Credits

Credits/Funding Source

The Bio-ITEST program was made possible by an Innovative Technology Experiences for Students and Teachers grant award from the National Science Foundation (NSF), DRL-0833779.

Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NSF, or NWABR’s consultants/advisory board members.

Authors and Contributors

Jeanne Ting Chowning, MS
Bio-ITEST Principal Investigator
Director of Education, Northwest Association for Biomedical Research

Dina Kovarik, MS, PhD
Scientific Curriculum Developer, Digital World Biology LLC

Joan Griswold, MIT
Education Outreach Coordinator,
Northwest Association for Biomedical Research

Sandra Porter, PhD
Bio-ITEST Co-Principal Investigator, President, Digital World Biology LLC

Jodie Spitze, NBCT
Science Teacher, Kent-Meridian High School

Carol L. Farris, PhD
Senior Fellow, Biomedical Health Informatics
University of Washington

Karen Peterson, MEd
Bio-ITEST Co-Principal Investigator
CEO, EdLab Group

Contributions, Editing, and Curriculum Design Services:

Kristen Clapper Bergsman, MEd
Laughing Crow Curriculum, LLC

Graphic Design:

Clayton DeFrate Design

Copyediting:

Polly Freeman, MA
Introduction

Up to this point, students have seen the BRCA1 protein represented in a linear, sequential form. In this lesson, students are introduced to the importance of a protein's three-dimensional structure. Students first engage in a short activity in which they use a pipe cleaner to perform a simple function, as an analogy for the relationship between a protein's structure and function. Students then learn to navigate between linear protein sequences and three-dimensional structures by using Molecule World™ to explore molecular models. Students begin by viewing and manipulating DNA—a familiar molecule—using Molecule World. When students are familiar with the program, they visualize parts of the BRCA1 protein to see how a specific mutation in the BRCA1 gene ultimately changes or destroys the protein's function. In Lesson Five, students learn how 3D animators might use bioinformatics tools in their careers.

Learning Objectives

At the end of this lesson, students will know that:

- Bioinformatics tools like Molecule World help scientists visualize proteins.
- A protein is a physical “thing” with a three-dimensional structure that determines its function.
- All proteins are comprised of amino acids linked together by covalent bonds and have the same general structure: a ‘beginning’ or N-terminus, which contains an amino group (NH₃⁺), and an ‘end’ or C-terminus, which contains a carboxyl group (COO⁻).
- Each amino acid has a chemical group that is unique to it, called an R-group (also known as a side-chain). The chemistry of the amino acid R-group is important for a protein’s shape and function.
- Mutations can impact the three-dimensional structure of proteins, and thus impact the protein’s function, as is the case with many genetic disorders.
- Bioinformatics tools are used and created by people in many careers, including 3D animators.

At the end of this lesson, students will be able to:

- Use Molecule World to view complex biological molecules in a three-dimensional format.
- Manipulate three-dimensional images in numerous ways to deepen understanding of molecular structure.

Key Concepts

- Genetic disorders are often caused by dysfunctional or absent proteins.
- Proteins are physical “things” with three-dimensional shapes. The shape of the protein is crucial to its function.
• All proteins are comprised of amino acids linked together by covalent bonds and have the same general structure: a ‘beginning’ or N-terminus, which contains an amino group (NH$_3^+$), and an ‘end’ or C-terminus, which contains a carboxyl group (COO$^-$).
• Each amino acid has a different R-group (also known as a side-chain). Each R-group has different chemical properties. The chemistry of the amino acid R-group is important for a protein’s shape and function.
• Mutations can cause changes in the three-dimensional shape of a protein. The change in shape can alter the function, resulting in genetic disorders, including cancer.
• Programs like Molecule World allow scientists (and students!) to view macromolecules in a number of different ways to enhance their understanding of the molecule’s structure and function.
• Bioinformatics tools are used and created by people in many careers, including 3D animators.

### Materials

<table>
<thead>
<tr>
<th>Materials</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copies of Student Handout—Careers in the Spotlight (handed out in Lesson One)</td>
<td>1 per student</td>
</tr>
<tr>
<td>Class set of Student Handout—Instructions for Using Molecular Models</td>
<td>1 per student (class set)</td>
</tr>
<tr>
<td>Student Handout—Using Molecular Models Worksheet</td>
<td>1 per student</td>
</tr>
<tr>
<td>Pipe cleaners</td>
<td>1 per student</td>
</tr>
<tr>
<td>Teacher Answer Key—Using Molecular Models</td>
<td>1</td>
</tr>
<tr>
<td>Teacher Resource—Teacher Demonstration: Comparing Molecular Structures</td>
<td>1</td>
</tr>
<tr>
<td>Teacher Resource—From Sequence to Structure</td>
<td>1</td>
</tr>
</tbody>
</table>

[Note: This worksheet is for students’ answers to lesson questions.]

### Computer Equipment, Files, Software, and Media

Computer or iPad with internet access and projector to display PowerPoint slides (or pdf document) and for the Teacher Demonstration described in Teacher Resource—Teacher Demonstration: Comparing 3D Structures.

**Alternative:** Print PowerPoint slides onto transparencies and display with overhead projector.

Either a classroom set of iPads or student iPads with Molecule World installed. Molecule World is available at https://itunes.apple.com/us/app/molecule-world/id863565223
It will be helpful to load structure files in Molecule World in advance (1NAJ, 1Y98, 1JNX, 1N5O). Otherwise, structure files can be downloaded during class. A note-taking or word processing program for the iPad is recommended. A program for annotating images, for example Skitch is also recommended.

**Lesson Five PowerPoint Slides—Learning to Use Molecular Models.** Available for download at: http://www.digitalworldbiology.com/dwb/gt/curriculum/introductory-bioinformatics-genetic-testing-MW.

A student version of lesson materials (minus teacher answer keys) is available from Digital World Biology at: http://www.digitalworldbiology.com/dwb/gt/curriculum/introductory-bioinformatics-genetic-testing-MW.

BRCA1 Animation. Available for viewing at: https://www.nwabr.org/sites/default/files/BRCA1_animation.swf.
Access to the Microsoft Word® or Google Docs document that students created in Lesson Four.

Music files (mp3 format) for Teacher Resource—From Sequence to Structure are available for download at: http://digitalworldbiology.com/dwb/bio-itest-genetic-testing.
Teacher Preparation

- Load the classroom computer with the Lesson Five PowerPoint slides.
- Download the Molecule World app from the iTunes store and install on student iPads. Molecule World is available in the iTunes store at: https://itunes.apple.com/us/app/molecule-world/id863565223

To maximize class time for lesson activities, it will be helpful to download the following structure files on each iPad in advance: 1NAJ, 1Y98, 1JNX, 1N5O.
- Make copies of Student Handout—Instructions for Using Molecular Models, one per student. This handout is designed to be reused as a class set.
- Make copies of Student Handout—Using Molecular Models Worksheet, one per student. The worksheet is used for students to write their answers to the lesson questions.

Procedure

WARM UP

1. As students enter the classroom, display Slide #1. This slide highlights 3D animator Beth Anderson.

2. Have students retrieve Student Handout—Careers in the Spotlight from Lesson One.

3. Students should think about, and write down, what kind of work a 3D animator might do (3D Animator Question #1). This will be revisited at the end of the lesson, including how a 3D animator might use bioinformatics in his or her job.

4. Tell students to keep their Careers in the Spotlight handout available for future lessons.
PART I: Structure Meets Function: Pencil Transferase

5. Explain to students the **aim of this lesson**. Some teachers may find it useful to write the aim on the board.

**Lesson Aim:**

- To understand that protein structure can impact protein function, using the bioinformatics tool Molecule World to visualize molecules.

Teachers may also wish to discuss the **Learning Objectives** of the lesson, which are listed at the beginning of this lesson.

6. Give each student a pipe cleaner. Tell students that their task is to **move a pencil from one desk to another desk** using the pipe cleaner.

7. The rules are:

   - Students must use a pencil (not a pen or other implement with a cap).
   - Students may not touch the pencil with their hands when they are moving the pencil across desks.
   - Students may bend the pipe cleaner as needed.
   - After the pencil has been successfully moved, students should **retain the shape of the pipe cleaner** that successfully moved the pencil.

8. Give students time to complete the task.

9. When students have finished, have them hold up their pipe cleaners for everybody to see. The pipe cleaners will likely be very similar shapes, with a single or double loop in the middle, and a handle on each end.

10. Ask students, “**What shape is your pipe cleaner? What shape is your neighbor’s pipe cleaner? Are there any shapes that are predominant in the class?**”

11. Draw the shape(s) on the board.

12. Finally, ask, “**Why do so many of the pipe cleaners have such similar shapes?**”

13. Students will recognize that the **shape of the pipe cleaner allows it to perform its function.**

14. Tell students that, like the pipe cleaner, proteins are folded into specific shapes to perform their functions.

15. Show **Slide #2**, which shows the structure of “Pencil Transferase.” Drawn on the board, the “Pencil Transferase” protein may look like the image shown in **Figure 1**.

- **-ase** is a common suffix for enzyme names.

---

**Figure 1**: “Pencil Transferase.”
16. Tell students that within “Pencil Transferase” there are areas with different structures that perform different functions. These units are called protein domains. Our “protein” has two handle domains and one pencil-binding domain.

17. Referring to Slide #3, show students the representation of the linear structure of “Pencil Transferase” on the board, as shown in Figure 2.

18. Point out that proteins are made up of amino acids that are held together by covalent bonds, and have a ‘beginning’ and an ‘end,’ labeled N-terminus and C-terminus respectively. The N-terminus or ‘beginning’ has an amino group (NH₃⁺), while the C-terminus or ‘end’ has a carboxyl group (COO⁻).

19. Show Slide #3 or draw the representation of the linear structure of BRCA1 on the board (see Figure 3). Tell students that the BRCA1 protein also has multiple domains, including a DNA binding domain and two BRCT domains. Their functions are:

- **DNA binding domain:** To bind DNA! Remember, the function of the BRCA1 protein is to repair damaged DNA. This region binds to damaged DNA to repair it.

- **BRCT domains:** BRCA1 needs to work with other proteins to repair DNA—it is one of many proteins that “cooperate” to do this. The BRCT domains facilitate protein-protein interactions involved in DNA repair. BRCT stands for Breast cancer C-Terminal domain (see Figure 3).

**Domain:** Specific area of a protein that performs a particular function.

**Figure 2:** “Pencil Transferase” domains.

[Note: Need to move a piece of chalk instead of a pencil? Your “chalk-moving protein” will likely have similar domains. The amino acid sequence that folds itself into a loop to carry the pencil may be the same amino acid sequence that folds itself into a loop to carry the chalk in a different protein. The recurring domain units found in different proteins that move pencils, chalk, pens, or other similar shapes are called conserved domains.]

**Covalent Bonds:** A type of chemical bond that is characterized by the sharing of a pair of electrons between atoms.

**N-terminus:** The ‘beginning’ of a protein, containing an amino group (NH₃⁺).

**C-terminus:** The ‘end’ of protein, containing a carboxyl group (COO⁻).
20. Inform students that the BRCA1 protein is comprised of 1,863 amino acids. In Lesson Four, students learned that the mutation that affects Deb (and other family members) occurs at amino acid number 1775. At position number 1775, a methionine amino acid is replaced with an arginine amino acid. This is abbreviated as a M1775R mutation.

21. Ask students, “Knowing the position of the mutation (at amino acid number 1775 out of a protein 1,863 amino acids long), which part of the protein is most likely affected by the mutation?”

22. Tell students that the M1775R mutation occurs near the intersection of the two BRCT domains. As such, when the class looks at the protein structure, the class will only be viewing the BRCT regions, not the whole protein. Proteins as large as BRCA1 are difficult for biochemists to work with whole, so they often crystallize one piece (i.e., one or two domains) at a time.

23. Draw students’ attention back to their pipe cleaners and tell them that different types and locations of mutations can affect the protein in different ways. For example, a substitution mutation (in which one amino acid is substituted for another) may not have a harmful effect if it is a silent mutation (does not result in a change in amino acid sequence), or if it happens inbetween functional domains, like at point (A), as shown in Figure 4.
A substitution mutation at a crucial point in a functional domain, such as where the amino acids link together to make a loop (point B), may destroy the shape of the protein altogether, thus destroying its function. A deletion or insertion mutation could have similar consequences, depending on where in the protein the mutation occurs and how crucial the correct amino acid is at that point.

24. Show Slide #4. Briefly review levels of protein structure:

- **1° Primary**: The linear order of amino acids (the sequence of amino acids along the pipe cleaner model).
- **2° Secondary**: The alpha-helices and beta-sheets (similar to the ‘loops’ in the pipe cleaner model).
- **3° Tertiary**: The whole conformation or shape of the protein (including the handle domains in the pipe cleaner model) – the way the whole model folds.
- **4° Quaternary**: If the protein has more than one subunit or chain. (This can be demonstrated by putting several pipe cleaner models together.)

**Part II: Using Molecular Models**

25. Tell students that, so far, they have been looking at the BRCA1 protein as a linear sequence of amino acids. Today they will be exploring the 3D shape of the BRCA1 protein to find out why the substitution of a single amino acid in the 1775th position has such serious consequences for the Lawler family.

26. The BRCA1 Animation illustrates the normal function of BRCA1, as well as the consequences of inherited mutations. Show the BRCA1 Animation to students. Alternatively, students can view the animation individually by visiting the NWABR website at: https://www.nwabr.org/sites/default/files/BRCA1_animation.swf.

27. Tell students that they will learn how to view molecules in 3D with a molecular modeling tool, Molecule World. They will learn about the program by viewing and manipulating a short DNA molecule.

[Note: What if a point mutation in a sequence of DNA results in a premature stop codon? This causes protein translation to stop early, leading to a shorter protein. Teachers can demonstrate a truncation by simply cutting off part of the pipe cleaner and asking, “How well do you think the protein would work now?”

Unfortunately, insertions and deletions may not occur in multiples of three bases, resulting in shifting of the reading frame during translation (i.e. frameshift mutations). This often introduces stop codons and premature truncation of the protein.]
28. When viewing the DNA molecule in Molecule World, students will see the anti-parallel strands, illustrated most clearly in Question #13 on Student Handout—Using Molecular Models. When viewing the DNA structure, students may be confused by the sequence of the DNA shown in the **Sequence viewer** below the **Structure**. DNA sequences are written 5’ to 3’ by convention. Therefore, both the 1NAJ-A and 1NAJ-B sequences are shown in the **Sequence viewer** as: cgcgaattcgcg. (See **Figure 5**.)

Figure 5: Molecule World Sequence viewer. Credit: Digital World Biology.

However, the two strands of DNA are complementary and anti-parallel in the actual DNA structure:

1NAJ_A: 5’ – cgcgaattcgcg – 3’

1NAJ_B: 3’ – gcgcttaagcgc – 5’

29. After students become familiar with Molecule World, students will view the “normal” (non-mutated) version of the BRCT region of the BRCA1 protein.

30. Tell students that neither students nor scientists can look at the entire BRCA1 protein, since its structure has not yet been “solved.” They can, however, look at the BRCT domains in that protein.

31. Remind students that BRCT stands for “**breast cancer carboxy-terminal.**” Although the BRCT domain is named for the **BRCA1** gene, it is present in many proteins that function in repairing DNA, in addition to BRCA1. Each BRCT domain is about 90-100 amino acids long and has a characteristic shape.

32. Pass out Student Handout—**Instructions for Using Molecular Models** and Student Handout—**Using Molecular Models Worksheet**. Allow time for students to work independently on the activity.
PART III: Teacher Demonstration—Comparing Structures in Molecule World

33. In some cases, scientists place molecular structures into orientations that make them easy to compare. When this is the case, we can see how some mutations change structures by viewing them one at a time. Using the instructions in Teacher Resource—Teacher Demonstration: Comparing Molecular Structures, teachers can demonstrate the structural changes between the non-mutated and mutated versions of the BRCA1 protein by highlighting the affected amino acids and alternating between the two protein structures.

Closure: Careers in the Spotlight

34. Today, students saw how a single amino acid substitution in a protein causes a very subtle shift in the protein’s shape and impacts its binding to another molecule. This small change, however, has significant implications for many families, including the Lawlers. There are many mutations to BRCA1 that can cause cancer aside from the one viewed in class today. Tomorrow, students will take a closer look at the BRCA test and other genetic tests to determine how useful they may or may not be.

35. Return to the picture of the 3D animator from the Careers in the Spotlight, Slide #5.

36. Show Slide #6, which provides job information for a 3D animator. Review this information with students.
37. Ask students, "What more do we know about 3D Animators after today’s lesson?" Point out that 3D animators in the biological sciences use computer programs like Molecule World or Cn3D and the information in protein crystal structures like the ones seen here for BRCA1 to create animations that help us visualize biological processes. Beth Anderson and her colleagues developed the BRCA1 animation seen in this lesson.

38. Ask students to answer 3D Animator Question #2 on their Careers in the Spotlight handout, which has students explain how this lesson has changed their understanding of the kind of work a 3D animator does.

39. Ask students to also answer 3D Animator Question #3 on their Careers in the Spotlight handout, which has students explain how a 3D animator might use bioinformatics in his or her work.

40. Tell students to keep their Careers in the Spotlight handout available for future lessons.

Homework

The following are suggested homework activities to follow this lesson:

A. Students should continue to prepare for the Socratic Seminar in Lesson Six using Student Handout—Categorizing Genetic Tests and/or Student Handout—Weighing the Risks and Benefits of Direct-to-Consumer Genetic Testing. Either Student Handout can also be passed out as homework to accompany the reading, which can be used as entry tickets for that class session. These are the reading and support materials for the Socratic Seminar students will participate in during Lesson Six. Students may need two days to prepare fully.

B. If students answered the Extension Questions #48-50 on Student Handout—Aligning Sequences with BLAST Worksheet in Lesson Four, they can expand upon their answers to these questions, describing what they have learned about how this single amino acid change can impact the structure of the BRCT domain of the BRCA1 protein. Students can also answer the questions posed on their Word® document created in Lesson Four and augmented in Lesson Five as homework (see Student Handout—Instructions for Using Molecular Models Part 2, Question #31).

C. As a reflective exercise, ask students to write about the activities they did in Lesson Four in their lab notebooks, on another sheet of paper, or in a word processing program like Microsoft Word® or Google Docs which they then provide to the teacher as a printout or via email. This can serve as an entry ticket for the following class. Have them complete these prompts:

   a. Today I learned that…

   b. An important idea to think about is…

   c. Something that I don’t completely understand yet is…

   d. Something that I’m really confident that I understand is…

[Note: Suggested scoring for reflection: +5 points if all 4 prompts are complete.]
Extension

- For a musical analogy with MP3 clips to describe protein structure and conserved domains, a draft extension activity for Lesson Five can be found at the end of this lesson plan on Teacher Resource—From Sequence to Structure.

Adaptation

- Teachers can work through all or part of Student Handout—Instructions for Using Molecular Models as a teacher-led demonstration, as desired.

Teacher Background

Protein Structures: Scientists know what DNA, a protein, or other macromolecules look like by using a number of different tools, including X-ray crystallography and nuclear magnetic resonance (NMR), to determine molecular structures. In the case of x-ray crystallography, proteins are crystallized before being bombarded with x-rays. By measuring how the crystals deflect or diffract the x-rays, scientists can determine or “solve” the protein’s structure. With NMR, molecules are subjected to a strong magnetic field that causes the natural magnets in atomic nuclei to spin in the same direction, just like iron filings line up in the presence of a magnet. Radio waves are used to disrupt this state, and scientists measure how each atom responds. The characteristics of the response are used to determine or “solve” the molecular structure by determining which other atoms are close by. These structures are then added to the Protein Data Bank (PDB), and later, to the structure database at the NCBI for others to view using special molecular-viewing programs like Molecule World. Molecular-viewing programs like Molecule World allow users to view and manipulate three-dimensional structures on a computer screen.

Assessment Suggestions

Students can be assessed on their answers to the questions posed on their Word® document created in Lesson Four and augmented in Lesson Five (see Student Handout—Using Molecular Models).

Glossary

Accession number: A unique identifier or code assigned to every entry in the National Center for Biotechnology Information (NCBI) databases. This unique code can be used to search the databases to find your gene, protein, or structure of interest.

Alpha helix (Plural: “alpha helices”): A common structure of proteins, characterized by a single, spiral chain of amino acids stabilized by hydrogen bonds.

Beta sheet: A structure that occurs in many proteins and consists of two or more parallel adjacent polypeptide chains arranged so that hydrogen bonds can form between the chains.

BRCT domain: The Breast Cancer C-Terminal domain, a protein domain in the BRCA1 protein located at the ‘end,’ or C-terminus, of the protein that is involved in protein-protein interactions.

Molecule World: A molecular modeling program from Digital World Biology. Molecule World displays three-dimensional molecular structures along with the sequences for biological molecules such as proteins, RNA, and DNA.

Covalent bonds: A type of chemical bond that is characterized by the sharing of a pair of electrons between atoms.

C-terminus: The ‘end’ of protein, containing a carboxyl group (COO-).
**Domain:** Specific area of a protein that performs a particular function.

**DNA binding domain:** Specific area of a protein (domain) that binds to DNA. The DNA binding domain in the BRCA1 protein is necessary for the BRCA1 protein to repair damaged DNA.

**Frameshift mutation:** A frameshift mutation is a genetic mutation caused by an insertion or deletion of a number of nucleotides not evenly divisible by three. Because codons are read as triplets, these insertions or deletions change the reading frame during protein translation. These reading frame shifts often create premature stop codons resulting in truncated proteins.

**Hydrophilic:** A substance that is attracted to water. From the Greek “hydro” which means water, and “philos” meaning love. A synonym for polar. The opposite of hydrophobic or non-polar.

**Hydrophobic:** A substance which repels water. From the Greek “hydro” which means water, and “phobos” which means fear. A synonym for non-polar. The opposite of hydrophilic or polar.

**NMR spectroscopy:** Nuclear Magnetic Resonance spectroscopy, usually abbreviated as “NMR,” is a technique used to determine the three-dimensional structure of molecules, including proteins. Molecules are subjected to a strong magnetic field that causes the natural magnets in atomic nuclei to spin in the same direction, just like iron filings line up in the presence of a magnet. Radio waves are used to disrupt this state, and scientists measure how each atom responds. The characteristics of the response are used to determine or “solve” the molecular structure by determining which other atoms are close by.

**Non-polar:** A substance that repels water. A synonym for hydrophobic.

**N-terminus:** The ‘beginning’ of a protein, containing an amino group (NH$_3$+).

**Polar:** A substance that is attracted to water. A synonym for hydrophilic. The opposite of hydrophobic or non-polar.

**Protein Data Bank (PDB):** A repository or collection of three-dimensional structures of large biological molecules, including proteins and nucleic acids, submitted by scientists from around the world. This data is typically obtained by X-ray crystallography or NMR spectroscopy.

**Stop codon:** A codon (series of thee nucleotides in a row) that terminates, or stops, protein translation.

**Substitution mutation:** When one amino acid is substituted for another as a result of mutation.

**Truncation:** To shorten, as if by cutting off. During translation, a growing protein chain is truncated if it encounters a premature stop codon.

**X-ray crystallography:** A method of determining the arrangement of atoms within a crystal, such as a crystal of a particular protein, in which a beam of X-rays strikes a crystal and deflects or diffracts into many specific directions. From the angles and intensities of these diffracted beams, a crystallographer can produce a three-dimensional picture of the density of electrons within the crystal, and thus calculate and estimate the three-dimensional shape of the molecule used to generate the crystals. This is often referred to as determining or “solving” the protein’s structure.

**Credit**

BRCA1 domain illustration from: http://www.biochemsctrans.org/bst/037/0597/bst0370597a01.gif.

BRCA1 Animation developed by Beth Anderson, Arkitek Studios, and Jill DelSordi.

Anderson, Beth. Personal Interview. 2 July 2010.

Photo of Beth Anderson provided by Doug Huff.


The authors wish to thank Wikimedia Commons for the definitions of some of the vocabulary terms found in the Glossary and throughout this lesson.
5

Instructions for Using Molecular Models

**Aim:** To understand that protein structure can impact protein function, using the bioinformatics tool Molecule World to visualize molecules.

**Instructions:** Write the answers to your questions on Student Handout—*Using Molecular Models Worksheet*, in your lab notebook, or on a separate sheet of paper, as instructed by your teacher.

---

**PART I: Viewing DNA Structure**

1. Open the Molecule World iPad app.

2. Tap the list icon in the top left-hand corner to view the list of pre-loaded structures.

3. If your structures have already been loaded, locate 1NAJ in the list of structures and touch the title to view it in the structure window.

   If 1NAJ is not in the list, tap the + icon to open the search menu. Select the MMDB database, type "1NAJ" in the search area, and touch the Search button to locate the structure in the NCBI's Molecular Modeling Database. Touch the title of the structure when it appears to view it in the structure window.

   We will begin our investigation of 3D structure by looking at a molecule you are already familiar with: DNA. "1NAJ" is the accession number for a file that contains structure information for a small piece of double-stranded DNA. An accession number is like a catalog number or bar code; it bears no resemblance to the product itself, but allows you to access information in databases such as the MMDB, PDB, or PubChem.

4. View 1NAJ, a DNA structure. Drag the structure with one finger to turn it around.

5. Drag the structure with two fingers to move the structure without turning it.

6. Pinch your fingers together to make the structure smaller or spread them apart to make the structure larger.

7. Experiment with controlling the movement and viewing the DNA molecule from multiple angles.

8. Touch the molecule icon to view possible drawing and coloring styles.

   Explore each of the options for drawing the DNA molecule.

   - **Ball & stick**
   - **Spacefill**
   - **Tubes**
Imagine you are teaching a class about DNA. Which drawing styles would you use to teach students about DNA structure? Explain the reasons for your decision.

9. Touch the molecule icon again to view different visualization styles. There are many different styles for coloring molecules. Try each of the coloring styles listed below. Answer each of the following questions.
   a. **Charge**: How did the charge coloring style change the color of the DNA? Use the palette icon to view the charge color key. What does the color key tell you about the the way DNA is charged?
   b. **Element**: How did the element coloring style change the color of the DNA? Use the palette icon to view the element color key. How many elements are in DNA? What elements are in DNA?
   c. **Rainbow**: ‘Rainbow’ uses the color red at the start (5’ end) and continues through the rainbow. Why are two regions in the DNA colored red?
   d. Explore a few combinations of drawing and coloring styles. Which coloring options do you find most useful? Why?

10. Touch the “Show sequence” button to open the **Sequence viewer**. Touch a letter in the DNA sequence to see what happens. Each letter represents a base in the DNA. Note that there are two DNA sequences: 1NAJ-A and 1NAJ-B. Each sequence corresponds to a different strand of the DNA.

![Figure 2: The Sequence viewer shows the sequences for nucleic acids or proteins. Credit: Digital World Biology, 2015](image)

11. Touch the molecule icon and the “Reset Appearance” button to restore the DNA to its original state.

12. In the sequence 1NAJ-A, touch the first “G” to select it. Next, touch the first “G” in sequence 1NAJ-B. Where are these two guanines located in the DNA molecule? Explain why.

13. Touch the molecule icon. Select the “Molecule” coloring style. Note the correspondence between the colors in the sequence and the colors in the structure.
a. What did changing the coloring style to Molecule do? Explain. How many molecules are in DNA?

b. Move the DNA so that the ring structure(s) of the selected nucleotides bases can be seen clearly. What do you see?

14. Open the Selection menu in the bottom right-hand corner. Choose “Clear Selection.”

15. Touch any base in either DNA sequence to select it.

16. Open the Show/Hide menu and choose “Hide Unselected Residues.” Answer the following questions:
   
   a. What do you see? Zoom in to view the base more easily.
   
   b. Challenge question: Is the base you selected a purine or a pyrimidine, and how do you know?
   
   c. In the Sequence viewer, touch the name of the chain that contains your base to see the complete strand.

PART II: Viewing the BRCA1 Protein

Now that you are familiar with Molecule World, we will view part of the BRCA1 protein and part of a second protein that it interacts with.

17. If your structures have been pre-loaded, locate 1Y98 in the structure list. Touch the title to view the structure.

   If the 1Y98 structure is not in the list, tap the + icon to open the search menu. Select MMDB as the database, type “1Y98” in the search area, and click the Search button to locate the structure. Touch the title of the structure after it appears to download the structure and view it in the structure window.

18. Turn the structure around to identify regions where a single chain shows multiple loops (alpha helices).

19. Locate regions where chains are organized in parallel lines (beta sheets). Both alpha helices and beta sheets are examples of secondary structure.

   Draw an alpha helix and a beta sheet.

20. The structure contains two BRCT domains. To see the domains more clearly:

   A. Touch the molecule icon and choose the “Rainbow” coloring style.
   B. Touch the Show sequence button to view the amino acid sequence.
   C. Then, touch 1Y98-A (the name of the top protein chain) to select it.
   D. Touch the Show/Hide button and choose “Hide unselected” to hide other objects.

   The rainbow coloring style will help you visualize the two BRCT domains. BRCT stands for Breast cancer C-Terminal domain. This domain is involved in protein-protein interactions. One domain will be colored in red, yellow, and green. The other will be colored in shades of blue and purple. If you turn the structure, you’ll see a space between the two domains.
21. Touch the molecule icon and choose “Reset appearance.”

22. Open the Sequence viewer and touch the name of the second protein chain, “1Y98-B,” to highlight the CtIP protein. This protein interacts with BRCA1 when damaged DNA is repaired. A phosphate group has been added to this protein.

23. Touch the chain name, 1Y98-B, again to deselect it.

24. Scroll through the amino acid sequences and look for the letter “M” in the top chain, “1Y98-A.” M stands for the amino acid methionine. Touch the M to select it and look at the PDB position to see where this amino acid would be located in the complete protein chain.

25. Touch the letter M again to deselect it.
26. Scroll through the sequence, selecting and deselecting M’s, until you find the M at PDB position 1775.

Amino acid position PDB 1775 is the location of the M1775R mutation in BRCA1 that affects the Lawler family. When the M (methionine) at PDB position “1775” has been selected, it will appear slightly brighter in the protein structure. The structure you are looking at does not have the mutation. (Recall that the mutation converts M to R.)

27. Touch the Show/Hide button and choose “All atoms in residue” to see the entire side chain.

28. Touch the molecule icon \[\text{molecule}\] and choose “Ball & stick” to see the methionine more clearly.

29. Touch the Selection button and choose “Select nearby” to select other molecules within a radius of six Angstroms. [Note: An Angstrom, Å, is 1x10^{-10} meters or 0.1 nanometer.]

30. Touch the Show/Hide button and choose “All atoms in residue,” then touch the molecule icon \[\text{molecule}\] again and choose Ball & stick.

Notice how close the side chains in the CtIP residues are to the methionine at the mutation site in BRCA1. You can imagine how a change in BRCA1 might impact its ability to interact with the CtIP protein which is also required for DNA repair.

31. Record your results by capturing an image of the interaction site between the BRCA1 and CtIP proteins:
   a. Touch the camera icon and choose “Save image” to save an image of the two proteins in Photos.
   b. Open a photo annotation program like Skitch and add an arrow or a circle to show the mutation site. Alternatively, you can add circles later in Word.
   c. Open the Word® document you created in Lesson Four. It should be labeled with your LASTNAME_BRCA1_NCBI.
   d. Insert the labeled image into your document.
   e. Type these instructions and questions at the bottom of your Word® document:
      1. Circle the location of the mutation in the picture of the BLAST alignment.
      2. Explain what the picture of the protein structure and the picture of the alignment represent.
      3. Explain how the pictures are connected to each other.
   f. Follow the instructions you typed above (1-3) and answer the questions in your Word® document. If you haven’t added arrows or circles to identify the mutation site already, you can use the drawing tools in Word® to draw circles around the mutation in the BLAST sequence window and the protein image.
   g. Save and close this document.

32. Challenge Question: What can you see now that you could not see before you annotated this structure? Does this help you understand the consequences of the Lawlers’ M1775R mutation?

33. Optional: If time permits, you may wish to experiment with protein structures 1JNX and 1N5O. Make sure that you enter capital O and not a zero for “1N5O.”

   1JNX: This is the accession number for the BRCT domains of the non-mutated version of the BRCA1 protein.
   1N5O: This is the accession number for the BRCT domains of the M1775R mutation of the BRCA1 protein.
Aim: To understand that protein structure can impact protein function, using the bioinformatics tool Molecule World to visualize molecules.

Instructions: Use Student Handout—Instructions for Using Molecular Models to complete this worksheet.

PART I: Viewing DNA Structure

8. There are many ways to view the DNA. Imagine you are teaching a class about DNA. What rendering option(s) would you use to teach students about DNA structure?

(circle one): Ball & stick Spacefill Tubes

Explain the reasons for your decision:

9a. How did the "Charge" coloring style change the coloring of the DNA? What does the color key tell you about the the way DNA is charged?

9b. How did the "Element" coloring style change the color of the DNA? How many elements are in DNA? What elements are in DNA?

9c. Now look at Rainbow: “Rainbow” uses the color red at the start (5’ end) and continues through the rainbow. Why are two regions in the DNA colored red?

9d. Which coloring option do you find most useful? Why?

12. When you select the two guanines that are in the same place for both DNA strands, where are they located on the DNA molecule? Explain why.
13a. You changed the **Coloring style** to "**Molecule**." What does that do? Explain. How many molecules are in DNA?

13b. You moved the DNA so that the ring structure(s) of the selected nucleotide's base is clearly seen. What did you see?

16a. After choosing the "**Hide Unselected Residues**" button, what do you see?

16b. **Challenge question:** Is the base you selected a purine or a pyrimidine, and how do you know?

**PART II: Viewing the BRCA1 Protein**


32. **Challenge Question:** What can you see now that you could not see before you annotated this structure? Does this help you understand the consequences of the Lawlers’ M1775R mutation?
PART I: Viewing DNA Structure

8. There are many ways to view the DNA. Imagine you are teaching a class about DNA. What rendering option(s) would you use to teach students about DNA structure?

(circle one): Ball & stick Spacefill Tubes

Explain the reasons for your decision:

Student answers will vary. Some students will prefer the simplified representation of the tubes, while others will prefer the additional details in ball & stick (which are both good for visualizing the rings of the DNA bases), or spacefilling views, which make it possible to view the full size of the atoms, by showing spheres to represent the electron clouds.

(+1 for selecting option; +1 for providing reasonable explanation.)

9a. How did the “Charge” coloring style change the coloring of the DNA? What does the color key tell you about the way DNA is charged?

The entire DNA molecule (the ‘object’) turns red. (+1)
DNA has a negative charge (+1)

9b. How did the “Element” coloring style change the color of the DNA? How many elements are in DNA? What elements are in DNA?

Each element appears in a different color. (+1)
DNA has five elements (hydrogen, phosphorus, carbon, nitrogen and oxygen) (+1)

9c. Now look at Rainbow: “Rainbow” uses the color red at the start (5’ end) and continues through the rainbow. Why are there two red regions of the DNA?

(+1 for ‘different strands’ or ‘two strands’ or ‘DNA is double stranded’.)

9d. Which coloring option do you find most useful? Why?

Student answers will vary. Some students will prefer the simplicity of the Object coloring, while others will find the Rainbow coloring useful to identify the 5’ and 3’ ends of each DNA strand.

(+1 for choosing option; +1 for providing reasonable explanation.)
13. When you select the two guanines that are in the same place for both DNA strands, where are they located on the DNA molecule? Explain why.

The two guanines are located on opposite strands of DNA, so in the three-dimensional DNA structure, one guanine is located on the top (near the 5’ end of one strand), and one is located on the bottom (near the 5’ end of the second strand).

(+1 for providing location; +1 for explaining that they are on opposite strands.)

14a. You changed the **Coloring style** to “**Molecule.**” What does that do? Explain. How many molecules are in DNA?

Each strand of DNA appears in a different color. (+1)
Each strand can be considered to be a separate molecule. (+1)
There are two separate molecules in DNA. (+1)

14b. You moved the DNA so that the ring structure(s) of the selected nucleotide’s base is clearly seen. What did you see?

You can visualize the structure of the base: the backbone, the ring, and the phosphodiester bond. (+1 for structure or shape of the base.)

16a. After choosing the “**Hide Unselected Residues**” button, what do you see?

Most of the DNA structure was hidden, and only the selected base was visible. (+1.)

16b. **Challenge question:** Is the base you selected a purine or a pyrimidine, and how do you know?

The selected base is either a purine (if it has two rings) or a pyrimidine (if it has one ring). (+1 bonus point for correct response; +1 bonus point for # of rings.)

**PART II: Viewing the BRCA1 Protein**


Each type of secondary structure should look somewhat like the drawing. (+1 for each object, +2 total)
31. You will need to open the document that you saved your BLAST Alignment to (the DNA and Protein sequence). Follow the instructions on Student Handout—Instructions for Using Molecular Models.

Students should include an image from Molecule World of the BRCA1 protein structure. Circles or arrows should be present and should identify the location of methionine 1775.

The image of the BLAST alignment should also have circles or arrows to identify the location of the M1775R mutation. Students should explain that the BLAST picture represents an alignment of the Lawler family protein sequences compared to a reference sequence, which shows the location of the mutation (a methionine to arginine mutation at position 1775). The protein image is a picture of part of the structure of the BRCA1 protein, which also has a circle around the location of the M1775R mutation found in the Lawler family.

(+1 for screen capture; +1 for circling mutation; +1 for explaining the BLAST picture is the protein sequence with mutation; +1 for explaining that the image represents the position where the mutation would be found in the protein structure. 4 points total.)

32. Challenge question: What can you see now that you could not see before you annotated this structure? Does this help you understand the consequences of the Lawlers’ M1775R mutation?

Students can now see the methionine impacted by the Lawler family mutation, as well as amino acids that are near methionine 1775 on the CtIP protein. If methionine is mutated to an arginine at position 1775, it may interfere with the ability of BRCA1 to interact with the CtIP protein.

(+1 bonus for methionine; +1 bonus for explaining multiple interactions.)
Teachers Demonstration: Comparing Molecular Structures

The molecular structures that modeling programs depict are like a kind of three-dimensional graph. These graphs all include x, y, and z coordinates that are used to describe the position of each atom in space. They may also include information about the relative diameter of each atom and kinds of bonds it forms with other atoms in the structure. Some programs such as VAST at the NCBI are able to align secondary structure elements to each other, so that structures can be superimposed. In other cases, the structural biologists and biochemists who have solved molecular structures use the same coordinate systems. In these cases, the structures almost appear to be aligned. When this occurs, we can gain insights into the ways that mutations change structures by viewing them one at a time.

In this demonstration, teachers can show the structural changes between the non-mutated and mutated versions of the BRCA1 protein by annotating two molecular structures. The accession numbers for the two proteins are:

1JNX—This shows the BRCA domains of the non-mutated version of the BRCA1 protein.
1N5O—This shows the BRCA domains of the M1775R mutation of the BRCA1 protein.

1. Open Molecule World.
2. Open the structure list and touch the + to open the search menu.
3. To obtain the first structure, type 1JNX in the search area and touch the Search button to find the structures. You will need to use an iPad connected to a wireless network to access structures from the internet.
4. Touch the name of the structure when it appears in the search results to download the structure.
5. Repeat this process with the mutant structure 1N5O. (Be sure to use an “Oh” and not a zero).
6. Return to the structure list and touch 1JNX to load it in the structure window.
7. Scroll through the sequence to find and select the M at PDB position 1775. (M represents methionine).
8. Touch the Show/Hide button and choose “Show all residues.”
9. Touch the Molecule icon and choose “Ball & Stick.”
10. Return to the structure list and locate 1N5O.
11. Scroll through the sequence until you can find and select the R at PDB position 1775. (R represents arginine).
12. Touch the Show/Hide button and choose “Show all residues.”
13. Touch the Molecule icon and choose “Ball & Stick.”
14. Touch the names of the structures in the list to alternate between them and show how the mutation changes the structure. (See Figure 1)

Lori’s protein (1JNX) contains a methionine, while Deborah’s protein (1N5O) contains an arginine at the same position (the M1775R mutation).
15. While toggling back and forth between structures, point out to students how the change in amino acids (from M to R) changes the protein structure.
**Figure 1**: The top image shows the portion of the BRCA protein bound to the phosphorylated CtIP peptide in structure 1Y98. Credit: Digital World Biology.

**Figure 2**: The middle image shows the normal, non-mutated version of the BRCA protein from structure 1JNX. Methionine (M), found in the non-mutated version of the protein, is a non-polar amino acid. Non-polar (hydrophobic) amino acids fold to the inside of the protein structure, away from the aqueous environment. Note the difference in shape in the absence of the CtIP peptide. Credit: Digital World Biology.

**Figure 3**: The bottom image shows the same portion of the BRCA protein with the M1775R mutation (structure 1N5O). Arginine (R), found in the mutated version of the protein, is a polar amino acid. Polar (hydrophilic) amino acids tend to rotate to the outside of the protein structure. Students can see arginine “pop” to the outside of the mutated structure, which affects the shape of the entire structure in a subtle—yet important—way. Varma, et al. propose that the arginine interferes with binding to the CtIP peptide. Credit: Digital World Biology.
Lesson 5 – Learning to Use Molecular Models

From Sequence to Structure

Introduction

This lesson extension uses music as an analogy for protein structure and function. Part I compares primary, secondary, tertiary, and quaternary protein structures to a musical composition. Part II relates the idea of conserved domains in proteins to the song “Happy Birthday.” All mp3 music files for this optional lesson extension can be found under the Resources tab on the Introductory curriculum webpage at: http://www.nwabr.org/curriculum/introductory-bioinformatics-genetic-testing.

PART I: Primary, Secondary, Tertiary, and Quaternary Protein Structure

A. Primary Structure

Music is made up of a sequence of notes, and proteins are made up of a sequence of amino acids. Knowing only the order of the amino acids in a protein tells us the primary structure for that protein. This does not, however, tell the whole story of the protein. In a musical composition, knowing the order of the notes is of course important but also does not tell the whole story. For example, the musical notes for the beginning of a piece of music are:

DABADABAAAABAAAABAGF

If we could search a large database containing music for individual notes (iTunes® is a large database containing music, but it cannot search for sequences of notes), we might be able to find out more about our piece of music. Are there many other songs that contain the same sequence? A musician might be able to make an educated guess that this piece is in the key of D. Does that musical key link it to other musical pieces? Of course, since this is a piece of music, we can actually listen to it.

And it would sound like this:
Play the “Primary Structure Song.”

It doesn’t sound like much, does it?

Now, let’s take a look at the beginning sequence of amino acids from the BRCA1 reference sequence that we used for the BLAST alignment in Lesson Four. The beginning sequence of amino acids for the BRCA1 gene is:

MDLSALRVEEVQNVINAMQKILECPICLE

Again, knowing the order of the amino acids is important but does not tell the whole story. Using the NCBI, we can search a database that might compare this sequence with other sequences, which could tell us how common the sequence is, or perhaps even whether it is linked to other proteins with similar sequences. But, again, this does not tell us very much about what the protein does.
B. Secondary Structure

The basic shape of a protein begins to emerge when we look at the **secondary structure** that is formed as the alpha helices and beta sheets take shape.

Let’s return to the music analogy. The order of the notes gives us an indication of what the music should sound like, but we need more information to make the music take shape. Music is given shape by rhythm, dynamics (how loudly or softly sections are played), accents, the key in which it is written, and many other factors. If we add rhythm, some tone, and accents to the string of notes we heard earlier, it begins to take shape and becomes much more recognizable:

Play the “Secondary Structure Song.”

How does the BRCA1 protein take shape? It has to do with the interactions between the different amino acids and their functional groups that constitute the protein—interactions such as hydrogen bonding, non-polar interactions, ionic bonding, and even covalent bonding. If you take a closer look at the string of amino acids from our primary structure—look at the chemical structure and composition of each amino acid—you will see that some of the amino acids have **polar** (hydrophilic, or water loving) side groups and some have non-polar (**hydrophobic**, or water fearing) side groups. The non-polar side groups shy away from the watery solution the molecule is in, and bend the protein so that they can be gathered together on the inside of the structure. Hydrogen bonds also form between the amino acids themselves, which twists the protein into an **alpha helix**. Sometimes the interactions between the amino acids cause the string to fold back and forth in a flattened zigzag, which is called a **beta sheet**.

Just as rhythm, tone, and accent give shape to music, these chemical interactions twist and fold our amino acid chain into a basic three-dimensional structure.

C. Tertiary Structure

**Tertiary structure** is completed when the attractions between alpha helices and beta sheets cause the protein to fold back on itself and make “bridges” between sections of the protein. This gives the protein its final, folded, and compacted shape.

Using the music analogy, we have seen the music take shape from a string of single notes to, perhaps, a recognizable tune. At the level of tertiary structure, additional connections and bridges are made within the tune itself. The original sequence showed (and you have only heard) the first 21 notes of a much longer composition. If the sequence were written out in its entirety, there may be places that bridge to and interact with the beginning sequence. It might sound like this:

Play the “Tertiary Structure Song.”

---

**Non-polar**: A substance that repels water. A synonym for hydrophobic.

**Covalent bonds**: A type of chemical bond that is characterized by the sharing of a pair of electrons between atoms.

**Polar**: A substance that is attracted to water. A synonym for hydrophilic. The opposite of hydrophobic or non-polar.

**Hydrophilic**: A substance that is attracted to water. From the Greek “hydro” which means water, and “philos” meaning love. A synonym for polar. The opposite of hydrophobic or non-polar.

**Hydrophobic**: A substance which repels water. From the Greek “hydro” which means water, and “phobos” which means fear. A synonym for non-polar. The opposite of hydrophilic or polar.

**Alpha helix** (Plural: “alpha helices”): A common structure of proteins, characterized by a single, spiral chain of amino acids stabilized by hydrogen bonds.

**Beta sheet**: A structure that occurs in many proteins and consists of two or more parallel adjacent polypeptide chains arranged so that hydrogen bonds can form between the chains.
The BRCA1 protein, likewise, has its basic shape but requires additional infrastructure to be fully functional. For many proteins, their structures are dictated by additional folding that occurs when important covalent bonds form between the sulfur-containing amino acid cysteine. The final structure looks something like this:


This final shape, or structure, allows the BRCA1 protein to perform its function. It is bent, folded, and twisted to make it just the right size and shape so that it fits along the DNA double helix and does its job of repairing broken strands of DNA.

D. Quaternary Structure

Some proteins have an additional level of structure, the quaternary structure. If the fully functional end-protein is composed of more than one protein chain, the assembly of this larger unit is the fourth level of structure.

In the music analogy, we need to look at the end product. The composition was not written for a single violin, but for orchestra and choir. The composer built additional musical “bridges” between the different instruments and choir to create one cohesive piece of music. When it is all put together, it sounds like this:

Play the “Quaternary Structure Song.”

Something to think about: There are many levels of organization in making a “functional” musical composition and in folding functional proteins. This humble protein may not look like the equivalent of the “Hallelujah Chorus,” but BRCA1 is found in many species and performs a crucial function in DNA repair. In fact, George Frideric Handel would not have existed to write the “Hallelujah Chorus” without functional BRCA1 proteins.

Sources:

3-D Molecular Designs
http://www.3dmoleculardesigns.com/news2.php

Survey and Summary: Structural Classification of Zinc Fingers
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC140525/?tool=pubmed

Wikipedia: Zinc Finger
http://en.wikipedia.org/wiki/Zinc_finger
PART II: Conserved Domains

Nature has versions of its own “greatest hits” collection. These are the sequences of amino acids that code for proteins or regions of proteins that have crucial functions for many organisms. If many organisms require the same function (i.e., DNA replication and repair, energy metabolism, cell division) they will need similar proteins, made from similar sequences of amino acids. These related sequences are said to be conserved sequences, or conserved domains.

Note that these conserved sequences are similar, not identical. They need to be similar enough that the resulting protein can do the required job for that organism, yet organisms have a certain amount of genetic diversity. How can living things make proteins to do a specific job, yet allow for genetic diversity?

Back to the music analogy: You need a song with a specific function. Let’s say your friend is having a birthday and you would like to celebrate it with a song. We have, of course, a song just for this function.

Play the “Happy Birthday” song.

This is perhaps the most well-known song in the world—if it were a protein, it would be highly conserved. Would your friend, however, recognize this version?

Play the “Happy Birthday Improvisation” song.

Perhaps we do not need the entire original song to perform our function of wishing your friend a happy birthday—maybe the first six notes are crucial to the song, and then some variation can occur. Perhaps, throughout our improvisations, we need to have a few notes along the way that anchor us back to the original melody. One can also imagine that too much improvisation could make the melody totally unrecognizable and your friend would not know that you are trying to wish him or her a happy birthday. A song like that would not fulfill its function, and would probably not become known by too many people.

In a similar way, amino acid sequences that code for a functional protein can vary from one organism to another, as long as the conserved portion of the gene continues to code for the crucial elements that make up that protein.