Detecting multiple oligomerization states by multidimensional analysis of FRET images

Kim Scott
Mentor: Henry Lester

Abstract: Fluorescence resonance energy transfer (FRET) microscopy is commonly used to measure distances between fluorophores or to qualitatively confirm interaction of proteins. Whereas multiple FRETing oligomerization states may be characterized clearly using single-molecule techniques, cellular applications suffer from wide and immeasurable variability in concentration of each species across pixels. Several examples are presented to demonstrate the inapplicability of Gaussian mixture model fitting to normalized FRET distributions to the problem of identifying populations of pixels of similar oligomerization state ratios. Instead, a direct clustering approach is developed in which each pixel is considered as a (FRET, acceptor, donor) triplet and assigned a probability of belonging to each of several clusters. The probabilities may be used to create an informative colored map of the cell indicating the locations of the various oligomer types.

Fluorescence resonance energy transfer

Fluorescence resonance energy transfer (FRET) occurs when a donor fluorophore in an excited state excites a donor fluorophore in the ground state through dipole-dipole interaction, without the transfer of a photon. The donor is excited by a laser near the peak of its absorption spectrum, and FRET is detected by recording fluorescence near the peak of the acceptor emission spectrum. Both donor and acceptor fluorescence are also measured so that the efficiency and not just the strength of interaction may be computed.

The distance at which resonance energy transfer is 50% effi**c**ient, called the Förster radius, is typically 20 to 90 Å. As the efficiency of energy transfer is proportional to R^{-6}, where R is the distance between donor and acceptor fluorophore, FRET microscopy is well-suited to the detection of oligomerization of proteins with attached fluorophores. Typical uses of this technique include measuring distances between fluorophores and qualitatively detecting covalent interaction (Lakowicz 2006).

Recent studies from the Lester lab at Caltech (Drenen et al. 2008, Son et al. 2009) have used FRET microscopy to study interactions among the subunits that make up nicotinic acetylcholine receptors (nAChRs), a class of pentameric ion channels. Whereas muscle nicotinic receptors have fixed composition, neuronal receptors may be composed of various combinations of subunits. Stoichiometry as well as presence of specific subunits determines sensitivity; for instance, the $(\alpha4)_2(\beta2)_3$ receptor has EC50 ~1μM, whereas the $(\alpha4)_3(\beta2)_2$ receptor is 100 times less sensitive. Upregulation of receptors in response to chronic nicotine exposure favors the high-sensitivity stoichiometry (Lester et al., 2009, review). A method is required for determining, from FRET images, the relative abundances of the various stoichiometries.

Current analytical techniques

The distributions of FRET efficiencies for multiple conformational states have been characterized for single-molecule FRET studies, which require a very dilute solution of fluorophores so that each pixel contains at most one FRET pair (Best 2007). In contrast, live cells may contain tens to hundreds of FRETing oligomers, as well as unpaired donors and acceptors. Partially assembled receptors (dimers and tetramers) may also be present in unknown geometries. Whereas the overall ratio of donor to acceptor may be measured in addition to FRET efficiency (Chen et al. 2006), the presence of multiple similar

stoichiometries of assembled receptors among unpaired and partially-assembled receptors has not been studied.

One standard approach that hopes to distinguish among stoichiometries involves normalizing the FRET intensities according to Xia and Liu (2001):

$$N_{FRET} = \frac{nF}{\sqrt{I_A \times I_D}} \quad \text{(Eqn 1)}$$

where the net FRET signal nF is defined by

$$nF = I_{FRET} - I_A \times BT_A - I_D \times BT_D \quad \text{(Eqn 2)}$$

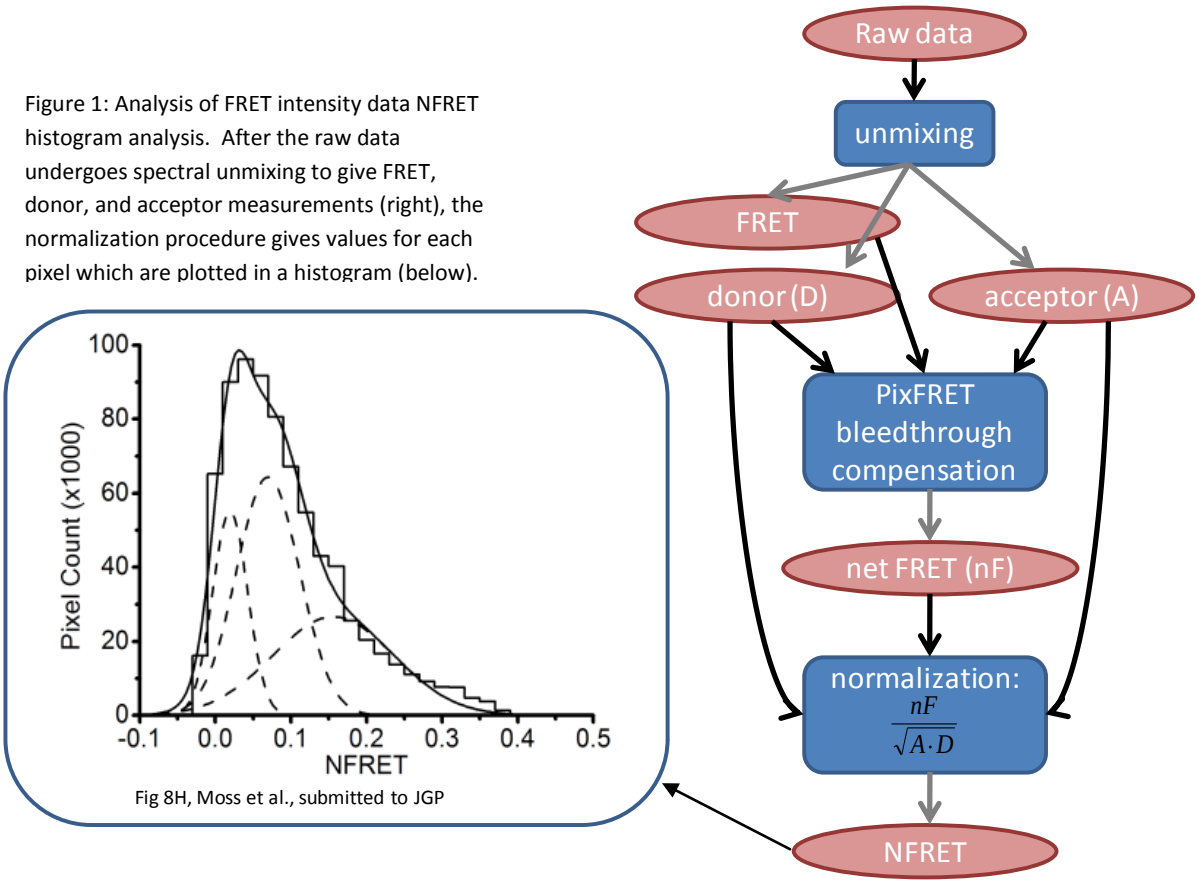


Figure 1: Analysis of FRET intensity data NFRET histogram analysis. After the raw data undergoes spectral unmixing to give FRET, donor, and acceptor measurements (right), the normalization procedure gives values for each pixel which are plotted in a histogram (below).

Fig 8H, Moss et al., submitted to JGP

I_{FRET}, I_A, and I_D are the intensities of FRET, acceptor, and donor fluorescence respectively; BT_A and BT_D are the bleedthrough coefficients for expected emission at the acceptor emission peak by the acceptor and donor due to excitation at the donor acceptance peak. The normalization procedure prevents the concentration of oligomers from influencing the NFRET value. Normalization is performed pixel-by-pixel using the ImageJ plugin PixFRET (Feige et al. 2005).

NFRET values are computed pixel-by-pixel and plotted in a histogram, sometimes monomodal, which is then fitted to a 2- or 3- component Gaussian mixture model (see Figure 1 for an overview of this approach). The components are assumed to represent collections of pixels which primarily contain one

oligomerization state, with overlap due to the various possible combinations of states in a pixel. This approach can distinguish several components of NFRET distributions (Moss et al., manuscript).

Fitting NFRET distributions to Gaussian mixture models, however, runs the risk of confusing skewness of the distribution due to the normalization method and distribution of oligomers with the presence of clearly-separated multiple populations. I present here an approach which relies directly on the three-dimensional data (nF, acceptor, and donor fluorescence intensities) to more accurately identify the stoichiometries present. This clustering approach can distinguish among mixed populations of receptors with subtly different ratios, even in the presence of unpaired fluorescence, and determine the appropriate number of clusters. It also automatically assigns pixels to populations, creating an informative map of the cell.

The case against fitting NFRET distributions to Gaussian mixture models

1. The NFRET value due to populations of two species with different pure NFRET values is generally nonlinear and not guaranteed to be monotonic. Given that a fraction *f* of the total FRETing constructs are of type A, the NFRET value

   $$T(f) = \frac{fnF_a + (1-f)\,nF_b}{\sqrt{(fA_a + (1-f)\,A_b)\,(fD_a + (1-f)\,D_b)}}$$ (Eqn 3)

   Figure 1 shows the total NFRET value in a single pixel with varying fractions of species A (NFRET value 0.1) and species B (NFRET value 0.2). Unless the denominator of T is constant, e.g. if $A_a = A_b$ and $D_a = D_b$, the NFRET value is nonlinear in f; additionally, for large differences between the donor:acceptor (D:A) ratio for species A and B, the NFRET value becomes nonmonotonic, i.e. the FRET value does not uniquely identify a ratio of species A to species B.

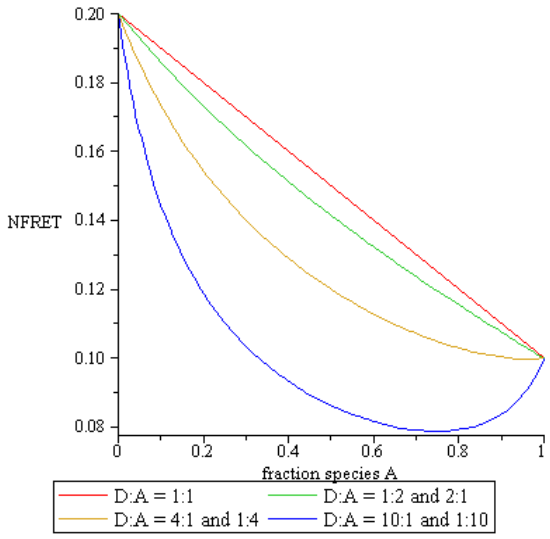


Figure 2 (left): NFRET value of a single pixel with varying contributions from species A and B.

   (See Figure 2).

2. Even with linear dependence on relative concentration, two populations binomially and independently distributed across the cell give skew NFRET distributions. The probability of a single pixel having NFRET value F is in this case

   $$\frac{\int_0^\infty P_a(Na)\,P_b((T^{\dagger}(F)-1)\,Na)\,dNa}{K_{norm}}$$ (Eqn 4)

The distribution of NFRET values is plotted for several cases in Figure 3. The skewness of the curves indicates that fitting to a Gaussian mixture model would be uninformative—a more skew curve requires more Gaussian components to fit, but the number and position of these components is not directly informative.

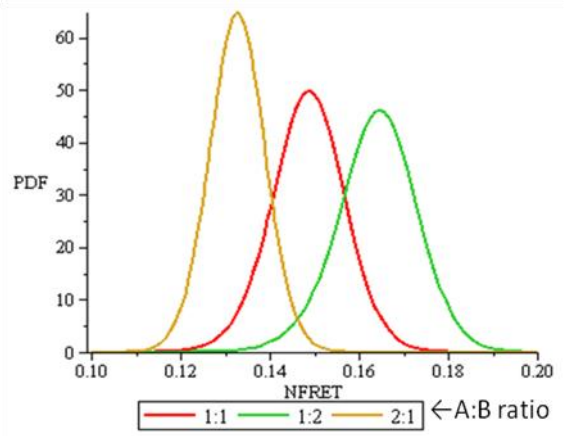


Figure 3 (left): NFRET value probability density function for several ratios of species A:B. Species A has NFRET value 0.1; species B, 0.2. Their acceptor to donor fluorescence ratios are both 1:1, so the skewness of the green and yellow curves is not due to nonlinear NFRET values as explained in (1).

3. Modeling the image as having three compartments—high A, low B; low A, high B; and "overlap" with equal low concentrations of A and B—demonstrates that the peaks of the NFRET distribution move in response to the concentration ratios (high:low), with the number of observable peaks dependent on the number of "compartments," the degree of overlap, and again the ratio fo high:low concentration. Figure 4 shows several such distributions, all with 1:1 ratios of total A:B expression and linear NFRET as described in point (1).

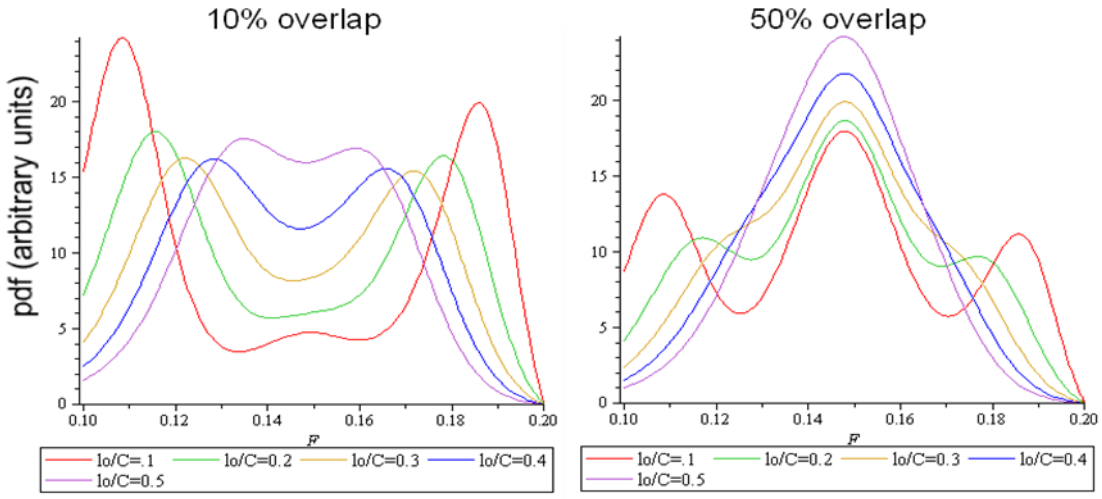


Figure 4 (above): NFRET probability distributions for a multiple-compartment model; each distribution is the sum of mostly-A, mostly-B, and "overlap" compartments.

4. Variation in single-species nF, acceptor, and donor fluorescence values yields highly skew distributions of NFRET values even for a *single* species, as shown in Figure 5.

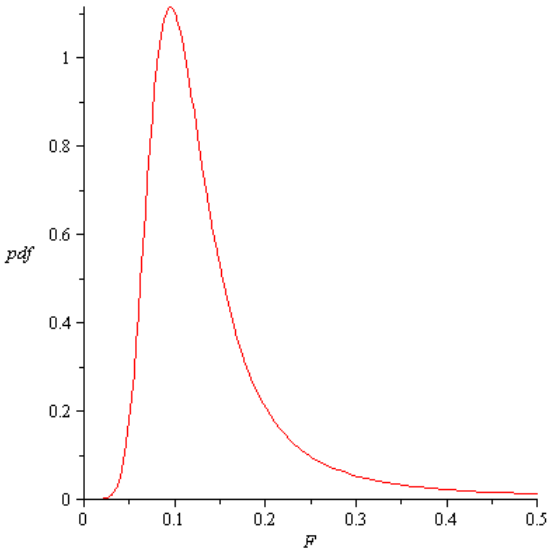


Figure 5 (left): NFRET pdf for a single-oligomer measurement when nF~N(1,0.2), A~N(10,4), and D~N(10,4).

NFRET is a poor normalization method for our purposes because it cannot uniquely identify stoichiometry ratios even under ideal circumstances. Fitting to Gaussians is dangerous because there are several sources of skewness inherent in the method. Additionally, it would be computationally prohibitive to compute the exact expected distribution of NFRET given all sources of variation.

The case for clustering in three dimensions

While normalization of FRET intensity to account for concentration is worthwhile, the NFRET procedure collapses a 3D datum at each pixel (nF, A, D) into a single datum, losing information. For instance, we would be powerless to make any distinction between two species with equal nF values but reversed A and D values using NFRET alone. It is, fortunately, unnecessary to normalize: instead, we should consider the (nF, A, D) triplets directly and attempt to cluster them along lines to identify populations of FRETing species.

Even when the NFRET histogram shows clear multiple peaks, Gaussian fitting cannot tell us which pixels belong to which distribution in the "overlap" region, which is often biologically relevant: for instance, are all the pixels of one distribution nearer the periphery? Clustering the pixels directly in 3-space gives us this information automatically. Ambiguous points may be assigned probabilities ("soft clustering") of belonging to three populations, and those probabilities mapped to color to produce a map of the cell.

Evaluation of clustering performance

Two clustering approaches were considered: fitting the (nF, A, D) triples directly to a Gaussian mixture model and using a projective k-means clustering algorithm to find lines (corresponding to specific stoichiometries or mixtures of stoichiometries) around which points were concentrated. Although both performed well in simple models (e.g. perfectly separated pure oligomer populations), Gaussian mixture (GM) clustering was more robust to unpaired fluorescence and other, more complex situations; further evaluation focused on this method. The Bayes Information Criterion (BIC) was used to identify the appropriate number of cluster centers. After identifying cluster means and variances in each dimension, posterior probabilities of each point belonging to the various clusters were computed and used to map

First, GM clustering was evaluated on a pure-population model in which each pixel expressed an average of 10 oligomers (binomially distributed) of one of two types, with varying fractional difference in nF mean. For each degree of separation, 25 "images" of 2000 pixels each were used. Figure 6 shows the mean and upper percentiles of the mistaken probability per pixel—the likelihood assigned by the algorithm that a given pixel belonged in the wrong cluster. At all fractional differences tested, two clusters were identified. Less than 5% of pixels were misclassified for fractional differences in nF > 0.3.

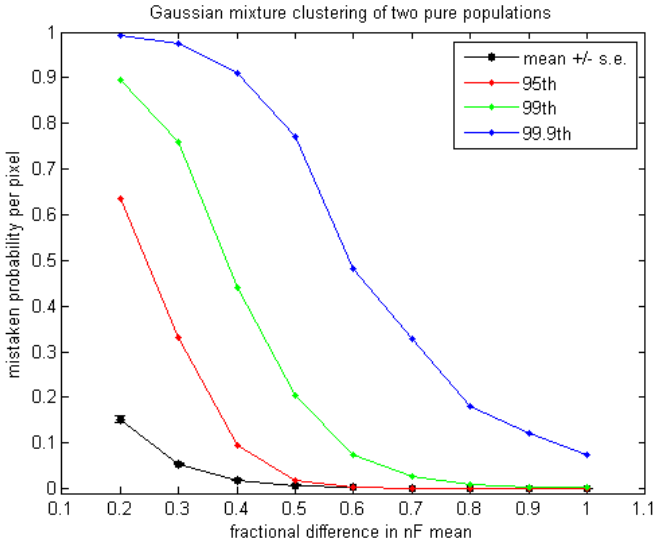
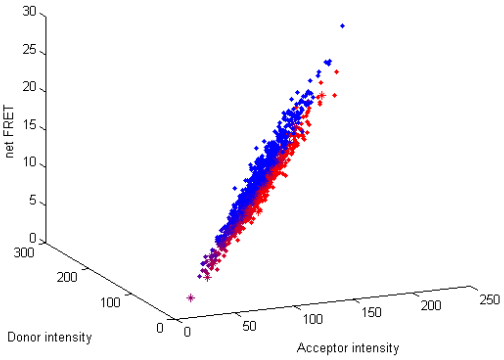


Figure 6 (left): Performance of GM clustering in the simplest case, two pure populations of oligomers with varying degrees of separation in NF values. Below: two populations identified, one in red and one in blue.



Next, GM clustering was tested on the segmented image model discussed in point (3). Figure 7 shows the mean and upper percentiles of mistaken probability per pixel in this case. Until the ratio of low to high concentration in each compartment exceeds 0.5, less than 5% of pixels are misclassified.

Note that normalization of concentration could eliminate one dimension on the plot; for instance, we could work only with nF/A and D/A instead of A, D, and nF. However, it is easier to work with approximately normally-distributed values of A, D, and nF than the distributions that result from moving a random variable to the denominator. Additionally, unpaired donor and acceptor fluorescence will simply shift a population on the 3-D plot without changing the slope, allowing us to focus on clustering around straight lines.

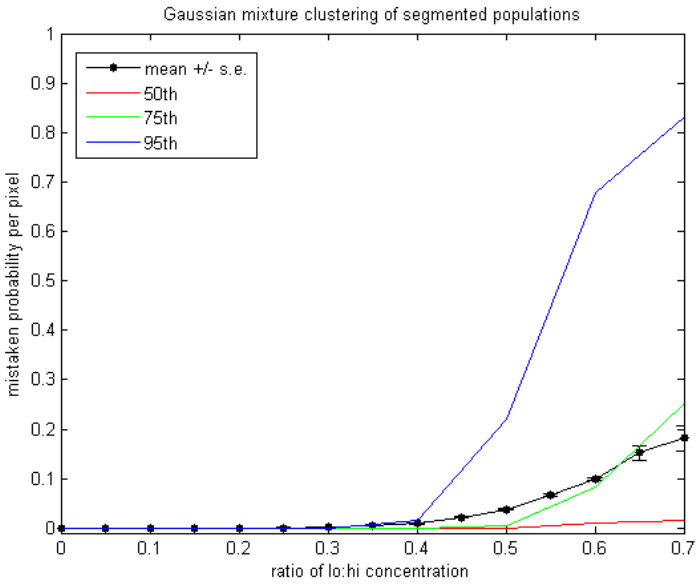


Figure 7(left): Performance of GM clustering in the case of a cell with two segments of equal size with a 2:3 ratio of nF values and varying ratios of relative concentration.

Discussion

While GM clustering is a promising approach for distinguishing among populations of varying stoichiometries of receptor throughout a cell, it requires rigorous testing on real datasets. In actual cells, membrane-specific and non-FRETing distributions may be used to calibrate expected cluster centers. Additionally, a method needs to be developed to allow the matching of identified clusters across cells with varying expression levels of each fluorescent construct due to individual differences as well as systematic variation of transfection ratios.

Acknowledgments

I would like to thank my mentor, Henry Lester, for guidance and support throughout the summer; Fraser Moss, Crystal Dilworth, Rigo Pantoja, and Rahul Srinivasan for discussions of analytical methods and the use of their experimental data; and the Amgen Foundation for financial support.

References

Best, R., K. Merchant, et al. (2007). "Effect of flexibility and cis residues in single-molecule FRET studies of polyproline." PNAS **104**(48): 18964-18969.

Chen, H., H. L. Puhl 3rd, et al. (2006). "Measurement of FRET efficiency and ratio of donor to acceptor concentration in living cells." Biophys J L39-L41.

Corry, B., D. Jayatilaka, et al. (2005). "A flexible approach to the calculation of resonance energy transfer efficiency between multiple donors and acceptors in complex geometries." Biophys J **89**: 3822-3836.

Drenan R. M., R. Nashmi, et al. (2008). "Subcellular trafficking, pentameric assembly, and subunit stoichiometry of neuronal nicotinic acetylcholine receptors containing fluorescently labeled alpha6 and beta3 subunits." Mol Pharmacol. **73**(1):27-41.

Feige, J. N., D. Sage, et al. (2005). “PixFRET, an ImageJ Plug-in for FRET calculation that can accommodate variations in spectral bleed-throughs.” Micros Res Techniq **68**:51-58.

Lakowicz, J. (2006). Energy transfer. Principles of fluorescence microscopy. Singapore, Springer**:** 443-528.

Lester, H., C. Xiao, et al. (2009). "Nicotine is a selective pharmacological chaperone of acetylcholine receptor number and stoichiometry. Implications for drug discovery." AAPS J **11**(1): 167-177.

Moss, F. J., P. I. Imoukhuede, et al. “Subcellularly resolved FRET explains GABA transporter function in N2a cells by detecting oligomerization states and plasma membrane anchoring.” Submitted to J Gen Physiol., 2009.

Okamoto, K. and M. Terazima (2008). "Distribution analysis for single molecule FRET measurement." J Phys Chem B **112**: 7308-7314.

Penton, R. and R. Lester (2009). "Cellular events in nicotine addiction." Semin Cell Dev Biol **20**: 418-431.

Son C. D., F. J. Moss, et al. (2009). “Nicotine normalizes intracellular subunit stoichiometry of nicotinic receptors carrying mutations linked to autosomal dominant nocturnal frontal lobe epilepsy.” Mol Pharmacol. **75**(5):1137-48.

Xia, Z. and Y. Liu (2001). "Reliable and global measurement of fluorescence resonance energy transfer using fluorescence microscopes." Biophys J **81**: 2395-2402.