Pipeline Pilot Cresset Package V2.6.0 User Guide

About Cresset Package for Pipeline Pilot

The Cresset Package incorporates key Cresset applications into Pipeline Pilot. Each component wraps the corresponding command line binary making it available within Pipeline Pilot.

The Cresset Package uses our proprietary [XED molecular mechanics force field](#) and gives realistic, low energy conformations and accurate molecular fields. However, the XED force field is optimized for small drug-like molecules and hence is sub-optimal for tasks that lie away from the study of bioactivity. In particular, support for non-organic elements is limited and transition states are not supported.

Forge™

[Forge](#) is Cresset's powerful computational suite for understanding Structure-Activity Relationship (SAR) and molecule design. Current functionality in Forge includes ligand-based alignment of multiple molecules to one or more reference compounds, 3D-QSAR model building using either Field QSAR or Machine Learning methods of regression and classification, activity cliff analysis with Activity Miner™, qualitative SAR analysis with Activity Atlas™, building of detailed field pharmacophores with FieldTemplater™, design of new molecules with the ideal activity and property profile, virtual screening of thousands of compounds.

Forge is available with both a Graphical User Interface (GUI) and as a series of command line binaries which are incorporated into the Pipeline Pilot components in this package.

The Forge components can be found under **Components > Chemistry > Cresset > Forge** and examples on how to use the components can be found under **Protocols > Examples > Chemistry > Cresset > Forge**.

Forge Align is a tool to align multiple molecules to a reference molecule in a predefined conformation. Molecules for alignment are not limited by 2D similarity to the reference molecule. When used on a congeneric series, the tool can help in library design and give a rationale for the prioritization of compounds for synthesis. Using Forge Align on a diverse set of active molecules can help define the requirements of the protein of interest, aiding the synthetic chemist in the design of new actives. Forge Align will read molecules from the input stream either as a series of conformations (such as the output of the XedeX component) or as flat representations that require conformational analysis. For each molecule in the input stream the best alignment to the reference molecule is sent to the output. Optionally, field and pharmacophore constraints and/or a protein structure (to be used as an excluded volume) can be specified to bias the alignment algorithm and penalize results which do not satisfy the constraint.

Activity Miner is designed to compare a dataset of compounds and output a matrix of similarities, distances or disparities for the molecules. In contrast to other similar methods the Activity Miner component can use either 2D or 3D metrics to calculate the matrix enabling the detection of activity cliffs in a biological context. Using Activity Miner highlights both the activity cliff and the reason for the observed change in activity.

Forge Build is a tool for generating 3D-QSAR models for activity from a set of aligned molecules. It determines a set of sampling positions around the molecules, based on their field points, which can be used to probe any molecule for the electrostatic potential or for the volume taken up by molecules. The data matrix is then processed by Partial Least Squares (PLS) or Machine Learning methods to derive an equation that describes activity or category. Optionally the data matrix can be exported for use with alternative data analysis methods.

Forge Score is a tool for calculating predicted activities or categories for molecules against a 3D-QSAR model. The molecules must be pre-aligned before scoring against the model.

The **Forge Model Viewer** (on Client) component is used to view the model generated by the Forge Build component. Forge needs to be installed and licensed on the client machine for this component to work.

FieldTemplater is a tool for comparing molecules using their electrostatic and hydrophobic fields to find common patterns. When applied to several structurally-distinct molecules with a common activity, FieldTemplater can determine the bioactive conformations and relative alignments of these molecules in the absence of protein information.

Forge Surface Writer generates the Positive, Negative, Van der Waals and Hydrophobic surfaces for the input molecules and saves them to a directory. The surfaces can be written in CCP4, Cube, Insight or MOE format.

Spark™

[Spark](#) finds biologically equivalent replacements for key moieties in your molecule and guides you through the process of growing and linking ligands. Given a 'starter molecule', preferably in the bioactive conformation, you can select a portion to replace and specify which databases to search. Spark will then present a list of biologically relevant replacements ranked using [Cresset's unique molecular field technology](#), or using [Lead Finder](#)™'s docking score. As well as replacing central parts of a molecule, Spark can suggest replacements for terminal groups. It can be used to grow ligands and fragments into unoccupied pockets of the target protein, carry out ligand joining and macrocyclization experiment, and to find fragments which displace a crystallographic water molecule near your ligand. Unlike other fragment replacement methodologies, Spark works in product space so all potential replacements are merged into the original starter molecule before scoring, leading to more diverse yet realistic results.

Spark is supplied with several databases of moieties generated from fragmentation of commercially-available compounds, active molecules reported in the literature, available reagents, and theoretical studies. An optional database generator module enables the creation of custom fragment databases.

Spark is available with both a GUI and as a series of command line binaries which are incorporated into the Pipeline Pilot components in this package.

The Spark components can be found under **Components > Chemistry > Cresset > Spark** and examples on how to use the components can be found under **Protocols > Examples > Chemistry > Cresset > Spark**.

Spark Database Search is used to find bioisosteric replacements for a region of the input starter molecule selected by the user.

The section of the molecule to be replaced can be specified by either using the 'Spark Fragment Selector (on Client)', a tag on the input molecule, or as a parameter for this component. The list of fragment databases to search can be specified by either using the 'Visual Spark Database Selector', a tag on the input molecule, or as a parameter for this component. A full list of databases can be viewed by using the 'Spark Database List Viewer'.

Spark result molecules can be scored by ligand similarity or docking into a protein structure. Optionally, field and pharmacophore constraints, docking constraints and/or a protein structure (to be used as an excluded volume with ligand similarity scoring) can be specified to bias the Spark search

and penalize results which do not satisfy the constraints. Additional reference molecules can be used to guide the experiment.

Spark Database Search (on Client) is the same as the Spark Database Search component except the section of the molecule to be replaced and the databases to search do not need to be specified. Instead the component will temporarily stop the running protocol and will ask the user to select the section of the molecule to be replaced using a 2D image of the molecule. It will also ask which databases should be searched. Spark needs to be installed and licensed on the client machine for this component to work.

Spark Database List Viewer shows which Spark databases have been installed on the Pipeline Pilot Server.

Generate Spark Database is used to create or update a Spark fragment database from a set of molecules. The database can then be used by the 'Spark Database Search' component.

Flare™

Flare is a structure-based design tool for the visualization and analysis of protein-ligand complexes.

Current capabilities in Flare include:

- Smart import and handling of protein-ligand complexes
- Accurate and reliable protein preparation
- Sequence alignment and superimposition
- Molecular Dynamics using OpenMM
- Minimization of protein ligand complexes
- Accurate docking using Lead Finder™
- Ligand-based conformation hunt and alignment
- Free Energy Perturbation (FEP) simulations
- Creation of custom parameters for the AMBER/GAFF and AMBER/GAFF2 force fields
- Protein interaction potentials
- Electrostatic complementarity™
- 3D-RISM water analysis
- WaterSwap analysis for ligand and binding site energetics
- The Flare Python™ API

Flare is available as a graphical user interface while the Flare Python API provides access to all the Flare functionality.

The Flare components can be found under **Components > Chemistry > Cresset > Flare** and examples on how to use the components can be found under **Protocols > Examples > Chemistry > Cresset > Flare**.

pyflare runs a Python script which has access to all of Flare functionality through the Flare Python API. The documentation for the Flare Python API can be found in the Flare installation directory.

Blaze™

Blaze is Cresset's ligand based virtual screening technology. Given an active molecule in its bioactive conformation, Blaze is used to search through databases of millions of molecules to find those that have the most similar electrostatic and shape. Blaze has extensive compound management features

that enable the control of compound collections and pre-population of conformations for all uploaded molecules.

Blaze can be accessed using a web browser or a REST API with the server located internally or in the cloud.

Components to access Blaze's functionally through the REST API can be found under **Components > Chemistry > Blaze**. To use the Blaze components, you will need an account for your internal Blaze server or for Cresset's software as a service offering of Blaze. To evaluate the Blaze functionality you can connect to the Blaze Demo server using a username and password from https://www.cresset-group.com/software/blaze-demo-server/.

New Blaze Search creates a new Blaze search.

Blaze Searches Viewer displays a brief overview of all your searches in Blaze.

Detailed Blaze Search Viewer displays a detailed report for a Blaze search.

Download Blaze Search Results downloads the results of a Blaze search.

Blaze Collections Viewer shows the Blaze collections which are available for searching.

Upload Molecules to Blaze Collection uploads molecules to a Blaze collection.

Blaze Status Viewer displays a report on the status of the Blaze server.

XedTools™

The [XedTools](#) components wrap key functionality of the XED force field and can be found in **Components > Chemistry > Cresset > XedTools**. Examples on how to use the components can be found in **Protocols > Examples > Chemistry > Cresset > XedTools**.

XedMin™ is a tool to minimize molecules using the [XED force field](#) from Cresset. XedMin uses a conjugate gradient to minimize the energy of a molecule using the XED molecular mechanics force field. All energies are calculated in vacuo. If the input structure is 2D or SMILES, then an initial conversion to 3D is undertaken before the minimization with the conjugate gradient minimizer. Optionally XedMin can be used to minimize just one fragment in a multi-fragment complex. However, XedMin can only process molecules with a maximum of 10,000 atoms (including all hydrogen and XED atoms) which roughly equates to a maximum of 2,500 heavy atoms. Optionally the XedMin component can load a protein structure (in mol2 format) that can be used as a context for the ligand minimization. Note that the 10,000-atom limit includes any protein structure that is loaded.

The **XedeX™** component provides an interface to the XedeX command line conformation generator. XedeX is a program to read in one or more molecules in multiple formats, perform a conformational hunt using the XED force field, and write the results to standard output. The conformational hunt process is designed to generate a highly diverse set of conformations, all of which are minimized, in as short a space of time as possible. XedeX is designed to find a representative set of minimum vacuum energy structures using the features of the XED force field. It is not designed to find all local minima over a given energy range. Rather, it has been designed to collect as broad a range of conformationally diverse structures as possible above the global minimum. This strategy is implemented deliberately to feed the requirements of the field overlay protocols (e.g. Forge Align) which need wide sampling of conformational space.

Cresset Molecule Writer converts molecules into Cresset's BXD file format. This can be used in Web Protocols to write out molecules to $(JobDir) directory as BXD files. Links to the molecules BXD files will then appear in the job result page, and when clicked the molecules will be opened in Forge.

Viewers

Viewer components can be found under **Components > Chemistry > Cresset > Viewers** and examples on how to use the components can be found under **Protocols > Examples > Chemistry > Cresset > Viewers**.

Cresset Molecule Viewer (on Client) is used to launch Forge for viewing molecules.

Flare Ligand and Protein Viewer (on Client) is used to launch Flare for viewing ligands and proteins.

Change Log

Cresset Package 2.6.0 for Pipeline Pilot 18.1: January 2021

Supports Flare 4.0, Forge 10.6, Spark 10.6.0, XedTools 3.2 and Blaze 10.3.

Enhancements to Existing Components

Components in **Components > Chemistry > Cresset** have been updated.

Cresset\Spark

Spark Database Search/Spark Database Search (on Client):

- New option to use docking instead of ligand similarity as the method for scoring the results molecules from a bioisostere replacement experiment in the Spark Database Search component.
- New option to set the docking region buffer size in the Spark Database Search component.
- New option to specify docking constraints in the Spark Database Search component.
- New option to set the maximum docking constraint penalty in the Spark Database Search component.
- Improved field and pharmacophore constraints, which can now be specified for both starter and reference molecules
- Improved fragment selection, now supporting the selection of a disconnected fragment in the starter molecule to perform and water replacement experiments.

New Protocol Examples

New example protocols have been added to **Protocols >Examples > Chemistry > Cresset**.

Cresset\Spark

Searches For Bioisosteric Replacements Using Spark Database Search with Docking (on Client):
Added example on how to use docking when performing a Spark search.

Cresset Package 2.5.0 for Pipeline Pilot 18.1: May 2019

Supports Forge 10.6, Spark 10.5.5, XedTools 3.2 and Blaze 10.2.

New Components

New components have been added to **Components > Chemistry > Cresset**.

Cresset\Forge

Forge Surface Writer: Generates the Positive, Negative, Van der Waals and Hydrophobic surfaces for molecules and writes them to a directory. The surfaces can be written in CCP4, Cube, Insight or MOE format.

Cresset\Flare

pyflare: Runs a Python script which has access to Flare functionality through the Flare Python API.

Cresset\Viewers

Flare Ligand and Protein Viewer (on Client): Displays ligands and proteins in Flare.

Enhancements to Existing Components

Components in **Components > Chemistry > Cresset** have been updated.

Cresset\Forge

Forge Align:

- Added option to perform substructure alignment with a SMARTS pattern.
- Added option to change the scoring metric to Tanimoto or Tversky.
- Added option to set pharmacophore constraints on the reference molecules.
- Added option to exclude conformations with boats.

Forge Build:

- Added support for building regression models using Machine Learning models: kNN, Random Forest, Relevance Vector Machine and Support Vector Machine.
- Added support for building classification models using Machine Learning models: kNN, Relevance Vector Machine and Support Vector Machine.
- New Weighted Sum Activity Cliffs Summary algorithm in Activity Atlas.
- Added option to export the surfaces of the Activity Atlas model.
- Added option to set which input tag contains the similarity data when building an Activity Atlas model.
- Added option to set the weighting scheme when building a kNN model.

Cresset\Spark

Spark Database Search/Spark Database Search (on Client):

- Added option to set field and pharmacophore constraints on the starter molecule.
- Added option to change the scoring metric to Tanimoto or Tversky.
- Added option to use multiple reference molecules to help guide the search.

New Protocol Examples

New example protocols have been added to **Protocols >Examples > Chemistry > Cresset**.

Cresset\Forge

Create surfaces for a molecule using Forge Surface Writer: example of how to use Forge Surface Writer to write molecule surfaces to a directory.

Several examples on how to use the new **Machine Learning regression and classification models**.

Cresset\Flare

A few examples on how to use the new **pyflare** component.

Cresset\Viewers

Flare Ligand and Protein Viewer: Example of how to use Flare Ligand and Protein Viewer (on Client) to view ligand and proteins in Flare.

Cresset Package 2.4.0 for Pipeline Pilot 9.0: October 2015

Supports Forge 10.4.2, Torch 10.4.2, Spark 10.3, XedTools 3.2 and Blaze 10.2.

New Components

New components have been added to **Components > Chemistry > Cresset**.

Cresset\XedTools

Cresset Molecule Writer: Writes an BXD file which can be opened in Forge or Torch directly from the Web Port.

Enhancements to Existing Components

Components in **Components > Chemistry > Cresset** have been updated.

Cresset\Forge

Forge Build: Added options for building Activity Atlas and k Nearest Neighbor (kNN) models.

Forge Score: Added support for Activity Atlas and k Nearest Neighbor (kNN) models.

Activity Miner: Added an option to output the results as a Forge project.

Cresset\Spark

Spark Database Search/Spark Database Search (on Client):

- Added an option to set the gradient cutoff.
- Added an option to output the results as a Spark project.

New Protocol Examples

New example protocols have been added to **Protocols >Examples > Chemistry > Cresset**.

Cresset\Forge

Generate an Activity Atlas Using Forge Build: Example of how to use Forge Build to create an Activity Atlas model.

Generate a k Nearest Neighbor (kNN) Model Using Forge Build: Example of how to use Forge Build to create a kNN model.

Calculate the Novelty of Molecules Using Forge Score with an Activity Atlas: Example of how to use Forge Score to fit molecules to an Activity Atlas model. The Activity Atlas model was created using Forge Build.

Calculate kNN Predicted Activities Using Forge Score: Example of how to use Forge Score to fit molecules to a kNN model. The kNN model was created using Forge Build.

New Web Port Protocol Examples

New example protocols have been added to **Protocols > Web Services > Web Port Examples > Chemistry > Cresset**.

Open Molecules in Forge or Torch: Example of how to create a Web Port Protocol which contains links which opens molecules in Forge or Torch.

Spark Database Search: Example of how to perform a Spark Database Search using the Web Port.

Cresset Package 2.3.0 for Pipeline Pilot 9.0: December 2014

Supports Forge 10.3, Torch 10.3, Spark 10.3, XedTools 3.1 and Blaze 10.2.

Enhancements to Existing Components

Components in **Components > Chemistry > Cresset** have been updated.

Cresset\Forge

Forge Align: Added advanced conformation hunt options.

FieldTemplater: Added advanced conformation hunt options.

Cresset\Spark

Spark Database Search/Spark Database Search (on Client):

- Added support for the Cresset Engine Broker.
- Added a filter for thioethers.

Generate Spark Database:

- Added support for the Cresset Engine Broker.
- Added an option to generate all reagent databases in one invocation.
- Added advanced conformation hunt options.

Cresset Package 2.2.0 for Pipeline Pilot 9.0: August 2014

Supports Forge 10.3, Torch 10.3, Spark 10.2, XedTools 3.1 and Blaze 10.2.

New Components

New components have been added to **Components > Chemistry > Cresset**.

Cresset\Forge

FieldTemplater: Create a pharmacophore from a set of molecules.

Cresset\Spark

Generate Spark Database: Creates a Spark fragment database from a set of molecules.

New Protocol Examples

New example protocols have been added to **Protocols > Examples > Chemistry > Cresset**.

Cresset\Forge

Create a pharmacophore using FieldTemplater: Example of how to use "FieldTemplater" to create a pharmacophore.

Cresset\Spark

Generate a Spark fragment database: Example of how to use "Generate Spark Database" to create a Spark fragment database.

Enhancements to Existing Components

Components in **Components > Chemistry > Cresset** have been updated.

Cresset\Forge

Forge Align:

- The conformation hunt can now be disabled.
- Added support for the **Cresset Engine Broker.** Refer to the Administration Guide for details on how to setup the Cresset Engine Broker.

Forge Build: Added support for the Cresset Engine Broker.

Activity Miner: Added support for the Cresset Engine Broker.

Cresset\Spark

Spark Database Search: The "Database" parameter now lists the databases available to search.

Cresset Package 2.1.0 for Pipeline Pilot 9.0: May 2014

Supports Forge 10.2, Torch 10.2, Spark 10.2, XedTools 3.0 and Blaze 10.2.

New Components

A new folder has been added under **Components > Chemistry > Cresset**. The new folder is named **Blaze** and contains the following new components.

New Blaze Search: Search chemical collections for molecules with similar properties to the input molecules.

Blaze Searches Viewer: Shows all the searches in Blaze.

Detailed Blaze Search Viewer: View detailed information on a Blaze search.

Download Blaze Search Results: Download the results of a Blaze search.

Upload Molecules to Blaze Collection: Uploads molecules to a chemical collection in Blaze.

Blaze Collections Viewer: Lists the collections in Blaze.

Blaze Status Viewer: View the status of the Blaze server and any running jobs.

Cresset Package 2.0.0 for Pipeline Pilot 9.0: January 2014

Supports Forge 10.2, Torch 10.2, Spark 10.2 and XedTools 3.0.

New Components

A new folder has been added under **Components > Chemistry**. The new folder is named Cresset and contains the following new components.

Cresset\Forge

Activity Miner: Generates a similarity or disparity matrix from a set of aligned molecules.

Forge Align: Align a set of molecules to reference molecules.

Forge Build: Generates a 3D QSAR model for activity from a set of aligned molecules and outputs the model as a Forge project.

Forge Model Viewer (on Client): Opens a Forge model created by Forge Build on the client in either Forge or Torch.

Forge Score: Calculates the predicted activities for molecules against a Forge 3D QSAR model.

Cresset\Spark

Spark Database Search: Searches the Spark databases for fragments to replace the specified section in the starter molecule.

Spark Database Search (on Client): Allows a user to select a section of the starter molecule to replace and searches the Spark databases for fragments to replace the section with.

Spark Database List Viewer: List all databases available to Spark.

Cresset\Spark\Utilities

Spark Fragment Selector (on Client): Allows a user to temporarily stop a protocol and select the region of the molecule to replace.

Visual Spark Database Selector: Allows a user to temporarily stop a protocol and select the databases Spark should search.

Cresset\Viewers

Cresset Molecule Viewer (on Client): Displays the molecule in Forge or Torch.

Cresset\XedTools

XedeX: Generate conformers for the input molecules.

XedMin: Minimise the input molecules.

New Protocol Examples

A new folder has been added under **Protocols > Examples > Chemistry**. The new folder is named **Cresset** and contains the following new protocol examples.

Cresset\Forge

Align Molecules Using Forge Align: Example of how to use Forge Align to align a set of molecules to a reference.

Calculate Predicted Activities Using Forge Score: Example of how to use Forge Score to fit molecules to a 3D QSAR model. The 3D QSAR model was created using Forge Build.

Generate a 3D QSAR Model For Activity Using Forge Build: Example of how to use Forge Build to create a 3D QSAR model for activity set of molecules.

Generate Disparity Matrix Using Activity Miner: Example of how to use Activity Miner to generate a disparity matrix from aligned molecules with activity data and display the results in a report.

Generate Similarity Matrix Using Activity Miner: Example of how to use Activity Miner to generate a similarity matrix from aligned molecules with activity data and display the results in a report.

Cresset\Spark

List Spark Databases: Example of how to use Spark Database List Viewer to view the available Spark databases.

Searches For Bioisosteric Replacements Using Spark Database Search: Example of how to use Spark Database Search to replace the specified section in the starter molecule.

Searches For Bioisosteric Replacements Using Spark Database Search (on Client): Example of how to use Spark Database Search (on Client) to replace the specified section in the starter molecule.

Cresset\XedTools

Generate Conformers Using XedeX: Example of how to use XedeX to generate conformers for a set of molecules.

Minimise Molecules Using XedMin: Example of how to use XedMin to minimise a set of molecules.

Support

Cresset	+44 (0)1223 858890
New Cambridge House	support@cresset-group.com
Bassingbourn Road	www.cresset-group.com
Litlington	Twitter: @cressetgroup
Cambridgeshire, SG8 0SS, UK	LinkedIn: Cresset group