

PSC-db: A structured and searchable 3D-database for plant secondary compounds

Alejandro Valdés-Jiménez †, Carlos Peña-Varas †, Paola Borrego-Muñoz, Lily Arrue, Melissa Alegría-Arcos, Hussam Nour-Eldin, Ingo Dreyer, Gabriel Nuñez-Vivanco, David Ramírez*

* Correspondence: david.ramirez@uautonoma.cl; Tel.: +56 2 23036667

† These authors equally contributed to this work.

Supplemental information

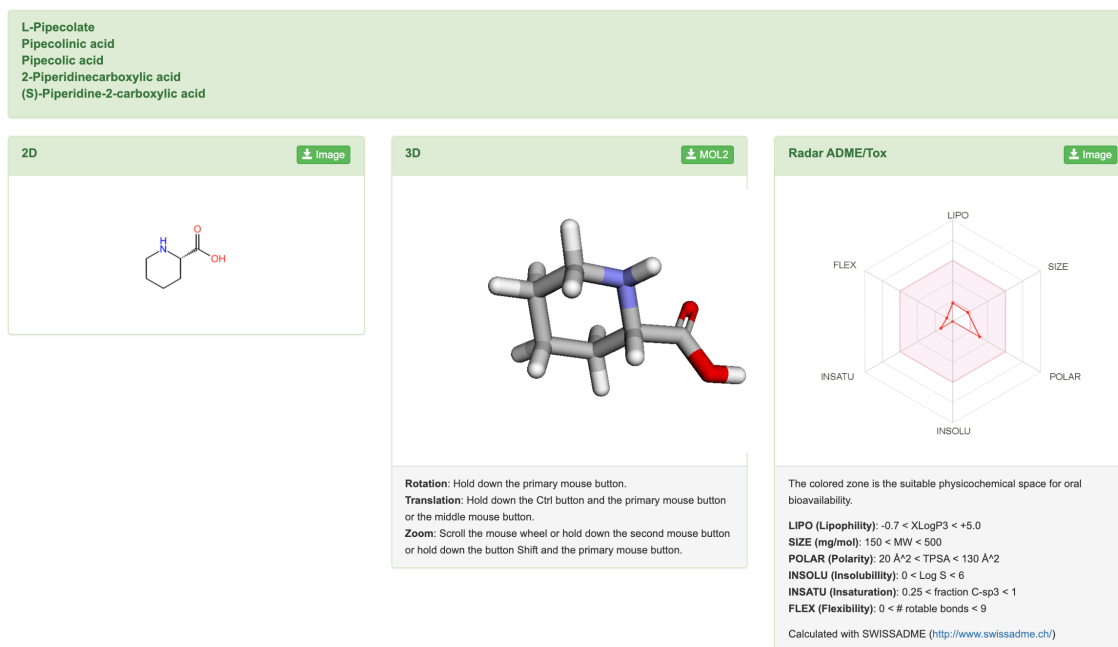


Figure S1. Illustrated record of PSC-db. A view of the 2D- and 3D-representation of a given molecule, as well as Bioavailability RADAR plotted with the predicted physicochemical and pharmaceutical features for the L-pipecolate alkaloid is shown.

<p>A</p> <p>L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid</p> <p>External Links Hierarchical Organization Physicochemical Properties Lipophilicity</p> <p>Water Solubility Pharmacokinetics Druglikeness Source Organism Biological Activity</p> <p>NIKKAJI PDB-CCD KNApSAcK PubChem KEGG ChEMBL ZINC</p>	<p>B</p> <p>L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid</p> <p>External Links Hierarchical Organization Physicochemical Properties Lipophilicity</p> <p>Water Solubility Pharmacokinetics Druglikeness Source Organism Biological Activity</p> <p>[-> Plant Secondary Compounds [-> Alkaloids [->> Alkaloids derived from lysine [->>> Piperidine alkaloids</p>
--	---

Figure S2. Record for the L-Pipecolate alkaloid is shown, with the names and alternative denominations, as well as the external links (A) and hierarchical organization (B).

<p>A</p> <p>L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid</p> <p>External Links Hierarchical Organization Physicochemical Properties Lipophilicity</p> <p>Water Solubility Pharmacokinetics Druglikeness Source Organism Biological Activity</p> <table border="1"> <tr><td>Number of heavy atoms</td><td>9</td></tr> <tr><td>Number of aromatic heavy atoms</td><td>0</td></tr> <tr><td>Fraction Csp3</td><td>0.83</td></tr> <tr><td>Number of rotatable bonds</td><td>1</td></tr> <tr><td>Number of H-bond acceptors</td><td>3</td></tr> <tr><td>Number of H-bond donors</td><td>2</td></tr> <tr><td>Molar Refractivity</td><td>37.33</td></tr> <tr><td>TPSA</td><td>49.33</td></tr> </table> <p>TPSA: Topological Polar Surface Area. Calculated from Ertl P. et al. 2000 J. Med. Chem. Calculated with SWISSADME (http://www.swissadme.ch)</p>	Number of heavy atoms	9	Number of aromatic heavy atoms	0	Fraction Csp3	0.83	Number of rotatable bonds	1	Number of H-bond acceptors	3	Number of H-bond donors	2	Molar Refractivity	37.33	TPSA	49.33	<p>B</p> <p>L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid</p> <p>External Links Hierarchical Organization Physicochemical Properties Lipophilicity</p> <p>Water Solubility Pharmacokinetics Druglikeness Source Organism Biological Activity</p> <table border="1"> <tr><td>lLOGP</td><td>1.16</td></tr> <tr><td>XLOGP3</td><td>-2.31</td></tr> <tr><td>WLOGP</td><td>-0.17</td></tr> <tr><td>MLOGP</td><td>-2.21</td></tr> <tr><td>SILICOS-IT</td><td>0.46</td></tr> <tr><td>Consensus Log P(o/w)</td><td>-0.61</td></tr> </table> <p>lLOGP: In-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model. XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model. MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev. SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com Consensus Log P(o/w): Average of all five predictions Calculated with SWISSADME (http://www.swissadme.ch)</p>	lLOGP	1.16	XLOGP3	-2.31	WLOGP	-0.17	MLOGP	-2.21	SILICOS-IT	0.46	Consensus Log P(o/w)	-0.61												
Number of heavy atoms	9																																								
Number of aromatic heavy atoms	0																																								
Fraction Csp3	0.83																																								
Number of rotatable bonds	1																																								
Number of H-bond acceptors	3																																								
Number of H-bond donors	2																																								
Molar Refractivity	37.33																																								
TPSA	49.33																																								
lLOGP	1.16																																								
XLOGP3	-2.31																																								
WLOGP	-0.17																																								
MLOGP	-2.21																																								
SILICOS-IT	0.46																																								
Consensus Log P(o/w)	-0.61																																								
<p>C</p> <p>L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid</p> <p>External Links Hierarchical Organization Physicochemical Properties Lipophilicity</p> <p>Water Solubility Pharmacokinetics Druglikeness Source Organism Biological Activity</p> <table border="1"> <tr><td>ESOL Log S</td><td>0.88</td></tr> <tr><td>ESOL Solubility (mg/mL)</td><td>981</td></tr> <tr><td>ESOL Solubility (mol/L)</td><td>7.59</td></tr> <tr><td>ESOL Class</td><td>Highly soluble</td></tr> <tr><td>ALI Log S</td><td>1.81</td></tr> <tr><td>ALI Solubility (mg/mL)</td><td>8340</td></tr> <tr><td>ALI Solubility (mol/L)</td><td>64.6</td></tr> <tr><td>ALI Class</td><td>Highly soluble</td></tr> <tr><td>SILICOS-IT Log S</td><td>-0.42</td></tr> <tr><td>SILICOS-IT Solubility (mg/mL)</td><td>48.5</td></tr> <tr><td>SILICOS-IT Solubility (mol/L)</td><td>0.376</td></tr> <tr><td>SILICOS-IT Class</td><td>Soluble</td></tr> </table> <p>ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model. Solubility class: Log S scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly ALI: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model. SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com Calculated with SWISSADME (http://www.swissadme.ch)</p>	ESOL Log S	0.88	ESOL Solubility (mg/mL)	981	ESOL Solubility (mol/L)	7.59	ESOL Class	Highly soluble	ALI Log S	1.81	ALI Solubility (mg/mL)	8340	ALI Solubility (mol/L)	64.6	ALI Class	Highly soluble	SILICOS-IT Log S	-0.42	SILICOS-IT Solubility (mg/mL)	48.5	SILICOS-IT Solubility (mol/L)	0.376	SILICOS-IT Class	Soluble	<p>D</p> <p>L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid</p> <p>External Links Hierarchical Organization Physicochemical Properties Lipophilicity</p> <p>Water Solubility Pharmacokinetics Druglikeness Source Organism Biological Activity</p> <table border="1"> <tr><td>Gastrointestinal absorption</td><td>High</td></tr> <tr><td>BBB permeant</td><td>No</td></tr> <tr><td>Cytochrome P450 1A2 inhibitor</td><td>No</td></tr> <tr><td>Cytochrome P450 2C19 inhibitor</td><td>No</td></tr> <tr><td>Cytochrome P450 2C9 inhibitor</td><td>No</td></tr> <tr><td>Cytochrome P450 2D6 inhibitor</td><td>No</td></tr> <tr><td>Cytochrome P450 3A4 inhibitor</td><td>No</td></tr> <tr><td>Skin permeation</td><td>-8.73</td></tr> </table> <p>Gastrointestinal absorption: according to the white of the BOILED-Egg. For details please refer to this article: A BOILED-Egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem (2016) 11(11):1117-1121. BBB permeant: Blood barrier brain permeation according to the yolk of the BOILED-Egg. For details please refer to this article: A BOILED-Egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem (2016) 11(11):1117-1121. Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91 Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87 Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.77 / AUC=0.81 Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87 Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86 Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res. Calculated with SWISSADME (http://www.swissadme.ch)</p>	Gastrointestinal absorption	High	BBB permeant	No	Cytochrome P450 1A2 inhibitor	No	Cytochrome P450 2C19 inhibitor	No	Cytochrome P450 2C9 inhibitor	No	Cytochrome P450 2D6 inhibitor	No	Cytochrome P450 3A4 inhibitor	No	Skin permeation	-8.73
ESOL Log S	0.88																																								
ESOL Solubility (mg/mL)	981																																								
ESOL Solubility (mol/L)	7.59																																								
ESOL Class	Highly soluble																																								
ALI Log S	1.81																																								
ALI Solubility (mg/mL)	8340																																								
ALI Solubility (mol/L)	64.6																																								
ALI Class	Highly soluble																																								
SILICOS-IT Log S	-0.42																																								
SILICOS-IT Solubility (mg/mL)	48.5																																								
SILICOS-IT Solubility (mol/L)	0.376																																								
SILICOS-IT Class	Soluble																																								
Gastrointestinal absorption	High																																								
BBB permeant	No																																								
Cytochrome P450 1A2 inhibitor	No																																								
Cytochrome P450 2C19 inhibitor	No																																								
Cytochrome P450 2C9 inhibitor	No																																								
Cytochrome P450 2D6 inhibitor	No																																								
Cytochrome P450 3A4 inhibitor	No																																								
Skin permeation	-8.73																																								
<p>E</p> <p>L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid</p> <p>External Links Hierarchical Organization Physicochemical Properties Lipophilicity</p> <p>Water Solubility Pharmacokinetics Druglikeness Source Organism Biological Activity</p> <table border="1"> <tr><td>Lipinski (Pfizer) filter violations</td><td>0</td></tr> <tr><td>Ghose filter violations</td><td>2</td></tr> <tr><td>Veber (GSK) filter violations</td><td>0</td></tr> <tr><td>Egan (Pharmacia) filter violations</td><td>0</td></tr> <tr><td>Muegge (Bayer) filter violations</td><td>2</td></tr> <tr><td>Abbott Bioavailability Score</td><td>0.55</td></tr> </table> <p>Lipinski (Pfizer) filter violations: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW ≤ 500 MLOGP ≤ 4.15 N or O ≤ 10 NH or OH ≤ 5 Ghose filter violations: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70 Veber (GSK) filter violations: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds ≤ 10 TPSA ≤ 140 Egan (Pharmacia) filter violations: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP ≤ 5.88 TPSA ≤ 131.6 Muegge (Bayer) filter violations: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 ≤ MW ≤ 600 -2 ≤ XLOGP ≤ 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5 Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem. Calculated with SWISSADME (http://www.swissadme.ch)</p>	Lipinski (Pfizer) filter violations	0	Ghose filter violations	2	Veber (GSK) filter violations	0	Egan (Pharmacia) filter violations	0	Muegge (Bayer) filter violations	2	Abbott Bioavailability Score	0.55	<p>E</p> <p>Lipinski (Pfizer) filter violations: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW ≤ 500 MLOGP ≤ 4.15 N or O ≤ 10 NH or OH ≤ 5 Ghose filter violations: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70 Veber (GSK) filter violations: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds ≤ 10 TPSA ≤ 140 Egan (Pharmacia) filter violations: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP ≤ 5.88 TPSA ≤ 131.6 Muegge (Bayer) filter violations: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 ≤ MW ≤ 600 -2 ≤ XLOGP ≤ 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5 Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem. Calculated with SWISSADME (http://www.swissadme.ch)</p>																												
Lipinski (Pfizer) filter violations	0																																								
Ghose filter violations	2																																								
Veber (GSK) filter violations	0																																								
Egan (Pharmacia) filter violations	0																																								
Muegge (Bayer) filter violations	2																																								
Abbott Bioavailability Score	0.55																																								

Figure S3. Record for the L-Pipecolate alkaloid is shown with the physicochemical and pharmaceutical properties calculated with SwissADME server (A-E).

L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid				A	L-Pipecolate Pipecolic acid Pipecolic acid 2-Piperidinecarboxylic acid (S)-Piperidine-2-carboxylic acid				B
External Links Hierarchical Organization Physicochemical Properties Lipophilicity Water Solubility Pharmacokinetics Druglikeness Source Organism					External Links Hierarchical Organization Physicochemical Properties Lipophilicity Water Solubility				
Biological Activity					Pharmacokinetics Druglikeness Source Organism Biological Activity				
Kingdom	Family	Species	Reference		Target	Target Type	ChEMBL ID	Reference	
Fungi	Arthrodermataceae	Arthroderma otae	CB5113488	D.Cook et al., <i>63 Gene Genomes Genetics</i> , 7, (2017), 1791-1797	Perisomal sarcosine oxidase	SINGLE PROTEIN	ChEMBL2254	10.1021/jp968331f	
Fungi	Arthrodermataceae	Trichophyton equinum	CB5127.97	D.Cook et al., <i>63 Gene Genomes Genetics</i> , 7, (2017), 1791-1797	Proton-coupled amino acid transporter 1	SINGLE PROTEIN	ChEMBL1914279	10.1016/j.bmc.2011.08.058	
Fungi	Dothideomycetes	Alternaria oxyptris	Raft River	D.Cook et al., <i>63 Gene Genomes Genetics</i> , 7, (2017), 1791-1797					
Fungi	Incertae sedis	Phoma medicaginis		Fan, Din., et al., <i>PLoS One</i> , 13, (2018), e0286641					
Plantae	Araceae	Lemna gibba		Fujisaka, <i>Plant Cell Physiol.</i> , 28, (1987), 995					
Plantae	Araceae	Lemna gibba		Fujisaka, <i>Plant Cell Physiol.</i> , 33, (1992), 419					
Plantae	Araceae	Lemna paucicostata		Fujisaka, <i>Plant Cell Physiol.</i> , 33, (1992), 419					
Plantae	Fabaceae	Phaseolus vulgaris		Harborne, <i>Phytochemical Dictionary Second Edition</i> , Taylor and Francis, (1999), Chapter 18					
Plantae	Fagaceae	Castanea sativa		Servillo, L. et al., <i>Food Chem.</i> , 106, (2016), 1381-1389.					

Source: KINQSDAR (http://www.kinqsadfamily.com/kinqsad_cmfhp.php)

Source: ChEMBL (<https://www.ebi.ac.uk/chembl/>)

Figure S4. Record for the L-Pipecolate alkaloid is shown with the source organism (A) and biological activity (B).

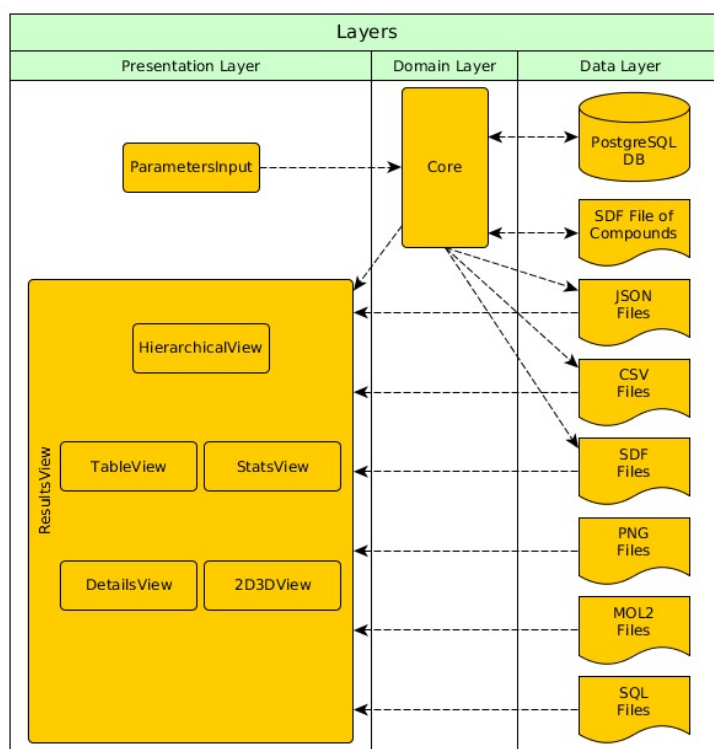


Figure S5. The architecture of the solution and the essential components of PSC-db. It consists of three main layers: the presentation, domain, and data layers, representing the interaction between the essential components of the solution.

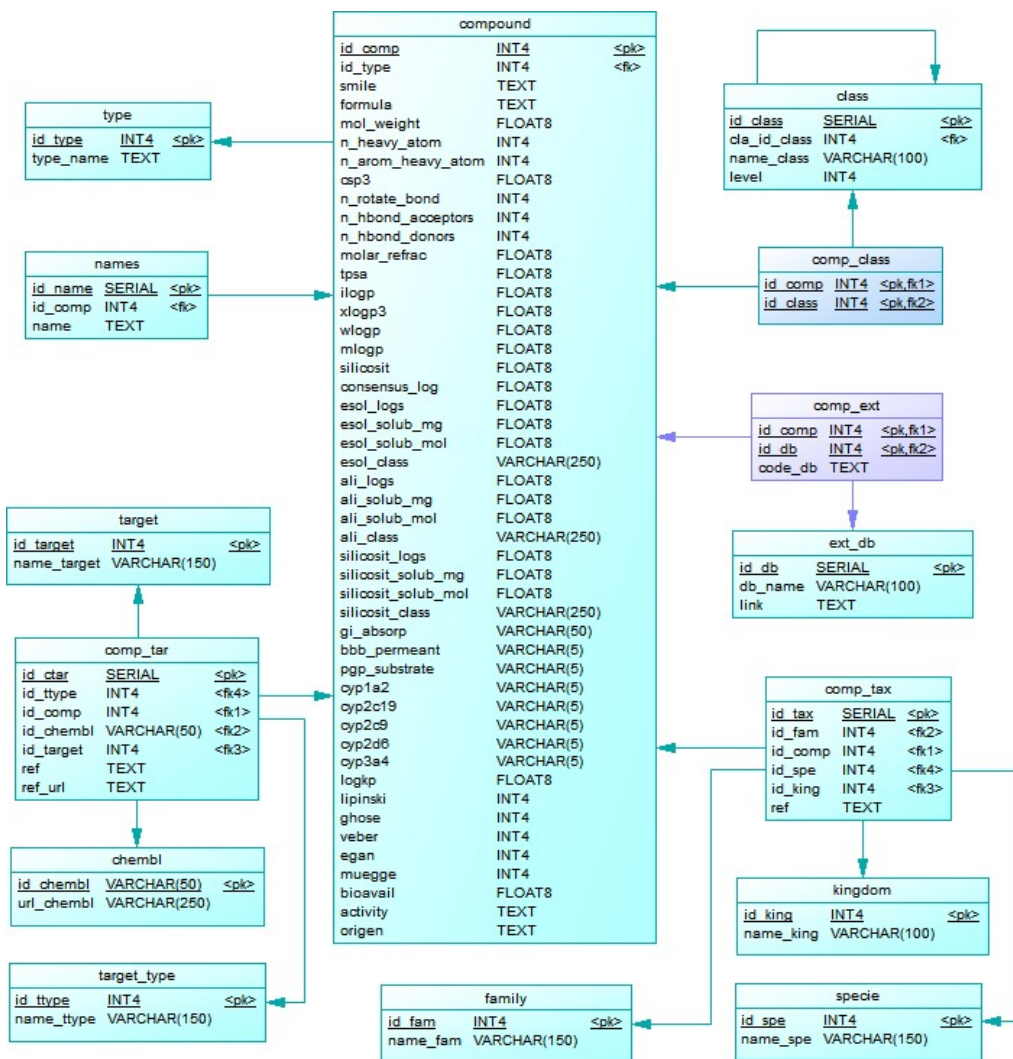


Figure S6. Physical data model implemented on PSC-db.