**Submission ID #: 67457**

**Scriptwriter Name: Debopriya Sadhukhan**

**Project Page Link:** [**https://review.jove.com/account/file-uploader?src=20574433**](https://review.jove.com/account/file-uploader?src=20574433)

**Title: Quantitative Structure—Activity Relationship, Activity Prediction, and Molecular Dynamics of Non-nucleotide Reverse Transcriptase Inhibitors**

**Authors and Affiliations:**

Vitalis Mbayo, Penny P. Govender, Ephraim F. Marondedze, Krishna K. Govender

Department of Chemical Sciences, Faculty of Science, University of Johannesburg

All author names and affiliations are correct (city/state/country information not included in video title page).

**Corresponding Authors:**

Krishna K. Govender krishnag@uj.ac.za

**Email Addresses for All Authors:**

Vitalis Mbayo vmbayo92@gmail.com

Penny P. Govender pennyg@uj.ac.za

Ephraim F. Marondedze [ephraimm18@gmail.com](mailto:ephraimm18@gmail.com)

Krishna K. Govender krishnag@uj.ac.za

## **Author Questionnaire**

**1. Microscopy**: Does your protocol require the use of a dissecting or stereomicroscope for performing a complex dissection, microinjection technique, or something similar? **No**

If **Yes**, can you record movies/images using your own microscope camera?

**No**

Ifyour protocol involves microscopy but you are not able to record movies/images with your microscope camera, JoVE will need to use our scope kit.

If your microscope does not have a camera port, the scope kit will be attached to one of the eyepieces and **you will have to perform the procedure using one eye**.

**Not applicable**

If a dissection or stereo microscope is required for your protocol, please list all shots from the script that will be visualized using the microscope (shots are indicated with the 3-digit numbers, like 2.1.1, 2.1.2, etc.).

**Click here to list microscope shots, using the shot numbers from the protocol section of the video script.**

**2. Software:** Does the part of your protocol being filmed include step-by-step descriptions of software usage? **Yes (this will be done by the student that did the work and he will make use of OBS to record everything)**

If **Yes**, we will need you to record using screen recording software.

We recommend using the screen capture program [OBS](https://obsproject.com/). JoVE’s tutorial for using OBS Studio is provided at this link: <https://review.jove.com/v/5848/screen-capture-instructions-for-authors?status=a7854k>

As these files are necessary for finalizing your script, please upload all screen-captured video files to your project page as soon as possible.

**3. Filming location:** Will the filming need to take place in multiple locations?  **No**

If **Yes**, how far apart are the locations? **Click to enter distance between locations.**

To ensure that your **script can be filmed in one day**, the protocol sections are cumulatively restricted to**55 shots** (shots are the 3-digit numbers like 2.1.1, 2.1.2…etc)

**Current Protocol Length**

Number of Steps: 23

Number of Shots: 48

# Introduction

***Videographer: Obtain headshots for all authors available at the filming location.***

Answers to these questions will become interview statements that you will deliver on camera.

* Answer the **1st REQUIRED** question and **at least 2 other questions (1.2 – 1.10)** below. Up to 5 interview statements will be included in the video.
* Enter the **full name** of the author who will deliver the statement.
* If possible, each author should deliver **no more than two statements**.
* Answer in full sentences, in a style suitable for being spoken aloud.
* Limit the length of each statement to **30 words or fewer**.
* Answers will be edited for length, clarity, and consistency with journal style guidelines.

**REQUIRED:** What is the scope of your research? What questions are you trying to answer?

* 1. Krishna K. Govender: Our research focuses on computer aided drug design where we look to either repurpose existing drugs for different diseases. In addition, we look to use computational methods to design new innovative drugs that can be used for treatment of various diseases that are prominent in Africa.

What are the most recent developments in your field of research?

* 1. Krishna K. Govender**:** Numerous tools have been developed in the field to allow for sampling of larger number of potential drug candidates using machine learning algorithms. These are tools that we wish to apply within our research group in the future.

What technologies are currently used to advance research in your field?

* 1. Krishna K. Govender**:** Within the field of computational chemistry, it is imperative to make use of high-performance computing technologies due to the size of the problems being investigated. These technologies are constantly advancing in terms of central processing units (CPUs) and graphical processing units (GPUs), allowing for fast simulation times and thereby giving rise to possible solutions within timeframes that are much shorter than expected from conventional experimental methods.

What are the current experimental challenges?

* 1. Krishna K. Govender**:** The field of computational chemistry is still relatively new within South Africa and students often lack the necessary background to ensure that the computational simulations are carried out in a manner that can be compared back to experiment. As a result, it does often take a lot longer than expected to identify potential drug hits when running the simulations.

What significant findings have you established in your field?

* 1. Krishna K. Govender**:** With some of our more recent work we have managed to isolate sets of natural products that are indigenous to Africa which show potential for treatment of SAR-CoV-2.

What research gap are you addressing with your protocol?

* 1. Krishna K. Govender**:** The protocol allows for bridging the gap between experiment and theory and it is hoped that this will allow more experimental chemists to add the computational aspects to their analysis process to make things more efficient.

What advantage does your protocol offer compared to other techniques?

* 1. Krishna K. Govender**:** It provides for direct comparison to experimentally available data, while also allowing for prediction of new uninvestigated compounds.

How will your findings advance research in your field?

* 1. Krishna K. Govender**:** It will allow us to come up with new drug candidates that can be taken to the experimentalist for synthesis, characterization and testing.

What new scientific questions have your results paved the way for?

* 1. Krishna K. Govender**:** Design of drugs is a complicated process, using the methods in this work we can make changes to existing scaffolds which give rise to better drug candidates.

What research questions will your laboratory focus on in the future?

* 1. Krishna K. Govender**:** We would like to create and sample large (millions or billions) ligand libraries to ensure that most of the drug phase space is utilized when discovering new drugs.

***Videographer: Obtain headshots for all authors available at the filming location.***

**Testimonial Questions (OPTIONAL):**

Answers to these questions **will not appear in the video** but may be featured in our journal's promotional materials.

* Enter the **full name** of the author who will deliver the statement.
* Answer in full sentences, in a style suitable for being spoken aloud.
* Limit the length of each statement to **50 words or fewer**.

What motivated you to choose JoVE for publishing your research?

* 1. Krishna K. Govender**:** I was approached by Tasaduq Wani and after a discussion with him I decided to submit a manuscript to the journal.

How does the research community benefit from video publications as compared to standard text publications?

* 1. Krishna K. Govender**:** It helps the community to visually see what is being done in the field and specifically for computational chemistry this is crucial as we work primarily with 3D representations of systems.

# Protocol

**Please review this section to make sure that it accurately describes your protocol. Use Track Changes when making edits or revisions.**

* The two-digit **steps** (e.g., 2.1., 2.2.) are the narration. **JoVE is responsible for the narration of the protocol and results.**
* *Red italics* are pronunciation guides indicating how the word will be spoken.
* Filming should take no more than 10 minutes per step. If a step takes more than 10 minutes, prepare the product for that step in advance.
* The three-digit **shots** (e.g., 2.1.1., 2.2.2.) are the actions that the videographer will capture.

1. **Retrieving and Importing Protein Structures**

**Demonstrator:** Vitalis Mbayo

If the same person is the demonstrator throughout, mention them once here and remove the "Demonstrator" field from the other sections; if the demonstrator changes, retain the field in the respective sections.

* 1. To begin, go to the **Window** icon on the monitor screen and click on it **[1]**. Select **All Apps** and scroll down to locate the **Schrodinger** folder. Open the folder and click on the **Maestro** icon. Select **Open** to launch the software **[2]**.
     1. WIDE: Talent sitting in front of a computer screen, moving the cursor to the **Window** icon and clicking on it. *Videographer: Make sure the computer screen is clearly visible in the frame.*
     2. SCREEN: To be provided by authors: **All Apps** menu being selected, the list being scrolled to locate the **Schrodinger** folder, the folder being opened, the **Maestro** icon followed by **Open** being clicked to launch the software.

Authors: For the shots labelled as SCREEN, please create the screencapture videos and a summary following our guidelines and upload them to your project page: <https://review.jove.com/account/file-uploader?src=20574433>

* 1. To retrieve the protein structure, click on the **File** tab and select **Get PDB** from the pop-up menu. Enter the PDB code of choice in the text box and click the **Download** button **[1]**. The selected PDB file will appear in the project window **[2].**
     1. SCREEN: To be provided by authors: The **File** tab being clicked, the **Get PDB** option being selected from the pop-up menu, the PDB code being entered in the text box, and the **Download** button being clicked.
     2. SCREEN: To be provided by authors: The downloaded PDB file appearing in the project window.
  2. Alternatively, download the protein from the **Protein Data Bank** by entering the PDB ID in the search box and clicking **Download** **[1]**. In **Maestro**, navigate to the **File** tab and select **Import Structures** **[2]**. In the Import interface, locate the downloaded PDB file and click **Import** **[3]**.
     1. SCREEN: To be provided by authors: **Protein Data Bank** website being opened. The PDB ID being entered in the **Protein Data Bank** search box, and the **Download** button being clicked.
     2. SCREEN: To be provided by authors: The **File** tab being clicked in **Maestro**, and **Import Structures** being selected.
     3. SCREEN: To be provided by authors: The downloaded PDB file being located in the **Import** interface and the **Import** button being clicked
  3. Now, select the protein structure and right-click on it **[1]**. Select the prepared protein, right-click the mouse button, choose the Split option and split into ligands, water, and other **[2]**.
     1. SCREEN: To be provided by authors: The protein structure being selected and right-clicked.
     2. SCREEN: To be provided by authors: The **Split Ligand** option being selected, followed by choosing the option to split into ligands, water, and other components.

1. **Ligand Preparation**

**Demonstrator:** Vitalis Mbayo

* 1. Open the **PubChem** database and type the compound name in the search bar to download the chemical compound **[1]**. Review the available structures and select **3D structures** **[2]**. Click **Download** to save the structure coordinates as a **Structured Data File** or SDF *(S-D-F)* **[3-TXT]**.
     1. SCREEN: To be provided by authors: The **PubChem** database being opened. The compound name being entered in the **PubChem** search bar.
     2. SCREEN: To be provided by authors: A list of available structures being reviewed, with **3D structures** being selected.
     3. SCREEN: To be provided by authors: The **Download** button being clicked and the structure being saved in **SDF** format. **TXT: If no 3D structure is available, download 2D and convert it**
  2. Click on the **File** tab in **Schrodinger** and select **Import Structures**. Navigate to the location where the **SDF** is saved and load the compound **[1]**.
     1. SCREEN: To be provided by authors: The **File** tab being clicked in **Schrodinger**, and **Import Structures** being selected. The SDF file location being accessed, and the **SDF** file being loaded.
  3. Click on **Task** in **Schrodinger**, type **LigPrep** in the search bar, and select it **[1]**.
     1. SCREEN: To be provided by authors: The **Task** button being clicked in the **Schrodinger** software. **LigPrep** being typed into the search bar. **LigPrep** being selected from the right window.
  4. Click **Use Structures From** to choose files from the Workspace or Project Table **[1]**. Select the preferred options in the **LigPrep** window and save the file to the local computer **[2]**. Click **Run** to submit the job for ligand preparation **[3]**.
     1. SCREEN: To be provided by authors: The **Use Structures From** option being selected and then the files being chosen from the Workspace or Project Table.
     2. SCREEN: To be provided by authors: The preferred option being selected from the **LigPrep** window and then the file being saved.
     3. SCREEN: To be provided by authors: **Run** button being clicked to start ligand preparation.

1. **Geometry** **and Optimization of Ligands**
   1. Open the software for geometry optimization of the structures **[1]**. Navigate to the **File** tab and select **Open** to choose the downloaded **SDF** from **PubChem** **[2]**.
      1. SCREEN: To be provided by authors: The software being opened for geometry optimization.
      2. SCREEN: To be provided by authors: The **File** tab being clicked, and the **Open** option being selected. The **SDF** file being loaded.
   2. Navigate to the **Calculate** tab and select **Gaussian Calculation Setup**. In the **Job Type** tab, choose **Optimization** or **Opt+Freq** *(optimization plus frequency)* **[1]**. Authors: How do you want to pronounce Opt+Freq? Optimization plus frequency?
      1. SCREEN: To be provided by authors: The **Calculate** tab being clicked, and **Gaussian Calculation Setup** being selected. The **Job Type** tab being accessed, and **Optimization** or **Opt+Freq** being selected.
   3. Now, navigate to the **Method** tab and select the quantum chemistry method. Choose the **Kohn–Sham global-hybrid exchange-correlation density functional**, **basis set, charge**, and **spin of choice** from the dropdown menus **[1]**.
      1. SCREEN: To be provided by authors: The **Method** tab being clicked, and a quantum chemistry method being selected. The **Kohn–Sham** **global-hybrid exchange-correlation density functional**, basis set, charge, and spin being chosen from dropdown menus.
   4. Go to the **Title** tab and enter a name for the compound under investigation **[1].** Navigate to the **Link 0** *(zero)* tab and specify the **Memory Limit** and **Shared Processors**. Untick the **Full Path** boxes **[2]**.
      1. SCREEN: To be provided by authors: The **Title** tab being accessed, and a name being entered for the compound.
      2. SCREEN: To be provided by authors:
   5. Click the **Edit** button at the bottom to save the Gaussian input file **[1]**. Save the file in the preferred location with a file name of choice as a Gaussian job file **[2]**.
      1. SCREEN: To be provided by authors: The **Edit** button being clicked to save the **Gaussian input file.**
      2. SCREEN: To be provided by authors: The file being named and saved as a **Gaussian job file (GJF)** in the preferred location.
2. **Receptor Grid Generation and Molecular Docking**
   1. Navigate to **Tasks** and select **Receptor Grid Generation** to detect the protein's active site bound to the core-crystal ligand **[1]**. Click **Pick to identify the ligand** and check for the presence of a co-crystallized ligand in the top pop-up notification **[2]**.
      1. SCREEN: To be provided by authors: The **Tasks** menu being accessed and **Receptor Grid Generation** being selected.
      2. SCREEN: To be provided by authors: The **Pick** **to identify the ligand** is checked , and the notification being checked for a co-crystallized ligand.
   2. For molecular docking, go to **Tasks**, select **Docking**, and then choose **Ligand Docking (Glide Docking)** toload the protein and the prepared ligands **[1]**. Then, load the grid file **[2]** and select ligands from the workspace using the **Use Ligand From** option **[3]**. Authors: How do you want to pronounce **Ligand Docking (Glide Docking)?**
      1. SCREEN: To be provided by authors: The **Tasks** menu being accessed, and **Docking > Ligand Docking (Glide Docking)** being selected.
      2. SCREEN: To be provided by authors: The **Grid File** being loaded.
      3. SCREEN: To be provided by authors: Ligands being selected from the workspace using **Use Ligand From** option.
   3. Choose the preferred docking precision method from the **Settings** tab **[1-TXT]**. Set the **Force Field** to **OPLS4** *(O-P-L-S-four)* for accurate modeling of molecular interaction **[2]**. In the **Constraints** tab, configure constraints such as hydrogen bonds **[3]**.
      1. SCREEN: To be provided by authors: The **Settings** tab being accessed, and the docking precision method being selected. **TXT: The default docking precision is SP (standard precision)**
      2. SCREEN: To be provided by authors: The **Force Field** being set to **OPLS4**.
      3. SCREEN: To be provided by authors: The **Constraints** tab being accessed, and constraints such as hydrogen bonds being configured.
   4. After reviewing all settings, save the docking job or file **[1]**. Click **RUN** to start the docking process **[2]**.
      1. SCREEN: To be provided by authors: Docking settings being reviewed and saved.
      2. SCREEN: To be provided by authors: The **RUN** button being clicked to initiate docking.
3. **Enumeration, Molecular Dynamics and Molecular Mechanics with Generalized Born and Surface Area (MM-GBSA)**
   1. Select a pair of the docked protein and the ligand complex from the **Workspace Navigator** **[1]**. Click **Analyze Workspace** in the **Ligand Designer** window **[2]**.
      1. SCREEN: To be provided by authors: A docked protein-ligand complex being selected in the **Workspace Navigator**.
      2. SCREEN: To be provided by authors: The **Analyze Workspace** button being clicked in the **Ligand Designer** window.
   2. To generate and evaluate new ligands, select **Isostere Scanning** from the workflow list **[1]**, which implies the growing method that extends the ligand by adding fragments to existing molecular structures **[2]**.
      1. SCREEN: To be provided by authors: **Isostere Scanning** being selected from the workflow list.
      2. SCREEN: To be provided by authors: Ligand structure being extended by adding molecular fragments.
   3. Click on the **Task** button and select **Desmond System Builder** **[1].**
      1. SCREEN: To be provided by authors: The **Task** button being clicked and **Desmond System Builder** being selected.
   4. Open the trajectory file and play the trajectory **[1]**. Visualize where the protein-ligand complex is equilibrated and note the number of frames **[2]**. Submit the job via the terminal **[3-TXT]**.
      1. SCREEN: To be provided by authors: The trajectory file being opened and played.
      2. SCREEN: To be provided by authors: The protein-ligand complex being visualized with the equilibrated position and frame count displayed.
      3. SCREEN: To be provided by authors: The job being submitted via the terminal. **TXT: Perform MMGB-SA simulation for protein-ligand complex**
   5. View the output file contents to analyze the generated results **[1]**. Read the **ΔG** *(delta-G)* Average value to determine the binding free energy of the protein-ligand complex **[2]**. Then, download the CSV*(C-S-V)*file **[1]**.
      1. SCREEN: To be provided by authors: The **output file** being opened and displayed.
      2. SCREEN: To be provided by authors: The **ΔG Average** value being displayed.
      3. SCREEN: To be provided by authors: The **CSV file** being downloaded.
   6. Open the CSV file and take note of the binding energy **[1-TXT]**. Finally, use the shown equation and calculate the free binding energy of the complex by averaging the binding energy values determined for each snapshot within the MD simulation **[2]**.
      1. SCREEN: To be provided by authors: The CSV file being opened and the value of the binding energy being shown. **TXT: See text for other thermodynamics and desolvation parameters**
      2. TEXT on PLAIN BACKGROUND:

# Results

**Please review this section to make sure that it accurately reflects your findings.**

* This section **will not be recorded** by the videographer. It only includes the figures/tables from your manuscript (called LAB MEDIA).
* Use Track Changes when making edits or revisions. Ensure the voiceover length is below 200 words. Current word count: 157.
* Please note that the video **cannot** include voiceover without an accompanying visual.

1. **Representative Results** 
   1. A scatter plot shows the observed activity versus the predicted activity for class 1 of the QSAR *(Q-S-A-R)* model **[1]**.The graph represents the fitting between class 1 as the training set **[2]** and the non-nucleotide reverse transcriptase inhibitors as the test set **[3]** to give a predictive activity value **[4]**.
      1. LAB MEDIA: Figure 15. *Video Editor: Highlight Activity (observed) in the x-axis when the VO says “the observed activity” and highlight Activity (predicted) in the y-axis when the VO says “the predicted activity”.* **TXT: QSAR: Quantitative Structure-Activity Relationship**
      2. LAB MEDIA: Figure 15. *Video Editor: Highlight the blue points and the legend training set.*
      3. LAB MEDIA: Figure 15. *Video Editor: Highlight the red points and the legend test set.*
      4. LAB MEDIA: Figure 15.
   2. The training set aligned well with the regression line **[1]**, while the test set had minor deviations **[2].**
      1. LAB MEDIA: Figure 15. *Video Editor: Highlight the blue points and the red straight line.*
      2. LAB MEDIA: Figure 15. *Video Editor: Highlight the red points and the red straight line.*
   3. The forces of interactions between the protein and different ligands are shown here **[1]**. These diagrams revealed hydrogen bonds with LYS101 in all ligands **[2]**. Authors: How do you want to pronounce LYS101? Lysine 101
      1. LAB MEDIA: Figure 16 (A, B, D, E).
      2. LAB MEDIA: Figure 16 (A, B, D, E). *Video Editor: Highlight LYS101 and the purple arrows.*
   4. Molecular dynamics simulations of the free protein stabilized after around 60 nanoseconds at an RMSD *(R-M-S-D)* of around 3.5 angstroms, confirming protocol reliability **[1-TXT]**.
      1. LAB MEDIA: Figure 17. *Video Editor: Highlight the red and blue lines in the graph for the region ~60 (in x-axis) and above when the VO says, “after around 60 nanoseconds at an RMSD of around 3.5 angstroms”*. **TXT: RMSD: Root Mean Square Deviation**
   5. Enumerated Etravirine, which stabilized at 3.5 angstroms, exhibited stronger and more stable binding at the active site of HIV-1 reverse transcriptase **[1]** compared to Etravirine, which stabilized at 4.5 angstroms **[2].**
      1. LAB MEDIA: Figure 17B, 17C. *Video Editor: Zoom in 17C*
      2. LAB MEDIA: Figure 17B, 17C. *Video Editor: Zoom in 17B*.
   6. The contact timeline of enumerated Etravirine also indicated stronger and more stable interactions over time **[1]**.
      1. LAB MEDIA: Figure 18B, Figure 19B. *Video Editor: Highlight the darkest orange line in 19B around A-LYS-101.*