Real time observations of single bacteriophage λ DNA ejections *in vitro*.

Paul Grayson, Lin Han, Tabita Winther, and Rob Phillips

April 10, 2007

Supplement A. Image processing.

Recorded movies of DNA ejection experiments were analyzed in two steps: First, ejections judged "good" (no DNA sticking to the slide, overlapping, or obvious photodamage) were manually selected from the movies, and 20 s of video, starting from the beginning of ejection, was converted into individual cropped image files. Second, these files were analyzed by a computer subroutine that automatically measured the length of the DNA using a *Difference-of-Gaussians* (DOG) filter [1, 2].

The DOG filter is used as a convenient approximation to the *Laplacian-of-Gaussian* (LOG) filter, an edge-detection algorithm that works as follows: We start with a raw image $\mathcal{I}$ that contains a certain amount of noise. The image is smoothed with a Gaussian filter, which we denote by $\mathcal{G}(\sigma)$. The standard deviation of the filter, σ, must be selected so that the filter erases most of the noise. Then, the Laplacian $\mathcal{L} = \nabla^2$ is applied to compute the curvature, which we denote by $\mathcal{C}$. Mathematically,

$$\mathcal{C} = \mathcal{L}(\mathcal{G}(\sigma) * \mathcal{I}) = (\mathcal{L}\,\mathcal{G}(\sigma)) * \mathcal{I}, \tag{1}$$

where $*$ represents the convolution operation. The final form of this expression follows because both $\mathcal{L}$ and $\mathcal{G}(\sigma)$ are linear: thus the LOG filter can be represented as a convolution of a single function, $\mathcal{L}\,\mathcal{G}(\sigma)$ with the image.

This function is closely approximated by the difference of two Gaussians with slightly different values of σ

$$\mathcal{L}\,\mathcal{G}(\sigma) \approx \mathcal{G}(\sigma) - \mathcal{G}(1.6\,\sigma), \tag{2}$$

which we call the DOG filter. In Fourier space, this convolution can be computed more efficiently as a product:

$$\tilde{\mathcal{C}} = \left(\tilde{\mathcal{G}}(\sigma) - \tilde{\mathcal{G}}(1.6\,\sigma)\right) \times \tilde{\mathcal{I}}. \tag{3}$$

The value of $\mathcal{C}$ is expected to change sign at edges, so by thresholding, the shape of the DNA may be extracted from the image. The following code, written in the Octave language, was used to apply the DOG filter to images and find the length of the given piece of DNA.

```
function [mylength] = find_length(img,sigma)
  w = size(img)(2);
  h = size(img)(1);

  ## generate the filter function
  g1 = zeros(h,w);
  g2 = zeros(h,w);

  for i=1:h
    for j=1:w
      ii = min(i-1,h+1-i);
      jj = min(j-1,w+1-j);
```

```
      g1(i,j) = exp(-(ii**2+jj**2)/(2*sigma**2));
      g2(i,j) = exp(-(ii**2+jj**2)/(2*(sigma*1.6)**2));
    end
  end
  g1 /= sum(sum(g1));
  g2 /= sum(sum(g2));
  dog = g1-g2;

  ## compute the curvature, C
  dog_f = fft2(dog);
  img_f = fft2(img);
  C = real(ifft2(img_f.*dog_f));

  ## compute the thresholded image, T
  cutoff = 0.2 * max(max(C));
  T = curvature > cutoff;

  ...

  mylength = rightedge - leftedge;
end
```

The removed section `...` finds the left and right edges of the largest region in the image. Figure 1 shows an example of the effect of the DOG filter, applied to image series from the text. As the figure shows, the size of small pieces of DNA is slightly exaggerated by a filter with a large value of σ, and the smallest pieces were entirely lost. We found that by reducing σ iteratively for smaller pieces of DNA, these problems could be eliminated.

Supplement B. Effect of flow.

In this section we present a brief theoretical treatment of the effect of flow and an additional plot in support the claim that the dynamics of DNA translocation is determined primarily by internal pressure rather than force from the flow.

As discussed in the text, one model for the state of a tethered piece of DNA in a shear flow is that there is a ball of unstretched DNA of length L_0 at the free end. The ball experiences a force from the flow; this force is what causes the remained $L - L_0$ of the DNA to be stretched out. We can use the Stokes formula to approximate this force:

$$F_{\text{flow}} = 6\pi\eta r v\,, \tag{4}$$

where r is the radius of the ball and v represents the average flow velocity over the ball. This force stretching out the DNA is balanced against its tendency to form a random coil: approximately $1\,k_{\text{B}}T$ of free energy is required for each persistence length ξ of DNA. Balancing the forces, we find

$$F_{\text{flow}} \approx 1\,k_{\text{B}}T/\xi \approx 0.1\text{pN}\,, \tag{5}$$

independent of the size of the ball. This force is trivial compared to the 10–40 pN of internal force found in λ, so we do not expect it to make a significant difference.

Figure 2 shows that when flow velocity is reduced by a factor of four, there is no significant change in the ejection process, indicating that the presence of a flow does not have an important effect on ejection. Additionally, Figure 3 compares the velocity of ejection under both flow rates, binned according to the method described in the text. What Figure 3 shows is that the translocation velocity is not increased under a stronger flow. In fact, for several data points, the translocation appears faster under the weaker flow. We believe that the faster points are due to the high fluctuations that are observed in a weak flow: in particular, the DNA tether calibration at 14 s^{-1} did not fit our model as well as it did at 57 s^{-1} (data not shown.) Our conclusion is that $57\ s^{-1}$ is a value that allows the DNA to be stretched out sufficiently to limit fluctuations, but without significantly affecting the translocation process.

References

[1] R. Fisher, S. Perkins, A. Walker, and E. Wolfart. Spatial filters - Laplacian/Laplacian of Gaussian. http://homepages.inf.ed.ac.uk/rbf/HIPR2/log.htm, 2003.

[2] David Young. Gaussian masks, scale space and edge detection. http://www.cogs.susx.ac.uk/users/davidy/teachvision/vision3.html, 1994.
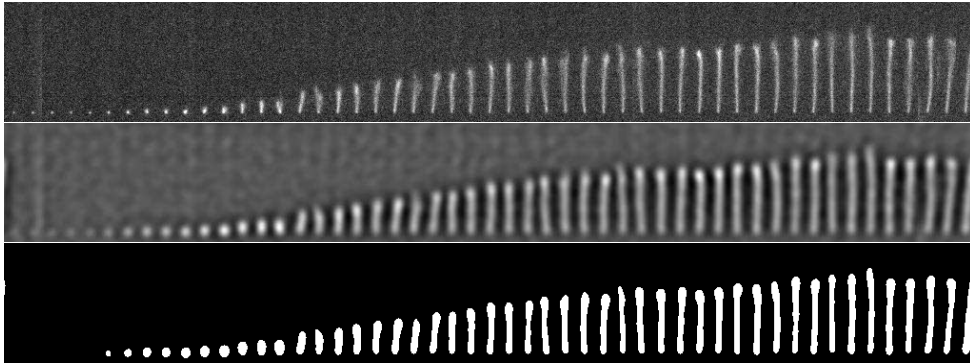
List of Figures

Figure 1: Three steps in analyzing images of ejected DNA fragments. Top: the original image captured on the camera. Middle: the result of applying the DOG filter, referred to as image C in the source code. A smoothing factor $\sigma \approx 0.5 \ \mu$m was used. Bottom: the thresholded image, image T in the source code.
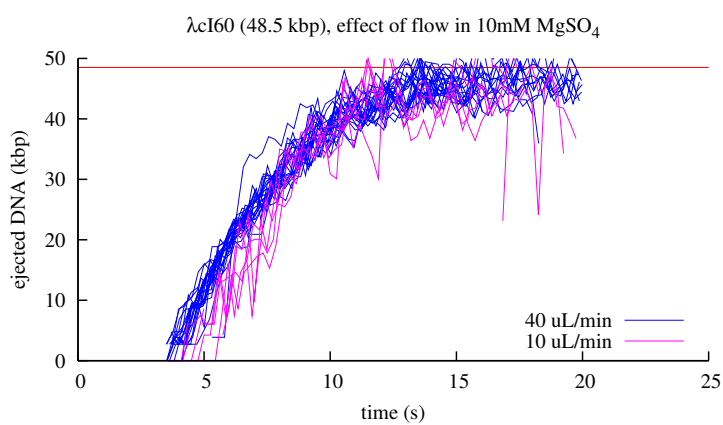
Figure 2: Graphs of the single ejection trajectories, showing the effect of the flow on the DNA. A four-fold change in the flow velocity appeared to have no effect on the ejection of DNA in $MgSO_4$ buffer. However, the data with a slower flow has an increased noise due to greater fluctuations of the DNA.
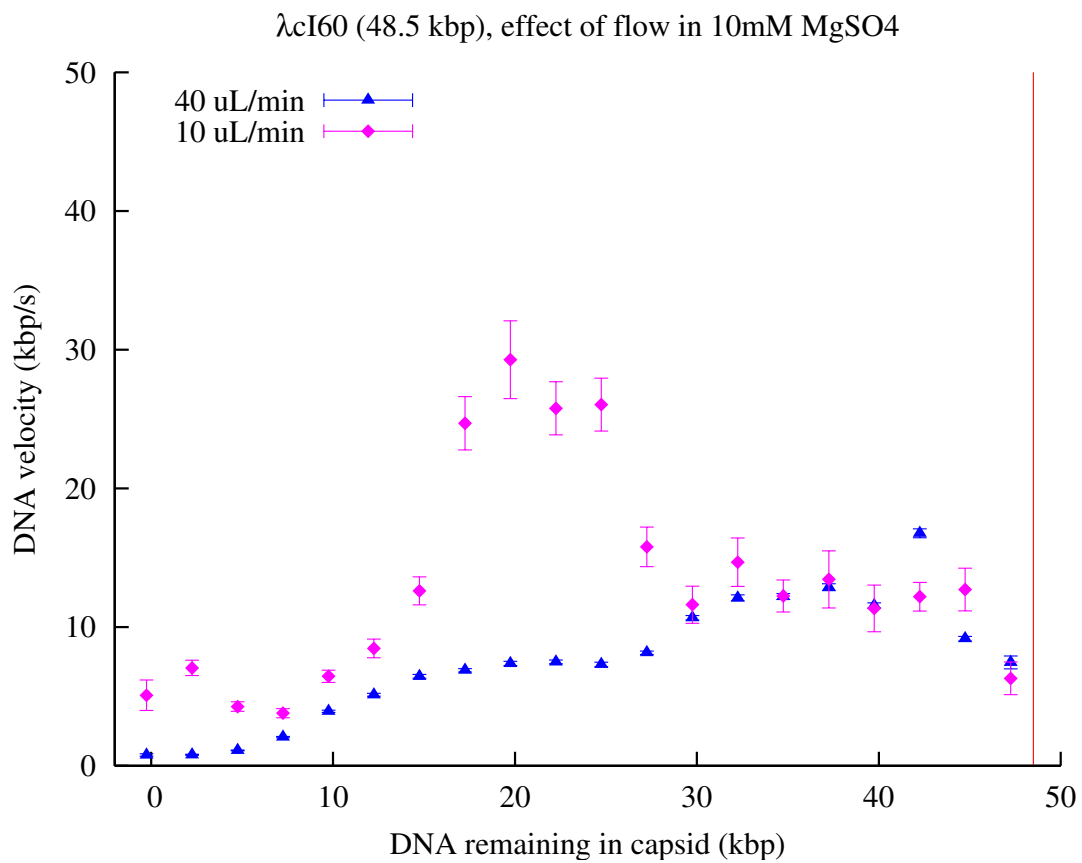
Figure 3: Comparison of the velocity of DNA translocation at two different flow rates: 10 and 40 $\mu L/min$, corresponding to shear flows of 14 and 57 s^{-1}, respectively. Except for several points during the middle of ejection, the velocities at both flows correspond closely, and even for the points that do not match, translocation appears faster in the slower flow. The red line indicates the position of the full-length genome, 48.5 kbp.

List of Tables

buffer	flow (μL/min)	shear (s^{-1})	L_0 (kbp)
Mg	10	14	45
Na	10	14	25
Mg	40	57	18
Na	40	57	8

Table 1: Calibration for various flow rates and buffers. The length L_0 of DNA that stretches out to 37% of its contour length was measured as described in the text. Mg buffer had an L_0 about twice as high as Na buffer, while the flow rate also had a large effect on the value. Additionally, it was noted that other buffers containing Mg^{2+} (TM buffer, buffer A) were equivalent to Mg buffer, indicating that it is the presence of Mg^{2+} rather than the absence of Na^+ that causes the DNA to be more flexible.