Online Atomic and Molecular Databases

Christian Hill
UCL, Quantemol LTD

IAEA Technical Meeting on Technical Aspects of Atomic and Molecular Data Processing and Exchange
(24th Meeting of the A+M Data Centres)

4–6 September 2017, IAEA Headquarters, Vienna, Austria
Summary

- **VAMDC** – connecting diverse databases through a common query interface and data exchange standard

- **HITRAN** – a database of molecular spectroscopic properties implemented as an online service with a relational backend

- **ExoMol** – methods for the storage and manipulation of large data sets

- **QuantemolDB** – Data on plasma processes implemented as an online service with user interaction

- **Challenges and future directions** for atomic and molecular databases
VAMDC

• VAMDC = Virtual Atomic and Molecular Data Centre

• http://vamdc.eu/

• EU FP7-funded 5-year project to build a consortium of research institutes and lay down a framework of standards for the interchange of atomic and molecular (AM) data:
  
  • VSS2: an SQL-like query language
  
  • XSAMS: an XML schema for storing and transmitting AM data.

• VAMDC portal: an online resource for simultaneously querying multiple databases.
VAMDC

VSS2 Query Language

- Based on a subset of SQL (Structured Query Language)
- Used as a parameter to a “GET” query in a URL, e.g.
  
  http://www.example.com/?query=SELECT States WHERE StoichiometricFormula="CO2"

- Defines “keywords” for types of data (returnables) and search parameters (restrictables)
VAMDC

XSAMS

• An XML-based format: a valid XSAMS document must conform to the XSAMS Schema

• Data from different databases can be compared, aggregated, etc.

• Data provenance: Requires data to have **units** and **sources** (for citation)

• Encourages specification of errors and uncertainties: data integrity
VAMDC

XSAMS

• Standard overseen by the IAEA (Atomic and Molecular Data Division)
• Verbose (because it’s XML) ...
• .... but compresses well (typically 40x)
• “self-describing”
• “self-validating” (to some extent)
Example XSAMS snippet: the self-describing XSAMS format helps verify quantum number assignments, e.g. $^{16}$O$^{17}$O:

```
<hundb:QNs>
  <hundb:ElecStateLabel>X</hundb:ElecStateLabel>
  <hundb:Lambda>0</hundb:Lambda>
  <hundb:S>1.0</hundb:S>
  <hundb:v>0</hundb:v>
  <hundb:J>32.0</hundb:J>
  <hundb:N>32</hundb:N>
  <hundb:F nuclearSpinRef="O2">29.5</hundb:F>
</hundb:QNs>
```

Validation, e.g.
\[
\begin{align*}
\nu & \geq 0, \\
N & \geq \Lambda, \\
|N-S| & \leq J \leq N+S \\
|J-I| & \leq F \leq J+I \\
(I & = 5/2 \text{ for } ^{17}\text{O, identified as "O2"})
\end{align*}
\]
Other VAMDC standards

- Standards for identifying species - i.e. specific isotopes of atoms or ions; specific isotopologues of molecules

- InChI = *IUPAC International Chemical Identifier*, e.g. $^{13}$C$^{16}$O$_2$ is:

  $\text{InChI}\!\!=\!\!1S/\text{CO}_2/c2\!-\!1\!-\!3/i1\!+\!1$

- The InChIKey is a 27-character hash on the species’ InChI identifier: this InChI becomes

  CURLTUGMZLYLIDIOUBTZVSYSA-N

- Allows unambiguous and unique identification of species
VAMDC

Other VAMDC standards

• Services exist (within VAMDC and on the wider www e.g. ChemSpider, OpenBabel) to translate between chemical identifier formats: e.g. for CO2:
  • SMILES: C(=O)=O
  • InChI: InChI=1S/CO2/c2-1-3
  • InChIKey: CURLTUGMZLYLDI-UHFFFAOYSA-N
  • CAS Registry Number: 124–38–9
  • CML
  • ...
VAMDC

Data Permanence

• Many databases simply replace old data with new
• Old data is therefore “lost”
• Can be impossible to reproduce results using old data
• VAMDC standards allow the use of the “ValidOn” keyword
• Old data can be “expired” but not deleted or replaced
The VAMDC portal (portal.vamdc.org/) links nodes to allow simultaneous querying of databases.

A VAMDC database node is an online service which:

- Can understand the VAMDC query language, VSS2
- Will return data in valid XSAMS format
- Is listed with the VAMDC registry
Some VAMDC database nodes:

- **VALD** - a major database of atomic and small-molecule spectroscopic line parameters (focus: astronomy and astrophysics)
- **HITRAN** - the major database of molecular spectroscopic line parameters for planetary atmospheres
- **BASECOL** - database of atomic and molecular collisional cross sections
- **CDMS / JPL** - microwave/FIR line parameters for interstellar spectroscopy
- **KIDA** - astrochemical kinetic database
VAMDC Portal

User makes query at portal website

VAMDC Portal

Databases
VAMDC Portal

User

VAMDC Portal

VSS2 queries

Databases
VAMDC Portal

User

VAMDC Portal

Data count responses

[No Data]

Databases
VAMDC Portal

User

Retrieve data

VAMDC Portal

XSAMS data files

[No Data]

Databases
**VAMDC Portal**

User

- e.g. csv data files

- XSAMS data files

- [No Data]

**Portal Data Processor**

**Databases**
**VAMDC Portal**

![VAMDC Portal](image)

### XSAMS processors
- Bibtex from XSAMS
- Table views of XSAMS
- Xsams2SME

### Database Table
<table>
<thead>
<tr>
<th>Name</th>
<th>Response</th>
<th>Download</th>
<th>Species</th>
<th>States</th>
<th>Processes</th>
<th>Radiative</th>
<th>Collisions</th>
<th>Non Radiative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water internet Accessible Distributed Information System</td>
<td>OK</td>
<td>XSAMS</td>
<td>1</td>
<td>201</td>
<td>627</td>
<td>627</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HITRAN-UCL resource</td>
<td>OK</td>
<td>XSAMS</td>
<td>1</td>
<td>82</td>
<td>189</td>
<td>189</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cologne Database for Molecular Spectroscopy: VAMDC-TAP service</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Carbon Dioxide Spectroscopic Databank - 1000K</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Carbon Dioxide Spectroscopic Databank - 296K</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MeCaSDa - Methane Calculated Spectroscopic Database</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ECaSDa - Ethene Calculated Spectroscopic Database</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GSMA Reims S&amp;MPO</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VALD sub-set in Moscow (obs)</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VALD (atoms)</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### VAMDC Portal

**Table of XSAMS Processes**

<table>
<thead>
<tr>
<th>Name</th>
<th>Response</th>
<th>Download</th>
<th>Species</th>
<th>States</th>
<th>Processes</th>
<th>Radiative</th>
<th>Collisions</th>
<th>Non Radiative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water internet Accessible Distributed Information System</td>
<td>OK</td>
<td>XSAMS</td>
<td>1</td>
<td>201</td>
<td>627</td>
<td>627</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HITRAN-UCL resource</td>
<td>OK</td>
<td>XSAMS</td>
<td>1</td>
<td>82</td>
<td>189</td>
<td>189</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cologne Database for Molecular Spectroscopy: VAMDC-TAP service</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Carbon Dioxide Spectroscopic Databank - 100K</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Carbon Dioxide Spectroscopic Databank - 296K</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MeCaSDa - Methane Calculated Spectroscopic Database</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ECaSDa - Ethene Calculated Spectroscopic Database</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GSC Reims S&amp;MPO</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VALD sub-set in Moscow (obs)</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VALD (atoms)</td>
<td>EMPTY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Line-list view of XSAMS

(Switch to view of **states** or **collisions**)

Assume* all wavelengths are in vacuum and [Convert to air](#) (at IAU STP).

Assume all wavelengths are in air (at IAU STP) and [Convert to vacuum](#).

Show table in CSV format

<table>
<thead>
<tr>
<th>Species</th>
<th>( \nu/\nu/n/E )</th>
<th>Probability</th>
<th>Upper state</th>
<th>Lower state</th>
<th>Broadening/shifting</th>
</tr>
</thead>
<tbody>
<tr>
<td>(12C)(16O)</td>
<td>n=2002.115 1/cm</td>
<td>A=28.85</td>
<td>5582.0901 1/cm — X ( \nu = 2 ) J = 26</td>
<td>3579.9751 1/cm — X ( \nu = 1 ) J = 27</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2003.6679 1/cm</td>
<td>A=14.46</td>
<td>4027.0394 1/cm — X ( \nu = 1 ) J = 31</td>
<td>2023.3715 1/cm — X ( \nu = 0 ) J = 32</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2006.7836 1/cm</td>
<td>A=29.08</td>
<td>5484.3687 1/cm — X ( \nu = 2 ) J = 25</td>
<td>3477.5851 1/cm — X ( \nu = 1 ) J = 26</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2008.5254 1/cm</td>
<td>A=14.57</td>
<td>3909.6562 1/cm — X ( \nu = 1 ) J = 30</td>
<td>1901.1308 1/cm — X ( \nu = 0 ) J = 31</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2011.4211 1/cm</td>
<td>A=29.29</td>
<td>5390.3747 1/cm — X ( \nu = 2 ) J = 24</td>
<td>3378.9536 1/cm — X ( \nu = 1 ) J = 25</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2013.3524 1/cm</td>
<td>A=14.68</td>
<td>3796.0148 1/cm — X ( \nu = 1 ) J = 29</td>
<td>1782.6624 1/cm — X ( \nu = 0 ) J = 30</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2016.0272 1/cm</td>
<td>A=29.51</td>
<td>5300.1117 1/cm — X ( \nu = 2 ) J = 23</td>
<td>3284.0845 1/cm — X ( \nu = 1 ) J = 24</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2018.1488 1/cm</td>
<td>A=14.79</td>
<td>3686.1197 1/cm — X ( \nu = 1 ) J = 28</td>
<td>1667.9709 1/cm — X ( \nu = 0 ) J = 29</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2020.6018 1/cm</td>
<td>A=29.73</td>
<td>5213.5832 1/cm — X ( \nu = 2 ) J = 22</td>
<td>3192.9814 1/cm — X ( \nu = 1 ) J = 23</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2022.9144 1/cm</td>
<td>A=14.9</td>
<td>3579.9751 1/cm — X ( \nu = 1 ) J = 27</td>
<td>1557.0607 1/cm — X ( \nu = 0 ) J = 28</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2025.1447 1/cm</td>
<td>A=29.95</td>
<td>5130.7926 1/cm — X ( \nu = 2 ) J = 21</td>
<td>3105.6479 1/cm — X ( \nu = 1 ) J = 22</td>
<td>Detail</td>
</tr>
<tr>
<td>(12C)(16O)</td>
<td>n=2027.6491 1/cm</td>
<td>A=15.02</td>
<td>3477.585 1/cm — X ( \nu = 1 ) J = 26</td>
<td>1449.9359 1/cm — X ( \nu = 0 ) J = 27</td>
<td>Detail</td>
</tr>
</tbody>
</table>
VAMDC
Guided Queries
VAMDC

Guided Queries
HITRAN

• A database of molecular spectroscopic data for atmospheric applications (climate modelling, remote sensing, etc.)

• Initiated by the US Air Force in the 1960s

• 5,000,000+ lines, 100 MB+ of cross sections

• Increasingly used for planetary atmospheres

• Now an provided by a web service, HITRANonline

• Also acts as a VAMDC node
HITRAN
Evolution: 1960s / 1970s
| 21 | 2291.946330 | 4.940E-29 | 3.414E-02 | 0.06820.086 | 3595.59060.78...0.007000 | 4 0 0 02 | 2 1 1 03 | Q 32f | 4455501221 | 1 1 1 4 | 65.0 65.0 |
| 21 | 2291.947360 | 1.660E-28 | 3.650E+02 | 0.06980.093 | 5204.946170.8000040 | 0 4 4 21 | 0 4 4 11 | R 24f | 4455501221 | 1 1 1 4 | 51.0 49.0 |
| 21 | 2291.953028 | 4.614E-27 | 2.522E+00 | 0.04610.073 | 3633.909170.76000000 | 0 4 4 11 | 0 4 4 01 | Q 49f | 4545501221 | 1 1 1 4 | 55.0 55.0 |
| 23 | 2291.957370 | 6.450E-27 | 1.504E+02 | 0.08930.123 | 2501.496970.50000000 | 2 0 0 13 | 2 0 0 01 | R 73e | 4555501221 | 1 1 1 4 | 298.0 294.0 |
| 21 | 2291.959750 | 4.946E-28 | 4.906E-01 | 0.06820.086 | 3634.144070.80000700 | 2 1 1 12 | 2 1 1 02 | Q 32f | 4455501221 | 1 1 1 4 | 65.0 65.0 |
| 22 | 2291.967650 | 4.920E-28 | 1.880E+02 | 0.06030.060 | 4140.440160.66...0.007100 | 1 1 1 12 | 1 1 1 01 | R 73e | 4555501221 | 1 1 1 4 | 51.0 49.0 |
| 22 | 2291.974450 | 3.600E-29 | 1.790E+02 | 0.06600.075 | 5496.458170...0.004490 | 0 4 4 11 | 0 4 4 01 | R 73e | 4555501221 | 1 1 1 4 | 51.0 49.0 |
| 22 | 2291.977720 | 4.040E-30 | 1.946E+02 | 0.06640.074 | 3386.790070...0.004430 | 0 4 4 11 | 0 4 4 01 | R 73e | 4555501221 | 1 1 1 4 | 51.0 49.0 |
| 21 | 2291.986760 | 4.920E-28 | 1.880E+02 | 0.06030.060 | 4140.440160.66...0.007100 | 1 1 1 12 | 1 1 1 01 | R 73e | 4555501221 | 1 1 1 4 | 51.0 49.0 |
| 21 | 2291.974450 | 3.600E-29 | 1.790E+02 | 0.06600.075 | 5496.458170...0.004490 | 0 4 4 11 | 0 4 4 01 | Q 49f | 4545501221 | 1 1 1 4 | 93.0 91.0 |
| 21 | 2291.977720 | 4.040E-30 | 1.946E+02 | 0.06640.074 | 3386.790070...0.004430 | 0 4 4 11 | 0 4 4 01 | Q 49f | 4545501221 | 1 1 1 4 | 93.0 91.0 |
HITRAN
Evolution: 2015 – Present

The HITRAN Database

HITRAN is an acronym for high-resolution transmission molecular absorption database. HITRAN is a compilation of spectroscopic parameters that a variety of computer codes use to predict and simulate the transmission and emission of light in the atmosphere.

News

Articles describing HITRANonline, HAPI and new line shape representations

Video tutorial for using HITRANonline. Recording from 10.26.2015 webinar

All inquiries can now be made to HITRAN's support team at info@hitran.org!

Database Updates

Update of $v_9$ and $v_8$ band regions and addition of data for $v_6 + v_8$ band region of diacetylene ($C_4H_2$)

Global update of the CO linelist, including addition of new bands up to the sixth overtone. Broadening and shift by CO$_2$, H$_2$ and He is also included for the first time.

Update of the line list for the principal isotopologue of ozone ($^{16}$O$_3$)
HITRAN

HITRANonline

- Web browser-based application with a relational database backend (MySQL) for searching the HITRAN database.

- [www.hitran.org](http://www.hitran.org)

- Features:
  - User management (security, accounting, profiles)
  - Per-user, customisable output formats
  - Species metadata (partition sums, masses, ...)
  - Interactive charts (for cross sections)
  - Acts as a VAMDC database node
ExoMol

- *Ab initio* calculated molecular line lists for high-temperature spectroscopic applications:
  - Flames
  - Cool stars
  - Exoplanet atmospheres
- ERC-funded, 5-year project at UCL (PI: Jonathan Tennyson)
- ~50 molecules in all major isotopologues
- Line positions (cm\(^{-1}\)) and Einstein A-coefficients
ExoMol

- Larger molecules have a huge number of transitions (e.g. CH$_4$ $10^{10}$ lines in preliminary form)
- At elevated temperature opacity due to all transitions must be included
- Too much data for a RDBMS, so:
  - Separate states and transitions into separate files
  - Break large transitions files into spectral regions
  - Provide software tools for conversion to other formats (e.g. HITRAN)
  - Provide online services for the calculation of $p$- and $T$-dependent absorption cross sections
### ExoMol

High temperature molecular line lists for modelling exoplanet atmospheres

**Data / By Molecule**

**Molecules**

<table>
<thead>
<tr>
<th>Ions</th>
<th>Metal Oxides</th>
<th>Triatomic Molecules</th>
<th>Larger Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>LiH⁺, H₂⁺, HeH⁺, H₃⁺</td>
<td>VO, AIO, TiO, SiO, CaO</td>
<td>H₂O, CO₂, SO₂, HCN, N₂O, H₂S</td>
<td>CH₄, NH₃, HNO₃, H₂O₂, H₂CO, PH₃, PH₃, SO₃, SiH₄</td>
</tr>
<tr>
<td>Other Oxides</td>
<td>Other Diatomics</td>
<td>Metal Hydrides</td>
<td></td>
</tr>
<tr>
<td>CO, NO, PN, KCl</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[exomol.com](http://exomol.com)
ExoMol

exomol.com

• Free-to-use, no sign-up required
• Search by molecule / isotopologue or
• Search by data type (species energies, line lists, partition functions, cooling functions, spectroscopic model, \textit{ab initio} calculation details, pressure-broadening coefficients, etc.)
• Bibliographic database of published journal articles relating to small molecules
QuantemolDB
quantemoldb.com

• Comprehensive database of data relating to low-temperature plasmas:
  • Energy-dependent collisional cross sections
  • Arrhenius-like parameters
  • Sticking coefficients for surface processes
  • Surface reaction probabilities
  • ...

• Relational database backend with a user-friendly web-based frontend
QuantemolDB

- Flexible data model allows for
  - Generic parameter data
  - Generic tabular data
  - Bibliographic references
  - Building “Chemistries” of related reactions / collisions

- Software libraries provide tools for:
  - Identifying and providing metadata on species and their quantum states
  - Manipulating the units of physical quantities
Quantemol-DB

Trusted chemistries for plasma research.

LATEST NEWS
- 10 Aug 2017
  We are moving! >
- 28 Jul 2017
  We will see you at ISPC! >
- 10 Jul 2017
  AVS 17th International Conference on Atomic Layer Deposition (ALD 2017) >
- 02 Jun 2017
  New NF3/O2/Ar chemistry added to QDB! >
- 02 Jun 2017
  Nitrogen Trifluoride (NF3) Cross Section Calculated and Confirmed >

Current Status
12545 Reaction data sets

Website by Quantemol
Reactions Search

CO2

Reactants

33 reactions found for reactants matching CO2.

To compare up to 5 reaction cross sections, click on the icon in the table below to add it to the "Selected Cross Sections" clipboard. All data sets for each reaction will be displayed.

<table>
<thead>
<tr>
<th>Species</th>
<th>Reaction</th>
<th>Process</th>
<th>Data available</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO2</td>
<td>e⁺ + CO₂ → e⁺ + e⁺ + CO₂⁺</td>
<td>EIN</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=1]</td>
<td>e⁺ + CO₂ → O⁺ + CO</td>
<td>EDA</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=1]</td>
<td>CO₂ + CO⁺ → CO₂⁺ + CO</td>
<td>HIR</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=2]</td>
<td>e⁺ + CO₂ → e⁺ + O + CO</td>
<td>EDS</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=2]</td>
<td>e⁺ + CO₂ → e⁺ + O + CO</td>
<td>EDS</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=3]</td>
<td>e⁺ + CO₂ → e⁺ + CO₂⁺ [v=1]</td>
<td>EVX</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=3]</td>
<td>e⁺ + CO₂ → e⁺ + CO₂⁺ [v=2]</td>
<td>EVX</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=4]</td>
<td>e⁺ + CO₂ → e⁺ + CO₂⁺ [v=3]</td>
<td>EVX</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=5]</td>
<td>e⁺ + CO₂ → e⁺ + CO₂⁺ [v=5]</td>
<td>EVX</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=6]</td>
<td>e⁺ + CO₂ → e⁺ + CO₂⁺ [v=6]</td>
<td>EVX</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=7]</td>
<td>e⁺ + CO₂ → e⁺ + CO₂⁺ [v=7]</td>
<td>EVX</td>
<td>✓</td>
</tr>
<tr>
<td>CO₂ [v=8]</td>
<td>e⁺ + CO₂ → e⁺ + CO₂⁺ [v=8]</td>
<td>EVX</td>
<td>✓</td>
</tr>
</tbody>
</table>
QuantemolDB

Datasets

D5373: $e^+ + CO_2 \rightarrow e^- + e^- + O + CO$
D5374: $e^- + CO_2 \rightarrow e^- + e^- + O_2 + C^+$
D5375: $e^- + CO_2 \rightarrow e^- + e^- + O^+ + CO$

B383: Y. Itikawa, "Cross Sections for Elect..."
QuantemolDB

- Flexible data model allows for
  - Generic parameter data
  - Generic tabular data
  - Bibliographic references
  - Building “Chemistries” of related reactions / collisions

- Software libraries provide tools for:
  - Identifying and providing metadata on species and their quantum states
  - Manipulating the units of physical quantities
QuantemolDB

• Complex demands for user-management:
  • AAA (Authentication, Authorization, Accounting)
  • Some users can upload data:
    • Such contributions must be validated, logged and attributed
    • Incoming data must have the correct format and physical units
  • Security concerns
• Users can evaluate data
• Use of the API (next slide) requires a key (token)
QuantemolDB

API

- quantemoldb.com exposes an API for the automated retrieval of data by software applications
- Data is returned in one of several formats specified by a keyword as part of the API request
- Entire Chemistries can be downloaded ready-made or constructed by specifying the species involved, e.g.

Challenges for AM databases

• Sustainability is a problem for any non-commercial database (= €, $, £, ¥, ₩)

• Security concerns and database / server maintenance

• A database cannot be sustainable unless it is easy to upload data to it (even if only by its administrators or their proxies)

• The problem of a common query language and output format across a wide range of applications has still not been solved
Future Directions

• A lightweight, generic relational database model for AM data?

• Software tools for conversion between different data formats

• Data attribution through simple identifiers (DOI or similar)

• Time stamping and versioning

• Distributed database hosting?
  • Subscription-based?
  • Blockchain technology? (Ethereum?)
Acknowledgements

- **VAMDC**
  - “This work is supported by the VAMDC project. VAMDC is funded under the “Combination of Coordination and Support Actions” Funding Scheme of The Seventh Framework Program. Call topic: INFRA-2012-3.3 Scientific Data Infrastructure. Grant Agreement number: 313284).”
  - SUP@VAMDC is supported by EU in the framework of the "Research Infrastructures - FP7 - INFRA-2012-3.3 - Scientific Data Infrastructures" initiative.”

- **HITRAN**
  - NASA Earth Observing System
  - NASA Planetary Atmospheres Program
  - Cooperative Research Development Foundation
  - Larry Rothman and Iouli Gordon

- **ExoMol**
  - “ExoMol is funded by the ERC under the Advanced Investigator Project 267219”
  - Jonathan Tennyson

- **QuantemolDB**
  - Quantemol LTD
  - Jonathan Tennyson and Dan Brown