

Working with Nucleic Acids

Part of the Jmol Training Guide from the MSOE Center for BioMolecular Modeling

Interactive version available at <http://cbm.msoe.edu/teachingResources/jmol/jmolTraining/nucleics.html>

Introduction

Most .pdb and .mol files will include protein structures and small molecules. Some, however, include **nucleic acids (DNA and/or RNA)**. If your structure file contains nucleic acid, there are a variety of suggested ways to display them. As with all of the **display formats** in Jmol, the nucleic acid formats can be mixed and matched to highlight different features of the nucleic acids.

Note that unless otherwise indicated, this section of the Jmol Training Guide uses the protein **Zinc Finger** paired with a short section of **DNA** based on the .pdb file **1zaa.pdb**. Please see the **Getting Started in Jmol** section of the Jmol Training Guide for information on how to download and open .pdb files.

Selecting and Displaying Nucleic Acids

Selecting Nucleic Acids

If you want to select **all** of the nucleic acid in your structure file, you can use the select nucleic command. However, nucleic acids also have a variety of other identifiers, including chain names (just like protein chains have), nucleotide numbers, and even base type ("g" for guanine, "c" for cytosine, "a" for adenosine, "t" for thymine and "u" for uracil).

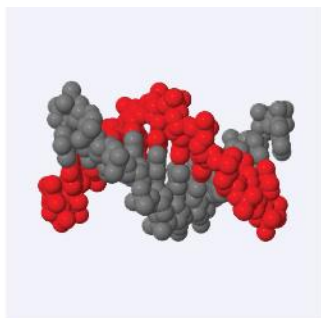
If you do not know the nucleotide number, chain letter, nucleotide type, or atom number of the item you want to select, you can click on the structure in the **display window**. Jmol will provide information on the **console window** regarding what you clicked on. This information can then be used with the select command.

Note that you can practice these commands yourself using your own copy of Jmol running on your desktop, or with the Jmol display to the right and the console to the bottom left. The nucleics in the file 1zaa.pdb will be shown with a spacefill of 1.5 and a wireframe of 1.0 - none of the protein structure in the file is shown. To reset the visualization to its default selection and representation, use the commands:

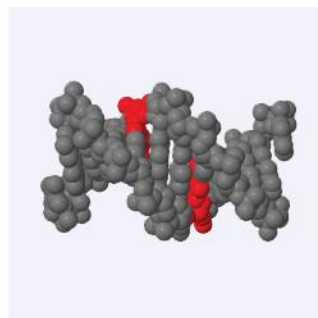
```
select all  
color gray
```



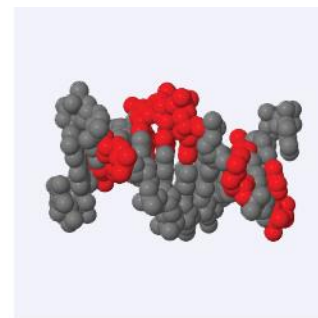
```
select nucleic  
color red
```



```
select :b  
color red
```



```
select 5  
color red
```



```
select g  
color red
```

Display Formats for Nucleic Acids

There are a variety of suggested display formats for nucleic acid sections of a molecular structure, each with their own strengths and weaknesses. Below is a description of each along with a Jmol button to demonstrate the commands using the Jmol display window shown to the right.

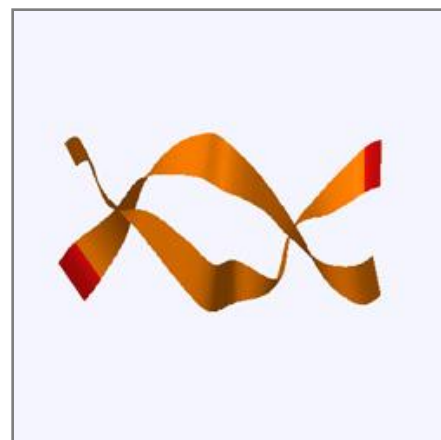
Note that we have used the **color cpk** color scheme for these display format examples.

Cartoon Format

This format shows the strand of nucleic acids as if it were traced with a ribbon. This version may be familiar as it is often seen in textbooks and creates a clearly identifiable double helix structure if used to represent double stranded DNA. Note that while cartoon is a nice way to display nucleic acids for molecular illustrations or interactive web pages, it **cannot** be used to build a physical model using 3D printing.

To turn cartoon format on:

```
cartoon on
```

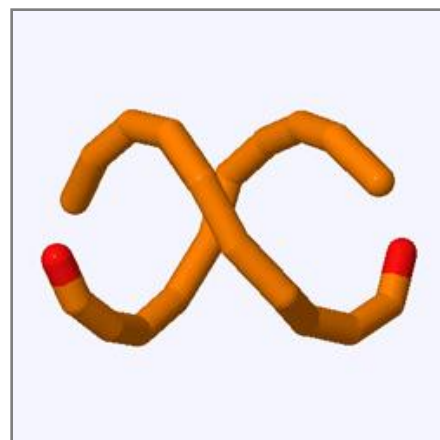


Backbone Format

For nucleic acids, the backbone format only displays the position of the backbone phosphorus atom in each nucleotide. All of the other atoms within the nucleotide are not displayed. Note that while backbone is a nice way to display nucleic acids for molecular illustrations or interactive web pages, it **cannot** be used to build a physical model using 3D printing **unless** you include **hydrogen bonds**, which will be discussed in more detail below.

To turn backbone on:

`backbone 1.5`



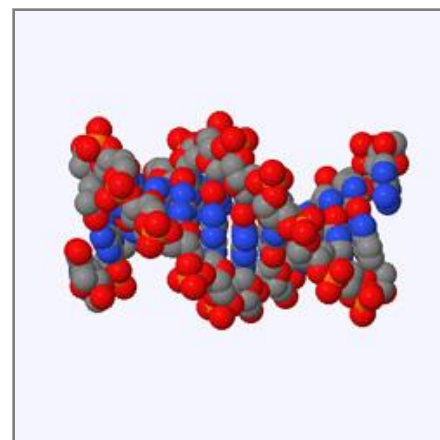
Ball and Stick Format

This format is a combination of **wireframe format** and **spacefill format** and creates a classic "molecule" appearance, with a sphere at each atom and a cylinder representing bonds between the atoms. Note that you can adjust the ball and stick format sizes to create the exact appearance you are interested in, but should make sure the spacefill is at least 1.5 angstroms and the wireframe is at least 1.0 angstroms if you plan to 3D print your structure.

To turn ball and stick format on:

`spacefill 1.5`

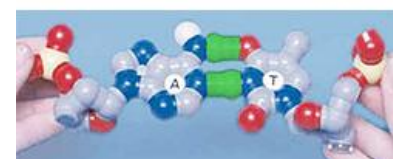
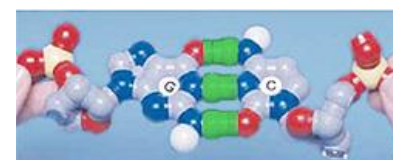
`wireframe 1.0`



Hydrogen Bonds in Nucleic Acids

Nucleotide bases pair together (A=T and G=C) with **hydrogen bonds** (shown in **green** in the image to the right). Because this base pairing is such an important part of nucleotide structure, it is often a feature you want to highlight when creating a Jmol representation.

There are two main ways to show hydrogen bonds in a section of nucleic acids, by connecting them to the atoms in the nitrogenous base of the nucleotide or by connecting them to the phosphorus kink in the backbone of the nucleotide. Which way you choose will depend largely on which display format you have used to represent your nucleics.



Hydrogen Bonds Between Nitrogenous Bases

The commands needed to add **hydrogen bonds** to a section of nucleics are similar to the commands used to add them to a protein structure. You begin by selecting the area of the structure you want to add the hydrogen bonds and then use the calculate hbonds command.

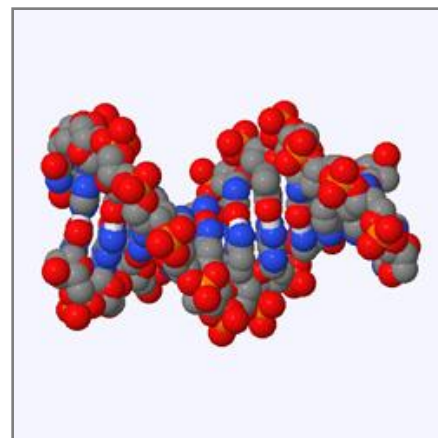
Hydrogen bonds, like wireframe, backbone, and spacefill, can be thickened by placing a number after the hbonds command. The standard size we suggest for building a physical model using 3D printing is 1.0 Ångströms.

The default display for hydrogen bonds is a dashed line. You will need to change this into a solid cylinder for building a physical protein model using 3D printing. You can do this using the set hbonds solid command.

Hydrogen bonds can also be colored using the same color command and color formats as other areas in a molecular structure. We suggest that you use a subtle color for your hydrogen bonds so they are not visually distracting or overwhelming.

To display hydrogen bonds on nitrogenous bases in nucleics:

```
select nucleic  
calculate hbonds  
hbonds 1.0  
set hbonds solid  
color hbonds white
```



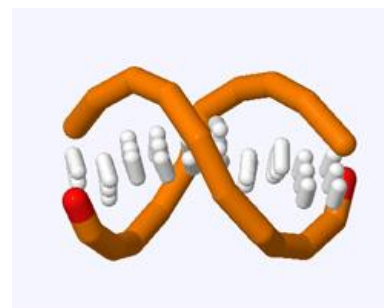
Hydrogen Bonds Connecting the Backbone

If you are using a representation of your nucleic acids that does not show the nitrogenous bases, such as **backbone format**, then the default format for hydrogen bonds will not work. The bonds will be positioned where the basepair atoms would be, but appear to float in the air because these basepair atoms are not displayed.

To correct this problem, you will need to apply the set hbond backbone command to pin the hydrogen bonds to the phosphorus backbone. To reverse this and return the hydrogen bonds to their original location, you can apply the set hbond sidechain command.

To display hydrogen bonds on the backbone of nucleics:

```
set hbonds backbone
```



Display Colors for Nucleics

Many of the common **display colors** that apply to protein structures also work on nucleic acids, including color cpk, color chain, and color temperature. You can also apply specific custom colors using the Jmol approved color names such as color red or you can apply colors using [R, G, B] values such as color [215, 105, 123].

Coloring by Base Type

One useful color format that we find often works well is to use four different colors to represent the four types of bases (**adenosone**, **thymine**, **guanine** and **cytosine**). For this color format, we keep the nucleic backbone white so that the overall shape of the structure can still be seen.

To color the nucleotides by base type:

```
select a
color red
select t
color yellow
select g
color green
select c
color blue
select backbone and nucleic
color white
```

