Protein Molecular Surface Calculator (ProMS)
Version 1.0

McGill University
Email: chomatt10@gmail.com
Website: coming soon

User Manual & Installation Guide

Date Created: April 1st, 2022
Date Updated: May 3rd, 2022

Table of Contents

1. About

Protein molecular surface calculator (ProMS) is an integrated application which calculates the surface properties of a given protein, while also providing a built-in visualization module. It is the successor to Nicolau's "Proteins Surface Property Calculator (PSPC)" [1]. The program was written mostly in Python. It uses Connolly's msRoll executable program to generate the molecular surface, and then Nicolau's Method 1 and Method 2 algorithms to calculate the surface properties [2, 3]. Additionally, Schrodinger's PyMOL is used as a molecular manipulation tool and to perform the visualizations [4].

The major advantage of ProMS is that it streamlines many processes into a single application. Through its easy-to-use graphical user interface (GUI), the user can select a protein and their desired parameters, and the protein will be subsequently "cleaned" of any unwanted molecules (solvent, hydrogens, heteroatoms, etc.). Its surface properties will then be calculated and there is also the option to output the 3-dimensional structure of the protein surface through the built-in custom visualization modules. This is all done "under the hood" at the click of a button.

By inputting a PDB ID, ProMS will calculate the following primary properties: total surface area, positive/negative area, hydrophilic/hydrophobic area, total positive/negative area, and total hydrophilic/hydrophobic area. Furthermore, secondary properties, such as density, specific density, and area extent are also outputted. These calculations are dependent on the probe radius and the pH, which are parameters specified by the user in the GUI prior to computation. ProMS was specifically developed to study surface hydrophobicity and hydrophilicity. Via the PyMOL interface, the molecular surface, custom-defined in terms of relative hydrophobicity, can be visualized and manipulated. There is the option to perform hydrophobicity calculations and visualizations at atomic or at amino acid resolution. Additionally, the application includes a "batch mode", which enables the autonomous computation of multiple proteins (with identical or different parameters) consecutively. Lastly, ProMS allows for compatibility with "custom PDBs", meaning it functions with PDBs that may have been modified by the user or that are not stored in an online database.

The runtime of SURFace calculator is limited by the execution speed of Connolly's msRoll program. The generation of the molecular surface is linear in the number of atoms in the molecule, and quadratic in the probe radius. A small protein using a probe radius of 1.4A can be run in a few seconds, whereas a larger protein using a large probe radius of 20A will take several hours to run to completion. Once the msRoll program has finished, the visualization in PyMOL (if enabled) can take a few minutes depending on the size of the protein, but it is generally quite fast.

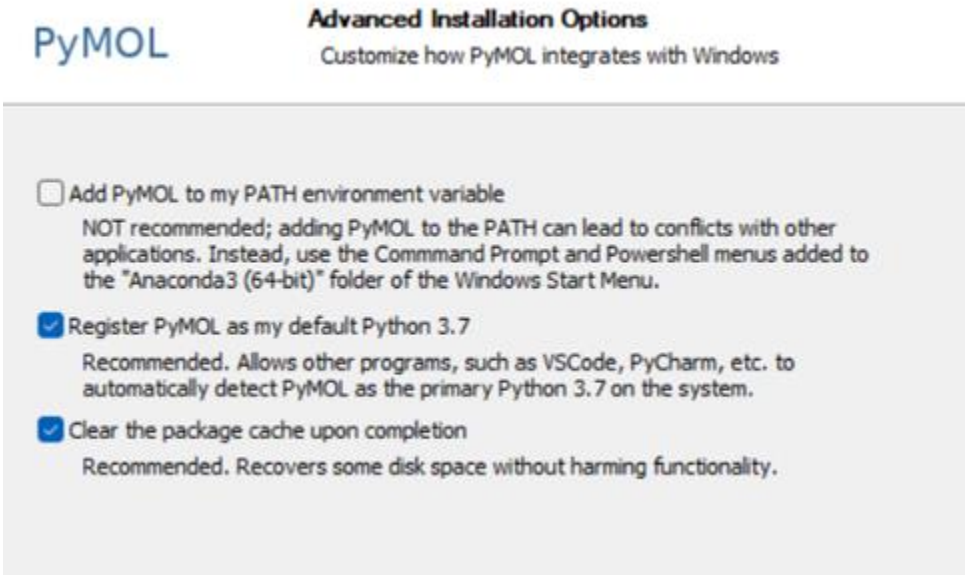
2. Installation

This is the quick & easy installation guide for the ProMS application. Installing it this way allows for the usage of the app and all its features. ProMS is restricted to use on Windows operating systems only due to limitations with Connolly's msRoll. This guide does not include instructions regarding accessing or modifying the source code. There is a video guide that can be viewed in conjunction with these instructions.

1. Download the source code file, "ProMS.zip". Do not unzip.

2. Download and install Schrodinger's PyMOL from their website, https://pymol.org/2/

- The easiest way to install is through the EXE installer, and use the settings shown below

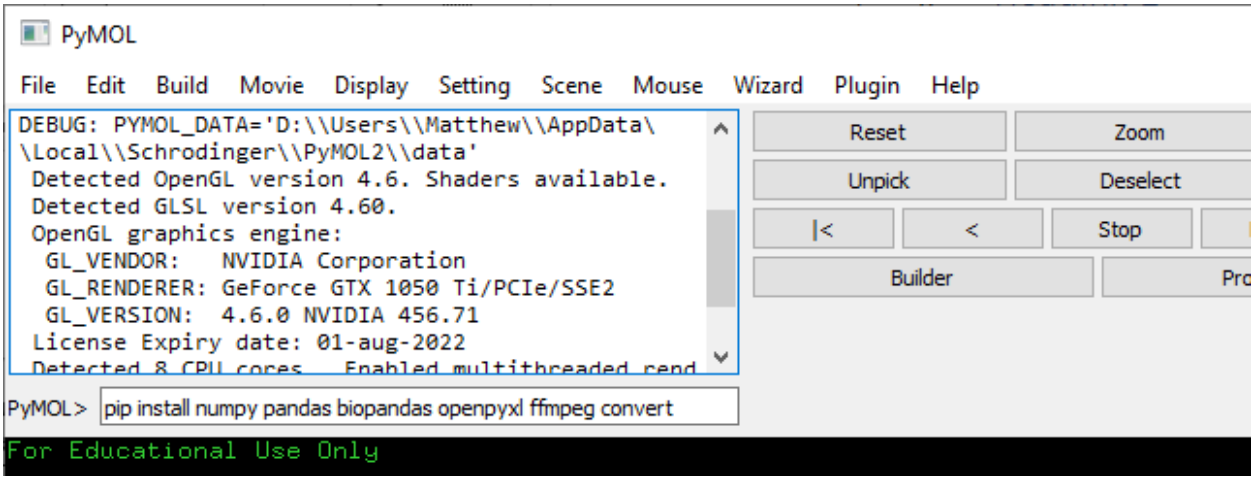


- PyMOL requires a license file. If you are a student, this file can be obtained for free from their website (https://pymol.org/edu/).

3. Once installed, launch PyMOL.

4. In the PyMOL command line, paste in the following text (without the bullet point):

- pip install numpy pandas biopandas openpyxl ffmpeg convert
- See image below for a reference



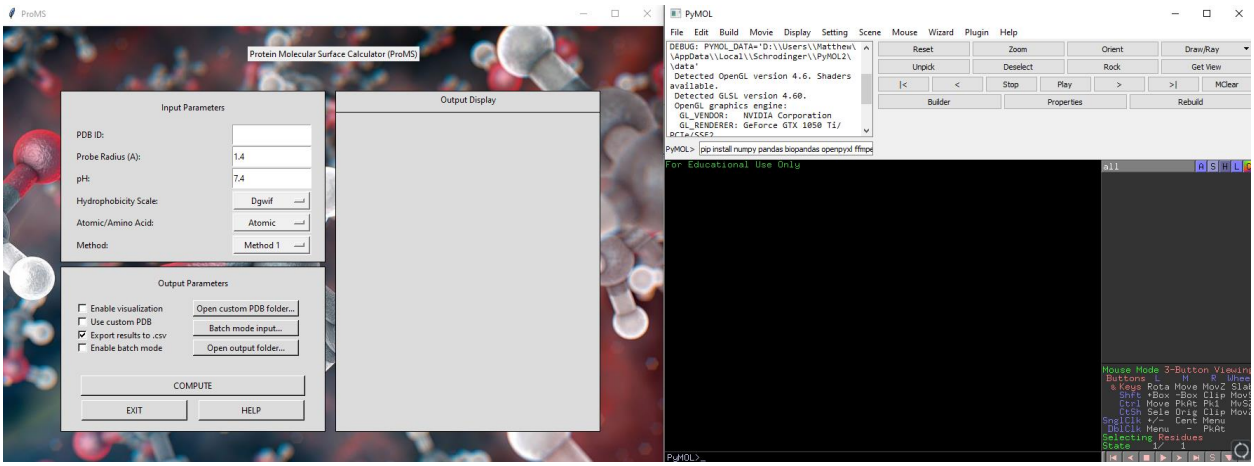
5. Navigate to Plugin (top menu) → Plugin Manager → Install New Plugin.

6. Under “Install from local file”, click on “Choose file…”.

7. When prompted, select the source code zip file “ProMS.zip” that was downloaded in step 1.

8. After a few seconds, a popup saying “Plugin installed successfully” show appear.

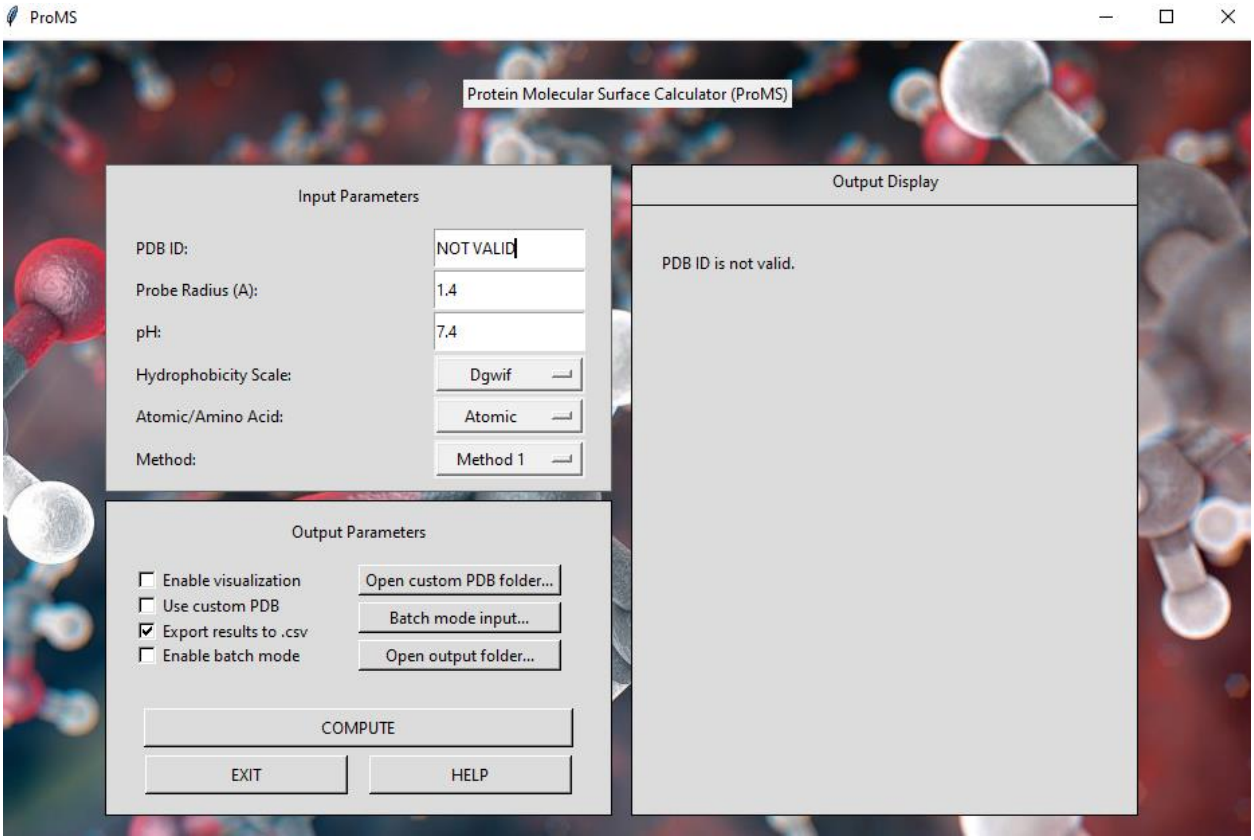
9. To launch ProMS once installed, return to the PyMOL default screen. From there, navigate to Plugin → Legacy Plugins → ProMS. Upon clicking “ProMS”, the GUI should launch (see image below).



* If experiencing issues installing ProMS, please contact the email address provided with the error you are encountering as well as the stack trace.

3. Usage

This section describes how to use ProMS. There is also a reference video which goes over the main features. Upon launch of ProMS as a PyMOL plugin, a separate application will open, as seen below.



ProMS GUI

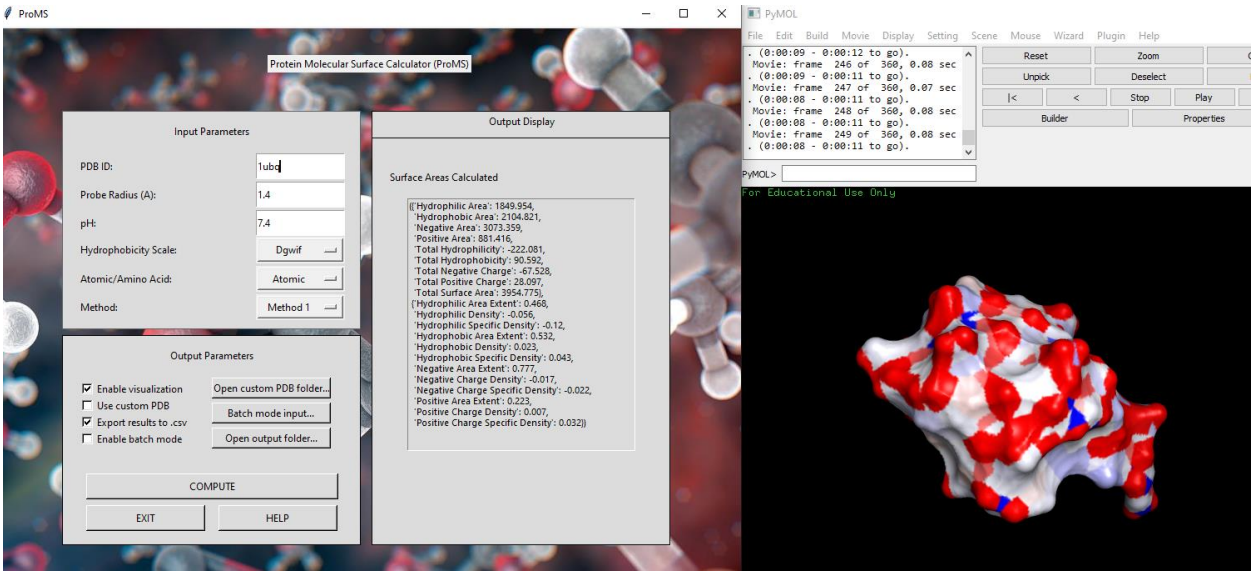
Input Parameters:

The input parameters are entered by the user and will directly affect the calculations and output generated by the application. The PDB ID input field takes the unique 4-letter/digit code of the protein (or molecule). The capitalization of the PDB ID is not necessary. Next, the user must enter a probe radius and pH, which are set by default to 1.4A and 7.4, respectively. The probe radius must be between 1.4 and 20 angstroms, and the pH between 0 and 14. Please note that the radii values must be discrete if not using 1.4A (i.e., 2, 3, 4, …, 20), and the pH values can have at most one decimal value. If the provided PDB ID is invalid, or the probe radius and/or pH are outside the allowed ranges, no output will be generated. However, the app will clearly display the error and prompt the user to re-enter valid inputs, as seen in the image above. The last three input parameters open drop-down lists when clicked. The current iteration of ProMS has two

hydrophobicity scales built-in: Dgwif and ST. The hydrophobicity values can be calculated via "Method 1" or "Method 2", as discussed previously. Lastly, the user has the choice to generate the output at atomic or amino acid resolution. Once the input parameters are all chosen, the "COMPUTE" button can be clicked, and the app will perform the protein surface calculations.

Output Parameters:

The second section of the GUI allows the user to customize the output. If visualization is enabled, a PyMOL window will open alongside the GUI, displaying the molecular surface computed from the input parameters, as seen in the image below. Once computed, several output files are created and saved by default. Every time a molecule is run successfully, the following files are generated: an atomic area .csv file, a volume .txt file, and a .txt and a .csv file containing the calculated surface properties. If visualization is enabled, the PyMOL session will also be saved, and a ~12 second movie (.mov) file will also be created of the protein's molecular surface being rotated so that every side is visible.



Ubiquitin (1UBQ) protein surface properties and visualization via ProMS

Batch Mode:

An important feature of this application is batch mode. This gives the user the opportunity to run many molecules with different input parameters autonomously and continuously. To enable this option, select "Enable batch mode" and click the "Batch mode input" button. This will prompt the opening of a .csv file named "BatchModeInput". The user must then enter their desired input parameters, as seen in the image below. If batch mode is enabled, the values in the "Input Parameters" frame are ignored. Of note, visualization cannot function in conjunction with batch mode. Therefore, if the protein is to be visualized, it must be entered manually in "Input Parameters".

	A	B	C	D	E	F
1	PDB ID	Radius	pH	Scale	Method	Resolution
2	1LYZ	1.4	7.4	Dgwif	1	AA
3	1LYZ	2	7.4	Dgwif	1	AA
4	1LYZ	3	7.4	Dgwif	1	AA
5	1LYZ	4	7.4	Dgwif	1	AA
6	1LYZ	1.4	7.4	ST	1	Atomic
7	1LYZ	2	7.4	ST	1	Atomic
8	1LYZ	3	7.4	ST	1	Atomic
9	1LYZ	4	7.4	ST	1	Atomic

BatchModeInput.csv with sample proteins to be run via “batch mode”

Custom PDBs:

ProMS also allows for custom PDBs to be run. Custom PDBs refer to PDBs that are not found on online databases or that have been modified by the other. To use this feature, first checkmark the “Use custom PDB” box. Then, click on “Open custom PDB folder…” and drag the custom PDB(s) into the folder that pops up. Lastly, click “COMPUTE”, and when prompted, select the custom PDB that is to be analyzed.

4. Future Work

ProMS has been developed as an open-source plugin which extends Schrodinger’s PyMOL. This is the first iteration of the app (v1.0) and therefore we expect we will be adding new features to improve functionality and user experience in the future. Specifically, future plans include adding more hydrophobicity scales, visualization modules, and enabling visualization within batch mode. Furthermore, the open-source nature enables the opportunity to obtain valuable input and feature suggestions from the community. The source code has been purposely written such that new components can be seamlessly integrated within the application.

5. References

[1] D. V. Nicolau Jr, F. Fulga, and D. V. Nicolau Sr, "A new program to compute the surface properties of biomolecules," in *Proceedings of the First Asia-Pacific bioinformatics conference on Bioinformatics 2003-Volume 19*, 2003, pp. 29-34.

[2] M. L. Connolly, "The molecular surface package," *Journal of molecular graphics*, vol. 11, no. 2, pp. 139-141, 1993.

[3] D. V. Nicolau Jr, E. Paszek, F. Fulga, and D. V. Nicolau, "Mapping Hydrophobicity on the Protein Molecular Surface at Atom-Level Resolution," *PLOS ONE*, vol. 9, no. 12, p. e114042, 2014, doi: 10.1371/journal.pone.0114042.

[4] W. L. DeLano, "Pymol: An open-source molecular graphics tool," *CCP4 Newsl. Protein Crystallogr*, vol. 40, no. 1, pp. 82-92, 2002.