Virtual Exploration of the chemical universe up to 11 atoms of C, N, O, F: assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physico-chemical properties, compound classes and drug discovery

Tobias Fink and Jean-Louis Reymond*

Department of Chemistry and Biochemistry, University of Berne, Freiestrasse 3, CH-3012 Berne, Switzerland

FAX: + 41 31 631 80 57

E-mail: jean-louis.reymond@ioc.unibe.ch

Supplementary Information

Figure S1, Table S1 and Figure S2.

Figure S1. Distribution of the highest atomic contribution (HAC) of ringsystems involved in the database construction process. The atomic contribution to the total steric energy of the molecule in a MM2 force field expresses how much this specific atom is distorted in respect to the ideal geometry. For graphs with multiple stereoisomers and conformers, the lowest HAC in the different stereoisomers and conformers was selected for each atom. Only one of the rejected ring systems (HAC +22.0 kcal/mol) appears in the literature with a phosphorous atom at the distorted bridge-head: CC(O4)(C)OCC4C2O[PH]13OCC(C2O3)O1. Munoz, A.; Rochal, A. Macrocyclic spirophosphanes prepared from isopropylidene mannitols. *Phosphorus, Sulfur and Silicon and the Related Elements* **2001**, 174, 177-192.

Table S1. List of filters applied during the structure generation process. Each filter consists of a rejected functional group and a list of exceptions.

Filter	Rejected feature	Exceptions
Acid Taut		-
Aminal		-
C=N		
Decarboxy		
Enamine		
Enol		
FC		
Geminal		-
Hemiacetal		-
Hemiaminal		-

Filter	Rejected feature	Exceptions
Het-Het	$\text{O}-\text{O}$ $\text{N}-\text{N}$ $\text{F}-\text{O}$ $\text{F}-\text{N}$ $\text{N}-\text{O}$ $\text{N}=\text{N}$ $\text{N}=\text{O}$ 	
Het-Het-Het	$\text{N}=\text{N}$ $\text{N}-\text{O}$ 	
Hetero-SR	 	
IntraMol	 	
Mixed	 	
Ortho		-
Polycyclic		-

Figure S2. Descriptor value distribution profile for **GDB** (=SCU, red), **RDB** (green) and **DMU** (=SMU, blue).