Molecular Modeling in Medicinal Chemistry

The process of developing new drugs involves many types of researchers over a long period of time (close to 20 years). Most of the drugs target proteins, the molecules that carry out the different functionality in living organisms and virus. The first step when dealing with a new disease is to identify the cause: is it an organism such bacteria or virus? Or is it some internal pathway that is misregulated (e.g., in cancer)? From there on the goal is to identify some target protein that is vital for the disease to progress (e.g., replication of new virus/bacteria) and inhibit them.

Inhibiting them means coming up with a molecule that can displace the native interaction from taking place. To develop such molecules one could try out different molecules and hope one works — but the chemical space is so large that the chances of this are slim. Thus, in the Drug Discovery stage, molecular modelers typically identify the possible binding sites in the target protein, and the chemical footprint required for a small molecule to interact with this target. Nowadays people rely on docking of massive virtual libraries of potential compounds. However, in the old days, docking was performed by hand — through expert molecular modelers that had developed chemical intuition about which groups (hydrophobic, polar, charged) would be best to interact and certain sites. Even today, once an initial pro-drug scaffold is found, medicinal chemists are able to formulate new improved molecules (solubility, adsortion, …) by just looking at it.

In this exercise, you will take on the role of a molecular modeler using VR as your modeling tool.

By the end of this modeling activity you should be able to:

- Navigate the protein data bank (PDB) to identify protein targets of interest.
- Use Nanome to open, visualize, and handle molecular systems.
- Draw and design new molecules in Nanome.
- Compute hydrogen bonds, steric clashes, electrostatic potential maps and ligand scores.
- Mutate protein residues.
- Put your chemical knowledge in action to design peptide and small molecule ligands.
- Understand the role of molecular modeling in drug design.

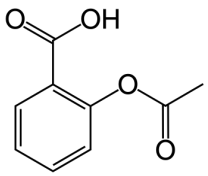
PART I: Building Molecules from Scratch

Rather than loading a new molecule, we are first going to learn how to draw our own. This will give us a chance to familiarize with the commands. For more details about the working, Nanome features several interactive tutorials within Nanome that you can use (Find them on the bottom left corner of the Main menu using 🎓).

Also make sure to check the Controls page on Canvas to see the functionality associated with each button on the controller.

1)

You will draw the molecule of the aspirin molecule (if you do not know it, search for it on a web browser. You can access a web browser inside Nanome from the menu on your left wrist).



To draw it press and hold the *B* button on the right controller to bring up the *Tools Menu*, then select *MedChem*.

You will see that you have the ability to "build" new molecules. Aspirin has an aromatic ring, so go ahead and grab a benzene. To draw it into your space you will just have to press the trigger button on the right hand.

Once it is loaded: grab it, rotate it, scale it up and down, be comfortable with those commands. They are mostly intuitive, but you can always look at the demos for help. Notice that the same way that you can move molecules, you can move menus about the 3D space (just check that the lock button at the top right corner of the menu is not on).

Now continue to add the other chemical groups into the molecule by simply pointing to the atom and dragging it from one atom to another. Notice that for some C—O you will need to add a double bond. As you add the different atoms, you will notice that the program optimizes the geometry (sp,sp2,sp3, and distance between atoms).

Going to the prep menu on your left hand —> auto add hydrogens will add hydrogens to your molecule. If you don't see all of them, go to the bottom menu (three dots) —>more settings —> show all hydrogens.

2)

Let's check some basic analysis functionality with this molecule.

Go to the modify —> tools —> measure functionality

Using your right joistick you can decide what type of measurement: distance, angle, dihedral, delete…

So if you set distance, you can measure the distance on all your C—C bonds, C—O. Go ahead and write them all out according to your measurements.

(https://www.youtube.com/watch?v=LC6SFHCXz_A&ab_channel=Nanome)

Then calcualte angles for sp2 and sp3 atoms, do they have the expected geometry?

Finally, select two dihedrals: one where all four atoms are inside the ring, and one where two atoms belong to the ring and the other two correspond to the longer R chain.

Now that you have measured them, select the modify—> tools—>dihedral functionality.

As you pass your right hand through different bonds in your molecule you will see either a red or white circle. The red circle corresponds to dihedrals that you cannot modify it (they belong to the ring), white ones you can freely rotate about. As you do so, you will note that some of the oxygens are coming very close to each other — this does not look like a good conformation!

3)

Let's compute some properties on your molecule:

Modify —> tools —> compute. We will calculate the different available properties: clashses, hydrogen bonds, and ESP (electrostatic) potential.

Starting with clashes, rotate the dihedral around and you will see that when atoms are too close to each other, there is an steric clash (represented by a yellow wall).

You can always carry out a minimization of the system to alliviate some of these issues.

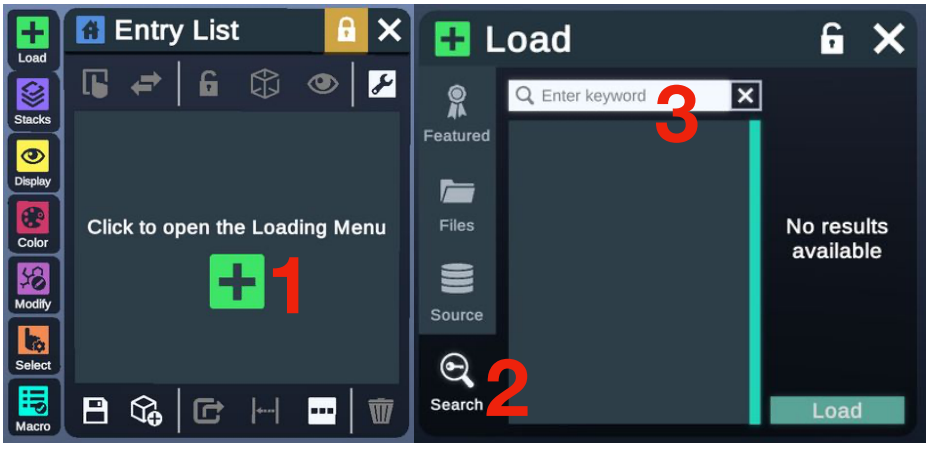
For this system it is unlikely that any h-bonds are present, but you can nonetheless compute them — we will use this later when we work on proteins.

Finally, the ESP calculation will create a surface representation of your molecule, where the color code is indicative of the electrostatic potential at that site on the surface. You will see that for this system you have white (hydrhopobic, no charge) sites and red (negative potential). This is useful to determine the ligand fingerprint. You can already see that this molecule is unlikely to interact with proteins that have highly negative binding sites — on the other hand it will likely interact at sites that are partially buried (hydrophobic interactions) and where there are possible positively charged residues such as arginine or lysine.

In order to display the ESP surface, you might have to select it from the *main menu*, then go to *display*, and select surface representation.

4)

Let's put this molecule on the side for a second and see a different way to create an aspirin system in nanome. Go to the main menu —> load new molecule, on the search bar type aspirin. And you should be able to load its structure from the drug bank. This is already optimized, so you can see how it compares to the conformation you were drawing.



PART II: Drug Design

Tasks in VR:

Since we were working with aspirin in the previous part, let's see if there are any proteins in the PDB that have aspirin bound to it. There are several, we have chosen 1OXR since the associated paper makes an interesting connection between aspirin's anti-inflammatory response and this protein-aspirin interaction.

1. Use the load menu to fetch the 1OXR structure from the PDB:

as soon as you load this molecule you see several things: first of all you have a representation that it's partially what we call new cartoon. You can see the beta strands in yellow and alpha helices in red. Additionally, you will see atomistic details (ball and sticks representation) of both the aspirin and residues that are interacting with the aspirin.

This might be a little bit overwhelming, if you go to the main panel, you will see that there is a list of units created 1OXR is the main one (the pdb), and it decomposes into protein, ligand and solvent. By hovering over any of those selections you can change how each is displayed or colored and it will also glow so that you can identify what selection you are working on.

2. From the *Entry List*, select the protein and create a surface representation, color it by atomic properties — it should show red and blue for Oxygens and Nitrogens. Can you identify any cavities in the protein surface? Small molecule ligands typically interact in these cavities. You can see aspirin is in one of them.
3. Calculate different properties like clashes, hydrogen bonds and electrostatic potential. Are there interactions between the protein and aspirin? Describe them.
4. Let's take a look at some helices and strands… what hydrogen bonds do you see? How are they stabilizing the system? Are they all the same length? Write out at least 6 h-bond distances in your system (helices and strands). Change your pointer to selection mode to identify what residues are involved in the hydrogen bonds (selecting them in turn). The residue numbering will allow you to identify if it is an alpha helix, a 3-10 helix or pi-helix. For the beta strands, notice the alternating pattern of interfacing atoms between the hairpins, explain this in terms of the Ramachandran plot for phi/psi dihedrals (You have already seen the Ramachandran plot in task 4).

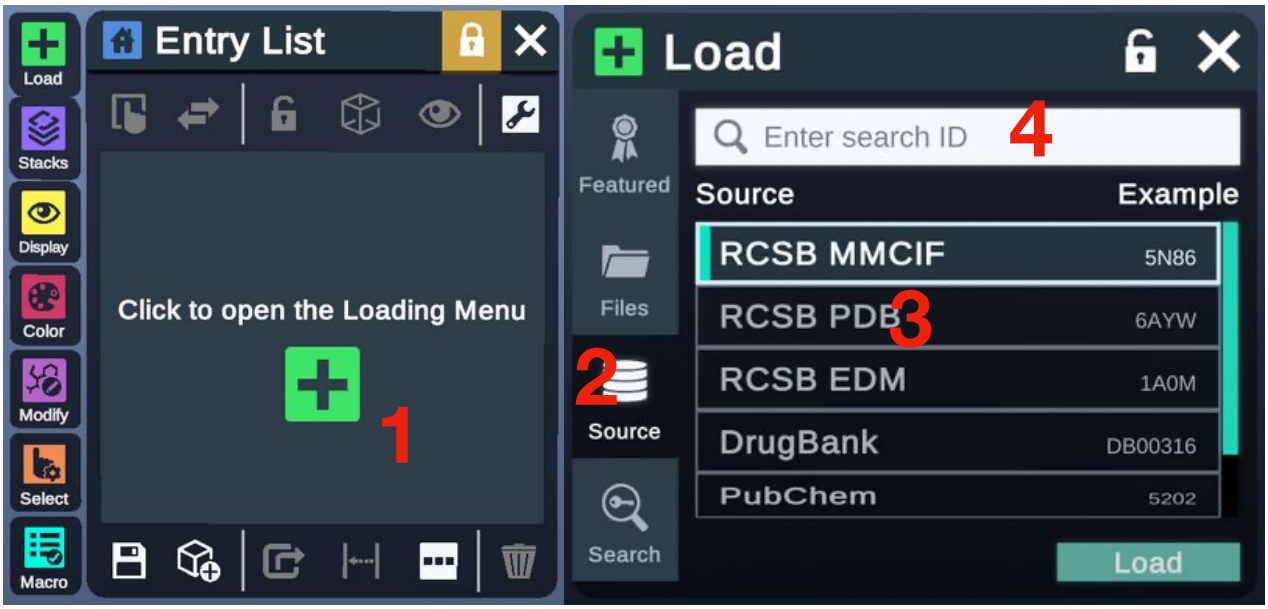
You might see some dashed lines For hydrogen bonds that seem to go nowhere… in representations look for ligands or solvent/ions. They might not be displayed in your current setup. Can you select them and change their representation to some atomistic representation so we can "see" them? Now all dashed lines should start/end in an atom. You might not readily recognize water molecules as they will likely only be represented by the oxygen (unless you have added hydrogens).

5. Since proteins are often surrounded by water, why do you think there are so few waters present here? Provide a feasible explanation. Hint: think on an experimental technique such as x-ray crystallography, which relies on well defined electron density (atoms static in one place in space).
6. Before moving to the next exercise, we will introduce some new functionality. From your dealings with proteins you might have heard that some residues, such as proline, are considered helix breakers. Let's mutate one residue from a helix in the protein into a

proline. First use the selector set to residue and select a residue you would like to mutate. Once it is selected (it should have a glow around it) look to your left wrist and select mutate —>proline. That's it!

7. Calculate clashes on the structure and you will see several, as well as a lost hydrogen-bond interaction. Even if you try to minimize the structure, the clashes will likely remain. The proline sidechain makes a cycle of 5 involving the backbone and heavily restricts available conformational preferences (phi/psi) for this residue to avoid clashes.
8. Design activity: now that you have seen how to navigate with nanome, and you are becoming familiar with the structure and protein-ligand complementarity, it is time to put your chemical knowledge to the test. You will be designing a new ligand that fits better than aspirin in the binding site.

Rather than start modifying our current aspirin molecule, we will create a duplicate of it, so we can always go back to compare to the original.



9. In the *Entry List*, select the ligand row, then go to menu —> split. This will allow you to move the aspirin independently from the protein (try it!). Now, with the aspirin selected we are going to hit the duplicate option. We now have a second aspirin molecule. Use the editing commands you learnt at the start to add/modify the functional groups and assess the quality of your new structure to bind the protein receptor.
10. If at any point you want to superpose the ligands onto the protein, go to the home menu, select the protein and two ligands and then hit the alignment button at the bottom of the menu. From the stacks menu you can even score how well the different ligands fit in the cavity (only with plugins??)

Report on your best designed structure. What made it a better drug candidate?