Instruction Manual and Supplementary Information

**Database for the kinetics of the gas-phase atmospheric reactions
of organic compounds**

Compiled by

W. P. L. Carter, M. R. McGillen, A. Mellouki, J. J. Orlando,
B. Picquet-Varrault, and T. J. Wallington

November 28, 2019

Summary

We present here a freely available, digital and searchable compilation of evaluated rate constants of gas-phase reactions of organic compounds with OH, Cl, and NO_3 radicals and with O_3 at atmospheric pressure. This includes recommendations for rate constants at 298K and their temperature dependences where sufficient data are available. The background, objectives, compilation and review methods, and a general discussion of the database is given in the manuscript "Database for the kinetics of the gas-phase atmospheric reactions of organic compounds", which is being submitted for publication (McGillen et al, 2019). The database is distributed in an Excel spreadsheet named AtmVOCkin-n, where "n" refers to the version number. The database and its associated documentation is available at its website at <https://data.eurochamp.org/data-access/kin/>. Our intention is to update this database as new information becomes available, so the database's website should be checked periodically to see if there is an updated version.

This document contains supplementary information for this database and is part of its overall documentation. It provides a user manual on how to search and extract selected information from the database and describes how the data are organized in the Excel spreadsheet file. The current document describes the first version of the distributed database, but will be updated as needed when new versions of the database become available.

Contents

List of Tables

List of Figures

Use of the Database File

The current version of the database is entirely contained in the Excel spreadsheet file AtmVOCkin-n.xls, where "n" is the version number. This has two worksheets with general information, three with inputs and/or outputs for macros to search and extract data, and five with various tables of data. The worksheets and macros that the user would normally use to search and extract output from the dataset are described in this section. The contents and format of the sheets and tables with the data and supporting information are given in the following section.

Note that it may be necessary to revise the security settings in Excel for the macros to be enabled for searching and extracting the data. Generally, you will get a warning when you open the file if your security settings won't allow the macros to run. For example, in Excel 2016, you can enable the macros through the file > options > Trust Center > Trust Center Settings > Macro Settings menus, then selecting "Enable all macros". This may differ between Excel versions. Note that Excel will warn that this is not recommended because potentially dangerous code can be run, so those who are uncomfortable with remaining in this mode should change the settings back when done with searching and extracting the desired data.

Information contained in the "Info" Sheet

This worksheet contains information about the database, including the version, log of updates, current values of the database size parameters, summary of worksheets and their contents, and information about macros to search and output the data. Much of this information is also contained in this document, but it also has database size information that is applicable only to the current version.

Finding Compounds where Data are Available

The "Find Compounds" sheet provides a means to search for compounds in the database so selected information for those compounds can be output. **Figure 1** shows a screenshot of relevant portions of this worksheet showing examples of the four possible search methods. These are discussed below.

Search by name or alias. The use of the search by name or alias method is shown on **Figure 1**a. As shown there, the user enters the part of the name or alias in the cell to the right of the "Part of Name or Alias" control, and clicks on that control to search for the name as a substring of the names or aliases assigned for the compound. In addition to names or aliases, InChIKey strings can also be used, and, for at least some compounds, the CAS number. See the discussion of the "Names DB" table, in the following section for more information on the names, aliases or other designations that can be used. This search is not case sensitive.

Search by Atom Numbers. The use of search by atom numbers is shown in **Figure 1**b. Put the atom numbers, without embedded spaces and with atom symbols case sensitive (e.g., C6H13NO3), in the cell next to the "Atom Nos" control. It doesn't matter how the atoms are ordered, but the macro will replace your entry with the standard order if applicable.

Search by Compound Type. The use of search by compound type is shown in **Figure 1**c. The current database classifies most (but not all) compounds by type (e.g., "normal alkane", "hydroxy nitrate", etc.) and right hand of the "Find Compounds" sheet has a table giving the code numbers used for each type (see "Table C" in **Figure 1**c). Note that any given compound

is assigned only one type. The user looks up the compound type number in "Table C" and enters it in the cell next to the "Compound Type" control and runs the macro to get all compounds assigned this type.
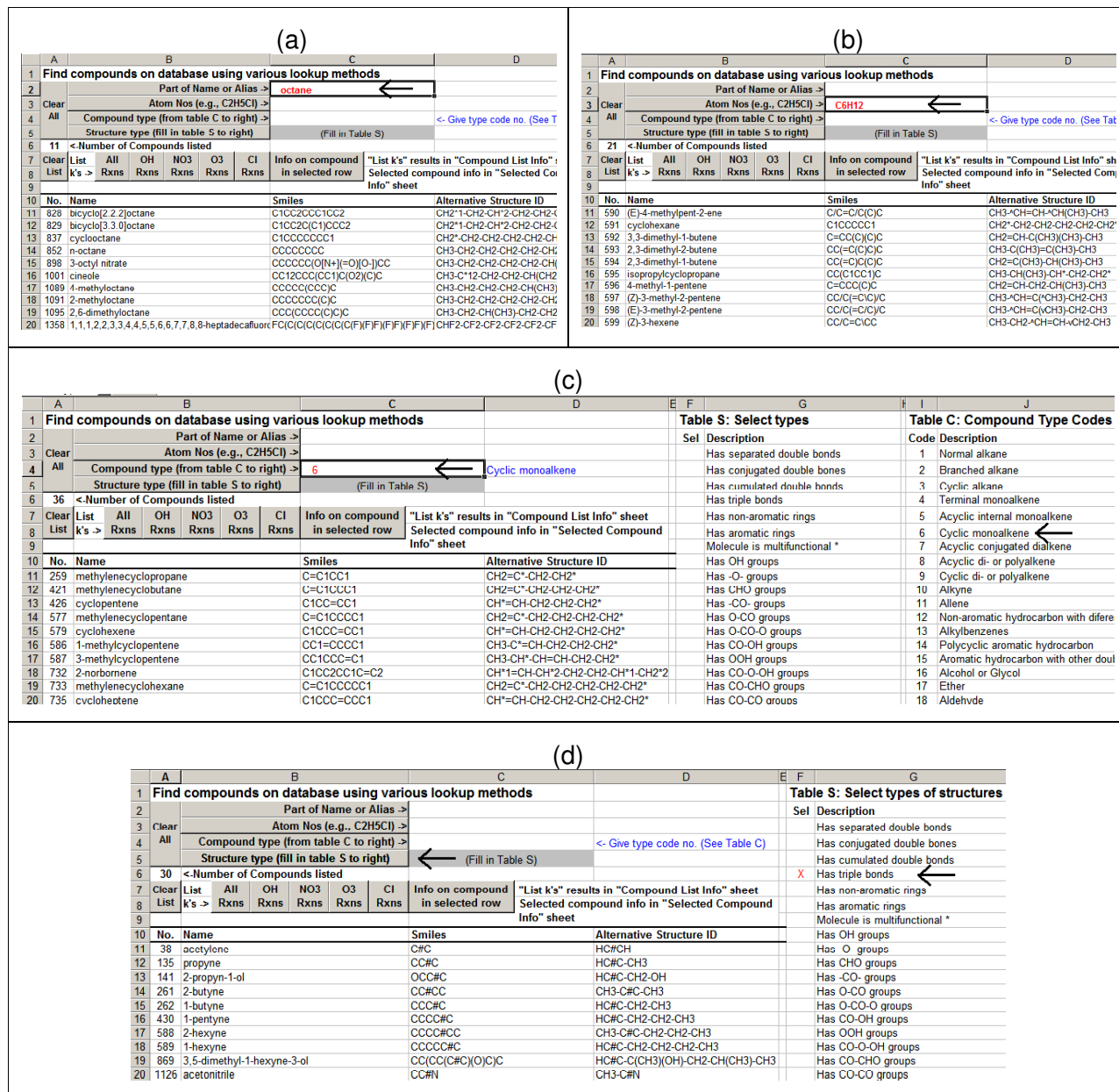
(a)

No.	Name	Smiles	Alternative Structure ID
	Part of Name or Alias -> octane		
Clear All	Atom Nos (e.g., C2H5Cl) ->		
	Compound type (from table C to right) ->		<- Give type code no. (See T
	Structure type (fill in table S to right)	(Fill in Table S)	
11	<-Number of Compounds listed		
828	bicyclo[2.2.2]octane	C1CC2CCC1CC2	CH2*1-CH2-CH*2-CH2-CH2-(
829	bicyclo[3.3.0]octane	C1CC2C(C1)CCC2	CH2*1-CH2-CH*2-CH2-CH2-(
837	cyclooctane	C1CCCCCCC1	CH2*-CH2-CH2-CH2-CH
852	n-octane	CCCCCCCC	CH3-CH2-CH2-CH2-CH2-CH2
898	3-octyl nitrate	CCCCCC(O[N+](=O)[O-])CC	CH3-CH2-CH2-CH2-CH2-CH(
1001	cineole	CC12CCC(CC1)(O2)(C)C	CH3-C*12-CH2-CH2-CH(CH2
1089	4-methyloctane	CCCCC(CCC)C	CH3-CH2-CH2-CH2-CH(CH3)
1091	2-methyloctane	CCCCCCC(C)C	CH3-CH2-CH2-CH2-CH2-CH2
1095	2,6-dimethyloctane	CCC(CCCC(C)C)C	CH3-CH2-CH(CH3)-CH2-CH2
1358	1,1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8-heptadecafluoro	FC(C(C(C(C(C(C(F)F)F)F)F)F)F)(F)F	CHF2-CF2-CF2-CF2-CF2-CF

Clear List / List k's -> / All Rxns / OH Rxns / NO3 Rxns / O3 Rxns / Cl Rxns / Info on compound in selected row / "List k's" results in "Compound List Info" Selected compound info in "Selected Compound Info" sheet

(b)

No.	Name	Smiles	Alternative Structure ID
	Part of Name or Alias ->		
Clear All	Atom Nos (e.g., C2H5Cl) -> C6H12		
	Compound type (from table C to right) ->		<- Give type code no. (See Tal
	Structure type (fill in table S to right)	(Fill in Table S)	
21	<-Number of Compounds listed		
590	(E)-4-methylpent-2-ene	C/C=C/C(C)C	CH3-^CH=CH-^CH(CH3)-CH3
591	cyclohexane	C1CCCCC1	CH2*-CH2-CH2-CH2-CH2-CH2*
592	3,3-dimethyl-1-butene	C=CC(C)(C)C	CH2=CH-C(CH3)(CH3)-CH3
593	2,3-dimethyl-2-butene	CC(=C(C)C)C	CH3-C(CH3)=C(CH3)-CH3
594	2,3-dimethyl-1-butene	CC(=C(C)C)C	CH2=C(CH3)-CH(CH3)-CH3
595	isopropylcyclopropane	CC(C1CC1)C	CH3-CH(CH3)-CH*-CH2-CH2*
596	4-methyl-1-pentene	C=CCC(C)C	CH2=CH-CH2-CH(CH3)-CH3
597	(Z)-3-methyl-2-pentene	CC/C=C\C)/C	CH3-^CH=C(^CH3)-CH2-CH3
598	(E)-3-methyl-2-pentene	CC/C=C/C)/C	CH3-^CH=C(\CH3)-CH2-CH3
599	(Z)-3-hexene	CC/C=C\CC	CH3-CH2-^CH=^CH-vCH2-CH3

(c)

No.	Name	Smiles	Alternative Structure ID		Sel	Description		Code	Description
	Part of Name or Alias ->					Has separated double bonds		1	Normal alkane
Clear All	Atom Nos (e.g., C2H5Cl) ->					Has conjugated double bones		2	Branched alkane
	Compound type (from table C to right) -> 6	Cyclic monoalkene				Has cumulated double bonds		3	Cyclic alkane
	Structure type (fill in table S to right)	(Fill in Table S)				Has triple bonds		4	Terminal monoalkene
36	<-Number of Compounds listed					Has non-aromatic rings		5	Acyclic internal monoalkene
						Has aromatic rings		6	Cyclic monoalkene
						Molecule is multifunctional *		7	Acyclic conjugated dialkene
259	methylenecyclopropane	C=C1CC1	CH2=C*-CH2-CH2*			Has OH groups		8	Acyclic di- or polyalkene
421	methylenecyclobutane	C=C1CCC1	CH2=C*-CH2-CH2-CH2*			Has -O- groups		9	Cyclic di- or polyalkene
426	cyclopentene	C1CC=CC1	CH*=CH-CH2-CH2-CH2*			Has CHO groups		10	Alkyne
577	methylenecyclopentane	C=C1CCCC1	CH2=C*-CH2-CH2-CH2-CH2*			Has -CO- groups		11	Allene
579	cyclohexene	C1CCC=CC1	CH*=CH-CH2-CH2-CH2-CH2*			Has O-CO groups		12	Non-aromatic hydrocarbon with differe
586	1-methylcyclopentene	CC1=CCCC1	CH3-C*=CH-CH2-CH2-CH2*			Has O-CO-O groups		13	Alkylbenzenes
587	3-methylcyclopentene	CC1CCC=C1	CH3-CH*-CH=CH-CH2-CH2*			Has CO-OH groups		14	Polycyclic aromatic hydrocarbon
732	2-norbornene	C1CC2CC1C=C2	CH*1=CH-CH*2-CH2-CH2-CH*1-CH2*2			Has OOH groups		15	Aromatic hydrocarbon with other doubl
733	methylenecyclohexane	C=C1CCCCC1	CH2=C*-CH2-CH2-CH2-CH2-CH2*			Has CO-O-OH groups		16	Alcohol or Glycol
735	cycloheptene	C1CCC=CCC1	CH*=CH-CH2-CH2-CH2-CH2-CH2*			Has CO-CHO groups		17	Ether
						Has CO-CO groups		18	Aldehyde

Table S: Select types
Table C: Compound Type Codes

(d)

No.	Name	Smiles	Alternative Structure ID		Sel	Description
	Part of Name or Alias ->					Has separated double bonds
Clear All	Atom Nos (e.g., C2H5Cl) ->					Has conjugated double bones
	Compound type (from table C to right) ->		<- Give type code no. (See Table C)			Has cumulated double bonds
	Structure type (fill in table S to right)	(Fill in Table S)			X	Has triple bonds
30	<-Number of Compounds listed					Has non-aromatic rings
						Has aromatic rings
						Molecule is multifunctional *
38	acetylene	C#C	HC#CH			Has OH groups
135	propyne	CC#C	HC#C-CH3			Has O groups
141	2-propyn-1-ol	OCC#C	HC#C-CH2-OH			Has CHO groups
261	2-butyne	CC#CC	CH3-C#C-CH3			Has -CO- groups
262	1-butyne	CCC#C	HC#C-CH2-CH3			Has O-CO groups
430	1-pentyne	CCCC#C	HC#C-CH2-CH2-CH3			Has O-CO-O groups
588	2-hexyne	CCCC#CC	CH3-C#C-CH2-CH2-CH3			Has CO-OH groups
589	1-hexyne	CCCCC#C	HC#C-CH2-CH2-CH2-CH3			Has OOH groups
869	3,5-dimethyl-1-hexyne-3-ol	CC(CC(C#C)(O)C)C	HC#C-C(CH3)(OH)-CH2-CH(CH3)-CH3			Has CO-O-OH groups
1126	acetonitrile	CC#N	CH3-C#N			Has CO-CHO groups
						Has CO-CO groups

Table S: Select types of structures

Figure 1. Screenshots of the "Find Compounds" sheet, showing selected input and output of the four search macros. Arrows point to the input cells and the controls used to run the macros.

Search by Structure Type. The use of search by structure type is shown in **Figure 1**d. The current database classifies most (though not all) compounds by the types of non-hydrocarbon groups they contain, and other aspects of their structures such as the presence of multiple bonds. Table "S" in the "Find Compounds" sheet has the structural characteristics that can be used, as indicated on **Figure 1**d. To do this type of search, the user puts a nonblank character in the "Sel" column of "Table S" and then runs the "Structure type" macro to get all

compounds with this structural feature. These nonblank characters in Table S do not affect other types of searches.

Note that currently only one type of search can be run at a time, i.e., it is not currently possible to use two or more criteria when doing searches. This capability can be added in future versions if there is sufficient need and interest.

As shown in **Figure 1**, The operations of these macros will produce a list of compounds in the "Find Compound" sheet that satisfy the search criteria selected, which may be anywhere between none or all of the compounds in the database. The list gives the primary name and SMILES string for the compound, and, for most compounds, the structural identifier string used by the SAPRC mechanism generation system (MechGen) (Carter, 2019). As discussed below, MechGen is used to derive the compound types and structural information used in the searches, and the format of its structural codes may be easier for some to understand than SMILES strings. These results can then be used to obtain rate constant information of the listed compounds, as discussed below.

Two types of macros can be run within the "Find Data" sheet to obtain rate constant information for the selected compounds. Control buttons for running these macros are located below those used to search the compound list, as shown on **Figure 1**. These two types of macros are discussed in the following section

Listings of recommended kinetic parameters in the "Compound List Info" sheet

Listings of recommended 298K rate constants, temperature dependences, overall uncertainty estimates and in some cases comments about availability of additional data or estimates can be output for all or selected compounds and reactions in the "Compound List Info" sheet. This sheet has columns for the type of reaction (OH, O_3, etc), primary name, primary SMILES string, $k(298)$ (in cm^3 molecule^{-1} s^{-1}), the relative uncertainty of k(298), the "A", "B", and "n" extended Arrhenius temperature dependence parameters [where $k(T)$=A exp(-B/T) $(T/300)^n$], the temperature range, and in some cases comments noting availability of data or estimates of $k(298)$ by extrapolation.

When distributed, this sheet has all the reactions in the database, and this can be re-generated by running the "List Data for All Compounds" macro on that sheet. This consists of thousands of columns of data, so it might not be the most convenient listing if one is interested in only one or a few compounds. A (usually) more convenient method is to use the "Find Compounds" sheet to get a listing of the compounds of interest, and then running the "All Rxns", "OH Rxns, etc. macros to produce the listings of recommended kinetic data. **Figure 2**a shows a screenshot of the "Find Compound" sheet after a search using "chloroethene" and indicates the macro to list all the reactions, and **Figure 2**b shows the resulting information in the "Compound List Info" sheet.

In the example shown on **Figure 2**b there was information on the OH, NO_3, O_3, and Cl rate constants for all the reactions listed. However, had there been no information for one of these reactions, then the list would not include the reaction. For example, had one run "All Rxns" for a list of alkanes, the list would not include any O_3 rate constants because no alkane + O_3 reactions are in our database.

(a)

	A	B	C	D
1		**Find compounds on database using various lookup methods**		
2		Part of Name or Alias ->	chloroethene	
3	Clear	Atom Nos (e.g., C2H5Cl) ->		
4	All	Compound type (from table C to right) ->		<- Give type code no. (See Table C)
5		Structure type (fill in table S to right)	(Fill in Table S)	
6	6	<-Number of Compounds listed		
7	Clear List	List All OH NO3 O3 Cl	Info on compound	"List k's" results in "Compound List Info" sheet
8		k's -> Rxns Rxns Rxns Rxns Rxns	in selected row	Selected compound info in "Selected Compound Info" sheet
9				
10	No.	Name	Smiles	Alternative Structure ID
11	55	chloroethene	ClC=C	CH2=CH-Cl
12	79	(Z)-1,2-dichloroethene	Cl/C=C\Cl	Cl-^CH=CH-vCl
13	80	(E)-1,2-Dichloroethene	Cl/C=C/Cl	Cl-^CH=CH-^Cl
14	81	1,1-dichloroethene	ClC(=C)Cl	CH2=C(Cl)-Cl
15	104	trichloroethene	ClC=C(Cl)Cl	Cl-CH=C(Cl)-Cl
16	121	tetrachloroethene	ClC(=C(Cl)Cl)Cl	Cl-C(Cl)–C(Cl)-Cl

(b)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
	List data for all compounds		Get Info on Selected Compound														
1	From the compilation of McGillen et al, "Database for the kinetics of the gas-phase atmospheric reactions of organic compounds," in preparation (2019) (Tentative title, subject to change)																
2	List created from AtmVOCkin-1.1.xls on 11/7/2019. Data shown for the 6 selected compounds.																
3				cm3 molec-1 s-1			k(T) = A * exp(-B/T) * (T/300)^n										
4	Rxn	Name	Smiles	k(289)	Unc.	A	B	n	Low T	High T							
5	OH	chloroethene	ClC=C	7.55E-12	20%	2.54E-12	-325		289	600							
6	OH	(Z)-1,2-dichloroethene	Cl/C=C\Cl	2.60E-12	10%	2.04E-12	-70		240	400							
7	OH	(E)-1,2-Dichloroethene	Cl/C=C/Cl	2.30E-12	10%	1.06E-12	-230		240	720							
8	OH	1,1-dichloroethene	ClC(=C)Cl	1.05E-11	10%	2.17E-12	-470		240	750							
9	OH	trichloroethene	ClC=C(Cl)Cl	2.00E-12	26%	3.00E-13	-565		230	300							
10	OH	tetrachloroethene	ClC(=C(Cl)Cl)Cl	1.60E-13	26%	3.50E-12	920		290	420							
11																	
12	NO3	chloroethene	ClC=C	3.90E-16	26%	1.40E-13	1740		266	367							
13	NO3	(Z)-1,2-dichloroethene	Cl/C=C\Cl	1.50E-16	40%												
14	NO3	(E)-1,2-Dichloroethene	Cl/C=C/Cl	1.11E-16	75%	7.80E-13	2640		293	373							
15	NO3	1,1-dichloroethene	ClC(=C)Cl	1.53E-15	25%												
16	NO3	trichloroethene	ClC=C(Cl)Cl	3.70E-16	49%	8.10E-13	2295		278	367							
17	NO3	tetrachloroethene	ClC(=C(Cl)Cl)Cl	-							Information is available but no recommendations.						
18																	
19	O3	chloroethene	ClC=C	2.45E-19	100%												
20	O3	(Z)-1,2-dichloroethene	Cl/C=C\Cl	4.90E-21	30%	1.22E-15	3704		315	382							
21	O3	(E)-1,2-Dichloroethene	Cl/C=C/Cl	8.00E-20	30%	9.30E-16	2788		301	394							
22	O3	1,1-dichloroethene	ClC(=C)Cl	4.90E-21	41%												
23	O3	trichloroethene	ClC=C(Cl)Cl	-		1.60E-10	7842		334	388	Estimated k(298)= 5.96E-22, extrapolated from 334 (Factor of 17.1 higher).						
24	O3	tetrachloroethene	ClC(=C(Cl)Cl)Cl	-		6.00E-04	14060		369	409	Estimated k(298)= 1.94E-24, extrapolated from 369 (Factor of 8761.5 higher)						
25																	
26	Cl	chloroethene	ClC=C	1.27E-10	10%												
27	Cl	(Z)-1,2-dichloroethene	Cl/C=C\Cl	9.10E-11	25%	6.30E-12	-799		290	340							
28	Cl	(E)-1,2-Dichloroethene	Cl/C=C/Cl	7.40E-11	24%	2.40E-12	-1018		298	348							
29	Cl	1,1-dichloroethene	ClC(=C)Cl	1.34E-10	25%	1.07E-11	-758		298	348							
30	Cl	trichloroethene	ClC=C(Cl)Cl	7.70E-11	25%	4.60E-12	-846		298	348							
31	Cl	tetrachloroethene	ClC(=C(Cl)Cl)Cl	6.60E-11	15%												
32																	

Figure 2. Screenshots illustrating use of the "All Rxns" macro to get listings or recommended rate parameters.

There are a number of reactions with rate constant information in our database but for which we have made no quantitative recommendations. These could be reactions that have only upper or lower limit measurements or where the data are too uncertain to make recommendations. Although these "Rxn" macros do not output kinetic information for such reactions, they are included in the listings with the comment "Information is available but no recommendation". The user can see what information is available for this and the other reactions of the compound by running an "info on compound" macro, discussed below.

There are also a small number of cases where we found no measurements of the rate constant around 298K, but where temperature dependence parameters have been derived from rate constants measured at higher temperatures. If the extrapolation is too great to be considered experimental data (see McGillen et al, 2019 for a discussion of the review methods), then we make no recommendation for k(298), while still making recommendations for the temperature dependence parameters in the measured temperature range. However, there may be cases where rate constants estimated by extrapolation, however uncertain, may be useful, so in those cases a comment is added giving the extrapolated rate constant as an estimate. The

6

comment also indicates the extent of the extrapolation, in terms of both temperature and rate constant, that was needed. Examples of this are shown on Figure 2b.

Data for a Single Compound in the "Selected Compound Info" Worksheet

The listings in the "Compound List Info" sheet discussed above do not contain all the information in the database about the reactions. Information not shown include the reference citation for the recommendation and associated comments about it, other entries in the database about the reaction that were considered when making the recommendation, and also information about reactions for which we made no recommendations. This is too much information to include in listings of multiple reactions without excessive clutter.

Instead, users can run macros to output to the "Selected Compound Info" worksheet essentially all information in the database about a selected compound and its reactions. The compound can be selected in several ways, as follows. The same output is produced regardless of how the compound is selected.

- From the "Find Compounds" sheet: After doing a search, select a row listing the compound of interest, and run the "Info on Compound in Selected Row" macro, as shown in **Figure 3**a.

- From the "Compound List Info" sheet: Select a reaction in the sheet of the compound of interest and run the "Get Info on Selected Compound" macro. It doesn't matter what the reaction is – only the compound.

- From the "Compounds" or "k-Data" sheet: Select a row listing the compound or the reactions of a compound and then run the "Get Info on Compound" control in the sheet.

Examples of output in the "Selected Compound Info" sheet are shown in **Figure 3**b and **Figure 4**. The primary identifiers and type information for the compound are given at the top of the sheet and additional information, such as alternate names and aliases that can be used in searches and structural characteristics are at the bottom. Information about the reactions is given in a table below the primary identifiers. Each row is an entry in the database concerning kinetic information from a reaction of the compound from some study or review, including, but not limited to, the entry with the recommended kinetic parameters. In addition to the same rate parameter information that is output in the "Compound List Info" sheet for multiple reactions, the rows give a citation number and all the comments associated with the entry. The entries with recommended data for each reaction are shown in bold font, and the "note" column indicates which of the entries have the recommendations. Note that some reactions will only have upper or lower limits or approximate data, and there may not be recommendations.

For most reactions with temperature-dependent information, we obtained the recommended temperature-dependent parameters from the literature. However, for some reactions we found that existing evaluations did not incorporate all available results and we had to fit the data ourselves to get temperature-dependent parameters that reflect all the available data. An example is the reaction of OH with methoxymethane, where we obtained the temperature dependence by fitting 102 data points from 9 references [see also Figure 4 of McGillen et al. (2019)]. In those cases, the output includes tables that give the individual rate constants that were fit and also the reference that they came from. **Figure 4** shows portions of the "Selected Compound Info" output showing the information for this reaction.

(a)

Find compounds on database using various lookup methods

Part of Name or Alias ->	chloroethene	
Atom Nos (e.g., C2H5Cl) ->		
Compound type (from table C to right) ->		<- Give type code no. (See Table C)
Structure type (fill in table S to right)	(Fill in Table S)	

Clear All

6 <-Number of Compounds listed

Clear List | List k's -> | All Rxns | OH Rxns | NO3 Rxns | O3 Rxns | Cl Rxns | Info on compound in selected row

2. Press

"List k's" results in "Compound List Info" sheet
Selected compound info in "Selected Compound Info" sheet

No.	Name	Smiles	Alternative Structure ID
55	chloroethene	ClC=C	CH2=CH-Cl
79	(Z)-1,2-dichloroethene	Cl/C=C\Cl	Cl-^CH=CH-vCl
80	(E)-1,2-Dichloroethene	Cl/C=C/Cl	Cl-^CH=CH-^Cl
81	1,1-dichloroethene	ClC(=C)Cl	CH2=C(Cl)-Cl
104	trichloroethene	ClC=C(Cl)Cl	Cl-CH=C(Cl)-Cl
121	tetrachloroethene <-- 1. Select	ClC(=C(Cl)Cl)Cl	Cl-C(Cl)=C(Cl)-Cl

(b)

tetrachloroethene

Smiles	ClC(=C(Cl)Cl)Cl
MG ID	Cl-C(Cl)=C(Cl)-Cl
Atoms	C2Cl4
Type	Unsaturated halogen-containing hydrocarbon (halogens separated)

Ask for Cmpd No.

Rxn with | $k(T) = A * \exp(-B/T) * (T/300)^n$ | Rate constants in cm3 molec-1 s-1

Rxn with	k298	Unc.	A	B	n	Low T	High T	Ref	Note
OH	1.60E-13	26%	3.50E-12	920		290	420	1	Recommended.
OH	1.70E-13	20%	4.70E-12	990		296	720	2	Not the recommendation.
NO3	1.00E-16							3	Upper limit.
O3			6.00E-04	14060		369	409	4	Recommended. Estimated k(298)= 1.94e-24. This was extrapolated from 369 K, where k(369) is a factor of 8761.5 higher. Unusual temperature dependence has not been explained.
O3	1.60E-24	38%						3	Rejected
Cl	6.60E-11	15%						5	Recommended.

Reference List

1. IUPAC 2017. Task group on Atmospheric Chemical Kinetic Data Evaluation. Evaluated data sheets available at http://iupac.pole-ether.fr/. Datasheets last updated November 1.
2. NASA 2015.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies. Evaluation Number 18," JPL Publication 15-10, Jet Propulsion Laboratory, Pasadena, California, October https://jpldataeval.jpl.nasa.gov
3. Calvert, J. G., J. J. Orlando, W. R. Stockwell, and T. J. Wallington. 2018 "The Mechanisms of Reactions Influencing Atmospheric Ozone," Oxford University Press, New York. Last entries dated 12/2018. Accessed early January, 2019.
4. Leather, K.E., McGillen, M.R., Ghalaieny, M., Shallcross, D.E., and Percival, C.J. 2011. "Temperature-Dependent Kinetics for the Ozonolysis of Selected Chlorinated Alkenes in the Gas Phase", Int. J. Chem. Kinet. 43. 120-129. https://doi.org/10.1002/kin.20533
5. Herath, Thushani N., Eric C. Clinch, Ivan Orozco, Erin L. Raign, and Paul Marshall. 2016. "Relative Rate and Product Studies of the Reactions of Atomic Chlorine with Tetrafluoroethylene, 1,2-Dichloro-1,2-Difluoroethylene, 1,1-Dichloro-2,2-Difluoroethylene, and Hexafluoro-1,3-Butadiene in the Presence of Oxygen", J. Phys. Chem. A 120. 7311-7319. https://doi.org/10.1021/acs.jpca.6b05305

Alternative names for tetrachloroethene

Name	1,1,2,2-tetrachloroethene
Name	1,1,2,2-tetrachloroethylene
Name	1,1,2,2-tetrakis(chloranyl)ethene
Name	perchloroethylene
Name	tetrachloroethene
Name	tetrachloroethylene
Name	C2Cl4
Name	C2Cl4
Smiles	ClC(=C(Cl)Cl)Cl
CAS	127-18-4
InChi	InChI=1S/C2Cl4/c3-1(4)2(5)6
iKey	CYTYCFOTNPOANT-UHFFFAOYSA-N
SDBID	C127-18-4
MGID	Cl-C(Cl)=C(Cl)-Cl
DMS	CL4-ETHE

No.	Structure Information for tetrachloroethene
1	Number of separated double bonds
2	Number of hydrocarbon groups with 2 or more non-hydrocarbon groups bonded to them
4	Number of Cl atoms

Figure 3. Screenshot illustrating the use of the "Info on Compound" macro to get information in the database about a given compound and its reactions.

	A	B	C	D	E	F	G	H	I	J
1						methoxymethane				
2	Smiles	COC								
3	MG ID	CH3-O-CH3								
4	Atoms	C2H6O								Ask for Cmpd No.
5	Type	Ether								
6										
7	Rxn				$k(T) = A * exp(-B/T) * (T/300)^n$					Rate constants in cm3 molec-1 s-1
8	with	k298	Unc.	A	B	n	Low T	High T	Ref	Note
9	OH	2.83E-12	10%	1.02E-12	-308	2.09	195	1470		Recommended rate parameters fitted manually. See below.
10	OH	2.70E-12	10%	9.70E-13	380		256	333	1	Not the recommendation.
11	OH	2.80E-12	20%	5.70E-12	215		230	300	2	Not the recommendation.
12	OH	2.70E-12	15%	5.70E-12	222				3	Not the recommendation.
13										
14	NO3	1.87E-16	60%	8.99E-13	2526		273	373		Recommended. Rate constants as a function of temperature derived by averaging k298 from [Ref#4,1], [Ref#4,2], [Ref#5] and T-dependences for T=273-373 from [Ref#4,2]. A adjusted by a factor of 0.661 to fit k298.
15	NO3	1.95E-16	40%						4	Not the recommendation.
16	NO3	2.60E-16	40%	1.36E-12	2526		273	373	4	Not the recommendation.
17	NO3	1.07E-16	10%						5	Not the recommendation.
18	NO3	2.55E-16	50%	2.30E-12	2714		273	373	3	Rejected
19										
20	Cl	1.80E-10	15%	1.80E-10	0		200	500	3	Recommended.
21										
22	T	k(T)	Unc	k(fit)	%Err				Ref	Manualy fitted data for reaction with OH
23	295	2.95E-12	4%	2.80E-12	-5				6	
24	332	3.40E-12	4%	3.19E-12	-6				6	
25	377.5	3.81E-12	4%	3.73E-12	-2				6	
26	442	4.52E-12	4%	4.60E-12	2				6	
27	292	2.44E-12	30%	2.77E-12	13				7	
				...						
115	750	1.17E-11	20%	1.04E-11	-11				12	
116	800	1.27E-11	20%	1.16E-11	-8				12	
117	850	1.43E-11	20%	1.29E-11	-10				12	
118	298	2.80E-12		2.83E-12	1				2	
119	256	2.10E-12	15%	2.44E-12	16				1	
120	269	2.40E-12	15%	2.55E-12	6				1	
121	280	2.50E-12	15%	2.65E-12	6				1	
122	297	2.70E-12	15%	2.82E-12	4				1	
123	313	2.90E-12	15%	2.98E-12	3				1	
124	333	3.00E-12	15%	3.20E-12	7				1	
125										
126										Reference List
									1	Kawade, Monali, Asmita Sharma, D. Srinivas, Ankur Saha, Hari P. Upadhyaya, Awadhesh Kumar, and Prakash D. Naik. 2018. "Rate Coefficients

Figure 4. Selected screenshot showing results of running the "Info on Compound" macro on a compound where we manually fitted the data to obtain the temperature dependence recommendations. Only selected portions of the screen are shown.

Obtaining Information from the Primary Data Tables Without Using Macros

It is not necessary to enable and use the macros discussed above to obtain information from this database. The primary information in the database is contained in tables in the "Compounds", "k-Data", "KT-Data", "Refs" and "Names DB" sheets, which are described in detail in the next section. Examples of how desired information can be extracted manually are as follows:

- The summary of recommended rate constants and availability of other data output by the macros in the "Compound List Info" sheet can be obtained from the appropriate columns in the "Compounds" sheet (see **Table 2**, below).

- All the kinetic information in a database (other than data points used in fitting) that are output in the "Selected Compound Info" sheet are in the "k-Data" sheet (see **Table 3**,

below), and the information output about the compound itself are in the "Compounds" sheet (see **Table 2**).

- A more complete listing of recommended rate constants, including comments and reference codes, can be obtained from the "k-Data" sheet by extracting all records (rows) that contain 1, 2, or 3 as "Type" codes (see **Table 3**, below)

- To find information about a compound, including recommended rate constants, go to the "Compounds" sheet then do a search for the compound primary name, SMILES, InChi, or InChiKey codes. (If you only have the SMILES code, the "OpenBabel" program can be used to derive the canonical SMILES code [use the "Make Dative Bond" option]). Note that the "Compounds" sheet does not have all the names for the compound, so if a search by a name is unsuccessful, you can go to the "Names DB" sheet, search for the name or designation, then read the compound number from the row with the name. The compound number will locate it on the "Compounds" sheet.

Using either of these methods requires a knowledge of the contents and formats of the data in the primary data tables. These are given in the various tables in the following section.

Contents of the Database File

Table 1 lists all the worksheets in the database file and briefly summarizes their contents. The "Info", "Find Compounds", "Compound List Info" worksheets are primarily for the user's convenience and were discussed above. The other sheets, which contain the database itself or information it uses, are summarized below.

The "Compounds" sheet contains gives information about all the compounds whose rate constants are given in the database. It also has summaries recommended rate constants and temperature dependences for the four types of reactions (OH, O_3, etc.) for the compound where data are available, though not the additional information about the reactions that is in the "k-Data" or "kT-Data" sheets. The entries in this sheet are ordered by compound number, but the macros should operate properly if they are rearranged, as long as there are no blank rows before the row with the last compound. The columns of data are summarized on **Table 2** and the types of recommended kinetic data output are those with labels underlined on **Table 3**.

Table 1. Worksheets in the Database Excel File

Sheet Name	Description
Information worksheets – not user modifiable	
Info	Information about the database and summary of worksheets and macros
Codes	List of various codes used in the data worksheets
Worksheets for User Interactions - Modifiable by macros	
Find Compounds	Use various methods to search for and list compounds in the database
Compound List Info	Lists of recommended kinetic parameters for all or selected types of reactions for all or selected compounds
Selected Compound Info	All information in the database about a selected compound and its reactions
Worksheets with primary data – not user modifiable	
Compounds	Information about the compounds
k-Data	Primary kinetic parameter records; one record with information for a single reaction from a single reference or review.
kT-Data	Rate constants at different temperatures used for manual fitting of temperature-dependence parameters for some of the reactions.
Refs	List of all the references cited in the "k-Data" or "kT-Data" sheets
Names DB	List of all the names, aliases, and other identifiers that can be used to search for compounds.

The "k-Data" sheet contains the table of kinetic results for single reactions from various sources. These sources could include the primary experimental literature, various reviews, and our own assessments of the data. If there is only a single study or review of the reaction in our database, then there may be only one entry for the reaction in this table, and, if the results are considered reasonable, the data in that entry is recommended for use. However, if there are multiple studies or reviews in the database, then there can be multiple entries for a reaction, though no more than one can be recommended for use. The entries are ordered first by oxidant then by compound number. Note that the macros assume that entries for each reaction are together, and may not operate properly if the data in this table are reordered, or records are inserted or deleted. The columns of data on this sheet are summarized on **Table 3**, and the codes used to categorize the data are summarized on **Table 4**.

The "kT-Data" sheet contains the rate constants at different temperatures for reactions where data were manually fitted to obtain our recommended temperature-dependence parameters. The entries are in no particular order except that all the data for a single reaction must be together, with no embedded blank rows, or the "info on compound" macros will not output all the data for the applicable reactions. The columns of data in this sheet are summarized in **Table 5**.

The "Refs" sheet contains a listing of all the references cited in the "k-Data" and "kT-Data" sheets, defining the codes that are used in these sheets. The ordering of references is essentially arbitrary, though "Average" and "Fitted" should come first, and there should be no embedded blank rows. "Average" refers to averages that we derived from data from various entries in "k-Data" and "Fitted" refers to kinetic parameters we derived by manual fitting from data in the "kT-Data" sheet. The contents of this sheet are summarized in **Table 6**.

The "Names DB" sheet associates various names, aliases, and designations to compound ID numbers, to aid searching the database. The entries are ordered by compound number and then type of designation, with the primary name given first. However, the order of entries in the database is not important as long as there are no embedded blank rows. The information in this sheet is summarized in **Table 7**.

The "Codes" sheet contains three tables defining codes that are used in the data tables. It is not strictly part of the database, but may be modified in future versions if codes are added or modified. The three tables are as follows: The General Structural Information table contains the codes used in the "Compounds" sheet to specify structural characteristics of the molecule. This is the same as "Table S" in the "Find Compounds" sheet. The Atoms table list the atomic numbers, weights, and names for all the types of atoms of compounds in the database (but no others). These are given in the order when output as atom number strings such as "C6H6O". The Use Codes table defines the "Type" codes used in the "k-Data" and "Compounds" sheets, as also given in **Table 4**.

Table 2. Columns of data in the "Compounds" sheet, giving information about compounds in the database.

Column Label(s)	Contents
Cmpd No	A positive integer uniquely identifying the compound in the database. This is referred as to the compound's "ID number" or "C.no" in the subsequent discussion. This starts at 1 (for methane) but is in no particular order, and there are some gaps that occurred when duplicate entries were removed. Note that the number can refer to different geometrical isomers if there are data for the individual isomers, or to mixtures of isomers if the isomers are unspecified when the measurement is reported.
Primary Name	The primary name for the compound that is given when data are output. This and other names are in the "names DB" sheet, described below.
Smiles	A unique canonical SMILES string giving the structure of the compound. Note that many compounds can have more than one valid Smiles string, so in this case unique canonical SMILES strings are created using the OpenBabel software (O'Boyle et al, 2011) using the "Make Dative Bond" option.
InChi; InChiKey	InChi and InChiKey strings provide another method to uniquely identify compounds without ambiguities associated with SMILES (IUPAC, 2019). InChi strings uniquely identify each compound, isomer, or mixture of isomers, and InChiKey strings are shortened versions of InChI that are unique in almost all cases, including all compounds in the current database. These can be obtained given the structure using a variety of software packages such as OpenBabel, ChemSketch, ChemDraw, etc.
Mwt; Atoms	Columns giving molecular weights and numbers of each type of atom
Alternative Structure ID	This gives the structure of the compound in the SAPRC mechanism generation system (MechGen, Carter, 2019), This provides a structural identifier that might be easier for some chemists to understand than SMILES, and also allows MechGen to be used as a tool to systematically get the compound type and structural information for the compound, discussed below. This is blank and the compound and structural information may be missing if the compound is not supported by MechGen.
Recommended kOH, kO3, etc.	Columns giving the recommended kinetic parameters for the reactions of the compound with OH, O_3, NO_3, and Cl. This is a subset of the kinetic information about this reaction in the "k-Data" sheet, described in **Table 3**, with its inclusion in the "Compounds" sheet being indicated by the label being underlined. This includes the relevant kinetic data but excludes the comments and reference information that are in the "k-Data" sheet.

Table 3. Columns of information in the "k-Data" sheet, giving information on kinetic data from the various sources.

Label(s) [a]	Contents
Rxn	A code giving the type of reaction, which could be "OH", "NO3", "O3" or "Cl".
C.no.	The ID number for the compound that is reacting. This must correspond to the ID number given in the entry for the compound in the "Compounds" sheet.
Name	The name used for the compound as used in the work where the data are reported, or the primary name if the entry gives recommendations from this work. Each name is also associated with the compound number in the "Names DB" sheet.
Ref, n	A reference code that uniquely identifies a reference citation or source of the data. This could refer to a paper reporting the measurement data, a review, or a code indicating that the entry is a result of our evaluations. These codes are defined in the "Refs" sheet. In a very small number of cases, there is more than one entry for this reaction with the same reference (generally a review). In these cases, a nonblank "sequence number" (1, or 2) is included in the "n" column so that each row will have a unique set of {Rxn, No, Ref, n} values. It is not needed by any of the current macros.
Type	A code indicating the type of record and its role in the database or database development, as summarized on **Table 4**. Note that this column on the "Compound" sheet can only have codes used for recommended data (1-3), or "info" to indicate that there are non-recommended data available in "k-Data".
k298	The rate constant at 298K in cm^3 $molec^{-1}$ s^{-1}. May be blank if not measured or recommended.
Unc	The estimated relative uncertainty of k(298), either taken from the citation or estimated by us. Note that these uncertainty assignments were not always done consistently in the current database, so these should be used with caution.
A, B, n	The temperature dependence parameters, where $k(T)=A \exp(-B/T) (T/300)^n$, where A is in cm^3 $molec^{-1}$ s^{-1}, B is in deg K, and n is unitless. May be blank if no data or recommendation.
Low, High	The low and high end of the temperature range from which the A, B, and n parameters were derived or are valid. Blank if no such parameters or if not entered from the citation used. Always present if the record is used to give recommended temperature dependence parameters.
n, comment lines	The number of comments to be output when information is given about the record, or blank if no comments. The comments themselves are in the subsequent columns.

[a] Labels for columns of data that are copied from the "k-Data" sheet to the "Compounds" sheet for recommended kinetic parameters are underlined.

Table 4.   Summary of type codes used in the "k-Data" and "Compounds" sheets.

Code(s)	Description
1, 2, or 3	"Compounds" or "k-Data" sheets: This record contains recommended data. A code of 1 or 2 means that the k(298) value is recommended, while a code of 2 or 3 means that the temperature dependence parameters are recommended.
Info	"Compounds" sheet only: We have no recommendations for this reaction but kinetic information is available for it in the "k-Data" sheet.
blank	"Compounds" sheet: There is no information about this reaction "k-Data" sheet: Same as "0"; see below.
	The following codes are only in the "k-Data" sheet
0	This entry contains results considered valid for information purposes, but which are not the recommendations. In some cases these data were included in the averages used to derive the recommendations, and in others the results are included in a review that we used.
-1	These data are not recommended for use and the temperature dependence is judged to be invalid, but the k(298) value is considered appropriate for averaging to obtain the recommended value.
a	This is a review but is not used in our recommendations because there are data (usually from more recent studies) that the review did not consider
x, ?, ~	Rejected, questionable or approximate data
Est	The k(298) is an estimate given in a review. Not recommended for use because our database is only for experimental data.
>, >>, <, <=	The k(298) given is an upper or lower limit.

Table 5.   Contents of the "kT-Data" sheet, containing rate constants at different temperatures that were used for manual fitting of temperature-dependence parameters.

Label	Description
Rxn	A code giving the type of reaction, which could be "OH", "NO3", "O3" or "Cl".
C.no	The ID number for the compound that is reacting.
RefID	The reference ID code for the source of the rate constant. These are the same as used on the "k-Data" sheet and are defined in the "Refs" sheet.
T	The temperature of the rate constant, in K.
k	The rate constant at that temperature in cm^3 $molec^{-1}$ s^{-1}.
Unc	An estimate of the uncertainty in cm^3 $molec^{-1}$ s^{-1}.

Table 6. Contents of the "Refs" sheet listing the references cited in the database

Label	Description
Code	The code used for the reference in the "k-Data" and "kT-Data" tables. Note that the special code "Avg" is used to refer to averages of data from other entries that we derived and "NA" refers to cases where we fitted the data using data on the "kT-Data" tables. The others refer to various reviews or publications reporting the measurements. The numbers or codes are arbitrary and have no particular meaning.
Short Citation	The citation in the first authors (year) format. An "a", "b", etc. is appended if there is more than one citation with the same first authors and year.
Full Citation	The full citation for the reference, in a format suitable for a journal.
Link	A URL where the citation can be obtained, usually a link to it at doi.org. Blank if the URL has not been determined.

Table 7. Contents of the "Names DB" sheet used for name lookups of compounds

Label	Description	
Name	The name, alias, or designation.	
C.no	The compound number associated with the name or designation. Must match a compound number in the "Compounds" sheet.	
Type	The type of designation, as follows:	
	Name	A name or alias, including the primary name and other names used in publications reporting the kinetic data, or other names found for the compound in the NCI database.
	iKey	The InChiKey string. This is given for all compounds
	CAS	The CAS number for the compound if we determined this. Note that we have not attempted to determine CAS numbers for all compounds, and some compounds do not have well-defined CAS numbers that can be used in searches.
	MGID	The structure identity string used for this compound by the SAPRC mechanism generation system (Carter, 2019)

References

Carter, W. P. L. (2019): "Gateway to the SAPRC-16 Mechanism Generation System," Available at http://mechgen.cert.ucr.edu/.

IUPAC (2019): "The IUPAC International Chemical Identifier (INCHI)," web site at https://iupac.org/who-we-are/divisions/division-details/inchi/. Copyright 2019.

McGillen, M. R., W .P. L. Carter, W. Mellouki, J. J. Orlando, B. Picquet-Varrault, and T.J. Wallington (2019): "Database for the kinetics of the gas-phase atmospheric reactions of organic compounds," submitted to Earth System Science Data.

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T. and Hutchison, G. R. (2011): "Open Babel: An open chemical toolbox," Journal of Cheminformatics, 3(1), doi:10.1186/1758-2946-3-33, 2011.