Matrix-assisted diffusion-ordered NMR spectroscopy with an invisible, tuneable matrix


Mixtures of species of similar sizes can be analysed with $^{1}$H NMR using diffusion-ordered spectroscopy (DOSY), by adding a co-solute with which different analytes interact to different extents (‘matrix-assisted DOSY’). A matrix containing aqueous perfluoroctanoate micelles has the dual advantages of invisibility to $^{1}$H NMR, and interactions tuneable by adjusting the pH.

Introduction

Diffusion-ordered spectroscopy\textsuperscript{4–3} is a powerful tool for the analysis of intact mixtures by NMR, but has two significant limitations: the signals of different species can normally only be resolved if the species differ significantly in size, and hence in rate of diffusion; and then only if the individual NMR signals do not overlap. The requirement that species have different sizes can be circumvented by adding a co-solute to which different analytes bind to different extents, so that their average diffusion coefficients in the matrix of solvent and co-solute are different. This matrix-assisted DOSY approach (MAD, also known as chromatographic DOSY) has been widely applied,\textsuperscript{4–18} but suffers from the disadvantage that the NMR signals of the co-solute can overlap with and obscure analyte signals. This is a particular problem with one of the most successful classes of MAD co-solute, micellar surfactants.\textsuperscript{9,11,12} These typically have long alkyl chains and hence have wide footprints in the aliphatic region of the $^{1}$H spectrum. Perfluorinated micromulsions have been successfully used previously for MAD/NMR chromatography.\textsuperscript{9} Here it is shown that perfluorinated micelles, which are also invisible to $^{1}$H NMR, can be very effective at resolving the signals of test mixtures of amino acids, and that the versatility of MAD can be greatly enhanced by varying pH to modulate charge–charge interactions between analytes and co-solute.

A variety of different types of species have been used to modify the diffusion matrix in MAD. Some of the earliest experiments used solids,\textsuperscript{13–16} either in slurries or as suspensions. These have the advantage of comparative invisibility, with little if any $^{1}$H signal visible in DOSY experiments, but suffer from poor resolution (even when magic angle spinning is used) and/or sample instability (because of slow sedimentation). Polymer co-solutes can be chosen to have few signals,\textsuperscript{17} but in general the most successful and versatile co-solutes, such as micellar surfactants and cyclodextrins, tend to have big spectral footprints. One way to avoid this problem is to use a high level of deuteriation,\textsuperscript{18} but this tends to be prohibitively expensive.

Perfluorosurfactants such as sodium perfluoroctanoate (PFO) are an attractive alternative, as they are cheap and readily available, and having no $^{1}$H signal provide an ‘invisible’ matrix\textsuperscript{9,19} (just as the signals of conventional hydrogenated surfactants are invisible in $^{19}$F matrix-assisted DOSY\textsuperscript{20,21}). A less obvious advantage is that the forces that drive analyte–matrix association can be both more predictable and easier to control than is often the case with alkyl chain surfactants. Common surfactants such as sodium dodecyl sulfate (SDS) and cetyl trimethylammonium bromide (CTAB),\textsuperscript{22,23} and nonionic surfactants such as the Brij family,\textsuperscript{24} have been extensively used in matrix-assisted DOSY for the separation of species on the basis of differences in hydrophobicity and amphiphilicity. Very recently, it has been shown that very high ionic strength CTAB solutions can be used as invisible matrices if strong T$_2$ weighting is used in DOSY experiments.\textsuperscript{25} However, analytes such as the dibydroxybenzenes, that show strong differential interactions with SDS and CTAB micelles, show negligible binding to PFO under similar conditions. Because perfluorosurfactant micelles have lower affinity for hydrophobic species than alkyl chain surfactants,\textsuperscript{25} their binding of analytes with ionisable groups tends to be dominated by charge–charge interactions. Adjusting the pH then allows this binding to be varied, tuning the matrix to achieve the separation required.

Here the use of PFO micelles at different pHs to separate the signals of aqueous mixtures of amino acids is demonstrated, and the relevant equilibria are analysed to rationalise the binding seen. PFO micelles should allow charge-based separation of the signals of a wide range of ionic analytes, and those separations will be tuneable in just the same way for other amphiprotic analytes.
Results and discussion

Fig. 1 compares the DOSY spectrum of an approximately neutral aqueous (D$_2$O) mixture of three amino acids with the spectrum measured in the presence of 100 mM sodium PFO at similar pH. As expected, all three amino acids show similar diffusion coefficients in the absence of PFO, and the DOSY spectrum is unresolved in the diffusion dimension. This remains the case over a wide pH range. The diffusion dimension is also unresolved in the presence of PFO at high pH, but as the pH is reduced each amino acid in turn shows a fall in measured diffusion coefficient, as it begins to bind to the PFO micelles, until at very low pH all three diffuse more slowly. For this mixture good resolution in the diffusion domain, with the signals of the three different amino acids clearly separated, is obtained around neutral pH, as shown in Fig. 1.

This behaviour can be rationalised quite straightforwardly in terms of the Coulomb interaction. At high pH the net charge on all the amino acid species is negative, and the charge–charge interactions between these species and the negatively charged PFO micelles tend to discouragement binding. As the pH is lowered, the charge on the amino acids reduces and they become neutral zwitterions, but binding remains weak. Only when the net charge on the amino acid species is positive do the charge–charge interactions with the negative micelles become favourable and lead to substantial binding. The sequence in which the individual amino acids begin to bind as the pH is lowered is governed by the ordering of the individual pK$_a$S, notably those of the side chains.

The effect of charge–charge driven binding on amino acid diffusion and speciation can be modelled quantitatively, given reasonable approximations, by solving the interlocking equilibria between free amino acid species, aqueous protons, and micelle-bound species to determine the fraction $x_i$ of a given amino acid that is bound to micelles. The observed amino acid diffusion coefficient $D_{av}$ is then given by Lindman’s law:

$$D_{av} = D_f (1 - x_b) + D_b x_b$$  

(1)

where the bound and free diffusion coefficients are $D_b$ and $D_f$ respectively. The former is well approximated by the micellar diffusion coefficient $D_m$, which in turn can be deduced from the diffusion coefficient $D_b$ measured for PFO at a concentration [PFO] that is above the critical micelle concentration cmc, using the diffusion coefficient $D_{PFO}$ of free PFO (as measured at concentrations below the cmc):

$$D_m = (D_b [PFO] - D_{PFO} cmc) / ([PFO] - cmc)$$  

(2)

An amino acid with $n$ protonation equilibria will have $n + 1$ different net charge states, e.g. $-1$, $0$, and $+1$ for valine, which has only amino and carboxylic protonation equilibria. The bound and free concentrations $c_b$ and $c_f$ for a given charge state $i$ are linked as follows:

$$K_a^i = c_f [D]^i / c_i^{i+1}$$  

(3)

$$K_b^i = c_b ([S_m] / c_f^i)$$  

(4)

$$x_b^i = c_b / c_{tot}$$  

(5)

where $[D]^i$ is the hydrogen ion concentration, $[S_m]$ is the concentration of micellar surfactant, $x_b^i$ is the bound fraction of species $i$, and $c_{tot}$ is the total amino acid concentration. The total amino acid bound fraction $x_b$ in eqn (1) is then just the sum over all $x_b^i$, and the free fraction $1 - x_b$.

Using the measured amino acid diffusion coefficients in the absence of PFO, the experimental micellar diffusion coefficient from eqn (2), and literature values for the pK$_a$S of all the ionisable groups, it is possible to fit measured amino acid diffusion coefficients in the presence of PFO to obtain estimates of the binding constants of the individual charged states of the different amino acids. A number of simplifying assumptions are needed. These include ideal behaviour, a single binding constant for each amino acid species that is independent of the amount bound and of any other species that may be bound, only one significant species per charge state, a $D_m$ that is unaffected by pH or by amino acid binding, a $D_{free}$ that is substantially independent of pH, and pH measurement that is unaffected by the solutes present.

The effect of using D$_2$O rather than H$_2$O may be allowed for by using the apparent pH measured in D$_2$O solution using a standard glass electrode, pH$_a$, in place of the deuterio analogue of pH, pD, and adjusting the apparent pK$_a$ to allow for the change in solvent.

In an amino acid solution containing PFO micelles, three types of equilibria occur simultaneously, see Fig. 2. The first is amino acid protonation. Depending on the type of side chain, one amino acid may have up to four different protonation states of net charge $z_i$, with up to three corresponding equilibrium constants $K_a^i$. The second set of up to four equilibria, with constants $K_b^i$, is between the free and the micelle-bound forms of the individual amino acid charge states. The third and final equilibrium is that between free and micellar NaPFO, which requires that the concentration of NaPFO be greater than its critical micelle concentration (cmc, 31 mM at 25 °C (ref. 29)).

To investigate the nature and extent of the binding of different amino acid species to PFO micelles as a function of pH, two different mixtures were studied, containing respectively

![Fig. 1](https://example.com/fig1.png)

400 MHz $^1$H DOSY spectra of approximately neutral mixtures of 1 mM each of histidine, lysine, glutamic acid, and sodium trimethylsilylpropanoate-d$_4$ in D$_2$O, in the absence (left) and in the presence (right) of 100 mM sodium perfluorooctanoate. Apparent pH values, corrected for the use of D$_2$O, were 7.4 and 6.9 respectively.
1 mM each of histidine (His), glutamic acid (Glu) and lysine (Lys) (mixture 1), and 1 mM each of phenylalanine (Phe), tyrosine (Tyr) and valine (Val) (mixture 2). Samples of mixtures 1 and 2 were made up in D2O containing 1 mM sodium trimethylsilylpropanoate-d3 (TMSP) as reference, with and without 100 mM sodium perfluorooctanoate (NaPFO). The choice of amino acid concentration here is a compromise between the need for sufficient signal-to-noise ratio to allow the determination of diffusion coefficients in a reasonable time, and the need to keep the ratio of PFO to total amino acid concentration high to minimise any effect of amino acid binding on the binding constants $K_i$ and on the micellar diffusion coefficient $D_m$. Similarly, the number of amino acids per mixture is a compromise between the need to resolve sufficient signals of individual amino acids to allow reliable determination of diffusion coefficients, and the need to minimise the number of samples and the duration of the experiments required.

Fig. 3 summarises the results of measurements of diffusion coefficients as a function of pH for the four samples studied. For each amino acid the experimental diffusion coefficient in the presence and in the absence of PFO is plotted as a function of pH, together with solid lines showing the results of fitting and the literature values obtained for acid dissociation constants. If the equilibria were unperturbed by the binding, the points of inflection in the red curves would coincide with the $pK_a$ values. Where the binding is weak, at high pH, this is a good approximation (see e.g. Fig. 3d and e). However where a significant fraction of the amino acid is bound, as at low pH in all the species studied except glutamic acid and tyrosine, the removal of much of the more protonated species from free solution shifts the overall equilibrium, so that the points of inflection lie to the right (to higher pH) of the $pK_a$ values.

The results of Fig. 3 provide significant reassurance about the approximations used and the probity of the model. The experimental diffusion coefficients measured for PFO show little variation with pH, confirming that within the pH range studied the micellisation is substantially unperturbed either by pH, or by amino acid binding at the low concentrations used. Indeed the quality of the fits seen in Fig. 3 is remarkably good given the small number of variable parameters used (just the three or four binding constants, not all of which are significant), and the assumptions and approximations involved.
Fig. 3  (Upper parts of a to f) Experimental results for measurements of diffusion coefficient \( D \) (dots with vertical error bars) as a function of measured pH (pH*) for the six amino acids shown in samples of mixtures 1 and 2 with (blue) and without (purple) 100 mM NaPFO. Solid horizontal purple lines show the averages of \( D \) measurements in the absence of NaPFO, and dashed horizontal purple lines those averages corrected for the estimated effects of the obstruction caused by surfactant in the samples containing NaPFO. Black dots with vertical error bars indicate the diffusion coefficients measured by \(^{19}\)F NMR for PFO, with their averages shown by solid horizontal black lines, and the micellar diffusion coefficients \( D_m \) calculated from eqn (2) indicated by dashed horizontal black lines. Error bars indicate plus and minus twice the standard error estimated in fitting, and do not take into account systematic errors. Literature values for \( pK_a \) (ref. 27) are shown by dashed vertical black lines. The results of fitting the experimental amino acid diffusion coefficients \( D \) to eqns (A1) and (A2) for the equilibria shown in Fig. 1, using literature values for \( pK_a \) corrected for D_{2}O solvent^{28} and treating the binding constants \( K_b \) as variable parameters, are shown as solid red lines. (Lower parts of a to f) Speciation plots showing bound (thick lines) and total (thin lines) mole fractions for the three or four different protonation states of each amino acid, with net charge becoming more positive in the order green, blue, purple, red, calculated using the fitted values for \( K_b \).
Table 1  Binding constants (in the form $pK_a = -\log_{10} K_a$) obtained by fitting, and literature dissociation constants,\textsuperscript{27} for the components of samples 1 and 2. Error ranges quoted are twice the standard errors estimated in the least squares fitting, and do not take into account systematic errors.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Species</th>
<th>$pK_a^1$</th>
<th>$pK_a^2$</th>
<th>$pK_a^3$</th>
<th>$pK_b^1$</th>
<th>$pK_b^2$</th>
<th>$pK_b^3$</th>
<th>$pK_b^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Glu</td>
<td>9.58</td>
<td>4.15</td>
<td>2.16</td>
<td>$-0.53 \pm 0.13$</td>
<td>$-0.22 \pm 0.18$</td>
<td>$-0.18 \pm 0.82$</td>
<td>\textsuperscript{a}</td>
</tr>
<tr>
<td>1</td>
<td>His</td>
<td>9.09</td>
<td>6.04</td>
<td>1.70</td>
<td>$-0.36 \pm 0.14$</td>
<td>$-0.28 \pm 0.30$</td>
<td>$-0.20 \pm 0.11$</td>
<td>\textsuperscript{a}</td>
</tr>
<tr>
<td>1</td>
<td>Lys</td>
<td>10.67</td>
<td>9.16</td>
<td>2.15</td>
<td>$-0.36 \pm 0.16$</td>
<td>$-0.86 \pm 0.13$</td>
<td>$-1.51 \pm 0.04$</td>
<td>$-3.21 \pm 0.20$</td>
</tr>
<tr>
<td>2</td>
<td>Phe</td>
<td>9.09</td>
<td>2.18</td>
<td></td>
<td>$0.09 \pm 0.24$</td>
<td>$2.23 \pm 0.08$</td>
<td>$2.07 \pm 0.10$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Tyr</td>
<td>10.10</td>
<td>9.04</td>
<td>2.24</td>
<td>$-0.02 \pm 0.6$</td>
<td>$a$</td>
<td>$a$</td>
<td>$a$</td>
</tr>
<tr>
<td>2</td>
<td>Val</td>
<td>9.52</td>
<td>2.27</td>
<td></td>
<td>$a$</td>
<td>$-1.80 \pm 0.15$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} No statistically significant binding constant estimate found (estimated standard error $> 0.5$).

One of the most significant of these qualifications is the assumption that the amino acid ionisation equilibria can be approximated by using pH* values as a proxy for $[D^+]$ with appropriate adjustments to the $H_2O$ solution values for the acid dissociation constants $K_a$. The adjustment used is based on extensive measurements made at 0.1 mol ionic strength.\textsuperscript{28} At first sight this seems a good match to the experimental conditions used (dominated by the 0.1 mol dm$^{-3}$ NaPFO), but the effective ionic strength here is rather lower because two thirds of the NaPFO is in micellar form. This, and the potential for the surfactant to perturb the performance of the standard glass membrane pH electrode used, does introduce significant uncertainty and it is gratifying that only one dataset, that for histidine, shows significant systematic deviations between experimental data and fitted line, and that even these deviations disappear if the histidine side chain $pK_a$ is changed by 0.4 of a unit.

The two groups of amino acids used here were chosen because their similar sizes mean that they have very similar diffusion coefficients in free neutral solution, and hence pose a particular challenge for DOSY separation. The ability to achieve good diffusion separation by controlling the pH in PFO solution depends on the existence of a significant difference in $pK_a$. Thus Phe and Tyr above are more difficult to separate than His and Lys.

Experimental

All materials were available commercially and used as received from Sigma Aldrich and Cambridge Isotopes Laboratories, Inc. pH was adjusted by sequential addition of NaOD/DCl in D$_2$O. Diffusion coefficient measurements for $^1$H and $^{19}$F were carried out on a Varian INOVA 400 spectrometer at 298 K, using the Oneshot DOSY pulse sequence.\textsuperscript{29,32} Data were acquired with an array of 16 nominal gradient amplitudes, ranging from 3.0 to 27.0 G cm$^{-1}$, using 16 transients, 32 768 complex data points, a total diffusion-encoding gradient duration δ of 2 ms and a diffusion time Δ of 0.25 s. Reference deconvolution\textsuperscript{33,34} with a Gaussian target line shape, 5 or 6 Hz wide as appropriate, was chosen to give the best compromise between sensitivity and spectral resolution for $^1$H measurements. DOSY spectra were constructed using correction for the effective ionic strength here is rather lower because two thirds of the NaPFO is in micellar form. This, and the potential for the surfactant to perturb the performance of the standard glass membrane pH electrode used, does introduce significant uncertainty and it is gratifying that only one dataset, that for histidine, shows significant systematic deviations between experimental data and fitted line, and that even these deviations disappear if the histidine side chain $pK_a$ is changed by 0.4 of a unit.

Conclusions

In matrix-assisted DOSY, sodium perfluorooctanoate micelles provide an “invisible” matrix that can be tuned by adjusting the pH to allow the signals of amphiprotic analytes such as amino acids to be separated. The binding of the amino acids studied is dominated by the Coulomb interaction, suggesting that it should be possible to predict in advance the conditions required to allow the separation of signals of ionic species in a given mixture.

Appendix

Analytical expressions for amino acid diffusion coefficients

Eqns (1)–(5) can be assembled to yield analytical expressions for average amino acid diffusion coefficient $D_{av}$ in the presence of a total surfactant concentration $S$ and a deuterium ion concentration $D^-$. For the cases where there are two and three ionization constants respectively these are

$$D_{av} = \frac{D_0(S - c_m)(K_a K_b S + D^+(K_a K_b + D^+(K_a + D^+)) + D_0(K_a K_b S + D^+(K_a + D^+)))}{K_a K_b (S - c_m) + 1} + D^+(K_a K_b (S - c_m) + 1) + D^+(K_a S - c_m) + 1)} \tag{A1}$$

and

$$D_{av} = \frac{D_0(S - c_m)(K_a K_b S + D^+(K_a K_b + K_a + D^+(K_a + D^+))) + D_0(K_a K_b S + D^+(K_a + D^+)))}{K_a K_b (S - c_m) + 1} + D^+(K_a K_b (S - c_m) + 1) + D^+(K_a S - c_m) + 1)} \tag{A2}$$
where the successive ionization constants are $K_{a1}$ to $K_{a2}$ ($K_{a3}$) and the binding constants $K_{b1}$ to $K_{b3}$ ($K_{b4}$) for the two (three) ionization step case, and the surfactant critical micelle concentration is cmc.

**Acknowledgements**

This work was supported by the EPSRC (grant numbers EP/E05899X/1 and EP/H024336/1). AH thanks the Mexican Scientific and Technology Council (CONACyT) for financial support. MP thanks the Analytical Chemistry Trust Fund for a summer studentship (ACSS 12/012).

**Notes and references**