AIMS User Guide

Version 1, Late update: June 19, 2024

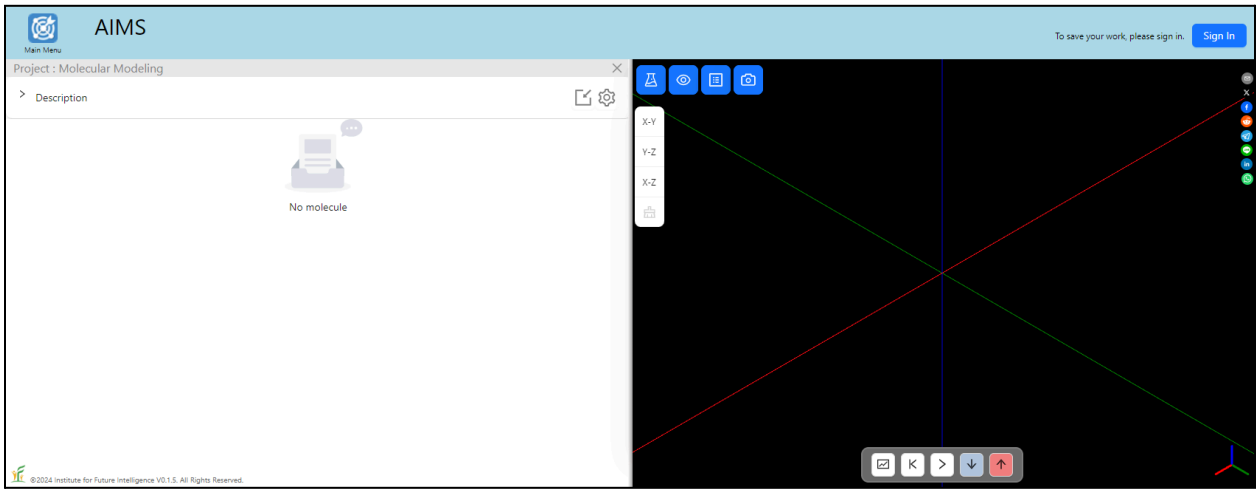
What is AIMS?

AIMS (Artificial Intelligence for Molecular Sciences) is a Web app that incorporates data science as a new pillar of scientific inquiry and engineering design in strategically important frontiers such as drug discovery and materials discovery based on the science of atoms and molecules. In addition to building the bridge between science education and scientific research, AIMS supports three-dimensional (3D) learning as envisioned in the Next Generation Science Standards (NGSS). For example, AIMS provides technology that enables the implementation of the 3D learning about the topic of Forces and Interactions. With its applications to drug discovery, it also opens the possibilities in K-12 educational settings to prepare a next generation workforce capable of addressing rare diseases occuring in underserved communities that may have been historically neglected.

This User Guide will walk you through the steps needed to integrate AIMS into your lesson plans through simple examples that are familiar to most chemistry and biology teachers.

Opening AIMS

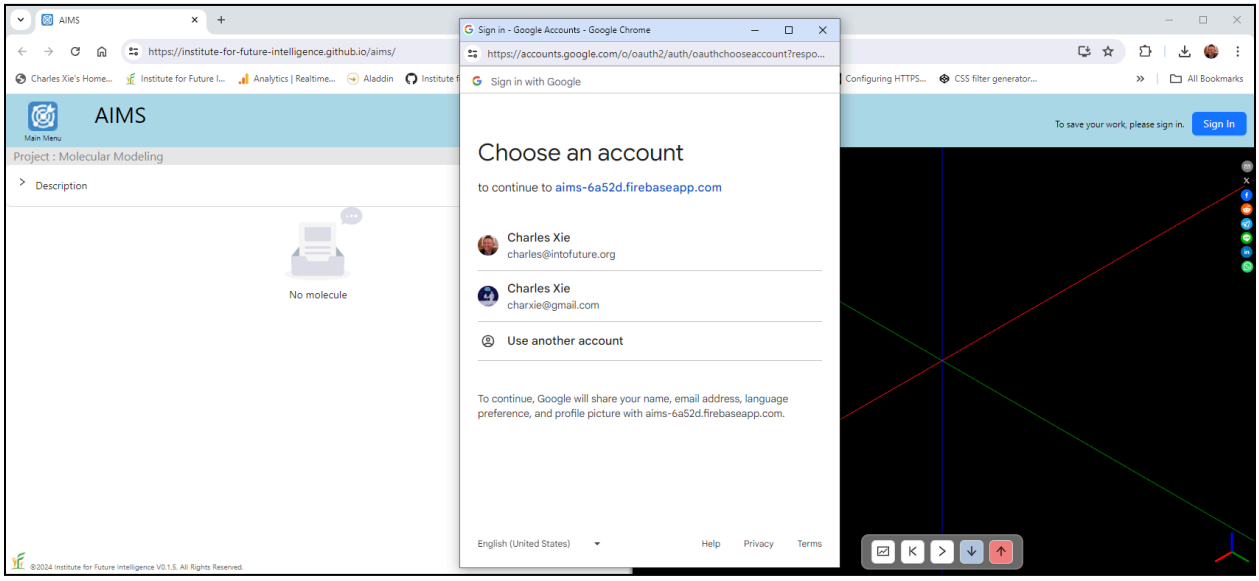
AIMS is a free, open-source Web app that can be opened by going to its homepage and clicking on the Entrance link at the upper-right corner (no download and installation is needed). By default, that link opens the following screen:



With regards to which browser to use, Firefox is not recommended at this point as AIMS seems to run much more slowly on it.

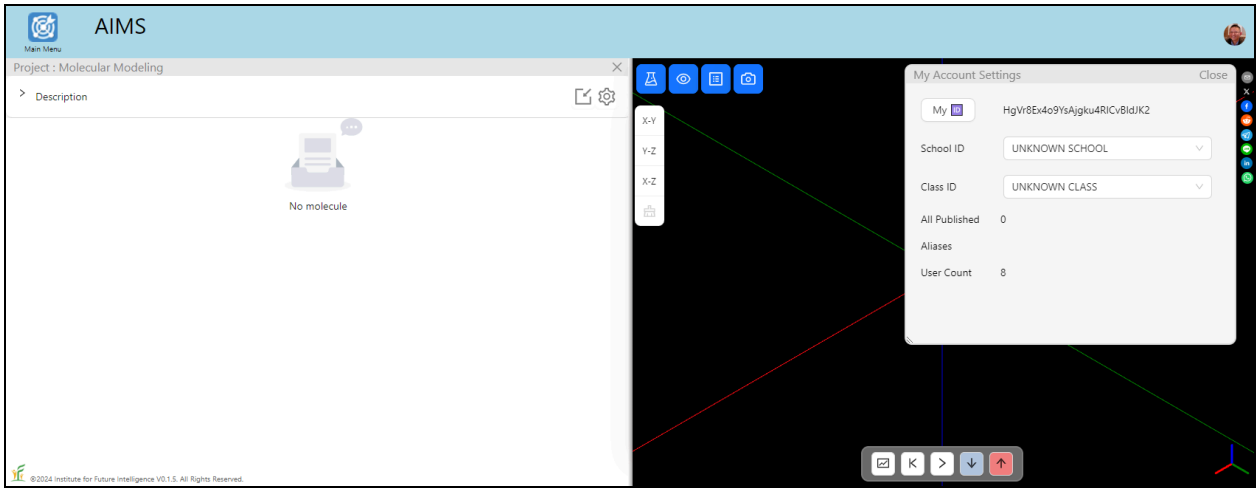
Creating an account

Although an AIMS project can be viewed by a user without creating an account, you must create one if you want to start an AIMS project of your own and save your work. To create an account, click the blue “Sign In” button at the upper-right corner. You will be prompted to choose one of your Google accounts that your browser is aware of.



Important note: Despite the message in the popup window above (“To continue, Google will share your name, email address, language preference, and profile picture with aims-6a52d.firebaseapp.com.”), AIMS does **not** store any of this information in its cloud database. All it collects is an alphanumeric string assigned by Google as a key to identify an AIMS user, such as “HgVr8Ex4o9YsAjgku4RICvBldJK2.” This key is not human readable and not related to any other apps or websites. Therefore, it cannot be used to log into any other accounts of the user.

Once you select a Google account to provide this service, you will see that your Google thumbnail photo appears at the upper-right corner in place of the “Sign In” button. If you click your photo, a popup will appear that allows you to sign out or view the accounting settings as follows (Note: For the time being, the class and school IDs are only needed for students who participate in our educational research).

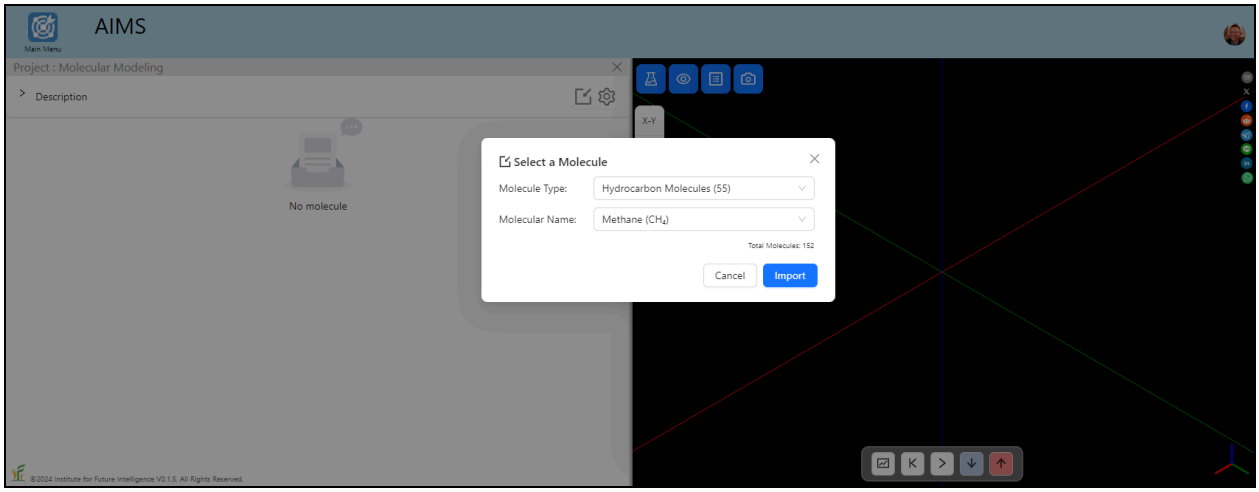


Creating a project

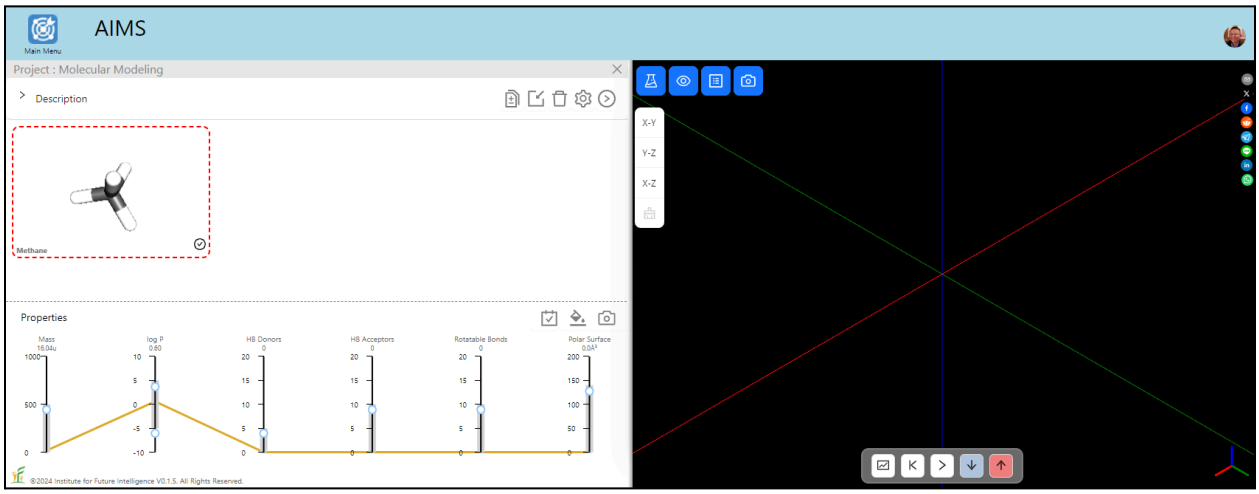
Users can create any number of AIMS projects. In AIMS, a project is represented by a visual workspace in which users perform data analytics and molecular simulation in a concerted fashion. Currently, AIMS supports two types of projects: Molecular Modeling (default) and Drug Discovery (under development). Let's use Molecular Modeling as a starting point. To create a new project, select "Main Menu > Project > Create New Project…." and follow the instructions in the popup window. If you need to copy an existing project, use "Main Menu > Project > Save Project As…". **Remember to save your project as frequently as needed using "Main Menu > Project > Save Project", as AIMS does not save your changes automatically (unlike other Web apps such as Google Docs)**.

The workspace of a project consists of two parts: The left part is the molecule gallery for displaying a collection of molecules for investigation, whereas the right part is the reaction chamber for conducting molecular simulations (i.e., virtual experiments) to test hypotheses.

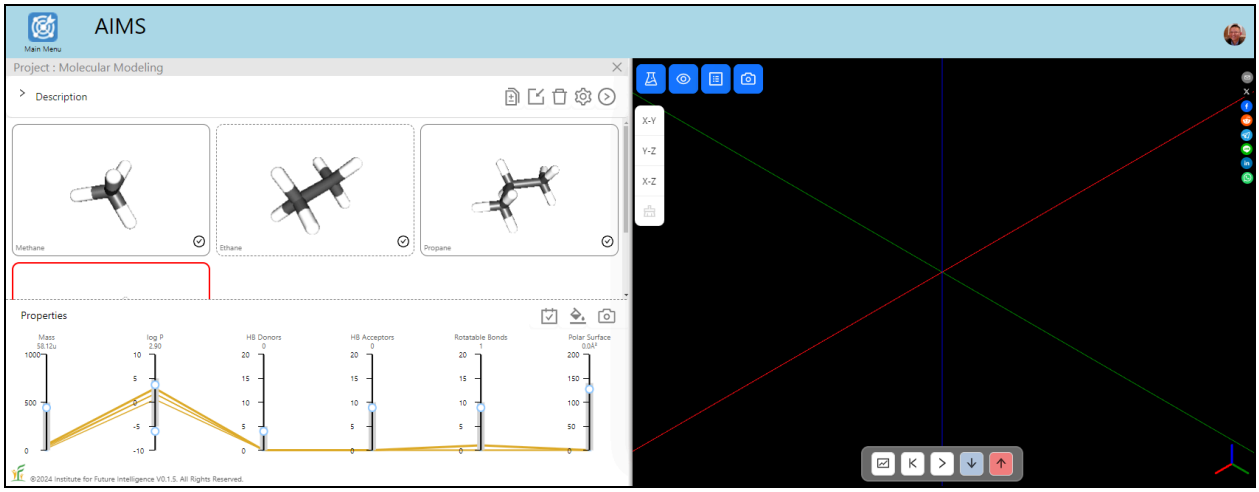
To add a molecule to the gallery, click the Import button on the toolbar above the gallery. You should see the following dialog window:



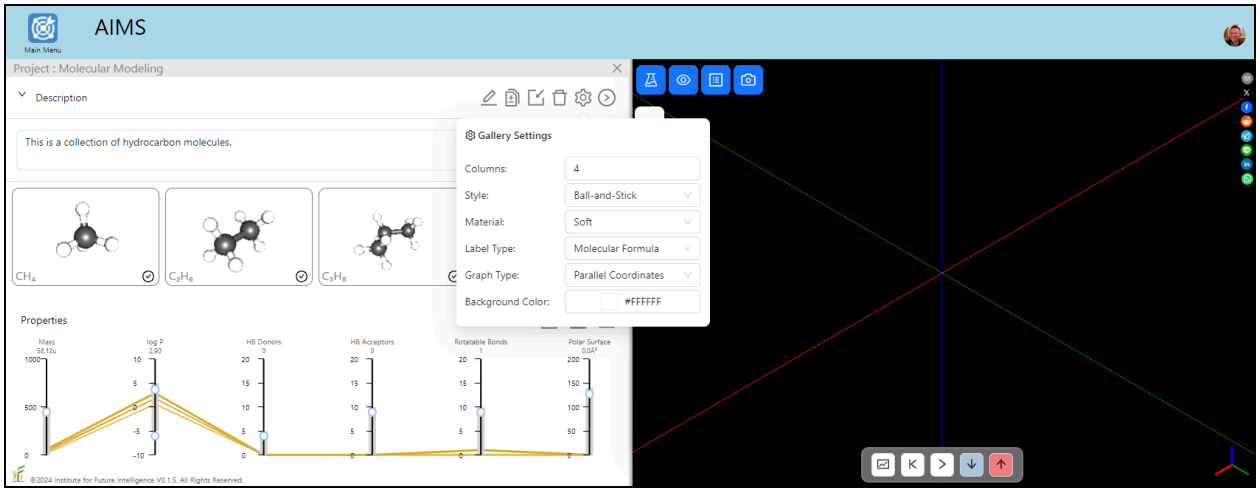
The dialog window provides you with hundreds of molecules that we have carefully curated from public databases such as PubChem and ChemSpider to ensure that no critical data are missing to pose unnecessary difficulties to beginners. These preselected molecules are categorized into different types. Let's choose Hydrocarbon Molecules from the Molecule Type dropdown menu and select Methane from the Molecule Name dropdown menu. You should see that a small window containing a methane molecule is now added to the gallery. You can interact with it with your mouse (e.g., rotate, pan, or zoom).



Repeat the above steps to add a few more hydrocarbon molecules.



You can customize the presentation view of the gallery by using the buttons above the gallery. This includes adding a brief description about the project. You can state the goals, rationale, and procedures of the project in this area. This information will be helpful when you revisit the project in the future or when you share your project with others.

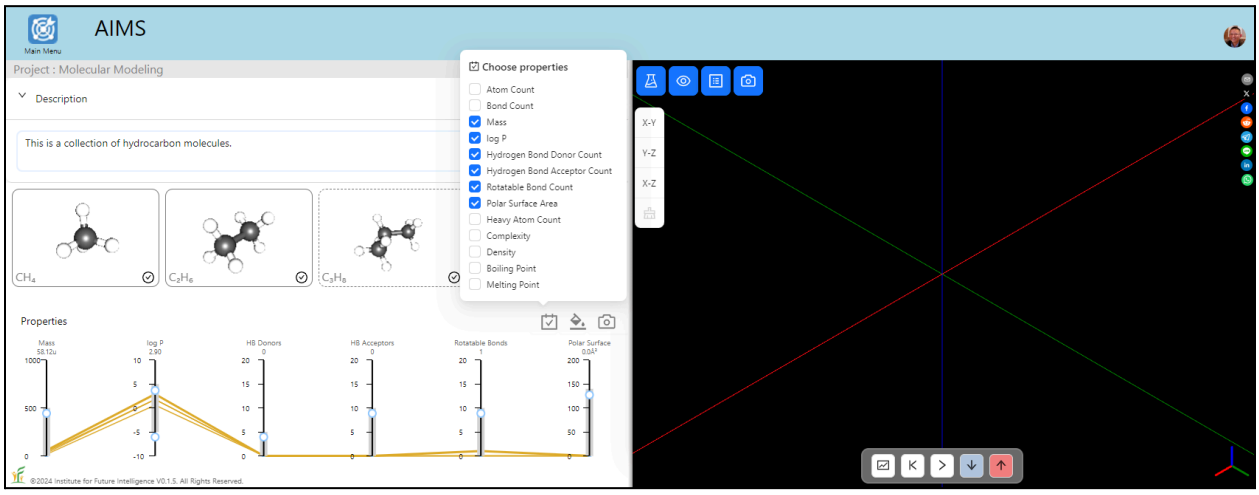


Analyzing the properties of molecules

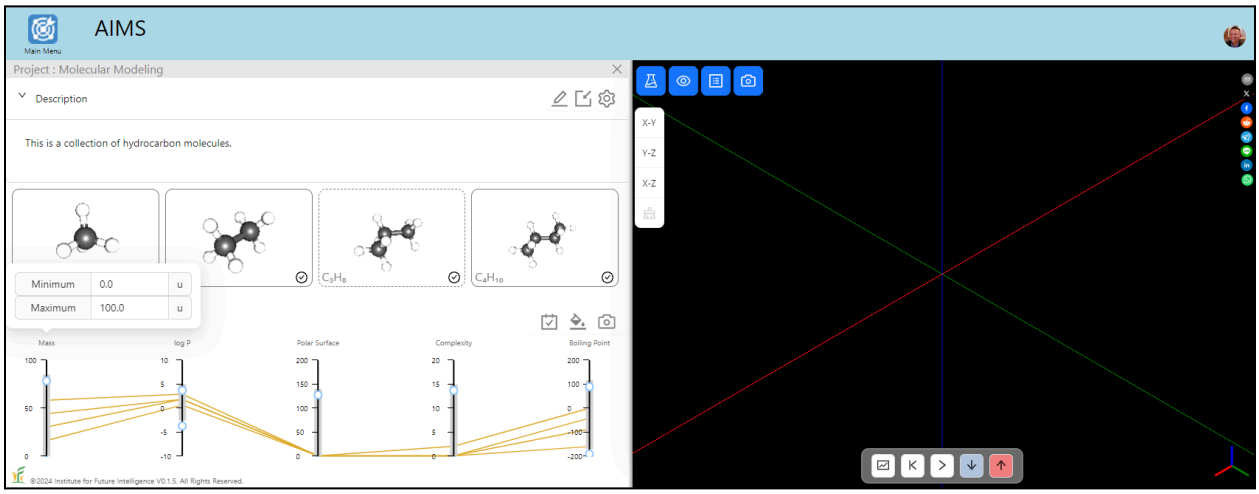
Once you have created a collection of molecules in the gallery, you may want to compare and analyze their physical and chemical properties. This is where data science techniques such as regression and classification can be applied. There are currently two visual analytics tools available in AIMS for you to perform these analyses without writing a single line of code.

Parallel coordinate plot

The first one is the parallel coordinate plot for visualizing high-dimensional datasets. AIMS provides a set of properties typically available in common chemical databases. You can choose the properties that you want to display in the parallel coordinate plot, as shown below.



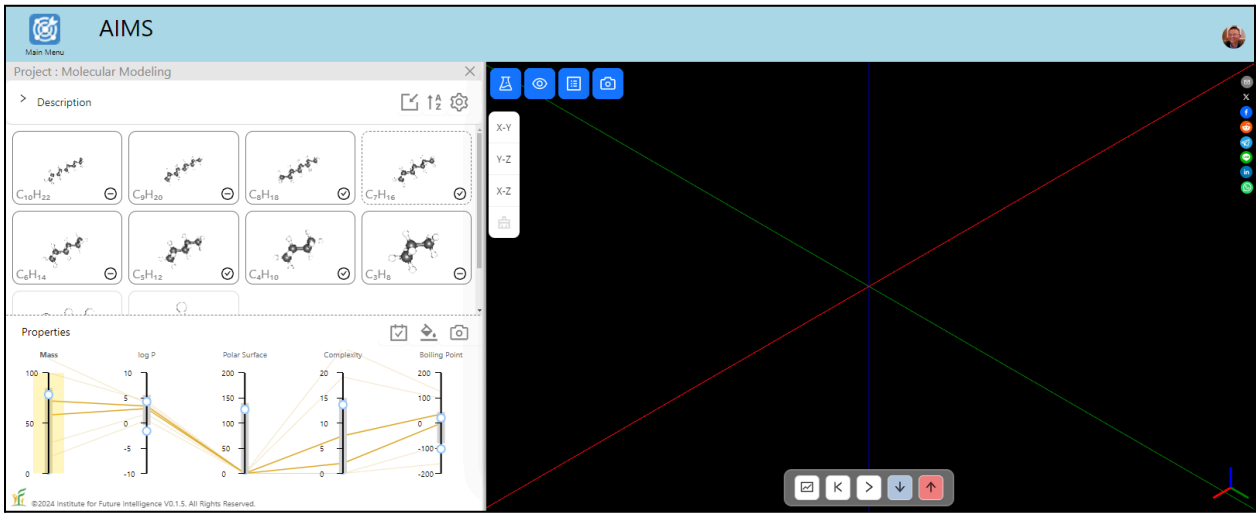
To adjust the range of an axis in the parallel coordinate plot, move the mouse over the name above the axis. A popup window will appear to allow you to set the upper and lower bounds, respectively.



Each axis is also equipped with a slider with two knobs that you can drag to set the filter range (note that the range for a filter is different from the range for the axis). The molecules with properties outside the ranges of the corresponding filters are grayed out in the gallery. With this tool, you can visually select molecules that have properties within certain scopes. This is a typical procedure in molecule screening.

Dynamic data linking and brushing features between the gallery and the plot are extensively supported. You can observe these features by moving your mouse over the molecules in the gallery or the lines in the plot. The two complementary representations (molecular structures in the 3D space and molecular properties in an abstract hyperspace) are synchronized so that you can tell which molecule corresponds to which line and vice versa.

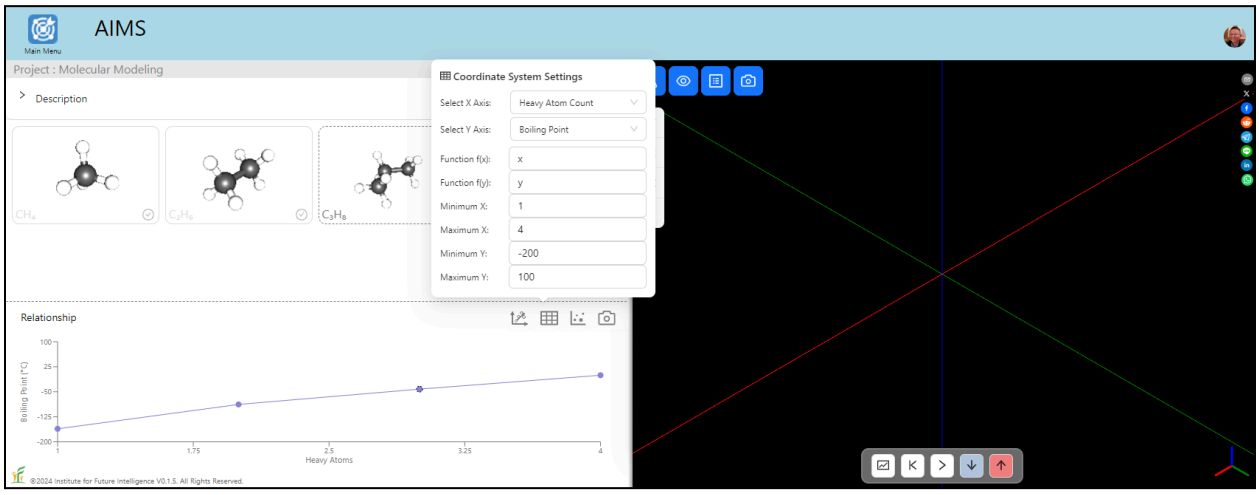
To sort the molecules in the gallery by a property, you can click on the name of the property above its axis in the plot. The axis will be highlighted, indicating that the molecules are sorted by the corresponding property. If needed, the sorting order can be reversed by clicking the sorting button on the toolbar above the gallery.



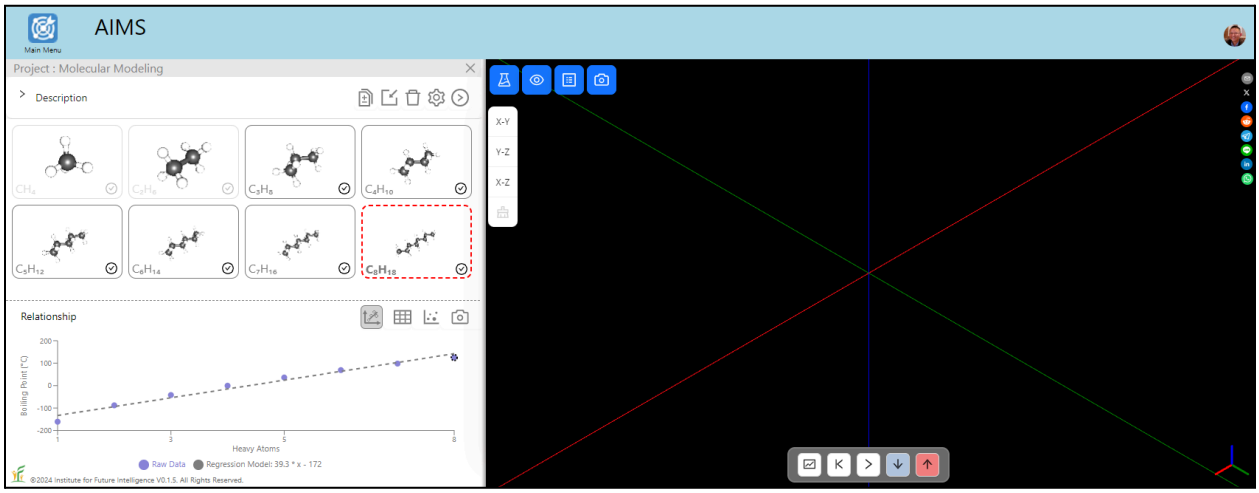
To return to the original order, just click on the name above the axis again to deselect it.

Scatter chart

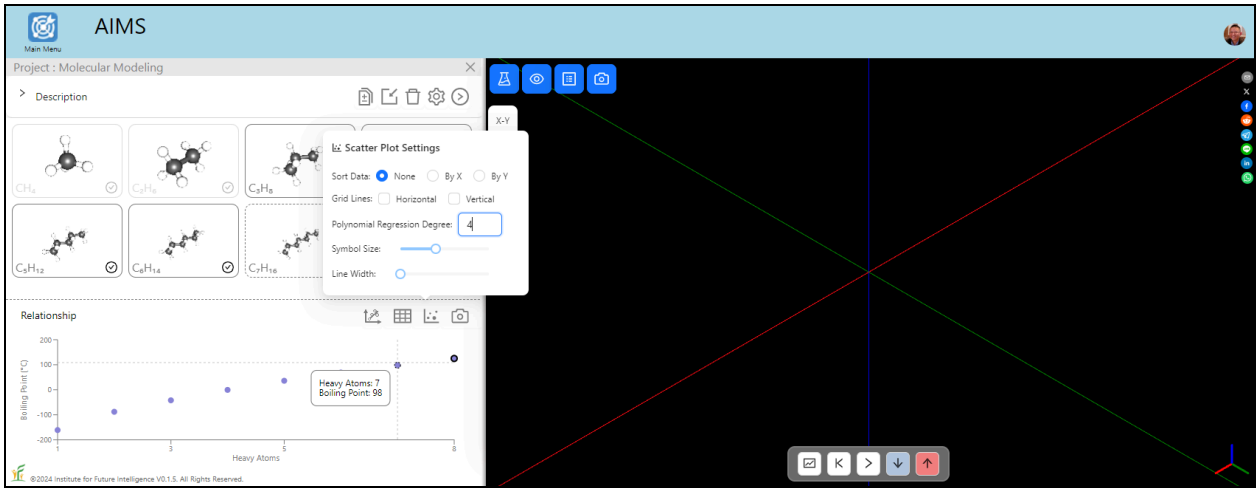
The other graph is the scatter chart that allows you to select two properties and study their relationship in a graphical way. You can customize the scatter chart by setting the ranges of its X and Y axes, the size of the symbols, the width of the line, and so on.



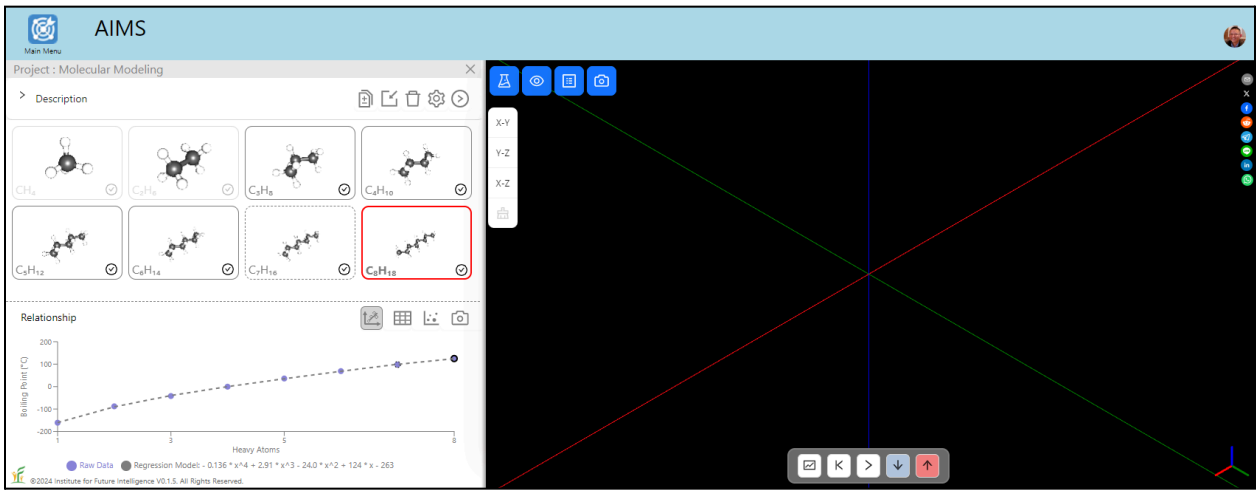
The scatter chart is equipped with a regression tool. To show how to use this built-in regression tool, let's add some more molecules to the gallery. Once you click the regression button on the toolbar above the scatter chart, a dashed line will appear that shows the result of the regression. The default is linear regression, which generates a straight line, as shown below (to show the result more clearly, the width for the line that connects the original data points has been set to zero, so the original data points are represented only by the symbols).



You can use the following popup window to increase the degree of the polynomial regression to get a better result.



For example, the degree of four yields a nearly perfect fitting of the data:



But is this regression model good enough? Let's find out in the next section.

Learning basic ideas of machine learning

The regression tool in the scatter chart can be used to teach some basic ideas about machine learning. According to Wikipedia,

> *"Machine learning is a field of study in artificial intelligence concerned with the development and study of statistical algorithms that can learn from data and generalize to unseen data, and thus perform tasks without explicit instructions."*

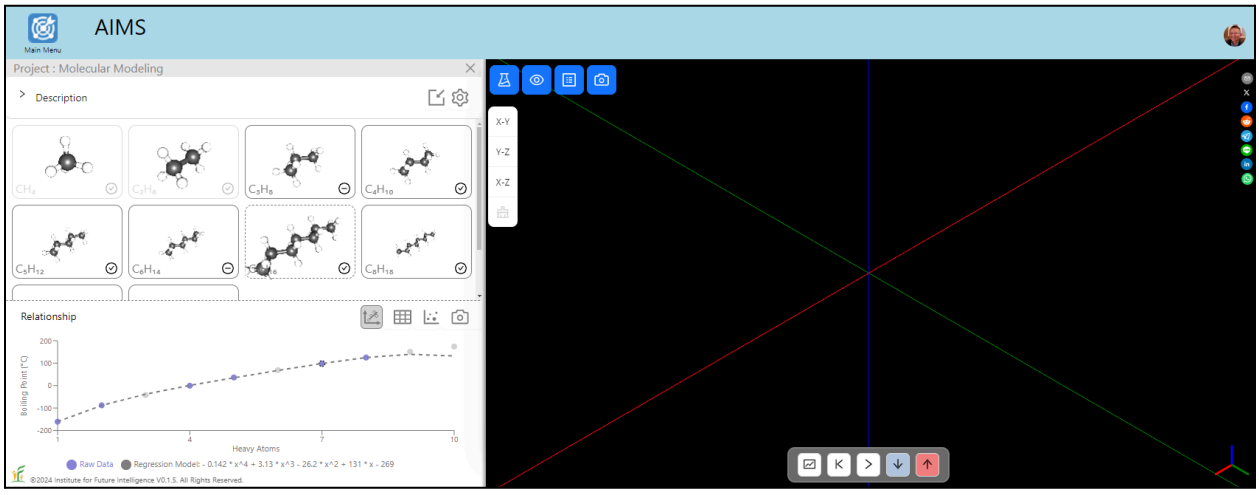
Machine learning involves two datasets: a training set and a test set. In the gallery of AIMS, you can choose a subset of molecules to form the training set and another subset of molecules to form the test set, build a regression model from the training set, and then validate the regression

model with the test set. A graphical user interface that streamlines these otherwise complex steps is available in AIMS so that you do not need to write any computer code to practice this type of machine learning.

With this easy-to-use interface in AIMS, machine learning can be employed to enhance inquiry-based learning through the following procedures:

1. Ask a question related to a chemical property or biological activity of molecules;
2. Collect information about relevant molecules from databases to prepare a training set;
3. Find patterns in the training set and build a mathematical model to represent them;
4. Use the model to predict the properties of other molecules;
5. Validate the results with a test set;
6. Repeat steps 2-5 to refine the model as needed.

The molecules in the gallery can be selected or deselected using the check circles at the lower-right corners of their corresponding display windows. The selected molecules are automatically included in the training set. Those unselected can be used as the test set. In the following screenshot that features the study of the relationship between the boiling point and the number of carbon atoms, eight molecules are included in the training set and four others are used as the test set. Among the four test molecules, two have a boiling point within the range of the training set and two have a boiling point outside the range. As you can see, the regression model can satisfactorily predict the former two but the prediction starts to deviate when it tries to predict the latter two (the one closer to the training set is better than the one further away from the training set).



Conducting molecular simulations to make sense of data

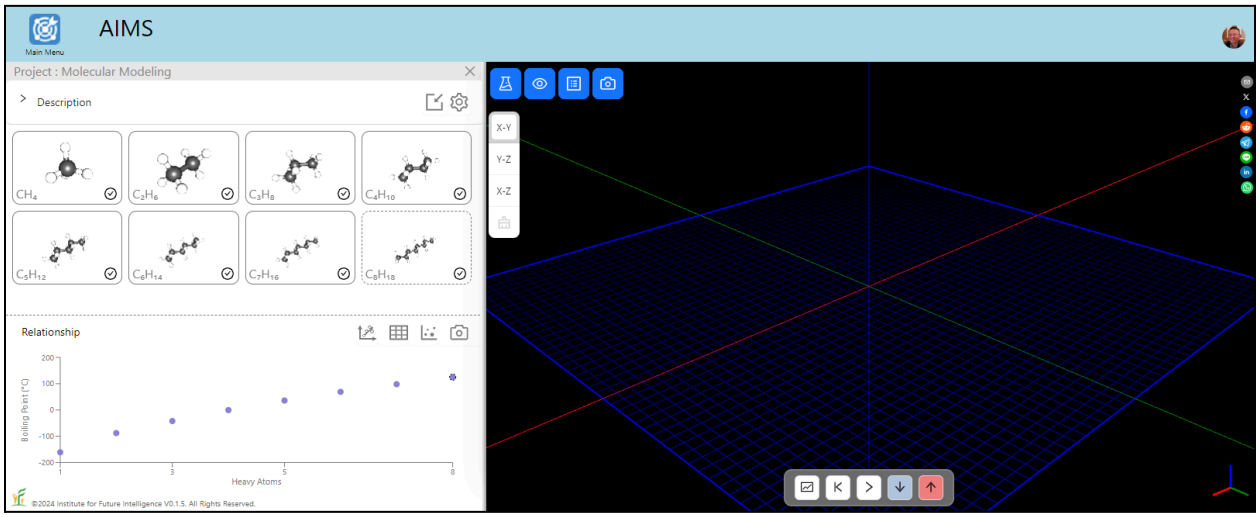
In the above sections, you have learned how to use data analytics to process and visualize molecular information. However, data analytics alone cannot explain the findings – you don't

know whether a prediction is correct or not until you design and conduct an experiment to check it. To do so, you need to understand the fundamental concepts in chemistry behind the data. Even though you have grasped those concepts, it is not always feasible to perform an experiment in the real world, especially in high schools. This is where molecular simulations come in. They provide a powerful way to learn and teach chemistry by empowering users to design and conduct virtual experiments when it is impossible or impractical to do real ones.

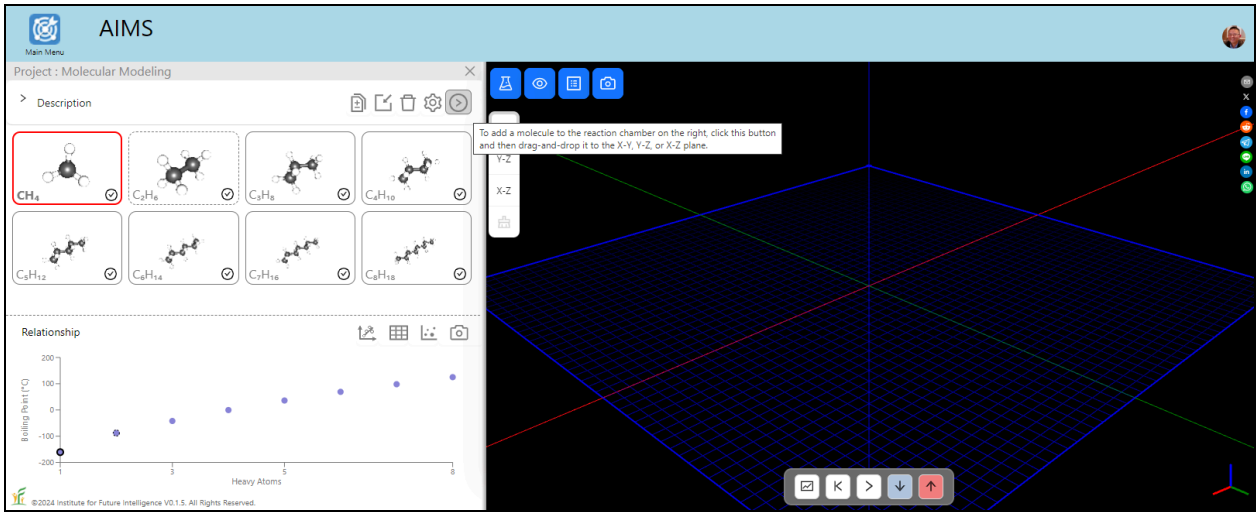
Let's go to the reaction chamber on the right of the screen to see how you can experiment with molecules with simulations.

Adding molecules to the reaction chamber

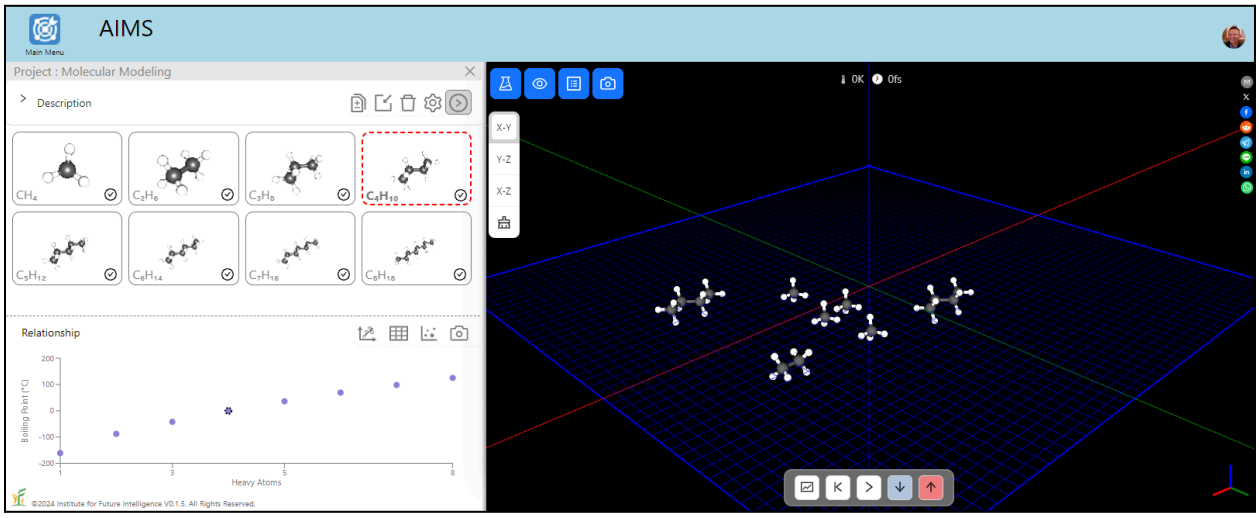
To add a molecule to the chamber, you need to set up a plane first. For instance, you can turn on the X-Y plane by clicking the X-Y button of the toolbar on the left edge of the chamber.



Next, press the rightmost button of the toolbar above the gallery.

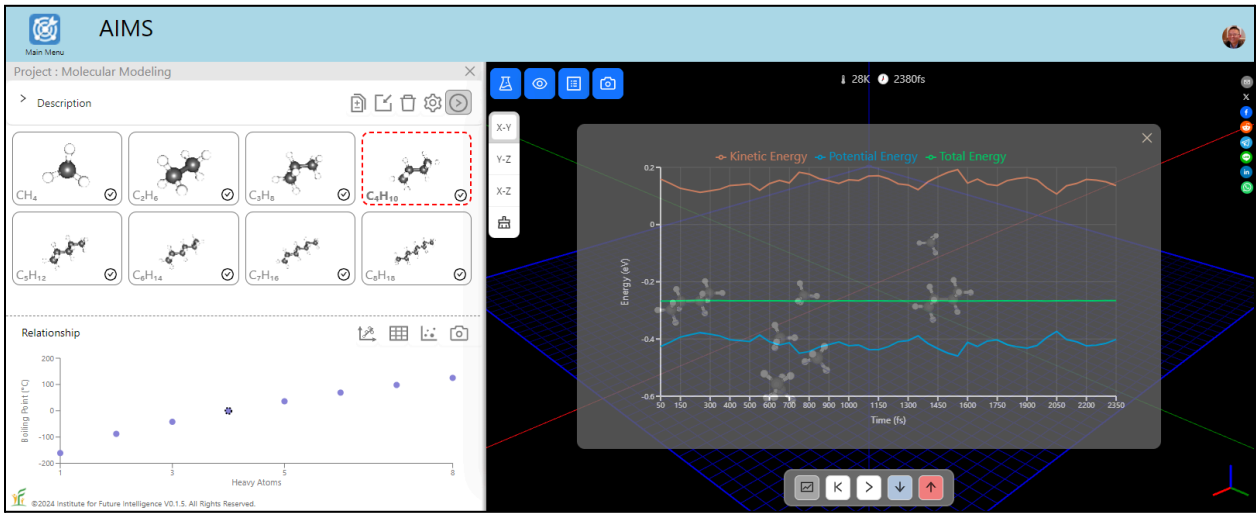


This switches on the drag-and-drop mode for adding a molecule. When that button is pressed, you can drag a molecule in the galley on the left of the screen and drop it onto the X-Y plane on the right of the screen (note that if you release the dragged molecule outside the X-Y plane or any other planes that are turned on, no molecule will be added). You can add as many molecules to the chamber as needed, as shown in the following screenshot.



Performing molecular dynamics simulations

After adding some molecules, you can immediately run a molecular dynamics simulation by clicking on the right-arrow button on the control panel at the bottom of the reaction chamber. The control panel also provides a few tools for heating, cooling, resetting, and so on. For example, clicking on the graphing button opens a graph window that allows you to monitor the time evolution of the kinetic, potential, and total energy of the molecules as the molecular dynamics simulation is under way. For molecules that are isolated in the simulation box, the total energy remains a straight line, representing the conservation of energy.

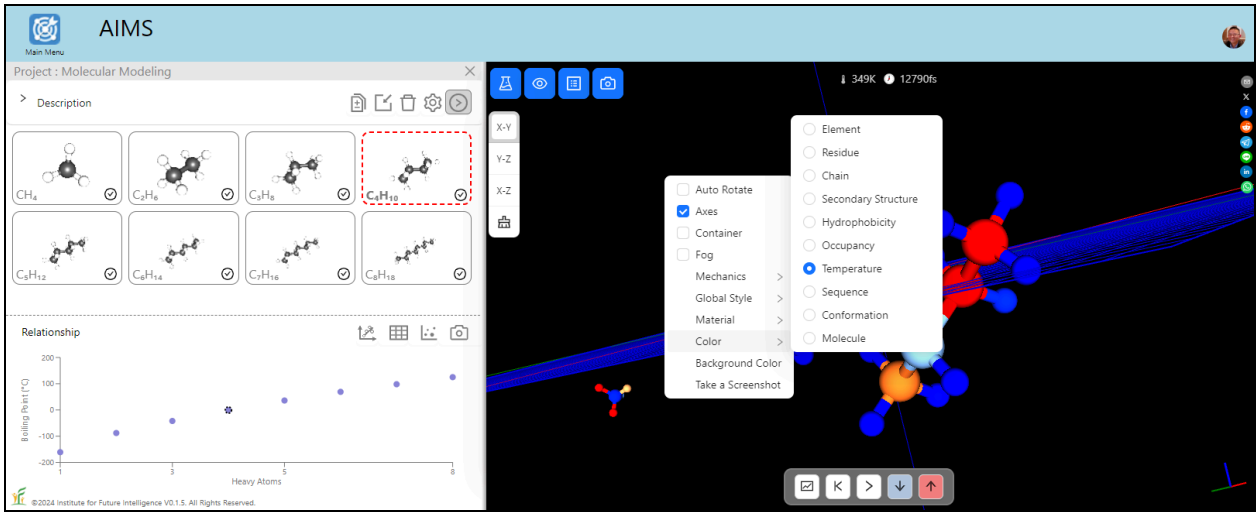


With molecular dynamics, you can design and perform virtual experiments to verify the relationship between the boiling point and the size of a molecule. Note that at this point, the results may be only qualitatively correct. In other words, you can observe that smaller hydrocarbon molecules have lower boiling points than larger ones, but the predicted values may not agree exactly with the experimental values. That said, qualitative results may be good enough for learning and teaching purposes.

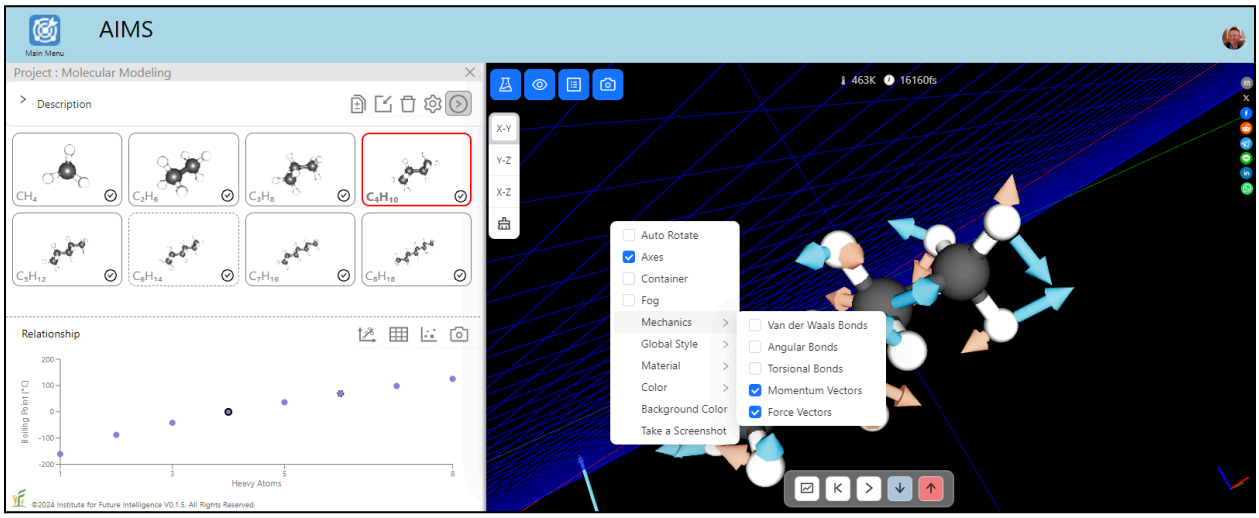
Observing and analyzing molecular simulations

An advantage of visual and interactive molecular simulations conducted in real time is that you can observe what happens among the molecules when the temperature increases or decreases (or other conditions change). The reaction chamber also provides a set of tools to enhance the simulation results, which are available in the menu that pops up when you click on an empty spot in the chamber or select “Main Menu > View > Mechanics”. In the following, you can see the effects of some of these enhancements.

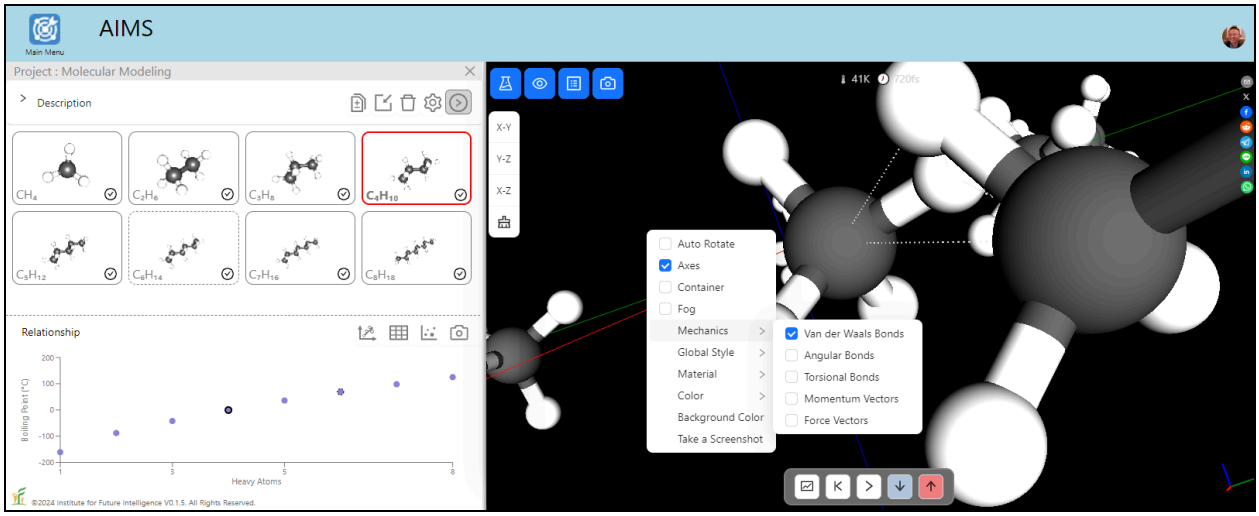
Kinetic energy shading displays the atoms in the colors representing their respective kinetic energy. Blue indicates lower kinetic energy (cold) and red indicates higher kinetic energy (hot).



As their names suggest, **momentum and force vectors** show the magnitudes and directions of the momenta and forces of the atoms. These representations may be useful when teaching the Kinetic Molecular Theory in introductory chemistry.



The **van der Waals Bonds** show the intermolecular interactions among the molecules as dashed lines. Different from the stable covalent bonds that are represented by solid sticks, the van der Waals bonds may form and break during a simulation. They are the main intermolecular forces associated with melting and boiling of hydrocarbons.



Other control parameters of molecular dynamics simulations

The reaction chamber also provides a set of tools for configuring a molecular dynamics simulation. For example, you can change the time step, impose a heat bath to maintain constant temperature, adjust the size of the simulation box, and so on. As with other molecular viewers, AIMS offers a set of molecular visualization options, such as ball-and-stick, space-filling, contact surface, and so on.

Acknowledgements

This project is supported by the National Institutes of Health (NIH) under grant number R25GM150143. Any opinions, findings, and conclusions or recommendations expressed in this material, however, are those of the authors and do not necessarily reflect the views of NIH.