



Visualizzazione di Chemical Structures

Alessandro Grottesi, Ph.D.



Outlook

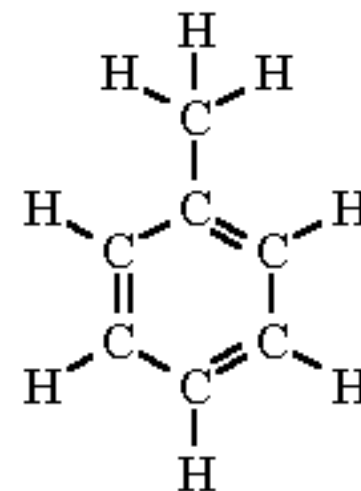
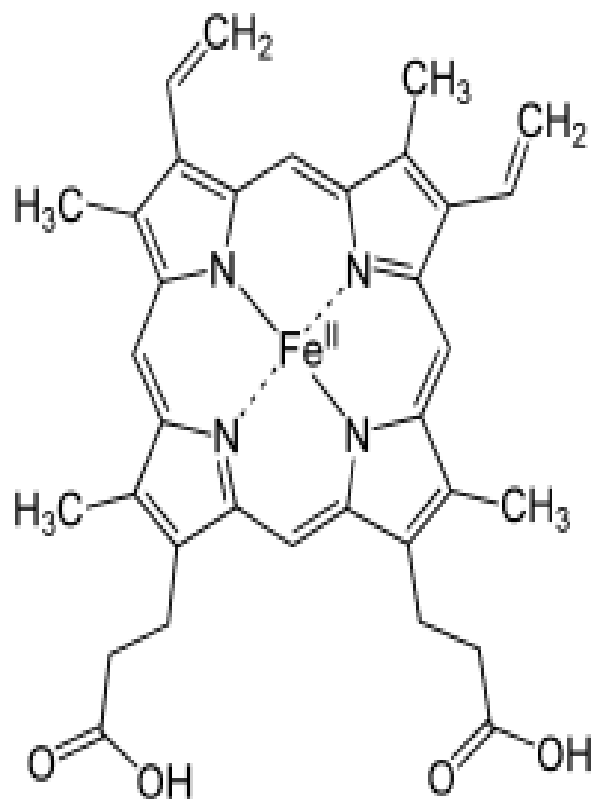
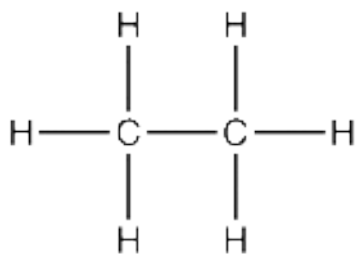
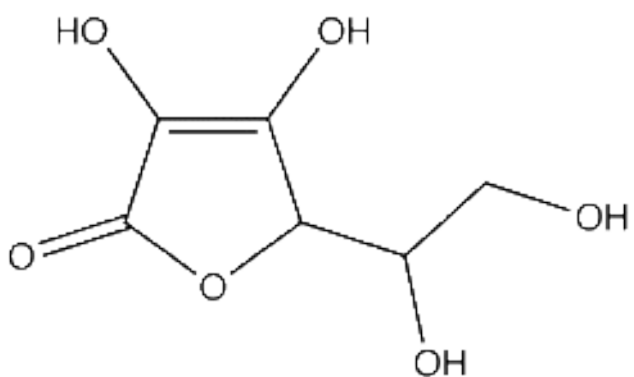
(of concepts and hands-on)



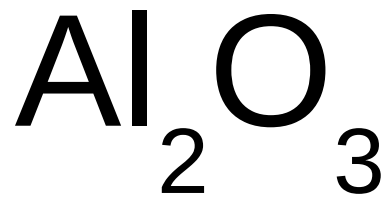
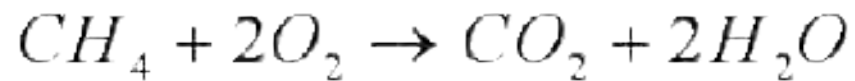
- How to visualize a molecular structure
- What type of info I can get
- Data insights
- Various type of visualizzazione
- Molecular structures in a simulation (trajectories)
- Rendering of chemical structures



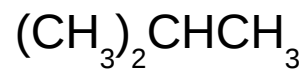
Structural Formula



Chemical Formula



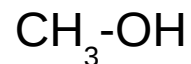
IUPAC Nomenclature



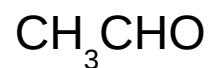
Iso-butane



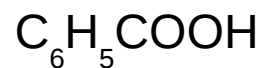
Methane



Methanol



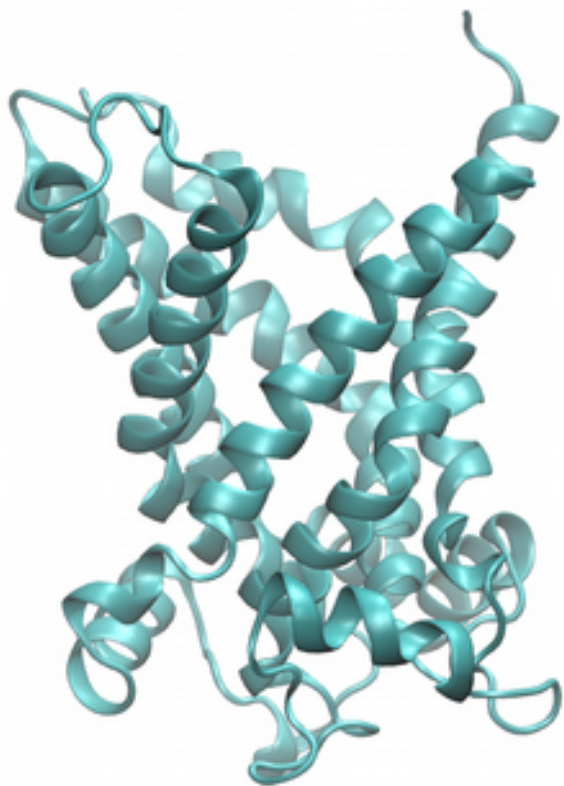
Acetaldehyde



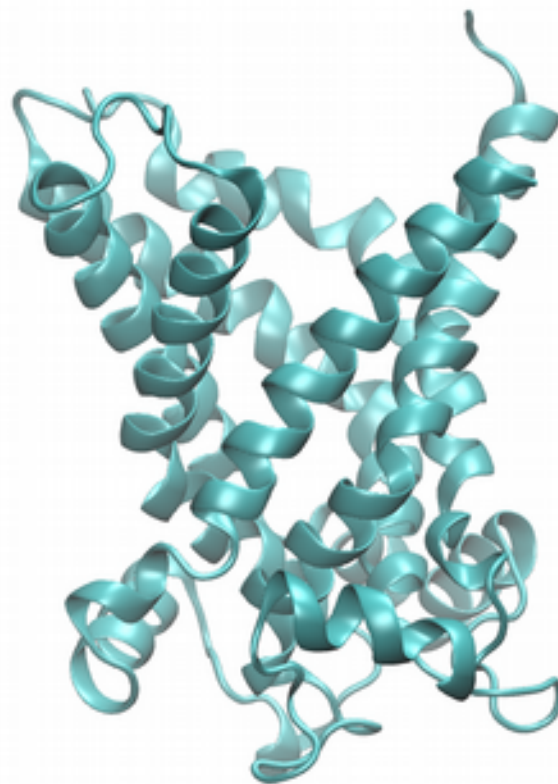
Benzoic acid



Visualizzazione di biomolecole



Low res.



High res.



Visualization of chemical structures



Historical background

- 1960's - 70's: Earliest Computer Representations
 - 1) ORTEP software
- 1980: Teaching Aids for Macromolecular Structure
 - stereoview glasses
- 1980-1990: Evans & Sutherland Computers
 - FRODO: software for visualizing X-ray structures
- 1992: David & Jane Richardson's Kinemage
 - MAGE and Prekin software for biomolecules
- 1993: Roger Sayle's RasMol
 - Rasmol released: high quality images and rendering
- 1996: MDL's Chime
 - First plugin for browser (Netscape): Chemical MIME



Drawing format for chemical structures



GIF supports up to 8-bit colour (256 possible colours) and stores custom palette with the image. GIF offers several options, including transparent backgrounds and lossless compression, which means that the restored image looks exactly like the original, therefore it is excellent for chemical sketches or cartoon-like images.

JPEG (JPG) supports up to 32-bit true colour (4.2 billion colours) and is an excellent option for photographs and scanned or rendered images – for images with intricate details and shades. JPEG algorithm uses lossy compression, providing high quality images with a high level of compression. On saving in this format, you can choose the display quality, from high quality to very low quality reproductions. The higher the image quality, the larger the file size.

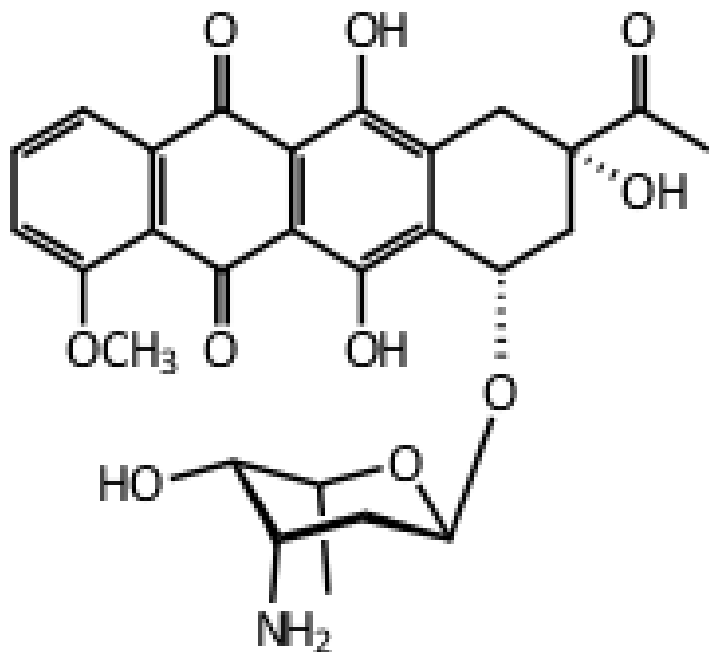
PNG is a relatively new format that was developed as an alternative to GIF and supports true colour graphics and lossless compression. The quality is superb, the size of the file is usually larger than that of it in JPG format.

TIFF and **TGA** are similar formats with lossless compression, sometimes useful when exporting into graphical applications.

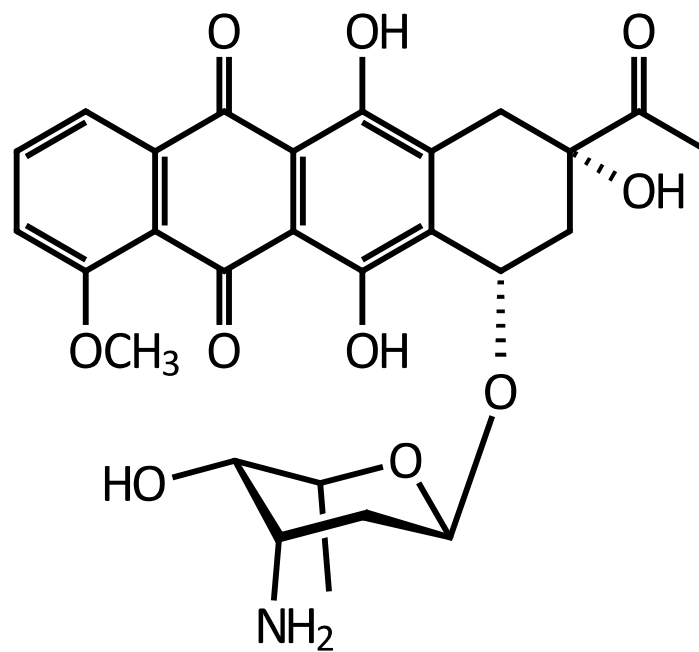
Vector graphics for the web: SVG



SVG (Scalable Vector Graphics) is an XML-based vector image format for 2D graphics with support for interactivity and animation. SVG was developed by the World Wide Web Consortium (W3C) and all modern browsers support it. When creating SVG files, text should be rendered as shapes (option for example in ChemDoodle). Examples of available software: [Inkscape](#) (open source)



PNG format



SVG format



Drawing of chemical (and biochemical) structures



Ball-and-sticks

CPK

Surface (VdW, solvent, isosurface)

Diagram

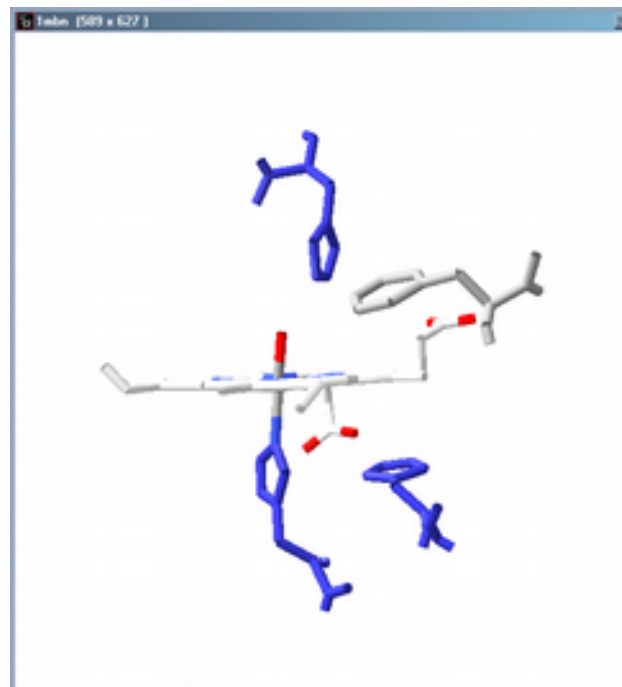
Ribbon representation

Cartoon



Structure data – informative views

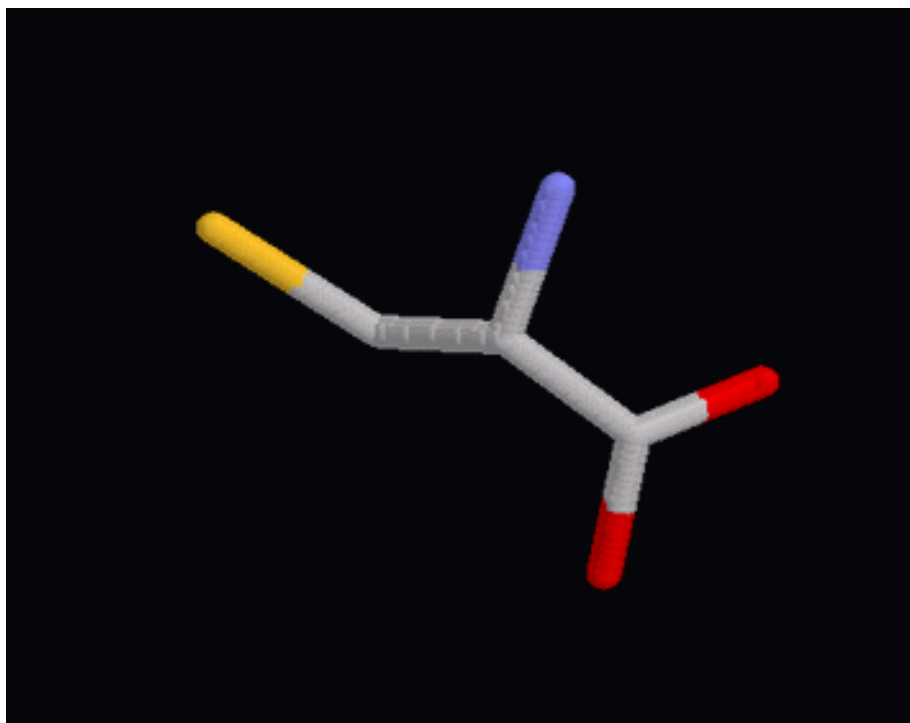
- Informative views highlighting important features
- First image:
 - Fold of the myoglobin molecule and its prosthetic group
- Second image:
 - residues within 3 Å of the heme group



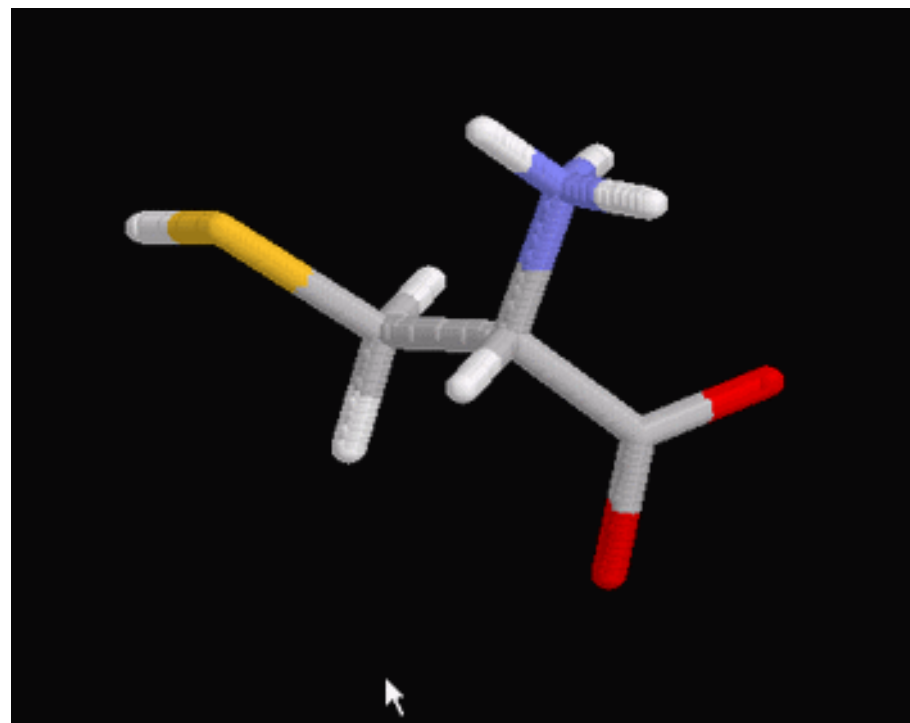
Visual styles - Wireframe



- Example: Cysteine
- Stick/Wireframe: shows individual bonds and corresponding angles



Without hydrogen atoms



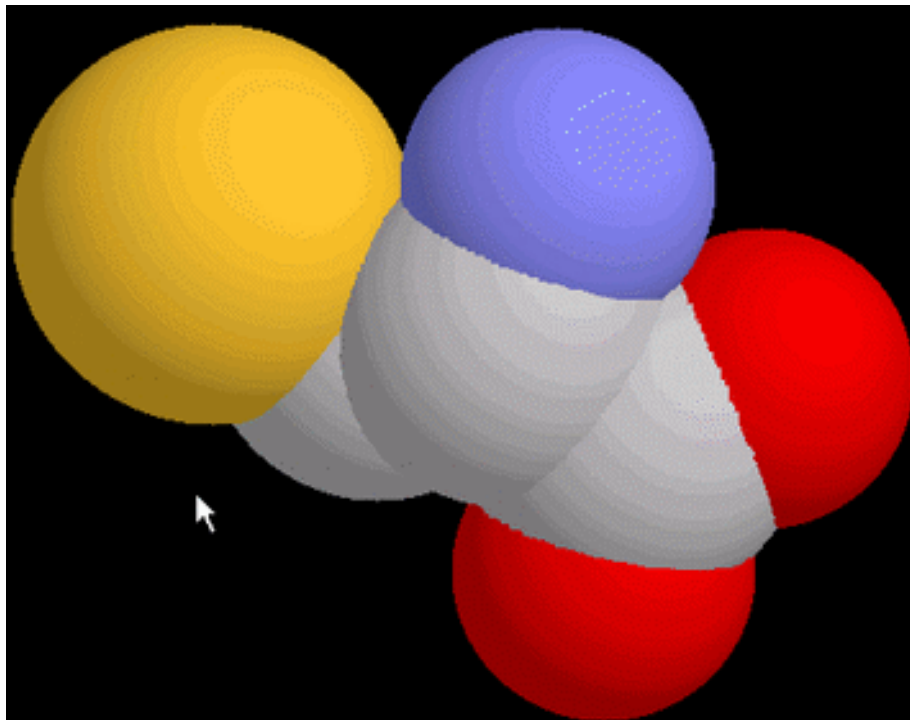
With hydrogen atoms



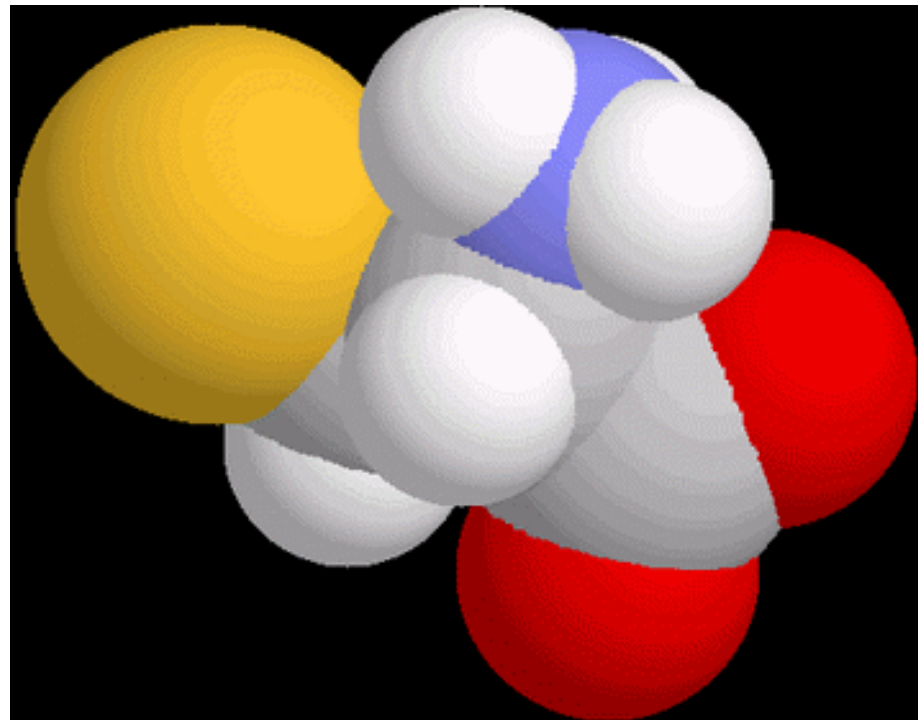
Visual styles - Spacefill



- Spacefilling:
 - Spheres represent the electron cloud around nuclei of atoms
 - Relative sizes of atoms and groups show up clearly
 - Merged spheres represent share electrons (covalent bond)



Without hydrogen atoms

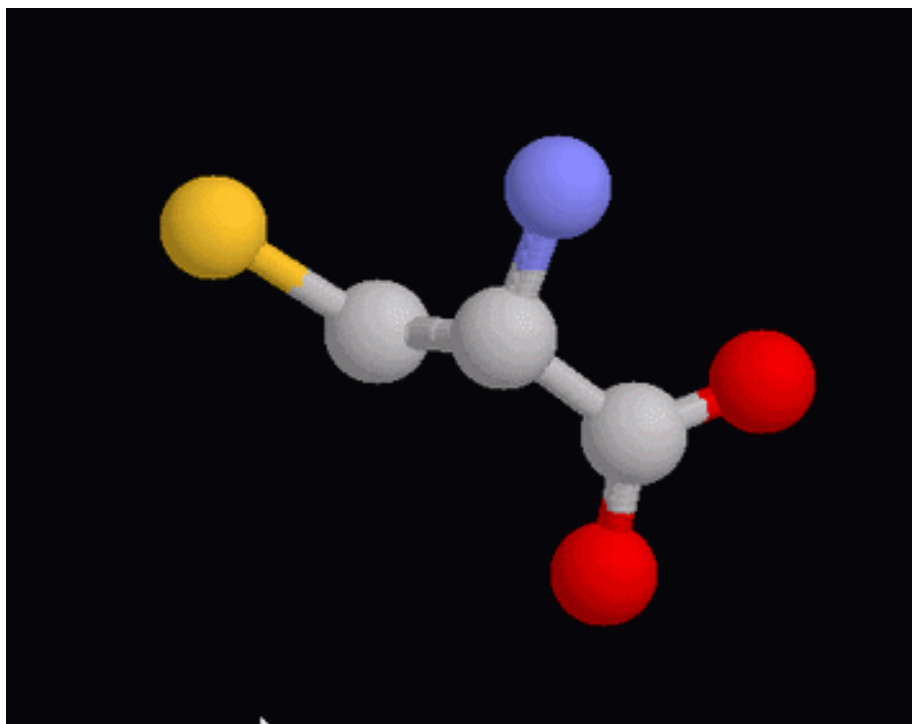


With hydrogen atoms

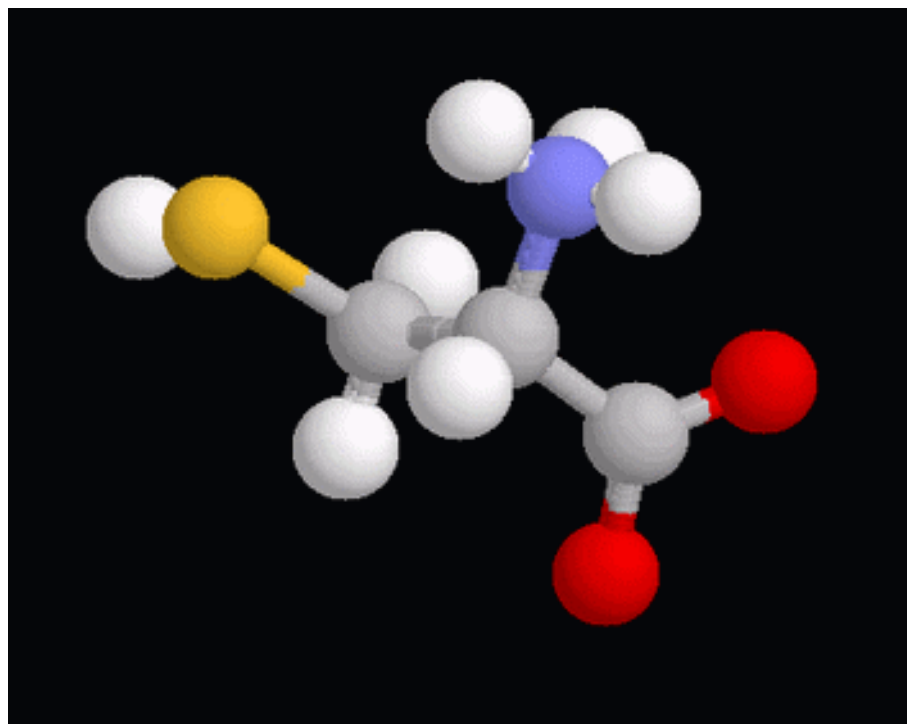
Visual styles – Ball and stick



- Ball and stick:
 - Allows structure to be seen
 - shows atomic position and bonds, and a portion of volume – small spheres are centred at the nuclei of atoms



Without hydrogen atoms

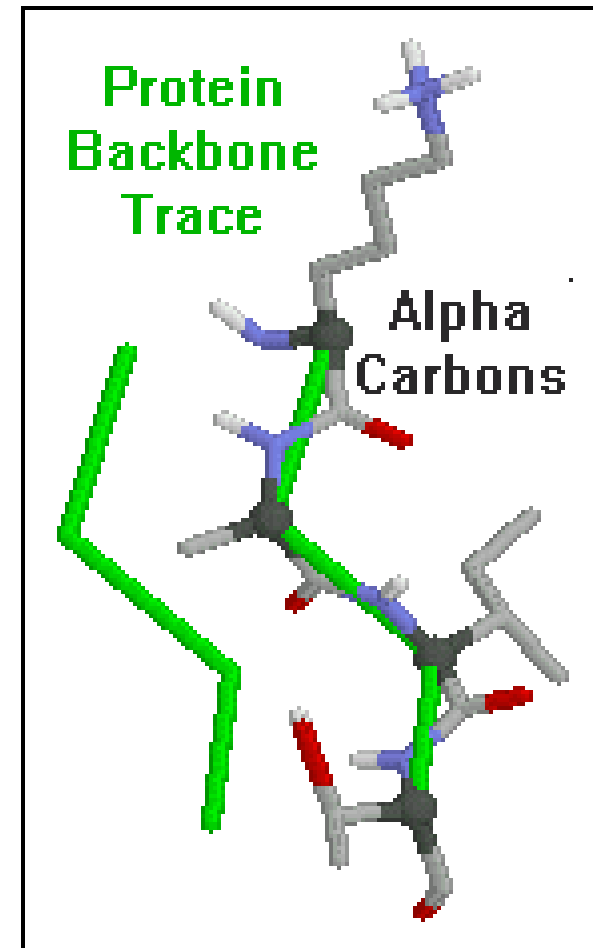


With hydrogen atoms

Visual styles – Backbone



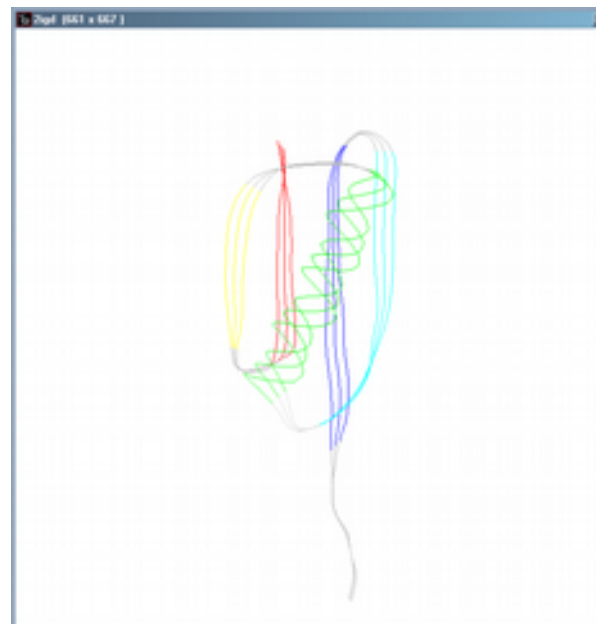
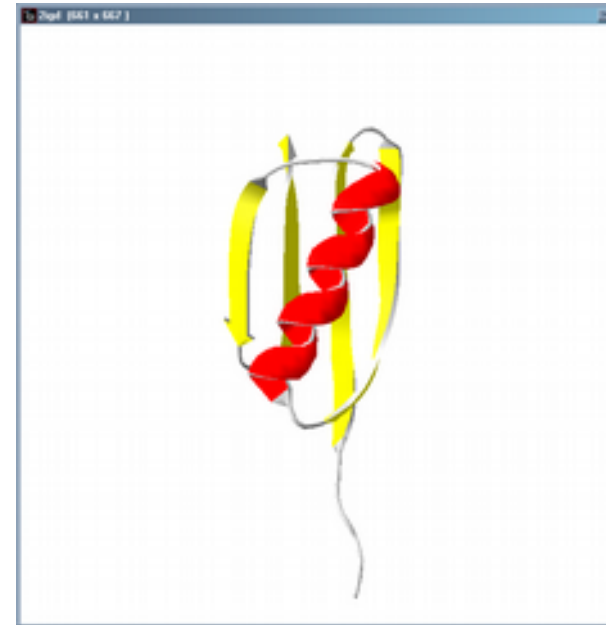
- **Backbone:**
 - Four amino acids shown (Lys, Ala, Ile, Thr from top to bottom)
 - Green line is a backbone trace: all side chains removed, rods connect alpha carbon positions (or phosphate atoms)
 - Backbone does not follow any actual bonds
 - Right image shows backbone superimposed on tetrapeptide
- Used e.g. for structural superimpositions



Visual styles - Ribbon



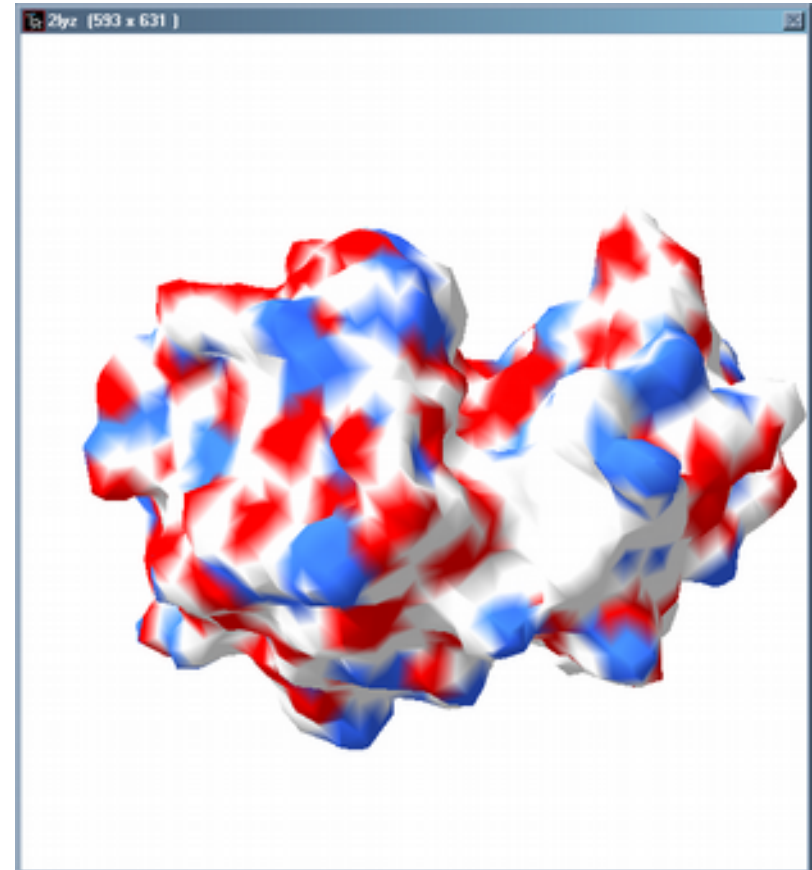
- Ribbon
 - Smooth ribbon drawn through peptide units – may either be solid ('ribbons') or parallel wires ('strands')
 - Display the fold of a protein
- Colour schemes shown
 - by secondary structure
 - by secondary structure succession
 - N-terminus blue
 - C-terminus red



Visual styles - Surface



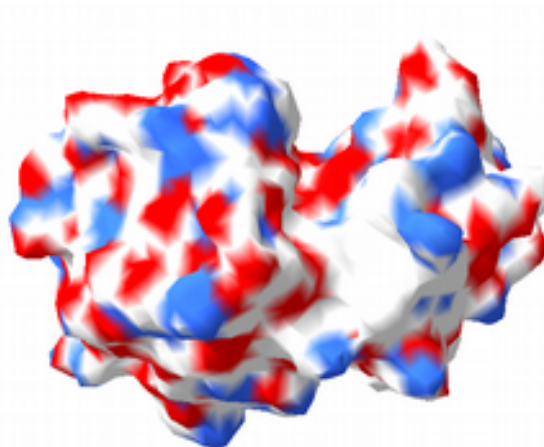
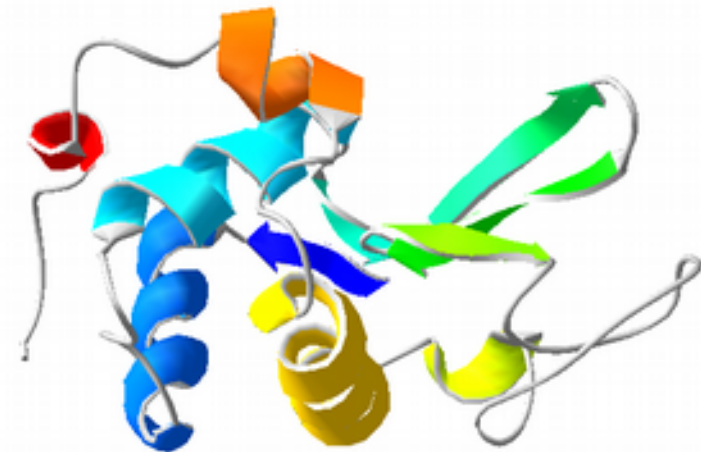
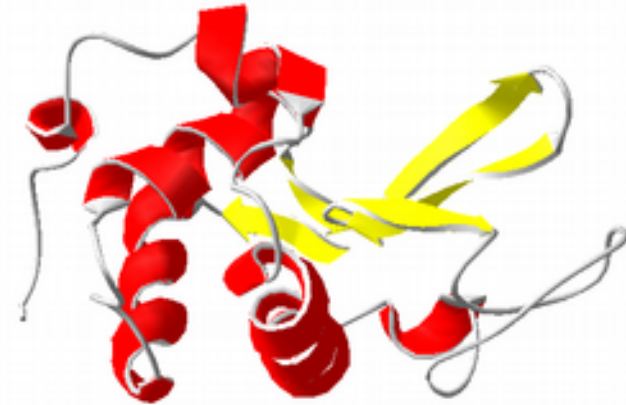
- Shows the surface accessible to water molecules
- Most useful for studying the properties of the active sites
- Molecular surface of lysozyme
 - Active site in a groove
 - Charged, hydrophilic surface



Colour schemes



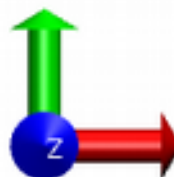
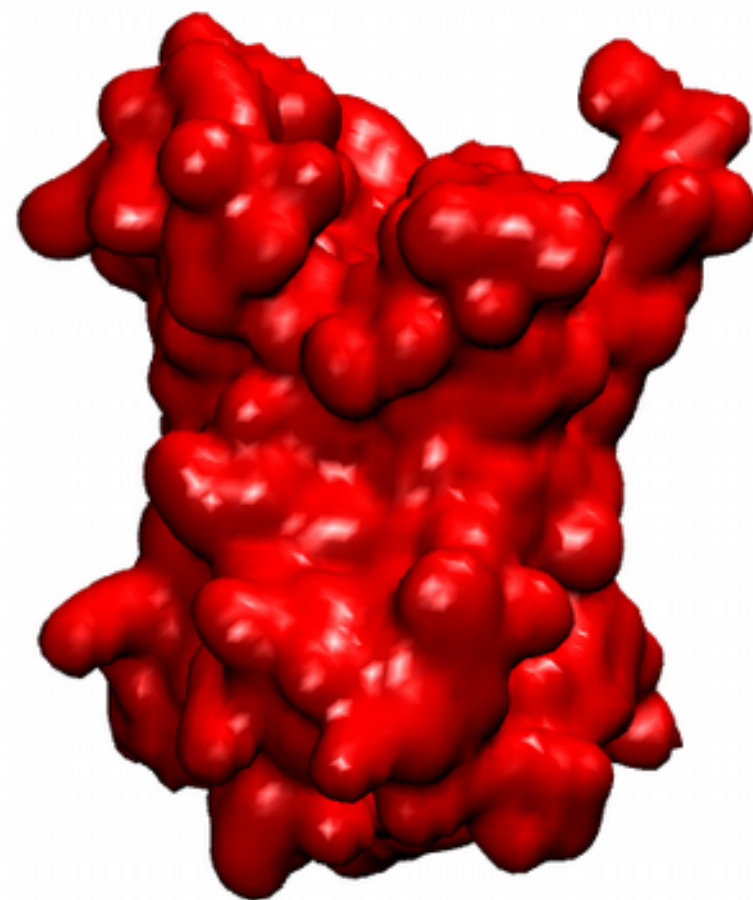
- Colour by:
 - secondary structure
 - secondary structure succession
 - B-factor (tells the quality of the model): red \square least certain, blue
 - Accessibility to solvent
 - red: most accessible, violet: least accessible
- Calculate molecular surface
 - color by: atom type (CPK), cavity or electrostatic potential



Visualizing molecular surface



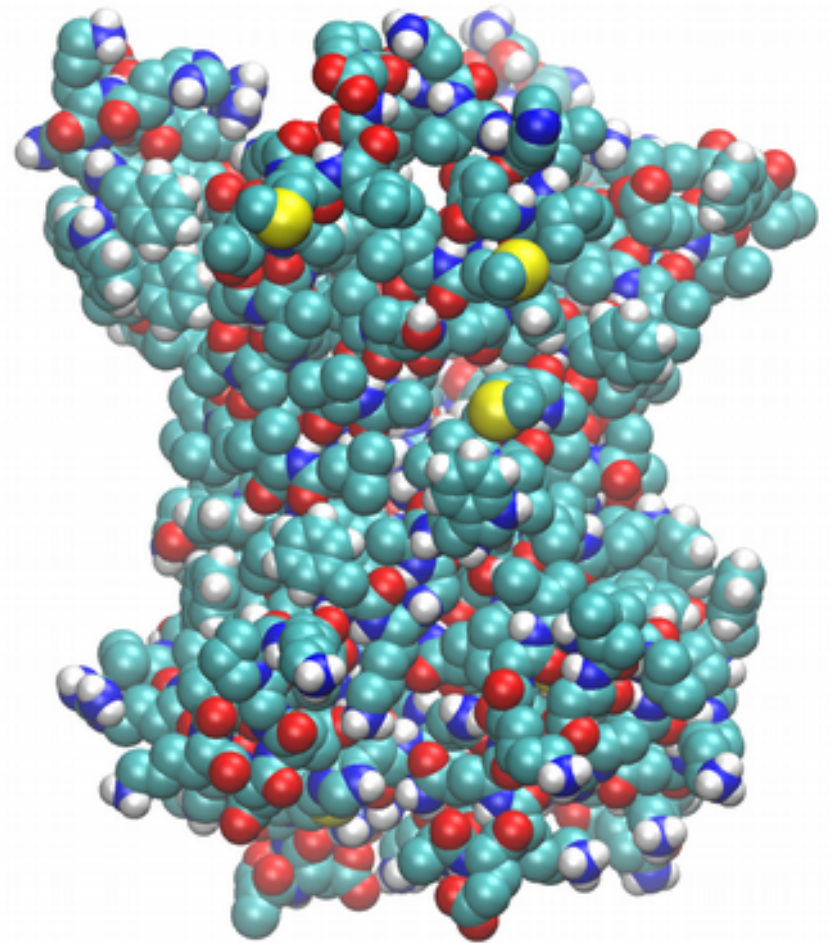
- Van der Waal surface
- Accessible surface area
- isosurface



Van der Waals surface



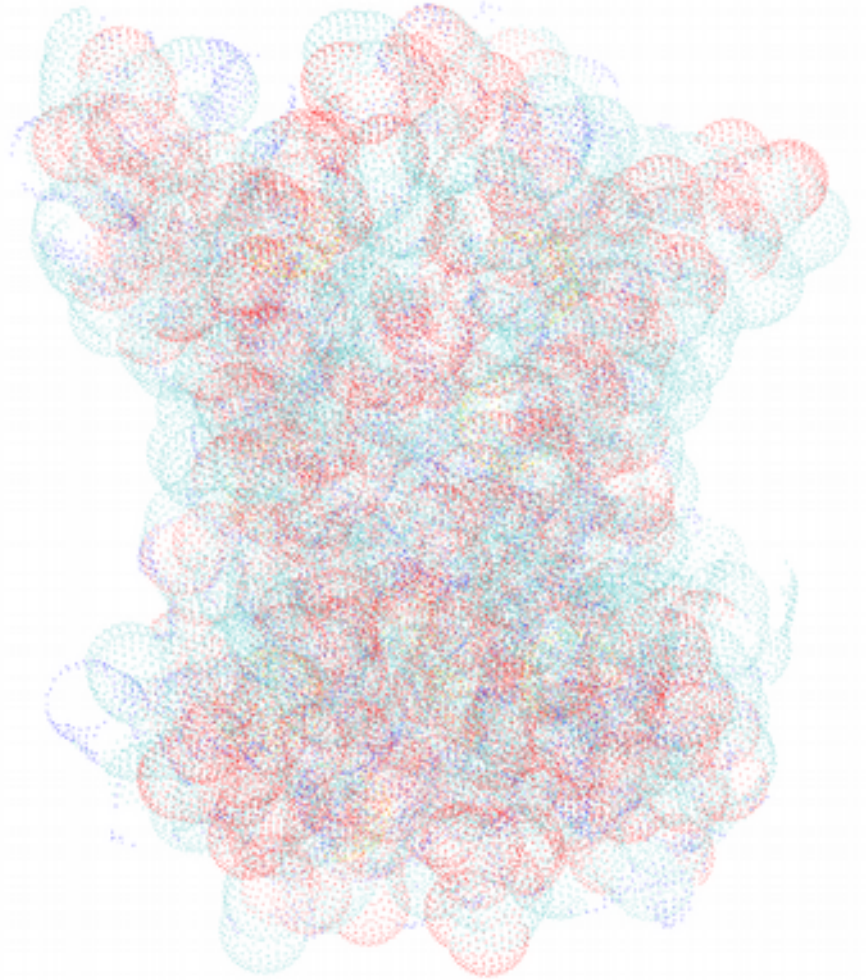
The van der Waals surface of a molecule is an abstract representation or model of that molecule, illustrating where, in very rough terms, a surface might reside for the molecule based on the hard cutoffs of van der Waals radii for individual atoms, and it represents a surface through which the molecule might be conceived as interacting with other molecules.



Solvent Accessible Surface Area



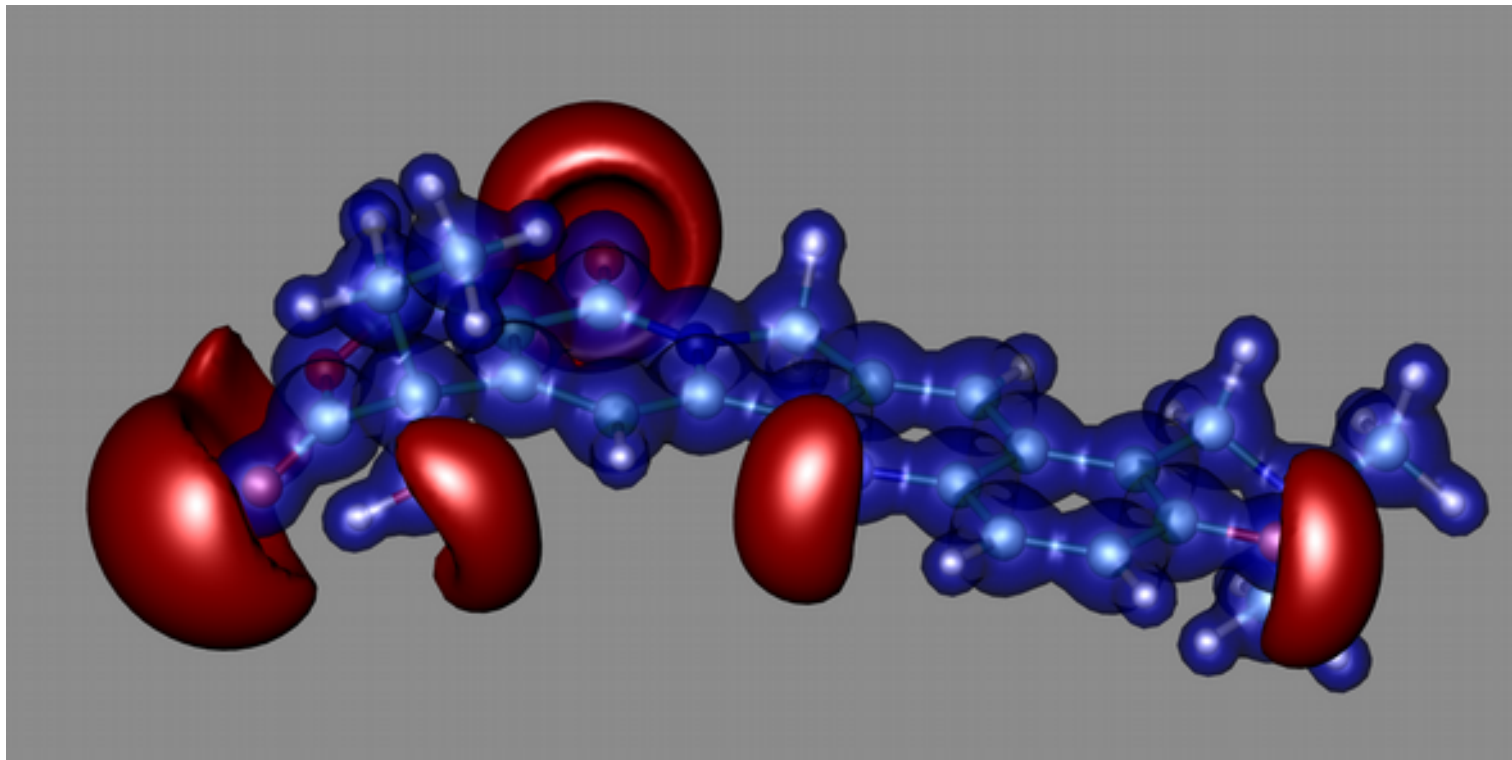
The accessible surface area (ASA) or solvent-accessible surface area (SASA) is the surface area of a biomolecule that is accessible to a solvent. Measurement of ASA is usually described in units of square ångstroms (a standard unit of measurement in molecular biology).



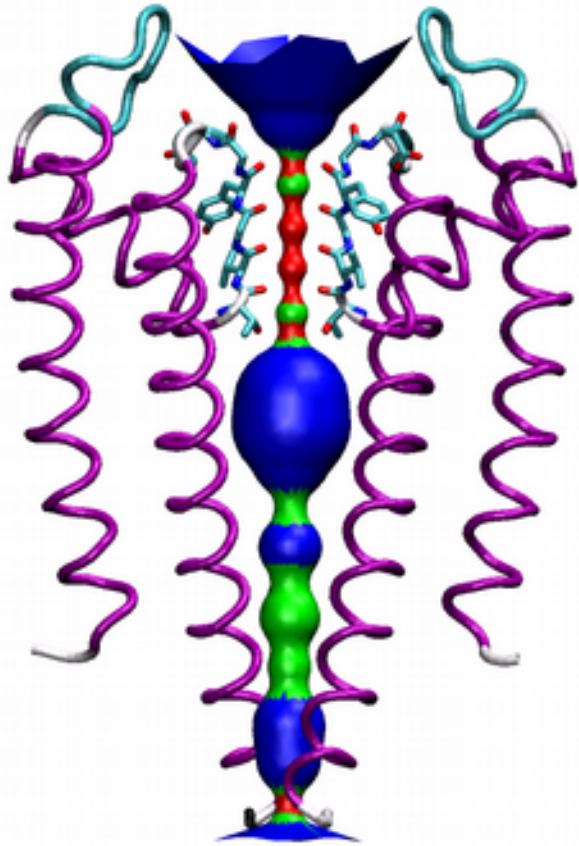
Isosurface representation



An isosurface is a three-dimensional analog of an isoline. It is a surface that represents points of a constant value (e.g. pressure, temperature, velocity, density) within a volume of space; in other words, it is a level set of a continuous function whose domain is 3D-space.

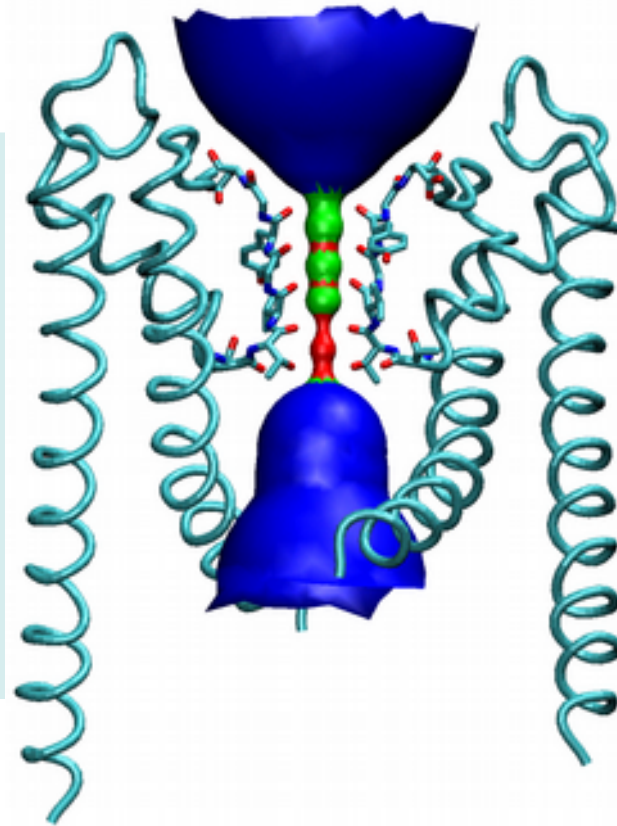


Closed (KcsA)



Ramachandran plot: 93.1% core
6.9% allow 0.0% gener 0.0% disall

Open (KvAP)



Ramachandran plot: 92.2% core
6.7% allow 1.1% gener 0.0% disall



Software for chemical graphics



Mostly used software (free license) for molecular graphics include:

- VMD
- Rasmol
- Pymol
- Chimera
- Molden
- Jmol
- Molscrip
- Avogadro
- What If
- gOpenMol
- OpenStructure
- Yasara
- Maestro
- Spartan
- Discovery
- GaussianView

Molecules in motion



Static representation of chemical structure provides a limited picture and somewhat reduced level of data insight as they represent static representation of a time-averaged observable or simply do not represents the emerging properties of the molecule.

Molecular simulations represent a valid tool for describing the structural and thermodynamics properties of a molecular system. Trajectories are a collection of chemical structures as a function of time. Typically, a potential energy function and equation of motions generate a time evolution of a chemical structure

Trajectories can be generated using specific software running on HPC clusters and specific graphics tools can be used to manage and analyze trajectory frames



Software for visualizing molecules: VMD

VMD provides user-editable "materials" which can be applied to molecular geometry. These material properties control the details of how VMD shades the molecular geometry, and how transparent or opaque the displayed molecular geometry is. With this feature, one can easily create nice looking transparent surfaces which allow inner structural details to be seen within a large molecular structure. The material controls can be particularly helpful when rendering molecular scenes using external ray tracers, each of which typically differ slightly.

Software for visualizing molecules: VMD



VMD Main

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
0	T	A	D	F	steep.gro	66337	1	0

Molecule File Browser

Load files for: 0: steep.gro

Filename: Browse...

Determine file type: Automatically Load

Frames: First: 0 Last: -1 Stride: 1

Load in background Load all at once

Graphical Representations

Selected Molecule:

Delete Rep

Color Selection

Structure all

Selected Atoms: all

Draw style: Selections: Trajectory: Periodic:

Material: Opaque

Spline Style: Catmull-Rom

Resolution: 10

Apply Changes Automatically Apply

```
vmd > Info) Using plug
Info) Using plugin gro
Info) Determining bond
Info) Analyzing struct
Info) Atoms: 66337
Info) Bonds: 46516
Info) Applied: 0.00
```

Software for visualizing molecules: Chimera

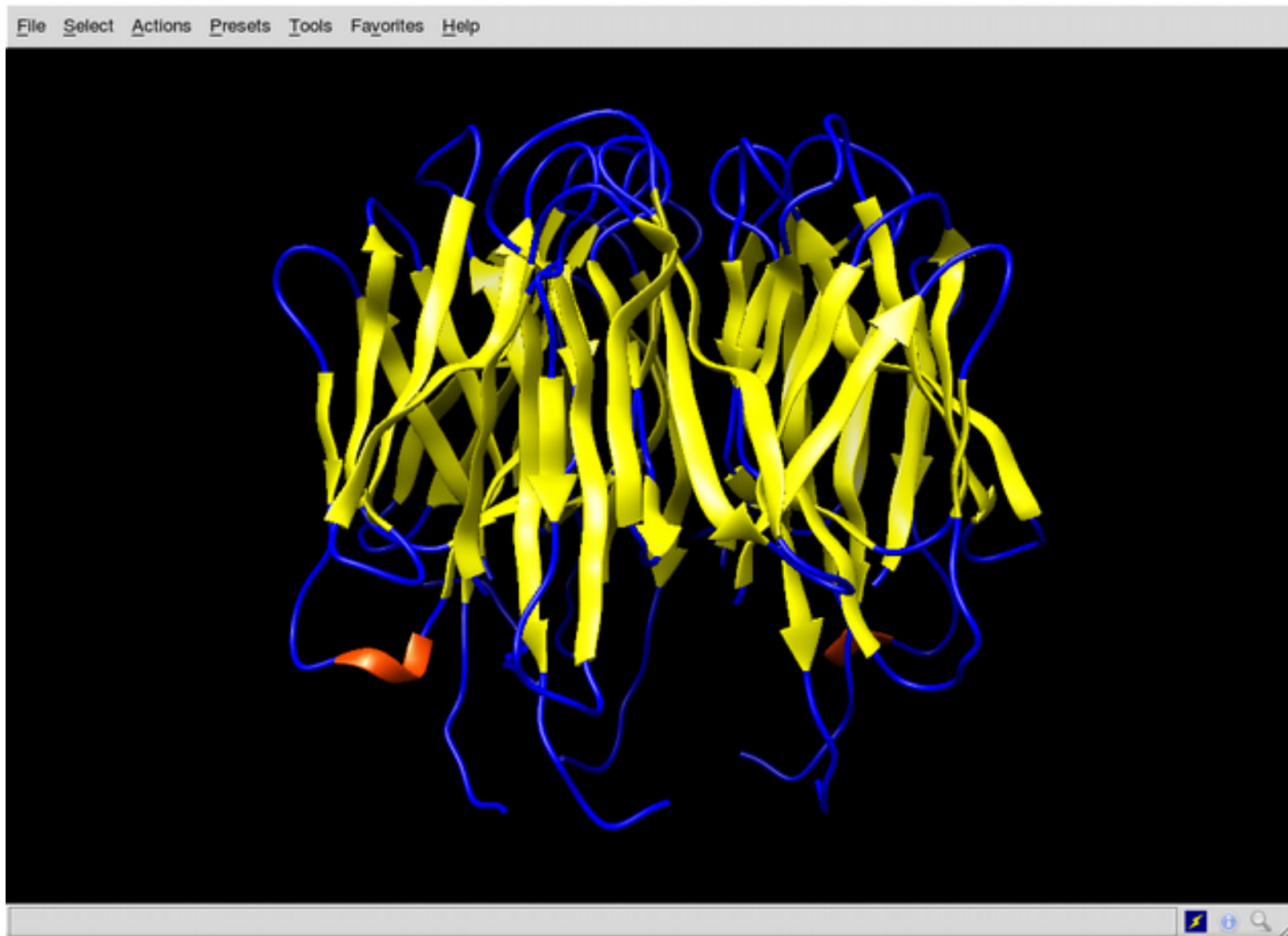


UCSF Chimera is a highly extensible program for interactive visualization and analysis of molecular structures and related data, including density maps, supramolecular assemblies, sequence alignments, docking results, trajectories, and conformational ensembles. High-quality images and animations can be generated.

<https://www.rbvi.ucsf.edu/chimera/>



Software for visualizing molecules: Chimera



Rasmol



RasMol is a molecular graphics program intended for the visualisation of proteins, nucleic acids and small molecules. The program is aimed at display, teaching and generation of publication quality images.

The program reads in molecular coordinate files and interactively displays the molecule on the screen in a variety of representations and colour schemes. Supported input file formats include Protein Data Bank (PDB), Tripos Associates' Alchemy and Sybyl Mol2 formats, Molecular Design Limited's (MDL) Mol file format, Minnesota Supercomputer Center's (MSC) XYZ (XMol) format, CHARMM format, CIF format and mmCIF format files.

The loaded molecule can be shown as wireframe bonds, cylinder 'Dreiding' stick bonds, alpha-carbon trace, space-filling (CPK) spheres, macromolecular ribbons (either smooth shaded solid ribbons or parallel strands), hydrogen bonding and dot surface representations.

<http://www.openrasmol.org/>

<https://www.umass.edu/microbio/rasmol/>



Software for visualizing molecules: Rasmol



```
agrotttes@nagrotttes220458: ~  
File Edit View Search Terminal Help  
agrotttes@nagrotttes220458:~$ rasmol-classic apoa1_prot.pdb  
RasMol Molecular Renderer  
Roger Sayle, August 1995  
Copyright (C) Roger Sayle 1992-1999  
Version 2.7.5.2 May 2011  
Copyright (C) Herbert J. Bernstein 1998-2011  
*** See "help notice" for further notices ***  
[32-bit version]  
  
Secondary Structure ... No Assignment  
Number of Chains ..... 2  
Number of Groups ..... 392  
Number of Atoms ..... 6410  
Number of Bonds ..... 6464  
  
RasMol>  
Number of H-Bonds ..... 293  
Number of Helices ..... 17  
Number of Strands ..... 2  
Number of Turns ..... 24  
RasMol> axes on  
RasMol> 
```

PyMol



The screenshot shows a Linux desktop environment with a terminal window on the left and a PyMol Viewer window on the right. The terminal window displays the output of the PyMol installation script, including component status checks and an installation error message. The PyMol Viewer window shows a 3D molecular model of a protein structure, rendered in green and blue, with a control panel on the right side.

Terminal Window Output:

```

agrottes@nagrottes220458: ~
File Edit View Search Terminal Help
- Ksdssp (installed)
- MODELLER (not installed)
- BLAST databases (installed)

#####
# Components installable with the install_all.py script #
#####
- The following components are missing:
  - MODELLER
- Checking permissions on target install directories.
- All permissions are granted.

#####
# Initializing installation of PyMod components #
#####

#####
# Installation Error #
#####
- Installation failed because of the following error: 'pymol_

#####
# PyMod installer script failed #
#####
- PyMod installer script quitting.

HEADER METAL TRANSPORT 18-JUL-11
TITLE CRYSTAL STRUCTURE OF THE G PROTEIN-GATED INWARD RECTI
TITLE 2 GIRK2 (KIR3.2) R201A MUTANT
COMPND G protein-activated inward rectifier potassium channe
ObjectMolecule: Read secondary structure assignments.
ObjectMolecule: Read crystal symmetry information.
Symmetry: Found 8 symmetry operators.
CndLoad: "/data/alegrot/GIRK_Cycle/3SYP/3SYP.pdb1" loaded as
  
```

PyMol Viewer Window:

The PyMol Viewer window displays a 3D molecular model of a protein structure, rendered in green and blue. The control panel on the right side includes a list of objects (All, 3SYP) and a Mouse Mode 3-Button Viewing section with various keyboard shortcuts for navigation and manipulation.

MOLDEN



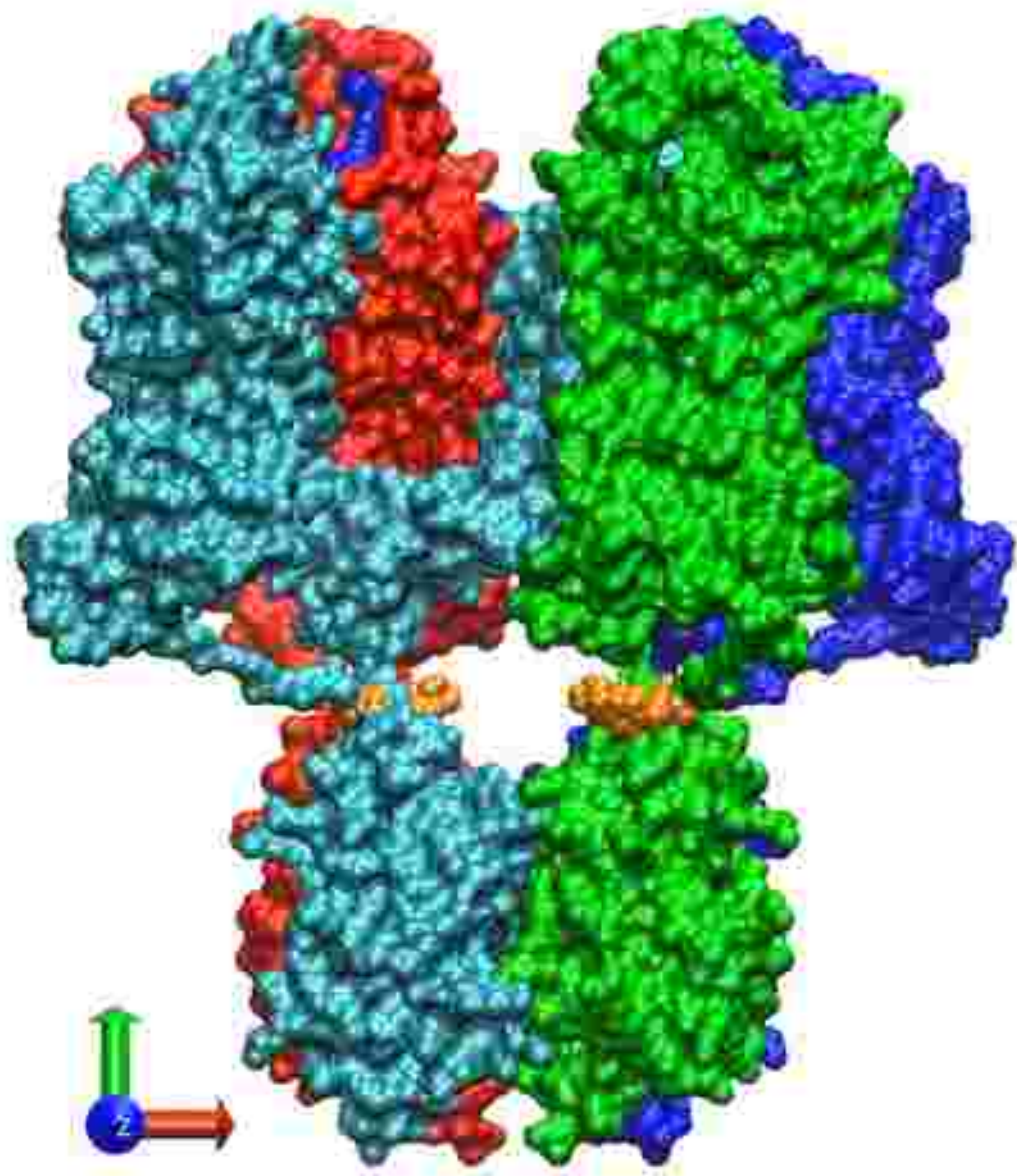
Molden is a package for displaying Molecular Density from the Ab Initio packages GAMESS-UK, GAMESS-US and GAUSSIAN and the Semi-Empirical packages Mopac/Ampac, it also supports a number of other programs via the Molden Format.

Molden reads all the required information from the GAMESS / GAUSSIAN outputfile. Molden is capable of displaying Molecular Orbitals, the electron density and the Molecular minus Atomic density. Either the spherically averaged atomic density or the oriented ground state atomic density can be subtracted for a number of standard basis sets.

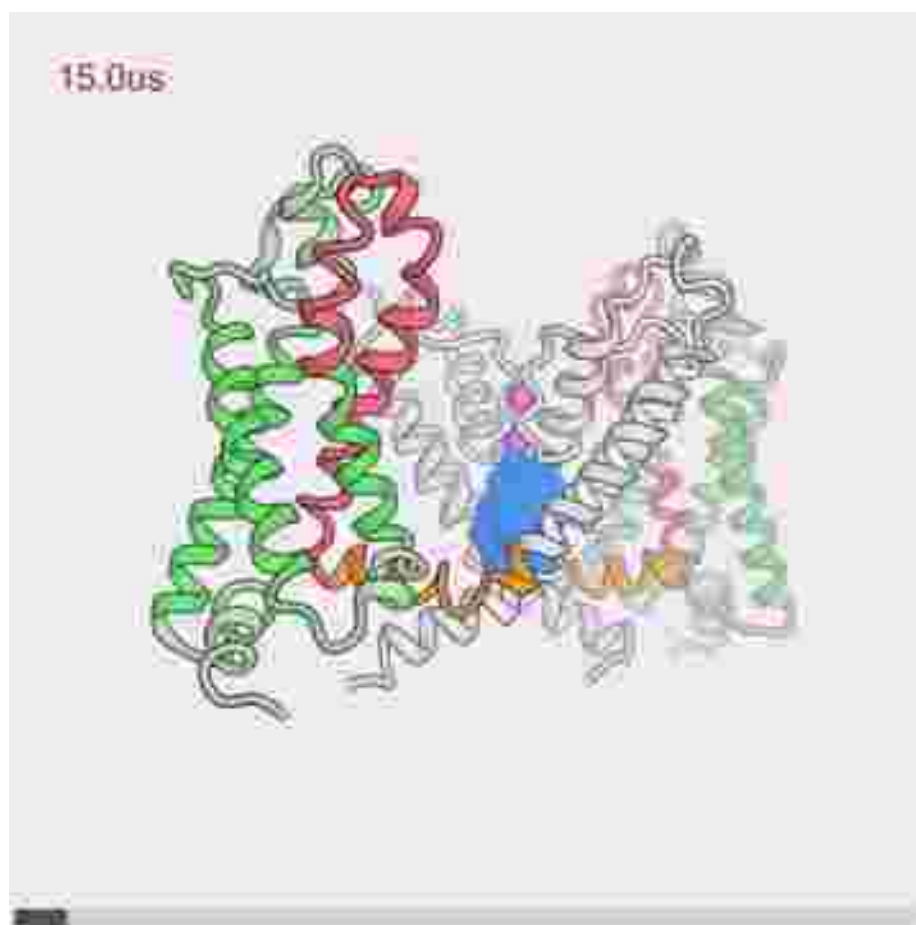
Molden supports contour plots, 3-d grid plots with hidden lines and a combination of both. It can write a variety of graphics instructions; postscript, XWindows, VRML, povray, OpenGL, tekronix4014, hpgl, hp2392 and Figure



Protein movies



Protein movies



Molecular Structure of Ion Channels

