VIRTUAL CHEMIST Suite 2019

FINDERS, REACT2D, CONSTRUCTS AND ACE

Tutorial

Department of Chemistry
McGill University
Montréal, Québec, Canada

Molecular Forecaster Inc.
Montréal, Québec, Canada

29/03/2019

Table of Contents

I. What will be Covered in this Tutorial

This example demonstrates the full design and evaluation of asymmetric organocatalysts using our VIRTUAL CHEMIST platform. You will search a catalog for reagents (using FINDERS) to be used in an esterification reaction to virtually synthesize proline analogs (using REACT2D). These analogs will then be built onto a three-dimensional transition state (TS) template of the Diels Alder reaction (using CONSTRUCTS). The TSs will be investigated by ACE to predict the enantioselectivity.

Why use this workflow? This workflow is akin to small molecule docking but for small molecule catalyst discovery, rather than drug discovery. Our software will "dock" catalyst candidates to a given reaction transition state (like a protein) and then model different conformations to predict the most likely "stable" conformation leading to each possible product (R/S). The difference in calculated energy is used to predict the ratio of one stereoisomer to the other (enantioselectivity).

How long will the tutorial take? The tutorial itself is run at your own pace. The workflow takes an hour or two to complete, depending on your computing hardware and if you run on multiple compute cores (see below). A standard laptop should handle the experiment in two hours. The majority of this time is for ACE, whereas the other elements of the workflow should complete in a minute or two.

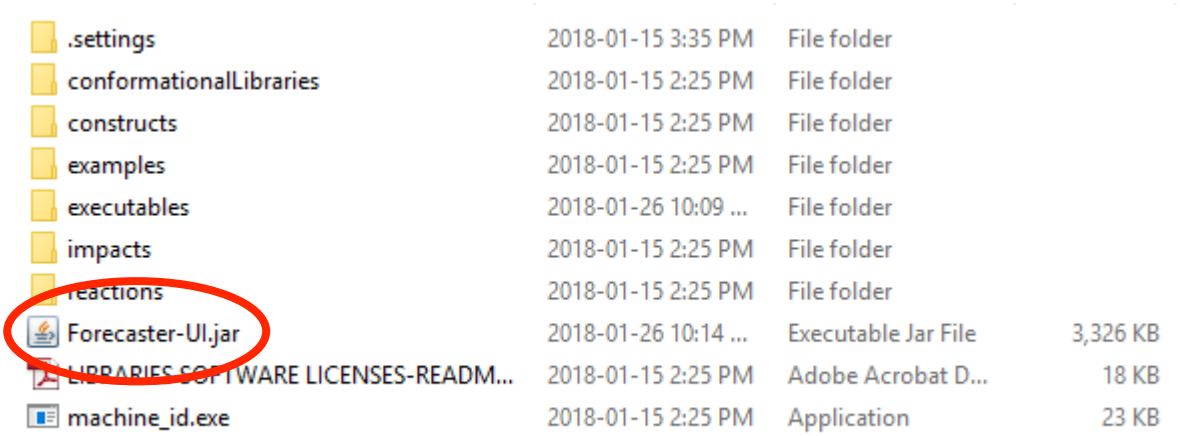
FAQ. ACE uses molecular mechanics (MM) and not quantum mechanics (QM), which allows the computation to complete relatively quickly. We have done extensive validation on numerous chemical reactions to demonstrate that our MM approach is viable regardless of the energy calculation shortcomings. Relatively, we can accurately rank catalysts for their selectivity.

II. Running FORECASTER with the User Interface

The user interface (UI) can be started by double clicking on the `Forecaster-UI.jar` file in the `Forecaster` folder. This will open the main window.

Under Linux and Mac OSX, it is recommended to launch it from a terminal window by typing the command below. Make sure that you are located in the folder where this jar file is.

```
Forecaster@Linux/Forecaster:~$ java -jar Forecaster-UI.jar
```



The first step is to set the working directory. Click **Browse**, under **Working Directory**. You will be prompted to navigate to the folder where you will save various files while working with FORECASTER. For this tutorial, go to “examples” and then to “virtual-chemist-AsymmetricCatalysis”.

All the necessary files will be written there. The molecules can be drawn using the FORECASTER sketcher and the molecule/reaction files will be written there as well. If you are planning to run calculation from an external library (e.g., a catalog of chemicals, a file from the ZINC database,…), a copy should be in your working directory.

III. Asymmetric Catalyst Discovery

When your working directory has been selected, click the **Start Forecasting** button to expand the virtual chemist workflow.

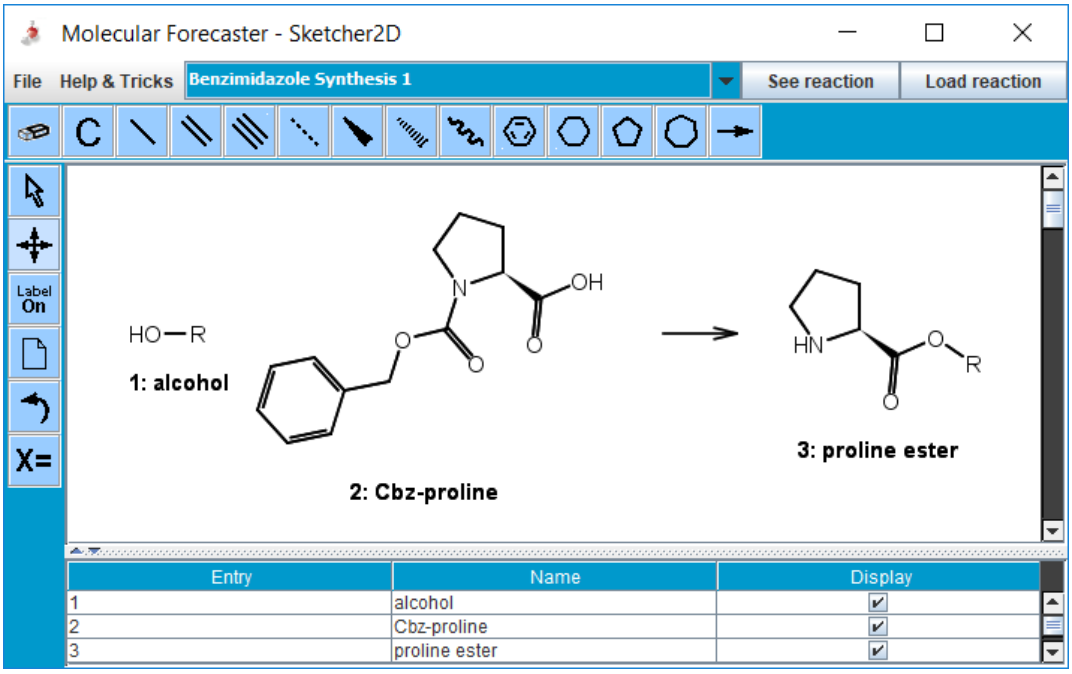
1. SKETCHER2D: Drawing the chemical reaction

Using the sketcher, you will now need to define the reaction. Click on the gear icon to open SKETCHER2D. Dozens of predefined reactions can also be used as described throughout this tutorial.

The drawing the reaction should be done following some rules to ensure the programs will encode them properly.

The esterification reaction shown below is provided in the tutorial folder, but you may draw it again to gain experience. Alternatively, you can use **File**, **Open reaction scheme (rxn)**.

If you choose to draw, the reaction is drawn from left to right with the reactant(s) and product(s) separated by an arrow. The simplest substructure needs to be drawn and points of diversity encoded with the R, G or Ar groups (for definition click on the "Help & tricks" menu). In our example, Cbz-Proline is esterified with various alcohols. For more information about these features, see tutorial on FINDERS and REACT2D.

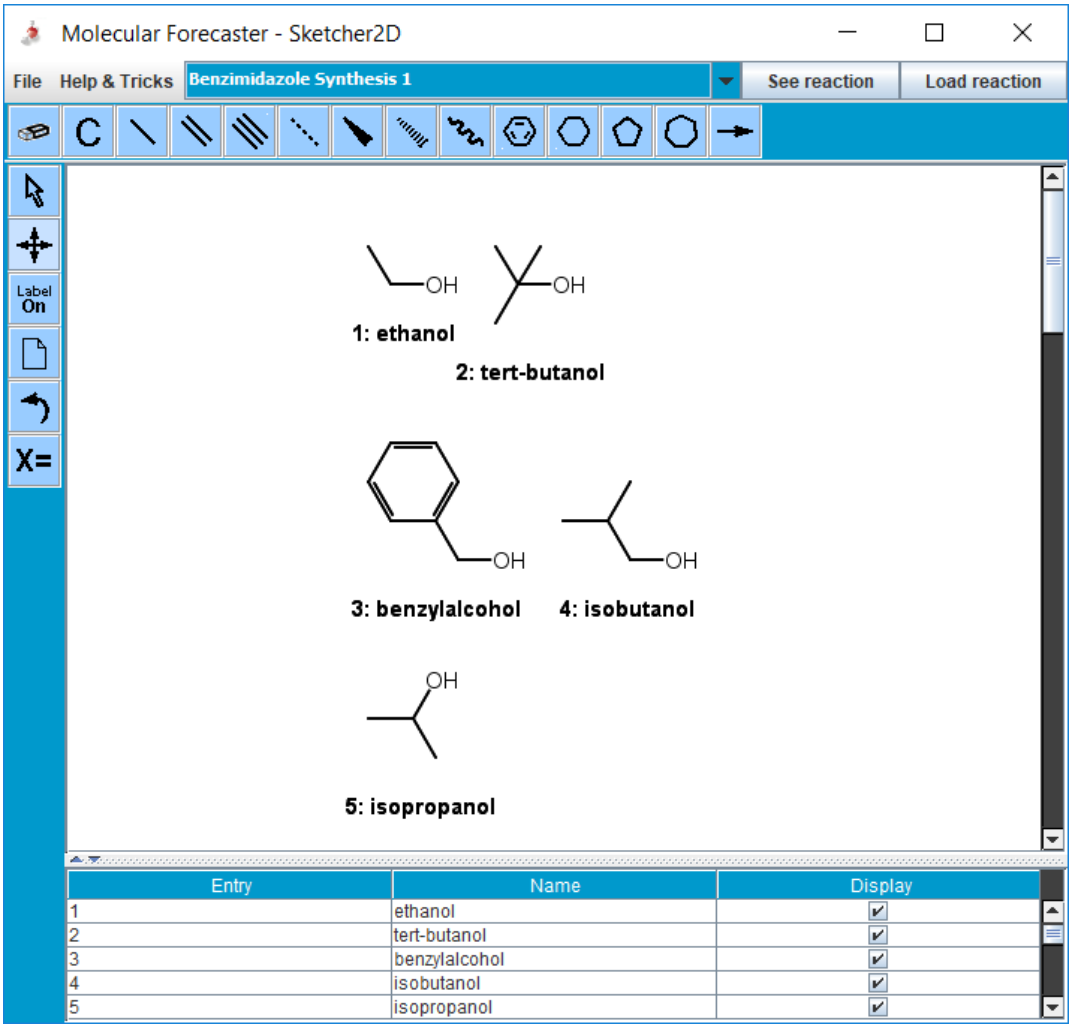


Once the reaction scheme is complete, the reaction can be saved using the **File**, **Save reaction (rxn)** option from the sketcher and providing the `esterification.rxn` filename.

For any other application, you may also load one of the ca. 60 reactions already listed using the drop-down menu at the top of the sketcher and then click **Load reaction**.

A small library of alcohols is provided in the tutorial folder. You can view it by using **File**, **Open structure file (sdf)**. If you'd like more drawing experience, you can complete the next step.

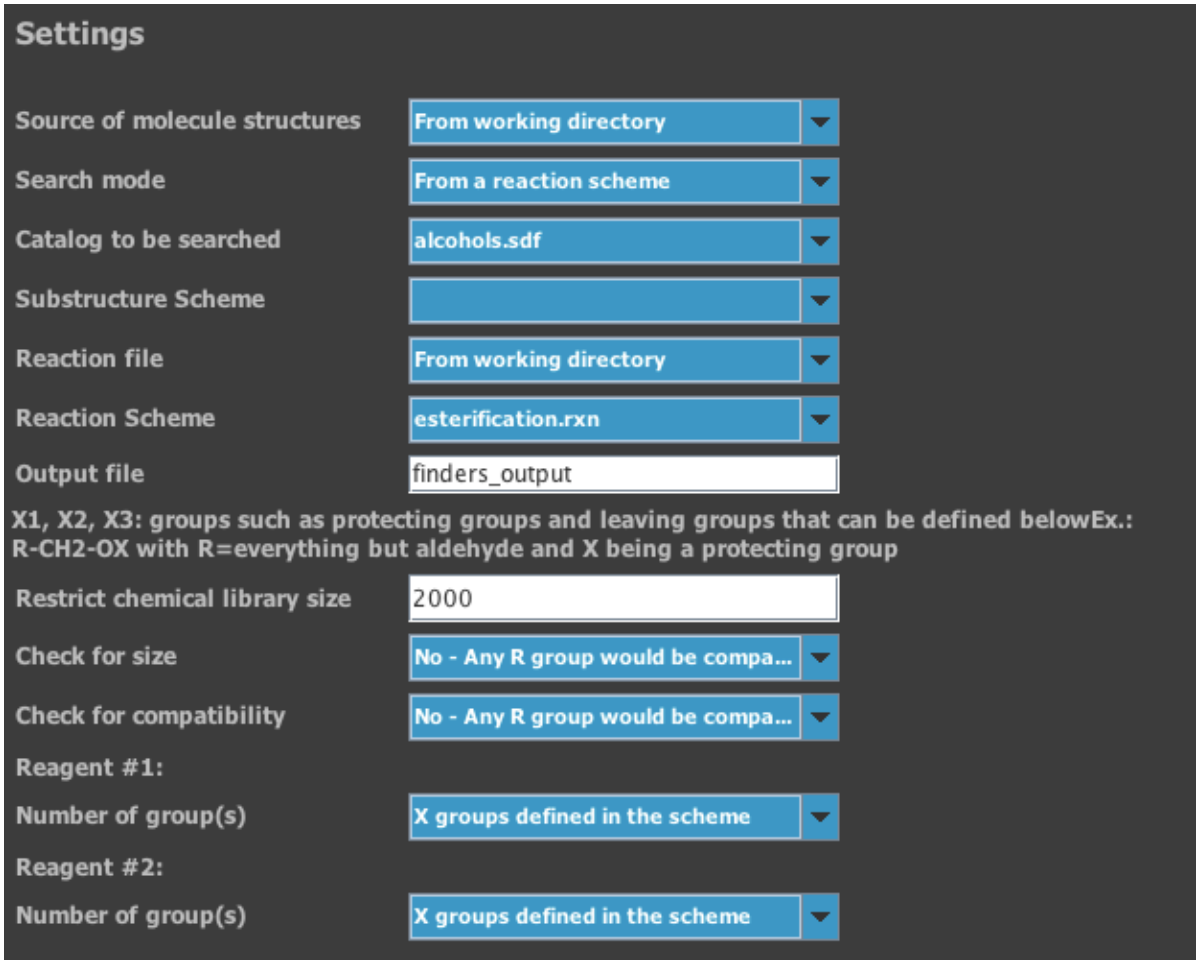
Using the same sketcher (page must be cleaned using the blank page icon on the left), a small library of alcohols is built and saved using the **File**, **Save structure (sdf)** option and providing alcohols.sdf as a filename. In other cases, large libraries of chemicals may be used (sdf format).



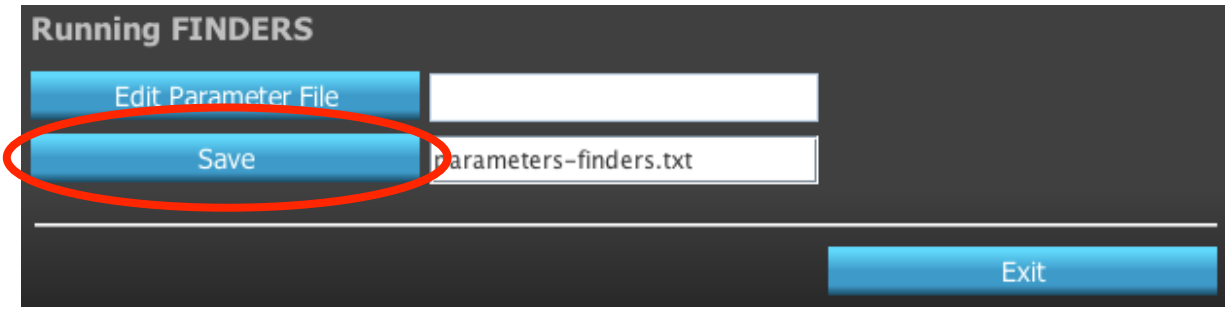
2. FINDERS: Searching libraries for compatible chemicals

Once the reaction scheme is defined, clicking on the gear icon of the FINDERS box will open the parameters section.

The **Catalog to be searched** should be "`alcohols.sdf`", the **reaction file** "From working directory", and the **Reaction Scheme** "`esterification.rxn`".



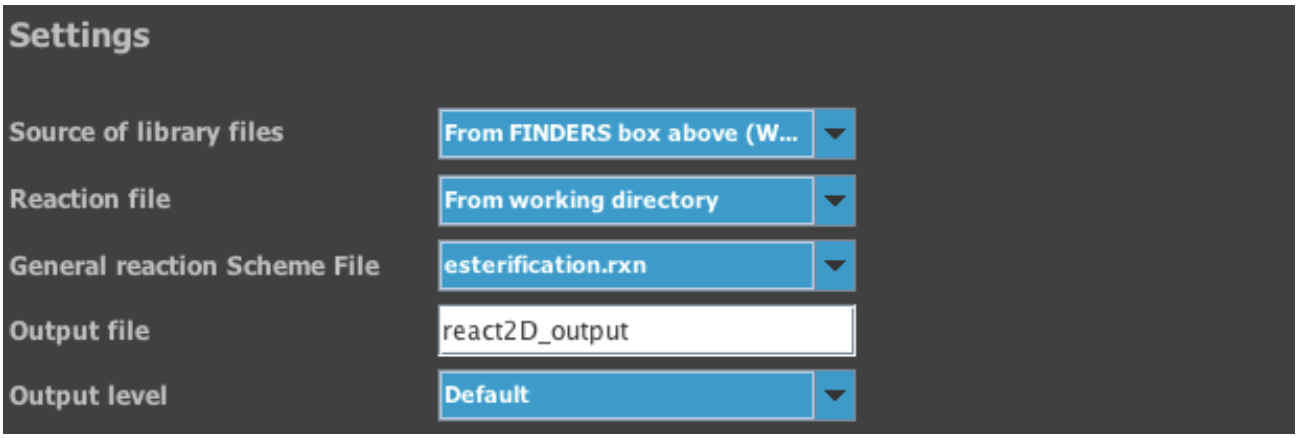
For this tutorial, the other parameters will remain defaulted. The parameter file needs to be written by clicking the **Save** button. Clicking the **Exit** button will close the FINDERS parameters section and return to the main workflow.



A green check-mark should now appear in the FINDERS box.

3. REACT2D: In silico combinatorial chemistry

Once the FINDERS settings are set, clicking on the gear icon of the REACT2D box will open the parameters section. The **Source of library files and reaction file** should be "From FINDERS box above". The **output file** can be filled with anything, "`react2D_output`" is suggested.

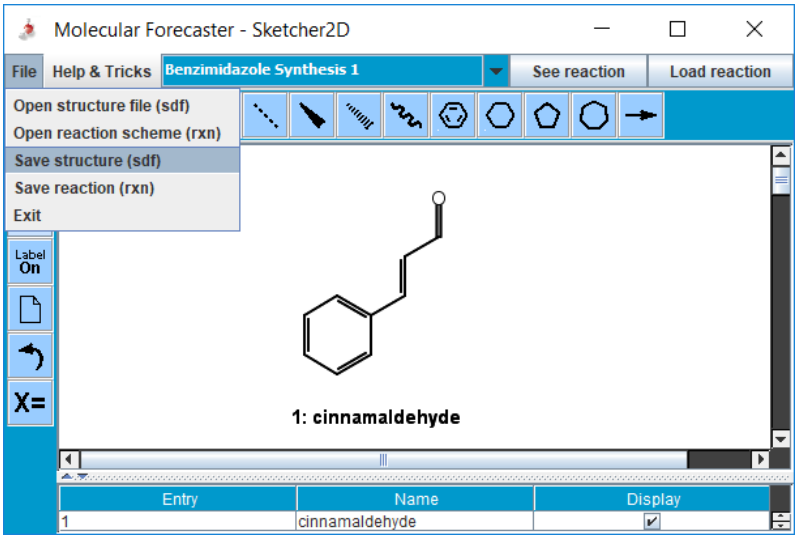


Once all the parameters are set, the parameter file needs to be written by clicking the **Save** button. Clicking the **Exit** button will close the REACT2D parameters section and return to the main workflow. A green check should now appear in the REACT2D box.

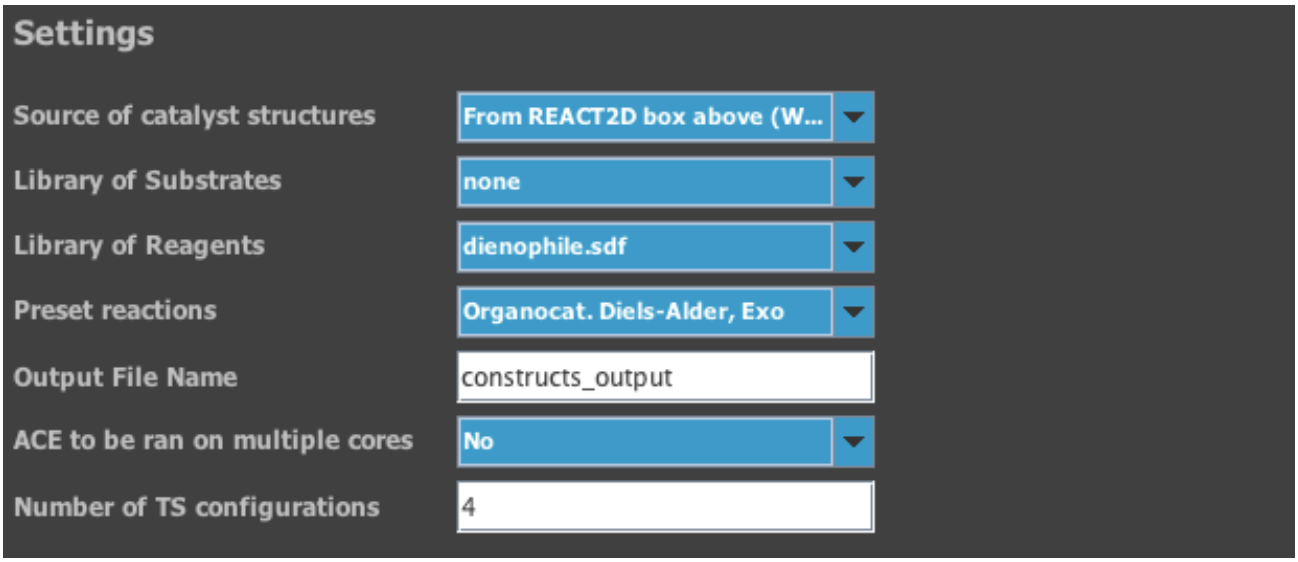
4. CONSTRUCTS: Building transition state structures

For the organocatalyzed Diels Alder reaction, we need a diene (cyclopentadiene already integrated as part of the transition state), a dienophile (which we will load or draw) and a catalyst (which was built through the FINDERS/REACT2D programs as discussed above).

For this example, the dienophile will be cinnamaldehyde and can be drawn using the sketcher and saved as "`dienophile.sdf`" or simply loaded from the tutorial folder.



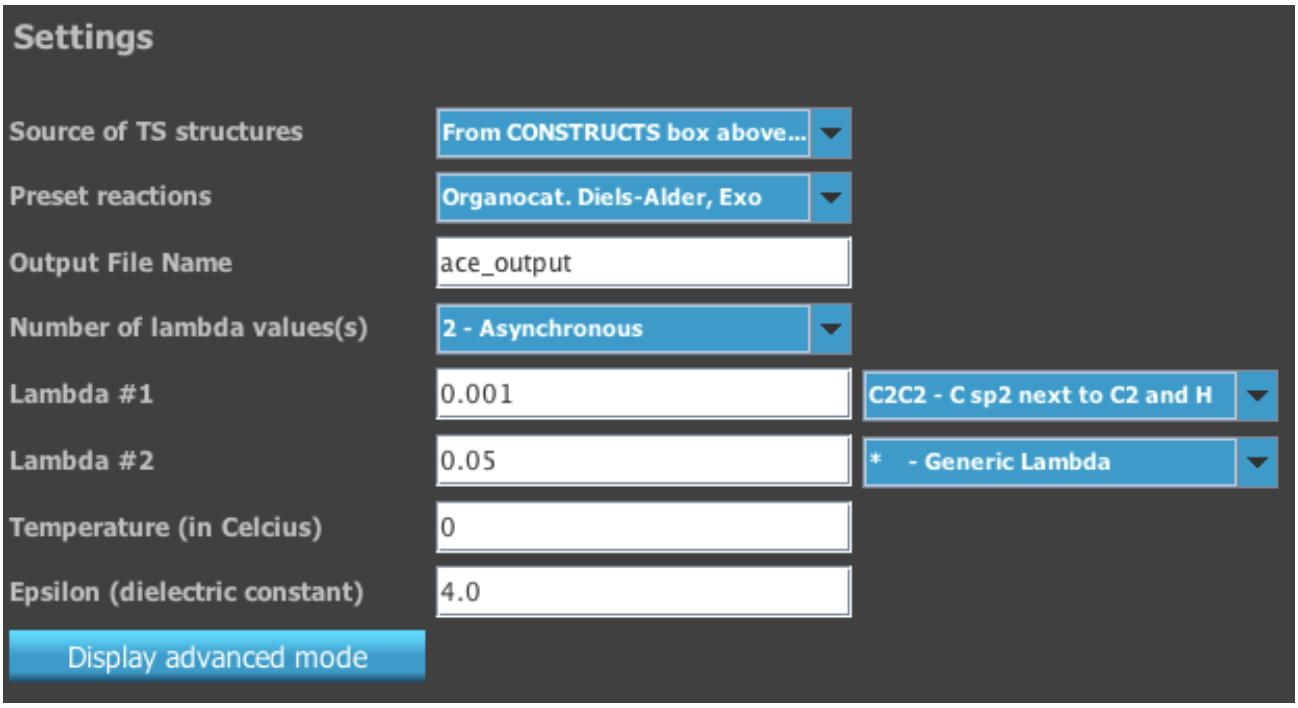
As soon as this dienophile is ready, click on the gear icon of the CONSTRUCTS box to open the parameters section. The **Source of catalyst structures** should be "From REACT2D box above"; **Library of substrates**: "none" (the actual substrate, cyclopentadiene, is already part of the pre-set reaction) and **Library of reagents**: "`dienophile.sdf`". The "Organocatalyzed Diels Alder, Exo" reaction is chosen. The **output file** can be filled with anything; "`constructs-output`" is suggested. A number of cores for the calculations can be changed. Here calculations are performed on a single core. A Library of substrates and reagents (dienes and dienophiles in a Diels Alder) can be given to CONSTRUCTS. With the preset reaction used here, cyclopentadiene is used as a reagent by default and only substrates should be given.



The parameter file will next be written by clicking the **Save** button. Clicking the **Exit** button will close the CONSTRUCTS parameters section and return to the main workflow. A green check should now appear in the CONSTRUCTS box. The program will not run until the complete workflow is ready.

5. ACE: Optimizing TS structures and computing stereoselectivity

Click on the gear icon of the ACE box to open the parameters section. The **Source of TS structures** should be "From CONSTRUCTS box above". The **output file** can be filled with anything, "`ace_output`" is suggested. As it is a pre-set reaction, the other boxes are filled. For new reactions, contact nicolas.moitessier@molecularforecaster.com for assistance.



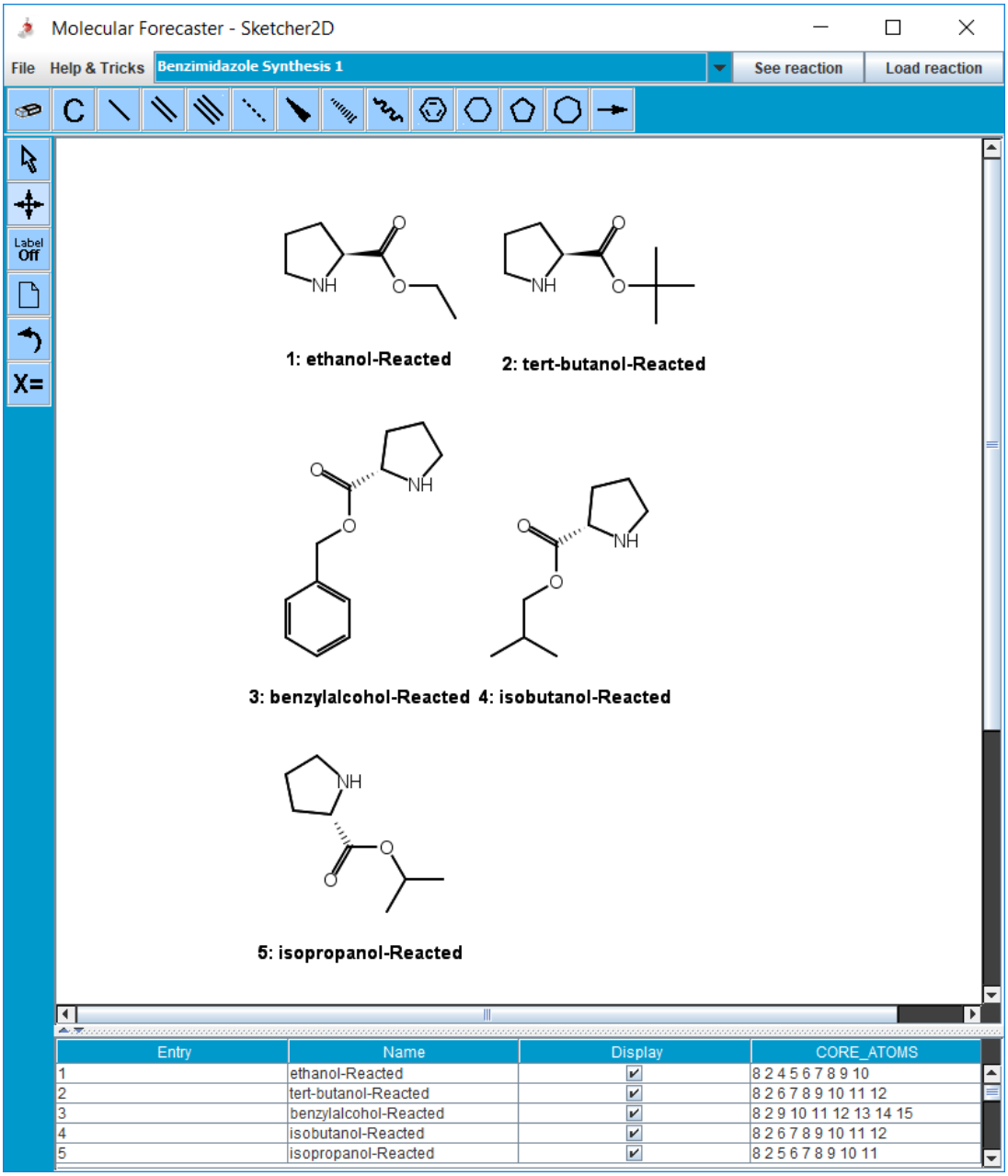
The parameter file will next be written by clicking the **Save** button. Clicking the **Exit** button will close the ACE parameters section and return to the main workflow. A green check should now appear in the ACE box.

6. Workflow: Running the VIRTUAL CHEMIST

Once all the included steps are ready, the workflow can be executed by clicking the **Run workflow** button. The programs run in a terminal (e.g. dos) and once the complete workflow is complete, the terminal window will close.

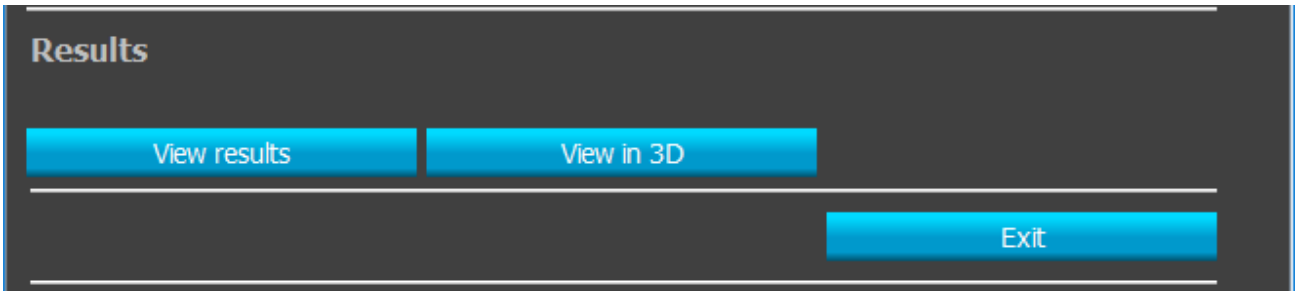
IV. Analyzing the results

Once the workflow is complete, multiple files will be created with the names assigned in each parameter box. For instance, the files generated with REACT2D are `react2D-output.out` (output file) and `react2D-output.sdf` (sdf file with the new library). The `react2D-output.sdf` file contains the entire library with the 2D structure of the various catalysts. It can be visualized in your favourite graphical program. Alternatively, the output library can be visualized by clicking the magnifier icon in the REACT2D box. Clicking the **2D structure** button will open the library within the sketcher (see Figure below) while the **Text file** button will open the library as a text file.

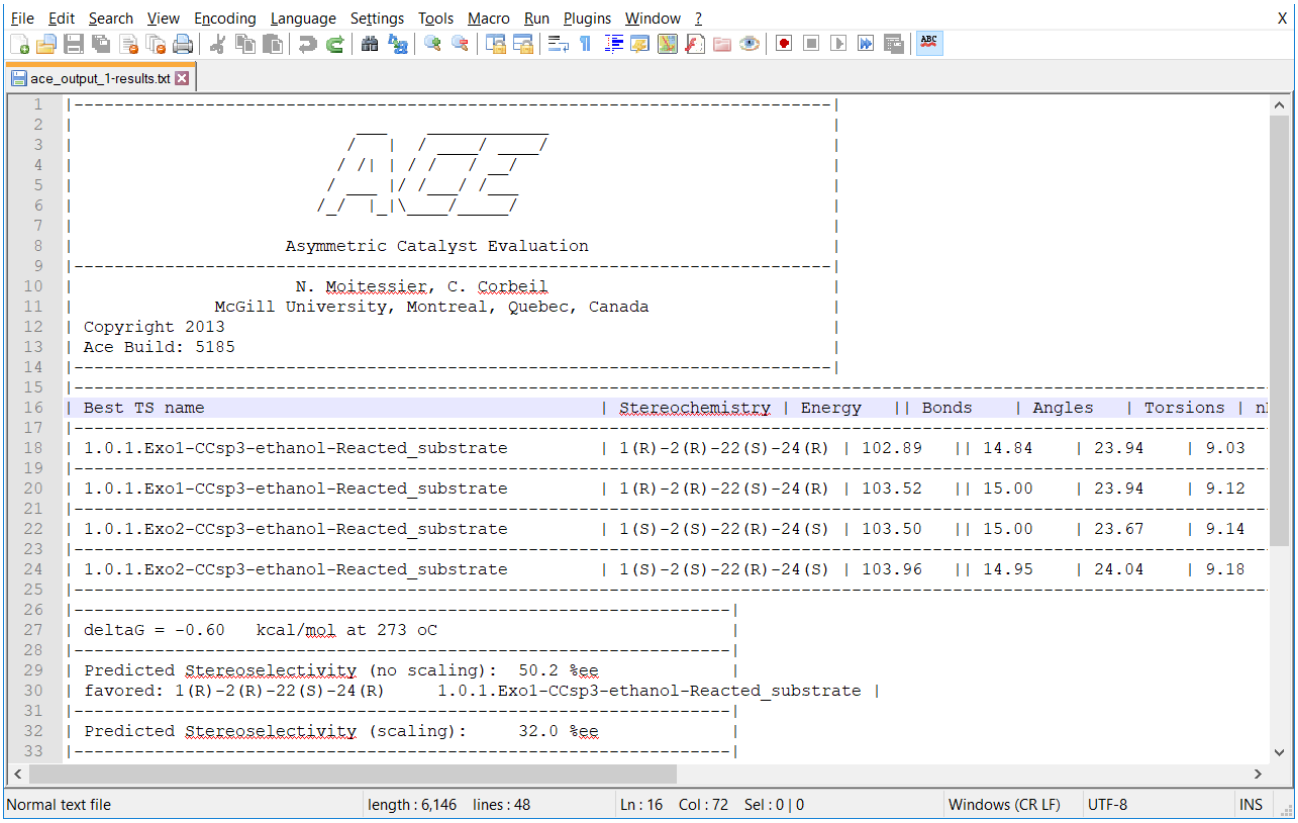


The ACE results can be found in `ace_output_1-results.txt` (calculations in the first core) and in the `mol2` files.

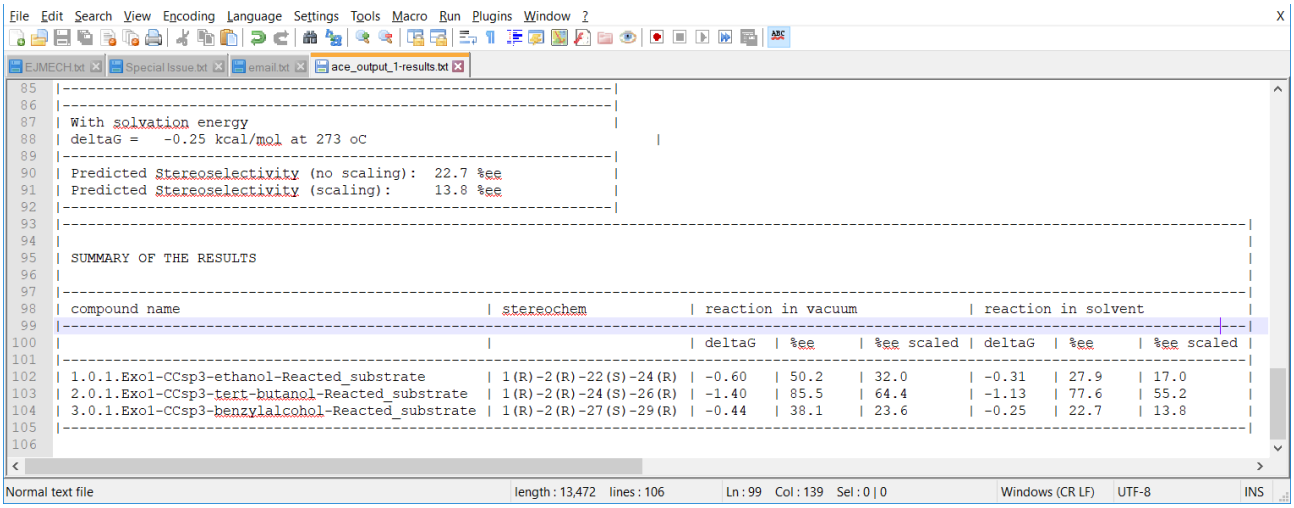
The platform can also be used. Clicking on the magnifying glass will open a new window.



By clicking on **View results**, you will open the first result file (if multiple jobs are running simultaneously, you will have to open the other result files with your favorite text editor, here Notepad++).



If you scroll down to the bottom of the file, a summary is given. The atom numbers used in the stereochemistry of each of the transitions states (e.g., 1(*R*)-2(*R*)-22(*S*)-24(*R*)) does not reflect the IUPAC nomenclature. Rather these numbers are the atom numbers in the files.



```
|-------------------------------------------------------------|
|-------------------------------------------------------------|
| With solvation energy                                       |
| deltaG =   -0.25 kcal/mol at 273 oC                         |
|-------------------------------------------------------------|
| Predicted Stereoselectivity (no scaling):  22.7 %ee         |
| Predicted Stereoselectivity (scaling):     13.8 %ee         |
|-------------------------------------------------------------|

|-------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                               |
| SUMMARY OF THE RESULTS                                                                                                        |
|                                                                                                                               |
|-------------------------------------------------------------------------------------------------------------------------------|
| compound name                                  | stereochem           | reaction in vacuum            | reaction in solvent           |
|------------------------------------------------------------------------------------------------------------------------------|
|                                                |                      | deltaG  | %ee     | %ee scaled | deltaG  | %ee     | %ee scaled |
|-------------------------------------------------------------------------------------------------------------------------------|
| 1.0.1.Exo1-CCsp3-ethanol-Reacted_substrate     | 1(R)-2(R)-22(S)-24(R) | -0.60   | 50.2    | 32.0       | -0.31   | 27.9    | 17.0       |
| 2.0.1.Exo1-CCsp3-tert-butanol-Reacted_substrate | 1(R)-2(R)-24(S)-26(R) | -1.40   | 85.5    | 64.4       | -1.13   | 77.6    | 55.2       |
| 3.0.1.Exo1-CCsp3-benzylalcohol-Reacted_substrate | 1(R)-2(R)-27(S)-29(R) | -0.44   | 38.1    | 23.6       | -0.25   | 22.7    | 13.8       |
|-------------------------------------------------------------------------------------------------------------------------------|
```

Clicking on **View in 3D** will open the first 3D transition state structure (not necessarily the lowest in energy). In the current version, if multiple jobs of ACE are run, multiple files will be created and only the first one will be open. You may use JMol within the platform or any visualizer to display the other transition state structures.