PUBLICATION

MUSTANG

A MUltiple Space and Time scale Approach for the quaNtification of deep saline formations for CO_2 storaGe

Project Number:227286

AUTHORS: *Mario Schaffer, Tobias Licha*

TITLE: *A guideline for the identification of environmentally relevant, ionizable organic molecule species*

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007/2013) under grant agreement n° [227286]

Status	AUTHOR VERSION
Date	2014
Publisher	Elsevier
Reference	Chemosphere, Vol. 103, pp. 12-25, 2014

Submitted, accepted and published by ELSEVIER in *Chemosphere*
(http://dx.doi.org/10.1016/j.chemosphere.2013.12.009)

A guideline for the identification of environmentally relevant, ionizable organic molecule species

Mario Schaffer, Tobias Licha*

Citation:

Schaffer, M., Licha, T., 2014. A guideline for the identification of environmentally relevant, ionizable organic molecule species. *Chemosphere* 103, 12–25.

*Corresponding author: Tobias Licha

Tel. +49 551 39 12861

Fax +49 551 39 9379

Tobias.Licha@geo.uni-goettingen.de

http://www.uni-goettingen.de/en/8483.html

Geoscience Centre, Dept. Applied Geology, University of Göttingen, Goldschmidtstr. 3, 37077 Göttingen, Germany

Abstract

An increasing number of organic compounds detected today in the aquatic environment are ionizable and, therefore, partially or permanently charged (ionic) under the pH conditions encountered in these systems. For evaluating their environmental behavior, which strongly depends on the charge state, the identification of functional groups together with their correct assignment of the respective acidic or basic dissociation constants (pK_a) is essential. Despite the growing concern and increasing awareness for ionizable compounds, contradicting and/or confusing information regarding their acid/base properties can be regularly found in the literature, especially when complex structures are encountered. Therefore, we provide a simplified, general, and comprehensive guideline for the identification of ionizable functional groups in organic compounds combined with the correct assignment of their respective pK_a values. Beside the explicit definition of basic terms, several tables with more than 30 of the most frequently encountered ionizable compound classes, including their typical pK_a value ranges are the centerpiece of the proposed procedure. The straight forward application of the guideline is successfully shown for several environmentally relevant compounds as example.

Keywords: Organic compounds, Dissociation, Organic acid, Organic base, Zwitter, pK_a and pK_b value

1. Introduction

Within the last decades, the focus in the field of environmental sciences dealing with organic compounds has been changed from classical non-ionizable compounds, such as chlorinated solvents or polycyclic aromatic hydrocarbons (PAH), to more or less readily ionizable compounds, such as many pharmaceuticals and pesticides (e.g., Højberg et al., 2005; Scheytt et al., 2006; Burgos and Pisutpaisal, 2006; Fatta-Kassinos et al., 2011; Banzhaf et al., 2012; Jin and Peldszus, 2012; Nödler et al., 2013). Franco et al. (2010) emphasized the relevance of ionizable compounds by evaluating a random sample of 1510 chemicals underlying the EU regulation REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals). This study came to the conclusion that about one half (49%) of the evaluated compounds bear structural moieties, which are at least partially ionizable under aqueous, environmentally relevant conditions (pH = 4–10). 27% of the chemicals could be attributed to the group of organic acids, 14% to organic bases, and 8% to zwitter-ionics and amphoteres, respectively. Another analysis from Manallack (2007) with 582 pharmaceuticals even showed that 77.5% of the compounds from this group contain water-relevant ionizable moieties (pK_a = 2–12). An extended study from the same author with 907 pharmaceuticals confirmed the relatively high proportion of ionizable compounds, since 64% of the investigated compounds contained relevant acidic and/or basic groups (Manallack, 2009).

Significant implications for their (pH-dependent) transport behavior in the environment can be derived, since numerous physicochemical properties (e.g., water solubility, hydrophobicity, volatility) and processes (e.g., sorption, partitioning, bioavailability) are strongly associated with the molecule's charge state (e.g., Lorphensri et al., 2007; Tülp et al., 2009; Vasudevan et al, 2009; Schaffer et al., 2012a; 2012b; Niedbala et al., 2013). Therefore, the acid/base properties of a molecule of interest have to be known and considered in order to evaluate or predict its environmental fate in the subsurface (e.g., degradation, mobility, transport). This requires exact

59 knowledge on the location (molecule moiety) and type (protonation, deprotonation) of ionization
60 as well as on its pH-dependent species distribution.
61 The thermodynamic ionization equilibrium is expressed by means of the logarithmic dissociation
62 constants pK_a for acids and pK_b for bases, respectively. Despite the definition of pK_b, merely pK_a
63 is usually used in environmental sciences even for alkaline compounds due to its more direct
64 relation to the pH. However, the acidic or basic character of a functional group is not always
65 obvious and very often not indicated in the literature or in chemical databases. Consequently, the
66 distinction and, if more than one pK_a is given, the assignment of pK_a to a certain substructure can
67 be very challenging. This is especially true for substances containing less common functional
68 groups or forming ionizable molecule moieties due to tautomerization equilibria.
69 The chemical database/property predictor SciFinder (http://scifinder.cas.org), for instance, is a
70 useful tool that provides pK_a values for a vast number of compounds and even indicates the
71 character of the functional groups. For compounds with several ionizable groups, however, only
72 the “most acidic” and “most basic” pK_a values are provided and the location of the respective
73 pK_a is not explicitly allocated. Another example for potentially incomplete data sources is the
74 medicinal or pharmaceutical literature, where often solely physiologically relevant pK_a values
75 are listed (Avdeef, 2003; Prankerd, 2007). As a consequence, a deeper understanding in organic
76 chemistry is required for determining the acid/base behavior of a molecule. This applies
77 especially for the identification of all decisive structures and the correct assignment of the
78 respective pK_a values to these functional groups.
79 This article provides a practical guide for the identification and assignment of ionizable
80 functional groups in organic compounds. A proposed procedure simplifies detecting decisive
81 acid/base properties of the compound of interest and, thus, the determination of charge states and
82 co-existing molecule species for specific pH conditions of the considered aquatic system. This
83 includes the recognition of (partially) ionic compounds and the assignment to the group of
84 anions, cations, or zwitter-ions. For this purpose, a short introduction to the theory of ionization

85 (Brønsted-Lowry theory) is given and the environmentally most relevant molecular structures,
86 which are ionizable under aqueous conditions (pK_a = 0-14), were collected and listed in
87 systematic tables. Furthermore, the suggested procedure was tested on the example of four
88 ionizable pharmaceuticals with different molecular features and acid/base properties.
89

2. Theory - Ionization constants: pK_a *vs.* pK_b

Dissociation constants are compound-specific measures for the location of the chemical equilibrium of acid/base-reactions. Generally, two constants can be defined for organic compounds: 1) the dissociation constant of organic acids K_a and 2) the association constant of organic bases K_b. These thermodynamic constants can be derived from the law of mass action for the ionization equilibrium of a mono-protic organic acid *A-H* and a mono-protic organic base *B* with their respective conjugate base A^- and acid $B\text{-}H^+$ in the aqueous phase with:

$$A\text{–}H + H_2O \rightleftharpoons A^- + H_3O^+ \quad \text{with} \quad pK_s = -\lg\frac{[A^-]\cdot[H_3O^+]}{[A\text{–}H]} \tag{1}$$

$$B + H_2O \rightleftharpoons B\text{–}H^+ + OH^- \quad \text{with} \quad pK_b = -\lg\frac{[B\text{–}H^+]\cdot[OH^-]}{[B]} \tag{2}$$

The lower the pK_a and pK_b, respectively, the stronger is the acidity or basicity of the ionizable molecule. From Eq. (1) and Eq. (2), the degree of dissociation α and the degree of association β, respectively, can be calculated with:

$$\alpha = \frac{[A^-]}{[A^-]+[A\text{–}H]} = \frac{10^{pH-pK_a}}{1+10^{pH-pK_a}} \quad \text{and} \quad 1-\alpha = \frac{[A\text{–}H]}{[A^-]+[A\text{–}H]} = \frac{1}{1+10^{pH-pK_a}} \tag{3}$$

$$\beta = \frac{[B\text{–}H^+]}{[B]+[B\text{–}H^+]} = \frac{1}{1+10^{pH-(pK_w-pK_b)}} \quad \text{and} \quad 1-\beta = \frac{[B]}{[B]+[B\text{–}H^+]} = \frac{10^{pH-(pK_w-pK_b)}}{1+10^{pH-(pK_w-pK_b)}} \tag{4}$$

where pK_w is the negative decadic logarithm of the self ionization constant (protonation capacity) of water. As a result, the pH-dependent species distributions can be derived (Fig. 1). Due to the differently charged species, significant differences in the physico-chemical properties (e.g., log K_{OW}, solubility) associated with a different environmental behavior can be expected for each species (Kah and Brown, 2008; Franco et al., 2010; MacKay and Vasudevan, 2013; Schaffer et al., 2013). The pK_w value of pure water at 25°C is about 14 (groundwater: pK_w = 14.5 at 10°C) and represents the sum of pK_a and pK_b. This relation is important, since only pK_a values are very

111 often provided in chemical data sources (especially for organic bases). Furthermore, it can be
112 seen from Eq. (3) that at pH = pK_a, the fraction of the charged equals the fraction of the neutral
113 species in solution (Fig. 1). Additionally, one of the species can usually be neglected when
114 |pH − pK_a| > 2, since their fractions become <1% (Fig. 1). Due to these direct relations between
115 pH and pK_a, the sole use of pK_a instead of pK_b is usually more meaningful from the practical
116 point of view. In the following, therefore, this system was adopted and acidic pK_a values are
117 indicated as p$K_{a,acid}$ and basic pK_a values are indicated as p$K_{a,base}$.
118 For compounds with more than one pK_a value, the estimation of the species distribution is rather
119 complex. This applies especially for poly-functional compounds with more than one acidic
120 and/or basic functional group, such as ampholytes or poly-protic acids and bases. As a
121 consequence, several acid-base equilibria (Fig. 3, Fig. 4, and Fig. 6) have to be considered
122 simultaneously (for details see Comer, 2007). Due to the environmental relevance of macro pK_a
123 values and the very challenging experimental determination of micro pK_a values, solely macro
124 pK_a values are considered in the following. In order to maintain the clarity of the article,
125 therefore, all pK_a given are macro pK_a values and only major microspecies are shown (e.g., in
126 Table 4).
127

3. Method for the identification of relevant molecule features

Since acidic or basic functionalities of functional groups are often not trivial, the identification and assignment of dissociation constants (pK_a values) to respective structural elements can be a complex and challenging task. Therefore, the comprehensive procedure presented in Figure 2 is recommended in order to assess potentially ionizable structures and factors influencing the dissociation equilibrium. As a result, the proposed method simplifies the determination of charge states and occurring molecule species for the pH value of interest. The application can be divided into five major procedural steps:

1) The structural formula for the compound of interest (neutral form preferred, no ions or salts) has to be found and should be cross-checked with additional references.

2) The structure is scrutinized for the presence of N, S and O atoms. If the molecule does not bear any of these atoms, it is non-ionizable (neutral) in the aquatic environment. If the molecule bears exclusively S and/or O atoms it is potentially acidic and if it bears only N atoms it is potentially basic. In case of potential ampholytes, groups containing S and/or O atoms co-exist with groups containing only N atoms and, thus, have to be evaluated separately.

3) In order to collect all environmentally relevant pK_a values (p$K_{a,acid} < 12$; p$K_{a,base} > 2$) and to verify their consistency, several literature/database resources or physicochemical property predictors have to be browsed. Special attention is required when only “most acidic” and “most basic” pK_a or pK_b values are provided, because additional, relevant pK_a values might exist and/or pK_b values have to be converted in p$K_{a,base}$ values *a priori*. Furthermore, it should be considered that pK_a values given for the same moiety may differ between two references. Conversely, pK_a values of different ionizable moieties might be very similar. When no pK_a value can be found or calculated for a certain molecule moiety, it is most probably non-ionizable. Hence, the simple presence of N, S or O atoms does not automatically imply acidic or basic functionalities (e.g., ethers, nitriles, nitro compounds).

4) The pK_a values have to be assigned to the acidic or basic functional groups listed in Table 1 and 2. Possible pK_a shifts due to inductive or mesomeric effects of the substituents (Fig. 5) need further consideration. Especially for complex molecule structures with more than one pK_a value (e.g., certain N-heterocycles, zwitters, and ampholytes), the additional use of computer programs is recommended. Software packages, such as ACD/pK_a, ChemAxon/pK_a, MoKa, and CompuDrug/pKalc can be very helpful for predicting pK_a values and assigning them correctly to the identified ionizable centers of the respective molecule. One should be aware, however, that the predictive power of computational methods is not always sufficient. This applies especially for new and/or complex molecules (Comer, 2007; Prankerd, 2007). Therefore, the consistency of all predictions has to be critically scrutinized.

5) When all relevant groups and pK_a values have been identified and assigned correctly, the charge state and major species of the considered molecule can be assessed for the pH value of interest. Alternatively, the species distribution can be calculated by means of Eq. (3) and Eq. (4) or for more complex cases (e.g., poly-protic compounds, ampholytes) by using software packages, such as Hyperquad Simulation and Speciation HySS (http://www.hyperquad.co.uk/hyss.htm; Alderighi et al., 1999).

3.1 Identification of ionizable acidic functional groups

Acidic compounds bearing solely functional groups, which are able to donate protons from their neutral form are referred to as organic acids. Due to dissociation (deprotonation), negatively charged organic anions can be formed. Organic acids that are able to donate more than one proton from one (e.g., phosphonic acids) or several functional groups (e.g., di- and tri-carboxylic acids) have several p$K_{a,acid}$ values and are called poly-protic acids (Fig. 3).

Acidic compounds with p$K_{a,acid}$ < 4 are mainly negatively charged (anions) and with p$K_{a,acid}$ > 10 neutral in the environmentally relevant pH range (pH = 4–10). Hence, the consideration of

178 negative p$K_{a,acid}$ values is also very important since permanent anions are formed in water (e.g.,
179 sulfonic acids with strongly electronegative substituents).
180 Table 1 lists the most common acidic reacting substructures in the relevant pK_a range of water
181 ($0 < pK_{a,acid} < 14$). As a rule of thumb, S- or O-atoms are always part of acidic functional groups.
182 Molecules not containing these atoms generally do not have acidic functionalities. To the
183 authors' knowledge, the only exception from this rule are certain N-heterocyclic ring systems,
184 such as several azoles (Table 3, No. 1-3), where the NH-group can dissociate under
185 environmentally relevant pH conditions.
186

187 ## *3.2 Identification of ionizable basic functional groups*

188 Alkaline compounds containing solely functional groups, which are able to accept protons to
189 their neutral form, are referred to as organic bases. Due to association (protonation), positively
190 charged organic cations can be formed. Some organic bases are poly-protic and are able to
191 accept more than one proton (several p$K_{a,base}$ values). In contrast to acids, poly-protic bases bear
192 always more than one ionizable functional group. That means only one proton can be attached
193 per N atom. Basic compounds with $pK_{a,base} > 10$ occur in the aquatic environment (pH = 4–10)
194 mainly in their cationic and with $pK_{a,base} < 4$ mainly in their neutral form.
195 Table 2 lists the most common basic reacting substructures in the relevant pK_a range of water
196 ($0 < pK_{a,base} < 14$). In general, N-atoms are a substantial part of basic functional groups. Thus,
197 molecules that do not contain N-atoms are not able to form basic functionalities. All ionizable
198 molecules containing exclusively N-atoms (no S and no O atoms) react basic. Again, several N-
199 heterocycles (e.g., Table 3, No. 1-3) are exceptional, since their NH-group can donate a proton
200 under environmental pH conditions.
201

3.3 Identification of ampholytes including zwitters

Compounds bearing both, acidic and basic reacting functional groups (p$K_{a,acid}$ and p$K_{a,base}$ values), are referred to as organic ampholytes or amphoteres. Depending on the pK_a values of the functional groups, two types of ampholytes can be distinguished (Table 3).

When p$K_{a,base}$ > p$K_{a,acid}$, zwitter-ionic species are formed in the pH range between p$K_{a,acid}$ – 2 and p$K_{a,base}$ + 2 (Fig. 4A). That means the molecule bears two oppositely charged moieties, but is formally neutral. The pH where the net charge becomes zero (sum of negative charges and sum of positive charges are equal) is referred to as isoelectric point and the mean of p$K_{a,acid}$ and p$K_{a,base}$. Amino acids are a very common example for zwitter-ions (Table 3, No. 6-8; Fig. 4A).

Vice versa, compounds with p$K_{a,base}$< p$K_{a,acid}$ form neutral molecule species in the transition zone between both pK_a values (Fig. 4B). 4-Aminophenol and 4-Aminobenzoic acid are provided as examples in Table 3 (No. 4 and 5). Table 3 lists selected compounds that include the combination of acidic and basic substructures in the relevant pK_a range of water (0 < pK_a < 14).

3.4 Factors leading to pK_a shifts

Depending on the chemical environment in the direct neighborhood to an acidic or basic group on an aliphatic or non-aromatic heterocyclic structure or on the same aromatic ring, pK_a values given in the Tables 1 to 3 can deviate considerably due to inductive (I) or mesomeric (M) effects of certain substituents (Fig. 5). Substituents with –I and –M effects increase the acid strength (lower p$K_{a,acid}$ values) by stabilizing the negative charge of the formed anion. *Vice versa*, bases become weaker and p$K_{a,base}$ values also decrease.

For example, due to the –I inductive effect of the nitro group and hydroxyl group, respectively, the p$K_{a,acid}$ value of nitrophenols (p$K_{a,acid}$ = 7−8) is significantly lower than for unsubstituted phenols (p$K_{a,acid}$ = 9−10) and the p$K_{a,base}$ of hydroxylamine (p$K_{a,base}$ = 5−6) is considerably lower than for other amines (p$K_{a,base}$ = 9−10).

227

228 4. Results and Discussion - Exemplary application of the 229 identification procedure

230 The recommended procedure in Figure 2 was exemplarily applied for four non-trivial, ionizable
231 pharmaceutical compounds (Table 4). A strongly pH-dependent subsurface transport can be
232 expected for some of these compounds due to the occurrence of different molecule species with
233 different transport properties (Schaffer et al., 2012).
234 The first example compound is the anticonvulsant phenobarbital (Table 4, No.1), which bears
235 three O and two N atoms in a barbituric ring structure. The keto-groups (O atoms) alone are not
236 able to donate protons under aqueous conditions and secondary amines are usually able to accept
237 protons (Table 2, No. 5). However, Table 1 (No. 8) shows that two ionizable, acidic OH groups
238 can be formed due to a lactam/lactim tautomerization between the NH- and keto group (5
239 possible tautomers in total). Consequently, p$K_{a,acid}$ values can be found. As an example,
240 SciFinder gives pK_a (most acidic) = 7.6, which is in accordance with other references (e.g.,
241 Avdeef, 2003) and implies that the molecule is partially anionic at pH > 5 (Fig. 3). However, an
242 additional p$K_{a,acid}$ exists, since two N atoms are in tautomerization equilibrium. This pK_a value is
243 around 12 (ChemAxon/pK_a, www.chemicalize.org; Prankerd, 2007). Thus, the double negatively
244 charged species is not relevant at pH typically encountered in the environment. The complete
245 species distribution for phenobarbital is shown in Figure 3. Similar pK_a values are also most
246 likely for all other barbituric acid derivatives (barbiturates).
247 The second considered compound is the beta-blocker atenolol (Table 4, No. 3). The structure
248 contains three O and two N atoms. In the literature one p$K_{a,base}$ = 9−10 and one not relevant
249 p$K_{a,acid}$ > 13 can be found. The p$K_{a,base}$ originates from the secondary amine (Table 2, No. 5) and
250 the p$K_{a,acid}$ from the very weakly acidic aliphatic hydroxy group (Table 1, No. 16). The

251 remaining N and two O atoms have no ionizable functionality, since they belong to an ether
252 bridge and a carboxamide group, respectively. Thus, atenolol is predominantly cationic at
253 pH < 9.5 (Fig. 1). The same applies for further, structurally very similar beta-blockers with a
254 secondary amine group (e.g. metoprolol, bisoprolol, esmolol, propanolol).
255 The third exemplary compound is the anticonvulsant carbamazepine (Table 4, No. 4). The
256 molecule bears two N atoms and one O atom. Table 1 and Table 2 give no evidence for the
257 existence of environmentally relevant pK_a values. This is confirmed by SciFinder and other
258 sources (e.g., ChemAxon/pK_a, www.chemicalize.org), which give two pK_a values (NH_2:
259 p$K_{a,acid}$ = 14; O: p$K_{a,base}$ = -0,5) for the carboxamide group (similar to atenolol). Thus, the
260 molecule is neutral over the entire environmentally relevant pH range (Fig. 6). Very similar pK_a
261 values can be also found for other environmentally relevant carboxamides, such as the herbicide
262 diuron.
263 The last explicitly considered compound is the antihistamine cetirizine (Table 4, No. 5) with two
264 N atoms and three O atoms. In the literature and chemical databases/predictors (e.g., van Balen,
265 2001; ChemAxon/pK_a, www.chemicalize.org), three pK_a values (pK_a = 2, 3 and 8) can be found.
266 Again, not all pK_a values can be found in SciFinder, since only the most acidic pK_a = 3.5 and the
267 most basic pK_a = 6.7 are given. The two N atoms belong to the cyclic piperazine substructure
268 with two pK_a (Table 2, No. 4) and two O atoms belong to the carboxy group (Table 1, No. 6).
269 The third O atom belongs to an ether group (see atenolol) without acidic/basic functionality.
270 However, this O atom causes a strong –I effect (Fig. 5) on the other functional groups.
271 Accordingly, the expected pK_a values from Table 1 and 2 are decreased significantly. Since one
272 p$K_{a,base}$ > p$K_{a,acid}$, ceterizine occurs mainly as a zwitter in the neutral pH range (Fig. 6).
273 The major species of six further ionizable and environmentally relevant chemical compounds are
274 given in Table 4 as additional examples. This includes three of the most commonly used
275 pesticides (glyphosate, ethephon, and mancozeb), two artificial sweeteners (aspartame and
276 cyclamate), and one stimulant (nicotine). Glyphosate is a herbicide with a phosphonic acid group

277 (Table 1, No. 3), a carboxyclic acid group (Table 1, No. 6), and a secondary amine group
278 (Table 2, No. 5) resulting in four pK_a values. The plant growth regulater ethephon is a
279 phosphonic acid (Table 1, No. 3) with two respective pK_a values. Mancozeb is a fungizide with
280 two dithiocarbamate groups (Table 1, No. 4) and a predominantly double negatively charged
281 anion at environmental pH conditions. The structure of the artificial sweetener aspartame
282 contains one relevant primary amine (Table 2, No. 5) as well as a carboxyclic acid group
283 (Table 1, No. 6). In contrast, cyclamate bears a sulfonic acid group (sulfamic acid group) with a
284 respective low pK_a value (Table 1, No. 1). The last exemplarily considered compound is the
285 stimulant nicotine, which is an organic cation at environmental pH due to the protonation of the
286 tertiary amine (Table 2, No. 5) in the pyrrolidine ring.
287
288
289 **Acknowledgments**
290 This research has received funding from the European Community's 7th Framework Programme
291 FP7/2007–2013, within the MUSTANG project (grant agreement No. 227286) and from the
292 German Research Foundation (DFG), within the GEOCAT project (project No. LI 1314/3-1).
293 Further, we acknowledge the German Federal Ministry of Education and Research (BMBF)
294 (project AGRO under promotional reference No. 02WRS1277A and the German Ministry for
295 Environment (BMU) (project REAKTHERM under grant No. 0325417) for the financial
296 support.
297
298

References

Alderighi, L., Gans, P., Ienco, A., Peters, D., Sabatini, A., Vacca, A., 1999. Hyperquad simulation and speciation (HySS): a utility program for the investigation of equilibria involving soluble and partially soluble species. Coord. Chem. Rev. 184 (1), 311-318.

Avdeef, A., 2003. Absorption and Drug Development: Solubility, Permeability and Charge State. John Wiley and Sons, Hoboken.

Banzhaf, S., Nödler, K., Licha, T., Krein, A., Scheytt, T., 2012. Redox-sensitivity and mobility of selected pharmaceutical compounds in a low flow column experiment. Sci. Total Environ. 438, 113-121.

Box, K., Bevan, C., Comer, J., Hill, A., Allen, R., Reynolds, D., 2003. High-throughput measurement of pK_a values in a mixed-buffer linear pH gradient system. Anal. Chem. 75 (4), 883-892.

Burgos, W.D., Pisutpaisal, N., 2006. Sorption of naphthoic acids and quinoline compounds to estuarine sediment. J. Contam. Hydrol. 84 (3-4), 107-126.

Caron, G., Steyaert, G., Pagliara, A., Reymond, F., Crivori, P., Gaillard, P., Carrupt, P.-A., Avdeef, A., Comer, J., Box, K.J., Girault, H.H., Testa, B., 1999. Structure-lipophilicity relationships of neutral and protonated β-blockers; Part I: Intra- and intermolecular effects in isotropic solvent systems. Helv. Chim. Acta 82, 1211-1222.

Chamberlain, K., Evans, A.A., Bromilow, R H., 1996. 1-octanol/water partition coefficient (K_{ow}) and pK_a for ionisable pesticides measured by a pH-Metric Method. Pestic. Sci. 47 (3), 265-271.

Comer, J.E.A., 2007. 5.16 - Ionization constants and ionization profiles, in: Taylor, J.B., Triggle, D.J. (Eds.), Comprehensive Medicinal Chemistry II, Elsevier, Oxford, 357-397.

Fatta-Kassinos, D., Meric, S., Nikolaou, A., 2011. Pharmaceutical residues in environmental waters and wastewater: current state of knowledge and future research. Anal. Bioanal. Chem. 399 (1), 251-275.

Franco, A., Ferranti, A., Davidsen, C., Trapp, S., 2010. An unexpected challenge: ionizable compounds in the REACH chemical space. Int. J. Life Cycle Assess. 15 (4), 321-325.

Guo, Z.X., Cai, Q., Yang, Z., 2007. Ion chromatography/inductively coupled plasma mass spectrometry for simultaneous determination of glyphosate, glufosinate, fosamine and ethephon at nanogram levels in water. Rapid Commun. Mass Spectrom. 21 (10), 1606-1612.

Hidber, P.C., Graule, T J., Gauckler, L.J., 1996. Citric acid-a dispersant for aqueous alumina suspensions. J. Am. Ceram. Soc., 79 (7), 1857-1867.

Højberg, A.L., Engesgaard, P., Bjerg, P.L., 2005. Pesticide transport in an aerobic aquifer with variable pH - Modeling of a field scale injection experiment. J. Contam. Hydrol. 78 (3), 231-255.

Jin, X., Peldszus, S., 2012. Selection of representative emerging micropollutants for drinking water treatment studies: A systematic approach. Sci. Total Environ. 414, 653-663.

Kah, M., Brown, C.D., 2008. Log *D*: Lipophilicity for ionisable compounds. Chemosphere 72 (10), 1401-1408.

Lorphensri, O., Sabatini, D.A., Kibbey, T.C.G., Osathaphan, K., Saiwan, C., 2007. Sorption and transport of acetaminophen, 17α-ethynyl estradiol, nalidixic acid with low organic content aquifer sand. Water Res. 41 (10), 2180–2188.

MacKay, A.A., Vasudevan, D., 2013. Polyfunctional ionogenic compound sorption: challenges and new approaches to advance predictive models. Environ. Sci. Technol. 46 (17), 9209-9223.

Manallack, D.T., 2007. The pK_a distribution of drugs: Application to drug discovery. Perspect. Medicin. Chem. 1, 25-38.

Manallack, D.T., 2009. The acid-base profile of a contemporary set of drugs: implications for drug discovery. SAR QSAR Environ. Res. 20 (7-8), 611-655.

Niedbala, A., Schaffer, M., Licha, T., Nödler, K., Börnick, H., Ruppert, H., Worch, E., 2013. Influence of competing inorganic cations on the ion exchange equilibrium of the monovalent organic cation metoprolol on natural sediment. Chemosphere 90 (6), 1945–1951.

Nödler, K., Hillebrand, O., Idzik, K., Strathmann, M., Schiperski, F., Zirlewagen, J., Licha, T. 2013. Occurrence and fate of the angiotensin II receptor antagonist transformation product valsartan acid in the water cycle - A comparative study with selected β-blockers and the persistent anthropogenic wastewater indicators carbamazepine and acesulfame. Water Research 47 (17), 6650-6659.

Prankerd, R.J., 2007. Critical compilation of pK_a values for pharmaceutical substances, in: Brittain, H.G. (Ed.), Profiles of Drug Substances, Excipients and Related Methodology 33. Academic Press, London, 1-33.

Schaffer, M., Boxberger, N., Börnick, H., Licha, T., Worch, E., 2012a. Sorption influenced transport of ionizable pharmaceuticals onto a natural sandy aquifer sediment at different pH. Chemosphere 87 (5), 513–520.

354 Schaffer, M., Börnick, H., Nödler, K., Licha, T., Worch, E., 2012b. Role of cation exchange processes on the
355 sorption influenced transport of cationic β-blockers in aquifer sediments. Water Res. 46 (17), 5472–5482.

356 Schaffer, M., Maier, F., Licha, T., Sauter, M., 2013. A new generation of tracers for the characterization of
357 interfacial areas during supercritical carbon dioxide injections into deep saline aquifers: Kinetic interface-
358 sensitive tracers (KIS tracer). International Journal of Greenhouse Gas Control 14, 200–208.

359 Scheytt, T.J., Mersmann, P., Heberer, T., 2006. Mobility of pharmaceuticals carbamazepine, diclofenac, ibuprofen,
360 and propyphenazone in miscible-displacement experiments. J. Contam. Hydrol. 83 (1-2), 53-69.

361 Skwierczynski, R.D., Connors, K.A., 1993. Demethylation kinetics of aspartame and L-phenylalanine methyl ester
362 in aqueous solution. Pharm. Res. 10 (8), 1174-1180.

363 Tülp, H.C., Fenner, K., Schwarzenbach, R.P., Goss, K.-U., 2009. pH-dependent sorption of acidic organic chemicals
364 to soil organic matter. Environ. Sci. Technol. 43 (24). 9189–9195.

365 van Balen, P.G., Caron, G., Ermondi, G., Pagliara, A., Grandi, T., Bouchard, G., Fruttero, R., Carrupt, P. A, Testa,
366 B., 2001. Lipophilicity behaviour of the zwitterionic antihistamine cetirizine in phosphatidylcholine
367 liposomes/water Systems. Pharmaceut. Res. 18 (5), 694-701.

368 Vasudevan, D., Bruland, G.L., Torrance, B.S., Upchurch, V.G., MacKay, A.A., 2009. PH-dependent ciprofloxacin
369 sorption to soils: interaction mechanisms and soil factors influencing sorption. Geoderma 151, 68–76.

370 Williams, R.T., 1971. The metabolism of certain drugs and food chemicals in man. Ann. N.Y. Acad. Sci. 179 (1),
371 141-154.

372 Yasuda, M., Ota, T., Morikawa, A., Mawatari, K.I., Fukuuchi, T., Yamaoka, N., Kaneko, K, Nakagomi, K., 2013.
373 Simultaneous determination of nicotine and cotinine in serum using high-performance liquid
374 chromatography with fluorometric detection and postcolumn UV-photoirradiation system. J. Chromatogr.
375 B 934, 41-45.

Figure Captions

Fig. 1: Visualization of the species distribution for selected mono-protic compounds: (**A**) mono-protic carboxylic acid with $pK_{a,acid} = 4.5$ (e.g., acetic or benzoic acid); and (**B**) mono-protic base with $pK_{a,base} = 9.5$ (atenolol; Table 4, No. 3) at 25 °C ($pK_w = 14$) using the software HySS 2009, V. 4.0.31 (http://www.hyperquad.co.uk/hyss.htm; Alderighi et al., 1999).

Fig. 2: Recommended procedure for the identification of ionizable molecules structures and the correct assignment of pK_a values.

Fig. 3: Visualization of the species distribution for selected poly-protic acids (Table 4, No. 1 and No. 2): (**A**) di-protic acid phenobarbital with $pK_{a,acid} = 7.6$ and 12.0; and (**B**) tri-protic citric acid with $pK_{a,acid} = 3.1$, 4.8, and 6.4 at 25 °C ($pK_w = 14$) using the software HySS 2009, V. 4.0.31 (http://www.hyperquad.co.uk/hyss.htm; Alderighi et al., 1999).

Fig. 4: Visualization of the species distribution for selected ampholytes (Table 3, No. 8 and 4): (**A**) γ-amino acid with $pK_{a,acid} = 4.5$ and $pK_{a,base} = 11$; and (**B**) aminophenol with $pK_{a,acid} = 10$ and $pK_{a,base} = 5$ at 25 °C ($pK_w = 14$) using the software HySS 2009, V. 4.0.31 (http://www.hyperquad.co.uk/hyss.htm; Alderighi et al., 1999).

Fig. 5: Inductive (I) and mesomeric (M) effects caused by different substituents in the vicinity of the functional groups.

Fig. 6: Visualization of the species distribution for selected pharmaceuticals (Table 4, No. 4 and 5): (**A**) carbamazepine; and (**B**) cetirizine at 25 °C ($pK_w = 14$) using the software HySS 2009, V. 4.0.31 (http://www.hyperquad.co.uk/hyss.htm; Alderighi et al., 1999).

Figure 1

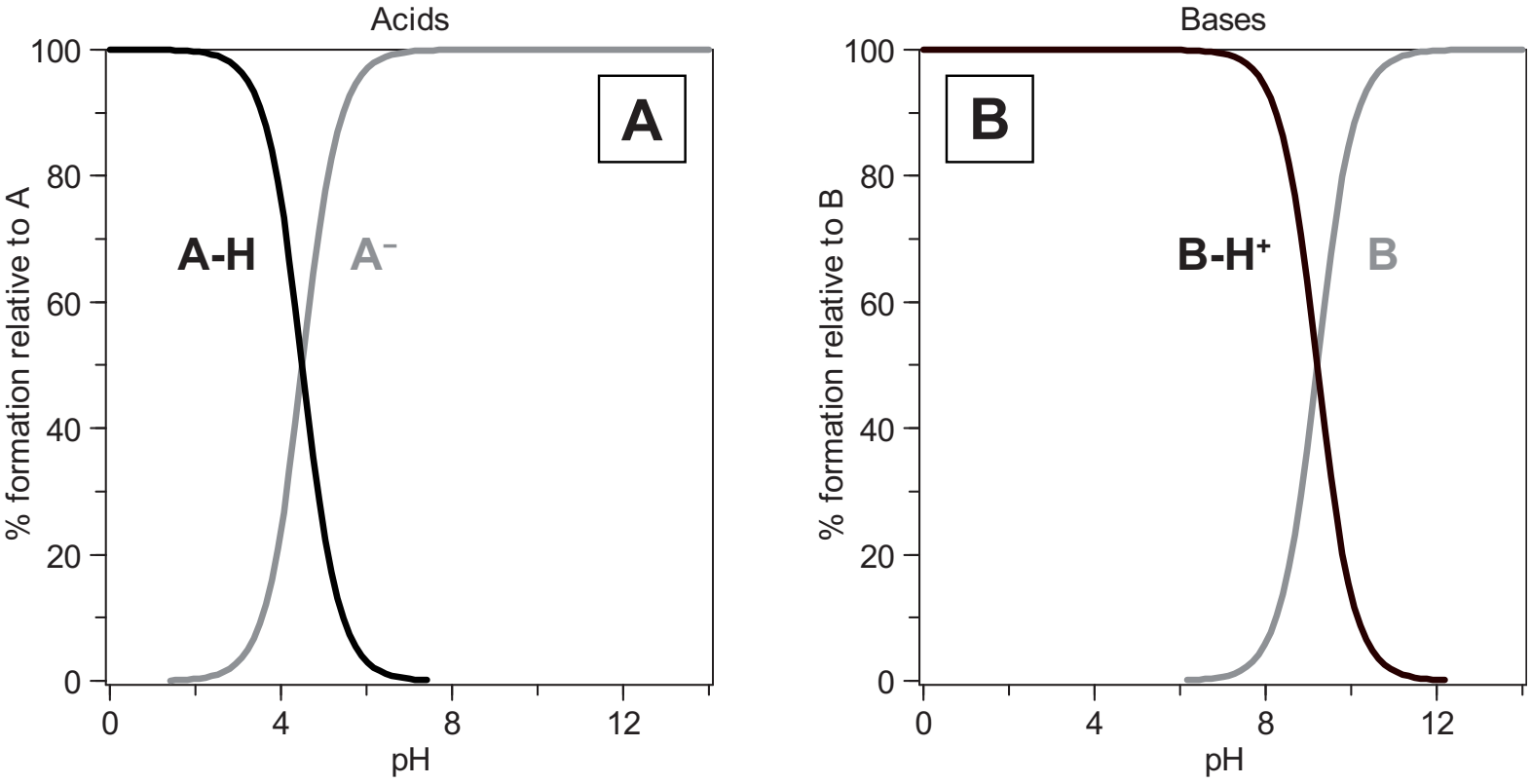
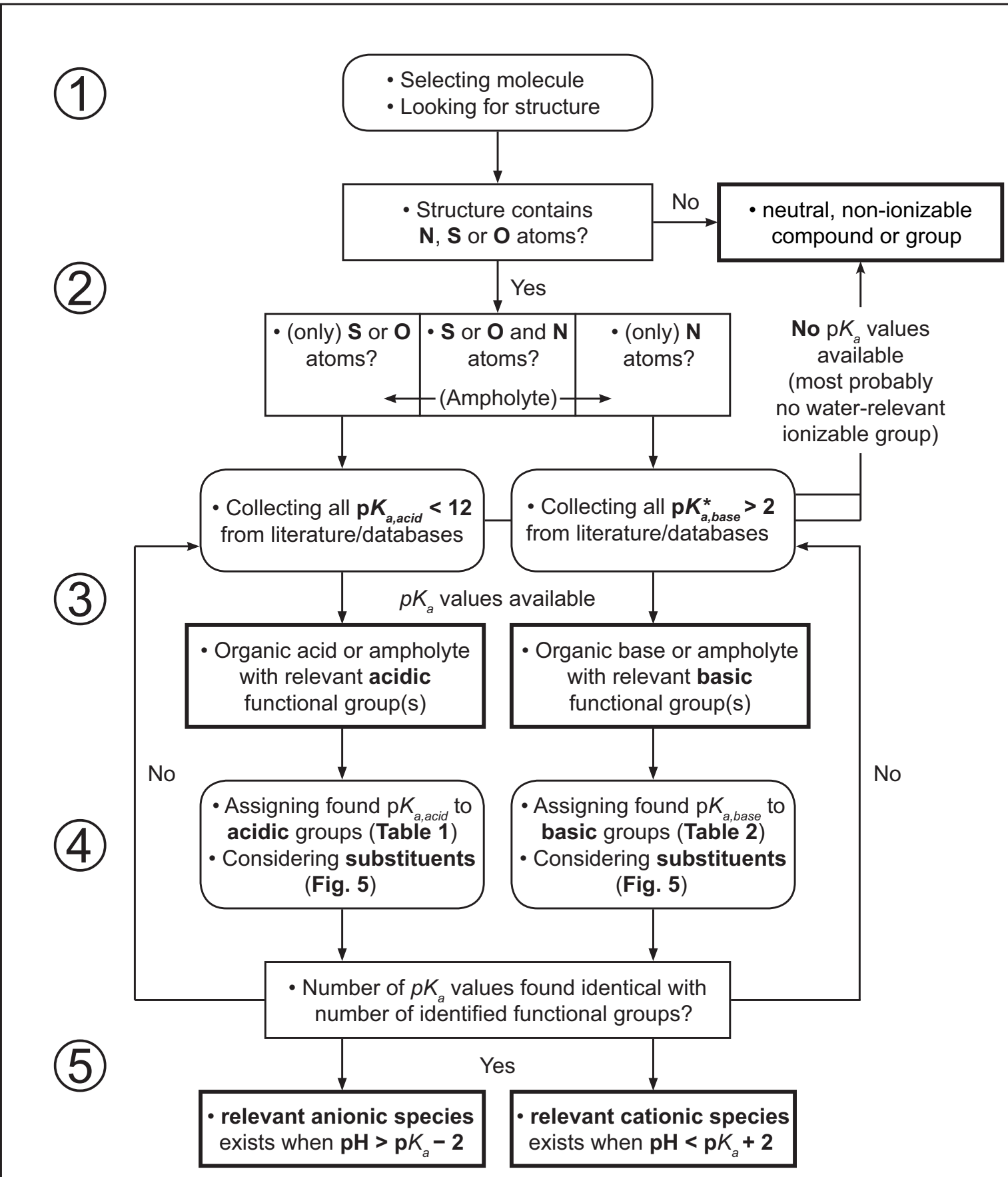

Acids
A
A-H
A⁻
% formation relative to A
pH
Bases
B
B-H⁺
B
% formation relative to B
pH

Figure 2



* Exception: Some N-containing heterocycles may also react acidic (p$K_{a,acid}$), e.g. azoles in Table 3.

Figure 3

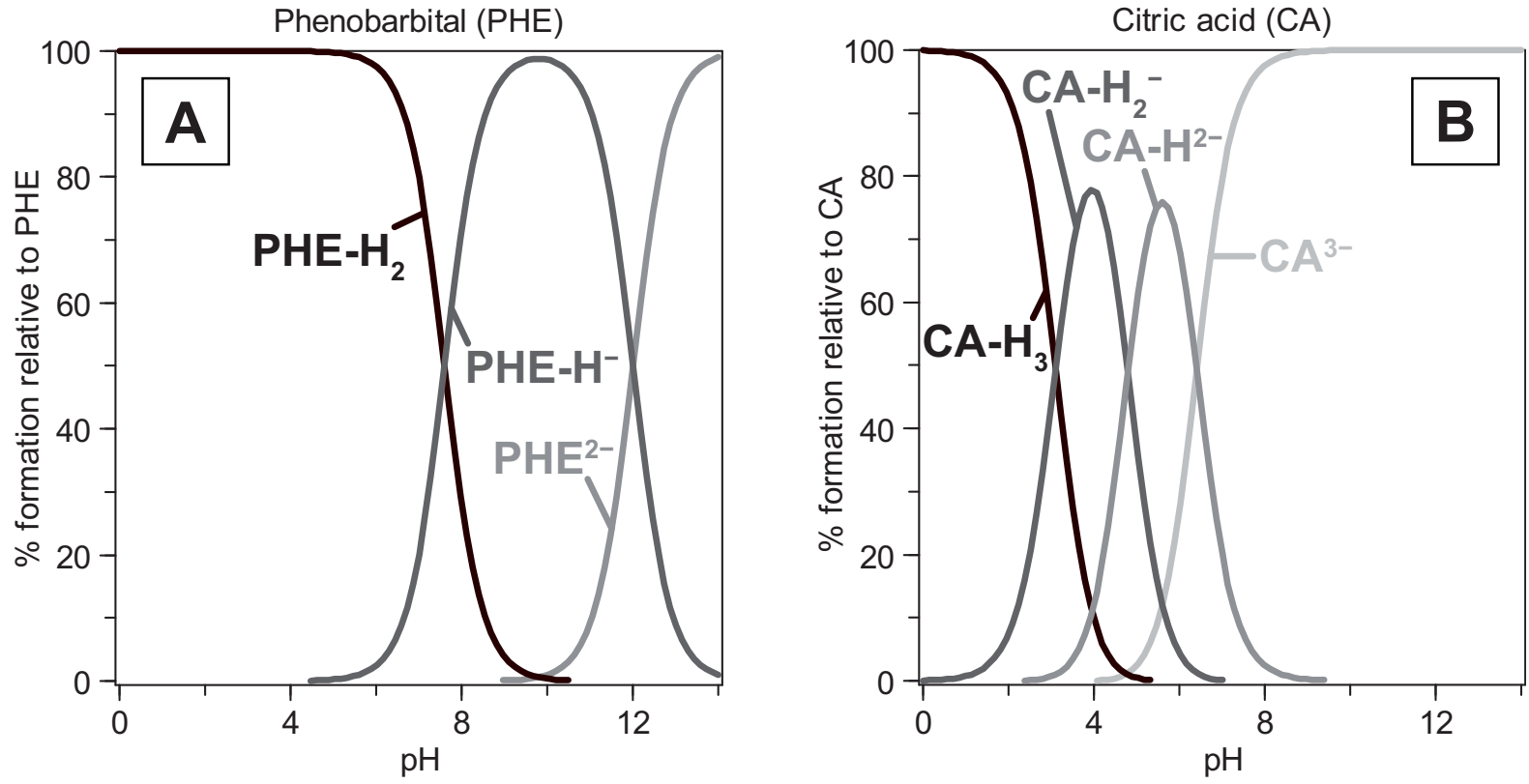

Phenobarbital (PHE)
A
PHE-H2
PHE-H−
PHE2−
% formation relative to PHE
pH
Citric acid (CA)
B
CA-H2−
CA-H2−
CA3−
CA-H3
% formation relative to CA
pH

Figure 4

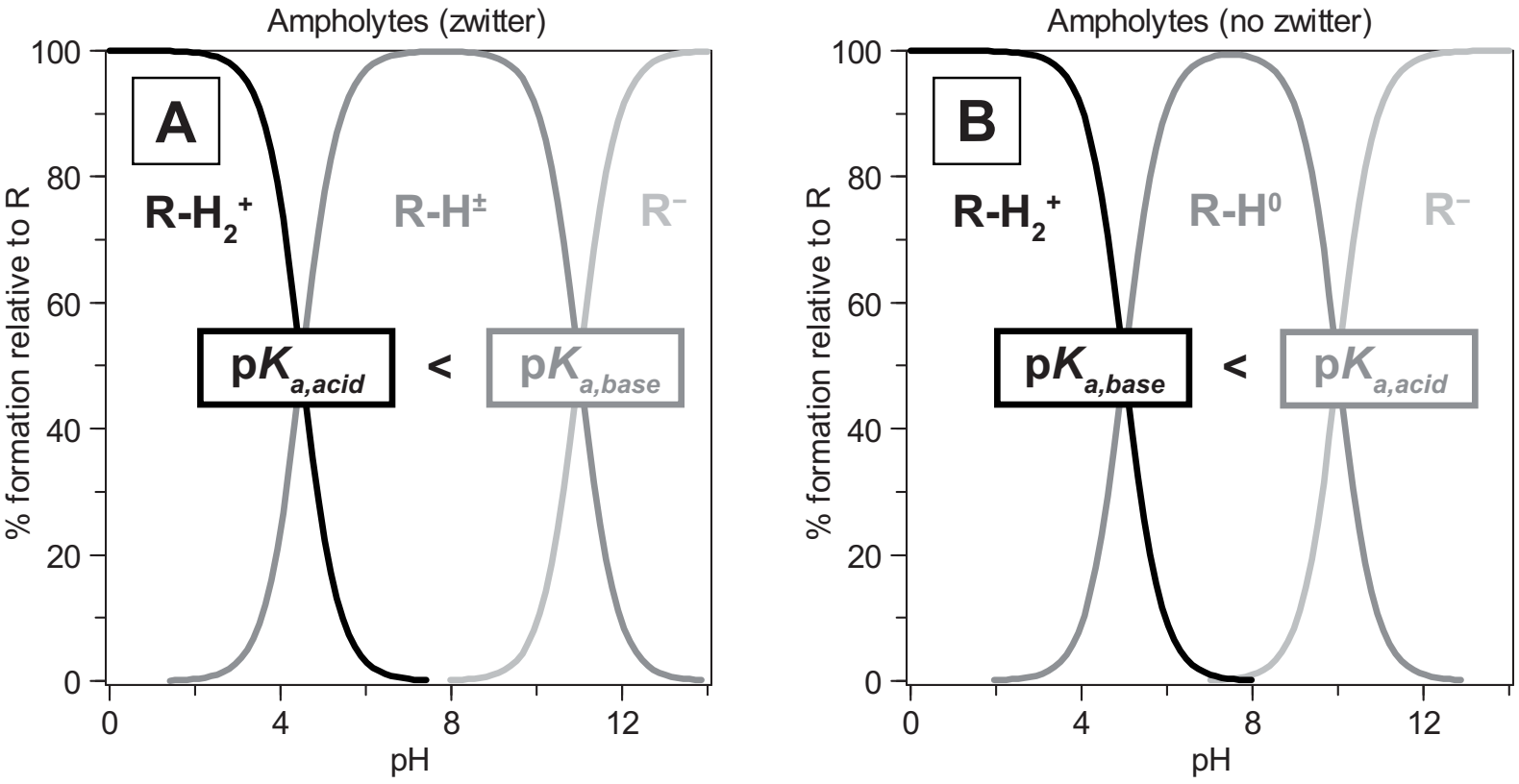

Ampholytes (zwitter)
A
R-H2+
R-H±
R−
pKa,acid
<
pKa,base
% formation relative to R
pH
Ampholytes (no zwitter)
B
R-H2+
R-H0
R−
pKa,base
<
pKa,acid
% formation relative to R
pH

Figure 5

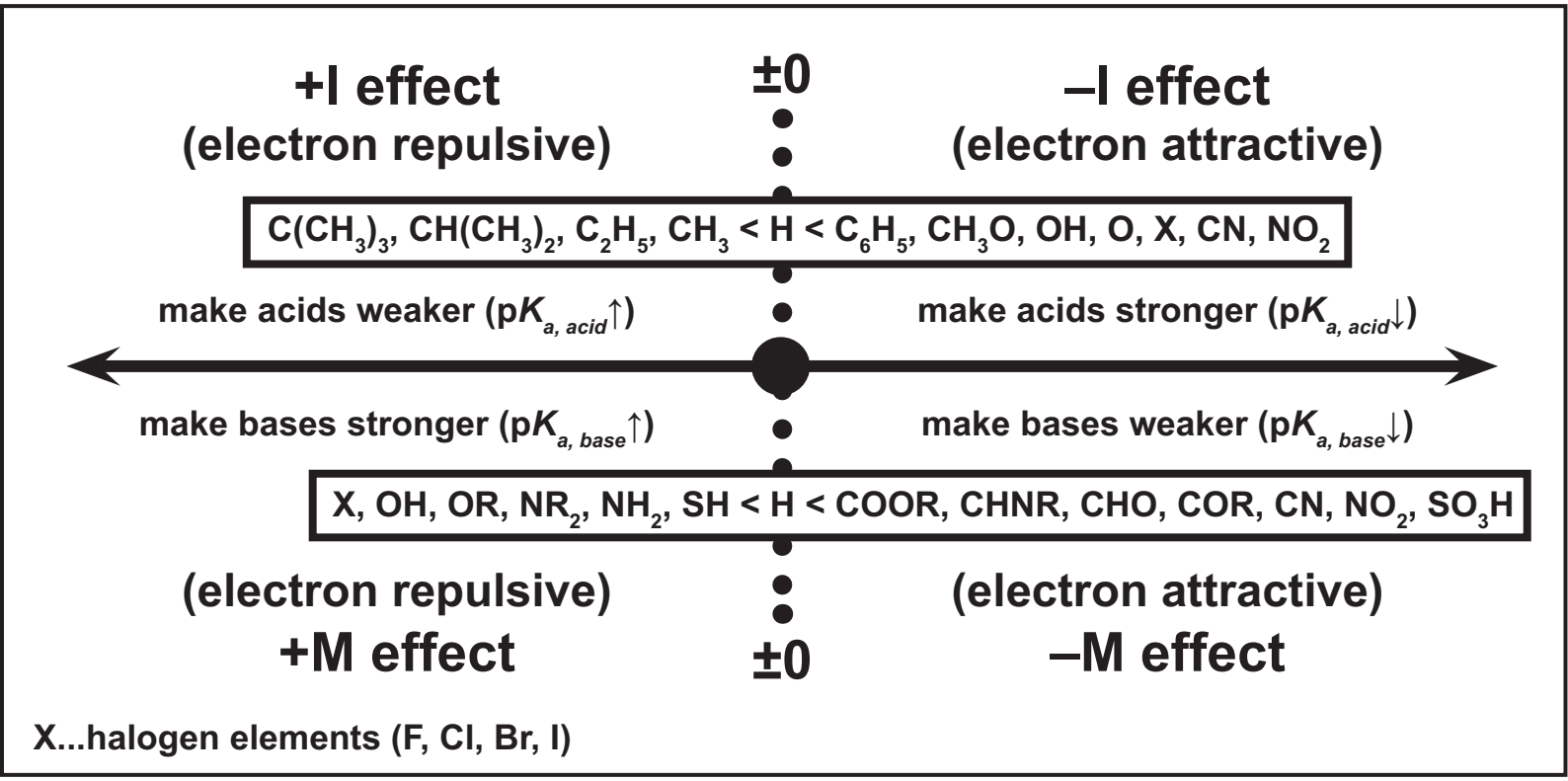

+I effect
(electron repulsive)
±0
–I effect
(electron attractive)
C(CH3)3, CH(CH3)2, C2H5, CH3 < H < C6H5, CH3O, OH, O, X, CN, NO2
make acids weaker (pKa, acid↑)
make acids stronger (pKa, acid↓)
make bases stronger (pKa, base↑)
make bases weaker (pKa, base↓)
X, OH, OR, NR2, NH2, SH < H < COOR, CHNR, CHO, COR, CN, NO2, SO3H
(electron repulsive)
+M effect
±0
(electron attractive)
–M effect
X...halogen elements (F, Cl, Br, I)

Figure 6

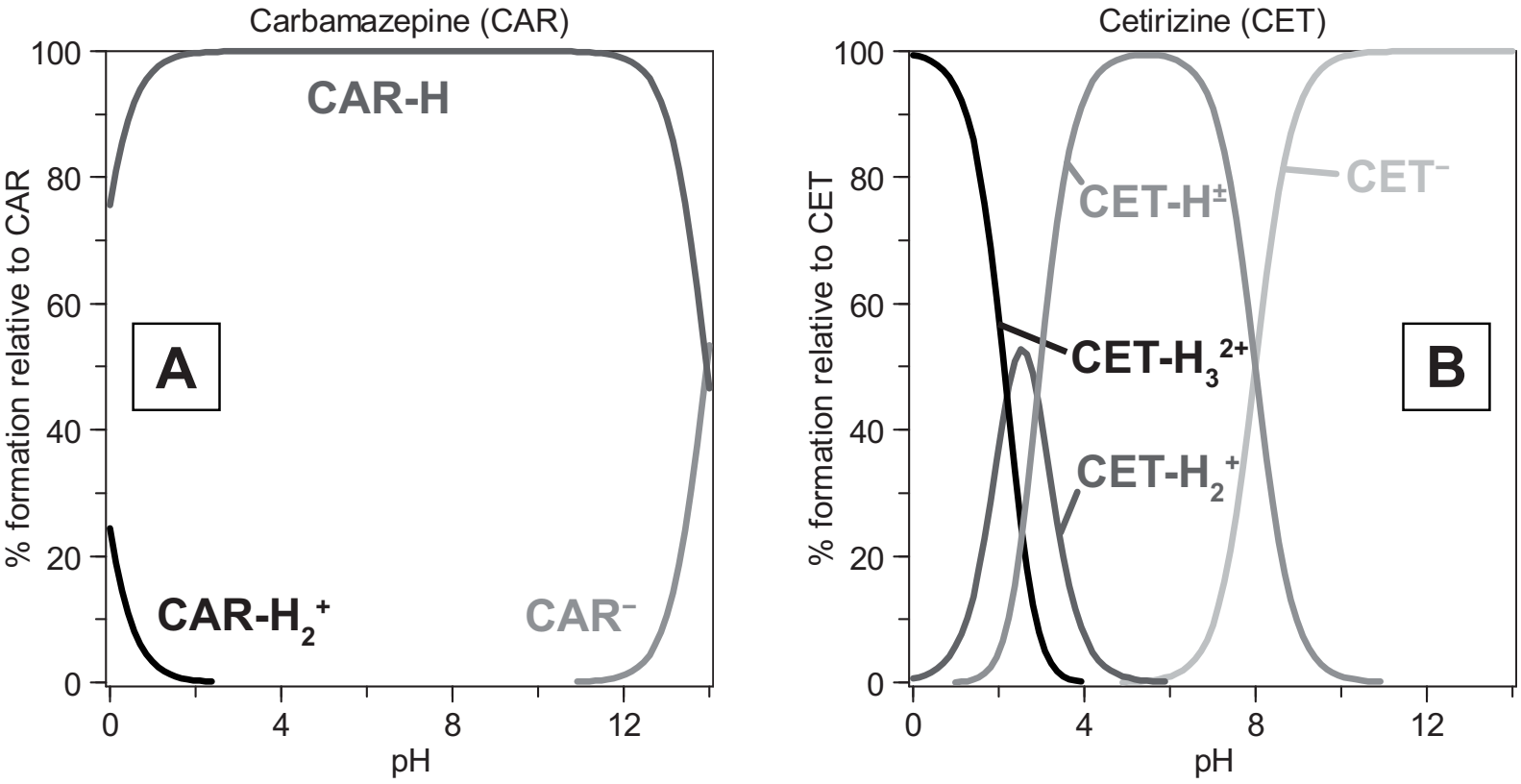
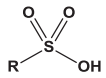
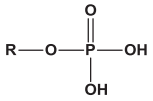
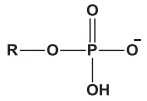
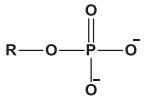
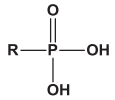
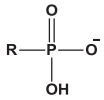
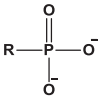
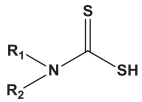
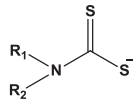
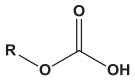
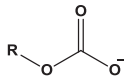
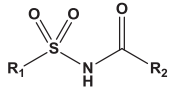
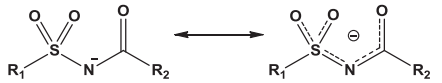
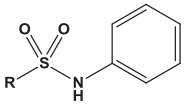
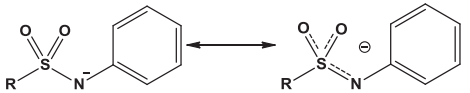
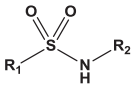
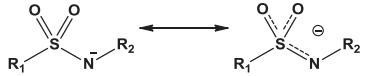
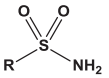
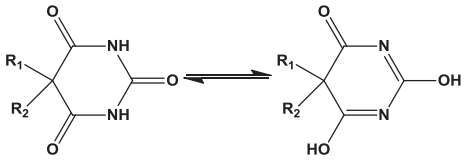
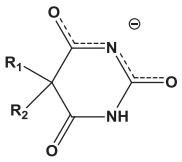
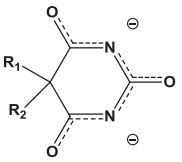
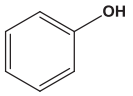
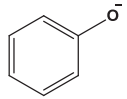
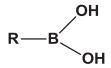
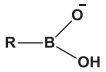
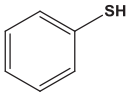
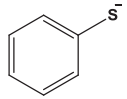
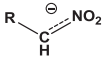
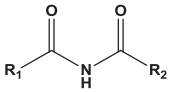
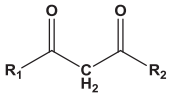
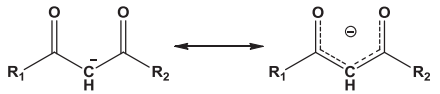
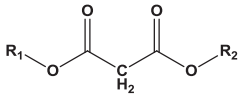
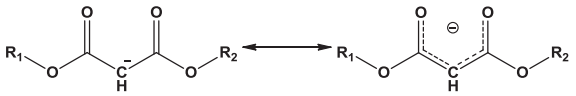

Carbamazepine (CAR)
100
80
60
40
20
0
% formation relative to CAR
CAR-H
A
CAR-H2+
CAR−
0
4
8
12
pH
Cetirizine (CET)
100
80
60
40
20
0
% formation relative to CET
CET-H±
CET−
CET-H32+
B
CET-H2+
0
4
8
12
pH

Table 1

Table 1
Organic acids.

No.	Compound class	Acid 1	* p$K_{a,acid1}$	Base 1 / Acid 2	* p$K_{a,acid2}$	Base 2
1	**Sulfonic acids**		<1			
2	**Phosphoric acid esters**		1-2		6-7	
3	**Phosphonic acids**		2-3		7-8	
4	**Carbamodithioic acids**		2-4			
5	**Carboxylic acid esters**		3-4			
6	**Carboxylic acids**		4-5			
7a)	† **Sulfonamides**		5			
b)			6-8			
c)			10			
d)			10-12			
8	†‡ **Barbiturates**	 Lactam/lactim-tautomerism	7-8		11-12	

No.	Compound class	Acid 1	* p$K_{a,acid1}$	Base 1 / Acid 2	* p$K_{a,acid2}$	Base 2
9	**Phenols**		10			
10	**Boronic acids**		8-10			
11a)	**Thiols**		7			
b)			10-11			
12	† **Nitroalkanes**	 Nitro/aci-nitro-tautomerism	7-9			
13	† **Imides**		10			
14	† **1,3-Diketones**		9-11			
15	† **Malonic acid esters**		10-12			
16	**Alcohols**		> 13			

* Given values represent the typical pK_a range for selected compound classes. Depending on the specific compound structure these values may differ by several log units.

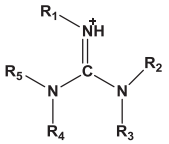
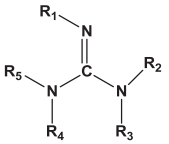
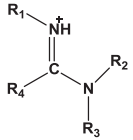
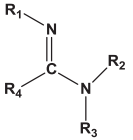
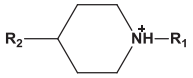
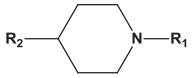
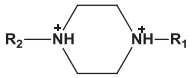
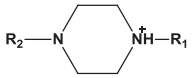
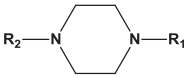
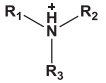
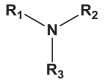
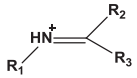
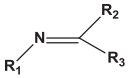
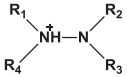
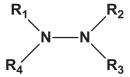
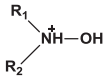
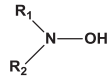
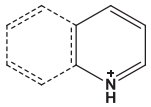
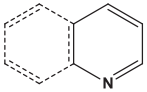
† Dashed lines represent the resonance structure of the anion due to charge delocalization (mesomerism). By analogy with barbiturates (No. 8) and the fact that N atoms have usually three and C atoms four covalent bonds, respectively, the deprotonation takes place most likely over the O atoms.

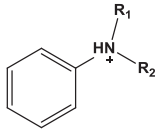
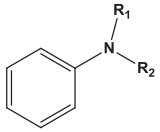
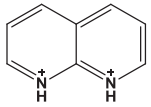
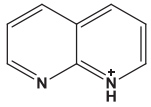
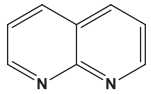
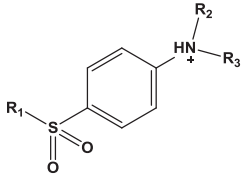
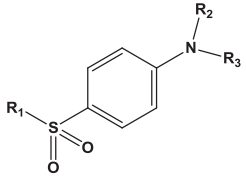
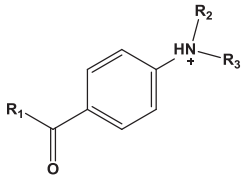
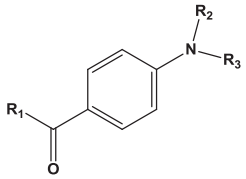
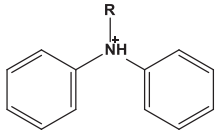
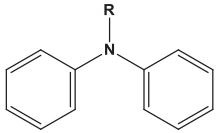
‡ Not all possible (five) tautomers of the barbituric ring shown.

⇌ Acid/base or tautomerization equilibrium.

⟷ Mesomersim.

Table 2
Organic bases.

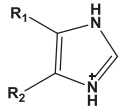
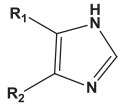
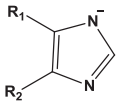
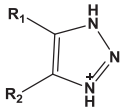
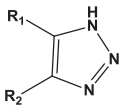
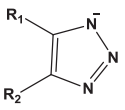
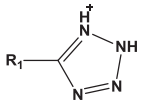
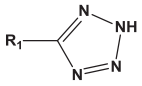
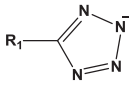
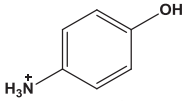
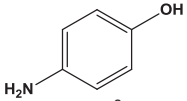
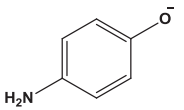
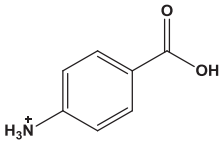
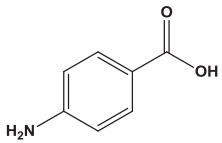
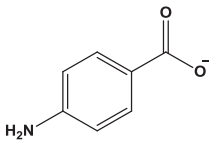
No.	Compound class	Acid 1	* p$K_{a,base1}$	Base 1 / Acid 2	* p$K_{a,base2}$	Base 2
1	**Guanidines**		Aliphatic: 13-14 Aromatic: 11-12			
2	**Amidines**		Aliphatic: 11-12 Aromatic: 9-10			
3	**Piperidines**		9-10			
4	**Piperazines**		5		10	
5	**Amines (primary, secondary, tertiary)**		9-10			
6	**Imines**		7-10			
7	**Hydrazines**		6-7			
8	**Hydroxylamines**		5-6			
9	**Pyridines and chinolines**		5			

No.	Compound class	Acid 1	* p$K_{a,base1}$	Base 1 / Acid 2	* p$K_{a,base2}$	Base 2
10	Aromatic amines		4-5			
11	Naphthyridines		1-2		2-4	
12	Phenylogic sulfonamides		2			
13	Phenylogic amides		1-2			
14	Diphenylamines		<1			

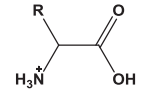
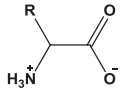
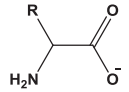
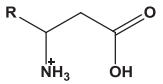
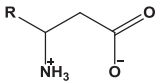
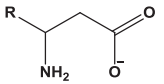
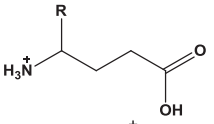
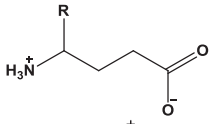
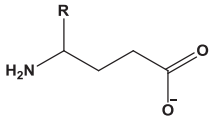
* Given values represent the typical pK_a range for selected compound classes. Depending on the specific compound structure these values may differ by several log units.
⇌ Acid/base equilibrium.

Table 3
Selected ampholytes.

A) Selected amphoteres ($pK_{a,base} < pK_{a,acid}$)

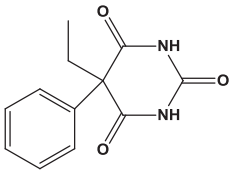
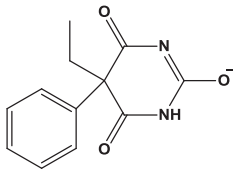
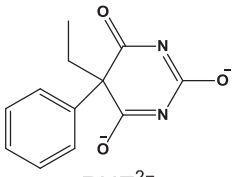
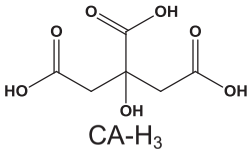
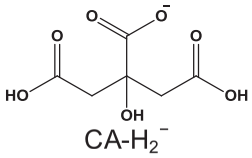
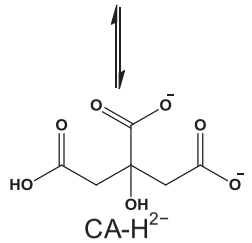
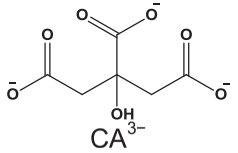
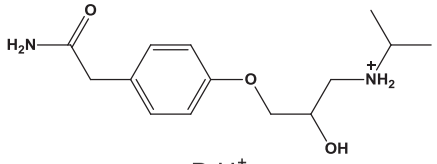
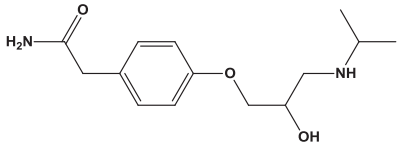
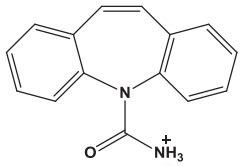
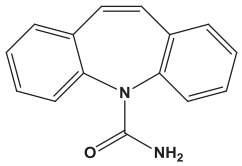
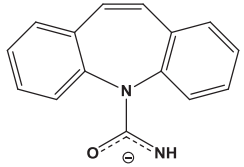
No.	Compound	Acid 1	* $pK_{a,base}$	Base 1 / Acid 2	* $pK_{a,acid}$	Base 2
1	‡ Imidazoles		8		15	
2	‡ Triazoles		1-2		9-11	
3	‡ Tetrazoles		0-1		4-5	
4	Aminophenol (Fig. 4B)	 R-H_2^+	5	 R-H^0	10	 R^-
5	Aminobenzoic acid (major species)		2.5		5	

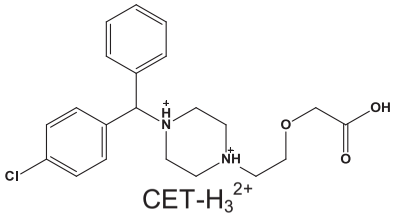
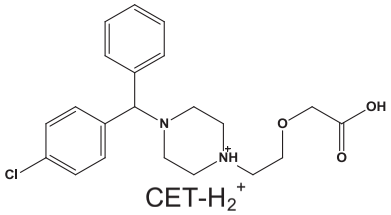
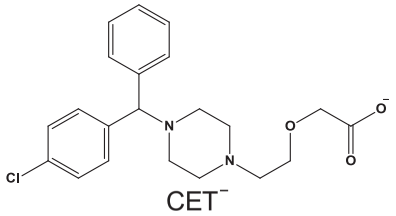
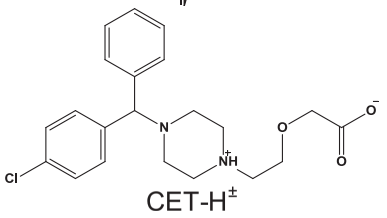
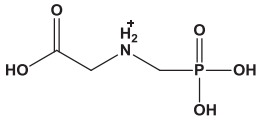
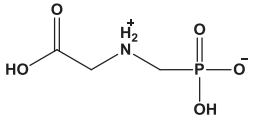
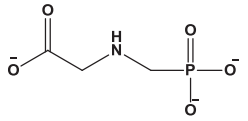
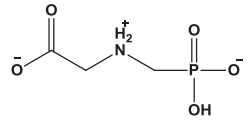
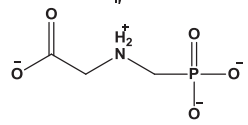
B) Selected zwitters ($pK_{a,acid} < pK_{a,base}$)

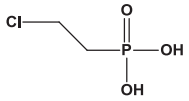
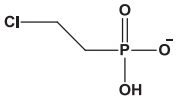
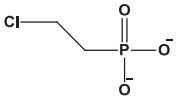
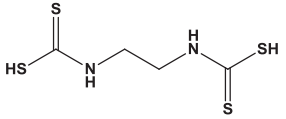
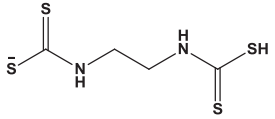
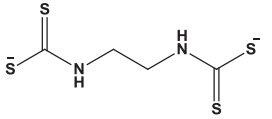
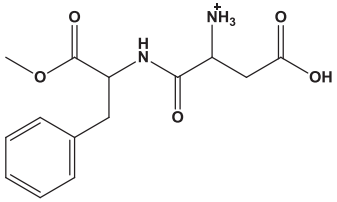
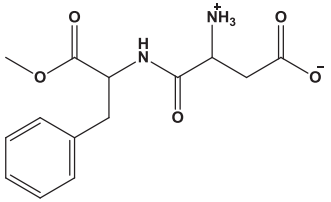
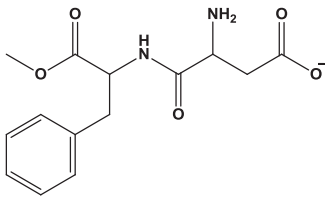
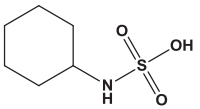
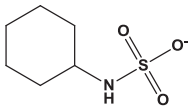
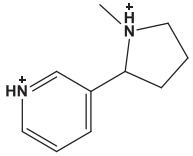
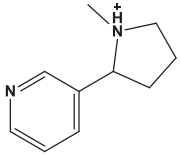
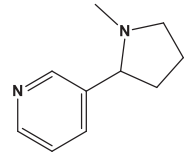
No.	Compound	Acid 1	* $pK_{a,acid}$	Base 1 / Acid 2	* $pK_{a,base}$	Base 2
6	α-Amino acids		2.5		9-10	
7	β-Amino acids		3.5		10-10.5	
8	γ-Amino acids (Fig. 4A)	 $R\text{-}H_2^+$	4-5	 $R\text{-}H^{\pm}$	11-11.5	 R^-

* Given values represent the typical pK_a range for selected compound classes. Depending on the specific compound structure these values may differ by several log units.
‡ Not all possible tautomers of the heterocyclic ring system shown.
⇌ Acid/base equilibrium.

Table 4
Example compounds.

No.	Compound		pK_a		pK_a	
1	‡ **Phenobarbital (acidic, Fig. 3A)**	 PHE-H_2	[a] $pK_{a,acid1}$ 7.5 ⇌	 PHE-H^-	[a] $pK_{a,acid2}$ 12.0 ⇌	 PHE^{2-}
2	**Citric acid (acidic, Fig. 3B)**	 $CA-H_3$	[b] $pK_{a,acid1}$ 3.1 ⇌  $CA-H_2^-$  $CA-H^{2-}$ (major species) [b] $pK_{a,acid2}$ 4.8 ⇌		[b] $pK_{a,acid3}$ 6.4 ⇌	 CA^{3-}
3	**Atenolol (basic, Fig. 1B)**	 $B-H^+$	[c] $pK_{a,base}$ 9.5 ⇌	 B		
4	† **Carbamazepine (neutral, Fig. 6A)**	 $CAR-H_2^+$	[d] $pK_{a,base}$ −0.49 ⇌	 CAR-H	[d] $pK_{a,acid}$ 13.9 ⇌	 CAR^-

No.	Compound		pK_a		pK_a	
5	**Cetirizine (zwitter, Fig. 6B)**	 CET-H_3^{2+}	[e] p$K_{a,base1}$ R-NH^+: 2.2	 CET-H_2^+	[e] p$K_{a,base2}$ R-NH^+: 8.0	 CET^-
			[e] p$K_{a,acid}$ R-COOH: 2.9			
			(major species)	 CET-$H^{\pm}$		
6	**Glyphosate (zwitter)**		[f] p$K_{a,acid1}$ R-H_2PO_3: 0.8		[f] p$K_{a,base}$ R-NH_2^+: 11.0	
			[f] p$K_{a,acid2}$ R-COOH: 2.3			
			[f] p$K_{a,acid3}$ R-HPO_3^-: 6.0			
			(major species)			

No.	Compound		pK_a		pK_a	
7	**Ethephon (acidic)**		[g] p$K_{a,acid1}$ $R\text{-}H_2PO_3$: 2.5 ⇌		[g] p$K_{a,acid2}$ $R\text{-}HPO_3^-$: 7.2 ⇌	
8	**Mancozeb (acidic)**		[h] p$K_{a,acid1}$ 1.7 ⇌		[h] p$K_{a,acid2}$ 2.3 ⇌	
9	**Aspartame (zwitter)**		[i] p$K_{a,acid1}$ 3.2 ⇌		[i] p$K_{a,base}$ 7.9 ⇌	
10	**Cyclamate (acidic)**		[j] p$K_{a,acid1}$ 1.9 ⇌			
11	**Nicotine (basic)**		[k] p$K_{a,base1}$ 3.1 ⇌		[k] p$K_{a,base2}$ 8.0 ⇌	

[‡] Not all possible (five) tautomers shown (compare Table 1, No. 8).

[†] Dashed lines represent the resonance structure of the anion due to charge delocalization (mesomerism). By analogy with barbiturates (Table 1, No. 8) and the fact that N atoms have usually three and C atoms four covalent bonds, respectively, the deprotonation takes place most likely over the O atoms.

[a] Prankerd (2007).
[b] Hidber et al. (1996).
[c] Caron et al. (1999).
[d] SciFinder Scholar 2007 predicted values, calculated using Advanced Chemistry Development (ACD/Labs) Software V11.02 (1994–2012 ACD/Labs).
[e] Van Balen et al. (2001).
[f] Chamberlain et al. (1996).
[g] Guo et al. (2007).
[h] Calculated using ChemAxon/pK_a (www.chemicalize.org).
[i] Skwierczynski and Connors (1993).
[j] Williams (1971).
[k] Yasuda et al. (2013).
⇌ Acid/base equilibrium.