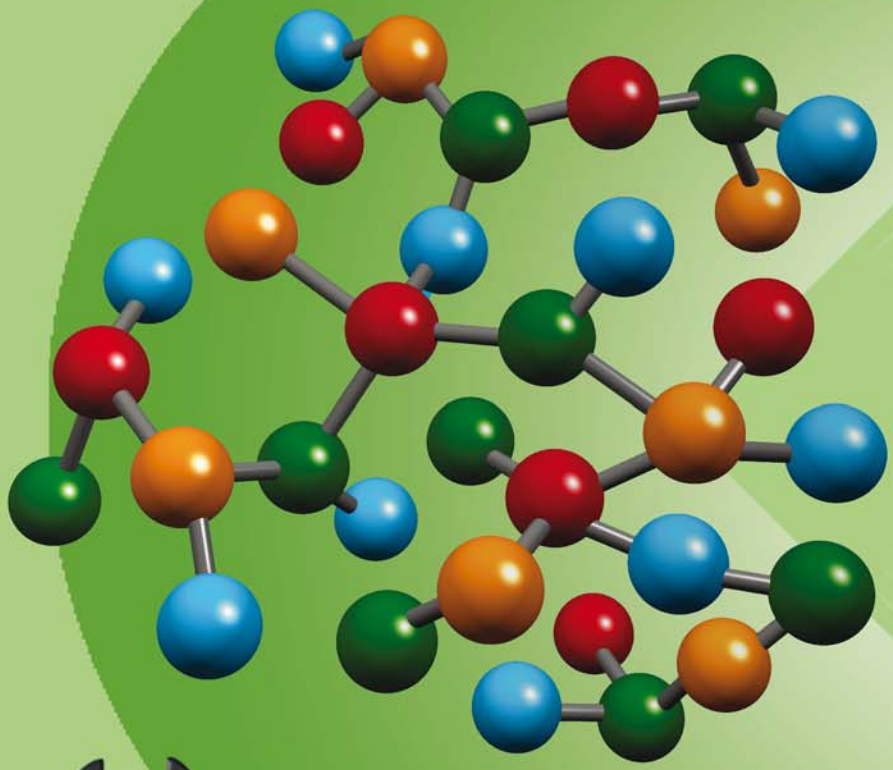
CEPOS INSILICO
SCIENCE FOR SCIENTISTS

1.0

USER MANUAL



CADDLE®

© 2013. COPYRIGHT CEPOS INSILICO LTD

Impressum

Copyright © 2013 by Cepos InSilico Ltd.

The Old Vicarage
132 Bedford Road
Kempston
BEDFORD, MK42 8BQ
United Kingdom
www.ceposinsilico.com

Manual David Whitley, Timothy Clark

Layout **www.eh-bitartist.de**

TABLE OF CONTENTS

© Cepos InSilico Ltd. 2013

1 INTRODUCTION

CADDLE.® is a new molecular modeling suite that runs entirely on a remote or local web server and that requires no software installation at all on the user's desktop. All that is required is a suitable browser (please see this **web page** for a list of supported browsers). This innovative architecture allows schools, colleges, university department or companies to install CADDLE.® on a single central server in order to make it available to all users. This flexibility does not come at the cost, for instance, of reduced graphics quality. The molecular visualization offered in CADDLE.® was developed by Cepos InSilico's partner **Molcad GmbH** and offers the state-of-the-art visualization that Molcad users expect. CADDLE.® is conceived to link with a number of modeling programs, the first of which is Cepos InSilico's **EMPIRE™**, a modern, parallel semiempirical molecular orbital (MO) program.

© Cepos InSilico Ltd. 2013

2 THE PRINCIPLES BEHIND CADDLE.®

The CADDLE.® workflow is based on a series of ***projects***, each of which contains one or more molecules and a set of calculations performed on those molecules. Initially, ***projects*** are assigned to the user who created them but can later be shared with other users. Molecules can be added to ***projects*** or new ***projects*** created and earlier calculations on a given project can be retrieved for further analysis (including visualization of molecular properties or orbitals). ***Projects*** are administered using the ***Projects*** tab on any of the CADDLE.® pages. Calculations are performed by selecting a compute engine (for instance ***EMPIRE***) from the main menu and the results are then visualized in the browser directly. No dedicated visualization program is required. Both the molecular and the visualization calculations are performed on the web server (which may be a central high-performance server or the local desktop machine) so that the calculations can be scheduled optimally by CADDLE.®. Visualization is achieved by calculating three-dimensional WebGL scenes that can be manipulated (rotated, translated, zoomed) using a suitable web browser. Not all web browsers are equally suitable and we recommend that you test your combination of machine and browser (click here for the **test web page**). CADDLE.® includes a user-management facility; the CADDLE.® administrator can add or delete CADDLE.® users, change passwords *etc*.

© Cepos InSilico Ltd. 2013

3 GETTING STARTED

If your system manager has set up CADDLE.® to require user names and passwords, your first view of CADDLE.® when you access the URL provided by your system administrator will be the login page shown in **Figure 1**, which is self-explanatory.

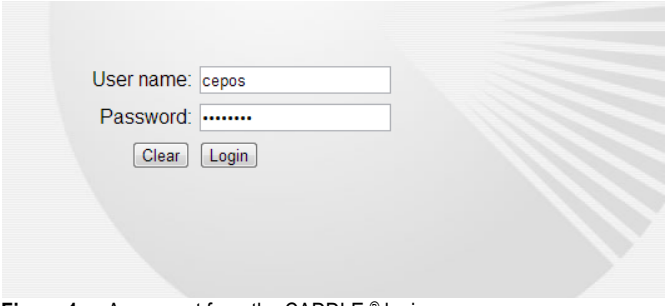


Figure 1 An excerpt from the CADDLE.® login page

Once you have logged into the system, CADDLE.® displays your personal **Home** page, which contains some user details, a link to change your CADDLE.® password and a list of recent ***projects*** (see **Figure 2**).The personal menu bar at the top right of the page provides links to view all available ***projects***, to access the online help and to logout from CADDLE.®.

Figure 2 The CADDLE.® home page

© Cepos InSilico Ltd. 2013

3.1 Creating a new project

To create a project, select the **New Project** button below the list of recent ***projects*** on the **Home** page (see **Figure 2**). As explained in the introduction, a project may be either a list of compounds that will all be treated in the same way or it may also be a single molecule. Enter the details of the new project (a name and an optional brief description) in the appropriate boxes, as shown in **Figure 3**, then click the **Create Project** button to add the project to CADDLE.®.

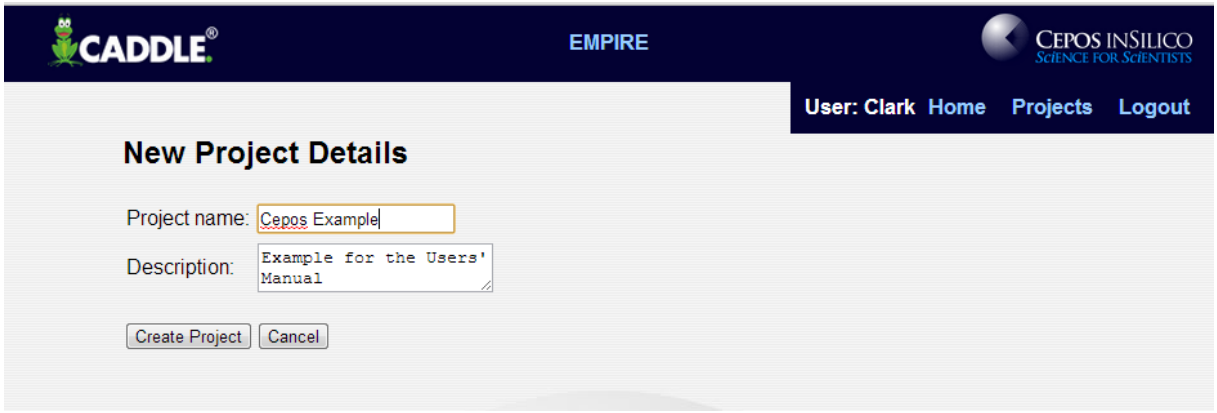


Figure 3 Creating a new CADDLE.® project

A **Project Details** page is then displayed (see **Figure 4**). Select the **Add Molecules** button to access the molecule input page, which provides three ways to add molecules to the project (see **Figure 5**.)

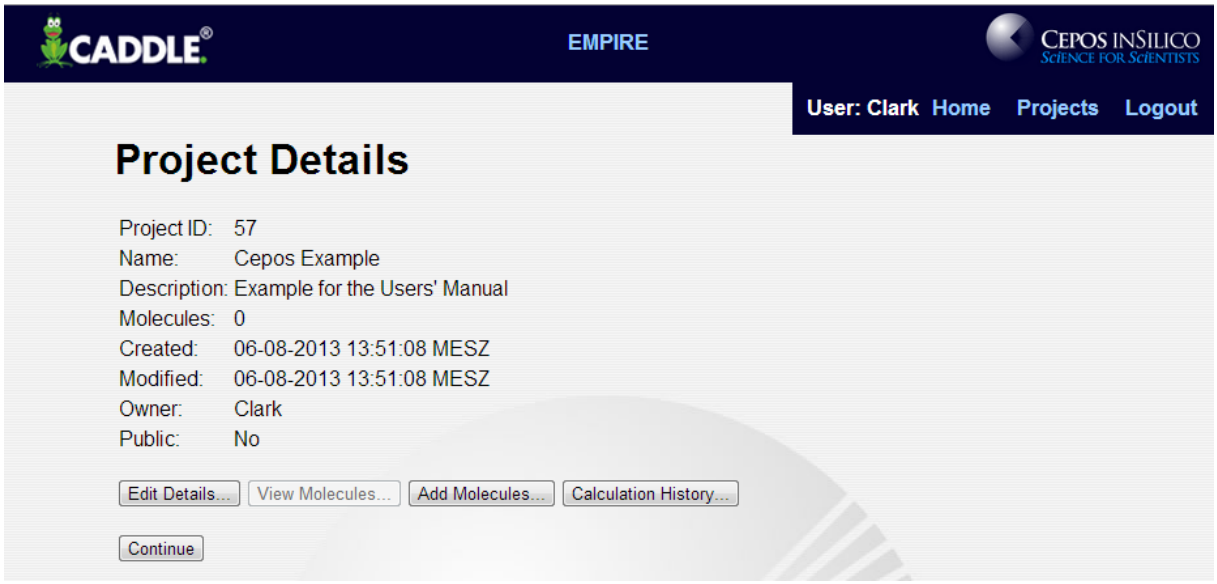


Figure 4 A CADDLE.® project details page

© Cepos InSilico Ltd. 2013

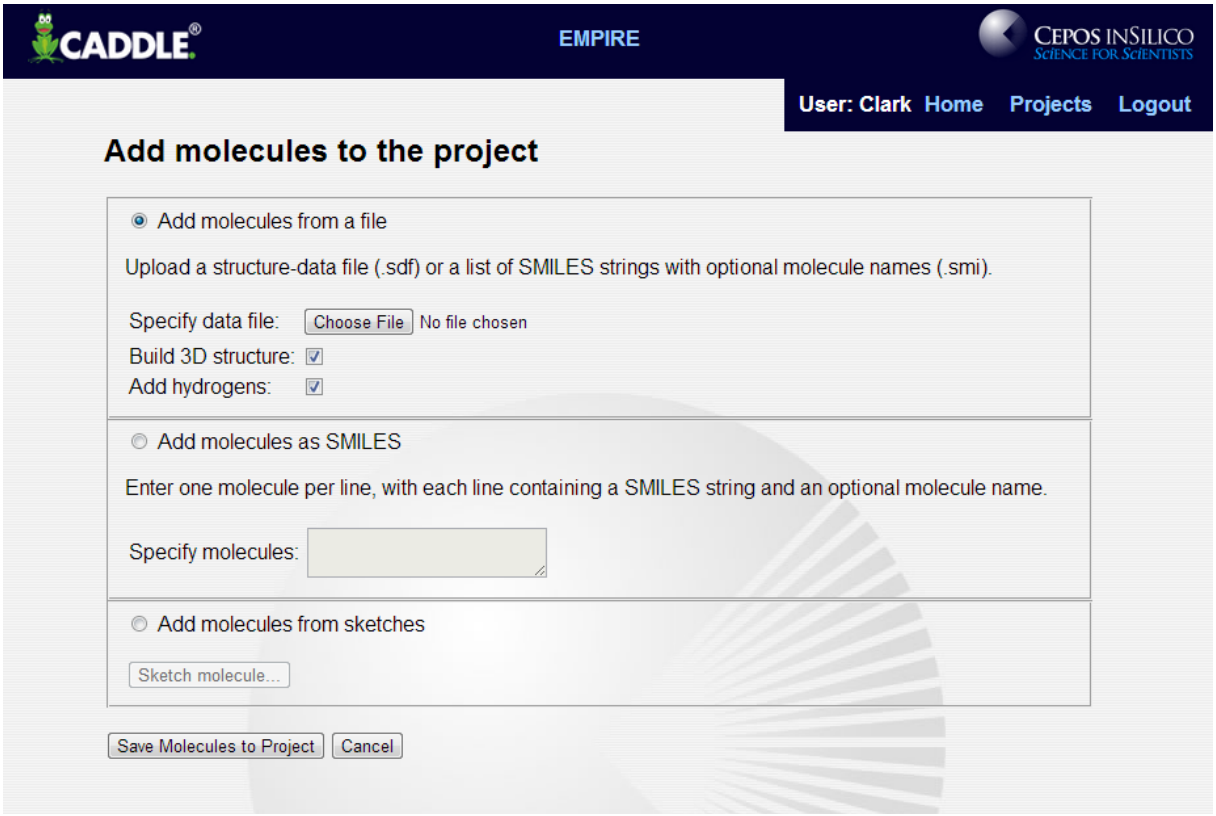


Figure 5 The CADDLE.® molecule input page

3.2 Reading molecules from a file

Molecules can be read as input for the project from structure-data (SDF) files or from lists of SMILES strings. SMILES input files should contain one SMILES per line, followed by an optional molecule name, and must have a `.smi` file extension. Molecules entered in SMILES format will have hydrogens added and 3D coordinates generated automatically by CADDLE.®. For molecules entered in SDF format, CADDLE.® can read two-dimensional structures and convert them to 3D automatically or it can read 3D structures produced by another program without changing them. In order to read a list of molecular structures from a file, select the **Add molecules from a file** option, choose the input file in the file browser and request the appropriate options to generate 3D coordinates or add hydrogens, as shown in **Figure 5**. CADDLE.® will load the molecules into the current project when you click the **Save Molecules to Project** button.

© Cepos InSilico Ltd. 2013

3.3 Defining molecules using SMILES

A single molecule or a list of molecules can also be loaded directly into CADDLE.® using SMILES notation. To do this, select the **Add molecules as SMILES** option and type (or cut-and-paste) a list of SMILES strings into the text entry area. Enter one SMILES per line, with an optional molecule name. In this case, the 3D coordinates will be generated by CADDLE.®. Once again, clicking the **Save Molecules to Project** button loads the molecule(s) into CADDLE.®.

3.4 Sketching molecules

The third method for inputting molecules into CADDLE.® is to use the built-in 2D sketcher. In order to access the sketcher, select the **Add molecules from sketches** option and click the **sketch molecule** button shown in **Figure 5**. This brings up the Sketcher window shown in **Figure 6**.

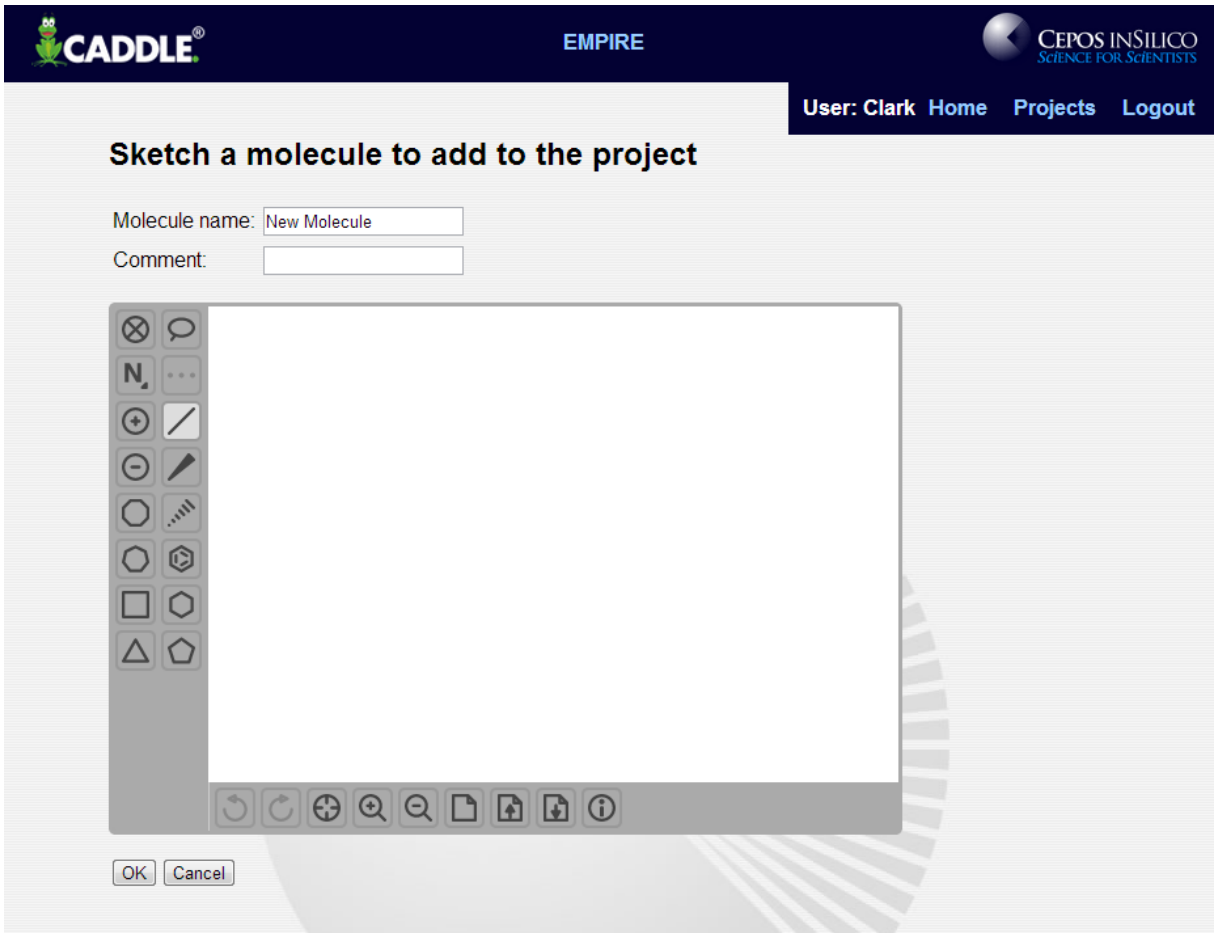


Figure 6 The CADDLE.® 2D sketcher

Using the sketcher is very intuitive. Ring systems can be added by selecting the appropriate symbol on the left hand side of the sketcher window, single bonds by selecting the bond symbol and then clicking

© Cepos InSilico Ltd. 2013

on the atom to which the bond is to be added and double bonds by selecting the bond symbol and clicking on an existing single bond. Further details on using the sketcher are provided in the online help.

When the molecule sketch is complete, click **OK** (**Figure 7**) to exit the sketcher and return to the molecule input page, where the sketched molecule is shown as a small 2D structure (see **Figure 8**). Further molecules can be sketched in the same way and they will be added to the list of sketched molecules. As for the other methods of adding molecules to the ***project***, clicking the **Save Molecules to Project** button loads the sketched molecule(s) into CADDLE.®.

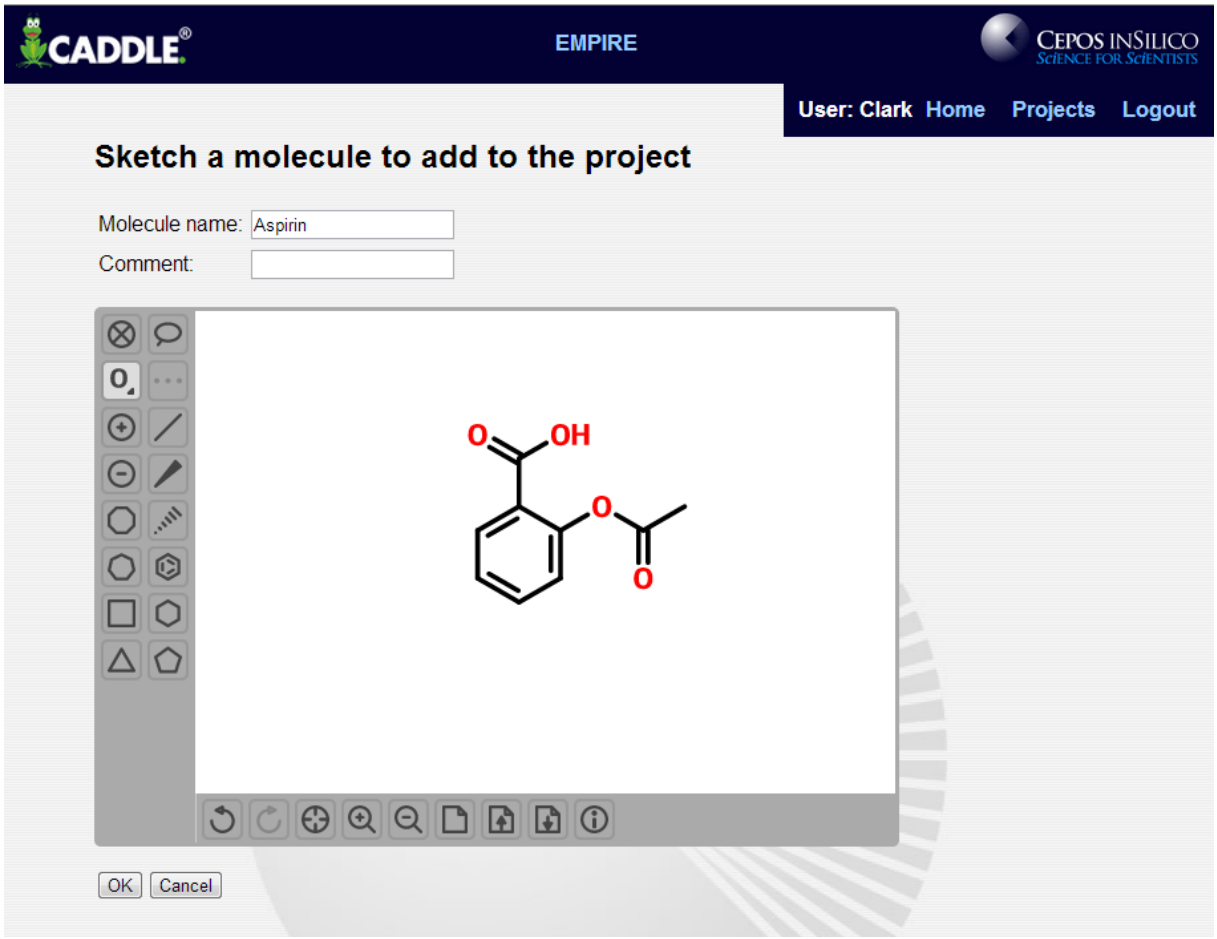


Figure 7 A molecular structure created using the CADDLE.® 2D sketcher

© Cepos InSilico Ltd. 2013

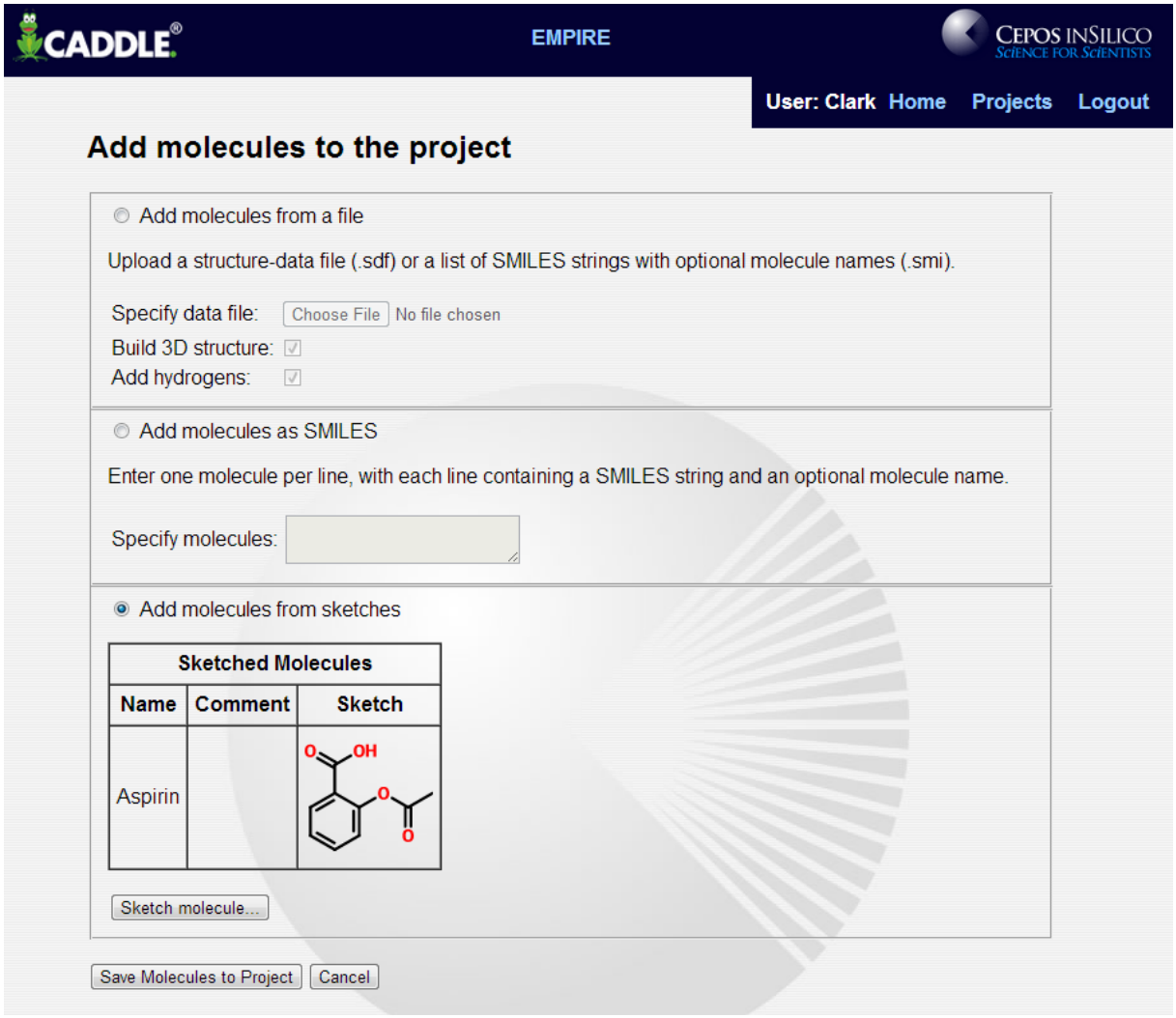


Figure 8 The CADDLE.® molecule input page with the molecule added from the 2D sketcher

3.5 Viewing input structures

Once the molecules have been added to the project, their input structures can be viewed using the **View molecules** link on the project details page (**Figure 4**). The input structures are then displayed in the **Molecule Browser**, shown in **Figure 9**.

© Cepos InSilico Ltd. 2013

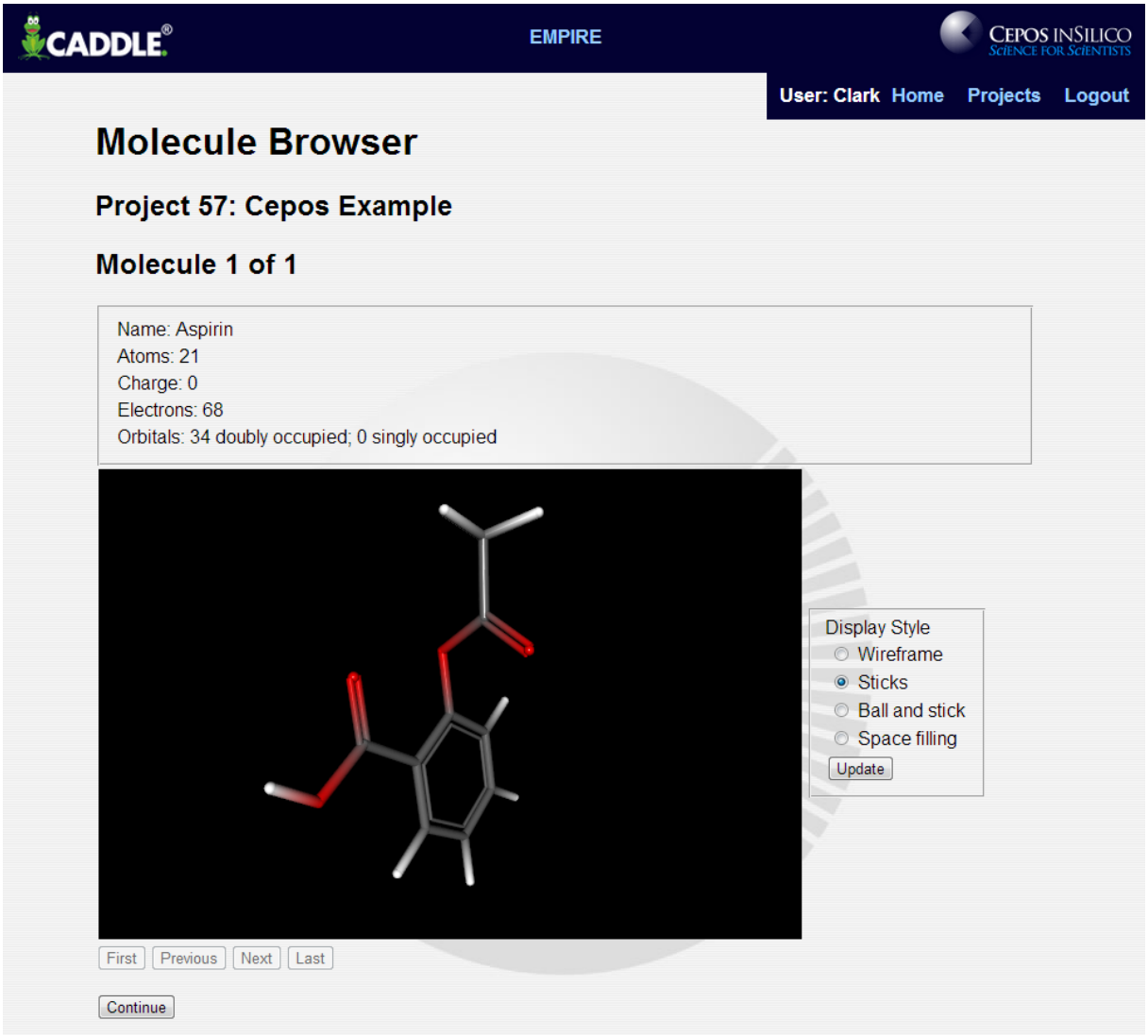


Figure 9 The CADDLE.® molecule browser

The structures of the molecules can be viewed in the interactive 3D viewer window and the structures rotated, shifted or zoomed using the mouse button combinations shown in **Table 1**.

Table 1 Mouse commands for the molecule viewer in common web browsers. For an up-to-date summary of suitable browsers, see the **test web page**

Action	Chrome	Firefox	Internet Explorer	Safari
Rotate	Left button	Left button	Left button	Left button
Translate	Central button or <ctrl> right button	<ctrl> right button	<ctrl> right button	<ctrl> right button
Zoom	Mouse wheel or right button	Mouse wheel or right button	Mouse wheel or right button	Right button

Molecules can be displayed as wireframe, stick, ball and stick or space-filling depictions by selecting the appropriate option and then clicking **Update**.

© Cepos InSilico Ltd. 2013

3.6 Accessing an existing project

Existing ***projects*** may be accessed from the list of recent projects on the **Home** page, or by selecting the **Projects** link on your personal menu bar, which produces a list of all your projects, together with projects made publicly available by other users (see **Figure 10**). From the project list the input structures of a project's molecules may be downloaded as an SDF file and the project may be deleted by its owner.

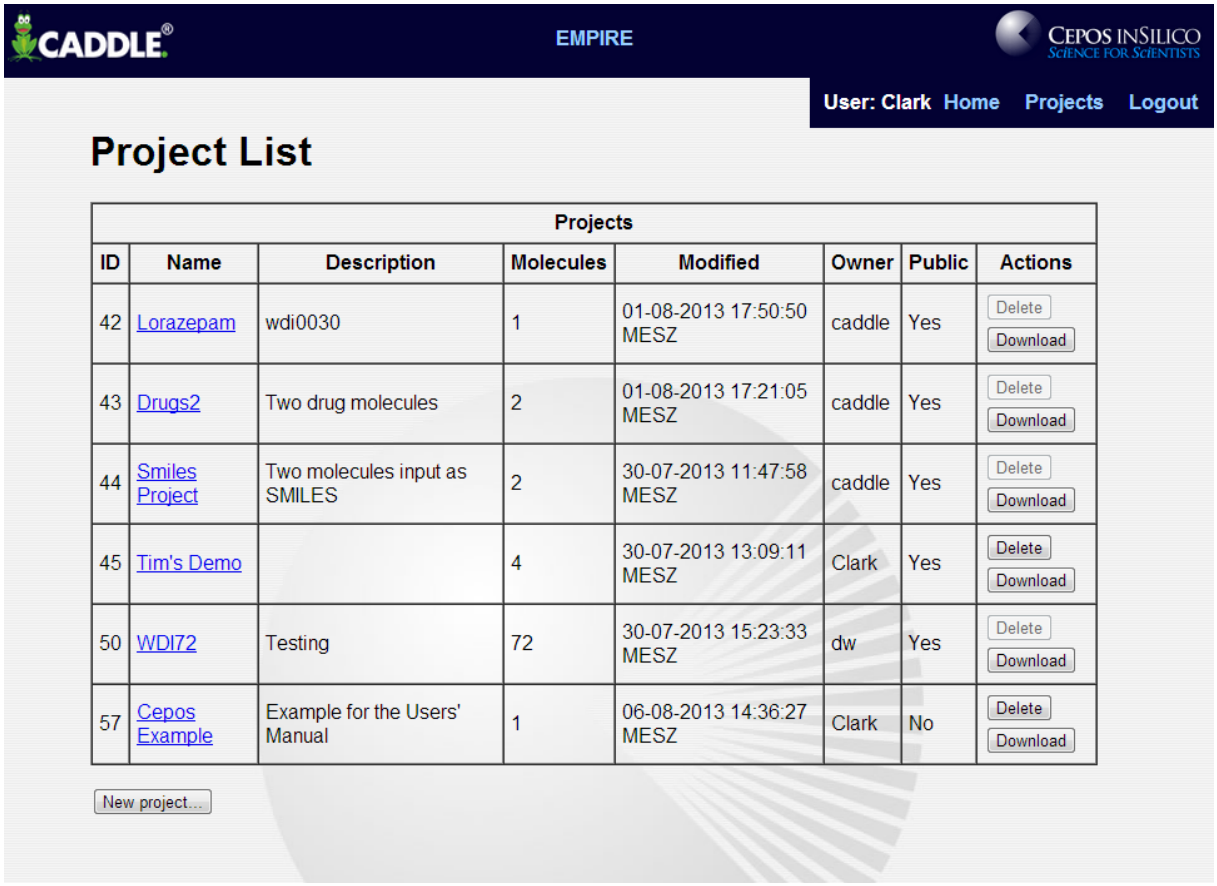


Figure 10 The CADDLE.® project browser

To select a project, click on its name in the project list. This will display the **Project Details** page (**Figure 4**), which allows the project owner to edit the project details, add molecules to the project, view the project molecules and access previous calculations.

3.7 Editing project details

The **Edit Details** button on the **Project Details** page produces the **Project Editor** shown in **Figure 11**. This allows the project name, description and visibility to be changed. Specifying a project to be ***Public*** grants read-only access to the project to all CADDLE.® users. A project that is not ***Public*** is visible only to the user who created it.

© Cepos InSilico Ltd. 2013

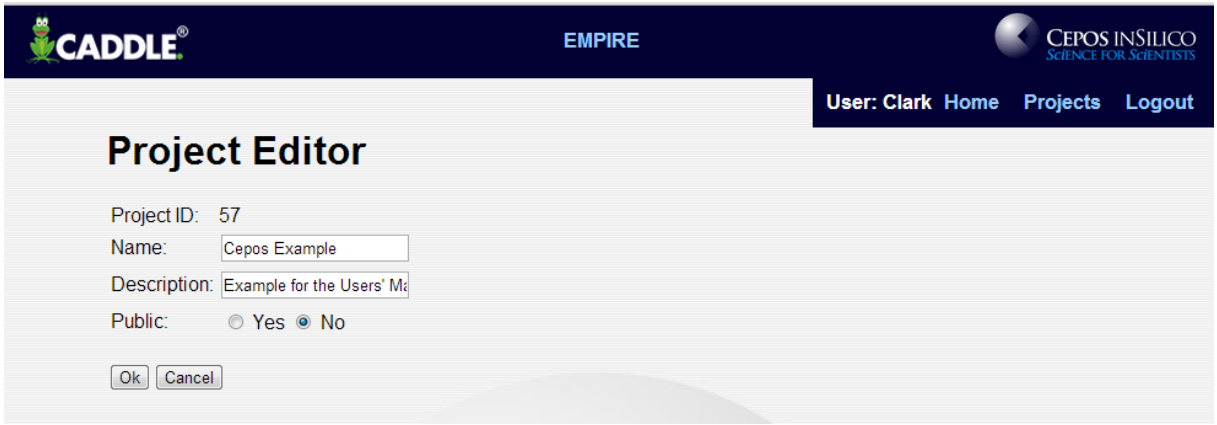


Figure 11 The CADDLE.® project editor

3.8 Viewing previous calculations

The **Calculation History** button on the **Project Details** page retrieves a list of previous calculations for the project (see **Figure 12**). To display the results of a specific calculation, click the calculation's ID in the first column of the table. The calculation list also allows the output files for a calculation to be downloaded or a calculation to be deleted.

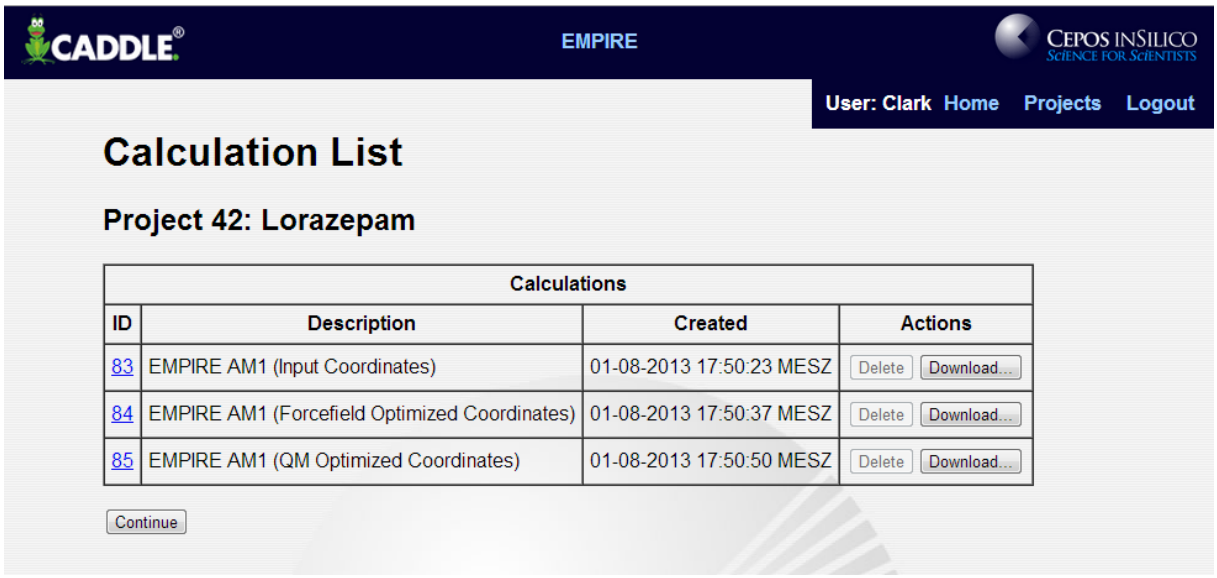


Figure 12 The CADDLE.® project calculation list

© Cepos InSilico Ltd. 2013

4 EMPIRE™ CALCULATIONS

EMPIRE™ is Cepos InSilico's new parallel semiempirical molecular orbital [1] program. To run a calculation, select the **EMPIRE** link in the main CADDLE.® menu to display the project selection page (**Figure 13**). Select a project for the calculation from the drop-down list and click **Next** to move to the EMPIRE control page shown in **Figure 14**.

Figure 13 The project selection page

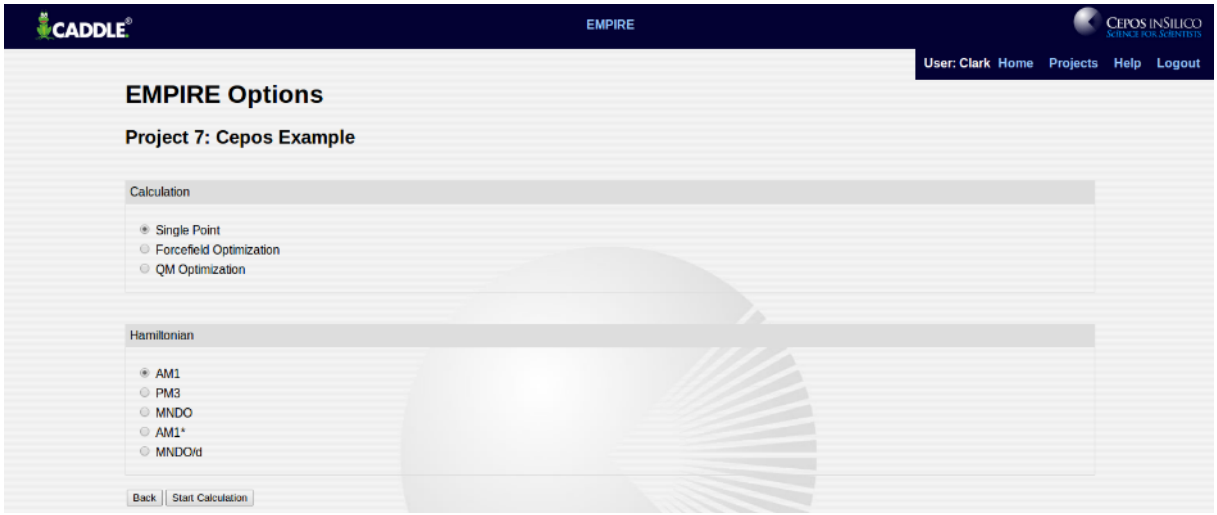


Figure 14 The EMPIRE™ control window

All molecules in the ***project*** are calculated in the same way, so that the **EMPIRE** options are limited to the choice of Hamiltonian (calculational method) to be used (currently AM1 [2], PM3 [3], MNDO [4], AM1* [5] and MNDO/d [6]) and the geometry to be used. Calculations can be performed on the input geometry without optimization (single point), on the optimized geometry obtained using the Merck Force Field [7] or the structures can be fully optimized using the semiempirical Hamiltonian.

Clicking the **Start calculation** button sends the calculations to the server. The results will appear on the **EMPIRE Browser** page automatically as soon as the calculations are finished. Details and options for the EMPIRE™ program are available in the **EMPIRE™ Users' Manual**.

© Cepos InSilico Ltd. 2013

5 VISUALIZATION

High quality visualization of the results of the calculations in a web browser is a major feature of CADDLE.®. The results from an EMPIRE™ calculation are displayed in an **EMPIRE Browser** page, shown in **Figure 15**.

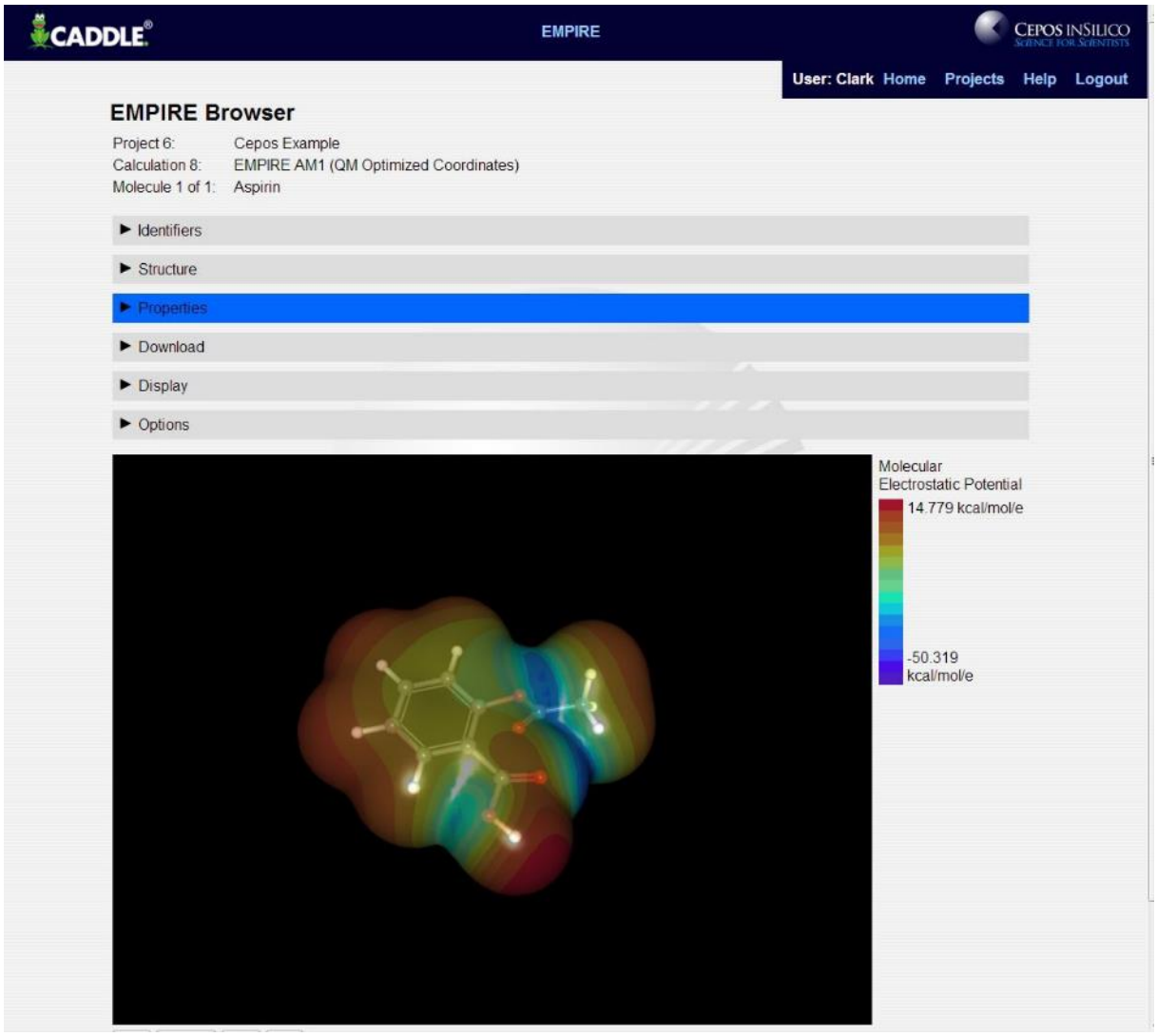


Figure 15 The EMPIRE™ browser window

Most of the information on the **EMPIRE™ Browser** page is contained in expanding sections whose contents are displayed when the section header bar is clicked. These sections include the molecule identifiers, a summary of the molecule's structural information and a list of properties calculated by EMPIRE™.

The **Download** section contains a list of output files from the EMPIRE™ calculation. These files can be downloaded to the user's computer in a zip archive by selecting the files required and clicking the **Download** button. The content of these files is described in the **EMPIRE™ Users' Manual**.

© Cepos InSilico Ltd. 2013

The **Display** section controls the contents of the interactive 3D viewer, which displays either a local property (the molecular electrostatic potential [8], the local ionization energy [9], the local electron affinity [10], electronegativity [10a] or hardness [10a]) projected onto an isodensity surface [11] for each molecule, or a molecular orbital. In each case the molecular structure may be displayed as a wireframe, sticks or ball-and-sticks model. The quality of the surface drawn can also be specified. It is recommended that, initially, low quality surfaces are drawn for local properties and high quality surfaces for molecular orbitals. Changes to the options in this section take effect when the **Update** button is clicked.

The **Options** section controls the size, transparency and background color of the 3D display. These options are applied immediately to the 3D viewer.

Figure 16 shows an orbital display for aspirin with the **Display** and **Options** sections expanded.

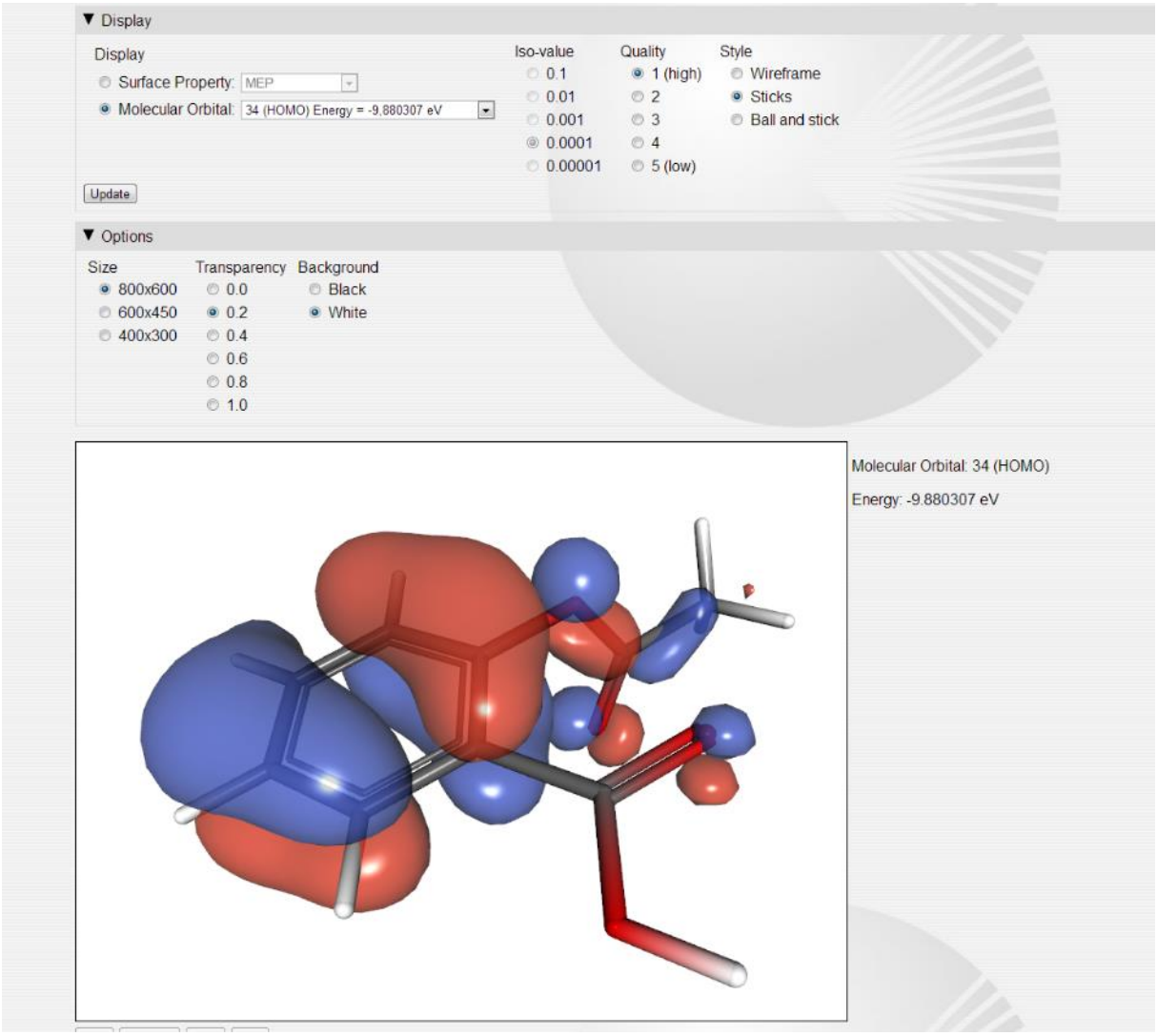


Figure 16 An orbital display demonstrating the use of the **Display** and **Options** facilities

© Cepos InSilico Ltd. 2013

6 SUPPORT

6.1 Contact

Questions regarding CADDLE.® should be sent directly to:

support@ceposinsilico.com

6.2 Cepos InSilico GmbH

Computer-Chemie-Centrum (CCC)
Nägelsbachstr. 25
91052 Erlangen
Germany

support@ceposinsilico.com
Tel. +49 (0)9131 970 4910
Fax. +49 (0)9131 970 4911

www.ceposinsilico.com/contact

© Cepos InSilico Ltd. 2013

7 LIST OF TABLES

© Cepos InSilico Ltd. 2013

8 LIST OF FIGURES

© Cepos InSilico Ltd. 2013

9 REFERENCES

[1] T. Clark, J. J. P. Stewart, *MNDO-like Semiempirical Molecular Orbital Theory and its Application to Large Systems, in Computational Methods for Large Systems, Vol. 8* Wiley, Chichester, **2011**.

[2] a) M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model, *Journal of the American Chemical Society* **1985**, *107*, 3902-3909. doi:10.1021/ja00299a024; b) A. J. Holder. AM1, in *Encyclopedia of Computational Chemistry, Vol. 1* (Eds.: P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer, III, P. R. Schreiner), Wiley, Chichester, **1998**, pp. 8-11

[3] a) J. J. P. Stewart. Optimization of parameters for semiempirical methods I. Method, *Journal of Computational Chemistry* **1989**, *10*, 209-220. doi:10.1002/jcc.540100208; b) J. J. Stewart. PM3, in *Encyclopedia of Computational Chemistry, Vol. 1* (Eds.: P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer, III, P. R. Schreiner), Wiley, Chichester, **1998**, pp. 2080-2086

[4] a) M. J. S. Dewar, W. Thiel. Ground states of molecules. 38. The MNDO method. Approximations and parameters, *Journal of the American Chemical Society* **1977**, *99*, 4899-4907; 4907-4917. doi:10.1021/ja00457a004; b) W. Thiel. in *Encyclopedia of Computational Chemistry, Vol. 1* (Eds.: P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer, III, P. R. Schreiner), Wiley, Chichester, **1998**, p. 1599

[5] a) P. Winget, A. C. Horn, C. Selçuki, B. Martin, T. Clark. AM1* parameters for phosphorus, sulfur and chlorine, *Journal of Molecular Modeling* **2003**, *9*, 408-414. doi:10.1007/s00894-003-0156-7; b) P. Winget, T. Clark. AM1* parameters for aluminum, silicon, titanium and zirconium, *Journal of Molecular Modeling* **2005**, *11*, 439-456. doi:10.1007/s00894-005-0236-y; c) H. Kayi, T. Clark. AM1* parameters for copper and zinc, *Journal of Molecular Modeling* **2007**, *13*, 965-979. doi:10.1007/s00894-007-0214-7; d) H. Kayi, T. Clark. AM1* parameters for bromine and iodine, *Journal of Molecular Modeling* **2009**, *15*, 295-308. doi:10.1007/s00894-008-0419-4; e) H. Kayi, T. Clark. AM1* parameters for vanadium and chromium, *Journal of Molecular Modeling* **2009**, *15*, 1253-1269. doi:10.1007/s00894-009-0489-y; f) H. Kayi, T. Clark. AM1* parameters for cobalt and nickel, *Journal of Molecular Modeling* **2010**, *16*, 29-47. doi:10.1007/s00894-009-0503-4; g) H. Kayi, T. Clark. AM1* parameters for manganese and iron, *Journal of Molecular Modeling* **2010**, *16*, 1109-1126. doi:10.1007/s00894-009-0614-y; h) H. Kayi, T. Clark. AM1* parameters for palladium and silver, *Journal of Molecular Modeling* **2011**, *17*, 2585-2600. doi:10.1007/s00894-010-0940-0

[6] a) W. Thiel, A. A. Voityuk. Extension of the MNDO formalism to d orbitals: integral approximations and preliminary numerical results, *Theor. Chim. Acta* **1992**, *81*, 391-404; b) W. Thiel, A. A. Voityuk. Extension of MNDO to d orbitals: parameters and results for the halogens, *Int. J. Quant. Chem.* **1996**, *93*; c) W. Thiel, A. A. Voityuk. Extension of MNDO to d orbitals: Parameters and results for the halogens, *Int. J. Quant. Chem.* **1992**, *44*, 807-829; d) W. Thiel, A. A. Voityuk. Extension of MNDO to d orbitals: parameters and results for the silicon, *J. Mol. Struct.* **1994**, *313*, 141-154; e) W. Thiel, A. A. Voityuk. Extension of MNDO to d Orbitals: Parameters and Results for the Second-Row Elements and for the Zinc Group, *The Journal of Physical Chemistry* **1996**, *100*, 616-626. doi:10.1021/jp952148o; f) W. Thiel. in *Encyclopedia of Computational Chemistry, Vol. 3* (Eds.: P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer, III, P. R. Schreiner), Wiley, Chichester, **1998**, pp. 1604-1606

[7] a) T. A. Halgren. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94, *J. Comput. Chem.* **1996**, *17*, 490-519; b) T. A. Halgren. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, *J. Comput. Chem.* **1996**, *17*, 520-552; c) T. A. Halgren. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94, *J. Comput. Chem.* **1996**, *17*, 553–586; d) T. A. Halgren. Merck molecular force field. IV. conformational energies and geometries for MMFF94, *J. Comput. Chem.* **1996**, *17*, 587-615; e) T. A. Halgren. Merck

© Cepos InSilico Ltd. 2013

molecular force field. V. Extension of MMFF94 using experimental data, *J. Comput. Chem.* **1996**, *17*, 616-641

[8] *Chemical Applications of Atomic and Molecular Electrostatic Potentials. Reactivity, Structure, Scattering, and Energetics of Organic, Inorganic, and Biological Systems*, Plenum Press, New York, NY, **1981**.

[9] P. Sjoberg, J. S. Murray, T. Brinck, P. Politzer. Average local ionisation energies on the molecular surfaces of aromatic systems as guides to chemical reactivity, *Can. J. Chem.* **1990**, *68*, 1440-1443

[10] a) B. Ehresmann, B. Martin, A. H. Horn, T. Clark. Local molecular properties and their use in predicting reactivity, *Journal of Molecular Modeling* **2003**, *9*, 342-347. doi:10.1007/s00894-003-0153-x; b) T. Clark. The local electron affinity for non-minimal basis sets, *Journal of Molecular Modeling* **2010**, *16*, 1231-1238. doi:10.1007/s00894-009-0607-x

[11] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory* Oxford University Press, Oxford, **1990**.

© Cepos InSilico Ltd. 2013