

## Biochemistry—Proteins

### A. CBS proteins, 3D representation

The protein we will consider today and periodically throughout the term is the yeast cystathionine beta synthase (CBS) protein. Today we will see gels visualizing wild-type and mutants versions of the yeast CBS protein. Yeast that lack CBS protein can not synthesize the amino acid cysteine. Mutant versions of the human CBS protein can lead to a very serious disease called homocystinuria. Some symptoms include mental retardation, early strokes and heart disease.

First consider a crystallographic model of a form of CBS protein. Open the CBS structure html file in Internet Explorer. Rotate the molecule and get an idea of the three dimensional structure of CBS.

The true 3-dimensional structure of CBS can be observed using the “**Show Spacefill**” button, and a trace of the peptide backbone can be observed using the “**Show Ribbon**” button.

1. What is the primary structure of a protein? What information does the modeling program provide us about the primary structure of CBS?

*The primary structure of a protein is the linear sequence of amino acids. It is impossible to determine this for CBS with the representation that has been provided. We could look up the sequence in the Protein Data Bank and it would read something like (...Trp-Ile-Arg-Pro-Asp-Ala-Pro-Ser...) through all 430 amino acids.*

2. What is the secondary structure of a protein? What information does the modeling program provide us about the secondary structure of CBS?

*The secondary structure is described as  $\alpha$ -helical or  $\beta$ -sheet. Remember that secondary structure is controlled by backbone hydrogen bonds between amino acids. CBS contains both types of secondary structure. The majority of the secondary structure elements are  $\alpha$ -helices. There are 4  $\beta$ -sheets.*

3. What is the tertiary structure of a protein? What information does the modeling program provide us about the tertiary structure of CBS?

*The tertiary structure is determined by the side-chain interactions between amino acids in adjacent secondary structures. Although it is hard to resolve the desired interactions from this view, it is clear that amino acids in both the  $\alpha$ -helices and  $\beta$ -sheets are interacting with amino acids in other secondary structure elements to contribute to the overall tertiary structure.*

4. What is the quaternary structure of a protein? What information does the modeling program provide us about the quaternary structure of CBS? (Hint: Use the “Color Ribbon by Protein Subunit” button)

*The quaternary structure of a protein is created by a number of distinct interacting amino acid chains. If we color CBS by chain we find that there are two amino acid strands indicating that the protein is a homodimer.*

## B. Protein gels

1. Yeast cells produce enough CBS to fulfill their needs depending on the conditions in which they are growing. In order to procure large quantities of the enzyme, we created special bacterial cells that mostly produce CBS. After we grew up a large number of these bacterial cells, what did we have to do to end up with pure CBS for loading on the gel?

*We had to grow up a lot of cells, break them all open, and then go through a gradual process of removing all parts of the cell we didn't need. We removed membranes, nucleic acids, carbohydrates, and small molecules. We then removed the majority of other proteins that were in the cell at the time of harvest, leaving us with mostly pure CBS.*

Some technical issues:

2. What is a gel?  
*A gel is a jelly-like matrix. Bigger objects have more trouble fitting through the matrix, and, therefore, travel slower through the gel.*
3. How does gel electrophoresis work?  
*Our sample preparations and the buffer used in the gel set-up ensure that unfolded proteins have negative charge overall. When the electric current is applied, they move toward the anode.*
4. If you run two proteins on the same gel, and one travels farther than the other, what can you say about the relative sizes of these proteins?  
*The one that travels farther has lower molecular weight.*
5. With gel electrophoresis, two different conditions are often used.
  - a. Denaturing gel: detergent is added to the gel and a reducing agent is added to the protein sample. Reducing agent converts the S-S disulfide bonds to SH and SH.
  - b. Native gel: Physiological pH conditions where no detergent or reducing agent is added.

Compare the 1°, 2°, 3°, and 4° structure of CBS for each of these conditions:

	Denaturing gel	Native gel
1°	<i>Primary structure is not disrupted. Same as that for CBS proteins present in a cell and for those run on native gels.</i>	<i>Primary structure is not disrupted. Same as that for CBS proteins present in a cell and for those run on denaturing gels.</i>
2°	<i>Secondary structure lost due to presence of detergent and reducing agent.</i>	<i>Secondary structure is not disrupted. Same as that for CBS proteins present in a cell.</i>
3°	<i>Tertiary structure lost due to presence of detergent and reducing agent.</i>	<i>Tertiary structure is not disrupted. Same as that for CBS proteins present in a cell.</i>
4°	<i>Unable to form quaternary structures due to presence of detergent and reducing agent.</i>	<i>Quaternary structure is not disrupted. Same as that for CBS proteins present in a cell.</i>

6. Information about what level(s) of protein structure do you expect to gain from a denaturing gel?

*Adding detergent and a reducing agent ensures that the protein will lose its secondary and tertiary structure, so the information we get from running such a denaturing gel is the size of each subunit of the protein complex.*

7. Information about what level(s) of protein structure do you expect to gain from a native gel?  
*Not adding detergent and a reducing agent allows the protein to enter the gel as a 3D structure, so we would be able to gain information about its apparent 3D size.*
8. What could the native gel tell you that the denaturing could not?  
*The native gel would provide information on the number of subunits involved in the protein complex.*

Let's now talk about the proteins themselves.

In the 3D representation of the CBS protein, consider a single subunit of CBS. Now consider the whole CBS protein again.

9. Which of these forms would we be visualizing with each kind of gel?  
*A native gel would give us an opportunity to visualize the whole CBS protein, while the denaturing gel would tell us the size of the individual subunit. These two pieces of information put together would tell us how many subunits are in each complete CBS protein.*

Imagine we are the researchers who first isolated the wild type and mutant versions of the protein. What kinds of things might we want to know about these proteins?

The five questions from journalism school are very much applicable here: Who? (What is this protein?) What? (What does it do in the cell?) When? (When does it act?) Where? (Where does it act?) How? (How does it perform its function? What does it interact with?). Note that these are all basic questions that can be asked from first principles.

10. Which of these things can we figure out by running the purified proteins on the gels and which ones can't we? Why?  
*We can only get a part of the answer to the last question from the gel. This is because we can not definitively answer questions about the protein's function and mode of operation in the cell without looking in the cell.*
11. What other techniques might you want to employ in your research of the purified protein? What would you hope to find out from these experiments?  
*If we determined by other means that this is an enzyme, we could try to determine what reaction the protein catalyzes by seeing whether any of the known or suspected substrates will be acted upon. We could also determine whether the enzyme needs a cofactor, such as any of the known small molecule co-factors.*

### **C. Analyzing the data and interpreting the results**

Take a look at the denaturing gels. One of our proteins migrates a shorter distance than the other.

1. What does this mean?

*It means that the protein that migrates a shorter distance is longer, i.e. has more amino acids in its primary structure.*

2. As we said earlier, we ran both the wild-type and mutant protein on the denaturing gel. Which do you expect to be longer—wild-type or mutant protein? Why?

*It could be either. A number of different mutations in the DNA sequence coding for a protein can create a shorter protein, by prematurely creating a stop instruction. On the other hand, a mutation that disrupts an existing stop instruction for this protein can result in creating a longer protein.*

*In this particular case, the mutant protein travels farther on the gel, and we, therefore, conclude that the mutant protein is shorter than the wild-type protein.*

Take a look at the photocopy of a native gel.

3. What can you say about the difference between the wild type and the mutant enzymes from the native gels?

*Native gel data indicates that the mutant protein forms dimers, while the wild-type protein forms tetramers and octamers.*

4. Would you expect the same section of the monomer subunit of CBS to be used in forming dimers as in forming tetramers? Why or why not?

*No. The mutant subunit is shorter, so we conclude that it is missing a portion of the amino acid chain with respect to the wild-type protein. The mutant can still form dimers, but not tetramers and octamers. Therefore, we expect different parts of the monomer subunit to be involved in establishing each of these interactions.*

5. Do you think the 3D representation we are working with today is of the wild-type or mutant protein? Why?

*The representation is of the mutant protein because it shows dimer as the quaternary structure of the enzyme, and not a tetramer or an octamer.*

Let's now look at the dimer interface that forms between the two chains. There are several important interactions that hold the two subunits together. We will look at some residues involved in keeping the interface together.

#### I. Choose “**Central Interaction**”

6. What are the residues involved in this interaction?

Hint: It may help to change the representation of the residues to ball and stick by clicking in the box labeled “Ball and Stick on/off” below the 3D window

*Phenylalanine from one chain interact with phenylalanine from the other chain.*

7. What types of amino acids are these residues (acidic, basic, polar, or hydrophobic)?

*Phenylalanine is hydrophobic.*

8. What types of interactions occur between these residues?

*The two phenylalanines interact via a Van Der Waals interaction.*

9. Would the CBS dimer interface be more, less or equally as stable if we changed one of the phenylalanines to a glutamic acid? An alanine? A tryptophan?

*A glutamic acid in this position would probably destabilize the interface because it is very polar and positively charged, while phenylalanine is a hydrophobic bulky ring.*

*An alanine in this position would also probably destabilize the protein, because although it is hydrophobic, it is very small whereas phenylalanine is quite bulky. The alanine would introduce a hole or gap in the dimer interface due to its small size.*

*Introduction of tryptophan in this position would be the residue most likely not to destabilize the dimer interface. It is large, aromatic, and hydrophobic, and therefore, retains a majority of the qualities of phenylalanine.*

## II. Choose “**Interaction I**”

10. We have highlighted two identical clusters at opposite ends of the CBS structure. Each individual cluster contains an alanine and a leucine from chain A and an alanine and a leucine from chain B. What type of amino acids are alanine and leucine?

*Alanine and leucine both have hydrophobic side chains.*

11. What type(s) of interaction(s) are the amino acids in these clusters involved in?

*These residues interact via hydrophobic interactions and van der Waals interactions.*

*Hydrophobic interaction refers to a situation where a number of hydrophobic amino acids are clustered together, unlike a case where only two hydrophobic amino acids are involved in the interaction. In the latter case we say that the strongest interaction is van der Waals (see question 8 above). Notice that when hydrophobic interaction is happening, van der Waals forces are still present, but are not the strongest interaction.*

12. Would you expect the side chains in this cluster to be exposed to the cytoplasm? Why or why not?

*We would not expect the side chains to be exposed to the cytoplasm because the cytoplasm is primarily composed of water and contains many other polar molecules. An interaction between the nonpolar, hydrophobic side chains of alanine and leucine and water would be highly unfavorable. We expect most hydrophobic clusters to be located on the inside of globular proteins.*

## III. Choose “**Interaction II**”

13. We have highlighted two identical clusters at far ends of the CBS structure. Each individual cluster contains a glutamine and an aspartic acid on one chain, and a lysine on the other. What types of amino acids are these residues?

*Glutamine has a polar side chain, aspartic acid has an acidic side chain, and lysine has a basic side chain.*

14. What type(s) of interaction(s) are found in this cluster?

*Glutamine may interact with either lysine or aspartic acid via a hydrogen bond. Lysine and aspartic acid bond together via an ionic bond.*

15. Select “**Interaction II with all atoms shown.**” This view shows all the atoms of CBS (not including hydrogens) as transparent dots. Are the amino acids involved in Interaction III found on the exterior or interior of the protein? Does this make sense if CBS is a soluble protein?

*These amino acids are on the exterior of surface of CBS. It would be energetically favorable for this cluster to be exposed to a hydrophilic solvent like the cytoplasm or water because the amino acid side chains involved in this cluster could hydrogen bond with other polar molecules.*

16. Choose “**Show Surface Hydrophobics.**” This colors all the hydrophobic residues of CBS pink. Are there a lot or a few surface exposed hydrophobic amino acids? Does this make sense? For what purpose could the protein use surface exposed hydrophobic residues?

*There a quite a few hydrophobic residues on the surface of CBS. Most of them are scattered throughout the surface and are not found in distinct clusters. This is typical of most soluble proteins. It allows them to remain stable, and to sometimes interact with other proteins or to bind hydrophobic ligands.*

*CBS is normally found in a tetrameric or octameric conformation. The wild-type CBS protein uses some of the hydrophobic amino acids on its surface, along with the domain missing in the mutant protein, to bind other CBS molecules.*

Back to the gel data.

17. Do you expect the mutant enzyme to be active? Why or why not?

*Unclear from data. It depends on what the function(s) are of the domain of the protein that was lost in the mutant.*

18. The in-gel activity assay actually indicates that the mutant protein is active. What do you think is the function of the missing domain?

*It is possible that the missing domain is only responsible for making tetramers and octamers. It could also have function that we have not analyzed yet.*

19. Suppose a colleague in the lab next door has figured out that the mutant protein is active in the cell all the time, while the wild type protein is active only some of the time. What conclusion might you draw from this information in combination with your data?

*The domain missing in the mutant appears to also have regulatory function that allows the protein to be in either active or inactive conformation.*