Published in final edited form as:

J Phys Chem B. 2010 April 29; 114(16): 5412–5419. doi:10.1021/jp9097827.

Observation of Pyridine Aggregation in Aqueous Solution Using Neutron Scattering Experiments and MD Simulations

Philip E. Mason[†], George W. Neilson[‡], Christopher E. Dempsey^{**}, David L. Price⁺, Marie-Louise Saboungi[#], and John W. Brady^{*,†}

[†]Department of Food Science Stocking Hall, Cornell University Ithaca, NY 14853. USA

[‡]H.H.Wills Physics Laboratory University of Bristol, BS8 1TL, UK

^{**}Department of Biochemistry, School of Medical Sciences, University Walk, Bristol BS8 1TD, UK

[#]Centre de Recherche sur la Matière Divisée 1 bis rue de la Férollerie 45071 Orléans, FRANCE

⁺Conditions Extrêmes et Matériaux: Haute Température et Irradiation 1D avenue de la Recherche Scientifique 45071 Orléans, FRANCE

Abstract

Neutron Diffraction with Isotopic Substitution (NDIS) experiments have been used to examine the structuring of aqueous solutions of pyridine. A new method is described for extracting the structure factors relating to intermolecular correlations from neutron scattering experiments on liquid solutions of complex molecular species. This approach performs experiments at different concentrations and exploits the intramolecular coordination number concentration invariance (ICNCI) to separate the internal and intermolecular contributions to the total intensities. The ability of this method to deconvolute molecular and intermolecular correlations is tested and demonstrated using simulated neutron scattering results predicted from molecular dynamics simulations of aqueous solutions of the polyatomic solute pyridine in which the inter- and intramolecular terms are known. The method is then implemented using neutron scattering measurements on solutions of pyridine. The results confirm that pyridine shows a significant propensity to aggregate in solution and demonstrate the prospects for the future application of the ICNCI approach to the study of large polyatomic solutes using next-generation neutron sources and detectors.

Introduction

Many polyatomic organic molecules have both hydrophobic and hydrophilic functional groups, and often impose significant structure on surrounding water molecules when they are in aqueous solutions.^{1,2} Thus, even if they are quite soluble in water, the structuring imposed on nearby water molecules by their non-polar surface components may be unfavorable, promoting limited aggregation beyond what might be expected from random encounters. Such association in water has recently been reported for a number of small organic solutes, both in experiments and in computer simulations.³⁻⁶ The nature of these interactions depends on the topological details of each solute molecule. In some cases such interactions can be through hydrophobic association, which can be particularly interesting for the case of planar molecules, as was found for guanidinium ions, while in other cases the association can be dominated by strong polar interactions, such as preferential hydrogen bonding. The specific details of such interactions are thought in some cases to play a role in

* author to whom correspondence should be addressed, jwb7@cornell.edu. .

the Hofmeister ranking of ions in terms of their effects on protein solubility and conformational stability.

The aromatic heterocyclic compound pyridine is of particular interest in this respect. Unlike benzene, pyridine is soluble in water in all proportions, even though the majority of the molecule is made up of non-polar C-H functional groups as in benzene. While the nitrogen atom of pyridine, lacking a proton, is unable to make a hydrogen bond as a donor, its lone pair can serve as a hydrogen bond acceptor, making it an interesting probe of water-solute hydrogen bonding in aqueous solutions. In addition, its relatively simple and symmetric planar structure make it a more tractable subject for study than more complex solutes.

Although several experimental techniques for studying structure in aqueous solutions are available,⁷⁻⁹ one of the most powerful methods has been neutron diffraction experiments. However, in solutions as complex as pyridine in water, the structure factors produced by these experiments are not only difficult to obtain but are difficult to interpret as well. This difficulty arises in part because the measured scattering intensities are the result of the summed scattering of all atoms in the system. For example, even for a simple solution like NaCl in water, there are four types of nuclei, O, H, Na, and Cl, producing 10 different pairwise radial correlation functions, so that a traditional neutron diffraction measurement gives a weighted summation of all 10 of these distribution functions. As a result, the collected data present formidable challenges to interpreting the scattering in terms of the individual contributions. As the number of elements in the sample gets larger, these correlations become increasingly more difficult to untangle.

The first order difference method, neutron diffraction with isotopic substitution (NDIS), was developed to simplify the interpretation of the structural data obtained from neutron scattering experiments.^{10,11} This approach involves performing paired experiments in which one atom type is replaced by an isotope of the same element with a different neutron scattering length, under the important assumption that the overall solution structure is not altered by such a substitution. Subtracting the structure factors for the two experiments then eliminates any correlations that do not involve the substitution-labeled atom. In the NaCl example cited above, a first difference experiment could be performed using a ^{35/37}Cl substitution in the salt, reducing the 10 correlations in the scattering to just 4: Cl-H, Cl-O, Cl-Cl, and Cl-Na, with the prefactors for the ion-ion terms being so small that they can often be ignored. Such experiments have been instrumental in the study of the structure of solutions of many solutes such as rare gases,¹² mono-atomic ions,¹³⁻¹⁷ or approximately spherical molecular solutes,¹⁸ and even peptides.¹⁹ For biological applications the NDIS method is also limited by the possible isotopes available for substitution, since the differences in the scattering lengths of the isotopes of oxygen and carbon are too small to be exploited. For pyridine, however, ^{15/14}N substitution offers one possible choice. H/D substitution of non-exchangeable hydrogen atoms also makes a good subject for NDIS experiments in the pyridine case due to the high contrast of these nuclei. This approach could be easily extended to other biological molecules due to the fact that a wide range of deuterated compounds are relatively cheap and readily available due to their common usage in NMR spectroscopy.

In recent years, NDIS methods have been applied to much more complex molecular species, including *tert*-butyl alcohol,^{20,21} ethanol,²² phenol,²² D-glucose,²² amino acids,²³⁻²⁵ and polymer electrolytes.^{26,27} However, the level of detail that can be extracted from these experiments, particularly concerning intermolecular structuring, is still limited by the extreme complexity of the scattering.

While the radial correlations functions extracted from NDIS experiments are still very complicated, many of their features can be assigned and interpreted with the assistance of molecular dynamics (MD) simulations of the same systems under similar conditions, as has recently been done for labeled D-glucose²⁸ and D-xylose samples.^{29,30} Unfortunately, for complex polyatomic solutes, the structure factors and corresponding radial distribution functions produced by NDIS experiments remain too complex to be interpreted, since they contain contributions both from intramolecular correlations with the other atoms of the solute as well as intermolecular contributions from correlations with the water solvent and other solute molecules. Even though the solute-water correlations constitute the majority of the scattering prefactors, they are dominated at short distances in the total radial correlation functions by intramolecular terms, which obscure the former. The separation of the intermolecular contributions to this radial distribution function from the more pronounced intramolecular contributions that would allow the characterization of the solution structure around the labeled position has proven to be difficult.

This paper outlines a procedure for effecting this separation based on the fact that the multiplicative prefactors for the intermolecular and intramolecular contributions to the total radial distribution function have different concentration dependences. That is, while the intermolecular contribution varies with concentration, the intramolecular contribution remains constant. Exploiting this intramolecular coordination number concentration invariance (ICNCI), a pair of first order NDIS experiments of natural and perdeuterated pyridine in aqueous solution (Figure 1) were carried out at two different concentrations and used to extract the intermolecular radial distribution function around these hydrogen/deuterium atoms. MD simulations of the system under the same conditions were carried out to guide in the interpretation of the peaks. This is the first report utilizing the concentration invariance of molecular correlations to separate out contributions related to intramolecular and intermolecular structure.

Procedures

Experimental

Neutron diffraction experiments were carried out at 20° C on aqueous solutions of pyridine at two different concentrations. The samples of *d*5 - pyridine (>99% enrichment) and natural abundance pyridine were purchased from Sigma-Aldrich. Two aqueous solutions of pyridine, at concentrations of 1.0 and 5.0 molal, were prepared by direct gravimetric addition of pyridine (dried over a common batch of molecular sieve) to a common batch of null water (in which the proportions of H₂O and D₂O are such that their scattering exactly cancels). The gravimetric accuracy of these solutions (by addition using a microsyringe) was better than ±1mg. A 1 ml aliquot of each sample was then placed in the neutron cell (a null scattering Ti/Zr alloy) and diffraction measurements were carried out over a period of about 4 hours under ambient conditions. All data were corrected for multiple scattering and absorption, and normalized versus a vanadium rod using standard procedures.^{31,32} Identical experimental conditions were applied for both the deuterium-labeled and natural abundance pyridine samples. All measurements were performed on the D4C diffractometer at ILL (Grenoble, France).

Data Analysis

The total measured intensity for the sample ($F(Q)$) is given by

$$F(Q) = \sum_{\alpha} \sum_{\beta} c_{\alpha} c_{\beta} b_{\alpha} b_{\beta} (S_{\alpha\beta}(Q) - 1) \quad (1)$$

where c_α is the atomic concentration of species α , b_α is the coherent neutron scattering length, and the sums are over all of the atomic species in solution. $S_{\alpha\beta}(Q)$ is the partial structure factor for atoms α and β . The partial radial distribution function $g_{\alpha\beta}(r)$ is derived from $S_{\alpha\beta}(Q)$ by Fourier transformation:

$$g_{\alpha\beta}(r) = \frac{1}{2\pi^2\rho r} \int (S_{\alpha\beta}(Q) - 1) \sin(Qr) Q dQ + 1 \quad (2)$$

The complexity of the present pyridine-water system is clear, since there are 5 types of nuclei: Hex (exchangeable hydrogen atoms on water); Ow (water oxygen); C (carbon in pyridine); N (nitrogen on pyridine) and Hnon (the non-exchangeable hydrogen atoms of pyridine). Therefore, the $S(Q)$ of this solution is the sum of 15 weighted partial structure factors. In each case the prefactors are numerically expressed as $2c_\alpha c_\beta b_\alpha b_\beta$, or $c_\alpha c_\beta b_\alpha b_\beta$ when $\alpha = \beta$, where c is the atomic concentration of nuclei of type α and b is the coherent neutron scattering length of that nucleus type.

First order difference scattering experiments were conducted on two solutions of pyridine of chemically-identical composition, one using the natural abundance molecule and the other using the sample of perdeuterated pyridine, or d_5 -pyridine, on the assumption that these two solutions are structurally identical. If a difference is taken between the $F(Q)$ s of these two solution, then all of the correlations exactly cancel except those that involve structure factors containing the substituted non-exchangeable hydrogen nuclei. As this experiment is performed in null water (the average scattering length of Hex = 0, rendering it essentially 'neutron invisible' from the point of view of coherent scattering), there are only four nuclei in this system, resulting in 10 correlations in $F(Q)$, which is reduced to 4 in the first order difference function $\Delta S_{H_{non}}^X(Q)$,

$$\begin{aligned} \Delta S_{H_{non}}^X(Q) &= S_{d_5\text{-pyridine}}(Q) - S_{\text{natural-pyridine}}(Q) \\ &= A [S_{H_{non}Ow}(Q) - 1] + B [S_{H_{non}C}(Q) - 1] + C [S_{H_{non}N}(Q) - 1] + D [S_{H_{non}H_{non}}(Q) - 1] \\ &= \sum \alpha_{H_{non}j} [S_{H_{non}j}(Q) - 1] \end{aligned} \quad (3)$$

where $j=Ow, C, N$ and H_{non} . The superscript X indicates that the function contains only correlations from the substituted nuclei H_{non} to all other nuclei in the system. The transform of this function gives the total radial distribution function $G_{H_{non}}^X(r)$,

$$G_{H_{non}}^X(r) = A g_{H_{non}Ow}(r) + B g_{H_{non}C}(r) + C g_{H_{non}N}(r) + D g_{H_{non}H_{non}}(r) - E \quad (4)$$

where $E = A+B+C+D$ and $g_{H_{non}X}(r)$ is the radial distribution function for atoms X around the substitution-labeled positions, and A - D are the neutron scattering prefactors for each atom type (which in each case is equal to $2c_{H_{non}}c_X\Delta b_H b_X$ ($c_{H_{non}}c_X\Delta b_H b_X$ when $X = H_{non}$), where c_X is the atomic concentration of each atom, b_X is the coherent neutron scattering length of each atom type X , and Δb_H is the contrast in the coherent scattering lengths of the substituted nuclei. In this case $\Delta b_H = b_D - b_H$ (deuterium and protium, respectively)).

Unfortunately, even this first order difference function $G_{H_{non}}^X(r)$ suffers from the same problems as the total measured $G(r)$, in that the intermolecular correlations are still essentially hidden under large and sharp intramolecular correlations.²⁹ The sharpness and separation of the intramolecular peaks makes them useful for exploring the molecular structure of the solute in water, particularly when some of the interatomic distances change as a function of conformation, a feature that has already been exploited to extend neutron

diffraction experiments to the study of sugar conformations in solution.²⁸⁻³⁰ For the study of the solvent structure around the labeled positions, however, their presence is very inconvenient. Clearly a method of removing these intense, sharp peaks to reveal the information on the intermolecular correlations is required.

Intramolecular Coordination Number Concentration Invariance (ICNCI)

In spite of the apparent intractability of the problem outlined above, it is possible to remove the intramolecular contributions to the labeled atom radial distribution functions using experiments performed at different concentrations. The method introduced here exploits the difference in how the measured neutron scattering multiplicative prefactors for the intramolecular and intermolecular contributions depend on concentration. The principal assumption made is that the conformational structure of the solute molecule does not change as a function of concentration, which is almost certainly true for the rigid pyridine molecule, but which may not hold in every case, particularly at very high concentrations of flexible molecules like peptides or sugars. In the present case the paired natural abundance/deuterium labeled difference experiments were carried out at two different pyridine concentrations, 1 and 5 molal, and the radial distribution function $G_{H_{non}}^X(r)$ was determined at each concentration. The neutron scattering prefactors A - D for first order differences at these concentrations are given in Table 1.

Consider initially the function $G_{H_{non}}^X(r)$ only in a small region around 1.1 Å, since in this range it will contain only one peak, due to the intramolecular H_{non} -C correlation for the carbon atom to which the labeled hydrogen is covalently bound. Therefore, for this region,

$$g_{H_{non}C}(r) = \frac{G_{H_{non}}^X(r)}{D} = \left(\frac{G_{H_{non}}^X(r)}{c_{H_{non}} c_C \Delta b_H b_C} \right) + 1 \quad (5)$$

Since this is an intramolecular correlation, the coordination number (CN) within this limited radius does not depend upon concentration (i.e. the coordination number for carbon around H_{non} is 1, regardless of the solution concentration), so that for the range 0 to 1.1 Å

$$4\pi\rho_1 c_{C_1} \int g_{H_{non}C}(r)_1 r^2 dr = 4\pi\rho_2 c_{C_2} \int g_{H_{non}C}(r)_2 r^2 dr \quad (6)$$

where ρ_1 and ρ_2 are the number densities of the solution at concentrations 1 and 2, c_{C_1} and c_{C_2} are the atomic fractions of carbon at concentrations 1 and 2, and $g_{H_{non}C}(r)_1$ and $g_{H_{non}C}(r)_2$ are the radial distribution functions for carbon around H_{non} at the concentrations 1 and 2, respectively.

From equation (6) it can be seen that for molecular correlations, where the coordination number is identical at both concentrations, the radial distribution function must have different-sized peaks to represent these identical coordination numbers. From (5) and (6) we have (for the range 0-1.1 Å)

$$\frac{4\pi\rho_1 c_{C_1} E_1}{c_{H_{non}1} c_{C_1} \Delta b_H b_C} + \frac{4\pi\rho_1 c_{C_1}}{c_{H_{non}1} c_{C_1} \Delta b_H b_C} \int G_{H_{non}}^X(r)_1 r^2 dr = \frac{4\pi\rho_2 c_{C_2} E_2}{c_{H_{non}2} c_{C_2} \Delta b_H b_C} + \frac{4\pi\rho_2 c_{C_2}}{c_{H_{non}2} c_{C_2} \Delta b_H b_C} \int G_{H_{non}}^X(r)_2 r^2 dr \quad (7)$$

which simplifies to

$$\frac{E_1\rho_1}{c_{H_{\text{non}1}}} + \frac{\rho_1}{c_{H_{\text{non}1}}} \int G_{H_{\text{non}}}^X(r)_1 r^2 dr = \frac{E_2\rho_2}{c_{H_{\text{non}2}}} + \frac{\rho_2}{c_{H_{\text{non}2}}} \int G_{H_{\text{non}}}^X(r)_2 r^2 dr \quad (8)$$

Since the constants $\frac{E_1\rho_1}{c_{H_{\text{non}1}}}$ and $\frac{E_2\rho_2}{c_{H_{\text{non}2}}}$ essentially constitute a structureless offset, it can be seen that the ratio of the peak sizes for a molecular correlation in the experimental measurement at two concentrations is dependent only upon the atomic fraction of the substituted nucleus and the number density of that solution. This example was chosen since in this case a single peak can be conceptually isolated in order to demonstrate the mathematical validity of the technique. However the method is valid whether the intramolecular peaks overlap with others or not, since in each case the measurement, which is a summation of several pairwise radial distribution functions, can be deconvoluted into the separate radial distributional components, which may in turn be deconvoluted into intra and intermolecular terms. It can thus be seen that the ICNCI method will effectively remove all intramolecular components of each radial distribution function, and that therefore the summation of all these pairwise radial distribution functions in the experimental ICNCI measurement will contain no intramolecular terms, regardless of whether the intramolecular peaks overlap with each other or not.

Practically this means that a simple weighted subtraction of two first order differences ($\Delta G_{H_{\text{non}}}^X(r)$), ignoring the offset, will eliminate all intramolecular correlations,

$$\Delta G_{H_{\text{non}}}^X(r) = \frac{\rho_2}{c_{H_{\text{non}2}}} G_{H_{\text{non}}}^X(r)_2 - \frac{\rho_1}{c_{H_{\text{non}1}}} G_{H_{\text{non}}}^X(r)_1 \quad (9)$$

Since $\frac{\rho_2}{c_{H_{\text{non}2}}} \Delta G_{H_{\text{non}}}^X(r)_2$ may be back-transformed into Q space to yield $\frac{1}{c_{H_{\text{non}2}}} \Delta S_{H_{\text{non}}}^X(Q)_2$, and similarly for solution 1, the difference function $\Delta G_{H_{\text{non}}}^X(r)$ can be expressed in Q space as:

$$\Delta \Delta S_{H_{\text{non}}}^X(Q) = \frac{1}{c_{H_{\text{non}2}}} \Delta S_{H_{\text{non}}}^X(Q)_2 - \frac{1}{c_{H_{\text{non}1}}} \Delta S_{H_{\text{non}}}^X(Q)_1 \quad (10)$$

where $\Delta \Delta S_{H_{\text{non}}}^X(Q)$ represents the experimental information that contains no correlations that do not vary with concentration; that is, only intermolecular correlations. More specifically, this residual function will contain no correlations whose coordination number is dependent on concentration. In Equation (9), both of the scaling factors contain the number density of the solution, while in Equation (10) neither do. This result is due to this factor canceling with the number density term that occurs in the Fourier transformation process (Equation 2). This procedure provides an opportunity to completely remove all intramolecular terms and allows insight into intermolecular structural terms.

Simulation

Molecular dynamics simulations were carried out on aqueous pyridine solutions under identical conditions as those in the experiments to provide assistance in the interpretation of the observed scattering. Neutral and periodic cubic systems were created at 1 and 5 molal concentrations containing a number of independent pyridine and water molecules interacting according to the CHARMM small molecules force field and the modified TIP3P model^{33,34}

for water. All simulations were performed with the CHARMM program,³⁵ with chemical bonds to hydrogen atoms kept fixed using SHAKE³⁶ and with a time step of 1 fs. The starting coordinates for the 1 molal system were created by placing 22 randomly orientated pyridine molecules in a 34 Å cube. These coordinates were superimposed on a box of 1296 water molecules, and solvent molecules that overlapped with the solutes were deleted, using an arbitrary separation cutoff, to produce the correct concentration (22 pyridine molecules and 1222 TIP3P water molecules, 1.000 molal). Finally, the box length was rescaled to 34.0148 Å, which yielded the correct physical number density (0.0993 atoms Å⁻³). The 5 molal solution, created in the same manner, contained 85 pyridine molecules and 944 TIP3P water molecules, giving a concentration of 5.002 molal. The box length was then rescaled to 33.9696 Å, which yielded the correct physical number density (0.0961 atoms Å⁻³). All van der Waals and electrostatic interactions were smoothly truncated on an group basis using switching functions from 13 to 15 Å. Initial velocities were assigned from a Boltzmann distribution (300 K) followed by 10 ps of equilibration dynamics with velocities being reassigned every 0.2 ps. The simulation was then run as an NVE ensemble for 1.0 ns with no further velocity modification. The first 0.33 ns of the trajectory was used as equilibration, and the remaining 0.67 ns for analysis.

Results and Discussion

The raw scattering data from the experiments, corrected for multiple scattering and absorption, and normalized versus a vanadium rod, produces the total scattering patterns ($F(Q)$ s) shown in Figure 2. The most obvious observation from these data is that the solutions that contain more H have a higher scattering level, and appear to be on a higher slope. The higher level is a consequence of the very high incoherent scattering cross section of H, while the slope is due to the inelastic recoil of the proton nucleus (the Placzek effect^{6,28}). For both the 1 and the 5 *m* solution the first order difference was obtained by the subtraction of the $F(Q)$ for natural-abundance pyridine solution from the $F(Q)$ for the perdeuterated pyridine solution. Both of these first order differences, $\Delta S_{H_{non}}^X(Q)$, are shown in Figure 2. There are two conspicuous differences between these two $\Delta S_{H_{non}}^X(Q)$. First, the oscillations in the higher concentration function are larger due to the higher contrast of this difference. In addition, the higher concentration function has a greater slope. This second point is due to the Placzek effect, since one of the substituted nuclei is hydrogen. While these artifacts may seem large, it should be noted that the Placzek effect yields an extremely broad feature (wavelengths ~ 80 Å⁻¹) and does not contribute to the higher frequency components that are of interest in this study (wavelengths 0-10 Å⁻¹). The experimental difference function $\Delta\Delta S_{H_{non}}^X(Q)$ is also shown in Figure 2. This plot also includes this difference function as calculated from MD simulations (discussed below) for comparison. These difference curves contain only the partial structure factors for the intermolecular correlations. Figure 3 shows the real space representation of $\Delta\Delta S_{H_{non}}^X(Q)$ obtained by transformation, the radial distribution function $\Delta G_{H_{non}}^X(r)$. The noise bar in these plots serves as a measure to estimate the ratio of the signal to the random counting statistics in these measurements. For the first order difference measurements, the noise is negligible relative to the signal, while in the second order function $\Delta\Delta S_{H_{non}}^X(Q)$, although the noise is greater, the signal is about an order of magnitude larger than the counting statistics of a state of the art machine.

The principal remaining problem with data such as is shown in Figure 3 is one of interpretation. The principal feature in $\Delta G_{H_{non}}^X(r)$, the peak at 7.5 Å, is primarily the result of the aggregation of pyridine at higher concentration, but this is not immediately obvious from

examination of Figure 3. However, parallel MD simulations on the same system can be invaluable for explaining the observed features. For this purpose, MD simulations were also performed for pyridine at 1 and 5 *m*. As a means of illustrating the difficulties involved in the analysis of the experimental data, the total atomic radial distribution function for a 1 molal pyridine solution as calculated from these simulations is shown in Figure 4. Unlike the case for the experimental scattering, the simulation function can be explicitly resolved into its various components. The obvious complexity of this function highlights the fact that the principal difficulty with neutron scattering measurements, particularly those of molecular solutes, is interpretation. The predicted experimental measurement contains several sharp peaks due to the intramolecular correlations of pyridine, which obscure the information relating to the intermolecular correlations. Even though the solute-water correlations constitute the majority of the scattering prefactors, they are dominated at short distances in $G(r)$ by the intramolecular terms. For these reasons the intermolecular structuring has been difficult to extract from the NDIS experiments.

The functions $G_{H_{\text{non}}}^X(r)$ from the simulations at both concentrations and the function $\Delta G_{H_{\text{non}}}^X(r)$ are shown in Figure 5. The function $\Delta G_{H_{\text{non}}}^X(r)$ has had all features that do not vary with concentrations removed by the ICNCI. The fact that no intramolecular correlations occur in this function is a demonstration of the validity of this method. The contrast (the sum of the scattering prefactors *A* through *D* (see Table 1) in the 1 *m* solution NDIS experiment is 12 mb, while it is found that the contrast of the function $\Delta G_{H_{\text{non}}}^X(r)$ is about 10 mb (estimated from the high *r* limit from Figure 3), easily within the limits of what can be experimentally measured, both with respect to the accuracy with which solutions of the correct atomic concentrations can be made, and the statistical counting error using current diffractometers. It is clear that the assumption made here that the molecular conformation does not change with concentration is very good, as would be expected for a rigid molecule like pyridine, in that the intramolecular terms are almost completely removed in the function $\Delta G_{H_{\text{non}}}^X(r)$. It is topologically inevitable that as the concentration of the solute is increased, the solute-solute coordination number will increase, while the solute-solvent coordination number will decrease. Therefore, the function $\Delta G_{H_{\text{non}}}^X(r)$ contains information about how the solute-solute and solute-solvent interactions vary with concentration.

The MD simulations are invaluable for examining the constitution and general features of such functions. Since the components of $\Delta G_{H_{\text{non}}}^X(r)$ are additive, we can examine each component from the simulations separately (see Figure 6). It was found that the $H_{\text{non}}\text{Ow}$ correlation is negatively weighted, indicating there is a preferential exclusion of water from the pyridine as the concentration is increased, while the positively weighted correlations for $H_{\text{non}}\text{C}$, $H_{\text{non}}\text{N}$ and $H_{\text{non}}H_{\text{non}}$ indicate that there is a preferential pyridine-pyridine interaction as concentration increases. From these partial contributions, it can be seen that the peak at about 4 Å is primarily due to the exclusion of water, while the peak at 7 Å is primarily due to the association of pyridine molecules.

This point can be further expanded upon by examination of the density maps for the structuring around pyridine (Figure 7). It can be seen that the main interaction of water with pyridine is via hydrogen bonds with the lone pairs of the nitrogen atom of the pyridine. In the 1 molal solution this nitrogen makes 1.97 hydrogen bonds with water, while in the 5 molal solution it is only somewhat reduced to 1.77 bonds, showing that interactions with other pyridine molecules do not significantly compete for this hydrogen bond (as would be expected due to the lack of hydrogen bond donors on pyridine). The coordination numbers for any atoms within 4.5 Å of a pyridine C atom are shown in Table 2. It was found that

there is a significant enhancement of pyridine concentration at the pyridine surface for both the 1 and 5 *m* solutions. The bulk ratio of Ow to C is 11.1 for the 1 *m* solution and 2.22 for the 5 *m* solution. Near the pyridine surface this ratio is 7.66 and 1.18 for the 1 and 5 *m* solutions respectively, indicating in both cases a propensity for the pyridine molecules to aggregate. Thus, for the 1 molal solution there is a 45% increase of the pyridine concentration at the surface compared to a 20% increase for the 5 molal solution. However, it should be noted that the surface concentration of other pyridine molecules at the pyridine surface is about 4 times higher for the 5 molal solution than for the 1 molal solution. From these results it becomes clear that the pyridine primarily displaces water from the faces and edges of the pyridine, while pyridine is essentially ineffective at displacing the waters hydrogen bonded to the N of pyridine. From examination of the density map for H_{non} and C around pyridine (Figure 8), it becomes evident that the main from of pyridine-pyridine interaction is via an edge:face style interaction, in that the cloud for H_{non} density is found nearer to the pyridine than C cloud, whereas for the face:face style interaction the H_{non} cloud would be found within the C cloud. This propensity of the pyridine to aggregate results in the significant feature in $\Delta G_{H_{non}}^X(r)$, where a long range augmentation of this function over the lengthscale of about 2 nm, (primarily due to the enhanced aggregation of pyridine at higher concentration), is observed. As expected, this results in a significant signal at lower *Q* in $\Delta\Delta S_{H_{non}}^X(Q)$ (Figure 2).

Conclusions

The technique of ICNCI has two principal benefits. The first of these is that it is a differential technique, and hence has a tendency to eliminate systematic errors that can be very difficult to isolate and eliminate in such techniques as neutron scattering. This is intrinsic to the nature of these experiments, where typically, due to the cost of running neutron scattering facilities, the user only gets access to the machine for enough time to perform the experiment once on a unique machine. The second benefit is that via the differential subtraction, all the intra-molecular correlations are eliminated. This is important since these correlations tend to be very pronounced, even at low concentrations, and tend to conceal the intermolecular correlations of interest.²⁹ It has also been observed in previous experiments that there is a significant discrepancy in the broadness of these peaks when the bond contains a hydrogen atom. It is unclear if this is a problem with the neutron scattering properties of hydrogen, or perhaps is due to the more pronounced quantum nature of hydrogen, that is inadequately modeled by classical Newtonian dynamics. However, the present approach offers significant advantages over second order difference NDIS experiments on H_{non}, which contain a very large and uninteresting signal from intramolecular pyridine correlations which is difficult to separate from the interesting intermolecular correlations.

Molecular dynamics simulations of pyridine in water show that the number of hydrogen bonds that the nitrogen atom of pyridine makes to water is relatively insensitive to the concentration of pyridine, only decreasing by about 20% as the concentration of pyridine is increased from 1 to 5 *m*. Meanwhile the nature and magnitude of the pyridine-pyridine interactions scales relatively linearly with concentration of pyridine. Pyridine principally interacts with itself via a T-type edge-face interaction, and not via face-face interactions as has been previously observed for Gdm⁺ ions.

The method of ICNCI shows significant promise as a mechanism for testing and refining the ability of MD simulations to successfully replicate the correct balance of solute-solute and solute-water interactions. Of course, MD simulations have limitations, such as the accuracy of water and solute force fields,³⁷ the treatment of long range interactions,³⁸ etc. Nevertheless, MD modeling has proven to be very effective in helping to guide assignments

of peaks in experimental radial distribution functions. However, the use of concentration-dependent experiments presented here represents a significant advantage over second order NDIS experiments alone in that it can be used to remove the strong intramolecular correlations which tend to dominate the experimental measurement from the intermolecular contributions in the experimental data without ever having to calculate or model them. While the resulting composite function is difficult to interpret directly, it nonetheless presents a stringent criterion that has significant value in the appraisal of simulation models, particularly where aggregation of molecular solutes may occur in aqueous media. Poor agreement between the modeling and experimental results would be a strong indication of the need for further improvements in the simulation force fields.

Acknowledgments

This project was supported by grant GM63018 from the National Institutes of Health and a grant of beam time from the Institute Laue-Langevin.

References

- (1). Kossiakoff AA, Sintchak MD, Shpungin J, Preta LG. *Proteins*. 1992; 12:223–236. [PubMed: 1557350]
- (2). Liu Q, Brady JW. *J. Am. Chem. Soc.* 1996; 118:12276–12286.
- (3). Dixit S, Crain J, Poon WCK, Finney JL, Soper AK. *Nature*. 2002; 416:829–832. [PubMed: 11976678]
- (4). Mason PE, Neilson GW, Enderby JE, Saboungi M-L, Dempsey CE, MacKerell AD, Brady JW. *J. Am. Chem. Soc.* 2004; 126:11462–11470. [PubMed: 15366892]
- (5). Mason PE, Dempsey CE, Neilson GW, Brady JW. *J. Phys. Chem. B*. 2005; 109:24185–24196. [PubMed: 16375411]
- (6). Mason PE, Neilson GW, Dempsey CE, Brady JW. *J. Am. Chem. Soc.* 2006; 128:15136–15144. [PubMed: 17117865]
- (7). Fecko CJ, Eaves JD, Loparo JJ, Tokmakoff A, Geissler PL. *Science*. 2003; 301:1698–1702. [PubMed: 14500975]
- (8). Nishida Y, Ohrui H, Meguro H. *Tetrahedron Lett.* 1984; 25:1575–1578.
- (9). Nishida Y, Hori H, Ohrui H, Meguro H. *J. Carb. Chem.* 1988; 7:239–250.
- (10). Enderby JE. *Chem. Soc. Rev.* 1995; 24:159–168.
- (11). Neilson GW, Enderby JE. *J. Phys. Chem.* 1996; 100:1317–1322.
- (12). Sullivan DM, Neilson GW, Fischer HE. *J. Chem. Phys.* 2001; 115:339–343.
- (13). Enderby JE, Neilson GW. *Rep. Prog. Phys.* 1981; 44:593–653.
- (14). Neilson GW, Enderby JE. *Proceedings of the Royal Society of London*. 1983; A 390:353–371.
- (15). Soper AK, Weckström K. *Biophysical Chemistry*. 2006; 124:180–191. [PubMed: 16698172]
- (16). Mancinelli R, Botti A, Bruni F, Ricci MA, Soper AK. *J. Phys. Chem. B*. 2007; 111:13570–13577. [PubMed: 17988114]
- (17). Sandstrom M, Neilson GW, Johansson G, Yamaguchi T. *Journal of Physics C: Solid State Physics*. 1985; 18:1115–1121.
- (18). Polydorou NG, Wicks JD, Turner JZ. *J. Chem. Phys.* 1997; 107:197–204.
- (19). McLain SE, Soper AK, Daidone I, Smith JC, Watts A. *Angewandte Chemie-International Edition*. 2008; 47:9059–9062.
- (20). Bowron DT, Finney JL, Soper AK. *J. Phys. Chem. B*. 1998; 102:3551–3563.
- (21). Bowron DT, Soper AK, Finney JL. *J. Chem. Phys.* 2001; 114:6203–6219.
- (22). Sidhu KS, Goodfellow JM, Turner JZ. *J. Chem. Phys.* 1999; 110:7943–7950.
- (23). Kameda Y, Ebata H, Usuki T, Uemura O, Misawa M. *Bulletin of the Chemical Society of Japan*. 1994; 67:3159–3164.

- (24). Sasaki M, Kameda Y, Yaegashi M, Usuki T. Bulletin of the Chemical Society of Japan. 2003; 76:2293–2299.
- (25). Kameda Y, Sugawara K, Usuki T, Uemura O. Bulletin of the Chemical Society of Japan. 2003; 76:935–943.
- (26). Mao GM, Saboungi M-L, Price DL, Armand MB, Howells WS. Physical Review Letters. 2000; 84:5536–5539. [PubMed: 10990988]
- (27). Mao G, Saboungi M-L, Price DL, Badyal YS, Fischer HE. Europhysics Letters. 2001; 54:347–353.
- (28). Mason PE, Neilson GW, Enderby JE, Saboungi M-L, Cuello G, Brady JW. J. Chem. Phys. 2006; 125:224505. (nine pages). [PubMed: 17176147]
- (29). Mason PE, Neilson GW, Enderby JE, Saboungi M-L, Brady JW. J. Am. Chem. Soc. 2005; 127:10991–10998. [PubMed: 16076206]
- (30). Mason PE, Neilson GW, Enderby JE, Saboungi M-L, Brady JW. J. Phys. Chem. B. 2006; 110:2981–2983. [PubMed: 16494297]
- (31). Barnes AC, Enderby JE, Breen J, Leyte JC. Chem. Phys. Lett. 1987; 142:405–408.
- (32). Barnes AC, Hamilton MA, Beck U, Fischer HE. J. Phys. Condens. Mat. 2000; 12:7311–7322.
- (33). Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. J. Chem. Phys. 1983; 79:926–935.
- (34). Neria E, Fischer S, Karplus M. J. Chem. Phys. 1996; 105:1902–1919.
- (35). Brooks BR, Bruccoleri RE, Olafson BD, Swaminathan S, Karplus M. J. Comput. Chem. 1983; 4:187–217.
- (36). van Gunsteren WF, Berendsen HJC. Mol. Phys. 1977; 34:1311–1327.
- (37). Mason PE, Brady JW. J. Phys. Chem. B. 2007; 111:5669–5679. [PubMed: 17469865]
- (38). Lague P, Pastor RW. J. Phys. Chem. B. 2004; 108:363–368.

Figure 1. Pyridine possesses five non-exchangeable aromatic protons which were labeled with deuterium in the present experiments. The molecule is planar and isoelectronic with benzene, but contains a polar nitrogen atom.

Figure 2.

Top left, the $F(Q)$ s for 1 m pyridine in null water (black, natural pyridine; grey, $d5$ pyridine); top right, the $F(Q)$ s for 5 m pyridine in null water (black, natural pyridine; grey, $d5$ pyridine). Lower left, the first order difference functions ($\Delta S^X_{H_{\text{non}}}(Q)$) for the previous two plots (grey, 1 m pyridine; black, 5 m pyridine). Lower right, the function $\Delta\Delta S^X_{H_{\text{non}}}(Q)$ as calculated from MD (red), and directly determined by experiment (grey). The black line shows the experimental data after the removal of the Placzek effect.

Figure 3.

On the left is shown the function $\Delta\Delta S_{H_{non}}^X(Q)$; on the right the function $\Delta G_{H_{sub}}^X(r)$. In both cases, the blue curve is the function predicted from MD, and the raw experimental data is shown in black, while the experimental function after the removal of the Placzek effect, the application of a smoothing function, and the cropping of the data at $8 \text{\AA}^{-1}$ in reciprocal space, is shown in red. Note in particular that the peak at $7.5 \text{\AA}$, primarily related to the longer range aggregation of pyridine, is reproduced by the molecular dynamics.

Figure 4.

On the left is shown the predicted total radial distribution for 1 *m* natural-abundance pyridine in null water (grey) and perdeuterated pyridine in null water (black). Even the total distribution functions at modest concentration are dominated by the intramolecular correlations of the polyatomic solute. Shown in red is the first order difference method applied to these two solutions. This factor contains fewer terms and is significantly easier to interpret than the total scattering pattern. However, it too is clearly still dominated by the intramolecular terms. On the right is shown the expansion of the MD prediction for the first order difference, along with the contribution the hydrating water makes to this function (blue).²⁸

Figure 5.

The MD predictions for the function $G_{H_{non}}^X(r)$ at pyridine concentrations of 1 and 5 m (lower grey and black respectively). Upper black is shown the function $\Delta G_{H_{non}}^X(r)$ (translated for clarity by 0.14 barns atom⁻¹ str⁻¹).

Figure 6.

The function $\Delta G_{H_{\text{non}}}^X(r)$ as predicted from MD (upper black), with its positively and negatively weighted components. The negatively weighted component result from the $H_{\text{non}}\text{Ow}$ correlation (lower red), while the positively weighted components result from $H_{\text{non}}\text{C}$, $H_{\text{non}}\text{N}$ and $H_{\text{non}}H_{\text{non}}$ correlations (upper grey and blue and green lines, respectively). The long distance fluctuations, highlighted by a nanometer lengthscale rise above the high r limit (black horizontal lines), are primarily due to the enhanced aggregation of pyridine at higher concentration. This long scale structure is also primarily responsible for the small angle signal observed in the experimental results.

Figure 7.

Both the density maps shown are for the 5 molal solution. Left (yellow) is shown the density map of C around pyridine ($0.04 \text{ atoms } \text{\AA}^{-3}$, 3.7 times the average number density of C). Right (red) is shown the density map of Ow around pyridine ($0.07 \text{ atoms } \text{\AA}^{-3}$, 2.9 times the average number density of Ow). The band over the nitrogen clearly shows that the water preferentially interacts with pyridine by a hydrogen bond, while the aggregation of pyridine molecules takes place through stacking of the rings.

Figure 8. Two views of the density maps of Hnon (yellow) and C (orange) around pyridine in the 5 *m* solution (both contour are at a number density of $0.035 \text{ atoms } \text{\AA}^{-3}$, 3.2 times the average number density of each nucleus, respectively). As the Hnon cloud lies closer to the pyridine than the C cloud, this indicates that pyridine molecules preferentially interact with each other in a T-type fashion.

Table 1

The neutron scattering prefactors for the first order differences of pyridine solutions substituted with deuterium at the H_{non} positions for the 1 and 5 molal solutions in null water. Notice the difference in behavior of these prefactors with concentration. The H_{non}Ow prefactor increases by about a factor of 3 while the H_{non}C prefactor increases by a factor of 16.

Prefactor	Correlation	Value in mb	
		1 molal	5 molal
A	H _{non} Ow	10.6	34.2
B	H _{non} C	1.09	17.6
C	H _{non} N	0.309	4.96
D	H _{non} H _{non}	0.242	3.89
Total (G)		12.29	60.65

Table 2

The coordination of each atom type within 4.5 Å of a carbon on pyridine.

Atom type	1 molal pyridine	5 molal pyridine
N	0.38	1.28
C	2.33	7.21
H _{non}	2.23	7.06
O _w	17.9	13.6
H _{ex}	34.0	25.8