J Fluoresc (2006) 16:511–524
DOI 10.1007/s10895-006-0095-7

ORIGINAL PAPER

Information Content of Fluorescence Polarization and Anisotropy

Gabor Mocz

Received: 14 November 2005 / Accepted: 27 March 2006 / Published online: 28 June 2006

Abstract The equality of information content in fluorescence polarization and emission anisotropy is a common assumption and the two quantities are used according to practical considerations. However, an information-theoretic analysis presented here reveals that their information content is substantially different. A scaling relation exists between polarization and anisotropy, and normalization allows their direct comparison. Various measures of information such as the absolute, relative, differential, and potential entropies all appear larger for anisotropy over part or all of its normalized overlap with the polarization function. The larger information content coincides with the signal range where the emitted light is polarized mostly in the parallel direction. Polarization takes on larger absolute entropy only when the emission is about perpendicular to the incident light and when the differential entropy is considered over the entire physical domain. The additional information locally afforded by polarization appears to be related to its larger signal range whereas the extra information in anisotropy may be attributed to a second perpendicular emission plane in its definition, which is oriented along the axis of propagation of light and takes the contribution of all degrees of rotational freedom into account. Thus anisotropy may be considered as a more accurate and more informative representation of the underlying physical phenomena. Some practical aspects relevant to studies of protein–ligand interactions are also discussed.

Keywords Anisotropy · Entropy · Fluorescence · Information · Polarization · Spectroscopy

G. Mocz (✉)
Pacific Biosciences Research Center, University of Hawaii, Honolulu, Hawaii 96822
e-mail: gmocz@hawaii.edu

G. Mocz
University of Hawaii, Biotechnology Program, Gilmore Hall 411, 3050 Maile Way, Honolulu, Hawaii 96822

Introduction

Analytical techniques based on polarization of fluorescence are used in almost every aspect of molecular analysis, especially in chemistry and biology, as well as in clinical, medical and environmental sciences [1, 2]. The directional characteristic or polarization of the emitted light carries temporal and spatial information on the electronic and molecular structure of the fluorophore. When linear polarized light passes through a fluorescent sample it is possible to obtain information on the orientation of the transition dipole and the incident light with respect to the molecular coordinates. Thus, fluorescence polarization measurements can provide analytical information on size, shape and flexibility of molecules, on molecular orientation and mobility, and on processes that modulate the phenomenon [3, 4].

Experimentally, the degree of polarization is determined from measurements of fluorescence intensities parallel and perpendicular to the plane of linearly polarized excitation light [5, 6]. The polarization state of fluorescence is characterized either by the polarization ratio (or degree of polarization) or the emission anisotropy [7, 8]. The two measures are highly interrelated and used according to practical considerations. For example, in the study of radiative transfer the polarization ratio is preferred whereas emission anisotropy is better suited to equilibrium binding studies, though both deliver similar information [9–11].

Nevertheless, there are characteristic differences between the two quantities. The polarization ratio and emission anisotropy are not on the same scale and the latter covers a smaller range of values. The polarization ratio is measured with reference to the fluorescent intensity in the direction of observation while the emission anisotropy is determined in relation to the total emission intensity. The anisotropy value is a simple function of the angular displacement between the excitation and emission dipoles while a more complex relationship exists between the polarization ratio and the angle between the two dipoles. In addition, the time dependence of the polarization ratio is determined by two factors, the fluorescence lifetime and the rotational motion whereas the time dependence of anisotropy is determined only by the rotational motion of the fluorophore [12, 13].

Despite these differences, the general view has been that the information content in the polarization function and the anisotropy function is identical. However, it should not be inferred from their close relatedness that the information content is the same. In fact, on closer examination one will find that it is substantially different.

This paper is concerned with the use of information theory to find the amount of information that can be obtained by performing fluorescence polarization and anisotropy measurements. This can be done by employing such different measures as those introduced by Shannon, Kullback, Danzer, and others [14–22]. Concepts, definitions, and information measures transferred from communication theory [23] have been widely used in a variety of disciplines from chemistry [24] and analytics [25, 26] to ecology [27] and quantum mechanics [28]. Special applications of information analysis such as maximum entropy methods have been employed in obtaining temporal and spatial information on molecular orientation and mobility as well as in a number of other applications, notably fluorescence lifetime and size distribution analysis [29–33]. Here the information content of the components of fluorescence signals that characterizes the measured polarization and anisotropy values in the fundamental definition of the physical phenomena and the information content that characterize the analytical system are analyzed. Although the informational theoretical analysis of the polarization state of fluorescence would appear beneficial to better understanding of these two related quantities, no studies of this type have been reported to date.

Theory and methodology

Basic relations

The polarization of fluorescence at right angles to the incident light is determined by the value of the angle between the direction of the absorption and the emission transition dipoles. Mathematically, the degree of polarization, or polarization ratio, is defined as the fraction of the light which is linearly polarized [5, 6]:

$$P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}} \tag{1}$$

where $I_{\parallel}$ and $I_{\perp}$ are the intensities of the vertically and horizontally polarized components of the emission and the exiting light is polarized in the vertical direction. The denominator represents the fluorescent intensity in the direction of observation.

The theoretical limits of polarization are -1 and $+1$ which would sometimes occur with oriented polymers when the emitted light is totally polarized in the perpendicular or in the parallel direction, respectively. In solutions these values are not attained. For one-photon excitation the real limits of polarization can be derived from its dependence on the angular displacement (δ) between the absorption and emission moments such that $P = (3\cos^2\delta - 1) \times (\cos^2\delta + 3)^{-1}$ [7]. Thus, the maximal value of polarization possible under any circumstances in a macroscopically isotropic solution is $P = 1/2$ for coincident absorption and emission dipoles ($\delta = 0^{[\mathrm{rad}]}$). The other extreme value of polarization occurs with perpendicular absorption and emission dipole ($\delta = \pi/2^{[\mathrm{rad}]}$). In this case $P = -1/3$. When the emission dipole loses all memory of the excitation polarization, $P = 0$ ($\delta = 3^{-1/2[\mathrm{rad}]}$.

Another measure of the polarization of fluorescence is the emission anisotropy, A or r, which is defined as the ratio of the polarized component to the total intensity and is valid for vertically polarized excitation [7, 8]:

$$A = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \tag{2}$$

The theoretical limits of anisotropy are $-1/2$ and $+1$ whereas its real limits can be derived from its dependence on the angle between the two dipole oscillators in the form of $A = (3\cos^2\delta - 1) \times 5^{-1}$ [7]. For one-photon excitation, its maximal value in solution is $A = 2/5$, and its minimal value is $A = -1/5$.

The emission anisotropy and polarization ratio can be interconverted by the relationship $A = 2P \times (3 - P)^{-1}$ or $P = 3A \times (2 + A)^{-1}$ which can be derived from Eqs. (1) and (2). It is important to bear in mind, however, that the two quantities cover two ranges of values that intersect, i.e. they are not on the same scale. There is a scaling relation between them and therefore they cannot be directly compared. Polarization and anisotropy are dimensionless physical quantities and are correlated by a power law given by the generic equation $y = cx^{\lambda}$, where λ is the scaling factor and c is a constant. Specifically, with $\lambda = -1$, $c = (2 \times 3^{-1})$, $y = A$ and $x = (P^{-1} - 3^{-1})$, the scaling

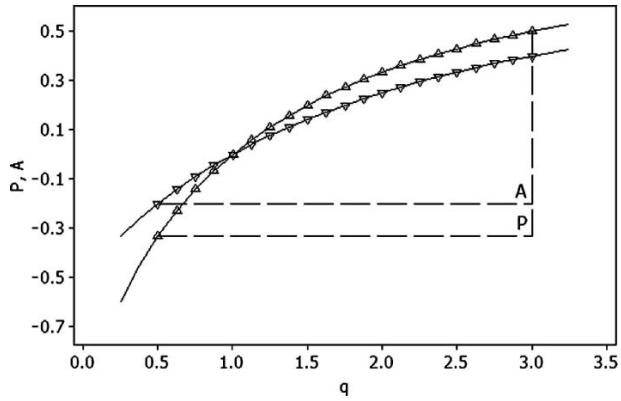

Fig. 1 Plot of fluorescence polarization (P) and anisotropy (A) versus the ratio of the vertically and horizontally polarized components of the emitted light (q), calculated by Eqs. (3) and (4), respectively. Although drawn on the same axes, the two quantities are not on the same scale since $P_{\text{theor}} \in [-1, 1]$ and $A_{\text{theor}} \in [-1/2, 1]$ as follows from the mathematical definition of these functions. Their physically meaningful domains for macroscopically isotropic solutions are indicated by *dashed lines*

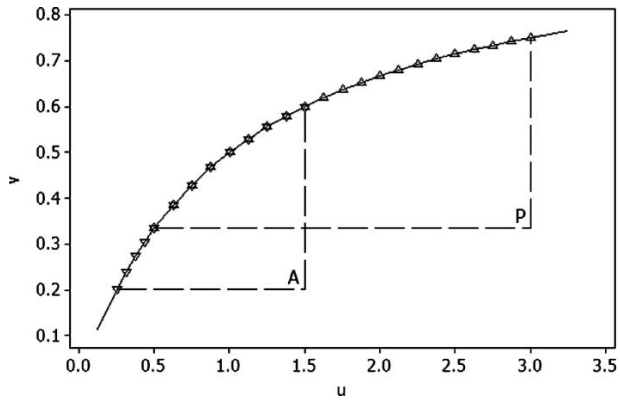

Fig. 2 Plot of fluorescence polarization (P) and anisotropy (A) in normalized form of a rectangular hyperbola. The intensity function v and the recorded quantity u are calculated by Eqs. (5) and (6), respectively. The two functions are on the same scale ($v_{\text{norm}} \in [0, 1]$) permitting their direct comparison. As in Fig. 1, the physically meaningful domains of polarization and anisotropy are indicated by dashed lines

relation is $A = (2 \times 3^{-1}) \times (P^{-1} - 3^{-1})^{-1}$ which is an often used alternative and more instructive formulation for the interconversion of polarization and anisotropy. As it will be shown later, the information content of continuous variables is not scale invariant; therefore, one has to put these functions in a form suitable for calculating their discrete and continuous entropies. Scale normalization allows investigating the scale-independent relations of the two quantities.

Transformation to normalized rectangular hyperbola

Equations (1) and (2) can be rewritten in terms of the ratio of the intensities of the vertically and horizontally polarized components of the emission that will be more useful in this work:

$$P = \frac{q-1}{q+1} \tag{3}$$

and

$$A = \frac{q-1}{q+2} \tag{4}$$

where $q = I_{\parallel}/I_{\perp}$. The minimum and maximum value of q is 1/2 and 3, respectively, both for polarization and anisotropy measurements as it follows from Eqs. (3) and (4).

Figure 1 shows the physically meaningful domains of polarization and anisotropy as a function of q. Although it is not evident, Eqs. (3) and (4) are the equations of rectangular hyperbolae. One can rewrite these equations in standard normalized form for consistent scaling. Rearrangement of Eqs. (3) and (4) gives

$$v = \frac{1+P}{2} = \frac{q}{1+q} = \frac{u}{1+u}, \quad (u \equiv q) \tag{5}$$

and

$$v = \frac{1+2A}{3} = \frac{q/2}{1+q/2} = \frac{u}{1+u}, \quad (u \equiv q/2) \tag{6}$$

The normalized variables u and v are dimensionless and can be plotted now on the same scale for direct comparison. The intensity function v is constrained within the limits 0 and 1, and the accessible theoretical range for signal u is $(0, \infty)$. Figure 2 shows the reduced, normalized rectangular hyperbola with the real domains of polarization in macroscopically isotropic solution which will be used to determine the information content of these two quantities. The transformed functions both lie in the first quadrant, unlike their original definitions, and the origin in the reduced frame coincides with the vertex of Eqs. (5) and (6).

Absolute information (Shannon entropy)

First, the information content of fluorescence signals in the context of experimental situation is explored. Suppose signal X is a (discrete or quantized) random variable which takes on a finite set of values in any one of a number of states according to a probability distribution $\Pr(X)$. Then, an appropriate measure of the information content of the system can be expressed in terms of Shannon's absolute (discrete) entropy [14, 15] as

$$H[X] = -\sum_{i=1}^{r} \mu(x_i) \log_2 \mu(x_i) \tag{7}$$

where X takes the value x_i, r is the number of levels of the variables, and $\mu(x_i)$ is the probability of the elementary event subject to the constraint $\sum_{i=1}^{r} \mu(x_i) = 1$. In fact, $H[X]$ represents the *average* information amount of a set of possible outcomes (mean entropy per elementary event). Sometimes it is convenient to use a quantity called *total* information which is directly related to Shannon's entropy but characterizes the information content of the entire system and is defined by $I[X] = N \times H[X]$ for a group of N individual components. [16, 17].

Since the fluorescence emission signal is normally characterized only by one state possessing two recorded attributes with probabilities $\mu(x)$ and $1 - \mu(x)$ for $I_{\parallel}$ and $I_{\perp}$, respectively, Eq. (7) reduces to

$$H[X] = -\mu(x)\log_2 \mu(x) - (1-\mu(x))\log_2(1-\mu(x)) \quad (8)$$

If there are m-out-of-n signal components possessing one of the two attributes, the best available estimate of $\mu(x)$ is m/n. The information content that corresponds to this signal composition may be written as

$$H[N] = -\frac{m}{n}\log_2\frac{m}{n} - \frac{n-m}{n}\log_2\frac{n-m}{n} \quad (9)$$

From Eqs. (5) and (6) one finds that the normalized signal components are represented by u and 1, respectively. Making the necessary substitutions one has $m = u$ and $n = 1+u$. Equation (9) may now be written as

$$H[U] = -\frac{u}{1+u}\log_2\frac{u}{1+u} - \frac{1}{1+u}\log_2\frac{1}{1+u} \quad (10)$$

Continuous information (differential entropy)

The absolute entropy may be generalized to the continuous case [34–36], i.e. when a signal is recorded by scanning the entire signal region (or part of it). Let X be a continuous random variable with a cumulative distribution $F(x) \equiv \Pr(X \le x)$. Then, the probability density function is given by $\rho(x) \equiv F'(x) \equiv \Pr(X)$, provided that $F(x)$ is continuous and $\rho(x)$ is normalized. Thus, if $\rho(x) > 0$ for $x \in [x1, x2]$ and $\rho(x) = 0$ outside the interval, the information content (or differential entropy) of the continuous variable X is denoted by

$$H[X] = -\int_{x1}^{x2} \rho(x)\log_2 \rho(x)\mathrm{d}x \quad (11)$$

with $\rho(x)$ being normalized so that $\int_{x1}^{x2} \rho(x)\mathrm{d}x = 1$. While Eq. (11) is mathematically correct and often used in scientific literature, it is seen to be dimensionally incorrect in physical applications [37, 38]. In Eq. (11), the argument of the logarithm has a dimension of a continuous probability density, i.e. length^{-1}, because here x represents a distance on the abscissa from a fixed reference point. A possible correction makes use of an invariant reference density $\rho_0(x)$ which represents the maximum information obtainable within the duration of the experiment [38]:

$$\begin{aligned} H[X] &= -\int_{x1}^{x2} \rho(x)\log_2\frac{\rho(x)}{\rho_0(x)}\mathrm{d}x \\ &= -\int_{x0-ct}^{x0+ct} \rho(x)\log_2(2ct\rho(x))\mathrm{d}x \end{aligned} \quad (12)$$

where x_0 is the location of the scanning device and $2ct$ denotes the length of interval that can be scanned in finite time t (c is the speed of light). In this paper x_0 is placed in the center of the physically meaningful interval of the recorded quantities. With this choice Eq. (12) reduces to

$$H[X] = -\int_{x1}^{x2} \rho(x)\log_2\left(\frac{x_2 - x_1}{2}\rho(x)\right)\mathrm{d}x \quad (13)$$

One may then extend Eq. (10) to the continuous domain and define an appropriate signal function that takes the form $\rho(u) = b^{-1} \times [u \times (1+u)^{-1} - a]$. The dimensionally correct information amount for the polarization ratio and emission anisotropy is written such that

$$\begin{aligned} H[U] = &-\int_{u1}^{u2}\left(\frac{1}{b} \times \frac{u}{1+u} - \frac{a}{b}\right) \\ &\times \log_2\left(\frac{u_2 - u_1}{2b} \times \frac{u}{1+u} - \frac{u_2 - u_1}{2} \times \frac{a}{b}\right)\mathrm{d}u \end{aligned} \quad (14)$$

where the integration is understood to be over the support set of U with respect to the physical limits of the two quantities. Here the signal function may be normalized either to $\lim_{u\to+\infty} \rho(u) = 1$, the theoretical upper limit of its standard reduced form, or to the union of the physically meaningful domains of the polarization and anisotropy function, whichever is preferred for a particular application. The normalization constants are $a = \min \rho(u) = \min u/(1 + \min u)$ (i.e. $a_P = 1/3$ and $a_A = 1/5$) and $b = \max H_c[U] = \int_0^{\max u} 1\mathrm{d}u = 3$ (theoretical limit), or $b = C(U_A) + C(U_P) - C(U_{A\cap P}) \cong 5/6$ (physical limit), respectively, where $C[U] = \int_{u1}^{u2} [u/(1+u) - a]\mathrm{d}u$

Relative information (Kullback–Leibler entropy)

A convenient way to measure the divergence or distance between two probability distributions $\Pr(X)$ and $\Pr_0(X)$ has been introduced by Kullback [18, 19] as follows:

$$D[X||X_0] = \sum_{i=1}^{r} \mu(x_i) \log_2 \frac{\mu(x_i)}{\mu_0(x_i)} \tag{15}$$

This measure is also known as the information gain or relative entropy [39], Eq. (15) is asymmetrical in its arguments, i.e. the distance $D[X||X_0]$ is not equal to $D[X_0||X]$. If the distributions are not too dissimilar, the difference between $D[X||X_0]$ and $D[X_0||X]$ is small, approaching 0. For distributions of a continuous variable this measure can be written as

$$D_c[X||X_0] = \int_{x1}^{x2} \rho(x_i) \log_2 \frac{\rho(x_i)}{\rho_0(x_i)} \mathrm{d}x \tag{16}$$

which is the gain of information obtained if the reference function $\rho_o(x)$ is replaced by $\rho(x)$. Then the relative information content of polarization with respect to anisotropy can be derived from Eqs. (10) and (15) as

$$\begin{aligned} D[U] = & -\frac{u_P}{1+u_P} \log_2 \left(\frac{1+u_A}{1+u_P} \times \frac{u_P}{u_A} \right) \\ & -\frac{1}{1+u_P} \log_2 \frac{1+u_A}{1+u_P} \end{aligned} \tag{17}$$

Replacing the probability densities in Eq. (16) with the appropriate variables from Eq. (14), one gets

$$\begin{aligned} D[U] = & -\int_{u1}^{u2} \left(\frac{1}{b} \times \frac{u_P}{1+u_P} - \frac{a_P}{b} \right) \log_2 \left(\frac{u_{2P}-u_{1P}}{u_{2A}-u_{1A}} \right. \\ & \left. \times \frac{1+u_A}{1+u_P} \times \frac{u_P - a_P u_P - a_P}{u_A - a_A u_A - a_A} \right) \mathrm{d}u \end{aligned} \tag{18}$$

Reversing the roles of polarization and anisotropy in Eqs. (17) and (18), analog expressions can be easily derived which would yield a different set of values depending on the similarity of the underlying distribution. Eqs. (17) and (18) lead to the information theoretic estimator of gain for polarization versus anisotropy, the alternative leads to the estimator of the same quantity where the gain is then for anisotropy versus polarization.

Maximum potential information (Eckschlager–Danzer entropy)

Next, the information content in the context of analytical information is explored. Generally, signal functions are obtained by instrumental methods of analysis as two dimensional information, $y = f(z)$. For continuous probability distributions, the maximum potential information amount (Eckschlager–Danzer information) of a sequence of signals that one can obtain by an analytical method is given in [26] as

$$M_{ZY} = N_Z \log_2 N_Y \tag{19}$$

with N_Z and N_Y being the signal's resolution power and its intensity resolution, respectively. Information expressed in this way is a specific case of the Brillouin measure of information [16, 17] which is understood in such a way that a set characterized by M_{ZY} can be decomposed into N_Z subsets each of N_Y elements.

The resolution power, i.e. the numbers of distinguishable signal positions within the interval of the measurement, is given by

$$N_Z = \int_{z1}^{z2} \frac{\mathrm{d}z}{\Delta z} \tag{20}$$

where z_1 and z_2 are the lower and upper limits, respectively, within which the signal is recorded. Δz represents the signal resolution which should be approximately double the signal half-width $\Delta z_{1/2}$, i.e. it is the smallest difference that one can analyze. In general, Δz is a function of the recorded quantity z in the form $\Delta z^{-1} = f(z)$.

In Eq. (19), the intensity resolution is evaluated as

$$N_Y = \frac{(y_2 - y_1) n_p^{1/2}}{2 s_y t(\alpha, f)} \tag{21}$$

where y_1 and y_2 are the lower and upper limits within which the signal is anticipated, s_y is the standard deviation for the signal measurements that characterizes the average precision of determining the intensity of a signal, and n_p is the number of parallel analyses. $t(\alpha, f) = n_\mathrm{p}^{1/2}$ denotes the critical value for a quantile of Student distribution. For a uniform *a priori* distribution and a normal *a posteriori* distribution the denominator of N_Y can be expressed as $2 s_y t(\alpha, f) = 2\Delta y \times n_\mathrm{p}^{1/2}$ where Δy is the confidence interval. Then, for $y_2 \to \bar{y} = (y_2 - y_1)/2$ and $y_1 \to 0$, one has:

$$N_Y = \frac{\bar{y}}{2\Delta y} \approx \frac{S}{N} \tag{22}$$

which corresponds to the signal-to noise-ratio (*S/N*) by which the information content is evaluated in information science. The signal to-noise ratio can also be written as $S/N = \bar{y}/s_y$ which is the criterion used to decide the precision of analytical instruments and procedures [26].

One can approximate the signal resolution by a function of the recorded quantity, $z \to u$ that corresponds to constant intensity resolution Δv. Following Eqs. (5) and (6), the signal resolution is obtained as $\Delta z \equiv \Delta u = \Delta v(1+u)^2 \times (1 - \Delta v - u\Delta v)^{-1}$, where $\Delta v \equiv \Delta p/2 \equiv 2\Delta A/3$, and ΔP and ΔA are the precision of polarization and anisotropy analysis, respectively, as usually practiced. One may write the approximate noise intensity y_N as the difference between the signal intensities at double half-width signal resolution with statistical risk of error $\alpha = 0.01$, i.e. $y_N \approx \Delta v/5$. The above considerations lead to a specific definition of the maximal analytical information amount for polarization and anisotropy as

$$M_{UV} = \left(\int_{u1}^{u2} \frac{1 - \Delta v(1+u)}{\Delta v(1+u)^2} \mathrm{d}u \right) \log_2 \left(\frac{v_2 - v_{1\backslash}}{2\Delta v} \right) \tag{23}$$

which connects the information content with the recorded quantity u and its intensity function v (at variable signal and constant intensity resolution).

Linking information measures to binding parameters

To link the information content of polarization and anisotropy functions to parameters that can be obtained from the measurement of the polarized fluorescence intensities, a simulated study of ligand–protein binding is considered. Using either polarization or anisotropy data, one can calculate the fraction of ligand bound (x) at any protein concentration. The dissociation constant (K_d) is related to x by the expression [11]:

$$K_d = \frac{(1-x)(nP_T - xL_T)}{x} \tag{24}$$

where p_T and L_T are the total concentration of protein and ligand, respectively, and n is the number of identical binding sites without interaction (most commonly $n = 1$). In terms of anisotropy, the expression relating the observed or simulated anisotropy to x is given by [11]

$$x = \frac{A - A_f}{A_b - A_f + (g-1)(A_b - A)} \tag{25}$$

with g representing the quantum yield enhancement factor and A_b and A_f denoting the value of bound and free anisotropy, respectively. An analogous expression exists for polarization measurements [11]:

$$x = \frac{(3 - P_b)(P - P_f)}{(3-P)(P_b - P_{f)} + (g-1)(3 - P_f)(P_b - A)} \tag{26}$$

where P_f is the polarization of the ligand free in solution, P_b is the polarization of the bound ligand, and P is the observed or simulated polarization. Substituting P and A into Eqs. (5) and (6) yields u, which in turn can be used in any of the information measures given above. Thus u effectively links x to the information content in determination of K_d. The joint dependence of entropy on these two quantities can be conveniently presented using contour plots.

Correlation of information content with time-resolved fluorescence

A critical consideration that impacts on studies of protein–ligand interactions is the fluorescent lifetime of the fluorophore. The observed polarization will depend not only on the rotational rates of the system but also on the excited state lifetime. For a simple system with a single lifetime and a single rotational correlation time, the polarization can be written in a rearranged form of the well-known Perrin equation [39]:

$$\frac{1}{P} - \frac{1}{3} = \left(\frac{1}{P_0} - \frac{1}{3} \right) \times \left(1 + \frac{\tau}{\phi} \right) \tag{27}$$

In this equation P is the observed polarization, P_0 is the limiting polarization in the absence of rotation, τ is the excited state lifetime, and ϕ is the rotational correlation time which is related to the rotational relaxation time (ρ) by the expression $\rho = 3\phi$. For a more complex system, the polarization may be determined by integration of the usually intricate total polarization decay function, which is beyond the scope of this paper. Perrin's formula is often stated in terms of anisotropy as

$$A = A_0 \times \left(1 + \frac{\tau}{\phi} \right)^{-1} \tag{28}$$

where A_0 is the limiting anisotropy. From Eqs. (27) and (28), P and A are obtained to correlate τ and ϕ to the amount of information via Eqs. (5) and (6) and the equations of appropriate information functions.

It remains to mention that all integrals used in this work have been verified by numerical integration using a robust, extended trapezoidal rule [40].

Results and discussion

A quantity computed according to Eq. (7) and all expressions that may be derived from it can be interpreted either as a measure of entropy or as a measure of information [41]. The first interpretation is appropriate when one deals with a system before an experiment is carried out on it. Thus H measures the uncertainty concerning the results of the experiment. On the other hand, when one deals with a system after the experiment, H measures the amount of information obtained in the experiment. The case of fluorescence polarization and emission anisotropy easily falls into the latter category. The experiment reducing the uncertainty of the system is the process of excitation while all the information is contained in the polarized intensities in relation to the absorption transition moment, leaving the emission transition moment invariant. The emission transition moment remains the same whatever the excited state reached by the fluorophore upon excitation because of internal conversion towards the first singlet state.

Shannon's measure for discrete signals

When applying Eq. (10) to polarization and anisotropy signals it is useful to compare the two phenomena both at identical values of the normalized variable u, and at identical values of the recorded quantity q, i.e. when q, the ratio of the intensities of the vertically and horizontally polarized components is the same in both functions. Recall from Eqs. (5) and (6) that u can be written as $u_P \equiv q$ and $u_A \equiv q/2$. For instance, if $q_P \equiv q_A = 3$ then $u_P = 3$ and $u_A = 3/2$, both representing identical state of the emitted light. Figure 3 shows a plot of the information content of polarization and anisotropy signals as a function of u for various signal compositions. In the overlapping interval of the two quantities, the information content appears to be the same. This result is not surprising since the two quantities are given by the same function and the information content is unchanged by an isomorphic transformation of a variable. However, the relationship between the information amount obtained from polarization and anisotropy can be more conveniently displayed when the same data are remapped as function of q (Fig.4). Since the normalized quantity u for anisotropy is half that observable for polarization under the same conditions (identical value of q), it easily follows from Eq. (10) that the information content of the anisotropy signal is larger than that of the polarization signal when $q > 2^{1/2}$ or $A > 0.121$, which is the physically more interesting recording range. Conversely, the information content of the polarization signal is larger when $P < 0.172$. The maximum information content that a polarization or anisotropy experiment can produce is 1 *bit* per signal component which occurs when $q = 1\,(P = 0)$ for polarization, and $q = 2\,(A = 1/4)$ for anisotropy. The

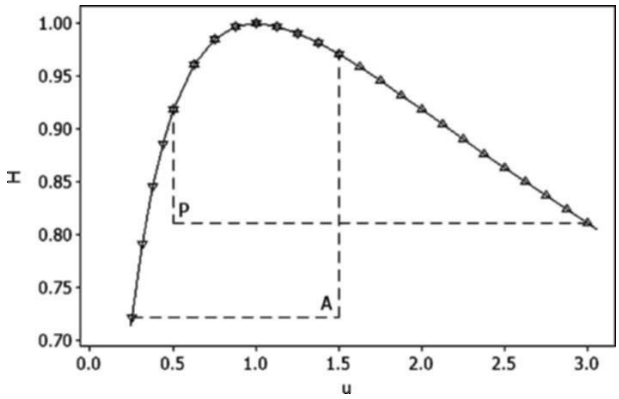

Fig. 3 Absolute (discrete) information content (H) of fluorescence polarization (P) and anisotropy (A) as a function of u. The lines are given by Eq. (10). The overlapping interval $u \in [1/2, 3/2]$ represents the isoinform domain

corresponding angular displacement between the absorption and emission dipoles is $\delta = 3^{-1/2[\mathrm{rad}]} \approx 54.7^\circ$ and $\delta = 3^{1/2} \times 2^{-1[\mathrm{rad}]} = 30^\circ$, respectively.

As an example, determine the information content for the maximum signal intensity in a polarization or anisotropy experiment. Substituting $u = 3$ $(q = 3)$ for polarization and $u = 3/2$ $(q = 3)$ for anisotropy into Eq. (10), one gets (in *bits* per normalized intensity component): $H[U_P] = -3 \times (1+3)^{-1} \times \log_2(3 \times (1+3)^{-1}) - (1+3)^{-1} \times \log_2((1+3)^{-1}) \cong 0.81$ and $H[U_A] = -3 \times (2+3)^{-1} \times \log_2(3 \times (2+3)^{-1}) - 2 \times (2+3)^{-1} \times \log_2(2 \times (2+3)^{-1}) \cong 0.97$, respectively.

Quite usefully, one may find the same result from simple physical considerations by counting the number photons emitted in the parallel and perpendicular directions, or equivalently, averaging over the appropriate intensity distribution functions of the fluorescence light. Using standard formulas for integration, relative values are easily obtained for the two observed components like $I_\parallel / I_\perp = 3/1$ at their respective maximum [41, 42]. This implies that the

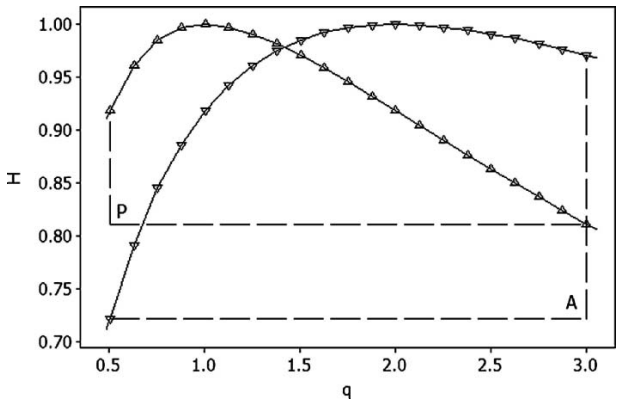

Fig. 4 The same information amount as in Fig. 3, remapped against q. The polarization and anisotropy functions coincide at one point, $u = 2^{1/2}$, above which anisotropy appears more informative

(normalized) total intensities are $n \equiv I_{\parallel} + I_{\perp} = 3 + 1 = 4$ and $n \equiv I_{\parallel} + 2I_{\perp} = 3 + 2 \times 1 = 5$ for polarization and anisotropy, respectively, whereas the signal component possessing the parallel attribute is the same in both cases: $m \equiv I_{\parallel} = 3$. This can be used to calculate the information content. From Eq. (9), one obtains the result in *bits* per intensity component that $H[N_A] = -3 \times 5^{-1} \times \log_2(3 \times 5^{-1}) - (5-3) \times 5^{-1} \times \log_2((5-3) \times 5^{-1}) \cong 0.97$ and $H[N_P] = -3 \times 4^{-1} \times \log_2(3 \times 4^{-1}) - (4-3) \times 4^{-1} \times \log_2((4-3) \times 4^{-1}) \cong 0.81$ which is the same outcome as that obtained from the more general function defined by Eq. (10). Then the information content of the population of n signals can be written as $I[N_P] = 4 \times H[N_P] \cong 3.25$ *bits* and $I[N_A] = 5 \times H[N_A] \cong 4.85$ *bits* for polarization and anisotropy, respectively, i.e. the anisotropy signal yields about 1.6 *bits* more information at the maximum value of the two signal components when $u = 3$. In a similar way as above, the information content can also be calculated at the minimum value of the intensity function, i.e. when $q = 1/2$. Readily, one has for the mean information entropy (in *bits* per intensity component) $H[N_P] \cong 0.92$ and $H[N_A] \cong 0.72$ which implies that the total information amount in *bits* is $I[N_P] \cong 3.67$ and $I[N_A] \cong 3.61$, respectively. Some practical aspects of the unequal information content of the two functions will be discussed later.

Scanning the whole signal range

Now turn attention to the continuous situation where the signal is recorded by scanning the entire signal region. The differential entropy acts as a measure of information of accurate (unbiased) direct measurements which is noise free and exact, i.e. one measures the amount of information as a decrease of uncertainty in the experiment and not as the absolute information of the signals. In general, the differential entropy cannot be obtained as a limiting case of the absolute entropy and H_c behaves differently in many ways than its finite counterpart. From the viewpoint of this work, the most relevant difference is in the scaling behavior of the two entropies. Rescaling a variable does change the differential entropy such that $H_c[\gamma \times X] = H_c[X] + \log|\gamma|$, whereas the absolute entropy remains invariant. Therefore, two variables need to be on the same scale, or be simultaneously rescaled by the same factor for correctly measuring their information contents in relation to each other. This characteristic requires the use of the standard normalized forms of the polarization and anisotropy functions instead of their original definitions.

First, calculate the information content at a finite scale, for instance at the level of signal resolution $\Delta z \equiv \Delta u \in [0.004, 0.049]$ (details of calculating Δz will be given later under Analytical Information). From Eq. (14), one gets (in *bits* per component) $H_P \cong 0.024$ and $H_A \cong 0.035$,

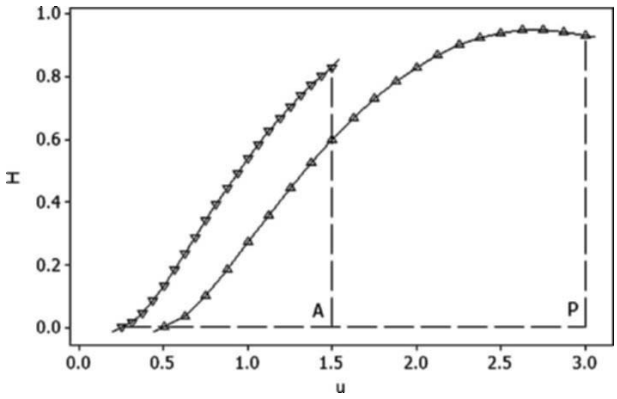

Fig. 5 Differential (continuous) information content (H) of fluorescence polarization (P) and anisotropy (A) as a function of u. The integration is carried out over the interval $u_P \in [1/2, u)$ and $u_A \in [1/4, u)$, respectively, according to Eq. (14)

respectively, when the integration is carried out from $u = 1$ to $1 + \Delta u$. At this fine scale the information content is small, but larger for anisotropy than for polarization. For larger values of u the behavior of the information amount is similar, but the actual difference between the two quantities is larger. Figure 5 shows the continuous information content of the polarization and anisotropy functions *versus* u by integrating over an extended range from $u_P = 1/2$ (or $u_A = 1/4$) to u as indicated on the abscissa. The amount of information appears on average 30% larger for anisotropy over the whole section of the normalized interval $u \in [1/2, 3/2]$ that the two functions share. This is a remarkable result because the two functions are on identical scale, and Eq. (15) makes an allowance for the physically meaningful domains of polarization and anisotropy. However, performing the integration over the respective entire domains of the two quantities, the information content of the polarization function appears larger. This may be explained by realizing that the continuous entropy is the extent of inequality between two items of information. Accordingly, the apparent extra information is due to the more extended domain of the normalized polarization function. Figure 5 also shows that the maximum continuous information content is attained at $u_P \approx 8/3$ ($q \approx 8/3$) and $u_A = 3/2$ ($q = 3$) which yields (in *bits* per component) $H_P \cong 0.95$ and $H_A \cong 0.83$, respectively if the scanning device is placed at the center of the interval. For comparison with the finite case, recall that the maximum absolute information content is 1 *bit* both for polarization and anisotropy which is obtained when $u = 1$.

It is worth noting that these considerations apply only to continuous, linearly polarized fluorescent light. For nonlinear optical waves, the classic definition of the degree of polarization is not relevant [43, 44] and one would need to proceed more carefully to calculate the information content from all the statistical moments in such medium which is beyond the scope of this paper.

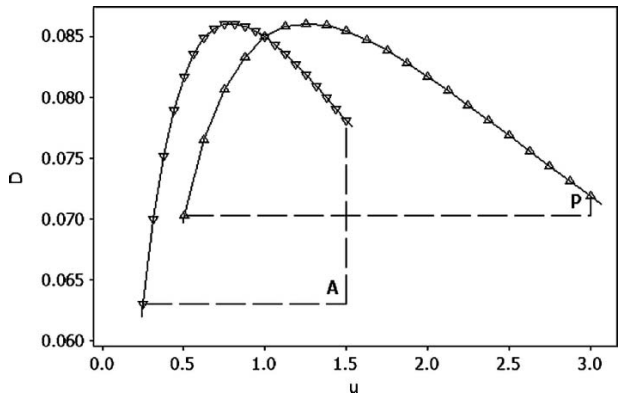

Fig. 6 Plots of both relative entropies $D[P\|A]$ and $D[A\|P]$ versus u for the discrete case. The two quantities are calculated by Eq. (17) and its reverse (in argument), respectively

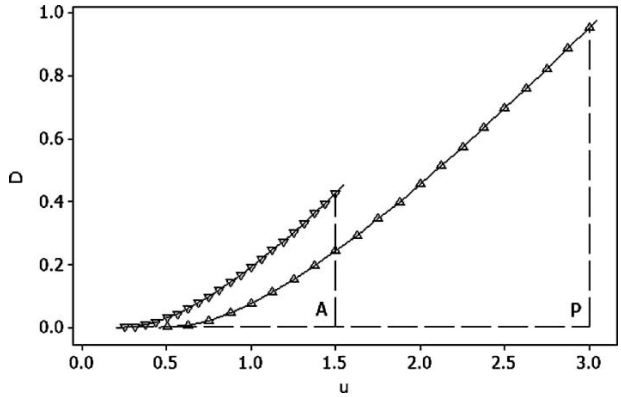

Fig. 7 Plots of both relative entropies $D[P\|A]$ and $D[A\|P]$ versus u for the continuous case. The two quantities are calculated by Eq. (18) and its reverse (in argument), respectively

Relative information gain

While the absolute entropy is used to assess the information content of analytical signals, the relative entropy is used to assess the effective information or information gain of analytical results. The Kullback–Leibler information measure allows quantitative treatment of the connection between entities or characteristics of the same type, possibly with respect to some prior knowledge or some unbiased distribution. The present work relates the normalized polarization and anisotropy functions with reference to each other in order to characterize the relative divergence in their signal functions. Figure 6 shows a plot of relative discrete information for polarization *versus* anisotropy and anisotropy *versus* polarization, respectively. As one can see in this figure the qualitative behavior of the two functions is similar but they only coincide at one point, $u = 1$ where both distributions maximally match. Generally, the difference between the two distributions is small and thus they mutually correlate to certain extent, both containing information on the other. The information gain appears maximum when $u_P = 2^{1/2}$ and $u_A = 2^{-1/2}$, i.e. when $q_P \equiv q_A = 2^{1/2}$ which corresponds to an angular displacement of $\delta \approx 43^\circ$.

For the continuous case, the relative entropy is plotted in Fig. 7. Here, the two distributions appear similar again; however their difference is larger than in the finite case displaying larger independence from each other. Over their normalized overlap, the information gain emerges greater for anisotropy with respect to polarization, while the information gain for polarization is naturally larger outside this range. Unlike with the discrete function, the information gain for polarization increases with increasing u. Maximum gain relatively to the chosen coordinate system is obtained when $q_P \equiv q_A = 3$.

Analytical information content

Now, consider the case of analytical information content, i.e. the maximum potential information amount or Eckschlager–Danzer measure of Brillouin information, for polarization and anisotropy measurements. This is conceptually different from the information measures discussed above as it derives the information content from the *difference* between two resolvable signal components. Therefore, it is characteristic of the analytical system (how signal is measured) rather than of the signal itself.

First note that very precise measurements ($P \pm 0.003$ or $A \pm 0.002$) are readily obtainable with modern instrumentation. Thus, from the relation $\Delta z \equiv \Delta u = \Delta v(1+u)^2 \times (1 - \Delta v - u\Delta v)^{-1}$ (see Theory and Methodology) one finds that typical values of the signal resolution fall in the range $\Delta z \equiv \Delta u_P \in [0.0042, 0.0168]$ and $\Delta z \equiv \Delta u_A \in [0.0068, 0.0486]$, respectively. A corresponding calculation for the noise intensity yields $\Delta v_P \equiv \Delta P/2 = 0.003$ and $\Delta v_A \equiv 2\Delta A/3 = 0.00267$, respectively. On substituting these values into Eq. (23), first one can calculate the maximum potential information amount for polarization and anisotropy measurements over the narrow interval of signal resolution. In both cases, the information amount obtained is about 1 *bit*; more exactly $M_P = 0.99399$ and $M_A = 0.99465$, respectively. At first sight this is a small difference in favor of anisotropy. However, the analytical information content takes on considerably high values when the entire signal range is considered, resulting in an amplification of the difference. Figure 8 shows the dependence of the maximum potential information on u. If the information content of the two quantities (P and A) is compared over the whole physically meaningful domain, one finds (in *bits*) that $M_p = 844$ and $M_A = 929$, respectively. Constraining the calculations to the same normalized interval

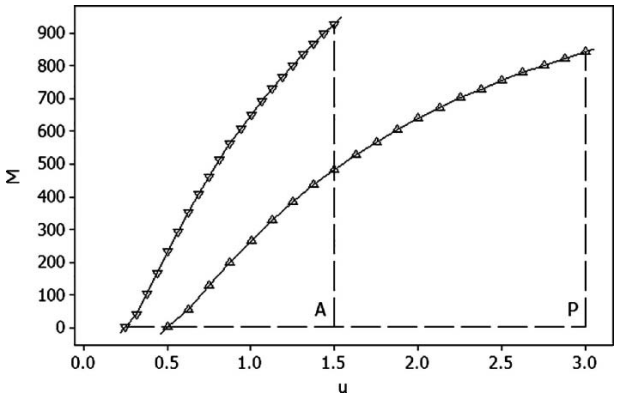

Fig. 8 Maximum potential information amount (M) of fluorescence polarization (P) and anisotropy (A) plotted against u. The intervals $u_P \in [1/2, u)$ and $u_A \in [1/4, u)$ are used to calculate the analytical information with fixed intensity resolution according to Eq. (23)

$u \in [1/2, 3/2]$ for the two quantities, greater analytical information is obtained for anisotropy, specifically $M_P = 484$ and $M_A = 561$, respectively. Unlike the case with differential entropy, the larger interval for polarization does not contribute to the increase of potential information.

The calculated potential information appears to be on the same order that other spectrometric methods furnish. A compilation of such values for various methods is listed in [26]. Just to cite two, UV-VIS spectrometry yields 500 *bits* of information and optical emission spectrometry provides 600 *bits* (by quantometers).

Practical considerations

Maximum absolute information is attained when the components of the (normalized) polarized fluorescence intensities are the same both in parallel and perpendicular directions. The distribution of entropy data values around the maximum is not symmetrical though. Values on one side of the distribution tend to be further from the maximum than values on the other side. Figure 4 shows that it may be more advantageous (from the information theoretic point of view) to use the anisotropy function in the higher end of the positive signal range where the emitted light is polarized mostly in the parallel direction. In this respect, polarization is better suited to be used in the negative signal range (negative initial polarization) or in the low positive range where the emitted light is polarized to a larger extent in the perpendicular direction.

As a practical example, consider how various information measures reflect into the determination dissociation constants in studies of ligand–protein interactions. The dissociation constant that characterizes a reversible equilibrium between a protein and a fluorescent ligand ($PL \leftrightarrow P + L$) is given by Eq. (24). To determine K_d, one must be able to measure the concentration of the ligand-protein complex i.e. the fraction of the ligand bound if the total ligand and total protein concentrations are known. Figure 9 shows typical titration curves for polarization and anisotropy data as a function of increasing total protein concentration while the ligand concentration is kept constant. In the simulated experiment, the observed polarization and anisotropy values are obtained from Eqs. (25) and (26) which are subsequently normalized via Eqs (5) and (6) to allow their use in expressions of various information measures derived in this work. One can then investigate how the information content is correlated to the fundamental experimental parameters x and u. Figures 10 and 11 show contour plots of absolute information obtained from anisotropy and polarization titrations in a simulated binding experiment as a joint function of x and u. The information content associated with K_d appears generally greater for anisotropy than for polarization. The most informative experimental range is approximately $x \in [0.4, 1.0]$ using anisotropy and $x \in [0.0, 0.6]$ using polarization. Figures 12 and 13 show the maximum potential information amount (analytical information) for the same experiment. In both cases, the larger information content is linked with larger x values of the bound ligand. Here, the most informative experimental range is about $x \in [0.3, 1.0]$ for anisotropy and $x \in [0.7, 1.0]$ for polarization. Thus, the information obtained from anisotropy measurements appears to be greater over a considerably larger range of the experimental u–x space. Accordingly, information measures may help limit the range of values for analysis. Note, however, that using extreme values of x (i.e. $x<0.1$ or $x>0.9$) may lead to large inaccuracies in the determination of dissociation constants. The experimental errors in x propagate most strongly into K_d at low and high values of x as described previously [11]. Selecting the optimal range of parameters is especially important in large scale screening experiments where a number of unknown proteins are tested simultaneously.

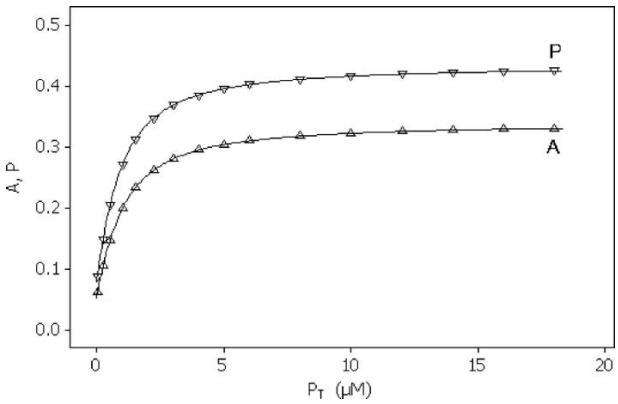

Fig. 9 Titrations of a fluorescent ligand with a protein: anisotropy (A) and polarization (P) vs. total protein concentration P_T. Data are simulated using Eqs. (24), (25) and (26). The parameters used are: $K_d = 1\ \mu M$, $L_T = 1\ \mu M$, $A_f = 0.060$, $A_b = 0.340$, $P_f = 0.087$, $P_b = 0.436$, and $g = 1.6$. P_T is varied from $0.055\ \mu M$ to $18\ \mu M$

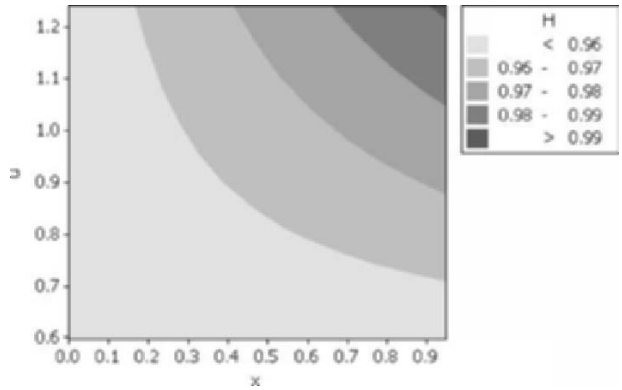

Fig. 10 Absolute information content (H) of fluorescence anisotropy (A) in a simulated binding experiment with respect to u and x (the fraction of ligand bound). In an ideal experiment, a preset value of K_d ($1\ \mu M$) is considered. Other parameters are as in Fig. 9. H is calculated by Eq. (10). The contour plot shows the information content reflected into the determination of K_d from the measurements of the polarized fluorescence intensities as a joint function of u and x

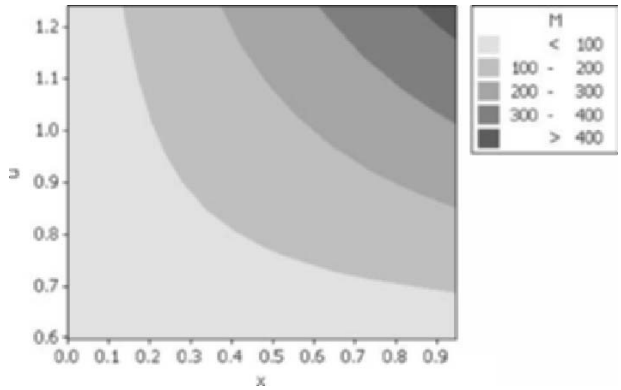

Fig. 12 Maximum potential information amount (M) of fluorescence anisotropy (A) in a simulated binding experiment with respect to u and x. M is calculated using Eq. (23). Parameters are as in Fig. 9

Some aspects of time-resolved fluorescence must also be examined as the differences between the timescales of the depolarization motions and the fluorescence lifetime of the fluorophore have direct relevance to studies of ligand–protein interactions. If the rotational correlation time (ϕ) associated with a depolarizing process is much shorter than the fluorescence lifetime (τ) the depolarization may be too rapid to be resolved. On the other hand, if a depolarizing process is much slower than the lifetime of the fluorophore, little depolarization will occur before emission. The approximate range of resolvable rotational correlation time is $\tau/10 \leq \phi \leq 10\tau$ [45]. Figure 14 shows the absolute information content of polarization and anisotropy as a function of the ratio of τ and ϕ for a simple system with only one rotational correlation time and one fluorescence lifetime. The information content for polarization appears to increase and level off with increasing lifetime whereas for anisotropy it is not a monotonous function but has a maximum. If one takes the limiting anisotropy to its greatest extent ($A_0 = 2/5$) the maximum occurs at $\tau/\phi = 3/5$, i.e. when the fluorophore's lifetime is 60% of the rotational correlation time. With smaller values of A_0, the maximum is shifted to smaller τ/ϕ ratios. Suppose a spherical protein of molecular weight 40 kDa has a rotational correlation time of 17 ns. One then calculates $\tau = 0.6 \times 17 \approx 10$ ns (at maximum limiting anisotropy) for the lifetime of an *ideal* fluorophore that would give maximum absolute information in an anisotropy decay experiment.

As presented on Fig. 15, the analytical information content diminishes as τ/ϕ increases both for polarization and anisotropy. This is a direct consequence of the fact that polarization and anisotropy are inversely proportional to the ratio of lifetime and rotation correlation time, see Eqs. (27) and (28). This inverse correlation, however, cancels in the numerator and denominator of absolute information function; therefore, Shannon's entropy does not show a monotonous decrease. The rate of decay in the analytical information function is different for polarization and anisotropy, the latter displaying a larger decrease from its higher initial value.

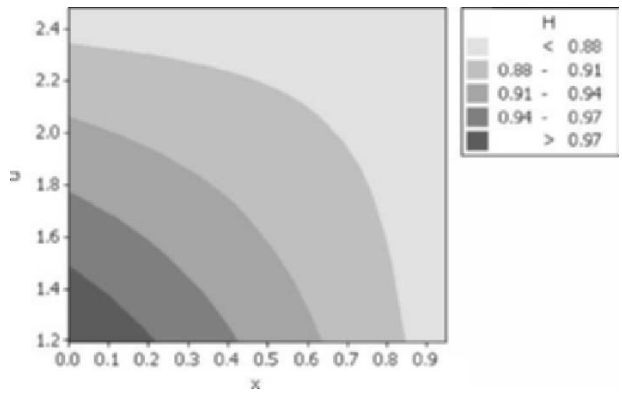

Fig. 11 Absolute information content (H) of fluorescence polarization (P) in a simulated binding experiment with respect to u and x. H is calculated by Eq. (10). Parameters are as in Fig. 9

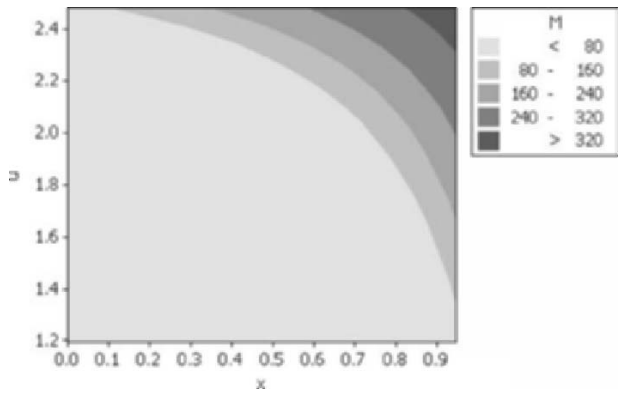

Fig. 13 Maximum potential information amount (M) of fluorescence polarization (P) in a simulated binding experiment with respect to u and x. M is calculated using Eq. (23). Parameters are as in Fig. 9

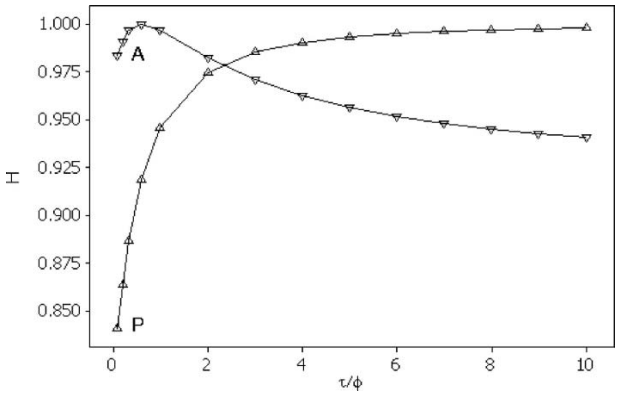

Fig. 14 Time dependence of the absolute information content (H) for polarization (P) and anisotropy (A). A simple spherical system with one fluorescence lifetime (τ) and one rotational correlation time (ϕ) is assumed. H is calculated using Eqs. (27), (28), and (10) via normalized values of P and A. The limiting polarization is $P_0 = 0.5$

The two functions yield the same information when $\tau/\phi \approx 1$ and the limiting polarization and anisotropy are at their respective maximum. With smaller values of P_0 and A_0, the isoinform point is shifted to lower τ/ϕ ratios. Since most extrinsic and intrinsic fluorophores used in ligand–protein studies have lifetimes less than 5 ns while proteins suitable for analysis typically have 5–20 ns rotational correlation times, the higher information content of the anisotropy function is effectively contained in the physically most useful τ/ϕ range. Consider, however, that some fluorescent ligands have very long lifetimes; for instance ruthenium can be hundreds of ns and certain pyrene probes can be more than a hundred ns. For such cases, the polarization function provides more information, both analytical (Fig. 15) and absolute (Fig. 14), but these probes could give very low polarization values.

Physical origin of the extra information

It follows from the preceding analysis that the larger differential information content of the polarization function is related to its larger normalized measurement range. The principal finding, however, is that anisotropy provides more absolute and potential information when polarization and anisotropy share the same normalized range. The absolute information in anisotropy is found to be larger when the ratio of the parallel and perpendicular signal components exceeds $2^{1/2}$ and it is larger for polarization below this threshold. The maximum potential information amount takes on a larger value for anisotropy in the entire overlapping interval and, in fact, delivers more information than polarization over its whole domain.

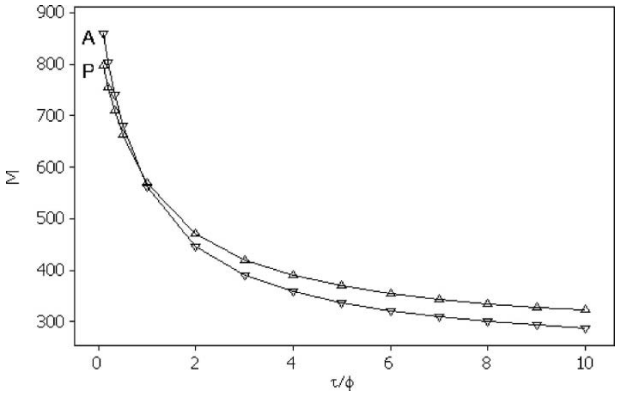

Fig. 15 Plots showing the analytical information amount (M) as a function of τ and ϕ for the system from Fig. 14. M is calculated using Eq. (23)

The extra information in the anisotropy function may be attributed to the presence of a second perpendicular intensity term in the denominator of its definition which determines the characteristic scale for anisotropy. The information amount depends not only on the information content of the signal but also on the way the signal is processed and the analysis is carried out. The second perpendicular emission plane is oriented along the axis of propagation of the polarized light. Therefore the anisotropy functions takes the contribution of all degrees of rotational freedom into consideration reflecting the orientations of molecules in the entire signal space via the photoselection process. When macroscopically isotropic systems are considered, both intensities that correspond to the perpendicular orientation of the polarized emission are the same. Thus anisotropy effectively gives proper weight to the perpendicular components that represent two-dimensional information, whereas the information from the single perpendicular component of polarization is one-dimensional. In other words, anisotropy accounts for molecular randomness in a larger signal space than polarization, and—by definition—information content is also a measure of randomness.

Conclusions

The information content of fluorescence polarization and anisotropy using various entropy measures was examined. While it may seem surprising that the information content for the two quantities are different as one measure can be converted to the other and vice versa, the difference mathematically originates from the fact that a scaling relation exists between them. Behind their formal definitions lies a physical point of view which is based on methodological differences regarding how the components of a fluorescent signal are measured. Scale normalization, however, allows one to investigate the scale-independent relations of the two quantities, as shown in this paper.

Although some details may be different, all information measures furnished concordant results. From the perspective of this work, perhaps the most accessible and most useful information-theoretic index is Shannon's entropy because

it does not depend on a particular reference distribution or on the analytical system; therefore it gives the most insight. Concluding this work, an opinion may be expressed that anisotropy is a more accurate representation of the underlying physical phenomena. As a whole, the differential information amount in fluorescence polarization measurements is larger over the whole recording range, but more importantly, the absolute and potential information obtained from anisotropy over its normalized common range with polarization attains a larger value. It should be emphasized that the information content of polarization is referenced to the polarized light in the direction of observation whereas that of the anisotropy function is related to the total amount of polarized light. The information content may reflect the real properties of the polarized intensity components, and is likely to be homomorphic to the Curie symmetry principle of the emitted fluorescent light in a broad sense. The absolute and differential information content can mathematically manifest the difference in the properties of the polarization and anisotropy function as well as the way the polarized signal is analyzed. These observations could be of use in further analysis of the quantitative aspects of two related quantities of the polarization state of fluorescence as a means in selecting the most informative function for a particular application. Future studies will be aimed at specific details of practical use.

Acknowledgements The author wishes to thank Professor David M. Jameson for his thoughtful comments and valuable discussions. This work was supported, in part, by a grant from the National Institute of Health (NIH/NCRR/RCMI G12-RR03061: Selective Research Excellence in Biomedicine and Health).

References

1. Lakowitz JR (1999) Principles of fluorescence spectroscopy, 2nd edn. Kluwer Academic(Plenum Publishers, New York
2. Sharma A, Schulman SG (1999) Introduction to fluorescence spectroscopy. Techniques in analytical chemistry series. John Wiley and Sons, Inc, New York
3. Valeur B (2002) Molecular fluorescence. Principles and applications. Wiley-VCH Verlag GmbH, Weinheim, Germany
4. Jameson DM, Croney JC, Moens PD (2003) Fluorescence, basic concepts, practical aspects and some anecdotes. Methods Enzymol 360:1–43
5. Weber G (1952) Polarization of the fluorescence of macromolecules. 1. Theory and experimental method. Biochem J 51:145–155
6. Weber G (1953) Rotational Brownian motion and polarization of fluorescence of solutions. Adv Protein Chem 8:415–459
7. Jablonski A (1960) On the notation of emission anisotropy. Bull Acad Polon Sci Ser Sci Math Astr et Phys 6:259–264
8. Weber G (1966) Polarization of the fluorescence of solutions. In: Hercules D (ed) Fluorescence and phosphorescence. Wiley, New York, pp 217–240
9. Jameson DM, Sawyer WH (1995) Fluorescence anisotropy applied to biomolecular interactions. Methods Enzymol 246:283–300
10. Jameson DM, Croney JC (2003) Fluorescence polarization, past present and future. Comb High Throughput Chem 6:167–173
11. Jameson DM, Mocz G (2005) Fluorescence polarization(anisotropy approaches to study protein-ligand interactions, Effects of errors and uncertainties. In: Nienhaus GU (ed) Methods in molecular biology. Protein–ligand interactions: methods and applications, vol 305. Humana Press Inc., Totowa, NJ, pp 301–322
12. Yguerabide J (1972) Nanosecond fluorescence spectroscopy of macromolecules. Methods Enzymol 26:498–578
13. Wahl Ph (1975) Fluorescence spectroscopy. In: Chen RF, Edelhoch H (eds) Biochemical fluorescence concepts, vol. 1. Marcel-Dekker, New York, pp 1–41
14. Shannon CE (1948) A mathematical theory of communication. Bell Syst Technol J 27(379–423):623–656
15. Shannon CE, Waever W (1949) Mathematical theory of communication. University of Illinois, Urbana
16. Brillouin L (1956) Science and information theory. Acad Press, New York
17. Brillouin L (1964) Scientific uncertainty and information. Acad. Press, New York
18. Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22(1):79–86
19. Kullback S (1959) Information theory and statistics. John Wiley & Sons, Inc., New York
20. Khinchin AI (1957) Mathematical foundations of information theory. Dover Publications, Inc., New York
21. Eckschlager K, Stepanek V (1979) Information theory as applied to chemical analysis. John Wiley and Sons, New York
22. Danzer K, Hopfe V, Marx G (1982) Möglichkeiten der Erhörung der Informationsmenge spectroskopischer Analysenmethode mit Hilfe der Rechentechnik. Z Chem 22:332–338
23. Hauser MD (1996) The evolution of communication. The MIT Press, Cambridge, Massachusetts
24. Bonchev D (1983) Information theoretic indices for characterization of chemical structures. Research Studies Press, Chichester, England
25. Eckschlager K, Stepanek V (1985) Analytical measurement and information. Advances in the information theoretic approach to chemical analyses. Research Studies Press Ltd. Letchworth, Hertfordshire, England
26. Eckschlager K, Danzer K (1994) Information theory in analytical chemistry. John-Wiley & Sons, Inc., New York
27. Magurran A (1988) Ecological diversity and its measurement. Princeton University Press, Princeton, New Jersey
28. Guiasu S (2001) Quantum mechanics. Nova Science Publishers, Huntington, New York
29. Alcala JR, Gratton E, Prendegrast FG (1987) Interpretation of fluorescence decays in proteins using continuous lifetime distributions. Biophys J 51:925–936
30. Lakowicz JR, Gryczynski I, Wiczk W, Johnson ML (1994) Distributions of fluorescence decay times for synthetic melittin in water-methanol mixture and complexed with calmodulin, troponin C, and phospholipids. J Fluoresc 4:169–177
31. Bronchon JC (1994) Maximum entropy method of data analysis in time-resolved spectroscopy. Methods Enzymol 240:262–311
32. van der Heide UA, Hopkins SC, Goldman YE (2000) Maximum entropy analysis of protein orientation using fluorescence polarization data from multiple probes. Biophys J 78:2138–2150
33. Sengupta P, Garai K, Balaji N, Periasamy N, Maiti S (2003) Measuring size distribution in highly heterogeneous systems with fluorescence correlation spectroscopy. Biophys J 84:1977–1984
34. Ash RB (1965) Information theory. Interscience, New York, New York
35. Levine RD, Tribus M (eds) (1979) The maximum entropy formalism. MIT Press, Cambridge, Massachusetts

36. Cover TM, Thomas JA (1991) Elements of information theory. John Wiley & Sons, Inc.
37. O'Neill EL (1963) Introduction to statistical optics. Addison-Wesley, Reading, MA
38. Smith JDH (2001) Some observations on the concepts of information-theoretic entropy and randomness. Entropy 3:1–11
39. Perrin F (1926) Polarisation de la lumière de fluorescence. Vie moyenne des molécules dans l'etat excitè. J de Phys. VIe Sèrie 7:390–401
40. Press WH, Flannery BP, Teukolsky AA, Vetterling WT (1989) Numerical recipes in pascal. The art of scientific computing. Cambridge University Press, NY
41. Renyi A (1965) On the foundations of information theory. Rev Intern Stat 3:1–14
42. Desper CR, Kimura I (1967) Mathematics of the polarized-fluorescence experiment. J Appl Phys 38:4225–4233
43. Badley RA (1976) Fluorescent probing of dynamic and molecular organization of biological membranes. In: Wehry EL (ed) Modern fluorescence spectroscopy. Plenum Press, New York, pp 91–168
44. Picozzi A (2004) Entropy and degree of polarization for nonlinear optical waves. Opt Lett 29:1653–1655
45. Wahl Ph (1979) Analysis of fluorescence anisotropy decays by a least squares method. Biophys Chem 10:91–104
46. Reflegler P (2005) Polarization degree of optical waves with non-Gaussian probability density functions, Kullback relative entropy-based approach. Opt Lett 30:1090–1092