

IRACEMA: A Database Management System for Bioactive Compounds Characterized by Brazilian Researchers

XXXI Symposium on Bioinformatics and Computer-Aided Drug Discovery
(BCADD-2025)

Thais A. Lourenço, Renan P. O. Costa, Luis Felipe M. Melo, Uilames A. Ferreira, Eduardo Henrique P. Alves, Luciano P. Oliveira-Filho, Luciana Scotti, Marcus Tullius Scotti, and Gustavo Henrique G. Trossini

October 22nd, 2025

Bioactive Compounds

Definition and Importance

Bioactive compounds are a vast class of chemical substances found in nature or synthesized in a laboratory that can cause biological (pharmacological or toxicological) responses in cells, tissues, or organisms

National Landscape

The research and development of bioactive compounds, both in Brazil and worldwide, has driven a substantial increase in the amount of available chemical and biological data,



Cheminformatics

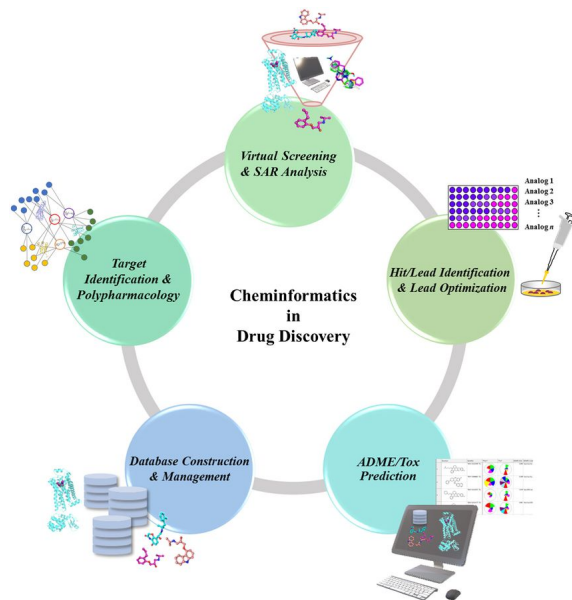


Figure 1. Role of Cheminformatics in the drug discovery process.

Cheminformatics is involved in almost every step of the drug discovery pipeline due to the employment and analysis of available data to translate into valuable knowledge, which can in turn be used as a data for further studies. DOI:10.3389/fphar.2018.00128

“A Cheminformatics is the mixing of those information resources to transform data into information and information into knowledge for the intended purpose of making better decisions faster in the area of drug lead identification and optimization.”

— Frank Brown, 1998

Databases

Definition

A database is a system optimized for efficient data search, query, and management, enabling the analysis of structured data and the relationship between information.



BrNPDB

Brazilian Biodiversity Natural Products Database





IRACEMA (Innovative Research, Analysis and Computational Exploration of Molecules Assembled in Brazil)

- The first database of its kind to focus on synthetic compounds in the country
- Provides an intuitive, user-friendly platform for molecular visualization and analysis by integrating cheminformatics tools
- Give national and international visibility to Brazilian laboratories and facilitate collaboration within research groups



Methodology - Data Curation

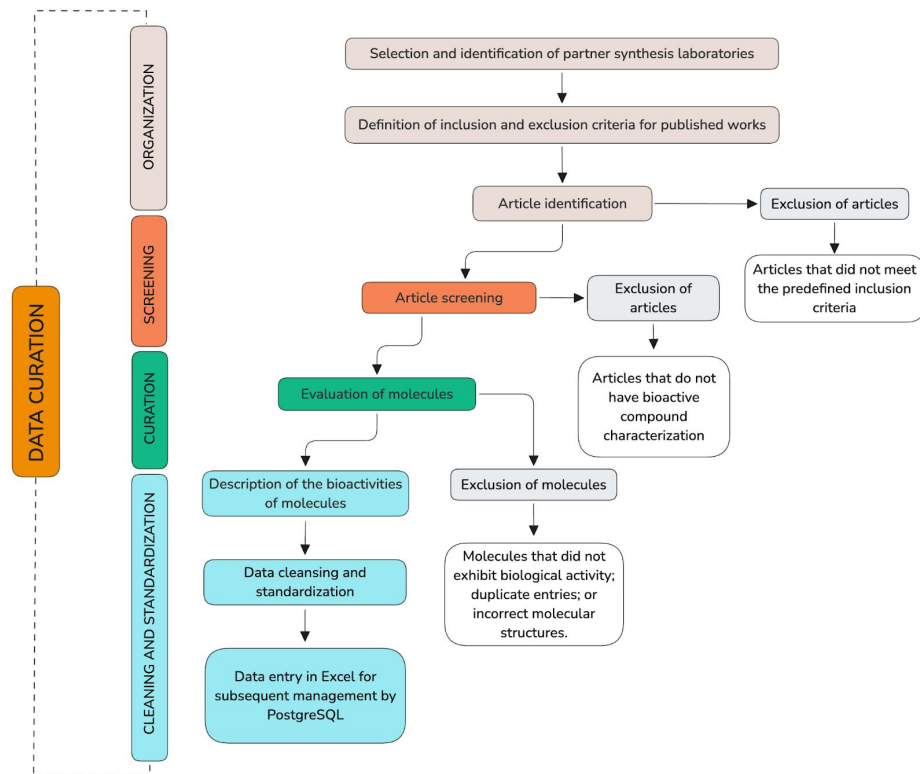


Figure 2. Flowchart of the Data Curation Methodology.

Methodology - System Architecture

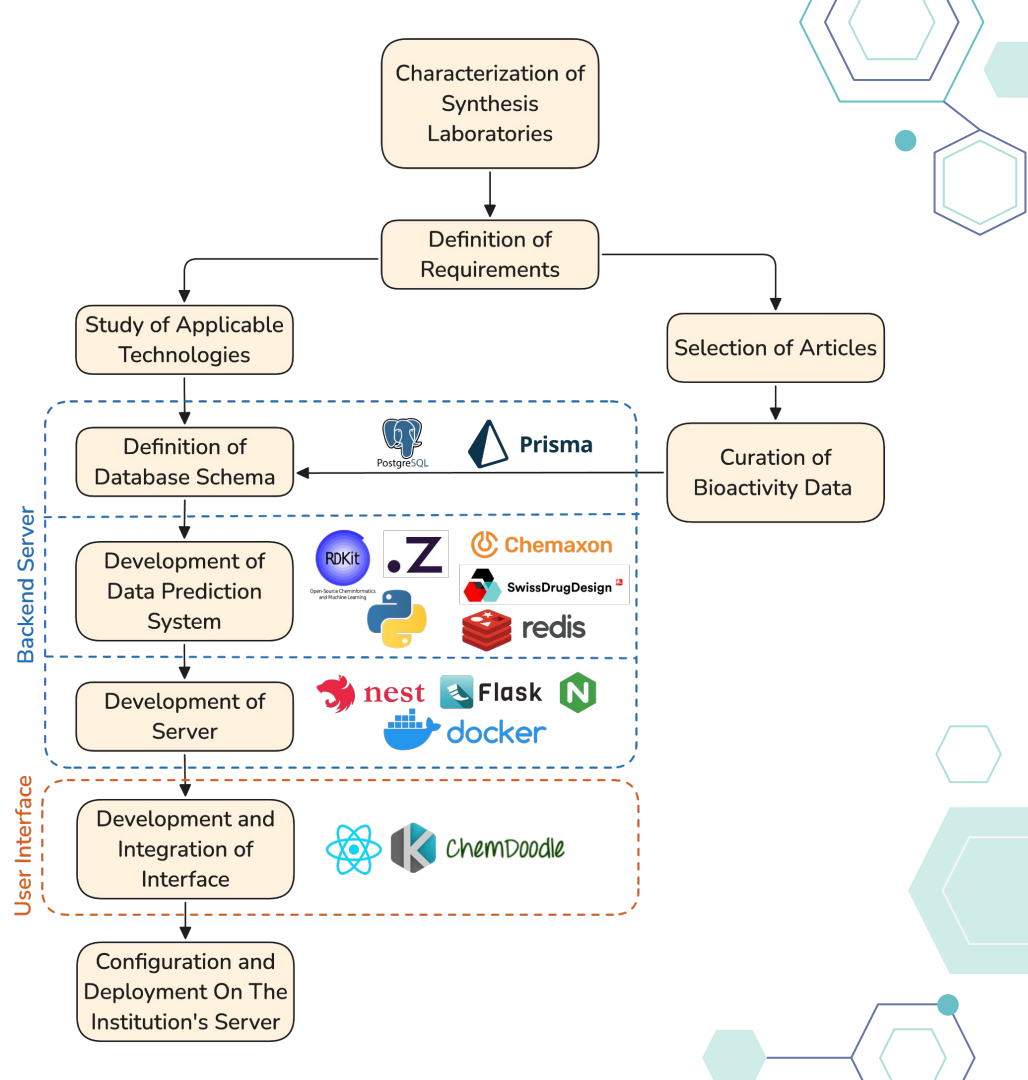


Figure 3. Workflow for compound and bioactivity data curation, alongside the architectural scheme of the IRACEMA system integrating frontend, backend, database, and cheminformatics microservices.

Visual Identity



Figure 4. Moodboard and elements that served as inspiration for visual identity

IRACEMA platform

IRACEMA

Home About Team Contact Us Docs

Welcome to IRACEMA database

Innovative Research, Analysis and Computational Exploration of Molecules Assembled in Brazil

Similarity > Search by Similarity

Draw Molecule

Background image showing chemical structures and a phase diagram (P vs T) with labels for solid, liquid, and vapor phases, and a critical point.

Figure 5. The IRACEMA homepage, with a navigation menu at the top and a search bar in the center.

IRACEMA Search

Figure 6. Interface of the Ketcher applet on IRACEMA. The interface allows for both substructure and molecular similarity searches (using the Tanimoto index).

The image displays the IRACEMA database interface and the Ketcher applet. The IRACEMA website (iracema.fcf.usp.br) features a navigation bar with links to Home, About, Team, Contact Us, and Docs. The main heading reads "Welcome to IRACEMA database" with the tagline "Innovative Research, Analysis and Computational Exploration of Molecules Assembled in Brazil". A search bar is present, and a dropdown menu is open, showing options: Similarity, Substructure, Class, Organism, Target, and Laboratory. The background of the website shows a complex molecular structure. The Ketcher applet, titled "Draw your Molecule", is overlaid on the right. It includes a toolbar with various drawing tools, a central canvas showing a chemical structure of a substituted benzene ring, and a vertical legend on the right listing elements: H, C, N, O, S, P, F, Cl, Br, I, and PT. At the bottom of the applet, there is a search bar with "Similarity" selected, a Tanimoto Coefficient set to 0.3, and buttons for "SEARCH" and "CANCEL".

IRACEMA

Home About Team Contact Us Docs

Welcome to
IRACEMA database

Innovative Research, Analysis and Computational Exploration of Molecules Assembled in Brazil

Similarity
Substructure
Class
Organism
Target
Laboratory

Draw your Molecule

H
C
N
O
S
P
F
Cl
Br
I
PT

Search By: Similarity Tanimoto Coefficient: 0.3

SEARCH CANCEL

IRACEMA Results Page

The screenshot displays the IRACEMA web interface. The top navigation bar includes the IRACEMA logo, a 'Laboratory' dropdown, and a search bar. The main content area is titled 'Search Results' and shows three molecules in a gallery view. Each molecule card includes a chemical structure, a name (AS12/15, IK-01, JN-11), a molecule ID, molecular formula, and weight, along with a 'VIEW MOLECULE' button. A 'FILTERS' button is highlighted, which opens an overlay window. This overlay contains two sections: 'Properties' with sliders for Molecular Weight, Exact Molecular Weight, LogP, and TPSA; and 'Categories' with dropdown menus for Class, Organism, and Target. The overlay also features 'CANCELAR' and 'APLICAR' buttons.

IRACEMA

Laboratory Search by Laboratory

Home > Search Results

Search Results

Sort by Similarity

AS12/15
Molecule ID: 1
Mol Formula: C16H17N3O2S
Mol Weight: 315.398
VIEW MOLECULE

IK-01
Molecule ID: 2
Mol Formula: C21H19N3O3
Mol Weight: 361.401
VIEW MOLECULE

JN-11
Molecule ID: 3
Mol Formula: C21H15N3O4S
Mol Weight: 405.435
VIEW MOLECULE

Filters

Properties

Molecular Weight

Exact Molecular Weight

LogP

TPSA

Categories

Class Select a Class

Organism Select a Organism

Target Select a Target

CANCELAR APLICAR

Figure 7. The results page displays the compounds in either gallery or list format, providing a brief description of each molecule.

IRACEMA Molecule Page

The screenshot displays the IRACEMA Molecule Page for molecule AS12/15. The page is divided into three main sections: a left sidebar for navigation, a central area for the molecule's structure, and a right sidebar for detailed information.

Left Sidebar (Molecule ID):

- Structures
- Names and Identifiers
- Spectral Information
- Biological Activity
- Pharmacokinetics
- Medicinal Chemistry

Central Area (Structures):

2D and 3D Molecule Structure

The 2D structure is shown as a chemical diagram. It features a benzene ring fused to a five-membered ring containing a carbonyl group (C=O) and a hydrazine group (NH-NH-C(=S)-NH₂). A side chain is attached to the five-membered ring, consisting of a double bond (C=C) and a methyl group (CH₃).

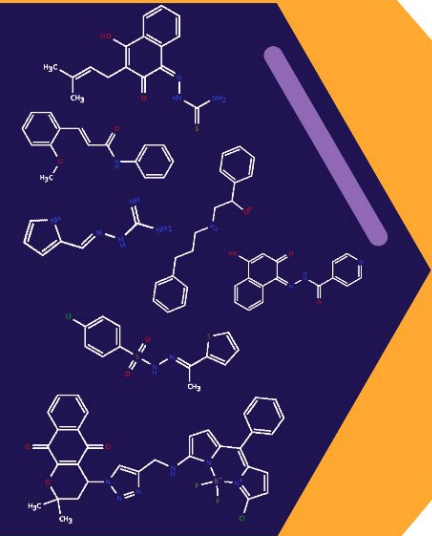
Right Sidebar (AS12/15):

Informations About the Molecule

IRACEMA ID	1
Common Name	AS12/15
IUPAC Name	{{[(1Z)-4-hydroxy-3-(3-methylbut-2-en-1-yl)-2-oxo-1,2-dihydronaphthalen-1-ylidene]amino}thiourea
Molecular Weight	315.398 g/mol
Molecular Formula	C16H17N3O2S
Class	Hydrazone

Figure 8. A detailed view of a selected molecule. The molecule page features 2D and 3D structural visualizations, manually curated biological activities, and predicted molecular properties.

BIOACTIVE MOLECULES



CURATION



PROCESSING

```
{
  "data": {
    "mol": "<rdkit.Chem.rdchem.Mol object at 0x7f44fd2d57b0>",
    "canon_smiles": "O=C(O)c1ccccc1",
    "inchi_str": "InChI=1S/C7H6O2/c8-7(9)6-4-2-1-3-5-6/h1-5H,(H",
    "inchikey": "WPYMKLBDIGXBTP-UHFFFAOYSA-N",
    "logp": 1.3848,
    "tpsa": 37.3,
    "mol_formula": "C7H6O2",
    "exact_mol_wt": 122.036779432,
    "mol_wt": 122.12299999999996,
    "coordinates_2d": "string",
    "coordinates_3d": "string"
  }
}
{
  "results": [
    {
      "smiles": "OC(=O)C1=CC=CC=C1",
      "Physicochemical Properties": {
      },
      "Lipophilicity": {
      },
      "Water Solubility": {
      },
      "Pharmacokinetics": {
      },
      "Druglikeness": {
      }
    }
  ]
}
```



VISUALIZATION

IRACEMA

Home > Search Results

Search Results

FILTERS RESET DOWNLOAD

3h

Molecule ID: IRA-55

Mol Formula: C14H15NO

Mol Weight: 213.280

VIEW MOLECULE >

5a

Molecule ID: IRA-56

Mol Formula: C15H15N

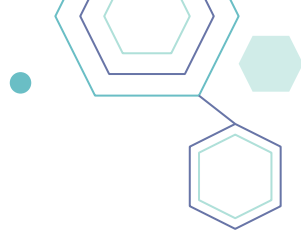
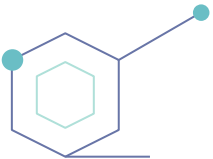
Mol Weight: 209.292

VIEW MOLECULE >

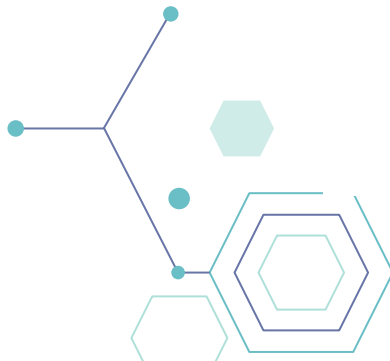
Freely available at <https://iracema.fcf.usp.br/>



IRACEMA
Database



Thank You!



References

- GUAËDAOUI, A. et al. What is a bioactive compound? A combined definition for a preliminary consensus. *International Journal of Nutrition and Food Sciences*, v. 3, n. 3, p. 174–179, 2014.
- SHRINET, K. et al. Bioactive compounds and their future therapeutic applications. In: *Bioactive Compounds: Health Benefits and Potential Applications*. p. 337-362, 2021.
- DE MENEZES, R. P. B.; SCOTTI, L.; SCOTTI, M. T. MACHINE LEARNING APLLIED TO QSAR. *Quimica Nova*, v. 47, n. 7, 2024.
- DING, Y.; XUE, X. Medicinal Chemistry Strategies for the Modification of Bioactive Natural Products. *Molecules*, v. 29, 2024.
- LEACH, A. R.; GILLET, V. J. *An Introduction to Chemoinformatics*. Revised ed. Dordrecht: Springer, 2007.
- ARAÚJO, V. R. et al. O crescimento do Brasil na pesquisa e desenvolvimento de novas moléculas sintéticas com propriedades terapêuticas desejadas: uma revisão. *Revista Brasileira de Ciências Farmacêuticas*, v. 58, n. 1, p. 1-15, jan./mar. 2023.
- SCOTTI, M. T. et al. SistematiX, an Online Web-Based Cheminformatics Tool for Data Management of Secondary Metabolites. *Molecules* 2018, Vol. 23, Page 103, v. 23, n. 1, p. 103, 3 jan. 2018.
- COSTA, R. P. O. et al. The SistematiX Web Portal of Natural Products: An Update. *Journal of Chemical Information and Modeling*, v. 61, n. 6, p. 2516–2522, 28 jun. 2021.

IRACEMA: A Database Management System for Bioactive Compounds Characterized by Brazilian Researchers

XXXI Symposium on Bioinformatics and Computer-Aided Drug Discovery
(BCADD-2025)

Thais A. Lourenço, Renan P. O. Costa, Luis Felipe M. Melo, Uilames A. Ferreira, Eduardo Henrique P. Alves, Luciano P. Oliveira-Filho, Luciana Scotti, Marcus Tullius Scotti, and Gustavo Henrique G. Trossini

October 22nd, 2025