

Using HOESY NMR spectroscopy to characterise pre-nucleation aggregates

Raitis Bobrovs^{*a}, Andrievs Auseklis Auzins^a, Laura Drunka^a, Rimants Metlans^a, Ruslans Muhamadejevs^a and Kristaps Jaudzems^{a,b}

^a Latvian Institute of Organic Synthesis, Aizkraukles 21, LV-1006, Riga, Latvia

^b Faculty of Chemistry, University of Latvia, Jelgavas 1, LV-1004 Riga, Latvia

Pre-nucleation aggregates are important species on the crystallisation pathway. Here we combine heteronuclear Overhauser effect spectroscopy (HOESY) and molecular dynamics calculations to study the self-association of a model system – benzoic acid. Our findings indicate that heteronuclear Overhauser effects arise from diffusion-limited pre-nucleation aggregates and that self-association is solvent dependant.

Introduction

Solute molecule self-association into diffusion-limited pre-nucleation aggregates is the first step in the formation of crystal nuclei during crystallisation. In order to understand the fundamental processes behind crystallisation and crystal nucleation, it is important not only to study the properties of the bulk crystalline phase obtained at the end of the crystallisation, but also to investigate the preceding steps¹ which play a decisive role in determining the crystal properties and structure. There is, however, a limited number of techniques that allow to probe the pre-nucleation aggregates and solute molecule self-associates. Some insight into the self-association process has been obtained using nuclear magnetic resonance (NMR)^{2–4}, Fourier-transform infra-red (FTIR) and ultra-violet and visible light (UV-Vis) spectroscopies^{4,5}, small- and wide-angle X-ray scattering (SAXS/WAXS)⁶, cryogenic transmission electron microscopy (cryo-TEM)⁷, atomic electron tomography (AET)⁸, and molecular dynamics (MD) simulations^{9,10}. Nevertheless, information at molecular level on the complexes formed by solute molecules during self-association, especially for the larger aggregates, is still poor.

NMR spectroscopy is one of the most powerful and versatile techniques, used to investigate structural and dynamic properties of molecules in solution. Besides providing information on through-bond connectivity of atoms, various NMR spectral parameters (e.g. chemical shifts, nuclear relaxation rates, nuclear Overhauser effects) report on the surrounding chemical environment of each nucleus in the studied molecule. While most of these parameters offer indirect information about molecular interactions, the nuclear Overhauser effect (NOE) can provide direct evidence of short *inter-* nuclear distances and is particularly valuable for probing through-space *inter-* molecular contacts. NOE arises from magnetisation transfer through space via cross relaxation between nuclei that are usually within ~5 Å. However, *inter-* molecular NOEs have a different distance dependence dominated by longer-range components and they can extend beyond 10 Å¹¹. Such *inter-* atomic contacts are formed during the solute molecule self-association, therefore, NOE could be an important tool to characterise the geometry of such complexes. The heteronuclear Overhauser effect (HOE) is a cross relaxation process similar to the NOE, but leading to through-space magnetisation transfer between nuclei of two different elements. It has been widely applied to probe *inter-* molecular contacts between small molecules in different systems such as ionic liquids¹², surfactant micelles¹³ metal-organic frameworks¹⁴ and nucleoside hydrates¹⁵. Although the intermolecular HOEs are weaker than NOEs, they allow measurement of unambiguous contacts between self-interacting molecules.

Here we explore the possibility to characterise small molecule self-association in pre-nucleation aggregates by exploiting the heteronuclear Overhauser effect. We apply the 2D HOESY NMR technique to study the self-association of benzoic acid as a model system, which has been thoroughly characterised in previous studies^{16–18}. The information obtained from HOEs is complemented with molecular dynamics (MD) calculations to predict the self-associate structure.

Experiment design

In order to detect *inter-* molecular NOEs between solute molecules, several conditions need to be satisfied. First, cross-relaxation between solute molecules has to occur faster than the self-

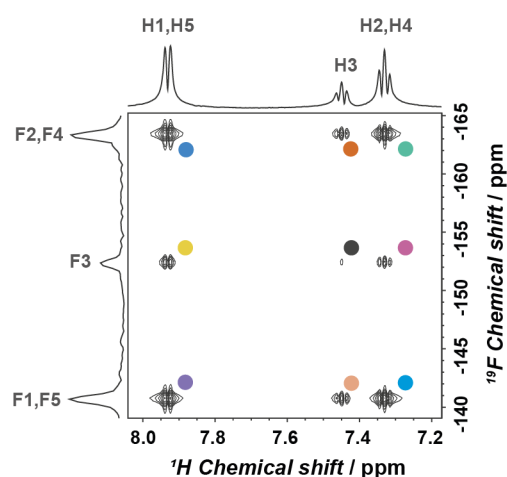
associate dissociation. Since small solute molecules have short correlation times, the cross-relaxation rates are usually slow, and as a consequence NOEs may be too small to be observed. However, cross-relaxation between solute molecules in pre-nucleation aggregates is facilitated by their restricted mobility and increased solution viscosity. Second, the nuclei involved in *inter-molecular* contacts need to be separated spectroscopically and *inter-molecular* contacts need to be distinguishable from *intra-molecular* contacts. This rules out the possibility to study single component systems in absence of isotope labelling, because in such systems *intra-* and *inter-molecular* interactions are indistinguishable. To overcome this obstacle, we chose a system of two structurally very similar compounds – benzoic acid (BA) and pentafluorobenzoic acid (pFBA) that are able to hetero-associate. Both compounds share the same functional groups – an aromatic ring and a carboxylic group, and their aromatic substituents (protons and fluorine atoms) are of similar size. Both compounds are also known to crystallise as stacked centro-symmetric carboxylic dimers, just like their co-crystal¹⁹ (see Fig. SI1). Additionally, the aromatic protons and fluorines allow separation of the NMR signals of the two molecules as well as discrimination between *ortho*, *meta* and *para* sites. Furthermore, ¹⁹F to ¹H cross-relaxation rate is almost as fast as from ¹H to ¹H. Thus, this model system perfectly meets the requirements for measuring *inter-molecular* contacts and is also expected to give a rather accurate approximation of the self-association process that occurs in single component systems.

Concentration dependant self-association

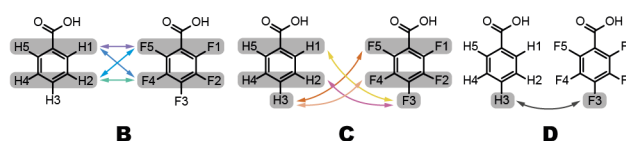
Initial HOESY spectra were measured for high concentration (~2 M) BA:pFBA 1:1 solution in methanol-D₄. The obtained spectra showed several HOEs, confirming that aromatic protons of BA are in close proximity to fluorine atoms of pFBA (Fig. 1.A, see SI for experimental details). The experimental parameters were optimized using this sample and used for all other samples.

Fig. 1.B-D shows all the possible *inter-molecular* HOEs between non-labile atoms of BA and pFBA. In case of totally random self-associates the relative ratio of these HOEs is approximately 4(B) : 2(C) : 1(D), which is equal to the ratio of equivalent spin pairs giving a particular HOE. Our experimental data show deviations from this ratio (Fig. 1.E), which indicates presence of preferred interactions between solute molecules. In addition, the HOE signal ratio markedly depends on the solution concentration. No HOE cross peaks were detected for low concentration solutions below 0.1 M, most likely, because of minor fraction of pre-nucleation aggregates.

At highest BA:pFBA concentrations the HOE signal ratio shows relatively small deviation from random self-association (red dashed line in Fig. 1.E). This is in-line with our previous studies²⁰ showing a great variety of self-associates present in *para*-aminobenzoic acid solutions in organic solvents, with comparable fractions of various π - π stacked and hydrogen bonded self-associates. Such scenario is likely in this system and would appear as a random distribution. At lower concentrations HOEs arising from *para*- site interactions, and to lesser extent also from *meta*-sites, become proportionally less intense. This suggests that the equilibrium of self-associates is shifted to complexes where both solute molecules interact mostly through carboxylic groups and/or are stacked in an arrangement where *ortho*- and *meta*- sites are in close proximity, while *para*-sites are not.



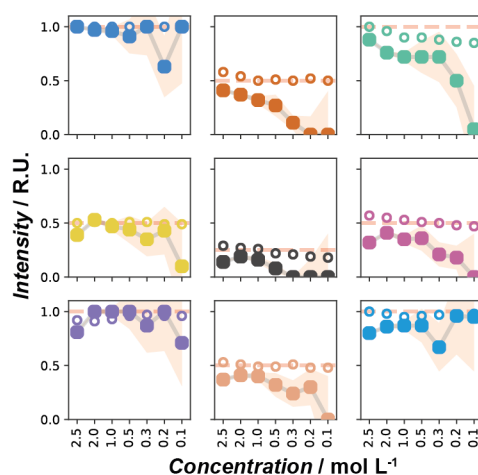
A



B

C

D



E

Fig. 1. **(A)** ^{19}F - ^1H HOESY spectrum of 2 M BA:pFBA 1:1 solution in methanol- D_4 . Proton and fluorine atom labelling is given below. **(B, C and D)** Possible ^{19}F - ^1H HOEs in BA:pFBA solution. **(E)** Relative intensity of experimentally observed (filled circles) and theoretically calculated (empty circles, see SI for computational details) HOEs as a function of 1:1 BA:pFBA methanol- D_4 solution concentration. Red dashed line represents the relative HOE intensity for totally random self-associates. Light orange area indicates the uncertainty of experimental data. Colour coding matches the subfigures A-D.

Our MD simulation data support these observations. BA and pFBA molecules in 2 M methanol solution are densely packed (see MD simulation snapshot in Fig. SI3.B), thus, satisfying the ^{19}F - ^1H *inter-* nuclear distance requirements necessary to observe the HOEs. The solute molecule self-associate size distribution (Fig. 2.A) indicates that BA and pFBA monomers are the energetically favoured species in most solutions, except those near saturation. This explains rather limited NMR experiment sensitivity at low concentrations, as the concentration of self-associates satisfying HOE conditions are several times lower than the actual solution concentration. The preferred solute molecule interactions in all systems studied (the same as experimentally) are π - π stacking (see Table SI1). There is, however, a large diversity of these aggregates – *T-shaped*, *face-to-face*, *offset* and various intermediate π - π stacked species (see Table SI1). The considerable aggregate conformational heterogeneity is also evident from the BA and pFBA molecule relative orientation and intermolecular distance free energy surface (FES, Fig. 2.B) and BA pFBA dimer spatial distribution function (SDF, Fig. 2.C). The FES shows two distinct basins – one corresponding to

parallel *face-to-face* and *offset* π - π stacked solute molecules, and second, considerably wider and more populated one, corresponding to other π - π stacking and hydrogen bonding interactions. MD simulations suggest that hydrogen bonding is less common than π - π stacking, besides it is mostly asymmetric, as previously determined using neutron-scattering¹⁶. While there are considerable differences in the fraction of solute molecules involved in self-association upon changes in concentration, the ratio of π - π stacked and hydrogen bonded self-associates changes only slightly with concentration (see Table SI1).

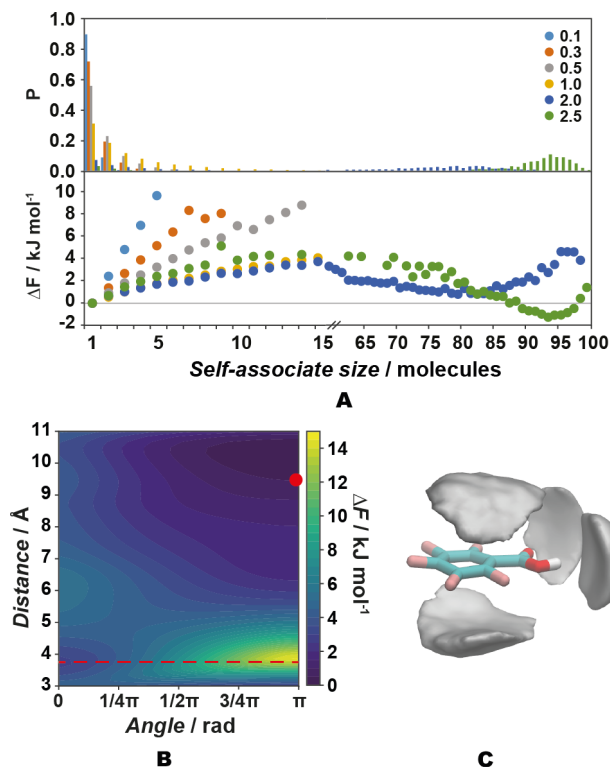


Fig. 2. (A) Probability of the solute molecule to be part of the cluster of specific size (top), and energetic penalty associated with the formation of this self-associate (bottom). (B) Free energy surface of self-associated solute molecule dimers in 0.5 M methanol solution as a function of intermolecular distance and relative orientation. Red dashed line represents intermolecular distance in face-to-face π - π stacked dimers. Red circle indicates location of hydrogen bonded dimer found in BA pFBA co-crystal. (C) The spatial distribution function of BA molecules around pFBA molecule in 0.5 M methanol solution.

The calculated time averaged HOE signal ratios match experimentally observed ratios for high concentration BA pFBA solutions (down to ~ 0.5 M, see Fig. 1.E). The pattern of HOE signal ratio changes with solution concentration generally follows the trend observed in the experiments, however, the magnitude of the changes in calculated HOE signal ratios is noticeably less pronounced. The main discrepancy is slight preference for self-associates with intensive HOEs between *ortho*- and *meta*- substituents in our experimental data.

An analysis of the aggregates average lifetime shows that all self-associated species possess a dynamic character. The hydrogen bonded interactions between two BA and pFBA are nearly twice as long as π - π stacking (Table SI1), however, it is several orders of magnitude lower than the molecular tumbling correlation times. Therefore, for low concentration solutions, the observed HOEs are more likely to represent time-averaged intermolecular contacts within smaller aggregates (dimers, trimers), while for high concentration solutions large self-associated aggregates with longer lifetimes are observable directly and are the main species contributing to the HOEs.

Formation of large, partly ordered diffusion limited aggregates is also supported by the observation that the maximum HOE intensity is measured after some equilibration time, which we attribute to the time necessary for solute molecule self-association into larger aggregates. This indicates that

the HOEs measured arise from large diffusion limited pre-nucleation aggregates alongside two-step nucleation pathway¹, rather than small short-lived aggregates like dimers, trimers, etc.

Solvent role in self-association

Solute molecule self-association is solvent dependant, and solvent used for the crystallisation is considered to be one of the key factors affecting crystallisation outcome for polymorphic systems^{17,20}. Although there are no reported BA polymorphs, the system studied can be used as a proof of concept, to characterise molecular synthons found in various solvents. To evaluate the solvent role on the self-association, HOE signal intensities were measured for 0.5 M BA pFBA solutions in cyclohexene- D_{12} , benzene- D_6 , chloroform- D , methanol- D_4 , pyridine- D_5 and acetic acid- D_4 (see Fig. 3). The relative intensity of nearly all HOE signals increases with solvent polarity, indicating that the average distances between all measured spin pairs decrease. This is likely a result of π - π stacking interactions becoming more favoured than hydrogen-bonded interactions in polar, hydrogen bond donor solvents (*inter*- nuclear distances are longer in hydrogen bonded dimers – see Fig. SI4 and SI5). Thus, our data support the assumption that solute molecule carboxylic hydrogen bonding will dominate in non-polar solvents, as there are no other hydrogen bond donors in the solution that might compete for hydrogen bonding.

HOE signal ratios calculated for the respective systems from MD simulations are in-line with the experimentally obtained data, however, again, the magnitude of the ratio differences between the solvents is slightly underestimated (see Fig. 3). Hydrogen bonding is favoured in low polarity solvents (see Table SI2.), where a large fraction of solute molecules is involved in hydrogen bonded oligomer formation, and only small fraction of solute molecules exist as monomers. Moreover, the fraction of symmetric hydrogen bonding is more than 2 times higher in non-polar solvents due to lack of competitive hydrogen bonding from solvent. The extensive hydrogen bonding results in considerably higher aggregate lifetimes in non-polar solvents, however, nowhere near the HOESY mixing times. There is no clear correlation between the fraction of π - π stacked molecules and solvent polarity.

This example demonstrates that it is possible to use HOE measurements to estimate the dominant self-associates in various solvents and, thus, rationalise the correlation between the solute molecule self-associates in the solution and molecular synthons observed in solid crystalline material. Nevertheless, as shown by the supporting MD calculations, there is a considerable variety of self-associates present in the solution besides the dominant aggregate species.

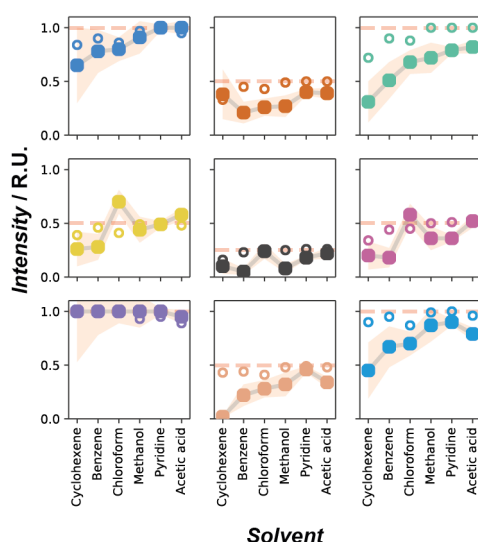


Fig. 3. The relative intensity of experimentally observed (filled circles) and theoretically calculated (empty circles) HOEs of 0.5 M 1:1 BA pFBA solution as a function of solvent used. Solvents are arranged according to their polarity and hydrogen bond donating/accepting properties²¹ – aprotic non-polar solvents are on the left, polar hydrogen bond donor/acceptor solvents on the right. Red dashed line represents the relative NOE intensity for totally random self-associates. Light orange area indicates the uncertainty of experimental data.

Conclusions

In this work, the relative ^{19}F - ^1H HOE intensity measurements were combined with MD calculations to analyse the self-association of benzoic acid. Here we show that the HOEs measured are mainly due large diffusion limited pre-nucleation aggregates, and the HOE signal ratios measured give the structural description of these self-associates. The obtained results demonstrate that there is a great variety of interactions present in the solution and self-association is both, concentration and solvent dependant. We believe that the method can also be applied for studying the potential self-associate rearrangement in a cooling crystallisation experiment. Methods that rely on HOE measurements are scarcely exploited to study small molecule self-association and crystallisation processes, however, we believe that by wisely selecting systems that allow to decouple *intra*- and *inter*- molecular interactions, such method could yield significant insights on self-association and crystallisation process.

Acknowledgement

This research is funded by the Latvian Council of Science, project "From molecules to crystals: molecular self-assembly in crystal nucleation from solution", project No. Izp-2018/2-0234.

Notes and references

- 1 D. Erdemir, A. Y. Lee and A. S. Myerson, *Acc. Chem. Res.*, 2009, **42**, 621–629.
- 2 C. E. Hughes and K. D. M. Harris, *Chem. Commun. (Camb)*., 2010, **46**, 4982–4.
- 3 A. Spitaleri, C. A. Hunter, J. F. McCabe, M. J. Packer and S. L. Cockroft, *CrystEngComm*, 2004, **6**, 489–493.
- 4 R. Bobrovs, L. Seton and N. Dempster, *CrystEngComm*, 2015, **17**, 5237–5251.
- 5 S. A. Kulkarni, E. S. McGarrity, H. Meekes and J. H. ter Horst, *Chem. Commun. (Camb)*., 2012, **48**, 4983–5.
- 6 S. Chattopadhyay, D. Erdemir, J. M. B. Evans, J. Ilavsky, H. Amenitsch, C. U. Segre and A. S. Myerson, *Cryst. Growth Des.*, 2005, **5**, 523–527.
- 7 E. M. Pouget, P. H. H. Bomans, J. A. C. M. Goos, P. M. Frederik, G. de With and N. A. J. M. Sommerdijk, *Science (80-.)*., 2009, **323**, 1455–1458.
- 8 J. Zhou, Y. Yang, Y. Yang, D. S. Kim, A. Yuan, X. Tian, C. Ophus, F. Sun, A. K. Schmid, M. Nathanson, H. Heinz, Q. An, H. Zeng, P. Ercius and J. Miao, *Nature*, 2019, **570**, 500–503.
- 9 G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen and A. Michaelides, *Chem. Rev.*, 2016, **116**, 7078–7116.
- 10 M. Salvalaglio, M. Mazzotti and M. Parrinello, *Faraday Discuss.*, 2015, **179**, 291–307.
- 11 S. Gabl, O. Steinhauser and H. Weingärtner, *Angew. Chemie - Int. Ed.*, 2013, **52**, 9242–9246.
- 12 P. A. Martin, E. Salager, M. Forsyth, L. A. O'Dell and M. Deschamps, *Phys. Chem. Chem. Phys.*, 2018, **20**, 13357–13364.
- 13 D. Canet, N. Mahieu and P. Tekely, *J. Am. Chem. Soc.*, 1992, **114**, 6190–6194.
- 14 A. Pastor and E. Martínez-Viviente, *Coord. Chem. Rev.*, 2008, **252**, 2314–2345.
- 15 B. Ancian, *NMR Studies for Mapping Structure and Dynamics of Nucleosides in Water*, Elsevier Ltd., 1st edn., 2010, vol. 69.
- 16 R. C. Burton, E. S. Ferrari, R. J. Davey, J. L. Finney and D. T. Bowron, *J. Phys. Chem. B*, 2010, **114**, 8807–8816.
- 17 A. J. Cruz-Cabeza, R. J. Davey, S. S. Sachithananthan, R. Smith, S. K. Tang, T. Vetter and Y. Xiao, *Chem. Commun.*, 2017, **53**, 7905–7908.
- 18 W. Tang, M. Zhang, H. Mo, J. Gong, J. Wang and T. Li, *Cryst. Growth Des.*, 2017, **17**, 5049–5053.
- 19 M. Gdaniec, W. Jankowski, M. J. Milewska and T. Połośki, *Angew. Chemie - Int. Ed.*, 2003, **42**, 3903–3906.
- 20 R. Bobrovs, L. Drunka, A. A. Auzins, K. Jaudzems and M. Salvalaglio, Preprint. 10.26434/chemrxiv.12943106.v1
- 21 C. Reichardt and T. Welton, *Solvents and Solvent Effects in Organic Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010, vol. 82.