Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 5740–5747, June 1994

Review

Sorting single molecules: Application to diagnostics and evolutionary biotechnology

Manfred Eigen and Rudolf Rigler†*

**Max-Planck-Institut für Biophysikalische Chemie, Göttingen, Federal Republic of Germany; and †Karolinska Institutet, Department of Medical Biophysics, Stockholm, Sweden*

ABSTRACT **A method is described that provides for detection and identification of single molecules in solution. The method is based on fluorescence correlation spectroscopy, which records spatio-temporal correlations among fluctuating light signals, coupled with devices for trapping single molecules in an electric field. This technique is applied to studies of molecular evolution, where it allows fast screening of large mutant spectra in which targets are labeled by specific fluorescent ligands. The method expands the horizon in molecular diagnostics by making it possible to monitor concentrations down to (less than) 10^{-15} M without any need for amplification.**

Fast and efficient screening for particular properties exhibited by single particles at both the genotypic and the phenotypic level has become increasingly important in molecular biological research and for its applications in medicine and biotechnology. The "particles" may be single cells or parts of their surfaces, cell organelles, viruses, single genes, proteins (enzymes, receptors), or even small molecular entities such as peptide hormones or other oligomeric compounds. The need for detecting minute quantities may arise from the rarity of a desired structure in the presence of a huge number of alternatives—e.g., in the screening for antibodies, antigens, receptors, or for single mutants in evolutionary optimization; or it may be required to diagnose a disease in its early stages—e.g., the direct detection of a virus in an early phase of infection or evidence for cell transformation at an early stage of tumor development.

In this paper we describe a method to solve problems of detecting minute quantities of particles. Rather than trying to get hold of minute amounts of specific substances by collecting and concentrating them or—if present in the form of replicable molecules—by amplifying them as in PCR, we trace them at the level of the single entities. This tracing is done by detecting the quanta of fluorescence emitted by a probe that specifically binds to a target entity that is diffusing through an illuminated cavity. The corresponding signal can be discriminated from the background noise by auto- or cross-correlating the fluorescent light intensity fluctuations belonging to the emission from discrete molecules. The method is called fluorescence correlation spectroscopy (FSC), and its concepts were formulated in the early seventies (1–3). However, it was not until recently that the full potential of the method and its extreme sensitivity, thanks to new developments in laser technique and microscopy, became evident (4–7). This method is combined with screening and separation techniques used in applied molecular evolution (8–10) for which automated machines have been developed. In fact, it was the need for screening and early detection and selection of mutants during the process of molecular evolution that unified both groups in the research described here.

The results thus far are so encouraging that we expect revolutionizing consequences for three fields of research: (*i*) large-scale screening techniques to detect rare structures and activities; (*ii*) molecular diagnostics that may ultimately attain the greatest sensitivity possible for recognizing and identifying parasites, and (*iii*) evolutionary biotechnology that allows a biased adaptation to a variety of new functions.

This paper deals with three different methodological problems: single molecule detection, large number screening, and evolutionary optimization requiring the amplification of functionally identified rare structures. Additional applications are described, and we conclude with a section on applications of this molecular technology.

THE PRINCIPLE OF THE METHOD

Fluorescence is a highly specific and sensitive property that is frequently used in chemistry for identifying certain molecules. It can be characterized as a biphasic reaction between light quanta and resonant molecules. In the first phase of the reaction a light quantum of wavelength λ_1 is taken up by the resonant molecule producing an (electronically) excited state that has a certain lifetime τ_F. In the second phase, another light quantum of wavelength λ_2—that can be considered the reaction product—is emitted, upon which the molecule (like a catalyst) returns to its original state. Because the fluorescence lifetime τ_F is usually as short as 10^{-8} to 10^{-9} s, this process can be iterated many times before the vulnerable excited state will eventually fall victim to a decomposing (e.g., oxidation) reaction. Mainly because efficient fluorescent dyes are used as specific indicators in a wide variety of reactions, this method has become a powerful analytical tool in organic chemistry—particularly in biological chemistry. By linking a dye to a primer, any RNA or DNA sequence can be tagged specifically. Likewise a protein, such as a receptor, enzyme, or antibody, may be tagged via its labeled complement: its ligand, substrate, or antigen.

The usual way of recording fluorescence signals is to illuminate a sufficiently large volume (Fig. 1*a*) and to separate carefully the emitted fluorescence from scattered excitation light and other interfering luminescence. In all methods where an average intensity is monitored, there is eventually a finite limit of resolution set by the background noise. This is true, particularly if a fluorescent dye is used as an indicator that requires discernible differences between the wavelengths and/or the lifetimes of the fluorescence of the target-bound and the free dye molecules. In practice, it is difficult to measure concentrations $<10^{-9}$ M by these methods. However, 10^{-9} M still means nearly 10^{12} target particles per ml.

Is there a way to eliminate this principal limitation, a way that may even allow the detection of one single particle per volume element? An average intensity is, in fact, composed of events produced by single particles. However, it is the superposition of the concomitant "noise" and the desired signal while averaging over larger domains of space and time that masks the desired singular events. Hence, to detect the singular events we must avoid averaging over space and

Abbreviation: FCS, fluorescence correlation spectroscopy.

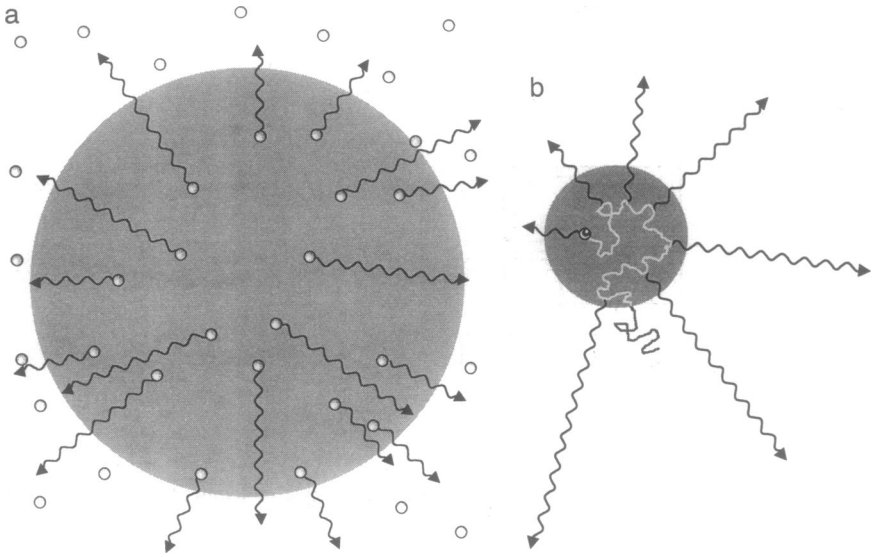

FIG. 1. Detecting the fluorescence of excited molecules. (*a*) Volume of light cavity (V_c) ≫ volume of particle territory (V_T), as is required for obtaining large mean intensities. (*b*) $V_c \ll V_T$, as is required for producing large fluctuations.

time. In other words, we must record signals derived from space elements that are small enough to host only single particles, and we must record the fluctuations of the signals with temporal resolution.

Let us return to the picture of a "reaction" between light quanta and molecules that characterizes fluorescence. We do not want to follow the progress of this reaction—as is usually done in chemical kinetics—as an integral over all elementary processes occurring in a larger domain of space and time. As suggested, we instead subdivide the space into small cells that are intensely illuminated by exciting light, and we record in discrete time intervals with a temporal resolution down to the fluorescence lifetime τ_F. How small must such space elements be? The answer is simple: the volume of the cell should be smaller than what we call the "territory" of a single target molecule. This so-called territory is simply the reciprocal of the particle concentration. For a 10^{-9} M solution we have 10^{15} particles per liter, accordingly a particle territory of 10^{-15} liter or 1 fl. Likewise a 10^{-15} M solution has a particle territory of 1 nl. These numbers clearly show that it suffices to attain spatial compartmentation of the order of magnitude of femtoliters. On the other hand, focusing light into space elements of 1 fl (and even smaller) has become possible in recent years using stable laser light sources in combination with confocal optics. Furthermore, time-resolved recording down to nanoseconds and faster can be easily achieved with modern techniques.

The process to be recorded can be described as a diffusion-controlled reaction of an individual fluorescent particle with an intensely illuminated discrete space cell (light cavity) of 0.1–1 fl in volume. This process is illustrated in Fig. 1*b*. Most of the time the light cavity is free of fluorescent particles; hence only noise is recorded (Fig. 2*a*). Occasionally, however, a fluorescent particle diffuses into the light cavity and then emits a burst of fluorescence quanta (compare the occasional burst in Fig. 2*b*). If recorded in a time-resolved manner, the quanta that belong to one fluorescing molecule can easily be identified by autocorrelating the time-resolved signals. For this purpose one correlates (i.e., multiplies) the intensity recorded at time t with that recorded at time $(t + \delta t)$, integrating the product $I(t)I(t + \delta t)$ over a finite time interval Δt, normalizing by the intensity and time of acquisition. If Δt is sufficiently large, the result is the average of the fluctuating amplitudes (rather than the mean intensity as in classical methods).

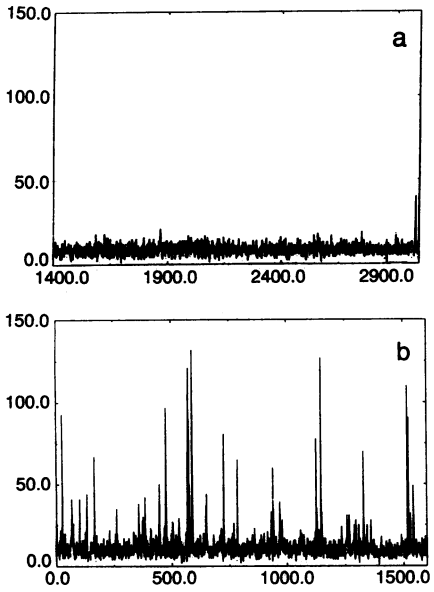

FIG. 2. Quantum bursts from single rhodamine molecules diffusing through a Gaussian laser beam. Channel time = 1 ms; diffusion time = 1 ms. The number of molecules in the light cavity is $N = 0.04$; the average count rate = 30 counts per channel. (*a*) Noise background. (*b*) Burst signals.

Quantitatively, the theory of diffusion-controlled reactions (11–14, 30) tells us that the encounter frequency of a particle to enter a sphere-like light cavity of radius R_c is given by

$$\tau_R^{-1} = 4\pi R_c D_i n_i, \qquad [1]$$

where D_i is the particle diffusion coefficient (usually given in cm^2/s) and n_i is the particle density (i.e., the number of particles per cm^3). If R_c is of the order of magnitude of 10^{-4} cm (1 μm) and if $D_i = 10^{-6}$ cm^2/s, one obtains for a concentration of 10^{-12} M $\tau \approx 1$ s. It is immediately apparent that concentrations down to 10^{-15} M (observation times of 15 min) are detectable in this way. By the same method we can calculate the average time a particle spends inside the light cavity to be

$$\tau_D = R_c^2/3D_i \qquad [2]$$

(12) or, with the above data: $\tau_D \approx 3 \times 10^{-3}$ s. This is more than enough time to yield a burst of several thousand fluorescent quanta during one encounter. The ratio of τ_D and τ_R

$$\tau_D/\tau_R = \frac{4\pi}{3} R_c^3 n_i \qquad [3]$$

is simply the ratio of the volumes of the light cavity and of the average particle territory yielding the probability for finding a particle inside the light cavity.

This opens up two avenues of approach. For rapidly diffusing particles it is advantageous to keep the light cavity at rest and let the particles diffuse into the laser beam. The signals from faster and slower diffusing particles can be separated by the process of autocorrelation. For slowly diffusing or immobile particles (viruses or cells in gels) one can move the light cavity relative to the sample. Scanning a cavity volume at a rate of 1 fl per ms permits one to *spot* one labeled particle in a 10^{-15} M solution within ≈15 min.

The important aspect of this method is that it focuses on single particles rather than on averages over large numbers. During the short time the particle is spotted it produces a large fluctuation in the temporally and spatially resolved signal that can be easily distinguished from the background noise. Averaging only over those fluctuations that can be identified by auto- or cross-correlation (and that refer to the average time the particle

spends inside the light cavity) eliminates the majority of unwanted noise. On the other hand, in combination with FCS, techniques for concentrating or separating particles such as electric focusing can be used, as will be described here. In this way a lower limit for the concentration of particles cannot be set; it depends on the skill of the experimentor and on the special properties of the particles to be observed.

THE FCS METHOD

Small Volume Elements with High Quantum Flux. By the use of diffraction-limited laser beams in combination with confocally imaged pin holes or fiber optics, a cylindrically shaped volume element with a radius of 200 nm and a length of 2000 nm can be intensely illuminated. The volume of such a "light cavity" is ≈0.2 fl. One femtoliter is approximately the size of an *Escherichia coli* cell, and one molecule in such a volume corresponds to a concentration of 10^{-9} M. The intensity of the cross section in the light cavity has a Gaussian distribution, and the longitudinal plane has a Lorentzian distribution. With a laser intensity of 0.5 mW a photon density of $\approx 1 \times 10^{24}$ photons per $cm^{-2}s^{-1}$ is achieved.

The use of a small volume for the light cavity is of paramount importance for detecting single molecules. The practical realization has become possible in recent years through innovations in confocal optics and laser technology. There are three reasons why a small volume of the light cavity is required: (*i*) The *signal-to-noise ratio* is allowing detection of *single quantum bursts* of fluorescent light. (*ii*) The *diffusion times* of the target molecules within the light cavity are kept short enough to prevent *bleaching* of the dye. (*iii*) It limits the *fraction of target molecules* simultaneously present in the light cavity and hence makes it possible to associate quantum bursts with *single molecules*. A schematic diagram of the experimental set-up is shown in Fig. 3.

Detection of Quantum Bursts by FCS. The ability to detect single molecules depends on the density of light quanta emitted. In the ideal case the number of quanta emitted per time interval is solely limited by the transition rate between the excited (singlet) state and the ground state. For rhodamine 6G with an absorption cross section of 1.4×10^{-16} cm^2 per molecule and a photon density of 1×10^{24} photons per $cm^{-2}s^{-1}$ a singlet excitation rate of 1.4×10^8 s^{-1} (per molecule) is reached. With a singlet decay rate of 2.5×10^8 s^{-1} in the absence of nonradiative transitions (quantum yield equal to 1), ≈25% of the ground state is depleted. Under these conditions we have been able to detect photoelectrons at a rate >100,000 per s per molecule. Taking into account the present detection efficiency and the band width of detection, this value corresponds to a rate of 2×10^6 emitted quanta per s. The discrepancy between the maximum possible and the measured emission rate is partially due to the existence of triplet states (15) and partially due to photochemical destruction during the excitation–deexcitation process.

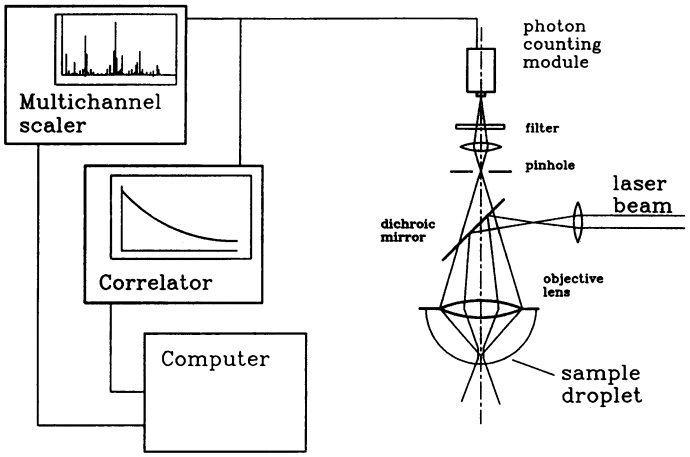

FIG. 3. Schematic of the instrumentation set up for single molecule detection and fluorescence correlation measurements.

An important parameter for single molecule detection is the signal-to-background ratio. The background is composed of scattered laser light, Raman scattering from the solvent, as well as of spurious fluorescence from the solvent and the optics. This background emission is approximately proportional to the third power of the radius of the light cavity (16). Although Rayleigh and Raman scattering can be suppressed by time-gated photodetection (17, 18) long-lived fluorescent background components will still contaminate the signal if they emit within the same wavelength range. We found that the best way to increase the signal-to-background ratio is by using small volume elements and continuous wave laser excitation. For the smallest volume element realized, a signal/background ratio of 1700 was obtained when the rotational Raman bands of water were suppressed (16); otherwise, a signal/background ratio of 40 was observed, indicating that the main part of the background emission is due to Raman scattering.

Random fluctuations of the intensity of fluorescence emission from individual molecules that have been excited by a stationary light source provide information on important molecular properties such as rotational (2) and translational diffusion (1), chemical kinetics (19), as well as on the lifetime of the excited state (2). The analysis of the signal is based on calculating the correlation function of the intensity fluctuations, as was shown by Elson and Magde in the case of translational diffusion and chemical kinetics, and by Ehrenberg and Rigler for the case of rotational motion and its coupling to the decay of the excited state. The success of the analysis rests on the ability to detect a sufficiently large photon flux per molecule (3). Photochemical destruction of the molecular species under observation has been a limiting factor for fluorescence correlation analysis.

The signal $I(t)$ generated by molecules traversing the volume element will fluctuate around a mean value $\langle I \rangle$ with the deviation $\delta I(t)$. The correlation function $G(t)$ is defined as

$$G(t) = \lim \frac{I}{2T} \int_{-T}^{T} I(t + t_1) I(t) dt_1$$

$$= \langle I(t) I(0) \rangle. \quad [4]$$

$G(t)$ can be divided into two parts: a constant term $\langle I \rangle^2$ and a time-dependent term $\langle \delta I(t) \delta I(0) \rangle$. Considering the (Poisson-distributed) statistics of small numbers and normalizing $G(t)$ by $\langle I \rangle^2$ shows that the limit of the amplitude of the time-dependent term at $G(t = 0)$ is given by the inverse number of molecules N within the volume element

$$\lim [\langle \delta I(t) \delta I(0) \rangle / \langle I \rangle^2] = 1/N. \quad [5]$$

The essential point is that we determine *correlated fluctuations* of emitted light quanta rather than average intensities. Intensities become largest when maximum numbers of light-emitting molecules are present, while fluctuations become negligible under these conditions. On the other hand, correlated fluctuations within a given time interval become greatest if all the light comes from a single source of emission. Autocorrelation was pioneered by Norbert Wiener as a powerful mathematical tool for noise reduction.

Involvement of Translational Diffusion. It is important to realize that diffusion is a "space-filling" process; this is due to its squared distance/time relationship. In

particular, this has consequences when comparing spaces of different dimensions. In spherical coordinates the theory of diffusion-controlled reaction yields for the encounter of diffusing particle "i" with a sphere-like cavity of radius R_c the frequency given by Eq. 1.

At first sight it may seem surprising that a recombination frequency is related to the radius R_c of the cavity rather than to its volume or some cross section, as in the case of collisions of gaseous molecules. It might seem even more surprising that a cylindrically shaped cavity with a very large ratio of its long half axis (l_c) to its radius (r_c) (or long and small half axes of a corresponding prolate ellipsoid) has a similar magnitude of the encounter frequency, provided that we compare the radius of the sphere R_c with the long axis of the cylinder l_c (13, 14). The actual relationship is

$$\frac{1}{\tau_R} = 4\pi D_i l_c n_i / \ln[2l_c/r_c]. \qquad [6]$$

In other words, a thread-like sink of length $2l_c$ and of comparatively small diameter ($l_c \gg r_c$) is—apart from the factor $\ln(2l_c/r_c)$, which may reach the magnitude of 10—as efficiently "hit" by a diffusing particle as a voluminous sphere that has a radius as large as the half-length of the thread. The explanation is that the trajectory of diffusion is space filling, and therefore the situation is strikingly different from the trajectory of a projectile. A better understanding of the above expressions can be obtained if we express, according to Einstein (20, 31), the (three-dimensional) translational diffusion coefficient by

$$D = \frac{\langle x^2 \rangle}{6\tau_x}, \qquad [7]$$

where $\langle x^2 \rangle$ is the mean square shift and τ_x is the time in which such a displacement occurs. If we choose for x a length that equals R_c, we obtain an encounter frequency of

$$\frac{1}{\tau_R} = \frac{1}{2\tau_{Rc}} (V_c/V_i), \qquad [8]$$

where τ_{Rc} is the time required to produce an average mean square shift of a diameter equal to R_c, V_c is the volume of the (spherical) cavity $4\pi R_c^3/3$, and V_i is the volume of the territory belonging to particle "i"—i.e., $1/n_i$. Physically, this expression means the following: if we subdivide the territory of each particle (volume V_i) into compartments of volume V_c, the chance of finding "i" in any one of the k compartments next to the cavity is $(V_c/V_i) \times k$ and the frequency of hopping into it is $1/(2k\tau_c)$ (i.e., k canceling out). This interpretation holds for the cylindrical shape as well. If $l_c \gg r_c$, there are many more cylindrical compartments of cross section $r_c^2\,\pi$ that fit into V_i than spheres of radius R_c, but the time for a mean square shift of size r_c^2 is correspondingly shorter than τ_{Rc}. Hence both encounters are of comparable efficiency. This interpretation has the advantage that we can easily transfer it to spaces of other dimensionality. The probability of a particle being in a certain space element is always the corresponding ratio of space element to territory (volume, area, length), and the frequency to move into it is the corresponding unidirectional "hopping" frequency. In all cases, the corresponding reverse process—i.e., hopping out of the cavity—is simply given by the same hopping frequency. These relations, of course, only hold in the stationary case, where for the total volume dn_i/dt is zero.

The above relations are important in the context of this paper because a decision must be made whether the laser beam should be focusing into a fixed volume element of space observing the particles as they diffuse in and out, or whether the entire space should be scanned.

In two-dimensional space, the squared dependence of diffusion on the distance is exactly balanced by the squared dependence of the particle distribution. For very slowly diffusing particles, scanning will always be advantageous. However, in scanning the diffusion coefficient is no longer available for differentiating particles according to their sizes. One may compensate for this loss by increasing the sophistication of the spotting procedures. Whenever one spots a particle, the scanning can be halted to determine the diffusion coefficient. Likewise, cross-correlation, as described below, may be employed to differentiate fluorescent labels that are bound in different states.

The correlation function for a molecule diffusing through a three-dimensional Gaussian intensity profile $I(x, y, z)$ with half axes r_1 and l_2 with

$$I_{x,y,z} = I_o e^{2(x^2+y^2)/r^2} \cdot e^{-2z^2/l^2} \qquad [9]$$

yields

$$G(t) = 1 + \frac{1}{N}\left(\frac{1}{1+4Dt/r^2}\right)\left(\frac{1}{1+4Dt/l^2}\right)^{1/2} \qquad [10]$$

and

$$G(t) = 1 + \frac{1}{N}\left(\frac{1}{1+t/\tau_{Diff}}\right) \times \left(\frac{1}{1+(r_2/l_1)^2 t/\tau_{Diff}}\right)^{1/2}, \qquad [11]$$

where τ_{Diff} is of the form $r^2/4D$.

A typical correlation curve for an 18-base primer for M13-DNA labeled with the fluorescent tag Bodipy is shown in Fig. 4*a*. The diffusion time of the primer is ≈0.16 ms and increases to 2.6 ms when the primer is bound to the M13-DNA (Fig. 4 *b* and *c*).

We now consider a diffusing label that interacts chemically with some target molecule. Specific interactions are characterized by high-stability constants (10^9–10^{15} M^{-1}) and slow dissociation rate constants (10^{-1}–10^{-6} s^{-1}). For this situation chemical relaxation times are much larger than the characteristic diffusion times $\tau_{chem} \gg \tau_{Diff}$, and the interaction of a fast diffusing ligand (*diffusion time* τ_{Lig}) and a slowly diffusing target (*diffusion time* τ_{Tar}) can be described as

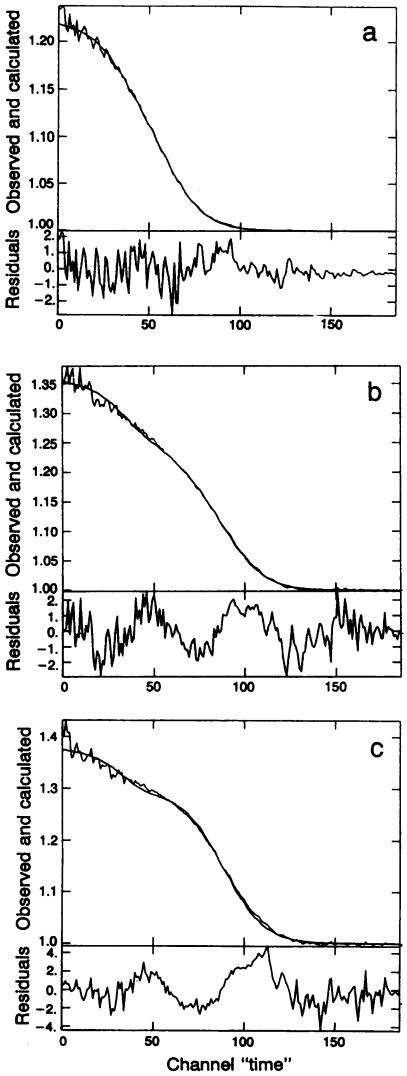

FIG. 4. Autocorrelation function of an M13 mp 18 (*t*) DNA primer, labeled with Bodipy (Molecular Probes), free and complexed with M13-DNA: (*a*) Free (30 nM, N = 4.5, τ_D = 0.166 ms), (*b*) 74% complexed, and (*c*) 100% complexed (τ_D = 2.6 ms). Lower panels: Residual between simulated and recorded correlation function.

$$G(t) = 1 + \frac{1}{N}\left(\frac{1-y}{1+t/\tau_{Lig}} + \frac{y}{l+t/\tau_{Tar}}\right), \quad [12]$$

where y is the fraction of bound ligand. Because the working concentrations of FCS are between 10^{-9} and 10^{-15} M, FCS is ideally suited for analyzing specific interactions. In addition, no physical separation of bound from unbound ligands is required to determine the binding isotherm. A detailed study on the acetylcholine receptor from *Torpedo marmorata* and its interaction with agonists (acetylcholine) and inhibitors (α-bungarotoxin) has been done (21), proving the validity of this analysis. Because the detection volume is of the order of femtoliters, only minute quantities are needed for the analysis.

Cross-Correlations of Stochastic Processes. The functions treated above describe the self- or auto-correlation of the fluctuating signal. Hence, all components with defined correlation times are observed. Instead of correlating the fluctuations of one signal, the fluctuations of two signals with different characteristics—e.g., different colors—can be correlated. In the case of specific interactions the cross-correlation function $G_x(t)$ will contain only those terms for which both signals contribute simultaneously to the fluctuation. In the case of specific interactions between tagged DNA primers and DNA sequences two primers with different tags (empty and filled circles in Fig. 5) can be used to interact with different sequences on the same DNA chain. The time-dependent part of the cross-correlation function will be nonzero only for those molecules that contribute to the random fluctuation and carry both primers (R, G). The cross-correlation of the translational diffusion of doubly tagged M13-DNA $G(t)$ will be

$$G^{R,G}(t) = \langle i^R \rangle \langle i^G \rangle + \langle \delta i^R(t) \delta i^G(0) \rangle. \quad [13]$$

In practice there will be some spillover between the detection channels for both primers, causing a contribution of the unbound primer to the correlation function. Its magnitude will depend on the properties of the band-pass filters used. With the available band-pass filters we estimate that doubly tagged DNA molecules can be detected in more than a thousandfold excess of ligand. The double-beam apparatus used in the Göttingen laboratory is shown in Fig. 6.

Requirements and Limitations. The results obtained in the preceding sections demonstrate well the range of applicability of this method. Diffusion coefficients of 10^{-5} to 10^{-7} cm^2/s allow concentrations of 10^{-13}–10^{-11} M to be monitored within 1 s. If measurements are extended over hours, concentrations $<10^{-15}$ M may be used. Usually a single molecule is involved in the process—e.g., at a concentration of 10^{-15} M, one molecule is contained in a volume of 1 nl, which is a million times the volume of the "light cavity" (<1 fl). Although these numbers compare extremely favorably with the limits of other analytical tools, we might ask whether they indicate absolute boundaries. If we can detect a single molecule, why not try to find it in any volume? Here we come to the really exciting aspect of this tool. For instance, how can one molecule be detected among a large number of alternative ones, and how can one guide and separate the detected molecule from the others? The numbers quoted do not represent principal limits; rather our method of detection must be combined with other techniques for screening, separation, and amplification.

TRACING SINGLE MOLECULES

The Problem of Relative Scarcity. A straightforward application of FCS would be to trace a particular gene, messenger, or virus with the help of a suitable (unique) antisense sequence that is labeled with a fluorescent dye. After annealing, the labeled primer would bind to its antipode (Fig. 5) and—after diffusing into the "light cavity"—it could be detected by FCS and identified by its diffusion time, which greatly differs from that of the unbound primer. Similarly, the

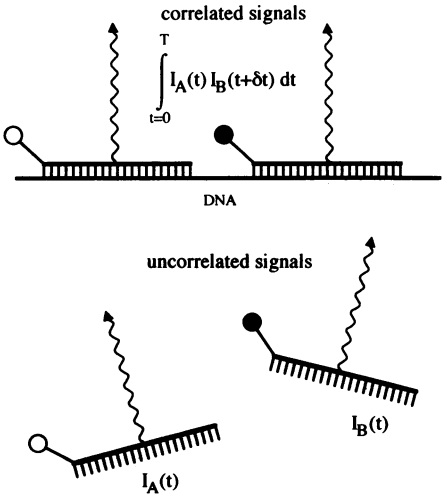

FIG. 5. Cross-correlation of viral DNA labeled with two primers emitting at different wavelengths.

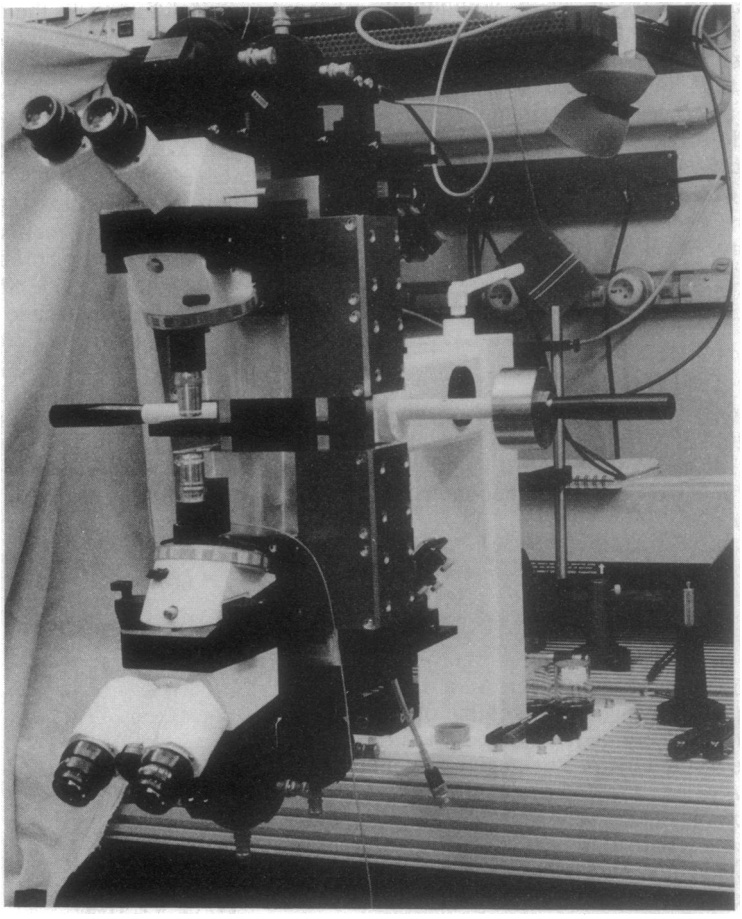

FIG. 6. Instrument for cross-correlation analysis (4π detection geometry). The sample is excited with two lasers at wavelengths λ_1 and λ_2. The emission is detected at two different wavelengths with detectors D_1 and D_2.

method could be used for screening particular hormone receptors with the help of their labeled hormones. There are many analogous cases. Under appropriate conditions one may "find" the rare molecule one is searching for; but how does one trap it and separate it from the other unwanted molecules?

A particular problem is differentiating between bound and unbound ligands if one is present in large excess. Recombination rate constants for enzymes with their substrates or for single-stranded nucleic acids with their antisense primers hardly exceed 10^7 $M^{-1}\cdot s^{-1}$, and these rates are often considerably lower. To bind a reporter molecule to its target within reasonable times concentrations must be at least 10^{-10} M or higher. Even then, one may have to wait for hours or days until equilibration is complete. On the other hand, the molecule searched for may be only one in a pico- (10^{-12} M), nano- (10^{-15} M), or microliter (10^{-18} M). In this case it would be difficult to detect the complex versus the overwhelming background of the unbound fluorescent label.

One way out of this problem has already been mentioned: cross-correlation spectroscopy (Fig. 5). It increases the sensitivity more than a thousandfold. Nevertheless, there is again a limit that cannot be exceeded.

Another way to lower the relative detection limits is to separate bound and unbound label. For this the binding constant must be very high, as is the case for specific oligonucleotide primers of a length of ≈20. Even if one would completely separate the unbound label, reestablishment of equilibrium would require a long time; this time would suffice to carry out the FCS measurements.

In the next section (*Electrical Trapping*) we describe electrophoretic methods and electrical traps that offer solutions to these problems. Traps can also be used to separate the traced molecule from its competitors. This is another problem of tracing particles present in small relative amounts. Tracing small relative numbers means handling large numbers of particles that are to be screened for the rare appearance of a particular singular structure. In *Screening of Large Numbers* we deal with screening procedures, and in the subsequent section we evaluate their merits.

Electrical Trapping. The goal is to separate the unbound fluorescent label from its macromolecular, viral, or cellular target with the help of electric forces. This type of separation is most efficient if the label and the target are oppositely electrically charged; this can usually be effected for protein ligand binding. However, what about the priming of nucleic acids, where both sense and antisense sequences are negatively charged?

A convenient and effective procedure is to use peptide nucleic acid rather than RNA or DNA strands for the fluorescent primer. Peptide nucleic acid is a polymer in which the phosphate sugar backbone of nucleic acids is replaced by a peptide-like backbone, based on the monomer 2-aminoethyleneglycin carrying any of the four nucleobases: A, T, G, C (22). This polymer, in contrast to RNA or DNA, is electrically neutral. To positively charge it, one could add, in addition to the fluorescent label, some lysine or any positively charged residues at one of the termini. The essential point is that this polymer binds favorably to RNA or DNA sequences as compared with corresponding antisense sequences made up of RNA or DNA (i.e., the RNA or DNA hybrid with peptide nucleic acid is more stable than the homo-double strand), (compare with ref. 22). Below the melting point a (positively charged) peptide nucleic acid antisense sequence of ≈20 nucleotides binds to its target RNA or DNA within a suitable equilibration time and remains bound to this target sequence for days, even if the excess unbound primer is removed (i.e., if the equilibrium is heavily disturbed again). Because the target RNA or DNA sequence labeled with primer carries an excess negative charge, it migrates toward the anode, in contrast to the unbound label.

We use this electric separation effect in two different versions to extend FCS to very low (relative) target concentrations.

In the first method FCS is coupled with capillary electrophoresis. The equilibrated sample is enclosed in a capillary containing a volume of ≈100 μl. One target molecule in this volume means a concentration of 10^{-20} M. The capillary has an outlet with an opening of ≈10^{-3} cm in diameter at one end, and the electrodes are arranged such that negative ions migrate toward this outlet with suitable speed. (A velocity of 0.1 cm/s requires a field strength of ≈1 kV/cm). The outlet of the capillary, in particular, poses a high barrier for positively charged ions due to the very large field gradient. However, the space outside of the outlet tip is critical. The outlet tip is the negative electrode for the static field within the capillary; simultaneously, another field is spanned between this tip and another sharp tip. These two tips are separated by a distance of only 10^{-3} cm, and an alternating electric field is applied between them. The light cavity is positioned between these two tips, closer to the sharp second tip (with a radius of curvature of only 10^{-4} cm). The alternating field drives a cloud of (collected) target molecules through the light cavity. In Fig. 7 we show fluorescence amplitudes from single molecules when such a cloud of fluorescent molecules oscillates between the two electrode tips that are separated by a distance of 10^{-3} cm (the applied voltage is up to 10 V). Two effects are observed: (*i*) a large amplitude of fluorescence is observed when the cloud of labeled targets traverses the light cavity; (*ii*) the fluorescence increases to saturation within the light cavity (which is placed between the electrodes) due to the formation of a diffuse double layer (or ion cloud) between the electrodes (the diffuse double layer is fed from the larger volume surrounding the field space between the electrodes) (unpublished data).

There are several problems associated with this method: (*i*) The capillary must be carefully shaped, and the electrodes

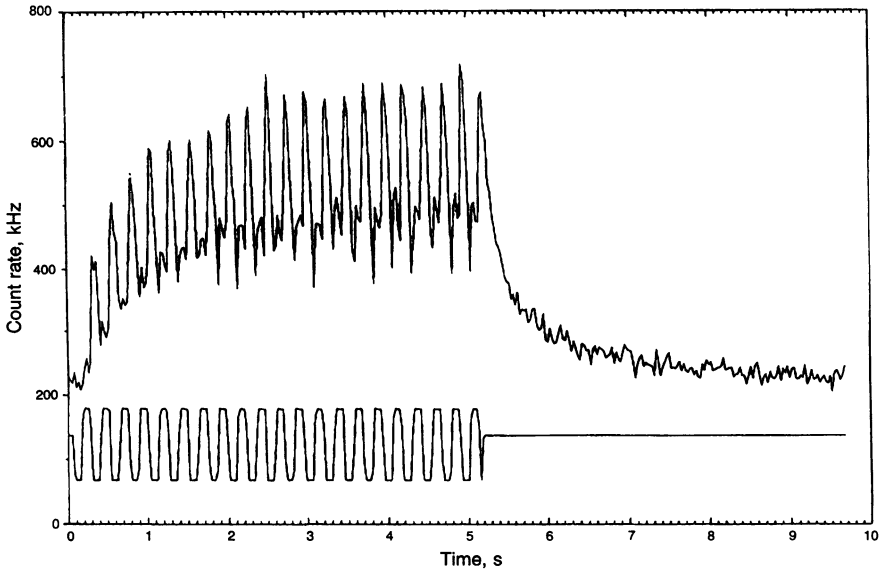

FIG. 7. Forced oscillations of rhodamine molecules between two electrodes separated by a distance of ≈10 μm in H_2O. (Upper trace) Signal from rhodamine-labeled dUTP. (Lower trace) Periodic electric field, 2 V (peak to peak), 4 Hz. The field is switched on at $t = 0$ and is switched off at $t = 5$ s.

must be arranged such that the potential at the outlet of the capillary is defined (e.g., ground) with respect to the other two electrodes, which are to be controlled. (*ii*) The material for the capillary must be selected carefully to establish opposite fluxes of positively and negatively charged carriers (preventing electroosmotic effects of water flow). (*iii*) The heat production and conduction in the capillary must be controlled when working with media of relatively large ionic strength (e.g., blood samples).

The method appears "electrophoretic" in nature. The critical feature, however, is the trapping of single molecules in the light cavity outside the static field that is superimposed on the separation process.

The second method of electrical trapping is even more sophisticated compared with conventional separation techniques. It is based on a real trap provided by a multipolar electrode arrangement with automatic voltage control, guided by the movement of the fluorescent molecule inside the light cavity. In this case, the light cavity is monitored with a bundle of parallel optical detectors (like the ommatidia in an insect eye) that detect the geometric position of the fluorescent molecule. The signals from these detectors are used to control the voltage at the multipolar electrodes by a feedback mechanism to keep the fluorescent molecule focused inside the (extended) light cavity (compare with Fig. 8). Here we face an open-ended development that will depend on the availability of new techniques for multiple parallel recording of light quanta through confocal optics. (The development of charge-coupled devices is only one step in this direction.) The screening techniques described in the following section will profit equally from any technological progress in this direction.

Screening of Large Numbers. Essentially, electric separation is a screening technique that allows the identification of a specifically marked compound among a large number of alternatives that do not respond to the specific marker. However, the purpose of screening is not just identification. It is important to isolate the identified compound and—even more important—to amplify it faithfully.

Isolation can be effected in two different ways. One follows immediately from the method described above. If we have identified a compound in the light cavity that is positioned in an electric trap, we could use a similar device—e.g., a capillary near the trap—to isolate it from the other components. With such a capillary we could suck up a pico- or femtoliter sample either mechanically—as in a vacuum cleaner—or eclectically, by using the charge of the particle.

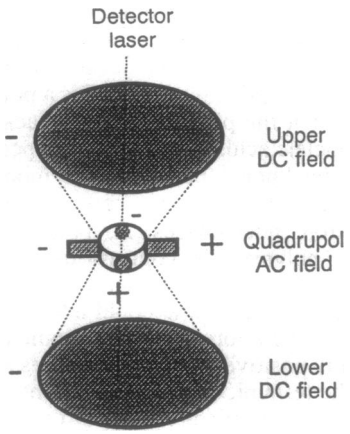

FIG. 8. Schematic of the quadruple trap for concentrating and analyzing single molecules. DC, direct current; AC, alternating current.

The other method is time screening of a sample that is spread on a plate, a method that we have used extensively in applied molecular evolution (23). The locally separated samples may be as small as pico- to femtoliters. Because recording times for FCS can be as short as 1 ms per sample, we can scan easily a million samples within fractions of an hour. The sampling can be done either stepwise (with picoliter-droplets containing mobile targets) or by continuous scanning of thin films containing immobilized particles. In stepwise scanning we can use wafer plates produced by modern etching techniques containing wells in the micrometer range. Such a device has been built and is being tested (24). Special care must be taken to prevent evaporation of the liquid. Carriers immobilized in the form of thin films can be handled more easily. They are especially interesting for screening substances that are exposed at the surface of viruses or bacteria (ref. 25 and B. Lindemann and M.E., unpublished data). In both cases the scanning device has to be addressable so that identified samples can be picked up easily.

The third requirement for screening is the amplification of an isolated sample. In the case of RNA or DNA this can be effected by replication. Because the samples contain minute amounts of replicable particles, one may even dispense with specific amplification techniques such as PCR. For protein screening, phage or bacterial-display systems offer an elegant solution. We have developed a bacterial-display method that is particularly adapted to FCS for application of this technology to the study of molecular evolution (B. Lindemann and M.E., unpublished data).

APPLICATION: NEW HORIZONS

Molecular Diagnostics. The main applications of single-molecule detection are to be expected in the fields of analytics and diagnostics. FCS represents an alternative to RIA or ELISA, especially when the concentrations are in the subnanomolar range, where other methods of direct analysis cease to function. With suitable instruments adapted to this purpose the recording times can be kept extremely short. The fact that single molecules are detected renders amplification such as PCR or other similar techniques superfluous. The use of primers as fluorescent labels guarantees the same specificity. Sample handling is considerably simplified. Diagnostics of virus diseases is carried to new levels. Fig. 9 presents an example of applying FCS to identify single virus particles.

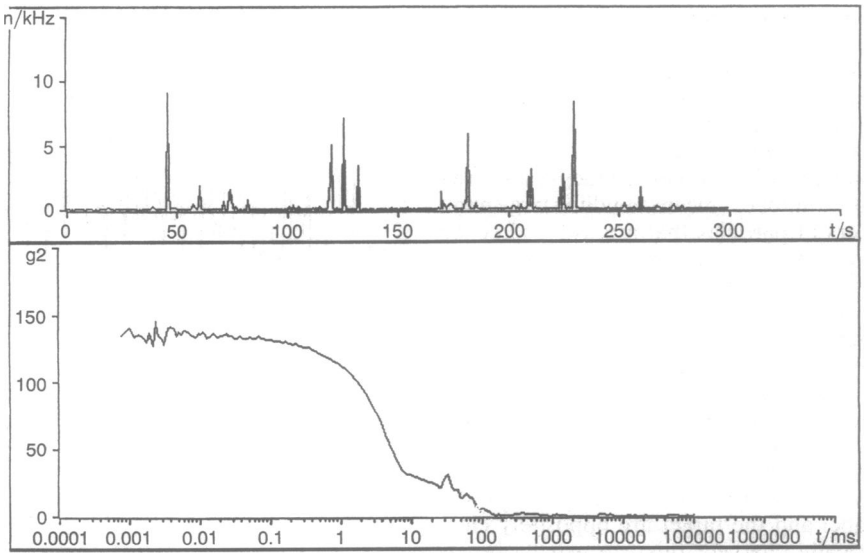

FIG. 9. Autocorrelation function of rhodamine-labeled M13-DNA molecules. (*Upper*) Transients from single M13-DNA molecules. (*Lower*) Autocorrelation function [46 pM, N (average number of molecules found in the light cavity) = 0.007, τ_D = 5 ms]. n, frequency of emitted fluorescent light quanta.

Screening Large Numbers for Rare Types in Molecular Evolution. What do we mean by "molecular evolution"? Superficially speaking, it is an extension of Darwinian algorithms to self-organization at the molecular level—i.e., variation and selection. This type of self-organization requires self- or complementary replicating molecular species, such as RNA or DNA. A theoretical treatment of such systems led to the concept of "molecular quasispecies" (26–29). The quasispecies is a broad distribution of mutants centered around one or several neutral master copies, usually showing a defined consensus sequence. The quasispecies is the target of selection. It is stable below an error threshold; the mutant distribution is close to this threshold and is widely dispersed, yielding optimal conditions for evolutionary adaptation.

Experimental studies of molecular evolution and its application to an evolutionary optimization of ribozymes, enzymes, catalytic antibodies, and other functional molecular entities require fast parallel screening of mutant spectra to find, isolate and amplify the best adapted molecular copy (8–10). The detection of single RNA or DNA genotypes as well as phenotypes, expressed and displayed at the surface of phages or bacteria, is one of our objectives in applying FCS. In fact, we started our joint project with this perspective in mind. The combination of FCS with large-scale screening (e.g., millions of clones in fractions of hours) coupled with genotypically controlled phenotype-display systems really opens up the horizons in this new field of research and technology.

DNA and RNA Sequencing. Another problem that triggered our interest is the use of FCS in RNA or DNA sequencing. The fact that single molecules can be recorded within milliseconds opens up entirely different approaches to this problem. A single sequence fixed in a very thin capillary may be degraded sequentially with the help of an exonuclease. The monomers are guided into the light cavity by using an especially adapted electric field trap. The bases of the molecule to be sequenced must be substituted by fluorescent analogues. It is sufficient to use just one fluorescent monomer for each pair if time-resolved detection of both the plus and the minus strand is possible. Improving the speed of sequence analysis to nearly one base per ms (limited by the turnover number of the exonuclease) extends the horizon for the human genome project and similar undertakings.

The joint research described in this paper was initiated by an Alexander v. Humboldt Award to R.R. who spent a sabbatical leave at the Max Planck Institute at Goettingen reviving a collaboration dating back to the late sixties. Several researchers from both granting institutions actively participated in the work reported in this review. We name, in particular, Dr. Masataka Kinio, Ulo Mets, and Jerker Widengren from Stockholm and Michael Brinkmeier, Dr. Bjoern Lindemann, and Petra Schwille from Goettingen. We further acknowledge cooperation with Evotec–Biosystems, Hamburg, Germany regarding applications in biotechnology. Research grants by the Max Planck Society, the Karolinska Institute, the A. v. Humboldt Foundation, the K. and A. Wallenberg Foundation, the German Ministry of Research and Technology, and the German Academic Exchange Service are acknowledged with gratitude. M.E. enjoyed the hospitality of Dr. Richard Lerner at the Scripps Research Institute at La Jolla, CA while preparing major parts of the manuscript.

1. Elson, E. L. & Magde, D. (1974) *Biopolymers* **13,** 1–27.
2. Ehrenberg, M. & Rigler, R. (1974) *J. Chem. Phys.* **4,** 390–401.
3. Koppel, D. E. (1974) *Phys. Rev. A* **10,** 1938–1945.
4. Rigler, R. & Widengren, J. (1990) *Bioscience* **3,** 180–183.
5. Rigler, R., Widengren, J. & Mets, U. (1992) in *Fluorescence Spectroscopy*, ed. Wolfbeis, O. J. (Springer, Berlin), pp. 13–24.
6. Rigler, R., Mets, U., Widengren, J. & Kask, P. (1993) *Eur. Biophys. J.* **22,** 169–175.
7. Rigler, R. & Mets, U. (1992) *Soc. Photo-Opt. Instrum. Eng.* **1921,** 239–248.
8. Eigen, M. & Gardiner, W. C. (1984) *Pure Appl. Chem.* **56,** 967–978.
9. Bauer, G. J., McCaskill, J. S., Otten, H. & Schwienhorst, A. (1989) *Nachr. Chem. Tech. Lab.* **37,** 484–488.
10. Eigen, M. (1993) *Gene* **135,** 37–47.
11. Smoluchowsky, M. V. (1916) *Physik. Z.* **17,** 557.
12. Eigen, M. (1954) *Z. Phys. Chem.* **1,** 176–200.
13. Eigen, M. (1974) in *Quantum Statistical Mechanics in the Natural Sciences*, eds. Kursunoglu, B., Mintz, S. L. & Widmayer, S. M. (Plenum, New York), pp. 37–61.
14. Eigen, M. & Richter, P. (1974) *Biophys. Chem.* **2,** 255–263.
15. Widengren, J., Rigler, R. & Mets, U. (1994) *J. Fluoresc.*, in press.
16. Mets, U. & Rigler, R. (1993) *J. Fluoresc.*, in press.
17. Rigler, R., Claesens, F. & Lomakka, G. (1984) in *Ultrafast Phenomena*, eds. Auston, D. H. & Eisenthal, K. B. (Springer, Berlin), Vol. 4, pp. 472–476.
18. Brooks-Shera, E., Seitzinger, N. K., Davis, L. M., Keller, R. A. & Soper, S. A. (1990) *Chem. Phys. Lett.* **174,** 553–557.
19. Magde, D., Elson, E. L. & Webb, W. W. (1974) *Biopolymers* **13,** 29–61.
20. Einstein, A. (1906) *Ann. Physik.* **19,** 289.
21. Rauer, B. (1993) Thesis (Univ. of Bielefeld, Bielefeld, Germany).
22. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1991) *Science* **254,** 1497–1500.
23. Köhler, M., Schober, A. & Schwienhorst, A. (1994) SENSOR, in press.
24. Schober, A. (1994) Thesis (Univ. of Braunschweig, Braunschweig, Germany).
25. Lindemann, B. F. (1992) Thesis (Univ. of Braunschweig, Braunschweig, Germany).
26. Eigen, M. (1971) *Naturwissenschaften* **58,** 456–524.
27. Eigen, M. & Schuster, P. (1979) *The Hypercycle: A Principle of Natural Self-organization* (Springer, Berlin).
28. Eigen, M., Schuster, P. & McCaskill, J. (1988) *J. Phys. Chem.* **92,** 6881–6891.
29. Eigen, M., Schuster, P. & McCaskill, J. (1987) *Adv. Chem. Phys.* **75,** 149–263.
30. Smoluchowsky, M. V. (1916) *Physik. Z.* **17,** 585.
31. Einstein, A. (1906) *Ann. Physik.* **19,** 371.