Using Voroprot to detect void spaces in molecules

Valeri Alexiev

1. General Overview of Voroprot and How to Install It:

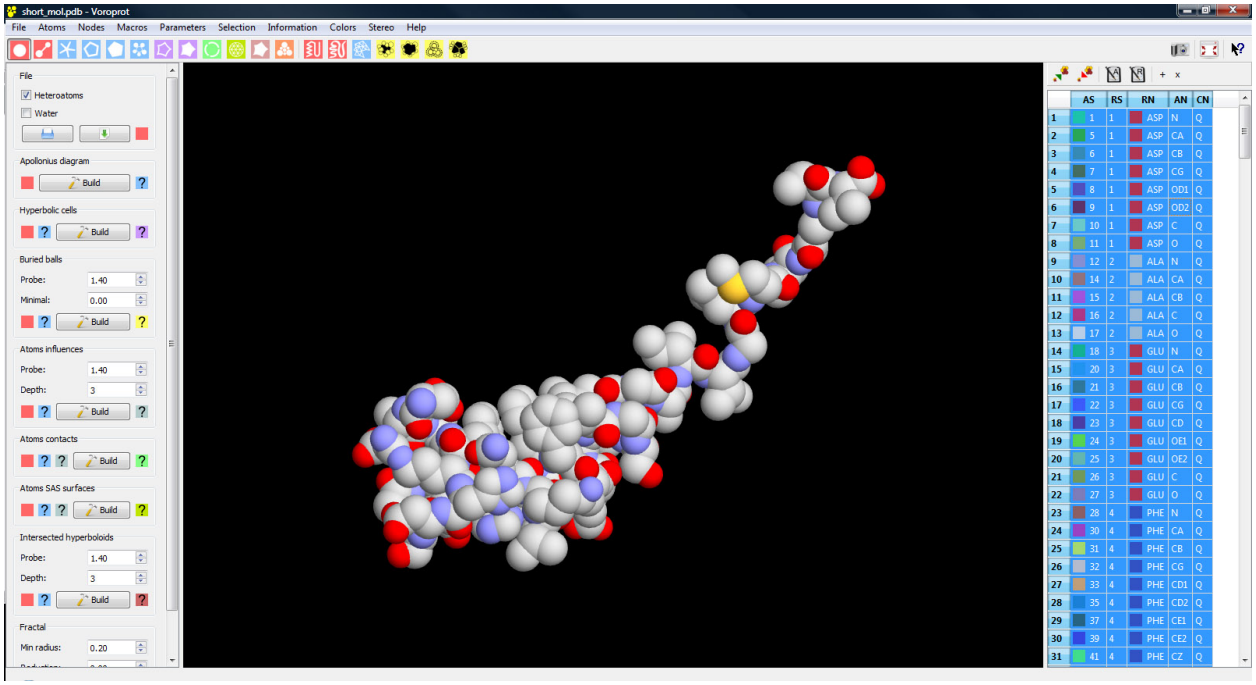
Voroprot is a software tool for constructing and visualizing Apollonius diagrams (additively weighted Voronoi diagrams), that includes a number of useful tools such as atom-by-atom visualization and functionality for finding cavities and pockets in the structures of protein. Voroprot is written with C++ and uses OpenGL for 3D rendering. Possible drawbacks of Voroprot is the fact that it skips hydrogen atoms for optimization purposes and, even then, is too resource intensive. Therefore, when working on general purpose computers, it is recommended that smaller molecule systems are used (for example, including only protein atoms).

On the Voroprot downloads page [2], one can download instructions for installing Voroprot on Windows, Mac and Linux. The source code of Voroprot is also available, if one wants to compile it. Unfortunately, the code lacks any documentation or comments, which makes it very hard to extend or modify.

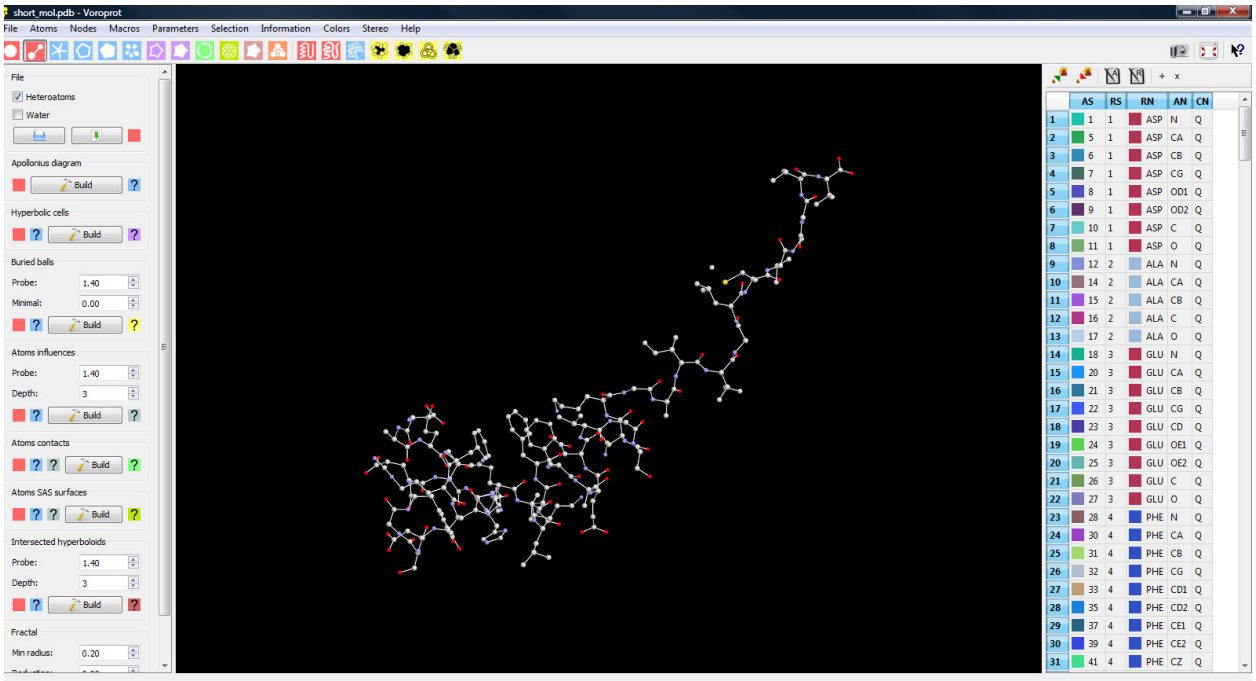
2. Basic Functionality:

On the Voroprot website, there is a detailed tutorial [3] that includes function-by-function descriptions and illustrations and videos that show the program in action. Nevertheless, in this section, I'll cover some useful functions.

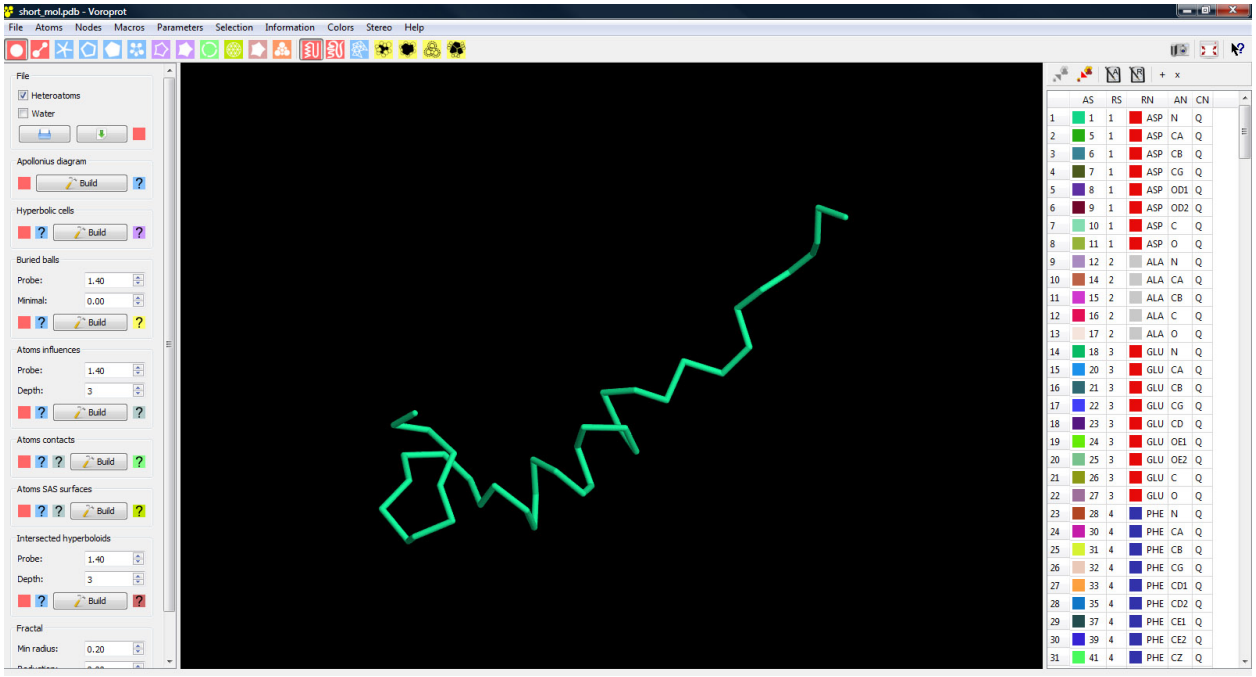
PDB files are loaded in Voroprot using File->Open and selecting the appropriate file. Once the file is loaded, a list of atoms will be displayed on the right of the screen as shown in the picture below. There are several ways to visualize atoms in Voroprot. One of them is the "Draw atoms" function, that has been selected in this image. It visualizes selected atoms as colored spheres, with each color corresponding to an atom type. Atoms can be selected individually from the list, several at a time using <Ctrl + mouse click> (in Windows) or all at a time using <Ctrl + A>.



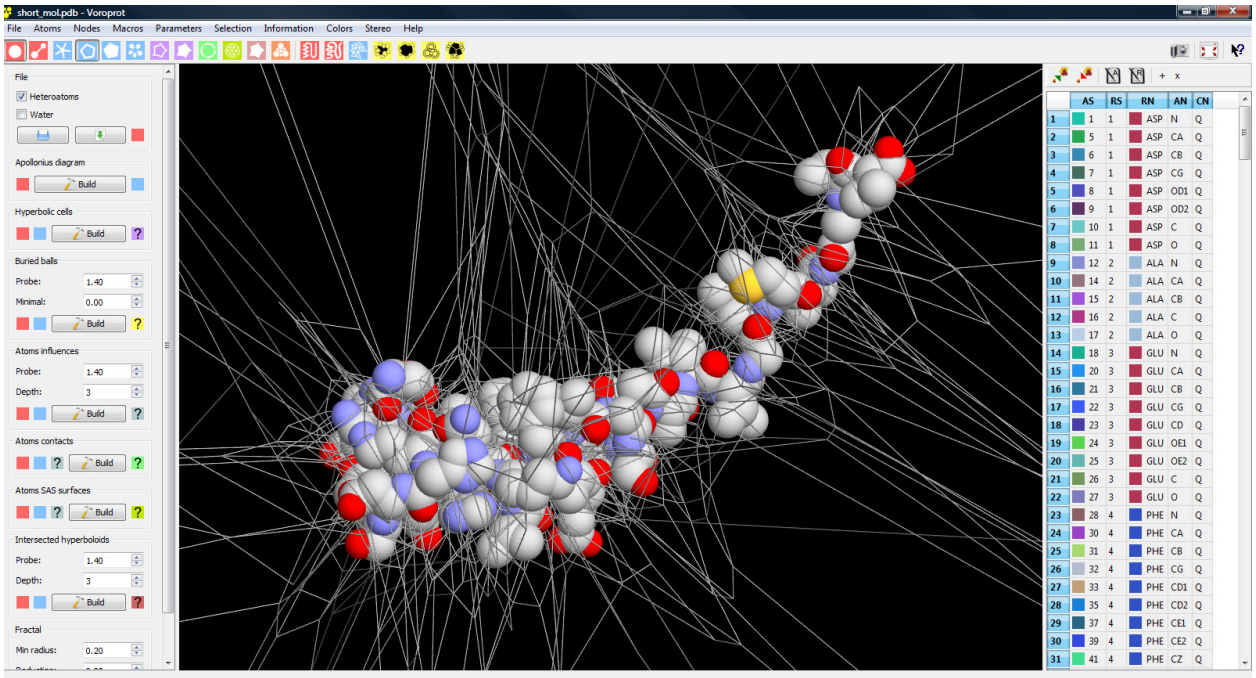
The next image shows another visualization option, "Draw sticks representation". It is used to show more clearly the bonds between atoms.



Proteins can also be visualized using the "Draw C-alpha trace" option. The C-alpha trace visualization is achieved by only representing the alpha carbons (the central carbons) of the amino acids, connected by straight lines. It is used to show the general shape of a protein.

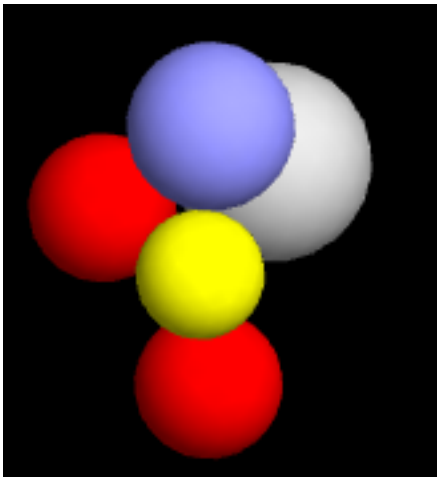


Another important functionality for this project is the ability to construct a Voronoi tessellation. Voroprot uses an additively weighted Voronoi diagram (also called Apollonius diagram). By clicking the "Build" button in the Apollonius diagram section on the left of the screen, the tessellation is constructed. There are several ways to visualize the diagram, which can be selected by clicking one of the blue buttons, underneath the main menu. Below, we can see the Voronoi Cell wireframe. Notice that the cells on the outside of the protein extend to infinity, as this system only includes the protein atoms.

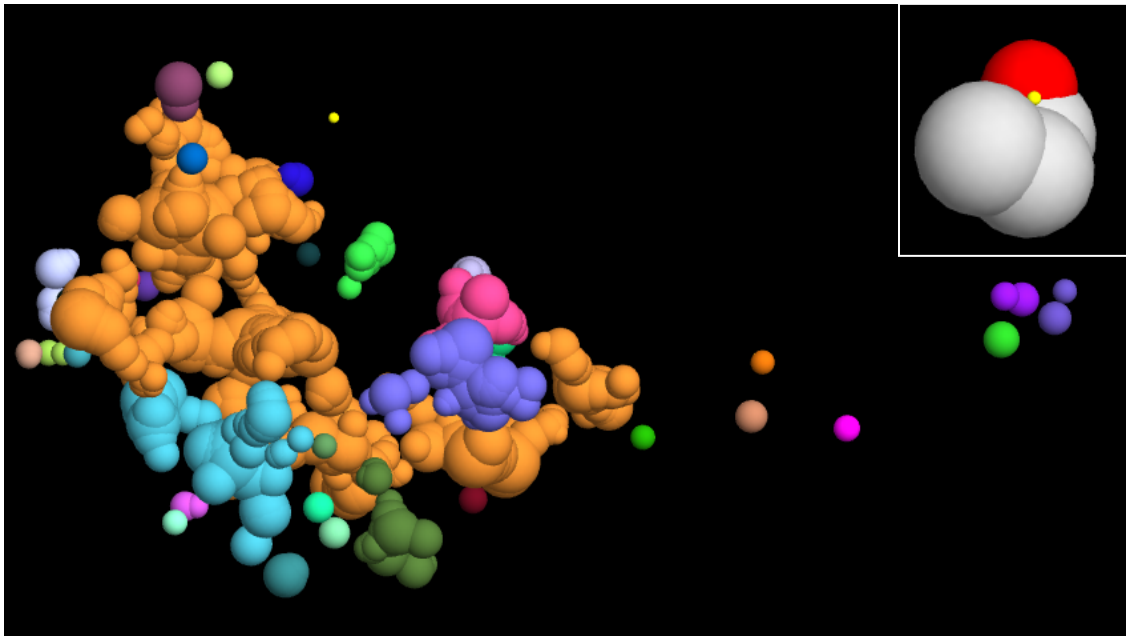


3. Finding Buried Balls and their Connection to Void Spaces:

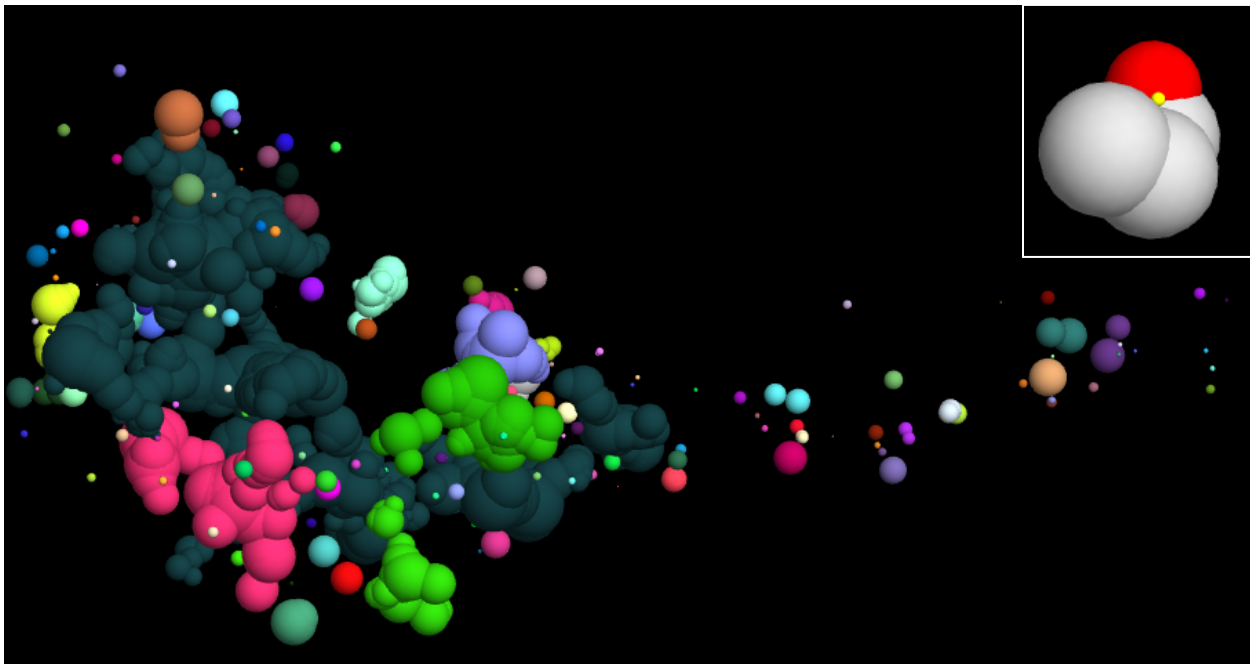
In order to construct the buried balls of a protein, click the "Build" button in the Burried Balls section on the left of the screen. There are two parameters for the Buried Balls function: the minimal buried ball radius and the rolling probe radius, both of which will be discussed later. Voroprot finds cavities by subdividing the protein in quadruples of atoms. A quadruple is a part of a cavity only if its tangent sphere cannot be accessed from the outside by a probe of the size of the solvent molecule. The following image shows a quadruple of atoms with its tangent sphere (in yellow).



The minimal buried ball radius is one of the parameters in the buried ball search. It defines the minimal radius of a tangent sphere needed for a sphere to be included in the buried ball search. Below, we see the result of running a buried ball search with parameters: Minimal Buried Ball Radius: 0.4 Å, Rolling Probe Radius: 1.4 Å. The yellow sphere is the tangent sphere of the quadruple shown on the right and is not included in the buried balls, as its radius is less than the minimum.

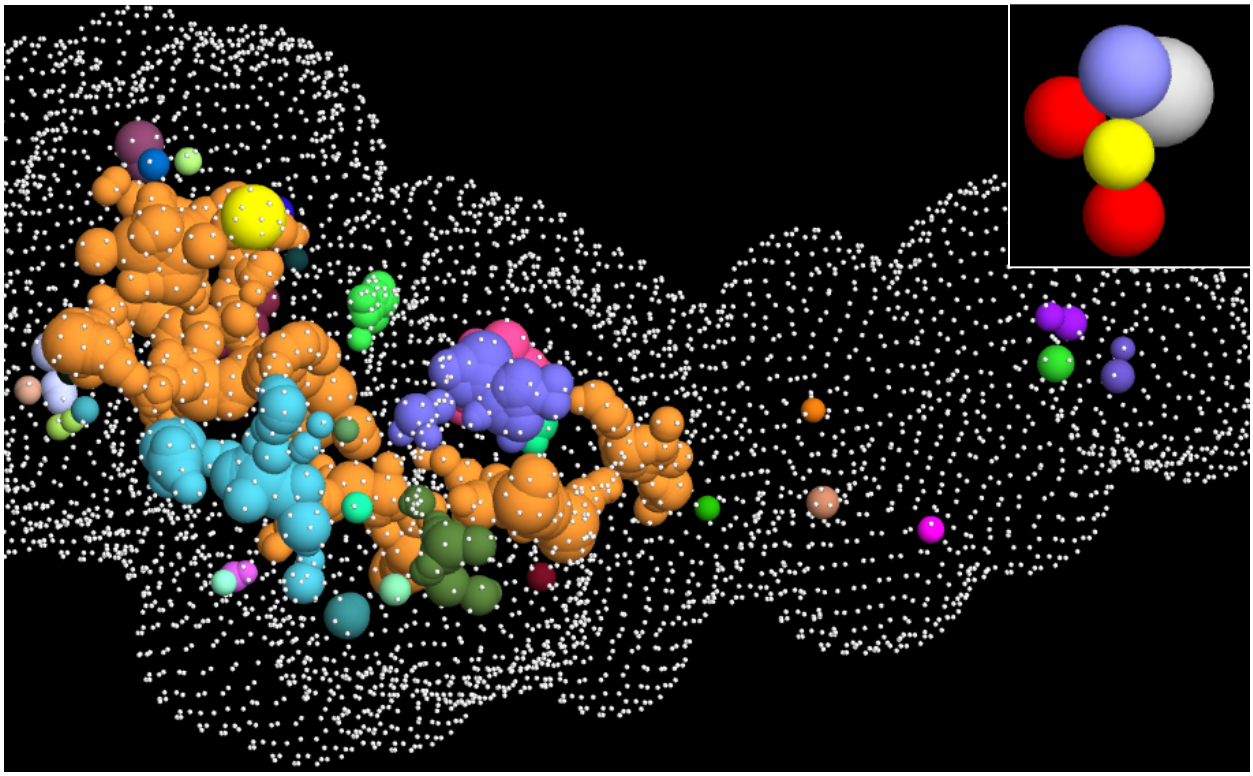


Next are the results of a search with parameters: Minimal Buried Ball Radius: 0.0 Å, Rolling Probe Radius: 1.4 Å. As can be seen, a lot more of the tangent spheres are now considered part of cavities, including the tangent sphere of the quadruple on the right.

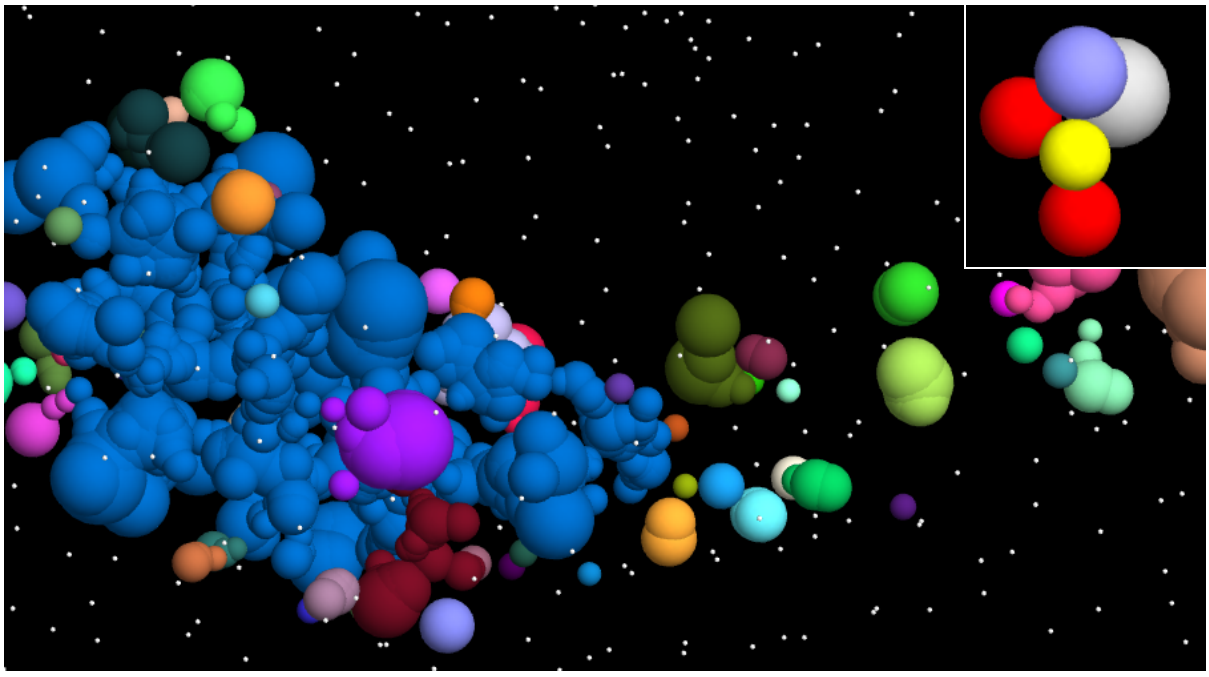


The second parameter of the buried balls function is the size of the rolling probe radius. The rolling probe is what Voroprot uses to determine whether a tangent sphere is part of a cavity or not. Thus, its size is crucial to the buried balls algorithm and has a huge influence on the result. For example, larger probe sizes can be even used to find pockets on the surface of

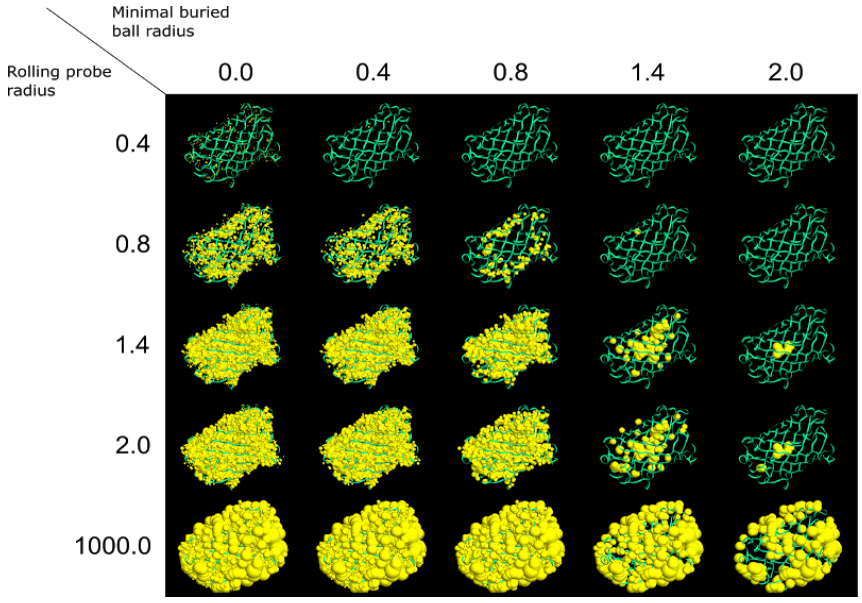
the protein, rather than cavities inside it. Below we see the result of running a search with parameters: Minimal Buried Ball Radius: 0.0 Å, Rolling Probe Radius: 1.4 Å. The white points in the image are the probe trace points as it explores the protein cavities. They were shown using the "Draw probe trace points" button below the menu. As before, the yellow sphere indicates the tangent sphere of the quadruple on the right.



The next image shows the buried balls with parameters: Minimal Buried Ball Radius: 0.0 Å, Rolling Probe Radius: 10 Å. We can see that the probe trace points are much farther away from the protein. Many more tangent spheres are included in this image, including the tangent sphere of the example quadruple.



Unfortunately, the cavities found using the Buried Balls method are not equivalent to the void spaces this projects is trying to examine. This is because, as can be seen in the examples before and the table below, the buried balls that are found vary quite drastically with the parameters. However, it should be possible to find the right parameters that will discover void spaces. For example, the minimal buried ball radius should be probably set to 0 Å as we want to detect even the smallest void spaces. In addition, the size of the rolling probe should be limited, in order to exclude pockets on the surface of the molecule from the search for void spaces. However, finding the exact parameters for the Buried Balls function would most likely require biochemical knowledge of the expected behavior of void spaces and their properties.



References and Useful Links:

1. http://www.ibt.lt/bioinformatics/voroprot/

2. http://code.google.com/p/baltymus/wiki/Installation?tm=2

3. http://code.google.com/p/baltymus/wiki/Tutorial?tm=6

4. https://wiki.cmbi.ru.nl/index.php/C-alpha_trace

5. https://wiki.cmbi.ru.nl/index.php/C-alpha

6. http://www.cgal.org/Manual/latest/doc_html/cgal_manual/Apollonius_graph_2/

   Chapter_main.html