Manual

QSAR Lapp is an application designed to conveniently create Gaussian 16 inputs, combine crystalline objects (e.g., nanomaterials) with conventional molecules (e.g., amino acids), create structures crystalline material – dipole in four different geometrical orientations (placement of the finite dipole over the material) and calculate original descriptors (developed by QSAR Lab team) based on Gaussian output file. The application was written entirely in python (version 3.11) using popular libraries such as NumPy (version 1.25.1) and Pandas (version 1.5.2).

The algorithm for placing a dipole above the NM and combining crystalline objects with conventional particles at distance 5 Å is based on simple symmetry operations and linear transformations using quaternion algebra, which are characterized by an efficient numerical procedure, significantly minimizing the time of rotation and translation operations in 3D space.

1. Gaussian Input Creator

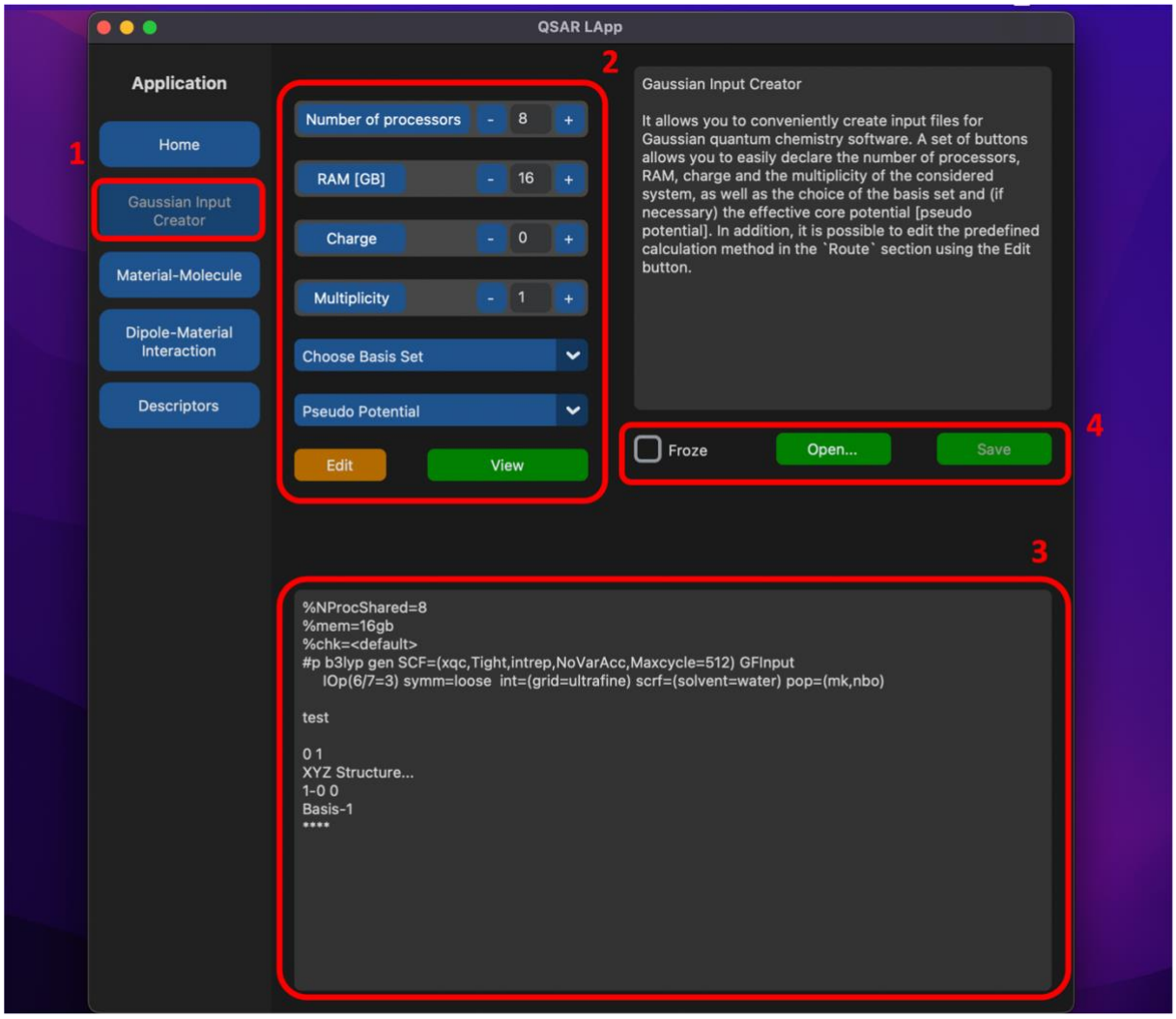

*Figure 1 Active section - Gaussian Input Creator [1], Settings buttons [2], a console with a preview of settings and computational methods [3] and buttons to load *xyz structures files, save results Gaussian input files, and Froze checkbox to lock atoms in input file [4].*

Gaussian Input Creator allows you to conveniently create input files for Gaussian 16 quantum chemistry software. A set of buttons allows you to easily declare the number of processors, RAM, charge, and the multiplicity of the considered system, as well as the choice of the basis set and (if necessary) the effective core potential. In addition, it is possible to edit the predefined calculation method in the `Route` section using the Edit button.

The active Gaussian Input Creator section is shown in Figure 1 below (number 1). In the panel marked No. 2, you can declare the method and computational resources, which will appear in the Gaussian input as a result. In the console (number 3) you can see a preview of the currently selected values and a place where you can load *.xyz coordinates (marked by 'XYZ structure'). If you want to personalize the calculation method, use the 'Edit' button in the panel marked with number 2 – you can easily change, for example, the default B3LYP functional to another (see Figure 2).

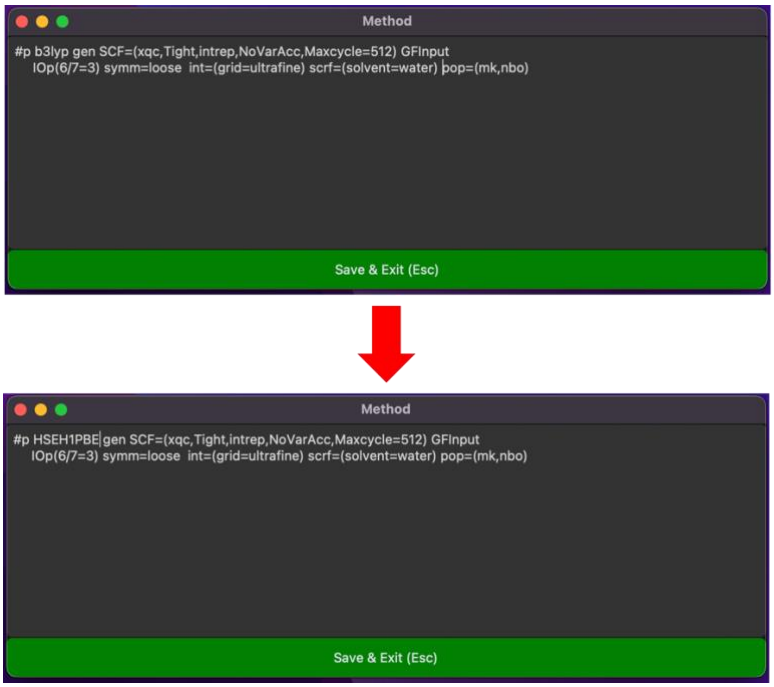

Figure 2 Computational method editing window - you can easily personalize the method by editing the route line. The change of the B3LYP functional is presented in the figure.

To load structure or structures in xyz Cartesian format file select the ‘Open…’ option in the panel marked with the number 4 in Figure 1. You can load one or more xyz files, and then you can set all of them as required (panel 1) and preview in the console by pressing the 'View' button. The process of loading files containing Cartesian coordinates with the change of the number of processors, RAM memory, basis set and B3LYP functional to HSEH1PBE presented below. You can then use the 'Save' button to create and save all input files in *.com format suitable for Gaussian 16 in the specified location. If necessary, you can lock the atoms before saving using the ‘Froze’ checkbox.

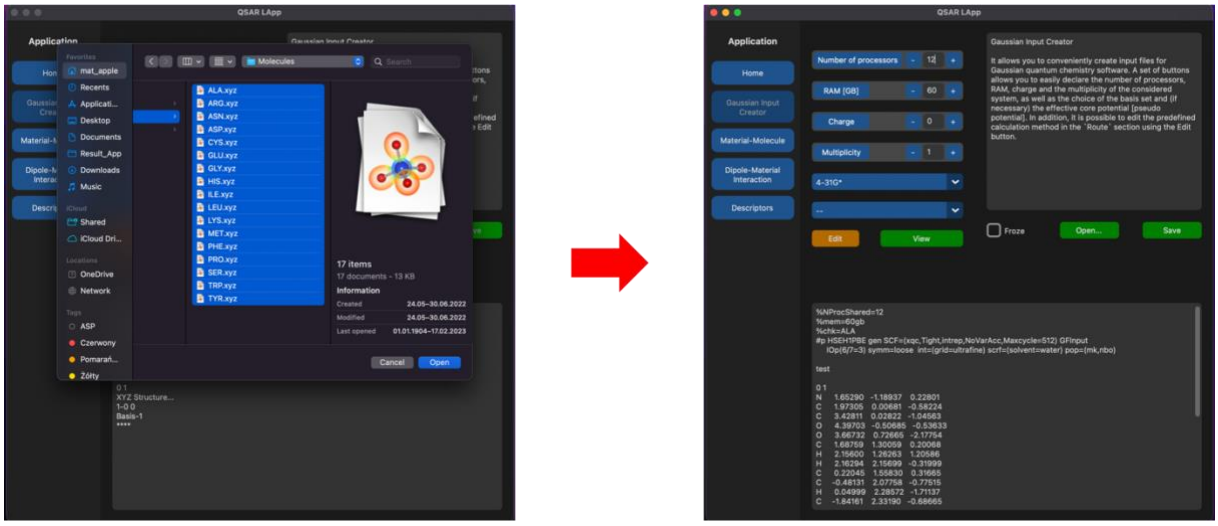

Figure 3 The process of loading structures and setting the computational method.

2. Material-Molecule Creator

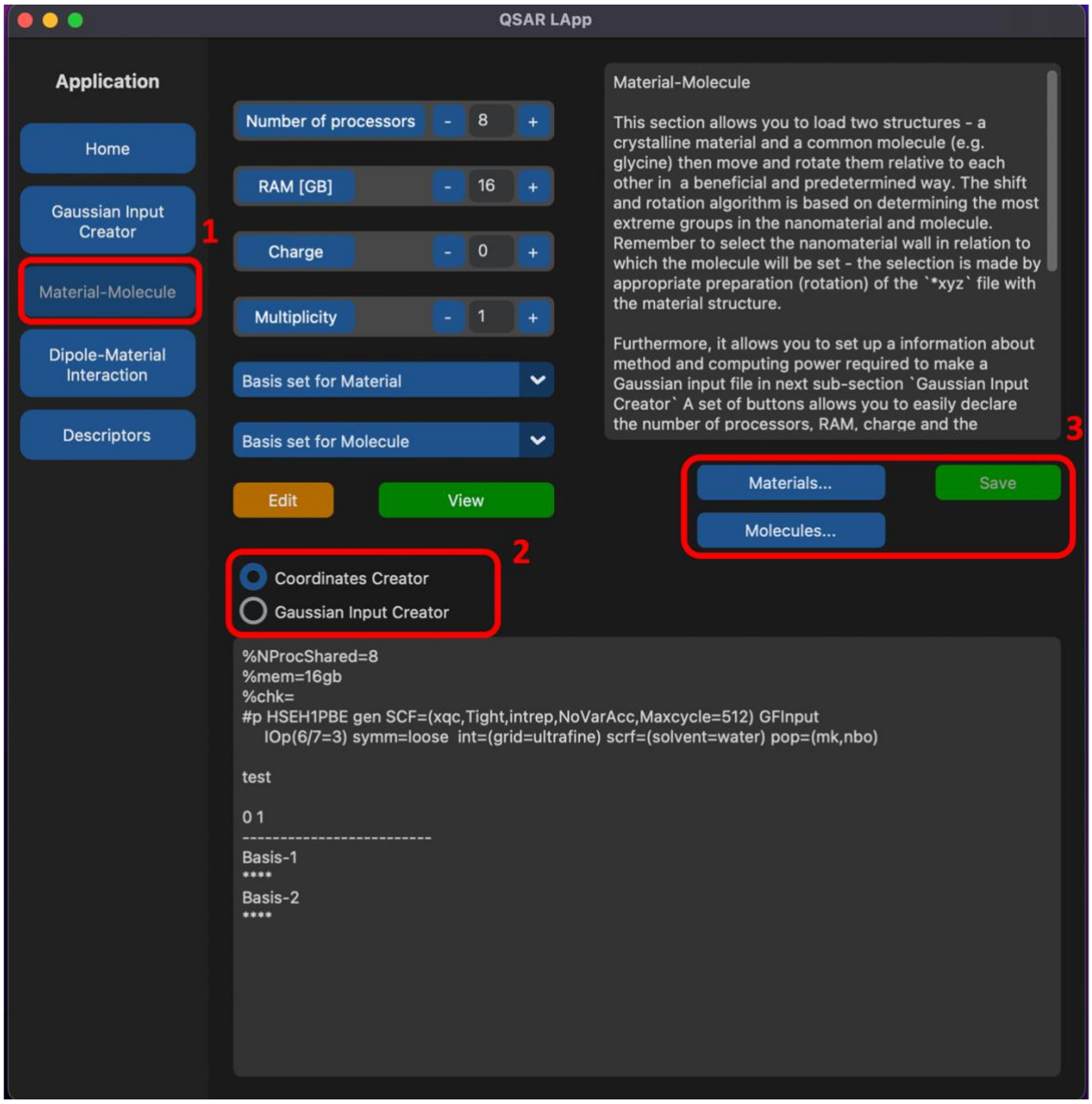

*Figure 4 Material-Molecule section [1], selection panel [2] and a panel for loading the structures of the material and a conventional molecule with a button for saving new (adjusted) *.xyz coordinates [3].*

This section allows you to load two structures - a crystalline material and a common molecule (e.g., glycine) then move and rotate them relative to each other in a beneficial and predetermined way. The shift and rotation algorithm is based on determining the most extreme groups in the nanomaterial and molecule. Remember to select the nanomaterial wall in relation to which the molecule will be set - the selection is made by appropriate preparation (rotation) of the `*xyz` file with the material structure.

If you want to combine two files with a structure in such a way that as a result you will get rotated and shifted structures with a fixed (already predefined) distance between them – use this tool. First, you need to properly prepare the file with the structure of the material, so that the movement of the molecule is performed within a specific wall (see Figure 5 below). In case you want to produce a full set consisting of different materials and different molecules, you only need to select a set of materials and molecules ('Materials' and 'Molecules' button in panel number 3), then press the 'Save' button and choose a location. All material-molecule

Figure 5 The process of creating a material-molecule system. 1 - Selection of the wall by appropriate positioning (rotation) of the material, 2 - Automatic placement of the molecule over the material (phenylalanine in this case).

combinations will be created and saved in the appropriate directories as files with the *.xyz extension. In addition to the *.xyz files, each directory contains the atom_info and method_info files, which contain information on where structure 1 and structure 2 start and end, as well as information on the selected method with computing power settings.

To create Gaussian 16 input files, go to the "Gaussian Input Creator" section in panel 2, then select the directories where the previously created *xyz, atom_info, method_info files are located. If necessary, you can lock the atoms before saving using the ‘Froze’ checkbox.

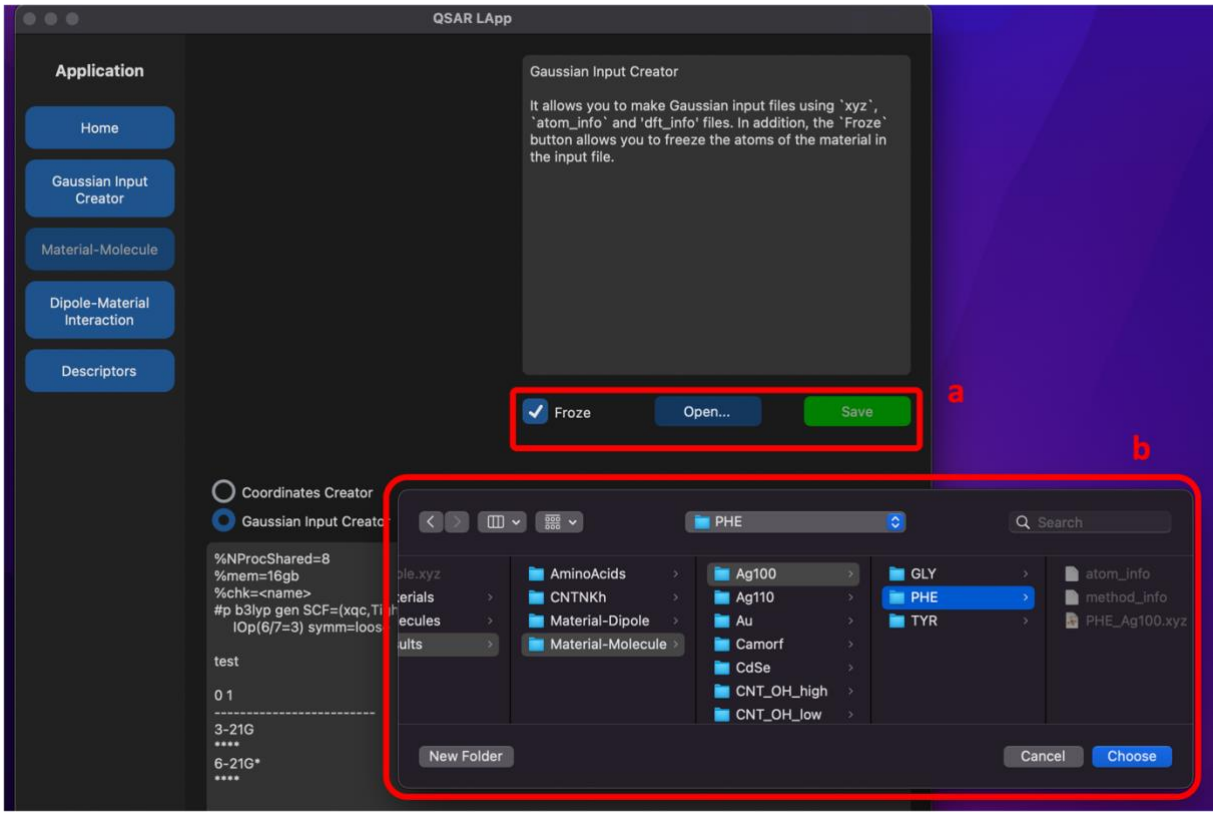

Figure 6 The Gaussian Input Creator. In panel 'a' by pressing 'Open...' you can open selection window 'b'. In order to properly create input files, select the parent directory marked on the screen as 'Material-Molecule'.

In order to easily trace the creation of such systems, sample files with structures (*.xyz files for material and molecule) have been attached.

3. Dipole-Material Section

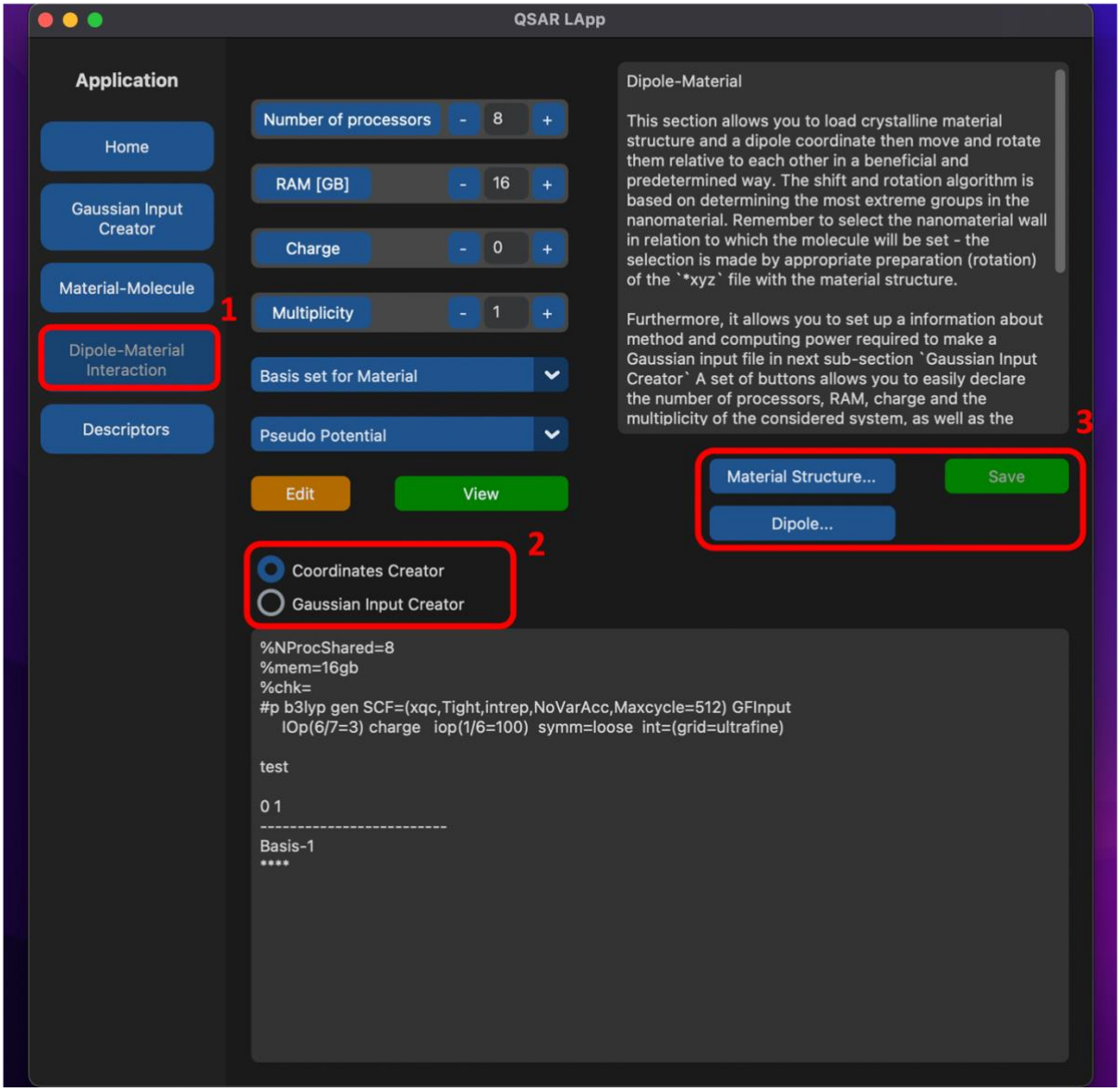

*Figure 7 Dipole-Molecule section [1], selection panel [2] and a panel for loading the structures of the material and a finite dipole coordinates with a button for saving new (adjusted) *.xyz coordinates [3].*

This section allows you to load crystalline material structure and a dipole coordinate then move and rotate them relative to each other in a beneficial and predetermined way. The shift and rotation algorithm is based on determining the most extreme groups in the nanomaterial. Remember to select the nanomaterial wall in relation to which the molecule will be set - the selection is made by appropriate preparation (rotation) of the `*xyz` file with the material structure.

If you want to combine material structure (e.g.: nanostructure) with finite dipole structure (4 orientations of dipole) in such a way that as a result you will get rotated and shifted structures with a fixed distance (5 Å) between them – use this tool. First, you need to properly prepare the file with the structure of the material, so that the movement of the molecule is performed within a specific wall (see Figure 8 below). In case you want to produce a full set consisting of different materials and finite dipole in four orientations, you only need to select

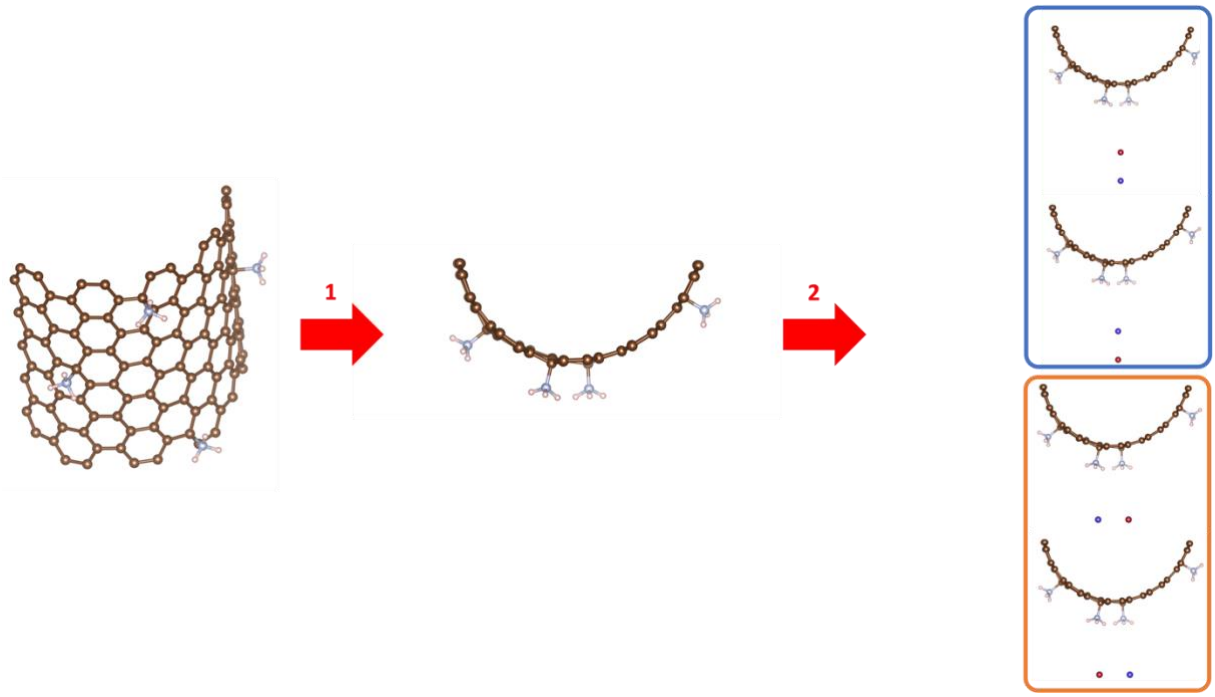

Figure 8 The process of creating a material-dipole systems. 1 - Selection of the wall by appropriate positioning (rotation) of the material, 2 - Automatic placement of the dipole over the material (four different orientations). Vertical orientation is marked in blue, orange is a horizontal orientation.

a set of materials and dipole coordinates ('Materials Structure…' and 'Dipole…' button in panel number 3), then press the 'Save' button and choose a location.

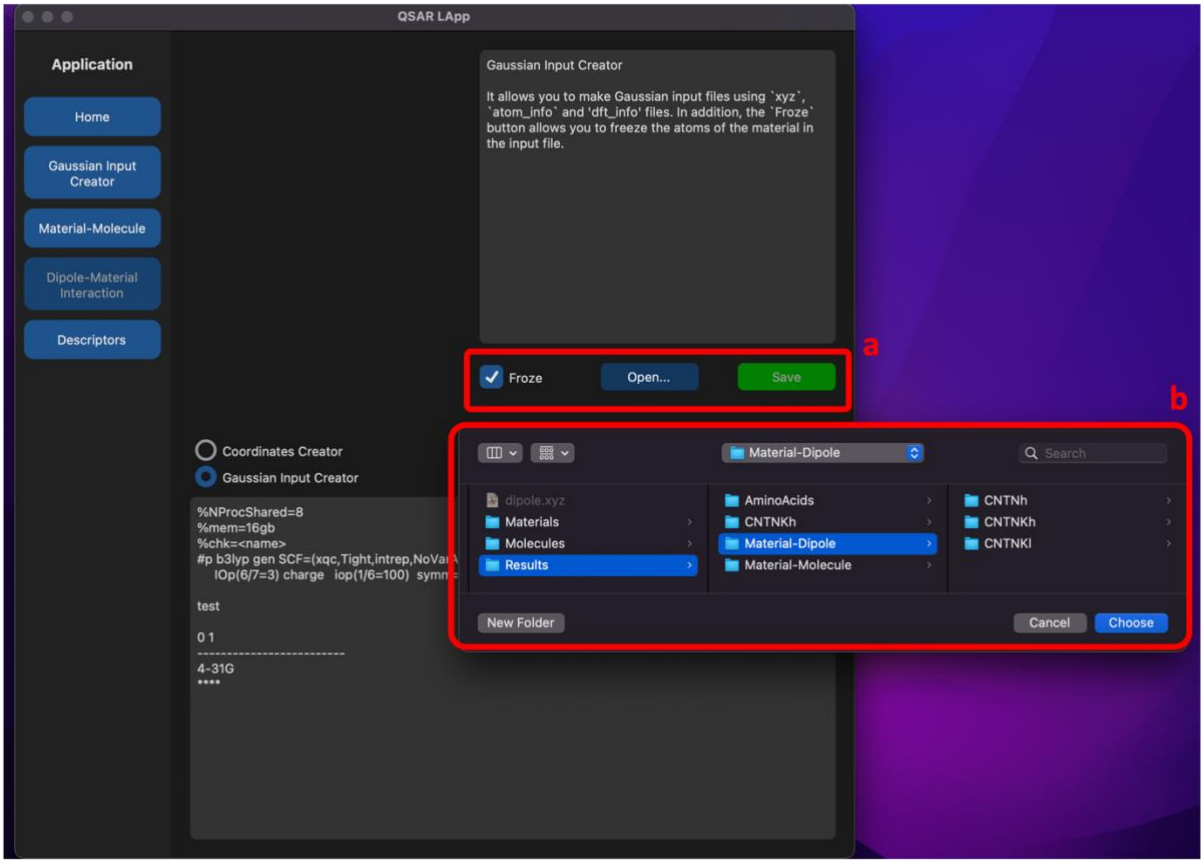

Figure 9 The Gaussian Input Creator. In panel 'a' by pressing 'Open...' you can open selection window 'b'. In order to properly create input files, select the parent directory marked on the screen as 'Material-Dipole'.

In addition to the *.xyz files, each directory contains the atom_info and method_info files, which contain information on where structure 1 and dipole start and end, as well as information on the selected method with computing power settings. All material-dipole combinations will be created and saved in the appropriate directories as files with the *.xyz extension.

To create Gaussian 16 input files, go to the "Gaussian Input Creator" section in panel 2, then select the directories where the previously created *xyz, atom_info, method_info files are located. If necessary, you can lock the atoms before saving using the ‘Froze’ checkbox. In order to easily trace the creation of such systems, sample files with structures (*.xyz files for material and dipole) have been attached.

4. Descriptors section

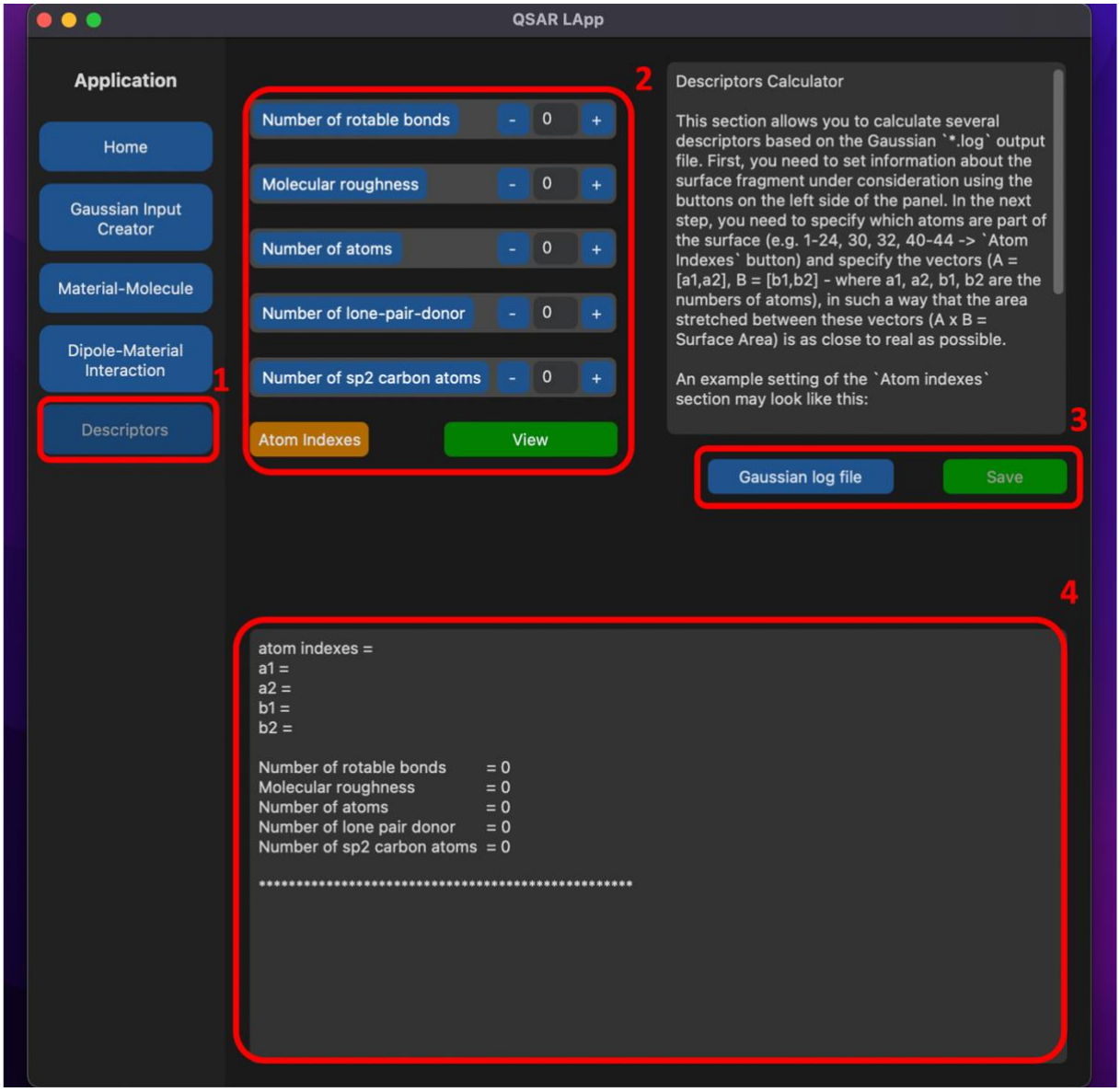

Figure 10 11 Descriptors section [1], settings buttons [2] a panel for loading a file upload and save panel [3] and a console with a preview of settings [4].

This section allows you to calculate several descriptors based on the Gaussian `*.log` output file. First, you need to set information about the surface fragment under consideration using the buttons on the left side of the panel 2. In the next step, you need to specify which atoms are part of the surface e.g.: 1-24, 30, 32, 40-44, by clicking on `Atom Indexes` button and specify the vectors (A = [a1,a2], B = [b1,b2] - where a1, a2, b1, b2 are the numbers of atoms), in such a way that the area stretched between these vectors (A x B = Surface Area) is as close to real as possible (see figure 12).

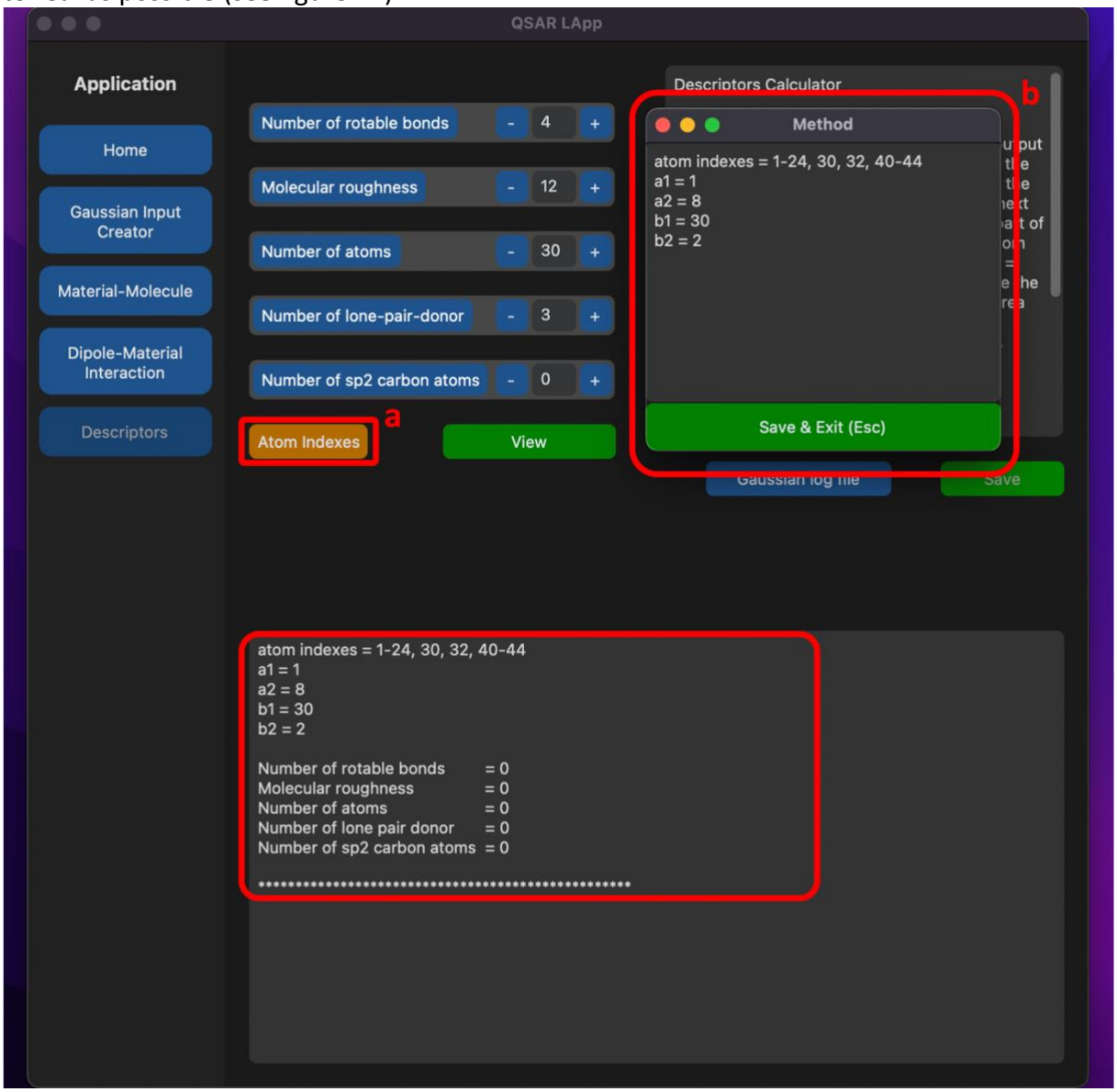

Figure 12 Determining the numbers of atoms that are part of the surface area and determining which atoms span the surface of interest to us. 'a' - opens the edit window 'b', where you need to enter information about which atoms make up the surface area.

If we are dealing with a cylindrical shape of the material (see Figure 13), we define the surface of interest as the largest flat fragment (we assume that such a fragment at the molecular level is involved in the adsorption process - more information in paper).

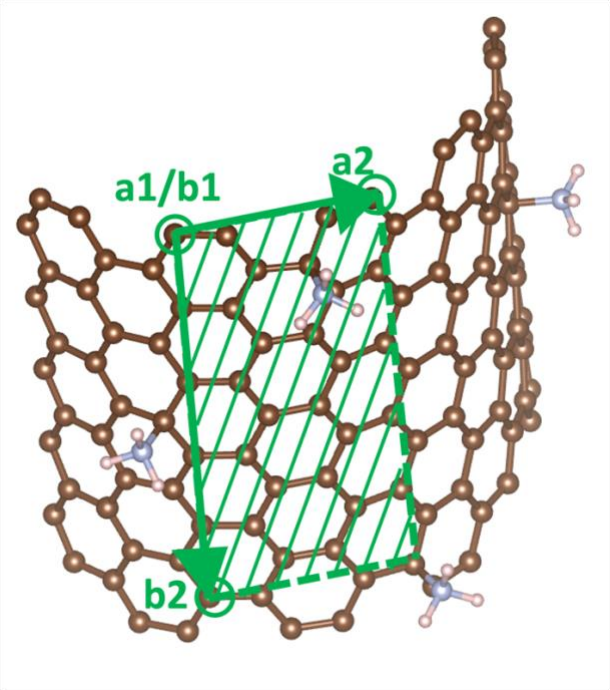

Figure 13 Determination of the fragment involved in the adsorption process (more precisely: the fragment interacting with finite dipoles). In the figure, the atoms a1, a2, b1, b2 that span the surface area (hatched area) are marked. Atoms entering the selected area are important for us and their 'atom indexes' should be given in 'method' shown in Figure 12 'b'.

The result `qsarlab.descr` file contains information about the descriptors values calculated per surface area. This file contains information about such descriptors as:

Number of rotable bonds	Number of single bonds (i.e., not involved in a ring structure) on the NM's surface per Å^2.
Molecular roughness	Number of functional groups sticking out of the NM's surface per Å^2.
Number of atoms	Number of atoms on the NM's surface per Å^2.
Number of electron-lone-pair-donor atoms	Number of atoms capable of donating their electron lone pair on the NM's surface per Å^2.
Number of sp^2 carbon atoms	Number of sp^2-hybridized carbon atoms on the NM's surface per Å^2.
Σ Pauling electronegativity	Pauling electronegativity values summed over the atoms on the NM's surface per Å^2.
Σ q_NBO	Natural Bond Orbital partial atomic charges (in a.u.) summed over the atoms on the NM's surface per Å^2.
Σ q_Mull	Mulliken partial atomic charges (in a.u.) summed over the atoms on the NM's surface per Å^2
Σ q_ESP	Merz-Kollman (i.e., fitted to the electrostatic potential) partial atomic charges (in a.u.) summed over the atoms on the NM's surface per Å^2.