A more realistic approach to speciation using the IUPAC Stability Constants Database

Leslie Pettit[‡] and Gwyneth Pettit

Academic Software, Sourby Old Farm, Timble, Otley, LS21 2PW, UK

Abstract: The IUPAC Stability Constants Database (SC-Database), designed to contain all significant published constants, provides the most complete and accessible route to stability constants in the literature. Collection of new constants is becoming less significant since most are now measured to confirm a mechanism or to identify bonding centers, not to provide data for general use. As a result, constants are often measured less rigorously and are of lower accuracy. What are required now are critical evaluations of existing data, coupled with accurate study of a number of important and superficially simpler systems (e.g., complexes with some inorganic ligands).

Calculation of species distribution curves is a major use of stability constants. Ways of reflecting possible errors in the calculated curves are now required. Historically, curves have been drawn as sharp lines, but these could only result from using the exact stability constants with an accurate model for the system. Two techniques for demonstrating the effect of errors are outlined.

Constants are dependent on ionic strength and temperature changes. Specific ion interaction theory (SIT) is the most general method of compensating for ionic strength changes up to about 5 molal. Software to correct constants for ionic strength changes, prepared under recent IUPAC projects, is described.

Keywords: critical evaluation; errors; ionic strength corrections; speciation; stability constants.

INTRODUCTION

Metal-ligand stability constants [1] are the key to the understanding and quantitative use of equilibria of metal ions in solution (e.g., in biological systems, industrial processes, natural waters, and effluent) and are essential for the calculation of species distribution curves. They are increasingly being used by scientists who are not knowledgeable in thermodynamics or quantitative equilibrium chemistry—in fact, many are not even chemists by training. As a result, the constants are frequently misused. Inappropriate constants are often selected, regardless of ionic background and temperature, and the influence of errors in published constants is under-estimated. IUPAC has been active in providing assistance to these scientists by collecting literature, preparing critical evaluations of constants for important metal-ligand systems, and encouraging the design of software to assist in calculating the effects of ionic strength and temperature changes.

THE IUPAC STABILITY CONSTANTS DATABASE (SC-DATABASE)

IUPAC has been active in the collection and publication of stability constants since the 1950s. Four very successful book volumes were published up to 1974. However, data collection and publication in book form was becoming too expensive so no further volumes were planned. In 1982, the IUPAC Commission on Equilibrium Data agreed to examine the possibility of continuing data collection while storing the data on computer. Early work assumed the use of a mainframe computer. In 1986, the authors agreed to design a system which would be compiled and stored on a PC. IUPAC specified that the data and storage should correspond as closely as possible to the book volumes and that it should be possible to distribute the resulting database on 360-K floppy disks. The first version was written in FOR-TRAN but this was soon replaced by compiled *Clipper* for MS-DOS using .DBF-compatible files. When MS-Windows arrived, this version was rewritten using *CA-Visual Objects* and then using *Delphi-Pascal* with *Apollo* database handling routines. All data from the four book volumes were added during the 1990s with assistance from the Royal Society of Chemistry and the Japanese Ministry of Education. Clearly, the specification has extended considerably since the initial design; the database (about 80 MB) is now distributed on CD-ROM. However, for backwards compatibility it is still constrained by the field structure defined in 1986.

Data are accumulated into a relational database of five *master* files holding reference data, ligand data, ligand structures in mol-file format, metal ion data, and experimental data. The latest version of the data (Vn 4.71) contains about half a million constants from over 23 200 references covering over 9500 ligands held in over 111 000 records. The dates of the published constants range from 1877 to 2006/7.

The *distribution* files are prepared from the master files by software written by Academic Software [2], to allow easy and fast searching (using over 50 index files) and a choice of output formats. The database may be searched on any fragment of ligand name or empirical formula, metal ion, author, or reference details or any experimental conditions using interrogation software which has also been written by Academic Software. In particular, ligands may be searched by substructure using any fragment of the 2D structure, *fuzzy*-searching of names is possible, and empirical formulae are searched intelligently. Stability constants and solubility constants are stored as $\log_{10}$ values but quoted errors are not included. Users are recommended always to consult the original references, to satisfy themselves of the accuracy of the constants. Where available, IUPAC *Recommended* and *Provisional* values are identified. The data held for a single record are illustrated in Fig. 1. Constants can be transferred interactively to the speciation program provided with the package. Distribution files, in a compressed format for installation on a user's PC, require about 15 MB.
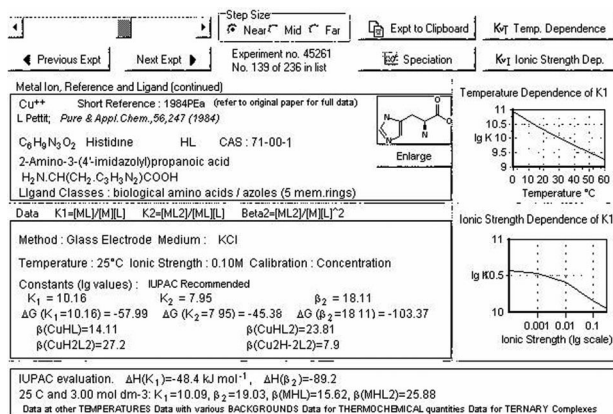

Fig. 1 Information stored in a record of SC-Database.

New stability constants reported in the literature are now, in general, not designed for general use by other scientists. Rather, they are measured to confirm a reaction mechanism or to identify bonding centers. The ligands studied are becoming more esoteric (e.g., calixarenes, or long-chain peptides), and the constants are often measured less rigorously and to a lower precision than earlier measurements. As a result, the constants are frequently of less general use and are less appropriate for inclusion in the IUPAC Stability Constants Database. The abstraction and compilation of new literature data will, therefore, probably cease during 2009 but the database will continue as an invaluable and irreplaceable resource, which should be maintained and made available into the future. Ideally, the interrogation software should be rewritten as a more future-proof Internet database and freed from the field structure defined in 1986.

SPECIES DISTRIBUTION CURVES

A major use of stability constants is the calculation of species distribution curves (speciation curves) which show graphically the distribution of the complexed species (including partially soluble species) formed between metal ions and ligands as a function of pH or pL. For the curves to be meaningful, calculations must be based on an accurate model for the system. Constants are frequently calculated on simplified models, often omitting partially hydrolyzed, mixed, or polynuclear species. The resulting speciation curves are usually valueless.

The curves are normally drawn as lines of notional (infinitesimal) thickness. Figure 2a shows the curves for the simplest equilibrium, $H^+ + L^- = HL$ with lg $K = 7.00$. These could only result from knowledge of the exact stability constant (7.00 ± 0.00), using the exact model for the system. Standard deviations in lg K (σ values) calculated from very careful measurements should be around 0.01. However, there are generally systematic errors; measurements carried out by different research groups using different equipment generally show standard deviations of from ±0.02 up to ±0.1 or even more. These error ranges should be reflected in the calculated speciation curves. Error bars could be included with a range of, e.g., 3σ, but these do not reflect the varying probability. A better approach would be to draw the curves as bands, with a width of $\pm 3\sigma$ and with the intensity of color dropping to zero following a normal distribution between -3σ and $+3\sigma$. These more realistic curves are demonstrated in black and white in Fig. 2b. Software is required to perform these calculations in the general case with color intensity following a normal distribution. This is more difficult than may be expected. Realistic standard deviations on published constants can be calculated (or estimated), but correlation coefficients are more difficult to obtain and these will influence the width of calculated bands considerably.
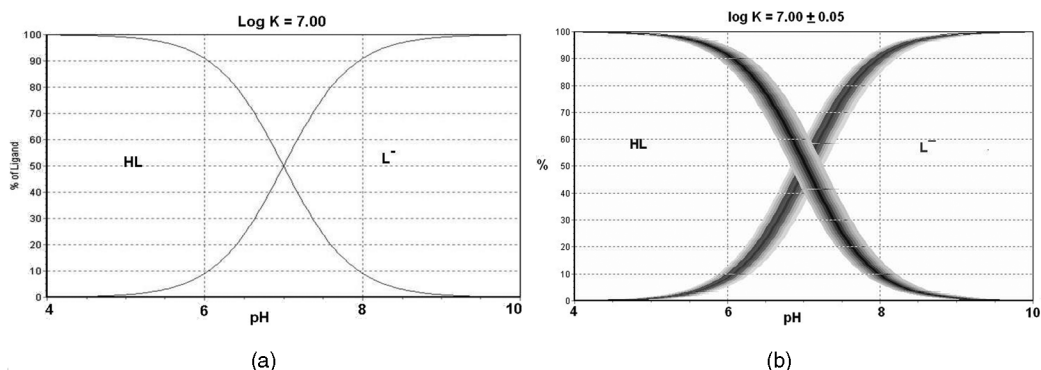

Fig. 2 Typical speciation curves: (a) as normally drawn, (b) proposed format.

One approach tested was to ignore correlation coefficients and, for each constant, to calculate a number of constants (e.g., 10 or more) which followed a normal distribution about the starting value with its standard deviation, σ. Speciation calculations were then performed on all combinations of the calculated ranges of constants. The results of these calculations for the formation of H_2L (lg K_{HL} = 5.00 ± 0.10 and lg β_{H2L} = 8.00 ± 0.03) and for M + HL (lg K_{HL} = 8.00 ± 0.03, lg K_{ML} = 3.00 ± 0.10) are shown in Fig. 3. Ideally, a much larger number of values within the error range of each constant should be selected, but such calculations would become prohibitively slow when a large number of constants are also included. Graphical representation of the large number of resulting curves would also be problematical.
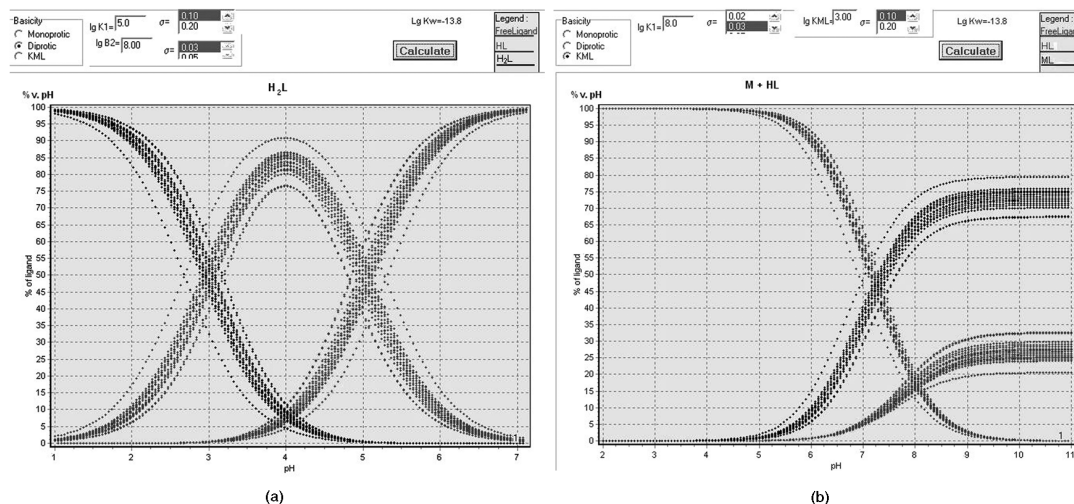

Fig. 3 Experimental speciation curves showing error ranges: (a) H_2L, (b) HL + M.

A more general, but less rigorous, approach has been used by modifying the speciation program provided with SC-Database to allow for real-time adjustment of stability constants used for the calculations. If the calculated curves are retained while a constant is adjusted (e.g., by ±2σ) the curves for that and neighboring species will be broadened. Aluminum hydrolysis is particularly important industrially and environmentally. At intermediate pH it is readily precipitated as the hydroxide, but this redissolves at higher pH. Just before precipitation, a number of polymerized species have been identified, several existing in significant concentrations. The ion $[Al_{13}(OH)_{32}]^{7+}$ [lg $\beta(Al_{13}H_{-32})$ = –98.7] precedes formation of insoluble $Al(OH)_3$ [lg $*Ks(AlH_{-3})$ = 10.0] and is particularly important. Figure 4a shows the calculated speciation curves for a 5-μM solution over the pH range of 4–8. Standard deviations in these constants would be at least ±0.04. Figure 4b demonstrates the effect of changing these constants by ±0.08 lg units—this is probably an under-estimate of the error range expected.

From Fig. 4a, the point of precipitation can be predicted accurately (about pH 5.5). However, when errors are taken into account (Fig. 4b), the pH of precipitation is much more difficult to predict; it could have any value between 5.2 and 5.8. This will have significant implications when using speciation to predict the state of aluminum in solution.
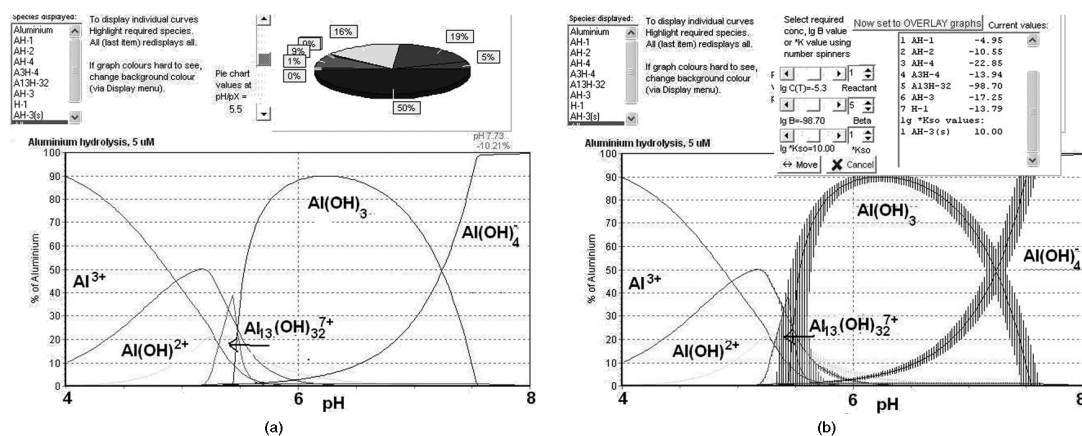
Fig. 4 Speciation curves for Al^{3+} hydrolysis: (a) omitting errors, (b) including small errors in lg $\beta(Al_{13}OH_{32}^{7+})$ and lg $*Ks[Al(OH)_3]$.

IONIC STRENGTH CORRECTIONS

SIT is the most general way of correcting stability constants for changes in ionic strength. However, it is technically complicated and not easily accessible to general scientists who use stability constants. To perform the corrections, it is necessary to know the SIT parameters for each of the interacting ions in the medium used (e.g., NaCl). A number of years ago, IUPAC approved a project to write a program which would assist in these calculations by incorporating a compilation of known SIT parameters and calculation routines for applying SIT corrections to stability constants for the general reaction

$$x M^{n+} + y L^{m-} + z H_2O = [M_xL_y(OH)_z]^{(xn-ym-z)+} + z H^+$$
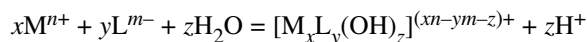
in a range of aqueous media. The program (SIT.EXE) is generally available as part of the Aqueous Solutions package [3] and with SC-Database. It may also be used to calculate some SIT parameters from a set of lg β values against ionic strength.

Further programs were then added to create a package of programs of general use to scientists working quantitatively with solution equilibria in aqueous solutions, Aqueous Solutions. This package contains in addition to SIT.EXE, the speciation program outlined above and a titration simulation program which interactively plots a titration curve, speciation curves, and pie chart, as constants and concentrations are varied. Drs. Igor Sukhno and Vladimir Buzko have added significantly to this package by extending the SIT program. They have also included programs to correct stability constants for ionic strength changes using approaches based on several adaptations of the Pitzer equation in a wide range of media, especially natural waters of varying salinity, a program to calculate the solubility of dioxygen, dinitrogen, argon, and air in saline waters and a program to correct stability constants for temperature changes using seven extensions of the original van't Hoff equation. The complete package is in the public domain and may be downloaded from the sites given in ref. [3].

REFERENCES

1. For definitions of terms used, see (a) IUPAC. *Compendium of Chemical Terminology*, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: http://goldbook.iupac.org (2006–) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins; (b) IUPAC Analytical Chemistry Division. *Compendium of Analytical Nomenclature*, 3rd ed. (the "Orange Book"); Prepared for publication by J. Inczédy, T. Lengyel, A. M. Ure. Blackwell Science, Oxford (1998).
2. L. D. Pettit, K. J. Powell. The IUPAC Stability Constants Database (SC-Database), current data version 4.71 (2008), Academic Software, Otley, UK, <http://www.acadsoft.co.uk>.
3. The Aqueous Solutions package may be downloaded from: (a) <http://www.acadsoft.co.uk>, (b) <http://public.kubsu/ru/aquasolsoft/>, or (c) <http://www.iupac.org/projects/2000/2000-003-1-500.html>.