Single-enzyme kinetics with fluorogenic substrates: lessons learnt and future directions

Petri Turunen, Alan E. Rowan*, Kerstin Blank*

Radboud University Nijmegen, Institute for Molecules and Materials, Department of Molecular Materials, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

Abstract

Single-molecule fluorescence techniques have developed into powerful tools for studying the kinetics of biological reactions at the single-molecule level. Using fluorogenic substrates, enzymatic reactions can be observed in real-time with single-turnover resolution. The turnover sequence contains all kinetic information, giving access to reaction substeps and dynamic processes such as fluctuations in the reaction rate. Despite their clearly proven potential, the accuracy of current measurements is limited by the availability of substrates with 1:1 stoichiometry and the signal-to-noise ratio of the measurement. In this review we summarize the state-of-the-art and discuss these limitations using experiments performed with α-chymotrypsin as an example. We are further providing an overview of recent efforts aimed at the improvement of fluorogenic substrates and the development of new detection schemes. These detection schemes utilize nanophotonic structures such as zero mode waveguides or nanoantennas. Nanophotonic approaches reduce the size of the effective detection volume and are a powerful strategy to increase the signal-to-noise ratio. We believe that a combination of improved substrates and novel detection schemes will pave the way for performing accurate single-enzyme experiments in biologically relevant conditions.

© 2014 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

Enzymes are dynamic molecules with the ability of catalyzing biochemical reactions with high rate accelerations and specificity. The origin of this dynamic nature is a multidimensional energy landscape that does not only define the 3D structure but also the catalytic function as well as regulation processes. Over the past few decades, single-molecule techniques, especially single-molecule fluorescence spectroscopy (SANS) and microscopy, have evolved into powerful tools for studying enzymatic reactions. Advances in measurement technology and the development of fluorescence-based reporter systems [1–4] have improved the sensitivity and the time resolution to an extent such that dynamic processes during the catalytic reaction can be monitored in real-time, directly giving access to the sequence of events. Temporal fluctuations of enzyme behavior as well as heterogeneities between different enzymes in the population can be determined directly in single-molecule experiments. This information is inaccessible in ensemble measurements.

Conformational changes can, for example, be monitored utilizing Förster Resonance Energy Transfer (FRET). FRET is a sensitive spectroscopic ruler that reports on the distance between two fluorophores in close proximity. The energy transfer efficiency is a direct readout of the fluorophore distance and allows for detecting distance changes in the range of several nanometers [5–12]. At the single-molecule level (smFRET), when the donor and acceptor fluorophores are attached to specific positions for studying enzymatic reactions, smFRET can also be used for following reaction substeps and intermediate states of enzymes containing an optically active cofactor, such as in many oxidoreductases. The enzymatic turnover cycle can be observed directly when the cofactor cycles between an oxidized and a reduced state with different spectral properties.
A cofactor that absorbs light in only one of its oxidation states can be used as a FRET acceptor. Labeling the enzyme with a donor fluorophore in close proximity to the cofactor will consequently lead to changes in the donor fluorescence intensity as a function of the oxidation state of the cofactor. FRET labeling is not required when the cofactor itself is fluorescent in its oxidized state. In this case the redox cycling of the cofactor can be followed directly.

One of the most intriguing discoveries of single-molecule enzymology has been the observation of temporal fluctuations in the activity of enzymes, a phenomenon named dynamic disorder. These temporal fluctuations in the rate-limiting step have been the basis for the ‘fluctuating enzyme model’ that was developed to explain these observations. This model relates fluctuating rate constants with conformational changes, assuming that each conformation has its own specific activity. A reliable identification of dynamic disorder requires the detection of a large number of individual enzymatic turnovers. Sufficient statistics is required to accurately detect the underlying dynamic processes that may span a large range of timescales. The acquisition of the required data is only possible with a suitable reporter system that allows for long measurement times. The reporter systems mentioned above each possess certain technical limitations preventing such long measurement times and consequently the accurate analysis of dynamic heterogeneities in the catalytic activity.

Fluorogenic substrates, which are converted by the enzyme into fluorescent product molecules, are in principle the only reporter systems that permit sufficiently long measurements. Fluorogenic substrates are the ideal system for studying enzyme kinetics as every fluorophore is produced during every enzymatic turnover. In this paper, we review the current state of single-enzyme experiments focusing on fluorogenic substrates as reporter systems. We will first describe the technological requirements and the basic experimental setup used for performing these experiments. In the following, we will describe a recent single-molecule study of \( \alpha \)-chymotrypsin as an example to highlight the unique information that can be obtained in such an experiment. Using this same example, the current limitations of the substrate design and the measurement setup will be discussed together with recent developments aimed at overcoming these limitations. These include improved fluorogenic substrates as well as single-molecule detection schemes. These technological advances have the potential to improve the signal-to-noise ratio of the measurement and will ultimately help to perform single-enzyme experiments under biologically relevant conditions while at the same time allowing for a more accurate kinetic analysis.

### 2. Single-turnover detection

The ideal single-turnover measurement requires the direct observation of every enzymatic turnover with a signal-to-noise ratio (SNR) high enough for accurate data interpretation. Fluorogenic substrates provide this direct readout of individual enzymatic turnovers. They consist of a fluorophore linked to a functional group that is recognized by the enzyme. This functional group alters the photochemistry of the fluorophore, making it non-emissive. After enzymatic cleavage of the functional group, the fluorophore recovers its fluorescence. To be able to follow the sequence of turnovers, the enzyme needs to be immobilized on the surface of a glass cover slip. The enzyme itself is also fluorescently labeled so that every individual enzyme can be located on the surface. Upon the addition of substrate, the fluorescence signal at the position of one enzyme is recorded. Enzyme activity results in the formation of fluorescent product molecules at this specific position on the surface. Using a confocal microscope, every product molecule is detected as a fluorescence burst above the background.

**Fig. 1.** Single-turnover detection with fluorogenic substrates. (a) The laser of a confocal microscope is focused onto the position of an immobilized enzyme on a glass cover slip. Every fluorophore produced by the enzyme is recorded in real-time with single-photon sensitivity yielding the arrival time (“macro-time”) of every individual photon. Threshold analysis (b) or change point analysis (c) are used to assign the ON- and OFF-times to the photon arrival time trace.
noise. The product fluorophore is detectable only for a short time window after it has been created by the enzyme. It quickly exits the detection volume by diffusion. As every turnover yields a new fluorophore, the measurement time is not limited by photo-bleaching of the fluorophore but by the accumulation of fluorescent product molecules, which results in a reduced SNR.

The choice of fluorophore is critical since its brightness directly determines the SNR. The measurement setup does not only need to be sensitive enough to detect every single product molecule, but also needs to have a high temporal resolution. These criteria are fulfilled by a fluorescence microscope with confocal optics and avalanche photodiode (APD) detectors [1–4]. In a confocal microscope that is optimized for single-turnover measurements, a collimated laser beam is focused into a diffraction-limited spot using a microscope objective with a high magnification and a high numerical aperture. In the most commonly used epi-fluorescence configuration, fluorescence emission from the sample is collected with the same objective, spectrally separated from the excitation light with optical filters and guided to the detector(s). Before reaching the detector, the emitted light is focused through a small pinhole. Consequently, only fluorescence originating from the small focal volume is collected, reducing the size of the detection volume to approximately 1 femtoliter. This small detection volume facilitates the required SNR for detecting individual fluorophores. The APD detector provides single photon sensitivity and the required time resolution.

Using the setup described above, only the absolute photon arrival time (“macro-time”) is recorded, which gives the fluorescence intensity (photons/time) of every individual product fluorophore. Extra information about the catalytic process can be obtained when the fluorescence lifetime is determined simultaneously in what is called time-resolved measurement [30,32,33]. For these measurements the generated fluorophore is excited periodically with a picosecond-pulsed laser. The emitted photons are detected with time-correlated single-photon counting (TCSPC) [34,35]. When using TCSPC, the arrival time relative to the last laser pulse (the “micro-time”) is recorded for each photon. Once a sufficient number of micro-times have been collected, the corresponding micro-time histogram allows for determining the lifetime of the fluorophore. The time resolution of modern, fast APD and photon counting electronics is in the picosecond range, yielding lifetimes with high accuracy.

The data analysis procedure for extracting kinetic information starts with the recorded photon arrival time trace (macro-time trace). The first and most important goal is to obtain a binary ON–OFF trace (ON–OFF assignment). This binary trace is the basis for the kinetic analysis of the system. In this binary trace, the time intervals characterized by high (ON) photon count rates represent the enzymatic turnovers, while the low-intensity intervals represent the background noise (OFF). The Poison nature of single-photon detection and the background noise consisting of product diffusion and inelastic scattering adds challenges to the ON–OFF assignment. Two principally different approaches are available for the ON–OFF assignment. The most commonly used approach is the so-called threshold method [Fig 1b]. When using this method, the photon arrival time trace is first binned to obtain a photon count (intensity) time trace. A threshold is then applied to the photon count time trace to separate the ON- and OFF-levels. An alternative and more recently developed method is change point analysis [29,36–38]. Change point analysis is a statistical method for identifying changes in the photon count rate directly in the photon arrival time trace [Fig 1c]. It does not require binning. Instead, the photon statistics before and after each photon is compared using a statistical hypothesis test. This test is performed along the complete time trace and evaluates the log-likelihood ratio of the probability for each photon to be an intensity change point. The photon with the highest probability is assigned to be a change point. The time trace is then cut into two fragments at the change point photon and the algorithm is used for both fragments of the time trace. The process is repeated recursively until no intensity change points are found in the time trace anymore.

Once the ON–OFF binary trace has been obtained, multiple strategies are possible to extract the desired kinetic information [3,23,29–41]. The kinetic constants can be determined by fitting the probability distribution function of the OFF-time histogram with a model describing the kinetic process. In the case of a reaction with a single rate constant the OFF-time histogram follows a mono-exponential decay (straight line in a log-linear plot). In the case of a dynamic enzyme with multiple active states, the contributions from multiple reaction rates result in a multieponential decay profile. For most experimental data obtained so far OFF-time histograms with a concave shape have been obtained [16,22–25,27]. It was found that these histograms were fitted best with a stretched exponential function. This “stretched exponential” was used to describe an enzymatic reaction that exhibits a distribution of rate constants (dynamic disorder). Instead of using the OFF-time distribution, also the autocorrelation function of the OFF-times can be used to extract information about the underlying kinetics. More complex kinetic behavior, such as dynamic disorder, leads to correlations between consecutive OFF-times, indicating that OFF-times of similar length are more likely to directly follow each other. The difficulty in obtaining and interpreting the kinetic information is discussed in the example below.

3. Single-molecule study of α-chymotrypsin

α-Chymotrypsin is a well-studied digestive protease. Catalyzing the hydrolysis of peptide bonds, its main function is the breakdown of proteins into amino acids. Its activity is tightly regulated and highly pH-dependent [42,43]. The enzyme exists in a conformational equilibrium involving an active and an inactive conformation [43]. The related conformational changes occur on a time scale from milliseconds to seconds allowing the conformational dynamics to be detected in single-molecule experiments. A detailed single-molecule study of α-chymotrypsin has been performed in a series of papers [29,30,44]. The goal of this study was to systematically investigate the performance of the currently used techniques described above and to improve the accuracy of these methods.

One key feature of many commonly used fluorogenic substrates is the presence of two functional groups. These double-substituted fluorogenic substrates require two enzymatic cleavage steps before the fully fluorescent product is obtained. In most cases the intermediate shows fluorescence in the same wavelength range, however, its brightness is lower. This significantly complicates its detection in a single-enzyme measurement. Turnovers might be missed preventing an accurate kinetic analysis. For the α-chymotrypsin experiments described here the fluorogenic substrate was based on the commonly used fluorophore Rhodamine 110 (Rh110). The substrate carries two short peptides coupled to the two amine groups of the fluorophore to yield the fluorogenic substrate (sucinyl-AlaAlaProPhe)2-Rhodamine 110 (sAAPF2-Rh110). The peptide sequence was chosen to match the amino acid preference of α-chymotrypsin. In the case of Rh110, the intermediate is not only bright enough to be detected [30] but also has a different fluorescence lifetime than the product [Fig 2] [30,44].

The activity of single α-chymotrypsin molecules, immobilized in an agarose gel, was measured using a confocal detection scheme equipped with TCSPC. After performing the ON–OFF assignment with change point analysis, the fluorescence lifetimes were determined for each individual ON- and OFF-state. This lifetime analysis
revealed that only the reaction intermediate was produced in the enzymatic reaction as only the corresponding lifetime of 2.7 ns was found for the ON-states. Also the OFF-states showed an average lifetime of 2.5 ns. This result suggests that the background photons mostly originate from intermediate molecules diffusing through the detection volume. The measurement was performed at a substrate concentration of 30 μM, a factor 4x higher than the apparent $K_M$ value of the two-step reaction. Under these conditions the substrate concentration is far higher than the intermediate concentration in a single-enzyme experiment. As both substrates and intermediate compete for binding to the enzyme, the hydrolysis of the intermediate is very unlikely. Formation of the Rh110 product would only be possible if the intermediate would be converted to product directly after it has been formed (intermediate channeling) [45]. The results do not show any evidence of channeling, enabling a direct insight into the reaction mechanism of the studied two-step reaction. This result allowed the system to be treated as a reaction with 1:1 stoichiometry, which is ultimately required for studying the kinetics of the conformational change and the possible presence of dynamic disorder.

The ability of detecting the intermediate clearly simplifies the kinetic analysis. At the same time, however, it complicates the ON–OFF assignment, which precedes the kinetic analysis. The problem is the low brightness of the intermediate ($\Phi = 0.31$) that causes a low SNR. In the case of the described α-chymotrypsin experiments the SNR was only ~2.5. At SNRs below 5 the intensity distributions of the ON- and OFF-states overlap, hampering an accurate ON–OFF assignment. For such low SNR cases, the accuracy of the commonly used binning and thresholding method has been systematically compared with change point analysis [29]. Both methods were tested using large sets of simulated data with different SNR and intensity levels. Binning and thresholding generally overestimates the number of short OFF-times. In contrast, the number of short OFF-times is underestimated when using change point analysis. This problem is especially dominant when the photon count rate is low (<4000 cps for background). As change point analysis is based on a statistical analysis of the photon arrival time trace, its accuracy depends on the total photon count rate. For both methods the calculated OFF-time histograms were well reproduced when the SNR was high (SNR > 5). When the SNR was low (SNR < 3), both methods failed to reproduce the predicted OFF-time histograms accurately.

For the experimental data with a SNR of 2.5 the different methods gave different turnover rates (26.3 s$^{-1}$ with the threshold method and 14.5 s$^{-1}$ with change point analysis). In addition, they produced significantly different OFF-time histograms and autocorrelation results (Fig. 3). The largest difference in the OFF-histograms is observed in the short OFF-time region where threshold analysis produces a larger number of short OFF-states (Fig. 3a). This over-representation of short OFF-times resulted in the typically observed concave shape of the OFF-histogram that was fitted with a stretched exponential in previous research. The over-representation of short OFF-times also explains the higher turnover rate determined with the threshold method.

More interesting is the appearance of correlations between consecutive short OFF-times in the 2D-correlogram when the threshold method is used (Fig. 3b). These correlations are absent in the case of change point analysis. These results indicate that the observation of these correlations and the over-representation of short OFF-times are artifacts originating from threshold analysis. The fact that change point analysis does neither yield correlations between consecutive events nor stretched exponential OFF-time histograms questions the interpretation of previous single-enzyme experiments where similar results were obtained [16,22–25]. This systematic analysis clearly highlights that only data with a high SNR will allow a more detailed kinetic analysis of α-chymotrypsin.
and other enzymatic reactions in general. Only then can any conclusions be drawn about the presence or absence of dynamic disorder.

In summary, the single-molecule experiments performed with α-chymotrypsin illustrate the unique information that can be obtained with the current measurement technology. Novel insight into the order of reaction substeps has been obtained: no product molecules are formed by intermediate channeling when using double-substituted substrates. The low SNR resulting from the low brightness of the intermediate is the main bottleneck that limits the information that can be extracted from these measurements. Advanced data analysis methods alone cannot solve this problem. Even change point analysis, as an objective method, cannot detect the intensity change points reliably enough if the SNR of the data is not sufficiently high. The absence of dynamic disorder in the α-chymotrypsin measurements does not exclude its existence. New developments addressing both the fluorogenic substrate reporter systems and the measurement technology are required before single-turnover measurements can facilitate a more detailed kinetic analysis and ultimately become a standardized tool for biochemistry.

Fluorogenic substrates combining 1:1 stoichiometry with a higher brightness are obviously needed. Despite the high time resolution and single-molecule detection efficiency, confocal microscopy has its limitations when considering the experimental conditions and the biological relevance of single-enzyme measurements. α-chymotrypsin is a unique example allowing single-turnover detection at substrate concentrations above the (apparent) $K_M$ value of the reaction. This is not possible for a large number of other enzymes. Single-molecule occupancy in a diffraction-limited, femtoliter-sized detection volume corresponds to picomolar to nanomolar fluorophore concentrations. This often prevents the use of substrate concentration above the $K_M$ value. The use of nanophotonic structures is one strategy that has the potential to overcome these limitations. These structures facilitate a drastic reduction of the size of the detection volume by either confining or locally enhancing the excitation light. In the following sections we will review recent efforts aimed at the development of next-generation fluorogenic substrates and of detection schemes based on nanophotonic structures. When combined, these developments will ultimately lead to single-enzyme measurements with drastically improved SNRs under biologically relevant conditions.

### 4. Next-generation fluorogenic substrates

Fluorogenic substrates are very powerful reporter systems. The experiments described above clearly show their potential for monitoring long reaction sequences. They also highlight one of the key problems of current substrate designs. Rhodamine 110 (Rh110) and fluorescein are commonly used for the synthesis of fluorogenic substrates. They are both bright fluorophores that are easily detected in a single-molecule experiment. But they contain two functional groups that can be substituted during substrate synthesis. Mono-substituted derivatives of Rh110 are fluorescent and bright enough to be detected in a single-turnover experiment, however, with a much lower SNR [30,46]. In contrast, mono-substituted fluorescein derivatives appear to have a very low quantum yield in the range from 0.0003 to 0.22 [47–50]. Even though no data is currently available, it appears highly unlikely that mono-substituted fluorescein derivatives can be detected at the single-molecule level.

For both Rh110 and fluorescein, next-generation substrates with only one enzyme-cleavable group have been developed in recent years. The first step towards improved Rh110-based substrates with 1:1 stoichiometry was the development of Rh110 substituted with one urea-group, such as morpholinecarbonyl-Rhodamine 110 (MC-Rh110; Fig. 4a) [44,51]. The brightness of mono-substituted Rh110 derivatives depends on the electronegativity of the substituent [44]. Substituents with a higher electronegativity cause a stronger distortion of the symmetry of the xanthene unit of the fluorophore, reducing both its extinction coefficient and its quantum yield. When comparing MC-Rh110 with a mono-functionalized peptide-Rh110, MC-Rh110 shows a significantly higher extinction coefficient ($ε = 52,000$ M$^{-1}$cm$^{-1}$) and quantum yield ($Φ = 0.60$) than the peptide-derivative ($ε = 24,600$ M$^{-1}$cm$^{-1}$; $Φ = 0.31$; Fig. 2). As the urea-bond is not cleaved enzymatically, the relatively bright MC-Rh110 can be used as a new fluorophore. The remaining amino group of MC-Rh110 can be coupled to a peptide to yield a fluorogenic substrate with 1:1 stoichiometry (Fig. 4a).
Another strategy to obtain a fluorophore with only one amino group for coupling the peptide is the direct replacement of the amino group with a different functional group. This strategy has been followed when synthesizing the fluorophore Singapore Green (SG; Fig. 4b) [52]. SG is a hybrid of Rh110 and the fluorescein analogue Tokyo Green carrying one amino and one hydroxyl group at the xanthene ring system. SG has a quantum yield of $\Phi = 0.5$ and an extinction coefficient of $\varepsilon = 28500 \text{ M}^{-1} \text{ cm}^{-1}$. These values are similar to mono-substituted peptide-Rh110 derivatives so that single-molecule detection should be possible. A more recent approach is based on a mono-substituted hydroxymethyl derivative of Rh110 (HMRG, Fig. 4c) [53]. Coupling of only one peptide to HMRG induces a spirocyclization reaction that disrupts the conjugated electron system of the xanthene unit. In this way a non-fluorescent compound is obtained. After enzymatic cleavage of the peptide, the fluorescent form of HMRG ($\varepsilon = 57 000 \text{ M}^{-1} \text{ cm}^{-1}; \Phi = 0.81$) is recovered.

Besides yielding 1:1 stoichiometry, the Rh110 derivatives MC-Rh110 and SG have one additional advantage. Depending on the amino acid sequence of the peptide coupled, the substrate might become less water-soluble, limiting the highest concentration that can be used in the experiment. In the case of MC-Rh110, the morpholinecarbonyl group can be readily replaced with a different and more water-soluble urea-substituent. In the case of SG, the hydroxyl group can, for example, be functionalized with a short, water-soluble poly(ethylene glycol) chain. Overall, asymmetric fluorophores provide numerous possibilities for introducing additional functionality into these fluorogenic substrates, improving their properties.

In parallel to the above efforts aimed at the development of mono-substituted Rh110 derivatives, several mono-substituted fluorescein derivatives have been synthesized. In analogy to MC-Rh110, one hydroxyl group of fluorescein can be converted into a methoxy group to yield the fluorophore 3-O-methyl fluorescein. This fluorescein derivative has been used to synthesize the phosphatase substrate 3-O-methyl fluorescein phosphate [54]. To best of our knowledge nothing is known about the brightness of this fluorophore in comparison to other weakly fluorescent mono-substituted fluorescein derivatives [47–50]. A more controlled strategy for obtaining fluorescein derivatives with 1:1 stoichiometry is based on tuning the redox potential between the xanthene unit and the benzene ring [55]. This redox potential has a direct influence on the photophysical properties of fluorescein derivatives. Making use of the fluorescein derivative Tokyo Green, it was possible to tune the system in a way that the coupling of an enzyme-cleavable group changed the redox potential such that the quantum yield was lowered to $\Phi = 0.01$. After enzymatic cleavage the redox potential is shifted back to its original value and the fluorescence was recovered ($\Phi = 0.85$). In this way, enzyme substrates for $\beta$-galactosidase [55] and alkaline phosphatase [56] have been designed.

None of the mono-substituted Rh110 and fluorescein derivatives described above have been tested in single-molecule experiments so far. Even though they appear to have a high potential, it is currently not clear if the substrates are indeed ‘dark’ and the products sufficiently bright for single-turnover experiments. In order to obtain a high SNR, the residual fluorescence of the substrate is crucial and critically determines the highest useable substrate concentration. In currently used confocal detection schemes it is, for example, impossible to use the mono-functionalized peptide-Rh110 intermediate as a substrate, even though this is possible in ensemble experiments. In a single-molecule experiment, several of these substrate molecules are always present in the detection volume. Even though the brightness of these molecules is low, they would still generate a strong fluorescence background reducing the SNR. This problem can be overcome with the use of improved detection schemes based on nanophotonic metal structures that reduce the effective detection volume.

5. Nanophotonic structures for single-enzyme experiments

The maximum concentration that still allows single fluorophore occupancy in the detection volume of diffraction-limited optics is in the low nanomolar range. This is demonstrated routinely in fluorescence correlation spectroscopy (FCS) experiments where molecular diffusion can be measured with a high SNR [57,58]. These maximally allowed concentrations could become a problem in single-enzyme experiments with fluorogenic substrates. Even the most pure substrates contain at least 0.1% of fluorescent product molecules. This is a direct consequence of the synthetic procedure used to prepare these substrates, where the fluorophore is one of the starting materials. Following coupling of the functional groups, the obtained fluorogenic substrate needs to be separated from the starting material (fluorophore). Even with the best purification procedure, trace amounts of fluorophore always remain in the newly synthesized substrate. Autohydrolysis of the substrate can further lead to an increase in the fluorophore concentration during the experiments, thereby causing a decrease in the SNR over time. Whereas trace amounts of impurities are usually not a problem in ensemble measurements, they have a significant effect on the SNR in single-molecule experiments. Assuming 0.1% of fluorescent impurities, it is not possible to use substrate concentrations above 10 $\mu$M as the substrate will contain 10 nM fluorescent impurities. As the average $K_M$ value of an enzyme is in the range between 100 $\mu$M and 1 mM (Fig. 5), single-turnover measurements are generally limited to substrate concentrations below the $K_M$ value. In this range, substrate diffusion to the active site might become the rate-limiting step and determine the kinetic behavior of the enzyme.

Reducing the size of the effective detection volume overcomes these problems, but is a technologically challenging solution. Many technological developments during the past decade have facilitated a reduction of the effective size of the detection volume [59]. When using Total Internal Reflection (TIR) illumination the size of the excited volume is defined by an evanescent wave that decays exponentially from the sample surface. In this way the excited volume is reduced in the z-dimension. As a result, the effective detection volume is approximately 2.5-fold smaller than for confocal detection schemes [60,61]. The Near-Field Scanning Optical Microscope (NSOM) also makes use of an evanescent wave to excite fluorescence on a sample surface [62,63]. The excited volume is spatially confined at a scanning probe tip that contains an aperture with a typical diameter of 50 nm. As the light intensity is localized at the aperture, the size of the detection volume is reduced by approximately 1000-fold. NSOM is mainly an imaging technique that is not easily applicable in single-turnover experiments. In parallel to these efforts aimed at reducing the effective detection volume, various other approaches have been taken. For example, diffraction-limited optical schemes based on nanophotonic metal structures have been demonstrated experimentally [64]. These structures reduce the effective detection volume by approximately 100-fold. In this way, single-turnover experiments can be performed with significantly improved SNR.

Fig. 5. Distribution of $K_M$ values containing all enzyme-substrate pairs in the BRENDA database. Data taken from http://www.brenda-enzymes.info/.
experiments. Tip fabrication is complex and the tip would have to be placed close to the enzyme for the whole duration of the measurement.

5.1. Zero mode waveguides

As demonstrated by the NSOM, nanophotonic structures can offer a powerful strategy for confining the excitation light. Zero mode waveguides (ZMWs) use the same fundamental operating principle as NSOM, but are more easily implemented for single-turnover experiments. ZMWs are nanophotonic structures that consist of an array of mostly round holes milled in an opaque metallic film supported by a glass surface (Fig. 6). When the hole diameter is well below the excitation wavelength, no light above a certain cut-off wavelength can propagate through the hole thus creating a decaying evanescent wave at the entrance of this aperture. This cut-off wavelength and the attenuation factor of the propagating modes depend on the shape and size of the aperture [64–66]. In theory, the excited volume in the apertures of a ZMW array is defined by the diameter of the aperture and the characteristic decay length of the evanescent wave. ZMW arrays are usually fabricated using electron-beam lithography or focused ion-beam milling. First, a 150–400 nm thick metal film is deposited on a microscope cover slip. Aluminum or gold is generally used as a material. In the next step, apertures with a typical diameter ranging from 30 to 250 nm are milled into the metal. The size of these apertures ensures cut-off for wavelengths in the 450-700 nm range, which is the wavelength range required for the most commonly used organic fluorophores.

The potential of ZMW arrays for reducing the size of the detection volume was first demonstrated in an FCS measurement where the diffusion of Rh110-labeled dCTP nucleotides was observed. In this experiment apertures with a diameter of 43 nm were fabricated in an aluminum film. For this aperture size the effective detection volume was estimated to be approximately 20 zL, which is three orders of magnitude smaller than with diffraction-limited detection schemes [64]. In the FCS experiment, this small size of the detection volume facilitated single-molecule occupancy at a fluorophore concentration of 10 μM. The FCS measurement further showed a significant reduction of the diffusion time of Rh110-dCTP to only a few microseconds. This is one order of magnitude faster than in a diffraction-limited detection volume and further confirms the reduced size of the detection volume.

The major advantage of using ZMWs instead of NSOM for single-turnover experiments is that they act as a nanoaperture and a reaction vessel at the same time (Fig. 6). Enzymes can be immobilized on the glass surface at the bottom of the nanostructure where the confined excitation intensity is highest. This approach has been used for following the activity of T7 DNA polymerase at the single-enzyme level [64]. The enzyme was immobilized in 43 nm-sized apertures in an aluminum waveguide structure. The enzyme concentration was adjusted such that statistically only a fraction of apertures contained a single enzyme. Polymerase activity was then measured using the intrinsically fluorescent substrate coumarin-dCTP at a concentration of 7.5 μM. Fluorescence bursts were detected and could be clearly distinguished from the background caused by the diffusion of fluorescent nucleotides. Enzymatic turnovers could further be discriminated from diffusion events based on the residence time of the fluorophore in the detection volume, which was significantly longer in the case of a nucleotide incorporation event. The ability to use an intrinsically fluorescent substrate based on a coumarin dye highlights the unique potential of nanostructures for single-enzyme experiments. Coumarin derivatives have a low excitation coefficient, which is not sufficient for single-molecule measurements in diffraction-limited volumes [67]. In the ZMW structure, however, the significantly smaller detection volume facilitates the discrimination of this fluorophore from the fluorescent background.

Following these proof-of-principle experiments, the measurement setup was developed further to facilitate DNA sequencing [68]. To be able to monitor nucleotide incorporation in real-time, the four different nucleotides were labeled with spectrally distinguishable fluorophores. In this way, nucleotide incorporation could be monitored for thousands of single polymerase molecules simultaneously using a large ZMW array and an advanced parallel confocal multichannel detection scheme [69]. Each nucleotide incorporation event was detected as a fluorescent burst in one of the spectral channels with a high SNR (SNR = 25) allowing for the specific assignment of the different DNA bases. In addition to obtaining the nucleotide sequence, the data also contained kinetic information, i.e. the duration of each nucleotide incorporation event. This information has, for example, been used for extracting differences in the polymerase reaction rate for different methylation states [70].

Despite their high potential, ZMWs are still only rarely used for studying single-enzyme kinetics. One important drawback, limiting the attractiveness of the ZMW-based detection system, is the stochastic nature of the enzyme immobilization procedure. Considering the theoretical Poisson limit, maximally 37% of the apertures can be ‘filled’ with exactly one single enzyme when adding an enzyme solution to the nanostructure. Moreover, the immobilized enzymes cannot be specifically placed in the center of the aperture. Individual enzyme molecules are located at different distances from the metal wall of the aperture resulting in heterogeneities in the fluorescence excitation and emission properties. To overcome these problems, an atomic force microscope (AFM) based “cut-and-paste” technique was recently developed [71,72]. With this technique an individual molecule can be picked up with the AFM cantilever, transported and subsequently placed into a ZMW aperture. The cut-and-paste process can be repeated until all apertures are filled with exactly one molecule. Resulting from the high positional accuracy of the AFM, this method also ensures
that every molecule is always placed in the center of the aperture so that heterogeneities in the fluorescence properties are reduced. In a very recent report, DNA origami was introduced as an alternative strategy to improve the process of placing single molecules into the apertures [73]. With DNA origami [74,75] well-defined DNA nanostructures can be assembled that allow for the positioning of fluorophores, nanoparticles and biomolecules. On a DNA origami scaffold the position of each oligonucleotide is known with nanometer precision. Oligonucleotides modified with functional groups can therefore be integrated into the origami structure at predefined positions. Using this approach, disc-shaped nanoantennas were designed that matched the size of the apertures. The nanoantennas were immobilized on the surface of the ZMW via a biotin-neutravidin interaction, thereby ensuring that maximally one nanoantenna was present per aperture. The nanoantennas were further labeled with a single ATTO647N dye to allow for their detection. In this way, single-molecule occupancy was observed in 60% of the apertures, thereby overcoming the theoretical Poisson limit of 37%.

5.2. Fluorescence enhancement by metal nanostructures

One characteristic feature of ZMWs is that they do not only confine the excitation volume. The nanostructure can also enhance the signal detected from a fluorophore thereby improving the SNR even further [76–78]. An enhancement of 6.5-fold was, for example, reported for the fluorophore Rhodamine 6G in the 150 nm holes of an aluminum ZMW [76]. A 12-fold enhancement was found for Alexa Fluor 647 in a gold structure with 120 nm holes [77,78]. In these examples, the enhancement originates from the energetic coupling of the fluorophore with electron oscillations (surface plasmons) on the metal surface. If these surface plasmons are resonant with the excitation or emission wavelength of the fluorophore, they can enhance the fluorescence [65,79].

The presence of surface plasmons affects both fluorescence excitation and emission. The plasmons can increase the excitation intensity by locally enhancing the optical field inside the aperture [76]. The surface plasmons further influence the de-excitation pathways causing an increased quantum yield and a shorter fluorescence lifetime [78]. The enhancement is stronger for low quantum yield fluorophores. It allows the detection of these fluorophores in a ZMW structure with a SNR improved by one order of magnitude [80], eventually facilitating their use for fluorogenic substrates in single-enzyme experiments. The enhancement factor does not only depend on the fluorophore but also on the metal and the geometry of the nanostructure. Compared to aluminum, gold possesses stronger surface plasmon resonances in the visible range. This directly leads to a stronger fluorescence enhancement in the 600–700 nm range, but can also influence the fluorescence signal in an unfavorable way at lower wavelengths. The geometry of the aperture and the surrounding metal structure further influence the emission directivity. Ideally using specifically designed nanostructures, the emitted photons can be directed towards the detector [65,81]. A fluorescence enhancement of 120-fold combined with a 13.6-fold increase in emission directionality was recently demonstrated for the fluorophore Alexa Fluor 647 when the ZMW aperture was surrounded by periodic corrugations [82]. These results show that an optimization of the nanostructure parameters is a promising route towards increasing the enhancement factor and the potential of ZMW structures in general.

5.3. Optical nanoantennas

ZMWs make use of nanoapertures for reducing the effective detection volume and for enhancing the local excitation field. This can also be obtained with nano-sized optical antennas that are specifically designed to harness plasmonic effects. In contrast to ZMWs, the antenna geometry is designed such that it optimizes the local excitation field. It concentrates the excitation intensity from the surroundings and creates a localized excitation hotspot [83,84]. A metal nanoparticle can be considered as the simplest optical nanoantenna. A 10-fold fluorescence enhancement has, for example, been reported for the fluorophore Nile Blue in the immediate vicinity of gold or silver nanoparticles [85]. The enhancement factor is strongly dependent on the particle-fluorophore distance. At very short distances (<15 nm) also fluorescence quenching can become dominant [86]. Intensive research efforts are currently directed at understanding the interaction between fluorophores and metal nanoantennas with the goal of optimizing the fluorescence enhancement [84]. The local field enhancement can be significantly improved when the excitation hotspot is created between two point sources [87]. This requires fabrication of the nanoantenna structure with high precision in the nanometer range as the field enhancement critically depends on the gap between these point sources. Focused ion beam lithography and electron beam lithography are typically used for antenna fabrication. In the following three typical antenna designs that show the biggest potential for biological applications such as single–enzyme experiments will be highlighted.

The first successful design that showed a high fluorescence enhancement for single fluorophores utilized a bowtie-shaped nanoantenna structure [88]. The bowtie nanoantenna was fabricated on a quartz cover slip and consists of two triangle-shaped gold tips that point towards each other (Fig. 7a). The excitation hotspot is located in the gap between the two tips. The terylene derivative TPQDI was used for the quantification of the fluorescence enhancement in the excitation hotspot. The TPQDI molecules were randomly immobilized in a PMMA layer that covered the whole antenna array. When a single TPQDI molecule was located in the antenna hotspot, an impressive 1340-fold fluorescence enhancement was observed for this low quantum yield (Φ = 0.025) fluorophore. The enhancement factor was dependent on the gap size (5–80 nm). Antennas with the smallest gap sizes yielded the highest enhancement.

When considering the use of this design for a single-enzyme experiment, the bowtie design suffers from one important drawback: freely diffusing fluorophores contribute to the background signal. In contrast to ZMWs, the excited volume surrounding the small antenna hotspot is not geometrically confined and fluorescence from a larger diffraction-limited volume is collected. Compared to the antenna hot spot, this diffraction-limited volume contains a significantly larger number of fluorophores. Even though these fluorophores are not affected by the fluorescence enhancement, they will still contribute to the detected signal, significantly reducing the SNR when high fluorophore concentrations are used.

This problem was overcome with the use of a so-called ‘antenna-in-box’ design (Fig. 7b) [89]. In this design the nanoantenna was placed inside a ZMW aperture. The excitation hotspot was created in the gap between two half-spheres with a diameter of 76 nm. These half-spheres were located inside a rectangular aperture with the dimensions of 290 nm x 100 nm. The performance of this design was characterized with FCS measurements of freely diffusing Alexa Fluor 647 fluorophores. A fluorescence enhancement of 1100-fold was observed, combined with an effective detection volume of 58 zL, which is almost four orders of magnitude smaller than a diffraction-limited volume.

For both the bowtie and the “antenna-in-box” designs the controlled immobilization of the molecule of interest in the hotspot is difficult. It requires advanced nanopositioning approaches such as the AFM cut-and-paste strategy used for ZMWs [72]. Self-assembled nanoantennas based on DNA origami [74,75] are an
alternative strategy to tackle the immobilization problem. This nanoscale assembly technique has been used to create a DNA nanopillar carrying two gold nanoparticles with a gap distance of 23 nm (Fig. 7c) [87]. This DNA nanopillar was further functionalized with several biotin molecules allowing its upright attachment to a neutravidin functionalized microscope cover slip. Most importantly, the nanopillar contained a special docking site for the molecule of interest in the antenna hotspot between the nanoparticles. When one ATTO647N fluorophore was positioned in the hotspot, a 117-fold fluorescence enhancement was observed. Subsequently, the conformational dynamics of the Holliday junction were visualized with this setup. Unfortunately, no measurements with freely diffusing fluorophores have been carried out yet, so that no information is available about the SNR that could be achieved in single-enzyme measurements.

In summary, nanophotonic structures have the potential to overcome current bottlenecks in single-enzyme experiments. The confinement of the effective detection volume allows the use of higher substrate concentrations while ensuring a high SNR. The fluorescence enhancement effect further facilitates the use of substrates based on fluorophores with a low brightness. In combination, this gives access to a larger number of substrate designs with 1:1 stoichiometry that can be used in the biologically relevant concentration range. ZMW nanoapertures have already developed into a powerful detection scheme for single-turnover detection. Optical nanoantennas, designed to optimize the local field enhancement, are a possible alternative approach. They are under intensive development and a number of designs have been proposed in the past few years. Initial results with biological systems are very promising and their implementation for single-enzyme experiments can be expected in the near future.

6. Conclusions

Single-enzyme experiments allow for studying the kinetics of enzymatic reactions in real-time, giving access to possible temporal heterogeneities in the reaction rate. Fluorogenic substrates are ideal reporter systems. They allow for following the catalytic reaction for extended periods of time thereby providing the time sequence (and statistics) of a large number of individual enzymatic turnovers. Current bottlenecks, such as the stoichiometry of the substrates and the low SNR of diffraction-limited detection schemes, complicate the data analysis and limit the information that can be extracted from single-enzyme experiments. New developments need to focus on strategies that facilitate the accurate detection of every individual turnover with a high SNR in order to facilitate a detailed kinetic analysis. The development of next-generation fluorogenic substrates with 1:1 stoichiometry and the implementation of new detection schemes based on ZMW nanoapertures and optical nanoantennas are crucial improvements. The ideal single-enzyme experiment combines the developments from both fields, allowing single-enzyme measurements to be performed with a drastically improved SNR under biologically relevant conditions. These strategies will push single-enzyme experiments to the next level and ultimately answer the question if enzyme kinetics is or is not characterized by dynamic disorder.

Acknowledgements

The authors thank Johan Hofkens, Chun-Biu Li and Tatjana Ternetjeva for many fruitful discussions. This work was supported by the Dutch National Research School Combination Catalysis Controlled by Chemical Design (NRSCC grant 2009-100168, A.E.R.), The Netherlands Organisation for Scientific Research (NWO; VIDI grant 700.58.430, K.B.; Gravitation program 024.001.035) and the Human Frontier Science Program Organization (HFSP; Long-Term Fellowship LT00866/2007-L, K.B.; Young Investigator Grant RGY0058/2010, K.B.).

References


