GLOVOCS - Master compound assignment guide for Proton Transfer Reaction Mass Spectrometry users

A.M. Yañez-Serrano[1,2], I. Filella[1,2], J. LLusia[1,2], A. Gargallo-Garriga[1,2,3], V. Granda[1], E. Bourtsoukidis[4], J. Williams[4], R. Seco[5,6], L. Cappellin[7], C. Werner[8], J. de Gouw[9], J. Peñuelas[1,2].

[1]{CREAF, E08193 Bellaterra (Cerdanyola del Vallès), Catalonia, Spain}

[2]{CSIC, Global Ecology Unit, CREAF-CSIC-UAB, E08193 Bellaterra (Cerdanyola del Vallès), Catalonia, Spain}

[3]{Global Change Research Institute, The Czech Academy of Sciences, Belidla 986/4a, CZ-60300 Brno, Czech Republic}

[4]{Atmospheric Chemistry Department, Max Planck Institute for Chemistry, P. O. Box 3060, 55128, Mainz, Germany}

[5]{Terrestrial Ecology Section, Department of Biology, University of Copenhagen, Copenhagen, Denmark}

[6]{Center for Permafrost, Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark}

[7]{Dipartimento di Scienze Chimiche, Università degli Studi di Padova, Padova, Italy and Research and Innovation Centre, Fondazione Edmund Mach, S. Michele a/A 38010, Italy}

[8]{Ecosystem Physiology, University of Freiburg, 79098 Freiburg, Germany}

[9]{Cooperative Institute for Research in Environmental Sciences & Department of Chemistry, University of Colorado, 216 UCB, CIRES, Boulder, Colorado 80309, USA}

Corresponding author: ayanezserrano@yahoo.es

1. Highlights

- We have created, GLOVOCS, a master compound assignment guide that can be referred to by PTR-MS practitioners.
- GLOVOCS is aimed to help in advancing science of VOCs by facilitating the research of multiple groups using PTR-MS.
- GLOVOCS, accessible at http://glovocs.creaf.cat, is a collaborative tool where users can both consult and contribute to.

2. Abstract

The richness of measurements obtained by Proton-Transfer Reactions Mass Spectrometry (PTR-MS) has opened a new paradigm for the quantification of volatile organic compounds (VOCs). A wide range of compounds can be monitored, however, each detected signal is subject to a compound assignment instead of actual identification because PTR techniques are mass-selective and isomers cannot be separately measured. Thus, rapid development in the field requests continued community efforts to identify compounds. In this study we have reviewed the available literature and created a master compound assignment guide called GLOVOCS that can be referred to by PTR-MS practitioners. GLOVOCS is aimed to help in advancing science of VOCs by facilitating the research of multiple groups using PTR-MS to monitor VOCs and to disentangle the physical, chemical and biological mechanisms underlying their production, emission and impact on environment and organisms from bacteria to humans. The guide is freely accessible at http://glovocs.creaf.cat as a collaborative tool, where users can both consult and contribute to the identification of VOCs by providing possible candidates for all chemical formulas from 18 to 330 atomic mass units. When available, we indicate if there is evidence for biogenic or anthropogenic VOC origin, as well as grouping the compounds based on the Classyfire chemotaxonomic classification (Djoumbou Feunang et al., 2016). While GLOVOCS aims to facilitate the first assessment and consistent classification of compounds, we highly recommend further cross-validation for verifying compounds when using PTR-MS techniques.

Keywords: VOC, PTR-MS, PTR-TOF-MS, quadrupole, time of flight, compound assignment

3. Introduction

46   The study of volatile organic compounds (VOCs) has increased considerably over the past two decades,
47   expanding to many disciplines such as atmospheric science (Karl et al., 2018), plant science (Fasbender
48   et al., 2018), ecology (Achotegui-Castells et al., 2016), microbiology (J. Peñuelas et al., 2014; Josep
49   Peñuelas et al., 2014), global change science (Peñuelas and Staudt, 2010), medical science (Beauchamp
50   et al., 2013), food science (Zardin et al., 2014), and even human physiological responses to emotions
51   (Stönner et al., 2018; Williams et al., 2016). Studies of VOCs are challenging, because these compounds
52   are commonly present at very low concentrations, and their measurement and identification are
53   complicated. Gas chromatographic techniques can accurately identify compounds but they have the
54   disadvantage of requiring sample pre-concentration and generally time-consuming offline analysis.

55   Analytical techniques such as proton transfer reaction mass spectrometry (PTR-MS; Lindinger et al.
56   (1998), Ionicon Analytik, Austria) and selected-ion flow-tube mass spectrometry (SIFT-MS; Smith and
57   Španěl, (2005), Syft Technologies, New Zealand) were developed and became commercially available
58   two decades ago. While first versions were based on quadrupole mass analysers, more sophisticated
59   versions appeared later, such as proton transfer reaction time-of-flight mass spectrometry (PTR-ToF-
60   MS; Jordan et al. (2009), Ionicon Analytik, Austria), VOCUS PTR-ToF-MS (Krechmer et al. (2018),
61   TOFWERK AG, Switzerland) and PTR3-ToF-MS (Breitenlechner et al., 2017). These instruments
62   perform rapid and highly sensitive measurements using soft chemical ionisation to simultaneously
63   monitor a range of VOCs in air. This low-energy ionisation results in relatively low levels of molecular
64   fragmentation. The technique is based on the protonation of VOC molecules by hydronium ions (H_3O^+)
65   formed at high density in an ion source with a hollow cathode. VOCs in an air sample are then
66   protonated in reactions with the hydronium ions in a drift tube under the influence of an electric field,
67   producing water clusters as a by-product. This process, however, may differ amongst instruments
68   (Holzinger et al., 2019). The protonated ions are mass separated and detected, which can also vary
69   amongst instruments. The quadrupole and ToF are the most common mass spectrometers used in PTR-
70   MS instruments, differing mainly in mass resolving power and the number of ions that can be measured
71   simultaneously with high time resolution. The resolution of PTR-MS systems is given by the ratio of
72   the mass-to-charge ratio (m/z) to its full width at half maximum (FWHM) (Stark et al., 2015). The
73   quadrupole mass spectrometer has a resolution of 300 m/Δm (i.e. can separate one full mass at a time),
74   while the ToF-MS can currently resolve up to 14000 m/Δm, depending on the instrument (i.e. resolving
75   masses up to the fourth decimal digit).

76   These analytical techniques measure a wide range of compounds, but with some limitations. Only
77   compounds with proton affinities higher than water can be properly protonated. However, this is an
78   advantage for measuring air samples: the sampled air acts as the carrier gas of the sample, because its
79   major components (nitrogen, oxygen, argon, carbon dioxide and methane) have a lower proton affinity
80   than water. Fragments or water clusters can be formed on multiple occasions, adding complexity to data
81   interpretation. The most important limitation of these techniques is nonetheless their mass selectivity,
82   since compounds with the same mass contribute to the same spectral peak. This limitation inhibits the
83   identification of compounds. Many compounds, like monoterpenes, often have the same molecular
84   formula, including isomers and enantiomers. Assigning compounds identities thus becomes difficult,
85   with the need to cross-validate the techniques, such as Gas Chromatography Mass Spectrometry.
86   Additionally, it is worth mentioning that PTR-MS technology is capable of using other ionization modes
87   with primary ions, such as NO^+ or O_2^+, that can resolve some mass identities (Liu et al., 2013), but we
88   restrict this dataset to H_3O^+.

89   The purpose of this study is to construct a guide, named GLOVOCS, to provide a collaborative tool in
90   the form of a master list of assigned compounds for users of PTR-MS systems to aid with mass
91   assignment and therefore data interpretation. We additionally provide classification of the compounds
92   based on the Classyfire database in order to standardise VOC groups. This database provides chemical
93   taxonomy up to 11 levels, such as Kingdom, SuperClass, Class and SubClass, by using a computable

94   structure-based chemical taxonomy (ChemOnt), and the computer program (ClassyFire) that uses only
95   chemical structures and structural features (Djoumbou Feunang et al., 2016).

4. Data sources

97   We reviewed all available PTR-MS compound lists to date and merged several published tables for
98   mass assignment that used PTR techniques. Each available compound and molecular formula was
99   cross-checked against the NIST Webbook (https://webbook.nist.gov/chemistry/). Molecular formulas
100 or compounds not available on this platform were verified using either PubChem
101 (https://pubchem.ncbi.nlm.nih.gov/) and/or ChemSpider (http://www.chemspider.com/).

102 The guide provides a list of protonated ion masses (M.H+ in the guide) related to protonation of H3O+
103 primary ions, along with neutral formulas. For each formula, we provide options from single to multiple
104 assignments for compound candidates, including typically used synonyms. For example, acetone
105 (C_3H_6O, M.H+ 59.05) is also known as 2-propanone, which is stated in the "synonyms" column.

106 We additionally classify each ion based on their nature of being a parent ion (protonated compound
107 without further), a fragment ion (signals from compound fragment from fragmented compounds upon
108 protonation) or a water cluster (compounds that have bonded with a water molecule upon protonation).
109 For example, total monoterpenes can be monitored both at the parent mass (m/z 137.13) and the
110 fragment mass (m/z 81.07) (Tani, 2013); or alcohols such as the diterpenoid cis-abienol ($C_{20}H_{34}O$) tend
111 to dehydrate when protonated, so they are monitored as $M.H^+$ - H_2O. Thus, cis-abienol major derived
112 signal is detected at $(C_{20}H_{32})H^+$ (Jud et al., 2016). We also included some masses normally monitored
113 as instrument performance parameters, such as m/z 29.99, which measures nitrogen oxide impurities in
114 the system, and is typically below 0.05% of the primary ions, or m/z 32.99, which measures the oxygen
115 impurities and is typically below 5% of the primary ion.

116 Furthermore, the diverse classifications of VOCs provided by different authors is a recurrent problem
117 in the literature. We include International Chemical Identifier key (InChlKey) for each VOC and
118 provide a classification based on the Classyfire chemotaxonomic library (Djoumbou Feunang et al.,
119 2016) to standardise VOC grouping. This information, however, is only an extension to our mass
120 identification guide. By providing bibliographic records of biogenic and anthropogenic measurements
121 of such compounds, when available, we aim to further assist the user in the mass identification.

5. GLOVOCS structure

123 The complete master compound assignment guide is available at http://glovocs.creaf.cat. GLOVOCS
124 contains >2000 entries up to 329.84 Da, the parent ion for diiodobenzene, a compound not typically
125 found in nature that is currently used in some PTR-MS instruments as an internal mass standard for
126 PTR-ToF-MS (Yáñez-Serrano et al., 2018).

127 GLOVOCS includes information of recent efforts, such as the recent mass list by Pagonis et al., (2019)
128 summarizing publications that made a conclusive connection between a specific VOC and a detected
129 ion. Moreover, GLOVOCS provides a wide range of proven and possible compound assignments and
130 includes considerably more compounds. As such, it can form the basis for the interpretation of mass
131 spectra and for the construction of a mass list for peak integration in high-resolution mass spectrometry.
132 When available, the proposed compound identification is referenced for both PTR-MS and PTR-TOF-
133 MS techniques by a hyperlink with the DOI information. Additional references can be found appended
134 to GLOVOCS. There are cases where literature is not available, such as for compounds that have not
135 been reported previously. In such cases, identification must performed with care, using the guide as a
136 merely orientative approach. In addition, as GLOVOCS targets instrument users rather than developers,
137 we provide further information about the possible origin (i.e. biogenic or anthropogenic) of the
138 compound. This guide provides at least one possible contributor for each compound formula but does
139 not contain all possible available compounds (Figure 1).

140 The key feature of GLOVOCS is its collaborative aim, where users can contribute and expand it at
141 http://glovocs.creaf.cat. On this website, the guide itself, information about GLOVOCS and a contact
142 form can be found. In particular, the contact form is for users to add or modify the entries in the guide.
143 This addition will be done by the webpage administrator after careful bibliographic verification,
144 creating a version based name of the guide for each update.

145 While consulting GLOVOCS, the user should consider possible heavier isotopes, fragments or water
146 clusters and other possible abundant compounds. For example, monitoring toluene and monoterpenes
147 simultaneously with PTR-ToF-MS is analytically challenging. For example, p-cymene ($C_{10}H_{14}$) is an
148 aromatic monoterpene that fragments to $C_7H_9^+$, coinciding with the toluene signal. Compound cross-
149 validation with other available techniques such as Gas Chromatography Mass Spectrometry will thus
150 remain important to confirm that the mass monitored with PTR-MS corresponds to the reported
151 compound.

152 The volatility of a compound is another factor to consider when using this guide. The common physical
153 state of a compound may not be an indicator of its detectability using PTR techniques. For instance, the
154 common state of the diterpene kaurene ($C_{20}H_{32}$) is solid, but gaseous emissions have been measured
155 (Yáñez-Serrano et al., 2018). In this way, new discoveries of compound emissions can be made.

156

157 Figure 1: Example of the guide. The information here shown is provided for all masses in the range
158 from 18 to 330 Da. The blue rows represent water clusters, the purple row represent fragments and the
159 green rows represent masses normally monitored as instrument performance parameters.

160 GLOVOCS also counts with a chemical taxonomy classification of the VOCs reported in order to
161 standardise VOC groups. It classifies compounds by using a computable structure-based chemical
162 taxonomy (ChemOnt), and the computer program (ClassyFire) that uses only chemical structures and
163 structural features (Djoumbou Feunang et al., 2016). Figure 2 shows the different contributions to the
164 Superclass level (i.e. second level) for all the different compounds reported in the guide, including
165 compounds with the same formula. The most abundant groups are the benzenoid and organic
166 heterocyclic compounds (19 % each), followed by the organic oxygen compounds (16%).
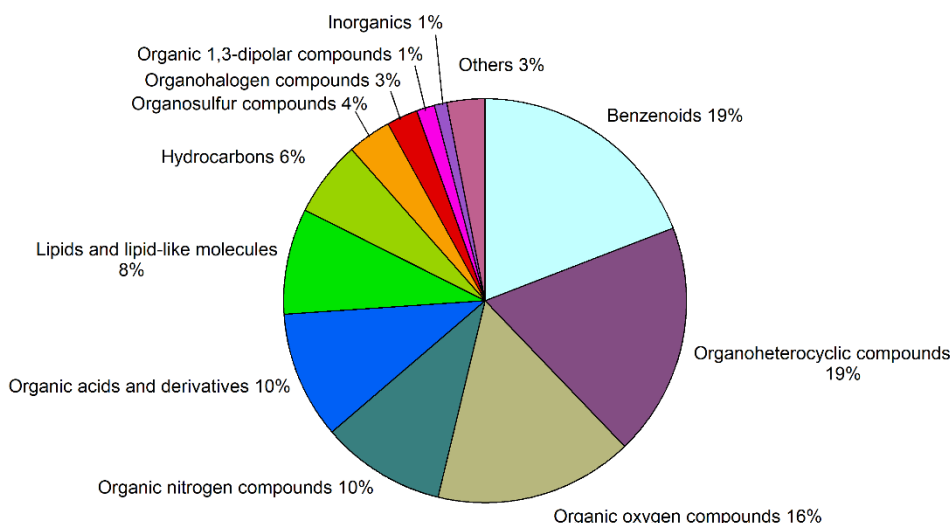
Figure 2: Pie chart of the Superclass level compound types from Classifyre available in the Master compound assignment guide for PTR users. The groups found are: benzenoids compounds, organic heterocyclic compounds, organic oxygen compounds, organic nitrogen compounds, organic acid compounds, lipid compounds, hydrocarbon compounds, organic sulfur compounds, organic halogen compounds, organic 1,3-dipolar compounds, inorganics and others.

6. Conclusions

GLOVOCS is a master compound assignment guide for PTR-MS users that aims to assist the analysis and interpretation of PTR-MS spectra. Cross-validation is highly desirable when performing PTR-MS analysis. We established an open, freely accessible database for compound assignment, where compounds can be added or corrected after verification, as a community collaborative tool. A continually growing, well-referenced guide for compound assignment can thus be a useful tool, with worldwide environmental applications in an open-access format.

7. Acknowledgments

8. Supplementary information

The guide, together with its bibliography and related comments can be found in the supplementary information. The updated guide can also be accessed at http://glovocs.creaf.cat.

9. Additional Information

Competing Interests: The authors declare no competing interests.

10. References

Achotegui-Castells, A., Della Rocca, G., Llusià, J., Danti, R., Barberini, S., Bouneb, M., Simoni, S., Michelozzi, M., Peñuelas, J., 2016. Terpene arms race in the Seiridium cardinale - Cupressus sempervirens pathosystem. Sci. Rep. 6. https://doi.org/10.1038/srep18954

196   Beauchamp, J., Herbig, J., Dunkl, J., Singer, W., Hansel, A., 2013. On the performance of proton-
197   transfer-reaction mass spectrometry for breath-relevant gas matrices. Meas. Sci. Technol. 24,
198   125003. https://doi.org/10.1088/0957-0233/24/12/125003

199   Breitenlechner, M., Fischer, L., Hainer, M., Heinritzi, M., Curtius, J., Hansel, A., 2017. PTR3: An
200   Instrument for Studying the Lifecycle of Reactive Organic Carbon in the Atmosphere. Anal.
201   Chem. 89, 5824–5831. https://doi.org/10.1021/acs.analchem.6b05110

202   Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., Fahy, E.,
203   Steinbeck, C., Subramanian, S., Bolton, E., Greiner, R., Wishart, D.S., 2016. ClassyFire:
204   automated chemical classification with a comprehensive, computable taxonomy. J. Cheminform.
205   8, 61. https://doi.org/10.1186/s13321-016-0174-y

206   Fasbender, L., Yáñez-Serrano, A.M., Kreuzwieser, J., Dubbert, D., Werner, C., 2018. Real-time
207   carbon allocation into biogenic volatile organic compounds (BVOCs) and respiratory carbon
208   dioxide (CO2) traced by PTR-TOF-MS, 13CO2 laser spectroscopy and 13C-pyruvate labelling.
209   PLoS One 13, e0204398. https://doi.org/10.1371/journal.pone.0204398

210   Holzinger, R., Acton, W.J.F., Bloss, W.J., Breitenlechner, M., Crilley, L.R., Dusanter, S., Gonin, M.,
211   Gros, V., Keutsch, F.N., Kiendler-Scharr, A., Kramer, L.J., Krechmer, J.E., Languille, B.,
212   Locoge, N., Lopez-Hilfiker, F., Materić, D., Moreno, S., Nemitz, E., Quéléver, L.L.J., Sarda
213   Esteve, R., Sauvage, S., Schallhart, S., Sommariva, R., Tillmann, R., Wedel, S., Worton, D.R.,
214   Xu, K., Zaytsev, A., 2019. Validity and limitations of simple reaction kinetics to calculate
215   concentrations of organic compounds from ion counts in PTR-MS. Atmos. Meas. Tech. 12,
216   6193–6208. https://doi.org/10.5194/amt-12-6193-2019

217   Jordan, A., Haidacher, S., Hanel, G., Hartungen, E., Märk, L., Seehauser, H., Schottkowsky, R.,
218   Sulzer, P., Märk, T.D., 2009. A high resolution and high sensitivity proton-transfer-reaction
219   time-of-flight mass spectrometer (PTR-TOF-MS), International Journal of Mass Spectrometry.
220   https://doi.org/10.1016/j.ijms.2009.07.005

221   Jud, W., Fischer, L., Canaval, E., Wohlfahrt, G., Tissier, A., Hansel, A., 2016. Plant surface reactions:
222   an opportunistic ozone defence mechanism impacting atmospheric chemistry. Atmos. Chem.
223   Phys. 16, 277–292. https://doi.org/10.5194/acp-16-277-2016

224   Karl, T., Striednig, M., Graus, M., Hammerle, A., Wohlfahrt, G., 2018. Urban flux measurements
225   reveal a large pool of oxygenated volatile organic compound emissions. Proc. Natl. Acad. Sci.
226   U. S. A. 115, 1186–1191. https://doi.org/10.1073/pnas.1714715115

227   Krechmer, J., Lopez-Hilfiker, F.D., Koss, A.R., Hutterli, M.A., Stoermer, C., Deming, B.L., Kimmel,
228   J.R., Warneke, C., Holzinger, R., Jayne, J.T., Worsnop, D.R., Fuhrer, K., Gonin, M., de Gouw,
229   J., 2018. Evaluation of a New Vocus Reagent-Ion Source and Focusing Ion-Molecule Reactor
230   for use in Proton-Transfer-Reaction Mass Spectrometry.
231   https://doi.org/10.26434/CHEMRXIV.6502652.V1

232   Lindinger, W., Hansel, A., Jordan, A., 1998. On-line monitoring of volatile organic compounds at
233   pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) medical
234   applications, food control and environmental research. Int. J. Mass Spectrom. Ion Process. 173,
235   191–241. https://doi.org/10.1016/S0168-1176(97)00281-4

236   Liu, Y.J., Herdlinger-Blatt, I., McKinney, K. a., Martin, S.T., 2013. Production of methyl vinyl ketone
237   and methacrolein via the hydroperoxyl pathway of isoprene oxidation. Atmos. Chem. Phys. 13,
238   5715–5730. https://doi.org/10.5194/acp-13-5715-2013

239   Pagonis, D., Sekimoto, K., de Gouw, J., 2019. A Library of Proton-Transfer Reactions of H3O+ Ions
240   Used for Trace Gas Detection. J. Am. Soc. Mass Spectrom. 30, 1330–1335.
241   https://doi.org/10.1007/s13361-019-02209-3

242   Peñuelas, J., Asensio, D., Tholl, D., Wenke, K., Rosenkranz, M., Piechulla, B., Schnitzler, J.P.P.,

243   2014. Biogenic volatile emissions from the soil. Plant. Cell Environ. 37, 1866–91.
244   https://doi.org/10.1111/pce.12340

245   Peñuelas, Josep, Farré-Armengol, G., Llusia, J., Gargallo-Garriga, A., Rico, L., Sardans, J., Terradas,
246   J., Filella, I., 2014. Removal of floral microbiota reduces floral terpene emissions. Sci. Rep. 4,
247   6727. https://doi.org/10.1038/srep06727

248   Peñuelas, J., Staudt, M., 2010. BVOCs and global change. Trends Plant Sci. 15, 133–144.
249   https://doi.org/10.1016/j.tplants.2009.12.005

250   Smith, D., Španěl, P., 2005. Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace
251   gas analysis. Mass Spectrom. Rev. 24, 661–700. https://doi.org/10.1002/mas.20033

252   Stark, H., Yatavelli, R.L.N., Thompson, S.L., Kimmel, J.R., Cubison, M.J., Chhabra, P.S.,
253   Canagaratna, M.R., Jayne, J.T., Worsnop, D.R., Jimenez, J.L., 2015. Methods to extract
254   molecular and bulk chemical information from series of complex mass spectra with limited mass
255   resolution. Int. J. Mass Spectrom. 389, 26–38. https://doi.org/10.1016/j.ijms.2015.08.011

256   Stönner, C., Edtbauer, A., Derstroff, B., Bourtsoukidis, E., Klüpfel, T., Wicker, J., Williams, J., 2018.
257   Proof of concept study: Testing human volatile organic compounds as tools for age classification
258   of films. PLoS One 13, e0203044. https://doi.org/10.1371/journal.pone.0203044

259   Tani, A., 2013. Fragmentation and Reaction Rate Constants of Terpenoids Determined by Proton
260   Transfer Reaction-mass Spectrometry. Environ. Control Biol. 51, 23–29.
261   https://doi.org/10.2525/ecb.51.23

262   Williams, J., Stönner, C., Wicker, J., Krauter, N., Derstroff, B., Bourtsoukidis, E., Klüpfel, T.,
263   Kramer, S., 2016. Cinema audiences reproducibly vary the chemical composition of air during
264   films, by broadcasting scene specific emissions on breath. Sci. Rep. 6.
265   https://doi.org/10.1038/srep25464

266   Yáñez-Serrano, A.M., Fasbender, L., Kreuzwieser, J., Dubbert, D., Haberstroh, S., Lobo-do-Vale, R.,
267   Caldeira, M.C., Werner, C., 2018. Volatile diterpene emission by two Mediterranean Cistaceae
268   shrubs. Sci. Rep. 8, 6855. https://doi.org/10.1038/s41598-018-25056-w

269   Zardin, E., Tyapkova, O., Buettner, A., Beauchamp, J., 2014. Performance assessment of proton-
270   transfer-reaction time-of-flight mass spectrometry (PTR-TOF-MS) for analysis of isobaric
271   compounds in food-flavour applications. LWT - Food Sci. Technol. 56, 153–160.
272   https://doi.org/10.1016/J.LWT.2013.10.041

273