ACD/Labs Technical Document

An Introduction to the Acid Dissociation Constant (pK_a)

Advanced Chemistry Development, Inc.
Toronto, ON, Canada
www.acdlabs.com

pK_a is a measure of the tendency of a molecule or ion to keep a proton, H^+, at its ionization center(s). It is related to the ionization capabilities of chemical species. The more likely that ionization occurs, the more likely a species will be taken up into aqueous solution, because water is a very polar solvent (dielectric constant, ε20 = 80). If a molecule does not readily ionize, then it will tend to stay in a non-polar solvent such as cyclohexane (ε20 = 2) or octanol (ε20 = 10). In biological terms, pK_a is thus an important concept in determining whether a molecule will be taken up by aqueous tissue components or the lipid membranes. It is also closely related to the concepts of pH (the acidity of solution) and log*P* (the partition coefficient between immiscible liquids).

The equilibrium acid ionization constant, K_a, expresses the ratio of concentrations for the reaction:

$$HA + H_2O \rightarrow H_3O+ + A^- \qquad K_a = [H_3O+]\,[A^-] / [HA]$$

where, by convention, it is assumed that the concentration of water is constant, and is absorbed into the K_a definition.

The acid ionization constant varies by orders of magnitude. For example, at 25°C:

acetic acid: $K_a = 1.8 \times 10^{-5}$ phenol: $K_a = 1.0 \times 10^{-10}$

It is easier to refer to such extreme numbers on a logarithmic scale and, again by convention, "p" is used to denote the negative logarithm (base 10): $pK_a = -\log(K_a)$

The K_a values of the compounds above are then easily converted to pK_a values:

acetic acid: $pK_a = -\log(1.8 \times 10\text{-}5) = 4.756$ phenol: $pK_a = -\log(1.0 \times 10\text{-}10) = 10.0$

There is an essential difference between interpreting the pK_a values for molecules vs. ions. A molecule which loses a proton ionizes:

$$HA + H_2O \rightarrow H_3O+ + A^-$$

A low pK_a value, therefore, denotes good aqueous solubility.

An ion which loses a proton, however, de-ionizes:

$$HB+ + H_2O \rightarrow H_3O^+ + B$$

and so a high pK_a value denotes good aqueous solubility.

There is no reason why pK_a values should fall between 0 and 14. For example, sulfuric acid, H_2SO_4, has a negative pK_a for the loss of its first proton:

$$H_2SO_4 \rightarrow HSO_4^- + H^+ \ (pK_a < 0)$$

although normally experiment can only measure pK_a between 1 and 13.

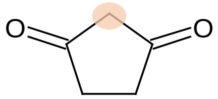
Ionization Centers

The pK_a determination depends on the presence of heteroatoms such as oxygen or nitrogen. Although in principle a pK_a value could be calculated for any atomic center, including carbon, in practice the extrapolation is poor for systems which have a very low amount of ionization. For example, the C–H bonds in methane have such highly covalent character that

$$CH_4 + H_2O \rightarrow CH_3^- + H_3O^+$$

has a vanishingly small probability of occurring.

Some C-H bonds do have measurable ionic character, and these are calculated by ACD/pK_a. For example, the C–H bond of the methylene group at the 2-position in 1,3-cyclopentanedione is highly polarized; its pK_a is predicted to be about 8.9:



Normally, however, a heteroatom is part of the ionization center, and ACD/pK_a is designed to test for the presence of heteroatoms which are capable of forming bonds with sufficient ionic character to have measurable pK_a values, thus enabling reasonable prediction of pK_a for related compounds.

Statistical Factor

Approximated calculation of constants yields a statistical factor that takes into account identical protonation sites. Here is how the statistical factor is defined:

"When a polybasic acid has *n* groups, each of which has an equal probability of losing a proton, the observed pK_a will be less—by (log *n*)—than the pK_a of a closely related monobasic acid. This 'statistical effect' arises because there are *n* equivalent ways of losing a proton but only one site to which the proton can be restored. Similarly, for second proton loss, the correction becomes (log [(*n* – 1) / 2]), then (log [(*n* – 2) / 3]), and so on. Thus, for a molecule such as butanedioic acid ($HOOC–CH_2–CH_2–COOH$), which has two identical acidic groups, loss of a proton from either group leads to the same monoanion. The consequence is that the first ionization constant, pK_{a1}, for the dibasic acid is twice as large as that for the closely related monobasic acid, that is, the observed pK_{a1} is 0.3 (= log2) units less than would be expected from a consideration of factors other than probability. Conversely, the monoanion has only one ionizable proton whereas the dianion has two identical sites for proton addition, so that the second ionization step, pK_{a2}, appears to be weaker by a factor of two, and the observed pK_{a2} to be greater by 0.3 than anticipated. Similarly, for a base with *n* basic centers, the measured pK_a* of greatest magnitude, pK_{aN}, will be greater than anticipated by log *n*, and so on."

D. D. Perrin, Boyd Dempsey and E. P. Serjeant. (1981). Molecular Factors that Modify pKa Values. In *pK_a Prediction for Organic Acids and Bases* (16–17). Netherlands: Springer.

*Referred to as apparent pK_a in ACD/pK_a

Experimental Measurement of pK_a

When comparing calculated pK_a values with experimentally determined data, it is wise to bear in mind how these measurements are carried out.

The determination of pK_a is based on pH measurements for a series of mixtures of the acid and its salt. For pK_a values in the range 2–12, this is frequently done by titrimetric methods. The pH is converted to proton molality, and then K_a is determined by measuring (or estimating) the activity coefficients of species in solution. The temperature, ionic strength, and reference solutions used in these determinations can influence the measured pK_a substantially. For example, benzoic acid was determined to have a pK_a of 4.2 by one experimental group and 4.0 by another.

Another standard method is the spectrophotometric determination of pK_a. This is particularly recommended for very small quantities of sample or for poorly soluble sample. A refinement of this method requires an estimate of the spectra for each form from the data. pK_a values are determined by nonlinear curve fitting, assuming good initial estimates can be chosen. In theory, any kind of spectral data can be used—UV-Vis, IR, NMR, etc., provided that the pH of the solution in which the spectrum was obtained can be measured. A plot of absorbance versus pH will show asymptotes at the absorbance of the conjugate acid and base forms of the molecule. Each wavelength gives different asymptotes, but the same inflection point. Data at enough wavelengths will generate the spectra of the conjugate acid and base forms, even if they can't be measured experimentally, say, for molecules with pK_a outside of the range 2–12. The (common) inflection point is the pK_a. For molecules with multiple ionization sites, a sum of S-shaped curves that need to be deconvolved is obtained. Without good initial estimates, the calculations can be tedious. The better the initial estimate, the faster the convergence. ACD/pK_a can provide good initial estimates for these calculations.

Just as there are aspects of experimental design which affect the accuracy of a pK_a determination, there are also aspects to the physical solution which can lead to apparent disagreement between the calculated and measured pK_a. For example, the presence of a non-negligible tautomeric ratio is one factor that may cause discrepancy between calculated and experimentally measured pK_a values.