

Version: 0.1a

GU Drug-Pro Toolkit

User Manual

Developed by Aaryesh Deshpande
@ GITAM (Deemed to be) University.

For feedback, support, or more information, contact adeshpan@gitam.in



GU Drug-Pro Toolkit

- Project author: Aaryesh Deshpande
- Project guide: Prof. Sekhar Talluri
- Development Facility: GITAM (Deemed to be) University

Disclaimer

All information and features in this software are provided "as is" without any warranty of any kind, either expressed or implied, including, but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. Under no circumstances shall Aaryesh Deshpande or GITAM be liable for any damages whatsoever, including direct, indirect, incidental, consequential, loss of business profits, or special damages, even if Aaryesh Deshpande or GITAM has been advised of the possibility of such damages. The information and functions included in this software may be updated or changed without prior notice. While we strive to provide accurate and up-to-date information, Aaryesh Deshpande and GITAM cannot guarantee the absolute accuracy of the software's contents or functionalities.

The GU Drug-Pro Toolkit is proprietary software developed by Aaryesh Deshpande. The graphical user interface, underlying code, and integrated algorithms are protected by copyright laws and treaties around the world. All rights reserved. Any modifications to the libraries are listed and documented in accordance with the original licenses. Modifications are solely for the purpose of compatibility and functionality within the GU Drug-Pro Toolkit.

Software Acknowledgments, Trademarks and licensing

All product names, logos, brands, trademarks, and registered trademarks are the property of their respective owners. All company, product, and service names used in this document are for identification purposes only. Use of these names, logos, and brands does not imply endorsement. Portions of the code are derived from open-source software under their respective licenses. The use of the following libraries is acknowledged:

The GU Drug-Pro Toolkit is developed with Python, utilizing open-source libraries under various licenses: PIL (Pillow) and matplotlib for image processing and plotting, NumPy and scikit-learn for numerical computations and machine learning, and TensorFlow for advanced machine learning applications. The user interface is enhanced by Custom_tkinter, with cx_Freeze used for executable creation. Data manipulation is supported by pandas, and chemical data retrieval is powered by pubchempy and rdkit. HTTP requests are handled via the request library, and tkinter provides the standard GUI toolkit. Additionally, export_info and props_ptr are proprietary modules for data exportation and descriptor calculations, respectively, while reportlab facilitates PDF generation.

All libraries are attributed to their respective owners and are incorporated within the licensing agreements. Modifications to any library are documented as per open-source license stipulations. For a complete list of contributors, please refer to the documentation or repositories of each library. For questions regarding third-party licenses, please contact the respective owners. For inquiries related to the GU Drug-Pro Toolkit, please reach out to the support provided in the contact section of this manual.

This software, including all its source code, documentation, and associated materials, is licensed under the MIT License. This means that anyone is free to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the software, provided that the above copyright notice and this permission notice are included in all copies or substantial portions of the software. The MIT License is a permissive free software license originating at the Massachusetts Institute of Technology (MIT). It is designed to promote the widespread use and contribution to the software while protecting the rights of the author. The license ensures that the software remains free and open, enabling innovative and collaborative developments. By using this software, you agree to abide by the terms of the MIT License. For more details, refer to the LICENSE file included with this distribution or visit <https://opensource.org/licenses/MIT>.

Table of Contents

Forward	5
Specifications	6
System Requirements	6
Quick Installation Guide	6
Introduction	7
Understanding ADMET	7
The Role of Physicochemical Properties	7
Introducing the GU Drug-Pro Toolkit	7
Getting Started	8
Main Interface	8
Upload and Input Section	8
Information Display Section	8
Tabs for Molecular Data	9
Menu Bar Functions	9
Features and Functionalities	10
Molecular Visualization	10
Database Linking	10
Molecular Property Calculation	11
Druglikeness Evaluation	11
Pharmacokinetic Prediction	11
Statistical Analysis	11
Exporting Data	11
Properties and Interpretation	12
Names and Identifiers	12
Database links	12
Structural Info	13
Record Data	14
Physicochemical Info	15

Pharmacokinetics	16
Druglikeness	21
Stats.....	24
Bioavailability Spiderweb Chart.....	24
Troubleshooting	25
Frequently Asked Questions FAQs	26
Contact and Support	27

Forward

Welcome to the GU Drug-Pro Toolkit, an integrated suite of tools designed to revolutionize the way chemists, pharmacologists, and researchers analyze molecular structures and properties. This user manual is your comprehensive guide to harnessing the full potential of our software, ensuring a seamless experience as you explore the vast functionalities at your fingertips.

Embarking on this journey, you will discover the intricate details of molecular compounds, predict their pharmacokinetic behaviors, and visualize their structures, providing valuable insights into their drug-likeness and potential as therapeutic agents. Our commitment to innovation and excellence is reflected in every aspect of the toolkit, aiming to empower your research and development endeavors.

We are thrilled to have you on board and are confident that the GU Drug-Pro Toolkit will be a valuable addition to your scientific toolkit.

Happy Discovering!

Specifications

Welcome to a seamless experience of molecular discovery with the GU Drug-Pro Toolkit! Before you delve into the world of molecular analysis with the GU Drug-Pro Toolkit, please ensure that your system meets the following requirements to facilitate a smooth and efficient user experience.

System Requirements

- Operating System: Windows 10 or higher
- Processor: Intel Core i5 or equivalent.
- RAM: 4 GB minimum (8 GB or more recommended for optimal performance).
- Disk Space: At least 550 MB of free space for installation, with additional space for data storage.
- Display: 1280x720 screen resolution or higher.
- Internet Connection: Required for initial setup and to access external databases like PubChem.
- Additional Software: A PDF reader for viewing the manual and reporting tools.

Once your system is prepared, follow the installation instructions detailed in the next section of this page to install the GU Drug-Pro Toolkit. After installation, you can launch the application and begin by inputting a molecular structure using the SMILES notation, chemical name, common name etc. or by uploading a compatible molecular file.

Quick Installation Guide

Prepare for Installation: Close other applications. Ensure system meets the "Getting Started" requirements.

- Run Installer: Double-click the downloaded MSI file for the GU Drug-Pro Toolkit.
- Setup Wizard: Follow the on-screen instructions, clicking "Next" to proceed through the setup wizard.
- Accept License: Read and accept the license agreement terms to continue.
- Choose Folder: Select the installation directory or use the default location provided.
- Install: Click "Install" to begin the installation process.
- Progress: Wait for the progress bar to indicate completion.
- Restart (if prompted): Restart your computer to finalize the installation.
- Launch: Open the GU Drug-Pro Toolkit via the executable file.

If you encounter any issues or have questions, please refer to the Troubleshooting and FAQ sections of this manual.

Introduction

Understanding ADMET

ADMET stands for Absorption, Distribution, Metabolism, Excretion, and Toxicity, encompassing the essential pharmacokinetic and toxicological properties of a compound. These properties determine a drug's overall efficacy and safety profile, predicting how it behaves in a biological system. The ability to accurately assess ADMET characteristics early in the drug development process is crucial for identifying potential issues and streamlining the path to clinical trials.

The Role of Physicochemical Properties

Physicochemical properties of a compound, including molecular weight, lipophilicity (logP), solubility (LogS), and polar surface area (TPSA), are fundamental predictors of a drug's ADMET profile. They influence a drug's absorption rate, bioavailability, and the likelihood of reaching its intended target within the body. Understanding these properties is vital for designing compounds with optimal pharmacokinetic attributes.

Introducing the GU Drug-Pro Toolkit

The GU Drug-Pro Toolkit is a state-of-the-art software designed to bridge the gap between complex molecular analysis and practical pharmacological applications. This toolkit equips researchers with a comprehensive suite of tools to calculate and visualize both physicochemical and ADMET properties with precision. It simplifies the analysis process through an intuitive interface and robust computational algorithms, allowing for:

- **Rapid Calculation:** Quickly computes a wide range of molecular descriptors and properties.
- **ADMET Prediction:** Utilizes advanced algorithms to predict ADMET profiles, aiding in the assessment of drug-like qualities.
- **Data Integration:** Connects with chemical databases for information enrichment and verification.
- **Visualization:** Offers detailed molecular visualizations to aid in structure-activity relationship studies.
- **Exportable Reports:** Generates reports and exports data for further analysis and sharing of findings.

With the GU Drug-Pro Toolkit, researchers can confidently navigate the complex landscape of drug discovery, making informed decisions that contribute to the development of safe and effective therapeutic agents.

Getting Started

Upon starting the application, you will be greeted with a splash screen followed by the primary interface. Familiarize yourself with the layout, which includes a molecule upload area, input fields for chemical identifiers, and multiple tabs for displaying various molecular properties.

Main Interface

Upload and Input Section

- Upload Molecule: Click to browse and upload a .mol or .sdf file.
- SMILES Entry: Type directly to enter a SMILES string, chemical name, or other identifiers.
- Calculate Button: Processes the molecule data and populates the result areas.
- Clear Button: Clears the current session, resetting all fields and outputs.

Main Interface Window

Molecule Input, either by Uploading a molecule file in (.mol or .SDF format) or entering the SMILES, Inchi key, or any chemical name associated with the compound

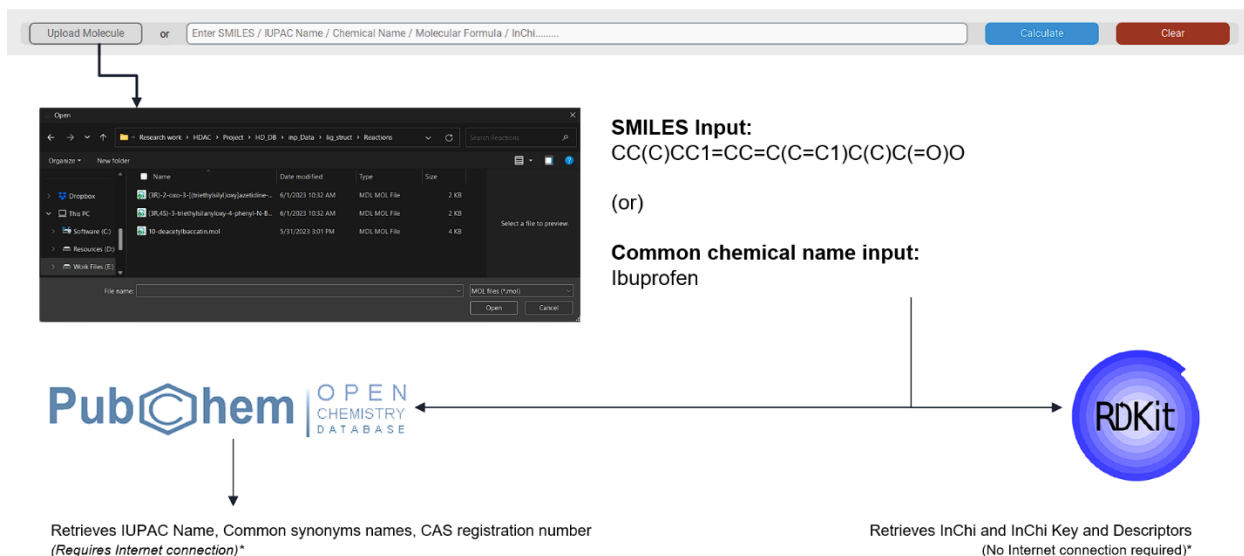
The screenshot shows the 'GUI Drug-Pro Toolkit' window. At the top, there's a menu bar (File, Tools, Settings, Help, About) and a toolbar with 'Upload Molecule', 'or', and 'Calculate' buttons. Below the toolbar is a text input field containing 'Ibuprofen'. To the left of the input field is a chemical structure visualization of Ibuprofen. Below the input field is a 'Names and Identifiers' section containing SMILES, InChI, InChIKey, IUPAC Name, Synonyms, and CAS Number. Below this is a 'Database Search and Cross search' section with buttons for PubChem, ZINC, ChEMBL, ChemSpider, Drug Central, ChEMBL, and Drug Bank. At the bottom, there's a 'Property Tab View' with tabs for 'Structural Info', 'Record Data', 'Physicochemical Info', 'Pharmacokinetics', and 'Toxicity'. The 'Structural Info' tab is active, displaying molecular formula, weight, mass, composition, atom counts, and ring counts.

Annotations in the image point to the following features:

- Display of Molecular Structure**: Points to the chemical structure visualization of Ibuprofen.
- Name and Identifier Info Display window**: Points to the 'Names and Identifiers' section.
- Database Search and Cross search**: Points to the database search buttons.
- Property Tab View**: Points to the 'Structural Info' tab.
- Molecular Property Display Window**: Points to the molecular property data displayed in the 'Structural Info' tab.

Information Display Section

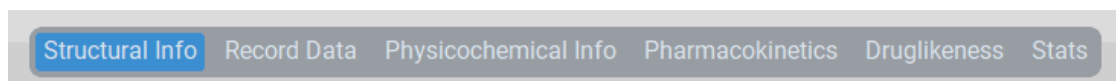
- Molecule Visualization: Presents a graphical depiction of the molecule's structure.
- Names and Identifiers: Lists synonyms, CAS numbers, and other identifiers.
- Database Links: Provides quick access to external databases for more information.



Tabs for Molecular Data

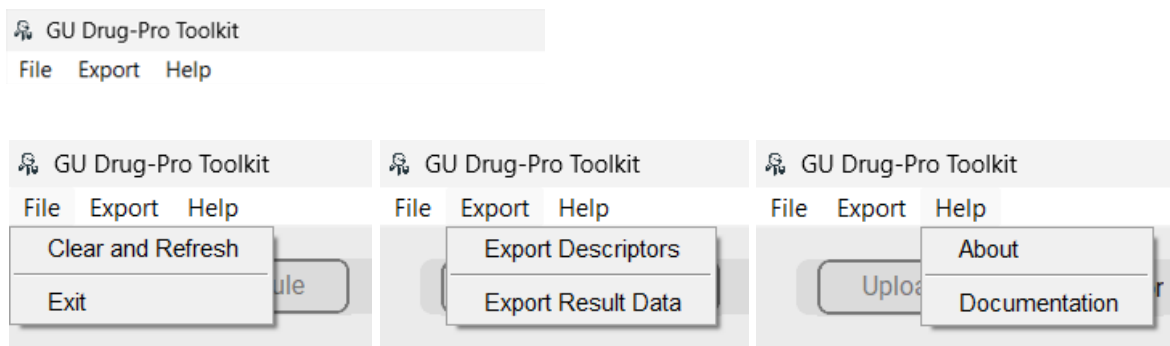
Detailed molecular information is organized across several tabs:

- **Structural Info:** Shows the molecule's structural data, including atom and bond information.
- **Record Data:** Lists the molecular formula, weight, exact mass, and related parameters stored in record databases like PubChem and DrugBank.
- **Physicochemical Info:** Displays calculated properties like LogP, solubility, and pKa etc.
- **Pharmacokinetics:** Shows predicted ADME (Absorption, Distribution, Metabolism, and Excretion) data.
- **Druglikeness:** Assesses the molecule's drug-like properties against various rules and scores, like Lipinski's Rule of Five etc.
- **Stats:** Presents graphical data representation like bioavailability radar charts.



Menu Bar Functions

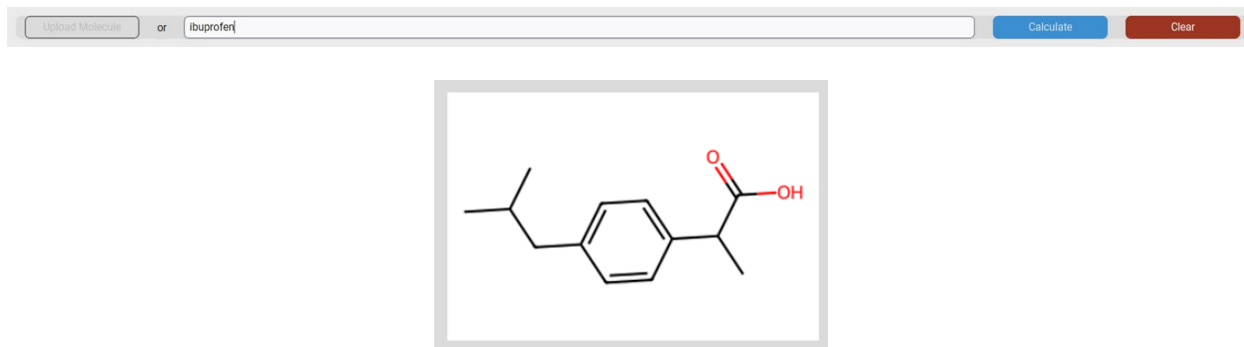
Includes options for file operations, data export which includes descriptors export option and result data export option, and accessing the documentation and about section.



Features and Functionalities

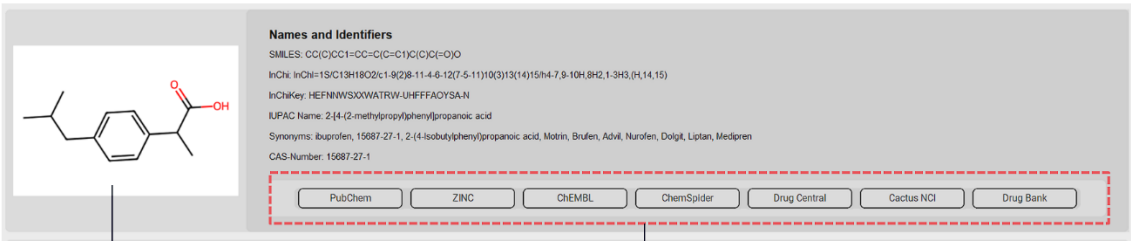
Molecular Visualization

Displays a 2D image of the molecular structure, automatically generated from the uploaded or inputted data.



Database Linking

Provides quick links to databases such as PubChem, ZINC, ChEMBL, ChemSpider, Drug Central, Cactus NCI, Drug Bank for additional information retrieval.



The interface shows the chemical structure of ibuprofen on the left and a list of identifiers on the right. Below the identifiers, there are buttons for various databases: PubChem, ZINC, ChEMBL, ChemSpider, Drug Central, Cactus NCI, and Drug Bank. Arrows point from the structure and the database buttons to the corresponding database screenshots below.

Names and Identifiers

SMILES: CC(C)CC1=CC=C(C(=C1)C(C)C(=O)O)
InChI: InChI=1S/C13H18O2/c1-9(2/8-11-4-6-12)(7-5-11)/0(3)13(14)15/n4-7,9-10H,8H2,1-3H3,(H,14,15)
InChIKey: HEFNNWSXOWATRW-UHFFFAOYSA-N
IUPAC Name: 2-[4-(2-methylpropyl)phenyl]propanoic acid
Synonyms: ibuprofen, 15687-27-1, 2-[4-(4-isobutylphenyl)propanoic acid, Motrin, Brufen, Advil, Nurofen, Dolgit, Liptan, Medipren
CAS-Number: 15687-27-1

PubChem ZINC ChEMBL ChemSpider Drug Central Cactus NCI Drug Bank

Molecular Structure generated based on SMILES to Molecule coordinates constructed by using Rdkit module.

Cross references the input molecule to chemical databases to search for molecular input information and characterized similar search of unknown molecular entry. Opens up custom webpage links in the browser to the result (*Requires internet connection)

PubChem Database ZINC Database Chem Spider Drug Bank

References and retrieves links about information on the molecule or similar/ homologous molecule by cross referencing across seven databases.

Molecular Property Calculation

Calculates basic molecular properties, including molecular weight, bond counts, and atom stereo counts and other properties.

Druglikeness Evaluation

Assesses the molecule's compliance with established druglikeness rules such as Lipinski's Rule of Five and other drug evaluation properties.

Pharmacokinetic Prediction

Predicts ADMET properties using machine learning models to estimate the molecule's pharmacokinetics.

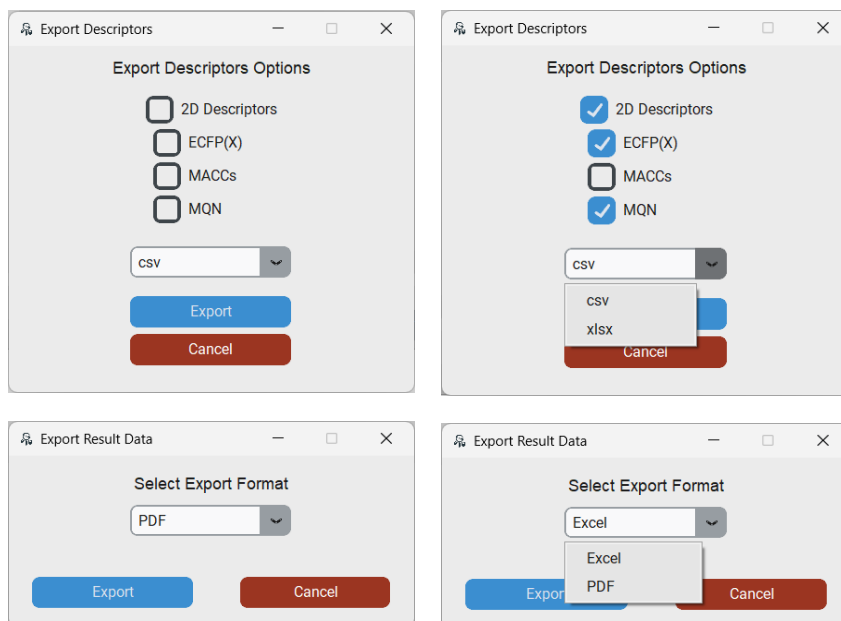
Statistical Analysis

Generates bioavailability radar charts comprising of important physicochemical properties for visual analysis.

Exporting Data

The application allows exporting data in various formats:

- **Export Descriptors:** Allows exporting of molecular descriptors like 2D Descriptors, ECFP, and MQN in CSV or Excel formats.
- **Export Result Data:** Export the analyzed data and graphical representations as PDF or Excel files for documentation.



Properties and Interpretation

Names and Identifiers

- **SMILES** (Simplified Molecular Input Line Entry System): A notation that allows a user to represent a chemical structure in a way that can be used by the computer.
- **InChi** (International Chemical Identifier): A textual identifier that provides a standard way to encode molecular information and facilitate the search for such information in databases.
- **InChiKey**: A condensed digital representation of the InChi, designed for easier searching and indexing of chemical compounds.
- **IUPAC Name**: The systematic name given to a compound according to the rules of the International Union of Pure and Applied Chemistry (IUPAC).
- **Synonyms** (other chemical names): Alternative names for a compound that may be used in different contexts or by different scientific communities.
- **CAS-Number** (Chemical Abstracts Service Number): A unique numerical identifier assigned to every chemical substance described in the open scientific literature, widely used for database searches.

Database links

- **PubChem**: A free database of chemical molecules and their activities against biological assays, providing a comprehensive resource for chemical information and compound structures.
- **ZINC**: A free database for commercially available compounds for virtual screening and chemical biology studies, offering over 35 million chemical records.
- **ChEMBL**: An open large-scale bioactivity database containing information on drug-like bioactive molecules, their biological activities, and the targets these molecules affect, crucial for drug discovery and development research.
- **ChemSpider**: A free chemical structure database providing access to over 58 million molecules, integrating and linking compound data across the web, serving as a comprehensive resource for chemists and researchers.
- **Drug Central**: An online drug compendium that provides up-to-date regulatory and pharmacological information on drugs approved for human use, serving as a critical resource for clinical, pharmaceutical, and academic institutions.
- **UniChem**: A web resource for connecting chemical compound identifiers across multiple databases, facilitating the mapping of chemical structures to their identifiers in other databases, supporting data integration and interoperability.
- **DrugBank**: A comprehensive, authoritative resource for information on drugs and drug targets, combining detailed drug data with comprehensive drug target (sequence, structure, and pathway) information, essential for pharmacology and drug discovery research.

Structural Info

- **Molecular Formula:** The notation representing the types and quantities of atoms in a molecule. Significance: Fundamental for understanding the composition and structure of a compound. No specific range.
- **Molecular Weight:** The sum of the atomic weights of all atoms in a molecule. Significance: Influences solubility, bioavailability, and overall pharmacokinetics. Acceptable Range: Generally, drug-like molecules are considered to have a molecular weight under 500 Da.
- **Exact Molecular Mass:** The calculated mass of a molecule using the exact mass of each isotope of each atom. Significance: Crucial for identifying compounds in mass spectrometry. No specific range.
- **Composition:** The breakdown of a molecule into its constituent elements. Significance: Helps in understanding the elemental makeup and potential reactivity. No specific range.
- **Atom Count:** The total number of atoms present in a molecule. Significance: Reflects the size and complexity of the molecule. No specific range.
- **Heavy Atom Count:** The number of non-hydrogen atoms in a molecule. Significance: Used in assessing molecular size and complexity, important for drug-likeness criteria. Acceptable Range: Drug-like molecules often have between 15 and 35 heavy atoms.
- **Asymmetric Atom Count:** The count of chiral (asymmetric) atoms within a molecule. Significance: Important for understanding stereochemistry and potential biological activity. No specific range.
- **Rotatable Bond Count:** The number of bonds that allow for free rotation within the molecule. Significance: Affects the flexibility of the molecule, which in turn influences binding with biological targets. Acceptable Range: Typically, drug-like molecules have fewer than 10 rotatable bonds.
- **Ring Count:** The total number of ring structures in a molecule. Significance: Rings often contribute to structural stability and can influence biological activity. No specific range.
- **Aromatic Ring Count:** The number of aromatic rings present. Significance: Aromatic rings are common in drugs due to their stability and ability to participate in π - π interactions. No specific range.
- **Hetero Ring Count:** The count of rings containing at least one non-carbon atom. Significance: Heterocycles are prevalent in bioactive molecules, influencing chemical reactivity and biological interactions. No specific range.
- **FSP3 (Fraction of SP3 Hybridized Carbons):** The ratio of SP3 hybridized carbons to the total carbon count. Significance: An indicator of three-dimensionality, higher values suggest greater complexity and potential for specific interactions. Acceptable Range: No fixed range, but higher values (>0.3) are often seen in molecules with specific chiral centers.
- **Hydrogen Bond Donor Count:** The number of atoms that can donate a hydrogen bond. Significance: Influences the molecule's ability to bind to biological targets. Acceptable Range: For drug-like properties, typically less than 5.
- **Hydrogen Bond Acceptor Count:** The number of atoms that can accept a hydrogen bond. Significance: Important for molecular recognition and binding. Acceptable Range: Typically, less than 10 for drug-like molecules.
- **Formal Charge:** The electrical charge of the molecule. Significance: Affects solubility, permeability, and interaction with charged biological molecules. Acceptable Range: Usually close to neutral for most drug-like molecules.
- **Topological Polar Surface Area (TPSA):** The surface area occupied by oxygen and nitrogen atoms including their attached hydrogens. Significance: Provides insight into a compound's ability to interact with the aqueous environment and permeate membranes. Acceptable Range: Drug-like molecules typically have a TPSA less than 140 \AA^2 for optimal oral bioavailability.

Record Data

- **Monoisotopic Mass:** The mass of a molecule calculated using the most abundant isotope of each element. Significance: Vital for precise mass spectrometry analysis. No specific range.
- **Isotope Atom Count:** Number of atoms of each isotope present in a molecule. Significance: Important for studying isotopically labeled compounds and their behavior. No specific range.
- **Atom Stereo:** The spatial arrangement of atoms around a chiral center. Significance: Critical for determining the 3D orientation of molecules, affecting biological activity. No specific range.
- **Bond Stereo:** Configuration of double bonds or stereocenters involving bonds. Significance: Influences molecular shape and interactions with biological targets. No specific range.
- **Covalent Unit:** Number of distinct covalently bonded entities within a molecule. Significance: Reflects the complexity and potential modular nature of compounds. No specific range.
- **XLogP:** A prediction of the compound's octanol-water partition coefficient. Significance: Estimates lipophilicity, influencing absorption and distribution. Acceptable Range: Typically, drug-like molecules have XLogP values between -0.4 and 5.6.
- **Volume3D:** The spatial volume occupied by a molecule. Significance: Affects how the molecule fits into binding pockets on targets. No specific range.
- **Charge:** Overall electrical charge of the molecule. Significance: Influences solubility and interaction with charged biological molecules. Acceptable Range: Typically, close to neutral for drug-like molecules.
- **Complexity (Index):** A measure of the molecular structure's complexity. Significance: Higher complexity may indicate a more specific mode of action. No specific range.
- **Coordinate Type:** The nature of the coordinates used to describe the molecular structure (e.g., 2D, 3D). Significance: Indicates the dimensionality of the molecular representation. No specific range.
- **Defined Atom Stereo Count:** Number of atoms with a defined stereochemistry. Significance: Reflects the specificity of molecular interactions. No specific range.
- **Defined Bond Stereo Count:** Number of bonds with a defined stereochemistry. Significance: Important for understanding precise molecular geometry. No specific range.
- **Undefined Atom Stereo Count:** Number of atoms with undefined stereochemistry. Significance: Indicates potential isomers or stereoisomers. No specific range.
- **Undefined Bond Stereo Count:** Number of bonds with undefined stereochemistry. Significance: Suggests flexibility or ambiguity in molecular structure. No specific range.
- **Effective Rotor Count 3D:** Number of rotatable bonds affecting 3D conformational flexibility. Significance: Impacts molecular dynamics and binding properties. No specific range.
- **Conformer ID 3D:** Unique identifier for a specific three-dimensional conformer. Significance: Allows distinction between different spatial arrangements of the same molecule. No specific range.
- **Conformer RMSD 3D:** Root-mean square deviation of atomic positions between conformers. Significance: Measures the similarity between different 3D structures of a molecule. No specific range.
- **Feature Self Overlap 3D:** Measure of the overlap of molecular features in 3D space. Significance: Can indicate the compactness or spread of functional groups. No specific range.
- **MMFF94 Energy 3D:** The energy of a molecule calculated using the MMFF94 force field. Significance: Provides insight into the stability of a conformer. No specific range.
- **Multipoles 3D:** The distribution of charge across the molecule represented by multipoles. Significance: Affects the electrostatic interactions with biological targets. No specific range.

- **Pharmacophore Features 3D:** Defined elements within a molecule that are believed to interact with a specific biological target. Significance: Essential for understanding and predicting molecular binding and activity. No specific range.
- **Shape Fingerprint 3D:** A digital representation of the molecular shape used for comparison. Significance: Useful for similarity searching and matching molecules by shape. No specific range.
- **Shape Self Overlap 3D:** Measure of how much a molecule's shape overlaps with itself in 3D space. Significance: Indicates conformational consistency and stability. No specific range.

Physicochemical Info

- **LogS (ESOL):** Estimated solubility in water, expressed as a logarithm. Significance: Predicts how easily a compound dissolves in water, critical for bioavailability. Acceptable Range: Generally, higher values indicate better solubility.
- **Molar Solubility (mol/l):** The amount of substance that can dissolve in a liter of solvent to form a saturated solution. Significance: Directly influences the dosage form and delivery method of a drug. Acceptable Range: Depends on drug application; higher solubility is usually preferred for oral drugs.
- **Solubility (mg/ml):** The mass of compound that can be dissolved in a milliliter of solvent. Significance: Important for formulation development and predicting bioavailability. Acceptable Range: Varied; higher values indicate better solubility, enhancing oral uptake.
- **Solubility Class:** Categorization of a compound's solubility based on defined thresholds (e.g., very soluble, soluble, slightly soluble). Significance: Guides formulation strategies and predicts challenges in drug delivery. No specific range but classified as per USP solubility definitions.
- **WLogP:** The predicted partition coefficient between water and octanol. Significance: Reflects compound's lipophilicity, influencing absorption and distribution. Acceptable Range: Drug-like molecules typically possess WLogP values between -0.4 and 5.6 for optimal balance between solubility and permeability.
- **LogD:** The distribution coefficient at a specific pH, representing the compound's ionization state. Significance: Affects solubility and permeability across biological membranes. Acceptable Range: Like WLogP, balanced values are preferred for drug candidates.
- **Nature:** Acidic or Basic with **pKa:** The acid/base nature of a compound and the pH at which it is 50% ionized. Significance: Influences solubility, absorption, and stability under physiological conditions. Acceptable Range: Depends on the target therapeutic window and formulation requirements.

Pharmacokinetics

Ames Mutagenicity (AMES Toxicity)

- **Definition & Significance:** The Ames test assesses a compound's mutagenicity, which is closely linked to carcinogenicity. This assay is widely used for testing the mutagenic potential of compounds, with a significant impact on evaluating their safety profile.
- **Optimal Value Range & Tags:** A compound is considered non-mutagenic (negative, hence positive for use) if it falls within Category 0 (probability of being toxic from 0 to 0.3), indicating excellent safety. Compounds in Category 1 (probability of being toxic from 0.7 to 1.0) are considered mutagenic (positive for Ames test, hence negative for use), indicating potential safety concerns.

Blood-Brain Barrier Penetration (BBB Penetration)

- **Definition & Significance:** BBB penetration measures a drug's ability to cross the blood-brain barrier, crucial for CNS-targeted therapies and avoiding unwanted CNS side effects for peripherally acting drugs.
- **Optimal Value Range & Tags:** Compounds with a probability of BBB penetration from 0 to 0.3 are considered to have low/no penetration (excellent, green), suitable for peripherally targeted drugs. High penetration, indicated by a probability of 0.7 to 1.0, is needed for CNS-targeted drugs but could be concerning for drugs intended to act outside the CNS.

Bioavailability (Ma) (F30%)

- **Definition & Significance:** Bioavailability reflects the proportion of an orally administered drug that reaches systemic circulation, crucial for its therapeutic effect.
- **Optimal Value Range & Tags for F30%:** For F30%, a probability of 0 to 0.3 for being F30%+ (bioavailability < 30%) indicates excellent bioavailability. Probabilities from 0.7 to 1.0 (++) for both F20% and F30%) suggest poor bioavailability, which could impact the drug's effectiveness.

CYP Inhibition (CYP1A2, CYP2C19, CYP2C9, CYP2D6, CYP3A4)

- **Definition & Significance:** Cytochrome P450 (CYP) enzymes play a vital role in drug metabolism. Inhibiting these enzymes can lead to interactions and adverse effects due to altered drug clearance rates.
- **Optimal Value Range & Tags:** For all CYP inhibitors mentioned, a probability of being an inhibitor from 0 to 0.3 suggests non-inhibition (excellent, green), indicating a lower risk of drug-drug interactions. A probability from 0.7 to 1.0 indicates significant inhibition (poor, red), raising concerns about potential interactions.

Drug-Induced Liver Injury (DILI)

- **Definition & Significance:** DILI is a critical safety concern, representing the leading cause of drug withdrawal from the market. It signifies the hepatotoxic potential of a drug.
- **Optimal Value Range & Tags:** A probability of DILI from 0 to 0.3 is considered excellent (green), indicating low risk for liver injury. Probabilities from 0.7 to 1.0 represent a high risk of DILI (poor, red), marking the compound as potentially dangerous to liver health.

Human Intestinal Absorption (HIA)

- **Definition & Significance:** HIA estimates the extent to which a drug is absorbed from the intestine, crucial for oral drug efficacy.
- **Optimal Value Range & Tags:** A probability range of 0 to 0.3 for being HIA+ (poor absorption, less than 30%) indicates excellent absorption (green), crucial for oral drugs' effectiveness. Probabilities from 0.7 to 1.0 (++) on the other hand, suggest poor absorption (red), potentially undermining the drug's therapeutic utility.

NR-AR-LBD: Nuclear Receptor AR Ligand Binding Domain

- **Definition & Significance:** The ligand-binding domain (LBD) of the Androgen Receptor (AR) is crucial for binding androgens, which influences AR-dependent diseases like prostate cancer. It's important for understanding how chemicals interact with this domain, potentially affecting hormonal balance and cancer progression.
- **Optimal Value Range & Tags:** A probability of 0 to 0.3 for being active indicates excellent safety (green), implying low or no affinity for the AR LBD, reducing the risk of disrupting hormonal balance. Probabilities from 0.7 to 1.0 indicate strong affinity (poor, red), raising concerns about potential hormonal disruption.

NR-AR: Nuclear Receptor AR

- **Definition & Significance:** The Nuclear Receptor AR plays a critical role in the development and maintenance of male characteristics. Interactions with this receptor can influence prostate cancer and other androgen-related diseases.
- **Optimal Value Range & Tags:** A probability of 0 to 0.3 for being active against NR-AR indicates good safety (green), suggesting minimal interaction with the AR that could disrupt hormonal functions. A high probability (0.7-1.0) suggests strong interaction (poor, red), potentially leading to adverse effects related to hormonal imbalances.

NR-AhR: Nuclear Receptor AhR

- **Definition & Significance:** The Aryl hydrocarbon Receptor (AhR) mediates responses to environmental pollutants and interacts with other signaling pathways. Its activation has implications for detoxification processes and potentially adverse cellular responses.
- **Optimal Value Range & Tags:** Probabilities of 0 to 0.3 for being activators indicate low likelihood of adverse activation (excellent, green), while probabilities from 0.7 to 1.0 signify high potential for problematic interactions (poor, red).

NR-Aromatase: Nuclear Receptor Aromatase

- **Definition & Significance:** Aromatase is key for converting androgens to estrogens, maintaining hormonal balance. Inhibiting aromatase can lead to disrupted steroid hormone levels, affecting reproductive health and hormone-dependent cancer progression.
- **Optimal Value Range & Tags:** Probabilities of 0 to 0.3 for being active aromatase inhibitors suggest a lower risk of disrupting hormonal balance (excellent, green). High probabilities (0.7-1.0) indicate significant inhibition risk (poor, red), potentially leading to hormonal imbalances.

NR-ER-LBD: Nuclear Receptor ER Ligand Binding Domain

- **Definition & Significance:** The ER LBD is vital for estrogen binding, impacting functions in development, reproductive health, and ER-dependent cancers. Chemical interactions with this domain can influence estrogenic activity and related health outcomes.
- **Optimal Value Range & Tags:** Probabilities of 0 to 0.3 indicate low affinity for ER LBD (excellent, green), suggesting minimal estrogenic activity disruption. High probabilities (0.7-1.0) imply strong affinity (poor, red), potentially affecting hormonal balance and cancer risks.

NR-ER: Nuclear Receptor ER

- **Definition & Significance:** The Estrogen Receptor (ER) is crucial for mediating estrogenic effects on growth, reproduction, and health. Compounds interacting with ER can have significant implications for hormonal balance and the development of estrogen-dependent diseases.
- **Optimal Value Range & Tags:** A probability of 0 to 0.3 for being active suggests minimal interaction (excellent, green), indicating lower risk of hormonal disruption. Probabilities of 0.7 to 1.0 indicate a higher interaction risk (poor, red), with potential estrogenic or anti-estrogenic effects that could influence health outcomes.

NR-PPAR-gamma: Nuclear Receptor PPAR-gamma

- Definition & Significance: PPAR-gamma plays a role in lipid metabolism and glucose homeostasis. It's targeted for diabetes treatment due to its role in regulating insulin sensitivity. Compounds interacting with PPAR-gamma can influence metabolic processes and have therapeutic potential.
- Optimal Value Range & Tags: Probabilities of 0 to 0.3 for being activators indicate a low likelihood of metabolic disruption (excellent, green). High probabilities (0.7-1.0) suggest significant interaction (poor, red), with implications for metabolic diseases and therapeutic interventions.

PAMPA_NCATS: Parallel Artificial Membrane Permeability Assay (NCATS)

- Definition & Significance: PAMPA assesses a compound's permeability across a lipid bilayer, predicting oral bioavailability and absorption. It's crucial for screening compounds for potential oral administration.
- Optimal Value Range & Tags: High permeability is indicated by Papp values greater than 10×10^{-6} cm/s (Excellent, Green), suggesting good potential for oral bioavailability. Moderate permeability, with values between 2×10^{-6} and 10×10^{-6} cm/s (Acceptable, Yellow), may require optimization for improved absorption. Low permeability, less than 2×10^{-6} cm/s (Poor, Red), signals challenges in achieving sufficient oral bioavailability without significant modification or alternative delivery methods.

Pgp_Broccatelli: P-glycoprotein Substrate (Broccatelli)

- Definition & Significance: P-glycoprotein (Pgp) is a critical efflux transporter affecting drug absorption and brain penetration. Being a Pgp substrate can influence a drug's distribution and elimination.
- Optimal Value Range & Tags: A probability of being a Pgp substrate below 0.3 suggests minimal interaction with Pgp, indicating safer pharmacokinetic profiles (Excellent, Green). Conversely, probabilities exceeding 0.7 indicate strong substrate potential, suggesting a higher likelihood of Pgp-mediated drug interactions and reduced bioavailability (Poor, Red).

SR-ARE: Stress Response - Antioxidant Response Element

- Definition & Significance: ARE signaling is key for managing oxidative stress, linked to diseases like cancer and neurodegeneration. Compounds activating ARE may offer protective benefits against oxidative damage.
- Optimal Value Range & Tags: Probabilities of 0 to 0.3 for being activators suggest beneficial potential for managing oxidative stress (excellent, green). High probabilities (0.7-1.0) could indicate strong activators (poor, red), with implications for therapeutic use or toxicity.

SR-ATAD5: Stress Response - ATAD5

- Definition & Significance: ATAD5 plays a role in DNA damage response, crucial for genomic stability. Compounds influencing ATAD5 activity may affect cell cycle regulation and DNA repair mechanisms.
- Optimal Value Range & Tags: Probabilities of 0 to 0.3 for being activators suggest minimal disruption of DNA repair processes (excellent, green), while high probabilities (0.7-1.0) indicate a potential for increased DNA damage risk (poor, red).

SR-HSE: Stress Response - Heat Shock Response Element

- Definition & Significance: HSE mediates the heat shock response, essential for cellular protection against stress. Activating this pathway can confer resilience to thermal and other stresses, influencing cell survival.
- Optimal Value Range & Tags: Probabilities of 0 to 0.3 for activation indicate low stress response activation (excellent, green), beneficial under normal conditions. Probabilities of 0.7 to 1.0 suggest strong activation (poor, red), which may be protective or potentially disruptive under certain conditions.

SR-MMP: Stress Response - Mitochondrial Membrane Permeability

- Definition & Significance: MMP changes are indicative of mitochondrial health and can signal apoptosis. Monitoring MMP can help understand a compound's impact on cell viability and energy production.
- Optimal Value Range & Tags: Probabilities of 0 to 0.3 for affecting MMP suggest minimal impact (excellent, green), indicating preservation of mitochondrial function. High probabilities (0.7-1.0) imply significant effects (poor, red), potentially indicating cytotoxicity.

SR-p53: Stress Response - p53 Response Element

- Definition & Significance: p53 is a tumor suppressor protein activated by DNA damage. Its activation leads to cell cycle arrest, DNA repair, or apoptosis, playing a critical role in preventing cancer development.
- Optimal Value Range & Tags: Probabilities of 0 to 0.3 for activation indicate low DNA damage risk (excellent, green), preserving genomic stability. High probabilities (0.7-1.0) suggest strong activation (poor, red), indicating potential DNA damage or stress response activation.

Skin Reaction

- Definition & Significance: Skin reactions encompass a variety of responses from the skin upon exposure to a compound, ranging from irritation and redness to severe allergic reactions. Understanding a compound's potential to cause skin reactions is crucial for safety assessments, especially for topically applied drugs or chemicals.
- Optimal Value Range & Tags: Traditionally quantified with a numerical range where lower the better is the case but categorized qualitatively. Negative reactions (e.g., irritation, dermatitis) would warrant a negative tag, while a lack of adverse skin reactions would be tagged positively.

hERG Inhibition

- Definition & Significance: Inhibition of the human Ether-à-go-go-Related Gene (hERG) potassium channels can lead to long QT syndrome, arrhythmia, and potentially fatal cardiac events. Assessing hERG inhibition is a critical part of the safety evaluation for new drugs.
- Optimal Value Range & Tags: $IC_{50} > 10 \mu M$ (indicating weak or no inhibition) would typically be considered safe (green tag), whereas $IC_{50} < 10 \mu M$ (strong inhibition) suggests a risk of cardiotoxicity (red tag).

Caco-2 Permeability (Wang)

- Definition & Significance: Caco-2 cell permeability assay predicts intestinal absorption of drugs, with higher permeability indicating potential for better absorption. This assay is critical for oral drug development.
- Optimal Value Range & Tags: Permeability $> -5.15 \log \text{ cm/s}$ indicates good permeability (excellent, green), while values $< -5.15 \log \text{ cm/s}$ suggest poor permeability (poor, red).

Clearance (Hepatocyte and Microsome, AstraZeneca)

- Definition & Significance: Hepatocyte clearance estimates how quickly a drug is metabolized by liver cells, while microsomal clearance focuses on metabolism within a subset of liver cells' components. Both are crucial for understanding a drug's metabolism and potential interactions.
- Optimal Value Range & Tags: Clearance rates $> 15 \text{ mL/min/kg}$ (high clearance) suggest rapid elimination (green for hepatocyte and microsome), while rates $< 5 \text{ mL/min/kg}$ (low clearance) indicate slow elimination, potentially leading to accumulation and toxicity (red).

Half-Life (Obach)

- Definition & Significance: The half-life of a drug indicates how long it takes for its plasma concentration to reduce by half, influencing dosing frequency and potential for accumulation.
- Optimal Value Range & Tags: A half-life < 3 hours is often considered short, leading to a more frequent dosing schedule (green if desirable for dosing flexibility), while a half-life > 24 hours is long, which may be beneficial for once-daily dosing but risks accumulation (yellow to red, depending on context).

Hydration Free Energy (FreeSolv)

- Definition & Significance: Hydration free energy relates to a compound's solubility and interaction with water, influencing its solubility, absorption, and distribution.
- Optimal Value Range & Tags: Generally, more negative values indicate better solubility and interaction with water (green), whereas positive values suggest poor solubility (red).

LD50 (Zhu)

- Definition & Significance: LD50, the lethal dose required to kill 50% of a test population, is a measure of acute toxicity. Lower LD50 values indicate higher toxicity.
- Optimal Value Range & Tags: LD50 values > 2000 mg/kg are considered relatively safe (green), while values < 50 mg/kg indicate high toxicity (red).

Lipophilicity (AstraZeneca)

- Definition & Significance: Lipophilicity measures a compound's ability to dissolve in fats, oils, and non-polar solvents, affecting its absorption, distribution, and interaction with biological membranes.
- Optimal Value Range & Tags: LogP values between 0 to 3 are considered optimal for good balance between solubility and permeability (green), while values > 5 suggest potential issues with solubility and bioavailability (red).

Plasma Protein Binding (AstraZeneca)

- Definition & Significance: The degree to which drugs bind to plasma proteins affects their free (active) concentration, distribution, and elimination.
- Optimal Value Range & Tags: < 90% binding is generally considered acceptable, with more drug available in free form (green), while > 90% suggests high binding that may limit drug availability (red).

Solubility (AqSolDB)

- Definition & Significance: Solubility is crucial for drug absorption and bioavailability, with higher solubility generally improving oral drug performance.
- Optimal Value Range & Tags: Solubility > 100 mg/mL is excellent (green), indicating high solubility, while < 1 mg/mL indicates poor solubility (red), potentially hindering absorption.

Volume of Distribution (VDss, Lombardo)

- Definition & Significance: VDss describes the distribution of a drug throughout the body's compartments, influencing its concentration at the site of action and required dosage.
- Optimal Value Range & Tags: VDss in the range of 0.04-20 L/kg is considered ideal, suggesting effective distribution without excessive tissue sequestration (green). Values outside this range may indicate limited distribution or excessive tissue accumulation (yellow to red).

Druglikeness

Synthetic Accessibility Score

- **Definition & Significance:** Synthetic Accessibility Score (SAS) quantifies the ease of synthesizing a compound, crucial for drug discovery and development. It evaluates synthetic complexity, guiding medicinal chemists towards more feasible compound designs. Numerical Range: Typically measured on a scale from 1 to 10.
- **Optimal Range & Criteria:** A lower SAS score (e.g., below 6) is generally preferred, suggesting compounds that are easier to synthesize (Excellent, Green). Higher scores (e.g., above 8) may indicate compounds with challenging synthetic routes (Poor, Red), potentially slowing down drug development efforts.

QED Score

- **Definition & Significance:** The Quantitative Estimate of Drug-likeness (QED) Score assesses a compound's drug-likeness based on various molecular properties. QED, a measure of drug-likeness based on desirability, integrates eight drug-related properties such as molecular weight, log P, number of hydrogen bond donors and acceptors, TPSA, number of rotatable bonds, the number of aromatic rings, and undesirable functional group alerts. Utilizing average descriptor weights, QED score calculation involves taking the geometric mean of individual desirability functions.
- **Optimal Range & Criteria:** Results interpretation suggests that compounds with a mean QED of 0.67 are deemed attractive, those with 0.49 are considered unattractive, and those below 0.34 are seen as excessively complex. Empirically, a QED score exceeding 0.67 is rated as excellent (green), while scores at or below 0.67 are deemed poor (red). Numerical Range: Typically ranges from 0 to 1, with higher scores indicating greater drug-likeness.

PAINS Alerts

- **Definition & Significance:** PAINS (Pan-Assay Interference Compounds) Alerts identify substructures commonly associated with assay interference or false-positive results in high-throughput screening assays. Avoiding compounds with PAINS alerts helps researchers focus on genuine hits during drug discovery. Numerical Range: Not applicable; typically represented as a binary result (presence or absence of PAINS alerts).
- **Optimal Range & Criteria:** Compounds without PAINS alerts are preferred (Excellent, Green), as they are less likely to interfere with screening assays. Compounds containing PAINS alerts (Poor, Red) may require further evaluation or structural modification to mitigate assay interference.

Lead Likeness

- **Definition & Significance:** Lead Likeness assesses a compound's similarity to known drug leads, guiding the selection of promising candidates in drug discovery. It considers physicochemical properties and structural features associated with successful drug development. Numerical Range: Often represented qualitatively as "lead-like" or "non-lead-like" based on specific criteria.
- **Optimal Range & Criteria:** Compounds classified as lead-like are preferred (Excellent, Green), suggesting they possess properties conducive to successful drug development. Non-lead-like compounds (Poor, Red) may require significant optimization to meet lead-like criteria. Criteria: $250 \leq MW \leq 350$, $XLOGP \leq 3.5$, Num. rotatable bonds ≤ 7

Lipinski's Rule

- **Definition & Significance:** Lipinski's Rule, also known as the Rule of Five, evaluates a compound's drug-likeness based on four key physicochemical properties: molecular weight (MW), lipophilicity (expressed as log P), number of hydrogen bond donors (NHBD), and number of hydrogen bond acceptors (NHBA). It helps identify compounds with favorable pharmacokinetic properties and guides drug discovery efforts by predicting a compound's likelihood of success in oral administration.
- **Criteria:** A compound is considered more likely to be orally bioavailable if it adheres to Lipinski's Rule, which states that a compound is drug-like if it has no more than 5 hydrogen bond donors, no more than 10 hydrogen bond acceptors, a molecular weight under 500 Daltons, and a log P value less than 5. Content: $MW \leq 500$; $\log P \leq 5$; $H_{acc} \leq 10$; $H_{don} \leq 5$

Ghose's Rule

- **Definition & Significance:** Ghose's Rule is a guideline used in drug discovery to evaluate a compound's drug-likeness based on its physicochemical properties. It assesses the distribution of molecular weight, log P, number of atoms, and molar refractivity to predict a compound's likelihood of being orally active. Ghose's Rule aims to identify compounds with optimal physicochemical properties for drug development.
- **Criteria:** Ghose's Rule suggests that orally active compounds typically have molecular weights between 160 and 480 Daltons, log P values between -0.4 and 5.6, atom counts between 20 and 70, and molar refractivity values between 40 and 130. Content: $160 \leq MW \leq 480$, $-0.4 \leq WLOGP \leq 5.6$, $40 \leq MR \leq 130$, $20 \leq \text{atoms} \leq 70$

Veber's Rule

- **Definition & Significance:** Veber's Rule is a rule of thumb used in drug discovery to assess the likelihood of a compound being orally bioavailable based on its rotatable bond count and polar surface area. It helps prioritize compounds with favorable pharmacokinetic properties for further development.
- **Criteria:** According to Veber's Rule, compounds with fewer than 10 rotatable bonds and a polar surface area less than 140 \AA^2 are more likely to be orally bioavailable. Content: Rotatable bonds ≤ 10 , TPSA $\leq 140 \text{ \AA}^2$

Egan's Rule

- **Definition & Significance:** Egan's Rule is a guideline used in drug discovery to evaluate a compound's likelihood of being orally bioavailable based on its physicochemical properties. It helps prioritize compounds with optimal characteristics for absorption and distribution.
- **Criteria:** Egan's Rule suggests that bioavailable compounds typically have log P values less than or equal to 5.88, and total polar surface area less than 131.6 \AA^2 . Content: $WLogP \leq 5.88$, $TPSA \leq 131.6 \text{ \AA}^2$

Muegge's Rule

- Definition & Significance: Muegge's Rule is a rule of thumb used in drug discovery to evaluate a compound's drug-likeness based on its physicochemical properties. It helps prioritize compounds with optimal characteristics for oral bioavailability and pharmacokinetics.
- Criteria: According to this rule, compounds are considered favorable if they meet specific thresholds for molecular weight (MW) within the range of 200 to 600 Daltons, XLOGP values between -2 and 5, total polar surface area (TPSA) not exceeding 150 Å², up to 7 rings, more than 4 carbon atoms, more than 1 heteroatom, no more than 15 rotatable bonds, no more than 10 hydrogen bond acceptors, and at least 5 hydrogen bond donors. Content: $200 \leq MW \leq 600$, $-2 \leq XLOGP \leq 5$, $TPSA \leq 150$, Num. rings ≤ 7 , Num. carbon > 4 , Num- heteroatoms > 1 , Num. rotatable bonds ≤ 15 , H-bond acc ≤ 10 , H-bond don ≥ 5

Pfizer Rule

- Definition & Significance: The Pfizer Rule is a set of guidelines used in drug discovery to assess a compound's likelihood of being orally bioavailable based on its physicochemical properties. It helps prioritize compounds with favorable pharmacokinetic profiles for further development.
- Criteria: The Pfizer Rule suggests that orally bioavailable compounds typically have molecular weights less than 500 Daltons, log P values between -0.4 and 6, Topological surface area TPSA less than 75 Å². Content: $LogP > 3$; $TPSA < 75 \text{ Å}^2$.

GSK Rule

- Definition & Significance: The GSK Rule is a set of guidelines used in drug discovery to assess a compound's drug-likeness based on its physicochemical properties. It helps prioritize compounds with optimal characteristics for oral bioavailability and pharmacokinetics.
- Criteria: The GSK Rule suggests that orally active compounds typically have molecular weights less than or equal to 400 Daltons, log P value less than or equal to 4. Content: $MW \leq 400$; $LogP \leq 4$

Golden Triangle

- Definition & Significance: The Golden Triangle is a concept used in drug discovery to visualize the balance between molecular size, lipophilicity, and polarity of compounds. It helps identify compounds with optimal characteristics for oral bioavailability and pharmacokinetics.
- Criteria: This law suggests that compounds that fall within the molecular weight (MW) range of 200 to 500 Daltons and possess logarithm of the distribution coefficient (LogD) values ranging from -2 to 5. The Golden Triangle suggests that compounds with moderate molecular size, distribution lipophilicity, and polarity are more likely to exhibit favorable pharmacokinetic properties compared to compounds that deviate significantly from this balance. Content: $200 \leq MW \leq 500$; $-2 \leq LogD \leq 5$.

Stats

Bioavailability Spiderweb Chart

Overview

The Bioavailability Spiderweb Graph is an innovative visualization tool used to assess the oral bioactivity profile of a compound. This graphical representation allows researchers to quickly ascertain how well a compound is expected to perform in biological systems with regards to its oral uptake. The graph plots various physicochemical properties critical to oral bioavailability, providing a comparative view against established optimal ranges.

Graph Properties

The graph examines the following properties, essential for determining the oral bioactivity of a compound:

- **Lipophilicity (LogP):** Measures the ability of a compound to dissolve in fats, oils, and lipids, which is crucial for passage through lipid cell membranes. Optimal range for lipophilicity is an XLOGP3 score between -0.7 and $+5.0$.
- **Molecular Weight (MW):** Reflects the size of the compound, which affects its ability to be absorbed. An optimal molecular weight range is between 150 and 500 g/mol.
- **Polar Surface Area (TPSA):** Indicates the area of a molecule that is polar, affecting the compound's solubility and permeability. The optimal range for polarity is a TPSA between 20 and 130 \AA^2 .
- **Solubility (LogS):** A measure of how well the compound dissolves in aqueous solutions, impacting its ability to be absorbed orally. Optimal solubility is indicated by a log S value not higher than 6.
- **Saturation (FSP3):** Represents the fraction of carbons in sp^3 hybridization, giving an indication of the compound's three-dimensional structure. A higher sp^3 fraction (not less than 0.25) generally suggests better bioavailability.
- **Flexibility (Rotatable Bonds):** Relates to the number of bonds that can freely rotate in a molecule, with fewer rotatable bonds typically leading to better bioavailability. Flexibility is considered optimal when there are no more than 9 rotatable bonds.

Interpreting the Graph

The graph presents two key areas:

- **Optimal Range (Green Area):** The shaded green area within the graph depicts the optimal ranges for each property. The extent to which a compound's characteristics fall within this area is an indicator of its likely oral bioavailability.
- **Compound Properties (Red Polygon):** The red polygon overlaid on the graph shows the actual values for the compound being assessed. The closer this shape aligns with the green area, the more likely the compound is to have good oral bioavailability.

Troubleshooting

Encountering issues while using the GU Drug-Pro Toolkit is not uncommon. This section provides solutions to some of the most frequently encountered problems, helping you navigate through and resolve them efficiently.

Application Does Not Start

- **Check System Requirements:** Ensure your computer meets the minimum system requirements specified in the "Getting Started" section.
- **Administrator Rights:** Try running the application with administrator rights, especially on Windows, to ensure it has the necessary permissions.

Unable to Load or Save Files

- **File Permissions:** Check if the file or directory you're trying to access has the appropriate read/write permissions.
- **Supported Formats:** Ensure the file format is supported by the toolkit. Refer to the user manual for a list of compatible file types. If you encounter issues with molecule uploads, ensure the file format is correct.
- **Path Issues:** Verify the file path for typos or incorrect directory references.

Poor Performance or Slow Response

- **System Resources:** Close unnecessary applications to free up system resources. Consider upgrading your hardware if you consistently encounter performance issues.
- **Large Datasets:** Working with very large molecules can slow down the application. Try processing smaller batches of data.

Errors in Molecular Visualization

- **Graphics Drivers:** Ensure your graphics drivers are up to date, as outdated drivers may cause rendering issues in GUI.
- **Data Integrity:** Check the integrity of the molecular data being visualized. Incorrect or malformed data can lead to errors in visualization.

Incorrect or Unexpected Results

- **Data Input:** Double-check the accuracy and format of the input data. Minor errors or discrepancies can lead to significant variances in results.
- **Software Bugs:** Despite thorough testing, software bugs can occur. Note the steps leading to the issue and report it to the support team for investigation.

Network or Database Access Issues

- **Internet Connection:** Ensure you have a stable internet connection if accessing online databases or features. For issues with internet-based features, check your network connection and PubChem availability.
- **Firewall/Antivirus Settings:** Adjust your firewall or antivirus settings if they're blocking the application's internet access.

Updating the Software

- **Check for Updates:** Regularly check for and install updates to ensure you have the latest features and bug fixes.
- **Compatibility:** Ensure that updates are compatible with your operating system and hardware specifications.

Frequently Asked Questions FAQs

Provides answers to common questions about the application's use and functionalities.

- Q1: Can the GU Drug-Pro Toolkit be used on any operating system?
A1: The GU Drug-Pro Toolkit is currently designed to be compatible with only Windows 10 and above operating system. However, ensure you meet the specific system requirements detailed in the "Getting Started" section.
- Q2: Do I need an internet connection to use the toolkit?
A2: An internet connection is not mandatory for all features but is required for functions that access online databases or require software updates.
- Q3: How can I update the software to the latest version?
A3: Updates will be made available through our official GitHub repository for the project. Always ensure you have the latest version for optimal performance and features.
- Q4: How do I report a bug or suggest a feature?
A4: Feedback is invaluable to us. Please report bugs or suggest features via the contact information provided in the manual or through the support section on our website.
- Q5: Can I use the toolkit for commercial purposes?
A5: Please refer to the licensing agreement provided with the software or contact us for detailed information on commercial use.
- Q6: What should I do if I'm having trouble installing the software?
A6: Refer to the "Troubleshooting" section of this manual for common installation issues and solutions. If the problem persists, please contact our support team for assistance.
- Q7: How often are the drug databases and models updated?
A7: Online drug databases accessed by the toolkit are updated regularly by their respective providers. We strive to ensure the toolkit interfaces with the latest versions of these databases. We also ensure to update our machine learning models with great precision based on the latest advancements and optimization techniques currently available.
- Q8: Can the toolkit predict the success of a drug in clinical trials?
A8: While the toolkit provides valuable insights into a compound's properties and potential efficacy, predicting clinical trial outcomes involves many factors beyond the scope of the software.
- Q9: Does the toolkit support collaborative projects?
A9: While the current version focuses on individual use, we are exploring features that will facilitate collaboration in future releases.
- Q10: What kind of support does the GU Drug-Pro Toolkit offer?
A10: We provide email support, an online knowledge base, and, for enterprise users, dedicated account management. Our goal is to ensure you have the support you need to use the toolkit effectively.

We hope these FAQs address your initial queries regarding the GU Drug-Pro Toolkit. For more detailed inquiries or assistance, don't hesitate to reach out to our support team.

Contact and Support

The GU Drug-Pro Toolkit team is dedicated to providing exceptional service and ensuring your complete satisfaction. If you encounter any issues, have questions, or wish to offer feedback, we're here to provide the support you need.

How to Reach Us

For all your inquiries, including technical support, licensing, sales, and general questions:

Email: adeshpan@gitam.in

Hours of Operation: Monday to Friday, from 9:00 AM to 5:00 PM (IST)

To ensure we address your concerns efficiently, please provide a detailed description of the issue, any relevant error messages, the troubleshooting steps you've attempted, and details about your system setup. This information is invaluable for a swift and effective resolution.

Seeking Information?

Whether you're inquiring about general information and licensing options, exploring partnerships, or have content related questions, your inquiries are important to us. Don't hesitate to reach out.

Your Feedback Matters

We are committed to continuous improvement and value your feedback. If you have suggestions or comments about the GU Drug-Pro Toolkit, please email us at adeshpan@gitam.in. Your insights help us enhance our software and serve you better. Your experience with the GU Drug-Pro Toolkit is our top priority.

Reporting Bugs, Mistakes, or Corrections

Encountering a bug or noticing a mistake within our software can be frustrating. We strive for excellence and accuracy in all we do, and your reports play a crucial role in achieving this goal. If you discover a bug, error, or any aspect of the software that requires correction, please let us know. Include as much detail as possible—such as steps to reproduce the issue, screenshots, and the specific environment (OS version, toolkit version) where the issue occurred. This detail will greatly aid our development team in diagnosing and addressing the problem promptly. Send your reports to adeshpan@gitam.in with the subject line "Bug Report" or "Correction Needed" to ensure it is directed to the appropriate team for quick action.

Your experience with the GU Drug-Pro Toolkit is our top priority. We appreciate the opportunity to support your work and look forward to assisting you.