# Data structures for photoadsorption within the ExoMol project

Jonathan Tennyson [ID],★ Marco Pezzella, Jingxin Zhang [ID] and Sergei N. Yurchenko

*Department of Physics and Astronomy, University College London, Gower Street, WC1E 6BT London, UK*

## ABSTRACT

The ExoMol data base currently provides comprehensive line lists for modelling the spectroscopic properties of molecules in hot atmospheres. Extending the spectral range of the data provided to ultraviolet (UV) wavelengths brings into play three processes not currently accounted for in the ExoMol data structure, namely photodissociation, which is an important chemical process in its own right, the opacity contribution due to continuum absorption, and predissociation which can lead to significant and observable line-broadening effects. Data structures are proposed which will allow these processes to be correctly captured and the (strong) temperature-dependent effects predicted for UV molecular photoabsorption in general and photodissociation in particular to be represented.

**Key words:** data methods – photoabsorption – exoplanets.

## 1 INTRODUCTION

The analysis of light as function of wavelengths provides a major window on the Universe. Interpretation of these signals requires appropriate laboratory data. The ExoMol project (Tennyson & Yurchenko 2012) was established to provide molecular line lists for hot astronomical atmospheres. Up until recently the ExoMol project, in keeping with similar projects such as HITEMP (Rothman et al. 2010; Hargreaves et al. 2019), TheoReTS (Rey et al. 2016), and NASA Ames (Huang, Schwenke & Lee 2021), has presented results as (large) lists of transitions or spectral lines. Implicitly this assumes that all lines are discrete transitions between bound states (bound–bound transitions) even though in some cases the upper states may actually lie in the continuum so that these transitions, while still appearing as lines, actually represent part of the bound – free spectrum. The ExoMol project uses a well-defined format for line lists (Tennyson, Hill & Yurchenko 2013) which has been further developed as part of its data releases (Tennyson et al. 2016b, 2020).

However, a number of recent developments, discussed below, have extended the scope of the ExoMol project and therefore the data it provides. This has caused us to consider how to generalize the ExoMol data structure to accommodate both the increased range of data and also its different uses as bound – free data are important not only for opacities and spectroscopic models but also, in that they represent a route to photodissociation, which is an important process for chemical models. Sharp transitions to states lying above the dissociation limit are already starting to be captured by ExoMol as part of standard line lists (Qu, Yurchenko & Tennyson 2021; Owens et al. 2022). However, once occupied these above dissociation states can decay either by emitting a photon, such as ultraviolet (UV) fluorescence which is an important astrophysical process (Lupu et al.

2011; Gérard et al. 2023), or they can dissociate. The aforementioned dissociation region also contains a continuum component to the photoabsorption caused, e.g. by excitation to dissociative electronically excited states. Currently neither ExoMol nor indeed the aforementioned data bases, capture this component. At the same time we have started to consider the role of photodissociation (Pezzella, Yurchenko & Tennyson 2021, 2022) which itself usually comprises sharp lines sitting on top of a continuum. Photoabsorption into the continuum cannot be represented by the current ExoMol line format. We note that continuous opacities on a grid of temperatures and wavelengths for molecules were generated by Kurucz, van Dishoeck & Tarafdar (1987) and tabulated experimental vacuum ultraviolet (VUV) cross-sections of molecules are collected in the MPI-Mainz UV/VIS Spectral Atlas (Keller-Rudek et al. 2013) as a continuous spectrum.

A third issue is the representation of the so-called predissociation which occurs when transitions to excited electronic states which lie above the dissociation limit spontaneously undergo a further process (usually a curve crossing or tunnelling) which leads to dissociation. The resulting lines are observed to be broadened, often significantly, due to the shorter lifetimes associated with predissociating states. So far while the ExoMol project has included states in its line lists which are predissociative, it has ignored the important line-broadening effects which result from the reduced lifetime associated with predissociative states. A recent study by Pavlenko et al. (2022) of the spectrum AlH in M-dwarf star Proxima Cen highlights problems with this approach. Pavlenko *et al.* used the ExoMol AlHambra line list for AlH (Yurchenko et al. 2018b) to model this spectrum. For the majority of transitions, which do not show any effects due to predissociation, this line list worked well but it proved to be less accurate for transitions to predissociating states. Importantly, only the lines which showed broadening, often by as much as 5 cm$^{-1}$, due to predissociation are not saturated in the stellar spectra meaning that it was only by analysing these broadened lines that Pavlenko *et al.* were able to retrieve abundances of AlH. It would therefore clearly

---

★ E-mail: j.tennyson@ucl.ac.uk
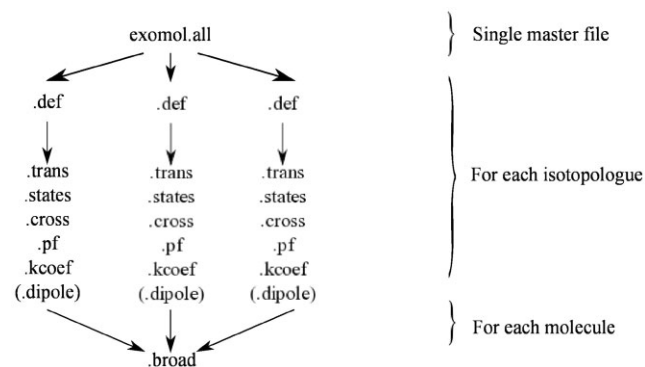
**Figure 1** Summary of the current ExoMol data structure.

be advantageous to include consideration of predissociation effects in the ExoMol data base.

In this research note we propose a generalization of the current ExoMol format to allow for the various processes discussed above. At the same time we draw a clear distinction between the photoabsorption data, needed for spectral and opacity models, and for the data needed for modelling the chemical consequences of photodissociation.

## 2 THE PRESENT EXOMOL DATA MODEL

Fig. 1 gives a simplified ExoMol data structure; a complete specification of the file types is given in Table 1. The master file exomol.all (https://www.exomol.com/exomol.all) gives an overview of the en-

tire data base and points towards the .def files which characterize the recommended line lists for each isotopologue for which data are available. The .def file contains specification of the data set in terms of what is available, e.g. uncertainties or lifetimes, quantum numbers used in the states file, and file sizes. It also gives a version number in yyyymmdd date format. The core of the data base are the .states and .trans files which provide a compact form of the line list data.

Table 2 gives the specifications for the mandatory part of the .states file. These specifications include the optional components: uncertainties in the term values, the state lifetime, and Landé *g*-factor. The specification of term value uncertainties was introduced as part of the last data release (Tennyson et al. 2020) and the aim is to make their inclusion compulsory once the available data sets have all had uncertainties added. The lifetimes column has thus far contained radiative lifetimes computed using the Einstein A coefficients available in the transitions file (Tennyson et al. 2016a); so far lifetime effects due to other processes such as predissociation have not been considered. As discussed below we propose changing this.

After the mandatory fields, the states files contain data on the quantum numbers and other metadata used to specify each state. The number and format of these quantum numbers is specified in by the .def file associated with that data set. Other metadata associated with level can also be included in this section. Table 3 shows the start of the states file for the recently update AlHambra line list for $^{27}Al^{16}O$ (Bowesman et al. 2021). Note that we have taken the opportunity to update our quantum number specifications to conform with the PyValem python package for parsing, validating, and manipulating quantum states labels of atoms, ions, and small molecules (Hill & Hanicinec 2022; Hill et al., in preparation). In general, this change

**Table 1.** Specification of the ExoMol file types. (Contents in brackets are optional.) The final three entries are new file types, see Tables 5, 6, and 7, respectively, for specifications.

| File extension | $N_{files}$ | File DSname | Contents |
|---|---|---|---|
| .all | 1 | Master | Single file defining contents of the ExoMol data base |
| .def | $N_{tot}$ | Definition | Defines contents of other files for each isotopologue |
| .states | $N_{tot}$ | States | Energy levels, quantum numbers, lifetimes, (Landé *g*-factors, uncertainties) |
| .trans | *a* | Transitions | Einstein *A* coefficients, (wavenumber) |
| .broad | $N_{mol}$ | Broadening | Parameters for pressure-dependent line profiles |
| .cross | *b* | Cross-sections | Temperature or temperature and pressure-dependent cross-sections |
| .kcoef | *c* | *k*-Coefficients | Temperature and pressure-dependent *k*-coefficients |
| .pf | $N_{tot}$ | Partition function | Temperature-dependent partition function, (cooling function) |
| .dipoles | $N_{tot}$ | Dipoles | Transition dipoles including phases |
| .super | *d* | Superlines | Temperature-dependent superlines (histograms) on a wavenumber grid |
| .nm | *e* | VUV cross-sections | Temperature and pressure-dependent VUV cross-sections (wavelength, nm) |
| .fits, .h5, .kta | *f* | Opacities | Temperature and pressure-dependent opacities for radiative-transfer applications |
| .overview | $N_{mol}$ | Overview | Overview of data sets available |
| .readme | $N_{iso}$ | Readme | Specifies data formats |
| .model | $N_{iso}$ | Model | Model specification |
| .cont | $N_{iso}/0$ | Continuum | Continuum contribution to the photoabsorption |
| .pdef | $N_{tot}$ | Definition | Defines contents of .photo files for each isotopologue |
| .photo | $N_{iso}$ | Photodissociation | Photodissociation cross-sections for each lower state |

*Notes.* $N_{files}$ is the total number of possible files; $N_{mol}$ is the number of molecules in the data base; $N_{tot}$ is the sum of $N_{iso}$ for the $N_{mol}$ molecules in the data base; and $N_{iso}$ is the number of isotopologues considered for the given molecule.

*a* There are $N_{tot}$ sets of .trans files but for molecules with large numbers of transitions the .trans files are subdivided into wavenumber regions.

*b* There are $N_{cross}$ sets of .cross files for isotopologue.

*c* There are $N_{kcoef}$ sets of .kcoef files for each isotopologue.

*d* There are $N_T$ sets of *T*-dependent superlines.

*e* There are $N_{VUV}$ sets of VUV cross-sections.

*f* Set of opacity files in the format native to specific radiative-transfer programs.

**Table 2.** Specification of the mandatory part of the states file with extra data options unc, $\tau$, and $g$.

| Field | Fortran format | C format | Description |
|---|---|---|---|
| $i$ | I12 | %12d | State ID |
| $E$ | F12.6 | %12.6f | State energy in cm$^{-1}$ |
| $g_{tot}$ | I6 | %6d | State degeneracy |
| $J$ | I7/F7.1 | %7d/%7.1f | $J$-quantum number (integer/half-integer) |
| (unc) | F12.6 | %12.6f | Uncertainty in state energy in cm$^{-1}$ (optional) |
| ($\tau$) | ES12.4 | %12.4e | Lifetime in s (optional) |
| ($g$) | F10.6 | %10.6f | Landé $g$-factor (optional) |

*Notes.* ID: State identifier – a non-negative integer index, starting at 1. $J$: Total angular momentum quantum, excluding nuclear spin.
Fortran format, $J$ integer: (I12,1x,F12.6,1x,I6,I7,1x,ES12.4,1x,F10.6)
or $J$ half-integer: (I12,1x,F12.6,1x,I6,F7.1,1x,ES12.4,1x,F10.6).

**Table 3.** An excerpt from the recently updated states file for $^{27}$Al$^{16}$O, see Bowesman et al. (2021).

| $i$ | $\tilde{E}$ (cm$^{-1}$) | $g$ | $J$ | unc (cm$^{-1}$) | $\tau$ (s) | $+/-$ | e/f | State | $v$ | $|\Lambda|$ | $|\Sigma|$ | $|\Omega|$ | Source label |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.000 000 | 12 | 0.5 | 0.000 001 | inf | + | e | X(2SIGMA+) | 0 | 0 | 0.5 | 0.5 | EH |
| 2 | 965.416 878 | 12 | 0.5 | 0.001 651 | 3.6193e+01 | + | e | X(2SIGMA+) | 1 | 0 | 0.5 | 0.5 | PS |
| 3 | 1916.827 286 | 12 | 0.5 | 0.010 519 | 9.1761e+00 | + | e | X(2SIGMA+) | 2 | 0 | 0.5 | 0.5 | PS |
| 4 | 2854.162 814 | 12 | 0.5 | 0.022 148 | 4.2147e+00 | + | e | X(2SIGMA+) | 3 | 0 | 0.5 | 0.5 | PS |
| 5 | 3777.464 572 | 12 | 0.5 | 0.019 275 | 2.3874e+00 | + | e | X(2SIGMA+) | 4 | 0 | 0.5 | 0.5 | PS |
| 6 | 4686.658 929 | 12 | 0.5 | 0.016 235 | 1.2041e+00 | + | e | X(2SIGMA+) | 5 | 0 | 0.5 | 0.5 | PS |
| 7 | 5346.089 546 | 12 | 0.5 | 0.009 700 | 2.0501e-04 | + | e | A(2PI) | 0 | 1 | 0.5 | 0.5 | Ma |
| 8 | 5581.908 884 | 12 | 0.5 | 0.014 747 | 4.3685e-02 | + | e | X(2SIGMA+) | 6 | 0 | 0.5 | 0.5 | PS |

*Notes.* $i$: State counting number; $\tilde{E}$: Term value (in cm$^{-1}$); $g_{tot}$: Total state degeneracy; $J$: Total angular momentum quantum number; unc: Estimated uncertainty of energy level (in cm$^{-1}$); $\tau$: Lifetime (in s$^{-1}$); $+/-$: Total parity; e/f: Rotationless parity; State: Electronic term value; $v$: Vibrational quantum number; $|\Lambda|$: Absolute value of the projection of electronic angular momentum; $|\Sigma|$: Absolute value of the projection of the electronic spin; $|\Omega|$: Absolute value of the projection of the total angular momentum; and Source label: Method used to generate term value (McKemmish et al., in preparation).

**Table 4.** Specification of the transitions file.

| Field | Fortran format | C format | Description |
|---|---|---|---|
| $i$ | I12 | %12d | Upper state ID |
| $f$ | I12 | %12d | Lower state ID |
| $A$ | ES10.4 | %10.4e | Einstein $A$ coefficient in s$^{-1}$ |
| $\tilde{v}_{fi}$ | E15.6 | %15.6e | Transition wavenumber in cm$^{-1}$ (optional) |

*Note.* Fortran format: (I12,1x,I12,1x,ES10.4,1x,ES15.6).

only affects electronic state designations for which X2SIGMA+, A2PI, and so forth are updated to X(2SIGMA+), A(2PI), etc. This update adds two characters to the electronic state field but otherwise should be transparent to users; however, it means that all state labels can now be parsed using PyValem which is important for some uses of the data base (Owens et al., in preparation).

Table 4 gives the specification of the simpler but generally much larger .trans file.

## 3 PROPOSED UPDATED EXOMOL DATA MODEL

There are three new aspects that need to be included in an updated ExoMol data structure: predissociation, the continuum contribution to the opacity, and photodissociation. Fig. 2 illustrates the various photoabsorption processes for the case of a diatomic molecule.
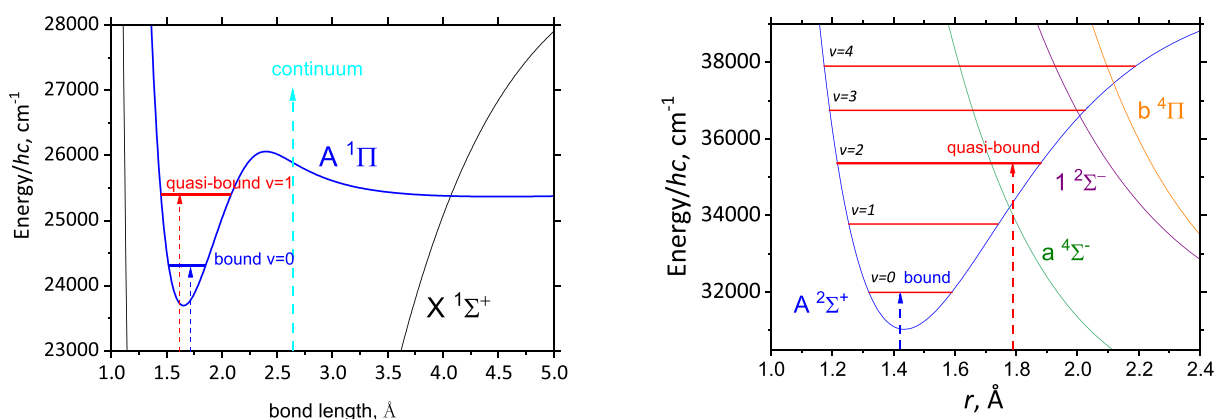
### 3.1 Predissociation

Fig. 3 illustrates the main mechanism leading to predissociation; for AlH it is caused by tunnelling through a small barrier to a dissociation, while for SH it is caused by couplings to dissociative

states crossing the state to which the transition goes to. The effect of predissociation can be included by a minor adjustment to the current ExoMol data structure. Predissociation manifests as a shortened lifetime which leads to enhanced natural line-broadening of any transition to (or from) the state concerned. We therefore propose to generalize the definition of the lifetime, $\tau$, given in the .states file. For most line lists, where predissociation is not important, $\tau$ is defined as the natural lifetime due to radiative decay. In cases where predissociation is considered, $\tau$ will represent the natural life due to both radiative decay and predissociation. For example, the radiative lifetime of $v = 1$ A $^2\Sigma^+$ state of SH was calculated by Gorman, Yurchenko & Tennyson (2019) as 5.13 ns, while the predissociative lifetimes was measured as 5.45(24) ps ($N = 0$) (Wheeler, Orr-Ewing & Ashfold 1997). A similar example of the predissociation effects in the spectra of AlH is the predissociative lifetime of the $J = 23$, $v = 1$, A $^1\Pi$ state of $^{27}$AlH, which was measured by Baltayan & Nedelec (1979) as 4.5 ns, while the radiation lifetime is calculated to be 101.7 ns (Yurchenko et al. 2018b). The reduced lifetime affects the line-broadening of the corresponding transitions and therefore is an important factor in retrievals of AlH abundance from stellar spectra, as was recently demonstrated by Pavlenko et al. (2022).
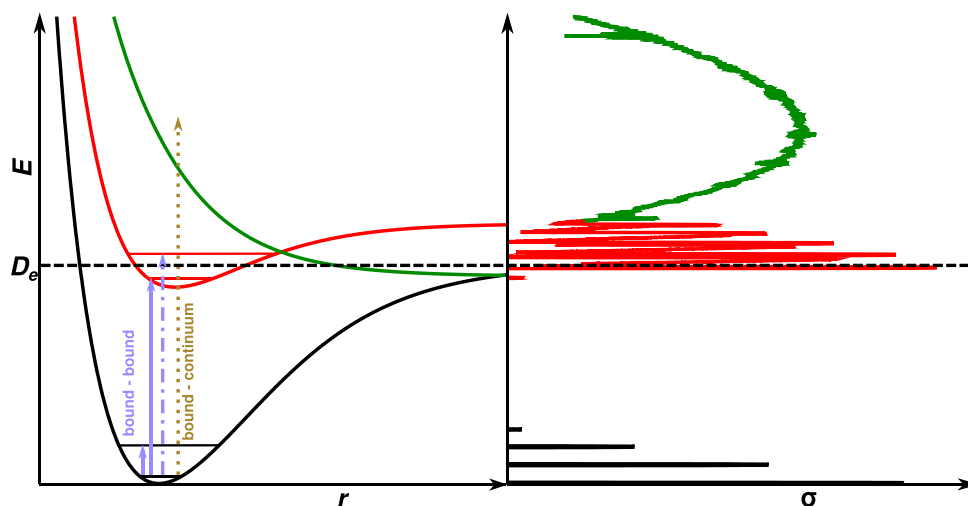
**Figure 2** Schematic representation of photoabsorption for a diatomic molecule showing line and continuum spectra. The transitions within the ground state electronic state (in black) lead to the rovibrational line spectrum (also in black) represented by transitions to excited electronic states (in red); transitions to the repulsive electronic state produce continuum spectra (in green). The line arrows denote sharp, line transitions; the dot interrupted arrow goes to states above the dissociation limit ($D_e$) which can exhibit predissociative effects. The golden dotted arrow shows photoabsorption which we represent using bound–continuum cross-sections. Photoabsorption above $D_e$ can contribute to photodissociation.



**Figure 3** Left-hand panel: Potential energy curves for AlH due to Yurchenko et al. (2018b) showing the predissociation region. The $v = 0$ vibrational state of the A $^1\Pi$ electronic state lies below the AlH dissociation limit: Transitions to this state are sharp as they do not show effects due to predissociation. Conversely, states in the $v = 1$ can predissociate by tunnelling through the barrier to dissociation and show pronounced effects due to lifetime broadening. Right-hand panel: The potential energy curves (bound and dissociative) and predissociated states of SH.

Cases where predissociation effects are included in this lifetime will be marked by a new flag, `prediss`, included in the `.def` file; `prediss` will default to 0 (none) and be set to 1 when the effects are present.

In principle, the natural lifetime provides a contribution to the line profile which sits alongside the temperature-dependent effects of Doppler broadening and pressure-dependent collisional broadening. In practice, the natural lifetime usually makes a minimal contribution to the line profile and is thus neglected. For predissociating states this is no longer the case. However, including the effect of lifetime broadening within a standard Voigt profile is straightforward. Lifetime broadening leads to a Lorentzian line shape like pressure broadening where $\gamma_\tau$, the half-width in cm$^{-1}$ due lifetime broadening, is given by

$$\gamma_\tau = \frac{\hbar}{2\tau h c}. \tag{1}$$

This half-width can simply be added to the pressure-broadening half-width, $\gamma_p$, to give the total Lorentzian component of the line profile. Given that a Voigt profile is already being used, this has little computational impact on a calculation which suggests that routine use of $\gamma_p + \gamma_\tau$ for half-width would avoid the need to worry about whether predissociation needs to be considered or not. To this end, for states with non-negligible predissociative lifetimes, the radiative values of $\tau_{\rm rad}$ in the ExoMol States files will be replaced by $\tau_{\rm prediss}$. In the example of the $J = 23$, $v = 1$, A $^1\Pi$ state of $^{27}$AlH, the ExoMol value $\tau_{\rm rad} = 1.0169 \times 10^{-7}$ s can be replaced by the experimental value $\tau_{\rm prediss} = 4.5 \times 10^{-9}$ s (Baltayan & Nedelec 1979), otherwise calculated values will be used.

## 3.2 Continuum absorption

Both continuum absorption and photodissociation need to be represented as cross-sections rather than lines as they are contin-

**Table 5.** Specification of the .cont continuum photoabsorption cross-section file format.

| Field | Fortran format | C format | Description |
|---|---|---|---|
| $\tilde{\nu}_i$ | F12.6 | %12.6f | Central bin wavenumber, cm$^{-1}$ |
| $\sigma_i$ | ES14.8 | %14.8e | Absorption cross-section, cm$^2$ molec$^{-1}$ |

*Note.* Fortran format: (F12.6,1x,ES14.8).

uum processes. However, in our proposed data model continuum absorption will be represented as a function of wavenumber (cm$^{-1}$) to retain consistency with the line spectra, while photodissociation cross-sections will be stored as function of wavelength (nm).

When considering photoabsorption by molecules to states lying above the dissociation limit, the spectrum can be thought of broadly dividing into two classes: line spectra comprising what looks like sharp bound–bound transitions, and absorption directly into the continuum. Predissociation spectra form an intermediate between these two cases and, as discussed above, will be treated as line spectra. The line spectra can be represented using the form of the standard line lists (line positions, Einstein A coefficients, lower/upper state energies, and other state descriptions), as captured by the .states and .trans files. However, the bound–continuum photoabsorption is best represented as temperature-dependent photoabsorption cross-sections. These continuum photoabsorption cross-sections, whose data specification is given below, will be stored as part of the standard line list data base as separate files for each isotopologue. The temperature-dependent photoabsorption spectrum is then obtained by adding the appropriate line and continuum cross-sections using software such as ExoCross (Yurchenko, Al-Refaie & Tennyson 2018a) or PyExoCross (Zhang, Tennyson & Yurchenko, in preparation).

Continuum molecular absorption due to collision induced absorption (CIA) (Richard et al. 2012; Karman et al. 2019) is already routinely considered as part of astronomical models; it has also been recently recommended that absorption due to the so-called 'water continuum' be considered in model atmospheres for water-rich exoplanets (Anisman et al. 2022).

The data model we propose for including continuum absorption due to simply photoabsorption into continuum states is given in Table 5. Again these cross-sections will be temperature-dependent but, unlike CIA and the water continuum, it is probably safe to assume that this continuum is not strongly pressure-dependent. The presence of a continuum absorption contribution to the photoabsorption will be indicated by a new flag, continuum, included in the .def file; continuum will default to 0 (none) and be set to 1 when the effects are present.

We note that our proposal involves providing photodissociation data on a wavelength grid (in nm) while continuum absorption cross-sections will be provided on a wavenumber grid (in cm$^{-1}$). This latter choice ties in closely with line lists which already prove transitions wavenumber (in cm$^{-1}$). The data structure of continuum absorption cross-sections is presented in Table 5. The file names have the following structure: '<ISO-SLUG>__<DSNAME>__<RANGE>__T<TEMP>K__P<PRESSURE>bar__<STEP>.cont', where ISO-SLUG is the iso-slug molecule name (Tennyson et al. 2020), DSNAME is the name of the line list, RANGE is the wavenumber range in cm$^{-1}$, TEMP is the temperature in K, PRESSURE is the pressure in bar, and STEP is the wavenumber step in cm$^{-1}$.

### 3.3 Photodissociation

Photodissociation cross-sections are separated from the line list and form a distinct section in the ExoMol data base. At present this section contains calculated cross-sections for HCl and HF (Pezzella et al. 2022) and measured cross-sections for CO, $H_2O$, $CO_2$, $SO_2$, $NH_3$, $H_2CO$, and $C_2H_4$ due to Fateev et al. (2023) and $CO_2$ due to Venot et al. (2018). The immediate plan is also to add temperature-dependent cross-sections due to Qin and co-workers who have performed photodissociation calculations on MgO (Bai, Qin & Liu 2021), AlH (Qin, Bai & Liu 2021a), AlCl (Qin, Bai & Liu 2021b), AlF (Qin, Bai & Liu 2022a), and $O_2$ (Hu, Qin & Bai, in preparation), as well as HF and HCl (Qin, Bai & Liu 2022b). In due course a structure of photodissociation .pdef files will be implemented to aid the navigation of this section of the data base.

As the photodissociation cross-sections form a distinct part of the ExoMol data base, we have added a new photodissociation definition file (.pdef) file to the data structure; the proposed format of this file is given in Table 6. This gives the necessary metadata to access and interpret the recommended photodissociation cross-sections. The information section mirror those given in the .def for the same system. For completeness we have added two more flags to the master file, line and photo, which define whether the line spectra and photodissociation cross-sections are present (=1) or not (=0). The default values are line=1 and photo=0 which aligns with structure of previous master files which assumed all data was in the form of a line list.

A file structure for photodissociation was already proposed by Pezzella et al. (2022); however, this is updated and extended here to align with the one proposed for VUV spectra in Tennyson et al. (2020); Table 7 gives the formal specification of the file structure. As a file naming convention we adopt the following:

'<ISO-SLUG>__<DSNAME>__<RANGE>__T<TEMP>K__P<PRESSURE>bar__<STEP>.photo', where ISO-SLUG is the iso-slug molecule name, DSNAME is the name of the line list, RANGE is the wavelength range in nm, TEMP is the temperature in K, PRESSURE is the pressure in bar, and STEP is the wavelength step in nm. For example, the nm file of the photodissociation cross-sections for HF the filename: 1H-19F__PTY__90.0-400.1__T200K__P0bar__0.1.nm , see Table 8.

Pezzella et al. (2022) found that their cross-sections depended strongly on the temperature of the molecule and proposed presenting these data for 34 temperatures between $T = 0$ and $T = 10\,000$ K. This data model implicitly assumes that the molecule is in local thermodynamic equilibrium (LTE). We discuss issues with treating non-LTE effects and other issues with this data model in the next section. These data are all implicitly at zero pressure as pressure-broadening effects are neglected. Data from other sources will have different temperature and pressure grids.

Experimental cross-sections of molecules covering the VUV region has been curated by the MPI-Mainz UV/VIS Spectral Atlas (Keller-Rudek et al. 2013) using a similar, wavelength (in nm) format. In Fig. 4, we illustrate photodissociation cross-sections of HCl from

**Table 6.** Photodissociation definition (`.pdef`) file format; each entry starts on a new line.

| Field | Fortran format | C format | Description |
|---|---|---|---|
| **Header Information** | | | |
| ID | A10 | %10 s | Always the ASCII string 'EXOMOL.pdef' |
| IsoFormula | A27 | %27 s | Isotopologue chemical formula |
| Iso-slug | A160 | %160 s | Isotopologue identifier, see text for details |
| DSName | A10 | %10 s | Isotopologue data set name |
| V | I8 | %8 d | Version number with format YYYYMMDD |
| MolKey | A27 | %27 s | Standard inchi key of the molecule |
| $N_{atom}$ | I4 | %4 d | Number of atoms |
| **Atom definition** (The following 2 lines occur $N_{atom}$ times) | | | |
| $I_{atom}$ | I3 | %3 d | Isotope number |
| Atom | A3 | %3 s | Element symbol |
| **Isotopologue information** | | | |
| $m_{Da}$  $m_{kg}$ | F12.6,1X,ES14.8 | %12.6f %14.8e | Isotopologue mass in Da and kg |
| $I_{sym}$ | A6 | %6 s | Molecular symmetry group (if $N_{atom} = 2$ then C or D) |
| $N_{irrep}$ | I4 | %4 d | Number of irreducible representations |
| **ExoMol information** | | | |
| FileType | A27 | %27 s | Data source: ExoMol, Expt., etc. |
| $T_{max}$ | F8.2 | %8.2f | Maximum temperature for cross-sections |
| $N_{temp}$ | I3 | | No. of temperatures available |
| $N_{pres}$ | I3 | | No. of pressures available |

**Table 7.** Specification of the `.photo` photodissociation cross-section file format.

| Field | Fortran format | C format | Description |
|---|---|---|---|
| $\tilde{\lambda}_i$ | F10.3 | %10.3f | Central bin wavelength, nm |
| $\sigma_i$ | ES13.6 | %13.6e | Photodissociation cross-section, $cm^2$ $molec^{-1}$ |

*Note.* Fortran format: (F10.3,1x,ES13.6)

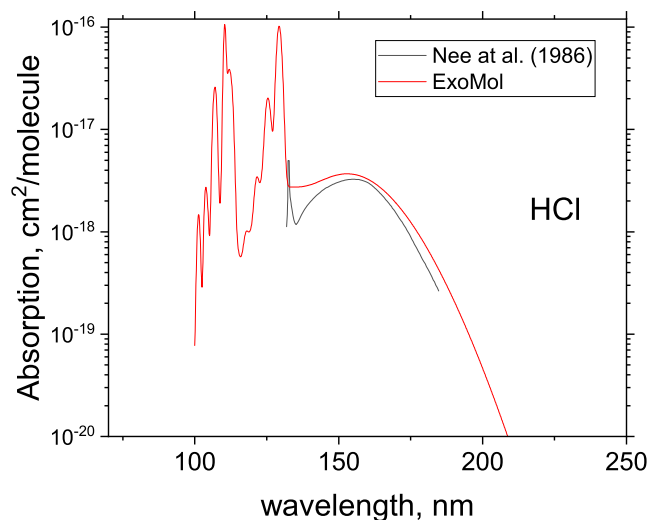**Table 8.** An excerpt from the `.nm` photodissociation file for $^1H^{19}F$, see Pezzella et al. (2022).

| $\tilde{\lambda}_i$ | $\sigma_i$ |
|---|---|
| 90.00 | 4.48427452E-19 |
| 90.10 | 4.50720456E-19 |
| 90.20 | 4.52964494E-19 |
| 90.30 | 4.55158741E-19 |
| 90.40 | 4.57293432E-19 |
| 90.50 | 4.59377114E-19 |
| 90.60 | 4.62624329E-19 |
| 90.70 | 1.68471832E-18 |

*Notes.* $\tilde{\lambda}_i$: Central bin wavelength (nm) and $\sigma_i$: Photodissociation cross-section ($cm^2$ $molec^{-1}$).

ExoMol (Pezzella et al. 2022) and by Nee, Suto & Lee (1986) at room temperature as provided by the MPI-Mainz UV/VIS Atlas.



**Figure 4** Photodissociation spectra of HCl from ExoMol (Pezzella et al. 2022) (theoretical, using natural abundance) and by Nee et al. (1986) (298 K) as provided by the MPI-Mainz UV/VIS Atlas.

## 4 OMISSIONS FROM THE UPDATED DATA MODEL

The assumption of LTE for a molecule undergoing photodissociation may not be valid in all cases. Our method of calculating these cross-sections does indeed involve computing the initial/final-state dependent data which would be required for a non-LTE representation of photodissociation. However, given that even for diatomic molecules a large number of initial states have to be considered, even considering the vibronic states only, the volume of these data is large. As yet no one has asked us for non-LTE photodissociation cross-sections so at present we do not propose including them in the standard data distribution; if they are required they can be obtained from the present authors. Examples of the state-dependent non-LTE cross-sections include the continuous opacities of CH and OH provided by Kurucz et al. (1987) as well as the CH data produced by Popa et al. (2023) and used in their non-LTE analysis of CH in metal-poor stellar atmospheres.

Another issue with our data model for photodissociation is that at present it provides no information of dissociation products. In comparison with the Leiden data base (Heays, Bosman & van Dishoeck 2017), which provides low-temperature photodissociation cross-sections for molecules of astrophysical interest, gives the dissociation products which are given at the threshold to photodissociation but does not provide information on other possible photodissociation products as they may arise at shorter wavelengths. Although our methodology is capable of providing the partial cross-sections (or branching ratios) associated with dissociation to different products, so far our models have not been constructed to produce these data. While the initial step in photodissociation generally involves dipole allowed transition to an electronically excited state, the subsequent dissociation step may involve crossings to states which cannot be reached by dipole transitions from the ground state such as ones with different spin multiplicity. Allowing for these extra states represents a significant complication in the calculation and again, so far, no one has asked for these data. However, should these partial cross-sections be needed it would be relatively simple to extend our proposed data model to accommodate them; we note that the International Atomic Energy Agency's `CollisionDB` data base (Hill 2023; Hill et al., in preparation) already used `PyValem` to address this issue in collision cross-sections.

## 5 CONCLUSION

This research note sets out how we propose to extend the current ExoMol data model to accommodate photoabsorption processes which occur at shorter wavelengths where the possibility of either direct or indirect absorption into the continuum can occur. Broadly these processes are classified as predissociation, continuum absorption, and photodissociation contribution to the opacity. While both predissociation and continuum absorption can be accommodated by generalizing our current line; a new branch starting form a photodissociation definition file has been added for the photodissociation cross-sections.

Given that these various processes considered are not mutually exclusive as, e.g. photodissociation also provides a contribution to the opacity, some care is required in defining data structures to facilitate the use of our extended data. We believe our proposals satisfy this requirement but that further expansion will be required to allow for pressure-dependent continuum absorption, photodissociation of molecules not in LTE, or to account for the possibility that photodissociation might result in creation of a variety of different photodissociation products.

## ACKNOWLEDGEMENTS

## DATA AVAILABILITY

The data discussed in this article are available from the ExoMol data base which can be accessed at www.exomol.com.

## REFERENCES

Anisman L., Chubb K. L., Elsey J., Al-Refaie A., Changeat Q., Yurchenko S. N., Tennyson J., Tinetti G., 2022, J. Quant. Spectrosc. Radiat. Transf., 278, 108013

Bai T., Qin Z., Liu L., 2021, MNRAS, 505, 2177

Baltayan P., Nedelec O., 1979, J. Chem. Phys., 70, 2399

Bowesman C. A., Shuai M., Yurchenko S. N., Tennyson J., 2021, MNRAS, 508, 3181

Fateev A., Clausen S., Wang Y., Owens A., Tennyson J., Yurchenko S. N., 2023, J. Quant. Spectrosc. Radiat. Transf., 251, 107027

Gérard J.-C., Soret L., Hubert B, Neary L, Daerden F, 2023, Icarus, in press

Gorman M., Yurchenko S. N., Tennyson J., 2019, MNRAS, 490, 1652

Hargreaves R. J. et al., 2019, J. Quant. Spectrosc. Radiat. Transf., 232, 35

Heays A.-N., Bosman A.-D., van Dishoeck E.-F., 2017, A&A, 602, A105

Hill C., 2023, CollisionDB, a database of plasma collisional processes. Atomic and Molecular Data Unit, IAEA. https://db-amdis.org/collisiondb

Hill C., Hanicinec M., 2022, A Python package for managing simple chemical species and states, version 2.5.12. https://github.com/xnx/pyvalem

Huang X., Schwenke D. W., Lee T. J., 2021, Acc. Chem. Res., 54, 1311

Karman T. et al., 2019, Icarus, 328, 160

Keller-Rudek H., Moortgat G. K., Sander R., Sörensen R., 2013, Earth Syst. Sci. Data, 5, 365

Kurucz R. L., van Dishoeck E. F., Tarafdar S. P., 1987, ApJ, 322, 992

Lupu R. E., Feldman P. D., McCandliss S. R., Strobel D. F., 2011, ApJ, 732, 37

Nee J. B., Suto M., Lee L. C., 1986, J. Chem. Phys., 85, 719

Owens A., Dooley S., McLaughlin L., Tan B., Zhang G., Yurchenko S. N., Tennyson J., 2022, MNRAS, 511, 5448

Pavlenko Y., Tennyson J., Yurchenko S. N., Jones H. R. A., Lyubchik Y., Suárez Mascareño A., 2022, MNRAS, 516, 5655

Pezzella M., Yurchenko S. N., Tennyson J., 2021, Phys. Chem. Chem. Phys., 23, 16390

Pezzella M., Yurchenko S. N., Tennyson J., 2022, MNRAS, 514, 4413

Popa S. A., Hoppe R., Bergemann M., Hansen C. J., Plez B., Beers T. C., 2023, A&A, 670, A25

Qin Z., Bai T., Liu L., 2021a, ApJ, 917, 87

Qin Z., Bai T., Liu L., 2021b, MNRAS, 508, 2848

Qin Z., Bai T., Liu L., 2022a, MNRAS, 510, 3011

Qin Z., Bai T., Liu L., 2022b, MNRAS, 516, 550

Qu Q., Yurchenko S. N., Tennyson J., 2021, MNRAS, 504, 5768

Rey M., Nikitin A. V., Babikov Y. L., Tyuterev V. G., 2016, J. Mol. Spectrosc., 327, 138

Richard C. et al., 2012, J. Quant. Spectrosc. Radiat. Transf., 113, 1276

Rothman L. S. et al., 2010, J. Quant. Spectrosc. Radiat. Transf., 111, 2139

Tennyson J., Yurchenko S. N., 2012, MNRAS, 425, 21

Tennyson J., Hill C., Yurchenko S. N., 2013, AIP Conf. Proc. Vol. 1545, 6th International Conference on Atomic and Molecular Data and their Applications (ICAMDATA-2012). AIP, New York, p. 186

Tennyson J., Hulme K., Naim O. K., Yurchenko S. N., 2016a, J. Phys. B: At. Mol. Opt. Phys., 49, 044002

Tennyson J. et al., 2016b, J. Mol. Spectrosc., 327, 73

Tennyson J. et al., 2020, J. Quant. Spectrosc. Radiat. Transf., 255, 107228

Venot O. et al., 2018, A&A, 609, A34

Wheeler M. D., Orr-Ewing A. J., Ashfold M. N. R., 1997, J. Chem. Phys., 107, 7591

Yurchenko S. N., Al-Refaie A. F., Tennyson J., 2018a, A&A, 614, A131

Yurchenko S. N., Williams H., Leyland P. C., Lodi L., Tennyson J., 2018b, MNRAS, 479, 1401

This paper has been typeset from a TEX/LATEX file prepared by the author.