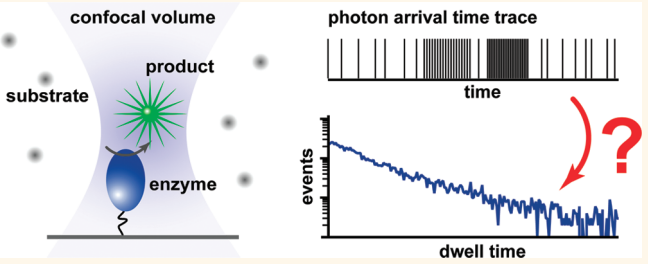
Dynamic Disorder in Single-Enzyme Experiments: Facts and Artifacts

Tatyana G. Terentyeva,[†] Hans Engelkamp,[‡] Alan E. Rowan,[‡] Tamiki Komatsuzaki,[§,*] Johan Hofkens,[†,*] Chun-Biu Li,[§,*] and Kerstin Blank[‡,*]

[†]Photochemistry & Spectroscopy, Department of Chemistry, Katholieke Universiteit Leuven, Leuven, Belgium, [‡]Institute for Molecules and Materials, Radboud University Nijmegen, Nijmegen, The Netherlands, and [§]Molecule & Life Nonlinear Sciences Laboratory, Research Institute for Electronic Science, Hokkaido University, Sapporo, Japan

ABSTRACT



Using a single-molecule fluorescence approach, the time series of catalytic events of an enzymatic reaction can be monitored, yielding a sequence of fluorescent "on"- and "off"-states. An accurate on/off-assignment is complicated by the intrinsic and extrinsic noise in every single-molecule fluorescence experiment. Using simulated data, the performance of the most widely employed binning and thresholding approach was systematically compared to change point analysis. It is shown that the underlying on- and off-histograms as well as the off-autocorrelation are not necessarily extracted from the "signal" buried in noise. The shapes of the on- and off-histograms are affected by artifacts introduced by the analysis procedure and depend on the signal-to-noise ratio and the overall fluorescence intensity. For experimental data where the background intensity is not constant over time we consider change point analysis to be more accurate. When using change point analysis for data of the enzyme α-chymotrypsin, no characteristics of dynamic disorder was found. In light of these results, dynamic disorder might not be a general sign of enzymatic reactions.

KEYWORDS: single-molecule fluorescence · enzyme kinetics · change point analysis · photon arrival time series · dynamic disorder · protein dynamics

Enzymes, and proteins in general, are molecules whose function is dynamically coordinated by their complex energy landscapes. The observation of individual enzyme molecules enables us to capture dynamic processes that are buried in ensemble measurements. It is one of the ultimate goals of single-molecule experiments to relate a time series of events to the energy landscape of an enzyme and to construct a kinetic scheme of the catalytic reaction.[1−5]

Single-molecule fluorescence spectroscopy is particularly well suited to study the dynamic behavior of enzymes.[6,7] A variety of different fluorescent reporter systems have been developed to study conformational changes during the catalytic reaction cycle or the kinetics of the catalytic reaction itself. Due to its strong distance sensitivity, single-molecule Förster resonance energy transfer (smFRET) is ideal for monitoring conformational changes on various time scales.[8−10] smFRET has further been used to indirectly measure the oxidation state of a cofactor during the catalytic reaction.[11,12] Alternatively, if the cofactor is fluorescent itself, alterations in cofactor fluorescence provide a direct readout of the catalytic turnover cycle.[13] A different and more general approach is the use of fluorogenic substrates that are converted by the enzyme into fluorescent product molecules.[14−18]

In many experiments where the time sequence of enzymatic turnovers was analyzed, the data suggested that the catalytic activity was not constant over time.[13−18] This observation was explained by the pres ence of different enzyme conformations leading to parallel reaction pathways that are each characterized by different rate constants for the catalytic reaction. In such a case the rate constant for the rate-limiting step might become a function of time, a phenomenon called dynamic disorder.[19] Dynamic disorder has been concluded for different enzymes and has been discussed to be a general property of enzymes.

To faithfully identify dynamic disorder in single-enzyme experiments and to analyze the underlying process on the energy landscape, data sets with a large number of enzymatic turnovers are required.[20,21] Fluorogenic substrates are ideal reporter systems, as every enzymatic turnover reaction releases a new fluorescent dye molecule that can, for example, be detected with a

* Address correspondence to k.blank@science.ru.nl, cbli@es.hokudai.ac.jp, Johan.Hofkens@chem.kuleuven.be, tamiki@es.hokudai.ac.jp.

Received for review September 24, 2011 and accepted December 1, 2011.

Published online December 01, 2011
10.1021/nn203669r

confocal microscope. One of the challenges of fluorogenic substrates, however, is the very short residence time of the fluorescent product molecules in the confocal volume, resulting in very short fluorescent on-states. Since the dye is not associated with the enzyme, it leaves the detection volume quickly after it has been produced. Short on-states separated by long off-states are difficult to identify accurately during subsequent data analysis.

The most frequently used data analysis procedure relies on binning the data followed by the application of a threshold that separates the short high-intensity on-states from the off-states. In a recent systematic study on quantum dot data it was shown that both the choice of the bin size and the threshold position have a large influence on the obtained off-histograms.[22] Statistical methods to identify intensity change points are an interesting alternative, as they do not require any subjective input.[23−25] So far, change point detection methods have rarely been applied to experimental data.[10,12,26] We have performed a systematic comparison of the commonly used binning/thresholding method with change point analysis. We have analyzed the performance and the limitations of both methods using simulated single-enzyme time traces with different signal-to-noise ratios and different intensity levels. The analysis results have been compared using the dwell time histograms of the on- and off-times as well as the off-time autocorrelation, as dynamic disorder has so far been linked to a stretched exponential off-time distribution and the presence of correlations between consecutive off-times. We have further compared the performance of both methods using experimental data of the enzyme α-chymotrypsin.

RESULTS AND DISCUSSION

Generation of the Simulated Data Sets. To simulate single enzyme data, we have chosen a kinetic scheme that can be considered as the simplest Michaelis−Menten-based scheme capable of producing dynamic disorder.[27] The kinetic scheme shows two conformations of the enzyme with different rates for the catalytic reaction (Figure 1). These conformations can interconvert at every step in the enzymatic reaction cycle with the rate constants α, β, and γ, respectively. To simplify the scheme, we have chosen identical rate constants for the forward and the reverse reactions of the conformational transitions. Using the Markovian property of the kinetic scheme, the time series of "on"- and "off"-events was simulated with a Monte Carlo algorithm for two sets of rate constants (see SI). In model 1 the rate constants for the conformational changes were chosen such that they are similar to the rate constants of substrate binding at a given substrate concentration (10^2−10^3 s^{-1}) but faster than those of the catalytic reactions ($k_{2a}/k_{2b} \approx 10$ s^{-1}). In model 2 the rate

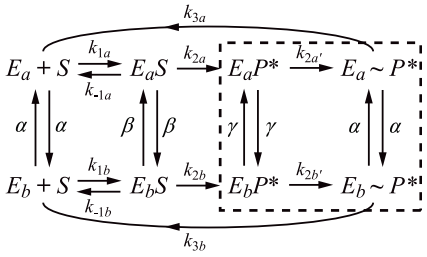


Figure 1. Kinetic scheme for the enzymatic conversion of a fluorogenic substrate into a fluorescent product at the single-molecule level. The scheme describes a typical reaction following Michaelis−Menten kinetics. The first step resembles substrate binding. Here $k_{1a} = \kappa_{1a}[S]$ and $k_{2a} = \kappa_{2a}[S]$, where κ_{1a} and κ_{2a} are the rate constants of substrate binding and [S] is the substrate concentration. The second step corresponds to the chemical conversion of substrate into product (ES → EP*). The third step (EP* → E∼P*) was introduced to account for the residence time of the generated fluorescent product molecule in the detection volume. Both enzyme conformations (E_a and E_b) are able to catalyze the reaction, but with different rate constants. They can interconvert at any state along the reaction pathway with the rate constants α, β, and γ, respectively. The states within the box contain the fluorescent product P* and represent the "on"-state of the reaction. The other states are considered as the "off"-state.

constants for the conformational transitions (0.5−6 s^{-1}) are slower than the rate constants for the catalytic reactions ($k_{2a}/k_{2b} \approx 10$ s^{-1}). Overall, the rate constants for the conformational change differ by a factor of 100−1000 between model 1 and model 2. Model 2 is adequate to describe a process with dynamic disorder, as the rate constants for the conformational transitions are smaller than any other rate constant in the kinetic scheme.

The corresponding dwell time histograms as well as the off-time autocorrelation graphs are shown in Figure S1. For model 1 the off-histogram shows a maximum, and no correlations between the off-times exist. In contrast, the off-histogram for model 2 appears slightly concave, and correlations between consecutive off-times are observed. These correlations are expected for model 2, as the conformational transitions are the slowest process.

In the next step, Poisson statistics were used to simulate photon arrival time traces (see SI). Poisson noise is a reasonable assumption for both the off- and the on-state, although it needs to be considered that the experimental photon-counting statistics of the on-state are also influenced by other factors such as dye blinking, rotation, and diffusion. In this way 16 data sets with different signal-to-noise ratios (S:N = 1.5, 2.5, 3.5, or 4.5) and different background intensity levels (2000, 4000, 6000, or 8000 photons/second) were generated.

Analysis of the Data with the Binning/Thresholding Approach. Binning and thresholding are performed by first binning the photon arrival time trace to obtain an intensity time trace (Figure 2). Subsequently, an intensity histogram is constructed and a cutoff value (threshold) is identified to separate the bins with an intensity above

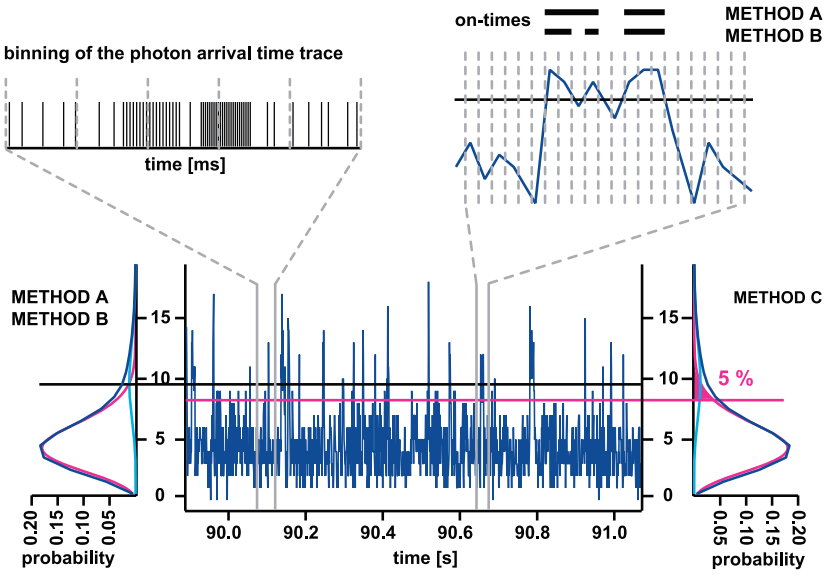


Figure 2. Data analysis with binning and thresholding. The photon arrival time trace is first binned to yield a fluorescence intensity (number of photons/time) time trace. In order to identify the threshold position, an intensity histogram is constructed. This histogram is a sum of the Poisson distribution of the background (off-state) and of the signal of the generated fluorescent dye (on-state). For overlapping histograms the intersection of these Poisson distributions is identified and the threshold is put at this position (methods A and B). As a noisy on-state might be separated into several on-states, an "interpeak" correction was further performed in method A: off-states of only one (2 ms) or four (5 ms) bins are not counted as "off" but instead are treated as part of the on-state (see top right). For method B no such correction was performed. In method C the threshold was determined such that less than 5% of the "off"-bins are counted as "on".

the threshold ("on"-bins) from those with an intensity below ("off"-bins). Finally, consecutive "on"-bins are combined to yield an on-time and consecutive "off"-bins are combined to yield an off-time.

When applied to experimental data, there is no unique solution to determine the bin size and the threshold position, and the values are often chosen arbitrarily. For simulated data the intensity levels of both the off-state (background) and the on-state (signal) are known. It is therefore possible to calculate exactly the off-Poisson and the on-Poisson distributions underlying the intensity distribution, facilitating a systematic comparison. Determining these distributions for different bin sizes can provide a basis for setting the threshold (Figure 2).

For our analysis we have chosen three different bin sizes and three different threshold methods. For methods A and B the threshold was placed at the intensity value where the calculated background distribution and the signal distribution intersect. It can be easily shown that the smallest amount of bins will be assigned incorrectly at this threshold position. If the distributions are not overlapping (*i.e.*, for good S:N ratios), the threshold can be put at any position between the two distributions. For method A a correction was used to prevent the separation of "on"-states into several shorter ones if their intensity fluctuates around the threshold value (Figure 2, top right). For method C the threshold was set at a position ensuring that less than 5% of the "off"-bins were counted as "on"-bins. Data sets where signal and noise intensity distributions were not overlapping have not been analyzed with method C, as this would artificially cause the detection of 5% false "on"-bins even if the actual overlap between the histograms yields less than 5% false "on"-bins.

We observe that both the bin size and the threshold position have a large influence on the shape of the off- and the on-histograms (Figure 3a and b). Only with the smallest bin size of 1 ms is the overall slope of the underlying off-histogram reproduced, while short off-times are overestimated, resulting in a deviation from single-exponential kinetics (see the inset of Figure 3). The origin of this artifact in the short off-time regime that might be misinterpreted as a stretched exponential is explained as follows: fluctuating on-levels are indeed segmented into several short on- and off-states depending on the position of the threshold (Figure 2 and SI). The number of off-events in this short time regime consequently depends on the correction used to account for on-state intensity fluctuations (*cf.* methods A and B in Figure 3b). Likewise, the number of short on-times is overestimated and depends on the use of the correction. The result that the threshold segments on-times into several short on- and off-times is further supported by the observation that the deviations from the underlying true histogram become smaller for the higher S:N ratios (Figure 3c). A similar improvement is observed when increasing the overall intensity (Figure 3d). Even for a S:N ratio of 2.5 an increase in the total number of photons leads to a reduction of the number of short dwell times. A more detailed

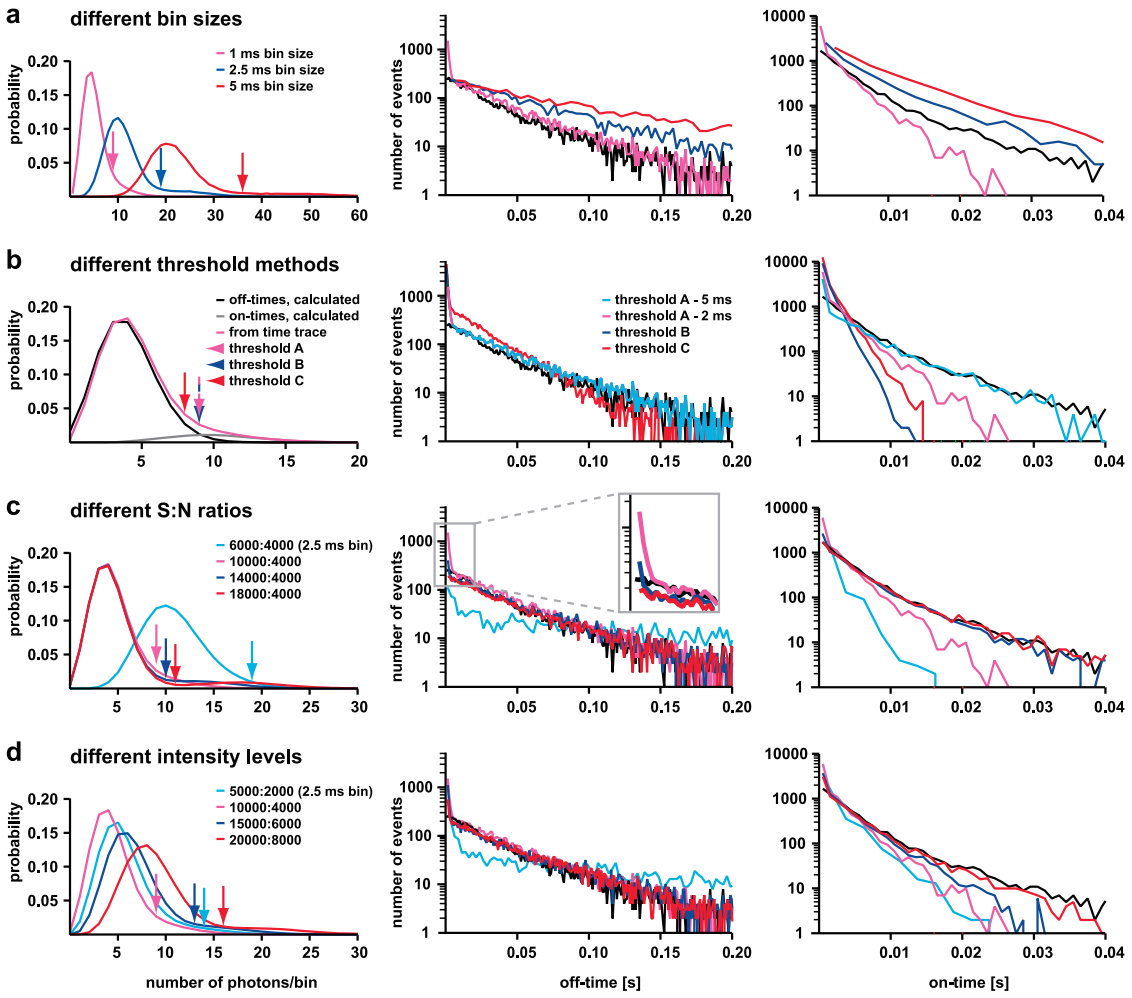


Figure 3. Analysis of simulated data (model 2) using threshold analysis. The dwell time histograms have been obtained for both the off-states and the on-states of the enzymatic reaction and compared to the underlying histograms (black line). The left column shows the intensity distributions including the threshold values, whose positions are indicated by arrows. The middle and the right columns show the off- and on-histograms, respectively. (a) Bin size dependence for one simulated data set (S:N = 10 000:4000 photons/s; method A with 2 ms interpeak correction). (b) Different threshold methods for the same simulated data set (S:N = 10 000:4000 photons/s; 1 ms bin size). (c) Different S:N ratios analyzed with method A (2 ms interpeak correction; 1 ms bin size). (d) Different intensity levels for a S:N ratio of 2.5 analyzed with method A (2 ms interpeak correction; 1 ms bin size). Note that in all cases the difference is bigger for the on-histograms than for the off-histograms. For the off-histograms the main difference is observed for the number of short off-times (inset).

discussion of the performance of the binning/thresholding approach is given in the SI.

The use of the above criteria for choosing the binning and thresholding parameters (smallest possible bin size and threshold value at the intersection between the signal and the background distribution) eliminates some of the arbitrariness of threshold analysis. It remains difficult, however, to determine the threshold position in experimental data because the underlying intensity levels of the on- and off-states are not known *a priori*. Furthermore, even if the smallest possible bin size is used, binning generally introduces an artificial time scale, and the shortest on- and off-times that can be resolved are limited by the bin size used. Bins flanking the on-states might have an intermediate intensity value as the intensity changes within the bin. This complicates the analysis even further.

Analysis of the Data with Change Point Analysis. Change point analysis has the potential to overcome these limitations.[23−25] Change point analysis statistically evaluates the likelihood for each photon to be an intensity change point. This is achieved by comparing the photon statistics before and after each photon. A photon is assigned to be an intensity change point based on the following statistical hypothesis test: the log-likelihood ratio (LLR) of the probability of being a change point and the probability of not being a change point is larger than a chosen confidence level. In order to detect multiple change points in the photon trace, the above procedure is applied recursively by binary segmentation at the change point photon. As a result, change point detection provides the time sequence of photons where intensity changes are most likely to occur (see SI for details). The hypothesis test scheme

proposed by Watkins *et al.*[23] is chosen in this work, as it has performed better practically in identifying change points in our simulated data even in low S:N cases. A detailed comparison of the performance of these methods will be discussed elsewhere.

Change point detection is usually followed by a clustering algorithm to determine the number of distinct intensity levels and to assign each of the detected change point intervals to the appropriate intensity level. Figure 4a and b exemplify these change point intervals plotted together with their corresponding intensities. The clustering of the change point intervals could be performed by introducing a certain boundary on the intensity-interval time plane: for the data with the larger S:N ratio (Figure 4b) one can see approximately two different peaks that might allow the assignment of a plausible boundary. This introduces a subjective step into the analysis, however. Also in the case of low S:N ratios (Figure 4a) no clear boundary can be identified. Here, we therefore introduce a new, efficient and simple clustering algorithm for the assignment of the "on"- and "off"-levels. This new algorithm subjects all change point intervals to a second change point detection step. In this way, also fluctuations in the background and signal intensity are taken into account naturally (see SI). The results of the clustering are indicated by the red and blue colors in Figure 4a and b. Although not relevant for our data, it should be noted here that this change point based clustering algorithm can also be applied to multistate trajectories in exactly the same way. For multistate trajectories more than one change point would be identified separating the intensity levels, allowing for a generalization of our approach (see SI for details).

The change point analysis results obtained for all S:N ratios show that short off-times and short on-times are underestimated in the respective histograms compared with the underlying histograms (Figure 4c and d). For increasing S:N ratios the difference between these histograms is becoming smaller: for the S:N ratios of 14 000:4000 and 18 000:4000 photons/s both the on- and the off-histograms match the underlying histograms very well. Similar results are obtained for an increase in the intensity levels (Figure 4e and f). This result implies that a minimum number of photons is required to identify a change in the intensity level with a sufficient statistical significance. Short on- and off-times with a small number of photons have a chance to be missed with change point analysis.

When comparing the results obtained with the binning/thresholding and the change point analysis method, it is evident that both methods perform almost equally well for high S:N ratios such as 14 000:4000 photons/s (*cf.* Figure 3c and Figure 4c). For lower S:N ratios the number of short on- and off-times becomes inaccurate, as the number of photons is too small. The inaccuracies in the short dwell time regime have a different sign in the deviation: while the threshold approach overestimates the number of short dwell times, the change point analysis underestimates it. This result is further supported by the results obtained for model 1, where threshold analysis fails completely in determining the maximum in the off-histogram (Figure S4). The accuracy of both methods can be improved by increasing the overall number of photons.

In addition to the dwell time histograms, correlations between consecutive off-times are an important characteristic of dynamic disorder. Figure 5 shows the autocorrelations for several data sets. Independent of the data analysis method, the estimated correlations are always lower than the underlying ones. Data sets with higher S:N ratios show slightly stronger correlations, but even for the largest S:N ratio (*e.g.*, 36 000:8000 photons/s) the estimated correlations are only about half compared to the underlying ones. By taking into account the frequent occurrence of false-positive events (incorrectly assigned "on") and false-negative events (incorrectly assigned "off") determined with both methods (Table S4), the loss of correlations is an expected result. Despite the underestimated correlation strength, the number of correlated events captured is similar to those of the underlying autocorrelation. Most importantly, as correlations between turnovers are only observed for model 2 and not for model 1 (not shown here), it can be concluded that both data analysis procedures only reduced the extent of the correlations but did not introduce artificial ones in the simulated data sets.

Analysis of Experimental Data from the Enzyme α-Chymotrypsin. Having established the best parameters for the two respective data analysis methods, we have applied them to analyze a photon arrival time trace recorded for the enzyme α-chymotrypsin. This time trace is 1000 s long and has a S:N ratio of approximately 2.5 estimated from the average "on" and "off" intensities from both the thresholding and change point analyses. The maximum of the noise intensity distribution is located around 3500 photons/s. This falls into the parameter range where the accurate separation of on- and off-times is most difficult.

The analysis of the experimental data is complicated by the fact that the on-level is not purely Poisson distributed. Processes such as dye blinking contribute to the signal, and the number of photons might decrease gradually when the dye leaves the detection volume by diffusion. In addition, the background intensity might fluctuate and also increase with time as fluorescent dye molecules are produced by the enzymatic reaction. As a result, no constant threshold can be applied to the complete time trace. Instead the time trace needs to be analyzed in fragments. Here, intervals of 3 s were chosen to be able to apply threshold method A (2 ms interpeak time).

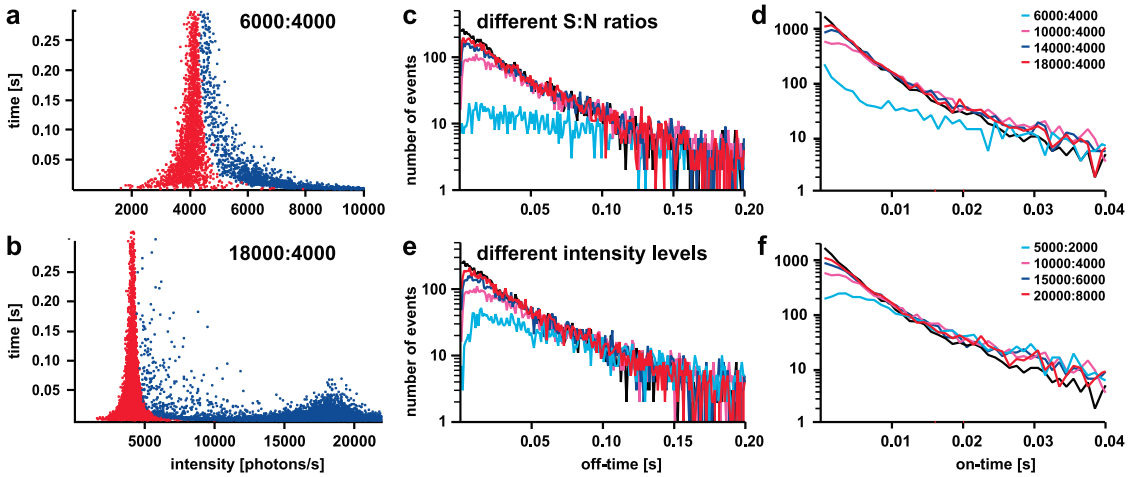


Figure 4. Analysis of simulated data (model 2) using change point analysis. On the left the durations of the change point intervals are plotted *vs* their intensity for two different data sets: (a) 6000:4000 photons/s and (b) 18 000:4000 photons/s. The blue and red data points represent on-states and off-states, respectively. The middle and the right columns show the off- and on-histograms. The analysis has been performed for different S:N ratios (c, d) and different intensity levels (e, f). The black lines represent the underlying histograms.

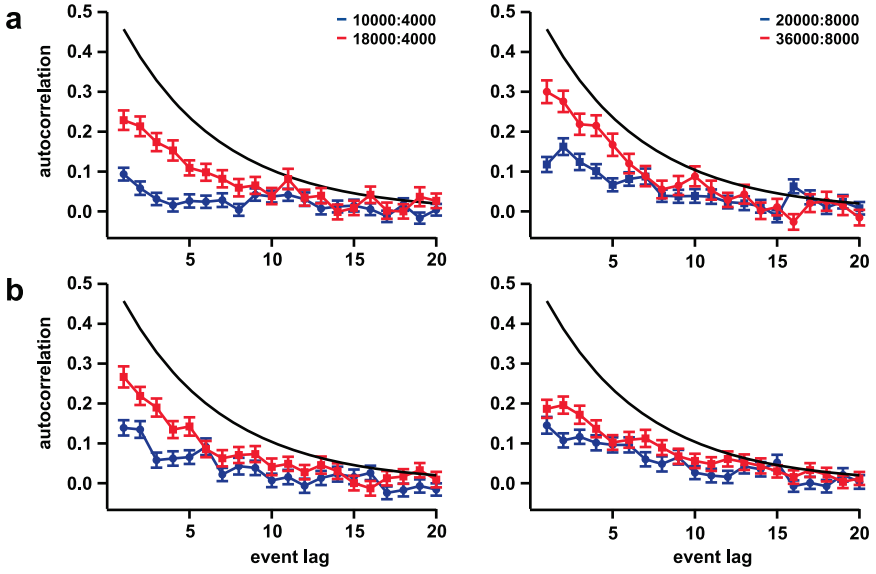


Figure 5. Off-time autocorrelation (simulated data: model 2). The autocorrelation of the off-times was determined for off-times obtained with threshold analysis (a) and with change point analysis (b) and compared with the underlying autocorrelation (black line). Autocorrelations are shown for two different S:N ratios and intensity levels.

We obtained 25 556 enzymatic turnovers with $\langle t_{off} \rangle$ = 0.038 s and $\langle t_{on} \rangle$ = 0.0015 s. The off-histogram has a very similar overall shape to the histogram obtained for the simulated data and shows a large number of short off-times (Figure 6). Using change point analysis the total number of enzymatic turnovers was 13 761 with $\langle t_{off} \rangle$ = 0.069 s and $\langle t_{on} \rangle$ = 0.0033 s. Now the off-histogram shows a very small number of short off-times with a maximum just as observed for the simulated data sets (Figure 4). The on-time histogram obtained with threshold analysis shows a smaller number of long on-times, resulting from the segmentation of on-times by the threshold (Figure 6). This artificial segmentation of the on-times is further highlighted in the autocorrelation analysis. No correlations are seen for the off-times determined with change point analysis, whereas the threshold method yields correlations in the short off-time region, as can be seen in the 2D-correlation plots (insets in Figure 6). Correlations between subsequent off-times as visualized in these 2D-correlation plots have been observed for other enzymes and have been considered as the most relevant proof for the existence of dynamic disorder.[13] Clearly, for α-chymotrypsin the observed correlations are artificially introduced and are directly related to the overrepresentation of short off-times when using threshold analysis.

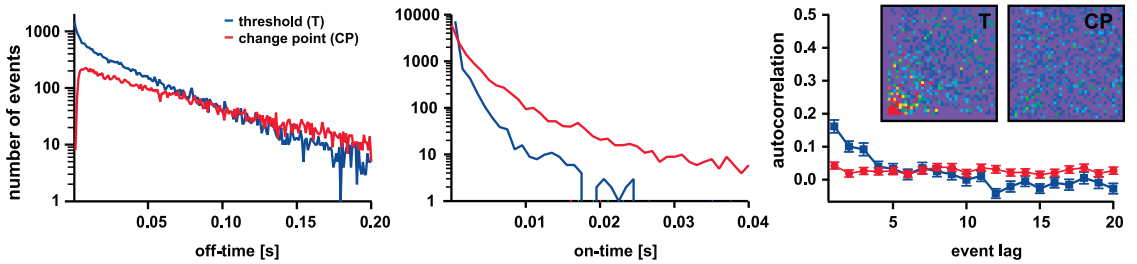


Figure 6. Analysis results for experimental data of the enzyme α-chymotrypsin. The time trace was analyzed with both threshold analysis (T, blue; 1 ms bin time, threshold method A, 2 ms interpeak correction) and change point analysis (CP, red). The off-time histograms show significant differences between the two analysis methods mostly in the short off-time region. Also the on-time histograms deviate from each other. Correlations are observed in the autocorrelation graph obtained with the threshold method. A 2D correlation graph indicates that these correlations are caused by the short off-times (inset; duration of the $i+1$th off-time plotted against the ith off-time with the correction $f(i, i+1) - f(i)f(i+1)$; the time scale on both axes ranges from 0 to 40 ms).

The observation that threshold analysis introduces artifical correlations when analyzing experimental data appears to be inconsistent with the autocorrelation results obtained for the simulated data sets of model 1, where no artifical correlations were introduced by the threshold method. This apparent discrepancy most likely results from the fact that the experimental on-states show much more fluctuations than the simulated data. Consequently, more and different errors are made in determining the presence and the duration of the on-states.

On the basis of the results obtained for both the simulated and the experimental data sets, we conclude that change point analysis is more accurate even when the S:N ratio is high enough to "lift" any overlap between the signal and the background intensity distributions. The following aspects support this conclusion: (i) the on-histograms are generally much better reproduced with change point analysis, as no binning is required; (ii) the overall number of events detected as false "on" or false "off" is smaller with change point analysis (Table S4); (iii) for experimental data, it is difficult to accurately determine the intersection value that is needed for threshold analysis; (iv) the on-state in experimental data might fluctuate much more than in our simulated data set; and (v) the diffusion of the fluorophore when leaving the confocal volume leads to a gradual rather than an instantaneous decrease in the intensity as we assumed for the simulated data. Although the latter two experimental complications lead to deviations from Poisson statistics in the on-state, we have not observed any negative influence on the performance of change point analysis.

Not only just comparing the results from a methodological point of view but also looking into their biological relevance, we observe that the shapes of the dwell time histograms obtained with the different methods are different. This affects the conclusions drawn from these histograms about the number of exponentials in the underlying kinetics. The concave shape of the off-histogram as obtained with threshold analysis has frequently been fitted with a stretched exponential and interpreted as one of the most important hallmarks of dynamic disorder. On the basis of the results presented here, this interpretation should be reconsidered as a possible artifact of the data analysis procedure. It should be noted that it is generally difficult to accurately determine the number of exponentials underlying the histograms if the number of exponentials needed for the fit is more than 3. Further, many different kinetic schemes can yield very similar histograms so that only little information about the rate constants can be obtained from fitting dwell time histograms.[28,29] This has recently been exemplified for kinetic schemes that contain a mixture of sequential and parallel steps along the reaction pathway.[29]

Correlations between consecutive off-times represent another characteristic of dynamic disorder, and it has been shown that the calculation of the autocorrelation of the off-times can already yield a clear indication for a rather small number of turnovers (>250).[20] As shown here, correlations may be introduced artificially by threshold analysis, however. For α-chymotrypsin it is clear that the correlations observed were artificially introduced by the threshold method. Consequently, the correlations seen for other enzymes may need to be revised critically, although no clear conclusions can be drawn yet.

Under our experimental conditions no indication can be found for the presence of dynamic disorder, which has been previously attributed to a stretched exponential off-histogram and off-time correlations. This does not exclude the existence of parallel reaction pathways (Figure 1) and the occurrence of conformational changes during an enzymatic reaction, however. The dwell time histograms might just not be the appropriate representation to describe dynamical processes especially if sequential and parallel steps occur in a given kinetic scheme.[29] More experiments with α-chymotrypsin and other enzymes are needed to answer the question if and how dynamic disorder contributes to enzymatic reactions and what will be the best experimental approach to detect it. Better S:N

ratios would greatly improve the accuracy of determining the on- and off-times, but can hardly be achieved with current fluorescent reporter systems and standard confocal microscope setups. New developments in both of these areas are required until a more systematic analysis will become feasible.

CONCLUSION

We have systematically analyzed simulated photon arrival time traces with the goal to identify the best method for extracting the on- and off-times of a single-molecule enzymatic reaction. For both methods the S:N ratio should be at least 2.5 and the background intensity level should be at least 4000 photons/s, as a minimum number of photons is required to reproduce the underlying true histograms accurately. For experimental data where the signal and background levels as well as the duration of the on-states are not known, change point analysis is more appropriate. Change point analysis of the data measured for the enzyme α-chymotrypsin does not yield a typical stretched exponential, and no correlations between consecutive off-times are seen. To draw conclusions about the presence of dynamic disorder from single-enzyme experiments, better data are required and higher order correlations need to be considered when analyzing the data. Besides enzymatic turnovers, change point analysis is also applicable to other systems, characterized by a sequence of "on"- and "off"-states such as quantum dots, ion channels, and carbon nanotube field effect transistors. Ultimately, it can also be used for trajectories with more than two states, making it a generic and objective procedure for a broad range of applications.

METHODS

Single-Molecule Measurements of α-Chymotrypsin. Single turnovers of the enzyme α-chymotrypsin were monitored with a confocal microscope setup as described by De Cremer *et al.*[18] α-Chymotrypsin hydrolyzes the fluorogenic substrate analogue $(suc\text{-}AAPF)_2$-rhodamine 110 in a two-step reaction, yielding the highly fluorescent dye rhodamine 110 as the final product.[30] Enzymes were immobilized in an agarose matrix on a clean glass coverslip by spin coating a solution of 0.05 nM enzyme and 1% (w/v) agarose in MQ water at 30 °C. Immediately after spin coating, the coverslip was placed in a sample holder and 1 mL of PBS (0.01 M phosphate pH 7.4, 138 mM NaCl, 2.7 mM KCl) was added on top of the coverslip. In order to localize individual enzyme molecules on the coverslip, the enzymes had been labeled with NHS-carboxyfluorescein (Invitrogen, mixed isomers; degree of labeling: 2 dyes/enzyme).

Utilizing the piezoelectric translation stage (Physik Instrumente) of the confocal microscope the surface was scanned to obtain an image of the fluorescence intensity at each position of the sample. Individual fluorescently labeled enzymes were visible as bright spots in the scanned images. After positioning the confocal spot at the position of an enzyme, the catalytic reaction was started by adding 1 mL of PBS containing 60 μM of the fluorogenic substrate analogue (30 μM final concentration).

Photon arrival time series of the enzymatic turnovers were recorded as described by Vosch *et al.*[31] Briefly, the sample was excited with 488 nm pulsed laser light (8.13 MHz repetition rate) from a Ti;sapphire laser (Tsunami, Spectra Physics). The excitation beam was focused through an oil immersion objective (Zeiss, 1.3 N.A., 100×) of an inverted microscope (Olympus IX70). The excitation power at the entrance port of the microscope was adjusted from 1 to 10 μW. Fluorescence was collected by the same objective, filtered by a band-pass 520/40 filter, and focused via a 100 μm pinhole onto an avalanche photodiode (SPCM-AQR-15, PerkinElmer). Photon arrival times were registered using a time-correlated single-photon counting computer card (Becker&Hickl GmbH, SPC 630).

Data Analysis with Binning and Thresholding. The data were binned with a bin size of 1 ms and subsequently analyzed with threshold method A (interpeak correction of 2 ms). For experimental data the background intensity increases over time because fluorescent product molecules accumulate as a result of the enzymatic and autohydrolysis of the substrate. A constant background as used for the simulated data is therefore not applicable for the analysis of experimental data. To facilitate the analysis, the time trace was separated into constant intervals of 3 s, and for every interval a new threshold value was calculated. This interval size is short enough to account for the nonconstant background intensity and long enough to build the intensity histograms. It should be noted that even with this procedure the intersection between the signal and background intensity distributions was very difficult to determine, as the distributions hardly showed any separation at all.

Data Analysis with Change Point Analysis. Before detecting the intensity change points of the experimental photon arrival trace, negative interphoton times (*i*th photon is detected by the avalanche photodiode (APD) before the $(i-1)$th) were removed. Photons having a microtime smaller than the microtime of the instrument response function were removed as well. Subsequently, change point analysis was performed as described in the SI (2.3).

Acknowledgment. The authors thank Haw Yang for supplying the code of the change point analysis program. T.G.T. and J.H. acknowledge financial support of the "Fonds voor Wetenschappelijk Onderzoek FWO" (grants G.0402.09, G.0413.10, G.0697.11), the K. U. Leuven Research Fund (GOA 2011/03), the Flemish government (long-term structural funding: Methusalem funding CASAS METH/08/04), and the Federal Science Policy of Belgium (IAP-VI/27). H.E., A.E.R., and K.B. acknowledge support from a long-term fellowship from the Human Frontier Science Program (HFSP) (K.B.), VIDI (K.B.), and VICI (A.E.R.) grants from The Netherlands Organisation for Scientific Research (NWO) as well as funding from the National Research School Combination Catalysis Controlled by Chemical Design (NRSCC). C.-B.L. and T.K. acknowledge support from JSPS, HFSP, and Grant-in-Aid for Research on Priority Area `Innovative nanoscience', MEXT.

Supporting Information Available: Additional results as well as simulation details are described. This material is available free of charge *via* the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

1. Baba, A.; Komatsuzaki, T. Construction of Effective Free Energy Landscape from Single-Molecule Time Series. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19297–19302.
2. Li, C. B.; Yang, H.; Komatsuzaki, T. Multiscale Complex Network of Protein Conformational Fluctuations in Single-Molecule Time Series. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 536–541.
3. Flomenbom, O.; Silbey, R. J. Toolbox for Analyzing Finite Two-State Trajectories. *Phys. Rev. E* **2008**, *78*, 066105.
4. Barkai, E.; Brown, F. L. H.; Orrit, M.; Yang, H. *Theory and Evaluation of Single-Molecule Signals*; World Scientific Publishing: Singapore, 2008.

5. Komatsuzaki, T.; Kawakami, M.; Takahashi, S.; Yang, H.; Silbey, R. J. *Advances in Chemical Physics*. Vol. 146. *Single Molecule Biophysics: Experiments and Theories*; John-Wiley & Sons, Inc.: New York, 2011.
6. Blank, K.; De Cremer, G.; Hofkens, J. Fluorescence-Based Analysis of Enzymes at the Single-Molecule Level. *Biotechnol. J.* **2009**, *4*, 465–479.
7. Claessen, V. I.; Engelkamp, H.; Christianen, P. C.; Maan, J. C.; Nolte, R. J.; Blank, K.; Rowan, A. E. Single-Biomolecule Kinetics: The Art of Studying a Single Enzyme. *Annu. Rev. Anal. Chem.* **2010**, *3*, 319–340.
8. Ha, T.; Ting, A. Y.; Liang, J.; Caldwell, W. B.; Deniz, A. A.; Chemla, D. S.; Schultz, P. G.; Weiss, S. Single-Molecule Fluorescence Spectroscopy of Enzyme Conformational Dynamics and Cleavage Mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 893–898.
9. Chen, Y.; Hu, D. H.; Vorpagel, E. R.; Lu, H. P. Probing Single-Molecule T4 Lysozyme Conformational Dynamics by Intramolecular Fluorescence Energy Transfer. *J. Phys. Chem. B* **2003**, *107*, 7947–7956.
10. Hanson, J. A.; Duderstadt, K.; Watkins, L. P.; Bhattacharyya, S.; Brokaw, J.; Chu, J. W.; Yang, H. Illuminating the Mechanistic Roles of Enzyme Conformational Dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 18055–18060.
11. Kuznetsova, S.; Zauner, G.; Aartsma, T. J.; Engelkamp, H.; Hatzakis, N.; Rowan, A. E.; Nolte, R. J.; Christianen, P. C.; Canters, G. W. The Enzyme Mechanism of Nitrite Reductase Studied at Single-Molecule Level. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 3250–3255.
12. Goldsmith, R. H.; Tabares, L. C.; Kostrz, D.; Dennison, C.; Aartsma, T. J.; Canters, G. W.; Moerner, W. E. Redox Cycling and Kinetic Analysis of Single Molecules of Solution-Phase Nitrite Reductase. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 17269–17274.
13. Lu, H. P.; Xun, L.; Xie, X. S. Single-Molecule Enzymatic Dynamics. *Science* **1998**, *282*, 1877–1882.
14. Velonia, K.; Flomenbom, O.; Loos, D.; Masuo, S.; Cotlet, M.; Engelborghs, Y.; Hofkens, J.; Rowan, A. E.; Klafter, J.; Nolte, R. J.; *et al*. Single-Enzyme Kinetics of CALB-Catalyzed Hydrolysis. *Angew. Chem., Int. Ed.* **2005**, *44*, 560–564.
15. Flomenbom, O.; Velonia, K.; Loos, D.; Masuo, S.; Cotlet, M.; Engelborghs, Y.; Hofkens, J.; Rowan, A. E.; Nolte, R. J.; Van der Auweraer, M.; *et al*. Stretched Exponential Decay and Correlations in the Catalytic Activity of Fluctuating Single Lipase Molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2368–2372.
16. English, B. P.; Min, W.; van Oijen, A. M.; Lee, K. T.; Luo, G.; Sun, H.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. Ever-Fluctuating Single Enzyme Molecules: Michaelis-Menten Equation Revisited. *Nat. Chem. Biol.* **2006**, *2*, 87–94.
17. Hatzakis, N. S.; Engelkamp, H.; Velonia, K.; Hofkens, J.; Christianen, P. C.; Svendsen, A.; Patkar, S. A.; Vind, J.; Maan, J. C.; Rowan, A. E.; *et al*. Synthesis and Single Enzyme Activity of a Clicked Lipase-Bsa Hetero-Dimer. *Chem. Commun.* **2006**, 2012–2014.
18. De Cremer, G.; Roeffaers, M. B.; Baruah, M.; Sliwa, M.; Sels, B. F.; Hofkens, J.; De Vos, D. E. Dynamic Disorder and Stepwise Deactivation in a Chymotrypsin Catalyzed Hydrolysis Reaction. *J. Am. Chem. Soc.* **2007**, *129*, 15458–15459.
19. Zwanzig, R. Rate-Processes with Dynamic Disorder. *Acc. Chem. Res.* **1990**, *23*, 148–152.
20. Shi, J.; Gafni, A.; Steel, D. Simulated Data Sets for Single Molecule Kinetics: Some Limitations and Complications of Data Analysis. *Eur. Biophys. J.* **2006**, *35*, 633–645.
21. Dan, N. Understanding Dynamic Disorder Fluctuations in Single-Molecule Enzymatic Reactions. *Curr. Opin. Colloid Interface Sci.* **2007**, *12*, 314–321.
22. Crouch, C. H.; Sauter, O.; Wu, X.; Purcell, R.; Querner, C.; Drndic, M.; Pelton, M. Facts and Artifacts in the Blinking Statistics of Semiconductor Nanocrystals. *Nano Lett.* **2010**, *10*, 1692–1698.
23. Watkins, L. P.; Yang, H. Detection of Intensity Change Points in Time-Resolved Single-Molecule Measurements. *J. Phys. Chem. B* **2005**, *109*, 617–628.
24. Kalafut, B.; Visscher, K. An Objective, Model-Independent Method for Detection of Non-Uniform Steps in Noisy Signals. *Comput. Phys. Commun.* **2008**, *179*, 716–723.
25. Ensign, D. L.; Pande, V. S. Bayesian Detection of Intensity Changes in Single Molecule and Molecular Dynamics Trajectories. *J. Phys. Chem. B* **2010**, *114*, 280–292.
26. Zhang, K.; Chang, H.; Fu, A.; Alivisatos, A. P.; Yang, H. Continuous Distribution of Emission States from Single CdSe/ZnS Quantum Dots. *Nano Lett.* **2006**, *6*, 843–847.
27. Kou, S. C.; Cherayil, B. J.; Min, W.; English, B. P.; Xie, X. S. Single-Molecule Michaelis-Menten Equations. *J. Phys. Chem. B* **2005**, *109*, 19068–19081.
28. Kienker, P. Equivalence of Aggregated Markov Models of Ion-Channel Gating. *Proc. R. Soc. London B: Biol. Sci.* **1989**, *236*, 269–309.
29. Floyd, D. L.; Harrison, S. C.; van Oijen, A. M. Analysis of Kinetic Intermediates in Single-Particle Dwell-Time Distributions. *Biophys. J.* **2010**, *99*, 360–366.
30. Terentyeva, T. G.; Van Rossom, W.; Van der Auweraer, M.; Blank, K.; Hofkens, J. Morpholinecarbonyl-Rhodamine 110 Based Substrates for the Determination of Protease Activity with Accurate Kinetic Parameters. *Bioconjugate Chem.* **2011**, *22*, 1932–1938.
31. Vosch, T.; Cotlet, M.; Hofkens, J.; Biest, K. V. D.; Lor, M.; Weston, K.; Tinnefeld, P.; Sauer, M.; Latterini, L.; Müllen, K.; *et al*. Probing Förster Type Energy Pathways in a First Generation Rigid Dendrimer Bearing Two Perylene Imide Chromophores. *J. Phys. Chem. A* **2003**, *107*, 6920–6931.