Biophysical Chemistry 21 (1985) 145–156
Elsevier



BPC 00906

# NUCLEOTIDE AGGREGATION IN AQUEOUS SOLUTION

## A MULTICOMPONENT SELF-DIFFUSION STUDY

Roger RYMDÉN and Peter STILBS

*Institute of Physical Chemistry, Uppsala University, Box 532, S-751 21 Uppsala, Sweden*

Received 12th June 1984
Revised manuscript received 21st September 1984
Accepted 4th October 1984

*Key words: Self-diffusion; Stacking; Aggregation; Nucleotide; NMR*

The self-aggregation of the mononucleotides (AMP, CMP, GMP and UMP) and caffeine up to their solubility limit in $^2H_2O$ has been monitored through self-diffusion measurements, using the Fourier transform NMR pulsed-gradient spin-echo self-diffusion technique. The data were iteratively fitted to a number of aggregation models. It was concluded that the best agreement between simulations and experiment for the mononucleotides was obtained for a 'semi-isodesmic', indefinite aggregation model (also known as a Type III SEK or cooperative indefinite self-association model), where the first (dimerization) aggregation constant is a magnitude lower than those for the higher aggregation steps. Typical values were 0.4 and 6 l $mol^{-1}$, respectively. Under these conditions, the main fraction of solute is monomeric throughout the concentration range and the distribution of higher oligomers is very broad. Caffeine self-aggregation is clearly different and is consistent with several aggregation models. The mixed aggregation of caffeine (at a low total concentration) and the mononucleotides was successfully monitored in an extension of the basic study. It was found that caffeine binding to mononucleotide aggregates increases in the series UMP, CMP, GMP and AMP.

## Introduction

The self-association of nucleic acid bases, nucleosides and nucleotides in aqueous solution is generally recognized as a key biophysical process. Such aggregation phenomena have extensively been studied over the past decades using numerous techniques [1–24]. The vast amount of experimental data accumulated to date is consistent with the assumption that aggregation occurs via vertical stacking of the planar bases.

Different techniques provide different types of information about aggregation processes. Information from thermodynamic methods, such as osmometry and vapour pressure measurements for example, is related to the overall solute activity or to the total extent of aggregation. Spectroscopic techniques, designed for aggregation monitoring, sample the aggregation process in a different manner from thermodynamic measurements * and may, in addition, provide information on the molecular level on the the actual mode of association. The most valuable spectroscopic method in this context has been NMR. Proton and carbon chemical shift changes upon aggregation have been extensively utilized to obtain association constants and structural information [14–18]. The methodological approach for structural deductions by NMR in the present context is based on the analysis of relative orientations of molecular fragments, assuming that chemical shift changes upon aggrega-

* Of course, information from as large a number of different methods as possible should be utilized and combined in the final description of the aggregation process under consideration.

tion originate from 'ring-current' effects from the aromatic moieties. To a first approximation, a point-dipole type shielding field around the aromatic rings is induced [25]. Deeper analyses of the origin of these shift effects have been made recently [26,27], possibly necessitating reassessment of literature data and structural conclusions on the molecular level, also in the nucleotide field.

More sophisticated NMR experiments, aiming at structural deductions about aggregation phenomena in these systems, have also been presented [19–24].

NMR-based quantification of the extent of aggregation is commonly based on time-averaged chemical shifts ($\delta_i$) in different environments (differences being ascribed to effects of ring-current fields), according to:

$$\delta_{obs} = \Sigma p_i \delta_i \quad (1)$$

where the different $\delta_i$ typically become iteration parameters in a computer simulation of eq. 1, linked to the equilibrium equations for the association model under consideration (which provides the $p$ values; the relative fractions of aggregating monomers in different aggregates; $\Sigma\ p_i = 1.00$). Usually some analytic relation between the shift values for the different species can be made plausible (see, e.g., ref. 16), although it may not be valid or proven (see section 5).

An identical measurement approach in outline can be based on time-averaged molecular self-diffusion data. The analog of eq. 1 then becomes:

$$D_{obs} = \Sigma p_i D_i \quad (2)$$

Self-diffusion coefficients provide a rather direct link to information on molecular aggregation processes, since the $D_i$ are closely dependent on aggregate size. Self-diffusion techniques for monitoring aggregation phenomena have been relatively little used, probably due to a lack of convenient experimental techniques. (One recent self-diffusion investigation in the present field is ref. 28; monomer diffusion was investigated in dilute solution). The experimental situation has recently changed drastically with the development of the Fourier transform [29] modification pulsed-gradient [30] spin-echo [31] NMR technique (FT-PGSE) into a practical tool [32,33]. It now offers a convenient, rapid and general means for measuring individual self-diffusion coefficients in complex systems, providing a general method for the study of aggregation phenomena. FT-PGSE techniques have previously been successfully applied for the investigation of aggregation in a number of systems, e.g., in micellar [33–35], microemulsion [35–38], vesicular [39], polymeric polyelectrolyte [40,41] and cyclodextrin solutions [42]. In general, the self-diffusion approach is particularly favourable for the monitoring of the binding of small molecules to supramolecular aggregates or macromolecules.

We found it of interest to explore the applicability of the FT-PGSE technique in connection with the association of molecules containing nucleic acid bases. The particular characteristics of the method as compared to previous techniques will be discussed below. In the present work, we have studied the concentration dependence of the self-diffusion coefficients of caffeine and a number of nucleotides. The results have been tested against four different association models.

One unique advantage of the FT-PGSE approach for the investigation of aggregation phenomena is its selective nature. The feasibility for monitoring the mixed association of caffeine with different nucleotides was, for illustration, investigated in an extension of the basic study.

## 2. Experimental

### 2.1. NMR measurements

The diffusion measurements were performed on a JEOL FX 100 Fourier Transform NMR spectrometer, as described previously [32,33]. Internal field/frequency lock on $^2H_2O$ was used throughout. All measurements were made at 24.7 $\pm$ 0.2°C. Briefly, the technique entails Fourier transformation of the second half of the pulsed-gradient spin-echo following the second field gradient pulse, keeping the pulse interval ($\Delta$) fixed for all durations of the gradient pulses ($\delta$). Under these conditions $J$ modulation as well as $T_2$ effects become constant and need not be further considered. The signal amplitude of a nucleus in the

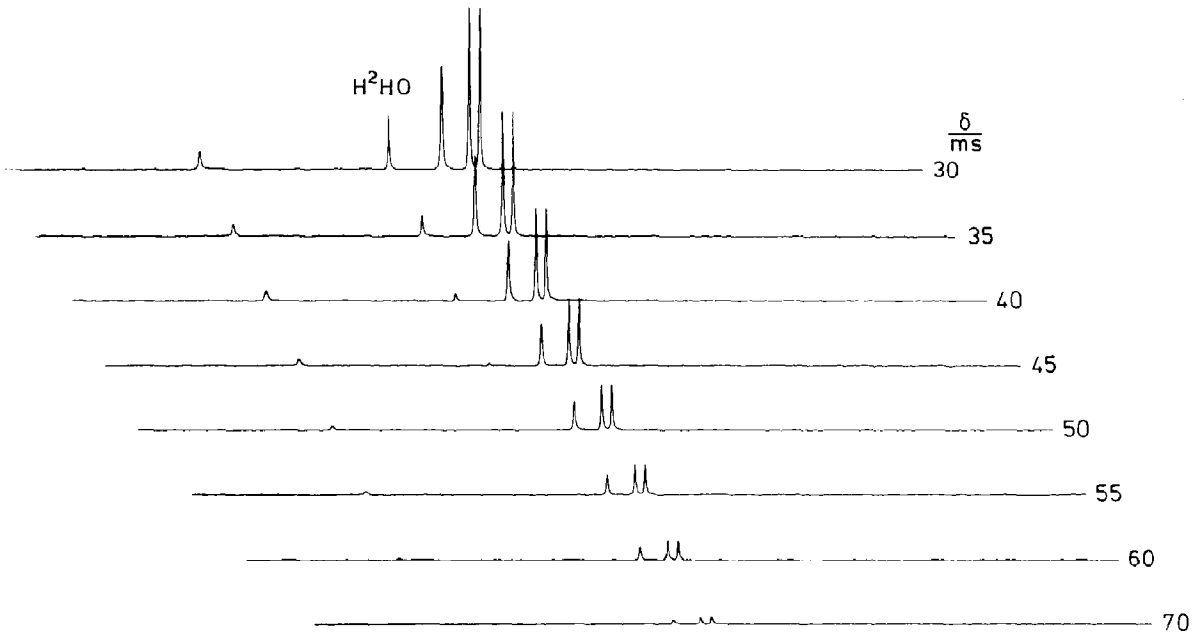


Fig. 1. A typical sequence of FT-PGSE spectra for caffeine ($C = 0.0335$ M). The durations of the applied magnetic field gradient pulses range from 30 to 70 ms. The 90–180° radiofrequency pulse interval was kept constant at 140 ms.

spin-echo spectrum is related to $\delta$ through the relation

$$A_i \alpha \exp\left(-\gamma^2 G^2 \delta^2 D_i (\Delta - \delta/3)\right) \qquad (3)$$

where $\gamma$ is the magnetogyric ratio of the nucleus, $G$ the strength of the applied pulsed magnetic field gradient, $\Delta$ the common 90–180° radiofrequency and gradient pulse interval and $\delta$ the duration of the applied magnetic field gradient pulses (the notation is established in spin-echo NMR and should not be confused with the identical symbol used for proton and carbon chemical shifts). In the present investigation $\Delta$ was 140 ms while $\delta$ ranged from 30 to 70 ms. A typical series of FT-PGSE spectra from the present investigation is illustrated in fig. 1.

### *2.2. Materials and sample preparation*

The nucleotides (all in the disodium form), adenosine 5′-monophosphate (type 11, from yeast), guanosine 5′-monophosphate (from yeast), cytosine 5′-monophosphate (from yeast), uracil 5′-monophosphate (from yeast), and also caffeine were obtained from Sigma Chemical Co. They were used without further purification. The solvent, $^2H_2O$, was purchased from Norsk Hydro, Rjukan, Norway. Stock solutions were prepared volumetrically, directly in the calibrated precision bore thin-wall 5-mm NMR tubes.

## 3. Measurement and calculation approach

### *3.1. General comments on self-diffusion in aggregating systems*

The present technique of studying molecular association is particularly suitable and straightforward for the investigation of the binding of small substrates to large macromolecules. This is easily seen from eq. 2; the fraction of each species is weighted according to its diffusion coefficient. The partial binding of a small molecule to a macromolecule or supramolecular species will therefore manifest itself in a strongly receptive manner in its

time-averaged self-diffusion coefficient. In applying the self-diffusion technique for the investigation of aggregation phenomena of the present kind, one notes that the method necessarily will be biased to contributions from monomers and lower oligomers. It can be noted, in contrast, that the 'chemical-shift' approach mentioned above intrinsically gives essentially equal weight to contributions from all species.

The binding constants involved in nucleotide association are known to be rather small, making it necessary to monitor a relatively large concentration span in order to induce significant changes in observed diffusion coefficients.

A complication arising from the increase in total concentration is the assessment of obstruction effects; the diffusional path of a particular species will increase due to the presence of aggregates and result in a decrease of the diffusion coefficient.

### 3.2. Calculations

The observed diffusion coefficient, $D_{obs}$, in a system of associating molecules subject to the condition of rapid exchange is given by eq. 2 By combining eq. 2 with the appropriate assumed association model (providing the $p$ values as a function of concentration), the experimental data can be fitted to different sets of association parameters through computer simulation. A linear obstruction correction, similar to the Wang relation [43] for the self-diffusion of small molecules in macromolecular solution was employed throughout on all raw self-diffusion data:

$$D_{obs} = D_{free}(1 - 1.5\phi) \quad (4)$$

where $\phi$ is the volume fraction of solute. The Wang relation in the above form was derived for substrate diffusion among spherical macromolecules. It has the same functional form for other obstruction geometries and is expected to serve as a first approximation for the systems under investigation in the present study. Typically, the assumed linear relation was such that the self-diffusion of all species in solution was reduced 20% in a 1 M solution. Numerous tests were made with other choices of the actual value of this linear correction term.

The fact that the aggregates are charged will lead to additional retardation effects on molecular transport in solution; we have chosen to assume that the concentration dependence of that factor is similar to, and can be included in, the general obstruction term just mentioned.

A series of iterative non-linear least-squares computer programs (minimizing the sum over all $(D_{obs} - D_{calc})^2$ under the constraints of the aggregation model and eq. 2) were developed for the simulation of self-diffusion data within the different aggregation models. Details of algorithms and computational procedures will be published elsewhere [44].

### 3.3. Model 1, a monomer-single n-mer equilibrium

The first model involves the formation of $n$-mers according to:

$$n\mathrm{M} \rightleftarrows \mathrm{M}_n \quad (5)$$

The computational problem in fitting the observed diffusion data to this aggregation model has been solved in ref. 45, for example, and a computer program ASSOCO was written, as based on the described algorithm. A series of $n$ values were tested. Self-diffusion coefficients of monomers and aggregates were considered simulation parameters.

### 3.4. Model 2, n-merization, followed by m-merization of these aggregates

A computer program ASSONM was written and acquired data were tested against several typical cases, e.g., tetramerization and subsequent dimerization of the tetramers with a different aggregation constant, etc. Self-diffusion coefficients of monomers and aggregates were considered simulation parameters.

### 3.5. Model 3, indefinite aggregation, all steps equal

A suitable algorithm has been presented by Chun [46], which formed the basis for a computer program ASSINF. A generalized algorithm for

evaluating the equilibrium concentrations in a series of coupled equilibria was later found, and formed the basis for a more general program, ASGEIG (see section 3.5.1). A special version of ASGEIG, ASGE11, simulates indefinite aggregation, with all equilibrium constants set equal.

*3.5.1. Model 4, indefinite aggregation, dimerization step unique*

The following model was considered:

$$\begin{aligned} &A + A \rightleftarrows A_2 \\ &A_2 + A \rightleftarrows A_3 \\ &A_3 + A \rightleftarrows A_4 \\ &A_4 + A \rightleftarrows A_5 \\ &A_5 + A \rightleftarrows A_6 \\ &A_6 + A \rightleftarrows A_7 \ldots . \end{aligned} \quad (6)$$

where the equilibrium constant for the first (dimerization) step represents an independent parameter. All higher equilibrium constants were set equal and were varied independently from the dimerization constant. Influence from all species, up to 20-mers, were considered in the simulations. A special version (ASGE30) of ASGEIG was written to account for species up to 30-mers in the case of AMP aggregation at increased concentrations.

In Models 3 and 4 it is necessary to include some relation between aggregation number and self-diffusion coefficient of the different species. We found plausible, and have chosen, a model based on hydrodynamic corrections for changing geometry (oblate-spherical-prolate) upon $n$-merization. The relation used had the following form:

$$D(n) = D(1)/\left[(n^{0.3333})\ (\text{polynomial in } `n')\right] \quad (7)$$

where the $n^{0.3333}$ term corrects for the general increase in hydrodynamic radius (see, e.g., ref. 46) and the polynomial function was based on least-squares fitting to tabulated coefficients of the so-called Perrin $F$ factors (ratios between frictional factors for spheres and ellipsoids of the same volume) [47] and an assumption that the oblate-spherical-prolate crossover point within a vertical stacking model occurs for dimers. Typically, the estimated $D$ in Models 3 and 4 decreases to approx. 25% of its original value, going from monomer to 20-mer, when using this hydrodynamic model basis.

*3.6. Caffeine binding-in*

Since the total concentration of caffeine was very much lower than that of the mononucleotide the following (simplest-possible) model was considered to serve as a good first approximation to this, considerably more complicated, aggregation problem (A, mononucleotide; C, caffeine):

$$\begin{aligned} &A + C \rightleftarrows AC \\ &A_2 + C \rightleftarrows A_2C \\ &A_3 + C \rightleftarrows A_3C \\ &A_4 + C \rightleftarrows A_4C \\ &A_5 + C \rightleftarrows A_5C \\ &A_6 + C \rightleftarrows A_6C \ldots . \end{aligned} \quad (8)$$

where the equilibrium constants for the first two steps represent independent parameters. All higher equilibrium constants were set equal and were varied independently from the first two. Influence from all species, up to 20-mers, was considered in the simulations.

The calculations were made by a special program ASGCOF, in a manner where data like those depicted in figs. 8–11 were first simulated (using the ASGEIG program) with regard to mononucleotide aggregation (providing the concentrations of mononucleotide oligomers at the different concentrations, followed by a second simulation on the model (eq. 8) on these data, assuming that the general mononucleotide aggregation is unperturbed upon binding of caffeine. Diffusion coefficients for aggregates $A_nC$ were assumed to be the same as those derived for the $A_{n+1}$ aggregates in the first cycle of the simulation.

It should be pointed out that the series of equilibria, eq. 8, are parallel ones, unlike those in eq. 6 which are coupled equilibria.

## 4. Results

*4.1. Model 1*

This association model would be appropriate if the distribution of aggregate sizes was sharply

peaked around a fixed aggregation number. The results of fitting the experimental self-diffusion data to eq. 2 and 5 for different values of $n$ are presented in table 2 (fits based on the limited concentration interval below 0.25 M) and in figs. 2–6. The agreement between observed and calculated values is quite good. Inspection of the fitted parameters corresponding to the aggregation number which gives a minimum in the error square sum reveals, however, that the diffusion coefficient of the aggregate becomes unphysically low for the purine nucleotides, i.e., AMP and GMP. (The general features of these data simulations are insensitive even to large obstruction correction terms.) The self-diffusion coefficients of the pyrimidine nucleotide aggregates turn out more reasonable, on the other hand, as seen in table 1.

Caffeine data, over the limited accessible solubility interval, fit this association model well, although the fit is insensitive to the actual value of the aggregation number (see table 2 and fig. 6).

### 4.2. Model 2

A large number of combinations of $m$ and $n$ values were tested with the aid of the ASSONM program. Since this model is, in effect, an extension of Model 1, the fit is generally better. The fitted diffusion coefficients for the aggregates were unphysically low, however. Some numeric instability in parameter values (due to covariance effects) was also noted in this rather complex simulation model.

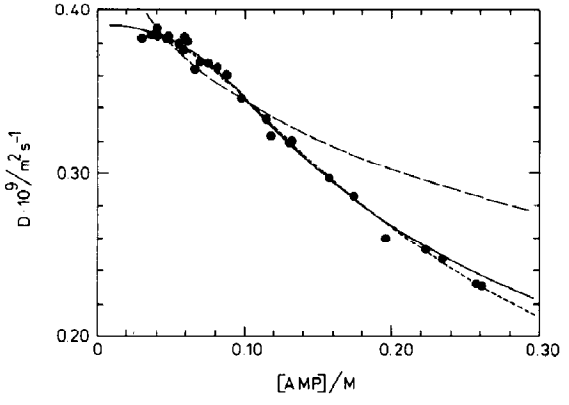


Fig. 2. The concentration dependence of the self-diffusion coefficient of AMP. Filled circles are experimental points. Curves are calculated from the fitted parameters for the ASSOCO (- - - - -), ASGE30 (———) and ASGE11 (— — —) computer simulations (Models 1, 4 and 3, respectively). The parameters used in the calculation of the ASSOCO curve pertain to the aggregation number giving the best fit to the experimental data (see table 2).

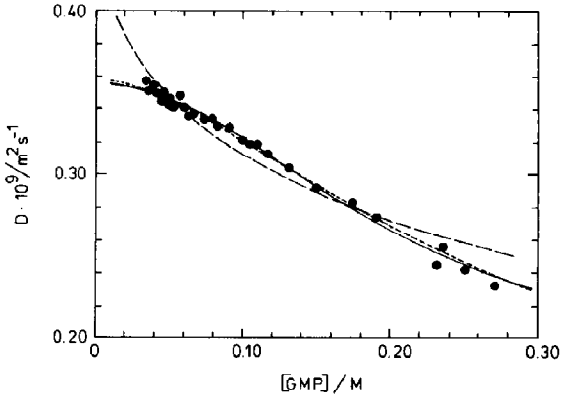


Fig. 3. The concentration dependence of the self-diffusion coefficient of GMP. Filled circles are experimental points. Curves are calculated from the fitted parameters for the ASSOCO (- - - - -), ASGEIG (———) and ASGE11 (— — —) computer simulations (Models 1, 4 and 3, respectively). The parameters used in calculating the ASSOCO curve pertain to the aggregation number giving the best fit to the experimental data (see table 2).

#### 4.2.1. Models 3 and 4

Here self-diffusion coefficients calculated from

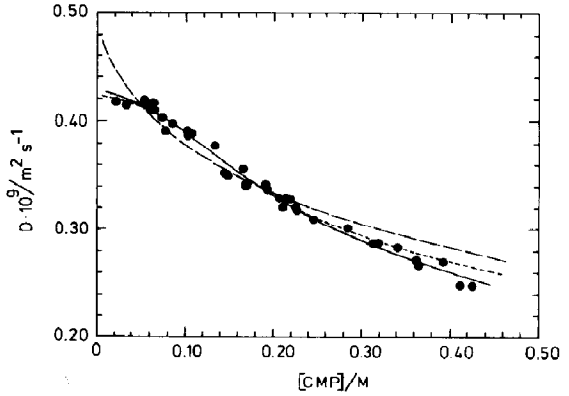


Fig. 4. The same data as in fig. 3, but for CMP.

hydrodynamics enter into the calculations. These are, of course, physically reasonable a priori; it is the aggregation model itself that is subject to fitting procedures.

With the exception of caffeine, which fits both models equally well, Model 4 provides a signifi-

Table 1

Self-diffusion coefficients of nucleotides and caffeine in $^2H_2O$

Diffusion coefficients are expressed in units of $10^{-9}$ $m^2$ $s^{-1}$.

| UMP | | GMP | | AMP | | CMP | | Caffeine | |
|---|---|---|---|---|---|---|---|---|---|
| $C/M$ | $D_{obs}$ | $C/M$ | $D_{obs}$ | $C/M$ | $D_{obs}$ | $C/M$ | $D_{obs}$ | $C/M$ | $D_{obs}$ |
| 0.0396 | 0.410 | 0.0344 | 0.357 | 0.0303 | 0.383 | 0.0212 | 0.417 | 0.00538 | 0.635 |
| 0.0407 | 0.412 | 0.0361 | 0.351 | 0.0371 | 0.385 | 0.0322 | 0.414 | 0.00645 | 0.638 |
| 0.0453 | 0.407 | 0.0402 | 0.354 | 0.0404 | 0.389 | 0.0559 | 0.415 | 0.00689 | 0.645 |
| 0.0528 | 0.411 | 0.0421 | 0.349 | 0.0406 | 0.384 | 0.0567 | 0.417 | 0.00886 | 0.634 |
| 0.0634 | 0.410 | 0.0446 | 0.345 | 0.0477 | 0.383 | 0.0580 | 0.416 | 0.00926 | 0.638 |
| 0.0683 | 0.410 | 0.0459 | 0.350 | 0.0485 | 0.384 | 0.0616 | 0.412 | 0.0107 | 0.631 |
| 0.0739 | 0.401 | 0.0470 | 0.350 | 0.0556 | 0.380 | 0.0636 | 0.417 | 0.0108 | 0.630 |
| 0.0743 | 0.406 | 0.0482 | 0.343 | 0.0587 | 0.376 | 0.0638 | 0.410 | 0.124 | 0.633 |
| 0.0821 | 0.399 | 0.0501 | 0.346 | 0.0592 | 0.384 | 0.0729 | 0.403 | 0.0135 | 0.622 |
| 0.0853 | 0.397 | 0.0507 | 0.342 | 0.0613 | 0.381 | 0.0769 | 0.391 | 0.0155 | 0.614 |
| 0.0860 | 0.395 | 0.0523 | 0.341 | 0.0660 | 0.364 | 0.0850 | 0.398 | 0.0183 | 0.609 |
| 0.0933 | 0.392 | 0.0573 | 0.348 | 0.0700 | 0.369 | 0.102 | 0.391 | 0.0202 | 0.596 |
| 0.0958 | 0.389 | 0.0600 | 0.341 | 0.0754 | 0.368 | 0.103 | 0.388 | 0.0228 | 0.591 |
| 0.105 | 0.388 | 0.0625 | 0.336 | 0.0817 | 0.366 | 0.107 | 0.389 | 0.0239 | 0.590 |
| 0.115 | 0.382 | 0.0660 | 0.337 | 0.0880 | 0.361 | 0.134 | 0.378 | 0.0239 | 0.589 |
| 0.137 | 0.366 | 0.0734 | 0.334 | 0.0980 | 0.346 | 0.146 | 0.351 | 0.0247 | 0.593 |
| 0.143 | 0.357 | 0.0792 | 0.334 | 0.115 | 0.333 | 0.148 | 0.350 | 0.0279 | 0.583 |
| 0.173 | 0.347 | 0.0825 | 0.330 | 0.118 | 0.323 | 0.166 | 0.357 | 0.0295 | 0.581 |
| 0.184 | 0.342 | 0.0909 | 0.329 | 0.131 | 0.319 | 0.169 | 0.341 | 0.0333 | 0.582 |
| 0.216 | 0.335 | 0.100 | 0.321 | 0.132 | 0.320 | 0.171 | 0.343 | 0.0365 | 0.565 |
| 0.225 | 0.324 | 0.105 | 0.319 | 0.157 | 0.297 | 0.191 | 0.342 | 0.0365 | 0.573 |
| 0.249 | 0.318 | 0.110 | 0.319 | 0.174 | 0.286 | 0.192 | 0.337 | 0.0381 | 0.570 |
| 0.290 | 0.311 | 0.117 | 0.312 | 0.176 | 0.299 | 0.207 | 0.329 | 0.0382 | 0.569 |
| 0.338 | 0.300 | 0.131 | 0.304 | 0.196 | 0.260 | 0.211 | 0.321 | 0.0405 | 0.571 |
| 0.405 | 0.269 | 0.150 | 0.292 | 0.223 | 0.253 | 0.213 | 0.330 | 0.0456 | 0.563 |
| 0.450 | 0.248 | 0.175 | 0.283 | 0.234 | 0.247 | 0.218 | 0.329 | 0.0521 | 0.558 |
| 0.579 | 0.202 | 0.191 | 0.274 | 0.257 | 0.232 | 0.226 | 0.321 | 0.0557 | 0.558 |
| 0.675 | 0.156 | 0.232 | 0.245 | 0.261 | 0.231 | 0.227 | 0.318 | 0.0636 | 0.546 |
| 0.810 | 0.128 | 0.236 | 0.256 | 0.321 | 0.204 | 0.246 | 0.309 | 0.0763 | 0.541 |
| | | 0.250 | 0.242 | 0.367 | 0.200 | 0.284 | 0.301 | | |
| | | 0.271 | 0.232 | 0.428 | 0.180 | 0.314 | 0.287 | | |
| | | 0.295 | 0.217 | 0.514 | 0.147 | 0.319 | 0.287 | | |
| | | 0.308 | 0.207 | | | 0.341 | 0.283 | | |
| | | 0.322 | 0.208 | | | 0.361 | 0.271 | | |
| | | 0.361 | 0.185 | | | 0.365 | 0.267 | | |
| | | 0.363 | 0.174 | | | 0.393 | 0.270 | | |
| | | | | | | 0.411 | 0.248 | | |
| | | | | | | 0.424 | 0.248 | | |
| | | | | | | 0.462 | 0.233 | | |
| | | | | | | 0.507 | 0.217 | | |
| | | | | | | 0.562 | 0.184 | | |
| | | | | | | 0.630 | 0.163 | | |
| | | | | | | 0.716 | 0.132 | | |
| | | | | | | 0.830 | 0.100 | | |

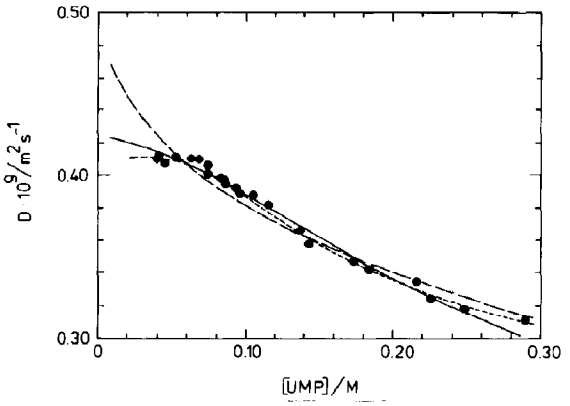


Fig. 5. The same data as in fig. 3, but for UMP.

cantly better agreement between experiment and simulation than Model 3, as evident from a visual inspection of figs. 2–6. The general trend from simulations of Model 4 is that the equilibrium constant for dimerization is a magnitude lower than those for the higher aggregation steps (see table 3).

### 4.3. *Caffeine binding-in*

As evident in figs. 8–11, caffeine diffusion decreases at clearly different rates upon increasing the nucleotide concentration; a qualitative measure of the degree of binding. A quantitative analysis by the simplest-possible binding-in model is shown. The agreement with simulation is clearly

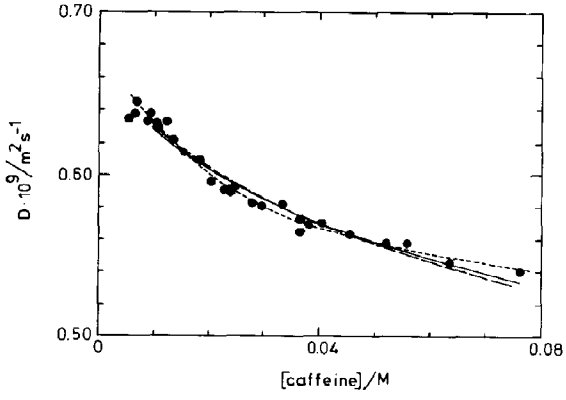


Fig. 6. The same data as in fig. 3, but for caffeine.

Table 2

Analysis of self-association of nucleotides and caffeine, according to a monomer-*n*-mer formation model

$D_1$ and $D_n$ are the fitted monomer and *n*-mer diffusion coefficients, respectively. $D_{n,est}$ is the hydrodynamically estimated *n*-mer diffusion coefficient, as obtained from the ASGEIG program. Diffusion coefficients are expressed in units of $10^{-9}$ $m^2$ $s^{-1}$. FOBJ represents the sum of the squares of the residuals and *K* refers to the single-step values in the aggregation.

| *n* | $D_1$ | $D_n$ | $D_{n,est}$ | $K/M^{-1}$ | FOBJ |
|---|---|---|---|---|---|
| AMP | | | | | |
| 2 | 0.4638 | 0.0000 | 0.3308 | 2.403 | 0.02670 |
| 3 | 0.4067 | 0.0000 | 0.2994 | 2.751 | 0.01025 |
| 4 | 0.3924 | 0.0072 | 0.2753 | 3.369 | 0.00647 |
| 5 | 0.3893 | 0.0780 | 0.2537 | 4.491 | 0.00656 |
| 6 | 0.3876 | 0.1130 | 0.2335 | 5.335 | 0.00691 |
| 8 | 0.3856 | 0.1457 | 0.1976 | 6.475 | 0.00783 |
| GMP | | | | | |
| 2 | 0.3932 | 0.0000 | 0.3004 | 1.358 | 0.00978 |
| 3 | 0.3597 | 0.0000 | 0.2718 | 2.000 | 0.00329 |
| 4 | 0.3538 | 0.0919 | 0.2500 | 3.287 | 0.00362 |
| 5 | 0.3517 | 0.1380 | 0.2303 | 4.371 | 0.00428 |
| 6 | 0.3506 | 0.1611 | 0.2120 | 5.191 | 0.00495 |
| 8 | 0.3494 | 0.1823 | 0.1794 | 6.286 | 0.00610 |
| CMP | | | | | |
| 2 | 0.4617 | 0.0000 | 0.3608 | 1.088 | 0.01365 |
| 3 | 0.4285 | 0.0346 | 0.3265 | 1.879 | 0.00705 |
| 4 | 0.4237 | 0.1654 | 0.3002 | 3.419 | 0.00639 |
| 5 | 0.4213 | 0.2080 | 0.2766 | 4.484 | 0.00621 |
| 6 | 0.4199 | 0.2282 | 0.2546 | 5.237 | 0.00625 |
| 8 | 0.4181 | 0.2481 | 0.2154 | 6.262 | 0.00656 |
| UMP | | | | | |
| 2 | 0.4529 | 0.0000 | 0.3583 | 0.897 | 0.00457 |
| 3 | 0.4239 | 0.0546 | 0.3243 | 1.799 | 0.00208 |
| 4 | 0.4194 | 0.1861 | 0.2982 | 3.369 | 0.00159 |
| 5 | 0.4172 | 0.2274 | 0.2748 | 4.465 | 0.00130 |
| 6 | 0.4160 | 0.2474 | 0.2429 | 5.272 | 0.00114 |
| 8 | 0.4147 | 0.2681 | 0.2140 | 6.437 | 0.00103 |
| 10 | 0.4141 | 0.2784 | 0.1830 | 7.239 | 0.00104 |
| 12 | 0.4137 | 0.2846 | 0.1598 | 7.816 | 0.00110 |
| Caffeine | | | | | |
| 2 | 0.6803 | 0.4229 | 0.5787 | 15.016 | 0.00189 |
| 3 | 0.6550 | 0.4848 | 0.5237 | 26.805 | 0.00145 |
| 4 | 0.6470 | 0.5039 | 0.4815 | 33.749 | 0.00118 |
| 5 | 0.6436 | 0.5134 | 0.4437 | 38.894 | 0.00103 |
| 6 | 0.6419 | 0.5193 | 0.4083 | 43.022 | 0.00095 |
| 7 | 0.6406 | 0.5229 | 0.3755 | 46.066 | 0.00092 |
| 8 | 0.6399 | 0.5257 | 0.3456 | 48.755 | 0.00092 |
| 9 | 0.6393 | 0.5276 | 0.3189 | 50.783 | 0.00093 |
| 10 | 0.6390 | 0.5292 | 0.2955 | 52.804 | 0.00095 |
| 12 | 0.6384 | 0.5315 | 0.2581 | 55.799 | 0.00099 |
| 14 | 0.6381 | 0.5331 | 0.2316 | 58.228 | 0.00103 |

Table 3

Association parameters pertaining to the ASGEIG, ASGE30 and ASGE11 computer simulations

$K_1$, $K_2$ and $K$ are the respective fitted association constants according to the association models indicated. $D_1$ represents the fitted monomer self-diffusion coefficient in units of $10^{-9}$ m$^2$ s$^{-1}$. FOBJ represents the sum of the squares of the residuals.

| ASGE30 | | | | ASGE11 | | |
|---|---|---|---|---|---|---|
| $K_1/M^{-1}$ | $K_2/M^{-1}$ | $D_1$ | FOBJ | $K/M^{-1}$ | $D_1$ | FOBJ |
| AMP | | | | | | |
| 0.324 | 9.49 | 0.392 | 0.00225 | 36.1 | 0.473 | 0.01045 |
| ASGEIG | | | | | | |
| GMP | | | | | | |
| 0.515 | 7.91 | 0.356 | 0.00386 | 48.2 | 0.447 | 0.02131 |
| CMP | | | | | | |
| 0.708 | 7.13 | 0.428 | 0.00677 | 27.9 | 0.503 | 0.02230 |
| UMP | | | | | | |
| 0.819 | 6.61 | 0.425 | 0.00201 | 22.6 | 0.493 | 0.00801 |
| Caffeine | | | | | | |
| 55.1 | 33.9 | 0.686 | 0.00228 | 24.1 | 0.666 | 0.00267 |

good. In general, the simulation leads to fitted parameters for the first binding steps that are very low, essentially zero. This need not necessarily reflect physical reality; it can be an evaluational effect, inherent in the relation between model and the actual aggregation process. The apparently very low binding constants for the lower steps merely reflect that inclusion of diffusion contributions from lower oligomers has an adverse effect on the least-squares sum in the simulation. As stated above, one should note that the binding-in model is based on parallel equilibria, not on coupled ones like Models 1–4. Therefore, the higher binding steps are essentially unaffected by the conditions in the lower steps.

The best fits in these simulations (see figs.

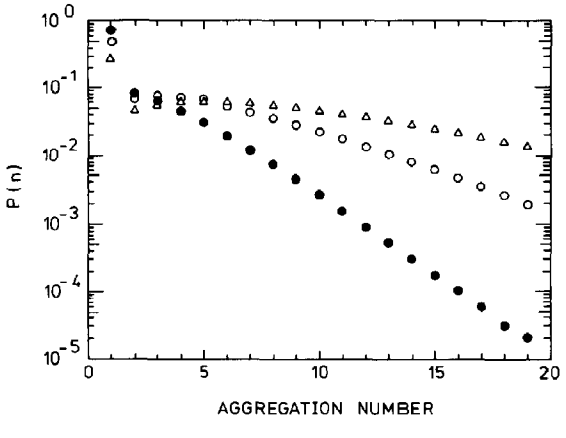


Fig. 7. The fraction of aggregate species vs. aggregation number at three different concentrations of CMP, calculated according to Model 4. (△) 0.411 M, (○) 0.207 M, (●) 0.102 M.

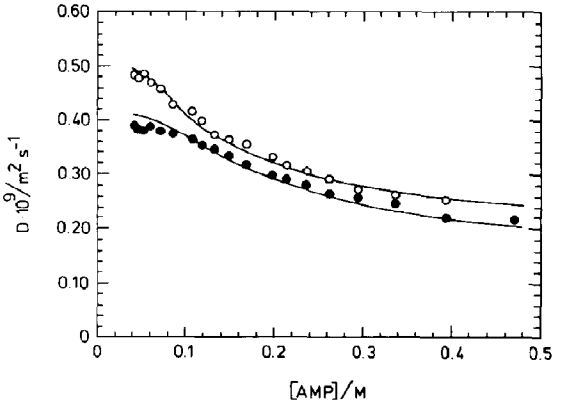


Fig. 8. The concentration dependence of the self-diffusion coefficients for the mixed system AMP+caffeine. The concentration of caffeine was kept constant at 0.025 M while the concentration of AMP was varied. (●) AMP, (○) caffeine. The solid curves represent the results of the ASGCOF computer simulation (see text).

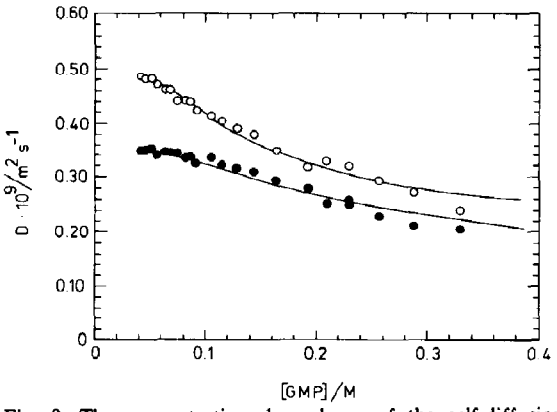


Fig. 9. The concentration dependence of the self-diffusion coefficients for the mixed system GMP + caffeine. The concentration of caffeine was kept constant at 0.022 M while the concentration of GMP was varied. (●) GMP, (○) caffeine. The solid curves represent the results of the ASGCOF computer simulation.

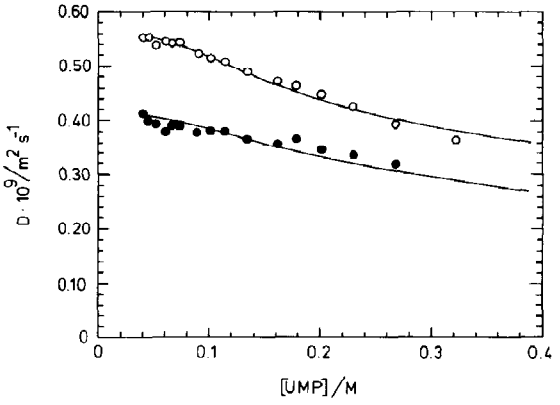


Fig. 11. The concentration dependence of the self-diffusion coefficients for the mixed system UMP + caffeine. The concentration of caffeine was kept constant at 0.022 M while the concentration of UMP was varied. (●), UMP, (○) caffeine. The solid curves represent the results of the ASGCOF computer simulation.

8–11) were obtained for zero values of the first and second binding constants and common values of 40.3, 62.8, 96.7 and 133 l mol$^{-1}$ for the higher binding steps for UMP, CMP, GMP and AMP, respectively. Caffeine binding is consequently much stronger to the purine nucleotides than to the pyrimidine types. This result would probably also be valid for more complex binding models.

## 5. Discussion

Nucleotide stacking has been thoroughly discussed in numerous papers in the past decades (see, e.g. refs. 1–24). The present investigation was not undertaken to provide a definite answer to the

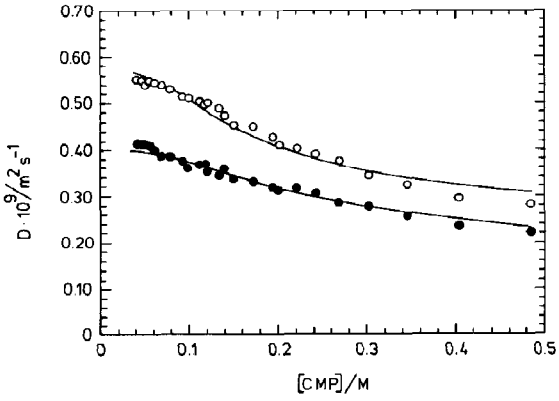


Fig. 10. The concentration dependence of the self-diffusion coefficients for the mixed system CMP + caffeine. The concentration of caffeine was kept constant at 0.015 M while the concentration of CMP was varied. (●), CMP, (○) caffeine. The solid curves represents the ASGCOF computer simulation.

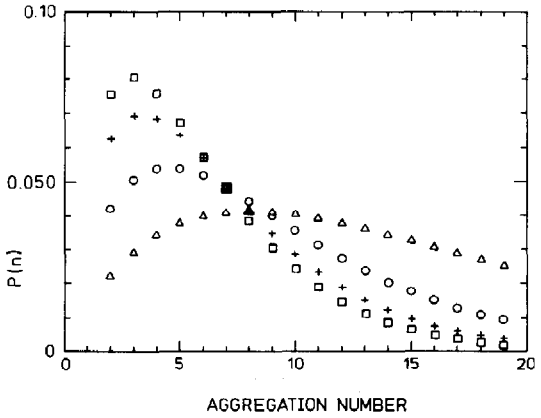


Fig. 12. The fraction of aggregate species vs. aggregation number at 0.25 M concentration, as based on the best fit of Model 4 to the observed data, using the ASGEIG program. (□) UMP, (+) CMP, (○) GMP, (△) AMP. The remaining fraction of mononucleotide in monomeric form at this concentration was 0.430, 0.423, 0.405 and 0.363, respectively.

details of the process. The main task has been to investigate the picture that emerges from self-diffusion information on nucleotide systems in aqueous solution, and the general feasibility of the new multicomponent self-diffusion approach for the quantification of aggregation phenomena of that kind. Of course, several extensions of the tested models are obvious and have previously been considered in the literature. At the present stage we have not found it justified to continue the search for better agreement between model and experiment because the meaningful limit has been reached (considering the precision of the data and the fits between model and experiment).

Association parameters reveal throughout that purine nucleotides exhibit stronger stacking ability than pyrimidines, in general agreement with previous results in the field [1,16]. It has also been concluded in the literature that the application of a cooperative aggregation model leads to better agreement with experimental results [3,48,49], as compared to the simple isodesmic model.

With regard to quantitative results and their relation to literature data, it should be pointed out that equilibrium constants for different aggregation models are, of course, not directly inter-comparable. In general, however, our numerical data are somewhat higher than previously published values for the same aggregation model. One fundamental reason for a general discrepancy is that, unlike activity-based data (like those obtained from osmometry and vapour pressure measurements, for example) the self-diffusion approach monitors concentration-related quantities, with no direct influence from non-ideal interactions in the thermodynamic sense. With regard to a comparison with NMR chemical shift-based data (cf. eq. 1) one should note that there is no a priori reason to expect chemical shifts for different oligomers to follow a regular and asymptotic trend with increasing aggregation number as normally done in the literature; especially for low aggregation numbers one cannot ignore chemical shift effects that may vary in a highly irregular fashion with the aggregation number, possibly making the analysis of the data rather incorrect. It is only for an aggregation process of the type 'monomer-single-*n*-mer' (Model 1) that chemical shifts become strictly meaningful, unbiased and justified iteration parameters. With regard to the formally identical measurement approach employed here, we feel that the relation between self-diffusion coefficients and aggregation number (eq. 7) rests on much safer ground. In the case of a monomer-single-*n*-mer model, of course, the self-diffusion coefficient of the *n*-mer again becomes an iteration parameter. We feel that the points just mentioned (as well as section 3.1 of the present paper) should be considered when relating the results in the present paper to data obtained by other techniques.

In summarizing the results, one finds that the only basic aggregation model that provides good agreement throughout between simulation and experiment with physically acceptable parameters is Model 4; indefinite aggregation with a low, first, binding constant *, suggesting that the aggregation process is characterized by cooperative interactions (Model 4 is also known in the literature as a Type III SEK or cooperative indefinite self-association model).

It is interesting to note that the UMP and CMP data also fit reasonably well to Model 1; a monomer-*n*-mer equilibrium. This does not necessarily contradict the main conclusion about Model 4; a closer examination of the simulations underlying Model 4 reveal that the oligomer distribution patterns for the nucleotides are significantly different, as illustrated in fig. 12. The UMP and CMP data analyses in terms of model 4 lead to relatively narrow aggregation number distributions, centred at a value not far from that found for Model 1.

Model 4 (eq. 6) also offers some self-consistency; a large fraction of monomer remains in

* Although any approach for the analysis of diffusion data on aggregating systems necessarily entails a number of assumptions, we must point out that the general conclusions of the present paper (quantitative and qualitative) are not in any way altered by moderate and physically reasonable changes in the model parameters. This was confirmed through a number of simulations. Also, the simulation results were found to be rather insensitive to the actual value of the obstruction correction factor (see above). The corrections are all physically justified; they were not introduced for the purpose of serving as additional fitting parameters.

solution (fig. 7). Exchange and equilibration can therefore occur with that monomeric fraction in all higher steps, as required in the physical realization of the Model. The exchange process can be visualized to occur by exchange of monomers at the ends of the stack.

## Acknowledgements

This work has been financially supported by the Swedish Natural Sciences Research Council. We thank Sture Forsén for stimulating discussions as well as for suggesting the current investigation. We are also grateful to the referees for constructive criticism.

## References

1 P.O.P. Ts'o, in: Basic principles of nucleic acid chemistry, vol. 1, ch. 6, (Academic Press, New York, 1974).
2 F.E. Evans and R.H. Sarma, Biopolymers 13 (1974) 2117.
3 E. Plesiewicz, E. Stepien, K. Bolewska and K.L. Wierzchowski, Biophys. Chem. 4 (1976) 131.
4 H. Fritzche, I. Petri, H. Schütz, K. Weller, P. Sedmera and H. Lang, Biophys. Chem. 11 (1980) 109.
5 R. Buchet and C. Sandorfy, J. Phys. Chem. 87 (1983) 275.
6 W.E. Ferguson, C.M. Smith, E.T. Adams, Jr and G.T. Barlow, Biophys. Chem. 1 (1974) 325.
7 Y. Matsuoka, B. Nordén and T. Kurucsev, J. Phys. Chem. 88 (1984) 971.
8 T. Solie and J.A. Schellman, J. Mol. Biol., 33 (1968) 61.
9 P.O.P. Ts'o and S.I. Chan, J. Am. Chem. Soc. 86 (1964) 4176.
10 A.A. Maevsky and B.I. Sukhorukov, Nucleic Acids Res. 8 (1980) 3029.
11 M.P. Heyn and R.P. Bretz, Biophys. Chem., 3 (1975) 35.
12 P. Hemmes, A. Mayevski, V.A. Buckin and A.P. Sarvazyan, J. Phys. Chem. 84 (1980) 699.
13 A. Skauge and P.I. Vestues, Acta Chem. Scand. A37 (1983) 47.
14 S.I. Chan, M.P. Schweizer, P.O.P. Ts'o and G.K. Helkamp, J. Am. Chem. Soc. 86 (1964) 4182.
15 M.P. Schweizer, A.D. Brown, P.O.P. Ts'o and D.P. Hollis, J. Am. Chem. Soc. 90 (1968) 1042.
16 K.H. Scheller, F. Hofstetter, P.R. Mitchell, B. Prijs and H. Sigel, J. Am. Chem. Soc. 103 (1981) 247.
17 K.H. Scheller and H. Sigel, J. AM. Chem. Soc. 105 (1983) 5891.
18 J.-L. Dimicoli and C. Hélene, J. Am. Chem. Soc. 95 (1973) 1036.
19 H.M. Schwarz, M. MacCoss and S.S. Danyluk, J. Am. Chem. Soc. 105 (1983) 5901.
20 M. Borzo, C. Detellier, P. Laszlo and A. Paris, J. Am. Chem. Soc. 102 (1980) 1124.
21 C. Detellier and P. Laszlo, J. Am. Chem. Soc. 102 (1980) 1135.
22 C.L. Fisk, E.D. Becker, H.T. Miles and T.J. Pinnavaia, J. Am. Chem. Soc. 104 (1982) 3307.
23 W.E. Egan, J. Am. Chem. Soc. 98 (1976) 4091.
24 S.B. Petersen, J.J. Led, E.R. Johnston and D.M. Grant, J. Am. Chem. Soc. 104 (1982) 5007.
25 J.A. Pople, J. Chem. Phys. 24 (1956) 111.
26 C.J. Jameson, in: Specialist periodical reports, nuclear magnetic resonance, vol. 12, ch. 1 (The Chemical Society, London, 1983).
27 P. Lazaretti and R. Zanasi, J. Chem. Phys. 75 (1981) 5019.
28 K. Nishida, Y. Ando and H. Kawamura, Colloid Polym. Sci., 261 (1983) 70.
29 R.L. Vold, J.S. Waugh, M.P. Klein and D.E. Phelps, J. Chem. Phys. 48 (1968) 3931.
30 E.O. Stejskal and J.E. Tanner, J. Chem. Phys. 42 (1965) 288.
31 E.L. Hahn, Phys. Rev. 80 (1950) 580.
32 P. Stilbs and M.E. Moseley, Chem. Scr. 15 (1980) 176, 215.
33 P. Stilbs, J. Colloid Interface Sci. 87 (1982) 385.
34 P. Stilbs and B. Lindman, J. Phys. Chem. 85 (1981) 2587.
35 P. Stilbs, J. Colloid Interface Sci. 89 (1982) 547.
36 B. Lindman, P. Stilbs and M.E. Moseley, J. Colloid Interface Sci. 83 (1981) 569.
37 P. Stilbs, K. Rapacki and B. Lindman, J. Colloid. Interface Sci. 95 (1983) 583.
38 B. Lindman, T. Ahlnäs, O. Söderman, H. Walderhaug, K. Rapacki and P. Stilbs, Faraday Disc. (Concentrated Colloidal Dispersion) 76 (1983) 317.
39 P. Stilbs, G. Arvidson and G. Lindblom, Chem. Phys. Lipids 35 (1984) 309.
40 P. Stilbs and B. Lindman, J. Magn. Resonance 48 (1982) 132.
41 B. Lindman, M.-C. Puyal, N. Kamenka, R. Rymdén and P. Stilbs, J. Phys. Chem. 88 (1984) 5048.
42 R. Rymdén, J. Carlfors and P. Stilbs, J. Inclusion Phenomena 1 (1983) 157.
43 J.H. Wang, J. Am. Chem. Soc., 76 (1954) 4755.
44 P. Stilbs, to appear.
45 B.-O. Persson, T. Drakenberg and B. Lindman, J. Phys. Chem. 83 (1979) 3011.
46 P Chun, Biophys. Chem. 2 (1974) 170.
47 C.R. Cantor and P.R. Schimmel, Biophysical Chemistry (W.H. Freeman, San Francisco, 1980), p. 560.
48 H. Sterk and H. Gruber, J. Am. Chem. Soc. 106 (1984) 2239.
49 F. Garland and S.H. Christian, J. Phys. Chem. 79 (1975) 1247.