

## Recommendations for standardizing nomenclature for dietary (poly)phenol catabolites

Colin D Kay,<sup>1</sup> Michael N Clifford,<sup>2,3</sup> Pedro Mena,<sup>4</sup> Gordon J McDougall,<sup>5</sup> Cristina Andres-Lacueva,<sup>6,7</sup> Aedin Cassidy,<sup>8</sup> Daniele Del Rio,<sup>4</sup> Nikolai Kuhnert,<sup>9</sup> Claudine Manach,<sup>10</sup> Gema Pereira-Caro,<sup>11</sup> Ana Rodriguez-Mateos,<sup>12</sup> Augustin Scalbert,<sup>13</sup> Francisco Tomás-Barberán,<sup>14</sup> Gary Williamson,<sup>3</sup> David S Wishart,<sup>15</sup> and Alan Crozier<sup>16,17</sup>

<sup>1</sup>Plants for Human Health Institute, Food Bioprocessing and Nutrition Sciences Department, North Carolina State University, Kannapolis, NC, USA; <sup>2</sup>School of Bioscience and Medicine, University of Surrey, Guildford, United Kingdom; <sup>3</sup>Department of Nutrition, Dietetics and Food, Monash University, Notting Hill, Victoria, Australia; <sup>4</sup>Department of Food and Drugs, University of Parma, Parma, Italy; <sup>5</sup>The James Hutton Research Institute, Invergowrie, Dundee, United Kingdom; <sup>6</sup>Department of Nutrition, Food Science and Gastronomy, University of Barcelona, Barcelona, Spain; <sup>7</sup>CIBER Frailty and Healthy Aging (CIBERfes), Institute of Health Carlos III, Barcelona, Spain; <sup>8</sup>Institute for Food Security, Queen's University, Belfast, United Kingdom; <sup>9</sup>Department of Life Sciences and Health, Jacobs University, Bremen, Germany; <sup>10</sup>Université Clermont Auvergne, INRAE, Clermont-Ferrand, France; <sup>11</sup>Department of Food Science and Health, Andalusian Institute of Agricultural and Fisheries Research and Training, Cordoba, Spain; <sup>12</sup>Department of Nutritional Sciences, Kings College, London, United Kingdom; <sup>13</sup>International Agency for Research on Cancer, WHO, Lyon, France; <sup>14</sup>Food and Health Laboratory, CEBAS-CSIC, Espinardo University Campus, Murcia, Spain; <sup>15</sup>Department of Biological Sciences, University of Alberta, Edmonton, Canada; <sup>16</sup>School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, United Kingdom; and <sup>17</sup>Department of Nutrition, University of California, Davis, Davis, CA, USA

### ABSTRACT

There is a lack of focus on the protective health effects of phytochemicals in dietary guidelines. Although a number of chemical libraries and databases contain dietary phytochemicals belonging to the plant metabolome, they are not entirely relevant to human health because many constituents are extensively metabolized within the body following ingestion. This is especially apparent for the highly abundant dietary (poly)phenols, for which the situation is compounded by confusion regarding their bioavailability and metabolism, partially because of the variety of nomenclatures and trivial names used to describe compounds arising from microbial catabolism in the gastrointestinal tract. This confusion, which is perpetuated in online chemical/metabolite databases, will hinder future discovery of bioactivities and affect the establishment of future dietary guidelines if steps are not taken to overcome these issues. In order to resolve this situation, a nomenclature system for phenolic catabolites and their human phase II metabolites is proposed in this article and the basis of its format outlined. Previous names used in the literature are cited along with the recommended nomenclature, International Union of Pure and Applied Chemistry terminology, and, where appropriate, Chemical Abstracts Service numbers, InChIKey, and accurate mass. *Am J Clin Nutr* 2020;112:1051–1068.

**Keywords:** dietary (poly)phenols, microbial catabolites, phase II metabolites, nomenclature, food/diet metabolome

### Introduction

The complexity of foods and the accompanying limited information on dietary phytochemicals in metabolomics and dietary

guidelines has recently been recognized (1). Although a number of databases contain information on phytochemicals of the plant metabolome, they are not directly relevant to human health or metabolomics because many of these compounds undergo substantial metabolism within the body following ingestion. Chemical/metabolite databases such as Phenol-Explorer, FooDB, ChEBI, and PubChem, metabolomics data repositories including HMDB and MetaboLights, and mass spectra databases such as Mass Bank, ReSpect for Phytochemicals, Golm, and Metlin contain fragmentary information that is far from a comprehensive overview of human metabolites resulting from consumption of phytochemicals. Plant-derived dietary compounds have a much greater metabolomic complexity in vivo than is currently recognized and, as a consequence, the metabolomes derived from plant foods are undercharacterized. This is evident with

CDK was supported by the USDA National Institute of Food and Agriculture (Hatch/Kay-Colin; 1011757). This work was partially supported by the JPI-HDHL FoodPhyt Project (CM by ANR-19-HDH2-0002-01 and PM and DDR by MIPAAF ID 1160) and partially by EU H2020 Project “PREVENTOMICS” (grant 818318) to PM and DDR. CA-L was supported by ISCIII (AC19/00096) and CIBERFES, co-funded by ERDF, “A way to make Europe,” ICREA Academia award 2018, and AGAUR (2017SGR1546).

Supplemental Table 1 is available from the “Supplementary data” link in the online posting of the article and from the same link in the online table of contents at <https://academic.oup.com/ajcn/>.

Address correspondence to A Crozier (e-mail: [alan.crozier44@gmail.com](mailto:alan.crozier44@gmail.com)).

Received May 5, 2020. Accepted for publication June 29, 2020.

First published online September 16, 2020; doi: <https://doi.org/10.1093/ajcn/nqaa204>.

the highly abundant dietary (poly)phenols, which are among the phytochemicals that are recognized as important regulators of human health (2, 3). To further complicate matters, there are areas of confusion regarding the absorption and metabolism of dietary (poly)phenols, partially because of the variety of nomenclatures and trivial names used to describe compounds that arise in the gastrointestinal tract as the result of microbial catabolism and human phase II metabolism. This confusion is perpetuated in online chemical and metabolomic databases.

One of the current goals of (poly)phenol research is to establish dietary guidance including a recommended intake or number of servings of (poly)phenol-rich foods to promote optimal health or reduce disease risk (4, 5). The aim of the proposed nomenclature standardization is to advance basic and applied (poly)phenol research, and associated methods such as metabolome pathway analysis tools (6–8), to provide evidence in support of observational studies (epidemiological evidence) that form the foundation for Dietary Guidelines (National Nutrition Monitoring and Related Research Act). Databases and pathway tools/systems operate using compound recognition, which requires standardized nomenclature and machine-readable identifiers such as InChIKey, and Chemical Abstracts Service (CAS) numbers. It is therefore paramount for (poly)phenol researchers to standardize nomenclature and reporting practices. In addition, there is a requirement for a reliable system of translation, permitting incorporation of pre-existing data described using older or nonstandard terminology into modern databases and bioinformatics tools such as PubChem and HMDB. This need to adopt universally accepted annotations, classification schemes and ontologies, facilitating data interoperability in metabolomics, has recently been emphasized (9, 10).

Calls for standardizing reporting practices in (poly)phenol research are not novel. In 2015 a special article was published in *The American Journal of Clinical Nutrition*, in which recommendations were proposed on reporting requirements for flavonoids in research (11). This report focused primarily on standardization of reporting relevant to dietary assessment, food composition characterization, food administration, human and animal intervention study design, and analytical reporting. Subsequently, a joint International Union of Pure and Applied Chemistry (IUPAC)/International Union of Biochemistry and Molecular Biology working group published recommendations for general nomenclature of (poly)phenols and provided examples of acceptable trivial names, together with semisystematic and fully systematic names (12). However, both these publications focused on phytochemical nomenclature used to characterize secondary plant metabolites in foods and did not cover catabolites arising from the gut microbiome. The current special article provides standardized nomenclature and reporting practices for human metabolites of dietary (poly)phenols encompassing gut microbial catabolites and their human phase II metabolites, which were overlooked in the previous reports. Our aim is to eliminate inconsistencies in nomenclature/terminology, which create confusion and ultimately affects interpretation of data surrounding the biological effects of plant-based foods in epidemiology and nutrition intervention studies.

## Standardization

A call for standardized nomenclature and reporting practices is timely because it will undoubtedly help researchers establish the untapped potential of the non-nutrient phytochemicals in food and diet to modulate human health. Phytochemical structural data need to be captured in ways that have future utility, and this will become of increasing importance because national funding bodies are beginning to require studies make their source data publicly available using open access data repositories (8, 13). Progress is currently impeded by confusion in the nutrition and food science literature because of the previously mentioned inconsistent chemical nomenclatures applied to phenolic catabolites, which make it very difficult to integrate data for review and meta-analysis. The magnitude of this problem is clearly illustrated by the results of Web of Science Core database searches (February 13, 2020; <https://clarivate.com/webofsciencegroup/solutions/web-of-science>) using known variations in nomenclature for a common (poly)phenol catabolite as the search “topic.” This revealed that even differences in hyphenation have significant impact on how many references may be identified when searching for a single compound. A search using our recommended name “3-(3'-hydroxyphenyl)propanoic acid” (Table 1) yielded fifty four references, but its associated synonyms yielded additional references as follows: 3-hydroxyphenylpropionic acid (thirty references), 3-hydroxy-phenylpropionic acid (one reference), 3-hydroxy-phenyl-propionic acid (two references), and 3-(3'-hydroxy-phenyl)-propionic acid (three references). Similarly, a further search using the recommended name “3-hydroxy-3-(3'-hydroxyphenyl)propanoic acid” (Table 1) produced only one reference, whereas its synonyms 3-hydroxy-(3'-hydroxyphenyl)propionic acid yielded eight references, 3-hydroxyphenyl-hydracrylic acid three references, and 3-hydroxyphenylhydracrylic acid one reference. The nomenclature system for phenolic catabolites provided herein will help clarify confusion in the nutrition literature and provide a harmonized approach for online databases.

Biological activity, or its absence, is a function of the 3D structure of a molecule. The goal of a nomenclature system is to describe this 3D structure unambiguously allowing translation of text to a 2D illustrated structure, in a manner as accurate as the available experimental data allow. It should be easy to use and understand, and accommodate multiple forms of isomerism, to permit distinction of isomeric structures with ease. There is, however, no perfect system capable of achieving this objective in a convenient manner that also accommodates all compounds in all situations. Accordingly, this article focuses on the commonly reported isomeric forms of (poly)phenol catabolites and their metabolites, recommending a system of nomenclature and also providing a thesaurus to facilitate translation between different systems of nomenclature. This is important because with many compounds, five and as many as ten synonyms and styles of nomenclature, some incorrect, appear in the nutrition literature (Table 1). In the proposed nomenclature, IUPAC-based names are used, but in instances in which IUPAC terminology is inconsistent (e.g., phenylpropanoic and phenylpropionic acids) or is overly complex and/or where alternative names have become accepted

TABLE 1 Examples of recommended nomenclature for (poly)phenols, their microbial catabolites, and human phase II conjugates<sup>1</sup>

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey, <sup>6</sup> monoisotopic mass g/mol)
1. Benzene diols and triols (C <sub>6</sub> -C <sub>0</sub> )			
Benzene-1,2-diol	1,2-Dihydroxybenzene Catechol Pyrocatechol Benzocatechin 1,3-Dihydroxybenzene Resorcinol	120-80-9	Benzene-1,2-diol (YCIMNLLNPGFGHC-UHFFFAOYSA-N; 110.036779 g/mol)
Benzene-1,3-diol	1,2,3-Trihydroxybenzene Pyrogallol	108-46-3	Benzene-1,3-diol (GHMLBKR AJCXXBS-UHFFFAOYSA-N; 110.036779 g/mol)
Benzene-1,2,3-triol	1,3,5-Trihydroxybenzene Phloroglucinol	87-66-1	Benzene-1,2,3-triol (WQGWDDDDVZFFDIG-UHFFFAOYSA-N; 126.031694 g/mol)
Benzene-1,3,5-triol		108-73-6	Benzene-1,3,5-triol (QCDYQDY XPDABM-UHFFFAOYSA-N; 126.031694 g/mol)
2. Benzaldehydes (C <sub>6</sub> -C <sub>1</sub> )			
4-Hydroxybenzaldehyde	4-Formylphenol	123-08-0	4-Hydroxybenzaldehyde (RGHHSNMVTDWUBI-UHFFFAOYSA-N; 122.036779 g/mol)
3,4-Dihydroxybenzaldehyde	Protocatechualdehyde	139-85-5	3,4-Dihydroxybenzaldehyde (IBGBGRVKPALMCQ-UHFFFAOYSA-N; 138.031694 g/mol)
2,4,6-Trihydroxybenzaldehyde	Phloroglucinaldehyde	487-70-7	2,4,6-Trihydroxybenzaldehyde (BTQAIGSMXCDDAL-UHFFFAOYSA-N; 154.026609 g/mol)
3-Hydroxy-4-methoxybenzaldehyde	Isovanillin	621-59-0	3-Hydroxy-4-methoxybenzaldehyde (JVTZY YHCGSXIV-UHFFFAOYSA-N; 152.047344 g/mol)
4-Hydroxy-3-methoxybenzaldehyde	3-Methoxy-4-hydroxybenzaldehyde Vanillin	121-33-5	4-Hydroxy-3-methoxybenzaldehyde (MWOOGOJBHARFG-UHFFFAOYSA-N; 152.047344 g/mol)
3. Benzoic acids (C <sub>6</sub> -C <sub>1</sub> )			
3-Hydroxybenzoic acid	3-Hydroxybenzene carboxylic acid	99-06-9	3-Hydroxybenzoic acid (UFXRHURBJZNAO-UHFFFAOYSA-N; 138.031694 g/mol)
4-Hydroxybenzoic acid	4-Hydroxybenzene carboxylic acid	99-96-7	4-Hydroxybenzoic acid (PIKROLUGYXJWQN-UHFFFAOYSA-N; 138.031694 g/mol)
2,5-Dihydroxybenzoic acid	Gallic acid 5-Hydroxy-salicylic acid Hydroquinone carboxylic acid Protocatechuic acid	490-79-9	2,5-Dihydroxybenzoic acid (WXTMDXOMEHIXQO-UHFFFAOYSA-N; 154.026611 g/mol)
3,4-Dihydroxybenzoic acid	$\alpha$ -Resorcylic acid	99-50-3	3,4-Dihydroxybenzoic acid (YQUVCSBIEUQKSH-UHFFFAOYSA-N; 154.026609 g/mol)
3,5-Dihydroxybenzoic acid	Gallic acid	99-10-5	3,5-Dihydroxybenzoic acid (UYEMGAFJZZIFP-UHFFFAOYSA-N; 154.026609 g/mol)
3,4,5-Trihydroxybenzoic acid		149-91-7	3,4,5-Trihydroxybenzoic acid (LNTHTITQWFMADLM-UHFFFAOYSA-N; 170.021523 g/mol)
2-Amino-5-hydroxybenzoic acid	5-Hydroxyanthranilic acid	394-31-0	2-Amino-5-hydroxybenzoic acid (HYNQTSZBTOTFKH-UHFFFAOYSA-N; 153.042593 g/mol)

(Continued)

TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey, <sup>6</sup> monoisotopic mass g/mol)
Benzoic acid-4-glucuronide	—	—	—
Benzoic acid-4-sulfate	—	—	—
3-Hydroxybenzoic acid-4-glucuronide	[Protocatechuic acid-4-glucuronide]	—	—
4-Hydroxybenzoic acid-3-glucuronide	[Protocatechuic acid-3-glucuronide]	—	—
3-Hydroxybenzoic acid-4-sulfate	[Protocatechuic acid-4-sulfate]	—	—
4-Hydroxybenzoic acid-3-sulfate	[Protocatechuic acid-3-sulfate]	—	—
3-Hydroxy-4-methoxybenzoic acid	Isovanillic acid	645-08-9	3-Hydroxy-4-methoxybenzoic acid (LBKFGYZQBSGRHY-UHFFFAOYSA-N; 168.042259 g/mol)
4-Methoxybenzoic acid-3-glucuronide	[Isovanillic acid-3-glucuronide]	—	—
4-Methoxybenzoic acid-3-sulfate	[Isovanillic acid-3-sulfate]	—	—
4-Hydroxy-3-methoxybenzoic acid	Vanillic acid	499-76-3	4-Hydroxy-3-methylbenzoic acid (LTFHKKUKQYVHDX-UHFFFAOYSA-N; 152.047344 g/mol)
3-Methoxybenzoic acid-4-glucuronide	[Vanillic acid-4-glucuronide]	—	—
3-Methoxybenzoic acid-4-sulfate	[Vanillic acid-4-sulfate]	—	—
3,5-Dimethoxy-4-hydroxybenzoic acid	Syringic acid	530-57-4	4-Hydroxy-3,5-dimethoxybenzoic acid (JMSYCTWVEWCHDZ-UHFFFAOYSA-N; 198.052823 g/mol)
3,4-Dimethoxybenzoic acid	Veratric acid	93-07-2	3,4-Dimethoxybenzoic acid (DAUAQNGYDSHRET-UHFFFAOYSA-N; 182.057909 g/mol)
4-Cinnamic acids (C <sub>6</sub> -C <sub>3</sub> unsaturated) <sup>7</sup>	<i>trans</i> -Cinnamic acid	621-82-9, 140-10-3	(2 <i>E</i> )-3-phenylprop-2-enoic acid (WBYWAXJHAXSINI-VOTSOKGWSA-N; 148.052429 g/mol)
Cinnamic acid			
2'-Hydroxycinnamic acid	2-Hydroxycinnamic acid <sup>8</sup>	614-60-8	(2 <i>E</i> )-3-(2-hydroxyphenyl)prop-2-enoic acid (PMOWTIVN WZYFF-AATRIKPKSA-N; 164.047344 g/mol)
3'-Hydroxycinnamic acid	<i>o</i> -Coumaric acid	588-30-7	(2 <i>E</i> )-3-(3-hydroxyphenyl)prop-2-enoic acid (KKSJDGJDHZE WEP-SNAWJCMRSA-N; 164.047344 g/mol)
4'-Hydroxycinnamic acid	<i>trans</i> -2-Coumaric acid	501-98-4	(2 <i>E</i> )-3-(4-hydroxyphenyl)prop-2-enoic acid (NGSWKAQJWESNS-ZZZKWWIFSA-N; 164.047344 g/mol)
	<i>trans</i> -3-Coumaric acid		
	3-(3-Hydroxyphenyl)acrylic acid		
	4-Hydroxycinnamic acid <sup>8</sup>		
	Coumaric acid		
	<i>p</i> -Coumaric acid		
	4-Coumaric acid		
	<i>trans</i> -4-Hydroxycinnamic acid		
	(4-Hydroxyphenyl)acrylic acid		
	3,4-Dihydroxycinnamic acid <sup>8</sup>		
	Caffeic acid	331-39-5, 501-16-6	(2 <i>E</i> )-3-(3,4-dihydroxyphenyl)prop-2-enoic acid (QAIPRVGONGVQAS-DUXPYHPUSA-N; 180.042259 g/mol)
	<i>trans</i> -Caffeic acid		

(Continued)

TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey; <sup>6</sup> monoisotopic mass g/mol)
Cinnamic acid-4'-glucuronide	3-(3,4-Dihydroxyphenyl)acrylic acid Cinnamic acid-4-glucuronide <sup>8</sup> [p-Coumaric acid-4'-glucuronide]		
4'-Hydroxy-3'-methoxycinnamic acid	4-Hydroxy-3-methoxycinnamic acid <sup>8</sup> Ferulic acid <i>trans</i> -Ferulic acid 3-(4-Hydroxy-3-methoxyphenyl)acrylic acid 3-Methoxycinnamic acid-4-glucuronide <sup>8</sup> [Ferulic acid-4-glucuronide]	537-98-4	(2 <i>E</i> )-3-(4-hydroxy-3-methoxyphenyl)prop-2-enoic acid (KSEBMYQB YZTDHS-HWKANZROSA-N; 194.057909 g/mol)
3'-Methoxycinnamic acid-4'-glucuronide	3-Methoxycinnamic acid-4-sulfate <sup>8</sup> [Ferulic acid-4'-sulfate]	—	—
3'-Methoxycinnamic acid-4'-sulfate	[Ferulic acid-4-sulfate]	—	—
3'-Hydroxy-4'-methoxycinnamic acid	3-Hydroxy-4-methoxycinnamic acid <sup>8</sup> Isoferulic acid Hesperetate	537-73-5	(2 <i>E</i> )-3-(3-hydroxy-4-methoxyphenyl)prop-2-enoic acid (QURCYMIEKCOAJU-HWKANZROSA-N; 194.057909 g/mol)
4'-Methoxycinnamic acid-3'-glucuronide	3-(3-Hydroxy-4-methoxyphenyl)acrylic acid 4-Methoxycinnamic acid-3-glucuronide <sup>8</sup> [Isoferulic acid-3'-glucuronide]	—	—
4'-Methoxycinnamic acid-3'-sulfate	[Isoferulic acid-3'-glucuronide] 4-Methoxycinnamic acid-3-sulfate <sup>8</sup> [Isoferulic acid-3'-sulfate] [Isoferulic acid-3-sulfate]	—	—
5. Phenylpropanoic acids (C <sub>6</sub> -C <sub>3</sub> )			
3-Phenylpropanoic acid	3-(Phenyl)propionic acid [Dihydrocinnamic acid]	501-52-0	3-Phenylpropanoic acid (XMIIGOLPHOKFCH-UHFFFAOYSA-N; 150.06808 g/mol)
3-(3'-Hydroxyphenyl)propanoic acid	3-(3-Hydroxyphenyl)propionic acid <sup>8</sup> 3-(3-Hydroxyphenyl)propionic acid 3-Hydroxy-dihydrocinnamic acid <i>m</i> -Hydroxy-dihydrocinnamic acid 3-(4-Hydroxyphenyl)propionic acid <sup>8</sup> 3-(4-Hydroxyphenyl)propionic acid 4-Hydroxy-dihydrocinnamic acid <i>m</i> -Hydroxy-dihydrocinnamic acid Phloretic acid	621-54-5	3-(3-Hydroxyphenyl)propanoic acid (QVWAEZIXDYOKEH-UHFFFAOYSA-N; 166.062994 g/mol)
3-(4'-Hydroxyphenyl)propanoic acid	3-(4-Hydroxyphenyl)propionic acid 3-(4-Hydroxyphenyl)propionic acid 4-Hydroxy-dihydrocinnamic acid <i>m</i> -Hydroxy-dihydrocinnamic acid Phloretic acid	10516-71-9	3-(4-Hydroxyphenyl)propanoic acid (NMHMNPHRMINGLLB-UHFFFAOYSA-N; 180.078644 g/mol)
3-(3',4'-Dihydroxyphenyl)propanoic acid	3-(3,4-Dihydroxyphenyl)propionic acid <sup>8</sup> 3-(3,4-Dihydroxyphenyl)propionic acid Dihydrocaffeic acid 3,4-Dihydroxybenzenepropanoic acid 3,4-Dihydroxy-dihydrocinnamic acid 3-(3,5-Dihydroxyphenyl)propanoic acid <sup>8</sup> 3-(3',5'-Dihydroxyphenyl)propionic acid 3,5-Dihydroxybenzenepropanoic acid 3,5-Dihydroxy-dihydrocinnamic acid 3-(Phenyl)propanoic acid-4-glucuronide <sup>8</sup> 3-(Phenyl)propionic acid-4-glucuronide [4-Hydroxy-dihydrocinnamic acid-4-glucuronide] [p-Hydroxy-dihydrocinnamic acid-glucuronide]	1078-61-1	3-(3,4-Dihydroxyphenyl)propanoic acid (DZAUWHDUNRCTF-UHFFFAOYSA-N; 182.057909 g/mol)
3-(3',5'-Dihydroxyphenyl)propanoic acid	3-(3,5-Dihydroxyphenyl)propionic acid 3-(3,5-Dihydroxyphenyl)propionic acid 3-(Phenyl)propanoic acid-4-glucuronide <sup>8</sup> 3-(Phenyl)propionic acid-4-glucuronide [4-Hydroxy-dihydrocinnamic acid-4-glucuronide] [p-Hydroxy-dihydrocinnamic acid-glucuronide]	26539-01-5	3-(3,5-Dihydroxyphenyl)propanoic acid (TPPHIKQWNDGDLG-UHFFFAOYSA-N; 182.057909 g/mol)
3-(Phenyl)propanoic acid-4'-glucuronide		—	—

(Continued)

TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey; <sup>6</sup> monoisotopic mass g/mol)
3-(3'-Methoxyphenyl)propanoic acid	3-(3-Methoxyphenyl)propanoic acid <sup>8</sup> 3-(3-Methoxyphenyl)propionic acid 3-Methoxy-dihydrocinamic acid	1798-09-0	3-(3-Methoxyphenyl)propanoic acid (BIJQJLOZWBEZGA-UHFFFAOYSA-N; 166.062994 g/mol)
3-(4'-Methoxyphenyl)propanoic acid	3-(4-Methoxyphenyl)propanoic acid <sup>8</sup> 3-(4-Methoxyphenyl)propionic acid 4-Methoxy-dihydrocinamic acid	1929-29-9	3-(4-Methoxyphenyl)propanoic acid (FIUFLISGGHNSM-UHFFFAOYSA-N; 180.078644 g/mol)
3-(3'-Hydroxy-4'-methoxyphenyl)propanoic acid	3-(3-Hydroxy-4-methoxyphenyl)propanoic acid <sup>8</sup> Dihydro-isoferulic acid	1135-15-5	3-(3-Hydroxy-4-methoxyphenyl)propanoic acid (ZVIJQFTLXXGJA-UHFFFAOYSA-N; 196.073559 g/mol)
3-(4'-Hydroxy-3'-methoxyphenyl)propanoic acid	3-(4-Hydroxy-3-methoxyphenyl)dihydrocinamic acid 3-(4-Hydroxy-3-methoxyphenyl)propanoic acid <sup>8</sup> 3-(4-Hydroxy-3-methoxyphenyl)propionic acid	1135-23-5	3-(4-Hydroxy-3-methoxyphenyl)propanoic acid (BOLQJTPHPSDZHR-UHFFFAOYSA-N; 196.073559 g/mol)
3-(3'-Hydroxyphenyl)propanoic acid-4'-glucuronide	3-(4-Hydroxy-3-methoxyphenyl)dihydrocinamic acid 3-(3-Hydroxyphenyl)propanoic acid-4-glucuronide <sup>8</sup> 3-(3-Hydroxyphenyl)propionic acid-4-glucuronide <i>[Dihydrocaffeic acid-4'-glucuronide]</i>	—	—
3-(3'-Hydroxyphenyl)propanoic acid-4'-sulfate	<i>[Dihydrocaffeic acid-4-glucuronide]</i> 3-(3-Hydroxyphenyl)propanoic acid-4-sulfate <sup>8</sup> 3-(3-Hydroxyphenyl)propionic acid-4-sulfate	—	—
3-(4'-Hydroxyphenyl)propanoic acid-3'-glucuronide	<i>[Dihydrocaffeic acid-4-sulfate]</i> <i>[Dihydrocaffeic acid-4-sulfate]</i> 3-(4-Hydroxyphenyl)propionic acid-3-glucuronide <sup>8</sup> 3-(4-Hydroxyphenyl)propanoic acid-3-glucuronide	—	—
3-(4'-Hydroxyphenyl)propanoic acid-3'-sulfate	<i>[Dihydrocaffeic acid-3'-glucuronide]</i> <i>[Dihydrocaffeic acid-3-glucuronide]</i> 3-(4-Hydroxyphenyl)propanoic acid-3-sulfate <sup>8</sup> 3-(4-Hydroxyphenyl)propionic acid-3-sulfate	—	—
3-(3'-Methoxyphenyl)propanoic acid-4'-glucuronide	<i>[Dihydrocaffeic acid-3'-sulfate]</i> <i>[Dihydrocaffeic acid-3-sulfate]</i> 3-(3-Methoxyphenyl)propanoic acid-4-glucuronide <sup>8</sup> 3-(3-Methoxyphenyl)propionic acid-4-glucuronide	—	—
3-(3'-Methoxyphenyl)propanoic acid-4'-sulfate	<i>[Dihydroferulic acid-4'-glucuronide]</i> <i>[Dihydroferulic acid-4-glucuronide]</i> 3-(3-Methoxyphenyl)propanoic acid-4-sulfate <sup>8</sup> 3-(3-Methoxyphenyl)propionic acid-4-sulfate	—	—
3-(4'-Methoxyphenyl)propanoic acid-3'-glucuronide	<i>[Dihydroferulic acid-4-sulfate]</i> <i>[Dihydroferulic acid-4-sulfate]</i> 3-(4-Methoxyphenyl)propanoic acid-3-glucuronide <sup>8</sup> 3-(4-Methoxyphenyl)propionic acid-3-glucuronide	—	—
3-(4'-Methoxyphenyl)propanoic acid-3'-sulfate	<i>[Dihydro-isoferulic acid-3'-glucuronide]</i> 3-(4-Methoxyphenyl)propanoic acid-3-sulfate <sup>8</sup> 3-(4-Methoxyphenyl)propionic acid-3-sulfate <i>[Dihydro-isoferulic acid-3'-sulfate]</i> <i>[Dihydro-isoferulic acid-3-sulfate]</i>	—	—

(Continued)

TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey; <sup>6</sup> monoisotopic mass g/mol)
6. Hydroxy-3-(phenyl)propanoic acids (C <sub>6</sub> -C <sub>3</sub> ) <sup>9</sup> 3-Hydroxy-3-(phenyl)propanoic acid	3-(Phenyl)-3-hydroxypropionic acid 3-(Phenyl)hydracrylic acid	3480-87-3	3-Hydroxy-3-phenylpropanoic acid (AYOLELPCNDV/ZKZ-UHFFFAOYSA-N; 166.062994 g/mol)
3-Hydroxy-3-(3'-hydroxyphenyl)propanoic acid	3-Hydroxy-3-(3-hydroxyphenyl)propanoic acid <sup>8</sup> 3-(3-Hydroxyphenyl)-3-hydroxypropionic acid 3-Hydroxy-3-(3'-hydroxyphenyl)propionic acid 3-Hydroxy-3-(3-hydroxyphenyl)propionic acid <i>β,m</i> -Dihydroxyphenylpropionic acid 3-(3-Hydroxyphenyl)hydracrylic acid 3-(3'-Hydroxyphenyl)hydracrylic acid <i>m</i> -Hydroxyphenylhydracrylic acid	3247-75-4	3-Hydroxy-3-(3-hydroxyphenyl)propanoic acid (KHTAGVZHYUZYM-F-UHFFFAOYSA-N; 182.057909 g/mol)
3-Hydroxy-3-(3'-hydroxy-4'-methoxyphenyl)propanoic acid	3-Hydroxy-3-(3-hydroxy-4'-methoxyphenyl)propanoic acid <sup>8</sup> 3-Hydroxy-3-(3-hydroxy-4'-methoxyphenyl)propionic acid 3-(3-Hydroxy-4'-methoxyphenyl)-3-hydroxypropanoic acid 3-(3-Hydroxy-4'-methoxyphenyl)-3-hydroxypropionic acid 3-(3'-Hydroxy-4'-methoxyphenyl)hydracrylic acid 3-(3-Hydroxy-4'-methoxyphenyl)hydracrylic acid 3-(Phenyl)-2-hydroxypropanoic acid <sup>8</sup> 3-(Phenyl)-2-hydroxypropionic acid 3-(Phenyl)lactic acid	—	3-Hydroxy-3-(3-hydroxy-4'-methoxyphenyl)propanoic acid (JEXBTMWMYGBBHO-UHFFFAOYSA-N; 212.068473 g/mol)
2-Hydroxy-3-(phenyl)propanoic acid	2-Hydroxy-3-(phenyl)propionic acid 3-(Phenyl)-2-hydroxypropanoic acid 3-(Phenyl)lactic acid	828-01-3	2-Hydroxy-3-phenylpropionic acid (VOXXWSYKYCBWHO-UHFFFAOYSA-N; 166.062994 g/mol)
2-Hydroxy-3-(4'-hydroxyphenyl)propanoic acid	2-Hydroxy-3-(4-hydroxyphenyl)propanoic acid <sup>8</sup> 3-(4-Hydroxyphenyl)-2-hydroxypropanoic acid 2-Hydroxy-3-(4-hydroxyphenyl)propionic acid 2-Hydroxy-3-(4-hydroxyphenyl)propionic acid 3-(4'-Hydroxyphenyl)-2-hydroxypropionic acid 3-(4'-Hydroxyphenyl)lactic acid 3-(4-Hydroxyphenyl)lactic acid	306-23-0	2-Hydroxy-3-(4-hydroxyphenyl)propanoic acid (JVGVDSSUAVXRDY-UHFFFAOYSA-N; 182.057909 g/mol)
2-Hydroxy-3-(3',4'-dihydroxyphenyl)propanoic acid	2-Hydroxy-3-(3,4-dihydroxyphenyl)propanoic acid <sup>8</sup> 3-(3,4-Dihydroxyphenyl)-2-hydroxypropanoic acid 2-Hydroxy-3-(3',4'-dihydroxyphenyl)propionic acid 3-(3,4-Dihydroxyphenyl)-2-hydroxypropionic acid 3-(3',4'-Dihydroxyphenyl)lactic acid 3-(3,4-Dihydroxyphenyl)lactic acid	23028-17-3	3-(3,4-Dihydroxyphenyl)-2-hydroxypropanoic acid (PAFLSMZLRSPALU-UHFFFAOYSA-N; 198.052823 g/mol)
7. Phenylacetic acids (C <sub>6</sub> -C <sub>2</sub> ) Phenylacetic acid	Phenylethanoic acid 2-Phenylethanoate	103-82-2	2-Phenylacetic acid (WLVXDMOQOQPHL-UHFFFAOYSA-N; 136.052429 g/mol)
3'-Hydroxyphenylacetic acid	3-Hydroxyphenylacetic acid <sup>8</sup> 3-Hydroxyphenylethanoic acid	621-37-4	2-(3-Hydroxyphenyl)acetic acid (FVMDYYGIDFPZAX-UHFFFAOYSA-N; 152.047344 g/mol)
3'-Methoxyphenylacetic acid	3-Methoxyphenylacetic acid <sup>8</sup> 3-Methoxyphenylethanoic acid	1798-09-0	2-(3-Methoxyphenyl)acetic acid (LEGPZHPSPYIO-UHFFFAOYSA-N; 166.062994 g/mol)

(Continued)





TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey <sup>6</sup> monoisotopic mass g/mol)
3'-Hydroxyhippuric acid	3-Hydroxyhippuric acid <sup>8</sup> 2-(3-Hydroxybenzamido)acetic acid N-(3-hydroxybenzoyl)glycine	1637-75-8	2-(3-Hydroxybenzoyl)aminoacetic acid (XDOFWFNMYJRHEW-UHFFFAOYSA-N; 195.053158 g/mol)
4'-Hydroxyhippuric acid	4-Hydroxyhippuric acid <sup>8</sup> 2-(4-Hydroxybenzamido)acetic acid N-(4-hydroxybenzoyl)glycine	2482-25-9	2-(4-Hydroxybenzoyl)aminoacetic acid (ZMHLUFWWWPBTU-UHFFFAOYSA-N; 195.053158 g/mol)
4-Methoxyhippuric acid	4-Methoxyhippuric acid <sup>8</sup> 2-(4-Ethoxybenzamido)acetic acid N-(4-methoxybenzoyl)glycine	—	2-[(4-Methoxybenzoyl)amino]acetic acid (SIEIOUWSTGWJGE-UHFFFAOYSA-N; 209.068808 g/mol)
10, Phenyl- $\gamma$ -valerolactones (C <sub>6</sub> -C <sub>5</sub> ) <sup>9,10</sup>			
5-(3',4'-Dihydroxyphenyl)- $\gamma$ -valerolactone	5-(3,4-Dihydroxyphenyl)- $\gamma$ -valerolactone <sup>8</sup> $\delta$ -(3,4-Dihydroxyphenyl)- $\gamma$ -valerolactone	21618-92-8 1108192-01-3 (4S)	5-(3,4-Dihydroxyphenyl)methylloxolan-2-one (ZNXXTWTPQHVLMT-UHFFFAOYSA-N; 208.073559 g/mol) <sup>9</sup>
5-(4'-Hydroxyphenyl)- $\gamma$ -valerolactone-3'-glucuronide	5-(4-Hydroxyphenyl)- $\gamma$ -valerolactone-3'-glucuronide [M6] [M6-glucuronide]	191666-22-5 (4R)	3,4,5-Trihydroxy-6-[2-hydroxy-5-[(5-oxoxolan-2-yl)methyl]phenoxy]oxane-2-carboxylic acid (UVGDTVGPBRLMQY-UHFFFAOYSA-N; 384.105647 g/mol)
5-(3'-Hydroxyphenyl)- $\gamma$ -valerolactone-4'-glucuronide	5-(3-Hydroxyphenyl)- $\gamma$ -valerolactone-4'-glucuronide <sup>8</sup> [M6-glucuronide]	—	3,4,5-Trihydroxy-6-[(2-hydroxy-4-[(5-oxoxolan-2-yl)methyl]phenoxy]oxane-2-carboxylic acid (OTBIYBOGMPICIK-UHFFFAOYSA-N; 384.105647 g/mol) <sup>9</sup>
5-(4'-Hydroxyphenyl)- $\gamma$ -valerolactone-3'-sulfate	5-(4-Hydroxyphenyl)- $\gamma$ -valerolactone-3'-sulfate <sup>8</sup> [M6-sulfate]	—	[2-Hydroxy-5-[(5-oxoxolan-2-yl)methyl]phenyl] hydrogen sulfate (YAXFYVDUIDAQPTJ-UHFFFAOYSA-N; 288.030374 g/mol)
5-(3'-Hydroxyphenyl)- $\gamma$ -valerolactone-4'-sulfate	5-(3'-Hydroxyphenyl)- $\gamma$ -valerolactone-4'-sulfate <sup>8</sup> [M6-sulfate]	—	[2-Hydroxy-4-[(5-oxoxolan-2-yl)methyl]phenyl] hydrogen sulfate (WAXYAOJFDCCESK-UHFFFAOYSA-N; 288.030374 g/mol)
5-(3'-Methoxyphenyl)- $\gamma$ -valerolactone-4'-sulfate	5-(3-Methoxyphenyl)- $\gamma$ -valerolactone-4'-sulfate <sup>8</sup> [Methyl-M6-sulfate]	—	[2-Methoxy-4-[(5-oxoxolan-2-yl)methyl]phenyl] hydrogen sulfate (FYRRHCSCZYSADR-UHFFFAOYSA-N; 302.046024 g/mol)
5-(3'-Methoxyphenyl)- $\gamma$ -valerolactone-4'-glucuronide	5-(3'-Methoxyphenyl)- $\gamma$ -valerolactone-4'-glucuronide <sup>8</sup> [Methyl-M6-glucuronide]	—	3,4,5-Trihydroxy-6-[2-methoxy-4-[(5-oxoxolan-2-yl)methyl]phenoxy]oxane-2-carboxylic acid (NGMVEYPOYAIGEL-UHFFFAOYSA-N; 398.121297 g/mol)
5-(Phenyl)- $\gamma$ -valerolactone-3'-sulfate-4'-glucuronide	5-(Phenyl)- $\gamma$ -valerolactone-3'-sulfate-4'-glucuronide <sup>8</sup> [M6-sulfate-glucuronide]	—	—
5-(3',4'-Dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulfate-4'-glucuronide	5-(3',4'-Dihydroxyphenyl)- $\gamma$ -valerolactone-3'-sulfate-4'-glucuronide <sup>8</sup>	—	—

(Continued)

TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey, <sup>6</sup> monoisotopic mass g/mol)
11. 4-Hydroxy-5-(phenyl)valeric acids (C <sub>6</sub> -C <sub>5</sub> ) <sup>9</sup>			
4-Hydroxy-5-(3',4'-dihydroxyphenyl)valeric acid	4-Hydroxy-5-(3,4-dihydroxyphenyl)valeric acid <sup>8</sup> 5-(3,4-Dihydroxyphenyl)-4-hydroxyvaleric acid 4-Hydroxy-5-(3,4-dihydroxyphenyl)pentanoic acid 5-(3,4-Dihydroxyphenyl)-4-hydroxypentanoic acid [5-(3,4-Dihydroxyphenyl)-γ-hydroxyvaleric acid] [5-(3,4-Dihydroxyphenyl)-γ-hydroxypentanoic acid]	—	5-(3,4-Dihydroxyphenyl)-4-hydroxypentanoic acid (JDBYFCLHVYVXCX-UHFFFAOYSA-N; 226.084124 g/mol)
4-Hydroxy-5-(4'-hydroxyphenyl)valeric acid-3'-glucuronide	4-Hydroxy-5-(4-hydroxyphenyl)valeric acid-3-glucuronide <sup>8</sup> 5-(4-Hydroxyphenyl)-4-hydroxyvaleric acid-3-glucuronide 4-Hydroxy-5-(4-hydroxyphenyl)pentanoic acid-3-glucuronide 5-(4-Hydroxyphenyl)-4-hydroxypentanoic acid-3-glucuronide [5-(3,4-Dihydroxyphenyl)-4-hydroxyvaleric acid-3-glucuronide] 4-Hydroxy-5-(3-hydroxyphenyl)valeric acid-4-glucuronide <sup>8</sup> 5-(3-Hydroxyphenyl)-4-hydroxyvaleric acid-4-glucuronide 4-Hydroxy-5-(3-hydroxyphenyl)pentanoic acid-4-glucuronide 5-(4-Hydroxyphenyl)-4-hydroxyvaleric acid-4-glucuronide [5-(3',4'-Dihydroxyphenyl)-4-hydroxyvaleric acid-4'-glucuronide]	—	6-[5-(4-Carboxy-2-hydroxybutyl)-2-hydroxyphenoxy]-3,4,5-trihydroxoxane-2-carboxylic acid (BKJQCPDRWDNBN-UHFFFAOYSA-N; 402.116212 g/mol)
4-Hydroxy-5-(3'-hydroxyphenyl)valeric acid-4'-glucuronide	4-Hydroxy-5-(3-hydroxyphenyl)valeric acid-4-glucuronide 5-(3-Hydroxyphenyl)-4-hydroxyvaleric acid-4-glucuronide 4-Hydroxy-5-(3-hydroxyphenyl)pentanoic acid-4-glucuronide 5-(4-Hydroxyphenyl)-4-hydroxyvaleric acid-4-glucuronide [5-(3',4'-Dihydroxyphenyl)-4-hydroxyvaleric acid-4'-glucuronide]	—	6-[4-(4-Carboxy-2-hydroxybutyl)-2-hydroxyphenoxy]-3,4,5-trihydroxoxane-2-carboxylic acid (LLKUARGHNZMSRT-UHFFFAOYSA-N; 402.116212 g/mol)
4-Hydroxy-5-(4'-hydroxyphenyl)valeric acid-3'-sulfate	4-Hydroxy-5-(4-hydroxyphenyl)valeric acid-3-sulfate <sup>8</sup> 5-(4-Hydroxyphenyl)-4-hydroxyvaleric acid-3-sulfate 4-Hydroxy-5-(4-hydroxyphenyl)pentanoic acid-3-sulfate 5-(4-Hydroxyphenyl)-4-hydroxypentanoic acid-3-sulfate [5-(3',4'-Dihydroxyphenyl)-4-hydroxyvaleric acid-3'-sulfate]	—	4-Hydroxy-5-(4-hydroxy-3-sulfooxyphenyl)pentanoic acid (HROSNTXKMPHTSL-UHFFFAOYSA-N; 306.040939 g/mol)
4-Hydroxy-5-(3'-hydroxyphenyl)valeric acid-4'-sulfate	4-Hydroxy-5-(3-hydroxyphenyl)valeric acid-4-sulfate <sup>8</sup> 5-(3-Hydroxyphenyl)-4-hydroxyvaleric acid-4-sulfate 4-Hydroxy-5-(3-hydroxyphenyl)pentanoic acid-4-sulfate 5-(3-Hydroxyphenyl)-4-hydroxypentanoic acid-4-sulfate [5-(3',4'-Dihydroxyphenyl)-4-hydroxyvaleric acid-4'-sulfate]	—	—
4-Sulfoxy-5-(3',4'-dihydroxyphenyl)valeric acid	4-Sulfoxy-5-(3,4-dihydroxyphenyl)valeric acid <sup>8</sup> 5-(3,4-Dihydroxyphenyl)-4-sulfoxyvaleric acid 4-Sulfoxy-5-(3,4-dihydroxyphenyl)pentanoic acid 5-(3,4-Dihydroxyphenyl)-4-sulfoxypentanoic acid [5-(3-Hydroxyphenyl)-4-sulphoxyvaleric acid]	—	5-(3,4-Dihydroxyphenyl)-4-sulfooxypentanoic acid (NBKHZVHBUVNGQF-UHFFFAOYSA-N; 306.040939 g/mol)
4-Hydroxy-5-(3'-methoxyphenyl)valeric acid-4'-sulfate	4-Hydroxy-5-(3-methoxyphenyl)valeric acid-4-sulfate <sup>8</sup> 5-(3-Methoxyphenyl)-4-hydroxyvaleric acid-4-sulfate 4-Hydroxy-5-(3-methoxyphenyl)pentanoic acid-4-sulfate 5-(3-Methoxyphenyl)-4-hydroxypentanoic acid-4-sulfate [4-Hydroxy-5-(3,4-dihydroxyphenyl)-valeric acid-3-methoxy-4-sulfate]	—	4-Hydroxy-5-[3-methoxy-4-(sulfooxy)phenyl]pentanoic acid (GUEAZORXKKSMA-UHFFFAOYSA-N; 320.056589 g/mol)

(Continued)

TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey, <sup>6</sup> monoisotopic mass g/mol)
4-Hydroxy-5-(4'-methoxyphenyl)valeric acid-3'-sulfate	4-Hydroxy-5-(4-methoxyphenyl)valeric acid-3-sulfate <sup>8</sup> 5-(4-Methoxyphenyl)-4-hydroxyvaleric acid-3-sulfate 4-Hydroxy-5-(4'-methoxyphenyl)pentanoic acid-3'-sulfate 5-(4-Methoxyphenyl)-4-hydroxypentanoic acid-3-sulfate [4-Hydroxy-5-(3,4-dihydroxyphenyl)-valeric acid-3-methoxy-4-sulfate]	—	—
4-Sulfoxy-5-(4'-hydroxy-3'-methoxyphenyl)valeric acid	4-Sulfoxy-5-(4-hydroxy-3-methoxyphenyl)valeric acid <sup>8</sup> 5-(4-Hydroxy-3-methoxyphenyl)-4-sulfoxyvaleric acid 4-Sulfoxy-5-(4-hydroxy-3-methoxyphenyl)pentanoic acid 5-(4-Hydroxy-3-methoxyphenyl)-4-sulfoxy-pentanoic acid [4-Hydroxy-5-(3,4-dihydroxyphenyl)-valeric acid-3-methoxy-4-sulfate]	—	5-(4-Hydroxy-3-methoxyphenyl)-4-(sulfoxy)pentanoic acid (JQISKEAFUDWLCF-UHFFFAOYSA-N; 320.056589 g/mol)
4-Hydroxy-5-(3'-methoxyphenyl)valeric acid-4'-glucuronide	4-Hydroxy-5-(3-methoxyphenyl)valeric acid-4-glucuronide <sup>8</sup> 5-(3-Methoxyphenyl)-4-hydroxyvaleric acid-4-glucuronide 4-Hydroxy-5-(3-methoxyphenyl)pentanoic acid-4-glucuronide 5-(3'-Methoxyphenyl)-4-hydroxypentanoic acid-4-glucuronide [4-Hydroxy-5-(3',4'-dihydroxyphenyl)-valeric acid-3'-methoxy-4'-glucuronide]	—	6-[4-(4-Carboxy-2-hydroxybutyl)-2-methoxyphenoxy]-3,4,5-trihydroxyoxane-2-carboxylic acid (WOBDOUCWWPCYBL-UHFFFAOYSA-N; 416.131862 g/mol)
12. 5-Phenylvaleric acids (C <sub>6</sub> -C <sub>5</sub> )			
5-(3',5'-Dihydroxyphenyl)valeric acid	5-(3,5-Dihydroxyphenyl)pentanoic acid <sup>8</sup> [5-(Dihydroxyphenyl)valeric acid] [Dihydroxy-benzenepentanoic acid]	74356-41-5	5-(3,5-Dihydroxyphenyl)pentanoic acid (QHXNIMVPAFCPR-UHFFFAOYSA-N; 210.089203 g/mol)
5-(3',4',5'-Trihydroxyphenyl)valeric acid	5-(3,4,5-Trihydroxyphenyl)valeric acid <sup>8</sup> [5-(Trihydroxyphenyl)valeric acid] [Trihydroxy-benzenepentanoic acid]	—	5-(3,4,5-Trihydroxyphenyl)pentanoic acid (MASHUJSDLKLZLA-UHFFFAOYSA-N; 226.084124 g/mol)
13. Urolithins (6H-Dibenzo[b,d]pyran-6-one)			
3-Hydroxy-urolithin	Urolithin B	1139-83-9	3-Hydroxy-6H-benzo[c]chromen-6-one (WXUQMTRHPNOXBV-UHFFFAOYSA-N; 212.047344 g/mol)
Urolithin-3-glucuronide	Urolithin B-3-glucuronide [Urolithin B-glucuronide]	823806-74-2	5-(6-oxo-6H-benzo[c]chromen-3-yl)oxyoxane-2-carboxylic acid (MHBWCULXQBVPQT-KSPMYQCISA-N; 388.079432 g/mol)

(Continued)

TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey; <sup>6</sup> monoisotopic mass g/mol)
Urolithin-3-sulfate	Urolithin B-3-sulfate [ <i>Urolithin B-sulfate</i> ]	—	(6-Oxobenzol[ <i>c</i> ]chromen-3-yl) hydrogen sulfate (LRRVKQQFGIEMSA-UHFFFAOYSA-N; 292.004159 g/mol)
3,8-Dihydroxy-urolithin	Urolithin A	1143-70-0	3,8-Dihydroxy-6H-benzol[ <i>c</i> ]chromen-6-one (RIUPLDUFZCXCHM-UHFFFAOYSA-N; 228.042259 g/mol)
8-Hydroxy-urolithin-3-glucuronide	Urolithin A-3-glucuronide [ <i>Urolithin A-glucuronide</i> ] [3,8-Dihydroxy-urolithin 3-glucuronide]	—	3,4,5-Trihydroxy-6-(8-hydroxy-6-oxo-6H-benzol[ <i>c</i> ]chromen-3-yl)oxyoxane-2-carboxylic acid (KXBXNRRJGUDTIQS-UHFFFAOYSA-N; 404.074347 g/mol)
3-Hydroxy-urolithin-8-glucuronide	Urolithin A-8-glucuronide [ <i>Urolithin A-glucuronide</i> ] [3,8-Dihydroxy-urolithin 8-glucuronide]	1365982-52-0	(2S,3S,4S,5R,6S)-3,4,5-trihydroxy-6-(3-hydroxy-6-oxobenzol[ <i>c</i> ]chromen-8-yl)oxyoxane-2-carboxylic acid (QMPHAAUHRNZSL-KSPMYQCISA-N; 404.074347 g/mol)
Urolithin-3,8-digluconide	Urolithin A-3,8-digluconide [ <i>Urolithin A-digluconide</i> ] [3,8-Dihydroxy-urolithin-digluconide]	—	(2S,3S,4S,5R,6S)-6-[3-[(2S,3R,4S,5S,6S)-6-carboxy-3,4,5-trihydroxyoxan-2-yl]oxy-6-oxobenzol[ <i>c</i> ]chromen-8-yl]oxy-3,4,5-trihydroxyoxane-2-carboxylic acid (SXMJSEFKPOZNAT-ILCXFEASA-N; 580.106435 g/mol)
8-Hydroxy-urolithin-3-sulfate	Urolithin A-3-sulfate [ <i>Urolithin A-sulfate</i> ] [3,8-Dihydroxy-urolithin 3-sulfate]	—	(8-Hydroxy-6-oxobenzol[ <i>c</i> ]chromen-3-yl) hydrogen sulfate (WMPNAWQWWZEFJTQ-UHFFFAOYSA-N; 307.999074 g/mol)
3-Hydroxy-urolithin-8-sulfate	Urolithin A-8-sulfate [ <i>Urolithin A-sulfate</i> ] [3,8-Dihydroxy-urolithin 8-sulfate]	—	(3-Hydroxy-6-oxo-6H-benzol[ <i>c</i> ]chromen-8-yl)oxidanestulfonic acid
Urolithin-3-sulfate-8-glucuronide	[ <i>Urolithin A-sulfate-glucuronide</i> ] [3,8-Dihydroxy-urolithin 3-sulfate-8-glucuronide]	—	—
3,9-Dihydroxy-urolithin	Isourolithin A	174023-48-4	3,9-Dihydroxy-6H-benzol[ <i>c</i> ]chromen-6-one (WDGSXHQUPZEHA-UHFFFAOYSA-N; 228.042259 g/mol)
9-Hydroxy-urolithin-3-glucuronide	Isourolithin A-3-glucuronide [ <i>Isourolithin A-glucuronide</i> ] [3,9-Dihydroxy-urolithin-3-glucuronide]	—	—
3-Hydroxy-urolithin-9-glucuronide	Isourolithin A-9-glucuronide [ <i>Isourolithin A-glucuronide</i> ] [3,9-Dihydroxy-urolithin 9-glucuronide]	—	—

(Continued)

TABLE 1 (Continued)

Recommended names <sup>2</sup>	Synonyms and incorrect nomenclature <sup>2,3</sup>	CAS no. <sup>4</sup>	IUPAC <sup>5</sup> (InChIKey; <sup>6</sup> monoisotopic mass g/mol)
3,8,9-Trihydroxy-uro lithin	Urolithin C [Hydroxyuro lithin A] [Trihydroxyuro lithin] Urolithin C-3-glucuronide [Urolithin C-glucuronide]	165393-06-6  1268248-76-5	3,8,9-Trihydroxy-6H-benzo[c]chromen-6-one (HHXMEZVPJFAJ-UHFFFAOYSA-N; 244.037173 g/mol) 2-(5S,4S,5R,6S)-6-((8,9-dihydroxy-6-oxo-6H-benzo[c]chromen-3-yl)oxy)-3,4,5-trihydroxyoxane-2-carboxylic acid (DDAQYQCCOWZGDO-KSPMYQCISA-N; 420.069261 g/mol)
3,4,8,9-Tetrahydroxy-uro lithin	Urolithin D	131086-98-1	3,4,8,9-Tetrahydroxy-6H-benzo[c]chromen-6-one (NEZDQSKPNRYAW-UHFFFAOYSA-N; 260.032088 g/mol)
3,8,9,10-Tetrahydroxy-uro lithin	Urolithin M6	1006683-97-1	3,8,9,10-Tetrahydroxy-6H-benzo[c]chromen-6-one (LGXFTZDSEIQMMP-UHFFFAOYSA-N; 260.032088 g/mol)
3,8,10-Trihydroxy-uro lithin	Urolithin M7	531512-26-2	3,8,10-Trihydroxy-6H-benzo[c]chromen-6-one (AKJHSPSPAOUJFT-UHFFFAOYSA-N; 244.037173 g/mol)
14. Others (C <sub>6</sub> -C <sub>2</sub> )			
2-(3'-Hydroxyphenyl)ethanol	2-(3-Hydroxyphenyl)ethanol <sup>8</sup> 3-Hydroxytyrosol 3,4-Dihydroxyphenylethanol 3,4-Dihydroxyphenethyl alcohol 4-Hydroxyphenethyl alcohol <sup>8</sup> Tyrosol	10597-60-1     501-94-0	4-(2-Hydroxyethyl)-1,2-benzenediol (JUUBCHWRXWPFPH-UHFFFAOYSA-N; 154.062994 g/mol)  4-(2-Hydroxyethyl)phenol (YCCILYSKPBXVIP-UHFFFAOYSA-N; 138.06808 g/mol)

<sup>1</sup>This table was established based on cross-referencing the following databases: PubChem (<https://pubchem.ncbi.nlm.nih.gov>), NIST Chemistry WebBook (<https://webbook.nist.gov>), ChemIDPlus (<https://chem.nlm.nih.gov/chemidplus>), HMDB (<http://www.hmdb.ca>), ChemSpider (<http://www.chemspider.com>), ChemicalBook (<https://www.chemicalbook.com>), and ChEMBL (<https://www.ebi.ac.uk>). Column 1, proposed nomenclature; column 2, equally accurate alternative "nonprime" nomenclature indicated by superscript 8, followed by common or trivial names, and, finally, often reported inaccurate nomenclature shown in italics and brackets. Table 1 is available for download (Excel file; [Supplemental Table 1](https://www.cas.org/support/documentation/chemical-substances)). CAS-RN, Chemical Abstracts Service Registry Number; DOPAC, 3',4'-dihydroxyphenylacetic acid; IUPAC, International Union of Pure and Applied Chemistry.

<sup>2</sup>Glucuronic acid-oxygen conjugations are of  $\beta$ -D-configuration.

<sup>3</sup>Column 2 contains synonyms along with incorrect nomenclature that are in brackets. Both should be avoided to prevent further confusion in the literature and in online databases.

<sup>4</sup>CAS Registry Number (CAS-RN) is a numeric identifier that can contain  $\leq 10$  digits, divided by hyphens into 3 parts, and each number is a unique numeric identifier that can link to information about the substance to which it refers, but it has no chemical significance per se (see <https://www.cas.org/support/documentation/chemical-substances>). Not all compounds have a CAS number, typically because they have a limited or no commercial source.

<sup>5</sup>The IUPAC is described as the world authority on chemical nomenclature and terminology (<https://iupac.org/who-we-are>).

<sup>6</sup>InChIKey: IUPAC standard textual unique chemical identifier.

<sup>7</sup>Cinnamic acids have *cis* (Z) and *trans* (E) geometric isomers. In nature, the *trans* isomer is more common.

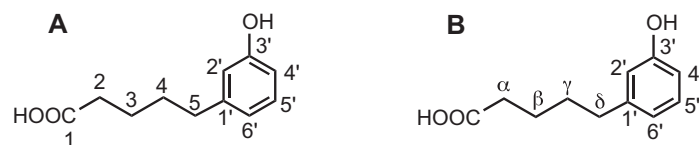
<sup>8</sup>Alternative "nonprime" nomenclature.

<sup>9</sup>Compounds that occur as *R*- and *S*-isomers.

<sup>10</sup>In rare instances, compounds appear to have 2 different InChIKey formulas in online databases which are generally associated with different CAS or registry numbers and possibly reflect uncharacterized isomeric configuration. For example, 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone or IUPAC 5-[(3,4-dihydroxyphenyl)methyl]oxolan-2-one is listed as having 2 different CAS RNs: 21618-92-8 and 191666-22-5; 2 different standard 27-character InChIKeys: ZNXXWTPQHVLMQOT-UHFFFAOYSA-N and ZNXXWTPQHVLMQOT-MRVPVSSYSA-N; and sharing the same simplified molecular-input line-entry system (SMILES) formula: C1CC(=O)OCC2=CC(=C(C=C2)O)O. In this case,

each structure has a unique full InChI (using the SHA-256 algorithm): InChI=IS/C11H12O4/c12-9-3-1-7(6-10(9)13)5-8-2-4-11(14)15-8/h1,3,6,8,12-13H,2,4-5H2 and IS/C11H12O4/c12-9-3-1-7(6-10(9)13)5-8-2-4-11(14)15-8/h1,3,6,8,12-13H,2,4-5H2(18-19)/s1, which should be used in cases in which 2 isomers require distinction. Similarly, 5-(3'-hydroxyphenyl)- $\gamma$ -valerolactone-4'-glucuronide or

IUPAC 3,4,5-trihydroxy-6-(2-hydroxy-4-(5-oxoxolan-2-yl)methyl)phenoxy)oxane-2-carboxylic acid has two different standard InChIKeys—OTBJYBQGMPCIK-UHFFFAOYSA-N and OTBJYBQGMPCIK-GHPVWUPISA-N—and different simplified molecular-input line-entry system (SMILES) formula and full InChIKey.



**FIGURE 1** Numbering of the phenolic ring and its side chain with the current practice (A) and the older less fashionable use of lowercase Greek letters for the side chain (B).

and widely used (e.g., hippuric acids, phenyl- $\gamma$ -valerolactones, phenylvaleric acids, and urolithins), modifications are used.

The phenolic compounds that are covered in [Table 1](#) are those originating from acyl-quinic acids, ellagitannins, and the main dietary flavonoids—namely anthocyanins, flavones, flavonols, flavanones, and flavan-3-ols—as well as the avenanthramides and alkyl-resorcinols characteristic of the staple foods wheat, rye, and quinoa. It does not cover catabolites of minor dietary (poly)phenols or those with a limited distribution, such as isoflavones, stilbenes, phenylpropenes, lignans, and iridoids (oleuropeins). Catabolites of these (poly)phenols also require a standardized nomenclature, which is beyond the scope of this article.

### Nomenclature of Phenolic Catabolites

(Poly)phenolic catabolites found in blood, tissues, urine, and feces can be easily subdivided using a shorthand system to describe their molecular skeleton by concisely defining the number of carbon atoms on the phenyl ring and its side chain and whether the side chain is unsaturated (i.e., containing double bonds) and/or contains nitrogen or substituents such as a hydroxyl group. In this shorthand system,  $C_6$  identifies a phenyl ring with its 6 carbon atoms ([Figure 1](#)), and the number of carbon atoms in the side chain is defined using  $C_0$ ,  $C_1$ ,  $C_2$ , etc., as necessary—that is,  $C_6-C_0$ ,  $C_6-C_1$ ,  $C_6-C_2$ , etc. The most extensively studied catabolites have a side chain with a terminal carboxyl group that is designated C-1. Further description of the molecule becomes more complicated because the literature contains trivial names, some of which are synonyms, such as acetic acid and ethanoic acid. For a  $C_6-C_3$  structure, the synonyms cinnamic, propenoic, and acrylic all signify an unsaturated  $C_3$  side chain with a double bond between C-2 and C-3. The synonyms propionic, propanoic, and dihydrocinnamic acid denote a saturated  $C_3$  side chain. Longer side chains include pentanoic ( $C_6-C_5$ ), also described as valeric, and the less common heptanoic ( $C_6-C_7$ ) and nonanoic ( $C_6-C_9$ ). The term phenyl-pentenoic identifies a  $C_6-C_5$  catabolite with an unsaturated side chain but does not specify the location of the double bond. This can be defined by identifying the lowest-number carbon associated with it—for example, pent-2-enoic or pent-3-enoic—and, if possible, specifying whether it is of a *cis* (*Z*) or *trans* (*E*) configuration. The current practice for numbering the carbon atoms in these compounds is shown in [Figure 1A](#), with the former less fashionable practice of using lowercase Greek letters shown in [Figure 1B](#), and illustrates the proposed framework for simplifying phenolic metabolite nomenclature.

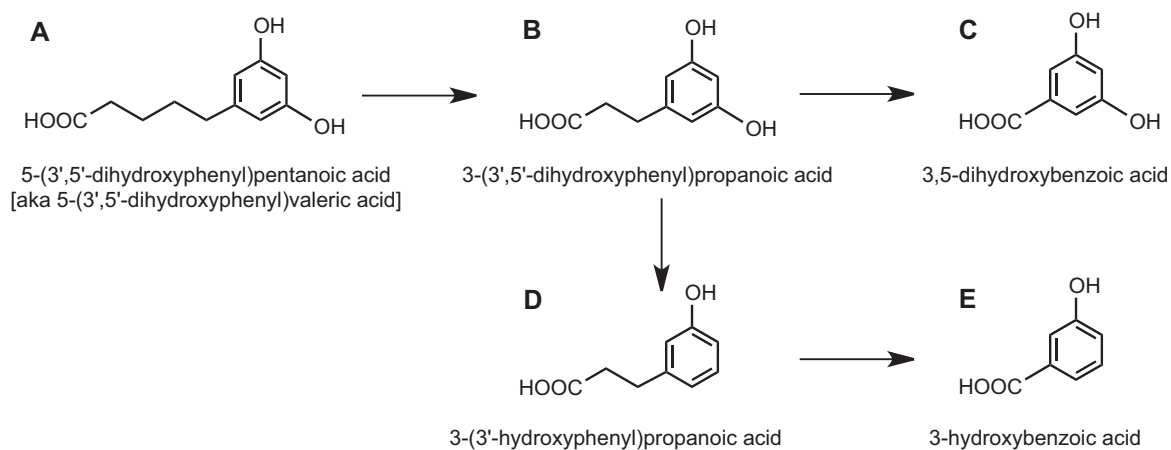
[Figure 1](#) also illustrates how the ring carbons can be distinguished from the side-chain carbons using a number bearing a prime—for example, 3'—rather than simply 3. This use of a prime is of value when otherwise a structure would contain

two atoms that might reasonably be described as “3,” as in 3-(3'-hydroxyphenyl)propanoic acid, in which 3 refers to the number of carbons in the side chain (counting from the carboxyl carbon, COOH = 1) and 3' refers to the position of substituent groups on the phenyl ring, with 1' being the carbon on the phenyl ring bearing the primary side chain. This use of prime numbers may not be essential by IUPAC convention, and seems superfluous to seasoned chemists, but it does not require an extensive understanding of nomenclature to facilitate its use. The system recommended here is designed to be easily accessible to nonchemists.

For  $C_6-C_0$  or  $C_6-C_1$  structures such as phenols and benzoic acids, the use of prime numbers is not required because the phenyl ring is the only site for substituent groups. This is illustrated in [Figure 2](#) with the conversion of 3-(3',5'-dihydroxyphenyl)propanoic acid ([Figure 2B](#)) to 3,5-dihydroxybenzoic acid ([Figure 2C](#)) and 3-hydroxybenzoic acid ([Figure 2E](#)).

The advantage of the prime symbol is clear with a structure such as 3-(3'-hydroxyphenyl)propanoic acid, but one could argue it is not entirely necessary to use a prime number for 3-(4'-hydroxyphenyl)propanoic acid because there are only 3 carbons in the side chain and no risk of assigning a bond position twice. However, this can become confusing when there are  $\geq 2$  substituents in the ring, as in 3-(3',4'-dihydroxyphenyl)propanoic acid, or if 1 of the hydroxyls in the phenolic ring is a biological conjugate, as in 3-(3'-hydroxyphenyl)propanoic acid-4'-sulfate, a feature discussed more extensively later. [Table 1](#) outlines the proposed and alternatively accepted nomenclature for reference and for integration of published works and databases. Examples of the proposed nomenclature are provided in column 1 of [Table 1](#), with equally accurate but alternative “nonprime” nomenclature, often found in online databases, presented in column 2 (superscript 8), followed by common or trivial names, and, finally, in italics and brackets, the often reported nomenclature for phase II conjugates that is inaccurate because of the double assigning of carbon atoms, as further discussed later. The use of trivial names and inaccurate nomenclature should be avoided to prevent further confusion in the literature and online databases. Column 3 contains CAS numbers (<https://www.cas.org/support/documentation/chemical-substances>), and column 4 includes the IUPAC nomenclature followed in parentheses by standard InChIKey chemical identifiers (<https://iupac.org/who-we-are>) and monoisotopic mass values.

[Figure 3](#) illustrates how the numbering of the ring carbons might change following metabolism. For example, when the 2'-OH is removed from the 3-(2',4',5'-trihydroxyphenyl)propanoic acid ([Figure 3A](#)) and the 3-carbon side chain is shortened, this produces 3,4-dihydroxybenzoic acid ([Figure 3B](#)) because in assigning numbers to the carbons bearing hydroxyl groups, it



**FIGURE 2** Nomenclature of metabolites involved in the microbial conversion of C<sub>6</sub>-C<sub>5</sub> to C<sub>6</sub>-C<sub>3</sub> and C<sub>6</sub>-C<sub>1</sub>.

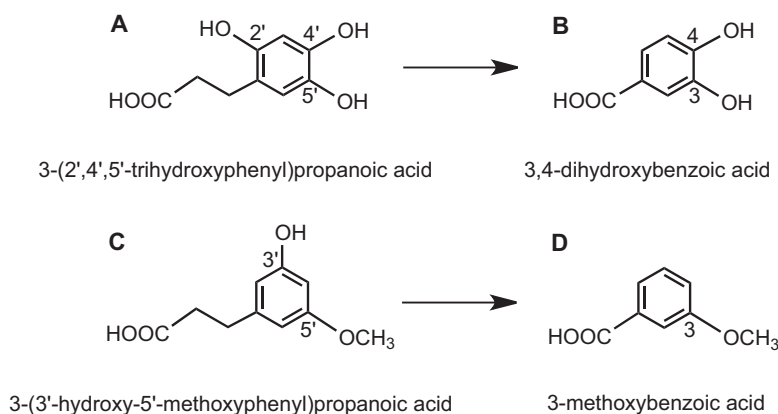
is necessary to take the shortest route from the carbon bearing the carboxyl moiety. This also applies to 3-(3'-hydroxy-5'-methoxyphenyl)propanoic acid (Figure 3C), in which the numbering of the hydroxyl is given priority over the methoxy group. When the hydroxyl is removed and the side chain shortened, the C<sub>6</sub>-C<sub>1</sub> product is 3-methoxybenzoic acid (Figure 3D) rather than 5-methoxybenzoic acid.

The terms hydracrylic acid and lactic acid describe isomeric saturated C<sub>3</sub> side chains bearing a hydroxyl, at C-3 for the hydracrylic acid side chain but at C-2 for a lactic acid side chain. These compounds are referred to as 3-hydroxy-3-(phenyl)propanoic acids and 2-hydroxy-3-(phenyl)propanoic acids, respectively. The alternative non-IUPAC names that are often used are 3-(phenyl)-3-hydroxypropionic and 2-hydroxy-3-(phenyl)propionic acids, respectively (Table 1, sections 6 and 8). For both of these types of compounds, there are *R*- and *S*-isomers determined by the orientation of the side-chain hydroxyl group.

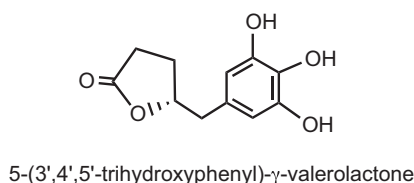
Phenyl- $\gamma$ -valerolactones, the main microbiota-derived metabolites of flavan-3-ols, are chemically related to 4-hydroxy-5-(phenyl)pentanoic acids. The most common species include 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone, 5-(3',5'-dihydroxyphenyl)- $\gamma$ -valerolactone, and 5-(3',4',

5'-trihydroxyphenyl)- $\gamma$ -valerolactone (Figure 4). They have traditionally been abbreviated as M6, M6', and M4, respectively, but this is now not recommended because it complicates accurate structural assignment. More details about standardized nomenclature for phenyl- $\gamma$ -valerolactones can be found in a recent review article by Mena et al. (14).

Urolithins, microbiota-mediated products derived from catabolism of ellagitannins, possess a 6H-dibenzo-*[b,d]*pyran-6-one structure (Figure 5), making them particularly difficult to assign appropriate structural nomenclature. They represent a special case because they possess 2 phenyl rings and lack a side chain, and therefore a slightly different numbering convention is required. The accepted system for numbering (15) is presented in Figure 5. The adoption of this system avoids issues associated with an oversimplified nomenclature often adopted in the literature, based on assigning "A," "B," or "C" to a structure, such as urolithin-A (3,8-dihydroxy-urolithin), -B (3-hydroxy-urolithin), and -C (3,8,9-trihydroxy-urolithin). The use of the A, B, C nomenclature is not recommended (Table 1) because it easily leads to inaccurate assignments, particularly when describing the positions of human phase II methyl, glucuronide, or sulfate conjugates.



**FIGURE 3** An illustration of how numbering of substituents may change during catabolism.



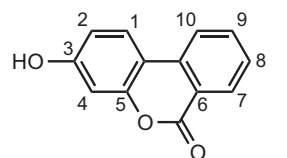
**FIGURE 4** Structure of 5-(3',4',5'-trihydroxyphenyl)- $\gamma$ -valerolactone.

## Phase II Metabolites of Phenolic Catabolites

The correct designation of mammalian phase II conjugates where a hydroxyl has been conjugated with a sulfate, glucuronide, or methyl group during the metabolism of dietary (poly)phenols, often causes confusion. For example, if 5-(3',4'-dihydroxyphenyl)valeric acid is conjugated with a glucuronide moiety, the substrate effectively loses one hydroxyl and becomes 5-(3'-hydroxyphenyl)valeric acid-4'-glucuronide or 5-(4'-hydroxyphenyl)valeric acid-3'-glucuronide, depending on which hydroxyl is conjugated. It is essential to be able to distinguish such isomers. To use 5-(3',4'-dihydroxyphenyl)valeric acid-3'-glucuronide would be inaccurate and is not recommended nomenclature because the 3' position is assigned twice. However, in order to facilitate interpretation and translation of pre-existing publications and databases, such “double assignments” are shown, italicized and in brackets in column 2 of Table 1, because they are common in the literature. “Double assignment” is also encountered when catabolites are described using trivial names—for example, dihydrocaffeic acid, dihydrocaffeic acid-3-glucuronide, and dihydrocaffeic acid-4-glucuronide. This should also be avoided, and we recommend 3-(3',4'-dihydroxyphenyl)propanoic acid, 3-(4'-hydroxyphenyl)propanoic acid-3'-glucuronide, and 3-(3'-hydroxyphenyl)propanoic acid-4'-glucuronide, respectively. Again, for the purposes of translation, trivial names are shown in column 2 of Table 1.

It is common when describing conjugated molecules to define the nature of the atom to which the conjugating moiety is attached—for example, *O*-glucuronide, *O*-sulfate, and *O*-methyl. This is necessary in any situation in which alternative points of attachment might be encountered, but as far as the authors are aware, there are no reported human glucuronide or sulfate metabolites of dietary (poly)phenols that are conjugated to a carbon atom. Accordingly, the use of “*O*” is redundant, albeit not incorrect, but use of the lowercase “*o*” is incorrect. We also recommend that methyl conjugates be referred to as “methoxy-” rather than “*O*-methyl” derivatives—for example, 3-(3'-hydroxy-4'-methoxyphenyl)propanoic acid.

The issue of conjugation is often further complicated by commercial vendors describing synthetic reference



3-hydroxy-uroolithin (uroolithin B)

**FIGURE 5** Urolithin nomenclature based on Gonzalez-Barrío et al. (15).

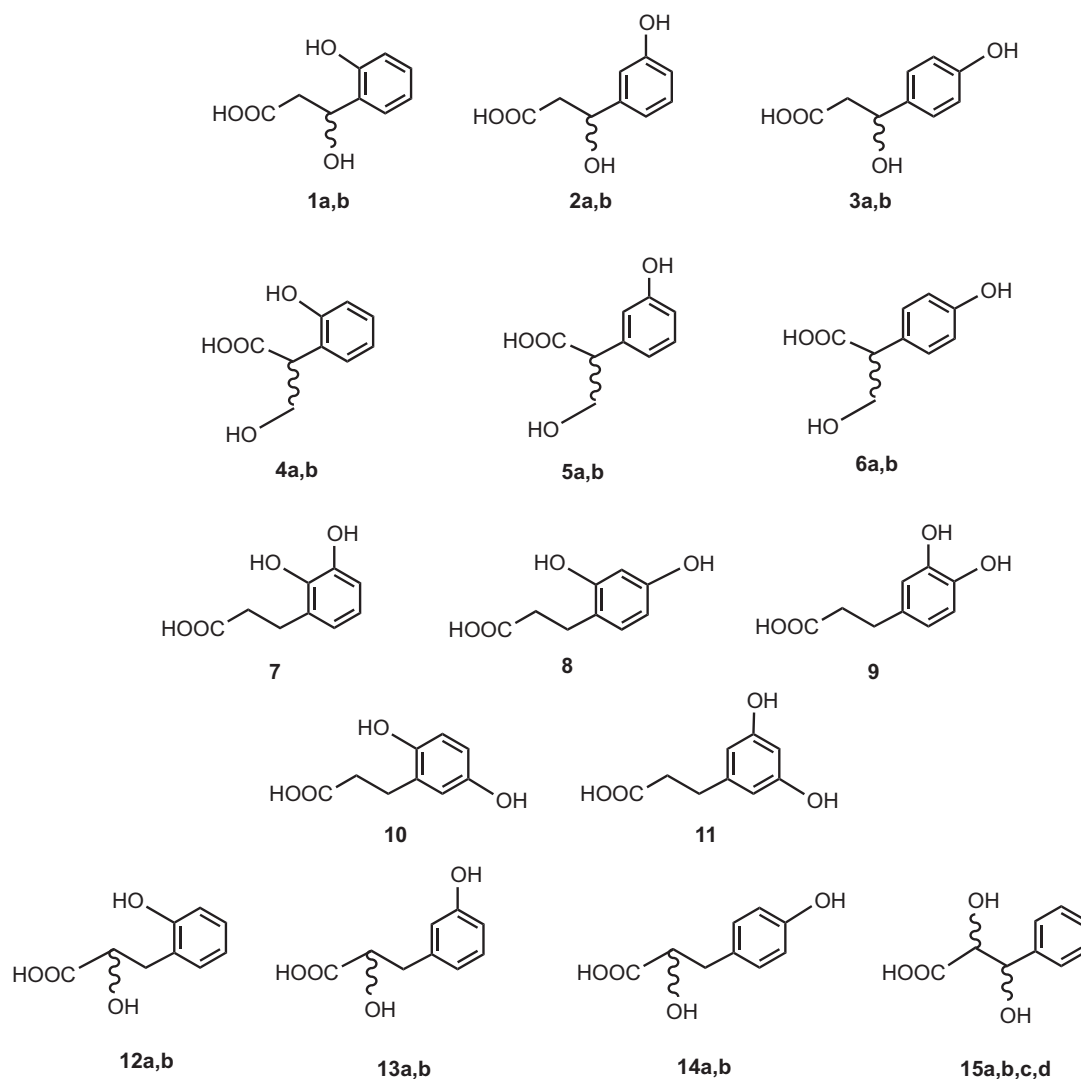
standards using inaccurate nomenclature. For example, many companies and online databases present the phase II metabolite of 3-(3',4'-dihydroxyphenyl)propanoic acid as 3-(3',4'-dihydroxyphenyl)propanoic acid-3'-glucuronide. As noted previously, a bond position for (poly)phenols and phenolic metabolites should not be assigned twice in structural nomenclature. The accurate representation for the structure under the proposed nomenclature is 3-(4'-hydroxyphenyl)propanoic acid-3'-glucuronide. As another example, phenyl- $\gamma$ -valerolactone conjugates are often inaccurately annotated as 5-(3',4'-dihydroxyphenyl)- $\gamma$ -valerolactone-4'-glucuronide, although the correct nomenclature is 5-(3'-hydroxyphenyl)- $\gamma$ -valerolactone-4'-glucuronide (14). The same applies to the naming of sulfate and methoxy metabolites. In addition, Greek symbols such as “ $\alpha$ ,” “ $\beta$ ,” and “ $\gamma$ ” may be presented as their textual equivalents “alpha,” “beta,” and “gamma” in some databases in which symbols may not be compatible with certain software or programs.

## Isomers

Dietary (poly)phenols, their catabolites, and phase II conjugates can display multiple forms of isomerism, and because biological activity is a function of the 3D structure, it is important to discriminate between isomers and describe their structure clearly and unambiguously. It is recommended that after determining the general nature of a metabolite, the number of possible isomers is calculated before assigning a more precise structure. The possible number may be surprisingly high, and if there are insufficient data to allow unequivocal discrimination between the possible structures, the description applied must make this clear. There may be good reasons to eliminate some of these possibilities and favor others, and this can be made clear in discussion, but overprecise assignment must be avoided. Previous detailed recommendations on this topic, which we endorse, were provided by Sumner et al. (16).

It is easy to overlook just how many isomers there might be even without the added complication of conjugation and geometric isomerization, because there may be regio- and stereo-isomers. Although perhaps not immediately obvious, there are  $\geq 27$  C<sub>6</sub>–C<sub>3</sub> compounds having two hydroxyls and an exact mass of 182.0579 Da (Figure 6), and although they are not all typical urinary catabolites, it is important to be able to distinguish between them because each of these structures could have very different biological activities. Furthermore, two stereo-isomers are also possible for the C<sub>6</sub>–C<sub>3</sub> phenylpropanoic acids, having a side-chain hydroxyl group at C-2 or C-3 (Figure 6, structures 1–6 and 12–15). Racemization—that is, conversion of an *S*-enantiomer to an *R*-enantiomer or the reverse—is possible at least for structures 1–3 in Figure 6 and is known to occur during mitochondrial  $\beta$ -oxidation of fatty acids, generally producing a small excess of the *R*-enantiomer (17). If it is uncertain which isomer is present, or if it is likely to be an unresolvable mixture, the structure can be described as, for example, 3*R/S*, and the structure can be drawn with the “wiggly” bond as shown in structures 1–6 and 12–15 (Figure 6). The *R* and *S* configuration can be critical in determining the interaction of a compound with a protein, transporter, or enzyme and hence is critical for biological activity. As an example of the importance of this





**FIGURE 6** C<sub>6</sub>–C<sub>3</sub> compounds with 2 hydroxyl groups and an exact mass of 182.0579 Da. 1a, b, 3*R/S*-hydroxy-3-(2'-hydroxyphenyl)propanoic acid; 2a, b, 3*R/S*-hydroxy-3-(3'-hydroxyphenyl)propanoic acid; 3a, b, 3*R/S*-hydroxy-3-(4'-hydroxyphenyl)propanoic acid; 4a, b, 2*R/S*-hydroxymethyl-2-(2'-hydroxyphenyl)acetic acid; 5a, b, 2*R/S*-hydroxymethyl-2-(3'-hydroxyphenyl)acetic acid; 6a, b, 2*R/S*-hydroxymethyl-2-(4'-hydroxyphenyl)acetic acid; 7, 3-(2',3'-dihydroxyphenyl)propanoic acid; 8, 3-(2',4'-dihydroxyphenyl)propanoic acid; 9, 3-(3',4'-dihydroxyphenyl)propanoic acid; 10, 3-(2',5'-dihydroxyphenyl)propanoic acid; 11, 3-(3',5'-dihydroxyphenyl)propanoic acid; 12a, b, 2*R/S*-hydroxy-3-(2'-hydroxyphenyl)propanoic acid; 13a, b, 2*R/S*-hydroxy-3-(3'-hydroxyphenyl)propanoic acid; 14a, b, 2*R/S*-hydroxy-3-(4'-hydroxyphenyl)propanoic acid; 15a–d, 2*R/S*,3*R/S*-dihydroxy-3-(phenyl)propanoic acid.

structural discrimination, the *S*-enantiomer of ibuprofen is active as a painkiller but the *R*-enantiomer is completely inactive (18).

It is clear that the multiplicity of isomers is a problem, not only of nomenclature but also for the identification of metabolites in biological samples, because without the use of appropriate reference standards, the potential for misidentification is considerable. As noted previously, commercial vendors can often be lax with regard to nomenclature, and it is relatively common for samples to be impure or even incorrectly described. Even with appropriate reference standards, exact identification may not be possible if isomers are not chromatographically resolved. In such cases, extreme caution should be used when describing the metabolite, and using “tentative identification” or “partial identification,” rather than “identification,” is recommended. A feature of the older literature that is still valid today when the

exact structure of a particular isomer is not known is to use the empirical formula (C<sub>9</sub>H<sub>10</sub>O<sub>4</sub>), a nominal molecular mass (182 Da), and a general name—for example, dihydroxy C<sub>6</sub>–C<sub>3</sub> metabolite. In this example, all the compounds in Figure 5 would be encompassed because no positional assignment is provided for the hydroxyl groups. In such circumstances, the metabolic standards initiative proposals of Sumner et al. (16) are of value. From the previous account, it is clear that the multiplicity of synonyms and variations in nomenclature can be confusing for even the seasoned biochemist, and standardization is essential.

## Conclusion

The proposals made in this article will help establish a convenient, clear, and unambiguous nomenclature that is

relevant to studies on microbiota-mediated breakdown of dietary (poly)phenols and their phase II metabolites. In publications, the use of fully characterized accurate names is recommended along with, whenever possible,  $\geq 2$  confirmatory identifiers taken from CAS numbers, InChIKey, and the IUPAC name, in addition to monoisotopic mass, because this allows others to locate the structure in question using recognized online databases. Furthermore, where analytical reference standards are not available and precise identification has not been possible, the metabolite should be described as fully as possible and the assignment downgraded to “tentative identification” or “partial identification.”

The proposed standardization of nomenclature will be of value to researchers because national funding bodies are beginning to require studies make their source data publicly available using open access data repositories. It will also be of value in literature searches for meta-analysis/systematic reviews. In such circumstances, standardized nomenclature will undoubtedly help researchers establish the untapped potential of (poly)phenols to human health.

The authors' responsibilities were as follows—Alan Crozier, CDK and MNC: conceived the study and wrote the first draft of the manuscript; PM, GJM, AR-M, and GW: edited the first draft of the manuscript; and all authors: contributed to the text and read and approved the final manuscript. The authors report no conflicts of interests. Where authors are identified as personnel of the International Agency for Research on Cancer/WHO, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/WHO.

## References

- Barabási A-L, Menichetti G, Loscalzo J. The unmapped chemical complexity of our diet. *Nat Food* 2020;1:33–7.
- Mendonça RD, Carvalho NC, Martin-Moreno JM, Pimenta AM, Lopes ACS, Gea A, Martinez-Gonzalez MA, Bes-Rastrollo M. Total polyphenol intake, polyphenol subtypes and incidence of cardiovascular disease: the SUN cohort study. *Nutr Metab Cardiovasc Dis* 2019;29:69–78. (doi: <https://doi.org/10.1016/j.numecd.2018.09.012>).
- Williamson G, Kay CD, Crozier A. The bioavailability, transport, and bioactivity of dietary flavonoids: a review from a historical perspective. *Compr Rev Food Sci Food Saf* 2018;17:1054–112.
- Pinto P, Santos CN. Worldwide (poly)phenol intake: assessment methods and identified gaps. *Eur J Nutr* 2017;56:1393–408.
- Williamson G. The role of polyphenols in modern nutrition. *Nutr Bull* 2017;42:226–35.
- Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, Mahendrakar T, Williams M, Neumann S, Rocca-Serra P, et al. MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res* 2013;41(Database issue):D781–6.
- Qiu Y-Q. KEGG pathway database. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H, editors. *Encyclopedia of systems biology*. New York: Springer; 2013. p. 1068–9.
- Sud M, Fahy E, Cotter D, Azam K, Vadelu I, Burant C, Edison A, Fiehn O, Higashi R, Nair KS, et al. Metabolomics Workbench: an international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Res* 2016;44:D463–70.
- Maruvada P, Lampe JW, Wishart DS, Barupal D, Chester DN, Dodd D, Djoumbou-Feunang Y, Dorrestein PC, Dragsted LO, Draper J, et al. Perspective: dietary biomarkers of intake and exposure—exploration with omics approaches. *Adv Nutr* 2020;11:200–15.
- Gao Q, Pratico G, Scalbert A, Vergeres G, Kolehmainen M, Manach C, Brennan L, Afman LA, Wishart DS, Andres-Lacueva C, et al. A scheme for a flexible classification of dietary and health biomarkers. *Genes Nutr* 2017;12:34.
- Balentine DA, Dwyer JT, Erdman JW Jr, Ferruzzi MG, Gaine PC, Harnly JM, Kwik-Uribe CL. Recommendations on reporting requirements for flavonoids in research. *Am J Clin Nutr* 2015;101:1113–25.
- Rauter A, Ennis M, Hellwich K-H, Herold B, Horton D, Moss G, Schomburg I. Nomenclature of flavonoids (IUPAC recommendations 2017). *Pure Appl Chem* 2018;90:1629.
- National Academies of Sciences, Engineering, and Medicine E. *Open science by design: realizing a vision for 21st century research*. Washington (DC): National Academies Press.
- Mena P, Bresciani L, Brindani N, Ludwig IA, Pereira-Caro G, Angelino D, Llorach R, Calani L, Brighenti F, Clifford MN, et al. Phenyl- $\gamma$ -valerolactones and phenylvaleric acids, the main colonic metabolites of flavan-3-ols: synthesis, analysis, bioavailability, and bioactivity. *Nat Prod Rep* 2019;36:714–52.
- Gonzalez-Barrio R, Truchado P, Ito H, Espin JC, Tomas-Barberan FA. UV and MS identification of urolithins and nasutins, the bioavailable metabolites of ellagitannins and ellagic acid in different mammals. *J Agric Food Chem* 2011;59:1152–62.
- Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW, Fiehn O, Goodacre R, Griffin JL, et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 2007;3:211–21.
- Jin SJ, Hoppel CL, Tserng KY. Incomplete fatty acid oxidation: the production and epimerisation of 3-hydroxy fatty acids. *J Biol Chem* 1992;267:119–25.
- Neupert W, Brugger R, Euchenhofer C, Brune K, Geisslinge G. Effects of ibuprofen enantiomers and its coenzyme A thioesters on human prostaglandin endoperoxide synthases. *Br J Pharmacol* 2009;122:487–92.