University of East Anglia

School of Pharmacy

Determining the dissociation and binding dynamics of molecules using ^{1}H NMR chemical shift imaging techniques

Thesis submitted in accordance with the requirements of the University of East Anglia for the degree of Doctor of Philosophy

By

Haider Hussain

January 2026

Abstract

This PhD thesis introduces and applies new methodologies based on Nuclear Magnetic Resonance (NMR) chemical shift imaging (CSI) to measure acid dissociation constants (pK_a) and metal–ligand binding dynamics in situ. The research addresses longstanding challenges in characterizing systems that are difficult to analyse by conventional means (such as polymers with broad NMR signals or systems requiring numerous titrations) by leveraging controlled chemical gradients within a single NMR tube. In one approach, a derivative-based fitting method was developed to determine pK_a from ^{1}H NMR titration data. By analysing the derivative of solution pH with respect to the observed chemical shift (δ_{obs}) of an analyte, this method circumvents the need to fit limiting chemical shift values and avoids convergence issues associated with traditional nonlinear regression. It enables rapid, robust pK_a determination via a simple polynomial fit, yielding values in excellent agreement with literature standards while reducing the data requirements and time cost of titrations.

The subsequent chapter expands NMR pK_a measurements to *NMR-invisible* species (e.g. macromolecules or ions whose proton signals are unobservable or too broad) by coupling CSI with indicators that measure the quantity of protons transferred to them from the analyte. Using a controlled pH gradient in a single “one-shot” CSI experiment, the quantity of protons transferred from an “invisible” analyte to an NMR-visible indicator is measured across space, enabling calculation of the analyte’s pK_a and concentration without direct observation of its own signals. This technique was validated on model systems including a high-molecular-weight polymer (polyacrylic acid) and a protein (wheat germ agglutinin), successfully yielding pK_a values consistent with their known behaviour. The method proved effective even when analyte NMR peaks were severely broadened or overlapping, demonstrating its versatility for complex mixtures. Furthermore, the thesis extends CSI-based titration to heterogeneous solvent systems: by establishing a solvent (water-Dimethyl sulfoxide) gradient, spatially resolved pH measurements were used to determine solvent dependent pK_a values (p$_s K_a$) of organic compounds across different solvent compositions. These data were extrapolated (via the Yasuda Shedlovsky method) to estimate the true aqueous pK_a, providing a novel route to obtain dissociation constants of sparingly water-soluble molecules.

Finally, a single-tube CSI method for metal ligand binding is presented. Here, a concentrated salt (calcium or magnesium acetate) is layered beneath an aqueous solution of a ligand (such as a polymer or small organic acid), generating an upward diffusion gradient of metal ions. ^{1}H CSI is then used to map spatial profiles of free metal concentration $[M^{2+}]_f$ (monitored via weak binding indicator molecules whose chemical shifts report on M^{2+} binding) and the total diffusant (acetate) concentration. From the mismatch between the acetate profile and the free M^{2+} profile, the local amount of metal bound to the ligand (B) is quantified. This approach was tested on strong, multivalent binders like oxalate or alginate which showed pronounced metal uptake, essentially sequestering Ca^{2+} and Mg^{2+} until saturation of their binding sites. In contrast, weak binders such as carboxymethyl cellulose (CMC) and gallic acid exhibited negligible binding, yielding free ion profiles almost overlapping with the control (no ligand) case. These findings quantitatively confirm expectations (e.g. alginate’s well-known Ca^{2+} crosslinking behaviour and CMC’s minimal interaction with divalent cations) and provide new insights such as the precise point of Ca^{2+} induced cellulose nanocrystal aggregation.

Overall, the thesis demonstrates that CSI can be harnessed to obtain detailed thermodynamic parameters such pK_a values and complexation capacities directly within an NMR tube. The developed techniques require only standard high-field NMR hardware and small sample quantities, yet achieve spatially resolved, multi point measurements in a single run. The

outcomes not only corroborate traditional measurements but also enable analyses previously impractical or impossible with conventional titrations and assays. This work significantly broadens the applicability of NMR in physical chemistry and chemical biology, offering efficient new tools for characterizing acid–base equilibria and metal–ligand interactions in complex or inhomogeneous systems.

Acknowledgements

Much thanks and gratitude for Professor Yaroslav Khimyak and Dr Matthew Wallace for their support, guidance and feedback throughout the PhD for which I am greatly grateful for. Their support was experienced both through their vast knowledge and experience and also through their friendship. I feel incredibly humbled and lucky to have had them as supervisors and couldn't have asked for better supervisors. Much thanks also for my fellow postdoc colleagues in the NMR office including Dr Trey Koev, Dr Serena Monaco and Dr Ignacio Delso for their support and guidance and feedback along with much thanks for the joyful working environment as well. I also am really thankful for UEA with their generous grant which funded the project.

List of abbreviations

Ac	Acetic acid
δ_{obs}	Observed chemical shift
CSI	Chemical shift imaging
DMSO	Dimethyl sulfoxide
DSS	2,2-dimethyl-2-silapentane-5-sulfonate
PFG	Pulsed field gradient
FOR	Formic acid
IUPAC	International Union of Pure and Applied Chemistry
K_{Ca}	Effective Ca^{2+} binding constant
K_{Mg}	Effective Mg^{2+} binding constant
MPA	Methylphosphonate
NMR	Nuclear magnetic resonance
2,6-DHB	2,6-dihydroxybenzoic acid
2-MI	2-methylimidazole
4-CN	4-cyanophenol
$CDCl_3$	Deuterated chloroform
CMC	Carboxymethyl cellulose

CNC	Cellulose nanocrystals
DS	Degree of substitution
D_2O	Deuterium oxide
MRI	Magnetic resonance imaging
PAA	Polyacrylic acid
T_1	Longitudinal (spin–lattice) relaxation time
T_2	Transverse (spin–spin) relaxation time
t_1	Preparation time
t_2	Detection time
WGA	Wheat germ agglutinin
δ_L	Deprotonated limiting chemical shift
δ_H	Protonated limiting chemical shift
STD	Saturation transfer difference
K_a	Acid dissociation constant

Table of Contents

Chapter 1: Introduction

1.1 Principles of nuclear magnetic resonance

Nuclear magnetic resonance (NMR) is a powerful spectroscopic technique that gives local information on the structure of chemical compounds in a non-invasive manner.[1] The technique is based on the fact that nuclei have an intrinsic angular momentum (I) which generates a magnetic moment (μ) that is proportional to I.[2] When those nuclei are placed inside an external magnetic field, an energy splitting called the Zeeman splitting occurs that favours nuclei precessing with the magnetic field (spin up) versus those precessing against the magnetic field (spin down) which can be described via the Boltzmann distribution which gives the energy difference between the two states (see figure 1.1).[3-5] Given the insights of Max Planck that energy of an electromagnetic wave is proportional to the frequency of the wave therefore measuring the frequency at which magnetic moment of nuclei precess around the external magnetic field will provide information about the energy of the system of interest.[6] The frequency at which the magnetic moment of nuclei precess around an external magnetic field is called the Larmor frequency (ω_0) and is dependent on the strength of the external magnet (B_0) and the gyromagnetic ratio (γ) (ratio of magnetic moment of a nucleus to its spin angular momentum). [7, 8]

$$\mu = \gamma I \tag{1.1}$$

$$\omega_0 = -\gamma B_0 \tag{1.2}$$

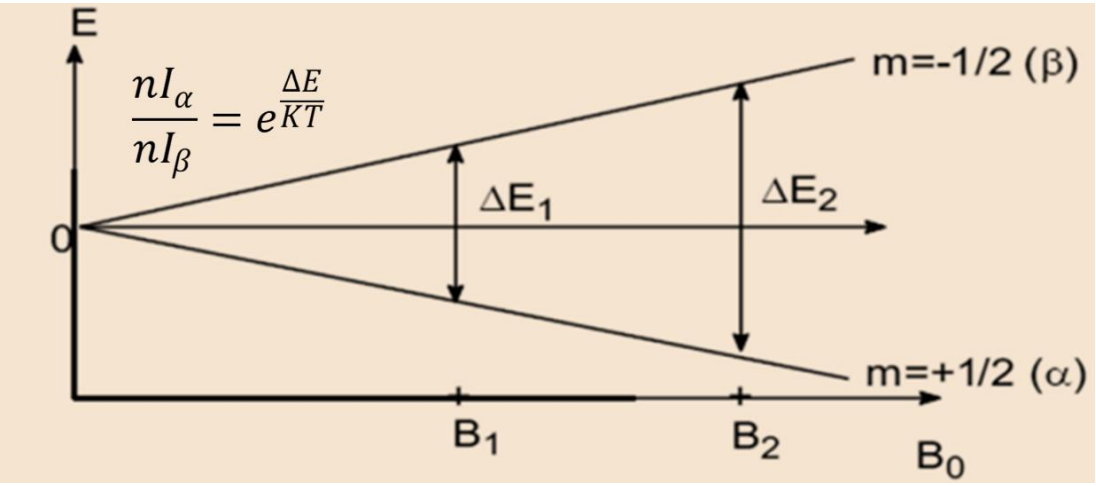


Figure 1.1. Graphical representation of Boltzmann distribution. When placed in an external magnetic field the nuclei are split into a low energy state α (also known as spin up) and the high energy state β (also known as spin down). The ratio of the two populations is given via the Boltzmann equation (top left) where I_α is the spin up state and I_β is the spin down state, in the respective populations, ΔE the energy difference between the two states, T is temperature and K is Boltzmann's constant

The nature and magnitude of the energy split is determined by the strength of the external magnet and the specific nature of the nuclei along with its local chemomagnetic environment.[9] As a result NMR has had a wide range of applications into structural elucidation of organic compounds[10], protein characterisation[11], metabolomics[12], purity analysis[13], drug discovery[14] and material science amongst other areas.[15]

1.2 Development of pulsed field gradients

One major step that was taken in the development of NMR was the use of pulsed field gradients (PFG) which are temporary magnetic fields that vary linearly with position, that exist in addition to the main external magnetic field B_0.[16] This causes the magnetic field around the NMR tube (B_z) to vary in the following manner:

$$B_z = B_0 + zG_z \tag{1.3}$$

Where B_0 is the static external magnetic field, z is position along the z axis and G_z is gradient strength along the z axis. Given that the Larmor frequency of a nucleus is dependent on the external magnetic field in the context of a field gradient the position dependent Larmor frequency (ω_z) is proportional to:

$$\omega_z = \gamma B_z = \gamma(B_z + zG_z) = \underbrace{\gamma B_0}_{\text{Gradient independent}} + \underbrace{zG_z\gamma}_{\text{Gradient dependent}} \tag{1.4}$$

The gradient independent term can be eliminated formally by representing the magnetisation in a rotating reference frame of γB_0.[17] Pulsed field gradients are created using gradient coils that are in Maxwell configuration (see figure 1.2) where two coils are placed carrying current in opposite directions.[18] This produces an incremental field that is zero at the magnetic isocentre but increases linearly in both the +z and -z direction. This field is added to the static magnetic field to produce a variable overall field. The size of the gradient can be controlled by varying the current and the sign by altering the direction of the current flow.

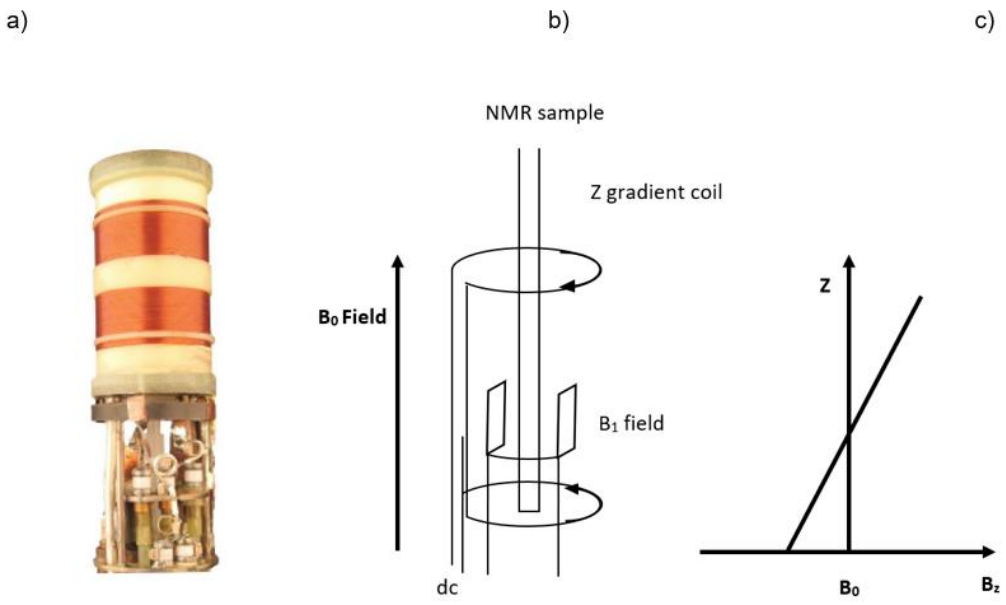


Figure 1.2. (a) NMR PFG coils (photograph taken from reference 19), (b) diagram of PFG coils (the two coils spin in opposite direction to give opposite impact to the external magnetic field B_0, (c) graph showing how PFG coils make the magnetic field to vary across the z spatial dimension with no change occurring at the isocentre.[19]

PFGs are extensively utilised in NMR spectroscopy now due to the many advantages it can bring. One is that it can eliminate unwanted magnetisations by inducing a phase shift (φ) to the magnetisation which at any position in the sample is given via equation 1.5:

$$\varphi_z = G_z z \delta \gamma \qquad (1.5)$$

where z is position from null point in NMR experiments (point where no phase shift occurs) and δ is gradient duration.[20] This means that the phase of any spin system will be different across the NMR tube hence they will "decohere" and will not be added to form an overall intensity and instead will cancel out (See figure 1.3).

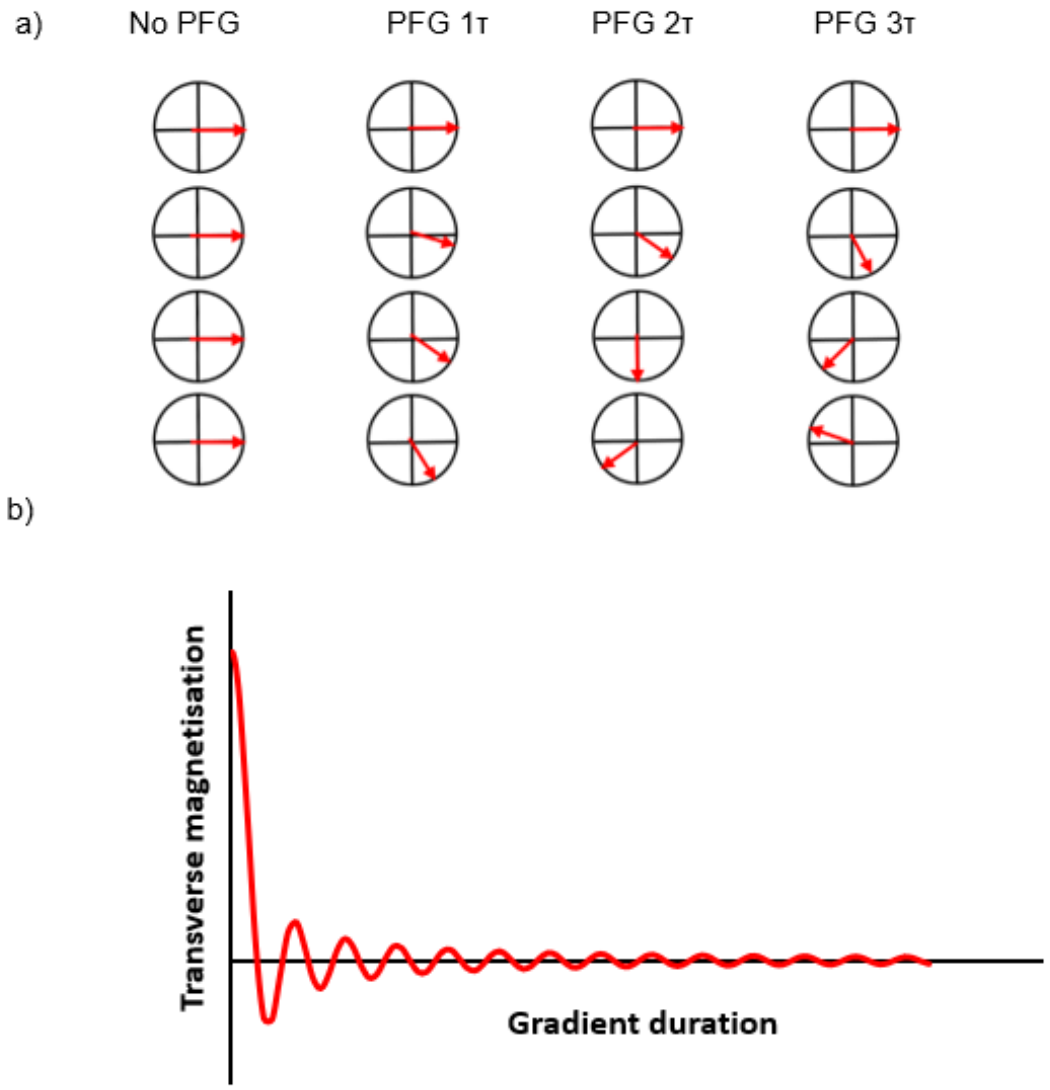


Figure 1.3. (a) vector representation of how a PFG dephases and removes signal (τ stands for unit time in which PFG is applied). As time progresses the different spins across the NMR tube start decohering with each other leading to loss of magnetisation in the transverse plane, (b) graph representing the loss of transverse magnetisation as the gradient is applied [21]

This technique, sometimes called homospoil gradient, is widely used to remove unwanted signals in the transverse plane.[22] It is also an important part of solvent suppression pulse programs like WATERGATE.[23] However, dephasing under PFG is a defined, "coherent" dephasing, not a random dephasing. The destructed coherence can be brought back before diffusion randomizes it spatially. This is done by implementing a second PFG which "refocuses" the magnetisation by removing the phase shift that was induced by the first PFG.[24] This means that any remaining decoherence is due to diffusion of the spins between the two gradient pulses. This allows for the implementation of Diffusion order spectroscopy (DOSY) experiments that are used for measuring the diffusion coefficient of compounds via observing the different phase shifts that compounds accrue after the implementation of the dephasing and rephasing gradients respectively.[25-27]

This fact is also quite useful in that it allows PFGs to be used for selecting desired magnetisation whilst suppressing unwanted magnetisation. This can be explained by showing that equation 1.5 assumes that the spins in the experiment have the same coherence order (coherence order being a parameter that says how many energy levels change when a coherence evolves). Specifically this means that the difference in magnetic quantum numbers between the two quantum states is the same. However, in situations where the system has spins which have different coherence orders then equation 1.5 needs to be modified to be:

$$\varphi_z = -pG_z z\delta\gamma \tag{1.6}$$

Where p is coherence order, Different coherence orders can arise due to different spin transitions such as single quantum coherence which has a coherence of ± 1 which represent normal spin-flip transitions or zero quantum coherence (when two spin states change together without changing their net quantum number) or ± 2 where involves two simultaneous spin flips (in this case phase shift doubles because both spins evolve together). This has really important implications because this means that the phase shift that spins acquire changes with respect to the coherence order that the spins have. This is utilised in NMR by setting coherence order -1 as the only observable coherence order. This means that all spins which have an initial coherence order p_1 need to have a final coherence order p_2 of -1. This sets up the following refocusing condition:

$$\frac{G_{z,1}\delta_1}{G_{z,2}\delta_2} = \frac{-p_2}{p_1} \tag{1.7}$$

Thus, by controlling the strength or duration of the second gradient one can control which coherence order get refocused and which gets suppressed. Additionally, in experiments that contain spins of different nuclei another possibility presents itself where spins that have a particular coherence order of one nuclei can be selected whilst spins that have the same coherence order of another nuclei can be suppressed this is because the phase shift that spins acquire is proportional to the gyromagnetic ratio of its respective nuclei.[28] Thus in heteronuclear experiments the refocusing equation 1.7 gets reformulated to obtain:

$$\frac{G_{z,1}\delta_1}{G_{z,2}\delta_2} = \frac{\gamma_I}{\gamma_S - \gamma_I} \tag{1.8}$$

Where γ_I and γ_S are the gyromagnetic ratios of the abundant and rare spin respectively. The use of this method is illustrated in the double quantum filtered COSY (DQF COSY) experiment where only magnetisation from the double quantum terms is selected for detection in the t_2 time (the time during which the signal is detected by the spectrometer) via the use of PFGs.[29, 30] This allows for obtaining of COSY spectra that are of higher intensity and have better phasing (See figure 1.4).[31]

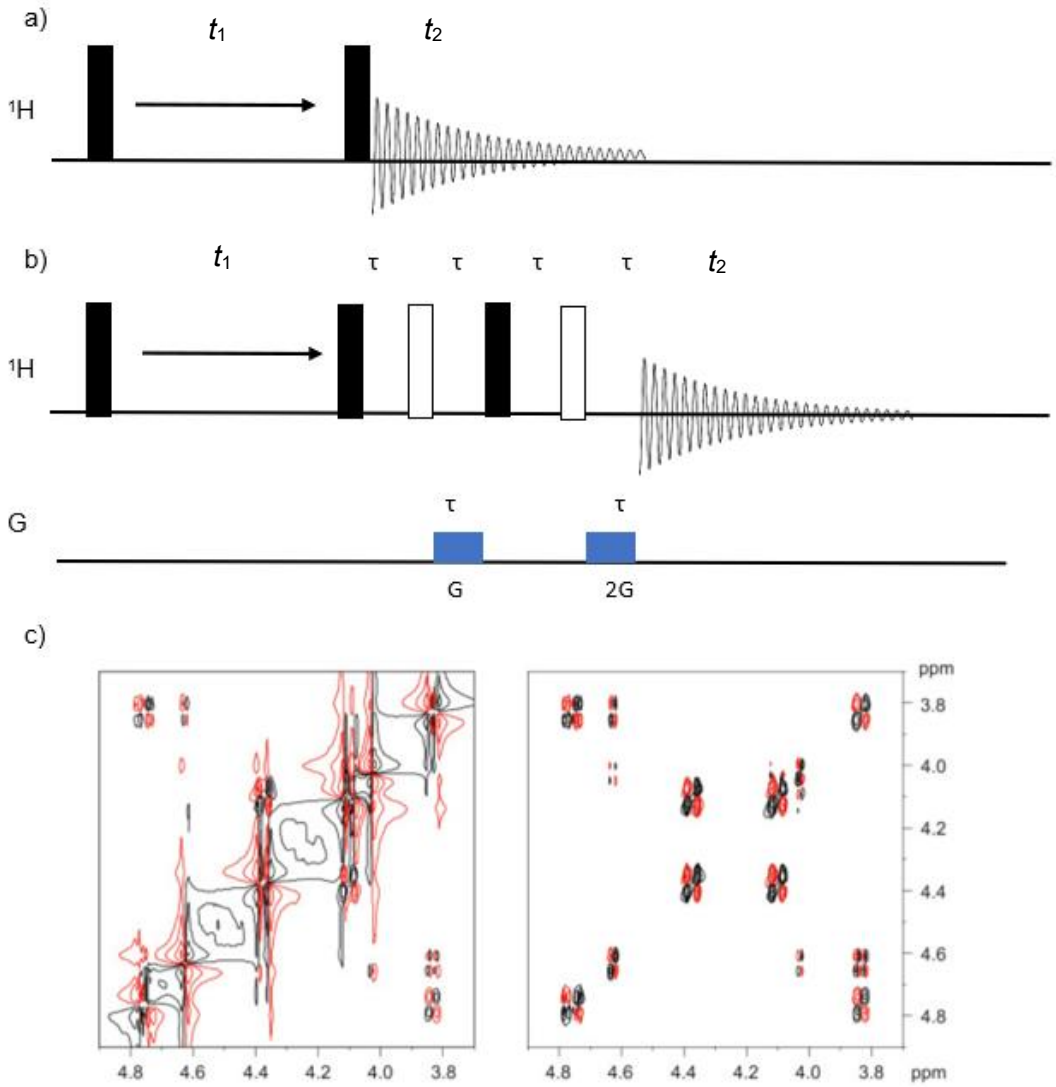


Figure 1.4. (a) NMR pulse sequence for COSY (black filled rectangular pulse is 90° hard pulse, white filled rectangular pulse is 180° hard pulse and PFGs are represented as blue rectangles, t_1 time also known as evolution period is the time during which the spins evolve after an initial preparation period and before the next step of the pulse sequence and t_2 time also known as the detection period is the period in which FID is being acquired) , (b) NMR pulse sequence for DQF COSY , the second gradient having twice the gradient strength as the first one whilst keeping the duration equal ensures only double quantum magnetisation is detected (the additional pulses in the proton channel are there to remove any unwanted phase errors that might accumulate during the experiment), (c) display of COSY 2D spectrum (left) and DQF COSY spectrum (right), highlighting the improved phasing and sensitivity of DQF COSY compared to COSY, taken from Reference [31,30]

1.3 Use of pulsed field gradients for spatially resolved NMR

The potential for PFGs to encode NMR information was noticed early on in the history of NMR. In 1965 E. O. Stejskal and J. E. Tanner first demonstrated that pulsed magnetic field gradients could be used in a spin-echo experiment to measure diffusion, introducing the now-classic

pulsed-gradient spin-echo (PGSE) method.[32] This was a pivotal moment showing that gradient pulses can encode information (in that case, molecular motion) in NMR signals. By the early 1970s gradients were beginning to be used to encode spatial information. Lauterbur's 1973 experiment was the first to obtain a 2D image using NMR by applying field gradients, effectively creating an NMR image of tubes of water (Initially, he called the technique Zeugmatography: "Because the interaction may be regarded as a coupling of the two fields by the object, I propose that image formation by this technique be known as zeugmatography, from the Greek ζενγμα, that which is used for joining").[33, 34] Lauterbur showed that by applying a gradient and correlating the NMR frequencies with position, one could reconstruct an image (see figure 1.5).

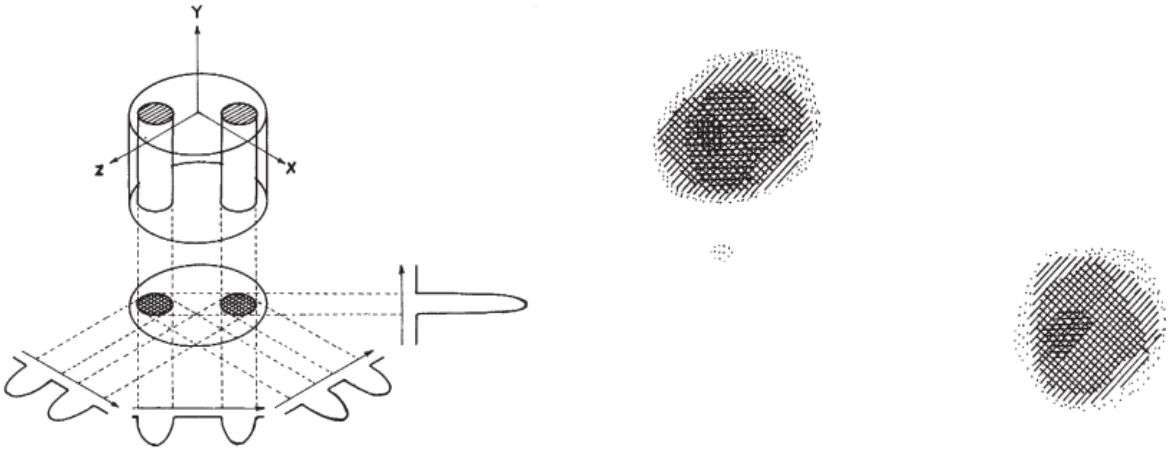


Figure 1.5. (Left) Lauterbur's original schema for NMR imaging where two small tubes of water placed millimetres apart inside a magnetic field gradient, Lauterbur repeated the measurement multiple times each time rotating the direction of the gradient, collecting a new frequency-encoded projection of the sample from a different angle, (right) reconstructed two dimensional ^{1}H NMR image of the samples [34]

Around the same time as Lauterbur, Peter Mansfield developed techniques for rapid NMR imaging (such as echo-planar imaging) using oscillating gradients.[35] Additionally he came up with the idea of "slice selection" whereby a linear field gradient could be used to localize the NMR signal on a slice-by-slice basis. Mansfield's experimental setup involved stacking multiple 1-mm-thick sheets of solid camphor into the bore of an NMR spectrometer. Applying a magnetic field gradient perpendicular to the sheets, Mansfield measured the transient NMR signal response to an applied RF-pulse. Interference peaks similar to those seen in x-ray diffraction were observed, which when inverse Fourier transformed revealed discrete layers of the camphor sample (see figure 1.6). [36]

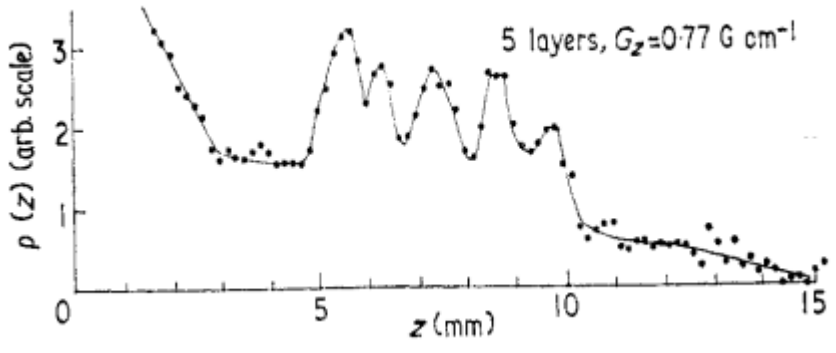


Figure 1.6. From Mansfield (1973) Five peaks corresponding to five stacked blocks of solid camphor [36]

Lauterbur and Mansfield's breakthroughs laid the foundation for MRI, earning them the 2003 Nobel Prize in Physiology or Medicine.[37] In magnetic resonance imaging (MRI) the pulsed field gradient is used to spatially encode the NMR signal to form an image.[38] This approach is used to give location and relaxation properties and to differentiate different cell and tissue types and is used to identify tumour.[39,40] It was also used in non-medical cases such to do non-invasive imaging of phosphorus metabolites in biological systems.[41] However the pulsed field gradients used in MRI modify the external magnetic field to a much greater extent than the chemical shift shielding such that in conventional MRI the presence of a field gradient during signal acquisition results in loss of the chemical shift spectral information. [42]

In the 1980s efforts were undertaken to develop methods that use pulsed field gradients to encode spatial information whilst also obtaining the chemical shift information traditionally obtained via NMR (eventually the term "chemical shift imaging" was used to refer to such techniques).[43] Some of the earliest chemical shift imaging experiments were ones developed by Brown *et al.* where they described a method to obtain the 3D distribution of chemical species by combining phase-encoded imaging with spectroscopic readout. Allowing them to map phosphorus metabolite concentration in vivo. [44] Around the same time Maudsley *et al.* proposed a method of obtaining spatially resolved chemical shift information via a "four-dimensional" ^{31}P NMR experiment where the method integrates multiple magnetic field gradients in the evolution period of the NMR experiment to obtain data from different spatial regions (x, y, and z axes), creating a four-dimensional dataset. The data is processed through a Fourier transform to reveal both spatial information and high-resolution NMR spectra.[45] Peter Mansfield also published two chemical shift imaging methods that allow for the spatial and chemical shift information to be simultaneously observed via what he termed "echo-planar" NMR imaging which is an ultrafast MRI technique that enables image acquisition in milliseconds to seconds. EPI uses rapid gradient switching, alternating the readout gradient's polarity while collecting multiple echoes. This allows full 2D image data acquisition in 20 –100 ms for single-shot EPI, or slightly longer for multi-shot variants. [46]

Early CSI experiments were limited by hardware – gradients were relatively weak and slow by modern standards – but they proved the principle that PFGs could localize spectra in space. Throughout the 1980s and 1990s, as stronger gradient coils and better amplifiers became available, NMR imaging transitioned from laboratory experiments to widespread applications, and CSI techniques advanced in tandem. Gradients began to be incorporated into high-resolution NMR spectrometers (not just whole-body MRI machines), enabling chemists to perform spatially resolved NMR on samples in ordinary NMR tubes.

Nowadays, modern CSI employs three primary methodologies to measure spatially dependent components and, consequently, determine the position of spins within a sample:

1. **Frequency Encoding**: During acquisition, the application of a gradient allows for the frequencies situated inside the free induction decay (FID) to vary as a function of position encoding the frequencies of nuclei with spatial information. However, this method becomes complex when multiple signals are present in the corresponding spectra, as variations in chemical shifts may be misinterpreted as spatial displacements.[47]
2. **Selective Excitation**: Gradients induce shifts in Larmor frequency, allowing for selective excitation through shaped pulses (pulses that have a variable phase and amplitude compared to normal 'hard' pulses and typically have a longer duration and thus have a narrow and tuneable excitation or inversion bandwidth). Since each scan captures signals from only a single slice, full sample measurement necessitates multiple experiments in a 2d manner.[48]
3. **Phase Encoding**: By applying a gradient during for an incremented time period or by incrementing gradient strengths during a fixed delay, it is possible to perform a 2D experiment, capturing the gradient-induced frequency in the indirect dimension.[49]

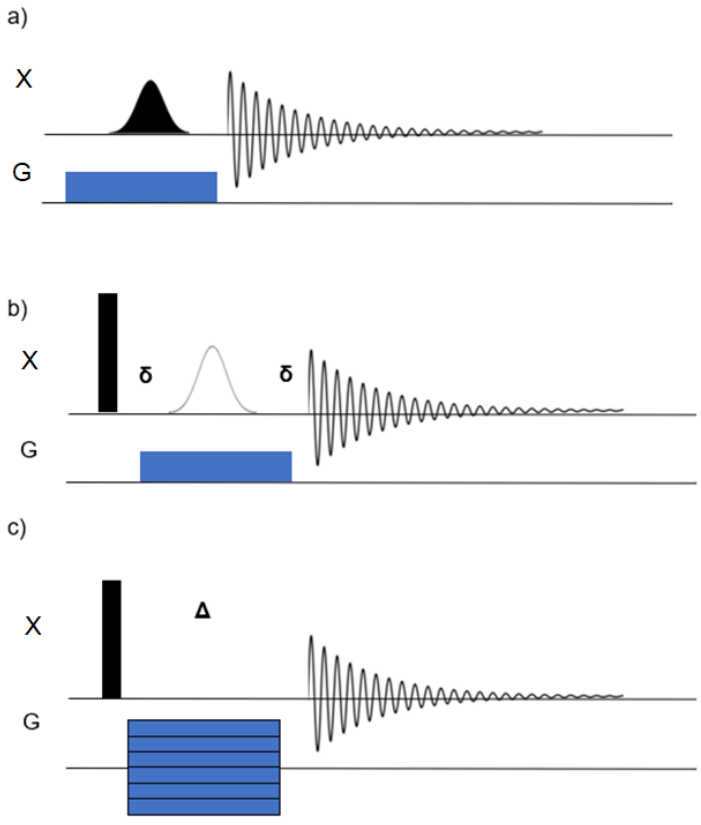


Figure 1.7. Different z-axis imaging sequences used in magnetic resonance imaging and spectroscopy (X stands for any NMR active nucleus): (a) This sequence employs a frequency-selective RF pulse in the presence of a z-axis gradient to excite only a specific region of the sample. The black filled Gaussian curve represents the 90° selective excitation pulse; (b) In this sequence, a broadband RF pulse is initially applied to excite the entire sample. Following this, a space-selective refocusing pulse is applied in the presence of a gradient, which selectively refocuses spins in a specific region (The white filled Gaussian curve represents the

180° selective excitation pulse and the δ represent spin echo delays)[50]; (c) In this sequence phase encoding is achieved by applying a gradient along the z-axis with varying strength increments (Δ represents gradient duration). The spins acquire different phases depending on their position along the z-axis, which allows for the capture of spatial information.[51]

Experiments that utilise the modern selective excitation method of using shaped pulses to excite spins of a sample along a partial axial "slice" were proposed by Kozminski in 2000, using a double pulsed field gradient spin echo (DPFGSE) element to selectively acquire high resolution ^{1}H NMR spectra in a two-phase $CDCl_3/D_2O$ sample in the same NMR tube.[52] Another example is the application of chemical shift imaging to simultaneously acquire a total of sixteen 1D ^{1}H NMR spectra from a 16-capillary bundle inserted in a standard 10 mm probe.[53]

More recently, CSI experiments that utilise the phase encoding method proved to be very useful in titration experiments where the titrant is allowed to diffuse across the NMR tube in a controlled manner and where chemical shift imaging allows for the obtaining of spatially sensitive spectra that are connected to a specific location on the NMR tube and hence to specific quantity of titrant across the NMR tube. This approach significantly reduces experimental time and material cost by performing the equivalent of multiple titration experiments in one go. Niklas *et al.* first proposed this method and applied it on the titration of lithium perchlorate with 12-crown-4 ether (a common crown ether that coordinates lithium ions) inside a single NMR tube by making a 12-crown-4 concentration gradient, obtaining smooth ^{7}Li chemical shift titration curve which displayed the formation of $[Li(12\text{-crown-}4)]^{+}$ and $[Li(12\text{-crown-}4)_2]^{+}$ complexes.[54] The method was subsequently applied by Mitrev *et al*. to determine the binding constant (K_a) and free energies (ΔG) of β-cyclodextrin with paracetamol through the use of agar gels as a porous matrix to create stable concentration gradients for studying molecular interactions via slice-selective NMR spectroscopy.[55]. Wallace *et al.* used CSI to study the binding of calcium ions and the resulting gelation of peptide assemblies at alkaline pH by varying the calcium concentration across the NMR tube and using metal ion-binding molecules to measure degree of calcium binding on the surface of gel fibres.[56] Shortly thereafter Wallace *et al.* varied the pH across the NMR tube and performed ^{1}H NMR CSI experiments using a set of NMR pH indicator compounds to measure the pH of a sample as a function of position and obtain the pK_a of the compound of interest in a "single shot" experiment (see figure 1.8).[57]

This approach was also incorporated into ligand binding analysis by incorporating CSI into saturation transfer difference (STD) which is an NMR experiment that is used in the study of ligand-receptor interactions and is particularly useful for identifying and characterizing binding sites and mapping the interactions of small molecules (ligands) with larger biomolecules (receptors, such as proteins or nucleic acids). [58] Monaco *et al.* combined STD with CSI to perform an imaging saturation transfer difference NMR experiment where controlled concentration gradients of small molecule ligands are developed to determine protein-ligand dissociation constants and to assess the binding specificity in a single NMR tube, avoiding time consuming titrations (see figure 1.9). [58] Compared to the traditional method where the ligand/protein ratio is changed manually, this approach allows for determining binding epitopes in a fraction of the conventional experimental time. Imaging STD can also be used to screen for non-specific binders, by monitoring any variation of the binding epitope map at increasing ligand/protein ratios. [58]

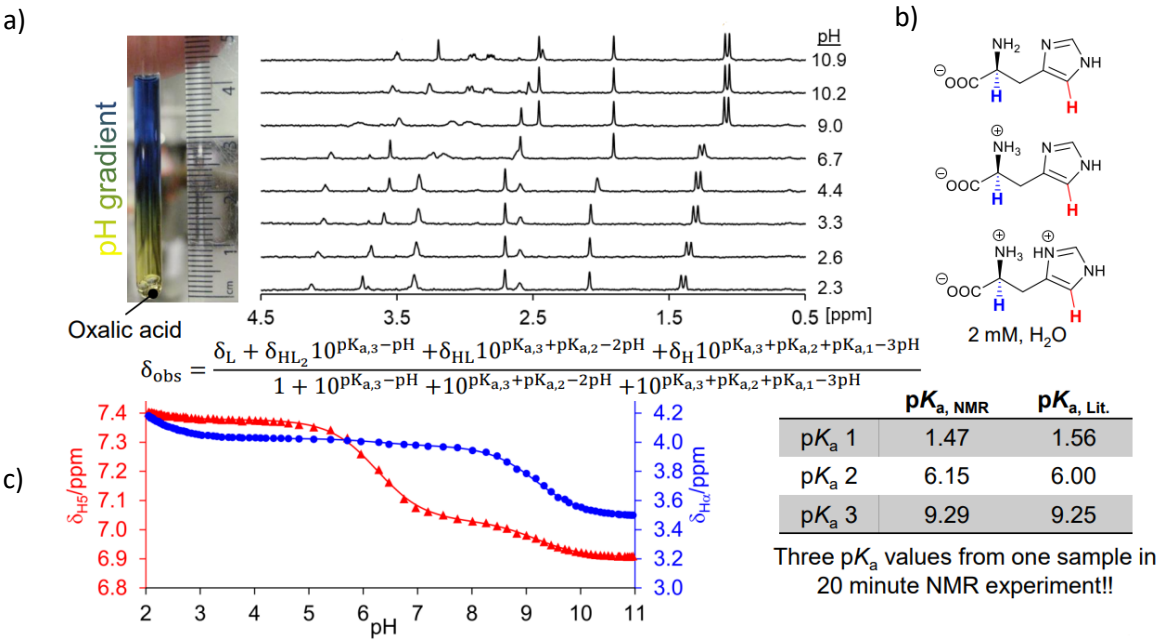


Figure 1.8. Determination of pK_a of histidine via pH gradient and ^{1}H CSI:[59] (a) pH gradient is made by putting a diffusing base or acid (oxalic acid in this experiment) powder at the bottom of the NMR tube allowing it to diffuse. Obtaining 1D spectra that have unique pH values across the tube, (b) the different ionisation states that histidine can take, (c) chemical shift of histidine as a function of pH and the fitted pK_a values of histidine associated with each ionisation state along with their equivalent literature values

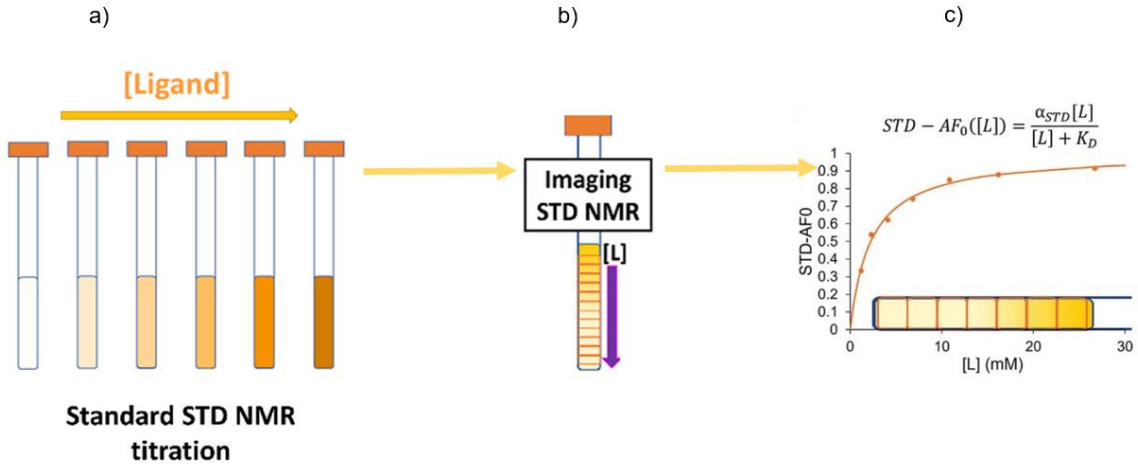


Figure 1.9. Scheme of imaging saturation transfer difference experiment: (a) Conventional method of running STD NMR where multiple NMR samples are prepared with different ligand concentrations, (b) imaging STD NMR with ligand inserted at top of NMR tube, allowed to diffuse down and spatially resolved NMR experiment provides “slices” across the NMR tube each with a different ligand concentration, (c) Fitting of STD data versus ligand concentration across NMR tube allows for determination of binding constant (K_d) [58]

The CSI methodology used throughout the thesis works by having a phase encoding gradient where the signals of everything are dephased except the one that is at the centre of the gradient. This variation in intensity as a function of position is encoded in an indirect dimension as a cosine function:

$$I_{nmr} = \int_{-\infty}^{\infty} I_z[\cos(\gamma G_z \delta z) + i \sin(\gamma G_z \delta z)] \, d_z \tag{1.9}$$

where z is position from null point in NMR experiments (Point where no phase shift occurs) and δ is gradient duration. This provides a way of encoding position dependent NMR spectra by encoding the position of the NMR spectra as a cosine function in the indirect dimension.[51]

Shown below is the basic outline of the pulse sequence utilised for our CSI experiments. They are based on the pulse sequences showcased in Trigo-Mouriño *et al.*[51] They include a homospoil gradient at the start to remove any transverse magnetisation from previous scans and then a water suppression sequence called excitation sculpting in which a dephasing PFG is done after a 90° pulse is given, causing for transverse magnetisation to dephase. Subsequently, a selective 180° pulse is done that acts on the solvent peak followed by a hard 180° pulse.[60] A rephasing PFG is then implemented which with respect to the non-solvent peaks will recohere the signals as previously described whilst for the solvent signal the dephasing will continue to increase leading to a reduction of solvent signal (sequence repeated twice for improved solvent suppression). What follows is the phase encoding sequence as described previously (delays and pulse lengths were omitted from description for simplicity. Please see Section A1 for delays, pulse lengths and phases used).

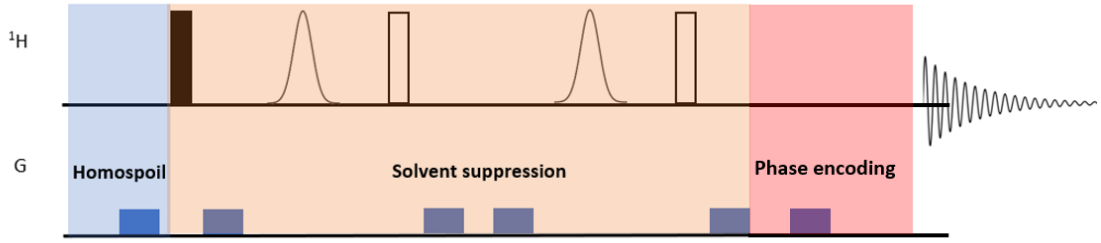


Figure 1.10. Chemical shift imaging sequence utilised in this thesis. Solid black rectangles represent 90° hard pulses. Open rectangles represent 180° hard pulses and open gaussian shapes correspond to 180° shaped pulses. It is broken down into three sections: (blue) homospoil gradient to destroy transverse magnetisation, (orange) solvent suppression sequence through excitation sculpting, (red) phase encoding sequence.

A series of FIDS, each one at a different strength of the encoding gradient are obtained and then Fourier transformed in both dimensions to get a 2D NMR spectra (see figure 1.11). The FID can be thought of as a combination of multiple 1D FIDs with each one having a different phase encoding associated with it.

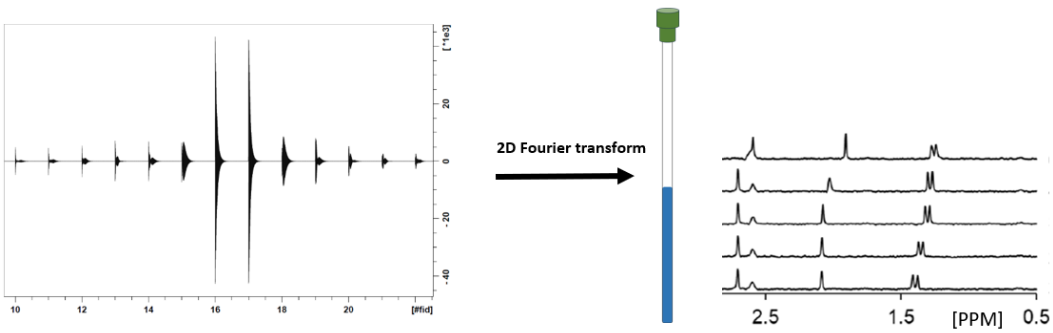


Figure 1.11. processing of CSI FID data using 2D Fourier transform to obtained spatially resolved 1D spectra

Despite its advantages, CSI is not without limitations. One of the primary challenges in CSI experiments is the sensitivity of phase encoding to magnetic field inhomogeneities. Unlike standard NMR spectroscopy, where signal acquisition occurs in a relatively uniform field, CSI relies on precise magnetic field gradients to encode spatial information.[61] Any deviation from a uniform field can lead to chemical shift distortions, peak broadening, and signal loss, particularly when working with samples that contain multiple phases or interfaces between different media.[62, 63] To mitigate this issue, shimming is required to optimize field homogeneity before CSI data acquisition.[64] In cases where field drift occurs over time, reference compounds such as 2H-labeled solvents (e.g., $CDCl_3$ or D_2O) can be used to provide an external lock signal that allows for continuous correction of field fluctuations.[65] When running a sample unlocked it is useful to perform a 1D experiment with a solution of $CDCl_3$ before the analyte's sample is run as that helps to correct and reset the field strength for the effect of magnetic drift accrued and thereby allowing peaks to appear in their correct position.

In addition, because each phase-encoded spectrum has a lower signal intensity due to the division of available magnetization across multiple spatial slices, CSI inherently has lower sensitivity compared to traditional NMR.[66] Hence, having adequate concentration of the analytes studied is important.

Another limitation is the inherently long acquisition time associated with phase encoding, as multiple gradient increments must be recorded to achieve high spatial resolution. [67] This can be mitigated by optimizing pulse sequences and leveraging modern data processing techniques, such as compressed sensing and non-uniform sampling.[68] Another challenge in CSI experiments is spectral overlap, particularly in complex mixtures. Since each 1D spectrum corresponds to a different spatial position, overlapping peaks can complicate data interpretation. Deconvolution methods and multidimensional NMR experiments can be employed to address this issue, but they require advanced processing and additional experimental time.

Additionally, because the radiofrequency coil used in NMR experiments does not produce a perfectly uniform magnetic field across the entire length of the NMR tube, the intensities of the sample's signal will vary across the NMR tube. Therefore, it is important to have an integral reference in the sample to have reliable integral information that can be compared between different slices. If the recycle delay is too short to allow for full relaxation of longitudinal magnetisation of all signals, then a correction factor needs to determine between the integral reference and the integrals of the analyte.

1.4 Comparison of chemical shift imaging NMR to conventional methods for pK_a determination

It's instructive to compare CSI to other techniques in terms of sensitivity, accuracy, sample needs, and practical constraints. Below CSI is contrasted with several common methods – potentiometric titrations, UV-Visible spectroscopy, fluorescence assays, isothermal titration calorimetry, and other NMR techniques – For the sake of simplicity and relevance to the results chapter CSI is compared to other techniques with respect to pK_a and divalent ion binding determination, highlighting the pros and cons of each.

Potentiometric titration. Potentiometric titration is a classical and widely used method to determine pK_a values. In a potentiometric pK_a measurement, one monitors pH (using a glass electrode) while incrementally adding acid or base to the sample (see figure 1.12).[69]

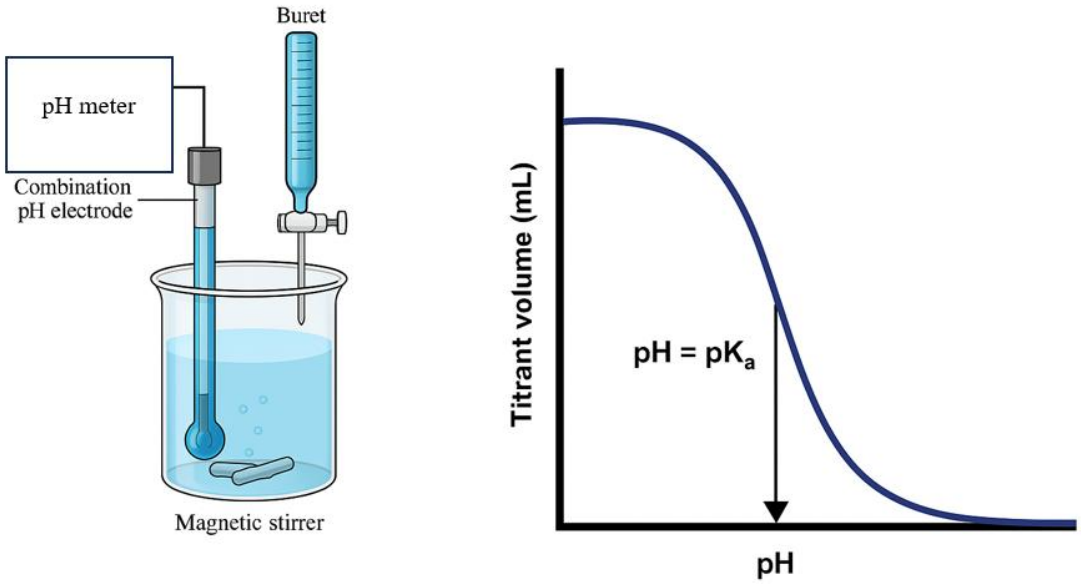


Figure 1.12. Schematic representation of potentiometric titration, (left) typical potentiometric titration setup, which includes a beaker with a magnetic stirrer, a combination pH electrode connected to a pH meter (with millivolt readout), and a burette delivering titrant into the analyte's solution. The right panel shows a titration curve plotting pH against titrant volume. The inflection point of the sigmoidal titration curve indicates the pK_a of the analyte

The inflection point of the pH vs. volume curve yields the dissociation constant. The major advantages of potentiometric titration are its simplicity, low cost, and direct measurement of H^+ concentration. [70] [69] However, this method has limitations. It generally requires relatively high analyte concentrations and sample volume for reliable results. This lower concentration limit arises because the electrode potential change at very low analyte concentrations becomes too small to measure accurately. [70] If one's sample is scarce or poorly soluble, potentiometric titration can be impractical for example, many sparingly soluble or non-aqueous systems cannot be directly titrated without special techniques. [71] Another limitation is that potentiometry measures the bulk solution pH, so it cannot easily deconvolute mixtures. If a solution contains multiple acids or bases, the electrode reports an overall titration curve; distinguishing individual pK_a values can be complex or impossible without assumptions. In contrast, NMR-based methods like CSI can target specific molecules or functional groups even in mixtures, since each chemical shift corresponds to a particular species.[69] Potentiometric titration also requires calibration and careful control of conditions (ionic strength, temperature) to get accurate absolute pK_a values errors like junction potentials or "alkaline error" of glass electrodes must be managed.[72]

Compared to CSI: potentiometry typically offers high accuracy and does not require expensive instrumentation (just a pH meter), but it needs more sample (millimolar concentrations, larger volumes) and provides less molecular detail (no spectra). CSI, on the other hand, can work with smaller quantities and simultaneously observe multiple species, albeit with the need for an NMR spectrometer.

UV Spectroscopy. UV spectroscopy is another method to measure pK_a or binding events, provided the species of interest has an optical absorbance that changes with its protonation or binding state. In a pK_a context, one can record the UV-Vis spectrum of an indicator or the analyte itself at varying pH and track changes in absorbance at a particular wavelength to

determine the pK_a.[73] The absorption is measured using two wavelengths across a pH range, where the second wavelength is used to provide line of reference. The ratio of absorption at the two wavelengths is plotted against the pH of the solution. The plot is in the form of a sigmoidal curve and the pK_a of the functional group can be estimated from the inflection point (see figure 1.13). [70]

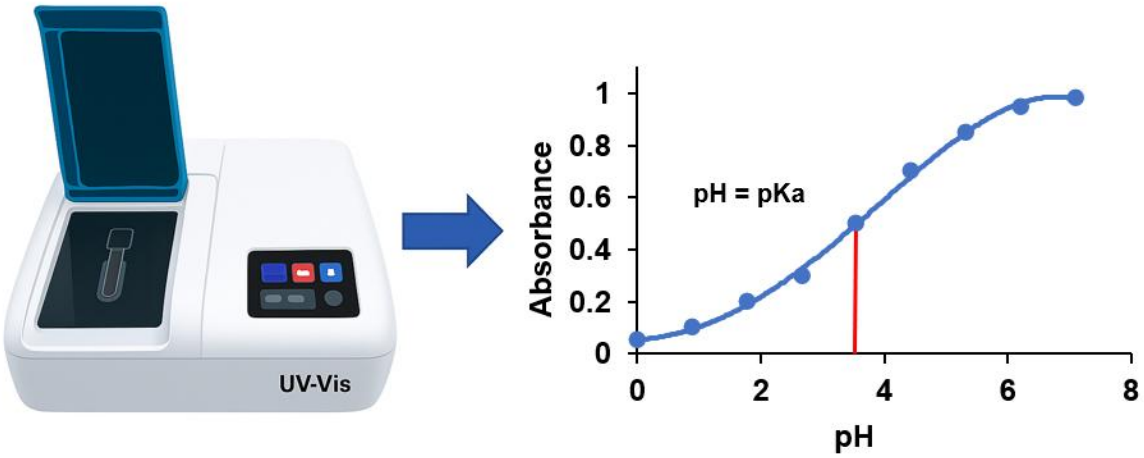


Figure 1.13. pK_a determination using UV-Visible spectroscopy, (left) UV spectrophotometer, (right) absorbance vs. pH curve. As the pH of the solution is varied, changes in absorbance are measured at a specific wavelength corresponding to the protonated and deprotonated forms of the analyte. The resulting sigmoidal curve allows determination of the pK_a value

The advantages of UV methods are their sensitivity and throughput. UV can often detect micromolar or even lower concentrations of analyte, far beyond the sensitivity of NMR-based methods. Indeed, where potentiometric titration fails at low concentration, UV spectrophotometry can still yield accurate pK_a values. [74] It's also relatively quick: obtaining a full titration curve by UV absorbance (especially with an automated titrator or multi-well plate reader) is fast and can be high-throughput.[69] However, UV methods require that the molecule (or a proxy indicator) has a distinct absorbance for each state.[75] In comparison to CSI, UV lacks the ability to directly distinguish multiple components in one solution unless their spectra are known and resolved. It also provides no structural information beyond absorbance changes. Sample requirements for UV-Vis are usually modest (a few mL or less of micromolar solution), which is an advantage over many other techniques. One limitation is that the solution generally must be optically clear; turbid or strongly coloured solutions (or biological samples with high background absorbance) can interfere.[76] In summary, UV-Vis is highly sensitive and rapid, ideal for many small-molecule pK_a measurements, but it is constrained to systems with an optical signal and may require external indicators and has difficulty doing accurate measurements in complex mixtures. CSI, by contrast, can probe systems with no chromophores (since NMR detects NMR-active nuclei inherently) and can do so in opaque or heterogeneous samples, but needs higher concentrations (typically millimolar) and longer times.

Fluorometry. Fluorometry operates on a similar principle to UV but instead of UV-absorption fluorometry measures pK_a based on the difference in the fluorescence spectrum between a free acid or base and its conjugated form.[69] Many biochemical properties (pH changes, metal binding, protein-ligand interactions) can be monitored via fluorescent dyes or intrinsic fluorescence changes. For instance, Ca^{2+} binding is commonly measured with fluorescent Ca^{2+} indicators like Fura-2, Fluo-4, or genetically encoded calcium sensors.[77] [78, 79]

These indicators change their fluorescence intensity or wavelength upon binding calcium. The detection limits of fluorescence methods are extremely low nanomolar to picomolar concentrations of analyte can often be detected due to the high signal-to-noise of fluorescence measurements. As an example, the calcium dye Fura-2 has a dissociation constant around 145 nM for Ca^{2+}, making it suitable to detect free Ca^{2+} in the nanomolar range.[80]

Such sensitivity is better than NMR or UV, meaning fluorescence can be used for trace analyses or very small sample volumes (including single-cell measurements in microscopy). Fluorescence assays are also generally fast and can be multiplexed (e.g., multi-well plate readers scanning dozens of samples in minutes). However, fluorescence methods require a responsive probe, which can be a limitation. If the molecule of interest is not inherently fluorescent or doesn't change fluorescence upon the event of interest, one must introduce a fluorescent tag or use a coupled indicator reaction.[81] This can complicate the experimental design for example, adding a fluorescent Ca^{2+} indicator to measure Ca^{2+} binding to a polymer might work, but one has to ensure the indicator itself does not perturb the binding equilibrium significantly.[82] Another consideration is specificity: fluorescence signals can be affected by other quenching or background fluorescence in the sample.[82] For instance, biological samples often auto fluoresce, which can raise the background. Careful calibration and controls are needed to ensure the fluorescence change is truly due to the binding or dissociation of interest. In terms of accuracy, fluorescence titrations can be very accurate in determining binding constants or pK_a (often comparable to other methods), but one must calibrate the fluorescence response to concentration changes (e.g., via known standards or inner-filter effect corrections). [69] Compared to CSI, fluorescence assays are more sensitive but they do not provide the rich chemical information that NMR does. CSI can directly observe the chemical species and shifts, whereas fluorescence gives a one- or two-dimensional readout (intensity ± wavelength) linked to the process. In scenarios where applicable, fluorescence is often the method of choice for measuring low-level Ca^{2+} binding or pH changes (for example, in live cell studies or rapid kinetics), whereas CSI might be chosen for detailed analysis in vitro when one wants to see the actual chemical environment changes (and can afford the sample quantity).

Isothermal Titration Calorimetry. Isothermal titration calorimetry (ITC) is a label-free technique that measures the heat change associated with a binding event or reaction, widely regarded as a standard for determining binding thermodynamics see (figure 1.14).[83] In an ITC experiment, one titrates a ligand (e.g., Ca^{2+} or a proton, via acid/base) into a solution of the binding partner and directly measures heat absorbed or released. The result is a binding isotherm from which the binding constant (K_a or K_d), reaction enthalpy (ΔH), and often entropy (ΔS) can be derived.[84] The biggest strength of ITC is its comprehensive thermodynamic information – it not only gives an affinity but also the enthalpy change, which no spectroscopic method directly provides. [85]. Despite these advantages, ITC has notable limitations. It is considered a low-sensitivity technique in terms of concentration – it typically requires relatively high concentrations of the interacting species and large sample volumes for a single titration .[69] Another challenge with ITC is that it measures aggregate heat: any side reactions or dilution heats will be included, so careful buffer matching and controls are required.[86, 87] It's also a relatively slow method, as one titration experiment can take an hour or more.[88] Both potentiometric and spectroscopic titrations usually provide only the binding constant (and maybe estimate ΔH if van Hoff analysis is done), whereas ITC directly gives ΔH. This extra information is a big plus for ITC in biochemical studies. In terms of equipment, ITC is a specialized instrument and not as commonly available as UV spectrometers or NMR machines, and it requires careful operation.

By comparison, CSI uses standard NMR hardware (which many labs have, though NMR itself is expensive equipment). If we compare ITC and CSI: CSI can determine an equilibrium

constant by NMR chemical shift changes, but it does not directly give enthalpy or entropy. CSI might require less sample volume, but NMR has much higher instrument costs and complexity in data analysis. ITC is highly accurate for affinity but is laborious and time consuming whereas NMR CSI experiments can obtain binding or pK_a information in a much faster, less material costly manner.

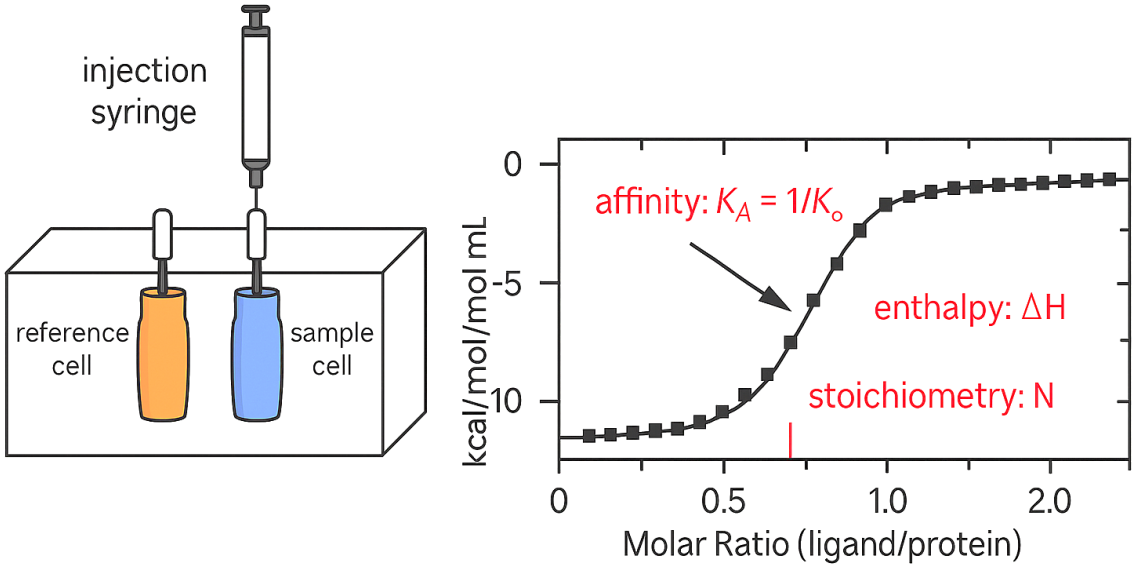


Figure 1.14. (left) Isothermal titration calorimetry setup, where a ligand is titrated into a sample cell containing protein, alongside a reference cell, with heat changes monitored upon each injection, (right) binding isotherm plot of the heat change (kcal/mol ligand) against the molar ratio of ligand to protein with key thermodynamic parameters K_a the enthalpy change (ΔH) of binding, and binding stoichiometry (N)

Other NMR-Based Techniques

Apart from CSI, there are other NMR methods to measure pK_a and binding properties. The most direct is a simple NMR titration: one acquires a series of 1D ^{1}H NMR spectra (or ^{13}C, ^{15}N, etc., depending on the nucleus of interest) at different conditions (see figure 1.15 for example of determining pK_a using NMR titrations).[57, 89] By observing chemical shift changes of particular signals, one can infer the pK_a or binding constant. For example, pK_a values have long been determined by measuring the pH-dependent shift of a proton resonance and fitting it to the Henderson-Hasselbalch equation. This NMR titration approach is quite reliable and was shown to agree well with standard methods. It has the advantage of assigning the change to a specific functional group (the proton giving that NMR signal). Moreover, NMR can probe multiple sites at once: if a molecule has two ionizable groups with well-separated NMR signals, one can potentially extract both pK_a from the shifts. Similarly, for binding, one can follow the shift of a ligand proton or a protein backbone amide (in ^{1}H-^{15}N HSQC spectra, for example) as ligand is added and fit the binding isotherm to get K_d. NMR titrations for binding are commonly used in drug discovery to validate hits and measure their affinity to a protein.[69, 90]

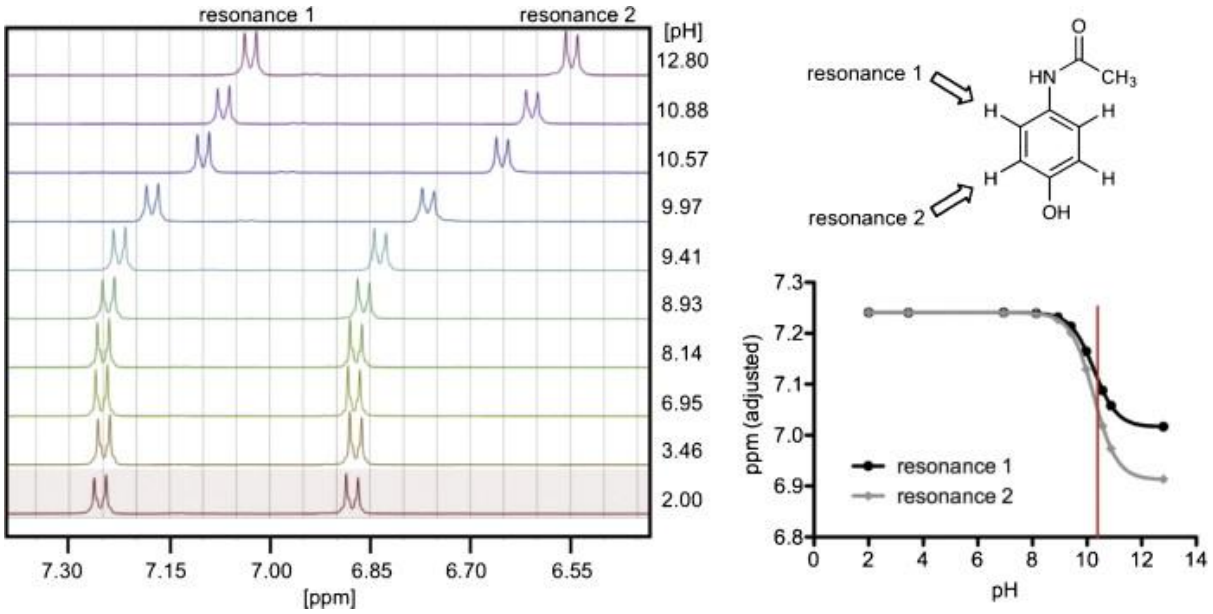


Figure 1.15. (left) Stacked ^{1}H NMR spectra collected across a pH range, showing the chemical shift behaviour of two distinct aromatic proton resonances (resonance 1 and resonance 2) in a compound. (right) plots the chemical shift (ppm) versus pH for both resonances, revealing a sigmoidal transition. The inflection point of these curves corresponds to the pK_a of the titratable group [89]

Compared with CSI, a conventional NMR titration is serial (one condition at a time) whereas CSI is parallel (all conditions at once in one experiment). CSI thus can be more efficient in terms of instrument time and avoids issues of sample-to-sample variability (the gradient ensures all conditions are part of one continuous sample). On the other hand, a classic NMR titration does not require setting up a physical gradient in the sample; it may be easier to implement with standard equipment (just prepare multiple samples or adjust pH stepwise). In summary, other NMR-based approaches share with CSI the advantage of molecular specificity and the ability to work in complex mixtures, but they can be sample- and time-intensive. CSI's innovation is to streamline this by encoding an entire titration in one go.

To conclude the introduction, chemical Shift Imaging (CSI) stands at the intersection of spatial resolution and molecular specificity, offering chemists a powerful, non-invasive means to study chemical distributions and equilibria in situ. Originally adapted from the foundational developments in NMR imaging and pulsed field gradients, CSI has evolved into a laboratory technique capable of mapping chemical shift information across space, enabling innovative applications such as single shot pK_a determination and metal ion binding analysis. Compared to traditional NMR methods and non-NMR imaging techniques, CSI offers unique advantages in analysing complex, heterogeneous, and opaque samples without requiring labels or destructive sampling. However, CSI also presents significant technical challenges, including reduced sensitivity due to voxel partitioning, susceptibility to diffusion and convection artifacts, and difficulties in quantitative analysis and spectral reconstruction. As the field advances with innovations in gradient hardware, sequence design, hyperpolarization, and computational acceleration, CSI continues to mature into a more practical and informative analytical tool. Its integration with complementary spectroscopic and imaging methods further broadens its utility.

1.5 Aims of the project

This chapter has provided a general overview of the chemical shift imaging experiment, situated in its context as a natural extension of the development of pulsed field gradients in NMR spectroscopy. The chapter also described the recent development of “titrated” CSI where a parameter of interest is varied across the NMR sample and hence allowing for the titration-based study of systems in a much more cost-effective and time- efficient manner. The chapter also described some implementations of the titrated CSI technique on determining pK_a of small molecules and in determination of protein-ligand dissociation constants and in assessing binding specificity, showing the potential benefits this technique has to offer.

The thesis sits inside the research program of developing new ways of using the titrated CSI NMR method to gain information on the binding and dissociation dynamics of molecules. In chapter 2 it will describe a new way of obtaining pK_a of ^{1}H NMR visible molecules using derivative based fitting, chapter 3 will then progress discussing how pK_a of ^{1}H NMR “invisible” molecules can be determined by measuring the quantity of protons transferred to pH chemical indicators via ^{1}H CSI, chapter 4 then will discuss obtaining the pK_a of water-insoluble compound by performing a “ solvent gradient “ ^{1}H CSI and finally in chapter 5 have a discussion of further extensions that CSI can have on calcium binding determination of polymers and small compounds.

Chapter 2: Derivative-based fitting as an alternative method to determine pK_a from ^{1}H NMR titration data

2.1 Chapter summary

The pK_a is a parameter that is related to stability and solubility of compounds. The use of NMR to determine pK_a has long been a potential method used to overcome some of the limitations of more traditional methods like potentiometric titrations. pK_a through NMR is typically determined by running multiple experiments with the analyte under different pH environments and correlating the measured chemical shift of the compound with pH and fitting pK_a through NMR modified Henderson-Hasselbalch equation. Recent developments with NMR chemical shift imaging allowed for the production of a pH gradient across the NMR tube and obtaining spatially selective 1D spectra each with a unique pH and determining pK_a through a "single shot" experiment.[57, 58] This chapter builds on the CSI method providing an alternative fitting method for obtaining pK_a that depends on measuring the derivative of pH with respect to the chemical shift of the analyte. This fitting method circumvents the need for fitting the limiting chemical shifts of the analyte and avoids issues of local maxima fitting convergence associated with the more traditional fixed point iteration method.

2.2 Introduction

The pH of a solution is a crucial factor that dictates the chemical and biophysical behaviour of the species it contains.[91, 92] For instance, even slight changes in pH can lead to significant alterations in a muscle's structure and functionality.[93, 94] Additionally, pH is a key consideration in the development of host–guest interactions and the formation of self-assembled materials, where it can influence the stability and dynamics of the systems involved.[95, 96] Understanding and controlling how the pH of a solution impacts the substances present in the solution is therefore essential in various fields, from biochemistry to materials science.[97, 98] For that goal determining the acid dissociation constant K_a (typically represented by its negative logarithm pK_a) which characterises the acidity of a compound can be useful in determining the behaviour of the compound as a function of pH. K_a characterises the acidity of compounds by representing the ratio of the compound in its dissociated form and undissociated form where $[H^+]$ is the concentration of hydrogen ions (protons), $[A^-]$ is the concentration of the conjugate base and [HA] is the concentration of the undissociated acid (see equation 2.1). pK_a is used extensively to understand and direct pH dependent substances and processes. For example it is used to prepare buffer solutions that maintain a specific pH for experiments, biochemical reactions, and industrial processes.[99] pK_a is also used in the pharmaceutical industry to determine the ionization state of a drug, thus its solubility in bodily fluids due to the fact that charged molecules are more soluble in water because water is a polar solvent and they interact strongly with water molecules through ion-dipole interactions.[100] Additionally, pK_a is used in predicting the pH dependent absorption of drugs as it describes the degree to which a drug is in its non nionised form within a certain pH and hence gives information to the degree of membrane permeability that the drug has.[101] Determination of pK_a is highly sought after as a result.[69]

$$K_a = \frac{[H^+][A^-]}{[HA]} \quad (2.1)$$

NMR spectroscopy can be a useful tool to measure the pK_a and pH dependent behaviour of compounds as it can be used to measure the chemical shift of the analyte of interest as a function of pH. Given that the chemical shift varies with pH, valuable information about the pK_a, intermolecular interactions and self-assembly of compounds can be obtained.[57, 102] Work by Wallace *et al.* displayed how pH gradients across a single sample can be produced and spatially resolved NMR spectra can be recorded along the NMR tube using chemical shift imaging techniques, circumventing the traditional methods, be they potentiometric or NMR, where multiple series of NMR experiments are run separately with the pH of each sample being adjusted between experiments thus significantly reducing the experimental time needed to perform a titration by NMR.[57]

In conventional NMR pK_a determination, experimental chemical shift versus pH data are typically fitted to the NMR modified Henderson–Hasselbalch equation using nonlinear regression (see appendix B1 for derivation of NMR modified Henderson–Hasselbalch equation). In practice, this fitting is commonly performed using a non-linear generalised reduced gradient (GRG) optimisation algorithm, such as that implemented in spreadsheet-based solvers (e.g. Microsoft Excel) where pK_a and the limiting chemical shifts δ_H and $\delta_{L\ are}$ treated as a free parameters (δ_H being the chemical shift the compound has when it is fully protonated and δ_L when it is fully deprotonated). The GRG algorithm looks at the gradient of the objective function as the input values are changed and determines whether the partial derivatives are equal to the values desired.[103] A general difficulty that occurs when using this method is that the value of the fit changes depending on the initial input values. This is because varying the initial input values for pK_a, δ_H, and δ_L alters the initial conditions of the optimisation, which can cause the GRG algorithm to converge to local extrema rather than the chemically meaningful global solution.[104, 105]

This causes variation in the fitting and makes it difficult in some instances to obtain objective pK_a values. To overcome this limitation, in this chapter a method is presented by which pK_a is determined by measuring the derivative of pH with respect to the chemical shift of the analyte and then fitting the data to a second order polynomial. Crucially this involves fitting a second order polynomial which has only one minima or maxima hence simplifying the fitting process whilst also increasing the objectivity of the method by removing errors accruing due to local maxima/minima distortions.

The method works as per previous method of Wallace *et al.* (see figure 2.1) by preparing a pH gradient across an NMR tube and measuring the pH of the solution across the NMR tube using indicators with known pK_a and limiting chemical shift values.[57] The manner in which pH is measured in the NMR tube has been developed and discussed previously.[106] As proof of concept, I select calibrate a number of NMR indicator compounds to accurately measure the pH of a solution over a wide range of pH values, demonstrate the accurate determination of the pK_a values of small organic molecules and validate the results by comparing them to literature values.

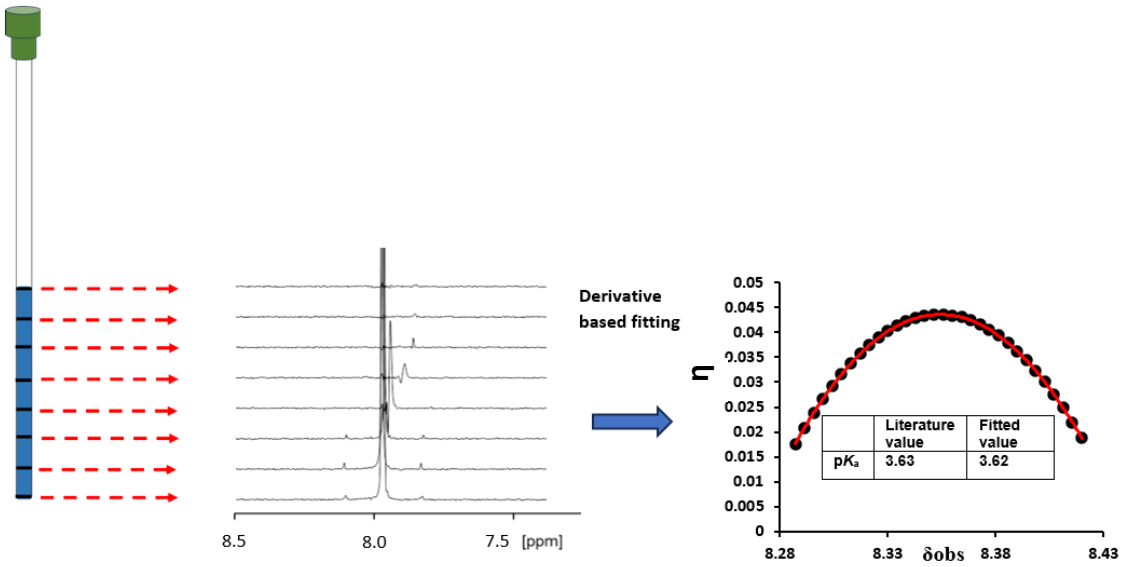


Figure 2.1. Scheme of how the derivative based pK_a determination works, CSI experiment is run with multiple 1d spectra obtained each having a unique pH due to the pH gradient formed by adding an acid or base source at the bottom of the NMR tube and derivative based data is obtained from the spectra and then fitted using a second order polynomial to obtain pK_a, δ_H and δ_L parameters

2.3 Experimental section

Materials. All chemicals were purchased from Fisher Scientific or Sigma-Aldrich and used as received. Stock H_2O (18.2 MΩ.cm) solution of the NMR indicators and analytes were prepared and used throughout the study. Compounds used were 2-methylimidazole (2-MI), sodium formate, sodium acetate,4-cyanophenol (4-CN), monohydrogen methylphosphonate ($MPAH^-$), benzoic acid, mandelic acid, propanoic acid, maleic acid and glycolic acid. Sodium formate, glycolic acid, sodium acetate, propanoic acid, 2-methylimidazole and 4-cyanophenol were used both as analytes and indicators in this study whilst mandelic acid, benzoic acid, monosodium maleate and monohydrogen methylphosphonate were used as analytes only. Stock solutions of the analytes in H_2O were prepared at 10 mM concentration of each. Formate, acetate and $MPAH^-$ were included with sodium as their cation. The stock solution contained 0.2 mM DSS which acted as chemical shift reference. All experiments were performed in 5 mm Wilmad 528-PP NMR tubes. To establish a pH gradient, 4-5 mg of solid indicator compound was weighed into the tube. Four, 2 mm diameter glass beads were placed on top of the base. An aliquot of the analyte solution was drawn up in a 9″ Pasteur pipet and gently layered on top of the glass beads to a height of 40-50 mm from the base of the NMR tube. The NMR tubes were stored in the sample changer rack at 21-22 °C until the time for running NMR experiments.

NMR. Experiments were performed on a Bruker Avance III 500 MHz standard bore spectrometer operating at 500.21 MHz for ^1H with a 5 mm broadband observe probe. The probe was equipped with z-axis pulsed field gradients. The temperature of the samples was maintained at 298±0.5 K. the 90° hard pulse was 10 µs with a respective of 22.677 Watt. The relaxation delay was 2 seconds. The spectral width was 16.02 ppm and the number of data points were 32k, giving an acquisition time of 2.045 seconds. CSI experiments were performed

using a gradient phase encoding sequence based on that of Trigo-Mouriño *et al.*[54] and incorporating double echo excitation sculpting for water suppression (Section A1). The encoding gradient pulse (172 μs, smoothed square) was varied between −18.8 and 18.8 G/cm in 64 steps. 4 ms Gaussian 180° pulses were employed for water suppression. Four scans were acquired at each step giving an acquisition time of 20 min and a spatial resolution of 0.41 mm. Sixteen dummy scans were acquired prior to signal acquisition. NMR data was processed in Bruker TopSpin 3.6.5. Each row of the CSI data set was automatically phase- and baseline-corrected using an automation script written in house. The ^{1}H chemical shifts of all the analytes and their respective indicators were extracted from each row using a custom script. Chemical shift data were exported to the spreadsheet accompanying this work where referencing and processing were done automatically. Scripts for the automated acquisition and processing of CSI data sets are provided in Section A3-A5.

2.4 Results and Discussion

Determining spatial location of 1D spectra. Determination of spatial location is done via running a biphasic sample comprising mineral oil and H_2O containing 0.5 M sodium acetate. Where mineral oil is situated on top of the H_2O solution. The boundary between the two phases is clear and allows for direct measurement via ruler (see figure 2.2). The sample boundary for the two experiments is 26 mm and 10 mm from base of NMR tube respectively. For both samples, a ^{1}H CSI experiment is run. The middle of the spectral width in the indirect dimension gives the centre of the NMR active region of the NMR tube thus the centre of the NMR active region as measured from the bottom of the NMR tube is:

$$Centre = Boundary\ distance - distance\ from\ centre \tag{2.2}$$

Figure 2.2. (left) NMR tube containing mineral oil and H_2O with boundary situated at 10 mm from base of NMR tube, (right) NMR tube containing mineral oil and H_2O with boundary situated at 26 mm from base of NMR tube

Two 2D ^{1}H CSI experiments were run for the two biphasic solutions for confirmation and their average was taken to be the centre of the NMR tube. The z-range (the total distance of NMR active region in the NMR tube in the z-direction) is equal to 26 mm and the number of points in the indirect dimension for these particular experiments was 32 hence the spatial resolution of the is spectrum is 0.81 mm. Hence, by determining the number of 1D spectra that separate the centre of the z-range from the boundary between the biphasic sample the distance of the centre of NMR active window from the bottom of the NMR tube is determined.

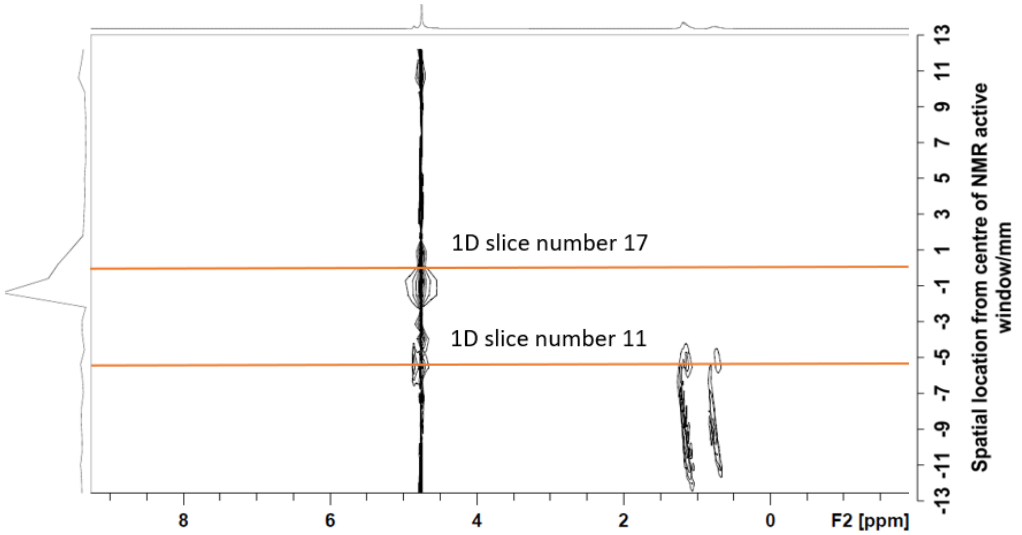


Figure 2.3. 2D CSI ^{1}H of biphasic sample with boundary situated at 26 mm from the bottom of the NMR tube with the indirect dimension corresponding to spatial location across the NMR tube (bottom of the figure corresponds to top of the NMR tube), slice 17 represent the middle of the NMR active window and slice 11 represents the boundary between the mineral oil phase and the H_2O phase and is situated vertically **above** the centre of the NMR active window

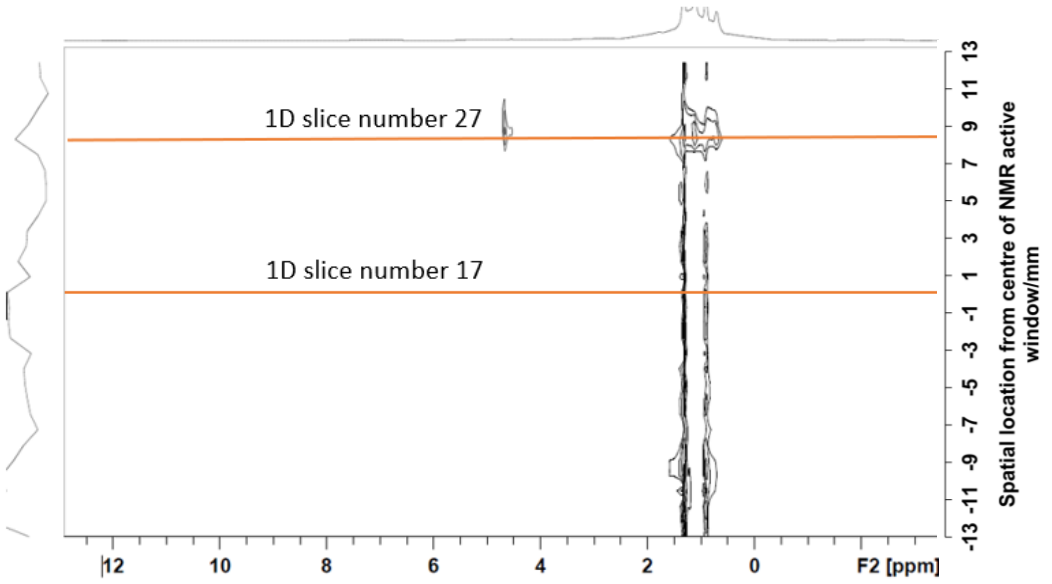


Figure 2.4. 2D CSI ^{1}H of biphasic sample with boundary situated at 10 mm from the bottom of the NMR tube with the indirect dimension corresponding to spatial location across the NMR

tube (bottom of the figure corresponds to top of the NMR tube) , slice 17 represent the middle of the NMR active window and slice 27 represents the boundary between the mineral oil phase and the H_2O phase and is situated vertically **below** the centre of the NMR active window

Once the centre the of NMR tube is obtained the method by which to relate the experiment number of the 1D slices to the spatial location from the bottom of the NMR tube is through equation 2.3:

$$Spatial\ location = Centre\ of\ NMR\ tube + (slice\ number - \frac{Number\ of\ slices}{2} - 0.5) \times \frac{Zrange}{Number\ of\ slices} \quad (2.3)$$

This means to say that spatial location of an NMR experiment is equivalent to the location of the centre of the NMR active regions ± the number of slices away from the centre (and their associated height). The centre of the NMR active region was thus determined to be 20 mm from the bottom of the NMR tube.

Determination of optimum time to run experiment. To calculate the time at which the optimum pH gradient develops the following calculations are made. The bases used in this work are highly water-soluble and dissolve within minutes of placing the analyte solution on top. Mathematically, the base behaves as though it were diffusing from a plane source at the bottom of the NMR tube. The concentration, C_Z, of diffusing base at a height Z from the bottom of the NMR tube after a time, t, is therefore given by equation 2.4:

$$C_z = N_{(t)}\, e^{(\frac{-(Z-h)^2}{4Dt})} \quad (2.4)$$

Where D is the diffusion coefficient of the base and N(t) is a time dependent parameter. h is the thickness of the solid base when placed at the base of the NMR tube (2 mm).

To obtain the total mass m of base that has diffused into the liquid, equation 2.4 is integrated over the volume of the sample. Since the NMR tube is approximately cylindrical with radius r, the differential volume element is $dV = \pi r^2 dZ$. Carrying out this integration yields equation 2.5:

$$\mathrm{m} = \pi \mathrm{r}^2 \mathrm{a} \mathrm{C}_0 \mathrm{M}_\mathrm{r} \sqrt{\pi \mathrm{D} \mathrm{t}_\mathrm{op}}\, \mathrm{e}^{(\frac{(\mathrm{Z_a}-\mathrm{h})^2}{4\mathrm{D}\mathrm{t}_\mathrm{op}})} \quad (2.5)$$

Where r is the inner radius of the NMR tube and Mr is the molecular weight of the base, a is the number of equivalents of diffusing basic indicator (pre-defined as 1 mM for this work) at height Z_a (18 mm) and C_0 is the concentration of acidic species. For a full derivation of this integration method can be found in standard references.[107] To determine t_{op}, for an experiment, t_{op} was adjusted iteratively until m matched the mass of indicator weighed into the NMR tube (4-5 mg).

Additionally, the diffusion coefficient of the indicators used needs to be known in order to reliably predict the formation of the pH gradient and to know the optimum time to perform the experiment. Diffusion coefficients of the indicators used are displayed in table 2.1 in order to calculate t_{op}. Diffusion coefficients were approximated using SEGWE diffusion coefficient calculator.[108] All diffusion coefficients were calculated at 25°C. The solvent assumed in the calculator is D_2O and whilst the diffusion coefficient of a compound in H_2O is different compared to D_2O for the purposes of the chapter the differences are negligible.[109]

Table 2.1. Diffusion coefficient and molecular weight values of indicators used

Indicator	Molecular weight/(g/mol)	Diffusion coefficient/ (m^2s^{-1})
2,6-lutidine	107.15	$7.12*10^{-10}$
Sodium formate	68.10	$8.87*10^{-10}$
2-methyl imidazole	82.10	$8.09*10^{-10}$
Glycolic acid	76.05	$8.40*10^{-10}$
Sodium acetate	82.03	$8.09*10^{-10}$
4-cyanophenol	119.12	$6.78*10^{-10}$

Analytes were run at different time points from t_{op} to showcase robustness of method in regards to obtaining pK_a values of analytes within a time window of ± 20% of t_{op} as in previous methods by Wallace *et al*.[57] Figures 2.5-2.7 showcase pH profiles of multiple compounds along with the pK_a values obtained for each pH gradient (pK_a values were obtained by fitting the chemical shift of the indicator as function of pH to the NMR modified Henderson–Hasselbalch equation (equation 2.4)). Results show that the method still obtains pK_a values close to literature within the previous 20% limits with accuracy decreasing outside these time windows. Overall, this section shows how pH gradients are robust to choice of diffusion time.

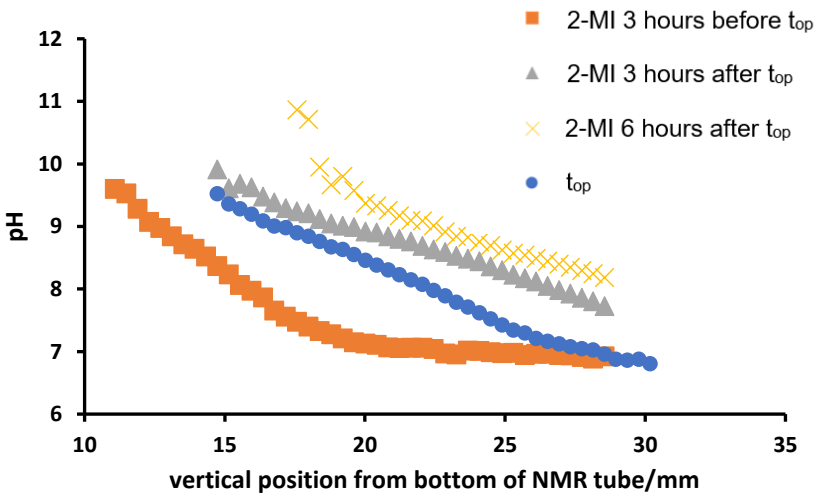


Figure 2.5. Plot of pH gradients for 2-methylimidazole (2-MI). Different colours correspond to different experiments done on the same sample (pH plotted beyond indicator limits is not the actual pH, but the lower or higher limit the indicator can measure). Experiments were done by inserting 4-5 mg of 4-cyanophenol powder on the bottom of the NMR tube then adding four, 2 mm diameter glass beads on top of the powder then adding the analyte solution and leaving the gradient to develop for 2.5, 5.5 (t_{op}), 8.5 hours and 11.5 hours from sample preparation

Table 2.2. Compilation of fitted pK_a values of 2-MI obtained with different diffusion times

Time of experiment	pK_a measurement
3 hours before t_{op}	8.0
t_{op}	8.14
3 hours after t_{op}	7.83
6 hours after t_{op}	8.89*

* Outside the limits of acceptable time of experiment hence the decrease in accuracy of measurement

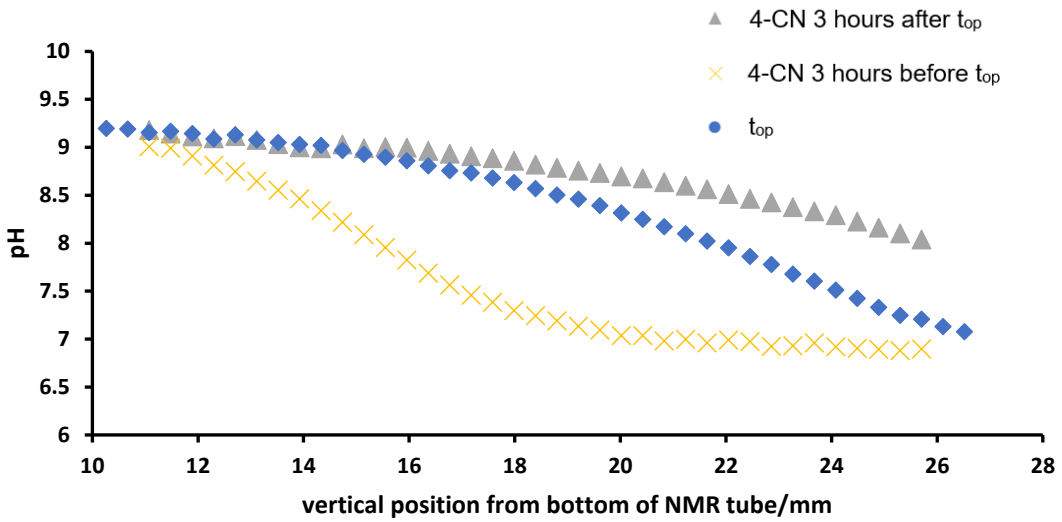


Figure 2.6. Plot of pH gradients for 4-cyanophenol (4-CN). Different colours correspond to different experiments done on the same sample (pH plotted beyond indicator limits is not the actual pH, but the lower or higher limit the indicator can measure). Experiments were done by inserting 4-5 mg of 2-methylimidazole powder on the bottom of the NMR tube then adding four, 2 mm diameter glass beads on top of the powder then adding the analyte solution and leaving the gradient to develop for 5.3, 8.3 (t_{op}) and 11.3 hours from sample preparation

Table 2.3. Compilation of fitted pK_a values of 4-CN obtained with different diffusion times

Time of experiment	pK_a measurement
3 hours before t_{op}	7.94
t_{op}	8.0
3 hours after t_{op}	8.18
6 hours after t_{op}	7.49*

* Outside the limits of acceptable time of experiment hence the decrease in accuracy of measurement

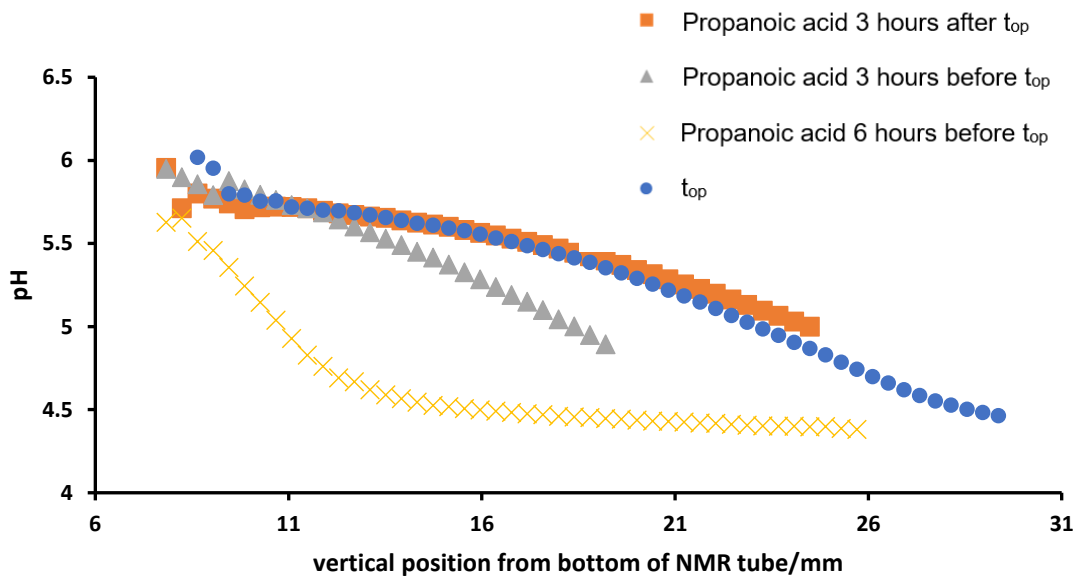


Figure 2.7. Plot of pH gradients for the analyte propanoic acid. Different colours correspond to different experiments done on the same sample (pH plotted beyond indicator limits is not the actual pH, but the lower or higher limit the indicator can measure). Experiments were done by inserting 4-5 mg of sodium acetate trihydrate powder on the bottom of the NMR tube then adding four, 2 mm diameter glass beads on top of the powder then adding the analyte solution and leaving the gradient to develop for 2.1, 5.1, 8.1 (t_{op}) and 11.1 hours from sample preparation

Table 2.4. Compilation of fitted pK_a values of propanoic acid obtained with different diffusion times

Time of experiment	pK_a measurement
6 hours before t_{op}	4.77
3 hours before t_{op}	5.04
t_{op}	4.81
3 hours after t_{op}	5.05

Determination of the limiting chemical shifts of indicator compounds to enable determination of pH *via* equation 2.4. NMR pH measurement centres on the fact that the chemical shift of any compound under conditions of fast exchange is an average of chemical shifts of the populations of said compound in its protonated states and deprotonated states.[110] Given that the fractional populations of the compound in its protonated and deprotonated states is intimately linked to the pH of the solution allows for the derivation of a an equation relating the chemical shift of a compound to the pH of the solution.[111] This relation is represented in the NMR modified Henderson-Hasselbalch equation:

$$pH = pK_a + \log_{10}\left(\frac{\delta_{obs} - \delta_H}{\delta_L - \delta_{obs}}\right) \quad (2.6)$$

Where δ_H and δ_L are the limiting chemical shifts of pH indicator when it is fully protonated and fully deprotonated respectively. If the limiting chemical shifts of the indicators are not known then they are to be determined experimentally by preparing a solution of the indicator (along with 0.2 mM DSS to act as a chemical shift reference compound) and adding aliquots of either HCl or NaOH to the solution and measure the pH using a pH probe.[57] For every addition of aliquot, a portion of the solution is taken for 1D ^{1}H NMR experiment where the referenced chemical shift of the indicator is measured. This allows for the production of chemical shift/pH titration curve of the indicator compound where the NMR modified Henderson-Hasselbalch equation (equation 2.6) is fitted using generalised reduced gradient solving method via Excel with δ_H and δ_L as free parameters (See figure 2.8).

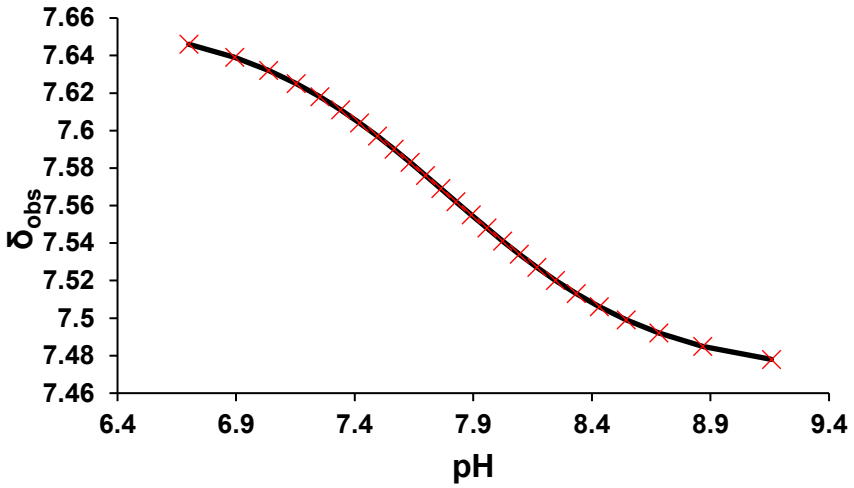


Figure 2.8. δ_{obs} versus pH graph of 4-cyanophenol the red crosses are the measured pH and δ_{0bs} values using the pH probe and NMR respectively and the black line is the chemical shift predicted using the indicator parameters obtained by fitting δ_L, δ_H and pK_a as free parameters in the NMR-modified Henderson-Hasselbalch equation

Determination of pK_a of analyte via derivative-based fitting. The pH of a solution is determined from the chemical shift of an indicator molecule, δ_{obs} by equation 2.7:

$$pH = pK_a + \log_{10}\left(\frac{\delta_{obs} - \delta_H}{\delta_L - \delta_{obs}}\right) \tag{2.7}$$

The derivative of pH with respect to δ_{obs} can be represented algebraically via equation 2.8.

$$\frac{dpH}{d\delta_{obs}} = \frac{d\left[pK_a + \log_{10}\left(\frac{\delta_{obs} - \delta_H}{\delta_L - \delta_{obs}}\right)\right]}{d\delta_{obs}} \tag{2.8}$$

Since pK_a is a constant equation 2.8 becomes:

$$\frac{dpH}{d\delta_{obs}} = \frac{d\left[\log_{10}\left(\frac{\delta_{obs} - \delta_H}{\delta_L - \delta_{obs}}\right)\right]}{d\delta_{obs}} \tag{2.9}$$

Applying the chain to equation 2.9 and using the log identity equation 2.10 is obtained:

$$\frac{dpH}{d\delta_{obs}} = \frac{1}{\ln(10)} \times \frac{\delta_L - \delta_{obs}}{\delta_{obs} - \delta_H} \times \frac{d\left[\frac{\delta_{obs} - \delta_H}{\delta_L - \delta_{obs}}\right]}{d\delta_{obs}} \tag{2.10}$$

Applying the quotient rule 2.10 becomes:

$$\frac{dpH}{d\delta_{obs}} = \frac{1}{\ln(10)} \times \frac{\delta_L - \delta_{obs}}{\delta_{obs} - \delta_H} \times \frac{\delta_L - \delta_H}{(\delta_L - \delta_{obs})^2} \tag{2.11}$$

Cancelling one factor of $\delta_L - \delta_{obs}$ and some rearrangement provides equation 2.12:

$$\frac{dpH}{d\delta_{obs}} = \frac{\delta_H - \delta_L}{\ln(10)(\delta_{obs} - \delta_H)(\delta_{obs} - \delta_L)} \tag{2.12}$$

The inverse of the derivative multiplied by the natural log of 10 is then represented as a function of the chemical shift of the analyte in equation 2.13:

$$\left(\ln(10)\,\frac{dpH}{d\delta_{obs}}\right)^{-1} = \frac{{\delta_{obs}}^2}{\delta_H - \delta_L} + \delta_{obs} + \frac{\delta_H \delta_L}{\delta_H - \delta_L} \qquad (2.13)$$

The formula $(ln(10)\frac{dpH}{d\delta_{obs}})^{-1}$ will be represented as η for reason of brevity. Thus equation 2.13 becomes:

$$\eta = \frac{{\delta_{obs}}^2}{\delta_H - \delta_L} + \delta_{obs} + \frac{\delta_H \delta_L}{\delta_H - \delta_L} \qquad (2.14)$$

The derivative is computed numerically using the central difference method. In this method, a given function f(x) and its derivative f' are represented by equation 2.15:

$$f' \approx \frac{f(x+h) - f(x-h)}{2h} \qquad (2.15)$$

Where h is a small step size. This approach uses function values on both sides of the point x, making it less sensitive to one-sided variations in the function and provides a second-order accurate approximation compared to the first-order accuracy of forward and backward differences.[112, 113] The derivative values are then inputted into equation 2.13 which is then fitted using a second order polynomial fitting by least squares regression in microsoft excel (see figure 2.9).

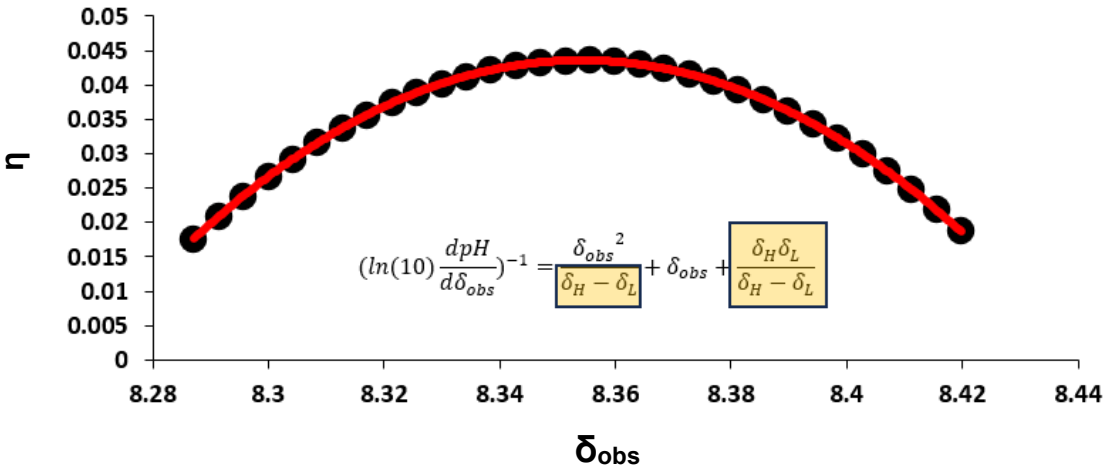


Figure 2.9. Plot of η versus δ_{obs} of formic acid containing simulated data (black circles) with derivative-based fitting (red line) to equation 2.13. Shown for illustrative purposes only; no experimental data are presented).

This equation is now a second order polynomial with δ_{obs} as the independent variable. Measuring the chemical shift of the compound of interest and the pH of the solution in a 2D CSI ^{1}H NMR experiment allows for polynomial fitting of the equation using least squares regression method to obtain two parameters that contain limiting chemical shift information see (see figure highlighted yellow for the two parameters and fitting process for simulated data of formic acid).

The two parameters can be solved simultaneously to obtain δ_L and δ_H :

Let $\frac{1}{\delta_H - \delta_L} = A$ and $\frac{\delta_H \delta_L}{\delta_H - \delta_L} = B$ it follows that $\delta_H - \delta_L = A^{-1}$ and that $A^{-1} \times B = \delta_H \delta_L$

Further analysis shows that $\delta_H = A^{-1} + \delta_L$ this can be substituted into the previous equation to get:

$$A^{-1} \times B = (A^{-1} + \delta_L)\delta_L = A^{-1}\delta_L + {\delta_L}^2 \quad (2.16)$$

If $A^{-1} \times B$ is then treated as the dependent variable and δ_L as the independent variable then δ_L can be obtained via the quadratic formula:

$$\delta_L = \frac{-A^{-1} \pm \sqrt{A^{-2}}}{2} \quad (2.17)$$

δ_H can then be obtained by solving equation 2.18:

$$A^{-1} + \delta_L = \delta_H \quad (2.18)$$

Finally in order to get the pK_a of the compound the newly acquired limiting chemical shifts are incorporated into the NMR modified Henderson-Hasselbalch equation and the equation is solved for pK_a

$$pH - \log_{10}(\frac{\delta_{obs} - \delta_H}{\delta_L - \delta_{obs}}) = pK_a \quad (2.19)$$

Therefore, by measuring the chemical shift of the compound of interest and knowing the pH of the solution using an indicator compound the pK_a of the compound can be directly calculated. This process is done for every slice of the CSI experiment and pK_a values are then averaged to obtain the final pK_a measurement. Additionally, the mean of the limiting chemical shifts of a particular compound corresponds to the chemical shift the compound will have when the pH of the solution is equal to its pK_a (this is because the average of the limiting chemical shift is the chemical shift at which the compound is 50% protonated and deprotonated).

One other important point to consider is that the condition in which the pH of the solution would equal pK_a is also the condition in which maximum buffering capacity is present in the system because the system would resist changes in pH most strongly when both acid and base forms are present in roughly equal amounts. This means that the derivative of pH with respect to δ_{obs} would be at its minimum at that pH. This then allows for clear graphical illustration when pH of the solution is equal to pK_a of the compound by observing for when the parabolic curve reaches its extremum and this allows for a much clearer view of when the pH range reached in the experiment is adequate to perform reliable fitting as opposed to the sigmoidal shape (see figure 2.10).

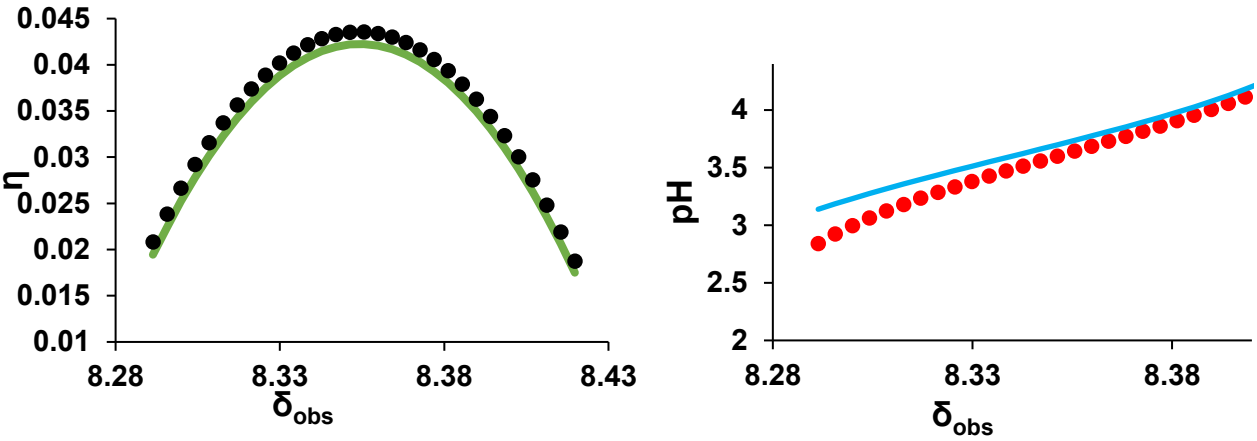


Figure 2.10. (left) Comparison of formate *in silico* η data (black) with respective derivative based fitting (green curve), (right) formate standard Henderson-Hasselbalch *in silico* data (red) with its respective fitting to equation 2.7 (blue)

For comparison fitting with respect to the derivative method and standard Henderson-Hasselbalch equation is shown table 2.5. For effective pK_a determination a well-developed pH gradient that changes smoothly is desired and is best obtained when running it at t_{op} however acceptable pH gradients are attained ± 2 hours on either side of t_{op}.[57] Smooth pH gradients are also necessary to reduce error in the approximation of derivative values. A complete set of the parameters obtained for all analytes is provided in table 2.6.

Table 2.5. Comparison of derivative based and Henderson-Hasselbalch based fitting of δ_L, δ_H, and pK_a of formic acid using *In silico* data of Figure 2.10

	Derivative based	Henderson-Hasselbalch based	Actual value
pK_a	3.65	3.60	3.63
δ_L	8.43	8.44	8.44
δ_H	8.27	8.24	8.26

Table 2.6. Fitted δ_L, δ_H, and pK_a of compounds studied to experimental data with their respective literature values using the derivative-based fitting method

Analyte	Fitted pK_a	Literature pK_a	Fitted δ_H/ppm	Measured δ_H/ppm	Fitted δ_L/ppm	Measured δ_L/ppm
4-cyanophenol	8.06	7.95	7.662	7.663	7.462	7.470
2-methylimidazole	8.04	7.96	7.273	7.270	6.956	6.958
$MPAH^-$	7.89	7.75	1.233	1.281	1.099	1.071
Monosodium maleate	6.11	6.20	6.324	6.337	5.996	5.989
Glycolic acid	3.92	3.89	4.146	4.213	3.941	3.932
Formic acid	3.68	3.75	8.221	8.267	8.447	8.441
Acetic acid	4.91	4.76	2.083	2.083	1.863	1.906

Mandelic acid	3.72	3.41	5.164	5.168	4.984	4.989
Benzoic acid	4.39	4.20	8.043	8.047	7.859	7.866
Propanoic acid	5.03	4.87	2.392	2.380	2.167	2.170

The method was tested on a variety of analytes and indicators with varying pK_a values between 2 and 10. The thermodynamic pK_a values of the analytes obtained agree with literature values to within ±0.3 units which is comparable to what is obtained via other techniques such electrophoresis, potentiometric titrations, conductometry, UV spectroscopy and more conventional NMR titrations.[73, 89, 114-119].

Determining pK_a using multiple indicators. In situations where there is a need for a large pH range (due to having multiple analytes of varying pK_a values or because of high uncertainty in regards to the pK_a of the compound) multiple pH indicators can be inserted into the solution to be able to measure pH values accurately over a wide pH range. A homogeneous solution of the indicators was prepared and 2-3 mg of oxalic acid dihydrate crystals were then added at the bottom of the NMR tube. Oxalic acid is then diffused along the NMR tube. Due to the fact that the pH would span over a large range, care as to how to use the indicators to measure pH as the accuracy of the measurement decreases as the pH of the solution is further away from the pK_a of the indicator. As a result, pH is measured as sensitivity-weighted average of the pH measurement of all the indicators:[111, 120]

$$pH = \frac{\sum_{i=1}^{n} S_i pH_i}{\sum_{i=1}^{n} S_i} \tag{2.20}$$

Where S_i is sensitivity of the indicator, pH_i is pH as it is measured from that indicator and n is the number of indicators used. The sensitivity can be defined as the first derivative of the indicator chemical shift with respect to pH and is represented by equation 2.21:[111, 120]

$$S_i = \frac{d\delta_{obs}}{dpH} = \frac{(\delta_l - \delta_{obs})(\delta_{obs} - \delta_h)}{\delta_h - \delta_l} \tag{2.21}$$

S_i was calculated for each indicator, and the apparent pH of the solution was calculated for each indicator. The pH of the solution could then be obtained as the average of the apparent pH values reported by each indicator in the pair, weighted by S_i.

To highlight the effectiveness of using multiple pH indicators experiment is run with multiple monoprotic analytes and the pK_a of all these compounds are measured using the derivatives method. Showcasing the robustness of the method in being able to increase the efficiency of the pK_a determination and decrease time costs.2D ^{1}H NMR CSI experiments were run with an H_2O solution containing 15 mM of 2-MI, 2,6-lutidine and sodium formate as pH indicators along with 1 mM of 4-CN, monosodium maleate and glycolic acid as analytes with 3 mg of Oxalic acid dihydrate inserted at the bottom of the NMR tube to vary the pH (difference in concentration between analytes and indicators is in order to clearly distinguish an indicator from an analyte on the NMR spectra).

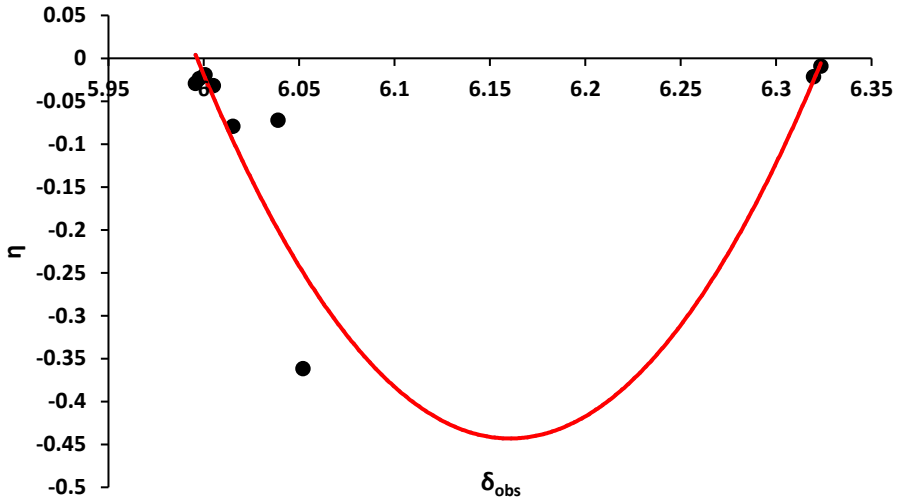


Figure 2.11. Plot of η versus δ_{obs} of Monosodium maleate with its respective fitting

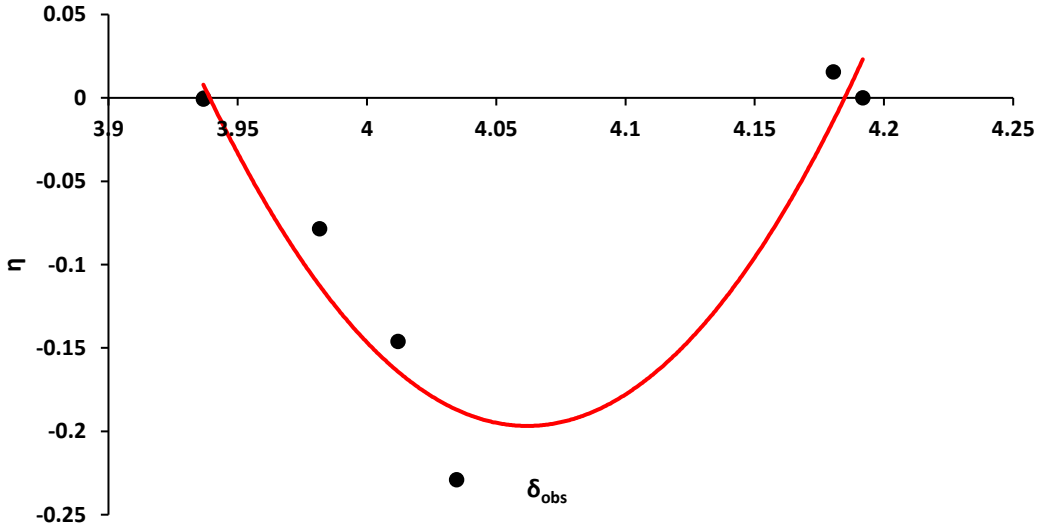


Figure 2.12. Plot of η versus δ_{obs} of glycolic acid with its respective fitting

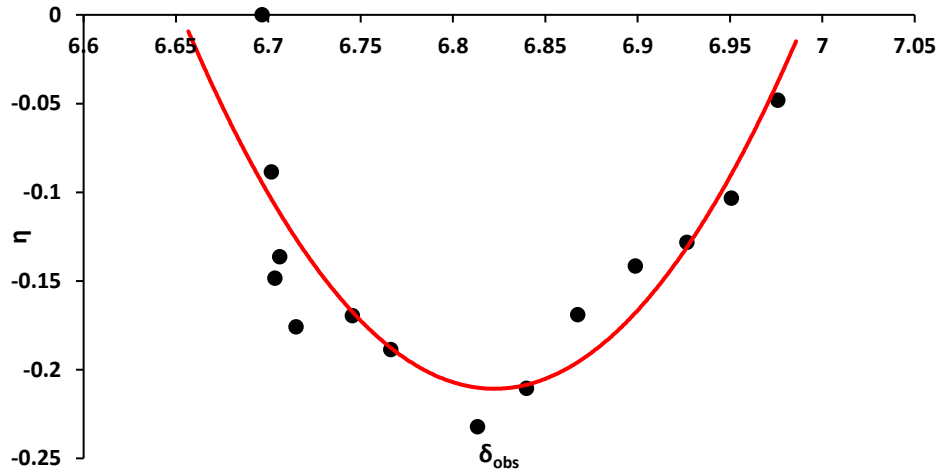


Figure 2.13. Plot of η versus δ_{obs} of 4-cyanophenol with its respective fitting

In this experiment the rapid change in pH (see figure 2.14) caused for relatively few data points to do derivative-based fitting. Additionally, because of relatively slow exchange the chemical shift of all compounds was significantly broad and due to the presence of multiple peaks there was significant spectral overlap. Nevertheless, accurate pK_a, δ_H and δ_L values were obtained for each of the analytes (see table 2.7) showing the capacity of this method to work in the presence of significant line broadening and spectral overlap. This method provides for a larger pH range as compared to when one pH indicator is used (see figure 2.15)

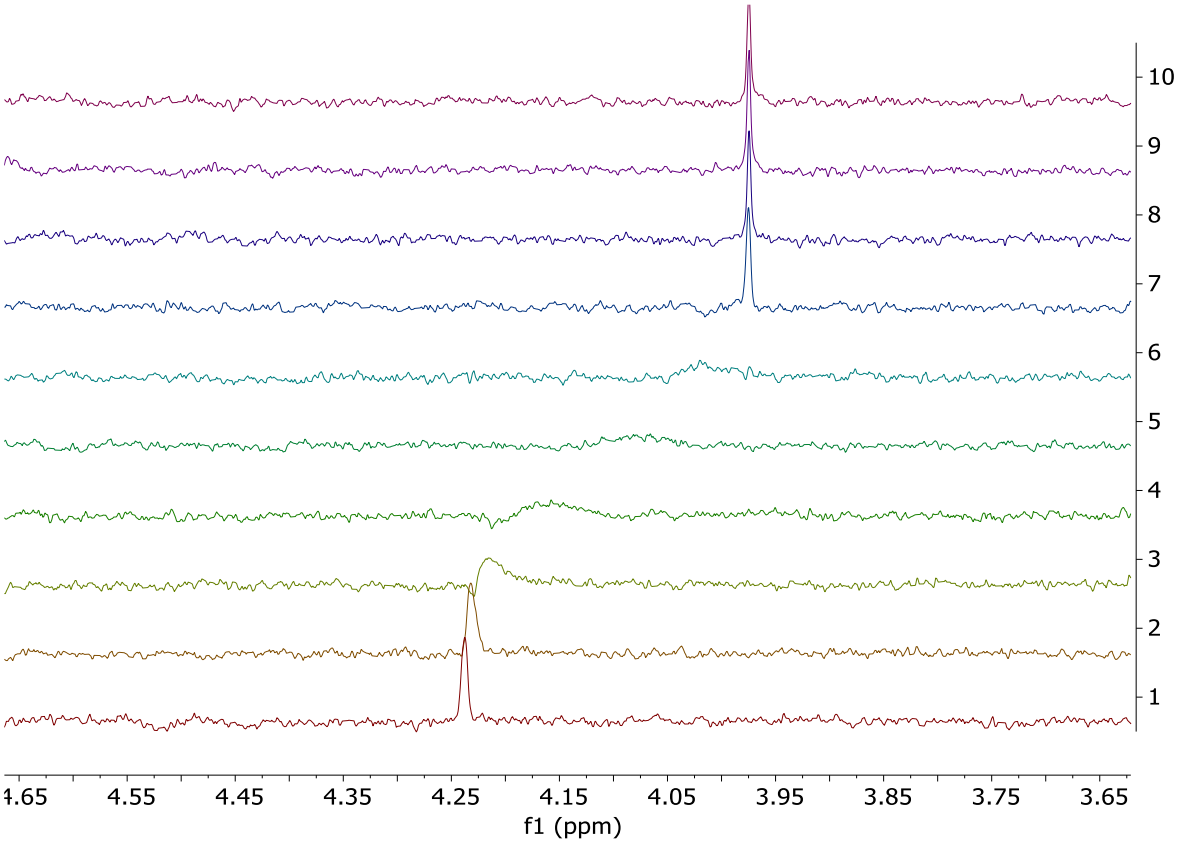


Figure 2.14. stacked spectra of glycolic acid showing rapid transition of its respective chemical shift

Table 2.7. Fitted δ_H, δ_L, and pK_a of analytes studied using multiple pH indicators

Analyte	Fitted pK_a	Literature pK_a	Fitted δ_H	Measured δ_H	Fitted δ_L	Measured δ_L
4-cyanophenol	7.88	7.95	6.991	6.995	6.653	6.642
Monosodium maleate	6.11	6.20	6.324	6.337	5.996	5.989
Glycolic acid	3.95	3.89	4.184	4.213	3.939	3.932

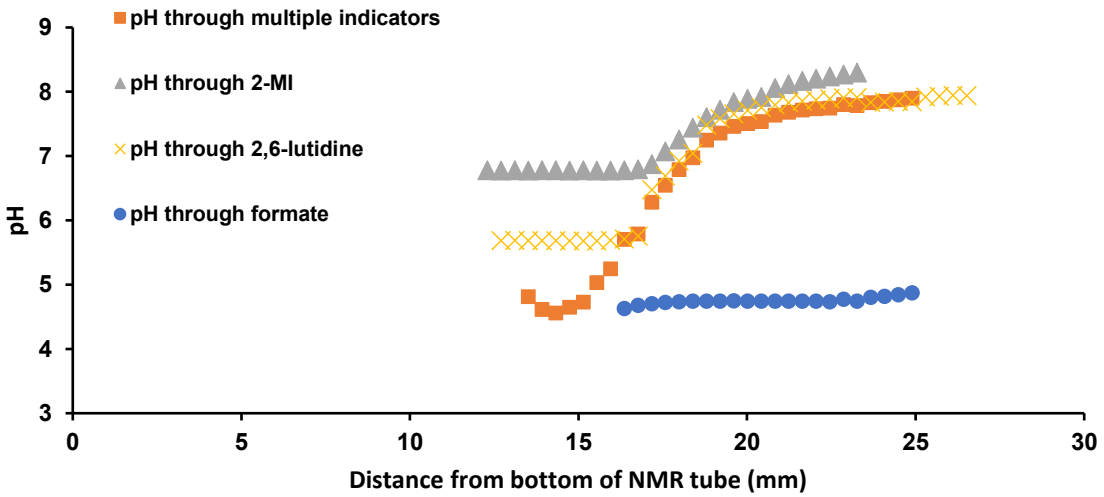


Figure 2.15. Comparison of pH ranges reported by indicators when indicators are used individually or collectively

2.5 Conclusions

This chapter has shown an alternative method of fitting the pK_a values of compounds by utilizing the methods developed by Wallace *et al.*[57] of preparing pH gradients inside an NMR tube and running chemical shift imaging experiments to obtain spatially selective spectra of the analyte across varying pH values and then using those spectra to obtain information on the derivative of pH with respect to δ_{obs} of the analyte. By second order polynomial fitting using least squares regression and simple algebraic manipulation the pK_a value of the analyte is obtained without the need of GRG solving method or estimating the limiting chemical shift of the compound and without the risk of running into local extrema whilst also improving the objectivity of the method. As proof of concept this method was tested on a variety of analytes of varying pK_a values and accurate pK_a determined is obtained for all of them within ± 0.3 units of literature values. This method could be of future use in drug screening as it greatly accelerates the pK_a measurement by avoiding cumbersome titrations and robustly fits the data obtained without worry about local extrema error. It also provides a useful graphical illustration of the pH range obtained in the gradient and provides a simple way of visualizing the adequacy of the pH gradient and whether the experimental data is appropriate for fitting or otherwise.

Chapter 3: Determining the pK_a and concentration of NMR invisible molecules using NMR spectroscopy

3.1 Chapter summary

NMR spectroscopy is a very powerful tool for measuring the dissociation constants (pK_a) of molecules, requiring smaller quantities of sample of lower purity relative to potentiometric or conductometric methods. However, current approaches are limited to those molecules possessing favorable pH-dependent NMR properties. Typically, a series of 1D experiments at varying pH are performed and the pK_a is obtained by fitting the observed chemical shift of the analyte as a function of pH using non-linear routines. However, the majority of polymers, biomolecules and inorganic species do not present favorable NMR resonances. Either the resonances are not observable, too broad, or the unambiguous interpretation of the NMR data is impossible without resorting to complex 2D experiments due to spectral overlaps. To overcome these fundamental limitations, this chapter presents a method to obtain the pK_a values and concentrations of acidic species without their direct observation by NMR. The new method instead determines the quantity of acidic protons removed from the species along a concentration gradient of an organic base in a single ^{1}H chemical shift imaging experiment that can be run under automation. The pK_a values are determined *via* simple linear plots, avoiding complex and potentially unreliable non-linear fitting routines.

3.2 Author contributions

Chapter 3 comprises a paper, “Determination of the pK_a and Concentration of NMR-Invisible Molecules and Sites Using NMR Spectroscopy”, which was published in *Analytical chemistry* as a technical note in 2024.[121] Haider Hussain’s contribution to the paper are as follows: Haider Hussain conceived the concept of the paper, designed the experiments, prepared all the samples and performed all of the NMR experiments. Matthew Wallace provided helpful guidance throughout the project. Haider Hussain wrote the manuscript, aided by helpful suggestions from Matthew Wallace and Yaroslav Khimyak.

3.3 Introduction

pK_a is a fundamental property which can be used to predict whether a molecule or ion will be protonated or deprotonated under different conditions.[122] As a result, it attracts immense interest in the areas of food science, pharmaceutical chemistry, and organic synthesis among other areas.[123] In drug discovery, pK_a values are used to predict drug-target interactions and drug solubility.[124] In materials science, the pK_a of polymers determines their self-assembly and complexation properties. For example, an understanding of the pK_a value of nucleic acid-polymer complexes is needed to prepare stable nucleic acids for therapeutic use.[125] Additionally, inorganic ions such as phosphorus or nitrogen species play an important role in biological systems while proteins exhibit a pH-dependent charge which determines their solubility, stability and separation properties.[126]

NMR spectroscopy is a valuable technique to measure pK_a as it offers the advantages of studying analytes using small volumes with equipment that is available in most research institutions.[89] Furthermore, we demonstrated how pK_a can be determined using a combination of pH gradients and 2D chemical shift imaging (CSI) ^{1}H NMR. This allows for a "single shot" determination of pK_a in one NMR experiment, saving material and time.[57, 120, 127] However, a restriction on all current NMR methods is that they require the analyte to display chemical shifts which can be clearly observed to change in a predictable manner as a function of pH. However, not all systems of interest can be characterised in this manner. A striking case are polymeric systems with molecular weights above 20,000 g/mol that typically display very broad resonances.[128] Additionally even for systems which have a lower molecular weight, given enough complexity, the spectral overlap will make the resolution of separate resonances impractical. The ChEMBL database displays 1,913,280 small molecules that are preclinical, of which approximately 20% have a molecular weight over 600 g/mol and are therefore likely to exhibit overlapping resonances on a standard ^{1}H NMR spectrum.[129] Moreover, there are 1305 small proteins according to the protein data bank that can be investigated using solution-state NMR but for which 2D NMR experiments are needed to assess their pH dependent behaviour.[130] Such "NMR restricted" molecules and ions are of great interest in a variety of fields and would otherwise benefit from pK_a determination by NMR.

Here, this chapter presents a more general NMR method that has the capacity to encompass these restricted molecules, expanding the range of molecules that can be probed by NMR hence providing avenues for exploring their pK_a associated activity that were not previously possible whilst also enabling the analysis of species whose toxicity, volatility or limited availability may preclude the use of conventional potentiometric titrations. The new method determines the pK_a of acidic NMR restricted molecules by measuring the quantity of protons transferred to a basic indicator along pH gradients in 5 mm NMR tubes. As proof of concept the measurement of the pK_a values of a range of small organic and inorganic molecules in excellent agreement with literature values. We further test this method on more complex systems by measuring the effective pK_a of polyacrylic acid and wheat germ agglutinin.

3.4 Experimental section

Materials. All chemicals were purchased from Fisher Scientific or Sigma-Aldrich and used as received. Stock H_2O (18.2 MΩ.cm) solution of the NMR indicators (Table 3.1) were prepared and used throughout the study. Indicators used were 2-methylimidazole (2-MI), 1,2,4-triazole, formate, acetate, monohydrogen methylphosphonate ($MPAH^-$) and 2,6-lutidine. Formate, acetate and $MPAH^-$ were included with sodium as their cation. The analytes studied were boric acid, 4-cyanophenol (4-CN), phosphoric acid (H_3PO_4), sodium dihydrogen phosphate (NaH_2PO_4), glycine, hydroxylammonium chloride (NH_3OHCl), ammonium chloride (NH_4Cl), benzoic acid and glycolic acid. Stock solutions of the analytes in H_2O were prepared at 10 mM concentration of each except phosphoric acid which was prepared at 60 mM to explore the scope of the methodology. Polyacrylic acid (PAA, Mw 240 kDa) and lectin from *Triticum vulgaris* (also known as wheat germ agglutinin (WGA)) were used as model analytes for measuring pK_a of polymers and proteins, respectively. The stock solution contained 0.2 mM DSS which acted as chemical shift reference and 0.01 vol % of either DMSO or 1,4-dioxane which acted as integral reference (choice was based on which reference compound had least spectral overlap with the analyte).

Limiting chemical shifts for the indicators (Table 3.1) were obtained by measuring the chemical shift of the indicator in a homogeneous solution containing HCl (0.01 M) with a concentration

gradient of the basic indicator. Concentration of protons transferred from HCl to indicator was assumed equal to the concentration of HCl as it is a strong acid. All experiments were performed in 5 mm Wilmad 528-PP NMR tubes. To establish a pH gradient, 4-5 mg of solid base was weighed into the tube. Four, 2 mm diameter glass beads were placed on top of the base when 2,6-lutidine was used as a diffusant, an aqueous solution (2 M, 30 µL) was pipetted on top of a 500 µL analyte solution where it floated due to its lower density. An aliquot of the analyte solution was drawn up in a 9″ Pasteur pipet and gently layered on top of the glass beads to a height of 40-50 mm from the base of the NMR tube. The NMR tubes were stored in the sample changer rack at 21-22 °C until the optimum time for running NMR experiments. Experiments were run within 20 minutes of the optimum time, t_{op}, at which the gradient was predicted to have developed.[57]

NMR. Experiments were performed on a Bruker Avance III 500 MHz spectrometer operating at 500.21 MHz for ^{1}H with a 5 mm broadband observe probe. The probe was equipped with z-axis pulsed field gradients. The temperature of the samples was maintained at 298±0.5 K. the 90° hard pulse was 10 µs with a respective of 22.677 Watt. The relaxation delay was 2 seconds. The spectral width was 16.02 ppm and the number of data points were 32k, giving an acquisition time of 2.045 seconds. CSI experiments were performed using a gradient phase encoding sequence based on that of Trigo-Mouriño *et al.*[54] and incorporating double echo excitation sculpting for water suppression (Section A1). The encoding gradient pulse (172 µs, smoothed square) was varied between −18.8 and 18.8 G/cm in 64 steps.4 ms Gaussian 180° pulses were employed for water suppression. Four scans were acquired at each step giving an acquisition time of 20 min and a spatial resolution of 0.41 mm. Sixteen dummy scans were acquired prior to signal acquisition. All spectra were referenced to DSS (0 ppm). NMR data was processed in Bruker TopSpin 3.6.5. Scripts for the automated acquisition and processing of NMR data are provided in Section A3-A5.

3.5 Results and Discussion

Derivation of κ parameter. κ is defined as the molar concentration of protons transferred from the acidic analyte to the basic indicator. Conversely it is defined as the molar concentration of protons the base obtains from the acid. Hence, κ can be described by equation 3.1 and 3.2:

$$\kappa = C_{indicator}\, f_{H\ indicator} \tag{3.1}$$

$$\kappa = C_{analyte}\, f_{L\ analyte} \tag{3.2}$$

Where $C_{indicator}$ is the total concentration of basic indicator, $f_{H\ indicator}$ is the fraction of it protonated, $C_{analyte}$ is the total concentration of acidic analyte and $f_{L\ analyte}$ is the fraction of it deprotonated. To obtain $f_{H\ indicator}$ we use equation 3.3:

$$\delta_{\mathrm{obs}} = \delta_{\mathrm{L}} \mathrm{f}_{\mathrm{L\ indicator}} + \delta_{\mathrm{H}} \mathrm{f}_{\mathrm{H\ indicator}} \tag{3.3}$$

Equation 3.3 describes how the chemical shift of observed molecule is the average of the chemical shifts of the protonated and deprotonated states. Equation 3.1 and 3.3 are combined to give equation 3.4:

$$\kappa = C_{indicator}\ \frac{\delta_{obs}-\delta_L}{\delta_H-\delta_L} \tag{3.4}$$

This allows for determination of κ given knowledge of δ_L and δ_H of indicator (all respective indicator values used in this work are found in table 1). However, when working with analytes with relatively low pK_a (such as phosphoric acid) requires the use of a relatively weak base (such as 1,2,4-triazole) and due to the weak basicity of the base the quantity of protons transferred from the acidic analyte to water becomes significant and cannot be ignored. Determining the protons transferred from the acidic analyte requires measuring the quantity of protons transferred to both the basic indicator and H_2O. Determining the quantity of protons transferred to H_2O is done simply by measuring the concentration of hydronium ions (which is just 10^{-pH}). This reformulates equation 3.4 to be:

$$\kappa = C_{indicator}\ \frac{\delta_{obs}-\delta_L}{\delta_H-\delta_L}\ +\ 10^{-pH} \tag{3.5}$$

$f_{L\ analyte}$ can be obtained by manipulating the equilibrium dissociation constant K_a to obtain equation 3.6:

$$K_{\mathrm{a}} = \frac{[H^+]\mathrm{f}_{\mathrm{L\ analyte}}}{1-\mathrm{f}_{\mathrm{L\ analyte}}} \tag{3.6}$$

consequently equation 3.2 gets reformulated to obtain:

$$\kappa = \frac{\mathrm{C}_{\mathrm{analyte}}}{1+10^{\mathrm{pKa-pH}}} \tag{3.7}$$

Rearranging 3 by taking the inverse of κ gives equation 3.8:

$$\kappa^{-1} = \frac{1}{\mathrm{C}_{\mathrm{analyte}}} + \frac{10^{-pH}}{K_a \mathrm{C}_{\mathrm{analyte}}} \tag{3.8}$$

Determination of activity coefficient γ and ionic strength I. K_a obtained from equation 3.8 would be the apparent K_a. The thermodynamic K_a is the K_a at infinite dilution with respect to ionic strength. This is represented by the thermodynamic association constant equation:

$$K_{\mathrm{a},0} = \frac{\gamma_{H+}[H^+]\gamma_{A-}[A^-]}{\gamma_{HA}[HA]} \tag{3.9}$$

Where γ is the activity coefficient relating apparent K_a to ideal $K_{a,0.}$ thermodynamic K_a is obtained by determining the activity coefficients of H^+ and A^- (γ_{HA} is assumed to equal 1 for analytes which are neutral in their protonated form and γ_{A-} is assumed to equal 1 for analytes which are positive in their protonated form). Activity coefficients are calculated via the Davies equation:

$$\gamma = 10^{-0.51z^2\left(\frac{\sqrt{I}}{1+\sqrt{I}}-0.2I\right)} \tag{3.10}$$

Where z is the charge of the ion and I is the ionic strength, γ_{H+} and γ_{A-} are determined as the average of the activity coefficients of all the NMR slices. Hence, γ_{H+} and γ_{A-} need to be

calculated for accurate calculation of pH and pK_a. To do this the ionic strength (I) is calculated at each point along the pH gradient based on the concentration and charge of the indicator and ionic strength of the solution.

Ionic strength is calculated using the standard ionic strength equation:[50]

$$I = 0.5 \sum C_i {z_i}^2 \tag{3.11}$$

Where C is concentration in mol/L and z is charge. This is done for each slice of the CSI experiment. Both the background ionic strength of the analyte solution, comprising the DSS and other potential buffers or indicators added to the solution (in present work for simplicity no buffers or other indicators were added) and the ionic strength contributed by the base indicator is calculated using equation 3.12:

$$I = I_0 + I_{base} + I_{analyte} \tag{3.12}$$

Where I_0 is background ionic strength and I_{base} is ionic strength contribution from base indicator and $I_{analyte}$ is ionic strength contribution from acidic analyte. Concentration of base indicator is calculated by measuring the integral of one of its peaks and relating it to a reference integral then determining what fraction of that concentration contributes to ionic strength (fraction protonated for nitrogenous bases and fraction deprotonated for non-nitrogenous bases). f is estimated using a modified Henderson–Hasselbalch equation:

$$\mathrm{pH} = \mathrm{pK_{a,0}} + \log_{10}\left(\frac{f_{L\ indicator}}{f_{H\ indicator}}\right) = \mathrm{pK_{a,0}} + \log_{10}\left(\frac{f_{L\ indicator}}{1 - f_{L\ indicator}}\right) \tag{3.13}$$

Obtaining $f_{L\ indicator}$ or $f_{H\ indicator}$ and multiplying them by the total concentration of the indicator at that slice gives a reasonable approximation of the concentration of non-nitrogenous base and nitrogenous base that contributes to ionic strength respectively. $I_{analyte}$ is determined to be the concentration of ionically active analyte species which is approximately equivalent to the apparent proton transferred for a non-nitrogenous analyte or the difference between the concentration of the acid and protons transferred if it is a nitrogenous analyte.

γ used for fitting was taken to be the average value of γ calculated for each useable row of the CSI dataset using Equation 3.10.

Optimum time determination. In previous work Wallace *et al.* demonstrated how to calculate the optimum time for running the experiment, t_{op}, as follows:[57] The bases used in this work are highly water-soluble and dissolve within minutes of placing the analyte solution on top. Mathematically, the base behaves as though it were diffusing from a plane source at the bottom of the NMR tube. The concentration, C_Z, of diffusing base at a height Z from the bottom of the NMR tube after a time, t, is therefore given by equation 3.14:

$$C_z = N_{(t)}\, e^{\left(\frac{-(Z-h)^2}{4Dt}\right)} \tag{3.14}$$

Where D is the diffusion coefficient of the acid and N(t) is a time dependent parameter. h is the thickness of the solid base when placed at the base of the NMR tube (2 mm).

Integrating equation 3.14 over the length of the sample gives the equation 3.15:

$$m = \pi r^2 a C_0 M_r \sqrt{\pi D t_{op}}\, e^{\left(\frac{(Z_a - h)^2}{4Dt_{op}}\right)} \quad (3.15)$$

Where r is the inner radius of the NMR tube and Mr is the molecular weight of the base, a is the number of equivalents of diffusing basic indicator to acid at height Z_a (18 mm). C_0 is the concentration of acidic species. Values for t_{op} and D are shown in table 3.1. D values were estimated *via* Stokes–Einstein Gierer-Wirtz Estimation (SEGWE) method.[51] To determine t_{op}, for an experiment, t_{op} was adjusted iteratively until m matched the mass of indicator weighed into the NMR tube (4-5 mg).

Table 3.1. D and t_{op} values of indicators used

Analyte	Indicator	D of indicator/10^{-9} m^2s^{-1}	t_{op}/hours
Glycine hydrochloride	1,2,4-triazole	0.88	4.73
H_3PO_4	1,2,4-triazole	0.88	4.73
4-cyanophenol, boric acid, NH_4Cl, NaH_2PO_4	2-methylimidazole	0.81	5.53, 5.39, 5.33, 5.30[a]
NH_3OHCl	2,6-lutidine	0.71	6.62
Benzoic acid	Acetate	0.81	5.32
Glycolic acid	Formate	0.90	4.62

[a] Analytes that were studied with 2-methylimidazole were grouped together with their respective t_{op}

Determining limiting chemical shifts of indicators. The limiting chemical shifts of formate, acetate, 1,2,4-triazole and 2-methylimidazole in table 3.1 were determined by running an NMR experiment in which 4-5 mg of solid base indicator was weighed into the tube (remaining indicator's limiting chemical shifts were obtained from literature). Four, 2 mm diameter glass beads (Assistent, Germany) were then placed on top of the base. An aliquot of an H_2O solution containing a 10 mM concentration of HCl, 0.2 mM DSS and 0.01 vol % DMSO was drawn up in a 9″ Pasteur pipet and gently layered on top of the glass beads to a height of 40-50 mm from the base of the NMR tube. A ^1H 2D CSI experiment was then run on the sample after 4-6 hours.

DSS and DMSO are used as chemical shift reference and integral reference respectively. Equation 3.7 approximates κ to $C_{analyte}$ as pK_a decreases (This is true only when $C_{indicator}$ >> $C_{analyte}$ as κ cannot be greater than $C_{indicator}$). Equation 3.7 can be reorganised to produce equation 3.16 where a linear plot $\kappa/C_{indicator}$ versus δ_{obs} is produced (Figure 3.1) with the slope as $(\delta_H-\delta_L)^{-1}$ and the intercept as $(-\delta_L/(\delta_H-\delta_L))$. Hence, through simple algebraic manipulation δ_H and δ_L can be calculated.

$\kappa = \frac{C_{analyte}}{1+10^{pKa-pH}} \approx C_{analyte}$ when HCl is the acidic analyte (pK_a << 0)

$$\frac{\kappa}{C_{indicator}} = \frac{\delta_{obs}}{\delta_H - \delta_L} - \frac{\delta_L}{\delta_H - \delta_L} \quad (3.16)$$

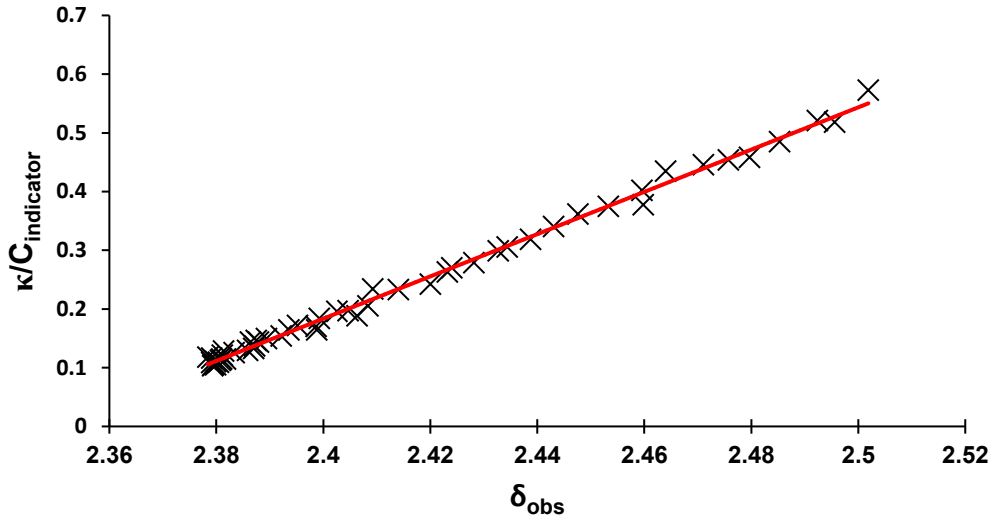


Figure 3.1. Plot of $\kappa/C_{indicator}$ versus δ_{obs} for 2-methylimidazole with 10 mM HCl in H_2O solution.

Determination of concentration of indicators. To calculate the concentration of the base indicator across the NMR tube, a known concentration of either dioxane or DMSO are inserted into the analyte solution and the integral of one of the peaks of the base indicator is referenced to the peaks of DMSO or dioxane. A correction factor needs to consider the different number of protons for each respective peak along with imperfections due to incomplete T_1 relaxation. The correction factor is measured by preparing an H_2O solution with a known concentration of 10 mM of the base indicator and 0.01 vol % of DMSO or dioxane with DSS as a chemical shift reference. A ^{1}H 2D CSI experiment is then run on the sample immediately and the integrals of DMSO or dioxane and the base indicator are measured. Correlating integral information and concentration is obtained using equation 3.17:

$$C_x = \frac{I_x}{I_{cal}} \times \frac{N_{cal}}{N_x} \times C_{cal} \times \rho \tag{3.17}$$

where I, N, ρ, and C are the integral area, number of nuclei, imperfections due to incomplete T_1 relaxation, and concentration of the compound of interest (x) and the calibrant (cal), respectively. All compounds used in the present work had ≥ 98 % purity.

$\frac{N_{cal}}{N_x} \times C_{cal} \times \rho$ are defined as correction factor λ and equation 3.17 can be rearranged to obtain λ:

$$C_x \times \frac{I_{cal}}{I_x} = \lambda \tag{3.18}$$

Table 3.2 displays correction factors for the indicators used in our experiments (Some indicators had only a correction factor of one of the reference compounds). Correction factor used was an average of the factors calculated for each experimental slice. Table 3.3 shows which integral references were used for each analyte choice was based on which integral reference gave least amount of spectral overlap.

Table 3.2. Indicator's correction factor for DMSO and dioxane when applicable

Indicator	λ **(Dioxane)**	λ **(DMSO)**
2-methylimidazole (aromatic)	4.69	5.43
2-methylimidazole (methyl)	2.51	2.86
2,6-lutidine (methyl)	1.37	N/A due to signal overlap
Acetate	2.14	2.90
Formate	9.45	12.73
1,2,4-triazole	5.74	5.59
$MPAH^-$	Not determined	1.84

Table 3.3. Table showing which integral reference was used for each analyte

Dioxane	**DMSO**
NaH_2PO_4	H_3PO_4
4-CN	glycine
Boric acid	Benzoic acid
NH_4Cl	Glycolic acid
NH_3OHCl	Lectin
PAA	

Determination of pH and κ. The pH of a solution is determined from the chemical shift of an indicator molecule, δ_{obs}, by equation 3.19:

$$pH = \mathrm{p}K_{\mathrm{a},0} + \log\frac{\delta_{\mathrm{obs}} - \delta_{\mathrm{H}}}{\delta_{\mathrm{L}} - \delta_{\mathrm{obs}}} + \Delta z^2\left(\frac{0.51\sqrt{\mathrm{I}}}{1+\sqrt{\mathrm{I}}} - 0.1\mathrm{I}\right) \quad (3.19)$$

where δ_H and δ_L are the limiting chemical shifts of the protonated and deprotonated species respectively, Δz^2 is the difference in the square of the charge of the indicator between the protonated and deprotonated states (-1 and +1 for non-nitrogenous and nitrogenous species in our work, respectively) and I is the ionic strength.[102,111,69] Basic indicators used in this work are provided in Table 3.4 (All p$K_{a,0}$, δ_H and δ_L values were obtained from literature sources except with 1,2,4-triazole where its p$K_{a,0}$, δ_H and δ_L was obtained experimentally using the method described in Chapter 2).[57, 131, 132]

Table 3.4. Basic NMR pH indicators used in this work

Indicator	p$K_{a,0}$	δ_H/ppm	δ_L/ppm	pH range[a]
2-MI	7.96	7.270	6.958	9-7
2,6-lutidine	6.75	2.707	2.456	8-6
Acetate	4.76	2.083	1.906	5.5-3.5

Formate	3.75	8.266	8.441	4.5-2.5
1,2,4-triazole	2.45	9.193	8.352	3-1
$MPAH^-$	2.12	1.510	1.281	3-1

[a]The pH range accessible when measuring the chemical shift of the indicator. The uncertainty in the measurement of chemical shift precludes accessing pH values beyond this range

Assuming protonation of the indicators from H_2O is negligible due to their low basicity, the concentration of protons transferred from the acidic analyte, κ, is represented through equation 3.5 and 3.7 derived earlier with equation 3.8 used for linear fitting:

$$\kappa = C_{indicator}\frac{\delta_{obs}-\delta_L}{\delta_H-\delta_L} + 10^{-pH} \tag{3.5}$$

$$\kappa = \frac{C_{analyte}}{1+10^{\mathrm{pKa-pH}}} \tag{3.7}$$

$$\kappa^{-1} = \frac{1}{\mathrm{C_{analyte}}} + \frac{10^{-pH}}{K_a \mathrm{C_{analyte}}} \tag{3.8}$$

κ is measured using equation 3.5 and the pH is measured using equation 3.19. Equation 3.8 forms a linear plot of κ^{-1} versus 10^{-pH} with a gradient of $1/K_aC_{analyte}$ and intercept of $1/C_{analyte}$ where K_a is of the 'mixed' type (the dissociation constant is written in terms of the concentration of the conjugate acid/base and activity of H^+).[115] Figure 3.2 showcases the scheme used in the experiment and the pK_a and $C_{analyte}$ obtained for boric acid and phosphoric acid.

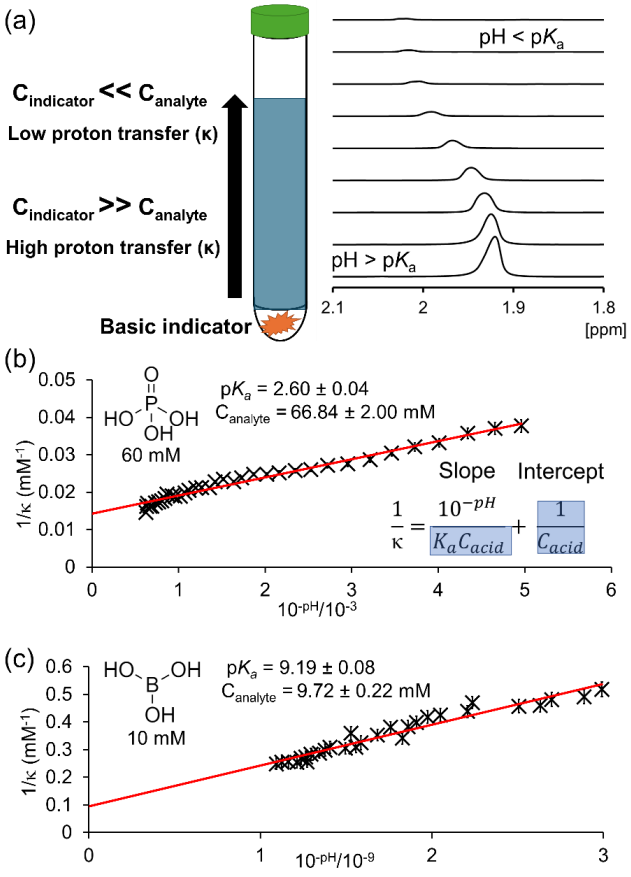


Figure 3.2. (a) A concentration gradient of a basic indicator is established in an NMR tube, allowing measurement of the quantity of protons transferred from acid to base as a function of pH; (b) plot of $1/\kappa$ versus 10^{-pH} for 60 mM H_3PO_4; (c) plot of $1/\kappa$ versus 10^{-pH} for 10 mM boric acid

$pK_{a,0}$ and $C_{analyte}$ were determined from the gradient and intercept of a plot of $1/\kappa$ versus 10^{-pH}. The values of pK_a obtained by linear fitting are within 0.5 units of the literature values (uncertainty values obtained are the 95% confidence intervals of the least squares regression line) across a wide range of analytes.[133-141] The indicators used were chosen to be within 2 units of the expected pK_a of the analyte and give pH ranges that cover ± 1 units across the pK_a of the indicator. Obtaining data at pH values more than one unit higher or lower than the pK_a of the indicator introduces a high degree of uncertainy (see appendix C.1). All analytes studied with their respective parameters and fitted values of pK_a and $C_{analyte}$ are listed in Table 3.5. pK_a measurements obtained were comparable in accuracy to standard methods.[142-145] To highlight the potential of the method to determine the acidities of polymeric systems, pK_a of polyacrylic acid was determined (Mw 240 kDa, 10 mM COOH groups) with sodium acetate as an indicator. A pK_a value of 4.83 ± 0.02 was obtained with an R^2 of 0.94, in good agreement with the literature value of 4.5, while the total concentration of acidic groups (9.65 mM ± 1.50) was also recovered by the fitting.[146]

Table 3.5. Literature and fitted values of pK_a and $C_{analyte}$

Compound	Fitted $pK_{a,0}$	Literature $pK_{a,0}$	$C_{analyte}$ /mM[a]	Indicator
Glycine	2.78 ± 0.23	2.40[138]	8.51 ± 1.50	$MPAH^-$
H_3PO_4*	2.60 ± 0.04	2.16[136]	66.84 ± 2.00	1,2,4-triazole
NaH_2PO_4*	7.36 ± 0.24	7.21[136]	8.95 ± 4.10	2-MI
4-CN	7.92 ± 0.03	7.95[133]	10.60 ± 0.25	2-MI
Boric acid*	9.27 ± 0.08	9.19[137]	9.72 ± 0.22	2-MI
NH_3OHCl*	5.76 ± 0.23	5.94[136]	8.92 ± 1.50	2,6-lutidine
NH_4Cl*	9.43 ± 0.27	9.25[139]	9.80 ± 3.40	2-MI

Benzoic acid	4.29 ± 0.04	4.20[134]	9.60 ± 0.24	Acetate
Glycolic acid	3.89 ± 0.06	3.89[140]	11.73 ± 1.02	Formate

[a]Concentrations were 10 mM, except H_3PO_4 (60 mM). *Compounds do not exhibit distinct ^1H resonances in H_2O due to fast exchange of NH and OH protons.

The results on PAA highlight the potential of the technique for the determination of the pK_a of ^1H inactive polymers which typically are hard to probe using NMR given their short T_2 values and considerable peak broadening effects (see appendix C.5).[147] The resonances of PAA disappear as the concentration of sodium acetate increases along a concentration gradient (Figure 3.3.b,c). Deprotonation of PAA will increase the hydrodynamic radius which is expected to cause a net migration of the polymer towards the diffusing base, in analogy to the diffusion of polygalacturonate towards diffusing calcium. [146, 148] The higher concentration and partial protonation of the PAA may lead to entanglement of the chains causing broadening of the PAA resonances beyond detection. We note that the resonances of PAA also decrease up a concentration gradient of NH_3, while no loss of signal is observed along a gradient of NaCl or in homogeneous solutions of PAA and sodium acetate (see appendix C.5). These results suggest that the loss of the PAA resonances is related to the presence of a concentration gradient of base rather than deprotonation alone or diffusiophoresis.[149] The method still returns accurate values of $C_{analyte}$ and pK_a. The method does not depend on the direct observation of the chemical shift of PAA and as a result is advantageous in situations where the analyte agglomerates or diffuses in solution after deprotonation has taken place.

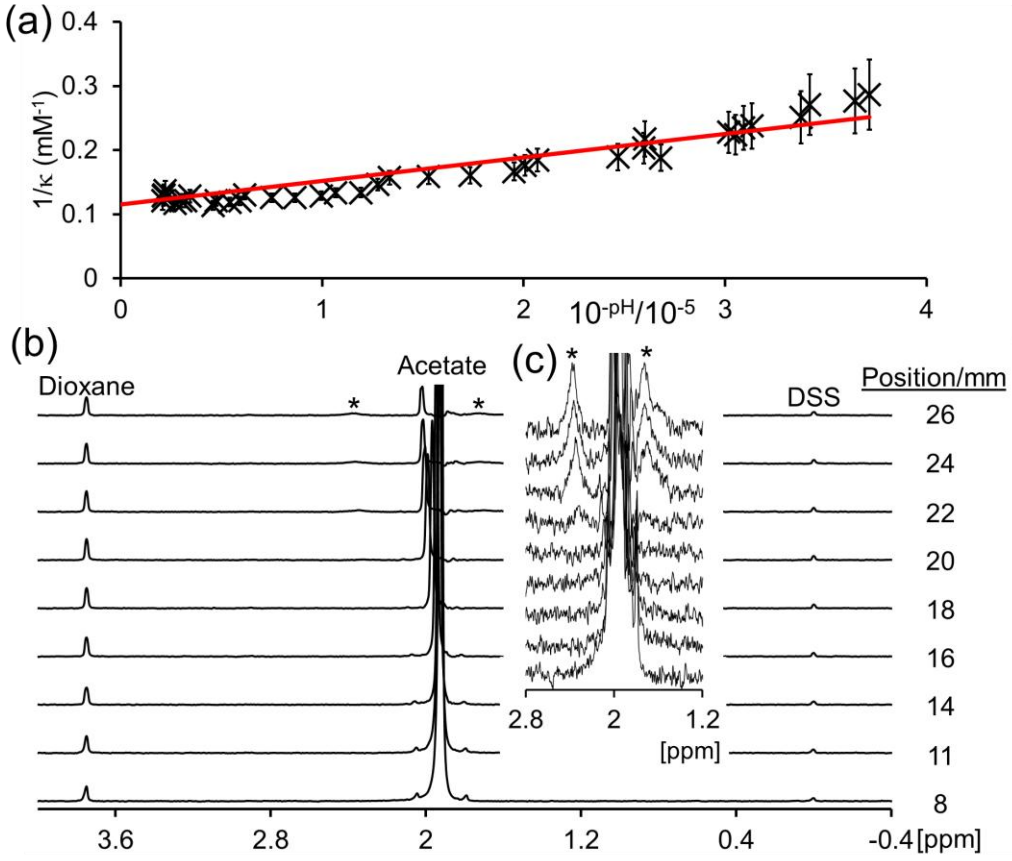


Figure 3.3. (a) Plot of $1/\kappa$ versus 10^{-pH} for polyacrylic acid (10 mM COOH groups) with sodium acetate as indicator**;** (b) Partial ^{1}H NMR spectra at different positions from the base of the NMR tube; (c) Expansion of the acetate and PAA region is presented to highlight the disappearance of the PAA resonances (*)

To highlight the potential of the method for studying the pH-dependent properties of proteins, simplifying the process and time needed to study these properties using solution state NMR, we determined the pH at which the net charge of a protein molecule is zero (defined as the isoelectric point p*I*).[150] of wheat germ agglutinin (WGA, 3.8 mg/mL) using 2-methylimidazole as the indicator. A p*I* value of 7.76 ± 0.4 was obtained in good agreement with the isoelectronic point (p*I*) of 7.70 (Figure 3.4).[151] $C_{analyte}$ was calculated as 2.38 ± 0.24 mM. Assuming a molecular weight of 21 kDa, the concentration of protein is 180 μM which would suggest 13 ± 1 acidic sites per protein. This value is consistent with the ca.10 acidic residues (glutamic acid or aspartic acid) reported in the sequences of WGA isolectins.[152]

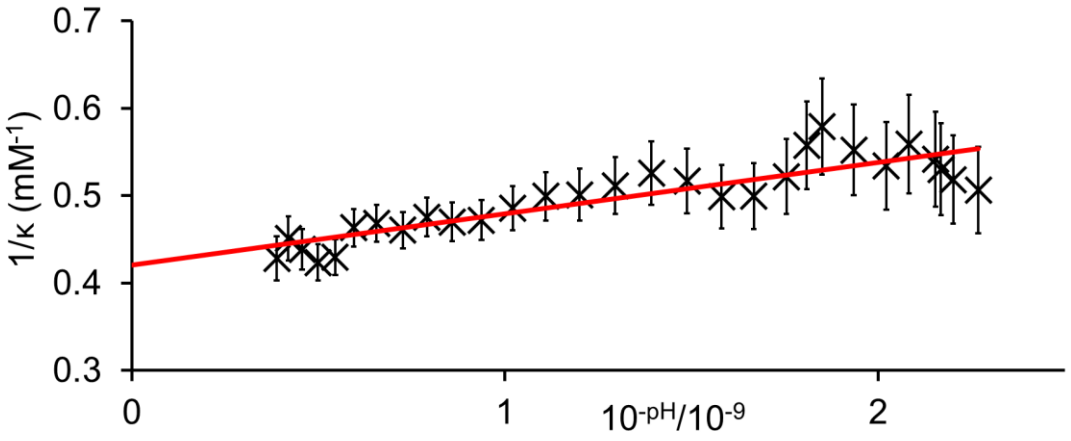


Figure 3.4. Plot of $1/\kappa$ versus 10^{-pH} for WGA with 2-methylimidazole as indicator

3.6 Conclusions

This chapter has shown how the pK_a of any substance can be determined in a single NMR experiment at a concentration of acidic sites as low as 2 mM. With a standard NMR sample volume of *ca*.500 μL, this approach provides a substantial saving in sample quantity and experimental time relative to conventional workflows based on 2D heteronuclear NMR experiments and manual adjustment of the sample pH. Furthermore, by avoiding completely the requirement for direct observation of the molecule of interest, any molecule or ion can be analysed including those that do not exhibit observable NMR resonances (NH_3OHCl), or else resonances that are too broad (PAA) or not pH responsive (WGA) for analysis based on the pH-dependence of their ^{1}H chemical shifts. Knowledge of pK_a of these large systems could be used to inform the design of polymer-drug conjugates or liposomal formulations.[153] This approach could potentially be extended to study the degree of proton transfer between acidic and basic partners at different volume fractions of organic solvent, greatly accelerating the determination of the pK_a of compounds with insufficient solubility for direct analysis in water.[154,155, 156]

Chapter 4: Aqueous pK_a Determination of Water Insoluble Analytes from Solvent Mixtures Using ^{1}H Chemical Shift Imaging NMR

4.1 Chapter summary

Determination of pK_a provides information on the degree of ionization of compounds and hence provides valuable information on the absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of pharmaceutical ingredients. Difficulties arise when measuring the pK_a of pharmaceutical compounds as a significant portion of them are water-insoluble and as a result their pK_a are measured using water-solvent mixtures and the pure water pK_a is predicted using Yasuda Shedlovsky equation (YS). This approach necessitates the preparation of multiple solvent mixtures which poses challenges in traditional potentiometric titrations, due to the need to perform lengthy and often inconvenient calibration of pH electrodes on solvent mixture of interest. This chapter introduces composition gradients of the solvent as a powerful new dimension for NMR analysis on both high-field and benchtop instruments. Taking advantage of the variation in the density of binary solvent mixtures as a function of their composition, two solutions are layered at opposite extremes of the compositional range in an NMR tube. Diffusion of the layers into each other establishes a continuous variation in solvent composition across the sample. Spatially resolved analysis of the sample using chemical shift imaging enables analysis of chemical systems as a function of the solvent composition. As proof of concept the pK_a of poorly water-soluble active pharmaceutical ingredients (APIs) across a wide composition range of dimethyl sulfoxide (DMSO)/water is measured with accurate extrapolations to the aqueous pK_a using the Yasuda Shedlovsky method. The hours of tedious experiments are thus condensed into a single 20-minute experiment.

4.2 Author contributions

Chapter 4 is in the process of being incorporated into a paper, “High and low-field NMR in binary solvent Gradients”, aimed to be published in *Analytical Chemistry* as a joint article with Paulina Putko and Krzysztof Kazimierczuk from Centre of New Technologies, University of Warsaw, and Dariusz Golowicz from institute of Physical Chemistry, Polish Academy of Sciences. The joint paper has a high field NMR section and a low field NMR section. The author’s contributions are as follows: Haider Hussain conceived of the concept of the high field section of the paper, designed the experiments, prepared all the samples, and performed all of the NMR experiments. Matthew Wallace provided helpful guidance throughout the project and determined the pK_a values of the indicators in 20, 50 and 80 vol% DMSO-d_6. Haider Hussain wrote the high field section of the paper aided by helpful suggestions from the other authors. The figures were made by Haider Hussain. Supporting information sections was also made by Haider Hussain.

4.3 Introduction

The solvent that a chemical system is present in is a fundamental component that needs to be taken into account in understanding the dynamics and kinetics of any chemical system.[157] A solvent is commonly understood as the component (or components) of a solution that forms the continuous phase of the system and is present in the largest molar fraction, while the remaining components are classified as solutes. The solvent itself may be a mixture of compounds; nevertheless, this mixture can be treated as a single solvent phase distinct from the solutes due to its dominant contribution to the physicochemical environment of the system.[158] Other typical features of a solvent is that solvents typically have the capacity to dissolve their respective solutes and typically the molar fraction of the solvent with respect to the whole solution is much larger compared with the solutes (in which case it would be called a dilute solution).[159, 160] From the definition of what a solvent is alone one can grasp the relative importance of it with regards to the chemical systems present given the fact the solvent is typically the chemical present in greatest quantity, that it surrounds all the solutes in way that maximises solvent-solute interactions and that given the fact that it is the medium in which the solutes "swim" in makes the solvent a medium between all the different solutes which impacts all the different exchange dynamics that occur between the different solutes such as molecular diffusion, ion transport and heat transport to and from chemical reactons.[158, 161, 162] As a result the effects that a solvent has upon the solubility, stability and reaction rates of chemical systems present within it is very high.[163, 164-165]

In the case of acid dissociation constant the solvent can have a significant impact on the degree of ionisation of compounds.[166, 167] This is because polar solvents tend to stabilise ionised species better than less polar solvents. This is due to their higher dielectric constant (ε) which allows for better solvation of charged particles.[168] Other non-electrostatic solvent effects such as hydrogen bonding, dipole-dipole interactions and van der Waals forces can also influence the acid dissociation constant.[166] Given that most pharmaceutically relevant drugs have low water solubility, measuring their pK_a in water is often impractical, and determining it in other solvents can yield values that differ significantly from those in water, limiting clinical relevance.[169]

As a result pK_a of water-insoluble pharmaceutical compounds is typically not determined directly by studying them in water rather they are determined by measuring the dissociation of constant of said compound in water-solvent mixtures (where typically the other solvent is dimethyl sulfoxide (DMSO) or methanol) and then varying the solvent composition of the mixture and recording the values of the solvent dependent pK_a obtained (labelled in the literature as $_s$pK_a meaning the pK_a measured at a specific solvent s) and then extrapolating from the values measured to what the pK_a of the compound would be in pure water.[170] Typically this has been undertaken using potentiometric titrations which requires inexpensive materials and are very simple to use.[171] Although inexpensive, this process has many disadvantages such as: laborious and time consuming due to the need to prepare multiple samples to measure their pK_a in order to have an adequate data sample to obtain reliable extrapolations additionally any small deviation in the electrolyte pH can alter the results of the titration, different calibrators can be toxic and lead to contamination during the electrolyte preparation causing deflections from the actual results.[127,120]

In this chapter a more efficient single shot experiment is presented which allows for establishing a solvent gradient across a single sample and record spatially resolved spectra across the length of the entire gradient using chemical shift imaging techniques.[57, 127] The solvent gradient established allows for the determination of pK_a values without the need for

traditional titrations. A solvent gradient is achieved by layering a DMSO rich solution with H_2O rich solution that contains both the analyte and its respective indicator and allowing the two solutions to mix in a single NMR tube and causing the solvent composition to vary in the NMR tube. This results in different slices which corresponds to different solvent compositions in the NMR tube which allows a unique pK_a value to be worked out for different DMSO/H_2O solvent compositions with the help of CSI. Using the Yasuda Shedlovsky (YS) an extrapolation can be done to find the pK_a of an analyte in pure water. This method tackles the issues of aqueous solubility and saves experiment time and material cost. Furthermore, this technique can work even if the analyte does not have visible ^{1}H NMR peaks. This is because the pK_a of the analyte is not measured directly but rather through the use of indicators that determine pK_a by measuring the quantity of protons the indicators obtained from the analyte as described in chapter 3.

4.4 Experimental section

Materials. All chemicals were purchased from commercial suppliers and used as received.1,2,4-triazole (TRI), Sodium formate, acetate, and sodium glycolate were used as proton transfer indicators to determine pK_a. Sodium methyl sulfonate (NaMSA) was used as a solvent composition indicator.2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used chemical shift (0 ppm) reference. Stocks of 20% H_2O and 80% DMSO, 80% H_2O and 20% DMSO solution of these compounds were prepared and used throughout the study. The analytes studied were naproxen, indomethacin, furosemide, quinine hydrochloride, salicylic acid, acetic acid, formic acid, sulfoacetic acid and glycolic acid. Stock solutions of the analytes in H_2O were prepared at 5 mM concentration of each except sulfoacetic acid which was prepared at 15 mM to explore the scope of the methodology.

NMR. Experiments were performed on a Bruker Avance III 500 MHz spectrometer operating at 500.21 MHz for ^{1}H with a 5 mm broadband observe probe. The probe was equipped with z-axis pulsed field gradients. The temperature of the samples was maintained at 298±0.5 K. the 90° hard pulse was 10 µs with a respective of 22.677 Watt. The relaxation delay was 2 seconds. The spectral width was 16.02 ppm, and the number of data points were 32 k, giving an acquisition time of 2.045 seconds. Experiments were locked to DMSO. CSI experiments were performed using a gradient phase encoding sequence based on that of Trigo-Mouriño *et al*.[51] and incorporating double echo excitation sculpting for water suppression. The pulse sequence was thus ES-τ1-g-τ2-acquire where ES is excitation sculpting and g is a gradient pulse and τ1 and τ2 are delays of 10 and 200 µs, respectively. Balancing delays were incorporated into the ES component. A spoil gradient (27 G/cm) was employed at the end of the signal acquisition period to destroy any transverse magnetization. The gradient pulse was 172 µs in duration and varied between −18.8 and 18.8 G/cm in 32 steps. The shape of the pulse was a smoothed square.4 ms Gaussian 180° pulses were employed for inverting magnetisation of the H_2O region. Four scans were acquired at each step giving a total acquisition time of 20 min and a theoretical spatial resolution of 0.81 mm. Sixteen dummy scans were acquired prior to signal acquisition.

NMR data was processed in Bruker TopSpin 4.6.5 and Mnova 14.2.0 and CSI images were processed in sensitive mode following the procedure of Trigo-Mouriño *et al*.[51] Following two-dimensional Fourier transformation of the CSI data sets, the individual spectra were automatically extracted, any residual phase errors corrected, and the spectra referenced to

DSS (0 ppm) using an automation script written in house. Analyte and indicator chemical shift data were imported from TopSpin and Mnova into Microsoft Excel and subsequently analysed.

4.5 Results and Discussion

Establishment of a solvent gradient. To establish a solvent gradient, two solutions were prepared containing the analyte, DSS, NaMSA and the indicator/indicators, one solution having 0.8 F_{dmso} and the other 0.2 $F_{dmso.}$ The concentration of the analyte was equal to the concentration of the indicator when one indicator was used. When multiple indicators were used the sum of their concentration equalled the concentration of the analyte.320µL of the 0.8 F_{dmso} solution was put into the NMR tube. This was followed by 320µL of the 0.2 F_{dmso} solution being gently layered on top using a 9 inch Pasteur pipette (see figure 4.1). The 0.8 F_{dmso} solution was put at the bottom of the NMR tube due to it being a denser solution.

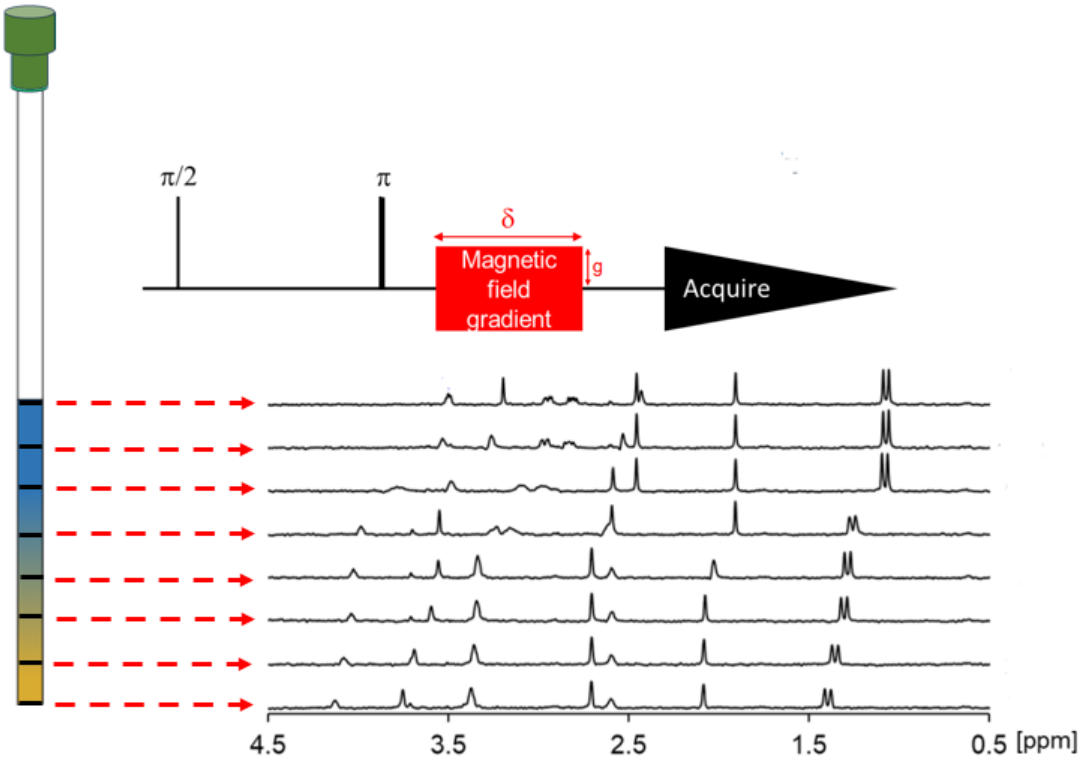


Figure 4.1. Scheme of how the solvent gradient experiment works two solutions are layered into the NMR tube orange is the 0.8 F_{DMSO} solution and the blue solution is the 0.2 F_{DMSO} and a gradient is formed between them across the NMR tube leading to changes to chemical shift based of solvent composition

Once a solvent gradient is produced.2D ^{1}H CSI experiments are run and multiple spectra of the analyte experiencing different solvent compositions across the NMR tube are obtained. The spectra are then used to obtain pK_a of analyte across the NMR tube and extrapolate to what the solvent dependent pK_a (p$_s$$K_a$) of the analyte will be in pure water using the Yasuda Shedlovsky extrapolation. The optimum time for running the solvent gradient experiment was determined by running ^{1}H 1D CSI NMR experiments with 320 µL of 80/20 DMSO H_2O fractional volume solution at the bottom of the NMR tube with 320 µL of 20/80 DMSO H_2O

fractional volume solution layered on top. Both solutions contain 0.2 mM of DSS, 20 mM of sodium methane sulfonate. Data was recorded every twenty minutes for an eight-hour period and another experiment was run 24 hours after preparation. Results show that within 3 hours 40 minutes to 8 hours from sample preparation the solvent gradient has a good compromise between it having a smooth linear change of fractional volume with respect to chemical shift of sodium methane sulfonate whilst also having the difference be significant enough to induce significant pH and hence analyte pK_a changes (see figure 4.2).

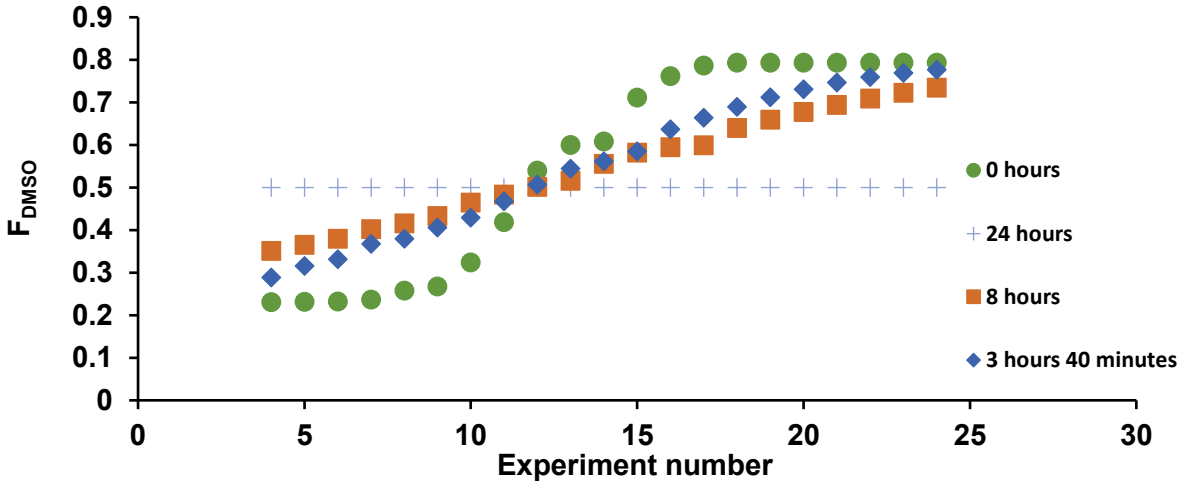


Figure 4.2. Plot of F_{DMSO} versus experiment number of solvent gradients at various times

Determination of dielectric constant. Determination of the value of the dielectric constant across the CSI experiment was done using the model utilized by Jouyban *et al.* which models the dielectric constant of a binary solvent mixture:[172]

$$\ln(\varepsilon) = \varphi_1 \ln(\varepsilon_1) + \varphi_2 \ln(\varepsilon_2) + \varphi_1 \varphi_2 \sum_{i=0}^{2} K_i (\varphi_1 - \varphi_2)^i \qquad (4.1)$$

Where ε_1 is dielectric constant of H_2O, $\varphi_1$is fractional volume of H_2O, ε_2 is dielectric constant of DMSO, φ_2 is fractional volume of DMSO and K_i are optimisation constants.

The model was calibrated for DMSO/H_2O solution mixtures by fitting the Jouyban *et al.* equation to the dielectric constants of DMSO/H_2O solvent mixtures measured by Płowaś *et al* to obtain the solvent dependent optimisation constants K_0, K_1 and K_2 (see figure 4.3).[173] Dielectric constant values obtained using the equation were compared with experimental values with reasonable accuracy obtained (see table 4.1). The dielectric constant across thus determined by measuring the fractional volume of DMSO and the calculating the dielectric constant using equation 4.8.

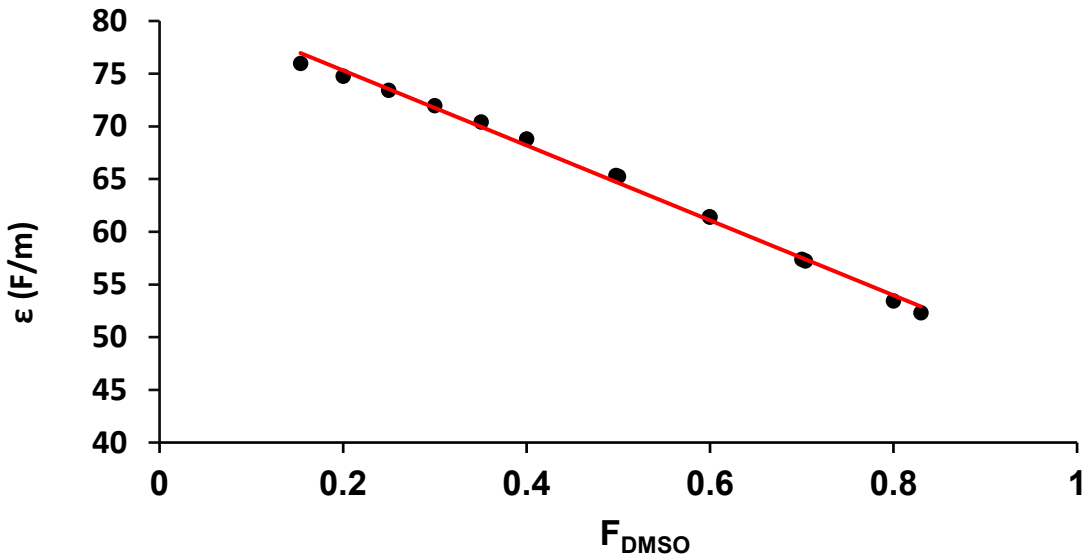


Figure 4.3. Plot of ε of water/DMSO binary mixtures versus F_{DMSO} of those mixtures with respective fitting

Table 4.1 Fitted and literature ε values of water/DMSO binary mixtures

F_{DMSO}	Literature ε	Fitted ε
0.1	77.14	77.33
0.1536	75.97	75.96
0.2	74.83	74.75
0.2496	73.52	73.41
0.3	71.86	71.96
0.3505	70.29	70.42
0.4	68.09	68.80
0.4975	65.13	65.34
0.5	67.93	65.25

F_{DMSO} measurement. Fractional volume of DMSO is measured using sodium methane sulfonate as solvent indicator. Sodium methane sulfonate was used due to its low pK_a value (-2.6) thus for the pH range of interest for our experiments the chemical shift of sodium methane sulfonate is independent of pH.[174] Thus by measuring the chemical shift of sodium methane sulfonate and seeing the change in chemical shift as a function of fractional volume of DMSO (see figure 4.4) an equation that relates the chemical shift of sodium methane sulfonate to fractional volume of DMSO is obtained. For that end a series of 1D ^{1}H NMR experiments containing 20 mM sodium methane sulfonate, 0.2 mM DSS and each containing different DMSO/H_2O solvent composition ranging from 80% fractional volume of DMSO to 20% fractional volume of DMSO.

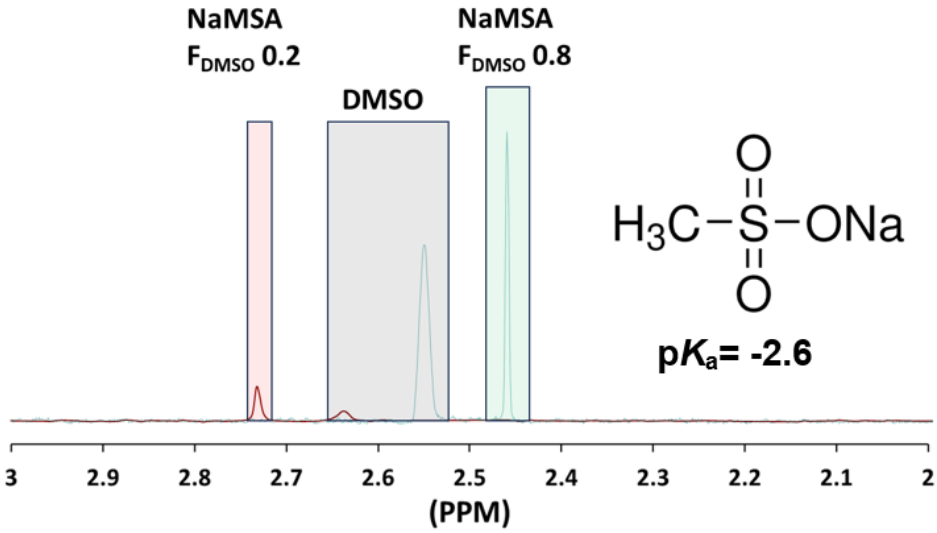


Figure 4.4. Spectra of NaMSA at 0.2 and 0.8 fractional volume of DMSO (DSS was used as reference compound for all experiments)

Polynomial fitting is done with respect to this data with the chemical shift of sodium methane sulfonate as the independent variable and the fractional volume of DMSO as the dependent variable and a 3rd order polynomial is obtained (see figure 4.5). In order to test the polynomial a 1D ^{1}H NMR experiment was run with 20 mM sodium methane sulfonate, 0.2 mM DSS with a fractional volume of DMSO of 0.45 and the polynomial predicted the value with reasonable accuracy (see table 4.2). Furthermore, to check for whether the indicator is insensitive to pH change the experiment was repeated with an additional 5 mM of HCl to it and the polynomial was used to predict the solvent composition and reasonable accuracy was obtained.

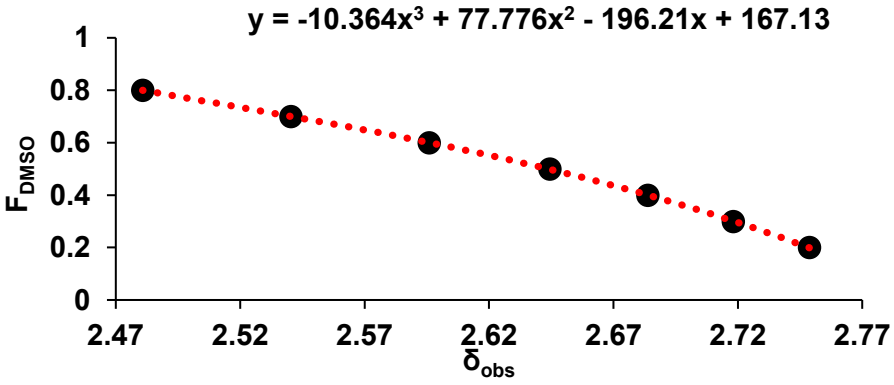


Figure 4.5. Plot of F_{DMSO} versus δ_{obs} of NaMSA with respective fitting

Table 4.2. F_{DMSO} measurement using NaMSA with HCl and no HCl solution

	F_{DMSO}	Measured F_{DMSO}
HCl	0.45	0.443

No HCl	0.45	0.446

Calibration of indicator compound.

For purposes of this experimental setup the limiting chemical shifts of the indicator along with its pK_a need to be known for any given DMSO/H_2O solvent composition each of the 1D slices in the experiment contains. This requires the obtaining of an equation that models the limiting chemical shift and pK_a of an indicator as a function of the solvent composition. For this purpose, the limiting chemical shifts and pK_a of the indicators are measured and then equations are fitted on the measurements to obtain reliable predictions of how the values of limiting chemical shift and pK_a change as a function of solvent composition.

Measurement of pK_a of indicator compounds across multiple DMSO/H_2O solvent compositions using the method of Wallace *et al.*[120] The method works by forming an indicator ladder where a 2D ^{1}H CSI NMR experiment is run with multiple indicators present, and fitting pK_a values in the order of increasing basicity, based initially on the pK_a values of 2,6-DHB. The pK_a of 2,6-DHB is previously determined for any solvent mixture by layering a solution of sodium 3-(trimethylsilyl)-1-propanesulfonate (DSS, 10 mM) on top of 8–9 mg of 2,6-DHB.). In the absence of other acids or bases, the fraction of 2,6-DHB in its deprotonated state, f_L, is given by equation 4.2:

$$f_L = \frac{-K_a + \sqrt{{K_a}^2 + K_a C}}{2C} \tag{4.2}$$

where K_a is concentration based and C is the total concentration of 2,6-DHB at each position along the sample. Assuming a fast exchange on the ^{1}H NMR time scale, the chemical shift of 2,6-DHB, δ_{DHB}, is given by:

$$\delta_{DHB} = f_L \delta_L + (1 - f_L)\delta_H \tag{4.3}$$

where δ_H and δ_L are the limiting chemical shifts of the protonated and deprotonated states, respectively. C is obtained by integrating the ^{1}H resonance of the 3,5-position of the aromatic ring of 2,6-DHB against the methyl resonance of DSS at each position along the sample using CSI. The molar ionic strength, I, of the solution at each position is calculated as the sum of the concentrations of DSS and dissociated 2,6-DHB. The activity coefficient of a univalent ion, γ, is obtained from eq 4.4:

$$\log_{10} \gamma = \frac{-A\sqrt{I}}{1 + B\sqrt{I}} \tag{4.4}$$

where A and B are parameters to modulate what the activity coefficient will be depending on the medium. A and B can be calculated from the equations of Schuman *et al.* :[175]

$$A = 1.826 \times 10^6 (\varepsilon \mathrm{T})^{-3/2} \tag{4.5}$$

$$B = R \times 50.3 \times (T\varepsilon)^{-\frac{1}{2}} \tag{4.6}$$

Where ε is dielectric constant, R is the radius of an ion which is assumed to be 6.5 Å and T is absolute temperature.[175] Therefore, A is taken as 0.52, 0.56 and 0.69 and B is taken as 1.32, 1,36 and 1.46 for 20%, 50% and 80% fractional volume of DMSO respectively.

In the absence of activity data ionic strength can be ignored and the method will yield K_a values uncorrected for ionic strength. The concentration-based K_a is calculated from the thermodynamic pK_a (p$K_{a,0}$) using eq 4.7:

$$K_a = \frac{10^{-pK_{a,0}}}{\gamma^2} \tag{4.7}$$

The p$K_{a,0}$ of 2,6-DHB is obtained by fitting the chemical shift of the 3,5-position to equation 4.3 with p$K_{a,0}$, δ_H, and δ_L as free variables. Once 2,6-DHB parameters are obtained the other indicators can be obtained by measuring the pH of the solution using 2,6-DHB using the NMR modified Henderson-Hasselbalch equation with the p$K_{a,0}$, δ_H, and δ_L as free variables. Once an indicators parameters are known other indicators of higher p$K_{a,0}$ value are also determined. Values obtained for indicators are provided in Appendix 4 D.4 .

Once the limiting chemical shifts are obtained polynomial fitting is then done on them with respect to fractional volume of DMSO (See figure 4.6). Thus, by measuring the fractional volume of DMSO the limiting chemical shifts of the indicator can be determined for each 1D spectra by putting the specific fractional volume of DMSO for that spectrum into the polynomial equation.

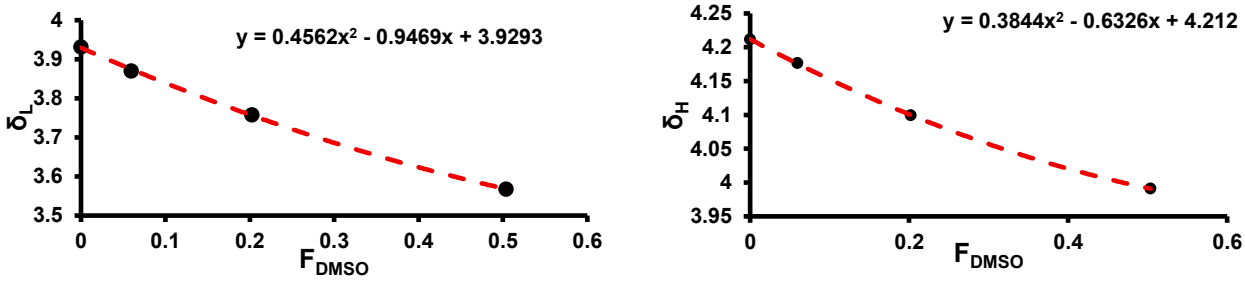


Figure 4.6. (left) Plot of δ_L versus F_{DMSO} of sodium glycolate with respective fitting, (right) plot of δ_H versus F_{DMSO} of sodium glycolate with respective fitting

Determination of equation for pK_a of the indicators across solvent composition is done by using the pK_a values of indicator already obtained for 80, 50 and 20 % fractional volume of DMSO using the method described and fitting them into the Yasuda Shedlovsky equation. By fitting the Yasuda Shedlovsky equation for the pK_a values already determined the slope and y-intercept of the linear equation is obtained. This allows for determination of pK_a of indicator as a function of the dielectric constant which can be determined via measuring the fractional volume of DMSO as described previously across the different 1D spectra by determining the dielectric constant specific for that spectra and inputting it into the Yasuda Shedlovsky equation (see figure 4.7 for example of fitting for glycolic acid and table 4.3 and figure 4.11 (e) for list of indicators and analytes measured by them).

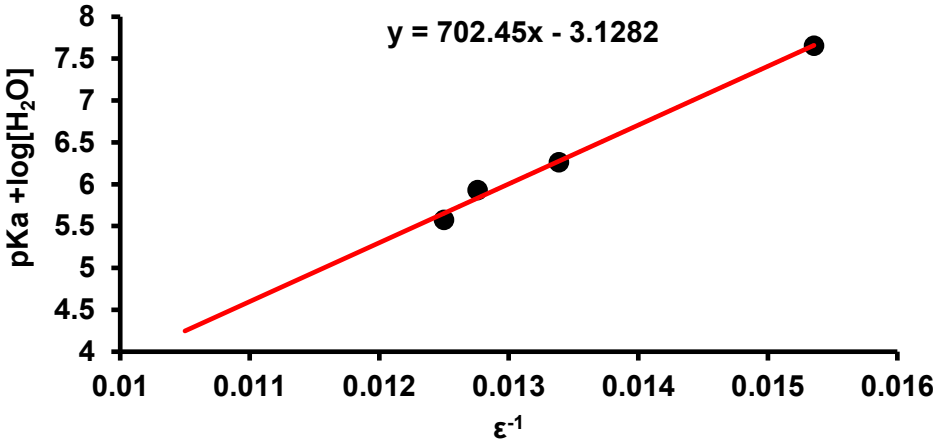


Figure 4.7. Plot of log[H_2O] + p$_s$$K_a$ versus ε^{-1} of sodium glycolate with respective fitting

Determination of concentration of H_2O. Concentration of H_2O across the NMR tube was done by assuming the density of water to be 1 g/mL and molar concentration of H_2O in mixture ($M_{H_2O,mixture}$) is calculated by multiplying the fraction volume of H_2O (F_{H_2O}) for each of the 32 slices of the CSI experiment (see figure 4.8) slice by the molar concentration of water in its pure form ($M_{H_2O,pure}$ which is 55.5 M).

$$M_{H_2O,mixture} = F_{H_2O} M_{H_2O,pure} \tag{4.8}$$

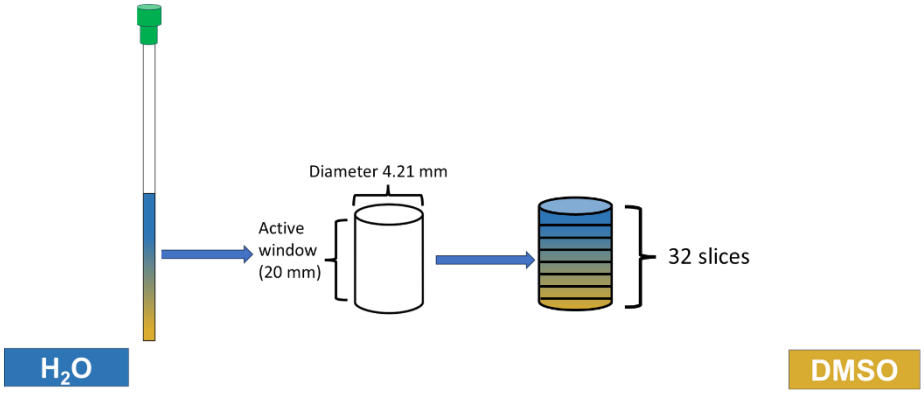


Figure 4.8. Scheme for calculating molar concentration of H_2O

Determination of p$_s$$K_a$ of analyte using proton transfer. Determination of p$_s$$K_a$ of analyte across the NMR tube is done using the κ parameter equations outlined in chapter 3. Given that the pK_a of the analyte is changing across the NMR tube due to solvent composition changes it is not possible to do linear regression as previously done. For this setup the pK_a of the analyte is calculated directly from the equation. This is done by calculating κ for any 1D spectrum using the indicator by measuring its chemical shift and solving for κ through equation 4.9:

$$\kappa = C_{indicator} \frac{\delta_{obs} - \delta_L}{\delta_H - \delta_L} \tag{4.9}$$

In this setup the concentration of the indicator, ($C_{indicator}$) needs to be known as well as the limiting chemical shift of the indicator (in our setup we made the concentration of the indicator and analyte the same which makes it possible to ignore concentration as they cancel out). This is done for each 1D slice of the 2D CSI NMR experiment. κ can also be represented from the perspective of the analyte via equation 4.10:

$$\kappa = \frac{C_{analyte}}{1 + 10^{p_sK_a - pH}} \tag{4.10}$$

Solving for p_sK_a the equation can be rearranged to obtain:

$$p_sK_a = log_{10}(\frac{C_{acid} - \kappa}{\kappa}) + pH \tag{4.11}$$

Where $C_{analyte}$ is concentration of analyte. pH of the solution throughout the different 1D spectra is determined using the indicator via the NMR modified Henderson–Hasselbalch equation:

$$pH = \log_{10}\left(\frac{\delta_{obs} - \delta_H}{\delta_L - \delta_{obs}}\right) + pK_a \tag{4.12}$$

Where δ_L and δ_H are the limiting chemical shifts of the indicator when it is fully protonated and fully deprotonated respectively. By determining pH, κ and $C_{analyte}$ the analyte's p_sK_a can be determined across the different 1D slices using equation 4.12.

The values of p_sK_a determined can then be inputted, along with its respective concentration of H_2O and dielectric constant ε into the Yasuda Shedlovsky equation:

$$p_sK_a + \log_{10}(H_2O) = \frac{A}{\varepsilon} + B \tag{4.13}$$

Where A and B are empirical constants. Linear regression analysis is then done to obtain A and B constants. p_sK_a of compound in pure H_2O is then calculated using the Yasuda Shedlovsky equation by inserting into the equation the dielectric constant of pure H_2O (80).

The method was tested on water-soluble compounds initially for method validation (figure 4.9 and 4.10) and their pure water pK_a values were determined using the method described within 0.4 units of literature pK_a values (see table 4.3 for comparison of analyte measured pK_a with literature equivalent). The uncertainty in pK_a measurements would increase as the difference between the pK_a of the analyte and the indicator increases, suggesting the use of indicators that have pK_a values that closely resemble the pK_a value of the analyte would decrease the uncertainty. Given that the pK_a of the analyte would be unknown the convenient way by which to assess the difference in pK_a is to observe whether the inflection point of the analyte or the indicator with respect to their chemical shifts are in a similar location across the NMR tube as approximately the closer they are the closer their pK_a values are. Linear fit was most closely resembling literature pK_a values when the data points were used for fitting were of $F_{DMSO} \leq 0.5$ (which is termed in the chapter as “water-rich “with data points that have an $F_{DMSO} \geq 0.5$ are termed as “DMSO-rich”). Figure 4.9 data points all were water-rich and for figure 4.10 and 4.11 some data points were water-rich and some DMSO-rich

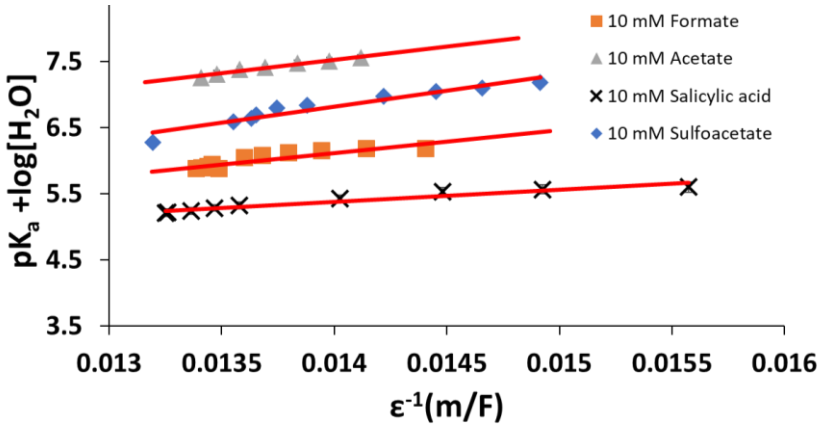


Figure 4.9. Plot of pK_a +log[H_2O] versus 1/ϵ of water-soluble analytes with respective fitting (quinine hydrochloride shown separately due to space constraint)

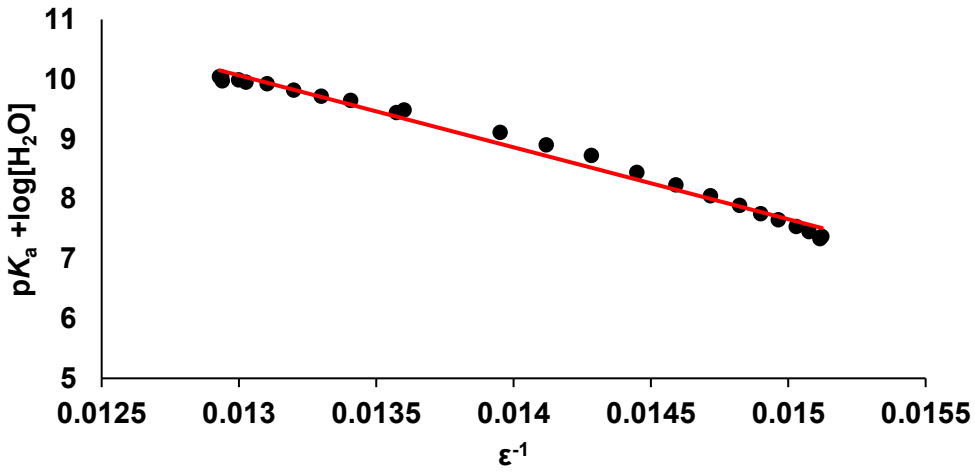


Figure 4.10 Plot of pK_a +log[H_2O] versus 1/ϵ of quinine hydrochloride

Table 4.3. Literature and fitted values of pK_a of water-soluble compounds

Analyte	Indicator	Literature pK_a	Fitted pK_a	Δ pK_a
Salicylic acid	1,2,4-triazole	3.00[156]	3.09	0.09
Formate	Sodium glycolate	3.75 [176]	3.92	0.17
Sulfoacetate	Sodium glycolate	4.07 [177]	4.37	0.30
Acetate	Sodium glycolate	4.77 [178]	5.16	0.39
Quinine hydrochloride	2-methylimidazole	8.55 [134]	8.53	0.02

When the water-insoluble analytes were studied their behaviour was remarkably different compared to the water-soluble compounds. Figure 4.11 shows how these compounds displayed a bimodal behaviour with one set of data points concentrating in a region of higher F_{DMSO} compared to the other data points (the data points with higher relative F_{DMSO} are termed ”DMSO-rich” and the ones with lower relative F_{DMSO} are termed ”water-rich”). The two domains can be represented as two lines with one for each domain. Extrapolation was done from both domains and interestingly the results obtained that were the closest with respect to literature pK_a values were the ones that were fitted on the DMSO-rich domain (see table 4.4 for comparison of analyte measured pK_a with literature equivalent). This is in stark contrast to the water-soluble compounds where fitting with respect to water-rich data points gave the results closest to literature pK_a value. This suggests the presence of aggregation dynamics in the water-rich domain that is not present in the DMSO-rich one for water insoluble compounds. The line width of naproxen, furosemide and indomethacin were investigated (see figure 4.11 b-d) and significant line broadening was being observed in the water-rich region as compared with the DMSO-rich region lending support to the idea that aggregation is happening in the water rich region.

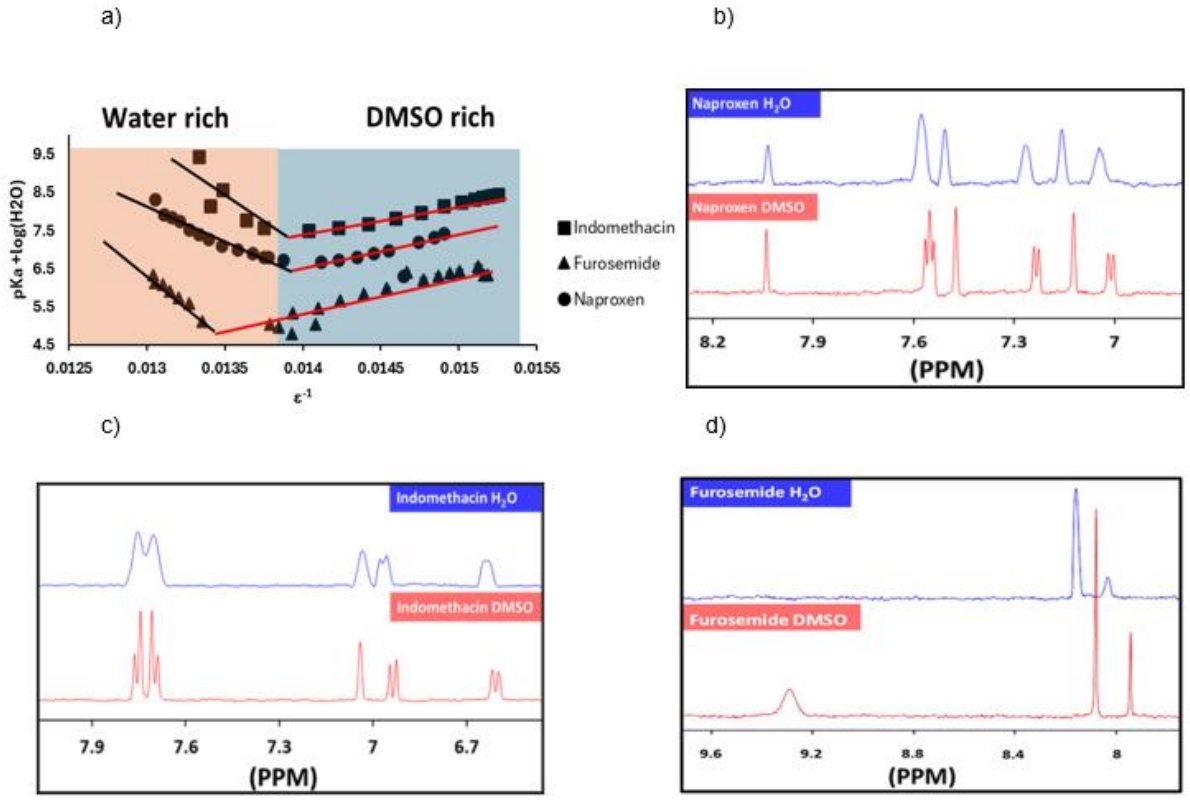


Figure 4.11. (a) Plot of pK_a +log[H_2O] versus ε^{-1} of water-insoluble analytes, (b) Line width analysis of naproxen signal in water rich domain (blue) and DMSO rich domain (red), (c) Line width analysis of indomethacin signal in water rich domain (blue) and DMSO rich domain (red), (d) Line width analysis of furosemide signal in water rich domain (blue) and DMSO rich domain (red)

Table 4.4. Literature and fitted values of pK_a of water-insoluble compounds

Analyte	Indicator	Literature pK_a	Fitted pK_a	Δ pK_a
Naproxen	Sodium formate	4.15[179]	4.11	0.04
Indomethacin	Sodium acetate	4.50 [180]	4.24	0.26
Furosemide	Sodium formate, sodium glycolate	3.34[181]	3.31	0.03

4.6 Conclusion

This chapter showcases a method of measuring pK_a of water-insoluble compounds using ^{1}H CSI NMR and a solvent gradient technique in a single shot experiment. The pK_a values obtained in these experiments can be extrapolated to pure water using the Yasuda Shedlovsky extrapolation. This method allows for a more robust and time efficient determination of pK_a water-insoluble compounds as it eliminates the issues associated with potentiometric titrations and simultaneously saves experimental time and cost. More work can be done to find concrete evidence to explain why it's better to extrapolate from DMSO rich regions for water-insoluble compounds compared to extrapolation from water rich regions. More work needs to be done to verify the reason why DMSO rich extrapolation for water-insoluble compounds is effective. Additional work can also be done to investigate how aggregation and binding change as a function of solvent composition and whether can give insights as to the self-assembly properties of compounds under different solvent mixtures.

Chapter 5: The measurement of calcium and magnesium binding via ^{1}H chemical shift imaging

5.1 Chapter summary

Binding of calcium and magnesium ions (M^{2+}) to polymers and macromolecules is a common occurrence in both chemical and biological systems. However, accurately determining the binding affinity of macromolecules for M^{2+} ions remains challenging. Conventional methods, such as potentiometric titrations and isothermal titration calorimetry, are often employed but require advanced equipment and can involve intricate measurement procedures that are not straightforward to analyse. In this chapter a new method is proposed based on ^{1}H NMR chemical shift imaging (CSI) that enables the binding affinity of compounds to be assessed in a single experiment on standard high-field NMR equipment. In this method, M^{2+} acetate salt is weighed into a standard 5 mm NMR tube and a solution of the analyte layered on top. As a result a magnesium or calcium gradient is produced across the NMR tube via diffusion. The concentration of free and bound calcium are measured at different positions along the sample by CSI. As proof of concept, the binding properties of, carboxylate functionalised nanocrystalline cellulose (CNC), carboxymethyl cellulose with 0.7 and 1.2 degree of substitution (CMC), gallate, oxalate and alginate were measured.

5.2 Author contributions

Chapter 5 incorporates work published in a paper, “Does It Bind? A Method to Determine the Affinity of Calcium and Magnesium Ions for Polymers Using ^{1}H NMR Spectroscopy ”, in *Analytical Chemistry* as an article in 2022.[102] The author contributions are as follows: Matthew Wallace conceived the work with input from Haider Hussain. Haider Hussain prepared the CNC and alginate samples and performed data analysis on them. Haider Hussain also wrote the corresponding sections of the manuscript. The ^{43}Ca and ^{25}Mg direct observation detection NMR experiments and CSI experiments done on CMC 0.7, CMC 1.2, gallate and oxalate are original work done by Haider Hussain that were not in the paper but use the same principles outlined in the paper.

5.3 Introduction

Divalent metal ions such as calcium (Ca^{2+}) and magnesium (Mg^{2+}) are central to a wide range of chemical, biological, environmental, and industrial processes.[182-184] Their ability to form coordinate bonds with negatively charged ligands makes them essential components in catalysis, molecular structure stabilization, biomineralization, and charge regulation.[185, 186] In biological systems, these cations serve as indispensable cofactors in enzymatic reactions, structural elements in membranes and cell walls, and critical signalling agents.[187, 188] For instance, calcium ions play a pivotal role in signal transduction, acting as secondary messengers that regulate various cellular activities including gene expression, exocytosis, muscle contraction, and metabolic control.[188, 189] Magnesium is also involved in over 300

enzymatic processes, including those related to DNA replication, ATP stabilization, and protein synthesis.[187, 190]

In environmental chemistry, the presence and mobility of Ca^{2+} and Mg^{2+} influence soil stability, water hardness, and contaminant transport. Water hardness, largely determined by the concentration of these ions affects not only ecosystem dynamics but also human industrial activities.[191, 192] In the context of water treatment, removal or control of divalent ions is often necessary to prevent scale formation or ensure the effectiveness of ion exchange systems.[193] Furthermore, Ca^{2+} and Mg^{2+} impact the stability and functionality of dissolved organic matter, colloids, and biofilms in natural waters.[194]

From a materials science perspective, calcium and magnesium are integral to the mechanical and physicochemical behaviour of numerous biomaterials. For example, alginate, a naturally derived polysaccharide relies on calcium-induced crosslinking to form hydrogels that are used in tissue engineering, wound dressings, and drug delivery matrices.[195, 196] The "egg box" model, describing cooperative Ca^{2+} binding between adjacent guluronate blocks, highlights how specific structural motifs and multivalent interactions enable gel formation.[197] Magnesium ions, while similar in charge, often fail to induce comparable gelation due to their smaller ionic radius, higher charge density, and stronger hydration shell, underscoring the subtleties of metal–ligand interactions.[198]

In food science and human nutrition, the binding of calcium and magnesium by dietary compounds directly affects the bioavailability of these essential minerals.[199-202] Many foods contain anti nutritional factors such as oxalic acid and phytic acid that form insoluble complexes with Ca^{2+} or Mg^{2+}, thereby inhibiting their absorption in the gastrointestinal tract. Oxalate, abundant in spinach and rhubarb, is notorious for its ability to precipitate calcium as calcium oxalate, a compound with extremely low solubility.[199, 203] This has implications not only for nutrition but also for pathology, particularly in the formation of kidney stones, where calcium oxalate is a major constituent.[204] Conversely, other dietary constituents like carboxymethyl cellulose (CMC) and gallic acid are presumed to have minimal interaction with divalent metals, yet quantitative confirmation of their binding potential remains limited in the literature.

Quantifying the binding affinity of Ca^{2+} and Mg^{2+} to various ligands is therefore a key objective in many fields. However, achieving reliable and high-resolution measurements remains challenging. Classical methods such as potentiometric titration, isothermal titration calorimetry (ITC), and atomic absorption spectroscopy (AAS) are valuable but suffer from significant drawbacks. Potentiometric titrations, for example, are prone to artefacts due to poor mixing, especially in viscous or gelling systems.[205] ITC, though thermodynamically precise, can be compromised by heat effects unrelated to binding (e.g. dilution, hydrolysis, oxidation).[85, 206] Fluorescence and colorimetric probes, while sensitive, often require specific calibration for each matrix and may introduce extraneous ligands that interfere with the system being studied.[207, 208]

To address these limitations, this study explores a Nuclear Magnetic Resonance (NMR) based technique using ^{1}H chemical shift imaging (CSI) as a means to map the binding of Ca^{2+} and Mg^{2+}. Metal acetate is introduced at the bottom of a 5 mm NMR tube, over which an aqueous analyte solution is gently layered. As the metal salt slowly dissolves and diffuses upward, a vertical concentration gradient of metal ions is established. By acquiring spatially resolved ^{1}H NMR spectra along this axis, the concentrations of metal and acetate at different heights in the tube are determined (see figure 5.1).

Two weakly binding indicator molecules glycolate and sulfoacetate serve as probes for the free metal concentration.[209] Their ^{1}H resonances exhibit predictable chemical shift changes

upon complexation with divalent cations, allowing $[M^{2+}]_f$ to be calculated at each position. Simultaneously, acetate concentration $[Ac^-]$ is measured by integrating its ^1H signal relative to an internal reference (typically DMSO or DSS). Because the total metal to acetate ratio in the diffusing region is 1:2, and because all other species are known, a mass balance equation can be constructed. This permits quantification of not only the free and indicator-bound metal but also the additional metal bound specifically to the analyte of interest.

The approach offers several key advantages. First, it is spatially resolved providing dozens of independent measurements from a single sample. Second, it avoids the need for incremental titration or separate sample preparation steps, reducing labour and minimizing sample usage.

This chapter applies the CSI-NMR method to a series of model analytes representing different types of metal-binding behaviour. These include sodium alginate, a strongly binding, gel-forming polysaccharide, citrate functionalized cellulose nanocrystals (CNC) with multiple carboxylate groups, carboxymethyl cellulose (CMC) with degrees of substitution 0.7 and 1.2: representing weakly charged, flexible polymers, gallic acid, a small phenolic molecule with potential chelation capability via carboxylic and hydroxyl groups and oxalic acid, a strongly binding small molecule known to precipitate calcium.

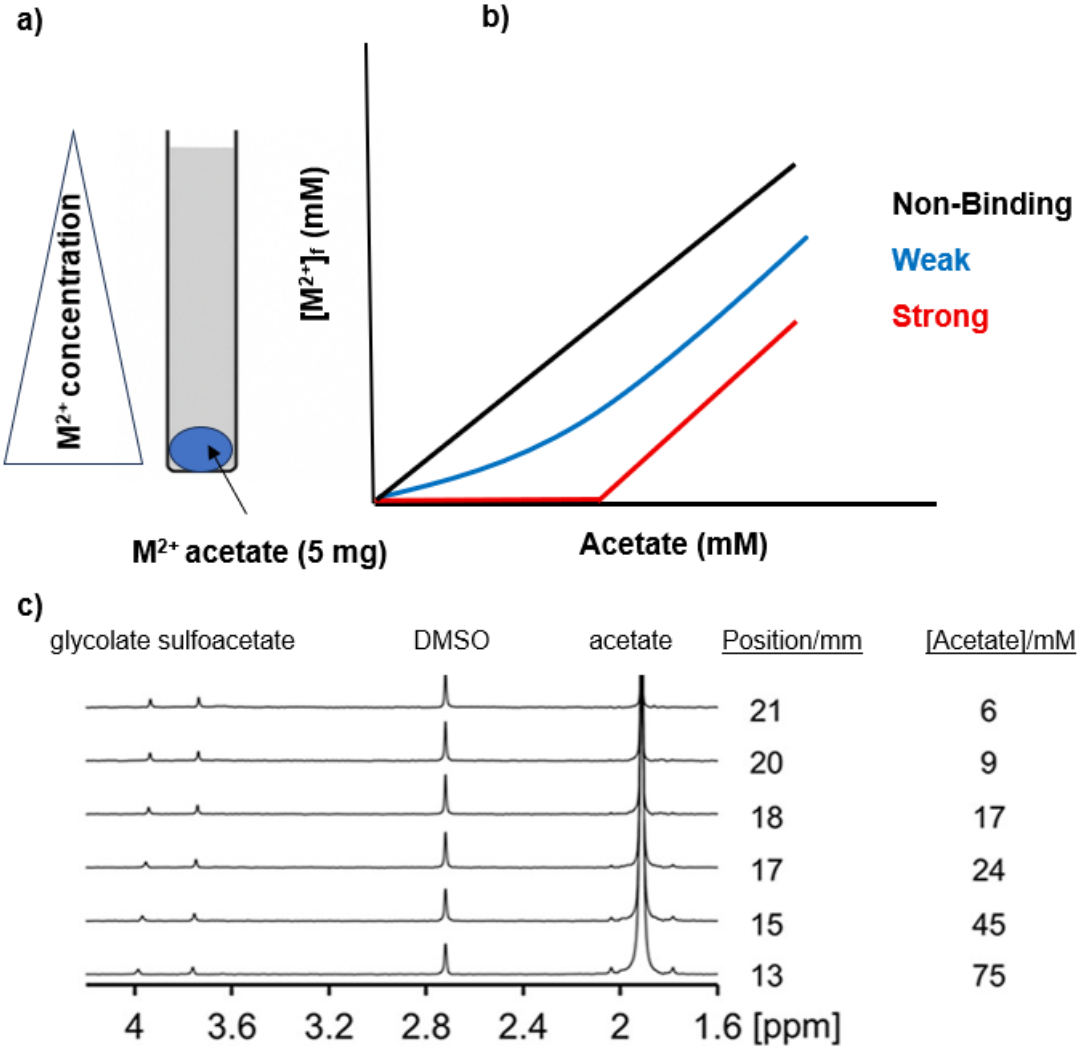


Figure 5.1. (a) A concentration gradient of M^{2+} acetate is established by layering a solution of polymer on top of solid acetate salt, (b) sketch of $[M^{2+}]_f$ versus acetate concentration for a strong, weak, and nonbinding polymer, (c) ^1H NMR spectra are recorded at different positions

along the sample using CSI techniques; $[M^{2+}]_f$ and the acetate concentration are determined from these spectra

5.4 Experimental section

Materials. All reagents were purchased from Merck or Fisher Scientific and used as received. Milli-Q water (18.2 MΩ cm) was used throughout the study. Calcium acetate and magnesium acetate tetrahydrate were the salts used to study ion binding. Glycolate and sulfoacetate (2 mM) were the indicators used to detect changes in ion binding.2-methylimidazole (2-MI, 2 mM) was the pH indicator used throughout the study to verify whether pH was too high and would generate metal hydroxide species, thus impacting calcium binding.0.01 vol % DMSO and DSS (0.2 mM) were used as the reference for all experiments. NaCl (50 mM) was also used in all experiments to account for the salt effect (see SI 5.7.6). Analytes studied were sodium alginate (16.2 mM, viscosity of 1 wt % solution in H_2O = 16 cps, mass loss on drying 13.7%), CMC 0.7 (10 mM), CMC 1.2 (10 mM), gallic acid ($C_6H_2(OH)_3CO_2H$, 10 mM), oxalic acid (HO−C(=O)−C(=O)−OH, 10 mM) and a 2 wt % stock dispersion of citrate functionalized CNC at pH 7.4 was prepared.[131] The concentration of deprotonated carboxylate groups, [COO^-], in a 1 wt % dispersion of CNC was determined as 3.1 ± 0.3 mM using the published method of Wallace *et al.*[131] CMC carboxylate concentration is likewise determined with 10 mM of CMC 0.7 having 6.71 mM of [COO^-] and 10 mM of CMC 1.2 having 8.76 mM of, [COO^-] (see appendix E.4). To prepare gradients of M^{2+} acetate for analysis by CSI, 4–5 mg of solid calcium acetate hydrate or magnesium acetate tetrahydrate was loaded into the tip of a 9″ Pasteur pipette by pressing into the solid salt. The salt was then transferred from the tip to the base of a 5 mm NMR tube (Wilmad 528-PP). Four, 2 mm diameter glass beads were placed on top of the acetate salt. Prior to use, the beads were rinsed with ethanol and water and dried. The solutions, prepared as above, were carefully layered on top of the glass beads to a height of 40 mm from the base of the NMR tube with a 9″ Pasteur pipette. Homogeneous samples of the analyte and M^{2+} were prepared directly in NMR tubes by combining a stock solution of M^{2+} acetate with the analyte and additives listed above. The concentration of the M^{2+} acetate stock was verified by integration of the acetate resonance against 0.5 M potassium hydrogen phthalate in D_2O. With the exception of the CNC samples, $[M^{2+}]_{tot}$ was based on the volume of acetate solution added and is assumed accurate to 3%. The CNC samples were prepared by addition of aliquots (<10 µL) of M^{2+} acetate solution to the CNC to conserve material and enable several values of $[M^{2+}]_{tot}$ to be measured with the same sample. $[M^{2+}]_{tot}$ in the CNC samples was determined from the ^1H integral of acetate. The calcium alginate and CNC samples were gently centrifuged (<500 rpm) on a Hettich 1011 hand centrifuge to drive solid material to the lower region of the tube.0.5 wt % CNC samples in hard and soft water were prepared by combining CNC stock with tap and Milli-Q water in different proportions. No glycolate, sulfoacetate, NaCl, or DMSO was included in these samples. All direct detection experiments had 0.1 M calcium acetate, 50 mM NaCl and 50 mM analyte.

NMR. All experiments were performed off lock in 100% H_2O at 298 K on a Bruker 500 MHz AVANCE III spectrometer with a 5 mm broad band observe probe. ^1H chemical shift images were acquired using a gradient-phase encoding sequence based on Trigo-Mouriño *et al.*[51] The sequence incorporated a double echo excitation sculpting component (Bruker library zgesgp) for water suppression (see appendix A1). Gaussian inversion pulses of 4 ms duration and 300 Hz peak power were applied to the H_2O resonance. The phase encoding gradient pulse (172 µs) was in the form of a smoothed square and was ramped from −18.8 to 18.8 G/cm in 64 steps, giving a theoretical spatial resolution of 0.41 mm along the *z*-axis. Four

scans were acquired at each gradient increment, with a signal acquisition time of 2 s and relaxation delay of 2 s. A spoil gradient (27 G/cm) was employed at the end of the acquisition period to destroy any remaining transverse magnetization.16 dummy scans were run prior to acquisition, giving a total acquisition time of 19 min ^{1}H spectra were acquired in 32 scans, 32 dummy scans, using the same excitation sculpting sequence and timings used for CSI but without the phase encoding gradient. For the 1D direct detection experiments A total of 32 scans and 32 dummy scans were taken. The total time taken for experiments were 4 minutes 21 seconds. The line broadening factor was 0.30 Hz. The pulse width had a duration of 10 μs and a respective power of 22.677 Watt. The relaxation delay was 2 seconds. The number of data points were 32 K. The spectral width was 16.02 ppm.

Data analysis. Prior to preparing M^{2+} acetate gradients, a ^{1}H spectrum of the samples was acquired to measure the ^{1}H chemical shifts of glycolate and sulfoacetate in the absence of M^{2+}.[209] The lower detection limits of $[M^{2+}]_f$ are 0.4 and 0.8 mM for Ca and Mg, respectively, below which $[M^{2+}]_f$ can be taken as zero within the uncertainty of the measurement.[102] CSI data sets were processed in a phase-sensitive mode with 32K points and an exponential line broadening factor of 3 Hz.1D spectra were processed with 64K points. Each row (64) of the CSI data set was automatically phase and baseline corrected using an automation script written in house. The ^{1}H chemical shifts of DSS, DMSO, glycolate, sulfoacetate, and 2MI and the ^{1}H integrals of DMSO and acetate were extracted from each row using a custom script $[Ac^-]$ was obtained as $[Ac^-] = kA/R$, where A and R are the integrals of acetate and reference (DMSO) where $k = 2.92$. All spectra were referenced to DMSO (2.72 ppm) or DSS (0.00 ppm). Appendix 5 provides the calculations of experimental uncertainty for the concentrations of bound and free M^{2+}, along with additional experimental details.

5.5 Results and Discussion

Determination of $[M^{2+}]_f$ by ^{1}H NMR. Measurement of Free Ca^{2+} and Mg^{2+} concentrations in the 1–100 mM range is performed via ^{1}H NMR by monitoring the chemical shifts of two low-affinity ligands: glycolate and sulfoacetate. These reporter ligands exhibit fast exchange with M^{2+} on the NMR timescale, and their chemical shifts are sensitive to the local concentration of free M^{2+} ions. Their observed chemical shift δ_{obs} is given by:

$$\delta_{obs} = \frac{[M^{2+}]_f K_M \delta_M + \delta_L}{1 + [M^{2+}]_f K_M} \tag{5.1}$$

where δ_M and δ_L are the limiting chemical shifts of the fully complexed and free ligands, respectively, and $[M^{2+}]_f$ is the free concentration of divalent metal ions. K_M is the complexation constant of a ligand of negative charge, n, given by

$$K_M = \frac{[ML^{2-n}]}{[M^{2+}]_f[L^{-n}]} \tag{5.2}$$

Where L is ligand and ML is the metal ligand complex. Equation 5.1 can be reformulated to get to make $[M^{2+}]_f$ the subject of the equation:

$$[M^{2+}]_f = \frac{\delta_L - \delta_{obs}}{K_M(\delta_{obs} - \delta_M)} \tag{5.3}$$

K_M is apparent (measured) complexation constant that includes the contribution of electrostatic interactions associated with binding a ligand of charge n in a given ionic environment. It can be related to K_0 which is the thermodynamic (intrinsic) binding constant, defined in terms of activities and independent of electrostatic work via equation 5.4:

$$K_0 = \frac{[ML^{n-2}]\gamma_{ML^{2-n}}}{\gamma_{M^{2+}}\gamma_{L^{n-}}[M^{2+}]_f[L^{-n}]} \quad (5.4)$$

where γ represents the activity coefficients of the different species. If the approximation is made that the activity coefficient of a species depends solely on the magnitude of its charge then K_M and K_0 may be interconverted using the relation:[210-212]

$$K_M = K_0\gamma_2 \quad (5.5)$$

for a ligand with n = 1. γ_2 is the activity coefficient of a divalent ion. For a ligand with n = 2:

$$K_M = K_0\gamma_2^2 \quad (5.6)$$

The activity coefficient can be determined without knowledge of the protonation states of all ligands by combining equation 5.3, 5.5 and 5.6 to get:

$$\gamma_2 = \frac{[K_1(\delta_{L2} - \delta_2)(\delta_1 - \delta_{M1})]}{[K_2(\delta_2 - \delta_{M2})(\delta_{L1} - \delta_1)]} \quad (5.7)$$

where the subscripts denote the charges of the ligands and Kthe thermodynamic binding constants thus by combining chemical shift data from both glycolate (charge −1) and sulfoacetate (charge −2), one can back calculate the free metal ion concentration and activity coefficient (γ_2) without prior knowledge of the full ionic composition of the sample.

Measuring binding to analytes Using M^{2+} Gradients. In realistic systems, the presence of additional cations (e.g., Na^+, K^+) alters the ionic strength and may interfere with indicator binding. In the presence of background electrolytes, the diffusion of divalent metal ions (M^{2+}) and their counterions, such as acetate, does not necessarily proceed at identical rates. This phenomenon, termed the “salt effect”, arises because co-diffusing ions experience different interactions with other ionic species in the solution. As a result, the expected 2:1 stoichiometric relationship between acetate and M^{2+} which would normally be preserved in an ideal binary salt system is disrupted. This leads to a measurable discrepancy between the local concentrations of acetate and free M^{2+} along a diffusion gradient. In the context of NMR chemical shift imaging, this effect must be accounted for in order to accurately determine metal binding. By introducing a correction factor that models the salt effect (denoted as N), it becomes possible to separate this physical diffusion imbalance from true chemical association of M^{2+} with polymer or ligand binding sites. The correction enables accurate extraction of free and bound metal ion concentrations even in complex media containing multiple ionic species. The way this is done is by representing the concentration of acetate [Ac^-] via equation 5.8:[102]

$$0.5[Ac] = [M^{2+}]_f + [M^{2+}]_L + B + N \quad (5.8)$$

Where $[M^{2+}]_L$ is concentration of divalent ions bound to the indicator ligands, B is the concentration of divalent ions bound to analyte. $[M^{2+}]_L$ is calculated using equation 5.9:[209]

$$[M^{2+}]_L = \frac{[L]_{total}\gamma_2^n K_0[M^{2+}]_f}{1 + \gamma_2^n K_0[M^{2+}]_f} \quad (5.9)$$

Where $[L]_{total}$ is the total concentration of the ligand. N may be calculated by assuming that the diffusion of M^{2+} and acetate follows separate Gaussian models as per the work of Wallace *et al:*[102]

$$N = \left(\frac{[Ac]}{2}\right) * \left[1 - \sqrt{\frac{D_{Ac}}{D_M, NaCl}} * \exp\left[\left(\frac{(Z - h)^2}{4t}\right) * \left(\frac{1}{D_{Ac}} - \frac{1}{D_{M,NaCl}}\right)\right]\right] \quad (5.10)$$

where $D_{M,NaCl}$ and D_{Ac} are the diffusion coefficients of the $M^{2+}$and acetate ions, respectively, measured in 50 mM NaCl in the absence of the polymer. D_{Ac} was obtained as 1×10^{-9} m^2 s^{-1},and $D_{M,NaCl}$ was obtained as 9.3×10^{-10} and 9.0×10^{-10} m^2 s^{-1} for Ca^{2+} and Mg^{2+}, respectively.[102] Z is the vertical distance from the absolute base of the NMR tube. t is the time elapsed since preparation of the sample. h is the height of the solid acetate layer when prepared (2 mm). The stated diffusion coefficients are used to calculate N in all experiments as 50 mM NaCl is used as a constant background medium in this work. B may be obtained by rearrangement of equation 5.8:

$$B = 0.5[Ac] - [M^{2+}]_f - [M^{2+}]_L - N \quad (5.11)$$

Using the chemical shift imaging (CSI) method, we obtained spatially resolved profiles of $[M^{2+}]_f$ and $[Ac^-]$ across each sample. Plotting $[M^{2+}]_f$ against $[Ac^-]$ (or equivalently examining *B* vs $[Ac^-]$) provides a visual indicator of the analyte's binding strength. For a strongly binding analyte, the free metal concentration remains near zero (and thus *B* rises nearly in step with total metal diffusing in) until the analyte's binding sites are saturated; only then does $[M^{2+}]_f$ begin to increase appreciably. In contrast, a weakly binding analyte shows a moderate increase in $[M^{2+}]_f$ even at low $[Ac^-]$, indicating that a substantial fraction of M^{2+} remains free (or bound only to low-affinity ligands) rather than associating with the analyte.[102] A non-binding system yields $[M^{2+}]_f$ rising proportionally with $[Ac^-]$ (essentially following the 1:2 diffusion ratio) and negligible *B*.[102] This method was applied to a series of analytes – alginate, citrate functionalized cellulose nanocrystals (CNC), carboxymethyl cellulose (CMC, with two degrees of substitution), gallic acid (gallate), and oxalic acid (oxalate) chosen to represent a range from strongly binding polyelectrolytes to weak or non-binding molecules. The results are discussed below by analyte. Throughout, Ca^{2+} and Mg^{2+} behaviour will be compared, as these two alkaline earth cations often exhibit markedly different binding affinities due to their distinct ionic properties. Notably, Ca^{2+} has a larger ionic radius (approx.1.00 Å six-coordinate) than Mg^{2+} (~0.72 Å) and a much lower hydration enthalpy (−1579 kJ/mol vs. −1926 kJ/mol).[213, 214] Consequently, Ca^{2+} is a "softer" divalent cation that is more easily shed of its hydration shell and more polarizable, tending to form more stable coordinate complexes, whereas the smaller Mg^{2+} (a "harder" cation) remains more tightly bound to water and often interacts with ligands more weakly.

Alginate (Sodium Alginate). Alginate is a natural polyanionic polysaccharide composed of β-D-mannuronate (M) and α-L-guluronate (G) residues. Divalent cations are known to crosslink guluronate-rich blocks of alginate chains, leading to gelation via the classic "egg box" model examined binding of Ca^{2+} and Mg^{2+} to sodium alginate at two concentrations (2 mg/mL and 4 mg/mL) using our CSI method.[215] The results (Figure 5.2-5.5) show a striking difference in behaviour between the two cations. For Ca^{2+}, a strong interaction with alginate is evident: the $[Ca^{2+}]_f$ remains near zero (within detection limits) over a substantial initial portion of the acetate gradient, and the binding term *B* increases essentially linearly with $[Ac^-]$. Only after *B* reaches a plateau corresponding to the stoichiometric capacity of the alginate do free Ca^{2+} ions begin to appear in solution. Quantitatively, 4 mg/mL alginate could sequester ~8 mM Ca^{2+} before saturation, while 2 mg/mL alginate bound ~4 mM Ca^{2+} (these values correspond to the total carboxylate site concentrations for each alginate dose). In other words, Ca^{2+} binding to alginate is essentially stoichiometric (one Ca^{2+} per alginate carboxylate group) until the alginate's sites are filled. In contrast, Mg^{2+} showed much weaker binding to alginate. Even at the outset, a measurable free Mg^{2+} concentration $[Mg^{2+}]_f$ was present and rose continuously

with increasing [Ac^-]. The binding capacity *B* for Mg^{2+} remained far below the theoretical site capacity; in fact, the alginate never became saturated with Mg^{2+} within the concentration range of the experiment. This indicates that alginate's affinity for Mg^{2+} is low. Most Mg^{2+} ions prefer to stay as free (hydrated) ions or bound only to the acetate/small ligands, rather than being strongly captured by alginate.

The results obtained for alginate are in line with its known gelation behaviour and binding thermodynamics. The fact that Ca^{2+} is tightly sequestered by alginate while Mg^{2+} is not explains why Ca^{2+} induces gelation of alginate (formation of crosslinked alginate networks), whereas Mg^{2+} cannot form a firm gel under similar conditions. In the "egg box" model of calcium alginate, Ca^{2+} ions fit into coordination pockets formed by G-blocks on adjacent alginate chains, bridging them like eggs in an egg crate.[215] Mg^{2+} ions, being smaller and more strongly hydrated, do not fit as well into these sites and tend to remain more diffusely associated (e.g. condensing along the negatively charged alginate without forming stable crosslinks).[102] This qualitative difference has been noted in literature: Ca^{2+} produces "egg box" junction zones with alginate (yielding hydrogels), while Mg^{2+} leads only to a loose association often described in terms of counterion condensation rather than true crosslinking. The average axial charge spacing along the alginate chain (~4.7 Å between carboxylate groups) is such that a divalent cation's electrostatic binding can be partly delocalized; for a high charge density ion like Mg^{2+}, this typically results in a diffuse binding mode (Manning condensation) rather than site-specific chelation.[216] Ca^{2+} being more polarizable and able to adopt higher coordination numbers, can overcome the charge repulsion and bind to alginate blocks, hence the observed strong, site-specific uptake.

The practical implications of Ca^{2+} vs. Mg^{2+} binding to alginate are significant in biomaterials and biotechnology. Calcium alginate gels are widely used for cell encapsulation and tissue engineering; in those applications the free Ca^{2+} concentration in the gel (as opposed to total Ca) is a critical parameter that affects cell viability and function. Our results confirm that in calcium alginate gels, $[Ca^{2+}]_f$ will remain very low until alginate's sites are saturated, meaning most Ca is bound in the crosslinked network. This is beneficial for gel stability but also means any cells embedded in a Calcium alginate matrix experience a low Ca^{2+} environment unless excess Ca is present. In contrast, the ineffectiveness of Mg^{2+} in binding alginate explains why Mg is rarely used to form alginate gels – even at high Mg^{2+} concentrations the alginate remains largely in solution or forms only weak, non-self-supporting structures. This weak interaction could, however, still contribute to osmotic and charge screening effects; for instance, Mg^{2+} may condense onto alginate enough to neutralize some charge (reducing electrostatic repulsion) without inducing the cooperatively crosslinked state that Ca produces.

In summary, sodium alginate exhibits strong, stoichiometric binding of Ca^{2+} (with effectively one Ca^{2+} bound per carboxylate group until gelation capacity is reached) but much weaker binding of Mg^{2+} consistent with the well-known fact that Ca^{2+} triggers alginate gelation whereas Mg^{2+} does not. These observations align with previous studies of alginate–divalent ion interactions and validate that our CSI method can distinguish high affinity vs. low affinity binding in situ.[102]

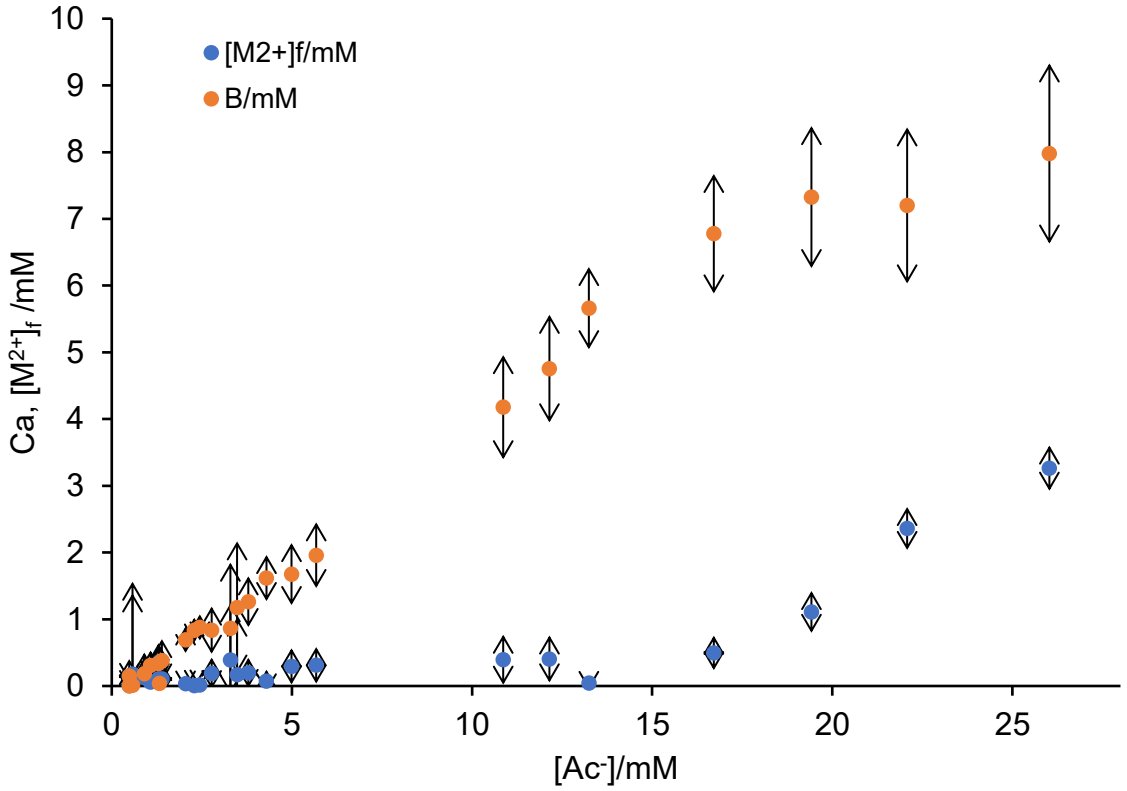


Figure 5.2. Plot of bound (orange) and free (blue) calcium versus [Ac⁻] concentration of 4 mg/mL alginate

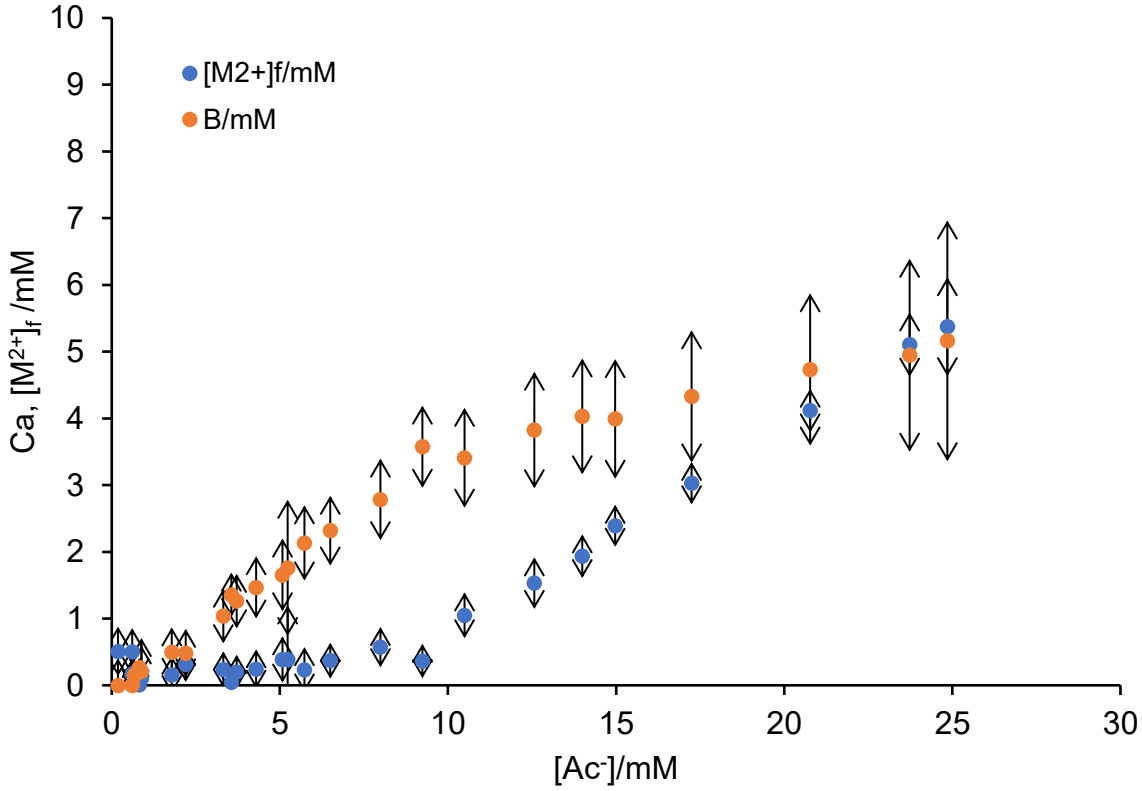


Figure 5.3. Plot of bound (orange) and free (blue) calcium versus [Ac⁻] concentration of 2 mg/mL alginate

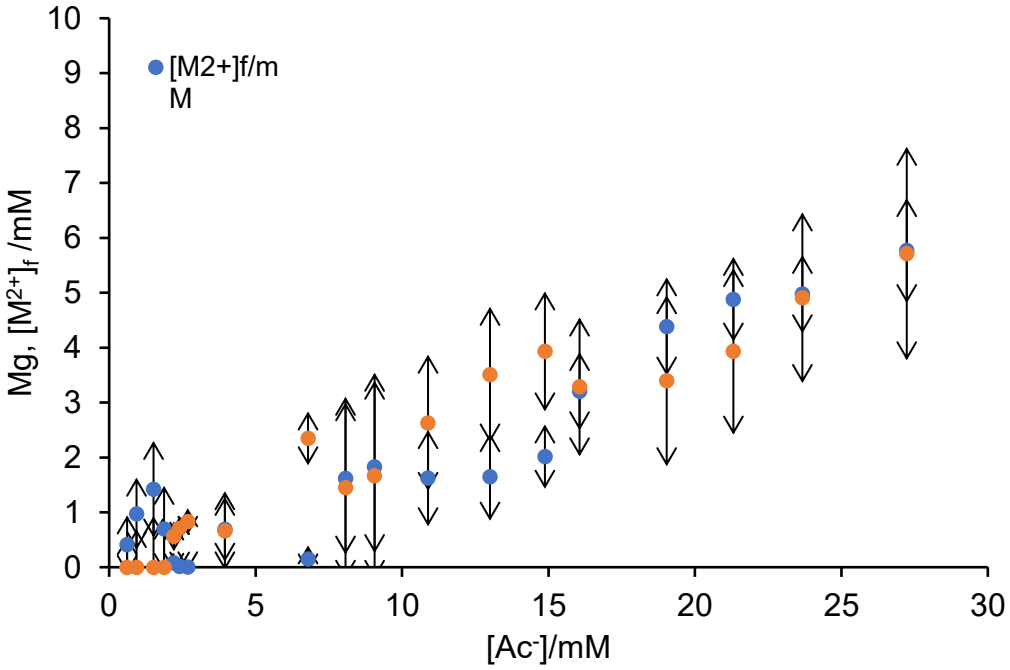


Figure 5.4. Plot of bound (orange) and free (blue) magnesium versus [Ac^-] concentration of 4 mg/mL alginate

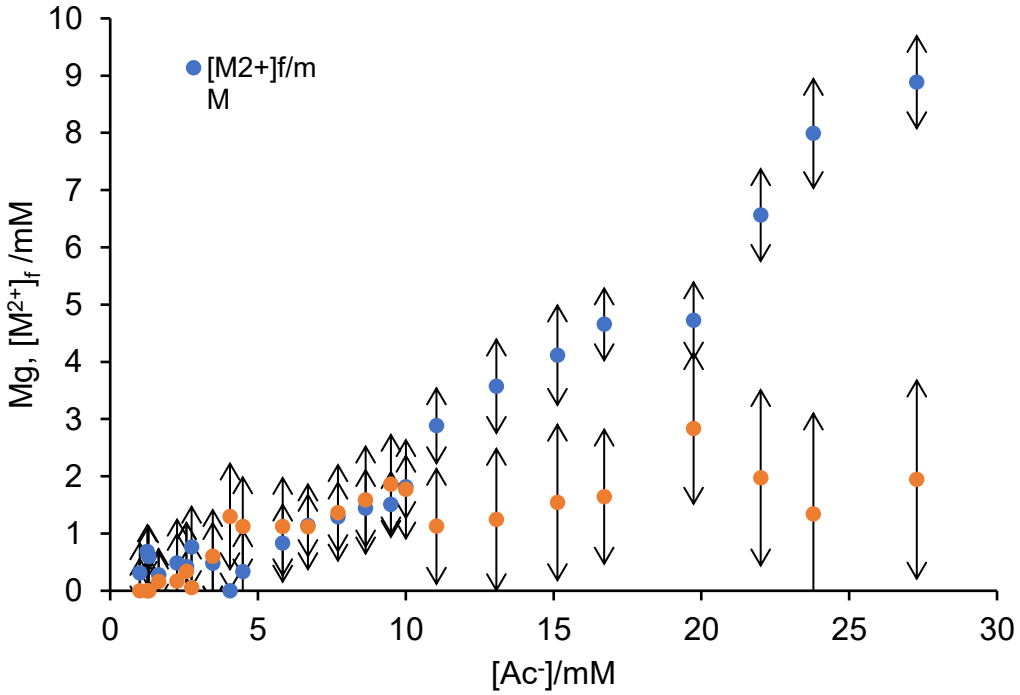


Figure 5.5. Plot of bound (orange) and free (blue) magnesium versus [Ac^-] concentration of 2 mg/mL alginate

Citrate Functionalized Cellulose Nanocrystals (CNC)

Cellulose nanocrystals (CNCs) functionalized with carboxylate groups (e.g. via surface grafting of citrate) represent another class of polyanionic substrate. In the experiments we used CNCs with a citrate functionalization such that a 1% w/v dispersion contains roughly 3 mM COO^- groups (thus our 2% w/v sample had ~6 mM carboxylate sites).[209] CNCs are high aspect ratio rod like nanoparticles, and divalent cations can bind to their negatively charged surfaces, potentially leading to aggregation (since a single Ca^{2+} can bridge two $–COO^-$ groups on different CNC rods, causing flocculation). This system is of interest for applications in colloids, nanocomposites, and water treatment.[217, 218] 2% CNC dispersion was analysed with this method (Figure 5.6-5.9). The data show that both Ca^{2+} and Mg^{2+} bind strongly to the citrate CNC, albeit with some differences in the details.

In the CSI profiles, the initial portion of the acetate gradient showed essentially no free Ca^{2+} or Mg^{2+} present virtually all diffusing $M^{2+}$was taken up by the CNC until a certain threshold was reached. For Ca^{2+}, $[Ca^{2+}]_f$began to rise noticeably only after *B* had approached ~6 mM, roughly the total carboxylate content of the CNC (matching the expectation of one Ca^{2+} per COO^- at saturation). Magnesium showed a very similar trend: $[Mg^{2+}]_f$ remained near zero until a comparable bound fraction (~6 mM) was achieved, and then free Mg^{2+} began to appear. This indicates that, within experimental uncertainty, Ca^{2+} and Mg^{2+} have comparable capacity to bind the CNC surface sites, at least in terms of how many ions can be bound. Notably, the affinity still differs. Ca^{2+} likely reaches that saturation a bit sooner (at a slightly lower $[Ac^-]$) than Mg^{2+}, but the difference is not large. Both cations exhibit what can be termed strong binding to CNC: neither is freely mobile in solution until the CNC's binding sites are largely occupied.

A dramatic consequence of this binding is CNC aggregation. Visually, it was observed that the originally stable CNC suspension became turbid and then formed flocs as M^{2+} diffused in, especially once the bound fraction approached saturation. In homogeneous titration tests, it was found that adding Ca^{2+} or Mg^{2+} at concentrations above a few millimolar caused the CNC dispersion to turn opaque and form visible precipitates. For example, at $[M^{2+}]_f \approx 3$ mM, both calcium and magnesium caused significant aggregation (loss of colloidal stability). This behaviour indicates that each divalent ion can neutralize and bridge the negatively charged CNC rods efficiently, collapsing the electrostatic repulsion that keeps them dispersed. The observation that Mg^{2+} induces aggregation at a concentration comparable to that of Ca^{2+} is noteworthy, given that Mg^{2+} is generally considered to exhibit weaker individual binding interactions with carboxylated surfaces. Nevertheless, aggregation was observed once the concentration of Mg^{2+} reached a level sufficient to significantly reduce the effective surface charge of the CNCs. This behaviour is consistent with previous studies on colloidal nanocellulose systems, which have shown that divalent cations promote aggregation at substantially lower concentrations than monovalent cations due to charge neutralisation, ion–ion correlation effects, and cation-mediated bridging between negatively charged sites.[219] In agreement with this general behaviour, NaCl at a background concentration of 50 mM did not induce CNC precipitation, whereas the addition of only a few millimolar calcium acetate or magnesium acetate resulted in aggregation.

The CSI data provide a quantitative underpinning for this: essentially all Ca^{2+} or Mg^{2+} introduced goes into a bound/bridging state until the point of incipient aggregation. Although both cations bind, Ca^{2+} may still have a slight edge in binding strength. It was observed in the spectra (Figure 5.8) that the proton signals of the CNC disappear from the NMR spectrum at a certain point in the titration, especially with Ca^{2+}. This disappearance is attributed to the formation of large or solid complexes (CNC aggregates) that are NMR invisible due to

broadening with Ca^{2+}. the CNC signals vanished at a somewhat lower total ion concentration compared to Mg^{2+} suggesting Ca induced aggregation earlier (i.e. slightly stronger binding, causing more rapid loss of NMR visibility). Nonetheless, Mg^{2+} also eventually caused loss of signal at higher concentrations, indicating that both can form insoluble CNC complexes. The end result for both cations was an essentially complete precipitation of CNC when M^{2+} was present in excess of the binding capacity. From an applications standpoint, the strong binding of Ca^{2+} and Mg^{2+} to carboxylated CNC has a dual significance. On one hand, it means CNCs can be used to scavenge these ions from solutions (for instance, in water treatment, carboxyl rich nanocellulose has been proposed as a bio-based flocculant or ion adsorbent). The fact that even Mg^{2+} which is relatively hard to bind can be substantially removed by CNC indicates a high capacity and affinity of the functionalized surface.

On the other hand, in any formulation where CNC is dispersed in a matrix containing divalent ions (e.g. a food colloid or a biomedical fluid), one must be mindful that even trace Ca/Mg hardness could cause the CNC to aggregate or gel. For example, our group observed that a 0.5% citrate CNC dispersion remained stable in "soft" water (low Ca^{2+} content) but rapidly formed aggregates in "hard" water (~2 mM Ca^{2+} as $CaCO_3$).[102] This susceptibility to hard water could limit the storage and handling of CNC based formulations or conversely be exploited intentionally for triggered gelation. In any case, the CNC results reinforce the idea that Ca^{2+} and even Mg^{2+} can strongly bind to and bridge polyanionic surfaces, in accordance with classical colloid science. [219]

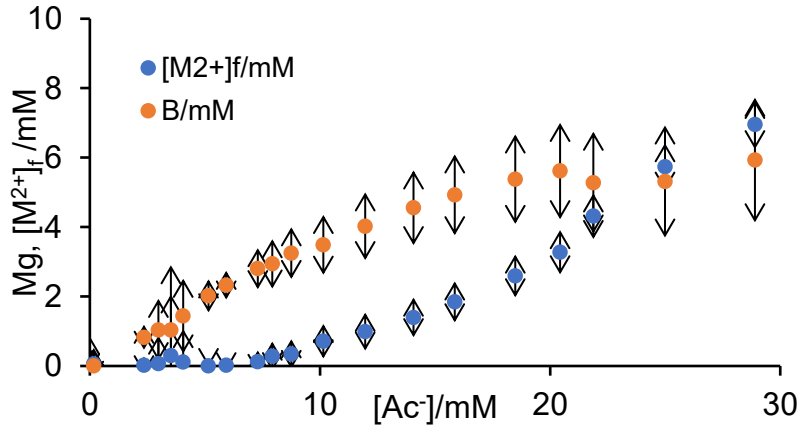


Figure 5.6. Plot of bound (orange) and free (blue) magnesium versus [Ac^-] concentration of 2 wt % citrate CNC

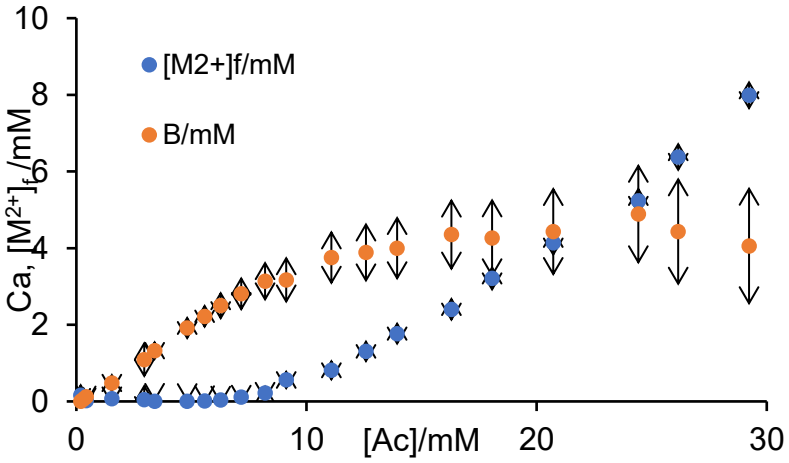


Figure 5.7. Plot of bound (orange) and free (blue) calcium versus [Ac^-] concentration of 2 wt % citrate CNC

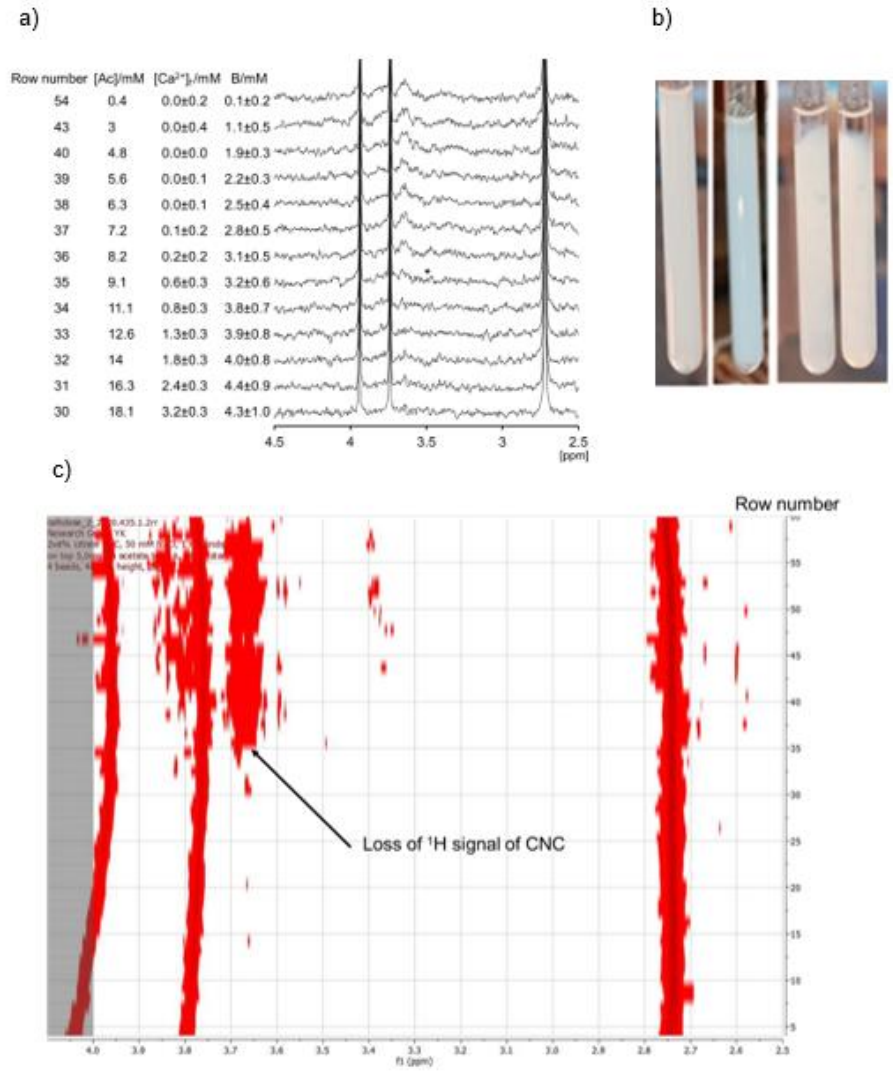


Figure 5.8. (a) rows of CSI dataset for diffusion of Ca^{2+} acetate into 2 wt% citrate functionalised CNC, (b) photographs of homogenous CNC samples with Ca^{2+} From left to right, $[Ca^{2+}]$total: 0, 1.1, 3.7, and 5.0 mM , (c) 2D plot (lower) of CSI dataset for diffusion of Ca^{2+} acetate into 2 wt% citrate functionalised CNC showing loss of signal of CNC due to aggregation

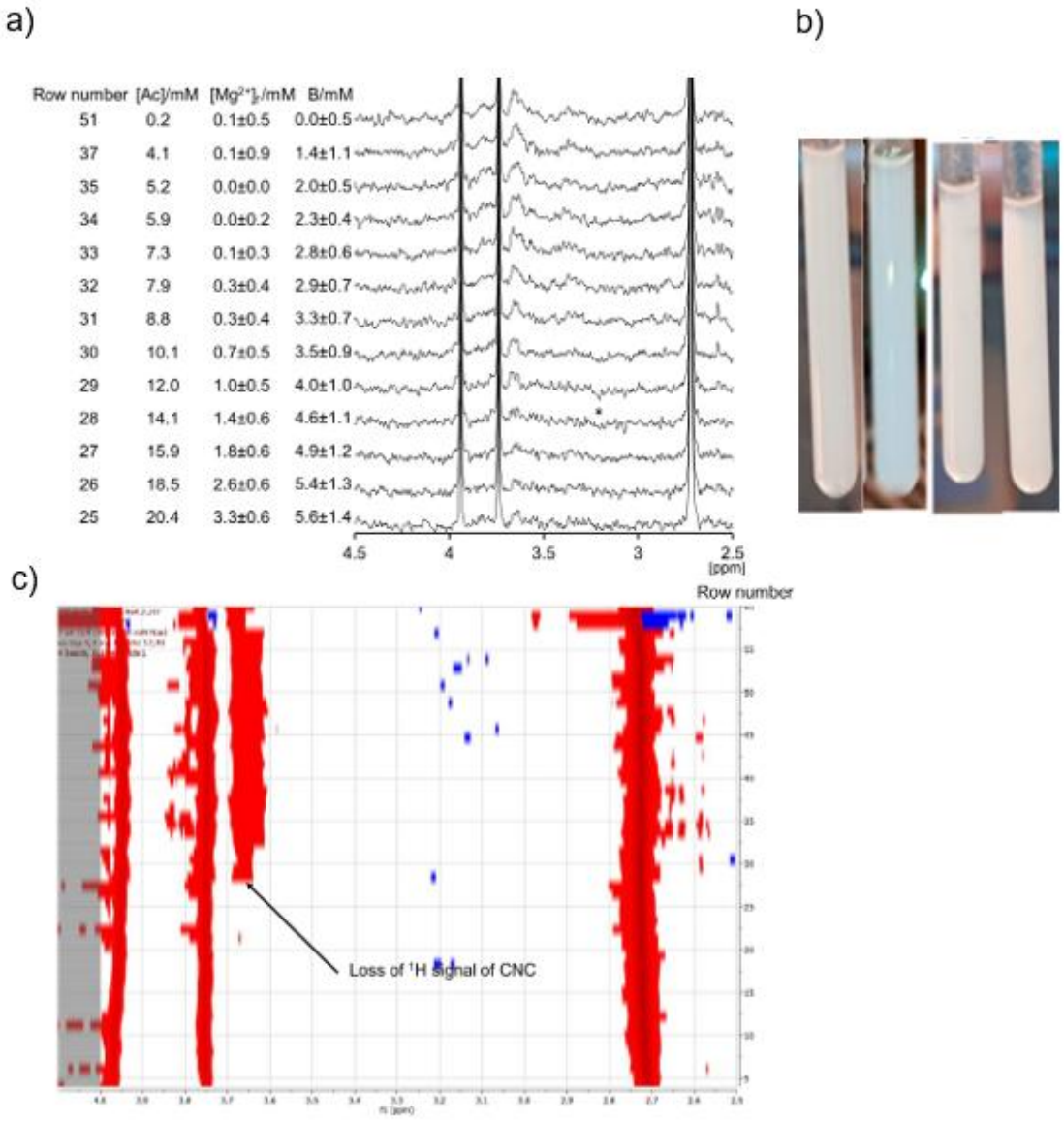


Figure 5.9. a) rows of CSI dataset for diffusion of magnesium acetate into 2 wt% citrate functionalised CNC, (b) photographs of homogenous CNC samples with Mg^{2+} From left to right, $[Mg^{2+}]$tot: 0, 1.1, 3.2, and 6.5 mM, (c) 2D plot (lower) of CSI dataset for diffusion of magnesium acetate into 2 wt% citrate functionalised CNC showing loss of signal of CNC due to aggregation

Carboxymethyl Cellulose (CMC)

Carboxymethyl cellulose is a water-soluble anionic polysaccharide derivative, where some of the hydroxyls on the glucose ring are substituted with CH_2–COO^- groups. We studied two CMC samples with degrees of substitution (DS) of 0.7 and 1.2, representing a medium range of carboxyl content (0.7 and 1.2 carboxyls per glucose unit, respectively). CMC is widely used as a food additive (thickener, stabilizer), in pharmaceuticals, and in other industries, often in environments that contain divalent ions.[220, 221] The efficacy of CMC (e.g. its viscosity or film-forming properties) can potentially be affected by interactions with calcium or magnesium, since binding of these cations could shield the charges and induce conformational or associative changes in the polymer.[222] For instance, Ca^{2+} induced aggregation of CMC chains

could increase solution viscosity or cause phase separation under some conditions. However, the CSI analysis indicates that CMC binds Ca^{2+} and Mg^{2+} only very weakly, especially in comparison to alginate or CNC.

Figure 5.10 displays the profiles of bound vs. free ion as a function of [Ac^-] for CMC. In both cases, the curves lie close to the behaviour of a non-binding system: at low to moderate acetate concentrations, essentially no significant *B* is detected the free metal ion concentration rises nearly proportionally with the amount of diffused salt. Only at very high added M^{2+} (approaching the upper limit of our experimental range, [Ac^-] > 100 mM, corresponding to $[M^{2+}]_f$ > ~50 mM) do we start to see a slight deviation indicating some calcium uptake by CMC . Statistically significant Ca^{2+} binding was only observed after ~100 mM acetate had diffused, and even then the magnitude of *B* was small (on the order of 1–2 mM bound out of 50 mM added). The Mg^{2+} profiles were similar or even weaker – essentially no measurable binding across the range, within error. In other words, both CMC samples behaved as very weak binders, and for practical purposes one can say Ca^{2+}/Mg^{2+} do not interact strongly with CMC in dilute solution. This was corroborated by visual observation: neither Ca nor Mg caused any precipitate or gel formation when added to CMC solutions, even up to high concentrations. The solutions remained clear, and no change in viscosity or appearance was evident, confirming the lack of extensive crosslinking or aggregation under these conditions.

Interestingly, increasing the degree of substitution from 0.7 to 1.2 had minimal effect on the binding behaviour. One might expect that CMC with more carboxylate groups (DS 1.2) would bind more Ca^{2+} simply due to higher site density, but our data showed that CMC 1.2 was not significantly different from CMC 0.7 in terms of Ca^{2+}/Mg^{2+} uptake. The *B* vs [Ac^-] curves for the two DS nearly overlap, and both indicate negligible binding. A possible explanation is that the distribution and spacing of the carboxymethyl groups on CMC limit multivalent binding. Unlike alginate, where blocks of adjacent carboxylated monomers (G-blocks) create well defined chelation sites for Ca^{2+}, CMC's carboxyls are substituents that may be irregularly spaced and partly hindered by the glucose ring. The polymer adopts a flexible coil conformation in water, and Ca^{2+}may not be able to simultaneously coordinate to two carboxyls on the same chain unless they are very close. If the average distance between carboxyls on CMC is larger or more random (especially for DS ~1), a Ca^{2+} ion might bind only electrostatically to one site at a time, which is a weak interaction comparable to Ca^{2+}binding to simple monocarboxylates. Essentially, the carboxyl groups on CMC serve to make the polymer highly water soluble and slightly anionic, but do not create the kind of multidentate pockets needed for strong chelation of Ca^{2+}/Mg^{2+}. This interpretation is supported by a recent study by Arumughan *et al.* (2022), who found that Ca^{2+} causes only modest changes in CMC chain dimensions (slight interchain association) at much higher concentrations and does not precipitate CMC out of solution. Our results concur: CMC remains essentially non crosslinked in the presence of Ca^{2+} or Mg^{2+} up to at least tens of millimolar of salt.[223]

From an applied perspective, the weak binding of CMC to dietary calcium or magnesium is actually a desirable feature. CMC is used as a food additive (E466) in many products, and one would not want it to sequester essential Ca^{2+}/Mg^{2+} ions in the diet. Our findings suggest that CMC (at typical usage levels) will not significantly interfere with the bioavailability of calcium or magnesium – it does not remove these ions from solution or form insoluble complexes with them. This supports the general assumption of CMC's physiological inertness with respect to mineral nutrients. Even at higher concentrations (e.g. in pharmaceutical formulations or eye drops containing CMC), the presence of divalent cations is unlikely to cause anything more than a slight increase in viscosity due to minor ionic crosslinking, with no precipitation. In fact, the substitution level of CMC in commercial products ranges roughly 0.4–1.0, and the results at 0.7 vs 1.2 DS indicate no dramatic change in binding across this range, meaning formulators

do not need to worry about calcium sensitivity even for highly substituted CMC. In summary, CMC can be considered essentially non-binding toward Ca^{2+} and Mg^{2+} under typical conditions, which is consistent with its failure to form gels like alginate does and with its safe use alongside minerals.

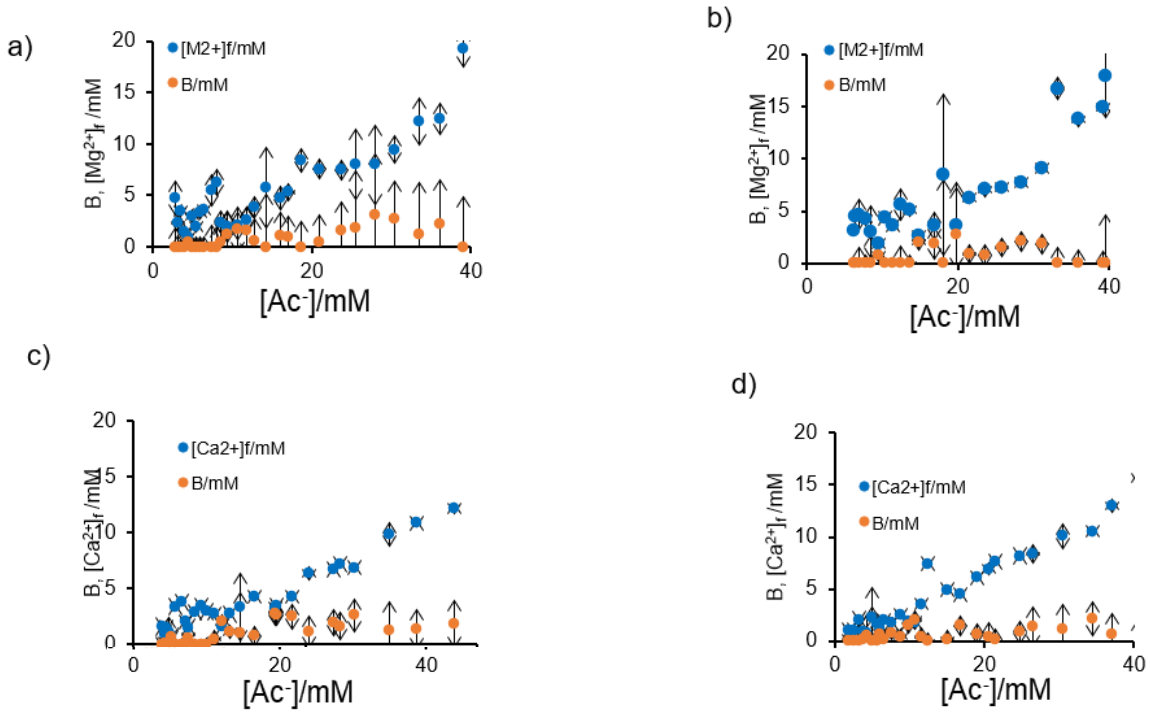


Figure 5.10. Plots of *B* and $[M^{2+}]_f$ versus $[Ac^-]$ of CMC: (a) CMC 0.7 with magnesium acetate, (b) CMC 1.2 with magnesium acetate, (c) CMC 0.7 with calcium acetate 0.7, CMC 1.2 with calcium acetate

Gallic Acid (Gallate)

Gallic acid (3,4,5-trihydroxybenzoic acid) is a small polyphenolic molecule with one carboxylic acid group and three phenolic –OH groups.[224] In neutral to slightly basic solution, it exists as the gallate anion (the carboxylate is deprotonated, and one or more phenolates may also deprotonate at higher pH). Gallic acid and its esters (such as propyl gallate) are used as food additives for their antioxidant properties, and gallic acid is also known to exhibit antimicrobial activity.[225] One mechanism by which gallic acid acts against Gram negative bacteria is by chelating the divalent cations (Ca^{2+}, Mg^{2+}) that stabilize the bacterial outer membrane, thereby disrupting the membrane integrity.[225] Given this known metal chelating ability, one might expect gallate to bind Ca^{2+} or Mg^{2+} to some extent in our experiments. However, the CSI results reveal that under our conditions (aqueous solution, pH ~7, 10 mM gallate), gallic acid behaves as a very weak binder to Ca^{2+} and Mg^{2+} effectively a non-binding analyte in the context of our measurement.

Figures 5.11 and 5.12 shows the binding curve for gallate. The plot of *B* vs $[Ac^-]$ for both Ca^{2+} and Mg^{2+} is nearly flat at $B \approx 0$ across the entire range, indicating no significant uptake of M^{2+} by gallate. Correspondingly, the free metal concentrations rise almost in lockstep with added acetate, essentially following the 1:2 ratio expected with no other binding. In statistical terms, any deviation of *B* from zero for gallate was within experimental uncertainty until the very highest $[Ac^-]$ points, where a barely perceptible binding might occur. This contrasts sharply with the strong-binding cases (alginate, CNC, oxalate) discussed elsewhere. Even when 20–30 mM Ca^{2+} had diffused into the gallate layer, the gallate did not appreciably sequester it free

Ca^{2+} accumulated freely. Likewise, Mg^{2+} showed no affinity for gallate; given Mg's general behaviour, this is unsurprising, but the absence of Ca binding is notable. Direct ^{43}Ca and ^{25}Mg NMR measurements were performed in samples of gallic acid (see figure 5.13 and 5.14) which can directly detect solid Ca and Mg compounds as binding of divalent ion into solid decreases the sensitivity and intensity of the observed calcium or magnesium signal (additionally, because ^{43}Ca and ^{25}Mg are quadrupolar nuclei their quadrupolar interactions in contexts such as ligand binding can cause rapid relaxation and severe line broadening of the signal hence only when ^{43}Ca and ^{25}Mg are in highly symmetric environment such as a fully hydrated ion in an aqueous solution can these ions give relatively sharp, observable signal). Those spectra showed only a very slight reduction in the intensity of the dissolved Ca^{2+} and Mg^{2+} signal in the presence of gallic acid, with no distinct solid precipitate signal. This confirms that gallate did not cause significant Ca^{2+} or magnesium precipitation (in contrast to oxalate, which produced a strong ^{43}Ca signal loss due to calcium oxalate precipitation, as discussed below).

The negligible binding of Ca^{2+} by gallate at pH 7 may at first seem inconsistent with gallate's known metal chelation effects. However, it can be rationalized by considering gallate's speciation and chelation requirements. At neutral pH, the gallate anion is monovalent (only the carboxylate is deprotonated; the phenolic hydroxyls have pK_a around 9 –10, so they remain largely protonated).[225] In this state, gallate offers essentially a single negatively charged binding site COO^- and possibly some polar OH groups, but it cannot form a stable multidentate chelate with Ca^{2+}. A bidentate chelation would require, for example, the adjacent phenolate and carboxylate both coordinating the metal, which would only occur if the phenolic oxygen were deprotonated and if the geometry is favourable. At pH 7, the phenolates are scarce, and Ca^{2+} binding to a single carboxylate (like acetate) is very weak (K ~1 M^{-1}. Therefore, it makes sense that gallate did not appreciably bind Ca^{2+} in our experiments. In the bacterial context, gallic acid's antimicrobial action may involve a locally higher pH or the combined effect of multiple gallate molecules complexing the Mg^{2+}/Ca^{2+} in the outer membrane; moreover, gallic acid's ability to intercalate into membranes could enhance its cation chelating impact on membrane stability. But in bulk solution at neutral pH, gallate behaves almost like a non-coordinating organic anion (comparable to benzoate or similar).

From a nutritional and formulation perspective, the extremely weak binding of Ca^{2+}/Mg^{2+} by gallic acid is reassuring. Gallic acid is present in various foods (tea, berries) and is used in additives; unlike oxalate, it is not considered a problematic anti-nutrient in terms of mineral absorption. Our data support this: gallic acid would not significantly sequester calcium or magnesium in the digestive tract. Even if gallic acid is co-ingested with a calcium source, it should not precipitate it or render it unavailable (indeed, gallate forms soluble complexes at most). This aligns with observations that gallic acid, despite its metal binding capability, does not severely impair mineral uptake in vivo a stark contrast to oxalic acid In food science, the primary role of gallic acid is as an antioxidant, and our findings indicate that any interaction with divalent cations in foods is minimal and likely limited to its antimicrobial mode of action (chelating surface bound cations on microbial cell envelopes). Thus, gallic acid can be regarded as essentially non-binding toward Ca^{2+} and Mg^{2+} under normal conditions, posing little risk of nutrient sequestration.

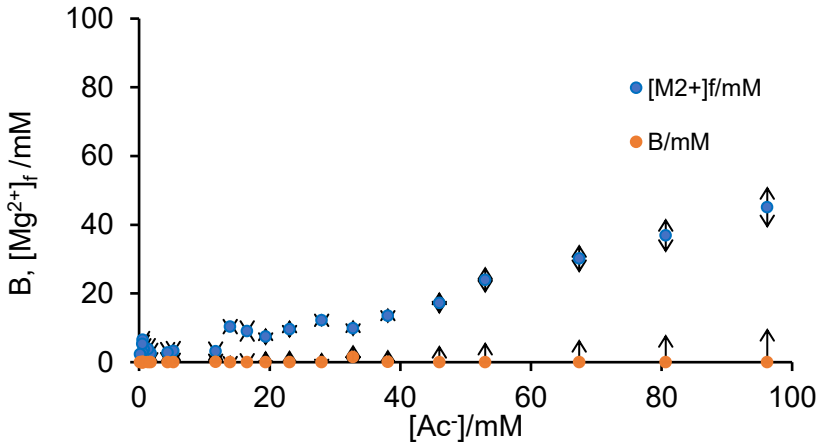


Figure 5.11. Plot of bound (orange) and free (blue) magnesium versus [Ac⁻] concentration of gallate

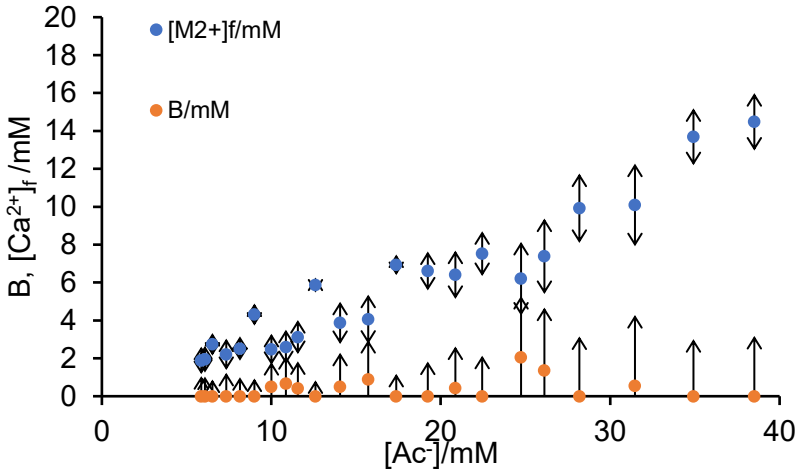


Figure 5.12. Plot of bound (orange) and free (blue) calcium versus [Ac⁻] concentration of gallate

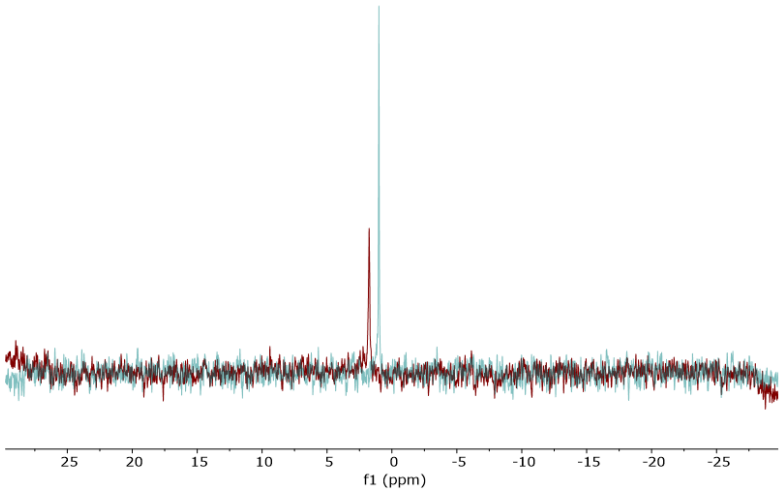


Figure 5.13. ^{43}Ca 1D NMR superimposition of 50 mM gallate (red) and reference sample with no analytes just 0.1 M calcium acetate (blue).

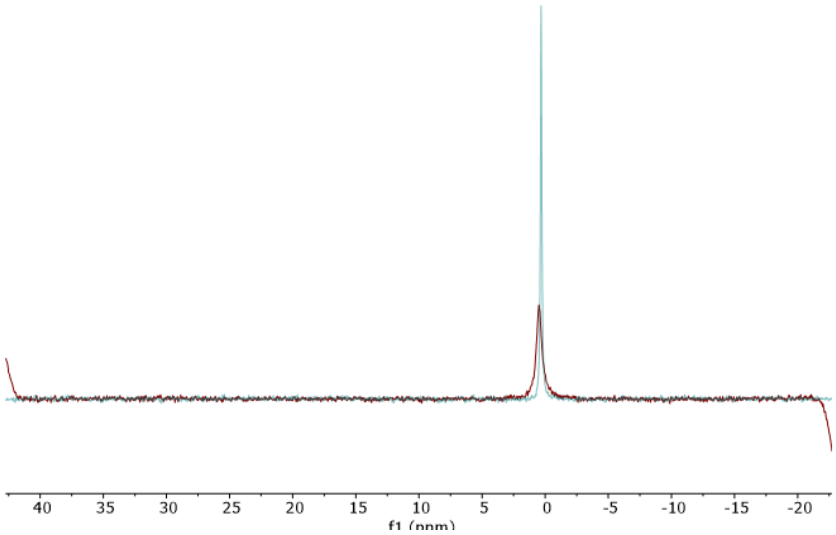


Figure 5.14.[25]Mg 1D NMR superimposition of 50 mM gallate (red) and reference sample with no analytes just 0.1 M magnesium acetate (blue).

Oxalic Acid (Oxalate)

Oxalate is a small dianion with two adjacent carboxylate groups.[226] It is well known for its strong chelation of divalent metal ions and the extreme insolubility of its calcium salt (calcium oxalate).[227] Oxalic acid is present in many plants (spinach, rhubarb, etc.) and is considered an antinutrient because it can bind dietary calcium and magnesium, forming unabsorbable precipitates. Interestingly, oxalic acid and oxalates are also used in certain food processing contexts for example, as a preservative or processing aid to sequester metal ions, prevent discoloration, and control texture in some vegetables.[228] oxalate binding was examined as a representative of a very strong-binding ligand. The CSI experiment was conducted with 10 mM sodium oxalate initially in the analyte layer. The results for oxalate are shown in figures 5.15-18.

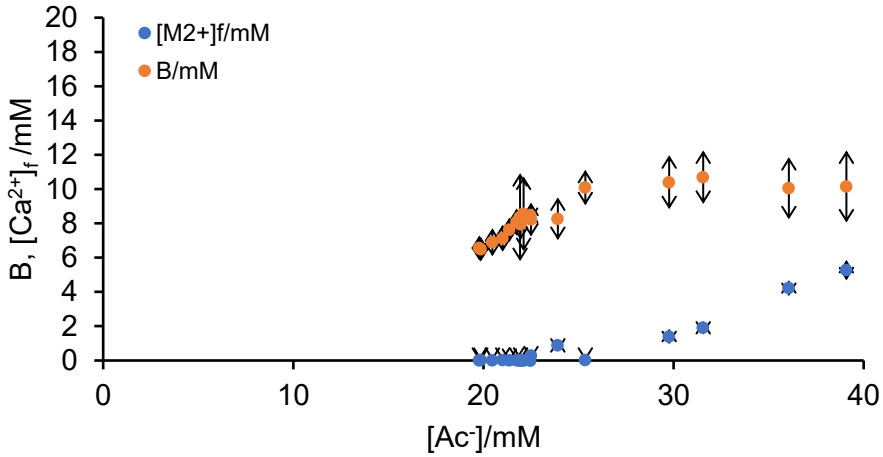


Figure 5.15. Plot of bound (orange) and free (blue) calcium versus [Ac^-] concentration of oxalate

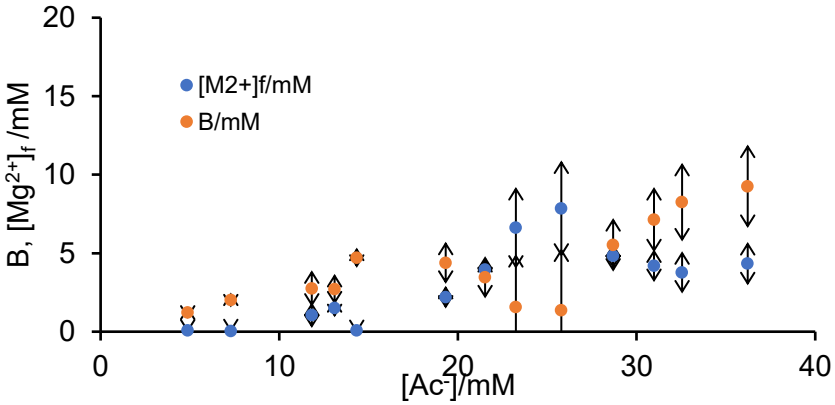


Figure 5.16. Plot of bound (orange) and free (blue) magnesium versus [Ac^-] concentration of oxalate

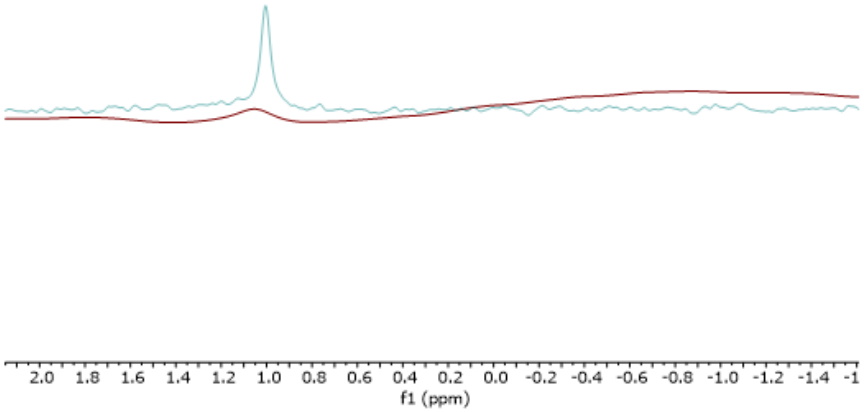


Figure 5.17. ^{43}Ca 1D NMR superimposition of 50 mM oxalate (red) and reference sample with no analytes just 0.1 M calcium acetate (blue). Baseline distortion possibly due to precipitation of calcium oxalate

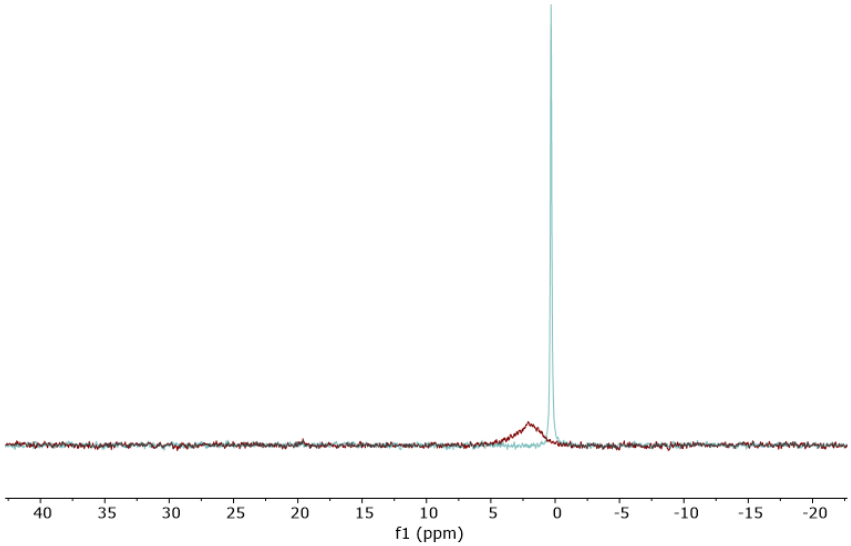


Figure 5.18. ^{25}Mg 1D NMR superimposition of 50 mM oxalate (red) and reference sample with no analytes just 0.1 M of magnesium acetate (blue).

oxalate scavenges Ca^{2+} almost completely until it is itself fully consumed. As Ca^{2+} diffuses upward, virtually none appears as free Ca^{2+} in solution until the amount of Ca^{2+} added exceeds the oxalate present. In fact, $[Ca^{2+}]_f$ remained near zero up to about $[Ac^-] \approx 20$ mM; only beyond that point did free Ca^{2+} begin to register, indicating that approximately 10 mM Ca^{2+} (half of 20 mM acetate) had been bound by the 10 mM oxalate (consistent with a 1:1 Ca:oxalate stoichiometry). Once the oxalate was saturated (i.e. converted to calcium oxalate), additional Ca^{2+} remained free in solution, causing $[Ca^{2+}]_f$ to rise. In terms of the binding parameter *B*, it essentially climbed to ~10 mM (the oxalate concentration) and then levelled off; *B* could not exceed the total oxalate since each oxalate can bind at most one Ca^{2+} (in a precipitated solid or complex). This behaviour is the hallmark of a strong, stoichiometric binder that precipitates its metal complex. Indeed, the formation of calcium oxalate was directly observed: the NMR tube developed a fine white precipitate in the oxalate region after the experiment, and in a homogeneous test, mixing Ca^{2+} acetate with sodium oxalate caused immediate precipitation of calcium oxalate (see figure 5.17). No such precipitate occurred with any other analyte (except CNC at high Ca, which formed flocs but not a crystalline solid).

Magnesium binding by oxalate was also evident, though to a lesser degree. In the same experiment, when magnesium acetate diffused into oxalate, we saw a suppression of , $[Mg^{2+}]_f$ relative to a blank, indicating that oxalate did bind some Mg^{2+}. However, the binding was not as complete as for Ca^{2+}. Free Mg^{2+} began to appear at a lower threshold of $[Ac^-]$ and *B* never reached the full 10 mM value; instead, a substantial fraction of Mg^{2+} remained unbound even before oxalate was saturated. This reflects the fact that magnesium oxalate has a higher solubility (or lower stability constant) compared to calcium oxalate. Quantitatively, literature reports that at physiological pH around 7, up to ~80% of Ca^{2+} can be precipitated as calcium oxalate, whereas only ~20% of Mg^{2+} might precipitate as magnesium oxalate under similar conditions.[229] Direct ^{25}Mg experiments on oxalate with magnesium acetate also showed significant binding of oxalate to magnesium albeit less strongly than to its calcium counterpart (see figure 5.18). The experimental observations are fully consistent with this: calcium oxalate formation was essentially complete (limited only by oxalate availability), while magnesium oxalate formation was partial. Chemically, this difference can be explained by the stronger lattice energy and coordination affinity in Calcium oxalate. The Ca^{2+} ion, being larger and more polarizable, can more easily accommodate the bidentate chelation by oxalate – it can adopt a coordination geometry that allows one oxalate to bind in a chelate ring (and indeed in the crystal Ca^{2+} coordinates to multiple oxalates). Mg^{2+}, with its smaller radius and stringent octahedral coordination, does not stabilize the oxalate chelate as effectively. Furthermore, Calcium oxalate is thermodynamically strongly favoured to remain in the solid phase relative to its dissociated ions in aqueous solution, as evidenced by its very low solubility product (K_{sp}) of $\sim 2.7\times10^{-9}$ at 25 °C (K_{sp} being the equilibrium constant for a solid ionic compound dissolving in water, representing the maximum product of its ions' concentrations in a saturated solution at a given temperature), whereas magnesium oxalate is orders of magnitude more soluble (K_{sp} $\sim 8.5\times10^{-5}$).[230] Thus, even when magnesium oxalate forms, a good portion can remain dissolved, contributing to $[Mg^{2+}]_f$ while calcium oxalate precipitates out almost completely, removing Ca^{2+} from the free pool. Direct ^{43}Ca NMR data for the oxalate sample showed a nearly total loss of the Ca^{2+} signal (since most of the calcium atoms went into solid calcium oxalate, which is NMR invisible), whereas for gallate the ^{43}Ca signal was largely retained, underscoring the contrast between strong and weak binders.

In practical terms, the oxalate results reinforce why oxalic acid is considered a potent anti-nutrient and a cause of pathological mineral deposition. Even a relatively small amount of oxalate (e.g.10 mM, which could correspond to eating a high oxalate food) can lock up an equivalent amount of calcium quantitatively, potentially leading to calcium deficiency if this occurs in the gut. In the context of kidney stone formation, our findings mirror the situation in urine: urine contains Ca^{2+} and Mg^{2+} and oxalate preferentially precipitates Ca^{2+} as calcium oxalate crystals, whereas magnesium oxalate remains more soluble. In fact, magnesium is sometimes given as a supplement to kidney stone patients because magnesium oxalate staying in solution means less oxalate available to precipitate Ca^{2+}.[230, 231] The strong binding of calcium by oxalate observed in our experiment is essentially the same phenomenon on a simpler laboratory scale.

On the other hand, the ability of oxalate to bind Ca^{2+} tightly can be useful in controlled scenarios. The food industry example mentioned (oxalate treatment of leafy greens) exploits this to inhibit spoilage: calcium is crucial for the activity of certain enzymes and microorganisms, so precipitating it with oxalate can slow decay and preserve texture.[232, 233] Our data confirm that adding oxalate would indeed eliminate free Ca^{2+} in a sample until oxalate is exhausted. However, one must be careful because any residual oxalate in food could then carry over as an antinutrient.

In summary, oxalate exhibits the strongest binding of all analytes studied, effectively sequestering Ca^{2+} in a 1:1 ratio and precipitating as calcium oxalate, and to a lesser extent binding Mg^{2+}. This behaviour is fully consistent with known chemistry (the low solubility of calcium oxalate) and with the need to manage oxalate intake to avoid nutritional and health issues. The CSI method was able to capture this precipitative binding process, distinguishing it clearly from the other systems.

Comparative Binding Trends and Practical Implications

Across the range of analytes tested from the biopolymers (CNC, CMC) to the small molecules (gallate, oxalate) a clear pattern emerges regarding Ca^{2+} vs. Mg^{2+} binding. In every case, Ca^{2+} showed equal or greater affinity for the analyte than Mg^{2+}. This general trend is grounded in the intrinsic properties of these cations: Ca^{2+} is larger, more polarizable, and has a lower charge density, making it an ion that can form stable complexes with multidentate ligands, whereas Mg^{2+} is a smaller, “harder” cation that remains heavily hydrated and typically coordinates fewer ligand atoms.[214] The results vividly illustrate this. In the case of strong-binding systems (alginate and oxalate), Ca^{2+} was bound so strongly that essentially no free Ca^{2+} appeared until the ligand was saturated, whereas Mg^{2+} binding was either much less complete (Mg–alginate showed only partial uptake; Mg–oxalate did not precipitate fully) or required higher concentrations to achieve similar uptake. For moderate-binding systems like CNC, both cations did bind substantially, but Ca^{2+} induced aggregation at a slightly lower threshold than Mg^{2+}, hinting at a mildly higher affinity. For weak/non-binding systems (CMC and gallate), neither cation bound strongly, but even there one could argue Ca^{2+} interacts a bit more than Mg^{2+} (for instance, CMC solution with calcium acetate appeared more viscous than CMC solution with magnesium acetate).

Another broad observation from the study is the importance of ligand structure in determining binding strength. Multidentate and polyanionic ligands (alginate, CNC, oxalate) show orders of magnitude greater uptake of M^{2+} than monodentate or sparsely substituted ligands (gallate, CMC). This is a manifestation of the chelate effect and cooperative binding. Alginate provides an extended array of carboxylate sites that can simultaneously engage one Ca^{2+} (especially

in G-blocks) once a Ca^{2+} is in the vicinity, it can interact with multiple residues, greatly enhancing the effective binding constant. Similarly, oxalate's two adjacent carboxylates allow it to form a 5-membered chelate ring with Ca^{2+}, which is thermodynamically favoured over two separate acetate like bindings. In contrast, gallate at pH 7 can only offer one strong binding site (its carboxylate) to a given metal ion, and CMC's carboxyls are too far apart or sterically hindered to chelate cooperatively. Thus, gallate and CMC behave essentially as a collection of independent single binding sites of low affinity, which is why their impact on free Ca^{2+}/Mg^{2+} is negligible.

5.6 Conclusion

The application of one-dimensional chemical shift imaging (CSI) in ^{1}H NMR spectroscopy provided a novel, robust, and quantitative method for analysing the binding of divalent cations, Ca^{2+} and Mg^{2+}, to a range of analytes including biopolymers and small organic molecules. By exploiting the spatial resolution inherent to CSI, we were able to construct detailed profiles of free and bound metal concentrations throughout a diffusion gradient, enabling the direct observation of binding phenomena without the need for extensive titration protocols or specialized instrumentation.

Our results clearly demonstrate that calcium ions exhibit stronger binding affinities across all studied analytes compared to magnesium ions, a trend consistent with fundamental principles of ion hydration, polarizability, and coordination chemistry. Polymers such as sodium alginate and citrate-functionalized cellulose nanocrystals (CNC) displayed strong binding with Ca^{2+}, with CNC also showing significant uptake of Mg^{2+}. In contrast, carboxymethyl cellulose (CMC) and gallic acid exhibited very weak to negligible binding under the studied conditions. Oxalate, as expected, showed extremely strong binding to Ca^{2+}, leading to stoichiometric precipitation of calcium oxalate, with a partial binding effect also observed for Mg^{2+}.

The study reinforced important concepts in macromolecular and solution chemistry. For instance, multivalent interactions, polyelectrolyte architecture, and ligand chelation geometry were identified as key determinants of binding strength. Furthermore, the practical implications of these findings extend to fields ranging from food science and water treatment to biomedical engineering and pharmaceutical formulation. Understanding the binding profiles of polymers and small molecules with divalent cations is crucial for designing materials with controlled mechanical properties, stability, and bioavailability.

Finally, this work illustrates the utility of CSI NMR as a rapid, high-resolution method for profiling ion binding behaviours under near physiological conditions. Given its minimal sample requirements, high throughput, and compatibility with standard NMR hardware, the approach developed herein offers a powerful tool for future studies of polymer–ion interactions, ligand screening, and the in situ monitoring of binding kinetics and equilibria. Future directions may include extending the method to trivalent metal ions (e.g., La^{3+}, Al^{3+}), competitive binding systems, and more complex matrices such as whole food extracts or biological fluids, thereby further expanding the applicability of CSI-based ion binding analysis.

Overall conclusions and outlook

This thesis has developed and demonstrated innovative nuclear magnetic resonance chemical shift imaging methodologies to study acid base equilibria and metal ligand interactions with unprecedented spatial and experimental efficiency. By generating controlled gradients of pH, solvent composition, or metal ion concentration within a single NMR tube, and using spatially resolved ^{1}H NMR spectroscopy, complex titration and binding studies were achieved without serial sample preparation. These methods drastically reduce time, material consumption, and manual effort, while enabling detailed, in situ analysis of chemical processes in homogeneous and heterogeneous environments alike.

Among the most significant contributions is the development of a derivative-based fitting method for pK_a determination, which offers a robust alternative to traditional curve fitting by utilizing the derivative of pH with respect to chemical shift. This polynomial based model mitigates convergence issues common in non-linear regression, while maintaining agreement with literature pK_a values to within 0.4 units. The method simplifies analysis and enhances reproducibility, providing a practical route to obtain high-quality thermodynamic data from a single CSI experiment.

The thesis also introduced a CSI based method for determining pK_a values of NMR-invisible species such as polymers or aggregated proteins by monitoring proton exchange with indicator molecules. This enabled successful pK_a determination even when the analyte itself could not be directly observed, broadening the scope of NMR-based acidity studies to challenging systems. Furthermore, the thesis demonstrated that spatial solvent gradients within a single sample could be used to determine solvent dependent pK_a values and extrapolate true aqueous pK_a via the Yasuda Shedlovsky method. This presents a valuable strategy for studying poorly water-soluble or amphiphilic compounds.

In later chapters, a single tube CSI approach for studying divalent ion binding (Ca^{2+} and Mg^{2+}) was introduced and applied to a variety of small molecules and macromolecules. This method yielded spatial binding profiles, enabling quantitative determination of free and analyte bound divalent ion concentrations. The results aligned well with known chemical behaviours, such as the strong Ca^{2+} binding affinity of alginate and the minimal interaction of CMC. Importantly, these techniques provided visual and quantitative insight into metal-induced aggregation phenomena such as CNC flocculation offering a powerful tool for materials science and colloid chemistry.

The broader implications of these developments span numerous disciplines. In pharmaceutical sciences, the ability to rapidly and accurately measure pK_a and divalent ion binding dynamics with minimal material is valuable for early-stage drug profiling. In food science and nutrition, the ability to quantify calcium sequestration by polysaccharides informs both dietary safety and formulation strategies. In polymer and materials chemistry, the methods offer new avenues to characterize ion-binding properties critical for hydrogel design, ion exchange, and responsive materials. The CSI-based methods are compatible with standard high-field NMR spectrometers, making them accessible and scalable for many research laboratories.

Looking forward, several promising directions emerge. The CSI approach could be extended to study trivalent ions (e.g., La^{3+}, Al^{3+}) or to competitive binding scenarios involving multiple ligands or ions. Coupling CSI with non-proton detection (e.g., ^{19}F, ^{31}P) could broaden the detectable species and offer more complex multiplexing. Gradient experiments involving other parameters such as ionic strength, temperature, or co-solvent polarity could provide multi-dimensional insight into chemical equilibria.

Moreover, applying these techniques to complex biological or industrial matrices presents an exciting opportunity. CSI has potential to analyse reactions in media that are opaque, heterogeneous, or poorly defined environments where traditional NMR or optical assays fail. Studying drug release in simulated gastric fluid, calcium binding in plant extracts, or charge interactions in colloidal suspensions could all benefit from the spatially resolved insights offered by CSI. These directions align with a broader scientific trend toward in situ, context-aware analysis and bring analytical chemistry closer to the real conditions under which biological and industrial processes occur. Additionally, CSI can be combined with other NMR experimental techniques such diffusion ordered spectroscopy (DOSY) in order to obtain spatially dependent information of how the diffusion coefficient of compounds vary along NMR tube. This approach could be useful to characterise polymers and porous media such as gels via the self-diffusion behaviour of imbibed molecular probes spatially resolved NMR measurements along the vertical axis of the tube are paramount to gain an accurate understanding of the internal architecture of gels.

In summary, this thesis establishes ^{1}H CSI as a powerful, versatile, and accessible method for measuring chemical equilibria. It expands the boundaries of what NMR can achieve by embedding rich spatial information into a single experiment. The techniques developed offer practical advantages, theoretical robustness, and broad applicability, laying the foundation for future research that is more efficient, informative, and connected to real-world systems.

Appendix NMR Processing and pulse program

All pulse programs were developed by Matthew Wallace and used with his permission and with his guidance.

A.1 2D pulse sequence for CSI (Bruker)

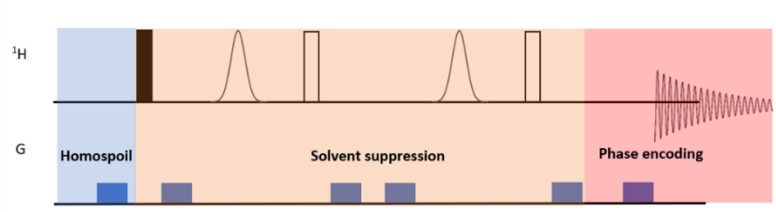


Figure A.1. Chemical shift imaging sequence utilised in this work. It is broken down into three sections: (blue) homospoil gradient to destroy transverse magnetisation, (orange) solvent suppression sequence through excitation sculpting, (red) phase encoding gradient pulse.

;Modified from: "Probing spatial distribution of alignment by deuterium NMR imaging"

;Chemistry - A European Journal, Volume 19, Issue 22, 27 May 2013, Pages 7013-7019

;2D sequence for z-imaging preserving chemical shift

;using a phase encoding gradient. Original sequence written by Christian Merle, Martin Koos

;Modified to be on 1H with Excitation sculpting for water suppression

;Matthew Wallace, 3/2022 (University of East Anglia, matthew.wallace@uea.ac.uk)

;This pulse program is not fully tested and comes without warranty.

;!!This version is for running unlocked with sweep etc. disabled

;!!For running locked, delete the (UN)BLKGRAMP and insert the (UN)BLKGRAD

;Set 1 SW to Z-range in mm (see cnst0) to get 1 Hz/mm scale in indirect dimension

;Make cnst0 bigger than actual sample size to avoid folding artefacts.

;Keep gpz6 at 100% and adjust cnst3 to get p30 to an acceptable length according to instrument (ca.150-300 us)

; 1H-Version

;$CLASS=HighRes

;$DIM=2D

```
;$TYPE=
;$SUBTYPE=
;$COMMENT=
prosol relations=<triple>
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>
"cnst2= 0.8914027" ; integralfactor of gradient shape SMSQ10.32
"cnst4= 267.52220" ; * 10^6 /Ts = gamma1H
"p30=(td1/cnst0)*(1/(cnst1*cnst2*cnst3))*(1/cnst4)*(2*3.14159265/1000)*0.5 s"
"l1=td1-1"
lgrad r1d = l1
"acqt0=0"
"DELTA1=p30+d16"
"TAU=de+p1*2/3.1416+50u"
"p2=p1*2"
"d12=20u"
"d4=d1-100m"
baseopt_echo
1 ze
2 30m ;BLKGRAD so locked during d1
d1
;spoil gradient from previous acquisition
3 50u UNBLKGRAMP ;UNBLKGRAD
p19:gp3
d16
;start of zgesgp
d12 pl1:f1
 p1 ph1
 50u
 p16:gp1
 d16 pl0:f1
 (p12:sp1 ph2:r):f1
 4u
```

```
 d12 pl1:f1
 p2 ph3
 4u
 p16:gp1
 d16
 TAU
DELTA1
 p16:gp2
 d16 pl0:f1
 (p12:sp1 ph4:r):f1
 4u
 d12 pl1:f1
 p2 ph5
 4u
 p16:gp2
 d16
p30:gp6*cnst3*r1d
S32
d16 BLKGRAMP ;remove this BLKGRAMP statement if running locked
go=2 ph31
30m
100m wr #0 if #0 zd igrad r1d
d4
lo to 3 times l1
goto 5
; run last increment:
4 30m ;BLKGRAD
d1
;spoil gradient from previous
5 50u UNBLKGRAMP ;UNBLKGRAD
p19:gp3
d16
 ;start of zgesgp
d12 pl1:f1
```

```
 p1 ph1
 50u
 p16:gp1
 d16 pl0:f1
 (p12:sp1 ph2:r):f1
 4u
 d12 pl1:f1
 p2 ph3
 4u
 p16:gp1
 d16
 TAU
DELTA1
 p16:gp2
 d16 pl0:f1
 (p12:sp1 ph4:r):f1
 4u
 d12 pl1:f1
 p2 ph5
 4u
 p16:gp2
 d16
p30:gp6*cnst3*r1d
d16 BLKGRAMP ;remove this BLKGRAMP statement if running locked
go=4 ph31
30m ;BLKGRAD
100m wr #0 if #0 zd
d3
exit
;Phase cycling as for zgesgp
ph1=0
ph2=0 1
ph3=2 3
ph4=0 0 1 1
```

ph5=2 2 3 3

ph31=0 2 2 0

;pl0 : 0W

;cnst0 : z

-Range in cm

;cnst1 : GCC (G/mm) from Gradpar

;cnst3 : set max to get acceptable length of GP of ca 150

-300us [0.95 max]

;pl1 : f1 channel

- power level for pulse (default)

;sp1 : f1 channel

- shaped pulse 180 degree

;p1 : f1 channel

- 90 degree high power pulse

;p12: f1 channel

- 180 degree shaped pulse (Gaus1.1000) [4 msec]

;p16: homospoil/gradient pulse for water suppression (1000 us)

;p19 : Spoil gradient pulse (1000 us)

;gpz6: 100% phase encoding gradient

;d1 : Relaxation delay

;d3: waiting time before next experiment in a time delay sequence [s] set to 0.1 if no delay

;d12: delay for power switching [20 usec]

;d16: standard eddy delay (200u)

;ns: 4 * n, total number of scans: NS * TD0

;ds: 2*m [16]

;td1: number of experiments

;FnMODE: QF

;use gradient ratio: gp 1 : gp 2

; 31 : 11

;for z

-only gradients:

;gpz1: 31%

;gpz2: 11%

;gpz3: 50%

;use gradient files:

;gpnam1: SMSQ10.100

;gpnam2: SMSQ10.100

;gpnam3: SMSQ10.100

;gpnam6: SMSQ10.32

;$Id: zgesgpimg,v 1.9 2012/01/31 17:49:32 ber Exp $

A.2 AU program to process raw CSI datasets

```
/*To produce phase corrected chemical shift image from gradient encoded data*/
/*Set 1st order phase correction for f1 dimension to 180*Td1 (11520 when 64 points in image), 0th order to 0*/
/*PH_mod should be set to PK in both dimensions*/
/*XFB to produce image*/
/*This is done automatically by script 13.2 above*/
/*With the 2D dataset selected, Run this Au*/
/*AU extracts each row in turn to a procno and automatically phase and baseline corrects/*
/*Reversal of F1 axis may be necessary, depending on NMR probe*/
/*This AU is not fully tested and comes without warranty.*/
/*The script works on Bruker Topspin 3.6.2 but has not been tested on other versions*/
/*Use kill command if all goes wrong*/
/*Matthew Wallace, 3/2022*/
/*University of East Anglia, matthew.wallace@uea.ac.uk*/
char disk1[32], user1[32], location[128], phtyp[8];
float abf1=8;
float abf2=6;
int phpno=1;
int w=1;
int np=64;
int pno=5;
GETCURDATA
int steno=expno;
strcpy(location,disk);
strcpy(phtyp,"k");
GETSTRING("Enter location of dataset",location)
phpno=procno;
GETINT("Enter experiment number to process",steno)
GETINT("Enter procno containing XFB processed 2D data :",phpno)
pno=phpno+5;
GETINT("Enter procno to write rows to phase and baseline correct (empty):",pno)
REXPNO(steno)
RPROCNO(phpno)
SETCURDATA
FETCHPAR1("SI",&np)
GETINT("Enter number of points in image (autodetects) :",np)
GETSTRING("APKS (s) or APK (k) or apkf (f) auto phase correction?",phtyp)
if(strcmp(phtyp,"f")==0)
{
GETFLOAT("Enter right limit for apkf and absf:",abf2)
GETFLOAT("Enter left limit for apkf and absf:",abf1)
}
w=1;
TIMES(np)
```

```
{
RPROCNO(phpno)
SETCURDATA
RSR(w,pno)
RPROCNO(pno)
SETCURDATA
        if(strcmp(phtyp,"s")==0)
        {
APKS
ABS
        }
        if(strcmp(phtyp,"k")==0)
        {
APK
ABS
        }
        if(strcmp(phtyp,"f")==0)
        {
                STOREPAR("absf1",abf1)
                STOREPAR("absf2",abf2)
APKF
ABSF
        }
        WSR(w,phpno,steno,name,user,location)
w++;
        }
        END
        QUITS
```

A.3 AU program to pick peak a chemical shift in CSI

```
/*Goes through a 2D image and writes file of peak position/ppm*/

/*The script works on Bruker Topspin 3.6.2 but has not been tested on other versions*/
/*Use kill command if all goes wrong*/
/*Matthew Wallace and Haider Hussain, 6/2024*/
/*University of East Anglia, matthew.wallace@uea.ac.uk*/
/*Will pick centre of a multiplet*/

FILE *find;

float min=0;

double f2pind=1.01;

double f1pind=1.6;

float ppsens=0.9;

double pc=0.1;

int steno=15;

int eno;

int ne=5;

int m=-1;

int rn=1;
```

```
double peakFreqHz, peakFreqPPM, peakIntensity, maxpsh, maxpsp, maxips,
sf,sfo1,so1p,mintpp,minpsp,peakppmneg,cent,ppmdif,maxpspneg;

int i, numPeaks;

int np=64;

int row=1;

int v=1;

int phpno=1;

GETCURDATA

steno=expno;

phpno=procno;

FETCHPAR1("td",&np)

GETINT("Enter starting experiment number: ",steno)

GETINT("Enter number of gradient values",np)

GETINT("Enter procno of 2D dataset",phpno)

GETDOUBLE("Indictaor right peak picking/ppm",f2pind)

GETDOUBLE("Indicator left peak picking/ppm",f1pind)

GETFLOAT("Enter peak picking senistitivty factor",pc)

GETDOUBLE("Enter satelite sensitivity factor",ppsens)

/*Reads in peak picking parameters for ind*/

REXPNO(steno)

SETCURDATA

RPROCNO(phpno)

SETCURDATA

STOREPAR("mi",min)

STOREPAR("pc",pc)


if ((find = fopen(ACQUPATH("Indicator chemical shift.txt"),"wt")) == 0)

   STOPMSG("Cannot create difflist")

TIMES(np)

{

RSR(v,5)

RPROCNO(5)

SETCURDATA

STOREPAR("f2p",f2pind)
```

```
STOREPAR("f1p",f1pind)
PP
 numPeaks = readPeakList(PROCPATH(0));
maxips=0.0;
   maxpsh=0.0;
   for (i=0; i<numPeaks; i++)
   {
          peakIntensity = getPeakIntensity(i);
          peakFreqHz = getPeakFreqHz(i);
          peakFreqPPM = getPeakFreqPPM(i);
          if (peakIntensity > maxips)
          {
             maxips = peakIntensity;
             maxpsh = peakFreqHz;
             maxpsp = peakFreqPPM;
          }
   }
   /*Pick most downfield side of multiplet*/
   mintpp=maxips*ppsens;
maxpsp=0.0;
   for (i=0; i<numPeaks; i++)
   {
                  peakIntensity = getPeakIntensity(i);
                  if(peakIntensity>mintpp)
                  {
          peakFreqPPM = getPeakFreqPPM(i);
          peakFreqHz = getPeakFreqHz(i);
          if (peakFreqHz >= maxpsh)
          {
             maxpsp = peakFreqPPM;
             maxpsh = peakFreqHz;
          }
   }
          }
```

```
/*Flips negative to choose most upfield peak of multiplet*/
for (i=0; i<numPeaks; i++)
   {
                peakIntensity = getPeakIntensity(i);
                if(peakIntensity>mintpp)
                {
        peakFreqPPM = getPeakFreqPPM(i);
peakppmneg=peakFreqPPM*m;
maxpspneg=maxpsp*m;
        if (peakppmneg >= maxpspneg)
        {
           minpsp = peakFreqPPM;
        }
   }
        }
   freePeakList();
/*writes centre of multiplet into text document*/
ppmdif=maxpsp-minpsp;
cent=minpsp+ppmdif*0.5;
   fprintf(find,"%f \n",cent);
v++;
rn++;
RPROCNO(phpno)
SETCURDATA
}
END
fclose(find);
QUIT
```

A.4 AU program to integrate a peak in CSI

```
/*Bruker AU script for integrating a CSI dataset*/
/*CSI dataset should have been fully processed in phase-sensitive mode*/
/*Crude chemical shift referencing in F2 also helps*/
```

```
/*The script works on Bruker Topspin 3.6.2 but has not been tested on other versions*/

/*Create a 1D integral file using the wmisc command*/

/*This should contain only the reference compound and indicator*/

/*Copy and paste this name into the text box when asked, or paste into strcpy(text,"indref"); line 46 below, in place of indref */

/*indref is the name of the integral range file used in this script by default*/

/*The script extracts each row in turn into the empty procno requested (will overwrite existing contents!!!)*/

/*Working in this procno, the spectrum is referenced to the chemical shift reference supplied*/

/*Default values here apply to DMSO, DMSO signal is downfield of indicator (ordans=d)*/

/*Change right and left peak picking limits if any risk of a non-reference peak being included in the referencing procedure*/

/*Having referenced the spectrum, the requested integral file is read in and the integral values exported to .txt documents*/

/*Each .txt document is stored in the procno directory of the 2D CSI dataset*/

/*The script will also by default save rows 15, 32 and 50 into procnos of the same number*/

/*This is so the work of the script can be checked - spectra are correctly referenced and integrals correctly applied*/

/*This AU is not fully tested and comes without warranty.*/

/*Use kill command if all goes wrong*/

/*Matthew Wallace and Haider Hussain, 6/2024*/

/*University of East Anglia, matthew.wallace@uea.ac.uk*/

FILE    *fpnt,*fref,*fac;

char savans[8],dummystr[256],intdir[256],location[128],ordans[8];

float min=0;

double f2pdss=-0.5;

double f1pdss=0.5;

double ref=0;

float ppsens=0.9;

double pc=0.1;

int m=-1;

int rn=1;

int sav1=15;

int sav2=32;

int sav3=50;

int steno;
```

```
double peakFreqHz, peakFreqPPM, peakIntensity, maxpsh, maxpsp, maxips, sf,sfn,sfo1,intgr;
double intnum,ppmdn,ppmup,intgrso1p,mintpp,minpsp,peakppmneg,cent,ppmdif,maxpspneg;
int i, numPeaks;
int np=64;
int row=1;
int v=1;
int wrpno=5;
int linenum=1;
int phpno=1;
strcpy(location,disk);
strcpy(savans,"y");
strcpy(text,"indref");
strcpy(ordans,"d");
GETCURDATA
steno=expno;
phpno=procno;
GETINT("Enter experiment number of 2D dataset",steno)
GETINT("Enter procno of 2D dataset",phpno)
REXPNO(steno)
RPROCNO(phpno)
SETCURDATA
FETCHPAR1("td",&np)
GETINT("Enter number of spectra in CSI image",np)
GETINT("Enter procno to extract rows into for integration (blank)",wrpno)
GETDOUBLE("Chemical shift reference right peak picking/ppm",f2pdss)
GETDOUBLE("Chemical shift reference left peak picking/ppm",f1pdss)
GETDOUBLE("Enter reference shift/ppm",ref)
GETFLOAT("Enter peak picking sensitivity factor",pc)
GETDOUBLE("Enter satellite sensitivity factor",ppsens)
GETSTRING("Save example spectra to check working OK? y/n",savans)
if(strcmp(savans,"y")==0)
{
GETINT("Enter 1st spectrum to save (put in pno of same#)",sav1)
GETINT("Enter 2nd spectrum to save (put in pno of same#)",sav2)
```

```
GETINT("Enter 3rd spectrum to save (put in pno of same#)",sav3)
}
GETSTRING("Which intrng file must be used?", text)
GETSTRING("Is the reference integral upfield (u) or downfield (d) of indicator?",ordans)
REXPNO(steno)
SETCURDATA
RPROCNO(phpno)
SETCURDATA
STOREPAR("mi",min)
STOREPAR("pc",pc)
STOREPAR("CURPRIN","Integrals.txt")
/*Create text files to hold integral data*/
/*Global scaling*/
STOREPAR("pscal",0)
if ((fref = fopen(ACQUPATH("Reference integral.txt"),"wt")) == 0)
   STOPMSG("Cannot create difflist")
if ((fac = fopen(ACQUPATH("Indicator integral.txt"),"wt")) == 0)
   STOPMSG("Cannot create difflist")
   /*No go through each row in turn*/
TIMES(np)
{
RSR(v,wrpno)
RPROCNO(wrpno)
SETCURDATA
sprintf(intdir,"%s\\%s\\%i\\pdata\\%i\\integrals.txt",location,name,expno,procno);
/*Reference the spectrum.*/
/*If the reference peak is split, the program will reference based on the average shift*/
/*of the two peaks in the requested peak picking range which are within ppsens of the largest peak found*/
STOREPAR("f2p",f2pdss)
STOREPAR("f1p",f1pdss)
PP
 numPeaks = readPeakList(PROCPATH(0));
maxips=0.0;
   maxpsh=0.0;
```

```
   for (i=0; i<numPeaks; i++)
   {
         peakIntensity = getPeakIntensity(i);
         peakFreqHz = getPeakFreqHz(i);
         peakFreqPPM = getPeakFreqPPM(i);
         if (peakIntensity > maxips)
         {
            maxips = peakIntensity;
            maxpsh = peakFreqHz;
            maxpsp = peakFreqPPM;
         }
   }
   /*Pick most downfield side of multiplet*/
   mintpp=maxips*ppsens;
maxpsp=0.0;
   for (i=0; i<numPeaks; i++)
   {
                  peakIntensity = getPeakIntensity(i);
                  if(peakIntensity>mintpp)
                  {
         peakFreqPPM = getPeakFreqPPM(i);
         peakFreqHz = getPeakFreqHz(i);
         if (peakFreqHz >= maxpsh)
         {
            maxpsp = peakFreqPPM;
            maxpsh = peakFreqHz;
         }
   }
         }
/*Flips negative to choose most upfield peak of multiplet*/
for (i=0; i<numPeaks; i++)
   {
                  peakIntensity = getPeakIntensity(i);
                  if(peakIntensity>mintpp)
```

```
                    {
        peakFreqPPM = getPeakFreqPPM(i);
peakppmneg=peakFreqPPM*m;
maxpspneg=maxpsp*m;
        if (peakppmneg >= maxpspneg)
        {
           minpsp = peakFreqPPM;
        }
   }
        }
   freePeakList();
/*References spectrum*/
ppmdif=maxpsp-minpsp;
cent=minpsp+ppmdif*0.5;
FETCHPAR("sf",&sf)
   sfn=sf+(cent-ref)*sf/(1e6);
   STOREPAR("sf",sfn)
/*Integrate spectra, read numbers and store integrals in text file*/
        RMISC("intrng", text)
        LI
fpnt=fopen(intdir, "r");
fgets(dummystr, sizeof(dummystr), fpnt);
while (fgets(dummystr, sizeof(dummystr), fpnt) != NULL)
{
        /*Need to selectively elimate rows, then scan for numbers*/
if(linenum>=5)
{
        (void) sscanf(dummystr,"%lf %lf %lf %lf",
                        &intnum,&ppmdn,&ppmup,&intgr);
                        /*DMSO is first, then indicator*/
                                if(linenum==5)
                        {
                                if(strcmp(ordans,"d")==0)
                                {
```

```
                        fprintf(fref,"%f\n",intgr);
                                }
                                        if(strcmp(ordans,"u")==0)
                                {
                        fprintf(fac,"%f\n",intgr);
                                }
                        }
                                if(linenum==6)
                        {
                if(strcmp(ordans,"d")==0)
                                {
                        fprintf(fac,"%f\n",intgr);
                                }
                                        if(strcmp(ordans,"u")==0)
                                {
                        fprintf(fref,"%f\n",intgr);
                                }
                        }
                        intnum=0;
                        ppmdn=0;
                        ppmup=0;
                        intgr=0;
                        linenum++;
}
        else
{
                        linenum++;
}
}
linenum=1;
fclose(fpnt);
        /*Save spectra in procnos if requested (will overwrite!!)*/
        if(strcmp(savans,"y")==0)
{
```

```
if(v==sav1)
{
WRP(sav1)
}
if(v==sav2)
{
WRP(sav2)
}
if(v==sav3)
{
WRP(sav3)
}
}
v++;
RPROCNO(phpno)
SETCURDATA
}
END
fclose(fac);
fclose(fref);
QUIT
```

A.5 AU program to extract multiple chemical shifts from CSI dataset (Bruker)

```
/*Bruker AU script for extracting peak positions from a CSI dataset*/
/*CSI dataset should have been fully processed in phase-sensitive mode*/
/*Rough chemical shift referencing in F2 also helps*/
/*The script works on Bruker Topspin 3.6.2 but has not been tested on other versions*/
/*The script extracts each row in turn into the empty procno requested (will overwrite existing contents!!!)*/
/*Will write peak positions of glycolate, SFA, 2MI, ref and acetate to separate .txt files/ppm  in procno directory of CSI dataset*/
/*These numbers go from first to final row of the dataset*/
/*Numbers can be copied and pasted into spreadsheet*/
/*Peak picking routine will find the centre of a multiplet (doublet, quartet, or singlet with bad shim)*/
/*Will treat two peaks as outer edges of a multiplet if their intensity is within ppsens of the biggest peak found in specified range*/
/*Adjust peak picking ranges as appropriate to avoid overlap*/
/*Default values here apply to DMSO as chemical shift reference*/
/*This AU is not fully tested and comes without warranty.*/
/*Use kill command if all goes wrong*/
/*Matthew Wallace, 3/2022*/
/*University of East Anglia, matthew.wallace@uea.ac.uk*/
```

```
FILE *fglyc,*fref,*fac,*fsfa,*fmi;
float min=0;
double f2pref=2.5;
double f1pref=2.8;
double f2pglyc=3.85;
double f1pglyc=4.2;
double f2pac=1.8;
double f1pac=2.18;
double f2mi=6;
double f1mi=9;
double f2sfa=3.65;
double f1sfa=3.85;
float ppsens=0.9;
double pc=0.1;
int steno=15;
int eno;
int ne=5;
int m=-1;
int rn=1;
double peakFreqHz, peakFreqPPM, peakIntensity, maxpsh, maxpsp, maxips, sf,sfo1,so1p;
double mintpp,minpsp,peakppmneg,cent,ppmdif,maxpspneg;
int i, numPeaks;
int np=64;
int row=1;
int v=1;
int wrpno=5;
int phpno=1;
GETCURDATA
steno=expno;
phpno=procno;
GETINT("Enter experiment number of CSI dataset",steno)
GETINT("Enter procno of CSI dataset",phpno)
REXPNO(steno)
RPROCNO(phpno)
SETCURDATA
FETCHPAR1("td",&np)
GETINT("Enter number of gradient values",np)
GETINT("Enter procno to extract rows into for peak picking (blank)",wrpno)
GETDOUBLE("Reference right peak picking/ppm",f2pref)
GETDOUBLE("Reference left peak picking/ppm",f1pref)
GETDOUBLE("Glycolate right peak picking/ppm",f2pglyc)
GETDOUBLE("Glycolate left peak picking/ppm",f1pglyc)
GETDOUBLE("Acetate right peak picking/ppm",f2pac)
GETDOUBLE("Acetate left peak picking/ppm",f1pac)
GETDOUBLE("Sulfoacetate right peak picking/ppm",f2sfa)
GETDOUBLE("Sulfoacetate left peak picking/ppm",f1sfa)
GETDOUBLE("2-methylimidazole right peak picking/ppm",f2mi)
GETDOUBLE("2-methylimidazole left peak picking/ppm",f1mi)
GETFLOAT("Enter peak picking sensitivity factor",pc)
GETDOUBLE("Enter satellite sensitivity factor",ppsens)
STOREPAR("mi",min)
STOREPAR("pc",pc)
/*Global scaling*/
STOREPAR("pscal",0)
/*Creates text files to hold peak positions*/
if ((fref = fopen(PROCPATH("Reference chemical shift.txt"),"wt")) == 0)
   STOPMSG("Cannot create file")
   if ((fglyc = fopen(PROCPATH("Glycolate.txt"),"wt")) == 0)
   STOPMSG("Cannot create file")
   if ((fac = fopen(PROCPATH("Acetate.txt"),"wt")) == 0)
```

```
   STOPMSG("Cannot create file")
    if ((fmi = fopen(PROCPATH("2-methylimidazole.txt"),"wt")) == 0)
   STOPMSG("Cannot create file")
    if ((fsfa = fopen(PROCPATH("sulfoacetate.txt"),"wt")) == 0)
   STOPMSG("Cannot create file")
  /*No go through each row in turn and extract into requested procno, find peak positions*/
TIMES(np)
{
RSR(v,wrpno)
RPROCNO(wrpno)
SETCURDATA
/*Extract chemcial shift of reference peak first*/
STOREPAR("f2p",f2pref)
STOREPAR("f1p",f1pref)
PP
 numPeaks = readPeakList(PROCPATH(0));
maxips=0.0;
   maxpsh=0.0;
   for (i=0; i<numPeaks; i++)
   {
          peakIntensity = getPeakIntensity(i);
          peakFreqHz = getPeakFreqHz(i);
          peakFreqPPM = getPeakFreqPPM(i);
          if (peakIntensity > maxips)
          {
             maxips = peakIntensity;
             maxpsh = peakFreqHz;
             maxpsp = peakFreqPPM;
          }
   }
   /*Pick most downfield side of multiplet*/
   mintpp=maxips*ppsens;
maxpsp=0.0;
   for (i=0; i<numPeaks; i++)
   {
                  peakIntensity = getPeakIntensity(i);
                  if(peakIntensity>mintpp)
                  {
          peakFreqPPM = getPeakFreqPPM(i);
          peakFreqHz = getPeakFreqHz(i);
          if (peakFreqHz >= maxpsh)
          {
             maxpsp = peakFreqPPM;
             maxpsh = peakFreqHz;
          }
   }
          }
/*Flips negative to choose most upfield peak of multiplet*/
for (i=0; i<numPeaks; i++)
   {
                  peakIntensity = getPeakIntensity(i);
                  if(peakIntensity>mintpp)
                  {
          peakFreqPPM = getPeakFreqPPM(i);
peakppmneg=peakFreqPPM*m;
maxpspneg=maxpsp*m;
          if (peakppmneg >= maxpspneg)
          {
             minpsp = peakFreqPPM;
          }
```

```
   }
        }
   freePeakList();
/*writes chemical shift into text document*/
ppmdif=maxpsp-minpsp;
cent=minpsp+ppmdif*0.5;
   fprintf(fref,"%f \n",cent);
/*Reads in for Glycolate*/
STOREPAR("f2p",f2pglyc)
STOREPAR("f1p",f1pglyc)
STOREPAR("mi",min)
PP
 numPeaks = readPeakList(PROCPATH(0));
maxips=0.0;
   maxpsh=0.0;
   for (i=0; i<numPeaks; i++)
   {
        peakIntensity = getPeakIntensity(i);
        peakFreqHz = getPeakFreqHz(i);
        peakFreqPPM = getPeakFreqPPM(i);
        if (peakIntensity > maxips)
        {
           maxips = peakIntensity;
           maxpsh = peakFreqHz;
           maxpsp = peakFreqPPM;
        }
   }
 /*Pick most downfield side of multiplet*/
   mintpp=maxips*ppsens;
maxpsp=0.0;
   for (i=0; i<numPeaks; i++)
   {
                peakIntensity = getPeakIntensity(i);
                if(peakIntensity>mintpp)
                {
        peakFreqPPM = getPeakFreqPPM(i);
        peakFreqHz = getPeakFreqHz(i);
        if (peakFreqHz >= maxpsh)
        {
           maxpsp = peakFreqPPM;
           maxpsh = peakFreqHz;
        }
   }
        }
/*Flips negative to choose most upfield peak of multiplet*/
for (i=0; i<numPeaks; i++)
   {
                peakIntensity = getPeakIntensity(i);
                if(peakIntensity>mintpp)
                {
        peakFreqPPM = getPeakFreqPPM(i);
peakppmneg=peakFreqPPM*m;
maxpspneg=maxpsp*m;
        if (peakppmneg >= maxpspneg)
        {
           minpsp = peakFreqPPM;
        }
   }
        }
   freePeakList();
```

```
/*writes chemical shift into text document*/
ppmdif=maxpsp-minpsp;
cent=minpsp+ppmdif*0.5;
   fprintf(fglyc,"%f \n",cent);
/*Reads in for Acetate*/
STOREPAR("f2p",f2pac)
STOREPAR("f1p",f1pac)
PP
 numPeaks = readPeakList(PROCPATH(0));
maxips=0.0;
   maxpsh=0.0;
   for (i=0; i<numPeaks; i++)
   {
          peakIntensity = getPeakIntensity(i);
          peakFreqHz = getPeakFreqHz(i);
          peakFreqPPM = getPeakFreqPPM(i);
          if (peakIntensity > maxips)
          {
             maxips = peakIntensity;
             maxpsh = peakFreqHz;
             maxpsp = peakFreqPPM;
          }
   }
 /*Pick most downfield side of multiplet*/
   mintpp=maxips*ppsens;
maxpsp=0.0;
   for (i=0; i<numPeaks; i++)
   {
                    peakIntensity = getPeakIntensity(i);
                    if(peakIntensity>mintpp)
                    {
          peakFreqPPM = getPeakFreqPPM(i);
          peakFreqHz = getPeakFreqHz(i);
          if (peakFreqHz >= maxpsh)
          {
             maxpsp = peakFreqPPM;
             maxpsh = peakFreqHz;
          }
   }
          }
/*Flips negative to choose most upfield peak of multiplet*/
for (i=0; i<numPeaks; i++)
   {
                    peakIntensity = getPeakIntensity(i);
                    if(peakIntensity>mintpp)
peakFreqPPM = getPeakFreqPPM(i);
peakppmneg=peakFreqPPM*m;
maxpspneg=maxpsp*m;
          if (peakppmneg >= maxpspneg)
          {
             minpsp = peakFreqPPM;
          }
   }
          }
   freePeakList();
/*writes chemical shift into text document*/
ppmdif=maxpsp-minpsp;
cent=minpsp+ppmdif*0.5;
   fprintf(fac,"%f \n",cent);
/*Reads in for sulfoacetate*/
```

```
STOREPAR("f2p",f2sfa)
STOREPAR("f1p",f1sfa)
PP
 numPeaks = readPeakList(PROCPATH(0));
maxips=0.0;
   maxpsh=0.0;
   for (i=0; i<numPeaks; i++)
   {
          peakIntensity = getPeakIntensity(i);
          peakFreqHz = getPeakFreqHz(i);
          peakFreqPPM = getPeakFreqPPM(i);
          if (peakIntensity > maxips)
          {
             maxips = peakIntensity;
             maxpsh = peakFreqHz;
             maxpsp = peakFreqPPM;
          }
   }
 /*Pick most downfield side of multiplet*/
   mintpp=maxips*ppsens;
maxpsp=0.0;
   for (i=0; i<numPeaks; i++)
   {
                    peakIntensity = getPeakIntensity(i);
                    if(peakIntensity>mintpp)
                    {
          peakFreqPPM = getPeakFreqPPM(i);
          peakFreqHz = getPeakFreqHz(i);
          if (peakFreqHz >= maxpsh)
          {
             maxpsp = peakFreqPPM;
             maxpsh = peakFreqHz;
          }
   }
          }
/*Flips negative to choose most upfield peak of multiplet*/
for (i=0; i<numPeaks; i++)
   {
                    peakIntensity = getPeakIntensity(i);
                    if(peakIntensity>mintpp)
                    {
          peakFreqPPM = getPeakFreqPPM(i);
peakppmneg=peakFreqPPM*m;
maxpspneg=maxpsp*m;
          if (peakppmneg >= maxpspneg)
          {
             minpsp = peakFreqPPM;
          }
   }
          }
   freePeakList();
/*writes chemical shift into text document*/
ppmdif=maxpsp-minpsp;
cent=minpsp+ppmdif*0.5;
   fprintf(fsfa,"%f \n",cent);
/*Reads in for 2-methylimidazole*/
STOREPAR("f2p",f2mi)
STOREPAR("f1p",f1mi)
PP
 numPeaks = readPeakList(PROCPATH(0));
```

```
maxips=0.0;
   maxpsh=0.0;
   for (i=0; i<numPeaks; i++)
   {
          peakIntensity = getPeakIntensity(i);
          peakFreqHz = getPeakFreqHz(i);
          peakFreqPPM = getPeakFreqPPM(i);
          if (peakIntensity > maxips)
          {
             maxips = peakIntensity;
             maxpsh = peakFreqHz;
             maxpsp = peakFreqPPM;
          }
   }
 /*Pick most downfield side of multiplet*/
   mintpp=maxips*ppsens;
maxpsp=0.0;
   for (i=0; i<numPeaks; i++)
   {
                    peakIntensity = getPeakIntensity(i);
                    if(peakIntensity>mintpp)
                    {
          peakFreqPPM = getPeakFreqPPM(i);
          peakFreqHz = getPeakFreqHz(i);
          if (peakFreqHz >= maxpsh)
          {
             maxpsp = peakFreqPPM;
             maxpsh = peakFreqHz;
          }
   }
          }
/*Flips negative to choose most upfield peak of multiplet*/
for (i=0; i<numPeaks; i++)
   {
                    peakIntensity = getPeakIntensity(i);
                    if(peakIntensity>mintpp)
                    {
          peakFreqPPM = getPeakFreqPPM(i);
peakppmneg=peakFreqPPM*m;
maxpspneg=maxpsp*m;
          if (peakppmneg >= maxpspneg)
          {
             minpsp = peakFreqPPM;
          }
   }
          }
   freePeakList();
/*writes chemical shift into text document*/
ppmdif=maxpsp-minpsp;
cent=minpsp+ppmdif*0.5;
   fprintf(fmi,"%f \n",cent);
v++;
RPROCNO(phpno)
SETCURDATA
}
END
fclose(fref);
fclose(fac);
fclose(fglyc);
fclose(fsfa);
```

```
fclose(fmi);
QUIT
```

A.6 AU program to extract integrals of acetate and reference from a CSI dataset (Bruker)

```
/*Bruker AU script for integrating a CSI dataset*/
/*CSI dataset should have been fully processed in phase-sensitive mode*/
/*Crude chemical shift referencing in F2 also helps*/
/*The script works on Bruker Topspin 3.6.2 but has not been tested on other versions*/
/*Create a 1D integral file using the wmisc command*/
/*This should contain only the reference compound and acetate*/
/*Copy and paste this name into the text box when asked, or paste into strcpy(text,"dmsim"); line 45 below, in place of dmsim */
/*dmsim is the name of the integral range file used in this script by default*/
/*The script extracts each row in turn into the empty procno requested (will overwrite existing contents!!!)*/
/*Working in this procno, the spectrum is referenced to the chemical shift reference supplied*/
/*Default values here apply to DMSO, DMSO signal is downfield of acetate (ordans=d)*/
/*Change right and left peak picking limits if any risk of a non-reference peak being included in the referencing procedure*/
/*Having referenced the spectrum, the requested integral file is read in and the integral values exported to .txt documents*/
/*Each .txt document is stored in the procno directory of the 2D CSI dataset*/
/*The script will also by default save rows 15, 32 and 50 into procnos of the same number*/
/*This is so the work of the script can be checked - spectra are correctly referenced and integrals correctly applied*/
/*This AU is not fully tested and comes without warranty.*/
/*Use kill command if all goes wrong*/
/*Matthew Wallace, 3/2022*/
/*University of East Anglia, matthew.wallace@uea.ac.uk*/
FILE    *fpnt,*fref,*fac;
char savans[8],dummystr[256],intdir[256],location[128],ordans[8];
float min=0;
double f2pdmso=2.65;
double f1pdmso=2.8;
double ref=2.72;
float ppsens=0.9;
double pc=0.1;
int m=-1;
int rn=1;
int sav1=15;
int sav2=32;
int sav3=50;
int steno;
double peakFreqHz, peakFreqPPM, peakIntensity, maxpsh, maxpsp, maxips, sf,sfn,sfo1,intgr;
double intnum,ppmdn,ppmup,intgrso1p,mintpp,minpsp,peakppmneg,cent,ppmdif,maxpspneg;
int i, numPeaks;
int np=64;
int row=1;
int v=1;
int wrpno=5;
int linenum=1;
int phpno=1;
strcpy(location,disk);
strcpy(savans,"y");
strcpy(text,"dmsim");
strcpy(ordans,"d");
GETCURDATA
```

```
steno=expno;
phpno=procno;
GETINT("Enter experiment number of 2D dataset",steno)
GETINT("Enter procno of 2D dataset",phpno)
REXPNO(steno)
RPROCNO(phpno)
SETCURDATA
FETCHPAR1("td",&np)
GETINT("Enter number of spectra in CSI image",np)
GETINT("Enter procno to extract rows into for integration (blank)",wrpno)
GETDOUBLE("Chemical shift reference right peak picking/ppm",f2pdmso)
GETDOUBLE("Chemical shift reference left peak picking/ppm",f1pdmso)
GETDOUBLE("Enter reference shift/ppm",ref)
GETFLOAT("Enter peak picking sensitivity factor",pc)
GETDOUBLE("Enter satellite sensitivity factor",ppsens)
GETSTRING("Save example spectra to check working OK? y/n",savans)
if(strcmp(savans,"y")==0)
{
GETINT("Enter 1st spectrum to save (put in pno of same#)",sav1)
GETINT("Enter 2nd spectrum to save (put in pno of same#)",sav2)
GETINT("Enter 3rd spectrum to save (put in pno of same#)",sav3)
}
GETSTRING("Which intrng file must be used?", text)
GETSTRING("Is the reference integral upfield (u) or downfield (d) of acetate?",ordans)
REXPNO(steno)
SETCURDATA
RPROCNO(phpno)
SETCURDATA
STOREPAR("mi",min)
STOREPAR("pc",pc)
STOREPAR("CURPRIN","Integrals.txt")
/*Create text files to hold integral data*/
/*Global scaling*/
STOREPAR("pscal",0)
if ((fref = fopen(PROCPATH("Reference integral.txt"),"wt")) == 0)
   STOPMSG("Cannot create difflist")
if ((fac = fopen(PROCPATH("Acetate integral.txt"),"wt")) == 0)
   STOPMSG("Cannot create difflist")
   /*No go through each row in turn*/
TIMES(np)
{
RSR(v,wrpno)
RPROCNO(wrpno)
SETCURDATA
sprintf(intdir,"%s\\%s\\%i\\pdata\\%i\\integrals.txt",location,name,expno,procno);
/*Reference the spectrum.*/
/*If the reference peak is split, the program will reference based on the average shift*/
/*of the two peaks in the requested peak picking range which are within ppsens of the largest peak
found*/
STOREPAR("f2p",f2pdmso)
STOREPAR("f1p",f1pdmso)
PP
 numPeaks = readPeakList(PROCPATH(0));
maxips=0.0;
   maxpsh=0.0;
   for (i=0; i<numPeaks; i++)
   {
         peakIntensity = getPeakIntensity(i);
         peakFreqHz = getPeakFreqHz(i);
         peakFreqPPM = getPeakFreqPPM(i);
```

```
        if (peakIntensity > maxips)
        {
           maxips = peakIntensity;
           maxpsh = peakFreqHz;
           maxpsp = peakFreqPPM;
        }
   }
   /*Pick most downfield side of multiplet*/
   mintpp=maxips*ppsens;
maxpsp=0.0;
   for (i=0; i<numPeaks; i++)
   {
                peakIntensity = getPeakIntensity(i);
                if(peakIntensity>mintpp)
                {
        peakFreqPPM = getPeakFreqPPM(i);
        peakFreqHz = getPeakFreqHz(i);
        if (peakFreqHz >= maxpsh)
        {
           maxpsp = peakFreqPPM;
           maxpsh = peakFreqHz;
        }
   }
        }
/*Flips negative to choose most upfield peak of multiplet*/
for (i=0; i<numPeaks; i++)
   {
                peakIntensity = getPeakIntensity(i);
                if(peakIntensity>mintpp)
                {
        peakFreqPPM = getPeakFreqPPM(i);
peakppmneg=peakFreqPPM*m;
maxpspneg=maxpsp*m;
        if (peakppmneg >= maxpspneg)
        {
           minpsp = peakFreqPPM;
        }
   }
        }
   freePeakList();
/*References spectrum*/
ppmdif=maxpsp-minpsp;
cent=minpsp+ppmdif*0.5;
FETCHPAR("sf",&sf)
   sfn=sf+(cent-ref)*sf/(1e6);
   STOREPAR("sf",sfn)
/*Integrate spectra, read numbers and store integrals in text file*/
        RMISC("intrng", text)
        LI
fpnt=fopen(intdir, "r");
fgets(dummystr, sizeof(dummystr), fpnt);
while (fgets(dummystr, sizeof(dummystr), fpnt) != NULL)
{
        /*Need to selectively elimate rows, then scan for numbers*/
if(linenum>=5)
{
        (void) sscanf(dummystr,"%lf %lf %lf %lf",
                        &intnum,&ppmdn,&ppmup,&intgr);
                        /*DMSO is first, then acetate*/
                                if(linenum==5)
```

```
                                {
                                        if(strcmp(ordans,"d")==0)
                                        {
                                fprintf(fref,"%f\n",intgr);
                                        }
                                                if(strcmp(ordans,"u")==0)
                                        {
                                fprintf(fac,"%f\n",intgr);
                                        }
                                }
                                        if(linenum==6)
                                {
                        if(strcmp(ordans,"d")==0)
                                        {
                                fprintf(fac,"%f\n",intgr);
                                        }
                                                if(strcmp(ordans,"u")==0)
                                        {
                                fprintf(fref,"%f\n",intgr);
                                        }
                                }
                                intnum=0;
                                ppmdn=0;
                                ppmup=0;
                                intgr=0;
                                linenum++;
}
        else
{
                                linenum++;
}
}
linenum=1;
fclose(fpnt);
        /*Save spectra in procnos if requested (will overwrite!!)*/
        if(strcmp(savans,"y")==0)
{
if(v==sav1)
{
WRP(sav1)
}
if(v==sav2)
{
WRP(sav2)
}
if(v==sav3)
{
WRP(sav3)
}
}
v++;
RPROCNO(phpno)
SETCURDATA
}
END
fclose(fac);
fclose(fref);
QUIT
```

A.7 AU program to extract areas of acetate and reference from a CSI dataset by lineshape deconvolution (Bruker)

```
/*Bruker AU script for deconvoluting acetate and reference signals in a CSI dataset*/
/*use if overlap with other signals prevents conventional integration*/
/*CSI dataset should have been fully processed in phase-sensitive mode*/
/*Crude chemical shift referencing in F2 also helps*/
/*Open the 2D dataset in Topspin as the active dataset before running this AU*/
/*The script works on Bruker Topspin 3.6.2 but has not been tested on other versions*/
/*Goes through a 2D image and writes file of peak areas*/
/*Will find the area of the most intense peak in specified range for reference and acetate*/
/*After deconvolution, peak areas will be exported to text files stored in procno folder of dataset*/
/*Default peak picking parameters are for DMSO as reference*/
/*This AU is not fully tested and comes without warranty.*/
/*Use kill command if all goes wrong*/
/*Matthew Wallace, 3/2022*/
/*University of East Anglia, matthew.wallace@uea.ac.uk*/
#include <inc/exptUtil>
FILE    *fac,*fref,*frat,*fpnt;
char dconvdir[256],location[256],dummystr[256];
double min=0.1;
double minincac=0.1;
double minincvac;
double minincref=0.1;
double minincvref;
double f2pac=1.8;
double f1pac=2.2;
double f2pref=2.65;
double f1pref=2.8;
double ppsens=0.45;
double fhz=0.0;
double hgt=0.0;
float pc=1;
float azfw=0.5;
double m=-1;
int wrpno=1;
int v=1;
double ppm,wppm,whz,area,chisq,minvref,minvac,ac,ref,refarea,rat,acarea;
int pknum=0;
int steno;
int np=64;
int linenum=1;
int lincnt=6;
int nmpex=1;
int nmiso=1;
int nmisocnt=1;
int nmisoit=0;
int phpno=procno;
int extrctno=phpno+6;
int i, numPeaks,nmpexv;
strcpy(location,disk);
GETCURDATA
steno=expno;
GETSTRING("Enter location of dataset",location)
GETINT("Enter experiment number of 2D dataset",steno)
GETINT("Enter procno of 2D dataset",phpno)
REXPNO(steno)
RPROCNO(phpno)
SETCURDATA
```

```
FETCHPAR1("td",&np)
GETINT("Enter number of gradient values",np)
GETDOUBLE("Right acetate peak picking limit/ppm",f2pac)
GETDOUBLE("Left acetate peak picking limit/ppm",f1pac)
GETDOUBLE("Right reference peak picking limit/ppm",f2pref)
GETDOUBLE("Left reference peak picking limit/ppm",f1pref)
GETFLOAT("Enter peak picking sensitivity factor",pc)
GETDOUBLE("Enter satellite sensitivity factor",ppsens)
wrpno=phpno+4;
GETINT("Enter procno to perform deconvolution (blank)",wrpno)
        if ((fac = fopen(PROCPATH("Acetate area.txt"),"wt")) == 0)
   STOPMSG("Cannot create file")
   if ((fref = fopen(PROCPATH("Reference area.txt"),"wt")) == 0)
   STOPMSG("Cannot create file")
   if ((frat = fopen(PROCPATH("Ratio of integral and reference.txt"),"wt")) == 0)
   STOPMSG("Cannot create file")
   TIMES(np)
{
RSR(v,wrpno)
RPROCNO(wrpno)
SETCURDATA
STOREPAR("mi",minvac)
STOREPAR("pc",pc)
STOREPAR("azfw",azfw)
/*Finds max peak in acetate region*/
STOREPAR("f2p",f2pac)
STOREPAR("f1p",f1pac)
/*Global scaling*/
STOREPAR("pscal",0)
minincvac=minincac;
nmpexv=nmpex;
PP
 numPeaks = readPeakList(PROCPATH(0));
 freePeakList();
while(numPeaks!=nmpexv)
{
STOREPAR("mi",minvac)
PP
/*Checks that expected number of peaks has been found, if not increases PC and repeats peak picking*/
/*If are in fact less than expected number will just pick noise, but this does not matter*/
numPeaks = readPeakList(PROCPATH(0));
freePeakList();
if(numPeaks>nmpex)
{
minvac=minvac+minincvac;
nmisocnt++;
if(nmisocnt>=20)
{
minincvac=minincvac/10;
nmisocnt=0;
/*Exit loop 0.001 seems sensible limit*/
if(minincvac<0.001)
{
nmpexv=numPeaks;
}
}
}
if(numPeaks<nmpex)
{
minvac=minvac-minincvac;
```

```
nmisocnt++;
if(nmisocnt>=20)
{
minincvac=minincvac/10;
nmisocnt=0;
/*Exit loop*/
if(minincvac<0.001)
{
nmpexv=numPeaks;
}
}
}
}
 freePeakList();
nmisocnt=1;
nmpexv=nmpex;
/*Having found peaks, now write peak positions into peaklist file for deconvolution*/
DATASET2(name,expno,procno+1000,disk,user)
PPP
    /*Now do deconvolution*/
MDCON
sprintf(dconvdir,"%s/%s/%i/pdata/%i/integrals.txt",location,name,expno,procno);
fpnt=fopen(dconvdir, "r");
fgets(dummystr, sizeof(dummystr), fpnt);
while (fgets(dummystr, sizeof(dummystr), fpnt) != NULL)
{
        /*Need to selectively eliminate rows from dconpeaks, then scan for numbers*/
if(linenum>=9)
{
(void) sscanf(dummystr,"%lf %lf %lf %lf %lf %lf %lf",
                        &ppm,&fhz,&wppm,&whz,&hgt,&area,&chisq);

if(linenum==10)
{
fprintf(fac,"%f\n",area);
acarea=area;
}
linenum++;
}
else
{
                        linenum++;
}
}
linenum=1;
        fclose(fpnt);
        /*Now do same for ref region*/
STOREPAR("f2p",f2pref)
STOREPAR("f1p",f1pref)
minincvref=minincref;
nmpexv=nmpex;
PP
 numPeaks = readPeakList(PROCPATH(0));
 freePeakList();
while(numPeaks!=nmpexv)
{
STOREPAR("mi",minvref)
PP
/*Checks that expected number of peaks has been found, if not increases PC and repeats peak picking*/
/*If are in fact less than expected number will just pick noise, but not matter*/
```

```
numPeaks = readPeakList(PROCPATH(0));
freePeakList();
if(numPeaks>nmpex)
{
minvref=minvref+minincvref;
nmisocnt++;
if(nmisocnt>=20)
{
minincvref=minincvref/10;
nmisocnt=0;
/*Exit loop 0.001 seems sensible limit*/
if(minincvref<0.001)
{
nmpexv=numPeaks;
}
}
}
if(numPeaks<nmpex)
{
minvref=minvref-minincvref;
nmisocnt++;
if(nmisocnt>=20)
{
minincvref=minincvref/10;
nmisocnt=0;
/*Exit loop*/
if(minincvref<0.001)
{
nmpexv=numPeaks;
}
}
}
}
  freePeakList();
nmisocnt=1;
nmpexv=nmpex;
PPP
    /*Now do deconvolution*/
MDCON
sprintf(dconvdir,"%s/%s/%i/pdata/%i/integrals.txt",location,name,expno,procno);
fpnt=fopen(dconvdir, "r");
fgets(dummystr, sizeof(dummystr), fpnt);
while (fgets(dummystr, sizeof(dummystr), fpnt) != NULL)
{
          /*Need to selectively elimate rows from dconpeaks, then scan for numbers*/
if(linenum>=9)
{
(void) sscanf(dummystr,"%lf %lf %lf %lf %lf %lf %lf",
                              &ppm,&fhz,&wppm,&whz,&hgt,&area,&chisq);
if(linenum==10)
{
fprintf(fref,"%f\n",area);
refarea=area;
}
linenum++;
}
else
{
                              linenum++;
}
```

```
}
linenum=1;
	fclose(fpnt);
	rat=acarea/refarea;
	fprintf(frat,"%f\n",rat);
	RPROCNO(phpno)
	SETCURDATA
	v++;
}
	END
	fclose(frat);
	fclose(fac);
	fclose(fref);
	QUIT
```

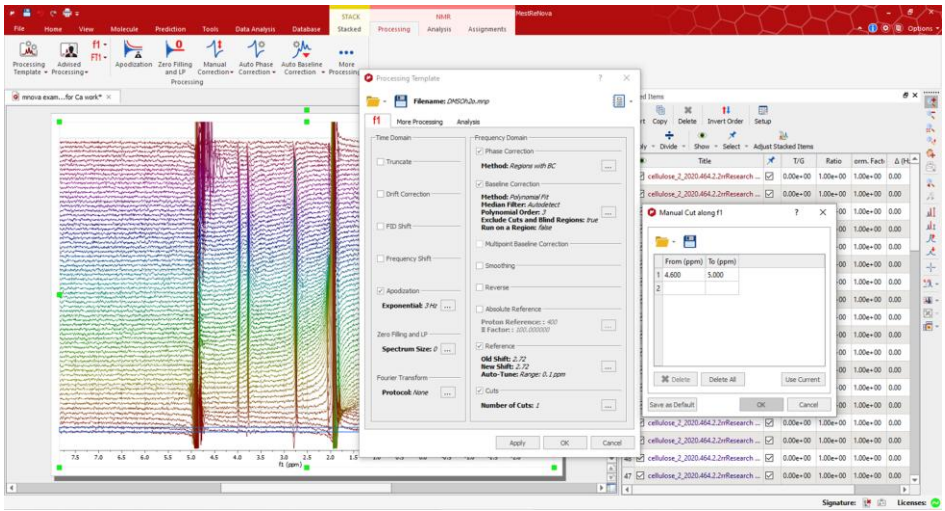
A.8 AU program to shim and find water suppression frequency when running under IconNMR (Bruker)

```
/*Bruker AU script for shimming, finding water supression frequency*/
/*and running CSI experiment through IconNMR*/
/*create macro for reading in default shim file*/
/*edmac rshim, then in macro text: rsh nameofshimfile*/
/*create macro for topshim on 1h:*/
/*edmac t1h, then in macro text: topshim 1h lockoff o1p=4.85 ordmax=3 convcomp*/
/*create a 1 scan proton parameter set (low rg) to find o1p of water, called H2O_SS*/
/*Remember to change peak picking regions in this PAR set to cover the expected range for water signal*/
/*create a CSI parameter set called 1hcsi (or equivalent), this runs the csi experiment*/
/*set parameter AUNM in this par set to the name of this script*/
/*Then add the par set onto your list of experiments in the IconNMR config menu*/
/*when running through Icon, disable locking and shimming by choosing the "IconNMR not responsible" option*/
/*The script works on Bruker Topspin 3.6.2 but has not been tested on other versions*/
/*This AU is not fully tested and comes without warranty.*/
/*Matthew Wallace, 5/2022*/
/*University of East Anglia, matthew.wallace@uea.ac.uk*/
 float peakFreqHz, peakFreqPPM, peakIntensity, maxpsh, maxpsp, maxips;
    char path[PATH_MAX];
double sf, sfo1;
int pscal_save, i, numPeaks;

GETCURDATA
/*Can set numbr of scans for water suppression in ICON*/
RPAR("H2O_SS","all")
/*STOREPAR("RG",1)*/
XMAC("rshim")
sleep(60);
XMAC("t1h")
sleep(90);
ZG
ERRORABORT
EF
ERRORABORT
APK
/*SREF*/
FETCHPAR("PSCAL",&pscal_save)
STOREPAR("PSCAL",0)
PP
```

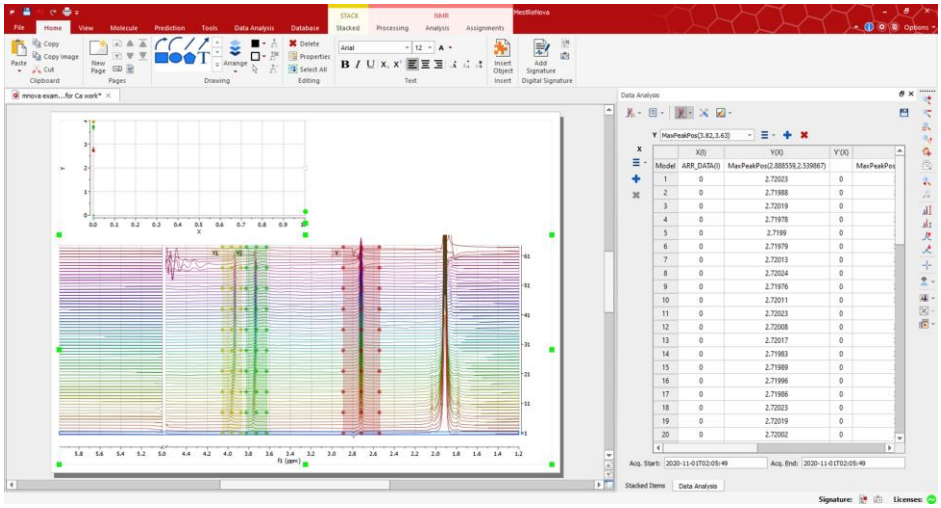
```
ERRORABORT
  strcpy(path, PROCPATH(0));
    numPeaks = readPeakList(path);

    maxips=0.0;
    maxpsh=0.0;
    for (i=0; i<numPeaks; i++)
    {
            peakIntensity = getPeakIntensity(i);
            peakFreqHz = getPeakFreqHz(i);
            peakFreqPPM = getPeakFreqPPM(i);
            if (peakIntensity > maxips)
            {
               maxips = peakIntensity;
               maxpsh = peakFreqHz;
               maxpsp = peakFreqPPM;
            }
    }
    freePeakList();

    FETCHPAR("SF",&sf);
    sfo1 = sf + maxpsh * 1.0e-6;
    STOREPAR("SFO1",sfo1);
    SETCURDATA
/*RGA can go wrong for water suppression*/
RPAR("1hcsi","all")
STOREPAR("SFO1",sfo1)
ZG
QUIT
```

A.9 Processing procedure for Mnova 14.2.0

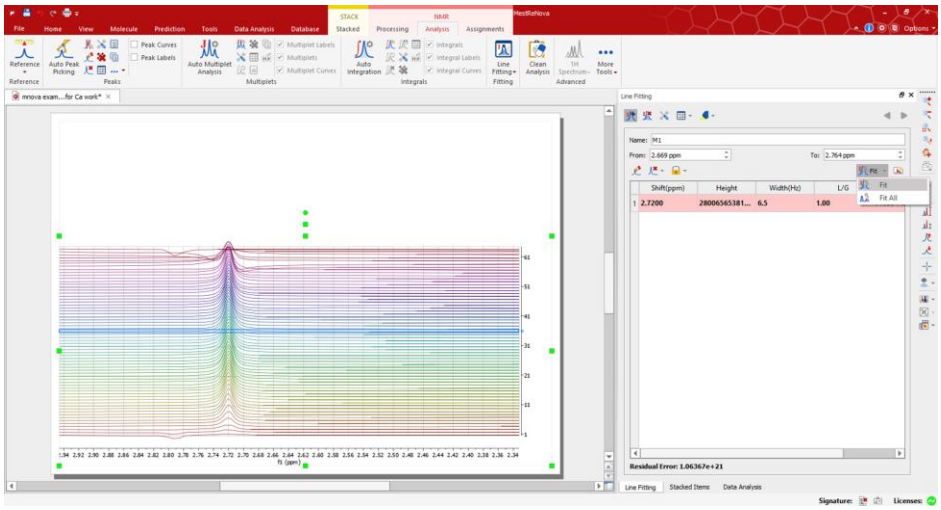
To import a CSI dataset into Mnova, run processing AU. Open 2rr file in procno directory of NMR dataset in Mnova. Set up and run a processing template as shown below to automatically phase, reference and baseline correct the data.



Extract peak positions and integrals using Data Analysis module.



To extract peak areas, create line fitting region and fit all rows in the CSI dataset. Then run script below to export areas:



```
/*****************************************************************************************
For finding area of singlet peak in stacked plot from a Chemical Shift Imaging (CSI) dataset
Matthew Wallace, 3/2022
University of East Anglia, matthew.wallace@uea.ac.uk

Define line fitting area with new fit region (clear all previous line fitting regions)
Click fit to fit all spectra in the stacked CSI dataset
Save this script in Mnova as PeakArea, and Run this script

Script will find the biggest area of a peak in a spectrum within the defined region and put in txt document
Area can then be copied and pasted into Excel

Based on Mnova script exportFitRegions
Copyright (C) 2014 Mestrelab Research S.L. All rights reserved.

This file is part of the Mnova scripting toolkit.

Authorized users of Mnova Software may use this file freely, but this file is provided AS IS
with NO WARRANTY OF ANY KIND, INCLUDING THE WARRANTY OF DESIGN,
MERCHANTABILITY AND FITNESS
FOR A PARTICULAR PURPOSE.
*****************************************************************************************/
```

```
/*globals settings, Dir, FileDialog, File, TextStream, Application, NMRSpectrum, print, Peak, MnUi*/
/*jslint plusplus: true, indent: 4*/

function PeakArea() {
	"use strict";

	function fitRegionToStream(aFitRegion, aFileStream, aNMRPeaks) {
/*		aFileStream.writeln("Area");*/
		var p, peak, tst, big
			tst=0;
			big=0;
			fitPeaksIds = aFitRegion.peaks;
		for (p = 0; p < fitPeaksIds.length; p++) {
			peak = new Peak(aNMRPeaks.byId(fitPeaksIds[p]));
			tst=peak.integral;
			if(tst>big)
			{
			big=peak.integral;
			}
			}
			aFileStream.write(big, "\n");
	}

   var fout, sout, spc, peakList, fitRegions, fr, oldCurSpecIndex, i,
		dirSettingsKey = "PeakArea/LastDir",
		saveDir = settings.value(dirSettingsKey, Dir.home()),
		dw = Application.mainWindow.activeDocument,
		spectra = dw.itemCount("NMR Spectrum"),
		specIndex = 0,
		fileName = FileDialog.getSaveFileName("ASCII Files (*.txt)", "", saveDir);

   if (!fileName.length) {
		return;
	}

   fout = new File(fileName);
   settings.setValue(dirSettingsKey, fout.absDirPath);
   if (!fout.open(File.WriteOnly)) {
		throw "Impossible to open file";
	}
   sout = new TextStream(fout);
   sout.precision = 10;

   while (specIndex < spectra) {
		spc = new NMRSpectrum(dw.item(specIndex, "NMR Spectrum"));
		specIndex++;
		if (!spc.isValid()) {
			throw "Invalid Spectrum";
		}
		oldCurSpecIndex = spc.curSpecIndex;
		for (i = 0; i < spc.specCount; i++) {
			spc.curSpecIndex = i;
			peakList = spc.peaks();
			fitRegions = spc.fitRegions();
			print(fitRegions);
			for (fr = 0; fr < fitRegions.length; fr++) {
				fitRegionToStream(fitRegions[fr], sout, peakList);
			}
		}
		spc.curSpecIndex = oldCurSpecIndex;
```

```
    }
    fout.close();
}

if (this.MnUi && MnUi.scripts_nmr) {
        MnUi.scripts_nmr.scripts_nmr_ExportASCIIFitRegions = PeakArea;
}
```

Appendix Chapter 2

B.1 Derivation of NMR modified Henderson-Hasselbalch equation

The use of NMR spectroscopy to determine the pK_a values of molecules is well-established and is discussed extensively.[89, 111, 234] The derivation is based on the facts one is that there is a dynamic equilibrium between an acid and conjugate base which can be expressed via standard Henderson-Hasselbalch equation (equation B.1) and that the chemical shift of an NMR such a compound if it is NMR active is in an average of the chemical shifts of the compound in its acid and conjugate base states (due to rapid exchange relative to NMR timescale).

$$pH = pKa + \log_{10}(\frac{[A^-]}{[HA]}) \tag{B.1}$$

In the derivation concentration fractions of base and acid for simplicity will be used instead. Given that chemical shifts are determined relative to the shift of a reference species (this is because the reference species itself possibly participates acid/base equilibrium reaction and hence needs to be taken into account). Assuming the acid/base reaction is at thermodynamic equilibrium the dissociation constant of the reference compound K_{ref} and indicator compound K_{ind} can be represented with the following equations:

$$K_{ref} = \frac{[H^+]X_{base}^{ref}}{X_{acid}^{ref}} \tag{B.2}$$

$$K_{ind} = \frac{[H^+]X_{base}^{ind}}{X_{acid}^{ind}} \tag{B.3}$$

Where H^+ is the concentration of protons, X_{base}^{ref} is fraction of base form of reference species, X_{acid}^{ref} is fraction of acid form of reference species, X_{base}^{ind} fraction of base form of indicator species and X_{acid}^{ind} is fraction of acid form of indicator species. Given that fractions add up to unity fraction can be represented by equation:

$$X_{acid}^{ref} = \frac{[H^+]X_{base}^{ref}}{K_{ref}} = \frac{[H^+](1 - X_{acid}^{ref})}{K_{ref}} = \frac{[H^+]}{[H^+] + K_{ref}} \tag{B.4}$$

$$X_{acid}^{ind} = \frac{[H^+]X_{base}^{ind}}{K_{ind}} = \frac{[H^+](1 - X_{acid}^{ind})}{K_{ind}} = \frac{[H^+]}{[H^+] + K_{ind}} \tag{B.5}$$

The exchange averaged chemical shift δ_{ave} for the reference and for the indicator species is:

$$\delta_{ave}^{ref} = X_{acid}^{ref}\delta_{acid}^{ref} + X_{base}^{ref}\delta_{base}^{ref} \tag{B.6}$$

$$\delta_{ave}^{ind} = X_{acid}^{ind}\delta_{acid}^{ind} + X_{base}^{ind}\delta_{base}^{ind} \tag{B.7}$$

The chemical shift of the base form of the reference compound is taken to be 0.00 PPM by definition (given that it is the reference in its non-protonated form) Thus equation of the δ_{ave}^{ref} can be simplified to be:

$$\delta_{ave}^{ref} = X_{acid}^{ref}\delta_{acid}^{ref} \tag{B.8}$$

The observed chemical shift is defined as the difference between the indicator shift and the reference shift:

$$\delta_{obs} = X_{acid}^{ind}\delta_{acid}^{ind} + X_{base}^{ind}\delta_{base}^{ind} - X_{acid}^{ref}\delta_{acid}^{ref} \tag{B.9}$$

Given the formula relating to fraction equation of δ_{obs} can be reformulated to be:

$$\delta_{obs} = \frac{[H^+]}{K_{ind} + [H^+]}\left(\delta_{acid}^{ind} - \delta_{base}^{ind}\right) + \delta_{base}^{ind}\frac{[H^+]}{K_{ref} + [H^+]}\delta_{acid}^{ref} \tag{B.10}$$

Under strongly basic conditions relative to K_{ref} the right hand most term dependent upon δ_{acid}^{ref} vanishes and it simplifies to:

$$\delta_{obs} = \frac{[H^+]}{K_{ind} + [H^+]}\left(\delta_{acid}^{ind} - \delta_{base}^{ind}\right) \tag{B.11}$$

This can be reformulated to the NMR modified Henderson-Hasselbalch equation

$$pH = \log_{10}\left(\frac{\delta_{obs} - \delta_{acid}^{ind}}{\delta_{base}^{ind} - \delta_{obs}}\right) + pK_a \tag{B.12}$$

Where δ_{base}^{ind} and δ_{acid}^{ind} are named δ_L and δ_H throughout the thesis.

B.2 Estimation of uncertainty in the determination of pK_a

The uncertainty in $\frac{dpH}{d\delta_{obs}}$ is obtained by propagation of uncertainty analysis of Equation 2.12:

$$\Delta\frac{dpH}{d\delta_{obs}} = \sqrt{\left(\frac{\partial\left(\frac{dpH}{d\delta_{obs}}\right)}{\partial\delta_{obs}}\right)^2\Delta{\delta_{obs}}^2 + \left(\frac{\partial\left(\frac{dpH}{d\delta_{obs}}\right)}{\partial\delta_H}\right)^2\Delta{\delta_H}^2 + \left(\frac{\partial\left(\frac{dpH}{d\delta_{obs}}\right)}{\partial\delta_L}\right)^2\Delta{\delta_L}^2} \tag{B.13}$$

$$\frac{dpH}{d\delta_{obs}} = \frac{\delta_H - \delta_L}{\ln(10)(\delta_{obs} - \delta_L)(\delta_{obs} - \delta_H)} \tag{B.14}$$

$$(\ln(10)\ \frac{dpH}{d\delta_{obs}})^{-1}\ = \frac{{\delta_{obs}}^2}{\delta_H - \delta_L} + \delta_{obs} + \frac{\delta_H \delta_L}{\delta_H - \delta_L} \tag{B.15}$$

$$\frac{\partial(\frac{dpH}{d\delta_{obs}})}{\partial\delta_{obs}}\ = \frac{(\delta_L - \delta_H)(2\delta_{obs} - \delta_L - \delta_H)}{\ln(10)(\delta_{obs} - \delta_L)^2(\delta_{obs} - \delta_H)^2} \tag{B.16}$$

$$\frac{\partial(\frac{dpH}{d\delta_{obs}})}{\partial\delta_H}\ = \frac{1}{\ln(10)(\delta_{obs} - \delta_H)^2} \tag{B.17}$$

$$\frac{\partial(\frac{dpH}{d\delta_{obs}})}{\partial\delta_L}\ = -\frac{1}{\ln(10)(\delta_{obs} - \delta_H)^2} \tag{B.18}$$

The uncertainty in $(\ln(10)\frac{dpH}{d\delta_{obs}})^{-1}$ is obtained by propagation of uncertainty analysis of Equation 2.13:

$$\Delta(\ln(10)\ \frac{dpH}{d\delta_{obs}})^{-1} = \sqrt{(\frac{\partial((\ln(10)\ \frac{dpH}{d\delta_{obs}})^{-1})}{\partial\delta_{obs}})^2{\Delta\delta_{obs}}^2 + (\frac{\partial((\ln(10)\ \frac{dpH}{d\delta_{obs}})^{-1})}{\partial\delta_H})^2{\Delta\delta_H}^2 + (\frac{\partial((\ln(10)\ \frac{dpH}{d\delta_{obs}})^{-1})}{\partial\delta_L})^2{\Delta\delta_L}^2} \tag{B.19}$$

$$\frac{\partial(\ln(10)\frac{dpH}{d\delta_{obs}})^{-1}}{\partial\delta_{obs}}\ = \frac{2\delta_{obs}}{\delta_H - \delta_L} + 1 \tag{B.20}$$

$$\frac{\partial(\ln(10)\frac{dpH}{d\delta_{obs}})^{-1}}{\partial\delta_H}\ = \frac{{\delta_{obs}}^2 + {\delta_L}^2}{(\delta_H - \delta_L)^2} \tag{B.21}$$

$$\frac{\partial(\ln(10)\frac{dpH}{d\delta_{obs}})^{-1}}{\partial\delta_L}\ = \frac{{\delta_{obs}}^2 + {\delta_H}^2}{(\delta_H - \delta_L)^2} \tag{B.22}$$

To estimate the uncertainty in the chemical shifts a homogenous solution of 0.1% 1,4-dioxane with 0.2 mM DSS was run and a 2D CSI ^{1}H NMR experiment was run and the referenced chemical shift of 1,4-dioxane was measured for 59 slices and the standard deviation was 0.0007 ppm thus $\Delta\delta_{obs}$, $\Delta\delta_L$, $\Delta\delta_H$ are taken as 0.0007 ppm in this work.

Error from derivative approximation was estimated by doing *in silico* work with compounds with known pK_a, δ_L, δ_H and running simulations on what the derivative values would be if done analytically using equation 2.10 (and what the value will be using the numerical derivative method). The *in silico* work also included simulating what the derivative values will be if the if the distance between the two points used for determining the numerical derivative are 0.04 ppm and 0.12 ppm. Subsequently the derivative values for the analytic, the 0.04 ppm and 0.12 ppm were used to determine pK_a, δ_L, δ_H using the method previously described. This error analysis was done with respect to multiple compounds with varied pK_a and limiting chemical

shift values to check for any differences in error margins between compounds shown in figure B.1 is example of analysis on 2,6-lutidine. Unsurprisingly the analytically derived derivative allowed for exact determination of pK_a and limiting chemical shifts of all compounds. For the 0.04 ppm numerical derivative the standard deviation with respect to literature pK_a and limiting chemical shift values of all compounds were 0.0097 units for pK_a, 0.00019 ppm for δ_L and 0.0032 ppm for δ_H. With regards to the 0.120 ppm numerical derivative the standard deviation with respect to literature pK_a and limiting chemical shift values were 0.048 units for pK_a, 0.0351 ppm for δ_L and 0.0361 ppm for δ_H.

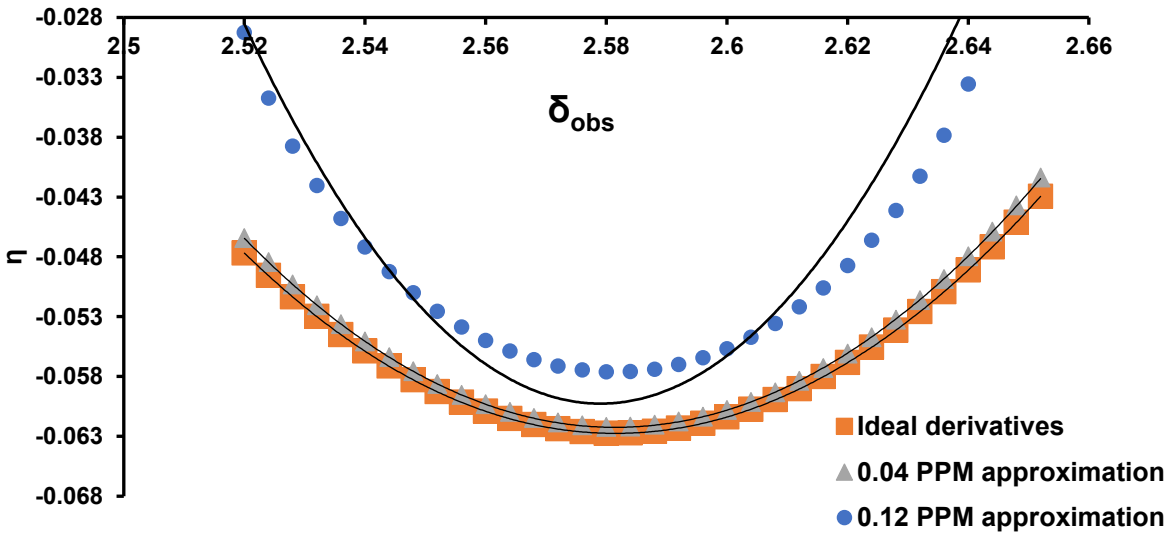


Figure B.1. Plot of η versus δ_{obs} of 2,6-lutidine and their respective fitting comprising ideal derivatives (orange, ideal derivatives being ones calculated directly from equation 2.10), numerical derivatives where the distance between the two points used is 0.04 PPM (grey), numerical derivatives where the distance between the two points used is 0.12 PPM (blue)

B.3 Comparison of derivative curves obtained with different analyte resonances

Experiments were performed to determine the pK_a of 2-MI and 4-CN using different respective resonances to verify robustness of method when the chemical shift is further from the donating site by measuring the chemical shift of 4-CN adjacent to the hydroxyl group and the one further away and comparing the pK_a determination. Additionally, experiments were run with 2-MI to check whether measuring chemical shift of a proton attached to an aromatic group compared to methyl group impacts pK_a determination. Analysis was using the standard and alternative method to determine robustness from both.

With regards to 4-CN experiment negligible impact was found between the two resonances pK_a fitting (see figures and table below), possibly due to them being part of a conjugated π-bond system, making the impact of the change in electron distribution shared and similar results were obtained in regards to the alternative method for measuring pK_a.

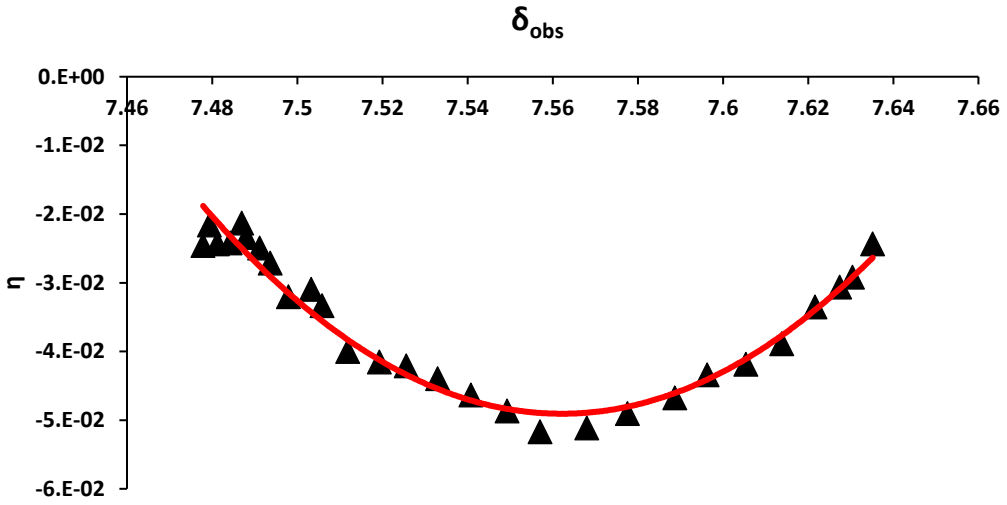


Figure B.2. Plot of η versus δ_{obs} of 4-CN downfield peak with its respective fitting

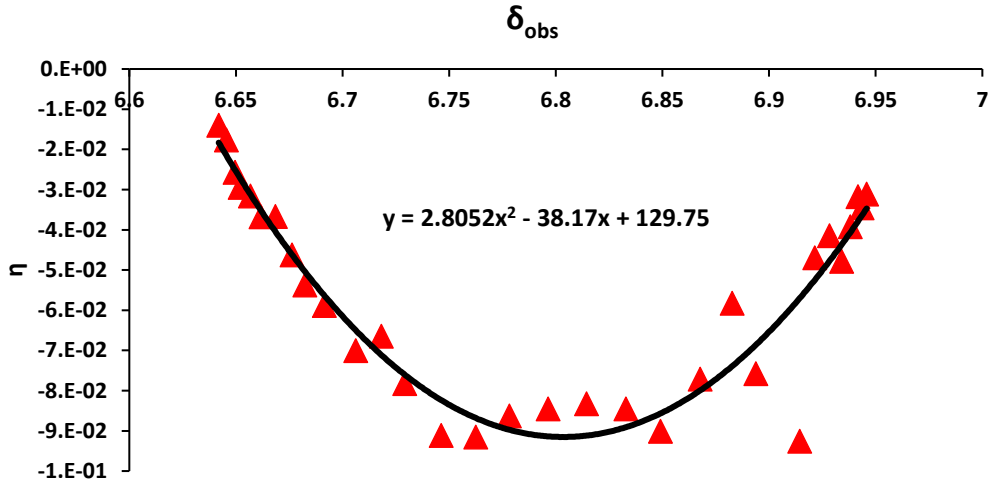


Figure B.3. Plot of η versus δ_{obs} of 4-CN upfield peak with its respective fitting

Table B.1. Fitted δ_H, δ_L, and pK_a of 4-CN

4-CN resonance	Fitted δ_H	Fitted δ_L	Fitted pK_a
Downfield	7.67	7.45	7.99
Upfield	6.98	6.62	8.01

With the 2-MI experiments both the methyl and aromatic protons gave pK_a values close to literature however the aromatic proton was slightly closer to the literature value (see figure and table below). This could be due to increased sensitivity of the aromatic proton to deprotonation of acid group in the compound because the negative charge introduced upon deprotonation is often delocalized across the aromatic π-system which can significantly alter the electron density throughout the molecule.

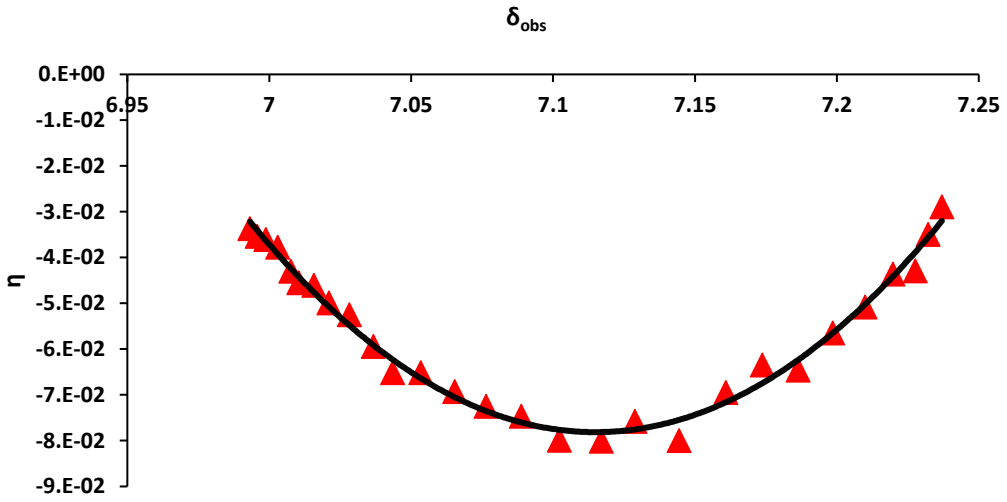


Figure B.4. Plot of η versus δ_{obs} of 2-MI aromatic peak with its respective fitting

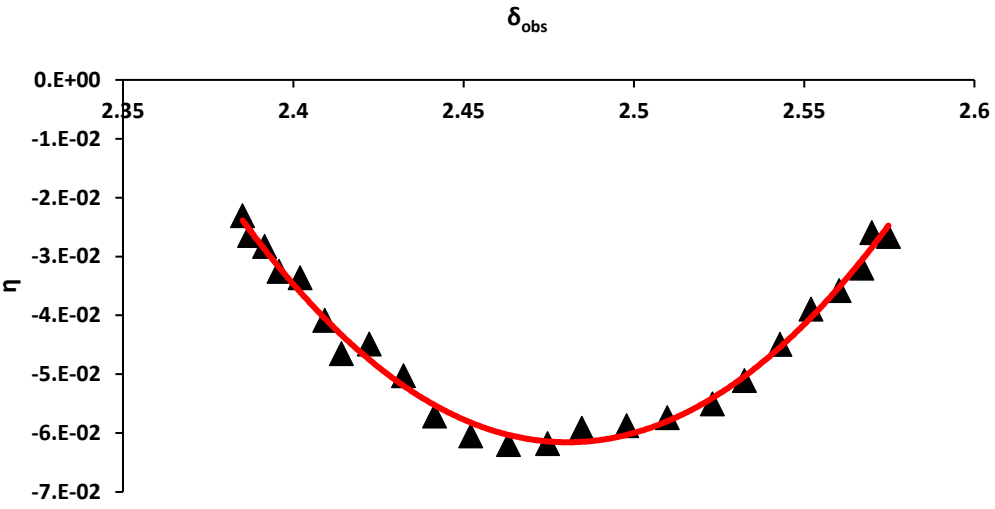


Figure B.5. Plot of η versus δ_{obs} of 2-MI methyl peak with its respective fitting

Table B.2. Fitted δ_H, δ_L, and pK_a of 2-MI

2-MI resonance	Fitted δH	Fitted δL	Fitted pK_a
Aromatic	7.27	6.96	8.05
Methyl	2.60	2.36	8.11

B.4 Derivative-based fitting plots of analytes

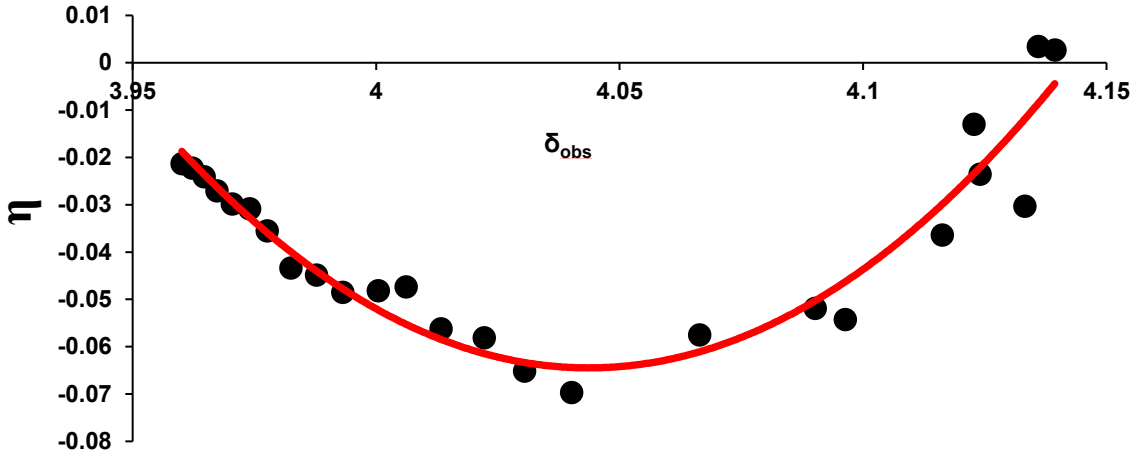


Figure B.6. Plot of η versus δ_{obs} for glycolic acid

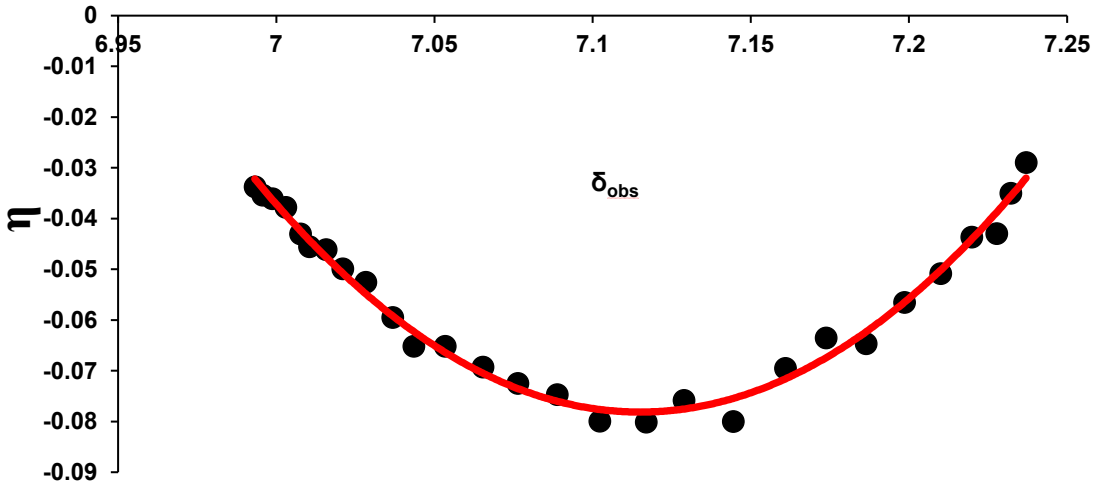


Figure B.7. Plot of η versus δ_{obs} for 2-methylimidazole

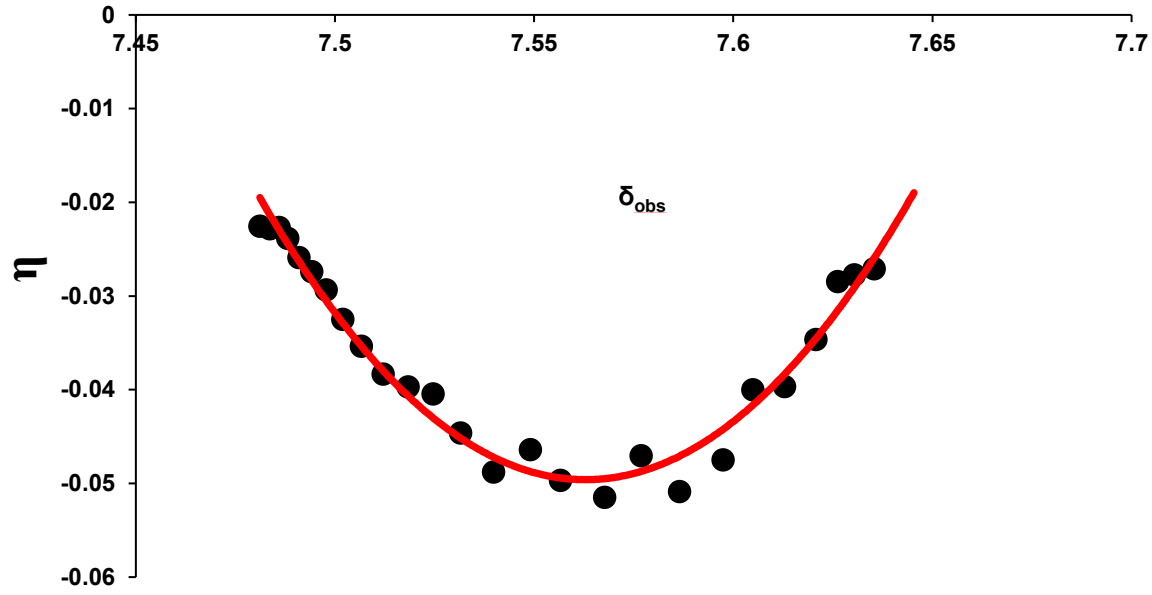


Figure B.8. Plot of η versus δ_{obs} for 4-cyanophenol

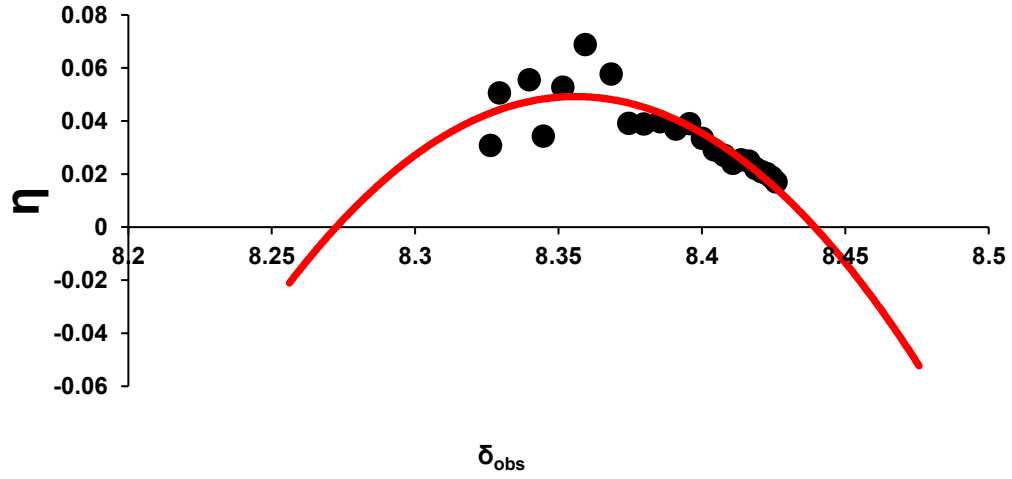


Figure B.9. Plot of η versus δ_{obs} for formic acid

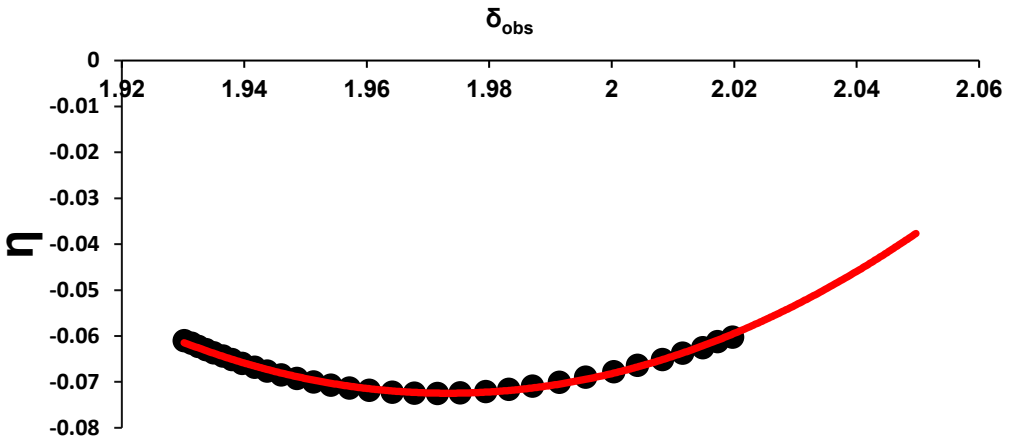


Figure B.10. Plot of η versus δ_{obs} for acetic acid

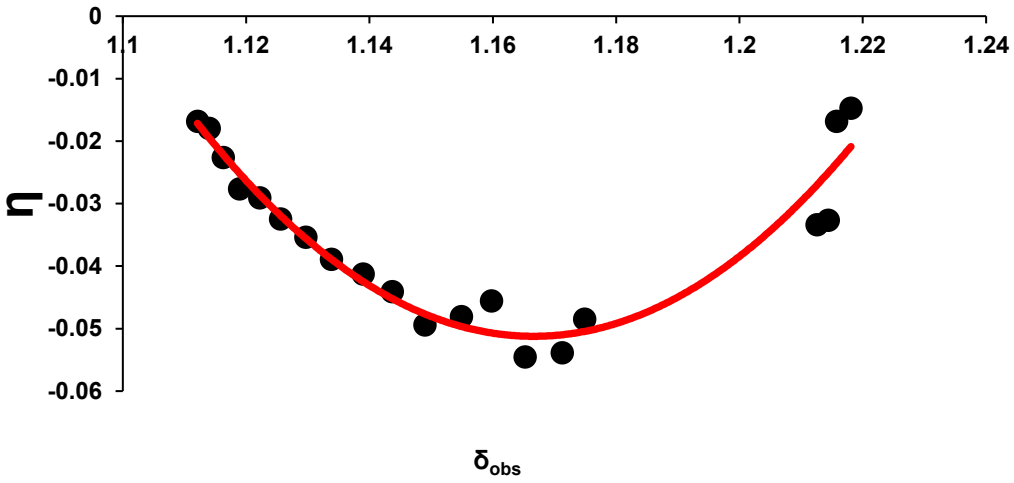


Figure B.11. Plot of η versus δ_{obs} for monohydrogen monophosphonate

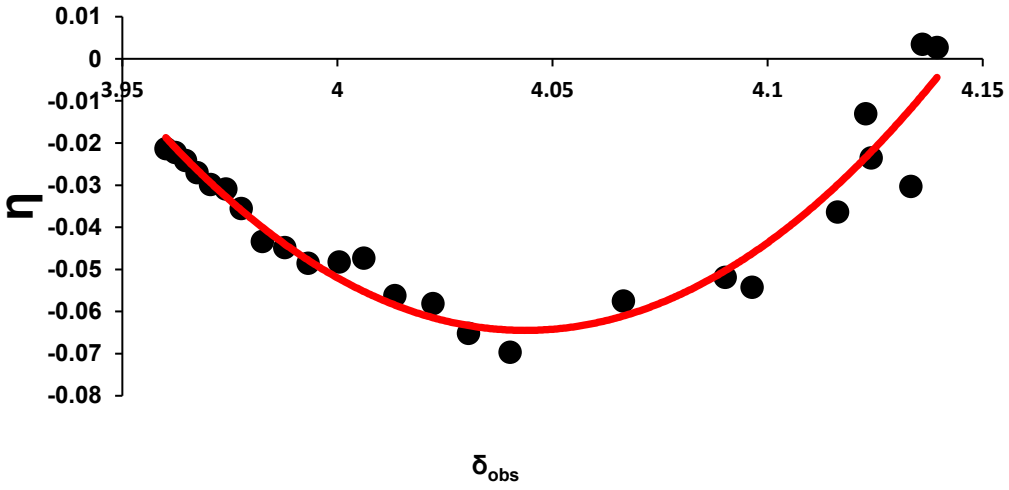


Figure B.12. Plot of η versus δ_{obs} for glycolic acid

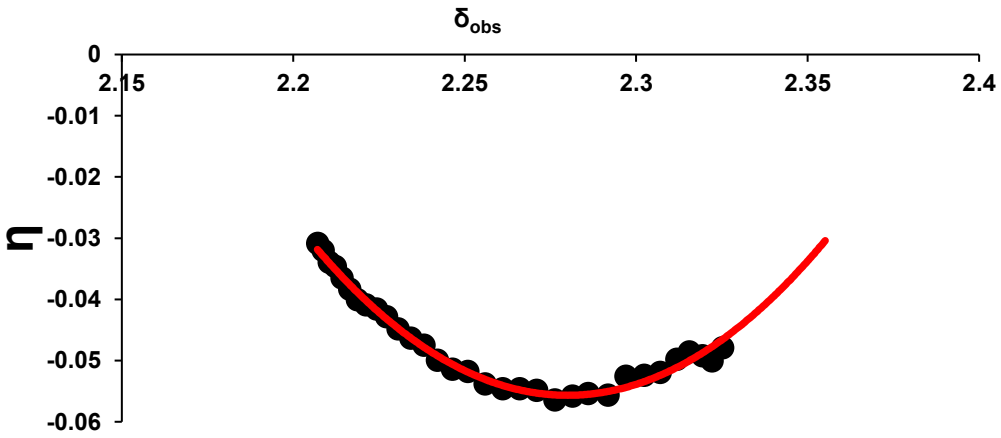


Figure B.13. Plot of η versus δ_{obs} for propanoic acid

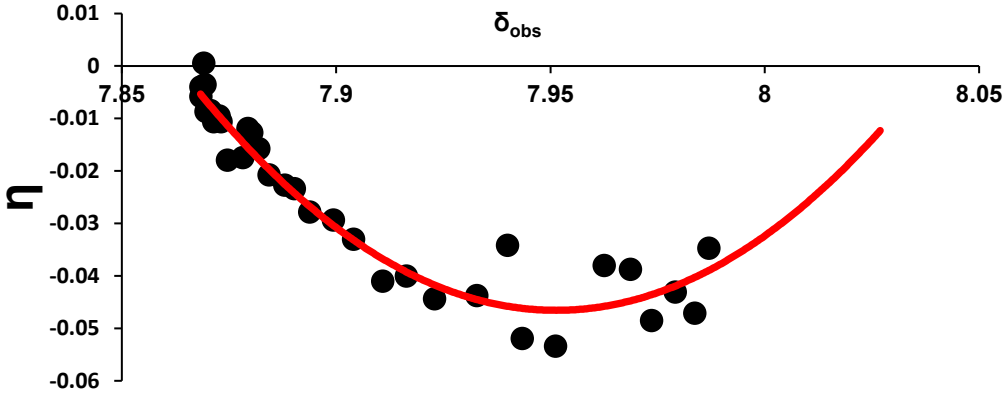


Figure B.14. Plot of η versus δ_{obs} for Benzoic acid

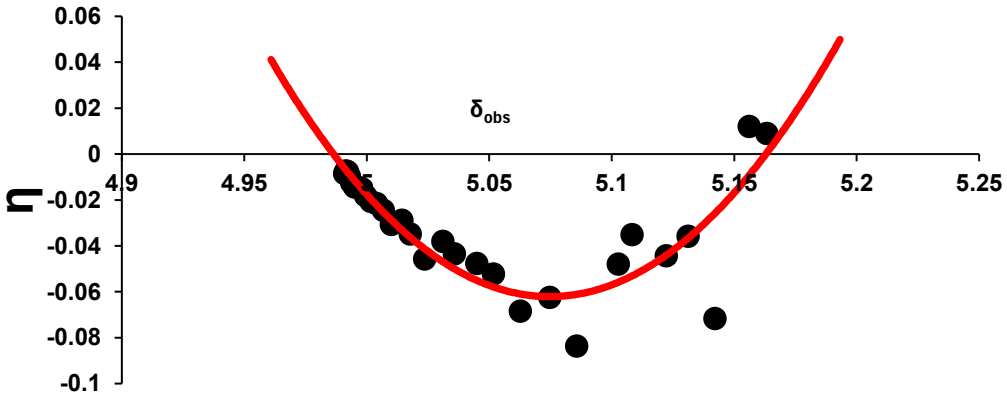


Figure B.15. Plot of η versus δ_{obs} for mandelic acid

Appendix Chapter 3

C.1 Uncertainty in pH determined by NMR

Chemical shift measurements have some degree of uncertainty due to factors such as slight magnetic field inhomogeneity or temperature fluctuations amongst other factors. We performed CSI experiments on a homogenous solution of DSS and DMSO and observed the chemical shift uncertainty as 0.0007 ppm by calculating the standard deviation of the chemical shift of DMSO (referenced to DSS) along the 64 datasets of the homogenous solution. We assume the uncertainty to be 0.0007 ppm in this paper. We also simulated a pH gradient of sodium acetate and performed uncertainty analysis on pH measurement with respect to δ_{obs} of sodium acetate (figure C.1) and observed that uncertainty increases significantly once pH is ± 2 units of the pK_a value of the indicator thus providing a clear window of when the indicator's pH measurement is reliable.

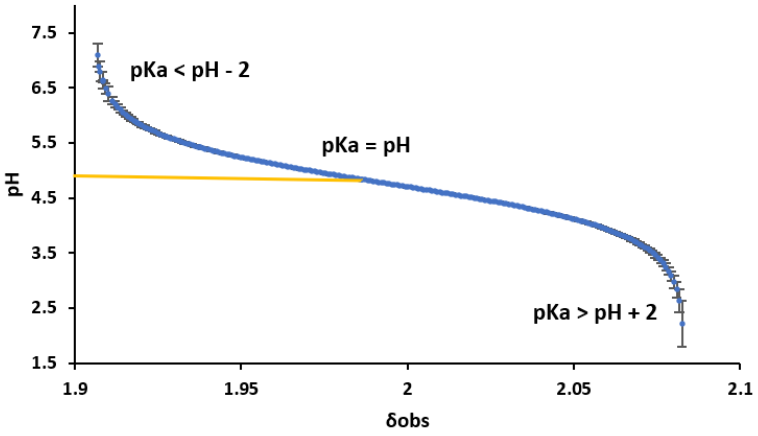


Figure C.1. Plot of pH calculated using equation 3.19 versus ^{1}H δ_{obs} of acetate (p$K_{a,0}$ = 4.76), yellow line crosses curve at p$K_{a,0}$ of acetate.

Table C.1. δ_{obs} and pH values calculated from the ^{1}H chemical shift of sodium acetate and the comparison of pH difference when pH ≈ pK_a of acetate (4.756) and when the pH is two units higher.

δ_{obs}/ppm	pH	δ_{obs}^*/ppm	pH*	ΔpH
1.9790	4.90971	1.9797	4.902633	0.007078
1.9073	6.886828	1.9080	6.698008	0.188820

* δ_{obs} + 0.0007 ppm and the resultant change in pH

C.2 Estimation of uncertainty in the determination of κ

The uncertainty in κ^{-1} is obtained by propagation of uncertainty analysis of Equation 3.5:

$$\Delta\kappa = \sqrt{\left(\frac{\partial\kappa}{\partial C_{indicator}}\right)^2 \Delta C_{base}{}^2 + \left(\frac{\partial\kappa}{\partial\delta_{obs}}\right)^2 \Delta\delta_{obs}{}^2 + \left(\frac{\partial\kappa}{\partial\delta_H}\right)^2 \Delta\delta_H{}^2 + \left(\frac{\partial\kappa}{\partial\delta_L}\right)^2 \Delta\delta_L{}^2} \quad \text{(C.1)}$$

Where Δ indicates the uncertainty in the variable. Differentiating equation 3.19 with respect to each variable:

$$\frac{\partial\kappa}{\partial C_{indicator}} = \frac{\delta_{obs} - \delta_L}{\delta_H - \delta_L} \quad \text{(C.2)}$$

$$\frac{\partial\kappa}{\partial\delta_{obs}} = \frac{C_{indicator} \times \delta_L}{\delta_H - \delta_L} \quad \text{(C.3)}$$

$$\frac{\partial\kappa}{\partial\delta_H} = \frac{C_{indicator} \times (\delta_L - \delta_{obs})}{(\delta_H - \delta_L)^2} \quad \text{(C.4)}$$

$$\frac{\partial\kappa}{\partial\delta_L} = -\frac{C_{indicator} \times (\delta_H - \delta_{obs})}{(\delta_L - \delta_H)^2} \quad \text{(C.5)}$$

$$\Delta\kappa^{-1} = \sqrt{\left(\frac{\partial\kappa^{-1}}{\partial C_{indicator}}\right)^2 \Delta C_{base}{}^2 + \left(\frac{\partial\kappa^{-1}}{\partial\delta_{obs}}\right)^2 \Delta\delta_{obs}{}^2 + \left(\frac{\partial\kappa^{-1}}{\partial\delta_H}\right)^2 \Delta\delta_H{}^2 + \left(\frac{\partial\kappa^{-1}}{\partial\delta_L}\right)^2 \Delta\delta_L} \quad \text{(C.6)}$$

Differentiating the inverse of equation 3.5 with respect to each variable:

$\Delta\delta_{obs}$, $\Delta\delta_L$, $\Delta\delta_H$ are taken as 0.0007 ppm in this work. $\Delta C_{indicator}$ is taken as 5% of expected

$$\frac{\partial\kappa^{-1}}{\partial\delta_H} = \frac{1}{(\delta_{obs} - \delta_L) \times C_{indicator}} \quad \text{(C.7)}$$

$$\frac{\partial\kappa^{-1}}{\partial\delta_L} = \frac{\delta_H - \delta_{obs}}{C_{indicator} \times (\delta_L - \delta_{obs})^2} \quad \text{(C.8)}$$

$$\frac{\partial\kappa^{-1}}{\partial\delta_{obs}} = \frac{\delta_L - \delta_H}{C_{indicator} \times (\delta_{obs} - \delta_L)^2} \quad \text{(C.9)}$$

$$\frac{\partial\kappa^{-1}}{\partial C_{base}} = \frac{\delta_L - \delta_H}{(\delta_{obs} - \delta_L) \times C_{indicator}{}^2} \quad \text{(C.10)}$$

concentration. As the pH in our experiments is within one unit of the pK_a of the indicator, we can ignore uncertainty in pH as this is expected to be negligible in comparison to $\Delta\kappa^{-1}$.

C.3 Alternative method for determining pK_a and concentration of analyte

pK_a and $C_{analyte}$ can be determined directly by obtaining κ using equation 3.5 and then fitting the measured κ using equation 3.7 where pK_a and $C_{analyte}$ are used as free parameters with pH measured using indicators and equation 3.19 (displayed are all the equations relevant). The

fitting is done using excel through nonlinear generalised reduced gradient function (GRG), which looks at the gradient of the objective function as the input values are changed and determines whether the partial derivatives are equal to the values desired. In the context of pK_a measurement, the objective function is set to equal 1 and the objective function is an R^2 comparison between the pH obtained from the modified Henderson-Hasselbalch equation from an indicator compound and the analyte in question (Example is shown in figure C.2 for the method and its respective result).

$$pH = \mathrm{p}K_{\mathrm{a},0} + \log\frac{\delta_{\mathrm{obs}} - \delta_{\mathrm{H}}}{\delta_{\mathrm{L}} - \delta_{\mathrm{obs}}} + \Delta z^2\left(\frac{0.51\sqrt{\mathrm{I}}}{1+\sqrt{\mathrm{I}}} - 0.1\mathrm{I}\right) \tag{3.19}$$

$$\kappa = C_{indicator}\,\frac{\delta_{obs}-\delta_L}{\delta_H-\delta_L} + 10^{-pH} \tag{3.5}$$

$$\kappa = \frac{C_{analyte}}{1+10^{\mathrm{p}K\mathrm{a}-\mathrm{pH}}} \tag{3.7}$$

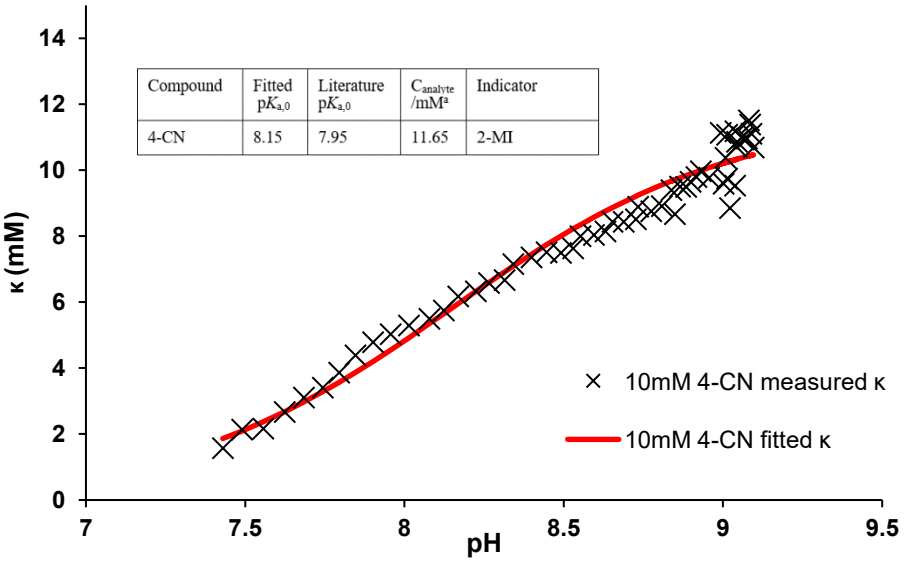


Figure C.2 κ versus pH graph of 10 mM 4-CN with 2-MI as indicator, both measured κ and fitted κ are displayed (R^2 of 0.96 is obtained)

C.4 Impact of H_2O on protonation of indicators

Taking 2-methylimidazole as the indicator with highest pK_a and hence most susceptible to protonation from H_2O. Calculation of fraction of indicators protonated f_H from H_2O we use the NMR modified Henderson-Hasselbalch equation 3.13 and 3.19:

$$\mathrm{pH} = \mathrm{pK_{a,0}} + \log_{10}\left(\frac{f_{H\ indicator}}{f_{L\ indicator}}\right) = \mathrm{pK_{a,0}} + \log_{10}\left(\frac{f_{H\ indicator}}{1 - f_{H\ indicator}}\right) \tag{3.13}$$

$$\mathrm{pH} = \mathrm{pK_{a,0}} + \log_{10}\left(\frac{\delta_{\mathrm{obs}} - \delta_{\mathrm{H}}}{\delta_{\mathrm{L}} - \delta_{\mathrm{obs}}}\right) + \Delta z^2\left(\frac{0.51\sqrt{\mathrm{I}}}{1 + \sqrt{\mathrm{I}}} - 0.1\mathrm{I}\right) \tag{3.19}$$

Nine separate 1D proton NMR experiments were run each containing 2-methylimidazole in H_2O with respective concentrations of 1, 2, 4, 5, 10, 20, 50, 100 and 200 mM. pH of solution in all experiments was measured using equation 2. f_H of 2-methylimidazole for each experiment was calculated using C.11:

$$f_{H\ indicator} = \frac{1}{1 + 10^{\mathrm{pK_{a,0}} - \mathrm{pH}}} \tag{C.11}$$

Figure C.3 demonstrates how the fraction protonated of 2-methylimdazole from H_2O is 3.22 % when concentration of 2-methylimidazole is 1 mM and f_H decays in a sigmoidal fashion as concentration of 2-methylimidazole increases. This implies that impact of protonation from H_2O to 2-methylimidazole is negligible and hence impact of protonation from water to indicators that are of lower pK_a than 2-methylimidazole is negligible as well.

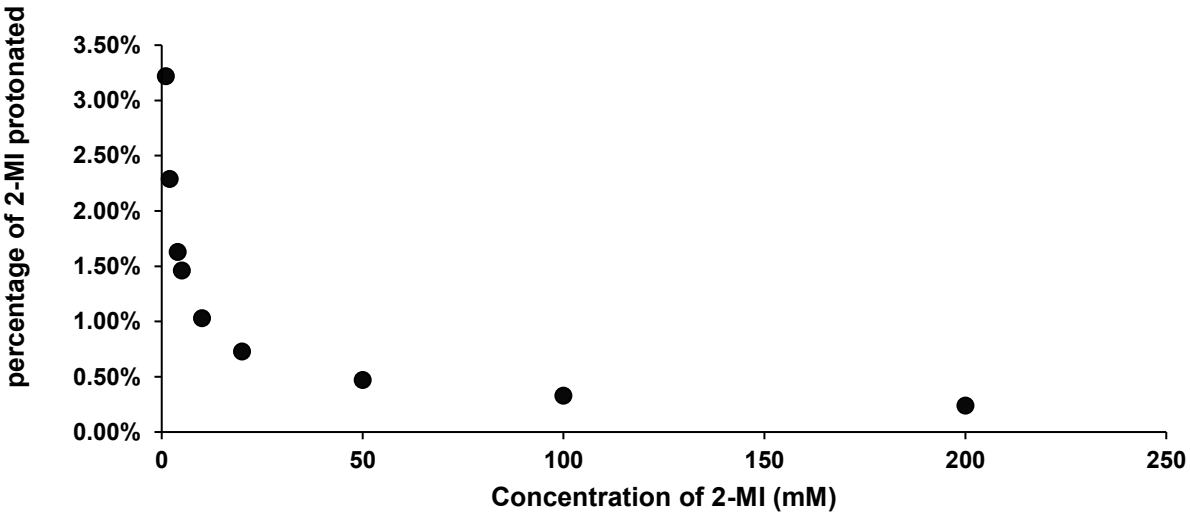


Figure C.3. Figure showing how f_H of 2-MI changes in water as a function of its concentration

C.5 ^{1}H NMR spectra of homogenous polyacrylic acid with sodium acetate, and concentration gradients of NH_3 and NaCl

In homogenoues mixtures of PAA (10 mM COOH groups) and sodium acetate, the ^{1}H resonances of PAA exhibit no difference in intensity between 5 and 200 mM sodium acetate (Figure C.4), although the resonances are shifted slightly upfield at the higher concentration due to deprotonation. These samples contained 0.2 mM DSS and 0.01 vol % dioxane and were otherwise identical to the sample of Figure 3.3. The loss of the PAA resonances observed on Figure 3.3 is thus attributable to an effect of the concentration gradient of sodium acetate. To further investigate the origin of this effect, a sample was prepared by layering 60 μL of 0.5 M NH_3 on top of 500 μL of PAA solution to create a concentration gradient of NH_3, with a higher concentration of base towards the top of the sample, in contrast to the sodium acetate. Significant loss of intensity of PAA is observed towards the top of the sample (Figure C.5). A sample prepared analgously to Figure 3.3, but with NaCl in place of sodium acetate, exhibits no change in the intensity of the PAA resonances along its length, suggesting diffusiophoresies along a salt gradient is not the source of signal loss of the PAA (Figure C.6).

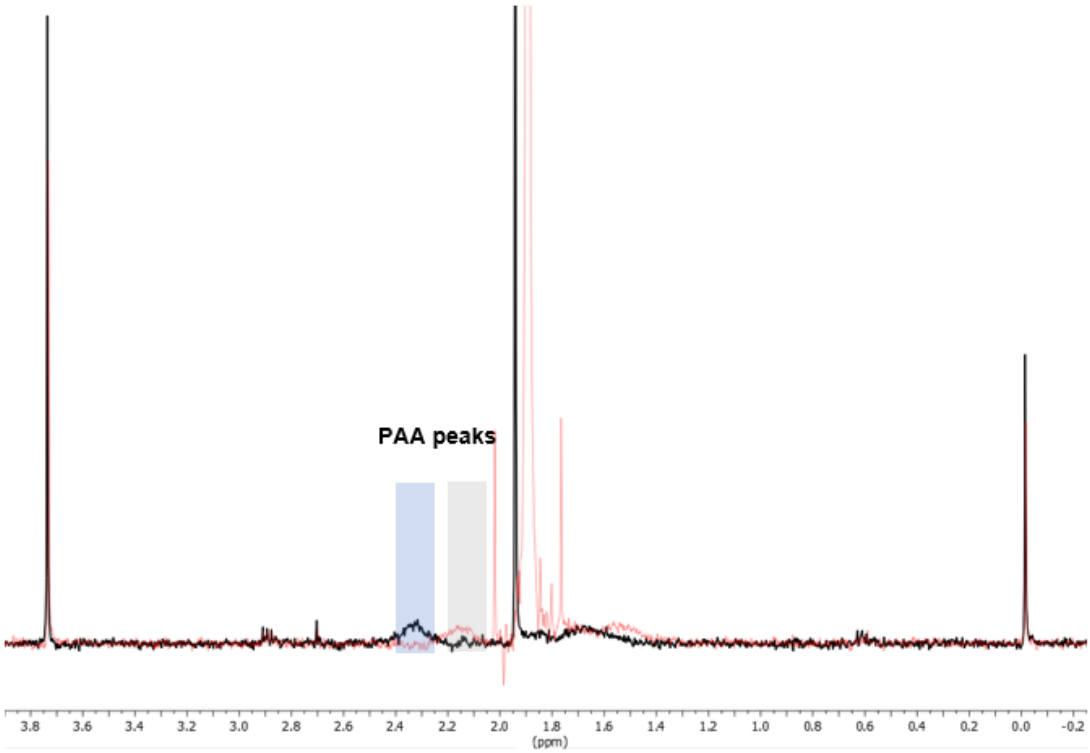


Figure C.4 Stacked ^{1}H NMR spectra of homogenous polyacrylic acid solution (9.10 mM COOH groups) with 200 mM sodium acetate (red spectrum), 5 mM sodium acetate (black spectrum).

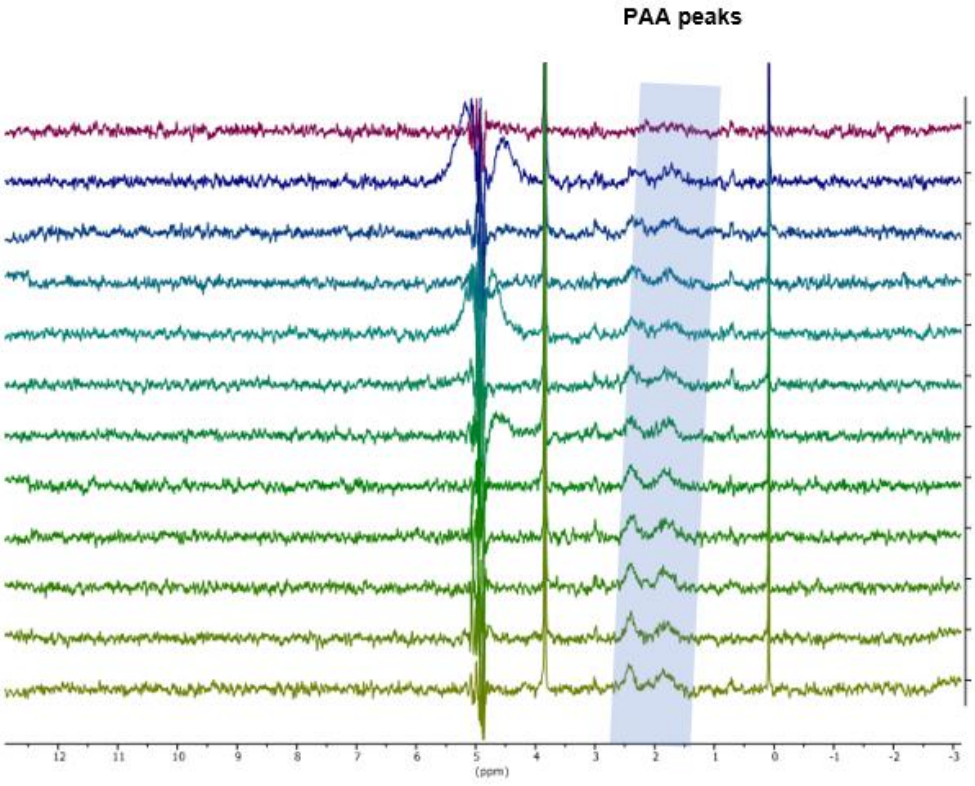


Figure C.5. Stacked ^{1}H NMR spectra of 10 mM polyacrylic acid solution with concentration gradient of ammonia (the top slices correspond to the top of the NMR active region of the sample tube, closer to the source of the ammonia).

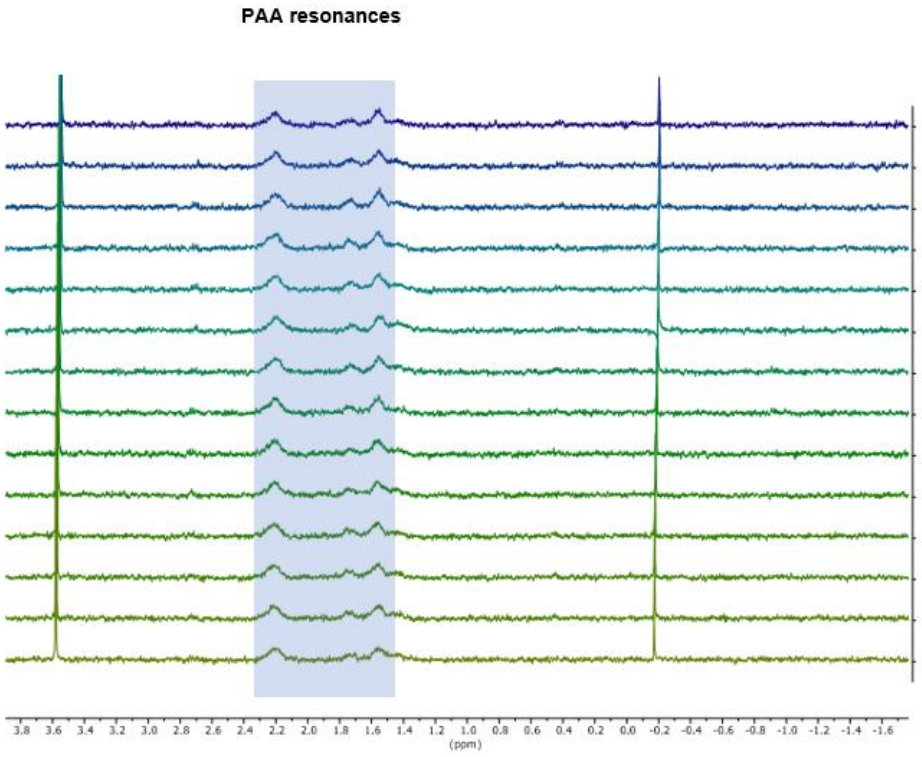


Figure C.6. Stacked ^{1}H NMR spectra of 10 mM polyacrylic acid solution with NaCl gradient (the top slices correspond to the top of the NMR active region of the sample tube, with a lower concentration of NaCl at the top).

C.6 κ profiles of analyte's pH gradient over time (all analytes)

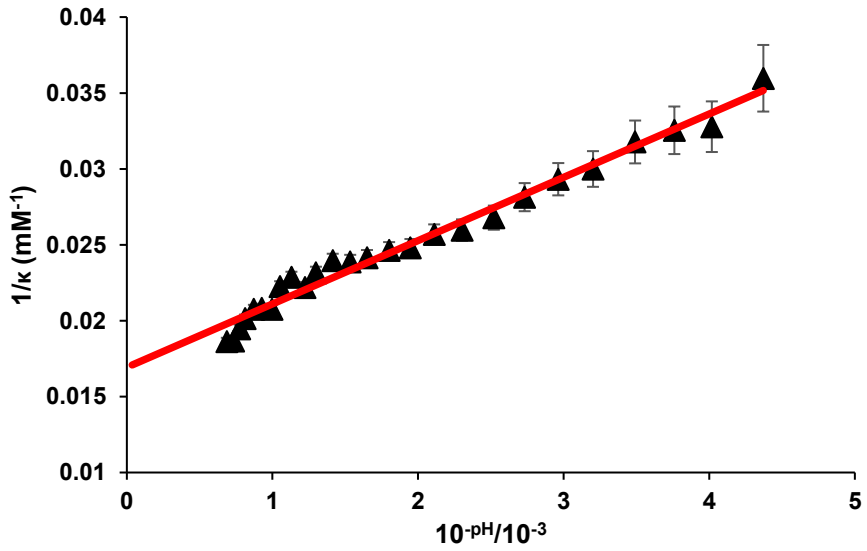


Figure C.7. Plot of $1/\kappa$ versus 10^{-pH} for H_3PO_4 with 1,2,4-triazole as indicator

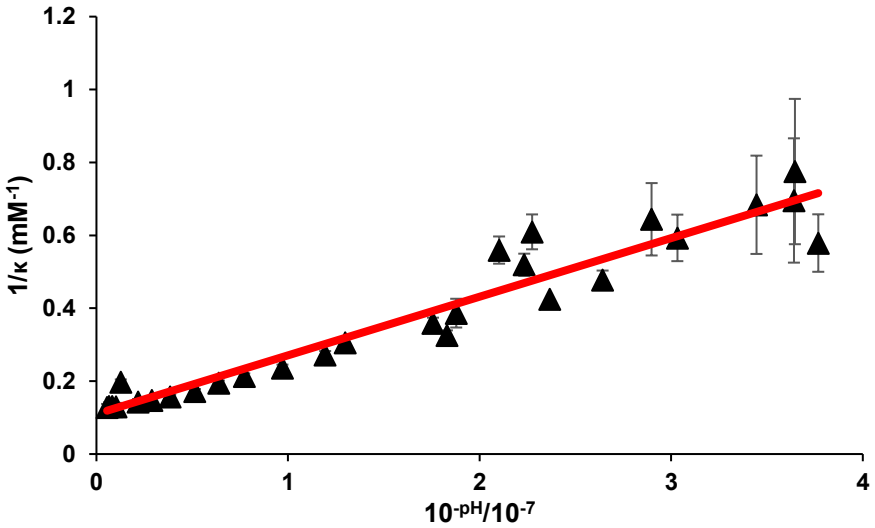


Figure C.8. Plot of $1/\kappa$ versus 10^{-pH} for NaH_2PO_4 with 2-methylimidazole as indicator

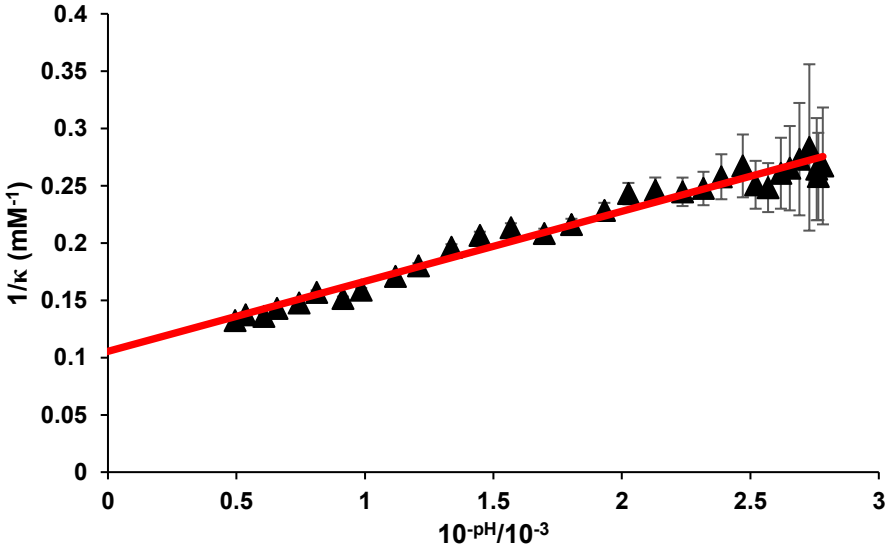


Figure C.9. Plot of $1/\kappa$ versus 10^{-pH} for glycine with 1,2,4-triazole as indicator

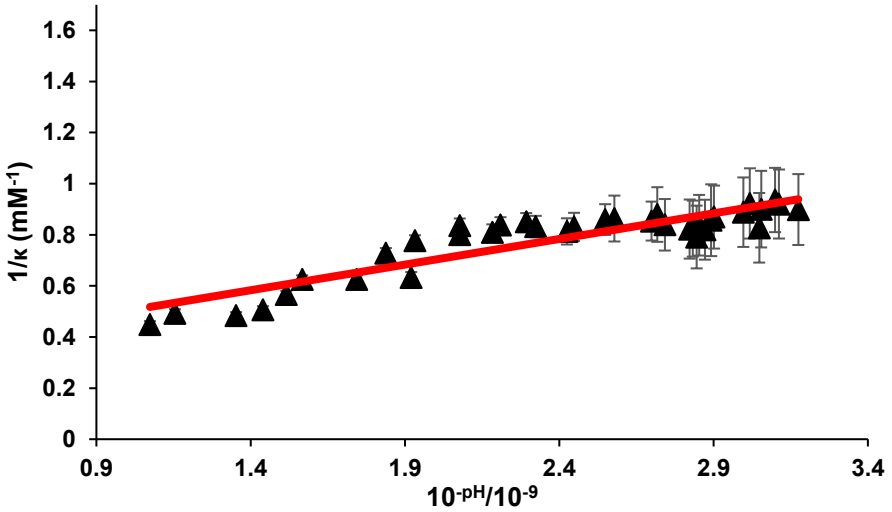


Figure C.10. Plot of $1/\kappa$ versus 10^{-pH} for NH_4Cl with 2-methylimidazole as indicator

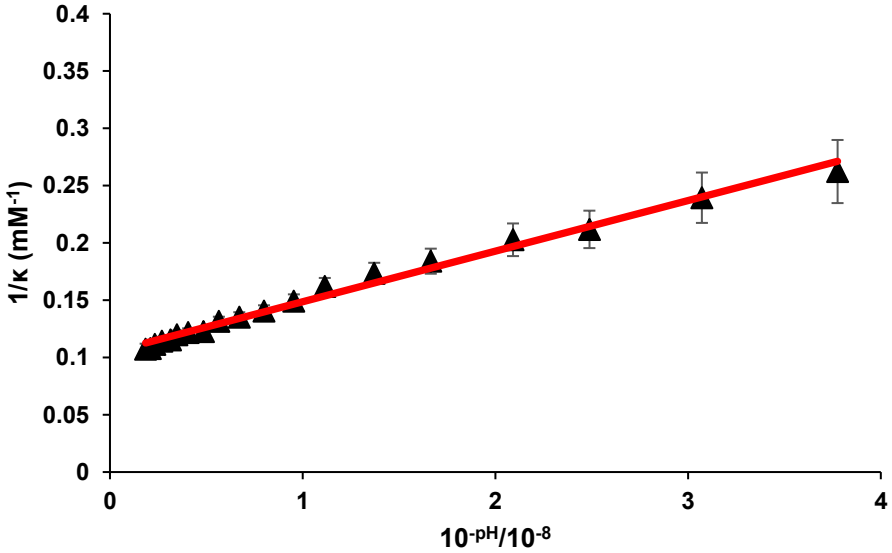


Figure C.11. Plot of $1/\kappa$ versus 10^{-pH} for 4-cyanophenol with 2-methylimidazole as indicator

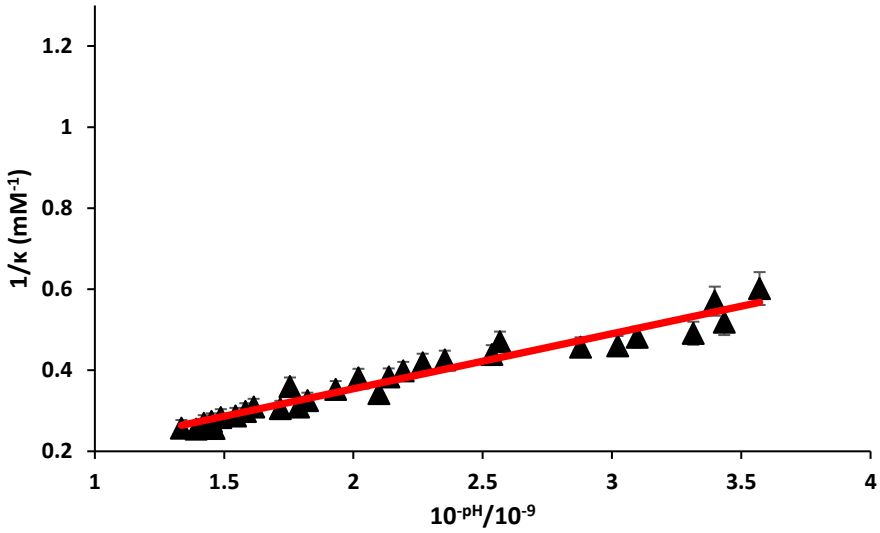


Figure C.12. Plot of $1/\kappa$ versus 10^{-pH} for boric acid with 2-methylimidazole as indicator

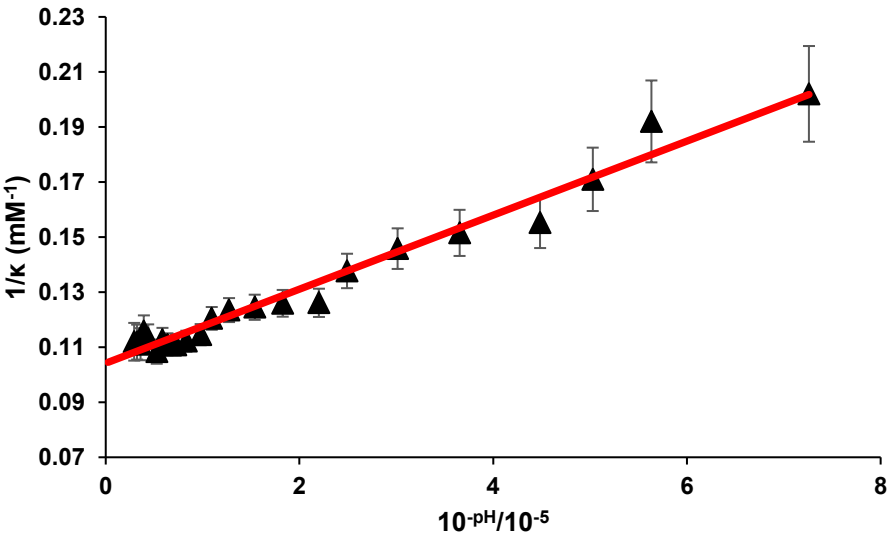


Figure C.13. Plot of $1/\kappa$ versus 10^{-pH} for benzoic acid with sodium acetate as indicator

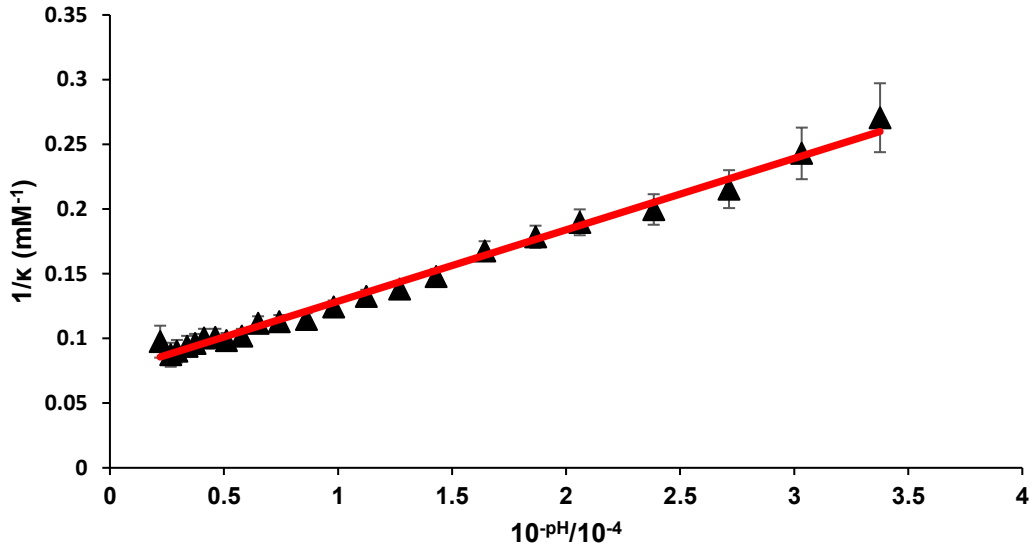


Figure C.14. Plot of $1/\kappa$ versus 10^{-pH} for glycolic acid with sodium formate as indicator

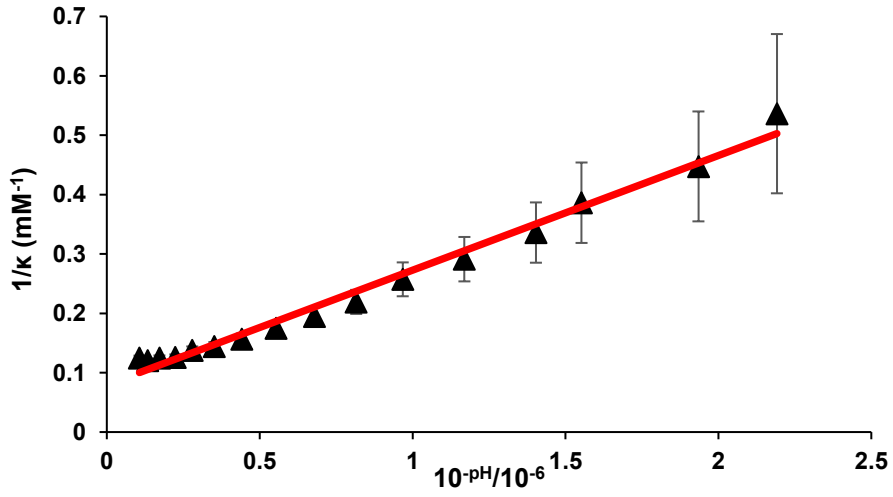


Figure C.15. Plot of $1/\kappa$ versus 10^{-pH} for NH_3OHCl with 2,6-lutidine as indicator

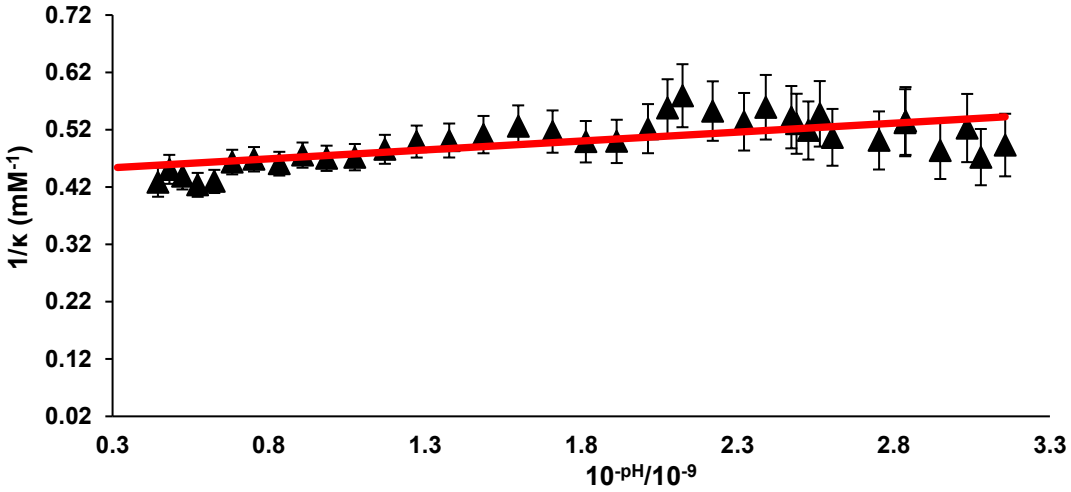


Figure C.16. Plot of $1/\kappa$ versus 10^{-pH} for lectin with 2-methylimidazole as indicator

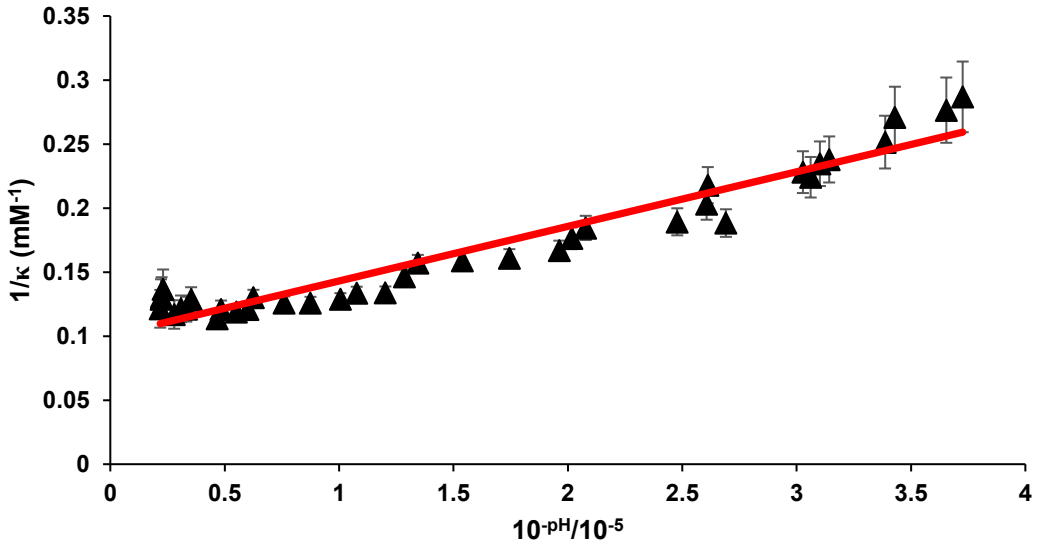


Figure C.17. Plot of $1/\kappa$ versus 10^{-pH} for polyacrylic acid with sodium acetate as indicator

Appendix Chapter 4

D.1 Using multiple indicators to determine pH and analyte p_sK_a

Using one indicator can be impractical as chances of indicator's pK_a being similar to the analyte's is not high. To increase the practicality of this technique, it is essential that more than one indicator could be used simultaneously as large differences between the indicator and analyte's pK_a can lead to larger uncertainty in measurement. Developing multiple indicators with varying pK_a values can cover a larger "pKa window". Minimising the contribution from high uncertainty indicators. The experimental setup is such that the sum of the protons transferred to the indicators is equal to the proton transferred by the analyte. This means that by measuring the κ of all the indicators and summing the total protons transferred by the analyte can be determined. This is represented by equation D.1:

$$\frac{C_{analyte}}{1+10^{pKa-pH}} = \sum_{i=1}^{n} C_i \times \frac{\delta_{obs,i} - \delta_{L.i}}{\delta_{H,i} - \delta_{L.i}} \tag{D.1}$$

Where n is the number of indicators. Given that the indicator's uncertainty with respect to pH measurement increases as pH is further from the indicator's pK_a it is necessary to use a sensitivity weighted average pH value when calculating κ and p_sK_a. This is done by calculating the sensitivity S_i of each indicator with respect to pH through equation D.2:

$$S_i = \frac{(\delta_L - \delta_{obs})(\delta_{obs} - \delta_H)}{\delta_H - \delta_L} \tag{D.2}$$

The pH value at each position along the sample is calculated as the sensitivity-weighted average of pH:

$$pH = \frac{\sum_{i=1}^{n} S_i\, pH_i}{\sum_{i=1}^{n} S_i} \tag{D.3}$$

Where pH_i is the pH as measured from the ith indicator

D.2 Estimation of uncertainty in the determination of aqueous pK_a

The uncertainty in analyte pK_a across different slices is obtained by propagation of analysis of equation 4.12:

$$\Delta p_sK_a = \sqrt{(\frac{\partial p_sK_a}{\partial pH})^2 \Delta pH^2 + (\frac{\partial p_sK_a}{\partial \kappa})^2 \Delta \kappa^2 + (\frac{\partial p_sK_a}{\partial C_{analyte}})^2 \Delta C_{analyte}{}^2} \tag{D.4}$$

Where Δ indicates the uncertainty in the variable. Differentiating equation 4.12 with respect to each variable:

$$\frac{\partial p_s K_a}{\partial pH} = 1 \tag{D.5}$$

$$\frac{\partial p_s K_a}{\partial \kappa} = \frac{C_{analyte}}{ln(10)\kappa(\kappa - C_{analyte})} \tag{D.6}$$

$$\frac{\partial p_s K_a}{\partial C_{analyte}} = \frac{1}{ln(10)(C_{analyte} - \kappa)} \tag{D.7}$$

Additionally uncertainty analysis with respect to κ is done:

$$\Delta\kappa = \sqrt{(\frac{\partial \kappa}{\partial C_{indicator}})^2 \Delta {C_{indicator}}^2 + (\frac{\partial \kappa}{\partial \delta_{obs}})^2 \Delta {\delta_{obs}}^2 + (\frac{\partial \kappa}{\partial \delta_H})^2 \Delta {\delta_H}^2 + (\frac{\partial \kappa}{\partial \delta_L})^2 \Delta {\delta_L}^2} \tag{D.8}$$

Differentiating equation 4.10 with respect to each variable:

$$\frac{\partial \kappa}{\partial C_{indicator}} = \frac{\delta_{obs} - \delta_L}{\delta_H - \delta_L} \tag{D.9}$$

$$\frac{\partial \kappa}{\partial \delta_{obs}} = \frac{C_{indicator}\delta_L}{\delta_H - \delta_L} \tag{D.10}$$

$$\frac{\partial \kappa}{\partial \delta_H} = \frac{C_{indicator}\,(\delta_L - \delta_{obs})}{(\delta_H - \delta_L)^2} \tag{D.11}$$

$$\frac{\partial \kappa}{\partial \delta_L} = \frac{C_{indicator}\,(\delta_H - \delta_{obs})}{(\delta_L - \delta_H)^2} \tag{D.12}$$

For uncertainty estimation of fractional volume of DMSO measurement 2D ^{1}H CSI experiments were with homogenous solutions containing varied fractional volumes of DMSO/H_2O ranging from 80/20 DMSO/H_2O to 20/80 DMSO/H_2O with their fractional volume of DMSO measured for each slice of the experiment and the standard deviation calculated for all the different solvent compositions and averaged. The standard deviation obtained was ± 0.27% fractional volume of DMSO.

D.3 Determination of pK_a, δ_L and δ_H of formate across solvent composition

For determination of limiting chemical shifts and pK_a of formate proton NMR CSI experiments were run with analyte solution containing 10 mM of sodium formate, 0.2 mM of DSS and 20 mM of sodium methanesulfonate. A solution of acetic acid (2 M, 30 µL) was then pipetted at the bottom of the NMR tube and four glass beads were placed in the solution. An aliquot of the analyte solution was then drawn up in a 9 in. Pasteur pipet and

gently layered on top to a height of 40-50 mm from the base of the NMR tube. This approach allows for an acid concentration gradient to be produced in the NMR tube causing for a variation in pH across the NMR tube. pH a is then measured across the NMR tube using equation D.13 with the acid as the pH indicator:

$$pH = p_sK_a + \log_{10}\frac{\delta_{obs} - \delta_H}{\delta_L - \delta_{obs}} \quad \text{(D.12)}$$

rearranging equation D.12 to make δ_{obs} the subject of the equation:

$$\delta_{obs} = \frac{\delta_H 10^{(p_sK_a - pH)} + \delta_L}{1 + 10^{(p_sK_a - pH)}} \quad \text{(D.13)}$$

δ_L , δ_H and pK_a of formate are then fitted as free parameters for equation S11 with δ_{obs} and pH measured experimentally across the NMR tube (see figure D.1 for example of fitting). This is done for 20/80, 50/50 and 80/20 H_2O/DMSO solvent composition (see Section D.4 for fitting plots of δ_L , δ_H and pK_a for all indicators).

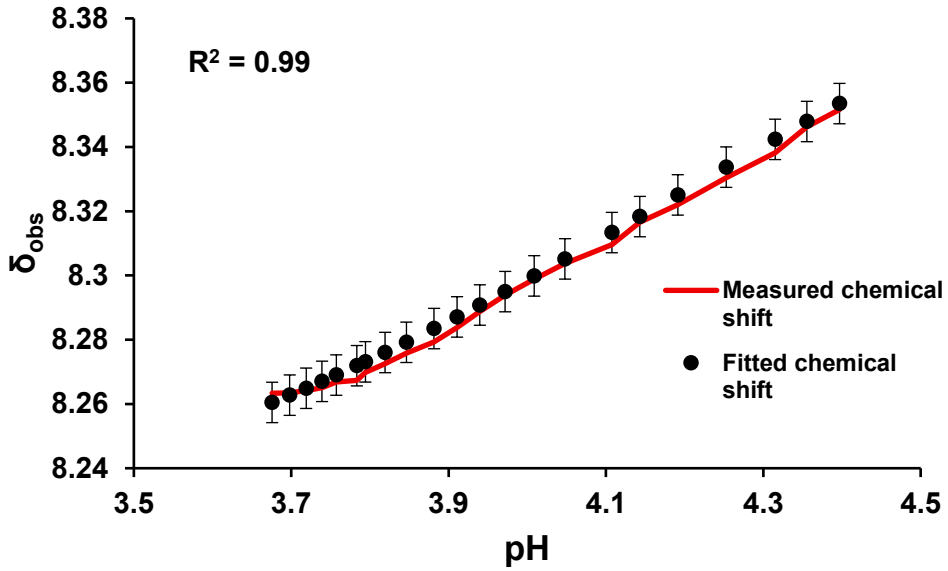


Figure D.1. Plot of δ_{obs} versus pH for formate in 80/20 H_2O/DMSO solvent composition with respective fitting

D.4 pK_a, δ_L, and δ_H values of indicators at 20/80, 50/50 and 80/20 H_2O/DMSO solution mixture

Table D.1. Parameters of indicators at 20/80 H_2O/DMSO

Indicator	δ_L	δ_H	pK_a	p$K_{a\ Lit.}$
1,2,4-Triazole	8.3271	9.3484	1.42±0.13	-
Sodium glycolate	3.5680	3.9910	6.61±0.16	6.94[235]
Sodium formate	8.4434	8.1612	6.57±0.17	6.42[236], 6.51[176]
Imidazole	7.0800	7.6140	5.37±0.13	-
2-methylimidazole	2.3179	2.5789	6.34±0.17	-

Sodium acetate	1.7127	1.9688	7.86±0.17	7.27[156], 8.00[237]
Dimethylglycinate (CH_2)	2.7400	3.4668	9.78±0.18	-

Table D.2. Parameters of indicators at 50/50 H_2O/DMSO

Indicator	δ_L	δ_H	pK_a	$pK_{a\ Lit.}$
1,2,4-Triazole	8.3621	9.1934	1.77	-
Sodium glycolate	3.7583	4.0994	4.82	-
Sodium formate	8.4508	8.1857	4.62	4.73[176]
2-methylimidazole	2.3181	2.5790	6.66	
Sodium acetate	1.8048	2.0293	5.72	

Table D.3. Parameters of indicators at 80/20 H_2O/DMSO

Indicator	δ_L	δ_H	pK_a	$pK_{a\ Lit.}$
Maleic acid	6.3334	6.4074	2.10±0.06	-
1,2,4-Triazole	8.3590	9.1079	2.42±0.27	-
Sodium glycolate	3.8702	4.1765	4.28±0.27	-
Sodium formate	8.4209	8.2163	3.85	4.13[176]
2-methylimidazole	2.3359	2.5915	7.76±0.28	
Sodium acetate	1.8674	2.0651	5.16±0.28	4.90[238]

D.5 $\log[H_2O] + p_sK_a$ profiles of analytes

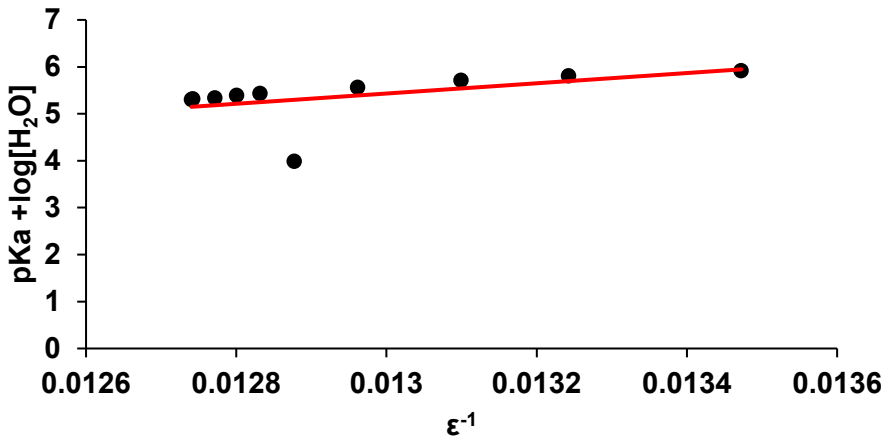


Figure D.2. Plot of $\log[H_2O] + p_sK_a$ versus 1/ε for salicylic acid

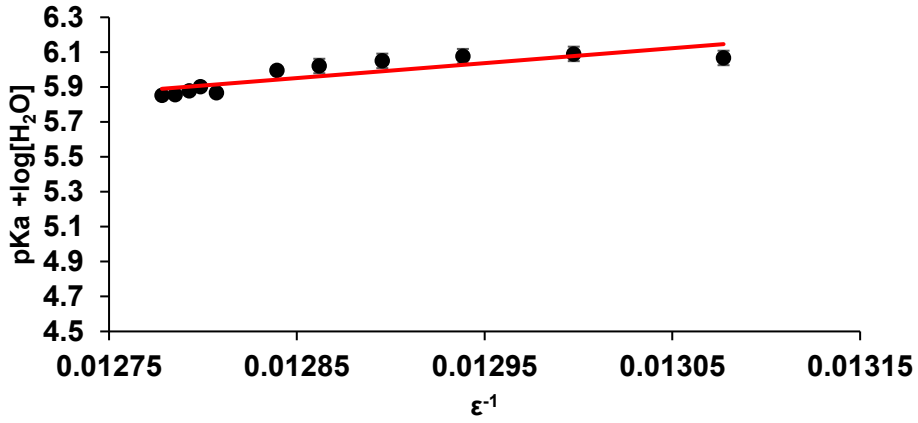


Figure D.3. Plot of log[H_2O] + p_sK_a versus 1/ε for formic acid

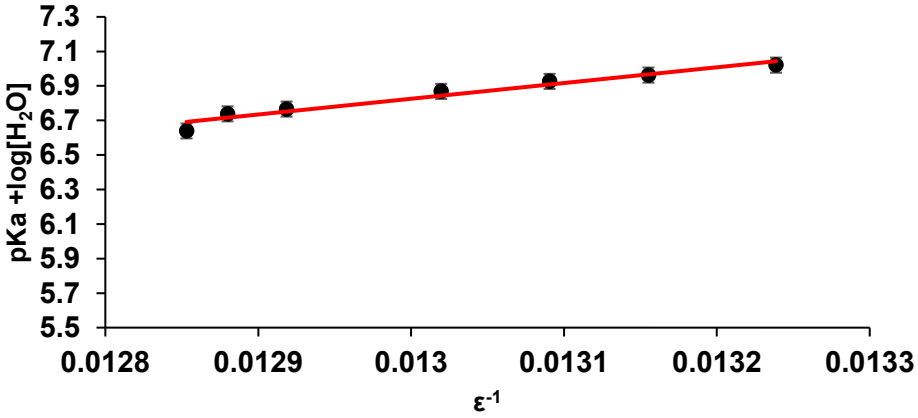


Figure D.4. Plot of log[H_2O] + p_sK_a versus 1/ε for sulfoacetic acid

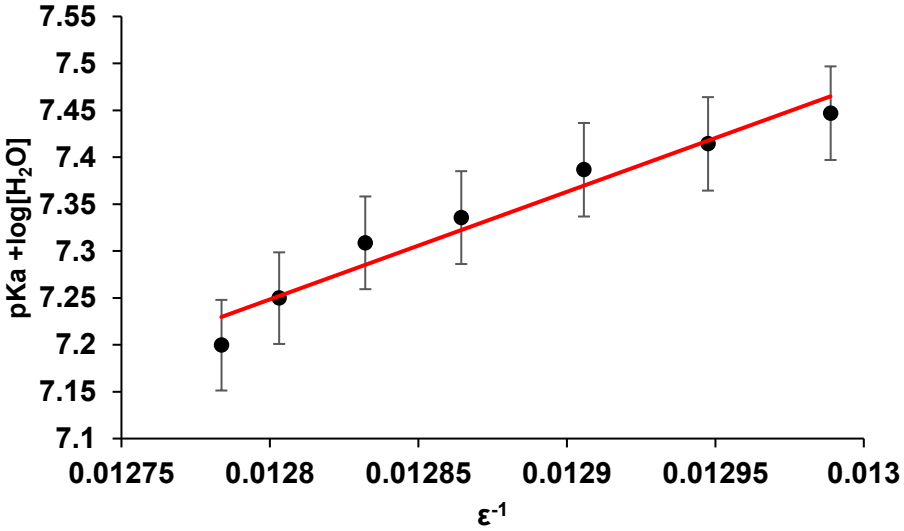


Figure D.5. Plot of log[H_2O] + p_sK_a versus 1/ε for acetic acid

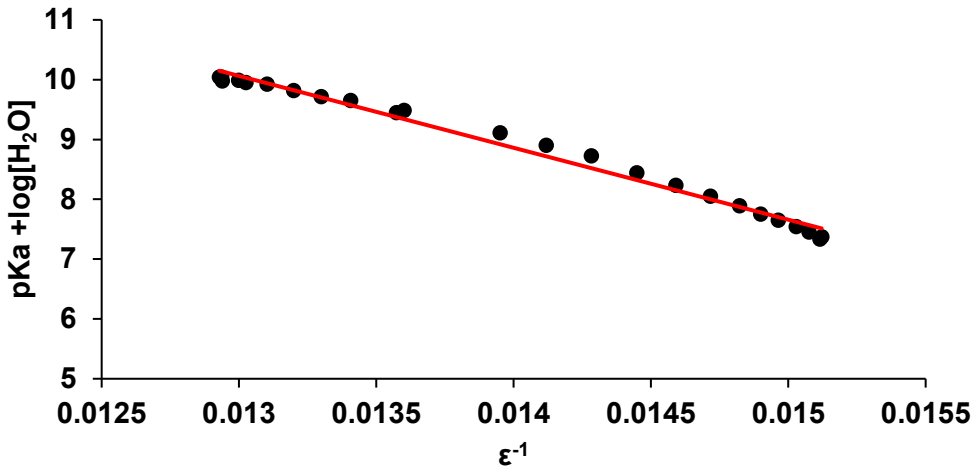


Figure D.6. Plot of log[H_2O] + p_sK_a versus 1/ε for quinine hydrochloride

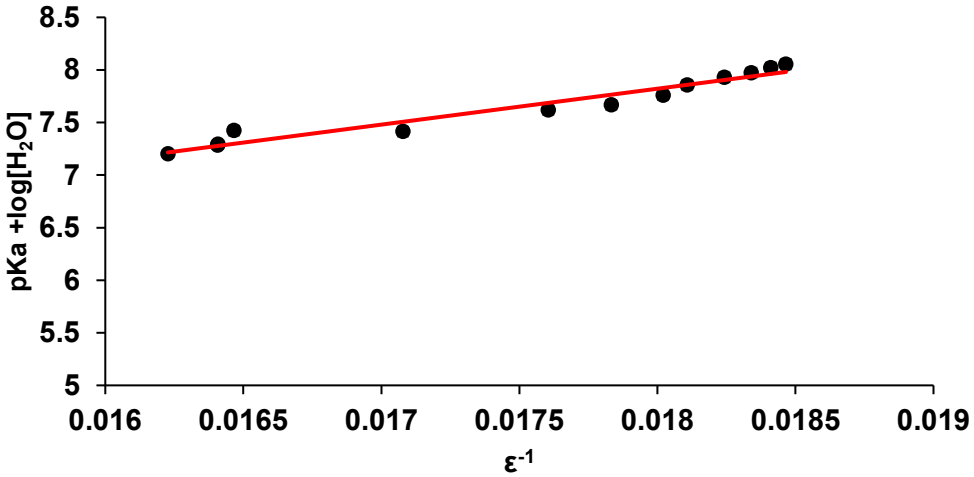


Figure D.7. Plot of log[H_2O] + p_sK_a versus 1/ε for naproxen

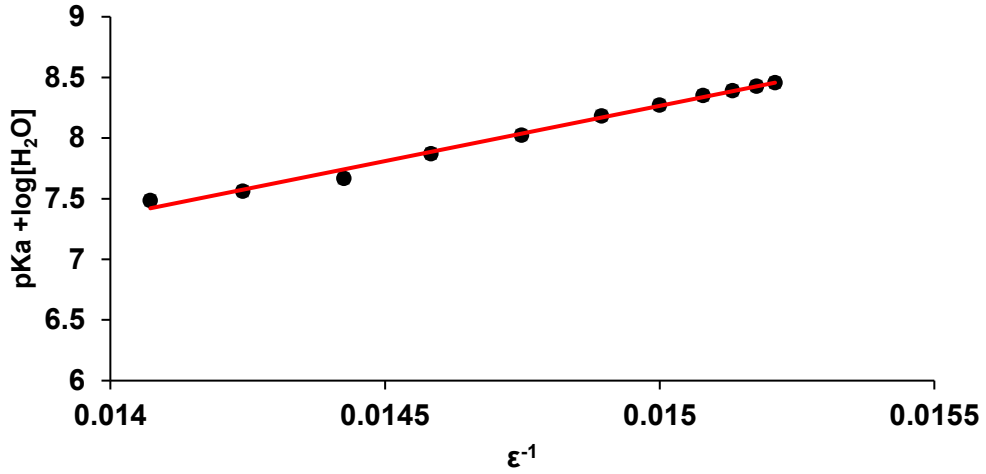


Figure D.8. Plot of log[H_2O] + p_sK_a versus 1/ε for indomethacin

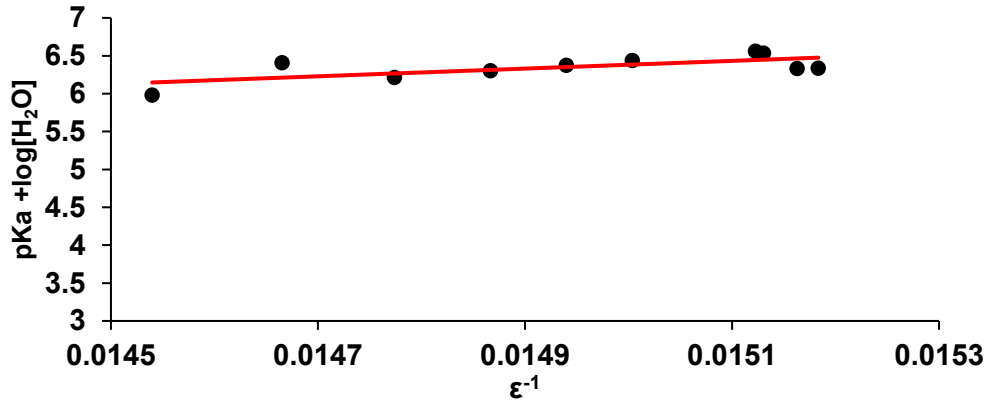


Figure D.9. Plot of log[H_2O] + p_sK_a versus 1/ε for furosemide

D.6 pK_a, δ_L, and δ_H profiles of solvent gradient indicators

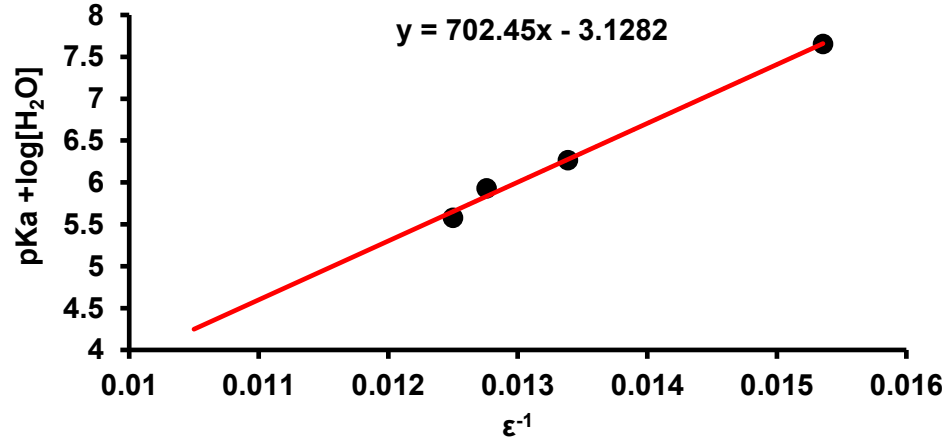


Figure D.10. Plot of log[H_2O] + p_sK_a versus 1/ε for sodium glycolate with respective fitting

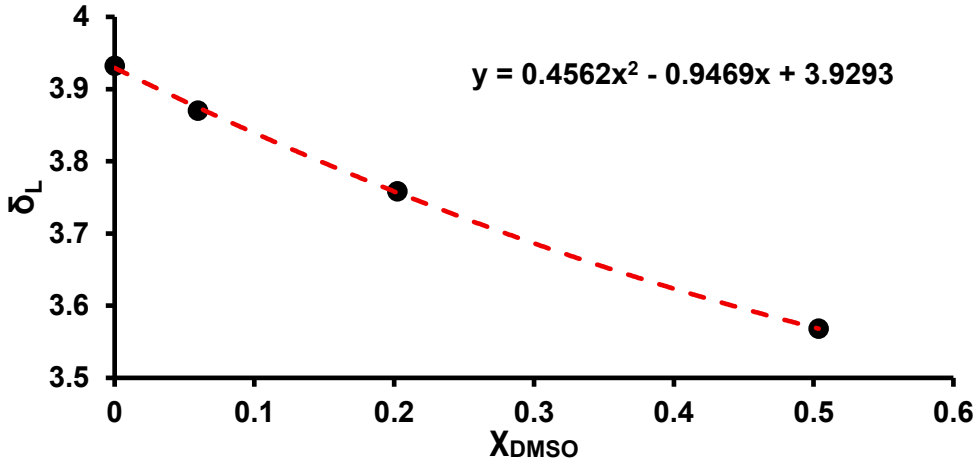


Figure D.11. Plot of δ_L of sodium glycolate as a function of mole fraction of DMSO with the respective fitting and equation

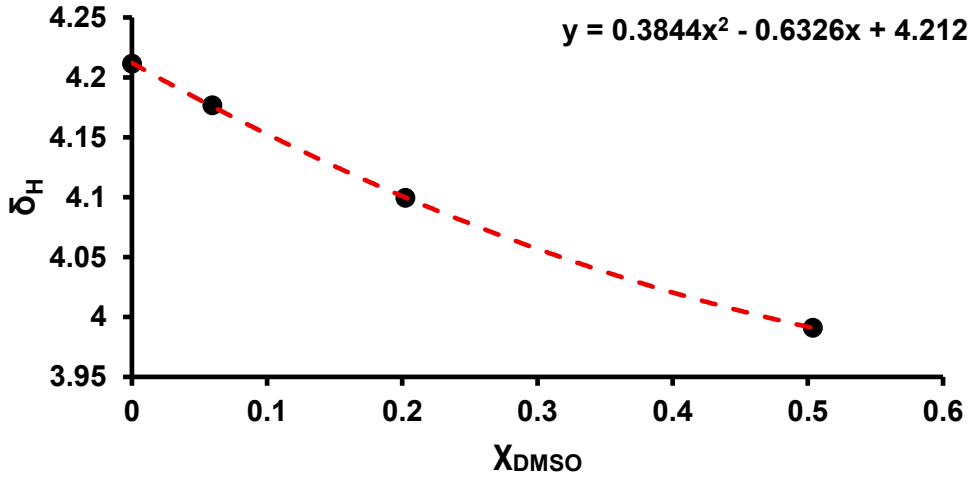


Figure D.12. Plot of δ_H of sodium glycolate as a function of mole fraction of DMSO with the respective fitting and equation

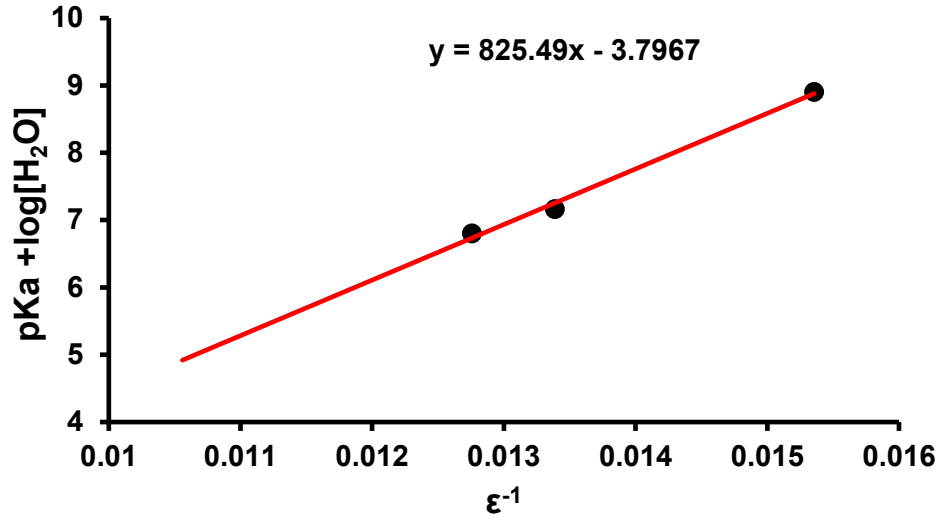


Figure D.13. Plot of log[H_2O] + p_sK_a versus 1/ε for sodium acetate with respective fitting

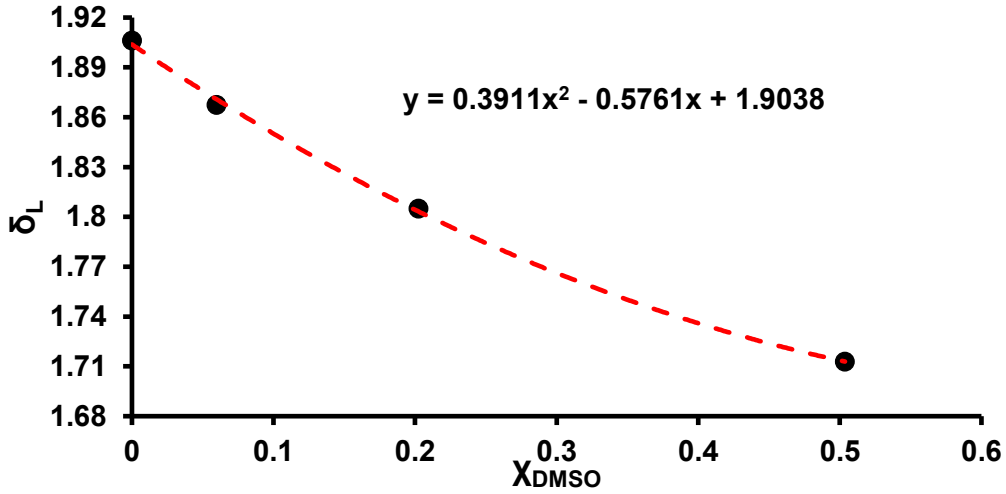


Figure D.14. Plot of δ_L of sodium acetate as a function of mole fraction of DMSO with the respective fitting and equation

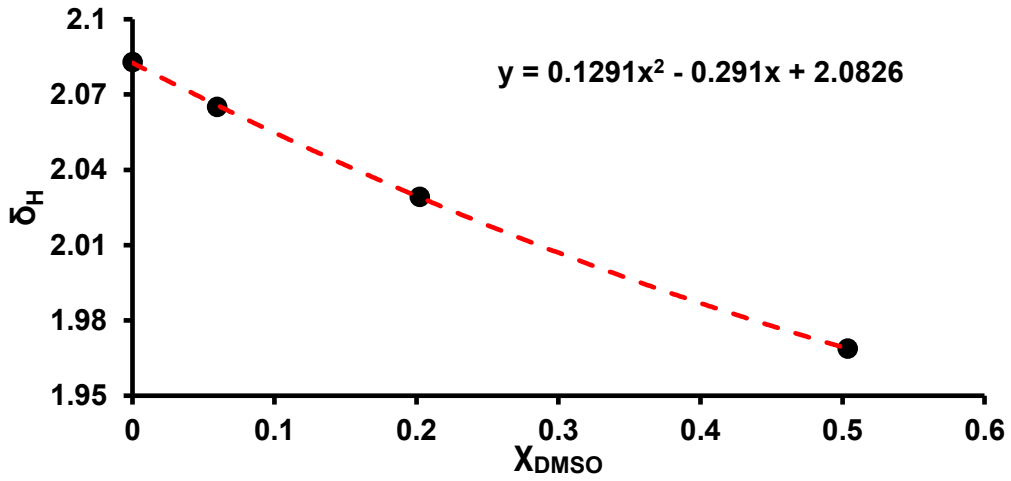


Figure D.15. Plot of δ_H of sodium acetate as a function of mole fraction of DMSO with the respective fitting and equation

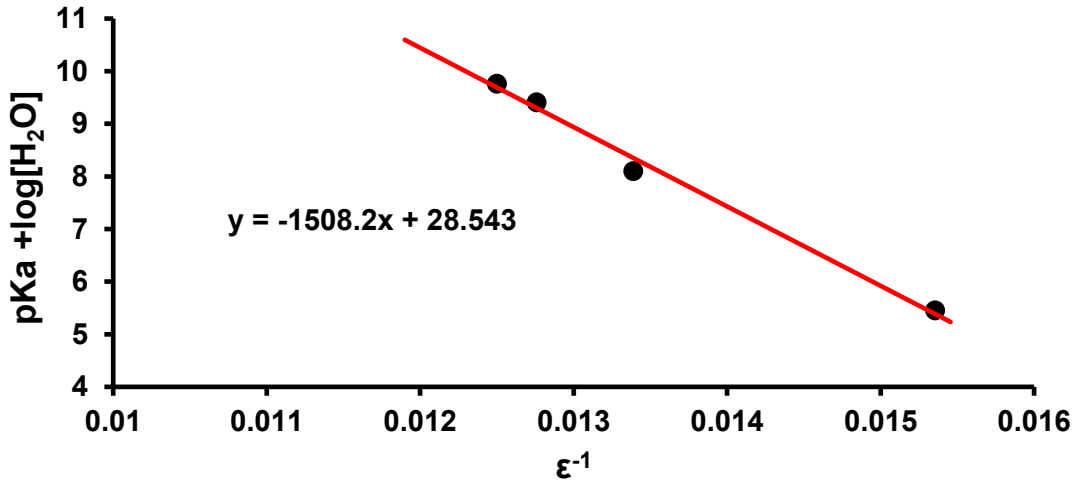


Figure D.16. Plot of $\log[H_2O] + p_sK_a$ versus $1/\varepsilon$ for 2-methylimidazole with respective fitting

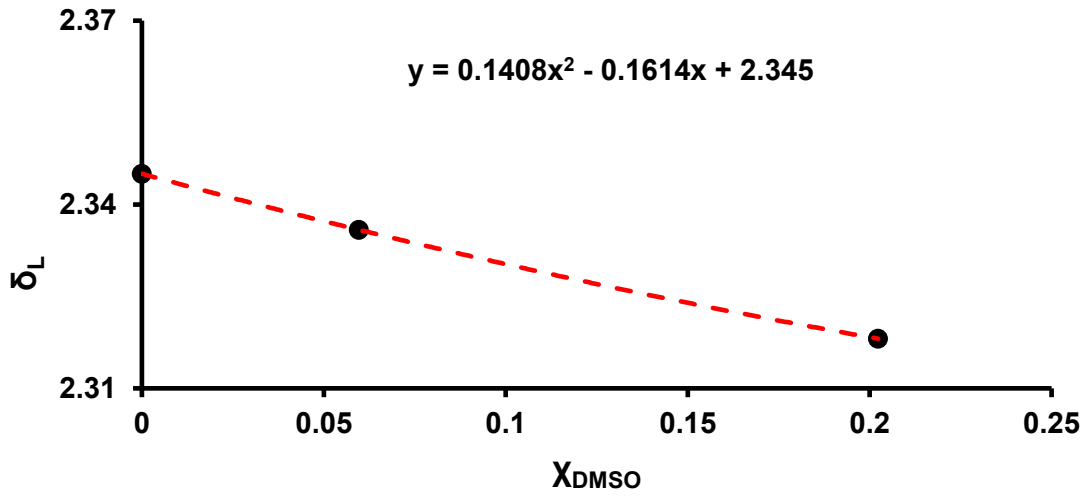


Figure D.17. Plot of δ_L of 2-methylimidazole as a function of mole fraction of DMSO with the respective fitting and equation

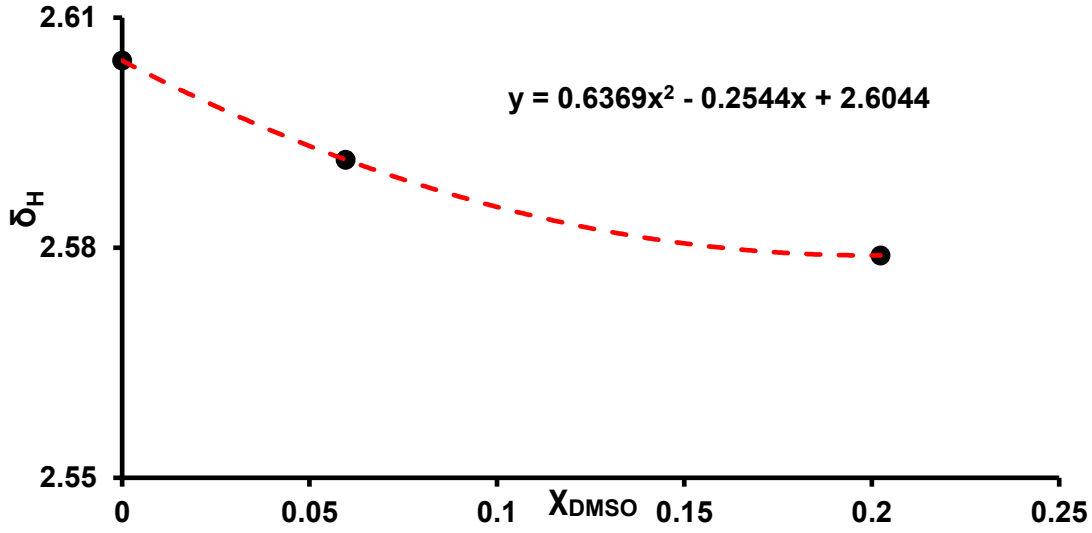


Figure D.18. Plot of δ_H of 2-methylimidazole as a function of mole fraction of DMSO with the respective fitting and equation

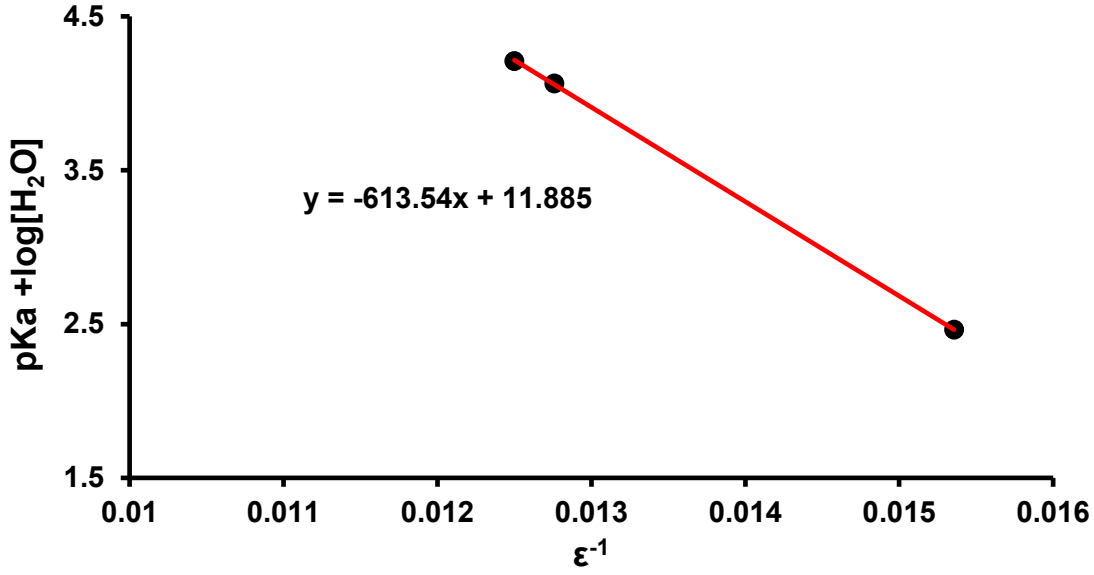


Figure D.19. Plot of log[H_2O] + p_sK_a versus 1/ε for 1,2,4-triazole with respective fitting

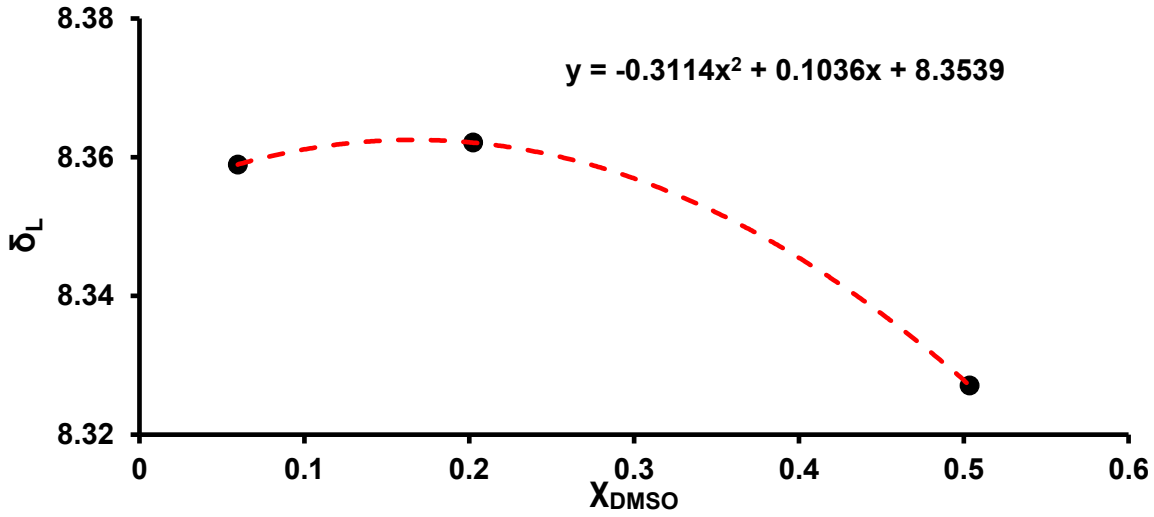


Figure D.20. Plot of δ_L of 1,2,4-triazole as a function of mole fraction of DMSO with the respective fitting and equation

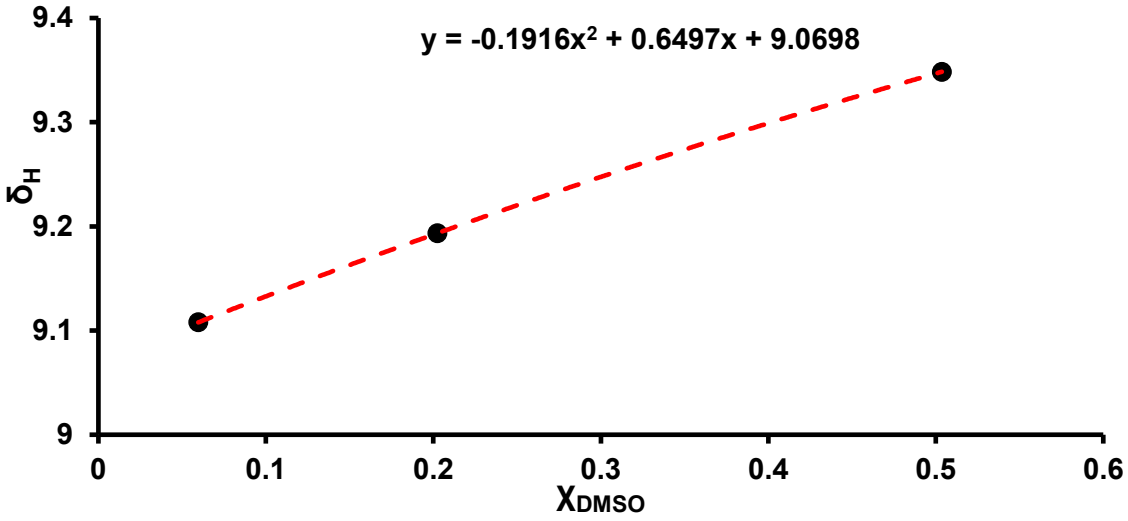


Figure D.21. Plot of δ_H of 1,2,4-triazole as a function of mole fraction of DMSO with the respective fitting and equation

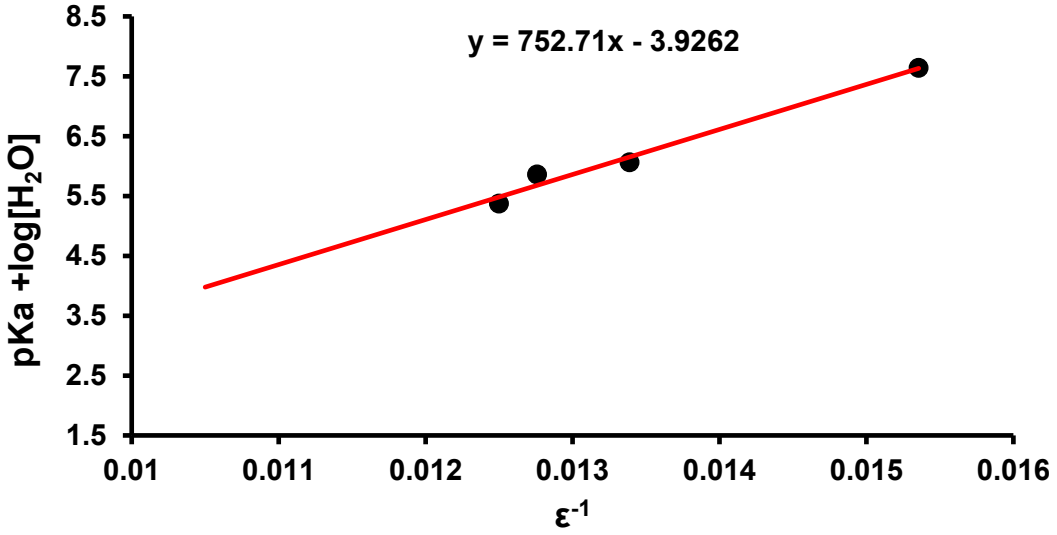


Figure D.22. Plot of $\log[H_2O] + p_sK_a$ versus $1/\varepsilon$ for sodium formate with respective fitting

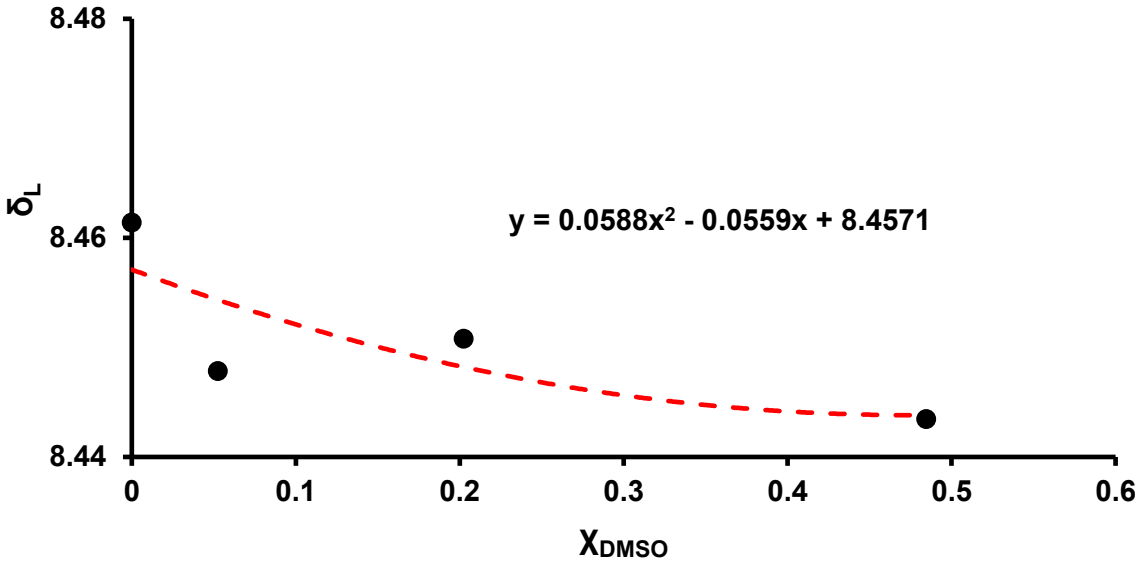


Figure D.23. Plot of δ_L of sodium formate as a function of mole fraction of DMSO with the respective fitting and equation

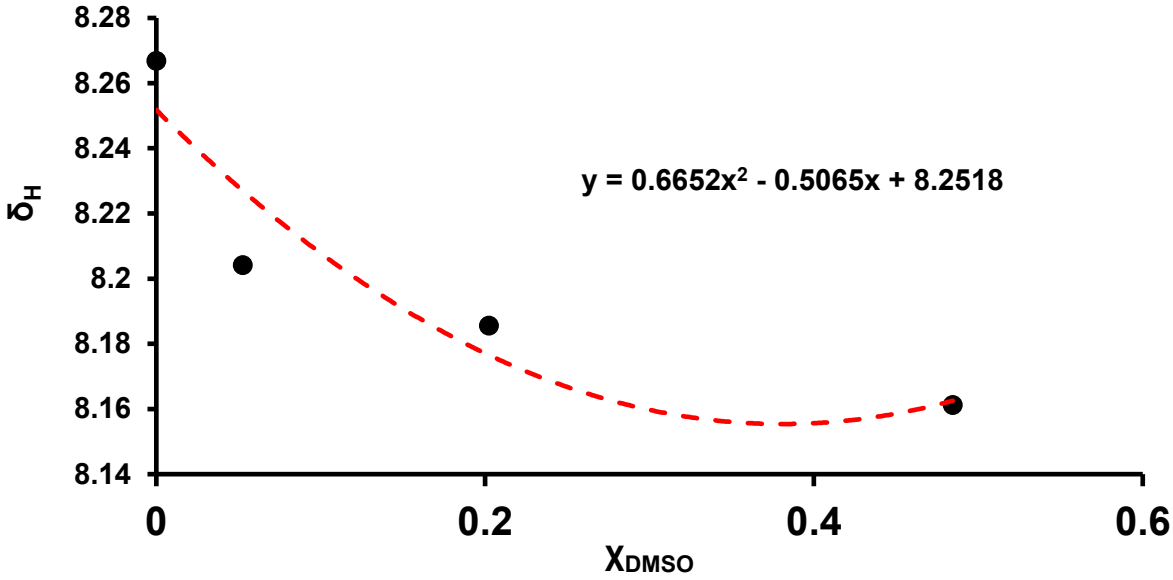


Figure D.24. Plot of δ_H of sodium formate as a function of mole fraction of DMSO with the respective fitting and equation

Appendix Chapter 5

E.1 Estimation of uncertainty in the determination of $[M^{2+}]_f$ and B, with spreadsheet calculation

The uncertainty in γ_2, $\Delta\gamma_2$, is estimated by comparing the value obtained from Equation 5.7 with the value obtained from the reported $[M^{2+}]_f$:[209]

$$\Delta\gamma_2 = \gamma_2 - 10^{-2.04\left[\frac{\sqrt{3[M^{2+}]_f+I_0}}{1+\sqrt{3[M^{2+}]_f+I_0}}-0.3(3[M^{2+}]_f+I_0)\right]} \quad \text{(E.1)}$$

where I_0 is the ionic strength of the sample prior to addition of M^{2+}. The uncertainty in $[M^{2+}]_f$, $\Delta[M^{2+}]_f$, is obtained by propagation of uncertainty analysis of Equation 5.3:

$$\Delta[M^{2+}]_f = \sqrt{\left(\frac{\partial[M^{2+}]_f}{\partial\delta_{obs}}\right)^2 \Delta\delta_{obs}{}^2 + \left(\frac{\partial[M^{2+}]_f}{\partial\delta_L}\right)^2 \Delta\delta_L{}^2 + \left(\frac{\partial[M^{2+}]_f}{\partial\delta_M}\right)^2 \Delta\delta_M{}^2 + \left(\frac{\partial[M^{2+}]_f}{\partial K_M}\right)^2 \Delta K_M{}^2} \quad \text{(E.2)}$$

where Δ indicates the uncertainty in the variable.

Differentiating Equation 5.3 with respect to each variable:

$$\frac{\partial[M^{2+}]_f}{\partial\delta_L} = \frac{1}{K_M(\delta_{obs}-\delta_M)} \quad \text{(E.3)}$$

$$\frac{\partial[M^{2+}]_f}{\partial\delta_{obs}} = \frac{\delta_M - \delta_L}{K_M(\delta_{obs}-\delta_M)^2} \quad \text{(E.4)}$$

$$\frac{\partial[M^{2+}]_f}{\partial K_M} = \frac{\delta_{obs} - \delta_L}{K_M{}^2(\delta_{obs}-\delta_M)} \quad \text{(E.5)}$$

$$\frac{\partial[M^{2+}]_f}{\partial\delta_M} = \frac{\delta_L - \delta_{obs}}{K_M(\delta_{obs}-\delta_M)^2} \quad \text{(E.6)}$$

$\Delta\delta_L$, $\Delta\delta_{obs}$ and $\Delta\delta_M$ are taken as 0.0005 ppm in this work.[131]

E.2 Ligand binding constants and limiting chemical shifts

Table E.1. Chemical shifts of free (d_L) and fully complexed (d_{Ca}, d_{Mg}) glycolate and sulfoacetate ligands and binding constants used in this work, taken from Reference [209]. Chemical shifts are reported relative to DMSO at 2.72 ppm in 100% H_2O. Shifts relative to DSS and methanol, and in 90% H_2O/10% D_2O are provided in our previous work. a: Reference [209]. b: Reference [239]. c: Glycolate and sulfoacetate are unable to distinguish between free Mg^{2+} ions and Mg-Cl ion pairs and the binding constant is set to zero in the present work for simplicity.[209]

Ligand	**δ_{Ca}/ppm**	**δ_{Mg}/ppm**	**δ_L/ppm**	**$K_{0,Ca}/M^{-1}$**	**$K_{0,Mg}/M^{-1}$**
Glycolate	4.0977	4.0471	3.9353	34.7±0.06	25.4±0.6
Sulfoacetate	3.8038	3.7812	3.7332	144.4±0.7	117.4±2.4
Acetate[a]	-	-	-	5.89	1.10
Chloride	-	-	-	2.63[b]	0[c]

E.3 CSI analysis of M^{2+} gradients at different times since preparation

The time at which a CSI experiment was run since preparation of the sample was rounded to the nearest half hour. This time was used in the calculation of N (Equation 5.10).

Table E.2. Time elapsed between preparation and analysis of CSI samples

Sample	Time since preparation/hour
CMC 1.2 (Mg)	8
CMC 1.2 (Ca)	8.5
CMC 0.7 (Mg)	8
CMC 0.7 (Ca)	8.5
Oxalate (Mg)	9
Oxalate (Ca)	8.5
Gallate (Mg)	7
Gallate (Ca)	6.5
4 mg/mL sodium alginate (Mg)	7.5
4 mg/mL sodium alginate (Ca)	7
2 mg/mL sodium alginate (Mg)	8.5
2 mg/mL sodium alginate (Ca)	8.5
Citrate CNC (Mg)	8.5
Citrate CNC (Ca)	8
NaCl (Mg)	5.5
NaCl (Ca)	5

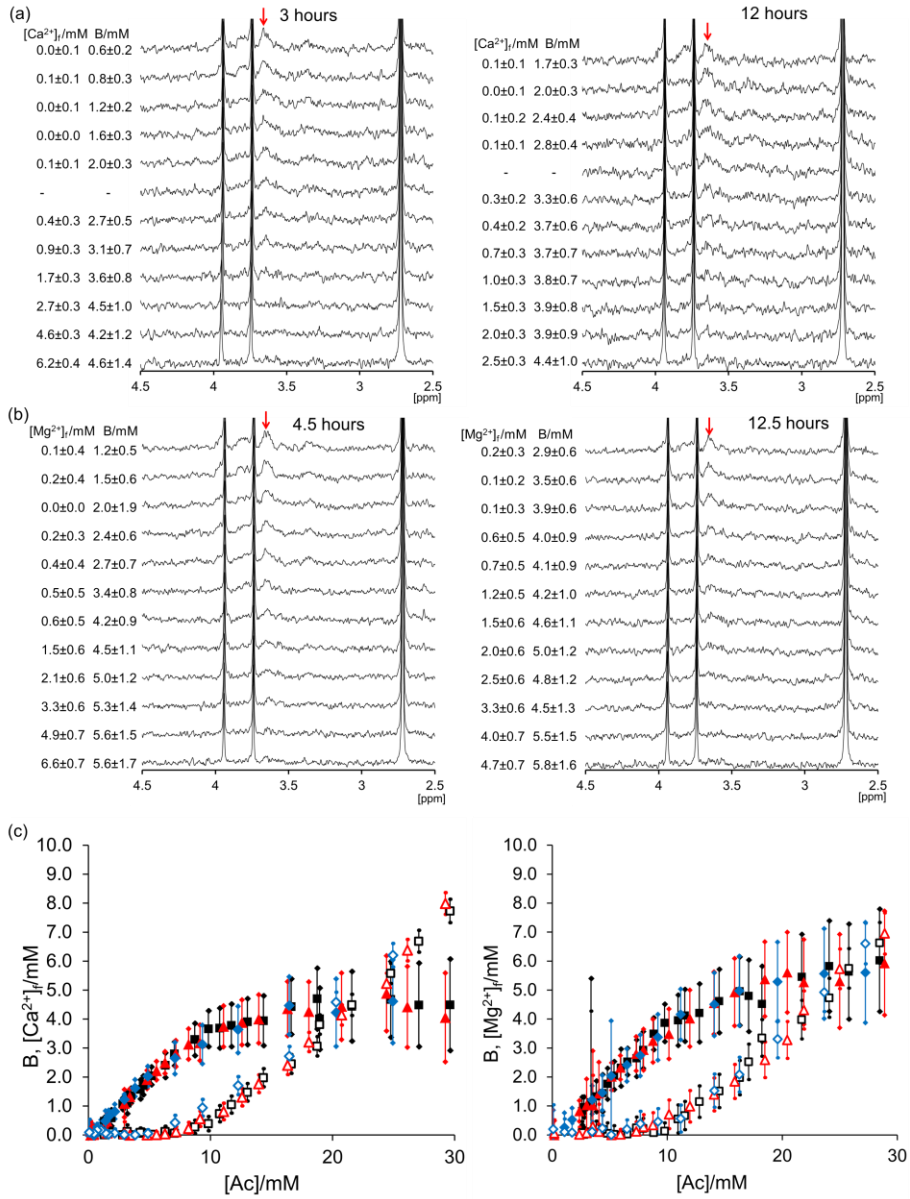


Figure E.1.2 wt% citrate CNC: ^{1}H spectra extracted from citrate functionalised CNC CSI datasets measured on sample of Figure 5.6 and 5.7 at time indicated since preparation. (a) Ca, (b) Mg. (c) Plot of B (solid symbols) and $[M^{2+}]_f$ (open) when calcium acetate (left) or magnesium acetate (right) was diffused into solution of citrate CNC. The sample was analysed at the following times since preparation: Ca: 3 hours (blue diamond), 8 hours (red triangle), 12 hours (black square). Mg: 4.5 hours (blue diamond), 8.5 hours (red triangle), 12.5 hours (black square).

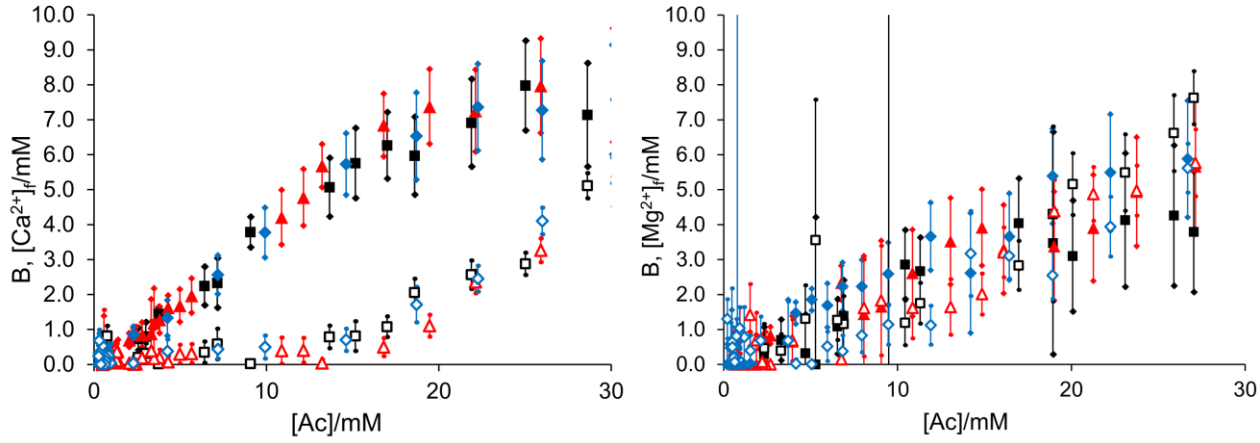


Figure E.2. Plot of B (solid symbols) and $[M^{2+}]_f$ (open symbols) when calcium acetate (left) or magnesium acetate (right) was diffused into 4 mg/mL sodium alginate. The sample was analysed at the following times since preparation: Ca: 4.5 hours (blue diamond), 7 hours (red triangle), 10 hours (black square). Mg: 4.5 hours (blue diamond), 7.5 hours (red triangle), 10.5 hours (black square).

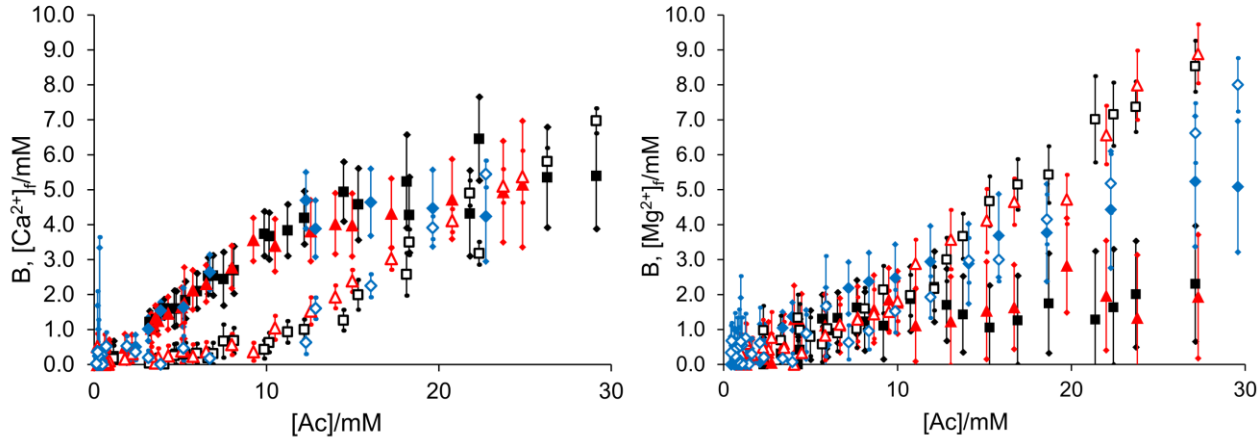


Figure E.3. Plot of B (solid symbols) and $[M^{2+}]_f$ (open symbols) when calcium acetate (left) or magnesium acetate (right) was diffused into 2 mg/mL sodium alginate. The sample was analysed at the following times since preparation: Ca: 4.5 hours (blue diamond), 8.5 hours (red triangle), 12.5 hours (black square). Mg: 5 hours (blue diamond), 8.5 hours (red triangle), 11.5 hours (black square).

E.4 Determination of carboxylate concentration of CMC 0.7 and 1.2

Carboxylate content was to be determined via the method previously made by Wallace *et al* .[131] Briefly, disodium methylphosphonate (MPA, 2 mM) was acidified with HCl (20 mM) and a ^{1}H NMR experiment was ran with MPA on its own to obtain a reference chemical shift (δ_{ref}) of the indicator without the presence of CMC. Another ^{1}H NMR experiment was done with CMC (0.7 or 1.2, 10 mM), HCl (20 mM), MPA (0.2 mM) and H_2O to calculate the dissociation constant of MPA (K_{eff}) using the equation below, where δ_L and δ_H are the deprotonated and protonated limiting chemical shifts of MPA and [HCl] is total concentration of HCl added to the reference and $[MPA]_{total}$ is the total concentration of MPA.

$$K_{eff} = \left([HCL] - [MPA]_{total}\left[1 + \frac{\delta_{ref} - \delta_L}{\delta_H - \delta_L}\right]\right)\left(\frac{\delta_H - \delta_{ref}}{\delta_{ref} - \delta_L}\right) \quad \text{(E.7)}$$

From the value of K_{eff}, the carboxyl content of the two substitutions of CMC can be found using equation D.8 below: [131]

$$COO(H) = K_{eff}\left(\frac{\delta_{ref} - \delta_L}{\delta_H - \delta_{ref}} - \frac{\delta_{obs} - \delta_L}{\delta_H - \delta_{obs}}\right) \quad \text{(E.8)}$$

The values for total concentration of carboxyl sites of the substitutions were calculated and are tabulated below.

Table E.3. Carboxylate concentration of CMC 0.7 and CMC 1.2

	CMC 0.7	*CMC 1.2*
Analyte concentration/mM	10	10
Total COOH concentration/mM	6.71	8.76

(1) Tampieri, A.; Szabó, M.; Medina, F.; Gulyás, H. A brief introduction to the basics of NMR spectroscopy and selected examples of its applications to materials characterization. *Physical Sciences Reviews* **2021**, *6* (1). DOI: doi:10.1515/psr-2019-0086 (accessed 2024-07-22).
(2) Levitt, M. H. *Spin dynamics: basics of nuclear magnetic resonance*; John Wiley & Sons, 2008.
(3) Hayashi, H. *Introduction to dynamic spin chemistry: magnetic field effects on chemical and biochemical reactions*; World Scientific, 2004.
(4) Keeler, J. *Understanding NMR spectroscopy*; John Wiley & Sons, 2010.
(5) Vitha, M. F. *Spectroscopy: Principles and instrumentation*; John Wiley & Sons, 2018.
(6) Ward, P. L. On the Planck-Einstein Relation. 2020.
(7) Rigden, J. S. Quantum states and precession: The two discoveries of NMR. *Reviews of Modern Physics* **1986**, *58* (2), 433–448. DOI: 10.1103/RevModPhys.58.433.
(8) Haris, R. Nuclear Magnetic Resonance Spectroscopy: A Physiochemical View. *Pitman, London, 1983á SearcháPubMed á* **1986**, 67.
(9) Kaushik, D.; Kaushik, M. K. *Modern Pharmaceutical Analytical Techniques: A Textbook*; Shashwat Publication, 2024.
(10) Breitmaier, E. *Structure elucidation by NMR in organic chemistry: a practical guide*; John Wiley & Sons, 2002.
(11) Purslow, J. A.; Khatiwada, B.; Bayro, M. J.; Venditti, V. NMR methods for structural characterization of protein-protein complexes. *Frontiers in molecular biosciences* **2020**, *7*, 9.
(12) Reo, N. V. NMR-based metabolomics. *Drug and chemical toxicology* **2002**, *25* (4), 375–382.
(13) Mahajan, S.; Singh, I. P. Determining and reporting purity of organic molecules: why qNMR. *Magnetic Resonance in Chemistry* **2013**, *51* (2), 76–81. DOI: https://doi.org/10.1002/mrc.3906.
(14) Pellecchia, M.; Sem, D. S.; Wüthrich, K. Nmr in drug discovery. *Nature Reviews Drug Discovery* **2002**, *1* (3), 211–219. DOI: 10.1038/nrd748.
(15) Chmelka, B. F. Materializing opportunities for NMR of solids. *Journal of Magnetic Resonance* **2019**, *306*, 91–97. DOI: https://doi.org/10.1016/j.jmr.2019.07.051.
(16) Parella, T. Pulsed field gradients: a new tool for routine NMR. *Magnetic Resonance in Chemistry* **1998**, *36* (7), 467–495. DOI: https://doi.org/10.1002/(SICI)1097-458X(199807)36:7<467::AID-OMR325>3.0.CO;2-S.

(17) Pengra, D. B. Exploring the Rotating Reference Frame in Off-Resonance Pulsed NMR. *signal 2*, 23.
(18) Hidalgo-Tobon, S. S. Theory of gradient coil design methods for magnetic resonance imaging. *Concepts in Magnetic Resonance Part A* **2010**, *36A* (4), 223–242. DOI: https://doi.org/10.1002/cmr.a.20163.
(19) Scientific, D. *PFG/Diffusion Liquids Z Gradient Probes-Pulsed Field Gradient Coils 0ver 3,300 G/cm (33 T/m)*. Resonating eNewsletter, 2024. https://dotynmr.com/products/diffusion-pfg-probes/ (accessed 2025 26/02/2025).
(20) Roberts, J. D. *ABC's of FT-NMR*; University Science Books, 2000.
(21) Keeler, J. *Understanding NMR spectroscopy*; John Wiley & Sons, 2010.
(22) Kay, L. E. Pulsed field gradient multi-dimensional NMR methods for the study of protein structure and dynamics in solution. *Progress in Biophysics and Molecular Biology* **1995**, *63* (3), 277–299. DOI: https://doi.org/10.1016/0079-6107(95)00007-0.
(23) Price, W. S.; Hayamizu, K.; Arata, Y. Optimization of the water-PRESS pulse sequence and its integration into pulse sequences for studying biological macromolecules. *J Magn Reson* **1997**, *126* (2), 256–265. DOI: 10.1006/jmre.1997.1166 From NLM.
(24) Claridge, T. D. W. Chapter 5 - Correlations through the chemical bond I: Homonuclear shift correlation. In *Tetrahedron Organic Chemistry Series*, Claridge, T. D. W. Ed.; Vol. 27; Elsevier, 2009; pp 129–188.
(25) Groves, P. Diffusion ordered spectroscopy (DOSY) as applied to polymers. *Polymer Chemistry* **2017**, *8* (44), 6700–6708, 10.1039/C7PY01577A. DOI: 10.1039/C7PY01577A.
(26) Johnson, C. S. Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications. *Progress in Nuclear Magnetic Resonance Spectroscopy* **1999**, *34* (3), 203–256. DOI: https://doi.org/10.1016/S0079-6565(99)00003-5.
(27) Morris, K. F.; Johnson, C. S., Jr. Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. *Journal of the American Chemical Society* **1992**, *114* (8), 3139–3141. DOI: 10.1021/ja00034a071.
(28) Keeler, J. *Understanding NMR spectroscopy*; John Wiley & Sons, 2010.
(29) Keeler, J. *Understanding NMR spectroscopy*; John Wiley & Sons, 2010.
(30) Lindon, J. C. NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY TECHNIQUES | Multidimensional Proton. In *Encyclopedia of Analytical Science (Second Edition)*, Worsfold, P., Townshend, A., Poole, C. Eds.; Elsevier, 2005; pp 350–357.
(31) Claridge, T. D. W. Chapter 6 - Correlations Through the Chemical Bond I: Homonuclear Shift Correlation. In *High-Resolution NMR Techniques in Organic Chemistry (Third Edition)*, Claridge, T. D. W. Ed.; Elsevier, 2016; pp 203–241.
(32) Stejskal, E. O.; Tanner, J. E. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics* **1965**, *42* (1), 288–292.
(33) Padhani, A. R.; Liu, G.; Koh, D. M.; Chenevert, T. L.; Thoeny, H. C.; Takahara, T.; Dzik-Jurasz, A.; Ross, B. D.; Van Cauteren, M.; Collins, D.; et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia* **2009**, *11* (2), 102–125. DOI: 10.1593/neo.81328 From NLM.
(34) Lauterbur, P. C. Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature* **1973**, *242* (5394), 190–191. DOI: 10.1038/242190a0.
(35) Mansfield, P. Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics* **1977**, *10* (3), L55. DOI: 10.1088/0022-3719/10/3/004.
(36) Mansfield, P.; Grannell, P. K. NMR'diffraction'in solids? *Journal of Physics C: solid state physics* **1973**, *6* (22), L422.
(37) Slavkovsky, P.; Uhliar, R. The nobel prize in physiology or medicine in 2003 to paul c. lauterbur, peter mansfield magnetic resonance imaging. *Bratislavske lekarske listy* **2004**, *105* (7/8), 245–249.
(38) Tušar, K.; Serša, I. Use of nonlinear pulsed magnetic fields for spatial encoding in magnetic resonance imaging. *Scientific Reports* **2024**, *14* (1), 7521. DOI: 10.1038/s41598-024-58229-x.

(39) Dechambre, S.; Duprez, T.; Lecouvet, F.; Raftopoulos, C.; Gosnard, G. Diffusion-weighted MRI postoperative assessment of an epidermoid tumour in the cerebellopontine angle. *Neuroradiology* **1999**, *41* (11), 829–831. DOI: 10.1007/s002340050850.
(40) Trigo-Mouriño, P.; Merle, C.; Koos, M. R. M.; Luy, B.; Gil, R. R. Probing Spatial Distribution of Alignment by Deuterium NMR Imaging. *Chemistry – A European Journal* **2013**, *19* (22), 7013–7019. DOI: https://doi.org/10.1002/chem.201300254 (accessed 2024/06/04).
(41) Santos-Díaz, A.; Noseworthy, M. D. Phosphorus magnetic resonance spectroscopy and imaging (31P-MRS/MRSI) as a window to brain and muscle metabolism: A review of the methods. *Biomedical Signal Processing and Control* **2020**, *60*, 101967. DOI: https://doi.org/10.1016/j.bspc.2020.101967.
(42) Rodriguez, A. Principles of magnetic resonance imaging. *Revista mexicana de física* **2004**, *50* (3), 272–286.
(43) Brink, H. F.; Buschmann, M. D.; Rosen, B. R. NMR chemical shift imaging. *Computerized Medical Imaging and Graphics* **1989**, *13* (1), 93–104. DOI: https://doi.org/10.1016/0895-6111(89)90081-5.
(44) Brown, T. R.; Kincaid, B.; Ugurbil, K. NMR chemical shift imaging in three dimensions. *Proceedings of the National Academy of Sciences* **1982**, *79* (11), 3523–3526.
(45) Maudsley, A. A.; Hilal, S. K.; Perman, W. H.; Simon, H. E. Spatially resolved high resolution spectroscopy by "four-dimensional" NMR. *Journal of Magnetic Resonance (1969)* **1983**, *51* (1), 147–152. DOI: https://doi.org/10.1016/0022-2364(83)90113-0.
(46) Mansfield, P. Spatial mapping of the chemical shift in NMR. *Magn Reson Med* **1984**, *1* (3), 370–386. DOI: 10.1002/mrm.1910010308 From NLM.
(47) Mansfield, P.; Maudsley, A. A. Medical imaging by NMR. *Br J Radiol* **1977**, *50* (591), 188–194. DOI: 10.1259/0007-1285-50-591-188 From NLM.
(48) Evans, R.; Sandhu, A.; Bridgwater, T.; Chong, K. Slice-Selective NMR: A Noninvasive Method for the Analysis of Separated Pyrolysis Fuel Samples. *Energy & Fuels* **2017**, *31* (4), 4135–4142. DOI: 10.1021/acs.energyfuels.7b00223.
(49) Edelstein, W. A.; Hutchison, J. M.; Johnson, G.; Redpath, T. Spin warp NMR imaging and applications to human whole-body imaging. *Phys Med Biol* **1980**, *25* (4), 751–756. DOI: 10.1088/0031-9155/25/4/017 From NLM.
(50) Jung, B. A.; Weigel, M. Spin echo magnetic resonance imaging. *Journal of Magnetic Resonance Imaging* **2013**, *37* (4), 805–817. DOI: https://doi.org/10.1002/jmri.24068.
(51) Trigo-Mouriño, P.; Merle, C.; Koos, M. R.; Luy, B.; Gil, R. R. Probing spatial distribution of alignment by deuterium NMR imaging. *Chemistry* **2013**, *19* (22), 7013–7019. DOI: 10.1002/chem.201300254 From NLM.
(52) Koźmiński, W. Application of spatially resolved NMR spectroscopy for high resolution spectra of heterogeneous samples. *Polish Journal of Chemistry* **2000**, *74*, 1185–1189.
(53) Ross, A.; Schlotterbeck, G.; Senn, H.; von Kienlin, M. Application of Chemical Shift Imaging for Simultaneous and Fast Acquisition of NMR Spectra on Multiple Samples. *Angewandte Chemie International Edition* **2001**, *40* (17), 3243–3245. DOI: https://doi.org/10.1002/1521-3773(20010903)40:17<3243::AID-ANIE3243>3.0.CO;2-F.
(54) Niklas, T.; Stalke, D.; John, M. Single-shot titrations and reaction monitoring by slice-selective NMR spectroscopy. *Chemical Communications* **2015**, *51* (7), 1275–1277.
(55) Mitrev, Y.; Simova, S.; Jeannerat, D. NMR analysis of weak molecular interactions using slice-selective experiments via study of concentration gradients in agar gels. *Chemical Communications* **2016**, *52* (31), 5418–5420, 10.1039/C6CC01853J. DOI: 10.1039/C6CC01853J.
(56) Wallace, M.; Cardoso, A. Z.; Frith, W. J.; Iggo, J. A.; Adams, D. J. Magnetically Aligned Supramolecular Hydrogels. *Chemistry – A European Journal* **2014**, *20* (50), 16484–16487. DOI: https://doi.org/10.1002/chem.201405500.
(57) Wallace, M.; Adams, D. J.; Iggo, J. A. Titrations without the Additions: The Efficient Determination of pKa Values Using NMR Imaging Techniques. *Analytical Chemistry* **2018**, *90* (6), 4160–4166. DOI: 10.1021/acs.analchem.8b00181.

(58) Monaco, S.; Angulo, J.; Wallace, M. Imaging Saturation Transfer Difference (STD) NMR: Affinity and Specificity of Protein–Ligand Interactions from a Single NMR Sample. *Journal of the American Chemical Society* **2023**, *145* (30), 16391–16397. DOI: 10.1021/jacs.3c02218.
(59) Wallace, M.; Sharpe, J. M.; Baj, K.; Ngwube, M.; Thirlway, J.; Kerigan Higgs, P. L.; Stephenson, G. R.; Iggo, J. A.; Storr, T. E.; Richards, C. J. Programmable wide-range pH gradients for NMR titrations: application to antibody–drug conjugate linker group modifications. *Analyst* **2025**, 10.1039/D5AN00406C. DOI: 10.1039/D5AN00406C.
(60) Hwang, T. L.; Shaka, A. J. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *Journal of Magnetic Resonance, Series A* **1995**, *112* (2), 275–279. DOI: https://doi.org/10.1006/jmra.1995.1047.
(61) Brown, M. A.; Semelka, R. C. *MRI: basic principles and applications*; John Wiley & Sons, 2011.
(62) Cho, Z. H.; Kim, D. J.; Kim, Y. K. Total inhomogeneity correction including chemical shifts and susceptibility by view angle tilting. *Medical Physics* **1988**, *15* (1), 7–11. DOI: https://doi.org/10.1118/1.596162.
(63) Park, H.; Kim, Y.; Cho, Z. Fast gradient-echo chemical-shift imaging. *Magnetic resonance in medicine* **1988**, *7* (3), 340–345.
(64) Callaghan, P. T. *Translational dynamics and magnetic resonance: principles of pulsed gradient spin echo NMR*; Oxford University Press, 2011.
(65) Claridge, T. D. *High-resolution NMR techniques in organic chemistry*; Elsevier, 2016.
(66) Callaghan, P. T.; Eccles, C. D. Sensitivity and resolution in NMR imaging. *Journal of Magnetic Resonance (1969)* **1987**, *71* (3), 426–445. DOI: https://doi.org/10.1016/0022-2364(87)90243-5.
(67) Frydman, L.; Scherf, T.; Lupulescu, A. The acquisition of multidimensional NMR spectra within a single scan. *Proc Natl Acad Sci U S A* **2002**, *99* (25), 15858–15862. DOI: 10.1073/pnas.252644399 From NLM.
(68) Mobli, M.; Hoch, J. C. Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2014**, *83*, 21–41. DOI: https://doi.org/10.1016/j.pnmrs.2014.09.002.
(69) Reijenga, J.; Van Hoof, A.; Van Loon, A.; Teunissen, B. Development of methods for the determination of pKa values. *Analytical chemistry insights* **2013**, *8*, ACI. S12304.
(70) Alkhzem, A. H. Medicinal Chemistry of Aminoglycosides. University of Bath, 2018.
(71) Avdeef, A. pH-metric log P. II: Refinement of partition coefficients and ionization constants of multiprotic substances. *J Pharm Sci* **1993**, *82* (2), 183–190. DOI: 10.1002/jps.2600820214 From NLM.
(72) Analytical, R. *pH theory and practice*; Villeurbanne, France, 2007. https://www.radiometer-analytical.com/GE3.
(73) Reijenga, J.; van Hoof, A.; van Loon, A.; Teunissen, B. Development of Methods for the Determination of pKa Values. *Anal Chem Insights* **2013**, *8*, 53–71. DOI: 10.4137/aci.S12304 From NLM.
(74) Pathare, B.; Tambe, V.; Dhole, S.; Patil, V. An update on various analytical techniques based on UV spectroscopy used in determination of dissociation constant. *Int. J. Pharm* **2014**, *4* (1), 278–285.
(75) Chemicals, E. C. f. E. a. T. o. *ECETOC Technical Report No. 123: Environmental Toxicity Testing of Difficult Substances and Mixtures*; ECETOC, 2014. https://www.ecetoc.org/technical-report-123/appendix-a-measurement-of-acidity-pka/ (accessed 26 March 2025).
(76) Wang, H.; Xiang, H.; Xiong, T.; Feng, J.; Zhang, J.; Li, X. A straightforward approach utilizing an exponential model to compensate for turbidity in chemical oxygen demand measurements using UV-vis spectrometry. *Front Microbiol* **2023**, *14*, 1224207. DOI: 10.3389/fmicb.2023.1224207 From NLM.
(77) Harroun, S. G.; Lauzon, D.; Ebert, M. C. C. J. C.; Desrosiers, A.; Wang, X.; Vallée-Bélisle, A. Monitoring protein conformational changes using fluorescent nanoantennas. *Nature Methods* **2022**, *19* (1), 71–80. DOI: 10.1038/s41592-021-01355-5.
(78) Martínez, M.; Martínez, N. A.; Silva, W. I. Measurement of the Intracellular Calcium Concentration with Fura-2 AM Using a Fluorescence Plate Reader. *Bio Protoc* **2017**, *7* (14), e2411. DOI: 10.21769/BioProtoc.2411 From NLM.

(79) Wu, N.; Nishioka, W. K.; Derecki, N. C.; Maher, M. P. High-throughput-compatible assays using a genetically-encoded calcium indicator. *Scientific Reports* **2019**, *9* (1), 12692. DOI: 10.1038/s41598-019-49070-8.
(80) Paredes, R. M.; Etzler, J. C.; Watts, L. T.; Zheng, W.; Lechleiter, J. D. Chemical calcium indicators. *Methods* **2008**, *46* (3), 143–151. DOI: 10.1016/j.ymeth.2008.09.025 From NLM.
(81) Hong, S.; Pawel, G. T.; Pei, R.; Lu, Y. Recent progress in developing fluorescent probes for imaging cell metabolites. *Biomed Mater* **2021**, *16* (4). DOI: 10.1088/1748-605X/abfd11 From NLM.
(82) Robson, A. L.; Dastoor, P. C.; Flynn, J.; Palmer, W.; Martin, A.; Smith, D. W.; Woldu, A.; Hua, S. Advantages and Limitations of Current Imaging Techniques for Characterizing Liposome Morphology. *Front Pharmacol* **2018**, *9*, 80. DOI: 10.3389/fphar.2018.00080 From NLM.
(83) Panalytical, M. *Isothermal Titration Calorimetry (ITC)*. 2024. https://www.malvernpanalytical.com/en/products/technology/microcalorimetry/isothermal-titration-calorimetry (accessed 2025 March 27).
(84) Buurma, N. J.; Haq, I. Advances in the analysis of isothermal titration calorimetry data for ligand–DNA interactions. *Methods* **2007**, *42* (2), 162–172. DOI: https://doi.org/10.1016/j.ymeth.2007.01.010.
(85) Velazquez-Campoy, A.; Freire, E. Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat Protoc* **2006**, *1* (1), 186–191. DOI: 10.1038/nprot.2006.28 From NLM.
(86) Sciences, H. I. o. t. L. *Isothermal Titration Calorimetry – Technical Guides*. 2024. https://www.huck.psu.edu/core-facilities/automated-biological-calorimetry-facility/technical-guides/isothermal-titration-calorimetry (accessed 2025 March 27).
(87) Rajarathnam, K.; Rösgen, J. Isothermal titration calorimetry of membrane proteins - progress and challenges. *Biochim Biophys Acta* **2014**, *1838* (1 Pt A), 69–77. DOI: 10.1016/j.bbamem.2013.05.023 From NLM.
(88) Minikel, E. *Isothermal titration calorimetry*. 2016. https://www.cureffi.org/2016/09/12/isothermal-titration-calorimetry/ (accessed 2025 March 27).
(89) Bezençon, J.; Wittwer, M. B.; Cutting, B.; Smieško, M.; Wagner, B.; Kansy, M.; Ernst, B. pKa determination by ^{1}H NMR spectroscopy - an old methodology revisited. *J Pharm Biomed Anal* **2014**, *93*, 147–155. DOI: 10.1016/j.jpba.2013.12.014 From NLM.
(90) Barile, E.; Pellecchia, M. NMR-based approaches for the identification and optimization of inhibitors of protein-protein interactions. *Chem Rev* **2014**, *114* (9), 4749–4763. DOI: 10.1021/cr500043b From NLM.
(91) Roos, A.; Boron, W. F. Intracellular pH. *Physiol Rev* **1981**, *61* (2), 296–434. DOI: 10.1152/physrev.1981.61.2.296 From NLM.
(92) Pérez-Sala, D.; Collado-Escobar, D.; Mollinedo, F. Intracellular Alkalinization Suppresses Lovastatin-induced Apoptosis in HL-60 Cells through the Inactivation of a pH-dependent Endonuclease (∗). *Journal of Biological Chemistry* **1995**, *270* (11), 6235–6242. DOI: https://doi.org/10.1074/jbc.270.11.6235.
(93) Bullock, A. J.; Duquette, R. A.; Buttell, N.; Wray, S. Developmental changes in intracellular pH buffering power in smooth muscle. *Pflügers Archiv* **1998**, *435* (4), 575–577. DOI: 10.1007/s004240050555.
(94) Chin, E. R.; Allen, D. G. The contribution of pH-dependent mechanisms to fatigue at different intensities in mammalian single muscle fibres. *J Physiol* **1998**, *512 (Pt 3)* (Pt 3), 831–840. DOI: 10.1111/j.1469-7793.1998.831bd.x From NLM.
(95) Beer, P. D.; Gale, P. A. Anion Recognition and Sensing: The State of the Art and Future Perspectives. *Angew Chem Int Ed Engl* **2001**, *40* (3), 486–516. From NLM.
(96) Morris, K. L.; Chen, L.; Raeburn, J.; Sellick, O. R.; Cotanda, P.; Paul, A.; Griffiths, P. C.; King, S. M.; O'Reilly, R. K.; Serpell, L. C.; et al. Chemically programmed self-sorting of gelator networks. *Nature Communications* **2013**, *4* (1), 1480. DOI: 10.1038/ncomms2499.

(97) Du, C.; Rammal, M.; Rivard, M. D.; Harrington, M. J.; Mauzeroll, J. Mussels Put Their Best "pH"oot Forward: Importance of pH in Formation of Biological and Bio-Inspired Materials. *Advanced Functional Materials* **2024**, 2409463. DOI: https://doi.org/10.1002/adfm.202409463.
(98) Matthew, J. B.; Gurd, F. R. N.; Garcia-Moreno, B. E.; Flanagan, M. A.; March, K. L.; Shire, S. J. pH-Dependent Processes in Protein. *Critical Reviews in Biochemistry* **1985**, *18* (2), 91–197. DOI: 10.3109/10409238509085133.
(99) Samuelsen, L.; Holm, R.; Lathuile, A.; Schönbeck, C. Buffer solutions in drug formulation and processing: How pKa values depend on temperature, pressure and ionic strength. *International Journal of Pharmaceutics* **2019**, *560*, 357–364. DOI: https://doi.org/10.1016/j.ijpharm.2019.02.019.
(100) Babić, S.; Horvat, A. J. M.; Mutavdžić Pavlović, D.; Kaštelan-Macan, M. Determination of pKa values of active pharmaceutical ingredients. *TrAC Trends in Analytical Chemistry* **2007**, *26* (11), 1043–1061. DOI: https://doi.org/10.1016/j.trac.2007.09.004.
(101) Orme, M. DRUG ABSORPTION IN THE GUT. *BJA: British Journal of Anaesthesia* **1984**, *56* (1), 59–67. DOI: 10.1093/bja/56.1.59 (accessed 12/4/2024).
(102) Wallace, M.; Holroyd, J.; Kuraite, A.; Hussain, H. Does it bind? A method to determine the affinity of calcium and magnesium ions for polymers using 1H NMR spectroscopy. *Analytical Chemistry* **2022**, *94* (31), 10976–10983.
(103) Lasdon, L. S.; Fox, R. L.; Ratner, M. W. Nonlinear optimization using the generalized reduced gradient method. *Revue française d'automatique, informatique, recherche opérationnelle. Recherche opérationnelle* **1974**, *8* (V3), 73–103.
(104) Lasdon, L. S. W., A. D.; Jain, A.; Ratner,. *Design and Testing of a Generalized Reduced Gradient Code for Nonlinear Optimization*; Case Western Reserve University, 1975.
(105) El Mouatasim, A. Two-Phase Generalized Reduced Gradient Method for Constrained Global Optimization. *Journal of Applied Mathematics* **2010**, *2010* (1), 976529. DOI: https://doi.org/10.1155/2010/976529.
(106) Orgován, G.; Noszál, B. Electrodeless, accurate pH determination in highly basic media using a new set of 1H NMR pH indicators. *Journal of Pharmaceutical and Biomedical Analysis* **2011**, *54* (5), 958–964. DOI: https://doi.org/10.1016/j.jpba.2010.11.022.
(107) Crank, J. The Mathematics of Diffusion, 21–24 Oxford University Press. London: 1975.
(108) Evans, R.; Deng, Z.; Rogerson, A. K.; McLachlan, A. S.; Richards, J. J.; Nilsson, M.; Morris, G. A. Quantitative Interpretation of Diffusion-Ordered NMR Spectra: Can We Rationalize Small Molecule Diffusion Coefficients? *Angewandte Chemie International Edition* **2013**, *52* (11), 3199–3202. DOI: https://doi.org/10.1002/anie.201207403.
(109) Holz, M.; Weingartner, H. Calibration in accurate spin-echo self-diffusion measurements using 1H and less-common nuclei. *Journal of Magnetic Resonance (1969)* **1991**, *92* (1), 115–125. DOI: https://doi.org/10.1016/0022-2364(91)90252-O.
(110) Kukić, P.; Farrell, D.; Søndergaard, C. R.; Bjarnadottir, U.; Bradley, J.; Pollastri, G.; Nielsen, J. E. Improving the analysis of NMR spectra tracking pH-induced conformational changes: Removing artefacts of the electric field on the NMR chemical shift. *Proteins: Structure, Function, and Bioinformatics* **2010**, *78* (4), 971–984. DOI: https://doi.org/10.1002/prot.22621 (accessed 2025/01/27).
(111) Ackerman, J. J.; Soto, G. E.; Spees, W. M.; Zhu, Z.; Evelhoch, J. L. The NMR chemical shift pH measurement revisited: Analysis of error and modeling of a pH dependent reference. *Magnetic resonance in medicine* **1996**, *36* (5), 674–683.
(112) Olver, P. J. *Introduction to Partial Differential Equations*; Springer, 2014.
(113) Djellouli, A. *Central Difference Method*. 2022. https://adamdjellouli.com/articles/numerical_methods/3_differentiation/central_difference (accessed 2025 February 27).
(114) Poole, S. K.; Patel, S.; Dehring, K.; Workman, H.; Poole, C. F. Determination of acid dissociation constants by capillary electrophoresis. *Journal of Chromatography A* **2004**, *1037* (1), 445–454. DOI: https://doi.org/10.1016/j.chroma.2004.02.087.

(115) Irving, H. M.; Miles, M. G.; Pettit, L. D. A study of some problems in determining the stoicheiometric proton dissociation constants of complexes by potentiometric titrations using a glass electrode. *Analytica Chimica Acta* **1967**, *38*, 475–488. DOI: https://doi.org/10.1016/S0003-2670(01)80616-4.
(116) Qiang, Z.; Adams, C. Potentiometric determination of acid dissociation constants (pKa) for human and veterinary antibiotics. *Water Research* **2004**, *38* (12), 2874–2890. DOI: https://doi.org/10.1016/j.watres.2004.03.017.
(117) Belcher, D. The Conductance and Ionization Constants of Propionic and Normal Butyric Acids in Water at 25°. *Journal of the American Chemical Society* **1938**, *60* (11), 2744–2747. DOI: 10.1021/ja01278a058.
(118) Papadopoulos, N.; Avranas, A. Dissociation of salicylic acid, 2,4-, 2,5-and 2,6-dihydroxybenzoic acids in 1-propanol-water mixtures at 25°C. *Journal of Solution Chemistry* **1991**, *20* (3), 293–300. DOI: 10.1007/BF00650443.
(119) Box, K.; Bevan, C.; Comer, J.; Hill, A.; Allen, R.; Reynolds, D. High-Throughput Measurement of pKa Values in a Mixed-Buffer Linear pH Gradient System. *Analytical Chemistry* **2003**, *75* (4), 883–892. DOI: 10.1021/ac020329y.
(120) Wallace, M.; Abiama, N.; Chipembere, M. Measurement of the pKa Values of Organic Molecules in Aqueous–Organic Solvent Mixtures by 1H NMR without External Calibrants. *Analytical Chemistry* **2023**, *95* (42), 15628–15635. DOI: 10.1021/acs.analchem.3c02771.
(121) Hussain, H.; Khimyak, Y. Z.; Wallace, M. Determination of the pKa and Concentration of NMR-Invisible Molecules and Sites Using NMR Spectroscopy. *Analytical Chemistry* **2024**, *96* (50), 19858–19862. DOI: 10.1021/acs.analchem.4c03596.
(122) Patel, P.; Ibrahim, N. M.; Cheng, K. The Importance of Apparent pKa in the Development of Nanoparticles Encapsulating siRNA and mRNA. *Trends Pharmacol Sci* **2021**, *42* (6), 448–460. DOI: 10.1016/j.tips.2021.03.002 From NLM.
(123) Manallack, D. T. The pK(a) Distribution of Drugs: Application to Drug Discovery. *Perspect Medicin Chem* **2007**, *1*, 25–38. From NLM.
(124) Manallack, D. T.; Prankerd, R. J.; Yuriev, E.; Oprea, T. I.; Chalmers, D. K. The significance of acid/base properties in drug discovery. *Chem Soc Rev* **2013**, *42* (2), 485–496. DOI: 10.1039/c2cs35348b From NLM.
(125) Santa Chalarca, C. F.; Dalal, R. J.; Chapa, A.; Hanson, M. G.; Reineke, T. M. Cation Bulk and pKa Modulate Diblock Polymer Micelle Binding to pDNA. *ACS Macro Letters* **2022**, *11* (4), 588–594. DOI: 10.1021/acsmacrolett.2c00015.
(126) Guckeisen, T.; Hosseinpour, S.; Peukert, W. Isoelectric Points of Proteins at the Air/Liquid Interface and in Solution. *Langmuir* **2019**, *35* (14), 5004–5012. DOI: 10.1021/acs.langmuir.9b00311.
(127) Schenck, G.; Baj, K.; Iggo, J. A.; Wallace, M. Efficient pKa Determination in a Nonaqueous Solvent Using Chemical Shift Imaging. *Analytical Chemistry* **2022**, *94* (23), 8115–8119. DOI: 10.1021/acs.analchem.2c00200.
(128) Novoa-Carballal, R.; Martin-Pastor, M.; Fernandez-Megia, E. Unveiling an NMR-Invisible Fraction of Polymers in Solution by Saturation Transfer Difference. *ACS Macro Letters* **2021**, *10* (12), 1474–1479. DOI: 10.1021/acsmacrolett.1c00628.
(129) Zdrazil, B.; Felix, E.; Hunter, F.; Manners, E. J.; Blackshaw, J.; Corbett, S.; de Veij, M.; Ioannidis, H.; Lopez, D. M.; Mosquera, Juan F.; et al. The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Research* **2023**, gkad1004. DOI: 10.1093/nar/gkad1004 (accessed 11/24/2023).
(130) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res* **2000**, *28* (1), 235–242. DOI: 10.1093/nar/28.1.235 From NLM.
(131) Wallace, M.; Lam, K.; Kuraite, A.; Khimyak, Y. Z. Rapid Determination of the Acidity, Alkalinity and Carboxyl Content of Aqueous Samples by 1H NMR with Minimal Sample Quantity. *Analytical Chemistry* **2020**, *92* (19), 12789–12794. DOI: 10.1021/acs.analchem.0c02594.

(132) Satchell, J. F.; Smith, B. J. Calculation of aqueous dissociation constants of 1,2,4-triazole and tetrazole: A comparison of solvation models. *Physical Chemistry Chemical Physics* **2002**, *4* (18), 4314–4318, 10.1039/B203118C. DOI: 10.1039/B203118C.
(133) Kim, S.-I.; Kim, E.-H.; Um, I.-H. Kinetics and Mechanism of Nucleophilic Displacement Reactions of Y-Substituted Phenyl Benzoates with Cyanide Ion. *Bulletin of the Korean Chemical Society* **2010**, *31* (3), 689–693.
(134) Settimo, L.; Bellman, K.; Knegtel, R. M. Comparison of the accuracy of experimental and predicted pKa values of basic and acidic compounds. *Pharmaceutical research* **2014**, *31*, 1082–1095.
(135) Pandit, N. Introduction to the Pharmaceutical Sciences. 2006. Lippincott Williams & Wilkins.
(136) Lide, D. CRC Handbook of Chemistry and Physics 81st Edition, 4-99. CRC Press, New York: 2000.
(137) Kasture, A.; Wadodkar, S. *Pharmaceutical chemistry-I*; Pragati Books Pvt. Ltd., 2015.
(138) Guranda, D. T.; Ushakov, G. A.; Yolkin, P. G.; Švedas, V. K. Thermodynamics of phenylacetamides synthesis: Linear free energy relationship with the pK of amine. *Journal of Molecular Catalysis B: Enzymatic* **2012**, *74* (1-2), 48–53.
(139) Farrow, J.; Ritchie, I.; Mangano, P. The reaction between reduced ilmenite and oxygen in ammonium chloride solutions. *Hydrometallurgy* **1987**, *18* (1), 21–38.
(140) Bulemela, E.; Trevani, L.; Tremaine, P. R. Ionization Constants of Aqueous Glycolic Acid at Temperatures up to 250◦ C Using Hydrothermal pH Indicators and UV-Visible Spectroscopy. *Journal of solution chemistry* **2005**, *34*, 769–788.
(141) Cook, M. J.; Katritzky, A. R.; Linda, P.; Tack, R. D. Aromaticity and tautomerism. Part II. The 4-pyridone, 2-quinolone, and 1-isoquinolone series. *Journal of the Chemical Society, Perkin Transactions 2* **1973**, (7), 1080–1086.
(142) Simmons, J. R.; Murza, A.; Lumsden, M. D.; Kenward, C.; Marsault, É.; Rainey, J. K. Simultaneous ligand and receptor tracking through NMR spectroscopy enabled by distinct 19F labels. *International Journal of Molecular Sciences* **2019**, *20* (15), 3658.
(143) Gocker, S. J. The Essential Physics of Medical Imaging. by Jerrold T. Bushberg, J. Anthony Seibert, Edwin M. Leidholdt, Jr., and John M. Bonne. *Medical Physics* **1995**, *22* (8), 1355.
(144) Garau, A.; Picci, G.; Bencini, A.; Caltagirone, C.; Conti, L.; Lippolis, V.; Paoli, P.; Romano, G. M.; Rossi, P.; Scorciapino, M. A. Glyphosate sensing in aqueous solutions by fluorescent zinc (II) complexes of [9] aneN 3-based receptors. *Dalton Transactions* **2022**, *51* (22), 8733–8742.
(145) Cardellini, F.; Tiecco, M.; Germani, R.; Cardinali, G.; Corte, L.; Roscini, L.; Spreti, N. Novel zwitterionic deep eutectic solvents from trimethylglycine and carboxylic acids: characterization of their properties and their toxicity. *Rsc Advances* **2014**, *4* (99), 55990–56002.
(146) Swift, T.; Swanson, L.; Geoghegan, M.; Rimmer, S. The pH-responsive behaviour of poly (acrylic acid) in aqueous solution is dependent on molar mass. *Soft matter* **2016**, *12* (9), 2542–2549.
(147) Foster, M. P.; McElroy, C. A.; Amero, C. D. Solution NMR of large molecules and assemblies. *Biochemistry* **2007**, *46* (2), 331–340.
(148) Kogon, R.; Faux, D.; Assifaoui, A.; Bodart, P. Advanced insight on the water dynamics of anisotropic hydrogels by field-cycling nuclear magnetic resonance: Application of 3-Tau model. *Carbohydrate polymers* **2023**, *314*, 120922.
(149) Shim, S. Diffusiophoresis, Diffusioosmosis, and Microfluidics: Surface-Flow-Driven Phenomena in the Presence of Flow. *Chemical Reviews* **2022**, *122* (7), 6986–7009. DOI: 10.1021/acs.chemrev.1c00571.
(150) Tokmakov, A. A.; Kurotani, A.; Sato, K. I. Protein pI and Intracellular Localization. *Front Mol Biosci* **2021**, *8*, 775736. DOI: 10.3389/fmolb.2021.775736 From NLM.
(151) Kronis, K. A.; Carver, J. P. Specificity of isolectins of wheat germ agglutinin for sialyloligosaccharides: a 360-MHz proton nuclear magnetic resonance binding study. *Biochemistry* **1982**, *21* (13), 3050–3057. DOI: 10.1021/bi00256a003 From NLM.
(152) Smith, J. J.; Raikhel, N. V. Nucleotide sequences of cDNA clones encoding wheat germ agglutinin isolectins A and D. *Plant Molecular Biology* **1989**, *13* (5), 601–603. DOI: 10.1007/BF00027321.

(153) Michaels, A.; Morelos, O. Polyelectrolyte adsorption by kaolinite. *Industrial & Engineering Chemistry* **1955**, *47* (9), 1801–1809.
(154) Rosés, M.; Bosch, E. Influence of mobile phase acid-base equilibria on the chromatographic behaviour of protolytic compounds. *J Chromatogr A* **2002**, *982* (1), 1–30. DOI: 10.1016/s0021-9673(02)01444-9 From NLM.
(155) Avdeef, A.; Box, K. J.; Comer, J. E. A.; Gilges, M.; Hadley, M.; Hibbert, C.; Patterson, W.; Tam, K. Y. PH-metric logP 11. pKa determination of water-insoluble drugs in organic solvent–water mixtures. *Journal of Pharmaceutical and Biomedical Analysis* **1999**, *20* (4), 631–641. DOI: https://doi.org/10.1016/S0731-7085(98)00235-0.
(156) Mchedlov-Petrossyan, N. O.; Mayorga, R. S. Extraordinary character of the solvent influence on protolytic equilibria: inversion of the fluorescein ionization constants in H2O–DMSO mixtures. *J. Chem. Soc., Faraday Trans.* **1992**, *88* (20), 3025–3032.
(157) Orozco, M.; Luque, F. J. Theoretical Methods for the Description of the Solvent Effect in Biomolecular Systems. *Chemical Reviews* **2000**, *100* (11), 4187–4226. DOI: 10.1021/cr990052a.
(158) Reichardt, C. Solvents and Solvent Effects: An Introduction. *Organic Process Research & Development* **2007**, *11* (1), 105–113. DOI: 10.1021/op0680082.
(159) Hein, M.; Arena, S.; Willard, C. *Foundations of college chemistry*; John Wiley & Sons, 2016.
(160) Lu, J. X.; Tupper, C.; Gutierrez, A. V.; Murray, J. Biochemistry, Dissolution and Solubility. In *StatPearls*, StatPearls Publishing

(161) Kappas, U.; Fischer, W.; Polymeropoulos, E. E.; Brickmann, J. Solvent effects in ionic transport through transmembrane protein channels. *J Theor Biol* **1985**, *112* (3), 459–464. DOI: 10.1016/s0022-5193(85)80014-x From NLM.
(162) Chan, T. C.; Lee, I.; Chan, K. S. Effect of Solvent on Diffusion: Probing with Nonpolar Solutes. *The Journal of Physical Chemistry B* **2014**, *118* (37), 10945–10955. DOI: 10.1021/jp505141z.
(163) Parker, A. J. Protic-dipolar aprotic solvent effects on rates of bimolecular reactions. *Chemical Reviews* **1969**, *69* (1), 1–32. DOI: 10.1021/cr60257a001.
(164) Mota, F. L.; Carneiro, A. P.; Queimada, A. J.; Pinho, S. P.; Macedo, E. A. Temperature and solvent effects in the solubility of some pharmaceutical compounds: Measurements and modeling. *European Journal of Pharmaceutical Sciences* **2009**, *37* (3), 499–507. DOI: https://doi.org/10.1016/j.ejps.2009.04.009.
(165) Abraham, M. H. Solvent effects on reaction rates. *Pure and Applied Chemistry* **1985**, *57* (8), 1055–1064. DOI: doi:10.1351/pac198557081055 (accessed 2025-03-06).
(166) Hemdan, S. S.; Al Gebali, A. M.; Ali, F. K. Investigations of solvent parameters on acidity dissociation constants of bromocresol purple in water-dioxane media. *Journal of Chemical Technology and Metallurgy* **2023**, *58* (4), 690–699.
(167) Lal De, A.; Atta, A. K. Solute–solvent effects in the dissociation of thymolsulphonephthalein (an uncharged acid) in aqueous mixtures of protic (methanol) and dipolar aprotic (dimethylformamide) solvents. *Journal of the Chemical Society, Perkin Transactions 2* **1986**, (9), 1367–1370, 10.1039/P29860001367. DOI: 10.1039/P29860001367.
(168) Gutbezahl, B.; Grunwald, E. The Effect of Solvent on Equilibrium and Rate Constants. II. The Measurement and Correlation of Acid Dissociation Constants of Anilinium and Ammonium Salts in the System Ethanol—Water1,2a. *Journal of the American Chemical Society* **1953**, *75* (3), 559–565. DOI: 10.1021/ja01099a014.
(169) Budiman, A.; Rusdin, A.; Aulifa, D. L. Current Techniques of Water Solubility Improvement for Antioxidant Compounds and Their Correlation with Its Activity: Molecular Pharmaceutics. *Antioxidants (Basel)* **2023**, *12* (2). DOI: 10.3390/antiox12020378 From NLM.
(170) Albishri, A.; Cabot, J. M.; Fuguet, E.; Rosés, M. Determination of the aqueous pKa of very insoluble drugs by capillary electrophoresis: Internal standards for methanol-water extrapolation. *Journal of Chromatography A* **2022**, *1665*, 462795. DOI: https://doi.org/10.1016/j.chroma.2021.462795.

(171) Zayas, H. A.; McCluskey, A.; Bowyer, M. C.; Holdsworth, C. I. Potentiometric determination of acid dissociation constants of novel biaryl monomers. *Analytical Methods* **2015**, *7* (19), 8206–8211, 10.1039/C5AY01673H. DOI: 10.1039/C5AY01673H.
(172) Jouyban, A.; Soltanpour, S.; Chan, H.-K. A simple relationship between dielectric constant of mixed solvents with solvent composition and temperature. *International Journal of Pharmaceutics* **2004**, *269* (2), 353–360. DOI: https://doi.org/10.1016/j.ijpharm.2003.09.010.
(173) Płowaś, I.; Świergiel, J.; Jadżyn, J. Relative Static Permittivity of Dimethyl Sulfoxide + Water Mixtures. *Journal of Chemical & Engineering Data* **2013**, *58* (6), 1741–1746. DOI: 10.1021/je400149j.
(174) Yao, W.; Yu, Z.; Wen, S.; Ni, H.; Ullah, N.; Lan, Y.; Lu, Y. Chiral phosphine-mediated intramolecular [3 + 2] annulation: enhanced enantioselectivity by achiral Brønsted acid. *Chemical Science* **2017**, *8* (7), 5196–5200, 10.1039/C7SC00952F. DOI: 10.1039/C7SC00952F.
(175) Yang, R.; Schulman, S. G. An operational pH in aqueous dimethylsulfoxide based upon the acidity dependence of the rate of a simple ionic recombination reaction in the lowest excited singlet state. *Talanta* **2003**, *60* (2), 535–542. DOI: https://doi.org/10.1016/S0039-9140(03)00112-7.
(176) Mollin, J.; Pavelek, Z.; Navrátilová, J.; Recmanová, A. Effect of medium on dissociation of carboxylic acids. *Collection of Czechoslovak chemical communications* **1985**, *50* (12), 2670–2678.
(177) BANKS, C. V.; ZIMMERMAN, J. The Ka2 Values of Sulfoacetic and 3-Sulfopropionic Acids. *Journal of Organic Chemistry* **1956**, *21* (12), 1439–1440.
(178) Hoefnagel, A.; Wepster, B. Substituent effects. 13. Anomalous dissociation constants in water-organic solvent mixtures: Alicyclic and aliphatic carboxylic acids. *Recueil des Travaux Chimiques des Pays-Bas* **1990**, *109* (9), 455–462.
(179) Bustamante, P.; Peña, M. A.; Barra, J. Partial-solubility parameters of naproxen and sodium diclofenac. *J Pharm Pharmacol* **1998**, *50* (9), 975–982. DOI: 10.1111/j.2042-7158.1998.tb06911.x From NLM.
(180) O'Brien, M.; McCauley, J.; Cohen, E. Indomethacin. In *Analytical Profiles of Drug Substances*, Florey, K. Ed.; Vol. 13; Academic Press, 1984; pp 211–238.
(181) Winiwarter, S.; Bonham, N. M.; Ax, F.; Hallberg, A.; Lennernäs, H.; Karlén, A. Correlation of human jejunal permeability (in vivo) of drugs with experimentally and theoretically derived parameters. A multivariate data analysis approach. *J Med Chem* **1998**, *41* (25), 4939–4949. DOI: 10.1021/jm9810102 From NLM.
(182) Deweese, J. E.; Osheroff, N. The use of divalent metal ions by type II topoisomerases. *Metallomics* **2010**, *2* (7), 450–459. DOI: 10.1039/c003759a (accessed 5/20/2025).
(183) Son, H.; Park, J.; Choi, Y. H.; Jung, Y.; Lee, J.-W.; Bae, S.; Lee, S. Exploring the dynamic nature of divalent metal ions involved in DNA cleavage by CRISPR–Cas12a. *Chemical Communications* **2022**, *58* (12), 1978–1981, 10.1039/D1CC04446J. DOI: 10.1039/D1CC04446J.
(184) Gameel, A.; Farag, H. A.; El-Gayar, D.; El-Sayed, E. M. Effective removal of calcium and magnesium Ions from boiler feed water using Polyacrylonitrile supported Titanium Tungstovanadate composite ion exchange nanofiber. *GLOBAL NEST JOURNAL* **2024**, *26* (10).
(185) Holm, R. H.; Kennepohl, P.; Solomon, E. I. Structural and Functional Aspects of Metal Sites in Biology. *Chemical Reviews* **1996**, *96* (7), 2239–2314. DOI: 10.1021/cr9500390.
(186) Misra, V. K.; Hecht, J. L.; Yang, A.-S.; Honig, B. Electrostatic Contributions to the Binding Free Energy of the λcI Repressor to DNA. *Biophysical Journal* **1998**, *75* (5), 2262–2273. DOI: https://doi.org/10.1016/S0006-3495(98)77671-4.
(187) Cowan, J. A. Structural and catalytic chemistry of magnesium-dependent enzymes. *Biometals* **2002**, *15* (3), 225–235. DOI: 10.1023/a:1016022730880 From NLM.
(188) Clapham, D. E. Calcium signaling. *Cell* **2007**, *131* (6), 1047–1058. DOI: 10.1016/j.cell.2007.11.028 From NLM.
(189) Berridge, M. J.; Bootman, M. D.; Roderick, H. L. Calcium signalling: dynamics, homeostasis and remodelling. *Nat Rev Mol Cell Biol* **2003**, *4* (7), 517–529. DOI: 10.1038/nrm1155 From NLM.
(190) Romani, A. M. Cellular magnesium homeostasis. *Arch Biochem Biophys* **2011**, *512* (1), 1–23. DOI: 10.1016/j.abb.2011.05.010 From NLM.

(191) Dontsova, K.; Darrell, N. Effects of exchangeable Ca: Mg ratio on soil clay flocculation, infiltration and erosion. **2001**, 580–585.
(192) Qadir, M.; Schubert, S.; Oster, J. D.; Sposito, G.; Minhas, P. S.; Cheraghi, S. A. M.; Murtaza, G.; Mirzabaev, A.; Saqib, M. High-magnesium waters and soils: Emerging environmental and food security constraints. *Sci Total Environ* **2018**, *642*, 1108–1117. DOI: 10.1016/j.scitotenv.2018.06.090 From NLM.
(193) Breck, W. Water chemistry (Snoeyink, VL; Jenkins, D.). ACS Publications: 1981.
(194) Wilkinson, K.; Lead, J. R. *Environmental Colloids and Particles: Behaviour, Separation and Characterisation*; 2007. DOI: 10.1002/9780470024539.
(195) Lee, K. Y.; Mooney, D. J. Alginate: properties and biomedical applications. *Prog Polym Sci* **2012**, *37* (1), 106–126. DOI: 10.1016/j.progpolymsci.2011.06.003 From NLM.
(196) Augst, A. D.; Kong, H. J.; Mooney, D. J. Alginate Hydrogels as Biomaterials. *Macromolecular Bioscience* **2006**, *6* (8), 623–633. DOI: https://doi.org/10.1002/mabi.200600069.
(197) Grant, G. T.; Morris, E. R.; Rees, D. A.; Smith, P. J. C.; Thom, D. Biological interactions between polysaccharides and divalent cations: The egg-box model. *FEBS Letters* **1973**, *32* (1), 195–198. DOI: https://doi.org/10.1016/0014-5793(73)80770-7.
(198) Topuz, F.; Henke, A.; Richtering, W.; Groll, J. Magnesium ions and alginate do form hydrogels: a rheological study. *Soft Matter* **2012**, *8* (18), 4877–4881, 10.1039/C2SM07465F. DOI: 10.1039/C2SM07465F.
(199) Weaver, C.; Heaney, R. *Calcium in Human Health*; 2006. DOI: 10.1007/978-1-59259-961-5.
(200) Bourassa, M. W.; Abrams, S. A.; Belizán, J. M.; Boy, E.; Cormick, G.; Quijano, C. D.; Gibson, S.; Gomes, F.; Hofmeyr, G. J.; Humphrey, J.; et al. Interventions to improve calcium intake through foods in populations with low intake. *Ann N Y Acad Sci* **2022**, *1511* (1), 40–58. DOI: 10.1111/nyas.14743 From NLM.
(201) Kesteloot, H.; Joossens, J. V. The relationship between dietary intake and urinary excretion of sodium, potassium, calcium and magnesium: Belgian Interuniversity Research on Nutrition and Health. *J Hum Hypertens* **1990**, *4* (5), 527–533. From NLM.
(202) Martinez, M. E.; Salinas, M.; Miguel, J. L.; Herrero, E.; Gomez, P.; Garcia, J.; Sanchez-Sicilia, L.; Montero, A. Magnesium excretion in idiopathic hypercalciuria. *Nephron* **1985**, *40* (4), 446–450. DOI: 10.1159/000183517 From NLM.
(203) Holmes, R. P.; Kennedy, M. Estimation of the oxalate content of foods and daily oxalate intake. *Kidney Int* **2000**, *57* (4), 1662–1667. DOI: 10.1046/j.1523-1755.2000.00010.x From NLM.
(204) Curhan, G. C. Epidemiology of stone disease. *Urol Clin North Am* **2007**, *34* (3), 287–293. DOI: 10.1016/j.ucl.2007.04.003 From NLM.
(205) Garnier, C.; Axelos, M. A.; Thibault, J.-F. Selectivity and cooperativity in the binding of calcium ions by pectins. *Carbohydrate Research* **1994**, *256* (1), 71–81.
(206) Sinn, C. G.; Dimova, R.; Antonietti, M. Isothermal titration calorimetry of the polyelectrolyte/water interaction and binding of Ca2+: effects determining the quality of polymeric scale inhibitors. *Macromolecules* **2004**, *37* (9), 3444–3450.
(207) Warttinger, U.; Krämer, R. Quantification of sulfated polysaccharides in urine by the Heparin Red mix-and-read fluorescence assay. *arXiv preprint arXiv:1811.09115* **2018**.
(208) Li, H.; Chen, J.; Huang, B.; Kong, L.; Sun, F.; Li, L.; Peng, C.; Cai, H.; Hou, R. A Rapid Fluorescence Sensor for the Direct Quantification of Rongalite in Foodstuffs. *Foods* **2022**, *11* (17). DOI: 10.3390/foods11172650 From NLM.
(209) Wallace, M.; Hicks, T.; Khimyak, Y. Z.; Angulo, J. Self-Correcting Method for the Measurement of Free Calcium and Magnesium Concentrations by 1H NMR. *Analytical Chemistry* **2019**, *91* (22), 14442–14450, Article. DOI: 10.1021/acs.analchem.9b03008 Scopus.
(210) Nancollas, G. 147. Thermodynamics of ion association. Part II. Alkaline-earth acetates and formates. *Journal of the Chemical Society (Resumed)* **1956**, 744–749.
(211) Bers, D. M.; Patton, C. W.; Nuccitelli, R. A practical guide to the preparation of Ca2+ buffers. In *Methods in cell biology*, Vol. 99; Elsevier, 2010; pp 1–26.

(212) Sun, M. S.; Harriss, D. K.; Magnuson, V. R. Activity corrections for ionic equilibria in aqueous solutions. *Canadian Journal of Chemistry* **1980**, *58* (12), 1253–1257.
(213) Lide, D. R. *CRC handbook of chemistry and physics: a ready-reference book of chemical and physical data*; CRC press, 1995.
(214) Smith, D. W. Ionic hydration enthalpies. *Journal of Chemical Education* **1977**, *54* (9), 540.
(215) Cao, L.; Lu, W.; Mata, A.; Nishinari, K.; Fang, Y. Egg-box model-based gelation of alginate and pectin: A review. *Carbohydrate Polymers* **2020**, *242*, 116389. DOI: https://doi.org/10.1016/j.carbpol.2020.116389.
(216) Donati, I.; Asaro, F.; Paoletti, S. Experimental evidence of counterion affinity in alginates: the case of nongelling ion Mg2+. *The Journal of Physical Chemistry B* **2009**, *113* (39), 12877–12886.
(217) Ruiz-Caldas, M.-X.; Apostolopoulou-Kalkavoura, V.; Hellström, A.-K.; Hildenbrand, J.; Larsson, M.; Jaworski, A.; Samec, J. S. M.; Lahtinen, P.; Tammelin, T.; Mathew, A. P. Citrated cellulose nanocrystals from post-consumer cotton textiles. *Journal of Materials Chemistry A* **2023**, *11* (13), 6854–6868, 10.1039/D2TA09456H. DOI: 10.1039/D2TA09456H.
(218) Lam, E.; Hemraz, U. D. Preparation and Surface Functionalization of Carboxylated Cellulose Nanocrystals. *Nanomaterials (Basel)* **2021**, *11* (7). DOI: 10.3390/nano11071641 From NLM.
(219) Benselfelt, T.; Kummer, N.; Nordenström, M.; Fall, A. B.; Nyström, G.; Wågberg, L. The Colloidal Properties of Nanocellulose. *ChemSusChem* **2023**, *16* (8), e202201955. DOI: https://doi.org/10.1002/cssc.202201955.
(220) Stephen, A. M.; Phillips, G. O. *Food polysaccharides and their applications*; CRC press, 2016.
(221) Rahman, M. S.; Hasan, M. S.; Nitai, A. S.; Nam, S.; Karmakar, A. K.; Ahsan, M. S.; Shiddiky, M. J. A.; Ahmed, M. B. Recent Developments of Carboxymethyl Cellulose. *Polymers (Basel)* **2021**, *13* (8). DOI: 10.3390/polym13081345 From NLM.
(222) Pensini, E.; Yip, C.; O'Carroll, D.; Sleep, B. Carboxymethyl cellulose binding to mineral substrates: Characterization by atomic force microscopy-based Force spectroscopy and quartz-crystal microbalance with dissipation monitoring. *Journal of colloid and interface science* **2013**, *402*. DOI: 10.1016/j.jcis.2013.03.053.
(223) Arumughan, V.; Özeren, H. D.; Hedenqvist, M.; Skepö, M.; Nypelö, T.; Hasani, M.; Larsson, A. Anion-Specific Adsorption of Carboxymethyl Cellulose on Cellulose. *Langmuir* **2023**, *39* (42), 15014–15021. DOI: 10.1021/acs.langmuir.3c01924.
(224) Fitzpatrick, L. R.; Woldemariam, T. 5.16 - Small-Molecule Drugs for the Treatment of Inflammatory Bowel Disease. In *Comprehensive Medicinal Chemistry III*, Chackalamannil, S., Rotella, D., Ward, S. E. Eds.; Elsevier, 2017; pp 495–510.
(225) Kahkeshani, N.; Farzaei, F.; Fotouhi, M.; Alavi, S. S.; Bahramsoltani, R.; Naseri, R.; Momtaz, S.; Abbasabadi, Z.; Rahimi, R.; Farzaei, M. H.; et al. Pharmacological effects of gallic acid in health and diseases: A mechanistic review. *Iran J Basic Med Sci* **2019**, *22* (3), 225–237. DOI: 10.22038/ijbms.2019.32806.7897 From NLM.
(226) Dean, P. The Oxalate Dianion, C2O42-, Planar or Nonplanar? *Journal of chemical education* **2012**, *89*, 417–418. DOI: 10.1021/ed200202r.
(227) Zayed, A.; Adly, G. M.; Farag, M. A. Management Strategies for the Anti-nutrient Oxalic Acid in Foods: A Comprehensive Overview of Its Dietary Sources, Roles, Metabolism, and Processing. *Food and Bioprocess Technology* **2025**, *18* (5), 4280–4300. DOI: 10.1007/s11947-024-03726-0.
(228) Salgado, N.; Silva, M. A.; Figueira, M. E.; Costa, H. S.; Albuquerque, T. G. Oxalate in Foods: Extraction Conditions, Analytical Methods, Occurrence, and Health Implications. *Foods* **2023**, *12* (17). DOI: 10.3390/foods12173201 From NLM.
(229) Kikunaga, S.; ISHII, H.; IMADA, S.; Takahashi, M. Correlation between the bioavailability of magnesium, other minerals and oxalic acid in spinach. *Journal of Home Economics of Japan* **1995**, *46* (1), 3–9.
(230) University of Rhode Island, D. o. C. *Solubility Product Constants near 25 °C*. 2025. https://www.chm.uri.edu/weuler/chm112/refmater/KspTable.html (accessed).

(231) Massey, L. Magnesium therapy for nephrolithiasis. *Magnes Res* **2005**, *18* (2), 123–126. From NLM.
(232) Ahvazi, B.; Boeshans, K. M.; Idler, W.; Baxa, U.; Steinert, P. M. Roles of Calcium Ions in the Activation and Activity of the Transglutaminase 3 Enzyme *. *Journal of Biological Chemistry* **2003**, *278* (26), 23834–23841. DOI: 10.1074/jbc.M301162200 (accessed 2025/06/12).
(233) Hasan, M. U.; Singh, Z.; Shah, H. M. S.; Kaur, J.; Woodward, A.; Afrifa-Yamoah, E.; Malik, A. U. Oxalic acid: A blooming organic acid for postharvest quality preservation of fresh fruit and vegetables. *Postharvest Biology and Technology* **2023**, *206*, 112574. DOI: https://doi.org/10.1016/j.postharvbio.2023.112574.
(234) Szakács, Z.; Hägele, G.; Tyka, R. 1H/31P NMR pH indicator series to eliminate the glass electrode in NMR spectroscopic pKa determinations. *Analytica chimica acta* **2004**, *522* (2), 247–258.
(235) DeMaria, P.; Fini, A.; Guarnieri, A.; Varoli, L. Thermodynamic dissociation constants of 3-substituted 3-(4 biphenylyl)-3-hydroxypropionic acids in aqueous DMSO. *Arch Pharm (Weinheim)* **1983**, *316*, 559–563.
(236) Eliason, R.; Kreevoy, M. M. Comparison of general-acid-catalyzed ethyl vinyl ether hydrolysis in 80% dimethyl sulfoxide with that in water. *The Journal of Physical Chemistry* **1974**, *78* (26), 2658–2659.
(237) Baughman, E. H.; Kreevoy, M. M. Determination of acidity in 80% dimethyl sulfoxide-20% water. *The Journal of Physical Chemistry* **1974**, *78* (4), 421–423.
(238) Ledenkov, S. F.; Sharnin, V. A.; Isaeva, V. A. *Russian Journal of Physical Chemistry* **1995**, *69* (6), 900–901.
(239) Daniele, P. G.; Foti, C.; Gianguzza, A.; Prenesti, E.; Sammartano, S. Weak alkali and alkaline earth metal complexes of low molecular weight ligands in aqueous solution. *Coordination Chemistry Reviews* **2008**, *252* (10-11), 1093–1107, Review. DOI: 10.1016/j.ccr.2007.08.005 Scopus.