The smfBox: an open-source platform for single-molecule FRET

*Benjamin Ambrose[1†], James Baxter[1†], John Cully[1†], Matthew Willmott[1], Elliot Steele[2], Benji C. Bateman[3], Marisa L. Martin-Fernandez[3], Ashley Cadby[2], Jonathan Shewring[1], Marleen Aaldering[1], Timothy D. Craggs[1]**

[1] Sheffield Institute for Nucleic Acids, Department of Chemistry, University of Sheffield, Sheffield, UK

[2] Department of Physics, University of Sheffield, Sheffield, UK

[3] Central Laser Facility, Research Complex at Harwell, Rutherford Appleton Laboratory, Oxford, UK

[†] *Equal contribution.*

**Corresponding author: t.craggs@sheffield.ac.uk*

Abstract

Single-molecule Förster Resonance Energy Transfer (smFRET) is a powerful technique capable of resolving both relative and absolute distances within and between structurally dynamic biomolecules. High instrument costs, and a lack of open-source hardware and acquisition software have limited smFRET's broad application by non-specialists. Here, we present the smfBox, a cost-effective confocal smFRET platform, providing detailed build instructions, open-source acquisition software, and full validation, thereby democratising smFRET for the wider scientific community.

Introduction

FRET is a photophysical process which results in the transfer of excitation energy from a donor fluorophore to an acceptor chromophore[1]. The efficiency of this transfer process scales inversely with the sixth power of the distance between the two chromophores. Therefore, by measuring the FRET efficiency (e.g. by observing the emission of the two fluorophores under excitation of the donor), spatial information can be determined in the 3-10 nm range, making FRET a 'spectroscopic ruler'[2] well matched to the dimensions of biomolecules such as nucleic acids and proteins[3]. In ensemble measurements this can be used to detect on-off / relative distance changes such as binding and cleaving in bimolecular interactions, or conformational changes (e.g. opening and closing) in unimolecular processes. At the single-molecule level FRET is sensitive to heterogeneous subpopulations, can measure kinetics of processes at equilibrium[4], and as demonstrated by a recent inter-laboratory benchmarking study[5], absolute FRET efficiencies can be used to infer precise distances for biomolecular structure determination[6–11]. Despite these advantages, smFRET

is currently rarely used outside specialist labs, largely due to the high costs of commercial instruments and lack of self-build, easy to use alternatives.

Results and Discussion

Here we present the **smfBox**[12], a cost-effective confocal-based platform capable of measuring the FRET efficiency between dye pairs on freely diffusing single molecules, using variable alternating laser excitation (ALEX)[13] for verification of correct dye stoichiometry and the determination of accurate FRET correction factors[5,14,15]. The smfBox (Fig. 1a,b) is constructed from readily-available optics and optomechanical components, replacing an expensive microscope body with machined aluminium, which forms a light-tight box housing the excitation dichroic, objective, lenses, and pinhole (see supplementary materials and online[12] for the complete parts list, build and alignment protocols). When assembled, the smfBox is sufficiently light-tight to allow safe and effective operation under ambient light conditions, as a Class I laser product (eliminating the need for user laser-safety training). The smfBox can be operated with either customisable LabVIEW acquisition software, or a stand-alone user interface written in C++. Both versions of the software provide all the necessary functionality for setting up the microscope (alignment and focusing) and recording data (Fig. 1c). The raw data, comprising of photon arrival times and detector ID, are saved in the open-source photon-HDF5 data format[16], and can subsequently be analysed either with the FRETBursts python module[17] using the Jupyter Notebooks provided (Supplementary Materials 6 and[12]), or with the MATLAB package PAM[18].

The smfBox detects both donor and acceptor emission from single molecules freely diffusing through a confocal spot under alternating laser excitation (Fig. 1 d and e). Emission under green excitation is used to determine the FRET efficiency (Supplemental Equation 1), whilst the response to the red laser confirms the presence of an active acceptor on the molecule, and allows the calculation of the stoichiometry parameter (S - Supplemental Equation 2) and the other correction parameters required for accurate FRET determination (Supplemental Equations 3-5)[5,13,15]. The precise ALEX cycle of the smfBox can be fully customised, allowing for faster or slower cycles, periodic acceptor excitation (PAX)[19], or an asymmetric ALEX scheme, which we show can reduce the width of FRET histograms, thereby increasing the resolution of different FRET species (see Supplementary Note 1). Furthermore, the design of the smfBox includes a 10:90 beam splitter in the excitation path, directing excitation light to a photodetector (Fig. 1a) to allow for precise monitoring of both laser powers in real time during the experiment, which can be saved into the HDF5 file. Our default optimal values of the laser powers, iris width, and ALEX cycle are provided for this setup (see online methods).

To test the performance of the smfBox we measured the FRET efficiencies of three DNA standards (Fig. 2a), which were recently characterised by multiple labs using a range of commercial and home-built microscopes[5]. Using the published correction procedures implemented in our open-source python analysis (Jupyter notebooks - see Supplementary Methods 6), we obtained data in excellent agreement with those from the other labs in the blind study. This provides both an excellent validation of the smfBox, but also a useful diagnostic for users to test their own builds of this instrument, as the successful reproduction of these data means that all hardware, acquisition and analysis software must be working correctly.

Furthermore, we have demonstrated the capability of the smfBox to recover rates of molecular conformational dynamics, using DNA hairpins as a test system (Fig. 1b). DNA hairpins have been shown to interconvert between a closed and open state, with rates that are dependent on NaCl concentration[20,21], within time scales accessible to smFRET experiments (see Supplementary Note 2). First, we reproduced data for a hairpin used in a recent study[20], using the dynamic photon distribution analysis (dPDA)[4] to determine opening and closing rates at a series of NaCl concentrations. Next, we analysed two further hairpins (Fig. 1c), identical in DNA sequence, but with a smaller and greater inter-fluorophore distance, to test the effects of the magnitude of FRET efficiency change, on the precision of the recovered kinetic parameters. As might be expected, the rates of interconversion for the hairpin with a higher FRET efficiency closed state could be determined more precisely, whereas analysis of the lower-FRET hairpin produced more variable results (Fig. 1d). In cases where the static (low-FRET and high-FRET) species have considerable overlap with each other (and therefore with the dynamic population) the dPDA model is less well constrained, leading to a greater variation in the values of recovered rates for a given sample size. These results have implications for the optimal positioning of FRET dyes when designing dynamic experiments.

Whilst the microscope we describe is built for ALEX confocal smFRET, its modular design makes for easy expansion to a number of related techniques. Without any hardware adaptations, the smfBox is capable of fluorescence correlation spectroscopy (FCS – using a single continuous-wave laser) and fluorescence cross-correlation spectroscopy (FCCS – using both lasers). The addition of one or more *pulsed* lasers and time-correlated single-photon counting (TCSPC) electronics will enable fluorescence lifetime correlation spectroscopy (FLCS) and pulsed-interleaved excitation (PIE) experiments[22]. The further addition of polarisation filters and two additional APDs would constitute a full multi-parameter

fluorescence detection (MFD) set up[23]. Furthermore, the addition of an XY-stage to the Z-positioning stage already included would facilitate any number of imaging techniques by scanning the sample. A list of recommended components for such expanded applications can be found in Supplementary Note 3.

To conclude, we have provided all necessary instructions and software required for the construction and operation of a cost-effective and open-source smFRET microscope with competitive capabilities[12]. We demonstrate the smfBox can determine absolute FRET efficiencies with the same accuracy as other instruments used by the community[5], and can recover biomolecular intercoversion kinetics in the range ~50-500 s^{-1}, in agreement with previous studies[20]. Finally, we have experimentally assessed the ability to determine kinetic rates of interconversion using dPDA, for systems with differing magnitudes of FRET efficiency change, providing useful information for the design of dynamic smFRET experiments.

Methods

The smfBox

Precise details of the construction and operation of the smfBox are described in the Supplementary Material and online[12]. Briefly; the smfBox alternates two lasers (515 nm – 52 mW and 635 nm – 15 mW, Omicron LuxX plus, powers measured at the laser) by TTL-controlled modulation of electronic shutters. The beams are coupled into a multi-mode fibre-bundle before being collimated (10 mm) and cropped (5 mm), then directed into a custom built aluminium microscope body. A dichroic mirror (Chroma ZT532/640 rpc 3mm) directs the beam into an objective (Olympus UPLSAPO 60× NA = 1.35 oil immersion), and the same objective collects the emission, which is focussed onto a 20 µm pinhole and split (Chroma NC395323 – T640lpxr) to two avalanche photodiodes (SPCM-AQRH-14 and SPCM-NIR-14, Excelitas), where photon arrival times are recorded by a national instruments card (PCIe-6353).

Accurate FRET experiments

Three duplex DNA constructs (referred to as 1a, 1b and 1c) labelled with Atto550 (donor) and Atto647N (acceptor), and were provided by the Hugel Lab as part of the blind, multi-lab FRET study[5] (for sequences see Supplemental Methods 5). DNAs were diluted to approximately 100 pM in observation buffer 1 (20□mM $MgCl_2$, 5□mM NaCl, 5□mM Tris, pH 7.5), and ~50 □l placed on a coverslip passivated with 1 mg/ml BSA, and data were

acquired by the smfBox. Analysis was done in Anaconda 5.3.0, with Jupyter Notebooks using the FRETBursts python module[17] (version 0.6.5), briefly: Background in each channel was estimated by means of an exponential fit of inter-photon delays. Bursts were identified using an all photon sliding window algorithm previously described[17,24] with L=10 and F=45 for both channels, and background was subtracted. Spectral cross talk factors were found by combining data from all standards and extracting bursts with a stoichiometry >0.95 as the donor only population and <0.175 for the acceptor only population to calculate α and δ. A dual channel burst search (DCBS) was then used to extract doubly labelled bursts from each 30-minute acquisition, and used to find E and S with single Gaussian fits. 2D Gaussian positions of each data set were then plotted together and fitted to obtain γ and β. Corrected FRET efficiencies of all doubly labelled bursts were then obtained using all four accurate FRET parameters as previously described[5]. See Supplementary Methods 6 for more details. Jupyter Notebooks and hdf5 files are attached and can also be found on the smfBox github[12]

Hairpin dynamics

DNA hairpins were made from a self-complementary oligonucleotide labelled with Cy3B annealed to a short oligonucleotide with Atto647N at different positions (Supplementary Methods 5). Oligonucleotides were purchased from LGC Biosearch (UK), with internal amino modified–dT bases which were labelled with Cy3B and Atto647N NHS esters purchased from GE Healthcare (US) and ATTO-TEC (Germany) respectively. Labelled DNAs were purified via polyacrylamide gel electrophoresis and annealed in annealing buffer (50 mM NaCl, 10 mM Tris HCl, pH 7.5, 1 mM EDTA), by heating to 95 °C (5 mins) followed by overnight cooling. Hairpins were then diluted to approximately 100 pM in Hairpin buffer (Tris 50 mM pH 7.5, BSA 0.1 mg/ml, EDTA 1 mM, Glycerol 5%, DTT 1mM) with additional NaCl where specified, placed into a chamber made of two coverslips and a silicone gasket (to enable >2 hour acquisitions with no sample evaporation, therefore maintaining a constant salt concentration), and data were acquired by the smfBox. Data were analysed using the MATLAB software package PAM[18]. Bursts were selected using a sliding window dual channel burst search, with a 50 photon threshold and a 500 μs window size. Doubly labelled bursts were selected between 0.2 and 0.85 S, bursts were cut into 0.5, 1, and 1.5 ms lengths. To access the precision of the kinetic parameters, acquisitions were split into subsets of 2000 molecules before further analysis. Dynamic PDA[4] was then used to fit a two state model to the data using the histogram library method. All hdf5 files for the hairpin experiments can be found on the smfBox github.[12]

Acknowledgements

We thank Gary Turner for his help in the mechanical workshop, in refining and producing the aluminium pieces for the smfBox. We are grateful for funding from the following sources: EPSRC studentships (B.A. and E.S.); STFC (B.C.B., M.L.M.-F. and T.D.C); EU Erasmus+ (M.A.); The University of Sheffield Start-up grant (T.D.C.); and the Royal Society (RGS\R2\180405, T.D.C).

Contributions

T.D.C designed the research; B.A., J.B., J.C., J.S. and T.D.C. built the initial instrument; B.A. and M.W. performed all validation experiments and data analysis. B.A., J.C., E.S. and B.C.B wrote acquisition and analysis software. J.B. and M.A. performed CAD design and rendering. M.L.M.-F., A.C. and T.D.C. provided supervision. B.A., J.B., J.C. and T.D.C wrote the manuscript with edits from all authors.

References

1. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann Phys* **437**, 55–75 (1948).
2. Stryer, L. & Haugland, R. P. Energy transfer: a spectroscopic ruler. *Proc. Natl. Acad. Sci. U. S. A.* **58**, 719–726 (1967).
3. Lerner, E. *et al.* Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. *Science* **359**, eaan1133 (2018).
4. Kalinin, S., Valeri, A., Antonik, M., Felekyan, S. & Seidel, C. A. M. Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. *J. Phys. Chem. B* **114**, 7983–7995 (2010).
5. Hellenkamp, B. *et al.* Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods* **15**, 669 (2018).
6. Hellenkamp, B., Wortmann, P., Kandzia, F., Zacharias, M. & Hugel, T. Multi-domain structure and correlated dynamics determined by self-consistent FRET networks. *Nat. Methods* **14**, 174–180 (2017).
7. Craggs, T. D. *et al.* Substrate conformational dynamics facilitate structure-specific recognition of gapped DNA by DNA polymerase. *Nucleic Acids Res.* **47**, 10788–10800 (2019).
8. Tsytlonok, M. *et al.* Dynamic anticipation by Cdk2/Cyclin A-bound p27 mediates signal integration in cell cycle regulation. *Nat. Commun.* **10**, 1676 (2019).
9. Nagy, J. *et al.* Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. *Nat. Commun.* **6**, 6161 (2015).
10. LeBlanc, S. J. *et al.* Coordinated protein and DNA conformational changes govern mismatch repair initiation by MutS. *Nucleic Acids Res.* **46**, 10782–10795 (2018).

11. Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225 (2012).

12. smfBox - Craggs Lab. https://craggslab.github.io/smfBox/.

13. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci.* **101**, 8936–8941 (2004).

14. Lee, N. K. *et al.* Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophys. J.* **88**, 2939–2953 (2005).

15. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).

16. Ingargiola, A., Laurence, T., Boutelle, R., Weiss, S. & Michalet, X. Photon-HDF5: Open Data Format and Computational Tools for Timestamp-based Single-Molecule Experiments. *Proc. SPIE-- Int. Soc. Opt. Eng.* **9714**, (2016).

17. Ingargiola, A., Lerner, E., Chung, S., Weiss, S. & Michalet, X. FRETBursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET. *PLOS ONE* **11**, e0160716 (2016).

18. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D. C. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophys. J.* **114**, 1518–1528 (2018).

19. Doose, S., Heilemann, M., Michalet, X., Weiss, S. & Kapanidis, A. N. Periodic acceptor excitation spectroscopy of single molecules. *Eur. Biophys. J.* **36**, 669–674 (2007).

20. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17**, 27862–27872 (2015).

21. Tsukanov, R. *et al.* Detailed Study of DNA Hairpin Dynamics Using Single-Molecule Fluorescence Assisted by DNA Origami. *J. Phys. Chem. B* **117**, 11932–11942 (2013).

22. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed Interleaved Excitation. *Biophys. J.* **89**, 3508–3522 (2005).

23. Widengren, J. *et al.* Single-Molecule Detection and Identification of Multiple Species by Multiparameter Fluorescence Detection. *Anal. Chem.* **78**, 2039–2050 (2006).

24. Nir, E. *et al.* Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments. *J. Phys. Chem. B* **110**, 22103–22124 (2006).

Figure Legends

Figure 1: The smfBox and smFRET. **a**: Schematic of the smfBox with parts labelled according to Supplementary Tables 1, 2 and 4. Lasers are collimated (L1), cropped (Iris) and steered by two mirrors (M1, M2) onto a 10:90 beam splitter (BS1). 10% of the beam is focused on to a photodiode for continuous power measurement and alternation cycle monitoring. 90% of the beam is directed via dichroic mirror 1 (DC1) to the objective. Light

from the back reflection is reflected by DC1 and BS1 onto a CCD camera for accurate focusing. Fluorescence emission from the sample passes through the excitation dichroic (DC1) and is focused onto a pinhole (P1) to remove out of focus light, before being split by colour (DC2) onto one of two avalanche photodiodes (APD0, APD1). **b**: 3D model of the completed smfBox, with the front panel of the microscope body removed. **c:** A flowchart of the smfBox platform showing the functionality of the acquisition software. **d**: Single molecules diffuse through a confocal volume (residence time ~1 ms) constructed by focusing the lasers into a near-diffraction-limited spot and using a pinhole to section the emission light in a thin focal plane. Lasers (515 nm and 638 nm) are alternated (20 kHz) to unsure multiple excitations of the donor and acceptor dyes for each molecule. **e**: A typical time trace for an smFRET experiment. Fluorescently-labelled molecules diffuse through the confocal volume emitting bursts of fluorescence. Individual photons generated by emission from the donor under donor excitation (DD – green) are recorded on APD0. APD1 records photons emitted by the acceptor either under donor excitation (DA – red) or direct acceptor excitation (AA – purple). **f**: 2D ES histogram showing uncorrected FRET efficiency (E*) and stoichiometry (S). Donor only molecules appear with low E* but high S, and acceptor only molecules appear with low S. Doubly labelled molecules appear with intermediate S.

Figure 2: Experiments validating the smfBox **a:** Fully corrected FRET efficiency histograms of three doubly-labelled DNA standards (1a, 1b and 1c, cartoons with dye accessible volumes) measured using the smfBox (grey). Vertical black lines and curves show Gaussian fits of our data, E=0.17±0.07, E=0.57±0.1, E=0.77± 0.07 (mean ± sd) , compared to the results from 20 other labs as part of a multi-lab benchmarking study[5] (red crosses). **b:** Proximity ratio (uncorrected FRET efficiency) histograms of a DNA hairpin at indicated salt concentrations. **c:** Salt dependent rates for hairpin opening (k_{open}) and closing (k_{close}) determined by dynamic photon distribution analysis (dPDA)[4]. **d:** Proximity ratio histograms of High-, Mid- and Low-FRET hairpins (at 400 mM NaCl). Data (grey) were fit using dPDA (black) to a two-state model, comprising a closed population (blue), open population (orange) and interconverting dynamic population (yellow). **e:** Plot of rates determined from dPDA of nine data sets for each hairpin, each containing 2000 molecules, quoting the mean and standard deviation across the data sets, with the mean chi-squared of the fits plotted to the right.