Single-molecule two-colour coincidence detection to probe biomolecular associations

Angel Orte*, Richard Clarke† and David Klenerman†1
*Department of Physical Chemistry, Faculty of Pharmacy, Campus Cartuja, Granada, 18071, Spain, and †Department of Chemistry, Cambridge University, Lensfield road, Cambridge, CB2 1EW, U.K.

Abstract
Two-colour coincidence detection (TCCD) is a form of single-molecule fluorescence developed to sensitively detect and characterize associated biomolecules without any separation, in solution, on the cell membrane and in live cells. In the present short review, we first explain the principles of the method and then describe the application of TCCD to a range of biomedical problems and how this method may be developed further in the future to try to monitor biological problems in live cells.

Single-molecule fluorescence
Fluorescence analysis of single molecules or complexes, one by one, has been developed and applied widely to biomolecules over the last decade to probe biomolecular structure, heterogeneity, dynamics and function [1–5]. Fluorescence provides sufficient sensitivity to detect and characterize different types of molecules or molecular complexes present in a sample, with a good signal to noise ratio and without the need for any separation, in contrast with bulk experiments which just measure the ensemble average. Single-molecule fluorescence experiments can be performed either in solution or on surface-immobilized molecules. Solution-based studies have the advantage that a large number of molecules can be probed and the disadvantage that the observation time is limited to typically 1 ms. Conversely, surface-immobilized studies have much longer observation times of up to 1–2 min, limited by photobleaching, but are performed on a smaller number of molecules. Solution experiments are normally performed using a tightly focused Gaussian laser beam and confocal detection in order to achieve the smallest possible probe volume with a sample solution at picomolar concentrations. Molecules can take different paths as they diffuse through the laser focus, giving rise to variation in the excitation rate of the fluorophore and hence the fluorescence intensity detected. In order to address this issue, ratiometric methods have been developed where two different fluorophores are attached to the same biomolecule and the ratio of their fluorescence intensities is measured as they diffuse across the laser-excited volume [6]. If the two fluorophores are sufficiently close and have appropriate spectral overlap, then it is possible to use only a single laser to excite one fluorophore, the donor, and get FRET (fluorescence resonance energy transfer) to the acceptor fluorophore. Such experiments are performed to measure the conformation of biomolecules, since the FRET efficiency depends mostly on the donor–acceptor separation [6]. In the more general case, the two fluorophores are excited independently by two different spatially overlapped lasers and coincident fluorescent photons are detected as the molecule diffuses across the laser-excited volume [7]. TCCD (two-colour coincidence detection) is based on detecting fluorescence photons emitted at two distinct frequency bands arising from continuous excitation of the confocal volume by one or two lasers. Although, in some cases, one laser is sufficient (for example, if the two fluorophores exhibit high FRET), it is usually very advantageous to be able to excite the two fluorophores with two independent lasers, typically one blue and one red. This allows the fluorophores to be placed at any convenient positions on the biomolecule without any requirement to place them close for FRET, which may not be possible if there is no information on the structure of the complex. In addition, a broader application of TCCD allows us to sensitively detect associated molecules when the fluorophores are placed in different subunits of a biomolecular complex. Analysis of the frequency and intensity of coincident fluorescence bursts on both channels allows us to quantify and determine the amount and stoichiometries of the associated molecules. One important advantage of TCCD is that it removes the constraints on the labelling of biological molecules, allowing single-molecule methods to be applied to systems whose structure and stoichiometry is unknown, and this has allowed the application of TCCD to a wide range of biological and biomedical problems. In addition, TCCD allows ultrasensitive detection of biomolecules and the measurement of intramolecular dynamics [8]. TCCD has also been shown to be less sensitive to background fluorescence than single-colour detection and to be capable of detecting femtomolar levels of molecular complexes [7].

TCCD
There are two major problems that became apparent as we started to apply TCCD to biological samples. First,
the formation of biomolecular complexes is often under conditions where the majority of the labelled components are non-associated or the complexes start to dissociate when diluted to picomolar concentrations for single-molecule analysis, producing non-associated molecules. These species increase the probability of chance coincident events that arise when two non-associated molecules happen to enter the probe volume at the same time. These events need to be removed from the data so we can just determine and analyse the true coincident events that arise from associated molecules. This is critical to obtaining high-sensitivity detection since, for low fractions of associated molecules, below 1%, the chance coincident events can be more frequent than the real coincident events. Since we could not determine the number of chance coincident events from reproducible control experiments, we devised two ways to do this directly from the experimental data alone [9]. One method is theoretical and is based on the event rates of blue- and red-labelled molecules. The second is based on red and blue time traces taken synchronously and desynchronizing them so that the red trace is paired with a blue trace taken at a different time. Any coincident events that are determined from these traces is purely due to chance.

The second problem is to again use the experimental data alone to determine the correct threshold level above which we should count events. If set too high, then we will miss events and need to collect data for longer, and if set too low, we will have problems with noise giving rise to events that are counted. The correct setting of the threshold is a particular problem with complex biological preparations or experiments on cells where the background level may change. To address this problem, we found it practical to optimize the background-corrected fraction of coincident events as a function of the thresholds used to identify events in the red and blue channels [10]. This again allows us to use the experimental data alone to determine the correct threshold values for data analysis.

Coincident events are analysed by a method that we have developed to extract the amount and stoichiometry of the complex studied. The frequency of coincident fluorescence bursts provides quantification of the complex, via the association quotient [9]. The stoichiometry is then obtained by forming ratio histograms of 

\[ \ln(I_{\text{red}}/I_{\text{blue}}) \]

where \( I_{\text{red}} \) and \( I_{\text{blue}} \) are the numbers of photons detected during the coincident burst in the red and the blue channel respectively [9,11]. As shown in Figure 1, if just one species is present, then the histogram is fitted by a Gaussian function, and the peak position depends on the stoichiometry of the complex. In the case that more than one species is present, then the histogram can be fitted to work out the fraction of each species present and their stoichiometry. There are cases that require a deeper analysis to extract the stoichiometry. For instance, both a 1:1 and 2:2 complex would give a very similar ratio histogram, so we also measure the mean brightness of the coincident events and found that the 2:2 complex is indeed twice as bright as the 1:1 complex, allowing us to distinguish between these two alternatives [12].

The most recent development we have made is the use of TCCD for samples where the red and blue fluorophores can also exhibit FRET [13]. We showed that when we excited the sample with two lasers, the sample detection efficiency was improved, and the FRET efficiency is also accessible by coincident burst analysis. Importantly, this allowed us to monitor the unfolding pathway of a labelled fluorescent protein that showed high FRET efficiency when folded, but that we could still detect upon unfolding if we excite with two lasers [14].

**In vitro experiments**

We have applied TCCD to samples of increasing complexity and larger size over the last few years (shown in Figure 2). TCCD has been used to study complexes including Protein G–IgG [15] and the dimerization of human telomerase RNA [16]. TCCD has also been used to determine the turnover rate and processivity of telomerase directly without amplification [11] and to determine the minimum stoichiometry of functional human telomerase [12]. We have also applied TCCD to single viruses of herpes simplex virus to look at the variation in virus assembly [17]. In the area of proteins we have applied TCCD to a GFP (green fluorescent protein) mutant labelled with an acceptor dye to follow the unfolding process and detected parallel unfolding pathways via an intermediate [14].

One of our goals is to apply TCCD to important biomedical problems. Hence we have also used the TCCD method to study the initial process of oligomerization of proteins, for instance the early stages of neuroserpin polymerization [18], related to familial encephalopathy, or to examine the assembly of oligomeric species on the pathway to formation of amyloid fibrils of the SH3 (Src homology 3) domain of PI3K (phosphoinositide 3-kinase) [19]. The mechanism of amyloid fibril formation is extremely challenging to study by conventional techniques because of the transient and heterogeneous character of the species formed [19]. The single-molecule experiments show that the species formed at the stage of the reaction where aggregates have been found previously to be maximally cytotoxic are a heterogeneous ensemble of oligomers with a median size of 38±10 molecules. Our experiments provided direct evidence for a general mechanism of amyloid aggregation in which the stable cross-\( \beta \)-structure emerges via internal reorganization of disordered oligomers formed during the lag phase of the self-assembly reaction. We have also employed TCCD in other biomedical problems. For instance, TCCD revealed how dyskeratosis-congenita-related mutations on the dyskerin protein hinder the formation of a stable complex with human telomerase RNA [20]. Furthermore, we have recently identified a new way in which the BRCA2 (breast cancer early-onset 2) protein functions during the DNA repair process, a process that goes wrong in cancer. We showed that BRCA2 plays a dual role, promoting the binding of the key enzyme RAD51 to single-stranded DNA in order to mediate repair and also inhibiting RAD51 binding to double-stranded DNA [21].
**Figure 1 | Principles of TCCD**

A red and blue laser are overlapped and focused to a diffraction limited spot. Non-associated molecules do not give rise to coincident events (A), apart from chance events that occur when two non-associated molecules enter the probe volume at the same time. In contrast, associated molecules such as dimers (B) and trimers (C) give rise to coincident events. Ratio histograms give Gaussian distributions whose peak position depends on the stoichiometry of the complex.

**Extension to live cells**

Generally, there is a paucity of methods for characterizing the subunit compositions of native protein complexes *in situ* that are useful in the context of the relatively low levels of protein expression often observed *in vivo*. We have exploited TCCD to detect the oligomerization state of native proteins on live cells and studied the composition of protein receptor complexes on the surface of live T-cells by tagging single subunits with mixtures of Fab fragments carrying either fluorophore [22]. We have shown that if the receptor of interest consists of two or more identical subunits, it gives coincident bursts of fluorescence as dual-colour-labelled receptors diffuse through the beam. In contrast, for a receptor comprising a single subunit, only non-coincident single colour bursts are detected, apart from occasional chance coincident events as occur in solution experiments. Control experiments with cells expressing the known homodimer CD28 gave a larger fraction of coincident events than cells expressing the monomer, CD86. We then showed that the TCR (T-cell receptor) comprises single αβ heterodimers, an important finding, since it constrains possible models for T-cell triggering. Moreover, this established TCCD as a general method to determine the oligomerization state of proteins on the surface of live cells and that it can be applied to endogenous proteins without the need for overexpression.

We have recently extended the concept of TCCD to two dimensions by using a CCD (charge-coupled device) camera to obtain simultaneous videos of the fluorescence from two distinct fluorophores, using TIRF (total internal reflection fluorescence) geometry [23]. The advantages of tracking molecules in two dimensions is that we can make measurements at lower expression levels (<1000 molecules per cell) and that there is no need to correct the data for chance coincidence events, where non-associated red and
blue molecules enter the same volume at the same time. Whereas for confocal measurements we used a statistical method to calculate the number of chance events and subtract them to determine the number of real coincident events, in two dimensions we can simply track the molecules for a number of frames to see whether the trajectories separate following their chance crossing at one point, or stay together, showing that the molecules are associated. In addition, we can detect immobile or slowly diffusing molecules on the cell surface. This method is applicable to both Fab and autofluorescent protein labelling and we also developed a method to extract maximum information from our videos of single molecules using a Bayesian-based analysis method [24]. In our initial experiments, we again validated the method by performing measurements again on T-cells. We used as a positive control the homodimer CD28 and as a negative control the monomeric protein CD86. These were labelled with both YFP (yellow fluorescent protein) and mCherry and we determined the fraction of all tracks that were coincident, showing there was a clear difference between these two controls. In addition, the TCR gave a coincidence level below the value for CD86, confirming once again that it is monomeric.

**Outlook**

TCCD has turned out to be widely applicable to biology. This is because many biological processes take place by changes in association between molecules, and, since these interactions are non-covalent, there will often be both fully formed and partially formed complex present. TCCD allows all of the associated molecules to be analysed without the need for any separation of the different complexes and non-associated molecules and allows their stoichiometry to be determined. However, even with the removal of the restrictions on the position of fluorophore labels, one of the slow steps in the application of the method, in our experience, is often the labelling of the biomolecules of interest without significant loss of biological function. Although the possibility of using protein fusion products with fluorescent proteins is a step forward, the development of easier labelling methods with single-molecule-compatible fluorophores would be a great help here.

TCCD and other variants of single-molecule fluorescence will be continued to be applied to larger macromolecular complexes, but one big area in the future will be to studying processes on or in live cells [25]. In these cases, it is not only the association and stoichiometry of the complexes formed, but also where they are located in two dimensions on the cell membrane or three dimensions in a cell, and how the associated molecules and their location change with time. This raises a new set of challenges for the detection methods, particularly due to the high autofluorescence in cells and the lack of very bright fluorophores which can easily be used as labels. However, the short history of single-molecule fluorescence shows that, within a decade, experiments that seemed extremely difficult are now routine in many laboratories all over the world, so these problems are not
insurmountable. The dream to follow key biological process such as transcription, translation and signal transduction at the level of single molecules in live cells in real time is still very much alive.

References