Tutorial UFZ-LSER Database

1. Searching experimental compound descriptors in the database

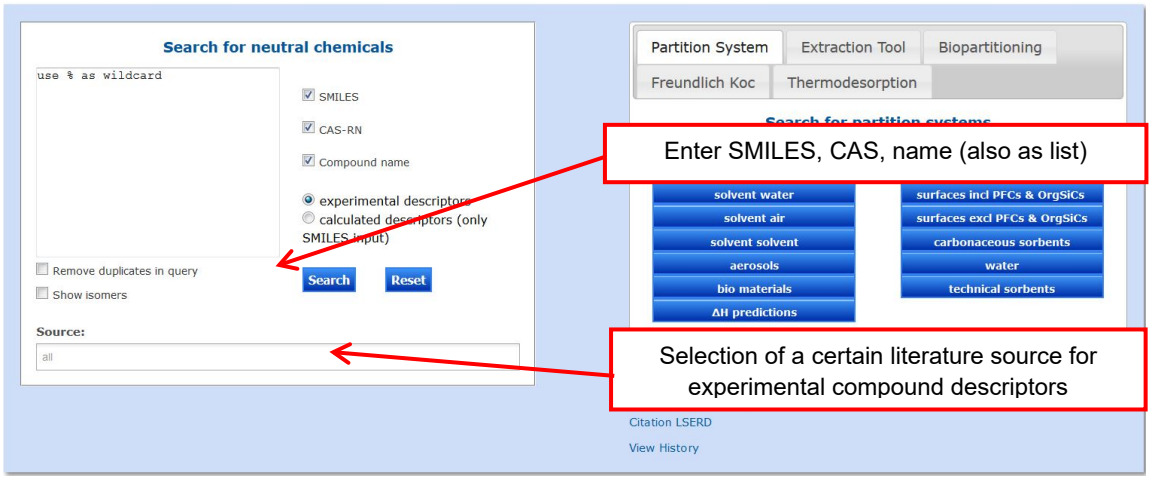


Figure 1: Search of experimental compound descriptors

In the field „Search for neutral chemicals“ experimental compound descriptors for single compounds or lists of compounds can be searched for. Queries can be performed using SMILES, CAS-numbers and compound names[1], searching for a list of compounds all three formats can be used at the same time. The wildcard symbol „%“ can be used to perform a less specific query. The query „%benzene%“ e.g. shows all compounds, which include the term „benzene“ within their name. Results are shown in a table below the query form. Using the filter function above the results table (see Figure 2), query results can further be reduced. For example, using „fluoro“ as a search term will reduce the table automatically to compounds, which include the term „fluoro” (in this case, it is not necessary to click the button „Search“ again). Using these options, one can find compound classes with similar functionalities and structural properties (e.g. search „propyl“) in the database. The filter function can also be applied to CAS numbers, SMILES and literature sources.

[1]Often the database contains only trivial names, thus it is helpful to use the CAS number or the SMILES instead.

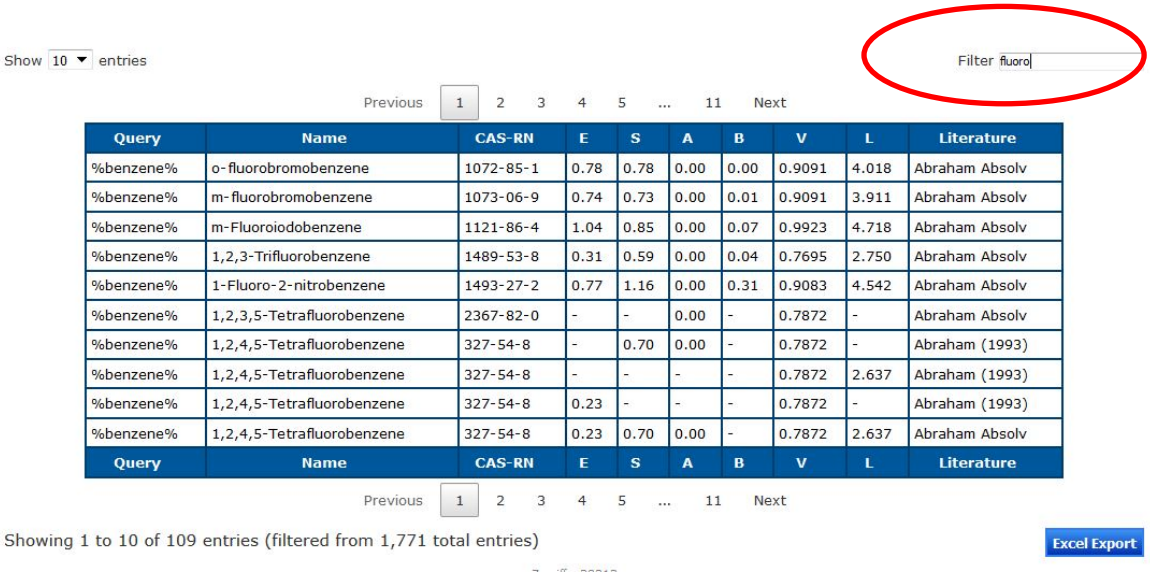


Figure 2: Filter function of the database

Queries can be restricted to certain literature sources. In this case the respective literature source has to be selected. Should all available compound descriptors from literature be displayed, the option „All“ has to be selected in the literature list. If preferred values of the UFZ should be used, the selection „UFZ preselected published values“ has to be clicked. In the latter case only one set of descriptors will be displayed for any compound. Compound descriptors can be exported out of the database with the button “Excel Export”.

2. Calculation of compound descriptors

Compound descriptors can also be calculated by the help of a QSPR model (quantitative structure property relationship). In this case one has to select „calculated descriptors“ in the query. In order to calculate the compound descriptors for a chemical, the SMILES code has to be entered in the query, CAS or name as input is not possible. SMILES codes can be in any format. Note: only compound descriptors of neutral chemicals can be calculated. The time, which is needed to perform the calculation, is given below the input field. Results can be found in a table below the query mask. Please check the application domain before using the calculated compound descriptors. Detailed information on the application domain are given in the excel export file. For each query a CSS (chemical similarity score) is given for the respective compound descriptor. This value contains the information whether the structure of the chemical is inside the application domain or not.

3. Edit function

After finishing a query, experimental descriptors can be edited by the use of the edit function.

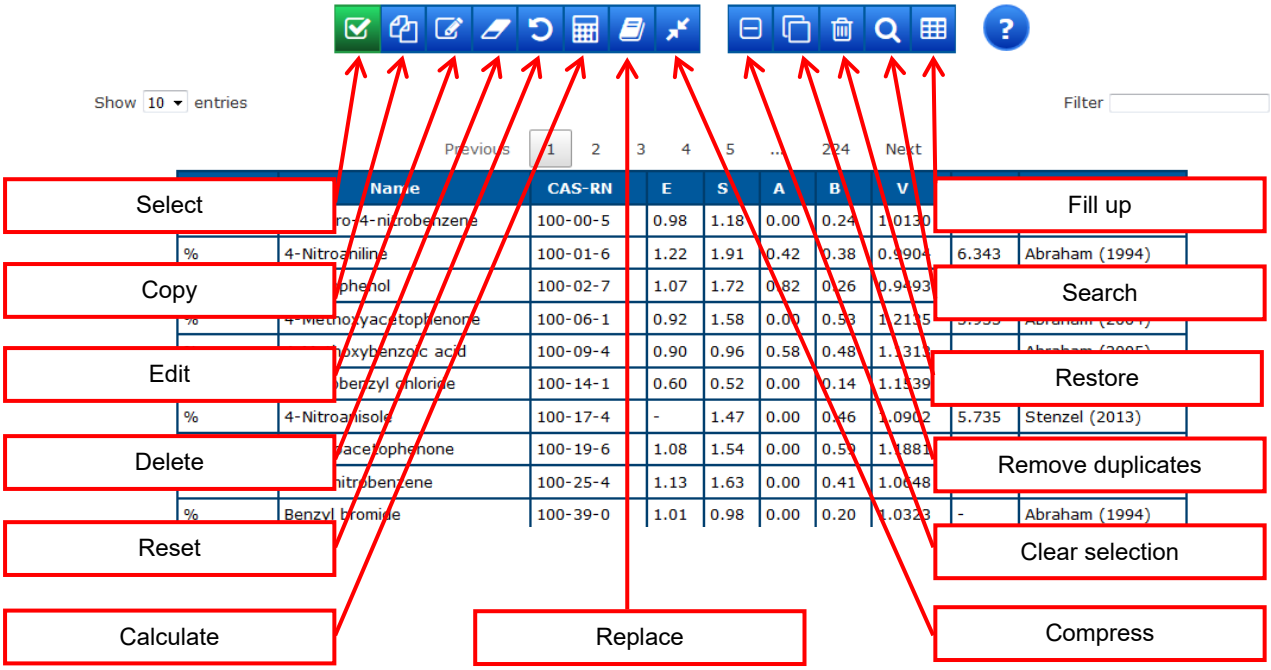
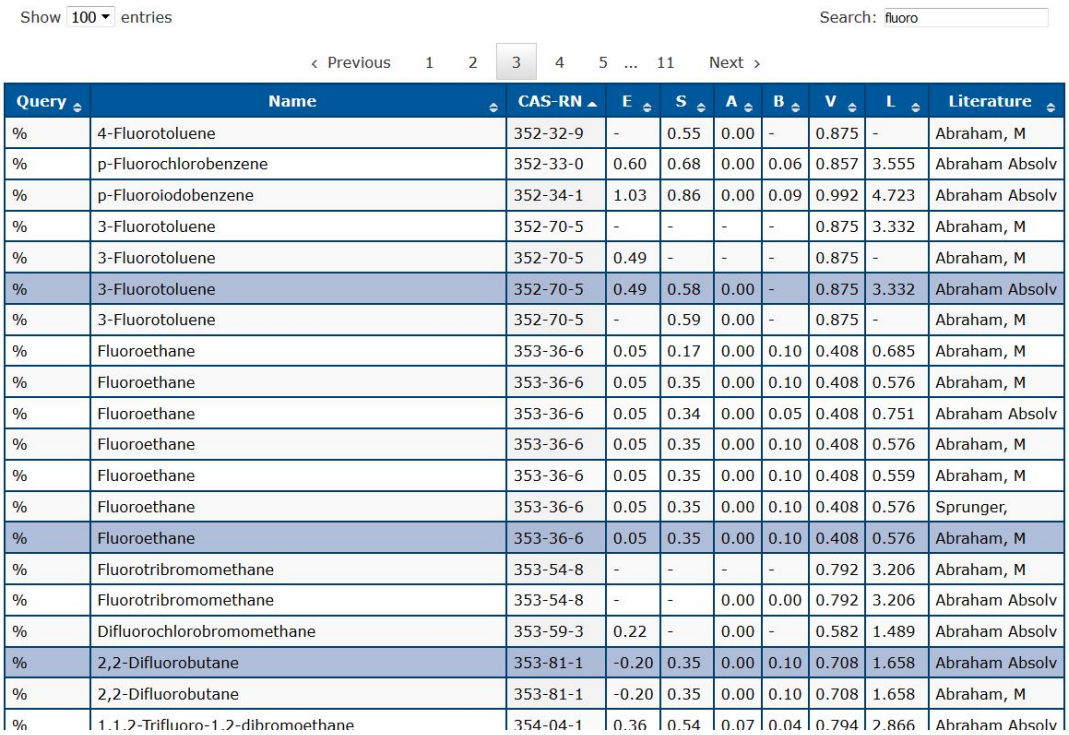


Figure 3: Different options in the edit function

It is possible to edit complete datasets or single compound descriptors (cells in the table). If only single cells should be edited, first the mode has to be selected e.g. “Calculate”, second the cell has to be selected, where the calculation should be performed. If different cells are to be selected for the calculation, the whole area can be marked. In this case all empty cells will be edited and filled by calculated values from the QSPR. All changes will be marked in yellow and information on the calculation is also given in the excel export file.
If complete datasets are to be edited e.g. „Remove duplicates“, all lines will be processed. In this case the results of the query in the table will be reduced to all cases, where at least one descriptor in the whole dataset is different compared to all other sets for the respective chemical. Deleted datasets can be found in the trash and can be restored if needed.

4. Calculation of partition coefficients

For the calculation of partition coefficients all compounds of the query or single compounds can be selected (e.g. if more than one descriptor set is available for one compound). If calculations are only needed for a few compounds, these compounds can be marked by clicking in the table.



Show 100 entries Search: fluoro

‹ Previous 1 2 3 4 5 … 11 Next ›

Query	Name	CAS-RN	E	S	A	B	V	L	Literature
%	4-Fluorotoluene	352-32-9	-	0.55	0.00	-	0.875	-	Abraham, M
%	p-Fluorochlorobenzene	352-33-0	0.60	0.68	0.00	0.06	0.857	3.555	Abraham Absolv
%	p-Fluoroiodobenzene	352-34-1	1.03	0.86	0.00	0.09	0.992	4.723	Abraham Absolv
%	3-Fluorotoluene	352-70-5	-	-	-	-	0.875	3.332	Abraham, M
%	3-Fluorotoluene	352-70-5	0.49	-	-	-	0.875	-	Abraham, M
%	3-Fluorotoluene	352-70-5	0.49	0.58	0.00	-	0.875	3.332	Abraham Absolv
%	3-Fluorotoluene	352-70-5	-	0.59	0.00	-	0.875	-	Abraham, M
%	Fluoroethane	353-36-6	0.05	0.17	0.00	0.10	0.408	0.685	Abraham, M
%	Fluoroethane	353-36-6	0.05	0.35	0.00	0.10	0.408	0.576	Abraham, M
%	Fluoroethane	353-36-6	0.05	0.34	0.00	0.05	0.408	0.751	Abraham Absolv
%	Fluoroethane	353-36-6	0.05	0.35	0.00	0.10	0.408	0.576	Abraham, M
%	Fluoroethane	353-36-6	0.05	0.35	0.00	0.10	0.408	0.559	Abraham, M
%	Fluoroethane	353-36-6	0.05	0.35	0.00	0.10	0.408	0.576	Sprunger,
%	Fluoroethane	353-36-6	0.05	0.35	0.00	0.10	0.408	0.576	Abraham, M
%	Fluorotribromomethane	353-54-8	-	-	-	-	0.792	3.206	Abraham, M
%	Fluorotribromomethane	353-54-8	-	-	0.00	0.00	0.792	3.206	Abraham Absolv
%	Difluorochlorobromomethane	353-59-3	0.22	-	0.00	-	0.582	1.489	Abraham Absolv
%	2,2-Difluorobutane	353-81-1	-0.20	0.35	0.00	0.10	0.708	1.658	Abraham Absolv
%	2,2-Difluorobutane	353-81-1	-0.20	0.35	0.00	0.10	0.708	1.658	Abraham, M
%	1,1,2-Trifluoro-1,2-dibromoethane	354-04-1	0.36	0.54	0.07	0.04	0.794	2.866	Abraham Absolv

Figure 4: Selection of single compounds for the calculation of the partition coefficients. The selected compounds are marked in blue.

After selection of the respective compounds, the partition systems have to be selected.

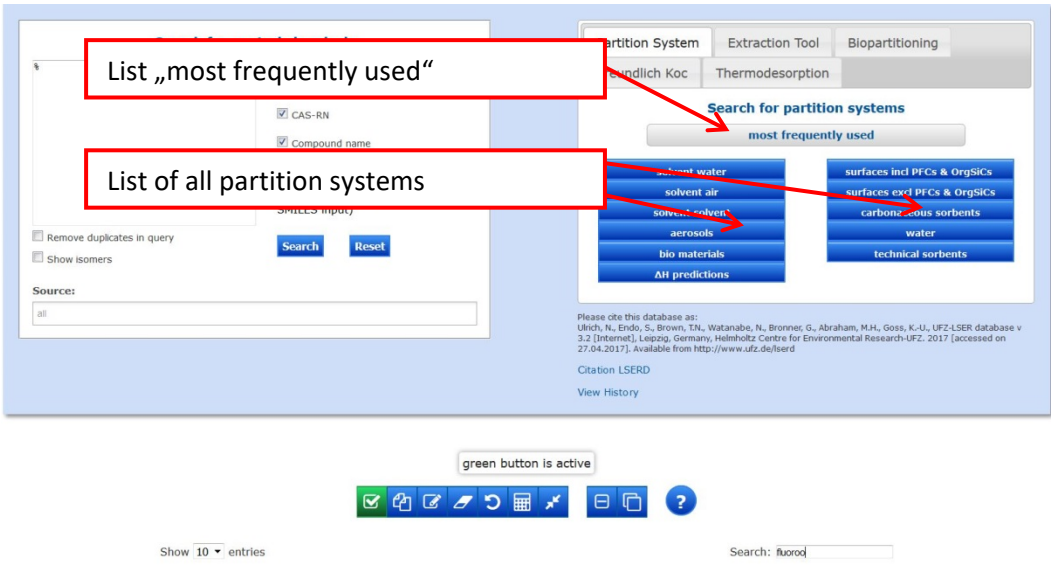
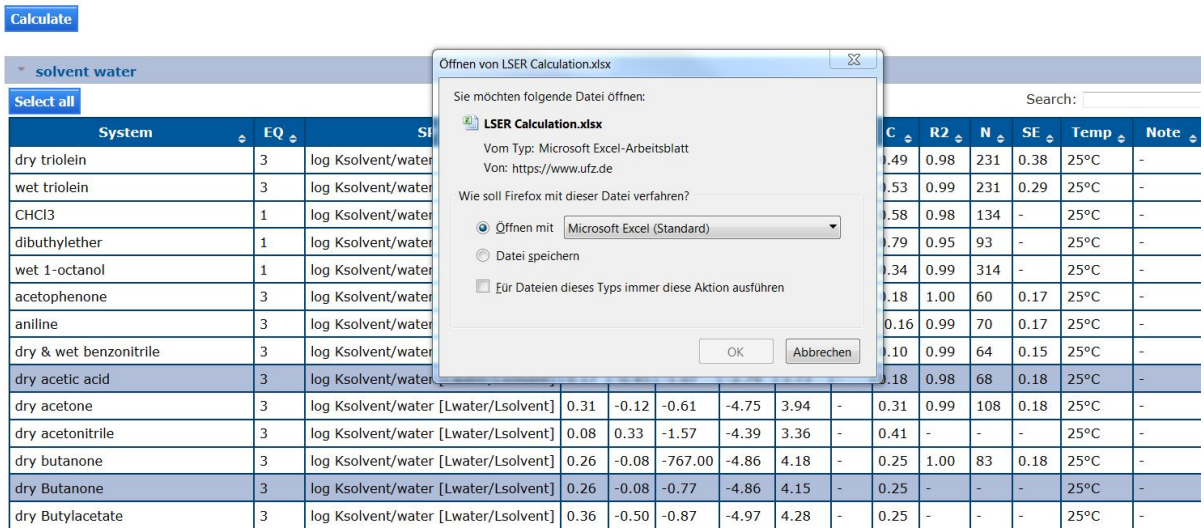


Figure 5: Selection of the partition systems

In the menu „most frequently used“ some often used partition systems are given as e.g. octanol-water, air-water, or different biosystems (e.g. BSA-water (Bovine Serum Albumin) and storage lipid-water). The systems can be selected, clicking the button „Calculate“ starts a calculation of the respective partition coefficients. Then, an excel file is generated, which contains the compound descriptors,

which were used for the calculation, and the calculated partition coefficients with the respective literature sources.

As an alternative to the listed systems in „most frequently used“, one can select from all systems which are stored in the database. First one has to select the category, e.g. „solvent water“, second one has to click on the systems of interest in the table, which one wants to be included in the calculation. After selection of all desired partition systems (also from different categories) an excel file with the results of the calculations is generated by clicking the button „Calculate“.



System	EQ	SI						C	R2	N	SE	Temp	Note
dry triolein	3	log Ksolvent/water						.49	0.98	231	0.38	25°C	-
wet triolein	3	log Ksolvent/water						.53	0.99	231	0.29	25°C	-
CHCl3	1	log Ksolvent/water						.58	0.98	134	-	25°C	-
dibuthylether	1	log Ksolvent/water						.79	0.95	93	-	25°C	-
wet 1-octanol	1	log Ksolvent/water						.34	0.99	314	-	25°C	-
acetophenone	3	log Ksolvent/water						.18	1.00	60	0.17	25°C	-
aniline	3	log Ksolvent/water						0.16	0.99	70	0.17	25°C	-
dry & wet benzonitrile	3	log Ksolvent/water						.10	0.99	64	0.15	25°C	-
dry acetic acid	3	log Ksolvent/water						.18	0.98	68	0.18	25°C	-
dry acetone	3	log Ksolvent/water [Lwater/Lsolvent]	0.31	-0.12	-0.61	-4.75	3.94	-	0.31	0.99	108	0.18	25°C	-
dry acetonitrile	3	log Ksolvent/water [Lwater/Lsolvent]	0.08	0.33	-1.57	-4.39	3.36	-	0.41	-	-	-	25°C	-
dry butanone	3	log Ksolvent/water [Lwater/Lsolvent]	0.26	-0.08	-767.00	-4.86	4.18	-	0.25	1.00	83	0.18	25°C	-
dry Butanone	3	log Ksolvent/water [Lwater/Lsolvent]	0.26	-0.08	-0.77	-4.86	4.15	-	0.25	-	-	-	25°C	-
dry Butylacetate	3	log Ksolvent/water [Lwater/Lsolvent]	0.36	-0.50	-0.87	-4.97	4.28	-	0.25	-	-	-	25°C	-

Figure 6: Selection of partition systems

The excel file contains different sheets: „General information“, „Equations“ and the resulting partition coefficients for the selected partition systems. In the first sheet general information about the ppLFER equations (*poly parameter linear free energy relationships*) are given. In the sheet “Experimental descriptors” the respective compound descriptors, which were used in the calculation, are displayed. In the sheet „Equations“ all equations can be found, which were used for the calculation of the partition coefficients. In the following sheets all calculated partition coefficients are given in combination with the compound descriptors used for the calculation. If experimental compound descriptors were used for the calculation, the literature source of these descriptors is given in the last column. The literature source of the system parameters is given in the second line below the respective system. Please note that **logarithmic** partition coefficients are presented. Units of the partition coefficients are given in the tables. These units can differ for different partition coefficients and can also be different compared to literature values.

Other Applications: after calculation of the partition coefficients for a compound a user will typically continue on his own work. For some standard applications we offer the opportunity to further use these partition coefficients to get further practical and helpful information.

5. Biopartitioning

   With the help of this tool fractions of a chemical in different compartments (serum albumin, muscle proteins, storage lipids, membrane lipids) of an organism can be calculated. First the composition of the organism has to be defined in volume percentage. Compartments such as carbohydrates and minerals are not accounted for in the calculations and are comprised as “Other”. Also an already defined dataset for rainbow trout can be used for the calculation. By clicking the button „Calculate“ an excel export file is generated, which is structured as already known from the other excel files. Two result sheets (Biopartitioning acc. To Eq. 1 and 3) are generated based on the equation type 1 or 3. In these sheets one can find the following information: the logarithmic organism-water partition coefficient (log $K_{organism/water}$), which is the basis for the calculation of the respective fractions in the compartments and the fractions for each compartment.

6. Freundlich K_{oc}

   The tool can be used to calculate the non-linear sorption for soil organic carbon. The pp-LFER equations for Koc provided under “carbonaceous sorbents” are strictly valid only in the concentration for which they have been calibrated. Using the tool “Freundlich Koc” provides the option to calculate Koc values for a range of concentrations from 0.0001 mmol/L to 1 mmol/L based on the assumption that the Freundlich n is typically in the range of 0.8. It is also possible to enter a specific water concentration of interest for which the Koc value is needed. Results are shown in an excel file after clicking the button „Calculate“. In this excel file the concentration of sorbed chemical, the Freundlich coefficient and the Freundlich exponent are given in the sheets Freundlich isotherm Eq. 1 and 3.

7. Extraction Tool

   With the help of this tool extraction efficiencies for selected compounds from water by different solvents can be calculated. First experimental (or calculated) compound descriptors for a chemical need to be generated, second the solvents of interest need to be selected. Selection of solvent can be performed by clicking on the solvent of interest in the displayed table. Three different options are available. Under „Calculate the fraction in the solvent for a given solvent volume“ the volume of the water sample and the volume of solvent can be defined. As an alternative the calculation for different solvent volumes can be performed under „Calculate the fraction in the solvent for different solvent volumes“ to find the best solvent volume for the approach. Results are given in an excel file by clicking the button „Calculate“. In the excel file different information can be found: the first sheet contains information about general equation forms

(analogue to the partition systems). In the sheet „Equations“ all relevant equations used for the calculations are listed. The sheet "Solvent Water“ contains compound descriptors, which are used for the calculation of the partition coefficients and the respective partition coefficients. Results fraction of the compounds in the solvents at equilibrium can be found in the sheet „Calculation of Fraction“.
As an alternative the desired fraction of the analyte in the solvent phase can be defined and the respective solvent volume is calculated for a given water volume. To represent real situations in a better way in the calculations we offer two options where additives such as organic matter or lipids and proteins in the water sample can be considered in the calculations. Concentrations of the respective additives can be entered at the section „Water contains:“. Two more options are available which can be used to optimize the extraction: at the section „Water contains:“ a salt concentration can be entered, in this case the extraction efficiency is calculated considering the salting out effect. Further it is possible to enhance the extraction efficiencies by repeated solvent extraction. In this case it is assumed that a second extraction step is performed by adding the same volume of new solvent to the water sample after a first extraction.

8. Thermodesorption

   With the help of this tool thermodesorption approaches can be optimized. There are different calculators, which can be used for the optimization of the experiments. For a given sampling volume and sampling temperature, the amount of sorbent, which should be used in the experimental approach, can be calculated. Further the 50% breakthrough volume can be calculated for this setup. The second calculator determines the 50% breakthrough volume based on the amount of sorbent used in the approach. This can be also calculated for a certain temperature using the third calculator. Clicking the button „Calculate“ one gets an excel file, which contains all results in the tab Thermodesorption for four sorbents: Tenax TA, Chromosorb 106, Poropak N, and Carbopak F.

9. Blow down

   This tool should enable the user to calculate analyte loss during pre-concentration of a solvent extract in a nitrogen gas stream. First, one has to search for the experimental descriptors of all analytes in the database or use the calculated descriptors from the QSPR approach. Second, one has to select the solvent of interest and enter the volume of the sample. Third, the solvent volume after blow down has to be entered. An excel sheet, which contains all results in the tab “Calculation of maximal loss” is generated by clicking the button “Calculate”.

10. Permeability through a Caco-2 monolayer

This tool calculates the passive apparent permeability through a monolayer of Caco-2 or MDCK cells grown on a permeable filter support. The calculation considers various transport resistances (unstirred water layers (UWL), cytosol, membranes, filter) as well as two parallel pathways: transcellular and paracellular. Details of the approach are described in ref. K. Bittermann and K. U. Goss, “Predicting apparent passive permeability of Caco-2 and MDCK cell-monolayers: A mechanistic model,” *PLoS One*, vol. 12, no. 12, pp. 1–20, 2017, doi: 10.1371/journal.pone.0190319. Deviating from this reference we have simplified the calculation of the paracellular transport by not distinguishing between neutral and ionic species for this pathway. This simplifies the required data input and has little effect on the overall performance of the tool, because setup specific differences in paracellular transport are expected to be much higher than speciation effects. Passive permeability through Caco-2/MDCK cell membranes is calculated from the solute’s hexadecane/water partition coefficient using the solubility-diffusion model and a correlation established between permeability in black lipid membranes and Caco-2/MDCK, as described in ref. C. Dahley, T. Böckmann, A. Ebert, and K. U. Goss, “Predicting the intrinsic membrane permeability of Caco-2/MDCK cells by the solubility-diffusion model,” *Eur. J. Pharm. Sci.*, vol. 195, no. January, 2024, doi: 10.1016/j.ejps.2024.106720.

When users provide the neutral fractions (f_n must be > 0) in aqueous layers for ionizable compounds in the exported Excel file, the tool considers pH-dependent concentration-shift effects in the UWL, filter, and cytosol, as described in ref. C. Dahley, K. U. Goss, and A. Ebert, “Revisiting the pKa-Flux method for determining intrinsic membrane permeability,” *Eur. J. Pharm. Sci.*, vol. 191, no. July, p. 106592, 2023, doi: 10.1016/j.ejps.2023.106592. If the iso-pH method (same pH at the apical and basolateral side) is used, the total UWL thickness is required. If the gradient-pH method (different pH at the apical and basolateral side) is used, the individual UWL thicknesses on both the apical and basolateral side are necessary for the calculation. If only total UWL is provided, a symmetrical distribution is assumed.

The calculated permeability refers to the total concentration of the chemical, assuming that the ionic species does participate in the permeation through the unstirred water layer and the paracellular pathway but not to the transport through the membrane itself. Note that the modeled membrane permeability is based on water and may not accurately represent scenarios where the adjacent phases are not pure water. Possible active transport or retention effects are not considered.

11. C_{free}

    For toxicity assays or the effect of active ingredients of pharmaceuticals it is essential to know the freely dissolved concentration or fraction of a solute, because it is this freely dissolved concentration that can directly and quantitatively be linked to an effect. However, often only the total/nominal concentration is known. A measurement of the freely dissolved concentration or fraction is often not possible or affordable. As an alternative one can calculate the freely dissolved concentration or fraction if one assumes that partition equilibrium is reached (usually this is the case) and if one knows the content of sorbing components (lipids and proteins) and the respective partition coefficients. Here we offer a calculation of C_{free} in blood-plasma based on the assumption that plasma contains 7 vol% albumin and 1 vol% (phospho)lipids. Another option, the calculation of C_{free} in various cellular assays is described in detail in Fischer, F. C., Henneberger, L., Konig, M., Bittermann, K., Linden, L., Goss, K. U., Escher, B. I. (2017). "Modeling Exposure in the Tox21 in Vitro Bioassays." Chemical Research in Toxicology 30(5): 1197-1208. It is important to note that all results for C_{free} only refer to neutral species. C_{free} for ionic species or ionizable chemicals cannot be calculated with these tools.

Citation

Please cite the database as follows:

UFZ-LSER database v 4.0 [Internet], Leipzig, Germany, Helmholtz Centre for Environmental Research-UFZ. 2024 [accessed on 31.07.2024]. Available from http://www.ufz.de/lserd

A ris-file for the import of the citation in EndNote, Citavi, … is available online as a download (link citation LSERD).

FAQ

There is no button for the calculation of partition coefficients.

Please check if you selected any compounds for the calculations. The button is only visible if compounds are selected.

The partition coefficients for my selected partitioning system were not calculated.

This case often occurs if not all experimental compound descriptors are available in the database. In this case it is useful to check whether there are other literature sources with complete compound descriptor sets in the database. Further an alternative equation can be used (e.g. with *vV* term instead of *IL* term if the *L d*escriptor is not available) or data gaps in the descriptor set can be filled with the QSPR tool.

Are the results valid for ions?

No, neither the QSPR model for the prediction of compound descriptors nor the calculation of the partition coefficients are applicable for ions. They should only be used for neutral chemicals.

In some cases results were given in two different excel tabs with the label "acc. to Eq. 1" or "acc. To Eq. 3" – what does this mean?

The results are calculated based on the two different equation types (Eq. type 1 or 3 – see "General information" tab in the excel sheet). Based on the different descriptors used for the calculations some differences may result in the calculated partition coefficients.