more accessible compared to the minor groove.



### b) Unspecific backbone interactions

List the amino acids frequently found to mediate unspecific binding of the protein to the sugar-phosphate backbone of the DNA. Which type of non-covalent interaction forces are involved?





**Key amino acids:**

* Positively charged residues: Lysine, Arginine (see image above), Histidine
* Polar residues: Serine, Threonine, Asparagine, Glutamine

**Non-covalent interactions involved:**

1. Electrostatic interactions between positively charged amino acid side chains and negatively charged phosphate groups
2. Hydrogen bonds between polar residues and phosphate oxygens
3. Van der Waals interactions

### c) DNA distortion and biological function

What could be differences in the biological functions of DNA-binding proteins that make some proteins strongly distort the DNA upon specific binding while others hardly affect the DNA conformation?

**Proteins that strongly distort DNA:**

* Transcription factors with an architectural function (e.g., CAP (catabolite activator protein) or TBP (TATA-box binding protein) or the histone octamer in the nucleosome that has the DNA wrapped around it
* Function requires structural changes for:
  + Creating a platform to assembe the transcription machinery (TBP)
  + Bringing distant regulatory elements into proximity by inducing DNA curvature (CAP)
  + Exposing binding sites for other proteins
  + Initiating specific biochemical processes
  + Compacting and organizing the genome (histone octamer)

**Proteins with minimal DNA distortion:**

* Sequence-specific binding of transcription factors like GCN4
* Purpose: Maintain DNA structural integrity while performing functions like:
  + Transcription activation through recruitment of other factors
  + Sequence-specific recognition and binding
  + Gene regulation without requiring DNA conformational changes

## Question 2: Nucleosome Structure Analysis

Examine the nucleosome crystal structure for which the pdb coordinates are given in the file “nucleosome.pdb” with a molecular viewer (see below for different viewers).

### a) Histone-DNA interactions

Evaluate interactions of the histone octamer protein core with the DNA. Do they occur mostly with DNA bases or with the sugar-phosphate backbone of the DNA? At which periodicity (expressed in the number of base pairs) are interactions between DNA and protein core present?

**Interaction type:**  
Predominantly with the sugar-phosphate backbone through:

* Electrostatic interactions between positively charged histone residues (Lys, Arg) and negatively charged phosphate groups
* Hydrogen bonding networks

**Periodicity:**

* Interactions occur every ~10 base pairs when the minor groove faces the histone surface
* More precise analysis shows a 5 bp periodicity due to alternating strand interactions

### b) DNA fragment characteristics

How long is the DNA fragment wrapped around the histone octamer protein core? Give the sequence of the 5 base pairs where the interaction starts and the 5 base pairs where it ends.

**Length:** 147 base pairs wrapped around the histone octamer in 1.7 turns

**Sequence boundaries:**

* Starting sequence: GCCCT
* Ending sequence: CATCC

### c) Unfolded regions

Does the histone octamer contain protein regions that are unfolded, and if so, where are they located?

The histone octamer contains unstructured N-terminal domains ("histone tails") that:

* Are present in all core histones (H2A, H2B, H3, and H4)
* Protrude from the globular core structure
* Serve as substrates for post-translational modifications (acetylation, methylation, phosphorylation)

## Question 3: Protein-DNA Binding Thermodynamics

The free energy ∆G of protein binding to DNA involves favorable and unfavorable entropy terms.

### a) Entropy contributions

Describe three different contributions to the entropy change that occurs if a protein binds to DNA.

1. **Hydrophobic effect (ΔS\_HE):**
   * Release of ordered water molecules from protein and DNA surfaces
   * Increases system entropy
   * Proportional to buried nonpolar surface area
   * Quantified as ΔS\_HE(T\_S) = 0.32ΔA\_np ln(T/386)
2. **Conformational entropy (ΔS\_conf):**
   * Reduction in protein and DNA flexibility
   * Decreases system entropy
   * Results from structural ordering upon binding
3. **Rigid body association entropy (ΔS\_rt):**
   * Loss of translational and rotational freedom
   * Decreases system entropy
   * Approximately -50 cal mol⁻¹ K⁻¹

### b) Impact on binding equilibrium

Explain which of these entropic changes drive binding and which would favor dissociation of the complex.

**Driving binding (favorable):**

* Hydrophobic effect (ΔS\_HE > 0)
* Polyelectrolyte effect (ion displacement) (ΔS\_PE > 0)

**Favoring dissociation (unfavorable):**

* Conformational entropy loss (ΔS\_conf < 0)
* Rigid body association entropy loss (ΔS\_rt < 0)

### c) Specific vs. unspecific binding

Which entropy term could significantly differ between a specific and unspecific protein-DNA complex and favor specific binding?

The hydrophobic effect (ΔS\_HE) shows the most significant difference between specific and unspecific binding:

**Specific binding:**

* Greater surface complementarity
* More extensive water displacement
* Larger favorable entropy contribution
* Compensates for larger conformational entropy loss

**Unspecific binding:**

* Less precise surface matching
* Fewer displaced water molecules
* Smaller entropic contribution
* Maintains more conformational flexibility