User's guide to SPECTCOL

Document Information

Editors:	Y. A. Ba, L. Nenadović
Author(s):	Y. A. BA, L. Nenadović
Contributor(s):	
Type of document:	Software documentation
Status:	release
Distribution:	public
Work package:	WP4 SUP@VAMDC
Version:	V12.07-r1
Document code:	
Directory and file name:	http://www.vamdc.org/documents/software/SPECTCOL_guide_v12.07-r1.pdf

Abstract: This document is a guide to the tool for the manipulation of VAMDC-XSAMS formatted data for the purpose of extracting and merging Einstein and rate coefficients from different sources.

Version History

Version	Date	Modified By	Description of Change
V11.09	01/06/11	Ljerka Nenadović	first draft
V12.07-r1	12/08/13	Yaye Awa Ba	

Disclaimer

All rights reserved

Acknowledgements

VAMDC is funded under the "Combination of Collaborative Projects and Coordination and Support Actions" Funding Scheme of The Seventh Framework Program. Call topic: INFRA-2008-1.2.2 Scientific Data Infrastructure. Grant Agreement number: 239108.

Table of Contents

Introduction

This document is a guide to the SPECTCOL tool for the manipulation of VAMDC-XSAMS formatted data for the purpose of extracting and merging Einstein and rate coefficients from different sources. For more information on the XML schema used visit the VAMDC Standards Documentation page at http://www.vamdc.org/documents/standards/

Configuration

The `database.properties` file contains configuration properties necessary to connect to the registry and query the databases. It comes with pre-set default values but the advanced users can update these when needed. The properties are as follows:

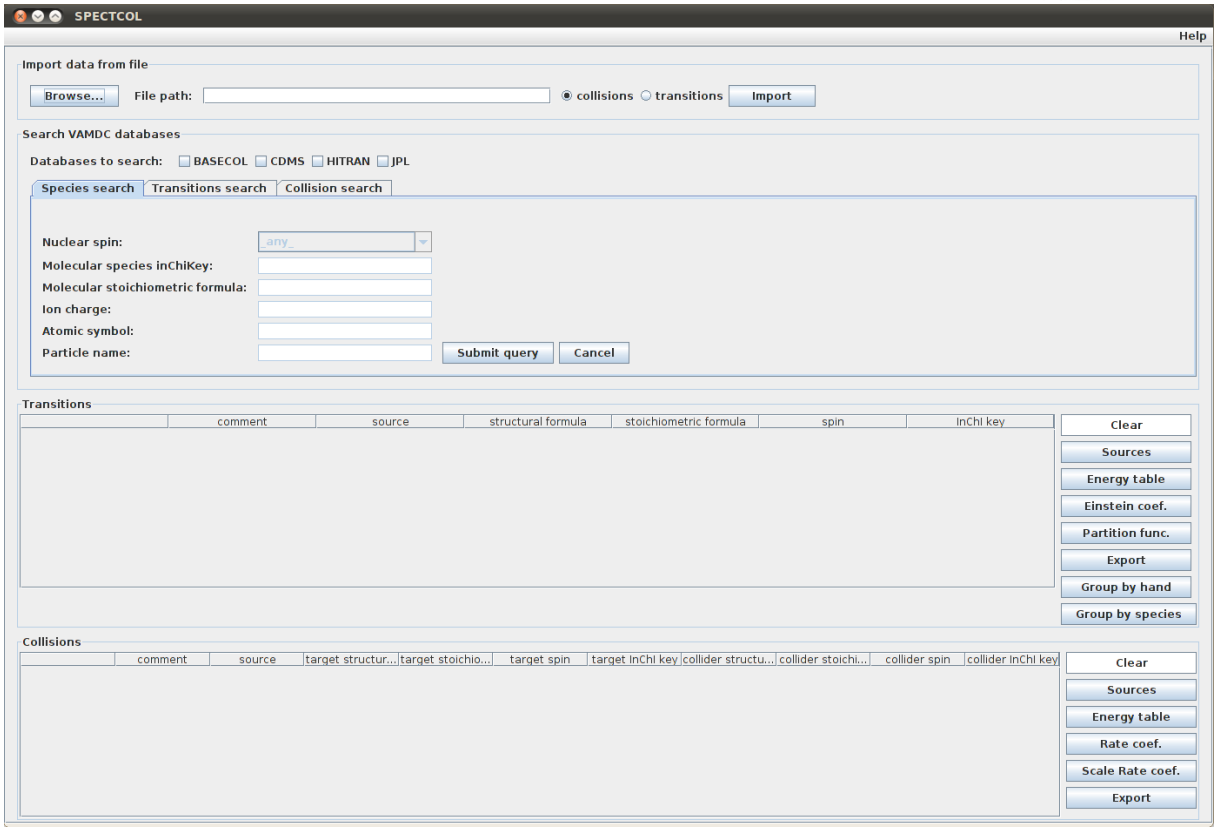
- `registry.endpoint` – The URL of the registry endpoint
- `<DATABASENAME>.identifier` – the URL based identifier of the resource in the following form:

  `ivo://{authorityid}/{resourcekey}`

  For the values of registry keys for relevant databases consult the registry. For more information about registry identifiers visit http://www.ivoa.net/Documents/REC/Identifiers/Identifiers-20070302.html

The Main Panel

The main panel consists of four sections: import, search, transitions and collisions summary tables with control buttons.



Import

Used for importing data from VAMDC-XSAMS files on hard drive or for importing VAMDC-XSAMS query URL.

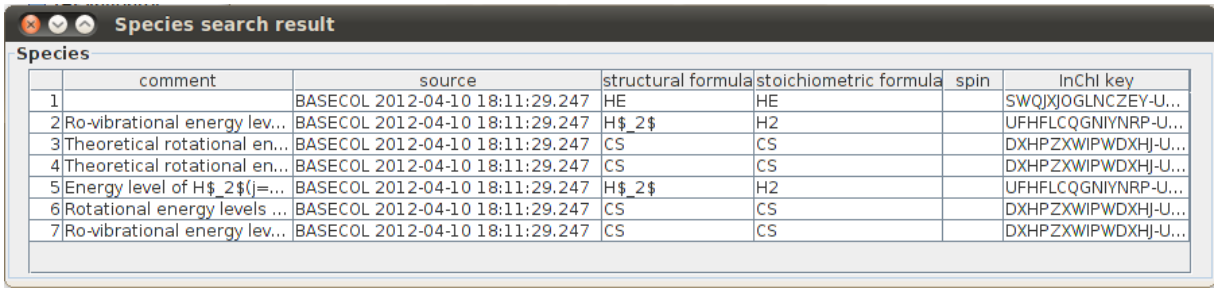
Search

Used to search selected databases on the fly based on the few specific keywords.

The search consists of database selection check boxes and is divided into three tabs: Species, Transitions and Collisions. These correspond to different types of searches.

Species search tab

The species search is intended as a preview of species available in the databases. It will return only the species information, without any states and processes data. The search results will pop up in a new window shown below.



	comment	source	structural formula	stoichiometric formula	spin	InChI key
1		BASECOL 2012-04-10 18:11:29.247	HE	HE		SWQJXJOGLNCZEY-U...
2	Ro-vibrational energy lev...	BASECOL 2012-04-10 18:11:29.247	H$_2$	H2		UFHFLCQGNIYNRP-U...
3	Theoretical rotational en...	BASECOL 2012-04-10 18:11:29.247	CS	CS		DXHPZXWIPWDXHJ-U...
4	Theoretical rotational en...	BASECOL 2012-04-10 18:11:29.247	CS	CS		DXHPZXWIPWDXHJ-U...
5	Energy level of H$_2$(j=...	BASECOL 2012-04-10 18:11:29.247	H$_2$	H2		UFHFLCQGNIYNRP-U...
6	Rotational energy levels ...	BASECOL 2012-04-10 18:11:29.247	CS	CS		DXHPZXWIPWDXHJ-U...
7	Ro-vibrational energy lev...	BASECOL 2012-04-10 18:11:29.247	CS	CS		DXHPZXWIPWDXHJ-U...

Transitions search tab

The transitions search is used to search for radiative transitions. The results of this search are grouped by the transition's initial state species reference. If the data returned from the database contains no transitions, the results are grouped by the available species information. The summary of the results is displayed in the Transitions Summary table.

Collisions search tab

This tab is used to search for collisions. A target restriction is required and the collider restriction is optional. The results of this search are grouped by collision sets. If the data returned form the database contains no collisions, the results are grouped by the available species information. The summary of the results is displayed in the Collisions Summary table.

Note: *With every new search, the results will be appended to the data already loaded. The amount of data typically returned by these queries is very large. It is advised that the user occasionally clears the search results using the Clear button.*

Control buttons

The buttons available for both summary tables are:

- **Clear** – clears selected rows described in the summary table.
  - To select a bloc of rows: select the first row of the bloc + click on “shift” button (maintain the click) + select the last row of the bloc.
  - To select non consecutive rows: select a row + click on “ctrl or pom for Mac” button (maintain the click) + select the other rows.
- **Get sources** – displays in a table format the sources of data relevant to the selected molecule/atom from the summary table.
- **Get energy table** – displays in a table format the energy levels and quantum numbers relevant to the selected molecule/atom from the summary table.
- **Export** – makes a subset of the original data based on the molecule selected in the summary table and exports it as VAMDC-XSAMS formatted XML file, to a user specified location.

Transitions buttons:

- **Get Einstein coefficients** – displays in a table format the Einstein coefficients relevant to the selected molecule from the transitions summary table, if such data is available.
- **Get Partition functions** – displays in a table format the partition functions relevant to the selected molecule from the transitions summary table, if such data is available.

- **Group by species** – based on the selected molecule/atom in the transitions summary table groups all species from species with the same InChiKey and presents the information in a new panel from which the user can perform further data extraction and merging.

- **Group by hand** – groups all species present in transitions summary table , collisions summary table and scaled collisions summary tables; and displays the information in a new panel from which the user can perform further data extraction and merging.

Collisions buttons:

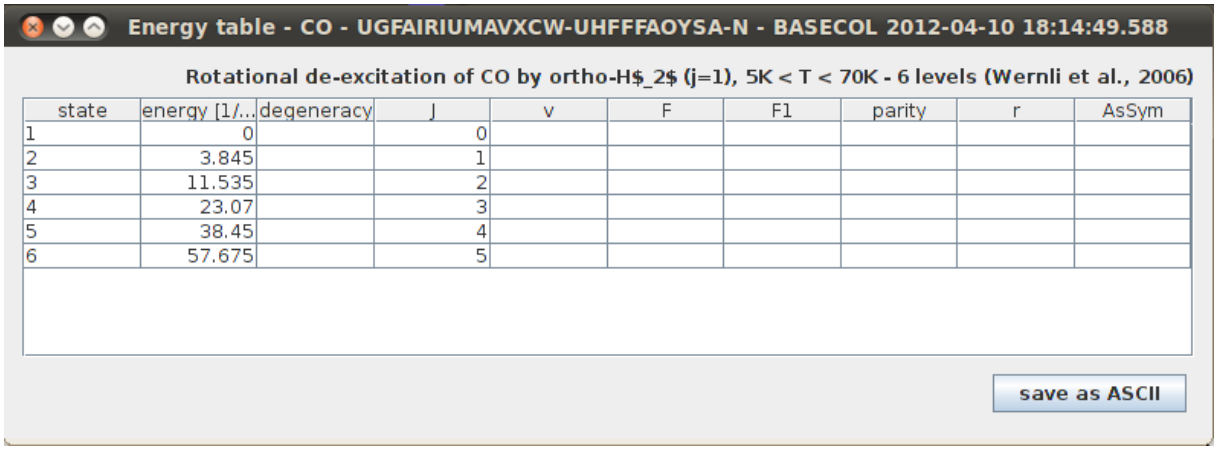
- **Get rate coefficients** – displays in a table format the rate coefficients relevant to the selected collisions set from the collisions summary table, if such data is available.

- **Scale rate coefficients** – displays in a table format the scaled rate coefficients relevant to the selected collisions and entered factor. It displays also the information in scaled collisions summary table.

Tabulated Data

Energy table

The energy table window displays the state energy and quantum number information for the selected species. The quantum numbers available vary depending on the selected species. For definition of quantum numbers and more information on cases see the VAMDC-XSAMS documentation.

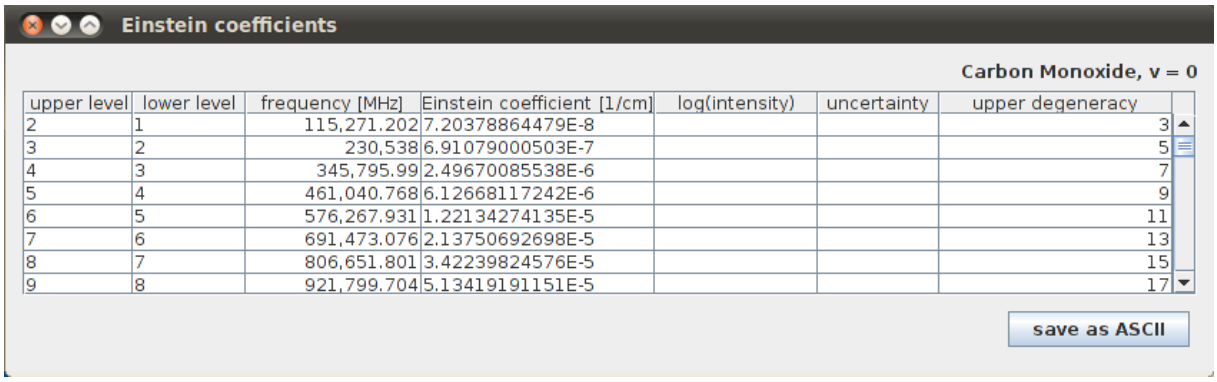


Einstein coefficients table

The Einstein coefficient data is displayed in the panel illustrated below. The columns of the table are:

- **upper level** – id of the initial state of a radiative transition. It can be used to identify the energy and quantum numbers from the energy table.
- **lower level** – id of the final state of a radiative transition. It can be used to identify the energy and quantum numbers from the energy table.
- **frequency**
- **Einstein coefficient**
- **log(Intensity)**
- **uncertainty**
- **upper state degeneracy**

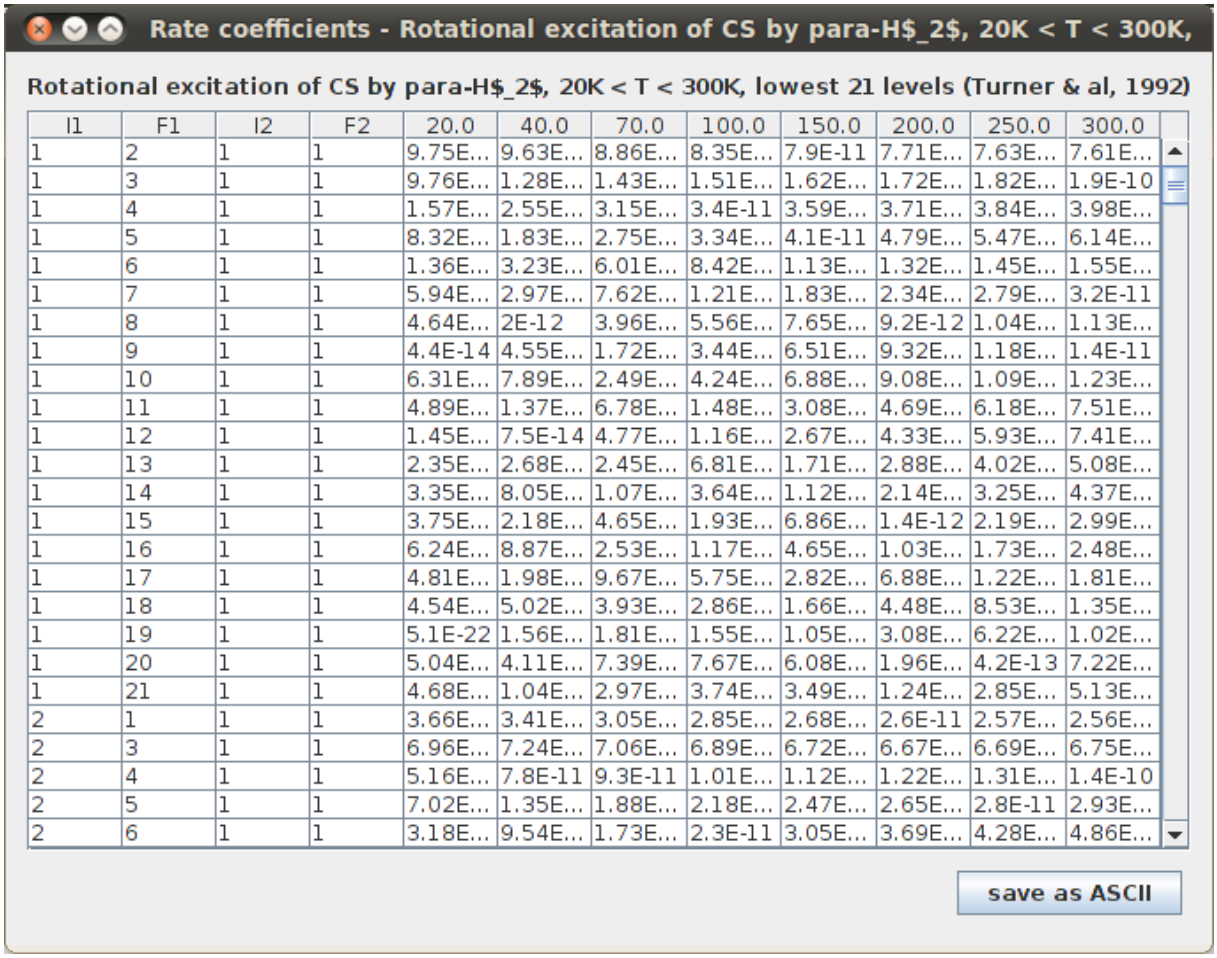
This information can be exported into a text file with comma separated values by clicking the "Save as ASCII" button.



Rate coefficients table

The rate coefficient data is displayed in the panel illustrated below.

This information can be exported into a text file with comma separated values by clicking the "Save as ASCII" button.

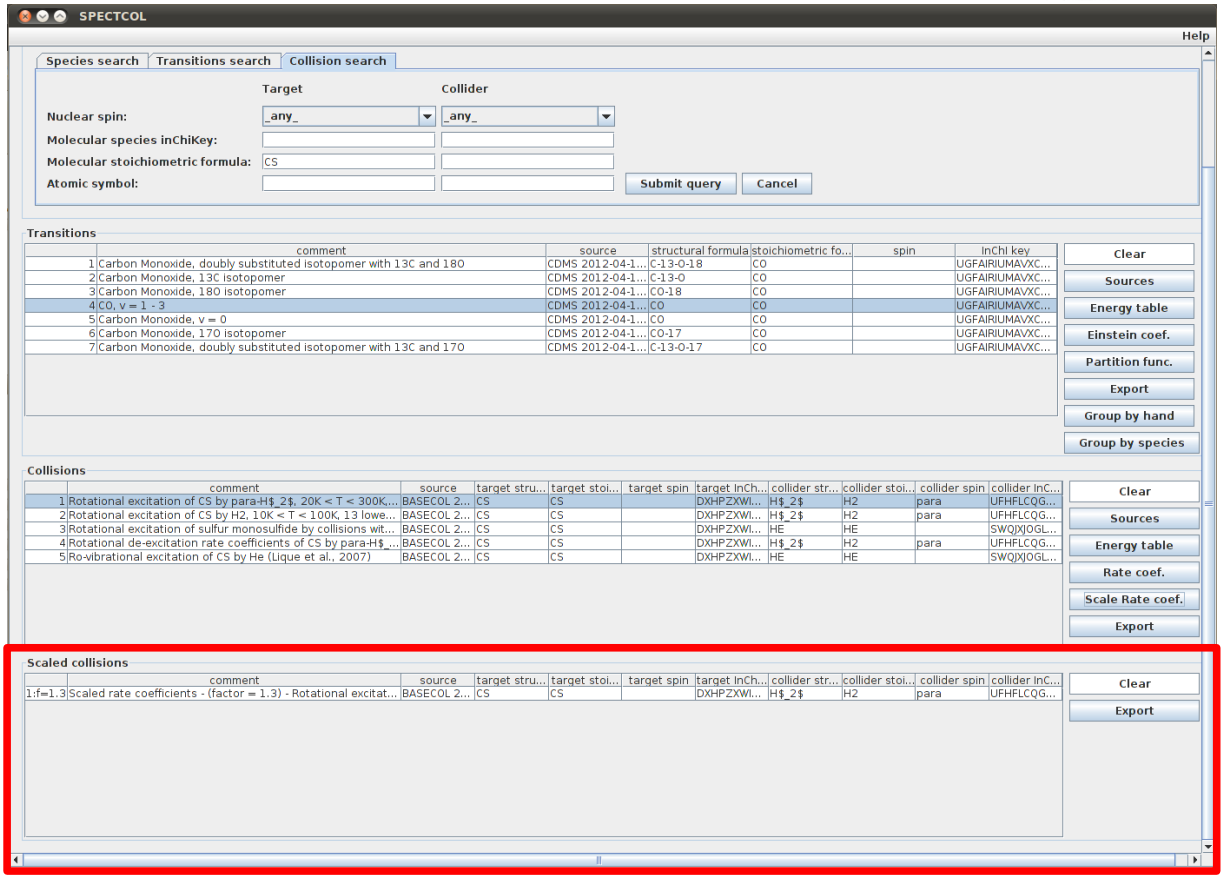


I1	F1	I2	F2	20.0	40.0	70.0	100.0	150.0	200.0	250.0	300.0
1	2	1	1	9.75E...	9.63E...	8.86E...	8.35E...	7.9E-11	7.71E...	7.63E...	7.61E...
1	3	1	1	9.76E...	1.28E...	1.43E...	1.51E...	1.62E...	1.72E...	1.82E...	1.9E-10
1	4	1	1	1.57E...	2.55E...	3.15E...	3.4E-11	3.59E...	3.71E...	3.84E...	3.98E...
1	5	1	1	8.32E...	1.83E...	2.75E...	3.34E...	4.1E-11	4.79E...	5.47E...	6.14E...
1	6	1	1	1.36E...	3.23E...	6.01E...	8.42E...	1.13E...	1.32E...	1.45E...	1.55E...
1	7	1	1	5.94E...	2.97E...	7.62E...	1.21E...	1.83E...	2.34E...	2.79E...	3.2E-11
1	8	1	1	4.64E...	2E-12	3.96E...	5.56E...	7.65E...	9.2E-12	1.04E...	1.13E...
1	9	1	1	4.4E-14	4.55E...	1.72E...	3.44E...	6.51E...	9.32E...	1.18E...	1.4E-11
1	10	1	1	6.31E...	7.89E...	2.49E...	4.24E...	6.88E...	9.08E...	1.09E...	1.23E...
1	11	1	1	4.89E...	1.37E...	6.78E...	1.48E...	3.08E...	4.69E...	6.18E...	7.51E...
1	12	1	1	1.45E...	7.5E-14	4.77E...	1.16E...	2.67E...	4.33E...	5.93E...	7.41E...
1	13	1	1	2.35E...	2.68E...	2.45E...	6.81E...	1.71E...	2.88E...	4.02E...	5.08E...
1	14	1	1	3.35E...	8.05E...	1.07E...	3.64E...	1.12E...	2.14E...	3.25E...	4.37E...
1	15	1	1	3.75E...	2.18E...	4.65E...	1.93E...	6.86E...	1.4E-12	2.19E...	2.99E...
1	16	1	1	6.24E...	8.87E...	2.53E...	1.17E...	4.65E...	1.03E...	1.73E...	2.48E...
1	17	1	1	4.81E...	1.98E...	9.67E...	5.75E...	2.82E...	6.88E...	1.22E...	1.81E...
1	18	1	1	4.54E...	5.02E...	3.93E...	2.86E...	1.66E...	4.48E...	8.53E...	1.35E...
1	19	1	1	5.1E-22	1.56E...	1.81E...	1.55E...	1.05E...	3.08E...	6.22E...	1.02E...
1	20	1	1	5.04E...	4.11E...	7.39E...	7.67E...	6.08E...	1.96E...	4.2E-13	7.22E...
1	21	1	1	4.68E...	1.04E...	2.97E...	3.74E...	3.49E...	1.24E...	2.85E...	5.13E...
2	1	1	1	3.66E...	3.41E...	3.05E...	2.85E...	2.68E...	2.6E-11	2.57E...	2.56E...
2	3	1	1	6.96E...	7.24E...	7.06E...	6.89E...	6.72E...	6.67E...	6.69E...	6.75E...
2	4	1	1	5.16E...	7.8E-11	9.3E-11	1.01E...	1.12E...	1.22E...	1.31E...	1.4E-10
2	5	1	1	7.02E...	1.35E...	1.88E...	2.18E...	2.47E...	2.65E...	2.8E-11	2.93E...
2	6	1	1	3.18E...	9.54E...	1.73E...	2.3E-11	3.05E...	3.69E...	4.28E...	4.86E...

Scaled rate coefficients

The scaled rate coefficients data is displayed as rate coefficients, the only differences are the title and the data of the temperatures which are multiplied by the entered factor.

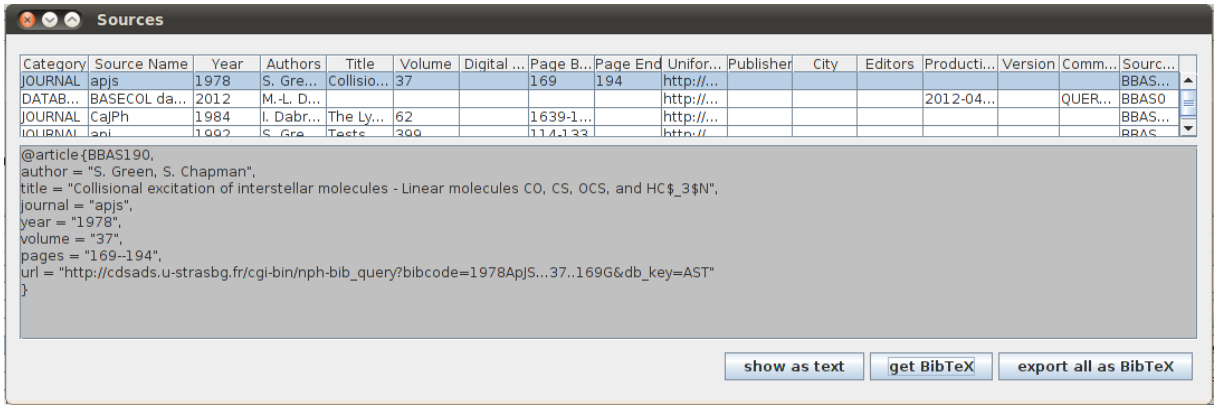
In the main panel of SPECTCOL, Scaled collisions summary table appears as below and is shown only if you use “Scaled Rate coeff.” button in collisions summary table.



Sources table

The source information is displayed in the panel illustrated below.

- “show as text” button displays in the text area data from the selected row as a formatted string.
- “get BibTeX” button displays in the text area the BibTeX data for the selected row if it is present in the VAMDC-XSAMS file. If no pre-formatted BibTeX data is available, a BibTeX entry can be constructed from the available fields.
- “export all as BibTeX” button exports all rows of the table in a single .bib file. Pre-formatted BibTeX data is used when available, otherwise a BibTeX record is generated based on the xsams source entry fields.



Importing VAMDC-XSAMS files or query URL

Importing existing VAMDC-XSAMS files

The application allows the user to import files already saved on their computer. The files must be correctly formatted XML files which conform to the VAMDC-XSAMS standard. For the more information about the VAMDC-XSAMS standard see http://www.vamdc.org/documents/standards/ .

To import an existing file:

1. click the “Browse” button
2. select the file of interest
3. select "transitions" or "collisions" to specify how to treat the imported data
4. click “Open”. This will fill the file path information.
5. click the “Import” button

Alternatively you can type the file path information manually.

Importing VAMDC-XSAMS query URL

The application allows the user to import URLs (VAMDC-XSAMS query URL). The URLs must lead to correctly formatted VAMDC-XSAMS documents.

To import an URL:

1. type or copy the URL in the file path information
2. select "transitions" or "collisions" to specify how to treat the imported data
3. click the “Import” button

Searching VAMDC databases

The application can be used to search selected VAMDC databases directly.

1. Select the databases to search by selecting the appropriate check boxes. Selecting a database will automatically disable the keyword options not available for the selected database. If several databases are selected only the keywords which can be used for all selected databases are available.
2. Fill in one of the keyword values for your search. If several fields are filled in only the topmost keyword will be used for the search.
3. Click “Submit Query”. Search results will appear in the summary table

Grouping data

By species

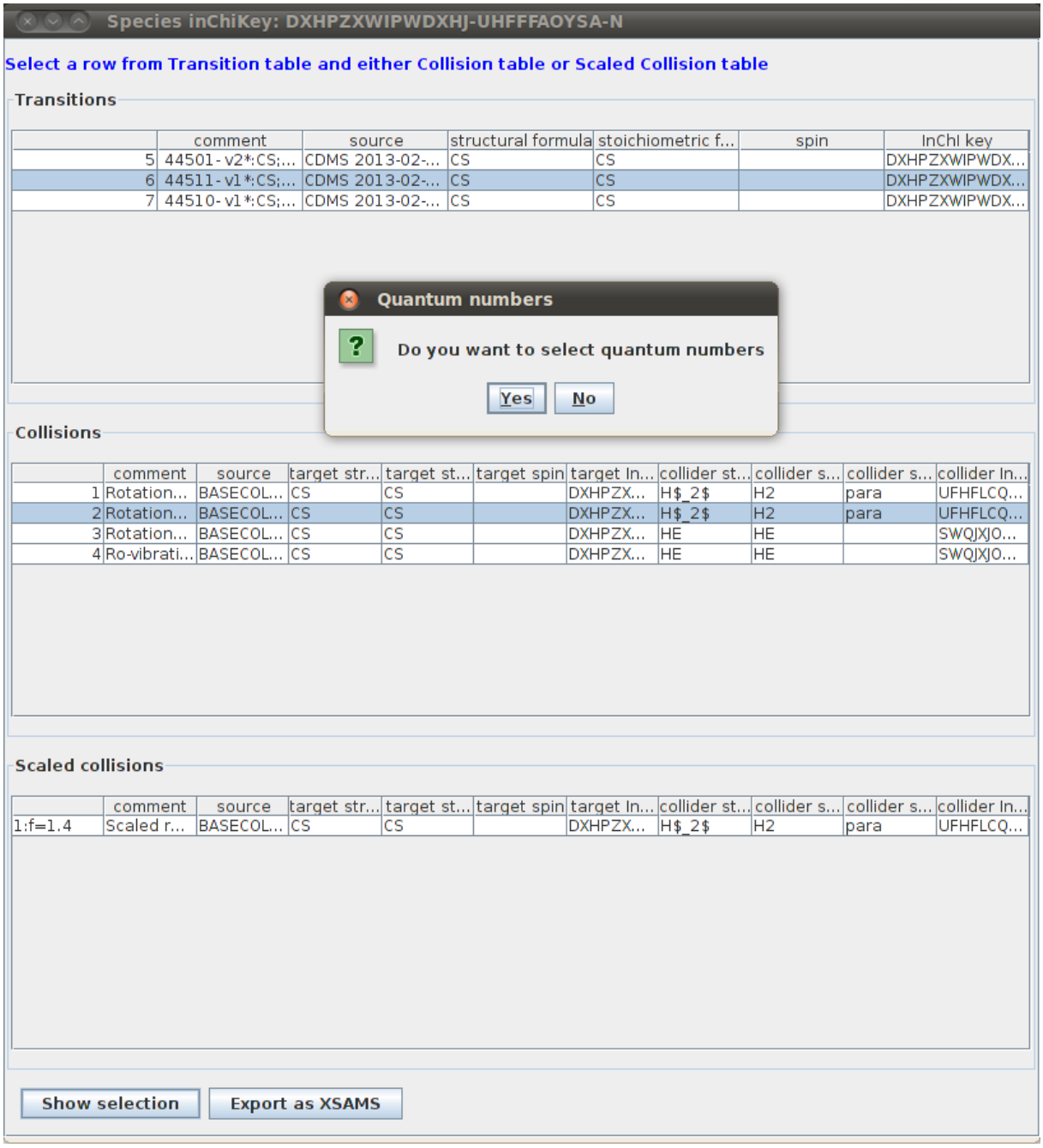
By selecting a row in the Transitions Summary table you identify the species you wish to use as a reference for data grouping. Clicking the "Group species" button groups the loaded data based on the selected species' inChiKey. For transitions this means all the information concerning species with the same inChiKey. For collisions and scaled collisions, this implies that the target element involved in the collision has the same inChiKey.

This is the first step in identifying what data can be merged.

By hand

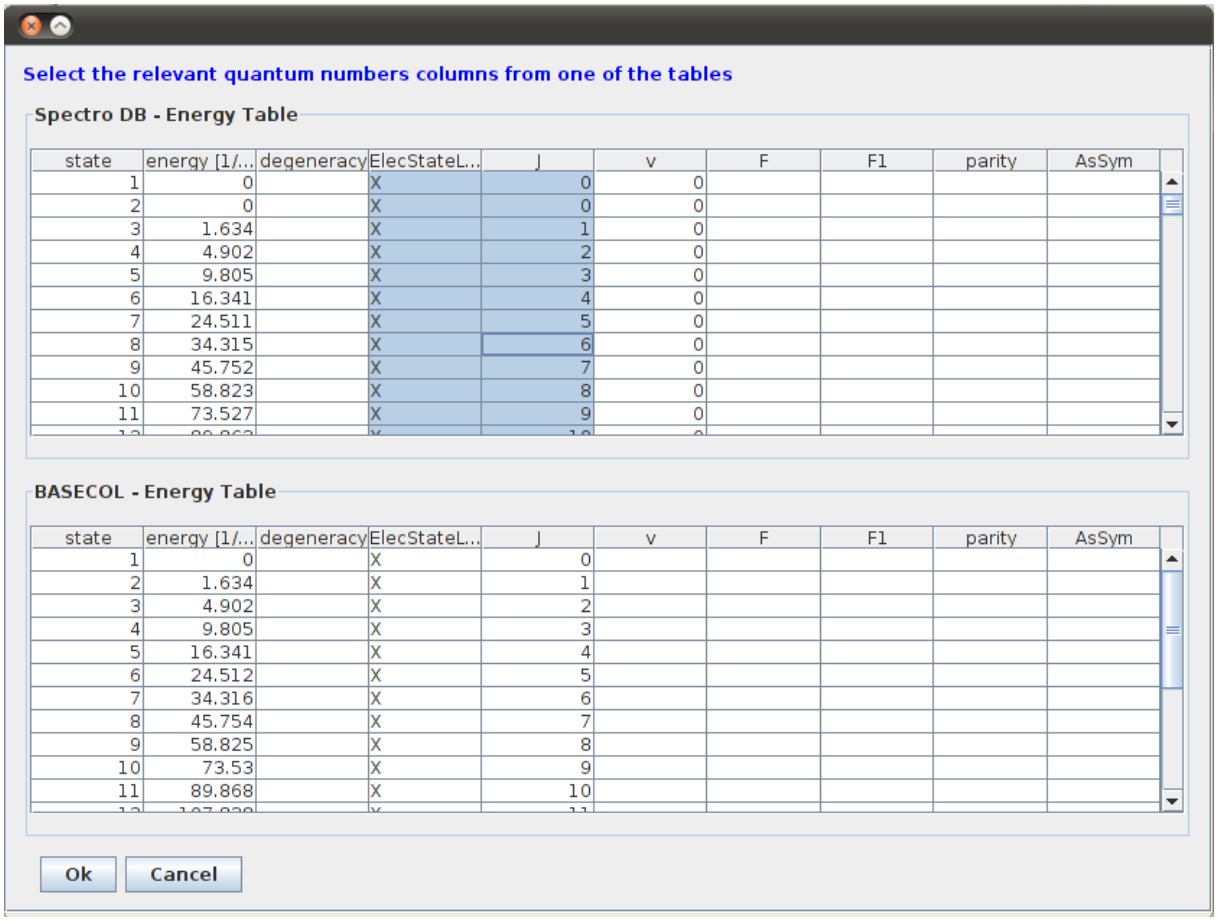
Without selecting a row in Transitions Summary table. Clicking the "Group by hand" button, groups all data present in Transitions summary table, Collisions Summary table and Scaled Collisions Summary table.

The selection is displayed in the Group Panel, shown below.



By selecting a row from Transitions table and either Collisions table or Scaled Collisions table, the user identifies two sets of data he wishes to merge. Sometimes VAMDC-XSAMS from two sources can be different even if it contains information about the same species, same cases and so on, because we can fill the schema in multiple different ways. So after “Show selection” another dialog window appears, allowing to choose if you want to select Quantum Numbers or not. Choosing “YES” option opens the Panel with the set of Quantum numbers available in the schema, as shown below.

The data can either be exported to file in VAMDC-XSAMS format, or viewed in the GUI.



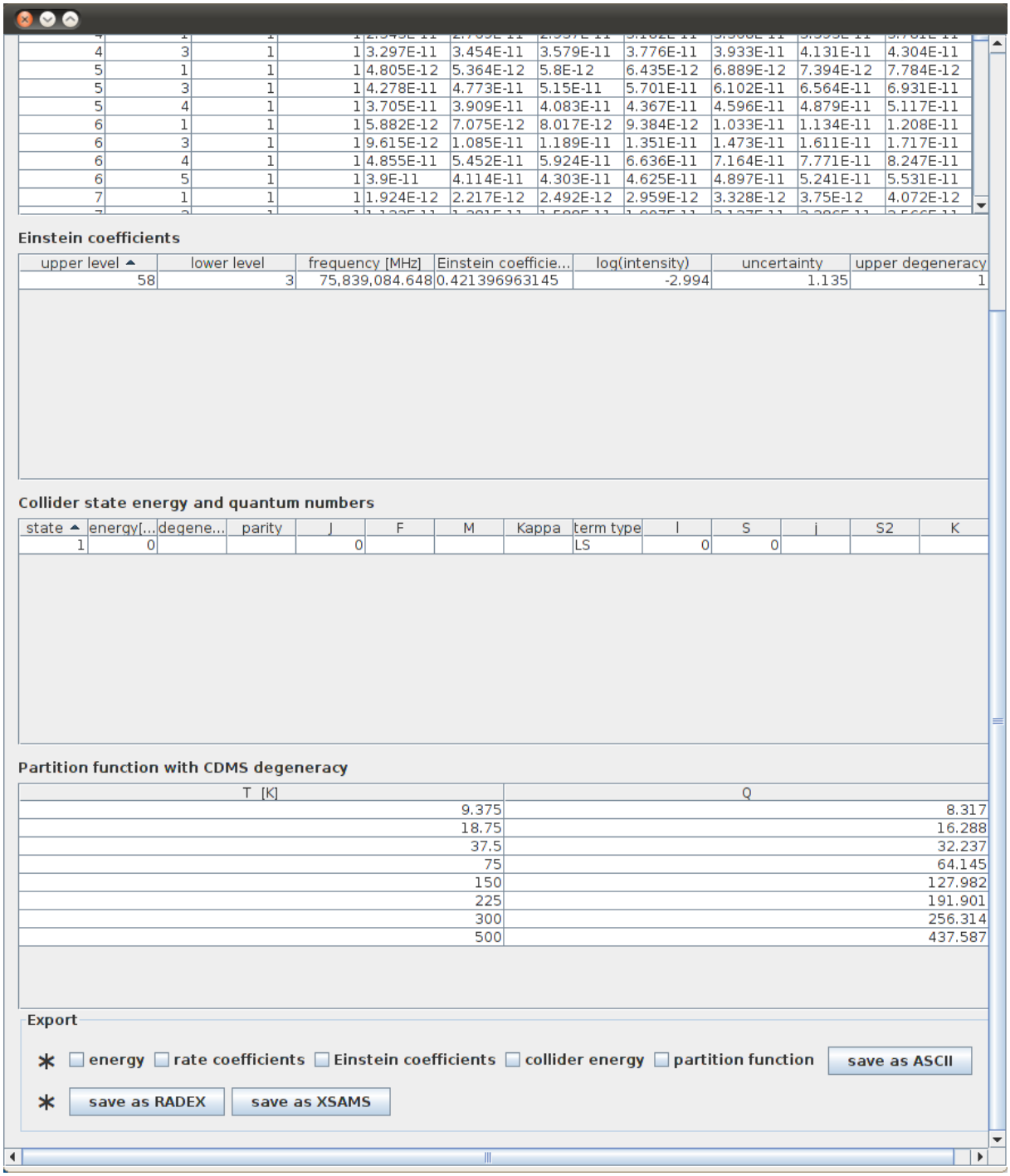
You can select one ore more quantum numbers you want to keep in the schema.

- When you click on one row – you select all the columns
- To select a block of quantum numbers – select first column + “Shift” button (maintain the click) + select the last column.
- To select non consecutive columns – select a column + “Ctrl or Pom in Mac OS” button (maintain the click) + select the other columns.

The merge is performed in the following way:

- the selection form the transitions table identifies the energy states
- for all the collisions in the selected collisions set, we replace the *target* state reference from the original data, with the equivalent state from the transition data
- two states are considered the same if they contain the same quantum numbers values
- if we can not find the equivalent state in the transition data, we drop the collision from the merge
- the state values of the collider elements remain the same
- all related source, method and function references are also added to the new data

The merged data can be displayed in a panel shown below.



We can export one or all parts of data from this window into a text file with comma separated values by selecting one or all checkboxes and clicking the “Save as ASCII” button.

We can also export these information into RADEX format file by clicking the “Save as RADEX” button or into XML format file by selecting “Save as XSAMS” button.

Extensions

SPECTCOL can be launched with arguments. These arguments must be VAMDC-

XSAMS query URLs. So the application appears with data filled into the involved summary tables.

Example

```
$ java -jar SPECTCOL-11-12-SNAPSHOT-jar-with-dependencies.jar http://vamdc.mssl.ucl.ac.uk/node/hitran/tap/sync?
LANG=VSS2&REQUEST=doQuery&FORMAT=XSAMS&QUERY=select+*+where+%28%28MoleculeStoichiometricFormula+
%3D+%27CO%27%29%29
```