

TU DRESDEN

ADVANCED PRACTICAL COURSE

LAB REPORT

Biomolecular Motors

Authors:

Toni EHMCKE
Christian SIEGEL

Supervisor:

Samata CHAUDHUR

Dresden, November 22, 2015

Date of experimental procedure: November 5, 2015

Contents

1	Introduction	2
1.1	Kinesin-1: Cell's workhorses	2
1.2	Fluorescence Microscopy	2
2	Experimental procedure	2
2.1	Making of the Kinesin-1-stepping assay	2
2.2	Stream acquisition	2
3	Data Analysis	3
3.1	Data evaluation of velocity	3
3.2	Data evaluation of run length	4
4	Discussion and conclusions	6
5	Appendix	6
6	References	7

1 Introduction

1.1 Kinesin-1: Cell's workhorses

1.2 Fluorescence Microscopy

2 Experimental procedure

2.1 Making of the Kinesin-1-stepping assay

2.2 Stream acquisition

First the experimental supervisor placed the flow cell holder on the microscope stage. We used an oil objective which has to touch the bottom of the flow cell. The objective has a magnification of 100x and a numerical aperture of 1,46. After the assay was fixed we used a Software called MetaMorph. We chose the camera button. We used a digital Camera by Andor which took greyscaled pictures. Its resolution is 2.6 MPixel with 512x512 pixels. The size of one pixel ist $256 \mu\text{m}^2$. According to the meta data of the movies and pictures the chosen additional magnification of the camera is 1.0x. This means, the pixel size is the original pixel size of the camera. If it had been 2.5x, we could have seen 2.5 times less, so a pixel would have been $A = 256 \mu\text{m}^2 / 2.5 = 102.4 \mu\text{m}^2$. So we took the TRITC filter and the lamp on the taskbar and watched the live images token by the camera. We searched and focused a cutout where we could see enough MTs. Unfortunately our assay showed that no MT's were fixed in it, so we could not take any pictures. Maybe this was caused by a too thin microtubule- or antibody-solution. We used the assay of our co-workers instead.

After we focused the MTs, we stopped live imaging and took pictures of the MTs luminescated by rhodamine. After that we take the GFP filter and choose the laser illumination. Then we choose again live imaging ("show live"). The TIRF angle will be adjusted. Again we choose the TRITC filter on the taskbar, the lamp illumination and take live imaging. We move to a new field of view, focus properly and take an image of the MTs such as in figure 1.

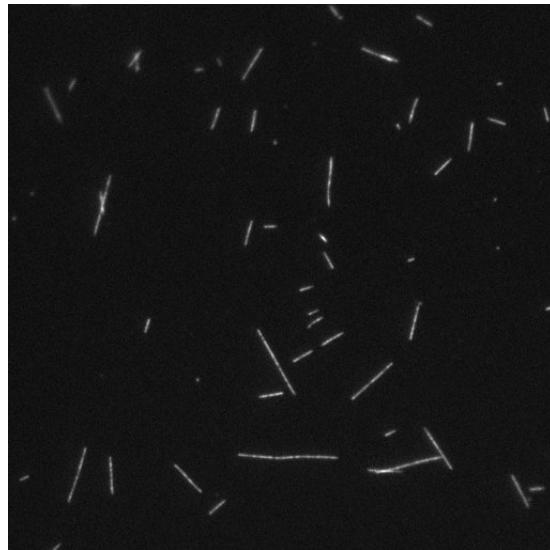


Figure 1: Example of an image, where we focused on the MTs

Then we save an image of this position. We now switched again to the GFP filter, selected the laser as light source and take live imaging. We can see moving motors. We stop live imaging and take a movie via the acquire button. The number of frames we took is 1000. For one frame the camera needs 150 ms, so at the end one we wait 150s per video stream. We took 8 movies. When all streams are collected, one can do the data processing with FIESTA. There you can mark the possible trajectories of the motor proteins (as seen in figure 2) and the software will calculate pieces

of the picture where the height is the time dimension and the width is the movement. If there is motion, one can mark up the lines and get the time which was needed for the marked distance (see figure 3).

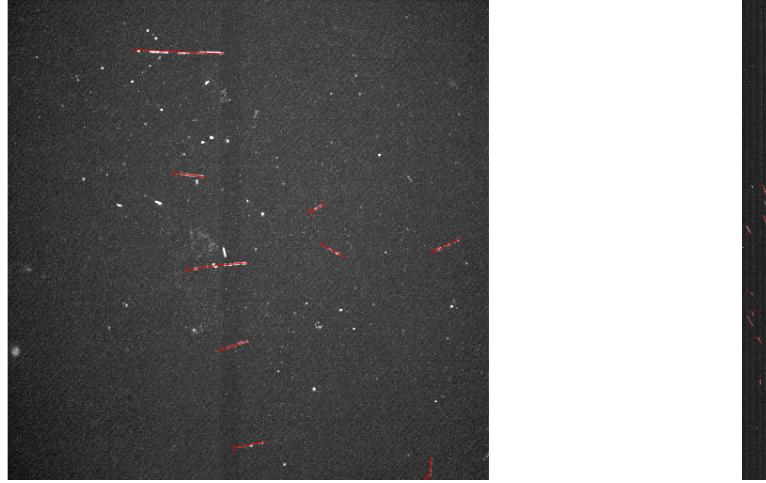


Figure 2: first step marking of the micro-tubules

Figure 3: marking of the trajectories of the motor proteins

3 Data Analysis

According to the data evaluation, we got $N = 204$ events. That means, we got 204 times a distance, a time and a velocity which are distributed.

3.1 Data evaluation of velocity

We think, the velocity is a gaussian distribution. This assumption can be proven by a student's t-test. But we only have one set of data.

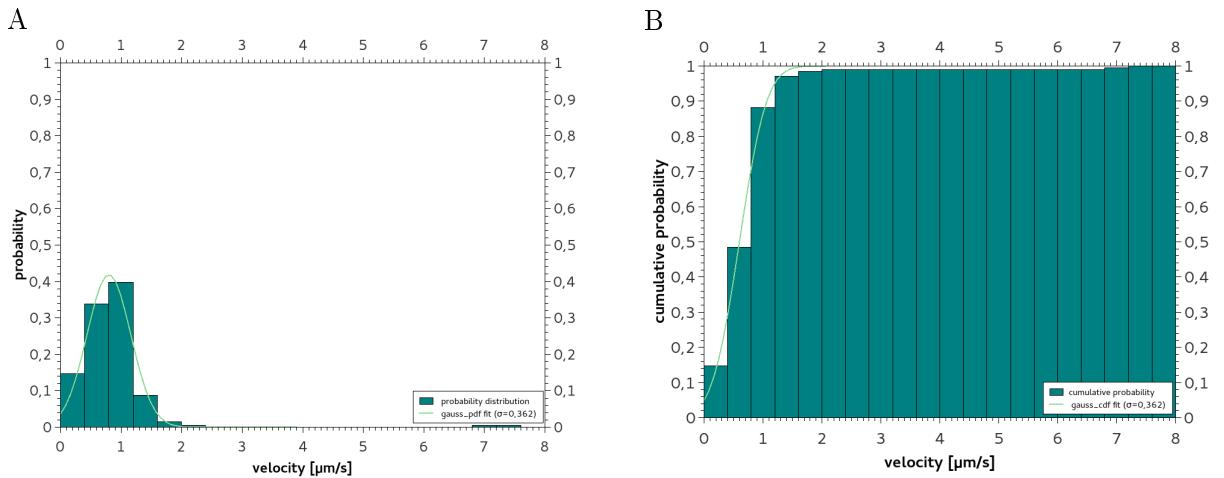


Figure 4: **A:** Histogram of the relative velocity distribution. **B:** Histogram of the cululative velocity distribution

The mean of the velocity \bar{v} , its standard deviation $\sigma_{\bar{v}}$ and the standard error of the mean velocity $\Delta \bar{v}$ can be calculated by a data analysis tool. We used an Origin like software called qtiplot, which calculates column statistics, which are not influenced by binning, because that are the measured

raw data. These include the searched values. In general one can calculate these values as following:

$$\bar{v} = N^{-1} \sum_{i=1}^N v_i, \quad \sigma_{\bar{v}} = \sqrt{\sum_{i=1}^N \frac{(v_i - \bar{v})^2}{N-1}}, \quad \Delta \bar{v} = \frac{\sigma_{\bar{v}}}{\sqrt{N}}$$

So we get:

$$\bar{v} = (0.87 \pm 0.05) \frac{\mu\text{m}}{\text{s}}, \quad \sigma_{\bar{v}} = 0.72 \frac{\mu\text{m}}{\text{s}}$$

To minimize the error one could easily take more measurements. According to the law of large numbers the measured mean velocity would converge to the expectation value. With an infinite number of measurements we would get the exact result. Also we could use software which marks the traces, which would create a unit law of marking. As a human being, one cannot see every trace and one can not mark every trace the way. This is caused by a diameter of the traces which is not infinite. The means, one could easily stretch the distance walked by the kinesin by marking from one edge to the diagonal opposite edge. An automatization would avoid these errors.

3.2 Data evaluation of run length

Now we are going to figure out what distance D a motorprotein covers on a single microtubule before releasing itself. For that we also use the data acquired by the streamed films. One frame of the stream corresponds to 150 ms. With that and the determined velocity $\bar{v} = (0.87 \pm 0.05) \mu\text{m}/\text{s}$ we can calculate the minimal distance we can measure with our streaming-system (and also the minimal bin-size): $d_{min} = \bar{v} \cdot 1 \text{ FRAME} = 0.12 \mu\text{m} \cong 0.2 \mu\text{m}$. Therefore all of the measured distances which are below that value cannot be reasonable and they will be ignored in our statistics - anyway they will appear as first bin in the histogram for the sake of completeness. We rounded the value of d_{min} up because we want to choose a smooth bin-width. Finally we decided the bin width to be $w_{bin} = 0.4 \mu\text{m} \cong 2 \cdot d_{min}$. We estimated this using the **Square-root choice**[02] which says that the number of bins is $k = \lfloor \sqrt{N} \rfloor$ where $N = 204$ is the number of samples.

Considering these facts we got the following results. At first the mean distance and its statistical error:

$$\bar{D} = (1.1 \pm 0.1) \mu\text{m}$$

The figures (5) and (6) visualise the distribution of the measured distances:

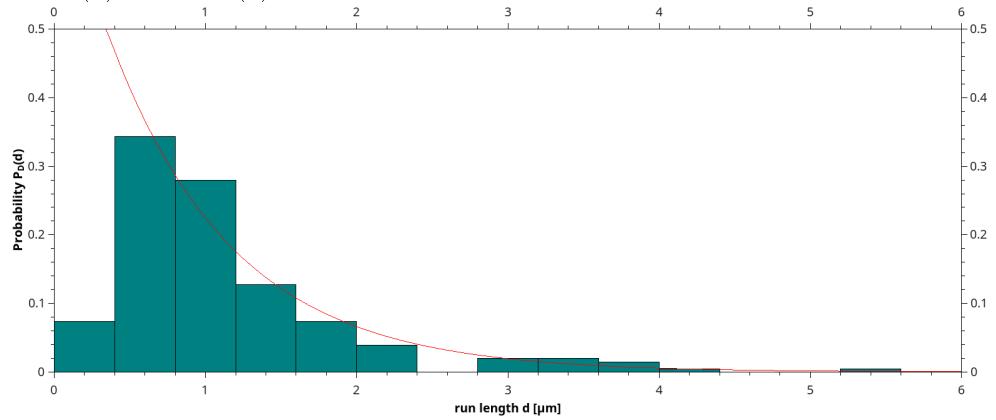


Figure 5: Probability distribution of the measured run length.

The very first bin is ignored from the exponential fitting

Because all invalid values fall into the first bin, it will be disregarded in the exponential fit of the probability function $P_D(d)$ and in the cumulative probability. We consider this fact by using an offset $d_0 = w_{bin} = 2 \cdot d_{min} = 0.4 \mu\text{m}$ in the exponential fit of the probability distribution. So the fit equation becomes:

$$P_D(d) = A \cdot e^{-\kappa \cdot (d - d_0)}$$

with the fit parameters:

$$\begin{aligned} A &= 0.37 \pm 0.02 \\ \kappa &= (1.2 \pm 0.1) \mu\text{m}^{-1} \\ d_0 &= 0.4 \mu\text{m} = \text{const.} \end{aligned}$$

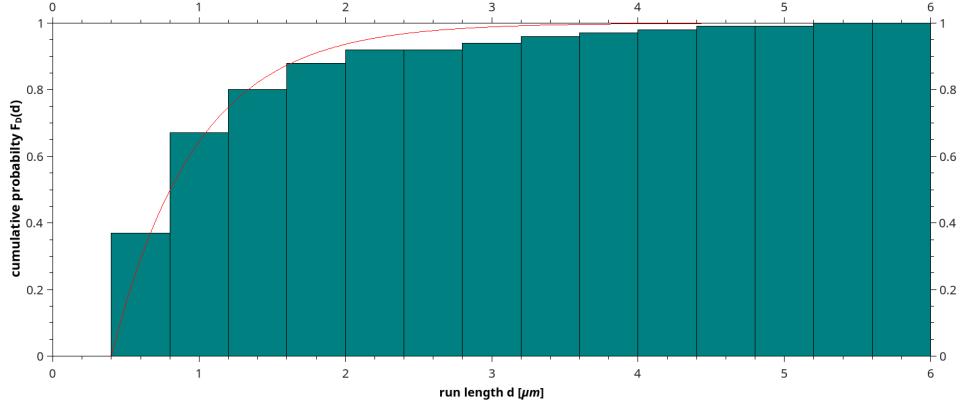


Figure 6: Cumulative probability disregarding the 'invalid' values with $d < d_0$

For the cumulative distribution function $F_D(d) = P_D(D \leq d)$ we use the fit equation:

$$F_D(d) = 1 - e^{-k \cdot (d - d_0)}$$

and extract the fit parameters:

$$\begin{aligned} k &= (1.7 \pm 0.4) \mu\text{m}^{-1} \\ d_0 &= 0.4 \mu\text{m} = \text{const.} \end{aligned}$$

With that equation we can choose an arbitrary confidence level $F_D(d) = \pi(d) \in [0, 1]$ which gives the probability of the run length D to be less or equal than given d . If we invert this function we get the *characteristic run length* $d = \pi^{-1}(\pi(d))$:

$$\begin{aligned} d(\pi) &= d_0 + \ln((1 - \pi)^{-1/k}) \\ \Delta d(\pi) &= \left| \frac{\partial d}{\partial k} \Delta k \right| = \left| \frac{\Delta k}{k} \cdot (d_0 - d) \right| \end{aligned}$$

We calculate this length for some characteristic values of π that correspond to the confidence levels of the Gaussian normal distribution for finding the value of the random variable x in an interval around the expected value: $x \in [-n\sigma + \mu, n\sigma + \mu]$, $n = 1, 2, 3$:

n	π	$d(\pi)$ [μm]	$\Delta d(\pi)$ [μm]
1	0.6827	1.1	0.2
2	0.9545	2.2	0.4
3	0.9973	3.9	0.8

We prefered the cdf-method to the pdf-method because we can check our result for given π just by looking at the histogram without an additional integration. The deviation of the pdf-fit with respect to the histogram is also larger than in the cdf. We could minimise the error Δd easily by taking more samples. We also should have avoided motorprotein-trajectories that are too 'short' to minimise the amount of values within the first bin. Also the magnification of the camera could be choosed larger to follow the trajectories better in the software

4 Discussion and conclusions

5 Appendix

6 References

References

- [01] Y.Wei; G.Wang: *An Intuitive Discussion on the Ideal Ramp Filter in Computed Tomography.* Iowa, 02/2004
- [02] <https://en.wikipedia.org/wiki/Histogram> [November 20, 2015]