International Journal of Pharmaceutics, 5 (1980) 49–58

COMPUTATION OF MULTICLASS DRUG–PROTEIN BINDING PARAMETERS

JOHN B. WHITLAM and KENNETH F. BROWN

Department of Pharmacy, University of Sydney, Sydney, N.S.W., 2006 (Australia)

(Received September 7th, 1979)
(Accepted January 9th, 1980)

SUMMARY

Normal scatter has been introduced into the free drug concentration variable, in sets of perfect binding data consisting of free (D_f) and corresponding total (D_t) drug concentrations. The standard deviation of error ranged from 0.1×10^{-4} M to 1×10^{-8} M over the free drug concentration range, 9×10^{-4} M to 1×10^{-8} M. These simulated data for 4 hypothetical drugs together with real equilibrium dialysis data for two non-steroidal anti-inflammatory drugs, ibuprofen and naproxen, have been used to evaluate various non-linear regression procedures and commonly used weighting schemes, for calculating the binding constants from the Scatchard multiclass drug–protein binding model.

It is statistically most desirable to regress unmanipulated primary experimental data, the independent variable being that measured most accurately. However, regression of unweighted D_f on D_t results in parameter estimates far removed from their true values. Appropriate data weighting is essential for this approach to be successful. Regression of either weighted or unweighted D_f on the molar binding ratio, $\bar{\nu}$, also produces poor estimates. It is shown that the computationally simpler regression of $\bar{\nu}$ on D_f performs as well as D_f on D_t (weighted 1/y), especially where good experimental data are obtained. This is of interest considering the greater computational sophistication required to regress D_f on D_t.

INTRODUCTION

In recent years, sophisticated computed technology and statistical techniques have been applied increasingly to determine drug–protein binding parameters. Interactions are most commonly described using the generalized Scatchard model,

$$\bar{\nu} = \sum_{i=1}^{i=i} \frac{n_i K_i D_f}{1 + K_i D_f} \qquad (1)$$

where $\bar{\nu}$ is the average number of moles of ligand bound per mole of protein, in equilibrium with free ligand, D_f and n_i is the number of sites in the i^{th} class of association constant K_i.

Estimations of stoichiometry (Σn_i) and association strength (ΣK_i) for particular drugs often disagree, between different laboratories. This has been attributed, at least in part, to choice of experimental technique and conditions. In addition, different albumin preparations may exhibit different binding tendencies for the same ligand (Nilsen and Jacobsen, 1976; Whitlam et al., 1979).

An equally important consideration, which may contribute substantial error to the estimation of binding parameters is the data reduction technique itself (Madsen and Robertson, 1974; Perrin et al., 1974; Boxenbaum et al., 1974; Plumbridge et al., 1978).

Where binding of a drug molecule occurs to only one class of sites there is little difficulty in obtaining relatively unbiased parameter estimates, either graphically or by least-squares, using one of the linear transformations of the Scatchard equation (Steinhardt and Reynolds, 1969). Even so, least-squares minimization of non-linearized data has been demonstrated to provide estimates of greater precision (Madsen and Robertson, 1974).

Traditionally, multiple class binding was expressed as curved Scatchard plots. These data were then treated by one of a variety of methods to estimate the binding parameters. From the pharmacodynamic standpoint it is usually only primary site binding which is important.

One approach was to ignore secondary sites and estimate n_1 and K_1 from the initial slope (Dearden and Tomlinson, 1970; Judis, 1972). However, contribution to the initial slope by secondary sites may result in substantially biased primary site estimates (Crooks and Brown, 1973; Chamness and McGuire, 1975). Other authors constrained n_1 and n_2 to be integers and then refined K_1 and K_2 until the regenerated curve adequately described the data (Eichman et al., 1962; Rosen, 1970). Crooks and Brown (1973) employed a method based on that described by Hart (1965), where the i^{th} association constant was expressed as the i^{th} root of a polynomial and the binding capacity was obtained from the solution of a system of i simultaneous linear equations.

Curve fitting by non-linear least-squares minimization of the correct model should provide the most reliable, mathematically correct, parameter estimates. Madsen and Robertson (1974), who considered single-class binding, stressed that selection of $\bar{\nu}$ or D_f as dependent variable depends on which has greatest associated error. In multiclass binding, the selection of D_f as dependent variable involves the added difficulty of expressing D_f explicitly in terms of $\bar{\nu}$.

This paper examines the treatment of binding data by non-linear regression for drugs that exhibit multiclass binding (Eqn. 1, for $i > 1$).

DATA TREATMENT

The fitting of correct equations to perfect data by least-squares minimization is, necessarily, exact. However, for real imperfect data, assumptions must be made regarding the relationship between the dependent and the independent variables and their respective errors (Daniel and Wood, 1971). It is assumed that the model is correct, the experimental data are typical of the population they represent, the independent variable is

known without error and that the dependent variable errors are: (a) statistically uncorrelated; (b) normally distributed; (c) of equal variance; and (d) not correlated with errors associated with the independent variable. The parameters evaluated will be most exact and unbiased only when these conditions are satisfied in the experimental and treatment design.

Practical determination of binding by separative techniques such as equilibrium dialysis, gel filtration and ultrafiltration results in primary data relating total drug concentration (D_t) and D_f [1]. Binding parameters have usually been obtained by manipulation of these variables to regress $\bar{\nu}$ on D_f, in accordance with Eqn. 1. Where extensive binding occurs and the free fraction is small it would seem more appropriate to regress D_f on $\bar{\nu}$ since the former variable would be measured with greatest uncertainty. Where multiclass models are employed, this is almost never considered because of the implicit nature of the function relating D_f in terms of $\bar{\nu}$. In addition, with data obtained using separative techniques either approach may be criticized since the bound drug concentration (D_b) is not determined directly but is calculated from the difference between D_t and D_f. Thus $\bar{\nu}$ is given by,

$$\bar{\nu} = \frac{D_t - D_f}{P_t} \tag{2}$$

where P_t is the total protein concentration. Both $\bar{\nu}$ and D_f contain the same errors associated with the free drug concentration. Thus the error in each will be correlated, resulting in biased parameter estimates.

Theoretically, a more correct approach is to regress the raw experimental data, the independent variable being that associated with least error. Since many drugs are strongly bound to albumin, D_f will usually be much smaller than D_t and its measurement in such cases is associated with greatest uncertainty.

Consider the case where the albumin binding of a high affinity drug is appropriately described by a two-class binding model. Eqn. 1 can then be expressed in terms of D_f and D_t,

$$D_t = \frac{P_t n_1 K_1 D_f}{1 + K_1 D_f} + \frac{P_t n_2 K_2 D_f}{1 + K_2 D_f} + D_f \tag{3}$$

In an expression such as this, where D_f is judged to be the dependent variable, D_f cannot simply be expressed explicitly as a function of the independent variable, D_t. However, by setting the equation equal to zero by re-arranging it to the form,

$$g(D_f, D_t, P) = 0 \tag{4}$$

where P contains the binding parameters, D_f may be solved for any particular values of D_t, ns and Ks numerically using Newton's method (Southworth and Deleeuw, 1965).

[1] Dynamic dialysis, which yields data relating the change in D_t to time, has been rigorously examined previously (Veng Pedersen et al., 1977a).

Determination of D_f is rapid since the iteration exhibits quadratic convergence. Newton's equation for the iteration may be written

$$D_{f'j+1} = D_{f'j} - \frac{g(D_f)}{g'(D_f)} \tag{5}$$

where

$$g(D_f) = \frac{P_t n_1 K_1 D_f}{1 + K_1 D_f} + \frac{P_t n_2 K_2 D_f}{1 + K_2 D_f} + D_f - D_t \tag{6}$$

and $g'(D_f)$ is the derivative of $g(D_f)$.

Experiments to estimate binding parameters usually require measurement of free and total drug concentrations spanning several orders of magnitude. It is most improbable that the variance associated with the analytical determination of the dependent variable will be constant. Thus consideration must be given to weighting the data. The function to be minimized (SS) is,

$$SS = \sum_{k=1}^{N} W_k(y_k \text{ obs.} - y_k \text{ calc.})^2 \tag{7}$$

where W_k is the weight of the k^{th} datum point. Frequently, data points are *unweighted*, i.e. $W_k = 1$. More correctly data should be weighted by the inverse of their associated variance, s^2. That is,

$$W_k = 1/s^2 \tag{8}$$

However, the nature of binding data is such that usually only one value of D_f is observed for each D_t, and consequently the variance is unknown. In such cases schemes such as the inverse of the observation or the inverse of the squared observation are commonly used.

EXPERIMENTAL

Perfect binding data (28 pairs of D_f and D_t) were generated using Eqn. 1 for 4 hypothetical drugs, each of which was assumed to bind to two classes of sites by various known binding parameters (Table 1). Normal scatter was then introduced into the *observed* dependent variable, D_f, using a standard program [2] and the following procedure. Values of D_f from largest to smallest were subdivided into groups of 3. The standard deviation of the mantissas of each group was varied by increments of 0.1 and ranged from 0.1 to unity. The likely experimental situation was thus more closely simulated by inducing greater error as the actual value of D_f decreased and the assumption of equal

[2] Texas Instruments TI58 programmable calculator and program number ML15, Texas Instruments, Dallas, Texas.

TABLE 1

VARIATION IN CALCULATED BINDING PARAMETERS FOR THE 4 HYPOTHETICAL DRUGS RESULTING FROM USE OF DIFFERENT REGRESSION PROCEDURES

Regression scheme (y vs x)		Weighting factor (W_k)	n_1	K_1 (M^{-1})	n_2	K_2 (M^{-1})	ω
Drug 1 [a]			1	1×10^6 [b]	7	1×10^4 [b]	
D_f	D_t	1	1.25	0.52	6.69	0.91	0.86
D_f	D_t	$1/y$	1.08	0.82	6.67	1.01	0.32
D_f	D_t	$1/y^2$	0.71	1.67	6.17	1.58	1.66
$\bar{\nu}$	D_f	1	1.00	0.89	6.62	1.08	0.24
D_f	$\bar{\nu}$	1	1.33	0.38	6.66	0.87	1.12
D_f	$\bar{\nu}$	$1/y$	1.19	0.67	6.70	0.93	0.63
D_f	$\bar{\nu}$	$1/y^2$	0.72	1.66	6.25	1.53	1.58
Drug 2 [a]			1	1×10^5 [b]	7	1×10^3 [b]	
D_f	D_t	1	1.42	0.52	8.21	0.61	1.46
D_f	D_t	$1/y$	1.07	0.91	5.06	1.56	1.00
D_f	D_t	$1/y^2$	1.23	0.79	8.92	0.65	1.06
$\bar{\nu}$	D_f	1	1.22	0.75	5.77	1.11	0.76
Drug 3 [a]			1	1×10^4 [b]	7	1×10^2 [b]	
D_f	D_t	1	1.23	0.88	2.60	1.12	1.10
D_f	D_t	$1/y$	1.07	1.29	4.87	0.92	0.74
D_f	D_t	$1/y^2$	0.97	2.03	5.06	1.08	1.42
$\bar{\nu}$	D_f	1	1.12	1.03	7.50	0.51	0.71
Drug 4 [a]			1	1×10^5 [b]	3	1×10^3 [b]	
D_f	D_t	1	0.56	1.50	2.72	2.25	2.28
D_f	D_t	$1/y$	0.76	1.54	3.98	0.90	1.21
D_f	D_t	$1/y^2$	0.80	1.53	5.29	0.61	1.88
$\bar{\nu}$	D_f	1	0.65	1.95	3.17	1.44	1.80

[a] Indicates actual binding parameters from which scattered data were generated.
[b] Data in column below should be multiplied by the appropriate exponent of 10.

variance in the observed data was forcibly violated. The standard deviation of error ranged from 0.1×10^{-4} M to 1×10^{-8} M over the free drug range 9×10^{-4} M to 1×10^{-8} M. The applicability of the data treatment was tested by regression of the scattered data, D_f against D_t using the 3 commonly used weighting schemes 1, $1/y$, $1/y^2$. For comparison, data were also treated by regression of D_f on $\bar{\nu}$ and also $\bar{\nu}$ on D_f. Newton's iteration (Eqn. 5) was used to establish the functional relationship between D_f and D_t or D_f and $\bar{\nu}$ using the interactive time-sharing program FUNFIT for non-linear least-squares regression (Veng Pedersen, 1977b).

Data generated for each of the 4 hypothetical drugs covered the same range of total drug concentration, D_t, and thus markedly different free concentrations. Because of this and the way normal scatter was induced in D_f, data for each hypothetical drug contained

error of different magnitude and acted as a separate test of the regression procedure.

The performance of each regression procedure was evaluated by calculating the sum of the absolute values of the normalized deviations of the parameters from true values (ω).

$$\omega = \sum_{i=1}^{4} \frac{|\bar{P}_i - P_i|}{\bar{P}_i} \tag{9}$$

where $\bar{P}_i$ is the i^{th} parameter from which the hypothetical data were generated and P_i is the estimated value. ω expresses numerically the collective deviation of the 4 parameters and may be used to rank the regression procedures by indicating better performance as $\omega \to 0$.

Binding parameter estimates for the interaction of the non-steroidal anti-inflammatory drug (NSAID), ibuprofen, with 1% HSA at 37°C obtained by equilibrium dialysis against 0.033 M phosphate buffer and subsequent regression of $\bar{\nu}$ on D_f have been reported previously (Whitlam et al., 1979). These data together with data obtained in similar fashion for the NSAID, naproxen, were treated by the various regression schemes to assess the applicability of each scheme for real data.

RESULTS AND DISCUSSION

Details of the binding parameter estimates n_1, K_1, n_2 and K_2 for the 4 hypothetical drugs are given in Table 1 together with the actual values from which the scattered data were generated.

It is important to note that regression of D_f on D_t, where data are *unweighted* ($W_k = 1$), results in parameter estimates far removed from the true values. This can be explained by the fact that D_f spans several orders of magnitude and very small values are effectively ignored in the minimization procedure. The squared deviations, for very small D_f values, have a negligible effect on the sum of squared deviations. Furthermore, in generating the data the assumption of equal variances of error in D_f was deliberately violated to approximate real data more closely.

Alternatively, parameters estimated by minimizing the *weighted* sum of squared deviations ($W_k = 1/y$ or $1/y^2$) indicate that weighting by $1/y$ generally results in much closer estimates, perhaps reflecting the nature of the error in these data. The poorer fits obtained using $1/y^2$ suggest that this weighting is inapplicable. The most appropriate weighting scheme can be chosen for replicated data depending on the relationship between the variance and the mean value of the dependent variable (Wagner, 1976). However, most binding experiments involve single-point or at best duplicate determinations. In these cases, an appropriate empirical weighting procedure can be selected by examination of residual plots. Such plots for the hypothetical drug, $n_1 = 1$, $K_1 = 10^5\ M^{-1}$, $n_2 = 7$ and $K_2 = 10^3\ M^{-1}$ are shown in Fig. 1 where unweighted normalized residuals are plotted against the negative log of the calculated value for various regression procedures. Normalized residuals are the ratio of the residual to y_{calc} and permit more meaningful comparison of the various fits. The dependent variable, y_{calc}, was expressed in logarithmic form for scaling purposes.

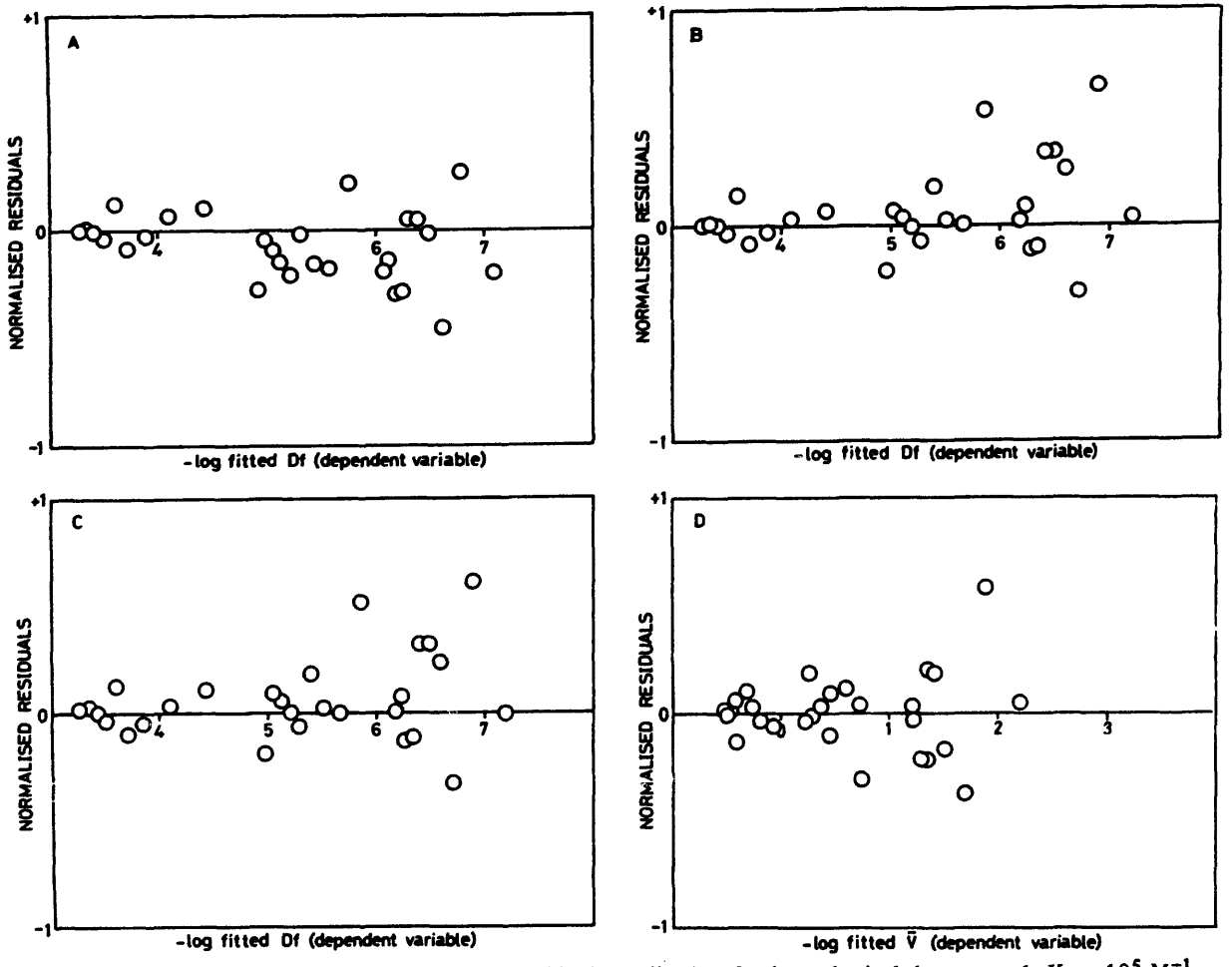

Fig. 1. Influence of regression procedure upon residual distribution for hypothetical drug; $n_1 = 1$, $K_1 = 10^5\ M^{-1}$, $n_2 = 7$, $K_2 = 10^3\ M^{-1}$. A = D_f vs D_t ($W_k = 1$); B = D_f vs D_t ($W_k = 1/y$); C = D_f vs D_t ($W_k = 1/y^2$); D = $\bar{\nu}$ vs D_f ($W_k = 1$).

As expected the distribution of residuals in each case is *wedge-shaped* (Daniel and Wood, 1971) consistent with the increasing variance in introduced error with a decrease in dependent variable. Weighting the data 1/y (Fig. 1B) provides a more even distribution of positive and negative residuals than either unity (Fig. 1A) or $1/y^2$ (Fig. 1C). A comparison of the calculated parameters, using the various regression procedures, with the actual values (Table 1) and subsequent ranking using ω indicates that the more usual technique of $\bar{\nu}$ on D_f (Table 1, Fig. 1D) is as good as regression of D_f on D_t weighted 1/y and both are clearly better than the other procedures investigated. This is of particular interest when considering the more sophisticated computational procedure required to regress D_f on D_t. Certainly the regression of appropriately weighted experimental values of D_f on D_t has a sounder theoretical basis than regressing $\bar{\nu}$ on D_f and would be expected to perform better as scatter in the data increases. Error in $\bar{\nu}$ is relatively small and unimportant when $\bar{\nu}$ is treated as a dependent variable. Since the variable spans a narrower range, the smaller data points assume greater significance in the minimization procedure. The appreciable error in D_f is ignored since the independent variable is assumed to be exact. Hence the fit obtained with this procedure is remarkably good without recourse to weighting. It has been suggested, (Madsen and Robertson, 1974) for drugs bound strongly to multiple classes of sites, that it is appropriate to regress D_f on $\bar{\nu}$. However, the data in Table 1 indicate that this is not always the case. Regression of either weighted or unweighted D_f on $\bar{\nu}$ produces poorer parameter estimates than either $\bar{\nu}$ on D_f or D_f on D_t (weighted 1/y). The only explanation for the poorer performance of D_f on $\bar{\nu}$ relative to D_f on D_t is that in the former case the errors between variables are correlated by calculation of $\bar{\nu}$.

Comparison of the binding parameter estimates obtained for the binding of ibuprofen and naproxen to 1% HSA at 37°C are shown in Table 2. Consistent with the findings for the hypothetical drugs (Table 1), regression of $\bar{\nu}$ on D_f results in parameter estimates remarkably close to those obtained by regression of D_f on D_t when weighted 1/y. This

TABLE 2

IBUPROFEN AND NAPROXEN BINDING PARAMETER ESTIMATES – INFLUENCE OF REGRESSION PROCEDURE

Regression scheme (y vs x)		Weighting factor (W_k)	n_1	K_1 ($M^{-1} \times 10^{-6}$)	n_2	K_2 ($M^{-1} \times 10^{-4}$)
Ibuprofen						
D_f	D_t	1	0.63	3.97	6.16	2.30
D_f	D_t	1/y	0.80	2.86	6.27	1.97
D_f	D_t	$1/y^2$	0.90	2.26	6.63	1.63
$\bar{\nu}$	D_f	1	0.80	2.73	6.27	1.95
Naproxen						
D_f	D_t	1	0.73	8.13	4.26	5.73
D_f	D_t	1/y	0.88	4.03	4.14	5.25
D_f	D_t	$1/y^2$	1.31	19.90	3.98	3.54
$\bar{\nu}$	D_f	1	0.94	3.40	4.12	4.98

provides strong supportive evidence for the selection of the weighting scheme 1/y as being the most appropriate and gives credence to our theoretical considerations.

It is clear that the choice of regression procedure and weighting scheme exert a significant influence on the values of the calculated binding parameters. When the major assumptions of non-linear least-squares minimization are not violated, regression of appropriately weighted D_f on D_t should provide unbiased estimates of greatest precision. However, this work suggests that the less rigorous regression of unweighted $\overline{\nu}$ on D_f should provide adequate estimates of binding parameters for most data and certainly better estimates than inappropriately weighted D_f on D_t or D_f on $\overline{\nu}$. Furthermore, use of this simpler procedure obviates the need for subjective judgements in the choice of a weighting scheme.

ACKNOWLEDGEMENTS

The Boots Pure Drug Co., Nottingham, U.K., and Syntex Research, Palo Alto, Calif., for gifts of unlabelled and [^{14}C]labelled samples of ibuprofen and naproxen, respectively.

REFERENCES

Boxenbaum, H.G., Riegelman, S. and Elashoff, R.M., Statistical estimations in pharmacokinetics. J. Pharmacokin. Biopharm., 2 (1974) 123–148.

Chamness, G.C. and McGuire, W.L., Scatchard plots: common errors in correction and interpretation. Steroids, 26 (1975) 538–542.

Crooks, M.J. and Brown, K.F., Binding of sulphonylureas to serum albumin. J. Pharm. Sci., 62 (1973) 1904–1906.

Daniel, C. and Wood, F.S., Fitting Equations to Data, Wiley-Interscience, New York, 1971, pp. 5–10 and pp. 27–28.

Dearden, J.C. and Tomlinson, E., Physico-chemical studies of analgesics. The protein binding of some *p*-substituted acetanilides. J. Pharm. Pharmacol., 22, Suppl. (1970) 53S–59S.

Eichman, M.L., Guttman, D.E., van Winkle, Q. and Guth, E.P., Interactions of xanthine molecules with bovine serum albumin., J. Pharm. Sci., 51 (1962) 66–71.

Hart, H.E., Determination of equilibrium constants and maximum binding capacities in complex in vitro systems: 1. The mammillary system. Bull. Math. Biophys., 27 (1965) 87–98.

Judis, J., Binding of sulphonylureas to serum proteins. J. Pharm. Sci., 61 (1972) 89–93.

Madsen, B.W. and Robertson, J.S., Improved parameter estimates in drug–protein binding studies by non-linear regression. J. Pharm. Pharmacol., 26 (1974) 807–813.

Nilsen, O.G. and Jacobsen, S., The influence of salts and differences in protein isolation procedure on the binding of quinidine to human serum albumin. Biochem. Pharmacol., 25 (1976) 1261–1266.

Perrin, J.H., Vallner, J.J. and Wold, S., An unbiased method for estimating binding parameters in a non-cooperative binding process. Biochim. Biophys. Acta, 371 (1974) 482–490.

Plumbridge, T.W., Aarons, L.J. and Brown, J.R., Problems associated with analysis and interpretation of small molecule/macromolecule binding data. J. Pharm. Pharmacol., 30 (1978) 69–74.

Rosen, A., The measurement of binding constants using circular dichroism – binding of phenylbutazone and oxyphenbutazone. Biochem. Pharmacol., 19 (1970) 2075–2081.

Southworth, R.W., and Deleeuw, S.L., Digital Computation and Numerical Methods, McGraw-Hill, New York, 1965, pp. 183–186.

Steinhardt, J. and Reynolds, J.A., Multiple Equilibria in Proteins, Academic Press, New York., 1969, p. 15.

Veng Pedersen, P., Crooks, M.J. and Brown, K.F., Method of obtaining drug–macromolecule binding parameters directly from dynamic dialysis data. J. Pharm. Sci., 66 (1977a) 1458–1461.

Veng Pedersen, P., Curve fitting and modeling in pharmacokinetics and some practical experiences with NONLIN and a new program FUNFIT. J. Pharmacokin. Biopharm., 5 (1977b) 513–531.

Wagner, J.G., Fundamentals of Clinical Pharmacokinetics, Drug Intelligence Publications, Ill., 1975, pp. 288–290.

Whitlam, J.B., Crooks, M.J., Brown, K.F. and Veng Pedersen, P., Binding of nonsteroidal anti-inflammatory agents to proteins-I. Ibuprofen–serum albumin interaction. Biochem. Pharmacol., 28 (1979) 675–678.