
ChemTastesDB

ChemTastesDB is a database that includes curated information of 2944 molecular tastants. ChemTastesDB constitutes a useful tool for the scientific community to expand the information of molecular tastants, which could assist in the analysis of the relationships between molecular structure and taste, as well as *in silico* (QSAR) studies for taste prediction by means of diverse machine learning approaches.

Molecules are sequestered to one of the five basic tastes (sweet, bitter, umami sour and salty), as well as to other non-basic classes such as: tasteless, non-sweet, multitaste and miscellaneous. ChemTastesDB provides the following information for each molecule: name, PubChem CID, CAS registry number, canonical SMILES string, class taste and the reference to the scientific sources from where data were retrieved. Moreover, the molecular structure in the HyperChem (*.hin*) format of each chemical is provided.

This is version 1.0 of the ChemTastesDB

CONTENT

Four files are provided:

1. "ChemTastesDB_readme.pdf" [datafile], containing a complete description of the ChemTastesDB.

2. "ChemTastesDB_database.xlsx" [datafile], containing information on the individual tastant, that is, molecular ID, name, PubChem CID, CAS registry number, canonical SMILES string, class taste and the reference to the scientific sources from where records were retrieved.

3. "ChemTastesDB_references.xlsx" [datafile], containing the comprehensive list of all scientific references with their extended details. This information is provided as a code including first author surname and year of the reference. For instance, Spillane2002 refers to the following citation: "Spillane, W. J., Feeney, B. G., & Coyle, C. M. (2002). Further studies on the synthesis and tastes of monosubstituted benzenesulfamates. A semi-quantitative structure-taste relationship for the meta-compounds. Food Chemistry, 79(1), 15-22." For molecules retrieved in books and book chapters, the page number was encoded in brackets [pp], while specific codes found in scientific references were set as (cod), for instance:

a) The *4-methylcyclohexylsulfamate* tastant is referred to Spillane2009b_(60A), that is, this compound is labelled as 60A in the paper "Spillane, W. J., & Thompson, E. F. (2009b). The effect on taste upon the introduction of heteroatoms in sulphamates. Food Chemistry, 114(1), 217-225".
b) The *ascorbic acid* tastant is referred as Wong2018_[371], that is, this molecule is available in page 371 of the book "Wong, D. W. S. (2018). Mechanism and theory in food chemistry (Second ed.): Springer".
c) The *(S)-morelid* tastant is referred as Suess2015_(8)/[335], that is, this compound could be found as code number 8 in page 335 of the chapter book "Suess, B., Festring, D., & Hofmann, T. (2015).

Umami compounds and taste enhancers. In J. K. Parker, J. S. Elmore & L. Methven (Eds.), Flavour development, analysis and perception in food and beverages, (pp. 331-351): Woodhead Publishing".

4. "ChemTastesDB_molecules.zip" [datafile], containing the *.hin* Hyperchem file of each tastant optimized by the mechanics force field (MM+). Files are named as the molecular IDs of the ChemTastesDB_database excel file.

Additional details on the content and curation pipeline can be found in the paper: Rojas C., Ballabio D., Pacheco Sarmiento K., Pacheco Jaramillo E., Mendoza M., García F., (2022), ChemTastesDB: A Curated Database of Molecular Tastants, Submitted

---

CONDITIONS

The database is freeware and may be used if proper reference is given to the authors. Preferably refer to the following paper:

Rojas C., Ballabio D., Pacheco Sarmiento K., Pacheco Jaramillo E., Mendoza M., García F., (2022), ChemTastesDB: A Curated Database of Molecular Tastants, Submitted

---

LICENCE

WEB CONTACTS

Write to us for comments, questions or if you find bugs we didn't see!

(Cristian Rojas, Grupo de Investigación en Quimiometría y QSAR, Facultad de Ciencia y Tecnología, Universidad del Azuay, crojasvilla@gmail.com)

HAVE FUN!!!
