A&A, 678, A22 (2023)
https://doi.org/10.1051/0004-6361/202347059

Astronomy & Astrophysics

Oxygen-bearing organic molecules in comet 67P's dusty coma: First evidence for abundant heterocycles

N. Hänni[1], K. Altwegg[1], D. Baklouti[2], M. Combi[3], S. A. Fuselier[4,5], J. De Keyser[6], D. R. Müller[1], M. Rubin[1], and S. F. Wampfler[7]

[1] Physics Institute, Space Research & Planetary Sciences, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
e-mail: nora.haenni@unibe.ch
[2] Institut d'Astrophysique Spatiale, Université Paris-Saclay, CNRS, 91405 Orsay, France
[3] Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA
[4] Space Science Division, Southwest Research Institute, San Antonio, TX, USA
[5] Department of Physics and Astronomy, The University of Texas at San Antonio, San Antonio, TX, USA
[6] Royal Belgian Institute for Space Aeronomy, BIRA-IASB, Brussels, Belgium
[7] Center for Space and Habitability, University of Bern, Gesellschaftsstrasse 6, 3012 Bern, Switzerland

Received 31 May 2023 / Accepted 26 July 2023

ABSTRACT

The puzzling complexity of terrestrial biomolecules is driving the search for complex organic molecules in the interstellar medium (ISM) and serves as a motivation for many in situ studies of reservoirs of extraterrestrial organics, from meteorites and interplanetary dust particles to comets and asteroids. Comet 67P/Churyumov-Gerasimenko (67P), the best-studied comet to date, has been visited and accompanied for 2 yr by the European Space Agency's Rosetta spacecraft. Around 67P's perihelion and under dusty conditions, the high-resolution mass spectrometer on board Rosetta has provided a spectacular glimpse into this comet's chemical complexity. For this work, we analyzed the O-bearing organic volatiles in unprecedented detail. Through a comparison of 67P's inventory with molecules detected in the ISM, in other comets, and in soluble organic matter extracted from the Murchison meteorite, we also highlight the (pre)biotic relevance of different chemical groups of species. We report first evidence for abundant extraterrestrial O-bearing heterocycles (with abundances relative to methanol often on the order of 10% and a relative error margin of 30–50%) and various representatives of other molecule classes, such as carboxylic acids and esters, aldehydes, ketones, and alcohols. As with the pure hydrocarbons, some hydrogenated forms seem to be dominant over their dehydrogenated counterparts. An interesting example is tetrahydrofuran, as it might be a more promising candidate for searches in the ISM than the long-sought furan. Our findings not only support and guide future efforts to investigate the origins of chemical complexity in space, but they also strongly encourage the study, in the laboratory as well as by modeling, of such topics as the ratios of unbranched versus branched species and hydrogenated versus dehydrogenated species in astrophysical ice analogs.

Key words. comets: general – comets: individual: 67P/Churyumov-Gerasimenko – instrumentation: detectors – methods: data analysis

1. Introduction

Along with meteorites (e.g., Jenniskens et al. 2000) and interplanetary dust particles (IDPs; e.g., Levasseur-Regourd et al. 2018), comets are considered to be a major source of the pristine organic matter delivered to the early Earth, and hence, they may have played an important role in prebiotic chemistry and the processes that led to carbon-based life on Earth (e.g., Chyba & Sagan 1992). Using observations at comet 67P/Churyumov-Gerasimenko (hereafter, 67P), Rubin et al. (2019b) have estimated that when bringing 22% of the xenon to Earth, as proposed by Marty et al. (2017), comets also bring about the terrestrial biomass equivalent (or more) in the form of organic molecules. Assuming full conservation of the cometary organics budget during impact, this corresponds to about 17 000 to 350 000 impacts of 67P-like objects. For the main building blocks of life, which are lipids, amino acids, nucleic acids, and sugars (see, e.g., Kwok 2016), the so-called heteroatoms (i.e., atoms other than C and H, which can be O, N, S, or P) are of outstanding importance. Because of its abundance in space and in biomolecules, including amino acids, fatty acids, and sugars, the heteroelement O has received special attention in astrochemistry. Extensive efforts have been made toward the detection and study of O-bearing molecules, with possible roles in prebiotic chemistry.

The list of O-bearing molecules detected in the interstellar medium (ISM) is relatively short. As of mid-2021, McGuire (2022) reported 22 neutral O-bearing organic molecules (radicals and disputed detections excluded) that are composed of four or more atoms. The most complex molecules among them are the structural isomers ethyl formate, methyl acetate, and hydroxyacetone, with a molecular weight of 74 Da and a chemical sum formula equal to $C_3H_6O_2$. In contrast, a plethora of O-bearing molecules and macromolecules are known to be present in meteorites and, as recently published, in samples from the near-Earth carbonaceous asteroid (162173) Ryugu (Naraoka et al. 2023; Yabuta et al. 2023) as well. Oxygen is the most abundant heteroelement in soluble organic matter (SOM; e.g., Botta & Bada 2002; Schmitt-Kopplin et al. 2010) as well as in insoluble organic matter (IOM; e.g., Alexander et al. 2017). Both in IOM and SOM, the heteroelement oxygen is roughly an order of magnitude more abundant than the heteroelement

nitrogen. Ryugu SOM seems to be an exception (Naraoka et al. 2023), as only two carboxylic acids (formic and acetic acid) were detectable, apart from various amines, amino acids, and N-containing heterocycles. Naraoka et al. (2023) explain the lack of O-bearing volatile molecules by the fact that the N-bearing species they find (with basic N functionality) may have been bound in the form of ammonium salts, as already reported for Ceres (de Sanctis et al. 2016) and comet 67P (Altwegg et al. 2020, 2022; Poch et al. 2020). However, as we discuss in more detail in Sect. 4, the extraction protocol may significantly influence quantitative results. Even less abundant than nitrogen in IOM and SOM is sulfur and phosphorus. A similar heteroelement abundance distribution relative to carbon has recently been reported for the volatile organics with molecular weights of up to 140 Da in comet 67P's dusty coma (Hänni et al. 2022). Those complex organics expose an average sum formula of $C_1H_{1.56}O_{0.134}N_{0.046}S_{0.017}$. Notably, no organic source of phosphorus has been identified in comet 67P to date.

We used the same high-resolution mass spectrometric dataset as in Hänni et al. (2022), which was collected on 3 August 2015 at comet 67P by the Rosetta Orbiter Spectrometer for Ion and Neutral Analysis (ROSINA) instrument suite (Balsiger et al. 2007), and we present an extensive study of this comet's inventory of volatile O-bearing organic molecules, which is still poorly constrained. Previous work focused on data from the time when the comet was far from its perihelion and activity was low, and hence, only a limited number of species – the rather small and abundant ones – were detected in 67P's coma (Le Roy et al. 2015; Altwegg et al. 2017; Schuhmann et al. 2019b; Rubin et al. 2019a). The conditions shortly before perihelion, which is when the dataset analyzed for this work was collected, were distinct. Comet 67P was reaching its perihelion at 1.24 au from the Sun in early August 2015, and the coma was dusty. Consequently, enhanced desorption of heavier and thus more complex molecules from ejected dust particles could have occurred. Decoupled from the heat sink of the cometary surface, the temperatures of ejected particles can rise up to a few hundreds of Kelvin (Lien 1990) in sunlight, whereas comet surface temperatures around that time were observed to not exceed 240 K (Tosi et al. 2019). For the most abundant subgroup of signals of pure hydrocarbon species in the 3 August 2015 dataset, Hänni et al. (2022) demonstrated a way to decompose the observed signals into contributions of individual neutral molecules. Following the same methodological approach, we present in this work an attempt to disentangle the second-most abundant group of cometary volatile organics, namely, those of the O-bearing $C_nH_mO_x$ hydrocarbons, where n = 1–8, m = 0–14, and x = 1–2. After giving the relevant information on data acquisition as well as data reduction and interpretation, we compare and contrast our findings with other comets, meteoritic organic matter, and the ISM.

2. Instrumentation and method

2.1. Data acquisition

ROSINA's high-resolution double focusing mass spectrometer (DFMS; Balsiger et al. 2007) was designed to detect the cometary volatiles originating from both the cometary bulk material and the ejected particles. In the neutral gas mode of the instrument, neutral species were allowed to enter the ionization chamber and were ionized by electrons impacting with energies of 45 eV. The electron impact ionization (EII) process usually yields singly charged cations of the parent molecules, which are also known as molecular ions (M). However, doubly charged species can form, and the parent species can fragment during the EII process into fragment species. We note that for better readability and to avoid confusion with ions naturally present in the cometary coma, we avoid indication of the positive charges induced by the EII process. All molecular ions and charge-retaining fragments were subsequently extracted from the DFMS ionization chamber and transferred via an electrostatic analyzer and a sector magnet, where separation according to the mass-to-charge ratios (m/z) occurs, onto a stack of two micro channel plates (MCPs). The MCPs were mounted in a Chevron configuration and released an electron cascade upon impact of analyte ions. The charge was then collected on two redundant rows of position-sensitive Linear Electron Detector Array (LEDA) anodes with 512 pixels each. To scan the whole mass range analyzed in this paper, the DFMS had to measure the m/z range around each integer mass number separately and sequentially, which, including overhead time, took roughly 30 s per single mass spectrum. The maximally accessible m/z range was 12–180 in the late mission phase, while on 3 August 2015 only an m/z range of 13–140 was scanned using the following combination of measurement modes: 222, 562, and 564. Those modes allowed scanning of the m/z ranges m/z = 44–80 between 15:03–15:20 UTC, m/z = 80–140 between 15:33–16:00 UTC, and m/z = 13–43 between 16:25–16:39 UTC. For $m/z > 70$, a post-acceleration potential of 1000 V was applied in front of the detector to increase the detection efficiency for heavy species.

Although observed at other instances (Altwegg et al. 2017, 2020), on 3 August 2015 there was no sign of particles sublimating inside or close to the DFMS ionization chamber, such as a sudden increase of the signal strengths or a blockage of electron or ion currents. We find it important to note, however, that even these dust events do not result in relevant energetic processing of the analyte prior to EII, as the particle's relative velocity is merely a few meters per second. The detected species are thought to be present in their free form either in the ice- or dust-dominated matrix from which they thermally desorb. Hänni et al. (2022), furthermore, argued in detail that the observed species are likely to have origins in the early history of our Solar System and are not decomposition products of cometary macromolecular matter (Fray et al. 2016).

2.2. Data reduction

The DFMS achieved a mass resolution of $m/\Delta m$ = 3000 considering the full width at 1% of the peak height of the molecular nitrogen peak at m/z = 28 (Balsiger et al. 2007). This resolution is sufficient to distinguish pure hydrocarbons from heteroatom-bearing hydrocarbons, which we exploit in this present analysis. As an example, Fig. 1 shows the mass spectrum collected around m/z = 105. The two registered signals, which are associated with C_7H_5O and C_8H_9, are well separated. See also the supplementary Fig. 2 in Hänni et al. (2022), which shows the mass spectrum around m/z = 109 with three well-separated signals of C_5H_3NS, C_7H_9O, and C_8H_{13}. The relative peak position is defined when two or more signals are present in the mass spectrum. The absolute peak position and hence the mass scale are defined by one correct assignment of a species (i.e., its exact mass) to a peak in the mass spectrum, usually of C_nH_m and C_nH_mX species, where X is a heteroatom. Continuation of the homologous series in the adjacent mass spectra as $C_nH_{m\pm1}$ and $C_nH_{m\pm1}X$ is frequently observed and helpful for verification of the mass scale. In addition, regularly measured major cometary species like H_2O, CO, CO_2, OCS, or CS_2 are used to

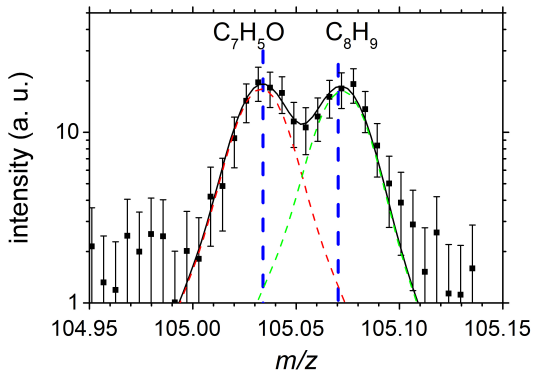

Fig. 1. DFMS mass spectrum collected on 3 August 2015 around integer mass 105 (black square markers). The peaks – here associated to C_7H_5O and C_8H_9 at the exact positions indicated by the dashed blue lines – are fitted with double-Gaussian functions (dashed red and green lines). The standard deviation of the fitted peak position is $\approx 6 \times 10^{-4}$.

define the mass scale. Selected candidate molecules are also verified by relative elemental and isotopic abundance considerations (Rubin et al. 2019a; Lodders 2021; Hänni et al. 2022). The signal intensities are extracted from fitting a double-Gaussian peak profile to each mass peak (see, e.g., Le Roy et al. 2015) and corrected for the mass-dependent instrument sensitivity. However, for low-intensity signals with amplitudes of less than an order of magnitude above background level, such as those of some heavier heteroatom-bearing hydrocarbon species, the second Gaussian disappears in the background. After fitting the spectra and summing both LEDA rows, the obtained intensities were corrected for a drift in the signal strength. As detailed in Hänni et al. (2022), this drift originates from the combination of spacecraft slew, comet rotation, and variability in the coma composition as well as thermal changes that occurred during the time period of interest. The drift was corrected using repetitively measured mass ranges. The estimated errors on the extracted intensities of C_nH_mO signals are ≈30%, and those of $C_nH_mO_2$ signals are ≈50%. This grouping is assumed for simplicity. However, there are slight differences for each individual species. These overall errors include: (1) a statistical error (for C_nH_mO species, which are more abundant, the statistical error is a minor contribution and can be neglected, while for $C_nH_mO_2$ species, which are less abundant, the statistical uncertainty is roughly 20%), (2) an error related to detector gain (all species were measured with the highest gain step, which means that for relative abundances, related uncertainties cancel out), (3) a fit error and an individual pixel gain error (together, they are estimated to be approximately 15%), and (4) a residual systematic error related to the drift (this error is estimated to be around 15%).

To derive the absolute abundance of the contributing molecules, their fragmentation patterns must be taken into account. Here, this was done based on the reference mass spectra from NIST, which was obtained for 70 eV electron impact. Therefore, the measured signals have to be corrected for the ionization cross section of the analyte molecules as well as a species-dependent sensitivity, accounting for the variable sizes, masses, and forms of the ions. Unfortunately, data only for a small set of simple molecules for the cross sections are available in the literature. The same limitations account for the species-dependent sensitivities, which for the DFMS have been studied only for a limited set of the molecules investigated in this work. As we do not consider these factors, we provide the individual abundances in arbitrary units. For the purpose of inter-comparison, it is most suitable to consider the relative abundances of a specific species rather than the absolute abundances. However, normalization with respect to water, the major cometary volatile, is not suitable for this investigation. The data we used here were collected during a time period where sublimation from ejected dust grains was strongly enhanced in comparison to sublimation from the cometary bulk material, and water is expected to be depleted. The complex organic species seem to be correlated with methanol, see Rubin et al. (in prep.) or Fig. 5 in Läuter et al. (2020). Consequently, we normalized our abundance values to methanol, which itself does have a negligibly small error. The methanol signal stands alone and is fitted well, and it led to a statistical error of less than 1%. The individual pixel gain error is roughly 5%. Combining the errors of the other species previously discussed, the errors on the abundances relative to methanol are approximately 30% for C_nH_mO molecules and 50% for $C_nH_mO_2$ molecules.

2.3. Deconvolution of fragmentation patterns

The so-called EII fragmentation pattern of a given molecule is the characteristic fingerprint of signals being registered when the *m/z* range is scanned while the parent and fragment ions are being collected. The pattern is usually graphed in terms of intensity on the *y*-axis and *m/z* on the *x*-axis. Such mass spectra are collected in reference databases such as that from the National Institute of Standards and Technology (NIST; Steins 2018), which is the most extensive, as far as we know. The electron energy used to acquire NIST standard mass spectra is usually 70 or 75 eV, which is different from the 45 eV used by the DFMS. However, the deviations in the fragmentation patterns resulting from this difference are minor (Schuhmann et al. 2019a,b), and the related error is small compared to other sources of uncertainty. Furthermore, the NIST reference spectra only come in unit-resolution. For certain signals, this results in ambiguities as to whether the O atom is retained in the fragment species (O-bearing fragment) or not (pure hydrocarbon fragment) because both fragment types could, and in some cases, do exist. In the following list, we briefly describe, for the most important O-bearing chemical functional groups, how the low-resolution mass spectra available from NIST are used to interpret our high-resolution data (we use '–' to indicate chemical bonds and '-' to indicate that an atom/group of atoms was lost from the molecular ion):

– Carboxylic acids (R–COOH, where R is a residual): Usually, carboxylic acids produce a detectable molecular ion signal (M signal). The acids commonly fragment under loss of the hydroxy group, leading to a prominent M-OH signal.

– Carboxylic acid esters (R–COO–R): Esters usually fragment due to bond cleavage next to the carboxyl function. That is, the alkoxy group (–OR) is lost while hydrogen atoms are rearranged. Despite being weak, M is mostly visible.

– Aldehydes (R–CHO): For aldehydes, M is observable, while M-H is mostly weak or even absent. Fragmentation typically occurs under retention of the O atom while a part of R is lost, for example, a methyl group leading to an M-CH_3 signal.

– Ketones (R–CO–R): For ketones, major fragments originate from cleavage of the CC bond adjacent to the carboxyl function. The M signal is usually observed but rather moderate.

– Alcohols (R–OH): When alcohols fragment, favorably, water is lost, leading to a pure hydrocarbon signal at M-H_2O. Especially for primary alcohols, the resonance-stabilized CH_3O

cation on $m/z = 31$ is energetically favorable and usually of high intensity.

– Ethers (R–O–R): For ethers, usually the main species are M and M-H.

The loss of a water molecule from the molecular ion, leading to an M-H_2O signal, is common not only for molecules containing a hydroxyl function, such as alcohols or carboxylic acids, but also for aldehydes as well as cyclic and straight chain aliphatic ketones (see Yeo & Williams 1969 and literature cited therein). The rules above may not perfectly capture the fragmentation behavior of every molecule with the respective functional group, but they may well serve as general guidelines to help decide cases where relevant signals are ambiguous in low-resolution NIST reference spectra.

When the analyte is a complex mixture of species, such as in the instance of the cometary coma, then the characteristic fragmentation patterns of all species present in that mixture overlap (i.e., multiple species can contribute to a given mass peak). As a result, a sum fragmentation pattern is observed. To determine which species have contributed to the result, the sum fragmentation pattern must be deconvolved. Hänni et al. (2022) demonstrated for the dominant subgroup of pure hydrocarbon species, C_nH_m, how the empirical principle of Occam's Razor (e.g., Clauberg 1654) can be employed to find a solution to the deconvolution problem. In this work, we applied the same method to the subgroup of $C_nH_mO_x$ species (parents and fragments), which depicts the second-most abundant group of species, which is roughly half as abundant as the pure hydrocarbons.

Such a deconvolution attempt comes with a number of caveats: First, the fragmentation pattern has been experimentally determined only for a fraction of the large number of complex organic molecules in our mass range of interest. Second, even if reference spectra are available, some molecules do not yield significant M signals and thus cannot be unambiguously identified. This is especially the case for longer carbon chain species such as pentanoic acid, pentanal, or 1-pentanol. Third, structural isomers possess exactly the same mass and sometimes have very similar fragmentation patterns, which could render them indistinguishable in mass spectrometry. It is clear that larger molecules have more structural isomers (including exotic ones), and not all the mass spectra of those isomers have been measured. Consequently, the selection of candidates becomes more ambiguous for larger molecules. It could be assumed, however, that the isomers with experimental data tend to be more stable (notably under terrestrial conditions) and, hence, might be more likely to be present in comets too.

Moreover, we have based our selection on arguments of plausibility. If a complex organic molecule has a very small M reference signal, the amount needed to explain the intensity observed by the DFMS may be very large (i.e., on the order of magnitude of very abundant cometary species like CO and CO_2). Even if such a molecule could not be discarded based on overshooting fragment signals, we still excluded the species based on the improbably high abundance. Even though additional molecules are likely present in the cometary coma beyond those captured by our deconvolution attempt, their total amount is limited by the uncertainties in the measured data. There are also cases where two structural isomers are very likely present in the cometary coma but their abundance ratio is badly constrained based on our data (e.g., glycolaldehyde and methyl formate; see Sect. 3).

To summarize, the solution we present in the following is the most likely solution obtained from applying Occam's razor to the available reference data while also taking knowledge on cometary molecules into consideration. However, we cannot exclude a scenario where many more molecules (i.e., the structural isomers of the candidates we selected) contribute to the observed overall fragmentation pattern. They could either add to the overall fragmentation pattern within the estimated error margins or (partially) replace some of the molecules, leading to decreased individual abundances of some of the larger candidate molecules. Consequently, our Occam's razor-based method might underestimate the diversity of O-bearing cometary species, especially those with higher molecular weights. We include these considerations when discussing our interpretation of the measured DFMS mass spectra in the next sections.

3. Results

Our Occam's razor-based deconvolution of the $C_nH_mO_{1-2}$ species observed on 3 August 2015 in comet 67P's dusty coma is shown in Figs. 2 and 3 and is explained subsequently. Table A.1 lists all selected candidate molecules, grouped according to their primary chemical functionality. Both CO and CO_2 are listed as reference species. To present our results in an understandable and compact way to the reader, we combined structural isomers and list them according to increasing molar masses. Moreover, we indicate which isomers have been observed in comets (Biver & Bockelée-Morvan 2019) and in the ISM (McGuire 2022) because we sometimes consider those identifications to support the interpretation of our data. We note that for simplicity, we do not reference the original literature, though it can nonetheless be found in the respective review papers. For each of the selected candidate molecules, we denote the characteristic signals we used for identification and list the estimated abundance relative to methanol (ARM) as a percentage. We discuss possible alternative isomers and assign every identification with a level-of-confidence indicator (LCI) according to the following scheme: 1 = maybe present, 2 = likely present, 3 = present with great certainty. However, our LCIs are based on the available reference data and knowledge on cometary and ISM species. For example, if a specific molecule is the only isomer with reference data in NIST and is in suitable agreement with our data, it will receive an LCI of 3 despite the fact that it is possible that other isomers without mass spectrum in NIST could produce an even better match.

After this procedure, some observed intensity remained unexplained (i.e., blanks in Figs. 2 and 3). This intensity is mostly associated with species with a low number of hydrogen atoms (compared to the fully saturated molecule): C_3H_2O, C_4HO, C_5H_3O, C_6H_5O, $C_2H_3O_2$, and C_3HO_2. Some of those species might be radicals for which no NIST reference data exists. Hänni et al. (2022) observed a similar situation for the pure hydrocarbon species and provided additional comments. Once the full cometary inventory of complex organics is done, these residual intensities can be investigated systematically.

3.1. C_nH_mO species

– CO ($m/z = 28$): Carbon monoxide (no. 1; LCI = 3; ARM = 292), here identified via M, was reported to be present in comet 67P previously (Le Roy et al. 2015; Rubin et al. 2019a), studied in situ in great spatial and temporal detail with multiple instruments (Biver et al. 2019; Läuter et al. 2020; Combi et al. 2020), and observed remotely in many comets (A'Hearn et al. 2012; Dello Russo et al. 2016) as well as in the ISM (McGuire 2022).

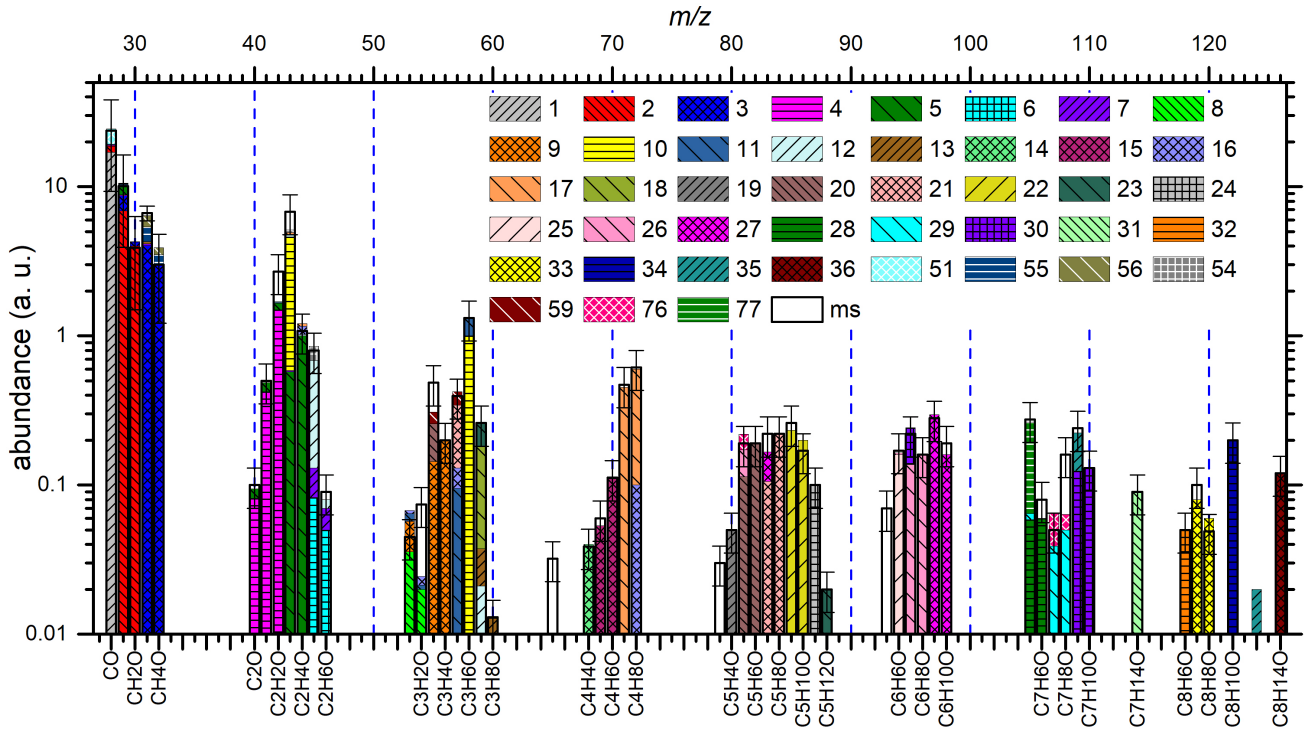

Fig. 2. Occam's razor-based deconvolution of the subset of signals associated with the C_nH_mO species as detected by the DFMS on 3 August 2015. The measured signals (ms) are given in arbitrary units (a. u.) with 30% error margins. With the color coding, we show the different contributions of the individual molecules to the observed sum fragmentation pattern of the C_nH_mO species (see the main text for details). The following 36 molecules were selected: (1) carbon monoxide, (2) formaldehyde, (3) methanol, (4) ketene, (5) acetaldehyde, (6) dimethyl ether, (7) ethanol, (8) 2-propynal, (9) 2-propenal, (10) acetone, (11) propanal, (12) isopropanol, (13) *n*-propanol, (14) furan, (15) 2,3-dihydrofuran, (16) butanal, (17) tetrahydrofuran, (18) 2-butanol, (19) 2,4-cyclopentadiene-1-one, (20) 2-methylfuran, (21) 2,3-dihydro-4-methylfuran, (22) tetrahydropyran, (23) 1-ethoxypropane, (24) 2-methoxy-2-methylpropane, (25) phenol, (26) 2,5-dimethylfuran, (27) 3-methoxycyclopentene, (28) benzaldehyde, (29) benzyl alcohol, (30) 2,3,5-trimethylfuran, (31) 2-methylcyclohexanol, (32) benzofuran, (33) 4-methylbenzaldehyde, (34) ethoxybenzene, (35) 6-methyl-3,5-heptadien-2-one, (36) 2,6-dimethylcyclohexanone, (51) carbon dioxide, (54) acetic acid, (55) glycolaldehyde, (56) methyl formate, (59) propanoic acid, (76) hydroquinone, (77) benzoic acid. For referencing and to distinguish the two populations, the selected C_nH_mO molecules have been given numbers below 50, while $C_nH_mO_2$ molecules were given numbers above 50. Some of the latter produce relevant C_nH_mO fragments, which we show in this figure and have thus listed.

– CH_2O ($m/z = 30$): Formaldehyde (no. 2; LCI = 3; ARM = 206) is identified here unambiguously via M and M-H, and it was previously reported for comet 67P (Schuhmann et al. 2019b). This molecule was also observed in other comets from the ground (Biver & Bockelée-Morvan 2019) and detected in many sources in the ISM (McGuire 2022).

– CH_4O ($m/z = 32$): Methanol (no. 3; LCI = 3; ARM = 100), unambiguously identified here via M and M-H, was detected in comet 67P previously (Schuhmann et al. 2019b). It was studied from the ground in various comets (Biver & Bockelée-Morvan 2019; Dello Russo et al. 2016) and is commonly detected in the ISM (McGuire 2022).

– C_2H_2O ($m/z = 42$): Ketene (no. 4; LCI = 3; ARM = 71) is identified here via M with high certainty. The isomers oxirene and ethynol do not have fragmentation patterns in NIST. Though oxirene could be produced in methanol-acetaldehyde matrices from isomerization of ketene (Wang et al. 2023), it has not yet been detected in extraterrestrial environments. Ketene was observed in other comets (Biver & Bockelée-Morvan 2019) and in the ISM (McGuire 2022).

– C_2H_4O ($m/z = 44$): Acetaldehyde (no. 5; LCI = 3; ARM = 150), identified here via M and M-H, has been reported for comet 67P previously (Schuhmann et al. 2019b). Oxirane, detected in the ISM (as reviewed by McGuire 2022), produces less intensity on $m/z = 43$ and is thus a less favorable candidate. However, we cannot completely rule out a marginal presence at the expense of some acetaldehyde. Ethenol does not have fragmentation data in NIST.

– C_2H_6O ($m/z = 46$): Dimethyl ether (no. 6; LCI = 3; ARM = 3.1) and ethanol (no. 7; LCI = 3; ARM = 4.1) must both be present in 67P's coma with high certainty to explain the observed sum fragmentation pattern. Acetone (no. 10) contributes to the same group of signals, too (see below). Ethanol produces signals on M, M-H, and M-CH_3 (i.e., CH_3O on $m/z = 31$), while dimethyl ether shows on M and M-H. Ethanol has been identified in comet 67P previously (Schuhmann et al. 2019b) and has been reported for other comets (Biver & Bockelée-Morvan 2019) and the ISM (McGuire 2022). In the ISM, dimethyl ether has been convincingly identified (McGuire 2022), whereas for comets, only an upper limit value is available (Biver & Bockelée-Morvan 2019).

– C_3H_2O ($m/z = 54$): 2-Propynal (no. 8; LCI = 2; ARM = 1.7) does not match the observed M/M-H ratio well, which makes this isomer less likely to be the sole isomer contributing to the observed sum fragmentation pattern analyzed in this work. However, other isomers do not have fragmentation data in NIST and hence cannot be included in this analysis. Propynal has been reported in the ISM (McGuire 2022).

– C_3H_4O ($m/z = 56$): 2-propenal (no. 9; LCI = 2; ARM = 29) is identified via M and M-H signals, but M-H cannot be explained solely by 2-propenal and needs contributions from

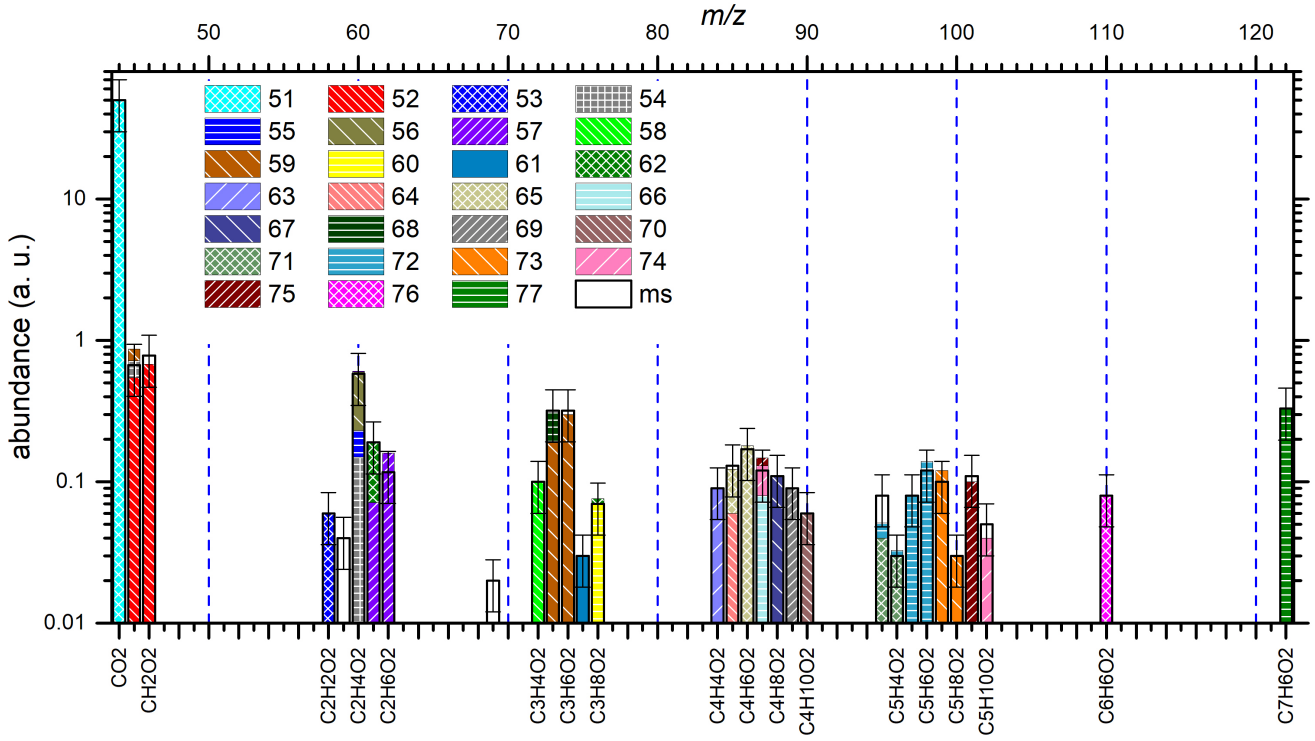

Fig. 3. Occam's razor-based deconvolution of the subset of signals associated with the $C_nH_mO_2$ species as registered by the DFMS on 3 August 2015. The measured signals (ms) are given in arbitrary units (a. u.) with 50% error margins. With the color coding, we show the different contributions of the individual molecules to the observed sum fragmentation pattern of the $C_nH_mO_2$ species (see details in the main text). The following 27 molecules have been selected: (51) carbon dioxide, (52) formic acid, (53) glyoxal, (54) acetic acid, (55) glycolaldehyde, (56) methyl formate, (57) ethylene glycol, (58) 2-propenoic acid, (59) propanoic acid, (60) 2-methoxyethanol, (61) methylal (62) 1,2-propanediol, (63) 2(3H)-furanone, (64) cyclopropanecarboxylic acid, (65) γ-butyrolactone, (66) 1,3-dioxane, (67) 1,4-dioxane, (68) butanoic acid, (69) 1-ethoxy-1-methoxyethane, (70) diethyl peroxide, (71) 3-furaldehyde, (72) 2-furanmethanol, (73) cyclopropanecarboxylic acid methyl ester, (74) propanoic acid ethyl ester, (75) 4-methyl-1,3-dioxane, (76) hydroquinone, and (77) benzoic acid.

other species with higher masses. In addition, structural isomers of 2-propenal might also make a contribution. However, other isomers do not have fragmentation data in NIST and hence cannot be included in this analysis. We note that 2-propenal has been reported in the ISM (McGuire 2022).

– C_3H_6O; $m/z = 58$: Acetone (no. 10; LCI = 3; ARM = 118) and propanal (no. 11; LCI = 3; ARM = 48) are both needed to explain the sum fragmentation pattern investigated in this work: propanal in order to explain the observed signal on $m/z = 57$ and acetone to explain that on $m/z = 43$. Propanal and/or acetone have been identified in 67P previously but could not be distinguished (Schuhmann et al. 2019b). Both of these isomers have been detected in the ISM (McGuire 2022), but to date, only acetone has been observed in comets from the ground (Biver & Bockelée-Morvan 2019). The ISM hosts yet another structural isomer, namely, propylene oxide, which is also known as 2-methyloxirane (McGuire 2022). Propylene oxide may or may not be marginally present at the expense of some acetone and propanal in comet 67P. Moreover, it is almost indistinguishable from methoxyethene based on its fragmentation pattern. Notably, no C_nH_mO species heavier than 58 Da have been firmly detected in the ISM as of mid-2021, which is the cutoff time of the census by McGuire (2022).

– C_3H_8O ($m/z = 60$): Both isopropanol (no. 12; LCI = 2; ARM = 19) and *normal*-propanol, or *n*-propanol, (no. 13; LCI = 3; ARM = 5.0) must be present to appropriately explain the observed sum signals. However, the fragmentation pattern of methoxyethane is similar to that of isopropanol, and its presence cannot be excluded. Methoxyethane has recently been proposed in the framework of a reanalysis of COSAC data collected from 67P's surface dust (Leseigneur et al. 2022), see details below. Though propanol has not been observed in any comet previously, it has been observed in the galactic center source Sagittarius B2 (Sgr B2), for which Belloche et al. (2022) reported the presence of both the normal and the iso species.

– C_4H_4O ($m/z = 68$): In this work, furan (no. 14; LCI = 2; ARM = 1.7) is identified with certainty via M, as most other isomers yield M and M-H (where no signal is observed). Only 3-butyn-2-one could partially contribute. However, M-CH_3 ($m/z = 53$) strongly limits its contribution. Furan has been searched for in various sources in the ISM, but to date only upper limits have been reported (see discussion below). Hänni et al. (2022) already mentioned the presence of furan based on the dataset evaluated here, but no other detections in comets are known to date.

– C_4H_6O ($m/z = 70$): 2,3-dihydrofuran (no. 15; LCI = 2; ARM = 8.3) was identified with high certainty via M and M-H. Most likely, both isomers of dihydrofuran contribute, 2,3-dihydrofuran and 2,5-dihydrofuran. Additional contributions of other isomers, such as 2-butenal, cyclopropanecarboxaldehyde, or (*Z*)-1,3-butadien-1-ol, cannot be fully ruled out because they have similar fragmentation patterns. In most cases, some of the fragments are limiting. Hänni et al. (2022) already reported the presence of dihydrofuran based on the same dataset evaluated in this work, but no other detections in comets are known to date.

– C_4H_8O ($m/z = 72$): Butanal (no. 16; LCI = 3; ARM = 12) and tetrahydrofuran (THF; no. 17; LCI = 3; ARM = 64)

are both present with high certainty to explain the observed signals. Butanal yields M but no M-H, while THF yields M and M-H. Ethyloxirane could replace them, but the abundance would have to be very high and would lead to an overshooting M-CH_3 fragment on $m/z = 57$. Moreover, in that case, oxirane should probably be dominant, too. Butanal (or isomers of it) has been identified in 67P previously (Schuhmann et al. 2019b), and THF has already been mentioned in Hänni et al. (2022). Although the furan group of molecules can explain the peak group $m/z = 68$–72 well, minor contributions of other structural isomers cannot be ruled out, and errors on the estimated relative abundances are considerable.

– $C_4H_{10}O$ ($m/z = 74$): 2-butanol (no. 18; LCI = 2; ARM = 8.3) is probably present and in this work mainly identified via its M-CH_3 fragment (M is very small). We note that, 1-Butanol may be contributing as well. However, it cannot be identified with high certainty, and if present, its contribution is limited.

– C_5H_4O ($m/z = 80$): 2,4-cyclopentadiene-1-one (no. 19; LCI = 2; ARM = ?) is the only molecule with this sum formula registered in NIST, but it does not have fragmentation data. However, a cyclic ketone is expected to produce a strong M signal (and likely an M-H signal too), which we assume here. We cannot rule out that the other isomers not listed in NIST may produce M and M-H signals in an appropriate ratio and hence might be more suitable. Due to the missing fragmentation pattern of 2,4-cyclopentadiene-1-one, we cannot deduce a relative abundance. However, since cyclopentadiene is a relatively abundant molecule among 67P's complex organics (cf. Hänni et al. 2022), we consider 2,4-cyclopentadiene-1-one a probable species.

– C_5H_6O ($m/z = 82$): We identified 2-methylfuran (no. 20; LCI = 3; ARM = 11) in this work with high certainty via M and M-H. The position of the methyl group can be 2 or 3 and is most probably a mixture of both. Other isomers with fragmentation data in NIST do not produce M-H. Unfortunately, for 2H- and 4H-pyran, no fragmentation pattern is available, but they are expected to be present, at least in low abundance, in analogy to the furan group of molecules.

– C_5H_8O ($m/z = 84$): We identified 2,3-dihydro-4-methylfuran (no. 21; LCI = 3; ARM = 20) via M and M-H signals with high certainty. Other isomers with fragmentation data available in NIST do not produce significant M-H signals and are hence less likely, according to Occam's razor.

– $C_5H_{10}O$ ($m/z = 86$): Tetrahydropyran (no. 22; LCI = 3; ARM = 25) is identified via M and M-H signals with high certainty. Other isomers with available fragmentation data in NIST do not produce comparable M-H signals and, hence, are less likely candidates based on Occam's razor.

– $C_5H_{12}O$ ($m/z = 88$): Both 1-ethoxypropane (no. 23; LCI = 3; ARM = 3.9) and 2-methoxy-2-methylpropane (no. 24; LCI = 3; ARM = 0.2) are needed to explain M (1-ethoxypropane) and M-H (2-methoxy-2-methylpropane). No isomer with fragmentation data available in NIST explains M and M-H simultaneously.

– C_6H_6O ($m/z = 94$): Phenol (no. 25; LCI = 2; ARM = 7.7) was identified in this work via strong M. Despite the fact that vinyl-furan has a very similar fragmentation pattern, a substantial contribution of this molecule is probably not likely because it would have to be more abundant than furan itself. However, Hänni et al. (2022) reported considerable abundances of benzene, which makes phenol a reasonable candidate species. We note, though, that many isomers do not have fragmentation data in NIST, and it is quite possible that one of them may fit the observed sum fragmentation pattern better or at least contribute to it.

– C_6H_8O ($m/z = 96$): With high certainty, 2,5-dimethylfuran (no. 26; LCI = 3; ARM = 11) can be identified via M, M-H, and M-CH_3. However, the positions of the two methyl groups are not clearly defined, and the contribution of isomers with other positions or an isomer mixture is possible.

– $C_6H_{10}O$ ($m/z = 98$): 3-methoxycyclopentene (no. 27; LCI = 3; ARM = 33) is identified via M, M-H, and M-CH_3 with high certainty. It is the only available isomer with a strong M-H signal and a reasonable candidate molecule, given the relatively high abundance of cyclopentene (Hänni et al. 2022).

– C_7H_6O ($m/z = 106$): Benzaldehyde (no. 28; LCI = 3; ARM = 4.1) is identified via strong M and M-H signals with high certainty and was already mentioned in Hänni et al. (2022). Isomers with fragmentation data available in NIST do not produce M-H and, hence, are less likely candidates.

– C_7H_8O ($m/z = 108$): Benzyl alcohol (no. 29; LCI = 1; ARM = 4.1), identified here via strong M and M-H in the suitable ratio, was already mentioned in Hänni et al. (2022). Benzyl alcohol is most likely part of a mixture of different structural isomers based on a substituted benzene ring, as these molecules show similar fragmentation patterns. The presence of anisole might especially add to the observed intensity on $m/z = 108$, which cannot be explained by benzyl alcohol alone.

– $C_7H_{10}O$ ($m/z = 110$): 2,3,5-trimethylfuran (no. 30; LCI = 2; ARM = 9.3) is identified via strong M and M-H signals in a suitable ratio. The presence of this molecule is probable because other isomers with fragmentation data available from NIST do not match the observed data as well.

– $C_7H_{14}O$ ($m/z = 114$): 2-methylcyclohexanol (no. 31; LCI = 1; ARM = 30) is identified via strong M and absent M-H. However, several other structural isomers expose a similar fragmentation pattern and, hence, the identification comes with a low level of certainty.

– C_8H_6O ($m/z = 118$): Benzofuran (no. 32; LCI = 3; ARM = 1.9) can be identified via strong M and absent M-H signals. As other structural isomers with fragmentation data available from NIST do not match well with the observed data, we consider benzofuran to be a very likely candidate.

– C_8H_8O ($m/z = 120$): 4-methylbenzaldehyde (no. 33; LCI = 2; ARM = 4.3) is identified via strong M and M-H signals. The position of the methyl group can be 2, 3, or 4, but it cannot be distinguished in the framework of this analysis because the fragmentation patterns of the different isomers are too similar. A mixture of multiple isomers is likely present.

– $C_8H_{10}O$ ($m/z = 122$): Ethoxybenzene (no. 34; LCI = 1; ARM = 19) is identified via strong M. However, the identification has a low certainty because other isomers with similar fragmentation patterns could and likely do contribute.

– $C_8H_{12}O$ ($m/z = 124$): We identified 6-methyl-3,5-heptadien-2-one (no. 35; LCI = 1; ARM = 8.3) via M and M-CH_3. Via a M-CH_3 fragment, this molecule contributes to the underpopulated $m/z = 109$ measured signal. However, some of the other isomers might also lose a methyl group and thus may have a similar effect. Hence, the identification has a low certainty.

– $C_8H_{14}O$ ($m/z = 126$): 2,6-dimethylcyclohexanone (no. 36; LCI = 1; ARM = 19) is a possible candidate identified via strong M. But some other cyclic isomers may and likely do contribute to the observed data.

3.2. $C_nH_mO_2$ species

– CO_2 ($m/z = 44$): This work identifies carbon dioxide (no. 51; LCI = 3; ARM = 1071) via its strong M and M-O (i.e., CO

on $m/z = 28$) signals. Like carbon monoxide, carbon dioxide has been identified in comet 67P previously (Le Roy et al. 2015; Migliorini et al. 2016; Fink et al. 2016; Rubin et al. 2019a) and was studied in great spatial and temporal detail (Combi et al. 2020; Läuter et al. 2020). Carbon dioxide is not observable via its rotational spectrum from the ground, as the molecule is missing a permanent dipole moment. Snodgrass et al. (2017) imaged 67P's CO_2 coma in the infrared in their Fig. 8. An overview of CO_2 measurements for other comets can be found, for example, in A'Hearn et al. (2012). In the ISM, CO_2 ice has been detected in absorption (McGuire 2022).

– CH_2O_2 ($m/z = 46$): Formic acid (no. 52; LCI = 3; ARM = 36), identified here with great certainty via M and M-H, was previously reported in comet 67P (Schuhmann et al. 2019b). It was also detected in various other comets (Biver & Bockelée-Morvan 2019) and the ISM (McGuire 2022).

– $C_2H_2O_2$ ($m/z = 58$): Identified via strong M, glyoxal (no. 53; LCI = 2; ARM = 2.5), also known as ethanedial, is possibly present, but the derived abundance is probably wrong. This is because the fragmentation pattern in NIST seems to be faulty. Specifically, no combination of atoms can explain the strong fragment signal on $m/z = 31$, which represents almost 90% of the intensity of the main signal on $m/z = 29$ (CHO cation). We hypothesize that the reference sample might have been contaminated with water-adduct or multimer species. However, the structural isomers, acetylenediol and acetolactone, do not have fragmentation data in NIST.

– $C_2H_4O_2$ ($m/z = 60$): Acetic acid (no. 54; LCI = 3; ARM = 11), glycolaldehyde (no. 55; LCI = 3; ARM = 44), and methyl formate (no. 56; LCI = 3; ARM = 44) are very likely present simultaneously. Acetic acid is clearly identified via strong M, M-CH_3, and M-OH and has been previously reported to be present in comet 67P by Schuhmann et al. (2019b). A combination of acetic acid and methyl formate and/or glycolaldehyde (identified via strong M) is necessary to explain M appropriately while not overshooting M-CH_3. Both methyl formate and glycolaldehyde have been identified in comets from the ground (Biver & Bockelée-Morvan 2019). Notably, glycolaldehyde was quite recently identified in comet C/2014 Q2 Lovejoy (Biver et al. 2015). However, based on the mass spectrometric data analyzed here, their abundances are not well constrained. This is because their fragmentation patterns are very similar, and we have artificially fixed their ratio to 1. Notably, the glycoladehyde-methyl formate ratio is about 0.2 in comet C/2014 Q2 Lovejoy (Biver & Bockelée-Morvan 2019).

– $C_2H_6O_2$ ($m/z = 62$): Ethylene glycol (no. 57; LCI = 3; ARM = 50) is identified with high certainty via M and M-H. A limited contribution of the structural isomer dimethyl peroxide is possible, but dimethyl peroxide alone cannot sufficiently explain M-H. Ethylene glycol has been identified in comet 67P previously (Schuhmann et al. 2019b) and was reported for other comets (Biver & Bockelée-Morvan 2019) and the ISM (McGuire 2022).

– $C_3H_4O_2$ ($m/z = 72$): In this work, we identified 2-propenoic acid (no. 58; LCI = 3; ARM = 6.3) via strong M. Other isomers with fragmentation data available from NIST do not have strong M signals and are thus less likely.

– $C_3H_6O_2$ ($m/z = 74$): Propanoic acid (no. 59; LCI = 3; ARM = 39), here clearly identified via M and M-H, was already mentioned by Hänni et al. (2022). Moreover, methyl acetate (or isomers of it) was reported to be present in comet 67P (Schuhmann et al. 2019b). In the ISM, ethyl formate, methyl acetate, and hydroxyacetone were detected (McGuire 2022). However, based on the data evaluated here, all three of these isomers are less favorable candidates than propanoic acid because they produce few to no M signals and no M-H signals. If present, their abundances cannot be constrained well by mass spectrometry alone.

– $C_3H_8O_2$ ($m/z = 76$): 2-methoxyethanol (no. 60; LCI = 3; ARM = 27), methylal (no. 61; LCI = 2; ARM = 4.0), and 1,2-propanediol (no. 62; LCI = 2; ARM = 24) are most likely present simultaneously. The combination of 2-methoxyethanol (yields M and a strong C_2H_5O fragment on $m/z = 45$), 1,2-propanediol (yields M-CH_3 but no relevant M or M-H signals), and methylal (yields M-H) provides a sufficient explanation of the observed data. No other isomers with fragmentation data in the NIST, apart from 1,2-propanediol, could contribute substantially to the observed signal on $m/z = 61$, and only 2-methoxyethanol yields relevant M. Consequently, the presence of both of these molecules is considered likely, with 1,2-propanediol being a bit less likely than 2-methoxyethanol because the M signal can be used for identification. Methylal is identified with less confidence too, as other isomers could yield M-H signals, but for methylal, the relative intensity of M-H is greatest.

– $C_4H_4O_2$ ($m/z = 84$): 2(3H)-furanone (no. 63; LCI = 2; ARM = 12) is identified via a strong M signal. For the position of the reduced C atom in the furan core structure, 3H or 5H are equally possible. Also, a relevant contribution of the structural isomer 2-butynoic acid cannot be ruled out based on the similarity of the reference fragmentation patterns. Unfortunately, the structural isomers 1,4-dioxine and 1,2-dioxine do not have fragmentation data in NIST. But in analogy to the furan and pyran groups of molecules, dioxines (more likely the non-peroxide isomer) may be present too, given that dioxanes are likely present (cf. no. 66 and no. 67 below).

– $C_4H_6O_2$ ($m/z = 86$): Cyclopropanecarboxylic acid (no. 64; LCI = 1; ARM = 12) is likely contributing to M-H, while γ-butyrolactone (no. 65; LCI = 1; ARM = 66) explains the observed M signal (and contributes to M-H). As cyclopropanecarboxylic acid does not have a relevant M signal, the reference data was normalized to M-H on $m/z = 85$. Other combinations of isomers are possible but slightly less likely, and many isomers with fragmentation data in NIST do not yield significant M-H.

– $C_4H_8O_2$ ($m/z = 88$): The combination of 1,3-dioxane (no. 66; LCI = 3; ARM = 6.2; strong M-H but no M) and 1,4-dioxane (no. 67; LCI = 3; ARM = 13; strong M but no M-H) explains the observed data well. 1,2-dioxane is a peroxide and a less stable structural isomer with no fragmentation data available from NIST. However, peroxides have been observed in comet 67P (see diethyl peroxide, no. 70, below), and we cannot rule out its contribution here. Butanoic acid (no. 68; LCI = 2; ARM = 13) is needed, in addition to the dioxane isomers, to explain M-CH_3 on $m/z = 73$ and is thus considered to be likely present as well.

– $C_4H_{10}O_2$ ($m/z = 90$): 1-ethoxy-1-methoxyethane (no. 69; LCI = 2; ARM = 19) does not yield relevant M and is identified via M-H (normalization to M-H on $m/z = 89$), while diethyl peroxide (no. 70; LCI = 2; ARM = 3.8) is identified via a strong M and does not yield significant M-H. Together, they explain the observed data well. We recall that dimethyl peroxide may also be present (see $C_2H_6O_2$ above). Methyl ethyl peroxide is not listed in NIST and hence cannot be clearly identified or ruled out.

– $C_5H_4O_2$ ($m/z = 96$): 3-furaldehyde (no. 71; LCI = 1; ARM = 2.2) is identified in this work via M and M-H signals. The structural isomer furfural exposes a very similar fragmentation pattern and is a possible alternative candidate.

– $C_5H_6O_2$ ($m/z = 98$): We identified 2-furanmethanol (no. 72; LCI = 3; ARM = 18) with high certainty via its strong M and

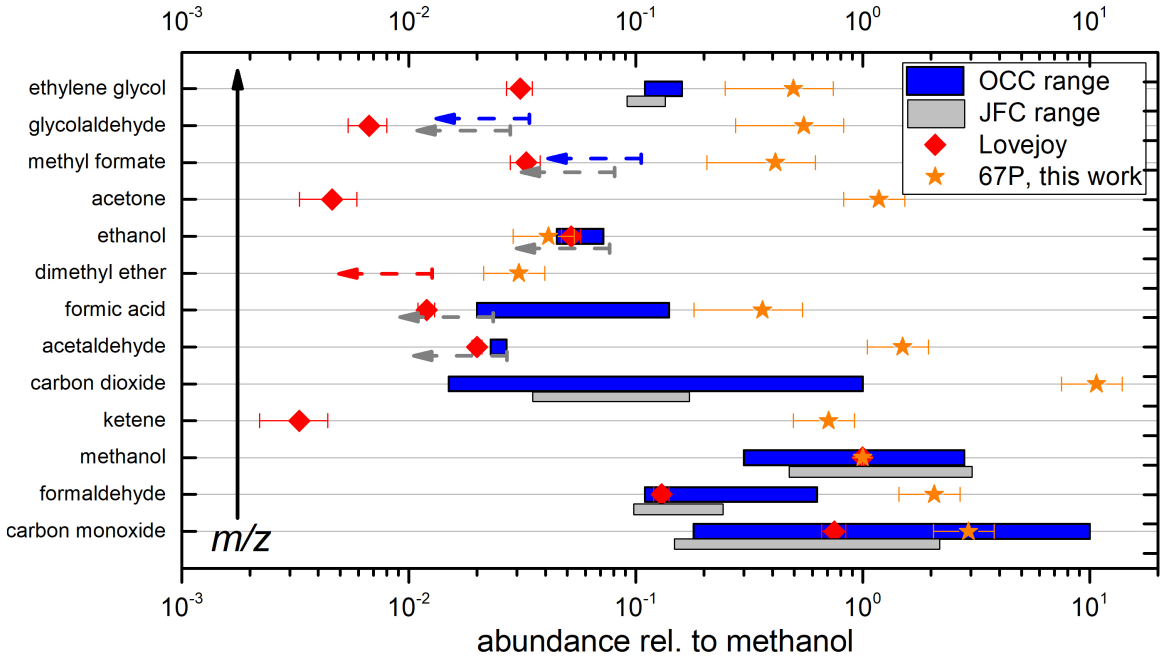

Fig. 4. Comparison to O-bearing organic molecules observed in comets remotely as reviewed by Biver & Bockelée-Morvan (2019). Cometary abundance data from Biver & Bockelée-Morvan (2019) were renormalized from water to methanol. In addition to ranges for both JFCs and OCCs, the values for the long-period OCC C/2014 Q2 (Lovejoy) are also included because glycolaldehyde was detected in this comet for the first time (Biver et al. 2015). The OCC and JFC ranges for CO_2 were taken from Harrington Pinto et al. (2022) and were normalized with the respective average value of methanol from Dello Russo et al. (2016). Dynamically new comets were not considered. Left-pointing arrows indicate upper limits that have been derived from non-detections of the respective species.

M-H signals. However, the position of the hydroxy group could be either two or three and is most likely a mixture of both.

– $C_5H_8O_2$ ($m/z = 100$): Cyclopropanecarboxylic acid methyl ester (no. 73; LCI = 2; ARM = 2.5) is identified via M and M-H. It is the only isomer with fragmentation data in NIST that yields significant M-H. Moreover, cyclopropanecarboxylic acid (no. 64 above) is likely present as well.

– $C_5H_{10}O_2$ ($m/z = 102$): While propanoic acid ethyl ester (no. 74; LCI = 2; ARM = 15) is identified via M, 4-methyl-1,3-dioxane (no. 75; LCI = 2; ARM = 20) yields M and M-CH_3. However, due to an absence of significant M, the latter fragmentation pattern had to be normalized to the M-H signal. The methyl group is more likely to be on position 4 than on position 2 because position 2 would yield too much intensity on M-CH_3, as compared to M-H. However, minor contributions of this isomer or other structural isomers without significant M and M-H signals cannot be ruled out. Pentanoic acid, for instance may or may not be present in very low abundance.

– $C_6H_6O_2$ ($m/z = 110$): Hydroquinone (no. 76; LCI = 2; ARM = 2.6) is identified via a strong M signal, but other isomers with the hydroxy group on different positions of the benzene ring are equally possible (i.e., resorcinol, also known as 1,3-benzenediol, and catechol, also known as 1,2-benzenediol). Less likely is 1-(2-furanyl)-ethanone, but a limited contribution, especially to better explain the $C_5H_3O_2$ fragment on $m/z = 95$, is possible.

– $C_7H_6O_2$ ($m/z = 122$): Benzoic acid (no. 77; LCI = 3; ARM = 25) is identified with high certainty via characteristic M and M-OH signals. This molecule was previously reported in Hänni et al. (2022).

4. Discussion

In the following, we discuss the different O-bearing molecules identified in the framework of the method applied in this work. Specifically, we compare comet 67P's inventory to other reservoirs of extraterrestrial organics, namely, other comets (Biver & Bockelée-Morvan 2019, Fig. 4), the ISM (McGuire 2022; Table 2), and SOM extracted from the Murchison meteorite (Botta & Bada 2002; Sephton 2002; Table 1). Before discussing the different chemical functional groups of O-bearing organic molecules in the reservoirs mentioned above (in order of decreasing priority of the functional group) while highlighting the relevance to biochemistry, prebiotic chemistry, and astrochemistry, we briefly comment on the reservoirs and the characteristics and caveats of the attempted comparison individually.

Figure 4 shows a comparison of our findings to ground-based observations of mostly the rotational spectra of O-bearing complex organics in the two comet populations, Jupiter Family Comets (JFCs; e.g., comet 67P) and Oort Cloud Comets (OCCs; e.g., comet C/2014 Q2 (Lovejoy)), as reviewed by Biver & Bockelée-Morvan (2019) and Harrington Pinto et al. (2022). Apart from water, CO and CO_2 are the major components of cometary ices (Mumma & Charnley 2011; Biver & Bockelée-Morvan 2019), and they have already been studied in detail in comet 67P previously (Läuter et al. 2020; Combi et al. 2020). Even around the comet's perihelion, CO and CO_2 originate mostly from the icy bulk material (Läuter et al. 2019), and hence, are largely unrelated to the desorption of complex O-bearing organics from ejected cometary grains. As cometary reference species, we included their abundances relative to methanol as observed on 3 August 2015 in Fig. 4. While the CO abundance lies just outside the upper end of the respective JFC range, the CO_2 one exceeds it by almost two orders of magnitude. The corresponding CO/CO_2 ratio is 0.27. We find that it is important to note, however, that this is a momentary value, and substantial variation of this ratio has been reported (Hässig et al. 2015). The variability is not only caused by variations in the outgassing itself but also by changes in the relative position of the spacecraft

Table 1. Selected molecules from this work compared to Murchison SOM in terms of (relative) sum abundances of functional groups.

Species	SOM conc. (ppm)	SOM rel. to acids	67P est. abundance	67P rel. to acids
Carboxylic acids	>300	1	142	1
Hydroxycarboxylic acids	15	0.05	xxx[(b)]	xxx[(b)]
Dicarboxylic acids and hydroxydicarboxylic acids	14	0.05	xxx[(b)]	xxx[(b)]
Esters linear	–	–	59	0.42
Esters cyclic/lactones[(c)]	–	–	78	0.55
Aldehydes	27	0.09	516	3.63
Ketones			217	1.52
Alcohols	11	0.04	223	1.57
Polyols	>8	0.03	76	0.53
Ethers linear	–	–	83	0.58
Ethers cyclic[(a)]	–	–	191	1.34
Peroxides	–	–	4	0.03

Notes. [(a)]Murchison SOM has been reviewed in Botta & Bada (2002) and Alexander et al. (2017) and literature cited therein. Here, we consider the values listed in Table 1 in Alexander et al. (2017). [(b)]Expected M signals of hydroxy-, di-, and hydroxydicarboxylic acids seem to be absent or below DFMS detections limits, which is indicated by XXX. [(c)]Cyclic esters and ethers are so-called O-bearing heterocycles.

with respect to the comet and by the comet's rotation. Notably, the methanol/water ratio can vary considerably within just a few hours, too. Except for ethanol, our relative abundance estimates are consistently higher than the other cometary abundances visualized in Fig. 4, sometimes by more than an order of magnitude. In addition, the specific measurement conditions probably also contribute to the observed deviations. In this work, we analyzed data from a very dusty period close to the comet's perihelion and heavy organics are expected to be released mostly from dust grains. During this period, dust was released in short-lived, local dust outbursts (Vincent et al. 2016), which means that the dust density and hence the density of heavy organics released from dust were by no means homogeneous on the sunward side of the nucleus. During perihelion passage, emissions were dominant from the south pole region (e.g., Läuter et al. 2019). Variable illumination conditions are expected to play a subordinate role for the sublimation of organics. Rubin et al. (2019a) have studied the comet's volatile bulk material farther away from the Sun when there was less sublimation from dust. They found, for instance, that formaldehyde is 152% (here: 206%), formic acid is 6.2% (here: 36%), and acetic acid is 1.6% (here: 11%) relative to methanol. Values from Rubin et al. (2019a) are closer to the values reviewed by Biver & Bockelée-Morvan (2019).

Notably, a recent study has provided reanalysis of the data collected by the time-of-flight mass spectrometer COSAC 25 minutes after the initial touchdown of Rosetta's Philae lander unit on 67P's surface. This new study, published by Leseigneur et al. (2022), updates previous results from Goesmann et al. (2015) and explains the observed overall fragmentation pattern ranging from $m/z = 12$ to 63 with a suite of 12 molecules. These molecules have been identified with a high likelihood based on a least-squares fitting algorithm applied to a set of 120 preselected and partially "hand-picked" reference spectra. Half of the molecules are O-bearing, complex organic species, namely, acetaldehyde, acetone, methoxyethane, ethylene glycol, 2-methoxypropane, and cyclopentanol. Acetaldehyde, acetone, and ethylene glycol have been securely identified in this work as well, whereas methoxyethane may be present but cannot be clearly distinguished from isopropanol, which shows a very similar fragmentation pattern, hampering unambiguous identification. The molecule 2-methoxypropane does not yield a significant M signal, and secure mass spectrometry-based detection is difficult. Cyclopentanol does yield a minor amount of M, but tetrahydropyran, with its strong M and M-H signals, must be favored according to our method of analysis. The COSAC mass spectrum evaluated by Goesmann et al. (2015) and Leseigneur et al. (2022) was collected in the so-called sniffing mode, and it was assumed that, during first contact with the cometary surface, some excavated material entered the exhaust region of the instrument and evaporated there at local temperatures of 12 to 15°C. Such conditions would be considerably different from those prevalent during the time when the DFMS data investigated in this work were collected, as ejected dust grains can reach temperatures of multiple hundreds of degrees Celsius in the coma (Lien 1990), and heavier species can sublimate.

Table 1 groups the cometary species selected in this work according to their chemical functionalities in order to allow comparison of the abundances of the groups in the two reservoirs, comet 67P (this work) and Murchison SOM (Alexander et al. 2017, updated from Botta & Bada 2002, updated from Cronin et al. 1988). The Murchison chondrite is one of the most organic-rich and best-characterized meteorites to date. Hänni et al. (2022) found striking similarities between Murchison SOM, as reported by Alexander et al. (2017), and comet 67P's complex organics in terms of average sum formula, which according to these authors, might indicate shared prestellar history. In this work, our candidate molecules have been grouped according to priority of the functional groups in cases where two different functionalities are present within the same molecule. This means that, for example, glycolaldehyde is assigned to the group of aldehydes and not to the group of alcohols despite its hydroxyl function. We expected that the relative sum abundances of the different groups are less error prone than the relative individual abundances, as uncertainties partially cancel each other out. Despite the fact that some species, such as straight chain molecules, are systematically underestimated due to absent or weak M signals, a rough comparison of these reservoirs of complex organics is still meaningful, given that this bias likely affects the different groups of molecules in a similar way.

The analysis of meteoritic SOM also comes with certain biases. First and foremost, small and hence highly volatile molecules have probably mostly been lost or transformed into

more complex molecules during aqueous alteration or metamorphism in the asteroid parent body, which would have occurred long before the meteorite was collected and laboratory analyses were initiated. Consequently, compared to organic species in comets, a shift in the mass-distribution of meteoritic organic molecules toward higher masses is expected. In untargeted high-resolution mass spectrometric investigations of meteoritic SOM, after simple solvent extraction, an extremely high chemical diversity has been demonstrated over a broad mass range (some species had masses greater than 400 Da; see, e.g., Schmitt-Kopplin et al. 2010). However, in targeted analyses, specific classes of compounds with a certain (pre)biotic or biological interest (such as amino acids, sugars, nucleobases, etc.) or molecules for which an established chemical protocol of detection is already known and optimized (carboxylic acids, aldehydes, aromatic compounds, etc.) are usually investigated. This is a further caveat of SOM investigations, as it is obvious that from such studies, SOM cannot be characterized comprehensively and in an unbiased manner. Partly related to this is another bias, namely, almost all studies of SOM follow an extraction protocol that includes an acid hydrolysis step (Botta & Bada 2002) during which esters can be hydrolyzed into an alcohol and a carboxylic acid moiety. Because water as a main solvent is found in large excess, the hydrolysis is total, and no ester molecules remain present. The hydrolysis products of the ester molecules then lead to an overestimation of the carboxylic acid and alcohol budget. Similarly, cyclic esters (lactones) are detected as hydroxycarboxylic acids under these conditions. Unfortunately, this has not been investigated thoroughly for the case of esters and lactones, to the best of our knowledge. However, the bias has been studied and quantified for amino acids (we note that amide functions are hydrolyzed into amines and carboxylic acids) by measuring their abundance in solvent extracts with and without hydrolysis (Martins et al. 2015). Such studies have helped to show that most of those species are not "free" molecules present in the meteorite bulk but that they are fragments of bigger (but still soluble) molecules. In that sense, it is not surprising that molecules with two or even three chemical functions (especially carboxylic acid, amine, and/or alcohol functions) are particularly abundant in SOM (targeted) findings. After an acid hydrolysis extraction, this may support the idea that those species were in fact originally fragments (or monomers) bound together in larger macromolecules (or polymers).

Lastly, Table 2 lists the 27 neutral organic O-bearing molecules composed of four or more atoms (radicals and disputed detections are not considered) that have been identified in the ISM to date. With a small number of exceptions, most of the molecules can be found in McGuire (2022). For each molecule, we indicate whether or not it can be detected in the mass spectrometric dataset analyzed for this work, and if yes, we list the LCI of our identification. However, the comparison between our in situ mass spectrometry-based work and the ISM detections usually performed with rotational-spectroscopic methods from the ground is subject to methodological biases. Despite the fact that both methods target the gas phase and rely on laboratory reference spectra, mass spectrometry does not always allow for distinction between structural isomers, and it needs the analyte species to produce a relevant M signal for a secure detection in a complex analyte while rotational spectroscopy can access only molecules with a permanent dipole moment. Notably, heteroatoms, like the oxygen atom in the focus of this work, usually introduce a dipole moment to the hydrocarbon molecule and thus help to make it accessible to spectroscopic studies from the ground. A more thorough discussion of the two different techniques and their intrinsic limitations can be found in Hänni et al. (2022).

4.1. Carboxylic acids (R–COOH) and carboxylate esters (R–COO–R)

Carboxylic acids are of biological interest, especially as the linear ones among them could form membranes and vesicles via hydrophobic interactions. Many of them have been identified to be part of meteoritic SOM (cf. Botta & Bada 2002 and literature referenced therein and our Table 1). Table 1 shows that Murchison SOM seems to be rich in carboxylic acids compared to comet 67P's volatile organics, where carboxylic acids are clearly not the dominant O-bearing functional group. As discussed previously, the extraction protocol can introduce a severe bias for the quantification of carboxylic acids as well as esters and amides, as the latter two may be hydrolyzed into acids, depending on the extraction conditions. This may lead to an overestimation of the abundance of carboxylic acids in meteoritic SOM. In Murchison, Yuen et al. (1984) reported acetic acid to be the most abundant monocarboxylic acid. From our data, the estimation of the effective total abundances of carboxylic acids is difficult as well because we know that they may be trapped in the form of ammonium salts (Altwegg et al. 2020; Poch et al. 2020; Altwegg et al. 2022). However, in the data from 3 August 2015 analyzed for this work, salts are not considered a major contributor. For the case of acetic acid, the coexistence of the two other isomers, namely, methyl formate and glycolaldehyde, introduces an additional factor of uncertainty. We also found clear indications for butanoic acid, but we cannot confirm pentanoic and higher carboxylic acids due to an absent M signal. In return, the abundance of propanoic acid might be overestimated because the O_2-bearing fragment on $m/z = 73$ ($C_3H_5O_2$), which is also the M-H fragment of propanoic acid, is characteristic to exactly those linear carboxylic acids with at least three carbon atoms that we cannot identify. Notably, the branched species are likely to contribute, but they expose very similar fragmentation patterns in terms of O-retaining fragments and, hence, cannot be distinguished clearly from the unbranched versions. From the ground, only formic acid has been observed in comets (Biver & Bockelée-Morvan 2019). Unlike in Murchison SOM, no signals of species with three or more O atoms were identified in our dataset that could be indicative of hydroxycarboxylic acids or dicarboxylic acids. The fact that no hydroxycarboxylic acids seem to be present in comets might be explained through the aqueous alteration step in meteoritic parent bodies. It is thought that α-hydroxycarboxylic acids are produced during parent body aqueous alteration alongside α-amino acids via Strecker-cyanohydrin synthesis from carbonyl compounds and hydrogen cyanide and in the presence of ammonia (Peltzer et al. 1984). This is a hypothesis, however, that still needs to be supported by systematic SOM analyses of a series of chondrites with different alteration degrees. Given the detection after acid hydrolysis, it is possible that these classes of compounds may, at least partly, be fragments of bigger molecules containing ester and amide moieties. In the ISM, acetic acid is the largest carboxylic acid detected (Mehringer et al. 1997). In our data, the largest and conclusively identified carboxylic acid is benzoic acid, which was already mentioned in Hänni et al. (2022).

In terms of esters, we state the following: Apart from methyl formate, the abundance of which is not well constrained from our data due to structural isomerism with glycolaldehyde and acetic acid, two lactones are likely present and make a relevant contribution, namely, furanone and its hydrogenated version

Table 2. O-bearing neutral and non-radical organic molecules containing four or more atoms identified in the ISM compared to the findings of this work, including the level-of-confidence indicator (LCI) of our identifications.

Molecule	Constitution formula	Mass (Da)	Reference	Presence in 67P (LCI)
Formaldehyde	H_2CO	30	McGuire (2022)	Yes (3)
Methanol	CH_3OH	32	McGuire (2022)	Yes (3)
Ketene	H_2CCO	42	McGuire (2022)	Yes (3)
Acetaldehyde	CH_3CHO	44	McGuire (2022)	Yes (3)
Ethylene oxide	c-C_2H_4O	44	McGuire (2022)	Maybe [a]
Vinyl alcohol	CH_2CHOH	44	McGuire (2022)	[b]
Formic acid	HCOOH	46	McGuire (2022)	Yes (3)
Dimethyl ether	CH_3OCH_3	46	McGuire (2022)	Yes (3)
Ethanol	CH_3CH_2OH	46	McGuire (2022)	Yes (3)
Propynal	HC_2CHO	54	McGuire (2022)	Yes (2)
Cyclopropenone	c-H_2C_3O	54	McGuire (2022)	[b]
Propenal	CH_2CHCHO	56	McGuire (2022)	Yes (2)
1-propen-1-on	CH_3CHCO	56	Fuentetaja et al. (2023)	[b]
Acetone	CH_3COCH_3	58	McGuire (2022)	Yes (3)
Propanal	CH_3CHCH_2O	58	McGuire (2022)	Yes (3)
Propylene oxide	CH_3CH_2OH	58	McGuire (2022)	Maybe [a]
Isopropanol	i-C_3H_7OH	60	Belloche et al. (2022)	Yes (2)
n-Propanol	n-C_3H_7OH	60	Belloche et al. (2022)	Yes (3)
Methyl formate	$HCOOCH_3$	60	McGuire (2022)	Yes (3)
Acetic acid	CH_3COOH	60	McGuire (2022)	Yes (3)
Glycolaldehyde	CH_2OHCHO	60	McGuire (2022)	Yes (3)
Z-1,2-ethenediol	HOCHCHOH	60	Rivilla et al. (2022)	[b]
Ethylene glycol	$HOCH_2CH_2OH$	62	McGuire (2022)	Yes (3)
3-hydroxypropenal [c]	OHCHCHCHO	72	Coutens et al. (2022)	[b]
Ethyl formate	C_2H_5OCHO	74	McGuire (2022)	Maybe [c]
Methyl acetate	CH_3COOCH_3	74	McGuire (2022)	Maybe [d]
Hydroxyacetone	CH_3COCH_2OH	74	McGuire (2022)	Maybe [d]

Notes. Original literature for the ISM species can be found in the 2021 census of McGuire (2022), and only more recent detections are referenced specifically. Disputed detections were not considered. [a]May or may not be present at the expense of the structural isomer(s). [b]No reference data available from NIST. [c]Identification tentative. [d]Molecule may or may not be present with low relative abundance, but the fragmentation pattern does not show any significant M or M-H signals for a clear identification.

γ-butyrolactone. The latter seems to be more abundant than the dehydrogenated furanone. However, relative abundance estimates come with a high uncertainty, and it is possible that they are generally overestimated because, in both cases, structural isomers with similar fragmentation patterns exist and could contribute. From the ground-based observation of other comets, only upper limits for methyl formate were obtained in OCCs other than C/1995 O1 (Hale-Bopp) and C/2014 Q2 (Lovejoy) and in JFCs (Biver & Bockelée-Morvan 2019). In the ISM, both methyl and ethyl formate were detected as well as methyl acetate (McGuire 2022). For the Murchison meteoritic SOM, esters have not been quantified to date to the best of our knowledge. As discussed in detail, a reason for this may be the conversion of esters into carboxylic acids under acid hydrolysis conditions. It might be worth trying to detect and quantify these compounds in chondritic SOM without prior hydrolysis and preferentially by using a solvent extraction protocol that does not involve water but, for instance, methanol.

4.2. Aldehydes (R–CHO) and ketones (R–CO–R)

Aldehydes and ketones are commonly found in meteorites and may also play an important role, for example, in the Strecker synthesis of amino acids (Botta & Bada 2002) and possibly in the abiotic synthesis of sugars. It has been hypothesized that sugars were prebiotically formed from formaldehyde via formose reaction either on the early Earth or even in parent bodies during aqueous alteration (see, e.g., Paschek et al. 2022 and literature referenced therein). The simplest sugar is the aldose glyceraldehyde ($C_3H_6O_3$) and its ketose isomer dihydroxyactone. Although it has been searched for in the ISM, glyceraldehyde has not been found yet (Hollis et al. 2004), and the detection of dihydroxyacetone in the galactic center source Sgr B2 by Widicus Weaver & Blake (2005) has been disputed (Apponi et al. 2006). However, a related molecule, glycolaldehyde ($C_2H_4O_2$), has been detected in the ISM (Hollis et al. 2000) as well as in comets (Biver et al. 2015), and we discuss it in more detail later in this subsection. Only recently, several sugars, including the pentose sugar ribose ($C_5H_{10}O_5$), have been identified in chondrites via gas chromatography coupled with mass spectrometry after acid hydrolysis (Furukawa et al. 2019). In its heterocyclic furanose form, ribose is present in the bio-macromolecules ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). For this reason, we made recurrent efforts toward the identification of O-bearing heterocycles. From a chemical point of view, O-bearing heterocycles are mostly cyclic ethers. We discuss further details in Sect. 4.4.

For Murchison SOM, aldehydes and ketones were found to make up a total of 27 ppm, which is roughly an order of magnitude less than the lower limit reported for carboxylic acids

(Botta & Bada 2002). For aldehydes, all possible structural isomers through C4 seem to be present with decreasing concentrations with increasing number of carbon atoms (cf. Jungclaus et al. 1976). A similar case has been reported ibidem for ketones. In addition, some higher molecular weight hydrocarbons with polar ketone functionalities, for example, ketone derivatives of simple polycyclic aromatic hydrocarbons (PAHs), have been tentatively identified (Krishnamurthy et al. 1992). From our data, we found a total of aldehydes and ketones that clearly outweighs the identified carboxylic acids by about a factor of five (cf. Table 1). However, it seems that carboxylic acids, rather than aldehydes and ketones, tend to not yield significant M signals, which might lead to an overestimation of this factor.

Formaldehyde and acetaldehyde are by far the two most abundant aldehydes identified from our data, followed by propanal and propenal. The relative abundance of glycolaldehyde is poorly constrained due to structural isomerism. In other comets, the following aldehydes have been observed from the ground: formaldehyde, acetaldehyde, and glycolaldehyde (Biver & Bockelée-Morvan 2019). We observed more formaldehyde as compared to methanol, while an inverse ratio was derived from the remote observational data for both OCCs and JFCs (cf. Fig. 4; Biver & Bockelée-Morvan 2019). Whether or not this might indicate a local and/or temporal dust source (Cottin et al. 2004; Cordiner et al. 2014) cannot be determined from this single dataset collected over a time period of just a few hours and at one specific cometocentric distance. Therefore, further investigation is necessary. In the ISM, the aldehydes propynal, propenal, and propanal have been listed as detected by McGuire (2022), in addition to the species that have been observed in comets. Formaldehyde, which seems to be the most abundant aldehyde, both in comets and the ISM, had already been detected in over a dozen interstellar sources in 1969 (Snyder et al. 1969).

Among the ketones, we found that acetone as well as ketene (standard name: ethenone) are very abundant. But there seems to be a small number of even larger ketones, as large as dimethylcyclohexanone. Both acetone and ketene have been identified in comets other than 67P through ground-based observations (Biver & Bockelée-Morvan 2019). In addition, the ketones cyclopropenone and hydroxyacetone have been reported in the ISM (McGuire 2022). While no reference mass spectrum is available from NIST for cyclopropenone, hydroxyacetone does not yield significant M and M-H signals. Consequently, even if present, we cannot clearly identify these ketones from our data.

4.3. Alcohols (R–OH)

In biomolecules, alcohols are mostly found in sugars, but they are also found in some amino acids and other molecules with biological relevance. The alcohols detected to date in the ISM (McGuire 2022) and in comets other than 67P (Biver & Bockelée-Morvan 2019) are methanol, ethanol, and ethanediol. In addition, vinyl alcohol has been reported for the ISM (cf. McGuire 2022 and literature referenced therein). Rivilla et al. (2022) identified Z-1,2-ethenediol, and Belloche et al. (2022) detected *n*-propanol and isopropanol. From our analysis, we found that methanol is abundant, while the higher homologous series alcohols (ethanol, propanol, etc.) are not. This is consistent with reports from analysis of Murchison SOM (Jungclaus et al. 1976), where methanol is present in absolute abundance of 5 µg/g, and abundances gradually decrease toward the butyl alcohols, which are the most complex alcohols found in that study. While isopropanol could be identified clearly, the identification of *n*-propanol was only tentative. If correct, *n*-propanol seemed to be slightly more abundant than isopropanol based on the gas chromatogram of the head space above the liquid extract. Our data indicate isopropanol to be more abundant than *n*-propanol by a factor of about four (subject to substantial uncertainty due to the possible presence of methoxyethane), while a recent study of the galactic center source Sgr B2(N) by Belloche et al. (2022) demonstrated a slight dominance of the *n*-propanol over the iso version. Modeling work accompanying that latter observation showed that the observed ratio might be a direct inheritance of the branching ratio of the OH radical addition to propylene in dust-grain ice mantles (driven by water photodissociation). However, results from modeling the branched carbon chain chemistry in Sgr B2(N) indicate that the degree of branching rises with molecular size (Garrod et al. 2017). Apart from 2-butanol in this study and butyl alcohols in Murchison SOM (Jungclaus et al. 1976), no higher homologous series alcohols can be identified. Admittedly, their mass spectrometry-based identification is hampered by the fact that especially larger and unbranched alcohols do not yield significant M signals (cf. Sect. 2.3). Nevertheless, it is possible to rule out the presence of large amounts of unidentified alcohols in 67P's coma on the basis of the resonance-stabilized CH_3O ion on $m/z = 31$, which is a common fragment of mostly primary alcohols. Our Occam's razor-based solution in Fig. 2 does explain the corresponding measured intensity well, and additional contributions are limited by the estimated 30% uncertainty of the signal. Beyond the homologous series alcohols, phenol and benzylacohol are also likely present in 67P's coma. We note that the latter has been reported already (Hänni et al. 2022). The abundances of the even larger alcohols (2-methylcyclohexanol, 2-methoxyethanol, and 2-furanmethanol) may be overestimated, as it is unlikely that among the many structural isomers of those molecules, only one exists. A possible reason could be that our Occam's razor approach is less appropriate for more complex molecules where it is more likely that multiple low-abundant isomers contribute to the observed sum fragmentation spectrum (cf. Methods section). Both in Murchison SOM as well as in our data, the sum abundance of aldehydes plus ketones outweighs that of alcohols. From Table 1, the abundance ratio of alcohols with respect to aldehydes plus ketones is roughly 0.7 for Murchison, while it seems to be lower for our data, about 0.4. However, if the underestimation of aliphatic species (see Methods section) is different for the different functional groups, the factor we estimate might be influenced. If, for example, aliphatic alcohols tend to be underestimated more than aliphatic ketones and aldehydes, then the effective ratio for comet 67P would be slightly underestimated and thus closer to the Murchison SOM value. Comet C/2014 Q2 (Lovejoy) shows a ratio of roughly 6.5 based on Fig. 4 and the data reviewed in Biver & Bockelée-Morvan (2019). However, in this comet, methanol is very abundant, almost a factor of 20 more abundant than ethanol, which is probably the main reason for the deviating ratio.

Interestingly, diols tend to be more abundant than the respective mono-alcohol. This becomes obvious from the reports of Biver & Bockelée-Morvan (2019) for ethanol versus ethylene glycol (standard name: 1,2-ethanediol) in both JFCs and OCCs (cf. Fig. 4). Comet C/2014 Q2 (Lovejoy), where ethanol is almost twice as abundant as ethylene glycol, seems to be a bit of an exception. From our data, we find an order of magnitude more of the diol, too. The same applies to the next higher homologous series members, as propanediol is more abundant than *n*-propanol by almost a factor of five. If the branched and unbranched isomer of the mono-alcohol are summed up, the same factor is only three. Methanediol (CH_4O_2) is not observed

in our data, that is, there is no signal at $m/z = 48$ where M would be expected. However, NIST does not list fragmentation data for this molecule, which might point toward increased instability. Indeed, it is known that geminal diols are prone to decay into water and the respective ketone or aldehyde, but the formation of methanediol from energetically processed methanol-oxygen ices and its gas phase stability have been recently demonstrated (Zhu 2022). The larger diols, namely, ethanediol and propanediol, are observed in the sterically more favorable vicinal 1,2-form. The isolated 1,3-propanediol loses water and produces strong signals on $m/z = 58$ and 57 but no relevant M signal. However, the corresponding measured intensity is already well explained by the presence of acetone and propanal, and hence possible contributions of the isolated diol are marginal. The largest diol that is very likely present in comet 67P's coma is the benzene derivative hydroquinone, and the structural isomers where the hydroxy functions are located at different positions on the benzene ring.

4.4. Ethers (R–O–R)

Due to their importance to biomolecules, Barnum et al. (2022) recently performed an extensive search of cyclic ethers and other heterocycles toward the Taurus Molecular Cloud (TMC-1). They used the GOTHAM (GBT Observations of TMC-1: Hunting Aromatic Molecules) collaboration survey and reported upper limits for three O-bearing heterocycles, namely, ethylene oxide (C_2H_4O), furan (C_4H_4O), and benzofuran (C_8H_6O). Ethylene oxide, also known as oxirane, was first identified in Sgr B2 by Dickens et al. (1997). Since then, it has been observed in several other interstellar environments. Also, its methylated form, propylene oxide – also known as 2-methyloxirane (c-$CH_3C_2H_2O$) – was detected toward Sgr B2 (McGuire et al. 2016). As summarized in Barnum et al. (2022), furan has been searched for in several sources previously, but only upper limits have been reported to date. The most constraining upper limit was derived from the GOTHAM line survey by Barnum et al. (2021) and corresponds to 1×10^{12} cm^{-2}. Barnum et al. (2022) interpreted their null detections of the searched heterocycles as pointing toward a unique chemistry of these species compared to the pure carbocycles and hypothesized that their absence in TMC-1 may be a matter of reaction dynamics, given their high terrestrial stability.

A very different picture presents itself in this work. For the first time, we can confirm the presence of significant amounts of extraterrestrial O-bearing heterocycles. We find the five-membered furan-derivatives to be especially abundant. Interestingly, dehydrogenated species seem to be less abundant than fully hydrogenated ones. Six-membered pyran-based species as well as six-membered rings with two O atoms (dioxanes) are very likely present but much less common. Out of the 19 identified ethers, 12 are heterocyclic molecules, and accordingly, the relative sum abundance of the cyclic ethers outweighs that of the linear ones by almost a factor of three, see Table 1. This imbalance is expected to be partially due to the methodological bias that chains tend to yield less stable M ions than rings and are hence more difficult to identify. However, our data leave limited room for the presence of additional species, given the 30% and 50% error margins of the observed intensities, respectively. There are additional heterocyclic molecules likely present in comet 67P that cannot be classified as ethers: two lactones (see above under carboxylic acids and carboxylate esters) and furanmethanol (see above under Alcohols). Interestingly, the functionalized and hydrogenated heterocycles seem to be more abundant than the simple dehydrogenated ones, such as furan and pyran. Our findings from in situ comet observations thus confirm the suggestion of Miksch et al. (2021) that the hydrogenated versions of heterocycles might be promising candidates for future interstellar searches. Based on theoretical work, these authors report that under interstellar conditions, many heterocycles should be slowly reduced by abundant hydrogen atoms. In terms of O-bearing heterocycles, especially 2,3-dihydrofuran and 2,5-dihydrofuran should be abundant, likely more abundant than furan itself. This is consistent with our findings, which further implies that the fully hydrogenated species, THF, is even more abundant than 2,3-dihydrofuran and 2,5-dihydrofuran. A similar situation was found for the two lactone heterocycles (see above under Sect. 4.1), where the fully hydrogenated γ-butyrolactone is more abundant than the dehydrogenated furanone. Also, tetrahydropyran was found to be abundant relative to methanol but only about half as abundant as THF. Unfortunately, missing reference data prevented the clear identification of pyran itself, and no direct comparison of the hydrogenated and the non-hydrogenated isomer was possible. Despite the fact that methylated or sometimes ethylated molecules seem to play a significant role, we stress that the locations of the alkyl groups are generally not well constrained by mass spectrometry, and the simultaneous presence of multiple isomers seems likely. This is true not only for the case of heterocycles but also for the other chemical groups of molecules discussed above. Moreover, we compared the relative abundance of heterocyclic species with their pure hydrocarbon counterparts. While Barnum et al. (2022) found significant depletion of heterocycles relative to the pure carbocycles in TMC-1, we cannot confirm this depletion based on our data from comet 67P. In TMC-1, benzofuran is at least ten times less abundant than indene. However, from the data analyzed for this work, both these species are similarly abundant in comet 67P. We found the same situation for furan and cyclopentadiene. The abundance of all four molecules is slightly above 1% relative to methanol.

To the best of our knowledge, ethers have not yet been studied in detail in the Murchison meteorite nor in any other chondrite.

4.5. Peroxides (R–OO–R)

From our data, diethyl peroxide seems to be present in 67P's coma, while dimethyl peroxide may or may not be present. In addition, inorganic hydrogen peroxide was observed based on DFMS data collected when the comet was still far from its perihelion (Bieler et al. 2015). Overall, the chemical class of peroxide does not seem to be abundant among complex organic molecules, which may be due to the low stability of the peroxide bond. To date, peroxides have not been identified in the ISM, in other comets, or in Murchison SOM, as far as we know. The low stability of the O–O bond possibly impairs the detection in meteoritic SOM extracts.

5. Summary and conclusions

In this work, we revisited 67P's inventory of O-bearing complex organic molecules previously studied by Schuhmann et al. (2019b), applying an Occam's razor-based approach to achieve deconvolution of the complex sum fragmentation pattern observed on 3 August 2015 by Rosetta's high-resolution mass spectrometer DFMS. Despite methodological limitations, such as the incompleteness of the NIST mass spectrometry database used as reference in this work and ambiguities related to structural isomerism, we derived a minimum of 63 $C_nH_mO_x$ candidate

molecules, roughly quintupling the number of species reported in Schuhmann et al. (2019b). However, additional contributions from low-abundance complex organics evading mass spectrometric detection are likely. From a thorough comparison of our findings to the well-studied Murchison SOM (Botta & Bada 2002; Schmitt-Kopplin et al. 2010; Alexander et al. 2017) as well as to the ISM (McGuire 2022) and other comets (Biver & Bockelée-Morvan 2019), we are able to conclude the following:

- Comet 67P's O-bearing complex organics expose diverse chemical functionalities, from carboxylic acids and carboxylate esters via alcohols to aldehydes to ketones to ethers. Deviations from the Murchison meteoritic inventory (Botta & Bada 2002; Schmitt-Kopplin et al. 2010; Alexander et al. 2017) can be understood when taking into account that the methodological bias resulting from the usually acidic meteorite sample work-up processes (Martins et al. 2015) mainly affects the quantification of carboxylic acids, esters, and alcohols. In addition, parent body (aqueous) alteration is thought to lead to aggregation of smaller ("more pristine") molecules or monomers into larger ("less pristine") (macro)molecular structures (e.g., Alexander et al. 2017), possibly explaining a depletion of small species in meteorites as compared to comets. For instance, we cannot confirm any signal to indicate the presence of molecules bearing three or more oxygen atoms (e.g., dicabroxylic acids, hydroxycarboxylic acids, or sugars) in comet 67P. Those species, which seem abundant in meteoritic SOM (Alexander et al. 2017), are hypothesized to form during aqueous alteration in parent bodies;
- Compared to other comets reviewed in Biver & Bockelée-Morvan (2019), $C_nH_mO_x$ species in comet 67P tend to be more abundant relative to methanol. However, this could be the effect of data selection (very dusty time period) and local enhancement due to dust sublimating near DFMS' ionization chamber. However, there are also parallels with the findings of previous studies: Our findings that non-geminal diols are more abundant than the corresponding mono-alcohols, for instance, are consistent with ground-based observations of other comets where ethanediol is consistently more abundant than ethanol (Biver & Bockelée-Morvan 2019). Despite the fact that the production of methanediol from energetically processed methanol-oxygen ices as well as gas phase stability has been demonstrated (Zhu 2022), and just like other comet observations (Biver & Bockelée-Morvan 2019), we cannot confirm its presence in 67P's coma. The reason may be this molecule's short lifetime and its proneness to undergo dehydration to form formaldehyde;
- We find significantly more hydrogenated than dehydrogenated heterocycles, which is consistent with theoretical work on heterocycles of Miksch et al. (2021) and has also been reported for the case of carbocycles in 67P's coma by Hänni et al. (2022). Our findings may have implications for future radiofrequency or infrared spectroscopic observational campaigns from both ground-based (e.g., with the Green Bank Telescope) and space-based facilities (e.g., with the *James Webb* Space Telescope). For example, the millimeter wave spectrum of furan in the frequency range of 75–295 GHz was investigated by Barnum et al. (2021) and used to derive an upper limit in the cold core TMC-1. If the ISM ratio of furan to its dehydrogenated counterpart THF is comparable to the one we derive in this work for comet 67P, a search for THF might be more promising. The rotational spectrum of THF was studied in Meyer et al. (1999) and Mamleev et al. (2001), and the pseudorotational band $n = 0 \rightarrow n = 2$ was investigated in the 170–360 GHz range by Engerholm et al. (1969) and Melnik et al. (2003). However, the rotational spectrum of THF is complex, and the identification of this molecule in the ISM may be complex as well. In favor of the radio line searches, THF does have a considerably larger dipole moment than its dehydrogenated counterpart;
- For the two cases where heterocycles (this work) are directly comparable to carbocycles (Hänni et al. 2022), we find that both variants of molecules exhibit similar relative abundances. In contrast, Barnum et al. (2022) have reported at least an order of magnitude difference between indene and its heterocyclic pendent benzofuran, the latter being the less abundant one. Overall, $C_nH_mO_x$ species are clearly less abundant than the pure hydrocarbons also in comet 67P's coma (Hänni et al. 2022);
- From our work, cyclic species seem to be the more suitable candidates for explaining the data than acyclic ones, which is in contrast to ISM studies, where clearly prolate molecules are dominantly found (McGuire 2022; McCarthy & McGuire 2021). Notably, while there is sometimes no fragmentation data on NIST for the acyclic species (e.g., polyynes) or the M signal tends to be either weak or absent (e.g., carbon chain species), their functionalized – commonly with a cyano group – variants possess a pronounced permanent dipole moment and are thus easy to target in space via their rotational spectra. This was discussed in more detail in Hänni et al. (2022). As far as we know, for meteoritic SOM, the overall ratio of cyclic to acyclic species has not been studied in detail;
- We also studied the abundance of branched versus unbranched alkyl chain species. In contrast to analysis of Murchison SOM (Jungclaus et al. 1976), we clearly identify the primary alcohol *n*-propanol, most likely together with its secondary counterpart isopropanol. The estimated abundance ratio of these two species is roughly four, suggesting that secondary alcohol is the dominant molecule. However, a possible contribution of methoxyethane could lower this estimated factor. Belloche et al. (2022) report a slight dominance of *n*-propanol over the iso version, while simulations of branched carbon chain chemistry in Sgr B2(N) seem to indicate that the degree of branching rises with molecular size (Garrod et al. 2017). To understand how this ratio might be variable for different classes of molecules in comet 67P, further investigations both in the laboratory and in space are required. In the future, having a full organic inventory accessible by Rosetta's high-resolution mass spectrometer DFMS on hand for comet 67P would allow for more detailed investigations (e.g., of branched versus straight chain species) and enable more detailed comparisons to modeling work done for the ISM (cf. Garrod et al. 2017).

It is still unclear whether the complex organics targeted in this study form from small precursors via bottom-up pathways or rather via top-down mechanisms from larger decomposing PAHs or macromolecular structures in carbonaceous dust grains, as discussed by Lee et al. (2019) or Burkhardt et al. (2021). Laboratory work suggests that PAHs embedded in astrophysical ice analogs are not only hydrogenated but also oxygenated when subjected to Ly-α radiation (Gudipati & Yang 2012). According to Bernstein et al. (1999), UV irradiation of PAHs in simulated interstellar ices also seems to provide probable formation pathways toward ethers as well as alcohols and quinones. However, our study of pristine cometary matter can only add to the

ongoing debate in combination with laboratory and modeling work. A combined effort is needed to tackle the formation of the complex organics in the ISM, including the ratios of hydrogenated versus dehydrogenated, straight versus branched, and cyclic versus acyclic species. While we are planning to further explore comet 67P's chemical complexity by extending our work to species containing heteroelements other than oxygen, we strongly encourage others to also pursue such efforts.

Acknowledgements. We gratefully acknowledge the work of the many engineers, technicians and scientists involved in the Rosetta mission and in the planning, construction, and operation of the ROSINA instrument in particular. Without their contributions, ROSINA would not have produced such outstanding data and scientific results. Rosetta is an ESA mission with contributions from its member states and NASA. Work at the University of Bern was funded by the Canton of Bern and the Swiss National Science Foundation (200020 182418). S.F.W. acknowledges financial support of the SNSF Eccellenza Professorial Fellowship PCEFP2_181150. M.R.C. acknowledges support from NASA grants 80NSSC18K1280 and 80NSSC20K0651. J.D.K. acknowledges support from the Belgian Science Policy Office. Especially, we thank Prof. Dr. B. McGuire for discussing our work with us from an ISM perspective.

References

A'Hearn, M. F., Feaga, L. M., Keller, H. U., et al. 2012, ApJ, 758, 29
Alexander, C. M. O. D., Cody, G. D., De Gregorio, B. T., Nittler, L. R., & Stroud, R. M. 2017, Chem. Erde/Geochemistry, 77, 227
Altwegg, K., Balsiger, H., Berthelier, J. J., et al. 2017, MNRAS, 469, S130
Altwegg, K., Balsiger, H., Hänni, N., et al. 2020, Nat. Astron., 4, 533
Altwegg, K., Combi, M., Fuselier, S. A., et al. 2022, MNRAS, 516, 3900
Apponi, A. J., Halfen, D. T., Ziurys, L. M., et al. 2006, ApJ, 643, L29
Balsiger, H., Altwegg, K., Bochsler, P., et al. 2007, Space Sci. Rev., 128, 745
Barnum, T. J., Lee, K. L. K., & McGuire, B. A. 2021, ACS Earth Space Chem., 5, 2986
Barnum, T. J., Siebert, M. A., Lee, K. L. K., et al. 2022, J. Phys. Chem. A, 126, 2716
Belloche, A., Garrod, R. T., Zingsheim, O., Müller, H. S. P., & Menten, K. M. 2022, A&A, 662, A110
Bernstein, M. P., Sandford, S. A., Allamandola, L. J., et al. 1999, Science, 283, 1135
Bieler, A., Altwegg, K., Balsiger, H., et al. 2015, Nature, 526, 678
Biver, N., & Bockelée-Morvan, D. 2019, ACS Earth Space Chem., 3, 1550
Biver, N., Bockelée-Morvan, D., Moreno, R., et al. 2015, Sci. Adv., 1, 1500863
Biver, N., Bockelée-Morvan, D., Hofstadter, M., et al. 2019, A&A, 630, A19
Botta, O., & Bada, J. L. 2002, Surv. Geophys., 23, 411
Burkhardt, A. M., Long Kelvin Lee, K., Bryan Changala, P., et al. 2021, ApJ, 913, L18
Chyba, C., & Sagan, C. 1992, Nature, 355, 125
Clauberg, J. 1654, Logica Vetus Et Nova (Kessinger Publishing)
Combi, M., Shou, Y., Fougere, N., et al. 2020, Icarus, 335, 113421
Cordiner, M. A., Remijan, A. J., Boissier, J., et al. 2014, ApJ, 792, L2
Cottin, H., Bénilan, Y., Gazeau, M.-C., & Raulin, F. 2004, Icarus, 167, 397
Coutens, A., Loison, J. C., Boulanger, A., et al. 2022, A&A, 660, L6
Cronin, J. R., Pizzarello, S., & Cruikshank, D. P. 1988, in Meteorites and the Early Solar System, eds. J. F. Kerridge, & M. S. Matthews (Springer), 819
Dello Russo, N., Kawakita, H., Vervack, R. J., & Weaver, H. A. 2016, Icarus, 278, 301
de Sanctis, M. C., Raponi, A., Ammannito, E., et al. 2016, Nature, 536, 54
Dickens, J. E., Irvine, W. M., Ohishi, M., et al. 1997, ApJ, 489, 753
Engerholm, G. G., Luntz, A. C., Gwinn, W. D., & Harris, D. O. 1969, J. Chem. Phys., 50, 2446
Fink, U., Doose, L., Rinaldi, G., et al. 2016, Icarus, 277, 78
Fray, N., Bardyn, A., Cottin, H., et al. 2016, Nature, 538, 72
Fuentetaja, R., Bermúdez, C., Cabezas, C., et al. 2023, A&A, 671, A6
Furukawa, Y., Chikaraishi, Y., Ohkouchi, N., et al. 2019, PNAS, 116, 24440
Garrod, R. T., Belloche, A., Müller, H. S. P., & Menten, K. M. 2017, A&A, 601, A48
Goesmann, F., Rosenbauer, H., Bredehöft, J. H., et al. 2015, Science, 349, 2.689
Gudipati, M. S., & Yang, R. 2012, ApJ, 756, L24
Hänni, N., Altwegg, K., Combi, M., et al. 2022, Nat. Commun., 13, 3639
Harrington Pinto, O., Womack, M., Fernandez, Y., & Bauer, J. 2022, Planet. Sci. J, 3, 247
Hässig, M., Altwegg, K., Balsiger, H., et al. 2015, Science, 347, aaa0276
Hollis, J. M., Lovas, F. J., & Jewell, P. R. 2000, ApJ, 540, L107
Hollis, J. M., Jewell, P. R., Lovas, F. J., Remijan, A., & Møllendal, H. 2004, ApJ, 610, L21
Jenniskens, P., Wilson, M. A., Packan, D., et al. 2000, Earth, Moon and Planets, 82, 57
Jungclaus, G. A., Yuen, G. U., Moore, C. B., & Lawless, J. G. 1976, Meteoritics, 11, 231
Krishnamurthy, R. V., Epstein, S., Cronin, J. R., Pizzarello, S., & Yuen, G. U. 1992, Geochim. Cosmochim. Acta, 56, 4045
Kwok, S. 2016, A&ARv, 24, 8
Läuter, M., Kramer, T., Rubin, M., & Altwegg, K. 2019, MNRAS, 483, 852
Läuter, M., Kramer, T., Rubin, M., & Altwegg, K. 2020, MNRAS, 498, 3995
Lee, K. L. K., McGuire, B. A., & McCarthy, M. C. 2019, Phys. Chem. Chem. Phys. (Incorp. Faraday Trans.), 21, 2946
Le Roy, L., Altwegg, K., Balsiger, H., et al. 2015, A&A, 583, A1
Leseigneur, G., Bredehöft, J. H., Gautier, T., et al. 2022, Angew. Chem. Int. Ed., 61, e202201925
Levasseur-Regourd, A.-C., Agarwal, J., Cottin, H., et al. 2018, Space Sci. Rev., 214, 64
Lien, D. J. 1990, ApJ, 355, 680
Lodders, K. 2021, Space Sci. Rev., 217, 44
Mamleev, A. H., Gunderova, L. N., & Galeev, R. V. 2001, J. Struct. Chem., 42, 365
Martins, Z., Modica, P., Zanda, B., & D'Hendecourt, L. L. S. 2015, Meteor. Planet. Sci., 50, 926
Marty, B., Altwegg, K., Balsiger, H., et al. 2017, Science, 356, 1069
McCarthy, M. C., & McGuire, B. A. 2021, J. Phys. Chem. A, 125, 3231
McGuire, B. A. 2022, ApJS, 259, 30
McGuire, B. A., Carroll, P. B., Loomis, R. A., et al. 2016, Science, 352, 1449
Mehringer, D. M., Snyder, L. E., Miao, Y., & Lovas, F. J. 1997, ApJ, 480, L71
Melnik, D. G., Gopalakrishnan, S., Miller, T. A., & De Lucia, F. C. 2003, J. Chem. Phys., 118, 3589
Meyer, R., López, J. C., Alonso, J. L., et al. 1999, J. Chem. Phys., 111, 7871
Migliorini, A., Piccioni, G., Capaccioni, F., et al. 2016, A&A, 589, A45
Miksch, A. M., Riffelt, A., Oliveira, R., Kästner, J., & Molpeceres, G. 2021, MNRAS, 505, 3157
Mumma, M. J., & Charnley, S. B. 2011, ARA&A, 49, 471
Naraoka, H., Takano, Y., Dworkin, J. P., et al. 2023, Science, 379, eabn9033
Paschek, K., Kohler, K., Pearce, B. K. D., et al. 2022, Life, 12
Peltzer, E. T., Bada, J. L., Schlesinger, G., & Miller, S. L. 1984, Adv. Space Res., 4, 69
Poch, O., Istiqomah, I., Quirico, E., et al. 2020, Science, 367, aaw7462
Rivilla, V. M., Colzi, L., Jiménez-Serra, I., et al. 2022, ApJ, 929, L11
Rubin, M., Altwegg, K., Balsiger, H., et al. 2019a, MNRAS, 489, 594
Rubin, M., Bekaert, D. V., Broadley, M. W., Drozdovskaya, M. N., & Wampfler, S. F. 2019b, ACS Earth Space Chem., 3, 1792
Schmitt-Kopplin, P., Gabelica, Z., Gougeon, R. D., et al. 2010, PNAS, 107, 2763
Schuhmann, M., Altwegg, K., Balsiger, H., et al. 2019a, A&A, 630, A31
Schuhmann, M., Altwegg, K., Balsiger, H., et al. 2019b, ACS Earth Space Chem., 3, 1854
Sephton, M. A. 2002, Natural Product Rep., 19, 292
Snodgrass, C., A'Hearn, M. F., Aceituno, F., et al. 2017, Philos. Trans. R. Soc. London Ser. A, 375, 20160249
Snyder, L. E., Buhl, D., Zuckerman, B., & Palmer, P. 1969, Phys. Rev. Lett., 22, 679
Steins, S. E. 2018, Chemistry WebBook, NIST Standard Reference Database Number, 69
Tosi, F., Capaccioni, F., Capria, M. T., et al. 2019, Nat. Astron., 3, 649
Vincent, J. B., A'Hearn, M. F., Lin, Z. Y., et al. 2016, MNRAS, 462, S184
Wang, J., Marks, J. H., Turner, A. M., et al. 2023, Sci. Adv., 9, eadg1134
Widicus Weaver, S. L., & Blake, G. A. 2005, ApJ, 624, L33
Yabuta, H., Cody, G. D., Engrand, C., et al. 2023, Science, 379, eabn9057
Yeo, A. N. H., & Williams, D. H. 1969, Org. Mass Spectrom., 2, 331
Yuen, G., Blair, N., Des Marais, D. J., & Chang, S. 1984, Nature, 307, 252
Zhu, C. 2022, PNAS, 119, 11938

Appendix A: Overview of candidate molecules

Table A.1. List of O-bearing neutral and non-radical organic molecules[(a)] identified as best candidates to explain the 3 August 2015 overall DFMS mass spectrum.

Molecule (identifier)	Sum formula	Mass (Da)	Level-of-confidence indicator	Abundance rel. to methanol (%)
Cometary O-bearing reference molecules				
Carbon monoxide (no. 1)	CO	28	3	292
Carbon dioxide (no. 51)	CO_2	44	3	1071
Carboxylic acids (R–COOH) and carboxylate esters (R–COO–R)				
Formic acid (no. 52)	CH_2O_2	46	3	36
Acetic acid (no. 54)	$C_2H_4O_2$	60	3	11
Methyl formate (no. 56)	$C_2H_4O_2$	60	3	44
2-Propenoic acid (no. 58)	$C_3H_4O_2$	72	3	6.3
Propanoic acid (no. 59)	$C_3H_6O_2$	74	3	39
2(3H)-Furanone (no. 63)	$C_4H_4O_2$	84	2	12
Cyclopropanecarboxylic acid (no. 64)	$C_4H_6O_2$	86	1	12
γ-Butyrolactone (no. 65)	$C_4H_6O_2$	86	1	66
Butanoic acid (no. 68)	$C_4H_8O_2$	88	2	13
Cyclopropanecarboxylic acid methyl ester (no. 73)	$C_5H_8O_2$	100	2	2.5
Propanoic acid ethyl ester (no. 74)	$C_5H_{10}O_2$	102	2	15
Benzoic acid (no. 77)	$C_7H_6O_2$	122	3	25
Aldehydes (R–CHO) and ketones (R–CO–R)				
Formaldehyde (no. 2)	CH_2O	30	3	206
Ketene (no. 4)	C_2H_2O	42	3	71
Acetaldehyde (no. 5)	C_2H_4O	44	3	150
2-Propynal (no. 8)	C_3H_2O	54	2	1.7
2-Propenal (no. 9)	C_3H_4O	56	2	29
Propanal (no. 11)	C_3H_6O	58	3	48
Acetone (no. 10)	C_3H_6O	58	3	118
Glyoxal (no. 53)	$C_2H_2O_2$	58	2	2.5
Glycolaldehyde (no. 55)	$C_2H_4O_2$	60	3	44
Butanal (no. 16)	C_4H_8O	72	3	12
2,4-Cyclopentadiene-1-one (no. 19)	C_5H_4O	80	2	?[(b)]
3-Furaldehyde (no. 71)	$C_5H_4O_2$	96	1	2.2
Benzaldehyde (no. 28)	C_7H_6O	106	3	4.1
4-Methylbenzaldehyde (no. 33)	C_8H_8O	120	2	4.3
6-methyl-3,5-heptadien-2-one (no. 35)	$C_8H_{12}O$	124	1	8.3
2,6-Dimethylcyclohexanone (no. 36)	$C_8H_{14}O$	126	1	19
Alcohols (R–OH)				
Methanol (no. 3)	CH_4O	32	3	100[(c)]
Ethanol (no. 7)	C_2H_6O	46	3	4.1
Isopropanol (no. 12)	C_3H_8O	60	2	19
n-Propanol (no. 13)	C_3H_8O	60	3	5.0
Ethylene glycol (no. 57)	$C_2H_6O_2$	62	3	50
2-Butanol (no. 18)	$C_4H_{10}O$	74	2	8.3
1,2-Propanediol (no. 62)	$C_3H_8O_2$	76	2	24
2-Methoxyethanol (no. 60)	$C_3H_8O_2$	76	3	27
Phenol (no. 25)	C_6H_6O	94	2	7.7
2-Furanmethanol (no. 72)	$C_5H_6O_2$	98	3	18
Benzyl alcohol (no. 29)	C_7H_8O	108	1	4.1
Hydroquinone (no. 76)	$C_6H_6O_2$	110	2	2.6
2-Methylcyclohexanol (no. 31)	$C_7H_{14}O$	114	1	30
Ethers (R–O–R)				
Dimethyl ether (no. 6)	C_2H_6O	46	3	3.1
Furan (no. 14)	C_4H_4O	68	2	1.7
2,3-Dihydrofuran (no. 15)	C_4H_6O	70	2	8.3
Tetrahydrofuran (no. 17)	C_4H_8O	72	3	64
Methylal (no. 61)	$C_3H_8O_2$	76	2	4.0
2-Methylfuran (no. 20)	C_5H_6O	82	3	11

2,3-Dihydro-4-methylfuran (no. 21)	C_5H_8O	84	3	20
Tetrahydropyran (no. 22)	$C_5H_{10}O$	86	3	25
1-Ethoxypropane (no. 23)	$C_5H_{12}O$	88	3	3.9
2-Methoxy-2-methylpropane (no. 24)	$C_5H_{12}O$	88	3	0.2
1,3-Dioxane (no. 66)	$C_4H_8O_2$	88	3	6.2
1,4-Dioxane (no. 67)	$C_4H_8O_2$	88	3	13
1-Ethoxy-1-methoxyethane (no. 69)	$C_4H_{10}O_2$	90	2	19
2,5-Dimethylfuran (no. 26)	C_6H_8O	96	3	11
3-Methoxycyclopentene (no. 27)	$C_6H_{10}O$	98	3	33
4-Methyl-1,3-dioxane (no. 75)	$C_5H_{10}O_2$	102	2	20
2,3,5-Trimethylfuran (no. 30)	$C_7H_{10}O$	110	2	9.3
Benzofuran (no. 32)	C_8H_6O	118	3	1.9
Ethoxybenzene (no. 34)	$C_8H_{10}O$	122	1	19
Peroxides (R–OO–R)				
Diethyl peroxide (no. 70)	$C_4H_{10}O_2$	90	2	3.8

Notes. [(a)] Molecules are subdivided into groups according to their chemical functionalities, as described in Section 4, starting always with the molecule with the smallest molecular weight. We always exclude CO and CO_2 because they are not organic and do not belong in any of the groups. [(b)] No NIST mass spectrum available for any of the structural isomers of with the sum formula C_5H_4O. The proposed molecule seems to be a plausible candidate, but no ARM can be estimated due to the missing fragmentation information. [(c)] Definition. Abundance estimates have been normalized relative to methanol.