The Astrochemistry Low-energy Electron Cross-Section (ALeCS) database I. Semi-empirical electron-impact ionization cross-section calculations and

A&A, 684, A41 (2024)
https://doi.org/10.1051/0004-6361/202348293

Astronomy & Astrophysics

The Astrochemistry Low-energy Electron Cross-Section (ALeCS) database

I. Semi-empirical electron-impact ionization cross-section calculations and ionization rates

Brandt A. L. Gaches[1,2], Tommaso Grassi[3], Stefan Vogt-Geisse[4], Giulia M. Bovolenta[4,5], Claire Vallance[6], David Heathcote[6], Marco Padovani[7], Stefano Bovino[8,9,10], and Prasanta Gorai[1,11]

[1] Department of Space, Earth and Environment, Chalmers University of Technology, Gothenburg 412 96, Sweden
e-mail: brandt.gaches@chalmers.se
[2] Center of Planetary Systems Habitability, The University of Texas at Austin, USA
[3] Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse 1, 85748 Garching bei München, Germany
[4] Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Concepción, Chile
[5] Atomistic Simulations, Italian Institute of Technology, 16152 Genova, Italy
[6] Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford OX1 3TA, UK
[7] INAF–Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy
[8] Departamento de Astronomía, Facultad Ciencias Físicas y Matemáticas, Universidad de Concepción Av. Esteban Iturra s/n Barrio Universitario, Casilla 160, Concepción, Chile
[9] INAF–Istituto di Radioastronomia, Via Gobetti 101, 40129 Bologna, Italy
[10] Chemistry Department, Sapienza University of Rome, P.le A. Moro, 00185 Rome, Italy
[11] Department of Chemistry and Molecular Biology, University of Gothenburg, 41296 Gothenburg, Sweden

Received 16 October 2023 / Accepted 8 February 2024

ABSTRACT

Context. Electron–molecule interaction is a fundamental process in radiation-driven chemistry in space, from the interstellar medium to comets. Therefore, knowledge of interaction cross sections is key. There have been a plethora of both theoretical and experimental studies of total ionization cross sections spanning from diatomics to complex organics. However, the data are often spread over many sources or are not public or readily available.

Aims. We introduce the Astrochemistry Low-energy Electron cross-section (ALeCS) database. This is a public database for electron interaction cross sections and ionization rates for molecules of astrochemical interest. In particular, we present here the first data release, comprising total ionization cross sections and ionization rates for over 200 neutral molecules.

Methods. We include optimized geometries and molecular orbital energies at various levels of quantum chemistry theory. Furthermore, for a subset of the molecules, we have calculated ionization potentials. We computed the total ionization cross sections using the binary-encounter Bethe model and screening-corrected additivity rule, and we computed ionization rates and reaction network coefficients for molecular cloud environments.

Results. We present the cross sections and reaction rates for >200 neutral molecules ranging from diatomics to complex organics, with the largest being $C_{14}H_{10}$. We find that the screening-corrected additivity rule cross sections generally significantly overestimate experimental total ionization cross sections. We demonstrate that our binary-encounter Bethe cross sections agree well with experimental data. We show that the ionization rates scale roughly linearly with the number of constituent atoms in the molecule.

Conclusions. We introduce and describe the public ALeCS database. For the initial release, we include total ionization cross sections for >200 neutral molecules and several cations and anions calculated with different levels of quantum chemistry theory, the chemical reaction rates for the ionization, and network files in the formats of the two most popular astrochemical networks: the Kinetic Database for Astrochemistry, and UMIST. The database will be continuously updated for more molecules and interactions.

Key words. astrochemistry – molecular data – molecular processes – cosmic rays – ISM: molecules

1. Introduction

Observational studies of molecular gas within the Milky Way have revealed a diverse zoo of about 300 molecules, from simple diatomics such as H_2 and CO to ever more complex molecules such as NC_4NH^+ (Agúndez et al. 2023), $NH_2C(O)CH_2OH$ (synglycolamide, Rivilla et al. 2023), 2-C_9H_7CN (2-Cyanoindene, Sita et al. 2022), and H_2CCCHC_3N (cyanoacetyleneallene, Shingledecker et al. 2021). Due to the cold temperatures of the interstellar medium (∼10–50 K), much of the gas-phase chemistry is driven through ion-neutral initiated chemistry (Larsson et al. 2012; Tielens 2013; van Dishoeck 2014). In regions shielded from external radiation, highly energetic charged particles, known as cosmic rays, in particular, protons, secondary electrons, and photons, provide the primary ionization source (Umebayashi & Nakano 1981). Finally, chemical models of ices irradiated by energetic particles or X-ray radiation, such as interstellar dust grains and comets, necessarily need to include

electron-impact ionization cross sections for as many molecules as possible in order to account for electron production and subsequent interactions (e.g. Shingledecker et al. 2020).

Astrochemical models describe the evolution of vast chemical networks over time for a wide range of physical environments, including the effects of gas density, temperature, atomic abundances, and radiation environments. Two key astrochemical databases contain the required reaction rate coefficients: KIDA[1] (Wakelam et al. 2012) and UMIST[2] (McElroy et al. 2013). However, for cosmic-ray ionization, there is a substantial paucity in reaction rate data, with KIDA and UMIST combined only reporting reaction rate coefficients for H, He, C, O, N, H_2, N_2, and CO. In astrochemical models, these reaction rates are expressed in the form

$$\zeta_{\mathrm{m,T}} = c_{\mathrm{m,T}}\zeta_{\mathrm{H_2,T}}, \tag{1}$$

where c_{m} is a scaling factor that relates the rate of the reaction of interest to the total ionization rate, $\zeta_{\mathrm{H_2,T}}$, of H_2. The coefficients are computed with respect to some reference total H_2 ionization rate, typically, $\zeta_{\mathrm{H_2,0}} \approx 3 \times 10^{-17}$ s^{-1}, which is often referred to as the "canonical" or "fiducial" ionization rate.

The coefficients $c_{\mathrm{m,T}}$ in KIDA and UMIST date back to early studies from the 1970s and 1980s (Cravens et al. 1975; Cravens & Dalgarno 1978; Glassgold & Langer 1974; Black 1975). For molecules containing atoms beyond hydrogen and helium, the computed ionization rates often use scaling relations between the high-energy cross sections of the molecules and assumed a Voyager-like proton cosmic-ray spectrum. However, there has been substantial development in the theoretical calculation and experimental measurements of these cross sections. Furthermore, it has been demonstrated that the spectrum of secondary nonthermal electrons is highly sensitive to the local proton cosmic-ray spectral shape (Ivlev et al. 2021). Observations of the ionization rate have demonstrated that the H_2 ionization rate within molecular clouds is more commonly around 10^{-16} s^{-1} (Caselli et al. 1998; van der Tak & van Dishoeck 2000; Neufeld et al. 2010; Indriolo & McCall 2012; Indriolo et al. 2015; Neufeld & Wolfire 2017; Sabatini et al. 2020; Luo et al. 2023a,b; Sabatini et al. 2023), with a decreasing trend toward an increasing column density. This is consistent with a combination of energy losses and possible diffusive transport (e.g. Padovani et al. 2009, 2018; Silsbee & Ivlev 2019; Phan et al. 2023).

A sizable number of studies have reported computed values for the total electron-impact ionization cross sections of a variety of molecules. Public databases such as PlasmaData[3] (Zhong et al. 2021), BEAMDB[4] (Marinković et al. 2017), the US-based National Institute Standards and Technology (NIST) Electron-Impact cross sections for Ionization and Excitation Database[5], and the Japan-based National Institute for Fusion Science (NIFS) datatabase[6] have compiled the cross sections for a wide range of molecules, often for use in plasma physics, medicine, or other industry applications. The latter of these databases comprises a significant number of experimental and evaluated ionization cross sections. Recently, Heathcote & Vallance (2018) and Zhou et al. (2019) presented large datasets of calculated and experimental cross sections, respectively, for a wide range of molecules of astrochemical interest. These datasets and databases represent a significant advancement in the availability of cross-section data for astrochemical use. However, they generally do not include molecules containing heavier atoms, and do not report on computed and recommended reaction rate coefficients either. Furthermore, databases comprising a large number of ionization cross sections often either do not use a standard evaluation process or do not provide the data in a readily accessible way.

A number of different electron-impact total ionization cross sections have been introduced in the literature (see e.g. Kim et al. 1997; Deutsch et al. 2000; Blanco et al. 2010). For large molecules, in which the computing electronic structure can become untenable, various different additivity rules have been introduced, which introduce a sort of algebra for adding atomic ionization cross sections and atomic or molecular orbits. One of the most important models is the binary-encounter dipole and its simplification, the binary-encounter Bethe (BEB) cross section (Kim & Rudd 1994; Kim et al. 1997), which we detail below in Sect. 2.2.3. It has performed well compared to experimental data (e.g., Kim et al. 1997; Zhou et al. 2019; Zhong et al. 2021) and has thus become a standard. All of the above databases also present the BEB cross sections as a major result.

In this paper, we present a new public database containing electron ionization cross sections and cosmic-ray ionization rate coefficients for over 200 molecules of astrochemical interest. We report the cross sections for low-energy electrons (10 eV $\leq E_{\mathrm{e}} \leq$ 5000 eV) for each of these molecules using both a screening-corrected additivity rule (SCAR) and the BEB model cross section, the molecular orbital binding and kinetic energies we used, and the recommended total reaction rate coefficients for molecular cloud environments. Crucially, the cross sections and coefficients are evaluated for nearly all molecules using the same procedure, providing homogeneous datasets to enable better comparisons and consistency.

2. Methods

We present here the methods for the different approaches of calculating the total electron-impact ionization cross sections. Then, quantum chemistry computations are performed, and the resulting ionization rates are calculated. In brief, we calculated three different models of the electron-impact total ionization rate: the SCAR, the BEB, and the damped-BEB (dBEB) model. In particular, we explored the accuracy of the SCAR method because it may become more applicable for larger molecules due to the computational expense required for accurately computing BEB-related cross sections. The latter models require knowledge of the electronic structure, while the former only requires knowledge of the geometry and a basis set of electron-atom ionization cross sections. For our molecule selection, we chose the primary molecules used in astrochemical networks with reaction rate data on the kinetics database KIDA.

We emphasize that the total ionization cross sections below are for single-ionization events. At high energies, multiple ionization events can occur, in particular, through Auger ionization. Nishimura et al. (1999) estimated the possible impact of multiple ionizations by doubling the contribution of the cross section from inner shell electrons and found that their inclusion shifted the peak by 5% and toward higher impact energies. The contributions of multiple ionization events are beyond the scope of this initial release, and they will be considered for future releases. Furthermore, molecules are assumed to be ionized from their ground state.

[1] https://kida.astrochem-tools.org/
[2] http://udfa.ajmarkwick.net/
[3] http://plasma.mathboylinlin.com/
[4] http://servo.aob.rs/emol/
[5] https://www.nist.gov/pml/electron-impact-cross-sections-ionization-and-excitation-database
[6] https://dbshino.nifs.ac.jp/nifsdb/

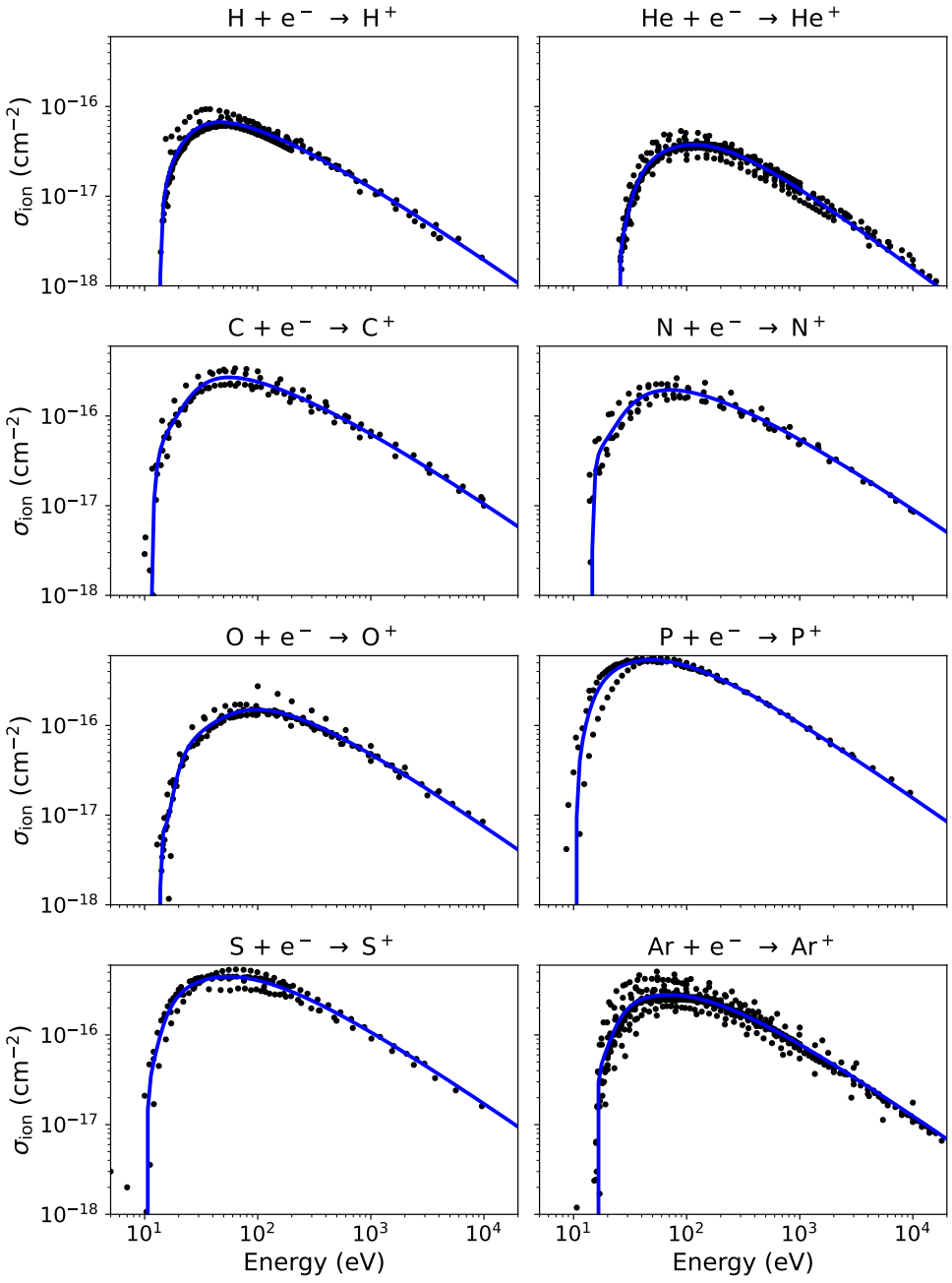


Fig. 1. Electron ionization cross sections. The solid lines give the fits from Eq. (2), and the black points show the data from the NIFS database.

2.1. Electron-atom ionization cross sections

The semi-empirical method we use requires a basis of electron-atom ionization cross sections. We obtain these cross sections by fitting a polynomial to experimental data from the NIFS AMIDIS-ION[5] database, where the electron-atom ionization cross section takes the form

$$\sigma_i(E) = \begin{cases} 0 & \text{for } E_e < \mathrm{IP}_i \\ a_0^2 \left(\frac{x-1}{x}\right)\left[c_0^2\left(\frac{\ln x}{x}\right) + \sum_{k=1}^{5} \frac{c_k}{x^k}\right] & \text{for } E_e \geq \mathrm{IP}_i, \end{cases} \quad (2)$$

where a_0 is the Bohr radius, E_e is the electron energy, IP_i is the ionization potential of the atomic species i, $x = E_e/\mathrm{IP}_i$, and c_k are fitting coefficients. Figure 1 shows the experimental data and the fit cross sections. Table 1 shows the results of the fits of Eq. (2) for each atom.

2.2. Semi-empirical cross-section calculation

2.2.1. Additivity rule

The simplest semi-empirical cross section is the simple additivity rule (AR). Here, the total molecule ionization cross section of species, m, is

$$\sigma_{m,\mathrm{AR}}(E) = \sum_i \sigma_i(E), \quad (3)$$

where the summation is over the different atoms, i. This ignores all possible overlap of the atomic cross sections and electronic structure and always provides an overestimate of the ionization cross section.

2.2.2. SCAR method

The cross sections are computed using the SCAR following Blanco & García (2003); Blanco et al. (2010). This method

Table 1. Electron-atom ionization cross-section fit coefficients for Eq. (2).

Atom	c_0	c_1	c_2	c_3	c_4	c_5
H	2.813905(0)	−1.101730(0)	1.838647(1)	−3.577323(1)	2.554819(1)	−2.358867(0)
He	1.313597(0)	1.261778(1)	−5.359803(1)	1.153092(2)	−1.239072(2)	5.145647(1)
C	7.575591(0)	−5.795490(1)	2.264768(2)	−6.500235(2)	9.067897(2)	−4.246887(2)
N	6.180613(0)	−2.659850(1)	6.022187(1)	−1.036028(2)	1.070141(2)	−1.959618(1)
O	5.400010(0)	4.102912(0)	−1.951275(2)	6.362247(2)	−7.754242(2)	3.372681(2)
P	8.463675(0)	3.806864(1)	−3.042532(2)	8.754706(2)	−1.029666(3)	4.429356(2)
S	9.674459(0)	−5.296533(1)	−8.623544(1)	6.587089(2)	−9.794655(2)	4.823979(2)
Ar	6.618024(0)	9.689090(−1)	−1.248942(2)	5.386474(2)	−7.977861(2)	4.114582(2)

Notes. Coefficients are given in the form $a(b) = a \times 10^b$.

enables the quick computation of molecular cross sections using just a precomputed basis of electron-atom ionization cross sections. SCAR is a generalized AR that takes into account the effective geometrical overlap of the atomic electron cross sections. The cross section for a specific molecule, m, is given by

$$\sigma_{\mathrm{m,SCAR}}(E) = \sum_i s_i \sigma_i(E), \tag{4}$$

where the sum is carried out over all constituent atoms, and $\sigma_i(E)$ is the electron-atom ionization cross section for an atom i. Under the simplest additivity rules, as shown above, $s_i = 1$. For the SCAR method, the additivity coefficients, s_i, are

$$s_i = \sum_{k=1}^{N_k} \frac{(-1)^{k+1}\, \varepsilon_i^{(k)}}{k!}, \tag{5}$$

where N_k is the number of perturbation terms included. The coefficients, $\varepsilon_i^{(k)}$, follow the recursive relation

$$\varepsilon_i^{(1)} = 1 \tag{6}$$

$$\varepsilon_i^{(k)} = \frac{N_k - k + 1}{N_k - 1} \sum_{\mathrm{j}(\neq \mathrm{i})} \frac{\sigma_j \varepsilon_j^{(k-1)}}{\alpha_{ji}}, k = 2, \ldots, N, \tag{7}$$

where $\alpha_{ji} = \max(4\pi r_{ij}^2, \sigma_i, \sigma_j)$, and r_{ij} is the distance between atoms i and j. The terms in Eq. (5) amount to higher-order correction factors accounting for the overlap of the cross sections of all the individual items. Due to the recursive nature of the coefficients (Eq. (6)), including higher-order correction terms leads to an exponential increase in computational cost. However, this is alleviated by building a dictionary cache during the recursion to avoid recomputing the same screening terms. This leads to a sublinear increase in computing time against maximum k-atom screening correction included.

2.2.3. BEB cross sections

The semi-empirical BEB cross section was developed as a simplification of the binary-encounter dipole cross section (Kim & Rudd 1994; Hwang et al. 1996). This cross section has been found to provide a reasonable match with experimental data, and it only requires knowledge of the molecular orbital energies. The BEB cross sections are the base of numerous cross section databases, in particular, the NIST database. The BEB cross section is defined as

$$\sigma_{\mathrm{m,BEB}} = \sum_\ell \left[\frac{S_\ell}{t_\ell + (u_\ell + 1)}\right]\left[\frac{\ln t_\ell}{2}\left(1 - \frac{1}{t_\ell^2}\right)\right. \tag{8}$$

$$\left. + \left(1 - \frac{1}{t_\ell} - \frac{\ln t_\ell}{t_\ell + 1}\right)\right], \tag{9}$$

where the sum is over orbitals, indexed ℓ, $t_\ell = E_\mathrm{e}/B_\ell$, $u_\ell = U_\ell/B_\ell$ and $S_\ell = 4\pi a_0^2 n_\ell \left(\frac{R_\infty}{B_\ell}\right)^2$. Here, B_ℓ and U_ℓ are the orbital binding and orbital kinetic energies of the ejected electron, respectively, n_ℓ is the orbital occupation number, and R_∞ is the Rydberg constant. The parameters B_ℓ and U_ℓ are generally computed using electronic structure methods.

We also propose a modified version of the BEB cross section (dBEB) to dampen the impact of orbitals with binding energies greater than the ionization potential. In dBEB cross section formulation, the B_ℓ is scaled by an exponential function such that the modified binding energy of orbital ℓ is given by

$$B'_\ell = B_\ell \mathrm{e}^{-(1 - B_\ell/\mathrm{IP})}. \tag{10}$$

As we demonstrate below, this prevents the BEB cross section from overestimating with respect to experimental values, which are generally upper limits. In general, BEB cross sections underpredict experimental results at higher energies since they only account for single ionizations, while experimental data, in which the measured signal is generally the ion current, include contributions from multiple ionization events that contribute more to the measured signal per event than single ionizations.

2.3. Quantum chemistry computations

We selected a set of 156 neutral species including atoms and molecules to calculate the BEB and dBEB cross sections. The initial structures were taken from NIST Computational Chemistry Comparison and Benchmark DataBase[7] (CCCBDB) database (Johnson 2022) or ChemSpider[8]. The equilibrium geometries were obtained at MP2/aug-cc-pVTZ method and basis set. To validate the accuracy of this level of theory, we also optimized a group of 50 structures using the highly accurate DF-CCSD(T)-F12/cc-pVDZ-F12 (U-CCSD(T)-F12/cc-pVDZ-F12 for radical species). The set of electron binding energy (eBE), B_ℓ, taken as the negative of the energy of orbital ℓ, was computed by means of electron propagator theory (EPT) in its P3+ implementation along with the aug-cc-pVTZ basis. In EPT, the electron correlation is taken into account, and it thus provides a more accurate value for B_ℓ than Hartree–Fock (HF) canonical orbital energies. To compute the BEB cross sections, we used the EPT-corrected orbital energies when the related pole strength was greater than 0.8, otherwise, we used the HF canonical orbital energies as an estimate of the B_ℓ. Since the lowest B_ℓ, namely the IP, according to Koopmans' theorem, is the predom-

[7] https://cccbdb.nist.gov/
[8] https://www.chemspider.com/

Table 2. Summary of the quantum chemistry methods and references.

Basis sets	aug-cc-pVXZ (X=T,Q) Complete basis set (CBS) extrapolation	Dunning (1989), Dunning et al. (2001) Helgaker et al. (1997)
Gaussian16	Møller-Plesset second order (MP2) CAM-B3LYP Grimme's dispersion (D3) Electron-propagator theory (EPT), P3+	Møller & Plesset (1934), Frisch et al. (1990a,b) Becke (1993), Yanai et al. (2004) Grimme et al. (2010) Migdal & Moszkowski (1968), Ortiz (2005)
PSI4	Coupled cluster (CCSD(T))	Raghavachari et al. (1989)
MOLPRO	Explicitly correlated density fitted CCSD(T) (DF-CCSD(T)-F12)	Gyõrffy & Werner (2018)

inant contribution to the BEB cross section, the accuracy of the eBE plays a significant role.

We computed vertical IPs for 148 molecules in our sample using the density functional theory (DFT) method CAM-B3LYP/aug-cc-pVQZ, employing the D3 empirical dispersion. Furthermore, to validate the accuracy of the IPs we compared the DFT results to highly accurate values that we computed at CCSD(T) paired with complete basis set (CBS) extrapolation using cc-pVXZ (X=D,T) basis functions for a subset of 56 species. This subset consists of molecules with two to four atoms.

The computations were performed using the software packages GAUSSIAN16[9] (Frisch et al. 2016), MOLPRO[10] (Werner et al. 2012, 2020), and PSI4[11] (Smith et al. 2020b, interfaced with the QCFRACTAL[12] infrastructure; see Smith et al. 2020a). Details of the different electronic structure methods are presented in Table 2.

Finally, we included molecular geometries and orbitals from Heathcote & Vallance (2018), which comprise 141 molecules, including 6 anions and 18 cations. These were computed at the HF/aug-cc-pVTZ level, with a small subset at the MP2/aug-cc-pVTZ level (C_4H, C_5H, C_5N, and C_6H; see Heathcote & Vallance 2018 for details). H_2 was computed at the CCSD(T)/aug-cc-pVTZ level. We also included the unpublished calculated data for cations and anions, which were previously excluded because there is a lack of experimental data for these molecules against wich to benchmark, and the method of calculation therefore could not be confirmed. We therefore excluded the anions and cations from the reaction rate plots, although they are included in the full database.

2.4. Low-energy electron spectrum and ionization rate

Low-energy electrons in molecular clouds are primarily generated as secondary particles produced by ionization of H_2 by primary cosmic-ray protons or electrons. We obtained depth-dependent electron spectra, for example, $j_e(E_e, N_{H_2})$, where N_{H_2} is the H_2 column density in the cloud following the new rigorous prescription in Ivlev et al. (2021), with corrections for H_2 excitation and new H_2 electron cross sections (Padovani et al. 2022).

[9] http://gaussian.com/
[10] https://www.molpro.net/
[11] https://psicode.org/
[12] https://github.com/MolSSI/QCFractal

The depth-dependent electron ionization rate of a molecule, m, is given by adding the contributions of primary protons, $\zeta_{p,m}(N_{H_2})$, using the approximation $m_e E_p \approx m_p E_e$,

$$\zeta_{p,m}(N_{H_2}) = 2\pi \int j_p(E_p, N_{H_2})\, \sigma_m \left[(m_e/m_p)E_p\right] \mathrm{d}E_p, \tag{11}$$

primary electrons,

$$\zeta_{e,m}(N_{H_2}) = 2\pi \int j_e(E_e, N_{H_2})\, \sigma_m(E_e)\, \mathrm{d}E_e, \tag{12}$$

and secondary electrons from both primary protons and electrons,

$$\zeta_{se,m}(N_{H_2}) = 4\pi \int j_{se}(E_e, N_{H_2})\, \sigma_m(E_e)\, \mathrm{d}E_e, \tag{13}$$

where *se* represents the secondary electrons. The factors of 2π and 4π account for fluxes, which are assumed to be plane parallel and isotropic, respectively. The total ionization rate for a molecule is then $\zeta_m = \zeta_{p,m} + \zeta_{e,m} + \zeta_{se,m}$. The factor of 4π comes from treating the secondary electrons as an isotropic local source.

We used the "High" primary proton and proton-induced secondary electron spectra as a function of column density from Padovani et al. (2022), who computed the electron spectrum down to 1 eV following the recent more rigorous theory of Ivlev et al. (2021). The High proton spectrum was calibrated to match diffuse gas observations of the H_2 ionization rate, which are not reproduced with a Voyager-like spectrum (Ivlev et al. 2015; Padovani et al. 2018). We also used an interstellar primary electron spectrum from Padovani et al. (2018) and their induced secondary electrons. The secondary electron computation assumes that the gas is fully molecular and includes energy losses from Coloumb interactions, as well as for H_2 ionizations and electronic and rovibrational excitations. We defined our column-dependent total ionization rate coefficient for species m, $c_{m,T}(N_{H_2})$ from

$$\zeta_{m,T}(N_{H_2}) = c_{m,T}(N_{H_2})\, \zeta_{H_2,T}(N_{H_2}), \tag{14}$$

where $\zeta_{H_2,T}(N_{H_2})$ is the total H_2 ionization rate including all primary, secondary, and tertiary processes. We report the column-density average coefficient, $\bar{c}_{m,T}$, for the cloud column density range, $N_{H_2} \approx 10^{20}$–10^{23} cm^{-2}, although we note that the coefficients only marginally scale with column density. Hereafter, we

Table 3. Unique molecules included in this database (202), sorted by their number of constituent atoms.

Atoms	2 atoms	3 atoms	4 atoms	5 atoms	6 atoms	7 atoms	8 atoms	9 atoms	10–15 atoms	15+ atoms
H	C_2	C_2H	C_2H_2	C_2H_2O	C_2H_4	C_2H_4O	$C_2H_3NH_2$	C_2H_5CN	$C_{10}H_2$	$C_{14}H_{10}$
He	CF^+	C_3	C_2N_2	C_2H_3	C_3H_3	C_2H_5	C_2H_6	C_2H_5OH	C_2H_5CHO	C_5H_{10}
Ar	CH	CCN	C_3H	C_2HNO	C_5H	C_6H	C_3H_5	C_3H_6	$C_2H_5OCH_3$	C_5H_{11}
C	CH^+	CCO	C_3H^+	C_3H_2	C_5N	CH_2CCH_2	C_4H_4	C_4H_5	C_3H_5CN	C_5H_{12}
N	CN	CH_2	C_3N	C_4H	CH_2CCO	CH_2CHCN	C_7H	C_5H_4	C_3H_7	C_6H_{10}
O	CO	CNO	C_3O	C_5	CH_2NH_2	CH_3C_2H	CH_2CHCHO	C_8H	C_3H_7CN	C_6H_{11}
S	CO^+	CO_2	C_4	CH_2NH	CH_3CN	CH_3CHO	CH_3C_3N	CH_2CCCCH_2	C_3H_8	C_6H_{12}
P	CP	H_2O	CH_2O	CH_2OH	CH_3NC	CH_3CHS	CH_3COCN	CH_3C_4H	C_4H_{10}	C_6H_{13}
	CS	H_2O^+	CH_3	CH_2PH	CH_3NH	CH_3NCO	CH_3OCH_2	CH_3CONH_2	C_4H_7	C_6H_{14}
	H_2	H_2S	CHNO	CH_2SH	CH_3OH	CH_3NCO	HC_6H	CH_3NHCHO	C_4H_8	C_6H_9
	HF	H_3^+	CO_2H^+	CH_3O	H_2C_3O	CH_3NH_2	$HCOOCH_3$	CH_3OCH_3	C_5H_5	C_7H_{10}
	HS	HCN	H_2CN	CH_3S	H_2CCNH	H_2CCHOH	NC_6N	HC_7N	C_5H_6	C_7H_8
	N_2	HCO	H_2CS	CH_4	HC_2CHO	HC_5N	NH_2CH_2CN		C_5H_7	
	NH	HCO^+	H_2O_2	H_2CCN	HC_3NH^+				C_5H_8	
	NO	HNC	H_3O^+	H_2CCS	HC_4H				C_5H_9	
	NO^+	HNO	HC_2O	H_2NCO^+	HC_4N				C_6H_4	
	NS	HO_2	HCCN	HC_2NC	HCOCHO				C_6H_5CN	
	O_2	HOC^+	$HCNH^+$	HC_3N	HNCHCN				C_6H_6	
	OH	HPO	HCNO	HC_3O	HNCHSH				C_6H_7	
	OH^+	N_2H^+	HCNS	HCCCO	N_2H_4				C_6H_8	
	PH	N_2O	HNCO	HCNCC	NC_4N				C_7H_4	
	PN	NH_2	HNCS	HCOOH	NH_2CHO				C_7H_7	
	PO	NO_2	HOCO	HNC_3	NH_2CHS				C_8H_2	
	S_2	O_3	HSCN	N_2H_3					C_8H_6	
	SO	OCN	HSSH	$NCCNH^+$					$CH_2CHCHCH_2$	
		OCS	NCSH	NH_2CN					CH_2OHCH_2OH	
		PH_2	NH_3	NH_4^+					CH_3C_5N	
		SO_2	NO_3	NHCNH					CH_3C_6H	
			SO_3						CH_3COCH_3	
									CH_3COOCH_3	
									CH_3OCH_2OH	
									$HC_{11}N$	
									HC_9N	
									$HCOOC_2H_5$	

denote $\bar{c}_{m,T}$ as $c_{m,T}$, due to the marginal scaling with column density. In general, since the BEB cross sections perform better than the SCAR data in comparisons against experimental data, we report here only the coefficients using the BEB cross section, although all cross sections are available in the database.

3. Results

In total, we have computed the structure, orbitals, and cross sections for 202 unique (by composition, not counting isomers) neutral molecules ranging in size from two to 24 atoms, shown in Table 3, including data augmented by the results from Heathcote & Vallance (2018). Figure 2 shows the distribution of the number of atoms for the molecules in our database. While most of the molecules have fewer than 6 atoms, we include some with up to 24 atoms ($C_{14}H_{10}$). When KIDA does not specify an isomer but multiple isomers exist, we took the most stable isomer. We detail a summary of the results below, with all of the data available online in the public database[13].

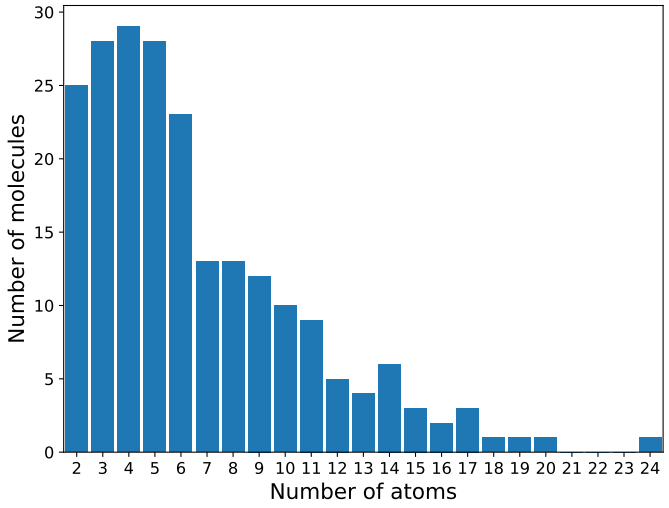


Fig. 2. Distribution of the number of atoms in each molecule for this data release.

3.1. Ionization potentials, electron binding energies, and electron kinetic energies

We present here the ionization potentials computed at both the DFT (CAM-B3LYP) and CCSD(T)/CBS levels, and we compare them to the experimental value recommended in NIST. Table A.1 shows the resulting ionization potentials in eV for the molecules. Figure 3 shows a comparison between the different calculations and the NIST database. In general, our calculations

[13] https://github.com/AstroBrandt/ALeCS

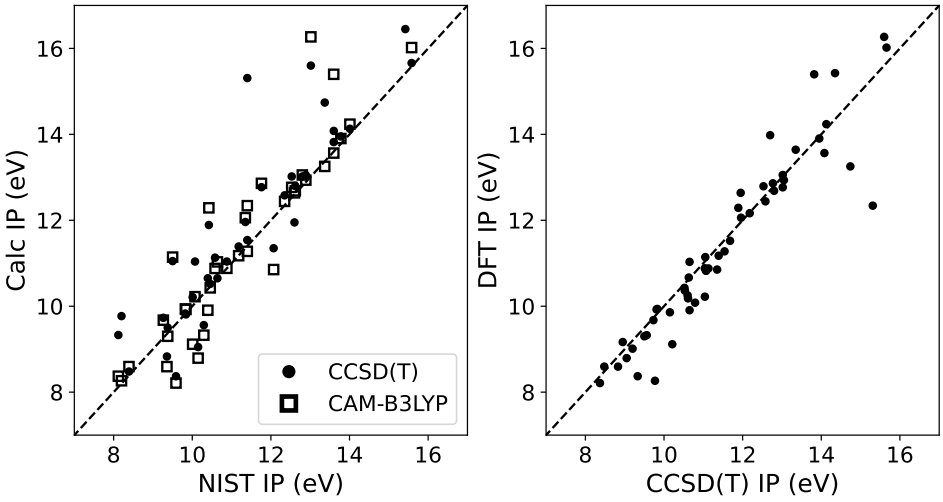


Fig. 3. Comparison of the ionization potentials presented here with the NIST database. Left: NIST database vs. CCSD(T) and CAM-B3LYP DFT calculations (solid and empty squares, respectively). Right: CCSD(T) vs. CAM-B3LYP DFT ionization potentials. The dashed line represents the one-to-one ratio.

for molecules with fewer than seven atoms are consistent with the values presented in the NIST database. For molecules with higher ionization potentials, our calculated IPs are generally higher. The agreement between the CAM-B3LYP and CCSD(T) calculations is relatively good, but there is a general trend that CAM-B3LYP produces smaller IPs compared to CCSD(T) for molecules with IP <10 eV.

As part of the database, we also present the molecular orbital binding and kinetic energies for each molecule. An example result of this for CO and H_2O is presented in Table 4 and Table 5. The tables show the eBEs computed from the canonical HF orbitals and those from the EPT theory. The eBE computed with the latter method are smaller than those obtained with HF since they are corrected to include the instantaneous electron–electron repulsion (electron correlation) in its energy, which increases the energy of the orbitals, and thus facilitates detachment for an electron. According to Koopmans' theorem, the lowest eBE is equal to the ionization potential of the molecule. The values of the EPT-eBE and the IP computed at our best level of theory (CCSD(T)/CBS) agree excellently. The discrepancies are about 0.2 eV and are much improved with respect to HF-eBEs. Given that the lowest eBE carries the most weight in the BEB cross section, it is paramount to ensure that it is accurate. Our results not only exhibit a marked improvement over canonical HF values, but also suggest that the error related to the electronic structure is very small. However, for the deep-lying orbitals, no reliable EPT-eBEs can be computed. We therefore chose to use the HF values for these orbitals.

Our HF-BE data and the NIST electron-impact cross-section database agree well. The database uses data from Hwang et al. (1996) computed at the HF/6-311-G level for the eBE. The agreement is expected because the only difference between the NIST and our HF orbital binding energies is a larger atomic orbital basis set in the latter case. However, the incorporation of EPT-BE into our database enhances the overall data quality compared to the existing NIST data. We thus conclude that our calculations of orbital binding energies are derived from a robust theoretical framework, offering the required level of accuracy given the underlying assumptions in our cross-section calculations.

Finally, we also included the optimized geometries at the MP2/aug-cc-pVTZ and DF-CCSD(T)-F12/cc-pVDZ-F12 levels presented here and the sample from Heathcote & Vallance (2018) in the database. The geometries are not appreciably different from each other. All geometries that represent minimum energy structures were verified by ensuring that no imaginary frequencies in the diagonalized Hessian matrix are present.

Table 4. Molecular orbitals of CO computed at the MP2/aug-cc-pVTZ level.

Orbital number	KE	BE-HF	BE-EPT	PS
1	794.32671	562.41699	***	***
2	436.29414	309.24685	***	***
3	78.26264	41.33263	***	***
4	71.85773	21.8887	***	***
5	53.94510	17.35920	17.015	0.903
6	53.94510	17.35920	17.015	0.903
7	42.78816	15.11922	14.266	0.910

Notes. All energies are in eV. The orbital kinetic energy (KE), the orbital binding energy based on canonical HF orbitals (BE-HF), the orbital binding energy based on EPT computation (BE-EPT), and the corresponding pole strength (PS) are displayed. *** denotes no data because EPT energies for pole strengths lower than 0.8 (electrons deep within the potential) are not considered.

Table 5. Same as Table 4, but for H_2O.

Orbital number	KE	BE-HF	BE-EPT	PS
1	794.31517	559.67367	***	***
2	71.17519	36.81820	***	***
3	48.69143	19.49103	18.872	0.933
4	58.43876	15.92808	14.802	0.925
5	60.77375	13.88698	12.554	0.921

3.2. Low-energy electron cross sections

As described earlier, we have computed the cross sections using three different approaches: the SCAR, the BEB model, and the new dBEB.

The SCAR method is inherently recursive, so that care must be taken to avoid exponential increases in computational time with the number of atoms. We used recursive caching to speed the calculations up, and we only computed to a maximum of the 10-atom screening correction. Figure 4 shows the individual

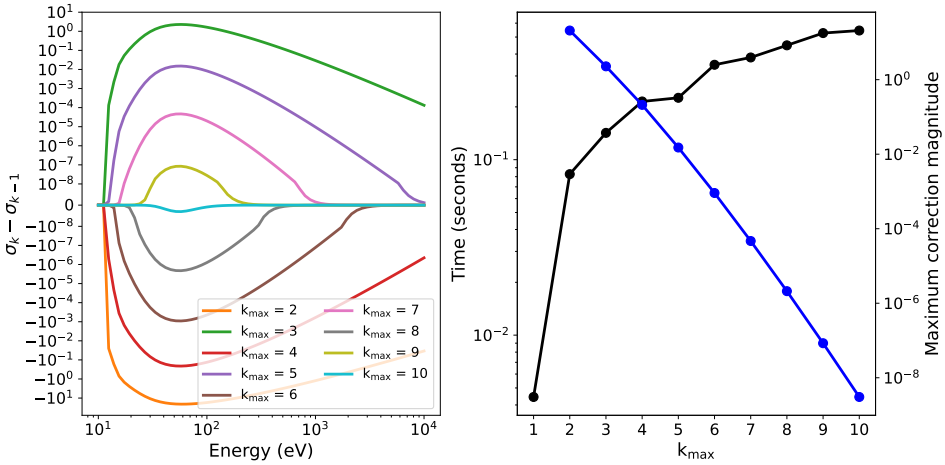


Fig. 4. Benchmark of the number of k-atom screening corrections included using C_5H_{10}. Left: k-atom screening corrections as a function of electron energy for C_5H_{10}. Right: time to compute $\sigma_{SCAR}(E_e)$ for a given maximum k-atom screening correction (black) and the magnitude of the maximum correction (blue) as a function of the maximum k.

screening corrections, the magnitude of the correction, and the total time as a function of the maximum screening terms kept. The screening terms past the 8-atom screening correction are negligible contributions to the total cross section, which is dominated by the first few correction factors.

We show results for a subset of the molecules, ranging from simple to complex, to compare the difference between the SCAR and BEB cross sections. Figure 5 shows them for a subsample of ten molecules, ranging in size from CO to c-C_6H_6. We find that the SCAR approach generally overestimates the cross section, although the impact is most pronounced in c-C_6H_6 because the molecular orbitals overlap. However, given the speed with which these cross sections were computed, they may be useful for first investigations when the optimized molecule geometry is known a priori. While we encourage the use of BEB cross sections, we include a Python script in the database to compute the SCAR cross section from a provided optimized geometry.

We investigated the model predictions for a subset of these molecules in more detail for which experimental data are available. Figure 6 shows a comparison with experimental data for a subset of molecules, including the simple molecule CO_2, symmetric ring c-C_6H_6, prebiotic species of interest NH_2CHO, and carbon chains. Experimental and theoretical cross section data are known to deviate quite significantly (see the discussion in Zhou et al. 2019). In particular, experimental cross sections include double ionizations and Auger ionizations, and they therefore overestimate the single-ionization cross section. The cross sections we present here are for single-ionization events. Multiple ionizations are left for future work. With the data we computed, the BEB slightly overesitmates the cross section in low-energy regions and underestimates it in higher-energy regions. Nonetheless, the values from the BEB calculation agree far better with experimental results than the SCAR or the AR model. Because the experimental curve should serve as an upper limit for the theoretical BEB model, and noting that even with highly accurate eBEs it still overestimates the cross section, we suggest a modified BEB model in which the contribution from deeper-lying electron ionizations are scaled, to better balance its weight in the total cross section. As expected, we find that the dBEB corss-section underpredicts the experimental value. Future calculations using the more complicated binary-encounter dipole model (Kim & Rudd 1994) will be investigated for future releases. In our further analysis, we use the BEB cross sections to calculate the reaction rates because toward higher energies, they tend to agree better with the data.

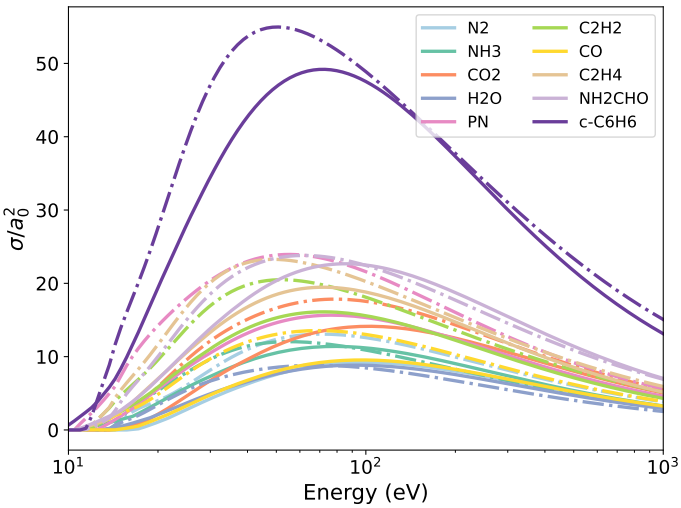
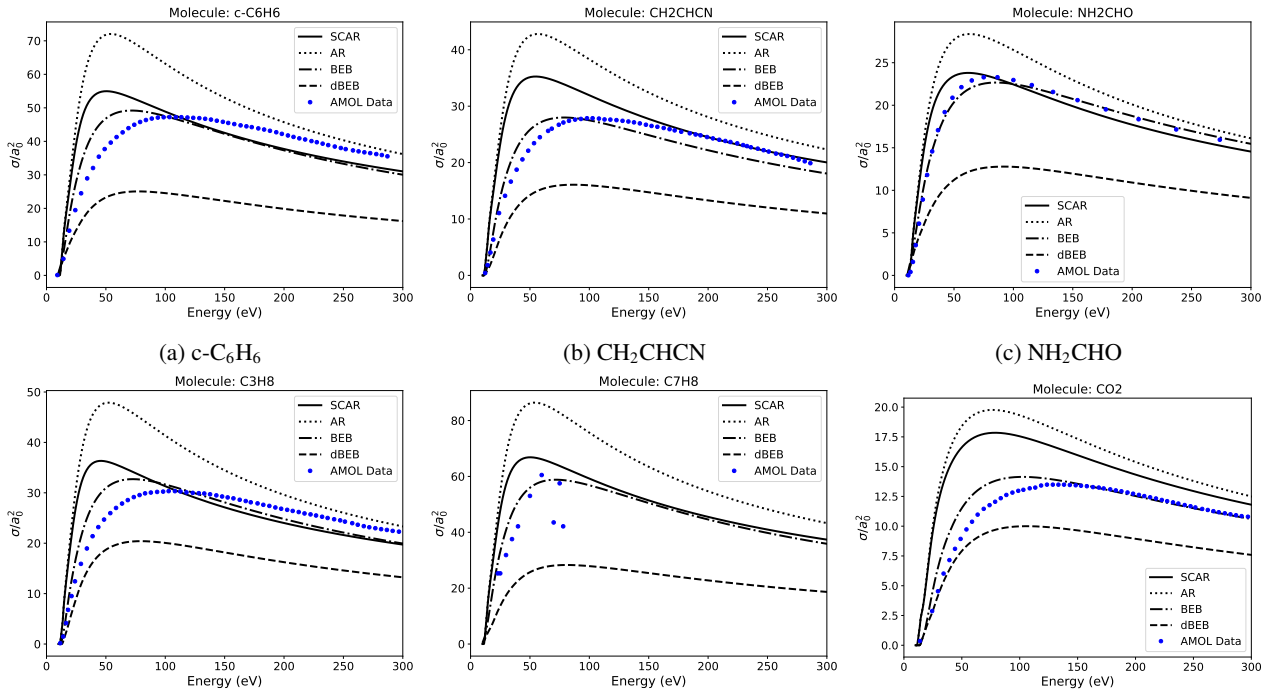


Fig. 5. Sample of the total electron ionization cross sections for a subsample of the molecules. The solid line denotes cross sections calculated using the BEB method, and the dotted line shows the cross sections calculated with the SCAR rule.

3.3. Molecular cloud rate coefficients

We discuss here the molecular cloud reaction rate coefficients, scaled to the total ionization rate (see Sect. 2.4). Figure B.1 show the coefficients $c_{m,T}$ for our database sample. There is a general trend of an increase in $c_{m,T}$ with the size of the molecule. However, for a given number of atoms, there is still substantial scatter due to the geometry and composition. Our reaction rate is marginally greater than that reported by UMIST for CO, originally from Black (1975), where they report $c_{CO,T} = 3$ whereas we find $c_{CO,T} = 3.476$. We report the mean coefficients in Table A.2, where the mean is taken of the coefficients for total hydrogen nucleus column densities ranging from 10^{20}–10^{23} cm^{-2}. We note that these reaction rates are tailored for molecular cloud-like environments, such as those whose prescribed external cosmic-ray spectrum matches the "High" model from Padovani et al. (2022) (see also Ivlev et al. 2015) and whose total hydrogen nucleus column densities lie between $10^{20} \leq N_H \leq 10^{23}$ cm^{-2}.

There is a substantial difference in the reaction rates found with the SCAR and BEB methods. While the additivity rules account for the geometric overlap of the atomic cross sections,



(a) c-C_6H_6 (b) CH_2CHCN (c) NH_2CHO

(d) C_3H_8 (e) C_7H_8 (f) CO_2

Fig. 6. Electron ionization cross sections computed with the AR (dotted), SCAR (solid), and BEB (dash-dotted) and dBEB (dashed) methods compared with experimental data from the NIFS database (blue dots).

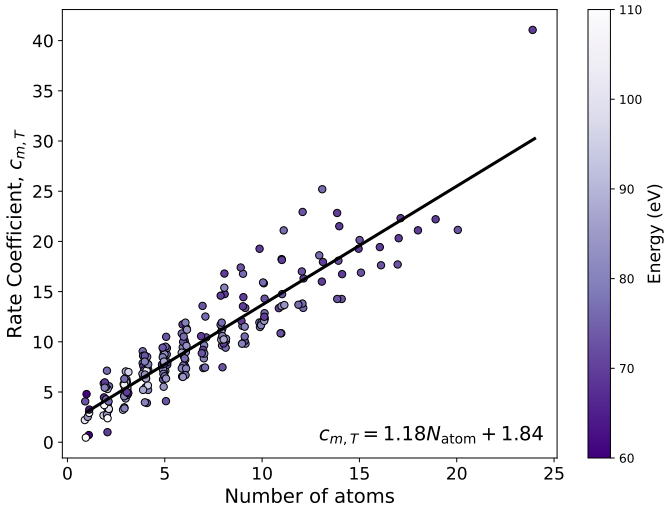


Fig. 7. Total ionization rate coefficient, $c_{m,T}$, as a function of the constituent number of atoms in each molecule. The color denotes the energy at the cross section maximum.

it does not account for the structure of the molecule orbitals nor for differences in eBEs due to molecular bonds. Therefore, we do not recommend the use of additivity rules to compute the cross sections for use in astrochemical modeling, unless it is not practical to compute molecular orbitals and use the BEB cross section.

Figure 7 shows the mean $c_{m,T}$ coefficient as a function of the number of atoms. We find a general increase in the reaction rate with the number of constituent atoms. The figure also shows a best-fit relation,

$$c_{m,T} = 1.19N_{atom} + 1.84, \quad (15)$$

where N_{atom} is the number of constituent atoms. The relation reproduces the general trend for $N_{atom} \leq 20$. We also fit the coefficients as a function of the number of valence electrons, which produces a slightly better fit,

$$c_{m,T} = 0.45N_{v,elec} - 0.64, \quad (16)$$

where $N_{v,elec}$ is the number of valence electrons in the molecule. This fit was only constrained for molecules with $9 \leq N_{v,elec} \leq 70$. In both cases, there is significant scatter of approximately a factor of 2 around the fit trends, so that caution should be used when these are used for molecules that are not listed in the database.

3.4. Database description and file format

We include a number of different formatted files for the total electron-impact cross sections, ionization potentials, molecular and atomic orbitals, and reaction rate coefficients. Cross sections are found in the `ion_xs/` directories for the SCAR, HF, MP2, and CCSD(T) calculations and atoms. The cross-section files are formatted as in the following example, `co.xs`

```
# CO cross section
256
1.00000e+01     0.00000e+00     0.00000e+00
1.04618e+01     0.00000e+00     0.00000e+00
1.09450e+01     0.00000e+00     0.00000e+00
1.14505e+01     0.00000e+00     0.00000e+00
1.19793e+01     0.00000e+00     0.00000e+00
1.25325e+01     0.00000e+00     0.00000e+00
1.31113e+01     0.00000e+00     0.00000e+00
1.37169e+01     0.00000e+00     0.00000e+00
                  {more data}
9.13659e+05     8.52230e-03     6.30487e-03
9.55855e+05     8.17497e-03     6.04801e-03
1.00000e+06     7.84171e-03     5.80155e-03
```

where the first row gives the number of electron energy bins. The later rows give the data in (two) three columns. For BEB cross sections, the columns provide electron impact energy (eV), BEB cross section in units of σ_0^2 and, if available, the damped-BEB cross section in units of σ_0^2. For SCAR cross sections, the three columns provide energy (eV) and SCAR cross section in units of σ_0^2.

The orbital information is stored in two directories containing the orbitals in the NIST format, `NIST_orbitals/` and including all orbitals, `full_orbitals/`. An example of the NIST formatted file, `co.norb`

```
#CO NIST Orbitals
#Orbital      B          U        N    Q
1         562.416986   794.326705  2    1
2         309.246847   436.294142  2    1
3          41.332626    78.262641  2    1
4          21.888866    71.857734  2    1
5          17.015000    53.945104  2    1
6          17.015000    53.945104  2    1
7          14.266000    42.788163  2    1
```

where the B column gives the eBE (eV), U gives the average electron kinetic energy (eV), N gives the orbital occupation number and Q is a scaling factor to include higher-order ionization effects. We also show an example of the full orbital file, `co.forb`

```
#CO Full Orbitals
#Alpha orbitals
#Orb.         KE        BE-HF    BE-P3+      Pole
   1   794.32671    562.41699       ***       ***
   2   436.29414    309.24685       ***       ***
   3    78.26264     41.33263       ***       ***
   4    71.85773     21.88887       ***       ***
   5    53.94510     17.35920    17.015     0.903
   6    53.94510     17.35920    17.015     0.903
   7    42.78816     15.11922    14.266     0.910
```

where the file will give the information for both alpha and beta orbitals for open-shell molecules that were computed with an unrestricted formalism. The KE and BE-HF columns give the average electron kinetic energy and eBE (eV) computed by the population analysis in GAUSSIAN16 following geometry optimization. The latter two columns include the results from EPT, where *** denotes orbitals where EPT was not computed. For orbitals where it was computed, we include the EPT-corrected HF canonical orbital energies (at P3+ level) results and the pole strength (PS).

We include the computed ionization potentials in the `ips/` directory for the compiled NIST and computed CAM-B3LPY and CCSD(T) molecules. These are two-column files with the molecule and ionization potential in eV. We also include two network files containing all the new molecular and atomic ionization rate coefficients in the `networks/` directory. These network files are in the KIDA and UMIST formats, `alecs.kida.in` and `alecs.umist.d`, respectively. We recommend users to only include ionization rates for molecules in which there are associated recombination rates for the ion. In our rate files, we assume that all ionizations occur in the manner $AB + e^- \longrightarrow AB^+ + e^- + e^-$, and we emphasize that potential users should check their chemical networks for other branches such as $AB + e^- \longrightarrow A + B^+ + e^- + e^-$. We leave these differences for future database releases. Finally, in the `geoms/` folder, we include all molecule geometries computed at the HF, MP2, and CCSD(T) levels. The MP2 and CCSD(T) geometries are formatted as PDB files and the HF geometries as XYZ.

Table 6. Initial abundances as a function of the total H nuclei abundance, i.e., $n_{\rm H,tot}$ (Hincelin et al. 2011).

Species	Abundance	Species	Abundance
H_2	5(–1)	P	2(–10)
He	9(–2)	N	6.2(–5)
C	1.7(–4)	O	2.4(–4)
S	8(–8)	Si	8(–9)
Fe	3(–9)	Na	2(–9)
Mg	7(–9)	Cl	1(–9)

Notes. We assume $a(b) = a \times 10^b$.

3.5. Astrochemical modeling

We include the new calculations in a model that uses the KIDA reactions framework, providing a zero-dimensional model as in Sect. 3.1 of Wakelam et al. (2015). It is important to note that in our context, this test has not been designed to quantify the impact of the new rates in an astrophysical environment, but to show that the new rate equations are compatible with previously established formats. Therefore, the cross sections may play a more important role in ice chemistry, disks, planetary atmospheres, exomoons, or cometary environments, for instance. However, this analysis is beyond the aims of the present paper, and it will be discussed in a forthcoming work.

We modeled a gas with total hydrogen density $n_{\rm H,tot} = 2 \times 10^4\,{\rm cm^{-3}}$, gas temperature $T = 10\,{\rm K}$, initial conditions as in Table 6, and no dust (e.g., see Hincelin et al. 2011; Loison et al. 2014). We computed the ionization rate $\zeta(N_{\rm H_2})$ including protons, primary and secondary electrons, and the secondary from primary electrons. The ionization rate is a function of the column density N (see Eq. (14)), and we assumed that the visual extinction (necessary for the photochemistry) is $A_V = 1.0638 \times 10^{-21}\,N$, where N is in units of ${\rm cm^{-2}}$. To model the time-dependent evolution of the chemical abundances for 10 Myr, we made use of KMARX (Grassi in prep., commit `58f6ac9`), a Python-based database that allowed us to solve the chemical ordinary differential equations with a standard BDF solver[14] (Hindmarsh 1983). We include the HTML output produced by the code in the database as a zip file in the `chem_models/` folder, where the reactions present in this work are listed under the class `CRReactionAdv`.

When using our chemical network, we replaced the ionization reactions from KIDA where possible (i.e., H, He, N, O, H_2, and CO) and added the missing ionization rates that affect the molecules listed in Table 6. We did not add the species that create sinks or sources in the chemical network, that is, species that only appear in the products or in the reactants. We also note that for the sake of comparison, the cosmic-ray reaction rates from KIDA were scaled so that the H_2 ionization matches the ionization rate of molecular hydrogen in our database. The aim of this scaling is to avoid discrepancies determined by different assumptions in the cosmic-ray spectra employed.

Figure 8 shows the results obtained with the KIDA database reaction rates (solid lines) and calculated with the rates present in this work (dashed lines). For the sake of clarity, we only plot the species that show a difference larger than one-tenth of an order of magnitude and reach at least $n(t) = 10^{-8} n_{\rm H,tot}$ during their evolution. Only a few species present a negligible discrepancy between the two databases. The extent of these variations

[14] https://github.com/Nicholaswogan/NumbaLSODA

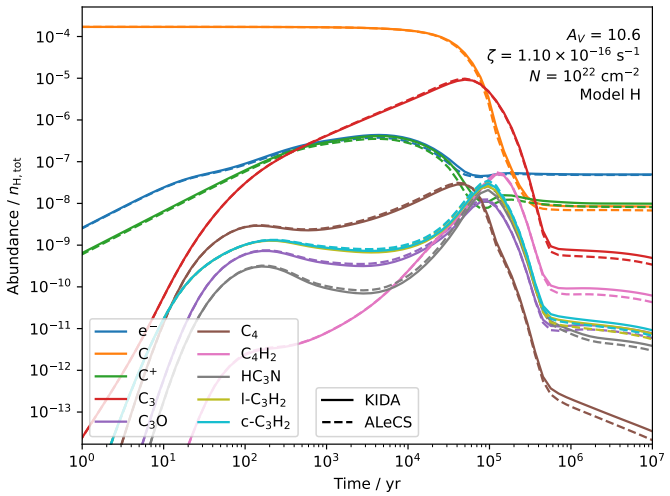


Fig. 8. Comparison of the chemical evolution using KIDA reaction rates (solid lines) and our rates (dashed lines). We report the species that present a difference of at least half an order of magnitude and reach at least $n(t) = 10^{-8} n_{\rm H,tot}$. The ionization rate of H_2 is 1.64×10^{-16} s^{-1}, and the rate for model "High" and with the assumptions described in the text corresponds to $A_V = 10.6$ and $N = 10^{22}$ cm^{-2}.

is very small because the critical cosmic-ray-driven reactions are very similar in ALeCS and in KIDA, but the results might be less interchangeable in different environments. Moreover, we note that we assumed that the constituents of the spectra match, that is, that both include protons, electrons, and secondary processes, hence the scaling mentioned above.

4. Conclusions

We presented the initial data release for the Astrochemistry Low-energy electron cross-section (ALeCS) database. In this release, we include the total ionization cross sections and ionization rate coefficients for over 200 neutral molecules of astrochemical interest calculated using three different semi-empirical methods: the SCAR (Blanco & García 2003; Blanco et al. 2010), the BEB model (Kim & Rudd 1994; Hwang et al. 1996), and a new dBEB presented here. The last model dampens orbitals deep within the potential well and was demonstrated to help prevent the BEB model from overestimating the ionization cross section when compared with experimental results. We also presented the ionization rate coefficients and molecular ionization rates scaled to a reference total H_2 ionization rate.

The database is fully public and will include the ionization data described above, along with the molecule orbitals and optimized geometries. In this current release, we only include the semi-empirical ionization cross sections and reference experimental data. Future releases will include more sophisticated ionization cross-section calculations along with excitation and momentum transfer. Finally, the database will include open-source software tools necessary to couple these processes to astrochemical codes.

Acknowledgements. We thank the referee, Jonathan Tennyson, for their useful comments which improved this work. The MP2 and DFT calculations presented here were performed on the VERA computing cluster, managed by the Chalmers Center for Computational Science and Engineering. BALG and PG are supported by the Chalmers Initiative on Cosmic Origins as Cosmic Origins Postdoctoral Fellows. TG acknowledges support of the DFG (German Research Foundation) Excellence Cluster ORIGINS – EXC-2094 – 390783311. S.B. is financially supported by ANID Fondecyt Regular (project #1220033), and the ANID BASAL projects ACE210002 and FB210003. CV and DH would like to thank the UK Engineering and Physical Sciences Research Council for funding via Programme Grants EP/V026690/1 and EP/T021675/1. S.V.G. is financially supported by VRID project 2022000507INV. GMB gratefully acknowledges support from ANID Beca de Doctorado Nacional 21200180. This is University of Texas Center for Planetary Systems Habitability (CPSH) contribution number #0073.

References

Agúndez, M., Cabezas, C., Marcelino, N., et al. 2023, A&A, 669, L1
Becke, A. D. 1993, J. Chem. Phys., 98, 5648
Black, J. H. 1975, PhD thesis, Harvard University, Massachusetts, USA
Blanco, F., & García, G. 2003, Phys. Lett. A, 317, 458
Blanco, F., Rosado, J., Illana, A., & García, G. 2010, Phys. Lett. A, 374, 4420
Caselli, P., Walmsley, C. M., Terzieva, R., & Herbst, E. 1998, ApJ, 499, 234
Cravens, T. E., & Dalgarno, A. 1978, ApJ, 219, 750
Cravens, T. E., Victor, G. A., & Dalgarno, A. 1975, Planet. Space Sci., 23, 1059
Deutsch, H., Becker, K., Matt, S., & Märk, T. D. 2000, Int. J. Mass Spectrom., 197, 37
Dunning, T. H. J. 1989, J. Chem. Phys., 90, 1007
Dunning, T. H. J., Peterson, K. A., & Wilson, A. K. 2001, J. Chem. Phys., 114, 9244
Frisch, M. J., Head-Gordon, M., & Pople, J. A. 1990a, Chem. Phys. Lett., 166, 275
Frisch, M. J., Head-Gordon, M., & Pople, J. A. 1990b, Chem. Phys. Lett., 166, 281
Frisch, M. J., Trucks, G. W., Schlegel, H. B., et al. 2016, Gaussian~16 Revision C.01 (Wallingford, CT: Gaussian Inc.)
Glassgold, A. E., & Langer, W. D. 1974, ApJ, 193, 73
Grimme, S., Antony, J., Ehrlich, S., & Krieg, H. 2010, J. Chem. Phys., 132, 154104
Győrffy, W., & Werner, H.-J. 2018, J. Chem. Phys., 148, 114104
Heathcote, D., & Vallance, C. 2018, J. Phys. B: At. Mol. Opt. Phys., 51, 195203
Helgaker, T., Klopper, W., Koch, H., & Noga, J. 1997, J. Chem. Phys., 106, 9639
Hincelin, U., Wakelam, V., Hersant, F., et al. 2011, A&A, 530, A61
Hindmarsh, A. C. 1983, IMACS Trans. Sci. Comput., 55
Hwang, W., Kim, Y. K., & Rudd, M. E. 1996, J. Chem. Phys., 104, 2956
Indriolo, N., & McCall, B. J. 2012, ApJ, 745, 91
Indriolo, N., Neufeld, D. A., Gerin, M., et al. 2015, ApJ, 800, 40
Ivlev, A. V., Padovani, M., Galli, D., & Caselli, P. 2015, ApJ, 812, 135
Ivlev, A. V., Silsbee, K., Padovani, M., & Galli, D. 2021, ApJ, 909, 107
Johnson, R. D. I. 2022, NIST Computational Chemistry Comparison and Benchmark Database, nIST Standard Reference Database Number 101, Release 22
Kim, Y.-K., & Rudd, M. E. 1994, Phys. Rev. A, 50, 3954
Kim, Y.-K., Hwang, W., Weinberger, N. M., Ali, M. A., & Rudd, M. E. 1997, J. Chem. Phys., 106, 1026
Larsson, M., Geppert, W. D., & Nyman, G. 2012, Rep. Progr. Phys., 75, 066901
Loison, J.-C., Wakelam, V., Hickson, K. M., Bergeat, A., & Mereau, R. 2014, MNRAS, 437, 930
Luo, G., Zhang, Z.-Y., Bisbas, T. G., et al. 2023a, ApJ, 942, 101
Luo, G., Zhang, Z.-Y., Bisbas, T. G., et al. 2023b, ApJ, 946, 91
Marinković, B. P., Jevremović, D., Srećković, V. A., et al. 2017, Eur. Phys. J. D, 71, 158
McElroy, D., Walsh, C., Markwick, A. J., et al. 2013, A&A, 550, A36
Migdal, A. B., & Moszkowski, S. A. 1968, Am. J. Phys., 36, 855
Møller, C., & Plesset, M. S. 1934, Phys. Rev., 46, 618
Neufeld, D. A., & Wolfire, M. G. 2017, ApJ, 845, 163
Neufeld, D. A., Goicoechea, J. R., Sonnentrucker, P., et al. 2010, A&A, 521, A10
Nishimura, H., Huo, W. M., Ali, M. A., & Kim, Y.-K. 1999, J. Chem. Phys., 110, 3811
Ortiz, J. V. 2005, Int. J. Quant. Chem., 105, 803
Padovani, M., Galli, D., & Glassgold, A. E. 2009, A&A, 501, 619
Padovani, M., Ivlev, A. V., Galli, D., & Caselli, P. 2018, A&A, 614, A111
Padovani, M., Bialy, S., Galli, D., et al. 2022, A&A, 658, A189
Phan, V. H. M., Recchia, S., Mertsch, P., & Gabici, S. 2023, Phys. Rev. D, 107, 123006
Raghavachari, K., Trucks, G. W., Pople, J. A., & Head-Gordon, M. 1989, Chem. Phys. Lett., 157, 479
Rivilla, V. M., Sanz-Novo, M., Jiménez-Serra, I., et al. 2023, ApJ, 953, L20
Sabatini, G., Bovino, S., Giannetti, A., et al. 2020, A&A, 644, A34
Sabatini, G., Bovino, S., & Redaelli, E. 2023, ApJ, 947, L18
Shingledecker, C. N., Incerti, S., Ivlev, A., et al. 2020, ApJ, 904, 189
Shingledecker, C. N., Lee, K. L. K., Wandishin, J. T., et al. 2021, A&A, 652, A12

Silsbee, K., & Ivlev, A. V. 2019, ApJ, 879, 14
Sita, M. L., Changala, P. B., Xue, C., et al. 2022, ApJ, 938, L12
Smith, D. G. A., Altarawy, D., Burns, L. A., et al. 2020a, WIREs Comput. Mol. Sci., e1491
Smith, D. G. A., Burns, L. A., Simmonett, A. C., et al. 2020b, J. Chem. Phys., 152, 184108
Tielens, A. G. G. M. 2013, Rev. Mod. Phys., 85, 1021
Umebayashi, T., & Nakano, T. 1981, PASJ, 33, 617
van der Tak, F. F. S., & van Dishoeck, E. F. 2000, A&A, 358, L79
van Dishoeck, E. F. 2014, Faraday Discuss., 168, 9
Wakelam, V., Herbst, E., Loison, J. C., et al. 2012, ApJS, 199, 21
Wakelam, V., Loison, J.-C., Herbst, E., et al. 2015, ApJS, 217, 20
Werner, H.-J., Knowles, P. J., Knizia, G., Manby, F. R., & Schütz, M. 2012, WIREs Comput. Mol. Sci., 2, 242
Werner, H.-J., Knowles, P. J., Manby, F. R., et al. 2020, J. Chem. Phys., 152, 144107
Yanai, T., Tew, D. P., & Handy, N. C. 2004, Chem. Phys. Lett., 393, 51
Zhong, L., Wu, B., Zheng, S., & Gu, Q. 2021, Phys. Plasmas, 28, 083505
Zhou, W., Wilkinson, L., Lee, J. W. L., Heathcote, D., & Vallance, C. 2019, Mol. Phys., 117, 3066

Appendix A: Additional tables

Table A.1: Ionization potentials in eV, computed at the CCSD(T)/CBS, CAM-B3LYP/aug-cc-pVQZ and from NIST.

Molecule	IP (CCSD(T))	IP (DFT)	IP[15] (NIST)
Diatomic	IP (CCSD(T))	IP (DFT)	IP (NIST)
C_2	15.31	12.34	11.40
CH	10.65	11.03	10.64
CN	13.82	15.40	13.60
CO	14.13	14.24	14.01
CP		11.79	
CS		11.45	
H_2	16.45		15.42
HS	11.89	12.29	10.42
N_2	15.66	16.02	15.58
NH	11.68	11.52	
NO	9.73	9.68	9.26
NS	8.95	9.17	
O_2	11.35	10.85	12.07
OH	15.60	16.27	13.02
PH	9.05	8.79	10.15
PN		11.89	
PO	8.48	8.59	8.39
S_2	8.83	8.59	9.36
SO	9.56	9.33	10.29
Triatomic	IP (CCSD(T))	IP (DFT)	IP (NIST)
C_3	11.95	12.64	12.60
CCN		11.21	
CCO	10.60	10.25	
CH_2	10.65	9.91	10.40
CNO	12.77	12.86	11.76
CO_2	13.95	13.90	13.78
H_2O	12.80	12.69	12.62
H_2S	10.52	10.43	10.46
HCN	14.08	13.56	13.60
HCO	9.33	8.37	8.12
HNC	12.18	12.16	
HNO	10.61	10.18	
HPO	10.53	10.36	
N_2O	13.05	12.93	12.89
NH_2	12.53	12.79	
NO_2	8.37	8.21	9.59
O_2H	11.96	12.06	11.35
O_3	13.02	12.77	12.53
OCN	13.35	13.64	
OCS	11.39	11.18	11.18
PH_2	9.83	9.94	9.82
SO_2	12.58	12.44	12.35
4-atomic	IP (CCSD(T))	IP (DFT)	IP (NIST)
C_2H_2	11.54	11.28	11.40
C_2N_2	14.74	13.25	13.37
C_3N	14.35	15.42	

Table A.1: continued.

Molecule	IP (CCSD(T))	IP (DFT)	IP (NIST)
C_4		10.56	
CH_3	9.81	9.92	9.84
H_2C_3O		10.59	
H_2CN	10.63	10.67	
H_2CO	11.04	10.88	10.88
H_2CS	9.50	9.30	9.38
HC_2N	10.79	10.08	
HC_2O	11.05	11.14	9.50
HCNS	9.20	9.01	
HNCS	10.15	9.86	
HOCO	9.77	8.27	8.20
HOOH	11.13	10.88	10.58
HSCN	11.07	10.83	
HSSH	10.21	9.12	10.01
NCSH	11.07	10.83	
NH_3	11.04	10.22	10.07
NO_3	12.70	13.98	
SO_3	13.02	13.05	12.80
5-atomic	IP (CCSD(T))	IP (DFT)	IP (NIST)
C_2H_3		8.73	
C_2H_3CHO		9.99	
C_4H		14.39	
CH_2NH		9.92	
CH_2OH		7.62	
CH_2PH		10.04	
CH_2SH		7.63	
CH_3O		8.89	
CH_3S		10.24	
CH_4		12.72	
H_2CCN		10.42	
H_2CCS		8.86	
HC_3N		11.48	
HC_3O		11.26	
HCCCO		7.57	
HCCNC		11.04	
HCNCC		11.09	
N_2H_3		7.81	
6-atomic	IP (CCSD(T))	IP (DFT)	IP (NIST)
C_2H_4		10.40	
C_3H_3		8.76	
CH_2CCO		9.02	
CH_2NH_2		6.28	
CH_3NCO		9.59	
CH_3NH		15.09	
CH_3OH		10.80	
HCOCHO		10.16	
HNCHSH		9.71	
N_2H_4		6.82	
NC_4N		11.67	

Table A.1: continued.

Molecule	IP (CCSD(T))	IP (DFT)	IP (NIST)
NH_2CHO		10.13	
NH_2CHS		8.54	
c-H_2C_3O		9.45	
l-HC_4H		9.96	
7-atomic	IP (CCSD(T))	IP (DFT)	IP (NIST)
C_2H_5		8.19	
CH_2CCH_2		7.24	
CH_2CHCN		10.72	
CH_3CHS		8.85	
CH_3NH_2		9.00	
8-atomic	IP (CCSD(T))	IP (DFT)	IP (NIST)
$C_2H_3NH_2$		7.98	
C_2H_6		11.59	
C_3H_5		7.77	
C_4H_4		9.32	
CH_3C_3N		10.59	
CH_3COCN		11.16	
CH_3OCH_2		7.05	
$HCOOCH_3$		10.74	
HNCCC		9.80	
NC_6N		10.69	
l-HC_6H		9.26	
9-atomic	IP (CCSD(T))	IP (DFT)	IP (NIST)
C_4H_5		7.96	
C_5H_4		10.03	
C_8H		8.81	
CH_2CCCCH_2		8.55	
CH_3C_4H		9.26	
CH_3OCH_3		9.85	
l-C_3H_6		9.57	
10-atomic and more	IP (CCSD(T))	IP (DFT)	IP (NIST)
$C_{10}H_2$		8.50	
C_2H_5CHO		9.92	
C_3H_5CN		10.00	
C_3H_7		10.99	
C_3H_8		10.99	
C_4H_{10}		10.89	
C_4H_7		7.58	
C_4H_8		10.89	
C_5H_5		7.80	
C_5H_6		8.74	
C_5H_7		7.42	
C_5H_8		8.35	
C_6H_4		8.74	
C_6H_7		7.31	
C_6H_8		8.70	
C_7H_4		7.98	
C_8H_2		8.81	
C_8H_6		8.65	

Table A.1: continued.

Molecule	IP (CCSD(T))	IP (DFT)	IP (NIST)
$CH_2CHCHCH_2$		8.80	
CH_3C_6H		8.73	
CH_3CCH		10.21	
CH_3CH_2OH		10.38	
CH_3COCH_3		9.64	
c-C_6H_6		9.17	
$C_{14}H_{10}$		7.21	
C_5H_{12}		10.76	
C_6H_{10}		8.27	
C_6H_{12}		9.29	
C_6H_{14}		10.77	
C_6H_9		6.97	
C_7H_8		9.09	

Table A.2: Chemical reaction rates, $c_{\mathrm{m,T}}$

Atom	$c_{\mathrm{m},T}$
H	0.713
He	0.459
C	3.252
N	2.522
O	2.208
P	4.796
S	4.064
Ar	2.919
Diatomic	$c_{\mathrm{m},T}$
C_2	4.061
CF^+	1.383
CH	3.190
CH^+	0.795
CN	3.881
CN^-	28.653
CO	3.476
CO^+	1.401
CP	5.960
CS	5.432
H_2	1.000
HF	2.388
HS	4.271
N_2	3.319
NH	2.740
NO	3.779
NO^+	1.316
NS	5.474
O_2	3.699
OH	2.666
OH^+	0.809
PH	4.465
PN	5.570
PO	5.855

Table A.2: continued.

Molecule	$c_{m,T}$
S_2	7.123
SO	5.300
Triatomic	$c_{m,T}$
C_2H	4.987
C_3	6.306
CCN	6.004
CCO	5.990
CH_2	3.473
CNO	5.790
CO_2	5.287
H_2O	3.232
H_2O^+	1.028
H_2S	4.710
H_3^+	0.306
HCN	4.357
HCO	4.662
HCO^+	1.643
HNC	4.574
HNO	4.783
HO_2	4.792
HOC^+	1.705
HPO	6.258
N_2H^+	1.622
N_2O	5.460
NH_2	3.366
NO_2	5.517
O_3	5.734
OCN	5.625
OCS	7.049
PH_2	4.887
SO_2	6.986
4-atomic	$c_{m,T}$
C_2H_2	5.463
C_2N_2	7.237
C_3N	7.753
C_3N^-	35.329
C_3O	8.028
C_4	8.618
CH_2O	5.212
CH_3	3.906
CHNO	6.348
CO_2H^+	2.713
H_2CN	5.206
H_2CS	7.162
H_2O_2	5.657
H_3O^+	1.281
HC_2O	6.741
HCCN	6.777
HCCO	6.820
$HCNH^+$	1.924

Table A.2: continued.

Molecule	$c_{m,T}$
HCNO	6.692
HCNS	9.062
HNCO	6.345
HNCS	8.507
HOCO	6.558
HSCN	7.877
HSSH	8.610
NCSH	7.878
NH_3	3.971
NO_3	7.138
SO_3	8.382
c-C_3H	7.137
l-C_3H	7.579
l-C_3H^+	2.881
5-atomic	$c_{m,T}$
C_2H_2O	7.272
C_2H_3	6.257
C_2HNO	8.029
C_4H	9.160
C_4H^-	47.898
C_5	10.496
CH_2NH	6.580
CH_2OH	6.342
CH_2PH	7.740
CH_2SH	7.959
CH_3O	5.522
CH_3S	7.711
CH_4	4.090
H_2CCN	7.107
H_2CCS	9.419
H_2NCO^+	4.804
HC_2NC	8.750
HC_3N	8.461
HC_3O	8.760
HCCCO	8.768
HCNCC	9.519
HCOOH	7.030
HNC_3	9.371
N_2H_3	6.805
$NCCNH^+$	3.890
NH_2CN	7.681
NH_4^+	1.469
NHCNH	7.828
c-C_3H_2	7.969
l-C_3H_2	8.009
6-atomic	$c_{m,T}$
C_2H_4	6.552
C_3H_3	8.596
C_4H_2	10.103
C_5H	11.728

Table A.2: continued.

Molecule	$c_{m,T}$
C_5N	11.910
C_5N^-	41.216
CH_2CCO	9.963
CH_2NH_2	7.559
CH_3CN	7.360
CH_3NC	7.644
CH_3NCO	10.829
CH_3NH	6.685
CH_3OH	6.502
H_2CCNH	8.355
HC_2CHO	9.454
HC_3NH^+	4.418
HCOCHO	8.916
HNCHCN	8.976
HNCHSH	9.831
N_2H_4	7.337
NC_4N	11.206
NH_2CHO	8.228
NH_2CHS	10.516
c-H_2C_3O	9.600
l-HC_4H	9.993
l-HC_4N	11.367
7-atomic	$c_{m,T}$
C_2H_5	7.383
C_6H	13.576
C_6H^-	50.847
CH_2CCH_2	9.199
CH_2CHCN	9.646
CH_3C_2H	8.860
CH_3CHO	8.352
CH_3CHS	10.436
CH_3NH_2	7.440
H_2CCHOH	8.824
HC_5N	12.515
c-C_2H_4O	10.563
8-atomic	$c_{m,T}$
$C_2H_3NH_2$	10.056
C_2H_6	7.462
C_3H_5	9.836
C_4H_4	11.289
C_6H_2	14.753
C_7H	16.800
CH_2CCHCN	11.912
CH_2CHCHO	10.575
CH_2OHCHO	10.422
CH_3C_3N	11.691
CH_3CHNH	9.417
CH_3COCN	11.037
CH_3COOH	10.210
CH_3OCH_2	9.636

Table A.2: continued.

Molecule	$c_{m,T}$
$HCOOCH_3$	10.238
NC_6N	15.387
NH_2CH_2CN	10.308
l-HC_6H	14.597
9-atomic	$c_{m,T}$
C_2H_5CN	10.534
C_2H_5OH	9.813
C_4H_5	12.083
C_5H_4	13.373
C_8H	17.401
C_8H^-	54.738
CH_2CCCCH_2	14.448
CH_3C_4H	13.542
CH_3CONH_2	11.505
CH_3NHCHO	11.590
CH_3OCH_3	9.817
HC_7N	16.778
l-C_3H_6	10.079
10-atomic and more	$c_{m,T}$
$C_{10}H_2$	22.934
C_2H_5CHO	11.476
$C_2H_5OCH_3$	13.377
C_3H_5CN	12.852
C_3H_7	10.849
C_3H_8	10.849
C_4H_{10}	14.267
C_4H_7	13.353
C_4H_8	14.267
C_5H_5	14.297
C_5H_6	14.750
C_5H_7	16.314
C_5H_8	15.996
C_5H_9	16.741
C_6H_4	15.829
C_6H_7	17.953
C_6H_8	18.079
C_7H_4	18.154
C_7H_7	22.830
C_8H_2	19.269
C_8H_6	21.514
$CH_2CHCHCH_2$	12.495
CH_2OHCH_2OH	12.313
CH_3C_5N	15.909
CH_3C_6H	18.239
CH_3COCH_3	11.656
CH_3COOCH_3	13.661
CH_3OCH_2OH	12.089
$HC_{11}N$	25.198
HC_9N	21.100
$HCOOC_2H_5$	13.549

Table A.2: continued.

Molecule	$c_{m,T}$
R−CH_3CHCH_2O	11.925
S−CH_3CHCH_2O	11.925
c-C_6H_5CN	18.613
c-C_6H_6	17.007
i-C_3H_7CN	13.695
n-C_3H_7CN	13.813
$C_{14}H_{10}$	41.065
C_5H_{10}	16.887
C_5H_{11}	17.624
C_5H_{12}	17.702
C_6H_{10}	19.434
C_6H_{11}	20.318
C_6H_{12}	21.104
C_6H_{13}	22.204
C_6H_{14}	21.143
C_6H_9	19.272
C_7H_{10}	22.314
C_7H_8	20.136

Appendix B: Additional figures

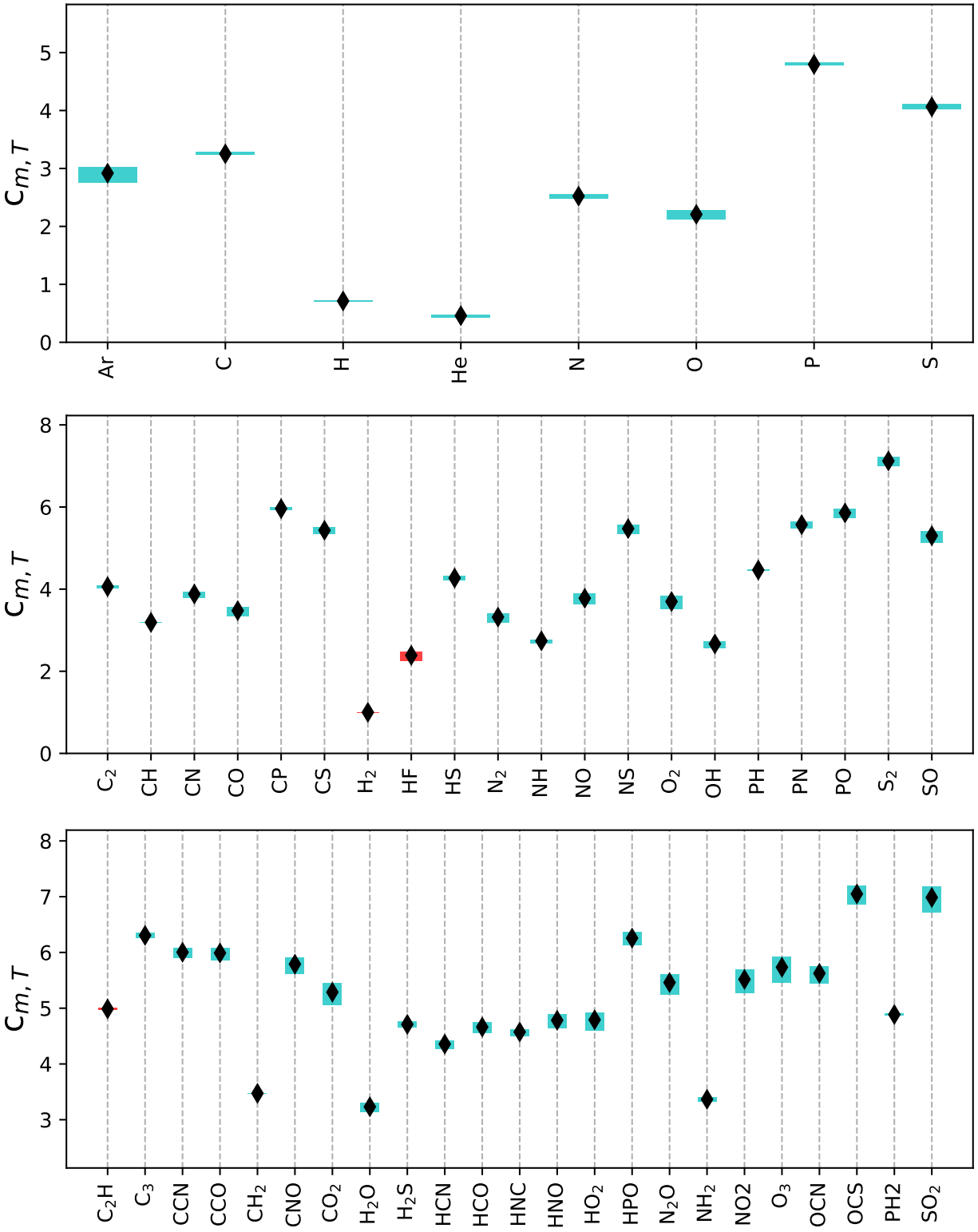


Fig. B.1: Reaction rate coefficients following Equation 14 for all neutral molecules in the database. The boxes show the minimum and maximum coefficients in the column density range 10^{20} cm$^{-2} \leq N_{H_2} \leq 10^{23}$ cm^{-2}, and plus signs show the mean. The cyan boxes denote molecules using the MP2 calculations, and the red boxes denote calculations from Heathcote & Vallance (2018).

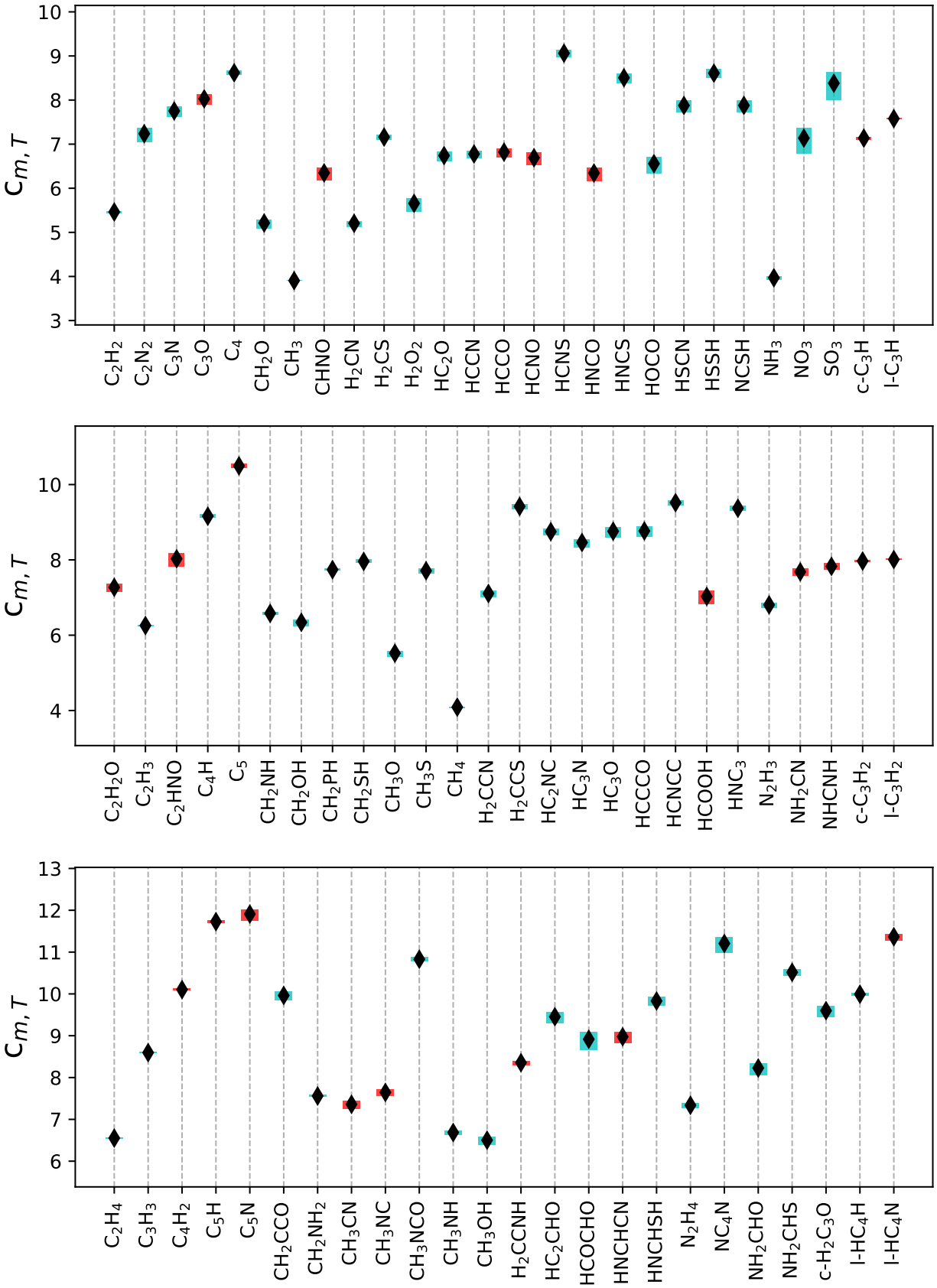

$C_{m,T}$
C_2H_2 C_2N_2 C_3N C_3O C_4 CH_2O CH_3 $CHNO$ H_2CN H_2CS H_2O_2 HC_2O HCCN HCCO HCNO HCNS HNCO HNCS HOCO HSCN HSSH NCSH NH_3 NO_3 SO_3 c-C_3H l-C_3H
C_2H_2O C_2H_3 C_2HNO C_4H C_5 CH_2NH CH_2OH CH_2PH CH_2SH CH_3O CH_3S CH_4 H_2CCN H_2CCS HC_2NC HC_3N HC_3O HCCCO HCNCC HCOOH HNC_3 N_2H_3 NH_2CN NHCNH c-C_3H_2 l-C_3H_2
C_2H_4 C_3H_3 C_4H_2 C_5H C_5N CH_2CCO CH_2NH_2 CH_3CN CH_3NC CH_3NCO CH_3NH CH_3OH H_2CCNH HC_2CHO HCOCHO HNCHCN HNCHSH N_2H_4 NC_4N NH_2CHO NH_2CHS c-H_2C_3O l-HC_4H l-HC_4N

Fig. B.1: cont.

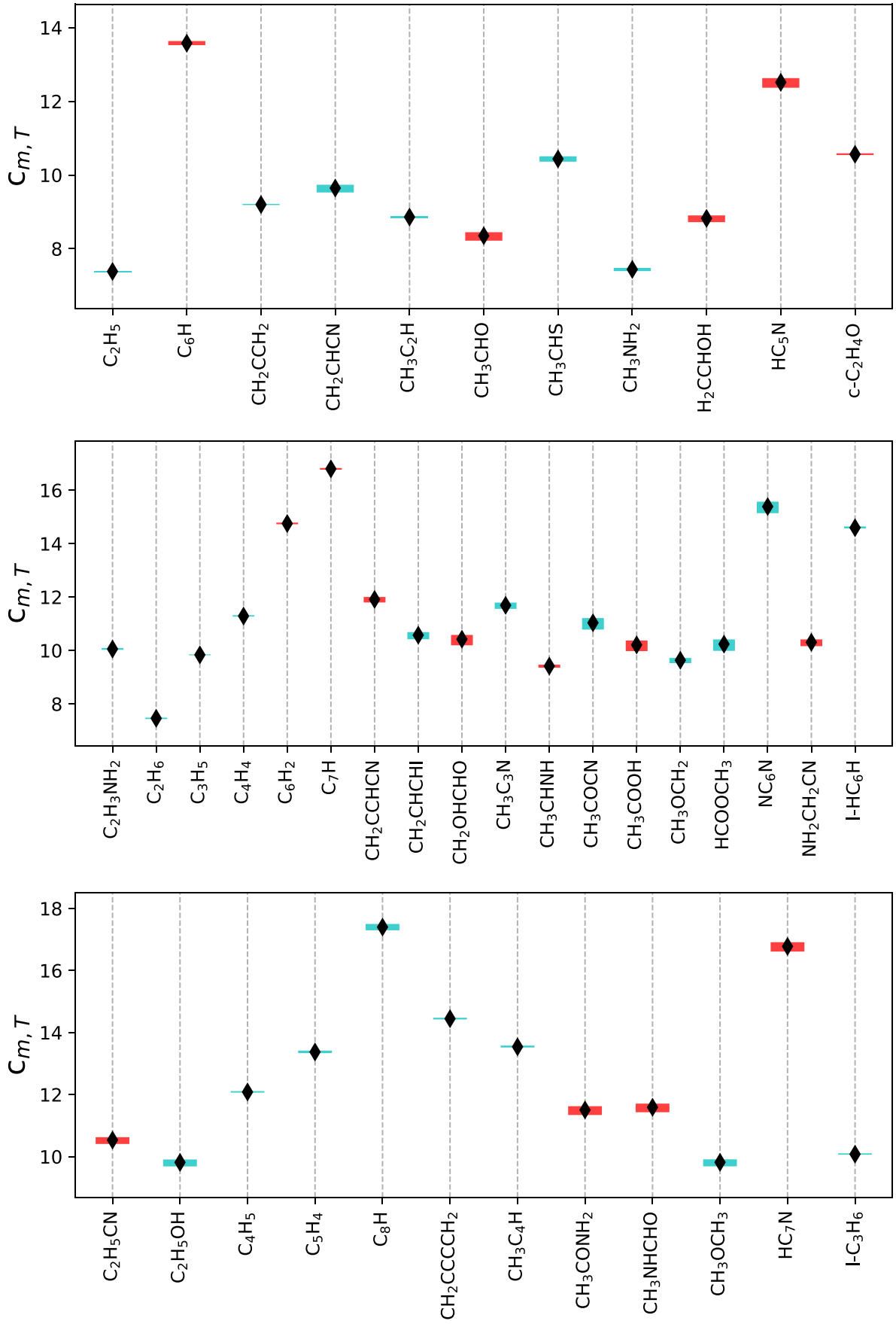

$C_{m,T}$
14
12
10
8
C_2H_5
C_6H
CH_2CCH_2
CH_2CHCN
CH_3C_2H
CH_3CHO
CH_3CHS
CH_3NH_2
H_2CCHOH
HC_5N
$c\text{-}C_2H_4O$
$C_{m,T}$
16
14
12
10
8
$C_2H_3NH_2$
C_2H_6
C_3H_5
C_4H_4
C_6H_2
C_7H
CH_2CCHCN
CH_2CHCHI
CH_2OHCHO
CH_3C_3N
CH_3CHNH
CH_3COCN
CH_3COOH
CH_3OCH_2
$HCOOCH_3$
NC_6N
NH_2CH_2CN
$I\text{-}HC_6H$
$C_{m,T}$
18
16
14
12
10
C_2H_5CN
C_2H_5OH
C_4H_5
C_5H_4
C_8H
CH_2CCCCH_2
CH_3C_4H
CH_3CONH_2
CH_3NHCHO
CH_3OCH_3
HC_7N
$I\text{-}C_3H_6$

Fig. B.1: cont.

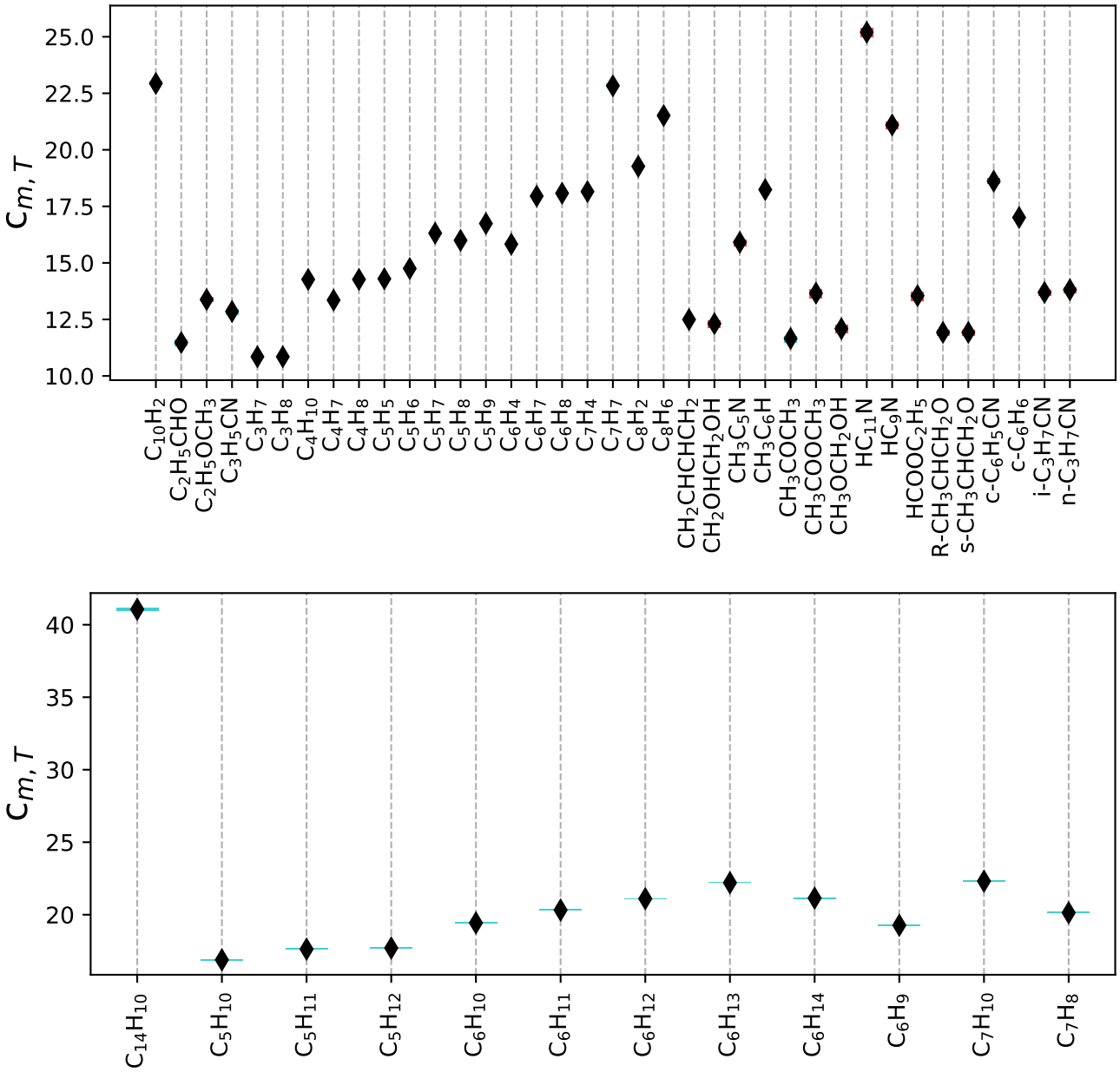

$C_{m,T}$
25.0
22.5
20.0
17.5
15.0
12.5
10.0
$C_{10}H_2$
C_2H_5CHO
$C_2H_5OCH_3$
C_3H_5CN
C_3H_7
C_3H_8
C_4H_{10}
C_4H_7
C_4H_8
C_5H_5
C_5H_6
C_5H_7
C_5H_8
C_5H_9
C_6H_4
C_6H_7
C_6H_8
C_7H_4
C_7H_7
C_8H_2
C_8H_6
$CH_2CHCHCH_2$
CH_2OHCH_2OH
CH_3C_5N
CH_3C_6H
CH_3COCH_3
CH_3COOCH_3
CH_3OCH_2OH
$HC_{11}N$
HC_9N
$HCOOC_2H_5$
R-CH_3CHCH_2O
s-CH_3CHCH_2O
c-C_6H_5CN
c-C_6H_6
i-C_3H_7CN
n-C_3H_7CN
$C_{m,T}$
40
35
30
25
20
$C_{14}H_{10}$
C_5H_{10}
C_5H_{11}
C_5H_{12}
C_6H_{10}
C_6H_{11}
C_6H_{12}
C_6H_{13}
C_6H_{14}
C_6H_9
C_7H_{10}
C_7H_8

Fig. B.1: cont.