For submission to J. Chem. Phys.

Shape-Independent Model (SHIM) Approach for Studying Aggregation by NMR Diffusometry

Adrian A. Hernandez Santiago [a], Anatoly S. Buchelnikov [b], Maria A. Rubinson [c], Semen O. Yesylevskyy [d], John A. Parkinson [e]*, Maxim P. Evstigneev [b,c]*

[a] Department of Physics and Mathematics, Faculty of Chemistry, Autonomous University of Puebla, Puebla, Mexico CP 72570

[b] Department of Biology and Chemistry, Belgorod State University, Belgorod 308015, Russia

[c] Department of Physics, Sevastopol State University, Universitetskaya str.33, Sevastopol 299053

[d] Department of Physics of Biological Systems, Institute of Physics of the National Academy of Sciences of Ukraine, Prospect Nauky 46, Kiev-28, 03680, Ukraine

[e] WestCHEM Department of Pure and Applied Chemistry, University of Strathclyde, 295 Cathedral Street, Glasgow G1 1XL, United Kingdom

Abstract

NMR diffusometry has been gaining wide popularity in various areas of applied chemistry for investigating diffusion and complexation processes in solid and aqueous phases. To date, the application of this method to study aggregation phenomena proceeding beyond the dimer stage of assembly has been restricted by the need for *a priori* knowledge of the aggregates' shape, commonly difficult to know in practice. We describe here a comprehensive analysis of aggregation parameter-dependency on the type and shape selected for modeling assembly processes, and report for the first time a shape-independent model (designated the SHIM-model), which may be used as an alternative in cases when information on aggregates' shapes are unavailable. The model can be used for determining equilibrium aggregation parameters from self-diffusion NMR data including equilibrium self-association constant and changes in enthalpy, ΔH, and entropy, ΔS.

Key words: NMR diffusometry, aggregation, self-diffusion, enthalpy, entropy.

Introduction

NMR diffusometry has become a popular routine method for characterizing molecular motion *via* translational diffusion in the solid and liquid states. The approach is extensively used in many areas of chemistry,[1-3] the field of research and development of associated methods and data treatments being active and vibrant.[4-7] Typical application of NMR diffusometry is to enable molecular aggregation and complexation phenomena to be quantified. So far this has been successfully applied in protein chemistry,[8] host-guest chemistry,[3] colloid chemistry,[9,10] inorganic chemistry,[11] supramolecular chemistry[12,13] and many other fields of chemical and materials research. A common approach makes use of the Einstein-Smoluchowski relation (eq 1) in order to link the translational diffusion coefficient, D, with the effective hydrodynamic radius (Stokes radius), R_{eff}, and the shape-factor (the so-called Perrin translational friction factor), f_P, which characterizes the deviation of the hydrodynamic shape of the studied object from an ideal sphere:

$$D = \frac{kT}{6\pi\eta R_{eff} f_P}, \qquad (1)$$

where k, T, η are the Boltzmann constant, absolute temperature and viscosity, respectively.

Equation 1 can only be used if an aggregate's exact shape is explicitly known, creating a major problem in the use of NMR diffusometry as a general method for studying aggregation phenomena, as discussed in detail here.

The magnitude of D is measured through NMR-based diffusion studies and embodies the aggregation parameters of interest. The Perrin translational friction factor, f_P, on the other hand contains information concerning the shape of the studied object. Once the link between f_P and the geometry of the object is established, eq 1 can be directly applied to fit experimental titration data (studied in the form of concentration dependency of D) and used for extracting relevant aggregation

parameters as *adjustable* quantities. In the basic cases of dimerization or 1:1 complex formation, the diffusion coefficients of the monomer, D_1, and dimer (or complex), D_2, commonly act as such adjustable quantities.[8,10,14] In these instances, knowledge of the exact form of f_P is not strictly required. Consequently, the overwhelming majority of known NMR diffusometry applications have successfully used such an approach (for reviews see references 1 and 3). The critical point of departure addressed by us in this article occurs if the aggregation process extends beyond the dimer stage. For such a condition, an explicit model is required describing the dependence of hydrodynamic shape on the dimensions of aggregates formed. Lack of knowledge associated with this dependency creates fundamental difficulty in applying any type of diffusometry for investigating aggregation phenomena. Indeed, the total number of papers dealing with aggregation beyond the dimer assembly stage is notably much smaller compared with simple dimerization or 1:1 complexation. Two main reasons are considered to be responsible for this.

Firstly, in practice the shape of aggregates is commonly unknown. Moreover, shape may change as a function of the increasing number of molecules responsible for forming an aggregate. Secondly, only a few classical shapes currently allow analytical equations to be written for the dependence between f_P and aggregate geometry (usually in the form of either a sphere, cylinder or oblate/prolate ellipsoid[2,13,15]). Any other shapes lead to significant difficulties in the computational implementation of the fitting procedure. This is probably the main reason why the majority of published papers introduce the simplest spherical shape to represent aggregates, with a very minor fraction of papers dealing with ellipsoid or other shapes.[13,16,17] It is also obvious that a spherical model shape used to represent an aggregate cannot cover the majority of probable shapes encountered in reality. Thus, the dependence of NMR diffusometry on a knowledge of the exact hydrodynamic shape of aggregates remains as the major bottleneck limiting the expansion of this approach towards the investigation of aggregation phenomena in general.

The aim of the present work is therefore to illustrate the shortcomings of modeling the dependence of the translational diffusion coefficient, *D*, measured *via* NMR diffusometry, on defined shape and to find a way to successfully bypass this shape dependency by introducing a modeling approach that is shape-independent (the SHIM approach). In this article NMR diffusometry is used to probe aggregation phenomena in terms of translational diffusion for different types of small molecules known to exert well-characterized aggregation tendencies in solution. To assist the reader, an explanation of the flow and structure of the article is provided as follows.

Firstly, a strategy detailing the rationale and criteria behind the choice of molecules for the investigation is laid out. Secondly, for those hydrodynamic shapes most widely encountered already within the literature, expressions are defined that allow equations to be derived for determining the translational diffusion coefficient for each type of shape (Table 1) for illustration and comparison purposes. Expressions for the diffusion coefficients of aggregates of each of these shapes follow from these definitions (viz. Equations 3). The expressions are then used to define the manner by which experimentally measured diffusion coefficients are treated and modeled: weighted averages of values from different sized aggregates are considered based on monomer and dimer diffusion coefficients for each shape separately resulting in Equations 5-8. Modeling of the measured diffusion coefficients for all molecules in the series is carried out with each of the shape-based models in turn to yield a matrix of results illustrative of the current approach adopted throughout the literature and that are treated according to five specific considerations (see Method of selection of the most appropriate model). The analysis of these results and the accompanying considerations are then used to guide the process by which the SHape Independent Model (SHIM) approach expressions are derived by highlighting the link between diffusion and the so-called friction coefficient. This yields expressions 12-14 for the new model, the latter providing a convenient form of the SHIM approach expressed using the hypergeometric function *F*. Finally the results of the analysis comparing the results from the SHIM

approach to each of the shape-dependent models are summarized (Table 3) and used for determining the fit between calculated thermodynamic parameters based on the SHIM-model and those reported in the literature for a subset of the molecules used in this study.

Results and Discussion

Strategy of investigation.

The target parameter of interest that most fully characterizes the equilibrium aggregation process is the equilibrium self-association constant, K (or Gibbs free energy change on aggregation).[27] The magnitude of K can be obtained from the dependence of the observable parameter (i.e. magnetization decay in NMR diffusometry data, directly transformed into D) on solute concentration, x_0, (i.e. *via* titration dilution experiments) by fitting these data with a certain model. The NMR-based diffusion aggregation model will always depend *a priori* on the chosen hydrodynamic shape of the aggregates. For the purposes of this work it was concluded that the shape dependence of the aggregation process be investigated through evaluation of the variation in magnitude of K (derived from the dependence of D on x_0) as a function of different models. As a reference K-value, it was proposed that the equilibrium constant derived from ^{1}H NMR titration data be used (i.e. the dependence of proton chemical shifts, δ, on x_0) recorded in parallel with NMR-based diffusion data on the same solutions. Such a strategy allows the well-known dependence of K on concentration range to be ruled out of influencing the investigation together with the type of experiment used to produce the titration curves (see ref. 28 for a full review).

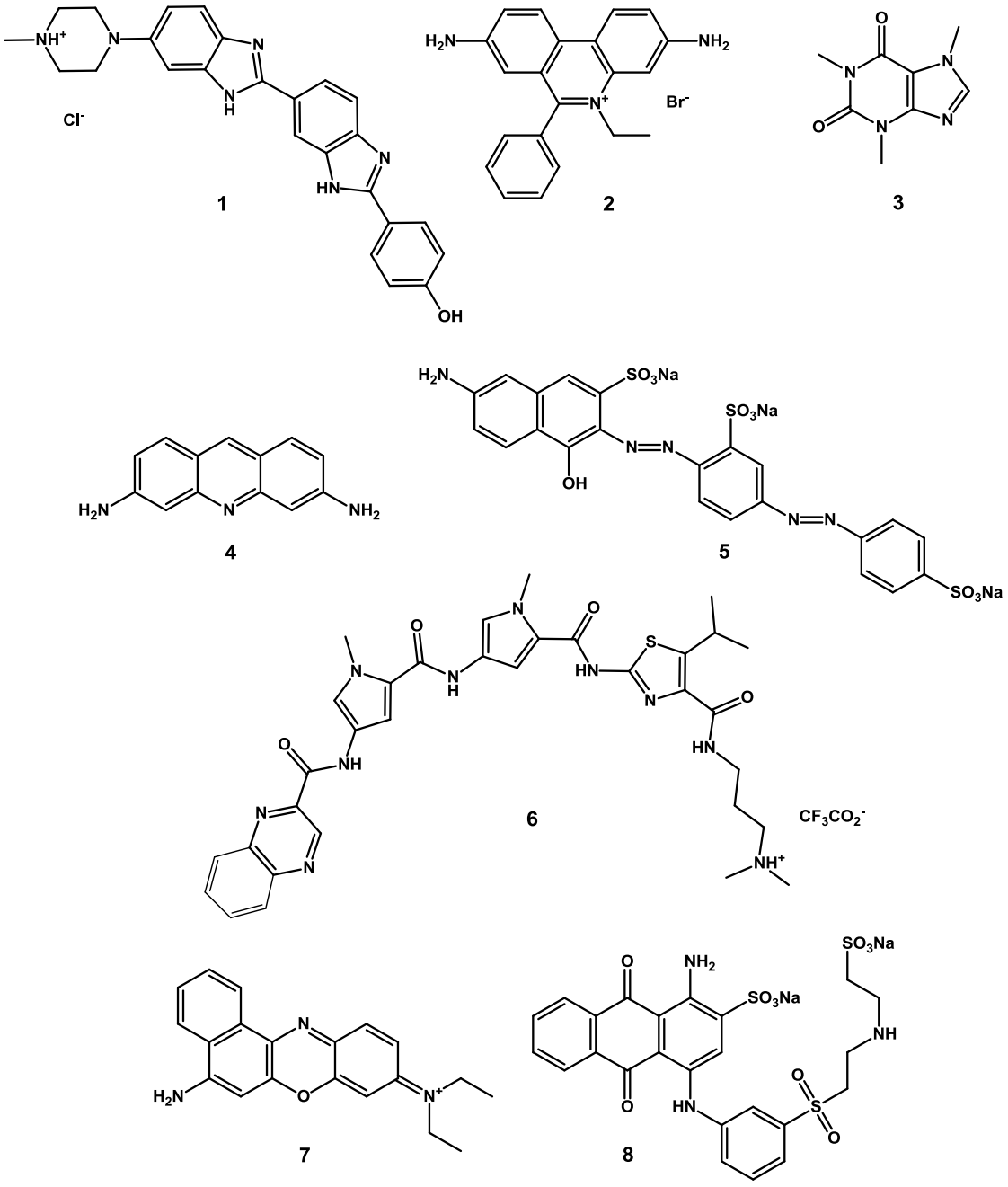


Fig. 1 Test molecules used for studying aggregation phenomena by means of NMR diffusometry.

Selection of the compounds for study (see Materials and Methods and **Figure 1**) was dictated by the following set of criteria:

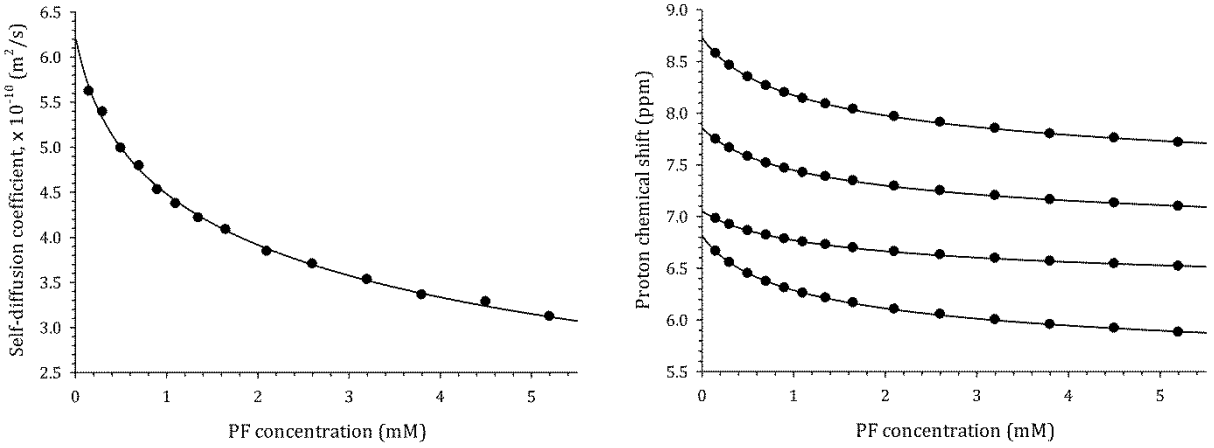
a) the molecules must feature different shapes in order to create differently shaped aggregates. However, the exact shape of any aggregate could not be predicted based on the shape of the

molecule alone and in each particular case must be discussed separately. In particular, the aromatic molecules not containing heavily branched side chains, *viz*. compounds **2**, **3**, **4** and **7** should follow a linear-type aggregation process, presumably matching cylindrical or ellipsoid shapes of aggregates, whereas for the rest of the molecules it is difficult to predict the aggregate's shape,

b) the aggregation tendency of the test compounds must vary in order to account for the dependence of the measured value of D on the magnitude of the self-association constant. The set of molecules selected feature a dispersion of K values spread over several orders of magnitude ranging from 11 M^{-1} (for **3**) up to 5600 M^{-1} (for **7**),

c) the test molecules must contain enough well-resolved non-exchangeable protons to allow reliable $D(x_0)$ and $\delta(x_0)$ curves to be established.

Experimental self-diffusion, $D_{obs}(x_0)$, and chemical shift, $\delta(x_0)$, data are shown in **Fig. 2** for compound **4** as a typical example. The data for the remaining compounds are provided within the Supporting Information. The behavior of the experimental curves is qualitatively similar for all of the molecules studied, *viz*. shift of the $\delta(x_0)$ curves to lower NMR frequency and shift of $D_{obs}(x_0)$ curves to lower values of diffusion coefficients on increasing the solute concentration. These features are typical of aggregation processes occurring by stacking of aromatic chromophores.[10,13,27] It is also worth noting that the concentrations of the test molecules used to obtain the titration curves fall into the low millimolar range, which is negligible compared with the concentration of the solvent molecules (D_2O). This allows any changes in viscosity of the solvent to be considered negligible and therefore capable of being ignored in the data treatment made here.



158

159 **Fig. 2** Experimental dependence of self-diffusion coefficient, $D_{obs}(x_0)$, and proton chemical shift, $\delta(x_0)$,

160 on concentration, x_0, for **4**, PF, taken as a typical example.

161

162 *Hydrodynamic shapes.*

163 As discussed in the preceding dialogue, there are three main types of shapes currently in use in

164 the majority of NMR diffusion studies concerning aggregation phenomena, namely the sphere, the

165 cylinder and the ellipsoid. Each of these general models can be further reduced to more specific shapes.

166 The link between the types of shape and the translational diffusion coefficient are detailed below.

167 Equation 1 can be re-written as:

168 $$D = \frac{kT}{r}, \qquad (2)$$

169 where $r = r_{sphere} f_P$ is the friction coefficient in which $r_{sphere} = 6\pi\eta R_{eff}$ is the coefficient of translational

170 resistance for the sphere. It should be noted that in the case of the ellipsoidal or cylindrical geometries

171 R_{eff} denotes the radius of the sphere of equivalent volume.[29] By evaluating the Perrin translational

172 friction factor, f_P, for a given shape, the final equation for diffusion coefficient can be obtained

173 directly from eq 1 according to the following discussion.

Let $p = a/b$ be the axial ratio where *a* and *b* are the major and minor semi-axes of an ellipsoid (or the half-length and radius of a cylinder). Note, if *a* = *b* then one gets the degenerate case of a sphere. Once these notations are introduced, the Perrin translational friction factors can be written in exact form.[29,30] **Table 1** summarizes all the formulas for the above-mentioned geometries. Evaluating R_{eff} and substituting it into the equation for the friction coefficient, *r*, along with Perrin factor, f_P, yields the final equations for translational diffusion coefficients in explicit form (last row in **Table 1**).

Table 1 Collection of formulas necessary to derive equations for translational diffusion coefficients for the most widely used geometrical shapes

Parameter	**Geometry**			
	Spheroid $a = b$ $(p = 1)$	**Oblate ellipsoid** $a < b$ $(p < 1)$	**Prolate ellipsoid** $a > b$ $(p > 1)$	**Cylinder** $a \neq b$ $(p \neq 1)$
Volume, *V*	$\frac{4}{3}\pi b^3$	$\frac{4}{3}\pi a^2 b = \frac{4}{3}\pi p^2 b^3$	$\frac{4}{3}\pi a b^2 = \frac{4}{3}\pi p b^3$	$2\pi a b^2 = 2\pi p b^3$
Effective hydrodynamic radius, $R_{eff} = \left(\frac{3V}{4\pi}\right)^{1/3}$	b	$p^{2/3} b$	$p^{1/3} b$	$\left(\frac{3}{2}\right)^{1/3} p^{1/3} b$
Perrin translational friction factor,	1	$\frac{\sqrt{1-p^2}}{p^{1/3} \arcsin\sqrt{1-p^2}}$	$\frac{\sqrt{p^2-1}}{p^{1/3} \ln\left(p+\sqrt{p^2-1}\right)}$	$\left(\frac{2}{3}\right)^{1/3} \frac{p^{2/3}}{\ln p + \nu}$

f_P				
Translational friction coefficient, $r = 6\pi\eta R_{eff}\, f_P$	$6\pi\eta b$	$6\pi\eta b \dfrac{p^{1/3}\sqrt{1-p^2}}{\arcsin\sqrt{1-p^2}}$	$6\pi\eta b \dfrac{\sqrt{p^2-1}}{\ln\left(p+\sqrt{p^2-1}\right)}$	$6\pi\eta b \dfrac{p}{\ln p + \nu}$
Translational diffusion coefficient, $D = kT/r$	$\dfrac{kT}{6\pi\eta b}$	$\dfrac{kT}{6\pi\eta b}\dfrac{\arcsin\sqrt{1-p^2}}{p^{1/3}\sqrt{1-p^2}}$	$\dfrac{kT}{6\pi\eta b}\dfrac{\ln\left(p+\sqrt{p^2-1}\right)}{\sqrt{p^2-1}}$	$\dfrac{kT}{6\pi\eta b}\dfrac{\ln p + \nu}{p}$

Note: In the case of an aggregate of cylindrical shape $\nu = 0.312 + 0.565/p - 0.100/p^2$ (a discussion of the parameter ν is detailed in the dialogue which follows later in this work).

Hydrodynamic models of aggregation.

The most common case of molecular aggregation is the growth of aggregates by sequential addition of monomers.[27,31] Hence, the geometrical parameters of any immediate aggregate (*a* and *b*) and, consequently, the diffusion coefficient, *D*, in eq 3, can be expressed via the number of molecules, *i*, in the aggregate.

For an oblate ellipsoid, $p < 1$ so that the major semi-axis, *a*, corresponds to the radius of the molecule (*d*/2, where *d* is the diameter), whereas the minor semi-axis, *b*, corresponds to half the sum of monomers constituting an aggregate: $a = d/2$, $b = Li/2$, $p = d/(Li)$, where *L* is the average thickness of a monomer unit. As an indicator, for molecules containing aromatic rings, it is common practice to take *L* = 0.34 nm, which is associated with the typical van der Waals distance between aromatic

surfaces.[15] In a prolate ellipsoid, $p > 1$ so that the major semi-axis, *a*, corresponds to half the sum of monomers constituting an aggregate, whereas the minor semi-axis, *b*, represents the radius of the molecule, similar to that in the cylindrical models: $a = Li/2$, $b = d/2$, $p = Li/d$. Considering an aggregate as a spheroid, the former is represented as a sphere of equivalent volume, which is the sum of equivalent volumes of constituent monomers. Thus, the equivalent radius, *b*, can be evaluated in terms of the monomer diameter, *d*: $b = i^{1/3} d/2$. Substitution of these relations into the equations from the last row of **Table 1** yields the diffusion coefficients of aggregates, D_i, for the standard set of shapes:

$$
\begin{aligned}
&\text{Sphere:} && D_i = \frac{kT}{3\pi\eta d i^{1/3}} \\
&\text{Oblate ellipsoid:} && D_i = \frac{kT}{3\pi\eta (Li)^{2/3}} \frac{\arcsin\sqrt{1-(d/(Li))^2}}{d^{1/3}\sqrt{1-(d/(Li))^2}} \\
&\text{Prolate ellipsoid:} && D_i = \frac{kT}{3\pi\eta} \frac{\ln\left(Li/d + \sqrt{(Li/d)^2 - 1}\right)}{\sqrt{(Li)^2 - d^2}} \\
&\text{Cylinder:} && D_i = \frac{kT}{3\pi\eta Li}\left(\ln(Li/d) + \nu(i)\right)
\end{aligned}
\qquad (3)
$$

Specifically for the cylindrical model a correction for the end effects is sometimes introduced in the form of a correction factor $\nu(i) = 0.312 + 0.565\, d/(Li) - 0.100\left(d/(Li)\right)^2$.[13,32]

Equations 3 provide explicit interrelation between D_i and *i* for basic shapes. It is, however, apparent that the shapes of aggregates at the monomer and dimer level may significantly deviate from those assumed for larger aggregates. Considering that the fraction of monomers and dimers typically dominate over other species in solution (if the aggregation process is not strongly cooperative), it is reasonable to introduce the diffusion coefficient of monomer, D_1, and dimer, D_2, as adjustable quantities. Such an approach will minimize the error from assigning basic shapes to the monomer

and/or dimer. Now, eq 3 may be used to express the experimentally observed translational diffusion coefficient obtained *via* NMR diffusion experiments, D_{obs}, as a weighted average of D_i:[9,33]

$$D_{obs} = \frac{1}{x_0}\sum_i D_i x_i \,, \tag{4}$$

where $x_i = i x_1 \left(K x_1\right)^{i-1}$ is the concentration of an aggregate containing *i* molecules.

Each model was used in two forms, *viz*. with variation of D_1, and with variation of D_1/D_2. Below are listed the set of final expressions used in the analysis of experimental NMR diffusometry data with the quantities in square brackets describing the adjustable parameters in the model.

SPHERICAL:

[D_1, D_2, K, d]
$$D_{\text{obs}} = \frac{x_1}{x_0}\left(D_1 + 2Kx_1D_2 + \frac{kT}{3\pi\eta d}\sum_{i=3}^{\infty} i^{2/3}\left(Kx_1\right)^{i-1}\right), \tag{5}$$

OBLATE ELLIPSOID:

[D_1, K, d]
$$D_{\text{obs}} = D_1\frac{x_1}{x_0}\sum_{i=1}^{\infty} i^{1/3}\left(Kx_1\right)^{i-1}\frac{\arcsin\sqrt{1-\left(d/\left(Li\right)\right)^2}}{\arcsin\sqrt{1-\left(d/L\right)^2}}\sqrt{\frac{1-\left(d/L\right)^2}{1-\left(d/\left(Li\right)\right)^2}}\,, \tag{6.1}$$

[D_1, D_2, K, d]
$$D_{\text{obs}} = \frac{x_1}{x_0}\left(D_1 + 2Kx_1D_2 + \frac{kT}{3\pi\eta L^{2/3}d^{1/3}}\sum_{i=3}^{\infty} i^{1/3}\left(Kx_1\right)^{i-1}\frac{\arcsin\sqrt{1-\left(d/\left(Li\right)\right)^2}}{\sqrt{1-\left(d/\left(Li\right)\right)^2}}\right), \tag{6.2}$$

PROLATE ELLIPSOID:

[D_1, K, d]
$$D_{\text{obs}} = D_1\frac{x_1}{x_0}\sum_{i=1}^{\infty} i\left(Kx_1\right)^{i-1}\frac{\ln\left(Li/d+\sqrt{\left(Li/d\right)^2-1}\right)}{\ln\left(L/d+\sqrt{\left(L/d\right)^2-1}\right)}\sqrt{\frac{L^2-d^2}{\left(Li\right)^2-d^2}}\,. \tag{7.1}$$

[D_1, D_2, K, d]
$$D_{\text{obs}} = \frac{x_1}{x_0}\left(D_1 + 2Kx_1D_2 + \frac{kT}{3\pi\eta}\sum_{i=3}^{\infty} i\left(Kx_1\right)^{i-1}\frac{\ln\left(Li/d+\sqrt{\left(Li/d\right)^2-1}\right)}{\sqrt{\left(Li\right)^2-d^2}}\right). \tag{7.2}$$

CYLINDRICAL:

[D_1, K, d]
$$D_{\text{obs}} = D_1 \frac{x_1}{x_0} \sum_{i=1}^{\infty} \left(Kx_1\right)^{i-1} \frac{\ln\left(Li/d\right) + \nu\left(i\right)}{\ln\left(L/d\right) + \nu\left(1\right)}, \tag{8.1}$$

[D_1, D_2, K, d]
$$D_{\text{obs}} = \frac{x_1}{x_0}\left(D_1 + 2Kx_1D_2 + \frac{kT}{3\pi\eta L} \sum_{i=3}^{\infty} \left(Kx_1\right)^{i-1} \left[\ln\left(Li/d\right) + \nu\left(i\right)\right]\right). \tag{8.2}$$

The monomer concentration, x_1, for all the models listed above takes the standard form for isodesmic aggregation:[9,15,17,27]

$$x_1 = \frac{1 + 2Kx_0 - \sqrt{1 + 4Kx_0}}{2K^2x_0}. \tag{9}$$

The self-diffusion data, $D_{obs}(x_0)$, were also treated using the dimer model of aggregation, which assumes that no aggregation proceeds beyond the dimer stage:[27]

DIMER:

[D_1, D_2, K]
$$D_{\text{obs}} = D_2 + \frac{2\left(D_1 - D_2\right)}{1 + \sqrt{1 + 8Kx_0}}. \tag{10}$$

The proton chemical shift titration data, $\delta(x_0)$, used as a reference, were treated according to the standard isodesmic model of self-association:[27]

^{1}H NMR ISODESMIC MODEL

[δ_1, δ_2, K]
$$\delta\left(x_0\right) = \delta_1 + \left(\delta_2 - \delta_1\right) \frac{2Kx_0 + 1 - \sqrt{4Kx_0 + 1}}{Kx_0}, \tag{11}$$

where δ_1, δ_2 are chemical shifts in monomer and dimer states, respectively.

Method of selection of the most appropriate model.

The following considerations have been taken into account when analyzing the results of computations over different models and different molecules:

1. All of the adjustable parameters must take physically meaningful positive values. Otherwise the model is considered inappropriate.
2. It is assumed that for a well-performing model, the magnitude of K should be as close as possible to the ^{1}H NMR derived constant obtained under similar solution conditions. However, it is known that different methods may yield different values of K and none of them may be considered as the most exact. This is also the case when comparing NMR diffusion and ^{1}H NMR-derived constants. It is accepted that if NMR diffusion and ^{1}H NMR-derived constants differ by an order of magnitude, the model is considered inappropriate.
3. The discrepancy function, Δ (or, alternatively, the goodness of fit, R^2), *i.e.* the mean square deviation of the theoretically calculated D values from the experimentally observed D_{obs} values, served as an additional criterion for selecting the best performing model, *viz.* the lower the value of Δ (or the higher the value of R^2), the better the model. One important point should be taken into account. Different models tested in the present work use different numbers of search parameters (between 2 and 4). Consequently the discrepancy of the model with a lower number of parameters may be slightly worse than that of the other models having larger numbers of parameters. This fact does not necessarily imply a poor model. However, if the discrepancy of a certain model in the analysis appears to be an order of magnitude worse than that of the others, it can serve as an indication that this model is not appropriate.
4. The magnitude of D_1 must always be higher than D_2. Taking the spherical model as an initial approximation, it follows that $D_1/D_2 \approx \sqrt[3]{2} \approx 1.26$.[29] This relationship was taken as a guess value for D_2 in data fitting. In order to estimate the meaningful range of D_1/D_2, variation in modeling

the self-diffusion process for the monomer and dimer for the selected set of molecules was performed (**Table 2**). It may be seen that on average the relation D_1/D_2 is rather close to the spherical approximation. The model which gives values outside the range $1 < D_1/D_2 < 2$ must be treated with caution.

5. The physically meaningful values of the *d* parameter in the models (5)-(8) are strongly dependent on the geometry of the molecule, but may be limited from the upper and lower side by taking into account the typical dimensions of aromatic heterocycles. For the set of the compounds studied in the present work it was assumed that the values of *d* falling outside the range 0.3 nm < *d* < 3 nm are erroneous.

Table 2 Magnitudes of monomer (D_1) and dimer (D_2) translational diffusion coefficients (10^{-10} m^2/s) in D_2O calculated by means of molecular dynamics simulation

Molecule	D_1	D_2	D_1/D_2 †
2	6.7	5.5	1.22
3	11.3	8.8	1.28
4	10.4	8.2	1.27

† Note: similar but higher values of D_1 and D_2 have been obtained in H_2O (data not shown), preserving virtually the same values of D_1/D_2 as those shown in the table.

Analysis of the results of calculations using various hydrodynamic models.

The result of computations covering the set of hydrodynamic models described above and applied in order to fit the $D_{obs}(x_0)$ titration (dilution) data, and the reference calculations of the self-association constant using $\delta(x_0)$ titration (dilution) data (see **Figure 2** and Supporting Information) are

294 presented in **Table 3** (Strategy 1) as qualitative representations and in the Supporting Information in a
295 quantitative form. The following conclusions may be drawn from inspection of these results (only for
296 Strategy 1 for now), omitting in the first instance the results obtained from the dimer model:

297 (i) The results for the molecules containing (**2**, **3**, **4**, **7**, **8**) and not containing (**1**, **5**, **6**) a rigid
298 aromatic chromophore do not show clear preference for a particular model suggesting that the
299 aggregation is relatively insensitive to the type of hydrodynamic model used. The latter may be
300 interpreted by the fact that the aggregation of these compounds in the concentration range studied
301 (limited by the solubility) is not pronounced, i.e. the contribution from aggregates of higher order
302 than dimer is relatively unimportant, thus attenuating the influence of the selection of the type of
303 shape in the model. The quality of fit of the diffusion data with various models for these
304 compounds is very similar and does not allow unambiguous selection of the best model by this
305 criterion;

306 (ii) The ellipsoid and cylindrical models with three adjustable parameters (i.e. eqs 6.1, 7.1, 8.1) for
307 the majority of molecules failed to describe the experimental data, whereas addition of D_2 as a
308 fourth adjustable parameter (i.e. eqs 6.2, 7.2, 8.2) enabled the data to be fitted with meaningful
309 outcomes. Hence, it is recommended that D_2 be always used in an explicit form when carrying
310 out numerical analysis of self-diffusion data for aggregation;

311 (iii) An apparent improvement of the performance of the cylindrical model is seen when the
312 correction for the end effects is introduced, which is in agreement with the current view; [13,32]

313 (iv) The spherical model with four parameters (eqs 5) showed the best performance as compared with
314 other models. It allows partial explanation as to why the spherical model has so far been applied
315 in the majority of cases for investigation of aggregation processes, as alluded to in the
316 introductory section of this article;

(v) Even though the shape-dependent models have, in general, shown good performance for different shapes of molecules, there remains a problem in verifying the reliability of the calculated magnitude of the parameter *d*, which is not possible to estimate based on the shape of the molecule or its dimer. Moreover, the results of calculations presented in the Supporting Information demonstrate high dispersion of *d* across the models studied. This result is difficult to interpret and is most likely unreliable. Hence, any use of spherical, ellipsoid or cylinder model must be treated with caution.

In summary, it is possible to establish initially that the aggregation processes of the test compounds appears not to be strongly related to the type of shape used in the hydrodynamic model. The additional test of this assumption was accomplished by varying D_1 and D_2 simultaneously such that the condition $D_1/D_2 \approx \sqrt[3]{2} \approx 1.26$ was always matched during the data fitting procedure, which is compliant with the results of molecular modeling (see above), and allows the number of adjustable parameters to be reduced. The results of these computations are shown (**Table 3**, Strategy 2). According to this approach, the spherical and cylindrical models (13 and 16) appear to be most appropriate for the largest number of molecules studied, suggesting that Strategy 2 (three adjustable parameters) may be recommended for the numerical analysis of self-diffusion data for self-aggregating systems using these models. However, the dispersion of *d* remains the most problematic issue.

In summary it may be concluded that the use of shape-dependent models (either spherical or cylindrical) with Strategies 1 or 2 is applicable only if some *a priori* information regarding an aggregate's shape is available enabling the value of *d* to be estimated. If such information is absent (which is the most likely scenario in practice), the present work shows that based on goodness of fit data alone, it is not possible to unambiguously select the most appropriate shape-dependent hydrodynamic model.

Development of shape-independent model (SHIM-model).

Taking into account i) the relative insensitivity of the aggregation parameters derived from diffusion NMR data to the shape selected in the model, ii) the difficulty in practice of predicting the shape of aggregates based only on the structure of monomer or dimer, and iii) the difficulty in *a priori* knowledge of the magnitude of the *d* parameter, the possibility of developing a model which does not introduce any assumptions about the type of shape and is free of the problem of the *d* parameter, is considered here as an alternative approach.

The key quantity in eq 2 is the friction coefficient, *r*, which appears in the standard equation for a resistance force in solution experienced by a molecule on moving with speed, *v*, *viz.* $F = -r \cdot v$. Force is an additive quantity. Hence, to a first approximation, this additive property can be transferred to *r* as well. Based on this assumption, it is possible to express the stepwise addition of a molecule to an aggregate in terms of a stepwise addition of the same quantity, Δr, to *r*, *i.e.* $r_i = r_1 + \Delta r\left(i-1\right)$, where *i* is the number of molecules in an aggregate. Diffusion and friction coefficients are linked to each other via eq 2, *i.e.*

$$D_i = \frac{kT}{r_i}; \text{ at } i = 2,\ D_2 = \frac{kT}{r_2}.$$

The latter allows the expression $\Delta r = \frac{kT}{D_2} - \frac{kT}{D_1}$ to be derived. Further use of this relation to derive the expression for the NMR observable self-diffusion coefficient follows as:

$$D_{obs} = \sum_{i=1}^{\infty} iD_i \frac{x_i}{x_0} = \frac{x_1}{x_0}\left[D_1 + \sum_{i=2}^{\infty} \frac{iD_1D_2\left(Kx_1\right)^{i-1}}{D_2 + \left(i-1\right)\left(D_1 - D_2\right)} \right], \qquad (12)$$

where x_1 is determined from eq 9 in a similar way to that from the shape-dependent models.

Equation 12 can finally be expressed in a more convenient form, representing the shape-independent model (the SHIM-model):

SHIM-model:

[D_1, D_2, K]

$$D_{\text{obs}} = \frac{x_1}{x_0}\alpha D_1 \sum_{i=0}^{\infty} \frac{i+1}{i+\alpha}\left(Kx_1\right)^i \text{, where } \alpha = \frac{D_2}{D_1 - D_2}. \quad (13)$$

Equation 13 can be further rewritten in more convenient form using the hypergeometric function, F, as follows:

$$D_{\text{obs}} = D_1 \frac{x_1}{x_0} F\left(2, \frac{D_2}{D_1 - D_2}; \frac{D_1}{D_1 - D_2}; Kx_1\right). \quad (14)$$

Such notation avoids the need for direct programming of the infinite summation in eq 13 being replaced instead with the standard hypergeometric function, available in the majority of mathematical software packages (e.g. MATLAB or MathCAD).

The results from computations using the SHIM-model are shown in **Table 3** for Strategies 1 and 2, and in the Supporting Information. Within Strategy 1, the SHIM-model with three adjustable parameters gives the same performance as the spherical model with four parameters (which is considered as the best over others) with nearly the same goodness of fit (see Supporting Information). Within Strategy 2 the SHIM-model has succeeded for all test molecules alike versus the spherical model. Recall that the SHIM-model is free of the problem of the *d* parameter discussed above, and gives nearly the same goodness of fit as the spherical model in both strategies but with lower number of adjustable parameters (4 vs. 3, or 3 vs. 2 parameters). It thus may be concluded that in cases when

385 the hydrodynamic shape of aggregates is unknown and the *d* parameter cannot be predicted, the SHIM-
386 model has an advantage over any other shape-dependent model.
387

388 **Table 3** Qualitative indication of when the model succeeded (shaded cell) or failed (blank cell) to fit

389 experimental data and/or to match the reference parameters

390

Models			**Molecules**							
No. of model in Supporting Information	type of the shape	number of adjustable parameters	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Strategy 1 (D_1 and D_2 are independent variables)										
1	Dimer model	3								
2	Spheroid	4								
3	Oblate ellipsoid	3								
4		4								
5	Prolate ellipsoid	3								
6		4								
7	Cylinder without correction	3								
8		4								
9	Cylinder with correction	3								
10		4								
11	SHIM-model	3								
Strategy 2 (fixed ratio D_1/D_2 = 1.26)										
12	Dimer model	2								
13	Spheroid	3								
14	Oblate ellipsoid	3								
15	Prolate ellipsoid	3								
16	Cylinder	3								
17	SHIM-model	2								

391

392

In order to provide additional reliability tests for the computational results obtained using the SHIM-model (specifically model 11 in **Table 3**) with respect to the number of experimental points measured, we recalculated the set of adjustable parameters by sequentially excluding one to three experimental data points randomly selected from the entire range of measured concentrations for each compound studied. The results are presented in the Supporting Information and clearly suggest that exclusion of even three data points does not change the magnitude of the adjustable parameters to any significant extent that could be considered to alter the conclusions formulated above regarding the comparison of different models.

Peculiarity of the dimer model with respect to self-diffusion data.

The use of the dimer model to treat self-diffusion data (intentionally omitted above) is linked to the fundamental problem associated with dimer and isodesmic models. These are indistinguishable from one another with respect to the goodness of fit of the titration data (see ref. 28 for a review). This must therefore be discussed separately. More simply put, it is not possible to distinguish between dimer and indefinite aggregation based on the magnitude of the discrepancy function, Δ, only. It has been shown[28] that this indistinguishability originates from the use of two basic assumptions in the model: (i) the observable is given as an additive quantity over the molecules forming an aggregate; (ii) the observable is influenced only by nearest neighbors in an aggregate. The majority of known experimental methods implicitly or explicitly use these assumptions in treating the aggregation process. Hence, the property of indistinguishability is intrinsic to many widespread physico-chemical methods such as NMR, spectrophotometry, microcalorimetry and so forth. It was also suggested[28] that any approach not meeting any of these two assumptions may potentially resolve the problem of indistinguishability. It is therefore worth considering whether this is possible within the diffusion NMR experiment.

The translational diffusion coefficient, *D*, is an additive quantity with respect to aggregates present in the system under the fast exchange regime on the NMR timescale. However, it is not an additive quantity with respect to the molecules forming an aggregate and has no relationship to nearest neighbor assumptions. Hence, in theory diffusion NMR data when treated according to either dimer or indefinite models should result in different goodness of fit values depending on whether the system aggregates beyond the dimer stage or not. **Table 3** shows that the dimer model has reliably succeeded for **3**, **8** and for the remaining systems the dimer model appears to be inappropriate. In fact this result highlights which category of aggregation state (dimer or extended aggregate) best matches each of the molecules studied. Although investigation of the dimer-to-indefinite aggregation by NMR diffusometry is a matter of special investigation, the preliminary results obtained in the present work suggest the potential ability of the technique to distinguish between the dimer and indefinite modes of aggregation and resolve the problem of indistinguishability.

Application of the SHIM-model to thermodynamic analysis of aggregation.

A common approach to determine changes in enthalpy, ΔH, and entropy, ΔS, of aggregation is to measure the temperature dependence of an experimental observable and then to fit it to an aggregation model (often the same one used to fit the titration data), in which the self-association constant is substituted with the van't Hoff relation[34,36]

$$K = \exp\left(\frac{\Delta S}{R} - \frac{\Delta H}{RT}\right), \qquad (15)$$

where R is the gas constant.

A similar approach can be used to obtain ΔH, ΔS from the dependence of D_{obs} on temperature by substituting eq 15 into eqs 5-11, 14 for either the shape-dependent models or the SHIM-model. However, for the self-diffusion data, the dependence of D_1 and D_2 on T must also be taken into account.

Let us designate D_1 and D_2 as $D_{1,2}$. Hence, eq 2 takes the form

$$D_{1,2} = \frac{kT}{r(T)}, \tag{16}$$

where $r(T)$ is the temperature-dependent coefficient of friction.

The dependence of r on T is due to the dependence of viscosity, η, on T, allowing eq 16 to be rewritten in the form:

$$D_{1,2} = C_{1,2}\frac{T}{\eta(T)}, \tag{17}$$

where $C_{1,2}$ is a temperature-independent constant.

The viscosity of D_2O depends on T as[13,37]

$$\lg \eta = -4.2911 - \frac{164.97}{174.24 - T} \tag{18}$$

and at T=298 η_{298}=0.0011 $kg \cdot m^{-1} \cdot s^{-1}$.

As long as the exact magnitudes of D_1 and D_2 are available from the analysis of titration data at fixed temperature (in the present work at T = 298 K, or 333 K for **6**), see above), i.e. $D_{1,2}^{(298)}$ is known, so the expression for $D_{1,2}$ at any temperature can be written as

$$D_{1,2} = D_{1,2}^{(298)} \cdot \frac{\eta_{298}}{\eta(T)} \cdot \frac{T}{298} = 3.691 \cdot 10^{-6} \cdot D_{1,2}^{(298)} \cdot \frac{T}{\eta(T)}. \tag{19}$$

It follows that the algorithm for obtaining thermodynamic parameters from self-diffusion data should occur by fitting the $D_{obs}(T)$ curve with the selected model (eqs 5-11, 14) in which the parameters K, D_1 and D_2 are replaced with eq 15 and eq 19. There are only two parameters in such an approach, *viz*. ΔH and ΔS, although in practice additional small variation of $D_{1,2}^{(298)}$ may also be introduced.

Equation 19 may be independently tested for appropriateness against the tetramethylammonium, used as a reference in all NMR experiments in the present work. If eq 19 is correct and if TMA does not complex with other species present in solution (a common assumption in

NMR), the temperature-dependent diffusion, $D_{obs}(T)$, for the TMA signal must be fitted with eq 19 with good quality having just one adjustable parameter, $D_{1,2}^{(298)}$. **Figure 3** shows the experimental $D_{obs}(T)$ curves for TMA in the self-aggregation studies for the two selected compounds **2** and **4**. The goodness of fit in all cases was not worse than R^2=0.99 indicating that eq 19 is appropriate in thermodynamic analyses using self-diffusion data.

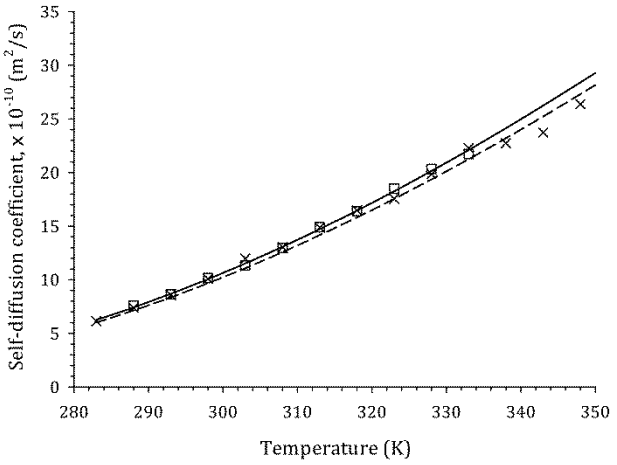


Fig. 3 Experimental $D_{obs}(T)$ curves for TMA in the self-aggregation studies and their fitting curves for **2**, EB (□ fitted with solid line) and **4**, PF (× fitted with dashed line)

Thermodynamic analysis of aggregation based on self-diffusion data has been performed in the present work taking as examples different structured compounds **1**, **2**, **3**, and **4** which have been thoroughly characterized previously in terms of the enthalpy and entropy of aggregation (for reviews see refs. 17, 34, 38). Experimental measurements as well as the numerical analysis were performed against two datasets namely $\delta(T)$ and $D_{obs}(T)$ measured in parallel for similar solutions. The computation of ΔH, ΔS from $\delta(T)$ was accomplished by using eq 11, and from $D_{obs}(T)$ by using eq 13 of the SHIM-model. The results are shown in **Table 4**. Good correspondence can be seen between the diffusion, ^{1}H chemical shift and literature data suggesting that NMR diffusometry with the SHIM-model can be used in thermodynamic analyses of aggregation phenomena.

Table 4 Changes in enthalpy ($kJ \cdot mol^{-1}$) and entropy ($J \cdot mol^{-1} \cdot K^{-1}$) upon aggregation

Data	**1**		**2**		**3**		**4**	
	$\Delta H°$	$\Delta S°$	$\Delta H°$	$\Delta S°$	$\Delta H°$	$\Delta S°$	$\Delta H°$	$\Delta S°$
^{1}H, $\delta(T)$	–31	–0.08	–26	–40	–25	–63	–38	–73
Diffusion, $D_{obs}(T)$	–40	–0.04	–29	–50	–21	–46	–41	–74
Literature[17,34,38]	–40	–0.06	–23	–31	–21	–50	–46	–101

Experimental Section

Chemicals

1 (4-(2'-(4-hydroxyphenyl)-1*H*,3'*H*-[2,5'-bibenzo[*d*]imidazol]-6-yl)-1-methylpiperazin-1-ium chloride, Hoechst 33258, purchased from Sigma-Aldrich), **2** (3,8-diamino-5-ethyl-6-phenylphenanthridin-5-ium bromide, ethidium bromide (EB) purchased from Sigma-Aldrich), **3** (1,3,7-trimethyl-1*H*-purine-2,6(3*H*,7*H*)-dione, caffeine (CAF) purchased from Sigma-Aldrich), **4** (acridine-3,6-diamine, proflavine (PF), purchased from Sigma-Aldrich), **5** (sodium 7-amino-4-hydroxy-3-((*E*)-(2-sulfonato-4-((*E*)-(4-sulfonatophenyl) diazenyl)phenyl)diazenyl)naphthalene-2-sulfonate, supplied as a gift), **6** (*N*-[5-({[5-({[4-({[3-(dimethylamino)propyl]amino}carbonyl)-5-isopropyl-1,3-thiazol-2-yl]amino}carbonyl)-1-methyl-1*H*-pyrrol-3-yl]amino}carbonyl)-1-methyl-1*H*-pyrrol-3-yl]-2-quinoxalinecarboxamide trifluoroacetate – AIK-18/52, supplied as a gift), **7** (*N*-(5-amino-9*H*-benzo[*a*]phenoxazin-9-ylidene)-*N*-ethylethanaminium chloride, Nile Blue (NB) – C. I. Basic Blue 12 purchased from Sigma-Aldrich) and **8** (sodium 1-amino-9,10-dioxo-4-((3-((2-((2-sulfonatoethyl)amino)ethyl)sulfonyl)phenyl)amino)-9,10-dihydroanthracene-2-sulfonate, supplied as a

499 gift) (**Figure 1**) were acquired and used without further purification. D_2O was supplied by Sigma-
500 Aldrich. Samples were prepared by making suitably concentrated stock solutions in D_2O and these then
501 used as the basis to create serially diluted samples for study by NMR spectroscopy. Measurements
502 were made by diluting samples within their NMR tubes to avoid issues encountered from experience
503 when samples are divided or when separate samples are used to generate a series of concentration-
504 dependent NMR data. Sample concentrations in each case are shown in the Supplementary
505 Information.
506
507 *NMR measurements.*
508 NMR spectra were acquired at a magnetic field strength of 14.1 Tesla using a Bruker Avance
509 II+ NMR spectrometer operating at a ^1H resonance frequency of 600.13 MHz and working under
510 TopSpin version 2.1 (Bruker Biospin, Karlsruhe, Germany) on an HP XW3300 workstation running
511 Windows XP. Typically all NMR spectra were acquired on the prepared samples using a broadband
512 observe probe-head equipped with a z-pulsed field gradient coil [BBO-z-atm].
513 1D ^1H NMR spectra were acquired over a frequency width of 12.3 kHz (20.55 ppm) centered at
514 a frequency offset equivalent to 6.175 ppm into 65536 data points during an acquisition time aq = 2.66
515 s with a relaxation delay d1 = 2 s for each of 32 transients. The assignment of proton signals was
516 accomplished with the aid of 2D heteronuclear [^1H, ^{13}C] HSQC and HMBC NMR data and 2D
517 homonuclear [^1H, ^1H] COSY, TOCSY and NOESY NMR data. All measurements have been
518 performed under the fast exchange regime on the NMR chemical shift timescale at T = 298 K with the
519 exception of specific variable temperature measurements, which were performed over a range of
520 temperatures from 278 K to 343 K. Chemical shifts were measured relative to an internal reference of
521 tetramethylammonium bromide (TMA) and recalculated with respect to (sodium 2,2 dimethyl 2-
522 silapentane-5-sulphonate, (DSS) according to $\delta_{DSS} = \delta_{TMA} + 3.178$ (ppm).

Diffusion measurements were carried out as previously described[18] using a bipolar gradient pulse program (Bruker pulse program ledbpgppr2s) in which presaturation was used to suppress residual solvent signal during the recycle delay. Typically 32 gradient increments were used by which the gradient strength was varied linearly in the range 2% to 95% of full gradient strength (54 G/cm with a rectangular gradient) using a sine-shaped gradient profile. Typically the gradient pulse duration was set to 1 ms and the diffusion period to 200 ms. With increasingly dilute samples, the number of transients was increased accordingly in order to allow for diffusion coefficients to be evaluated with a reasonable fit of the experimental data to theory (i.e. number of transients (ns) per FID varied in the range $32 \leq ns \leq 256$ for sample concentrations in the maximal range from 31 mM to 100 μM). Diffusion data were processed under TopSpin (version 2.1, Bruker Biospin) using the T_1/T_2 analysis module in order to fit the data to the standard expression of diffusion coefficient as a function of gradient strength.

Molecular modeling.

All simulations were performed using GROMACS 4.5.5 molecular dynamics package[19,20] with the GROMOS 53a6 force field.[21] The SPC water model was used with the bond lengths constrained by means of the SETTLE algorithm.[22] All other bonds were constrained using the LINCS[23] algorithm. Heavy water (D_2O) was simulated by doubling the masses of hydrogen atoms in the standard SPC water topology. An NVT ensemble was used. The temperature of 298 K was maintained by coupling the system to v-rescale thermostats with a relaxation time of 0.1 ps. Coulomb interactions were computed explicitly within a 1 nm cut-off range, while the Lennard-Jones interactions were computed within a 1.4 nm cut-off range. Long-range electrostatic interactions were computed using the PME method[20] with a grid spacing of 0.12 nm. A simulation step of 1 fs was used.

Topologies of the studied molecules were generated with the Automatic Topology Builder (ATB) server.[24] The charges associated with **2**, ethidium bromide, **3**, caffeine and **4**, proflavine were computed in the course of ATB topology generation on the B3LYP/6-31G* level of theory using ESP fitting of the Merz-Kollman charges. The dimers were constructed manually by positioning the planar ring systems of the monomer at a distance of 0.3 nm from each other and orientating any protruding chemical groups outside the center of the dimer. In the case of charged solutes, the necessary number of chloride counter ions was added to neutralize the system.

Six independent simulations of 2 ns each were performed for each system. Velocities of all atoms in the system were saved every 10 fs. Following this, the diffusion coefficients were computed using the Green-Kubo relations from velocity autocorrelation functions of the center of masses of solutes.[25] The recommended procedure for computing diffusion coefficients within the GROMACS software package was used.[1] The diffusion coefficients obtained from six independent runs were averaged.

Numerical analysis.

All computations were made in such a way that all models were subjected to similar input conditions, such as guess points, without any other restraints being introduced specifically to a particular model. The guess points were generated randomly within 10% variation of ^{1}H NMR- derived K and expected from $D(x_0)$ curve values of D_1 and D_2. We used MATLAB software in order to perform discrepancy (Δ) minimization. In order to ensure that the resultant minimum was reliable, we used three different algorithms of minimization incorporated in MATLAB, *viz*. 'trust-region dogleg', 'Gauss-Newton' and 'Levenberg-Marquardt'. The results of minimizations in MATLAB were also

[1] see http://www.gromacs.org/Documentation/How-tos/Diffusion_Constant

568 independently verified by performing calculations by means of alternative procedures used previously
569 in the analysis of large sets of self- and hetero-associations.[26]
570
571 **Associated Content – Supporting Information**
572 Graphs of concentration- and temperature-dependence of ^{1}H chemical shifts and concentration- and
573 temperature-dependence of self-diffusion coefficients measured by ^{1}H NMR spectroscopy for
574 compounds **1**-**8** (Figures S1-S28); list of model numbers with brief model description for 17 different
575 mathematical models (Table S1); calculated parameters K, D_1, D_2, d and R^2 from each of 17 models
576 tested for compounds **1**-**8** (Tables S2a-S9a); calculated parameter K, D_1, D_2 and R^2 for model number
577 11 tested for compound **1**-**8** following randomized exclusion of 1, 2 or 3 data points (Tables S2b-S9b).
578
579
580 **Conclusion**
581
582 The possibility of using NMR diffusometry for quantification of thermodynamic parameters of
583 aggregation (equilibrium self-association constant, changes in enthalpy and entropy) proceeding
584 beyond the dimer stage is currently very limited due to the necessity for *a priori* knowledge of the
585 hydrodynamic shape of aggregates, which is not always available in practice. In the present work we
586 have investigated the dependence of aggregation parameters on the type of aggregation model selected
587 and, based on this, developed a new shape-independent model (the SHIM-model, equation 13 and
588 expressed in the more convenient form of equation 14 using the hypergeometric function, F). It was
589 found that this approach enables experimental self-diffusion NMR data to be described with the same
590 quality or better (the goodness of fit and the correspondence of the aggregation parameters to a method
591 used as a reference) as compared with the shape-dependent models for the whole set of test compounds

(equations 5-8 in the current work). It is recommended that the SHIM-model be used in cases where the hydrodynamic shape of aggregates is unknown. An algorithm for using the self-diffusion data with the aim of determining enthalpy and entropy of aggregation was also developed. The results of this work open up in particular the possibility of using NMR diffusometry as a general method to study aggregation phenomena in solution.

Acknowledgements

The authors thank Dr. A. I. Khalaf for the gift of compound **6** and Dr. M. G. Hutchings for the gift of compounds **5** and **8**. This work was, in part, supported by Russian Fund for Basic Researches (project no.15-04-03119).

References

(1) Y. Cohen, L. Avram and L. Frish, *Angew. Chem. Int. Ed.* **44**, 520 (2005).

(2) A. Macchioni, G. Ciancaleoni, C. Zuccaccia and D. Zuccaccia, *Chem. Soc. Rev.* **37**, 479 (2008).

(3) J. Hu, T. Xu and Y. Cheng, *Chem. Rev.* **112**, 3856 (2012).

(4) S. Floquet, S. Brun, J.-F. Lemonnier, M. Henry, M.-A. Delsuc, Y. Prigent, E. Cadot and F. Taulelle, *J. Am. Chem. Soc.* **131**, 17254 (2009).

(5) D. Li, G. Kagan, R. Hopson and P. G. Williard, *J. Am. Chem. Soc.* **131**, 5627 (2009).

(6) A. A. Colbourne, G. A. Morris and M. Nilsson, *J. Am. Chem. Soc.* **133**, 7640 (2011).

(7) T. A. Shastry, A. J. Morris-Cohen, E. A. Weiss and M. C. Hersam, *J. Am. Chem. Soc.* **135**, 6750 (2013).

(8) S. L. Mansfield, D. A. Jayawickrama, J. S. Timmons and C. K. Larive, *Biochim. Biophys. Acta* **1382**, 257 (1998).

(9) I. Pianet, Y. Andrè, M.-A. Ducasse, I. Tarascou, J.-C. Lartigue, N. Pinaud, E. Fouquet, E. J. Dufourc and M. Laguerre, *Langmuir* **24**, 11027 (2008).

(10) P. S. Denkova, L. Van Lokeren, I. Verbruggen and R. Willem, *J. Phys. Chem. B* **112**, 10935 (2008).

(11) G. Consiglio, S. Failla, P. Finocchiaro, I. P. Oliveri, R. Purrello and S. Di Bella, *Inorg. Chem.* **49**, 5134 (2010).

(12) M. S. Kaucher, Y.-F. Lam, S. Pieraccini, G. Gottarelli and J. T. Davis, *Chem. Eur. J.* **11**, 164 (2005).

(13) A. Wong, R. Ida, L. Spindler and G. Wu, *J. Am. Chem. Soc.* **127**, 6990 (2005).

(14) C. Cabaleiro-Lago, M. Nilsson, A. J. M. Valente, M. Bonini and O. Söderman, *J. Colloid Interface Sci.* **300**, 782 (2006).

(15) M. P. Renshaw and I. J. Day, *J. Phys. Chem. B* **114**, 10032 (2010).

(16) I. V. Nesmelova and V. D. Fedotov, *Biochim. Biophys. Acta* **1383**, 311 (1998).

(17) N. J. Buurma and I. Haq, *J. Mol. Biol.* **381**, 607 (2008).

(18) D. Hazafy, M.-V. Salvia, A. Mills, M. G. Hutchings, M. P. Evstigneev and J. A. Parkinson, *Dyes Pigm.* **88**, 315 (2011).

(19) B. Hess, C. Kutzner, D. Van der Spoel and E. Lindahl, *J. Chem. Theory Comput.* **4**, 435 (2008).

(20) D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.* **26**, 1701 (2005).

(21) C. Oostenbrink, A. Villa, A. E. Mark and W. F. van Gunsteren, *J. Comput. Chem.* **25**, 1656 (2004).

(22) S. Miyamoto and P. A. Kollman, *J. Comput. Chem.* **13**, 952 (1992).

(23) B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.* **18**, 1463 (1997).

(24) A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink and A. E. Mark, *J. Chem. Theory Comput.* **7**, 4026 (2011).

(25) D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids* (Academic Press, London, 1990).

(26) M. P. Evstigneev, D. B. Davies and A. N. Veselkov, *Chem. Phys.* **321**, 25 (2006).

(27) R. B. Martin, *Chem. Rev.* **96**, 3043 (1996).

(28) M. P. Evstigneev, A. S. Buchelnikov, V. V. Kostjukov, I. S. Pashkova and V. P. Evstigneev, *Supramol. Chem.* **25**, 199 (2013).

(29) W. S. Price, *NMR studies of translational motion* (University Press, Cambridge, 2009).

(30) V. A. Bloomfield, *Survey of biomolecular hydrodynamics;* in *Separations and Hydrodynamics* (ed. T. M. Schuster) (Biophysical Society, 1980). www.biophysics.org

(31) M. P. Evstigneev, A. S. Buchelnikov and V. P. Evstigneev, *Phys. Rev. E* **85**, 061405 (2012).

(32) M. M. Tirado, C. L. Martínez and J. G. de la Torre, *J. Chem. Phys*. **81**, 2047 (1984).

(33) I. A. Kotzé, W. J. Gerber, J. M. McKenzie and K. R. Koch, *Eur. J. Inorg. Chem.* 1626 (2009).

(34) D. B. Davies, L. N. Djimant and A. N. Veselkov, *J. Chem. Soc., Faraday Trans.* **92**, 383 (1996).

(35) L. Tavagnacco, U. Schnupf, P. E. Mason, M.-L. Saboungi, A. Cesàro and J. W. Brady, *J. Phys. Chem. B* **115**, 10957 (2011).

(36) D. B. Davies, D. A. Veselkov, M. P. Evstigneev and A. N. Veselkov, *J. Chem. Soc., Perkin Trans. 2* 61 (2001).

(37) J. Lapham, J. P. Rife, P. B. Moore and D. M. Crothers, *J. Biomol. NMR* **10**, 255 (1997).

(38) D. B. Davies, D. A. Veselkov, L. N. Djimant and A. N. Veselkov, *Eur. Biophys. J.* **30**, 354 (2001).

Shape-Independent Model (SHIM) Approach for Studying Aggregation by NMR Diffusometry

Adrian A. Hernandez Santiago, Anatoly S. Buchelnikov, Maria A. Rubinson, Semen O. Yesylevskyy, John A. Parkinson and Maxim P. Evstigneev

Supplementary Information

Section A - Supplementary Figures

The following figures represent experimental NMR data (filled circles) along with their fits (solid lines). The well-known indefinite self-association model (eq 11 of the article) is used in order to fit the ^{1}H NMR data, namely:

$$\delta(x_0) = \delta_1 + (\delta_2 - \delta_1)\frac{2Kx_0 + 1 - \sqrt{4Kx_0 + 1}}{Kx_0}.$$

^{1}H diffusion NMR data were fitted according to the SHIM-model (eq 13 of the article):

$$D_{\text{obs}} = \frac{x_1}{x_0}\alpha D_1 \sum_{i=0}^{\infty} \frac{i+1}{i+\alpha}(Kx_1)^i \text{, where } \alpha = \frac{D_2}{D_1 - D_2}.$$

^{1}H VT and ^{1}H DOSY VT NMR data were fitted using the above equations in which the equilibrium constant *K* was substituted with the van't Hoff relation (eq 15 of the article):

$$K = \exp\left(\frac{\Delta S}{R} - \frac{\Delta H}{RT}\right).$$

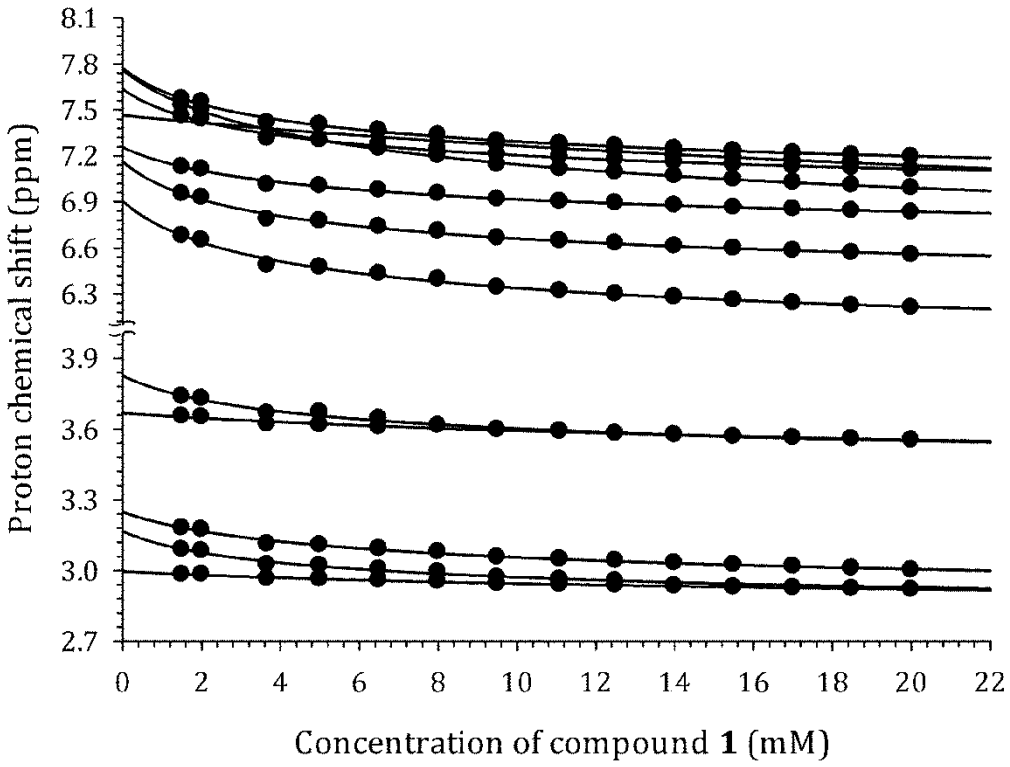


Figure S1: ^{1}H NMR chemical shifts as a function of solute concentration for **1**, Hoechst 33258 measured at T = 298 K.

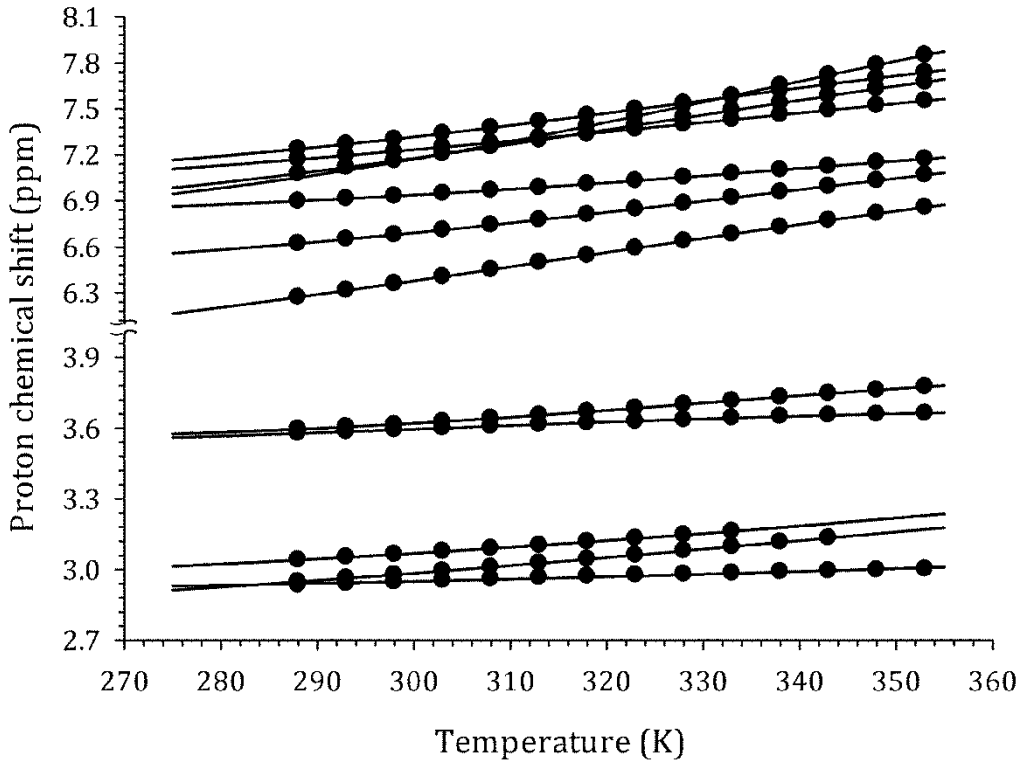


Figure S2: ^{1}H NMR chemical shifts as a function of temperature for **1**, Hoechst 33258, at a solute concentration of 3.5 mM.

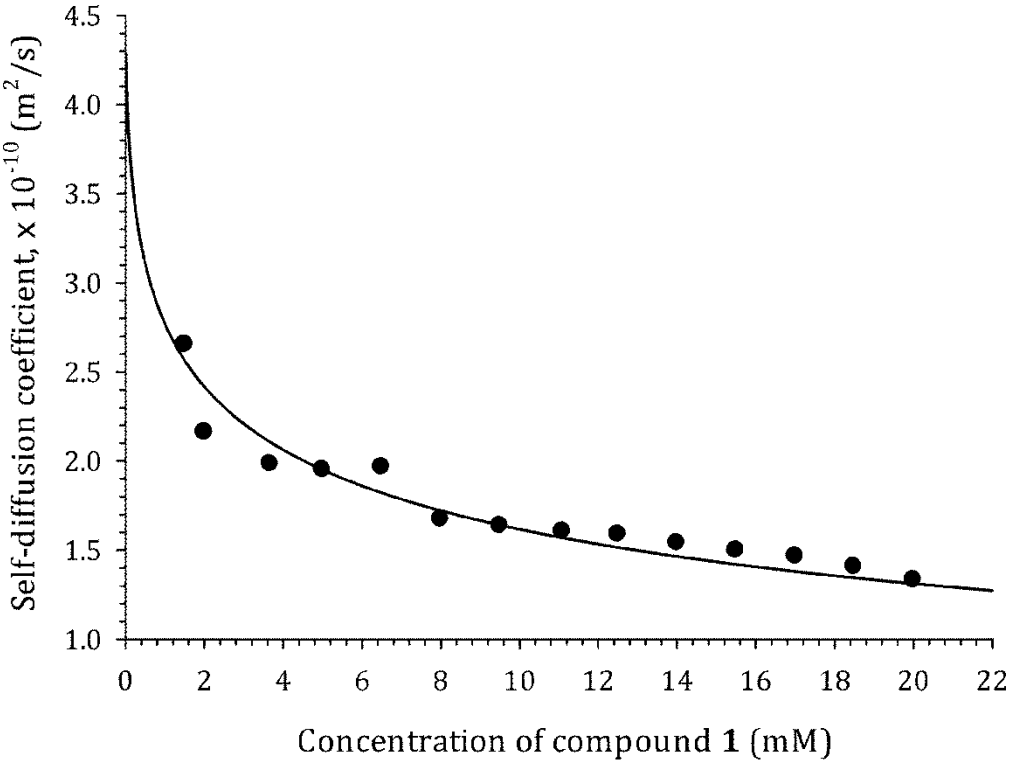


Figure S3: ^{1}H NMR-derived diffusion coefficient as a function of solute concentration for **1**, Hoechst 33258 at T = 298 K.

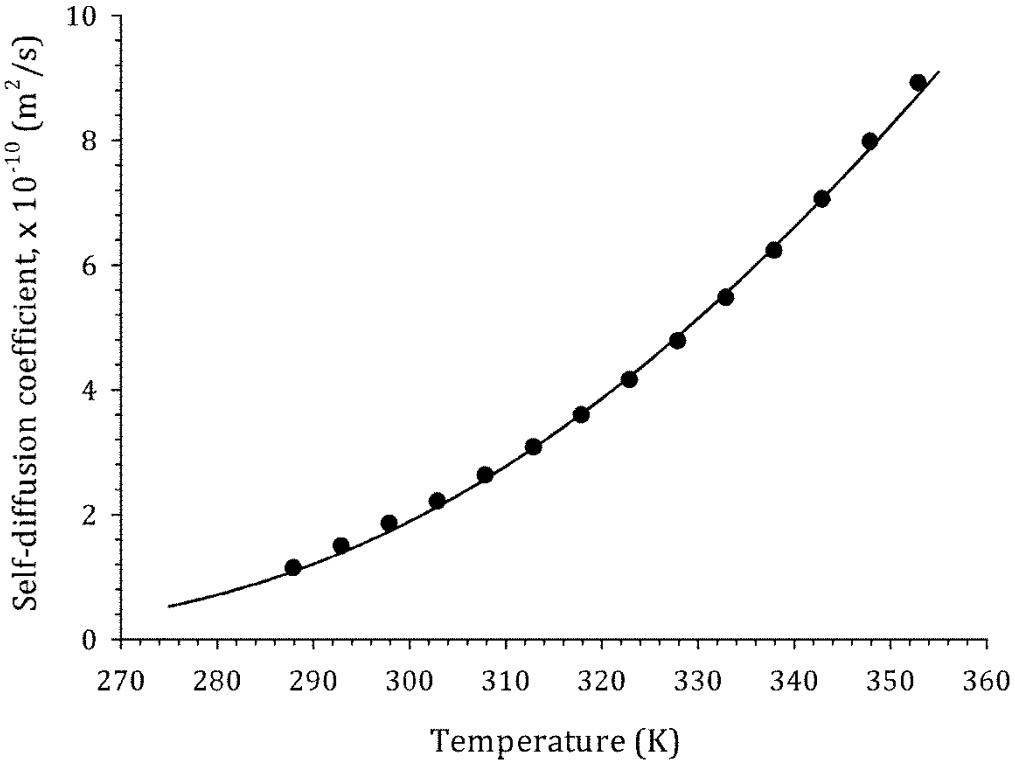


Figure S4: ^{1}H NMR-derived diffusion coefficient as a function of temperature for **1**, Hoechst 33258 at a solute concentration of 3.5 mM.

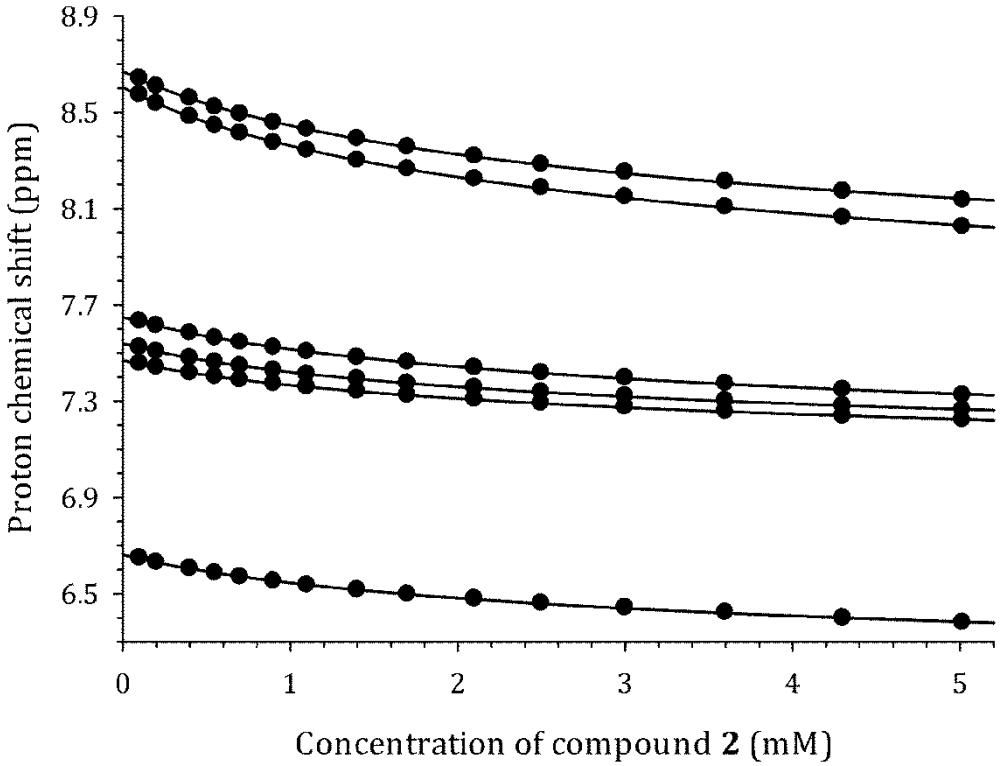


Figure S5: ^{1}H NMR chemical shifts as a function of solute concentration for **2**, Ethidium Bromide, measured at T = 298 K.

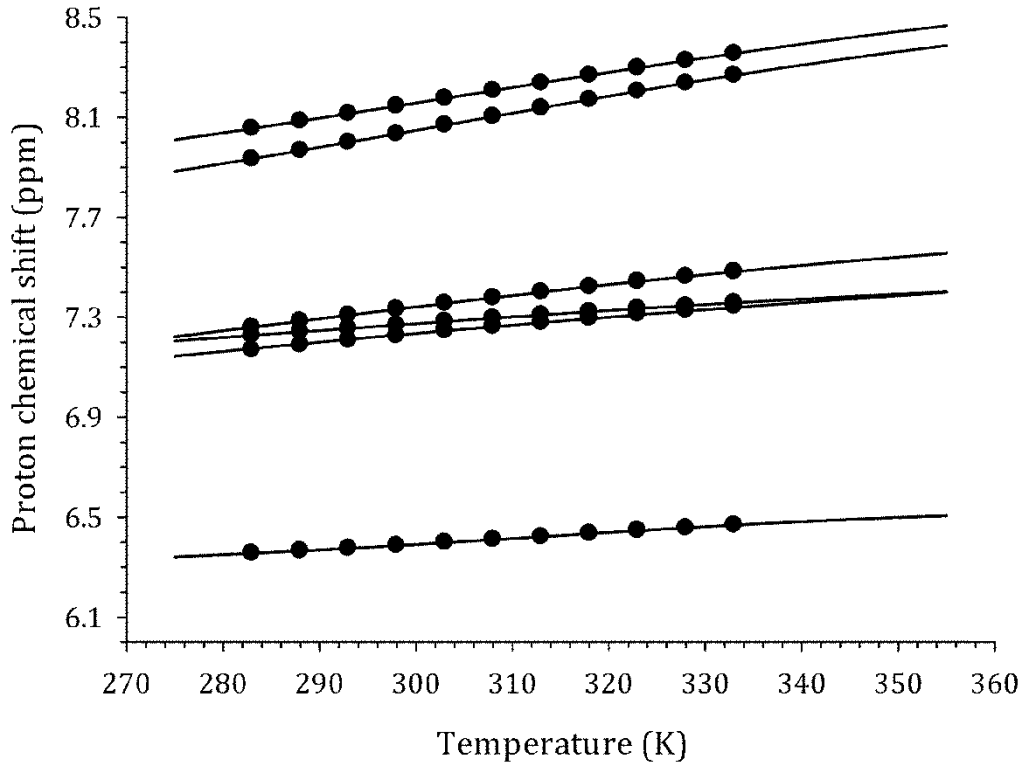


Figure S6: ^{1}H NMR chemical shifts as a function of temperature for **2**, Ethidium Bromide, at a solute concentration of 3.0 mM.

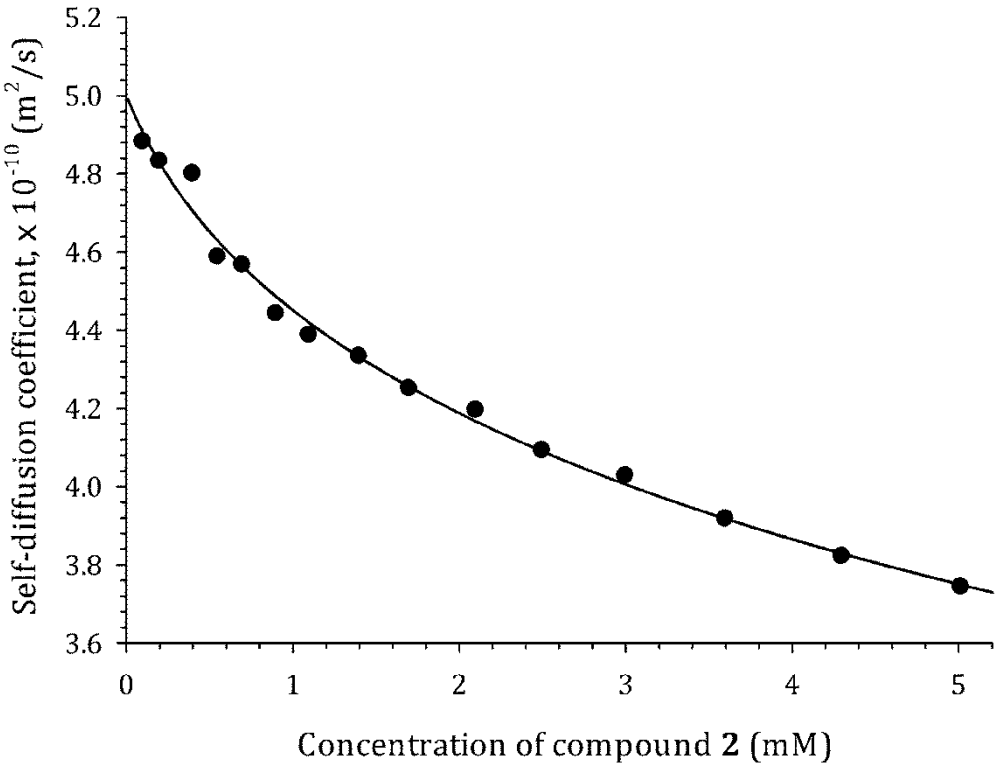


Figure S7: ^{1}H NMR-derived diffusion coefficient as a function of solute concentration for **2**, Ethidium Bromide, at T = 298 K.

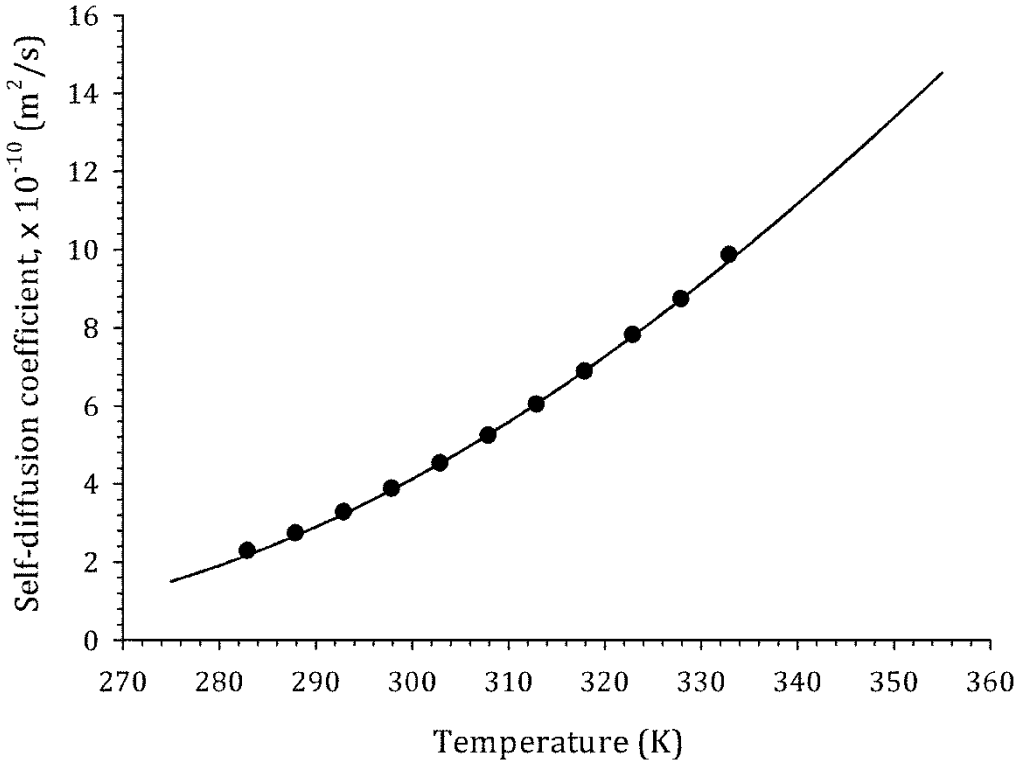


Figure S8: ^{1}H NMR-derived diffusion coefficient as a function of temperature for **2**, Ethidium Bromide, at a solute concentration of 3.0 mM.

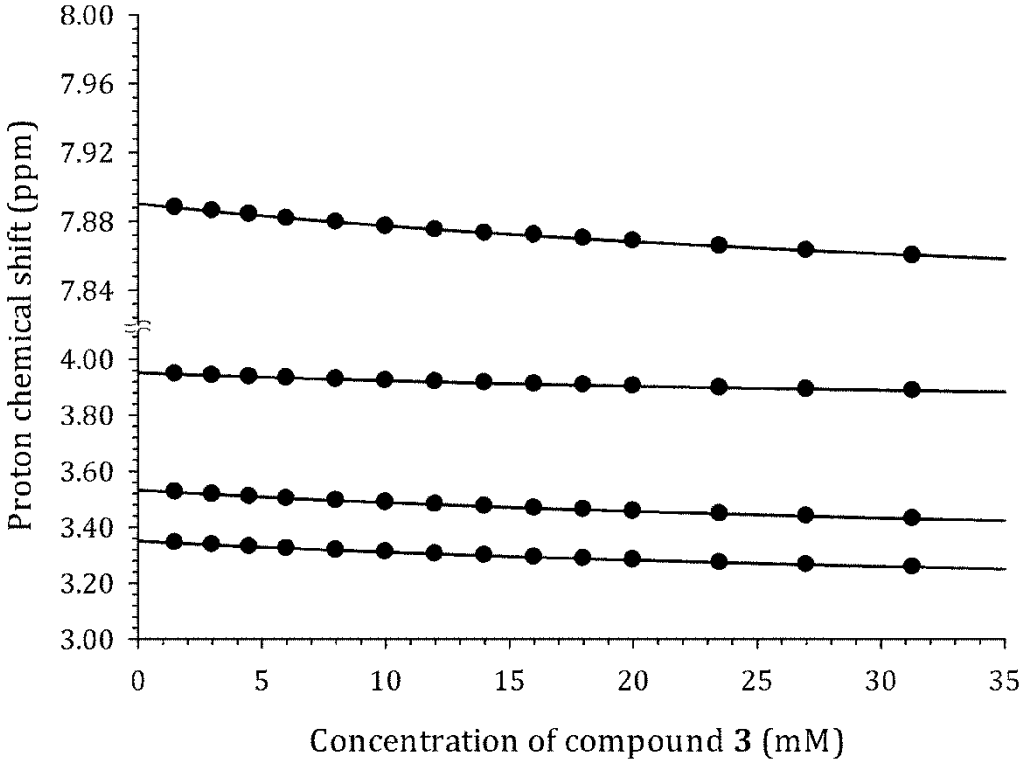


Figure S9: ^{1}H NMR chemical shifts as a function of solute concentration for **3**, Caffeine, measured at T = 298 K.

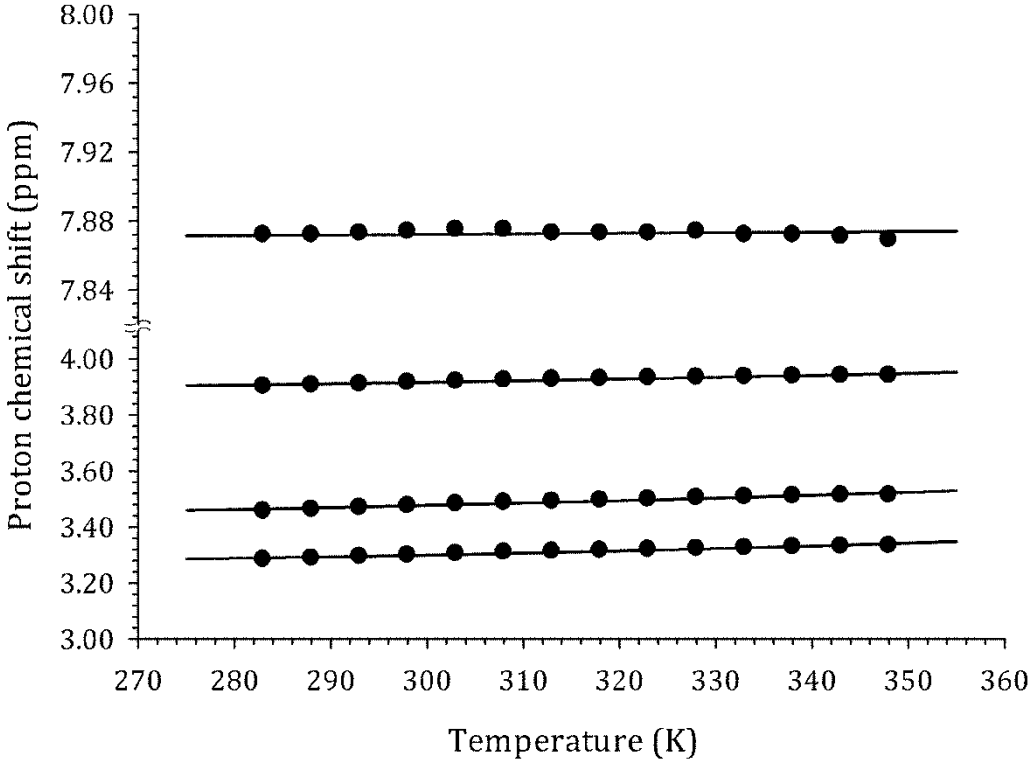


Figure S10: ^{1}H NMR chemical shifts as a function of temperature for **3**, Caffeine, at a solute concentration of 20.0 mM.

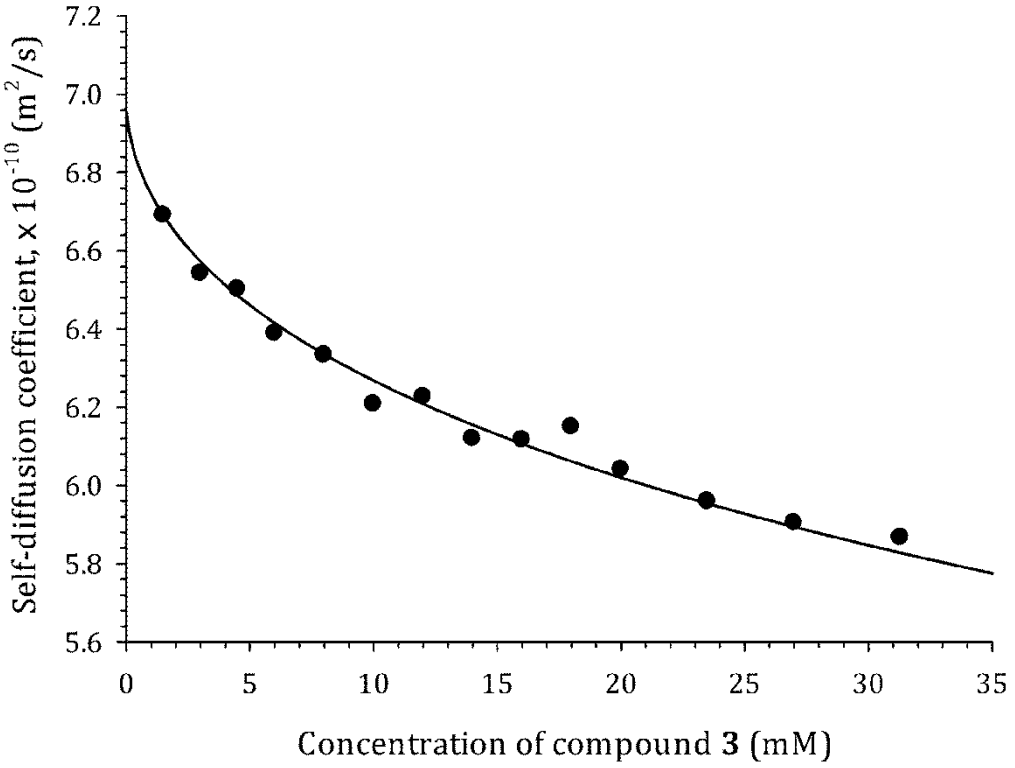


Figure S11: ^{1}H NMR-derived diffusion coefficient as a function of solute concentration for **3**, Caffeine, at T = 298 K.

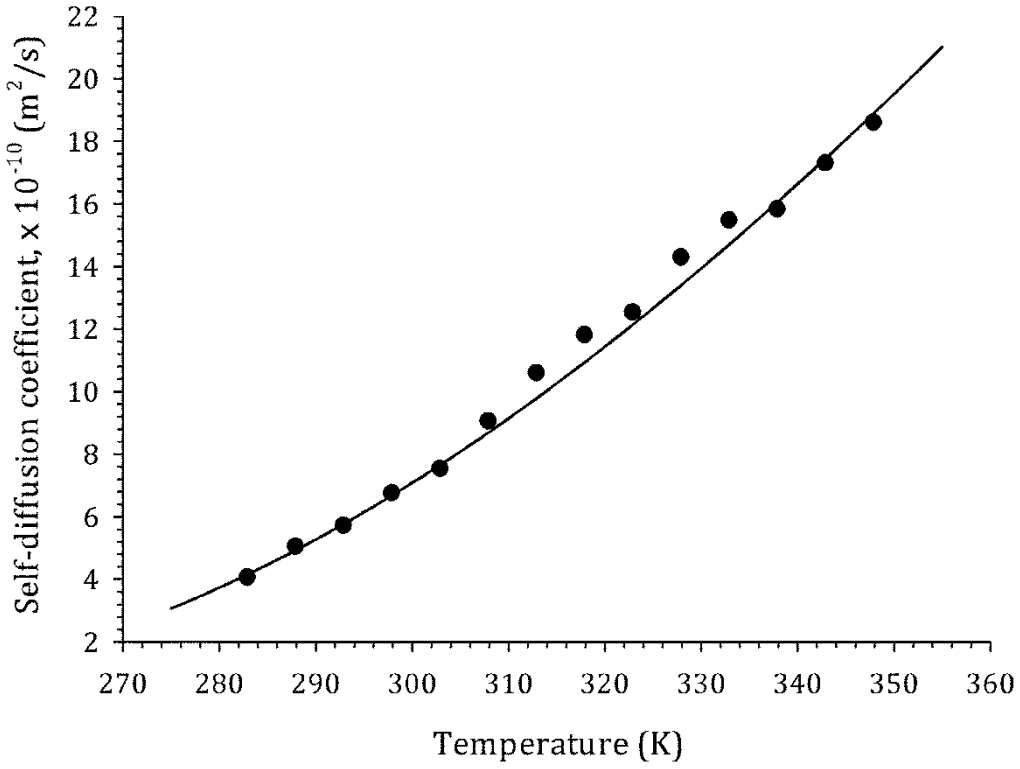


Figure S12: ^{1}H NMR-derived diffusion coefficient as a function of temperature for **3**, Caffeine, at a solute concentration of 20.0 mM.

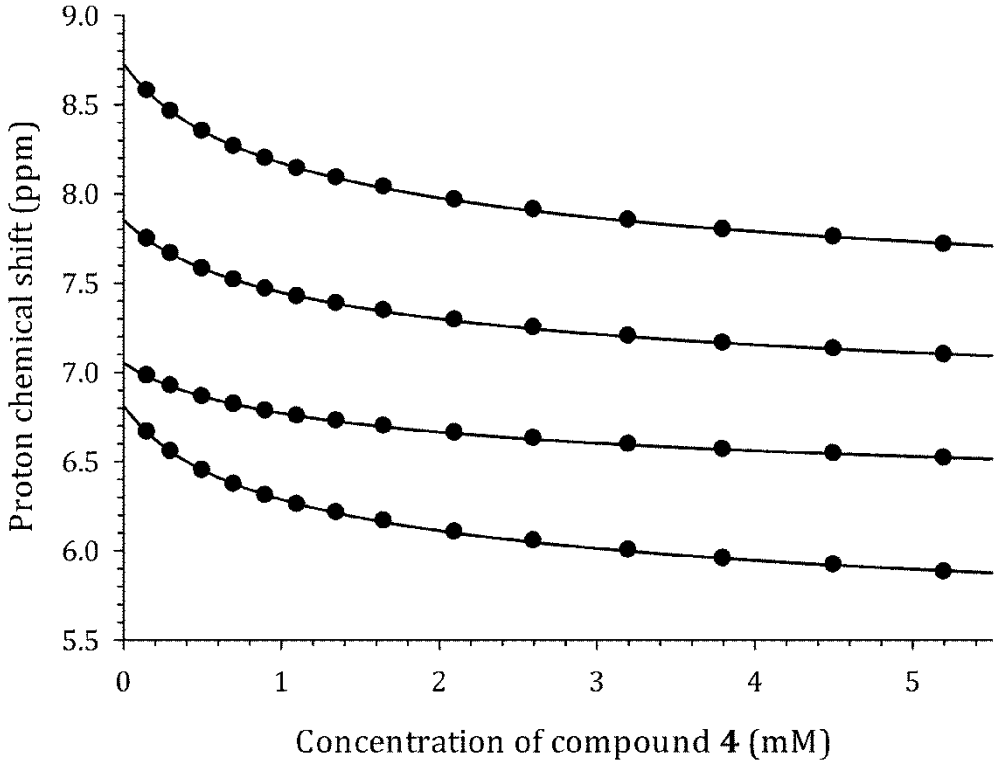


Figure S13: ^{1}H NMR chemical shifts as a function of solute concentration for **4**, Proflavine, measured at T = 298 K.

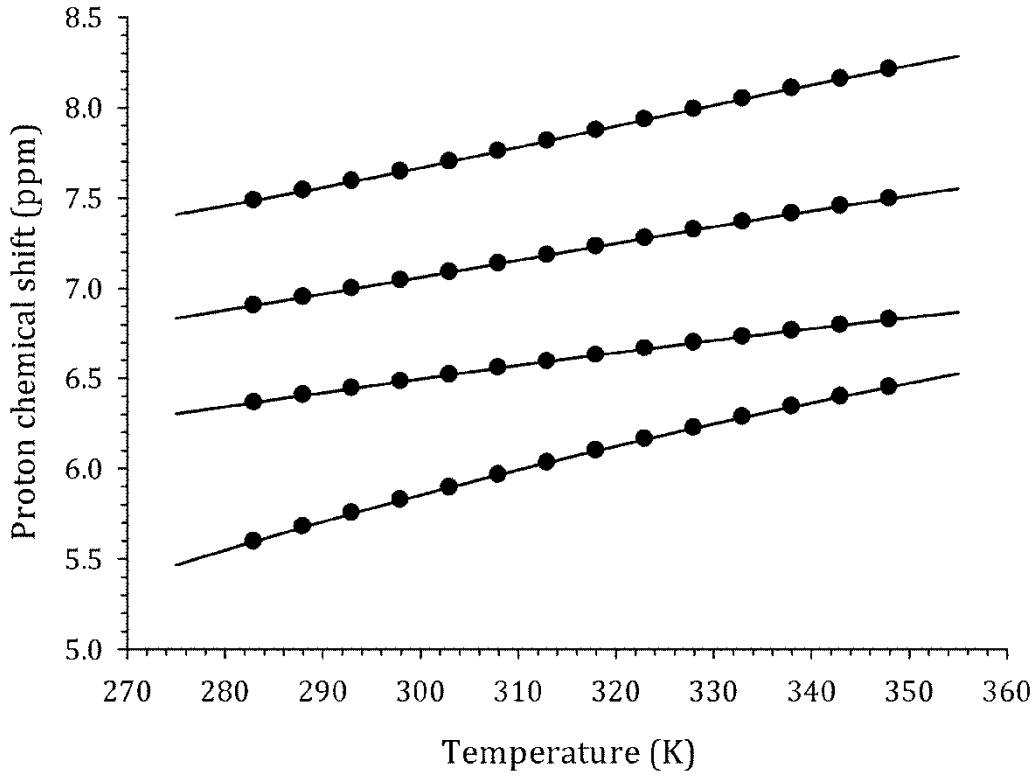


Figure S14: ^{1}H NMR chemical shifts as a function of temperature for **4**, Proflavine, at a solute concentration of 4.5 mM.

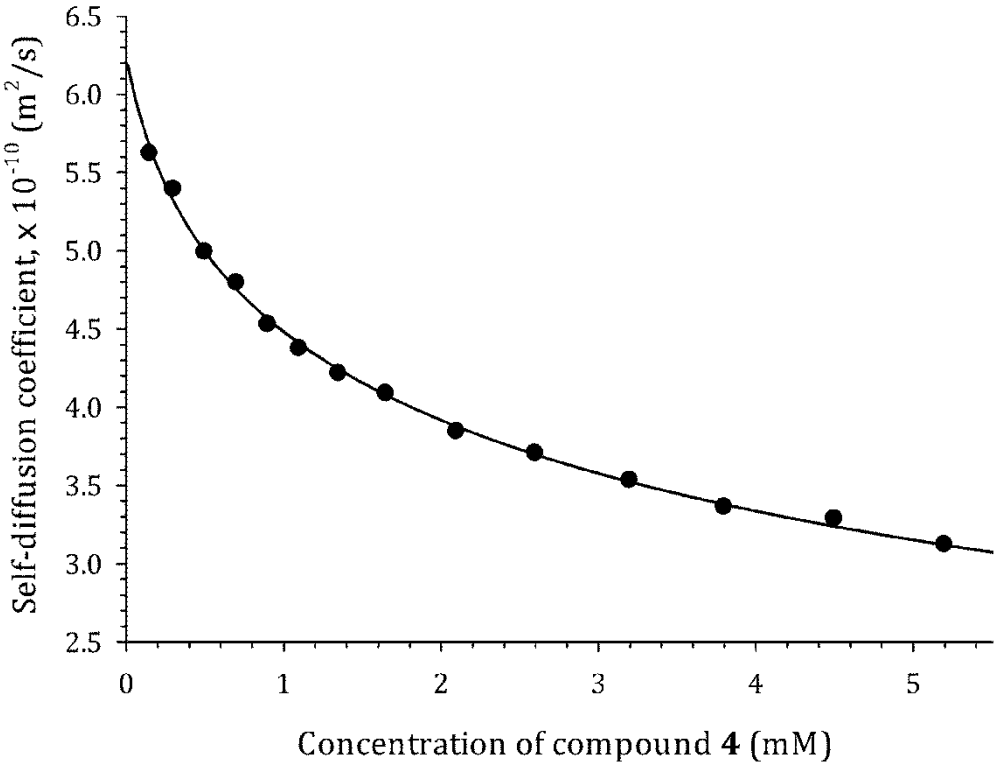


Figure S15: ^{1}H NMR-derived diffusion coefficient as a function of solute concentration for **4**, Proflavine, at T = 298 K.

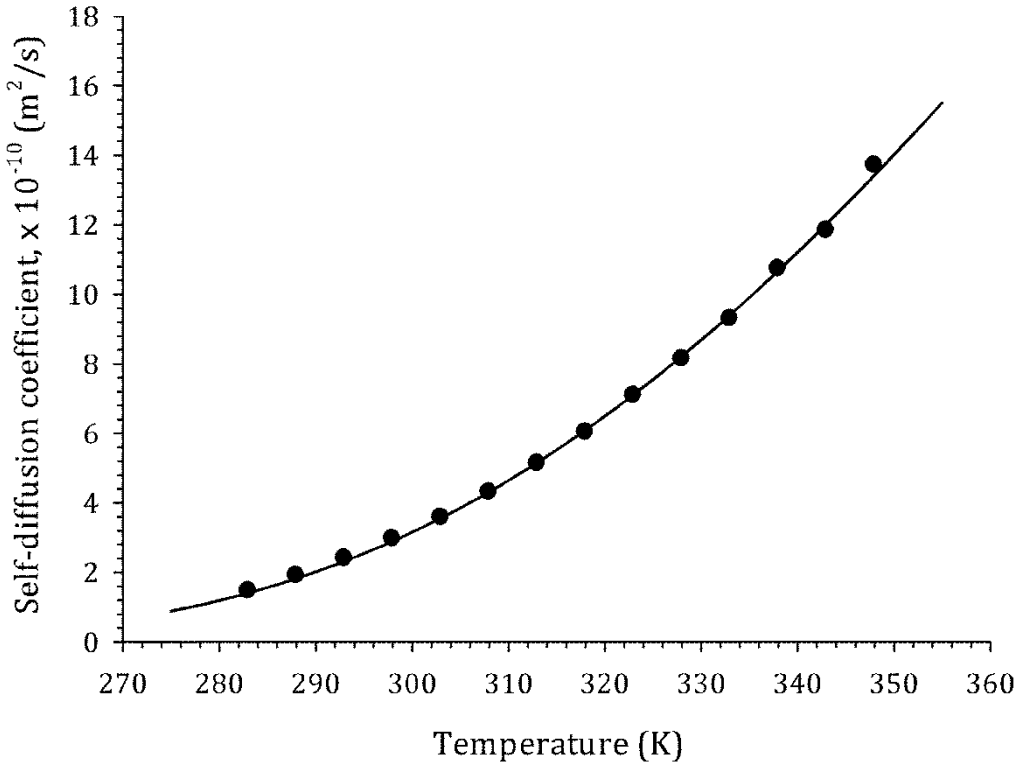


Figure S16: ^{1}H NMR-derived diffusion coefficient as a function of temperature for **4**, Proflavine, at a solute concentration of 4.5 mM.

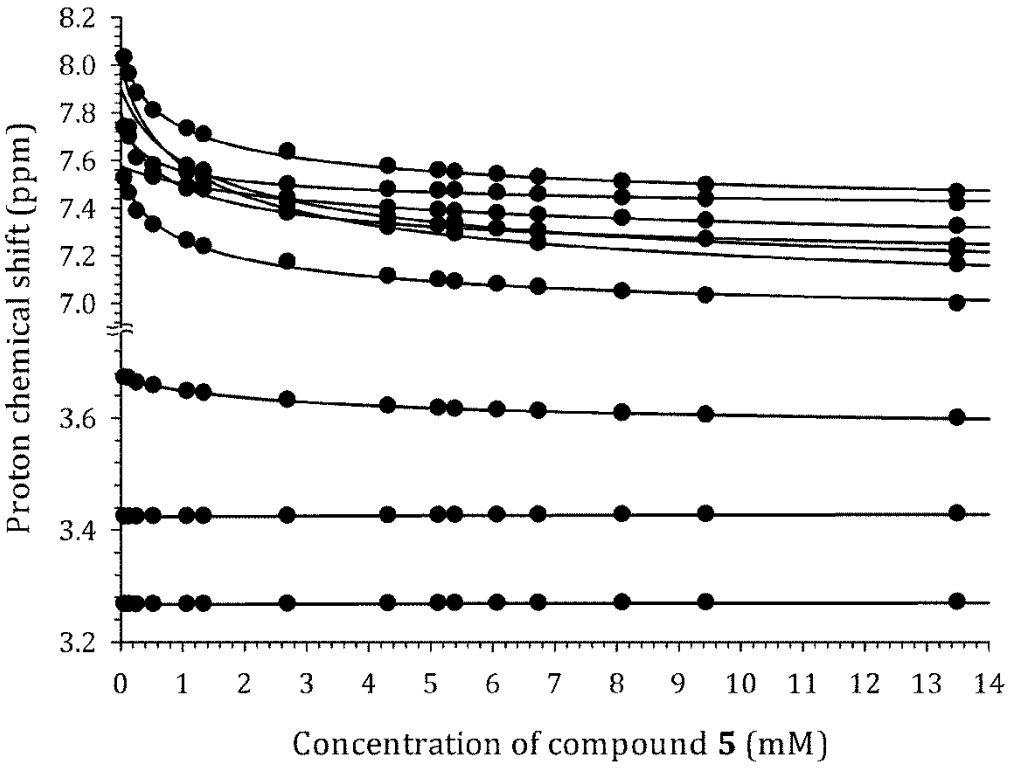


Figure S17: ^{1}H NMR chemical shifts as a function of solute concentration for **5** measured at T = 298 K.

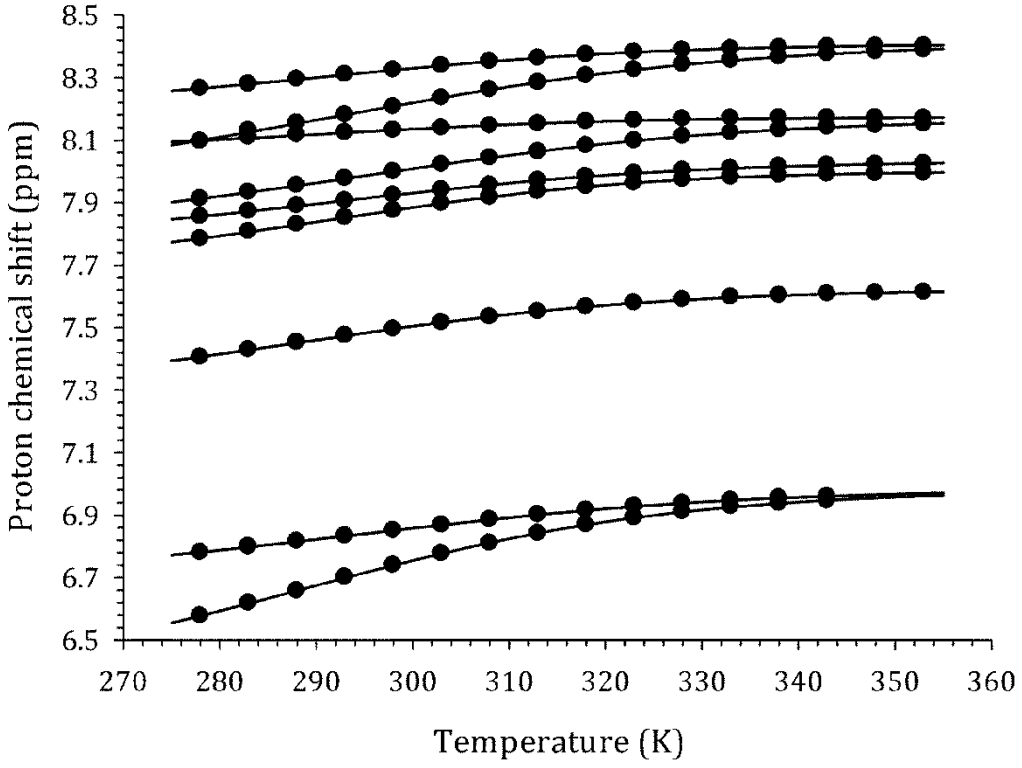


Figure S18: ^{1}H NMR chemical shifts as a function of temperature for **5**.

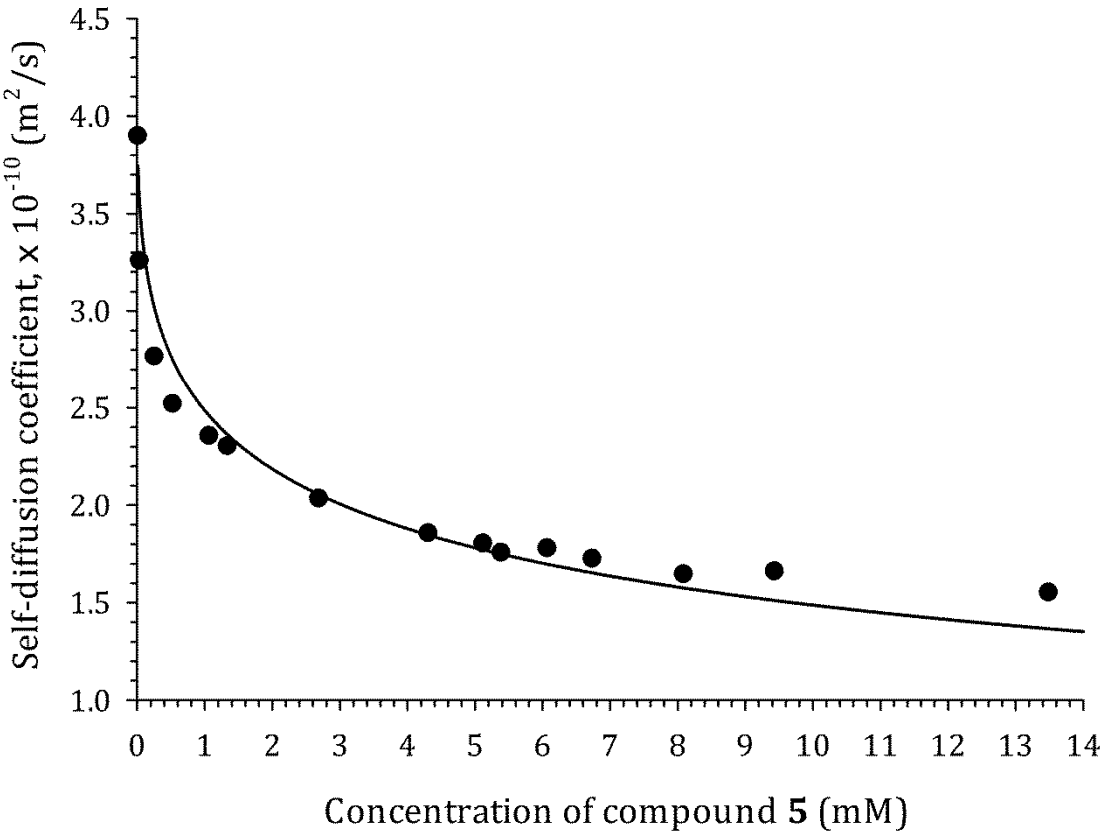


Figure S19: ^{1}H NMR-derived diffusion coefficient as a function of solute concentration for **5** at T = 298 K.

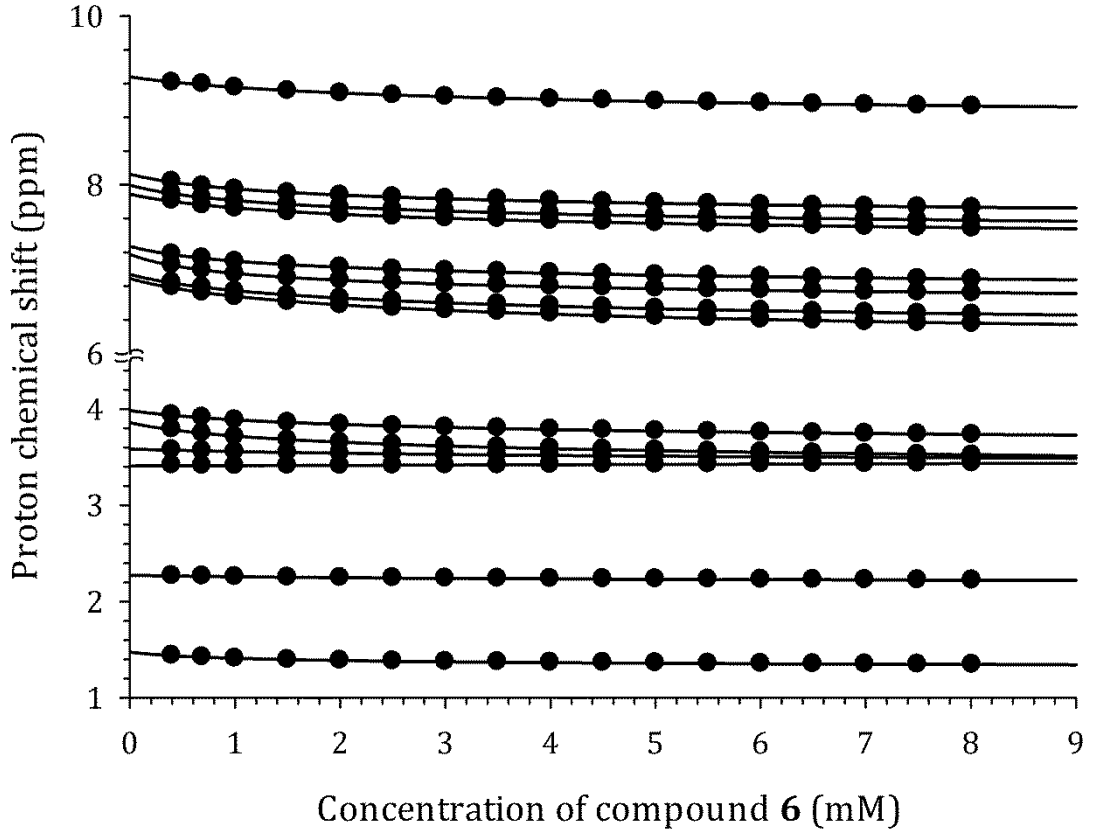


Figure S20: ^{1}H NMR chemical shifts as a function of solute concentration for **6**, AIK-18/52, at T = 298 K.

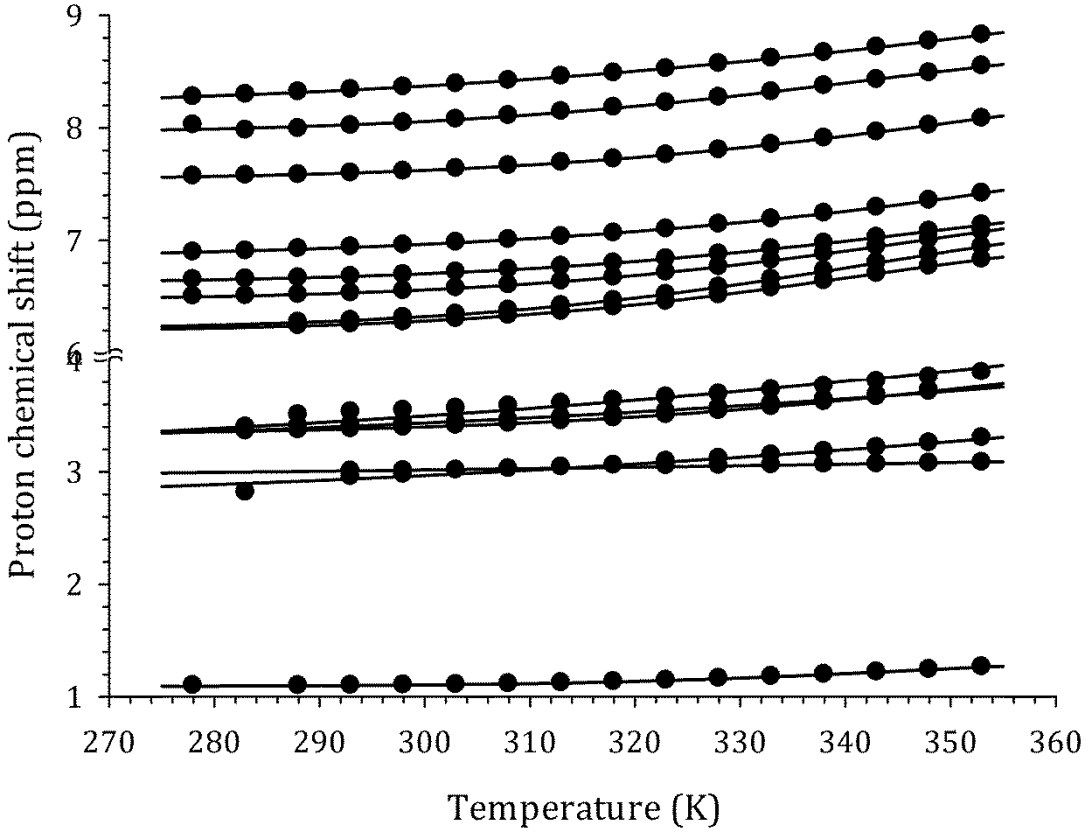


Figure S21: ^{1}H NMR chemical shift as a function of temperature for **6**, AIK-18/52.

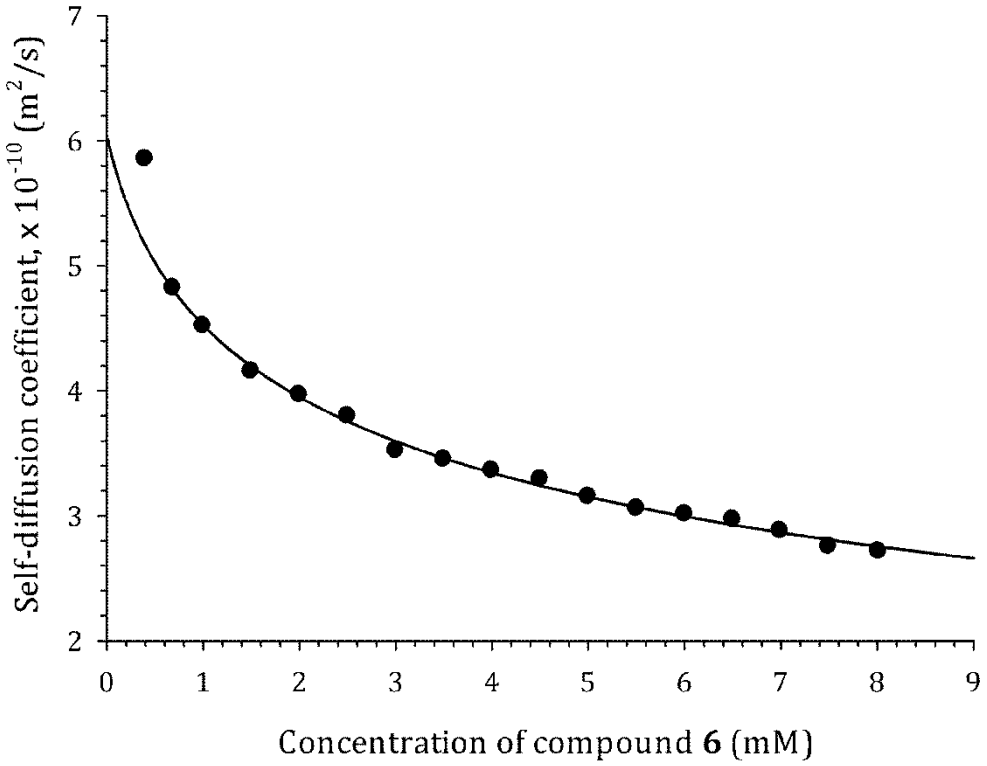


Figure S22: ^{1}H NMR-derived diffusion coefficient as a function of concentration for **6**, AIK-18/51.

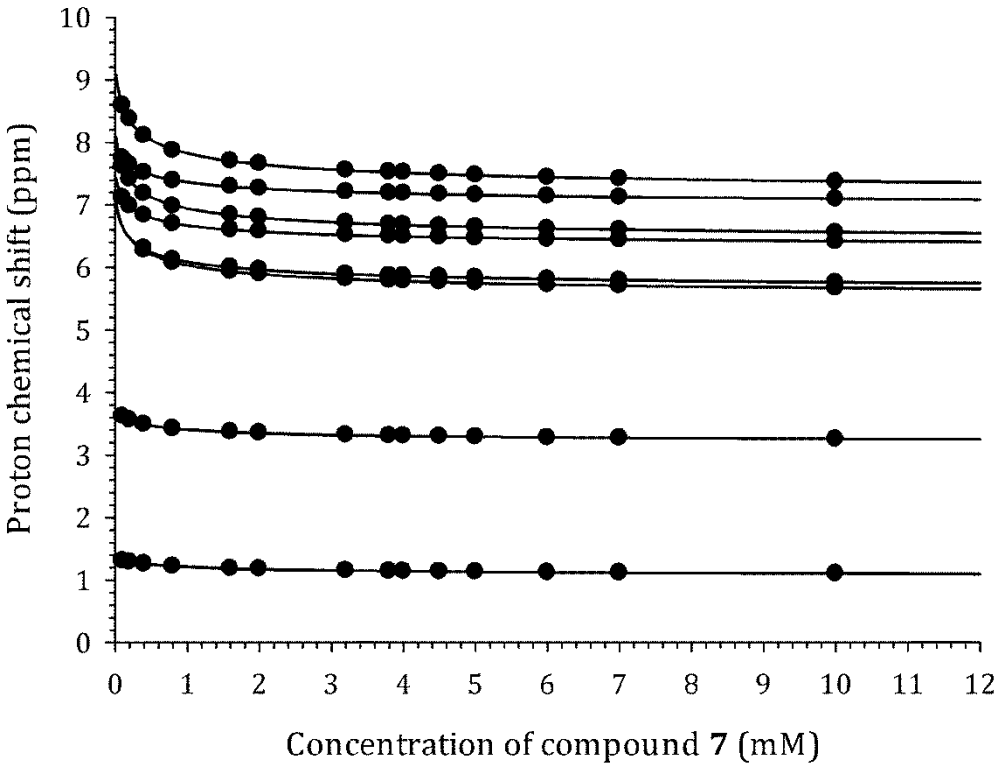


Figure S23: ^{1}H NMR chemical shift as a function of concentration for **7**, Nile Blue (C. I. Basic Blue 12).

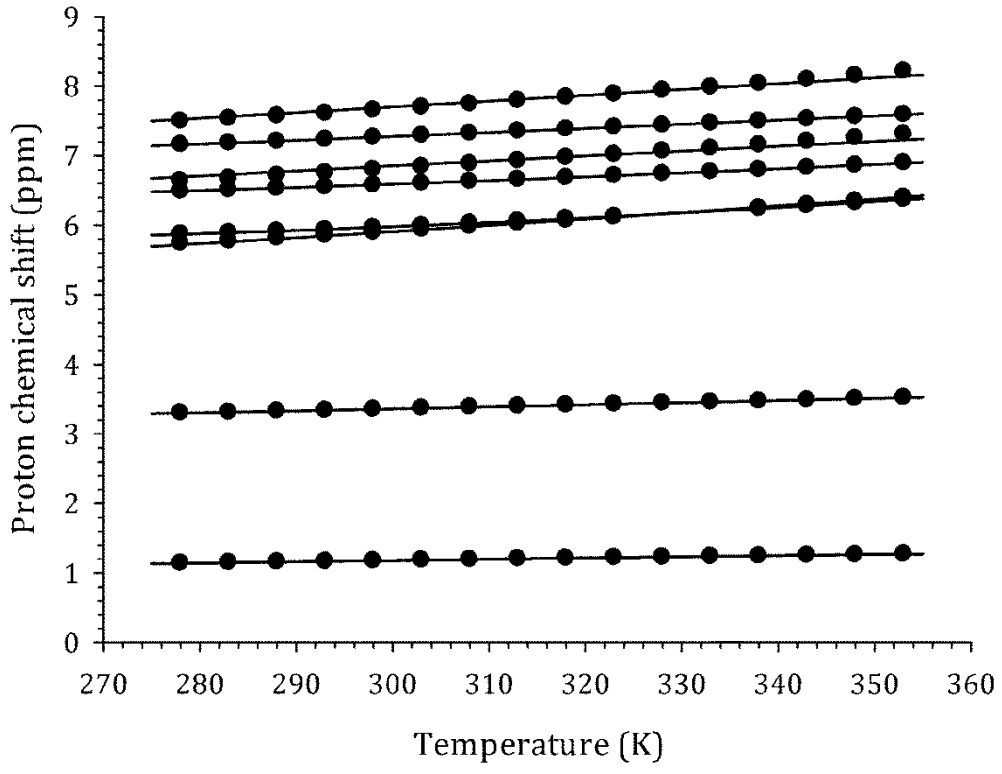


Figure S24: ^{1}H NMR chemical shift as a function of temperature for **7**, Nile Blue (C. I. Basic Blue 12).

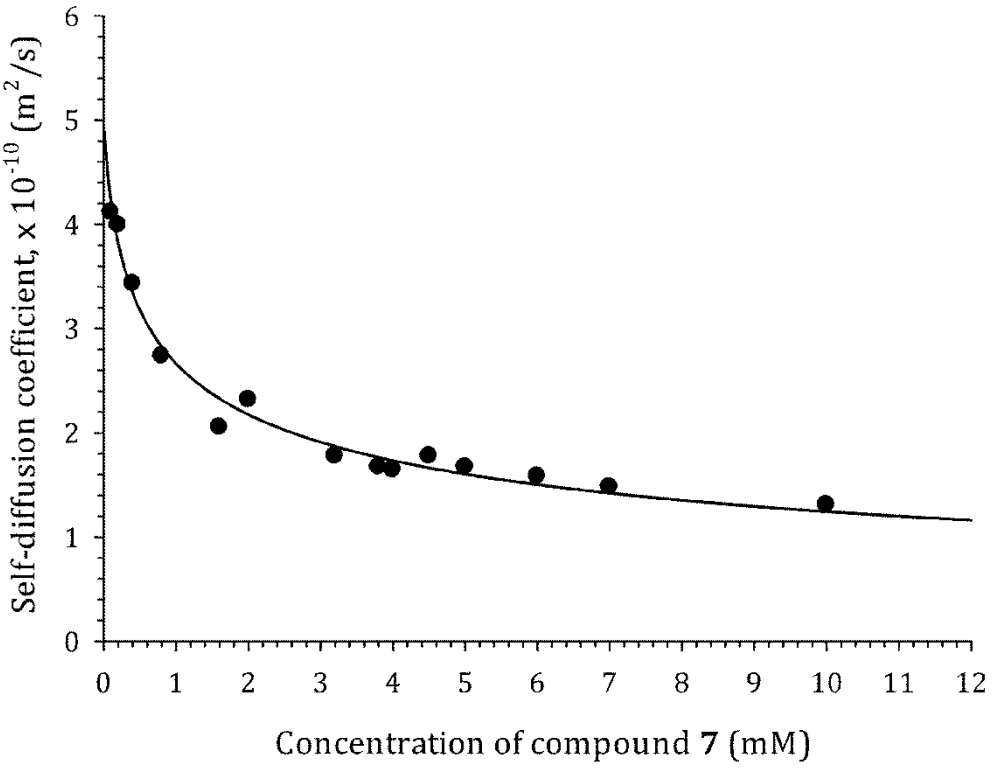


Figure S25: ^{1}H NMR-derived diffusion coefficient as a function of temperature for **7**, Nile Blue (C. I. Basic Blue 12).

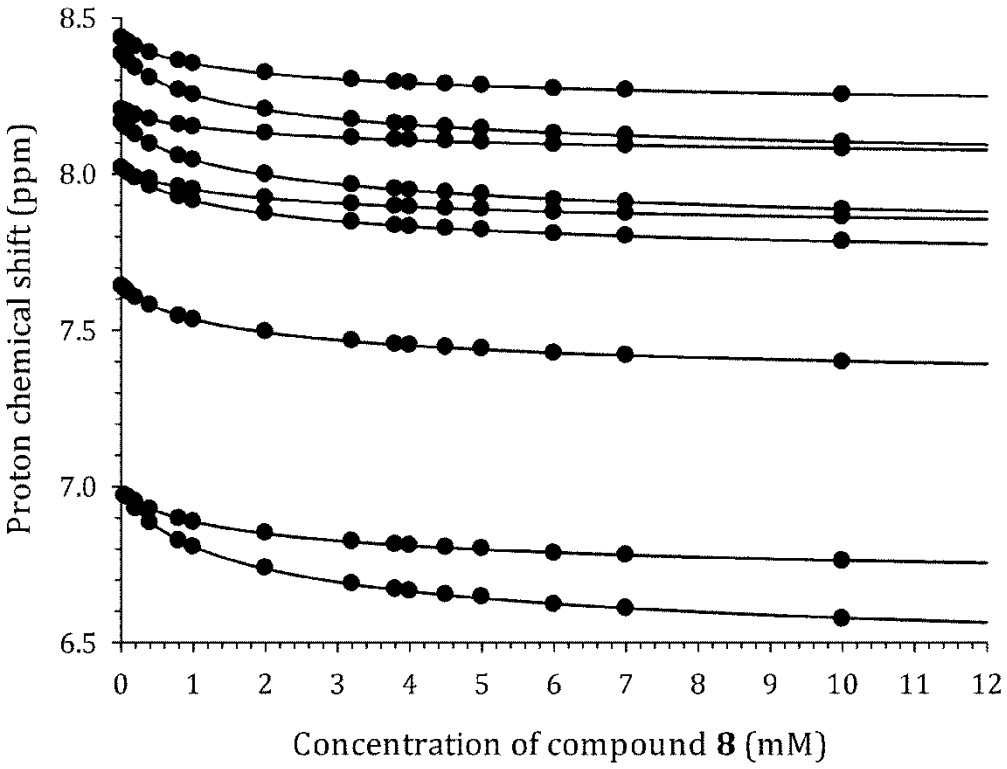


Figure S26: ^{1}H NMR chemical shift as a function of concentration for **8**.

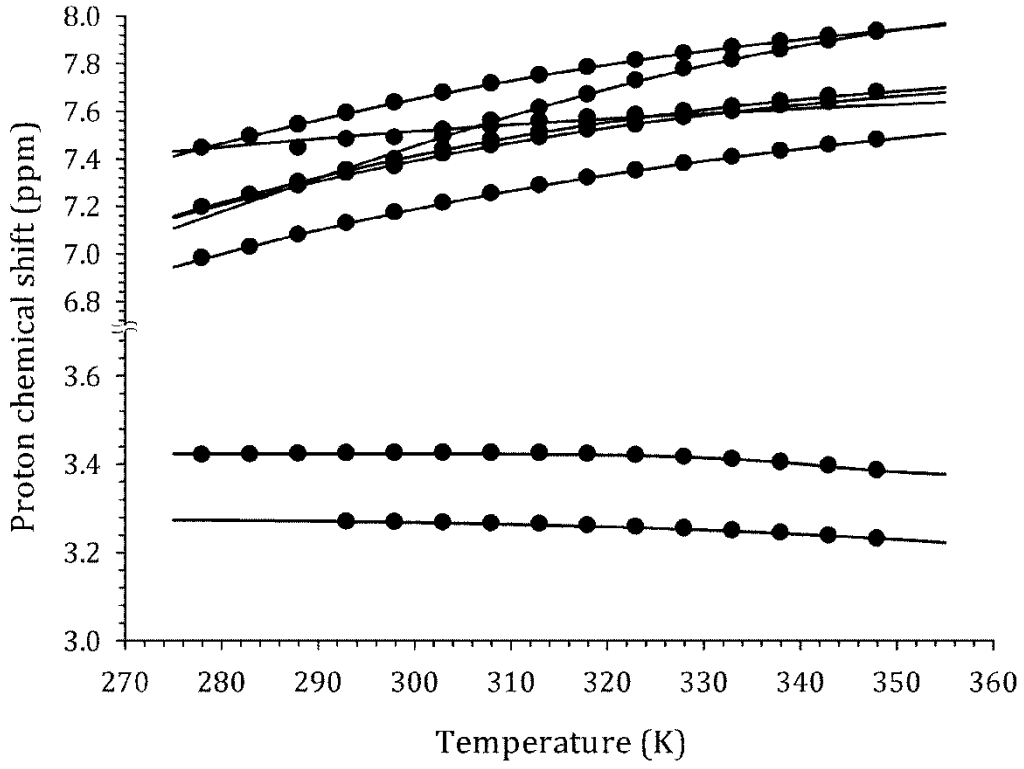


Figure S27: ^{1}H NMR chemical shift as a function of temperature for **8**.

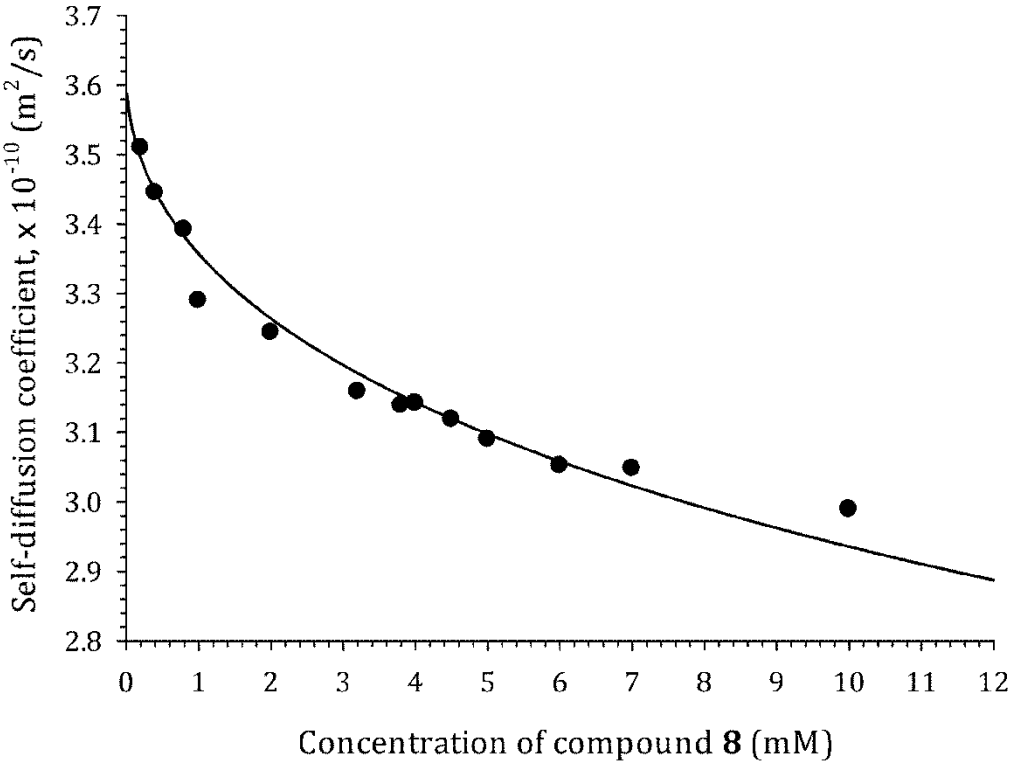


Figure S28: ^{1}H NMR-derived diffusion coefficient as a function of concentration for **8**.

Section B – Supplementary Tables

In the following tables of supporting information, the models referred to in the columns headed "**Model**" are as described in **Table S1**.

Table S1: Model definitions

Model Number	Model Definition
1	Dimer model with 3 adjustable parameters
2	Spherical with 4 adjustable parameters
3	Oblate ellipsoid with 3 adjustable parameters
4	Oblate ellipsoid with 4 adjustable parameters
5	Prolate ellipsoid with 3 adjustable parameters
6	Prolate ellipsoid with 4 adjustable parameters
7	Cylinder without correction for the end-effects with 3 adjustable parameters
8	Cylinder without correction for the end-effects with 4 adjustable parameters
9	Cylinder with 3 adjustable parameters
10	Cylinder with 4 adjustable parameters
11	SHIM-model with 3 adjustable parameters
12	Dimer model with fixed D1/D2=1.26
13	Spherical with 4 adjustable parameters with fixed D1/D2=1.26
14	Oblate ellipsoid with 4 adjustable parameters with fixed D1/D2=1.26
15	Prolate ellipsoid with 4 adjustable parameters with fixed D1/D2=1.26
16	Cylinder with 4 adjustable parameters with fixed D1/D2=1.26
17	SHIM-model with fixed D1/D2=1.26

The calculated parameters K - equilibrium self-association constant, D_1 – monomer self-diffusion coefficient, D_2 – dimer self-diffusion coefficient, d – molecule diameter and R^2 – goodness of fit are listed in each of the following tables associated with each of the eight test compounds used for experimental data collection according to the model type used as defined in detail in the main text of the article.

Table entries that are shown in **red** highlight inappropriate models that are identified through calculated parameters that lie outside the designated criteria defined for acceptable models according to the details described in the main text of the paper.

Table S2a: Parameter values calculated with each model for **1**, Hoechst 33258.[†]

Model	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	d, nm	R^2
Strategy 1					
1	0.043553541	2.460234012	0.191888597		0.983092316
2	0.153776503	2.618211206	1.538370793	2.461574486	0.979792921
3	0.130500814	2.470018707		8.52107E-05	0.979756703
4	0.139229442	2.384305991	2.055453795	6.803903687	0.98275724
5	0.101860199	2.472317542		0.010260918	0.983144996
6	0.077473885	2.435558086	1.324313333	8.029517172	0.982875093
7	0.027646527	2.295449999		1.272938940	0.975639108
8	0.03805445	2.391629478	0.585126387	0.974464786	0.981984815
9	0.020510286	1.275610557		2.527527960	0.973778624
10	0.045069831	2.403085871	0.795360449	2.897712792	0.982301779
11	0.117028721	2.479998219	1.541642647		0.983165918
Strategy 2					
12	0.108649847	1.851329557	1.469401243		0.319773425
13	0.134220652	2.412123444	1.914503647	3.021445057	0.982915452
14	0.122234534	2.398553323	1.903733034	7.097486425	0.982792606
15	0.123474243	2.399939289	1.904833076	7.219233700	0.982806095
16	0.084545577	2.352485503	1.867168981	3.095057349	0.982124298
17	0.462237541	2.679535935	2.126749081		0.981896694

[†] K determined by ^{1}H NMR chemical shift measurements = 0.183 mM^{-1}.

Table S2b: Parameter values calculated to specifically test model 11 on **1**, Hoechst 33258, using randomized exclusion of data points.

No. Points Excluded	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	R^2
Model 11 Test Data				
1	0.159833346	2.516724919	1.527207378	0.938631136
2	0.184974218	2.525461315	1.522140721	0.932333151
3	0.133122673	2.518705133	1.499285837	0.927128892

Table S3a: Parameter values calculated with each model for **2**, Ethidium Bromide.[†]

Model	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	d, nm	R^2
			Strategy 1		
1	0.147845555	4.971256910	2.28198288		0.991147866
2	0.426756552	4.967226170	4.024847726	0.804788644	0.991304599
3	0.972169175	4.995172917		1.311520958	0.991528788
4	0.542950625	4.990982973	4.116596734	1.115168211	0.991578672
5	0.859643778	4.995225350		1.264363099	0.991545769
6	0.544751659	4.991697415	4.114860655	1.112075615	0.991586824
7	0.358266584	4.912790012		0.119049209	0.986422985
8	1.019902278	4.968690233	4.648514404	0.381848727	0.991831676
9	0.034717579	3.279303620		2.292646506	0.957742133
10	0.571149352	4.991257431	4.158023741	0.788546424	0.991595908
11	1.161351576	5.000985502	4.533294518		0.991505321
			Strategy 2		
12	0.749950617	4.802114706	3.811440968		0.704039422
13	0.425479873	4.982592574	3.954686347	0.796254373	0.991439292
14	0.444129170	4.987954282	3.958941938	1.178979616	0.991562897
15	0.445437394	4.988325379	3.959236477	1.176740327	0.991570512
16	0.439972640	4.987277675	3.958404913	0.846271885	0.991562203
17	0.307547262	4.919849043	3.904886773		0.987359919

[†]K determined by ^{1}H NMR chemical shift measurements = 0.305 mM^{-1}.

Table S3b: Parameter values calculated to specifically test model 11 on **2**, Ethidium Bromide, using randomized exclusion of data points.

No. Points Excluded	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	R^2
		Model 11 Test Data		
1	1.224485488	5.003162516	4.550089471	0.989481574
2	1.371085468	5.007544928	4.584366135	0.986752109
3	1.498834251	5.010735387	4.609887227	0.983054016

Table S4a: Parameter values calculated with each model for **3**, Caffeine.[†]

Model	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	d, nm	R^2
Strategy 1					
1	0.030226266	6.846290095	4.851883806		0.990352792
2	0.041781352	6.859950220	5.295135494	0.468521362	0.990696236
3	0.008693528	6.657718445		0.013896633	0.952106126
4	0.035729084	6.859331966	5.056457823	0.48570649	0.990794680
5	0.006223601	6.644292501		1.37E-05	0.945929054
6	0.035162807	6.859139568	5.031673845	0.480168296	0.990805496
7	0.020736727	6.676471611		0.118707601	0.959576187
8	0.036888075	6.859698688	5.107545731	0.216959885	0.990790478
9	0.007737575	6.583200070		0.019975425	0.949554198
10	0.035012892	6.859169554	5.024045710	0.346002704	0.990808419
11	0.033101402	6.700092743	5.845994206		0.969402057
Strategy 2					
12	0.080519323	6.888191398	5.467161135		0.944891696
13	0.050523478	6.870550915	5.453159875	0.463693864	0.990637394
14	0.052086787	6.860483172	5.445169102	0.471811052	0.989497540
15	0.051010125	6.855094922	5.440892446	0.468905863	0.989211212
16	0.050726658	6.853541723	5.439659671	0.338431578	0.989111928
17	0.017599753	6.680842011	5.302587818		0.961207097

[†]K determined by ^{1}H NMR chemical shift measurements = 0.0118 mM^{-1}.

Table S4b: Parameter values calculated to specifically test model 11 on **3**, Caffeine, using randomized exclusion of data points.

No. Points Excluded	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	R^2
Model 11 Test Data				
1	0.051350766	6.772114960	5.814783017	0.984221765
2	0.032586202	6.828990743	5.840144198	0.981647519
3	0.046924641	6.891206859	5.049503202	0.978986961

Table S5a: Parameter values calculated with each model for **4**, Proflavine.[†]

Model	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	d, nm	R^2
Strategy 1					
1	0.373374914	6.105641956	1.213592072		0.998509356
2	0.483843874	6.065380531	2.702376484	1.100223953	0.998551455
3	1.183231316	6.166468101		0.149826988	0.998322845
4	0.521213144	6.068208664	2.920604894	1.921651147	0.998553508
5	1.417129078	6.189347033		0.255437974	0.998235202
6	0.518172247	6.067981706	2.904014086	1.923763970	0.998553828
7	3.569946136	6.140111259		0.146464701	0.998159085
8	0.683811041	6.012960022	4.010832633	0.622310194	0.998001832
9	1.442602425	4.660368453		0.193006529	0.998231796
10	0.532686417	6.060608144	3.032217946	1.338341737	0.998547980
11	2.218750179	6.239209341	5.005948387		0.998030288
Strategy 2					
12	0.757718943	4.707683604	3.736490953		0.364033400
13	1.126453431	6.034715103	4.789756552	0.961190953	0.998587350
14	1.151972815	6.044161655	4.797254285	1.560358266	0.998606131
15	1.155101111	6.045103492	4.798001822	1.556975477	0.998606838
16	1.120519901	6.036755839	4.791376285	1.122240056	0.998615079
17	2.022142217	6.215582602	4.933311181		0.998021098

[†]K determined by ^{1}H NMR chemical shift measurements = 0.698 mM^{-1}.

Table S5b: Parameter values calculated to specifically test model 11 on **4**, Proflavine, using randomized exclusion of data points.

No. Points Excluded	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	R^2
Model 11 Test Data				
1	2.201899788	6.237733180	4.999985943	0.997666762
2	1.699713453	6.188550047	4.784884710	0.997797944
3	1.694820300	6.188019570	4.782335824	0.997269671

Table S6a: Parameter values calculated with each model for azo-dye **5**.[†]

Model	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	d, nm	R^2
			Strategy 1		
1	0.814774689	3.376789024	1.152875028		0.993287527
2	5.743957393	3.503401660	2.886276327	1.121080161	0.995720014
3	47.61525164	3.370360952		3.809236483	0.991402446
4	6.978306342	3.774588385	2.336633379	1.741230738	0.997162812
5	138.1782144	3.537100301		7.312346333	0.993246578
6	6.373302198	3.732797739	2.359258424	1.786154600	0.997111120
7	1.826894809	3.355531622		0.110975420	0.949431006
8	10.98714523	3.403466689	3.749675782	0.613774378	0.996937735
9	1.931329908	0.313464104		0.436347647	0.971493303
10	6.623353588	3.709790519	2.484715707	1.260360264	0.997080186
11	52.42028279	3.406135723	3.287904248		0.988178627
			Strategy 2		
12	0.944574302	2.421324537	1.921806559		0.393609281
13	4.805021465	3.499558181	2.777601169	1.155161045	0.995715189
14	5.161511697	3.493436541	2.772742420	1.969922669	0.996467836
15	4.899663110	3.481855870	2.763550835	1.997612317	0.996361076
16	5.481799012	3.524538018	2.797427679	1.361039334	0.996792486
17	1.238155832	3.145330086	2.496450144		0.968873123

[†]K determined by ^{1}H NMR chemical shift measurements = 2.17 mM^{-1}.

Table S6b: Parameter values calculated to specifically test model 11 on azo-dye **5** using randomized exclusion of data points.

No. Points Excluded	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	R^2
		Model 11 Test Data		
1	47.06723738	3.465298485	3.318578623	0.938631136
2	42.25420345	3.453290089	3.304328100	0.932333151
3	40.89567265	3.435030210	1.499285837	0.927128892

Table S7a: Parameter values calculated with each model for **6**, AIK-18/52.[†]

Model	K, mM^{-1}	D_1, ×10^{-10}, m^2·s^{-1}	D_2, ×10^{-10}, m^2·s^{-1}	d, nm	R^2
			Strategy 1		
1	0.209639110	5.781284554	0.595121024		0.996516311
2	0.591342644	5.628452875	4.453836181	1.286582674	0.996127819
3	0.650997442	5.906972328		3.30666E-07	0.995795423
4	0.337106048	5.674188973	3.035127299	2.901069160	0.996204695
5	0.649338190	5.923320024		0.063733482	0.996708576
6	0.336937721	5.675343202	3.028617837	2.908853610	0.996208736
7	0.665016447	5.929957774		0.030919968	0.996711589
8	0.240702263	5.552033405	2.597541852	0.800342790	0.995635996
9	0.644195128	5.732643746		0.046641689	0.996706861
10	0.338695177	5.636624608	3.214647918	1.906771515	0.996071011
11	1.036705676	6.046585245	4.367102370		0.996796913
			Strategy 2		
12	0.294209829	3.879991412	3.079551225		0.283303479
13	0.565418615	5.589352877	4.436272318	1.318373914	0.996149472
14	31.50065745	1.051798072	0.834812683	0.211794821	0.996402137
15	11.85291914	5.734903174	4.551795666	0.390504857	0.996644169
16	11.44597394	5.769104338	4.578941147	0.284417588	0.996643467
17	2.233201951	6.386664094	5.069098651		0.996672818

[†]K determined by ^{1}H NMR chemical shift measurements = 0.406 mM^{-1}.

Table S7b: Parameter values calculated to specifically test model 11 on **6**, AIK-18/52, using randomized exclusion of data points.

No. Points Excluded	K, mM^{-1}	D_1, ×10^{-10}, m^2·s^{-1}	D_2, ×10^{-10}, m^2·s^{-1}	R^2
		Model 11 Test Data		
1	1.168653004	6.085534575	4.396858122	0.996771836
2	1.200712329	6.123898789	4.420620078	0.997408301
3	1.518471874	6.166341438	4.486926216	0.997133495

Table S8a: Parameter values calculated with each model for **7**, Nile Blue.[†]

Model	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	d, nm	R^2
			Strategy 1		
1	1.055904949	4.923606841	0.427253064		0.984235376
2	0.494400683	4.641945104	0.406717889	1.924396313	0.986816658
3	2.664518834	4.726658902		7.02E-07	0.975091906
4	10.07976888	8.67631E-05	12.10557216	2.539659149	0.988918579
5	2.934200894	4.968951164		0.045121216	0.982701086
6	13.60474208	-2.756213467	15.06168658	2.209800776	0.988941677
7	2.923036460	4.964765202		0.02128132	0.982688323
8	1.000502510	4.553515811	2.631494588	1.06510832	0.986431374
9	2.937628319	4.870873531		0.033295959	0.982713489
10	0.977358222	4.663569764	1.967678710	2.727277253	0.986646705
11	4.099913202	5.006336409	3.493836991		0.981821695
			Strategy 2		
12	0.911004405	2.660697983	2.111797389		0.432112336
13	2.364618177	4.676429813	3.711684802	1.860169941	0.986078166
14	2.164008373	4.644981713	3.686724429	4.18704887	0.986385098
15	2.178813915	4.647314075	3.688575626	4.193680907	0.986364220
16	1.669317955	4.563295810	3.621890285	2.724254445	0.987066078
17	11.43856976	5.393202115	4.280587355		0.981313039

[†]K determined by ^{1}H NMR chemical shift measurements = 5.6 mM^{-1}.

Table S8b: Parameter values calculated to specifically test model 11 on **7**, Nile Blue, using randomized exclusion of data points.

No. Points Excluded	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	R^2
		Model 11 Test Data		
1	3.703697324	4.919239189	3.373381427	0.979294696
2	5.216159935	5.047122264	3.868464401	0.961377294
3	4.799939790	5.153334523	3.924701274	0.926247262

Table S9a: Parameter values calculated with each model for **8**.[†]

Model	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	d, nm	R^2
			Strategy 1		
1	0.250529993	3.580656558	2.672039924		0.988586
2	0.249513088	3.593943886	2.601723527	0.840841979	0.989530963
3	0.559828344	3.480399211		2.623930758	0.949879622
4	0.230250078	3.590573587	2.56087916	1.307556673	0.989425858
5	0.026879131	3.430210893		0.006031920	0.893816722
6	0.229250451	3.590404572	2.558516052	1.307221909	0.989419370
7	0.084963034	3.451989829		0.120171510	0.914684091
8	0.347430487	3.594100379	2.873065869	0.477859130	0.989485469
9	0.010759925	0.222613331		7.856992706	0.869019714
10	0.235072836	3.590601925	2.580654903	0.932100693	0.989424097
11	0.185153254	3.467333592	3.110810143		0.941071433
			Strategy 2		
12	0.446497107	3.572883791	2.835799744		0.959669096
13	0.319219326	3.570070183	2.833566582	0.849781224	0.984552560
14	0.275974038	3.555837091	2.822269769	1.357522208	0.982679138
15	0.274423335	3.555241996	2.821797442	1.358477771	0.982564770
16	0.276830457	3.557848304	2.823866070	0.964189293	0.983716005
17	0.391405546	3.556128607	2.844943379		0.980996616

[†]K determined by ^{1}H NMR chemical shift measurement = 0.585 mM^{-1}.

Table S9b: Parameter values calculated to specifically test model 11 on **8** using randomized exclusion of data points.

No. Points Excluded	K, mM^{-1}	D_1, $\times10^{-10}$, $m^2 \cdot s^{-1}$	D_2, $\times10^{-10}$, $m^2 \cdot s^{-1}$	R^2
		Strategy 1		
1	0.223298855	3.490480412	3.169422631	0.977395466
2	0.242716096	3.515658038	3.172353229	0.979324397
3	0.207621645	3.562686925	3.279464710	0.975584142