<table>
<thead>
<tr>
<th>Grade: 11</th>
<th>Subject: Chemistry 20</th>
<th>Unit: Chemical Bonding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Can be used with modification in Biology 20</em> (protein structure) or <em>Biology 30</em> (Transcription and translation)</td>
<td><em>Building with molecular models.</em></td>
</tr>
</tbody>
</table>

**Title:** Exploration: Searching for Molecules at the Speed of Google

**Rationale**

Real World application of the use of **molecular models** in the development of a searchable molecular library.

**Background Information**

Research in the Derda Research Group at the University of Alberta is bridging synthetic Organic Chemistry and Genetically-Encoded Libraries (a technology for creating a large library of small molecules used in medicinal chemistry) to solve fundamental problems in drug discovery, cell biology, and diagnosis of disease.

One of the fundamental interests in the group is **accelerated discovery of functional molecules from genetically-encoded libraries of chemicals.** In other words, the group is looking to make a practical and rapid method of searching a vast library of molecules for specific molecules that can bind to specific proteins.

Currently the Derda Group is attempting to identify molecules that can block communication between proteins and carbohydrates in drug resistant cancer tumors by blocking a particular protein called galectin-3. Galectin-3 is a protein which is responsible for cancer cells growing and spreading.

If a molecule can bind to a foreign protein, it can stop its function and be an effective drug. This is of considerable interest to pharmaceutical companies. The question is, how can the molecules that bind to proteins be found quickly? Before this search can be done, the library has to be created, just like Google’s library of websites.

This activity will allow students to use **molecular model kits** to decode one molecule that has been added to the molecular library using a similar procedure that the Derda lab uses.

The libraries of molecules are created by allowing a virus, known as a bacteriophage, to produce a molecule made of amino acids. Bacteriophages are viruses that infect bacteria, and they contain a strand of DNA. Scientists can insert a piece of DNA into the virus DNA and
then the virus will convert that DNA sequence into a sequence of amino acids which becomes the molecule of research interest. Chemists can then modify the molecule by adding sugar molecules to it. This new molecule will be able to bind well with proteins such as lectin. Lectins are a type of protein that can bind to cell membranes. They are sugar binding molecules. Lectins offer a way for molecules to stick together without getting the immune system involved, which can influence cell-cell interaction.

**Curriculum Connections**

- 20–A2.3k relate electron pairing to **multiple and covalent bonds**
- 20–A2.4k draw electron dot diagrams of atoms and molecules, **writing structural formulas for molecular substances** and using Lewis structures to predict bonding in simple molecules
- 20–A2.5k apply VSEPR theory to **predict molecular shapes** for linear, angular (V-shaped, bent), tetrahedral, trigonal pyramidal and trigonal planar molecules
- 20–A2.6k **illustrate, by drawing or by building models**, the structure of simple molecular substances
- 20–A2.7k **explain intermolecular forces**, London (dispersion) forces, dipole-dipole forces and hydrogen bonding
- 20–A2.1s state a hypothesis and **make a prediction about the properties of molecular substances based on attractive forces**

**Lesson Objectives/Concepts**

- Students will translate a molecular formula into a three dimensional representation using molecular model kits
- Students will become familiar with current scientific research in organic chemistry
- Students will analyze molecular models to determine bond characteristics and intermolecular forces

**Time:** 1-2 class periods of 80 minutes

**Materials:**

- Vials with stoppers containing a “phage DNA sequence” and the inserted sequence for two additional amino acids - enough for one vial per two students. Cut out strips of paper containing 6 DNA letters that will code for two amino acids. There can be any sequence of DNA letters on either side of the 6 DNA letters. The surrounding letters represent the bacteriophage DNA; the 6 letters represent the inserted DNA coding the amino acids. The surrounding letters should be consistent for each student. Include “start” and “end” so students know which order to read from. For example, make strips of sequences such as “AACGT-start – TTAGGC – end – GTCAT” which represents
phage DNA, then the sequence to code for leucine and glycine, and then phage DNA again. Another example could be “AACGT – start – AGATAT – end-GTCAT” to code for arginine and tyrosine.

This link will provide all DNA sequences for all 20 amino acids:
http://sites.science.oregonstate.edu/genbio/otherresources/aminoacidtranslation.htm

- Molecular model kits
- DNA Translation Chart
- Amino Acid Chart

Introduction (10-15 minutes)

Pre-teaching: Ensure you have completed instruction on intramolecular bonds and intermolecular forces, VSPER shape predictions, and the solubility of molecular compounds. This activity is a culminating activity of many chemical bonding concepts.

1. Explain to students that they will be carefully examining molecules called amino acids and building them using molecular models. They will be determining the particular amino acids using a technique similar to one used at the University of Alberta where they will look inside a small virus and decode a small piece of DNA. The virus is called a bacteriophage and is very useful in biomedical research and can only affect bacteria so is safe to use.

2. Provide students a generalized summary on how the code of DNA is responsible for making protein. This website is useful in summarizing the process simplistically: http://www.ducksters.com/science/biology/proteins_and_amino_acids.php. DNA is the blueprint for a cell and will be “transcribed” to RNA which is “translated” to a series of amino acids connected together that eventually forms into a protein. DNA coding for amino acids can be found at https://en.wikipedia.org/wiki/DNA_codon_table

3. Show students the video on how bacteriophages can help make useful vector molecules. An example of this is when making antibody medicines. Antibodies are Y-shaped proteins that are produced by the immune system to help stop intruders from harming the body. This video shows how the DNA that codes for antibodies is inserted into a bacteriophage and then a library of antibody phage molecules is created.
### Activities/Procedure (50-80 minutes)

1. Obtain a stoppered vial, representing the bacteriophage virus, with a DNA sequence inside to code for two amino acids. Remove the DNA sequence containing 6 letters.
2. Three DNA letters represents one amino acid.
3. Consult the DNA translation chart to determine which amino acids will be produced. First you will “transcribe” DNA to RNA. Then you will “translate” the RNA to the amino acid. Start in the middle of the chart with the first of the three RNA letters, then move in sequence, to the outside with the remaining letters. The outside of the chart includes the amino acid abbreviation for that three sequence of letters.
4. Consult the Amino Acid chart to determine the structure of your two amino acids.
5. Build the amino acids with molecular model kits and identify bond characteristics - single, double, polarity, solubility and shapes around central atoms.
6. Record the information in the data table.
7. **Extension**: Combine the two amino acids together chemically by creating a **peptide bond** - condensation reaction. Then add a sugar molecule, glucose, to an amino acid by forming a **glycosidic bond** with the amino acid and a sugar molecule - dehydration synthesis.

### Summary

Ask students to write a “six-word story” about the chemical bonding in amino acids. The story can involve either covalent bonds or intermolecular forces. A six-word story is a summary strategy to condense a lesson’s material into literally six words. Ask for a number of student responses before dismissing class.

### Assessment

Students will use the provided peer assessment rubric to assess other amino acids that a classmate has created.
**Extensions/Connections:** IB or AP students, or students who have studied some biology or biochemistry:

- Students can chemically combine the amino acids through a condensation reaction. A condensation reaction is shown below:

  ![Condensation Reaction](https://goo.gl/6BxDVE)

- Students can create the glycosidic bond through dehydration synthesis. The students will create this bond with one of their amino acids (the OH group) with a sugar molecule shown below.

  ![Glycosidic Bond](https://en.wikipedia.org/wiki/Glycosidic_bond)

- Students can also choose different monosaccharides to add to the amino acids. Examples of monosaccharides students can research are glucose, mannose, galactose, and fructose.
Answer Key:

Example answers for amino acids alanine and asparagine.
DNA coding for amino acids can be found at https://en.wikipedia.org/wiki/DNA_codon_table

DNA Translation Chart:
To decode your DNA, you must first convert your DNA to RNA.
Write out your DNA Sequence: ___CGATTTG________________________________
Split your 6 letters into two 3 letter parts. Each 3 letters represent one amino acid.
DNA Sequence 1: _____CGA_________________________
DNA Sequence 2: _____TTG_________________________

For each 3 letter sequence complete the following:
Convert your DNA sequence to RNA Sequence by the process of “translation” where the DNA sequence changes as follows:
   “A” becomes “U”,
   “C” becomes “G”,
   “T” becomes “A”,
   “G” becomes “C”
RNA converted Sequence 1: _GCU_________________________________________
RNA converted Sequence 2: _AAC_________________________________________

Now, convert your RNA Sequence to the amino acid it codes for using the chart below. This is called translation. Start with the first letter of your sequence in the middle of the chart and then proceed to the outside. The amino acid you will build is represented by a three letter abbreviation. Check the amino acid chart to determine the name of your amino acid.
Start in the middle of the chart with the first of the three RNA letters, then move in sequence, to the outside with the remaining letters. The outside of the chart includes the amino acid abbreviation for that three sequence of letters.

**Amino Acid #1: GCU becomes Alanine (Ala)**

**Amino Acid #2: AAC becomes Asparagine (Asn)**

**Amino Acid Structures:**
Each amino acid has an amino end ($\text{NH}_3^+$) and a carboxylic acid end ($\text{COO}^-$). The distinguishing parts of amino acids are called “R” groups. The hydrogen bonded to the N can come off and bind to the COO as well. So NH$_2$ and COOH are also written. There are 20
amino acids each with common amino and acid groups but differing R groups. The R groups in the following image are separated from the common pink areas representing the unique amino and acid components.

Image: Cell Biology OLM, [http://goo.gl/2YRrC5](http://goo.gl/2YRrC5)
Molecule Characteristics:

<table>
<thead>
<tr>
<th>Amino Acid Name</th>
<th>Alanine</th>
<th>Asparagine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular Weight</td>
<td>89 g/mol</td>
<td>132 g/mol</td>
</tr>
<tr>
<td>Number of single bonds</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Number of double bonds</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Shapes around R group atoms</td>
<td>Tetrahedral around the carbon in the R group</td>
<td>Tetrahedral around the first carbon, trigonal planar around the second carbon. Trigonal pyramidal around the nitrogen.</td>
</tr>
<tr>
<td>Intermolecular Forces present in R groups</td>
<td>London Dispersion</td>
<td>London Dispersion, Dipole-Dipole and Hydrogen Bonding</td>
</tr>
<tr>
<td>Nature of R group (polar / non-polar)</td>
<td>Non-polar</td>
<td>Polar</td>
</tr>
<tr>
<td>Solubility of R group in water</td>
<td>Not soluble in water</td>
<td>Soluble in water</td>
</tr>
<tr>
<td>IB or AP Extension: Bond angles around R group atoms</td>
<td>Bond angle around carbon in R group: 109.5°</td>
<td>Bond angle around first carbon is 109.5°, second carbon is 120°, and nitrogen is 107°</td>
</tr>
</tbody>
</table>

Model of Alanine:  
Model of Asparagine

**Extension:** Molecules in the chemical library are modified by adding sugar molecules to the amino acids. After the bacteriophage produces the polypeptide (usually containing around
seven amino acids), chemists chemically alter the molecule by adding sugar molecules. These sugar molecules help by allowing the molecule to bind stronger to the protein of interest.

**Procedure:**

1. Using the model kits, combine the amino acids together through a condensation reaction as shown below.

![Diagram of peptide bond formation](http://goo.gl/eMJ30p)

   - **Amino acid (1)**
   - **Amino acid (2)**
   - **Peptide bond**
   - **Dipeptide**
   - **N-terminus**
   - **C-terminus**
   - **Water**

2. Build this sugar molecule, glucose, and chemically add it to one of the amino acids. You will need to research the reaction mechanism; however, use the name of the reaction to guide you. **The second molecule is representing the OH part of the amino acid.**

![Diagram of glucose addition](https://goo.gl/bSYv89)