



StarBiochem

DNA Glycosylase Exercise - Levels 1 & 2

Background

In this exercise, you will explore the structure of a DNA repair protein found in most species, including bacteria. DNA repair proteins move along DNA strands, checking for mistakes or damage. DNA glycosylases, a specific type of DNA repair protein, recognize DNA bases that have been chemically altered and remove them, leaving a site in the DNA without a base. Other proteins then come along to fill in the missing DNA base.

Learning objectives

We will explore the relationship between a protein's structure and its function in a human DNA glycosylase called human 8-oxoguanine glycosylase (hOGG1).

Getting started

We will begin this exercise by exploring the structure of hOGG1 using a molecular 3-D viewer called StarBiochem. In this particular structure, the repair protein is bound to a segment of DNA that has been damaged. We will first focus on the structure hOGG1 and then on how this protein interacts with DNA to repair a damaged DNA base.

- To begin using StarBiochem, please navigate to: <http://mit.edu/star/biochem/>.
- Click on the **Start** button Click on the **Start** button for **StarBiochem**.
- Click **Trust** when a prompt appears asking if you trust the certificate.
- In the top menu, click on **Samples** → **Select from Samples**. Within the **Amino Acid/Proteins** → **Protein** tab, select “**DNA glycosylase hOGG1 w/ DNA – H. sapiens (1EBM)**”. “1EBM” is the four character unique ID for this structure.

Take a moment to look at the structure from various angles by rotating and zooming on the structure.

- Instructions for changing the view of a structure can be found in the top menu, under **Help** -> **Structure viewing instructions**.

The current way you are viewing the structure is by seeing each atom and bond in the protein drawn as a ball and a line, respectively. This way of representing a structure is called the **ball-and-stick** model and is the default model in StarBiochem. The **ball-and-stick** model allows you to see how atoms in the structure bond together. However, the space each atom occupies IS NOT accurately represented. To see a more realistic representation of the atoms in the structure you can use the **space-filled** model, where each atom is drawn as a sphere, whose size represents the physical space an atom occupies.

You can switch from the **ball-and-stick** model to the **space-filled** model in StarBiochem by increasing the size of the atoms in the structure:

- Notice that different atoms are slightly different in size.
- Gray = Carbon, Blue = Nitrogen, Red = Oxygen, Yellow = Sulfur
- Click on the **Primary** tab. The default atom size is 20% (**ball-and-stick** model).
- Move the **Atoms Size** slider to 100% (**space-filled** model).

The last page, Reference page, contains a series of terms and useful information that you will refer to during this exercise.

Exercise

Part 1 - hOGG1's Protein Structure Basics

1 The hOGG1 structure contains both DNA and protein. Can you differentiate between the DNA and protein components? How did you distinguish the DNA from the protein?

Answer

2 hOGG1 contains multiple sulfur atoms.

a) Identify the name and sequence number of one of the amino acids in the structure that contains a sulfur atom.

- Using your mouse, point to a sulfur atom (yellow) in the structure. A small box appears on top of the mouse cursor indicating the name of the amino acid and its position in the amino acid sequence (ex: “[Ala]223:A CB # 2257” → amino acid: alanine; position: 223).

Answer

b) Is the sulfur atom located in the backbone or in the side chain of the amino acid?

- From the top menu select **View** → **View Specific Regions / Set Center of Rotation**. This will open a smaller window, which enables you to set specific region(s) of the structure visible and centered in the viewer. Follow the steps below to view only the amino acid with the sulfur atom identified above.
- In the **Protein** → **Primary** tab of the **View Specific Regions / Set Center of Rotation** window, select the amino acid identified above in the **Sequence Window**.
- Move the **VDW Radius** slider all the way to the left (1 Van der Waals radii). Zoom in as needed.
- First uncheck the **side-chain** box to hide the side-chain atoms and view only the atoms in the backbone. Next uncheck the **backbone** box and check the **side-chain** box to view only the atoms in the side-chain.

Answer

3 Next, we will explore the primary structure of the hOGG1 protein (Reference page – last page within the document). The hOGG1 protein consists of 325 amino acids. List the 13 amino acids numbered 105 through 117 in order.

- Reset the structure by clicking **Reset** → **Reset structure** in the top menu.
- In the **Protein** → **Primary** tab within the main StarBiochem window, scroll down in the **Sequence Window** to locate amino acids 105-117 in the sequence. Refer to the Reference page for the complete name of each amino acid.

Answer

105 _____	106 _____	107 _____	108 _____
109 _____	110 _____	111 _____	112 _____
113 _____	114 _____	115 _____	116 _____
117 _____			

4 Within a protein chain, amino acids form local structures called secondary structures (Reference page).

a) Explore the secondary structures found in hOGG1. Are helices, sheets or coils present in hOGG1? Describe the color that represents the secondary structures you observe.

- Click on the **Secondary** tab.
- To show the secondary structures one at a time, check the box beside the desired structure (ex: helices) and move the **Structures Size** slider to the right to increase the size.
- View additional secondary structures that may be present by checking the boxes next to each structure.

Answer

Helices _____ (Yes or No) Sheets _____ (Yes or No) Coils _____ (Yes or No)

b) Amino acids 105 through 117 fold into one of the secondary structures. Which secondary structure do they fold into?

- In the **View Specific Regions** window, go to the **Protein** → **Secondary** tab and click the **All** button.
- Within the amino acid **Sequence Window**, select amino acids 105-117: click on amino acid 105, hold down **Shift** and click on amino acid 117.
- Move the **VDW Radius** slider to the left to show only the selected amino acids in the viewer.

Answer

5 Now we will explore the relationship between DNA glycosylase's structure and one of the several types of amino acids that contribute to hOGG1's overall shape, its tertiary structure (Reference page).

a) Negatively charged amino acids are hydrophilic (Reference page). Are the negatively charged amino acids located on the inside (buried) or outside (exposed) of this protein? What does that suggest about the cellular environment surrounding this protein, is it hydrophobic or hydrophilic? Explain your answer.

- In the top menu, click **Reset** → **Reset structure**.
- Click on the **Tertiary** tab. Move the **Atoms Size** slider to the left decreasing size for all the atoms in the protein.
- Click on the **negatively charged/acidic** button. Move the **Atoms Size** slider to the right to increase the atom size for the negatively charged/acidic amino acids.

Answer

Part 2 - hOGG1's Interaction with DNA (Basic)

6 Now let's explore how DNA glycosylase interacts with DNA to recognize the damaged DNA base within its sequence. DNA is composed of four bases: Adenine (A), Thymine (T), Cytosine (C) and Guanine (G). In this particular structure, the hOGG1 protein is bound to a segment of DNA that contains an oxidized guanine. First, let's make the DNA segment more visible.

- In the top menu, click **Reset** → **Reset structure**.
- Click on the **Nucleic Acids** tab. Move the **Atoms Size** slider to the right (35%) to increase the size of the DNA segment.

- In the **Protein** → **Primary** tab move the **Atoms Size** slider completely to the left and the **Bonds Translucency** slider to the right (95%) to minimize the appearance of all the amino acids in the hOGG1 protein.

Second, let's take a closer look at how the DNA bases are oriented within the double helix. Each base is attached to a sugar and phosphate backbone to form a complete nucleotide. Within the double helix, a base within one strand pairs with another base from the opposite strand by hydrogen bonding forming a base pair: Adenine (A) base pairs with Thymine (T) and Cytosine (C) base pairs with Guanine (G).

a) How many DNA base pairs can you count within this double helix?

- Click on the **Nucleic Acids** tab.
- Uncheck **phosphates** and **sugars**. Move the **Atoms Size** slider to the right (70%) to increase the size of the bases while leaving the size of the phosphates and sugar nucleotide components intact. This will allow you to count the base pairs more easily.

Answer

b) How many DNA bases are unpaired (not paired to its partner on the other strand)?

Answer

c) Is the oxidized guanine base paired or unpaired? Describe the position of the oxidized guanine with respect to the hOGG1 protein and the double helix. What does this suggest about the mechanism that hOGG1 uses to identify damaged DNA bases?

- Click on the **Non-Peptide** tab. Move the **Atoms Size** slider completely to the right to increase the size of the atoms in the oxidized guanine ([8OG]25).

Answer

7 Certain amino acids within hOGG1 form contacts with the DNA and are able to recognize if a guanine base has been damaged by oxidation. Where are you more likely to find the amino acids that recognize damaged guanine bases within hOGG1, in Helix 1 or Helix 16? Explain why.

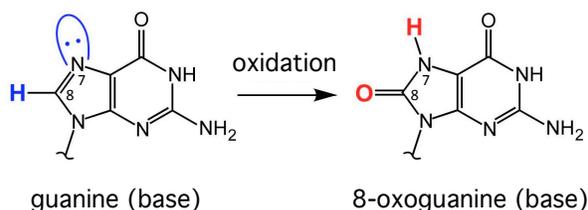
- In the **Protein** → **Secondary** tab, click the **All** button.
- In the **Sequence Window**, select the amino acids within Helix 1: click on the first amino acid of Helix 1, Shift+click on the last amino acid of Helix1. Increase the size of Helix 1 by moving the **Secondary Structures** size slider to the right. Note the location of Helix 1 with respect to the oxidize guanine.
- Now in the **Sequence Window**, select the amino acids within Helix 16 as you did for Helix 1. Note the location of Helix 16 with respect to the oxidized guanine.
- *Optional:* This set of steps will allow you to see if any of the side chains within Helix 1 or Helix 16 contact the oxidized guanine ([8OG]25). Click on the **Protein** → **Primary** tab. Select the amino acids that make up Helix 1 in the **Sequence Window**. Uncheck **backbone**. Move the **Atoms Size** slider to

the right to visualize the side chains of the amino acids Helix 1. Now select the amino acids found within Helix 16.

Answer

Part 3 - hOGG1's Interaction with DNA (Advanced)

8 hOGG1 is able to distinguish very efficiently between guanine and its oxidized counterpart, 8-oxoguanine. This represents a formidable task given that the oxidized nucleobase, 8-oxoguanine, differs by only two atoms from its normal counterpart, guanine (positions 7 and 8).



The following structures illustrate hOGG1 bound to either an 8-oxoguanine or a guanine nucleobase: “1YQR” (8-oxoguanine) and “1YQK” (guanine). We will compare these two structures to understand how hOGG1 can effectively distinguish between 8-oxoguanine and guanine.

- In the top menu, click on **Samples** → **Select from Samples**. Within the **Amino Acid/Proteins** → **Protein** tab, select “DNA glycosylase hOGG1 w/ DNA containing oxoG - H. sapiens (1YQR)”. “1YQR” is the four character ID for this particular structure.
- Repeat these steps to open the “DNA glycosylase hOGG1 w/ DNA containing G - H. sapiens (1YQK)” structure.

Glycine 42 in hOGG1 directly interacts with the base portion of 8-oxoguanine. This interaction is crucial for detection of oxidative damage of guanine bases. Examine the 8-oxoguanine and glycine 42 interaction in the “1YQR” structure.

a) Which atom in the base portion of 8-oxoguanine is directly contacting glycine 42? Draw this interaction (use the chemical structures within this question and the amino acids structures found in the Reference page).

- In the **Protein** → **Primary** tab move the **Atoms Size** slider completely to the left and the **Bonds Translucency** slider to completely to the right to minimize the appearance of all the amino acids in the hOGG1 protein.
- Click on the **Nucleic Acids** tab. Click the oxidized guanine ([8OG]23:C) and increase its size by moving the **Atoms Size** slider completely to the right.
- In the **Protein** → **Primary** tab, click glycine 42 and increase its size by moving the **Atoms Size** slider completely to the right.

Answer

b) What does the interaction between 8-oxoguanine and glycine 42 indicate about how hOGG1 differentiates between 8-oxoguanine and guanine?

Answer

c) Compare the general location and orientation of guanine and glycine 42 in the “1YQK” structure with that of 8-oxoguanine and glycine 42 in the “1YQR” structure. How do the location and orientation of these two nucleobases and glycine 42 differ in these structures? Explain how this comparison adds to your understanding of hOGG1’s ability to discriminate between 8-oxoguanine and guanine.

- Reset the “1YQR” structure by clicking on **Reset** → **Reset structure** in the top menu.
- In the **Protein** → **Primary** tab move the **Atoms Size** slider completely to the left and the **Bonds Translucency** slider to completely to the right (95%) to minimize the appearance of all the amino acids in the hOGG1 protein.
- Click on the **Nucleic Acids** tab. Click the oxidized guanine ([8OG]23:C) and increase its size by moving the **Atoms Size** slider completely to the right.
- In the **Protein** → **Primary** tab, click glycine 42 and increase its size by moving the **Atoms Size** slider completely to the right.
- Repeat these steps with the “1YQK” structure, but instead of increasing the size of the oxidized guanine, increase the size of guanine at position 23.

Answer

9 One intriguing question about hOGG1 and other DNA glycosylases is whether these enzymes search for DNA oxidation by extruding each base from DNA and presenting it for inspection to the active site or by recognizing DNA oxidation sites without extruding each base. It is known that the base pairing between the 8-oxoguanine nucleobase and cytosine (oxoG:C) does not result in a conformational change in the DNA double helix. Therefore, it seems unlikely that DNA glycosylases would be able to correctly detect the presence of an 8-oxoguanine nucleobase without extruding each base. Yet, DNA glycosylases need to survey approximately 6×10^9 base pairs in the human genome, making the extrusion of each base energetically prohibitive and extremely time-consuming.

Crystal structure evidence of a bacterial homolog of hOGG1, called MutM, suggests that a specific amino acid, phenylalanine 114, in MutM intercalates into the DNA helix at each nucleotide position as DNA glycosylase travels along the DNA - if you are curious, this structure is found within **Samples** and it is called "**DNA glycosylase MutM w/ DNA - G. stearothermophilus (2F5O)**". It has been hypothesized that phenylalanine 114 acts as a sensor to detect the presence of 8-oxoguanine nucleobases. To explain how intercalation of phenylalanine at the site of an oxoG:C would disrupt base pairing and promote extrusion of the 8-oxoguanine nucleobase, but not of A:T and G:C base pairs, it has been suggested that intercalation of phenylalanine 114 at the site of an oxoG:C base pair may sufficiently destabilize base pairing because oxoG:C base pairs are less stable than A:T and G:C pairs.

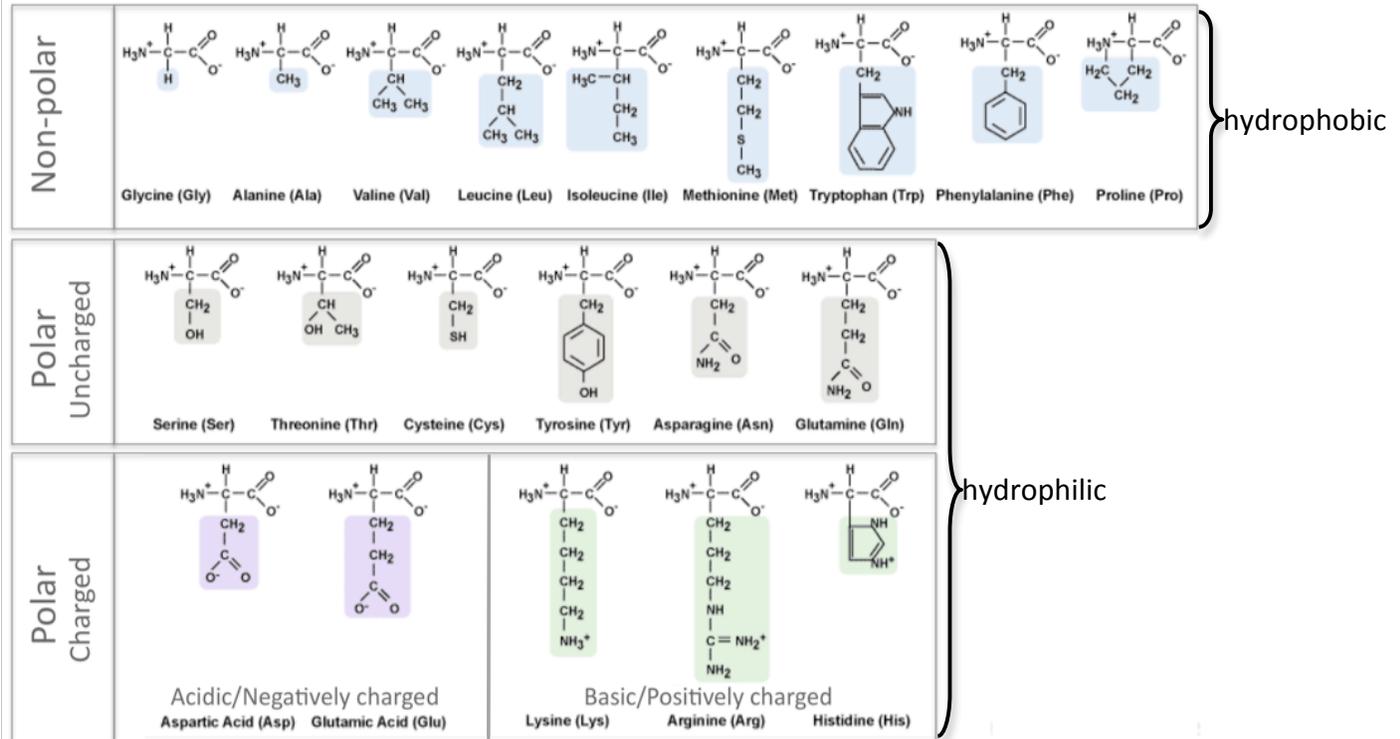
Propose an experiment to test whether phenylalanine 114 plays a role in the detection and repair of 8-oxoguanine nucleobases.

Answer

Reference

CHEMICAL STRUCTURES OF THE AMINO ACIDS

The 20 amino acids share a common backbone and are distinguished by different side chains, also called 'R' groups, highlighted by the various colors below.



PROTEIN STRUCTURE BASICS

All proteins have the following three levels of protein structure:

Primary structure

Describes the order of the amino acids in the protein chain but does not describe its shape.

Secondary structure

Describes shapes that form from local folding of regions within the amino acid chain. These smaller structures can be divided into two main types: helices and sheets. Coils are made of amino acids that do not form regular secondary structures (helices and sheets) but play important roles in protein folding.

Tertiary structure

Describes the entire folded shape of a protein chain.

In addition, some proteins interact with themselves or with other proteins to form larger protein structures. These proteins have an additional level of protein structure:

Quaternary structure

Describes how multiple protein chains interact and fold to form a larger protein complex.