frontiers

Protonation of apolar species: from Cl_2H^+ to (*E*)-NCCHCHCNH$^+$ through computational investigations

Silvia Alessandrini [1,*], **Luca Bizzocchi** [2], **Mattia Melosso** [3] **and Cristina Puzzarini** [2]

[1] *Scuola Normale Superiore, Pisa, Italy.*
[2] *Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna, Bologna, Italy.*
[3] *Scuola Superiore Meridionale, Naples, Italy.*

Correspondence*:
Silvia Alessandrini
Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy.
silvia.alessandrini@sns.it

ABSTRACT

Radioastronomy is a powerful tool for the discovery of molecules in space but it requires molecular species to be polar. The observation of apolar species can be however enabled by protonation, which occurs from reaction with the abundant H_3^+ ion whenever the proton affinity of the species under consideration is greater than that of H_2. This property can be easily investigated by computational chemistry and, in this work, it has been used to asses the potential protonation of simple homo diatomics, such as Cl_2H^+, P_2H^+, and Si_2H^+, as well as apolar species containing two equivalent CN moieties, such as diisocyanogen (CNNC) and (*E*)-1,2-dicyanoethene. Quantum chemistry has also been exploited to investigate the mechanisms of three protonation reactions of H_3^+ with Cl_2, P_2, and CNNC. To support laboratory measurements and astronomical observations of the resulting transient species, the rotational spectroscopic parameters were accurately computed together with their fundamental vibrational frequencies. For this purpose, we have employed CCSD(T)-based computational methodologies, which provide equilibrium structures with errors better than 0.001 Å and 0.1° for bond distances and angles, respectively. Such an accuracy is expected to lead to rotational constants predicted, in relative terms, with uncertainties better than 0.2%. Instead, the expected accuracy on vibrational frequencies is $\sim$10 cm^{-1}, thus being well suited to guide band assignments.

Keywords: astrochemistry, protonation, rotational spectroscopy, apolar molecules, ISM, proxy, quantum-chemistry

1 INTRODUCTION

The overwhelming majority of chemical species in the interstellar medium (ISM) have been discovered via radioastronomy, which exploits their rotational signatures. However, the observation of new astronomical species comes to a setback when the species of interest are weakly polar or apolar (Yamamoto, 2017). In the former case, a possible way out is offered by the line stacking technique (Yen et al., 2016), while for apolar molecules, one has to rely on the observation of the so-called molecular proxies. These are

polar systems whose presence in the astronomical object under consideration can be directly related to non-observable species. Molecular proxies can be obtained in several ways, with some of them being well established in the literature.

A first possibility is to claim the presence of an apolar species based on the observation of its cyano-derivatives, which usually have a significant electric dipole moment. Indeed, the high abundance of the CN radical together with its highly reactive nature leads to species containing the CN moiety through formation pathways that are usually accessible even in the conditions of the ISM. A significant example is provided by the observation of benzonitrile (C_6H_5CN) (McGuire et al., 2018), whose detection in the ISM has been undoubtedly related to the presence of benzene in the same environment. Indeed, different experiments have demonstrated the effectiveness of the reaction between the CN radical and benzene C_6H_6 (Lee et al., 2019; Cooke et al., 2020). Similarly, the observation of cyanonaphthalene towards TMC-1 (McGuire et al., 2021) suggests the presence of naphthalene in this cold core. Other than the CN radical, proxies can also be formed by the CCH radical. However, the observation of its derivatives is more challenging because the ethynyl group leads to molecular proxies with a smaller electric dipole moment when compared to the CN-containing counterpart (Cernicharo, J. et al., 2021; McCarthy et al., 2021).

Another important way to observe apolar species is to consider their protonated form. Protonation in the ISM occurs owing to the presence of H_3^+, which in turn is obtained from the reaction between H_2 and H_2^+, with the latter being formed via cosmic ray ionization (Yamamoto, 2017). The H_3^+ initiates the protonation via this general mechanism:

$$H_3^+ + A \rightarrow AH^+ + H_2 \ , \tag{1}$$

which proceeds only if the proton affinity (PA) of the generic species A is greater than that of H_2. Based on the PA values reported by Yamamoto (2017), H_3^+ is the primary source of N_2H^+ and HCO^+/HOC^+, but it does not lead to the formation of O_2H^+ because O_2 has a smaller PA than H_2.

In the present paper, we propose the use of protonated species as suitable proxy for the radio-observation of apolar species of astrochemical interest. To this end, their computed rotational spectra are reported here for the first time. This study will involve simple diatomic molecules, such as P_2, Si_2, and Cl_2, a linear centrosymmetric species, i.e., CNNC, and a C_{2h} species, the *E* isomer of NC−CH=CH−CN. The first three are simple diatomic molecules that can be considered as reservoirs of atomic P, Si, and Cl, respectively. This is particularly interesting because the chemistry of third-row atoms in astronomical environments is still in early stages, with only a few species containing P, Si, or Cl being detected so far (McGuire, 2022). Since both vibrational (infrared) and rotational spectra of homonuclear diatomics cannot be observed, the only way-out for their detection (if avoiding electronic spectroscopy) is offered by their protonated forms. As an example, the N_2H^+ molecule – observed for the first time in 1974 by Turner (1974) – is considered a good tracer of N_2, which is otherwise not observable at radio frequencies. For this reason, the study of the simple P_2H^+, Si_2H^+, and Cl_2H^+ ions can potentially provide useful data to guide laboratory and/or astronomical observations. While the P_2 and Si_2 precursors might be considered somewhat exotic, Cl_2 is a simple species stable under Earth conditions. To assess the accuracy of our work, the N_2H^+ and HCO^+ ions will be used as references to test our computational strategy.

Another interesting molecule is the linear species diisocyanogen (CNNC) because two isomers of its family, namely CNCN and NCCN, have already been detected in the ISM. While the former was observed directly via its rotational transitions (Agúndez et al., 2018), the presence of NCCN was inferred thanks to

the observation of its protonated counterpart, i.e NCCNH$^+$ (Agúndez et al., 2015). Thus, the computational spectroscopic characterization of CNNCH$^+$ surely represents a preliminary step toward its astronomical detection. In this case, the well-characterized NCCNH$^+$ species will be used as reference to understand the uncertainties associated with our methodology.

The last molecule considered is the (*E*) isomer of 1,2-dicyanoethene, which is the main product of the reaction between the CN radical and vinyl cyanide (CH_2=CH−CN, Marchione et al. (2022)) also known as acrylonitrile. While the (*Z*) isomer of 1,2-dicyanoethene, being a polar molecule, might be detected directly, without the use of any molecular proxy, the (*E*) form lacks of a permanent electric dipole moment. Thus, we aim at providing good estimates for the rotational parameters of its counterpart, i.e. (*E*)-NC−CH=CH−CNH$^+$.

This work will also report the fundamental vibrational frequencies and intensities of all the species mentioned above. This data could provide a reliable starting point for experimental measurements of their vibrational spectra as well as benchmark reference for future computations.

The manuscript is organized as follows. In the next section (§ 2), we detail the computational strategy adopted to describe the protonated species, with specific focus on the computation of proton affinity and formation routes as well as spectroscopic parameters. Then, the results are reported in § 3 where they are discussed together with those of chemically related species. After briefly addressing the PA values, the formation routes for the species considered are detailed, and then the computational spectroscopic characterization is considered. The last section (§ 4) will draw the main conclusions of the present study.

2 COMPUTATIONAL DETAILS

For most of the species considered in this work, experimental and computational data are missing, thus a preliminary study using density functional theory (DFT) was carried out to understand their stable form. For this purpose, the double-hybrid rev-DSD-PBEP86 functional (Santra et al., 2019), combined with the GD3BJ empirical dispersion (Grimme et al., 2011, 2010), was used in conjunction with the jun-cc-pVTZ basis set (Dunning Jr., 1989; Papajak and Truhlar, 2011). Overall, this level of theory is shortly denoted as revDSD/junTZ. Noted is that, for third-row elements, the d-augmented jun-cc-pV(T+d)Z sets have been employed (Dunning et al., 2001). The revDSD/junTZ level has been employed to derive preliminary values for equilibrium geometries, electronic energies and (harmonic) zero-point energy corrections (ZPEs).

For Cl_2H^+, P_2H^+, and Si_2H^+, equilibrium structures ranging from linear to cyclic were considered and the most stable forms of each species are shown in fig. 1, where their electronic state is also reported. The same figure also provides the equilibrium structure of CNNCH$^+$ and (*E*)-NC−CH=CH−CNH$^+$, which were somewhat more straightforward to derive since bent structures would be very unstable. After this preliminary study, additional computations were mandatory to derive accurate spectroscopic parameters, PA estimates and formation routes.

2.1 Proton Affinity and Formation Routes

To understand if a protonated molecule can be formed in the ISM starting from H_3^+, the PA value of each species has to be considered in comparison with that of the H_2 molecule. The PA at T=0 K is computed as (McNaught et al., 1997; Marino et al., 2000):

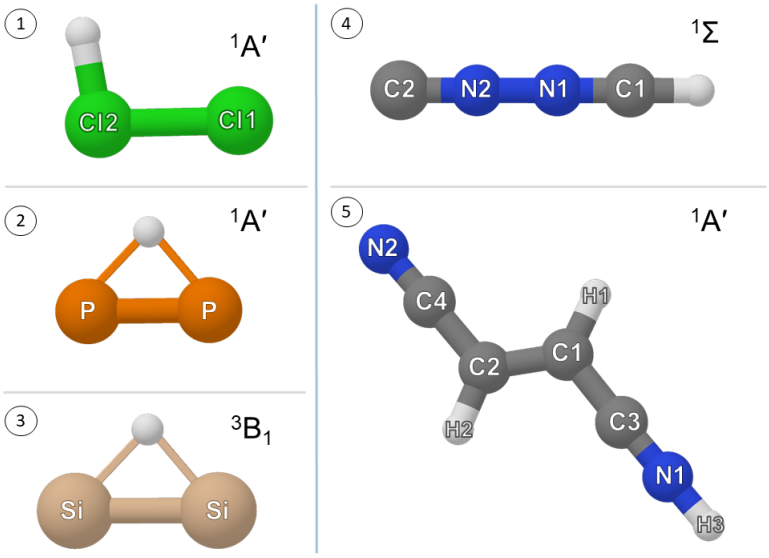

Figure 1. Equilibrium structure and electronic ground state of (1) Cl_2H^+, (2) P_2H^+, (3) Si_2H^+, (4) $CNNCH^+$, and (5) protonated (*E*)-1,2-dicyanoethene.

$$\mathrm{PA} = \Delta H_0^{\mathrm{neutral}} - \Delta H_0^{\mathrm{protonated}} \tag{2}$$

where both ΔH_0 values are obtained by combining the electronic energy obtained using the jun-Cheap composite scheme (hereafter junChS) (Alessandrini et al., 2019) with the anharmonic ZPE. The latter correction was computed at the ae-CCSD(T)/cc-pwCVQZ for protonated diatomics, while the cc-pwCVTZ basis set was used for $NCCNH^+$ and $CNNCH^+$. For protonated 1,2-dicyanoethene, the ZPE was evaluated at the revDSD/junTZ level. Here, ae-CCSD(T) stands for the coupled-cluster (CC) method including a full treatment of single and double excitations and a perturbative estimate of triples (Raghavachari et al., 1989), with all electrons (ae) correlated. Anharmonic computations were carried out on top of the corresponding reference geometry, on top of which junChS energy evaluations were also performed. Within the junChS approach, the electronic energy, $E(\mathrm{junChS})$, is obtained as follows (Alessandrini et al., 2019):

$$E(\mathrm{junChS}) = E(\mathrm{CCSD(T)/junTZ}) + \Delta E_{\mathrm{MP2}}^{\infty} + \Delta E_{\mathrm{MP2}}^{\mathrm{CV}} \quad . \tag{3}$$

Here, the first term on the right-hand side is the fc-CCSD(T)/jun-cc-pVTZ energy (with jun-cc-pV(T+d)Z being used for third-row elements), where fc indicates that the frozen-core approximation is used. The second term accounts for the extrapolation to the complete basis set (CBS) limit at the MP2 level, with MP2 denoting Møller-Plesset perturbation theory to the second order (Møller and Plesset, 1934). This extrapolation exploits the n^{-3} formula by Helgaker et al. (1997), and requires energy computations with two consecutive members of a hierarchical family of basis sets. In this case, the jun-cc-pV(T+d)Z and jun-cc-pV(Q+d)Z sets were used. The third term in eq. (3) is the core-valence (CV) correction and incorporates the effect of the correlation of inner-shell electrons, which is not included in the previous computations. This term is computed as the difference between all-electron and frozen-core calculations, both with the cc-pwCVTZ basis set (Peterson and Dunning Jr., 2002).

For three of the species considered, namely Cl_2H^+, P_2H^+, and $CNNCH^+$, the formation route was also investigated. To this end, the revDSD/junTZ level of theory was employed to locate minima (MINs) and

transition states (TSs) on reactive potential energy surface (PES). The correct assignment of two MINs to the connecting TS was assured by exploiting the intrinsic reaction coordinate (IRC) analysis (Fukui, 1981), performed at the same level of theory. To refine the electronic energy of each stationary point, the junChS approach was employed, while the harmonic ZPE was evaluated at the revDSD/junTZ level.

2.2 Spectroscopic Characterization

Focusing on rotational spectroscopy, the main parameters influencing the spectrum are the rotational constants (Barone et al., 2021). For a given inertia axis ($\gamma = a$, b, or c), in the framework of vibrational perturbation theory to second order (VPT2) (Mills, 1972), the rotational constant of the vibrational ground state (B_0^γ, corresponding to A_0 for $\gamma = a$, B_0 for $\gamma = b$, and C_0 for $\gamma = c$), can be written as:

$$B_0^\gamma = B_e^\gamma - \sum_r \frac{d_r}{2}\alpha_r^\gamma \ . \tag{4}$$

The first term on the right-hand side of eq. (4) is the equilibrium rotational constant, which entirely depends on the isotopic composition of the molecule and its equilibrium structure. Thus, B_e^γ is straightforwardly obtained from a geometry optimization and accounts for about 99% of the value of B_0^γ (Puzzarini et al., 2008, 2010; Alessandrini et al., 2018). The second term on the right-hand side of eq. (4) is the vibrational contribution to the rotational constant and it is obtained as the sum over all r vibrational modes of the vibration-rotation interaction constants (α_r^γ). In eq. (4), d_r denotes the degeneracy of the r-th vibrational mode. The vibrational correction represents a minor contribution to B_0^γ, but its consideration is crucial to reach accurate computational results, comparable with experimental ones (Puzzarini et al., 2008, 2010; Alessandrini et al., 2018).

Since B_e^γ is the major contribution to B_0^γ, it should be computed at the highest possible level of theory when aiming at high accuracy. For this reason, for Cl_2H^+, P_2H^+, Si_2H^+, and $CNNCH^+$, the equilibrium geometry was obtained by exploiting a composite scheme entirely based on coupled-cluster (CC) theory (Heckert et al., 2005, 2006), which requires the minimization of the following energy gradient:

$$\frac{dE_{\mathrm{CBS+CV+fT+fQ}}}{dx} = \frac{dE^\infty(\mathrm{HF-SCF})}{dx} + \frac{d\Delta E^\infty(\mathrm{CCSD(T)})}{dx} + \frac{d\Delta E(\mathrm{CV})}{dx} + \frac{d\Delta E(\mathrm{fT})}{dx} + \frac{d\Delta E(\mathrm{fQ})}{dx} \ . \tag{5}$$

This composite scheme is denoted as "CBS+CV+fT+fQ" and it consists of 5 different terms:

- The extrapolation to the CBS limit, which is obtained by the sum of the HF-SCF and fc-CCSD(T) extrapolated gradients, i.e. $\frac{dE^\infty(\mathrm{HF-SCF})}{dx}$ and $\frac{d\Delta E^\infty(\mathrm{CCSD(T)})}{dx}$, respectively. The former is evaluated using the three-point formula by Feller (1993) and the cc-pVTnZ basis sets, with n=T,Q,5. For the CCSD(T) correlation energy, ΔE(CCSD(T)), gradient the n^{-3} formula (Helgaker et al., 1997) is exploited in conjunction with the cc-pVTZ and cc-pVQZ basis sets.
- The CV contribution, which is obtained as difference between ae- and fc-CCSD(T)/cc-pwCVTZ computations.
- The fT term, which accounts for the full treatment of triples. This is obtained as difference between two computations carried out with the cc-pVTZ basis set, the first employing the CCSDT (Noga and Bartlett, 1987; Scuseria and Schaefer III, 1988) method (the acronym standing for CC singles, doubles and triples) and the second one with CCSD(T), both within the fc approximation.

- The fQ term, which is the contribution due to quadruple excitations. This is computed as difference between CCSDTQ (CC singles, doubles, triples and quadruples; Kállay and Gauss (2008); Kucharski and Bartlett (1991)) and CCSDT calculations, both with the cc-pVDZ basis set and within the fc approximation.

The CBS+CV+fT+fQ equilibrium rotational constants were then corrected for the vibrational contribution. The evaluation of vibration-rotation interaction constants requires a full cubic force field (Puzzarini et al., 2010), which was computed at the ae-CCSD(T)/cc-pwCVQZ level. For the CNNCH$^+$ ion, a smaller basis set, cc-pwCVTZ, was instead employed. For the (*E*)-NC$-$CH$=$CH$-$CNH$^+$ species, the computational approach was slightly different because of its size. The equilibrium rotational constants were obtained from the minimization of the "CBS+CV" energy gradient, thus retaining only the first three terms on the right-hand side of eq. (5) in conjunction with the same basis sets mentioned above. In addition, the anharmonic computation required for the vibrational corrections was carried out at the revDSD/junTZ level of theory.

For all the species considered, the VPT2 calculations performed on the computed anharmonic force field allowed for obtaining other spectroscopic parameters, such as quartic centrifugal distortion constants and fundamental vibrational frequencies. These latter can be obtained because the computations of the cubic force constants give access, without any other additional cost, to the semi-diagonal quartic force constants. For protonated 1,2-dicyanoethene, quartic centrifugal distortion constants were computed at a different level of theory than the vibrational corrections to rotational constants, indeed considering the harmonic force field at the fc-CCSD(T)/jun-cc-pVTZ level.

On top of the best geometry, the electric dipole moment and the nuclear quadrupole-coupling constants (NQCCs), which are both first-order proprieties, were obtained. While the electric dipole moment is required to predict the type and intensity of rotational transitions, the NQCCs shape the rotational spectrum (hyperfine structure). NQC is the interaction between the quadrupole moment of a nucleus and the electric gradient at the same nucleus. This determines a splitting of the rotational energy levels, thus causing a splitting of the rotational transitions, i.e. the so-called hypefine structure. Noted is that a nucleus is defined as quadrupolar whenever its nuclear spin is $I > 1/2$. The hyperfine structure can be useful for the assignment and analysis of rotational lines in the laboratory and in astronomical observations (Turner, 1974; Puzzarini et al., 2010; Melosso et al., 2021); thus, spectral simulations strongly benefit from the introduction of such interaction terms. Both NQCCs and electric dipole moment components were obtained at the ae-CCSD(T)/aug-cc-pwCVQZ level for protonated diatomics, while for CNNCH$^+$ the aug-cc-pwCVTZ basis set was used. For protonated (*E*)-1,2-dicyanoethene, the fc-CCSD(T)/jun-cc-pVTZ level of theory was employed. Additionally, the vibrational corrections to the NQCCs and the electric dipole moment were also considered and obtained from the anharmonic treatment mentioned above.

The computational strategy outlined above was also employed to describe three molecules that are experimentally well-characterized in the literature. These are N_2H^+ and HCO$^+$, which offer a good reference for the new protonated diatomics, and NCCNH$^+$ for the direct comparison with CNNCH$^+$.

All DFT and MP2 computations were carried out using the Gaussian16 suite of programs (Frisch et al., 2016), while the CFOUR quantum-chemistry program (Stanton et al., 2016; Matthews et al., 2020) was used for all computations based on the CC theory. Concerning VPT2 calculations, its generalized version (GVPT2) (Schiff, 1955; Piccardo et al., 2015) was employed whenever using the Gaussian16 package.

3 RESULTS

The main molecule driving protonation in the ISM being H_3^+, the first mandatory step is to investigate whether the species considered in this work can be produced by reaction 1. The first screening is based on PA. If PA of P_2, Si_2, Cl_2, CNNC, and NCCHCHCN is greater than that of H_2, the protonation occurs and the reaction proceeds at the Langevin rate (Yamamoto, 2017), i.e. without a strong temperature dependence. The data required for this comparison are reported in Table 1, which also includes the data for N_2, CO, and NCCN, whose protonated forms have already been observed in the ISM.

Table 1. Computed proton affinities (kJ·mol^{-1}) for the molecules considered in this work. Electronic energies and ZPE are in Hartrees.

Process	Neutral Species		Protonated Species		
	Energy[a]	ZPE[b]	Energy[a]	ZPE[b]	PA
$H_2 \longrightarrow H_3^+$	-1.17568	0.00993	-1.34493	0.01985	418.3
$N_2 \longrightarrow NNH^+$	-109.52407	0.00538	-109.72269	0.01604	493.5
$CO \longrightarrow HCO^+$	-113.30829	0.00494	-113.54486	0.01609	591.5
$P_2 \longrightarrow PPH^+$	-682.43110	0.00179	-682.69261	0.00816	669.9
$Si_2 \longrightarrow Si_2H^+$	-578.57543	0.00123	-578.90135	0.00685	841.0
$Cl_2 \longrightarrow Cl_2H^+$	-920.12253	0.00127	-920.33914	0.00933	547.6
NCCN $\longrightarrow$ NCCNH$^+$	-185.63331	0.01567	-185.89266	0.02642	652.7
CNNC $\longrightarrow$ CNNCH$^+$	-185.51579	0.01422	-185.80000	0.02505	717.8
NCCHCHCN $\longrightarrow$ NCCHCHCNH$^+$	-263.04052	0.04875	-263.33125	0.05935	735.5

[a] Electronic energy computed at the junChS level. See text. [b] Anharmonic ZPE corrections at the ae-CCSD(T)/pwCVQZ level of theory for diatomic species. The same method in conjunction with the cc-pwCVTZ basis set was used for NCCN and CNNC. For (*E*)-NCCHCHCNH$^+$ and (*E*)-NCCHCHCN, the ZPEs were obtained at the revDSD/junTZ level of theory.

Table 1 indicates a PA value of 418 kJ·mol^{-1} for H_2, in good agreement with what reported by Yamamoto (2017) and Hunter and Lias (1998), i.e. 422.61 kJ·mol^{-1}. A similar agreement is noted for N_2 and CO, with values of 493.5 kJ·mol^{-1} and 591.1 kJ·mol^{-1}, respectively, to be compared with 493.8 kJ·mol^{-1} and 594 kJ·mol^{-1} given in Yamamoto (2017). The PA values reported in this work refer to 0 K and are more accurate than those available in the literature. The differences noted can be ascribed to the (small) thermal correction, the literature values referring to a temperature of 298.15 K.

From Table 1, it is noted that a PA value greater than that of H_2 has been obtained for Cl_2, thus suggesting the possible formation of Cl_2H^+ in the ISM. A further support to this is provided by the fact that PA(Cl_2) is intermediate between PA(N_2) and PA(CO), with both N_2H^+ and HCO^+ detected in the ISM. In particular, our PA(Cl_2) value of 547.6 kJ·mol^{-1} is in good agreement with that established experimentally by Cacace et al. (1998), i.e. 548.1 $\pm$ 13 kJ·mol^{-1}.

The PA value of P_2 is 669.9 kJ·mol^{-1} and it is very close to that of NCCN, i.e., 652.7 kJ·mol^{-1}, thus confirming the plausible formation of the protonated form of P_2. Finally, large PA values have been obtained for CNNC, (*E*)-NCCHCHCN, and Si_2. For the first two species, the PA values are 717.8 kJ·mol^{-1} and 735.5 kJ·mol^{-1}, respectively, which are comparable with those of molecules like H_2O, H_2S, and H_2CO (Yamamoto, 2017). The Si_2 species shows the largest PA, 841.0 kJ·mol^{-1}, this result being likely due to its triplet electronic ground state. Indeed, a similar value (869 kJ·mol^{-1}) is obtained for the CCS radical in

the same electronic state (Yamamoto, 2017). Notably, $HCCS^+$ as well as H_3O^+ and H_2COH^+ have already been observed in the ISM (Cabezas, C. et al., 2022; Wootten et al., 1991; Bacmann et al., 2016; Ohishi et al., 1996)

3.1 Formation routes

While the discussion above on PA values suggests that all the species under consideration could be formed in the ISM, the next step is the investigation of their formation routes. For this purpose, three test cases have been chosen: Cl_2H^+, P_2H^+, and $NCCNH^+$. According to the literature (Klippenstein et al., 2010; Zhu et al., 2019), the reaction between any molecule and $H_3{}^+$ is barrierless and proceeds through the formation of a stable intermediate characterized by a weak interaction between the protonated species and the H_2 molecule. This also applies to the reactive PESs considered in this work, which are shown in fig. 2. While a full account on their energetics is provided in the Supporting Material (SM), in the following, only the junChS+ZPE energies will be considered for our discussion.

Let us start by considering the formation of Cl_2H^+ from the reaction between $H_3{}^+$ and Cl_2 (panel **A** in fig. 2). The first intermediate formed, denoted as MIN1, is located at -143.2 kJ·mol^{-1} with respect to the energy of the reactants. This minimum can interconvert to its spatial equivalent structure via TS1 (-85.7 kJ· mol^{-1}) and forms Cl_2H^+ and H_2 without any exiting barrier. A similar reaction mechanism is also observed for P_2, panel **B** of fig. 2. However, two different products can be formed in this case. Starting from the reactants the open routes are: a barrierless approach to MIN1 (-237.7 kJ·mol^{-1}) and a barrierless approach to MIN2 (-192.4 kJ· mol^{-1}). These two intermediates are connected by the transition state TS2, located at -191.8 kJ· mol^{-1}. In addition, MIN1 can also form its spatial equivalent structure through TS1. MIN1 leads to product Prod1 without any exiting barrier, with Prod1 being the most stable stationary point of the P_2H^+ + H_2 reactive PES. Instead, MIN2 forms Prod2, i.e. the linear isomer of P_2H^+, which might be potentially present in the ISM, even though it lies $\sim$50 kJ· mol^{-1} higher in energy than the cyclic form.

The last panel of fig. 2 (panel **C**) shows the reaction mechanism for the protonation of CNNC. In this case, there are three open barrierless approaches that lead to the formation of MIN1, MIN2 and MIN3. The first two intermediates are very close in energy: MIN1 lies at -305.5 kJ· mol^{-1} and MIN2 at -301.8 kJ· mol^{-1}. They can interconvert through TS3, which differs from MIN2 for the orientation of the H_2 moiety with respect to the linear backbone. The motion of H_2 is a large amplitude one, TS3 being indeed characterized by an imaginary frequency of only 14 cm^{-1}. Consequently, the two structures are very close in energy, i.e. they differ by 0.1 kJ· mol^{-1} at both the revDSD/junTZ and junChS levels of theory augmented by corresponding ZPEs. From fig. 2, it is noted that both MIN1 and MIN2 can form – without going through any barrier – Prod1, i.e. linear $CNNCH^+$ and H_2. TS2, lying at -46.4 kJ· mol^{-1}, represents the barrier to be overcome for distorting the linear chain and forming MIN3, which can then interconvert with its equivalent structure via TS1 or form Prod2, i.e. a bent form of $CNNCH^+$ + H_2.

The three protonation reactions of fig. 2 demonstrate that Cl_2H^+, P_2H^+ and $CNNCH^+$ can be formed in the ISM. The energetic characterization there reported is expected to have an accuracy of 1-5 kJ· mol^{-1} (Alessandrini et al., 2019, 2021b) and can thus be used to guide the kinetic analysis of these reactions. While there are no doubts about the isomeric form produced in the reaction between Cl_2 and $H_3{}^+$, for the other two reactions, two possible products are present. However, the stability of MIN1 and the small energy difference between MIN1 and Prod1 should favor the formation of cyclic P_2H^+ and linear $CNNCH^+$. Therefore, only these forms have been considered in the following spectroscopic characterization.

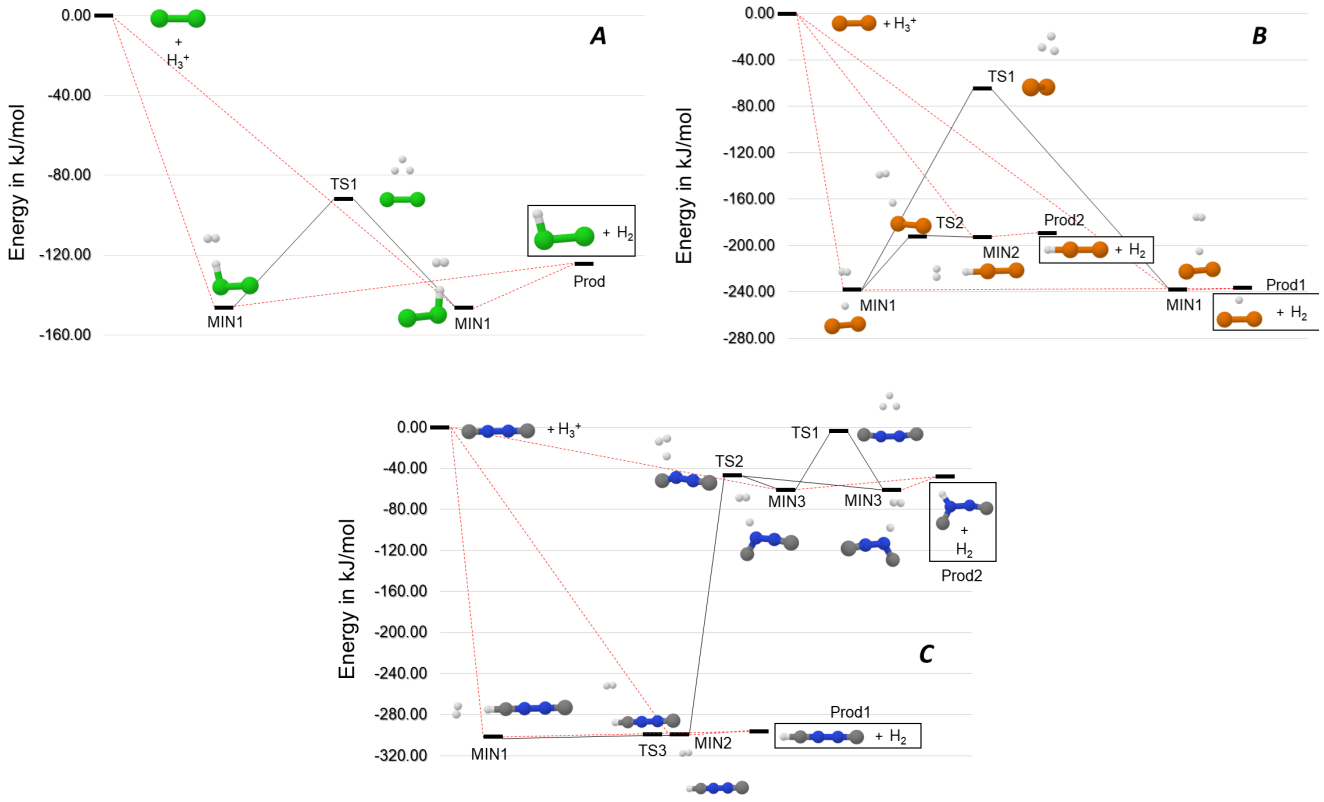

Figure 2. Reactive PES for: the Cl_2 + H_3^+ reaction (panel **A**), the P_2 + H_3^+ reaction (panel **B**), and the CNNC + H_3^+ reaction (panel **C**). The full list of the energies is reported in the SM.

3.2 Spectroscopic Characterization

To enable the experimental detection of the protonated species considered in this work, a computational spectroscopic characterization involving fundamental vibrational frequencies and rotational spectroscopy parameters has been carried out. In the following, the results are presented according to the class of species they belong: (*i*) protonated diatomics, where the experimental reference is offered by N_2H^+ and HCO^+; (*ii*) the linear $CNNCH^+$ cation compared with the $NCCNH^+$ counterpart; (*iii*) the protonated form of (*E*)-1,2-dicyanoethene. These species, with the only exception of the linear $CNNCH^+$ molecule, are planar asymmetric rotors. Their planarity leads to a null dipole moment component along the c axis ($\mu_c = 0$, with ab thus being the molecular plane). For P_2H^+ and Si_2H^+, the μ_a component is equal to zero as well because of symmetry reasons. For the linear rotor $CNNCH^+$, the only non-null dipole moment component is that along the molecular axis, and it is simply denoted as μ.

3.2.1 Protonated Diatomics

The rotational parameters of Cl_2H^+, P_2H^+, and Si_2H are reported in Table 2, while those of N_2H^+ and HCO^+ are provided in the SM. It should be pointed out that protonated CO and N_2 are linear rotor, and thus they are characterized by a single rotational constant and one quartic centrifugal distortion term. Differently, Cl_2H^+, P_2H^+ and Si_2H^+ are asymmetric rotors, very close to the prolate limit ($\kappa = 0.999$). Therefore, these species have three rotational constants and five quartic centrifugal distortion parameters.

The "CBS+CV+fT+fQ" level of theory augmented by vibrational corrections at the ae-CCSD(T)/cc-pwCVQZ level of theory, provides the rotational constant for N_2H^+ with an error as low as 0.01%, as shown

in Table S1. For HCO^+, the error is slightly larger, this being 0.07%. However, both values are within the uncertainty predicted by benchmark studies for the "CBS+CV+fT+fQ" approach (Puzzarini et al., 2008). Even though N_2H^+ and HCO^+ are linear molecules, the accuracy reached for such species should hold also for the rotational constants of the bent Cl_2H^+ cation and the cyclic P_2H^+ and Si_2H^+ species, at least when the mean error of the three rotational constants is considered.

Table 2. Rotational parameters (in MHz) of Watson's S reduced Hamiltonian (I^r representation) for Cl_2H^+, P_2H^+, and Si_2H^+. Dipole moment components (in debyes) are also reported.

Parameter [a]		Cl_2H^+	P_2H^+	Si_2H^+
A_e		312463.22	316752.63	326007.83
B_e		7178.27	8682.64	6894.94
C_e		7017.06	8450.99	6752.14
A_0		310826.35	308542.91	318623.35
B_0		7148.88	8648.59	6852.23
C_0		6980.01	8401.96	6699.21
$D_J \times 10^3$		5.51	4.99	7.05
D_{JK}		0.213	0.072	-0.224
D_K		22.2	45.1	66.27
$d_1 \times 10^4$		-1.19	-1.51	-2.01
$d_2 \times 10^6$		-6.50	-8.99	-5.15
1.5 χ_{aa}	(Cl1)	-205.5		
$(\chi_{bb} - \chi_{cc})$/4	(Cl1)	-2.0		
χ_{ab}	(Cl1)	1.14		
1.5 χ_{aa}	(Cl2)	-95.9		
$(\chi_{bb} - \chi_{cc})$/4	(Cl2)	-25.6		
χ_{ab}	(Cl2)	-17.3		
$\lvert\mu_a\rvert$ / D		2.28	0.0	0.0
$\lvert\mu_b\rvert$ / D		1.57	0.13	0.06

[a] Equilibrium rotational constants from the "CBS+CV+fT+fQ" approach. Dipole moments and NQCCs computed (CBS+CV+fT+fQ reference geometry) computed at the ae-CCSD(T)/aug-cc-pwCVQZ level. Vibrational corrections to the previous quantities and quartic centrifugal distortion constants at the ae-CCSD(T)/cc-pwCVQZ level of theory.

To accurately simulate rotational spectra, other parameters are required. These are the NQCCs of the N and Cl atoms together with quartic centrifugal distortion constants for all species. As mentioned in Computational details section, the latter parameters have been computed at the ae-CCSD(T)/cc-pwCVQZ level of theory, while NQCCs have been obtained, on top of the "CBS+CV+fT+fQ" geometry, by computing the electric-field gradient at the quadrupolar nuclei at the ae-CCSD(T)/aug-cc-pwCVQZ level. The NQCCs have also been corrected for the vibrational contribution evaluated at the ae-CCSD(T)/cc-pwCVQZ level.

For N_2H^+, the experimental D_J quartic centrifugal distortion constant is reproduced with a discrepancy of about 3%, while for HCO^+ the error is 2%. Similar uncertainties are expected for the quartic centrifugal distortion constants of the protonated species of Table 2. Moving to NQCCs, for N_2H^+, the discrepancy with respect to experiment is on average 1% and, thus, a similar accuracy is expected for the constants of Cl_2H^+, as found in other works in the literature (Cazzoli et al., 2006; Puzzarini et al., 2010).

Finally, to simulate the rotational spectrum, the electric dipole moment components were computed at the same level of theory as NQCCs. The obtained values indicate that P_2H^+ and Si_2H^+ have a (very) weak *b*-type rotational spectrum, the only non-null dipole moment component (μ_b) being equal to 0.01 D. In the case of P_2H^+, the intensity of the rotational lines is also affected by spin statistics. Indeed, ^{31}P has a nuclear spin I of $1/2$ (fermion) and the total wave function has to be anti-symmetric with respect to the exchange of the two P nuclei. The electronic and vibrational ground states being total symmetric, the overall symmetry of the total wave function is determined by the rotational and nuclear terms. Three *ortho* (symmetric) and one *para* (anti-symmetric) states arise for the nuclear wave function. Instead, the symmetry of the rotational contribution is determined by the $K_a + K_c$ sum[1] because the nuclei exchange occurs by rotation around the b axis. As a consequence, the rotational transitions associated with odd $K_a + K_c$ values are three times more intense than those with even $K_a + K_c$ values.

The Cl_2H^+ species has both a-type and b-type spectra, the corresponding dipole moment components being 2.28 D and 1.57 D, respectively. Thus, a-type transitions should have an intensity comparable to that of N_2H^+ and HCO^+, while b-type transitions are expected to be weaker. An example is shown in fig. 3, where the blue spectrum reports a μ_a transition of Cl_2H^+ at around 98307 MHz and a weaker μ_b transition predicted at 98352 MHz. These are compared with the $J = 1 \leftarrow 0$ transition of N_2H^+, for which both the theoretical simulation (this work, in red) and the experimental data (Cazzoli et al. (2012), in black) are considered. The experimental transition occurs at 93173 MHz and the theoretical simulation lies close in frequency, indeed being centered at 93181 MHz. In relative terms, this discrepancy should also apply to the simulation of Cl_2H^+. Conservatively, based on benchmark studies (Alessandrini et al., 2018), an uncertainty of 0.1% can be expected for each rotational constant of Cl_2H^+ . This would locate the $J_{K_a,K_c} = 7_{1,7} \leftarrow 6_{1,6}$ transition of fig 3 in the range 98307.6 ± 18 MHz (3σ deviation). Lastly, the two insets of fig. 3 show the hyperfine splittings based on the computed NQCCs. For N_2H^+, the structure reproduces very well the experimental one in terms of both intensity and shift with respect to the unperturbed frequency. The same should apply to Cl_2H^+.

For Cl_2H^+, P_2H^+, and Si_2H^+, vibrational frequencies have been obtained at the ae-CCSD(T)/cc-pwCVQZ level. The results are reported in Table 3, while those of N_2H^+ and HCO^+ are provided in the SM, where they are also compared with the experimental counterparts. To understand the accuracy of the predictions for the protonated species addressed in this work, we start by analyzing the results for N_2H^+ and HCO^+. For the former cation, the lowest fundamental frequency is due to the bending mode. This is predicted at 700.94 cm^{-1} within the harmonic approximation and about 9 cm^{-1} lower in energy at the anharmonic level. The experiment locates this band at 698.6353(14) cm^{-1} (Foster and McKellar, 1984). The other two bands of N_2H^+ were observed at 3233.96085 cm^{-1} (Nakanaga et al., 1990) and 2257.8667(13) cm^{-1} (Owrutsky et al., 1986). Both of them are well predicted by the VPT2 treatment, with an error of only 8 cm^{-1}. For HCO^+, the fundamental bands are predicted (incorporating anharmonicity) to lie at $\nu_1 = 3090.31$ cm^{-1}, $\nu_2 = 2187.67$ cm^{-1}, and $\nu_3 = 830.05$ cm^{-1}. These are in good agreement with the experimental data, with discrepancies of about $2 - 4$ cm^{-1}. In fact, the experimental observed frequencies are: $\nu_1 = 3088.7951(31)$ cm^{-1} (Amano, 1983), $\nu_2 = 2183.9496(6)$ cm^{-1} (Foster and McKellar, 1984), and $\nu_3 = 828.2305(9)$ cm^{-1} (Kawaguchi et al., 1985).

A similar accuracy is expected for the other protonated diatomics. The lowest fundamental band is observed for the stretching involving third-row atoms (ν_2 band for P_2H^+ and Si_2H^+, ν_3 band for Cl_2H^+).

[1] The asymmetric rotor energy levels are labeled using (in addition to the J quantum number) two "pseudo" quantum numbers: K_a and K_c, which correspond to the K values of the two limiting prolate (referred to the a symmetry axis) and oblate (referred to the c symmetry axis) cases of the symmetric rotor, respectively.

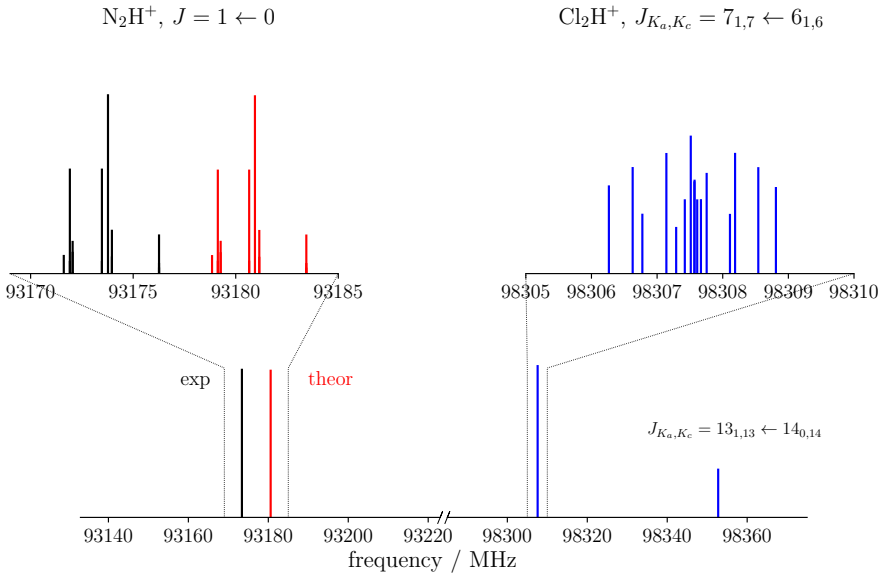

Figure 3. Theoretical (red) and experimental (black) $J = 1 \leftarrow 0$ transition of N_2H^+ compared with the simulated spectrum of Cl_2H^+ (blue) in the nearby frequency region. For the latter, both a-type ($J_{K_a,K_c} = 7_{1,7} \leftarrow 6_{1,6}$) and b-type ($J_{K_a,K_c} = 13_{1,13} \leftarrow 14_{0,14}$) transitions are visible. The two insets show the hyperfine structures due to nitrogen and chlorine nuclei for N_2H^+ and Cl_2H^+, respectively.

This vibrational mode lies at 518.0 cm^{-1} for Cl_2H^+, 728.05 cm^{-1} for P_2H^+, and 459.89 cm^{-1} for Si_2H^+. The ν_3 band for P_2H^+ and Si_2H^+, and the ν_2 band for Cl_2H^+ correspond to the asymmetric stretching involving the H atom, which is predicted at 949.02 cm^{-1}, 971.54 cm^{-1}, and 858.1 cm^{-1} for Si_2H^+, P_2H^+, and Cl_2H^+, respectively. The fundamental frequencies for symmetric stretching (ν_1) are, in the same order as before, 1435.82 cm^{-1}, 1659.00 cm^{-1}, and 2619.23 cm^{-1}. For the Cl_2H^+ species, the most intense band is ν_1, while for P_2H^+ and Si_2H^+ it is ν_3. In conclusion, in view of their accuracy – which is conservatively expected to be better than 10 cm^{-1} – the data obtained in this work can allow the investigation of the vibrational spectrum of these species as well as the ro-vibrational ones [2].

3.2.2 CNNCH$^+$ and NCCNH$^+$

The rotational parameters of CNNCH$^+$ are compared with those of NCCNH$^+$ in Table 4. The error associated with our computational strategy can be derived by considering the experimental values of NCCNH$^+$ (Gottlieb et al., 2000). For B_0 of NCCNH$^+$, the discrepancy between experiment and theory is about 4 MHz, which means 0.09% in relative terms; thus, the value of the rotational constant of CNNCH$^+$ lies within the interval 5225.59 MHz $\pm$4.7 MHz. For D_J of NCCNH$^+$, the deviation of the computed value (0.480 $\times 10^{-3}$ MHz) from experiment is 9.3% and the same confidence range is expected to apply to CNNCH$^+$. Moving to NQCCs, in both molecules, we have two nitrogen atoms which are denoted as the outer N and the inner N; while such denotation is rather straightforward for NCCNH$^+$, in the case of CNNCH$^+$, the outer N is the furthest from H, i.e. N2 in fig. 1. In NCCNH$^+$, the NQCC of the outer N is well reproduced, the error being 0.2%, while a discrepancy of about 12% is noted for the inner NQCC, the experimental value being 0.250 MHz and the computed one 0.221 MHz. However, it has to be noted that the absolute discrepancy is nearly the same: $\sim$50 kHz for χ(outer) and $\sim$30 kHz for χ(inner). Therefore, the hyperfine splittings are well reproduced as shown in the left inset in fig. 4, where the experimental

[2] For those interested in the prediction of the rotational constants of vibrational excited states, the required α_r^γ values are available in Table S2

Table 3. Harmonic and anharmonic vibrational frequencies (cm^{-1}) for Cl_2H^+, P_2H^+, and Si_2H^+. Anharmonic intensities in $km \cdot mol^{-1}$. All quantities are at the ae-CCSD(T)/cc-pwCVQZ level of theory.

		Cl_2H^+		
Mode[a]		Harm. Freq.	Frequency	Intensity
ν_1	A'	2740.60	2619.23	234.4
ν_2	A'	880.69	858.10	19.0
ν_3	A'	525.24	518.05	13.4
		P_2H^+		
		Harm. Freq.	Frequency	Intensity
ν_1	A_1	1793.54	1659.00	433.8
ν_2	A_1	737.32	728.05	0.71
ν_3	B_2	1095.77	971.54	1376.9
		Si_2H^+		
		Harm. Freq.	Frequency	Intensity
ν_1	A_1	1518.28	1435.82	1120.2
ν_2	A_1	468.18	459.89	10.2
ν_3	B_2	1063.83	949.02	1675.0

[a] The symmetry of the normal mode is given in parentheses.

and theoretical hyperfine structures of the $J = 3 \leftarrow 2$ transition of NCCNH$^+$ are depicted. Focusing on the unperturbed transition, the experimental frequency of NCCNH$^+$ is 24 MHz lower than the simulated one. This deviation can be considered as the uncertainty affecting the $J = 3 \leftarrow 2$ transition of CNNCH$^+$, which is shown in fig. 4 together with its hyperfine structure. Concerning the latter, an analogous accuracy to that observed for NCCNH$^+$ is expected. Indeed, in absolute terms, the uncertainty affecting NQCCs should be of the order of 50 kHz. Since the dipole moment of CNNCH$^+$ is large (about 6 D) and only 0.4 D smaller than that of NCCNH$^+$, the detection of this new protonated form is reasonable and future works will benefit from the present data.

Table 4 also reports the fundamental frequencies and intensities of the vibrational modes for both NCCNH$^+$ and CNNCH$^+$. In this case, no experimental reference is available in the literature, but vibrational frequencies computed at the ae-CCSD(T)/cc-pwCVTZ level are sufficiently accurate to guide the assignment of the (ro-)vibrational spectrum for both species.

3.2.3 NCCHCHCNH$^+$

The last molecule considered in this work is the *E* isomer of NCCHCHCNH$^+$. According to the computed electric dipole moment components, *a*-type transitions are the most prominent features of the rotational spectrum because the associated dipole component is as large as 10.1 D. Based on the literature on this topic (Puzzarini et al., 2008), in relative terms, the uncertainty affecting the equilibrium "CBS+CV" rotational constants is expected to be 0.6%, which drastically reduces to about 0.1% when vibrational corrections are incorporated. Actually, this uncertainty can be as small as 0.05%, as demonstrated, for example, by Melosso et al. (2022); Alessandrini et al. (2021a). For an accurate simulation of the rotational spectrum of

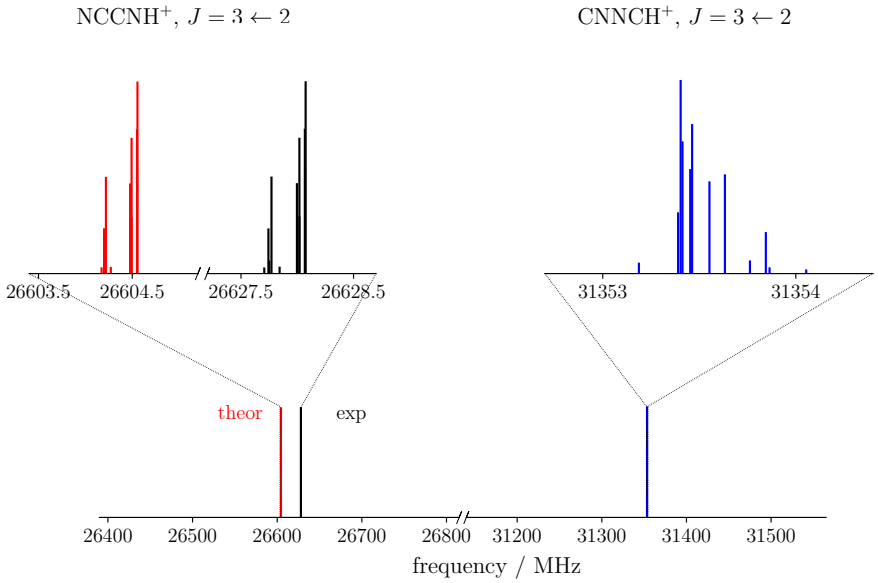

Figure 4. Theoretical (red) and experimental (black) $J = 3 \leftarrow 2$ transition of NCCNH$^+$ compared with the simulation of the same transition of CNNCH$^+$ (blue). The insets show in detail the corresponding hyperfine structures.

Table 4. Rotational parameters and fundamental vibrational frequencies of CNNCH$^+$ and NCCNH$^+$.

Parameter	CNNCH$^+$		NCCNH$^+$	
	Theory[a]		Theory[a]	Exp.[b]
B_e / MHz	5213.55		4431.93	-
B_0 / MHz	5225.59		4434.09	4438.01154(27)
$D_J \times 10^3$ / MHz	0.624		0.481	0.53063(21)
χ(outer) / MHz	4.038		-5.429	-5.480(3)
χ(inner) / MHz	1.698		-0.221	-0.250(8)
$\lvert\mu\rvert$ / D	5.86		6.26	-
Mode[c]	Freq. (cm^{-1})	I (km · mol^{-1})	Freq. (cm^{-1})	I (km · mol^{-1})
ν_1 (Σ)	3205.05	227.8	3451.15	862.7
ν_2 (Σ)	2286.31	95.8	2352.24	296.8
ν_3 (Σ)	1983.44	802.0	2162.15	198.8
ν_4 (Σ)	972.47	25.1	849.10	52.1
ν_5 (Π)	678.36	25.6	597.02	117.0
ν_6 (Π)	393.15	8.0	502.47	34.9
ν_7 (Π)	142.18	65.0	202.58	26.2

[a] Equilibrium rotational constants from the "CBS+CV+fT+fQ" approach. Dipole moment and NQCCs computed, on top of the CBS+CV+fT+fQ reference geometry, at the ae-CCSD(T)/aug-cc-pwCVTZ level. Vibrational corrections to the previous quantities and quartic centrifugal distortion constants at the ae-CCSD(T)/cc-pwCVTZ level of theory. Anharmonic frequencies and intensities at the ae-CCSD(T)/cc-pwCVTZ level of theory. [b] Experimental data taken from Gottlieb et al. (2000). [c] The symmetry of the vibrational modes is given in parentheses.

Table 5. Computed spectroscopic parameters of (*E*)-NCCHCHCNH$^+$.

Rotational Spectroscopy[a]		Vibrational Spectroscopy[b]		
Parameter (MHz)		Mode	Freq. (cm^{-1})	I (km · mol $^{-1}$)
A_e	43570.12	$\nu_1\,(A')$	3518.01	1027.2
B_e	1454.23	$\nu_2\,(A')$	3077.27	45.8
C_e	1407.26	$\nu_3\,(A')$	3050.92	2.3
A_0	43419.06	$\nu_4\,(A')$	2276.41	75.8
B_0	1451.13	$\nu_5\,(A')$	2214.97	199.9
C_0	1403.78	$\nu_6\,(A')$	1601.12	156.2
$D_J \times 10^3$	0.0497	$\nu_7\,(A')$	1300.07	10.2
D_K	0.118	$\nu_8\,(A')$	1257.23	5.9
D_{JK}	14.75	$\nu_9\,(A')$	1011.17	1.1
$d_1 \times 10^3$	0.0125	$\nu_{10}\,(A')$	1004.12	29.3
$d_2 \times 10^6$	-3.98	$\nu_{11}(A')$	608.03	118.3
1.5 χ_{aa}(N1)	0.162	$\nu_{12}\,(A')$	548.07	10.0
(χ_{bb}-χ_{cc})/4 (N1)	-0.027	$\nu_{13}\,(A')$	514.13	3.6
χ_{ab} (N1)	-0.110	$\nu_{14}\,(A')$	240.38	5.3
1.5 χ_{aa}(N2)	-6.876	$\nu_{15}\,(A')$	126.30	1.1
(χ_{bb}-χ_{cc})/4 (N2)	1.146	$\nu_{16}\,(A'')$	965.06	36.8
χ_{ab} (N2)	-1.542	$\nu_{17}\,(A'')$	828.78	16.8
$\lvert\mu_a\rvert$ / D	9.9	$\nu_{18}\,(A'')$	555.71	1.0
$\lvert\mu_b\rvert$ / D	0.82	$\nu_{19}\,(A'')$	471.74	129.3
		$\nu_{20}\,(A'')$	353.25	3.9
		$\nu_{21}\,(A'')$	123.49	0.6

[a] Values in MHz if not otherwise stated. Equilibrium rotational constants at the "CBS+CV" level of theory augmented by vibrational corrections at the revDSD/junTZ level. Quartic centrifugal distortion constants at the fc-CCSD(T)/junTZ level. Dipole moments (in debyes) and NQCCs computed at the same level of theory but on top of the "CBS+CV" geometry and vibrationally corrected at the revDSD/junTZ level. For further details see text. Watson's S reduction in I^r representation is used. [b] The symmetry of the normal mode is given in parentheses. Vibrational frequencies and intensity computed at the revDSD/junTZ level of theory using GVPT2.

this species, the NQCCs of both nitrogen atoms have been computed, on top of the "CBS+CV" geometry, at the fc-CCSD(T)/jun-cc-pVTZ level. Quartic centrifugal distortion constants have been obtained from a harmonic force field at the same level of theory. All these data, collected in Table 5, can guide future experimental measurements as well as astronomical search of this molecule in view of its large dipole moment (in particular that associated with a-type transitions) and the small uncertainties affecting our spectroscopic parameters. Notably, the observation of this species in the ISM would indirectly confirm the feasibility of the reaction between the CN radical and vinyl cyanide, as suggested in Marchione et al. (2022).

In Table 5, the vibrational fundamental frequencies of (*E*)-NCCHCHCNH$^+$ are reported. It is noted that the strongest band is the one associated with the stretching of the terminal H atom. The weakest band has an intensity of 0.6 km·mol^{-1} and it is the out-of-plane bending of the CN and CNH terminal groups.

According to the literature on this topic (Barone et al., 2015), double-hybrid functionals should reproduce the experimental fundamental frequencies with a mean error of about 10.4 cm^{-1}.

4 CONCLUSIONS

In the ISM conditions, protonation can occur easily thanks to the abundance of H_3^+. For this reason, protonated forms of "invisible" (to rotational spectroscopy and thus to radioastronomy) apolar species are considered tracers of the latter and can be used to derive an estimate of their abundance. In this work, the presence of five new apolar species in the ISM has been suggested based on the investigation of the corresponding protonated forms: P_2H^+, Cl_2H^+, Si_2H^+, CNNCH$^+$, and (E)-NCCHCHCNH$^+$. These can be considered good proxies of P_2, Cl_2, Si_2, CNNC, and (E)-NCCHCHCN, respectively. The most stable forms of these species have been preliminary investigated at the revDSD/junTZ level of theory. This pointed out that Cl_2H^+ has a bent structure, while P_2H^+ (in contrast with N_2H^+ which is linear) and Si_2H^+ have a cyclic structure. CNNCH$^+$ is linear in analogy to the NCCNH$^+$ isomer, which has been already characterized and detected in the ISM. Finally, a planar structure has been found for (E)-NCCHCHCNH$^+$.

State-of-the-art computational methodologies have been used to assess the PA values of Cl_2, P_2, Si_2, CNNC, and (E)-NCCHCHCN to be compared with that of H_2. Indeed, a PA greater than that of H_2 implies that protonation due to H_3^+ is favored. This is the case for all the species considered in this study, and the computed PA values have been found in good agreement with experimental and previous theoretical estimates available in the literature. To give an example, the PA value of Cl_2H^+ is in good agreement with the experimental counterpart reported by Cacace et al. (1998), the difference being only 0.5 kJ·mol^{-1}.

To provide more details on the protonation mechanisms, the formation routes leading to Cl_2H^+, P_2H^+, and CNNCH$^+$ have been explored. The junChS+ZPE energies, evaluated on top of revDSD/junTZ geometries, pointed out the presence of simple reaction paths starting from the corresponding apolar species and H_3^+. All these mechanisms are dominated by a barrierless step from reactants to adducts, but also from these latter to products. In all the three cases, the minimum energy path leads to the most stable protonated forms, namely bent Cl_2H^+, cyclic P_2H^+, and linear CNNCH$^+$. The first two formation routes are similar to those producing N_2H^+ and HCO$^+$; therefore, it can be safely assumed that a similar mechanism leads to Si_2H^+. We also expect that the formation of (E)-NCCHCHCNH$^+$ presents similarity with that of CNNCH$^+$.

Since the investigated protonated species are potentially present in the ISM, an accurate computational spectroscopic characterization has also been carried out. For protonated diatomics, the rotational constants are expected to reproduce the experimental counterparts with an error of 0.1%, or even smaller as in the case of HCO$^+$ and N_2H^+. For CNNCH$^+$, and (E)-NCCHCHCNH$^+$, because of the reduction in the level of theory employed, a larger uncertainty has been obtained, but still on the order of 0.1%. For all the species investigated, a full set of rotational parameters, also including quartic centrifugal distortion constants and NQCCs, has been obtained and used to estimate low-J rotational transitions. To give an example, the $J_{K_a,K_c} = 7_{1,7} \leftarrow 6_{1,6}$ rotational transition of Cl_2H^+ is expected to lie in the range 98307.6 ± 18 MHz. For CNNCH$^+$, the $J = 3 \leftarrow 2$ rotational transition is predicted to be centered at around 30353.5 MHz with an uncertinity of about 24 MHz. The spectroscopic characterization of (E)-NCCHCHCNH$^+$ is of particular interest because it results from protonation of the *E* isomer of 1,2-dicyanoethene, which in turn is the main product of the reaction between the CN radical and acrylonitrile (Marchione et al., 2022). This reaction is of great astrochemical relevance, but the only possible option for detecting (E)-NCCHCHCN is via its protonated proxy. This work reports – for the first time – the rotational constants of this species and points out that protonation leads to a large dipole moment of about 10 D, which further suggests the

detectability of (E)-NCCHCHCNH$^+$. Therefore, our simulations will be useful to support and complement experimental measurements as well as to guide tentative detections or assignments.

For all the aforementioned molecules, the fundamental vibrational frequencies have accurately obtained by exploiting a methodology that incorporates the effects of anharmonicity. Therefore, the expected accuracy of our predictions is 10 cm^{-1} or even better, which is largely sufficient for the assignment of experimental infrared bands.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

SA worked on conceptualization of the work as well as investigation. SA, LB and MM carried out the formal analysis and wrote a first draft of the manuscript. CP reviewed the manuscript and provided the funding. Final editing was carried out by CP, LB, MM and SA.

FUNDING

This work has been supported by MIUR (PRIN Grant Number 202082CE3T) and by the University of Bologna (RFO funds). The COST Action CA21101 "COSY - Confined molecular systems: from a new generation of materials to the stars" is also acknowledged. This work was supported by the Open Access Publishing Fund of the Scuola Normale Superiore.

ACKNOWLEDGEMENTS

The SMART@SNS Laboratory (`http://smart.sns.it`) is acknowledged for providing high-performance computing facilities.

SUPPLEMENTAL DATA

Supplementary Material collects the rotational parameters and vibrational frequencies of N_2H^+ and HCO^+ and their comparison with experiment. For these, also vibration-rotation interaction constants are provided together with those of Cl_2H^+, P_2H^+, and Si_2H^+. The optimized structural parameters at different level of theory are reported for all the molecules of this work. In the last section, the energy values of the stationary points involved in the formation routes of fig. 2 are given.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

REFERENCES

Agúndez, M., Cernicharo, J., De Vicente, P., Marcelino, N., Roueff, E., Fuente, A., et al. (2015). Probing non-polar interstellar molecules through their protonated form: Detection of protonated cyanogen (NCCNH$^+$). *Astron. Astrophys.* 579, L10

Agúndez, M., Marcelino, N., and Cernicharo, J. (2018). Discovery of interstellar isocyanogen (CNCN): Further evidence that dicyanopolyynes are abundant in space. *Astrophys. J. Lett.* 861, L22

Alessandrini, S., Barone, V., and Puzzarini, C. (2019). Extension of the "cheap" composite approach to noncovalent interactions: The jun-chs scheme. *J. Chem. Theory and Comput.* 16, 988–1006

Alessandrini, S., Gauss, J., and Puzzarini, C. (2018). Accuracy of rotational parameters predicted by high-level quantum-chemical calculations: Case study of sulfur-containing molecules of astrochemical interest. *J. Chem. Theory and Comput.* 14, 5360–5371

Alessandrini, S., Melosso, M., Jiang, N., Bizzocchi, L., Dore, L., and Puzzarini, C. (2021a). Conformational stability of cyclopropanecarboxaldehyde is ruled by vibrational effects. *Molecular Physics* 119, e1955988

Alessandrini, S., Tonolo, F., and Puzzarini, C. (2021b). In search of phosphorus in astronomical environments: The reaction between the cp radical ($X^2\Sigma^+$) and methanimine. *J. Chem. Phys.* 154, 054306

Amano, T. (1983). The ν_1 fundamental band of HCO$^+$ by difference frequency laser spectroscopy. *J. Chem. Phys.* 79, 3595–3595

Bacmann, A., García-García, E., and Faure, A. (2016). Detection of protonated formaldehyde in the prestellar core L1689B. *Astron. Astrophys.* 588, L8

Barone, V., Alessandrini, S., Biczysko, M., Cheeseman, J. R., Clary, D. C., McCoy, A. B., et al. (2021). Computational molecular spectroscopy. *Nat. Rev. Methods Primers* 1, 1–27

Barone, V., Biczysko, M., Bloino, J., Cimino, P., Penocchio, E., and Puzzarini, C. (2015). Cc/dft route toward accurate structures and spectroscopic features for observed and elusive conformers of flexible molecules: pyruvic acid as a case study. *J. Chem. Theory Comput.* 11, 4342–4363

Cabezas, C., Agúndez, M., Marcelino, N., Tercero, B., Endo, Y., Fuentetaja, R., et al. (2022). Discovery of the elusive thioketenylium, HCCS$^+$, in TMC-1. *Astron. Astrophys.* 657, L4

Cacace, F., de Petris, G., Pepi, F., Rosi, M., and Sgamellotti, A. (1998). Elemental chlorine and chlorine fluoride: Theoretical and experimental proton affinity and the gas phase chemistry of Cl_2H^+ and FClH$^+$ ions. *J. Phys. Chem. A* 102, 10560–10567

Cazzoli, G., Cludi, L., Buffa, G., and Puzzarini, C. (2012). Precise thz measurements of HCO$^+$, N_2H^+, and CF$^+$ for astrophysical observations. *Astrophys. J. Suppl. Ser.* 203, 11

Cazzoli, G., Puzzarini, C., Gambi, A., and Gauss, J. (2006). Rotational spectra of 1-chloro-2-fluoroethylene. i. main isotopologues and deuterated species of the trans isomer. *J. Chem. Phys.* 125, 054313

Cernicharo, J., Agúndez, M., Kaiser, R. I., Cabezas, C., Tercero, B., Marcelino, N., et al. (2021). Discovery of two isomers of ethynyl cyclopentadiene in TMC-1: Abundances of CCH and CN derivatives of hydrocarbon cycles. *Astron. & Astrophys.* 655, L1

Cooke, I. R., Gupta, D., Messinger, J. P., and Sims, I. R. (2020). Benzonitrile as a proxy for benzene in the cold ism: Low-temperature rate coefficients for CN + C_6H_6. *Astrophys. J. Lett.* 891, L41

Dunning, T. H., Peterson, K. A., and Wilson, A. K. (2001). Gaussian basis sets for use in correlated molecular calculations. x. the atoms aluminum through argon revisited. *J. Chem. Phys.* 114, 9244–9253. doi:10.1063/1.1367373

Dunning Jr., T. H. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* 90, 1007–1023

Feller, D. (1993). The use of systematic sequences of wave functions for estimating the complete basis set, full configuration interaction limit in water. *J. Phys. Chem.* 98, 7059–7071

Foster, S. and McKellar, A. (1984). The ν_3 fundamental bands of HN_2^+, DN_2^+, and DCO+. *J. Chem. Phys.* 81, 3424–3428

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). Gaussian16 Revision C.01 Gaussian Inc. Wallingford CT

Fukui, K. (1981). The path of chemical reactions - the IRC approach. *Acc. Chem. Res.* 14, 363–368

Gottlieb, C., Apponi, A., McCarthy, M., Thaddeus, P., and Linnartz, H. (2000). The rotational spectra of the $HCCCNH^+$, $NCCNH^+$, and CH_3CNH^+ ions. *J. Chem. Phys.* 113, 1910–1915

Grimme, S., Antony, J., Ehrlich, S., and Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* 132, 154104

Grimme, S., Ehrlich, S., and Goerigk, L. (2011). Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* 32, 1456–1465

Heckert, M., Kállay, M., and Gauss, J. (2005). Molecular equilibrium geometries based on coupled-cluster calculations including quadruple excitations. *Mol. Phys.* 103, 2109–2115

Heckert, M., Kállay, M., Tew, D. P., Klopper, W., and Gauss, J. (2006). Basis-set extrapolation techniques for the accurate calculation of molecular equilibrium geometries using coupled-cluster theory. *J. Chem. Phys.* 125, 044108

Helgaker, T., Klopper, W., Koch, H., and Noga, J. (1997). Basis-set convergence of correlated calculations on Water. *J. Chem. Phys.* 106, 9639

Hunter, E. P. and Lias, S. G. (1998). Evaluated gas phase basicities and proton affinities of molecules: an update. *J. Phys. Chem. Ref. Data* 27, 413–656

Kállay, M. and Gauss, J. (2008). Approximate treatment of higher excitations in coupled-cluster theory. II. Extension to general single-determinant reference functions and improved approaches for the canonical Hartree–Fock case. *J. Chem. Phys.* 129, 144101

Kawaguchi, K., Yamada, C., Saito, S., and Hirota, E. (1985). Magnetic field modulated infrared laser spectroscopy of molecular ions: The ν_2 band of HCO^+. *J. Chem. Phys.* 82, 1750–1755

Klippenstein, S. J., Georgievskii, Y., and McCall, B. J. (2010). Temperature dependence of two key interstellar reactions of H_3^+: $O(^3P)+ H_3^+$ and $CO + H_3^+$. *J. Phys. Chem. A* 114, 278–290

Kucharski, S. A. and Bartlett, R. J. (1991). Recursive intermediate factorization and complete computational linearization of the coupled-cluster single, double, triple, and quadruple excitation equations. *Theor. Chim. Acta* 80, 387–405

Lee, K. L. K., McGuire, B. A., and McCarthy, M. C. (2019). Gas-phase synthetic pathways to benzene and benzonitrile: a combined microwave and thermochemical investigation. *Phys. Chem. Chem. Phys.* 21, 2946–2956

Marchione, D., Mancini, L., Liang, P., Vanuzzo, G., Pirani, F., Skouteris, D., et al. (2022). Unsaturated dinitriles formation routes in extraterrestrial environments: A combined experimental and theoretical investigation of the reaction between cyano radicals and cyanoethene (C_2H_3CN). *J. Phys. Chem. A*

Marino, T., Russo, N., Sicilia, E., Toscano, M., and Mineva, T. (2000). Density functional computations and mass spectrometric measurements. can this coupling enlarge the knowledge of gas-phase chemistry? In *Advances in Quantum Chemistry* (Elsevier), vol. 36. 93–120

Matthews, D. A., Cheng, L., Harding, M. E., Lipparini, F., Stopkowicz, S., Jagau, T.-C., et al. (2020). Coupled-cluster techniques for computational chemistry: The CFOUR program package. *J. Chem. Phys.* 152, 214108

McCarthy, M. C., Lee, K. L. K., Loomis, R. A., Burkhardt, A. M., Shingledecker, C. N., Charnley, S. B., et al. (2021). Interstellar detection of the highly polar five-membered ring cyanocyclopentadiene. *Nat. Astron.* 5, 176–180

McGuire, B. A. (2022). 2021 census of interstellar, circumstellar, extragalactic, protoplanetary disk, and exoplanetary molecules. *Astrophys. J. Suppl. Ser.* 259, 30

McGuire, B. A., Burkhardt, A. M., Kalenskii, S., Shingledecker, C. N., Remijan, A. J., Herbst, E., et al. (2018). Detection of the aromatic molecule benzonitrile (c-C_6H_5CN) in the interstellar medium. *Science* 359, 202–205

McGuire, B. A., Loomis, R. A., Burkhardt, A. M., Lee, K. L. K., Shingledecker, C. N., Charnley, S. B., et al. (2021). Detection of two interstellar polycyclic aromatic hydrocarbons via spectral matched filtering. *Science* 371, 1265–1269

McNaught, A. D., Wilkinson, A., et al. (1997). *IUPAC. Compendium of Chemical Terminology*, vol. 1669 (Blackwell Science Oxford), 2nd edn.

Melosso, M., Bizzocchi, L., Dore, L., Kisiel, Z., Jiang, N., Spezzano, S., et al. (2021). Improved centrifugal and hyperfine analysis of ND_2H and NH_2D and its application to the spectral line survey of L1544. *J. Mol. Spectrosc.* 377, 111431

Melosso, M., Bizzocchi, L., Gazzeh, H., Tonolo, F., Guillemin, J.-C., Alessandrini, S., et al. (2022). Gas-phase identification of (Z)-1, 2-ethenediol, a key prebiotic intermediate in the formose reaction. *Chem. Comm.* 58, 2750–2753

Mills, I. M. (1972). Vibration-rotation structure in asymmetric-and symmetric-top molecules. *Vibration-Rotation Structure in Asymmetric-and Symmetric-Top Molecules in Molecular Spectroscopy: Modern Research* 1, 115

Møller, C. and Plesset, M. S. (1934). Note on an approximation treatment for many-electron systems. *Phys. Rev* 46, 618–622

Nakanaga, T., Ito, F., Sugawara, K., Takeo, H., and Matsumura, C. (1990). Observation of infrared absorption spectra of molecular ions, H_3^+ and HN_2^+, by FTIR spectroscopy. *Chem. Phys. Lett.* 169, 269–273

Noga, J. and Bartlett, R. J. (1987). The full CCSDT model for molecular electronic structure. *J. Chem. Phys.* 86, 7041–7050

Ohishi, M., Ishikawa, S.-i., Amano, T., Oka, H., Irvine, W. M., Dickens, J. E., et al. (1996). Detection of a new interstellar molecular ion, H_2COH^+ (protonated formaldehyde). *Astrophys. J.* 471, L61

Owrutsky, J., Gudeman, C., Martner, C., Tack, L., Rosenbaum, N., and Saykally, R. (1986). Determination of the equilibrium structure of protonated nitrogen by high resolution infrared laser spectroscopy. *J. Chem. Phys.* 84, 605–617

Papajak, E. and Truhlar, D. G. (2011). Convergent partially augmented basis sets for post-Hartree-Fock calculations of molecular properties and reaction barrier heights. *J. Chem. Theory Comput.* 7, 10–18

Peterson, K. A. and Dunning Jr., T. H. (2002). Accurate correlation consistent basis sets for molecular core-valence correlation effects: The second row atoms Al-Ar, and the first row atoms B-Ne revisited. *J. Chem. Phys.* 117, 10548–10560

Piccardo, M., Bloino, J., and Barone, V. (2015). Generalized vibrational perturbation theory for rotovibrational energies of linear, symmetric and asymmetric tops: Theory, approximations, and automated approaches to deal with medium-to-large molecular systems. *Int. J. Quantum Chem.* 115, 948–982

Puzzarini, C., Heckert, J., and Gauss, J. (2008). The accuracy of rotational constants predicted by high-level quantum-chemical calculations. I. Molecules containing first-row atoms. *J. Chem. Phys.* 128, 194108

590 Puzzarini, C., Stanton, J. F., and Gauss, J. (2010). Quantum-chemical calculation of spectroscopic
591 parameters for rotational spectroscopy. *Int. Rev. Phys. Chem.* 29, 273–367
592 Raghavachari, K., Trucks, G. W., Pople, J. A., and Head-Gordon, M. (1989). A fifth-order perturbation
593 comparison of electron correlation theories. *Chem. Phys. Lett.* 157, 479–483
594 Santra, G., Sylvetsky, N., and Martin, J. M. L. (2019). Minimally Empirical Double-Hybrid Functionals
595 Trained against the GMTKN55 Database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4.
596 *J. Phys. Chem. A* 123, 5129–5143
597 Schiff, L. I. (1955). *Quantum mechanics* (McGraw-Hill)
598 Scuseria, G. E. and Schaefer III, H. F. (1988). A new implementation of the full CCSDT model for
599 molecular electronic structure. *Chem. Phys. Lett.* 152, 382–386
600 Stanton, J. F., Gauss, J., Cheng, L., Harding, M. E., Matthews, D. A., and Szalay, P. G. (2016). CFOUR,
601 Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package With
602 contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, Y.J. Bomble,
603 O. Christiansen, F. Engel, R. Faber, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D.
604 Jonsson, J. Jusélius, T. Kirsch, K. Klein, W.J. Lauderdale, F. Lipparini, T. Metzroth, L.A. Mück, D.P.
605 O'Neill, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons,
606 S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J.
607 Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen,
608 and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see
609 http://www.cfour.de.
610 Turner, B. (1974). U93.174-a new interstellar line with quadrupole hyperfine splitting. *Astrophys. J.* 193,
611 L83–L87
612 Wootten, A., Mangum, J. G., Turner, B. E., Bogey, M., Boulanger, F., Combes, F., et al. (1991). Detection
613 of Interstellar H_3O^+: A Confirming Line. *Astrophys. J. Lett.* 380, L79
614 Yamamoto, S. (2017). Introduction to astrochemistry. *Editorial: Springer*
615 Yen, H.-W., Koch, P. M., Liu, H. B., Puspitaningrum, E., Hirano, N., Lee, C.-F., et al. (2016). Stacking
616 spectra in protoplanetary disks: detecting intensity profiles from hidden molecular lines in HD 163296.
617 *Astrophys. J.* 832, 204
618 Zhu, Y., Tian, L., Song, H., and Yang, M. (2019). Kinetic and dynamic studies of the $H_3^+ + CO \rightarrow H_2 +$
619 HCO^+/HOC^+ reaction on a high-level ab initio potential energy surface. *J. Chem. Phys.* 151, 054311