

# Analyst

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: R. Monge Neria, R. A. Saylor and L. Kisley, *Analyst*, 2026, DOI: 10.1039/D6AN00199H.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

## The non-linear effects of the number of stochastic single-molecule adsorption events on ensemble elution profiles

Ricardo Monge Neria<sup>a</sup>, Rachel A. Saylor<sup>b</sup>, and Lydia Kisley<sup>a, c\*</sup>

<sup>a</sup>Department of Physics and <sup>c</sup>Department of Chemistry, Case Western Reserve University, Cleveland, Ohio, USA 44106

<sup>b</sup>Department of Chemistry and Biochemistry, Oberlin College, Oberlin, OH, USA 44074

\*Corresponding author: lydia.kisley@case.edu

### Abstract

The stochastic model of chromatography mathematically represents the molecular mass transfer that occurs during a separation. Single-molecule microscopy allows for direct visualization of molecular analytes adsorbing within column materials. Experimental single-molecule data and the stochastic model can predict elution profiles using the adjustable variable  $\bar{r}_m$ , the average number of adsorptions per molecule. Previously, only a single  $\bar{r}_m$  value was used to match either peak location or shape, while the effects of adjusting  $n_p$ , the number of modeling points, have not been studied. Here, we systematically explore a wide range of these two variables in the stochastic model to determine if it is possible to optimize agreement between modeled single-molecule and high-performance liquid chromatography (HPLC) chromatograms. A metric to quantify chromatogram agreement is introduced by taking the weighted difference in the elution time and shape of the chromatograms. We determine the non-linear effects of  $\bar{r}_m$  and  $n_p$  on peak height, width, and asymmetry and link the observations to the molecular behavior. Applying our approach to experiments with variable flow rate shows that increased sampling of rare, long time adsorption events affects agreement between simulated and HPLC elution profiles. Finally, we make quantitative recommendations that the single-molecule experiments should sample available binding sites following an exponential association model and that  $n_p > 1.5 \cdot \bar{r}_m$  to achieve accurate results. Overall, we verify that the current form of the stochastic model based solely on mass transfer is unable to simultaneously match both peak location and shape and make recommendations for future improvements to the model by separation scientists.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Unported Licence  
AM  
Commercial 3.0  
Attribution-NonCommercial 3.0  
CC BY-NC  
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence  
5/22/2016 4:40:44 AM  
Downloaded on 01/14/2016  
Published on 01/14/2016  
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence

## 1. Introduction

The macroscopic<sup>1-4</sup> and microscopic<sup>5-8</sup> theories of chromatography offer two different perspectives on the same fundamental processes of separation. Traditional liquid chromatography (LC) models use the top-down macroscopic view, focusing on Fickian and similar mechanistic descriptions of diffusion and convection, defined in terms of bulk properties such as mass transfer and concentration zones. Macroscopic models are based on solving differential mass balance equations that account for all processes affecting the analyte concentration in a given space over time, describing the overall movement of solute bands down the column. Commonly used examples are the general rate (GRM), equilibrium dispersive (EDM), transport dispersive (TDM), and lumped rate (LRM) models.<sup>1-4</sup> Phenomenological, experimentally determined coefficients are also often incorporated to help characterize the rate of mass transfer between the mobile and stationary phases. Similarly, theoretical plates and van Deemter equation analysis are frequently used to relate column efficiency (defined by plate height  $H$ ) to the mobile phase velocity, and explain band broadening in terms of eddy diffusion, longitudinal diffusion, and resistance to mass transfer between phases.<sup>9-11</sup>

In contrast, microscopic models of chromatography view the separation process as the result of a statistical distribution of single-molecule-scale behaviors. This bottom-up approach, originally proposed by Giddings and Eyring,<sup>5</sup> models the behavior of individual solute molecules as they travel through the column. The microscopic (or stochastic) theory models the migration of single molecules (SM) as random walks, where molecules adsorb to the stationary phase and then desorb. The adsorption-desorption events are treated as a Poisson process, with exponentially distributed residence times in each phase. The final chromatographic peak shape is then the result of the summed behavior of a vast number of individual molecules, with the elution profile described by the probability density function of the retention times for all the molecules.

However, while both the macroscopic and microscopic approaches are complementary and, under linear conditions, can be shown to be mathematically equivalent,<sup>12-14</sup> the microscopic description is less widely used. The microscopic model has been expanded into a Lévy process representation,<sup>6</sup> and applied to study chromatographic retention in relation to experimental single-molecule microscopy observables<sup>15-19</sup> and Monte Carlo simulations.<sup>7,20</sup> Previous applications<sup>7,8,16,18,21</sup> have shown that rare, long-lasting adsorptions can have a significant influence on elution time and the skewness of chromatograms. However, these studies have struggled to obtain consistent agreement between the experimental on-column elution and the SM-based simulations, with some inconsistencies in the use of the model, as well as the adjustment and inclusion of free parameters.

Here, we explore if the Lévy process representation of the stochastic model can match both peak elution time and shape of HPLC on-column elution peaks from single-molecule adsorption data through the iterative variation of model inputs to reveal nonlinear behaviors. We first review the history, theory, and derivation of the Lévy process representation of the stochastic model and show how simulated chromatograms vary depending on the average number of SM adsorptions ( $\bar{r}_m$ ), as well as the number of modeling points ( $n_p$ ). Systematically scanning over a large range of variables that represent realistic column lengths, we quantify nonlinear trends in peak height, width, and asymmetry with respect to both parameters, as well as under sampling that can occur due to the choice of  $n_p$ . Making use of experimentally measured SM adsorptions on commercial stationary phase particles and associated HPLC on-column experiments,<sup>19</sup> we implement an error minimization approach to match stochastic modeled results to the experimental HPLC data and explain the sources of discrepancy. Finally, we discuss sampling issues that arise from the experimental SM conditions. Overall, we show how free parameters can be adjusted to match specific characteristics of the experimental HPLC elution profiles, while revealing nonlinearities in the model behavior with respect to the average number of adsorptions, as well as under sampling that can occur in both the model and SM experiments. We make recommendations for

total data collection and  $n_p$  to ensure adequate sampling in future single-molecule measurements and modeling of chromatography.

## 2. Theory and background

### 2.1 The stochastic model of chromatography.

As the majority of processes underlying a separation are stochastic in nature (*i.e.* diffusion and sorption behaviors), the resulting chromatogram shape can be explained by probability distributions. Adhering with this philosophy, in the 1950s the works of H. Eyring, J. C. Giddings, and R. A. Keller, initially in relation to gas and thin layer chromatography,<sup>5,22,23</sup> established the stochastic (or microscopic) model of chromatography. They modeled the random migration trajectories of single molecules across a chromatography column based on two fundamental processes: 1) molecular movement in the mobile phase and 2) adsorption-desorption of molecules in the stationary phase. The fast diffusion behavior of molecules in the mobile phase was originally defined as a one-dimensional random walk, and the slower adsorption-desorption process as an exponential Poisson distribution of the sojourn time of molecules traveling throughout both phases.<sup>5,24,25</sup>

The trajectories of molecules were defined in terms of probabilities of adsorption on the surface of the stationary phase and a constant mobile phase velocity that is maintained for the overall column. Molecules through the chromatographic bed spend an average fly time ( $\tau_m$ ) in the mobile phase between adsorption events and an average sojourn time ( $\tau_s$ ) adsorbed on the surface of the stationary phase. These two quantities ( $\tau_m$ ,  $\tau_s$ ) are randomly distributed and account for the unique residence times of each individual molecule with its own trajectory. From this, the chromatogram then becomes the probability-density function of retention times for individual molecular trajectories, modeled as a Poisson distribution.<sup>5,24</sup>

Adsorption-desorption kinetics are related to the sojourn times in terms of mass transfer rate constants. These are defined as the probability per unit time that a molecule enters ( $k_a = \frac{1}{\tau_m}$ , for adsorption) and leaves ( $k_d = \frac{1}{\tau_s}$ , for desorption) the stationary phase, with a retention factor ( $k = \frac{k_a}{k_d}$ ). Applying these to a Poisson distribution of molecules that spend time  $t_0$  (unretained, also known as void time) in the mobile phase, and time  $t_R$  (retained) in the stationary phase results in the probability density function  $P(t)$  that defines the chromatographic peak shape as:

$$P(t) = \frac{\sqrt{4k_a k_d t_0 t}}{2t} I_1(\sqrt{4k_a k_d t_0 t}) e^{-k_d t - k_a t_0} \quad (1)$$

where  $I_1$  is a modified Bessel function of the first kind and first order. The probability density function can be approximated by a Gaussian function when the number of mass transfer events is sufficiently large, as:

$$P(t) = \frac{k_d}{2\sqrt{\pi k_a t_0}} \exp\left[-\frac{(k_a t_0 - k_d t)^2}{4k_a t_0}\right] = \frac{1}{2\tau_s \sqrt{\pi n}} \exp\left[-\frac{(n - t/\tau_s)^2}{4n}\right] \quad (2)$$

where  $n$  corresponds to the average number of adsorption events,  $n = \frac{t_0}{\tau_m} = k_a t_0$ .<sup>5,24</sup>

Unfortunately, at the time when Giddings and Eyring first published their work (1955), the stochastic theory saw little application due to computation methods being incredibly laborious and it being experimentally impossible to measure the necessary molecular observables for practical implementation. By 1963, D. A. McQuarrie worked to expand the model, and provide a more usable

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

solution to the many sites version of the stochastic model.<sup>26</sup> Then, in 1999, the multiple-site model was further refined to account for more complex molecular interactions by using a characteristic function (CF) definition of the equations by A. Cavazzini et al.<sup>27</sup> Their work depicted the chromatogram as originally defined by the Giddings-Eyring expression in eqn 2 by the inverse Fourier transform of the CF and was successfully used to generate model elution curves. The same group, under A. Felinger, then continued to develop the model, including creating a novel Monte Carlo simulation based approach for nonlinear chromatography.<sup>28</sup> By 2003, L. Pasti et al.<sup>29</sup> provided experimental validation of the stochastic theory by applying it to size exclusion chromatography (SEC) column experiments eluting multiple sized polymer standards. However, at this point the CF-SEC was complex and difficult to decipher, and the lack of experimental single-molecule observables made it challenging to further bridge the gap between the ensemble-scale chromatograms and the theory.

### 2.2 The Lévy representation of the stochastic model applied to single-molecule experiments.

A direct connection between experimental single-molecule observables, the molecular dynamics on which the stochastic model is based, and on column results was made in 2005 by Pasti et al.<sup>15</sup> In this approach, the Poisson distribution CF description was replaced by a Lévy process probability distribution formalism to define the CF of molecular trajectories of the stochastic model. The Poisson process is the simplest jump-based Lévy process, which models events that occur at a constant average rate but at random intervals. By definition, a Poisson process is spatially discrete, with fixed jump distances. However, chromatographic migration down a column occurs on a continuum, and with random diffusion and dispersive mechanisms. This better aligns with a Lévy process description, which is a more general class of stochastic processes that represent random walks in continuous time, where sudden and independent jumps of varying amounts are allowed. In essence, this turns the chromatographic description into a more complete three components definition: 1) the constant solvent migration velocity, 2) the Brownian dispersive component due to diffusion behavior, and 3) the Poisson randomization of jumps/sorption times.<sup>6</sup>

Under the Lévy process formalism, the CF expression exists as a frequency ( $\omega$ ) distribution of different solute retention times and is re-written as a discrete sum to relate this formalism back to SM experimental observables, e.g. adsorption times. Defining the original Poisson probability distribution that during time  $t_m$  a given molecule will adsorb  $r_m$  times as:

$$P(r_m) = \frac{(\mu t_m)^{r_m}}{r_m!} e^{-\mu t_m} \quad (3)$$

where  $\mu$  is the frequency of adsorption events. Here, solute molecules migrating down the length of column  $l$ , with a mobile phase velocity  $v_m$ , spend time  $t_m = l/v_m$  in the mobile phase. Thus, the average value of  $r_m$  becomes:

$$\bar{r}_m = t_m \mu = t_m / \tau_m \quad (4)$$

where  $\tau_m = 1/\mu$ , the average time spent in the mobile phase between adsorptions. Then, for the distribution of random  $r_m$  values, the CF is defined in the frequency domain as:

$$\phi(r_m; \omega) = \sum_r e^{ir\omega} P(r_m) \quad (5)$$

For the Lévy process representation,<sup>15</sup> the CF is defined as:

$$\phi(t_s; \omega | t_m) = \exp \left\{ \bar{r}_m \sum_{i=1}^{i=m} (e^{i\omega\tau_{s,i}} - 1) \Delta F_S(\tau_{s,i}) \right\} \quad (6)$$

where  $t_s$  and  $t_m$  are times spent in the stationary and mobile phases, respectively (as defined above). The variable  $\tau_s$  is the random time spent during one sorption step, and  $\Delta F_S(\tau_{S,i})$  is the normalized frequency function that defines the probability of observing a sorption time in the range  $\tau_{S,i}$  and  $\tau_{S,i} + \Delta\tau_S/2$  (*i.e.* probability distribution for sorption times). The frequency function  $\Delta F_S(\tau_{S,i})$  is normalized as:

$$\sum_{i=1}^{i=m} \Delta F_S(\tau_{S,i}) = 1 \quad (7)$$

The CF function (an array of values) is joined by the mirror (*i.e.* flipped vector left-to-right, or a 180° rotation) of its complex conjugate, and an Inverse Fast Fourier Transform (IFFT) is applied to return the chromatogram distribution.

$$[\phi(t_s; \omega | t_m), R_{180}(\phi^*(t_s; \omega | t_m))] \xrightarrow{IFFT} f(t_s) \quad (8)$$

In application,  $\Delta F_S(\tau_{S,i})$  is defined by the measured (normalized) frequency distribution of observed single-molecule adsorption times ( $\tau_{S,i}$ ), and  $\omega$  is defined by the number of model sampling points ( $n_p$ ; *i.e.* number of points in the final chromatogram) as:

$$\omega = \frac{k\pi}{n_p \Delta\tau_S} \quad (9)$$

Importantly, only the real part of the IFFT is kept, and it is scaled by  $\sqrt{2n_p}$  to account for sampling points, with the time axis defined by steps of  $\Delta\tau_S$  (the time resolution of the measured single-molecule observations).

This Lévy process formulation was then applied by Pasti et al. to experimental single-molecule data, using eqn 8 with experimental measurements of adsorption-desorption times obtained by M. J. Wirth and D. J. Swinton<sup>30</sup> and S. H. Kang, M. R. Shortreed, and E. S. Yeung<sup>31</sup> through single-molecule spectroscopy on stationary phase-mimicking thin film surfaces. The model results were compared to Wirth and Yeung's associated on-column experimental elution results. The work by Pasti et al. represented the first successful application of the stochastic theory in direct connection to the experimental observables from single-molecule fluorescence microscopy. The stochastic model with the Lévy process formulation was an essential bridge between a molecular perspective of chromatography and ensemble separations via the model, single-molecule microscopy, and ensemble HPLC and electrophoretic data.

Agreement between the model and ensemble results were limited by differences in sample conditions and the inaccessibility of the parameter  $\bar{r}_m$  from single molecule experiments. In practice,  $\bar{r}_m$ , the average number of adsorptions throughout the column was left as an adjustable free parameter which had a heavy influence on the resulting elution time and shape of the model peaks. As such, Pasti et al. approached this disconnect by defining two different  $\bar{r}_m$  values to try to either match experimental elution peak shape (chosen by inference and trial-and-error), or elution time by defining  $\bar{r}_m$  in relation to the ratio of the experimental peak elution time ( $t_s$ ) and the average adsorption (dwell) time ( $\tau_s$ ) from the SM measurements as:

$$\bar{r}_m = \frac{t_s}{\tau_s} \quad (10)$$

However, it was not clear how these definitions of  $\bar{r}_m$  affect other aspects of the model peak, with Pasti et al. noting that “[t]wo simulations were thus performed aiming at attaining the best agreement, for one, in peak shapes and, for the other one, in peak positions.”<sup>15</sup> Further, the experimental single-molecule data was limited to model thin film surfaces which can have different chemistries and steric pore structure than



the more complex, commercially-prepared stationary phases packed in columns typically used in chromatography.<sup>32</sup>

Two open questions were left about the applicability of the Lévy representation of the stochastic model of chromatography. First, is it possible to optimize and achieve agreement between both shape (width and asymmetry) and position of a chromatogram and single-molecule measurement? And second, was the inability to match the elution profiles a result of experimental differences in the thin film samples compared to the chromatographic packing material, the adjustable variables  $\bar{r}_m$  and  $n_p$  used, or a limitation of the model itself? To tackle these questions, here we explore the stochastic model and accuracy in relation to packed bed (fully porous spherical particles) liquid chromatography columns and recent single-molecule microscopy measurements in the exact same stationary phase particles.<sup>19</sup>

### 3. Material and Methods

#### 3.1 Single-molecule (SM) microscopy experiments.

Single-molecule fluorescence microscopy of commercially available stationary phase particles was performed following previously developed methodology in Monge Neria et al.<sup>19</sup> Briefly, dry powder (unpacked) commercial stationary phase particles, Cellulose-B (cellulose tris(3,5 dimethyl phenyl-carbamate); 100 nm nominal pore diameter,  $5 \pm 0.8 \mu\text{m}$  particle diameter) were obtained from Regis Technologies Inc.. The particles were immobilized on silica glass for microscopy by relying on polar interactions, as previously described.<sup>32</sup> Microscopy slides (20 mm x 30 mm, #1.5, Fisherbrand) were prepared by submerging in a base-peroxide ( $\text{H}_2\text{O}_2 + \text{NH}_4\text{OH} + \text{H}_2\text{O}$ , at a 1:1:6 volume ratio, respectively) bath at 70 °C for 90 s, water rinsed, dried with nitrogen gas (5.0 grade, Airgas), then plasma cleaned in an  $\text{O}_2$  (Industrial grade, Airgas) plasma cleaner (PDC-32G, 115 V, Harrick Plasma) at 140-280 Torr, medium power, for 2 minutes. Different water suspensions of 0.1 wt % solids were used to deposit the stationary phases by drop casting 8  $\mu\text{L}$  volumes unto the glass slides and allowed to dry overnight. The slides were then covered by silicon flow cells (Hybriwell, 13 mm diameter, 0.15 mm depth, Grace Biolabs) and rinsed multiple times with water, then left to rehydrate for at least 2 hours before imaging.<sup>19</sup>

Single-molecule measurements of adsorption were achieved with highly inclined and laminated optical sheet (HILO) fluorescence imaging of nanomolar concentrations of fluorescent rhodamine 6G dye (99%, Fisher) under flow at ambient 22 °C, as previously described.<sup>32</sup> The instrument consists of an Olympus IX-73 inverted microscope, with a 100x magnification oil immersion objective (Olympus, 100x, NA 1.49, UAPON100XOTIRF), and a CNI diode laser (532 nm, MGL-III-100 mW) used for excitation. The laser light is set to a 77° exit angle from the objective to achieve HILO. Imaging was done with 1 nM rhodamine 6G solutions in 20 mM HEPES buffer (Biotang Inc, pH 7.33) passed through the flow cells using a syringe pump (NE-1000, New Era Pump Systems Inc.). Flow rates were varied between 1 to 100  $\mu\text{L}/\text{min}$ , corresponding to linear velocities 0.06 to 6.0 cm/min (Table S2). An EMCCD camera (iXon Life 897, Andor) collected the emission signal at 30 ms exposure, 400 electron multiplying gain, in Frame Transfer mode under 532 nm laser excitation of 10 mW (379 W/cm<sup>2</sup>) at the sample. Single molecules were identified and analyzed using a home-written, publicly available<sup>33</sup> MATLAB (2022b) code that uses 2D Gaussian fitting strategies, paired with radial symmetry<sup>34</sup> centroid position refinements, to obtain super-resolved, single-molecule locations and adsorption times (**Fig. S1**).

#### 3.2 HPLC bulk measurements of chromatographic elutions to match SM experiments.

Bulk chromatography experiments were performed on a Shimadzu liquid chromatography instrument (LC-20AD Prominence) with photodiode array detection (SPD-M20A, Prominence) with an isocratic mobile phase of 70% ethanol (EtOH, 200 proof, ACS grade from Pharmaco) and 30% HEPES (20 mM, pH 7.33, 2-[4-(2-hydroxyethyl)-1-piperazinyl]-ethane sulfonic acid, high purity from VWR). A single



REFLECT C-Cellulose B (5  $\mu\text{m}$ , 5 cm length x 4.6 mm diameter from Regis Technologies) column was rinsed and equilibrated with mobile phase (70% EtOH:30% HEPES) at 0.1 mL/min for 2 hrs prior to experiments.

For comparison to microscopy measurements, rhodamine 6G (3  $\mu\text{L}$ , 100  $\mu\text{M}$  in HEPES buffer, from Biotang Inc) was run with an isocratic mobile phase of 70% EtOH: 30% HEPES, and varying flow rates (see **Table S2**) up to 0.5 mL/min, and a column temperature of 30  $^{\circ}\text{C}$ . Detection was accomplished using a photodiode array over the range 200 – 600 nm at 5 Hz. All elution runs were done in triplicate. Resulting chromatograms (**Fig. S2**) were visualized in Shimadzu Lab Solutions software, extracted at 530 nm (4 nm bandwidth), exported, and further analyzed.

### 3.3 An error minimization approach to implement the stochastic model and adjust the free parameter $\bar{\tau}_m$ to match HPLC elution curve characteristics.

Single molecules were first identified and analyzed using a home-written, publicly available<sup>33</sup> MATLAB (2022b) code, where spatial grouping and shape cross-correlation are used to generate super-resolution images where sites are validated based on event density and binding specificity. From these validated locations, kinetic adsorption data is extracted by converting frame counts into precise dwell (defined by the continuous presence of stationary/bound molecules) and association (defined as the time between molecular adsorption in a given site) times (see SI section 2).

A modified version of the stochastic model implementation from Pasti et al. was implemented in MATLAB (2024a), which can be found at (*DOI on Zenodo will be added at time of publication*).<sup>18,19</sup> We systematically tested elution profile behavior under varying  $n_p$  (50 to 2000, variable step sizes of 10, 50, and 100) and  $\bar{\tau}_m$  (1 to 2000, variable step sizes of 1, 5, 10, and 100) and with different single-molecule dwell time (*i.e.* adsorption) distributions (see **Fig. S1**) obtained from microscopy experiments at varying flow rates (Table S2). An option to select  $\bar{\tau}_m$  between the theoretical value defined by **eqn 10**, or an error minimization fitting with respect to experimental chromatogram elution and shape between the model elution and HPLC data was implemented in our code. All results presented here use the error minimization approach unless stated otherwise, or if  $\bar{\tau}_m$  is manually selected/defined for the purposes of testing model outputs. The error minimization scales as three varied forms of error parameter,  $E$ , calculations for 1) elution time, 2) peak width, and 3) peak asymmetry defined as:

$$1) E_{time} = \left| \frac{(t_{R,LC} - t_{0,LC}) - t_{S,model}}{(t_{R,LC} - t_{0,LC})} \right| \quad (11)$$


$$2) E_{width} = \left| \frac{(w_{0.1,LC}) - (w_{0.1,model})}{w_{0.1,LC}} \right| \quad (12)$$

$$3) E_{asym} = \left| \frac{\left(\frac{b_{LC}}{a_{LC}}\right) - \left(\frac{b_{model}}{a_{model}}\right)}{b_{LC}/a_{LC}} \right| \quad (13)$$

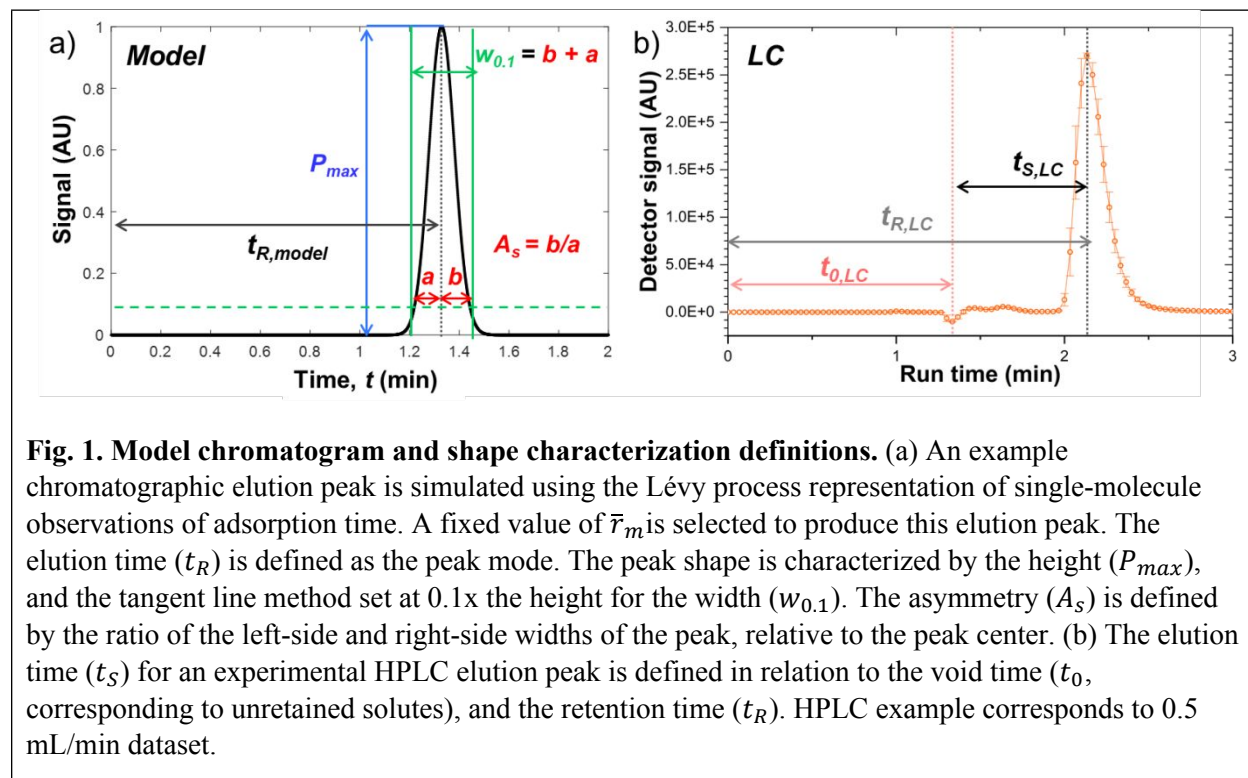
Here the  $LC$  subscript labels correspond to the HPLC bulk chromatography measurements compared to the *model* subscript referring to the single-molecule model run parameters. We characterized the elution peak time and shape as shown in **Fig 1a**. The asymmetry factors ( $A_s = b/a$ , red) are defined by the left ( $a$ ) and right ( $b$ ) peak widths with respect to the peak center (max), with the width ( $w_{0.1} = a + b$ , green) defined using the tangential method at 10% signal (0.1 of the normalized peak height,  $P_{max}$ , blue). For the experimental elution time ( $t_{S,LC}$ ), we defined it in relation to the void time ( $t_{0,LC}$ ), signaled by the dip caused by the unretained volume during sample injection (**Fig. 1b**), and the elution peak location ( $t_{R,LC}$ ). Thus, the corrected experimental elution time is defined by  $t_S = t_R - t_0$ , where time  $t_0$  (unretained, also known as void time) in the mobile phase, and time  $t_R$  (retained) are used. Meanwhile, the model peak

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Open Access Article. Published on 01 May 2024. Downloaded on 5/22/2024 4:40:44 AM.  
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.



elution time ( $t_{R,model}$ ) is defined as the peak mode, which coincides with the peak maximum for low asymmetry (near  $A_s = 1.0$ ) peaks. Note that the model peak is simulated using time units ( $\Delta\tau_s$ , from the time resolution of SM observations, 30 ms in our case) of ms, but the resulting time axis (*i.e.* the time delta between model points,  $n_p$ ) is defined by unit conversion (*e.g.*  $\frac{1 \text{ ms}}{100 \text{ sec}} \cdot \frac{1 \text{ sec}}{60 \text{ min}}$ ).



There is no void time to correct for in the model peaks. The net fitting error for minimization is defined as the sum of the desired fitting parameters, weighted by the parameter  $W_i$ , as:

$$Error_{net} = \sqrt{\sum_i W_i (E_i^2)} \quad (14)$$

where the weights are all set equal to one by default but are changed to prioritize a specific peak characteristic (see Results). Hereafter we report our weights in the format  $\mathbf{W}(W_{time}, W_{width}, W_{asym})$ . Unlike the Generalized Method of Moments or *l-moments*, we maintained only three ‘moments’ and only adjust the single parameter  $r_m$  when fitting to help avoid overparameterization, avoid making complex assumptions about the system (*e.g.* mass transfer rates and diffusivity), nor require multiple runs of mixed elutants.<sup>35–38</sup>

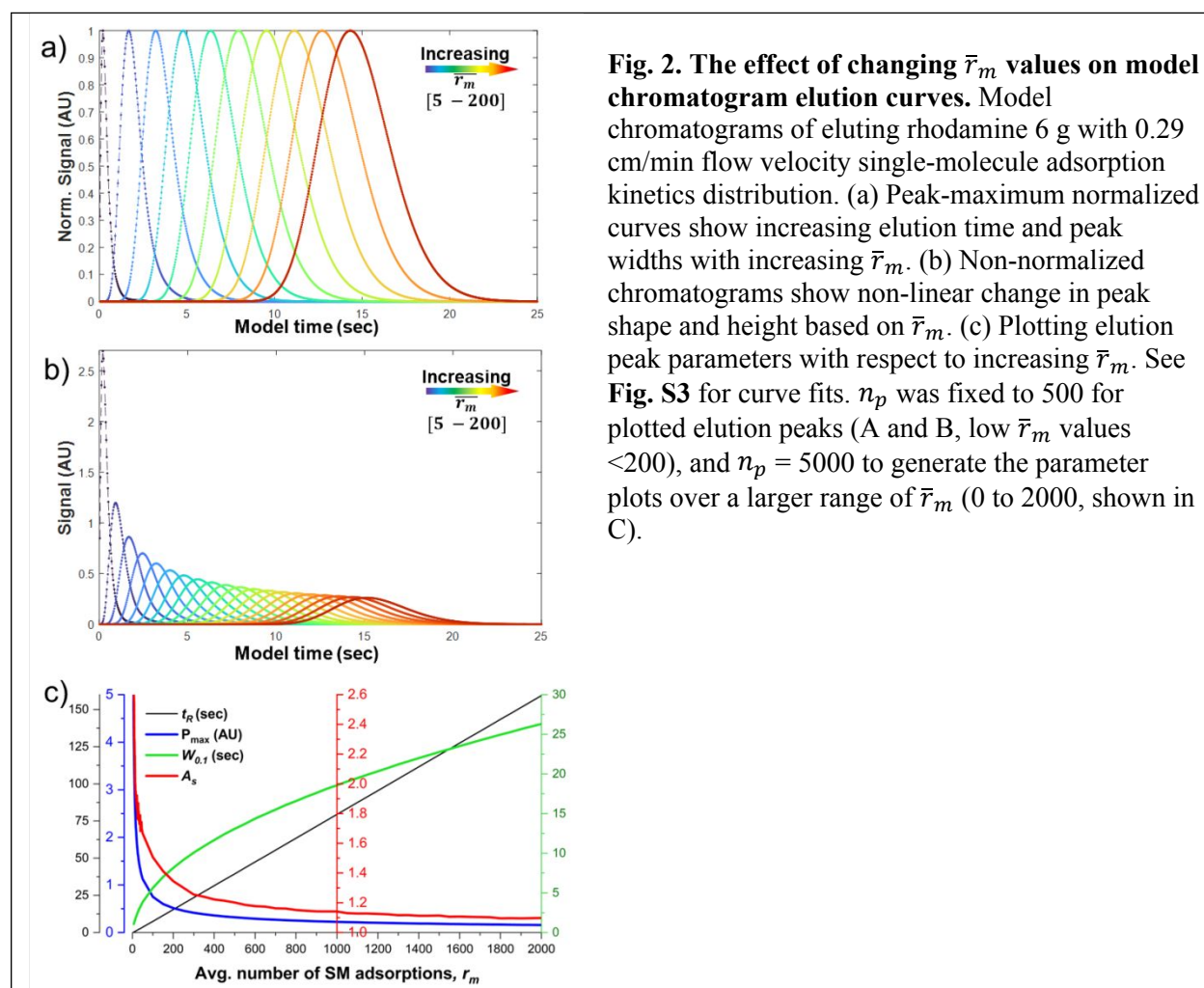
## 4. Results

### 4.1 Single-molecule data reveals non-linear model elution peak dependance on $\bar{r}_m$

We generated model elution chromatograms by applying the Lévy process formulation presented in **eqn 7** and **eqn 8** and manipulated  $r_m$  over hundreds of values to explore the effects on the model peak elution time ( $t_R$ ), peak asymmetry ( $A_s$ ), and the peak width ( $w_{0.1}$ ). We used our SM experimental dwell time distribution values of rhodamine 6g analyte adsorption to Regis Technologies REFLECT C-

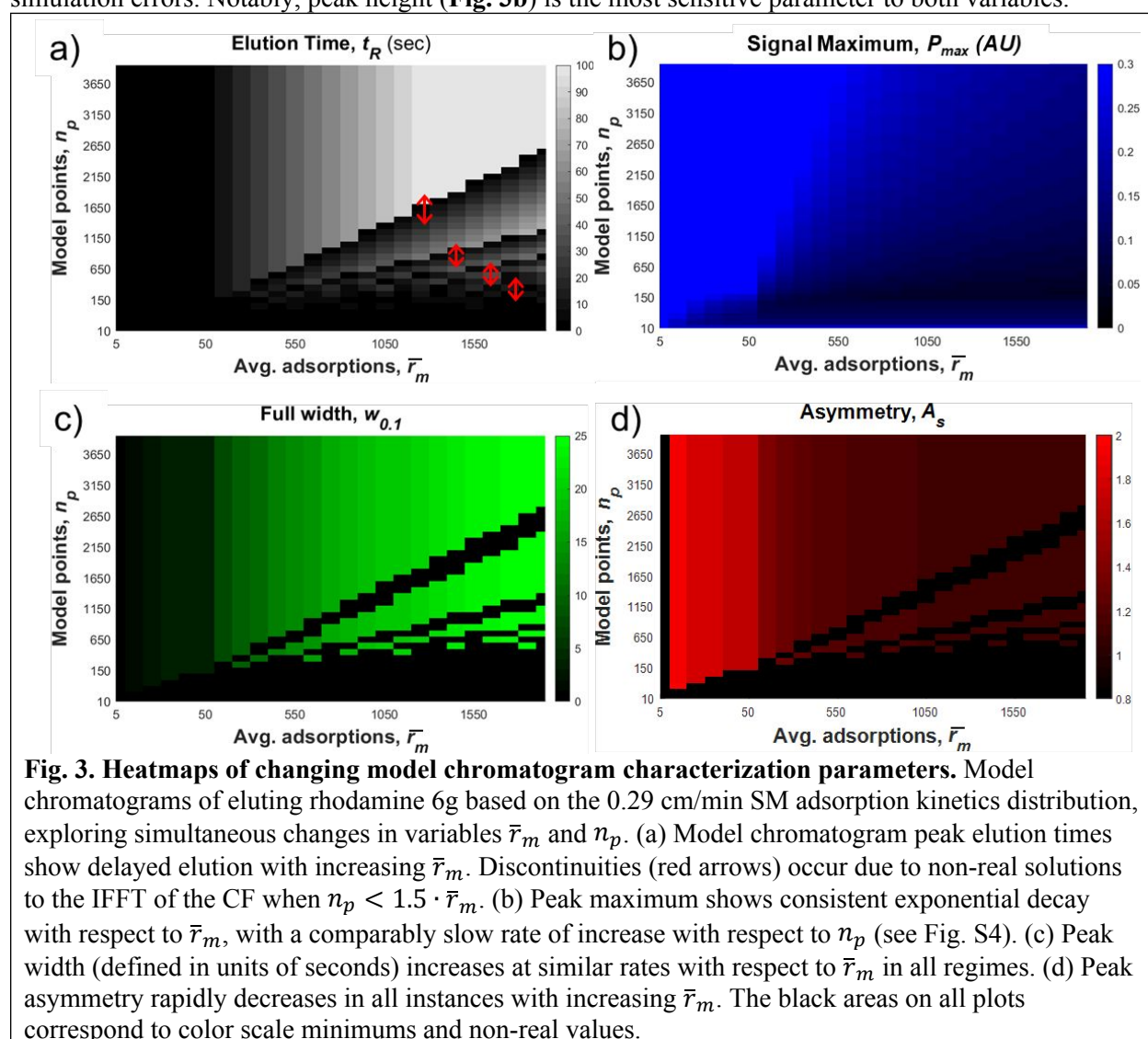
Cellulose B 5  $\mu\text{m}$  stationary phase particles (**Fig. S1**) to produce numerous elution profiles (**Fig. 2a, 2b**). We used a fixed large number of model points ( $n_p=5000$ ) relative to  $\bar{r}_m$  to avoid any model sampling issues (see section 4.2), and a full dataset of SM observations for our 0.29 cm/min flow velocity SM experiments.

Most of the peak characteristic parameters vary non-linearly in relation to increasing average number of single-molecule adsorptions. We observe a consistently increasing delay in elution time ( $t_R$ ), linearly proportional to the value of  $\bar{r}_m$  (**Fig. 2c**, black) and peak widening ( $w_{0.1}$ , **Fig. 2c**, green) following what matches a two-component exponential association curve shape (see **Fig. S3** for curve fits). Meanwhile, both the peak height ( $P_{max}$ , **Fig. 2c**, blue) and asymmetry ( $A_s$ , **Fig. 2c**, red) follow a two-component exponential decay (see Discussion). Initial observations from varying  $\bar{r}_m$  over a large range (5-2000) show that the elution time relation to  $\bar{r}_m$  is directly proportional, but suggest that wide and high asymmetry peaks are not accurately represented outside of low ( $\ll 200$ ) or very high ( $\gg 200$ ) average number of adsorptions. These factors indicate that a combination of long adsorption times and a high frequency of adsorptions can significantly dictate the peak shape.



4.2 Simultaneously representing the effects of  $n_p$  and  $\bar{r}_m$  on model elution peaks reveals a threshold for adequate sampling regimes and consistent  $\bar{r}_m$  dependence.

In addition to the adjustable variable  $\bar{r}_m$ , the sampling of the resulting elution profile is dictated by the user-defined variable,  $n_p$ . We simultaneously visualize how  $n_p$  and  $\bar{r}_m$  affect the elution time, peak asymmetry, width, and peak height of the modeled elution peaks (Fig. 3). The elution peak shape parameters demonstrate that in small  $n_p$  conditions ( $n_p \ll \bar{r}_m$ ), both elution time (Fig. 3a) and peak height (Fig. 3b) show variable rates of change with respect to  $\bar{r}_m$  depending on which region of the  $n_p$  model space sampling they are in. Closer inspection (see Fig. S4) suggests that the observed discontinuities (Fig. 3a, red arrows) correspond to undersampling that occurs in Fourier space, scaling as approximately  $1.5 \times \bar{r}_m$ , which results in non-real solutions to the inverse Fourier transform of the CF in eqn 8. Meanwhile, both peak width (Fig. 3c) and peak asymmetry (Fig. 3d) show no direct relation with  $n_p$  in comparison to  $\bar{r}_m$ , except for discontinuities in the under-sampled range ( $n_p > 1.5 \times \bar{r}_m$ ). All parameters show discontinuities except for peak height due to our definition as signal maximum to avoid simulation errors. Notably, peak height (Fig. 3b) is the most sensitive parameter to both variables.



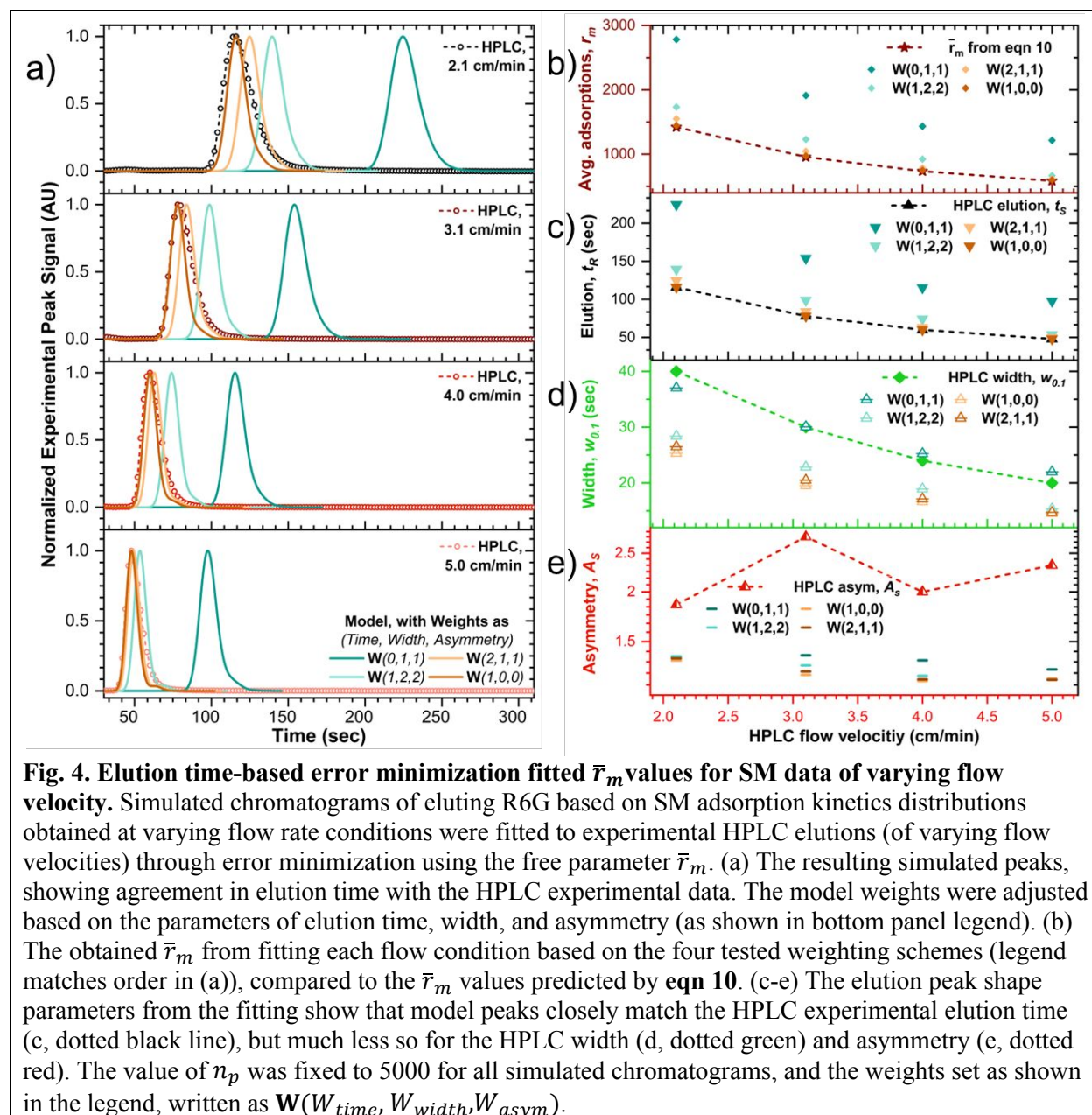
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

4.3 Error minimization can be used to select  $\bar{r}_m$  to accurately match HPLC elution times but does not capture peak width and asymmetry.

We test how the free parameter  $\bar{r}_m$  can be systematically adjusted to match experimental HPLC elution times with stochastic model data from measured SM adsorption time distributions. HPLC data was collected of the same rhodamine 6g analyte and Cellulose B stationary phase packed within a column as was imaged via SM microscopy. The  $\bar{r}_m$  values are determined through our error minimization fitting with respect to three chromatogram parameters (position, width, and asymmetry) compared between the model elution and HPLC data as detailed in section 3.3.

Using the SM dwell-time datasets of rhodamine 6g solution with flow velocities of comparable magnitude to the HPLC conditions (**Table S2**) through the stationary phases (**Fig. S1**), we test the model fitting results of using an error minimization approach, defined by **eqn 14** when we vary the weighting scheme of the error calculation. The weights were defined by the values of  $W_{time}$ ,  $W_{width}$ , and  $W_{asym}$  (which we hereto present in said order), with set values of 0, 1, or 2 to prioritize certain elution time vs shape characteristics (e.g. ' $W(2,1,1)$ ' corresponds to weights of  $W_{time} = 2$ ,  $W_{width} = 1$ , and  $W_{asym} = 1$ ). For error minimization fitting of the model, we tested the measured HPLC data in comparison to the SM data with the most similar flow velocity (**Fig. 4**). This corresponded to the  $2.9 \pm 0.5$  cm/min SM data for all HPLC conditions except for the  $5.0 \pm 1.0$  cm/min HPLC runs, which better coincide with the  $6.0 \pm 1.0$  cm/min SM experimental conditions (**Table S2**). By fitting the model with different parameter weights through adjustment of the free parameter  $\bar{r}_m$ , we obtain multiple simulated elution peaks that variedly match the HPLC elution time and shape (**Fig. 4a**).

The best agreement between the SM model and HPLC data (**Fig. 4a brown fits**) occurs when the model is weighted to prioritize elution time (i.e.  $W_{time} > W_{width}$  and  $> W_{asym}$ ). When only elution time is considered ( $W(1,0,0)$ ), the obtained  $\bar{r}_m$  values from error minimization without *a-priori* knowledge end up closely matching the theoretical values predicted by **eqn 10** (**Fig. 4b**), with good matching of the experimental elution times (**Fig. 4c**). The agreement based on elution time supports Pasti et al.'s derivation and parameter definitions (**eqn 10**, see **Figs. S5** and **S6**). By adjusting the error minimization weighting for width (**Fig. 4d**) or asymmetry (**Fig. 4e**), the HPLC peak shapes can be better approximated by the simulated chromatograms, but the elution time and shape could not be simultaneously matched by adjusting the value of  $\bar{r}_m$  (see **Fig. S7**). When the weights prioritize only peak shape, ignoring elution time, as  $W(0,1,1)$ , the resulting fits show high agreement in peak width (**Fig. 4d**, dark aquamarine), but less so for asymmetry (**Fig. 4e**). Interestingly, the gap between model and HPLC narrows for both width (**Fig. 4d**) and elution time (**Fig. 4c**) as the flow velocity increases, regardless of the chosen weighting scheme. Notably, asymmetry was the parameter most poorly captured by the model (**Fig. 4e**), with it having the least influence over the other weights. This corresponds with the inverse relation seen between these parameters in **Fig. 2C**.

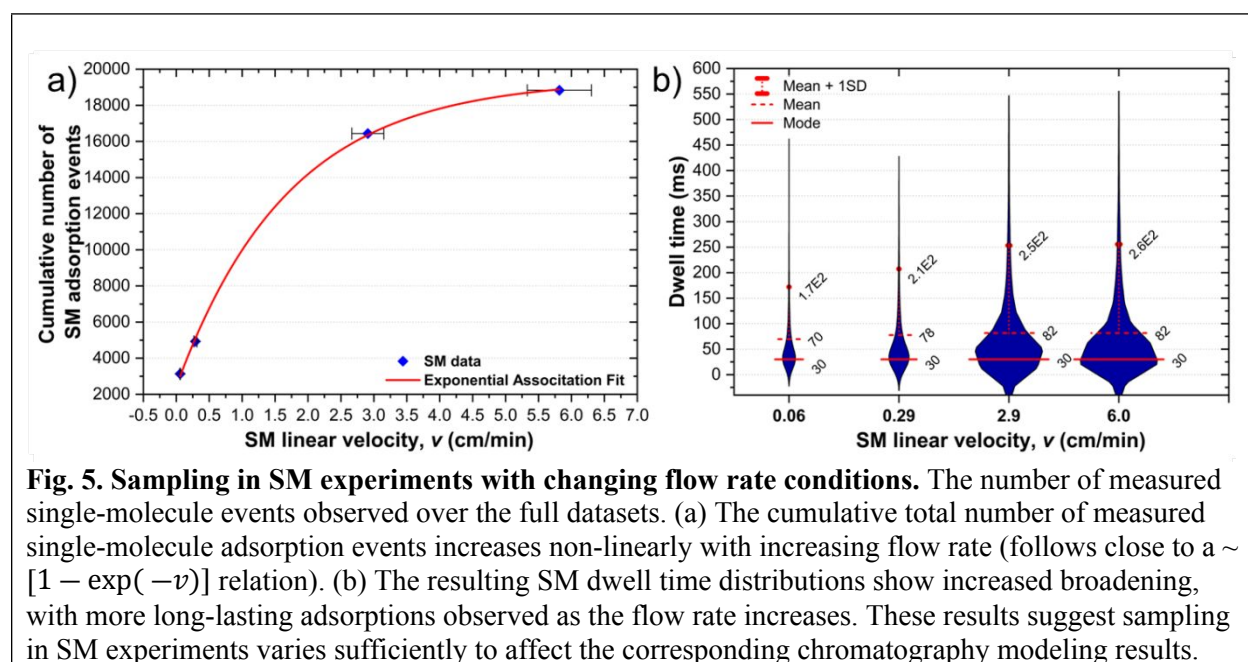


#### 4.4 Single-molecule experimental flow velocity conditions affect sampling of rare long-lasting adsorptions.

Increased sampling of rare, long desorption time events at high flow rates improves the observed agreements between HPLC and simulation trends with respect to flow rate (Figs. S5-8). All SM fluorescence microscopy measurements were done over the same observation time lengths (2000 frames per collection, for 6 collections, with 32 ms time resolutions), on the same 12 representative stationary phase particles (Fig. S1), the same buffer (20 mM HEPES buffer), at the same concentrations (1 nM rhodamine 6g), with only the varied flow velocity (Table S2) using a syringe pump. As the SM flow velocity increases, a larger population sample is obtained for the same total observation time. We observe that the growth in the number of SM observables increases following what can be described by an exponential association curve (Fig. 5a red fit;  $y = (2430 \pm 220) + (17000 \pm 300) \cdot$

( $1 - \exp\left(-\frac{x}{1.70 \pm 0.14}\right)$ ), with 0.999  $R^2$ ). This could be attributed to the increasing saturation of available adsorption sites on the imaged surfaces. As the number of sampled single molecules increases, the standard average and standard deviations in measured dwell times (**Fig. 5b**) also increase, despite the mode remaining constant. This would align with the wider and more asymmetric simulated peaks with increasing SM experimental flow rate (**Figs. S5-S8**). Furthermore, comparison of the linear velocities measured in the SM experiments and those in the HPLC column runs (**Table S2**) shows that different velocity regimes were approached, with the higher SM flow rates being closer to the HPLC conditions.

SM data collected at increased flow rates has increasing tailing more similar to the HPLC data. We observe that higher SM flow velocities result in wider peaks (**Fig. S5D** green, and **Fig. S7D** green) and increased asymmetry (**Fig. S5D** red, and **Fig. S7D** red), regardless of whether it is fitted by elution time or shape. The very distinct tailing visible in the simulated peaks that arises at the higher flow velocity datasets (**Fig. S5A, S5B, S7A, and S7B, blue lines**) demonstrate the significant influence that even rare SM adsorption times have on the model. We also find that these trends remain true across a wide range (5 to 2000) of  $\bar{\tau}_m$  values (**Fig. S8**), with datasets of higher SM flow velocities showing drastic differences in asymmetry when  $\bar{\tau}_m$  is large (**Fig. S8B**).



**Fig. 5. Sampling in SM experiments with changing flow rate conditions.** The number of measured single-molecule events observed over the full datasets. (a) The cumulative total number of measured single-molecule adsorption events increases non-linearly with increasing flow rate (follows close to a  $\sim [1 - \exp(-v)]$  relation). (b) The resulting SM dwell time distributions show increased broadening, with more long-lasting adsorptions observed as the flow rate increases. These results suggest sampling in SM experiments varies sufficiently to affect the corresponding chromatography modeling results.

## 5. Discussion

### 5.1 Non-linear behaviors in the stochastic model can be linked to experimental on-column and microscopic interactions.

Non-linear behaviors within the stochastic model, with a strong dependence on the average number of single molecule events ( $\bar{\tau}_m$ ), can be directly linked to physical, molecular behaviors occurring in a column. The elution time scales directly to  $\bar{\tau}_m$  (**Fig. 2c**, black), while peak width ( $w_{0.1}$ ) follows a two-component exponential association function (**Fig. 2c** and **S3**, green) with respect to  $\bar{\tau}_m$ . On a column, an increasing  $\bar{\tau}_m$  corresponds to an increasing number of available adsorption sites, which can occur with either an increased density of sites, increased accessibility, or a longer column. Then, the exponential

association behavior matches known peak broadening trends for longer columns<sup>39</sup> as well as recent Monte Carlo simulations<sup>8</sup> that show that increased site densities and column lengths can lead to similar non-linear peak broadening. In terms of equilibrium adsorption behavior, this increasing  $\bar{\tau}_m$  can be likened to increasing occupancy of adsorption sites, which scales non-linearly, as seen in Langmuir and similar isotherms.<sup>40</sup> Isotherms like the Jovanovic–Freundlich model combine kinetics with allowed surface heterogeneity and are often written in terms of exponential association functions of the form  $q(C) = q_0(1 - e^{-C^n})$ .<sup>41</sup>

In the transition from equilibrium modeling to adsorption kinetics, the Langmuir isotherm itself describes equilibrium (the steady state), while the approach to that equilibrium is often mathematically characterized by exponential functions, particularly in complex or heterogeneous systems.<sup>42</sup> Relatedly, we observe that simulated elution peak heights ( $P_{max}$ , **Fig. 2c** and **S3**, blue) follow a two-component decaying exponential function with increasing  $\bar{\tau}_m$ , which likely arises from the same number of measured single molecules (*i.e.* constant concentration) being spread out over a broader range of adsorption sites and total flight times – thus being inversely related to the peak width. In the field of single-molecule kinetics, it is not uncommon to describe dwell time distributions (**Fig. S1**) and similar datasets in terms of multi-component decay exponentials, which are attributed to the multiple heterogeneous forms of adsorption that can exist on a sample surface<sup>16,30,43,44</sup>.

The model peak asymmetry factor ( $A_s$ , **Fig. 2c** and **S3**, red) also follows a two-component decaying exponential behavior. Similar to peak broadening, this can be related back to the increasing number of sites leading to a lower degree of skewing due to non-mass transfer behavior (*e.g.* diffusion and convection) dominating the separation, with adsorption kinetics having a reduced effect. The total time spent in the column follows the relation  $t_R = \tau_m + \bar{\tau}_m \tau_s$ , which dictates that at low values of  $\bar{\tau}_m$ , the net elution time is dominated by time spent in the mobile phase.<sup>15</sup> Thus, as it can also be gathered from careful inspection of the CF (**eqn 6**), the initial degree of skewing is dominated by the frequency function of adsorption times at low number of average adsorptions.

### 5.2 Numerical discontinuities occur due to the mathematical construction for discretization of the model.

Through analysis of the model's output peaks with respect to the number of points used in a chromatogram, we reveal model discontinuities and a minimum value of  $n_p$  necessary to obtain consistent results. When sampling through  $n_p$  when  $\bar{\tau}_m$  is fixed, we observe that the model does not output singular, complete elution peaks at low numbers of model points (**Fig. S4**, **Fig. 3**) but then settles at a single elution time position at higher  $n_p$  values. By plotting the elution peak parameters (**Fig. S4c**, **Fig. 3**), it becomes evident that discontinuities can occur depending on the value of  $n_p$ , where the solution to the IFFT of the CF results in only non-real values. When simultaneously probing  $n_p$  and  $\bar{\tau}_m$  (**Fig. 3**) we observe that these discontinuities and inconsistencies occur throughout the whole variable space, until a threshold is passed, corresponding to approximately when  $n_p$  is at least 1.5 times greater than  $\bar{\tau}_m$ . The discontinuity at  $n_p < 1.5 \bar{\tau}_m$  likely directly results from the mathematical construction of the model (**eqn 6**) and the numerical limitations of the IFFT calculation itself, where the frequency term ( $\omega$ , in the Fourier domain) scales inversely to  $n_p$  (**eqn 9**). Given a sufficiently small  $n_p$  (relative to  $\bar{\tau}_m$ ), the value of the imaginary component of the IFFT (primarily  $\omega$  dependent) will dominate over the influence of the purely  $\bar{\tau}_m$  scaling components. Thus, we find that an important requirement ( $n_p > 1.5 \cdot \bar{\tau}_m$ ) exists in selecting a sufficiently high number of model points when applying this discrete form of the Lévy process representation of the stochastic model; a threshold and important caveat that to our knowledge was not discussed in the original report,<sup>15</sup> nor previous applications of the model.<sup>8,16,18</sup>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### 5.3 Weighted error minimization reveals the strengths and shortcomings of the model.

A moments-based error minimization approach (eqn 11-14) was implemented to accurately fit the model to HPLC elution data by adjusting the  $\bar{\tau}_m$  free parameter. By tuning the weighting scheme ( $W_i$  values in eqn 14) for the error minimization, we can prioritize fitting (Fig. 4) by elution time (i.e.  $W_{time} > W_{width}$  and  $> W_{asym}$ , see Fig. S5-6) or fitting by peak shape (i.e.  $W_{time} < W_{width}$  and  $< W_{asym}$ , see Fig. S7). We find that the model very accurately matches the HPLC elution time when  $\bar{\tau}_m$  is at or near the values predicted by the theory (Fig. 4c, eqn 10), corresponding to prioritizing elution time fitting, matching previous observations<sup>15,16</sup>, and demonstrating applicability across a range of different elution times with changing HPLC flow rates (Fig. S6). The agreement in elution time supports Pasti et al.'s original derivation and parameter definitions. Interestingly, we also observe that regardless of the chosen weighting scheme, the gap between model and HPLC narrows for both width (Fig. 4d) and elution time (Fig. 4c) as the flow velocity increases. However, both the elution peak shape and time cannot be fully matched simultaneously in this application of the model (Fig. S5D vs Fig. S7D). We find that when the weights prioritize only peak shape (e.g.  $W(0,1,1)$ ), the resulting fits show high agreement in peak width (Fig. 4d), but much less so for asymmetry (Fig. 4e), which never fully converges with the HPLC experiment. We suspect that the narrowing gap between model and HPLC with increasing flow velocity, and the inaccurate asymmetry point to the necessity to improve the model and SM experimental design as sources of peak broadening other than mass transfer (e.g. diffusion) are neglected in the derivation and formalism of the model, as well as in the SM experimental observations.

### 5.4 Limitations in experimental conditions between HPLC and SM can drive differences between the model and HPLC.

Experimental differences in SM and HPLC stationary phases that affect mass transfer cannot fully account for the inability of the stochastic model to accurately predict both elution time and shape. In Pasti et al.,<sup>15</sup> mismatches between single-molecule data and LC were attributed in part to the observed SM adsorption on a functionalized thin film being very different than the column materials. However, here we find that using the same stationary phases between SM and HPLC does not suffice to address the disagreement in SM model and LC data as significant diffusive differences in elution profiles can remain. To attempt to address this disconnect, Mabry et al. had included an additional fitting parameter to try to account for the non-observable, fast diffusion components in their SM experiments.<sup>16</sup> In a similar vein, additional terms and SM observables could be added to expand the model to include diffusional terms in future applications. We further note that  $\bar{\tau}_m$  is applied as a single average value in the discrete form of the CF; additional work could add a mathematically rigorous expansion to include a distribution of values that would be more representative of the on-column availability of adsorption sites.

SM experimental factors of sampling can affect agreement between simulated and HPLC elutions and can hint at sources of peak broadening in the on-column experiments. Higher SM flow velocity conditions (2.9 and 6.0 cm/min) show distinctively higher degrees of elution peak asymmetry (Fig. S5-S8), which also affects trends in  $\bar{\tau}_m$  behavior (Fig. S8B). The significantly lower sampling of SM adsorption events (Fig. 5a) that occurs at the lower flow conditions (0.06 and 0.29 cm/min) correspond to a lower likelihood of observing the rare long-lasting adsorption events (Fig. 1a, Fig. 5b) which disproportionately contribute to peak broadening and tailing,<sup>16</sup> and generate a wider distribution of SM elution times with more significant tailing (Fig. 5b). Peak tailing can also be linked to the effects that long observation times can have by introducing irreversible adsorption into the measured dwell time distributions (causing increasingly exponential cumulative distributions).<sup>45</sup> However, in our experiments we do not venture out to very long observation timescales (measurements are ~1 minute per collection), and our distributions do not show any dwell times >5.5 seconds (Fig. S1A), suggesting little to no observed irreversible adsorption. Further, here, the cellulose-based stationary phase exists as matrix that has been modified



with 3,5-dimethylphenyl carbamate groups at the hydroxy groups. With the choice of the small (~1 nm<sup>46,47</sup>) cationic rhodamine 6g dye analyte, the predominant interactions should be van der Waals interactions between hydrophobic moieties of the dye and the stationary phase and pi-pi-interactions, which result in primarily short-lived adsorption.<sup>48–50</sup> However, similarly to what has been observed in previous work<sup>44,51</sup> exposed silanol groups could remain on the modified silica surfaces that make up the solid support of the stationary phase, leading to the observed rare relatively long-lasting adsorptions.

Other sources of discrepancy between HPLC and our model results could be the effects of diffusion phenomena based on the mobile phase polarity and stationary phase porosity. While our HPLC and SM experiments both use the same stationary phase materials and analytes, they were ultimately performed under different mobile phase conditions due to experimental constraints, which can lead to significant differences in elution behavior.<sup>19</sup> Most notably, the addition of ethanol on columns that was necessary for any rhodamine analyte to reach the detector might lead to shorter overall adsorption times.<sup>52</sup> Increasing the eluent strength (primarily from addition of more polar solvent modifiers) has been shown to decrease analyte-stationary phase interactions, thereby increasing the effective contribution of analyte-mobile phase interactions (by saturating the stationary phase with mobile phase) for cellulose-type stationary phases.<sup>53,54</sup> The incorporation of the stronger eluent ethanol, as was done here, could result in increased peak broadening and asymmetric effects due to diffusion mechanisms, compared to the mass-transfer dominated SM measurements. As we have previously shown<sup>19</sup> the high degree of surface functionalization present on the cellulose-B “fully porous” stationary phases can limit accessibility to the porous network. Confined pores can cause flow channeling (**Fig. S1B**), which leads to flow heterogeneity and consequently peak broadening in the HPLC elution. However, as we demonstrated,<sup>19</sup> decreasing the degree of functionalization directly speeds up elution, despite the increased time spent diffusing within the porous stationary phases. This would mean that the longer retention on column can be directly attributed in large part to the high density of adsorptive sites causing more frequent adsorptions (i.e. higher average number of adsorptions,  $r_m$ ), matching the behavior presented by this model. Overall, differences between SM stochastic modeled elution and HPLC data can point to factors beyond simple adsorption-desorption mass transfer contributing to the elution.

Further consideration should be given to the diffusion mechanisms, driven by flow conditions, that affect on-column separations. In our microscopy flow cells, we are limited by pressure, which we estimate<sup>19</sup> to be well under 1,000 psi, with our shown flow rates (**Table S2**) corresponding to ~678, 681, 714, and 750 psi (or ~47 to 52 bar). Higher pressures in HPLC have been associated with improved separation factors, which is partly attributed to conformational changes in the case of proteins.<sup>55</sup> However, pressure related effects become less pronounced for small molecules, and indeed the primary components of the dwell times distributions (**Fig. S1A**) in our system –which are the primary data used within the model –remain mostly unaffected by increasing pressure (i.e. flow rate). Another consideration is the corresponding shearing rates (**Table S2**) that occur due to flow near the walls, which have been shown<sup>56</sup> to lead to increased  $k_a$  values (i.e. shorter association times), which we also observe (**Fig. S1C**).

Overall, we believe that future work should aim to better bridge these gaps between HPLC and SM experimental conditions, as well expand the model to more rigorously include diffusive information. At present we are working to better replicate the high-pressure packed column environment of HPLC columns in our microscopy set-up. We are working towards imaging at >30,000 psi using capillary-based approaches developed in biophysics<sup>57</sup>, trying to explore open questions about diffusive behaviors and the influence of back pressure. Future research should focus on quantifying Eddy diffusion in packed beds and membranes while pushing for higher operating pressures. By simultaneously enhancing microscopic resolution and speed, we can better connect individual molecular behaviors to full-scale column

performance. Simultaneously, the model should be expanded to present  $\bar{r}_m$  as a distribution of values (rather than an average) and expand the Brownian components of the Lévy process representation to better include diffusion information based on experimental conditions.

## 6. Conclusions

In this work, we systematically explored a wide range of the variables  $\bar{r}_m$  (the average number of single molecule adsorptions) and  $n_p$  (the number of modeling points) in the stochastic model of chromatography to determine if it is possible to optimize and achieve agreement between SM modeled and HPLC chromatograms from measurements carried out on the same commercial stationary phase materials under varying flow velocity conditions. We also introduced a metric to quantify the agreement between the single-molecule and HPLC by taking the weighted difference in the elution time and shape of the chromatograms and explored the effects that experimental sampling can have on the model outputs. We showed how elution peak shapes vary depending on the average number of SM adsorptions ( $\bar{r}_m$ ), as well as the number of modeling points ( $n_p$ ). Our results reveal two-component exponential association curve behaviors occur in peak broadening (width), and two-component decaying exponentials in peak height and asymmetries, as well as under sampling that can occur due to the choice of  $n_p$ , setting a minimum threshold value of  $n_p > 1.5 \cdot \bar{r}_m$ . An error minimization approach was also implemented and simulated elution times and peak shape can be independently matched to experimental HPLC data, which was applied to new experimentally measured SM adsorptions on commercial stationary phase particles and comparable HPLC on-column experiments.<sup>19</sup> Sampling issues that can occur in experimental SM measurements were also explored. We conclude, below, with some suggestions for future implementation of the model.

From our observations, we summarize the following recommendations for other separation scientists interested in applying the stochastic model to single-molecule data in the future. 1) The experimental SM observation time scales should be significantly longer than the scale of the adsorption times, with a high degree of sampling of single-molecule events. For this we recommend tailoring both analyte concentrations, flow conditions, number of available adsorptions sites (*e.g.* number of stationary phase particles in the FOV or increased number of samples), and total observation time to increase the robustness and statistics of the SM datasets. Collecting data at or near conditions of the plateau of exponential association curve, as in **Fig. 5a**, are advised. 2) The model free parameter  $\bar{r}_m$  can be fit with weighted error towards either elution time or shape, but not both simultaneously. Given that elution time has a more direct theory-based connection with  $\bar{r}_m$ , it is the recommended fitting target. If the SM observations are sufficiently robust as noted in the first point, agreement with elution peak shape should improve. 3) The number of model points should be scaled with the average number of adsorption events, following the rule of  $n_p > 1.5 \cdot \bar{r}_m$  at a minimum. For our application, we have kept  $n_p > 2 \cdot \bar{r}_m$  as a minimum to avoid any Fourier space undersampling. Beyond direct application of the model, future attempts to connect SM experimental observables and HPLC experiments should continue to bridge the gap between their experimental conditions, such as by matching mobile phase conditions, including pressure and linear velocities, more closely. Finally, further development of the stochastic model can be implemented by including terms to help add diffusional information into the microscopic theory of elution,<sup>16</sup> extending beyond the focus on mass transfer kinetics, as well as designing SM experiments to measure these diffusion parameters.


### Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### CRedit authorship contribution statement

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Open Access Article. Published on 01 May 2022. Downloaded on 5/22/2022 4:40:40 AM.  
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.



Ricardo Monge Neria: Conceptualization, Investigation, Visualization, Writing – original draft, Writing – review & editing.

Rachel A. Saylor: Investigation, Resources, Writing – review & editing.

Lydia Kisley: Conceptualization, Resources, Writing –review & editing, Project administration, Funding acquisition.

### Acknowledgements

Oberlin College provided financial support for RAS's contributions. The Research Corporation for Science Advancement Cottrell Scholar fund and 3M Non-tenured Faculty Award provided financial support for RMN, and we acknowledge funding from the US Department of Energy, Basic Energy Science, Separation Science (DE-SC0025607) for partial support of LK's contributions.

### Supplementary materials

Supporting information document (PDF), including Figs. S1-S5.

Matlab code package found in: [will add Zenodo DOI at time of publication]

### Data availability

Data will be made available on request.

### Appendix A


Table A1. Summary of important acronyms and variables and their meanings.			
SM	Single molecule	$\tau_s$	Time spent in the stationary phase
HPLC	High performance liquid chromatography	$\tau_m$	Time spent in the mobile phase
R6G	Rhodamine6G	$t_R$	Elution time
CF	Characteristic function	$t_0$ or $t_{void}$	Detector void time for unretained solutes. Only applies for experimental HPLC data.
FFT	Fast Fourier Transform	$P_{max}$	Elution peak height.
IFFT	Inverse Fast Fourier Transform	$w_{0.1}$	Peak width defined by tangential method at 10% peak maximum.
$\bar{r}_m$	Average number of single-molecule adsorptions	$A_s$	Peak asymmetry, defined as the ratio of the right over left widths of the peak.

## References

- 1 S. Leweke and E. von Lieres, *Comput. Chem. Eng.*, 2018, **113**, 274–294.
- 2 A. Felinger and A. Cavazzini, *Liq. Chromatogr. Fundam. Instrum. Second Ed.*, 2017, **1**, 17–37.
- 3 K. Miyabe, *TrAC Trends Anal. Chem.*, 2016, **81**, 79–86.
- 4 F. Rischawy, D. Saleh, T. Hahn, S. Oelmeier, J. Spitz and S. Kluters, *Comput. Chem. Eng.*, 2019, **130**, 106532.
- 5 J. C. Giddings and H. Eyring, *J. Phys. Chem.*, 1955, **59**, 416–421.
- 6 F. Dondi, A. Cavazzini and L. Pasti, *J. Chromatogr. A*, 2006, **1126**, 257–267.
- 7 L. D. C. Bishop, A. Misiura and C. F. Landes, *Analyst*, 2021, **146**, 4268–4279.
- 8 N. Kovalenko and C. F. Landes, *Anal. Chem.*, 2025, **97**, 20237–20246.
- 9 F. Gritti, *Anal. Chem.*, 2021, **93**, 5653–5664.
- 10 S. W. Foster, N. Wright, J. P. Grinias and L. M. Blumberg, *J. Chromatogr. A*, 2021, **1659**, 462645.
- 11 C. R. Bernau, M. Knödler, J. Emonts, R. C. Jäpel and J. F. Buyel, *Front. Bioeng. Biotechnol.*, 2022, **10**, 1009102.
- 12 A. Felinger, *J. Chromatogr. A*, 2008, **1184**, 20–41.
- 13 L. Pasti, N. Marchetti, R. Guzzinati, M. Catani, V. Bosi, F. Dondi, A. Sepsey, A. Felinger and A. Cavazzini, *Trends Anal. Chem.*, 2016, **81**, 63–68.
- 14 A. Felinger, A. Cavazzini and F. Dondi, *J. Chromatogr. A*, 2004, **1043**, 149–157.
- 15 L. Pasti, A. Cavazzini, A. Felinger, M. Martin and F. Dondi, *Anal. Chem.*, 2005, **77**, 2524–2535.
- 16 J. N. Mabry, M. J. Skaug and D. K. Schwartz, *Anal. Chem.*, 2014, **86**, 9451–9458.
- 17 L. D. C. Bishop and C. F. Landes, *Acc. Chem. Res.*, 2018, **51**, 2247–2254.
- 18 L. Kisley, J. Chen, A. P. Mansur, B. Shuang, K. Kourentzi, M. V. Poongavanam, W. H. Chen, S. Dhamane, R. C. Willson and C. F. Landes, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 2075–2080.
- 19 R. Monge Neria, M. Zeeshan, A. Kapoor, T. K. J. Kim, N. Hoven, J. S. Pigott, B. Gurkan, C. E. Duval, R. A. Saylor and L. Kisley, *Sci. Adv.*, 2023, **11**, 790.
- 20 N. A. Moringo, L. D. C. Bishop, H. Shen, A. Misiura, N. C. Carrejo, R. Baiyasi, W. Wang, F. Ye, J. T. Robinson and C. F. Landes, *Proc. Natl. Acad. Sci. USA*, 2019, **116**, 22938–22945.
- 21 M. J. Wirth, D. J. Swinton and M. D. Ludes, *J. Phys. Chem. B*, 2003, **107**, 6258–6268.
- 22 J. C. Giddings, in *Dynamics of Chromatography: Principles and Theory*, Marcel Dekker, Inc., New York, 1st edn., 1965, pp. 13–94.
- 23 R. A. Keller and J. C. Giddings, *J. Chromatogr. A*, 1960, **3**, 205–220.
- 24 A. Felinger and A. Cavazzini, *Liq. Chromatogr. Fundam. Instrum.*, 2013, 19–40.
- 25 F. Gritti and G. Guiochon, *J. Chromatogr. A*, 2013, **1302**, 1–13.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Open Access Article. Published on 01 May 2026. Downloaded on 5/22/2026 4:40:44 AM.  
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.



- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26 D. A. McQuarrie, *J. Chem. Phys.*, 1963, **38**, 437–445.  
27 A. Cavazzini, M. Remelli, F. Dondi and A. Felinger, *Anal. Chem.*, 1999, **71**, 3453–3462.  
28 A. Cavazzini, F. Dondi, A. Jaulmes, C. Vidal-Madjar and A. Felinger, *Anal. Chem.*, 2002, **74**,  
6269–6278.  
29 L. Pasti, F. Dondi, M. Van Hulst, P. J. Schoenmakers, M. Martin and A. Felinger,  
*Chromatographia*, 2003, **57**, S171–S186.  
30 M. J. Wirth and D. J. Swinton, *Anal. Chem.*, 1998, **70**, 5264–5271.  
31 S. H. Kang, M. R. Shortreed and E. S. Yeung, *Anal. Chem.*, 2001, **73**, 1091–1099.  
32 R. Monge Neria and L. Kisley, *Anal. Chem.*, 2023, **95**, 2245–2252.  
33 J. Chen, A. Bremauntz, L. Kisley, B. Shuang and C. F. Landes, GitHub -  
KisleyLabAtCWRU/SuperResKinetics: Super-resolution imaging and single molecule kinetics  
MATLAB analysis, <https://github.com/KisleyLabAtCWRU/SuperResKinetics>, (accessed 5 June  
2022).  
34 R. Parthasarathy, *Nat. Methods*, 2012, **9**, 724–726.  
35 F. Ghanem and K. M. Yenkie, *Ind. Eng. Chem. Res.*, 2025, **64**, 7527–7539.  
36 I. Kiss, I. Bacskay, F. Kilár and A. Felinger, *Anal. Bioanal. Chem.*, 2010, **397**, 1307–1314.  
37 K. Miyabe, A. Arai and M. Ishizuka, *Anal. Chem.*, 2021, **93**, 10365–10371.  
38 J. P. Foley and J. G. Dorsey, *Anal. Chem.*, 1983, **55**, 730–737.  
39 M. Naefe and I. Halász, *Anal. Chem.*, 1972, **44**, 76–84.  
40 Y. Liu, *Colloids Surfaces A Physicochem. Eng. Asp.*, 2006, **274**, 34–36.  
41 I. Quiñones and G. Guiochon, *J. Chromatogr. A*, 1998, **796**, 15–40.  
42 P. Patil, G. Jeppu, C. R. Girish and B. Mohan, *Sep. Sci. Technol.*, 2024, **59**, 373–394.  
43 R. Walder, M. Kastantin and D. K. Schwartz, *Analyst*, 2012, **137**, 2987–2996.  
44 M. J. Wirth and M. A. Legg, *Annu. Rev. Phys. Chem.*, 2007, **58**, 489–510.  
45 K. Lyu, H. Chen, J. Gao, J. Jin, H. Shi, D. K. Schwartz and D. Wang, *Biomacromolecules*, 2022,  
**23**, 4709–4717.  
46 Z. Chen, Y.-J. Tang, T.-T. Xie, Y. Chen and Y.-Q. Li, DOI:10.1007/s10895-007-0241-x.  
47 M. Chapman and W. B. Euler, *J. Fluoresc.*, 2018, **28**, 1431–1437.  
48 Y. Liu, X. Wang, J. Yu and X. Guo, *Electrophoresis*, 2021, **42**, 1461–1472.  
49 S. Cao, Q. Ma, Y. Liu, J. Zhang and Z. Wang, *New J. Chem.*, 2022, **46**, 9704–9709.  
50 Y. Liu, L. Cai, J. Lun, M. Zhao and X. Guo, *New J. Chem*, 2020, **44**, 18337.  
51 M. J. Wirth, M. D. Ludes and D. J. Swinton, *Anal. Chem.*, 1999, **384**, 3911–3917.  
52 S. Khater, Y. Zhang and C. West, *J. Chromatogr. A*, 2014, **1363**, 294–310.  
53 L. A. Papp, Z. I. Szabó, G. Hancu, L. Farczádi and E. Mircia, *Molecules*, 2024, **29**, 1346.



1  
2  
3 Data availability statement:

4  
5 Data needed to evaluate the conclusions in the paper are present at Zenodo DOI:  
6 [10.5281/zenodo.13697196](https://doi.org/10.5281/zenodo.13697196) and DOI: [10.1126/sciadv.ads0790](https://doi.org/10.1126/sciadv.ads0790).

7  
8  
9 Code for analyzing the raw microscopy data and simulating chromatograms are available  
10 at <https://github.com/KisleyLabAtCWRU/>.

11  
12 Any additional data or analysis requests can be directed to [lydia.kisley@case.edu](mailto:lydia.kisley@case.edu).

13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60