Imaging Analysis of Photoswitching Fluorophores Using Single-Molecule Microscopy

Katherine E. Binkley  
kbinkley@patriots.uttyler.edu

Caleb Griffin  
cgriffin25@patriots.uttyler.edu

Follow this and additional works at: https://scholar.smu.edu/jour

Recommended Citation
Available at: https://scholar.smu.edu/jour/vol6/iss2/1

This Article is brought to you for free and open access by SMU Scholar. It has been accepted for inclusion in SMU Journal of Undergraduate Research by an authorized administrator of SMU Scholar. For more information, please visit http://digitalrepository.smu.edu.
Imaging Analysis of Photoswitching Fluorophores Using Single-Molecule Microscopy

Katherine E. Binkley, Caleb M. Griffin
kbinkley@patriots.uttyler.edu, cgriffin25@patriots.uttyler.edu

Alexander R. Lippert¹

ABSTRACT
Single-molecule localization microscopy (SMLM) is a developing field of biological imaging that employs the use of photoswitching fluorophores to image sub-cellular biological structures at a higher resolution than was previously possible. These fluorophores are used for protein labeling, so that the sample can be imaged under fluorescence microscopy. This type of microscopy requires the use of many different types of fluorophores, which are fluorescent organic compounds that blink stochastically on and off. Thus, it is critical for developers in the field to have easy access to statistical models of the behaviors of different fluorophores. Here, we take AlexaFluor 647 and analyze it using a fluorescence microscope, taking data on its blinking behaviors and discerning its properties when immersed in a fluorescence-dampening buffer solution. We find that the compound behaves best in buffer solution, and we forge a new methodology for evaluating new fluorophores in a systematic fashion using readily available computer software.

1. INTRODUCTION
In conventional fluorescence microscopy, the resolution is restricted by the diffraction limit, a physical principle that refers to the minimum amount of space between two light-emitting species required in order for them to be distinguishable. [1] At the molecular level, this becomes problematic because fluorescent molecules do not have adequate spacing, resulting in low-resolution images. [2]

By labeling proteins in biological samples with molecules that spontaneously switch between a photon-emitting fluorescent state and a non-emitting off-state, the emissions can be localized with high precision, [3] allowing samples to be imaged with a resolution that oversteps the diffraction limit. [4, 5, 6] This is done by compiling hundreds of images taken over time as the fluorophores blink, resulting in a near nanometer-resolution product. [7] Visual representations of photoswitching, localization, and reconstruction can be found in Figure 1.

SMLM requires optimal fluorescent behavior of fluorophores. [8, 9] The dyes must blink stochastically and exhibit a high percentage of time spent in the off-state to prevent overlap of multiple emitting fluorophores. [10, 11] One goal set by developers in this field is to conjure a method of multicolor SMLM using multiple different dyes per sample, which would entail the availability of dyes that fluoresce with different wavelengths but do not send interfering signals. [12, 13] Because of this, in addition to the already burgeoning demand for SMLM dyes, new fluorophores must be easily assessed for their photoswitching properties in order to evaluate their potential for imaging. This requires an established and efficient method of analysis. [12] Ideally, such a proposed method should be possible before the dyes are conjugated to proteins, in order to eliminate unnecessary steps. Here, we developed such a method of analysis by evaluating AlexaFluor 647, a commercially available fluorophore, using ImageJ and Microsoft Excel. We obtained graphs of single-molecule emissions as well as rate constants.

1 Dr. Alexander R. Lippert is an Associate Professor in SMU’s Department of Chemistry.
2. METHODS

A. Microscopy and data adjustment

A slide was prepared containing Alexa Fluor 647 (Figure 2) immersed in a fluorescence-dampening buffer solution, on a thin film of polyvinyl alcohol. The fluorophores were photographed under an OMX SR NL5.120R super-resolution microscope every 5 seconds for about 10 minutes, totaling 116 images. These images were then collected into a stacks file, which operates similar to a video, with 116 distinguishable frames. The stack images were edited and analyzed using ImageJ, an imaging analysis software developed by the National Institutes of Health. When the stacks file was opened using ImageJ, it was converted into an 8-bit format from a 16-bit format (image > type > 8-bit). The contrast and brightness were then adjusted to sharply distinguish the dots from the background (image > adjust > brightness/contrast). The brightness was adjusted by positioning the stack to the first frame and selecting “Auto” in the “B&C” panel. Contrast was increased to maximum by moving the contrast slider completely to the right. The stacks were then overlaid with a 5x5 square grid (analyze > tools > grid). The area per point for the grid was changed from 20 to 67 square micrometers, and “center grid on image” was selected. The file was then saved and re-opened in Fiji, an updated version of ImageJ, in order to correct a camera-drift using a plug-in called “Manual drift correction.” In order to run the drift correction, a specific test dot was selected out of a zoomed-in area of the stack. This dot was measured for diameter using the “line” feature, in frame 1, and then added to the ROI manager (see explanation below). This same dot was then selected in the same way in several other frames throughout the stack, ending in the last frame. All ROI’s were then selected, and the manual drift correction was run, outputting the corrected stack (plugins > registration > manual drift correction). The file was saved again and re-opened in ImageJ to continue.

B. Systematic analysis

The grid squares were labeled by horizontal rows (A-E) and vertical columns (1-5). Beginning with A1, each square was magnified and analyzed to find four representative fluorophores and their graphs. To select a molecule, first the stacks were viewed as a gif with a frame rate of 24 frames per second, and molecules which exhibited characterized behavior were chosen for analysis. Once a fluorophore had been chosen for analysis, the stack gif was paused, and it was highlighted using ImageJ’s region-selecting tools. These include the “wand” feature, which auto-outlines a bright area; the freehand feature; and the circle feature (Figure 3).

Once the “region-of-interest” (ROI) had been outlined, it was saved to the ROI-manager (analyze > tools > ROI manager > Add; shortcut, “t”). Graphs were found from the ROI’s, were converted into list form (image > stacks > plot z-axis profile > list > copy), and were then pasted into Excel. This process was repeated for four fluorophores in each square, totaling 100 molecules.

C. Data Entry into Microsoft Excel

Each fluorophore in Excel was graphed and labeled according to the square it was from. From the graphs, the fluorophores were categorized based on their behavior, into one of 8 types: stays on the whole time (1), blinks on only once (2), blinks off only once (3), photobleaching (4), turns on and stays on (5), blinks off multiple times (6), blinks on multiple times (7), and uncharacterized (8). Above each data column, three calculation cells were added: off-frames, on-frames, and duty cycle. [12] Specific Excel commands can be found in Table 1.

<table>
<thead>
<tr>
<th>Calculation</th>
<th>Excel Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-Frames</td>
<td>=COUNTIF(Initial:Final, “&lt;=100”)</td>
</tr>
<tr>
<td>On-Frames</td>
<td>=(total frames)-(off frames)</td>
</tr>
<tr>
<td>Duty-Cycle</td>
<td>=(on-frames)/(total frames)</td>
</tr>
</tbody>
</table>

Table 1. The table of excel functions used to find the duty cycle of individual molecules.

The data columns were then analyzed for their average on-times. This was defined as the average duration of fluorescence between periods of no fluorescence. Because of this definition, a period of fluorescence at the beginning or end of the stack was omitted, since it could not be determined how long the molecule had fluoresced outside the duration of the stack. A molecule was fluorescing if the intensity was greater than 100 on the 8-bit image, which...
ranges in values from 0 to 255. Because Excel does not have a function that directly computes average durations, a program was written in Python to accomplish this task (see Supporting Information). Specifically, the program presented the average on-times into the Excel file as its output. Below these, the rate constants for each of the fluorophores were calculated by taking the inverse of the on-times. \[14\] The rate constants were averaged over all the fluorophores, as well as for each of the 8 types.

The entire above process was repeated for a second data set, taken of Alexa Fluor 647 without the fluorescence-quenching buffer solution and at 10x the original concentration. This data set only covered 30 frames at 5-second intervals, and also used the TIRF feature of the microscope. The data was logged into Excel and analyzed in a near identical fashion to the first data set. In order to obtain a fair comparison of the two data sets, the first set was shortened to 30 frames and re-analyzed for comparison against the second data set.

3. RESULTS AND DISCUSSION.

Each of the small white dots seen in Figure 4 represents a single molecule of a commercially available SMLM dye. The molecules are photoswitching fluorophores, which turn stochastically off and on over time. A slide was printed containing this compound, Alexa Fluor 647, immersed in a fluorescence-quenching buffer solution on a PVA film \[12\] and analyzed using an OXR super-resolution microscope. The slide was photographed 120 times every 5 seconds over a 10-minute period, resulting in a stack file containing 120 individual frames. Figure 4 shows three frames of the stack. From frame to frame, the molecules scintillate and blink, and in every dark space there are invisible molecules in the dark state. Since the blinking patterns vary greatly from molecule to molecule, it was determined that single molecules needed to be analyzed.

First, in order to more sharply distinguish the dots from the background, the stack file was raised in contrast and converted from a 16-bit format to an 8-bit format. The new image was overlaid with a 5x5 square grid, labelled by column and row, and analyzed for 4 data points per square in order to total 100 molecules.

In order to analyze a molecule, one was chosen and highlighted (Figure 5a), saved to the program’s ROI manager (Figure 5b), and plotted for the z-axis profile (Figure 5c), where the z-axis is the intensity of emission. Each graph was converted into list form and then copied into Excel for analysis. Once in Excel, each set of data was graphed again and then given a characterization as one of eight types (Figure 6). The eight different types represent different blinking patterns. These include those which stay on the whole time and do not turn off (Type 1); those which blink on for only one interval (Type 2); those which turn off for only one interval (Type 3); those which are on at first and then turn off permanently, i.e., photobleaching (Type 4); those which are off at first and then turn on for the remainder of the time (Type 5); those which blink off multiple times (Type 6); and those which blink on multiple times (Type 7); In Figure 6, we find a representative graph from each of the eight types.
Photobleaching occurs when a molecule permanently ceases to fluoresce; Type 4 molecules show this. This could happen for a number of reasons, and could potentially impact the quality of the dye if it becomes a prevalent type. Thus this is a problem that necessitates future exploration, although such exploration was not deemed relevant for a kinetics study.

Based on the intensity graphs, the molecules were analyzed for duty cycles, on-times and rate constants. Duty cycles are calculations of molecules’ percentage of time spent in the on-state, defined as frames for which the molecules are on divided by the total frames in the acquisition. [12] On-time was determined by averaging the number of consecutive frames a molecule was in an on state between periods of off states, and the number of on frames was multiplied by the time interval between frames. The rate constants were found by taking the reciprocal of each on-time, [12] thus converting seconds to s⁻¹. The average rate constant for all molecules from the first data set was found to be 0.0443 s⁻¹. Table 2 expresses the rate constants for the individual types, better reflecting the unique behaviors of the molecules rather than the compound as a whole. Since the rate constants were found by inverting the on-times, molecules with on-times of zero did not yield a rate constant.

### Table 2. The experimentally determined rate constants, organized by type, of the complete first data set.

<table>
<thead>
<tr>
<th>Type</th>
<th>All</th>
<th>2</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate constant (s⁻¹)</td>
<td>0.0443</td>
<td>0.0370</td>
<td>0.0545</td>
<td>0.0274</td>
<td>0.0968</td>
</tr>
</tbody>
</table>

Comparing the two data sets, which included data taken with and without the buffer solution, showed that the average rate constant of two sets differed by less than 1%, but the rate constants for each type vary widely. The buffer appeared to stabilize the solution, as it experienced less standard deviation in rate constants for each molecule (Table 3).

### Table 3. The rate constants of the first and second data sets, with and without the buffer solution, respectively.

<table>
<thead>
<tr>
<th>Type</th>
<th>All</th>
<th>2</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer (s⁻¹)</td>
<td>±0.0665</td>
<td>±0.035</td>
<td>±0.0482</td>
<td>±0.062</td>
<td></td>
</tr>
<tr>
<td>No (s⁻¹)</td>
<td>±0.0741</td>
<td>±0.0692</td>
<td>±0.0977</td>
<td>±0.0423</td>
<td>±0.037</td>
</tr>
</tbody>
</table>

Table 4. The on-times with standard deviations, listed per type.

<table>
<thead>
<tr>
<th>Type</th>
<th>All</th>
<th>2</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer (s)</td>
<td>±10.3</td>
<td>±2.61</td>
<td>±2.24</td>
<td>±15.9</td>
<td>±6.78</td>
</tr>
<tr>
<td>No (s)</td>
<td>±17.7</td>
<td>±17.9</td>
<td>±12.7</td>
<td>±28.9</td>
<td>±3.08</td>
</tr>
</tbody>
</table>

The on-times of the types was the measure directly found from the data, and as a result, also utilizes the zero values in the data. The difference between the standard deviation of the buffered and non-buffered data was much more significant. The standard deviation per type also seems to be less for the buffered solution, with Type 7 being the exception (Table 4).

The first data set included four preliminary data points which were selected while designing the techniques for analysis. These four extra points were used in the analysis of the first as well as the 100 main points. While the buffered solution contained a plurality of Type 2 molecules, the non-buffered solution had a majority of Type 2 molecules. The no-buffer data set did not exhibit any Type 3 or Type 8 molecules, while the buffered data set contained at least 2 examples of Types 2-8 (Table 5). Although both solutions contained Type 1 molecules, they were not selected for analysis, and therefore the quantities cannot be compared.

### Table 5. The quantity of data points collected for each type in the buffered and non-buffered data sets.

<table>
<thead>
<tr>
<th>Type</th>
<th>All</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffered</td>
<td>104</td>
<td>32</td>
<td>2</td>
<td>11</td>
<td>5</td>
<td>19</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>No</td>
<td>100</td>
<td>53</td>
<td>0</td>
<td>18</td>
<td>6</td>
<td>9</td>
<td>14</td>
<td>0</td>
</tr>
</tbody>
</table>

The most common occurrence of fluorophores across both data sets was Type 2, which are molecules that blink on for only one interval. This represents the most ideal behavior of the dye; however, it is speculated that many of the data obtained for Type 2 merely encompass “drifter” molecules rather than true blinking. Although the PVA films...
are designed so that the fluorophores bind to the surface for imaging. [12] PVA is soluble in water, so the aqueous buffer solution over time dissolves the films and thus loosens molecules from the surface, creating drifters. For ROI graphs for which the on-time duration is only one frame, it is most likely that the molecule drifted briefly into the ROI as it migrated across the slide. However, some of the Type 2 fluorophores remained on for several frames, as this is indicative of true on-time.

As seen in Figure 6, Type 1 did not have any on-time by definition, and therefore concluded no rate constants. Types 6 and 7 both strayed much further from the average value. Types 2 and 5 retain values closer to the average, and Types 3 and 4 did not yield any rate constants because they exhibited no on-time, by definition. See Table 2 for representative graphs of each type. Category 8 was reserved for molecules that exhibited behavior that was incongruent with other species - for example, molecule A3.3 (Figure 7) appears to begin in a blinking state between mostly on and sometimes off, but then at about frame 52, it turns off almost completely, only blinking on a couple more times for the rest of the file. This could be seen as switching from Type 6 to Type 7.

Another example of abnormal behavior, molecule B1.1, (Figure 8) is rather unique in that it appears to be “on” and blinking for a while at the expected intensity of ~255; however, at frame 67, its intensity drops to around 150. It then hovers around this value until the last four frames, where it drops to zero. This abnormal curve displays quantized emission levels very nicely, although the precise reason for the 150 bar is uncertain. It is speculated that this could represent two molecules either on top of or very closely side-by-side to one another, where one blinks off near frame 67, and the other blinks on around frame 112.

The duty cycle calculations proved to be much simpler. The entire first data set yielded an average duty cycle of 0.307, much greater than the 0.0138 found by Bittel et al. [12] The individual types were extremely varied in duty cycle, with Type 2 being the smallest and closest to the literature values. Conversely, Type 6 had the greatest duty cycle and was farthest from the literature values. (Table 6). Types 2 and 7 are similar by definition since Type 2 is, fundamentally, a singular form of Type 7. This similarity is coherent with the duty cycle results, as the two types produced the most similar duty cycles.

The buffered and non-buffered solution duty cycles were notably different; Type 5 was dramatically so. Type 5 is defined as those which are initially off and then “un-bleach,” per se; it is likely that the extreme difference in the two duty cycles is because the point at which the molecules turned on was in the portion of the data set which was removed to make the buffer and non-buffered data sets the same length. This is supported by comparing the duty cycle of the full-length buffered solution to the shortened solution. Type 5 yielded a duty cycle of 0.266 in the full length and 0.00667 in the shortened data set (Tables 6 and 7).

### Table 6. The list of duty cycles, organized by type, for the analysis of the complete first data set.

<table>
<thead>
<tr>
<th>Type</th>
<th>All</th>
<th>2</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duty cycle</td>
<td>0.307</td>
<td>0.0612</td>
<td>0.266</td>
<td>0.796</td>
<td>0.119</td>
</tr>
</tbody>
</table>

### Table 7. Table comparing the duty cycles of the buffered and unbuffered solutions, by type.

<table>
<thead>
<tr>
<th>Type</th>
<th>All</th>
<th>2</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffered</td>
<td>0.308</td>
<td>0.0208</td>
<td>0.00667</td>
<td>0.749</td>
<td>0.0778</td>
</tr>
<tr>
<td>No Buffer</td>
<td>0.239</td>
<td>0.0937</td>
<td>0.422</td>
<td>0.533</td>
<td>0.0690</td>
</tr>
</tbody>
</table>

Comparison of the data on the slide containing the buffer solution versus the one without shows that the buffer...
solution helps to produce ideal photoswitching behavior. The data from the non-buffered slide was erratic, harder to analyze, and exhibited a larger standard deviation in rate constants. It is noted that a true comparison between these two data sets in this respect is difficult to make, because the second data set also contained a higher concentration, fewer frames, and the TIRF feature. However, it is helpful to consider the comparison in light of the overarching goal of the project. The high-resolution of this type of microscopy requires fluorophores to spend more time in the off-state than in the on-state, since the localization of single molecules is dependent upon their temporal separation. Ideally, each fluorophore only spends a short interval on before staying off for the majority of the acquisition, so that no two fluorophores are in close proximity at any given time.

4. Conclusion

Duty-cycle calculations were comparatively higher from literature values, indicating that this particular sample of Alexa Fluor 647 remained in the “on” state for a higher percentage of time. One proposed reason for this could be the nature of the data collection—for this analysis, molecules were intentionally chosen based on their different types of behaviors, rather than for one specific behavior.

Another reason could be the length of the acquisition. The frame count for the first data set was reduced to 30 frames at 5 second intervals in order to match the second data set, whereas it was originally 120 frames. Furthermore, since the second data set’s on-times exhibited much higher standard deviations, it can be inferred that stable fluorophore behavior results from slides prepared with the fluorescence quenching buffer solution. This is necessary for optimal SMLM image quality since excessive fluorescence causes fluorophores to not be temporally separated.

The rate constants of Alexa Fluor 647 serve to provide the probability of a molecule to revert to the on-state. According to the units of the rate constants (s^{-1}), the observed reaction rate is first-order. However, the kinetics are rather complicated, since the buffer solution has been shown to be a necessary component, indicating a bimolecular process. For the purposes of this paper, we are assuming a first-order observed rate.

Some ideas for future experimentation include comparison of two data sets with only one variable that differs between the two. Here, the second data set analyzed contained a higher concentration of fluorophores, a shorter frame count, and used the microscope’s TIRF feature, making a true comparison to the first set difficult. Additionally, future work might exclude graphs of molecules which only turn on for one frame. As previously described, these are speculated to merely be drifter molecules; a better definition of Type 2 might require molecules to be on for at least 2 frames in order to be counted in the data collection. Lastly, although this data analysis included appreciable data with only 100 molecules, an ideal analysis would include all the molecules of a stack. This task proved difficult to do by way of program writing, and was judged to be nearly impossible to do manually considering the likelihood of the user to either skip or repeat molecules. However, if it were possible for future experimenters to find an automated way to find the ROI’s of all molecules without skipping or repeating any, the data obtained might be better and less subject to experimenter bias.

As the described methods have shown, the combination of ImageJ and Excel to select and analyze single molecules provides a prospective method for gathering data on a vast library of fluorophores, contributing to the need for quick access to fluorophore data for SMLM imaging. Protein conjugation requires multiple synthesis steps, which is highly time-consuming, especially if the dye properties can only be assessed post-conjugation. Previously, dye properties could only be assessed post-conjugation. Thus, as SMLM advances, it is critical that new dyes can be assessed for optimal properties prior to protein conjugation, because it is advantageous for SMLM developers to have convenient access to information on various dyes. This analysis of Alexa Fluor 647 demonstrates one efficient method that accomplishes this using user-friendly, easy-to-access software.
5. **SUPPORTING INFORMATION**

Section S1: Full data set, including time-intensity profiles of all single molecule observations.

![Intensity versus time plots for different sections](image-url)
Type 2 graphs of the complete first data set, buffered.

Type 3 graphs of the complete first data set, buffered.

Type 4 graphs of the complete first data set, buffered.
Type 5 graphs of the complete first data set, buffered.

Type 6 graphs of the complete first data set, buffered.
Type 7 graphs of the complete first data set, buffered.
Section S2: Python source code used for analysis.

```python
On-time program [python 3 script]
excel_file = "name.xlsx"
boundary_value = 100
interval = 5
data_list = []
import openpyxl
from statistics import mean

excel_data = openpyxl.load_workbook(excel_file)
raw_data = excel_data.active
height = raw_data.max_row
width = raw_data.max_column

for progress in range(1, width+1):
    for cell_height in range(1, height+1):
        cell_raw_data = raw_data.cell(row=cell_height, column=progress)
        value = cell_raw_data.value
        try:
            if value >= boundary_value:
                data_list.append(1)
            else:
                data_list.append(0)
        except:
            print("Unrecommended Data Format, Errors Are Possible")
            data_list.append(2)

print(data_list)

count = -1
length = 0
```

Type 8 graphs of the complete first data set, buffered.
on_time = []
for number in data_list:
    if number == 1 and data_list[count] == 2:
        length = -1
    elif number == 1 and length >= 0:
        length += 1
    elif number == 0 and length == -1:
        length = 0
    elif number == 0 and length != 0:
        on_time.append(length)
        length = 0
    elif number == 2:
        length = 0
        on_time.append(0)
        count += 1
print(on_time)

values_list = []
times_list = []
for value in on_time:
    if value != 0:
        values_list.append(value)
    elif value == 0:
        try:
            times_list.append(mean(values_list))
        except:
            times_list.append(0)
        values_list = []
converted_times_list = [element * interval for element in times_list]
print(converted_times_list)

all_times = []
for value in on_time:
    if value != 0:
        all_times.append(value)
for each in range(1, width+1):
    raw_data.cell(row=height+2, column=each).value = converted_times_list[each-1]
    raw_data.cell(row=height+3, column=2).value = mean(all_times) * interval
    raw_data.cell(row=height+1, column=1).value = "Average on-time for each column"
    raw_data.cell(row=height+3, column=1).value = "Total average on-time"
excel_data.save(excel_file)
print("Success!")

6. REFERENCES


“From single-molecule spectroscopy to super-resolution imaging of the neuron: a review.”


