



Automated R workflow for Batman peak fitting using unified equation and stochastic modeling

Arash Mirzahosseini^{a,b},^{*}, Annamária Sepsey^c, Ali Mhammad^{a,b}, Gergely Dombi^{a,b}, Simon Horváth^d, Eliza Tóth^{a,b,d}, Gábor Németh^d, Oliver Trapp^e, Attila Felinger^{c,f}, Gergő Tóth^{a,b}

^a Department of Pharmaceutical Chemistry, Semmelweis University, Budapest, Hungary

^b Center for Pharmacology and Drug Research & Development, Semmelweis University, Budapest, Hungary

^c Institute of Bioanalysis, Medical School, University of Pécs, Pécs, Hungary

^d Drug Substance Analytical Development Division, Egis Pharmaceuticals PLC, Budapest, Hungary

^e Department of Chemistry, Ludwig-Maximilians-Universität München, München, Germany

^f Department of Analytical and Environmental Chemistry and Szentágothai Research Center, University of Pécs, Pécs, Hungary

ARTICLE INFO

Keywords:

Batman peak
Enantiomers
Interconversion
Chiral separation
Rate constant

ABSTRACT

Many drugs are chiral, and their enantiomers often display distinct pharmacological activities. Accurate determination of enantiomerization kinetics during chromatographic separations is therefore of critical importance. Batman peaks (plateau regions arising between partially resolved enantiomer peaks) carry valuable information about on-column interconversion dynamics. We developed an automated R framework that integrates both the unified equation and a stochastic model to extract kinetic and thermodynamic constants from chromatographic data. The workflow applies robust optimization algorithms to enhance flexibility, accelerate fitting, and improve accuracy relative to manual approaches. Quetiapine chromatograms were analyzed over a wide range of flow rates (0.1 mL/min to 3.0 mL/min) and temperatures (10 °C to 40 °C). Both the unified equation and the stochastic model produced consistent rate constants, which were in good agreement with values obtained from optical rotation experiments. Notably, the stochastic model remained applicable even under extreme conditions where peaks coalesced. Linearized Eyring–Polányi analysis afforded the determination of thermodynamic parameters. The presented workflow provides a robust, flexible, and customizable platform for determining enantiomerization kinetics from Batman peaks. By combining the unified equation with stochastic modeling, it delivers reliable results across diverse chromatographic conditions. This framework is readily adaptable for routine analysis of dynamic enantiomerization in pharmaceutical and analytical chemistry applications.

1. Introduction

While enantiomers occur naturally in many biomolecules, most synthetic drugs are likewise chiral, a property crucial to their biological activity. The separation of chiral molecules is especially critical for biologically active substances, since enantiomers may have very distinct effects, the distomer may be inactive or even harmful [1]. So-called *Batman peaks* have long been observed in chromatographic data, though they are often misinterpreted or overlooked. The term Batman peak (appearing in the literature as early as 1994 [2]) arose informally due to the characteristic shape of these coalescing peaks. Although the phenomenon itself had been described earlier under other names (e.g., Schurig referred to this phenomenon as “enantiomerization” or “peak coalescence of the second kind” [3]), the “Batman” terminology

became popular because it provides an intuitive visual descriptor for this distinctive elution profile. During the injection of a racemic sample, these peaks appear with a plateau or a saddle between the two adjacent peaks because of the coexistence of both enantiomers: molecules that have interconverted during the chromatographic run. If interconversion is very fast relative to the separation timescale, only a single peak is observed. Conversely, when the interconversion rate is slow enough, the enantiomers can be separated on a chiral stationary phase. If the timescale of isomerization is comparable to the timescale of drug clearance through the body, this phenomenon will have pharmacokinetic relevance [4]. Importantly, the appearance of Batman peaks requires an intermediate regime in which interconversion is neither fully suppressed nor too rapid, and their prominence is often enhanced at

^{*} Correspondence to: Budapest H-1092, Högyes Endre utca 9., Hungary.
E-mail address: mirzahosseini.arash@semmelweis.hu (A. Mirzahosseini).

elevated temperatures. This phenomenon is most commonly observed in molecules exhibiting slow tautomerism or moderate configurational stability, such as atropisomers (conformational enantio- or diastereomers as a result of hindered bond rotation). Several comprehensive reviews have summarized on-column enantiomerization [5–11]. When Batman peaks are discernible the determination of interconversion rate constants from chromatographic data is possible, making it a faster alternative to other methods (e.g., optical rotation, circular dichroism, fluorescence resonance energy transfer (FRET)).

Since the early 1980s, Schurig and co-workers [12], and later Trapp and colleagues [6], have laid the theoretical foundations of such separations. They introduced approximation functions for calculating enantiomerization rate constants directly from peak parameters, followed by a unified equation [13] that relates chromatographic parameters — such as retention times, plate numbers, half-height peak widths, and plateau heights — to the rate constant. Software implementations have also been developed, most notably *DCXplorer* [14], which remains available. Sepsey et al. (2018) [15] extended this field by presenting a stochastic modeling approach capable of fitting merged peaks and yielding mechanistically meaningful kinetic constants. The first statistical description of on-column enantiomerization, however, dates back to the works of Eyring, Giddings [16], and Keller [17], who derived the relevant molecular distribution functions.

The present work aims to determine enantiomerization rate constants from Batman peaks using both the unified equation and the stochastic model within a single, automated, and customizable workflow. We apply this framework to the experimental data set studied by Sepsey et al. (2018), enabling direct comparison with their optical rotation experiments. The functions, implemented in R and available for customization, yield consistent results across the different methods, demonstrating their suitability for routine chromatographic analysis. Analyses were carried out on various HPLC columns to assess how the chiral selector influences enantiomer interconversion.

2. Theory

2.1. Peak picking and unified equation

The algorithm begins by receiving a folder path and reading the chromatograms provided as comma-separated values (CSV) files. If the raw chromatograms are stored in another format, the files can be converted to CSV with the *chromConverter* package [18], after the user had selected the optimal detector signal. The CSV file for each chromatographic run must contain at least two columns; the first two columns must correspond to time (in minutes) and detector intensity, as these are automatically selected for analysis. The filename encodes the metadata required by the workflow using the pattern `[compound]_[column]_[flow]_[temperature].csv`. The naming convention also allows for number of parallel to be appended after temperature (`_[parallel]`), if certain chromatographic runs were repeated. This convention allows the parser to automatically assign compound identity, stationary phase, flow rate, and temperature to each trace prior to peak picking and subsequent modeling. The CSV files are imported through a helper function that automatically determines the encoding, column separator, decimal format, and the presence or absence of a header row.

The algorithm then proceeds to identify the void time peak (if present in the chromatogram) and the characteristic Batman peak (peak A and peak B, the first and second eluted species), or a merged peak if enantiomeric signals are not clearly resolved. Prior to peak identification the algorithm applies a cubic baseline fitting and correction followed by a regular pattern recognition (of increase or decrease in time series) function from the *pracma* package [19], using a user-defined signal-to-noise threshold. In cases where a merged peak is detected, the calculation of the unified equation is bypassed. Once the relevant peaks are identified, the algorithm proceeds to determine

their retention time (t_i), full width at half maximum (w_i) and the concomitant standard deviation (s_i), peak height (h_i), and plateau relative height (h_p) directly from the chromatogram ($i \in \{A, B\}$). When peaks are not detected at the expected retention times, the algorithm can ingest a user-supplied peak-table guide, which raises the peak-picking success rate to 100%. Alternatively, if the void time marker peak is absent from the chromatograms, the user may provide a list of void times for each run; the algorithm then anchors retention time alignment to these values.

The unified equation is as follows [14].

$$k_{ue,f} = -\frac{1}{t_A} \left[\log \left(\frac{100(1 - A_0) + A_0 \left(100 - h_p \left(1 + \sqrt{\frac{2}{\pi N}} \right) \right)}{t_B - t_A} \right) - \log \left((1 - A_0) \left(\frac{h_p e^{-\frac{(t_B - t_A)^2}{2s_B^2}} - 100 e^{-\frac{(t_B - t_A)^2}{8s_B^2}}}{s_B \sqrt{2\pi}} + \frac{100}{t_B - t_A} \right) - A_0 \left(\frac{100 e^{-\frac{(t_B - t_A)^2}{8s_A^2}} - h_p}{s_A \sqrt{2\pi}} + \frac{h_p \left(1 + \sqrt{\frac{2}{\pi N}} \right) - 100}{t_B - t_A} \right) \right) \right] \quad (1)$$

where N is the number of theoretical plates, and A_0 is the fraction of enantiomer A (first-eluting species) in the injected sample (this quantity must be known *a priori*). Eq. (1) applies to cases when peak A is taller than peak B. In the opposite case, the unified equation takes on a slightly different form, which the algorithm selects automatically:

$$k_{ue,f} = -\frac{1}{t_A} \left[\log \left(\frac{100A_0 + (1 - A_0) \left(100 - h_p \left(1 - \sqrt{\frac{2}{\pi N}} \right) \right)}{t_B - t_A} \right) - \log \left(A_0 \left(\frac{h_p e^{-\frac{(t_B - t_A)^2}{2s_A^2}} - 100 e^{-\frac{(t_B - t_A)^2}{8s_A^2}}}{s_A \sqrt{2\pi}} + \frac{100}{t_B - t_A} \right) - (1 - A_0) \left(\frac{100 e^{-\frac{(t_B - t_A)^2}{8s_B^2}} - h_p}{s_B \sqrt{2\pi}} + \frac{h_p \left(1 - \sqrt{\frac{2}{\pi N}} \right) - 100}{t_B - t_A} \right) \right) \right] \quad (2)$$

The rate constant for the reverse direction is calculated using different expressions depending on whether the principle of microscopic reversibility is assumed to hold [13]. By default, microscopic reversibility is not enforced, i.e., an enantioselective stationary phase is allowed to produce forward and reverse rate constants independently. The corresponding equations for the reverse rate constant, for both cases (with and without microscopic reversibility), are provided below.

$$k_{ue,r} = k_{ue,f} \frac{A_\infty}{B_\infty} \frac{t_A}{t_B} \quad k_{ue,r} = k_{ue,f} \frac{A_\infty}{B_\infty} \quad (3)$$

Finally, the Eyring–Polányi linearization of k_{ue} vs T (absolute temperature) affords ΔH° and ΔS° values, which in turn provide ΔG° . The Eyring–Polányi regression is performed with both the forward and reverse kinetic constants. The Gibbs free energy can also be calculated directly from the unified equation as follows:

$$\Delta G^\circ = -RT \log \left(\frac{k_{ue} h}{\kappa k_B T} \right) \quad (4)$$

where R , h , k_B , and κ are the ideal gas constant, Planck's constant, Boltzmann constant, and transmission coefficient (≈ 0.5), respectively. Table 1 provides an overview of the chromatographic parameters estimated for use in the unified-equation treatment.

Table 1
List of parameters used for unified equation.

Quantity	Unit	Name	Estimation
t_i	s	retention time	estimated from peak mode
s_i	s	standard deviation of peak	estimated from FWHM
h_p		plateau relative height (%)	intensity at mid-point of 2 peaks relative to tallest peak intensity
A_0		fraction of enantiomer A in injected sample	provided <i>a priori</i>
$\frac{A_{\infty}}{B_{\infty}}$		ratio of peak area under curves	estimated from Batman peak area under curve split at mid-point
N		number of theoretical plates	obtained by moment analysis of the 2 peaks, then averaged
$k_{ue,f/r}$	s ⁻¹	kinetic constant (forward/reverse)	obtained from the unified equation

2.2. Stochastic model

The characteristic function approach to the stochastic theory of chromatography [20], [21] was used to model the separation process as a composite Poisson process with exponentially distributed sojourn (residence) times. The characteristic function, as the Fourier transform of the elution profile, contains all necessary information about the peaks:

$$\Phi(\omega) = \exp\left(n\left(\frac{1}{1 - i\tau\omega} - 1\right)\right) \quad (5)$$

where n is the average number of adsorption–desorption steps, τ is the average time spent by a molecule bound to the stationary phase (mean sojourn time), i is the imaginary unit, and ω is an auxiliary real variable (frequency). The parameters n and τ jointly determine both the retention time and the overall shape of the chromatographic peak. Both parameters influence the peak position; however, the peak symmetry is governed primarily by τ : larger τ values increase residence-time variability and lead to broader, more right-skewed peaks, whereas changes in n mainly scale the degree of compression or dispersion without strongly altering skewness. Enantiomers A and B injected with relative amounts α and $\beta = 1 - \alpha$ will each be in 3 possible states at end of run. Note that α corresponds directly to A_0 in the unified equation framework. For racemic mixtures $A_0 = \alpha = 0.5$, whereas for scalemic samples this value should be set to the actual enantiomeric ratio present in the injected mixture. For species A these 3 probabilities are:

$$\begin{aligned} P'_{AA} &= \mathbb{P}(\text{species}_{t=\infty} = A \mid \text{species}_{t=0} = A \wedge \text{no interconversion}), \\ P_{AA} &= \mathbb{P}(\text{species}_{t=\infty} = A \mid \text{species}_{t=0} = A \wedge \text{interconversion}), \\ P_{AB} &= \mathbb{P}(\text{species}_{t=\infty} = B \mid \text{species}_{t=0} = A \wedge \text{interconversion}). \end{aligned} \quad (6)$$

Analogously the same can be defined for starting as B: P'_{BB} , P_{BB} , P_{BA} . By taking the appropriate weighted probabilities according to the law of total probability, we arrive at the probabilities of observing the enantiomers at the end of the run:

$$\begin{aligned} P_A &= \mathbb{P}(\text{species}_{t=\infty} = A) = \alpha(P'_{AA} + P_{AA}) + \beta P_{BA}, \\ P_B &= \mathbb{P}(\text{species}_{t=\infty} = B) = \beta(P'_{BB} + P_{BB}) + \alpha P_{AB}. \end{aligned} \quad (7)$$

The peak profile of the Batman peaks can be obtained by the convolution of the elution curves resulting from Eq. (5) with probability densities of Eq. (7). The conditional probabilities in Eq. (6) are given by the Keller-Giddings distribution as follows:

$$\begin{aligned} P_{AA}(x) &= \sqrt{\frac{ab(1-x)}{x}} \exp(-a(1-x) - bx) I_1\left(\sqrt{4abx(1-x)}\right), \\ P_{AB}(x) &= a \exp(-a(1-x) - bx) I_0\left(\sqrt{4abx(1-x)}\right), \\ P'_{AA} &= \delta_A \exp(-a). \end{aligned} \quad (8)$$

where x is the fraction of time that a molecule spends as A, I_i the modified Bessel function of the first kind and i th order, δ_i is the Dirac delta, and a , b are the expected number of interconversions from A to B and vice versa. The same distributions can be defined for starting state B by exchanging the roles of a with b , and x with $1 - x$. In the context of on-column interconversion, the parameters a and b control the extent to which the two enantiomer peaks overlap and form a

Batman profile. As a and b increase, a progressively larger fraction of molecules switches enantiomeric form while migrating between the two peak maxima. When either parameter exceeds roughly 5, the probability of interconversion becomes so high that the characteristic saddle between the peaks nearly disappears, leading to near-complete peak coalescence and the loss of a resolvable Batman structure.

It should be noted that the original weighted sum of probability densities outlined by Keller and Giddings (analogous to Eq. (7)) contains an inaccuracy. The proper approach is to multiply the probabilities corresponding to a given initial state by its concomitant injected fraction (α or β). Fig. 1 illustrates this with two simulated Batman peaks: one generated with equal initial enantiomer ratios, and the other with the B enantiomer initially more abundant, while both cases were modeled with equal numbers of forward and reverse interconversions (a and b). The overall probabilities of observing species A or B at the end of the run are shown as solid red and blue lines, respectively, whereas the contributing conditional probabilities P_{AA} , P_{AB} , P_{BB} , P_{BA} are displayed as dashed or dotted lines. When $\alpha = 0.5$, the results are indistinguishable. However, when $\alpha = 0.2$, striking differences emerge in the probability densities, even though the change in the elution profile is minimal. In the original Keller–Giddings formulation, P_{AB} (species starting in state A and ending in B) is incorrectly assigned as the most dominant probability to the center of the peak, which is physically implausible. With the corrected formula, the dominant contribution arises from P_{BA} , consistent with the initial enantiomeric ratio. Interestingly, despite the B enantiomer being predominant at the start of the run, the final plateau region of the Batman peak is dominated by species in the A state.

Table 2 provides an overview of the parameters estimated for use in the stochastic model treatment.

2.3. Fitting algorithm

For fitting the Batman peaks, the convolution profile was simulated using initial parameter values estimated from peak picking results using the characteristic function moments:

$$\tau_i = \frac{\text{var}_i}{2t'_i}, \quad n_i = \frac{t'_i}{\tau_i}. \quad (9)$$

where t'_i is the reduced retention time, and var_i is the variance of the peak profile, which are approximated by the peak apex distances and $\text{var}_i = \frac{w_i^2}{8 \log 2}$, respectively. When initial parameter estimates were not available from peak-picking for a given temperature/flow rate condition, starting values were extrapolated from other chromatograms recorded at the same temperature, leveraging the facts that sojourn times are, in principle, independent of flow rate and that n varies approximately linearly with flow rate. If no suitable chromatograms existed at that temperature, initial estimates were taken from the nearest lower temperature, ensuring continuity and physical plausibility of the starting parameters.

For fitting the Batman peaks with the stochastic model, the chromatograms were first deconvoluted from the peak of the void time

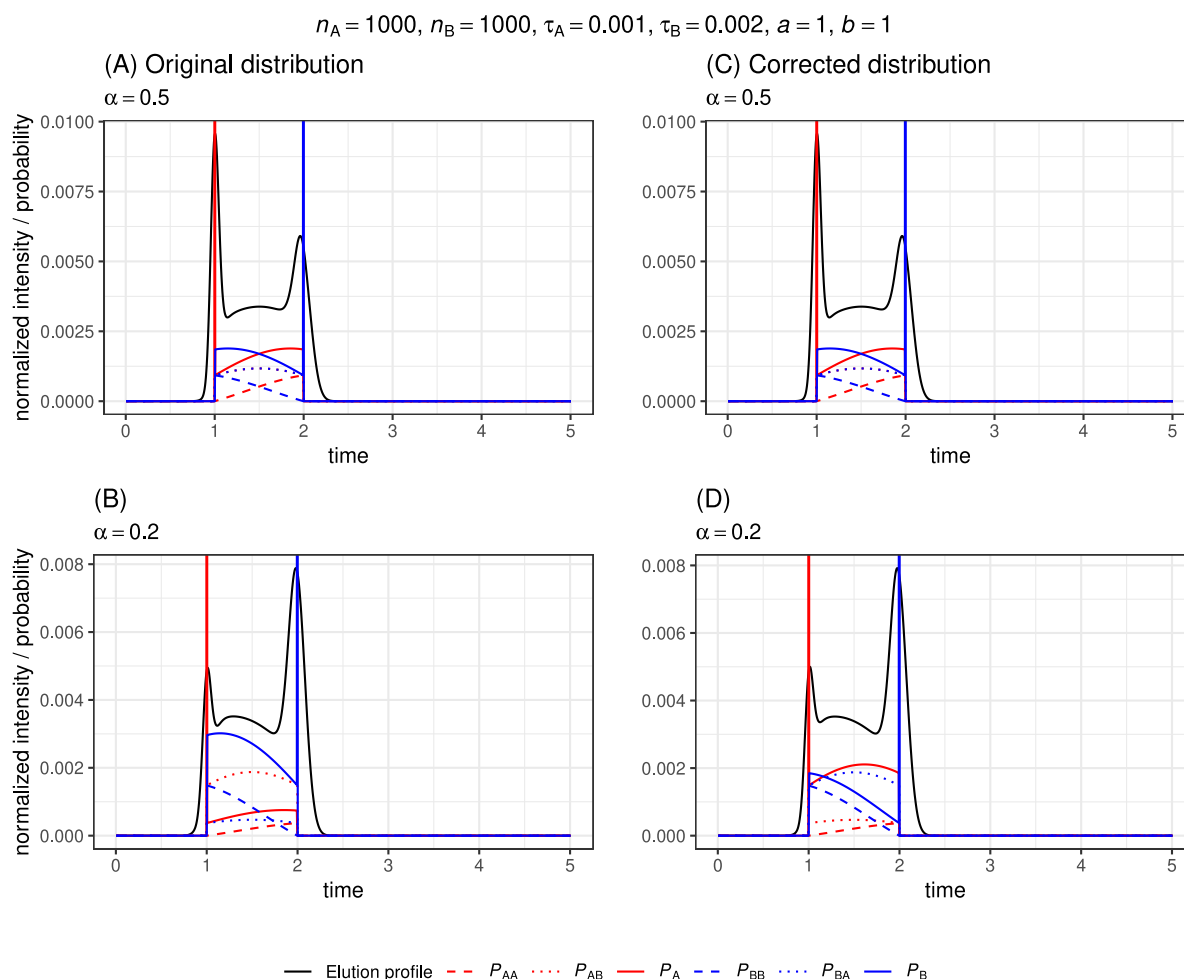


Fig. 1. Simulated Batman peaks with corrected probability weighting. Two cases are shown: equal initial enantiomer ratios in subplots (A) and (C) ($\alpha = 0.5$) and an excess of the B enantiomer in the injected sample (subplots (B) and (D), $\alpha = 0.2$). The left column (subplots (A) and (B)) shows stochastic modeling with the originally reported probability weighting, while the right column (subplots (C) and (D)) show the effect of corrected probability weighting reported in this work. Solid black lines show simulated chromatograms (using parameter values depicted in top of figure) with normalized intensity scale. Solid red and blue lines denote the overall probabilities of observing species A and B at the end of the run, respectively, while dashed and dotted lines indicate the contributing conditional probabilities. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

List of parameters used for stochastic modeling.

Quantity	Unit	Name	Estimation
n		number of adsorption–desorption steps	initially estimated from peak moments, then optimized
τ	s	sojourn time	initially estimated from peak moments, then optimized
ω	s^{-1}	frequency (Fourier domain)	sampled via simulation based on run time and number of points
α, β		ratio of enantiomers in injected sample	provided <i>a priori</i>
x		fraction of time a molecule spends as A	sampled via simulation on entire probability space
a, b		number of forward/reverse interconversions	optimized parameters

marker. This was achieved by fitting an exponentially modified Gaussian (EMG) function [22] (performed in log space to avoid box constraints) to the peak of the void time marker and subsequently subtracting the fitted contribution from the chromatogram. The fitting proceeded on the deconvoluted chromatogram in two stages: first, a differential evolutionary (DE) genetic algorithm [23] was applied for 200 iterations to refine the parameters toward the optimal region; this was followed by a Levenberg–Marquardt (LM) algorithm [24] to achieve precise convergence. Both the number of DE iterations and the maximum number of LM steps can be modified by the user within the function. This two-stage strategy was chosen heuristically: while LM achieves very tight convergence, it is highly sensitive to initial conditions and typically fails when started far from the optimum. In contrast,

DE converges reliably even from a cold start, albeit more slowly, and thus its progress is printed during computation for transparency. Because LM does not support parameter bounds whereas DE requires explicit lower and upper limits, sensible constraints are imposed during the DE phase: namely, half to twice the initial estimates for n and τ , and fixed bounds of 0 to 10 for the interconversion parameters (a and b). The upper limit for a and b , which represent the maximum number of forward and reverse interconversions, can be adjusted through a function argument in cases where more extensive interconversion is observed.

Throughout both optimization stages, the current parameter values (initial guesses and iteration-wise updates) were used to generate the characteristic function, which was subsequently transformed into

the corresponding peak profiles via the inverse fast Fourier transform `ifft()`. These simulated profiles were then convolved with the probability-weighted state distributions using `convolve()` (which is also carried out in Fourier space). Together, this hybrid approach provides robust global exploration via DE and precise local refinement via LM, yielding reliable convergence across a wide range of chromatographic conditions.

Finally, the linear regression of a or b versus t_A or t_B yields the kinetic rate constants in the forward and reverse directions, respectively [15]. This follows directly from the definitions of the interconversion counts, $a = k_f t$ and $b = k_r t$; i.e., the rate constant of interconversion corresponds to the slope of the fitted line when plotting the number of forward or reverse interconversions against time. Although, in principle, the void time could be used as the characteristic timescale, under the assumption that no interconversion occurs on the stationary phase, in practice this assumption does not hold for chiral stationary phases. Interconversion takes place in both the mobile and stationary phases, and the chromatographic residence time reflects both contributions. For this reason, the regression is performed using the retention times rather than the column void time, providing a more physically meaningful estimate of the apparent kinetic rate constants. It is noteworthy that the kinetic rate constants obtained from both the stochastic model and the unified equation represent apparent rate constants — weighted means of the individual rate constants in the mobile phase and in the chiral stationary phase, reflecting the combined dynamics of both environments during chromatographic separation.

3. Methods

3.1. Chemicals and HPLC measurements

Quetiapine fumarate as a raw material came from EGIS Pharmaceuticals PLC (Budapest, Hungary). 1,3,5-tri-*tert*-butylbenzene, HPLC grade acetonitrile and diethylamine were ordered from Merck (Merck KGaA, Darmstadt, Germany).

The chromatograms of quetiapine, a commercially available drug that undergoes interconversion, were analyzed. The racemic mixture was injected together with 1,3,5-tri-*tert*-butylbenzene as a non-retained reference component in polysaccharide column. Chromatographic separations were carried out using an Agilent 1100 chromatographic system consisting of an inline degasser (G1322A), quaternary pump (G1311A), automatic injector (G1329A) paired with sample thermostat (G1330A), column thermostat (G1316A) and a diode array detector (G1315A) controlled by Agilent Chemstation B04.03-SP2 software (Agilent Technologies, Waldbronn, Germany). Polysaccharide-type chiral column, like Lux Amylose-1 (based on amylose tris(3,5-dimethylphenylcarbamate)), Lux i-Amylose-1 (based on amylose tris(3,5-dimethylphenylcarbamate)), Lux Cellulose-2 (based on cellulose tris(3-chloro-4-methylphenyl carbamate)) and Lux Cellulose-4 (cellulose tris(4-chloro-3-methylphenyl carbamate)) with identical dimensions (150 × 4.6 mm; particle size: 5 μm) were the products of Phenomenex (Torrance, CA, USA). Lux Amylose-1 and Lux i-Amylose-1 contain the same chiral selector however, the process for retaining the selector on the surface of the silica differs. In the first case, the chiral selector is coated onto the surface of porous silica, while in the latter, it is covalently attached to it. The mobile phase consisted of acetonitrile containing 0.1% diethylamine. Samples were prepared at concentrations of 1.00 mg mL⁻¹ quetiapine and 0.50 mg mL⁻¹ 1,3,5-tri-*tert*-butylbenzene in the mobile phase, and 10 μL was injected for each run. UV detection was performed at 254 nm using a diode-array detector, as this non-chiroptical method provides identical responses for both enantiomers.

3.2. Data analysis and code script

All analyses were performed using custom scripts written in R version 4.5.1 [25] and developed in RStudio Version 2025.09.0+387 (Posit, Boston, MA, USA). The full code is openly available on GitHub: https://github.com/mirzahassemi-arash-semmelweis/Batman_chromat. The complete, fully commented code is also provided in the Supplementary Material, along with an additional script that guides the user through running the algorithm step by step, including practical tips and comments. Furthermore, representative sample data from the manuscript are supplied so that key analyses can be reproduced directly.

4. Results

4.1. Validation of the algorithm

Quetiapine chromatograms performed on Lux Cellulose-4 chiral stationary phase were recorded at flow rates ranging from 0.1 mL/min to 3.0 mL/min and at column temperatures between 10 °C to 40 °C. The schematic structure of quetiapine and the employed chiral stationary phases are depicted in Fig. 2. The resulting Batman peaks were analyzed with a custom R workflow, which enabled the determination of both kinetic and thermodynamic constants. Two representative chromatograms are presented in Fig. 3. When the Batman peaks could be clearly distinguished, the unified equation and the stochastic model produced closely matching values. In contrast, when the peaks merged into a single one, the stochastic model could still be fitted successfully by estimating near-optimal initial parameters from chromatograms measured at the same temperature.

After the exclusion of outliers, a comparison of the kinetic constants obtained in this work is provided in Table 3. The table also includes values calculated with the unified equation (employed by DCXplorer), as well as results from the stochastic fitting approach reported by Sepsey et al. (2018). For benchmarking, the optical rotation data from Sepsey et al. (2018) are listed alongside. The linearized Eyring–Polányi plot (Fig. 4) reveals a clear linear relationship between the transformed rate constants and temperature. From these linear fits, the relevant thermodynamic parameters were extracted, which in turn allowed the Gibbs free energy to be calculated at each investigated temperature. The standard enthalpy and entropy values for the various methods were as follows: optical rotation $\Delta H^\circ = 91.7 \pm 0.9 \text{ kJ mol}^{-1}$, $\Delta S^\circ = 0.12 \pm 0.03 \text{ kJ mol}^{-1} \text{ K}^{-1}$; unified equation $\Delta H^\circ = 67.2 \pm 5.2 \text{ kJ mol}^{-1}$, $\Delta S^\circ = -0.71 \pm 0.18 \text{ kJ mol}^{-1} \text{ K}^{-1}$; stochastic modeling $\Delta H^\circ = 72.2 \pm 4.1 \text{ kJ mol}^{-1}$, $\Delta S^\circ = -0.05 \pm 0.01 \text{ kJ mol}^{-1} \text{ K}^{-1}$.

4.2. Ablation study

An ablation study was conducted to identify essential algorithm elements and determine the minimal number of temperature and flow-rate conditions required to robustly estimate the stochastic-rate constant k and the Eyring–Polányi parameters. We systematically removed components of the workflow one at a time — baseline correction, exponentially modified Gaussian (EMG) deconvolution of the peak of the void time marker, and the optimization stages (Levenberg–Marquardt, LM; and evolutionary/genetic, EG) — to assess their impact on fit quality and parameter stability. Baseline correction and EMG deconvolution exerted only minor effects on the estimated parameters, whereas omitting the LM refinement yielded noticeably less accurate fits. Crucially, the EG stage proved essential for convergence; without it, the model failed to fit altogether. In addition to testing the impact of removing specific workflow components, we also performed a systematic data reduction experiment. Specifically, we progressively left out observations following a grid defined by decreasing temperature and flow-rate abundance, thereby probing the interaction between these two factors while simultaneously monitoring model performance. This

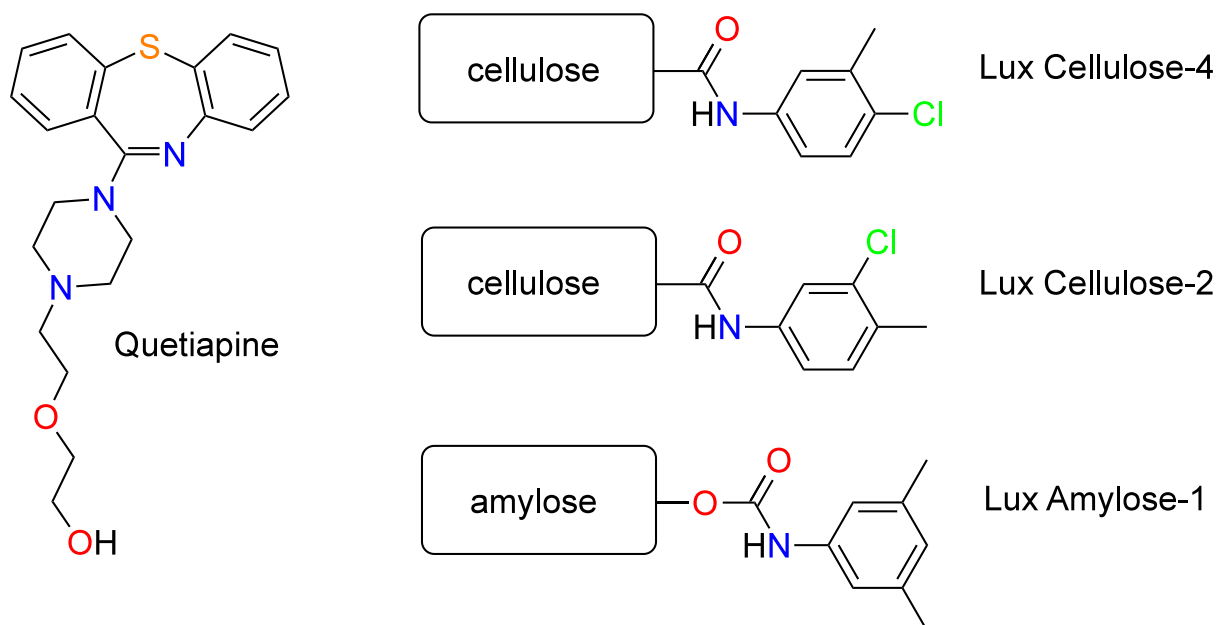


Fig. 2. The structure of quetiapine (left), which proved to be an appropriate model compound to study the phenomenon, as the seven-membered ring can give rise to atropisomers if the ring-flip is hindered. On the right the schematic structures of the examined chiral stationary phases are depicted. Lux Amylose-1 and immobilized Lux i-Amylose-1 have the same chemical modifications.

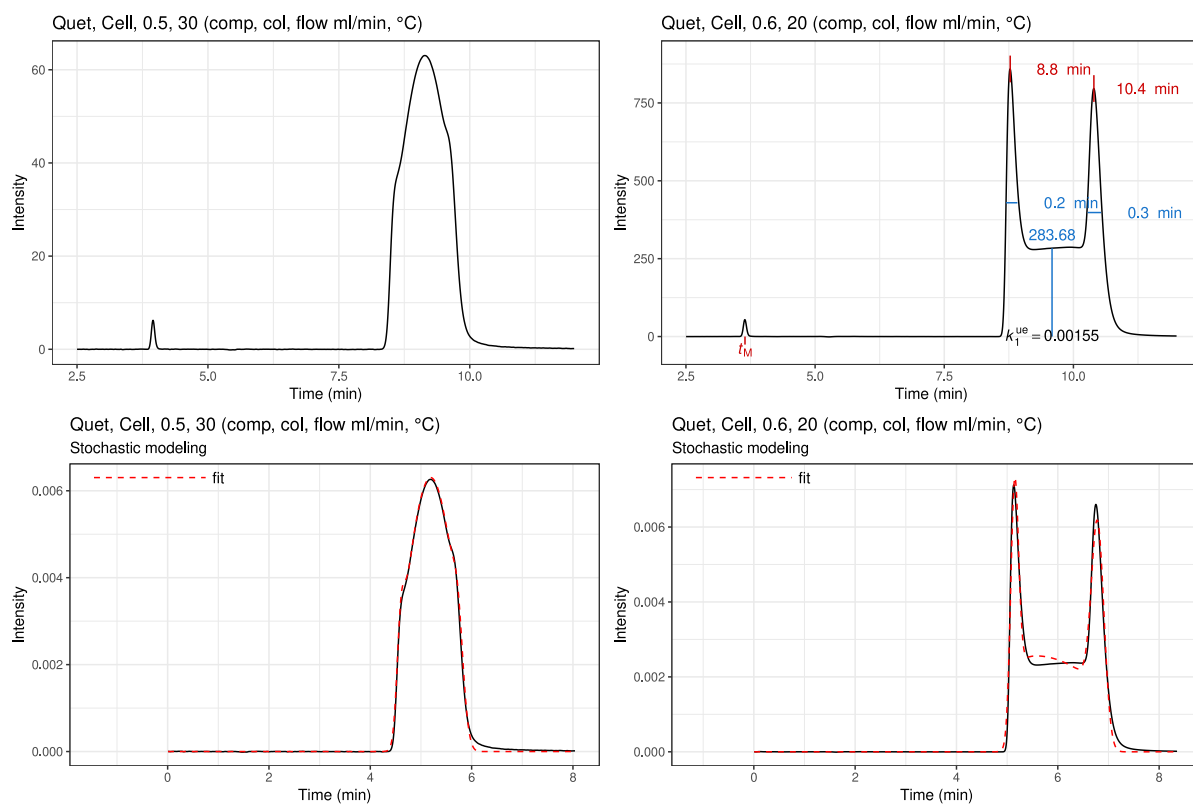


Fig. 3. Two representative chromatograms with merged peaks due to higher temperature (left column), and discernible Batman peaks (right column) measured on the Lux Cellulose-4 chiral stationary phase. In the case of Batman peaks the unified equation affords reliable rate constants by determining the plateau height (top right plot), however the merged peaks can only be fitted with the stochastic model. The bottom row shows the stochastic model fit on deconvoluted chromatograms (i.e., x-axis shifted by the void time).

Table 3

Rate constants (forward direction, in 1/s) of enantiomer interconversion determined using the unified equation (k_{uc}), stochastic modeling (k_{stoch}), and optical activities ($k_{opt}/2$). Note that racemization constants determined by optical rotation have to be divided by 2 to be comparable to enantiomerization constants. Asterisk denotes parameters determined by Sepsey et al. (2018). MoE denotes margin of error, i.e., the half width of the 95% confidence interval.

T (K)	k_{uc}	MoE_{uc}	k_{uc}^*	MoE_{uc}^*	k_{stoch}	MoE_{stoch}	k_{stoch}^*	MoE_{stoch}^*	$k_{opt}/2^*$	$MoE_{opt}/2^*$
283.15	6.0e-04	1.2e-05	1.6e-03	7.4e-05	7.2e-04	1.1e-04	1.1e-03	8.9e-05	NA	NA
288.15	9.9e-04	1.6e-05	2.4e-03	8.6e-05	1.0e-03	5.4e-05	1.7e-03	5.6e-04	6.3e-04	8.5e-06
293.15	1.6e-03	4.5e-05	3.8e-03	1.2e-04	1.9e-03	6.8e-05	3.2e-03	8.5e-04	1.2e-03	1.8e-06
298.15	2.7e-03	1.0e-04	5.2e-03	3.7e-04	2.5e-03	1.3e-04	5.9e-03	7.8e-04	2.3e-03	8.0e-06
303.15	3.9e-03	1.4e-04	7.5e-03	4.7e-04	5.4e-03	3.7e-04	8.7e-03	6.0e-04	4.5e-03	1.9e-05
308.15	5.4e-03	2.6e-04	1.0e-02	5.9e-04	6.5e-03	6.3e-04	9.1e-03	7.1e-03	8.1e-03	6.5e-05
313.15	NA	NA	NA	NA	1.2e-02	3.8e-03	NA	NA	NA	NA

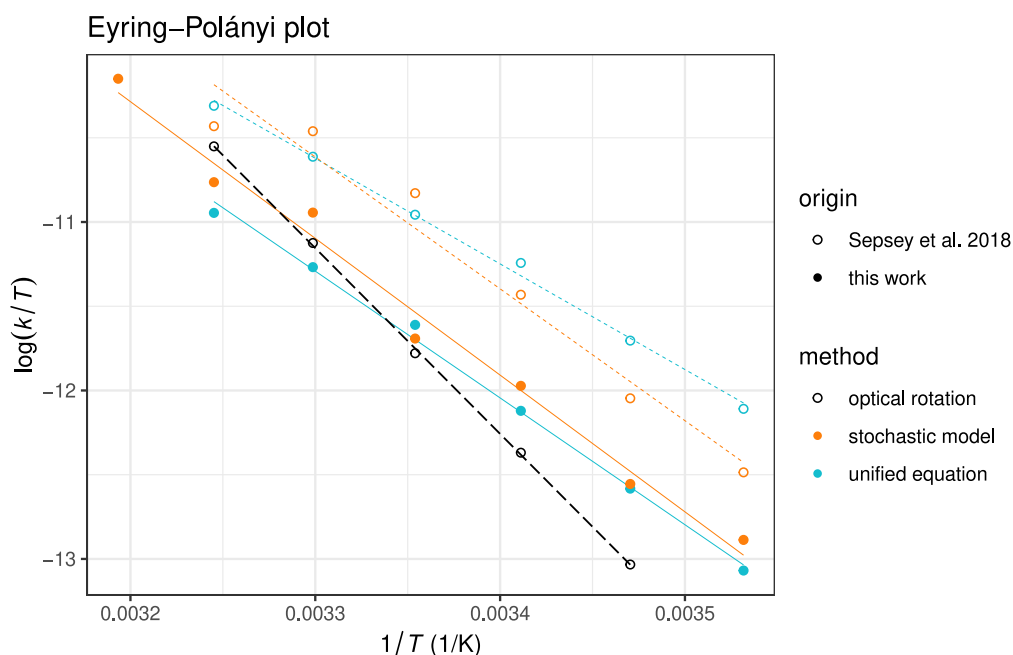


Fig. 4. The linearized Eyring–Polányi plot (left) of rate constants determined with optical rotation (black), unified equation (blue), or stochastic modeling (orange). The open circles are from the work of Sepsey et al. (2018), while filled circles are from this work. The plot on the right shows the Gibbs free energy determined at each temperature using the standard enthalpy and entropies in kJ/mol and kJ/molK, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

approach provided further insight into the robustness of the stochastic model and clarified the experimental balance required between temperature diversity and flow-rate sampling density. Fig. 5 summarizes how a key evaluation metrics change under these leave-out scenarios, providing practical guidance on the minimum experimental design needed for reliable kinetic and thermodynamic inference.

4.3. Measurement on other columns

To assess whether the kinetic and thermodynamic constants of quetiapine depend on the choice of chiral stationary phase, additional experiments were performed using three other polysaccharide-type chiral columns: Lux Cellulose-2, Lux Amylose-1, and Lux i-Amylose-1. The structure of the chiral selectors can be seen in Fig. 2. The resulting fits were evaluated using the same automated workflow as for the Lux Cellulose-4 column. The results are summarized in Table 4, where the fitted parameters from these alternative columns are compared against the benchmark values obtained on Lux Cellulose-4. Since the two methods of unified equation and stochastic modeling show excellent agreement, the best estimate of the kinetic and thermodynamic values is their mean, therefore reported values are averaged over the methods. This comparison enables a direct evaluation of how column chemistry influences the determination of interconversion kinetics and thermodynamics.

Table 4

Rate constants (forward direction, in 1/s) of enantiomer interconversion averaged over the determination methods (unified equation (k_{uc}) and stochastic modeling (k_{stoch})) depicted for the various temperatures (in K). The last 2 rows depict the standard enthalpies and entropies in kJ/mol and kJ/molK, respectively. The standard errors are on the order of 10^{-4} – 10^{-5} 1/s for the rate constants, 1.5–3.5 kJ/mol for standard enthalpies, and 0.005–0.01 kJ/molK for standard entropies.

T (K)	Lux Cellulose-4	Lux Cellulose-2	Lux Amylose-1	Lux i-Amylose-1
283.15	6.6e-04	9.0e-04	7.0e-04	1.0e-03
288.15	1.0e-03	9.0e-04	1.3e-03	1.6e-03
293.15	1.7e-03	1.3e-03	2.2e-03	1.6e-03
298.15	2.6e-03	3.2e-03	3.1e-03	2.7e-03
ΔH°	67.48	34.24	80.92	56.15
ΔS°	−0.07	−0.18	−0.02	−0.11

5. Discussion

Fig. 3 demonstrates that the stochastic model accurately reproduces the experimental chromatographic profiles. Model convergence was achieved for the majority of chromatograms, approximately 95% of the cases examined. The few chromatograms where fitting did not

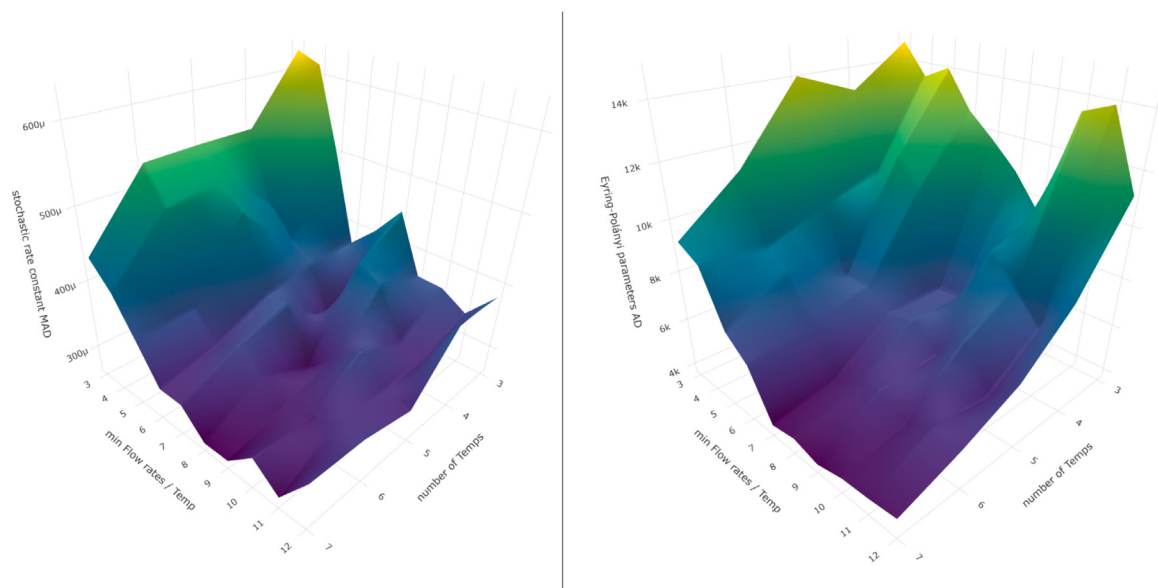


Fig. 5. Three-dimensional surface plots of the mean absolute deviation (MAD) of the stochastic rate constants (top) and absolute deviation (AD) of the Eyring–Polányi estimates (bottom) as a function of the number of temperatures and the minimum number of flow rates. The surface illustrates how experimental design influences the accuracy of parameter estimation, highlighting the dependence of fit quality on both temperature diversity and flow-rate sampling density. The random data holdout was performed with 50 repetitions and the results are averaged.

converge corresponded to conditions with an extremely high number of interconversions, typically arising at very low flow rates, very high temperatures, or under some combination of both. In contrast, the unified equation consistently yielded results when an appropriate peak detection guide was available, although its reliability decreased when the peaks approached coalescence, as reported previously. The comparison of rate constants summarized in Table 3 demonstrates that the unified equation and stochastic model are in very good agreement, with the key distinction that the stochastic model could still be applied under the most extreme temperature conditions employed. Both sets of rate constants also aligned closely with values obtained from optical rotation measurements, calculated by halving the experimentally available rate of racemization, according to the laws of first-order kinetics. The optical rotation data still provide the most accurate estimates of the rate constants; however, the unified equation and stochastic model implemented in this workflow often had better accuracy (narrower confidence intervals) compared to previous determinations. Moreover, the implementation of these analyses in an automated R framework offers distinct methodological advantages: the use of robust optimization algorithms improves flexibility, accelerates the fitting process, and ensures more accurate results than traditional manual fitting approaches. This is epitomized by the linearized Eyring–Polányi plots (Fig. 4), which show that the rate constants determined with our integrated framework fall slightly closer to the regression line derived from the optical rotation data, further supporting the accuracy and reliability of the combined approach. Interestingly, the slopes of all chromatographic models are in mutual agreement but clearly deviate from the slope obtained via optical rotation measurements. This observation reflects the apparent nature of the kinetic constants derived from chromatographic data — they represent an amalgamation of contributions from both the mobile and stationary phases. In contrast, optical rotation measurements probe only the kinetics in the mobile phase. At elevated temperatures, one may expect these lines to converge, as the influence of on-column effects diminishes when the residence time on the stationary phase becomes negligible.

Obvious limitations of the present approach should also be acknowledged. The model assumes a single type of interaction site on the stationary phase, thereby neglecting possible multi-site or heterogeneous binding effects. Likewise, dispersion in the mobile phase is not

explicitly accounted for. The interconversion is treated as a single-step process without intermediate species, although the method does not impose any constraint on whether the interconversion is reversible or irreversible. These simplifications define the operational scope of the model and should be considered when interpreting the results. Despite these limitations, the method offers significant practical advantages. The workflow is computationally efficient: the differential evolution (DE) component of the optimization is automatically parallelized, providing substantial speed gains when processing many chromatograms. The unified interface facilitates full automation of the analysis, enabling high-throughput evaluation with minimal user intervention. At the same time, the structure of the code allows users to override defaults and perform manual or guided fitting when needed by customizing the functions or optimization settings. This combination of speed, flexibility, and transparency makes the approach well suited for both routine and exploratory chromatographic kinetic analysis.

Progressive removal of conditions highlights the minimal experimental design needed for robust kinetic parameter estimation. The ablation study revealed, as expected, that the accuracy of rate constant determination depends predominantly on the diversity of flow rates rather than on temperature. As the number of available flow rates decreased, the mean absolute deviation (MAD) of the fitted kinetic constants increased steadily, underscoring the critical role of flow-rate variation in constraining the stochastic model. By contrast, the Eyring–Polányi fit showed progressive deviation from the baseline with reductions in both temperature and flow, although temperature clearly exerted the greater influence. Interestingly, the standard errors of both the rate constant and the derived thermodynamic estimates increased much more slowly, suggesting that the model retains statistical robustness even as accuracy declines. This behavior indicates a tendency toward systematic bias under reduced experimental diversity: parameter estimates remain precise but drift away from their true values. Taken together, these results highlight the complementary roles of temperature and flow-rate variation: while flow-rate diversity is essential for accurate kinetic fitting, sufficient temperature coverage is necessary to ensure robust thermodynamic parameter estimation. These findings provide practical guidance for designing efficient experiments that minimize measurement effort without compromising the robustness of the extracted constants: a minimum of four distinct temperatures (can

be fewer if emphasis is only on kinetic constants) combined with at least seven flow rates per temperature is recommended to achieve reliable and unbiased determination of both kinetic and thermodynamic parameters.

The comparative analysis of quetiapine interconversion across four different chiral stationary phases (Lux Cellulose-4, Lux Cellulose-2, Lux Amylose-1, and Lux i-Amylose-1) highlights both common trends and distinctive features in the kinetic and thermodynamic behavior. At lower temperatures (283.15 K to 288.15 K), all columns yielded rate constants in the range of 10^{-3} s^{-1} to 10^{-4} s^{-1} , consistent with relatively slow interconversion. With increasing temperature, the rate constants rose steadily, demonstrating the expected Arrhenius-type dependence. Among the tested phases, Lux Cellulose-4 and Lux Amylose-1 columns showed the most pronounced increase with temperature, reaching rate constants above $2 \times 10^{-3} \text{ s}^{-1}$ at 298.15 K. By contrast, Lux Cellulose-2 displayed consistently smaller rate constants across the full range, suggesting a lower efficiency in resolving the Batman peak profile. Lux i-Amylose-1 produced intermediate behavior, with higher values than Lux Cellulose-2 but not exceeding those of Lux Cellulose-4. Thermodynamic analysis provided further insight. The standard enthalpy (ΔH°) was highest for Lux Amylose-1 ($\approx 81 \text{ kJ mol}^{-1}$), followed by Lux Cellulose-4 ($\approx 67 \text{ kJ mol}^{-1}$) and Lux i-Amylose-1 ($\approx 56 \text{ kJ mol}^{-1}$), while Lux Cellulose-2 exhibited the lowest value ($\approx 32 \text{ kJ mol}^{-1}$). These differences suggest column-dependent energy barriers that modulate the interconversion process. The standard entropy (ΔS°) values were generally negative across all stationary phases, indicating that the transition state is more ordered relative to the ground state. The lowest ΔS° was observed for Lux Cellulose-2, supporting its reduced kinetic responsiveness.

The direct comparison of Lux Cellulose-4 and Lux Cellulose-2 is particularly informative, as these chiral selectors differ only in their substituent patterns. The markedly different activation enthalpies reflect their disparate efficiencies as selectors, which is mirrored in the observed differences in rate constants. Interestingly, Lux Cellulose-4 exhibited lower binding affinity (higher enthalpy), whereas Lux Cellulose-2 had higher binding affinity. The stronger binding to the stationary phase in Lux Cellulose-2 was associated with longer retention times and generally larger apparent rate constants, leading to more pronounced peak merging (Fig. 6). However, peak merging can also arise from reduced chiral selectivity of the stationary phase, not solely from accelerated interconversion. The observed relationship between binding strength and interconversion rate suggests that stronger binding may lower the energy of the transition state when this state is favorable for interconversion, thereby catalyzing the process. Conversely, if the transition state is not stabilized by the chiral selector, strong binding may hinder interconversion. Almost the same magnitude of difference is observed in the standard enthalpies of Lux Amylose-1 and its immobilized version Lux i-Amylose-1, which is particularly interesting given that the only difference between the two phases is the immobilization of the chiral selector. These observations illustrate how chiral selector chromatography can serve as a powerful model for enantiomerization in biological contexts, particularly in situations where enantiomers interact with receptors or proteins. The ability of stationary phases to modulate both the kinetics and thermodynamics of interconversion underscores the broader principle that enantiomerization in biological systems may be significantly influenced by binding interactions.

Taken together, the results demonstrate that while all tested chiral stationary phases can capture the enantiomerization of quetiapine, Lux Cellulose-4 and Lux Amylose-1 are most effective in providing reliable Batman peak resolution across a broad temperature window. These findings are consistent with earlier reports on polysaccharide-based stationary phases and confirm that both kinetic and thermodynamic factors must be considered when optimizing chiral separations.

6. Conclusions

This work demonstrates that enantiomerization rate constants can be reliably extracted from Batman peaks using a unified, automated R workflow that integrates both the unified equation and a stochastic modeling approach, with the former providing computational support to the latter. The two methods provided consistent results across a wide range of chromatographic conditions. The close agreement between chromatographic and optical rotation data validates the accuracy of the framework and confirms its suitability for quantitative enantiomerization studies. By leveraging robust optimization algorithms, the workflow affords flexibility, faster processing, and improved accuracy, while remaining fully customizable, enabling adaptation to different analytes, stationary phases, experimental setups, or file formats. Taken together, these results highlight the potential of the presented framework as a practical and reliable tool for routine chromatographic analyses of dynamic enantiomerization processes.

CRedit authorship contribution statement

Arash Mirzahassemi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Anamária Sepsey:** Writing – review & editing, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Ali Mhammad:** Methodology, Investigation, Data curation. **Gergely Dombi:** Validation, Methodology, Investigation, Conceptualization. **Simon Horváth:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Eliza Tóth:** Investigation, Data curation. **Gábor Németh:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Oliver Trapp:** Writing – review & editing, Validation, Software, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Attila Felinger:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Formal analysis, Conceptualization. **Gergő Tóth:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) did not use AI tools to substantially write or generate the manuscript text. Basic tools were employed solely for proofreading, grammar correction, and minor stylistic refinement.

Funding sources

This study was funded by the grant of the Higher Education Institutional Excellence Programme within the Framework of the Molecular Biology Thematic Programme of Semmelweis University, Hungary (TKP2021-EGA-24) (A.M.). This work was funded by the National Research, Development and Innovation Office, Hungary (grant: NKFIH FK 146930). This work was supported by the 2025-2.1.1-EKÖP-2025-00014 University Research Scholarship Programme of the Ministry for Culture and Innovation from the Source of the National Research, Development and Innovation Fund (G.D.), Hungary. This work was supported by the 2025-2.1.2-EKÖP-KDP Scholarship Programme of the Ministry for Culture and Innovation from the Source of the National Research, Development and Innovation Fund (E.T.), Hungary. This work was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (G.T.).

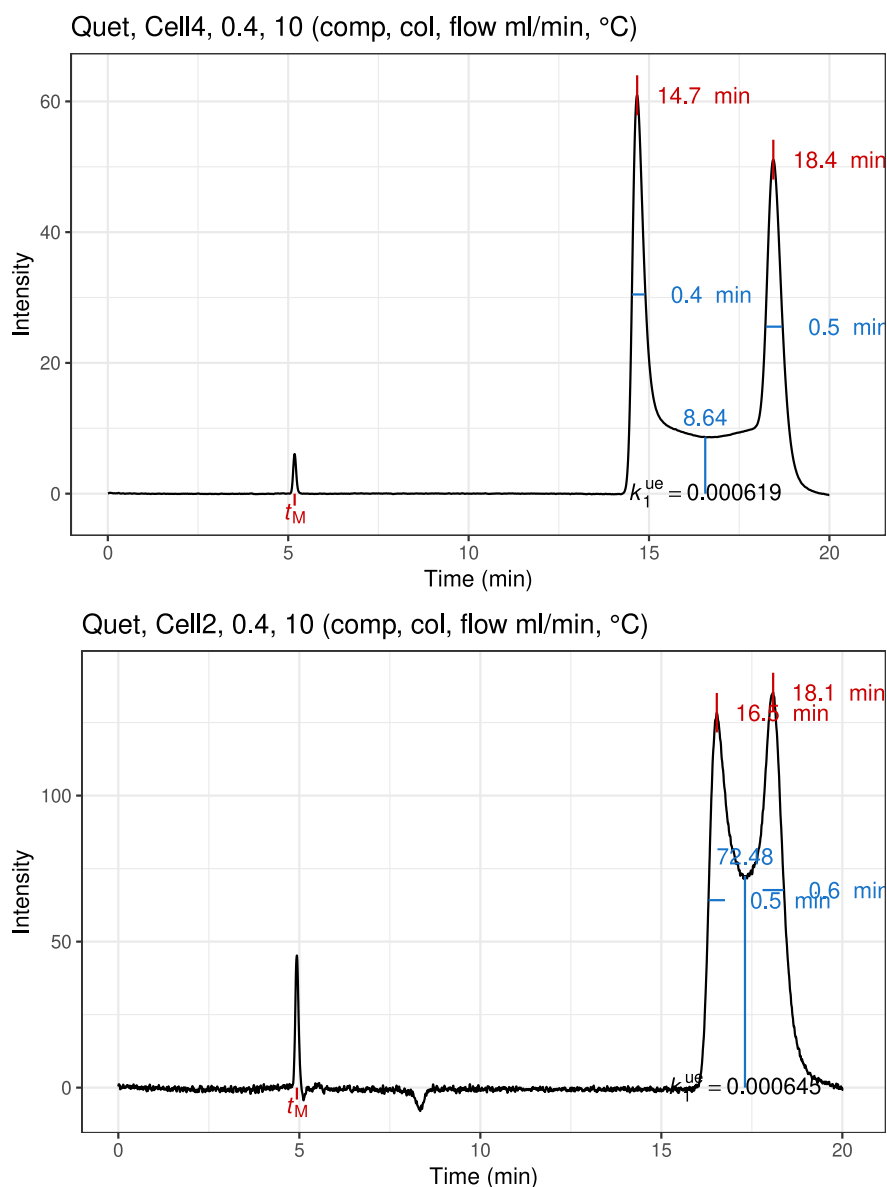


Fig. 6. Two representative chromatograms of quetiapine recorded under identical flow rate (0.4 mL/min) and temperature (283.15 K) conditions on Lux Cellulose-4 (top) and Lux Cellulose-2 (bottom) columns. In the case of Lux Cellulose-2, the chromatogram shows lower retention times accompanied by more clearly resolved peaks, indicating slower interconversion kinetics compared to Lux Cellulose-4.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Arash Mirzahassemi reports financial support was provided by Higher Education Institutional Excellence Programme. Gergo Toth reports financial support was provided by National Research, Development and Innovation Office, Hungary. Gergely Dombi reports financial support was provided by Ministry for Culture and Innovation from the Source of the National Research, Development and Innovation Fund. Toth Eliza reports financial support was provided by Ministry for Culture and Innovation from the Source of the National Research, Development and Innovation Fund. Gergo Toth reports financial support was provided by Hungarian Academy of Sciences. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The generous assistance of Zsolt Lang in the development of the convolution algorithm based on the probability distributions is highly appreciated.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.chroma.2025.466558>.

Data availability

Data will be made available on request.

References

- [1] J. Tomaszewski, M.M. Rumore, Stereoisomeric drugs: FDA's policy statement and the impact on drug development, *Drug Dev. Ind. Pharm.* 20 (1994) 119–139.
- [2] K.M. Kirkland, D.A. McCombs, Changes in chiral selectivity with temperature for an ovomucoid protein-based column, *J. Chromatogr. A* 666 (1994) 211–219.
- [3] V. Schurig, Gas chromatographic separation of enantiomers on optically active metal-complex-free stationary phases. New analytical methods (24), *Angew. Chem. Int. Ed. Engl.* 23 (1984) 747–765.
- [4] S.R. LaPlante, et al., Assessing atropisomer axial chirality in drug discovery and development, *J. Med. Chem.* 54 (2011) 7005–7022.
- [5] A. Felinger, A. Cavazzini, M. Remelli, F. Dondi, Stochastic-dispersive theory of chromatography, *Anal. Chem.* 71 (1999) 4472–4479.
- [6] O. Trapp, G. Schoetz, V. Schurig, Determination of enantiomerization barriers by dynamic and stopped-flow chromatographic methods, *Chirality* 13 (2001) 403–414.
- [7] I. D'Acquarica, F. Gasparrini, M. Pierini, C. Villani, G. Zappia, Dynamic HPLC on chiral stationary phases: A powerful tool for the investigation of stereomutation processes, *J. Sep. Sci.* 29 (2006) 1508–1516.
- [8] A. Felinger, Molecular dynamic theories in chromatography, *J. Chromatogr. A* 1184 (2008) 20–41.
- [9] J. Krupčík, J. Mydlová, P. Májek, P. Šimon, D.W. Armstrong, Methods for studying reaction kinetics in gas chromatography, exemplified by using the 1-chloro-2, 2-dimethylaziridine interconversion reaction, *J. Chromatogr. A* 1186 (2008) 144–160.
- [10] C. Wolf, *Dynamic Stereochemistry of Chiral Compounds: Principles and Applications*, RSC Publishing, 2008.
- [11] V. Schurig, Elaborate treatment of retention in chemoselective chromatography—the retention increment approach and non-linear effects, *J. Chromatogr. A* 1216 (2009) 1723–1736.
- [12] W. Bürkle, H. Karfunkel, V. Schurig, Dynamic phenomena during enantiomer resolution by complexation gas chromatography: A kinetic study of enantiomerization, *J. Chromatogr. A* 288 (1984) 1–14.
- [13] O. Trapp, Unified equation for access to rate constants of first-order reactions in dynamic and on-column reaction chromatography, *Anal. Chem.* 78 (2006) 189–198.
- [14] O. Trapp, A novel software tool for high throughput measurements of interconversion barriers: DCXplorer, *J. Chromatogr. B* 875 (2008) 42–47.
- [15] A. Sepsey, D.R. Németh, G. Németh, A. Felinger, Rate constant determination of interconverting enantiomers by chiral chromatography using a stochastic model, *J. Chromatogr. A* 1564 (2018) 155–162.
- [16] J.C. Giddings, H. Eyring, A molecular dynamic theory of chromatography, *J. Phys. Chem.* 59 (1955) 416–421.
- [17] R. A. Keller, J.C. Giddings, Multiple zones and spots in chromatography, *J. Chromatogr. A* 3 (1960) 205–220.
- [18] E. Bass, *Chromconverter: Chromatographic file converter*, 2025, <https://doi.org/10.5281/zenodo.6792521>.
- [19] H.W. Borchers, *Pracma: Practical numerical math functions*, 2023.
- [20] D.A. McQuarrie, On the stochastic theory of chromatography, *J. Chem. Phys.* 38 (1963) 437–445.
- [21] A. Cavazzini, M. Remelli, F. Dondi, A. Felinger, Stochastic theory of multiple-site linear adsorption chromatography, *Anal. Chem.* 71 (1999) 3453–3462.
- [22] R.A. Rigby, D.M. Stasinopoulos, Generalized additive models for location, scale and shape, (with discussion), *Appl. Stat.* 54 (3) (2005) 507–554.
- [23] K. Mullen, D. Ardia, D.L. Gil, D. Windover, J. Cline, DEoptim: An R package for global optimization by differential evolution, *J. Stat. Softw.* 40 (2011) 1–26.
- [24] T.V. Elzhov, K.M. Mullen, A.-N. Spiess, B. Bolker, Minpack.lm: R interface to the Levenberg–Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds, 2023.
- [25] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, 2023.