A novel two-site binding equation presented in terms of the total ligand concentration

Zhi-Xin Wang*, Ruo-Fan Jiang

National Laboratory of Biomacromolecules, Institute of Biophysics, Academia Sinica, Beijing 100101, China

Received 28 May 1996; revised version received 15 July 1996

Abstract For the most frequently used two-site model, an exact binding equation is presented in terms of the total ligand concentration. This equation has been extended to analyze the spectroscopic titration experiment where the dilution of protein solution cannot be neglected, the displacement study, and the effect of non-specific binding. Thus, with a non-linear regression program, all unknown binding parameters can be determined correctly by fitting these equations to the experimental data without any data transformation. As an example of the use of the new equations, the experimental data for receptor–insulin binding were taken from literature and reanalyzed by using a non-linear regression data analysis program.

Key words: Ligand–protein interaction; Non-linear regression; Dissociation constant; Binding equation; Displacement curve

1. Introduction

Studies of ligand binding to proteins are of great importance in a large number of scientific disciplines, e.g., endocrinology, immunology, enzymology, neurobiology, and protein physical chemistry. Probably the most commonly used models for analysis of ligand–receptor binding data are the one-site and two-site models. Graphical methods have traditionally been the principle means for estimation of parameters in these models [1–5]. For the one-site binding model, the common feature of these methods is that they rearrange an original non-linear equation into a linear form. The slope and intercept of the modified function are used to calculate the required parameters. However, such graphical methods will only produce correct values for the parameters in the absence of error. Since all the experimental measurements will inevitably be subject to some degree of imprecision, and the equation transformations will result in a non-standard error distribution, use of linearized equations will not in practice give the correct values for the binding parameters [6–8]. Let us consider the Scatchard plot as an example [4]. The objective of this plot is to transform a set of experimental data into a plot of the amount of bound ligand divided by the free ligand (Y axis) as a function of bound ligand (X axis). For a ligand-binding problem with a single class of non-interacting binding sites, this transformation will provide a straight line. The next step is to 'fit' a least-squares straight line to the transformed data points. The slope of this line is related to the ligand-binding affinity and the X-axis intercept is the binding capacity. Fitting a least-squares straight line to the transformed data assumes that the experimental uncertainties follow a random distribution and are parallel to the Y axis [7,8]. In a Scatchard plot, however, the uncertainties are nearly parallel to the Y axis at low fractional saturation and nearly parallel to the X axis at high fractional saturation because the experimental error appears on both the independent and dependent variables [9]. Thus, the use of a least-squares method is not valid for analysis of the Scatchard plot. When a protein molecule contains two classes of binding sites, the Scatchard plot becomes non-linear. Some investigators will draw two straight lines through the limiting slopes of the Scatchard plot and assume that these slopes reflect the binding affinities of the high- and low-affinity classes of sites. As pointed out by Johnson and Frasier [8], however, it is nearly impossible to perform such an operation with reasonable precision. Specially, when the two classes of sites have affinities which are reasonably close, such as within a factor of 10, even in the absence of experimental error the limiting slopes do not correspond to the individual binding affinities.

In the past two decades, improvements in computer software and advances in methods of analysis have paralleled computer hardware developments. Therefore, the best approach to analyze experimental binding data is to perform regression analysis on the original data, without any transformations, by a non-linear least-squares method. The most commonly used and commercially available non-linear least-squares regression program use the method of parameters optimization to achieve the best fit of the model to the experimental data. Analysis of experimental data requires that one assumes a mathematical relationship between the observed quantities, the dependent variables, and the independent variables. This relationship is the fitting function [7]. For the simplest one-site model, the mathematical expression for describing the equilibrium system is a quadratic equation, and the physically meaningful root can be easily identified [10–12]. For the two-site model, however, combining all partial equilibria and mass conservation, one obtains a cubic equation. Roots of polynomials degree $n > 2$ are usually extracted by methods of numerical analysis [13]. Since a polynomial of n degree has n roots, in the case of numerical methods one would be faced with the possible existence of multiple real roots [14].

In the present paper, the properties of the roots of the general cubic equation have been analyzed in detail. For the most frequently used two-site model, a unique physically meaningful root of the corresponding cubic equation has been identified unambiguously. Thus, this algebraically explicit equation can be put into a commercially available non-linear regression computer program, all unknown parameters can be determined correctly by fitting this equation to the experimental data without any data transformation.

*Corresponding author. Tel.: (86) 10-62020077, ext. 462; Fax: (86) 10-62022026

PII S0014-5793(96)00818-6

2. Theory

2.1. Exact analytical expression for describing two-site binding model

When a protein molecule has two classes of non-interacting binding sites, the concentration of bound ligand molecules, $[L]_b$, is given by

$$[L]_b = \frac{[R_1]_0[L]}{K_1 + [L]} + \frac{[R_2]_0[L]}{K_2 + [L]} \quad (1)$$

where $[L]$ is the free ligand concentration, K_1, $[R_1]_0$ and K_2, $[R_2]_0$ are the dissociation constants and the binding capacity for classes 1 and 2, respectively [15–17].

Substitution of the mass conservation equation, $[L]_b = [L]_0 - [L]$, into Eq. 1 and rearrangement yields

$$[L]^3 + a[L]^2 + b[L] + c = 0 \quad (2)$$

where

$$a = K_1 + K_2 + [R_1]_0 + [R_2]_0 - [L]_0$$

$$b = K_1K_2 + K_2[R_1]_0 + K_1[R_2]_0 - (K_1 + K_2)[L]_0$$

$$c = -K_1K_2[L]_0$$

Let $[L] = u - (a/3)$, we get the transformed equation

$$u^3 - \left(\frac{a^2}{3} - b\right)u + \left(\frac{2}{27}a^3 - \frac{1}{3}ab + c\right) = 0 \quad (3)$$

Since $\Delta < 0$, the three real roots of Eq. 3 are given by [18]

$$u_1 = \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{\theta}{3} \quad (4)$$

$$u_2 = \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{2\pi - \theta}{3} \quad (5)$$

$$u_3 = \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{2\pi + \theta}{3} \quad (6)$$

where

$$\theta = \arccos\frac{-2a^3 + 9ab - 27c}{2\sqrt{(a^2 - 3b)^3}} \quad (0 < \theta < \pi)$$

According to the definition of u and the physical conditions of the problem proposed, it can be verified that u_1 expresses the unique physically meaningful root of Eq. 3, and u_2 and u_3 have no reference to the problem proposed (see Appendix). Thus, the physically meaningful root of Eq. 2 can be written as

$$[L] = -\frac{a}{3} + \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{\theta}{3} \quad (7)$$

and the expression for the concentration of bound ligand molecules, $[L]_b$, is given by

$$[L]_b = [L]_0 - [L] = [L]_0 + \frac{a}{3} - \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{\theta}{3} \quad (8)$$

To test the validation of the new method, the theoretical data (○, Fig. 1) were artificially generated on the computer with Eq. 1, and then Eq. 8 was fitted to the data by using a commercially available non-linear regression program, SigmaPlot V.2.00 (——, Fig. 1). It can be seen from Fig. 1 that with the explicit analytical expression, the non-linear least-squares curve fitting of the binding parameters can give exact values of the parameters for the theoretical data (the standard errors for the all parameters equal to zero).

For the spectroscopic titration experiments, assuming that the change in spectroscopic signal, ΔF, is directly proportional to the concentration of bound ligand, $[L]_b$, the signal for protein–ligand interaction can then be written as

$$\Delta F = C[L]_b = C\left\{[L]_0 + \frac{a}{3} - \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{\theta}{3}\right\} \quad (9)$$

The constant C is a measure of the change in signal (ΔF) changes per unit complex concentration (molar signal coefficient). In the absorption spectroscopy, C is a molar difference absorption coefficient $\Delta\varepsilon$.

In the conventional spectroscopic titration experiments, the stock ligand concentration should ideally be about 100 to 500-fold that of the final ligand concentration to be tested. This will minimize dilution of the protein solution, which should not be diluted by more than 5%. In practice, however, due to some reasons, such as poor solubility of ligand and so on, it is often difficult to prepare highly concentrated ligand stock solution. In these situations, the conventional spectroscopic method cannot be applied to determine the dissociation constants of protein–ligand complexes. Some years ago, a novel spectroscopic titration method was developed to solve this problem [12]. With the analytical expression given above, this method can be easily applied to the case of the two-site model.

If the initial volume and concentrations of the binding site in solution are V', $[R_1]'$, and $[R_2]'$, respectively, and the ligand concentration in stock solution is $[L]'$, then the total concentrations of binding sites ($[R_1]_0$, $[R_2]_0$) and ligand ($[L]_0$) can be obtained by accounting for the total volume of the aliquot (V_c) added during titration experiment:

$$[R_1]_0 = \frac{[R_1]'V'}{V' + V_c}, [R_2]_0 = \frac{[R_2]'V'}{V' + V_c}, [L]_0 = \frac{[L]'V_c}{V' + V_c} \quad (10)$$

By substituting Eq. 10 into Eq. 9, one can obtain a explicit analytical relationship between ΔF and V_c.

2.2. Analytical expression for describing displacement analysis curve

There is an alternative method for the study of receptor–ligand binding. It involves incubation of a fixed amount of radio-labeled ligand and various concentrations of unlabeled ligand with the receptor, and measurement of bound concentration of labeled ligand as the function of increasing concentrations of unlabeled ligand. A major problem generally encountered during receptor–ligand binding studies is the presence of non-specific binding, defined as a low-affinity ligand binding to non-receptor domains that does not show any saturating behavior within the range of ligand concentration used [8,19]. As a consequence of such non-specific interaction the binding of radio-labeled ligand cannot fully be displaced by a homologous unlabeled ligand. Assuming that the radio-labeled and unlabeled ligands have the same binding affinities for the two classes of sites, the binding equation for describing

this equilibrium system can then be written as

$$[L]_b = \frac{[R_1]_0[L]}{K_1 + [L] + [L^*]} + \frac{[R_2]_0[L]}{K_2 + [L] + [L^*]} + N[L] \quad (11)$$

$$[L^*]_b = \frac{[R_1]_0[L^*]}{K_1 + [L] + [L^*]} + \frac{[R_2]_0[L^*]}{K_2 + [L] + [L^*]} + N[L^*] \quad (12)$$

in which the experimentally obtained value for the total amount of labeled bound ligand $[L^*]_b$ is a component and a non-specific component characterized by the constant N.

From Eq. 11 and 12, we have

$$\frac{[L]_b}{[L^*]_b} = \frac{[L]_0 - [L]}{[L^*]_0 - [L^*]} = \frac{[L]}{[L^*]} \quad (13)$$

and

$$\frac{[L^*]_0}{[L^*]} - 1 = \frac{[L]_0}{[L]} - 1 \text{ or } \frac{[L^*]}{[L^*]_0} = \frac{[L]}{[L]_0} \quad (14)$$

Let $[L_T] = [L] + [L^*]$ and $[L_T]_0 = [L]_0 + [L^*]_0$, one can obtain

$$[L^*] = \frac{[L^*]_0}{[L]_0 + [L^*]_0}[L_T] \quad (15)$$

From Eqs. 11 and 12, we have

$$[L_T]^3 + a[L_T]^2 + b[L_T] + c = 0 \quad (16)$$

where

$$a = \{(K_1 + K_2)(1 + N) + [R_1]_0 + [R_2]_0 - [L_T]_0\}/(1 + N)$$

$$b = \{(K_1K_2(1 + N) + K_2[R_1]_0 + K_1[R_2]_0 - (K_1 + K_2)[L_T]_0\}/(1 + N)$$

$$c = -K_1K_2[L_T]_0/(1 + N)$$

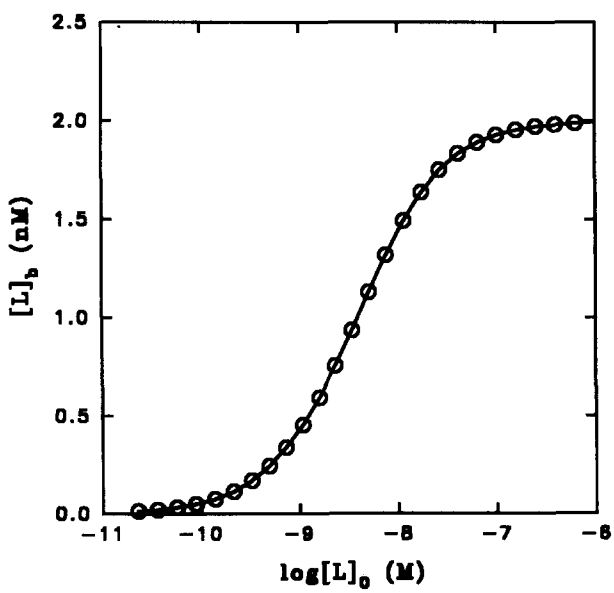

Fig. 1. A plot of a typical experiment for ligand binding to a protein with two classes of independent binding sites. The theoretical data were generated according to Eq. 1 with parameters: $[R_1]_0 = 0.5 \times 10^{-9}$ M, $[R_2]_0 = 1.5 \times 10^{-9}$ M, $K_1 = 5.0 \times 10^{-10}$ M, $K_2 = 4.762 \times 10^{-9}$ M (○). Solid line (——) represents the best-fitting curve. Exact values of parameters were obtained by fitting Eq. 8 to the theoretical data using the non-linear regression program, SigmaPlot V.2.00.

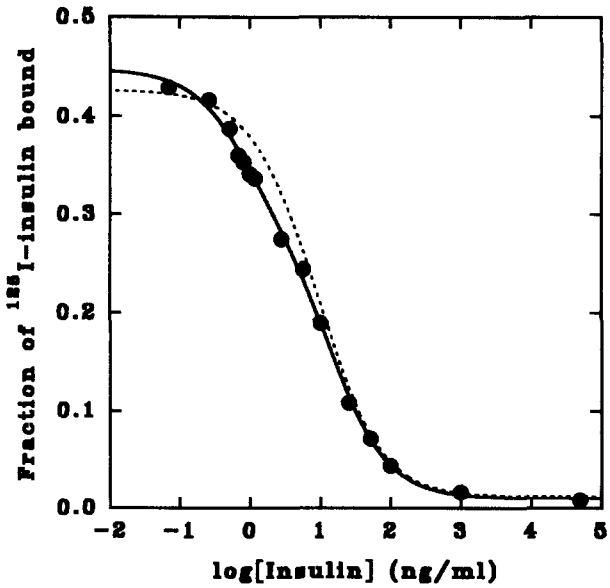

Fig. 2. Percent of total ^{125}I-insulin bound to liver membranes as a function of total unlabeled insulin concentration. Data were taken from Fig. 6 at 4°C of Kahn et al. [20]. The solid line (——) represents the best fit of the data using Eq. 18. The optimized values of K_1, K_2, $[R_1]_0$, $[R_2]_0$, and N were 0.066 ± 0.021 *nM*, 1.65 ± 0.33 *nM*, 0.092 ± 0.032 pmol/mg of membrane protein, 1.90 ± 0.20 pmol/mg of membrane protein, and 0.0102 ± 0.0046, respectively. The dotted line (·····) was generated with the parameters obtained from the reference [20].

Eq. 16 is identical in form with Eq. 2. Therefore, we have

$$[L_T] = -\frac{a}{3} + \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{\theta}{3} \quad (17)$$

The concentration of the bound radioactive ligand is given by

$$[L^*]_b = [L^*]_0 - \frac{[L^*]_0}{[L]_0 + [L^*]_0}\left\{\frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{\theta}{3} - \frac{a}{3}\right\} \quad (18)$$

Fig. 2 shows a example of analyzing a receptor–insulin binding experiment by the new method. Data were taken from Kahn et al. [20] and Eq. 18 was fitted to the data. The solid line was best-fitting curve obtained from the new method, and the optimized values of K_1, K_2, $[R_1]_0$, $[R_2]_0$, and N were 0.066 ± 0.021 nM, 1.65 ± 0.33 nM, 0.092 ± 0.032 pmol/mg of membrane protein, 1.90 ± 0.20 pmol/mg of membrane protein, and 0.0102 ± 0.0046, respectively. For comparison, the dotted line was generated with the parameters given by reference [20]. It can be seen that the result obtained from the non-linear regression method is much better than that obtained by traditional plotting method.

3. Discussion

As mentioned earlier, analysis of experimental data by the commonly used non-linear regression program requires an explicit fitting function. Most of the mathematical formulations for describing multiple equilibrium binding systems are presented in terms of the free ligand concentrations rather than total concentrations [21–23]. The use of these formulae are, therefore, restricted to cases in which either the bound concentrations of ligand are much smaller than the total con-

centrations so that the concentrations of free ligand can be treated to be equal to their total concentrations, or the free ligand concentrations can be determined accurately from experiments. Often, these conditions are not met. In a receptor–ligand binding experiment, the free ligand concentration is usually calculated as the total ligand concentration minus the bound concentration. When the concentrations of all the species in system are comparable with each other and relatively high with respect to the dissociation constants of corresponding complexes, bound concentration is a significant fraction of the total ligand concentration, and any small experimental error in bound concentration will generate an error in free concentration. In this case, since both the bound and free concentrations have the same associated errors, the error bar at the region of low ligand concentration will no longer be vertical. This violates a fundamental mathematical principle that imposes stochastic independence between the variables in order for any regression analysis to be applicable [6–8]. This problem can be approached in essentially two ways. (1) The experimental data can be fitted to an explicit mathematical expression presented in terms of the total ligand concentration [10–12]. (2) Numerical methods can be used to compute the free ligand concentrations from the total concentrations, and the bound ligand concentrations in turn can be calculated from explicit expressions [8,19]. When a protein molecule has two or more classes of independent sites, combining all partial equilibria and mass conservation, one obtains a polynomial of higher degree. The formulae for the roots of cubic and quartic polynomials are cumbersome, and there is no general expression for the roots of quintic and higher degree polynomials. Therefore, roots of polynomials degree $n > 2$ are usually extracted by methods of numerical analysis [13,14,18]. Munson and Rodbard have developed a numerical method for analysis of data from binding experiments [19]. However, since a polynomial of n degree has n roots, with an unfortunate starting value, the procedure will converge to the wrong root or perhaps enter a non-convergent cycle [24].

In the present paper, the unique physically meaningful root for the most frequently used two-site model has been identified unambiguously. Thus, with this exact analytical expression and a commercially available non-linear regression program, all unknown parameters can be correctly determined from the experimental data. Comparing to the numerical method developed by Munson and Rodbard, another major advantage of using the analytical expression is that it can be easily extended to analyze the results of the spectroscopic titration experiments where the dilution of protein solution cannot be neglected during the titration course. This is particularly useful if the ligand solubility becomes a problem, and the conventional spectroscopic titration method is not applicable to determine the binding constants of protein–ligand complexes [25].

Acknowledgements: The authors would like to acknowledge the help of Dr. Y. Tan for the data analysis by using the computer program, SigmaPlot V.2.00. This work was supported in part by Grant 39421003 from the China Science Foundation and the Pandeng Project of the National Commission of Science and Technology.

Appendix

In order to identify the physically meaningful root of Eq. 3, let us discuss some properties of angle θ first. If the angle θ be between 0 and π/2 or if its cosine be positive, then $(2\pi-\theta)/3$ and $(2\pi+\theta)/3$ are both greater than π/2 and their cosines are negative; but if θ be between π/2 and π or if its cosine be negative, then θ/3 and $(2\pi-\theta)/3$ are less than π/2 and $(2\pi+\theta)/3$ greater than π/2, that is

$$\begin{array}{ll} 0<\theta/3<\pi/3, & 0.5<\cos(\theta)/3<1 \\ \pi/3<(2\pi-\theta)/3<2\pi/3, & -0.5<\cos(2\pi-\theta)/3<0.5 \\ 2\pi/3<(2\pi+\theta)/3<\pi, & -1<\cos(2\pi+\theta)/3<-0.5 \end{array}$$

It can be seen from the discussion given above that $u_1 > u_2 > u_3$, and u_1 is always a positive root and u_3 a negative root of Eq. 3, whatever the value of θ may be. Therefore, the three roots of Eq. 2 are given by $[L]_i = u_i - (a/3)$, $i = 1, 2, 3$.

According to the definition of u and the physical conditions of the problem proposed, one can obtain

$$\begin{aligned} u &= (a/3) + [L] = \{3[L] + K_1 + K_2 + [R_1]_0 + [R_2]_0 - [L]_0\}/3 \\ &= \{2[L] + K_1 + K_2 + [R_1]_0 + [R_2]_0 - [L]_0\}/3 > 0 \end{aligned}$$

On the other hand, since $[L]$ is the concentration of free ligand, we have

$$[L] = u - (a/3) > 0 \text{ or } u > (a/3)$$

Therefore, a physically meaningful root of Eq. 3, u, must satisfy the following condition:

$$u > \max\{a/3, 0\} \qquad (A1)$$

That is, $u > (a/3)$ if $a > 0$, and $u > 0$ if $a < 0$. Since u_3 is always a negative root, it can be excluded first. For u_2, according to the relationships between the roots and coefficients, we have

$$\begin{aligned} [L]_1[L]_2[L]_3 &= \{u_1 - (a/3)\}\{u_2 - (a/3)\}\{u_3 - (a/3)\} \\ &= K_1 K_2 [L]_0 > 0 \end{aligned} \qquad (A2)$$

When $a > 0$, there must be two negative roots and one positive root since $\{u_3 - (a/3)\}$ is always negative. As we have seen, $u_1 > u_2$, it must follow therefore that $u_1 - (a/3) > 0$ and $u_2 - (a/3) < 0$, i.e., $u_2 < a/3$. On the other hand, when $a < 0$, $-2a^3 + 9ab - 27c > 0$, $\cos\theta > 0$. It can be seen from the discussion given above that in this case we have $u_2 < 0$ because $(2\pi-\theta)/3$ is greater than π/2. Thus, according to inequality (A1), u_2 should also be excluded and u_1 expresses the unique proper root of Eq. 3. u_2 and u_3 have no reference to the problem proposed. Therefore, the unique rational root of Eq. 2 is given by

$$[L] = -\frac{a}{3} + \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{\theta}{3} \qquad (A3)$$

References

[1] Lineweaver, H. and Burk, D. (1934) J. Am. Chem. Soc. 6, 658–666.
[2] Eadie, G.S. (1934) J. Biol. Chem. 146, 85–93.
[3] Hofstee, B.H.J. (1952) Science 116, 329–331.
[4] Scatchard, G. (1949) Ann. NY Acad. Sci. 51, 660–672.
[5] Stinson, R.A. and Holbrook, J.J. (1973) Biochem. J. 131, 719–728.
[6] Bevington, P. (1969) Data Reduction and Error Analysis for the Physical Science, McGraw-Hill, New York.
[7] Johnson, M.L. (1992) Anal. Biochem. 206, 215–225.
[8] Johnson, M.L. and Frasier, S.G. (1985) Methods in Enzymology, Vol. 117, pp. 301–342.
[9] Munson, P.J. and Rodbard, D. (1983) Science 220, 979–981.

[10] Green, N.M. (1965) Biochem. J. 94, 23–24.
[11] Inglese, J., Blatchly, R.A. and Benkovic, S.J. (1990) J. Med. Chem. 32, 937–940.
[12] Wang, Z.X., Kumar, N.R. and Srivastava, D.K. (1992) Anal. Biochem. 206, 376–381.
[13] Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1988) Numerical Recipes in C, Oxford Univ. Press, Cambridge.
[14] Arehbold, J.W. (1970) Algebra, 4th edn., pp. 174–194, Pitman, London.
[15] Fersht, A.R. (1985) Enzyme Structure and Mechanism, 2nd edn., pp. 185–191, W.H. Freeman, New York.
[16] Hammes, G.G. (1982) Enzyme Catalysis and Regulation, pp. 161–164, Academic Press, New York.
[17] Wang, Z.X. (1990) J. Theor. Biol. 143, 445–453.
[18] Burington, R.S. (1973) Handbook of Mathematical Tables and Formulas, 5th edn., pp. 12–13, McGraw-Hill, New York.
[19] Munson, P.J. and Rodbard, D. (1980) Anal. Biochem. 107, 220–239.
[20] Kahn, C.R., Freychet, P. and Roth, J. (1974) J. Biol. Chem. 249, 2249–2257.
[21] Adair, G.S. (1925) J. Biol. Chem. 63, 529–545.
[22] Tanford, C. (1961) Physical Chemistry of Macromolecule, Ch. 8, John Wiley, New York.
[23] Cantor, C.R. and Schimmel, P.R. (1980) Biophysical Chemistry, Part III, Freeman, San Francisco, CA.
[24] Wells, J.W. (1992) (in: Hukme, E.C., ed.), Receptor-Ligand Interactions, pp. 289–323, Oxford University Press, Oxford.
[25] Wang, Z.X., Killilea, S.D. and Srivastava, D.K. (1993) Biochemistry 32, 1500–1509.