

Yasara course – October/November 2012

Start with YASARA

The program is already on your desktop.

→ Click on the icon to start YASARA

Let's start with opening a PDB-file. The program can work with several different formats such as .job and .sce, but we will look at pdb-files first. Of course you will need to download a PDB-file first, but YASARA always has 1 test-file in its PDB-directory, which we will use now:

→ Go to "file", "load", "load PDB-file", which PDB-code do you see?

You will enter computer's file-system, you usually know where your files are, if you cannot see a file, make sure it has the correct extension (.pdb in this case) otherwise YASARA cannot find the file. You can browse through all your directories. Clicking on the little YASARA icon will bring you to YASARA's directory directly and here you can go to YASARA's own PDB-directory. (yasara/pdb)

→ Go to the YASARA-directory, go to the "pdb"-directory and choose "1crn", click "ok".

A small protein will appear on screen, this is crambin, our example-protein. We will use it as a test-case to try all our commands and actions. Now your screen should look something like this:

The protein: 1crn

Quick-buttons to clean the screen, load a new file, undo/redo, select atoms, rotate and center

Top menu

ATOM PROPERTIES

Number: 1
Name: C
Element: C
Occupancy: 1.000000 % BFactor: 0.000000
Residue:
Object:
Position: X = 0.000000 0.000000 Å
Y = 0.000000 0.000000 Å
Z = 0.000000 0.000000 Å
Speed: X = 0.0000000000 m/s
Y = 0.0000000000 m/s
Z = 0.0000000000 m/s
Total = 0.0000000000 m/s
Active X = 0.0000000000 FN
Forces: Y = 0.0000000000 FN
Z = 0.0000000000 FN
Total = 0.0000000000 FN
Bonds:
1) Type: to ()
Length: Å
2) Type: to ()
Length: Å
3) Type: to ()
Length: Å
4) Type: to ()
Length: Å
Marked Distance: Å
Marked Angle: °
Marked Dihedral: °

SCENE CONTENT

Obj	Name	Vis	Act	Atom
1	1crn	Yes	Yes	1
2		No	No	
3		No	No	
4		No	No	
5		No	No	
6		No	No	
7		No	No	
8		No	No	
9		No	No	
10		No	No	

60 Hz Mem98 100% L

YASARA *Simulation*

im Off Obj all 327 Atoms

EN 11:02 25-10-2012

ATOM-properties menu: shows the properties of the selected atom, such as element, coordinates, B-factor, etc

HUD: shows all contents of the soup

- Try the quick-buttons on the right-top of the screen, center the protein, make it rotate, select some atoms and undo/redo this action.

Basic skills

The basic manipulations of the protein are all done using your mouse. Remember this:

Left-click and move = grab and rotate the protein

Right-click and move = zoom in/out

Both mouse buttons/scrollwheel = grab the protein and move it in space

- Play with these manipulations, notice that you can always center the protein again by using the quick-button for “center” in the top-menu

Visualizations, scene-styles

The protein is still a collection of balls now and it is difficult to see the separate residues. YASARA has different options to visualize the protein. You can reach these so-called scene-styles using the top menu, but it is easier to use the F1-F8 keys. Remember:

*F1: **Ball**, Shows all atoms as little balls with approximately 75% of their Van der Waals radii*

*F2: **Ball & Stick**, shows the atoms as smaller, separate balls and the bonds between the atoms as little sticks*

*F3: **Stick**, Shows both atoms and bonds as sticks*

*F4: **Trace**, Shows a trace between the C-alpha atoms of the residues*

*F5: **Tube**, Shows the backbone of the protein as a smooth trace*

*F6: **Ribbon**, Shows the backbone of the protein as a trace, but with extra emphasis on the alpha-helices and B-sheets. (favorite style for pictures!)*

*F7: **Cartoon**, As F6 but shows the helices as closed cylinders*

*F8: **Side chain on and off**, shows the side chains of the residues either as ball&stick or stick only.*

- Try all these styles, try the combinations with F8, Which one do you like the most?
- Go to “View”, “Scene style” and choose any of the styles in the list. These options are exactly the same, but it’s usually faster to use the F-keys.

These basic actions, visualizations and even more are explained in the help-movie. Yasara’s help-movies are quite good and can always be found in the top-menu

- Play the help movie, go to “Help”, “Play help movie” and choose movie “1.1 Working with Yasara”.

The movie will explain once more how to move and rotate the protein. It will explain that you can also use your keyboard for this, but in practice you will almost never need that.

- ➔ Play the help movie about Molecular Graphics; go to “Help”, “Play help movie” and choose movie “2.1 The Basics”.
- ➔ Extra hint: when you click on 1 atom and press the control key, you can rotate the protein around that selected atom. So, the selected atom stays in the center all the time.

Download other PDB-files -> The easy way

1crn is a nice test-case but we want to use other PDB files too. You can load any .pdb-file saved on your computer in YASARA, but on these pc's YASARA can also directly download from the PDB.

- ➔ To clean the screen go to “file”, “new”, you can also use the fast-button in the right top-menu to do this of course
- ➔ Go to “file”, “load”, “load PDB from internet”. Enter a “1hbo” and click OK. The program will ask you from which PDB it should download the file, use the default setting.

Using this method you just downloaded hemoglobin directly from the PDB, the file is not saved to your computer. Your own YASARA at home will probably not support this function. That means that you download the files you need from the PDB (www.pdb.org), save them as .pdb-file and load them in YASARA. For this course we will use the easy method.

The soup, objects, molecules, residues and atoms

To perform any action on a residue or atom you will need to select these atoms first. To understand how to do this you need to understand how YASARA actually works. Remember the following:

Soup: The soup consists of all the atoms in active objects in YASARA. When you make a selection using “all” you usually select everything that's in the soup on which that command can be used. So, you can think of the soup as “all”.

- ➔ The F-keys work on everything in the soup, so your basic selection is always “all” here. Try the different options once more. Do you see the little extra red balls that show up when you use F8?

Objects: Files loaded in YASARA always consist of at least 1 object. These objects are shown in the menu on the right of the screen. You can switch these objects on and off by clicking on yes/no in the column with the name “Vis”(which stands for visibility). When you load a file in YASARA it is automatically active. (how to make an object inactive is beyond the scope of the course). Sometimes a PDB-file consists of separate monomers that show up as one object, in other cases the monomers end up in separate objects. This depends on the definition in the PDB file. The hemoglobin structure is loaded as 1 object, but in contrast to crambin you know that it consists of multiple protein-chains. (hemoglobin is a tetramer). Using the options in the “edit” menu, you can change this...

- ➔ Once PDB-file 1hbo is loaded, go to “edit”, “split” and choose “all” (which is in this case the same as “object” because there is only one object). And click OK to center the newly created

objects. Look at the list in the HUD and see that we now have 22 objects. Switch them on and off by clicking on the “Vis” column in the HUD for that object.

Using the split-option you will place the separate molecules in separate objects.

- ➔ If you want to show only one of the objects you don't have to switch all the other ones off, you can simply right-click on the visibility column for the object that you want to see. This will switch all the other ones off. Use this option to see which objects contain the real protein-chains. Do you have any idea what we can find in the other objects?

One thing you also need to know (because sometimes you will do this by accident):

- ➔ Normally, you grab all atoms in the soup and rotate/zoom them. You can also select 1 object and move it by pressing the shift-button. Press the shift key several times and see what happens in the lower right corner of the screen. You will see “obj all” change into “obj 1” and “obj 2”. See what happens when “obj 1” is shown and you move the object around. Try the different F-keys and notice that the different scene-styles still work for all objects in the soup.

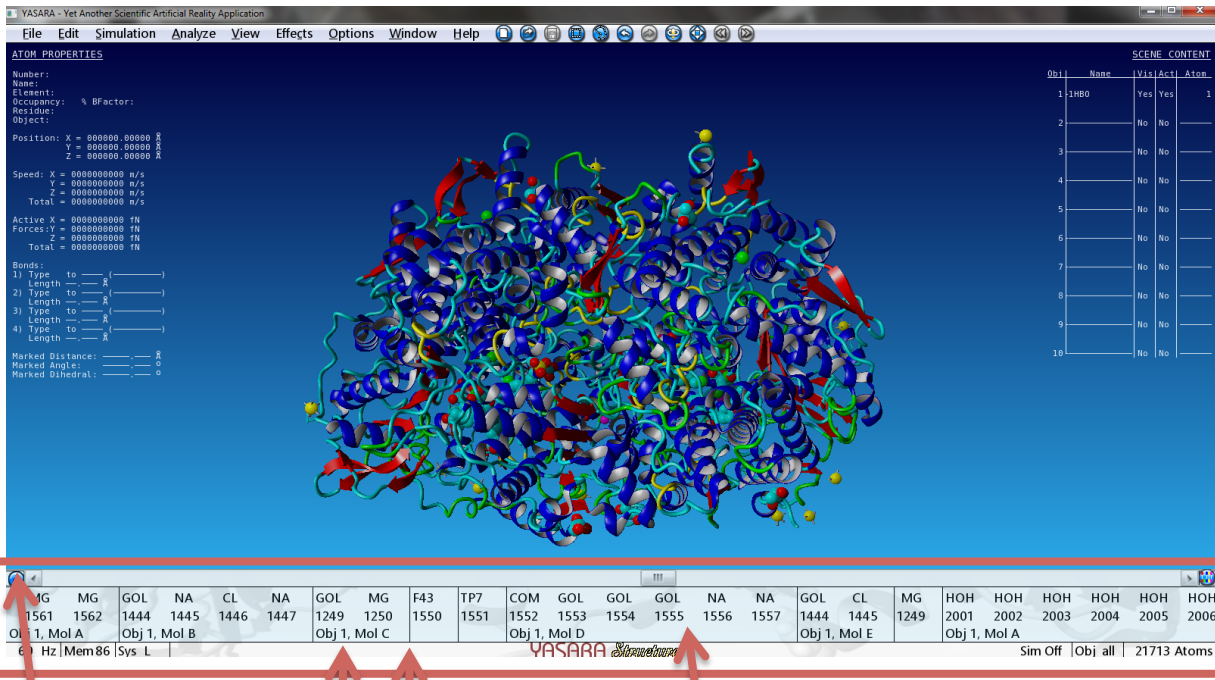
When one object is moving and the other ones are not...check the lower right corner of the screen, you might have hit the shift button by accident.

So, now you know how to manipulate the objects, let's have a look at molecules.

- ➔ Clean your screen and load PDB-file 1hbo again, show the side chains of the residues

Molecules: In general, structures such as protein-chains or ligands, that have different names are treated as different molecules in YASARA. They can be physically connected by a disulfide bond, as long as the chains have a different name, they are considered a different molecule. When a file contains 50 separate water molecules they are treated as 1 molecule because they all have the name “HOH”. Names of the molecules and residues are shown in the so-called sequence selector.

- ➔ Move the mouse to the lower part of the screen and the selector will appear as a ribbon. Pin it down. Now you can scroll through all the residues in the soup using the scrollbar.



Little button to pin down the selector

Sequence selector, shown as a ribbon, shows all residues in the soup

Number of the object and name of the molecule, separated by a single line

➔ Move the scrollbar in the sequence selector to the right and look at the names of the molecules in the lower left corner. The object number will always be 1 because we only have one object, the name of the molecule will change. Notice how the molecules are separated by a single line.

The names and order of molecules in a PDB-file is decided by the people that make and deposit the PDB-file in the PDB-database. At the end of 1hbo you will only find HOH-residues, which are of course water-molecules and are treated as a single molecule in YASARA.

- ➔ Let's delete all water molecules in hemoglobin. There is a shortcut for this: "Edit" "Delete", "Waters" will delete all HOH-molecules in the file.
- ➔ Split the object again ("Edit", "Split", "All"). Notice how the objects in the sequence selector now have different names and are separated by double lines.

Usually, the split function places different molecules in different objects although there are exceptions....

Residues: Consist of a continuous stretch of atoms that share the same residue name, number and molecule name. Peptide chains (molecules) consists of several residues. You are familiar with the 20 standard amino acids, but YASARA uses other names that are in the PDB-file are given to stretches atoms such as RET for retinol or LYR for a modified Lysine. And of course HOH for water. Sometimes the names for residues are unclear or misleading.

- ➔ Scroll through the residues in 1hbo. The normal peptide-chains have residues with the normal amino acid names. Notice how the residues all have their own number within that chain. A new chain (molecule) can start with residue-number 1 again.
- ➔ Try to find residue "F43", with number 1550. This F43 is the heme-group in hemoglobin and is actually forming the complete molecule (only one residue in this molecule). We can rename this residue with "Edit" "rename" "residue". You can select exactly the F43 at position 1550 in obj 7 in the list on the left, or you can select the residue type F43 in the middle. In the latter case, the F43 residue in object 12 will obtain the same name. Click OK and give the residue a new name. Notice how you can only use 3 characters, because residue names consist of 3 characters only.

Single atoms can sometimes be considered a residue in YASARA.

- ➔ Scroll to object 9, mol A and you will see COM, GOL and ZN, NA and MG. The latter "residues" are single metal-atoms, but they have a different name and are not connected, therefore they are different residues. Like objects you can delete residues separately or by residue-type. Zoom in on one of the Magnesium (MG) atoms. Go to "Edit" "Delete" "residue" and choose residue-type: MG. You can now delete all magnesium atoms in once.

Atoms: The smallest particles in the SOUP. In the F1 (Ball) visualization these atoms are all shown as separate balls with a color that is specific for the atom-type. Most used colors are: Cyan=carbon, Red=oxygen, Blue=nitrogen, Green=sulfur. Metal-atoms do also have their own colors, such as: Zn=magenta, Magnesium/Sodium=yellow, Chlorine=green. The ATOM properties are shown in the menu on the left-hand side of the screen.

- ➔ To delete an atom, such as a metal-atom that is not biologically relevant you can simply click on the atom to select it and press "delete".
- ➔ To scroll through all the atoms on by one you can select the first one and use the scroll wheel to go through all the atoms. You will see the little fireflies on screen jump from one atom to the other, see the ATOM-properties change accordingly and you will see the selected residue jump to the next one in the sequence selector.

You have seen that the sequence selector shows the contents of the soup. There is a 2nd method to see what's in the soup:

- ➔ Click on the object name in the HUD. It will expand to show all the different molecules in that object, click on one of these molecules and you will see a list of residues, click on a residue and you will see a list of atoms in that residue. Clicking on the object name will collapse the atom list again. Try this method to see the contents of hemoglobin.

You can use both methods to go directly to one residue/atom of interest.

- ➔ Find an atom or residue of interest in either the sequence selector or in the HUD. Click on the atom/residue name while pressing the control-button. YASARA will now go to that atom/residue directly and place it in the middle of the screen.

Selections

Many commands in YASARA apply to a selection. You can select everything in once, one of the objects, certain molecules, a special atom or residue-type, and more. Making correct selections can save you loads of work. Keep in mind YASARA's structure of objects/molecules/residues/atoms and think of what you actually want to do/see in the selected part of the soup.

There are several ways to make selections. Here we will use the color-command to demonstrate these methods, but in general this also works for showing/hiding/deleting/etc.

Method 1 – using the top-menu: In the top-menu you click on “View” , and go to “color”. You can now choose between “all” (which is easy because it will simply select everything), molecules, residues, etc. Click for example on “residue” and YASARA will show a list of all residues in the soup. Here, you can select your range of residues, such as residue 1-20, or all Tryptophan residues, or all residues with secondary structure helix. If you choose for “color”, “atom” you will see a list of all atoms in the soup. This gives you different options, such as coloring all residues in side chains, or all oxygen atoms.

- ➔ Clear your screen and load crambin again. Play with this method to color crambin magenta, all atoms in its side chains by element, the disulfide-bonds yellow and the first and last residue magenta.

With this method you can also use special selections.

- ➔ Choose “color”, “residue” and type in the empty bar: “all with distance < 10 from res 20” and choose a color. You will now color all residues that are located within a distance of 10 angstrom from residue 20. You can use this to color or show all residues within a certain range from a certain residue or a ligand.

Method 2-selection by clicking on an atom: You can also select and color atoms directly by selecting them in the soup on your screen.

- ➔ Look at crambin on your screen. Choose a residue and left-click on it. White fire-flies appear around the atom, indicating that this one is selected. Make a right mouse click on exactly the same atom and a menu will appear. In this menu you find options that can be applied to exactly that atom or to the residue/molecule in which that atom is located. Go to “color” and color that single atom.

Notice that it is not possible to color for example other residues with the same atom-type. You will need to use method 1. This is a quick way to select and manipulate only one atom. But you can repeat the command on other residues by pressing the “tab”-key and clicking on the other residues.

- ➔ Select and color residue 5 in crambin. Now hold the tab-key and click on another atom, you will see that the other residue will get the same color.

Method 3- Selection using the sequence selector/HUD: Sometimes there might be too many atoms on screen and it is difficult to find the one of interest. If you know the name or residue number you can use the sequence-selector or HUD.

- ➔ Go to residue 35 in the sequence selector or in the HUD and right-click on it. The same context-menu will appear as if you just used method 2. Notice that it is not possible to color a certain atom now, you can only use residue-selection because the sequence-selector works per residue. Color the residue.
- ➔ Repeat the command by holding the tab-key and clicking on other residues in the sequence selector/HUD.

Now you have seen 3 different methods to select atoms. We have used coloring as a test-command, but there are many more commands that need a selection of atoms. Let's try the different visualization options. Using the F1-F8 keys you can choose different visualization options. By default, these F-keys work on the soup. Let's try the visualization-options on a selection.

- ➔ Show crambin in Ribbon-style without its side chains. Go to "View" "Show atoms in", "Residue side chain &CA". Choose all residues with name "cys". You will now see that the side chains of all cysteines are shown...in crambin they form 3 cysteine bonds.

Showing a selection of residues with side chain is usually better than showing all side chains because your screen will be too crowded.

- ➔ Load PDB file 1tup. This is the DNA-binding domain of p53. Show the structure in ribbon-style and you will directly see which molecules are DNA. There are 3 protein-chains, which one is the best one to use for analysis? Delete the other ones. Delete the waters in the file.

DNA is always shown as a C-alpha trace in YASARA, simply because it does not have a secondary structure and therefore the ribbon/cartoon view would not be useful. In this view it does not really look like a protein/DNA interaction.

- ➔ Color the protein chain magenta, and change the scene-style to stick or ball&stick. You will now see all the atoms in the DNA, including the phosphate backbone. Now the interaction between the DNA and the protein side chains becomes obvious.

So, we want to show all atoms in the DNA and the protein in ribbon-style with some of the side chains that are close to the DNA. This is how it can be done...

- ➔ Show the scene in ribbon-style. Select an atom in the DNA molecule and right click to bring up the menu, choose "show atoms" and "molecule". You will now see all atoms in that molecule. Repeat this for the other DNA-strand (use tab). Check the Molecule names for the DNA strands. You want to show the side chain & CA for all the residues within 10 Angstrom from the DNA molecules. So, go to the "view" menu to "show atom in" "Residue side chain & CA" and try to show "all with distance < 10 from Mol E". You will have to repeat this step for all Mol F.

You will now see which residues are close to the DNA. Some of the residue are interacting with the basepairs in the DNA, others are interaction with the backbone phosphates. If you want to investigate these interactions closer you can show the hydrogenbonds

Hydrogenbonds

To show the hydrogenbonds we need hydrogens to be present. Most PDB-files are deposited in the PDB without hydrogens. This is not a problem because YASARA can calculate the correct positions of these protons. Let's add the hydrogens to our proteins.

- ➔ Go to "Edit", "Add", "Hydrogens to: All" You will see the hydrogens appear on all residue side chains in white.

Of course these hydrogens can fill your screen very fast. Let's hide the hydrogens. This means that we don't see them, but they are still present in the soup.

- ➔ Go to "View", "Hide atoms in", "hydrogens" and you will see that the hydrogens disappear.

Now we can add the hydrogenbonds

- ➔ Go to "View", "Show hydrogenbonds of", "All", and choose "Extend selection...". Now YASARA will show all hydrogenbonds.

The hydrogenbonds between backbone atoms are always shown. You can see the different hydrogenbonds between the turns of the helices and the hydrogenbonds that keep the strands together in a sheet. You can also see the hydrogenbonds from the side chains to the DNA-bases

- ➔ Zoom in on arginine 280. How many different hydrogenbonds does this residue make. How many of the hydrogenbonds are with the DNA? What about residue Lysine 120?
- ➔ Show all the side chains again, and hide the hydrogen-atoms. Look at the core of the protein. You see that only a few hydrogenbonds are formed between side chain-atoms in the core of the protein. This makes sense because most residues in the core of a protein are hydrophobic, so they cannot make hydrogenbonds at all.
- ➔ To hide all the hydrogenbonds you can either go to "View" "hide hydrogenbonds" "all", or simply delete all hydrogens by "Edit", "Delete" "hydrogens"

Saving pictures

Sometimes you want to show your work to others, or use your work in a presentation or on a website. Let's make some pictures. Of course you can use the old "print screen" trick to capture a screenshot....but YASARA has better ways to do it.

- ➔ Open a PDB-file or scene, color the protein grey, choose one residue to show the side chain and color this residue red. Rotate the protein in such a way that you have a nice view of the residue. Now go to "File", "save as", "normal screenshot". Choose a name and location, and check the box under "include menu". Save the file. Have a look at the picture.

You will indeed see a screenshot of the YASARA scene, including all the menu's. How to remove all the menu's?

- ➔ Use the same scene. Remove the sequence selector by clicking on the little pin-icon in the left top-corner of the ribbon. The sequence selector will disappear when you move the mouse. To remove the left and right-menu in your screen you press the "insert" key. Now the screen is

empty and a screenshot will not include these menu's. Go to "File", "save as", "normal screenshot". Choose a name and location, and uncheck the box under "include menu".

The screenshot will now show the protein only. The pictures will look more professional without the menu's. However, YASARA can do better. The normal screenshot-mode makes a 1-on-1 copy of the atoms visible on screen. Using a method called raytracing you can improve the pictures, add nice lighting/shadows, etc. YASARA uses an a program called POVray to do this (big movie companies like Pixar also use these programs.)

- ➔ Again, make sure you have something nice on your screen that you wish to save and show to someone else. Go to "file", "save as", "Ray-traced screenshot". A new menu with settings for POVray will pop-up. The default settings are OK, and you choose "Run POVray and save ray-traced image". Choose a location and name and save the file.
- ➔ If POVray is not yet installed on your computer you will receive an error message. The message will tell you how to install the program. You simply go to "help", "install program" and choose "POVray". Installing will only take a minute.

(of you want to read more about POVray you can always search in the YASARA documentation, go to "help", "search documentation" and search for povray.)

The POVray images usually look nicer than the screenshots and there are a few hints to improve them even more. Let's start with the background. In the POVray menu you can set the background to transparent by checking the box "transparent background". This will result in .png-files without background. However, sometimes a simple white background in your pictures is enough.

- ➔ Make sure you have a nice protein on your screen. Go to "View", "color", "background". And choose 1 color in the list or in the colorwheel. Click "OK".

The background is now white, and all pictures that you make using POVray will have this white background. For presentations and showing-off you might want to use a gradient.

- ➔ Go to "View", "color", "background". And choose 1 color in the list or in the colorwheel. Don't click "OK" but choose "continue with bottom color". Choose a different color and click "ok".

Your scene does now have a background with a gradient that runs from the top to the bottom. You can add up to 4 colors for the 4 different corners of the screen. If you want a gradient that runs from left to right, you give the top and bottom left corner the same color and the top and bottom right corner a different color.

- ➔ Play with this option, and make a picture with 4 different colors in the background, and make another picture with a nice black-grey gradient that runs from left to right.
- ➔ You can always reset the background. Go to "View", "color", "background" and choose "default".

Measurements in the soup

Sometimes you are interested in the geometry of a certain residue/atoms, such as the distance between two atoms, or the torsion angle over a bond. YASARA has easy options to measure (and change!) these features. What you need to know here is that you can select 1 to 4 atoms by clicking

on the atoms while pressing the control-key. You will see white, red, green and blue fireflies appear around the selected atoms. This option only works on separate atoms!

Distance

- ➔ Load crambin, color the protein grey and show the side chains, zoom in on Arginine 10. Select 1 residue in the side chain, now press the control key and select another atom in the side chain. The second atom will be marked with red fireflies and in the atom menu on the left you see a value appear for “marked distance” This is the distance in Angstrom between the two atoms. Try different atoms and notice that you can measure the distance between any two atoms.

You can also change the distance between two atoms. This only works for 2 atoms that are linked via 1 bond.

- ➔ Select two atoms in the side chain that are directly bonded. Right mouse-click on the first selected atom (the one with white fireflies) and the menu appears. Choose “Geometry” and “distance”. Now you can move the scrollbar in the box to increase and decrease the distance between the atoms. Try this on the atoms and also try to select the other atoms first.

Angle

Any 3 points form an angle. In YASARA you will need to select 3 atoms.

- ➔ Select 3 atoms in crambin by clicking on the atoms while pressing control. You will see white, red and green fireflies appear. The value of the angle in degrees will appear in the atom properties menu on the left side of the screen as “marked angle”. Try different atoms.

The value for the angle can also be changed. Again, the three atoms need to be linked by bonds.

- ➔ Select consecutive 3 atoms in the arginine 10 side chain. Right-click on the first selected residue to show the menu and go to “geometry”, choose “angle”. A new menu will pop-up in which you can move the scrollbar to change the angle.

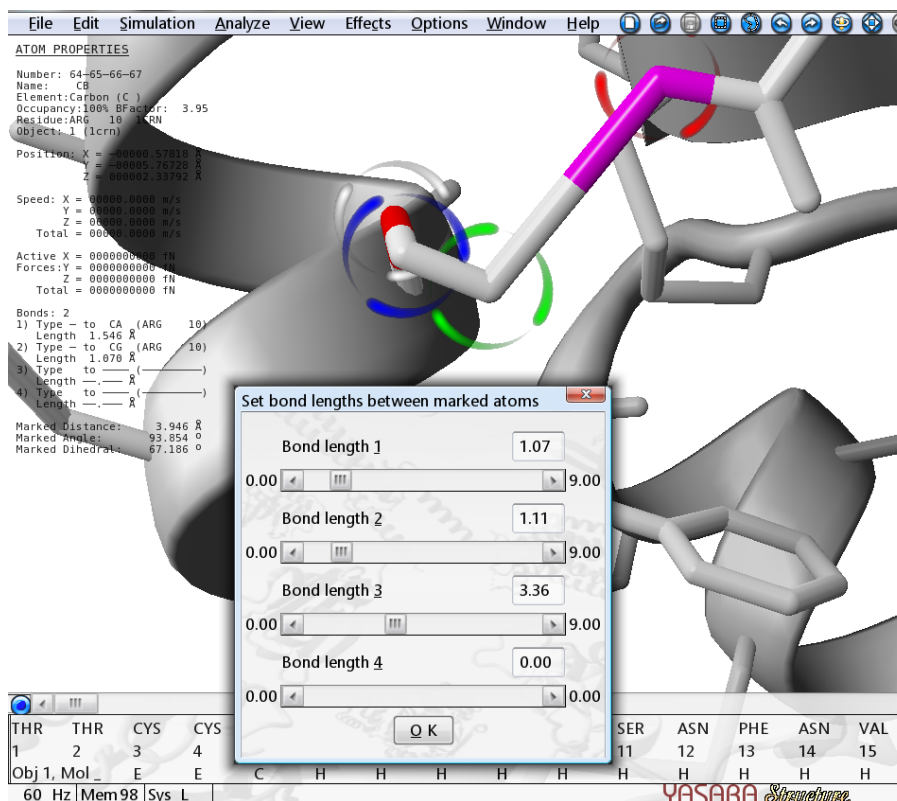
Dihedral/torsion angle

A dihedral is the angle between two planes. In proteins 4 residues in a row can make a dihedral because the first two residues and the last two residues make a plane. The dihedral angle is the angle between these two planes, seen over the bond between the 3th and 4th residue. To measure the dihedral angle, you need to select 4 consecutive atoms. The torsion angle will show up in the left menu and is measured over the bond between the green and blue fireflies.

- ➔ Select 4 consecutive atoms in the Arginine 10 side chain. Go to the “geometry” menu again and choose “dihedral. Use the scrollbar to change the torsion angle of the bond between residue 2 and 3. The value for the dihedral angle can also be found in the atom-properties menu as “marked dihedral”
- ➔ When multiple atoms are selected you can also change all the bond lengths and angles in once. Select 4 or even 5/6 atoms in the arginine side chain. Go to “geometry”, and choose

“bond lengths”. In the pop-up menu you can now change all the bondlengths between the atoms in once. (see pic)

Keep in mind that you can change all these values for distances, angles and dihedrals....even though they are biologically impossible. YASARA will use strict distances between atoms and only a limited number of rotamers to position sidechains. However, you can overrule these settings by hand using the described methods. The set of rules that YASARA uses to visualize biologically possible proteins is described in force fields. The optimal distance between two atoms for instance is a balance between repulsion between the two positive atom-cores and the attraction of electrons to the other core. Two hydrophobic residues have side chains that cannot occupy the same position, however, they want to stay close together because they want to exclude water. These , and many more rules are used for Energy Minimizations and Molecular Dynamics.



Surfaces

All scene-styles in YASARA are simplified views of reality. Even the F1 (ball) option shows balls with a radius that is smaller than the real Van der Waals-radius. If you want to see the “real” surface of a protein in YASARA you can use the surface-option.

➔ Load crambin again. Go to “view”, “Show surface of”, “all”. Here you can choose different types of surface.

Van der Waals surface: Outer surface per atom

Molecular surface: Is calculated using a water probe that rolls over the surface. The surface that can be touched by this water probe is called the molecular surface. In practice, this surface is bigger than the previous one.

Solvent accessible surface: The surface that can be reached by the center of a water-probe that rolls over the surface of a protein. In practice, this is bigger than the previous two.

If you want more explanation you can use YASARA's very extensive documentation ("Help", "Search documentation" and type in "surface") to read more about this.

- ➔ Choose "Solvent accessible surface", the other settings are all OK by default, and click "continue with surface color" (if you click "ok" in this menu the surface will be green) .

In the next menu you can choose a color for the outside of the surface. "Alpha" indicates the transparency of the surface (so, whether you can see the underlying structure or not). If you choose "continue with inside color" you can also choose the color for the inside of the surface, which you will only see when you zoom in until you are inside the surface.

- ➔ Play with these options. Try different colors, and different settings for alpha.

If you choose to color the surface by atom-color (atomcol in the list) the surface will receive the same color as the underlying atom. This can be very useful sometimes.

- ➔ Try to color the protein in one color, and a residue of interest in another color. See if you can color the surface by atom color so you can see the contribution of that separate atom to the surface.
- ➔ Switch off the surface by clicking "View", "Hide surface", "all"

You can also show the surface of a selections of one or more atoms, residues, molecules.

- ➔ Select one residue by selecting it directly on your screen. Right-mouse click on the selected residue and the menu will appear. In this menu you can also find "show surface". Try to show the surface for one residue, or one atom.

This option becomes very interesting when you are looking at a protein complex.

- ➔ Load 1tup again, delete the waters and the outer 2 protein-chains. Color the one remaining protein chain, and the two DNA molecules in different colors. Now you can see where the proteins touch each other.

YASARA does also have an option to calculate the buried surface area, but this is beyond the scope of this course.

Finding PDB-files

Although, the PDB contains many different PDB-files already, not every protein has a known 3D-structure yet. This is caused by the fact that it is very difficult to solve protein structures, whereas it becomes easier every day to find the sequence of proteins. How can you know whether your protein is solved and stored in the PDB?

The PDB is a database in which you can perform simple text-searches, such as “hemoglobin”. You will obtain a list of PDB-files that all contain the word “hemoglobin” in the file somewhere. This is not a guarantee that the protein actually is hemoglobin. We will now try to most secure way to find your protein of interest by performing a BLAST search against the PDB.

BLAST is a program that can compare your sequence of interest with many sequences in a database. This means that you need a sequence, go to a website that runs the BLAST program where you submit that sequence. The site will now compare your sequence with all the other sequences in its database and show a list of hits. The website that we will use is called MRS.

➔ Go to mrs.cmbi.ru.nl and choose BLAST in the top-menu.

You will see an empty box in which you can submit your sequence of interest. Let’s use the following sequence”

```
TTCCPSIVAR SNFNVCRLPG TPEALCATYT GCIIIPGATC PGDYAN
```

➔ Copy and paste this sequence in the empty box. Don’t forget to change “databank to search” into “PDB”. MRS can also BLAST against other databanks, but there is only one that contains 3D-structures. Click on the results and have a look at the results.

The list contains many hits for your sequence of interest. If you click on the red bar you will obtain more details about that hit. The most important thing for now is the identity score (see picture). A score of 100% means that the sequence that was found is 100% the same as your sequence of interest. In other words, you’re sequence has a solved protein structure.

➔ How many hits do have a 100% score for our query? Do they all have the same length?

Query Options Advanced Options

Enter one protein sequence in FastA format

```
>cram_craab
TCCPSIVARSNFNVCRLPGTPEALCATYTGCIIPGATCPGDYAN
```

Databank to search PDB Run Blast

Blast results

Nr	ID	Databank	Status
2	cram_craab	pdb	47 hits found

Blast results for cram_craab

Nr	ID	Coverage	Description	Hsps	BitScore	E-value
1	1yva.A		crambin; (PLANT PROTEIN); nmr solution structure of crambin in dpc micelles	1	93.2041	2.62e-27

Hsp nr	Alignment	Score	Bitscore	E-value	Length	Identity	Similarity	Gaps
1		230	93.2041	2.62e-27	46	46 (100%)	46 (100%)	0 (0%)

Q: 1 TCCPSIVARSNFNVCRLPGTPEALCATYTGCIIPGATCPGDYAN 46
 TCCPSIVARSNFNVCRLPGTPEALCATYTGCIIPGATCPGDYAN
 S: 1 TCCPSIVARSNFNVCRLPGTPEALCATYTGCIIPGATCPGDYAN 46

2	1yv8.A		crambin; (PLANT PROTEIN); solution structure of crambin in acetone/water mixed solvent	1	93.2041	2.62e-27
3	4fc1.A		crambin; (UNKNOWN FUNCTION); ultra-high resolution neutron structure of crambin at room-temperature	1	93.2041	2.62e-27
4	2fd7.A		crambin; (UNKNOWN FUNCTION); x-ray crystal structure of chemically synthesized crambin	1	93.2041	2.62e-27
5	2fd9.A		crambin; (UNKNOWN FUNCTION); x-ray crystal structure of chemically synthesized crambin- {alpha}carboxamide	1	93.2041	2.62e-27
6	2eyb.A		crambin; (PLANT PROTEIN); water refined solution structure of crambin in acetone/water	1	93.2041	2.62e-27
7	2eyc.A		crambin; (PLANT PROTEIN); dms0 refined solution structure of crambin in dpc micelles	1	93.2041	2.62e-27
8	2eyd.A		crambin; (PLANT PROTEIN); water refined solution structure of crambin in dpc micelles	1	93.2041	2.62e-27
9	2eya.A		crambin; (PLANT PROTEIN); dms0 refined solution structure of crambin in acetone/water	1	93.2041	2.62e-27
10	1ejg.A		crambin (pro22_ser22/leu25_ile25); (PLANT PROTEIN); crambin at ultra-high resolution: valence electron density.	1	93.2041	2.62e-27

Identity score

Using BLAST gives you an easy overview of the available protein structures that belong to your sequence. If you find multiple PDB-files that belong to your sequence you can have a look at all of them and compare the structures.

- ➔ Load 1ejg and 1crn in YASARA. The sequences of these PDB-files are exactly the same. We can “align” the two residues to see whether the structures are also exactly the same. Click “analyze”, “Align”, “Objects with MOTIF”. Select 1 object that will be aligned on top of another object, click OK, and select the other object on which the first one needs to be aligned. Click Ok and YASARA will align the two objects. Are the structures the same? If you cannot see the differences between the objects color them in a different color.

In this example the structures are as good as the same. But sometimes a protein is solved in an active and inactive state. Aligning the proteins gives insight in the movements that take place during activation. Proteins can also be solved in different complexes, or with different ligands. Aligning the proteins provides understanding of the movements that take place in the protein during ligand binding.

- ➔ Clear the screen and load PDB-file 1yv8. You will see 20 different objects of the same protein. This is a typical NMR structure.

NMR structure are solves as assemblies. These assemblies are useful to indicate the parts of the protein that have more freedom to move, and those parts that are stable. However, in general the quality of NMR structures is lower than that of X-ray structures and therefore we try to use X0ray structures when possible.

Saving your work: .pdb, .sce and .job

Imagine you studied a structure, you made some changes, deleted the junk and colored the file according to your wishes, but now it's 5 o'clock and you need to go home. There are several options to save your work.

- ➔ Load a protein of interest (try a random pdb-code for fun). Color the protein and use any visualization that you want. Now go to "File", "Save as", "PDB-file". Choose which of the objects you want to save (you can select more than 1). And choose a name and location.
- ➔ Reload the same pdb file and see what happened to your work.

Saving something as a PDB will only save the atoms and their locations. Colors and visualizations will be lost. Therefore, YASARA has other options. However, other file formats can only be used in YASARA. Other programs have their own formats or use the generic .pdb format.

- ➔ Repeat the last exercise. Make a nice protein with colors and some of the side chains shown. Save the protein using "File", "Save as", "YASARA scene" and choose a name and location. Clear the screen and reload this scene with "File", "load", "YASARA-scene"

The YASARA-scenes are saved with all options for colors and visualizations. So, continue your work in exactly, the same scene as you saved it.

- ➔ Make another scene by either downloading another file, or simply by giving the existing one a different color. Save this one too as a YASARA scene. Now, open one scene and try to open the second one too.

Only one YASARA scene can be opened at a time. You can add another PDB file to a scene, but not another scene to a scene.

- ➔ Add a PDB file to the scene. Make sure one of your saved YASARA-scenes is open and load another random PDB-file using "File", "load", "PDB-file from internet".

The PDB file is simply added to the scene as the first free object. YASARA has a second file-format that can only be used in YASARA: the YASARA-object.

- ➔ Color two of the objects on your screen, give one of them a nice surface, and go to "file", "save as", "YASARA object". Save the two objects with different names. Clear the screen and open your YASARA-scene again. Now load the two YASARA-objects using "file", "load", "YASARA object" (you can open 2 objects at the same time).

You will see that the objects are added to the scene. In contrast to the PDB-files the YASARA objects will have their own colors and styles. So:

YASARA-scenes: Contains all the objects in the soup, as well as all the visualization-settings, can only be used within YASARA, extension is .sce

YASARA-objects: Contains only one of the objects, including the style color and additional items (like labels), can only be used within YASARA, extension is .job

PDB-files: Contains the coordinates of the atoms only, can be used in other programs, extension is .pdb

Labeling

Sometimes you want to add a label or text to the scene to indicate a special point. .

- ➔ Clear the screen and load crambin again. Choose the ribbon-style and show all the side chains. Select residue serine 11 and rightclick on the selected residue to bring up the menu. Go to "label. Choose "residue". Now you can format the label. Check for example the boxes "Residue N (1 letter)", "residue number" and add as additional text "surface residue". Click OK.

The label will appear. It is "attached" to the C-alpha atom of the selected residue but you can select and move the label.

- ➔ Click on the label, fireflies will appear and now you can move the label with the left mouse-button and zoom in and out with the right mouse-button.

You can add anything you want in the label, change the colors and size.

- ➔ Select another residue to add a label, uncheck all the boxes but choose an additional text. Now you change the height of the label to 3.0 and choose "set label color". Choose a color and see what happens. Play around with the other options. Also try to label the object or molecule. Notice that in those cases it is not possible to label an object with residue names.
- ➔ Unlabel some of the residues but selecting the residue, right mouse-click "unlabel", "residue". Notice that the faster method is simply to select the label and press the "delete" button.

Like most options you can select any atom/residue that you want to label directly in the soup, using the right mouse-button. Sometimes you want a smarter selection.

- ➔ Remove all the labels. Go to "effects", "labels", "residue". Here you can select the residues that need to be labeled, such as "all", or residues 10-20. In our case we want to label the cysteines in crambin so choose residues with the name "cys" and click OK. In the next menu you choose to label these residues with their 1-letter code and number, make sure the labels are not too big and colored white.

So, labels can make it easy to find the most important features of a protein again. If you want to use the labels in a picture you have to make sure that they are positioned correctly. Of course, you can always choose to add the labels to the picture using a different program.

Animations

Sometimes you want to show all sides of a protein. YASARA has the option to save an animated gif of a rotating protein. You can use this animation on a website or in a presentation.

- ➔ Make sure the protein is centered in the middle of the screen, and switch of the menu's (insert-key). Go to "File", "save as" and "Animated GIF/MPEG clip". Choose a location and name for the files (it might be wise to make a separate directory). In the next menu you can choose the rotation-axis. Select "Y-axis", because this one is the most natural one. Do not

check “use ray-tracing” because this will severely slow down the process. Click OK and in the next menu you can change the settings for the animation. YASARA will rotate your protein a full 360 degrees and make a number of pictures as set in the “frames encoded” box. By default this number is set on 90 frames, meaning one picture per 4 degrees rotation. This results in a good animation, however 45 frames is also OK. For a real smooth movie you can increase the number of frames. You can change the X and Y resolution, but increasing the resolution will also increase file size. The default settings are usually OK, but set the number of frames on 45. Click OK.

Now YASARA will reduce the screen to the correct resolution (you can maximize the screen when the animation script is finished) and will make a bunch of pictures. If you saved the files to your desktop you will be able to follow this process. In the end YASARA will encode the animation which will convert all the separate pictures into one .gif file. If you look at the resulting .gif file you will see a rotating protein.

Energy Minimizations (EM)

A protein wants to fold into its lowest energy conformation. YASARA can perform real-time EM, so you can follow the process on your screen.

- ➔ Let’s change a protein-conformation in such a way that the protein will not be in its lowest energy conformation for sure. Change some of the bond distances and angles until the protein looks weird. Now go to “Options”, “Choose experiment”, “Energy minimization”. Follow the process on your screen, do you see what happens?

During an energy minimization YASARA will move the protein to a nearby lowest energy conformation. It is not possible to fold a peptide chain using this method. Do you see that YASARA moved the changed side chains back to their biological values? If you want to know more about this process, use YASARA’s documentation. Here, we will focus on the application of EM.

Imagine that you have a mutation and you want to study the effect of the mutation on the protein structure. In YASARA you can make the mutation and analyze the effect.

Swapping

- ➔ Clear the screen and load crambin again. Select residue Cysteine 4, right click on it and choose “swap”, “residue”. Choose a new residue type, for example “tyrosine” and click OK.

You just changed one amino acid type into another. In other words, you made an *in silico* mutation. You will directly see the effect on the side chain (bigger smaller, different conformation). But any side chain change will also affect the position of the backbone, other side chains, etc. To see the effect on the complete protein structure you will need to perform an EM.

- ➔ Go to “options”, “Experiment”, “Energy minimization”. See what happens. At first, only the protein side chains are moving, in a second step the backbone is moving too.

To compare the effect of the mutation with the wild type you can look at both proteins aligned.

- ➔ Load another crambin. Align the 2 objects using “Analyze”, “Align”, “Objects with motif”. Delete the hydrogens. (Edit/delete). Most of the protein structure will be exactly the same. Color everything grey, color the wild type side chain green and the new mutant side chain red. Do you see the changes in the protein due to the mutation?

Using this method you can make quick mutations in a protein structure and see the effect of the mutation on the conformation. This is very useful for researchers working in the genetics field where they are trying to understand the effect of mutations in a patient and the link to a certain syndrome or disease.

However, this only works when a PDB-file for the protein of interest is available. Even though, the PDB contains thousands of protein structures already only 20% of the known human genes has a solved protein structure.

Homology modeling

For more information about homology modeling you can simply google or visit the Wikipedia page: http://en.wikipedia.org/wiki/Homology_modeling The basics are:

- The structure of a protein enables the function of that protein. So, proteins that have the same conformation usually have the same function. (Exceptions exist of course)
- During evolution the sequence of a protein changes faster than the structure. This is caused by the fact that some amino acids are quite similar. A change of a Leucine into an Isoleucine changes the sequence at that position, however, the hydrophobic property of the residue is conserved and therefore the structure will not change (or only slightly).

In practice these rules imply that we can use the structure of one protein to predict the structure of another one. Imagine you have the sequence of a protein that you want to study but no experimentally solved structure is known. A homologous protein, one that looks similar to your protein, might be known, and we can use that homologous one to predict the structure of interest. Homology modeling is a multi-step process for which the details will be explained in another seminar. Here, we will show you how you can make homology models using YASARA.

To build a model in YASARA you always need at least a sequence of your protein of interest. This sequence needs to be saved in fasta-format in a separate directory. Let’s have a look at the following sequence.

```
>bacr3
MDPIALQAGYDLLGDGRPETLWLGIGTLLMLIGTFYFLVRGWGVTDKDAREYYAVTILVPGIASAAYLSMFF
GIGLTEVTVGGEMLDIYYARYADWLFTTPLLDDLALLAKVDRVTIGTLVGVDAIMIIVTGLIGALSHTAIAR
YSWWLFSTICMIVVLYFLATSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLR
LFMVLDVTAKVGFGLFILLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLR
```

- ➔ Make a separate directory (for example on your desktop) where YASARA can save all the files that will be generated during the modeling process. Save the sequence for BACR3 in this

directory. You can use a text-editor (use for example the terminal and "gedit" to open a text-editor) to save the a file. Make sure the file has extension ".fasta ". (A fasta-file contains a sequence in 1 letter-code, and on the first line you can always find ">name".)

YASARA will use this sequence to search for homologous proteins in the PDB. The program has an internal set of rules to decide which of the homologous proteins is the best one to use. Depending on your settings YASARA will use multiple templates to build different models. The models will be compared and the program will choose the best model...again, based on a set of internal rules. You can interfere with the process by, for example, choosing 1 template and forcing YASARA to use this one, or by submitting your own alignment. However, we will start by using the easiest option.

- ➔ Once you have made a separate directory and saved the sequence as a fasta-file you can clear the screen and go to "Options", "choose experiment", "homology modeling". YASARA will now ask you for the fasta-file with the model-sequence. Choose the fasta-file with your sequence and click "ok". In a next screen you can indicate the structure that the program should use as a template. We did not load a template in the soup, so the list is empty. This means that YASARA will perform a complete PSI-BLAST search to find templates. Click OK. Now you choose the location for all the files that will be generated. Choose the directory that you just made, and give the files a name like "model_". Click OK. The last menu contains the modeling parameters that can be changed to tweak the modeling process. However, the default settings are OK for now. Click OK, and the modeling process will begin.

YASARA will automatically perform all steps: template choosing, making the alignments, building new loops, energy minimizations, etc. You can follow the process live on your screen, but it can take a while before the model is finished. In practice it is wise to build models overnight.

YASARA will generate a modeling report in html-format. Which means that you can study them in your internet browser.

- ➔ Use your browser to open the file report.html to have a look at the modeling-process. Have a look at the list of templates, do you understand why YASARA choose to discard some of them? Have a look at the alignments and the different models that were build.

During the modeling process YASARA well extend the report until it makes the final model, which will be saved as model.pdb. This process can take a while. In the meantime you can have a look at the YASARA documentation for modeling. (If you cannot use the search function while YASARA is modeling you can open the file Bahm.html in the YASARA/doc/ directory in your browser directly. Another file with useful homology modeling hints is called "Uhmh.html"

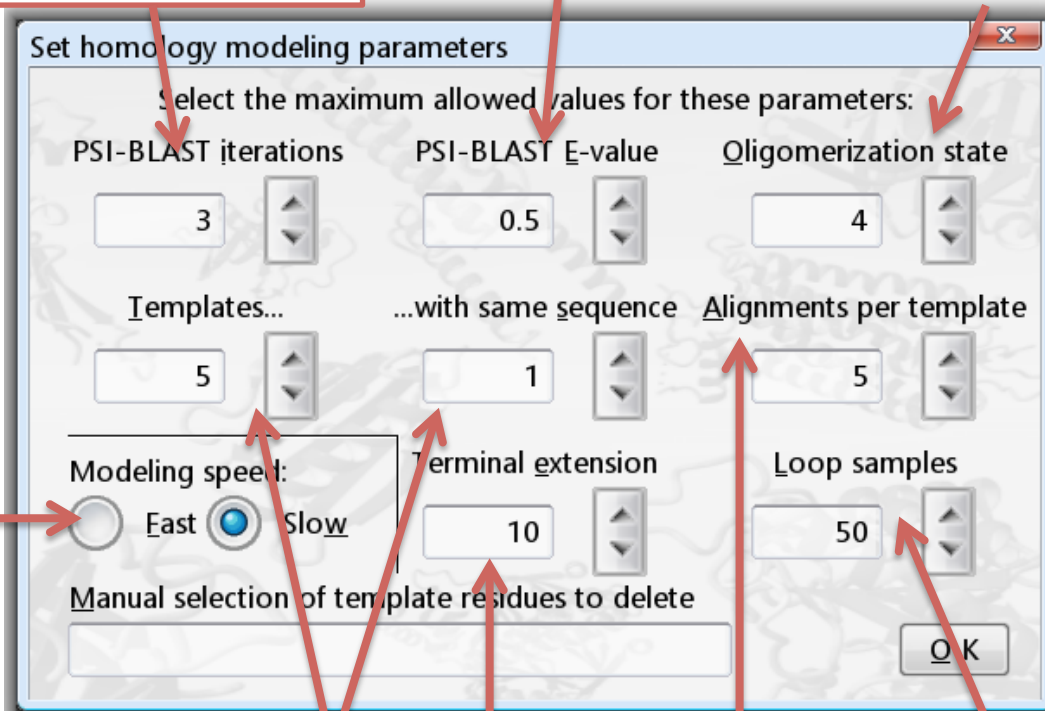
- ➔ Have a look at both files and see if you understand YASARA's modeling process.

As mentioned before, you can infer with the homology modeling process. In the modeling-parameter menu you can adjust some of the settings. Default settings are usually OK to produce a good model, but sometimes you want to speed-up the process, or model a multimeric state. Have a look at the menu in the picture:

PSI-BLAST iterations: increasing the number will enable YASARA to find more distant homologs, but will slow down the process.

E-value: indicates how distant the model and template can be. Increasing this value can help finding difficult targets.

Oligomerization state: by default YASARA maximally builds tetrameric states. If your protein consists of more subunits you need to increase this number.



Number of templates for which different models will be build.
By default only 1 template with a certain sequence will be considered. Increasing this number will increase the number of resulting models, will also slow down the process

By default a maximum of 5 different alignments will be tried per template.

YASARA can model residues without structural information in the template, by default max 10 residues will be added to the N- and C- termini of the model.

YASARA will try different possibilities to build new loops, by default max 50 different loops will be built.

YASARA can perform a fast modelling or a slow modelling process. In general you want to use the slow process because this will include an extensive template-search to make sure YASARA uses the correct template, and a long EM to remove errors. However, if you need the model directly, or you know high % sequence identity templates exist, you can use the fast modeling process. The PSI-BLAST search will be shorter, the EM will not be as extensive and you can still end up with a good model, or at least a good idea about the structure of the protein. Use with caution!

xAs mentioned before it can take a while for the model to be finished. In case the model-run finished you can have a look at the results. When the model-run is not finished you can stop and restart the process anytime you want. In case you want to restart the process, you simply start exactly the same model-run with the same sequence and parameters. YASARA will detect which files are already present and continue the run from there.

(for the next part: If your modeling run did not finish yet, you can have a look at our precalculated data. We have performed the same modeling run for you, the directory with the files can be found here: <http://www.cmbi.ru.nl/~hvensela/BACR3/>)

- ➔ Stop the modeling process in case the model-run is not finished yet, clear the screen. Have a look at the scene with the results. (your model.sce in the model-directory or the precalculated model.sce on the website). YASARA always ranks the models for you, the best one can be found in object 1. Open the scene in YASARA and study the different models.

Notice that the first model is not always the best one. This is caused by the fact that a single protein structure can be solved in complex with several other proteins, ligands and other conformations such as active/inactive or open/closed states. Therefore, it is always wise to look at the other generated models too. To get an idea about the quality of the model you can have a look at the modeling report which can be found in model.html.

- ➔ Open the modeling report in your browser. Scroll to the end of the report, here you can see how many different models were built, and how good YASARA thinks these models are. Which one is the best one? How many different templates were used, and how many different alignments were tried? Have a look at the best model, is there a big difference between the 1st ranked model and the 2nd and 3rd?

The end of the report summarizes the model process and indicates which model is the best one (according to YASARA). For more details per model, alignments and such, you need to scroll through the report.

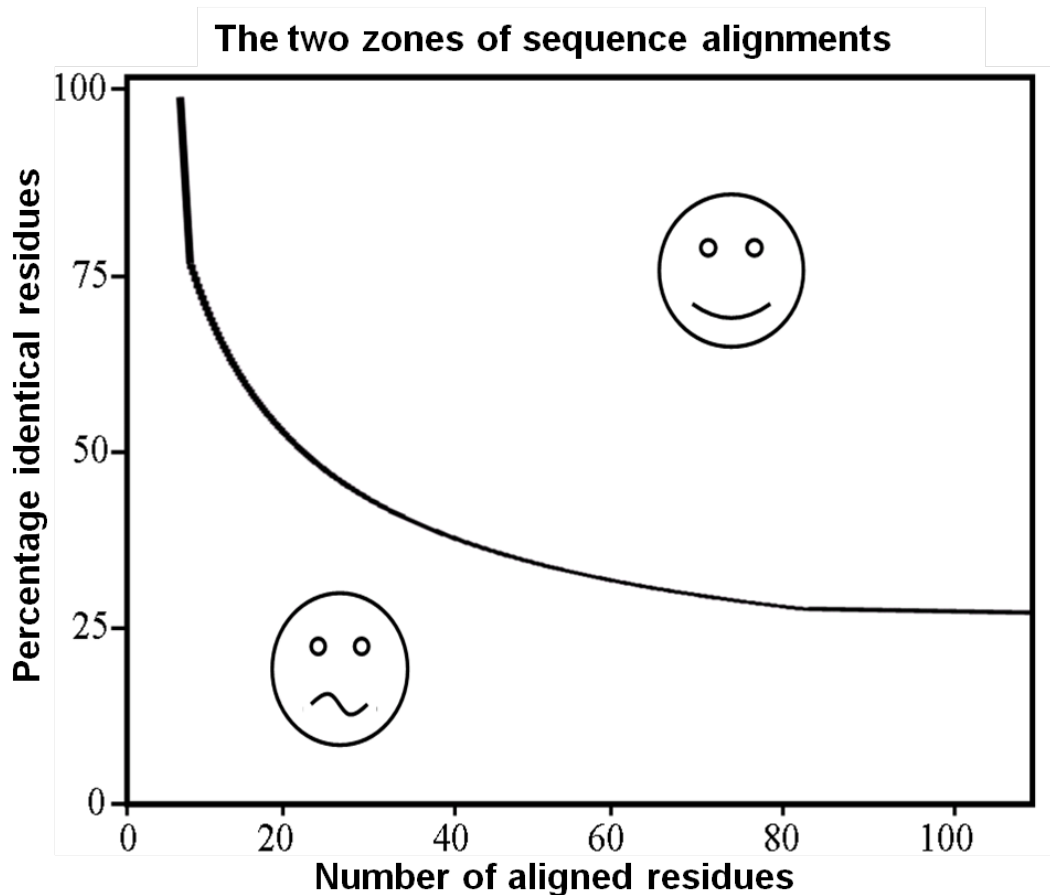
- ➔ Go to the paragraph that shows the details about the best model. Here you can find the alignments, the % sequence identity between the model and template, and the details for loop-modelling. Did YASARA build new loops for this model? And what is the % sequence identity?

The % sequence identity gives you a good indication about the quality of the model. More % identity between template and sequence means that the structure of the proteins are more likely to have the same conformation. 90% sequence identity will in general result in a very good model. But for longer alignments the sequence can be as low as 30%. Of course, in those cases the model can be totally wrong in some details, but you can still get a nice idea about the overall structure.

- ➔ The model for BACR3 can now be used to analyze mutations. Find residue 226, does it have a special function? Can you understand what happens when you mutate this residue?

When looking at models you always need to keep the %identity in mind. Models with lower % identity should not be used for very detailed analysis. To indicate when it is still safe to build a model we use the so-called Sander&Schneider plot. (see picture)

The Sander &Schneider-plot tells us when it is safe to build a homology model. Have a look at the length of your sequence alignment between the model and template sequences. A higher number of aligned residues means that you need a lower % identical residues. For example, an alignment of 80 residues needs 30% identical residues...if the number fall above the threshold shown in the plot you can assume that the model is OK.



Of course it is always wise to have a look at the alignment for the region you are interested in, for example the residues surrounding a mutation that you want to study.

YASARA's method for template-detection does not directly use this Sander&Schneider plot, although it includes the % identity in its decision. As a result YASARA sometimes builds models with very low % identity.... Be careful when you use these models.

It happens that YASARA cannot find a template at all. You can still force YASARA to build a model based on a template that you indicate. This happened once for the NDP protein. Sequence:

>NDP

MRKHVLAASFMSLSLLVIMGDTDSKTDSSFIMDS DPRRCMRHHYVDSISHPLYKCSSKMV
LLARCEGHCSQASRSEPLVFSFSTVLKQPFERSSCHCCRPQTSKALKALRLRCSGGMRLTATY
RYILSCHCEECNS

→ Make a new directory for this model, and ask YASARA to build a model for NDP.

In this case YASARA was not able to find a modeling template. You can force YASARA to use a template of your choice.

→ Load pdb-file 2kd3. This is an NMR structure, so you receive multiple objects, delete all objects except for the 1st one. Start the modeling process again. In the second menu you can now ask YASARA to use the pdb-file in the soup as a modeling template. Use the default settings, but set the modeling process on “fast” instead of “slow”.

In this case the modeling-run should finish quite fast. YASARA will skip the PSI-BLAST search and will only make models for 1 template. In this case we knew from literature that the PDB file could be used for modeling because other researchers studied to proteins and concluded that the structures should be similar, even though the sequences were not obviously homologous.

More about YASARA

You have seen many different options in YASARA, you can perform the basic skills that are necessary to study a protein in its 3D-conformation, you can make nice pictures/animations and you know how to do basic manipulations such as changing angles/distances and in silico mutations. We even showed you how to perform EM-simulations and homology modeling. There are many more options in YASARA which we did not cover yet. You can have a look at the help-movies to learn more about the other options too.

→ Go to “help”, “play help movie” and choose a topic that seems interesting to you. Follow the hints in the movie. For example, you can learn how to build small molecules (movie 4.1).

Don't forget that YASARA's documentation is very complete and detailed. In case you don't understand a term, or you simply want to know more about a topic use the “search-documentation” function under “help”.

That's it...good luck with YASARA!