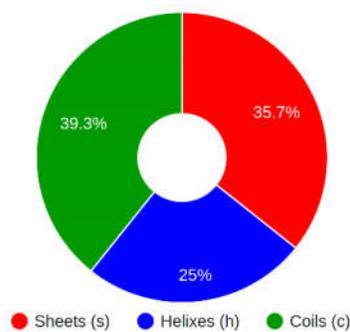
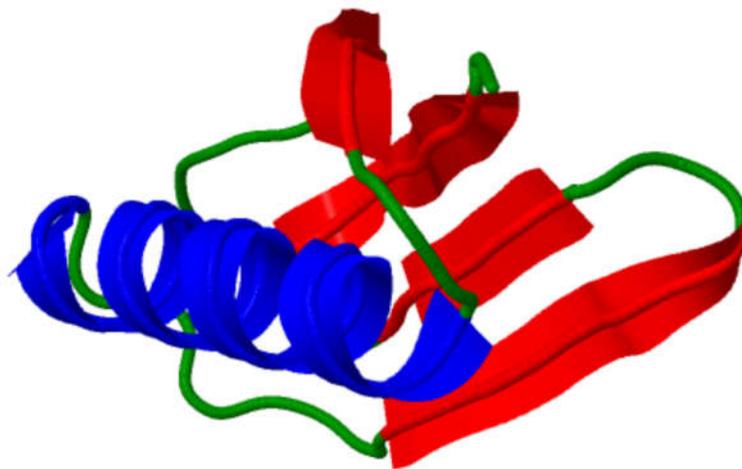


# ProSSA: (Hydrogen-Bond-Independent) Proteins' Secondary Structure Assignment

Concise User Guide



<http://bioinformatics.center/servers/ProSSA>

December, 2015

[1] To use ProSSA, visit <http://bioinformatics.center/servers/ProSSA>  
You should see a page that looks like the screenshot below.

The screenshot shows a web browser window with the URL [bioinformatics.center/servers/ProSSA](http://bioinformatics.center/servers/ProSSA). The page has a blue header with a navigation menu: Home, Tools, Web Servers, Research Interests, Portals, and Blogs. Below the header is a logo for the Bioinformatics Center and the title "Hydrogen-Bond-Independent Proteins' Secondary Structure Assignment".

Below the header, the page content includes:

- A breadcrumb trail: "You are here: Home >> Online Tools >> Hydrogen-Bond-Independent Proteins' Secondary Structure Assignment".
- A link: "You can log in or create a new account here."
- A sidebar menu with the following items: Home, Tools, RaFoSA, Web Servers, **ProSSA**, Research Interests, Cellular Automata, Computational Epidemiology, Computer Vision, Malaria, Web Indexing, Portals, User Account, and Blogs.
- A main content area with the title "ProSSA: (Hydrogen-Bond-Independent) Proteins' Secondary Structure Assignment".
- A section titled "Select an Option to see additional information, usage description, and examples." with four radio button options:
  - Upload PDB File
  - Enter Content in PDB Format
  - Specify PDB ID
  - Upload Trajectory in PSF and DCD Format (Save bandwidth: only the coordinates of Alpha Carbon Atoms, CA, are needed)
- A note: "NOTE: Secondary Structure (SS) assignment is done for **amino acids alone**, and only the **first chain of every frame** is read and processed so as to avoid confusing the user. If you must work with many chains, you must submit them separately (otherwise only the first chain will be processed)."

[2] Select the approach/format through which you intend to provide input for ProSSA.

There are four options available. You can either

[A] Upload PDB File

[B] Enter Content in PDB Format

[C] Specify PDB ID  or

[D] Upload Trajectory in PSF and DCD Format

Regardless of the approach/format you select,

[I] Secondary structure assignment is done for protein residues alone (i.e. for amino acids alone). DNA, RNA, water, or ligand molecules are ignored.

[II] Only the first protein chain in the data provided is used.

[III] All the frames/models are processed. In fact, we encourage submitting multiples frames whenever user has more than one frame/model for the same molecule of interest.

[IV] Only the information of Alpha Carbon atoms are needed and used regardless of whether the user submits all-atom information or not.

[V] We encourage users to submit information of Alpha Carbon atoms alone so as to save bandwidth, especially when submitting multiple jobs with multiple frames/models. We provide a tutorial on how to create PSF, DCD, or PDB file that contains only Alpha Carbon atoms here <http://bioinformatics.center/blogs/1001/1001>

[Option A] If the user chooses to submit a PDB File, a form for uploading PDB File is displayed such as in the screenshot below.

## ProSSA: (Hydrogen-Bond-Independent) Proteins' Secondary Structure Assignment

Select an Option to see additional information, usage description, and examples.

- Upload PDB File
- Enter Content in PDB Format
- Specify PDB ID
- Upload Trajectory in PSF and DCD Format  
(Save bandwidth: only the coordinates of Alpha Carbon Atoms, CA, are needed)

*NOTE: Secondary Structure (SS) assignment is done for **amino acids alone**, and only the **first chain of every frame** is read and processed so as to avoid confusing the user. If you must work with many chains, you must submit them separately (otherwise only the first chain will be processed).*

Choose PDB File to Upload

No file chosen

[Example: [sampleFile.pdb](#)]

[Option B] If the user chooses to Enter Content in PDB Format, a form for entering the PDB Content is displayed such as in the screenshot below.

Users who do not have enough knowledge about PDB Format should generally avoid using this option. If necessary user may refer to this <http://bioinformatics.center/blogs/1001/1002> short tutorial on PDB File Format.

## ProSSA: (Hydrogen-Bond-Independent) Proteins' Secondary Structure Assignment

Select an Option to see additional information, usage description, and examples.

- Upload PDB File
- Enter Content in PDB Format
- Specify PDB ID
- Upload Trajectory in PSF and DCD Format  
(Save bandwidth: only the coordinates of Alpha Carbon Atoms, CA, are needed)

*NOTE: Secondary Structure (SS) assignment is done for **amino acids alone**, and only the **first chain of every frame** is read and processed so as to avoid confusing the user. If you must work with many chains, you must submit them separately (otherwise only the first chain will be processed).*

Enter (or Copy and Paste) Coordinates\* of the Protein in the Space Below

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	

Submit ▶

\*Note 1: Entered Text Must be in PDB Format.

Note 2: If you do not know what PDB format it, we recommend that you do not use this option at all. If you must use this option, please, refer to this [short blog post on PDB Format](#).

Example: The following is a sample content.

```
1 MODEL 1
2 ATOM 1 CA LYS A 1 -11.635 -7.622 5.165 0.00 0.00
3 ATOM 2 CA VAL A 2 -8.492 -8.238 7.154 0.00 0.00
4 ATOM 3 CA PHE A 3 -7.627 -11.887 7.417 0.00 0.00
5 ATOM 4 CA GLY A 4 -6.122 -13.158 10.743 0.00 0.00
6 ATOM 5 CA ARG A 5 -2.682 -15.071 10.450 0.00 0.00
7 ATOM 6 CA CYS A 6 -4.180 -18.509 11.356 0.00 0.00
8 ATOM 7 CA GLU A 7 -7.491 -18.080 9.268 0.00 0.00
9 ATOM 8 CA LEU A 8 -5.284 -17.238 6.239 0.00 0.00
10 ATOM 9 CA ALA A 9 -2.969 -20.282 6.614 0.00 0.00
11 ATOM 10 CA SER A 10 -2.910 -20.500 8.150 0.00 0.00
12 ATOM 11 CA THR A 11 -2.910 -20.500 8.150 0.00 0.00
```

[Option C] If the user chooses to Specify PDB ID, a form for specifying the PDB ID is displayed such as in the screenshot below.

The PDB ID must be four characters in length and must be a valid identifier of a record in the Protein Data Bank. The record must contain protein. It does not matter whether the PDB record contains protein alone or protein alongside other biomolecules; but the record must contain protein.

## ProSSA: (Hydrogen-Bond-Independent) Proteins' Secondary Structure Assignment

Select an Option to see additional information, usage description, and examples.

- Upload PDB File
- Enter Content in PDB Format
- Specify PDB ID
- Upload Trajectory in PSF and DCD Format  
(Save bandwidth: only the coordinates of Alpha Carbon Atoms, CA, are needed)

*NOTE: Secondary Structure (SS) assignment is done for **amino acids alone**, and only the **first chain of every frame** is read and processed so as to avoid confusing the user. If you must work with many chains, you must submit them separately (otherwise only the first chain will be processed).*

Enter Four-Character\* PDB ID:

[Example: 2mqo]

Submit ID

\*Note: Entered Characters Must Represent a Valid PDB ID.

[Option D] If the user chooses to provide input in PSF and DCD Format, a form for uploading the PSF and DCD Files is displayed such as in the screenshot below.

We encourage users to submit information of Alpha Carbon atoms alone so as to save bandwidth, especially when submitting multiple jobs with multiple frames/models. We provide a tutorial on how to create PSF, DCD, or PDB file that contains only Alpha Carbon atoms here <http://bioinformatics.center/blogs/1001/1001>

## ProSSA: (Hydrogen-Bond-Independent) Proteins' Secondary Structure Assignment

Select an Option to see additional information, usage description, and examples.

- Upload PDB File
- Enter Content in PDB Format
- Specify PDB ID
- Upload Trajectory in PSF and DCD Format  
(Save bandwidth: only the coordinates of Alpha Carbon Atoms, CA, are needed)

*NOTE: Secondary Structure (SS) assignment is done for **amino acids alone**, and only the **first chain of every frame** is read and processed so as to avoid confusing the user. If you must work with many chains, you must submit them separately (otherwise only the first chain will be processed).*

Choose PSF\* File to Upload

Choose File No file chosen

[Example: [sampleFile\\_onlyCA.psf](#)]

Choose DCD\* File to Upload

Choose File No file chosen

[Example: [sampleFile\\_onlyCA.dcd](#)]

Submit ▶

\*Note: Only the coordinates of Alpha Carbon atoms (CA alone) are needed regardless of whether you submit all-atom information or not. Therefore, to save bandwidth, please upload PSF and DCD files that contain information for CA alone. If you do not know how to save PSF and DCD files that contain information of CA atoms alone, please, read [this](#) very short tutorial/blog post [ [Creating PSF and DCD Files That Contain Only Some Selected Atoms](#) ] we created.

[3] Once you have selected an option in [2] above and clicked submit, you will be provide a link to the result page and the calculations will be start immediately. (The available computation resources are shared among ongoing computations. So, computations may take a moment.) Follow the link provided. If the computation is not ready, you will be told how much of additional time is needed. You will need to refresh the result page once the time has elapsed.

[4] Once the SS assignment is finished, you will see a result page similar to the one shown in the screenshots below (Part 1 to Part 3 of Results).

[Part 1 of Results] This part of the results is always available. It shows the job's ID, and the assigned secondary structure (in both 3-class and 7-class systems). If you submitted more than one frame/model, the results for all the frames are shown. The results for a frame span a line.

## ProSSA: (Hydrogen-Bond-Independent) Proteins' Secondary Structure Assignment

**Result Page**  
Job ID: 568a5a98g

**Assigned Secondary Structure, 3-Class\* System** [ [Download](#) ]

Frames|Secondary Structures\* Assigned to the Residues

1	ccssssssccccccccssssccchhhhhhhhhhhhhccccccccssssccccsscc
2	cccssccccccccccccccccccccchhhccchhhcccccccccccccccccccc
3	ccccccccchhhccccccccssscchhhhhccccccccccccccccccccsscccc
4	ccccccccccccccccccccccccccccchhhhhcccccccccccccccccccccccc
5	cc
6	ccccccccchhhhhccccccccccccchhhhhcccccccccccccccccccccccc
7	ccccccccchhhhhccccccccccccchhhhhccccchhhcccccccccccccccc
8	ccccccccchhhhhccccccccccccccccchhhcccccccccccccccccccccc
9	ccccccccccccccccccccccccccccccccchhhhhcccccccccccccccccccc
10	ccccccccchhhhhccccccccccccccccchhhhhcccccccccccccccccccc
11	chhhccccchhhccccccccccccccccchhhhhhhcccccccccccccccccccc
12	ccccccccchhhhhccccccccccccccccchhhhhhhcccccccccccccccccccc

\* c = coil; h = helix; s = sheet

**Assigned Secondary Structure, 7-Class\*\* System** [ [Download](#) ]

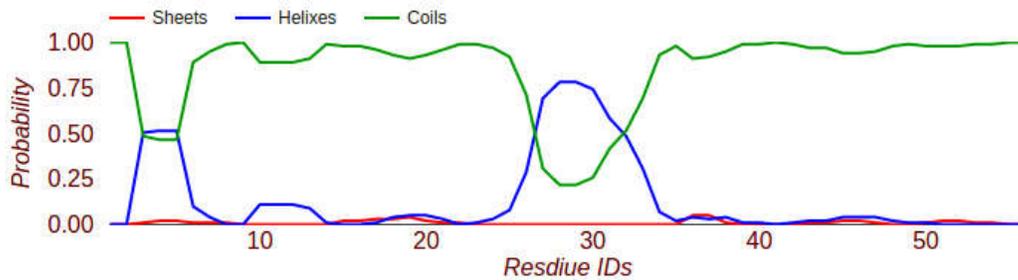
Frames|Secondary Structures\*\* Assigned to the Residues

1	--EEEEEE-SS--SEEEEE-SSHHHHHHHHHHHHHTT---EEEEETTTEEEE--
2	---EET---TT-SSEEE---S-STHHHTTSHHHHSSS-SS---TSTT-----
3	----ST-SSHHHSSS--EEEEESSHHHHHHS--TTTTT-----STTEE----
4	----S--SSTT-SS---EESSHHHHHTS--S-SST-TS---EESTT-----
5	----SS-STT--S--EEESTTSTTGGGT--TT-SSSS-TS--ST-TTTSS-S--
6	--TTTS--SHHHT---SSS-SSSTHHHH-S--SSSS-S--S--STTSSSS--
7	--STTT--SHHHH----S--SSSHHHH--S-SHHHTT---S---TT-ST----
8	---TTT--SHHHS-----SST--HHHTTTTSS-TT-S--S--TTTSS----
9	----TT--TTTTT-----SSS--HHHHHTTS-TT-S--S-TSTT---SS--
10	--TTTTT--HHHHTS-----TTTTHHHHHTT---TTTT--TT--TS-SSS---
11	--HHHTT-SHHHTS-----SSHHHHHHHTT---TTTT--TT-TTS--SS---
12	--GGGTT--HHHHTS-----S-SSSTHHHHHT---TTTT--TT-STTSSTS---

\*\* H = alpha helix; B = beta bridge; E = strand; G = helix-3;  
I = helix-5; T = turn; S = bend; - = unknown

[Part 2 of Results] This part of the results is available only if more than one frame was submitted. This Part 2 of Results shows the probability that a residue at a position in the protein is part of a sheet, helix, or coil.

### Distribution of the Assigned Secondary Structures, 3-Class System



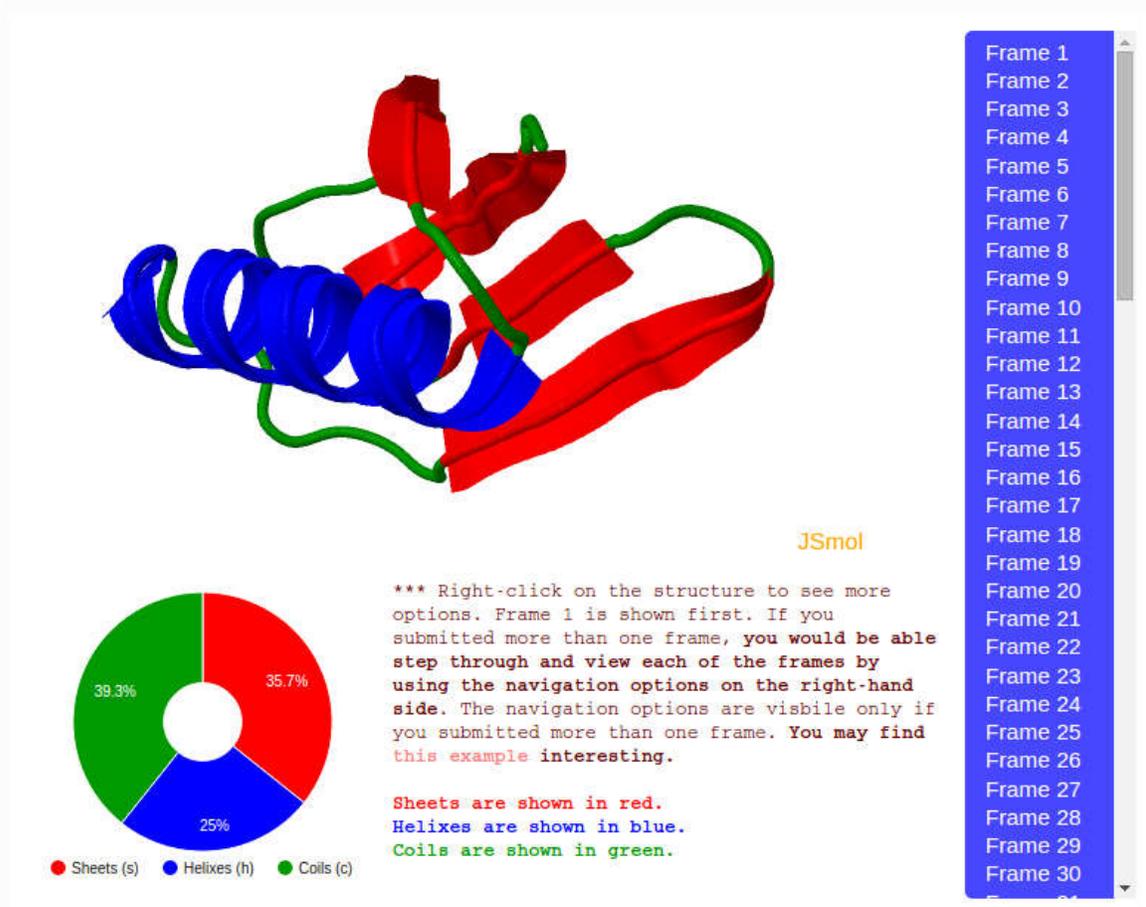
Position the mouse pointer on the chart to see each residue name, residue ID, frequency of sheets/helixes/coils [F(Sheet)/F(Helixes)/F(Coils)] for each residue, and probability of sheets/helixes/coils [P(Sheet)/P(Helixes)/P(Coils)] for each residue.

[Part 3 of Results] This part of the results is also always available, except the blue panel on the right-hand side which is available only if more than one frame/model was submitted. This part of the results shows the 3-D structure of the protein and the assigned secondary structure. Sheets are shown in red, helixes in blue, and coils in green. This part also shows the distribution of sheets, helixes, and coils in the protein based on the number of residues assigned to each of these three secondary structure types. This Part 3 of Results also allows the user to step through and visualize all the frames (if more than one frame was submitted). For example, we show the first three frames from job "568a5a98g" in the screenshots below.

[Frame 1 of Job 568a5a98g] For all jobs submitted to ProSSA, the results for Frame 1 is shown first. The screenshot below is from job of Job 568a5a98g available here <http://bioinformatics.center/servers/ProSSA/results/568a5a98g#ssVisualization> . The user can view any frame he/she wants to show using the blue panel shown on the right-hand side.

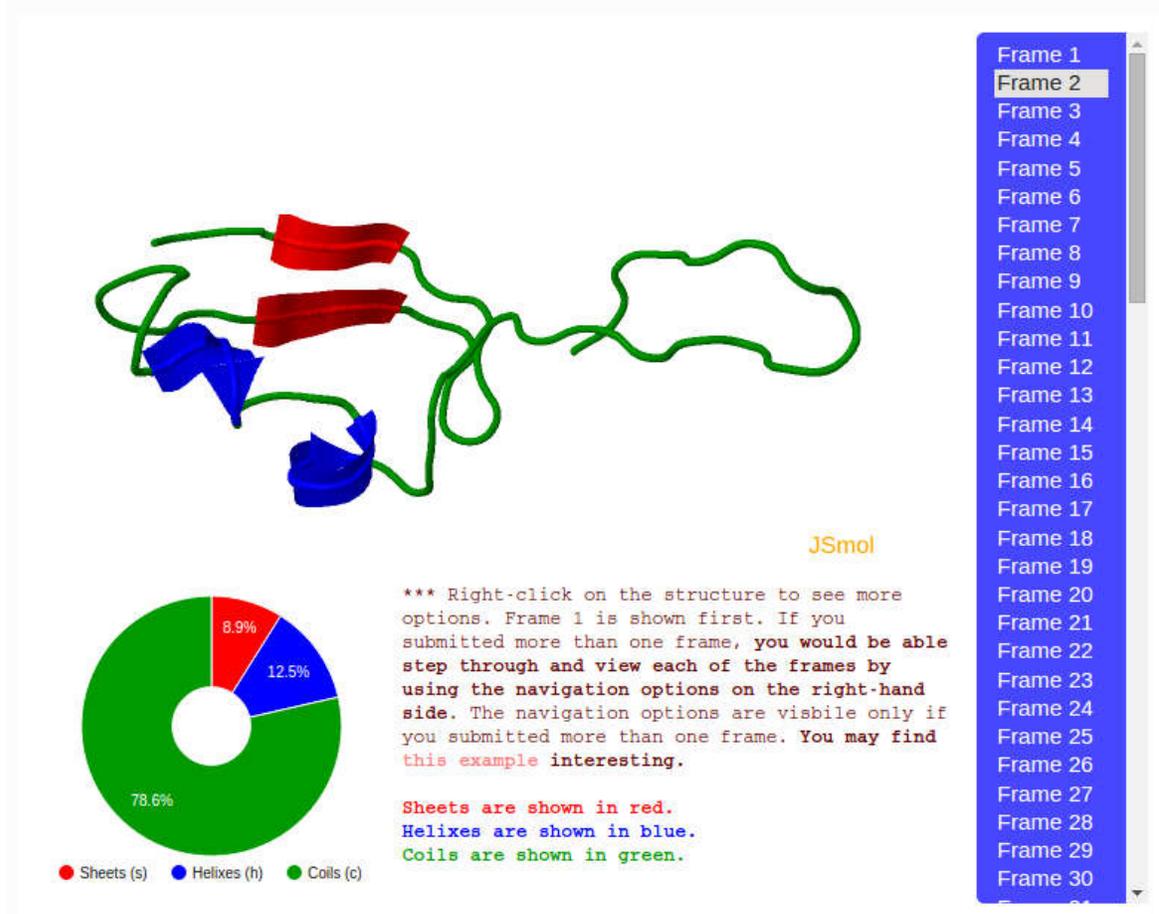
NOTE: The blue panel on the right-hand side is available only if more than one frame/model was submitted.

### Visualization of the Molecule Based on the Assigned Secondary Structures\*\*\*



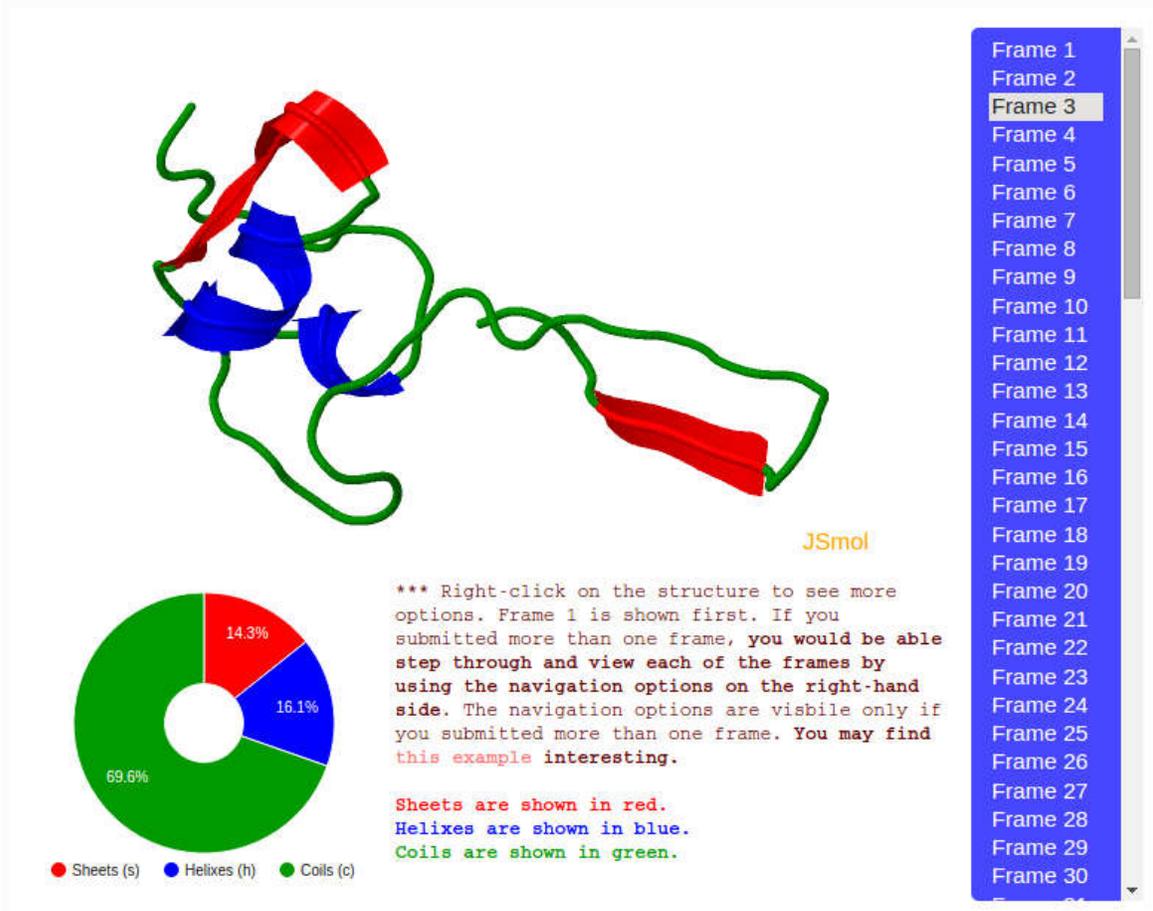
[Frame 2 of Job 568a5a98g] You can select "Frame 2" in the blue panel shown on the right to see the visualization of the assigned secondary structures for the submitted frame/model 2.

### Visualization of the Molecule Based on the Assigned Secondary Structures\*\*\*



[Frame 3 of Job 568a5a98g] You can select "Frame 3" in the blue panel shown on the right to see the visualization of the assigned secondary structures for the submitted frame/model 3.

### Visualization of the Molecule Based on the Assigned Secondary Structures\*\*\*



All the other frames can also be accessed.

Please, <http://bioinformatics.center/servers/ProSSA/results/568a5a98g#ssVisualization> go here to see this specific example.

If you need additional information or additional help, please do not hesitate to contact us.

Email: [emmanuel@gapp.nthu.edu.tw](mailto:emmanuel@gapp.nthu.edu.tw), [es13acp@herts.ac.uk](mailto:es13acp@herts.ac.uk)

Tel: +18133888836

<http://bioinformatics.center/servers/ProSSA>