EQNMR: A Computer Program for the Calculation of Stability Constants from Nuclear Magnetic Resonance Chemical Shift Data

Michael J. Hynes
Chemistry Department, University College, Galway, Ireland

A computer program has been elaborated which uses the complexation-induced displacements of NMR chemical shifts to calculate the stability constants for the general reaction (i) which gives the generalised stability constant (ii). The program can deal with data from a wide variety of reactions including proton

$$mM + nL + jH \rightleftharpoons M_mH_jL_n \quad \text{(i)}$$

$$\beta_{mjn} = [M_mH_jL_n]/[M]^m[H]^j[L]^n \quad \text{(ii)}$$

equilibria, metal-ion hydrolysis and metal–ligand interactions. It can also deal with situations where both ligand proton equilibria and complex-formation reactions must be considered.

The determination of equilibrium constants forms an important part of research in physical, inorganic and biochemistry. The general form of the equilibrium reaction and the generalised stability constant are shown in equations (1) and (2). The most

$$mM + nL + jH \rightleftharpoons M_mH_jL_n \quad (1)$$

$$\beta_{mjn} = [M_mH_jL_n]/[M]^m[H]^j[L]^n \quad (2)$$

widely used methods for the determination of stability constants are potentiometry and spectrophotometry and computer methods are well developed for the processing of the experimental data obtained.[1,2]

Stability constants can be determined by NMR spectrometry when the species are in rapid exchange on the NMR time-scale[3] and when there is a variation in the chemical shift of a suitable nucleus on formation of the complex species. However, the methodologies for the treatment of NMR chemical shift data are considerably less well developed than those for potentiometric and spectrophotometric procedures. Some of the procedures necessitate the use of a large excess of one of the reactants[4–6] while others require that both reactants be present in equimolar concentrations[7] and none appears to deal with NMR chemical shift data in the general way that SUPERQUAD[8] and SQUAD[9,10] respectively deal with potentiometric and spectrophotometric data. Recent textbooks[1,2] dealing with the computation of solution equilibria make no reference to computer programs for the calculation of stability constants from NMR chemical shift data. Nevertheless, NMR spectroscopy can prove very useful for the determination of equilibrium constants[11–15] and as seen above various authors have developed computational procedures. This method is particularly useful in non-aqueous solution where potentiometric methods may not be readily applicable. It should be noted however that its use requires concentrations which are generally much higher than those used in either potentiometric or spectroscopic work and this difference can influence both the stoichiometry and concentration of the species formed.

The program EQNMR can be used to evaluate equilibrium constants and chemical shifts in systems where all the species are in rapid equilibrium under the experimental conditions used and where the NMR chemical shift of some nucleus varies with the degree of complex formation. In this event only a single NMR resonance is observed for the nucleus ^{nn}X. Consider a reaction of the type shown in equation (3) where M is the

$$mM + nL \rightleftharpoons M_mL_n \quad (3)$$

nucleus whose chemical shift is being monitored by NMR spectroscopy. The chemical shift of this resonance is given by equation (4) where δ_{calc} is the weighted average of the chemical

$$\delta_{calc} = \sum_{m=1}^{m=i} \sum_{n=0}^{n=j} \frac{\delta_{mn} m[M_mL_n]}{[M]_{total}} \quad (4)$$

$$[M_mL_n] = \beta_{mn}[M]^m[L]^n \quad (5)$$

shifts of the various M-containing species present, M_mL_n, where M represents the 'free' uncomplexed form of M, L represents the ligand, and i and j represent the maximum values of m and n respectively. Substituting for M_mL_n in equation (4) from (5), gives (6). Thus the problem resolves itself into determining the

$$\delta_{calc} = \sum_{m=1}^{m=i} \sum_{n=0}^{n=j} \frac{\delta_{mn}\beta_{mn} m[M]^m[L]^n}{[M]_{total}} \quad (6)$$

optimum values for the various δ_{mn} and β_{mn} which best fit the experimental chemical shift data.

The program EQNMR can be used to calculate these parameters. The current version is model independent and all the information regarding the model and the manner in which the chemical shifts of both free and complexed species contribute to δ_{obs} is contained in the data file. The program contains four main sections. Section one reads the input data. This consists of: (*a*) the concentrations of the various reagents, pH values if relevant, together with the measured chemical shifts and the weighting factors for the data points; (*b*) details of the model in the form of the stoichiometric coefficients of the complex species present; (*c*) initial guesses for the various parameters (stability constants and chemical shifts) to be fitted; (*d*) details of any constraints, *i.e.* values for any parameters which are to be held constant or $\geqslant$ or $\leqslant$ to some value; and (*e*) details of the manner in which all the species (both free and complexed) contribute to the overall chemical shift of the nucleus being monitored.

Section two of the program consists of the custom-written

subroutine SAMPGEN which utilises the imput data from section one to perform two tasks. First, it calculates the concentrations of all species present in solution using either the initial guesses of the stability constants or the improved values as the refinement proceeds. The COMICS[16] algorithm is used to calculate the concentrations of all free and complexed species as it has been shown to give more reliable convergence than other alternatives, particularly in the case of complicated systems.[2] The program then computes the calculated chemical shift, δ_{calc}, using equation (6), taking account of both free and complexed species contributing to the overall chemical shift of the nucleus being monitored.

Section three of EQNMR contains the non-linear least-squares subroutines which carry out the refinement of the various parameters using the Levenberg–Marquardt method.[17,18] The coding used is a modification of the well proven NIHH22/NIHH23 routines developed by Fletcher and Shrager[19,20] which have been used in this laboratory for a number of years.

Section four of EQNMR contains the output routines. Output consists of printed tables of the input data, best-fit values of the chemical shifts and formation constants together with estimates of their errors. The values of δ_{calc} which are calculated using the 'best-fit' values of the parameters are also printed. A table of the concentrations of all species present at each experimental point in the titration is also produced. Output is also provided in the form of two plots. The first of these displays the variation in the concentrations of the various species present as the titration proceeds. The second consists of both the experimental and calculated chemical shifts against the concentration of the titrant, together with the residuals (in magnified form) against the titrant concentration.

The program EQNMR can deal with chemical shift data from a wide variety of reactions where the equilibria can be expressed in terms of equation (1). These include complex-formation reactions involving reaction of a ligand with either proton or metal ion, metal-ion hydrolysis and monomer–dimer equilibria. The nucleus the chemical shift of which is being monitored can be either the central metal ion M[11,12] or an atom located on one of the ligands.[13–15] Various models can be readily evaluated, *e.g.* models containing only species ML, ML_2 and M_2L or any combination of them. In most instances visual inspection of the graphical output is sufficient to determine the 'best-fit' model. The value of the graphical output of the raw and best-fit data together with the residuals cannot be over-emphasised as it highlights both systematic deviations between the experimental chemical shifts and those calculated using the 'best-fit' parameters and any deficiencies in the model. In addition, quantitative comparison of the fits is carried out using either the estimated errors in the various parameters or, more conveniently, using the 'merit function' shown in equation (7)

$$R = 100\left(\frac{\Sigma W_i(\delta_{obs} - \delta_{calc})^2}{\Sigma W_i(\delta_{obs})^2}\right)^{\frac{1}{2}} \quad (7)$$

where W_i is the weight attributed to observation i. Where the chemical shifts are of similar magnitude this function usually enables a choice to be made between potential models. Normally unit weights are used, although alternative weighting schemes can be implemented if required.

The program EQNMR does not suffer from the limitations of previous programs[7] in that it can readily deal with systems where the concentrations of the reactants are not equal and it does not require that one reactant be present in large excess. The concentrations used are dictated solely by the requirements of good experimental design. The chemical shifts of the various species can be constrained to directly measured experimental values where that is possible or they can be treated as variable parameters.

The program was extensively tested using synthesised model data. Suitable concentrations were chosen and shift data were calculated for three systems. System one contained only the species M and ML. In such a system a maximum of three parameters have to be refined, δ_M, δ_{ML} and β_{11}. Similar analyses were carried out for systems containing M, ML and ML_2 in which a maximum of five parameters have to be refined, δ_M, δ_{ML}, δ_{ML_2}, β_{11} and β_{12}. A further test involved analysis of a system containing the species M, ML and M_2L and the constants β_{11} and β_{21}. The best-fit parameters were independent of the initial starting parameters. In addition to the above, tests were carried out in order to ensure that the absolute magnitude of the measured chemical shifts had no effect on the calculated values of the stability constants. Constant values of from 100 to −100 ppm were added to the measured chemical shifts following which the stability constants were re-evaluated. No significant differences were obtained.

Software and Hardware Environment.—The program is written in FORTRAN 77 and consists of approximately 3000 lines. In order to simplify data input, free-format input has been used where possible. It has been run on VAX 8700 and 6000 computers without difficulty. The graphics package used is GINO-F. In order to facilitate conversion for use with other graphics packages, the data required to create the various plots are written to separate files. These are then used as input to a separate plotting module (FITPLT).

The FORTRAN listings of the EQNMR and FITPLT programs can be provided as ASCII files on an MS-DOS/PC-DOS formatted disk or diskette.

References

1 M. Meloun, J. Havel and E. Högfeldt, *Computation of Solution Equilibria*, Ellis Horwood, Chichester, 1988 and refs. therein.
2 D. J. Leggett, *Computational Methods for the Determination of Formation Constants*, Plenum, New York and London, 1985.
3 R. G. Bryant, *J. Chem. Educ.*, 1983, **60**, 933.
4 I. Armitage, G. Dunsmore, L. D. Hall and A. G. Marshall, *Can. J. Chem.*, 1972, **50**, 2119.
5 A. Arduini, I. M. Armitage, L. D. Hall and A. G. Marshall, *Carbohydr. Res.*, 1973, **31**, 255.
6 D. M. Rackham, *Spectrosc. Lett.*, 1981, **14**, 117.
7 I. Horman and B. Dreux, *Anal. Chem.*, 1983, **55**, 1219.
8 P. Gans, A. Sabatini and A. Vacca, *J. Chem. Soc., Dalton Trans.*, 1985, 1195.
9 D. J. Leggett and W. A. E. McBryde, *Anal. Chem.*, 1975, **47**, 1065.
10 D. J. Leggett, *Anal. Chem.*, 1977, **49**, 276.
11 M. J. Hynes, J. M. Keely and J. McManus, *J. Chem. Soc., Dalton Trans.*, 1991, 3427.
12 D. Cunningham, J. McManus and M. J. Hynes, *J. Organomet. Chem.*, 1990, **393**, 69.
13 P. G. Harrison, P. Brown, M. J. Hynes, J. M. Keely and J. McManus, *J. Chem. Res.*, 1991, *S* 174.
14 M. J. Hynes, J. K. Keely, E. E. Lee, P. McArdle, J. O'Callaghan and N. Gavin, *J. Chem. Soc., Perkin Trans. 2*, 1991, 363.
15 P. G. Harrison, P. Brown, J. McManus, M. J. Hynes and J. M. Keely, *Inorg. Chim. Acta*, 1991, **190**, 209.
16 G. Ginzburg, *Talanta*, 1976, **23**, 149.
17 K. Levenberg, *Q. Appl. Math.*, 1955, 164.
18 D. W. Marquardt, *J. Soc. Ind. Appl. Math.*, 1963, **2**, 431.
19 J. E. Fletcher, National Institute of Health, Bethesda, MD, 1973.
20 R. I. Shrager, *J. Assoc. Comput. Mach.*, 1970, **17**, 446.

Received 7th July 1992; *Paper* 2/03591J