Available online at www.sciencedirect.com
SCIENCE DIRECT®
Analytical Biochemistry 327 (2004) 35–44
ANALYTICAL BIOCHEMISTRY
www.elsevier.com/locate/yabio

Single-nucleotide polymorphism detection using nanomolar nucleotides and single-molecule fluorescence

Charles R. Twist,[a] Michael K. Winson,[a] Jem J. Rowland,[b] and Douglas B. Kell[a,c,*]

[a] *Institute of Biological Sciences, Cledwyn Building, University of Wales, Aberystwyth SY23 3DD, UK*
[b] *Department of Computer Science, University of Wales, Aberystwyth SY23 3DB, UK*
[c] *Department of Chemistry, UMIST, Faraday Building, PO Box 88, Sackville St., Manchester M60 1QD, UK*

Received 27 June 2003

Abstract

We have exploited three methods for discriminating single-nucleotide polymorphisms (SNPs) by detecting the incorporation or otherwise of labeled dideoxy nucleotides at the end of a primer chain using single-molecule fluorescence detection methods. Good discrimination of incorporated vs free nucleotide may be obtained in a homogeneous assay (without washing steps) via confocal fluorescence correlation spectroscopy or by polarization anisotropy obtained from confocal fluorescence intensity distribution analysis. Moreover, the ratio of the fluorescence intensities on each polarization channel may be used directly to discriminate the nucleotides incorporated. Each measurement took just a few seconds and was done in microliter volumes with nanomolar concentrations of labeled nucleotides. Since the confocal volumes interrogated are ~1 fL and the reaction volume could easily be lowered to nanoliters, the possibility of SNP analysis with attomoles of reagents opens up a route to very rapid and inexpensive SNP detection. The method was applied with success to the detections of SNPs that are known to occur in the BRCA1 and CFTR genes.

Keywords: SNP; Detection; FCS; Single molecule; Miniaturization; Assay

It is now widely considered that the human polymorphisms that underpin the differential susceptibilities of individuals to disease, and their responses to pharmaceuticals and other environmental challenges, occur mainly at the level of the single nucleotide [1–15]. It is therefore of interest to develop rapid, large-scale methods for detecting such single-nucleotide polymorphisms (SNPs)[1] [15]. Many methods for allelic discrimination rely on differential hybridization of oligonucleotides [16] or PCR primers [17,18], including the use of microarrays [19] and protection on solid supports [20], whereas other methods that have been proposed include the use of pyrosequencing [21–23], RNAse cleavage [24], and matrix-assisted laser desorption ionization mass spectrometry [25,26].

The differential primer-based incorporation of nucleotides catalyzed by some kind of polymerase is attractive as it is conceptually straightforward and works well [15,27–33]. The discrimination of the alleles of interest can be effected by differences in the primers used or by using the same primer and discriminating which base is added by the polymerase.

It is clear that we shall need knowledge of millions of SNPs just to form SNP maps to establish any disease associations [1,8,9,12,14,34–36] (using advanced computational methods capable of effecting the necessary nonlinear mapping [37–41]), and these are already being produced. Then when we have established which alleles or haplotypes [7,42,43] are of interest it will be necessary to screen many more millions of individuals to determine which polymorphisms they possess. Thus there is a need for methods of SNP detection that are both cheap and high throughput.

A major driver for the future will thus involve a continuing need to make nucleic acid sequencing—as with other biochemical assays [44]—both faster and

* Corresponding author. Fax: +44-161-200-4556.
E-mail address: dbk@umist.ac.uk (D.B. Kell).

[1] *Abbreviations used:* SNP, single-nucleotide polymorphism; TAMRA, 5-carboxy-tetramethylrhodamine; FCS, fluorescence correlation spectroscopy; FIDA, fluorescence intensity distribution analysis; APD, avalanche photodiode; 2D, two-dimensional.

0003-2697/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2003.12.023

more sensitive and, because of the huge decreases in reagent costs that result from the latter, consequently much cheaper [45]. Although its realization is by no means new [46], the ultimate in sensitivity is represented by measurements at the levels of the individual molecule (e.g. [33,47–54]), and what is therefore required is a method that will sequence individual DNA molecules rapidly and reliably. We here describe a suite of methods that reveal that nucleotides when used at nanomolar concentrations can be incorporated into templates and detected using single-molecule spectroscopies, thus providing the first step toward this ultimate goal.

Materials and methods

Determining incorporation of each fluorescently labeled dideoxy nucleotide

The template-directed incorporation of a chain terminator [28,55] was used to extend the 3′ extremity of a primer by a single nucleotide. The model template sequence 5′-CAA AAA TAA CXA GGA GGC ATC CAC GGG ATT-3′, where X is any one of the 4 bases, was analyzed to simulate a single-nucleotide polymorphism (see Fig. 1). A 19-base match between primer and template ensured a melting temperature ($T_m = 58.8\,°C$) that gave the double-stranded DNA stability at room temperature. Using a DNA polymerase, the primer was then extended with a fluorescently labeled base to act as probe in the single-molecule fluorescence experiments.

A 31-mer oligonucleotide containing a single-nucleotide polymorphism was synthesized on the basis of the CFTR sequence. The SNP is located within a coding region. The sequence was

CFTR-Seq1: 5′-ACT TCT AAT GRT GAT TAT GGG AGA ACT GGA G-3′,where R was either one of the two purine bases.

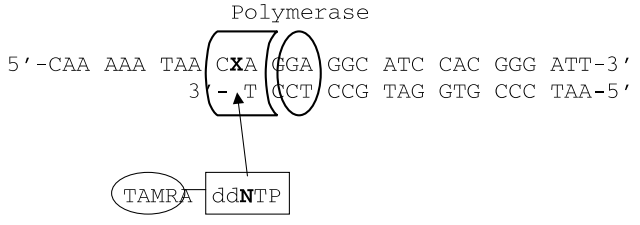

Fig. 1. Schematic representation of the process of template-directed incorporation (TDI). The 30-mer template includes an "unknown" base, denoted here as X; a complementary 19-mer primer is designed to anneal to the template up to and including the base immediately 3′ of X. A polymerase—in this case a modified bacteriophage T7 DNA polymerase—will then insert a dideoxy nucleotide (ddNTP) in the next position if N is complementary to X. A fluorescent dye, in this case TAMRA, is covalently attached to the ddNTP, so that the reaction may be monitored.

Two 31-mer oligonucleotides containing single-nucleotide polymorphisms were synthesized on the basis of the BRCA1 sequence. The SNPs in these sequences are within coding regions. The sequences were

BRCA1-Seq1: 5′-GTT TTT AAA GRA GCC AGC TCA AGC AAT ATT A-3′ and

BRCA1-Seq2: 5′-AGC GTC CAG ARA GGA GAG CTT AGC AGG AGT C-3′, where R was either one of the two purine bases.

Antisense oligonucleotides were synthesized to be complementary to the above strands between their 3′ extremities and the nucleotide just 3′ of R.

Enzyme and solutions

Dideoxy nucleotides labeled with 5-carboxy-tetramethylrhodamine (TAMRAddATP, TAMRAddGTP, TAMRAddCTP, and TAMRAddUTP—replacing TAMRAddTTP) were purchased from Perkin–Elmer Life Sciences (Hounslow, UK). Working solutions of each labeled base were prepared at a concentration of 100 nM. The DNA oligonucleotides were purchased in a lyophilized form from MWG-Biotech. Working solutions of each oligonucleotide were diluted in 10 mM Tris–HCl, pH 7.5, to a concentration of 100 μM. A 10 nM solution of 3′ TAMRA-labeled oligonucleotide was used to simulate the final incorporation product. A fluorescent replica of the primer oligonucleotide used above was annealed to the 30-mer model template. As the measured and calculated properties of the final incorporation product depend on the solution conditions, all measurements were carried out in the same buffers as the reaction.

The incorporation reaction used Sequenase 2.0 (Amersham Biosciences) (mass = 96 kDa; [56]) which was stored prior to use at 13 units μL^{-1} in 20 mM potassium phosphate buffer, pH 7.4, 1 mM dithiothreitol, 0.1 mM ethylenediamine tetraacetic acid, 50% glycerol at −20 °C. The procedure described in this paper is a modification of the procedure described in the enzyme manufacturer's documentation, although the buffer solutions remain unchanged. The reaction volume of 5 μL contained the three following elements: annealed DNA, labeled ddNTP, and enzyme. No unlabeled nucleotides were added.

Annealed DNA was obtained by mixing 1 μL of template (100 μM) and 1 μL of primer DNA (100 μM). To this was added 1 μL of buffer containing 200 mM Tris–HCl, pH 7.5, 100 mM $MgCl_2$, and 250 mM NaCl (reaction buffer). The primer and template were annealed by first heating to 85 °C and then cooling under controlled conditions over 30 min to 25 °C. The stock enzyme was diluted between 15- and 30-fold in a solution containing 50 mM Tris–HCl, pH 7.5, and 10 mM 2-mercaptoethanol (dilution buffer). (The polymerase concentration was chosen so as to avoid false positives caused by inter-

actions between the nucleotide and the enzyme.) Then 1 μL of diluted enzyme was added to the 3 μL of annealed DNA, followed by 1 μL of 100 nM labeled dideoxy nucleotide. The mix was incubated at 22 °C for 10 min prior to measurement.

Single-molecule fluorescence reader

All single-molecule fluorescence measurements were carried out with an Evotec Technologies (Düsseldorf) Insight single-molecule fluorescence reader. A 32×48-well microplate (purchased from Evotec) was used; each well holds 4 μL of solution. The duration of any one session was limited to 1 h to avoid evaporation problems. A He-Ne laser ($\lambda_{exc} = 543$ nm) was polarized in the vertical plane. It was then focused through an inverted confocal microscope, through the bottom of the sample plate, into the solution to a distance of 150 μm. The emitted light was collected through the same lens, filtered through a dichroic mirror, and split by a polarizer cube. The intensities of the parallel and perpendicular components of the polarized emitted light were measured by a pair of avalanche photodiodes in photon-counting mode. Laser power was set at 150 μW which strongly reduced the photobleaching and the triplet formation rates—the latter was estimated to be 10 to 15% with a decay time of 3 to 4 μs in all FCS analyses. All observations were carried out at room temperature.

The first analysis method is the long-standing [57,58] fluorescence correlation spectroscopy (FCS; see, e.g. [59–62]) which is a statistical analysis of the fluorescence intensity fluctuations. In this method, the autocorrelation is calculated for different times, τ, to yield the autocorrelogram $G(\tau)$. In the modern embodiment, and for particles subjected to Brownian motion, $G(\tau)$ is an exponential where the autocorrelation time corresponds to the typical translational diffusion time τ_{Diff} of a fluorescent particle through the confocal volume. This diffusion time is proportional to the cube root of the mass and to the solvent–solute friction. If the latter is assumed constant within the framework of one experiment, then this method provides an evaluation of the mass [59,63,64]. As the mass of the incorporated labeled base is far greater than the free labeled base, the FCS method should demonstrate primer extension. The value $G(0)$ is inversely related to the number of particles in the focal volume. FCS has been applied to DNA analysis in a number of ways [65–72] but never to sequencing by measuring dye incorporation in individual molecules directly.

The second analysis method is the fluorescence intensity distribution analysis (FIDA; [73–82]). This technique relies on a Poissonian analysis of the number of photons emitted during a specified dwell time, in these observations 50 μs, to reveal the number of photons emitted per fluorescent particle per second. By using two detectors, a two-dimensional form of FIDA can be implemented [79,81] to reveal the count rates corresponding to two different fluorescence properties under the same solution conditions. In these observations, the two photon detectors were set to measure the intensity fluctuations of the two components of the polarized fluorescence emission. The count rates on each photon detector can then be extracted and used to derive the anisotropy of the fluorescent dye.

Differences in the optical pathways to each detector needed to be corrected. As the incoming laser light is polarized only vertically, the correction factor G could not be calculated directly. Instead, a dye solution of known anisotropy is used, in this case, a 1 nM solution of TAMRA (in phosphate buffer saline plus 0.01% Tween 20), which has an anisotropy defined as $A_{True} = 0.023$. The count rates of the reference solution measured at each photon detector are used to extract a correction factor, G, by the equation

$$G = \frac{q_{\parallel} \times (1 - A_{True})}{q_{\perp} \times (1 + 2A_{True})}, \tag{1}$$

where $q_{\parallel}$ and $q_{\perp}$ are the count rates of the parallel and perpendicular components of the fluorescent light, respectively.

This value G is then used to correct the count rates obtained with a solution of unknown polarization, by the equation

$$A = \frac{q_{\parallel} - G \times q_{\perp}}{q_{\parallel} + 2(G \times q_{\perp})}, \tag{2}$$

where $q_{\parallel}$ and $q_{\perp}$ are as in Eq. (1).

Anisotropy depends both on the rotational motion of the fluorophore (and therefore its mass) and on its lifetime. 2D-FIDA measurements are therefore sensitive to either of these two and to any change in the count rate itself (which is related to the quantum yield). Any one of these parameters is thus potentially susceptible to change upon nucleotide incorporation.

The two types of measurement—FCS and FIDA—were carried out simultaneously using one optical configuration of the instrument. All measurements are the result of five scans, each lasting 10 s (although measuring times could have been reduced if desired—note that >100,000 assays per day is typical for high-throughput users of FCS/FIDA [59]). Measurements on a standard solution of 1 nM TAMRA (in phosphate-buffered saline plus 0.01% Tween 20) also allow the accompanying software to derive the volume parameters necessary to a proper FIDA treatment.

Results and discussion

To determine which base was incorporated at the 3′ extremity of the primer, four separate reactions were carried out, each with one TAMRA-labeled dideoxy

nucleotide, either A, G, C, or U. The fluorescence of the probe was then analyzed as described by either FCS or 2D-FIDA.

Fluorescence correlation spectroscopy

FCS was carried out on the perpendicular channel. The autocorrelation time for pure TAMRA in the standard solution was measured to be $205 \pm 5\,\mu s$. The autocorrelation time for 3′ TAMRA-labeled double-stranded DNA in the reaction solution was measured to be $755 \pm 28\,\mu s$. The autocorrelation time for each free TAMRA-labeled dideoxy nucleotide was measured in the reaction solution but using only single-stranded DNA; i.e., the antisense strand was omitted. The diffusion time τ_{Diff} was measured to be $318 \pm 11\,\mu s$ for TAMRAddATP, $278 \pm 10\,\mu s$ for TAMRAddGTP, $317 \pm 13\,\mu s$ for TAMRAddCTP, and $300 \pm 10\,\mu s$ for TAMRAddUTP. These values are longer than would be expected from a simple mass change between TAMRA and TAMRA-labeled dideoxy nucleotide. This increase is likely due to a change in solvent–solute friction, especially as mediated via the conformation of the nucleotides, and to the fact that the solutions are not identical. The slight variations in the diffusion time (which do not entirely follow the variations in mass) may be related to differences in the conformation of each labeled base.

Each dideoxy base was then tested for incorporation using the model template sequence. The resulting autocorrelograms are displayed in Fig. 2. When the base is incorporated, its autocorrelogram is very clearly shifted to longer times. These curves were then fitted according to a one-component autocorrelation function as embodied in the FCS method. It is not the object of the present work to present a quantitative time-dependent measurement of incorporation; therefore, a multiple-component analysis was not necessary. Furthermore, a one-component analysis presents the advantage that it can be fully automated, unlike the two-component fitting routine, which usually requires human intervention to fix certain parameter values. A one-component analysis can be carried out without a priori knowledge. Another advantage of one-component analyses is that they can be carried out on data of lesser quality than would be required for proper two-component analyses, thereby allowing shorter measurement times while preserving robustness. Finally, it should be noted that the fitting routine used in the analysis of the autocorrelogram assumes that the constitutive components share the same emission intensity; this will be shown not to be the case (vide infra). The diffusion times obtained from the one-component analyses of each case are given in Table 1.

For each dideoxy base (Table 1), it is clear when it is incorporated and when it is not and thus which complementary base or SNP is represented on the template. As can be seen, the apparent diffusion times of incorporated labeled bases are shorter than those of a pure 3′-labeled double-stranded oligonucleotide. This is probably due to the presence of unincorporated labeled base. When the enzyme was omitted from the reaction mixture the correlation time was the same as that for a sample in which the primer was omitted and close to that of the free dye (but slightly longer than it because of the oligonucleotide-induced change in solution viscosity).

Because the absolute diffusion times and the percentage of incorporation vary from one nucleotide to the next, it is slightly more difficult to compare the diffusion times for each primer (Table 1 rows)—which is the object of SNP detection—although the discrimination between the bases was very obvious. Although FCS is still valid as a method, since the characteristics of each dideoxy base are known when it is not incorporated, a method with even clearer contrast would be preferred.

Fluorescence intensity distribution analysis

FIDA was used in a two-dimensional form as there are two photon detectors; one each was used for each of the vertical and horizontal components of the polarized emitted fluorescence. The count rates (photons per fluorescent particle per second) on the vertical and horizontal channels for TAMRA alone in the standard solution are, respectively, 20.3 and 20.7 kHz. Its anisotropy is defined as 0.023. The correction factor G (Eqs. (1) and (2)) was thus calculated to be 0.92 for the data presented here although values were found to vary from 0.85 to 0.92 on different days according to the optical alignment.

The count rates for the TAMRA-labeled bases and the TAMRA-labeled double-stranded oligonucleotide in the reaction solution were measured. The results are given in Table 2. The count rates and the anisotropies are clustered according to the nature of the base, i.e., purine or pyrimidine. The count rates of the labeled free bases are approximately 10% lower than those of TAMRA, reflecting a slight degree of quenching. The fluorescence of the TAMRA-labeled double-stranded oligonucleotide, on the other hand, was strongly quenched, as a result of the close proximity of the nucleic acids in the oligonucleotide.

Each dideoxy base was then tested for incorporation using the model template sequence. A one-component analysis of the 2D-FIDA was chosen, as this routine is more robust than two-component fits due to the reduced number of variables. The values of the count rates were obtained and observed to be of the same order as that for the TAMRA-labeled oligonucleotide. The anisotropies for the sample solutions were then calculated and the resulting values are given in Table 3. As with the FCS analysis, incorporation of a dideoxy base is very

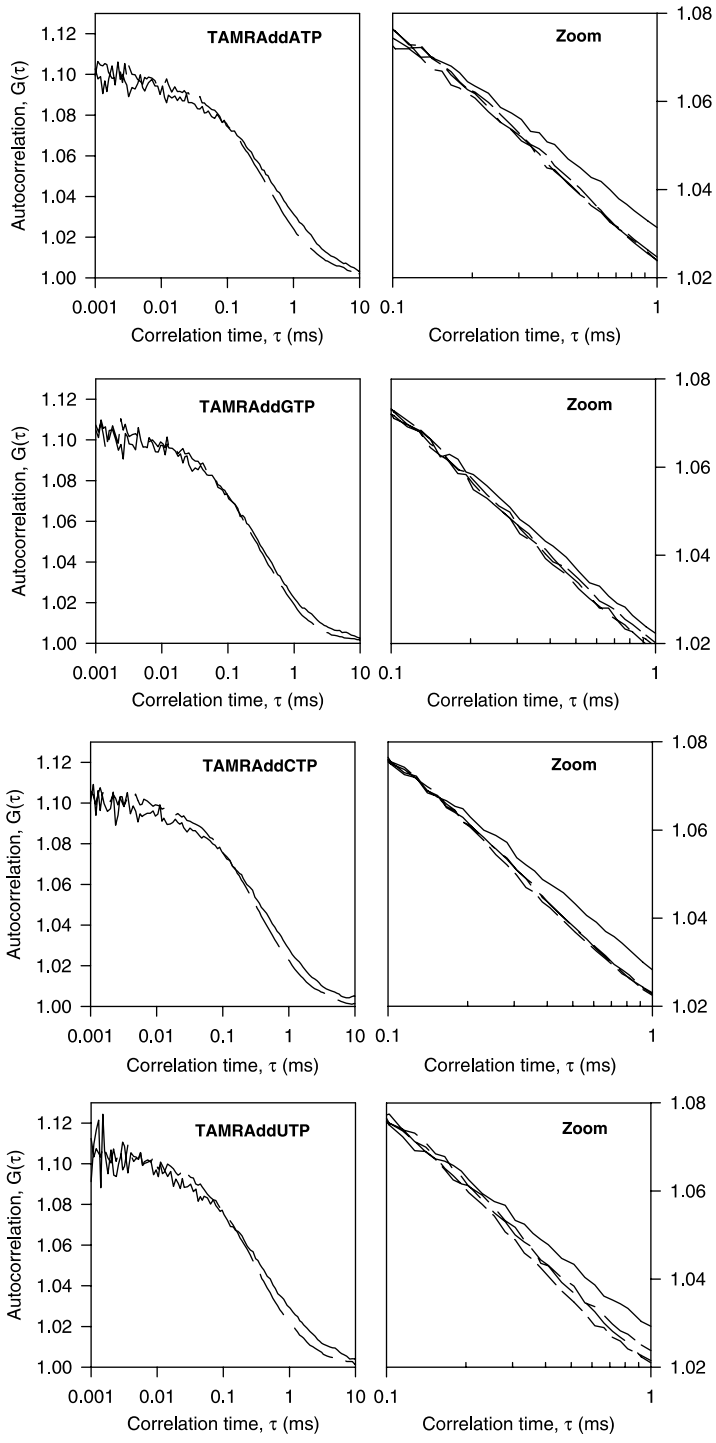

Fig. 2. Autocorrelograms of the four TAMRAddNTPs: TAMRAddATP, TAMRAddGTP, TAMRAddCTP, and TAMRAddUTP. In the left panes are the curves for the TAMRAddNTP where N is complementary to X (——), the other curve being the average autocorrelogram for the three cases where X is not complementary to N (- - -). As the concentration varies from one experiment to the next, a normalization is implemented in the graphs, such that all the autocorrelograms share the same area under the curve. The right-hand panes zoom in between 0.1 and 1 ms—the interval which covers the typical diffusion times. In these panes all four curves are given. In each case, the curve (——) corresponds to the template containing X complementary to N. The other curves (- - -) correspond to the noncomplementary bases in each case.

readily identified and that incorporated is easily discriminated from the other three possibilities. The presence of unincorporated labeled base does not affect these FIDA analyses as strongly as the FCS analyses, since anisotropy measurements are not distorted by the relative emission intensity of each fluorescent component in solution. 2D-FIDA therefore presents the possibility of readily testing each DNA strand for a point mutation.

Table 1
Diffusion time (μs) of the four different bases when reacted with templates containing either one of the four bases inposition X (see Fig. 1)

	TAMRAddATP	TAMRAddGTP	TAMRAddCTP	TAMRAddUTP
X = A	358	267	324	**448**
X = G	360	259	**469**	295
X = C	337	**313**	316	298
X = T	**511**	287	313	295

The TAMRAddNTPs are divided into columns while the four templates are divided into rows. The nature of X is given in the first column. Incorporation should occur only when X and N are complementary (values marked in boldface). The values given were extracted from the curves presented in Fig. 2. Five curves are best-fitted so that each value given is the average of five estimated values. The standard deviation was below 20 μs in all cases.

Table 2
Count rates (photons per fluorescent particle per second, kHz) and calculated polarization (milliA) for the four TAMRA-labeled ddNTPs, and for a TAMRA-labeled double-stranded 30-mer

	q (parallel)	q (perpendicular)	Anisotropy
TAMRA	20.3 ± 0.1	20.7 ± 0.1	23
TAMRAddATP	17.1 ± 0.1	16.0 ± 0.1	52 ± 4
TAMRAddGTP	15.7 ± 0.3	15.0 ± 0.1	45 ± 3
TAMRAddCTP	18.9 ± 0.1	18.7 ± 0.1	33 ± 2
TAMRAddUTP	18.3 ± 0.4	18.2 ± 0.4	31 ± 3
Labelled oligomer	13.0 ± 1.0	9.3 ± 0.7	148 ± 2

The anisotropy of TAMRA is defined as 0.023, which enabled the calculation of the correction factor, G. In this case, $G = 0.92$. The values were averaged over 20 scans taken with four separate optical adjustments. The count rates were corrected for any variation due to the laser power intensity and then averaged. Values given are mean ± SD.

Table 3
Anisotropy (in milliA) of the four different bases when reacted with templates containing either one of the four bases in position X (see Fig. 1)

	TAMRAddATP	TAMRAddGTP	TAMRAddCTP	TAMRAddUTP
X = A	58	55	33	**94**
X = G	56	49	**97**	32
X = C	55	**83**	32	33
X = T	**119**	48	39	33

The TAMRAddNTPs are divided into columns while the four templates are divided into rows. The nature of X is given in the first column. Incorporation should occur only when X and N are complementary (values marked in boldface). As in Table 1, each value is the average of five fitted values; the error margins are of the order of 10^{-4} (0.1 milliA). The error margins are smaller than those in Table 2 as the values are not averaged over several optical adjustments.

Since anisotropy as measured here as an ensemble is an extensive variable, it is possible from the above values (Table 3) to estimate the percentage of incorporation, $\%I = 100(A_{sample} - A_{Free})/(A_{Bound} - A_{Free})$. The extents of incorporation of TAMRAddCTP and of TAMRAddUTP under our conditions were of the same magnitude, being 56 and 54%, respectively. The values for TAMRAddATP and TAMRAddGTP were significantly different, being, respectively, 70 and 37%. Note that these numbers assume that A_{bound} are those measured on the labeled 30-mer; this does not take into account the fact that the 30-mer is labeled slightly differently from the moiety formed upon ddNTP incorporation. The extent of incorporation could only be approximated using the FCS method, since the sum of the count rates obtained from each channel by the FIDA method decreased upon incorporation.

According to the FCS method, the extent of incorporation for TAMRAddATP, TAMRAddGTP, TAMRAddCTP, and TAMRAddUTP were, respectively, 44, 7, 35, and 33%. The estimated rates were thus lower than those given by the more accurate anisotropy approach. This may be explained by the fact that the nonincorporated bases were brighter than their incorporated counterparts and were therefore overrepresented in the overall autocorrelogram.

The lower values for TAMRAddGTP may be related to the known ability of guanosine nucleotides to act as quenchers of fluorescence [83]. It was therefore very likely that the extents of incorporation indicated by the calculations based on the FCS method were indeed below those that had actually taken place. Finally, with regard to TAMRAddGTP, we note that the rate of incorporation of derivatives of G has on occasion been found to be lower than that for other bases [84].

The above FIDA analysis yields an overall value for $q_{\parallel}$ and $q_{\perp}$. The high incorporation rates justify one-component analyses. Rather than use $q_{\parallel}$ and $q_{\perp}$ to extract the anisotropy, a simpler analysis could merely use the total fluorescence intensity detected on each polarization

channel, $I_{\parallel}$ and $I_{\perp}$. The process can be further simplified by considering that:

$$A = \frac{(I_{\parallel}/I_{\perp}) - G}{(I_{\parallel}/I_{\perp}) + 2G}, \tag{3}$$

where $I_{\parallel}$ and $I_{\perp}$ are the fluorescence emission intensity on the parallel and perpendicular channels and G is the correction factor. Furthermore, in the range $G < I_{\parallel}/I_{\perp} < 2G$, the intensity ratio $I_{\parallel}/I_{\perp}$ is almost linearly related to the anisotropy A (regression factor = 0.995). This implied a very much simpler analysis relying uniquely on $I_{\parallel}/I_{\perp}$ to reveal the incorporated base.

The intensity ratio is dependent on the optical alignment of the instrument and will therefore vary from one setup to the next. However, it was found to be very stable in the course of any one session. Each APD receives light quasi-simultaneously, thus canceling out any variations in fluorophore concentration or in laser power. These advantages are conserved in the simplification process from anisotropy calculation to intensity ratio calculation. Furthermore, once the instrument is set up and calibrated, any reference to a standard solution of pure dye becomes irrelevant. In our measurements, the intensities in each channel were averaged over 10 s (although this could be lower if required), thus reducing still further any noise.

For pure TAMRA free in solution, the intensity ratio was 0.98 when $G = 0.92$. For TAMRA label covalently bound to the 3′ end of an oligonucleotide, the fluorescence polarization intensity ratio was 1.21. The ratio of vertical to horizontal component intensity for each DNA–base combination is given in Table 4. When incorporation occurred, the relative intensity of the parallel (vertical) channel increased between 12 and 20%. The values for the unincorporated bases were generally a little higher than those for pure TAMRA, reflecting their greater anisotropy. These data validate the present novel approach, which may now be applied to other DNA fragments.

Finally, we chose to study oligonucleotides representing polymorphisms in genes of medical interest, namely CFTR and BRCA1 (two sequences from BRCA1 are presented). The three oligonucleotides each have two possible SNPs. These were reacted with enzyme and each TAMRA-labeled base in turn, as described above. The results are shown in Fig. 3, left hand and center frames. Whenever the template contained a nucleotide complementary to the added labeled base, the relative ratio (defined as the measured ratio minus the ratio for the free base) rose sharply, thereby accurately determining the presence and nature of the SNP in that gene.

Equimolar mixtures of the two templates (which define the two possible variants within one heterozygous gene) were tested. The concentration of each oligonucleotide was 50 μM. The mixtures were tested for 3′ primer extension in four experiments, by adding each labeled base in turn. Again the relative ratio rose when the complementary base was present in the reaction.

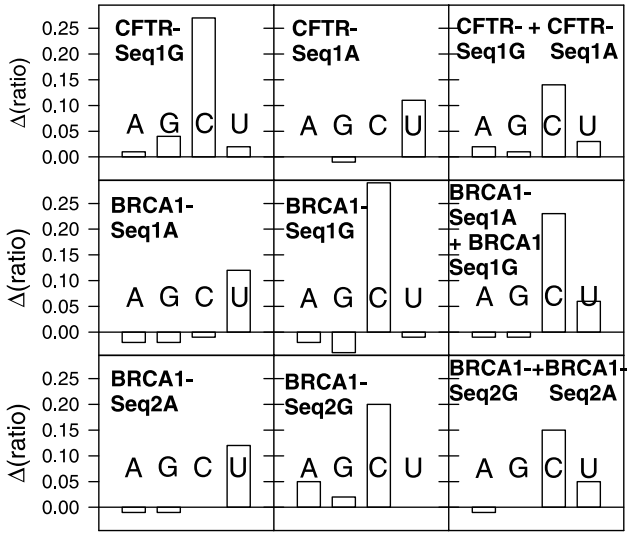

Fig. 3. Analysis of three oligonucleotides, each derived from a gene of medical interest: one from CFTR (CFTR-Seq1) and two from BRCA1 (BRCA1-Seq1 and BRCA1-Seq2). The ratio of intensity on the parallel channel to that on the perpendicular channel is calculated. The relative ratio Δ(ratio) is defined as the observed ratio minus the ratio for free TAMRAddNTP. Δ(ratio) is given for each labeled base added to the reaction, either A, G, C, or U. Each row represents a different gene. There are two possible single-nucleotide polymorphisms (SNPs) present in each gene; therefore two oligonucleotide templates were used. The results for each are shown in the left and center frames, as indicated. A solution containing both SNPs (both templates present at a concentration of 50 μM) was also analyzed; the results from this experiment are presented in the right frames.

Table 4
Ratios of the intensity on the parallel to intensity on the perpendicular channel for the four different bases when reacted with templates containing either one of the four bases in position X (see Fig. 1)

	TAMRAddATP	TAMRAddGTP	TAMRAddCTP	TAMRAddUTP
X = A	1.09	1.07	1.01	**1.20**
X = G	1.08	1.05	**1.22**	1.01
X = C	1.08	**1.17**	1.01	1.01
X = T	**1.29**	1.06	1.03	1.01

The TAMRAddNTPs are divided into columns while the four templates are divided into rows. The nature of X is given in the first column. Incorporation should occur only when X and N are complementary (values marked in boldface). The ratios are the average obtained with a single optical adjustment from five scans lasting 10 s each.

The extent of incorporation is evidently lower when two templates are present, more so for TAMRAddUTP than for TAMRAddCTP. This decrease may be explained by the dependence of the reaction on the enzyme concentration. When two templates are present, they compete for enzyme, such that the incorporation extent for that base is reduced.

Concluding remarks

The continuing need for rapid and inexpensive methods of SNP detection has led to the development of numerous methods in which single-stranded primers are used to discriminate SNPs via differential annealing/hybridization or via the differential incorporation of labeled nucleotides or nucleotide analogues. Others have sought to develop sequencing methods that rely on detecting single molecules (e.g., via exonuclease digestion [72,85–88] or by detecting the nature or existence of PCR products amplified therefrom [67–71]); recently, the incorporation of nucleotides has been observed on single DNA strands by fluorescence microscopy [89]. Future progress depends on the ability to observe primer extension with ever lower nucleotide concentrations.

We have here shown that the incorporation of dideoxy nucleotides is possible with 20 nM nucleotides in the reaction mix. Furthermore, using both FCS and 2D-FIDA (where the two dimensions are used to determine the polarization anisotropy), the presence of the product can be detected without any further preparative steps. In part because the measurements are carried out under quasi-stationary conditions, the simple ratio of intensities in the two channels is in fact sufficient alone to detect which nucleotide has been incorporated. This would in principle allow the method to be performed with a simple microscope, while a multichannel capability would allow several dye labels to be detected simultaneously. By labeling each base with a different dye, the four bases could be probed in turn, by using multiple laser excitation wavelengths and/or using four detectors for different emission wavelengths, such that one would be able to measure the extent of incorporation of all four bases in only one well.

In the present work we used microliter volumes for convenience, but the confocal volume that we interrogated was ~1 fL. It is evident that by reducing the reaction volume to nanoliters or even lower we could discriminate SNPs with attomoles or less of nucleotide reagent. The reaction time was held at 10 s, but we made no attempt to use spectral denoising techniques, such as those based on wavelets [90–92], which could have reduced the measurement time to less than 1 s while preserving the same signal:noise [93]. Especially since the method has been shown to work for SNPs in genes of medical interest, we are confident that the methods of single-molecule detection described in this paper hold out the hope of very rapid and inexpensive SNP analysis.

Finally, we note that the present work has used purified oligonucleotides without a specific cleanup or purification step that might be required were mixed DNA to be analyzed. This will be the subject of future work.

Acknowledgments

We thank the EBS Committee of the UK BBSRC and the UK EPSRC for financial support, Dr. Hywel Griffiths for assistance with molecular biology techniques, and Adrian Eichler, Dr. Leif Brand, and other staff at Evotec Technologies for support.

References

[1] J.J. McCarthy, R. Hilfiker, Nat. Biotechnol. 18 (2000) 505–508.
[2] J. Hoh, A. Wille, R. Zee, S. Cheng, R. Reynolds, K. Lindpaintner, J. Ott, Ann. Hum. Genet. 64 (2000) 413–417.
[3] D.A. Campbell, A. Valdes, N. Spurr, Drug Discov. Today 5 (2000) 388–396.
[4] L. Kruglyak, Nat. Genet. 22 (1999) 139–144.
[5] K. Lindblad-Toh, E. Winchester, M.J. Daly, D.G. Wang, J.N. Hirschhorn, J.P. Laviolette, K. Ardlie, D.E. Reich, E. Robinson, P. Sklar, N. Shah, D. Thomas, J.B. Fan, T. Gingeras, J. Warrington, N. Patil, T.J. Hudson, E.S. Lander, Nat. Genet. 24 (2000) 381–386.
[6] D.E. Reich, M. Cargill, S. Bolk, J. Ireland, P.C. Sabeti, D.J. Richter, T. Lavery, R. Kouyoumjian, S.F. Farhadian, R. Ward, E.S. Lander, Nature 411 (2001) 199–204.
[7] J.C. Stephens, J.A. Schneider, D.A. Tanguay, J. Choi, T. Acharya, S.E. Stanley, R.H. Jiang, C.J. Messer, A. Chew, J.H. Han, J.C. Duan, J.L. Carr, M.S. Lee, B. Koshy, A.M. Kumar, G. Zhang, W.R. Newell, A. Windemuth, C.B. Xu, T.S. Kalbfleisch, S.L. Shaner, K. Arnold, V. Schulz, C.M. Drysdale, K. Nandabalan, R.S. Judson, G. Ruano, G.F. Vovis, Science 293 (2001) 489–493.
[8] A.-C. Syvänen, Nat. Rev. Genet. 2 (2001) 930–942.
[9] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew, D.H. Huson, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L. Gabor Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M.E. Higgins, R.R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, et al., Science 291 (2001) 1304–1351.
[10] D.G. Wang, J.B. Fan, C.J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M.

Mittmann, M.S. Morris, N.P. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T.J. Hudson, R. Lipshutz, M. Chee, E.S. Lander, Science 280 (1998) 1077–1082.
[11] N.J. Schork, D. Fallin, J.S. Lanchbury, Clin. Genet. 58 (2000) 250–264.
[12] H.L. McLeod, W.E. Evans, Annu. Rev. Pharmacol. Toxicol. 41 (2001) 101–121.
[13] A. Vignal, D. Milan, M. SanCristobal, A. Eggen, Genet. Sel. Evol. 34 (2002) 275–305.
[14] P.Y. Kwok, Z.J. Gu, Mol. Med. Today 5 (1999) 538–543.
[15] P.Y. Kwok, Annu. Rev. Genomics Hum. Genet. 2 (2001) 235–258.
[16] J.A. Prince, L. Feuk, W.M. Howell, M. Jobs, T. Emahazion, K. Blennow, A.J. Brookes, Genome Res. 11 (2001) 152–162.
[17] L. Beaudet, J. Bedard, B. Breton, R.J. Mercuri, M.L. Budarf, Genome Res. 11 (2001) 600–608.
[18] M.V. Myakishev, Y. Khripin, S. Hu, D.H. Hamer, Genome Res. 11 (2001) 163–169.
[19] S. Drmanac, D. Kita, I. Labat, B. Hauser, C. Schmidt, J.D. Burczak, R. Drmanac, Nat. Biotechnol. 16 (1998) 54–58.
[20] S. Goto, A. Takahashi, K. Kamisango, K. Matsubara, Anal. Biochem. 307 (2002) 25–32.
[21] A. Ahmadian, B. Gharizadeh, A.C. Gustafsson, F. Sterky, P. Nyren, M. Uhlen, J. Lundeberg, Anal. Biochem. 280 (2000) 103–110.
[22] M. Ronaghi, E. Elahi, Comp. Funct. Genomics 3 (2002) 51–56.
[23] H. Fakhrai-Rad, N. Pourmand, M. Ronaghi, Hum. Mutat. 19 (2002) 479–485.
[24] M.M. Goldrick, Hum. Mutat. 18 (2001) 190–204.
[25] W. Pusch, J.H. Wurmbach, H. Thiele, M. Kostrzewa, Pharmacogenomics 3 (2002) 537–548.
[26] D. Di Giusto, G.C. King, Nucleic Acids Res. 31 (2003) E7-7.
[27] A.-C. Syvänen, Clin. Chim. Acta 226 (1994) 225–236.
[28] X. Chen, L. Levine, P.Y. Kwok, Genome Res. 9 (1999) 492–498.
[29] T.M. Hsu, X. Chen, S. Duan, R.D. Miller, P.Y. Kwok, Biotechniques 31 (2001) 560, 562, 564-8, passim.
[30] T.M. Hsu, S.M. Law, S. Duan, B.P. Neri, P.Y. Kwok, Clin. Chem. 47 (2001) 1373–1377.
[31] P.Y. Kwok, Hum. Mutat. 19 (2002) 315–323.
[32] B.D. Freeman, T.G. Buchman, S. McGrath, A.R. Tabrizi, B.A. Zehnbauer, J. Mol. Diagn. 4 (2002) 209–215.
[33] M.J. Levene, J. Korlach, S.W. Turner, M. Foquet, H.G. Craighead, W.W. Webb, Science 299 (2003) 682–686.
[34] K.U. Mir, E.M. Southern, Annu. Rev. Genomics Hum. Genet. 1 (2000) 329–360.
[35] R.D. Miller, P.Y. Kwok, Hum. Mol. Genet. 10 (2001) 2195–2198.
[36] The International SNP Map Working Group, Nature 409 (2001) 928–933.
[37] Ø. Braaten, O.K. Rodningen, I. Nordal, T.P. Leren, Comput. Methods Programs Biomed. 61 (2000) 1–9.
[38] K.M. Dipple, E.R.B. McCabe, Am. J. Hum. Genet. 66 (2000) 1729–1735.
[39] J. Ott, J. Hoh, Am. J. Hum. Genet. 67 (2000) 289–294.
[40] L. Kruglyak, M.J. Daly, M.P. Reeve-Daly, E.S. Lander, Am. J. Hum. Genet. 58 (1996) 1347–1363.
[41] D.B. Kell, Trends Genet. 18 (2002) 555–559.
[42] R. Judson, J.C. Stephens, A. Windemuth, Pharmacogenomics 1 (2000) 15–26.
[43] J.S. Bader, Pharmacogenomics 2 (2001) 11–24.
[44] D.B. Kell, Trends Biotechnol. 17 (1999) 89–91.
[45] D.B. Kell, Trends Biotechnol. 18 (2000) 186–187.
[46] B. Rotman, Proc. Natl. Acad. Sci. USA 47 (1961) 1981–1991.
[47] P. Ambrose, P. Goodwin, J. Jett, M. Johnson, J. Martin, B. Marrone, J. Schecker, W. Wilderson, R. Keller, Phys. Chem. 97 (1993) 1535–1542.
[48] S.A. Soper, Q.L. Mattingly, P. Vegunta, Anal. Chem. 65 (1993) 740–747.
[49] A. Rich, Proc. Natl. Acad. Sci. USA 95 (1998) 13999–14000.
[50] B.C. Stipe, M.A. Rezaei, W. Ho, Science 280 (1998) 1732–1735.
[51] G.K. Schenter, H.P. Lu, X.S. Xie, J. Phys. Chem. A 103 (1999) 10477–10488.
[52] X.S. Xie, H.P. Lu, J. Biol. Chem. 274 (1999) 15967–15970.
[53] S. Weiss, Science 283 (1999) 1676–1683.
[54] W.E. Moerner, M. Orrit, Science 283 (1999) 1670–1676.
[55] X.N. Chen, Combinatorial Chem. High Throughput Screen. 6 (2003) 213–223.
[56] S. Tabor, C.C. Richardson, Proc. Natl. Acad. Sci. USA 84 (1987) 4767–4771.
[57] E.L. Elson, D. Madge, Biopolymers 13 (1974) 1–27.
[58] D. Madge, E.L. Elson, W.W. Webb, Biopolymers 13 (1974) 29–61.
[59] M. Auer, K.J. Moore, F.J. Meyer-Almes, R. Günther, A.J. Pope, K.A. Stöckli, Drug Discov. Today 3 (1998) 457–465.
[60] J. Widengren, R. Rigler, Cell. Mol. Biol. 44 (1998) 857–879.
[61] H. Schürer, A. Buchynskyy, K. Korn, M. Famulok, P. Welzel, U. Hahn, Biol. Chem. 382 (2001) 479–481.
[62] L. Edman, J. Phys. Chem. A 104 (2000) 6165–6170.
[63] K.J. Moore, S. Turconi, S. Ashman, M. Ruediger, U. Haupts, V. Emerick, A.J. Pope, J. Biomol. Screen. 4 (1999) 335–353.
[64] A.J. Pope, U.M. Haupts, K.J. Moore, Drug Discov. Today 4 (1999) 350–362.
[65] T. Kral, M. Langner, M. Benes, D. Baczynska, M. Ugorski, M. Hof, Biophys. Chem. 95 (2002) 135–144.
[66] H. Xu, J. Frank, U. Trier, S. Hammer, W. Schroder, J. Behlke, M. Schafer-Korting, J.F. Holzwarth, W. Saenger, Biochemistry 40 (2001) 7211–7218.
[67] M. Kinjo, Biotechniques 25 (1998) 706.
[68] S. Bjorling, M. Kinjo, Z. Foldes-Papp, E. Hagman, P. Thyberg, R. Rigler, Biochemistry 37 (1998) 12971–12978.
[69] M. Kinjo, Anal. Chim. Acta 365 (1998) 43–48.
[70] F. Oehlenschläger, P. Schwille, M. Eigen, Proc. Natl. Acad. Sci. USA 93 (1996) 12811–12816.
[71] N.G. Walter, P. Schwille, M. Eigen, Proc. Natl. Acad. Sci. USA 93 (1996) 12805–12810.
[72] J. Stephan, K. Dorre, S. Brakmann, T. Winkler, T. Wetzel, M. Lapczyna, M. Stuke, B. Angerer, W. Ankenbauer, Z. Foldes-Papp, R. Rigler, M. Eigen, J. Biotechnol. 86 (2001) 255–267.
[73] T. Winkler, U. Kettling, A. Koltermann, M. Eigen, Proc. Natl. Acad. Sci. USA 96 (1999) 1375–1378.
[74] Y. Chen, J.D. Muller, P.T.C. So, E. Gratton, Biophys. J. 77 (1999) 553–567.
[75] U. Kettling, A. Koltermann, M. Eigen, Curr. Top. Microbiol. Immunol. 243 (1999) 173–186.
[76] P. Kask, K. Palo, D. Ullmann, K. Gall, Proc. Natl. Acad. Sci. USA 96 (1999) 13756–13761.
[77] P. Kask, K. Palo, N. Fay, L. Brand, U. Mets, D. Ullmann, J. Jungmann, J. Pschorr, K. Gall, Biophys. J. 78 (2000) 1703–1713.
[78] J.D. Müller, Y. Chen, E. Gratton, Biophys. J. 78 (2000) 474–486.
[79] K. Palo, U. Metz, S. Jager, P. Kask, K. Gall, Biophys. J. 79 (2000) 2858–2866.
[80] M. Rüdiger, U. Haupts, K.J. Moore, A.J. Pope, J. Biomol. Screen. 6 (2001) 29–37.
[81] K. Palo, L. Brand, C. Eggeling, S. Jager, P. Kask, K. Gall, Biophys. J. 83 (2002) 605–618.
[82] U. Haupts, M. Rüdiger, S. Ashman, S. Turconi, R. Bingham, C. Wharton, J. Hutchinson, C. Carey, K.J. Moore, A.J. Pope, J. Biomol. Screen. 8 (2003) 19–33.
[83] A.O. Crockett, C.T. Wittwer, Anal. Biochem. 290 (2001) 89–97.
[84] D.H. Densham, in Patent application WO 99/05315 (1999).
[85] M. Sauer, B. Angerer, W. Ankenbauer, Z. Foldes-Papp, F. Gobel, K.T. Han, R. Rigler, A. Schulz, J. Wolfrum, C. Zander, J. Biotechnol. 86 (2001) 181–201.
[86] Z. Foldes-Papp, B. Angerer, P. Thyberg, M. Hinz, S. Wennmalm, W. Ankenbauer, H. Seliger, A. Holmgren, R. Rigler, J. Biotechnol. 86 (2001) 203–224.

[87] P.M. Goodwin, H. Cai, J.H. Jett, S.L. IshaugRiley, N.P. Machara, D.J. Semin, A. VanOrden, R.A. Keller, Nucleosides Nucleotides 16 (1997) 543–550.
[88] L.M. Davis, F.R. Fairfield, C.A. Harger, J.H. Jett, R.A. Keller, J.H. Hahn, L.A. Krakowski, B.L. Marrone, J.C. Martin, H.L. Nutter, R.L. Ratliff, E.B. Shera, D.J. Simpson, S.A. Soper, Genet. Anal. Tech. Applic. 8 (1991) 1–7.
[89] I. Braslavsky, B. Hebert, E. Kartalov, S.R. Quake, Proc. Natl. Acad. Sci. USA 100 (2003) 3960–3964.
[90] D.L. Donoho, IEEE Trans. IT 41 (1995) 613–627.
[91] B.K. Alsberg, A.M. Woodward, D.B. Kell, Chemometr. Intell. Lab. Systems 37 (1997) 215–239.
[92] B.K. Alsberg, A.M. Woodward, M.K. Winson, J. Rowland, D.B. Kell, Analyst 122 (1997) 645–652.
[93] A.D. Shaw, M.K. Winson, A.M. Woodward, A. McGovern, H.M. Davey, N. Kaderbhai, D.I. Broadhurst, R.J. Gilbert, J. Taylor, É.M. Timmins, B.K. Alsberg, J.J. Rowland, R. Goodacre, D.B. Kell, Adv. Biochem. Eng. 66 (1999) 83–113.