

# Overlapping $pK_a$ of the Multiprotic Hemostyptic Eltrombopag using UV–Vis Multiwavelength Spectroscopy and Potentiometry

Milan Meloun<sup>1</sup>  · Lucie Pilařová<sup>1</sup> · Tomáš Pekárek<sup>2</sup> · Milan Javůrek<sup>3</sup>

Received: 25 April 2017 / Accepted: 17 August 2017 / Published online: 11 October 2017  
© Springer Science+Business Media, LLC 2017

**Abstract** Dissociation constants of the multiprotic hemostyptic Eltrombopag determined by pH-potentiometric and WApH-spectrophotometric titrations are compared. Hemostyptic and hemostatic Eltrombopag treats low blood platelet counts in adults with chronic immune idiopathic thrombocytopenia ITP. Eltrombopag has five protonatable sites in a pH range of 2–10, where only two  $pK$  are well separated ( $\Delta pK > 3$ ), while the other three are near dissociation constants of overlapping equilibria. According to the MARVIN prediction, in neutral medium Eltrombopag occurs in the slightly water soluble form  $LH_3$  that can be protonated to the soluble species  $LH_4^+$  and  $LH_5^{2+}$ . The molecule  $LH_3$  can be dissociated to still difficultly soluble species  $LH_2^-$ ,  $LH^{2-}$  and  $L^{3-}$ . Due to the limited solubility of Eltrombopag above pH = 9.5 the protonation was studied up to pH = 10. Five dissociation constants can be reliably determined with REACTLAB and SQUAD84 leading to the same values. From a dependence on ionic strength, the thermodynamic dissociation constants were estimated at 25 °C:  $pK_{a1}^T = 2.69$ ,  $pK_{a2}^T = 6.97$ ,  $pK_{a3}^T = 7.13$ ,  $pK_{a4}^T = 7.65$ ,  $pK_{a5}^T = 8.30$ . At pH values above 10 and pH below 5 a very fine precipitate of Eltrombopag, forming a slight opalescence, was observed; thus measurements of the potentiometric titration curve above pH = 9 and pH below 5 were excluded from the regression analysis to estimate  $pK_{a2} = 6.59(01)$ ,  $pK_{a3} = 7.56(04)$ ,  $pK_{a4} = 8.48(59)$  and  $pK_{a5} = 9.29(34)$  at 25 °C with ESAB.

---

✉ Milan Meloun  
milan.meloun@upce.cz

Lucie Pilařová  
lucie.pilarova91@centrum.cz

Tomáš Pekárek  
tomas.pekarek@zentiva.cz

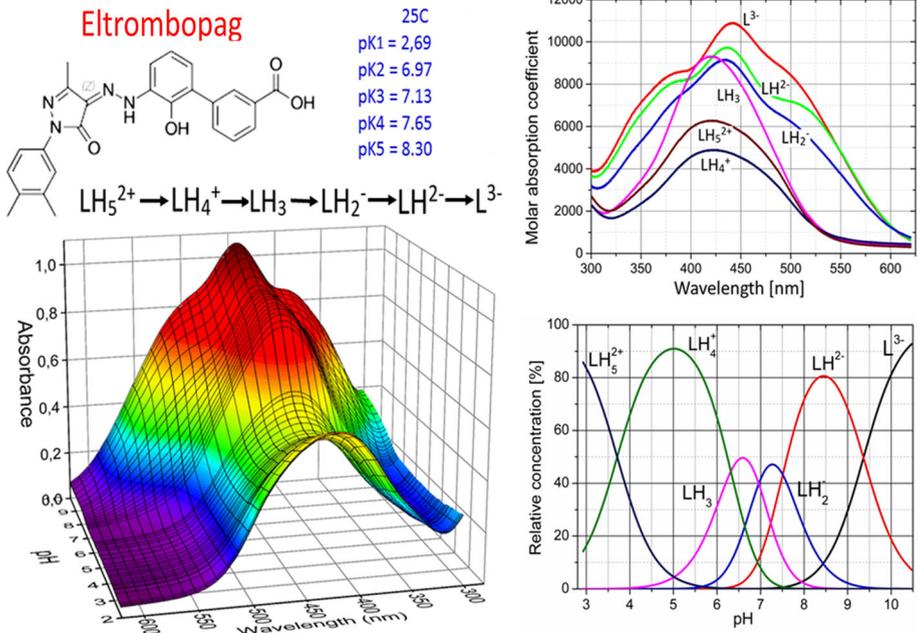
Milan Javůrek  
milan.javurek@upce.cz

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

<sup>2</sup> Zentiva k.s, U kabelovny 130, 102 37 Prague, Czech Republic

<sup>3</sup> Department of Process Control, University of Pardubice, 532 10 Pardubice, Czech Republic

Graphical Abstract



collaboration between GlaxoSmithKline and Ligand Pharmaceuticals [1]. Eltrombopag (code named SB-497115-GR, CAS number 496775-61-2, 496775-62-3, ATC code B02BX05, PubChem CID 9846180, ChemSpider 21106301) with the IUPAC name 3'-[(2Z)-2-[1-(3,4-dimethylphenyl)-3-methyl-5-oxo-1,5-dihydro-4H-pyrazol-4-ylidene]hydrazino]-2'-hydroxy-3-biphenylcarboxylic acid, is used to treat low blood platelet counts in adults with chronic immune idiopathic thrombocytopenia ITP, when other medicines, or spleen removal surgery, have not worked well enough. It works by causing the cells in the bone marrow to produce more platelets, [2]. Eltrombopag also increases the number of platelets in people who have hepatitis C, a viral infection [3] that may damage the liver so that they can begin and continue treatment with interferon (Peginterferon, Pegintron, others) and ribavirin (Rebetol).

Knowledge of the possible ionization states of a pharmaceutical substance, embodied in the logarithm of the mixed acid dissociation constant  $pK_a$ , is vital for understanding many properties essential to drug development [4]. As the majority of drugs are weak acids and/or bases, knowledge of the dissociation constant in each case helps in understanding the ionic form a molecule will take across a range of pH values and the level of general interest in such ionization phenomena is evident from the large number of recent publications on the topic [5–11].  $pK_a$  values can be either experimentally measured or theoretically predicted:

1. Many new substances are poorly soluble in aqueous solutions; conventional potentiometric determination of dissociation constants of these compounds can often be difficult [12]. Spectrophotometric  $pK_a$  determination is an alternative method to potentiometry provided that the compound is water soluble to the extent of  $10^{-6}$  mol·dm<sup>-3</sup> or more and provided the compound possesses pH-dependent light absorption due to the presence of a chromophore in proximity to the ionization center (cf. Ref. [13–17]). In previous work [18–27] the authors have shown that the multiwavelength spectrophotometric pH-titration method in combination with suitable chemometric tools (called as the WApH-method by Tam [28]) can be used for the determination of dissociation constants  $pK_a$  even for barely soluble drugs. Spectra are generally a superposition of spectra of the numerous compounds present. In many cases, however, the spectral responses of two and sometimes even more components overlap considerably and the analysis is no longer straightforward. Hard modelling methods, e.g., SQUAD84 [16, 29, 30], include traditional least-squares curve fitting approaches, based on a previous postulation of a chemical model, i.e., a set of species defined by their stoichiometric coefficients and formation constants, which are then refined by the least-squares minimization. The most relevant algorithms are SQUAD84 and REACTLAB [31] or its previous version SPECFIT32 [32]. Soft modelling techniques, as for example REACTLAB, such as multivariate curve resolution methods based on factor analysis, work without any assumption of a chemical model, and do not need to comply with the mass-action law. The molar absorptivities are usually not required for analysis. However, explicit equations for the equilibrium expression are necessary to rotate the eigenvectors to give the correct concentration profiles. It may be difficult to generalize these explicit equations for multistep ionization system [28].
2. Nine commercially available or free programs for predicting ionization constants were compared [4]. Meloun et al. [27] used the REGDIA regression diagnostics algorithm written in S-Plus [33] critically examine an accuracy of  $pK_a$  predictions with four programs ACD/pK [3, 6, 34, 35], Marvin Sketch [36, 37], PALLAS, and SPARC

[38, 39] and the best were considered ACD/Labs™ [40] and pK<sub>a</sub> Predictor 3.0 [41]. Balogh et al. [42] found the most predictive and reliable predictors to be MARVIN-Sketch and ACD/Percepta [34, 43, 44].

There are no systematic reports about the drug Eltrombopag concerning its UV–VIS spectra, pK<sub>a</sub> values and the distribution diagram of variously protonated species. The aim of our study was to examine and verify the UV-metric pK<sub>a</sub> determination (WApH-spectrophotometric titration) of the pH-absorbance matrix and to carry out the pH-metric pK<sub>a</sub> determination (pH-potentiometric titration) of the protonation model to find suitable conditions for a reliable regression determination of dissociation constants. Considering the role of pK<sub>a</sub> predictions in early phase discovery, we concluded that the selection of appropriate prediction tools for regular pharmaceutical chemistry use requires solid benchmarking studies. We are reporting our obtained results that were evaluated by two different LFER based pK<sub>a</sub> predictions tools, MARVIN-Sketch and ACD/pK software [42].

## 2 Theoretical

### 2.1 UV-metric pK<sub>a</sub> Determination (WApH-Spectrophotometric Titration)

The acid–base equilibrium of Eltrombopag studied is described in terms of the protonation of the Brönsted base L<sup>z-1</sup> according to the equation L<sup>z-1</sup> + H<sup>+</sup> ⇌ HL<sup>z</sup> [21]. The protonation equilibria between the species L (the charges are omitted for the sake of simplicity) of a drug and a proton H are considered to form a set of variously protonated species L, LH, LH<sub>2</sub>, LH<sub>3</sub>...etc., with the general formula LH<sub>r</sub> in a particular chemical model and which are represented by n<sub>c</sub> (the number of species), r<sub>i</sub>, i = 1, ..., n<sub>c</sub> where index i indicates their particular stoichiometry; the overall protonation (stability) constant of the protonated species, β<sub>r</sub>, may then be expressed as

$$\beta_r = \frac{[\text{LH}_r]}{[\text{L}][\text{H}]^r} = \frac{c}{lh^r} \quad (1)$$

where the free concentration [L] = l, [H] = h and [LH<sub>r</sub>] = c, [21]. For dissociation at constant ionic strength the “mixed dissociation constants” are defined as

$$K_{a,j} = \frac{[\text{LH}_{j-1}]a_{\text{H}^+}}{[\text{LH}_j]} \quad (2)$$

As each species is characterized by its own spectrum, for UV–Vis experiments and the i<sup>th</sup> solution measured at the j<sup>th</sup> wavelength, the Lambert–Beer law relates the absorbance, A<sub>i,j</sub>, defined as:

$$A_{i,j} = \sum_{n=1}^{n_c} \varepsilon_{j,n} c_n = \sum_{n=1}^{n_c} (\varepsilon_{r,j} \beta_r l h^r)_n \quad (3)$$

where ε<sub>r,j</sub> is the molar absorption coefficient of the LH<sub>r</sub> species with the stoichiometric coefficient r measured at the j<sup>th</sup> wavelength. The absorbance A<sub>i,j</sub> is an element of the absorbance matrix A of size (n<sub>s</sub> × n<sub>w</sub>) being measured for n<sub>s</sub> solutions with known total concentrations of n<sub>z</sub> = 2 basic components, c<sub>L</sub> and c<sub>H</sub>, at n<sub>w</sub> wavelengths [19]. The general procedure used to build the protonation model with SPECFIT32, REACTLAB or SQUAD84 was described previously [13, 21, 25, 45]. Determining the chemical model of

the drug protonation equilibria by a regression analysis of potentiometric titration data or by spectra seems to be dependent on user experience and the software used. A significant role is played by resolution hypotheses of the proposed regression model and distinguishability of the spectra of differently protonated chromophores in the molecule. Two different programs for the numerical analysis of spectra were used, REACTLAB and SQUAD84. These programs found the consensus in numerical parametric estimates and in a fitness of the predicted absorbance spectra through measured absorbance data.

## 2.2 pH-Metric $pK_a$ Determination (pH-Potentiometric Titration)

The overall protonation constant of the protonated species,  $\beta_{qr}$ , Eq. 1, and the mixed dissociation constants  $K_a$ , Eq. 2, are used whereas the mass balance equations are  $L = l + \sum \beta_{H_i} l h^i$  and  $H = h - \frac{K_w}{h} + j \sum \beta_{H_i} l h^i$ . Potentiometric readings obtained with the proton-sensitive glass and reference electrodes cell can be described by the equation:

$$E_{\text{cell}} = E^0 + \frac{fRT \ln 10}{F} \log_{10} a_{\text{H}^+} + j_a a_{\text{H}^+} - \frac{j_b K_w}{a_{\text{H}^+}} - E_{\text{ref}} = E^0 + S \log_{10} h \quad (3)$$

where  $E^0$  is the standard potential of a glass electrode cell containing some other constants of the glass electrode, including the asymmetry potential, etc., and  $a_{\text{H}^+} = [\text{H}^+]_{\text{yH}^+} = h y_{\text{H}^+}$ ; the liquid-junction potential  $E_j$  is expressed by the term  $E_j = j_a a_{\text{H}^+} - j_b K_w / a_{\text{H}^+}$ , and  $S = (fRT \ln 10) / F$  is the slope of the glass electrode for a Nernstian response,  $K_w$  is the operational ion product of water at temperature  $T$  [K], while the correction factor  $f$ , is taken as an adjustable parameter.

An explicit equation for the titration curve at constant ionic strength expresses the relationship between the volume of the added titrant  $V_i$  and the monitored emf  $E_{\text{cell},i}$  or  $p a_{\text{H}^+}$  with the vector of unknown parameters (**b**) being separated into the vector of common parameters (**K<sub>a</sub>**) and the vector of group parameters (**p**), i.e.  $V_i = f(E_{\text{cell},i}; \mathbf{b}) = f(E_{\text{cell},i}; \mathbf{K}_a, \mathbf{p})$ . The vector of common parameters  $\mathbf{K}_a = (K_{a,1}, \dots, K_{a,m})$  contains  $m$  dissociation constants of the acid  $\text{LH}_j$  while a vector of group parameters  $\mathbf{p} = (E^0, S, K_w, j_a, j_b, L_0, L_T, H_0, H_T)$  containing the two constants of Nernst's equation,  $E^0$  and  $S$ , and the total ligand concentration,  $L_0$ , and hydrogen ion concentration,  $H_0$ , of the titrand in the vessel, and the corresponding quantities of titrant,  $L_T$  and  $H_T$  in the burette [46–48]. Group parameters **p** can be refined simultaneously with the common parameters **K<sub>a</sub>**. Two independent regression approaches to a minimization of the sum of squared residuals have been applied:

(1) The program ESAB [46, 47] uses this strategy for treating  $p a_{\text{H}^+}$  data to find dissociation constants that give the “best” fit to experimental data. As the primary data contains the total concentration  $H_T$  of proton from the burette and the measured  $p a_{\text{H}^+}$ , one could trust  $p a_{\text{H}^+}$  and minimize the residual sum of squares  $(V_{\text{exp}} - V_{\text{calc}})^2$ . The residual  $e$  is formulated with the volume of added titrant  $V$  from burette so that  $e_i = (V_{\text{exp},i} - V_{\text{calc},i})$  and the resulting residual sum of squares  $U(\mathbf{b})$  is defined by:

$$U(\mathbf{b}) = \sum_{i=1}^n w_i (V_{\text{exp},i} - V_{\text{calc},i})^2 = \sum_{i=1}^n w_i e_i^2 \quad (4)$$

where  $w_i$  is the statistical weight usually set equal to unity while in ESAB it may be equal to:

$$\frac{1}{w_i} = s_i^2 = s_E^2 + \left[ \frac{dE_i}{dV_i} \right]^2 s_V^2 \quad (5)$$

and, with a good equipment, we have generally  $s_E = 0.01$  pH units and  $s_V = 0.0001$ – $0.0005$  cm<sup>3</sup>.

(2) In the program HYPERQUAD [49] the objective function is given in matrix notation  $U = \mathbf{e}^T \mathbf{W} \mathbf{e}$ , where  $\mathbf{e}$  is a vector of residuals measured in pH and  $\mathbf{W}$  is a matrix of weights. To minimize the objective function, the Gauss–Newton–Marquardt method is used. The

SIGMA criterion of a goodness-of-fit is defined as  $\text{SIGMA} = \sqrt{\frac{\sum_{i=1}^n w_i e_i^2}{n-m}}$  where the weights,  $w_i$ , are calculated from estimates of the error in pH and titer, the latter only being important in regions where the titration curve slopes more steeply. Sigma squared is also a Chi squared statistic.

### 2.3 Reliability of $pK_a$ Estimates Obtained

The detailed procedure of the graphical and numerical analysis of residuals is described in [21, 25]. The vector of residuals in each spectrum and finally in the entire absorbance matrix is statistically analyzed and the closest fit to the data is proven. The vector of residuals should exhibit a Gaussian distribution and the average of the absolute values of residuals should have a magnitude similar to the signal noise or instrumental standard deviation of absorbance  $s_{\text{inst}}(A)$ . The adequacy of a proposed regression model with experimental data and a reliability of found parameter estimates,  $b_j$ ,  $j = 1, \dots, m$ , may be examined by the goodness-of-fit test, cf. page 101 in Ref. [33].

#### 2.3.1 The Quality of the Parameter Estimates

The quality of the parameter estimates  $b_j$ ,  $j = 1, \dots, m$  [21] is considered according to their confidence intervals or according to their variances  $D(b_j)$ . Often an empirical rule is used: parameter  $b_j$  is considered to be significantly different from zero when its estimate is greater than 3 standard deviations,  $3\sqrt{D(b_j)} < |b_j|$ ,  $j = 1, \dots, m$ . Higher parameter variances are also caused by termination of a minimization process before reaching a minimum [33].

#### 2.3.2 The Quality of the Curve Fitting [21]

The adequacy of a proposed model and  $m$  parameter estimates found with  $n$  values of experimental data is examined by the goodness-of-fit test based on the statistical analysis of classical residuals. If a proposed model represents the data adequately, the residuals should form a random pattern having a normal distribution  $N(0, s^2)$  with the residual mean equal to zero,  $E(\hat{e}) = 0$ , and the standard deviation of residuals  $s(\hat{e})$  being near the noise, i.e. the experimental error  $\varepsilon$ . Systematic departures from randomness indicate that the model and parameter estimates are not satisfactory. The following statistics of residuals can be used for a numerical goodness-of-fit evaluation, cf. page 290 in Ref. [50]: (1) The residual bias, which is the arithmetic mean of residuals  $E(\hat{e})$ , should be equal to zero. (2) The mean of absolute values of residuals,  $E|\hat{e}|$ , and the square-root of the residuals variance  $s^2(\hat{e}) = U(\mathbf{b})/(n-m)$ , known as the estimate of the residual standard deviation,  $s(\hat{e})$ , should be both be of the same magnitude as the instrumental error of the regressed

variable, absorbance  $A$ ,  $s_{\text{inst}}(A)$ . Obviously, it should be also valid that  $s(\hat{\epsilon}) \approx s_{\text{inst}}(A)$ . (3) The residual skewness,  $g_1(\hat{\epsilon})$ , for a symmetric distribution of residuals should be equal to zero. (4) The kurtosis,  $g_2(\hat{\epsilon})$ , for normal distribution should be equal to 3.

### 2.3.3 Quality of the Molar Absorption Coefficients

The numerical estimates of the molar absorption coefficients [21] of differently protonated light-absorbing species in an equilibrium mixture as functions of wavelength  $\lambda$  represent another important result of the spectra regression analysis.

### 2.3.4 The Distribution Diagram

The distribution diagram [21] presents the relative concentrations of differently protonated light-absorbing species in the protonation equilibria and provides a specific image on the protonation model. It allows for the chemical interpretation of a proposed regression model, to perform its correction, to comment on the presentation of major and minor species in an equilibrium mixture, and to reveal which protonated species are present in the solution at a given pH. It represents the culmination of an interpretation of the regression analysis of the spectra [51].

### 2.3.5 The Deconvolution of Each Experimental Spectrum

The deconvolution of each experimental spectrum [21] into the spectra for the individual species shows whether the experimental design was efficient. If for a particular concentration range the spectrum consists of just a single component, further spectra for that range would be redundant though they should improve the precision. In ranges where many components contribute significantly to the experimental spectrum, several spectra should be measured.

## 3 Materials and Methods

### 3.1 Chemicals and Solutions

Eltrombopag was donated by ZENTIVA GROUP, Ltd. (Prague) with declared purity checked by a HPLC method and alkalimetrically, was always  $> 99\%$ . This drug has been weighed straight into a reaction vessel to reach a resulting concentration of about  $0.001 \text{ mol}\cdot\text{dm}^{-3}$ . Hydrochloride acid,  $1.044 \text{ mol}\cdot\text{dm}^{-3}$ , was prepared by diluting concentrated HCl (p.a., Lachema Brno) with redistilled water and standardization against HgO and KI with a reproducibility better than 0.2% according to the equation  $\text{HgO} + 4 \text{KI} + \text{H}_2\text{O} \rightleftharpoons 2 \text{KOH} + \text{K}_2[\text{HgI}_4]$  and  $\text{KOH} + \text{HCl} \rightleftharpoons \text{KCl} + \text{H}_2\text{O}$ . Potassium hydroxide,  $0.876 \text{ mol}\cdot\text{dm}^{-3}$ , was prepared from the exact weight of pellets p.a., Aldrich Chemical Company with carbon-dioxide free redistilled water. The solution was stored for several days in a polyethylene bottle under argon. This solution was standardized against a solution of potassium bi-phthalate using the derivative method with reproducibility 0.1%. All solutions were preserved from atmospheric  $\text{CO}_2$  by means of soda lime traps. Mercury oxide, potassium iodide and potassium chloride, p.a. Lachema Brno, were not extra purified. Grade A glassware and twice-redistilled water were employed in the preparation of all the solutions.

### 3.2 Apparatus and Procedure

The apparatus used and the WApH-spectrophotometric titration procedure have been described previously in detail [13, 25, 26]. The experimental and computation scheme to determine of the protonation constants of the multi-component system is taken from Meloun et al. (cf. page 226) in Ref. [33] and the five steps are described in detail elsewhere [26].

The free hydrogen-ion concentration  $h$  was measured on the digital voltmeter (Hanna HI 3220) with a precision of  $\pm 0.002$  pH units, using a Theta HC 103-VFR combined glass electrode. Titrations were performed in a water jacketed double-walled 100 mL glass vessel, closed with a Teflon bung containing the electrodes, an argon inlet, a thermometer, a propeller stirrer and the capillary tip from a micro-burette. All pH measurements were carried out at  $(25.0 \pm 0.1)$  °C. During the titrations, a stream of argon gas was bubbled through the solution both for stirring and for maintaining an inert atmosphere. The argon was passed through two vessels containing the titrand medium before entering the corresponding titrant solution. All titrations were performed using standardized  $1 \text{ mol}\cdot\text{dm}^{-3}$  HCl or  $1 \text{ mol}\cdot\text{dm}^{-3}$  KOH titrants. In general, sample solutions of 20 mL volumes were preacidified to a relatively low pH value (ca. 2–3) and then titrated alkalimetrically to an appropriate high pH value (ca. 10–11). The burettes used were syringe micro-burettes of 1250  $\mu\text{L}$  capacity (META, Brno) with a 25.00 cm micrometer screw [52]. The potentiometric titrations of drugs with potassium hydroxide were performed using a hydrogen activity scale. Standardization of the pH meter was performed using WTW standard buffers values 4.006, 6.865 and 9.180 at 25°C.

The ESAB program [46, 47] estimated the total proton concentration in a burette  $H_T$  and the total concentration of the drug in the titration vessel  $L_0$  from the actual titration of a mixture of the drug and hydrochloric acid with potassium hydroxide; some group parameters are given in the input data for ESAB, including the Nernstian slope and  $pK_w$ , which are both accessible from the literature [53]. With ESAB, two group parameters,  $L_0$  and  $H_T$ , were refined to give the best fit, while the fitness was examined by the goodness-of-fit criteria.

### 3.3 Computation and Software

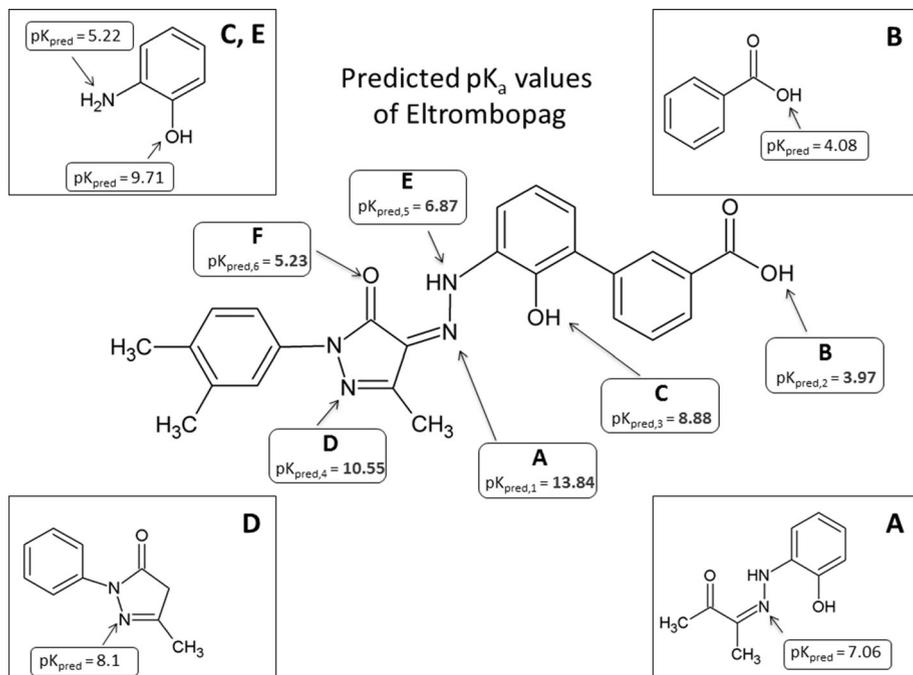
Computation relating to the determination of dissociation constants was performed by regression analysis of the UV–Vis spectra using the SQUAD84 and REACTLAB programs. A qualitative interpretation of the spectra, with the use of the INDICES program [51], aims to evaluate the quality of the dataset and remove spurious data and to estimate the minimum number of factors (i.e. contributing aqueous species) that are necessary to describe the experimental data and determine the number of dominant species present in the equilibrium mixture. Computation of the dissociation constants was performed by regression analysis of titration curves using the ESAB and HYPERQUAD programs [49]. Most graphs were plotted using ORIGIN 8 [50] and S-Plus [33]. ACD/pK [46] and MARVIN [36, 54] are programs for predictions based on the structural formulae of drug compounds. Entering the compound topological structure descriptors graphically,  $pK_a$  values of organic compound are predicted using hundreds of Hammett and Taft equations and quantum chemistry calculus.

## 4 Results and Discussion

### 4.1 Computational Prediction of the Protonation Scheme

The protonation scheme of the Eltrombopag has six functional groups (denoted with letters A, B, C, D, E, F in graphical order in Fig. 2) that can be associated to dissociation constants; two ionizations are associated to hydroxyl groups and other to the nitrogen atom. In the following text the charges of ions are omitted for the sake of simplicity and ions are denoted as variously protonated species.

The macro-dissociation constants of Eltrombopag were predicted according to the chemical structure analyzed with the use of two reliable  $pK_a$  prediction tools [42]: ACD/Percepta [34, 43, 44] was run using the GALAS model, which uses an internal training set of > 31,000 individual  $pK_a$  values for approximately 16,000 compounds in aqueous solution [55]. Marvin  $pK_a$  predictions are based on the calculated partial charge of the atoms located in the analyzed structure, using Hammett–Taft's approach. MARVIN and ACD/Percepta [34, 43] showed similar performance on the dataset and provide  $pK_a$  results for all the recognized ionization sites. Predicted  $pK_a$  values are assigned to the corresponding ionization sites by both tools, which is essential for compounds with multiple ionization sites (Fig. 2). The whole molecule of Eltrombopag was further subdivided into four auxiliary fragments containing functional groups on which protonation occurred (Fig. 2). These predicted  $pK_a$  values served to compare with predicted values throughout



**Fig. 2** Structural formula with protonated ionization sites in Eltrombopag. The whole molecule of Eltrombopag was subdivided into four auxiliary fragments containing functional groups on which protonation occurred. These predicted  $pK_a$  values served to compare with predicted values throughout the structure of the Eltrombopag molecule

the structure of the Eltrombopag molecule. In the ionization site B the predicted  $pK_a$  of the Eltrombopag molecule is close to the predicted  $pK_a$  values of its auxiliary fragment. Other ionization sites (A, C, D, E, F) are located at the site of the molecule that forms a rather complex conformation. The fragments A, C, D, E contain sites which are not affected by the strong electron field of the rest of Eltrombopag molecule, and therefore their predicted  $pK_a$  values differ significantly from those predicted for the whole Eltrombopag molecule (Table 1).

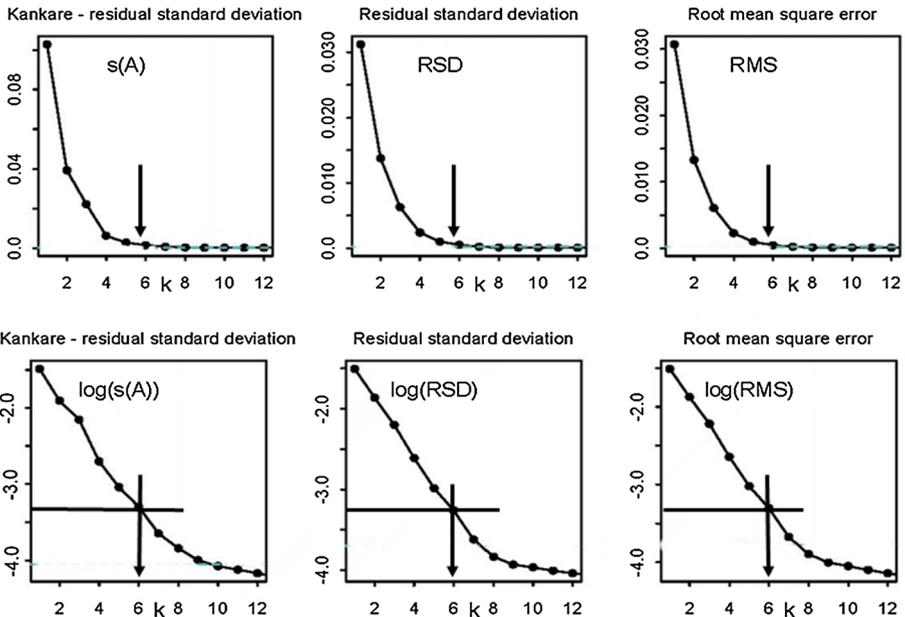
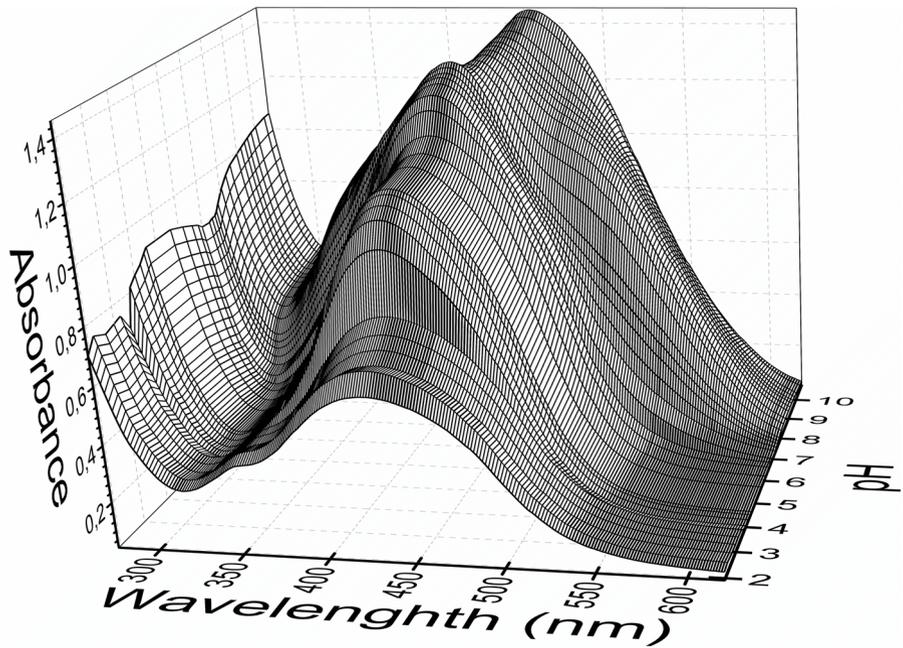
## 4.2 UV-Metric $pK_a$ Determination

Many drug compounds are sparingly soluble in water and a precise determination of their  $pK_a$  values pose a challenging problem for potentiometric titration, since the accuracy of this method is restricted by its detection limit of about  $10^{-4}$  mol·dm<sup>-3</sup> [56]. Spectroscopic titration has been utilized as an alternative to determine  $pK_a$  values of substances with large molar absorptivities, because of its high sensitivity at concentrations of substance as low as  $10^{-6}$  mol·dm<sup>-3</sup> [56]. The strategy for an efficient determination of dissociation constants followed by spectral data treatment, as described in [26], was used. Eltrombopag contains the complicated molecular structure shown in Figs. 1 and 2 and several protonation equilibria were monitored spectrophotometrically to analyze a spectra set in two steps: first, the spectral data in the form of a data matrix were subjected to principal component analysis to determine the number of independent light absorbing species using the INDICES algorithm [51], (Fig. 3).

The INDICES indicates the position of break points on the  $s_k(A) = f(k)$  curve in the scree plot using the three most reliable approaches (Kankare's  $s(A)$ , RSD and RSM, cf. Ref. [51]) and gives  $k^* = 6$  with corresponding co-ordinate  $s_6^*(A) = 0.4$  mAU. This value also represents the actual instrumental error  $s_{inst}^*(A) = 0.4$  mAU and  $\log_{10}(s_{inst}^*(A)) = -3.4$  of the spectrophotometer CINTRA 5 (GBC, Australia). The number of light-absorbing species is an aid to the establishment of a protonation model. It means that five dissociation constants will be preferred and six species  $LH_5^+$ ,  $LH_4^+$ ,  $LH_3$ ,  $LH_2^-$ ,  $LH^{2-}$  and  $L^{3-}$  are assumed to be present. Due to the large variations in the indicator values, these

**Table 1** Predicted dissociation constants at the suggested ionization sites A through F of the molecule Eltrombopag with the use of the program MARVIN

Tautomer, protonated ionization sites	$pK_{a1}$	$pK_{a2}$	$pK_{a3}$	$pK_{a4}$	$pK_{a5}$	$pK_{a6}$
1 A, B, C	0.13	3.97	9.36	–	13.13	–
2 A, B, C, D	4.70	3.61	9.40	–	14.56	–
3 B, C, D	0.51	3.97	7.20	6.22	–	–
4 B, C, F	–	3.94	12.17	1.53	–	5.23
5 B, C	–0.80	3.97	7.68	–	–	–
6 B, C, D, F	1.05	3.96	7.78	10.15	–	5.54
7 B, C, D	–	4.00	7.68	2.89	–	–
Auxiliary fragments	For B is 4.08	For E is 5.22	For A is 7.06	For D is 8.1	For C is 9.71	Not found
Site in molecule of Eltrombopag	For B is 3.97	For E is 6.87	For A is 13.84	For D is 10.55	For C is 8.88	For F is 5.23



**Fig. 3** The 3D-absorbance-response-surface for 71 measured absorption spectra for  $9.5 \times 10^{-5} \text{ mol-dm}^{-3}$  Eltrombopag at 25 °C, reflecting the dependence of protonation equilibria on pH. This represents the input for the SQUAD84 and SPECFIT/32 programs (upper part). The Cattell's scree plot of the Wernimont-Kankare procedure for the determination of the rank of the absorbance matrix of Eltrombopag  $k^* = 6$  leads to six light-absorbing species in the mixture,  $n_c = 6$ , with the use of Kankare's  $s(A)$ , RSD and RSM (lower part)

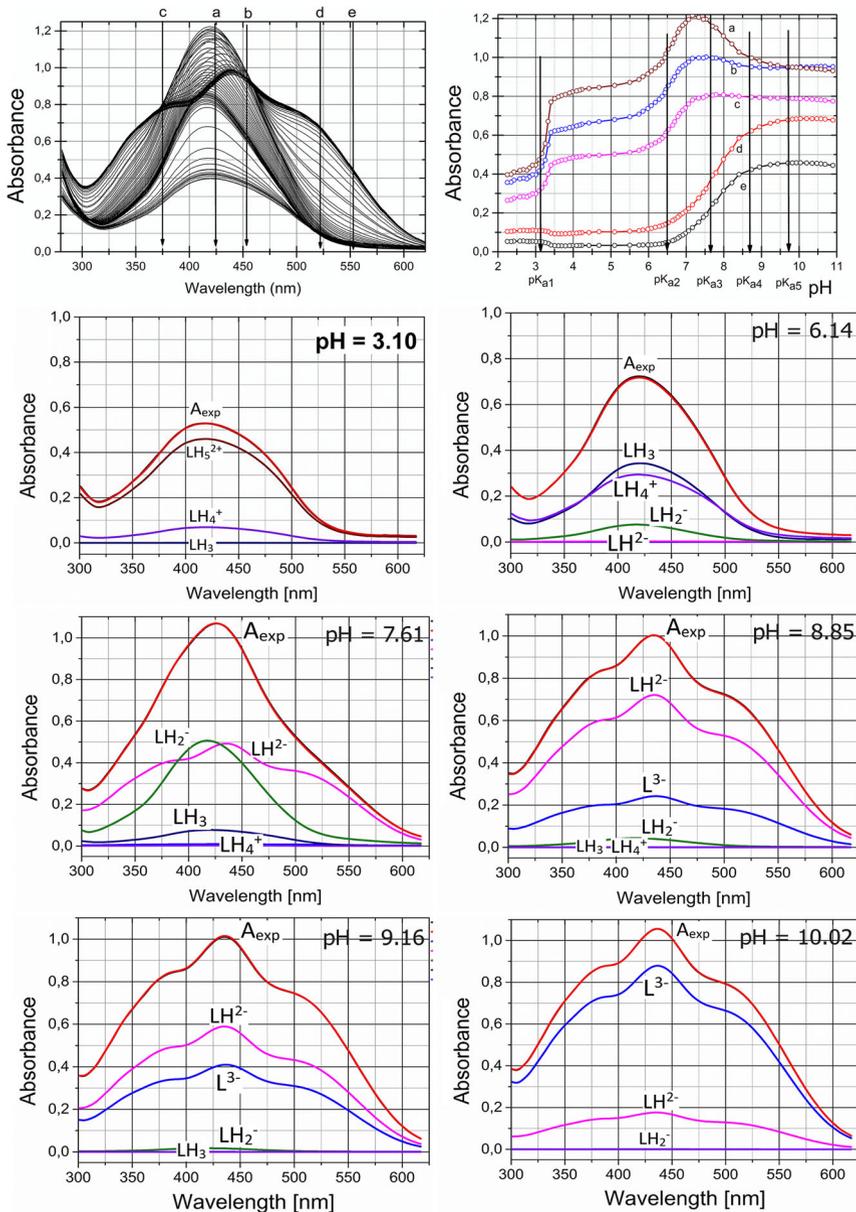
latter graphs are plotted on a logarithmic scale (Fig. 3) and the number of light absorbing species  $p$  can be predicted from the index function by finding the point  $p = k^*$  where the slope of index function  $PC(k) = f(k)$  changes, or by comparing  $PC(k)$  values with the instrumental error  $s_{\text{inst}}^*(A) = 0.4$  mAU when  $\log_{10} s_{\text{inst}}^*(A) = -3.4$ . This is common criterion for determining  $p$ . Low values of  $s_{\text{inst}}^*(A)$  prove the reliability of the spectrophotometer and experimental technique used [26].

In the spectra set in Fig. 4, the five analytical wavelengths (a) through (e) were those at which the absorbance-pH curves were analyzed. Six following graphs on Fig. 4 show the consecutive deprotonation response in the spectra, when each spectrum was deconvoluted into the spectra of differently protonated species. At pH = 3.10 the species  $\text{LH}_5^{2+}$  accompanied by species  $\text{LH}_4^+$  predominates in the solution. At pH = 6.14, together with the molecule  $\text{LH}_3$ , two species  $\text{LH}_2^-$ ,  $\text{LH}_4^+$  exhibit absorption bands at the same wavelength of absorption maximum  $\lambda_{\text{max}}$ . At pH = 7.61 the experimental spectrum is decomposed to three absorption bands concerning the species  $\text{LH}_3$  which dissociate to species  $\text{LH}_2^-$  and  $\text{LH}^{2-}$ . At pH = 8.85 and 9.16 the species  $\text{L}^{3-}$  occurs with  $\text{LH}^{2-}$ , while the concentration of  $\text{L}^{3-}$  increases up to pH 10.02.

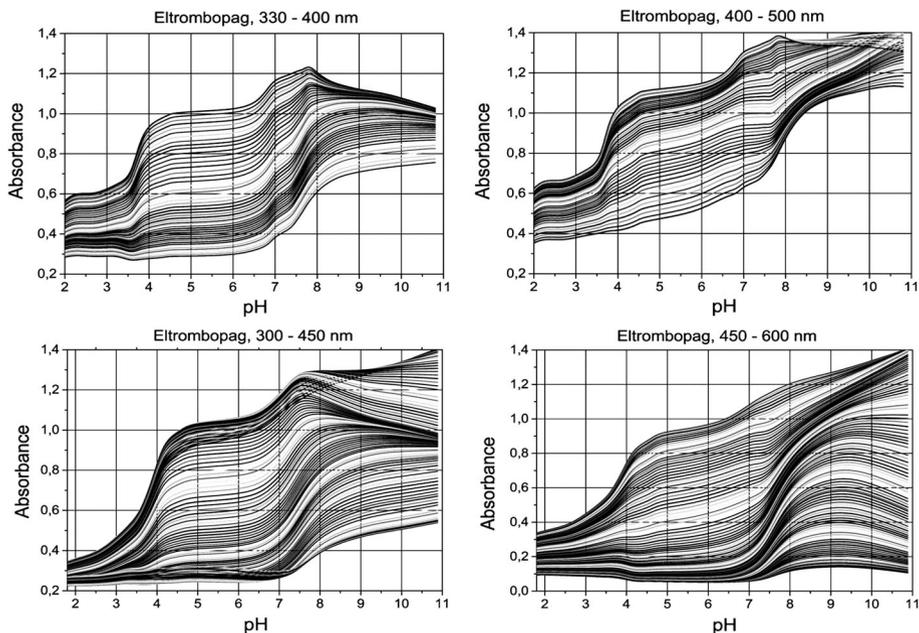
Attention should be paid to the sensitivity of the analytical wavelengths chosen to determine the  $\text{pK}_a$  values when overlapped  $\text{pK}_a$  values are observed in compounds in which the ionizable groups are in symmetrical positions and do not interact. Figure 5 shows four sets of pH-spectrophotometric titration spectra, monitored at various wavelengths, in which chromophore(s) are sensitive on pH change. The results of search for the best chemical model containing either 4 or 5 dissociation constants are shown in Table 2. Four useful wavelength ranges and the overall range were examined to determine the spectral range in which the actual chromophore is active and reflects protonation/dissociation of the molecule. The best regression model was determined by testing two working protonation models: the first involving four and the other with five dissociation constants. Criterion of reliability was the goodness-of-fit test. At the same time the estimates of the dissociation constants using SQUAD84 and REACTLAB were compared. The standard deviation of residuals and Hamilton  $R$ -factor of a relative fitness generally showed that the better fit of the calculated spectra was always for the protonation model with five dissociation constants.

Five dissociation constants  $\text{pK}_{a1}$ ,  $\text{pK}_{a2}$ ,  $\text{pK}_{a3}$ ,  $\text{pK}_{a4}$ ,  $\text{pK}_{a5}$  and six molar absorptivities of Eltrombopag  $\varepsilon_{\text{L}}$ ,  $\varepsilon_{\text{LH}}$ ,  $\varepsilon_{\text{LH}_2}$ ,  $\varepsilon_{\text{LH}_3}$ ,  $\varepsilon_{\text{LH}}$  and  $\varepsilon_{\text{LH}_5}$  were estimated using SQUAD84 and REACTLAB in the first run. The reliability of the parameter estimates may be tested with the following diagnostics:

The first diagnostic value indicates whether all of the parametric estimates  $\beta_r$  and  $\varepsilon_r$  have physical meaning and reach realistic values. As the standard deviations  $s(\log_{10} \beta_r)$  of  $\log_{10} \beta_r$  and  $s(\varepsilon_r)$  of  $\varepsilon_r$  are significantly smaller than their corresponding parameter estimates, all the variously protonated species are statistically significant at a significance level  $\alpha = 0.05$ . The absolute values of  $s(\beta_r)$ ,  $s(\varepsilon_r)$  gives information about the last  $\text{RSS}$ -contour of the hyperparaboloid in the neighborhood of the pit,  $\text{RSS}_{\text{min}}$ . For well-conditioned parameters, the last  $\text{RSS}$ -contour is a regular ellipsoid, and the standard deviations are reasonably low. High  $s$  values are found with ill-conditioned parameters and a “saucer”-shaped pit. The relation  $s(\beta_j) \times F_\sigma < \beta_j$  should be met where  $F_\sigma$  is equal to 3. The set of standard deviations of  $\varepsilon_r$  for various wavelengths,  $s(\varepsilon_r) = f(\lambda)$ , should have a Gaussian distribution; otherwise, erroneous estimates of  $\varepsilon_r$  are obtained. Upper part of Fig. 6 shows the estimated molar absorptivities of all of the protonated species,  $\varepsilon_{\text{L}}$ ,  $\varepsilon_{\text{LH}}$ ,  $\varepsilon_{\text{LH}_2}$ ,  $\varepsilon_{\text{LH}_3}$ ,  $\varepsilon_{\text{LH}}$  and  $\varepsilon_{\text{LH}_5}$ , of Eltrombopag as functions of wavelength. Three species  $\text{LH}_2^-$ ,  $\text{LH}^{2-}$  and  $\text{L}^{3-}$  exhibit similar spectra at  $\lambda_{\text{max}} = 440$  nm; the intensities of which decrease with increasing protonation.



**Fig. 4** In a spectra set, the five analytical wavelengths (*a*) through (*e*) were selected at which the absorbance-pH curves were plotted. The six following figures from pH = 3.10 through pH = 10.02 show the consecutive deprotonation response in spectra, when each spectrum was deconvoluted to the spectra of the differently protonated species present. At pH = 3.10 the species  $\text{LH}_5^{2+}$  predominates accompanied by  $\text{LH}_4^+$ . At pH = 6.14, together with the species  $\text{LH}_3$  two species  $\text{LH}_2^+$ ,  $\text{LH}_4^+$  exhibit absorption bands at the same wavelength of absorption maximum  $\lambda_{\text{max}}$ . At pH = 7.61 the experimental spectrum is decomposed to three absorption bands for the species  $\text{LH}_3$  which dissociate to species  $\text{LH}_2^-$  and  $\text{LH}^{2-}$ . At pH = 8.85 and 9.16 the species  $\text{L}^{3-}$  occurs with species  $\text{LH}_2^-$  and  $\text{LH}^{2-}$ , and concentration of  $\text{L}^{3-}$  in the solution increases up to pH = 10.02



**Fig. 5** Four wavelength regions of the 2D-absorbance-pH response spectra set for  $9.5 \times 10^{-5} \text{ mol}\cdot\text{dm}^{-3}$  Eltrombopag at 25 °C

The species  $\text{L}^{3-}$  exhibits the highest value  $\epsilon_{\text{max}}$  at  $\lambda_{\text{max}} = 440$  while the species  $\text{LH}_2^-$  has the lowest value  $\epsilon_{\text{max}}$ . The species  $\text{LH}_3$  exhibits a hypsochromic shift relative to  $\text{LH}_2^-$  to a lower value of  $\lambda_{\text{max}} = 420$  nm, which is the result of further protonation. Further protonation to the species  $\text{LH}_4^+$  results in a significant decrease in  $\epsilon_{\text{max}}$  at the same  $\lambda_{\text{max}}$  as  $\text{LH}_3$ . Further acidification to pH = 2 or 1 leads to species  $\text{LH}_5^{2+}$  whose spectrum increases. Interestingly, the spectra of  $\text{LH}_2^-$ ,  $\text{LH}^{2-}$  and  $\text{L}^{3-}$  are similar with a common  $\lambda_{\text{max}} = 440$  nm and the spectra of  $\text{LH}_3$ ,  $\text{LH}_4^+$  and  $\text{LH}_5^{2+}$  are similar in shape with the common  $\lambda_{\text{max}} = 420$  nm.

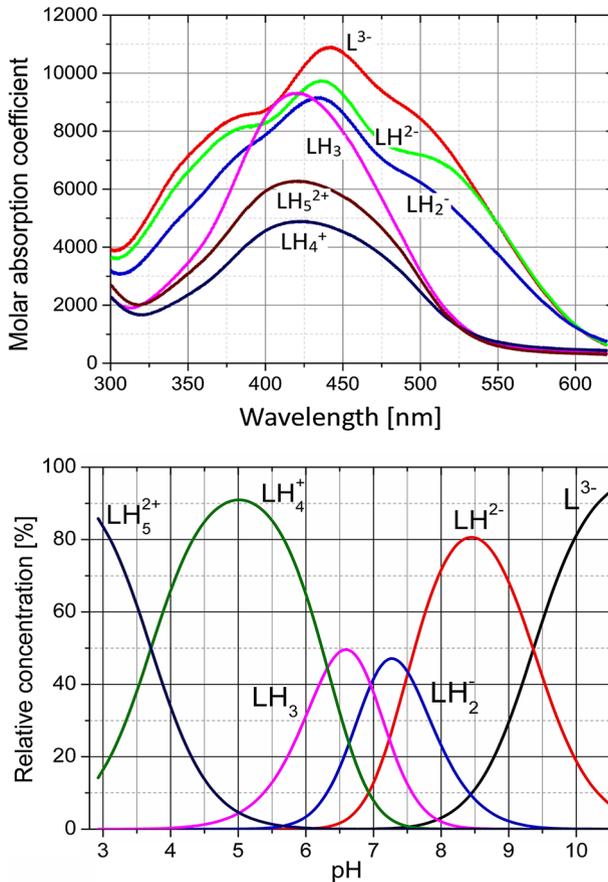
The second diagnostic examines whether all of the calculated relative concentrations of the variously protonated species in the distribution diagram have physical meaning, which proved to be the case (lower part of Fig. 6) [13]. The calculated free concentrations of the basic components and variously protonated species of the chemical model should show molarities down to about  $10^{-8} \text{ mol}\cdot\text{dm}^{-3}$ . Expressed in percentage terms, a species present at about 1% relative or less in an equilibrium behaves as numerical noise in a regression analysis. The distribution diagram in Fig. 6 makes it easier to judge the contributions of individual species to the total concentration. Since the molar absorptivities will generally be in the range  $10^3$ – $10^5 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ , species present at less than ca. 0.1% relative concentration will affect the absorbance significantly only if their  $\epsilon$  is extremely high. The distribution diagram shows the protonation equilibria of  $\text{LH}_5^{2+}$ ,  $\text{LH}_4^+$ ,  $\text{LH}_3$ ,  $\text{LH}_2^-$ ,  $\text{LH}^{2-}$  and  $\text{L}^{3-}$ . At neutral pH (5–8) Eltrombopag is predominantly  $\text{LH}_3$  and from pH = 6 to pH = 9, in the form of species  $\text{LH}_2^-$ . In the pH range of 6–10 the species  $\text{LH}_3$  deprotonates to the species  $\text{LH}_2^-$  and  $\text{LH}^{2-}$  and finally to  $\text{L}^{3-}$ . Acidification of  $\text{LH}_3$  solution gives firstly  $\text{LH}_4^+$ , which in solutions of pH = 3–7 predominates reaching 90% relative concentration. Further acidification from pH = 4 to pH = 1 yields the species  $\text{LH}_5^{2+}$ . At concentrations of  $10^{-4}$  to



**Table 2** continued

Residual-square-sum function <i>RSS</i>	SQUAD84	0.04213	0.03208	0.09593	0.08435	0.11700	0.09120	0.07408	0.04500	0.21250	0.15830
	ReactLab	0.04216	0.03224	0.09590	0.08413	0.11803	0.09081	0.07395	0.04550	0.23914	0.14765
Hamilton R-factor from SQUAD84 (%)	SQUAD84	0.48	0.42	0.47	0.44	0.50	0.44	0.58	0.45	0.55	0.47
	ReactLab	—	—	—	—	—	—	—	—	—	—

Solution of  $9.5 \times 10^{-5}$  mol·dm<sup>-3</sup> Eltrombopag at *I* = 0.0001 mol·dm<sup>-3</sup> at 25 °C, for *n<sub>s</sub>* spectra measured at *n<sub>w</sub>* wavelengths for *n<sub>z</sub>* = 2 basic components L and H forms *n<sub>c</sub>* = 6 variously protonated species. The standard deviations of the parameter estimates are in the last valid digits in parentheses. The resolution criterion and reliability of parameter estimates found are proven with goodness-of-fit statistics: the residual standard deviation by factor analysis *s<sub>k(A)</sub>* (mAU), the mean residual  $E|\bar{\epsilon}|$  (mAU), the standard deviation of absorbance after termination of the regression process *s(ĉ)* (mAU), the sigma *s(A)* (mAU) from REACTLAB, the residual square sum *RSS*, the Hamilton R-factor of relative fitness (%) from SQUAD84



**Fig. 6** The graphs of the molar absorption coefficients of six variously protonated species of Eltrombopag against wavelength (upper part). Corresponding distribution diagram of the relative concentration of the six variously protonated species for Eltrombopag (lower part), (SPECFIT, ORIGIN)

$10^{-6}$  mol·dm $^{-3}$ , Eltrombopag is sufficiently soluble that all of its dissociation constants can be spectrophotometrically determined.

The next diagnostic concerns the goodness-of-fit [20]. The goodness-of-fit achieved is easily seen by examination of the differences between the experimental and calculated absorbance values,  $e_i = A_{\text{exp},ij} - A_{\text{calc},ij}$ . Examination of the spectra and of the graph of the predicted absorbance response surface through all the experimental points should reveal whether the results calculated are consistent and whether any gross experimental errors have been made in the measurement of the spectra. One of the most important statistics calculated is the standard deviation of absorbance  $s(A)$ , calculated from a set of refined parameters at the termination of the minimization process. Although this statistical analysis of residuals [26] gives the most rigorous test of the degree-of-fit, realistic empirical limits must be used. The statistical measures of all residuals  $e$  prove that the minimum of the elliptic hyperparaboloid  $RSS$  is reached (Table 2): the residual standard deviation  $s(\hat{e})$  always has sufficiently low values, lower than 3 mAU, which is less than 0.2% of measured absorbance value.

Dissociation constants estimated with SQUAD84 and REACTLAB are in a good agreement. The SQUAD approach has the great advantage in rigorous goodness-of-fit test made by the statistical analysis of residuals. Reproducibility of four experimental spectra sets with the use of two regression programs shows that  $pK_{a3} = 7$ ,  $pK_{a4} = 7.5$  and  $pK_{a5} = 9.3$  are well-conditioned in the regression model and therefore their numerical evaluation is quite reliable. The first two dissociation constants  $pK_{a1} = 3.1$  and  $pK_{a2} = 6.5$  are ill-conditioned in the regression model; the hyperparaboloid on these two parameters is rather saucer shaped without a distinctive minimum. Numerical enumeration of all coordinates of this minimum is more difficult and the parameter estimates are therefore less reliable. Acidifying the solution of a species  $LH_3$  leading to species  $LH_4^+$  and  $LH_5^{2+}$  may be disturbed by precipitation of Eltrombopag, which manifests itself especially at higher concentrations in potentiometric determination. For this reason, it may not be at  $0.0001 \text{ mol}\cdot\text{dm}^{-3}$ , because Eltrombopag in acidic solution precipitates. Both programs gave the same estimates of all five dissociation constants and identical results for the goodness-of-fit test.

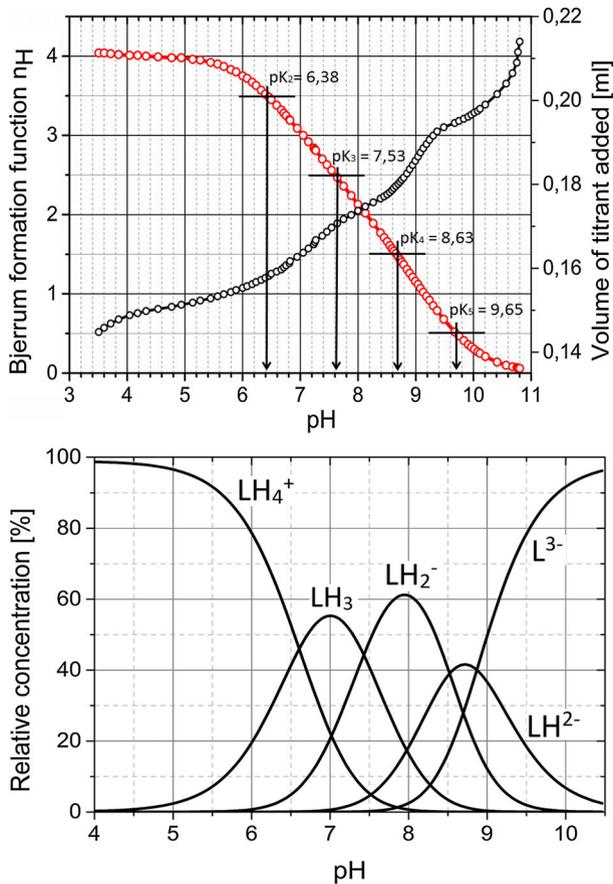
### 4.3 Potentiometric Titration Data Analysis

The potentiometric titration of a mixture of HCl and Eltrombopag with potassium hydroxide was carried out at  $25 \text{ }^\circ\text{C}$  at a constant ionic strength (Fig. 7). The initial tentative value of the dissociation constant of the drug, corresponding to the midpoint value in each plateau of the potentiometric titration curve, was refined using the ESAB and the HYPERQUAD programs.

Because Eltrombopag exhibits four close dissociation constants, their numerical estimation is rather difficult and impossible without the use of a computer assisted nonlinear regression. A regression analysis was employed with the use of a plateau of the middle part titration curve, for alkalized Eltrombopag titrated with hydrochloric acid, followed by a subsequent retitration with potassium hydroxide. The assessed titration curve was calculated as well the Bjerrum formation protonation curve function, which is shown in the graph in Fig. 7. On the Bjerrum formation curves the estimates of three or four dissociation constants  $pK_{a2}$ ,  $pK_{a3}$ ,  $pK_{a4}$ ,  $pK_{a5}$  are plotted. Since at pH above 9 and pH below 5 Eltrombopag forms a fine precipitate, which is observed as a slight opalescence, the titration data for pH above 9 and below 5 were not used in the regression analysis.

Because it is difficult, in regression analysis, to estimate such close overlapping dissociation constants, two computer programs, ESAB and HYPERQUAD, were used and the resulting  $pK$  estimates were compared. The programs differ in the definition of the sum of the squares of residuals. While in the ESAB the residuals are defined as the difference between the experimental and calculated titrant volume, in the HYPERQUAD the residuals are defined as differences between the experimental and calculated values of pH. The assumptions of the least-squares method require that the independent variable is not subject to significant random errors and that the regression analysis is of the dependent variable, which is carries random experimental errors. This assumption is met only with the program HYPERQUAD.

Table 3 shows the results of the ESAB and HYPERQUAD regression analyses of a selected part of the titration curve when the minimization process terminates. Both the common and the group parameters are refined and the best curve-fitting is proven by the results of a statistical analysis of the residuals. The reliability of the dissociation constant may be determined from the goodness-of-fit test in which an increasing number of group parameters are refined, a better fit is achieved and therefore a more reliable estimate of



**Fig. 7** Protonation equilibria of Eltrombopag analyzed with ESAB: the pH-potentiometric titration curve of acidified Eltrombopag plus HCl titrated with KOH is plotted with the Bjerrum protonation function indicating  $pK$  values (upper part). The distribution diagram of a relative presentation of variously protonated species  $L^{3-}$ ,  $LH_2^-$ ,  $LH_3$  and  $LH_4^+$  of Eltrombopag as functions of pH at 25 °C (lower part) (ESAB, HYPERQUAD, ORIGIN)

dissociation constants results. As further group parameters are refined the fit is improved. A quite sensitive criterion of the reliability of the dissociation constant is the mean of absolute values of residuals  $E|\hat{\epsilon}|$ . Comparing residuals with the instrumental noise,  $s_{\text{inst}}(y)$ , represented here by either  $s_{\text{inst}}(y) = s(V) = 0.0001 \text{ mL}$  or  $s_{\text{inst}}(y) = s(\text{pH}) = 0.01$ , an excellent fit is confirmed because the mean  $E|\hat{\epsilon}|$  and the residual standard deviation  $s(\hat{\epsilon})$  are nearly the same and lower than the experimental noise  $s_{\text{inst}}(y)$ . Here,  $E|\hat{\epsilon}| = 0.0001 \text{ mL}$  and  $s(\hat{\epsilon}) = 0.0002 \text{ mL}$  are similar and both are lower than the microburette error  $s(V) = 0.0001 \text{ mL}$ . As the bias  $E(\hat{\epsilon})$  is equal to  $-2.6 \times 10^{-6}$  in ESAB, which may be taken as near to zero, no systematic error in curve fitting is expected. All residuals oscillate between lower  $-0.0002 \text{ mL}$  and upper  $0.0001 \text{ mL}$  Hoaglin's inner bounds and therefore no outlying residuals lay outside these bounds. Residuals exhibit a normal distribution as confirmed by the Jarque–Berra normality test for combined sample skewness and kurtosis (cf. page 80 in Ref. [57]), and also by the skewness  $g_1(\hat{\epsilon})$  is near 0 (which is proving a

**Table 3** Four dissociation constants  $pK_{a2}$ ,  $pK_{a3}$ ,  $pK_{a4}$ ,  $pK_{a5}$  of Eltrombopag when their standard deviations in last valid digits are in parentheses

	HYPERQUAD	ESAB
Number of points $n$	39	39
$pK_{a2}$ ( $s_1$ ), $H_4L \rightleftharpoons H^+ + H_3L$	6.60(13)	6.59(01)
$pK_{a3}$ ( $s_2$ ), $H_3L \rightleftharpoons H^+ + H_2L^-$	7.38(12)	7.56(04)
$pK_{a4}$ ( $s_3$ ), $H_2L \rightleftharpoons H^+ + HL$	8.46(09)	8.48(59)
$pK_{a5}$ ( $s_4$ ), $HL \rightleftharpoons H^+ + L$	8.77(07)	9.29(34)
$L_0$ concentration of drug (mol·dm <sup>-3</sup> )	$4.6 \times 10^{-4}$	$1.7 \times 10^{-4}$
Goodness of fit test by the statistical analysis of residuals in pH (HYPERQUAD) or in $V$ (mL) (ESAB)		
Sigma in pH units (HYPERQUAD)	0.615 pH units	*
Bias, arithmetic mean of residuals $E(\hat{\epsilon})$	$-9.05 \times 10^{-3}$ pH units	$6.9 \times 10^{-5}$ mL
Residual mean, $E \hat{\epsilon} $	0.0228 pH units	0.0001 mL
Standard deviation of residuals, $s(\hat{\epsilon})$	0.0335 pH units	0.0004 mL
Residual skewness, $g_1(\hat{\epsilon})$	1.08	0.14
Residual kurtosis, $g_2(\hat{\epsilon})$	8.42	3.39
Jarque–Bera normality test	Rejected	Accepted

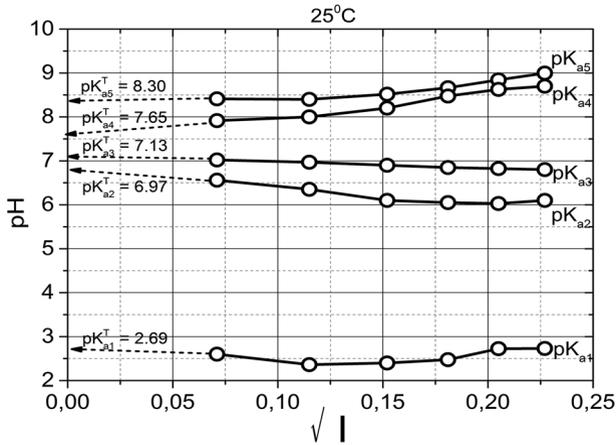
The reliability of parameter estimation is proven with a goodness-of-fit statistics: the sigma in pH units from HYPERQUAD, the bias or arithmetic mean of residuals  $E(\hat{\epsilon})$  (mL), the residual mean  $E|\hat{\epsilon}|$  (mL), the standard deviation of residuals  $s(\hat{\epsilon})$  (mL), the residual skewness  $g_1(\hat{\epsilon})$  and the residual kurtosis  $g_2(\hat{\epsilon})$  proving a Gaussian distribution and Jarque-Berra normality test. ESAB and HYPERQUAD refinement of common and group parameters for a titration of Eltrombopag with HCl and KOH were performed

Common parameters refined:  $pK_{a2}$ ,  $pK_{a3}$ ,  $pK_{a4}$ ,  $pK_{a5}$ . Group parameters refined:  $L_0$ . Constants:  $H_T = -0.8138$  mol·dm<sup>-3</sup>,  $t = 25.0$  °C,  $pK_w = 13.9799$ ,  $V_0 = 20.22$  mL,  $s(V) = s_{inst}(y) = 0.0001$  mL,  $s(pH) = s_{inst}(y) = 0.01$ ,  $I_0 = 0.004$  (in vessel),  $I_T = 0.8138$  (in burette KOH) or 1.0442 (in burette HCl)

\* Means that statistics is not available in the program's output

symmetric distribution), and the kurtosis  $g_2(\hat{\epsilon})$  is near 3 (which proves a symmetric normal distribution). Excellent fitness is indicated and the regression parameter estimates are considered sufficiently reliable. ESAB has reached constantly better fitness than HYPERQUAD and therefore it can be concluded that estimates of the dissociation constants estimated by ESAB are more reliable. The individual pair of one dissociation constant estimated with ESAB and HYPERQUAD, in which the same points of titration curve were used, differ mostly on the second decimal place. The curve-fitting is significantly improved using the refinement of the group parameter  $L_0$ , the concentration of the titrated drug Eltrombopag.

The ESAB program minimizing residuals  $e_i = (V_{exp,i} - V_{calc,i})$  reaches 0.1 or 0.2 microliters and HYPERQUAD minimizing  $e_i = (pH_{exp,i} - pH_{calc,i})$  reaches SIGMA value about 1 or less, thus proving an excellent fit. It may be concluded that the reliability of the dissociation constants of Eltrombopag was proven even when group parameters  $L_0$ ,  $H_T$  were ill-conditioned in a model. Their determination is uncertain and might lead to false estimate of common parameters  $pK_a$  and therefore make the computational strategy important. These group parameters can have great influence on a systematic error in the estimated  $pK_a$  and they should be refined together with common parameters  $pK_a$ . External calibration of pH of the glass electrode cell performed during titration is sufficiently accurate. Comparing two computational approaches, the ESAB and the HYPERQUAD programs, ESAB led to a better fitness of the potentiometric titration curve. The goodness-



**Fig. 8** Dependence of the mixed dissociation constants of Eltrombopag on the square root of the ionic strength for five dissociation constants at 25 °C

of-fit proved sufficient reliability of parameter estimates for four dissociation constants of drug Eltrombopag at 25 °C.

## 5 Conclusion

Spectrophotometric and potentiometric pH-titration allowed the measurement of five dissociation constants of Eltrombopag, but low solubility at pH above 9 and below 5, at micromolar Eltrombopag concentrations, limits the estimation of the  $pK_a$  above pH = 10 and in potentiometry lower than 5.

(1) At neutral pH, Eltrombopag occurs in the sparingly soluble form  $LH_3$ , which can be protonated to form the soluble species  $LH_4^+$ . The species  $LH_3$  can also dissociate into the water soluble species  $L^{3-}$ . Acid–base titration of the triprotic molecule  $LH_3$  with KOH leads to a mixture of six species  $H_3O^+$ ,  $OH^-$ ,  $LH_3$ ,  $LH_2^-$ ,  $LH^{2-}$ ,  $L^{3-}$  and  $K^+$ . The graph of molar absorption coefficients of variously protonated species against wavelength shows that the spectra of species  $LH_3$  and  $LH_2^-$  are of only slightly different. The same is true for the chromophores  $LH^{2-}$  to  $LH_2^-$ , while protonation of chromophore  $LH_2^-$  to  $LH_3$  has greater influence on chromophores in Eltrombopag and results in considerable spectral change.

(2) We have proven that in the range of pH = 2–10 five dissociation constants can be reliably estimated from the spectra when concentration of Eltrombopag are less than  $10^{-4}$  mol·dm $^{-3}$ . Although the change of pH somewhat less affected changes in the chromophore, five mixed dissociation constants at an ionic strength  $I = 0.005$  mol·dm $^{-3}$  can be reliably determined with REACTLAB and SQUAD84 reaching the similar values with both programs. From a dependence on ionic strength the thermodynamic dissociation constants were estimated at 25 °C (Fig. 8):  $pK_{a1}^T = 2.69$ ,  $pK_{a2}^T = 6.97$ ,  $pK_{a3}^T = 7.13$ ,  $pK_{a4}^T = 7.65$ , and  $pK_{a5}^T = 8.30$

(3) Four dissociation constants of Eltrombopag in concentration of  $5 \mu\text{mol}\cdot\text{dm}^{-3}$  were determined by regression analysis of potentiometric titration curves without adjusting the ionic strength  $I = 0.005$  mol·dm $^{-3}$  and using ESAB and HYPERQUAD, reaching the

similar values with both programs  $pK_{a2} = 6.59(01)$ ,  $pK_{a3} = 7.56(04)$ ,  $pK_{a4} = 8.48(59)$ ,  $pK_{a5} = 9.29(34)$  at 25 °C (Table 3). The standard deviations in the last valid unit number are in the parentheses.

(4) Prediction of the dissociation constants of Eltrombopag was performed using the MARVIN program to specify protonation locations to give the values in Table 1. Comparing two predictive with two experimental techniques it may be concluded that the prediction programs often vary considerably in the estimation of  $pK_a$ . It was proven that the most reliable regression estimate of the dissociation constants comes from the reliable experimental data. In the case of close dissociation constants a higher degree of uncertainty in estimates should be expected and therefore usually two independent instrumental methods should be applied and the results calculated using several independent programs to compare.

## References

1. Manallack, D.T.: The  $pK(a)$  distribution of drugs: application to drug discovery. *Perspect. Medic. Chem.* **1**, 25–38 (2007)
2. Garnock-Jones, K.P., Keam, S.J.: Eltrombopag. *Drugs* **69**, 567–576 (2009). doi:[10.2165/00003495-200969050-00005](https://doi.org/10.2165/00003495-200969050-00005)
3. Danish, F.A., Koul, S.S., Subhani, F.R., Rabbani, A.E., Yasmin, S.: Considerations in the management of hepatitis C virus-related thrombocytopenia with Eltrombopag. *Saudi J. Gastroenterol* **16**, 51–56 (2010). doi:[10.4103/1319-3767.58772](https://doi.org/10.4103/1319-3767.58772)
4. Liao, C.Z., Nicklaus, M.C.: Comparison of nine programs predicting  $pK(a)$  values of pharmaceutical substances. *J. Chem. Inf. Model.* **49**, 2801–2812 (2009)
5. Szakacs, Z., Noszal, B.: Protonation microequilibrium treatment of polybasic compounds with any possible symmetry. *J. Math. Chem.* **26**, 139–155 (1999). doi:[10.1023/A:1019133927929](https://doi.org/10.1023/A:1019133927929)
6. Milletti, F., Storchi, L., Sforma, G., Cruciani, G.: New and original  $pK_a$  prediction method using grid molecular interaction fields. *J. Chem. Inf. Model.* **47**, 2172–2181 (2007)
7. Settimo, L., Bellman, K., Knegtel, R.A.: Comparison of the accuracy of experimental and predicted  $pK_a$  values of basic and acidic compounds. *Pharm. Res.* **31**, 1082–1095 (2014)
8. Tam, K.Y.: Multiwavelength spectrophotometric determination of acid dissociation constants. Part VI. Deconvolution of binary mixtures of ionizable compounds. *Anal. Lett.* **33**, 145–161 (2000)
9. Tam, K.Y., Quere, L.: Multiwavelength spectrophotometric resolution of the micro-equilibria of cetirizine. *Anal. Sci.* **17**, 1203–1208 (2001)
10. Tam, K.Y., Takacs-Novak, K.: Multiwavelength spectrophotometric determination of acid dissociation constants: a validation study. *Anal. Chim. Acta* **434**, 157–167 (2001)
11. Tam, K.Y.: Multiwavelength spectrophotometric resolution of the micro-equilibria of a triprotic amphoteric drug Methacycline. *Mikrochim. Acta* **136**, 91–97 (2001). doi:[10.1007/s006040170073](https://doi.org/10.1007/s006040170073)
12. Purdie, N., Tomson, M.B., Riemann, N.: The thermodynamics of ionization of polycarboxylic acids. *J. Solution Chem.* **1**, 465–476 (1972)
13. Meloun, M., Bordovská, S., Vrána, A.: The thermodynamic dissociation constants of the anticancer drugs camptothecine, 7-ethyl-10-hydroxycamptothecine, 10-hydroxycamptothecine and 7-ethylcamptothecine by the least-squares nonlinear regression of multiwavelength spectrophotometric pH-titration data. *Anal. Chim. Acta* **584**, 419–432 (2007). doi:[10.1016/j.aca.2006.11.049](https://doi.org/10.1016/j.aca.2006.11.049)
14. Hartley, F.R., Burgess, C., Alcock, R.M.: *Solution equilibria*. Ellis Horwood, Chichester (1980)
15. Havel, J., Meloun, M.: *Computational methods for the determination of formation constants, general computer programs for the determination of formation constants from various types of data*. Plenum Press, New York (1985)
16. Leggett, D.J., McBryde, W.A.E.: General computer program for the computation of stability constants from absorbance data. *Anal. Chem.* **47**, 1065–1070 (1975)
17. Kankare, J.J.: Computation of equilibrium constants for multicomponent systems from spectrophotometric data. *Anal. Chem.* **42**, 1322–1326 (1970). doi:[10.1021/ac60294a012](https://doi.org/10.1021/ac60294a012)
18. Meloun, M., Ferenčíková, Z., Kašťánek, M., Vrána, A.: Thermodynamic dissociation constants of Butorphanol and Zolpidem by the least-squares nonlinear regression of multiwavelength spectrophotometric pH-titration data. *J. Chem. Eng. Data* **56**, 1009–1019 (2011). doi:[10.1021/jc101000t](https://doi.org/10.1021/jc101000t)

19. Meloun, M., Ferenčíková, Z., Vrána, A.: The thermodynamic dissociation constants of Methotrexate by the nonlinear regression and factor analysis of multiwavelength spectrophotometric pH-titration data. *Central Eur. J. Chem.* **8**, 494–507 (2010). doi:[10.2478/s11532-010-0023-1](https://doi.org/10.2478/s11532-010-0023-1)
20. Meloun, M., Ferenčíková, Z., Javůrek, M.: Reliability of dissociation constants and resolution capability of SQUAD(84) and SPECFIT/32 in the regression of multiwavelength spectrophotometric pH-titration data. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **86**, 305–314 (2012). doi:[10.1016/j.saa.2011.10.041](https://doi.org/10.1016/j.saa.2011.10.041)
21. Meloun, M., Nečasová, V., Javůrek, M., Pekárek, T.: The dissociation constants of the cytostatic Bosutinib by nonlinear least-squares regression of multiwavelength spectrophotometric and potentiometric pH-titration data. *J. Pharm. Biomed. Anal.* **120**, 158–167 (2016). doi:[10.1016/j.jpba.2015.12.012](https://doi.org/10.1016/j.jpba.2015.12.012)
22. Pathare, B., Tambe, V., Patil, V.: A review on various analytical methods used in determination of dissociation constant. *Int. J. Pharm. Pharm. Sci.* **6**, 26–34 (2014)
23. Reijenga, J., Hoof, A.V., Loon, A.V., Teunissen, B.: Development of methods for the determination of  $pK_a$  values. *Anal. Chem. Insights* **8**, 53–71 (2013)
24. Hernández, J.A., Hernández, A.R., Urbina, E.M.C., Rodríguez, I.M.D.L.G., Manzanares, M.V., Medina-Vallejo, L.F.: New chemometric strategies in the spectrophotometric determination of  $pK_a$ . *Eur. J. Chem.* **5**, 1–5 (2014)
25. Meloun, M., Syrový, T., Bordovská, S., Vrána, A.: Reliability and uncertainty in the estimation of  $pK_a$  by least squares nonlinear regression analysis of multiwavelength spectrophotometric pH titration data. *Anal. Bioanal. Chem.* **387**, 941–955 (2007). doi:[10.1007/s00216-006-0993-1](https://doi.org/10.1007/s00216-006-0993-1)
26. Meloun, M., Bordovská, S., Syrový, T., Vrána, A.: Tutorial on a chemical model building by least-squares non-linear regression of multiwavelength spectrophotometric pH-titration data. *Anal. Chim. Acta* **580**, 107–121 (2006). doi:[10.1016/j.aca.2006.07.043](https://doi.org/10.1016/j.aca.2006.07.043)
27. Meloun, M., Bordovská, S.: Benchmarking and validating algorithms that estimate  $pK_a$  values of drugs based on their molecular structures. *Anal. Bioanal. Chem.* **389**, 1267–1281 (2007). doi:[10.1007/s00216-007-1502-x](https://doi.org/10.1007/s00216-007-1502-x)
28. Tam, K.Y., Hadley, M., Patterson, W.: Multiwavelength spectrophotometric determination of acid dissociation constants, Part IV. Water-insoluble pyridine derivatives. *Talanta* **49**, 539–546 (1999)
29. Leggett, D.J.: Numerical analysis of multicomponent spectra. *Anal. Chem.* **49**, 276–281 (1977)
30. Leggett, D.J., Kelly, S.L., Shiue, L.R., Wu, Y.T., Chang, D., Kadish, K.M.: A computational approach to the spectrophotometric determination of stability constants-II. Application to metalloporphyrin-axial ligand interactions in non-aqueous solvents. *Talanta* **30**, 579–586 (1983). doi:[10.1016/0039-9140\(83\)80136-2](https://doi.org/10.1016/0039-9140(83)80136-2)
31. Maeder, M., King, P.: Analysis of chemical processes, determination of the reaction mechanism and fitting of equilibrium and rate constants. In: Varmusa, K. (ed): *Chemometrics in Practical Applications*. Intech (2012). <https://www.intechopen.com/books/chemometrics-in-practical-applications/analysis-of-chemical-processes-determination-of-the-reaction-mechanism-and-fitting-of-equilibrium-an>. Accessed 1 Oct 2017
32. Gamp, H., Maeder, M., Meyer, C.J., Zuberhuhler, A.D.: Calculation of equilibrium constants from multiwavelength spectroscopic data—II: SPECFIT: two user-friendly programs in basic and standard FORTRAN 77. *Talanta* **32**, 257–264 (1985)
33. Meloun, M., Havel, J., Högfeldt, E.: *Computation of Solution Equilibria: A Guide to Methods in Potentiometry, Extraction, and Spectrophotometry*. Ellis Horwood Series in Analytical Chemistry. Ellis Horwood, Chichester (1988)
34. Japertas, P., Lanevskij, K., Sazonovas, A.: ACD/Percepta structure design engine: virtual enumeration and screening of physchem properties for  $10^{16}$  compounds in real time. *Abstr. Pap. Am. Chem. Soc.* **248** (2014) (Open Access). <http://www.acdlabs.com/company/events/eventinfo.php?id=140>
35. Garnock-Jones, K.P.: Eltrombopag: a review of its use in treatment-refractory chronic primary immune thrombocytopenia. *Drugs* **71**, 1333–1353 (2011). doi:[10.2165/11207390-000000000-00000](https://doi.org/10.2165/11207390-000000000-00000)
36. Balogh, G.T., Tarcsay, A., Keseru, G.M.: Comparative evaluation of  $pK_a$  prediction tools on a drug discovery dataset. *J. Pharm. Biomed. Anal.* **67–68**, 63–70 (2012). doi:[10.1016/j.jpba.2012.04.021](https://doi.org/10.1016/j.jpba.2012.04.021)
37. ten Brink, T., Exner, T.E.:  $pK(a)$  based protonation states and microspecies for protein–ligand docking. *J. Comput. Aid Mol. Des.* **24**(11), 935–942 (2010)
38. SPARC On-line Calculator. <http://ibmlc2.chem.uga.edu/sparc/index.cfm> (2011)
39. Hilal, S.H., Carreira, L.A., Karickhoff, S.W.: Verification and Validation of the SPARC Model. *EPA* **600**(R-03/033) (2003)
40. Meloun, M., Javůrek, M., Bartoš, M.: Complexation equilibria of some sulfoazoxines. 8. Complexes of Snaozox with copper(II), lead(II) and cadmium(II) ions evaluated by regression-analysis of potentiometric data. *Analyst* **113**, 1357–1364 (1988). doi:[10.1039/An9881301357](https://doi.org/10.1039/An9881301357)
41. Kromann, J.C., Larsen, F., Moustafa, H., Jensen, J.H.: Prediction of  $pK_a$  values using the PM6 semiempirical method. *PeerJ* **4**, e2335 (2016). doi:[10.7717/peerj.2335](https://doi.org/10.7717/peerj.2335)

42. Balogh, G.T., Gyarmati, B., Nagy, B., Molnar, L., Keseru, G.M.: Comparative evaluation of in silico  $pK_a$  prediction tools on the gold standard dataset. *QSAR Comb. Sci.* **28**, 1148–1155 (2009)
43. Japertas, P., Sazonovas, A., Lanevskij, K.: ACD/Percepta portal: crowdsourcing in medicinal chemistry projects. *Abstr. Pap. Am. Chem. Soc.* **248** (2014) (Open Access). <http://www.acdlabs.com/products/percepta/portal/>
44. da Cunha, M.G., Francob, G.C.N., Franchin, M., Beutler, J.A., de Alencar, S.M., Ikegaki, M., Rosalen, P.L.: Prediction of pharmacokinetic and toxicological parameters of a 4-phenylcoumarin isolated from geopropolis: in silico and in vitro approaches. *Toxicol. Lett.* **263**, 6–10 (2016). doi:[10.1016/j.toxlet.2016.10.010](https://doi.org/10.1016/j.toxlet.2016.10.010)
45. Meloun, M., Syrový, T., Ghasemi, J.: Recent progress in the  $pK_a$  estimation of druglike molecules by the nonlinear regression of multiwavelength spectrophotometric ph-titration data. *SRX Pharmacol.* (2010). doi:[10.3814/2010/481497](https://doi.org/10.3814/2010/481497)
46. Rigano, C., Grasso, M., Sammartano, S.: Computer-analysis of equilibrium data in solution—a compact least-squares computer-program for acid-base titrations. *Ann. Chim. Rome* **74**(7–8), 532–537 (1984)
47. De Stefano, C., Princi, P., Rigano, C., Sammartano, S.: Computer analysis of equilibrium data in solution ESAB2 M: an improved version of the ESAB program. *Ann. Chim. Rome* **77**, 643–675 (1987)
48. Capone, S., De Robertis, A., De Stefano, C., Sammartano, S., Scarcella, R.: Ionic-strength dependence of formation-constants. 10. Proton activity-coefficients at various temperatures and ionic strengths and their use in the study of complex equilibria. *Talanta* **34**, 593–598 (1987)
49. Gans, P., Sabatini, A., Vacca, A.: Investigation of equilibria in solution. Determination of equilibrium constants with the HYPERQUAD suite of programs. *Talanta* **43**(10), 1739–1753 (1996). doi:[10.1016/0039-9140\(96\)01958-3](https://doi.org/10.1016/0039-9140(96)01958-3)
50. Meloun, M., Militký, J., Forina, M.: *Chemometrics for Analytical Chemistry, Volume 2: PC-Aided Regression and Related Methods*. Chemometrics for Analytical Chemistry. Ellis Horwood, Chichester (1994)
51. Meloun, M., Čapek, J., Mikšík, P., Brereton, R.G.: Critical comparison of methods predicting the number of components in spectroscopic data. *Anal. Chim. Acta* **423**, 51–68 (2000). doi:[10.1016/S0003-2670\(00\)01100-4](https://doi.org/10.1016/S0003-2670(00)01100-4)
52. Meloun, M., Říha, V., Žáček, J.: Piston microburette for dosing aggressive liquids. *Chem. Listy* **82**, 765–767 (1988)
53. May, P.M., Williams, D.R., Linder, P.W., Torrington, R.G.: The use of glass electrodes for the determination of formation-constants. 1. A definitive method for calibration. *Talanta* **29**, 249–256 (1982)
54. Meloun, M., Javůrek, M., Högfeldt, E.: A comparison of some linear transformations in the evaluation of spectrophotometric data. *Chem. Scr.* **28**, 323–329 (1988)
55. Manchester, J., Walkup, G., Rivin, O., You, Z.P.: Evaluation of  $pK_a$  estimation methods on 211 drug like compounds. *J. Chem. Inf. Model.* **50**, 565–571 (2010). doi:[10.1021/ci100019p](https://doi.org/10.1021/ci100019p)
56. Albert, A., Serjeant, E.P.: *The determination of ionization constants—a laboratory manual*. Chapman and Hall, New York (1984)
57. Meloun, M., Militký, J., Forina, M.: *Chemometrics for Analytical Chemistry, Volume 1: PC-Aided Statistical Data Analysis*. Chemometrics for Analytical Chemistry. Ellis Horwood, Chichester (1992)