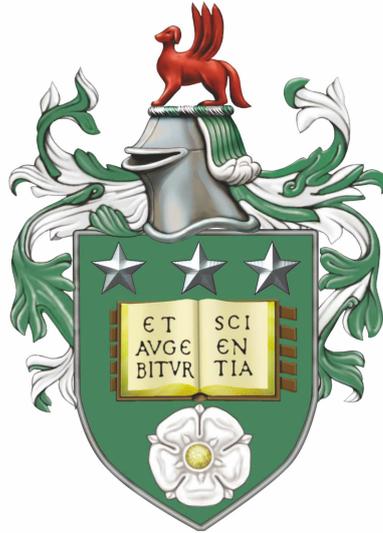


The Biological Role of Water in Extreme Conditions



Harrison Samuel Tarrent Laurent

School of Physics and Astronomy
The University of Leeds

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

October 2021

Declaration

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. Chapter 3 is based on neutron scattering data originally taken and corrected by Dr Alan Soper and Dr Kristian Weckström. Chapter 4 is based on experiments that were devised by Professor Lorna Dougan and the candidate, and completed by the candidate with the assistance of Dr Matthew Hughes, Dr Tristan Youngs, and Dr Tom Headen. Chapter 5 is based on experiments that were devised by Professor Lorna Dougan and the candidate, and completed by the candidate with the assistance of Dr Matthew Hughes, Dr Tristan Youngs, and Dr Tom Headen. Chapter 6 contains neutron scattering experiments that were devised by Professor Lorna Dougan and the candidate, and completed by the candidate with the assistance of Dr Matthew Hughes, Dr Tristan Youngs, and Dr Tom Headen. This section also contains studies on a model protein devised and completed by the candidate, synthesised with the assistance of Sophie Cussons, and studies on a model peptide devised by the candidate and synthesised with the assistance of Dr Martin Walko. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

©2021 The University of Leeds and Harrison Samuel Tarrent Laurent.

The right of Harrison Samuel Tarrent Laurent to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

For my Grandfather Henry Raymond Smith, who inspired my love of physics, continues to support me and my new family, and without whom I certainly wouldn't be in the wonderful place I am today.

*“If life can exist under such extreme conditions down here, then surely it could exist
somewhere out there”*

-David Attenborough, Blue Planet II

Acknowledgements

At the tender age of three I proudly stood up in front of my local playgroup and declared “my name is Harry and I am a genius”. I then spent the rest of my life trying to validate this hypothesis. At the end of my PhD I can confirm that with the exception of a few scientific anomalies, such as my attempting to steal a giant stuffed piglet from my local cinema when I was 15 and both planning a wedding and impregnating my new wife during the writing of this thesis, that my initial thoughts were essentially accurate, if not to the correct degree of magnitude.

But now that the modesty that was instilled into me by my parents is out of the way, I can begin to thank the people who were *actually* responsible for the success of my PhD and the results included in this thesis. Firstly I cannot thank my primary supervisor, Professor Lorna Dougan, enough. She obtained the joint University of Leeds and ISIS facility development studentship that funded my PhD. While I imagine I was most likely hired because during my interview for the position I managed to sooth her infant daughter to sleep, I can only hope that the remainder of my work for her has been as useful. She has provided an ever-present guidance to every aspect of the work I have managed to complete during my PhD, and has managed to curb my enthusiasm for my studies in the most useful direction. I consider her both a mentor and a role-model, and I am overwhelmingly grateful for the foundation to my scientific career she has helped me to build, and is continuing to help me to build.

I would also like to thank my other supervisor, Dr Alan Soper. Not only has he provided wisdom and insight into my research that can only come with being a giant in the field, he developed the primary tool for my PhD studies, EPSR, and helped me to develop my own analysis routines which run off its back. I am the last of Alan’s official PhD students, and while academia is not something one ever truly retires from, hence I am sure I will continue to run to him with questions, I hope I helped him to end his astounding official academic career on a good note. I must also thank the instrument scientists who helped me obtain the raw neutron scattering data and perform the post-acquisition corrections, Dr Tristan Youngs and Dr Tom Headen. Their oversight and

guidance during the 24/7 lifestyle that is ISIS beamtime experiments was invaluable, and without them none of my core work would have been possible. I would also like to thank Dr Matthew Hughes and Dr Benjamin Hanson in the Dougan research group for their help at beamtimes, useful scientific discussions throughout my PhD, and helping me practice for my viva. I have been very fortunate to have wonderful colleagues in the Dougan research group, and I look forward to our future work together.

The multidisciplinary nature of the study of extremophiles means that I have had to drop in and out of the biology and chemistry departments to allow me to perform my research. With this in mind it is vital that I thank Nasir Khan, Professor David Brockwell, Sophie Cussons and Dr Martin Walko in the Astbury Centre for Structural Molecular Biology and School of Chemistry for their help and patience in guiding me through concepts and experimental techniques that were outside anything I had come close to studying in my undergraduate physics degree. I must also thank Dr Michael Ries and Dr Daniel Baker from the soft matter physics department for their help with the NMR experiments and input into the resulting papers. The willingness of all these people to go the extra mile to help someone from outside their immediate department is an inspiration and demonstrates, in my opinion, how interdisciplinary research should be done. I am truly grateful.

Penultimately I wish to thank my family. My mother has had to put up with my ...enthusiasm for my passions for my entire life, and without her guiding hand and seemingly unlimited patience I certainly would not have ended up where I am today. My father has been a constant example of what one can achieve from the ground up with hard work and dedication. He has always supported me and my brothers in our interests, regardless of his own interest in the subject, and the sacrifices he has made by regular deployments to Afghanistan have meant that I have been able to receive the education and support I have needed. Both my parents have set an incredibly high standard, and now that I am about to become a father myself, I only hope that I can repay what they have done for me by raising my own children to love them as much as I do.

Finally, and perhaps most importantly, I need to thank my wonderful wife, Mollie Ducharme. I was lucky enough to meet the love of my life at the start of my 3rd year of undergraduate. Since then we have hardly spent a night apart and her wonderful gifts of patience, empathy, and kindness have supported me in every aspect of my life. It is no exaggeration to say that without her support I may not have finished my PhD, and it certainly would not have been to the same standard. We have built our lives together, and with the ending of my PhD it is the final piece of the foundation we needed to start our family together. I can never repay her for what she has already given to me, but it won't stop me trying for the rest of our lives.

Also I should probably thank our dog Hartley for getting me out of the house for some fresh air during the writing of this thesis to keep me sane. Woof woof I suppose.

Abstract

Despite exposure to extreme conditions, extremophilic organisms have developed a range of mechanisms to survive these detrimental perturbations to their solvent environment. In this thesis we study how extremes of pressure and salinity affect the aqueous environment with which extremophiles interact. To do this we employ a combination of neutron scattering with computational modelling to examine the perturbations to water structure, and nuclear magnetic resonance to examine the perturbations to water dynamics. We first examine perturbations to water structure and dynamics by simple monovalent model potassium halide salts and find that we can extract atom scale information which we can link to bulk measurable thermodynamic properties. We then use neutron diffraction to study the organic osmolyte TMAO in solution, which is used by high pressure adapted organisms to protect their biochemistry and find that it can preserve the structure of water against pressure induced perturbations. We then study aqueous magnesium perchlorate, a salt which is found in high concentrations in the Martian regolith, and likely in the surface lakes, and show that it induces a pressure-like effect on water structure. We then show that TMAO can also resist this pressure-like structural perturbation in the same way it can resist pressure induced structural perturbations. Finally, we investigate how magnesium perchlorate hinders biomolecular self-assembly by studying the amino acid glycine and the stability of the model protein I27. These fundamental insights help us to understand how observed adaptations in extremophilic organisms help them survive, with implications in medical and industrial settings.

Contents

Abstract	vii
Contents	viii
Publications	xii
List of Abbreviations	xiv
List of Tables	xviii
List of Figures	xx
1 Introduction	1
1.1 Protein Structure and Folding	2
1.1.1 The Covalent and Non-covalent Interactions that Govern Protein Structure	3
1.1.2 Free Energy Changes in Biomolecular Self-Assembly - Protein Folding	8
1.2 Adaptations to Extreme Environments	12
1.2.1 Adapting to a Cold Aqueous Environment - Pyschrophiles	13
1.2.2 Adapting to a Saline Aqueous Environment - Halophiles	14
1.2.3 Adapting to a High Pressure Aqueous Environment - Piezophiles .	17
1.3 The Liquid State	19
1.3.1 Monitoring Perturbations to Water Structure	19
1.3.1.1 The Water Molecule	20
1.3.1.2 Tetrahedrally Structured Liquid Water	21
1.3.1.3 The Two-State vs The Continuum Model of Liquid Water	23

1.3.1.4	Pressure and Temperature Variation	26
1.3.1.5	Salt Addition	29
1.3.1.6	Addition of Other Solute Species - Alcohols, Amino Acids, and Osmolytes	31
1.3.2	Monitoring Perturbations to Water Dynamics	35
1.3.2.1	The Extended Jump Model	35
1.3.2.2	Pressure and Temperature Variation	36
1.3.2.3	Salt Addition	37
1.3.2.4	Addition of Other Solute Species - Alcohols, Amino Acids, and Osmolytes	40
1.4	Aims and Objectives	42
2	Materials and Methods	44
2.1	Diffraction	45
2.1.1	Superposition	46
2.1.2	Reciprocal Space	49
2.1.3	Neutron Diffraction	52
2.1.4	The Total Structure Factor	54
2.1.5	The Neutron Scattering Experiment	62
2.1.6	Data Correction - Gudrun	68
2.2	Empirical Potential Structure Refinement	71
2.2.1	Building the Simulation	73
2.2.2	Monte Carlo	74
2.2.3	Reference Potential	74
2.2.4	Empirical Potential	77
2.2.5	Auxiliary Routines	81
2.2.6	EPSR Summary	85
2.3	Custom Analysis Routine	86
2.3.1	Reading in .ato File	86
2.3.2	Hydrogen Bonding	89
2.3.3	Interaction Strength Calculation	92

2.3.4	Bond and Dipole Angles	93
2.4	Nuclear Magnetic Resonance	95
2.4.1	Spin	95
2.4.2	Resonance Frequency	97
2.4.3	Chemical Shifts	100
2.4.4	Spin-Lattice Relaxation	103
2.4.5	Diffusion Measurements	107
2.5	Proteins	110
2.5.1	Media and Buffers	111
2.5.2	Transformation of Cells	112
2.5.3	Full-scale Expression	112
2.5.4	Cell Lysis	113
2.5.5	Purification	114
2.5.6	His Cleavage	116
2.6	Peptides	117
2.6.1	Peptide Synthesis	117
2.7	Fluorescence Spectroscopy of Proteins	118
2.8	Circular Dichroism	122
3	Perturbations to Water Structure, Dynamics, and Thermodynamics due to Simple Monovalent Ions	125
3.1	Introduction	125
3.2	Structure of Water - RDFs	127
3.3	Structure of Water - Dipole Angle Distributions	132
3.4	Thermodynamics of Hydration	135
3.5	Dynamic Perturbations to Water - Nuclear Magnetic Resonance	141
3.6	Conclusions	149
4	Perturbations to Water Structure due to Hydrostatic Pressure in the Presence of Trimethylamine N-Oxide	153
4.1	Introduction	153
4.2	Experimental Considerations	156

4.3	Structure of Water - EPSR	157
4.4	Structure of Water - Extended Analysis	163
4.5	TMAO Hydration - Hydrophilic	167
4.6	TMAO Hydration - Hydrophobic	172
4.7	Conclusions	175
5	Dissecting the Effects of Different Solutes in Tertiary Solution: Trimethylamine N-Oxide vs Mg(ClO₄)₂	177
5.1	Introduction	177
5.2	Experimental Considerations	179
5.3	Structure of Water - EPSR	184
5.4	Solute-Solvent and Solute-Solute Interactions - EPSR	188
5.5	Structure of Water - Extended Analysis	192
5.6	Nuclear Magnetic Resonance	195
5.7	Dynamic Fingerprint in EPSR - The Bifurcated Hydrogen	200
5.8	Quantification of Relative Perturbing Ability of Solutes	203
5.9	Conclusions	205
6	The Study of Biomolecules in Solution Under Extreme Conditions	209
6.1	Introduction	209
6.2	Experimental Considerations - Aqueous Glycine and Mg(ClO ₄) ₂	211
6.3	Water Structure in Aqueous Glycine and Mg(ClO ₄) ₂	213
6.4	Hydration of Glycine in Aqueous Mg(ClO ₄) ₂	214
6.5	Glycine - Mg(ClO ₄) ₂ Interactions	223
6.6	Glycine Self-Association in Aqueous Mg(ClO ₄) ₂	226
6.7	The Model Protein I27 in Aqueous Mg(ClO ₄) ₂ and TMAO	229
6.8	Neutron Scattering on Aqueous Peptides	237
6.9	Conclusions	240
7	Conclusions and Future Work	242
7.1	Concluding Remarks	242
7.2	Future Work	252

7.2.1	High Pressure NMR	253
7.2.2	A Model Biomolecule for Neutron Scattering: The β -Hairpin Protein CLN025	254
7.2.3	The Next Generation of Structural Refinement Software: Dissolve	256
7.2.4	Analysis Routine	257
7.3	Final Thoughts	258
A	Supplementary info for chapter 3	260
B	Supplementary info for chapter 4	268
C	Supplementary info for chapter 5	270
D	Supplementary info for chapter 6	272

Publications

[1] Laurent, H.; Soper, A.; Dougan, L. Biomolecular Self-Assembly under Extreme Martian Mimetic Conditions. *Mol. Phys.* 2019, 117 (22), 3398–3407. <https://doi.org/10.1080/00268976.2019.1649485>.

[2] Laurent, H.; Soper, A. K.; Dougan, L. Trimethylamine N-Oxide (TMAO) Resists the Compression of Water Structure by Magnesium Perchlorate: Terrestrial Kosmotrope vs. Martian Chaotrope. *Phys. Chem. Chem. Phys.* 2020, 22 (9), 4924–4937. <https://doi.org/10.1039/c9cp06324b>.

[3] Laurent, H.; Baker, D.; Soper, A.; Ries, M.; Dougan, L. Solute Specific Perturbations to Water Structure and Dynamics in Tertiary Aqueous Solution. *J. Phys. Chem. B* 2020, 124, 10983–10993. <https://doi.org/10.1021/acs.jpcc.0c07780>.

[4] Hanson, B. S.; Brown, C. P.; Laurent, H.; Hughes, M. D. G.; Dougan, L. Hierarchical Biomechanics: Student Engagement Activities with a Focus on Biological Physics. *Phys. Educ.* 2020, 55 (2). <https://doi.org/10.1088/1361-6552/ab687e>.

[5] Laurent, H.; Baker, D.; Soper, A. K.; Ries, M. E.; Dougan, L. Bridging Structure, Dynamics, and Thermodynamics: An Example Study on Aqueous Potassium Halides. *J. Phys. Chem. B* 2021. (Accepted Manuscript)

List of abbreviations

Experimental Technique

EPSR	Empirical potential structure refinement
NMR	Nuclear magnetic resonance
CD	Circular dichroism
2D IR	2 dimensional infrared spectroscopy
MD	molecular dynamics
XRD	x ray diffraction

Model/Properties

$F(Q)$	Total structure factor
$S(Q)$	Partial structure factor
b	Coherent neutron scattering length
RDF or $g(r)$	Radial distribution function
SDF	Spatial density function
ΔG	Change in Gibbs free energy
ΔH	Change in enthalpy
S	Entropy
ΔS	Change in entropy
P	Pressure
V	Volume

T	Temperature
Ω	Number of accessible states
ΔV	Change in volume
ΔE	Change in internal energy
ΔG_{un}	Difference in Gibbs free energy between the folded and unfolded state of a protein
ΔG_{TS}	Difference in Gibbs free energy between the folded and transition state of a protein
k_u	Rate of unfolding
k_f	Rate of folding
λ	wavelength of radiation
Q	The momentum transfer between scattering particle and the scattering object
LJ	Lennard - Jones
ε	The potential well depth in the Lennard - Jones potential
σ	The distance at which the Lennard - Jones potential becomes positive
q	charge
T_1	Spin-lattice relaxation time
D	Self-diffusion coefficient

Quantities

μg	microgram
mg	milligram
g	gram
kg	kilogram
ml	millilitre
μl	microlitre

l	litre
M	Molar
M_r	Molecular Weight
kbar	kilobar
Pa	pascal
kPa	kilopascal
GPa	gigapascal
Å	Ångstroms
nm	nanometre
μm	micrometre
mm	millimetre
cm	centimetre
m	metre
s	seconds
min	minutes
K	Kelvin
°C	degrees Celsius
°	degrees

Databases

PDB	Protein data bank
-----	-------------------

Biological molecules

I27	The 27th immunoglobulin domain of human cardiac Titin
MBP	Maltose binding protein
DNA	Deoxyribonucleic acid

Important Chemicals

H ₂ O	water (¹ H only)
D ₂ O	water (² H only)
HDO	water (equimolar mixture of H ₂ O and D ₂ O)
KF	potassium fluoride
KCl	potassium chloride
KBr	potassium bromide
KI	potassium iodide
Mg(ClO ₄) ₂	magnesium perchlorate
TMAO	trimethylamine N-oxide
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
NaOH	sodium hydroxide

Constants

k_b	Boltzmann's constant
h	Plank's constant

List of Tables

1.3.1	Properties of H ₂ O and H ₂ S	21
2.1.1	Key properties for selected subatomic particles	53
2.1.2	Elemental neutron scattering cross sections	59
2.4.1	Gyromagnetic ratio for selected elements	97
2.5.1	Protein buffers	111
3.2.1	Peak positions of O _w - O _w RDFs for aqueous potassium halides	130
3.4.1	Thermodynamic parameters of solvation of aqueous potassium halides	137
3.4.2	Calculated distance at which $\Delta H_{solv} = 0$ for aqueous potassium halides	141
3.5.1	Parameters of ion size, charge, and hydration for aqueous potassium halides	146
3.5.2	NMR B parameters from T_1^{-1} data for aqueous potassium halides . . .	147
3.5.3	NMR B parameters from D data on aqueous potassium halides	149
4.3.1	Densities of pure water and aqueous TMAO at 25 bar and 4 kbar . . .	158
4.3.2	Peak positions of O _w - O _w RDFs for water and aqueous TMAO at 25 bar and 4 kbar	160
4.3.3	O _w - O _w coordination numbers for water and aqueous TMAO at 25 bar and 4 kbar	162
4.4.1	Peak positions in bulk water - water dipole angle distributions for water and aqueous TMAO at 25 bar and 4 kbar	165
4.4.2	Average bulk water - water hydrogen bond interaction energy for water and aqueous TMAO at 25 bar and 4 kbar	166

4.5.1	O_T - O_w coordination numbers for aqueous TMAO at 25 bar and 4 kbar	169
4.5.2	Peak positions in O_T - water dipole angle distributions for aqueous TMAO at 25 bar and 4 kbar	169
4.5.3	Average O_T - water hydrogen bond interaction energy for aqueous TMAO at 25 bar and 4 kbar	171
4.6.1	C_T - O_w coordination numbers for aqueous TMAO at 25 bar and 4 kbar	173
5.2.1	Densities of solutions for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	184
5.3.1	EPSR box sizes for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	185
5.3.2	Peak positions in O_w - O_w RDFs for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	187
5.4.1	Peak positions in O_T - O_w and C_T - O_w RDFs for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	189
5.4.2	Peak positions in Mg - O_w and Cl_p - O_w RDFs for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	190
5.8.1	Calculated g values for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	205
6.3.1	Peak positions in O_w - O_w RDFs for aqueous $Mg(ClO_4)_2$ and glycine .	214
6.4.1	Peak positions in the N_a - O_w and N_a - H_w RDFs for aqueous $Mg(ClO_4)_2$ and glycine	218
6.4.2	Peak positions in C_c - O_w and C_c - H_w RDFs for aqueous $Mg(ClO_4)_2$ and glycine	220
6.4.3	Peak positions in C_{bk} - O_w and C_{bk} - H_w RDFs for aqueous $Mg(ClO_4)_2$ and glycine	222
6.5.1	Peak positions in N_a - Mg and N_a - Cl_p RDFs for aqueous $Mg(ClO_4)_2$ and glycine	224

A.0.1	Lennard Jones and q parameters for aqueous potassium halides	260
A.0.2	EPSR Box sizes for aqueous potassium halides	261
B.0.1	Lennard Jones and q parameters for aqueous TMAO	268
C.0.1	Lennard Jones and q parameters for aqueous TMAO and $\text{Mg}(\text{ClO}_4)_2$.	270
D.0.1	Lennard Jones and q parameters for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine .	272

List of Figures

1.1.1	Basic structure of an amino acid	3
1.1.2	Structures of 20 naturally occurring amino acids	4
1.1.3	Structural hierarchy in proteins	5
1.1.4	The hydrophobic effect	8
1.1.5	2 state protein folding energy landscape	11
1.1.6	Funnel shaped protein folding energy landscape	12
1.3.1	Structure of a water molecule	20
1.3.2	Tetrahedrally structured water	22
1.3.3	$O_w - O_w$ RDF for pure ambient water	23
1.3.4	Phase diagram of water	26
1.3.5	OWater - water RDFs with increasing pressure by XRD	28
1.3.6	Spatial density functions of water at pressure and with $\text{Mg}(\text{ClO}_4)_2$. .	30
1.3.7	$O_w - O_w$ RDFs for pure water, aqueous urea, and aqueous urea and TMAO	32
1.3.8	Segregation in aqueous methanol	33
1.3.9	Structure of TMAO	34
1.3.10	Extended jump model	36

1.3.11	Self-diffusion coefficient and viscosity of water with increasing temperature and pressure	37
1.3.12	Contributions to jump and frame reorientation for monovalent salts	39
1.3.13	Rotational correlation functions for water in aqueous TMAO	42
2.1.1	Diffraction of a plane wave through an aperture	45
2.1.2	Superposition	46
2.1.3	Bragg's Law	47
2.1.4	Young's double slit experiment	50
2.1.5	The scattering vector Q	54
2.1.6	Scattering of a plane wave from a fixed point	56
2.1.7	Fission	63
2.1.8	Spallation	64
2.1.9	ISIS	65
2.1.10	NIMROD	67
2.1.11	Neutron scattering sample holders	68
2.2.1	Lennard - Jones potential	76
2.2.2	EPSR weights matrix	80
2.2.3	Ambient water SDF	84
2.2.4	EPSR flowchart	85
2.3.1	Water hydrogen bond angle	89
2.3.2	Hydrogen bonding intermediates	92
2.3.3	Water dipole angle	94
2.4.1	Spin projection	96
2.4.2	Spin alignment	98
2.4.3	Nuclear shielding	100
2.4.4	Ethanol NMR	102
2.4.5	Spectral density function	104
2.4.6	T_1 relaxation	105

2.4.7	Spin echo pulse sequence	108
2.7.1	Jablonski diagram	119
2.7.2	I27	120
2.8.1	Circular dichroism	123
3.2.1	F(Q) for aqueous KF at 2.4 mol/kg H ₂ O	129
3.2.2	O _w - O _w RDFs for pure water and aqueous potassium halides at 2.4 mol/kg H ₂ O	129
3.2.3	Ion - O _w RDFs for aqueous potassium halides at 2.4 mol/kg H ₂ O . . .	131
3.2.4	Ion - H _w RDFs for aqueous potassium halides at 2.4 mol/kg H ₂ O . . .	131
3.3.1	Ion - water dipole angle distributions for aqueous potassium halides at 2.4 mol/kg H ₂ O	133
3.3.2	Water - water dipole angle distributions for aqueous potassium halides at 2.4 mol/kg H ₂ O	135
3.4.1	ΔH_{EPSR} for potassium halides as a function of concentration	137
3.4.2	Bulk water - water hydrogen bond interaction energy distributions in aqueous potassium halides at 2.4 mol/kg H ₂ O	139
3.4.3	Average bulk water - water hydrogen bond interaction energy in aqueous potassium halides	139
3.4.4	Average ion - water enthalpic interaction energy as a function of distance for aqueous potassium halides	140
3.5.1	T_1^{-1} decay times for water in aqueous potassium halides	142
3.5.2	Water self-diffusion coefficient for aqueous potassium halides	143
3.5.3	D^{-1} vs T_1^{-1} for water in aqueous potassium halides	147
4.1.1	TMAO concentration in muscle tissue of snailfish	154
4.3.1	O _w - O _w RDFs for water and aqueous TMAO at 25 bar and 4 kbar .	159
4.3.2	Water - water SDFs for water and aqueous TMAO at 25 bar and 4 kbar	161
4.4.1	Bulk water - water dipole angle distributions for water and aqueous TMAO at 25 bar and 4 kbar	164

4.4.2	Bulk water - water hydrogen bond interaction energy distributions for water and aqueous TMAO at 25 bar and 4 kbar	166
4.5.1	$O_T - O_w$ RDFs for aqueous TMAO at 25 bar and 4 kbar	168
4.5.2	O_T - water dipole angle distributions for aqueous TMAO at 25 bar and 4 kbar	170
4.5.3	O_T - water hydrogen bond interaction energy distributions for aqueous TMAO at 25 bar and 4 kbar	171
4.6.1	$C_T - O_w$ RDFs for aqueous TMAO at 25 bar and 4 kbar	173
4.6.2	$O_w - H_w$, $O_T - H_w$, and $C_T - H_w$ RDFs for water and aqueous TMAO at 25 bar and 4 kbar	174
5.2.1	Melting temperature of aqueous $Mg(ClO_4)_2$ and estimated induced pressure	180
5.2.2	Concentration of TMAO in the muscle tissue of bony fish and estimated pressure resting ability of TMAO	181
5.3.1	$O_w - O_w$ RDFs for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	186
5.3.2	Water - water SDFs for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	187
5.4.1	$O_T - O_w$ and $C_T - O_w$ RDFs for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	189
5.4.2	$Mg - O_w$ and $Cl_p - O_w$ RDFs for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	190
5.4.3	$O_T - Mg$ and $O_T - Cl_p$ RDFs for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	191
5.4.4	$C_T - Mg$ and $C_T - Cl_p$ RDFs for tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO	192
5.5.1	Bulk water - water hydrogen bond interaction energy distributions for binary and tertiary aqueous solutions of $Mg(ClO_4)_2$ and TMAO . . .	193

5.5.2	Average bulk water - water hydrogen bond interaction energies for binary and tertiary aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO . . .	193
5.6.1	^1H NMR peak shifts for binary and tertiary aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO	196
5.6.2	T_1^{-1} and D for water in binary and tertiary aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO	198
5.6.3	Calculations of T_1 for binary and tertiary aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO	199
5.6.4	T_1 vs D for water in binary and tertiary aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO	200
5.7.1	Abundance of bifurcated hydrogens in binary and tertiary aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO	202
5.8.1	Effective concentration master curves for binary and tertiary aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO	204
5.8.2	Determination of g value for binary and tertiary aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO	205
6.3.1	$O_w - O_w$ RDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine	215
6.3.2	Water - water SDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine	215
6.4.1	Structure of glycine	216
6.4.2	$N_a - O_w$ and $N_a - H_w$ RDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine . . .	217
6.4.3	$C_c - O_w$ and $C_c - H_w$ RDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine . . .	219
6.4.4	Glycine - water SDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine	221
6.4.5	$C_{bk} - O_w$ and $C_{bk} - H_w$ RDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine . . .	222
6.5.1	$N_a - \text{Mg}$ and $N_a - \text{Cl}_p$ RDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine . . .	223
6.5.2	$C_c - \text{Mg}$ and $C_c - \text{Cl}_p$ RDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine . . .	225
6.5.3	$C_{bk} - \text{Mg}$ and $C_{bk} - \text{Cl}_p$ RDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine . .	225
6.6.1	$H_a - O_c$ and $H_{bk} - H_{bk}$ RDFs for aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine . . .	227
6.6.2	Clustering of glycine in aqueous $\text{Mg}(\text{ClO}_4)_2$ and glycine	228

6.7.1	I27 with highlighted tryptophan	230
6.7.2	I27 CD spectra in aqueous $\text{Mg}(\text{ClO}_4)_2$	231
6.7.3	I27 structure with A' and G strands highlighted	232
6.7.4	Secondary structure fractions of I27 in aqueous $\text{Mg}(\text{ClO}_4)_2$	233
6.7.5	Folded fraction of I27 as a function of urea concentration in the presence of TMAO and $\text{Mg}(\text{ClO}_4)_2$	235
6.7.6	Folded fraction of I27 as a function of urea concentration in the presence of mixed TMAO and $\text{Mg}(\text{ClO}_4)_2$	236
6.8.1	CLN025	238
6.8.2	Corrected neutron scattering data on aqueous CLN025	239
7.1.1	The links between structure, dynamics, and thermodynamics in aqueous potassium halides	243
7.1.2	$O_w - O_w$ RDFs for pure water and aqueous TMAO at 25 bar and 4 kbar	246
7.1.3	Average bulk water - water hydrogen bond interaction energy in binary and tertiary solutions of TMAO and $\text{Mg}(\text{ClO}_4)_2$	248
7.1.4	Fraction of folded I27 as a function of urea concentration in mixed TMAO and $\text{Mg}(\text{ClO}_4)_2$ and free energy of unfolding	251
A.0.1	Pure water $F(Q)$	262
A.0.2	Aqueous potassium halides $F(Q)$	263
A.0.3	$O_w - O_w$ RDFs for aqueous potassium halides at 0.6 and 1.2 mol/kg H_2O	264
A.0.4	Ion - water dipole angle distributions for aqueous potassium halides at 0.6 and 1.2 mol/kg H_2O	264
A.0.5	Bulk water - water dipole angle distributions for aqueous potassium halides at 0.6 and 1.2 mol/kg H_2O	265
A.0.6	Bulk water - water hydrogen bond interaction energy distributions for aqueous potassium halides at 0.6 and 1.2 mol/kg H_2O	266

A.0.7	Average ion - water enthalpic interaction energy for aqueous potassium halides at 0.6 and 1.2 mol/kg H ₂ O	266
A.0.8	Calculation of distance at which enthalpic interactions between ions and water become 0	267
B.0.1	Pure water and aqueous TMAO at 25 bar and 4 kbar F(Q)	269
C.0.1	Tertiary solutions of TMAO and Mg(ClO ₄) ₂ F(Q)	271
D.0.1	Aqueous solutions of Mg(ClO ₄) ₂ and glycine F(Q)	273

Chapter 1

Introduction

When one begins to consider the subject of biology, one is immediately overwhelmed by the various different length scales this encompasses [6]. At its largest, the biological field of ecology considers how groups of organisms interact with their environment, and therefore its length scale can be considered to occupy the whole biosphere of Earth. At its intermediate length scale it considers individual plants and animals through the fields of zoology and botany, from the giants of blue whales and redwood trees to microscopic tardigrades and phytoplankton. Finally, at its smallest length scale, it considers the fundamental processes vital to life through the field of molecular biology, such as transcription and translation to build proteins from the genetic code built into DNA [7]. However, despite the vast array of sizes, there exists a single molecular species that underpins them all, and without which biological organisms and phenomena as we know them could not exist: water.

Indeed, liquid water is vital to life at all length scales [8, 9], performing a number of roles from providing the largest natural environment on Earth to the driving force behind biological molecule self-assembly [10]. In this thesis we will focus on its most fundamental relationship with biology: its interaction with molecules relevant to life, such as proteins, ions, and osmolytes. More specifically, and as the title would suggest, we will concern ourselves with how these interactions are perturbed in extreme envi-

ronments. The variety of extreme environments on Earth in which life is still capable of thriving is truly awe inspiring [11]. Even in the harshest of conditions, subjected to extremes of temperature, pressure, pH, salinity, radioactivity, often simultaneously, life has been able to adapt and persist. This has led to the classification of organisms that thrive under such extreme conditions as “extremophiles”. Not only do these present an exciting area of study for pure scientific intrigue, a detailed understanding of their biological adaptations has profound implications in both medical and industrial research [12–16]. Fundamental insight into the interactions between water and biomolecules in extreme environments is therefore an extremely important and timely issue, as the ever expanding reach of humanity means that now more than ever these unique environments are under threat.

This first chapter will outline some of the key principles and background theory of the work covered in this thesis. In section 1.1 we will outline the basic principles of protein formation and folding. This serves as a useful tool to outline various covalent and non-covalent interactions present in molecular biology and how they are intimately connected to water. We will then give examples of some of the mechanisms that extremophiles use to survive perturbations to their aqueous environments by extreme conditions through protein adaptations and the use of protecting organic osmolytes in section 1.2. Next we will discuss the nature of liquid water in section 1.3, before discussing ways of monitoring perturbations to water structure and dynamics in sections 1.3.1 and 1.3.2. Finally we will present the aims and objectives of this thesis in section 1.4. Following this chapter we will describe the background theory and application of the various experimental methods used in this research in chapter 2, before presenting the key results in chapters 3 - 6.

1.1 Protein Structure and Folding

One of the many mechanisms extremophiles employ to survive perturbations to their aqueous environments is by the modification of their proteins, which allows them to be both stable and suitably flexible in order to carry out their biological roles [17]. In order to understand the significance of these adaptations, we must first define proteins

themselves. Proteins are naturally occurring chains of molecules called amino acids or peptides which fold into clearly defined 3D structures due to their interaction with their aqueous environment [18]. This structure is determined by their amino acid sequence and defines the role of the protein. The various roles performed by biological proteins are extensive and can include, but are certainly not limited to: catalysis, mechanical support, force mediation, enzymatic activity, molecular transport, and replication [19].

1.1.1 The Covalent and Non-covalent Interactions that Govern Protein Structure

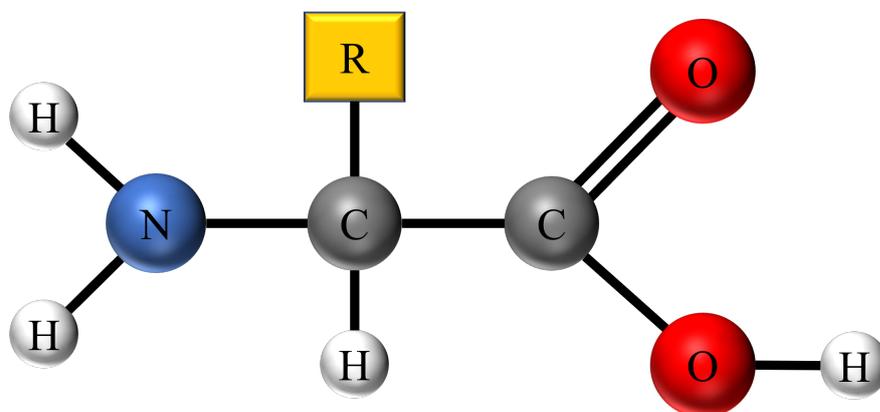


Figure 1.1.1: The structure of a single amino acid, featuring a common backbone consisting of an amine group (NH_2), a carbonyl group (COOH), and a central backbone carbon covalently bound to a single hydrogen and a side chain group (R). The chemical structure of this side chain group defines the species of the amino acid.

Amino acids themselves are relatively simple molecules which all consist of a common backbone comprising of an amine group (NH_2), a carbonyl group (COOH), and a central carbon which is covalently bound to a single hydrogen and side chain group (R). This basic structure is shown in figure 1.1.1. The chemical structure of this side chain then defines the species of the amino acid. There are 20 naturally occurring amino acids which go on to form every protein species in existence. A diagram displaying these 20 amino acids is presented in figure 1.1.2. Here the amino acids are displayed as they would occur as monomers in solution at neutral pH. Under these conditions amino acids take their zwitterionic form, hence the carbonyl group becomes deprotonated and donates a

hydrogen to the amine group, rendering the two oppositely charged.

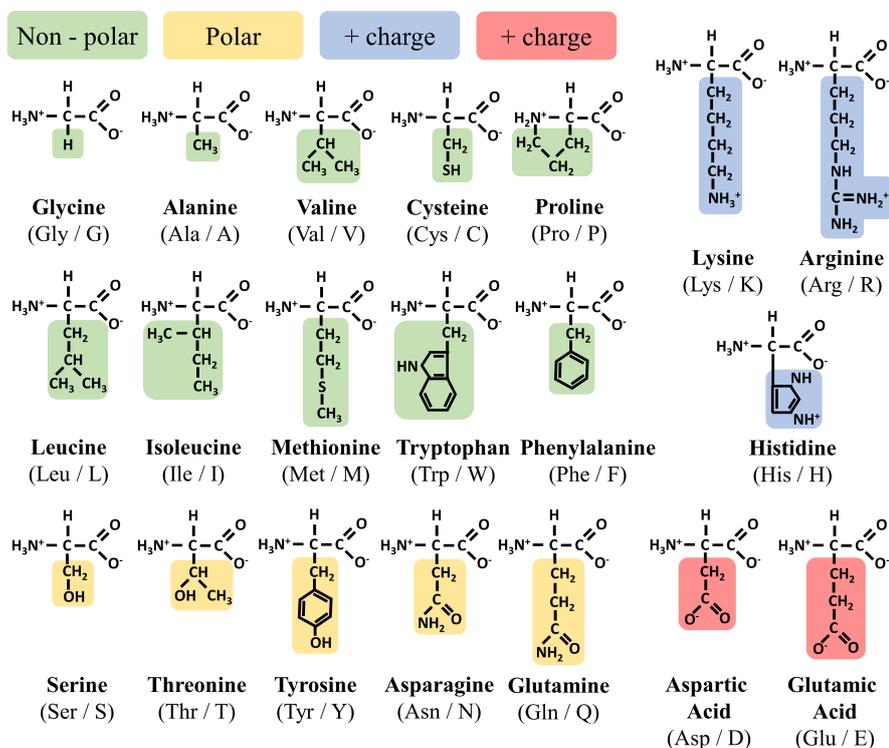


Figure 1.1.2: The 20 naturally occurring amino acids grouped according to the character of their side chains. Non-polar side chains are shown in green, polar side chains are shown in yellow, positively charged side chains are shown in blue, and negatively charged side chains are shown in red.

Proteins are then formed from amino acids through many dehydration synthesis reactions between the amine and carbonyl groups, which forms a peptide bond between the amino acids and releases a water molecule [20]. This forms the long chains of amino acids which defines the protein's primary structure. Once the proteins primary structure is formed the side chains and remaining backbone NH and CO groups can then interact with groups on other amino acids in their proximity to form secondary structure elements, including α -helices, and parallel/anti-parallel β -sheets. Secondary structures elements can then interact through hydrophobic interactions, ionic bonds, hydrogen bonds, and disulphide bridges, which causes the protein to fold into its tertiary structure. A diagram

of this process is displayed in figure 1.1.3. In the case of large proteins, tertiary structures can also go on to interact and form quaternary structures.

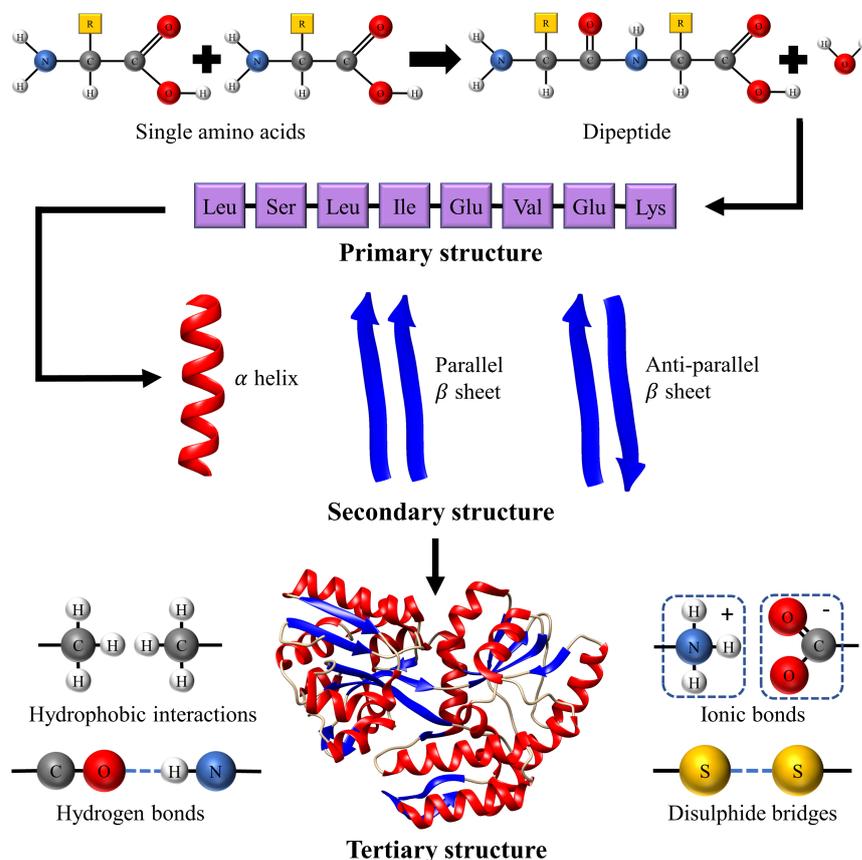


Figure 1.1.3: The classification of the structures present in folded proteins. Amino acids form a peptide bond between their amine and carbonyl groups, releasing a water molecule, to form peptides. The sequence of amino acids defines the proteins primary structure. Individual amino acids in the peptide chain then go on to interact with other nearby amino acids to form the protein secondary structure. These secondary structure elements then interact through various means to form the protein tertiary structure. Maltose binding protein is shown as an example protein which contains the three secondary structure elements described in this figure [21] (PDB code: 7MQ7).

Understanding the interactions between secondary structure elements which stabilise the folded structure of proteins is of crucial importance if one wishes to understand how extremophile proteins are able to maintain structure and flexibility in their aqueous environments, and as such they will now be overviewed. One of the most prevalent interactions that both stabilises the tertiary structure of the protein and the secondary structure elements themselves is hydrogen bonding. Hydrogen bonds occur when a

hydrogen is covalently bound to another host element that has a large electronegativity difference between it and the hydrogen, such as nitrogen or oxygen. As such the electrons in the bond are pulled closer to the host element, rendering the hydrogen more positively charged and the host element more negatively charged. The bond is therefore relatively polar. This positively charged hydrogen can then be electrostatically attracted to another negatively charged neighbouring atom type either within the same molecule or on another molecule in its proximity. The hydrogen can therefore be thought of as being shared between the host atom and the neighbouring atom, and behaves rather like a bond, hence it is called a hydrogen bond [22]. Water, containing two OH groups, can form a network of hydrogen bonds, and can form hydrogen bonds with solvent exposed CO, NH, OH, and SH groups on the protein. These have shown to be potentially crucial in protein folding in turn regions of protein structures, where the formation of a turn is mediated by a single water molecule which forms a hydrogen bond with two NH groups located either side of the apex of the turn [23]. The mediation of turns in this way also allows the remaining amino acid sequences either side of the turn to come into close enough proximity to form β -sheets.

In the case of α -helices, the backbone NH group on one amino acid forms a hydrogen bond with the backbone CO group on the amino acid four amino acids ahead in the sequence [24]. In the case of β -sheets, the backbone NH group of one amino acid forms a hydrogen bond with the backbone CO group of an amino acid on the neighbouring strand. Due to the large variety of different conformations this can take and atomic species involved, hydrogen bonds can occupy a broad range of bonding energies. They are often very short lived, on the order of picoseconds in the case of hydrogen bonds in liquid water [25–27], but can also persist for much longer, such as the hydrogen bonds that hold together DNA molecules which can last for hundreds of years [28, 29]. Hydrogen bonding between water molecules will be overviewed in greater detail in section 1.3.

Ionic bonds function through a similar mechanism. In this instance groups of atoms located on the side chains of amino acids that have an overall charge are electrostatically attracted to groups of atoms on a nearby side chain that have an overall opposite charge [22]. The greater difference in charge between the two interacting groups means that

they are stronger than hydrogen bonds, but occur less frequently. As both ionic and hydrogen bonds are driven by favourable electrostatic interactions, water molecules will also orient a single hydrogen towards a negatively charged group, in the same way it would towards a hydrogen bond accepting CO group, or away from a positively charged group, in the same way it would away from a hydrogen bond donating NH group. The strongest class of bonds within protein structures are disulphide bonds. These occur between the side chain SH groups of two appropriately placed cysteine amino acids. The two groups are then oxidized and form a covalent chemical bond between the two sulphur atoms [30].

These three classes of intrapeptide interactions would therefore, in principle, persist regardless of the environment of the protein once they are formed, as they depend only on the chemical makeup of the interacting groups. However, the fourth class of interactions, hydrophobic interactions, occur as a direct consequence of the interactions between the peptide chain and its aqueous environment. Consider a particle introduced into water. In order to accommodate this particle the water molecules must organise themselves around the particle, and as such are interacting with it through electrostatic and van der Waals interactions [31]. In the case where the particle is either positively or negatively charged, the surrounding water molecules will interact with the particle through favourable enthalpic electrostatic interactions by orienting their polar OH groups away from or towards the particle respectively. These favourable enthalpic electrostatic interactions mean that the particle can be introduced into water relatively easily, and as such the particle can be classified as “hydrophilic”. It is this reason why salts, which dissolve into their constituent charged ions when introduced into water, exhibit high solubilities [28].

Let us now consider the case where the particle is uncharged and nonpolar. Because of this characteristic, the water molecules around the particle will have no tendency to orient their polar OH groups towards or away from the particle, and there is very little in the way of favourable enthalpic electrostatic interactions. However, the water molecules must organise themselves around the particle nevertheless, and this induced ordering of the water molecules means that there is an associated entropy penalty which is not

compensated by enthalpic interactions. The importance of balances between enthalpy and entropy will be discussed shortly. This also comes at the expense of breaking hydrogen bonds that existed between water molecules prior to the introduction of the particle which are not compensated for by the formation of new favourable enthalpic electrostatic interactions between the particle and the surrounding water molecules [31–34]. The introduction of an uncharged non-polar particle into water is therefore energetically unfavourable. As such they exhibit low solubilities, and are referred to as “hydrophobic”. In reality this is a slight misnomer, as they are not really hydrophobic, so much as hydrononchalant. Hydrophobic solutes can therefore minimise this entropy penalty by clustering together in solution, as this requires fewer water molecules to hydrate the solutes. This is explained schematically in figure 1.1.4. This is the origin of the hydrophobic interaction. Amino acids with long non-polar side chains, such as those featuring many CH bonds which are non-polar due to the low electronegativity difference between carbon and hydrogen, cluster their side chains together in the core of the protein so as to minimise their contact with water and drive protein folding.

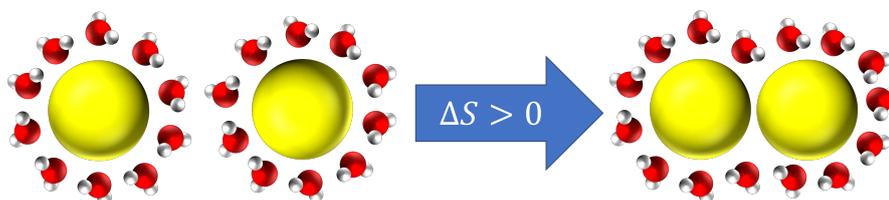


Figure 1.1.4: A schematic illustration of two hydrophobic solutes clustering together in solution to minimise the entropy penalty associated with their hydration.

1.1.2 Free Energy Changes in Biomolecular Self-Assembly - Protein Folding

The final folded structure of a protein, referred to as its native state, and its folding pathway are therefore defined by its amino acid sequence [20] and its interaction with its aqueous environment, however the mechanism by which a protein spontaneously adopts its native conformation based on its amino acid sequence is still not completely understood. The process of protein folding can be discussed in terms of the difference in Gibbs free energy between its folded and unfolded state. The Gibbs free energy is

a general measure of the thermodynamic stability of a given state [18], and therefore a process will occur spontaneously if the difference between the Gibbs free energy of the starting state and the final state is negative. The Gibbs free energy change associated with a particular transition is given by equation 1.1.1, where ΔG is the change in Gibbs free energy, ΔH is the change in enthalpy, T is absolute temperature, and ΔS is the change in entropy.

$$\Delta G = \Delta H - T\Delta S \quad (1.1.1)$$

The change in Gibbs free energy therefore contains two components: enthalpy and entropy. Enthalpy is a measure of the interactions between particles in a given state, such as the breaking or forming of hydrogen bonds between water molecules or within a protein. Like Gibbs free energy, it also contains two components: the change in internal energy of the system ΔE , and the work done on the system at constant pressure $P\Delta V$. The change in enthalpy is therefore expressed in equation 1.1.2.

$$\Delta H = \Delta E + P\Delta V \quad (1.1.2)$$

Entropy is a measure of disorder of a system, and is directly related to the total number of possible conformations a system can occupy. It is expressed in equation 1.1.3, where S is entropy, k is the Boltzmann constant ($1.381 \times 10^{-23} \text{ m}^2\text{kgs}^{-2}\text{K}^{-1}$ [28]), and Ω is the total number of accessible states.

$$S = k \ln \Omega \quad (1.1.3)$$

In general then, a process will occur spontaneously if it can form a greater number or more stable enthalpic interactions in its final state compared with its initial state, or it becomes more disordered, as these result in a more negative ΔG . We can now return to the concept of introducing a particle into water. The introduction of a particle results

in an ordered hydration shell as water molecules reorient to accommodate the solute. This ordering is entropically unfavourable, as the accessibility of different conformational states is reduced, and therefore $\Delta S < 0$. This means ΔG becomes more positive and the particle will not easily be introduced into water. If the particle is charged the water molecules can orient their hydrogens appropriately resulting in favourable electrostatic enthalpic interactions, and therefore $\Delta H < 0$. This compensates for the reduction in entropy, and makes ΔG more negative, and hence the particle will be easily introduced into water [31, 35]. When we consider protein folding this becomes a complex balancing act between the two quantities arising from several different areas. These include, but are not limited to: the enthalpically favourable formation of various intrapeptide bonds, the enthalpically favourable formation of hydrogen bonds between the protein and its solvent environment, the entropically favourable dehydration of hydrophobic side chains by clustering them together in the core of the protein, and the entropically unfavourable structuring of the protein, as an unfolded protein can adopt a greater range of conformations. This process can be simplified by considering protein folding/unfolding as a two state process, where the folded and unfolded state are separated by an energy barrier, as shown in figure 1.1.5.

Here we observe a simplified picture where a protein is forced from a folded to an unfolded state by following some reaction coordinate, such as temperature, through a rate limiting transition state. This transition state is the conformation with the highest free energy along the folding pathway and is typically more reminiscent of an expanded folded state than the unfolded state. In reality this process is much more complex, and many proteins, particularly above about 100 residues, help direct their folding pathways through a number of, usually thermodynamically unstable, intermediates [36, 37]. These can occasionally be well characterised as they are stabilized through strong interactions, such as disulphide bonds, however more often they are stabilized through hydrogen bonding, such as the mediation of turns by a hydrogen bonded water molecule [23], which cannot be trapped. A protein folding energy landscape can therefore be more accurately thought of as a funnel shape [38], where the protein begins at the top of the funnel in an unfolded and energetically unfavourable state, and moves down the funnel

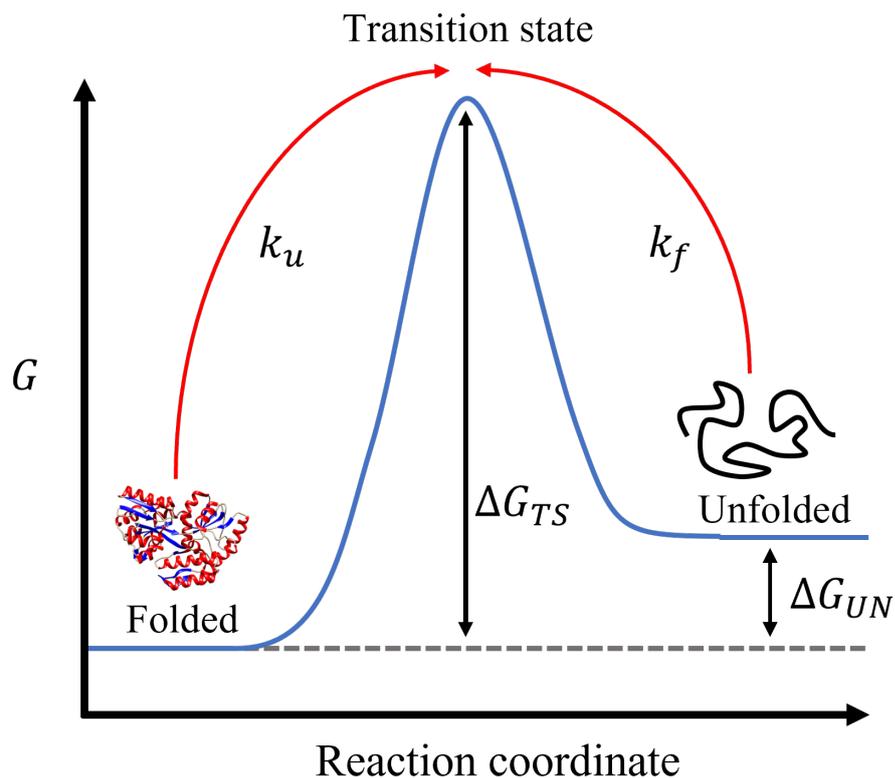


Figure 1.1.5: A schematic illustration of the change of free energy associated with protein unfolding as one follows a particular reaction coordinate (temperature, pressure, applied force, etc.). Here the folded and unfolded states are separated by an energy barrier of magnitude ΔG_{TS} corresponding to the formation of a transition state. The difference in free energy between the folded and unfolded state is then ΔG_{UN} . This unfolding transition is then classified by two rate constants, k_u and k_f , which correspond to the rate constants of unfolding and folding respectively.

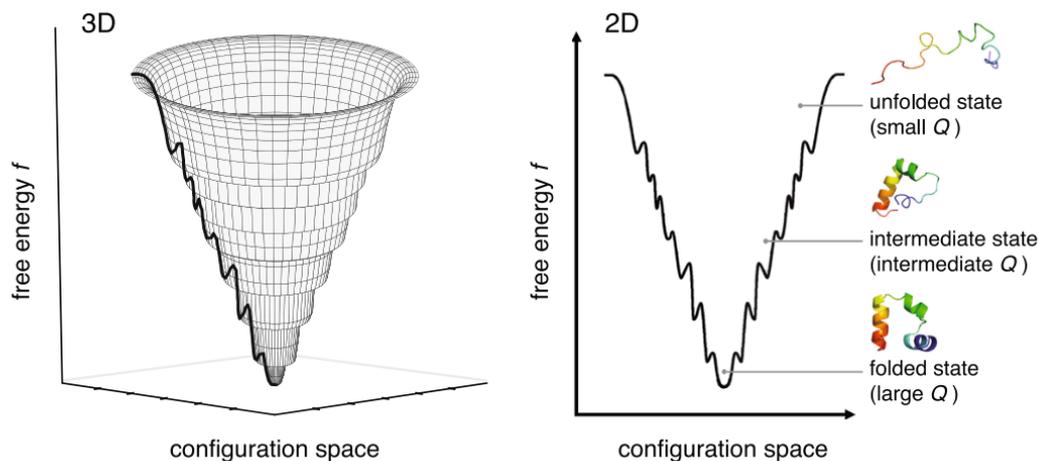


Figure 1.1.6: A 3D (left) and 2D (right) schematic illustration of the funnel shaped energy landscape associated with protein folding. Reproduced from reference [38] Nature Springer Ltd: Scientific Reports, ©2009.

through a series of folding intermediates. This is illustrated in figure 1.1.6. Through this process the available conformations the protein can adopt become increasingly restricted until the protein eventually arrives at its native folded state. The native state of the protein is considered to have the lowest possible free energy, however it is also possible that it is only the most stable conformation that is kinetically accessible.

1.2 Adaptations to Extreme Environments

The extremes of temperature, pressure, salinity, pH, etc. present in extreme environments means that water in extreme conditions is strongly perturbed, and therefore the subtle balance of enthalpy and entropy that result in folded protein stability is also strongly perturbed. Extremophiles therefore employ a range of strategies to keep their proteins folded and appropriately flexible. In this section we will focus on two of these methods: adaptations to the amino acid sequence of their proteins, and the accumulation of protecting organic osmolytes. More specifically we will focus on three classes of extremophiles that are particularly relevant to the results of this thesis: cold adapted extremophiles (psychrophiles), high salt concentration adapted extremophiles (halophiles), and pressure adapted extremophiles (piezophiles).

1.2.1 Adapting to a Cold Aqueous Environment - Psychrophiles

Strictly speaking, psychrophilic organisms are classified as those which grow below temperatures of 20 °C [17], however the true limit for psychrophilic organisms can be much lower than this, such as microbial communities living within porous rocks in Antarctic dry valleys which can reach -60 °C [39]. Low temperatures represent a number of challenges to the fundamental processes required for living organisms. The reduction in available thermal energy results in a less dynamic and more rigid solvent environment, and means that proteins are less able to sample their available conformational space and as such are less flexible [40]. This in turn means that classes of proteins such as enzymes are far less efficient at lower temperatures due to reduced flexibility and the increased difficulty in overcoming energy barriers related to catalysis [39].

Therefore the main challenge to be overcome by psychrophilic proteins is how to maximise their flexibility so as to use what little thermal energy is available in the solvent environment most efficiently. This is achieved through a number of modifications to the amino acid makeup of their proteins. In a broad sense, all of these adaptations result in a reduction of stabilising interactions within the protein. The first and perhaps most obvious consequence of this is that psychrophilic proteins are less thermodynamically stable and will unfold at lower temperatures than analogues from organisms that thrive at more standard conditions (mesophiles) [41], however the nature of their environment means that this disadvantage is of little concern. Another consequence of this is that unfolding in psychrophilic proteins appears increasingly two state as their folding funnel, as shown in figure 1.1.6 is effectively much shallower and folding intermediates are less populated or skipped entirely.

These stabilising interactions are reduced by various amino acid substitutions resulting in fewer or less stable enthalpic intraprotein interactions, such as: a reduction in the number of disulphide bridges, a reduction in the number of intraprotein electrostatic interactions such as hydrogen bonds and interactions between aromatic residues, a decreased arginine content to inhibit the formation of salt bridges, and fewer and weaker metal binding sites [42]. Stability is also reduced by increasing the entropy penalty asso-

ciated with interactions between hydrophobic side chains and the aqueous environment, and therefore a reduction in the hydrophobic effect. This is achieved through a reduction of the hydrophobicity of the protein core and an increased number of hydrophobic residues on the surface of the protein.

Psychrophilic proteins are also allowed to be more flexible by extending their surface loop regions through an increase in neutral/small amino acids and a reduction in proline content. This increase in flexibility is particularly important around protein binding sites. Increased flexibility of protein binding sites allows the solute of interest to bind with a greater degree of complementarity, resulting in a reduced activation of catalytic reactions and an increase in the rate of substrate turnover [41]. This can be achieved through an increase in glycine content at the binding site. The binding sites are also typically larger while maintaining the residues that are principally responsible for catalytic activity.

1.2.2 Adapting to a Saline Aqueous Environment - Halophiles

Halophiles are defined as organisms which thrive in salt concentrations above 8.8 wt% [11]. As introduced in section 1.1.1, salts dissolve into their constituent ions when dissolved in aqueous solution [28]. Proteins in such environments are therefore surrounded not only by water molecules, but by a sea of dispersed charged ions. The observation that salt species can cause proteins to aggregate and precipitate was made as early as 1888 by Franz Hofmeister [43], who showed that high salt content caused the precipitation of hen egg white lysozyme, and that the rate of precipitation was dependent on the salt species. This led to the first iteration of the “Hofmeister series”, which ranks anions and cations by their ability to destabilise folded proteins. It has been long suggested that this is as a result of the perturbation of the solvent environment due to the presence of the ions, and has resulted in a host of studies considering the effect of “Hofmeister salts” on the structure and dynamics of water [24, 44–52]. Since its first iteration it has been greatly expanded, and it is now accepted that a simple order which describes an ion’s effect on every protein or biomolecule based on its interaction with water is far from sufficient [44, 53]. The ability of a particular ion species to stabilise or destabilise a protein is

highly protein dependent and depends on the protein's surface chemistry. It therefore more appropriate to describe the Hofmeister series as a Hofmeister phase diagram, which depends on various parameters such as the charge density, hydrophobic character, pH, and ion concentration. A complete review of these phenomena goes outside the scope of this thesis and therefore will not be discussed further.

Broadly speaking, it can be stated that the main issues regarding protein structure and flexibility in highly concentrated salt environments primarily arise from two areas: the screening of favourable electrostatic interactions required for protein stability such as hydrogen bonding by the ions, and perturbations to hydrophobic interactions due to the structural and dynamic perturbations to the solvent environment [17]. To resist these issues, halophilic proteins typically contain a much higher proportion of acidic residues on their surface relative to their mesophilic analogues [16, 54]. The role of these acidic residues is not entirely clear, but a number of possibilities have been put forward. One possibility is that the ions in solution strongly bind water molecules, and the negative charge on the acid residues allows the protein to compete for hydration and maintain the protein - solvent interactions that are responsible for protein folding. It is also suggested that acidic surface residues increase protein surface hydration by binding hydrated cations, which helps maintain water - protein interactions. This also acts as a protective mechanism for stabilising electrostatic interactions in the protein. The binding of cations to acidic residues means that the cations cannot preferentially associate around hydrogen bonding or ionic interaction sites, and as such cannot screen these interactions and allow the protein to remain folded.

High salt concentrations have been demonstrated to either enhance or reduce the hydrophobic effect depending on the salt species [45, 49, 55]. For example, Kim *et al.* [56] demonstrated that the solubilities of the model hydrophobes methane (CH_4) and ethane (C_2H_6) are decreased by increasing NaCl concentration. The solubility of hydrophobic moieties is dependent on how energetically favourable their hydration is, as discussed in section 1.1.1, hence a reduction in solubility corresponds to an enhanced hydrophobic effect as they will be more inclined to form hydrophobically associated clusters to minimise their hydration. This eventually leads to aggregation and precipitation, referred

to as “salting out”. Halophilic proteins therefore combat this enhanced hydrophobic effect at high salt concentrations in a similar manner to psychrophiles, by utilising fewer and smaller hydrophobic residues in their hydrophobic core [16, 17]. This allows the proteins to remain suitably flexible in concentrated salt environments. The enhanced hydrophobic effect is also counteracted by the increase in acidic surface residues mentioned above, as electrostatic repulsion between negatively charged surface groups also helps to enhance protein flexibility.

Unlike environmental externalities like pressure and temperature, which cannot be escaped, high salt concentrations can be circumvented, and is a popular strategy utilised by halophilic organisms [54]. One potential method of achieving this would be an uptake of water into the cells, effectively reducing the salt concentration, however this is energetically costly and as such is not employed. It is however common for halophilic organisms to preferentially uptake potassium cations while expelling sodium cations. This allows the cells to maintain a positive osmotic pressure inside the cells which allows for cell expansion while avoiding detrimental effects of sodium cations to the stability of their proteins. An alternative is rather than the uptake of ions, is the uptake of “compatible solutes”, also known as osmolytes. These are organic molecules that stabilise the folded structure of proteins and therefore improve their activity against denaturing externalities [11, 57–61]. In a similar vein to the Hofmeister series, the observation that organic osmolytes can stabilise protein structure has prompted a host of studies into their perturbations to the aqueous environment [2, 3, 62–67]. There are a wide variety of protecting osmolytes, commonly employed examples include methylamines like trimethylamine N-oxide (TMAO) and betaine, polyols such as glycerol or sorbitol, and amino acids and amino acid derivatives such as glycine and ectoine. These function through preferential exclusion from the protein surface [67–73] and therefore stabilise the hydration shell of the protein so as to maintain the interaction with the solvent environment. These can either be incorporated into the cell by extracting them from their immediate aqueous environment, or synthesised within the cell itself. The fact that protecting osmolytes are preferentially excluded from the protein surface means that their stabilising effects must be mediated by several neighbouring water molecules, and therefore understanding their

interaction with water is of crucial importance. The mechanism of the organic osmolyte TMAO will be discussed in much greater detail in chapters 4 and 5.

1.2.3 Adapting to a High Pressure Aqueous Environment - Piezophiles

Piezophiles are those organisms which grow and function under high hydrostatic pressure, such as the bottom of the deep sea [74]. Water in this environment is less well structured, as neighbouring water molecules are forced closer together. However, within the range of pressures that can naturally exist on Earth, higher pressure also results in increased water dynamics, as increased electrostatic repulsion due to shorter intermolecular distances result in a decreased viscosity and an increased self diffusion coefficient [75]. The inaccessibility to their native environment due to high pressure and the difficulties associated with performing high pressure experiments, such as the need for robust sample holders which often compromise signal quality, means that piezophilic proteins remain some of the least investigated. The effect of high pressure on proteins causes them to compress. Compression under extreme pressure is relatively modest, with proteins only becoming approximately 1% more compact at 2 kbar [76, 77]. This is coupled with a tendency to reduce oligomerisation, separating structures consisting of multiple associated proteins. However beyond this, at approximately 4 kbar [78] high hydrostatic pressure causes the protein to unfold. This is primarily due to a shifted thermodynamic equilibrium where proteins under pressure occupy more water molecules in the hydrophobic core of the protein, disrupting the hydrophobic effect between side chains, and causing the protein to unfold [79]. It is observed that an increase of the size of internal cavities through site specific mutation of mesophilic proteins results in unfolding at comparatively lower pressures.

The nature of piezophilic organisms means that they are often polyextremophiles, adapted to deal with more than one extreme environmental stress, as they will inhabit the sea cold floor or will gather around hot hydrothermal vents [80]. As such it is difficult to comment on protein adaptations that are specific to piezophilic stability, nevertheless, there are trends. Pressure denaturation can be resisted through the elimination of cavi-

ties within the protein core [81], hence piezophilic proteins have a modified amino acid sequence which allows for a more densely packed hydrophobic core, makes the protein less compressible, and shifts the equilibrium back towards a totally dehydrated hydrophobic core state [74]. A greater abundance of smaller hydrogen bond forming amino acids is also observed, which is believed to aid in the formation of the dense hydrophobic core by forming more intraprotein bonds. This also likely helps maximise the effects of crucial protein - water hydrogen bonds to protein stability, as the increased dynamics in pressurised water will result in shorter lived hydrogen bonds. Piezophilic proteins are also observed to more frequently form multimers. This allows the individual protein monomers to become more compact, isolates them further from their aqueous environment to prevent hydration of the hydrophobic core, and stabilises hydrogen bonding within the protein monomers [17].

As discussed with halophiles, another popular mechanism by which piezophilic organisms are capable of thriving in their environment is by the accumulation of organic osmolytes [11, 82, 83], which have been shown to stabilise biomolecules against pressure induced denaturation [68, 84, 85]. A particular example of this is demonstrated by Yancey *et al.*, who observed that the concentration of the osmolyte TMAO in the muscle tissue of snailfish was linearly proportional to the depth from which the snailfish was harvested, and therefore the pressure of its environment [86]. As discussed in section 1.2.2, TMAO is preferentially excluded from the protein surface and therefore helps stabilise the protein - solvent interactions, and must do so through several mediating water molecules. It has been observed that TMAO can strengthen water - water hydrogen bonding and slow water molecule dynamics [3], counteracting the effects of pressure and rendering it more similar to ambient water, promoting protein stability. The effects of TMAO on water structure and dynamics will be discussed in much more detail in chapters 4 and 5.

The modifications to protein structures in extremophiles are therefore all made such that the protein can remain stable and flexible in response to perturbations to its solvent environment. Environmental perturbations resulting in enhanced water - water hydrogen bonding and slowed water dynamics are overcome by a reduction in protein stabilising

interactions. Environmental perturbations resulting in destabilised water - water hydrogen bonding and accelerated dynamics are overcome by an increase in protein stabilising interactions. The next step in understanding the biological role of water in extreme conditions is therefore to obtain a detailed understanding of the solvent environment, as will now be attempted in section 1.3.

1.3 The Liquid State

Section 1.2 overviewed how extremophilic organisms modify their proteins and biochemistry to help overcome the challenges presented by extreme conditions. As the changes in their environment through temperature, pressure, salinity, etc. are felt through their interactions with their aqueous environment, it is crucially important to their investigation that one understands how these extreme conditions perturb the water surrounding biomolecules. The study of water in general is an exceptionally long standing field of study, and can be thought of as dating back to the ancient Greeks [9]. They too recognised the importance of water and proposed that everything in the universe was made by the combination of four different elements: earth, air, fire, and water, each of which had their own geometric form. In the following two sections we will overview some of the more recent results than the ancient Greeks managed, and examine how water structure and dynamics are studied through a variety of techniques.

1.3.1 Monitoring Perturbations to Water Structure

The nature of liquid water means that its constituent molecules are always in relatively close contact and interact through weak intermolecular forces which can break and reform on short timescales, allowing it to flow [28, 87, 88]. The fact that the molecular matrix of water is constantly in motion means that trying to visualise its structure is conceptually difficult. Unlike crystalline solids, which exhibit long range order and can be described by considering an individual unit cell, the weak intermolecular interactions between water molecules means that it can only be described by its short range order [89].

1.3.1.1 The Water Molecule

To understand this let us first consider the structure of the water molecule itself, as displayed in figure 1.3.1. Its molecular formula is H_2O , and as such it consists of two hydrogen atoms covalently bound to a central oxygen atom. The OH bond length is determined to be 0.96 \AA and the HOH bond angle is calculated to be 104.5° [90]. Because oxygen is much more electronegative than hydrogen it means that the electrons in this OH bond are pulled closer to the oxygen, and as such the hydrogen is partially positively charged δ_+ and the oxygen is partially negatively charged δ_- [31]. These partial charges are estimated in the SPC/E model of water, which is shown to accurately model water dynamics and structure [91], to be $0.4238e$ (where e is the unit of elemental charge equal to the absolute charge on an electron/proton and calculated to be $1.602 \times 10^{-19} \text{ C}$ [28]) and $-0.8476e$ respectively. The valency of oxygen means that the formation of water results in two pairs of electrons located on the underside of the oxygen molecule. These electron pairs are positioned normal to the plane containing the HOH bond, and as such the charge distribution on the water molecule is roughly tetrahedral.

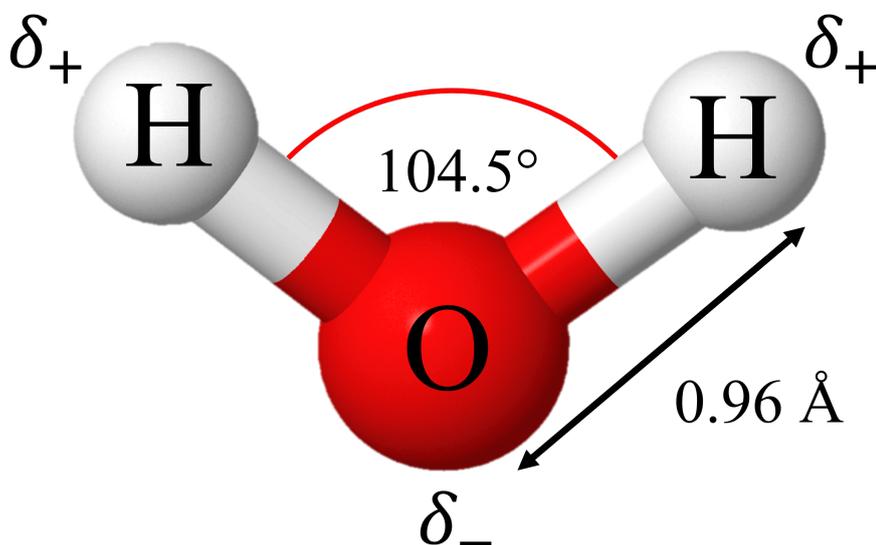


Figure 1.3.1: The molecular structure of the water molecule. Two hydrogen atoms are covalently bound to a central oxygen with a bond length of 0.96 \AA and a HOH bond angle of 104.5° . The electronegativity on the OH bond and the two pairs of lone electrons on the oxygen render the hydrogens partially positively charged δ_+ and the oxygen partially negatively charged δ_- .

Table 1.3.1: Properties of water (H₂O) and the structural analogue hydrogen sulphide (H₂S).

Property	H ₂ O	H ₂ S
Molar mass (g/mol)	18.02	34.081
Boiling point (K)	373.12	212.85
Melting point (K)	273.15	187.85
Specific heat capacity (constant volume) (J mol ⁻¹ K ⁻¹)	74.539	26

The localised positive charge on the hydrogen atom and the localised negative charge on the electron pairs on the oxygen atom mean that these can act as hydrogen bond donors and acceptors respectively [31, 88]. This is the key property that causes water to behave as it does and as such as been researched extensively [25, 75, 88, 92–98]. One of the simplest methods to illustrate the importance of these interactions is to compare some of the physical properties of water with the chemical H₂S, which has an identical chemical shape, but the lower electronegativity difference between hydrogen and sulphur means that it is less able to form hydrogen bonds. These differences are summarised in table 1.3.1 [31]. Here we observe that despite the increased size of H₂S, and therefore greater van der Waals interactions, the enthalpically favourable interactions through hydrogen bonding result in water having higher melting and boiling points, and a higher specific heat capacity as the presence of hydrogen bonds allows energy to be absorbed through many more avenues. Several definitions of the precise nature of a water - water hydrogen bond that consider various geometric, energetic, and topological criteria have been suggested [88, 93, 99], and these are discussed in more detail in section 2.3.2.

1.3.1.2 Tetrahedrally Structured Liquid Water

The fact that a hydrogen bond arises due to the electronegativity difference between two bound atoms means that they are highly directional. The roughly tetrahedral organisation of the hydrogen bond donating (hydrogen atoms) and hydrogen bond accepting (oxygen lone pairs) results in a roughly tetrahedral structure of liquid water. This is illustrated in figure 1.3.2. Here we observe that the average distance between a water molecule and a neighbour in its first hydration shell is approximately 2.8 Å and the

average distance between a water molecule and a neighbour in its second hydration shell is approximately 4.5 Å.

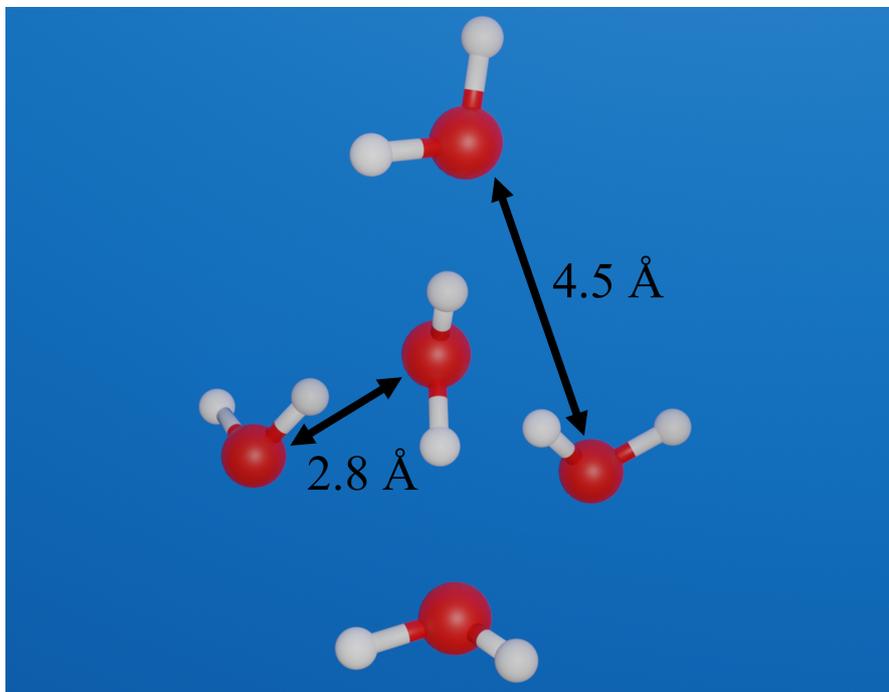


Figure 1.3.2: The tetrahedral structure of water arising from its hydrogen bonding interactions with its neighbours. The average distance from a water molecule to another water molecule in its first hydration shell calculated to be is 2.8 Å, and the average distance from a water molecule to another water molecule in its second hydration shell calculated to be is 4.5 Å [100].

The tetrahedral nature of water is well examined experimentally through neutron and x-ray scattering [100–102]. These techniques allow for the production of site - site radial distribution functions (RDFs or $g(r)$ s). Radial distribution functions describe the local density of a particular atomic species normalised to the bulk density as a function of distance from a central atom species. In the case of a perfectly crystalline solid this would therefore yield a series of evenly spaced sharp narrow peaks of constant intensity, however in liquids there is only short range order as a result of steric hinderance and weak intermolecular interactions, hence the RDFs will appear as a series of broader peaks of diminishing intensity with increasing distance. An example of the water oxygen - water oxygen ($O_w - O_w$) RDF calculated through neutron diffraction at 25 °C and ambient pressure is displayed in figure 1.3.3. Here the peaks corresponding to the first

and second hydration shells around a central molecule occur at approximately 2.8 and 4.5 Å respectively. This is indicative of the tetrahedral structure as it can be shown through trigonometry that for four edge points (E) arranged into a perfect tetrahedron with a single point at the centre (C) (and therefore an ECE angle of 109.5 °), the ratio of the distance between two edge points and the distance from the central point to an edge point is equal to $\sqrt{8/3}$. This is in close agreement to the ratio of the second hydration shell and first hydration shell distance.

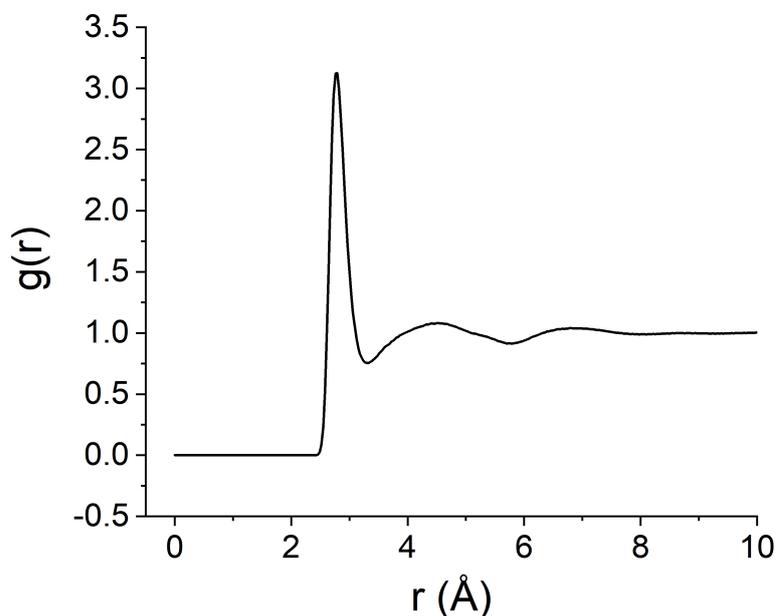


Figure 1.3.3: The $O_w - O_w$ RDF for liquid water at 25 °C and ambient pressure calculated through neutron scattering.

1.3.1.3 The Two-State vs The Continuum Model of Liquid Water

The various anomalous properties of water that arise from its hydrogen bonded network, such as its maximum density at 4 °C, high heat capacity, and low compressibility to name a few [31, 75], mean that there have been several suggested models that describe its structure [103]. Broadly speaking, these can be classed into two groups of models: the two state model of water structure, and the continuum model of water structure. The two state model, originally introduced by Wilhelm Röntgen in 1892 [104], describes

water as an equilibrium mixture between more highly ordered low density tetrahedrally hydrogen bonded domains of water molecules, reminiscent of the structure of hexagonal ice, dispersed in a matrix of higher density less ordered more weakly hydrogen bonded molecules [105]. By comparison the continuum model initially put forward by Bernal and Fowler [106] describes water as a fully interconnected network of water molecules interacting through hydrogen bonds of various strengths and geometries.

Experimental evidence does exist which supports the idea of the two state model of liquid water. Liquid water can be supercooled such that it remains liquid below its typical freezing temperature by slowly cooling ultrapure water [107]. In principle water can exist as a liquid to approximately 150 K, however this is difficult to observe experimentally as homogeneous nucleation usually takes place and the liquid freezes [31]. In supercooled water we observe that as temperature is reduced towards $-45\text{ }^{\circ}\text{C}$ that the isothermal compressibility and isobaric heat capacity become very large and seem to diverge towards infinite values. This can potentially be explained by a phase transition between the two states, however it can also be indicative of a spinodal point where liquid water can no longer exist. X-ray scattering experiments on supercooled water droplets were also suggested to point to the existence of a liquid-liquid transition in water, consistent with the two state model [108], however this has been subsequently disputed [109]. Raman spectroscopy and infrared spectroscopy studies, which are sensitive to the frequencies associated with bond stretching and bending in aqueous solutions observe the emergence of isobestic points in the spectra of liquid water with varying temperature [75, 110]. Isobestic points are defined as points at which a series of spectral lines cross, and are typically indicative of a shifting equilibrium between two states. Their existence can therefore be interpreted as the existence of a dynamic equilibrium between the two states of water. The deconvolution of the Raman spectrum of pure water into five Gaussian sub-peaks is also consistent with the two state model [111]. However, the existence of distinct regions of high and low density water described through the two state model would require density fluctuations in the liquid equal to approximately 25% compared with the bulk density, where experimental evidence from neutron scattering suggests that density variations in liquid water are on the order of 4% [112], and fails

to capture significant areas of low or high density. Recently Shi *et al.* used neutron scattering to suggest why such a disparity in the literature may exist and show that, while they are able to capture characteristics of a high density (HDL) and low density (LDL) phase of liquid water, that “LDL and HDL are the macroscopic phases of water that can exist only below the second critical point, so it is not appropriate to regard liquid water as a mixture of LDL and HDL, which is widely used in the literature” [113].

This lower density fluctuation is more consistent with the continuum model proposed by Bernal [106]. This is also well supported by molecular dynamics studies by Swiatla-Wojcik *et al.* [93], who tested a series of hydrogen bonding definitions between water molecules and demonstrated a smooth transition between hydrogen bonded clusters of water molecules. A similar study by Kumar *et al.* also tested several different hydrogen bonding definitions [88] and observed in all instances that water is better described by a continuum model than a two state model. Bandyopadhyay *et al.* [114] also used molecular dynamics to study variations in the void space (the space between the first and second hydration shells around a water molecule) of water molecules with different numbers of hydrogen bonded neighbours. This showed that tetrahedrally bonded molecules have a larger void space than water molecules with more neighbours, but that increasing temperature caused a smooth transition between the two regimes, consistent with the continuum model. It can also be demonstrated that the isobestic points observed in Raman spectroscopy, and thought to be indicative of a two state model, can be explained by a continuum model, provided that the fluctuations in temperature and pressure are less than 100 K and 10^4 atm [75]. A compromise between the two models is suggested by Nucci *et al.* [50] through temperature excursion infrared spectroscopy measurements on aqueous salt solutions. This suggests that all water molecules are hydrogen bonded and there is an equilibrium between water molecules with relatively linear hydrogen bonds and more bent hydrogen bonds. The applications of different externalities (pressure, temperature, solute addition, etc.) then shifts this equilibrium. This gives rise to a bimodal distribution of hydrogen bonding angles centred at roughly 12° and 50° . This corresponds well with the results of Kumar *et al.* [88].

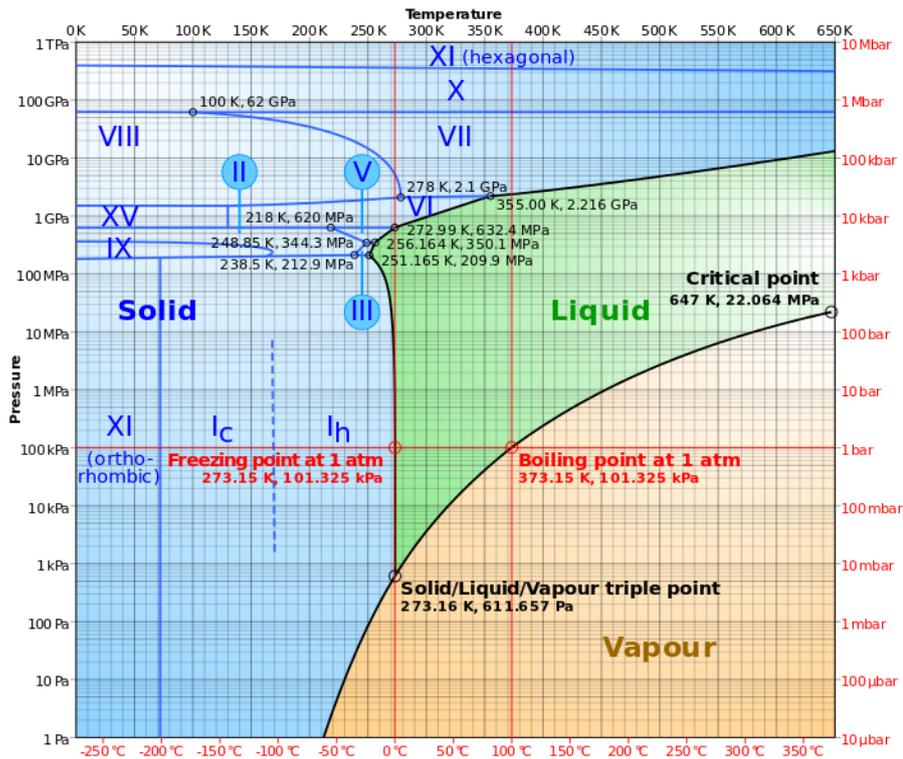


Figure 1.3.4: The phase diagram of water. Figure taken from reference [118].

1.3.1.4 Pressure and Temperature Variation

Regardless of which model one chooses to employ, perturbations to water structure as a result of different externalities are well observed through a variety of scientific techniques [75, 100, 115–117]. The nature of water can be described by its phase diagram, presented in figure 1.3.4 [118]. Here we observe several key features about pressure and temperature variation of water. Firstly, water can exist as either a solid, liquid, gas. In this thesis we limit ourselves to discussion of the liquid state of water. At ambient conditions (1 atmosphere of pressure), water exists in its liquid state between 0 °C (273.15 K) and 100 °C (373.15 K). However, with increasing pressure, water can remain in liquid state until its critical point at 647 K and 22.064 MPa, beyond which the liquid and vapour states are indistinguishable [28]. At low temperatures water will form solid ice, however the exact crystal structure of this ice can take a great many forms depending on the pressure conditions. These are also illustrated in figure 1.3.4.

Arguably the easiest perturbation to water structure to monitor experimentally is variations of temperature. Increasing temperature, unsurprisingly, has been shown to decrease the tetrahedrality of the hydrogen bonded network as more thermal energy is available to overcome hydrogen bonding between water molecules [93, 110, 111, 119]. As such water at higher temperatures exhibit fewer and weaker hydrogen bonds per molecule than water at lower temperatures [75, 93]. It is also shown to decrease the structural correlations as a function of distance, particularly beyond the third hydration shell [115, 120]. These perturbations are well observed through a variety of techniques including: neutron scattering [115], dielectric spectroscopy [75, 121], NMR [75, 122], molecular dynamics (MD) [93, 102], and x-ray scattering techniques [102, 122]. This temperature increase is coupled with a decrease in density and an increase in the distance between water molecules [119, 122].

An externality which is more difficult to access experimentally, but is equally important to the phase diagram of water, is pressure. Hydrogen bonding between water molecules results in it having an unusually low compressibility [31, 112], however its structure can still be perturbed by the application of sufficiently high pressures. The primary effect of pressure application is a compaction of the second hydration shell into the first, which has been observed by both neutron [115] and x-ray scattering [123]. X ray scattering data of liquid water at pressures from 1 bar up to 7.7 kbar by Okhulkov *et al.* are presented in figure 1.3.5. Here the collapse of the second hydration shell into the first is reflected by a gradual decrease in intensity and inward movement of the second peak in the water - water RDFs. MD simulations have shown that this results in an unchanged position of the first peak in the $O_w - O_w$ RDF due to steric effects, but that the orientational order of the water molecules in the first hydration shell is significantly diminished as the first hydration shell coordination is increased [64]. At sufficiently high temperatures as pressures, the compaction of water structure along with the increased thermal energy available to water molecules means that water begins to behave increasingly like a non-interacting conventional liquid [31, 124].

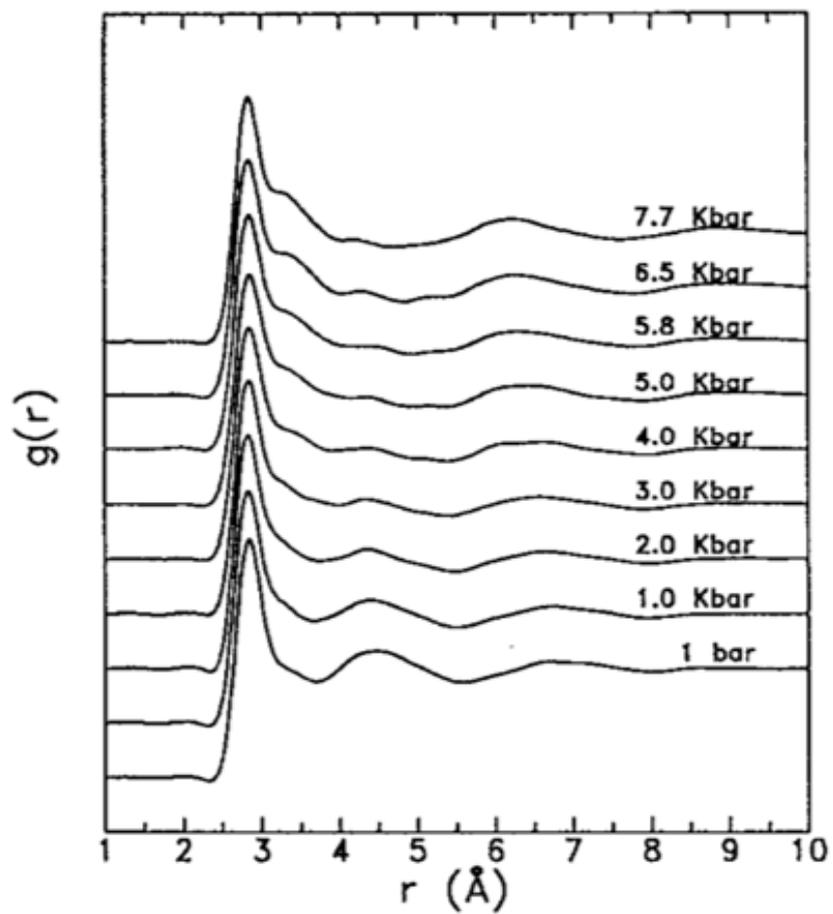


Figure 1.3.5: The water - water RDFs calculated by x-ray scattering at various pressures. RDFs vertically shifted for clarity. Figure taken from reference [123].

1.3.1.5 Salt Addition

As discussed in section 1.1.1, the introduction of solute particle into water means that the water network must reorganise so as to accommodate the particle [31, 125]. One of the most important and prevalent solute species to life on Earth are salts, which dissolve into their constituent ions in aqueous solution. The organisation of the water molecules around a charged ion is dependent on the sign of the charge (negative or positive), the magnitude of the charge, and the size of the ion. In the case of negatively charged anions, water molecules will achieve favourable enthalpic interactions by orienting a positively charged hydrogen towards a central anion, and in the case of positively charged cations they will tend to orient both hydrogens away from the cation. [126–129]. One of the key observations from the study of the structure of aqueous salt solutions is that the effect of ions on water structure is very similar to the effects of increasing pressure, namely an inward movement of the second peak in the O_w-O_w RDF [130, 131]. For example, a neutron study by Winkel *et al.* on aqueous LiCl suggested that this pressurising effect at a molar ratio of 1 LiCl ion pair to 40 water molecules resulted in a structural perturbation to water equivalent to the application of approximately 3 kbar [132]. A more extreme example of this, $Mg(ClO_4)_2$, at a concentration of 44 wt% can perturb water in a similar manner to the application of approximately 3 GPa. This is well illustrated by the neutron scattering study of Lenton *et al.* [133]. In figure 1.3.6 we report the spatial density functions (SDFs) for pure water at ambient conditions, for pure water at 400 MPa, and for pure water in 44 wt% $Mg(ClO_4)_2$. SDFs can be thought of a 3D visualisation of RDFs, and not only show the distances at which neighbouring water molecules can be found, but also their orientation with respect to a central water molecule. The yellow isosurfaces in figure 1.3.6 correspond to the 30% most likely positions for neighbouring water molecules around a central water molecule. Here we observe with increasing pressure or $Mg(ClO_4)_2$, that the second hydration shell begins to compress into the first hydration shell. This particular salt will be studied in much greater detail in chapters 5 and 6.

Broadly speaking, the effects of ions on the structure of water depend on the charge density of the ion, that is the charge of the ion per unit volume. Simple monovalent

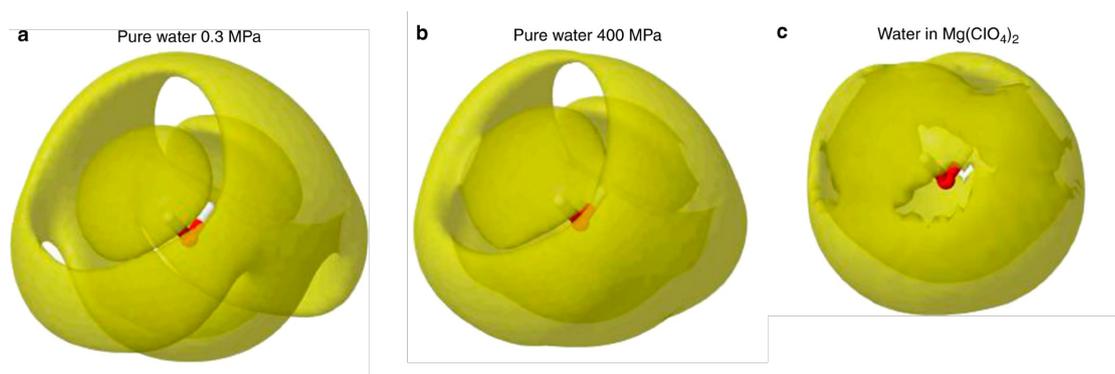


Figure 1.3.6: The water - water SDFs calculated by neutron scattering for pure water (a), water at 400 MPa (b), and water with 44 wt% $\text{Mg}(\text{ClO}_4)_2$. The yellow isosurface contains the 30% most likely positions for neighbouring water molecules around a central water molecule. Figure taken from reference [133].

ions such as Na^+ , K^+ , Cl^- , etc. all result in a modest perturbation to the tetrahedral structure of water as the ions are accommodated [126]. Here it is observed that the lower charge density cation K^+ shows a less well defined hydration structure compared with Na^+ , indicated by the heights of the first and second peaks in the ion - O_w RDFs and the distribution of dipole orientations of the surrounding water molecules. This also results in a reduced pressure like effect, with the second hydration shell of the O_w - O_w RDF being shifted inwards by lesser amounts in aqueous KCl than NaCl at equivalent concentrations [127]. The same is observed for water hydrating monovalent anions by studying the aqueous potassium halides KF, KCl, KBr, and KI. The hydration structure monitored through the ion - O_w RDFs and distribution of hydrating water dipole orientations becomes less well defined as the anion size increases as one moves from F^- to I^- [128]. These particular salts will be investigated in much more detail in chapter 3. Infrared spectroscopy studies on aqueous NaCl, NaBr, and NaI also suggested weakened structural correlations between water molecules and the anion with increasing anion size [134]. These are well reflected by thermodynamic measurements which suggest that the enthalpy of hydration becomes less negative with increasing ion size [27, 35]. Interestingly, opposite trends to increasing pressure like perturbation with increasing ion charge density are observed in the case of aqueous LiOH, NaOH, and KOH when compared with NaCl and KCl. Here the pressure like perturbation was shown to be greatest in the case of the KOH, rather than LiOH [131]. As all ions show a pressure

like effect, and this effect is relatively similar between similarly charged ions, it would seem that the specific order in which ions induce pressure like structural perturbations is dependent of the species of the counterion. In the case of these two studies the difference is likely due to the fact that OH^- ions can form strong hydrogen bonds due to the dipole moment on the ion, whereas Cl^- interacts comparatively weakly with surrounding water molecules [126, 127, 135].

In the case of extremely charge dense ions such as H^+ , which is a bare proton, the pressurising effect can become so severe that not only does the second peak in the O_w - O_w RDF shift inwards, but the first one does as well [136]. The effects of increasing charge density of the ion on water structure are also well demonstrated by using x-ray diffraction to study aqueous solutions containing divalent ions. Collins *et al.* examined perturbations to water structure due to Li^+ , Na^+ , K^+ , Mg^{2+} and Be^{2+} cations and F^- and Cl^- counterions [135]. Here the divalent cations show much more tightly bound water molecules, and that this binding is tighter with decreasing ion size.

1.3.1.6 Addition of Other Solute Species - Alcohols, Amino Acids, and Osmolytes

Other solute species induce similar changes to water structure. It has been shown through neutron scattering that the simple amino acid glycine forms hydrogen bonds with surrounding water molecules through its amine and carbonyl groups, and that the amine - water hydrogen bonds are likely weaker than the carbonyl - water hydrogen bonds [1]. Similar results are also observed for neutron scattering studies of aqueous single amino acids like proline [137], and glutamine [138, 139], as well as dipeptides [140], and are also well reflected by results from machine learning [97]. Urea is another well researched osmolyte due to its role as a protein denaturant, and has been shown to form favourable hydrogen bonds with surrounding water molecules, stronger than water - water hydrogen bonds, while leaving the bulk structure of water relatively unperturbed [62, 63, 141, 142]. This is demonstrated by the O_w - O_w RDFs for pure water, aqueous urea at a mol fraction of 0.05, and aqueous urea and TMAO at a mol fraction of 0.05

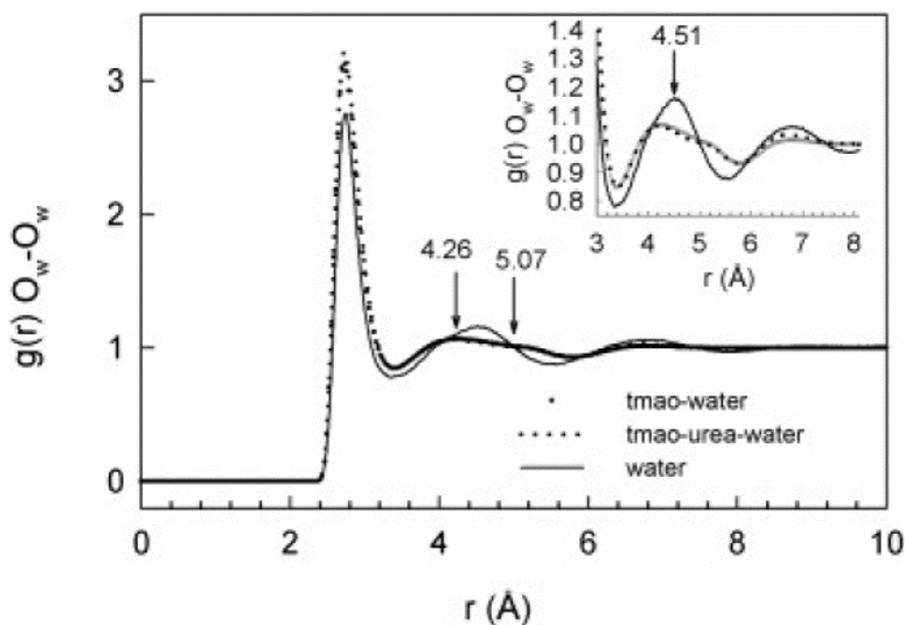


Figure 1.3.7: The $O_w - O_w$ RDFs in pure water, aqueous urea, and aqueous TMAO at a solute mole fraction of 0.05 and a TMAO : urea ratio of 1 : 1. Figure taken from reference [62].

and a urea : TMAO ratio of 1 : 1 calculated through neutron scattering by Meersman *et al.*, as shown in figure 1.3.7. Liquid - liquid hydrogen bonding has also been researched through neutron scattering studies on aqueous alcohols. Here it is shown that while both methanol and ethanol are capable of forming hydrogen bonds with water through their OH group, that the two liquid species tend to segregate, resulting in water rich and alcohol rich regions, as shown in figure 1.3.8. This segregation is observed to be increased with reducing temperature [143, 144].

A particular solute which is of core interest to the study of extremophiles and this thesis is TMAO, as pictured in figure 1.3.9. As mentioned in section 1.2, this protecting osmolyte is utilised by extremophiles to resist the negative consequences of their environment [57, 86, 145]. For this reason its effects in aqueous solution under ambient conditions [70, 146] and in the presence of other externalities, such as pressure [64, 147] and urea [62, 66], have been well studied. Various experimental methods, including neutron scattering [2, 3, 62], molecular dynamics [70], x-ray spectroscopy [142, 146], and dielectric spectroscopy [148] have identified that TMAO enhances the strength of hydrogen bonding between water molecules. This is likely the origin of its ability to

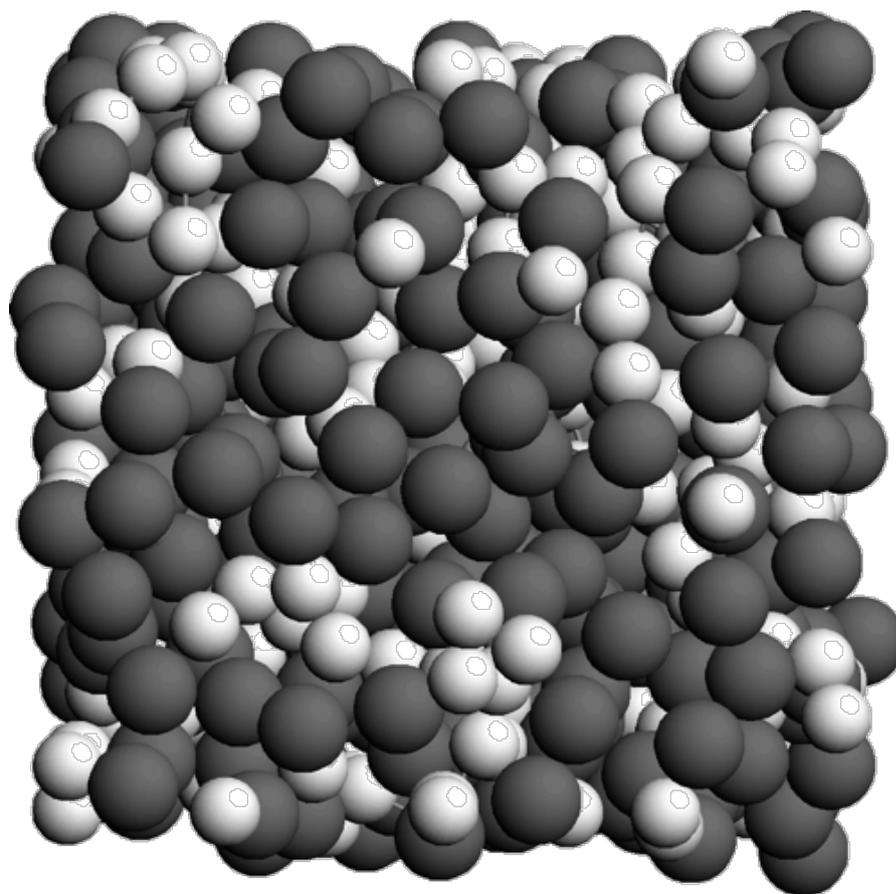


Figure 1.3.8: The structure of a mixed water (white spheres) - methanol (dark spheres) calculated through neutron scattering and computer modelling. The two liquids segregate into methanol rich regions and water rich regions, while remaining more mixed in others. Figure taken from reference [144].

stabilise biomolecules against denaturing externalities. This occurs through the formation of a more ordered structure and strong hydrogen bonds between the TMAO oxygen and approximately three neighbouring water molecules [64, 66, 70, 146, 149, 150] as a result of the large dipole moment across the NO bond [148]. The formation of these strong hydrogen bonds occurs while leaving the water structure relatively unperturbed [2, 62, 66, 72].

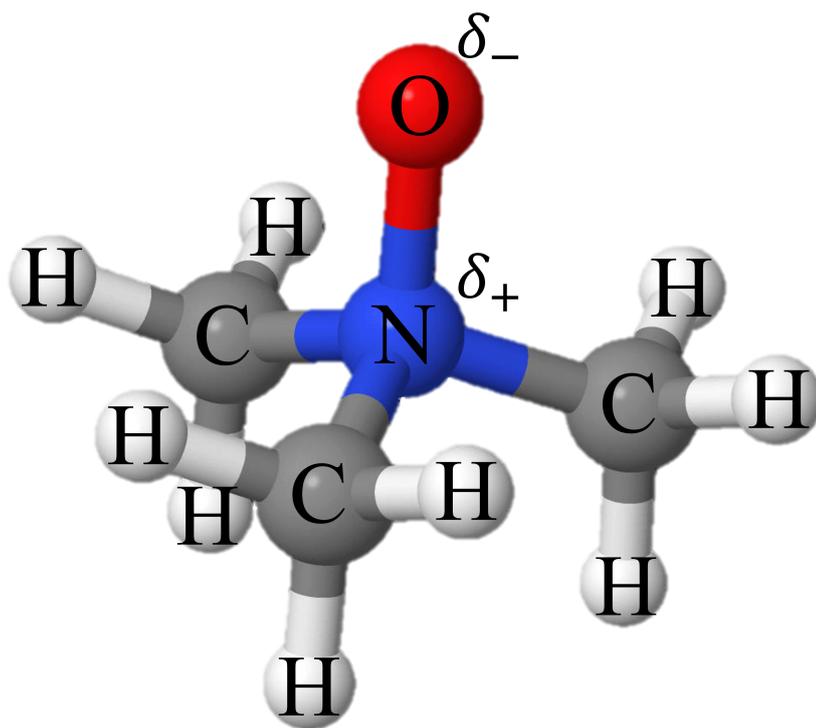


Figure 1.3.9: A schematic illustration of the TMAO molecule with nitrogen (N), oxygen (O), carbon (C) and hydrogen (H) atoms labelled. A large dipole moment exists on the TMAO molecule due to the localised negative charge δ_- on the oxygen and the localised positive charge δ_+ on the nitrogen.

Studies of aqueous TMAO under high hydrostatic pressure have demonstrated that while the structure of water is sensitive to pressure, with the second hydration shell collapsing into the first [115, 123], the hydration structure around the hydrophilic TMAO oxygen and the hydrophobic TMAO methyl groups are much less sensitive [64, 72, 147]. The main effect of pressure around the TMAO molecule is an overall compaction of water structure. This results in an increase in coordinated water molecules around

the hydrophobic groups [64, 72, 149], and an increase in coordinated water molecules around the hydrophilic group while maintaining strong hydrogen bonds, where four fold coordination becomes to become more common than ambient three fold coordination [147, 149, 150]. This relatively rigid hydration structure results in a less compressible overall solution [151] and suggests TMAO acts to resist large external pressures. This hypothesis will be investigated in much more detail in chapter 4.

A pressure resisting affect by TMAO has also been observed in aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$, which perturbs water structure in a similar manner to a large external pressure [2, 3]. Here water - water hydrogen bonding was shown to be more stable in the presence of TMAO at constant $\text{Mg}(\text{ClO}_4)_2$ concentration. These results will be discussed in much more detail in chapter 5. The competing effects on biomolecule stability by TMAO and urea mean that this tertiary solution has also been well studied [62, 63, 142, 152]. These studies suggest that urea and TMAO interact through both direct hydrogen bonding between the urea amine groups and the TMAO oxygen, as well as hydrophobic style association of urea around the TMAO methyl groups. It is therefore easy to conclude that the stabilisation of biomolecules by TMAO against urea is as a result of their direct interaction, leaving less urea available to interact with the biomolecule and cause denaturation. However, this conclusion must be taken with a pinch of salt, as TMAO and urea are thought to act on biomolecules through opposite mechanisms. Urea is thought to enrich at the protein surface [153, 154] whereas TMAO is thought to be excluded from it [68, 69, 73]. Therefore any statements regarding how TMAO and urea perturb biomolecule stability originating from studies that do not contain a model biomolecule must be carefully evaluated.

1.3.2 Monitoring Perturbations to Water Dynamics

1.3.2.1 The Extended Jump Model

As the structure of liquid water is constantly changing, it is equally important to consider the impact of different externalities of the dynamics of liquid water. In general dynamics

in liquid water can be described through the extended jump model proposed by Laage and Hynes [26, 155]. This states that water molecule dynamics consist of two components, large angular jumps as a water molecule switches its hydrogen bonding partner through a jump angle of approximately 68° called jump reorientation, and diffusion of pairs of water molecules while maintaining their hydrogen bonds called frame reorientation. The former relates to rotational motion of the water molecules in solution and the latter is related to the solution viscosity. This jump orientation proceeds through an intermediate conformation called the “bifurcated hydrogen”, where the switching water molecule will temporarily form a hydrogen bond with two neighbouring water molecules through a single hydrogen. A schematic illustration of this bond switching event taken from snapshots of molecular dynamics simulations by Laage *et al.* [155] is shown in figure 1.3.10. This is explained in greater detail in section 2.3.2.

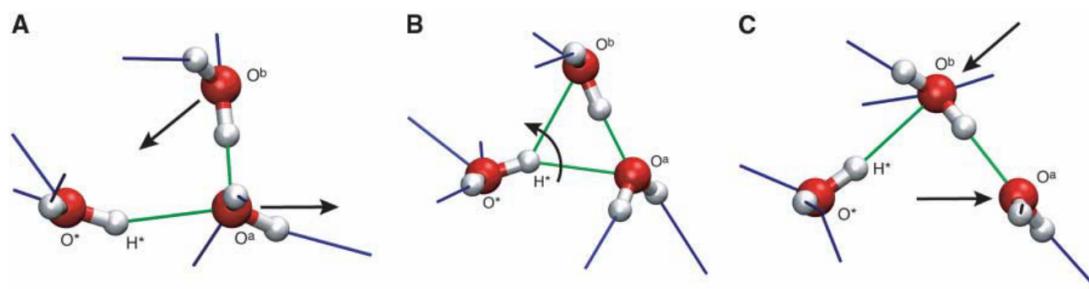


Figure 1.3.10: A series of screenshots from molecular dynamics simulations of a central water molecule switching its hydrogen bonding partner. This involves three neighbouring water molecules. Hydrogen bonds between the three water molecules are shown in green and hydrogen bonds between other water molecules in the simulation that are not shown are shown in blue. In A, a water molecule(*) is hydrogen bonded to another water molecule(a). In B, it then switches partners through the formation of a bifurcated hydrogen, where hydrogen H^* is shown to hydrogen bond to both O_a and O_b . In C, the hydrogen bond between hydrogen H^* and O_a is terminated and water(a) diffuses away. Figure taken from reference [155].

1.3.2.2 Pressure and Temperature Variation

Unsurprisingly, increasing temperature results in increased dynamics of water molecules due to the greater availability of thermal energy. This is reflected by a decreasing viscosity, increased rotational motions, increased self-diffusion coefficient with increasing temperature [75]. By comparison pressure variation on water dynamics is more complex.

It is experimentally observed at low temperatures (275.4 K) that the viscosity of water decreases with increasing pressure up to 100 - 200 MPa, before increasing again. At high temperatures (373.2 K) the relationship is much clearer, and viscosity increases linearly with pressure. Similar trends are observed for the self diffusion coefficient, where the self diffusion coefficient does through maximum between 100 - 200 MPa before decreasing again with increasing pressure, but at high temperature the self diffusion coefficient decreases linearly with increasing pressure. These results are displayed in figure 1.3.11. This relationship is explained by competition between two factors: increasing pressure results in increased overlap of close-contact electron orbitals, which weakens the hydrogen bonding structure and promotes flow. However, the compressed water structure means that the local fluctuations in density responsible for hydrogen bond switching events [26] are reduced, and jump reorientations occur less frequently, resulting in slowed dynamics [124].

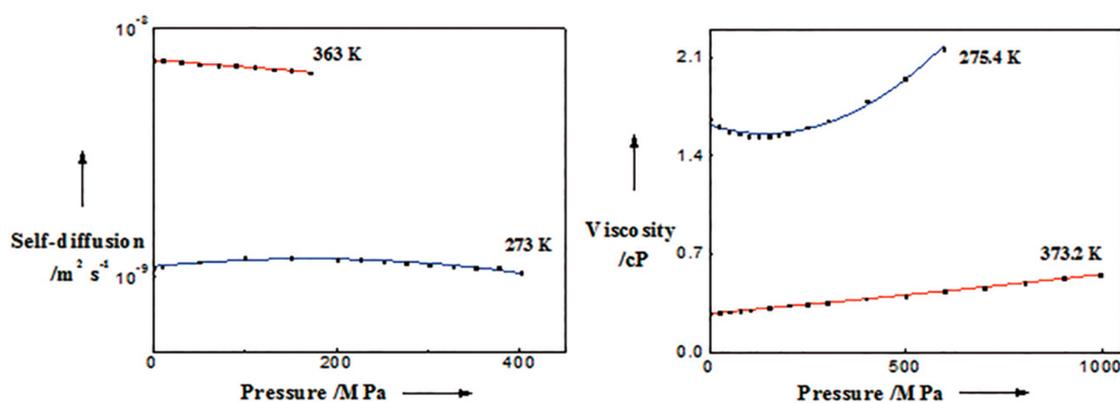


Figure 1.3.11: The variation of the self-diffusion coefficient of water (left) and the viscosity of water (right) as a function of pressure at temperatures of 275.4 K and 373.2 K. Figure taken from reference [75].

1.3.2.3 Salt Addition

The structural organisation of water around charged ions is reflected in the dynamic perturbations to water, and is again sensitive to ionic charge density. Perturbations to water dynamics in the vicinity of solutes can be thought of as the sum of two components: the strength of favourable electrostatic between the solute and the water molecules rela-

tive to the strength of water - water interactions, and the excluded volume effect of the solute [156]. In the case where the interactions between the solute and the surrounding water are stronger than that between water molecules, it means that a water molecule is in a more enthalpically stable state and therefore will be less likely to undergo a jump rotation or diffuse away from the solute. The excluded volume also effect reduces jump reorientations by reducing the likelihood that a water molecule will be appropriately positioned for a hydrogen bond switching event to occur. This is well demonstrated by viscosity measurements of aqueous solutions containing Zn^{2+} , Cu^{2+} , Fe^{3+} , and Cr^{3+} , where Ma *et al.* demonstrated that the solution viscosity decreased with increasing ionic charge density [117]. At high concentration, all ions are shown to retard water dynamics [25], however at lower concentrations the effects of ions on water dynamics are more varied. In the case of single ions, low charge density ions are shown to accelerate dynamics and high charge density ions are shown to retard dynamics [87, 157]. For example, Kim *et al.* observed a decreasing self diffusion coefficient for water in aqueous NaCl, but an increasing self diffusion coefficient in aqueous CsI [158]. It is also worth noting here that ions may perturb the two components of the extended jump mechanism differently. For example, LiCl is observed to retard both the jump and frame reorientations, resulting in reduced overall dynamics. However, NaBr is shown to marginally accelerate the jump reorientations while drastically slowing the frame reorientations. Hence, despite increased rotational motion, the overall viscosity is increased and dynamics are reduced. The opposite is true of NaI, where rotational motions are increased with a moderate increase in viscosity, resulting in accelerated overall dynamics [157]. These results are displayed in figure 1.3.12.

The simple rule of thumb that high charge density ions retard dynamics and low charge density ions accelerate dynamics is no longer applicable once one begins to consider larger polyatomic ions. 2D infrared spectroscopy measurements demonstrated that the time constant associated with total randomisation of the hydrogen bond network was increased from 1.7 ps to 4.8 ps in 6 M NaBr [159], consistent with NMR measurements [157]. This was also compared with the time constant for aqueous NaBF_4 . The absolute charge on the BF_4^- ion spread over 5 atoms would suggest that it has very low

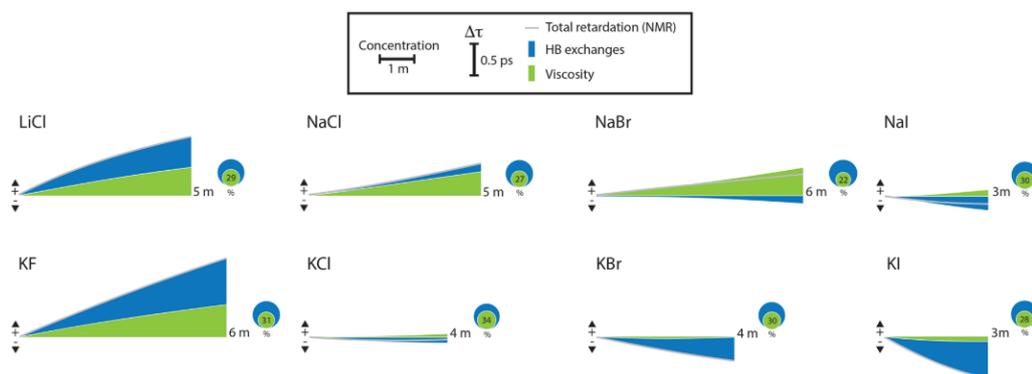


Figure 1.3.12: The relative changes in the jump reorientation (blue) and the frame reorientation (green) time in aqueous sodium and potassium halide salts. Slowed dynamics are shown by increased values in the + direction and *vice versa*. The circles represent the relative contributions of each component at the maximum studied concentration. Maximum studied concentration displayed beneath circles. Figure taken from reference [157].

charge density, however 6 M NaBF_4 was shown to increase the time constant to 6 ps, suggesting slower dynamics than in NaBr at equivalent concentration [159]. Conversely the similarly charged polyatomic ion ClO_4^- was shown to accelerate water dynamics [160, 161]. It is suggested that a key feature to consider for polyatomic ions that is not as important when one considers dynamic perturbations to water by single ions is polarisability effects. A particularly interesting example of aqueous polyatomic ions comes from the 2D infrared spectroscopy study of van der Post *et al.* [162], who studied aqueous tetraalkylammonium ($\text{N}(\text{alkyl})_4^+$) ions. These are ions that have an absolute charge of -1, but contain four variable length hydrocarbon alkyl chains covalently bonded to a central nitrogen atom. Here it was observed that increasing the excluded volume effect by increasing the alkyl chain length resulted in slowed dynamics with increasing chain length. This is a particularly important observation as the hydrophobic effect, as discussed in section 1.1.1, is shown to originate from unfavourable entropic interactions between a hydrophobic solute and water. This reduced dynamics around a hydrophobic entity therefore contributes to an increased entropy penalty, as reduced dynamics means that water molecules are less able to explore their available conformations.

1.3.2.4 Addition of Other Solute Species - Alcohols, Amino Acids, and Osmolytes

The different dynamics induced in surrounding water molecules by hydrophilic and hydrophobic entities can be studied using molecules that contain both hydrophilic and hydrophobic moieties, such as amino acids. In general it is observed that dynamics of hydrating water molecules are reduced around hydrophobic surfaces due to an excluded volume effect of the surface, but the added favourable enthalpic interactions between water molecules and a hydrophilic surface reduce dynamics yet further [163, 164]. As was the case for ions, the relative decrease in dynamics of the hydrating water molecules around hydrophilic groups is dependent on the group species. Sterpone *et al.* used molecular dynamics to investigate aqueous amino acids and determined that water rotational dynamics are slowed by a factor between 1.1 - 1.4 around the amine group, but slowed down by a factor of up to 3 around carbonyl groups [165]. This is also shown in other hydrogen bonding species such as aqueous alcohols [166] and urea [167], where solute - water hydrogen bonds are observed to be stronger than water - water hydrogen bonds, resulting in a retarded water rotational and diffusive dynamics as a function of solute concentration. The relative perturbations to dynamics has been shown to be crucially important to thermodynamic parameters. Okouchi *et al.* showed through ^{17}O NMR that the entropy of hydration of a particular solute correlates well with the “dynamic hydration number” (DHN), as shown in equation 1.3.1 where n_h is the coordination number of the species, τ_c^h is the rotational correlation time of a water molecule in the hydration shell of the species (roughly equal for the time taken for the average angular deflection of a water molecule to be one radian [168]), and τ_c^0 is the rotational correlation time of pure water. This observation is a particularly nice example of the importance of considering both structural and dynamic perturbations to water by solute addition [169]. This correlation between DHN and entropy of hydration is also observed in aqueous alcohols [170].

$$\text{DHN} = n_h \left(\frac{\tau_c^h}{\tau_c^0} - 1 \right) \quad (1.3.1)$$

The relative slowdown of water rotational motion in aqueous TMAO was well characterised in a combined molecular dynamics and terahertz spectroscopy study by Imoto *et al.* [148]. Here it is observed that the rotational relaxation time (related to but not equal to the rotational correlation times) for pure water is 1.76 ps. In aqueous TMAO the rotational relaxation time is retarded around the hydrophobic groups to 2.18 ps, and retarded yet further around the hydrophilic oxygen to 2.87 ps. This calculated through the rotational correlation function, which quantifies how predictable the orientation of a particular water molecule will be as a function of time. At time = 0 it therefore has a value of 1 and as time approaches ∞ it decays towards 0. Faster dynamics therefore correspond to a steeper gradient in $C_{rot}(t)$ and *vice versa*. This data is shown in figure 1.3.13. This again demonstrates that the combined effects of favourable enthalpic interactions and excluded volume effects typically result in reduced dynamics of hydrating water molecules around hydrophilic groups compared with water molecules hydrating hydrophobic groups, which have little in the way of favourable enthalpic interactions but maintain the excluded volume effect. These results are in good qualitative agreement with ab initio molecular dynamics simulations by Stirnemann *et al.* [65], who showed a reduction in rotational dynamics of water molecules around the hydrophobic groups of TMAO by a factor of 1.6, but a reduction in rotational dynamics of water molecules around the hydrophilic oxygen by a factor of 4 - 5. Similar results are also reported by the molecular dynamics studies of Usui *et al.* [171] and Xie *et al.* [167]. Interesting the combined study of Imoto *et al.* showed that water was not only retarded in the hydration shell of TMAO, but also in the bulk solution, where the rotational relaxation time slowed from 1.76 to 1.83 ps. This suggests, consistent with the results from various structural studies [2, 3, 70, 142, 146], that TMAO acts to enhance the hydrogen bonding between bulk water molecules. This effect will be studied in much greater detail in sections 4 and 5, and the different range over which solutes perturb water structure and dynamics will be commented on in chapters 3 - 5.

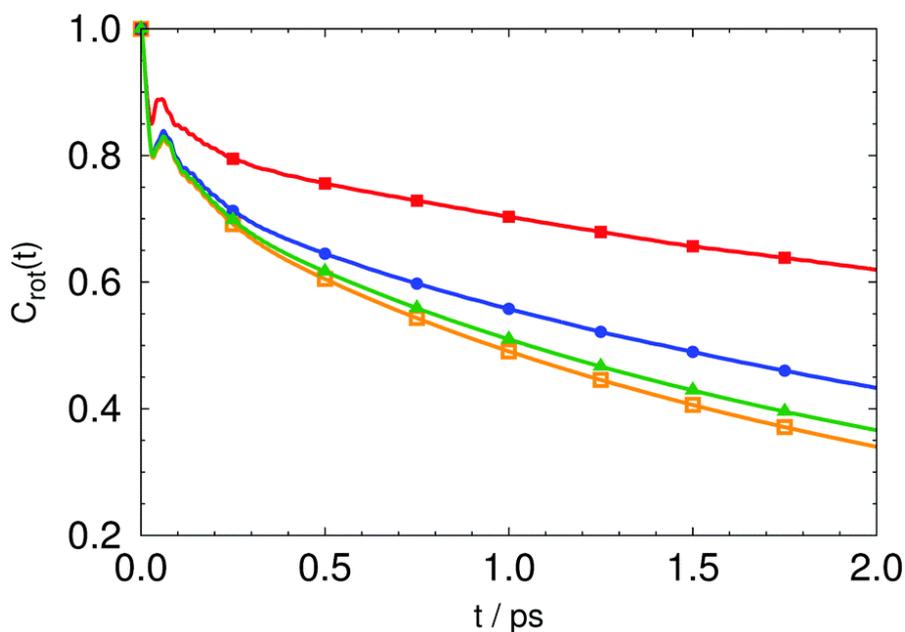


Figure 1.3.13: . The rotational correlation function for water molecules in pure solution (orange), bulk water in aqueous TMAO (green), water around the hydrophobic regions of TMAO (blue), and water around the hydrophilic TMAO oxygen (red). Figure taken from reference [148].

1.4 Aims and Objectives

One could easily spend a lifetime studying the perturbations to water and its interactions with biomolecules under these conditions, however over the course of a four year PhD one is unfortunately forced to narrow one's focus. Hence in this thesis we will primarily concern ourselves with two extreme environments: the high pressure environments present in the deep ocean on Earth [172], and model versions of the subsurface lakes recently discovered beneath the Martian south pole, which are suggested to contain the salt $\text{Mg}(\text{ClO}_4)_2$ [173, 174]. The aims and objectives of this thesis are therefore as follows:

- In chapter 3 we will use a combination of neutron scattering, NMR, and thermodynamics to obtain a detailed understanding of the perturbation to water structure and dynamics as a result of the addition of simple monovalent biologically relevant ions. These are studied as they are present in all known environments, regardless

of additional extreme [11, 175, 176], and their simplicity serves as a good basis on which to build the following results chapters.

- In chapter 4 we will then use neutron scattering to investigate the effects of large external pressure on water structure, and we will determine whether the rigid hydration structure and enhancement to water - water hydrogen bonding previously observed by TMAO addition acts to resist this pressure induced perturbation.
- In chapter 5 we will use neutron scattering and NMR to examine the pressure-like perturbation to water structure induced by $\text{Mg}(\text{ClO}_4)_2$, previously observed by Lenton *et al.* [133], and investigate whether TMAO can also resist this style of perturbation to water structure. We will also discuss the dynamic perturbations to water as a result of the two solutes, and comment on how one can deconvolute the roles of individual solute species in more complex tertiary aqueous solutions.
- Finally, we will consider the effects of $\text{Mg}(\text{ClO}_4)_2$ and TMAO on biomolecules in chapter 6. This is done by examining the ability of the model amino acid to self-associate in aqueous solution containing $\text{Mg}(\text{ClO}_4)_2$ through neutron scattering and measuring the effect of $\text{Mg}(\text{ClO}_4)_2$ and TMAO the Gibbs free energy of unfolding of the model β sheet protein I27.

In this way we will build a complete story that begins with the perturbations to water structure and dynamics by simple ions, investigates the perturbation to water structure and dynamics by extremes of pressure and more complex ions found in the Martian regolith, examines how the osmolyte TMAO already employed by organisms to combat extreme environments affects these perturbations, and finally demonstrate how the two solute species $\text{Mg}(\text{ClO}_4)_2$ and TMAO perturb the interactions between biomolecules.

Chapter 2

Materials and Methods

The ubiquity of liquid water on Earth and its vital relationship with molecular biology, along with countless other areas of scientific inquiry, means that it is difficult to think of an experimental technique that hasn't been applied to its study, as discussed in section 1.3. In this next chapter we will overview the fundamental theory and application of the principle techniques used in this thesis to monitor structural and dynamic perturbations to water, namely: neutron diffraction, computational modelling, and nuclear magnetic resonance. We will first discuss the phenomenon of diffraction and interference in general in section 2.1 and 2.1.1, before going on to discuss how this theory specifically applies to neutron diffraction in section 2.1.3. This will be expanded on further to overview the raw data produced through neutron diffraction, the total structure factor, in section 2.1.4, and its analysis through EPSR and additional analysis routines in sections 2.2.1-2.3.4. We will then go on to overview how nuclear magnetic resonance can shed further insight on aqueous systems by first overviewing the fundamental theory of NMR in sections 2.4.1-2.4.2. This technique can help verify the structural information gleaned through neutron scattering and computational modelling, as discussed in section 2.4.3, and can be used to investigate dynamic perturbations to water, as discussed in sections 2.4.4 and 2.4.5.

These perturbations go on to have significant effects to the structure and stability of biomolecules. In this thesis we will specifically focus on perturbations to the protein I27

and the 10 amino acid peptide CLN025. In sections 2.5.2-2.6 we will describe the production and purification of proteins and peptides, before going on to describe how their structure and stability is experimentally investigated through fluorescence spectroscopy in section 2.7 and circular dichroism in section 2.8.

2.1 Diffraction

Diffraction is a fairly broad term which refers to the the ability of wavefronts to bend around an obstacle or through an aperture. This occurs most effectively when the object/aperture is of comparable size (within a few wavelengths) to the wavelength λ of the incoming wave, as shown below in figure 2.1.1(a). In this case a parallel wave will diffract through the aperture and become a circular wave, and the aperture will effectively act as a new point wave source. However, in the case when the size of the aperture is much larger than the wavelength of the incoming wave, as shown in figure 2.1.1(b) only the part of the wavefront which is within a few wavelengths to the edge of the aperture will diffract and the remainder of the wave will continue unchanged [28, 177].

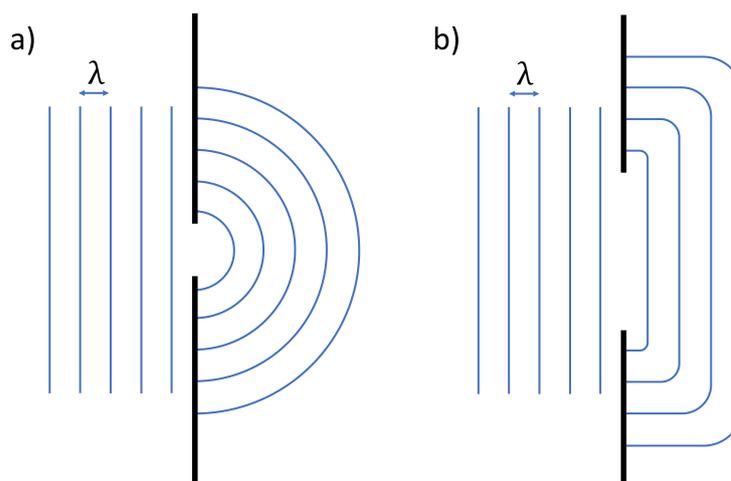


Figure 2.1.1: The diffraction of a parallel wave of wavelength λ passing through an aperture of size comparable to λ (a) and passing through an aperture of size much larger than λ (b).

It is for this reason that one can easily hear someone talking from around the corner of a building, as the wavelength of sound is on the order of metres, and therefore diffracts very effectively around an obstacle like a corner and a large portion of the wave reaches the listener. However, one is unable to see the other person, as the wavelength of visible light is only hundreds of nanometers and therefore does not diffract effectively around the corner, and practically no light reaches the viewer from the source. It is this relationship between the wavelength of the wave and the size of the object being studied that makes diffraction such a useful probe of structure, as we can tune the wavelength of our probing wave such that it is suitable to study structures of a given size. The following sections will explore this in more detail and how it applies to using neutron diffraction to study aqueous solutions.

2.1.1 Superposition

With the theory of how waves diffract around objects or through apertures firmly in hand we can now begin to address what happens when two waves come together to cause superposition. The principle of superposition states that: “when two or more waves overlap, the resultant wave is the algebraic sum of the individual waves” [28]. This is illustrated in figure 2.1.2.

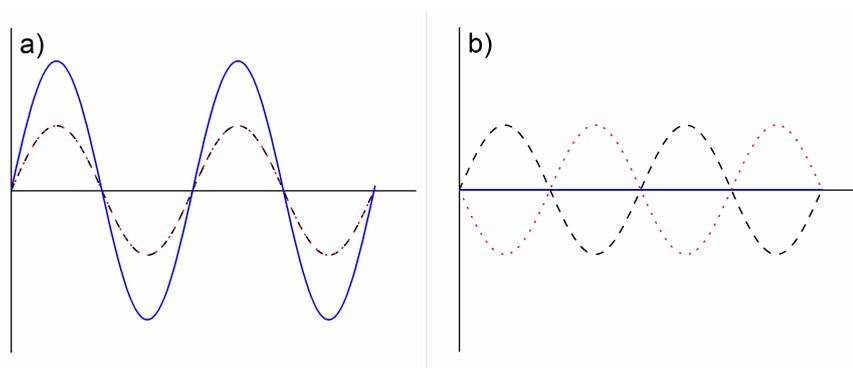


Figure 2.1.2: The algebraic sum (blue) of two individual waves (red dots and black dashes) that are perfectly in phase (a) resulting in constructive interference and perfectly out of phase (b) resulting in destructive interference.

In this figure we observe the two most extreme cases of superposition. In figure

2.1.2(a) we observe that if two waves of equal amplitude and frequency are perfectly in phase (coherent) then the resultant wave is one of twice the amplitude at the same frequency. This scenario is referred to as constructive interference. Conversely, in figure 2.1.2(b), if two waves of equal amplitude and frequency are perfectly out of phase, with a phase difference equal to π , then the resultant wave has 0 amplitude. This scenario is referred to as destructive interference.

We will now demonstrate how the phenomena of interference and superposition can be exploited to reveal structural information about a sample. In the first instance consider a crystal consisting of planes of atoms, shown in red, separated by a distance d , as shown in figure 2.1.3.

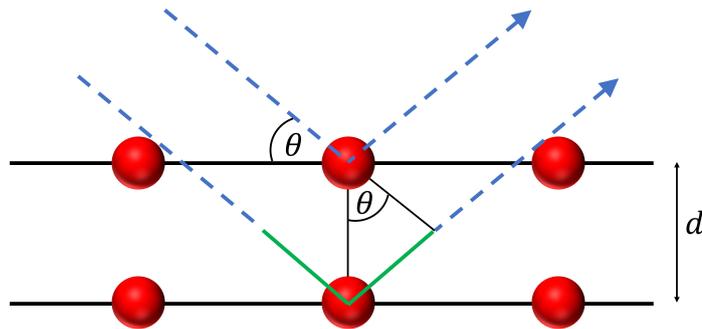


Figure 2.1.3: Scattering of two coherent waves (blue) at an angle θ off planes of atoms (red) separated by a distance d in a crystal. The extra distance travelled by the bottom wave is shown in green.

We now consider two parallel coherent waves, shown in blue, incident on the crystal at an angle θ relative to the atomic planes of the crystal. We observe that the top wave is reflected off the atom in the top plane, and the bottom wave is reflected off the atom in the bottom plane. We also notice that the bottom wave has had to travel an extra distance, called the path difference, shown in green compared to the top wave. We can show using trigonometry that the path difference of the bottom wave is equal to $2d \sin \theta$. We then consider how the two emerging waves will interact with one another. If the path

difference of the bottom wave is equal to an integer number of wavelengths λ , then the two emerging waves will remain in phase with one another and constructive interference will occur. This condition is represented algebraically in equation 2.1.1.

$$n\lambda = 2d \sin \theta \quad (2.1.1)$$

This equation is known as Bragg's law, named after the father and son team who first derived it in 1913 at the University of Leeds [178]. They first observed the phenomenon when directing coherent x-rays at mica crystals through an angle θ , and noticing that the reflected radiation intensity would spike at regular intervals of $\sin \theta$. This therefore allowed them to calculate the spacing between the atomic planes of the mica crystals, and is a technique underpins the field of crystallography to this day. This equation also helps us to understand the importance of tuning the wavelength of the incoming wave to the size of the structural motifs, in this case the distance between planes of atoms, if one wishes to gain meaningful insight into the sample. As an example let us consider molybdenum K_α x-rays ($\lambda = 0.709 \text{ \AA}$) incident on a sodium chloride crystal, where planes of Na and Cl atoms are separated by a distance of 2.81 \AA [28]. Applying Bragg's law we can determine that one will observe the first two instances of constructive interference at scattering angles of 7.25° and 14.61° . This is certainly a plausible experimental setup that would allow molybdenum x-rays to determine the atomic spacing in a sodium chloride crystal. One can now use Bragg's law to explore the upper limit of wavelengths that would be suitable for this experiment. In this case, with a lattice spacing 2.81 \AA , any wavelengths greater than 1.405 \AA would render Bragg's law unsolvable and would not allow for determination of the atomic spacing. Conversely if one were to use gamma rays ($\lambda \approx 10^{-12} \text{ m}$), then the first two instances of constructive interference would occur at scattering angles of 0.1° and 0.2° , which would require large separation between crystal and detector and renders the experiment highly impractical. The short version is this: if you want to look at things separated by distances of \AA , use something with a wavelength of \AA .

A second example of the use of the relationship between superposition of waves

and the structure of the diffracting objects can be found in Thomas Young's double slit experiment [179]. In figure 2.1.4 we observe a plane light wave of wavelength λ diffracting through two narrow slits separated by a distance d . As described in section 2.1, the plane wave diffracts through a narrow aperture resulting in two circular waves, hence each slit essentially acts as a new point source of circular waves. The resultant pair of waves then go on to interfere to create areas of constructive and destructive interference. The areas of constructive interference are shown in red. If a screen is placed in front of the resultant waves the areas of constructive interference will produce bright bands on the screen separated by a distance of $\frac{2\pi}{d}$. We therefore observe a reciprocal relationship between the spacing of the slits and the resultant diffraction pattern. When the slits are separated by a small distance, the resultant diffraction pattern is separated by larger distances, as shown in figure 2.1.4(a), and *vice versa*, as shown in figure 2.1.4(b).

2.1.2 Reciprocal Space

This reciprocal relationship between the structure of the scattering object, in this case a pair of slits, and the resultant diffraction pattern brings us neatly to the concept of reciprocal space, also referred to as k -space or Q -space. To do this we will now describe a wave by its wavevector k , given in equation 2.1.2.

$$k = \frac{2\pi}{\lambda} \tag{2.1.2}$$

The difference between the wave incident on the slits k_i and the total wave emerging from the other slits k_f is the exchanged wavevector Q , shown in equation 2.1.3.

$$Q = k_i - k_f \tag{2.1.3}$$

We can now say that the spacing between the slits is related to Q by:

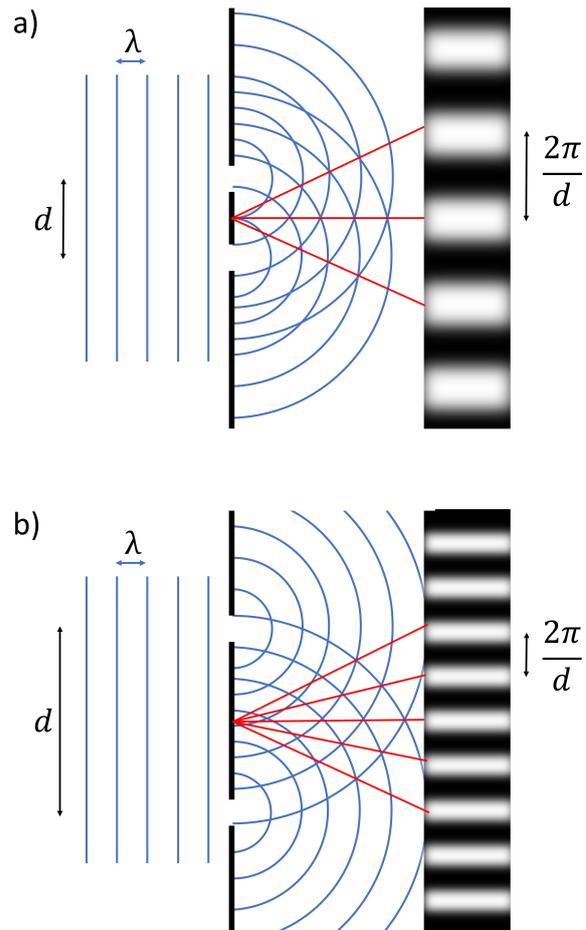


Figure 2.1.4: The diffraction of a wave of wavelength λ through two narrow slits separated by distance d . The resultant diffracted waves go on to interfere and produce an interference pattern on a screen. The interference pattern is a series of bands separated by a distance $\frac{2\pi}{d}$.

$$Q = \frac{2\pi}{d} \quad (2.1.4)$$

One can now use the observations gleaned from the double slit experiment to explain how one moves between the structure of the diffracting object to the eventual diffraction pattern [180]. A pair of slits separated by a distance d can be described mathematically by the aperture function, shown in equation 2.1.5. Here the pair of slits are described as an amplitude $A(x)$ equal to a pair of Dirac δ functions separated by a distance $\frac{d}{2}$ from the centre, where x represents the distance along the screen containing the slits.

$$A(x) = \delta\left(x - \frac{d}{2}\right) + \delta\left(x + \frac{d}{2}\right) \quad (2.1.5)$$

We can move from this equation to the eventual diffraction pattern by the use of a Fourier transform, shown in equation 2.1.6, where $\psi(Q)$ is the diffracted wave function, ψ_0 is a constant of proportionality, and i is the square root of -1.

$$\psi(Q) = \psi_0 \int_{-\infty}^{\infty} A(x) \exp(iQx) dx \quad (2.1.6)$$

Armed with the knowledge that the integral of a Dirac delta function from $-\infty$ to ∞ is equal to 1, we can rewrite equation 2.1.6 as equation 2.1.7.

$$\psi(Q) = \psi_0 \left[\exp\left(iQ\frac{d}{2}\right) + \exp\left(-iQ\frac{d}{2}\right) \right] = \psi_0 2 \cos\left(\frac{Qd}{2}\right) \quad (2.1.7)$$

One does not observe the diffracted wave function, but rather its intensity $I(Q)$, which is given in equation 2.1.8, where $\psi(Q)^*$ is the complex conjugate of $\psi(Q)$.

$$I(Q) = |\psi(Q)|^2 = \psi(Q)\psi(Q)^* \quad (2.1.8)$$

Hence we can finally say that the observed diffraction pattern as a result of a plane

wave diffracting through a pair of narrow slits separated by a distance d is given by equation 2.1.9.

$$I(Q) \propto \left[\cos \left(\frac{Qd}{2} \right) \right]^2 \propto 1 + \cos(Qd) \quad (2.1.9)$$

This predicts that the observed diffraction pattern will be a series of evenly spaced equally intense bands, as observed experimentally, and demonstrates that one can use Fourier transforms of observed diffraction patterns to extract structural details of a diffracting object.

2.1.3 Neutron Diffraction

We will now overview how neutrons can be used in diffraction experiments and why they are a sensible choice for the study of aqueous solutions. To do this we will first review a fundamental concept in physics: wave-particle duality. In 1923 Louis de Broglie [181] demonstrated that particles and waves were not in fact separate entities, but that a particle could exhibit wave like properties, and *vice versa*. The de Broglie hypothesis states that the wavelength of a wave or particle of any sort is inversely proportional to its momentum by a constant of proportionality equal to Planck's constant, as shown in equation 2.1.10.

$$\lambda = \frac{h}{p} = \frac{h}{mv} \quad (2.1.10)$$

With this equation one can make the rather entertaining thought of guaranteeing a strike in bowling, by making the ball diffract through the approximately 1 m opening at the end of a bowling lane like a plane wave and occupying the entire area in which the pins are situated. However, for a 5 kg bowling ball to have a wavelength of 1 m, and therefore diffract through the opening, equation 2.1.10 shows us that it would have to be rolled at $1.33 \times 10^{-34} \text{ ms}^{-1}$, and would therefore take 4.3×10^{27} years to traverse the 18 m bowling lane. In this sense, practice may be more practical than

patience.

While this hypothesis is clearly not useful in everyday life, it becomes vital when one wishes to study the structures on the molecular lengthscale, where structural motifs are on the orders of Å. In this case neutrons become a very practical diffraction probe, as the momentum of the neutron can be carefully tuned such that its wavelength is comparable to the Å scale separations between atoms in a liquid. However, its benefits extend beyond simply having a tunable wavelength, as we will now discuss.

The neutron is a subatomic particle that is found alongside protons in the nucleus of all atoms, with the exception hydrogen whose nucleus comprises of a single proton. It was discovered in 1932 by James Chadwick [182] through the reaction of Beryllium with alpha particles, resulting in the unstable carbon isotope ^{13}C , which then decays into ^{12}C and a single neutron, as shown in equation 2.1.11. Some of the key properties of the neutron, compared with protons and photons, are displayed in table 2.1.1.

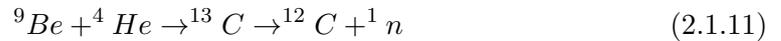


Table 2.1.1: Comparing key properties of neutrons, protons, and photons [28].

Property	Neutron	Proton	Photon
Mass (MeV/c ²)	939.6	938.3	0
Charge (e)	0	+1	0
Spin	$\frac{1}{2}$	$\frac{1}{2}$	1
Classification	Baryon	Baryon	Boson

It is through these properties that we can begin to elaborate on what makes neutrons a useful diffraction probe. They are electrically neutral, which allows them to penetrate deeply into a sample in a way that charged particles like protons cannot, as electrostatic interactions between charged particles and electron clouds around atoms and molecules would cause charged particles to lose kinetic energy very quickly. They have an intrinsic magnetic moment, given by their spin, which means they are sensitive to the magnetic

domains in a sample, however this property is not useful for the investigations covered in this thesis. Finally they are baryons, and therefore interact through the nuclear strong force. This is the same force that holds the nucleus of atoms together, as without it the strong electrostatic repulsion between the positively charged protons would cause the nucleus to fly apart. It also means that neutrons will scatter off the nuclei of atoms present in the sample, resulting in diffraction patterns which can then be related to the structure of the sample through Fourier transforms as described in section 2.1.2. Conversely photons, such as x-rays, will scatter off the electron clouds around the atoms and molecules present in a sample as the photon is the mediating particle for electrostatic interactions. In the next section we will explore neutron diffraction and obtaining the resultant diffraction patterns in rigorous mathematical detail.

2.1.4 The Total Structure Factor

We will begin this explanation by considering a scattering event shown in figure 2.1.5. Here an incident neutron with wavevector k_i is incident upon a sample, undergoes a scattering event and emerges at an angle θ and final wavevector k_f before going into a detector. The exchanged wavevector Q is then the difference between between the initial wavevector and the final wavevector.

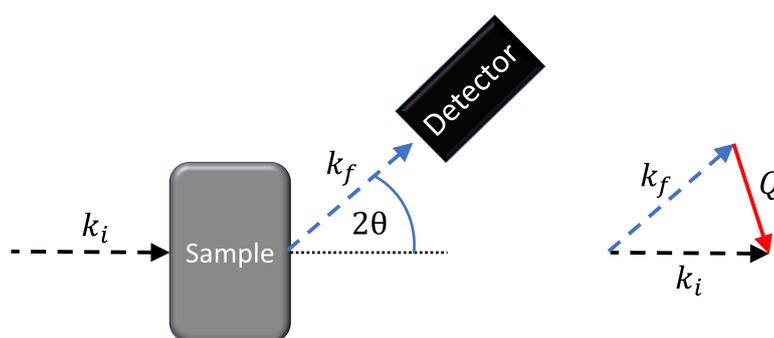


Figure 2.1.5: An incident wave of wavevector k_i (black) scatters off a sample and exits as wavevector k_f (blue) through an angle 2θ into a detector. The exchanged wavevector Q (red) is the difference between the incident and scattered wavevectors.

For a neutron diffraction experiment it is necessary that the scattering event of the incident neutron is completely elastic, which means that no energy is lost by the neutron to the sample, and therefore the modulus of the incident and scattered wavevectors are equivalent, as shown in equation 2.1.12:

$$|k_i| = |k_f| = \frac{2\pi}{\lambda} \quad (2.1.12)$$

In this case trigonometry can be used to demonstrate that the exchanged wavevector Q can be simply expressed in equation 2.1.13

$$Q = \frac{4\pi \sin \theta}{\lambda} \quad (2.1.13)$$

We will now turn our attention to the detection of neutrons, and therefore the actual measurable quantity of a neutron scattering experiment. The following derivations can be found in greater detail than will be covered in this thesis in reference [180]. This measurable quantity is known as the differential cross section $\frac{d\sigma}{d\Omega}$, and relates to the proportion of scattered neutrons per unit solid angle per unit time, and is expressed in equation 2.1.14 where R_{tot} is the rate of neutrons scattering through the solid angle (s^{-1}) and in the case of purely elastic scattering is equal to R_{el} , N is the number of scattering centres, and Φ is the incident flux ($s^{-1}m^{-2}$)

$$\frac{d\sigma}{d\Omega} = \frac{R_{tot}}{N\Phi d\Omega} \quad (2.1.14)$$

Let us now examine elastic scattering of a beam of neutrons of wavelength λ by a single fixed atom in more detail, as shown in figure 2.1.6.

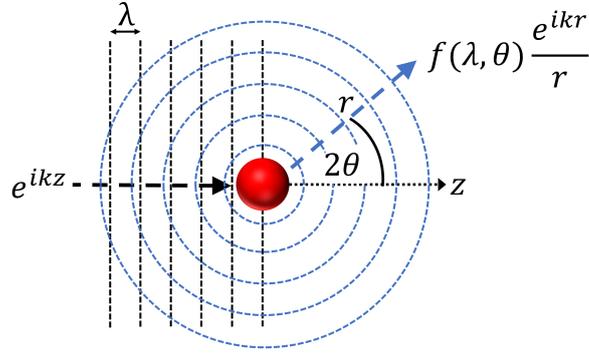


Figure 2.1.6: Elastic scattering of a beam of neutrons of wavelength λ off a fixed atom.

Here the incoming beam can be described as a complex plane wave in the form e^{ikz} moving in the z direction. Upon scattering a single neutron will move outwards radially, hence the component of the scattered wave in the r direction can be given by e^{ikr} . Because the wave is scattering radially the number of neutrons per unit area will decay with increasing distance r , hence the final scattered wave can be expressed as a wave function shown in equation 2.1.15, where ψ_0 is a constant of proportionality and $f(\lambda, \theta)$ is a function that describes the likelihood of a particle of wavelength λ being scattered at an angle θ .

$$\psi_f = \psi_0 f(\lambda, \theta) \frac{e^{ikr}}{r} \quad (2.1.15)$$

This $f(\lambda, \theta)$ function is better known as the “atomic form factor”, and is of vital importance if one wishes to meaningfully interpret the resultant diffraction pattern from a sample. In the case of neutron scattering this is very simple. As described in section 2.1.1, in order for a scattering experiment to be useful, the wavelength of the probe must be comparable to the structural motifs, which in the case of the samples studied in this thesis are on the order of \AA (10^{-10} m). As discussed in section 2.1.3, neutrons interact via the nuclear strong force, which operates on lengthscales of $\approx 10^{-14}$ m. Hence for

an appropriately tuned neutron of wavelength $\lambda \approx 10^{-10}$ m scattering off an atomic nucleus, the nucleus essentially acts as a point source and the neutron will scatter in all directions with equal probability. Hence we can write the atomic form factor as shown in equation 2.1.16, where b is a constant value known as the “scattering length” and the - sign is added for convention.

$$f(\lambda, \theta) = -b \quad (2.1.16)$$

The value of b is isotope specific, its magnitude represents the strength of scattering, and its sign represents whether the scattered wave is perfectly in phase or perfectly out of phase with the incident wave. It also varies with atomic number with no easily discernible trend. When this quantity is measured experimentally for a particular isotope one observes that it adopts an average value and an associated standard deviation, as shown in equation 2.1.17. The standard deviation originates from the relative orientation between the overall spin of the nucleus, and the spin of the incident neutron, hence nuclei with 0 overall spin show low standard deviations.

$$b = \langle b \rangle \pm \Delta b \quad (2.1.17)$$

The fact that the neutron form factor adopts a single value with practically no predictability for each isotope differs strongly from the form factor for x-ray diffraction, which is given in equation 2.1.18. The origin of this difference lies in the nature of x-ray scattering. X-rays, as photons, scatter off the electron cloud around an atom, whereas neutrons scatter off the nucleus. The x-ray form factor is therefore proportional to the atomic number Z , the classical radius of an electron r_e , and a function $g(Q)$ which has a value equal to unity at $Q = 0$ and decays towards 0 as $Q \rightarrow \infty$.

$$f(\lambda, \theta) = Zg(Q)r_e \quad (2.1.18)$$

With a knowledge of form factors we can now begin to recover the scattering cross section. If one pictures a spherical shell around an atom which is acting as a scattering centre and performs a surface integral over the entire sphere to determine the scattering rate, as shown in equation 2.1.19, one can then determine in the case of neutrons that the scattering cross section is proportional to the scattering length squared, as shown in equation 2.1.20.

$$R = \int_{2\theta=0}^{\pi} \int_{\phi=0}^{2\pi} |\psi_f|^2 dA \quad (2.1.19)$$

$$\sigma = 4\pi|b|^2 \quad (2.1.20)$$

One can then deconvolute this further by considering the standard deviation of the neutron scattering length. From equation 2.1.17 we can create the expression shown in equation 2.1.21

$$\langle b^2 \rangle = \langle b \rangle^2 + (\Delta b)^2 \quad (2.1.21)$$

This means that the scattering cross section σ can be deconvoluted into its coherent and incoherent parts, as shown in equation 2.1.22.

$$\langle \sigma \rangle = \sigma_{coh} + \sigma_{incoh} = 4\pi \langle b \rangle^2 + 4\pi(\Delta b)^2 \quad (2.1.22)$$

It is here that we can now begin to understand the true power of neutron scattering, compared with x-ray scattering, when one wishes to investigate the structure of aqueous samples. As shown in equation 2.1.18, the x ray scattering form factor is directly proportional to the atomic number Z , hence low molecular weight elements, such as hydrogen, will scatter very weakly in comparison to larger molecular weight elements, such as oxygen. Therefore if one was to perform an x-ray diffraction experiment on pure

water the signal would be dominated by the oxygen - oxygen correlations in the sample, and would not be able to resolve detailed structural information about the relative orientations of the hydrogen present in the sample. This is of crucial interest to us, as it is these orientations that are responsible for important phenomena in aqueous solutions such as hydrogen bonding.

We can now examine the coherent, incoherent, and total scattering cross sections, as shown in table 2.1.2 for the isotopic species studied in this thesis. Now we observe that hydrogen is in fact a very strong scatterer of neutrons, however the majority of this is incoherent scattering and will have to be corrected for post-experiment, as will be discussed further in section 2.1.6. This means that neutron diffraction is much more sensitive to hydrogen - oxygen and hydrogen - hydrogen correlations, and is therefore much more informative about the structure of aqueous solutions.

Table 2.1.2: Coherent neutron scattering length b_{coh} , coherent neutron scattering cross section σ_{coh} , incoherent neutron scattering cross section σ_{incoh} , and total neutron scattering cross section σ for hydrogen, deuterium, carbon, nitrogen, and oxygen taken from Sears [183]. Hydrogen differs from its isotope deuterium by a single neutron.

Isotope	b_{coh} (fm)	σ_{coh} (barn)	σ_{incoh} (barn)	σ (barn)
^1H	-3.74	1.76	80.27	82.03
^2H	6.67	5.92	2.05	7.64
^{12}C	6.65	5.56	0.00	5.56
^{14}N	9.37	11.03	0.50	11.53
^{16}O	5.80	4.23	0.00	4.23

Let us now return to the raw data of the neutron scattering experiment: the differential scattering cross section $\frac{d\sigma}{d\Omega}$ can be deconvoluted into its components arising from “self” scattering and “distinct” scattering, as shown in equation 2.1.23, where c_α is the concentration of atomic species α , b_α is the scattering length of atomic species α and $S_{\alpha\beta}(Q)$ is the partial structure factor between atomic species α and β [184].

$$\left(\frac{d\sigma}{d\Omega}\right) = \left(\frac{d\sigma}{d\Omega}\right)_{self} + \left(\frac{d\sigma}{d\Omega}\right)_{dist} = \sum_{\alpha} c_{\alpha} b_{\alpha}^2 + \sum_{\alpha\beta \geq \alpha} (2 - \delta_{\alpha\beta}) c_{\alpha} c_{\beta} b_{\alpha} b_{\beta} S_{\alpha\beta}(Q) \quad (2.1.23)$$

The “self” contribution arises from scattering by a single nucleus, and therefore provides no structural information and produces a featureless background signal which also needs removing post-experiment. The “distinct” contribution arises from scattering and interference from between pairs of atoms, and therefore does provide structural information about the sample. This component is proportional to the “total structure factor” $F(Q)$, which can now be more easily rewritten in equation 2.1.24.

$$F(Q) = \sum_{\alpha\beta} c_{\alpha} c_{\beta} b_{\alpha} b_{\beta} (S_{\alpha\beta}(Q) - 1) \quad (2.1.24)$$

From this equation one can see that the total structure factor is a weighted sum of all the partial structure factors present in the system. These partial structure factors are particularly useful when examining the structure in aqueous systems as $S_{\alpha\beta}(Q)$ is related to the radial distribution function between atoms α and β ($g_{\alpha\beta}(r)$) via Fourier transform, as shown in equation 2.1.25, where ρ is density.

$$S_{\alpha\beta}(Q) = \rho \int_0^{\infty} g_{\alpha\beta}(r) \exp(iQr) dr = 4\pi\rho \int_0^{\infty} g_{\alpha\beta}(r) \frac{\sin Qr}{Qr} dr \quad (2.1.25)$$

The radial distribution function $g_{\alpha\beta}(r)$ gives the local density of atomic species β at a given distance r from a central atomic species α normalised to the bulk density of β . Therefore by definition $g_{\alpha\beta}(r) = g_{\beta\alpha}(r)$. In principle we can now deconvolute of the total structure factor $F(Q)$ shown in equation 2.1.24 and the extract the pairwise radial distribution functions from the resulting partial structure factors $S(Q)$ shown in equation 2.1.25 to create a clear structural picture of a given sample by taking neutron scattering data. However, like any simultaneous equation, one requires as least as many different expressions describing the system as one has unknown variables in order to solve

it. In this case our unknown variables are the individual $S(Q)$ and the total expression is $F(Q)$. This therefore requires us to determine how many pairwise $S(Q)$ are there for a given aqueous system? It can be shown that for a system containing J different atomic species that one can calculate the number of different pairwise interactions N using equation 2.1.26.

$$N = \frac{J(J + 1)}{2} \tag{2.1.26}$$

One can then exploit the isotope specificity of neutron scattering to obtain multiple $F(Q)$ describing a single system. As shown in table 2.1.2, different isotopes can have markedly different coherent scattering lengths. As shown in equation 2.1.24, the $F(Q)$ is dependent on the coherent scattering lengths of the atomic species present in the sample. Therefore, by simply performing an isotope substitution such as exchanging hydrogen for its isotope deuterium, which is assumed to induce negligible structural change to an aqueous system, one can obtain a vastly different $F(Q)$ for an identically structured system.

To explain this point further let us consider the example of a neutron scattering experiment on pure water. Pure water contains two different atomic species, and therefore $N = J(J + 1)/2 = 3$. The individual pairwise interactions are therefore oxygen-oxygen, hydrogen-oxygen, and hydrogen-hydrogen. We therefore need 3 different $F(Q)$ in order to deconvolute all 3 pairwise $S(Q)$. This could be easily done in the case of pure water using hydrogen - deuterium substitution as one could prepare samples of H_2O , D_2O , and HDO (where H is hydrogen and D is the isotope deuterium). In principle then, all one needs to do to extract every individual $S(Q)$ for a given sample is perform experiments on N different isotope variants of the sample.

However, it is here where a problem arises. For more complex systems containing many different atomic species J , N rises exponentially, and to deconvolute $F(Q)$ would require increasingly large numbers of isotopic variants. This is expensive, time consuming, and in many cases likely impossible. Hydrogen deuterium substitution is widely used

in neutron scattering for the reason that their coherent scattering lengths b are vastly different, resulting in strongly differing $F(Q)$, and this exchange can be done reasonably easily. However, other isotope substitutions have significantly smaller effects and are significantly more difficult to achieve. It is also more than likely that in doing many different isotopic variants required for a complicated system one will end up with several $F(Q)$ that are essentially identical, and attempting to do a deconvolution over many highly similar $F(Q)$ will mean that the resultant $S(Q)$ will be highly unreliable. We therefore need a solution if we are not to be forever stuck investigating simple hydrogen rich liquids and luckily, one is presented in section 2.2.

2.1.5 The Neutron Scattering Experiment

As discussed in section 2.1.3, neutrons were first discovered by James Chadwick [182] through the reaction shown in equation 2.1.11. This was undoubtedly an incredible leap forward for physics and had countless implications, however it is not a good method for a neutron diffraction experiment as one requires a steady stream of neutrons of predictable wavelength. There two principle methods by which a stream of neutrons are produced for diffraction experiments: reactor sources and spallation sources [185].

Reactor sources produce neutrons via nuclear fission of heavy elements such as uranium, thorium, or plutonium. Fission occurs as a result of a large parent nucleus splitting into smaller daughter fragments due to the absorption of an incident neutron, as shown in figure 2.1.7. The combined rest mass of the daughter nuclei is less than that of the parent nucleus, and they both have higher binding energies per nucleon. This loss in mass is then converted to energy through the relation $E = mc^2$. The process also releases neutrons. In the case of ^{235}U an average of 2.7 neutrons are produced per fission event. The reaction can then become self sustaining one of the neutrons produced in the fission of a single ^{235}U nucleus goes on to cause a second fission of another ^{235}U nucleus. If more than a single neutron per fission event goes on to cause a second event the system will become “supercritical” and cause the reaction to run out of control. Conversely if less than a single neutron per fission event goes on to cause a second event the system

will become “subcritical” and the reaction will eventually stop. To help control this reactor sources make use of moderators, such as light water or graphite, which slow down neutrons to a speed where they will be more likely to cause a fission reaction.

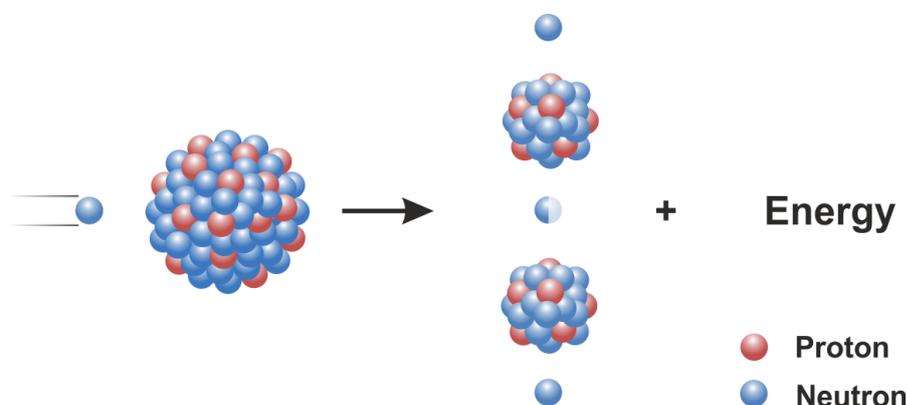


Figure 2.1.7: The process of nuclear fission. A parent nucleus absorbs a neutron and decays into daughter nuclei, neutrons, and energy. Figure adapted from [186].

The remaining neutrons escape the reactor to be used in diffraction studies through instruments situated around the reactor. The most powerful reactor source used for scientific inquiry is the Institut Laue-Langevin (ILL) in Grenoble, France. This uses a single element of uranium enriched with 93% ^{235}U to produce a thermal neutron flux of $1.5 \times 10^{15} \text{ cm}^{-2} \text{ s}^{-1}$ and 58 MW of heat. In the case of reactor sources neutrons are produced continuously over a range of energies. In order for the emerging neutrons to be useful in diffraction experiments they must therefore be strongly monochromated so they are of a predictable wavelength.

Spallation sources are more complex than reactor sources. In these cases a heavy metal target, such as tungsten, is bombarded by high energy protons accelerated using linear accelerators, cyclotrons, or synchrotrons, as shown in figure 2.1.8. This begins a nucleon cascade within the target nucleus, causing the ejection of high energy neutrons to be absorbed by other nuclei. This is then followed by a de-excitation which causes the evaporation of many more lower energy neutrons to be used in diffraction experiments. Through this reaction it is possible to produce 20-40 neutrons per incident particle depending on the target material.

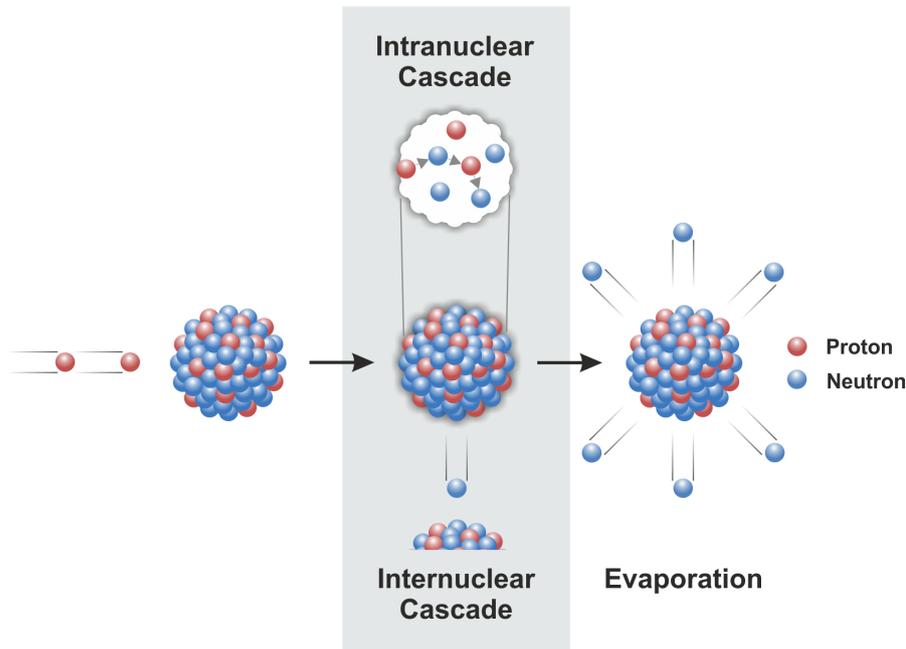


Figure 2.1.8: The process of spallation. A parent nucleus absorbs a high energy proton, first causing an internuclear cascade which releases high energy neutrons, then causing evaporation of more low energy neutrons. Figure adapted from [186].

The ISIS neutron and muon source, UK, is a synchrotron based spallation source where the experimental neutron scattering data presented in this thesis was gathered. A schematic of this facility is shown in figure 2.1.9. Here H^- ions are accelerated by a linear accelerator from 665 keV to 70 MeV. They then pass through a thin sheet of alumina which strips the electrons from the H^- ions and leaves bare protons. These are then introduced into the synchrotron which bends the protons into a circular path using focusing magnets and accelerates them further to 800 MeV. This also causes the beam to separate into two bunches, each containing 2.5×10^{13} protons equivalent to $200 \mu A$, on opposite sides of the synchrotron. Once completed the beam is magnetically “kicked” upwards by 1° out of the synchrotron and into the extracted proton beamlines. These then guide the beam to tungsten targets, where the spallation reaction occurs and the produced neutrons go outwards into the suites of instruments arranged radially around each target station. It is the focussing of the proton beam into bunches that trigger spallation events at regular intervals that makes ISIS a pulsed muon source.

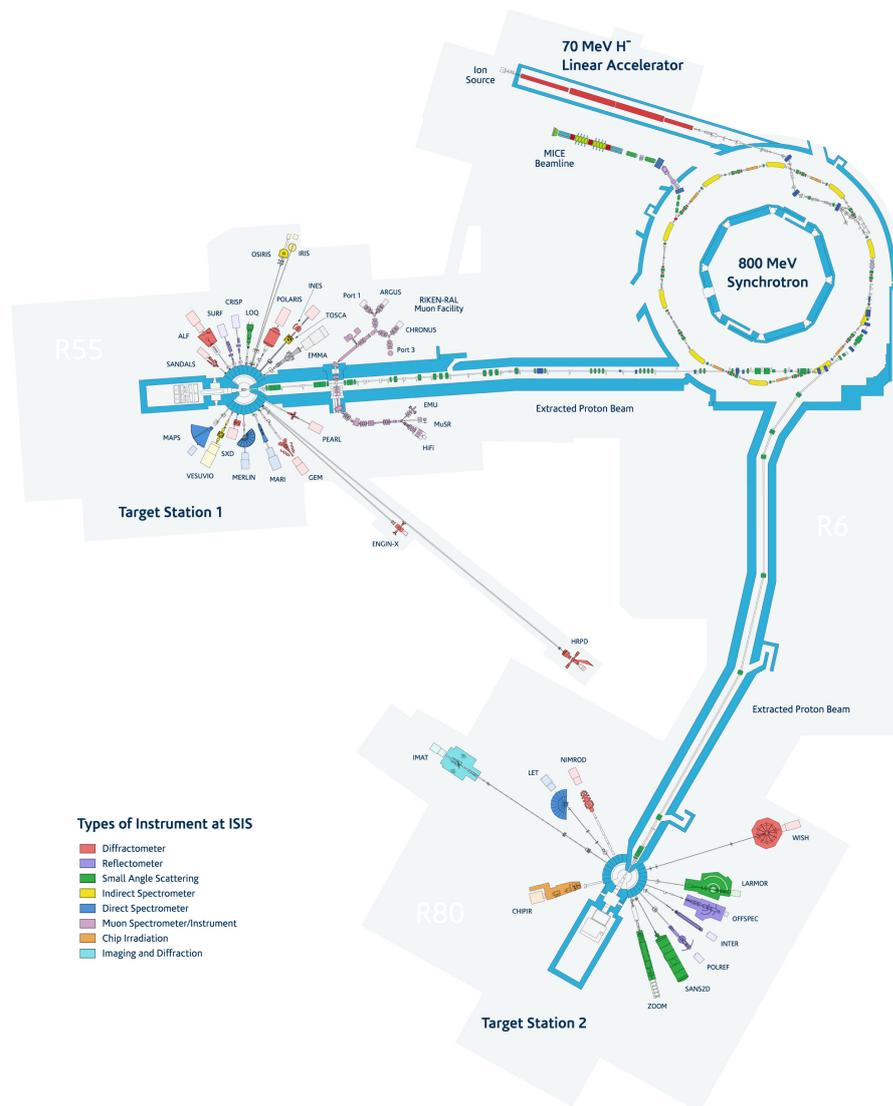


Figure 2.1.9: A schematic of the ISIS facility. Figure adapted from [187].

The advantage of generating pulses in this manner is that one now knows precisely the moment of neutron formation. One can therefore monitor the time between the neutron production at the target and the detection of an incident neutron at the instrument and calculate its velocity and therefore Q according to equation 2.1.27. Here h is Planck's constant, m is the mass of the neutron, t is the time between proton pulses, L is the distance from the target to the instrument, and t_0 and L_0 are instrument specific constants.

$$\lambda = \frac{h(t - t_0)}{m(L - L_0)} \quad (2.1.27)$$

This means that the beam does not require monochromation, as the Q value for each neutron is precisely known and the whole pulse can be used. The lag time between pulses also gives the target materials time to cool, and they are therefore more efficient. This is ultimately the limiting factor in power for both reactor and pulsed sources, as both sources require significant cooling to prevent meltdown. However, pulsed sources are far more complex and therefore less reliable than reactor sources, and at present generate less flux than reactor sources like the ILL.

All the raw neutron scattering data acquired in this thesis was taken on the Near and InterMediate Range Order Diffractometer (NIMROD) at the ISIS neutron and muon source, shown in figure 2.1.10. This instrument is able to collect neutron data over a wide Q range from 0.02-50.0 \AA^{-1} , corresponding to structures ranging from 0.1 to 300 \AA . It has therefore been used to investigate structural properties of a host of systems, including: deep eutectic solvents [188–190], porous ices [191, 192], 2D nano-materials [193, 194], catalysis [195], confined fluids [196, 197], and aqueous solutions containing biomolecules [1–3, 198].

This is made possible through 1098 detector elements covering scattering angles (2θ) from 3.5 to 40 $^\circ$. It also features a low angle detector bank containing a dartboard shaped array of a further 756 detector elements to cover scatter angles from 0.5 to 2.2 $^\circ$ [199]. It is this wide range of detector angles which makes NIMROD a versatile instrument well suited to the study of aqueous solutions containing water and larger biomolecules.

Each detector element consists of 22 layers of a ZnS lattice doped with Ag and interspersed with 6LiF . These are sandwiched between 23 layers of glass and viewed by two photomultiplier tubes. Upon the interaction of an incident neutron the reaction shown in equation 2.1.28 takes place, producing helium, the hydrogen isotope tritium, and energy.

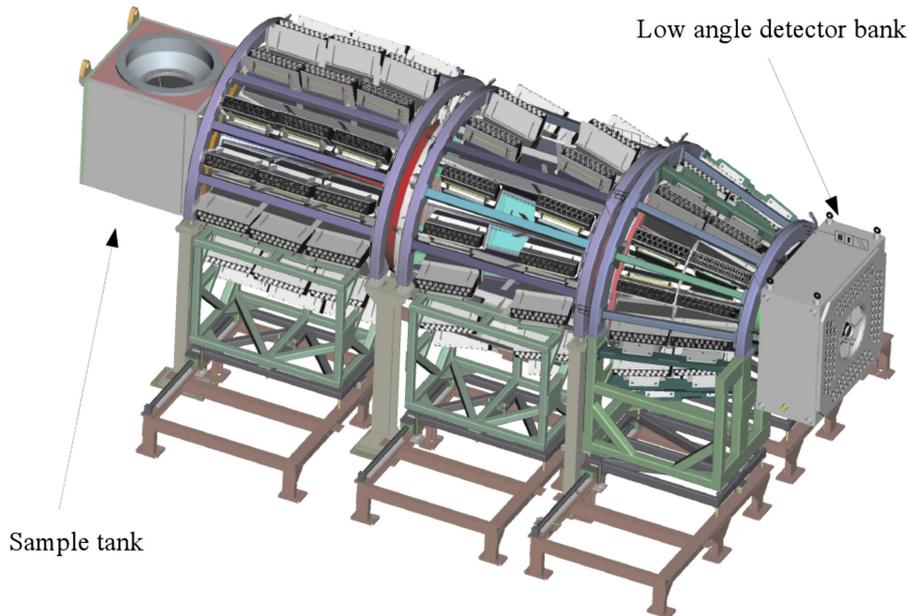


Figure 2.1.10: A schematic of the NIMROD instrument. Figure adapted from [199].



This energy then goes on to excite an electron and form an electron-hole pair. When this electron-hole pair then reaches an *Ag* in the *ZnS* lattice it recombines, releasing a photon [186]. This then travels into the photomultiplier tubes to increase the signal so it can be counted, and the raw data finally to be acquired.

The last consideration is the sample holder itself. In this research we made use of two sample containers: the flat plate cell and the high pressure cell depicted in figure 2.1.11. Both cells can be used to study aqueous samples at a range of temperatures by heating/cooling the cells while in the neutron beam using a circulating water/oil pump.

Both cells are made of an alloy of titanium (*Ti*) and zirconium (*Zr*) at a molar ratio of 7.16 : 3.438. This material is chosen as the coherent neutron scattering length of *Ti* is -3.44 fm, and of *Zr* is 7.16 fm [183]. The coherent neutron scattering length of the whole canister is then the weighted average of the coherent scattering lengths of the constituent elements, therefore at the molar ratio employed the coherent scattering

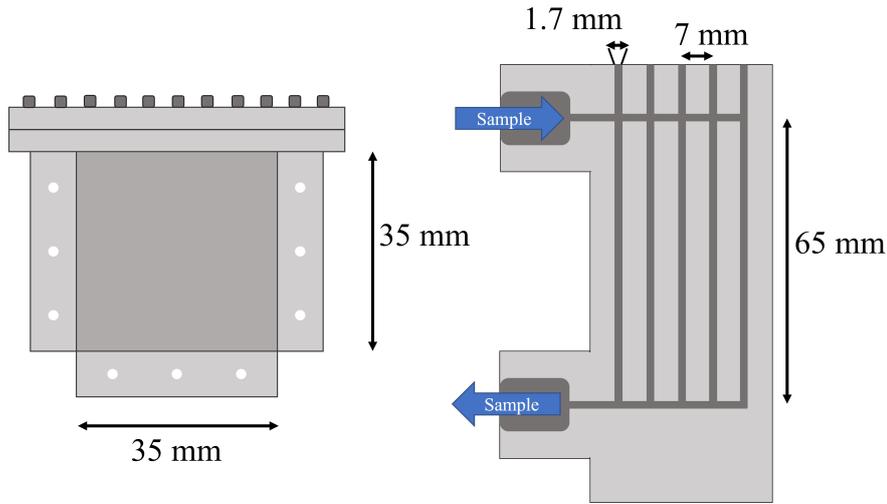


Figure 2.1.11: A schematic of the TiZr canisters used for neutron scattering experiments. Experiments at ambient pressure were made using the 35 x 35 x 1 mm cans (left). High pressure experiments were made using the high pressure cell (right), which features five cylindrical bores containing the sample arranged in a flat plane, each 1.7 mm wide and spaced at 7 mm intervals.

length of the container is essentially 0 and the canister is effectively a null scatterer. This minimises large signal arising from the sample container.

The main difference between the two cells is the flat plate cell is simply a hollow flat plate into which the sample is filled, whereas the high pressure cell contains 5 capillaries arranged in a plane into which the sample is filled so pressure can be applied. The high pressure cell is also much thicker so as to withstand the high pressures (up to 5 kbar) which can be applied to the sample. This means that while the high pressure cell can explore a greater range of externalities, the ratio between the volume of sample and the volume of the container in the beam is lower, and hence the experimental data will require greater correction. Hence unless one is particularly concerned with the effects of pressure on one's sample, it is a more sensible choice to use the flat plate cell.

2.1.6 Data Correction - Gudrun

We now have the means to acquire data, the theoretical understanding of the relationship between the observed diffraction pattern and the structure of the system, and the method

of using isotopic substitution to produce many identically structured samples that each produce a different $F(Q)$. However, there still exists a hurdle to overcome. As discussed in section 2.1.4, for the data to be meaningful only coherent scattering can be measured and all scattering events must be completely elastic. However, as the ZnS detectors detect all neutrons indiscriminately, it is inevitable that neutrons scattered through incoherent and inelastic processes will be introduced into the data. We also must discard any neutrons that have undergone multiple scattering events within the sample, as well as the featureless background signal produced by self scattering which contains most of the inelastic scattering. Finally we must remove any measured scattering arising from the sample container. We therefore require a method of correcting the data post acquisition to yield purely coherent elastic scattering data from the sample alone, and one such solution is Gudrun developed by Alan Soper *et al.* [184]. In this section we will now overview the method by which Gudrun functions. This is a highly complex process, and so will not be described in full detail here. A complete explanation containing rigorous mathematical detail can be found in reference [184].

In order to enable data corrections to take place Gudrun requires the following datasets:

- Background - taken in the absence of any sample or container
- Vanadium standard - Vanadium has a very low coherent scattering cross section (0.018 barns) but a high incoherent scattering cross section (5.08 barns) [183] and can therefore be used for calibration
- Container - each sample container to be used in the experiment without any sample present
- Sample - the experimental samples in the containers

These supply the raw material using which Gudrun can then begin correction. Its first step is to identify any detectors that may be malfunctioning resulting in over or underestimated counts. Gudrun achieves this by running PURGE. This begins by grouping

detectors together into 18 groups and measures the ratio between counts between detectors over a given timeframe, along with the standard deviation. Detectors that cause a high standard deviation above a certain threshold are over counting, and are eliminated, and detectors that cause a low standard deviation below a certain threshold are under counting, and are also eliminated. This process is iterated with increasingly restrictive thresholds until no more detectors are eliminated, and therefore the remaining data is only being derived from correctly functioning detectors.

Gudrun then seeks to correct for attenuation. To do this one requires knowledge of the total cross section for the materials of the sample and container. For most elements the total cross section of a material $\sigma^{(t)}(\lambda)$ is the sum of the scattering cross section $\sigma^{(s)}(\lambda)$ and the absorption cross section $\sigma^{(a)}(\lambda)$. So long as nuclear resonances are avoided, the absorption cross section is linearly dependent on wavelength, and if the constant of proportionality for each element is known, this can be easily calculated.

For most elements, and in particular larger elements, $\sigma^{(s)}(\lambda)$ is independent of the wavelength of the incident neutron λ , however this is not the case with light elements with significant inelastic scattering such as hydrogen and deuterium. In order to determine $\sigma^{(s)}(\lambda)$ we can monitor the transmitted beam via a monitor detector placed directly behind the sample. This signal is connected to $\sigma^{(t)}(\lambda)$ in the case of flat plane sample container geometries by equation 2.1.29, where ρ is the sample density and L is the plate thickness [184]. Gudrun can also do this correction without the use of a transmission monitor by using the scattering cross sections reported in the Sears tables [183].

$$TRANS(\lambda) = \exp(-\rho\sigma^{(t)}(\lambda)L) \quad (2.1.29)$$

Gudrun then corrects for the intrinsic deadtime of the detectors. This arises from the time resolution of the detectors. This means that two neutrons detected one after another over a very short time period can be read as a single signal. Gudrun corrects for this by monitoring the number of counts per μs and the number of pulses of neutrons received. Following all these procedures means that the background can now be subtracted from

the measured signal to yield the pure elastic coherent distinct scattering data from the sample alone.

The last step is to put the data on an absolute scale. To do this the data is normalised to the vanadium standard, as vanadium is a metal with a precisely known density that does not require a container with a very small coherent scattering length (-0.38 fm [183]). It is also a high mass element, meaning that inelastic scattering is fairly negligible, and its attenuation properties are well known. This normalisation is shown in equation 2.1.30 [186].

$$\textit{Corrected result} = \frac{\textit{Sample data} - \textit{Empty container}}{\textit{Vanadium standard} - \textit{Empty instrument}} \quad (2.1.30)$$

This whole process is then performed iteratively (typically 5 iterations) to yield the final datasets for analysis. The number of iterations required is typically highly sample dependent, where samples containing large quantities of hydrogen, that are therefore prone to significant inelastic scattering, often require a higher number of iterations. In order to determine the “correct” number of iterations the procedure is iterated a different numbers of times (3, 5, 7, 9...), and compared. When the data at low Q , where inelastic scattering is most prominent, no longer changes significantly if the number of iterations are increased, this is deemed to be sufficiently corrected without introducing errors due to over-correction.

2.2 Empirical Potential Structure Refinement

We have now finally arrived at the point where we have datasets on a few isotopic variants of our sample that have been corrected using Gudrun and so we have been left with coherent elastic distinct scattering data on the sample alone. However, as mentioned in section 2.1.4, we still have the problem that for a sample containing J distinct atomic species we have $N = J(J + 1)/2$ different correlations, and therefore

require N different $F(Q)$ to successfully deconvolute all the constituent $S(Q)$ and get an accurate depiction of the structure of the aqueous system. This is often impossible if the sample contains even a fairly modest amount of distinct isotopic species. To overcome this final problem we turn to the technique developed by Alan Soper [200, 201], Empirical Potential Structure Refinement (EPSR).

EPSR is a reverse Monte Carlo simulation based technique of analysing experimental diffraction data. In short, diffraction data are taken on several isotopic variants of a particular sample. For samples containing water and other components that do not contain hydrogen, and for which isotopic substitution is therefore likely to be costly and the isotopes may not have considerably different b values, one typically takes three datasets: H_2O + other components, D_2O + other components, and HDO + other components. If the other components do contain hydrogen, such as the biomolecules that will be studied in this thesis, one typically does seven datasets: H_2O + H-X, H_2O + D-X, D_2O + H-X, D_2O + D-X, HDO + H-X, HDO + D-X, and HDO + HD-X where X refers to the additional component and H/D/HD refers to whether the component contains fully hydrogen, fully deuterium, or an equal mixture of the two. This is only a rule of thumb, and the required datasets are more often decided through difficulties and cost associated with sample preparation.

A simulation is then built which matches the experimental components, concentration, and density. Theoretical scattering data can be produced from this simulation, and the simulation is then refined against all the experimental scattering datasets until the two match to an acceptable degree of accuracy. At this point the simulation is deemed to be an accurate representation of the experimental sample, and can therefore be studied to determine the structures present in the aqueous solution. This means that one does not need to take the full N datasets to still obtain all the different $S(Q)$ present in the sample. This process has been applied to understand the structure of a variety of liquids such as: water at various temperatures and pressures [115], ionic liquids [202], aqueous solutions of ions [126, 128, 131], alcohols [203, 204], and organic molecules such as osmolytes [2, 3, 62, 141], amino acids [1, 138], and lipids [198, 205] to name a few. It therefore serves as the core method for this thesis, and will be described in greater

detail in the following sections.

2.2.1 Building the Simulation

Once one has acquired and corrected the data on isotopic variants of ones sample, the next step is to build a simulation which matches the experimental sample. To do this one requires knowledge of the concentrations of any solutes, and the density of the sample. In this thesis density values were either taken from previous literature or measured experimentally using an Anton Parr DMA 4100 M densitometer. Then one can begin to build the molecules present in ones sample. EPSR does this by using Jmol, a Java based program designed for the building of molecular structures. Within Jmol it is now possible to draw out molecules, or open models acquired from literature sources, and load them into EPSR where bond lengths and angles can be further modified. It is at this point that the reference potentials describing each atom species are defined. This reference potential consists of a Lennard-Jones and a Coulomb component, and will be discussed in greater detail in section 2.2.3.

Next one has to fill the simulation box. The upper limit of ones box size is largely dependent on the computing power one has available. A larger box size will gather more statistics and better approximate an infinite system, but will run more slowly. EPSR also only calculates structural correlations over a maximum distance of half the box width, so if one wishes to investigate large scale structures, one will have to use a larger box size. To mitigate the issues caused by a finite box size EPSR uses periodic boundary conditions, where when a molecule exits one side of the box it immediately reenters from the opposite side, to approximate an infinite system. For this research it was determined that a reasonable compromise between box size and iteration time was approximately 5000 molecules, resulting in a cubic box of dimension approximately 55 Å.

Next one introduces the experimental data taken on the various isotopic variants of the sample, and generates additional files describing the isotopic ratios and whether the hydrogens are exchangeable. Finally EPSR writes a .inp file, which effectively serves as a master file that runs the simulation. Here various general parameters can be modified.

The simulation is now ready to proceed.

2.2.2 Monte Carlo

EPSR is a Monte Carlo based simulation procedure. This means that each molecule within the box is individually moved at random. These movements can be whole molecule translations, whole molecule rotations if the molecule is larger than a single atom, and motions within the molecule such as side chain rotations. Following this movement the potential energy of the system is calculated and compared with the potential energy of the system prior to the movement, as described in equation 2.2.1.

$$\Delta U = U_{after} - U_{before} \quad (2.2.1)$$

If $\Delta U < 0$ the move is accepted, but if $\Delta U > 0$ the move is accepted with a probability $\exp\left(-\frac{\Delta U}{kT}\right)$, where k is the Boltzmann constant and T is temperature. This is known as the Metropolis condition. Within EPSR the potential is the sum of two parts, the reference potential, and the empirical potential. Hence the potential energy between two atoms of type α and β would be expressed as shown in equation 2.2.2. We will now explore these two potentials in greater detail.

$$U_{\alpha\beta}(r) = U_{\alpha\beta}^{Emp}(r) + U_{\alpha\beta}^{Ref}(r) \quad (2.2.2)$$

2.2.3 Reference Potential

Within EPSR an intramolecular potential is used that controls the geometry of the individual molecules. Atoms within a molecule are assumed to interact via a harmonic potential, with each intramolecular distance controlled by the average distance $d_{\alpha\beta}$ and a width $w_{\alpha\beta}$. The total intramolecular energy of the system is then given by equation 2.2.3, where $r_{\alpha(i)\beta(i)}$ is the separation distance between two atoms of species α and β

within molecule i . The C term is a constant derived from the high Q data, and was kept at its default setting of $C/2 = 65 \text{ \AA amu}^{\frac{1}{2}}$.

$$U_{intra} = C \sum_i \sum_{\alpha\beta>\alpha} \frac{(r_{\alpha(i)\beta(i)} - d_{\alpha\beta})^2}{2w_{\alpha\beta}^2} \quad (2.2.3)$$

The width term of the potential depends on the set average distance between the atom species and the mass of the atoms and allows them to adopt a broader range. This allows for the natural fluctuation in bond lengths due to thermal fluctuations to be accounted for without the need for individual Debye-Waller factors, which are likely impossible to incorporate into EPSR. It is shown in equation 2.2.4.

$$w_{\alpha\beta}^2 = d_{\alpha\beta} \left(\frac{M_\alpha M_\beta}{M_\alpha + M_\beta} \right)^{-\frac{1}{2}} \quad (2.2.4)$$

This potential does not include any similar harmonic constraints on bond angles, but instead achieves this by employing the same distance potential. Atom B is bonded to both atom A and atom C , and therefore forms the apex of bond angle $\angle ABC$. The harmonic distance potential is applied to atoms A and C and therefore keeps the bond angle about a fixed value.

We will now move on to describe the interatomic potential. As stated in section 2.2.1, the reference potential for each atom consists of a Lennard-Jones component and a Coulomb component and can therefore be written as shown in equation 2.2.5.

$$U_{\alpha\beta}^{Ref}(r) = U_{\alpha\beta}^{LJ}(r) + U_{\alpha\beta}^q(r) = 4\varepsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r} \right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r} \right)^6 \right] + \frac{q_\alpha q_\beta}{4\pi\varepsilon_0 r} \quad (2.2.5)$$

The Lennard Jones component only significantly interacts at relatively short distances and is used to mimic the van der Waals interaction between atoms and molecules [28]. Its general form can be observed in figure 2.2.1.

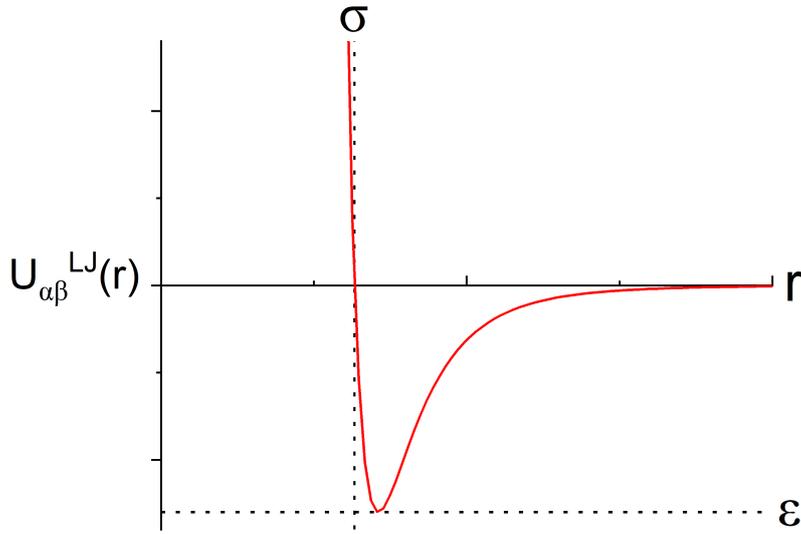


Figure 2.2.1: The general form of the Lennard Jones potential as a function of distance. The corresponding values of σ and ϵ are shown on the graph.

Here it is shown that ϵ corresponds to the potential well depth and has units of kJ/mol and σ corresponds to the distance below which the interactions between the atom/molecule species begin to become repulsive and has units of \AA . It is also known as the van der Waals radius. The origin of this sharp increase in potential below σ is due to steric repulsion between atoms/molecules due to the Pauli exclusion principle, and its inclusion therefore prevents atomic overlap. The favourable potential well after σ is due to the polarisation of the atoms/molecules resulting in favourable electrostatic interactions between areas of net positive and negative charge. It can be shown through differentiation that the most preferable distance with the lowest potential energy occurs at the minimum at a distance of $r = 2^{\frac{1}{6}}\sigma$. Beyond this point the function decays until at large distances it is negligible.

Within EPSR each atom species is assigned a σ and ϵ parameter that describe this potential. So for two atomic species α and β described by σ_α , ϵ_α , σ_β , and ϵ_β , the terms $\sigma_{\alpha\beta}$ and $\epsilon_{\alpha\beta}$ in $U_{\alpha\beta}^{LJ}(r)$ are calculated according to the Lorentz-Berthelot mixing rules, shown in equation 2.2.6 and 2.2.7.

$$\varepsilon_{\alpha\beta} = [\varepsilon_{\alpha}\varepsilon_{\beta}]^{\frac{1}{2}} \quad (2.2.6)$$

$$\sigma_{\alpha\beta} = \frac{1}{2}(\sigma_{\alpha} + \sigma_{\beta}) \quad (2.2.7)$$

The general form of the electrostatic component is simply Coulombs law [28], which states that the electrostatic potential between two charged objects is inversely proportional to the distance between the objects and directly proportional to the product of their charges. Therefore the potential between two oppositely charged atoms/molecules is attractive and for similarly charged atoms/molecules it is repulsive.

A Monte Carlo simulation using only this potential would evolve to look something like the experimental sample, and therefore serves as a good base, but must be refined yet further using an additional potential that describes the system more accurately. This is the role of the empirical potential.

2.2.4 Empirical Potential

The empirical potential is named as such because it is derived solely from the difference between the supplied diffraction data and the theoretical scattering data from the simulation. It does not derive from any physical principles. After trial and error, the current form of the empirical potential in EPSR is a series of Poisson functions shown in equation 2.2.8 and 2.2.9. Here the C_i values are real numbers and are derived from the difference between the simulated and experimental data. They are therefore referred to as “difference coefficients”. The iterative procedure of EPSR therefore improves the fit between the simulated and experimental data and refines the potential by modifying these values via the Monte Carlo route such that they attempt to approach 0. This will be discussed in more detail in the following section. Finally σ_r is a width function and ρ is the atomic number density of the system.

$$U^{EP}(r) = kT \sum_i C_i p_{n_i}(r, \sigma_r) \quad (2.2.8)$$

$$p_n(r, \sigma) = \frac{1}{4\pi\rho\sigma^3(n+2)!} \left(\frac{r}{\sigma}\right)^n \exp\left[-\frac{r}{\sigma}\right] \quad (2.2.9)$$

Equation 2.2.9 is also useful for EPSR as it has an exact Fourier transform which allows it to be expressed in Q space without introducing truncation artefacts. This is shown in equation 2.2.10, where $\alpha = \arctan(Q\sigma)$.

$$P_n(Q, \sigma) = \frac{1}{(n+2)(\sqrt{1+Q^2\sigma^2})^{(n+4)}} \left[2 \cos(n\alpha) + \frac{(1-Q^2\sigma^2)}{Q\sigma} \sin(n\alpha) \right] \quad (2.2.10)$$

Hence the values for C_i are produced by fitting a series as shown in equation 2.2.11 to the experimental Q data. EPSR then performs a second Monte Carlo step based off these values and the newly calculated empirical potential, and the process of fitting begins again.

$$U_{EP}(Q) = \sum_i C_i P_{n_i}(Q, \sigma_Q) \quad (2.2.11)$$

With these potentials now all determined we can give the final expression for the probability that a whole molecule move or internal molecule structure change such as a subgroup rotation is accepted. This is shown in equation 2.2.12. The intramolecular potential is included in this acceptance because while whole molecule moves should not in principle affect ΔU_{intra} , over a large number of moves round off errors can accumulate, hence the simulation is more accurate if it is readdressed at every step.

$$\text{Acceptance Probability} = \exp\left(-\left[\Delta U_{intra} + \frac{1}{kT}(\Delta U_{Ref} + \Delta U_{Ep})\right]\right) \quad (2.2.12)$$

We will now discuss the method of refining the potential in greater mathematical detail. As described in section 2.1.4, for J distinct atomic species there exists $N = J(J+1)/2$ different atomic correlations. We also showed that each total structure factor for each different isotopic variant of the sample can be expressed as the weighted sum of the partial structure factors, as shown in equation 2.2.13 where j represents a particular atomic pair and w_{ij} is the weighting of the particular atomic pair calculated through their relative concentration and neutron scattering length.

$$F_i(Q) = \sum_{j=1,N} w_{ij} S_j(Q) \quad (2.2.13)$$

Equating this to equation 2.1.23, we can say that the individual weights can be represented by equation 2.2.14 in the case of un-normalized data.

$$w_{ij} = (2 - \delta_{\alpha\beta}) c_\alpha c_\beta b_\alpha b_\beta \quad (2.2.14)$$

The weights will therefore form a matrix of dimension $M \times N$, where N is the number of distinct atomic correlations and M is the number of datasets. It is necessary to invert this matrix to generate a perturbation to each of the site-site potential functions and refine the simulation. This now presents a problem. The inversion of a non-square $N \times M$ matrix is not trivial. We also need to determine how reliable each of the individual datasets are. This issue is particularly problematic if one or several $S(Q)$ are very weakly weighted due to low concentration or neutron scattering lengths, and therefore the calculated $S(Q)$ will be highly unreliable. To address this EPSR introduces a confidence, or feedback, factor f , which is between 0 and 1 that modifies the weights. This also helps form a corresponding confidence factor in the simulation. This is described in equation 2.2.15 and essentially says that the data is accepted with confidence f and the simulation is accepted with confidence $(1 - f)$.

$$F(Q) = \begin{array}{c} \text{Data, } M \\ \hline \begin{matrix} fw_{11} & fw_{12} & \dots & & \dots & fw_{1N} \\ fw_{21} & fw_{22} & \dots & & \dots & fw_{2N} \\ \dots & \dots & & & & \dots \\ fw_{M1} & fw_{M2} & & & \dots & fw_{MN} \end{matrix} \\ \hline \begin{matrix} (1-f)\delta_{ij} & 0 & 0 & \dots & \dots & 0 \\ 0 & (1-f)\delta_{ij} & 0 & \dots & \dots & \dots \\ 0 & 0 & (1-f)\delta_{ij} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & (1-f)\delta_{ij} & 0 \\ 0 & \dots & \dots & \dots & 0 & (1-f)\delta_{ij} \end{matrix} \\ \hline \text{Simulation, } N \end{array} \times \begin{bmatrix} S_1(Q) \\ S_2(Q) \\ \dots \\ S_N(Q) \end{bmatrix}$$

Figure 2.2.2: Expressing $F(Q)$ using the modified weights matrix containing the confidence factor f for the M datasets and the N distinct interatomic correlations. Inverting this matrix can then yield each distinct $S(Q)$.

$$f(x) = \begin{cases} fw_{ij} & 1 \leq i \leq M \\ (1-f)\delta_{(i-M),j} & M < i \leq (M+N) \end{cases} \quad (2.2.15)$$

The final result is a complete over-determined matrix made up of the modified weights from the data and the simulation, and therefore has dimensions of $N \times (M + N)$. This is shown visually in figure 2.2.2. This matrix is then inverted using an iterative Monte Carlo procedure to allow for the determination of the difference coefficients following each simulation iteration, and the empirical potential to be refined.

This process is iterated until the difference coefficients become essentially 0 and the empirical potential no longer changes, or the amplitude of the empirical potential reaches a predefined limit. At this point the simulation iterations can be accumulated to gather meaningful statistics about the structure of the sample. The latter case is more often used in experimental work, as without capping the amplitude of the empirical potential it would become excessively large in its efforts to minimise the difference coefficients and result in physically unreasonable structures.

Within EPSR this is controlled by the value “EREQ”. EREQ simply measures the difference between the highest and lowest points of a particular site-site empirical potential and is therefore given units of kJ/mol. The choice of setting the value for EREQ can be informed by the Metropolis condition stated in equation 2.2.1. The acceptance of a particular molecular motion is dependent on the relative difference between the change in potential energy as a result of the move and kT . At 300K this is equivalent to approximately 2.5 kJ/mol. Hence if one were to set EREQ to be far lower than kT , it would be unlikely to have any effect of successfully changing the structure in the simulation. If it is much larger than kT , the simulation would begin to accumulate unreasonable structures. Therefore a good rule of thumb is to set EREQ to be on the same order as kT . For the research conducted in this thesis, EREQ values between 4 - 12 were typically employed as these are comparable to kT and yielded the highest quality fits in each case.

In EPSR the quality of the fit between the simulated and experimental data is measured by the “R factor”. This is described by equation 2.2.16, where M is the number of datasets, $D_i(Q)$ and $F_i(Q)$ are the experimental and simulated structure factor for the i th dataset, and $n_Q(i)$ is the number of Q values for the i th dataset. A perfect agreement between experimental and simulated structure factors therefore yields an R factor = 0.

$$R = \frac{1}{M} \sum_i \frac{1}{n_Q(i)} \sum_Q [D_i(Q) - F_i(Q)]^2 \quad (2.2.16)$$

2.2.5 Auxiliary Routines

We are now at the last step of our EPSR journey, where we have taken data, corrected it with Gudrun, built a simulation, allowed it to equilibrate using the reference potential, and then introduced the empirical potential to refine the simulation. After many iterations with the empirical potential active what we should now have is a simulated box of molecules whose structure is consistent with the experimental scattering data. This is the main caveat of EPSR. It does not yield a single unique solution that is guaranteed

to perfectly describe the experimental sample. What it produces is a simulation that closely matches the corrected experimental data and is based on sensible starting parameters. It is also based on a Monte Carlo simulation, and therefore molecular motions are performed at random and accepted with a probability based on the change in potential energy of the molecule at the new location, rather than in response to force fields generated by neighbouring molecules as they are in molecular dynamics. In this respect EPSR is weaker than molecular dynamics as individual molecular motions will more than likely not be realistic and EPSR cannot follow dynamic processes, only yield ensemble averaged structural information. However the advantage of EPSR is that the refinement process means that EPSR is continuously comparing itself to real experimental data. This means that, in principle, so long as the reference potential is vaguely sensible, the empirical potential should mean that a good and consistent solution is always found. Whereas molecular dynamics is based on force fields defined by the user which, while based on experimental observations, means that the final simulation can vary strongly with varying starting force fields.

At this point we can begin to accumulate statistics over thousands more iterations of the EPSR simulation to extract relevant structural data. This can include the main tool of EPSR, the radial distribution functions, which can now be generated by physically measuring the distance between atoms in the simulation over many iterations to generate smooth functions. However, a number of auxiliary routines are also included in EPSR that can yield other relevant structural information. The ones employed in this thesis are: COORD, CLUSTERS, and SHARM, which will now be explained.

The COORD routine calculates the coordination number of an atomic species β around a central atom species α , where α and β can be the same or different. The means of operation for this procedure is fairly simple. The user defines a minimum radius **rmin** and a maximum radius **rmax**. The routine then counts how many atoms of species β fall within this distance range around the central atom species α over the whole simulation and produces a distribution of coordination numbers and an associated standard deviation. In practice one tends to set **rmin** and **rmax** to correspond to the local minima in the associated $g_{\alpha\beta}(r)$ so one can examine the coordination number in

particular hydration shells.

The CLUSTERS routine is used to quantify association of molecules in solution in a manner not dissimilar to COORD. Consider a molecule x containing atom α and molecule y containing atom β . The user again sets an **rmin** and **rmax** for the distance between α and β which typically corresponds to the locations of minima in $g_{\alpha\beta}(r)$. If the two atoms fall within this distance then the molecules x and y are deemed to be part of the same cluster. The routine then identifies all molecules in the simulation which are also part of the same cluster and counts how large the cluster is. Finally the routine creates a distribution of cluster sizes found in solution by calculating the normalised probability of finding a cluster of a given size.

The final routine used in this thesis is SHARM. This is used to calculate the spacial density functions of molecules around a central molecule species. These are effectively the same as the radial distribution functions however they are expressed in 3D rather than in 2D and therefore allow the structures present in the sample to more easily visualised. In order to facilitate the description of this routine let us consider a water molecule. The first step in this routine is to define the molecules of interest by their rotational symmetry and molecular axes. In the case of a water molecule its rotational symmetry is 2. SHARM defines the molecular axes by the user stating a central atom type, in the case of a water molecule this would be the oxygen atom, which serves as the origin of the x, y, z coordinate system containing the molecule. The user then defines the yz plane which contains the central atom type and two other atom types located in the plane. In the case of a water molecule this would be the two hydrogen atoms. Finally the y axis is defined by stating an atom which the axis passes through, hence the y axis is parallel to the bond between the central atom and this atom. In the case of the water molecule this atom would be either of the two hydrogens. This then allows the x and z axis to be defined as all three axes must be orthogonal. This procedure is then repeated for the second molecule type of interest, for example choosing water as the second molecule type allows the structure of pure water to be visualised. With the molecular axes defined SHARM can then use the simulation to calculate the spherical harmonic coefficients. This is a complex mathematical procedure and will not

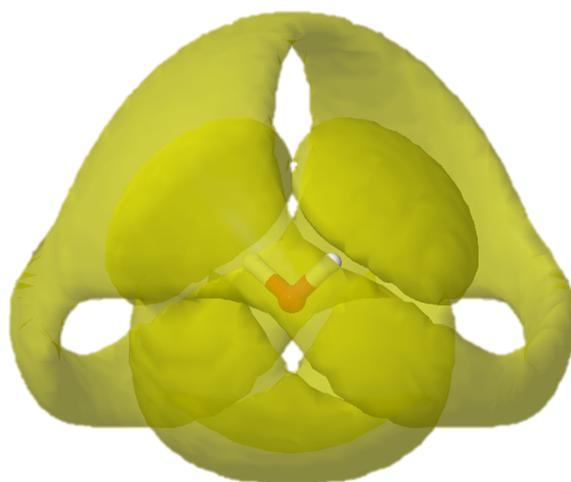


Figure 2.2.3: The spatial density function for pure water at ambient conditions. The yellow isosurface contains the 30% most likely positions for neighbouring water molecules around a central water molecule.

be explained in this thesis, however a full description can be found in reference [201]. The end result is the spatial density function, which shows the most likely positions of neighbouring molecules relative to a central molecule type. An example of this for pure water is shown in figure 2.2.3

Here the yellow areas are an isosurface which represent the 30% most likely positions for neighbouring water molecules to occupy around a central water molecule. This probability cutoff can however be tuned to whatever the user sees fit. In this example of pure water this allows the hydration structure to be more easily visualised and understood. We notice that the first hydration shell consists of two areas of high probability located directly above each hydrogen atom, and a second broader area of high probability located directly below the central oxygen. These three areas represent the first hydration shell. Using the water oxygen - water oxygen radial distribution function would also show this hydration shell, however as it is only a 2D representation of the system it would only be shown as a single peak and one would not be able to distinguish which areas around the water molecule are hydrated more strongly than others. The second shell then fills in at a larger distance in antiphase to the first hydration shell.

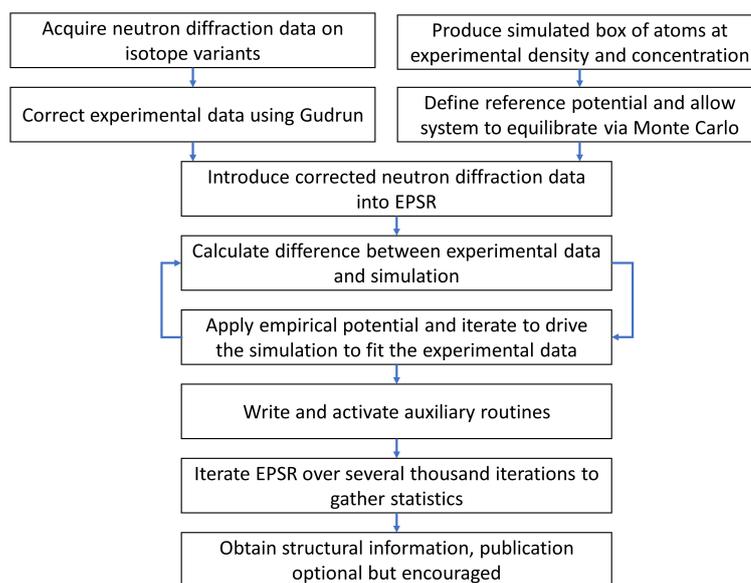


Figure 2.2.4: Flowchart depicting the general procedure followed when moving from the raw neutron diffraction data to atomic resolution structural information.

2.2.6 EPSR Summary

The journey from raw diffraction data to final structural information on the sample produced through EPSR is a long and arduous process that is bound to cause anyone who needs to understand or report it in detail to go mad. In an effort to simplify this understanding and avoid madness the process of EPSR can be shown in the flow chart in figure 2.2.4.

EPSR is a beautiful technique that has helped to address a host of systems, as described briefly in section 2.2. However, as versatile and detailed as it is in its current form, there are some things that EPSR is not capable of doing. Its main weakness is that it addresses every molecule in the simulation, so if one wishes to examine particular microenvironments in EPSR in detail, measure the abundance of particular conformations of bonded molecules, or quantify phenomena relevant to aqueous solutions such as hydrogen bonding, one requires an even more detailed tool. In the absence of one, I have had to come up with my own. Its function and capabilities will now be described in section 2.3 and the associated subsections.

2.3 Custom Analysis Routine

The custom analysis routine employed in this thesis was designed to investigate very particular conformations of molecules and microenvironments in EPSR simulations that EPSR itself is not capable of calculating [2, 3]. This allows us to investigate subtle differences between various aqueous solutions that would not be immediately obvious using conventional EPSR analysis. It was written in Python and functions by first reading in the .ato file produced by EPSR, which contains the Cartesian coordinates of every atom present in the system. It then uses a series of cutoff distances between particular atom types to populate lists that identify molecules in various conformations including: molecules that fall within the first hydration shell of a central molecule, molecules that are donating hydrogen bonds to a central molecule, and molecules that are hydrogen bonded in particular conformations. The routine then uses the lists to calculate bond angles, dipole angles, bond energies, and conformation abundances. Finally the routine exports the key results to .csv files to they can be fitted and plotted further in other programs such as Excel and Origin. This process will now be described in more detail in the following sections.

2.3.1 Reading in .ato File

The routine begins by first reading in the .ato file produced through EPSR after each iteration. All files produced by EPSR are relatively simple text files with different custom extensions only applicable to EPSR. The format of the .ato file is described below [201]. The first two lines contain parameters used to control the whole simulation, then the remaining lines describe the locations of individual molecules.

Line 1: Gives the number of molecules, width of the cubic box, and temperature

Line 2: Contains a series of parameters used for running the EPSR simulation, such as the value for $C/2$, step sizes for molecular motions, etc.

Line 3: Gives the x, y, z coordinates for the centre of mass of the first molecule,

along with other parameters that describe whether the molecule is tethered, how many unsuccessful moves it has had until this point, and the molecule number. This line is then followed by a series of other lines which describe the atoms in the molecule.

Line 4: Gives the atom type and its number within the molecule.

Line 5: Gives the x, y, z coordinates for the atom in the molecule relative to the centre of mass of the molecule.

Line 6: Details which other atoms within the molecule the atom is bonded to and their separations. Lines 4-6 are then repeated for every atom in the molecule.

Line 7-10: Details about rotational and dihedral groups present in molecule. Gives the number of groups, their constituent atoms, and the axis about which the rotation will take place or their dihedral angle as appropriate. For the small biological molecules used in this thesis no head-group rotations were permitted. The lines 3-10 are then repeated for every molecule in the system.

Line 11-15: Parameters for reference potentials of atom types, whether they are exchangeable, and atomic masses.

For the analysis code only lines 3 and 5 are important, as these allow for the calculation of the x, y, z coordinates for every atom present in the system. To extract these the values lists are produced containing the line numbers corresponding to lines 3 and 5 of each constituent molecule. Using these numbers as a guide the code then reads in the related x,y,z coordinates and appends these to new lists. The centre of mass absolute coordinates and the atomic coordinates relative to the centre of mass of the molecule are then added as appropriate to produce the final absolute atomic coordinates lists.

The nature of a simulated box of a molecules means that the molecules at the edge of the system will inevitably be undercoordinated. EPSR avoids this problem by employing periodic boundary conditions, however this is not included in this analysis routine. In order to avoid this distorting potential results the edge cases are removed by creating a new list only populated by water molecules whose x, y, or z coordinate does not lie within

a distance corresponding to the width of a water molecule hydration shell (calculated to be 3.38 Å through EPSR) from the edge of the box. Later in the routine when bond angles, hydration energies, etc. are being calculated only these water molecules are allowed to act as central molecules.

The code then calculates water molecules that lie within a central molecule's hydration shell. It does this by measuring the distance between a central atom and the water oxygen for every other water molecule in the system. If this is below a threshold value, typically corresponding to the width of the first hydration shell calculated through EPSR, then this is deemed to be a hydration water molecule. For single atoms and water molecules this can be done by specifying a single atom type. For molecules of a more complex shape, such as trimethylamine N-oxide, the hydration shell is identified by using the oxygen headgroup and methyl group carbons as central atoms, and then combining the resultant hydration shells. The final result is for every constituent central molecule, be it water, biological molecules, ions, etc., a list is produced containing the coordinates of every water molecule in its first hydration shell. It is then these coordinates that are used to calculate bond energies, angles, etc.

Water molecules that do not belong to the first hydration shell of another molecular species (TMAO, ions, etc.) are classified as “bulk” water molecules. These are differentiated to study whether the introduction of solutes perturbs the structure of water outside their first hydration shell. This is of particular importance when one is studying perturbations to water-water hydrogen bonding. If water molecules that lie within the first hydration shell of a solute molecule/atom are allowed to act as central molecules for further analysis, then strong interactions between the water molecule and the solute can mean that the central water molecule can be in energetically unfavourable conformations relative to other neighbouring water molecules. This means that the water-water interactions will be calculated as being less stable (higher interaction energy) than is representative of the system.

2.3.2 Hydrogen Bonding

In order to investigate hydrogen bonding in aqueous solution, one must first settle on a definition for a hydrogen bond. Hydrogen bonding occurs when a single positively charged hydrogen which is covalently bonded to a parent molecule is electrostatically attracted to another negatively charged atom on a neighbouring molecule [28, 75, 93]. Hydrogen bonds can take a broad range of bonding energies and are often very short lived, on the order of picoseconds [25–27], but can also persist for much longer, such as the hydrogen bonds that hold together DNA molecules which can last for hundreds of years [28, 29]. Because of these broad characteristics, pinning down an exact definition of a hydrogen bond is difficult [206]. During the writing of this analysis routine several definitions of hydrogen bonding were considered from previous literature.

Several simulation based techniques have used definitions of hydrogen bonding between water molecules which require the fulfilment of both distance and angular criteria. For example, Galamba *et al.* employed a definition that stated that two water molecules would be considered hydrogen bonded if the distance between them was less than 3.3 Å and the O-OH angle θ , as described in figure 2.3.1, was less than 30°. Other research groups have used similar criteria [93, 96, 99, 114, 119, 207].

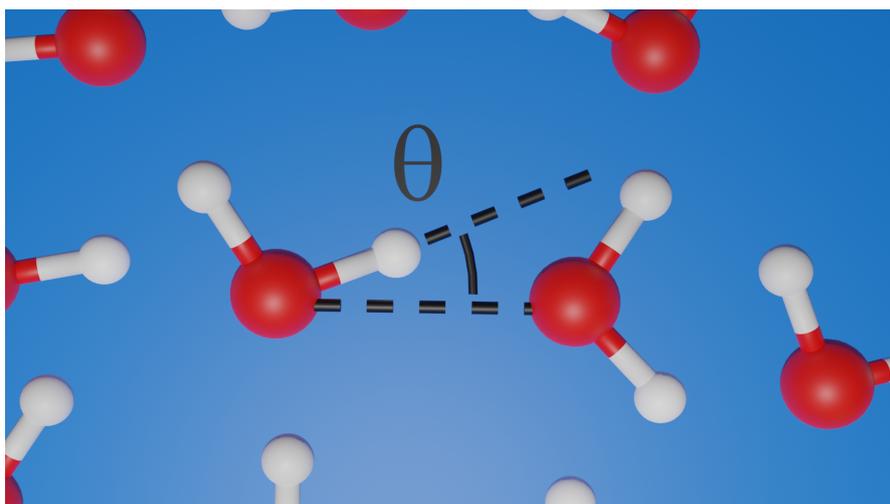


Figure 2.3.1: The O-OH angle used by Galamba *et al.* to define hydrogen bonding between water molecules.

Conversely rather than use geometric criteria, some have chosen to employ an energetic cutoff, where two water molecules are deemed to be hydrogen bonded if their total interaction potential energy is lower than a threshold value. Vibrational spectroscopy by Auer *et al.* suggests a energy threshold of -12.9 kJ/mol [95], however other energetic cutoffs have also been employed [27, 88, 93]. Others have also employed a topological argument, where a pair of water molecules are deemed to be hydrogen bonded if the distance between the oxygen on the central molecule and the hydrogen on the hydrating molecule satisfies two conditions: the oxygen must be the nearest non chemically bonded neighbour of the hydrogen, and the hydrogen is the first or second intermolecular near neighbour of the oxygen [94, 96].

The hydrogen bond definition used in the analysis routine is most similar to the topological definitions and the geometric condition employed by Bandyopadhyay *et al.* [114]. Here we say that a water molecule is hydrogen bonded to a central water molecule if it simultaneously satisfies two criteria: the distance between the central water molecule oxygen and the hydrating water molecule oxygen must be less than a distance corresponding to the first minimum in the $g_{OO}(r)$, and the distance between the central water molecule oxygen and the hydrating water molecule hydrogen must be less than a distance corresponding to the first minimum in the $g_{OH}(r)$. A similar definition is used when investigating hydrogen bonding between water molecules hydrating a central TMAO molecule. In this case, the TMAO oxygen is treated as the central atom type, rather than the central water molecule oxygen. This allows the definition of hydrogen bonding to be more flexible and sample specific, as it is informed by the EPSR simulation for each sample. The routine iterates this definition over every central molecule in the box to identify each water molecule that is donating a hydrogen bond to each central molecule in the system.

Once this condition is satisfied the analysis routine can count the number of hydrogen bonds per central molecule, and calculate hydrogen bond angle θ distributions according to the bond angle shown in figure 2.3.1. This is done according to equation 2.3.1, where r_{OH} is the OH bond distance (set at an average value of 0.96 Å within EPSR), r_{O-O} is the distance between the hydrogen bonded water molecule oxygens, and r_{O-H} is the

distance between the donor water molecule hydrogen and the central molecule oxygen. This is done for every hydrogen bonded molecule pair and the results binned. In the case of water-water hydrogen bonding, where many water molecules in each simulation box mean that large numbers of statistics are gathered relatively quickly, the bin width can be set as small as 0.5 °, but for TMAO-water hydrogen bonding, where far fewer TMAO molecules means that statistics are not gathered as quickly, the bin width is set to 2 °. These bin widths produce relatively smooth distributions over 20-30 iterations of the analysis routine. The binned results are then normalised to the total number of TMAO/water-water hydrogen bonds detected in the system so the resultant distributions represent the probability of finding a hydrogen bond of angle θ between two hydrogen bonded molecules.

$$\theta = \cos^{-1} \left(\frac{r_{O-O}^2 + r_{OH}^2 - r_{O-H}^2}{2r_{O-O}r_{OH}} \right) \quad (2.3.1)$$

The routine then identifies the relative abundance of specific hydrogen bonded conformations. These serve as another way of monitoring structural perturbations in detail, but are also included to try and obtain a fingerprint of what may be happening to the dynamics in the system. It is suggested that when a water molecule switches its hydrogen bonding partner that this goes through a less stable intermediate conformation [96, 155]. These routes that occur via hydrogen bonding intermediates are shown in figure 2.3.2. By monitoring how frequently these intermediates are occurring we can therefore gain potential insight into whether the dynamics of the system are increased or decreased, as faster dynamics will result in more frequent bond switching and therefore a greater abundance of hydrogen bonding intermediate conformations.

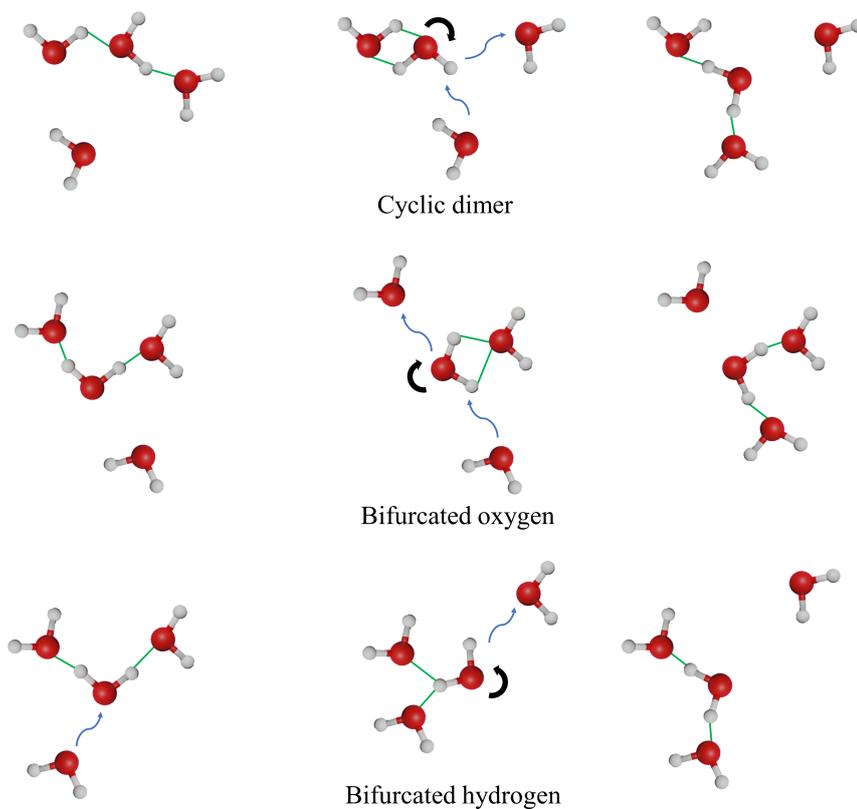


Figure 2.3.2: A central water molecule switching its hydrogen bonding partners through a cyclic dimer, bifurcated oxygen, and bifurcated hydrogen. Hydrogen bonds are shown in green, molecular rotations are shown by black arrows, molecular diffusion is shown by blue arrows.

2.3.3 Interaction Strength Calculation

The routine is then capable of calculating the interaction strength between molecules under the different conditions:

1. Pairs of molecules that are hydrogen bonded
2. Pairs of molecules that lie within each others first hydration shell
3. The total interaction strength between a central molecule and every other molecule in its first hydration shell

Once these cases are identified the interaction energy between two molecules is calculated using the reference potential used in EPSR. The interaction energy is therefore the sum of the van der Waals and Lennard Jones potential between each pair of atoms α and β in both molecules, as described in equation 2.3.2. This means that the interaction energy calculated is purely enthalpic.

$$U_{Int} = \sum_{\alpha \neq \beta} \left[U_{\alpha\beta}^{LJ}(r) + U_{\alpha\beta}^q(r) \right] = \sum_{\alpha \neq \beta} \left[4\epsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r} \right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r} \right)^6 \right] + \frac{q_{\alpha}q_{\beta}}{4\pi\epsilon_0 r} \right] \quad (2.3.2)$$

Applying this calculation to molecules that satisfy each of the three cases listed above therefore corresponds to a different type of interaction energy:

1. The total interaction energy of the hydrogen bond
2. The total interaction energy between a central molecule and another molecule in its hydration shell
3. The enthalpy of hydration

As previously, the final results are then collected, binned, and normalised for plotting with bin widths set to 1 kJ/mol.

2.3.4 Bond and Dipole Angles

The final calculation done by the routine is the dipole angle distribution for water molecules around a central molecule species. This is defined as shown in figure 2.3.3. The dipole on an individual molecule is the vector which points from the area of net negative charge to the area of net positive charge. In the case of the water molecule this means that the dipole is parallel to the vector which originates at the oxygen and bisects the two hydrogens, as shown by the white arrows in figure 2.3.3. The dipole angle is therefore the angle between the vector which points from the central atom to

the hydrating water molecule oxygen, and the vector corresponding to the hydrating water molecule dipole.

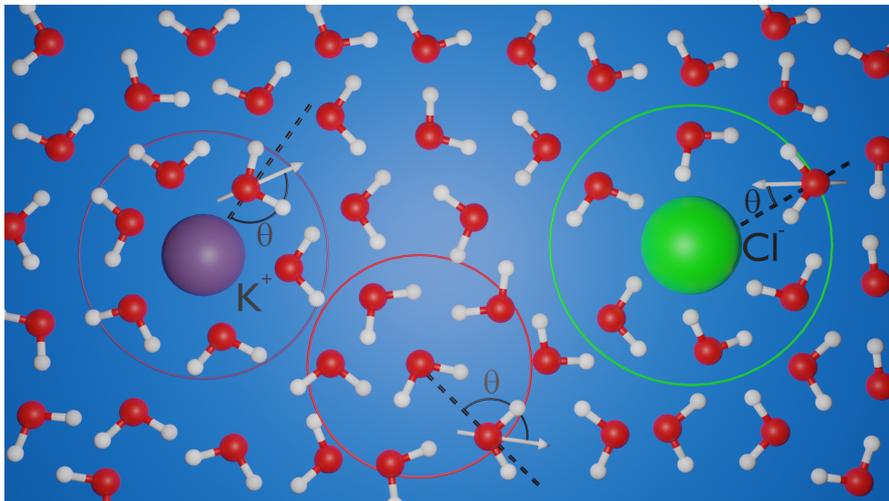


Figure 2.3.3: Definition of the dipole angle for a water molecule around a central molecule species shown for a potassium ion (purple), a chlorine ion (green), and a bulk water molecule (red).

This is calculated for each water molecule that lies within the first hydration shell of the central molecule species according to equation 2.3.3, where r_{X-O} is the distance between the central atom and the hydrating water oxygen (in the example in figure 2.3.3 X is either a central water oxygen, the K^+ ion, or the Cl^- ion), r_{OHH} is the distance between the hydrating water oxygen and the point exactly half way between the two hydrating water hydrogens, and r_{X-HH} is the distance between the central atom and the point exactly half way between the two hydrating water hydrogens.

$$\theta = \cos^{-1} \left(\frac{r_{X-O}^2 + r_{OHH}^2 - r_{X-HH}^2}{2r_{X-O}r_{OHH}} \right) \quad (2.3.3)$$

As always the results are then binned and normalised for plotting, in this case using bin widths of 2° .

2.4 Nuclear Magnetic Resonance

To validate the structural information gained by neutron scattering and EPSR and begin to consider dynamic perturbations to water alongside structural perturbations we employ the technique of nuclear magnetic resonance (NMR) spectroscopy. This is a technique which is sensitive to the magnetic moments of individual nuclei present in the sample and whose signal is perturbed by the local environment in which individual nuclei find themselves [168, 208]. It is also particularly sensitive to the signal generated by hydrogen nuclei. It is therefore another powerful method to monitor structural and dynamic perturbations on the atomic length scale and serves as a good compliment to EPSR. The background theory and relevant NMR techniques will now be overviewed in the following sections.

2.4.1 Spin

The phenomenon by which NMR is capable of function is due to the intrinsic angular momentum of nuclei, referred to as “spin”. This is a quantum property and does not relate to any physical movement of the particle itself [28]. Like energy, spin can be shown to be quantised, and it can be demonstrated that the magnitude of spin adopts discrete values according to equation 2.4.1 [168], where l is the “spin quantum number”.

$$\text{Magnitude of spin angular momentum} = \sqrt{l(l+1)}\hbar \quad (2.4.1)$$

In the case of fermions, such as protons, neutrons, and electrons, the l values can be half integers ($\frac{1}{2}$, $1\frac{1}{2}$, $2\frac{1}{2}$, etc.) whereas for bosons, such as photons and gluons, the l values can be integers (0, 1, 2, etc.). The spin angular momentum \vec{I} is a vector, and therefore has both magnitude and direction. In the case of protons and neutrons, where $l = \frac{1}{2}$ and they are therefore called spin $\frac{1}{2}$ particles, the magnitude of this vector given by equation 2.4.1 is $\frac{\sqrt{3}}{2}\hbar$. The projection of this vector onto an arbitrarily defined axis z is shown in figure 2.4.1. As the spin is quantised it is allowed to adopt $2l + 1$ projections.

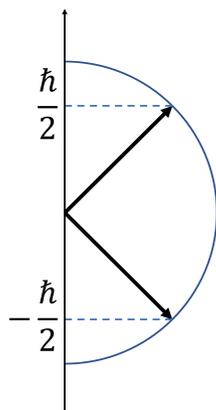


Figure 2.4.1: The projection of the spin angular momentum vector \vec{I} onto an arbitrarily defined z axis. For a spin with $l = \frac{1}{2}$ there are two allowed projections, whose projection onto the z axis I_z is equal to $\pm \frac{\hbar}{2}$.

This projection means that the spin for an individual proton/neutron can be $\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$, and are therefore either called “spin up” or “spin down” respectively. This can be simplified by saying that $I_z = m\hbar$ where m is the magnetic quantum number and has $2l + 1$ values equal to $l, l - 1, l - 2, \dots, -l$. As the nucleus is made up of protons and neutrons, the spin on the overall nucleus is equal to the sum of the spins on the individual protons and neutrons. Hence elements like hydrogen, whose nucleus comprises of a single proton, have an overall spin. However within the nucleus protons will pair with other protons with opposite spin, as will neutrons, and will cancel out. Therefore elements with an even number of protons and an even number of neutrons, such as ^{12}C and ^{16}O will show no overall spin and will therefore be invisible to NMR experiments. This factor is of little concern to us in this thesis as we will only concern ourselves with ^1H NMR.

The spin on a nucleus is then related to the magnetic moment of the nucleus by a constant of proportionality which is dependent on the species of the element, as shown in equation 2.4.2. This constant of proportionality γ is called the “gyromagnetic ratio”. This is therefore where the “nuclear magnetic” part of NMR comes from. Table 2.4.1 then compares the gyromagnetic ratio for NMR sensitive isotopes of elements relevant

to this thesis [168]. Here we see that hydrogen, like in neutron scattering, causes a particularly strong signal in NMR as a single proton has a large magnetic moment due to its large γ . The importance of this parameter will be explored further in the following sections.

$$\vec{\mu} = \gamma \vec{I} \quad (2.4.2)$$

Table 2.4.1: The gyromagnetic ratio for NMR sensitive isotopes of different elements. The negative value for ^{17}O is because the magnetic moment is in the opposite direction to the spin.

Isotope	γ ($10^7 T^{-1} s^{-1}$)
^1H	26.752
^2H	4.107
^{13}C	6.728
^{14}N	1.934
^{17}O	-3.628

2.4.2 Resonance Frequency

Now that we have defined and explained the quantum property of spin and its relation to the magnetic properties of a nucleus, we can discuss NMR. The first stage of any NMR experiment is the application of a large external magnetic field \vec{B}_0 on the order of T parallel to what will be defined as the z axis (one commonly employed magnetic field strength is $9.4 T$) [168, 208]. This has the effect of causing the magnetic moments on individual nuclei to align either largely parallel to \vec{B}_0 in the case of spin up nuclei, or antiparallel to \vec{B}_0 in the case of spin down nuclei, as shown in figure 2.4.2, where the magnetic moment μ in the nucleus forms an angle φ with \vec{B}_0 . In the absence of a magnetic field, all orientations of μ are equally likely and there is no overall magnetisation of the nuclei in the sample, however once a field is applied the population splits. This is known as the “Zeeman effect” [209].

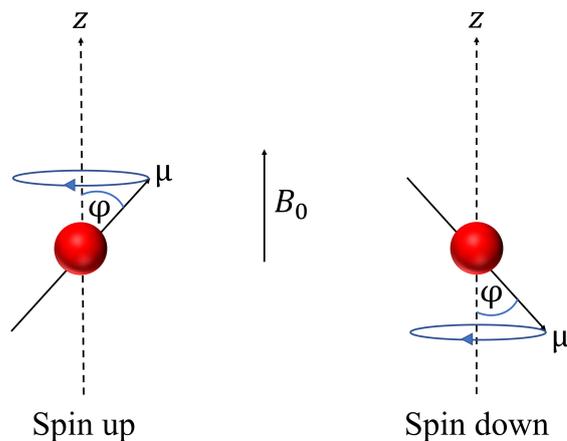


Figure 2.4.2: The alignment of the magnetic moment on a nucleus μ with the large external field \vec{B}_0 parallel to the z axis.

The resultant force felt by the magnetic moment as a result of the large external magnetic field is always perpendicular to the magnetic moment and the large external field [28], hence the magnetic moment will precess around the z axis at constant φ with an angular frequency ω_0 according to equation 2.4.3, where γ is the gyromagnetic ratio of the nucleus and B_0 is the magnitude of the large external magnetic field. This is known as the Larmor frequency. NMR spectrometers are then classified according to the Larmor frequency of a hydrogen nucleus (a single proton) at the operating magnetic field strength. In the case of a 9.4 T NMR, this would therefore be classified as a 400 MHz NMR.

$$\omega_0 = \gamma B_0 \quad (2.4.3)$$

Due to the Zeeman effect the nuclei split into two populations labelled spin up and spin down in figure 2.4.2. Initially these populations are equal and there is no net magnetisation on the sample. However the spin up state is slightly more energetically favourable, and there will therefore be a tendency towards greater abundance of spin up nuclei compared with spin down nuclei. This occurs through the process of spin-lattice relaxation and will be discussed further in section 2.4.4. The relative abundance

of these populations is then given by Boltzmann's law in equation 2.4.4, where $n(x)$ is the abundance of state x , ΔE is the difference in energy between the spin up and spin down state, k is Boltzmann's constant, and T is temperature.

$$\frac{n(\text{spin up})}{n(\text{spin down})} = \exp\left(\frac{\Delta E}{kT}\right) \quad (2.4.4)$$

Using the Planck-Einstein relation we can relate this energy difference ΔE to the angular frequency ω_0 and therefore the gyromagnetic ratio γ according to equation 2.4.5.

$$\Delta E = \hbar\omega_0 = \gamma\hbar B_0 \quad (2.4.5)$$

We will now introduce the concept that it is possible to swap the spin up and spin down states in the system by use of a second magnetic field. Upon the introduction of \vec{B}_0 , the magnetic moments on individual nuclei begin to precess with angular frequency ω_0 . If one then introduces a second magnetic field B_1 at a frequency ω_0 then the spin up states can absorb energy and become spin down states, and the spin down states can emit energy and become spin up states. The use of a second magnetic field at equal frequency to ω_0 is where the "resonance" part of NMR comes from. As the system began with an excess of spin up states, as shown by equation 2.4.4, there will be an overall absorption of energy. This absorption of energy of a particular frequency can then be detected, and this provides the final signal in the NMR experiment.

It is here that we can point out the power of hydrogen in the NMR experiment. As hydrogen has a large gyromagnetic ratio, equation 2.4.5 demonstrates that at a fixed external magnetic field strength B_0 then the energy difference between the spin up and spin down state will be large compared with other elements. This then means that the population of spin up relative to spin down will be greater in the case of hydrogen compared with other elements according to equation 2.4.4. A sample rich in hydrogen will therefore absorb a larger amount of the energy from the magnetic field B_1 , and will therefore generate a large signal. Hence NMR is ideal to study water and aqueous

solutions, as these will be extremely rich in hydrogen.

2.4.3 Chemical Shifts

We have now demonstrated that by the use of external magnetic fields we can generate a signal from a sample using NMR as it will absorb specific frequencies depending on the elements present in the sample. At first glance then NMR is useful for determining the elemental composition of an unknown sample, but not much else. However, much like everything else in physics, things are often much more complicated than they seem.

Another effect of the large external magnetic field \vec{B}_0 is that it will cause the electrons around a nucleus to circulate in their orbitals, as shown in figure 2.4.3. This will then result in the production of a second, far weaker (by a factor of $10^4 - 10^5$) magnetic field \vec{B}' which opposes \vec{B}_0 .

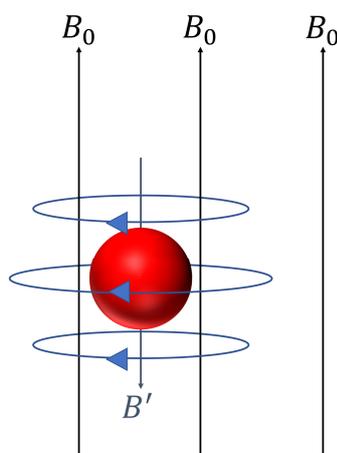


Figure 2.4.3: The application of the large external magnetic field \vec{B}_0 causes the electrons to circulate within their orbitals. This in turn induces a second weak magnetic field \vec{B}' which opposes \vec{B}_0 .

This means that the actual magnetic field experienced by the nucleus is not in fact B_0 , but is slightly smaller. This is expressed in equation 2.4.6, where B is the magnetic field strength experienced by the neutron as the result of the application of a magnetic field strength B_0 , and σ is known as the screening/shielding constant.

$$B = B_0(1 - \sigma) \quad (2.4.6)$$

This in turn means that the magnetic moment on the nucleus does not precess about z at ω_0 , but at the slightly lower frequency ω , and will therefore resonate at a slightly lower frequency. This phenomenon is known as the “chemical shift”. As the electron orbitals around a nucleus will depend on its immediate bonding partners and local chemical environment, this perturbation to ω_0 is therefore dependent on the local environment of the nucleus. Hence the NMR signal will not only be element specific, but now environment specific. This is the true power of NMR.

We now turn to plotting the signal produced by NMR. Because plotting the absolute resonance frequency and therefore the chemical shift is difficult in practice, the alternative is to plot the chemical shift relative to a known value. The most common choice is trimethylsilane, as this gives a single clear 1H peak. The NMR signal is then plotted as a series of peaks, corresponding to the resonance frequency of the constituent nuclei, normalised to this trimethylsilane peak. This is done according to equation 2.4.7, where δ is the location of the resonance peak, ω is the resonance frequency of the nucleus of interest, and $\omega_{0,TMS}$ is the resonance frequency of trimethylsilane. The factor of 10^6 is added by convention for convenience, and therefore the units are quoted as parts per million (ppm). The choice of trimethylsilane as the reference means that its peak occurs at $\delta = 0$. In this research we found it more practical to use dimethylsulfoxide, which shows a single clear peak at $\delta = 2.5$.

$$\delta = 10^6 \left(\frac{\omega - \omega_{0,TMS}}{\omega_{0,TMS}} \right) \quad (2.4.7)$$

This is best shown by an example 1H NMR spectrum of ethanol (C_2H_5OH), shown in figure 2.4.4 [210]. Here we observe three groups of peaks, each corresponding to hydrogens in different locations on the ethanol molecule (the outermost CH_3 group, the central CH_2 group, and the end OH group). The reason these peaks occur as groups rather than a single clear peak is due to spin-spin coupling, which is outside the scope of

this thesis. We can therefore use NMR to infer information about the local environment in which a particular nucleus sits, and use this to infer information about molecular structure.

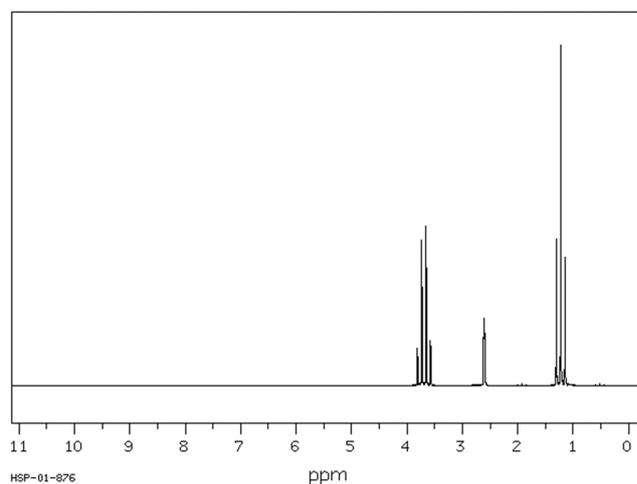


Figure 2.4.4: The NMR spectrum of ethanol. Figure taken from the Spectral Database for Organic Compounds [210].

The most useful origin of the peak shift effect for this thesis is its relation to hydrogen bonding. Stronger hydrogen bonds result in deshielding of the hydrogen nucleus and therefore a shift of the NMR peak downfield (to higher ppm) and *vice versa*. Classically this is explained by the movement of the hydrogen away from its host molecule due to a strong hydrogen bond, and therefore experiences less shielding from the electrons found in the bond between the hydrogen and its host molecule, however the true origin of this effect remains unclear. It is more likely to result from the distortion of the force field in which the proton sits, resulting in a shift of rotational and vibrational frequencies. We can therefore use NMR to examine the relative perturbing ability of different externalities, such as solute addition, to the hydrogen bonding ability of water molecules. This allows us to validate the structural results gained from neutron scattering, EPSR, and the analysis routine, which can measure the strength of hydrogen bonding between molecules as described in section 2.3.

2.4.4 Spin-Lattice Relaxation

Now that we have seen how NMR can be used to validate observations of structural perturbations in aqueous solutions, we will now demonstrate how it can also be used to reveal information about dynamic perturbations. One method by which this can be achieved is by a measurement of the spin-lattice relaxation time T_1 , a phenomenon without which NMR would be impossible. Spin lattice relaxation is the process which allows for the splitting of the nuclei into different spin states with a greater abundance of spin up nuclei as described in section 2.4.2.

Upon the application of \vec{B}_0 all the spins in the sample experience a force that causes their spins to align with (spin up) or against (spin down) the z axis. Initially these populations are equal, but the spin up state is slightly more energetically favourable. In the case of liquids while the magnetic moments on individual nuclei are all precessing about the z axis, the molecules are still rotating in solution. Due to frequent collisions with neighbouring molecules, the rate at which a single molecule rotates is sporadic and changes frequently, hence it is more commonly referred to as “tumbling”. The tumbling of a particular molecule is then quantified by the rotational correlation time τ_c . This is the average amount of time it takes for a single molecule to deflect by an angle equal to one radian and are typically on the order of 10^{-12} s (ps).

This tumbling movement means that individual spins are in motion and are therefore subject to time-dependent magnetic fields due to dipolar coupling with other spins. These time dependent magnetic fields will each occur at a frequency ω . As the molecular motions occur over a wide variety of timescales there also exist a wide variety of ω . The resultant concentration of fields at a particular ω is then called the spectral density function $J(\omega)$. This is given by equation 2.4.8 and shown in fig 2.4.5. Here we observe that the spectral density is largely flat and then decays to 0. This occurs because at small ω the value of $\omega\tau_c \ll 1$, and the expression simplifies to $J = 2\tau_c$, and at large ω the denominator grows exponentially. This initial flat portion is known as the extreme narrowing limit.

$$J(\omega) = \frac{2\tau_c}{1 + \omega^2\tau_c^2} \quad (2.4.8)$$

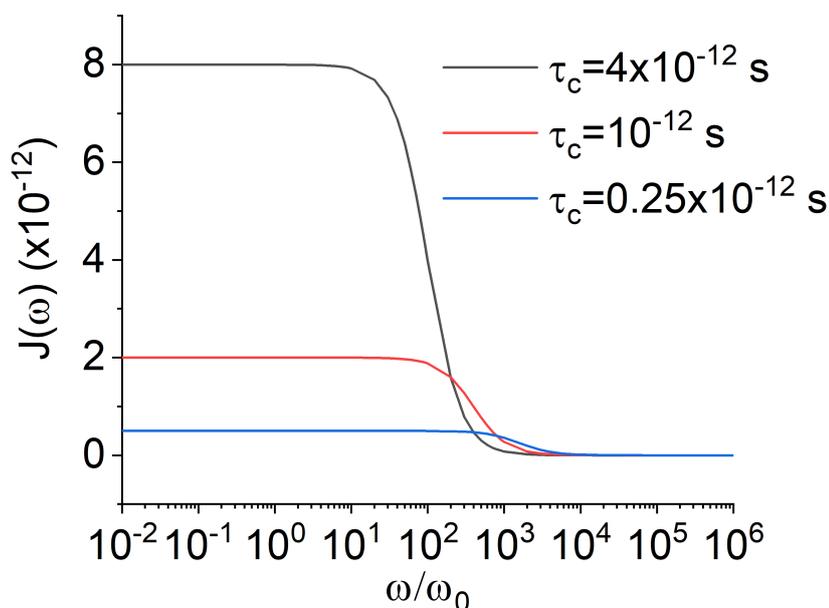


Figure 2.4.5: The spectral density function $J(\omega)$ for different values of τ_c where ω_0 corresponds to 400 MHz.

As explained in section 2.4.2, a spin can be made to flip by the application of a magnetic field of frequency ω_0 . Therefore if an individual spin experiences a time dependent magnetic field due to its relative motion between neighbouring spins at a frequency ω_0 , its spin can flip. This allows spin down electrons to flip into the energetically favourable spin up state by emitting energy that is dissipated through the lattice and the population of spin up electrons to build, as shown in equation 2.4.4. This makes NMR possible. The rate at which the population of spin up nuclei can build towards the equilibrium population difference Δn_{eq} is characterised by the spin-lattice relaxation time T_1 . The difference in population between spin up and spin down nuclei as a function of time $\Delta n(t)$ is given by equation 2.4.9 and shown in figure 2.4.6. It is for this reason that if one is to do an NMR experiment, one should wait for significantly longer than T_1 to make sure one has built up a significant population of spin up nuclei. In the case of

water $T_1 \approx 3$ s [211], hence this is hardly a major inconvenience.

$$\Delta n(t) = \Delta n_{eq} \left[1 - \exp\left(\frac{-t}{T_1}\right) \right] \quad (2.4.9)$$

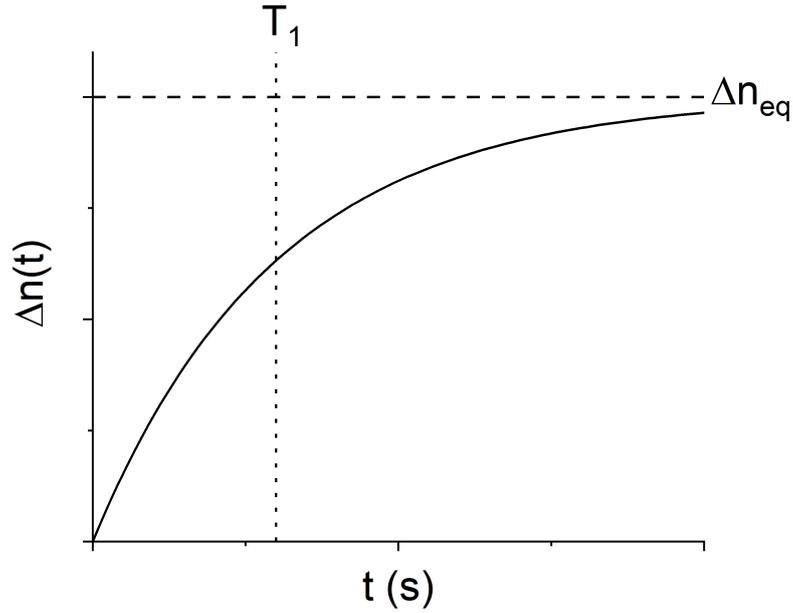


Figure 2.4.6: The population difference between spin up and spin down electrons as a function of time.

The rate constant for this process $1/T_1$ time will be proportional to the likelihood of experiencing a magnetic field of frequency ω_0 , equal to the spectral density at the Larmor frequency, $J(\omega_0)$. In the case of liquids τ_c is on the order of ps, hence $\omega_0\tau_c \ll 1$ and we are in the extreme narrowing limit. This means that $J(\omega) \approx 2\tau_c$, and therefore $1/T_1 \propto \tau_c$, as shown in equation 2.4.10. Hence if one measures the T_1 relaxation time for the hydrogen nuclei belonging to water molecules in an aqueous sample one can calculate the rotational correlation time of the water molecules, and therefore obtain a measure of whether rotational dynamics of water molecules are being accelerated or retarded by different externalities [87].

$$\frac{1}{T_1} \propto J(\omega_0) \propto \tau_c \quad (2.4.10)$$

In practice this is achieved by using the NMR technique of inversion recovery. At the beginning of the experiment the large external magnetic field \vec{B}_0 has been applied, and the greater abundance of spin up nuclei has been allowed to build to Δn_{eq} . The net magnetisation on the whole sample is therefore parallel to the positive z axis. An intense magnetic pulse is then applied close to ω_0 parallel to the y axis. This exerts a force on the spins and causes the net magnetisation of the whole sample to tip away from being parallel to the z axis, through the zy plane. The duration of this pulse therefore controls by how much the net magnetisation on the sample will tip. Appropriately done this pulse can therefore causes the population of spin up and spin down nuclei to invert, and the net magnetisation on the whole sample becomes parallel to the negative z axis.

Hence the signal generated by a particular nucleus also inverts as energy is now being overall emitted at a frequency ω_0 rather than absorbing at ω_0 . Once the pulse is complete the nuclei will again feel a force due to \vec{B}_0 and will gradually return to the original Δn_{eq} through spin-lattice relaxation, and the signal will also recover. One can therefore monitor the NMR signal of a particular nucleus following the inversion pulse to determine the T_1 relaxation time. The raw data is fit to equation 2.4.11, where $M_z(t)$ is the magnetisation in the positive z direction, $M_z(0)$ is the net magnetisation before the pulse, and t is time following the pulse.

$$M_z(t) = M_z(0) \left[1 - 2 \exp\left(\frac{-t}{T_1}\right) \right] \quad (2.4.11)$$

T_1 measurements of the aqueous samples presented in this research were measured at 27 °C on a Magritek Spinsolve 43 MHz NMR spectrometer.

2.4.5 Diffusion Measurements

We have now demonstrated that NMR can be sensitive to the rotational motion of water molecules in aqueous solutions, however it can also be made to be sensitive to diffusive motions in aqueous solutions. This gives us a second means by which to quantify dynamic perturbations to water and is achieved using the principle of pulsed field gradient spin echo [212], which will be overviewed in the following section.

Diffusive motions of individual molecules in aqueous solutions proceed via Brownian motion as a result of the thermal energy of the system. These are characterised by the diffusion coefficient D which relates the average root mean square displacement r_{RMS} of an individual molecule to time t , as shown in equation 2.4.12.

$$r_{RMS} = \sqrt{2Dt} \quad (2.4.12)$$

In the special case of spherical objects diffusing through a liquid the diffusion coefficient D can be expressed via the Stokes-Einstein relation as shown in equation 5.6.1, where k is the Boltzmann constant, T is the temperature, η is the liquid viscosity, and r is the radius of the diffusing object. This applies in situations where the diffusing object is much larger than the constituent molecules of the liquid. However as the size of the diffusing object becomes comparable to the solvent molecules, the factor of 6 must be amended.

$$D = \frac{kT}{6\pi\eta r} \quad (2.4.13)$$

In this thesis we monitor diffusive dynamic perturbations by using NMR to calculate the D value for individual water molecules in different aqueous solutions. Increased dynamics then corresponds to an increased D value and means that the water molecules are individually more mobile.

As stated previously, this is measured using pulsed field gradient spin echo. To

explain this we will first overview the principle of spin echo. This is explained visually in figure 2.4.7. Before the spin echo pulse sequence begins the individual spins are precessing about the z axis. An individual spin can be described by a “spin vector”, hence in this case the individual spin vectors are all parallel to the z axis, and due to the imbalance of the up and down spins so is the net magnetisation. A 90° magnetic pulse is then applied which tips the magnetisation of the sample from being parallel to the z axis to being parallel to the y axis, however the spins are still precessing around the z axis. Due to local inhomogeneities in the magnetic field each spin will precess with a slightly different frequency ω , and hence individual spins will start to dephase by a phase difference δ . This means that the magnitude of the magnetisation in the xy plane will gradually decrease. This is known as spin-spin or T_2 relaxation. After a time t a 180° magnetic pulse is applied which causes the individual spin vectors to flip around the y axis. This means that rather than dephasing, the individual spin vectors now begin to rephase, as the spin vectors with higher ω will start to catch up to the spin vectors with lower ω . After another time t the spin vectors will again be completely back in phase ($\delta = 0$), and the magnitude of magnetisation in the xy plane will return to its original value. This rephasing produces a signal that can be detected by the NMR, and is known as the spin echo.

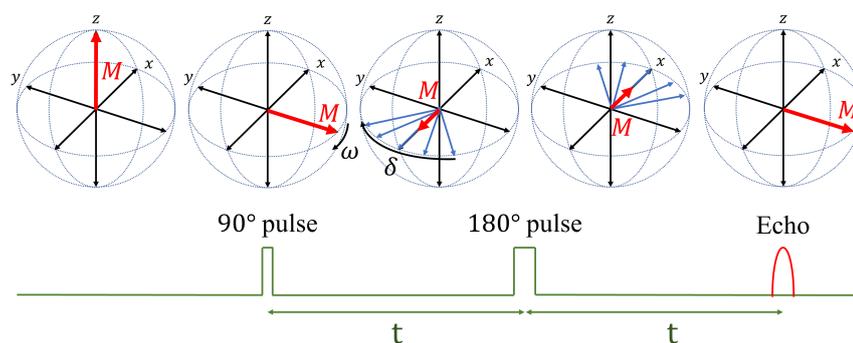


Figure 2.4.7: A visual representation of the spin echo pulse sequence. The net magnetisation is shown in red and individual spin vectors are shown in blue. The corresponding pulse sequence is shown in green, with the final spin echo shown in red.

A common analogy used to explain this sequence is a series of runners on a track. Before the pulse sequence all the runners start in the same place. The pulse sequence then begins at the 90° pulse and the runners set off each with a constant velocity. As the race progresses the faster runners will pull out ahead of the group and the slower runners will fall behind the group. This can be thought of as the group dephasing. After a time t the 180° degree pulse is applied and all the runners stop where they are, turn around, and run back towards the starting line. After another time t the runners should therefore all pass the starting point at the same time and produce a signal, which can be thought of as the spin echo.

Under the condition that \vec{B}_0 is uniform throughout the sample, and therefore all the nuclei of a given species should precess at ω_0 , all the individual spins should perfectly rephase following the spin echo pulse sequence. However, let us now consider the case where \vec{B}_0 is not uniform and instead depends linearly on the position along the z axis $B_0(z)$. This is known as a field gradient and is quantified by the value G , according to how strongly the field varies as a function of z . Therefore the Larmor frequency ω_0 of the individual nuclei is now dependent on their z position, and therefore each nuclei has effectively become “labelled” according to its physical position. Following the spin echo pulse sequence this means that the spins will only perfectly rephase if their z position is unchanged throughout the pulse sequence and they have experienced a constant B_0 and hence ω_0 . However if molecules are diffusing in the z direction this will not be the case and the spins will not perfectly rephase, and the final echo signal will be of a reduced intensity (attenuation). The attenuation of the echo signal will therefore depend on the field gradient G and how rapidly molecules in the sample are diffusing, measured by their diffusion coefficient D .

Returning to the analogy of the runners the effect of a field gradient can be thought of as rather than each of the runners running at a constant speed they are allowed to accelerate/decelerate throughout the race. After the 180° pulse the runners then turn around and run back to the start line, but their varying speeds mean that they will no longer all cross the start line at the same time.

The diffusion coefficient can then be related to the field gradient by the Stejskal Tanner equation [213], shown in equation 2.4.14, where I_G is the echo signal strength at gradient strength G , γ is the gyromagnetic ratio of the nuclei in question (hydrogen in this case), δ is the length of the pulses, and Δ is the time between pulses. The application of this principle means that the diffusion coefficient for individual water molecules in aqueous solutions can be calculated through NMR to examine translational dynamic perturbations.

$$I_G = I_{G=0} \exp \left[-(\gamma\delta G)^2 D \left(\Delta - \frac{\delta}{3} \right) \right] \quad (2.4.14)$$

2.5 Proteins

We have now discussed how we study the structure and dynamic perturbations to water as a result of different externalities in considerable detail, so we will now turn our attention to the study of biomacromolecules. In this thesis the considered biomacromolecules were the 93 amino acid β sheet protein I27 and the 10 amino acid peptide CLN025, which can be thought of as a “miniprotein” [214]. As outlined in section 1.1, proteins are natural polymers consisting of long chains of amino acids (residues) [18]. As they consist of both hydrophobic and hydrophilic residues upon exposure to water they attempt to minimise their free energy by folding into a shape that minimises the exposure of the hydrophobic residues to water and maximises the exposure of the hydrophilic residues to water. It is this shape that then dictates the proteins function.

In the following sections we will overview how the I27 proteins was produced, purified, and stored. I27 was selected as it is well suited to protein synthesis using *E. Coli*, it has been well previously characterised through mechanical [215–217] and thermodynamic studies [218–221], is relatively stable, exhibits reversible folding, and does not contain any disulphide bridges or cis prolines. In short, a pET vector containing the genetic code for I27, along with other genes to aid in protein production, was transformed into *E. Coli* cells. This was then grown up to suitable levels and induced to start producing

Table 2.5.1: Growth media and buffers used in the production and purification of I27.

Media/buffer	Components
Lysogeny broth (LB)	10 g bacto-tryptone, 5 g yeast extract, 10 g NaCl dissolved in 1 L of water
Terrific broth (TB)	12 g bacto-tryptone, 24 g yeast extract, 16.4 g K ₂ HPO ₄ , 2.3 g KH ₂ PO ₄ , 4 mL glycerol dissolved in 1 L of water
Lysis buffer	4.84 g Tris-HCl, 35.06 g NaCl, 300 μ L Triton X-100, 0.68 g Imidazole, 20 mL of 200 mM phenylmethylsulfonyl fluoride, 20 mL of 100 mM Benzamidine dissolved in 1.96 L of water corrected to pH 8
Wash buffer	4.54 g Tris-HCl, 35.06 g NaCl, 1.36 g Imidazole dissolved in 2 L of water corrected to pH 8
Elution buffer	1.21 g Tris-HCl, 8.77 g NaCl, 8.51 Imidazole dissolved in 500 mL of water corrected to pH 8
Cleavage buffer	3.03 g Tris-HCl, 29.22 g NaCl, 1 mL β mercaptoethanol dissolved in 1 L of water corrected to pH 8
TEV elution buffer	1.51 g Tris-HCl, 14.61 g NaCl, 0.5 mL β mercaptoethanol, 8.51 g Imidazole dissolved in 500 mL water corrected to pH 8
Investigation buffer	2.98 g HEPES dissolved in 500 mL water corrected to pH 7.4

the protein in large quantities [7]. The cells were then lysed to extract the protein, and subsequently purified. These steps will now be explained in more detail.

2.5.1 Media and Buffers

The various steps of I27 expression, purification, cleavage, and investigation require various different medias for the growing of bacterial cells and buffers for purification and study of proteins. The media and buffers used in this thesis are outlined in table 2.5.1 and are all dissolved in ultrapure Milli-Q water unless otherwise specified. pH values were corrected to the desired value using concentrated HCl or NaOH as appropriate. Buffers/growth media were prepared and then filtered using a 0.22 micron filter for sterilisation.

2.5.2 Transformation of Cells

The process of expressing, purifying, and cleaving I27 was done under close supervision by research technician Sophie Cussons. The first step is the transformation of the pET vector containing the gene which codes for the I27 monomer along with a gene which codes for ampicillin resistance, the lac I gene, the T7 RNA polymerase gene, and the lac operator supplied by Sophie Cussons [222]. The I27 gene contains the DNA sequence which codes for the I27 protein, along with a sequence that codes for a cleavage site between the protein and six concurrent His residues located at the N terminus. This is vital for later purification and cleavage. The cells are transformed by first defrosting 50 μL of competent *E. Coli* BL21(DE3)pLysS cells on ice. Cells had been previously made to be competent by suspending *E. Coli* cells in 100 mM CaCl_2 [223] and then transferring the cells into a solution of 100 mM CaCl_2 with 30% glycerol before chilling on dry ice and storing at $-80\text{ }^\circ\text{C}$. Cells prepared in this manner were generously donated by previous PhD students in the Dougan group. Once defrosted the cells were transferred to a sterile 0.5 mL eppendorf and 2-4 μL of the pET vector at a concentration of 100 $\mu\text{g}/\mu\text{L}$ was added. This is then left to incubate on ice for 20 minutes, before heat shocking the solution in a water bath at $42\text{ }^\circ\text{C}$ and returning to ice. The heat shock step is done to increase the permeability of the cell membrane which allows the *E. Coli* cells to uptake the pET vector.

2.5.3 Full-scale Expression

Once the cells contain the pET vector they must then be allowed to replicate. To do this 0.9 mL of sterile LB was added to the eppendorf containing the cells before incubating for 1 hour at $37\text{ }^\circ\text{C}$ while spinning at 200 rpm. The spinning is done to keep the solution well oxygenated and allow the cells to efficiently replicate. During this step the cells express the protein using the ampicillin resistant gene that allows them to be resistant to antibiotics. Following this 100 μL of the bacteria are spread onto an agar plate containing 100 $\mu\text{g}/\text{mL}$ of ampicillin using a sterile glass rod. The inclusion of the ampicillin in the agar plate means that any cells that have not successfully included the

pET vector, and hence cannot be used to express I27, are killed and only the useful cells remain. The agar plates are then incubated at 37 °C overnight. To be sure that the grow was successful the procedure was also carried out using pure water rather than the pET vector solution as a control. A successful grow is therefore one that yields roughly equally sized colonies on the sample plate and no colonies on the control plate.

A single colony is then removed from the agar plate and placed into a flask containing 20 mL of sterile LB, also containing carbenicillin at 100 µg/mL, to form a “starter culture”. This is incubated overnight at 37 °C and while being spun at 200 rpm. 10 flasks each containing 1 L of sterile TB are then prepared and each inoculated with 2 mL of the starter culture, which are again left to incubate at 37 °C while being spun at 200 rpm. During this incubation 1 mL of solution was removed every hour to measure the optical density using a UV spectrometer at 600 nm (OD_{600}). This is done as bacteria scatter light of wavelength 600 nm [224], and therefore by monitoring the OD_{600} it allows one to monitor the growth rate of the bacteria. This was performed until an OD_{600} reading of approximately 0.8 was reached, at which point the cells have reached the exponential phase of replication.

At this point 1 mL of 1 M isopropyl β -D-1-thiogalactopyranoside (IPTG) solution was added to each of the flasks such that the final concentration of IPTG is 1 mM. This causes the cells to over-express the I27 protein. The flasks are then left for a final incubation stage at 37 °C while being spun at 200 rpm for 18 hours. Finally the solutions are spun down in a centrifuge at 8000 rpm for 25 minutes at 4 °C. This forces the cells, now rich with the I27 protein, to the bottom of the solution where they form a pellet. The remaining supernatant is then discarded and the pellets retained.

2.5.4 Cell Lysis

Now the I27 protein must be extracted from the cells. To do this each pellet is resuspended in approximately 200 mL of lysis buffer. This causes the cells to break down and release the I27 protein, along with a number of other naturally occurring *E. Coli* proteins, into solution. Complete breakdown was further encouraged by homogenising

the solutions with an electronic disperser and leaving the solutions to incubate for 1 hour at room temperature. The resultant solution (lysate) was then passed through a cell disruptor operating at 30 kpsi and 25 °C. Finally the lysate was spun down in a centrifuge at 25000 rpm for 25 minutes at 4 °C. This causes the lysed cell debris to be forced to the bottom and form another pellet, while the proteins remain in the supernatant above. The supernatant was then collected and the pellets discarded.

2.5.5 Purification

We now have a supernatant containing the I27 along with a number of other proteins. In order to extract the desired I27 protein the lysate supernatant is loaded onto a 35 mL GE-nickel affinity chromatography column (HisTrap) using a peristaltic pump at a rate of 1 mL per minute overnight. The columns feature nickel (Ni^{2+}) bound to a matrix of highly cross linked agarose beads. It is here that the six His residues located at the N terminus of the protein come into play. Ni^{2+} ions have a strong affinity for imidazole, and so the six imidazole groups on the side chains of the six concurrent His residues will likely bind strongly to the column [225]. This means that the column retains the protein while the undesired proteins present in the lysate are washed through. The protein can then be knocked off the column by running through a concentrated imidazole solution which displaces the bound protein, given by the elution buffer.

The His column was then loaded onto an ÄKTA Prime. This provides two functions: the first is it allows two different buffers to be mixed at specified ratios before running them through the HisTrap such that the imidazole concentration can be well controlled by mixing the wash buffer with the elution buffer. The wash buffer contains a small concentration (10 mM) of Imidazole from the outset, this is included to displace any other weakly bound proteins that may contain a surface imidazole group. The second is that it monitors the absorbance of the solution exiting the His column at 280 nm (A_{280}). This is useful because proteins containing tyrosine or tryptophan, of which I27 contains one of each, absorb strongly at 280 nm. Therefore by monitoring the A_{280} it is possible to tell at which point proteins are being knocked off the HisTrap. As the ratio

of wash buffer to elution buffer is gradually increased the solution exiting the HisTrap is collected in a series of 5 mL vials. This means that when strong A_{280} is recorded by the ÄKTA Prime, one can isolate the solution containing that particular protein species.

The various protein species eluted must then be tested to isolate I27. This is done by analysing $10\mu\text{L}$ of each of the pooled protein species using Tris-tricine buffered sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE gel). A potential difference is applied across two different gels stacked on top of one another. This causes protein to migrate through the first gel (stacking gel) at a low current without entering the second gel (resolving gel) so that the protein samples all start from the same position. The current is then increased and the proteins travel through the resolving gel at a rate that is logarithmically dependent on their molecular weight. A sample containing proteins of calibrated weights (Mark12 Protein Standard, Invitrogen, UK) serves a scale by which to quantify the molar mass of each of the pooled protein samples. This allows one to isolate the sample containing I27, which can be further purified using size exclusion chromatography (SEC) if needed.

At this point the protein sample is still rich in salts from the elution buffer. To remove these the protein is then dialysed into pure water by loading the sample into Snakeskin dialysis tubing. This tubing is a semipermeable membrane which allows the exchange of small molecules, such as the various ions present in the sample, but retains the larger I27 protein molecules. The tube is then placed in a large amount of water and the ions allowed to equilibrate either side of the membrane due to osmotic pressure [28]. By repeatedly changing the water it is possible to reduce the concentration of ions in the protein sample to negligible amounts.

We now are left with a pure protein sample however it is at very low concentration. This can be measured by measuring the A_{280} and applying the Beer Lambert Law, as shown in equation 2.5.1, where ε is the molar absorptivity of I27, b is the length of the sample and C is the concentration in mol/L. The sample is therefore loaded into centrifugal concentrator tubes to increase the concentration. The sample is centrifuged at 4000 rpm until the concentration of protein has reached approximately 1 mg/mL.

Finally the sample is snap frozen using liquid nitrogen and lysopholized on a vacuum freeze drier, and stored at -80 °C. The final yield for this style of grow is typically approximately 200 mg of I27 protein.

$$A_{280} = \epsilon bC \quad (2.5.1)$$

2.5.6 His Cleavage

The final step is to remove the His tag so as to remove any effects this may induce in future studies of the protein. This is done with the application of the nuclear inclusion protease from tobacco etch virus (TEV protease), which recognises the cleavage site along a protein backbone by its peptide sequence ENLYFQ/G [226] and cleaves the protein between the Q and G residues. This was supplied by Sophie Cussons and expressed in a similar manner to that described by sections 2.5.1 - 2.5.5.

For this procedure 100 mg of I27 containing the cleavage site was resuspended in 10 mL of cleavage buffer. 8 mg of TEV protease, previously prepared and suspended to a concentration of 2 mg/mL in cleavage buffer and snap frozen in liquid nitrogen, was then thawed on ice before adding it to the I27 solution. This was then incubated at room temperature for one hour to allow the enzymatic reaction of the TEV protease to cleave the His tag from the I27. The resultant solution was then diluted in 150 mL of cleavage buffer before loading it onto a HisTrap and purifying using the ÄKTA Prime and SDS-PAGE gels as previously. This allows for approximately 85-90% of the protein to be successfully cleaved and recovered. The final cleaved protein solution was then dialysed into pure water using Snakeskin dialysis tubing and concentrated down to at least 0.5 mg/mL as previously. Finally this was placed into 1 mL aliquots, snap frozen using liquid nitrogen and stored at -80 °C.

2.6 Peptides

Peptides are short chains of amino acids shorter than approximately 20 residues, however there is no precise cutoff length. Just like larger proteins, peptides can also adopt secondary structure. The peptide used in this thesis was a 10 residue β hairpin peptide called CLN025 [214, 227]. This was selected as it was the smallest peptide that was found in the literature that adopted a stable secondary structure, which could be subsequently denatured. This denaturation behaviour means that CLN025 behaves as a “miniprotein”, however it is not biologically derived. This is important as a smaller peptide will be better suited for neutron scattering and EPSR experiments as it will have fewer interatomic correlations. The synthesis and purification of these peptides will now be overviewed.

2.6.1 Peptide Synthesis

Peptides were synthesised using microwave assisted solid-phase FMoc based synthesis in a CEM Liberty Blue peptide synthesizer. This is a method for growing peptides on a resin support one amino acid at a time [228]. This is extremely effective for growing smaller peptides, but becomes complicated if one wished to grow full proteins. FMoc is shorthand for the 9-fluorenylmethoxycarbonyl group that is used as a protecting agent during peptide synthesis. Solutions of the amino acids present in CLN025 with an FMoc group covalently bound to their N-termini were prepared at 0.2 M in dimethylformamide (DMF). A solution of resin beads is then also prepared in equal parts dichloromethane (DCM) and DMF which will act as the support for the growing peptide. The automated synthesiser then transfers these beads to the reaction vessel, where a 20% piperidine solution in DMF is added before a microwave step. The piperidine deprotects any FMoc groups on the resin such that a new amino acid can bind. This is then washed through. The amino acid is then added to the resin solution, along with a DMF solution of N,N'-Diisopropylcarbodiimide (DIC) and ethyl 2-cyano-2-(hydroxyimino)acetate (Oxyma). These chemicals act as a coupling agent which facilitates the formation of the peptide bond and a coupling agent to reduce racemization respectively. A microwave step is then included to promote bond formation, and the system is washed through to

remove waste. This procedure is iterated until the full peptide is synthesised. Finally the sample is removed from the synthesiser, washed through with DMF, resuspended in DMF, and 0.2 mL of acetic anhydride is added. The sample is then placed on a rotator for 30 mins, which allows the acetic anhydride to form a protecting acetyl cap on the unprotected N terminus.

At this point the peptide is synthesised but is still firmly attached to the resin. The peptide is then cleaved by washing the final resin through with DMF, DCM, and diethyl ether and dried under vacuum. 5 mL of cleaving cocktail, containing 92.5% trifluoroacetic acid (TFA), 2.5% water, 2.5% triisopropylsilane (TIPS), and 2.5% 3,6-dioxa-1,8-octanedithiol (DODT), is then added and the resin left on a rotator for 3 hours. This cleaves the peptide from the resin. The solution is then filtered to extract the peptide, and the resin is washed with TFA to wash off any remaining peptide. Finally 45 mL of diethyl ether at -20 °C is added to precipitate the peptide from the filtered solution. The separated solution is then spun down for 3 minutes at 6000 rpm, and the diethyl ether removed. The process is then repeated. Finally the diethyl ether is poured off and the peptide is resuspended in water to be freeze dried as described in section 2.5.5. These final peptides are then rehydrated and purified using high precision liquid chromatography coupled with mass spectrometry.

2.7 Fluorescence Spectroscopy of Proteins

Now that the CLN025 peptide and the I27 protein have been expressed and purified they can begin to be investigated. This is done by two methods in this thesis: fluorescence spectroscopy and circular dichroism. The former will now be outlined.

Fluorescence spectroscopy (in this case fluorescence emission spectroscopy) involves subjecting a sample to a beam of light at constant wavelength which acts as an excitation beam and monitoring the intensity of the emitted light as a function of wavelength to produce an emission spectrum [229]. The principle of fluorescence occurs as a result of the interaction between a photon of energy hf , where h is Planck's constant and f is

the frequency of the photon, and an incident atom/molecule. This is shown visually in figure 2.7.1. If the energy of the photon matches the energy difference between the ground state of an electron in the molecule S_0 and the first (S_1) or second (S_2) electronic vibrational state orbitals then the photon can be absorbed and the electron is excited into a higher energy state, as shown by the red arrows. Once in this excited state the electron can then relax via a number of mechanisms: These can be vibrational relaxations, which transfer energy between energy levels within a given electron orbital and are shown by the blue arrows, and the excess energy is released as heat. They can also be via internal conversion, in which the energy is transferred between electron orbitals with no loss of energy, and are shown as orange arrows. Finally the electron can spontaneously relax into its ground state orbital and the excess energy is emitted as a photon via fluorescence, shown by the green arrows. As the energy of the absorbed photon will be greater than that of the emitted photon due to the energy lost through internal relaxations, the wavelength of the emitted photon will be comparatively longer depending on which relaxation pathway was taken.

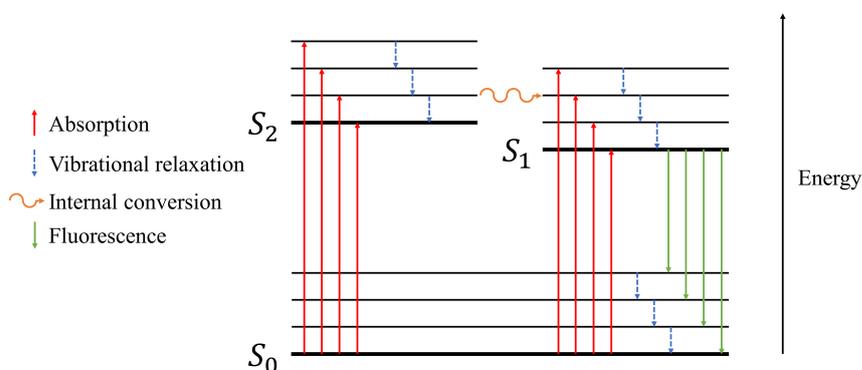


Figure 2.7.1: A Jablonski diagram demonstrating the different energy pathways of electron excitation and relaxation upon absorbance of an incident photon of appropriate energy.

When studying the I27 protein via fluorescence we choose an excitation wavelength of 280 nm, as tryptophan absorbs strongly at this wavelength and emits between 300-400 nm [230, 231]. The tryptophan on I27 is buried in the hydrophobic core of the

molecule [232], as shown in figure 2.7.2 and its fluorescence depends strongly on the chemical location of the side chain. As the protein denatures and the hydrophobic core of the molecule becomes hydrated its emission spectra will broaden and shift to longer wavelengths. Therefore by monitoring the emission spectra of I27 after excitation at 280 nm we can monitor protein unfolding. In this thesis fluorescence emission spectra were monitored on a PTI fluorescence spectrometer.

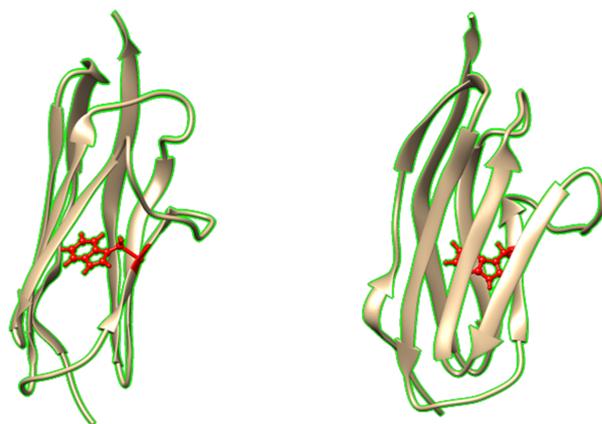


Figure 2.7.2: The cartoon structure showing the secondary β structure of I27 with the tryptophan atomic side chain structure highlighted in red.

In this research protein denaturation was induced through the addition of large quantities of urea. Urea preferentially associates at the protein surface and strongly reduces the protein-solvent van der Waals interaction energy [233], leading to protein denaturation [153, 234–236]. Increasing concentrations of urea will therefore cause the equilibrium of a system of aqueous proteins to tend towards an unfolded state. Urea also shows negligible contribution to the emission spectra at 300-400 nm at an excitation wavelength of 280 nm. If one therefore monitors the emission spectra of I27 at varying concentrations of urea one will observe a gradual change in the fluorescence signal as the population of unfolded protein relative to the population of folded protein begins to increase. This is therefore an equilibrium denaturation experiment. In the case of I27 there is a large difference between its emission signal in a folded and unfolded state at 320 nm. Therefore the emission signal was monitored at this wavelength taken a reading every second for

60 s to produce a signal to quantify protein unfolding.

Samples of I27 at varying urea concentrations were produced by first defrosting the I27 protein samples prepared as described in section 2.5.6 on ice. An investigation buffer stock + other solute and an investigation buffer + other solute + >8M urea stock were then prepared. The urea stock was made in excess of 8M because upon the addition of the I27 stock it would become diluted, and an upper concentration limit of 8 M was desired. The excess urea concentration was therefore calculated according to the concentration of the snap frozen I27. The other solutes are added at the desired concentration to monitor their perturbing effects on the stability of I27. From these 3 stock solutions a further 5 substock solutions were made of investigation buffer + 0.05 mg/mL I27 + 0, 2, 4, 6, 8 M urea by mixing the stock solutions accordingly. These substocks were then mixed to produce the final 0 - 8 M urea samples in increments of 0.2 M. The use of intermediate substocks reduces the error associated with mixing small quantities of stock solutions when one is producing the very high/low concentration samples, resulting in noisy data.

The resultant emission signal at 320 nm as a function of urea concentration is then analysed using the “two state model” of protein unfolding. As described in section 1.1.2, protein unfolding can be thought of as a two state process separated by an energy barrier with a difference between the Gibbs free energy of the folded vs the unfolded state of ΔG [37, 237]. This ΔG value can be calculated from the emission signal data according to equation 2.7.1, where I is the signal intensity, [denaturant] is the concentration of denaturant, in this case urea, in M , R is the ideal gas constant and T is absolute temperature. This equation essentially states that the emission signal of the totally folded and totally unfolded state are linearly dependent on urea concentration, hence the baseline at low urea concentration is fit to $(A[\text{denaturant}] + B)$, where A is the gradient of the pre-transition baseline and B is the intercept, and the baseline at high urea concentration is fit to $(C[\text{denaturant}] + D)$ where C is the gradient of the post transition baseline and D is the intercept. The folded signal is then weighted by the Boltzmann factor which itself scales with urea concentration by parameter m . This controls the steepness of the transition and is also shown to correlate with the difference between solvent accessible surface area in the folded and unfolded state.

$$I = \frac{(A[\text{denaturant}] + B) \exp\left(\frac{\Delta G - m[\text{denaturant}]}{RT}\right) + (C[\text{denaturant}] + D)}{1 + \exp\left(\frac{\Delta G - m[\text{denaturant}]}{RT}\right)} \quad (2.7.1)$$

Fluorescence spectroscopy therefore allows us to monitor the folded fraction of a protein as a function of urea and extract thermodynamic properties of the protein. By including other solutes at a constant concentration throughout this experiment we can then quantify their perturbations to protein stability by observing how they perturb ΔG . A protein stabilising agent would therefore increase ΔG and a protein destabilising agent would decrease ΔG . As it is only actually measuring the fluorescence of a single residue, it is only a global representation of the whole protein. If we wish to gain more specific insight into structural perturbations to the protein as a result of external stimuli, we need to turn to other optical techniques, such as circular dichroism.

2.8 Circular Dichroism

Circular dichroism spectroscopy is an optical technique which uses circularly polarised light and is sensitive to the chirality of structures present in the sample. This is useful for the study of proteins as the secondary structures present in a folded protein are chiral [238]. Circularly polarised light can be produced as shown in figure 2.8.1. This begins by polarising a monochromatic beam of unpolarised light using a linear polariser [28]. This means that rather than the electric and magnetic fields oscillating in all directions with equal probability, the electric field now oscillates in a single direction and the associated magnetic field oscillates perpendicular to it. We will now define the optic axis as parallel to the vertical direction. If the emergent light is at an angle to the optic axis, as shown in figure 2.8.1, where the emergent polarised light is at 45° to the optic axis, then it can be thought of as the sum of two waves. A polarised wave oscillating parallel to the optic axis, or extraordinary wave, or oscillating perpendicular to the optic axis, or ordinary wave. In the case of the polarised light shown in figure 2.8.1 these two waves are in phase

and perpendicular to each other.

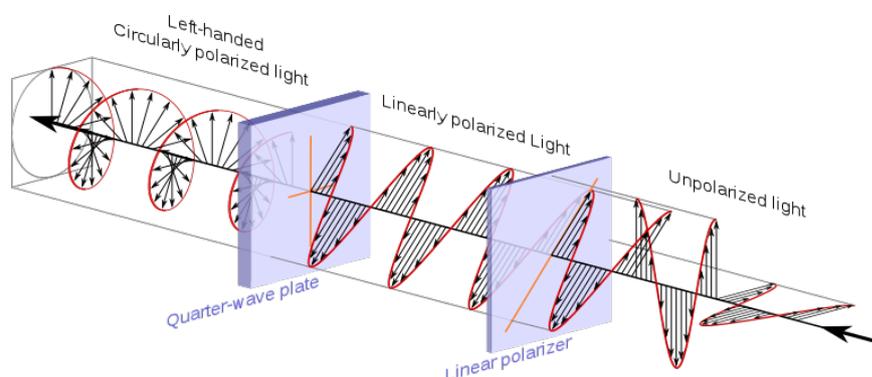


Figure 2.8.1: A schematic representation of the production of circularly polarised light from unpolarised light. Taken from reference [239].

The wave is then incident on a quarter wave plate. This is a birefringent material, which means that its refractive index is dependent on the orientation of the polarisation of the incident light. Light polarised parallel to the optic axis will be slowed to a different extent than light polarised perpendicular to the optic axis. The underlying mechanism of birefringent materials is due to the anisotropic nature of their molecular structure. This means that as the polarised light in figure 2.8.1 passes through the waveplate, the ordinary and extraordinary waves gradually become out of phase. If the plate is appropriately thick then the phase difference between the two waves will be 90° . This means that as the extraordinary wave intensity approaches 0, the ordinary wave intensity approaches its maximum, and *vice versa*. The vector therefore corresponding to the total electric field rotates around the direction of propagation, and the wave is said to be circularly polarised. The direction the electric field vector rotates then determines whether the emergent wave is right hand circularly polarised or left hand circularly polarised.

A CD spectrometer uses this principle to produce both left hand and right hand circularly polarised light. This is then incident upon a protein sample and the chiral structures in the protein will preferentially absorb left hand or right hand circularly polarised light depending on the structure and the wavelength of the incident light. If both components were absorbed equally the transmitted light would be plane polarised,

however in the case of preferential absorbance the transmitted light will be elliptically polarised. The spectrometer then records the difference in absorbance between the two components ($\Delta A = A_L - A_R$) and converts the result into an ellipticity value θ measured in degrees where $\theta = \tan^{-1}(b/a)$ and b and a are the major and minor axis values of the resulting ellipse. This is related to ΔA by $\theta = 32.98\Delta A$. A spectrum can then be taken by measuring θ as a function of wavelength. For I27 aqueous samples in the various media of interest were investigated at 0.2 mg/mL in investigation buffer using 1 cm quartz cuvettes on a Chirascan plus circular dichroism spectrometer (Applied PhotoPhysics). The final values determined from the experiment were converted to mean residue ellipticity according to equation 2.8.1 [238], where M is the molecular weight, R is the number of residues in I27, d is the depth of the sample, and c is the concentration in mol/L.

$$[\theta]_{mrw,\lambda} = \frac{M}{10 \times (R - 1) \times d \times c} \quad (2.8.1)$$

The resultant circular dichroism spectra (CD spectra) can be further deconvoluted to measure the relative abundance of each type of secondary structure within the protein. This is done using the Contin algorithm [240] within the online tool DichroWeb [241]. This works by refining supplied CD spectra against a large dataset of proteins of known secondary structure fractions using the locally linearised model. This helps to eliminate any structures that are dissimilar to the protein, which reduces the size of the dataset to a small number of similar proteins of known structure. These spectra are then convoluted to resemble the supplied data and measure the relative abundance of each secondary structure.

Chapter 3

Perturbations to Water Structure, Dynamics, and Thermodynamics due to Simple Monovalent Ions

3.1 Introduction

In this first results chapter I will investigate the effects of one of the simplest perturbations to water: the introduction of simple monovalent salts. Salts are groups of ions, in this case a pair between a halide anion and a potassium cation, that associate through an ionic bond. This bond occurs because a potassium atom can donate an electron to a halide atom, resulting in an oppositely charged pair which are electrostatically attracted to one another [28]. Salts then dissociate into their constituent ion when dissolved in water. This phenomenon has received widespread attention from a host of scientific communities due to the ubiquity of salt solutions in all areas of life. Most relevant to this thesis is the observation that different ion species can either stabilise or destabilise biomolecules. This was first observed by Franz Hofmeister in 1888 [43], who studied the relative ability of various ions to precipitate egg white lysozyme. This resulted in the first iteration of the “Hofmeister series”, which ranks ions by their ability to sta-

bilise/destabilise proteins and led to an explosion of literature in this area [24, 44–52].

In this section we will aim to achieve a detailed understanding of the effects of simple monovalent ions on water. Specifically we will seek to answer key questions concerning the study of aqueous salts, namely: over what range do ions perturb water structure, and can we meaningfully connect microscale perturbations present in salt solutions to their bulk properties? To do this we choose to investigate aqueous potassium halides. Halides are those elements which belong to group 7 of the periodic table, including fluorine, chlorine, bromine, and iodine. Potassium halides were chosen as they are ions with large biological relevance [175, 176], such as the uptake of K^+ ions in halophilic organisms to balance osmotic pressure [54] or the role of K^+ ions in insulin secretion [242], for which neutron scattering data is already available [128]. Choosing to investigate monovalent anions also provides us with the opportunity to understand relatively simple ion solutions in detail before the potentially more complex effects of multivalent/polyatomic ions are considered.

To achieve this understanding we will adopt a multi-scale approach. In section 3.2 neutron scattering and EPSR will be used to reveal atomistic scale structure present in aqueous potassium halides by producing site-site RDFs. In section 3.3 the custom analysis routine will be used to reveal more detailed structural information from the atomic coordinates produced through EPSR, namely dipole angle distributions of water molecules around bulk water molecules and ions. In section 3.4 we will then begin to consider thermodynamics of hydration of ions using the custom analysis routine. We will use this to calculate the enthalpy of hydration of the ions by considering particular microenvironments present in aqueous potassium halides and compare the findings to accepted literature values for the bulk solutions [35]. We will also monitor perturbations to bulk water hydrogen bonding, and calculate the distance over which ions have favourable enthalpic interactions with neighbouring water molecules. Finally, we will use NMR to quantify dynamic perturbations to water induced by potassium halide ions in section 3.5 by monitoring perturbations to the rotational correlation time and diffusion coefficient of the water molecules. We will also use the findings from sections 3.2 - 3.4 to justify the application of mathematical models [87] which deconvolute the ensemble averaged

dynamic perturbations to water produced by NMR into the specific contributions from the different microenvironments present in the solutions. By understanding the range over which simple monovalent salts like potassium halides perturb water structure and dynamics we can begin to comment on the mechanisms by which ions in solution interact with biomolecules and how this may lead to biomolecular stabilisation or denaturation.

3.2 Structure of Water - RDFs

In order to investigate structural perturbations to water due to the introduction of potassium halide ions, previous neutron scattering data on aqueous potassium halides taken by Weckstrom and Soper [128] on the Small Angle Neutron Diffractometer for Amorphous and Liquid Samples (SANDALS) instrument at the ISIS neutron and muon source was reanalysed using EPSR. In this study aqueous samples of KF, KCl, KBr, and KI were prepared at a concentration of 0.6, 1.2, and 2.4 mol/kg H₂O in H₂O, D₂O, and HDO. Isotopic variants of water were used such that EPSR analysis could be employed to deconvolute the resultant total structure factors into the site-site partial structure factors, as described in section 2.1.4. The raw data were then corrected for multiple scattering, attenuation, and inelasticity effects using Gudrun, as described in section 2.1.6.

Within EPSR a box containing 5000 water molecules and either 60, 120, or 240 potassium halide ion pairs were constructed for each of the four salt species at each of the three concentrations investigated such that the experimental concentrations were matched. At this box size and matching experimental densities this equates to a cubic box of dimension 53.2 - 55.2 Å. As described in section 2.2, EPSR calculates correlations up to a distance equal to half the box dimension, and therefore the several hydration shells would be observable through a box of this size. The analysis routines employed by EPSR also take averaged statistics over the entire simulation, hence a larger box size provides a larger sample size and more reliable data. The periodic boundary conditions employed by EPSR also mean that an infinite solution is better approximated with a larger box. This box size was therefore chosen as it was the largest possible box that

allowed the simulation to run in a reasonable timescale. The Lennard Jones potentials, Coulombic charges, solution densities, and sample temperature of 20 °C were taken from the original neutron scattering study and EPSR analysis on aqueous potassium halides by Soper *et al.* [128]. Full details of this are given in section A.

The total structure factors for the aqueous KF sample at 2.4 mol/kg H₂O for the experimental data $D(Q)$ and the EPSR simulation $F(Q)$ are shown in figure 3.2.1 as an example of the final results produced by EPSR for these solutions. The total structure factors for the other potassium halide salts and concentrations are shown in section A. From this data it is clear that the simulation is in good agreement with the supplied diffraction data. This is indicated by the R factor, as described in section 2.2.4, which for all four salt species at all concentrations studied has a value of approximately 0.0006. Previous EPSR simulations on aqueous L-glutamine and glutamine imino acid by Rhys [186] have had R factors between 0.0023 and 0.00016, and were all deemed acceptable quality fits. The R factors calculated for these simulations are within these bounds, hence the simulation is a good representation of the experimental sample and can be studied to provide meaningful results.

The resultant water oxygen - water oxygen RDFs produced through EPSR are shown in figure 3.2.2. Only the data for the highest concentration studied, 2.4 mol/kg H₂O, are shown for clarity. The remaining RDFs can be found in section A.

Here we can observe the tetrahedral structure of pure water, as indicated by a first peak located at approximately 2.8 Å and a second peak at approximately 4.5 Å. Upon the introduction of all four of the potassium halide salts studied here at a concentration of 2.4 mol/kg H₂O we observe a decrease in the height of the first peak as well as a movement inwards to approximately 2.7 Å. The second peak also moves inwards to 3.9 - 4.0 Å. The RDFs for the sample concentrations of 0.6 and 1.2 mol/kg H₂O are shown in section A and the peak positions for all studied concentrations are given in table 3.2.1.

This is similar to the perturbation to water structure induced by the application of an external pressure. The ability of ions to perturb water in this manner has been well examined previously [126–128, 130, 132, 133, 136] and is explained by strong favourable

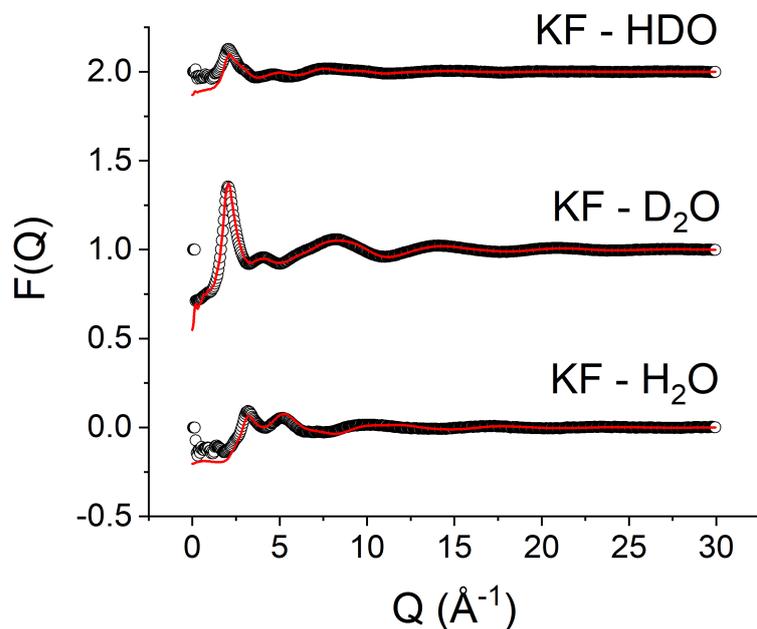


Figure 3.2.1: Supplied diffraction data $D(Q)$ from a previous study taken on SANDALS instrument [128] (hollow circles) and total simulated structure factor $F(Q)$ (red line) for aqueous KF at 2.4 mol/kg H_2O in H_2O , D_2O , and HDO . Each dataset vertically shifted by 1 for clarity.

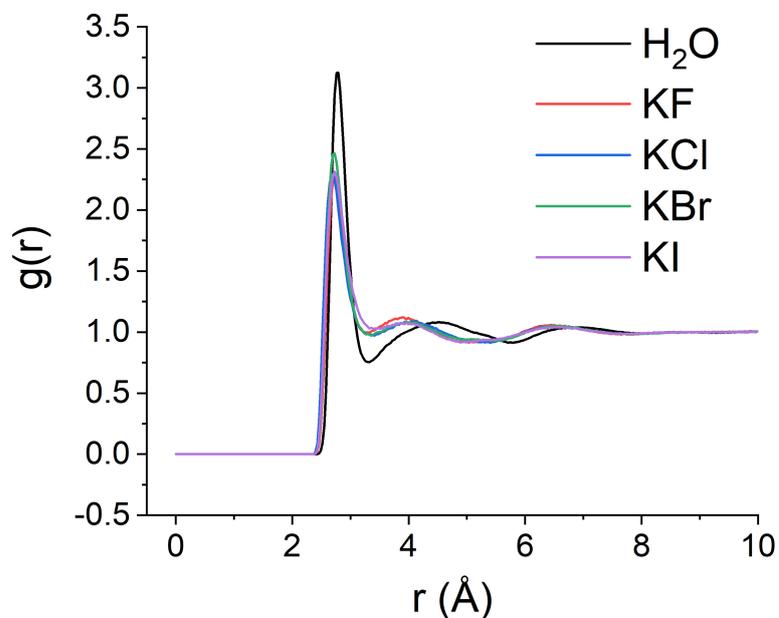


Figure 3.2.2: Water oxygen - water oxygen RDFs for pure water and aqueous KF, KCl, KBr, and KI. Only data for concentrations of 2.4 mol/kg H_2O shown for clarity.

Table 3.2.1: Peak positions of the first and second peaks in the $O_w - O_w$ RDFs for aqueous potassium halides calculated through EPSR.

System	First peak (Å)	Second peak (Å)
Pure H ₂ O	2.78	4.52
KF - 0.6 mol/kg H ₂ O	2.73	4.15
KF - 1.2 mol/kg H ₂ O	2.74	4.04
KF - 2.4 mol/kg H ₂ O	2.73	3.91
KCl - 0.6 mol/kg H ₂ O	2.74	4.33
KCl - 1.2 mol/kg H ₂ O	2.71	4.13
KCl - 2.4 mol/kg H ₂ O	2.69	4.05
KBr - 0.6 mol/kg H ₂ O	2.72	4.10
KBr - 1.2 mol/kg H ₂ O	2.72	4.10
KBr - 2.4 mol/kg H ₂ O	2.72	3.99
KI - 0.6 mol/kg H ₂ O	2.74	4.23
KI - 1.2 mol/kg H ₂ O	2.71	4.15
KI - 2.4 mol/kg H ₂ O	2.73	3.91

electrostatic interactions between the dissolved ions and the surrounding water molecules. Upon first glance it would seem that the different potassium halide species perturbed water in very similar manner to one another.

However, one sees much more ion specific perturbations to water structure if one considers the hydration shells of the individual ion species. Ions interact with neighbouring water molecules due to electrostatic interactions between the charged ion and the polar water molecule. This means that ions induce structure in the surrounding water molecules as they associate around the ion. In the case of negatively charged anions the surrounding water molecules will tend to orient a positively charged hydrogen inwards, whereas in the case of positively charged cations the surrounding water molecules will tend to orient both hydrogens outwards. The ion - water oxygen RDFs are shown in figure 3.2.3. Here one observes that as the size of the ion increases (F^- , K^+ , Cl^- , Br^- , I^-) and charge density therefore decreases that the first peak of the ion - water oxygen RDF shifts to larger distances, and becomes shorter and broader. The outward shift is simply explained by the increased radius of the central ion resulting in increased steric repulsion at short distances.

The reduction in peak height and increase in peak width is indicative of a less ordered

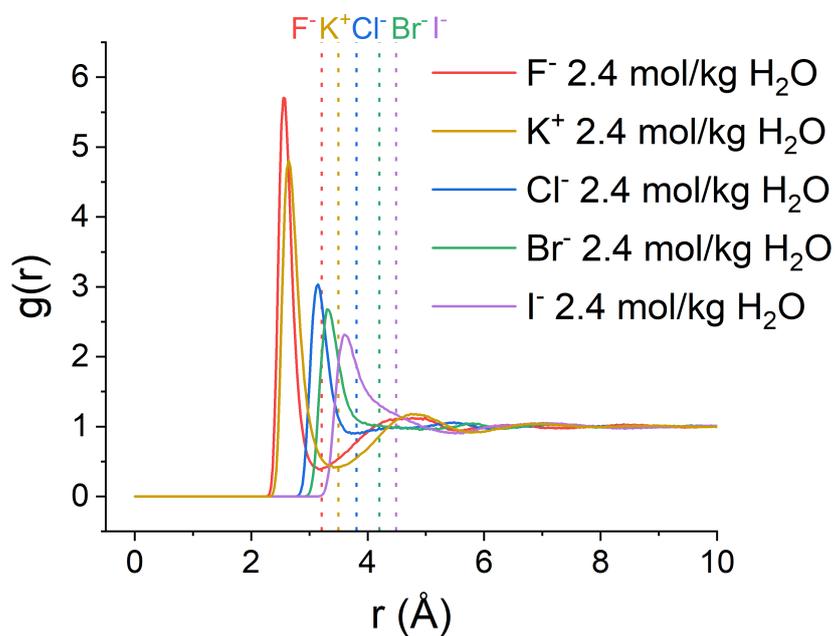


Figure 3.2.3: Ion - water oxygen RDFs for K^+ , F^- , Cl^- , Br^- , and I^- . Only data for concentrations of 2.4 mol/kg H₂O shown for clarity. Vertical lines indicate distance cutoff for water molecules in the first hydration shell of the ion.

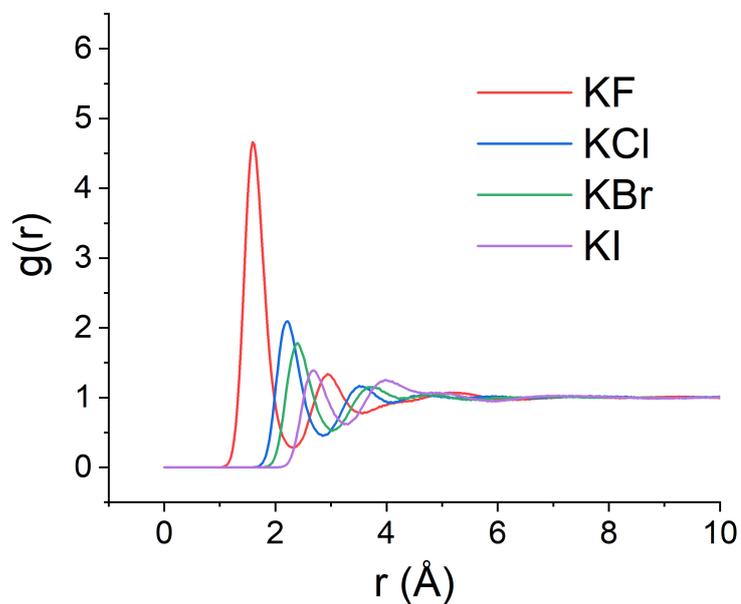


Figure 3.2.4: Halide anion - water hydrogen RDFs for F^- , Cl^- , Br^- , and I^- . Only data for concentrations of 2.4 mol/kg H₂O shown for clarity.

hydration shell of water molecules around the central ions. This trend is mimicked in the halide anion - water hydrogen RDFs, as shown in figure 3.2.4. Here the first peak occurs at shorter distances in the anion - water hydrogen RDF than the anion - water oxygen RDF for each of the four salt species, indicating the tendency of surrounding water molecules to orient a hydrogen towards the central anion. The reduction in height and increase in width of the first peak with increasing size of the halide anion suggests this tendency is reduced, and the anion hydration shell becomes increasingly disordered. The RDFs reported in this section are in excellent agreement with the results reported by Weckstrom and Soper [128].

3.3 Structure of Water - Dipole Angle Distributions

We can examine the degree to which the potassium cations and halide anions induce structure in their surrounding water molecules in more detail by considering the water dipole angle distribution in the first hydration shell of the ions using the custom analysis routine. The first hydration shell of an ion extends to a distance corresponding to the first minima in the ion - water oxygen RDF, as shown in figure 3.2.3. This method of monitoring structural perturbation has been previously applied to neutron scattering and molecular dynamics simulation techniques [128, 243]. As described in 2.3.4, the dipole angle for a water molecule hydrating a central ion is given by the angle between the vector which points from the ion to the hydrating water oxygen and the vector which is parallel to the dipole moment on the hydrating water molecule. The final normalised data (see section 2.3.2 for binning and normalisation procedure) is shown in figure 3.3.1(a) and figure 3.3.1(b) for the highest studied concentration of 2.4 mol/kg H₂O. The data for the 0.6 and 1.2 mol/kg H₂O samples are shown in section A. The resultant distributions were then fit using a function containing two summed Gaussian components, shown in equation 3.3.1 where y_0 is the vertical offset, A_x is the area under Gaussian peak x , w_x is the width of Gaussian peak x , and x_c is the peak location for Gaussian peak x .

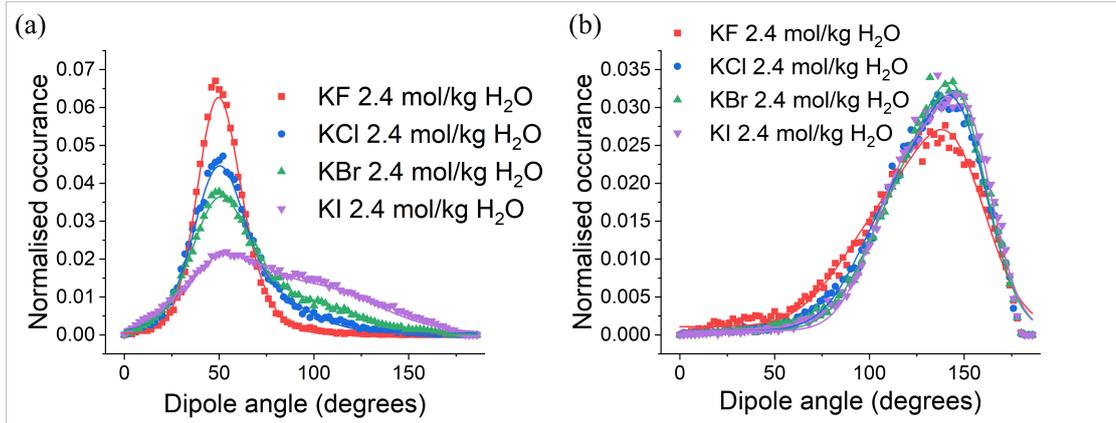


Figure 3.3.1: Distribution of dipole angles for water molecules hydrating a central halide anion (a) or potassium cation (b) within its first hydration shell. The dipole angle is defined as the angle between the vector pointing from the ion to the hydrating water oxygen and the vector corresponding to the hydrating water molecular dipole. Values normalized to the total number of data points. Only data for highest concentration 2.4 mol/kg H₂O samples shown for clarity.

$$y = y_0 + \frac{A_1}{w_1\sqrt{\frac{\pi}{2}}} \exp -2 \left(\frac{x - x_{c1}}{w_1} \right)^2 + \frac{A_2}{w_2\sqrt{\frac{\pi}{2}}} \exp -2 \left(\frac{x - x_{c2}}{w_2} \right)^2 \quad (3.3.1)$$

Figure 3.3.1(a) and figure 3.3.1(b) again show that all the ion species discussed in this work induce structure in the first hydration shell due to electrostatic interactions between the charged ions and the polar water molecules, as indicated by the presence of a single clear peak. This peak occurs at approximately 50 ° for water in the first hydration shell of halide anions, and at approximately 155 ° for water in the first hydration shell of the potassium cations. As described in section 3.2, this is due to the tendency of the hydrating water molecules to orient one hydrogen towards anions, and both hydrogens away from cations.

Within the halide anions (figure 3.3.1(a)), we observe that as one moves from F⁻ to I⁻ and ionic radius increases from 119 pm to 206 pm [244], that the peak in the halide anion water dipole angle distribution reduces in height and increases in width. We also observe the emergence of a second Gaussian component at higher angles, occurring at approximately 90 ° in the most extreme case of I⁻. Collectively this again demonstrates

an increasingly disordered hydration shell with increasing halide anionic radius.

When we consider the water dipole angle distribution around the potassium cations in each of the four aqueous salts (figure 3.3.1(b)), we observe that it remains essentially unchanged in the case of KCl, KBr, and KI, showing a single clear peak at approximately 155° . However, in the case of KF at 2.4 mol/kg H₂O we observe that the peak moves inwards to approximately 145° whilst simultaneously becoming shorter and broader. This only occurs at this highest studied concentration (see section A for 0.6 and 1.2 mol/kg H₂O data). At this concentration one can calculate that the average distance between ions to be 4.2 - 4.4 Å for all salt species (box size of approximately 168000 Å³ containing 240 ion pairs). As shown in figure 3.2.3, the first hydration shell of water around any of the ion species discussed in this chapter extends from 3.21 Å in the case of F⁻ to 4.49 Å in the case of I⁻. Therefore, at a potassium halide concentration of 2.4 mol/kg H₂O it is highly likely that a significant proportion of water molecules are in the first hydration shell of two or more separate ions, and are therefore experiencing strong electrostatic interactions from multiple sources, which in turn will perturb the water dipole angle distribution. In the case of F⁻ these electrostatic interactions are shown to be most impactful on water structure, as shown by the tall narrow peaks in the F⁻ - water oxygen RDF (figure 3.2.3) and the F⁻ water dipole angle distribution (figure 3.3.1(a)), hence the effects of overlapping F⁻ K⁺ hydration shells at high concentration causes the dipole angle distribution around potassium cations to move slightly inwards while becoming slightly shorter and broader.

We can also consider the water dipole angle distribution around a central bulk water molecule. As described in section 2.3.1, a bulk water molecule is defined as any water molecule that does not lie within the first hydration shell of any solute molecules, in this case potassium halide ions. Treating the EPSR simulation this way allows us to deconvolute different local microenvironments in a way that EPSR cannot. In the same manner as previously, the dipole angle is defined as the angle between the vector pointing from the central bulk water molecule oxygen to the hydrating water molecule oxygen and the vector corresponding to the hydrating water molecule dipole. This data is shown in figure 3.3.2 for the highest studied concentration of 2.4 mol/kg H₂O. The data for the 0.6

and 1.2 mol/kg H₂O samples are shown in section A. Here it is clear that the bulk water dipole angle distribution is negligibly altered by the presence of potassium halide ions, even at the highest studied concentration. This would indicate that the structural perturbations to water by potassium halide ions are short ranged, only significantly perturbing water structure within the first hydration shell. This is broadly in line with the original findings by Soper and Weckstrom who observed significant perturbations to the ion-water dipole angle distribution in the first hydration shell of potassium halide ions, and much weaker perturbations to water in the second hydration shell [128].

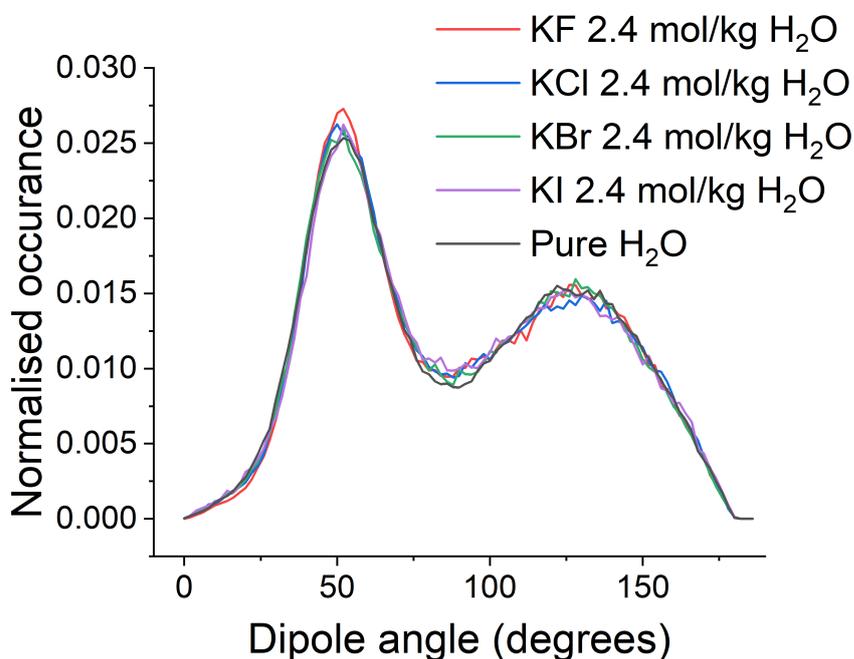


Figure 3.3.2: Distribution of dipole angles for water molecules hydrating a central bulk water molecule within its first hydration shell. The dipole angle is defined as the angle between the vector pointing from the central bulk water oxygen to the hydrating water oxygen and the vector corresponding to the hydrating water molecular dipole. Values normalized to the total number of data points. Only data for highest concentration 2.4 mol/kg H₂O samples shown for clarity.

3.4 Thermodynamics of Hydration

We can now begin to consider the thermodynamics of hydration of potassium halide ions. As a first step we will use the structural information from EPSR to calculate the

enthalpy of hydration ΔH_{solv} using the custom analysis routine to calculate the total interaction energy between a central ion and all the water molecules in its first hydration shell, as described in section 2.3.3. The enthalpy of hydration for an ion is defined as “the amount of heat released when a mole of the ion dissolves in a large amount of water forming an infinitely dilute solution in the process [245]”. In order to be consistent with this definition, ΔH_{solv} was calculated for each ion species at each of the three studied concentrations (0.6, 1.2, and 2.4 mol/kg H₂O) and the results were extrapolated back to 0 concentration to approximate an infinitely dilute solution. The individual data points for each salt at each concentration are reported with negligible error bars. This is because the large number of statistics gathered through a combination of 20-30 iterations of the analysis routine and the large EPSR box size means that the average total enthalpy of hydration for each ion in each circumstance can be calculated with negligible error. However, it is likely that these calculated values would differ somewhat if the cut-off distances for the first hydration shell were varied, or the empirical potential were considered in the calculations, which would allow for the calculation of error bars. As the purpose of this data is to extrapolate the calculated values back to 0 concentration, the calculated error is therefore the error in the y intercept of the fitted linear function. The addition of error bars to the data points would increase this error, hence the reported error in the final value for each ion is the minimum theoretical error given the data. The final results for infinitely diluted solutions of the potassium cation and the corresponding halide anion were then summed to allow the salt species to be addressed as a whole, ΔH_{EPSR} , as the enthalpy of hydration of an cation/anion is shown to depend weakly on the species of the corresponding anion/cation [35]. This process is shown in figure 3.4.1(a) for the halide anions and figure 3.4.1(b) the potassium cations.

The final determined values for ΔH_{EPSR} for each of the four salt species is summarised in table 3.4.1. Here we also report the values for the Gibbs free energy of hydration, ΔG_{solv} , and the enthalpy of hydration, ΔH_{solv} for potassium halides as calculated by Tissandier *et al.* [35]. These values were calculated by taking the difference between the free energy/enthalpy of formation of the corresponding salt and that of the gas-phase ions. The entropy of hydration multiplied by the experimental temperature

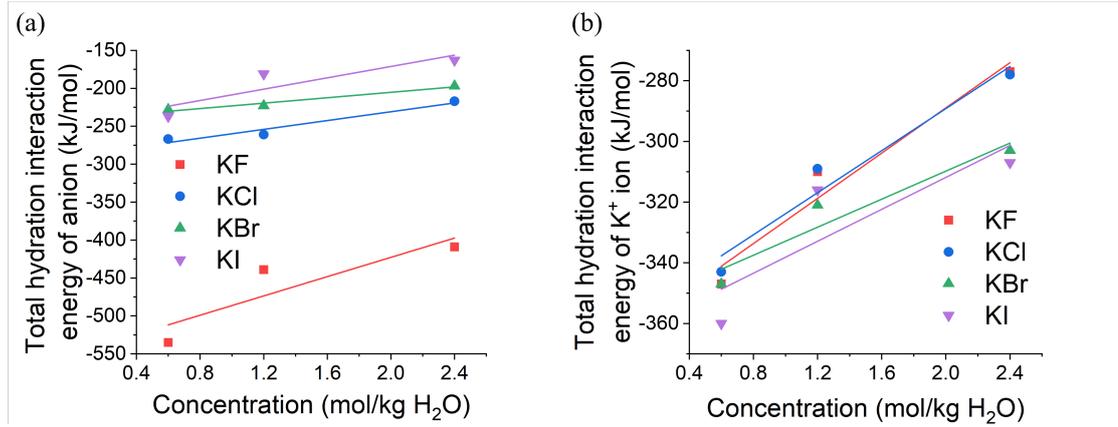


Figure 3.4.1: Total interaction energy between a central halide anion (a) or potassium cation (b) and every water molecule in its first hydration shell, ΔH_{solv} , at each of the three studied concentrations.

$T\Delta S_{solv}$ is also included and is calculated using equation 3.4.1.

$$\Delta G_{solv} = \Delta H_{solv} - T\Delta S_{solv} \quad (3.4.1)$$

Table 3.4.1: Gibbs free energy of hydration, ΔG_{solv} , enthalpy of hydration, ΔH_{solv} , and entropy of hydration multiplied by experimental temperature $T\Delta S_{solv}$ of potassium halides taken from previous literature [35] and enthalpy of hydration calculated using simulation details of EPSR, ΔH_{EPSR} .

Salt species	ΔG_{solv} (kJ/mol)	ΔH_{solv} (kJ/mol)	$T\Delta S_{solv}$ (kJ/mol)	ΔH_{EPSR} (kJ/mol)
KF	-780.8	-844	-63.2	-910 ± 60
KCl	-656.2	-700	-43.8	-650 ± 20
KBr	-629.4	-669	-39.6	-600 ± 20
KI	-592.1	-627	-34.9	-610 ± 60

Here we observe that as one moves down the halides group from F⁻ to I⁻ that the experimental Gibbs free energy, enthalpy, and entropy of hydration for the corresponding potassium halide salt calculated by Tissandier *et al.* [35] all become less negative. This indicates that with increasing ionic radius, and therefore decreasing charge density, the introduction of the ion into water becomes less energetically and enthalpically

favourable, but there is also a decreased entropy penalty. If we now compare ΔH_{soln} to ΔH_{EPSR} for each salt we can show that the results are in reasonable quantitative agreement. This demonstrates that by only considering water molecules in the first hydration shell of an ion that thermodynamic properties of the bulk solution can be recaptured in the case of aqueous potassium halides. This in turn suggests that the thermodynamics of hydration calculated for bulk solutions of aqueous potassium halides are dominated by short ranged interactions, in agreement with conclusions drawn from examining structural perturbations in section 3.3.

We can substantiate this claim further by monitoring the average hydrogen bond interaction energy between water molecules in the bulk. As previously, a bulk water molecule is defined as a water molecule that does not lie within the first hydration shell of a solute. The definition of hydrogen bonding between water molecules and method of data acquisition and plotting is described in section 2.3.2. The resultant normalised distributions of bulk water - water hydrogen bond interaction energies in pure water and in 2.4 mol/kg H₂O of the four salt species is shown in figure 3.4.2. Data for the four salt species at 0.6 and 1.2 mol/kg H₂O are shown in section A.

These distributions are fit with a Gaussian function to determine the average hydrogen bond interaction energy and associated uncertainty. These results are shown in figure 3.4.3. Here it can be seen that the average hydrogen bond interaction energy between two bulk water molecules is only negligibly altered in the presence of each of the four salt species, again indicating only short range perturbations to water structure and enthalpic interactions as a result of the ions in solution.

Until this point we have observed short ranged perturbations to water structure and enthalpic interactions as a result of potassium halide ion addition by monitoring the water dipole angle distribution around the ions and the enthalpy of hydration as calculated through EPSR, ΔH_{EPSR} . We can now begin to quantify the range over which ions perturb water structure. To do this we monitor the average enthalpic interaction between a central potassium halide anion and a surrounding water molecule as a function of distance, as shown in figure 3.4.4. Again only the data for the highest studied concen-

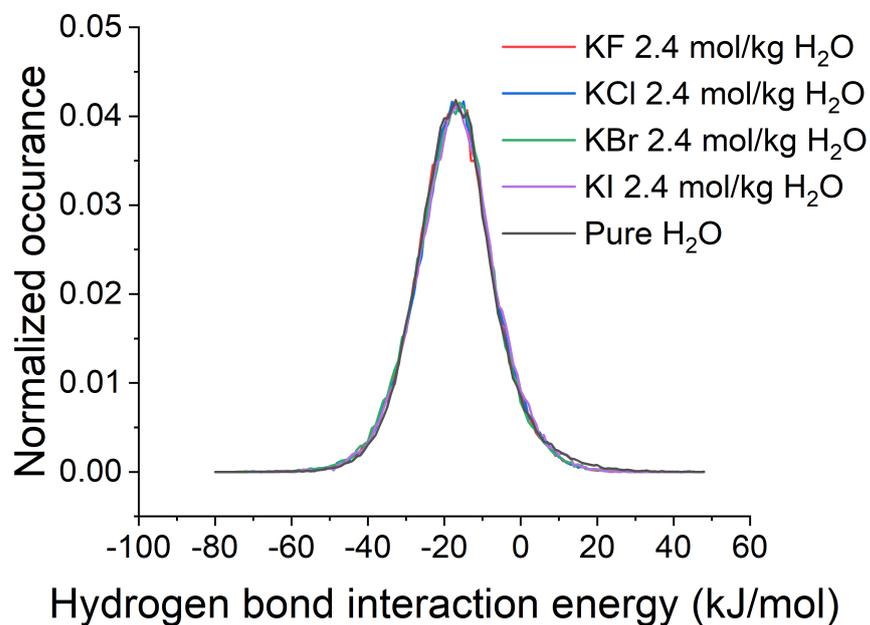


Figure 3.4.2: Bulk water - water hydrogen bond interaction energy distribution for pure water and each of the four studied potassium halide salts at 2.4 mol/kg H₂O.

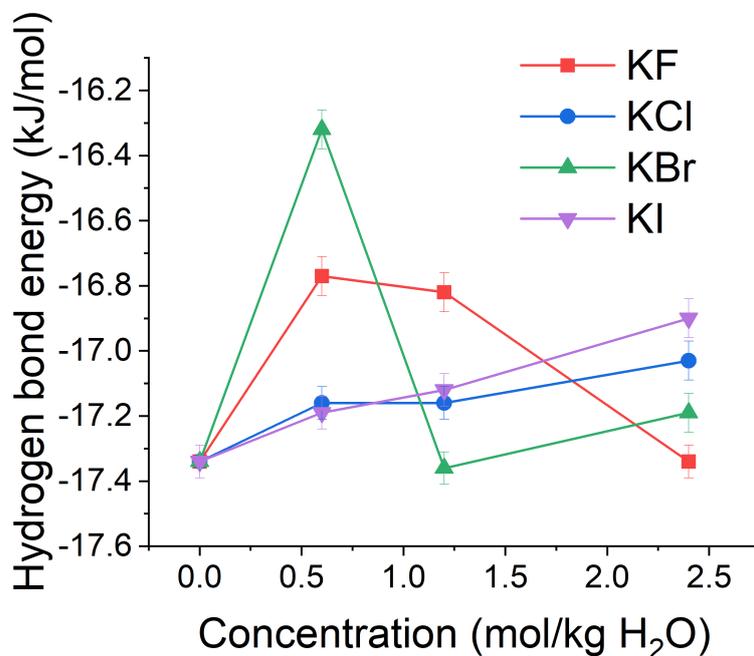


Figure 3.4.3: Bulk water - water hydrogen bond interaction energy for pure water and each of the four studied potassium halide salts as a function of solute concentration.

tration of 2.4 mol/kg H₂O are shown for clarity. The remaining datasets can be found in section A. From this data it is clear that at small distances, below approximately 3.5 Å there are significant favourable enthalpic interactions between the central anion and the surrounding water molecules, shown by the negative values of the average interaction energy, however at larger distances this quickly decays towards 0, indicating negligible enthalpic interactions between anions and water molecules in solution at distances above approximately 4.0 - 4.5 Å. This again demonstrates that the structural perturbations to water by potassium halide ions are short ranged.

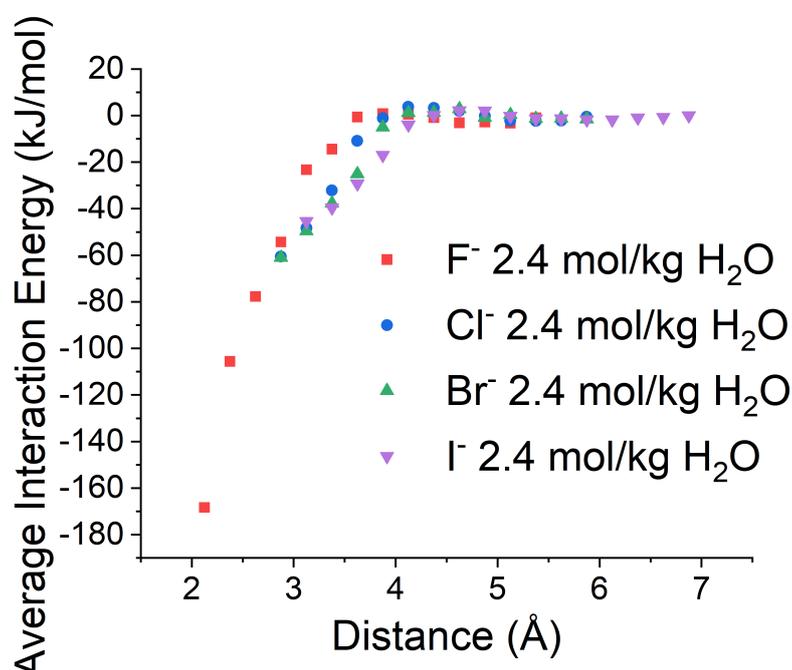


Figure 3.4.4: Average enthalpic interaction strength, ΔH between a central potassium halide anion and its surrounding water molecules at a given distance.

To quantify the distance over which the anion has a favourable enthalpic interaction with the surrounding water molecules a linear function was fit to the data shown in figure 3.4.4 at low distances and extrapolated to determine the distance at which $\Delta H = 0$. This procedure yields an enthalpic interaction distance for each potassium halide anion at each of the three studied concentrations. The calculated distances for each concentration in turn were then extrapolated back to 0 concentration to yield a final value for the enthalpic interaction distance $r_{\Delta H=0}$ for each potassium halide anion with associated uncertainty.

Table 3.4.2: Enthalpic interaction distances for each potassium halide anion species determined by monitoring the enthalpic interaction energy between the potassium halide anion and the surrounding water molecules as a function of distance. Final values reported are the enthalpic interaction distances extrapolated to 0 concentration for each potassium halide species $r_{\Delta H=0}$. Cut-off distances for first hydration shell for each potassium halide anion, as shown in figure 3.2.3, are included.

Potassium Anion	Halide	Enthalpic Interaction Distance $r_{\Delta H=0}$ (Å)	Cut-off distance for First Hydration shell (Å)
F ⁻		3.40 ± 0.02	3.21
Cl ⁻		4.03 ± 0.04	3.81
Br ⁻		4.21 ± 0.05	4.20
I ⁻		4.45 ± 0.03	4.49

An example of this procedure using the data for KF at 1.2 mol/kg H₂O is shown in section A. These results are shown in table 3.4.2.

These results show that with increasing ionic radius that $r_{\Delta H=0}$ also increases, however the total enthalpic interaction, shown in table 3.4.1, is weaker. Table 3.4.2 also compares these results to the cut-off distances for the first hydration shell around each potassium halide ion as described in figure 3.2.3. Here we observe that for the smaller anion species, F⁻ and Cl⁻, that $r_{\Delta H=0}$ extends outside the first hydration shell, however for the largest anion species, Br⁻ and I⁻, the two quantities are approximately equal. These results indicate that $r_{\Delta H=0}$ varies more weakly with anionic species than the ability of the anion to form ordered hydration shells.

3.5 Dynamic Perturbations to Water - Nuclear Magnetic Resonance

Now that we have considered perturbations to water structure induced by potassium halide ions, let us now consider perturbations to water dynamics. NMR can be used to probe the dynamics of water by measuring the T_1 relaxation time and the diffusion coefficient, D , of the water molecules in aqueous solution. As described in section 2.4.4, the inverse T_1 relaxation time is directly proportional to the rotational correlation time,

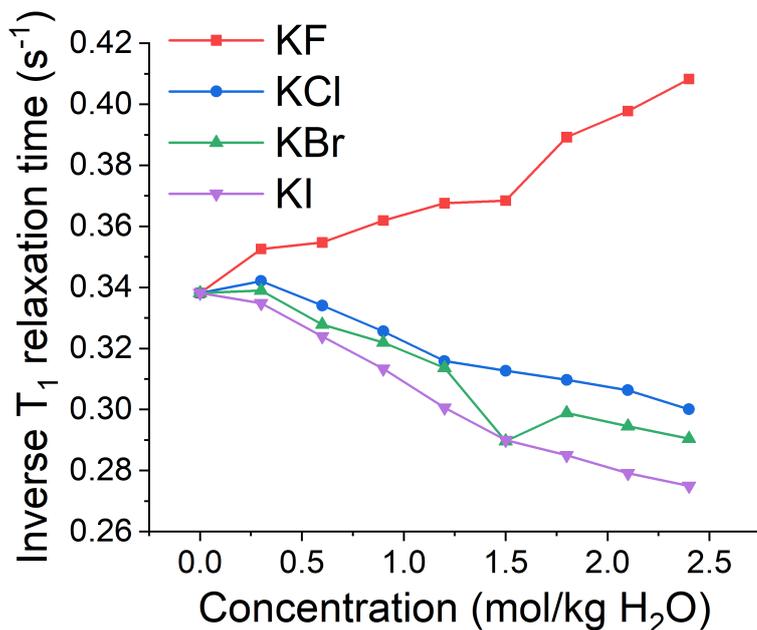


Figure 3.5.1: Inverse T_1 decay time from NMR experiments as a function of solute concentration for aqueous potassium halides.

τ_c , of the water molecule in the fast limit ($\omega\tau_c \ll 1$). The rotational correlation time is roughly equal to the time taken for the root-mean-square deflection of the water molecule to be approximately 1 radian, and therefore serves as a measure of the microscale rotational dynamics of the system [87, 168, 208]. The diffusion coefficient then measures the microscale diffusive dynamics of water molecules in the solutions.

To investigate the effects of potassium halides on water dynamics T_1 and D NMR experiments were completed on the four aqueous potassium halide salts KF, KCl, KBr, and KI over a concentration range of 0 - 2.4 mol/kg H₂O in 0.3 mol/kg H₂O increments. These results are shown in figure 3.5.1 and figure 3.5.2. As these results were obtained to serve as a supplement to the structural data produced through neutron scattering and EPSR, only a single measurement was taken for each salt at each concentration, and hence the data are displayed without error bars.

These results are in general agreement with NMR results on aqueous potassium halides recently reported by Laage *et al.* [157]. This demonstrates that the rotational dynamics of water molecules are retarded in aqueous KF, as shown by an increasing

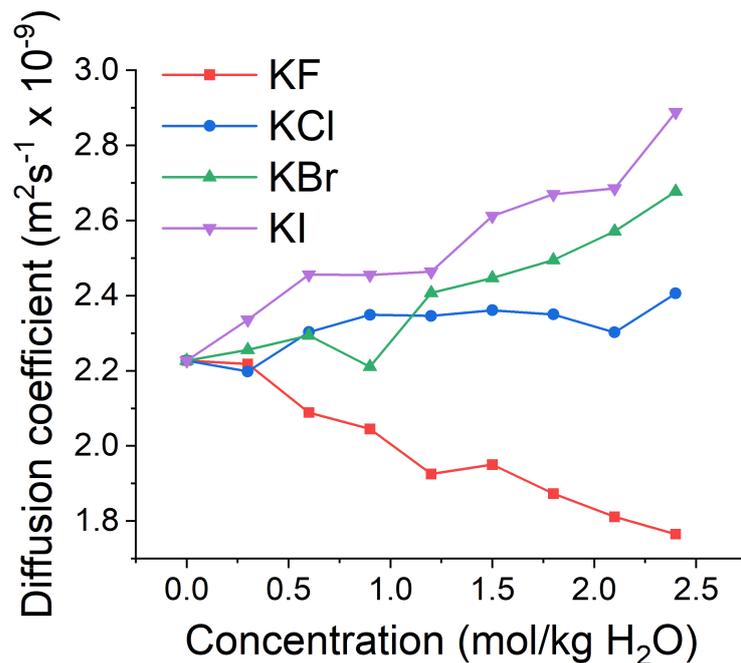


Figure 3.5.2: Diffusion coefficient D of water molecules from NMR experiments as a function of solute concentration for aqueous potassium halides.

inverse T_1 relaxation time and therefore an increased rotational correlation time with increasing KF concentration, but accelerated in the case of KCl, KBr, and KI, as shown by a decreasing inverse T_1 relaxation time with increasing salt concentration. Within these three salts KCl accelerates rotational dynamics the least, and KI accelerates rotational dynamics the most. The diffusive dynamics of water molecules are perturbed in the same manner: KF is shown to retard diffusive dynamics of the water molecules, whereas KCl, KBr, and KI accelerate diffusive dynamics, with KCl being the least accelerating and KI being the most accelerating.

The results gained from these styles of NMR experiments represent an ensemble average of all the rotational and diffusive characteristics of every water molecule in the entire sample. As demonstrated by the structural and thermodynamic analysis in sections 3.2, 3.3, and 3.4, solutions of potassium halides are heterogeneous in their structure, containing differently structured microenvironments. It is therefore a reasonable assumption that the water molecule dynamics are also heterogeneous in nature and will depend on the particular microenvironment in which a particular water molecule sits. As demon-

strated in the structural and thermodynamic analysis, potassium halide ions significantly perturb water in their first hydration shells, but these perturbations do not significantly extend into the bulk solution. Therefore water molecules in the hydration shells of ions will likely exhibit different rotational and diffusive characteristics than water in the bulk. We will now attempt to deconvolute these dynamic characteristics to determine the rotational and diffusive dynamics of water molecules in the hydration shells of the relevant ions.

In order to do this we will follow the procedure outlined by Engel and Hertz in 1968 [87]. This method assumes that within an aqueous solution containing two separate ionic species, such as the potassium halide salts discussed in this work, there exist three distinct rotational correlation times for water molecules. Each of these rotational correlation times is attributed to water molecules in each of three different microenvironments present in the solution: water molecules located in the first potassium cation hydration shell, τ_c^+ , water molecules located in the first halide anion hydration shell, τ_c^- , and the remaining bulk water molecules not located in the first hydration shell of either the potassium cation or the halide anion τ_c^0 that exhibit identical rotational dynamics to pure water. This assumption is therefore in good agreement with the structural and thermodynamic results described previously.

As the experimentally derived inverse T_1 relaxation time, directly proportional to τ_c , represents an ensemble average of the three different rotational dynamics of water molecules for each of the three distinct environments, we can construct an expression that describes this ensemble average. This is shown in equation 3.5.1, where one mole of salt dissolves into ν^\pm moles of cations/anions, both equal to unity in the simple case of potassium halides, n_h^\pm is the coordination number for the cation/anion, described in section 2.2.5 and c is the concentration of the salt in mol/kg H₂O. The factor of 55.5 originates from the molarity of pure water [87]. The coefficient for each of the three rotational correlation times described in equation 3.5.1 is therefore equivalent to the fraction of water molecules located in each of the three environments.

$$\frac{1}{T_1} \propto \left[1 - \frac{\nu^+ n_h^+ c}{55.5} - \frac{\nu^- n_h^- c}{55.5} \right] \tau_c^0 + \frac{\nu^+ n_h^+ c}{55.5} \tau_c^+ + \frac{\nu^- n_h^- c}{55.5} \tau_c^- \quad (3.5.1)$$

We now consider a form of the Jones-Dole equation [246], which describes the ratio between the viscosity of a solution to the viscosity of pure water as a polynomial function of concentration. A version of this equation as applicable to T_1 relaxation measurements is shown in equation 3.5.2, where $T_{1,0}$ is the relaxation time of pure water and c is the solute concentration in mol/kg H₂O.

$$\frac{\frac{1}{T_1}}{\frac{1}{T_{1,0}}} = 1 + Bc + Cc^2 + \dots \quad (3.5.2)$$

For salts at sufficiently low concentration such that the inverse T_1 relaxation time is linearly dependent on concentration, such as is the case for the salts and concentrations studied here (see figure 3.5.1), only terms up to and including Bc in equation 3.5.2 are necessary to sufficiently describe the solution. This in turn means we can also safely assume that the rotational correlation times for water molecules in the hydration shells of the potassium cation or the halide anion are independent of concentration in this concentration range.

We now allow the B coefficient to be expressed as the sum of the contributions from the rotational correlation times of water molecules in the potassium cation hydration shell B^+ and the rotational correlation times of water molecules in the halide anion hydration shell B^- , as shown in equation 3.5.3.

$$B = B^+ + B^- \quad (3.5.3)$$

Finally, we can combine equations 3.5.1, 3.5.2, and 3.5.3 to form equation 3.5.4, which describes the rotational correlation times for water molecules in the hydration shell of the potassium cation/halide anion normalised to the rotational correlation time of pure water.

Table 3.5.1: Comparison of ionic radius, enthalpy of hydration, water coordination number, and enthalpy of hydration per coordinated water molecule to determine which halide ion is most similar to the potassium cation.

Ion species	Ionic radius (pm)	ΔH_{EPSR} (kJ/mol)	Water coordination number	ΔH_{EPSR} per coordinated water (kJ/mol)
K^+	152	-360 ± 2	6.4	-56.3 ± 0.3
F^-	119	-550 ± 50	6.3	-87 ± 8
Cl^-	167	-290 ± 10	7.0	-41 ± 1
Br^-	182	-241 ± 5	8.7	-27.7 ± 0.6
I^-	206	-250 ± 30	10.1	-25 ± 3

$$\frac{\tau_c^\pm}{\tau_c^0} = 1 + \frac{55.5}{n^\pm} B^\pm \quad (3.5.4)$$

In order to use this expression one must already have knowledge of the B parameter for one of the ion species present in any of the potassium halide salts studied here. The simplest way of achieving this is to state that two ion species have identical B parameters. For this we again follow the procedure of Engel and Hertz [87] who state that $B(Cl^-) = B(K^+)$. Whilst this assumption is simplistic, we can consider several parameters relating to the two ion species to justify that this is the most reasonable assumption as the potassium cation and chloride anion are the most similar ions. These comparisons are shown in table 3.5.1. These parameters show that while the coordination number is most similar between K^+ ions and F^- ions, the ionic radius, ΔH_{EPSR} , and ΔH_{EPSR} per coordinated water molecule are the most similar between the K^+ and Cl^- ions. Therefore, K^+ and Cl^- are the most similar overall and therefore are likely to have similar B parameters.

It is now possible to fit the data displayed in figure 3.5.1 using equation 3.5.2, assuming that terms beyond Bc are negligible and also assuming that $B(Cl^-) = B(K^+)$ to yield a B parameter for each of the five ion species discussed in this work. EPSR can be used to determine the water coordination numbers calculated according to the cut-off distances shown in table 3.4.2 and figure 3.2.3. Applying equation 3.5.4 to these results then allows us to calculate the rotational correlation time for water molecules in

Table 3.5.2: Calculated B parameter for each of the five ion species discussed in this work derived from inverse T_1 relaxation data, water coordination number calculated through EPSR, and rotational correlation time for water molecules in the hydration shell of the ion normalised to the rotational correlation time for pure/bulk water calculated according to equation 3.5.4.

Ion species	B (kg H ₂ O/mol)	n_h	$\frac{\tau_c^\pm}{\tau_c^0}$
K ⁺	-0.0264	6.4	0.77
F ⁻	0.1072	6.3	1.94
Cl ⁻	-0.0264	7.0	0.79
Br ⁻	-0.0387	8.7	0.75
I ⁻	-0.0592	10.1	0.67

the first hydration shell of the various ions normalised to the rotational correlation time for pure/bulk water. These results are displayed in table 3.5.2.

These results show that water is retarded in the hydration shell of F⁻, but accelerated in the hydration shells of Cl⁻, K⁺, Br⁻, and I⁻ (listed in order of decreasing normalised rotational correlation time and therefore increasing acceleration).

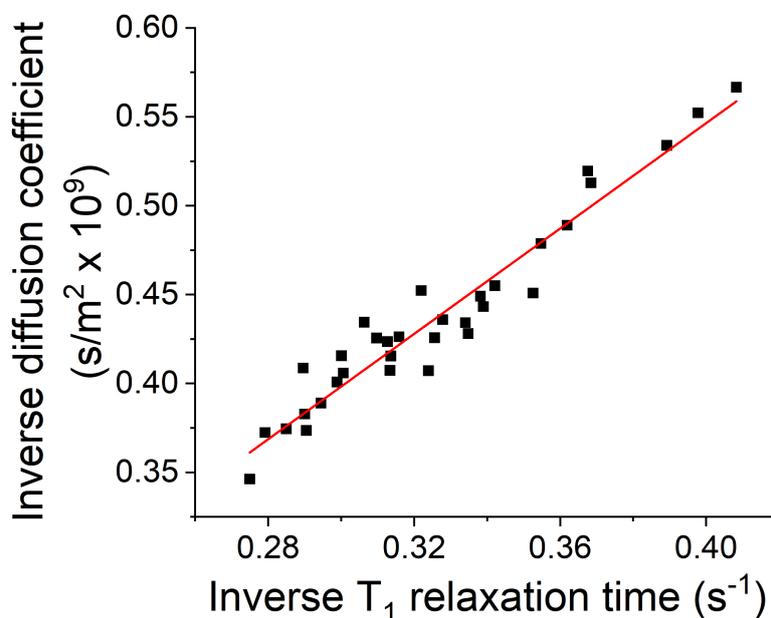


Figure 3.5.3: Inverse diffusion coefficient for each sample plotted against the corresponding inverse T_1 relaxation time for each sample to demonstrate directly proportional relationship. Straight line fitted as guide to the eye.

Now that we have deconvoluted the inverse T_1 relaxation data into its constituent parts as a result of the different microenvironments present in aqueous potassium halides, we can attempt to do the same for the diffusion data. This will allow us to predict the diffusion coefficient of water molecules in the first hydration shells of the five ion species presented here. To begin this process we start by validating the relation, also discussed by Engel and Hertz [87], that $\frac{1}{T_1} \propto \frac{1}{D}$. To do this we plot the inverse T_1 relaxation time for each sample (a particular salt species at a particular concentration) against the corresponding inverse diffusion coefficient for the sample, as shown in figure 3.5.3. These results clearly indicate that for aqueous potassium halides over the studied concentration range that $\frac{1}{T_1} \propto \frac{1}{D}$. Using this information we can now formulate a series of expressions in the same manner as presented in equation 3.5.1, equation 3.5.2, and equation 3.5.4

$$\frac{1}{T_1} \propto \frac{1}{D} \propto \left[1 - \frac{\nu^+ n_h^+ c}{55.5} - \frac{\nu^- n_h^- c}{55.5} \right] \frac{1}{D_0} + \frac{\nu^+ n_h^+ c}{55.5} \frac{1}{D^+} + \frac{\nu^- n_h^- c}{55.5} \frac{1}{D^-} \quad (3.5.5)$$

$$\frac{\frac{1}{D}}{\frac{1}{D_0}} = 1 + Bc + \dots \quad (3.5.6)$$

$$\frac{\frac{1}{D^\pm}}{\frac{1}{D_0}} = 1 + \frac{55.5}{n^\pm} B^\pm \quad (3.5.7)$$

Finally the data presented in figure 3.5.2 can be reanalysed using equation 3.5.7 to predict the diffusion coefficient of water molecules in the first hydration shell of the five ion species studied here. To facilitate understanding the final values predicted by equation 3.5.7 have been converted from $\frac{\frac{1}{D^\pm}}{\frac{1}{D_0}}$ to $\frac{D^\pm}{D_0}$ as this is conceptually easier to understand. The final values are reported in table 3.5.3.

These results show that the diffusive dynamics of water, just as in the case of the rotational dynamics, are retarded in the hydration shell of F^- , but accelerated in the hydration shells of Cl^- , K^+ , Br^- , and I^- (listed in order of increasing normalised diffusion coefficient and therefore increasing acceleration).

Table 3.5.3: Calculated B parameter for each of the five ion species discussed in this work derived from diffusion coefficient data, water coordination number calculated through EPSR, and diffusion coefficient for water molecules in the hydration shell of the ion normalised to the diffusion coefficient for pure/bulk water calculated according to equation 3.5.7.

Ion species	B (kg H ₂ O/mol)	n_h	$\frac{D^\pm}{D_0}$
K ⁺	-0.0133	6.4	1.13
F ⁻	0.1257	6.3	0.47
Cl ⁻	-0.0133	7.0	1.12
Br ⁻	-0.0583	8.7	1.59
I ⁻	-0.0693	10.1	1.62

3.6 Conclusions

The results presented in this chapter are in agreement with the growing consensus [44, 47, 87, 126–128, 133–135, 247–254] that simple monovalent ions, such as the potassium halide ions discussed in this work, only significantly perturb the structure of water within their first hydration shells. By considering the experimental values for the enthalpy of hydration of potassium halides determined by Tissandier *et al.* [35] we can show by using EPSR that only the first hydration shell around an ion needs considering to recapture this bulk property of the solution, indicating short ranged structural perturbations. This is reflected in the bulk water - water hydrogen bond interaction energy and the bulk water - water dipole angle distribution, which are only negligibly perturbed for all four salt species at all studied concentrations. We then quantify this short range interaction by considering the enthalpic interaction between ions and the surrounding water molecules, and show that significant enthalpic interactions between potassium halide ions and water molecules only extend to approximately 4.5 Å, closely corresponding to the cut-off distances for the first ionic hydration shells.

We can now begin to consider correlations between results using the various different experimental methods outlined in this work, as well as previous results from other experimental studies. Previous simulation and infrared spectroscopy studies have shown that the rotational dynamics of water in the hydration shell of potassium halide ions observes a trend from slow to fast with: F⁻, Cl⁻, Br⁻, K⁺, I⁻ [254, 255]. This order

correlates with the height of the peak in the water dipole angle distribution around the ion, as these also occur in the same order. If we consider the normalised rotational correlation time data 3.5.2 and the normalised diffusion coefficient data 3.5.3, we observe the order from slow to fast as, F^- , Cl^- , K^+ , Br^- , I^- . Considering these two orderings of the dynamic perturbations to water induced by potassium halides we can observe that within the halide anion series (F^- , Cl^- , Br^- , I^-), that water molecules are increasingly mobile, both in their rotational and diffusive motions, with increasing halide anion size when monitored using multiple techniques.

However, it is more difficult to confidently state where the dynamics of water molecules in the hydration shell of the potassium cation fall on these scales. This difficulty likely arises from the assumptions involved with different experimental techniques. Within the simulation and EPSR studies water molecules and ions are described with a relatively simple potential, consisting of a Lennard-Jones and coulomb component [128, 255]. This simplification is made to allow simulations and modelling to proceed in a reasonable timescale but is certain to introduce uncertainties in the final results. Within our NMR approach, introduced by Engel and Hertz [87], we make the assumption that the NMR B parameter derived from both inverse T_1 and inverse diffusion coefficient D data is identical for K^+ and Cl^- , as described previously. This is almost certainly not the case but is still a useful way of observing the relative perturbing ability of ions on the rotational and diffusive dynamics of water molecules within their hydration shells. We must also consider the difference in the nature of cationic vs anionic hydration. In positively charged cations, water molecules within the first hydration shell will tend to orient both hydrogens away from the central cation, whereas in negatively charged anions, water molecules within the first hydration shell will tend to orient one hydrogen towards the central anion. This will certainly result in subtle differences in the mechanism for hydrogen bond switching events resulting in diffusive and rotational motion of the water molecules within ionic hydration shells, as explored by Laage *et al.* [157]. With these subtle differences in hydration structure, the need for assumptions in modelling and simulations, and the overall similarity in results when we consider the height and width of the water dipole angle distribution around the Cl^- and Br^- anions, as

shown in figure 3.3.1(a), and the potassium cation, as shown in figure 3.3.1(b), and the perturbation to rotational dynamics of water molecules in the first hydration shell, as shown in table 3.5.2, and the perturbation to diffusive dynamics of water molecules in the first hydration shell, as shown in table 3.5.3 of the K^+ , Cl^- , and Br^- ions, it is little surprise that it is difficult to confidently rank the potassium cation in the order of which they perturb the dynamics of hydrating water molecules.

We can now consider the relation between structural perturbation and dynamic perturbation within the first hydration shell of halide anions. If we consider the water-dipole angle distribution (figure 3.3.1(a)) alongside the halide anion - water oxygen RDFs (figure 3.2.3) we observe that both the peak of the water dipole angle distribution and the first peak of the anion - water oxygen RDF become shorter and broader with increasing size of the halide anion. Simply put, it can therefore be said that water molecules are less structurally constrained in the hydration shells around the halide anions with increasing anion size and are freer to adopt a larger range of distances and orientations. We also observe that with increasing halide anion size that water molecules in the first hydration shell exhibit accelerated rotational and diffusive dynamics. It would therefore appear that how structurally constrained the water is within the hydration shell of an ion plays a key role in the dynamics of the water molecules. Less constraint over preferred orientation and distance of the hydrating water molecules relative to the central halide anion likely means that they can more often be found in an appropriate conformations to undergo hydrogen bond switching, and a reduced enthalpic interaction provides a lower energy barrier to carry out the hydrogen bond switching event, resulting in the increased dynamics.

This reduction of how structurally constrained water molecules are in the first hydration shell around halide anions is also reflected in bulk thermodynamic properties, namely the calculated entropy of hydration at the experimental temperature $T\Delta S_{solv}$ shown in table 3.4.1. We observe that with a less structurally constrained hydration shell with increasing halide anion size, shown by the shorter and broader peaks in the water dipole angle distribution around halide anions (figure 3.3.1(a)) and the anion - water oxygen RDFs (figure 3.2.3), and increased rotational (table 3.5.2) and diffusive

(table 3.5.3) dynamics, that there is a reduction in the entropy penalty associated with including the halide anion into the water network at constant temperature $T\Delta S_{solv}$, shown by less negative values of $T\Delta S_{solv}$ with increasing halide anion size (table 3.4.1). This is unsurprising, as less structurally constrained more dynamic water molecules are clearly free to explore a larger variety of different states, resulting in a reduced entropic penalty by definition.

The conclusions drawn here, that the effects of ions are local rather than global, clearly hold for simple monovalent electrolytes such as potassium halides. We have also observed that in the case of simple ions which only significantly perturb the structure and dynamics of water molecules in their first hydration shells that we can study these short ranged local environments to calculate bulk properties of the complete aqueous system, such as the enthalpy of hydration ΔH_{solv} , and reveal correlations between other bulk properties, such as the entropy of hydration ΔS_{solv} . Recently published results from Gaiduk *et al.* [243] suggest that while structural perturbations may be short ranged, that water molecules can experience polarizability changes due to simple monovalent ions in solution at larger distances, however this finding is inaccessible to EPSR as the employed force fields are non-polarizable. In the next chapter we will begin to explore the effects of more complex organic solutes on water at ambient and high pressures.

Chapter 4

Perturbations to Water Structure due to Hydrostatic Pressure in the Presence of Trimethylamine N-Oxide

4.1 Introduction

As discussed in section 1.2, life on Earth is capable of surviving at extremes of temperature, salinity, pH, and pressure [11]. It is this last extreme that will be the focus of this chapter. On Earth the highest pressure natural environment is located at the bottom of the deepest point of the ocean: the Mariana Trench [172]. This is almost 11 km below the surface of the ocean and the resultant pressure due to the column of water above it is approximately 1.1 kbar or 8 tons per square inch. At these depths and pressures it is difficult to imagine that living organisms could thrive, yet many organisms have adapted to live in these high pressure environments. Collectively they are referred to as piezophiles [80].

One of the many difficulties of living at high pressure is that high pressure has been shown to destabilise biomolecules such as proteins and, at pressures above approximately 4 - 5 kbar, causes them to denature [74, 81]. Therefore the various biochemical processes fundamental to life should not be possible under these conditions. However, piezophiles have adapted a range of strategies to overcome these issues. One of these strategies is the accumulation of organic osmolytes known as “piezolytes” [82, 83] which have been shown to stabilise biomolecules against pressure induced denaturation [68, 84, 85]. One particular piezolyte that has received significant attention is trimethylamine N-oxide (TMAO). This has been shown to accumulate in the muscle tissue of bony fish at a concentration that is linearly dependent on the depth from which the fish was harvested, and therefore the pressure at which the fish survives [86]. This data is reported in figure 4.1.1.

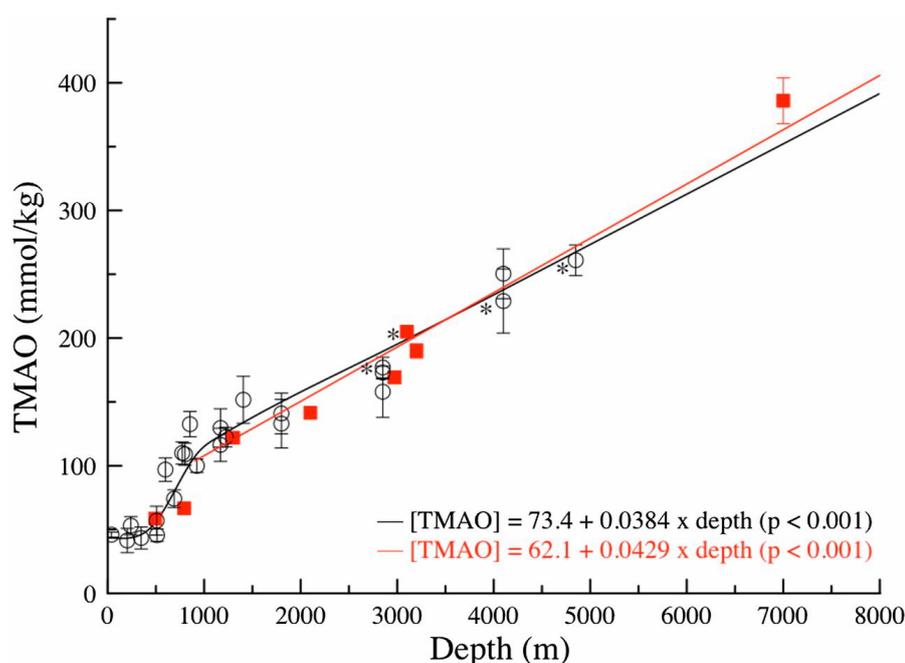


Figure 4.1.1: Data on the concentration of TMAO (measured in mol of TMAO per kg of wet muscle tissue) in the muscle tissue of bony fish harvested from different depths originally published by Yancey *et al.* [86].

However the usefulness of TMAO extends beyond its ability to stabilise proteins against pressure denaturation. It has also been shown to stabilise proteins against denaturation due to urea addition [69, 256], as well as other denaturing osmolytes [145, 257],

pressure [60, 64, 70, 258], and temperature [154]. It is this ubiquity of TMAO as a protein stabilising agent that has prompted such a depth of research into the origin of this effect. Like other protein stabilising agents, it is thought to be preferentially excluded from protein surfaces [67–73]. Therefore the origins of its stabilising properties will not lie in direct protein-TMAO interactions, but rather through indirect interactions as a result of how TMAO perturbs the solvent environment. For this reason aqueous TMAO solutions have been well investigated to determine how TMAO perturbs water structure and dynamics under ambient conditions [65, 66, 146, 171], under pressure [64, 72, 147, 148, 151, 259, 260] and in the presence of other solutes such as urea [62, 63, 73, 142, 152, 261, 262] and $\text{Mg}(\text{ClO}_4)_2$ [2, 3]. These results have all shown that TMAO acts to enhance hydrogen bonding in water, and that this helps resist perturbations to water structure induced by various externalities.

However, it must be kept in mind that when considering findings on binary solutions of water and TMAO under pressure that much of the previous literature stems from simulation/modelling studies [64, 72, 147, 148, 259, 260]. It is only recently that experimental studies have begun to be reported to validate these findings [149, 151]. The most recent such study by Kolling *et al.* [149] studied 1.5 M TMAO up to 12 kbar with a combination of ab initio molecular dynamics, force field molecular dynamics, and THz absorption spectroscopy. This demonstrated that the negatively charged TMAO oxygen coordinates approximately three water molecules under ambient conditions, but that this increases to four at 10 kbar, consistent with previous results from infrared spectroscopy [150]. These results also showed that the increased pressure results in a more compressed water structure around the hydrophobic methyl groups on the TMAO molecule, which is partially compensated for by an “on average weakened H-bond network due to the shift from a threefold to a fourfold H-bonded TMAO oxygen”. This study is certainly useful as it attempts to bridge the findings of molecular dynamics studies with experimental data, however neutron scattering and EPSR can offer an even more detailed insight into the structural perturbations present in aqueous TMAO. Therefore in this chapter we seek to answer whether TMAO is capable of resisting the pressure induced perturbation to water structure through neutron scattering and EPSR.

4.2 Experimental Considerations

For this experiment a total of 10 samples were prepared. These included 3 isotopic variants of pure water, H₂O, D₂O, and HDO, and 7 isotopic variants of aqueous TMAO at 2.0 mol/kg H₂O. These were produced by combining appropriate quantities of anhydrous TMAO containing only hydrogen (H-TMAO), purchased from Sigma Aldrich, or anhydrous TMAO containing only deuterium (D-TMAO), purchased from CK isotopes, with appropriate quantities of isotopic variants of water. Samples containing an equal molar ratio of H-TMAO and D-TMAO (HD-TMAO) were also produced. The final TMAO samples prepared were: H₂O + H-TMAO, H₂O + D-TMAO, H₂O + HD-TMAO, D₂O + H-TMAO, D₂O + D-TMAO, D₂O + HD-TMAO, HDO + HD-TMAO.

As we wish to investigate pressure induced structural perturbations to water we therefore need to choose a suitable pressure such that the structure of water is clearly perturbed but not so high that any counteracting effect of TMAO would be overwhelmed. In an ideal world this would involve investigating a range of pressures and TMAO concentrations, however due to the time constraints involved with neutron scattering experiments, one must unfortunately settle on a single high pressure measurement. Previous x-ray scattering data on pure water from 0 - 9 kbar by Okhulkov *et al.* demonstrated that increased hydrostatic pressure causes the second hydration shell of water to collapse into the first hydration shell [123]. From this data it was decided that a pressure of 4 kbar would be used for the present study, as at this pressure the second hydration shell is clearly beginning to collapse into the first hydration shell, as evidenced by the water oxygen - water oxygen ($O_w - O_w$) RDF, but can still be collapsed further. Any additional perturbations as a result of TMAO addition should therefore be clear in the $O_w - O_w$ RDF at 4 kbar. To put this into context this is approximately 4 times the pressure found at the bottom of the Mariana Trench, or approximately equal to the weight of a Boeing 747 placed on an area the size of a playing card. Pressure was then applied using a high pressure rig provided by the ISIS facility which can apply pressures up to 5 kbar to aqueous samples via a mechanical piston [263]. The pure water and aqueous TMAO samples at 4 kbar were then compared to data taken on the samples at 25 bar.

This pressure was chosen to be suitably high such that the pressure rig was definitely in contact with the solution and generated a signal on the high pressure rig, but not so high as to appreciably perturb water structure.

A concentration of 2.0 mol/kg H₂O was used based on estimations from previous research. In our 2020 publication on Physical Chemistry Chemical Physics [3], we used data on the concentration of TMAO in the muscle tissue of bony fish published by Yancey *et al.* [86] to estimate that TMAO at 1.0 mol/kg H₂O should resist a pressure equivalent to approximately 2 kbar. As we wish to evaluate whether TMAO is capable of resisting the pressure induced structural perturbations to water we therefore wish to choose a TMAO concentration that would effectively balance this perturbation. As we have selected a pressure of 4 kbar, it therefore makes sense to use a TMAO concentration of 2.0 mol/kg H₂O, as this should achieve the desired balance.

4.3 Structure of Water - EPSR

Structural perturbations to water at 20 °C were monitored by collecting neutron scattering data using the Near and InterMediate Range Order Diffractometer (NIMROD) instrument at the ISIS neutron and muon source, UK, covering a Q range of 0.01 - 30 Å⁻¹ [199], as described in section 2.1.5. The data was then corrected using Gudrun, as described in section 2.1.6 and analysed using EPSR and the custom analysis routine as described in sections 2.2-2.3.4. Within EPSR cubic boxes were made at the experimental concentrations and densities for pure water at 25 bar and 4 kbar and aqueous TMAO at 25 bar and 4 kbar. These contained 5000 water molecules in the case of pure water and 5550 water molecules and 200 TMAO water molecules in the case of aqueous TMAO. The reasons for selecting these general box sizes are overviewed in section 3.2. For aqueous TMAO these box sizes were chosen as this is a ratio that allows the general box size to be maintained and a concentration of 2.0 mol/kg H₂O to be well represented. The densities used were taken from previous literature [151, 264–266] and are outlined in table 4.3.1. The final R factors for the fits between the EPSR simulation and experimental scattering data, shown in section B, were between 0.0025 - 0.0006, corresponding to good quality

System	Density (g/cm ³)	Density (atoms/Å ³)
H ₂ O at 25 bar	0.998	0.1000
H ₂ O at 4 kbar	1.131	0.1134
Aqueous TMAO at 25 bar	1.002	0.0988
Aqueous TMAO at 4 kbar	1.124	0.1109

Table 4.3.1: Densities of pure water and aqueous TMAO at 2.0 mol/kg H₂O at 25 bar and 4 kbar taken from previous literature [151, 264–266].

fits compared with previous neutron scattering and EPSR results on aqueous systems [186]. These were higher than reported in chapter 3, likely due to the unconventional sample shape used in the high pressure sample holder, as shown in section 2.1.5, which requires greater data correction in Gudrun.

The reference potential used to describe water and TMAO was taken from previous neutron scattering studies on aqueous TMAO and urea by Meersman *et al.* [62, 63]. Due to the previously stated popularity of simulating aqueous TMAO there are other force fields available [147, 171], however it was decided to use the Meersman force fields for consistency with previous neutron scattering literature. For the aqueous TMAO sample at 4 kbar we also tested the high pressure $V3 - HP(p)$ TMAO force field published by Hölzl *et al.* [147]. This is a TMAO force field that depends continuously on pressure and has been shown to accurately describe aqueous TMAO at pressures up to 10 kbar. In this work we chose to use the force field values for the highest pressure of 10 kbar as this is the most extreme case and therefore allows us to most easily monitor the sensitivity of the TMAO force field employed on the final results. In principle this should only marginally effect the final results as any inaccuracies in the reference potential should be accounted for by the empirical potential, as described in section 2.2.4. The final fits to the experimental neutron scattering data and reference potentials can be seen in section B.

As stated in section 4.2, the effect of hydrostatic pressure is to force the second hydration shell of water to collapse into the first hydration shell. This is seen clearly in the $O_w - O_w$ RDF, shown in figure 4.3.1. Here we observe that at the low pressure of 25 bar that the position of the first and second peaks are located at 2.77 Å and 4.35 Å respectively. These occur at slightly shorter distances than the 2.78 Å and 4.5

Å peak positions reported in previous literature for ambient water, but a slight inward movement is consistent with what is observed in pressurised water, even at the relatively low pressure of 25 bar [115]. At the high pressure of 4 kbar we observe, consistent with the findings of Okhulkov *et al.* [123], that the second peak has begun to collapse completely into the first, and no longer exhibits a clearly discernible peak. The first peak also moves slightly inwards again, to a distance of 2.76 Å. This first peak is relatively insensitive to pressure as its position is limited by steric effects with the central water molecule. Conversely the relatively open tetrahedral nature of water structure [31] means that there is unoccupied space in the first hydration shell that water molecules in the second hydration shell can occupy, and therefore its position is much more sensitive to external stimuli.

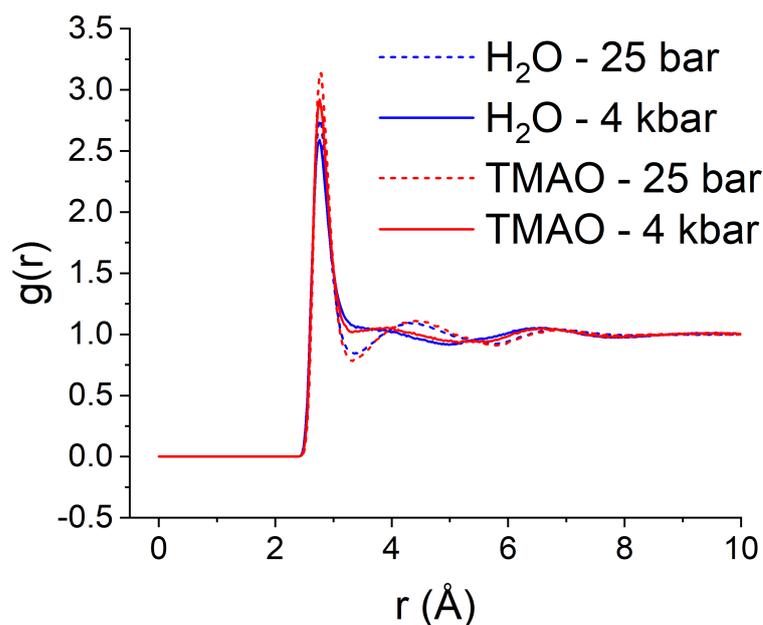


Figure 4.3.1: $O_w - O_w$ RDF for pure water at 25 bar (blue dotted) and 4 kbar (blue solid) and aqueous TMAO at 2.0 mol/kg H_2O at 25 bar (red dotted) and 4 kbar (red solid) calculated through EPSR.

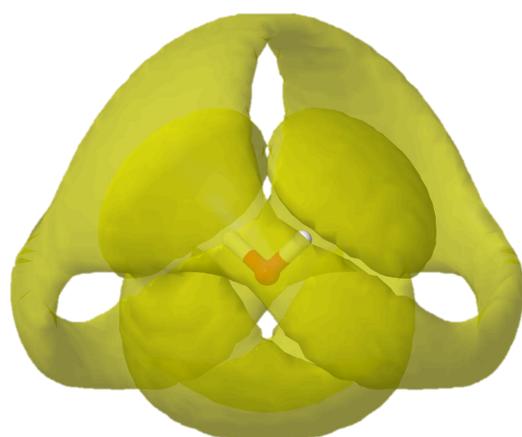
Upon the addition of TMAO at 2.0 mol/kg H_2O the structure of water then begins to resemble that of water at a lower pressure than the applied 25 bar or 4 kbar. At 25 bar the first peak position is shifted from 2.77 to 2.78 Å and the second peak position is

System	First peak (Å)	Second peak (Å)
H ₂ O at 25 bar	2.77	4.35
H ₂ O at 4 kbar	2.76	N/A
Aqueous TMAO at 25 bar	2.78	4.42
Aqueous TMAO at 4 kbar	2.77	3.85

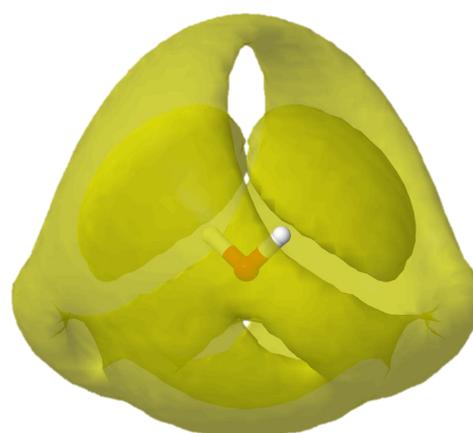
Table 4.3.2: Peak positions of the first and second peaks in the $O_w - O_w$ RDFs for pure water and aqueous TMAO at 2.0 mol/kg H₂O at 25 bar and 4 kbar calculated through EPSR.

shifted from 4.35 to 4.42 Å. This is more reminiscent of ambient water and supports the hypothesis that TMAO resists the pressure induced structural perturbations to water. This is again observed at 4 kbar, where the first peak shifts outwards from 2.76 to 2.77 Å and the second peak is now a clearly discernible peak with a position of 3.85 Å. The positions of the first and second peaks in these RDFs are summarised in table 4.3.2.

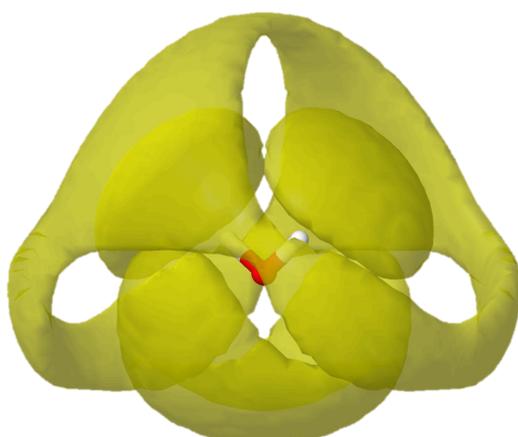
These perturbations to water structure can be visualised in 3D using the spatial density functions (SDFs) shown in figure 4.3.2. Here the yellow areas indicate the 30% most likely positions to find a neighbouring water molecule. The first hydration shell, corresponding to the first peak in the $O_w - O_w$ RDF, is therefore made up of two areas located directly above each water hydrogen and a broader area located beneath the water oxygen. The second hydration shell is then in antiphase around the first hydration shell at larger distances. For pure water at 25 bar both hydration shells are clearly visible, but at 4 kbar the second hydration shell begins to collapse into the first, moving to shorter distances and becoming less well separated. In the presence of TMAO at 2.0 mol/kg H₂O at 25 bar the SDF is similar to the SDF of pure water at 25 bar, however the area corresponding to the second hydration shell is noticeably narrower. This corresponds to a more ordered hydration shell and illustrates the power of TMAO to resist the pressure induced perturbation to water structure. The same is also true when comparing the SDFs of water and aqueous TMAO at 4 kbar. Here the second hydration shell is noticeably more separated from the first hydration shell and is overall narrower. The SDFs therefore support the findings derived from the RDFs: that TMAO acts to resist the pressure induced perturbation to water structure and in this instance behaves like a negative pressure.



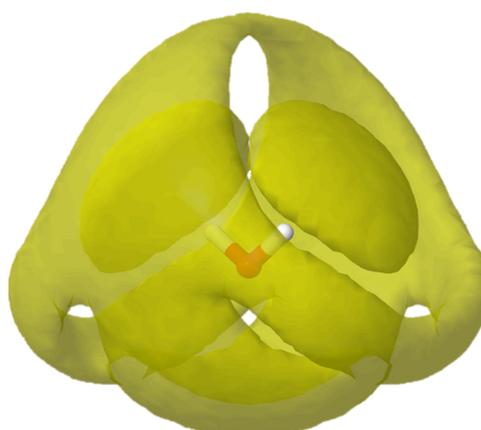
Water – 25 bar



Water – 4 kbar



Aqueous TMAO – 25 bar



Aqueous TMAO – 4 kbar

Figure 4.3.2: The spatial density functions for water and aqueous TMAO at 2.0 mol/kg H₂O at 25 bar and 4 kbar calculated through EPSR. The yellow isosurface contains the 30% most likely positions for neighbouring water molecules around a central water molecule.

System	Coordination number
H ₂ O at 25 bar	4.54
H ₂ O at 4 kbar	5.44
Aqueous TMAO at 25 bar	3.97
Aqueous TMAO at 4 kbar	4.64

Table 4.3.3: $O_w - O_w$ coordination numbers calculated over 3.38 Å by EPSR. Many statistics drawn from from > 1000 EPSR iterations mean that the coordination numbers can be determined with error values < 0.01 and are therefore not reported.

We can also examine the difference between the structure of water under pressure in/without the presence of TMAO by monitoring the $O_w - O_w$ coordination number, as described in section 2.2.5. As stated in this section, one traditionally selects a cutoff distance when one is calculating coordination numbers that corresponds to the locations of minima in the associated RDF such that specific hydration shells can be studied. In the case of pressurised water the second hydration shell begins to collapse into the first and it is no longer possible to clearly pick a cutoff distance between the first and second hydration shell. Therefore for consistency the cutoff distance for all four cases was set to 3.38 Å corresponding to the location of the first minimum in the $O_w - O_w$ RDF for pure water at 25 bar. The calculated coordination numbers are reported in table 4.3.3.

If TMAO is acting to resist the effects of large external pressure on water structure, this should be reflected by a lower water - water coordination number in aqueous TMAO compared with pure water at equivalent pressure. Table 4.3.3 shows the water - water coordination numbers in each of the four samples. This shows that as pressure increases from 25 bar to 4 kbar, the water - water coordination number increases due to the compression of the second hydration shell into the first. Upon the addition of TMAO, the water - water coordination number decreases compared with pure water at equivalent pressure, suggesting the resistive power of TMAO. However, it is difficult to comment on whether this truly supports the hypothesis, as the excluded volume effect of the TMAO molecule itself will also cause this effect regardless of how it is perturbing water structure. What is perhaps easier to comment on is the compressibility of the two solutions. An increase in coordination number from 4.54 - 5.44 in pure water when moving from 25 bar to 4 kbar is equivalent to an increase in coordination by approximately 20%. However in aqueous TMAO at 2.0 mol/kg H₂O, the increase from 25 bar to 4

kbar is only approximately 17%. This suggests that water in TMAO solutions is less compressible than pure water, which is consistent with the results found by Knierbein *et al.* [151]. Overall it is therefore clear by using the techniques available to EPSR that TMAO does indeed resist the structural perturbation to water induced by large external pressures.

4.4 Structure of Water - Extended Analysis

As overviewed in section 2.3, EPSR can be expanded using a custom built analysis routine to investigate the structure of aqueous solutions in more detail. This includes measuring parameters related to bulk water - water and TMAO - water hydrogen bonding and water dipole angle distributions for water hydrating a central TMAO oxygen/bulk water oxygen. The water dipole angle between a hydration water molecule and a central molecule species is defined as the angle between the vector which points from the hydrating water oxygen to the central TMAO oxygen/bulk water oxygen and the vector that originates at the hydrating water oxygen and bisects the two hydrating water hydrogens, parallel to the dipole moment of the water molecule, as illustrated in figure 2.3.3. This is measured over the whole box for water molecules hydrating bulk water molecules and TMAO oxygens and the results binned in increments of 2 °. Finally the bin heights are normalised to the total number of data points such that the final data represents the probability distribution. The location of the peaks are then determined by fitting the data with a function containing two summed Gaussian distributions. The results for water molecules hydrating a central bulk water molecule are shown in figure 4.4.1.

As seen in section 3.3, the water - water dipole angle distribution is defined by two peaks. The peak at shorter angles reflects water molecules that are likely donating a hydrogen bond to the central water molecule and the peak at larger angles reflects water molecules that are likely accepting a hydrogen bond from the central molecule. This data shows that upon increasing pressure from 25 bar to 4 kbar for pure water the dipole angle distribution becomes less well defined as the network is perturbed, with both peaks reducing in height and the minimum between the two peaks increasing in

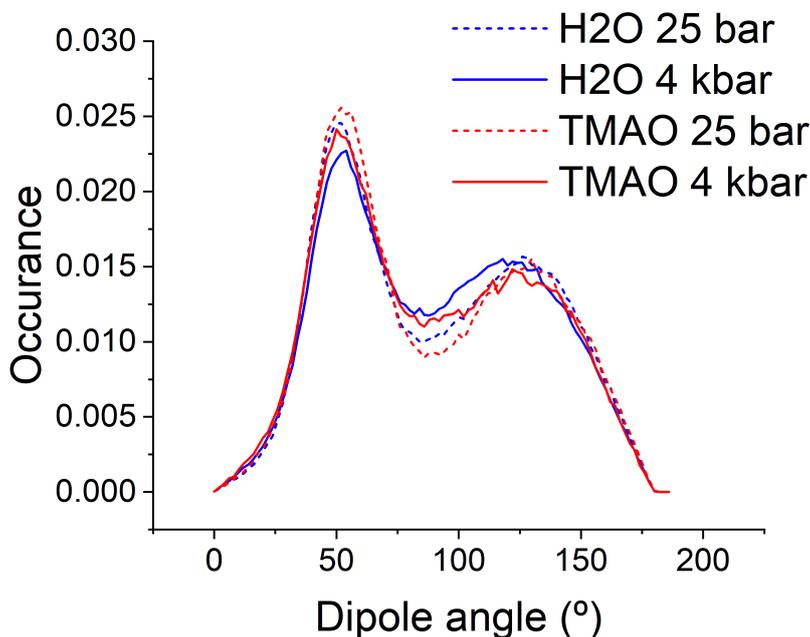


Figure 4.4.1: The dipole angle distribution for water molecules hydrating a central bulk water molecule.

height. The positions of the first and second peak for pure water at 25 bar are located at 50.9° and 127.2° respectively. Upon increasing pressure to 4 kbar the first peak moves outwards to 52.4° , and the second peak moves inwards to 120.1° . In aqueous TMAO, like in pure water, as one increases pressure from 25 bar to 4 kbar the heights of the two peaks decrease and the minimum becomes less shallow, however at equivalent pressure the peaks are taller and the minima are shallower in aqueous TMAO compared with pure water. This suggests a more well defined network overall in the presence of TMAO. The peaks also shift as well, however in this case the first peak shifts slightly inwards from 52.2° to 51.6° and the second peak shifts inwards from 127.9° to 123.1° . The peak positions are summarised in table 4.4.1.

These results again demonstrate the ability of TMAO to resist the pressure induced structural perturbation to water. At equivalent pressure the dipole angle distributions reflect a more defined overall network in aqueous TMAO compared with pure water, and the position of the second peak is much less sensitive to increasing pressure in the presence of TMAO. The position of the first peak shows no discernible trend.

System	First peak (°)	Second peak (°)
H ₂ O at 25 bar	50.9	127.2
H ₂ O at 4 kbar	52.4	120.1
Aqueous TMAO at 25 bar	52.2	127.9
Aqueous TMAO at 4 kbar	51.6	123.1

Table 4.4.1: Peak positions of the first and second peaks in the bulk water - water dipole angle distributions for pure water and aqueous TMAO at 2.0 mol/kg H₂O at 25 bar and 4 kbar. The smooth distributions as a result of many statistics drawn from from 30 iterations of the analysis routine and large EPSR box sizes mean that the peak positions can be determined with error values $< 0.1^\circ$ and are therefore not reported.

We can also monitor structural perturbation by measuring the bulk water - water hydrogen bond interaction energy, calculated as described in section 2.3.2. Here two water molecules are deemed hydrogen bonded if their oxygens are within a distance corresponding to the first minimum in the $O_w - O_w$ RDF, and a hydrogen of the hydrating water molecule must be within a distance corresponding to the first minimum in the $O_w - \text{water hydrogen } (H_w)$ RDF. Their interaction energy is then calculated using the reference potential from EPSR. Due to the issues raised previously when calculating the water - water coordination number, it is difficult to define the first minimum for pure water at 4 kbar. Therefore, again for consistency, the $O_w - O_w$ and $O_w - H_w$ cutoff distances were set at the values determined for pure water at 25 bar for all samples (3.38 Å and 2.41 Å respectively). The resultant probability distributions are shown in figure 4.4.2.

The distributions are all well fit to a single Gaussian distribution to calculate the peak position and the average hydrogen bond interaction energy. These results are summarised in table 4.4.2. This demonstrates that as the pressure is increased from 25 bar to 4 kbar and the system is perturbed that the average hydrogen bond interaction energy between two bulk water molecules becomes less stable, increasing from -17.46 ± 0.05 to -16.73 ± 0.05 . We also observe that the addition of TMAO at 2.0 mol/kg H₂O causes the water - water hydrogen bonding to be more stable at equivalent pressure, consistent with TMAO resisting the pressure induced perturbation to water structure. The hydrogen bond interaction energy between two bulk water molecules is -17.82 ± 0.05 at 25 bar and -17.23 ± 0.04 at 4 kbar. This increased stability of hydrogen bonding is also consistent with results from NMR studies on aqueous TMAO [2].

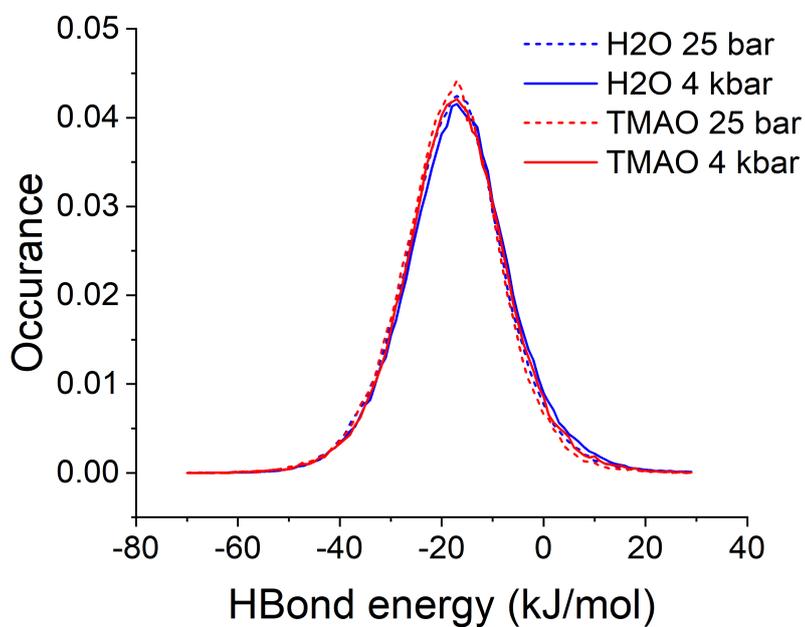


Figure 4.4.2: Hydrogen bond interaction energy between two hydrogen bonded bulk water molecules.

System	Hydrogen Bond Interaction Energy (kJ/mol)
H ₂ O at 25 bar	-17.46 ± 0.05
H ₂ O at 4 kbar	-16.73 ± 0.05
Aqueous TMAO at 25 bar	-17.82 ± 0.05
Aqueous TMAO at 4 kbar	-17.23 ± 0.04

Table 4.4.2: The peak positions in the bulk water - water hydrogen bond interaction energy distributions. Reported uncertainty is the uncertainty in the peak position calculated through performing a Gaussian fit to the raw distribution.

4.5 TMAO Hydration - Hydrophilic

We will now turn our attention to the hydration of the TMAO molecule itself so as to attempt to understand the origins of the pressure resisting effects of TMAO. To do this we will examine the analysis using EPSR and the analysis routine of neutron scattering data taken on aqueous TMAO at the two pressures. Within EPSR the 25 bar data is analysed using the TMAO reference potential taken from Meersman *et al.* [62], and the 4 kbar data is analysed using both the reference potential taken from Meersman *et al.* and the high pressure adapted reference potential V3-HP(10 kbar) developed by Hölzl *et al.* [147]. We will also examine the hydration of TMAO around the two key areas of the molecule: the hydrophilic and hydrophobic areas. The TMAO molecule has a strong dipole across the nitrogen - oxygen bond due to the large difference in charge ($N = 0.44e$ and $O = -0.65e$ in the Meersman potential and $N = 0.5932e$ and $O = -0.8599e$ in the Hölzl potential) between the two atoms. This strong dipole means that water molecules can interact favourably with the TMAO oxygen by orienting their positively charged hydrogens towards the negatively charged oxygen and donating a hydrogen bond. This favourable interaction renders the TMAO oxygen hydrophilic. By contrast the dipole moments present in the methyl groups of TMAO across the carbon - hydrogen bonds ($C = -0.26e$ and $H = 0.11e$ in the Meersman potential and $C = -0.2690e$ and $H = 0.1166e$ in the Hölzl potential) are far more modest. This means that hydrating water molecules will not interact as strongly via electrostatic forces as they will around the hydrophilic site, and that these methyl groups are therefore hydrophobic.

We will first examine the hydrophilic TMAO oxygen (O_T) - O_w RDF for aqueous TMAO at both pressures using both force fields. These are plotted in figure 4.5.1. Here we observe that unlike the $O_w - O_w$ RDF, the $O_T - O_w$ RDF is much less sensitive to pressure, likely as a result of the large dipole moment on the TMAO molecule. The peak positions stay relatively unchanged and the peak heights increase very slightly upon increasing pressure from 25 bar to 4 kbar. There is also very little difference between the Meersman potential and the Hölzl potential, leading only to a slight increase of the first peak height in the case of the Hölzl potential. This shows that the structure of water

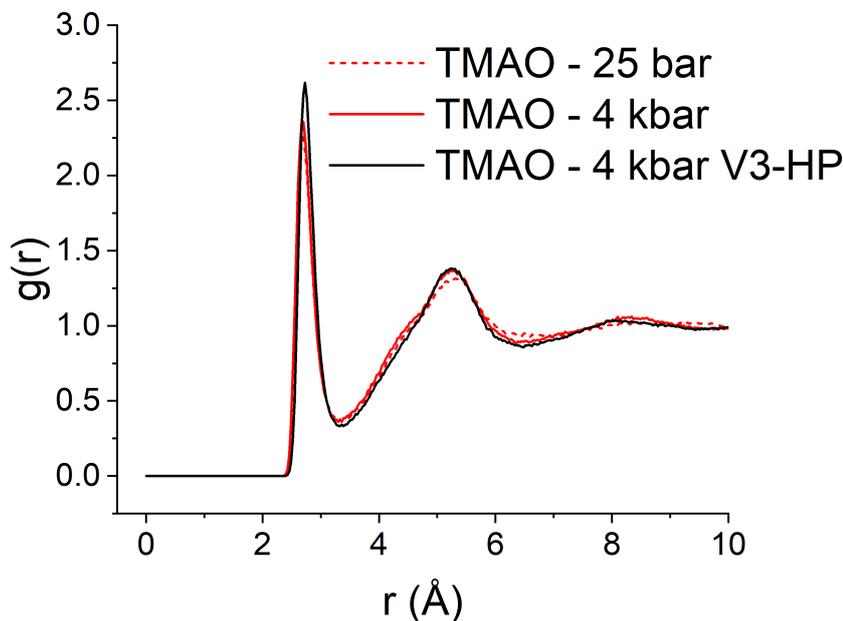


Figure 4.5.1: O_T - O_w RDF for aqueous TMAO at 2.0 mol/kg H_2O at 25 bar and 4 kbar using the Meersman reference potential (red) and the Hölzl reference potential (black).

around the hydrophilic regions of TMAO is relatively insensitive to pressure variations, and may even be slightly better defined at higher pressure, compared with bulk water which is certainly has a less well defined structure at higher pressure.

We can also monitor the $O_T - O_w$ coordination number, calculated over a range of 3.38 \AA which corresponds to the location of the first minimum in the $O_T - O_w$ RDF at 25 bar, for all samples. These results are summarised in table 4.5.1. Here we observe that at 25 bar the $O_T - O_w$ coordination number is 2.65, and that this increases to 3.04 in the case of the Meersman potential, and 3.09 in the case of the Hölzl potential. This increase in hydrophilic hydration is consistent with previous literature [64, 72, 149], which predicts an average coordination number of 3 at ambient pressure, and an increasing abundance of coordination number of 4 as pressure is increased to 10 kbar. This allows us to discuss compressibility as we did for the case of pure water. Upon increasing pressure from 25 bar to 4 kbar, the increase coordination number around the TMAO oxygen corresponds to an increase in approximately 15% using the Meersman force field and 17% in the case

System	$O_T - O_w$ coordination number
Aqueous TMAO at 25 bar	2.65
Aqueous TMAO at 4 kbar	3.04
Aqueous TMAO at 4 bar V3-HP	3.09

Table 4.5.1: $O_T - O_w$ coordination numbers calculated over 3.38 Å by EPSR. Many statistics drawn from from > 1000 EPSR iterations mean that the coordination numbers can be determined with error values < 0.01 and are therefore not reported.

System	Peak location (°)
Aqueous TMAO at 25 bar	49.4
Aqueous TMAO at 4 kbar	50.9
Aqueous TMAO at 4 bar V3-HP	49.3

Table 4.5.2: Location of the peak in the $O_T - O_w$ dipole angle distributions.

of the Hölzl force field. This suggests that not only is the structure of water around the TMAO oxygen less sensitive to pressure, it is also less compressible than bulk water consistent with the results found by Knierbein *et al.* [151].

We can also monitor structural perturbations to the hydrophilic hydration shell via the water dipole angle distribution around the TMAO oxygen. These are shown in figure 4.5.2. Unlike the water dipole angle distribution around a central bulk water molecule, the $O_T - O_w$ dipole angle distribution only shows one peak, as the central TMAO oxygen is only capable of accepting hydrogen bonds, whereas a central water molecule is capable of both donating and accepting hydrogen bonds. The peak position in each of the three cases is determined by fitting the same double Gaussian function used in section 4.4 and described in section 2.3.4. These results are summarised in table 4.5.2.

Here we observe that when using the Meersman force field that upon increasing pressure from 25 bar to 4 kbar that the water dipole angle distribution becomes slightly shorter, broader, and the peak location moves outwards from 49.4 ° to 50.9 °. This would correspond to a slightly less defined hydration shell around the TMAO oxygen. If one then uses the Hölzl force field to describe the 4 kbar aqueous TMAO data then the dipole angle distribution is essentially unchanged, with the peak moving very slightly inwards to 49.3 ° and the overall peak shape remaining unchanged. Overall, much like the observations from the $O_T - O_w$ RDFs, the dipole angle distribution is relatively

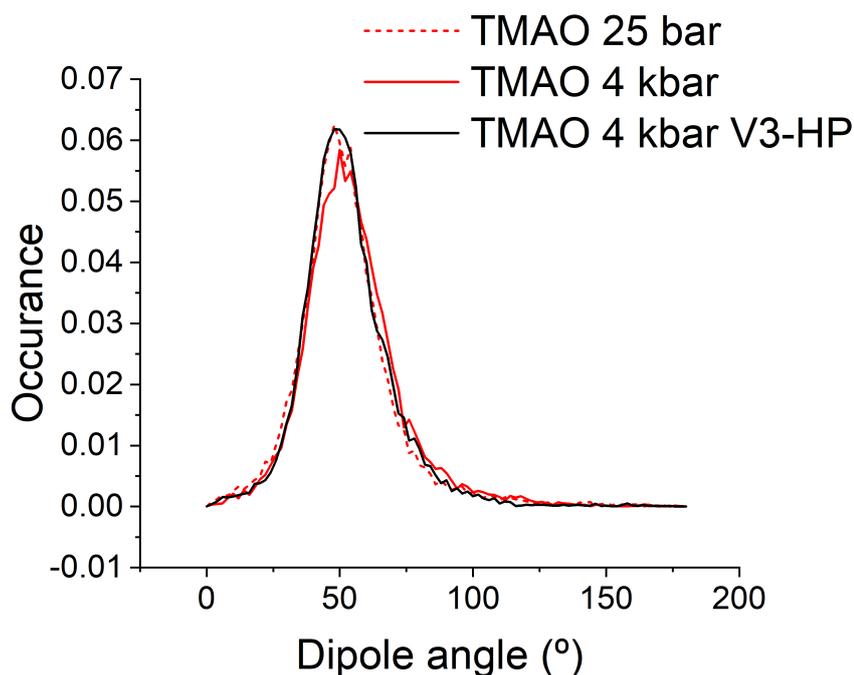


Figure 4.5.2: Water dipole angle distribution around a central TMAO oxygen in aqueous TMAO at 2.0 mol/kg H₂O at 25 bar and 4 kbar using the Meersman reference potential (red) and the Hölzl reference potential (black).

insensitive to pressure compared with the case for pure water, likely due to the strong dipole on the TMAO molecule. We also observe that the peak position is at slightly lower angles for water molecules hydrating the TMAO oxygen compared with water molecules hydrating a central bulk water molecule. This suggests more linear bonds between water and TMAO than between water and water and a stronger overall interaction, which we will now examine in more detail.

To do this we can consider TMAO - water hydrogen bonding. Here a hydrogen bond is defined as described in section 2.3.2 and as reiterated in section 4.4, where a water molecule is donating a hydrogen bond to a central TMAO oxygen if both the $O_T - O_w$ distance and $O_T - H_w$ are less than a distance corresponding to the first minima in the associated RDFs. In this case the $O_T - O_w$ cut off distance was set to be 3.38 Å and the $O_T - H_w$ cut off distance was set to be 2.48 Å calculated according to the RDFs from aqueous TMAO at 25 bar. These were again kept constant for all samples to allow for meaningful comparisons. The resultant distributions are shown in figure 4.5.3.

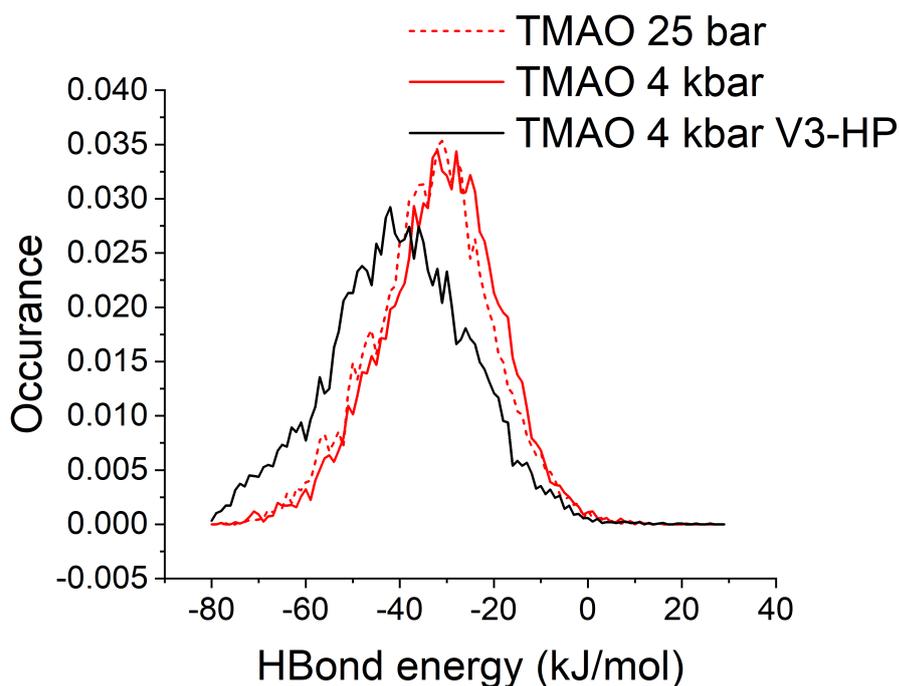


Figure 4.5.3: TMAO - water hydrogen bond energy calculated using the reference potential from EPSR in aqueous TMAO at 2.0 mol/kg H₂O at 25 bar and 4 kbar using the Meersman reference potential (red) and the Hölzl reference potential (black).

System	Peak location (kJ/mol)
Aqueous TMAO at 25 bar	-32.5 ± 0.2
Aqueous TMAO at 4 kbar	-30.6 ± 0.1
Aqueous TMAO at 4 bar V3-HP	-39.7 ± 0.2

Table 4.5.3: Location of the peak in the TMAO - water hydrogen bond energy distributions.

Here we again observe that the resultant distributions are well described by a single Gaussian function. Fitting a Gaussian function therefore allows the average hydrogen bond energy to be calculated by determining the peak position of the distribution. These results are summarised in table 4.5.3. What is immediately obvious is that the TMAO - water hydrogen bond is indeed much stronger than the water - water hydrogen bonds reported in section 4.4. This is well observed in previous literature [2, 65, 66, 70, 142, 171] and we suggest it is the key reason why the structure of water around the TMAO oxygen is much less sensitive to pressure than the structure of bulk water.

We also observe that the average hydrogen bond interaction energy calculated using

the Hölzl force field is significantly lower than the values calculated using the Meersman force field, corresponding to a much more stable hydrogen bond. However, this stark difference between the two force fields is not observed through either the $O_T - O_w$ RDFs or dipole angle distributions. The origin of this difference is therefore likely entirely due to the choice of using the solely the EPSR reference potential when calculating interaction energies. Within EPSR the potential between two atom species is the sum of the reference potential and the empirical potential. The high degree of similarity between the $O_T - O_w$ RDFs and dipole angle distributions regardless of which forcefield is employed suggests that the final sum of the reference potential and the empirical potential is essentially identical in both cases. Therefore if both potentials were used to calculate an interaction energy one would expect the final results to be highly similar. However, if one only uses the reference potential as is the case in this work, then for an essentially identically structured system using the Hölzl force field parameters rather than the Meersman force field parameters will predict much more stable hydrogen bonds as the electrostatic contributions will be more significant due to the increased dipole moment on the TMAO molecule in the Hölzl force field. This observation demonstrates the importance of carefully choosing ones reference potential and being mindful of this choice when analysing the results. It also suggests that the inclusion of the empirical potential may be an important feature to future versions of the extended analysis routine.

4.6 TMAO Hydration - Hydrophobic

We will finally consider hydration around the hydrophobic methyl groups of TMAO through the methyl carbon (C_T) - O_w and H_w RDFs and associated coordination numbers. The $C_T - O_w$ RDF is shown in figure 4.6.1. Here we observe a single clear peak located at 3.46 Å at a pressure of 25 bar. Upon increasing pressure to 4 kbar this peak shifts inwards to 3.42 Å using the Meersman potential and to 3.45 Å using the Hölzl potential while increasing in height. This change in height is also observed in the $O_T - O_w$ RDFs however the peak position is marginally more sensitive in the $C_T - O_w$ RDF.

This increase in peak height is the result of increased coordination of water molecules

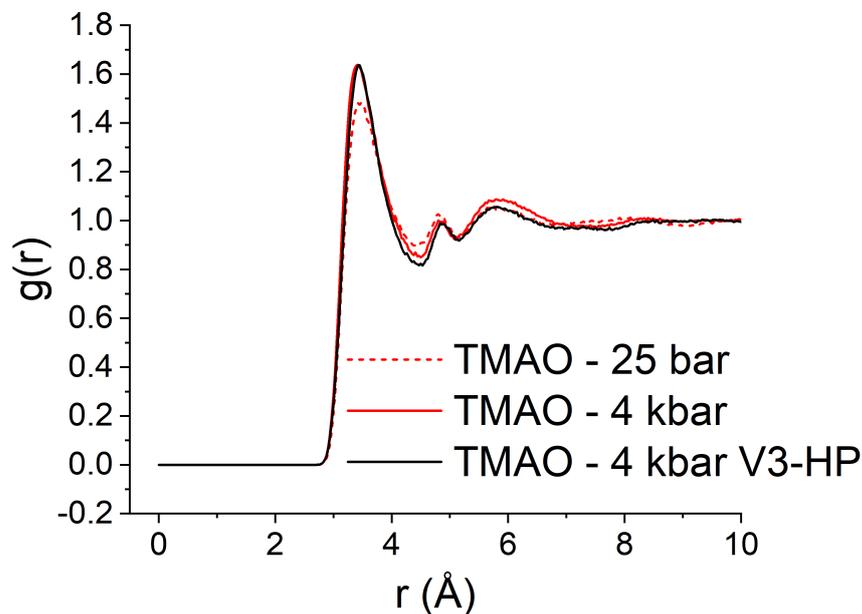


Figure 4.6.1: C_T-O_w RDF for aqueous TMAO at 2.0 mol/kg H_2O at 25 bar and 4 kbar using the Meersman reference potential (red) and the Hölzl reference potential (black).

System	$C_T - O_w$ coordination number
Aqueous TMAO at 25 bar	8.10
Aqueous TMAO at 4 kbar	9.44
Aqueous TMAO at 4 bar V3-HP	9.13

Table 4.6.1: $O_T - O_w$ coordination numbers calculated over 4.48 Å by EPSR. Many statistics drawn from from > 1000 EPSR iterations mean that the coordination numbers can be determined with error values < 0.01 and are therefore not reported.

around the methyl groups upon increasing pressure. These values are reported in table 4.6.1 and were measured using a cut off distance of 4.48 Å in all cases, corresponding to the first minima in the $C_T - O_w$ RDF. At 25 bar the methyl group coordinates a total of 8.10 water molecules. This increases to 9.44 water molecules using the Meersman potential and 9.13 using the Hölzl potential at 4 kbar, corresponding to a relative increase of 17% and 13% respectively. This is still less than the relative increase in coordination of bulk water over the same pressure range, and suggests that water around the methyl groups of TMAO is less compressible than bulk water, consistent with the results of Knierbein *et al.* [151], similar to water around the TMAO oxygen.

This means that water is less compressible around both the hydrophobic and hydrophilic parts of the molecule than it is in pure bulk water. The reduction in compressibility around the TMAO oxygen is as a result of strong hydrogen bonding between the TMAO oxygen and the surrounding water molecules, however the TMAO methyl groups have no such tendency. This can be shown by examining the $C_T - H_w$ RDFs in comparison to the $O_T - H_w$ and $O_w - H_w$ RDFs, as shown in figure 4.6.2. Here it is clear in the case of hydrophilic hydrogen bond forming central species, such as the TMAO oxygen or the water oxygen, that the resultant RDFs exhibit clear peaks and troughs, representative of the ability of the surrounding water molecules to preferentially orient their hydrogens to form hydrogen bonds with the central species. In the case of the water hydrating hydrophobic methyl groups, this tendency is reduced due to the lack of a strong dipole on the methyl group to allow for favourable electrostatic interactions, reflected by the comparatively featureless $C_T - H_w$ RDF.

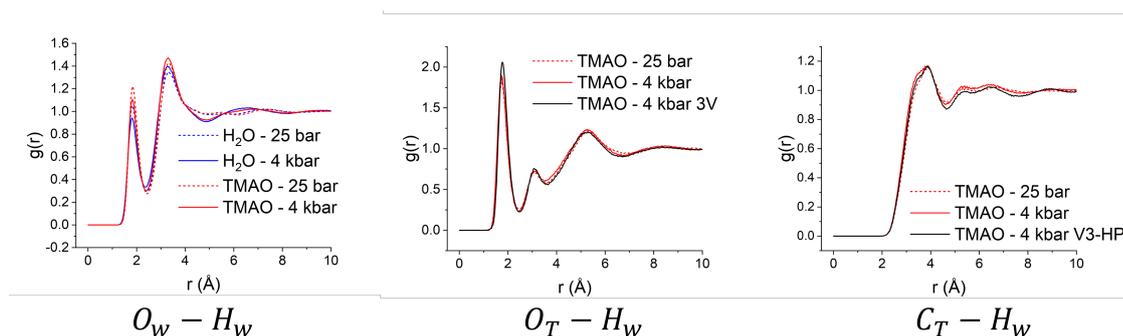


Figure 4.6.2: O_w-H_w , O_T-H_w , and C_T-H_w RDFs for aqueous TMAO at 2.0 mol/kg H_2O at 25 bar and 4 kbar. TMAO hydration is shown using the Meersman reference potential (red) and the Hölzl reference potential (black).

This means that non-hydrogen bond forming and hydrogen bond forming areas around the TMAO molecule both induce a reduction in compressibility of the surrounding water molecules relative to pure water. To resolve this apparent disparity we consider the structure of the water molecule. Water’s bent conformation, consisting of two positively charged hydrogens and two negatively charged electron pairs, is ideally suited to form a tetrahedral network via the donation of two hydrogen bonds and the acceptance of two hydrogen bonds [31]. To consider what water may look like in the absence of hydrogen bonding we can consider a liquid that does not have any propensity to hydro-

gen bond, such liquid argon just below its critical point of 150 K [267]. This also forms ordered hydration shells due to steric effects between argon atoms but has a coordination number in its first hydration shell of 6.0. This is much higher than the water - water coordination number, calculated at 25 bar in this research to be 4.54, and means in the absence of hydrogen bonding water would therefore likely be much denser.

The hydrogen bonded tetrahedral network of liquid water is therefore relatively open, and as such water molecules located in the second hydration shell of a central water molecule can be compressed into the first hydration shell. The lack of hydrogen bonding around hydrophobic methyl groups means that the location of the first hydration shell is more dependent on steric effects than the organisation of neighbouring water molecules to form hydrogen bonds, and that the immediate network will therefore be less compressible. As a reminder, water molecules around the TMAO oxygen are less compressible because the TMAO - water hydrogen bond is much stronger than that of water - water hydrogen bonding. Hence it is the open tetrahedral network of bulk water coupled with weaker hydrogen bonding than with TMAO that makes it more compressible than water in the hydration shell of either the hydrophobic or hydrophilic areas of TMAO.

4.7 Conclusions

In this chapter we have demonstrated through neutron scattering, EPSR, and extended analysis that TMAO clearly resists the structural perturbations to water as a result of a large external pressure. This is shown by the relative movement of the second peak in the $O_w - O_w$ RDF in figure 4.3.1, which moves inwards upon increasing pressure from 25 bar to 4 kbar, but that this effect is not as pronounced in the presence of TMAO at 2.0 mol/kg H_2O . This is then visualised using the $O_w - O_w$ SDF in figure 4.3.2, where the second hydration shell is narrower and better separated from the first hydration shell in the presence of TMAO at equivalent pressure.

Bulk water is also shown to be less compressible in the presence of TMAO. This is due in part to the strengthening of water - water hydrogen bonding from -17.46 ± 0.05

kJ/mol to -17.82 ± 0.05 kJ/mol at 25 bar and from -16.73 ± 0.05 to -17.23 ± 0.04 at 4 kbar by TMAO. This strengthening of water hydrogen bonding through TMAO addition is well reported in the literature [2, 65, 66, 146, 171]. It is also due to the hydration shell of water molecules around both the hydrophilic oxygen of TMAO and the hydrophobic methyl groups being less compressible than pure water.

The reduction in water compressibility around the TMAO oxygen is thought to arise from the formation of strong hydrogen bonds between the TMAO oxygen and the surrounding water molecules. Using the Meersman force field these are calculated to be -32.5 ± 0.2 kJ/mol at 25 bar and -30.6 ± 0.1 kJ/mol at 4 kbar. TMAO can therefore be thought of as providing a site within water structure that can form strong hydrogen bonds, which serves as an anchor point from which the tetrahedral network of water can build and become more stable. In the next chapter we will explore how TMAO can resist the perturbing ability of a different external stimuli, the addition of the salt $\text{Mg}(\text{ClO}_4)_2$.

Chapter 5

Dissecting the Effects of Different Solutes in Tertiary Solution: Trimethylamine N-Oxide vs $\text{Mg}(\text{ClO}_4)_2$

5.1 Introduction

As discussed in chapter 3, the introduction of ions into water results in a host of structural, dynamic, and thermodynamic perturbations to water. One of these perturbations, as shown in figure 3.2.2, is the inwards movement of the second peak in the $O_w - O_w$ RDF. This is highly similar to the perturbation arising from the application of an external pressure, as shown in chapter 4, and is well observed for a number of salt species [130]. One particular salt species for which this effect is particularly evident is the salt magnesium perchlorate ($\text{Mg}(\text{ClO}_4)_2$). Unlike the simple monovalent potassium halide salts discussed in chapter 3 which consist of a pair of oppositely charged ions each with a charge of magnitude e , $\text{Mg}(\text{ClO}_4)_2$ is made up of a single divalent Mg^{2+} cation ionically bound to two polyatomic ClO_4^- anions.

It has been previously reported using neutron diffraction and EPSR at 298 K at a concentration of 44 wt% (2.7 M), corresponding to the eutectic point (the concentration at which the freezing point of the solution is depressed to the lowest temperature of 206 K [268]), that the second hydration shell of water is compressed completely into the first hydration shell, resulting in a structure approximately equivalent to water under an external pressure of 3 GPa (30 kbar) [133]. By comparison the eutectic point of a simple monovalent salt like sodium chloride (NaCl) is 252 K at 23.3wt% [269] and the similar salt sodium perchlorate (NaClO_4) is 236 K at 52 wt% [268]. $\text{Mg}(\text{ClO}_4)_2$ is therefore clearly a very strong perturber of water structure.

The motivation behind the investigation of this particular salt and the effect it has on water in the context of life in extreme environments is due to its recent discovery in the Martian regolith [270]. Here it is present in surprisingly high quantities (Mg^{2+} at $3.3 \times 10^{-3}m$ and ClO_4^- at $2.4 \times 10^{-3}m$ where the concentrations are given in molality) and dominates the ionic makeup of the Martian regolith at the Phoenix lander site. This is coupled with the recent discoveries of periods of flowing surface water on Mars [271] and the presence of subsurface lakes [173, 174] approximately 1.5 km below the Martian south pole. The Phoenix lander site in the Vastias Borealis plains was originally chosen to study the history of water on Mars, therefore the presence of high concentrations of $\text{Mg}(\text{ClO}_4)_2$ allows us to speculate that this highly perturbing salt may also have been present in historic Matian water and may still be present in the subsurface lakes. On Earth there also exist subsurface lakes at a similar distance below the south pole which harbour life, such as Lake Vostok [272] and Lake Whillans [273]. The subsurface lakes on Mars therefore represent interesting aqueous environments that are potentially capable of harbouring life, likely rich in $\text{Mg}(\text{ClO}_4)_2$, similar to the high salt concentration environments on Earth such as salt lakes and solar salterns [274], that are protected from the combined bacteriocidal effects of $\text{Mg}(\text{ClO}_4)_2$ and ultraviolet radiation present at the Martian surface [275, 276].

Organisms living in such an environment would therefore require a method to overcome the pressurising effects of $\text{Mg}(\text{ClO}_4)_2$ on water structure in the same manner as organisms living in the deep sea on Earth. As discussed in chapter 4, one of the mech-

anisms by which marine organisms achieve this is through the accumulation of organic osmolytes like TMAO [82, 83, 86]. In this chapter we therefore seek to answer the question as to whether the organic osmolyte TMAO is capable of resisting the pressure like perturbation induced by $\text{Mg}(\text{ClO}_4)_2$ in the same manner as it was capable of resisting the perturbation to water by a physically applied pressure. To do this we will employ neutron scattering, EPSR, and extended analysis as previously to examine perturbations to water structure as a result of the addition of both solutes as in chapter 4. However, we will also expand this study to consider the dynamics of water molecules present in the various solutions through the use of NMR.

5.2 Experimental Considerations

Just as with the choice of pressure used in chapter 4, when looking at the resisting power of TMAO it is sensible to choose a $\text{Mg}(\text{ClO}_4)_2$ concentration that perturbs water in such a manner that it is clearly noticeable through neutron scattering, but not so high that any resisting power of TMAO would be overwhelmed. Unlike pressure, the effect of $\text{Mg}(\text{ClO}_4)_2$ on water structure as a function of concentration is not well characterised, hence we must come up with a manner of estimating this relationship. To do this we turn to data of the melting temperature of aqueous $\text{Mg}(\text{ClO}_4)_2$ as a function of concentration reported by Pestova *et al.* [277], reproduced in figure 5.2.1(a).

Here we observe that the melting temperature of solution shows a roughly exponential dependence on $\text{Mg}(\text{ClO}_4)_2$ concentration measured in molality until the eutectic point (shown as a star) is reached. This is well fit to an equation of form shown in equation 5.2.1, where A is an amplitude, y_0 is a y offset, x is the concentration of $\text{Mg}(\text{ClO}_4)_2$ and t_1 is a decay constant.

$$y = A \times \exp\left(\frac{-x}{t_1}\right) + y_0 \quad (5.2.1)$$

We now assume, as is the case with other salt solutions, that the density of $\text{Mg}(\text{ClO}_4)_2$

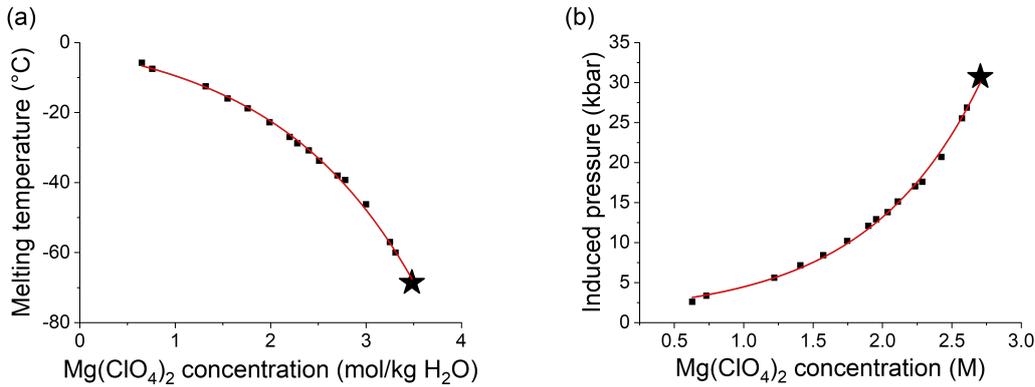


Figure 5.2.1: The melting temperature of $\text{Mg}(\text{ClO}_4)_2$ solution as a function of concentration (a) and the equivalent pressure on water structure induced by $\text{Mg}(\text{ClO}_4)_2$ estimated from the melting temperature data (b). Concentration previously studied through neutron diffraction by Lenton *et al.* [278] shown as a star in both (a) and (b).

solution is directly proportional to the wt% of the salt [279]. As we have experimentally measured the density of the solution at 44wt% (2.7 M) to be 1.3770 g/cm^3 , as shown in table 5.2.1, and the density of pure water is well established we can use this data point to convert the concentration from molality (mol/kg H₂O) in Pestova *et al.* to molarity M (mol/L). We then assume that the change in melting temperature of the solution is representative of the perturbation to water structure in the solution. This means we can say that the pressure like perturbation induced by $\text{Mg}(\text{ClO}_4)_2$ is proportional to the change in melting temperature. As experimental neutron scattering results have shown that 2.7 M $\text{Mg}(\text{ClO}_4)_2$ induces a pressure equivalent to approximately 30 kbar while reducing the melting temperature to 206 K, we can use this data point to calculate the equivalent pressure induced by all the other concentrations measured by Pestova *et al.*. This is shown in figure 5.2.1(b). Finally this is fit to an equation of the form 5.2.1 with determined fitting constants of $A = 1.07565$, x is now the concentration in molarity, $t_1 = -0.81949$ and $y_0 = 0.82427$.

Finally, we choose an induced pressure that does not cause any counteracting effects of TMAO to become overwhelmed. To do this we use the fitted equation from figure 5.2.1(b) to estimate the concentration of $\text{Mg}(\text{ClO}_4)_2$ required to induce a pressure equal to approximately 2 kbar. This reveals that 0.2 M $\text{Mg}(\text{ClO}_4)_2$ should induce a perturbation equivalent to 2.2 kbar, which is sufficient for our purposes. It is important to

note that because the melting temperature data presented by Pestova *et al.* does not extrapolate to 0 °C at 0 concentration, the induced pressure does not extrapolate to atmospheric pressure at 0 concentration. However, this is merely an estimate drawn from limited data, and hence is a reasonable enough approximation for our purposes. We are not attempting to pick a concentration of TMAO and $\text{Mg}(\text{ClO}_4)_2$ to cancel each other out, we simply wish to observe a measurable resisting effect of TMAO.

Next we must estimate the relative resisting ability of TMAO so as to pick a sensible concentration to resist 0.2 M $\text{Mg}(\text{ClO}_4)_2$, for which at the time of this work there also existed no direct experimental evidence. To estimate this we turn to measurements of the concentration of TMAO in the wet muscle tissue of bony fish as a function of depth reported by Yancey *et al.* [86] which is reproduced in figure 5.2.2(a).

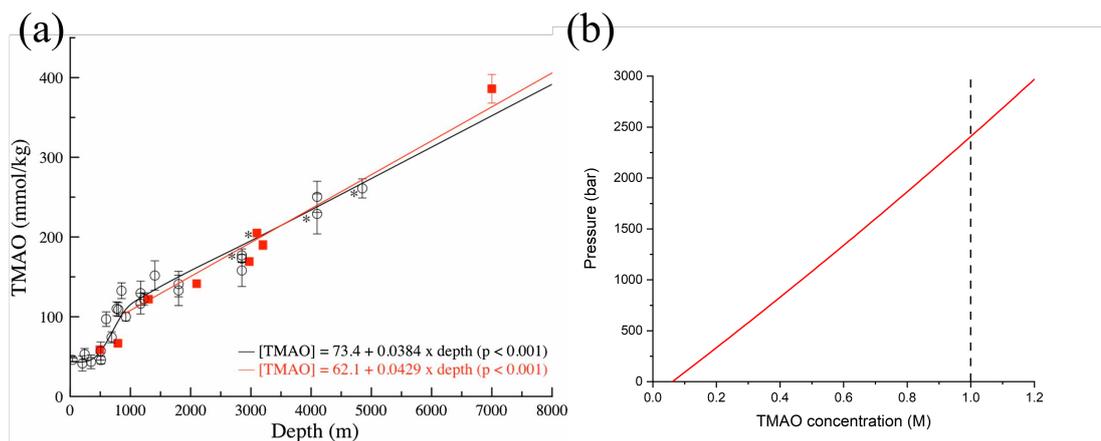


Figure 5.2.2: The concentration of TMAO in the muscle tissue of bony fish as a function of depth published by Yancey *et al.* [86] (a) and the estimated pressure resisting ability of TMAO as a function of concentration estimated from the Yancey data (b).

As it is well known that pressure in the ocean increases 1 atm (0.01325 bar) for every 10 m of depth, and we have experimentally validated that the addition of TMAO only negligibly affects the density of solution as shown in table 5.2.1, we can use this data to predict the pressure resisting ability of TMAO. For the reported Yancey data [86] the authors provide two fits to the experimental data. For our purposes we will choose the fit shown in red, as this better fits the data at higher depths and therefore pressures, as

shown in equation 5.2.2.

$$[TMAO] = 62.1 + 0.429 \times \text{depth} \quad (5.2.2)$$

We must again convert the concentration in molality into molarity. The definition molality is defined as the number of moles of a solute per 1000 g of solvent, the mass of a solution in grams m of molality x is therefore given in equation 5.2.3, where M_s is the molar mass of the solute, equal to 51.11 g/mol for TMAO.

$$m = xM_s + 1000 \quad (5.2.3)$$

We have also shown that the density of solution of TMAO is relatively independent of TMAO concentration and we can assume it remains fixed at the value for pure water, $\rho = 1000$ g/L. Therefore the volume in litres V of the solution mass m is given by equation 5.2.4.

$$V = \frac{m}{\rho} = \frac{75.11x + 1000}{1000} \quad (5.2.4)$$

Molarity is defined as the mols of solute per L of solution, therefore the concentration in molarity M from molality x is now given by equation 5.2.5

$$M = \frac{x}{V} = \frac{1000x}{75.11x + 1000} \quad (5.2.5)$$

This can be rearranged to make the molality the subject, such that it can be substituted back into equation 5.2.2. It is then multiplied by a factor of 1000 to account for the fact that the concentration in the Yancey expression is given in mmol/kg. The final expression is shown in equation 5.2.6

$$\frac{M \times 10^6}{1000 - 75.11M} = 62.1 + 0.0429 \times \text{depth} \quad (5.2.6)$$

Next we use the fact that pressure increases above atmospheric pressure by 1.01325 bar for every 10 m in depth, shown mathematically as $P = 1.01325 + 0.101325 \times \text{depth}$, to express equation 5.2.6 as a function of pressure P rather than depth. This is shown in equation 5.2.7.

$$\frac{M \times 10^6}{1000 - 75.11M} = 62.1 + 0.0429 \left[\frac{P - 1.01325}{0.101325} \right] \quad (5.2.7)$$

Finally we rearrange this equation to make P the subject, as shown in equation 5.2.8.

$$P = \left[\frac{M \times 10^6}{1000 - 75.11M} - 62.1 \right] \left[\frac{0.101325}{0.0429} \right] + 1.01325 \quad (5.2.8)$$

This function is plotted in figure 5.2.2(b). As stated previously, we have chosen an $\text{Mg}(\text{ClO}_4)_2$ concentration of 0.2 M, corresponding to an induced pressure of 2.2 kbar. Using equation 5.2.8, we can state that 1.0 M TMAO should resist a presence of 2.4 kbar (point shown by dotted line in figure 5.2.2). The estimations based on previous available data therefore suggest that a solution of 1.0 M TMAO and 0.2 $\text{Mg}(\text{ClO}_4)_2$ may provide balance between structural perturbations that would be observable through neutron diffraction. We therefore used these concentrations for the experimental study. We will also used a mixture of 1.0 M TMAO and 2.7 M $\text{Mg}(\text{ClO}_4)_2$ as this is a $\text{Mg}(\text{ClO}_4)_2$ that has been previously studied with neutron diffraction. Therefore to observe the resisting power of TMAO 6 different samples sets were prepared, these were: pure water, 0.2 M $\text{Mg}(\text{ClO}_4)_2$, 1.0 M TMAO, 0.2 M $\text{Mg}(\text{ClO}_4)_2$ + 1.0 M TMAO, 2.7 M $\text{Mg}(\text{ClO}_4)_2$, and 2.7 M $\text{Mg}(\text{ClO}_4)_2$ + 1.0 M TMAO.

To analyse the data with EPSR hydrogen - deuterium substitution was again employed on each of the sample sets. For the pure water sample and the samples containing only $\text{Mg}(\text{ClO}_4)_2$ as a solute species 3 isotopic variants were produced: H_2O , D_2O , and

System	Density (g/cm ³) \pm 0.0001
Pure H ₂ O	0.9970
1.0 M TMAO	0.9994
0.2 M Mg(ClO ₄) ₂	1.0281
1.0 M TMAO + 0.2 M Mg(ClO ₄) ₂	1.0275
2.7 M Mg(ClO ₄) ₂	1.3770
1.0 M TMAO + 2.7 M Mg(ClO ₄) ₂	1.3513

Table 5.2.1: Densities of solutions as measured at 25 °C using an Anton Parr DMA 4100 M densitometer.

HDO. For the samples containing TMAO 7 different isotopic variants were produced: H₂O + H-TMAO, H₂O + D-TMAO, D₂O + H-TMAO, D₂O + D-TMAO, HDO + H-TMAO, HDO + D-TMAO, and HDO + HD-TMAO, where H/D corresponds to the whether the hydrogens on the host molecule were in fully protinated or fully deuterated state, and HD samples were prepared using an equal molar ratio of H and D samples. This therefore yields a total of 30 individual samples. To inform the EPSR simulation the density of the fully protinated samples were measured using an Anton Parr DMA 4100 M densitometer, and are reported in table 5.2.1.

It is interesting to note that before neutron scattering measurements and EPSR analysis have even begun that simply measuring the densities of the solutions seem to offer support for the idea that TMAO may be able to resist the pressurising effect of Mg(ClO₄)₂. Here we observe that the addition of TMAO into pure water results in a slight increase in density due to the addition of a larger molecule. In the case of both 0.2 M Mg(ClO₄)₂ and 2.7 Mg(ClO₄)₂ the pressurising effect of Mg(ClO₄)₂ causes the density of the solutions to rise in both cases. However upon the subsequent addition of TMAO into both Mg(ClO₄)₂ solutions, the recorded density is shown to drop slightly. This would be consistent with TMAO resisting this pressurising effect.

5.3 Structure of Water - EPSR

To calculate the structural perturbations present in the system EPSR was performed on the raw neutron scattering data after correction with Gudrun, as described in section

System	Water Molecules	TMAO Molecules	Mg ²⁺ ions	ClO ₄ ⁻ ions	Box Dimension (Å)
Pure H ₂ O	5000	-	-	-	53.1329
1.0 M TMAO	4104	80	-	-	51.0434
0.2 M Mg(ClO ₄) ₂	4104	-	16	32	49.9840
1.0 M TMAO + 0.2 M Mg(ClO ₄) ₂	4104	80	16	32	51.2654
2.7 M Mg(ClO ₄) ₂	4224	-	260	520	54.4902
1.0 M TMAO + 2.7 M Mg(ClO ₄) ₂	4224	80	260	520	55.5774

Table 5.3.1: Number of molecules used in each simulation box and length of cubic box.

2.1.6. This started with the building of simulated boxes of molecules representative of the experimental samples. The box size and number of molecules of each species are displayed in table 5.3.1. As described in section 2.2.1, box sizes of approximately 5000 molecules are ideal to capture the desired features while maintaining a reasonable computing time. It is important here to note that the number of water molecules is higher in the samples containing 2.7 M Mg(ClO₄)₂. This is because Mg(ClO₄)₂ is highly hygroscopic and hence absorbs water molecules from the air. The amount of water absorbed was estimated from the predicted differential cross section in Gudrun, which is highly sensitive to H content and therefore serves as a good measure. The reference potentials for TMAO were taken from Meersman *et al.* [62, 63] and the reference potentials for water and Mg(ClO₄)₂ were taken from Lenton *et al.* [133]. These are summarised in section C. The final fits of the simulations to the experimental data for each of the 6 datasets are presented in section C. These produce R factors between 0.0004 - 0.0007, and can therefore be regarded as very high quality fits.

Just as in section 4.3, the simplest way to assess the resisting power of TMAO against the pressure like perturbation of Mg(ClO₄)₂ is to examine the O_w - O_w RDF in each of the six sample sets. These results are shown in figure 5.3.1. However, these results are not as clear as those determined from the O_w - O_w RDFs in section 4.3. In the case of aqueous TMAO under high pressure, the differences between the peaks in each of the four sample sets were clearly evident, however here it becomes difficult to clearly

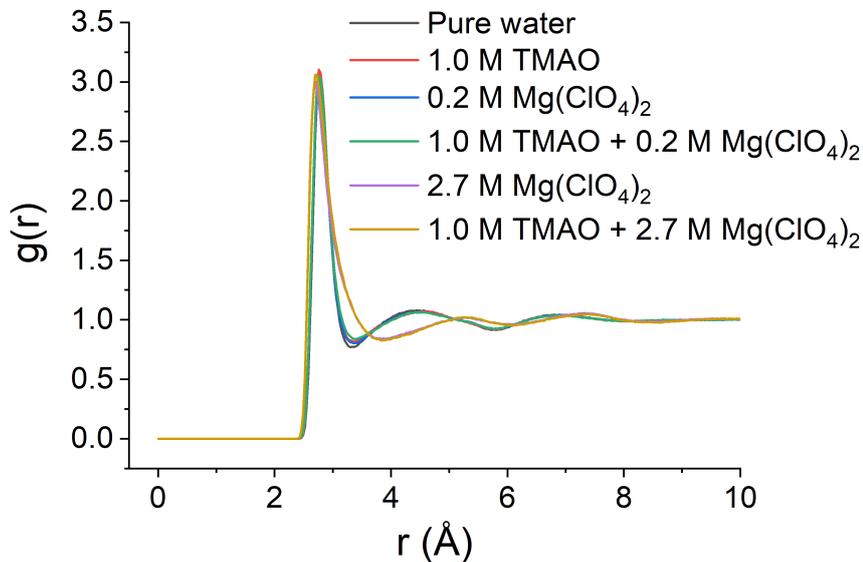


Figure 5.3.1: $O_w - O_w$ RDFs for the six sample sets studied in this section as calculated through EPSR.

distinguish the effects of TMAO at the different $Mg(ClO_4)_2$ concentrations.

In the case of pure water, the first and second peaks in the RDF are located 2.78 and 4.45 Å respectively, consistent with previous literature [89]. Consistent with the data reported in chapter 4, the first peak is relatively insensitive to TMAO addition, moving slightly inwards to 2.77 Å and the second peak moves outwards to 4.56 Å. Surprisingly, and contrary to the expectation that 0.2 M $Mg(ClO_4)_2$ would induce a pressurising effect, a similar shift is observed in 0.2 M $Mg(ClO_4)_2$, where the first peak is unchanged relative to pure water and the second peak moves slightly outwards to 4.49 Å. However the counteracting effects are observed if one considers the positions of the combined 1.0 M TMAO + 0.2 M $Mg(ClO_4)_2$ sample, where the first peak occurs at 2.77 Å, and the second peak occurs between the values calculated for the individual 0.2 M $Mg(ClO_4)_2$ and 1.0 M TMAO samples at 4.50 Å. Consistent with previous literature [133], the first peak in 2.7 M $Mg(ClO_4)_2$ shifts inwards to 2.71 Å, and the second peak compresses completely into the first peak. The third peak therefore becomes the new second peak and occurs at 5.27 Å. The ability of TMAO to resist the pressure-like perturbation to water structure induced by $Mg(ClO_4)_2$ is then exhibited again as in the combined 2.7 M

System	First peak (Å)	Second peak (Å)
Pure H ₂ O	2.78	4.45
1.0 M TMAO	2.77	4.56
0.2 M Mg(ClO ₄) ₂	2.78	4.49
1.0 M TMAO + 0.2 M Mg(ClO ₄) ₂	2.77	4.50
2.7 M Mg(ClO ₄) ₂	2.71	5.27
1.0 M TMAO + 2.7 M Mg(ClO ₄) ₂	2.72	5.29

Table 5.3.2: Peak positions of the first and second peaks in the $O_w - O_w$ RDFs.

Mg(ClO₄)₂ + 1.0 M TMAO sample the first peak only shifts inwards to 2.72 Å and the second peak now occurs at 5.29 Å. These peak positions are summarised in table 5.3.2.

The ability of TMAO to resist the pressure-like perturbation to water structure induced by Mg(ClO₄)₂ is also indicated if we consider the relative peak heights of the first peak in the $O_w - O_w$ RDFs. The addition of 1.0 M TMAO causes the first peak to increase in height relative to pure water. The addition of either 0.2 M or 2.7 M Mg(ClO₄)₂ causes a roughly equal reduction in peak height. However the peak height in the combined Mg(ClO₄)₂ and TMAO samples is relatively unchanged. The overall insensitivity of the $O_w - O_w$ RDFs to TMAO addition in Mg(ClO₄)₂ solutions is reflected in the water - water SDFs, as shown in figure 5.3.2. Here we observe, consistent with previous literature [133], that the structure of water is like that of water under a high pressure in the presence of 2.7 M Mg(ClO₄)₂. Upon the subsequent addition of 1.0 M TMAO the SDF moves marginally further outwards, but is rather unchanged.

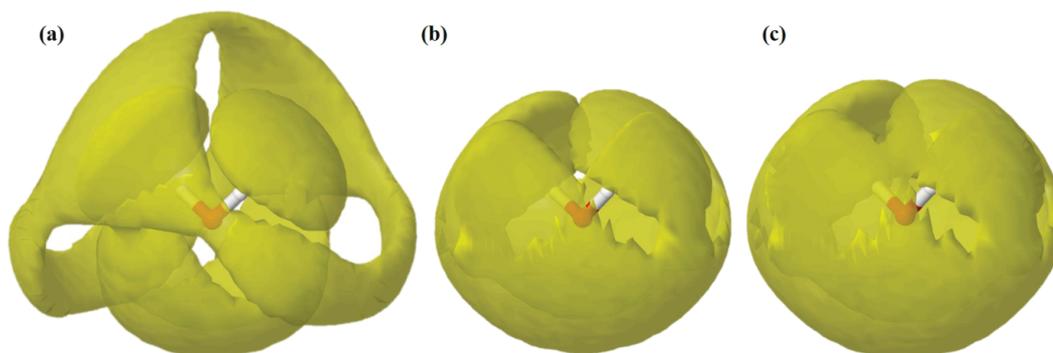


Figure 5.3.2: Water - water SDFs for pure water (a), 2.7 M Mg(ClO₄)₂ (b), and 1.0 M TMAO + 2.7 M Mg(ClO₄)₂. These surface contours contain the 30% probable locations of a neighbouring water molecule within 5 Å.

The ability of TMAO to resist the pressure-like perturbation to water structure induced by $\text{Mg}(\text{ClO}_4)_2$ as measured by the $\text{O}_w - \text{O}_w$ RDF and the water water SDFs is therefore observed, however it is very subtle and more detailed analysis is required to determine whether this is a real effect. This will be discussed in section 5.5.

5.4 Solute-Solvent and Solute-Solute Interactions - EPSR

Through EPSR we can also examine solute-solvent interactions to examine the hydration around the hydrophobic and hydrophilic parts of the TMAO molecule, as well as the hydration around the Mg^{2+} and ClO_4^- ions. The $\text{O}_T - \text{O}_w$, $\text{C}_T - \text{O}_w$, $\text{Mg}^{2+} - \text{O}_w$, and $\text{Cl}_p - \text{O}_w$ RDFs are presented in figures 5.4.1 and 5.4.2, where the T subscript represents TMAO as the host molecule and p subscript represents ClO_4^- as the host ion. As reported in chapter 4, TMAO strongly coordinates water molecules around its hydrophilic oxygen, as indicated by the tall narrow first peak in the associated RDF, but the hydration around the hydrophobic group is less well defined. Both TMAO - O_w RDFs are relatively insensitive to $\text{Mg}(\text{ClO}_4)_2$ addition when compared with the $\text{O}_w - \text{O}_w$ RDFs, as was the case for increasing pressure. Whereas the addition of 2.7 M $\text{Mg}(\text{ClO}_4)_2$ causes the second hydration shell of water around a central water molecule to be completely collapsed into the first, the addition of $\text{Mg}(\text{ClO}_4)_2$ only induces a very modest compaction of water structure around TMAO. In 1.0 M TMAO the first peak in the $\text{O}_T - \text{O}_w$ RDF occurs at 2.69 Å and the first peak in the $\text{C}_T - \text{O}_w$ occurs at 3.49 Å. Both these peaks shift slightly inwards to 2.68 and 3.44 Å upon the addition of 0.2 M $\text{Mg}(\text{ClO}_4)_2$, and shift further inwards to 2.63 and 3.40 Å upon the addition of 2.7 M $\text{Mg}(\text{ClO}_4)_2$. These results are summarised in table 5.4.1. Therefore, as was the case with increasing pressure, TMAO is able to resist the pressurising effect of $\text{Mg}(\text{ClO}_4)_2$ by preserving the structure of water in its immediate vicinity.

The Mg^{2+} and ClO_4^- ions are shown to very strongly coordinate water molecules, as evidenced by the first peaks in both the $\text{Mg}^{2+} - \text{O}_w$, and $\text{Cl}_p - \text{O}_w$ RDFs. This is due to the high charge of the ions resulting in very strong electrostatic interactions. When one compares the height of the first peak in the O_T (charge = -0.65e), Cl_p (charge on ion

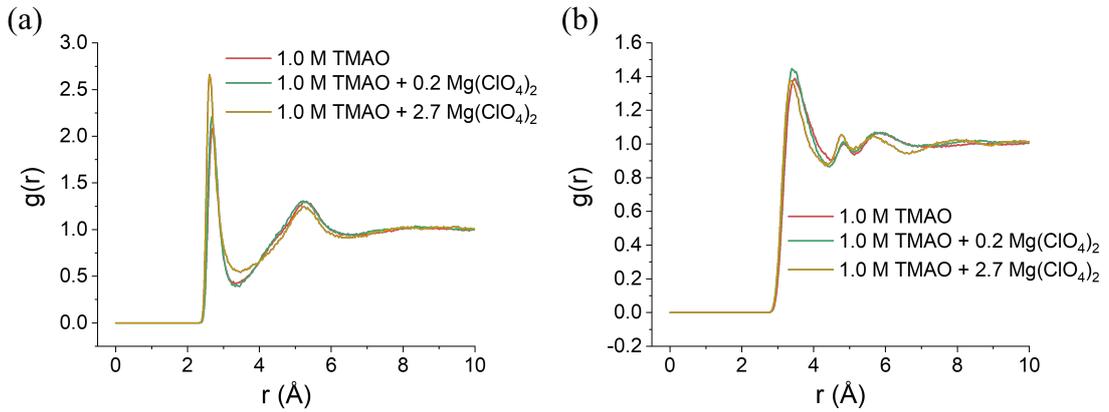


Figure 5.4.1: $O_T - O_w$ (a) and $C_T - O_w$ (b) RDFs for the three relevant sample sets studied in this section as calculated through EPSR.

System	$O_T - O_w$ peak (Å)	$C_T - O_w$ peak (Å)
1.0 M TMAO	2.69	3.49
1.0 M TMAO + 0.2 M $Mg(ClO_4)_2$	2.68	3.44
1.0 M TMAO + 2.7 M $Mg(ClO_4)_2$	2.63	3.40

Table 5.4.1: Peak positions of the first peaks in the $O_T - O_w$ and $C_T - O_w$ RDFs.

= -e), and Mg^{2+} (charge on ion = 2e) - O_w RDFs one observes that the height of the first peak increases with increasing charge of the central species. This is also coupled by a further reduced sensitivity of the peak position to concentration/TMAO addition with increasing charge of the central species. In the case of the first peak in the $Cl_p - O_w$ RDFs, which is recorded at 3.95 Å in 0.2 M $Mg(ClO_4)_2$, there is then a slight inward movement upon the addition of 1.0 M TMAO to 3.93 Å. In 2.7 M $Mg(ClO_4)_2$ with or without 1.0 M TMAO the first peak moves further inwards to 3.91 Å. In the $Mg^{2+} - O_w$ RDFs the position of the first peak is unchanged all all instances, occurring at 1.61 Å. These peak shifts are summarised in table 5.4.2. The hydration around the ions demonstrates again that $Mg(ClO_4)_2$ strongly perturbs water structure both in the bulk, as evidenced by the $O_w - O_w$ RDFs presented in figure 5.3.1 and in its immediate vicinity.

As we are now considering the competing effects of two different solute species in a tertiary solution, rather than the effect of one solute species in a binary solution that is subjected to an external pressure, we can begin to consider solute - solute interactions

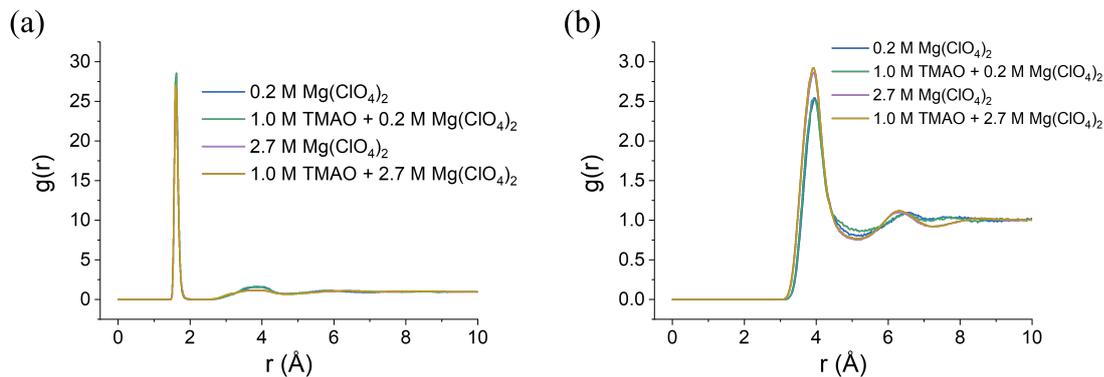


Figure 5.4.2: Mg^{2+} - O_w (a) and Cl_p - O_w (b) RDFs for the four relevant sample sets studied in this section as calculated through EPSR.

System	Mg^{2+} - O_w peak (Å)	Cl_p - O_w peak (Å)
0.2 M $\text{Mg}(\text{ClO}_4)_2$	1.61	3.95
1.0 M TMAO + 0.2 M $\text{Mg}(\text{ClO}_4)_2$	1.61	3.93
2.7 M $\text{Mg}(\text{ClO}_4)_2$	1.61	3.91
1.0 M TMAO + 2.7 M $\text{Mg}(\text{ClO}_4)_2$	1.61	3.91

Table 5.4.2: Peak positions of the first peaks in the Mg^{2+} - O_w and Cl_p - O_w RDFs.

in a way that we could not in chapter 4. The ionic association around the TMAO oxygen is shown in the O_T - Mg^{2+} and O_T - Cl_p RDFs in figure 5.4.3. Here it is clear that like water molecules, Mg^{2+} ions strongly coordinate the negatively charged TMAO oxygen at a distance of 1.54 Å due to strong favourable electrostatic interactions as shown by the tall narrow first peak. Surprisingly the height of the first peak is greater in 1.0 M TMAO + 0.2 M $\text{Mg}(\text{ClO}_4)_2$ compared with 1.0 M TMAO + 2.7 M $\text{Mg}(\text{ClO}_4)_2$, indicating a more well defined coordination. Unsurprisingly the coordination number calculated over 2.2 Å increases in the case of 2.7 M $\text{Mg}(\text{ClO}_4)_2$ compared with 0.2 M $\text{Mg}(\text{ClO}_4)_2$ from 0.02 to 0.20. This suggests that the more poorly defined hydration shell may be due to slightly more frequent occurrences of a single TMAO oxygen coordinating multiple Mg^{2+} ions, resulting in increased electrostatic repulsion between the similarly charged cations and causing the coordination shell to be more poorly defined. In both cases the coordination of cations, while well defined, occurs infrequently, as shown by associated coordination numbers much less than 1.0. Conversely the coordination of Cl_p ions around the TMAO oxygen is very weak, as shown by lack of clearly defined peaks

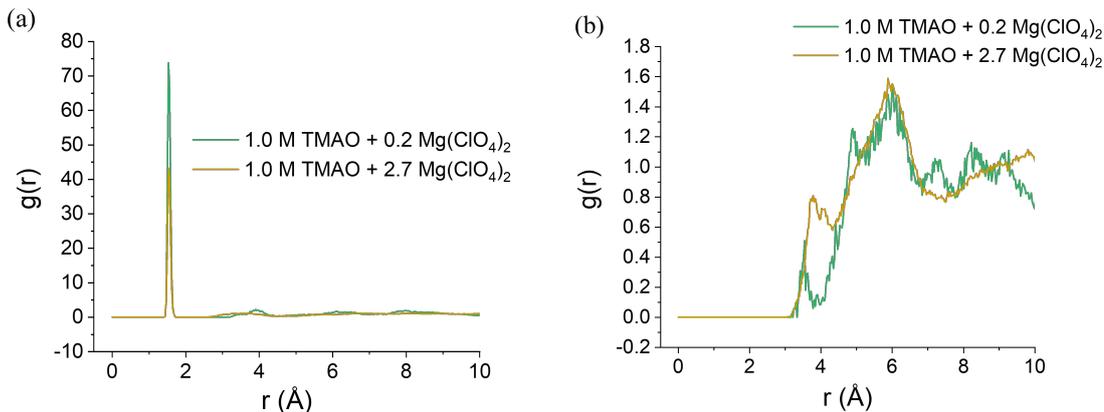


Figure 5.4.3: $O_T - Mg^{2+}$ (a) and $O_T - Cl_p$ (b) RDFs for the two relevant sample sets studied in this section as calculated through EPSR.

in the the $O_T - Cl_p$ RDF. This is likely due to unfavourable electrostatic interactions between the negatively charged TMAO oxygen and the Cl_p^- anion.

We can now also consider the coordination of ions around the hydrophobic regions of TMAO. This is shown via the $C_T - Mg^{2+}$ and $C_T - Cl_p$ RDFs in figure 5.4.4. Here we again observe that Mg^{2+} coordinates reasonably strongly around the TMAO methyl group, as shown by a tall narrow first peak, but far more weakly that it coordinates around the TMAO oxygen. It is again very infrequent, with a coordination number calculated over 4.0 Å equal to 0.03 in 1.0 M TMAO + 0.2 M Mg(ClO₄)₂ and 0.22 in 1.0 M TMAO + 2.7 M Mg(ClO₄)₂. Conversely the coordination of Cl_p^- ions around the TMAO methyl groups as a much more well defined series of hydration shells. This association is likely the result of two mechanisms. The first is the slight favourable interaction between the negatively charged perchlorate anion and the solvent exposed slightly positively charged methyl hydrogens (charge = 0.11e). The second is likely to be hydrophobic style association. As the perchlorate anion is polyatomic (containing 4 oxygen atoms tetrahedrally bonded to a central chlorine atom), it has a relatively large volume over which a charge of -e is distributed, hence it has a relatively low charge density. This means it behaves more hydrophobically than smaller more highly charged ions like Mg^{2+} [31], and can therefore associate via hydrophobic interactions with the TMAO methyl group. This results in a coordination number calculated over 5.30 Å of 0.12 in 1.0 M TMAO + 0.2 M Mg(ClO₄)₂ and 1.44 in 1.0 M TMAO + 2.7

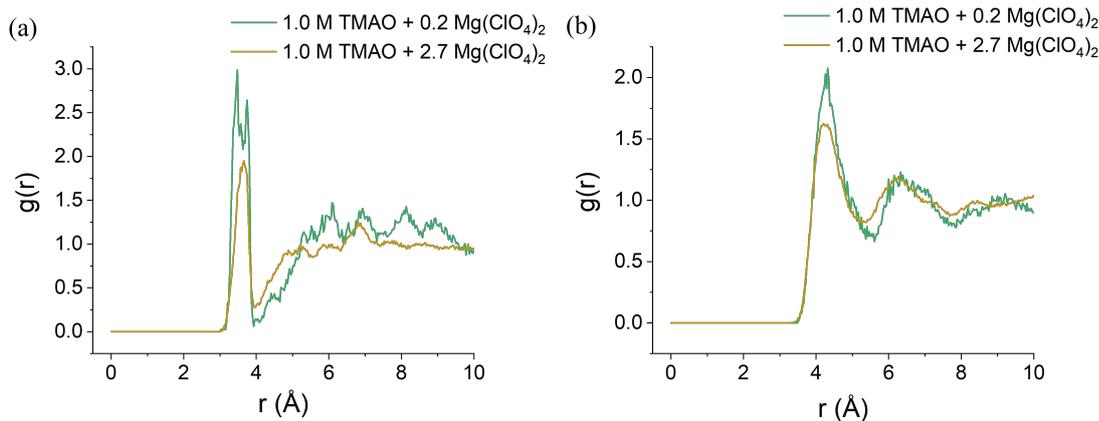


Figure 5.4.4: $C_T - \text{Mg}^{2+}$ (a) and $C_T - \text{Cl}_p$ (b) RDFs for the two relevant sample sets studied in this section as calculated through EPSR.

M Mg(ClO₄)₂. This more hydrophobic style hydration is therefore the most prominent solute-solute interaction present in the TMAO + Mg(ClO₄)₂ tertiary solutions.

5.5 Structure of Water - Extended Analysis

As discussed in section 5.3, the ability of TMAO to resist the pressure like perturbation to water structure induced by Mg(ClO₄)₂ is very subtle and difficult to identify using the means typically available through EPSR. Hence we will use the extended analysis routine to further examine structural perturbations present in the system. We will first consider bulk water - water hydrogen bonding calculated in the same manner as described in section 2.3. The final normalised distributions of bulk water - water hydrogen bond interaction energy calculated using the reference potentials of Lenton *et al.* [133] and Meersman *et al.* [62, 63] are shown in figure 5.5.1.

As previously, the raw distributions are well fit with a Gaussian distribution to determine the average value by determining the centre of the Gaussian distribution fit. These values along with the associated uncertainty calculated through the fitting procedure are shown graphically in figure 5.5.2.

By studying this data the resisting ability of TMAO is much easier to identify. In pure water the average hydrogen bond strength between two bulk water molecules is

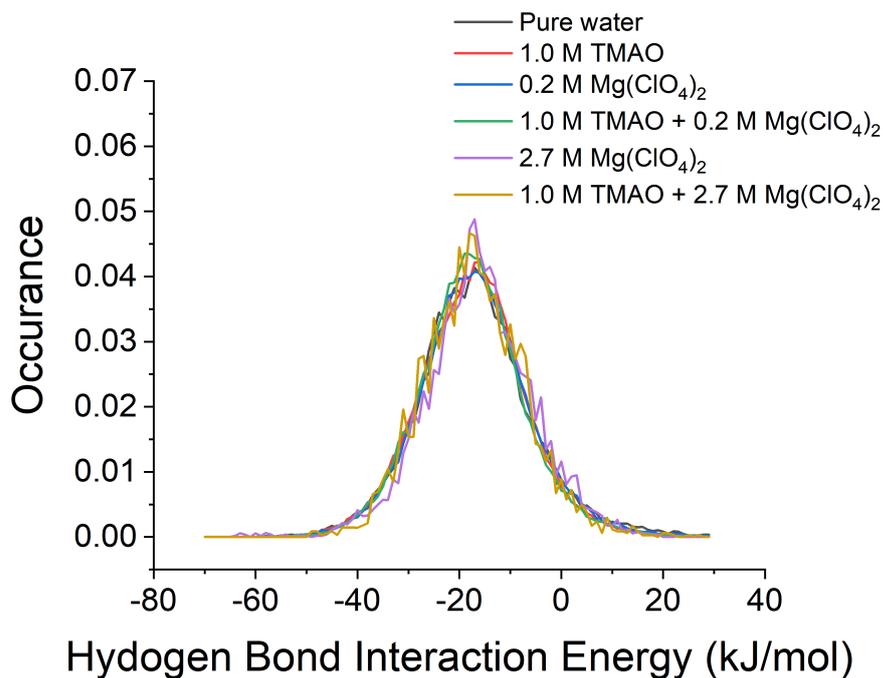


Figure 5.5.1: Normalised distribution bulk water - water hydrogen bond interaction energies calculated using the extended analysis routine.

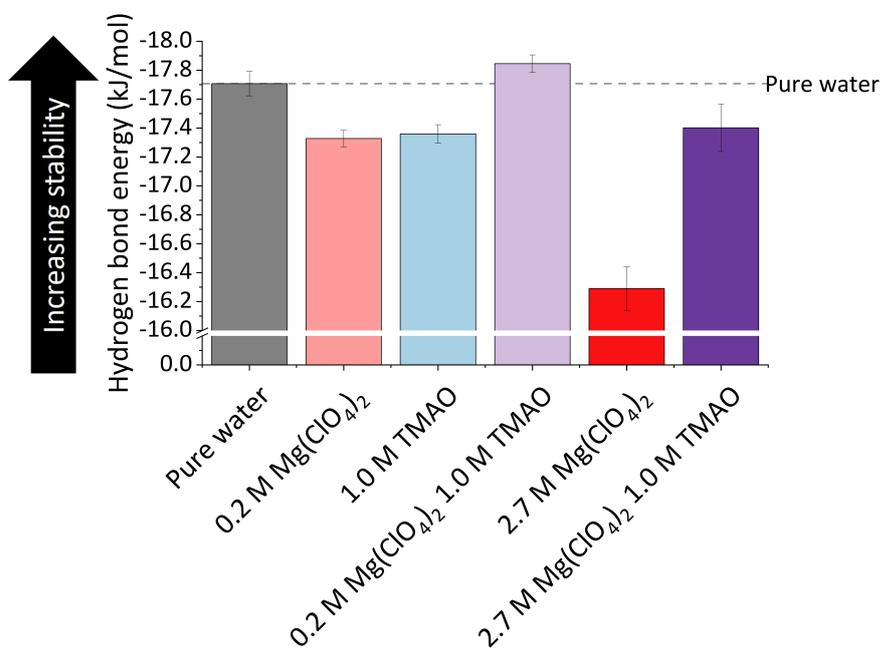


Figure 5.5.2: Average bulk water - water hydrogen bond interaction energies calculated by fitting a Gaussian distribution to the raw hydrogen bond interaction energy distributions.

calculated to be -17.71 kJ/mol. Unsurprisingly given the pressurising effect observed in 2.7 M $\text{Mg}(\text{ClO}_4)_2$, the addition of 0.2 M $\text{Mg}(\text{ClO}_4)_2$ results in a decrease in stability of the bulk water - water hydrogen bonding as the average hydrogen bond interaction energy increases to -17.33 kJ/mol. Surprisingly, and in contrast to the results displayed in chapter 4, the addition of 1.0 M TMAO results in a decrease in stability in bulk water - water hydrogen bonding relative to pure water to -17.36 kJ/mol. This is potentially due to the excluded volume effect of the TMAO molecule, or as a result of inadequate accumulations of EPSR iterations. As EPSR proceeds through a Monte Carlo simulation, it is possible for the simulation to become temporarily “trapped” in less energetically favourable states. Within EPSR the effect this has on data is minimised by performing a large number (at least 1000) iterations, however the analysis routine only samples 20-30 of these iterations due to limitations in computing time. It is therefore possible that the EPSR iterations used to calculate this data point may not have been perfectly representative of the system. It could also potentially be due to solely using the reference potential when calculating the interaction energy rather than also using the empirical potential, as overviewed in section 4.5.

However, regardless of this single data point, the resisting power of TMAO is well observed if we consider the average bulk water - water hydrogen bond interaction energy in the combined 1.0 M TMAO + 0.2 M $\text{Mg}(\text{ClO}_4)_2$ dataset. Here the stability of the bulk water - water hydrogen bond is increased relative to the 0.2 M $\text{Mg}(\text{ClO}_4)_2$ to -17.85 kJ/mol. This effect is even more pronounced if we consider the two datasets containing 2.7 M $\text{Mg}(\text{ClO}_4)_2$. Here in 2.7 M $\text{Mg}(\text{ClO}_4)_2$ the average bulk water - water hydrogen bond interaction energy becomes much less stable, increasing to -16.29 kJ/mol, however upon the subsequent addition of 1.0 M TMAO the stability is increased to -17.40 kJ/mol. From this data we can therefore say that TMAO resists the pressurising effect to water structure induced by $\text{Mg}(\text{ClO}_4)_2$, in the same manner as it resists perturbations due to a large external pressure.

As discussed in section 4.5 and 5.4, TMAO serves as a structural anchor point through an incompressible hydrophobic shell and a strongly hydrogen bond accepting oxygen. In 1.0 M TMAO the extended analysis routine predicts that each TMAO oxygen accepts

2.19 ± 0.1 hydrogen bonds with an average hydrogen bond interaction energy of -35.5 ± 0.3 kJ/mol. Upon addition of 0.2 M $\text{Mg}(\text{ClO}_4)_2$ this is reduced to 2.01 ± 0.01 hydrogen bonds and the average hydrogen bond is less stable at -31.8 ± 0.2 kJ/mol. The number of hydrogen bonds and stability are further reduced in 1.0 M TMAO + 2.7 M $\text{Mg}(\text{ClO}_4)_2$ to 1.33 ± 0.01 and -30.4 ± 0.4 kJ/mol respectively. In section 5.4 we suggested that Mg^{2+} was strongly, if infrequently, coordinated by the TMAO oxygen. This is consistent with the results observed through this analysis routine as association of Mg^{2+} around the TMAO oxygen would both limit the volume available for water molecules to occupy and donate a hydrogen bond, and electrostatically screen the negative charge on the TMAO oxygen inhibiting its ability to act as a hydrogen bond acceptor.

5.6 Nuclear Magnetic Resonance

The results of the extended analysis routine therefore suggest that TMAO can indeed preserve the hydrogen bond network of water against the pressure-like perturbation to water structure induced by $\text{Mg}(\text{ClO}_4)_2$. However, unlike experiments at high pressure which require specialised and often expensive equipment, these experiments on tertiary solutions at standard pressure can be verified and expanded using standard NMR techniques, as will be described in the following section.

As described in section 2.4, the position of the ^1H NMR peak corresponding to the water hydrogen is indicative of the strength of hydrogen bonding. A stronger hydrogen bond results in increased deshielding of the hydrogen bond, resulting in a higher Larmor frequency and a movement of the peak to higher ppm (downfield). The results from EPSR and the extended analysis routine would therefore predict a downfield peak shift with increasing TMAO concentration and an upfield peak shift with increasing $\text{Mg}(\text{ClO}_4)_2$ concentration. This is indeed the case, as shown by figure 5.6.1. Here a positive peak shift is induced due to increasing TMAO concentration which opposes the negative peak shift induced by $\text{Mg}(\text{ClO}_4)_2$. This is also consistent with the results observed in section 4.4, where increased water - water hydrogen bond stability was observed with increasing TMAO concentration. This result helps demonstrate the overall

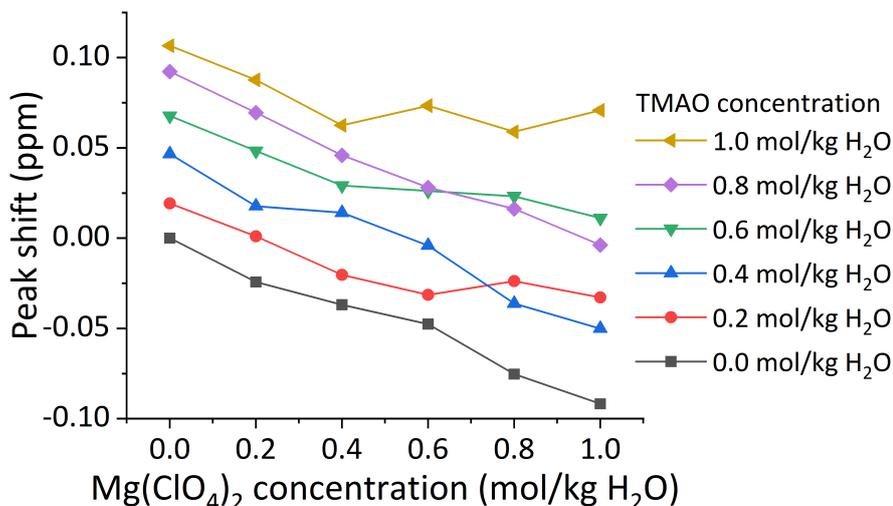


Figure 5.6.1: Relative peak position of the ^1H NMR peak of water in binary and tertiary solutions of TMAO and $\text{Mg}(\text{ClO}_4)_2$. Concentrations studied in 0.2 mol/kg H_2O increments from 0 to 1 mol/kg H_2O . As these results were obtained to serve as a supplement to the structural data produced through neutron scattering and EPSR, only a single measurement was taken for each salt at each concentration, and hence the data are displayed without error bars.

reliability of the analysis routine when studying binary and tertiary aqueous solutions.

However, as detailed in section 2.4 and 3.5, NMR is not limited to structural measurements. NMR can also be used to determine perturbations to the dynamic properties of the system. This is shown by two measurements in this thesis: the inverse T_1 relaxation time and the water self-diffusion coefficient D . In the fast limit ($\omega\tau_c \ll 1$) the inverse T_1 relaxation time is proportional to the rotational correlation time of the water molecules and serves as a good measure of perturbations to system rotational dynamics. The values determined across the same concentration range as figure 5.6.1 are shown for binary and tertiary solutions of TMAO and $\text{Mg}(\text{ClO}_4)_2$ in figure 5.6.2. Here we observe a rather different result than the results of the structural analysis. The inverse T_1 relaxation time and self diffusion coefficient of pure water are determined to be $T_1^{-1} \approx 3.0$ s and $D \approx 2.3 \times 10^{-9}$ m²/s, consistent with previous literature [211, 280]. Upon increasing $\text{Mg}(\text{ClO}_4)_2$ concentration both the rotational and diffusive dynamics of the water molecules are slowed as shown by the increasing inverse T_1 relaxation time and decreasing D . However, whereas TMAO opposed the structural perturbations to water induced by $\text{Mg}(\text{ClO}_4)_2$, here TMAO appears to act additively and slows the rotational

and diffusive dynamics yet further.

This is in good agreement with previous literature [156]. Theory states that solute species that interact more strongly with water molecules than water molecules interact with one another will result in slowed dynamics due to longer residence times and the increased energy barrier to be overcome by fluctuations in the local network that would allow a bond switching event to take place. Both TMAO and Mg^{2+} ions interact more strongly with neighbouring water molecules than water molecules interact with each other due to increased electrostatic interactions, hence a slowing in system dynamics is expected. Conversely, previous literature has shown that the interactions between water molecules and ClO_4^- anions are weaker than water - water interactions [156, 247, 281], and therefore one would expect increased dynamics in the vicinity of the anion. These results show that any accelerating effects are minimal in comparison to the retarding effects of the other solute species.

These results also suggest that unlike the dynamic and structural properties of aqueous potassium halides discussed in chapter 3, which exhibited highly heterogeneous micro environments between hydration and bulk water, the effects of $\text{Mg}(\text{ClO}_4)_2$ and TMAO are more global. This is in agreement with the results presented in section 5.5, where the bulk water hydrogen bond interaction energy was shown to perturb with increasing $\text{Mg}(\text{ClO}_4)_2$ and TMAO concentration. The raw data from the spin inversion recovery pulse sequence used to determine the T_1 relaxation time can be used to demonstrate this more global effect. Example data for this is shown in figure 5.6.3. Here we observe that the data is well fit to a single exponential, indicative of a single dominant rotational correlation time, and minimal effects from distinct micro environments with individual rotational correlation times which would require multiple exponential terms to fit.

We can also consider the self diffusion coefficient D data in conjunction with the inverse T_1 relaxation time data to demonstrate this more global effect. If the perturbations to water dynamics by TMAO and $\text{Mg}(\text{ClO}_4)_2$ are more global then it follows that the rotational and diffusive dynamics of the water molecule should be governed by the same microscale viscosity. The diffusion coefficient D is related to viscosity η by

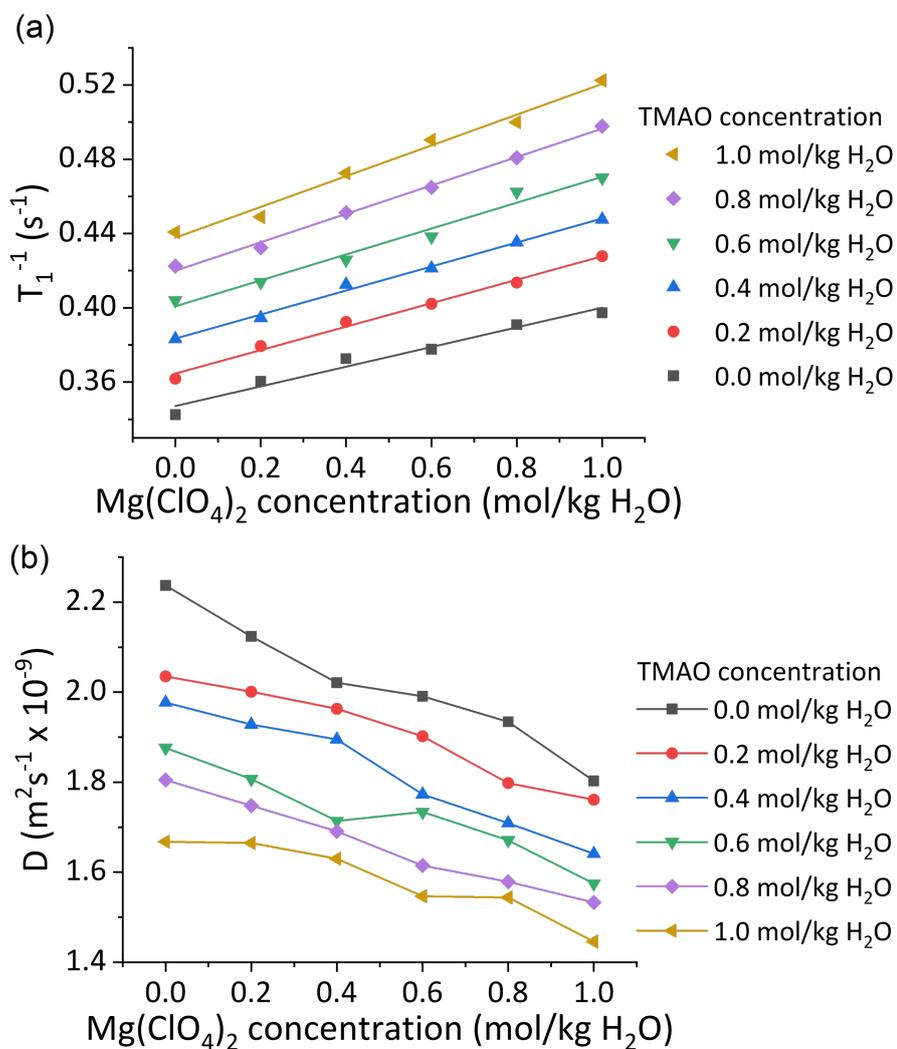


Figure 5.6.2: T_1 relaxation time (a) and self diffusion coefficient D of water in binary and tertiary solutions of TMAO and $\text{Mg}(\text{ClO}_4)_2$. Concentrations studied in 0.2 mol/kg H_2O increments from 0 to 1 mol/kg H_2O . As these results were obtained to serve as a supplement to the structural data produced through neutron scattering and EPSR, only a single measurement was taken for each salt at each concentration, and hence the data are displayed without error bars.

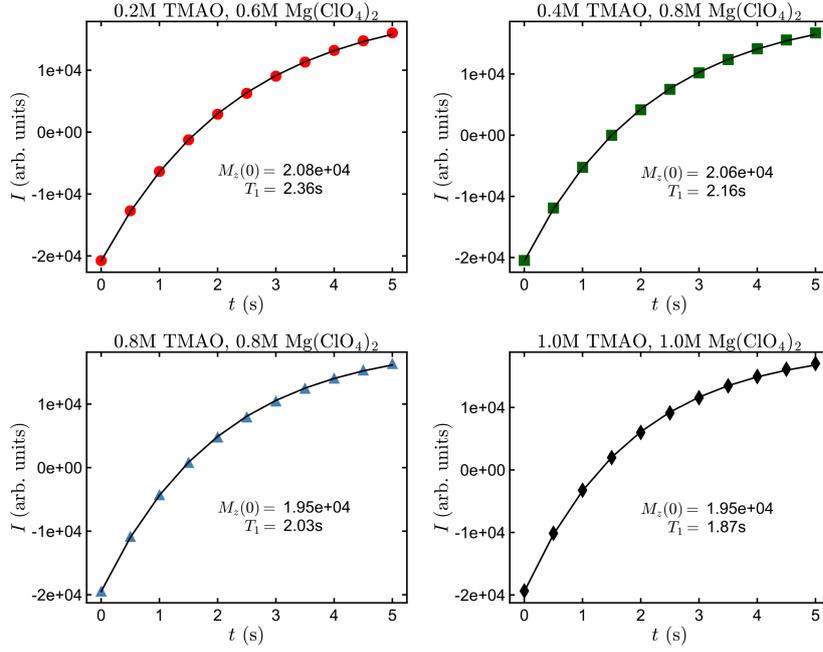


Figure 5.6.3: The raw data for the magnetisation in the z direction of the water hydrogens following the spin inversion pulse sequence for example concentrations of $\text{Mg}(\text{ClO}_4)_2$ and TMAO. Data fit to an equation of form 5.2.1 to determine the characteristic spin-lattice relaxation time T_1 .

the Stokes-Einstein relation, as shown in equation 5.6.1 where T is temperature, k_B is Boltzmann's constant, and r_h is the hydrodynamic radius of the molecule [282].

$$D = \frac{k_B T}{6\pi\eta r_h} \quad (5.6.1)$$

We can also relate the rotational correlation time τ_c of the water molecule, rather than its diffusive characteristics, to a viscosity through the Debye-Einstein relationship, shown in equation 5.6.2.

$$T_1^{-1} \propto \tau_c = \frac{4\pi\eta r_h^3}{3k_B T} \quad (5.6.2)$$

Therefore if the effects of the two solutes are more global and the translational and the rotational characteristics of the water molecule are governed by the same local microviscosity then $D \propto \tau_c^{-1} \propto T_1$ [283]. To validate this we can plot D against T_1 for

each studied sample and should in theory obtain a linear correlation between the two parameters. This is shown in figure 5.6.4 and does indeed demonstrate that the rotational and diffusive motions of water molecules in the tertiary solutions studied here are governed by the same microviscosity. This supports the idea of a more global perturbation to water structure and dynamics by TMAO and $\text{Mg}(\text{ClO}_4)_2$ and also supports the assumption that the T_1 relaxation time is dominated by rotational motion.

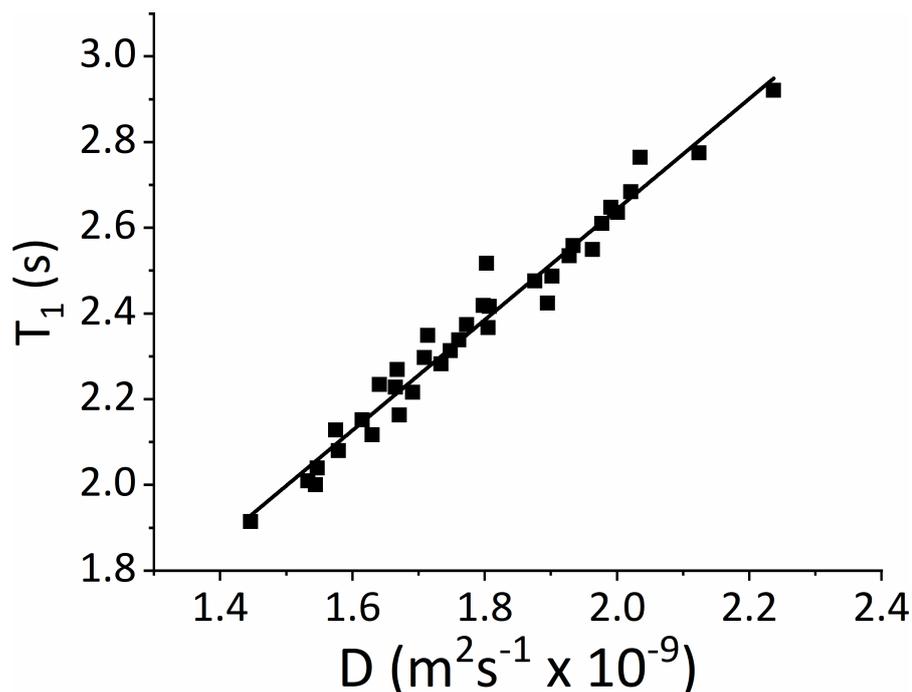


Figure 5.6.4: The T_1 relaxation time and self diffusion coefficient for water molecules in tertiary solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO. Straight line fit to help guide the eye and demonstrate a directly proportional relationship between the two parameters, demonstrating that the rotational and diffusive characteristics of water molecules are governed by the same local viscosity. Each data point corresponds to a single sample with no repeat measurements so error bars are not reported.

5.7 Dynamic Fingerprint in EPSR - The Bifurcated Hydrogen

We have now used NMR to validate the results on structural perturbations to water in tertiary $\text{Mg}(\text{ClO}_4)_2$ and TMAO solutions. The question now arises: can we now do the opposite and investigate dynamic perturbations demonstrated by NMR using EPSR? To

do this we consider the work of Laage and *et al.* [26, 65, 155–157, 164], who present a theory of molecular reorientations in aqueous solutions through the so called “extended jump model”. This model states that water molecule dynamics consists of two components, each with a characteristic lifetime: (i) Jump reorientation, where water molecules hydrogen bonded to an overcoordinated neighbour will experience a large angular jump ($\approx 68^\circ$ in pure water) to switch hydrogen bonding partners to an undercoordinated neighbour. This proceeds through an intermediate called a bifurcated hydrogen where the switching water molecule will temporarily hydrogen bond to both neighbours through a single hydrogen. This event is shown in figure 5.7.1(a). (ii) Frame reorientation, where pairs of water molecules diffuse through solution while maintaining their hydrogen bond. It therefore follows that a system with increased/decreased rotational dynamics should show a greater/lesser abundance of bifurcated hydrogens as bond switching events become more/less frequent. The abundance of this conformation can be monitored through the extended analysis routine and normalised to the total measured number of water - water hydrogen bonds. This data is shown in figure 5.7.1(b).

Here we observe that with increasing $\text{Mg}(\text{ClO}_4)_2$ concentration that the relative abundance of bifurcated hydrogens is reduced, and that this is increasingly true in the presence of 1.0 M TMAO. This is consistent with the retardation of rotational dynamics observed through the T_1 measurement in NMR. It is important to mention here that the perturbations to rotational motion monitored by the T_1 relaxation time are modelled on the assumption that rotation of molecules is a diffusive process, rather than the rapid jump like process we investigate through the bifurcated hydrogen abundance. This is due to the difference in scale between the two techniques. When using EPSR and the extended analysis routine we are considering a very small number of water molecules, and consider individual motions of individual molecules to be jump like. When using NMR we are considering an ensemble average over a much larger samples size, and ensemble average of many individual jump like motions would appear diffusive. Hence the two results can be meaningfully compared. This is the first instance of which I am aware of a dynamic model like the extended jump model has being applied to EPSR to infer results about dynamic perturbations to a system. It allows us to state that $\text{Mg}(\text{ClO}_4)_2$

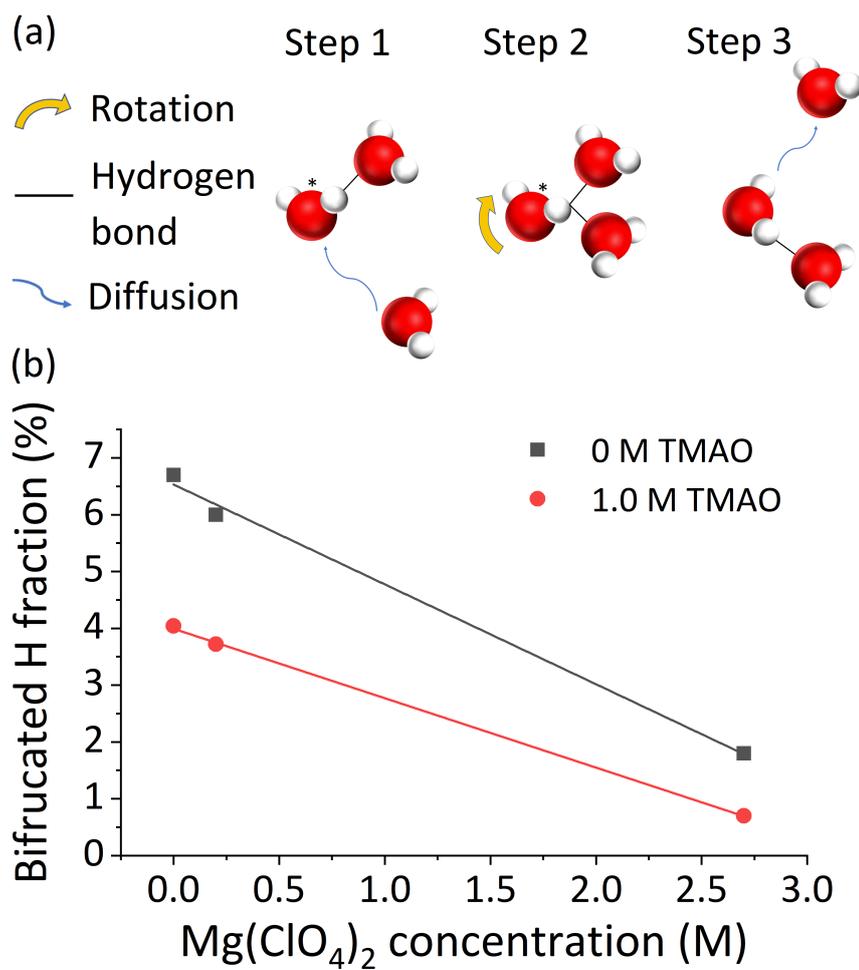


Figure 5.7.1: (a) A schematic representation of the bifurcated hydrogen conformation occurring as an intermediate during a hydrogen bond switching event. (b) The relative abundance of bifurcated hydrogen conformations as a function of solute concentration.

and TMAO perturb water structure in an opposing manner but perturb water dynamics in an additive manner through both EPSR and NMR and validate each technique with the other.

5.8 Quantification of Relative Perturbing Ability of Solutes

We can now consider quantifying each solute's perturbing ability relative to the other solute. To do this we will plot structural and dynamic properties of the tertiary solutions using a single effective total concentration that is determined from the concentrations of the individual solutes. This is described in equation 5.8.1, where $[x]$ is the concentration of species x and g is a weighting parameter that describes how much stronger TMAO is as a perturbing agent relative to $\text{Mg}(\text{ClO}_4)_2$.

$$[\text{Eff}] = [\text{Mg}(\text{ClO}_4)_2] + g[\text{TMAO}] \quad (5.8.1)$$

Data plotted in this manner will therefore lie on a single linear master curve when the value of g is a correct representation of the system. By performing a straight line fit on the data for varying values of g we can monitor the quality of the fit through the calculated r^2 value, which quantifies the fit on a scale of 0 (poor fit) to 1 (perfect fit). The g value with the highest r^2 value therefore corresponds to the correct relative perturbing ability of TMAO to $\text{Mg}(\text{ClO}_4)_2$. This procedure was conducted using five datasets: (i) the average bulk water - water hydrogen bond interaction energy, (ii) the inverse T_1 data, (iii) the ^1H NMR peak shift data, (iv) the diffusion coefficient data, (v) and the bifurcated hydrogen data. The resultant master curves using a $|g|$ value of 1.54 are plotted in figure 5.8.1. Here it is important to note that the structural measures of the system will demonstrate the opposing effects of both solutes through negative g values, and the additive dynamic measures will have positive g values, so the absolute value of g is plotted to allow for meaningful comparison.

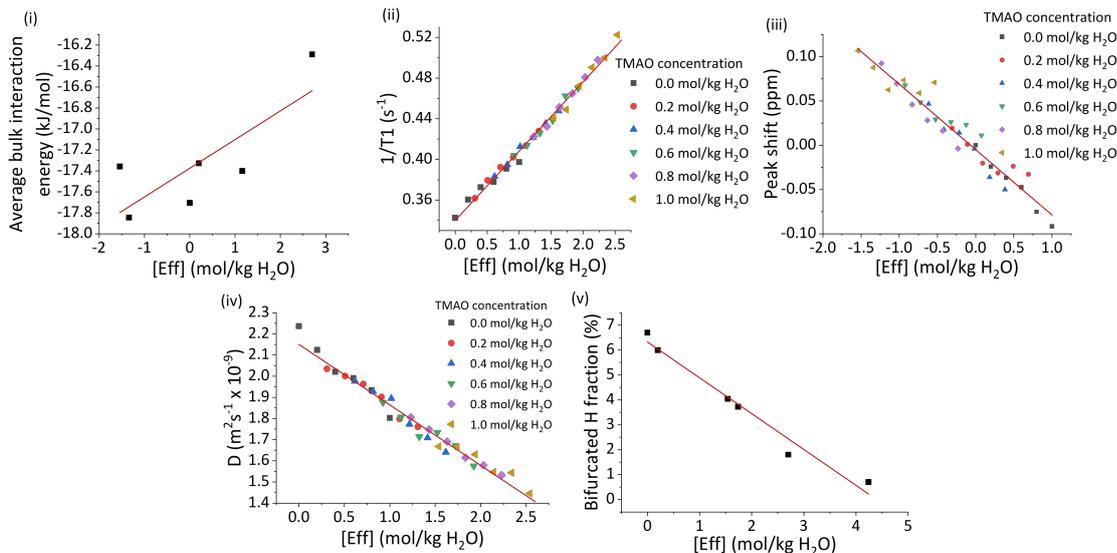


Figure 5.8.1: Structural and dynamic characteristics of the systems plotted using the effective total solute concentration using a $|g|$ factor of 1.54. Plotted values are: (i) the average bulk water - water hydrogen bond interaction energy, (ii) the inverse T_1 data, (iii) the ^1H NMR peak shift data, (iv) the diffusion coefficient data, (v) and the bifurcated hydrogen data. Straight line fits are shown.

The variation r^2 with $|g|$ for each dataset is then shown in figure 5.8.2. It is worth noting here that the r^2 values are much lower for the average bulk water hydrogen bond interaction energy data. This is likely due to the number of assumptions that are necessary to obtain this data. These include the intrinsic limitations of neutron diffraction and EPSR which require significant post acquisition correction through Gudrun as well as definitions of hydrogen bonding itself, as discussed in chapter 2. The optimum value of $|g|$ therefore corresponds to the peak position of the resultant data, which was determined by numerical differentiation.

The peak positions and average value are reported in table 5.8.1. This shows a reasonable consistency in the $|g|$ value when using two different experimental techniques and three different experimental apparatus to consider both structural and dynamic perturbations to the system. If one considers the similarity between the $|g|$ values determined for both the T_1 data and the D data, one also observes consistency in $|g|$ values across different time and length scales. The diffusion coefficient measurement uses a lag time of 60 ms, during which a water molecule will have a root-mean-square deviation of 20-30 μm according to the measured diffusion coefficients. This corresponds to approximately

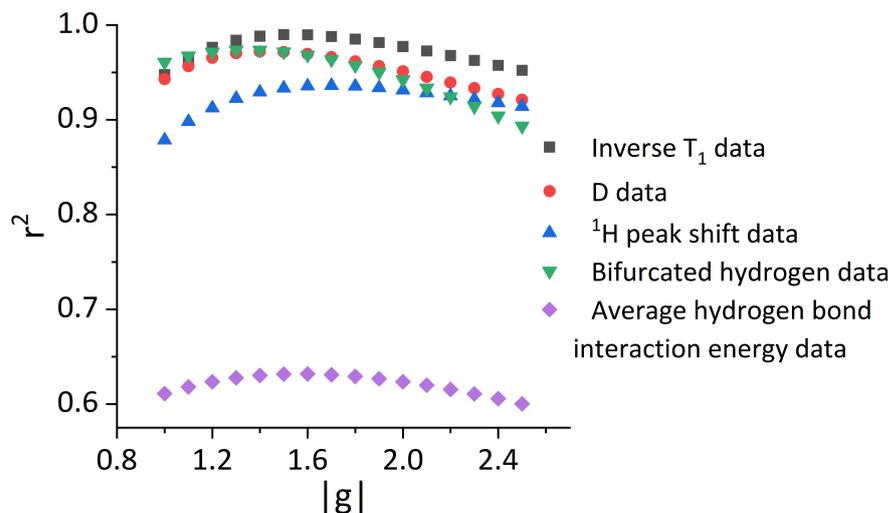


Figure 5.8.2: The r^2 value for the linear fits to structural and dynamic characteristics of the systems plotted using the effective total solute concentration $|g|$.

Data set	$ g $
T_1^{-1}	1.563
D	1.442
NMR peak shift	1.727
Bifurcated hydrogen	1.370
Bulk water - water hydrogen bond energy	1.579
Average value	1.54 ± 0.06

Table 5.8.1: Peak location in r^2 vs $|g|$ data to determine optimum $|g|$ value for each data set.

83,000 water molecule diameters. Conversely the T_1 measurements are measuring dynamic processes on the order of ps and angular jumps of approximately 68° . The overall consistency between techniques demonstrates that TMAO can be thought of as a more powerful perturbing agent of water in general than $\text{Mg}(\text{ClO}_4)_2$ by a factor of 1.54 ± 0.06 . The data sets plotted using this factor are shown in figure 5.8.1.

5.9 Conclusions

The results shown in this section using neutron scattering, EPSR, and NMR demonstrate that while TMAO and $\text{Mg}(\text{ClO}_4)_2$ act to perturb water structure in an opposing manner, they perturb water dynamics in an additive manner. This reduction in dynam-

ics and preservation of water structure in its immediate vicinity means TMAO can be well described as an “anchor” within water in aqueous solutions. We also grouped the results of this chapter together to quantify the relative perturbing ability of TMAO to $\text{Mg}(\text{ClO}_4)_2$. This determined that TMAO is 1.54 ± 0.06 more effective at perturbing water in general than $\text{Mg}(\text{ClO}_4)_2$.

In this chapter we also coupled EPSR to the “extended jump model” of Laage and Hynes [26]. This allowed us to infer information about dynamic perturbations from EPSR data by monitoring the relative abundance of bifurcated hydrogens. However, when employing the extended jump model to EPSR in this manner one must be mindful of three things: the excluded volume effect, the frame reorientation mechanism, and structural heterogeneity. The simple fact that solute species occupy volume means that there is consequentially less volume that water molecules could occupy. This therefore means that water molecules will be less likely to be appropriately positioned to adopt the bifurcated hydrogen conformation, and one might expect the fraction of bifurcated hydrogens to decrease regardless of solute species. If one then takes this as an indication of perturbations to system dynamics one could mistakenly conclude that that the solute is slowing molecular dynamics when this may not necessarily be the case. In order to ensure this is not responsible for the observed effects we can consider the relative size of the TMAO molecule and the Mg^{2+} and ClO_4^- ions. We can estimate this using the the solute - O_w RDFs. If we take the position of the first peak and estimate that the solute occupies a sphere of that radius, we can estimate that the a TMAO molecule has an occupied volume of approximately 387 \AA^3 . $\text{Mg}(\text{ClO}_4)_2$ dissociates into an Mg^{2+} ion (18 \AA^3) and two ClO_4^- ions (256 \AA^3) therefore a single $\text{Mg}(\text{ClO}_4)_2$ molecule occupies 530 \AA^3 . If the observed g parameter was a result of these excluded volumes then one would expect it to be equivalent to the ratio of the TMAO excluded volume to the total $\text{Mg}(\text{ClO}_4)_2$ excluded volume, calculated to be 0.73. As this is not the case, and the weighting parameter derived from EPSR is consistent with the NMR data, this suggests that the bifurcated hydrogen serves as a reliable tool to monitor dynamic perturbations in tertiary solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO.

Monitoring the frame reorientation contribution is unfortunately impossible in EPSR

because EPSR is a Monte Carlo based technique. This means that molecular motions are performed at random, and not as a result of forces experienced due to neighbouring molecules. This means that any two consecutive iterations of the simulation will not represent realistic motions of water molecules, and continuous dynamic processes such as diffusion cannot be observed, only instantaneous intermediates. For this system the inability to monitor frame reorientation is not a significant issue, as the directly proportional relationship between T_1 and D reported in figure 5.6.4 demonstrates that diffusive motions are dominated by individual rotational motions, hence the bifurcated hydrogen abundance serves as a good measure of perturbations to system dynamics. However for other aqueous systems, such as NaCl, where slowed dynamics due to frame reorientation is the dominant dynamic mechanism and even opposes the slight increase in dynamics due to the jump mechanism [157], simply measuring the abundance of bifurcated hydrogens would not be sufficient. Determining this timescale would necessitate using complementary techniques, such as molecular dynamics and bulk viscosity measurements.

Aqueous systems that exhibit heterogeneous environments, such as the aqueous potassium halides discussed in chapter 3, would also not be suitable for dynamic information to be inferred from the abundance of bifurcated hydrogens. Here we observe that water structure is strongly perturbed in the immediate vicinity of the ion, but that this perturbation does not significantly extend into the bulk. Water in the bulk is therefore relatively unaffected by the presence of the potassium halide ions, and so one would expect the frequency of bifurcated hydrogens to be more or less equivalent to that of pure water. Hence while perturbations to system dynamics are clearly observed in the case of aqueous potassium halides, likely due to strong perturbations to water molecule dynamics in the hydration shells of the ions, using the abundance of bifurcated hydrogens to validate these observations would not be suitable.

The results gleaned in this work on specific tertiary solutions also have wider implications to fundamental forces in biology, such as the hydrophobic effect. The hydrophobic effect was classically explained by hydrophobic solutes inducing a highly structured rigid hydration shell of water molecules in their immediate vicinity [9, 53]. This results in

a strong reduction in entropy. If hydrophobic solutes come together in solution they require fewer entropically unfavorable water molecules to build their ordered hydration shell, and hence hydrophobic association occurs spontaneously. However, this highly structured layer of water around hydrophobic solutes is not well observed in previous literature, nor in this research, as shown by the $C_T - O_w$ RDFs presented in figure 5.4.1(b). The origins of the hydrophobic effect may not be solely due to structural perturbations, but also dynamic perturbations around the hydrophobic solutes. It is therefore crucial to consider both aspects, as is done in this work.

Chapter 6

The Study of Biomolecules in Solution Under Extreme Conditions

6.1 Introduction

Finally, the last step of this thesis so as to complete the story promised in the title, is to examine how the structural and dynamic perturbations to water structure in extreme conditions go on to perturb biomolecular self-assembly and stability. To do this we will subject two model biomolecules to the extreme conditions we have suggested are likely to be found on Mars by preparing them in aqueous solution with $\text{Mg}(\text{ClO}_4)_2$. The two biomolecules studied in this manner are the simplest amino acid glycine, whose side chain comprises of a single hydrogen, and the β sheet rich globular protein I27. The rationale behind their selection will be detailed in the following chapter.

Using neutron scattering and EPSR we will monitor the self-association of glycine in aqueous $\text{Mg}(\text{ClO}_4)_2$ so as to better understand how hydrogen bonding between aqueous amino acids may be perturbed by the presence of this salt and speculate how the

stabilising interactions present in folded proteins may be perturbed by $\text{Mg}(\text{ClO}_4)_2$. It also allows us to further speculate how single aqueous amino acids present in solution may be able to self-associate under extreme Martian mimetic conditions to begin to form the precursor molecules for life. We will then build on the observations on self-association of glycine in aqueous $\text{Mg}(\text{ClO}_4)_2$ and study the larger protein I27 in aqueous $\text{Mg}(\text{ClO}_4)_2$. This will be done via two methods: (i) Circular dichroism will be used to monitor perturbations to the secondary structure of the protein induced by $\text{Mg}(\text{ClO}_4)_2$. (ii) Fluorescence spectroscopy will be used to monitor the free energy of unfolding of I27 in the presence of $\text{Mg}(\text{ClO}_4)_2$. Using this technique we will also investigate whether the ability of TMAO to preserve the hydrogen bonding structure of water against the pressure like perturbation of $\text{Mg}(\text{ClO}_4)_2$ is reflected in its ability to preserve protein structure against the pressure like perturbation of $\text{Mg}(\text{ClO}_4)_2$.

Finally we will present preliminary neutron scattering data on a 10 amino acid peptide which adopts a β hairpin conformation in solution, named CLN025. While 10 amino acids is very short in the context of proteins, this is a very large molecule to attempt to model using EPSR. Its size not only means that the number of distinct atomic correlations is very high, but that the number of conformations it is possible for the peptide to adopt is also very high. How many distinct atomic correlations exist within CLN025 can be estimated by considering the molecular dynamics study of CLN025 by McKiernan *et al.*, who used three forcefields to model the peptide [284]. The most successful of these (AMBER FF-15) contained 96 distinct atomic species, however several of these species are described by very similar force fields. By combining highly similar atom types while conserving charge the list can be condensed to 33 distinct atom types. As the number of atomic correlations for J distinct atomic species is equal to $J(J + 1)/2$, this means that even a condensed version of the peptide still results in 561 distinct atomic correlations. The number of conformations the peptide is capable of adopting can be estimated if we assume that the bond between two amino acids can adopt two conformations corresponding to dihedral angles of 0° or 180° (cis or trans). The total number of conformations is therefore 2^{N-1} where N is the number of peptides, hence the peptide can adopt 512 conformations, however if one includes side chain rotations

then this will certainly be much larger. The large size of CLN025 also means that the scattering due to large structures will be primarily at low Q . As low Q translates to large r , and EPSR only calculates correlations up to half the box size, large box sizes are needed to correctly utilise the low Q data. This large parameter space coupled with the need for a large simulation size means that the computing power required to run EPSR is large, and thus iterations proceed very slowly. The nature of this challenge alone means that other co-solutes, such as $\text{Mg}(\text{ClO}_4)_2$ or TMAO were not included in this experiment. Aqueous CLN025 in water with NaOH to help solubilise the protein was studied at three different temperatures. These temperatures were chosen so as to monitor thermal denaturation of the peptide.

6.2 Experimental Considerations - Aqueous Glycine and $\text{Mg}(\text{ClO}_4)_2$

As previously, the structure in aqueous solutions of glycine and $\text{Mg}(\text{ClO}_4)_2$ were determined through neutron scattering, followed by data correction with Gudrun and subsequent analysis through EPSR. To do this experimental samples were prepared of aqueous glycine at 1.8 mol/kg H_2O both with and without $\text{Mg}(\text{ClO}_4)_2$ at 2.7 M (3.3 mol/kg H_2O), corresponding to the concentration employed by Lenton et al [133] and used in the previous chapter. The glycine : water molar ratio was therefore 1 : 30 in aqueous glycine and the glycine : $\text{Mg}(\text{ClO}_4)_2$: water molar ratio was 1 : 1.89 : 31.5 in aqueous glycine with $\text{Mg}(\text{ClO}_4)_2$. Here the solution containing $\text{Mg}(\text{ClO}_4)_2$ have a larger water : amino acid ratio than the pure aqueous amino acid solution. This is again due to the hygroscopic nature of $\text{Mg}(\text{ClO}_4)_2$ introducing extra water into the system, the final amount of which was again estimated using Gudrun as described in section 5.3.

The choice of glycine as a model biomolecule to investigate biomolecular self-association in aqueous $\text{Mg}(\text{ClO}_4)_2$ was made due for several reasons. Firstly, its molecular simplicity means that the computational workload required by EPSR can be minimised and larger box sizes/greater number of iterations can be achieved to improve

the statistical significance of the results. Secondly it has a high solubility (249 mg/mL in water at 25 °C) [285], which means that more concentrated solutions can be made. This is vital as if we wish to reliably observe amino acid - amino acid correlations they must make up a significant weight of the total structure factor and this weight is linearly proportional to concentration, as described in equation 2.2.14. Lastly, as we wish to investigate aqueous solutions likely found on Mars with implications on the potential stability of biological organisms, it is sensible to choose a model biomolecule that has been previously detected in astronomical environments. Glycine fulfils this criteria and has been previously identified in astronomical environments through analysis of comets and meteorites [286–288].

The glycine concentration was chosen so as to be roughly in line with other studies on aqueous amino acids [137, 140, 289, 290]. This ensures that meaningful data can be obtained that will allow us to comment on amino acid - amino acid correlations. Scoppola *et al.* [289] investigated the tripeptide γ - L-glutamyl-L-cysteinylglycine in aqueous solution at a tripeptide : water molar ratio of 1 : 130, which can be thought of as an amino acid : water ratio of 1 : 43.3. Busch *et al.* [137] investigated aqueous proline at molar ratios of 1 : 20 and 1 : 5 proline : water. McLain *et al.* [140, 290] investigated three dipeptides in aqueous solutions at molar ratios of 1 : 20 dipeptide : water (1 : 10 amino acid : water). The molar ratio employed here of 1 : 30 is therefore roughly in line with these previous studies. The study on aqueous glutamine by Rhys *et al.* [138] was performed at at 30 mg/mL, corresponding to a molar ratio of 1 : 270 glutamine : water. This far more dilute solution was limited by the low solubility of the amino acid, however even at this low concentration amino acid - amino acid correlations were observed through neutron scattering and EPSR analysis.

Finally we will make use of the freezing temperature depression induced by $\text{Mg}(\text{ClO}_4)_2$ and investigate amino acid association at both ambient and low temperature. Aqueous glycine was studied at 25 °C and aqueous glycine with $\text{Mg}(\text{ClO}_4)_2$ was studied at 25 °C and -20 °C. These temperatures were chosen with the temperature limitations of potential future experiments on other extraterrestrial environments in mind. Liquid water is also suggested to be present in subsurface lakes on other extraterrestrial

bodies within our solar system. The icy moons of Saturn, Enceladus and Titan, are suggested to have large bodies of subsurface water potentially rich in ammonia [291–293]. The icy moon of Jupiter, Europa, is suggested to have subsurface water with a similar ionic makeup to ocean water on Earth [293, 294], and hence would be dominated by NaCl. The investigated temperatures were therefore chosen so as to be below the bubble point of aqueous ammonia at its eutectic point (23 wt%) [295] but above the freezing temperature of aqueous NaCl at its eutectic point, 252 K at 23.3wt% [269].

6.3 Water Structure in Aqueous Glycine and $\text{Mg}(\text{ClO}_4)_2$

To obtain the relevant site-site RDFs of aqueous glycine with and without $\text{Mg}(\text{ClO}_4)_2$ neutron scattering data was collected using the NIMROD instrument and corrected using Gudrun. Simulated boxes of the two solutions were then created in EPSR that matched the experimental concentrations and densities. In this instance the densities of solution were determined by weighing 1 mL of the solutions under ambient conditions and verified using a densitometer. The calculated densities were $0.0975 \text{ atoms}/\text{\AA}^3$ for binary aqueous glycine and 0.1030 for tertiary aqueous glycine and $\text{Mg}(\text{ClO}_4)_2$. It was also assumed that the density of the tertiary solution was insensitive to temperature due to the fact that the compression of water structure due to $\text{Mg}(\text{ClO}_4)_2$ is a far more dominant mechanism than density variations of water with temperature. For aqueous glycine the simulated box consisted of 5160 water molecules and 172 glycine molecules with a box dimension of 55.0677 \AA . For aqueous glycine and $\text{Mg}(\text{ClO}_4)_2$ the box consisted of 5418 water molecules, 172 glycine molecules, 325 Mg^{2+} cations and 650 ClO_4^- anions with a box dimension of 60.4608 \AA . As the pH of solution is relatively unaffected by the addition of either solute species, glycine was modelled in its zwitterionic form as it occurs at neutral pH [296]. This means that the amine group becomes protonated and has an overall positive charge (NH_2 to NH_3^+) and the carbonyl group becomes deprotonated and has an overall negative charge (COOH to COO^-).

The reference potential parameters for EPSR for water and $\text{Mg}(\text{ClO}_4)_2$ were set to agree with previous neutron diffraction experiments [133], however at the time of the

System	First peak (Å)	Second peak (Å)
Glycine at 25 °C	2.78	4.42
Glycine + Mg(ClO ₄) ₂ at 25 °C	2.69	5.50
Glycine + Mg(ClO ₄) ₂ at -20 °C	2.72	5.26

Table 6.3.1: Peak positions of the first and second peaks in the $O_w - O_w$ RDFs.

experiment no previously published parameters from a neutron diffraction and EPSR experiment were available for glycine. Therefore the bond angles, lengths, and reference potential parameters were estimated using previously published literature on similar amino acids/peptides [137, 138, 140, 289, 290, 297–299]. The final values for the reference potentials, as well as the final EPSR fits to the corrected scattering data, can be found in section D. The simulations were fit with an R factor between 0.0004 - 0.0006, and hence the simulated fits are in excellent agreement with the scattering data.

As in previous chapters, we will begin by discussing the perturbations to bulk water structure induced by the two solute species. These are best visualised through the $O_w - O_w$ RDFs, shown in figure 6.3.1. The introduction of glycine at 1.8 mol/kg H₂O results in a modest perturbation to water structure, with the first peak remaining at 2.78 Å, and the second peak moving inwards from 4.45 Å to 4.42Å. Glycine is therefore a negligible perturber of water structure. Consistent with the results observed for previous solutions containing 2.7 M Mg(ClO₄)₂, the first peak moves inwards to 2.69 Å and 2.72 Å at 25 °C and -20 °C respectively. The second peak again compresses completely into the first peak and the third peak becomes the new second peak, which is located at 5.50 Å and 5.26 Å at 25 and - 20 °C respectively. These peak positions are summarised in table 6.3.1. The corresponding SDFs which show the compression of the second hydration shell into the first hydration shell in the presence of 2.7 M Mg(ClO₄)₂ are shown in figure 6.3.2.

6.4 Hydration of Glycine in Aqueous Mg(ClO₄)₂

As this study is done with the aim of understanding biomolecule hydration and self-association in extreme Martian mimetic conditions, it is now of vital importance that we study the hydration around the glycine molecule. In this section we will discuss hydration

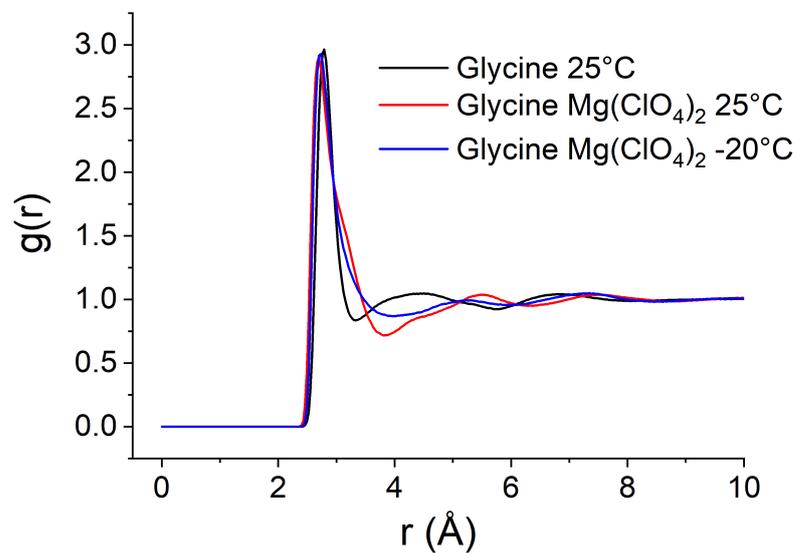


Figure 6.3.1: $O_w - O_w$ RDFs for the three sample sets studied in this section as calculated through EPSR.

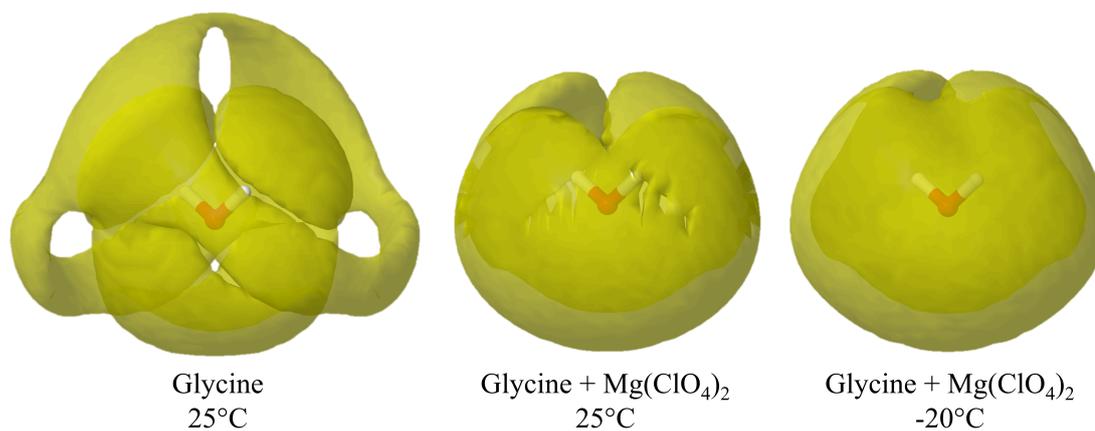


Figure 6.3.2: Water - water SDFs for aqueous glycine at 25 °C, aqueous glycine and $Mg(ClO_4)_2$ at 25 °C, and aqueous glycine and $Mg(ClO_4)_2$ at -20 °. These surface contours contain the 30% probable locations of a neighbouring water molecule within 5 Å.

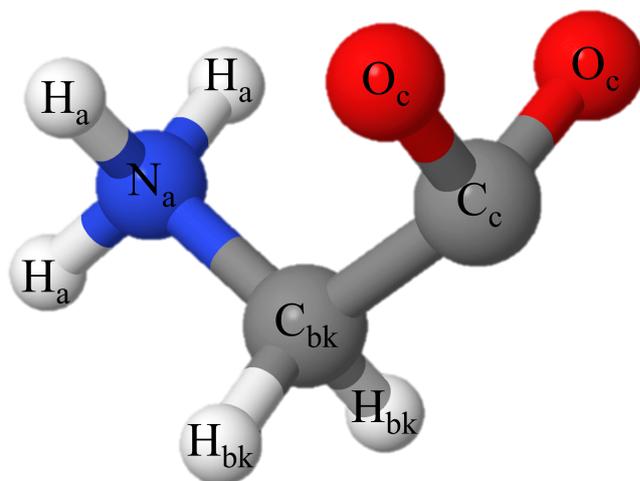


Figure 6.4.1: The structure of glycine in its zwitterionic form as modelled in this research with atom names displayed.

around three key areas of the glycine molecule: the end amine and carbonyl groups, which take their zwitterionic form and are hence hydrophilic, and the hydrophobic backbone hydrogens. The structure of glycine as modelled in this research and the atom naming conventions are shown in figure 6.4.1.

We will begin by considering the hydration around the positively charged amine group. This is done by examining the $N_a - O_w$ and $N_a - H_w$ RDFs, as shown in figure 6.4.2. Here we observe strong coordination of water molecules around the amine group as indicated by a clear first peak followed by a series of smaller peaks in the $N_a - O_w$ RDF. For aqueous glycine at 25 °C the peak occurs at 2.76 Å. This is coupled with a relatively strong single first peak in the $N_a - H_w$ RDF located at 3.39 Å, which suggests that the hydrating water molecules orient both hydrogens away from the amine group. To understand the implications of these two calculated distances let us consider an example water molecule hydrating a central atom species that is donating a strong hydrogen bond to the water molecule. In this case the water molecule will have a strong tendency to orient both hydrogens away from the central atom species. Using trigonometry it can be demonstrated that in the case of this strong hydrogen bond, where both hydrogens are oriented away from the central atom species at the greatest possible distance and both hydrogens are equidistant from the central species, for an atom - O_w distance of

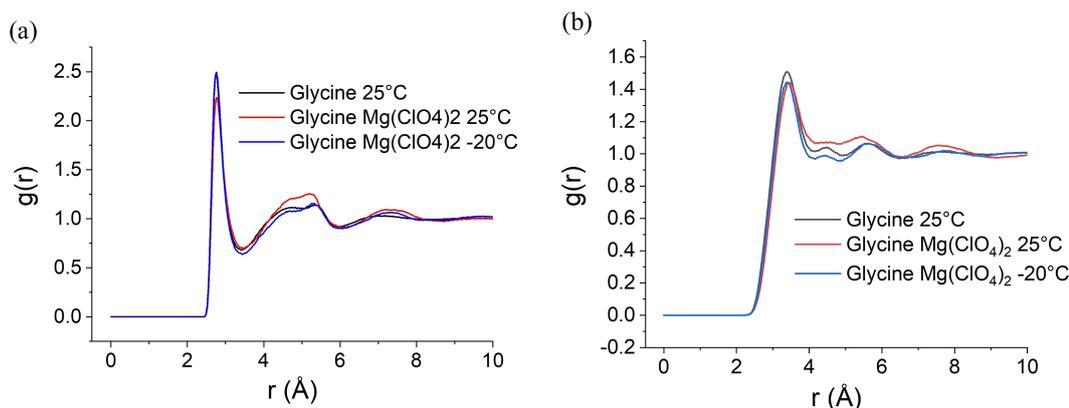


Figure 6.4.2: $N_a - O_w$ (a) and $N_a - H_w$ (b) RDFs for the three sample sets studied in this section as calculated through EPSR.

2.76 Å the atom - H_w distance should be 3.43 Å. Here the peak $N_a - H_w$ distance is calculated to be 3.39 Å, which is remarkably close to the idealised 3.43 Å. This suggests that the amine group is donating strong hydrogen bonds to surrounding water molecules via the three positively charged amine hydrogens.

Upon the addition of 2.7 M $Mg(ClO_4)_2$ at 25 °C we observe that the peak diminishes in intensity and shifts slightly outwards to 2.77 Å. We also observe that the intensity of the second peak increases. We observe similar perturbations in the $N_a - H_w$, where the first peak diminishes in intensity, moves outwards to 3.46 Å, and intensity is increased in the area immediately after the first peak (between approximately 4 - 6 Å). We therefore see that similar to the hydration around the hydrophilic oxygen of TMAO described in chapter 5, the hydration around the hydrophilic group of glycine is far less sensitive to $Mg(ClO_4)_2$ addition than the bulk water structure. This again suggests the tendency of the positively charged amine groups to form strong hydrogen bonds with neighbouring water molecules.

Upon decreasing temperature in tertiary aqueous glycine and $Mg(ClO_4)_2$ from 25 to -20 °C we see that the water structure around the glycine amine group returns to being more similar to the structure of water around the glycine amine group at 25 °C in the absence of $Mg(ClO_4)_2$. The first peak in the $N_a - O_w$ RDF recovers its intensity and moves back inwards to 2.75 Å, and the second peak reduces in height. The first

System	$N_a - O_w$ peak (Å)	$N_a - H_w$ peak (Å)
Glycine at 25 °C	2.76	3.39
Glycine + $Mg(ClO_4)_2$ at 25 °C	2.77	3.46
Glycine + $Mg(ClO_4)_2$ at -20 °C	2.75	3.40

Table 6.4.1: Peak positions of the first peaks in the $N_a - O_w$ and $N_a - H_w$ RDFs.

peak in the $N_a - H_w$ RDF does not recover its intensity, but does shift back inwards to 3.40 Å. The proposed hydrogen bond between the glycine amine group and surrounding water molecules is therefore likely perturbed by the addition of $Mg(ClO_4)_2$, as seen in TMAO - water hydrogen bonding in chapter 5, but this perturbation is reduced at lower temperatures. An enhancement of hydrogen bonding at lower temperatures is consistent with previous studies on pure liquid water [75, 102, 300] and alcohol - water mixtures [143, 301, 302], and therefore allows us to confidently suggest that the glycine amine group and surrounding water molecules form strong hydrogen bonds through the amine hydrogens. The first peak positions of these two sets of RDFs are summarised in table 6.4.1.

Next we can go on to consider hydration around the other hydrophilic region of glycine, the negatively charged carbonyl group, shown via the $C_c - O_w$ and $C_c - H_w$ RDFs in figure 6.4.3. In aqueous glycine at 25 °C we observe association around the carbonyl group, shown by the presence of a single clear peak at 3.53 Å. We also observe a tendency of the hydrating water molecules to orient a hydrogen towards the carbonyl group, as evidenced by the presence of two clear peaks in the $C_c - H_w$ RDF located either side of the position of the first peak in the $C_c - O_w$ RDF. These peaks are located at 2.65 and 4.03 Å. The tendency of the hydrating water molecule to orient a hydrogen towards the carbonyl group can be shown by the difference in position between the first peak in the two RDFs. This is equal to 0.88 Å, which is close to the OH bond length of the water molecule of 0.96 Å. This shows that this tendency is therefore quite strong and corresponds to fairly linear hydrogen bonds donated by the hydrating water molecules to the carbonyl oxygens.

As around the amine group, this hydrogen bond is disrupted by the addition of 2.7 M $Mg(ClO_4)_2$ at 25 °C, resulting in a slight increase in peak height in the $C_c - O_w$ RDF and

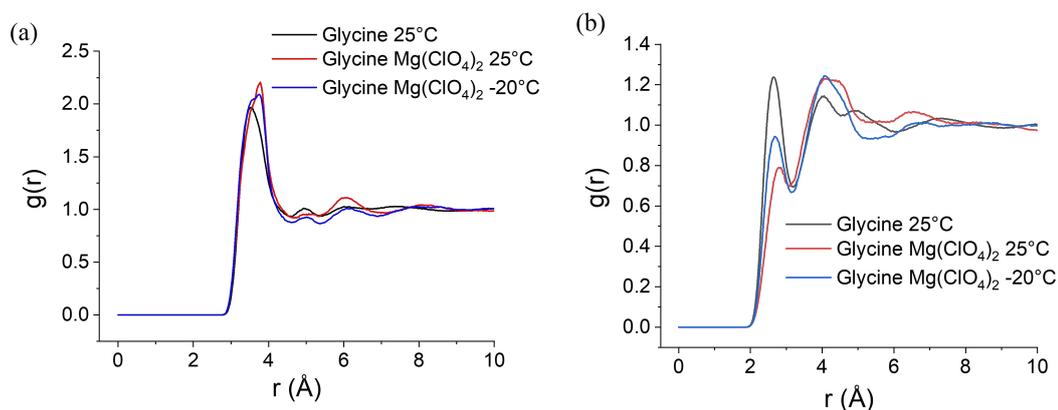


Figure 6.4.3: $C_c - O_w$ (a) and $C_c - H_w$ (b) RDFs for the three sample sets studied in this section as calculated through EPSR.

a movement outwards to 3.76 Å. The hydration structure is therefore also more stable against the addition of Mg(ClO₄)₂ than the bulk water structure, as was the case for the amine group. The disruption to hydrogen bonding is more clearly evidenced by a stark reduction in the first peak height of the $C_c - H_w$ RDF, coupled with a slight outward movement to 2.81 Å, suggesting a substantially decreased tendency of hydration water to orient a hydrogen towards the carbonyl group and a therefore weakened hydrogen bond. Again as with water hydrating the amine group, the water - carbonyl hydrogen bonding is apparently enhanced as temperature is reduced from 25 to -20 °C. This is demonstrated by a reduction of the peak height in the $C_c - O_w$ RDF back towards that of the peak height observed in aqueous glycine at 25 °C, and an inward movement of the peak to 3.64 Å. This is also clear in the $C_c - H_w$ RDF, where we observe an increase in the height of the first peak with reducing temperature and a movement back inwards to 2.70 Å, suggesting an increased tendency of the hydrating water molecule to orient a hydrogen towards the carbonyl group. The two charged end groups of the glycine molecule, the amine and carbonyl groups, can therefore be thought of as strong hydrogen bond donors and acceptors respectively, whose hydrogen bonding ability is reduced by the addition of Mg(ClO₄)₂, but much less so than bulk water. The peak shifts shown in this RDF are summarised in table 6.4.2.

We can visualise this hydrophilic hydration and hydrogen bond perturbations through the glycine - water SDFs around a central amine group or a central carbonyl group.

System	$C_c - O_w$ peak (Å)	$C_c - H_w$ peak (Å)
Glycine at 25 °C	3.53	2.65
Glycine + $Mg(ClO_4)_2$ at 25 °C	3.76	2.81
Glycine + $Mg(ClO_4)_2$ at -20 °C	3.64	2.70

Table 6.4.2: Peak positions of the first peaks in the $C_c - O_w$ and $C_c - H_w$ RDFs.

These SDFs are shown in figure 6.4.4. Here we observe that in aqueous glycine that water molecules preferentially locate themselves directly above each NH bond in the amine group, resulting in three areas of preferential hydration. Upon addition of 2.7 $Mg(ClO_4)_2$ at 25 °C these areas become less well defined as a result of perturbed hydrogen bonding between the amine group and the water molecules, and as hydrogen bonding is enhanced by reducing temperature the areas become more defined again. For water hydrating the carbonyl group we observe that water also preferentially hydrates around three areas: two areas either side of the carbonyl group and one above. Upon addition of 2.7 $Mg(ClO_4)_2$ at 25 °C we observe that the area above the carbonyl group reduces in size, and the two areas either side of the carbonyl group become more localised to be directly above the CO bonds. As hydrogen bonding is restored by reducing temperature the area above the carbonyl group becomes larger again, and the two areas directly above the CO bonds move back towards the position they adopt in aqueous glycine at 25 °.

Finally we can turn our attention to hydration around the hydrophobic area of the glycine molecule, the central backbone hydrogens, by considering the $C_{bk} - O_w$ and $C_{bk} - H_w$ RDFs, as shown in figure 6.4.5. As the dipole moment across the $C_{bk}H_{bk}$ bond is far weaker than the dipole moment across either the N_aH_a or C_cO_c bond, it is expected that electrostatic interactions between water and the backbone hydrogens will be far weaker than between water and the hydrophilic regions. The observed RDFs are consistent with this expectation, as here we observe that water is coordinated, but that this coordination is rather ill defined in comparison to hydration around the hydrophilic groups. This is demonstrated by the presence of two small overlapping peaks in the $C_{bk} - O_w$ RDFs, rather than a single clear peak as was observed around the hydrophilic areas. The first peak in the $C_{bk} - O_w$ RDF is located at 3.57 Å, which shifts outwards to 3.62 Å upon addition of 2.7 $Mg(ClO_4)_2$, and then back inwards to 3.56 Å upon reducing temperature

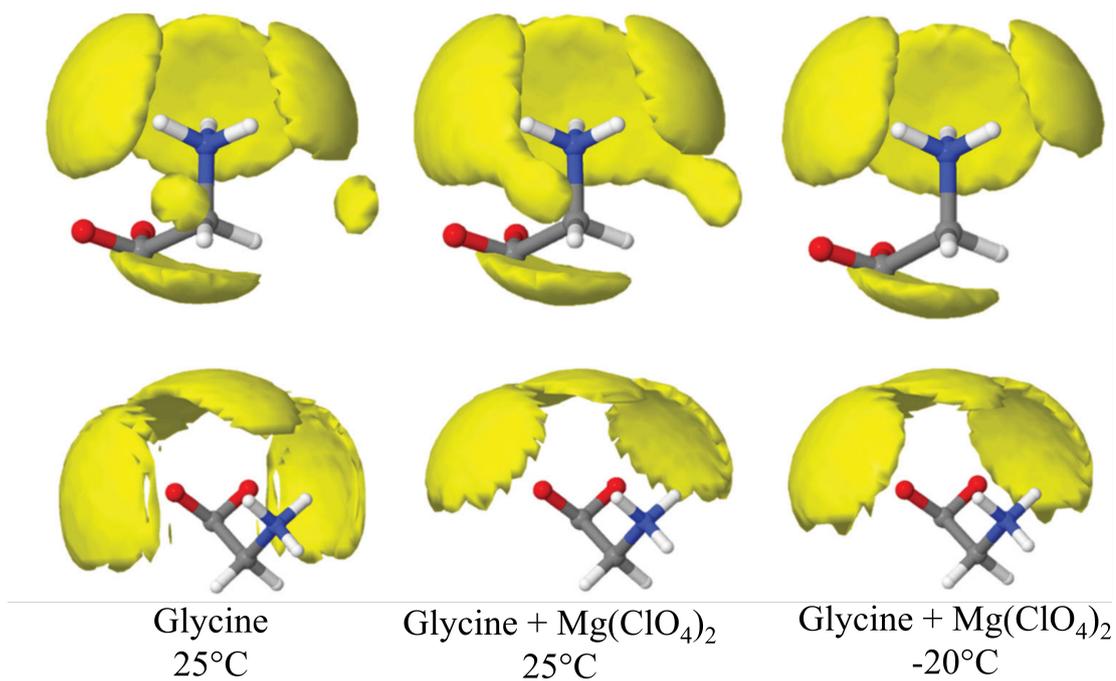


Figure 6.4.4: Spatial density functions of water around a central amine group (top) or a central carbonyl group (bottom) of aqueous glycine at 25 °C (left), aqueous glycine and Mg(ClO₄)₂ at 25 °C (middle), and aqueous glycine and Mg(ClO₄)₂ at -20 °C (right) from neutron diffraction data and EPSR analysis. These surface contours contain the highest 15% probability areas of finding a water molecule within a distance of 5 Å from the central atom.

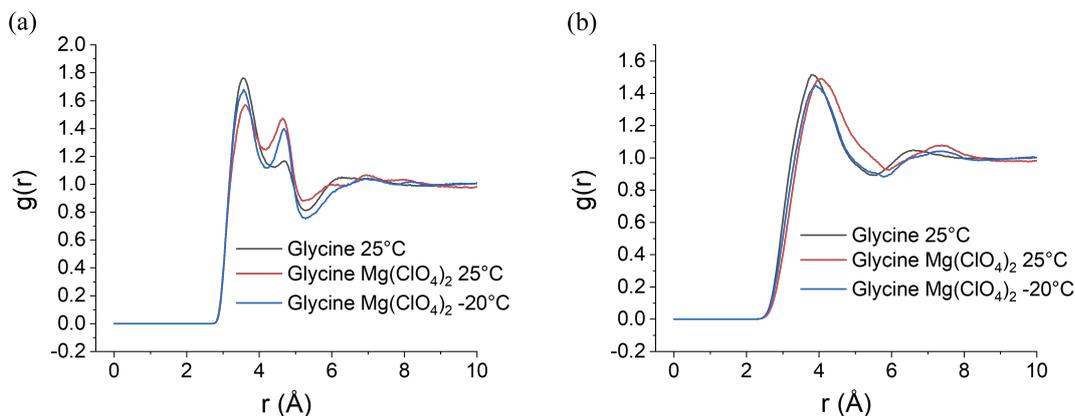


Figure 6.4.5: $C_{bk} - O_w$ (a) and $C_{bk} - H_w$ (b) RDFs for the three sample sets studied in this section as calculated through EPSR.

System	$C_{bk} - O_w$ peak (Å)	$C_{bk} - H_w$ peak (Å)
Glycine at 25 °C	3.57	3.85
Glycine + $Mg(ClO_4)_2$ at 25 °C	3.62	4.05
Glycine + $Mg(ClO_4)_2$ at -20 °C	3.56	3.93

Table 6.4.3: Peak positions of the first peaks in the $C_{bk} - O_w$ and $C_{bk} - H_w$ RDFs.

from 25 to -20 °C, in a similar manner to water hydrating the hydrophilic groups.

The reduced electrostatic interactions between water and the backbone hydrogens are more clearly evidenced by considering the $C_{bk} - H_w$ RDFs. Here we observe a single, but rather broad, peak in aqueous glycine at 25 ° located at 3.85 Å. The difference between the first peak locations in the $C_{bk} - O_w$ and $C_{bk} - H_w$ RDF is therefore 0.28, and therefore the hydrating water molecules show only a slight tendency to orient their hydrogens away from the backbone hydrogens. Upon addition of $Mg(ClO_4)_2$ the first peak in the $C_{bk} - H_w$ RDF moves slightly outwards to 4.05 Å, and moves back inwards to 3.93 Å upon cooling. This recovery of both the $C_{bk} - O_w$ and $C_{bk} - H_w$ upon cooling, as stated previously, is characteristic of hydrogen bonding between the backbone hydrogens and the hydrating water. However, the fact that the tendency for hydrating water molecules to orient their hydrogens away from the backbone hydrogens is relatively weak, suggests that these hydrogen bonds if present are far weaker than those between the hydrophilic areas and hydrating water molecules. The first peak positions in the two sets of RDFs are summarised in table 6.4.3.

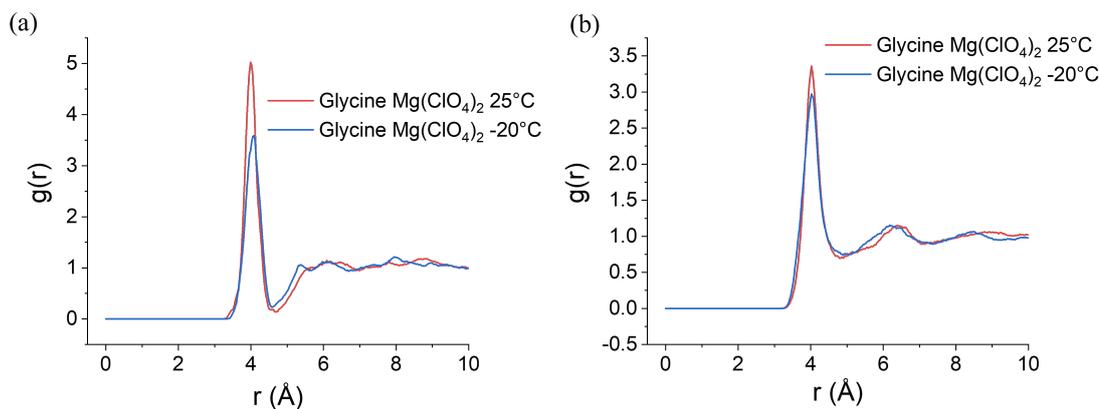


Figure 6.5.1: $N_a - Mg^{2+}$ (a) and $N_a - Cl_p$ (b) RDFs for aqueous glycine and $Mg(ClO_4)_2$ at 25 and -20 °C as calculated through EPSR.

6.5 Glycine - $Mg(ClO_4)_2$ Interactions

As we did in section 5.4 for TMAO - ion interactions, we will now consider interactions between glycine and Mg^{2+} or ClO_4^- ions in solution. We will first consider the amine - ion interactions through the $N_a - Mg^{2+}$ and $N_a - Cl_p$ RDFs shown in figure 6.5.1. Here we observe that the amine group, despite unfavourable electrostatic interactions between the positively charged amine group and the positively charged Mg^{2+} ion, strongly coordinates Mg^{2+} ions as shown by a single clear peak located at 4.00 and 4.06 Å at 25 and -20 °C respectively. This result is surprising, however if we consider the fact that the first peak in the $N_a - O_w$ RDF occurs at 2.77 and 2.75 Å at 25 and -20 °C respectively, it becomes clear that while Mg^{2+} ions coordinate around the amine group, this happens through a mediating water layer which screens the unfavourable electrostatic interactions. As the temperature is reduced, and amine - water hydrogen bonding is enhanced, the Mg^{2+} ions are further excluded from the amine group, resulting in the outward peak shift and reduction in intensity.

The amine group also coordinates ClO_4^- ions, as shown by a single clear peak in the $N_a - Cl_p$ RDF located at 4.02 and 4.03 Å at 25 and -20 °C respectively. While this also occurs at much larger distances than the $N_a - O_w$ RDF first peak, the ClO_4^- ion is polyatomic, and therefore much larger than the Mg^{2+} ion (the ClO bond distance in the ClO_4^- ion is equal to 1.42 Å). When this factor is taken into account, and one considers

System	$N_a - Mg^{2+}$ peak (\AA)	$N_a - Cl_p$ peak (\AA)
Glycine + $Mg(ClO_4)_2$ at 25 °C	4.00	4.02
Glycine + $Mg(ClO_4)_2$ at -20 °C	4.06	4.03

Table 6.5.1: Peak positions of the first peaks in the $N_a - Mg^{2+}$ and $N_a - Cl_p$ RDFs.

that the electrostatic interactions between the positively charged amine group and the negatively charged ClO_4^- ion will be favourable, it becomes likely that this interaction is direct and does not require a layer of mediating water to screen unfavourable electrostatic interactions. However, as was the case with Mg^{2+} association around the amine group, decreasing temperature results in a reduction in peak intensity and an outward shift. This is again likely due to enhanced water - amine hydrogen bonding resulting in ClO_4^- ions being excluded from the amine group. The ion - amine RDF peak positions are summarised in table 6.5.1.

Next we will consider ion association around the negatively charged carbonyl group through the $C_c - Mg^{2+}$ and $C_c - Cl_p$ RDFs shown in figure 6.5.2. Here we observe, unsurprisingly, that Mg^{2+} ions coordinate strongly around the carbonyl group due to favourable electrostatic interactions, as shown by a single clear peak at 2.67 \AA at both temperatures. This peak is also of much greater intensity than the first peak in the $N_a - Mg^{2+}$ peak due to this strong direct interaction. By comparison the ClO_4^- ion association around the carbonyl group is far weaker, showing a broad and poorly defined first peak at approximately 4.8 \AA . This is likely due to unfavourable electrostatic interactions between the carbonyl group and the ClO_4^- ion. The association of both ion species around the negatively charged carbonyl group is decreased as temperature is decreased, as shown by a reduction in intensity of the first peak of both ion RDFs. This is again likely due to enhanced hydrogen bonding between the carbonyl group and hydrating water molecules resulting in greater exclusion of ions from the carbonyl group.

Finally, we will consider ion association around the hydrophobic backbone hydrogens by studying the $C_{bk} - Mg^{2+}$ and $C_{bk} - Cl_p$ RDFs as shown in figure 6.5.3. Here we observe that Mg^{2+} ions again coordinate strongly around the backbone hydrogens, as evidenced by the single clear peak at 3.75 \AA at both temperatures. As was the case in chapter 5, the strength of Mg^{2+} association reduces as the electrostatic interactions between the ion

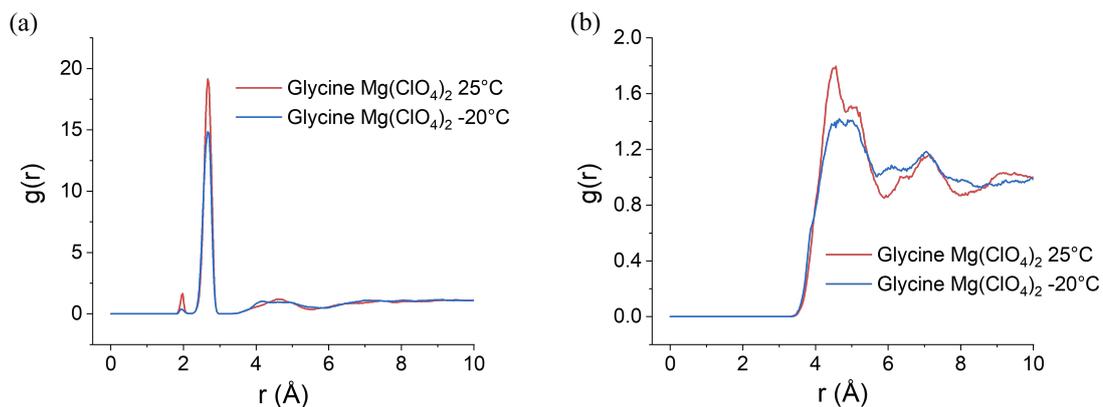


Figure 6.5.2: $C_c - \text{Mg}^{2+}$ (a) and $C_c - \text{Cl}_p$ (b) RDFs for aqueous glycine and $\text{Mg}(\text{ClO}_4)_2$ at 25 and -20 °C as calculated through EPSR.

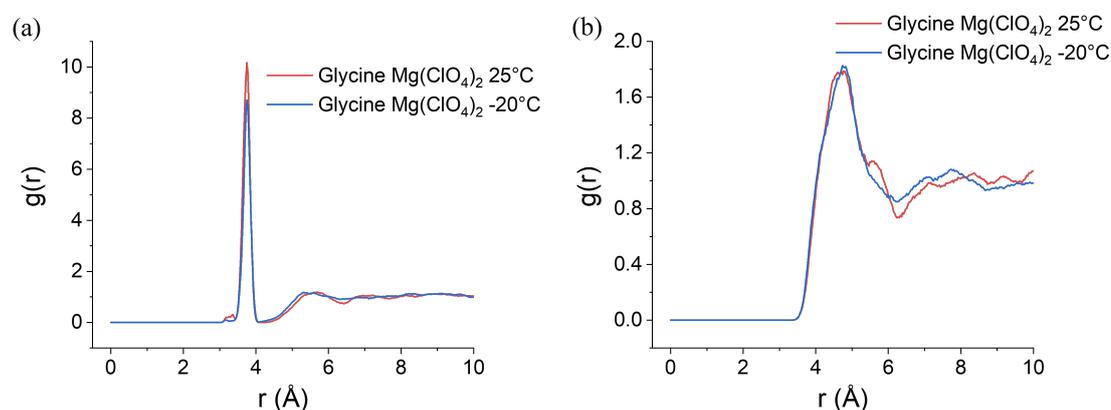


Figure 6.5.3: $C_{bk} - \text{Mg}^{2+}$ (a) and $C_{bk} - \text{Cl}_p$ (b) RDFs for aqueous glycine and $\text{Mg}(\text{ClO}_4)_2$ at 25 and -20 °C as calculated through EPSR.

and the central group become less favourable, as evidenced by the relative peak heights in the central group - Mg^{2+} RDFs ($\text{Mg}^{2+} - C_c$ peak $>$ $\text{Mg}^{2+} - C_{bk}$ peak $>$ $\text{Mg}^{2+} - N_a$ peak). This peak occurs at slightly larger distances than the $C_{bk} - O_w$ peak, and is therefore more likely to be due to favourable electrostatic interactions between the Mg^{2+} and the hydrating water molecules, rather than interactions with the backbone hydrogens itself. The same is likely true for the association of ClO_4^- ions around the backbone hydrogens, however these two species will also likely associate through hydrophobic style interactions due to the low charge density of the ClO_4^- ion, as discussed in section 5.4. This is shown by a rather poorly defined first peak in the $C_{bk} - \text{Cl}_p$ RDF at approximately 4.7 Å.

Interestingly, the change in peak intensity with decreasing temperature allows us

to comment on the strength of water - backbone hydrogen bonding. Here it can be observed that the relative reduction in peak height upon decreasing temperature in both the $C_{bk} - Mg^{2+}$ and $C_{bk} - Cl_p$ RDFs is less than the amine - ion or carbonyl - ion RDF peak reductions. The reduction in peak height in both the amine - ion and carbonyl - ion RDFs was explained by enhanced hydrogen bonding at lower temperatures between the hydrophilic groups on the glycine molecule and the surrounding water molecules, resulting in greater exclusion of the ions. Around the backbone hydrogens the reduction in peak height with decreasing temperature is less evident, hence the mechanism by which ions are being excluded from the backbone hydrogens with decreasing temperature must be less pronounced. It therefore follows that hydrogen bonding between the backbone hydrogens and surrounding water is much weaker, consistent with the results observed in section 6.4.

6.6 Glycine Self-Association in Aqueous $Mg(ClO_4)_2$

We will now finally turn our attention to the self-association of aqueous glycine in the presence of $Mg(ClO_4)_2$. To do this will concern ourselves with two potential mechanisms of glycine self-association: interactions between the oppositely charged hydrophilic regions, and hydrophobic style associations between the backbone hydrogens. These are best examined through the $H_a - O_c$ RDFs (hydrophilic) and the $H_{bk} - H_{bk}$ RDFs (hydrophobic), as shown in figure 6.6.1. From these RDFs the preferred method of glycine self-association is clear.

The $H_a - O_c$ RDFs exhibit two clear peaks as a result of favourable interactions between the oppositely charged hydrophilic groups. As we have established that these groups are capable of donating and accepting hydrogen bonds from neighbouring water molecules, it is likely that these favourable interactions between glycine molecules are hydrogen bonds donated from the amine group of one molecule and accepted by the carbonyl group of another. In aqueous glycine at 25 °C the two peaks occur at 1.65 and 3.01 Å. Upon addition of 2.7 M $Mg(ClO_4)_2$ the first peak is unmoved and the second peak shifts slightly outwards to 3.04 Å. Upon cooling from 25 to -20 °C both

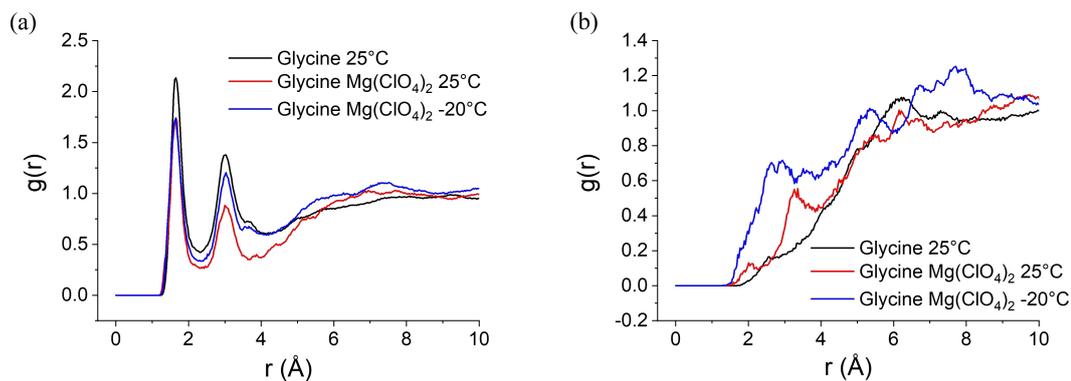


Figure 6.6.1: $H_\alpha - O_c$ (a) and $H_{bk} - H_{bk}$ (b) RDFs for the three sample sets studied in this section as calculated through EPSR.

peaks then move slightly inwards to 1.63 and 3.03 Å. Overall it can be stated that these peak positions are relatively insensitive to $Mg(ClO_4)_2$ addition or temperature, however the relative peak intensities are more telling. Upon addition of $Mg(ClO_4)_2$ the intensity of both peaks sharply reduces, indicative of reduced hydrogen bonding between glycine molecules. Consistent with the results observed in section 6.4, the intensity of the peaks then recover with decreasing temperature as hydrogen bonding is restored. Conversely, hydrophobic style association between the hydrophobic backbone hydrogens is not observed, demonstrated by the relatively featureless $H_{bk} - H_{bk}$ RDFs in all three samples.

Hydrophilic style association is therefore the dominant mechanism by which glycine molecules associate in solution, consistent with previous neutron scattering and EPSR based literature on other amino acids in solution [137–140, 289, 290]. A study on aqueous proline by Busch *et al.* [137] demonstrated this very effectively, and showed that aqueous proline typically forms dimers in solution as a result of direct interactions between its CO_2^- and NH_2^+ groups. In a similar manner, we can now take the observation of hydrophilic association between glycine molecules a step further and begin to quantify the clustering of glycine molecules in aqueous solution. This is done using the CLUSTERS routine within EPSR following the procedure described in section 2.2.5. Here two glycine molecules are deemed to be part of the same cluster if a H_α of one glycine molecule lies within a distance of the O_c of another glycine molecule corresponding to

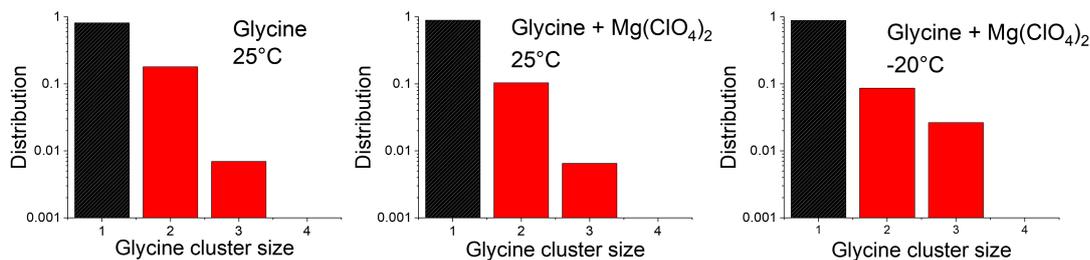


Figure 6.6.2: Cluster size distribution predicted from EPSR using a $H_a - O_c$ cutoff distance of 2.32 Å. Distribution plotted on a logarithmic scale for clarity. Proportion of clusters containing only one glycine molecule shown in black.

the first minimum in the $H_a - O_c$ RDF, established here to be equal to 2.32 Å in all datasets. This allows the proportion of molecules which belong to a cluster size n to be determined. The resulting distributions are shown in figure 6.6.2.

These distributions allow us to quantify cluster formation between glycine molecules. These results demonstrate that zwitterionic glycine will can form dimers and even trimers in solution in all cases, however it is more likely to find a glycine molecule as a monomer (not belonging to a cluster) than it is to find it belonging to a cluster. In aqueous glycine at 25 °C it is determined through EPSR that the proportion of glycine molecules that can be found in clusters containing two or three glycine molecules is equal to 0.320. Upon the addition of 2.7 M $Mg(ClO_4)_2$ this charged based hydrophilic association is reduced, likely due to screening by the charged ions. The proportion of clustered glycine molecules then reduces to 0.203. However, as the system is then cooled to -20 °C and hydrogen bonding is enhanced, this proportion recovers slightly to 0.221. This is consistent with observations of temperature dependent association of alcohols in aqueous solutions [143, 301, 302]. The most important implication of this finding is that while $Mg(ClO_4)_2$ is damaging to hydrogen bonding in aqueous solutions, as described in this chapter and chapter 5, it is only capable of reducing hydrogen bonding between biomolecules, not eliminating it entirely.

In this study we chose to investigate aqueous glycine for the reasons discussed in section 6.2. The simplicity of the glycine molecule suggests that the results of this experiment should be consistent for all zwitterionic amino acids, and amino acids should

be able to form intermolecular hydrogen bonds, albeit to a lesser extent, in concentrated $\text{Mg}(\text{ClO}_4)_2$. The fact that we observe charge based interactions as the dominant mechanism for glycine self-association also suggests that hydrogen bonding between amino acids which feature a charged side chain may be more resistant to the detrimental effects of $\text{Mg}(\text{ClO}_4)_2$. This is because the presence of charged side chains will provide more sites for ion - amino acid interactions and allow the remaining hydrophilic sites to continue to form hydrogen bonds. It is certainly true that halophilic organisms, which thrive in environments with high salt concentrations, tend to have a higher proportion of surface accessible amino acids with charged side chains in their proteins [16, 17]. These likely screen the effects of ions present in their aqueous environment and allow intraprotein hydrogen bonding to remain intact.

6.7 The Model Protein I27 in Aqueous $\text{Mg}(\text{ClO}_4)_2$ and TMAO

With a firm understanding of the effects of $\text{Mg}(\text{ClO}_4)_2$ on individual amino acid association, we can now turn our attention to the effects of $\text{Mg}(\text{ClO}_4)_2$ on complete protein stability. For this we choose to study the 27th immunoglobulin domain from the human muscle protein titin, I27 (PDB code: 1TIT). The backbone structure of this protein along with its buried tryptophan residue is shown in figure 2.7.2 and reproduced for convenience in figure 6.7.1. As described in section 2.5, I27 is well suited for the work covered in this thesis because it is well suited to expression and purification with standard protocols, but more importantly its structure is not stabilised through chemical bonds such as salt bridges or disulfide bonds. It is stabilised through intraprotein hydrogen bonds and the hydrophobic collapse of its residues with hydrophobic side chains [232]. This means that the results gained through the study of aqueous solutions of $\text{Mg}(\text{ClO}_4)_2$, TMAO, and glycine can, in theory, be applied to the study of this protein.

As mentioned in chapter 2, proteins in this thesis are studied through two experimental methods: circular dichroism (CD) and fluorescence spectroscopy. CD is sensitive

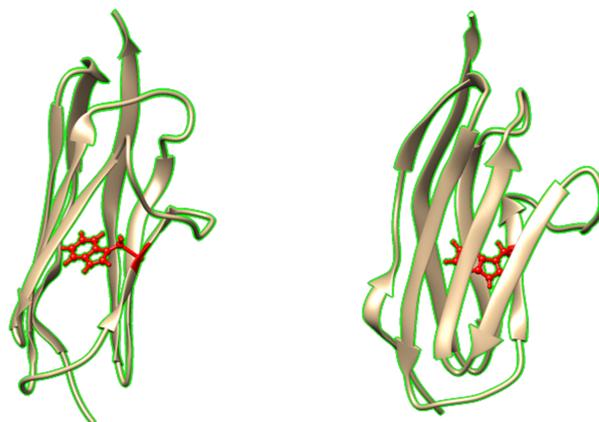


Figure 6.7.1: The cartoon structure showing the secondary β structure of I27 with the tryptophan atomic side chain structure highlighted in red (PDB code: 1TIT).

to the chirality of the sample, and is therefore capable of assessing the secondary structure components of proteins [238]. Fluorescence spectroscopy is sensitive to the chemical environment of particular amino acid side chains, in this case the buried tryptophan in I27. The fluorescence emission spectra of tryptophan changes as the protein denatures and the hydrophobic core becomes hydrated, hence fluorescence spectroscopy is sensitive to the tertiary structure of the protein [303]. We will first use these two methods to determine the effect of $\text{Mg}(\text{ClO}_4)_2$ on I27 in aqueous solution. To achieve this I27, produced and purified as described in section 2.5, was prepared to either 0.2 mg/mL for CD studies or 0.05 mg/mL for fluorescence studies in 25 mM HEPES buffer at pH of 7.4. $\text{Mg}(\text{ClO}_4)_2$ was also included as a cosolute at concentrations of 0 - 1 M in increments of 0.1 M for CD studies and 0.03 M for fluorescence studies. The resultant CD spectra from 195-260 nm and emission counts at 320 nm following excitation at 280 nm, normalised to the value at 0 M $\text{Mg}(\text{ClO}_4)_2$, are shown in figure 6.7.2.

The results of these two techniques allow us to demonstrate that $\text{Mg}(\text{ClO}_4)_2$ is having a strong effect on the structure of I27. Firstly, as shown in figure 6.7.2(a), the strong change in CD spectra as $\text{Mg}(\text{ClO}_4)_2$ concentration is increased from 0 to 1 M indicates that the secondary structure is strongly perturbed. This perturbation in secondary structure is accompanied by an unfolding of the protein. This is shown by a reduction in

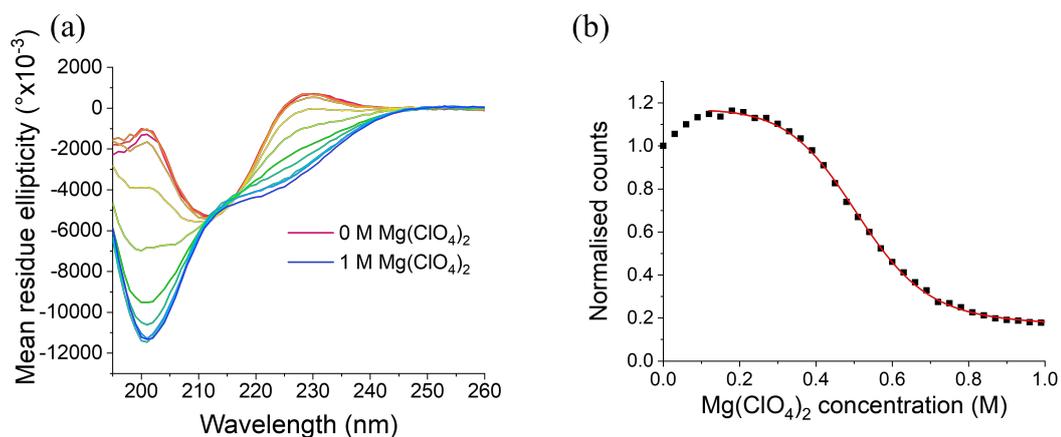


Figure 6.7.2: The CD spectra of I27 from 195 - 260 nm in the presence of 0 - 1 M $\text{Mg}(\text{ClO}_4)_2$ in increments of 0.1 M (a) and the normalised fluorescence emission counts of I27 in the presence of 0 - 1 M $\text{Mg}(\text{ClO}_4)_2$ in increments of 0.03 M.

the normalised emission counts at 320 nm with increasing $\text{Mg}(\text{ClO}_4)_2$ concentration in figure 6.7.2(b). The observation that the fluorescence counts tend towards a plateau, and that the CD spectra stop changing, at $\text{Mg}(\text{ClO}_4)_2$ concentrations above 0.8 M suggests that above this concentration the protein is completely unfolded.

This in turn suggests that the observed detrimental effects on water - water, amino acid - water, and amino acid - amino acid hydrogen bonding induced by $\text{Mg}(\text{ClO}_4)_2$ are also observed here. The structure of I27 has been experimentally demonstrated via atomic force microscopy to be stabilised by a set of six interstrand hydrogen bonds between β strands A' and G which form a mechanical clamp [217, 218, 221]. These two strands are shown in red in figure 6.7.3. Upon mutation induced removal of these hydrogen bonds, unfolding is shown to proceed at a much faster rate. It is also observed using high pressure NMR that when unfolding I27 using a large external pressure that the A'G region is the first region affected before the rest of the protein denatures [304]. It is therefore likely that in the same way it was disrupting hydrogen bonding of water molecules and single amino acids, that $\text{Mg}(\text{ClO}_4)_2$ is able to disrupt these six hydrogen bonds resulting in unfolding of the protein.

Based on the observed perturbations to amino acid and water hydrogen bonding induced by $\text{Mg}(\text{ClO}_4)_2$, and the position of both Mg^{2+} and ClO_4^- on the Hofmeister series

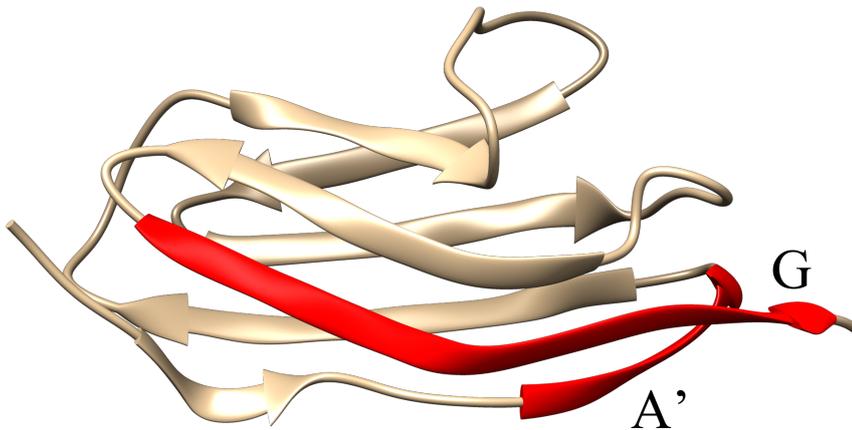


Figure 6.7.3: The secondary structure of I27 with the strands A' and G which form the mechanical clamp highlighted in red.

[47] which predicts they will both act as strong protein denaturants, it is perhaps unsurprising that $\text{Mg}(\text{ClO}_4)_2$ causes I27 to unfold. However, what is perhaps more surprising is that this unfolding does not simply result in the I27 protein going from a totally structured to a totally unstructured state. Using the Contin algorithm [240] and reference set 7 within the online tool DichroWeb [241], it is possible to deconvolute the secondary structure fractions within I27 as a function of increasing $\text{Mg}(\text{ClO}_4)_2$ concentration. The results of these deconvolutions are shown in figure 6.7.4. The Contin algorithm within Dichroweb is capable of deconvoluting a single CD spectra into six different fractions: ordered helix, distorted helix, ordered sheets, distorted sheets, turns, and disordered. For convenience the ordered and distorted helices and sheets have been summed in this figure.

Here we observe at 0 $\text{Mg}(\text{ClO}_4)_2$ concentration that the structure of I27 in its native state, in agreement with previous results [232], is dominated by β sheet structures with relatively large quantities of turns and disordered regions. Rather predictably, as $\text{Mg}(\text{ClO}_4)_2$ concentration is increased and the protein unfolds, the proportion of disordered regions increases at the expense of sheets and turns. However there are two more surprising observations: Firstly, while $\text{Mg}(\text{ClO}_4)_2$ causes the structure to unfold, it still retains a reasonably large degree of sheet and turn content, indicating that the denatured state of I27 induced by $\text{Mg}(\text{ClO}_4)_2$ still retains a degree of secondary struc-

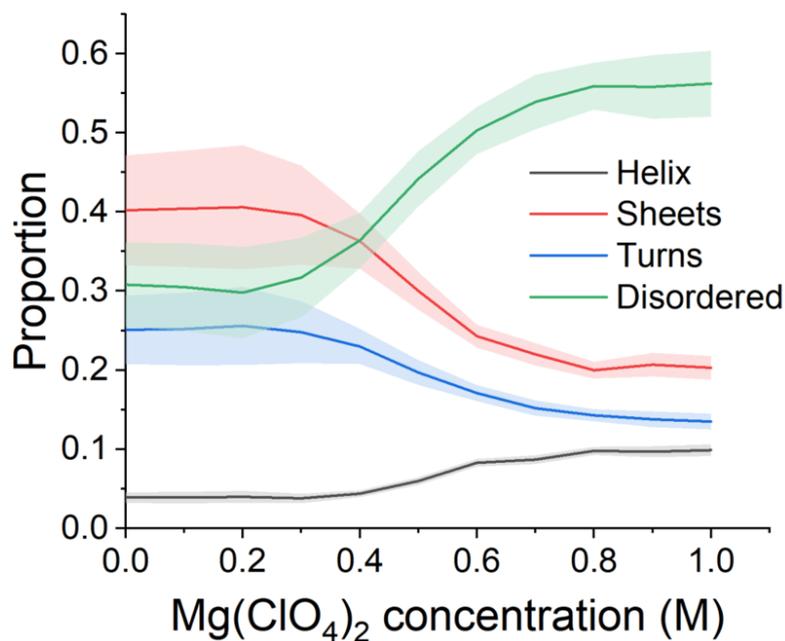


Figure 6.7.4: The secondary structure fractions of I27 as a function of $\text{Mg}(\text{ClO}_4)_2$ concentration as determined from the data displayed in figure 6.7.2(a) using Dichroweb. Associated uncertainties shown by shaded regions around lines.

ture. Secondly, upon increasing $\text{Mg}(\text{ClO}_4)_2$ concentration we observe the emergence of a small population helical structures from what is a fully β sheet protein in its native state. This suggests that the $\text{Mg}(\text{ClO}_4)_2$ unfolded state of I27 is perhaps more similar to a folding/unfolding intermediate, rather than a completely unfolded protein.

The increase of helical content of peptides upon the addition of various cosolute species has been previously observed in the literature. Peggion *et al.* [305] observed that for peptides consisting of lysine and phenylalanine, that NaClO_4 induced a change in peptide structure from a random coil to a helical conformation. Maison *et al.* [306] and Ascitutto *et al.* [307] also observed that ClO_4^- anions increase the helicity of alanine rich peptides. Crevenna *et al.* [24] studied three aqueous 12 amino acid peptides: a neutrally charged peptide $\text{A}(\text{EAAAK})_2\text{A}$, a negatively charged peptide $(\text{AE})_6$, and a positively charged peptide $(\text{AK})_6$ using the salts KF , KCl , NaCl , NaClO_4 , and GdmCl . These results demonstrated that for the neutral peptide, helical content was reduced with increasing salt concentration and temperature, however for the charged peptides

helical content increased with increasing salt concentration. This was determined to be due to the screening of unfavourable electrostatic interactions between similarly charged side chains which prevent helix formation.

The formation of helical domains within β rich proteins has been shown to play an important role in protein folding for certain proteins, most notably β lactoglobulin [308–312]. Here it is suggested that the formation of α helices is required to constrain protein conformations and prevent misfolding, as helices can be formed between adjacent amino acids and therefore form much more quickly, whereas β sheets require interactions between amino acids that are likely to be well separated. Helices are therefore formed in segments of the protein which then come together and allows for the formation of β sheets. It has been suggested since the 1970s that all globular proteins require the formation of helical intermediates to facilitate protein folding. Chen *et al.* [313] demonstrated that the use of trifluoroethanol (TFE), which has been shown to act as a strong stabiliser of helical conformations regardless of amino acid sequence, can be used to accelerate the folding of the protein RIFABP (rat intestinal fatty acid binding protein), which traditionally shows no helical intermediates using stopped flow CD. This result demonstrates the importance of helical intermediates.

These results, and the observations shown in this work, suggest that $\text{Mg}(\text{ClO}_4)_2$ may be a useful tool to study intermediate conformations of proteins. Unfortunately, regarding I27, it is impossible with the techniques discussed here to determine exactly which regions of the protein are undergoing a helical transition. This would require the synthesis of isotopically labelled (^{15}N) I27 for use in 2D NMR experiments, which falls outside the scope of this work [314].

However, a property we *can* quantify of I27 unfolding is the change in Gibbs free energy between the folded and the unfolded state, known as the free energy of unfolding. This is done through fluorescence spectroscopy and application of the two state model, as described in section 2.7. This allows us to quantify the fraction of folded protein as a function of denaturant concentration, in this case urea. By performing a urea induced denaturation of I27 in the presence of a set concentration of a cosolute, we can investigate

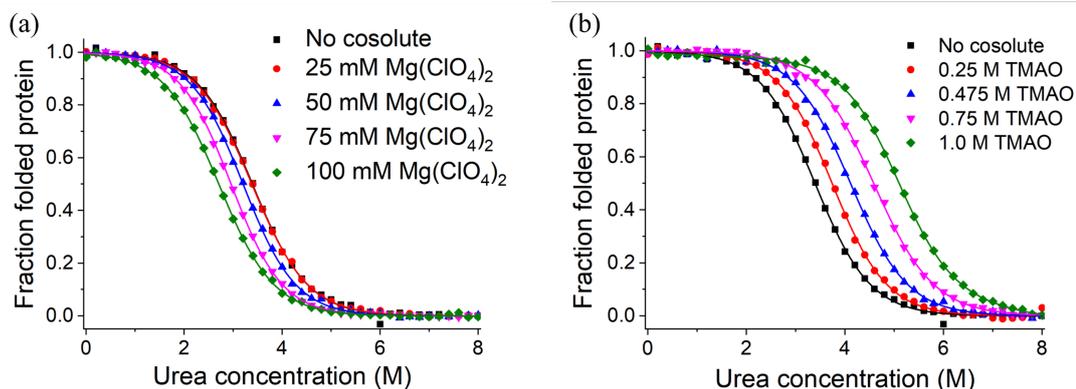


Figure 6.7.5: The folded fraction of I27 as a function of urea concentration in the presence of a set concentration of $\text{Mg}(\text{ClO}_4)_2$ (a) or TMAO (b). Data fit to the two state model shown in equation 2.7.1.

the effect of the cosolute on the free energy of unfolding. Here (rather predictably) we choose to investigate the two co-solutes that have been the central focus of this thesis: TMAO and $\text{Mg}(\text{ClO}_4)_2$. These results are presented in figure 6.7.5.

Here we observe that the urea induced denaturation of I27 is well described by the two state model of protein unfolding, as evidenced by the high quality fits of the two state model equation presented in equation 2.7.1 to the data. From this data the effect that the two cosolute species are having on the stability of I27 is clear. With increasing $\text{Mg}(\text{ClO}_4)_2$ concentration the data is shown to shift to lower urea concentrations, meaning the concentration of urea required to unfold the protein is shown to decrease, and I27 is becoming less stable. Conversely, with increasing TMAO concentration the data is shown to shift to higher urea concentrations as more urea is required to unfold the protein and I27 is becoming more stable. The opposing effects of $\text{Mg}(\text{ClO}_4)_2$ and TMAO on the stability of water - water hydrogen bonding discussed in chapter 5 are therefore mirrored here for the stability of I27. The predicted Gibbs free energies of unfolding for I27 in the presence of the various concentrations of each of the two solute species is presented in figure 6.7.6(a).

Here we observe that the free energy of unfolding for I27 in the presence of no other co-solutes is calculated to be 15 ± 1 kJ/mol, consistent with previous literature [315]. Upon the addition of $\text{Mg}(\text{ClO}_4)_2$ this value is shown to drastically reduce as the

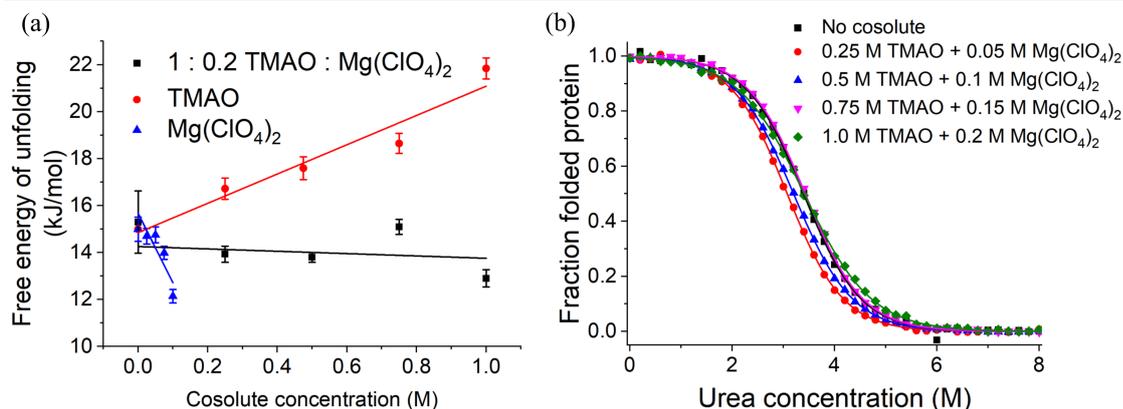


Figure 6.7.6: The Gibbs free energy of unfolding predicted through the two state model for I27 in the presence of Mg(ClO₄)₂ (blue), TMAO (red), and a 0.2 : 1 molar ratio mixture of Mg(ClO₄)₂ : TMAO (black) (a). Error bars calculated through the fitting of the two state model to the raw data. The raw data for the folded fraction of I27 as a function of urea concentration containing a 0.2 : 1.0 molar ratio of Mg(ClO₄)₂ : TMAO (b).

protein becomes less stable, and the opposite is true for TMAO addition. By performing linear fits to this data we can state that the concentration dependence of Mg(ClO₄)₂ or TMAO on the free energy of unfolding of I27 ($\Delta\Delta G$) is equal to -30 ± 9 kJ/(mol M) for Mg(ClO₄)₂ and 6 ± 1 kJ/(mol M) for TMAO. From these $\Delta\Delta G$ values it therefore follows that an I27 solution containing a 0.2 : 1.0 molar ratio of Mg(ClO₄)₂ : TMAO should leave the stability of I27 unchanged overall, and the folded fraction of I27 as a function of urea concentration should be essentially identical to I27 with no co-solutes. Figure 6.7.6(b) and the corresponding black data points in figure 6.7.6(a) show that this is indeed the case.

Interestingly this ratio of 0.2 : 1.0 Mg(ClO₄)₂ : TMAO concentration required for no overall net effect to protein stability is the same ratio that was estimated in section 5.2 from Mg(ClO₄)₂ melting temperature data [277] and the concentration of TMAO in the muscle tissue of bony fish as a function of depth [86]. However, it is in stark contrast to the results observed in chapter 5, which stated that TMAO is 1.54 times more effective at perturbing water structure and dynamics than Mg(ClO₄)₂. This leads us to what is arguably the most important observation of this thesis: simply studying the effects of different externalities (ion addition, pressure, temperature, etc.) on the solvent environment may only yield qualitatively informative results on how these externalities may

affect the stability of biomolecules. While the two are indisputably fundamentally linked, and we have shown throughout this thesis that in general the perturbations of different externalities to water - water hydrogen bonding are well reflected by the perturbations to biomolecule stability, for quantitative results it is vital that the biomolecules are included in the experimental sample. This is in good agreement with the findings of Ball and Hallsworth, who reviewed the validity of the terms “chaotrope” and “kosmotrope”. These have been used to describe a chemical as structure breaking or structure making with regards to its affect on biomolecular stability and water structure, however these two effects are often different depending on the biomolecule and precisely how one defines “structure” so one must use them with caution [53]. In the next section we will begin to address how one might begin to include biomolecules in neutron scattering studies.

6.8 Neutron Scattering on Aqueous Peptides

As discussed in the previous section, if we wish to understand the biological role of water in extreme conditions, we must prepare samples which reflect the extreme conditions and also contain a model biomolecule. As we wish to understand biomolecular stability in this context, the model biomolecule must behave similarly to naturally occurring biomolecules. In this respect single amino acids such as glycine, as studied in this chapter, are insufficient because while they are certainly naturally occurring biomolecules, they are too small to adopt a folded structure and therefore are not representative of naturally occurring biomolecules. The model protein I27 we have also studied in this chapter has the opposite problem. If we wished to examine the structure of water around biomolecules and co-solutes using neutron scattering and EPSR, the inclusion of such a large (93 amino acids) biomolecule would mean that the number of distinct inter-atomic correlations, as well as the need for increasingly large box sizes to account for the large scale structures as discussed in section 2.2, would mean that EPSR would quickly become overwhelmed. Hence we require a model biomolecule that serves as a middle ground between these two options.

For this we turn to the 10 amino acid peptide, CLN025 [214] (amino acid sequence:

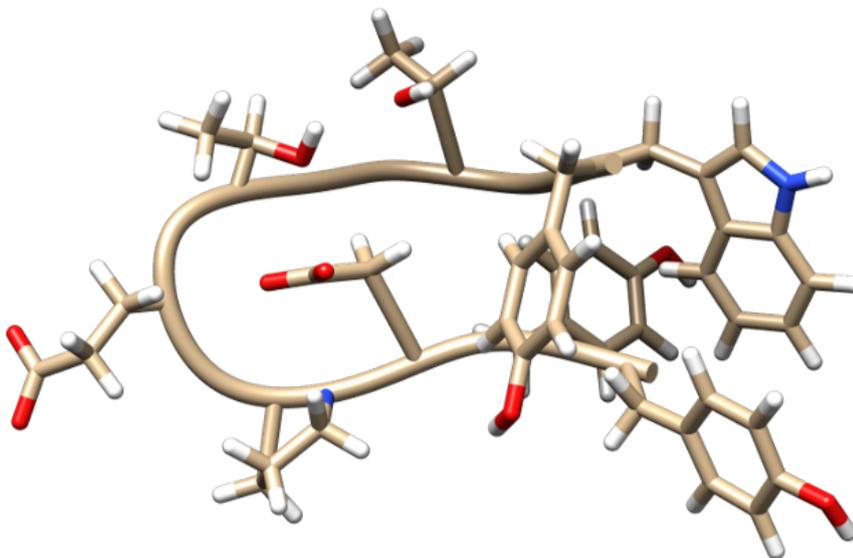


Figure 6.8.1: The β hairpin structure of CLN025. Peptide backbone shown as a smooth line and side chains shown with complete atoms.

YYDPETGTWY), shown in figure 6.8.1. This forms a simple β hairpin in solution and it has been previously shown using CD that it denatures in a similar manner and with comparable free energy of unfolding to complete proteins. It therefore behaves as a naturally occurring biomolecule, despite the fact it is completely synthetic, while being much smaller than naturally occurring biomolecules. It was also found to be capable of suspension at high concentrations in aqueous solution, which is a necessity for meaningful analysis with EPSR following the acquisition and correction of neutron scattering data. CLN025 is therefore ideally suited to the role of a model biomolecule for neutron scattering and EPSR experiments.

Before one can begin to investigate the effects of different externalities of biomolecular stability using neutron scattering and EPSR, one must first be able to meaningfully investigate CLN025 under ambient conditions using these techniques. Hence CLN025 was prepared as detailed in section 2.6 and suspended to 150 mg/mL in 250 mM NaOH in H₂O, D₂O, and HDO to neutralise the acidic residues on the peptide and promote its solubility. The three resultant samples were then investigated using NIMROD at three different temperatures: 25, 60, and 90 °C. These temperatures were chosen as increasing temperature is the simplest way to denature the peptide, hence in principle one could

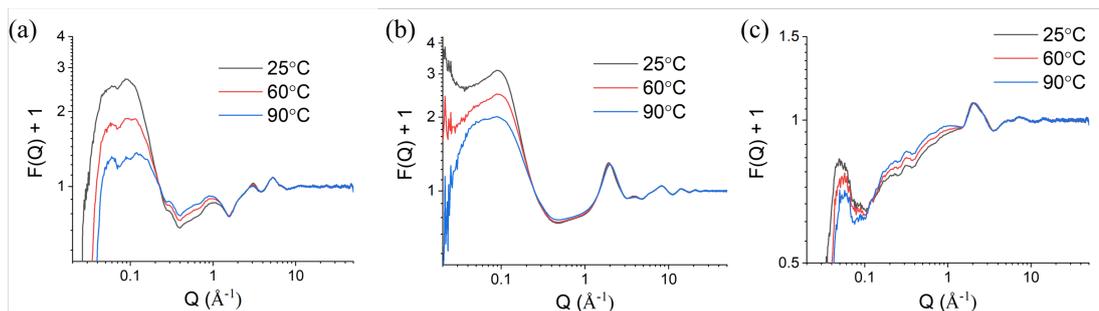


Figure 6.8.2: The corrected neutron scattering data obtained using the NIMROD instrument for CLN025 in H₂O (a), D₂O (b), and HDO (c) with 250 mM NaOH. Corrected $F(Q)$ vertically shifted by 1 to allow for plotting on a logarithmic axis.

follow the unfolding of CLN025 via EPSR. The temperatures therefore correspond to previous literature values for temperatures at which the peptide is fully folded, partially unfolded, and as unfolded as possible within safe working limits of the apparatus. The resultant raw scattering data, following correction by Gudrun, are presented in figures 6.8.2.

Analysing the data using EPSR is currently ongoing due to the series of complications related to including a molecule as complex as CLN025 into EPSR, however examination of the raw neutron scattering data can yield some preliminary results. This data appears to be comprised of two distinct areas, the scattering above $Q \approx 1$ which is due to smaller scale structures and the scattering below $Q \approx 1$ which is due to larger scale structures. Above $Q \approx 1$ we observe data highly reminiscent of that observed for simple aqueous solutions, suggesting that this portion of the data primarily corresponds to interatomic correlations between pairs of water molecules and water molecules hydrating both CLN025 and the two ion species present in the samples. Below $Q \approx 1$ we observe a large signal intensity in both the H₂O and D₂O sample. These correspond to highly structured large scale structures and therefore likely correspond to interatomic correlations within the folded CLN025 molecule.

This suggestion is strengthened further when we consider that the signal in this region in the HDO sample is relatively negligible. It can be predicted by considering the sums of the coherent scattering lengths for every atom in the CLN025 molecule [183] that its immersion into HDO at 150 mg/mL should render CLN025 effectively a null scatterer. As

no signal is observed in the CLN025 sample prepared in HDO it can be confidently stated that the large signal observed in both the H₂O and D₂O samples arises as result of large scale structures in the CLN025 molecule. We also observe that in all three samples, the intensity of this CLN025 region decreases with increasing temperature. This is consistent with what would be expected upon thermal denaturation of the peptide, as an unfolded peptide will certainly have fewer reoccurring structural motifs than a folded peptide. Preliminary observations from the corrected scattering data are therefore promising, however a complete EPSR analysis of the samples is still ongoing.

6.9 Conclusions

In this final results chapter we have presented neutron scattering data and EPSR analysis on glycine in aqueous solution and in aqueous Mg(ClO₄)₂. These results have helped expand the findings of chapter 5, which showed that Mg(ClO₄)₂ destabilises hydrogen bonds between bulk water molecules in aqueous solution, to show that Mg(ClO₄)₂ also destabilises hydrogen bonding between glycine molecules. This in turn goes on to hinder their clustering in solution. At 25 °C we observe that 32.0% of glycine molecules can be found in clusters containing two or three glycine molecules, but upon the addition of 2.7 M Mg(ClO₄)₂, this fraction drops to 20.3%. However, we also observe that this reduction in clustering can be partially restored by enhancing hydrogen bonding by decreasing temperature to -20 °C, at which point 22.1% of glycine molecules can be found in clusters. This observation has implications on the potential ability of life to form on Mars, as the building blocks for life would be less able to come together in solution and therefore less able to form the required chemical bonds which would result in functional biomolecules.

We then went on to examine how Mg(ClO₄)₂ perturbs the stability of complete proteins, rather than just single amino acids, by studying the model β sheet protein I27 using CD and fluorescence spectroscopy. Here we demonstrated that Mg(ClO₄)₂ causes the protein to denature, however even in its completely Mg(ClO₄)₂ denatured state it still retains a degree of secondary structure, and even gains a small amount of α helical

content. We then quantified this denaturing ability through fluorescence spectroscopy and determined that the $\Delta\Delta G$ value for $\text{Mg}(\text{ClO}_4)_2$ is equal to -30 ± 9 kJ/(mol M). However, just as it has been shown to stabilise proteins against a host of denaturing externalities [60, 64, 69, 70, 145, 154, 256–258], as well as preserving hydrogen bonding in bulk water against pressure as shown in chapter 4 and $\text{Mg}(\text{ClO}_4)_2$ as shown in chapter 5, the naturally occurring osmolyte TMAO is capable of resisting the $\text{Mg}(\text{ClO}_4)_2$ induced denaturation to I27. We determined that TMAO has a $\Delta\Delta G$ value of 6 ± 1 kJ/(mol M), and therefore a mixture of 0.2 : 1.0 $\text{Mg}(\text{ClO}_4)_2$: TMAO leaves the stability of the protein I27 relatively unaffected.

This is in close agreement to the predicted ratio at which $\text{Mg}(\text{ClO}_4)_2$ and TMAO cancel each other out in section 5.2, but strongly disagrees with the observations of perturbations to water structure and dynamics by $\text{Mg}(\text{ClO}_4)_2$ and TMAO presented in chapter 5. This disparity demonstrates the importance of including a model biomolecule if one wishes to gain a detailed understanding of the role different externalities play in perturbing water structure and dynamics and the resultant perturbations to biomolecule stability. We then go on to display the preliminary data of a potential model biomolecule, CLN025, in solution. While detailed EPSR analysis is yet to be completed, initial signs suggest that this is a good choice of model biomolecule. Its inclusion into more complex aqueous solutions should allow for atomic scale insight into structural perturbations arising from changes to its solvent environment and give us new fundamental insight into the intimate relationship between water and biomolecules.

Chapter 7

Conclusions and Future Work

In this thesis we have primarily used neutron scattering combined with the computational modelling technique EPSR to visualise perturbations to the structure of aqueous solutions as a result of different externalities, such as ion addition, pressure, and temperature. We have also been able to validate some of the findings of this method using NMR spectroscopy, as well as expand our structural results to also consider dynamic perturbations to aqueous solutions. In this final chapter we will summarise the main conclusions from the four results chapters in section 7.1 before going on to consider the various avenues this research has opened for future work in section 7.2.

7.1 Concluding Remarks

We began by using this combination of techniques to give fundamental insight into the structure, dynamics, and thermodynamics of aqueous systems containing simple monovalent potassium halide salts in chapter 3. This was investigated as ions are present in solution at every known length scale which is important to life, from the interactions between ions and biomolecules at the molecular level [1, 233, 316, 317] to the presence of ions in both salt and freshwater environments in which biological organisms thrive [318]. Regardless of any additional extremes in pressure, temperature, pH, etc., ions are

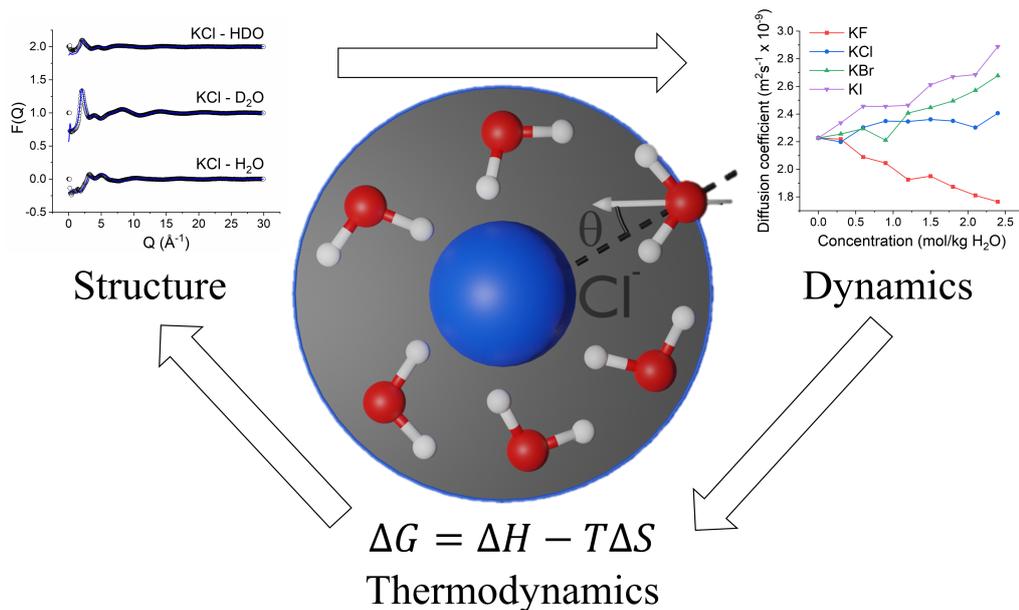


Figure 7.1.1: A diagram representing the links between structure, dynamics, and thermodynamics in aqueous solutions of potassium halides discussed in chapter 3.

present. It therefore serves as sensible starting point to begin to investigate the biological role of water in extreme conditions. Potassium halide salts serve as sensible model ions as they are monovalent salts with high biological relevance [175, 176] for which neutron scattering data was already available [128].

We observed that in the case of the simple aqueous potassium halides KF, KCl, KBr and KI, that the introduction of the ions into water results in an inward movement of the second peak in the $O_w - O_w$ RDF, similar to the perturbation that arises as a result of a large external pressure and consistent with previous studies on aqueous ions [126–128, 130, 132, 133, 136]. This inward movement was very similar for all four ion species, however the ion - O_w RDFs were far more ion specific. We observed that as the size of the ion increases, and therefore the charge density decreases, that the anion - O_w and anion H_w RDFs become shorter, broader, and move outwards to larger distances, indicative of less well defined hydration shells. This reduction in definition was also observed through the water dipole angle distributions around the halide anions, where we observe that the distributions become shorter and broader as the size of the ion increases from F^- to I^- .

However, while the structure of water around the ion varies strongly with ion species,

the same cannot be said for bulk water structure, as the dipole angle distributions and average hydrogen bond interaction energy between bulk water molecules are shown to be relatively unchanged for all four salts at every studied concentration. We demonstrate that enthalpic interactions between ions and surrounding water molecules become negligible at distances above approximately 3.5 - 4.5 Å depending on the ion species.

The tendency of potassium halide ions to induce structure in their first hydration shell is well reflected by their enthalpies of hydration, as calculated through EPSR and the extended analysis routine. Here we observe that with increasing anion size that the enthalpy of hydration becomes less negative, indicative of less stable interactions between the ions and the surrounding water molecules, and resulting in less well defined structures through the dipole angle distributions and anion - water RDFs. These results were shown to be in good agreement with experimental results on the Gibbs free energy, enthalpy, and entropy of hydration of potassium halides determined by Tissandier *et al.* [35]. The less well defined hydration structure with increasing ion size is also reflected by a reduced entropy penalty of hydration.

The increasingly unstructured hydration water molecules around a central ion species with increasing ion size are also coupled with a perturbation to water molecule dynamics. As the enthalpy of hydration becomes less negative, and the hydration layer becomes less structured, we observe that the water molecule diffusive and rotational dynamics become accelerated. This translates to increased freedom of a water molecule in the hydration layer of an ion to explore its available conformations with increasing ion size, and is another likely origin for the observed reduced entropy of hydration penalty with increasing ion size.

The key result from this chapter is that in the case of aqueous potassium halides, water structure and dynamics are strongly perturbed in the hydration shell of the ions and that these perturbations correlate well with observed thermodynamic measurements from bulk ensemble measurements. We also observe that these perturbations do not extend significantly into the bulk water structure. This allows us to speculate that perturbations to other solute species that may be present in aqueous solutions of potassium

halides would likely be due to local effects between the ions and the other solute species, rather than due to changes in the bulk water structure. The results of this chapter also help demonstrate the reliability of the custom built extended analysis routine, as the results gained from its application are in good agreement with results from other scientific techniques.

We then subjected water and aqueous solutions of the piezolyte TMAO to high pressures and studied structural perturbations with neutron scattering and EPSR to determine whether TMAO is capable of restoring the hydrogen bonded network of water against pressure induced perturbations in chapter 4. For this we chose to investigate pure water and aqueous TMAO at 2.0 mol/kg H₂O at a low pressure of 25 bar, chosen to ensure the high pressure rig was in contact with the sample while causing relatively negligible perturbations to water structure, and 4 kbar, chosen based on previous literature [123] such that pressure induced perturbations to water structure should be obvious in the O_w - O_w RDF while not being so severe that any resisting power of TMAO would be overwhelmed.

Here we clearly observed the ability of TMAO to resist the pressure induced structural perturbation to water structure through the O_w - O_w RDFs. At 25 bar, the second peak in the O_w - O_w RDF moved slightly inwards relative to pure water, and at 4 kbar it moves further inwards and begins to overlap with the first peak. In the presence of TMAO at 2.0 mol/kg H₂O at both pressures these shifts were not as significant and the structure of water is less perturbed by an external pressure. This is also reflected in the water - water SDFs, where the second hydration shell is shown to collapse inwards as pressure is increased from 25 bar to 4 kbar, but less so in the presence of TMAO.

Unlike the results of aqueous potassium halides, where the addition of the solute only causes significant perturbations in the first hydration shell of the solute and not in the bulk water, the addition of TMAO results in measurable affects to both the hydration water and the bulk water. The water dipole angle distributions around a central bulk water molecule were shown to be less perturbed by an external pressure in the presence of TMAO relative to pure water, and the bulk water - water hydrogen bond interaction

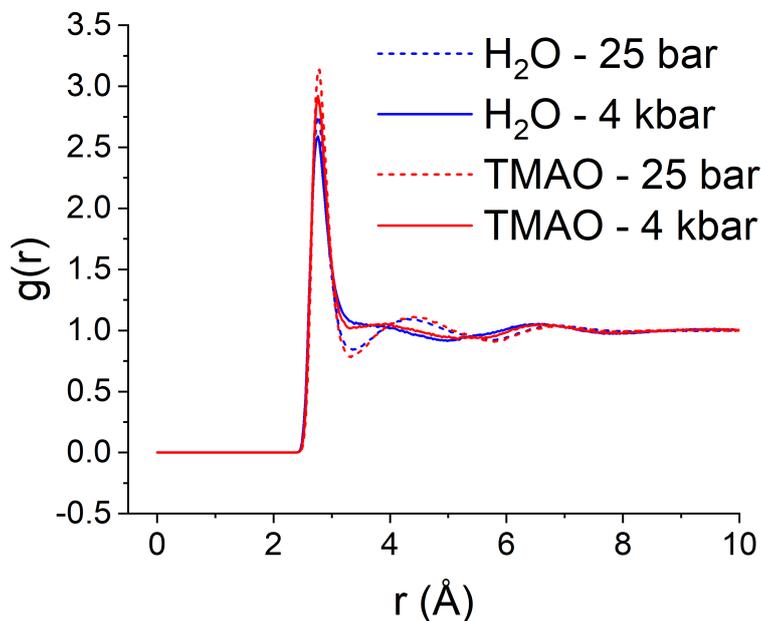


Figure 7.1.2: The $O_w - O_w$ RDFs for aqueous TMAO at 2.0 mol/kg H_2O at 25 bar and 4 kbar reported in chapter 4 that demonstrate the ability of TMAO to resist the pressure induced perturbation to water structure.

energy was shown to be more stable at both pressures relative to pure water in the presence of TMAO.

We attempt to understand the origins of this pressure resisting ability of TMAO by considering the structure of water hydrating the hydrophilic oxygen and the hydrophobic methyl groups. Here we observe that in both instances the $O_T - O_w$ and the $C_T - O_w$ RDFs are relatively insensitive to increasing pressure from 25 bar to 4 kbar. The same is observed for the water dipole angle distributions around the TMAO oxygen. We then went on to quantify the interaction energy of the TMAO - water hydrogen bond, and demonstrated that TMAO forms much stronger hydrogen bonds between water molecules and its oxygen than the hydrogen bonds formed between bulk water molecules. These TMAO - water hydrogen bonds were also relatively insensitive to pressure.

The picture that begins to emerge by considering all of these observations together is that TMAO forms a hydration structure around both its hydrophobic and hydrophilic groups that is less compressible than bulk water, and acts as an anchor point from which

water can build a more structured and stable network. We suggest that this is likely part of the mechanism by which TMAO is capable of preserving biomolecular stability against the denaturing effects of increased hydrostatic pressure, as has been previously observed [68, 84, 85]. The validity of this suggestion is strengthened when we consider that biomolecule-stabilising osmolytes like TMAO are thought to be preferentially excluded from the biomolecule surface [67–73], and therefore their ability to stabilise biomolecules must be as a result of perturbations to the bulk water network rather than direct interactions with the biomolecule.

We then went on to consider whether TMAO could also restore the hydrogen bonded network of water against the pressure-like perturbation induced by the salt $\text{Mg}(\text{ClO}_4)_2$ with an aim to understand potential mechanisms that life could use to survive in an extreme Martian-like environment in chapter 5. Here we chose to investigate tertiary solutions of $\text{Mg}(\text{ClO}_4)_2$ and TMAO at two different $\text{Mg}(\text{ClO}_4)_2$ concentrations. A concentration of 0.2 M $\text{Mg}(\text{ClO}_4)_2$ and 1.0 M TMAO was chosen as it was estimated from $\text{Mg}(\text{ClO}_4)_2$ solution melting point data [277] and previous neutron scattering data on aqueous $\text{Mg}(\text{ClO}_4)_2$ [133] that 0.2 M $\text{Mg}(\text{ClO}_4)_2$ should induce a similar perturbation to water structure as the application of approximately 2 kbar. Data on the concentration of TMAO in the muscle tissue of bony fish as a function of depth [86], and therefore pressure, also allowed us to estimate that 1.0 M TMAO should resist a pressure of approximately 2 kbar. The studied mix of 0.2 M $\text{Mg}(\text{ClO}_4)_2$ and 1.0 M TMAO should therefore render an unchanged bulk water structure. We also chose to study 2.7 M $\text{Mg}(\text{ClO}_4)_2$ and 1.0 M TMAO as this is a $\text{Mg}(\text{ClO}_4)_2$ concentration that had been previously studied by neutron diffraction and EPSR [133].

In these tertiary solutions the resisting ability of TMAO to the pressure like perturbation induced by $\text{Mg}(\text{ClO}_4)_2$ was again evident in the $\text{O}_w - \text{O}_w$ RDFs and the water - water SDFs. The inward shifts of the peaks in the $\text{O}_w - \text{O}_w$ RDFs induced by $\text{Mg}(\text{ClO}_4)_2$ were shown to be less significant in the presence of TMAO, but not to the same extent that they were in the high pressure aqueous TMAO results from chapter 4. The bulk water - water hydrogen bond interaction energy was also shown to be less stable upon $\text{Mg}(\text{ClO}_4)_2$ addition, but that this recovered in the presence of TMAO. This result is

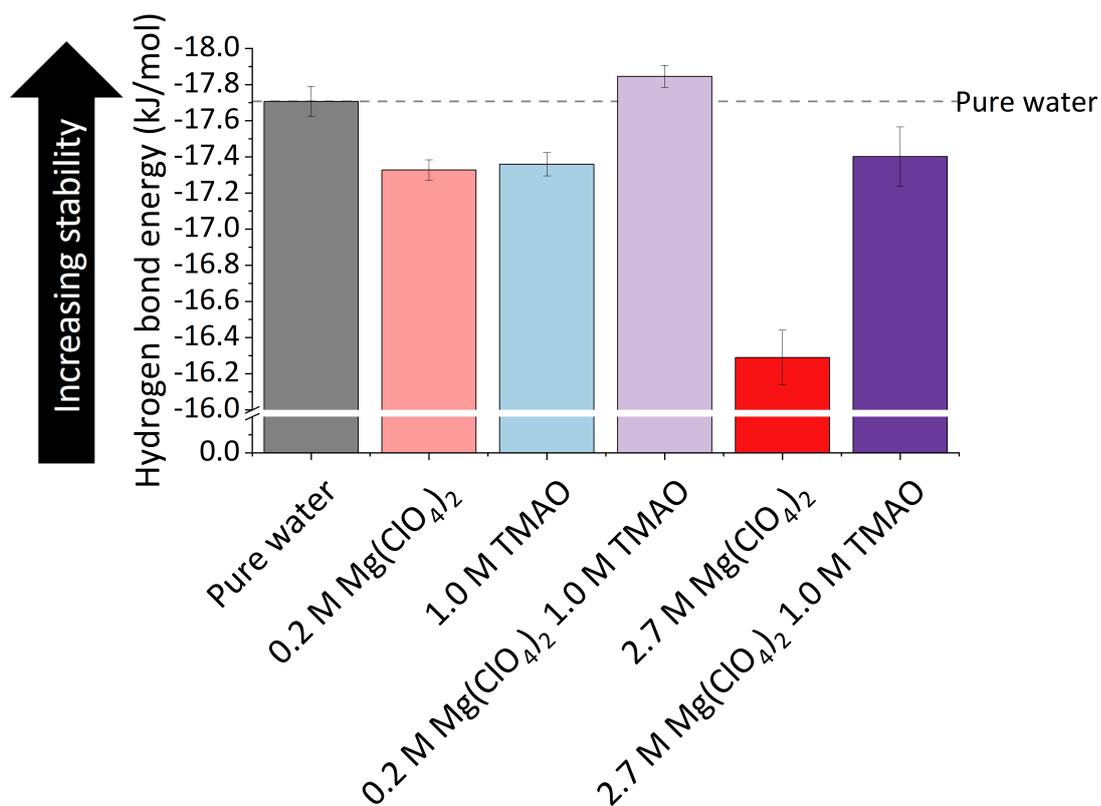


Figure 7.1.3: The average bulk water - water hydrogen bond interaction energy calculated for each of the six solutions studied in chapter 5. These demonstrate the ability of TMAO to resist the pressure-like structural perturbation to water induced by $\text{Mg}(\text{ClO}_4)_2$.

significant beyond examining the opposing effect of the two solutes, as it demonstrates that more complex salts like $\text{Mg}(\text{ClO}_4)_2$ can perturb water structure outside the first hydration shell of the ions, whereas the simple potassium halide salts discussed in chapter 3 cannot. It was also observed that, similar to the results of high pressure aqueous TMAO, the hydration structure of water around both the hydrophilic and hydrophobic regions of the TMAO molecule was less sensitive to $\text{Mg}(\text{ClO}_4)_2$ addition than the structure of bulk water. TMAO again seems to be serving as an anchor point in aqueous solution that can act to resist the perturbations to water structure by various externalities.

Unlike the results from high pressure aqueous solutions studied in chapter 4, the obtaining of which requires specialised equipment to handle the extreme pressures required, the results of tertiary solutions at ambient pressure can be relatively easily validated using more standardised experimental methods. Here we chose to use NMR to validate the result that $\text{Mg}(\text{ClO}_4)_2$ destabilises water - water hydrogen bonding, which is subsequently restabilised by the addition of TMAO. Tertiary solutions of TMAO and $\text{Mg}(\text{ClO}_4)_2$ were prepared from 0 to 1 mol/kg H_2O in 0.2 mol/kg H_2O increments, and the relative peak position of the ^1H NMR peak corresponding to the water hydrogen was monitored as this is indicative of the strength of hydrogen bonding. This allowed us to demonstrate, consistent with the results gained through neutron scattering and EPSR, that $\text{Mg}(\text{ClO}_4)_2$ does indeed destabilise water - water hydrogen bonding, and TMAO does indeed stabilise water - water hydrogen bonding. The effect of the two solutes on water structure therefore act in opposition to one another.

The use of NMR also allows us to expand the structural study of these tertiary aqueous solutions to consider perturbations to water molecule dynamics. Here we observe the interesting result that while TMAO and $\text{Mg}(\text{ClO}_4)_2$ act in opposition to each other with respect to the way they perturb water structure, they act to perturb water dynamics in an additive manner, both resulting in retarded water molecule rotational and diffusive dynamics. This effect can also be observed within the EPSR simulations, as the abundance of hydrogen bonded conformations of water molecules associated with hydrogen bond switching events (bifurcated hydrogens) are shown to decrease with increasing TMAO and $\text{Mg}(\text{ClO}_4)_2$ concentration.

By grouping all of these results together we can finally quantify the general perturbing ability of $\text{Mg}(\text{ClO}_4)_2$ to water structure and dynamics relative to TMAO. This is done by plotting five observed variables (^1H NMR peak shift, bulk water - water hydrogen bond interaction strength, T_1 relaxation time, the self-diffusion coefficient of water D , and the abundance of bifurcated hydrogens), as a function of effective total concentration. The effective total concentration of the two solutes is calculated by the sum of the concentration of $\text{Mg}(\text{ClO}_4)_2$ and the concentration of TMAO multiplied by a weighting factor g . This weighting factor therefore represents how much more effective a perturbing agent TMAO is relative to $\text{Mg}(\text{ClO}_4)_2$. By varying this weighting factor such that the observed variables all fall on a single master curve for each of the five data sets, we determine that TMAO is a more effective general water perturbing agent than $\text{Mg}(\text{ClO}_4)_2$ by a factor of 1.54 ± 0.06 . This method therefore represents a reliable way of quantifying the perturbing ability of one solute relative to another in tertiary solutions that could easily be applied to any number of other tertiary solutions.

Finally, in chapter 6, we consider how the externalities studied in the previous results chapters go on to affect biomolecule stability and self-association. To achieve this we used neutron scattering and EPSR to examine the effect of $\text{Mg}(\text{ClO}_4)_2$ on the ability of the amino acid glycine to form clusters in solution with an aim of understanding protein stability and formation in Martian like environments. This allows us to demonstrate that glycine molecules form clusters of up to three molecules in aqueous solution through the formation of hydrogen bonds between their amine and carbonyl groups. The subsequent introduction of $\text{Mg}(\text{ClO}_4)_2$ then causes the proportion of clustered glycine molecules to reduce as inter amino acid hydrogen bonds are disrupted by the salt. However, this clustering can be recovered by subsequently enhancing inter amino acid hydrogen bonding by reducing temperature.

We also expanded these results to consider the effect $\text{Mg}(\text{ClO}_4)_2$ has on the structure of the model protein I27 using CD by preparing samples of I27 in aqueous solution containing 0 - 1 M $\text{Mg}(\text{ClO}_4)_2$ in 0.1 M increments. These results demonstrated, in a similar manner to the detrimental effects $\text{Mg}(\text{ClO}_4)_2$ has on hydrogen bonding between water molecules and amino acids, that $\text{Mg}(\text{ClO}_4)_2$ causes I27 to unfold likely through the

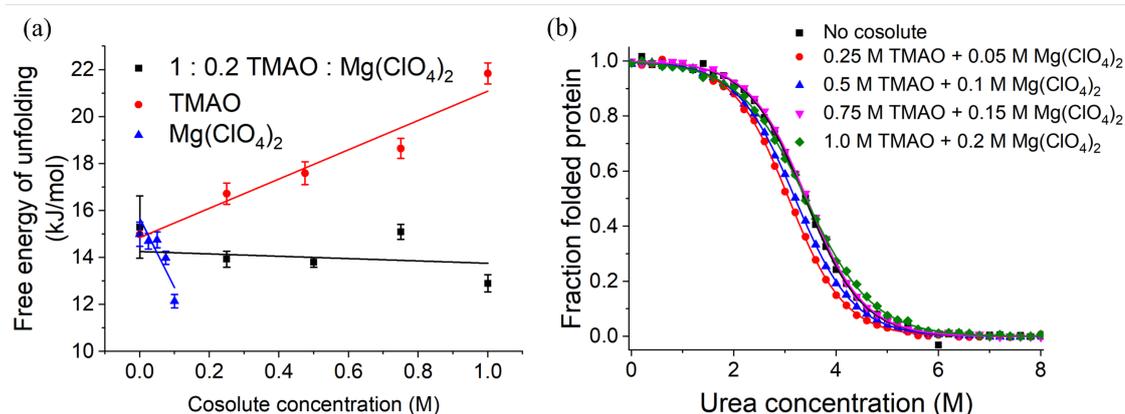


Figure 7.1.4: The effect of TMAO and $\text{Mg}(\text{ClO}_4)_2$ on the free energy of unfolding on the protein I27 both in binary and tertiary solutions (a) and the fraction of folded protein in mixed TMAO and $\text{Mg}(\text{ClO}_4)_2$ solutions as a function of urea concentration (b) discussed in chapter 6.

destabilisation of intraprotein hydrogen bonds. Surprisingly however, this does not result in complete unfolding of the protein. The deconvolution of the obtained CD spectra for I27 using Dichroweb [241] showed that with increasing $\text{Mg}(\text{ClO}_4)_2$ concentration that the proportion of disordered protein increases at the expense of β -sheets and turns. However, even in the fully $\text{Mg}(\text{ClO}_4)_2$ denatured state, the proportion of β -sheets and turns is still relatively high. We also observe the emergence of a small proportion of α -helical structures. The promotion of α -helical structures in peptides and proteins as a result of salt addition has been previously observed in the literature [24, 305, 306]. The formation of α -helical intermediates even plays an important role in the folding of β -rich proteins such as β lactoglobulin [308–312]. The addition of $\text{Mg}(\text{ClO}_4)_2$ could therefore potentially be used as a method of investigating protein folding intermediates by stabilising helical structures.

The detrimental effects of $\text{Mg}(\text{ClO}_4)_2$ to the thermodynamic stability of I27 were then quantified using fluorescence spectroscopy. Here it was observed that $\text{Mg}(\text{ClO}_4)_2$ causes the free energy of unfolding of I27 to reduce, and has a $\Delta\Delta G$ value of -30 ± 9 kJ/(mol M). Conversely it was also observed through fluorescence spectroscopy that TMAO, similar to the manner in which it promotes the hydrogen bonding stability of bulk water, causes the thermodynamic stability of I27 to increase, and has a $\Delta\Delta G$ value of 6 ± 1 kJ/(mol M). These values suggest that $\text{Mg}(\text{ClO}_4)_2$ is five times more

effective at reducing I27 stability than TMAO is at enhancing I27 stability. This was then validated, as mixtures of 0.2 : 1.0 mol $\text{Mg}(\text{ClO}_4)_2$: TMAO concentration solutions yielded a relatively unchanged I27 stability.

This ratio of the relative perturbing ability of $\text{Mg}(\text{ClO}_4)_2$ to TMAO determined through their perturbations to I27 stability is identical to the predicted ratio determined in section 5.2, however it is in stark contrast to the calculated value in section 5.8 determined through their relative perturbations to water structure and dynamics. The origin of this stark difference is likely due to the difference in mechanisms through which the two solutes perturb I27 stability. As a protein denaturant, $\text{Mg}(\text{ClO}_4)_2$ is likely enriched at the protein surface, whereas TMAO is likely to be excluded [67–73]. $\text{Mg}(\text{ClO}_4)_2$ is therefore able to decrease protein stability through direct interactions, whereas TMAO must promote protein stability through indirect interactions through several mediating water layers. It is therefore little surprise that while we demonstrate that TMAO is a more effective perturber of water structure and dynamics than $\text{Mg}(\text{ClO}_4)_2$, this is not the case when we consider their relative perturbations to protein stability.

This result therefore demonstrates that while observations of the perturbations to water structure and dynamics by a solute can help predict their likely effect on biomolecular stability, in order to validate these predictions it is vital that a model biomolecule is also included. We therefore displayed preliminary neutron scattering data of a potential model biomolecule, the 10 amino acid β hairpin peptide CLN025, which could potentially fulfill this role in future neutron scattering experiments. This is a suitably small molecule such that it could reasonably be incorporated into EPSR analysis, yet it has similar thermodynamic behavior to complete proteins. The analysis of this data is ongoing.

7.2 Future Work

The results chapters of this thesis represent a complete story that begins with the perturbations to water structure and dynamics as a result of the inclusion of simple ions, goes

through the perturbations to water structure and dynamics that would likely be found in extreme environments, like the deep ocean or in the Martian subsurface lakes, demonstrates how protecting osmolytes like TMAO can be used to overcome these extremes, and finally shows how these perturbations go on to affect the stability and self-association of biomolecules. However, much like any scientific endeavour, while we have managed to build a single complete story, there are still many sequels and subplots that could be explored. Some of the options for future research based on the findings of this thesis will now be discussed in the following section.

7.2.1 High Pressure NMR

In chapters 3 and 5 we demonstrated that NMR can be used to validate and expand the structural insight into aqueous solutions offered to us by neutron diffraction and EPSR. However, due to the lack of availability of the appropriate experimental apparatus, we did not validate or expand the findings concerning aqueous TMAO under high hydrostatic pressures. NMR spectrometers suitable to the investigation of solutions under high pressures certainly exist, and have been well used to monitor pressure induced denaturation of proteins and various other pressure related phenomena in aqueous solutions of biomolecules [79, 81, 319, 320]. It would therefore be helpful to investigate the same TMAO solution at 2.0 mol/kg H₂O at pressures up to 4 kbar using NMR spectroscopy.

This would serve several purposes: firstly, it would allow us to validate the findings presented in chapter 4, namely that increasing pressure results in less stable water - water hydrogen bonding and that TMAO resists this perturbation by enhancing water - water hydrogen bonding, by monitoring the water ¹H NMR peak position as a function of pressure. Secondly, the nature of NMR also means that it is a far higher throughput technique than neutron scattering and EPSR, hence a much greater range of TMAO concentrations and pressures could be explored. This would allow for a detailed understanding of the true pressure resisting ability of TMAO, rather than the crude estimation made in section 4.2. It would also allow for the identification of any particularly interesting concentration or pressure regions, which could then be further explored

using neutron diffraction and EPSR. Lastly, high pressure NMR would also allow us to comment on dynamic perturbations to the system resulting from pressure and TMAO addition in the same way as we have done in chapters 3 and 5.

7.2.2 A Model Biomolecule for Neutron Scattering: The β -Hairpin Protein CLN025

In section 6.8 we presented corrected neutron scattering data for the aqueous peptide CLN025 at 150 mg/mL at 25, 60, and 90 °C. The next step is therefore the analysis of this scattering data with EPSR. The successful analysis of this system would then pave the way for a series of future experiments concerning subjecting this model biomolecule to various externalities, such as co-solute addition and pressure, which would allow for fundamental insight into various important biological processes. However, this is far easier said than done for a number of reasons.

As discussed in section 2.2, the inclusion of large molecules into EPSR results in a variety of potential issues. Firstly, the number of distinct interatomic correlations for a system containing J distinct atom species is equal to $J(J + 1)/2$ [321]. Large molecules therefore require exponentially larger numbers of partial structure factors to describe the system. This means that EPSR simulations require increasingly long run times, and the acquisition of a significant number of iterations to yield statistically significant results becomes more difficult. The size of the molecule also means that the scattering data that describes the structure of the molecule is located at low Q . As EPSR can only calculate correlations up to half the box size [201], to use this low Q data the box must be made exponentially larger, again resulting in long run times.

There also exist substantial difficulties relating to the inclusion of a large structured molecule like CLN025 into EPSR. The size of the molecule means that it is described by an increasingly large number of bond lengths, angles, and dihedral angles, all of which need carefully defining such that they accurately represent the structure of the CLN025 molecule [214]. Fortunately, there does exist a shortcut to manually defining these parameters. The structure of CLN025 can be downloaded from sources such as

the protein data bank and edited using software such as Chimera. Following this the structure can be directly loaded into Jmol, and included into EPSR. This therefore allows the structure to be accurately loaded into EPSR, and any minor changes to bond angles can then be made afterwards.

We must then concern ourselves with defining the reference potential that describes the molecule. Molecular dynamics simulations of CLN025 become useful here, as reference potential parameters can be adopted from previous literature [284], such as the AMBER force field. However, these potentials are often extremely detailed. This results in a large number of distinct atom types, which EPSR is not equipped to handle. Fortunately, many of these atom types are described by highly similar reference potentials, hence one can use these potential parameters to produce average potentials that are capable of describing multiple distinct atom types in the original force field as a single atom type. This then allows the number of distinct reference potentials to be substantially reduced such that they can be handled by EPSR. However, one must be extremely careful while carrying out this process so as not to oversimplify the molecule, which would result in physically unreasonable results, and ensure that the net charge on the molecule or important molecular groups remains unchanged.

Finally we must concern ourselves with the degrees of freedom that we allow the CLN025 molecule to adopt. In the case of the neutron scattering data taken at 25 °C, this is of little concern, as we can confidently model the peptide as a single folded conformation. However, if we wish to gain insight into phenomena such as temperature induced peptide unfolding, as would be present in the 60 and 90 °C data, we must allow the peptide to unfold. Within EPSR this is achieved by allowing particular chemical groups to rotate around a given chemical bond [201]. As the Monte Carlo simulation proceeds the molecule would therefore be allowed to change conformation such as to replicate protein unfolding. However, this must again be done with caution. Allowing a large number of groups to rotate results in a large potential conformational space, which again comes at the expense of long run times. One must therefore carefully select which bonds are allowed to be rotational so that protein unfolding can be accurately replicated with the minimum number of rotational groups to minimise run times.

7.2.3 The Next Generation of Structural Refinement Software: Dissolve

In short, the inclusion of a large molecule like CLN025 into EPSR is reliant on two limiting factors: the patience of the experimenter and the computing power available. In principle then, a patient experimenter requires only a powerful computer to achieve the desired outcome of a meaningful EPSR simulation of the large peptide. However, there exists an alternative to brute force computing power. This requires a total rewrite of EPSR such that it runs more efficiently and larger systems can be investigated in a shorter time frame. Fortunately, this requirement has already been noticed and acted upon by Dr Tristan Youngs with the latest version of EPSR, Dissolve [322]. This is a highly parallelised version of EPSR which means that boxes containing up to 3×10^6 atoms can be simulated, and has been designed with the aim in mind of simulating large and complex systems previously inaccessible to EPSR. It also contains a significant new feature: whereas EPSR is simply a Monte Carlo based simulation, Dissolve also features a molecular dynamics component. This means that in principle it will be much better suited to simulate large molecules such as CLN025, as much more detailed force fields taken from molecular dynamics studies can be employed.

While Dissolve will certainly play an important role in the future of investigation of liquid and amorphous samples, it is still in its infancy, and as such still contains a number of bugs which make its application to neutron scattering data difficult. Dissolve has proved to deliver consistent results to EPSR when analysing the structure of pure water and liquid benzene, however more complicated systems are still to be explored. Future work would therefore require reanalysing previous neutron scattering data through Dissolve that has been previously analysed through EPSR to ensure that the results remain consistent. Once this can be done with confidence, Dissolve could be applied to new systems inaccessible to EPSR. This means that the size limitations of molecules will gradually become less restrictive, and range of potential experiments can grow.

7.2.4 Analysis Routine

Finally, the last area that requires further work is the custom built analysis routine featured in this thesis. This routine was built to examine particular microenvironments present in aqueous solutions that are inaccessible to the ensemble averaged measurements produced by EPSR. Since its first iteration, it has grown in complexity and reliability, and has allowed us to publish results in greater detail than would have previously been possible [2, 3]. However, it can be pushed much further, as will now be discussed.

As it stands, the analysis routine is only capable of identifying hydrogen bonds that are being donated to a central atom type by a neighbouring water molecule. It does this by identifying water molecules who simultaneously satisfy two different criteria: the oxygen of the donating water molecule must be within a predefined cut off distance, and a hydrogen of the donating water molecule must be within a predefined cut off distance. This works well for atom species that accept hydrogen bonds, such as the TMAO oxygen or water oxygen discussed in this work, however it would be inapplicable to molecular groups that donate hydrogen bonds to neighboring water molecules, such as the amine groups on glycine discussed in chapter 6. It therefore follows that another definition of hydrogen bonding must be included into the analysis routine if one wishes to investigate these styles of hydrogen bonds. Such a definition could be adopted from the work of Kumar *et al.* [88], who performed a rigorous study on various definitions of hydrogen bonds in liquid water. This resulted in a hydrogen bonding definition which focused on the acceptance of a hydrogen bond by a water molecule, defined by the satisfaction of various distance and angular criteria, rather than the donation of a hydrogen bond. This would therefore be well suited for inclusion into the analysis routine.

The analysis routine would also benefit from the inclusion of periodic boundary conditions, such as those employed by EPSR. This would mean that fewer edge case molecules would have to be removed from further consideration, and more statistics could be gathered per EPSR iteration. Simple extensions already featured in the routine would also be of benefit, such as the consideration of water dipole angle distributions around previously unconsidered areas, such as TMAO methyl groups or around areas of

interest on glycine molecules. Upon successful EPSR analysis of CLN025, the analysis routine could also be expanded to consider intrapeptide hydrogen bonding, particularly in the turn region of the peptide, to investigate the effects of different externalities on the peptide in detail. Finally, as the code was written for the benefit of a single user (me), it would likely be completely unusable by another user. The code would therefore benefit from the writing of a separate code, which only requires the inclusion of a few key variables by a user, which would then write the required sections and run the code in the background, producing only the significant results. This would allow the analysis routine to be accessible to the wider community, such that other experimentalists could benefit from its existence.

7.3 Final Thoughts

The majority of the results in this thesis have been derived from EPSR analysis and neutron scattering data and as such have all been limited by the same factors: the flux limitations of neutron scattering, the intrinsic limitations of EPSR, and the computing power available for analysis [201]. While the efficiency of reactor based sources have essentially been optimised, the same cannot be said for pulsed neutron sources like the ISIS facility [185]. As these technologies continue to improve, and neutron sources generate more flux, the statistical errors and experimental run times can be reduced. This will allow for a greater number of samples to be studied in greater detail, and will require lower solute concentrations. It will also allow for the study of samples that may not be stable for the hours required for current neutron scattering experiments. New technologies such as Dissolve [322] will also allow for the analysis of larger and more complex systems, such as the CLN025 peptide discussed in this work. The future for structural refinement of neutron scattering data therefore looks bright, and before much longer (5 - 10 years) I would expect detailed structural analyses of more relevant biomolecular systems to be published, and provide us with more rigorous insight into the host of debated questions that persist in this research area.

Through this thesis we have also demonstrated the power of collaborative approaches

using multiple scientific methodologies, principally neutron scattering and NMR. This has allowed us to comment on both structure and dynamics in biologically relevant systems, and gain a more complete understanding of than either observation could offer individually. However the collaborative approach of scientific disciplines certainly should not stop here. The subjects of biology and physics were once considered to be almost opposites, however modern science makes no such distinctions. The realisation is clear that different scientific disciplines are all deeply connected and interwoven, and as such I envisage an ever greater requirement for one to be interdisciplinary in both their scientific knowledge and experimental application. I firmly believe that it is only through such approaches that we can continue to resolve the discrepancy between different scientific disciplines and answer the fundamental questions that govern our very ability to exist in the Universe.

Appendix A

Supplementary info for chapter 3

Table A.0.1: The Lennard-Jones and coulomb parameters for aqueous potassium halides. Values taken from [128].

Atomic species	ε (kJ/mol)	σ (Å)	charge (e)
O_w	0.6500	3.166	-0.8476
H_w	0.0000	0.0000	+0.4238
K^+	0.5144	2.940	+1
F^-	0.5660	3.080	-1
Cl^-	0.5660	4.191	-1
Br^-	0.5660	4.538	-1
I^-	0.5660	5.094	-1

Table A.0.2: The number of potassium halide ion pairs included in each simulation at each studied concentration and the corresponding atomic number density. Values taken from [128].

Salt species	Concentration (mol/kg H ₂ O)	Number of ion pairs	Atomic Num- ber Density (atoms/Å ³)
KF	0.6	60	0.10020
	1.2	120	0.10040
	2.4	240	0.10020
KCl	0.6	60	0.09915
	1.2	120	0.09800
	2.4	240	0.09580
KBr	0.6	60	0.09870
	1.2	120	0.09720
	2.4	240	0.09419
KBr	0.6	60	0.09803
	1.2	120	0.09583
	2.4	240	0.09216

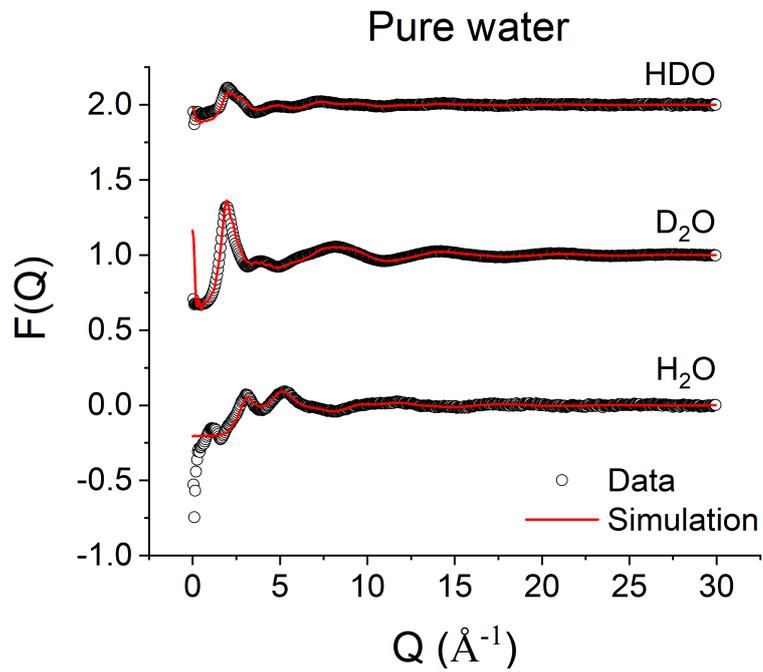


Figure A.0.1: Supplied diffraction data $D(Q)$ taken using NIMROD instrument (hollow circles) and total simulated structure factor $F(Q)$ (red line) for three isotopic variants of pure water. Each dataset vertically shifted by 1 for clarity.

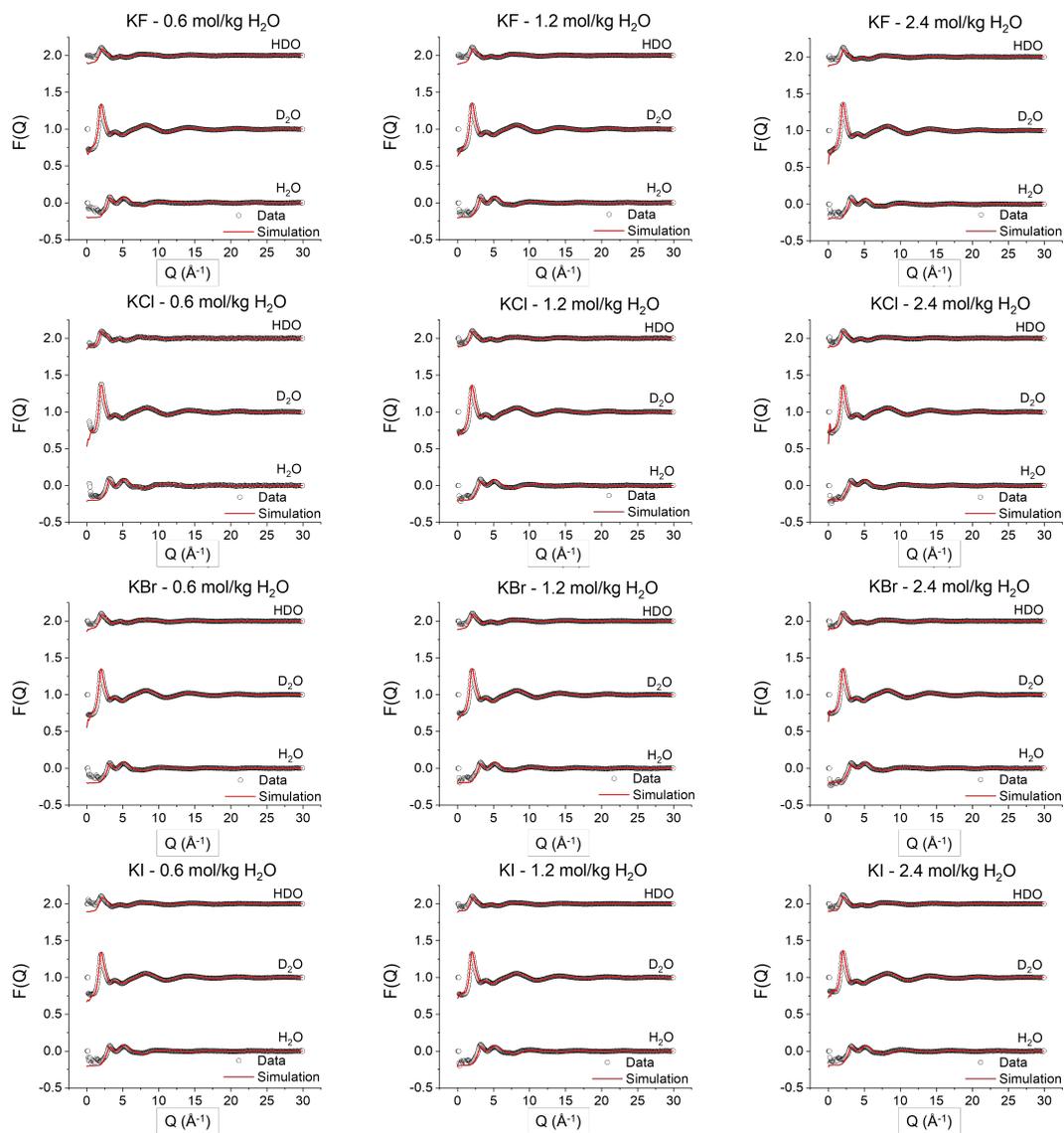


Figure A.0.2: Supplied diffraction data $D(Q)$ from a previous study taken on SANDALS instrument [128] (hollow circles) and total simulated structure factor $F(Q)$ (red line) for aqueous potassium halides in three isotopic variants water. Each dataset vertically shifted by 1 for clarity.

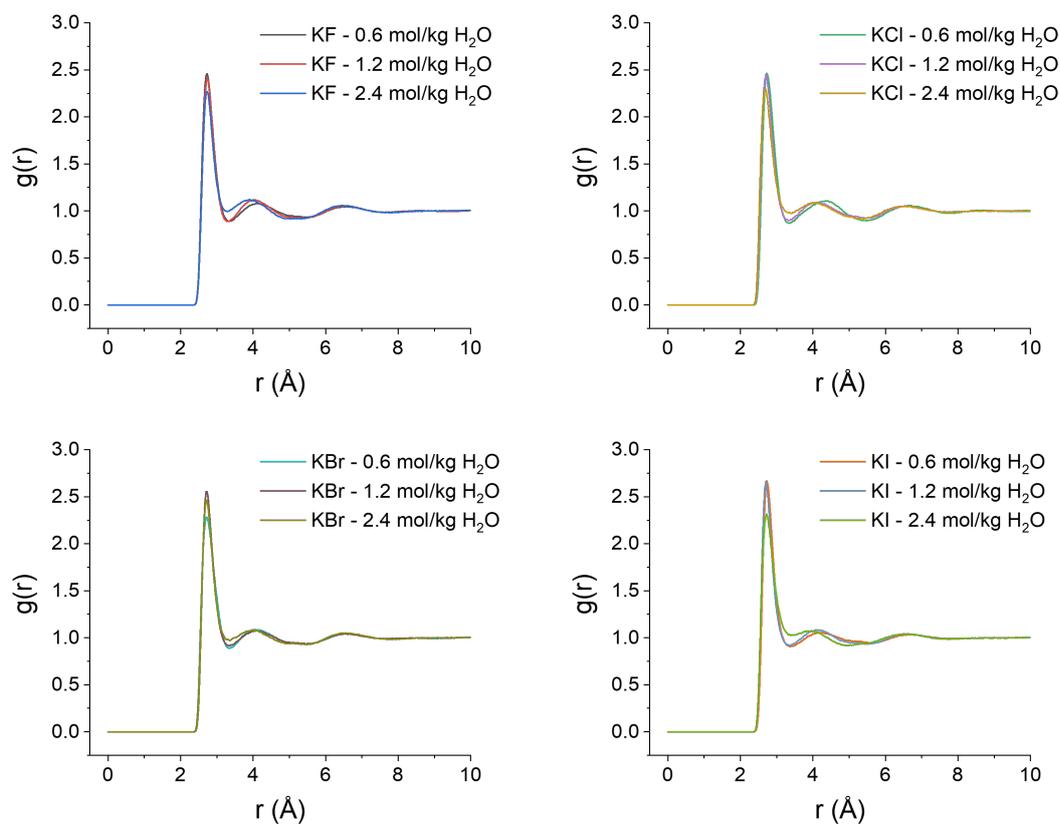


Figure A.0.3: Water oxygen - water oxygen RDFs for pure water and aqueous KF, KCl, KBr, and KI at 0.6, 1.2, and 2.4 mol/kg H₂O.

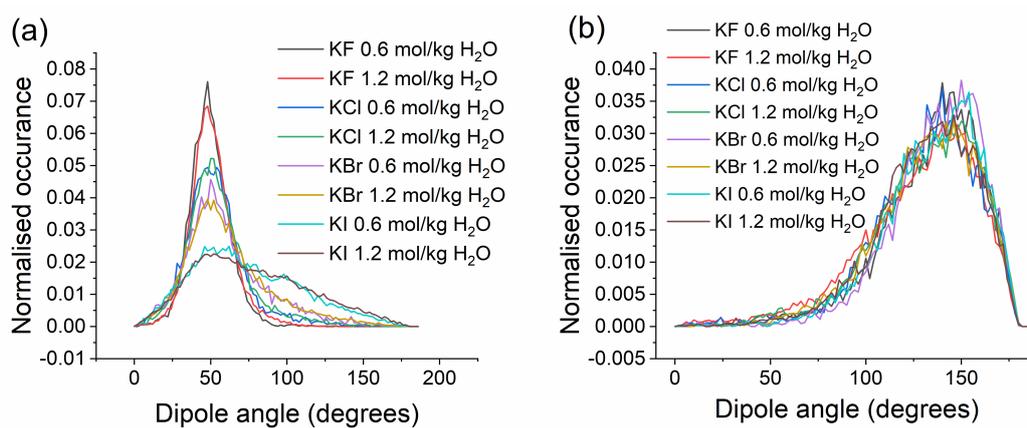


Figure A.0.4: Distribution of dipole angles for water molecules hydrating a central halide anion (a) or potassium cation (b) within its first hydration shell. The dipole angle is defined as the angle between the vector pointing from the ion to the hydrating water oxygen and the vector corresponding to the hydrating water molecular dipole. Values normalized to the total number of data points. Data shown for concentrations of 0.6 and 1.2 mol/kg H₂O samples.

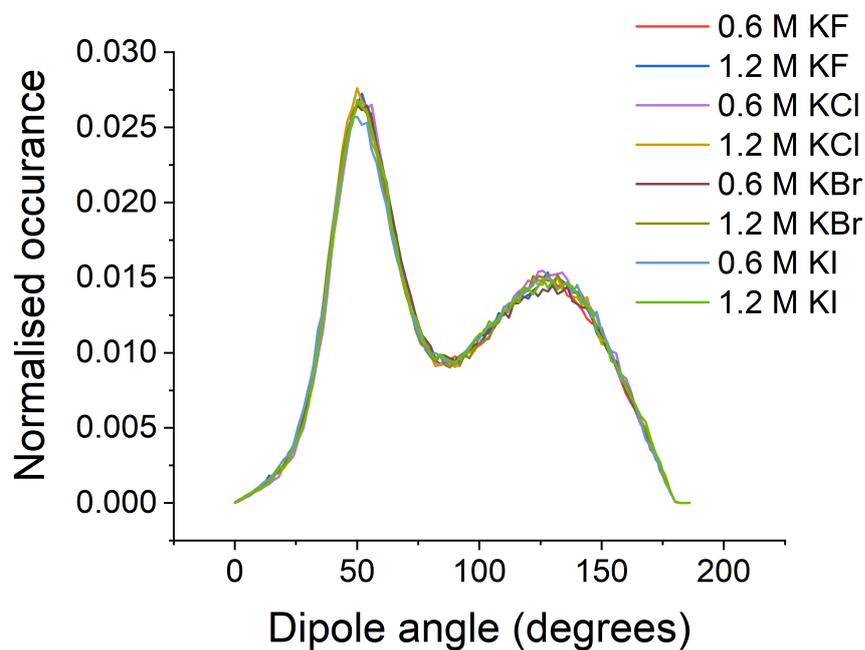


Figure A.0.5: Distribution of dipole angles for water molecules hydrating a central water molecule within its first hydration shell. The dipole angle is defined as the angle between the vector pointing from the central water oxygen to the hydrating water oxygen and the vector corresponding to the hydrating water molecular dipole. Values normalized to the total number of data points. Data shown for concentrations of 0.6 and 1.2 mol/kg H₂O samples.

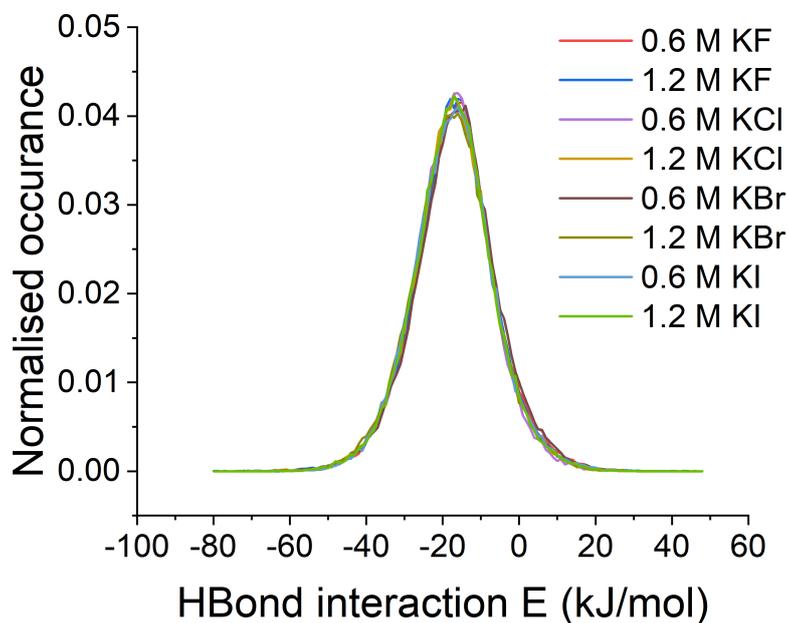


Figure A.0.6: Bulk water - water hydrogen bond interaction energy distribution each of the four studied potassium halide salts at 0.6 and 1.2 mol/kg H_2O .

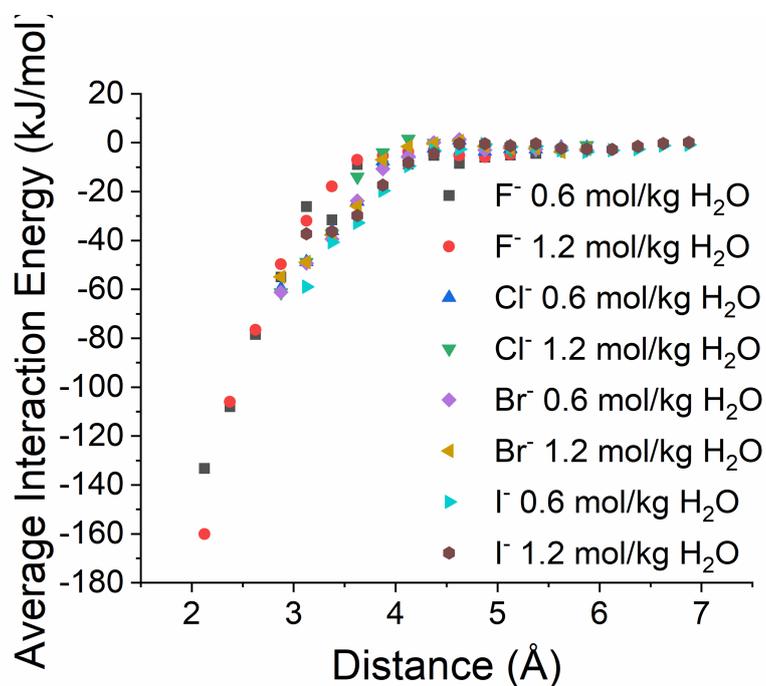


Figure A.0.7: Average enthalpic interaction strength, ΔH between a central potassium halide anion and its surrounding water molecules at a given distance.

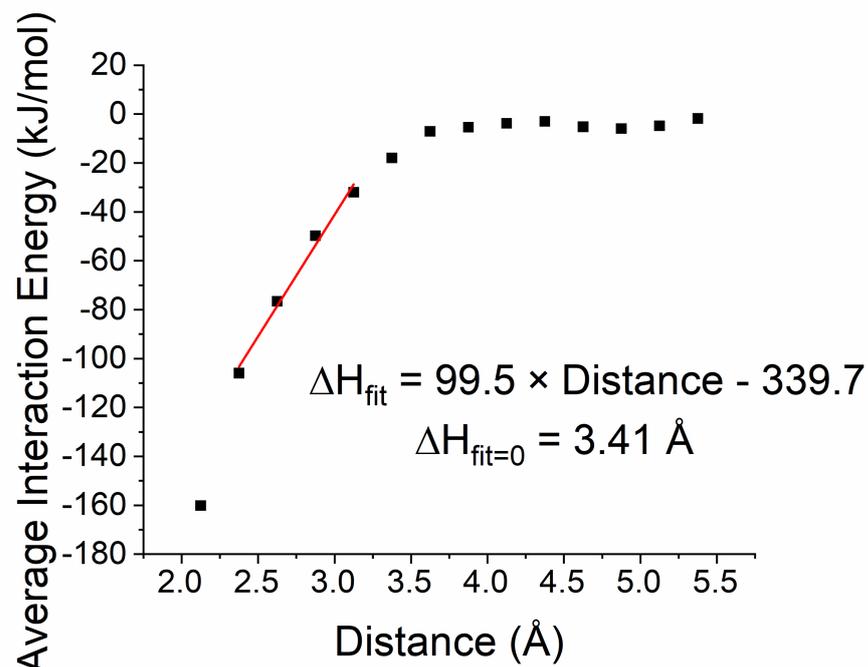


Figure A.0.8: Performing a linear fit to the low distance region of the average enthalpic interaction strength, ΔH between a central potassium halide anion and its surrounding water molecules at a given distance data to determine the point at which $\Delta H = 0$.

Appendix B

Supplementary info for chapter 4

Table B.0.1: The Lennard-Jones and coulomb parameters for aqueous TMAO. Values taken from [62] and [147]*.

Atomic species	ε (kJ/mol)	σ (Å)	charge (e)
O _w	0.6500	3.166	-0.8476
H _w	0.0000	0.0000	+0.4238
C _T	0.3900	3.7000	-0.2600
H _T	0.0650	1.8000	0.1100
N _T	0.7110	3.2500	0.4400
O _T	0.5850	3.0800	-0.6500
C _T *	0.2830	3.7070	-0.2609
H _T *	0.7750	2.1300	0.1166
N _T *	0.8374	2.9260	0.5932
O _T *	0.6389	3.2660	-0.8599

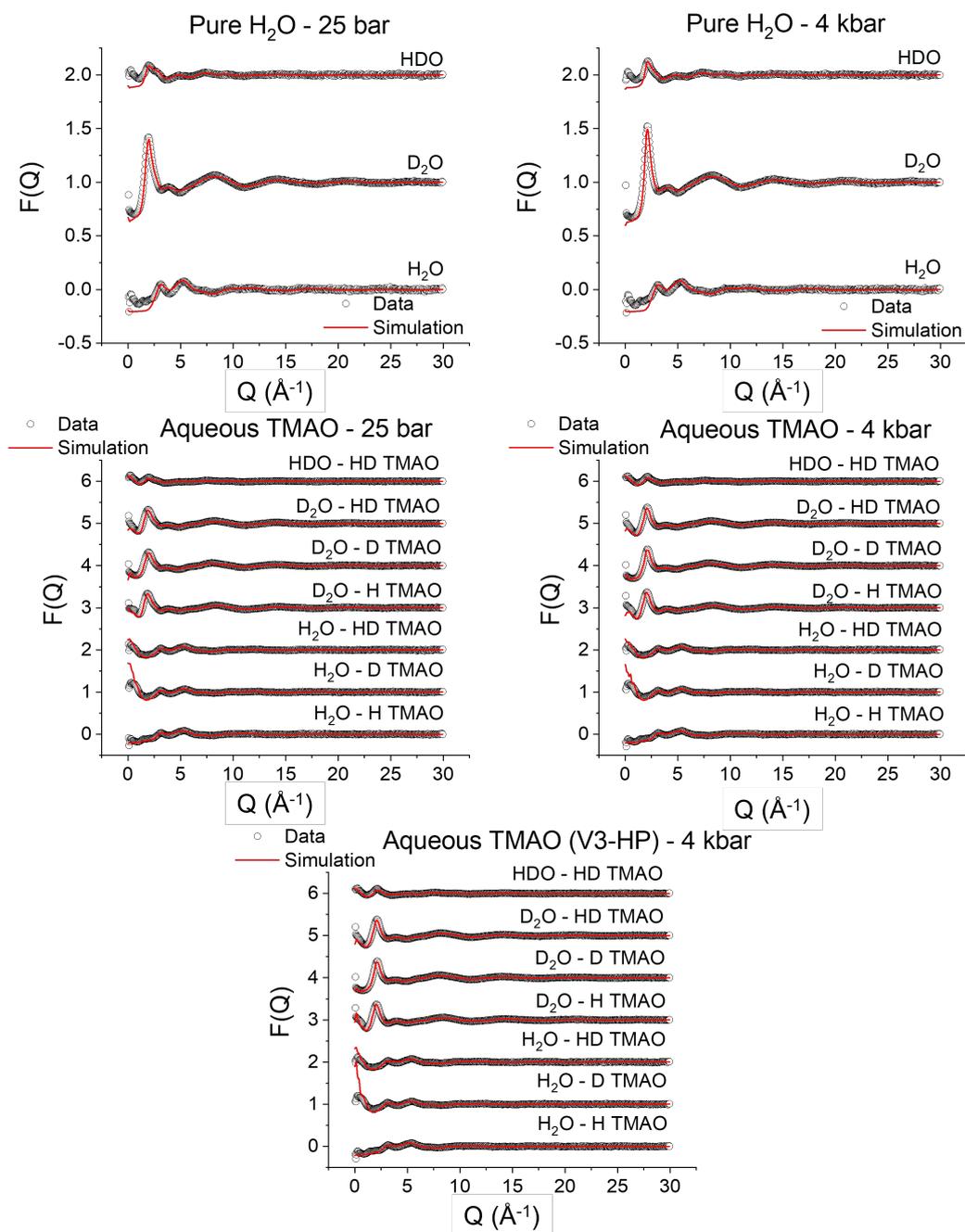


Figure B.0.1: Supplied diffraction data $D(Q)$ taken using NIMROD instrument (hollow circles) and total simulated structure factor $F(Q)$ (red line) for three isotopic variants of pure water and seven isotopic variants of aqueous TMAO at 25 bar and 4 kbar. EPSR was performed using the potentials published by Meersman *et al.* [62] for all four samples and performed using the potential published by Hölzl *et al.* for aqueous TMAO at 4 kbar [147]. Each dataset vertically shifted by 1 for clarity.

Appendix C

Supplementary info for chapter 5

Table C.0.1: The Lennard-Jones and coulomb parameters for aqueous TMAO and $\text{Mg}(\text{ClO}_4)_2$. Values taken from [62] and [133].

Atomic species	ε (kJ/mol)	σ (Å)	charge (e)
O_w	0.6500	3.166	-0.8476
H_w	0.0000	0.0000	+0.4238
C_T	0.3900	3.7000	-0.2600
H_T	0.0650	1.8000	0.1100
N_T	0.7110	3.2500	0.4400
O_T	0.5850	3.0800	-0.6500
Mg	0.4593	0.9000	2.0000
Cl_p	0.5660	4.1900	2.3904
O_p	0.6500	3.1660	-0.8476

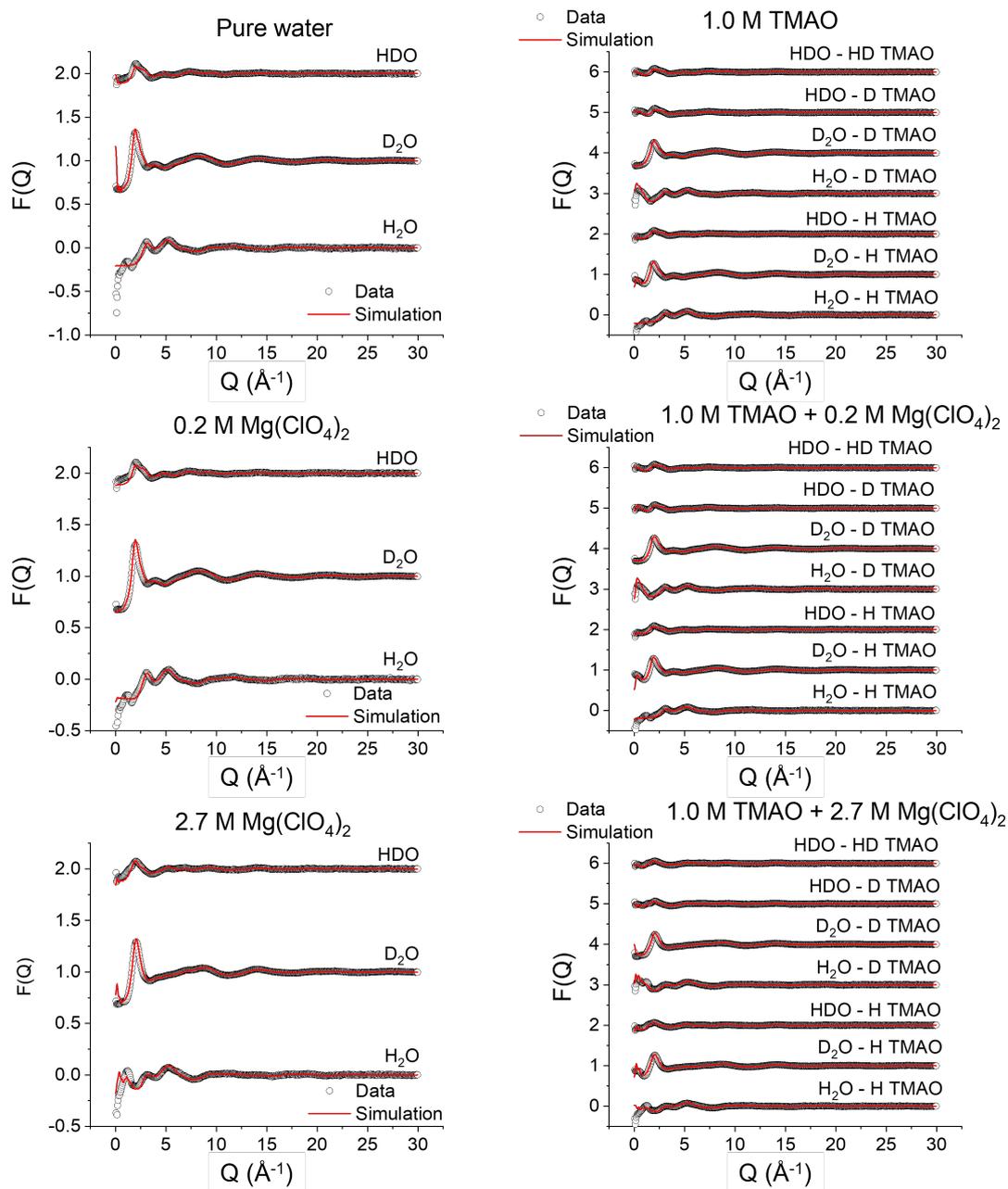


Figure C.0.1: Supplied diffraction data $D(Q)$ taken using NIMROD instrument (hollow circles) and total simulated structure factor $F(Q)$ (red line) for three isotopic variants of pure water and aqueous $\text{Mg}(\text{ClO}_4)_2$ and seven isotopic variants of aqueous TMAO with and without $\text{Mg}(\text{ClO}_4)_2$. EPSR was performed using the potentials published by Meersman *et al.* [62] and Lenton *et al.*. Each dataset vertically shifted by 1 for clarity.

Appendix D

Supplementary info for chapter 6

Table D.0.1: The Lennard-Jones and coulomb parameters for aqueous glycine and $\text{Mg}(\text{ClO}_4)_2$. Values for water and $\text{Mg}(\text{ClO}_4)_2$ taken from [133] and glycine parameters estimated from previous neutron diffraction studies on aqueous amino acids and peptides [137, 138, 140, 289, 290, 297–299].

Atomic species	ε (kJ/mol)	σ (Å)	charge (e)
O_w	0.6500	3.166	-0.8476
H_w	0.0000	0.0000	+0.4238
C_{bk}	0.4142	3.8000	0.09000
C_c	0.42932	3.7500	0.7000
H_{bk}	0.0000	0.0000	0.0600
N_a	0.71128	3.2500	-0.3000
O_c	0.87864	2.9600	-0.8000
H_a	0.0000	0.0000	0.3300
Mg	0.4593	0.9000	2.0000
Cl_p	0.5660	4.1900	2.3904
O_p	0.6500	3.1660	-0.8476

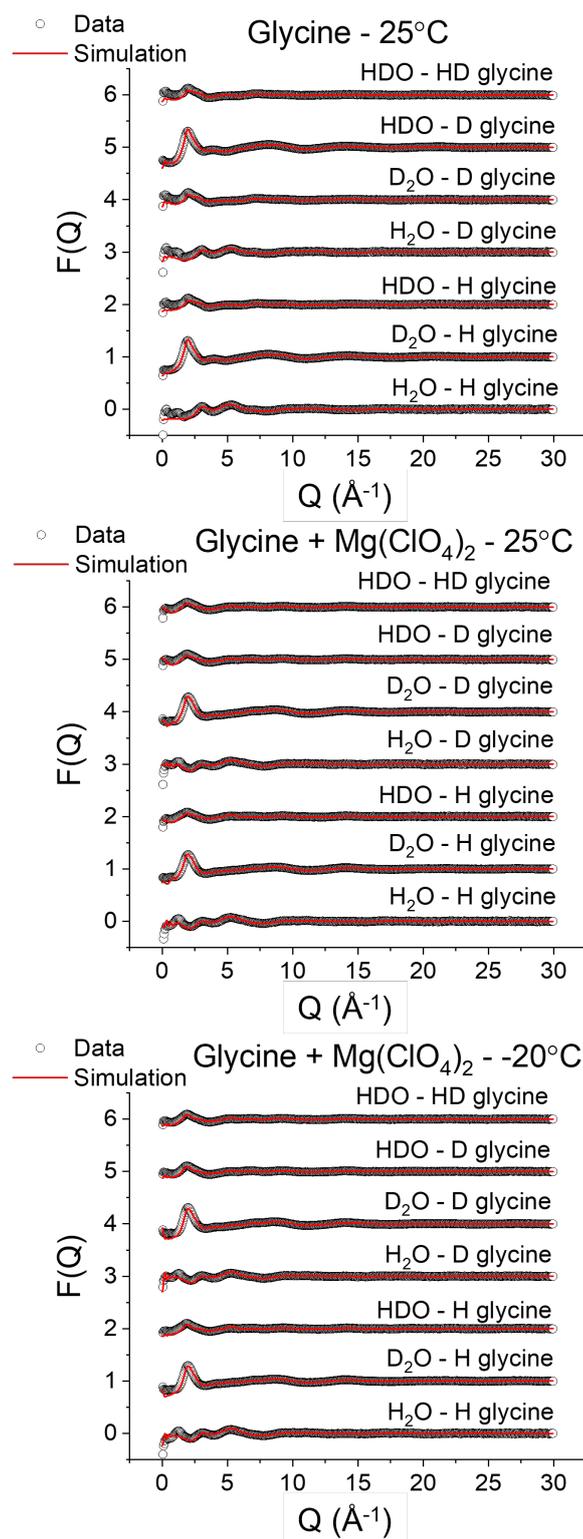


Figure D.0.1: Supplied diffraction data $D(Q)$ taken using NIMROD instrument (hollow circles) and total simulated structure factor $F(Q)$ (red line) for seven isotopic variants of aqueous glycine with and without $\text{Mg}(\text{ClO}_4)_2$. Each dataset vertically shifted by 1 for clarity.

Bibliography

- [1] Harrison Laurent, Alan Soper, and Lorna Dougan. Biomolecular self-assembly under extreme Martian mimetic conditions. *Molecular Physics*, 117(22):3398–3407, 2019.
- [2] Harrison Laurent, Daniel Baker, Alan Soper, Michael Ries, and Lorna Dougan. Solute Specific Perturbations to Water Structure and Dynamics in Tertiary Aqueous Solution. *Journal of Physical Chemistry B*, 124:10983–10993, 2020.
- [3] Harrison Laurent, Alan K. Soper, and Lorna Dougan. Trimethylamine: N -oxide (TMAO) resists the compression of water structure by magnesium perchlorate: Terrestrial kosmotrope vs. Martian chaotrope. *Physical Chemistry Chemical Physics*, 22(9):4924–4937, 2020.
- [4] Benjamin S. Hanson, Christa P. Brown, Harrison Laurent, Matt D.G. Hughes, and Lorna Dougan. Hierarchical biomechanics: student engagement activities with a focus on biological physics. *Physics Education*, 55(2), 2020.
- [5] Harrison Laurent, Daniel Baker, Alan K. Soper, Michael E. Ries, and Lorna Dougan.
- [6] Kara Rogers, Susan Heyner Joshi, and Edna R. Green. *Biology*, 2019.
- [7] Edward B. Ruth. Replication, Transcription, and Translation. *The American Biology Teacher*, 46(8):470–472, nov 1984.
- [8] Philip Ball. Water is an active matrix of life for cell and molecular biology. *Proceedings of the National Academy of Sciences*, 114(51):13327–13335, 2017.
- [9] Philip Ball. *H₂O A Biography of Water*. Weidenfeld and Nicolson, London, first edition, 2000.
- [10] Marie Claire Bellissent-Funel, Ali Hassanali, Martina Havenith, Richard Henchman, Peter Pohl, Fabio Sterpone, David Van Der Spoel, Yao Xu, and Angel E. Garcia. Water Determines the Structure and Dynamics of Proteins. *Chemical Reviews*, 116(13):7673–7697, 2016.
- [11] Nancy Merino, Heidi S. Aronson, Diana P. Bojanova, Jayme Feyhl-Buska, Michael L. Wong, Shu Zhang, and Donato Giovannelli. Living at the extremes: Extremophiles and the limits of life in a planetary context. *Frontiers in Microbiology*, 10(MAR), 2019.
- [12] James A. Coker. Extremophiles and biotechnology: Current uses and prospects. *F1000Research*, 5(396), 2016.

- [13] Prasanti Babu, Anuj K. Chandel, and Om V. Singh. *Extremophiles and Their Applications in Medical Processes*. Number October. 2015.
- [14] Sepideh Parvizpour, Nurulfarhana Hussin, Mohd Shahir Shamsir, and Jafar Razmara. Psychrophilic enzymes: structural adaptation, pharmaceutical and industrial applications. *Applied Microbiology and Biotechnology*, pages 899–907, 2021.
- [15] Angela Beckett and David Gervais. What makes a good new therapeutic l-asparaginase? *World Journal of Microbiology and Biotechnology*, 35(10):1–13, 2019.
- [16] Shiladitya DasSarma and Priya DasSarma. Halophiles and their enzymes: negativity put to good use. *Current Opinion in Microbiology*, 25(12):120–126, jun 2015.
- [17] Christopher J. Reed, Hunter Lewis, Eric Trejo, Vern Winston, and Caryn Evilia. Protein adaptations in archaeal extremophiles. *Archaea*, 2013:14, 2013.
- [18] A. V. Finkelstein and O. B. Ptitsyn. *Protein physics: a course of lectures*. Academic Press, London, first edition, 2002.
- [19] Alain J. Cozzzone. Proteins: Fundamental Chemical Properties. In *Encyclopedia of Life Sciences*. John Wiley and Sons, Ltd, Chichester, UK, apr 2010.
- [20] D. A. Waller. Protein folding. *Biochemical Education*, 21(1):50, jan 1993.
- [21] Amar Thaker, Luqmanal Sirajudeen, Chad R. Simmons, and Brent L. Nannenga. Structure guided identification of a peptide for bioenabled gold nanoparticle synthesis. *Biotechnology and Bioengineering*, page bit.27927, sep 2021.
- [22] Linus Pauling. *The Nature of the Chemical Bond*. Cornell University Press, New York, 3rd edition, 1960.
- [23] Sebastian Busch, Chrystal D. Bruce, Christina Redfield, Christian D. Lorenz, and Sylvia E. McLain. Water mediation is essential to nucleation of β -turn formation in peptide folding motifs. *Angewandte Chemie - International Edition*, 52(49):13091–13095, 2013.
- [24] Alvaro H. Crevenna, Nikolaus Naredi-Rainer, Don C. Lamb, Roland Wedlich-Söldner, and Joachim Dzubiella. Effects of Hofmeister ions on the α -helical structure of proteins. *Biophysical Journal*, 102(4):907–915, 2012.
- [25] Qiang Zhang, Zhijun Pan, Lu Zhang, Ruiting Zhang, Zhening Chen, Tan Jin, Tianmin Wu, Xian Chen, and Wei Zhuang. Ion effect on the dynamics of water hydrogen bonding network: A theoretical and computational spectroscopy point of view, 2018.
- [26] Damien Laage and James T. Hynes. On the molecular mechanism of water reorientation. *Journal of Physical Chemistry B*, 112(45):14230–14242, 2008.
- [27] N. Galamba. On the effects of temperature, pressure, and dissolved salts on the hydrogen-bond network of water. *Journal of Physical Chemistry B*, 117(2):589–601, 2013.
- [28] Paul A. Tipler and Gene Mosca. *Physics for Scientists and Engineers: With Modern Physics*. W. H. Freeman and Company, New York, sixth edition, 2008.

- [29] Morten E. Allentoft, Matthew Collins, David Harker, James Haile, Charlotte L. Oskam, Marie L. Hale, Paula F. Campos, Jose A. Samaniego, Thomas P. M. Gilbert, Eske Willerslev, Guojie Zhang, R. Paul Scofield, Richard N. Holdaway, and Michael Bunce. The half-life of DNA in bone: Measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences*, 279(1748):4724–4733, 2012.
- [30] Jacob Israelachvili. *Intermolecular and Surface Forces*. Academic Press, London, 3rd edition, 2011.
- [31] Emiliano Brini, Christopher J. Fennell, Marivi Fernandez-Serra, Barbara Hribar-Lee, Miha Lukšič, and Ken A. Dill. How Water’s Properties Are Encoded in Its Molecular Structure and Energies. *Chemical Reviews*, 117(19):12385–12414, 2017.
- [32] E. E. Meyer, K. J. Rosenberg, and J. Israelachvili. Recent progress in understanding hydrophobic interactions. *Proceedings of the National Academy of Sciences*, 103(43):15739–15746, oct 2006.
- [33] J. A. V. Butler. The energy and entropy of hydration of organic compounds. *Transactions of the Faraday Society*, 33:229, 1937.
- [34] Robert L. Baldwin. The new view of hydrophobic free energy. *FEBS Letters*, 587(8):1062–1066, apr 2013.
- [35] Michael D. Tissandier, Kenneth A. Cowen, Wan Yong Feng, Ellen Gundlach, Michael H. Cohen, Alan D. Earhart, James V. Coe, and Thomas R. Tuttle. The proton’s absolute aqueous enthalpy and Gibbs free energy of solvation from cluster-ion solvation data. *Journal of Physical Chemistry A*, 102(40):7787–7794, 1998.
- [36] Kazumasa Sakurai, Masanori Yagi, Tsuyoshi Konuma, Satoshi Takahashi, Chiaki Nishimura, and Yuji Goto. Non-Native α -Helices in the Initial Folding Intermediate Facilitate the Ordered Assembly of the β -Barrel in β -Lactoglobulin. *Biochemistry*, 56(36):4799–4807, 2017.
- [37] Joachim Seelig and Hans-Joachim Schönfeld. Thermal protein unfolding by differential scanning calorimetry and circular dichroism spectroscopy Two-state model versus sequential unfolding. *Quarterly Reviews of Biophysics*, 49, 2016.
- [38] Song Ho Chong and Sihyun Ham. Folding Free Energy Landscape of Ordered and Intrinsically Disordered Proteins. *Scientific Reports*, 9(1):1–9, 2019.
- [39] Georges Feller. Psychrophilic Enzymes: From Folding to Function and Biotechnology. *Scientifica*, 2013:1–28, 2013.
- [40] Georges Feller. Protein stability and enzyme activity at extreme biological temperatures. *Journal of Physics: Condensed Matter*, 22(32):323101, aug 2010.
- [41] Pieter De Maayer, Dominique Anderson, Craig Cary, and Don A. Cowan. Some like it cold: Understanding the survival strategies of psychrophiles. *EMBO Reports*, 15(5):508–517, 2014.
- [42] R. Cavicchioli, T. Charlton, H. Ertan, S. Mohd Omar, K. S. Siddiqui, and T. J. Williams. Biotechnological uses of enzymes from psychrophiles, 2011.
- [43] Franz Hofmeister. Zur Lehre von der Wirkung der Salze. *Archiv für experimentelle Pathologie und Pharmakologie*, 24:247–260, 1888.

- [44] Nadine Schwierz, Dominik Horinek, Uri Sivan, and Roland R. Netz. Reversed Hofmeister series - The rule rather than the exception. *Current Opinion in Colloid and Interface Science*, 23:10–18, 2016.
- [45] Pavel Jungwirth and Paul S. Cremer. Beyond Hofmeister, 2014.
- [46] Taylor P. Light, Karen M. Corbett, Michael A. Metrick, and Gina Macdonald. Hofmeister Ion-Induced Changes in Water Structure Correlate with Changes in Solvation of an Aggregated Protein Complex. *Langmuir*, 32:12, 2016.
- [47] Halil I. Okur, Jana Hladílková, Kelvin B. Rembert, Younhee Cho, Jan Heyda, Joachim Dzubiella, Paul S. Cremer, and Pavel Jungwirth. Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions, 2017.
- [48] Kim D. Collins and Michael W. Washabaugh. The Hofmeister effect and the behaviour of water at interfaces. *Quarterly Reviews of Biophysics*, 18(4):323–422, 1985.
- [49] Ellen E. Bruce, Halil I. Okur, Sina Stegmaier, Chad I. Drexler, Bradley A. Rogers, Nico F. A. Van Der Vegt, Sylvie Roke, and Paul S. Cremer. Molecular mechanism for the interactions of hofmeister cations with macromolecules in aqueous solution. *Journal of the American Chemical Society*, 142(45):19094–19100, 2020.
- [50] Nathaniel V. Nucci and Jane M. Vanderkooi. Effects of salts of the Hofmeister series on the hydrogen bond network of water. *Journal of Molecular Liquids*, 143(2-3):160–170, 2008.
- [51] Ashley K. Borkowski, Zeke A. Piskulich, and Ward H. Thompson. Examining the Hofmeister Series through Activation Energies: Water Diffusion in Aqueous Alkali-Halide Solutions. *Journal of Physical Chemistry B*, 125(1):350–359, 2021.
- [52] Virginia Mazzini, Guangming Liu, and Vincent S. J. Craig. Probing the Hofmeister series beyond water: Specific-ion effects in non-aqueous solvents. *Journal of Chemical Physics*, 148(22):222805, 2018.
- [53] Philip Ball and John E. Hallsworth. Water structure and chaotropicity: Their uses, abuses and biological implications. *Physical Chemistry Chemical Physics*, 17(13):8297–8305, 2015.
- [54] Nina Gunde-Cimerman, Ana Plemenitaš, and Aharon Oren. Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS Microbiology Reviews*, 42(3):353–375, 2018.
- [55] Ken A. Dill, Thomas M. Truskett, Vojko Vlachy, and Barbara Hribar-Lee. Modeling Water, the Hydrophobic Effect. *Annual Review of Biophysics and Biomolecular Structure*, 34:173–199, 2005.
- [56] Y. S. Kim, B. D. Lim, J. E. Lee, and C. S. Lee. Solubilities of Carbon Dioxide, Methane, and Ethane in Sodium Chloride Solution Containing Gas Hydrate. *Journal of Chemical Engineering Data*, 53(6):1351–1354, jun 2008.
- [57] Somayeh Mojtabavi, Nasrin Samadi, and Mohammad Ali Faramarzi. Osmolyte-induced folding and stability of proteins: Concepts and characterization. *Iranian Journal of Pharmaceutical Research*, 18(Special Issue):13–30, 2019.

- [58] Adam Bownik and Zofia Stepniewska. Ectoine as a promising protective agent in humans and animals. *Arhiv za Higijenu Rada i Toksikologiju*, 67(4):260–265, 2016.
- [59] M. Lever, J. W. Blunt, and R. G. A. R. Maclagan. Some ways of looking at compensatory kosmotropes and different water environments. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 130(3):471–486, 2001.
- [60] Michel W. Jaworek, Vitor Schuabb, and Roland Winter. The effects of glycine, TMAO and osmolyte mixtures on the pressure dependent enzymatic activity of α -chymotrypsin. *Physical Chemistry Chemical Physics*, 20(3):1347–1354, 2018.
- [61] Paul H. Yancey. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses, 2005.
- [62] Filip Meersman, Daniel Bowron, Alan K. Soper, and Michel H. J. Koch. Counteraction of urea by trimethylamine N-oxide is due to direct interaction. *Biophysical Journal*, 97(9):2559–2566, 2009.
- [63] Filip Meersman, Daniel Bowron, Alan K. Soper, and Michel H. J. Koch. An X-ray and neutron scattering study of the equilibrium between trimethylamine N-oxide and urea in aqueous solution. *Physical Chemistry Chemical Physics*, 13(30):13765–13771, 2011.
- [64] Sho Imoto, Harald Forbert, and Dominik Marx. Water structure and solvation of osmolytes at high hydrstatic pressure: pure water and TMAO solutions at 10 kbar versus 1 bar. *Phys. Chem. Chem. Phys.*, 17:24224, 2015.
- [65] Guillaume Stirnemann, Elise Duboué-Dijon, and Damien Laage. Ab Initio Simulations of Water Dynamics in Aqueous TMAO Solutions: Temperature and Concentration Effects. *Journal of Physical Chemistry B*, 121(49):11189–11197, 2017.
- [66] Tatsuhiko Ohto, Johannes Hunger, Ellen H. G. Backus, Wataru Mizukami, Mischa Bonn, and Yuki Nagata. Trimethylamine-N-oxide: Its hydration structure, surface activity, and biological function, viewed by vibrational spectroscopy and molecular dynamics simulations, 2017.
- [67] Apramita Chand and Snehasis Chowdhuri. A comparative study of hydrogen bonding structure and dynamics in aqueous urea solution of amides with varying hydrophobicity: Effect of addition of trimethylamine N-oxide (TMAO). *Journal of Molecular Liquids*, 242:70–81, 2017.
- [68] Seishi Shimizu and Paul E. Smith. How Osmolytes Counteract Pressure Denaturation on a Molecular Scale. *ChemPhysChem*, 18(16):2243–2249, 2017.
- [69] Pritam Ganguly, Pablo Boserman, Nico F. A. Van Der Vegt, and Joan Emma Shea. Trimethylamine N-oxide Counteracts Urea Denaturation by Inhibiting Protein-Urea Preferential Interaction. *Journal of the American Chemical Society*, 140(1):483–492, 2018.
- [70] Rahul Sarma and Sandip Paul. The effect of aqueous solutions of trimethylamine-N-oxide on pressure induced modifications of hydrophobic interactions. *The Journal of Chemical Physics*, 137:094502, 2012.
- [71] Pritam Ganguly, Timir Hajari, Joan Emma Shea, and Nico F. A. Van Der Vegt. Mutual exclusion of urea and trimethylamine n -oxide from amino acids in mixed solvent environment. *Journal of Physical Chemistry Letters*, 6(4):581–585, 2015.

- [72] Marina V. Fedotova, Sergey E. Kruchinin, and Gennady N. Chuev. Hydration structure of osmolyte TMAO: concentration/pressure-induced response. *New Journal of Chemistry*, 41(3):1219–1228, 2017.
- [73] E. D. Kadtsyn, A. V. Anikeenko, and N. N. Medvedev. Structure of Aqueous Solutions of Trimethylaminoxide, Urea, and Their Mixture. *Journal of Structural Chemistry*, 59(2):347–354, 2018.
- [74] Toshiko Ichiye. *Enzymes from piezophiles*, 2018.
- [75] Lin Zhao, Kai Ma, and Zi Yang. Changes of water hydrogen bond network with different externalities. *International Journal of Molecular Sciences*, 16(4):8454–8489, apr 2015.
- [76] David J. Wilton, Richard B. Tunnicliffe, Yuji O. Kamatari, Kazuyuki Akasaka, and Mike P. Williamson. Pressure-induced changes in the solution structure of the gb1 domain of protein g. *Proteins: structure, function, and bioinformatics*, 71(3):1432–1440, May 2008.
- [77] Huang, Tran, Rodgers, Bartlett, Hemley, and Ichiye. A molecular perspective on the limits of life: Enzymes under pressure. *Condensed Matter Physics*, 19(2):22801, mar 2016.
- [78] Catherine A. Royer. Revisiting volume changes in pressure-induced protein unfolding. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1595(1-2):201–209, mar 2002.
- [79] Julien Roche, Catherine A. Royer, and Christian Roumestand. Monitoring protein folding through high pressure NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 102-103:15–31, 2017.
- [80] Jiasong Fang, Li Zhang, and Dennis A. Bazylinski. Deep-sea piezosphere and piezophiles: geomicrobiology and biogeochemistry. *Trends in Microbiology*, 18(9):413–422, 2010.
- [81] Julien Roche and Catherine A. Royer. Lessons from pressure denaturation of proteins, 2018.
- [82] Deana Desmarais Martin, Douglas H. Bartlett, and Mary F. Roberts. Solute accumulation in the deep-sea bacterium *Photobacterium profundum*. *Extremophiles*, 6(6):507–514, 2002.
- [83] Christina M. Papini, Pranav P. Pandharipande, Catherine A. Royer, and George I. Makhatadze. Putting the Piezolyte Hypothesis under Pressure. *Biophysical Journal*, 113(5):974–977, 2017.
- [84] Satyajit Patra, Christian Anders, Nelli Erwin, and Roland Winter. Osmolyte Effects on the Conformational Dynamics of a DNA Hairpin at Ambient and Extreme Environmental Conditions. *Angewandte Chemie - International Edition*, 56(18):5045–5049, 2017.
- [85] Roland Hermann Alfons Winter, Jim-Marcel Knop, Satyajit Patra, Harish Balasubramanian, and Catherine Royer. The Deep Sea Osmolyte TMAO and Macromolecular Crowders Rescue the Antiparallel Conformation of the Human Telomeric G-Quadruplex from Urea and Pressure Stress. *Chem. Eur. J.*, 24:14346–14351, 2018.

- [86] Paul H. Yancey, Mackenzie E. Gerringer, Jeffrey C. Drazen, Ashley A. Rowden, and Alan Jamieson. Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. *Proceedings of the National Academy of Sciences of the United States of America*, 111(12):4461–4465, 2014.
- [87] G. Engel and H. Hertz. On the Negative Hydration. A Nuclear Magnetic Relaxation Study. *Berichte der Bunsengesellschaft für physikalische Chemie*, 72(7):808–834, 1968.
- [88] R. Kumar, J. R. Schmidt, and J. L. Skinner. Hydrogen bonding definitions and dynamics in liquid water. *Journal of Chemical Physics*, 126(20):204107, 2007.
- [89] A. K. Soper. The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure? *ISRN Physical Chemistry*, 2013:1–67, 2013.
- [90] Martin Chaplin. *Water Structure and Science*, 2019.
- [91] Pekka Mark and Lennart Nilsson. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *Journal of Physical Chemistry A*, 105(43):9954–9960, 2001.
- [92] Masakazu Matsumoto. Relevance of hydrogen bond definitions in liquid water. *Journal of Chemical Physics*, 126(5):054503, 2007.
- [93] Dorota Swiatla-Wojcik. Evaluation of the criteria of hydrogen bonding in highly associated liquids. *Chemical Physics*, 342(1-3):260–266, 2007.
- [94] A. D. Hammerich and V. Buch. An alternative near-neighbor definition of hydrogen bonding in water. *Journal of Chemical Physics*, 128(11):111101, 2008.
- [95] B. M. Auer and J. L. Skinner. Water: Hydrogen bonding and vibrational spectroscopy, in the bulk liquid and at the liquid/vapor interface. *Chemical Physics Letters*, 470(1-3):13–20, 2009.
- [96] Richard H. Henchman and Stuart J. Cockram. Water’s non-tetrahedral side. *Faraday Discussions*, 167:529–550, 2013.
- [97] Piero Gasparotto and Michele Ceriotti. Recognizing molecular patterns by machine learning: An agnostic structural definition of the hydrogen bond. *Journal of Chemical Physics*, 141(17):174110, 2014.
- [98] Abdullah Ozkanlar, Tiecheng Zhou, and Aurora E. Clark. Towards a unified description of the hydrogen bond network of liquid water: A dynamics based approach. *Journal of Chemical Physics*, 141(21):214107, 2014.
- [99] Diego Prada-Gracia, Roman Shevchuk, and Francesco Rao. The quest for self-consistency in hydrogen bond definitions. *Journal of Chemical Physics*, 139(8):084501, 2013.
- [100] Katrin Amann-Winkel, Marie Claire Bellissent-Funel, Livia E. Bove, Thomas Lortting, Anders Nilsson, Alessandro Paciaroni, Daniel Schlessinger, and Lawrie Skinner. X-ray and Neutron Scattering of Water. *Chemical Reviews*, 116(13):7570–7589, 2016.

- [101] A. K. Soper. The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure? *ISRN Physical Chemistry*, 2013:1–67, 2013.
- [102] Congcong Huang, K. T. Wikfeldt, D. Nordlund, U. Bergmann, T. McQueen, J. Sellberg, L. G. M. Pettersson, and A. Nilsson. Wide-angle X-ray diffraction and molecular dynamics study of medium-range order in ambient and hot water. *Physical Chemistry Chemical Physics*, 13(44):19997–20007, 2011.
- [103] Philip Ball and John E. Hallsworth. Water structure and chaotropy: their uses, abuses and biological implications. *Physical Chemistry Chemical Physics*, 17(13):8297–8305, 2015.
- [104] W. C. Röntgen. Über die Constitution des flüssigen Wassers. *Annalen der Physik*, 281(1):91–97, 1892.
- [105] Paola Gallo, Katrin Amann-Winkel, Charles Austen Angell, Mikhail Alexeevich Anisimov, Frédéric Caupin, Charusita Chakravarty, Erik Lascaris, Thomas Lorting, Athanassios Zois Panagiotopoulos, John Russo, Jonas Alexander Sellberg, Harry Eugene Stanley, Hajime Tanaka, Carlos Vega, Limei Xu, and Lars Gunnar Moody Pettersson. Water: A Tale of Two Liquids. *Chemical Reviews*, 116(13):7463–7500, 2016.
- [106] J. D. Bernal. A Geometrical Approach to the Structure Of Liquids. *Nature*, 183(4655):141–147, jan 1959.
- [107] Lars G. M. Pettersson and Anders Nilsson. The structure of water; From ambient to deeply supercooled. *Journal of Non-Crystalline Solids*, 407:399–417, 2015.
- [108] Kyung Hwan Kim, Alexander Späh, Harshad Pathak, Fivos Perakis, Daniel Mariedahl, Katrin Amann-Winkel, Jonas A. Sellberg, Jae Hyuk Lee, Sangsoo Kim, Jaehyun Park, Ki Hyun Nam, Tetsuo Katayama, and Anders Nilsson. Maxima in the thermodynamic response and correlation functions of deeply supercooled water. *Science*, 358(6370):1589–1593, dec 2017.
- [109] Frédéric Caupin, Vincent Holten, Chen Qiu, Emmanuel Guillerm, Max Wilke, Martin Frenz, José Teixeira, and Alan K. Soper. Comment on “Maxima in the thermodynamic response and correlation functions of deeply supercooled water”. *Science*, 360(6390):eaat1634, may 2018.
- [110] Pascal Larouche, Jean-Joseph Max, and Camille Chapados. Isotope effects in liquid water by infrared spectroscopy. II. Factor analysis of the temperature effect on and Isotope effects in liquid water by infrared spectroscopy. II. Factor analysis of the temperature effect on H₂O and D₂O. *The Journal of Chemical Physics The Journal of Chemical Physics The Journal of Chemical Physics The Journal of Chemical Physics Proteins Chinese Journal of Chemical Physics The Journal of Chemical Physics*, 129(10):184505–224512, 2008.
- [111] David M. Carey. Measurement of the Raman spectrum of liquid water. *Journal of Chemical Physics*, 108(7):2669–2675, 1998.
- [112] A. K. Soper. Is water one liquid or two? *Journal of Chemical Physics*, 150(23):234503, 2019.

- [113] Rui Shi and Hajime Tanaka. Direct Evidence in the Scattering Function for the Coexistence of Two Types of Local Structures in Liquid Water. *Journal of the American Chemical Society*, 142(6):2868–2875, 2020.
- [114] Dibyendu Bandyopadhyay, S. Mohan, S. K. Ghosh, and Niharendu Choudhury. Correlation of structural order, anomalous density, and hydrogen bonding network of liquid water. *Journal of Physical Chemistry B*, 117(29):8831–8843, 2013.
- [115] A. K. Soper. The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. *Chemical Physics*, 258(2-3):121–137, 2000.
- [116] Yizhak Marcus. Effect of ions on the structure of water: Structure making and breaking, 2009.
- [117] Kai Ma and Lin Zhao. The opposite effect of metal Ions on short-/long-range water structure: A multiple characterization study. *International Journal of Molecular Sciences*, 17(5):602–617, 2016.
- [118] Annette Faith Dexter (<https://physics.stackexchange.com/users/163219/annette-faith-dexter>). Phase diagram of water. Physics Stack Exchange. URL:<https://physics.stackexchange.com/q/346750> (version: 2017-07-21).
- [119] Arshad Khan. A Liquid Water Model: Density Variation from Supercooled to Superheated States, Prediction of H-Bonds, and Temperature Limits. *The Journal of Physical Chemistry B*, 104(47):11268–11274, 2000.
- [120] Congcong Huang, T. M. Weiss, D. Nordlund, K. T. Wikfeldt, L. G. M. Pettersson, and A. Nilsson. Increasing correlation length in bulk supercooled H₂O, D₂O, and NaCl solution determined from small angle x-ray scattering. *Journal of Chemical Physics*, 113(20):134504, 2010.
- [121] S. J. Suresh and V. M. Naik. Hydrogen bond thermodynamic properties of water from dielectric constant data. *The Journal of Chemical Physics*, 113(10), 2000.
- [122] Hitoshi Ohtaki. Effects of temperature and pressure on hydrogen bonds in water and in formamide. *Journal of Molecular Liquids*, 103-104(SPEC.):3–13, 2003.
- [123] A. V. Okhulkov, Yu. N. Demianets, and Yu. E. Gorbaty. X-ray scattering in liquid water at pressures of up to 7.7 kbar: Test of a fluctuation model. *The Journal of Chemical Physics*, 100(2):1578–1588, 1994.
- [124] A. Cunsolo, F. Formisano, C. Ferrero, F. Bencivenga, and S. Finet. Pressure dependence of the large-scale structure of water. *The Journal of Chemical Physics*, 131(19):194502, 2009.
- [125] Stephanie R. Dillon and Ralph C. Dougherty. Raman studies of the solution structure of univalent electrolytes in water. *Journal of Physical Chemistry A*, 106(34):7647–7650, 2002.
- [126] R. Mancinelli, A. Botti, F. Bruni, M. A. Ricci, and A. K. Soper. Hydration of sodium, potassium, and chloride ions in solution and the concept of structure maker/breaker. *Journal of Physical Chemistry B*, 111(48):13570–13577, 2007.
- [127] R. Mancinelli, A. Botti, F. Bruni, M. A. Ricci, and A. K. Soper. Perturbation of water structure due to monovalent ions in solution. *Physical Chemistry Chemical Physics*, 9(23):2959–2967, 2007.

- [128] Alan K. Soper and Kristian Weckström. Ion solvation and water structure in potassium halide aqueous solutions. *Biophysical Chemistry*, 124(3):180–191, 2006.
- [129] O. N. Pestova, Yu. P. Kostikov, and M. K. Khripun. X-ray phase analysis of structure of water-salt systems: NaCl-H₂O and KCl-H₂O. *Russian Journal of Applied Chemistry*, 77(7):1066–1069, 2004.
- [130] Jörg Holzmann, Ralf Ludwig, Alfons Geiger, and Dietmar Paschek. Pressure and salt effects in simulated water: Two sides of the same coin? *Angewandte Chemie - International Edition*, 46(46):8907–8911, 2007.
- [131] S. Imberti, A. Botti, F. Bruni, G. Cappa, M. A. Ricci, and A. K. Soper. Ions in water: The microscopic structure of concentrated hydroxide solutions. *Journal of Chemical Physics*, 122(19):194509, 2005.
- [132] K. Winkel, M. Seidl, T. Loerting, L. E. Bove, S. Imberti, V. Molinero, F. Bruni, R. Mancinelli, and M. A. Ricci. Structural study of low concentration LiCl aqueous solutions in the liquid, supercooled, and hyperquenched glassy states. *Journal of Chemical Physics*, 134(2):024515, 2011.
- [133] Samuel Lenton, Natasha H. Rhys, James J. Towey, Alan K. Soper, and Lorna Dougan. Highly compressed water structure observed in a perchlorate aqueous solution. *Nature Communications*, 8(1):1–5, 2017.
- [134] Chiara H. Giammanco, Daryl B. Wong, and Michael D. Fayer. Water dynamics in divalent and monovalent concentrated salt solutions. *Journal of Physical Chemistry B*, 116(46):13781–13792, 2012.
- [135] Kim D. Collins, George W. Neilson, and John E. Enderby. Ions in water: Characterizing the forces that control chemical processes and biological structure. *Biophysical Chemistry*, 128(2-3):95–104, 2007.
- [136] R. Mancinelli, A. Sodo, F. Bruni, M. A. Ricci, and A. K. Soper. Influence of concentration and anion size on hydration of H⁺ ions and water structure. *Journal of Physical Chemistry B*, 113(13):4075–4081, 2009.
- [137] Sebastian Busch, Christian D. Lorenz, Jonathan Taylor, Luis Carlos Pardo, and Sylvia E. McLain. Short-Range Interactions of Concentrated Proline in Aqueous Solution. *J. Phys. Chem. B*, 118:14267–14277, 2014.
- [138] N. H. Rhys, A. K. Soper, and L. Dougan. The hydrogen-bonding ability of the amino acid glutamine revealed by neutron diffraction experiments. *Journal of Physical Chemistry B*, 116(45):13308–13319, 2012.
- [139] Natasha H. Rhys, Alan K. Soper, and Lorna Dougan. Hydrophilic Association in a Dilute Glutamine Solution Persists Independent of Increasing Temperature. *Journal of Physical Chemistry B*, 119(51):15644–15651, 2015.
- [140] Sylvia E. McLain, Alan K. Soper, and Anthony Watts. Water structure around dipeptides in aqueous solutions. *European Biophysics Journal*, 37(5):647–655, 2008.
- [141] A. K. Soper, E. W. Castner, and Alenka Luzar. Impact of urea on water structure: A clue to its properties as a denaturant? *Biophysical Chemistry*, 105(2-3):649–666, 2003.

- [142] Christoph J. Sahle, Martin A. Schroer, Iina Juurinen, and Johannes Niskanen. Influence of TMAO and urea on the structure of water studied by inelastic X-ray scattering. *Physical Chemistry Chemical Physics*, 18(24):16518–16526, 2016.
- [143] Sam Lenton, Natasha H. Rhys, James Joseph Towey, Alan K. Soper, and Lorna Dougan. Temperature Dependent Segregation in Alcohol-Water Binary Mixtures Is Driven by Water Clustering, 2018.
- [144] Alan K. Soper, Lorna Dougan, Jason Crain, and John L. Finney. Excess entropy in alcohol-water solutions: A simple clustering explanation. *Journal of Physical Chemistry B*, 110(8):3472–3476, 2006.
- [145] Serena Cozzolino, Rosario Oliva, Giuseppe Graziano, and Pompea Del Vecchio. Counteraction of denaturant-induced protein unfolding is a general property of stabilizing agents. *Physical Chemistry Chemical Physics*, 20(46):29389–29398, 2018.
- [146] Yuu Sasaki, Yuka Horikawa, Takashi Tokushima, Kazumasa Okada, Masaki Oura, and Misako Aida. Hydration structure of trimethylamine N-oxide in aqueous solutions revealed by soft X-ray emission spectroscopy and chemometric analysis. *Physical Chemistry Chemical Physics*, 18(39):27648–27653, 2016.
- [147] Christoph Hölzl, Patrick Kibies, Sho Imoto, Roland Frach, Saba Suladze, Roland Winter, Dominik Marx, Dominik Horinek, and Stefan M. Kast. Design principles for high-pressure force fields: Aqueous TMAO solutions from ambient to kilobar pressures. *Journal of Chemical Physics*, 144(14):144104, 2016.
- [148] Sho Imoto, Harald Forbert, and Dominik Marx. Aqueous TMAO solutions as seen by theoretical THz spectroscopy: Hydrophilic: versus hydrophobic water. *Physical Chemistry Chemical Physics*, 20(9):6146–6158, 2018.
- [149] Inga Kolling, Christoph Hölzl, Sho Imoto, Serena R Alfarano, Hendrik Vondracek, Lukas Knake, Federico Sebastiani, Fabio Novelli, Claudius Hoberg, Jean-Blaise Brubach, Pascale Roy, Harald Forbert, Gerhard Schwaab, Dominik Marx, and Martina Havenith. Aqueous TMAO solution under high hydrostatic pressure. *Physical Chemistry Chemical Physics*, 23(19):11355–11365, 2021.
- [150] Sho Imoto, Patrick Kibies, Christopher Rosin, Roland Winter, Stefan M. Kast, and Dominik Marx. Toward Extreme Biophysics: Deciphering the Infrared Response of Biomolecular Solutions at High Pressures. *Angewandte Chemie - International Edition*, 55(33):9534–9538, 2016.
- [151] Michael Knierbein, Christoph Held, Christoph Hölzl, Dominik Horinek, Michael Paulus, Gabriele Sadowski, Christian Sternemann, and J. Nase. Density variations of TMAO solutions in the kilobar range: Experiments, PC-SAFT predictions, and molecular dynamics simulations. *Biophysical Chemistry*, 253(May):106222, 2019.
- [152] Sarah G. Zetterholm, Genevieve A. Verville, Leeann Boutwell, Christopher Boland, John C. Prather, Jonathan Bethea, Jordan Cauley, Kayla E. Warren, Shelley A. Smith, David H. Magers, and Nathan I. Hammer. Noncovalent Interactions between Trimethylamine N-Oxide (TMAO), Urea, and Water. *Journal of Physical Chemistry B*, 122(38):8805–8811, 2018.
- [153] Peter J. Rossky. Protein denaturation by urea: Slash and bond. *Proceedings of the National Academy of Sciences*, 105(44):16825–16826, nov 2008.

- [154] Deepak R. Canchi and Angel E. García. Cosolvent Effects on Protein Stability. *Annual Review of Physical Chemistry*, 64(1):273–293, 2013.
- [155] Damien Laage and James T. Hynes. A molecular jump mechanism of water reorientation. *Science*, 311(5762):832–835, 2006.
- [156] Guillaume Stirnemann, Erik Wernersson, Pavel Jungwirth, and Damien Laage. Mechanisms of acceleration and retardation of water dynamics by ions. *Journal of the American Chemical Society*, 135(32):11824–11831, 2013.
- [157] Damien Laage and Guillaume Stirnemann. Effect of Ions on Water Dynamics in Dilute and Concentrated Aqueous Salt Solutions. *Journal of Physical Chemistry B*, 123(15):3312–3324, 2019.
- [158] Jun Soo Kim, Zhe Wu, Andrew R. Morrow, Anand Yethiraj, and Arun Yethiraj. Self-diffusion and viscosity in electrolyte solutions. *Journal of Physical Chemistry B*, 116(39):12007–12013, 2012.
- [159] Michael D. Fayer, David E. Moilanen, Daryl Wong, Daniel E. Rosenfeld, Emily E. Fenn, and Sungnam Park. Water dynamics in salt solutions studied with ultrafast two-dimensional infrared (2D IR) vibrational echo spectroscopy. *Accounts of chemical research*, 42(9):1210–1219, 2009.
- [160] Minbiao Ji, Michael Odelius, and K. J. Gaffney. Large angular jump mechanism observed for hydrogen bond exchange in aqueous perchlorate solution. *Science*, 328(5981):1003–1005, 2010.
- [161] Qianshun Wei, Dexia Zhou, and Hongtao Bian. Negligible cation effect on the vibrational relaxation dynamics of water molecules in NaClO₄ and LiClO₄ aqueous electrolyte solutions. *RSC Advances*, 7(82):52111–52117, 2017.
- [162] Sietse T. Van Der Post, Stefan Scheidelaar, and Huib J. Bakker. Femtosecond study of the effects of ions on the reorientation dynamics of water. *Journal of Molecular Liquids*, 176:22–28, 2012.
- [163] Daniela Russo, Greg Hura, and Teresa Head-Gordon. Hydration Dynamics Near a Model Protein Surface. *Biophysical Journal*, 86(3):1852–1862, 2004.
- [164] Damien Laage, Thomas Elsaesser, and James T. Hynes. Water Dynamics in the Hydration Shells of Biomolecules. *Chemical Reviews*, 117(16):10694–10725, 2017.
- [165] Fabio Sterpone, Guillaume Stirnemann, James T. Hynes, and Damien Laage. Water hydrogen-bond dynamics around amino acids: The key role of hydrophilic hydrogen-bond acceptor groups. *Journal of Physical Chemistry B*, 114(5):2083–2089, 2010.
- [166] Ning Zhang, Zhuanglin Shen, Cong Chen, Gaohong He, and Ce Hao. Effect of hydrogen bonding on self-diffusion in methanol/water liquid mixtures: A molecular dynamics simulation study. *Journal of Molecular Liquids*, 203:90–97, 2015.
- [167] Wen Jun Xie, Seoncheol Cha, Tatsuhiko Ohto, Wataru Mizukami, Yuezhi Mao, Manfred Wagner, Mischa Bonn, Johannes Hunger, and Yuki Nagata. Large Hydrogen-Bond Mismatch between TMAO and Urea Promotes Their Hydrophobic Association. *Chem*, 4(11):2615–2627, 2018.

- [168] P. J. Hore. *Nuclear Magnetic Resonance*. Oxford University Press, Oxford, second edition, 2014.
- [169] Shoichi Okouchi, Tetsushi Moto, Yoshimasa Ishihara, Haruhiko Numajiri, and Hisashi Uedaira. Hydration of amines, diamines, polyamines and amides studied by NMR. *J. Chem. SOC. Faraday Trans.*, 92(11):1853–1857, 1996.
- [170] Yoshimasa Ishihara, Shoichi Okouchi, and Hisashi Uedaira. Dynamics of hydration of alcohols and diols in aqueous solutions. *Journal of the Chemical Society - Faraday Transactions*, 93(18):3337–3342, 1997.
- [171] Kota Usui, Johannes Hunger, Marialore Sulpizi, Tatsuhiko Ohto, Mischa Bonn, and Yuki Nagata. Ab Initio Liquid Water Dynamics in Aqueous TMAO Solution. *Journal of Physical Chemistry B*, 119(33):10597–10606, 2015.
- [172] National Geographic. The Mariana Trench, 2021.
- [173] R. Orosei, S. E. Lauro, E. Pettinelli, A. Cicchetti, M. Coradini, B. Cosciotti, F. Di Paolo, E. Flamini, E. Mattei, M. Pajola, F. Soldovieri, M. Cartacci, F. Cassenti, A. Frigeri, S. Giuppi, R. Martufi, A. Masdea, G. Mitri, C. Nenna, R. Noschese, M. Restano, and R. Seu. Radar evidence of subglacial liquid water on Mars. *Science*, 361(6401):490–493, 2018.
- [174] Sebastian Emanuel Lauro, Elena Pettinelli, Graziella Caprarelli, Luca Guallini, Angelo Pio Rossi, Elisabetta Mattei, Barbara Cosciotti, Andrea Cicchetti, Francesco Soldovieri, Marco Cartacci, Federico Di Paolo, Raffaella Noschese, and Roberto Orosei. Multiple subglacial water bodies below the south pole of Mars unveiled by new MARSIS data. *Nature Astronomy*, 2020.
- [175] Alexey Rozov, Iskander Khusainov, Kamel El Omari, Ramona Duman, Vitaliy Mykhaylyk, Marat Yusupov, Eric Westhof, Armin Wagner, and Gulnara Yusupova. Importance of potassium ions for ribosome structure and function revealed by long-wavelength X-ray diffraction. *Nature Communications*, 10(1):1–12, 2019.
- [176] C. B. Black, H. W. Huang, and J. A. Cowan. Biological coordination chemistry of magnesium, sodium, and potassium ions. Protein and nucleotide binding sites. *Coordination Chemistry Reviews*, 135-136(C):165–202, 1994.
- [177] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *Feynman: The Feynman Lectures on Physics*. Basic Books, New York, sixth edition, 2010.
- [178] W. H. Bragg and W. L. Bragg. The reflection of X-rays by crystals. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 88(605):428–438, jul 1913.
- [179] Thomas Young. *A Course of Lectures on Natural Philosophy and the Mechanical Arts: Volume 2*. Johnson, Joseph, London, 1807.
- [180] D. S. Silva. *Elementary Scattering Theory: For X-Ray and Neutron Users*. Oxford University Press, New York, fifth edition, 2017.
- [181] Louis De Broglie. Waves and Quanta. *Nature*, 112(2815):540–540, oct 1923.
- [182] James Chadwick. The Existence of a Neutron. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 136(830):692–708, 1932.

- [183] Varley F. Sears. Neutron scattering lengths and cross sections. *Neutron News*, 3(3):26–37, 1992.
- [184] A. K. Soper. Rutherford Appleton Laboratory Technical Report RAL-TR-2011-013. Technical Report July, 2011.
- [185] B. T. M. Willis and C. J. Carlile. *Experimental Neutron Scattering*. Oxford University Press, Oxford, first edition, 2009.
- [186] Natasha H. Rhys. *Exploring the Structural Properties of Natural and Synthetic Biological Molecules in Aqueous Solution*. PhD thesis, University of Leeds, 2015.
- [187] J. W. G. Thomason. The ISIS Spallation Neutron and Muon Source - The first thirty-three years. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 917(August 2018):61–67, 2019.
- [188] Oliver S. Hammond, Daniel T. Bowron, and Karen J. Edler. Liquid structure of the choline chloride-urea deep eutectic solvent (reline) from neutron diffraction and atomistic modelling. *Green Chemistry*, 18(9):2736–2744, 2016.
- [189] Oliver S. Hammond, Daniel T. Bowron, and Karen J. Edler. The Effect of Water upon Deep Eutectic Solvent Nanostructure: An Unusual Transition from Ionic Mixture to Aqueous Solution. *Angewandte Chemie - International Edition*, 56(33):9782–9785, 2017.
- [190] Oliver S. Hammond, Karen J. Edler, Daniel T. Bowron, and Laura Torrente-Murciano. Deep eutectic-solvothermal synthesis of nanostructured ceria. *Nature Communications*, 8(May 2016):1–7, 2017.
- [191] Catherine R. Hill, Christian Mitterdorfer, Tristan G. A. Youngs, Daniel T. Bowron, Helen J. Fraser, and Thomas Loerting. Neutron Scattering Analysis of Water’s Glass Transition and Micropore Collapse in Amorphous Solid Water. *Physical Review Letters*, 116(21):1–5, 2016.
- [192] Christian Mitterdorfer, Marion Bauer, Tristan G. A. Youngs, Daniel T. Bowron, Catherine R. Hill, Helen J. Fraser, John L. Finney, and Thomas Loerting. Small-angle neutron scattering study of micropore collapse in amorphous solid water. *Physical Chemistry Chemical Physics*, 16(30):16013–16020, 2014.
- [193] Carlos Cabrillo, Fabienne Barroso-Bujans, Ricardo Fernandez-Perea, Felix Fernandez-Alonso, Daniel Bowron, and Francisco Javier Bermejo. Adsorbate-induced ordering and bilayer formation in propanol-graphite-oxide intercalates. *Carbon*, 100:546–555, 2016.
- [194] V. Tudisca, F. Bruni, E. Scoppola, R. Angelini, B. Ruzicka, L. Zulian, A. K. Soper, and M. A. Ricci. Neutron diffraction study of aqueous Laponite suspensions at the NIMROD diffractometer. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 90(3):1–9, 2014.
- [195] Tristan G. A. Youngs, Haresh Manyar, Daniel T. Bowron, Lynn F. Gladden, and Christopher Hardacre. Probing chemistry and kinetics of reactions in heterogeneous catalysis. *Chemical Science*, 4(9):3484–3489, 2013.

- [196] K. L. Stefanopoulos, F. K. Katsaros, Th. A. Steriotis, A. A. Sapalidis, M. Thommes, D. T. Bowron, and T. G. A. Youngs. Anomalous Depletion of Pore-Confined Carbon Dioxide upon Cooling below the Bulk Triple Point: An in Situ Neutron Diffraction Study. *Physical Review Letters*, 116(2):1–6, 2016.
- [197] Konstantinos L. Stefanopoulos, Tristan G. A. Youngs, Richard Sakurovs, Leslie F. Ruppert, Jitendra Bahadur, and Yuri B. Melnichenko. Neutron Scattering Measurements of Carbon Dioxide Adsorption in Pores within the Marcellus Shale: Implications for Sequestration. *Environmental Science and Technology*, 51(11):6515–6521, 2017.
- [198] A. K. Soper and K. J. Edler. Coarse-grained empirical potential structure refinement: Application to a reverse aqueous micelle. *Biochimica et Biophysica Acta - General Subjects*, 1861(6):1652–1660, 2017.
- [199] D. T. Bowron, A. K. Soper, K. Jones, S. Ansell, S. Birch, J. Norris, L. Perrott, D. Riedel, N. J. Rhodes, S. R. Wakefield, A. Botti, M. A. Ricci, F. Grazzi, and M. Zoppi. NIMROD: The Near and InterMediate Range Order Diffractometer of the ISIS second target station. *Review of Scientific Instruments*, 81(3):033905, 2010.
- [200] A. K. Soper. Empirical potential Monte Carlo simulation of fluid structure. *Chemical Physics*, 202(2-3):295–306, 1996.
- [201] Alan K. Soper, Daniel Bowron, Helen Thompson, Fabio Bruni, Maria Antonietta Ricci, Sylvia McLain, Stefan Klotz, Thierry Straessle, Silvia Imberti, Rowan Hargreaves, Tristan Youngs, Sam Callear, and Christoph Salzmänn. Empirical Potential Structure Refinement A User’s Guide. Technical report, 2017.
- [202] Christopher Hardacre, John D. Holbrey, S. E. Jane McMath, Daniel T. Bowron, and Alan K. Soper. Structure of molten 1,3-dimethylimidazolium chloride using neutron diffraction. *Journal of Chemical Physics*, 118(1):273–278, 2003.
- [203] D. T. Bowron, J. L. Finney, and A. K. Soper. Structural investigation of solute-solute interactions in aqueous solutions of tertiary butanol. *Journal of Physical Chemistry B*, 102(18):3551–3563, 1998.
- [204] D. T. Bowron, A. K. Soper, and J. L. Finney. Temperature dependence of the structure of a 0.06 mole fraction tertiary butanol-water solution. *Journal of Chemical Physics*, 114(14):6203–6219, 2001.
- [205] Natasha H. Rhys, Imogen B. Duffy, Christopher L. Sowden, Christian D. Lorenz, and Sylvia E. McLain. On the hydration of DOPE in solution. *Journal of Chemical Physics*, 150(11):115104, 2019.
- [206] Frank Weinhold and Roger A. Klein. What is a hydrogen bond? Mutually consistent theoretical and experimental criteria for characterizing H-bonding interactions. In *Molecular Physics*, volume 110, pages 565–579, 2012.
- [207] Upayan Baul and Satyavani Vemparala. Ion hydration and associated defects in hydrogen bond network of water: Observation of reorientationally slow water molecules beyond first hydration shell in aqueous solutions of MgCl₂. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 91(1):12114, 2015.

- [208] J. B. Lambert and E. P. Mazzola. *Nuclear magnetic Resonance Spectroscopy: An Introduction to Principles, Applications, and Experimental Methods*. Pearson Education, Inc., Upper Saddle River, first edition, 2004.
- [209] P. Zeeman. Xxxii. on the influence of magnetism on the nature of the light emitted by a substance. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 43(262):226–239, 1897.
- [210] National Institute of Advanced Industrial Science and Technology. Sdbs no. 1300hsp-01-876, 1999.
- [211] V. Graf, F. Noack, and G. J. Béné. Proton spin T1 relaxation dispersion in liquid H2O by slow proton-exchange. *The Journal of Chemical Physics*, 72(2):861–863, 1980.
- [212] Claridge, Timothy D. W. *High-Resolution NMR Techniques in Organic Chemistry*. Elsevier, Oxford, third edition, 2016.
- [213] E. O. Stejskal and J. E. Tanner. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, 42(1):288–292, 1965.
- [214] Shinya Honda, Toshihiko Akiba, Yusuke S. Kato, Yoshito Sawada, Masakazu Sekijima, Miyuki Ishimura, Ayako Ooishi, Hideki Watanabe, Takayuki Odahara, and Kazuaki Harata. Crystal structure of a ten-amino acid protein. *Journal of the American Chemical Society*, 130(46):15327–15331, 2008.
- [215] Mu Gao, Hui Lu, and Klaus Schulten. Unfolding of titin domains studied by molecular dynamics simulations. *Journal of Muscle Research and Cell Motility*, 23(5-6):513–521, 2002.
- [216] Hu Chen, Guohua Yuan, Rickson S. Winardhi, Mingxi Yao, Ionel Popa, Julio M. Fernandez, and Jie Yan. Dynamics of Equilibrium Folding and Unfolding Transitions of Titin Immunoglobulin Domain under Constant Forces. *Journal of the American Chemical Society*, 137(10):3540–3546, 2015.
- [217] Susan B. Fowler, Robert B. Best, José L. Toca Herrera, Trevor J. Rutherford, Annette Steward, Emanuele Paci, Martin Karplus, and Jane Clarke. Mechanical unfolding of a titin Ig domain: Structure of unfolding intermediate revealed by combining AFM, molecular dynamics simulations, NMR and protein engineering. *Journal of Molecular Biology*, 322(4):841–849, 2002.
- [218] Mariano Carrion-Vazquez, Andres F. Oberhauser, Susan B. Fowler, Piotr E. Marszalek, Sheldon E. Broedel, Jane Clarke, and Julio M. Fernandez. Mechanical and chemical unfolding of a single protein: A comparison. *Proceedings of the National Academy of Sciences*, 96:3694–3699, 1999.
- [219] Susan B. Fowler and Jane Clarke. Mapping the folding pathway of an immunoglobulin domain: Structural detail from phi value analysis and movement of the transition state. *Structure*, 9(5):355–366, 2001.
- [220] Eric Botello, Nolan C. Harris, Jacob Sargent, Wei Hung Chen, Kuan Jih Lin, and Ching Hwa Kiang. Temperature and chemical denaturant dependence of forced unfolding of titin I27. *Journal of Physical Chemistry B*, 113(31):10845–10848, 2009.

- [221] Larissa Tskhovrebova and John Trinick. Properties of titin immunoglobulin and fibronectin-3 domains. *Journal of Biological Chemistry*, 279(45):46351–46354, 2004.
- [222] Novagen. *pET System Manual*. Biosystems, 2003.
- [223] M. Mandel and A. Higa. Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology*, 53(1):159–162, oct 1970.
- [224] Carlos F. Barbas, Dennis R. Burton, Jamie K. Scott, and Gregg J. Silverman. Quantitation of DNA and RNA. *Cold Spring Harbor Protocols*, 2007(11):pdb.ip47, nov 2007.
- [225] Joshua A. Bornhorst and Joseph J. Falke. Purification of proteins using polyhistidine affinity tags. In *Methods in Enzymology*, pages 245–254. 2000.
- [226] Yoav Peleg, Vadivel Prabahar, Dominika Bednarczyk, and Tamar Unger. Heterologous Gene Expression in E.coli. 1586:221–230, 2017.
- [227] Shinya Honda, Kazuhiko Yamasaki, Yoshito Sawada, and Hisayuki Morii. 10 Residue folded peptide designed by segment statistics. *Structure*, 12(8):1507–1518, 2004.
- [228] Amaury A. Ibarra, Gail J. Bartlett, Zsófia Hegedüs, Som Dutt, Fruzsina Horbor, Katherine A. Horner, Kristina Hetherington, Kirstin Spence, Adam Nelson, Thomas A. Edwards, Derek N. Woolfson, Richard B. Sessions, and Andrew J. Wilson. Predicting and Experimentally Validating Hot-Spot Residues at Protein - Protein Interfaces. *ACS Chemical Biology*, page acschembio.9b00560, sep 2019.
- [229] Joseph R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Springer US, Boston, MA, third edition, 2006.
- [230] F. W. J. Teale. The ultraviolet fluorescence of proteins in neutral solution. *Biochemical Journal*, 76(2):381–388, aug 1960.
- [231] E. A. Burstein, N. S. Vedenkina, and M. N. Ivkova. Fluorescence and the Location of Tryptophan Residues in Protein Molecules. *Photochemistry and Photobiology*, 18(4):263–279, oct 1973.
- [232] Sabina Improta, Anastasia S. Politou, and Annalisa Pastore. Immunoglobulin-like modules from titin I-band: Extensible components of muscle elasticity. *Structure*, 4(3):323–337, 1996.
- [233] Shota Murakami, Tomohiko Hayashi, and Masahiro Kinoshita. Effects of salt or cosolvent addition on solubility of a hydrophobic solute in water: Relevance to those on thermal stability of a protein. *Journal of Chemical Physics*, 146(5):055102, 2017.
- [234] Matthew Auton, Luis Marcelo F. Holthauzen, and D. Wayne Bolen. Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(39):15317–15322, 2007.

- [235] Lan Hua, Ruhong Zhou, C. D. Thirumalai, and B. J. Berne. Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proceedings of the National Academy of Sciences of the United States of America*, 105(44):16928–16933, 2008.
- [236] Deepak R. Canchi and Angel E. Garcíá. Backbone and side-chain contributions in protein denaturation by urea. *Biophysical Journal*, 100(6):1526–1533, mar 2011.
- [237] J. Ramprakash, V. Doseeva, A. Galkin, W. Krajewski, L. Muthukumar, S. Pulalarevu, E. Demirkan, O. Herzberg, J. Moulton, and F. P. Schwarz. Comparison of the chemical and thermal denaturation of proteins by a two-state transition model. *Analytical Biochemistry*, 374(1):221–230, 2008.
- [238] Sharon M. Kelly, Thomas J. Jess, and Nicholas C. Price. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1751(2):119–139, aug 2005.
- [239] Dave3457. *File: Circular.Polarization.Circularly.Polarized.Light Circular.Polarizer Creating.Left.Handed.Helix.View.svg*. Dec 2015.
- [240] Stephen W. Provencher and Jürgen Glöckner. Estimation of Globular Protein Secondary Structure from Circular Dichroism. *Biochemistry*, 20(1):33–37, 1981.
- [241] Andrew J. Miles, Sergio G. Ramalli, and B. A. Wallace. DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. *Protein Science*, (June):1–10, 2021.
- [242] John W. Rowe, Jordan D. Tobin, Robert M. Rosa, and Reubin Andres. Effect of experimental potassium deficiency on glucose and insulin metabolism. *Metabolism*, 29(6):498–502, jun 1980.
- [243] Alex P. Gaiduk and Giulia Galli. Local and Global Effects of Dissolved Sodium Chloride on the Structure of Water. *Journal of Physical Chemistry Letters*, 8(7):1496–1502, 2017.
- [244] R. D. Shannon. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallographica Section A*, 32(5):751–767, sep 1976.
- [245] Chung Peter Chieh. Hydration, 2021.
- [246] G. Jones and M. Dole. The Viscosity of Aqueous Solutions of Strong Electrolytes with Special Reference to Barium Chloride. *Journal of the American Chemical Society*, 51(10):2950–2964, 1929.
- [247] H. J. Bakker. Structural dynamics of aqueous salt solutions. *Chemical Reviews*, 108(4):1456–1473, 2008.
- [248] Hitoshi Ohtaki and Tamás Radnai. Structure and Dynamics of Hydrated Ions. *Chemical Reviews*, 93(3):1157–1204, 1993.
- [249] S. Ansell, A. C. Barnes, P. E. Mason, G. W. Neilson, and S. Ramos. X-ray and neutron scattering studies of the hydration structure of alkali ions in concentrated aqueous solutions. *Biophysical Chemistry*, 124(3):171–179, 2006.

- [250] Yekkoni L. Jeyachandran, Frank Meyer, Sankaranarayanan Nagarajan, Andreas Benkert, Marcus Bär, Monika Blum, Wanli Yang, Friedrich Reinert, Clemens Heske, Lothar Weinhardt, and Michael Zharnikov. Ion-solvation-induced molecular reorganization in liquid water probed by resonant inelastic soft X-ray scattering. *Journal of Physical Chemistry Letters*, 5(23):4143–4148, 2014.
- [251] R. Gulich, M. Köhler, P. Lunkenheimer, and A. Loidl. Dielectric spectroscopy on aqueous electrolytic solutions. *Radiation and Environmental Biophysics*, 48(1):107–114, 2009.
- [252] Masato Kondoh, Yasuhiro Ohshima, and Masaaki Tsubouchi. Ion effects on the structure of water studied by terahertz time-domain spectroscopy. *Chemical Physics Letters*, 591:317–322, 2014.
- [253] Jean-Joseph Max, Véronique Gessinger, Caroline Van Driessche, Pascal Larouche, and Camille Chapados. Infrared spectroscopy of aqueous ionic salt solutions at low concentrations. *The Journal of Chemical Physics*, 126(10):184507–114509, 2007.
- [254] Snehasis Chowdhuri and Amalendu Chandra. Dynamics of halide ion-water hydrogen bonds in aqueous solutions: Dependence on ion size and temperature. *Journal of Physical Chemistry B*, 110(19):9674–9680, 2006.
- [255] Qiang Zhang, Hailong Chen, Tianmin Wu, Tan Jin, Zhijun Pan, Junrong Zheng, Yiqin Gao, and Wei Zhuang. The opposite effects of sodium and potassium cations on water dynamics. *Chemical Science*, 8(2):1429–1435, 2017.
- [256] Prashant N. Jethva and Jayant B. Udgaonkar. The Osmolyte TMAO Modulates Protein Folding Cooperativity by Altering Global Protein Stability. *Biochemistry*, 57(40):5851–5863, 2018.
- [257] Marilisa Vigorita, Serena Cozzolino, Rosario Oliva, Giuseppe Graziano, and Pompea Del Vecchio. Counteraction ability of TMAO toward different denaturing agents. *Biopolymers*, 109(10), 2018.
- [258] Hsuan Lei Sung and David J. Nesbitt. High pressure single-molecule FRET studies of the lysine riboswitch: Cationic and osmolytic effects on pressure induced denaturation. *Physical Chemistry Chemical Physics*, 22(28):15853–15866, 2020.
- [259] Tim Pongratz, Patrick Kibies, Lukas Eberlein, Nicolas Tielker, Christoph Hölzl, Sho Imoto, Markus Beck Erlach, Simon Kurrmann, Paul Hendrik Schummel, Martin Hofmann, Oliver Reiser, Roland Winter, Werner Kremer, Hans Robert Kalbitzer, Dominik Marx, Dominik Horinek, and Stefan M. Kast. Pressure-dependent electronic structure calculations using integral equation-based solvation models. *Biophysical Chemistry*, 257(August 2019):106258, 2020.
- [260] Michael Knierbein, Maximilian Venhuis, Christoph Held, and Gabriele Sadowski. Thermodynamic properties of aqueous osmolyte solutions at high-pressure conditions. *Biophysical Chemistry*, 253(June):106211, 2019.
- [261] Christoph J. Sahle, Martin A. Schroer, Iina Juurinen, and Johannes Niskanen. Influence of TMAO and urea on the structure of water studied by inelastic X-ray scattering. *Physical Chemistry Chemical Physics*, 18(24):16518–16526, 2016.
- [262] Ewa Anna Oprzeska-Zingrebe and Jens Smiatek. Aqueous Mixtures of Urea and Trimethylamine-N-oxide: Evidence for Kosmotropic or Chaotropic Behavior? *Journal of Physical Chemistry B*, 123(20):4415–4424, 2019.

- [263] Isis high pressure.
- [264] Engineering Toolbox. Water - Density, Specific Weight and Thermal Expansion Coefficient, 2003.
- [265] T. Grindley and John E. Lind. PVT properties of water and mercury. *The Journal of Chemical Physics*, 54(9):3983–3989, 1971.
- [266] Dmitriy M. Makarov, Gennadiy I. Egorov, and Arkadiy M. Kolker. Density and volumetric properties of aqueous solutions of trimethylamine N-oxide in the temperature range from (278.15 to 323.15) K and at pressures up to 100 MPa. *Journal of Chemical and Engineering Data*, 60(5):1291–1299, 2015.
- [267] W. Brostow and Y. Sicotte. Coordination number in liquid argon. *Physica A: Statistical Mechanics and its Applications*, 80(5):513–522, 1975.
- [268] Vincent F. Chevrier, Jennifer Hanley, and Travis S. Altheide. Stability of perchlorate hydrates and their liquid solutions at the Phoenix landing site, mars. *Geophysical Research Letters*, 36(10):L10202, may 2009.
- [269] Liquid-Solid Phase Diagrams: Salt Solutions, 8 2020. [Online; accessed 2021-09-08].
- [270] G. M. Marion, D. C. Catling, K. J. Zahnle, and M. W. Claire. Modeling aqueous perchlorate chemistries with applications to Mars. *Icarus*, 207(2):675–685, 2010.
- [271] L. Daly, M. R. Lee, S. Piazzolo, S. Griffin, M. Bazargan, F. Campanale, P. Chung, B. E. Cohen, A. E. Pickersgill, L. J. Hallis, P. W. Trimby, R. Baumgartner, L. V. Forman, and G. K. Benedix. Boom boom pow: Shock-facilitated aqueous alteration and evidence for two shock events in the Martian nakhlite meteorites. *Science Advances*, 5(9):eaaw5549, 2019.
- [272] Martin J. Siegert, John C. Priscu, Irina A. Alekhina, Jemma L. Wadham, and W. Berry Lyons. Antarctic subglacial lake exploration: First results and future plans. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2059):20140466, 2016.
- [273] Brent C. Christner, John C. Priscu, Amanda M. Achberger, Carlo Barbante, Sasha P. Carter, Knut Christianson, Alexander B. Michaud, Jill A. Mikucki, Andrew C. Mitchell, Mark L. Skidmore, Trista J. Vick-Majors, W. P. Adkins, S. Anandakrishnan, S. Anandakrishnan, L. Beem, A. Behar, M. Beitch, R. Bolsey, C. Braneky, A. Fisher, N. Foley, K. D. Mankoff, D. Sampson, S. Tulaczyk, R. Edwards, S. Kelley, J. Sherve, H. A. Fricker, S. Siegfried, B. Guthrie, Timothy Hodson, R. Powell, R. Scherer, H. Horgan, R. Jacobel, E. McBryan, and A. Purcell. A microbial ecosystem beneath the West Antarctic ice sheet. *Nature*, 512(7514):310–313, 2014.
- [274] Bassam Oudh Aljohny. Halophilic bacterium-A review of new studies. *Biosciences Biotechnology Research Asia*, 12(3):2061–2069, 2015.
- [275] Jennifer Wadsworth and Charles S. Cockell. Perchlorates on Mars enhance the bacteriocidal effects of UV light. *Scientific Reports*, 7(1):4662, 2017.
- [276] Nisrine Jahmidi-Azizi, Rosario Oliva, Stewart Gault, Charles S. Cockell, and Roland Winter. The Effects of Temperature and Pressure on Protein-Ligand Binding in the Presence of Mars-Relevant Salts. *Biology*, 10(7):687, jul 2021.

- [277] O. N. Pestova, L. A. Myund, M. K. Khripun, and A. V. Prigaro. Polythermal study of the systems $M(\text{ClO}_4)_2 \cdot n\text{H}_2\text{O}$ ($M^{2+} = \text{Mg}^{2+}, \text{Ca}^{2+}, \text{Sr}^{2+}, \text{Ba}^{2+}$). *Russian Journal of Applied Chemistry*, 78(3):409–413, 2005.
- [278] Samuel Lenton, Natasha H. Rhys, James J. Towey, Alan K. Soper, and Lorna Dougan. Supplementary Information Highly compressed water structure observed in a perchlorate aqueous solution. *Nature Communications*, pages 3–5.
- [279] Engineering Toolbox. Density of aqueous solutions of inorganic sodium salts, 2017.
- [280] R. Mills. Self-diffusion in normal and heavy water in the range 1–45 Å. *Journal of Physical Chemistry*, 77(5):685–688, 1973.
- [281] Sungnam Park, Michael Odellius, and Kelly J. Gaffney. Ultrafast dynamics of hydrogen bond exchange in aqueous ionic solutions. *Journal of Physical Chemistry B*, 113(22):7825–7835, 2009.
- [282] Richard J. Foster, Robin A. Damion, Thomas G. Baboolal, Stephen W. Smye, and Michael E. Ries. A nuclear magnetic resonance study of water in aggregate solutions. *Royal Society Open Science*, 3(3):150705, 2016.
- [283] W. M. Ezzawam and M. E. Ries. Diffusion and Relaxometry to Study Carbohydrates Dissolved in Ionic Liquids. In Rongchun Zhang and Pingchuan Sun, editors, *NMR Methods for Characterization of Synthetic and Natural Polymers*, chapter Diffusion, pages 36 – 62. Royal Society of Chemistry, first edition, 2019.
- [284] Keri A. McKiernan, Brooke E. Husic, and Vijay S. Pande. Modeling the mechanism of CLN025 beta-hairpin formation. *Journal of Chemical Physics*, 147(10):104107, 2017.
- [285] Renzo Carta and Giuseppe Tola. Solubilities of L-Cystine, L-Tyrosine, L-Leucine, and Glycine in Aqueous Solutions at Various pHs and NaCl Concentrations. *Journal of Chemical and Engineering Data*, 41(3):414–417, 1996.
- [286] Hervé Cottin, Julia Michelle Kotler, Kristin Bartik, H. James Cleaves, Charles S. Cockell, Jean Pierre P. de Vera, Pascale Ehrenfreund, Stefan Leuko, Inge Loes Ten Kate, Zita Martins, Robert Pascal, Richard Quinn, Petra Rettberg, and Frances Westall. *Astrobiology and the Possibility of Life on Earth and Elsewhere...*, 2017.
- [287] Oliver Botta and Jeffrey L. Bada. Extraterrestrial organic compounds in meteorites. *Surveys in Geophysics*, 23(5):411–467, 2002.
- [288] Kathrin Altwegg, Hans Balsiger, Akiva Bar-Nun, Jean Jacques Berthelier, Andre Bieler, Peter Bochslers, Christelle Briois, Ursina Calmonte, Michael R. Combi, Hervé Cottin, Johan De Keyser, Frederik Dhooghe, Bjorn Fiethe, Stephen A. Fuselier, Sébastien Gasc, Tamas I. Gombosi, Kenneth C. Hansen, Myrtha Haessig, Annette Jäckel, Ernest Kopp, Axel Korth, Lena Le Roy, Urs Mall, Bernard Marty, Olivier Mousis, Tobias Owen, Henri Rème, Martin Rubin, Thierry Sémon, Chia Yu Tzou, James Hunter Waite, and Peter Würz. Prebiotic chemicals-amino acid and phosphorus in the coma of comet 67P/Churyumov-Gerasimenko. *Science Advances*, 2(5):e1600285, 2016.

- [289] Ernesto Scoppola, Armida Sodo, Sylvia E. McLain, Maria Antonietta Ricci, and Fabio Bruni. Water-peptide site-specific interactions: A structural study on the hydration of glutathione. *Biophysical Journal*, 106(8):1701–1709, 2014.
- [290] Sylvia E. McLain, Alan K. Soper, Isabella Daidone, Jeremy C. Smith, and Anthony Watts. Charge-based interactions between peptides observed as the dominant force for association in aqueous solution. *Angewandte Chemie - International Edition*, 47(47):9059–9062, 2008.
- [291] R. S. Taubner, J. J. Leitner, M. G. Firneis, and R. Hitzenberger. Modelling the Interior Structure of Enceladus Based on the 2014’s Cassini Gravity Data. *Origins of Life and Evolution of Biospheres*, 46(2-3):283–288, 2016.
- [292] Gabriel Tobie, Olivier Grasset, Jonathan I. Lunine, Antoine Mocquet, and Christophe Sotin. Titan’s internal structure inferred from a coupled thermal-orbital model. *Icarus*, 175(2):496–502, jun 2005.
- [293] H. Hussman. *Treatise on Geophysics, Volume 10, Volume 10*. Elsevier, London, first edition, 2007.
- [294] Tilman Spohn and Gerald Schubert. Oceans in the icy Galilean satellites of Jupiter?, 2003.
- [295] Friedrich Merkel and Franjo Bošnjaković. *Diagramme und Tabellen zur Berechnung der Absorptions-Kältemaschinen*. Berlin, first edition, 1929.
- [296] S. Roy, A. Hossain, K. Mahali, and B. K. Dolui. Thermodynamics and mechanisms of glycine solvation in aqueous NaCl and KCl solutions at 298.15 K. *Russian Journal of Physical Chemistry A*, 89(11):2111–2119, 2015.
- [297] Per-Gunnar Jonsson and Ake Kvick. Precision Neutron Diffraction Structure Determination of Protein and Nucleic Acid Components. III.* The Crystal and Molecular Structure of the Amino Acid α -Glycine. *Acta Cryst.*, B28:1827, 1972.
- [298] Hyung Kyu Lim, Young Sun Choi, and Seung Tae Hong. Magnesium perchlorate anhydrate, $\text{Mg}(\text{ClO}_4)_2$, from laboratory X-ray powder data. *Acta Crystallographica Section C: Crystal Structure Communications*, 67(6):i36–i38, 2011.
- [299] Attila G. Császár, Gábor Czakó, Tibor Furtenbacher, Jonathan Tennyson, Viktor Szalay, Sergei V. Shirin, Nikolai F. Zobov, and Oleg L. Polyansky. On equilibrium structures of the water molecule. *Journal of Chemical Physics*, 122(21):214305, 2005.
- [300] Limei Xu, Francesco Mallamace, Zhenyu Yan, Francis W. Starr, Sergey V. Buldyrev, and H. Eugene Stanley. Appearance of a fractional Stokes-Einstein relation in water and a structural interpretation of its onset. *Nature Physics*, 5:565–569, 2009.
- [301] L. Dougan, S. P. Bates, R. Hargreaves, J. P. Fox, J. Crain, J. L. Finney, V. Réat, and A. K. Soper. Methanol-water solutions: A bi-percolating liquid mixture. *Journal of Chemical Physics*, 121(13):6456–6462, 2004.
- [302] L. Dougan, R. Hargreaves, S. P. Bates, J. L. Finney, V. Réat, A. K. Soper, and J. Crain. Segregation in aqueous methanol enhanced by cooling and compression. *Journal of Chemical Physics*, 122(17):174517, 2005.

- [303] Maurice R. Eftink. The use of fluorescence methods to monitor unfolding transitions in proteins. *Biophysical Journal*, 66:482–501, 1998.
- [304] I. Herrada, P. Barthe, M. Van Heusden, K. DeGuillen, L. Mammri, S. Delbecq, F. Rico, and C. Roumestand. Monitoring Unfolding of Titin I27 single- and bi-domain with High-Pressure NMR Spectroscopy. *Unpublished*, 2018.
- [305] E. Peggion, A. Cosani, M. Terbojevich, and G. Borin. Conformational studies on polypeptides. The effect of sodium perchlorate on the conformation of poly-L-lysine and of random copolymers of L-lysine and L-phenylalanine in aqueous solution. *Biopolymers*, 11(3):633–643, 1972.
- [306] Wolfgang Maison, Robert J. Kennedy, and Daniel S. Kemp. Chaotropic anions strongly stabilize short, N-capped uncharged peptide helices: A new look at the perchlorate effect. *Angewandte Chemie - International Edition*, 40(20):3819–3821, 2001.
- [307] Eliana K. Ascitutto, Ignacio J. General, Kang Xiong, Sanford A. Asher, and Jeffrey D. Madura. Sodium perchlorate effects on the helical stability of a mainly alanine peptide. *Biophysical Journal*, 98(2):186–196, 2010.
- [308] Yutaka Kuroda, Daizo Hamada, Toshiki Tanaka, and Yuji Goto. High helicity of peptide fragments corresponding to β -strand regions of β -lactoglobulin observed by 2D-NMR spectroscopy. *Folding and Design*, 1(4):255–263, 1996.
- [309] Masamichi Ikeguchi. Transient non-native helix formation during the folding of β -lactoglobulin. *Biomolecules*, 4(1):202–216, 2014.
- [310] Daizo Hamada, Shin-ichi Segawa, and Yuji Goto. Non-native α -helical intermediate in the refolding of β -lactoglobulin, a predominantly β -sheet protein. *Nat. Struct. Biol.*, 3(10):868–873, 1996.
- [311] George Chikenji and Macoto Kikuchi. What is the role of non-native intermediates of beta-lactoglobulin in protein folding? *Proceedings of the National Academy of Sciences*, 97(26):14273–14277, 2000.
- [312] Takahiro Okabe, Toshiaki Miyajima, Kanako Nakagawa, Seiichi Tsukamoto, Kazuo Fujiwara, and Masamichi Ikeguchi. Effect of non-native helix destabilization on the folding of equine β -lactoglobulin. *Journal of Biochemistry*, 156(5):291–297, 2014.
- [313] Eefei Chen, Mary Lou Everett, Zoie E. Holzknecht, Robert A. Holzknecht, Shu S. Lin, Dawn E. Bowles, and William Parker. Short-Lived α -Helical Intermediates in the Folding of β -Sheet Proteins. *Biochemistry*, 49(26):5609–5619, 2010.
- [314] David G. Reid. *Protein NMR Techniques*, volume 60. Humana Press, New Jersey, jul 1997.
- [315] Megan L. Hughes. *Determining the importance of hydrogen bond strength on the mechanical stability of proteins* Megan Lianne. PhD thesis, University of Leeds, 2015.
- [316] Xin Chen, Tinglu Yang, Sho Kataoka, and Paul S. Cremer. Specific Ion Effects on Interfacial Water Structure near Macromolecules. *J. Am. Chem. Soc.*, 129:12272–12279, 2007.

- [317] Hiranmay Maity, Aswathy N. Muttathukattil, and Govardhan Reddy. Salt Effects on Protein Folding Thermodynamics. *Journal of Physical Chemistry Letters*, 9(17):5063–5070, 2018.
- [318] David Sutcliffe. The ionic composition of surface waters in the english lake district, related to bedrock geology, with some singular facts and speculation on the existence of mineral-rich groundwaters. 1998.
- [319] José A. Caro and A. Joshua Wand. Practical aspects of high-pressure NMR spectroscopy and its applications in protein biophysics and structural biology. *Methods*, 148:67–80, sep 2018.
- [320] Ryo Kitahara, Kazumi Hata, Hua Li, Mike P. Williamson, and Kazuyuki Akasaka. Pressure-induced chemical shifts as probes for conformational fluctuations in proteins. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 71:35–58, may 2013.
- [321] A. K. Soper. Inelasticity corrections for time-of-flight and fixed wavelength neutron diffraction experiments. *Molecular Physics*, 107(16):1667–1684, 2009.
- [322] Tristan Youngs. Dissolve: next generation software for the interrogation of total scattering data by empirical potential generation. *Molecular Physics*, pages 1–14, 2019.