

# REGRESSION DIAGNOSTICS IN THE $pK_a$ ESTIMATION BY THE NONLINEAR REGRESSION OF MULTIWAVELENGTH SPECTROPHOTOMETRIC DATA

M. Meloun<sup>1</sup>, Z. Ferenčíková<sup>1</sup> and A. Vrána<sup>2</sup>

<sup>1</sup>Department of Analytical Chemistry, University of Pardubice, CZ 532 10 Pardubice, Czech Republic,

<sup>2</sup>IVAX Pharmaceuticals, s.r.o., CZ 747 70 Opava, Czech Republic,

[milan.meloun@upce.cz](mailto:milan.meloun@upce.cz),

Recent development in the computational diagnostic tools for the  $pK_a$  estimation of druglike molecules carried out by the nonlinear regression of multiwavelength spectrophotometric pH-titration data are demonstrated. Most indices-based methods in the factor analysis of spectra predict the correct number of components even in the presence of a minor one or in cases of overlapping equilibria when the *signal-to-error ratio* *SER* is higher than 10. The Wernimont-Kankare procedure in the INDICES program performs a reliable determination of the instrumental standard deviation of the spectrophotometer used  $s_{inst}(A)$  and correctly predicts the number of light-absorbing components present. It can also solve an ill-defined problem with severe colinearity in the data set, with very small changes in spectra and overlapping equilibria. The mixed dissociation constants of an example, the drug silybin, at ionic strength  $I = 0.03$  and a temperature of 25°C were determined using two different programs, SPECFIT32 and SQUAD(84). A proposed experimental and computational strategy with the use of regression diagnostics for the determination of the dissociation constants is presented with results, *i.e.*  $pK_{a1} = 6.898(0.022)$  and  $6.897(0.002)$ ,  $pK_{a2} = 8.666(0.021)$  and  $8.667(0.012)$ ,  $pK_{a3} = 9.611(0.010)$  and  $9.611(0.004)$ ,  $pK_{a4} = 11.501(0.008)$  and  $11.501(0.007)$ . While great progress has been achieved in terms of the reliability of the protonation model estimation, among the most efficient diagnostics of the nonlinear regression of multiwavelength pH-spectra are the goodness-of-fit test, Cattel's scree plot of the factor analysis, spectra deconvolution, the signal-to-error *SER* ratio analysis and other tools of efficient spectra analysis. The second example, the mixed dissociation constants of *methotrexate* at various ionic strengths  $I$  of range 0.01 - 0.4, and at temperatures of 25°C and 37°C, were determined with the use of two different multiwavelength and multivariate treatments of spectral data resulting to three thermodynamic dissociation constants for methotrexate  $pK_{a1}^T = 2.895(13)$ ,  $pK_{a2}^T = 4.410(14)$ ,  $pK_{a3}^T = 5.726(15)$  at 25°C and  $pK_{a1}^T = 3.089(15)$ ,  $pK_{a2}^T = 4.392(12)$ ,  $pK_{a3}^T = 5.585(11)$  at 37°C, where the figure in brackets is the standard deviation in last significant digits.

## Reference:

1. M. Meloun, Z. Ferenčíková, A. Vrána, The thermodynamic dissociation constant of capecitabine by the regression analysis of spectrophotometric and potentiometric data, *J. Chem. Thermodynamics* **43**, 930 – 937 (2011).
1. M. Meloun, Z. Ferenčíková, A. Vrána, The thermodynamic dissociation constant of methotrexate by the regression analysis of spectrophotometric and potentiometric data, *Cent. J. Eur. Chem.* **8**, 494 - 507 (2010).

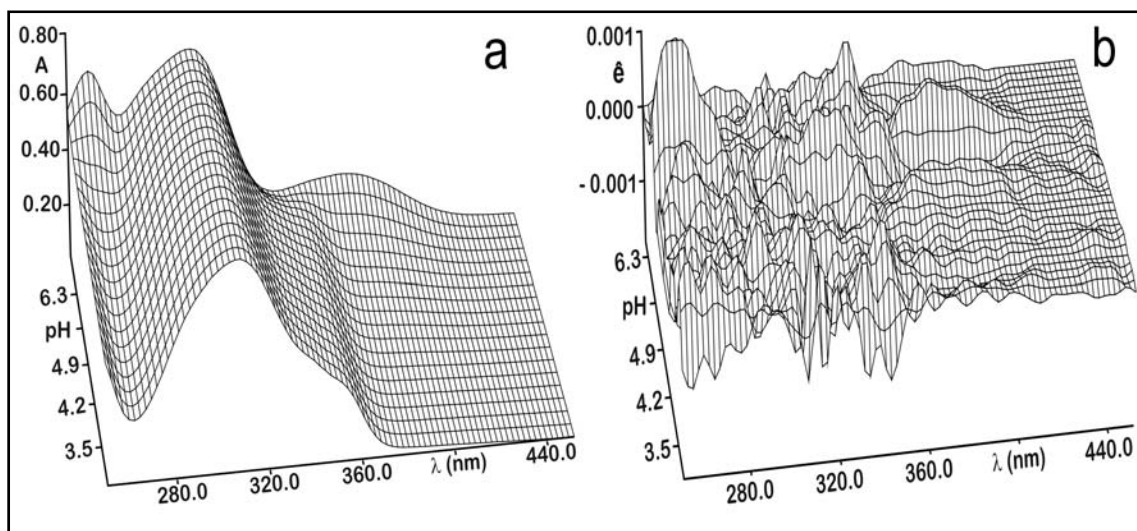


Figure 1 (a) The 3D-absorbance-response-surface representing the measured multiwavelength absorption spectra for methotrexate in dependence on pH at 25°C, (b) the 3D-residuals map after non-linear regression performed with SPECFIT and SQUAD, (S-Plus).

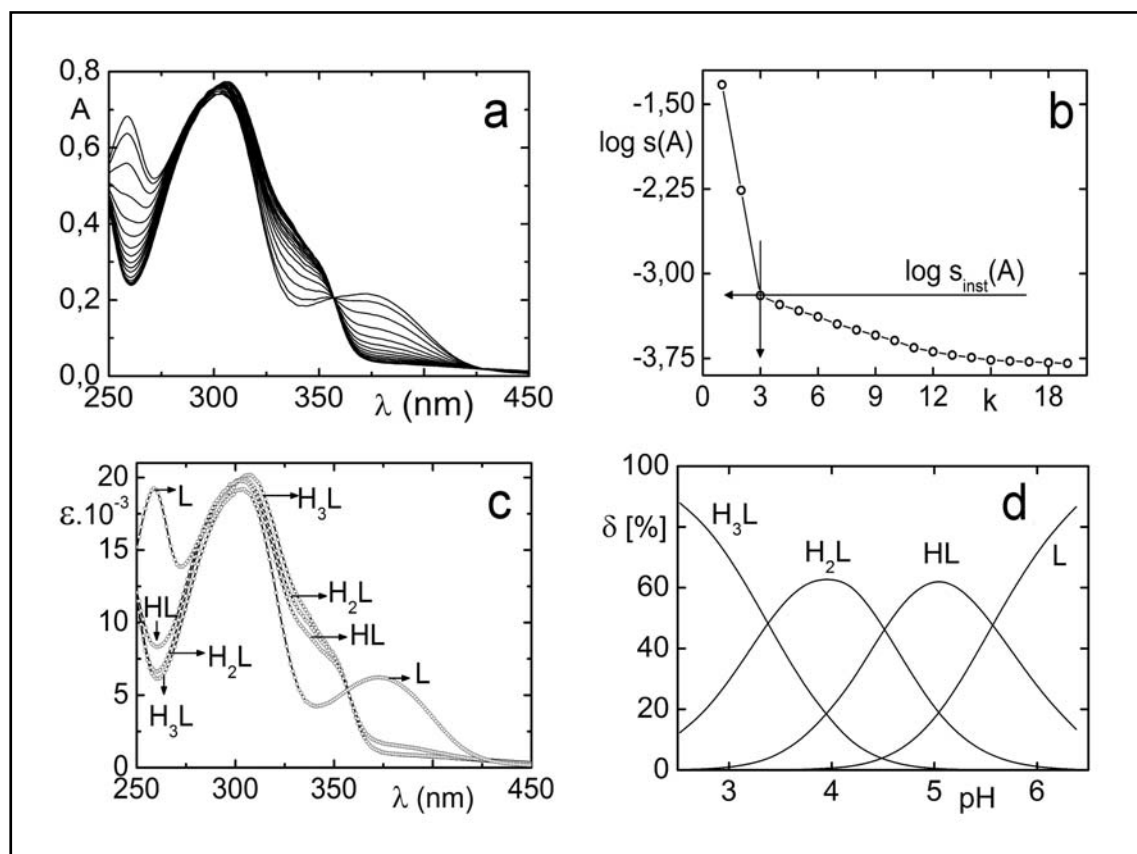


Figure 2 The non-linear regression analysis of the protonation equilibria model and factor analysis of methotrexate: (a) Absorption spectra of  $1.2 \times 10^{-5}$  M MTX in dependence on pH at 25°C; (b) Cattell's scree plot of the Wernimont-Kankare procedure for the determination of the number of light-absorbing species in the mixture  $k^* = 3$  leads to  $n_c = 3$  and the actual instrumental error of the spectrophotometer used  $s_{inst}(A) = 0.65$  mAU (INDICES in S-Plus); (c) pure spectra profiles of molar absorptivities vs. wavelengths for the variously protonated species L, HL, H<sub>2</sub>L and H<sub>3</sub>L; (d) distribution diagram of the relative concentrations of both variously protonated species L, HL, H<sub>2</sub>L and H<sub>3</sub>L of rasagiline base in dependence on pH at 25°C. The charges of species are omitted for the sake of simplicity (SPECFIT, ORIGIN).

The financial support of the Grant Agency IGA MZ ČR (No NS9831-4/2008) and of the Czech Ministry of Education (Grant No MSM0021627502) is gratefully acknowledged.