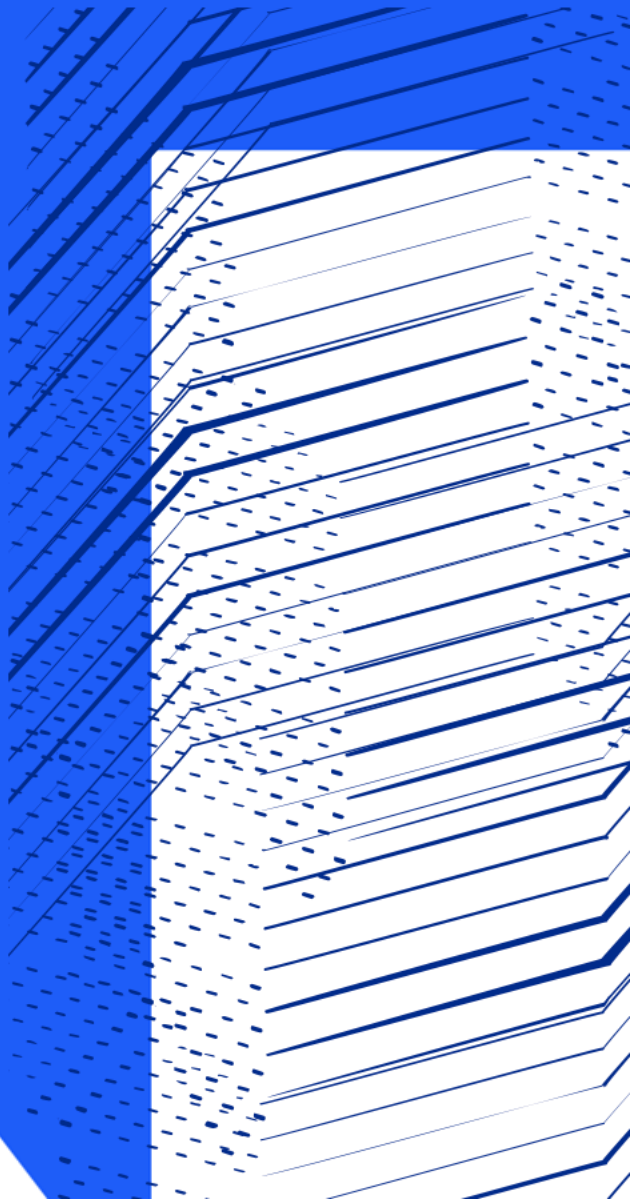
Science and Technology Facilities Council

ISIS Neutron and Muon Source

EPSR & Dissolve

Data-driven structural modelling of total scattering data

Dr Tristan Youngs
Disordered Materials Group
tristan.youngs@stfc.ac.uk



Science and Technology Facilities Council

ISIS Neutron and Muon Source

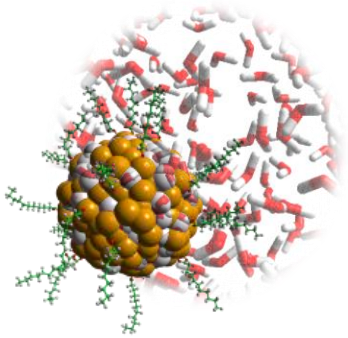
Disordered Materials and Total Scattering

Disordered Materials

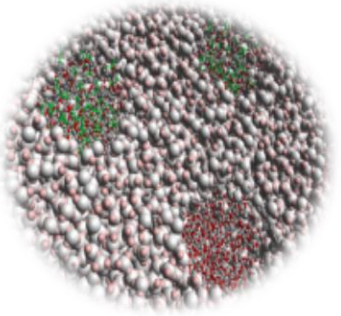
- Show predominantly local ordering (i.e. on the scale of atoms or molecules)
- Show primarily diffuse scattering (“soft” features in measured data)
- Are typically dynamic (e.g. liquids) – average structural picture is obtained
- May show long-range order (i.e. Bragg scattering)
- May contain large length-scale correlations (e.g. micelles, lamellae)

Liquids

Glasses

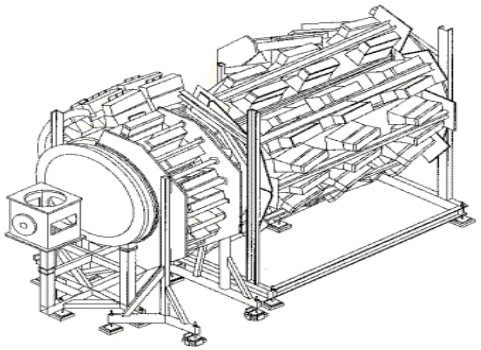


Solutions



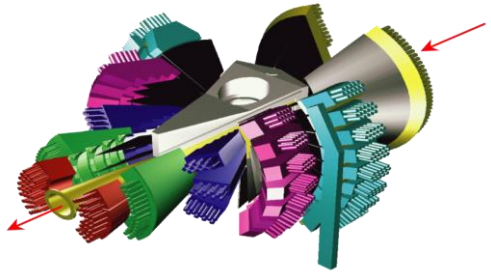
Condensed phases under confinement

Total Scattering Instruments @ ISIS



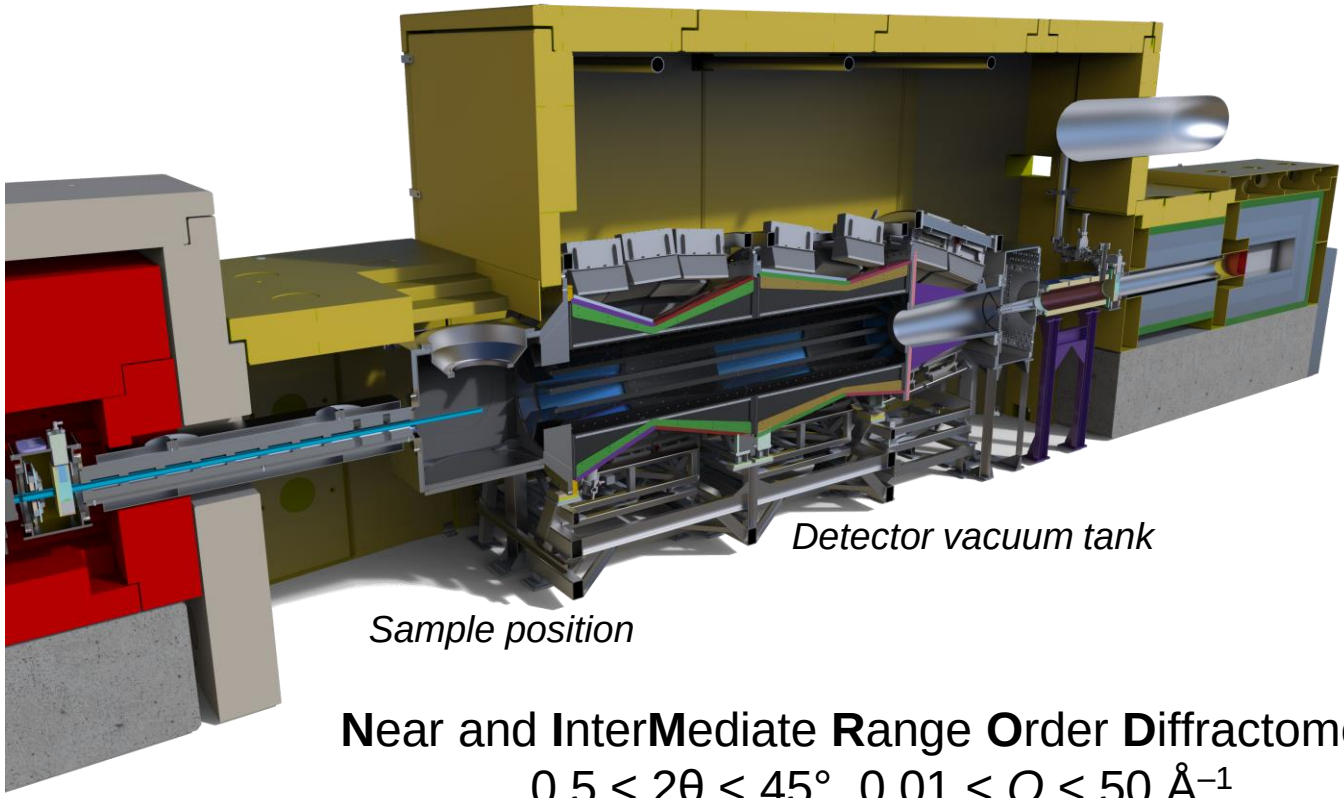
Small **A**ngle **N**eutron **D**iffractometer for **A**morphous and **L**iquid **S**amples

$3 < 2\theta < 38^\circ$, $0.1 < Q < 50$ Å^{-1}



General **M**aterials Diffractometer

$1.21 < 2\theta < 171.4^\circ$, $0.04 < Q < 50$ Å^{-1}



Near and **I**nter**M**ediate **R**ange **O**rder **D**iffractometer

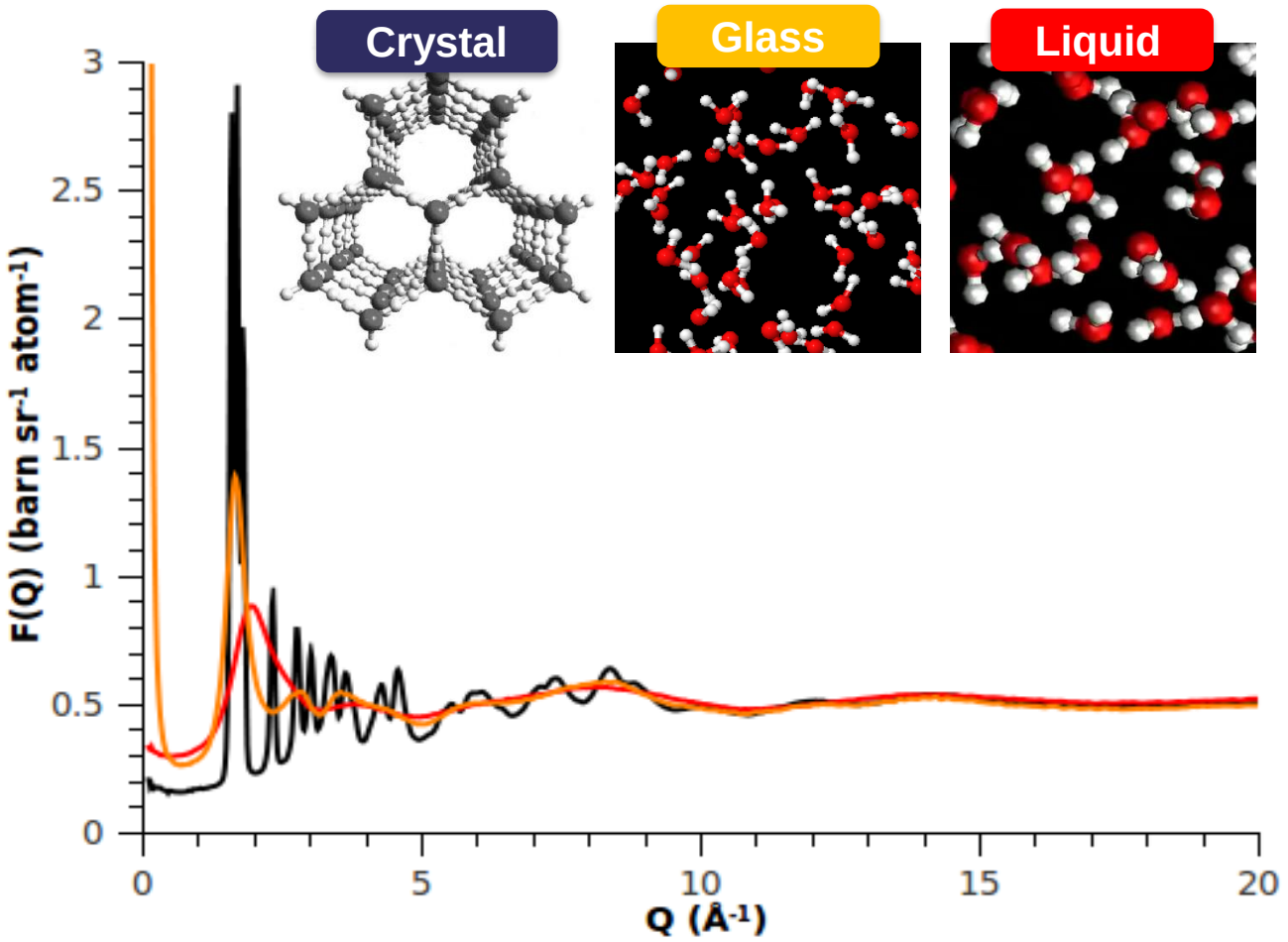
$0.5 < 2\theta < 45^\circ$, $0.01 < Q < 50$ Å^{-1}

All three instruments exist to probe and understand material structure

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Total Structure Factor F(Q)



- Contains all correlation information between all “objects” in the system
- Also includes Bragg scattering
- Also includes any SANS

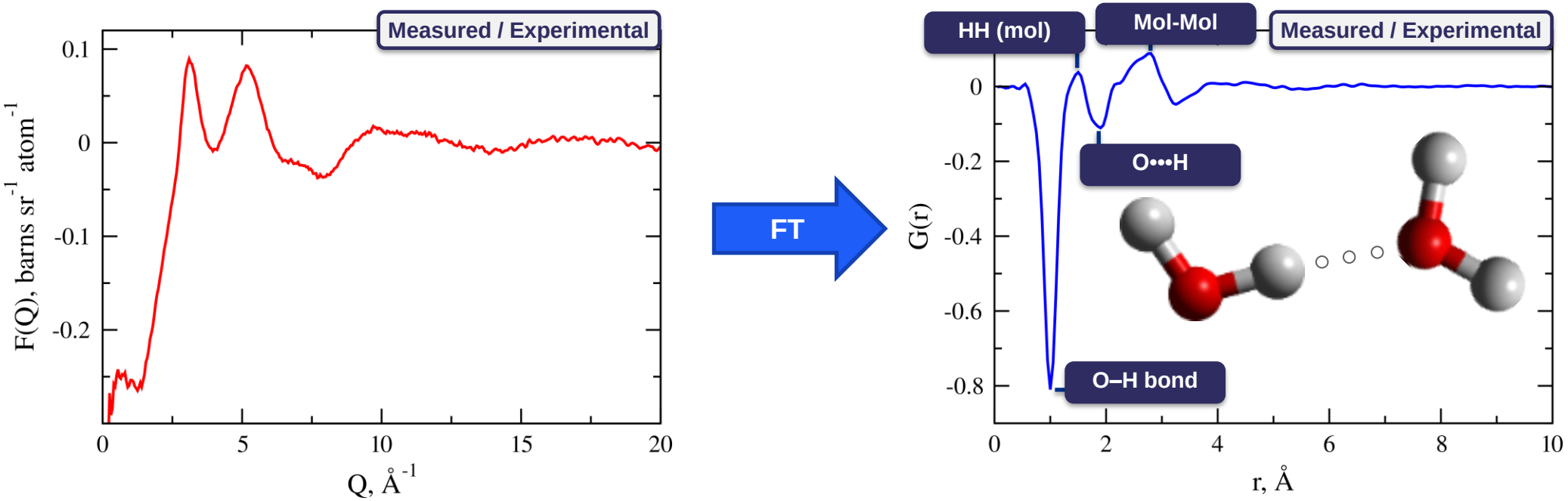
“A single dataset encompassing structural information on the target sample, no matter the phase, complexity, or composition of the system.”

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Interpreting F(Q)

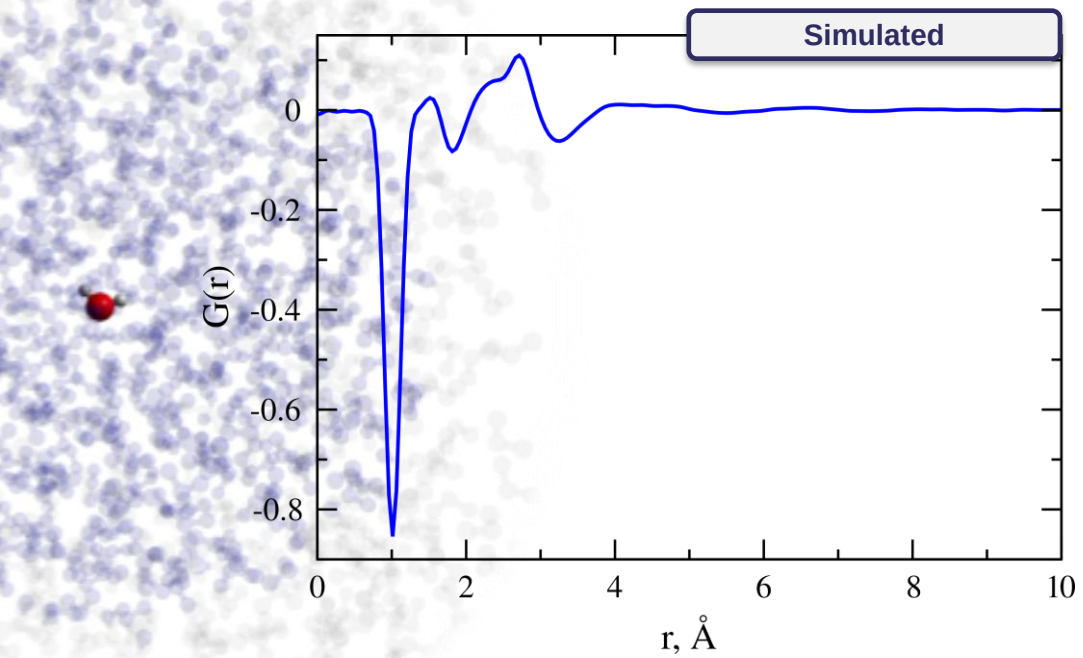
Can Fourier transform data from *Q*-space (instrument) to *r*-space (real)



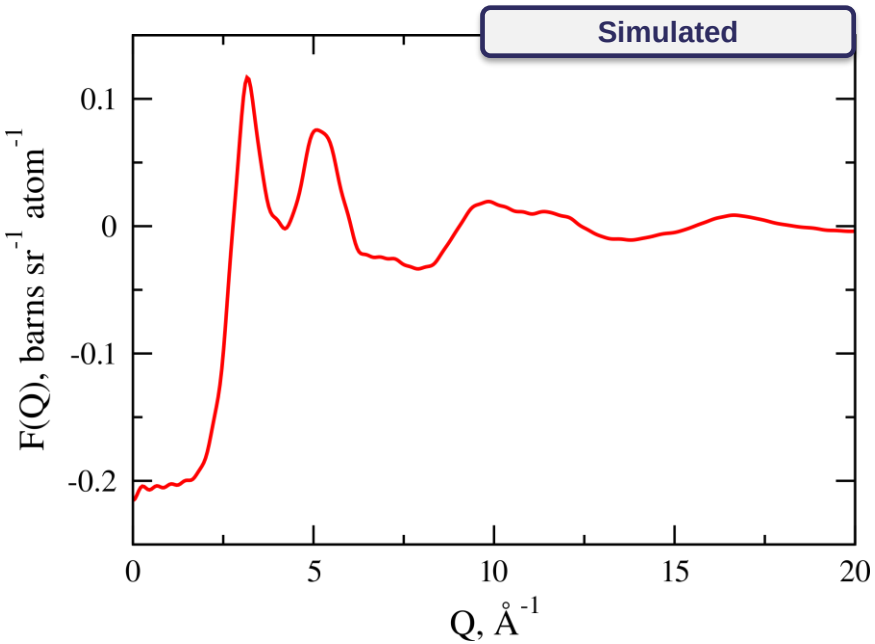
Non-trivial to analyse by inspection. Angular correlations? 3D structure?

Simulating F(Q)

Atomistic simulation – molecular dynamics, Monte Carlo – using off-the-shelf forcefield.



FT

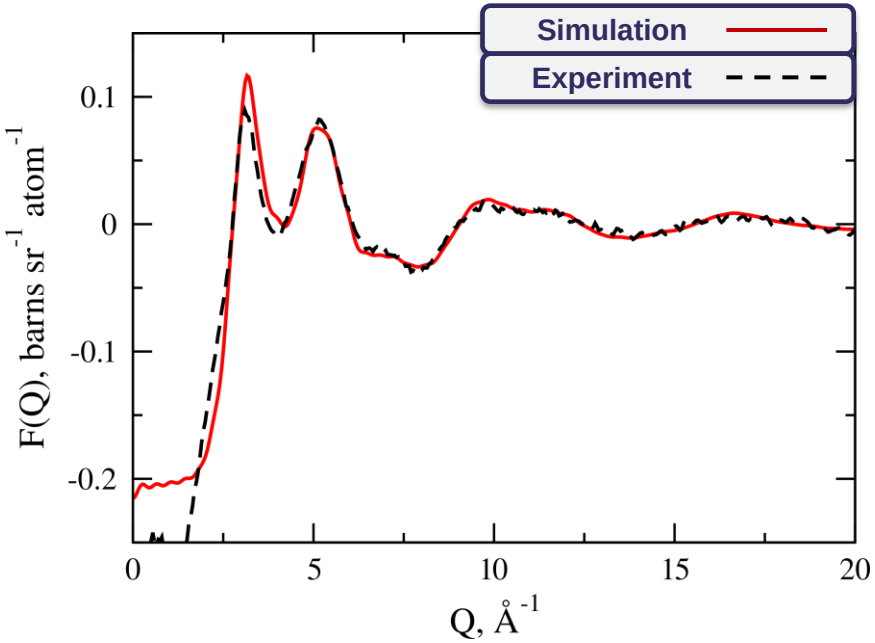


Can calculate any correlation I want from a simulation, but does it reflect reality?

Simulation vs Reality

Forcefields can give results close to experiment, but often show discrepancies.

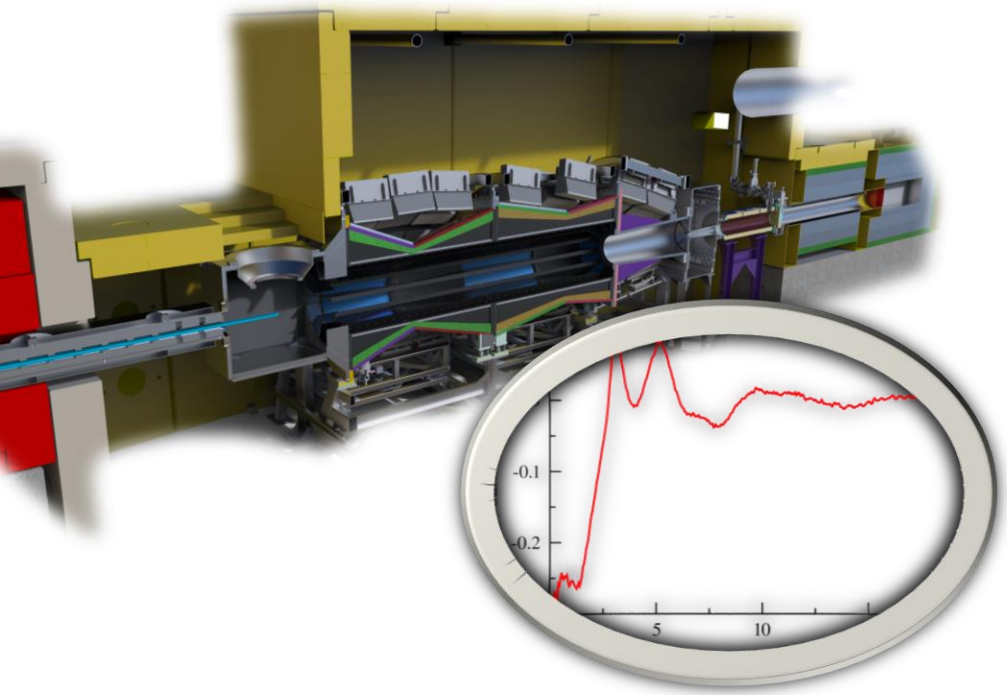
- Parameterised against phase data etc.
- Rarely against bulk structure data

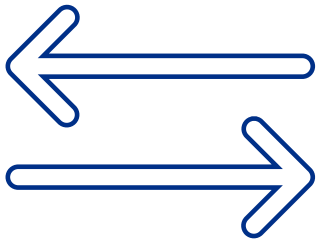


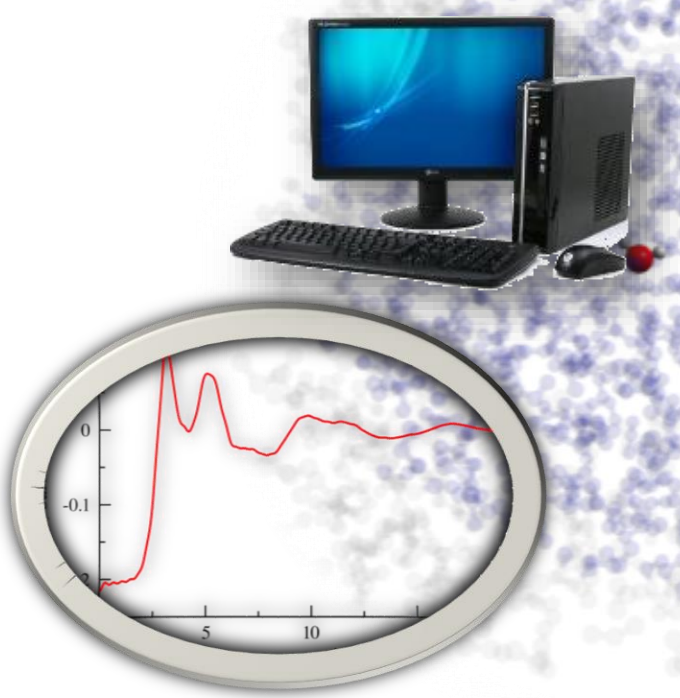
Solution?

- Modify the forcefield to improve it
- By hand? Tedious, impractical…
- Automatically, using the data? How?

Joining Simulation and Experiment







- How can I form a simulation consistent with the experimental data?
- How can I expand my experimental data to reduce underdetermination?

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Data-Driven Refinement

The Goal

- Make a simulation of a system of arbitrary complexity
- Compare available experimental F(Q) with simulated F(Q)

- **Adjust the underlying forcefield to improve agreement**

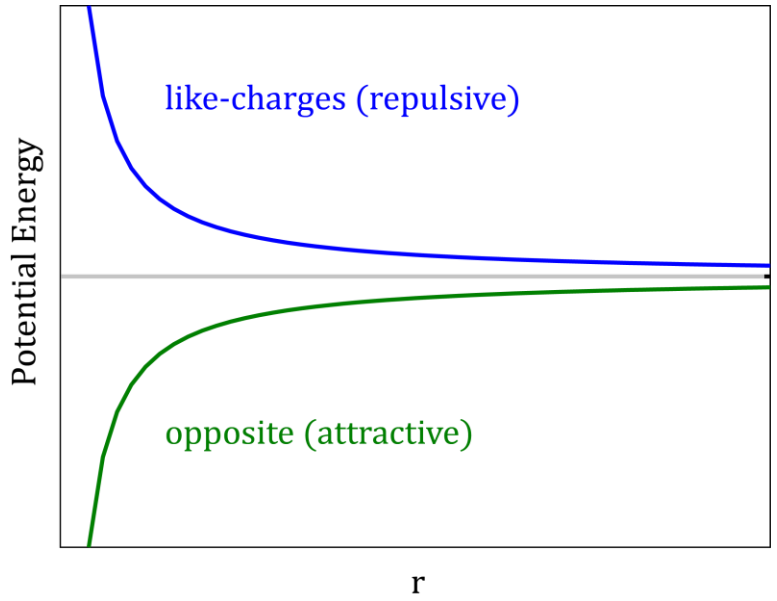
- Calculate structural properties of interest
- Write a paper
- Go home

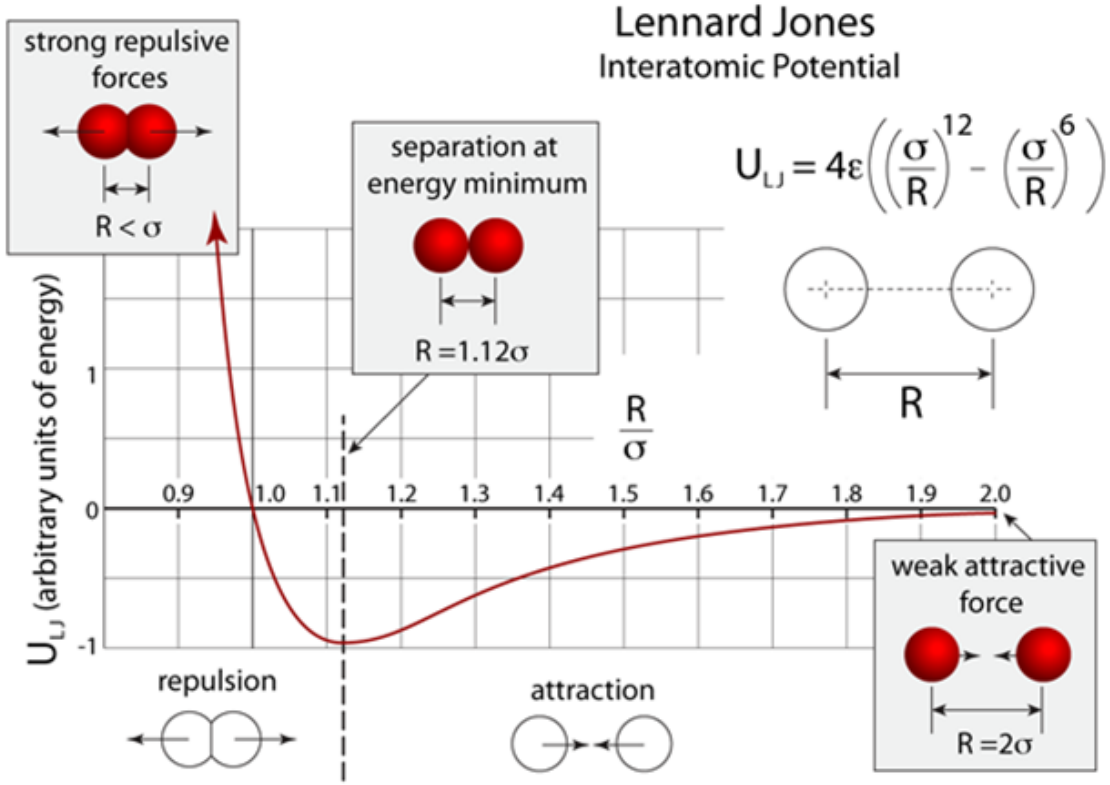
The Target: Pair Potentials

- Describes the interactions between atoms “through space”
- Parameters from:
  - Existing forcefields (LJ+q)
  - Calculated via QM / DFT (q)

Coulomb:

$$U = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}^2}$$





Science and Technology Facilities Council

ISIS Neutron and Muon Source

Constructing the F(Q)

$$F(Q) = \sum_{i,j} (2 - \delta_{ij}) c_i c_j b_i b_j S_{ij}(Q)$$

Total Structure Factor
(Experimental Observable)

$$c_i = \frac{n_i}{N} \qquad b_i = scattering\ length$$

$$S_{ij}(Q) = \rho \int_0^\infty 4\pi r^2 g_{ij}(r) \frac{\sin Qr}{Qr} \mathrm{d}r$$

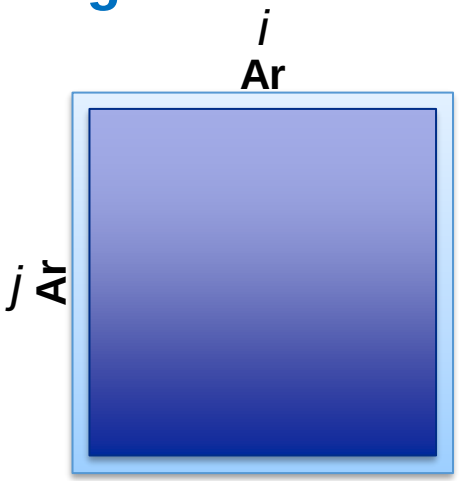
Partial Structure Factor

(Simulation Observable)

Atom Types

- A specific kind of atom *in the simulation*
  - Depends at least on the element. Can be split by chemical environment.
  - Does not depend on isotope...

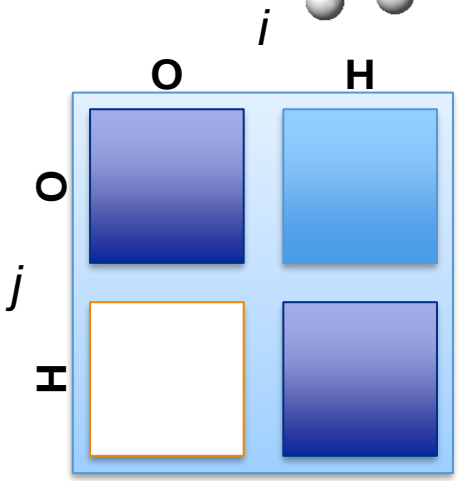
Argon



***N* = 1**

Also N_2, O_2…

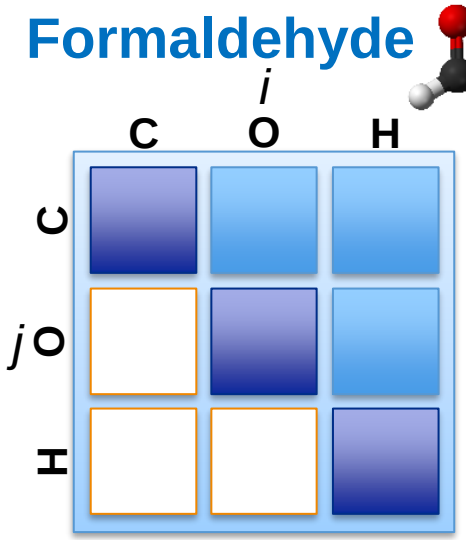
Water



***N* = 3**

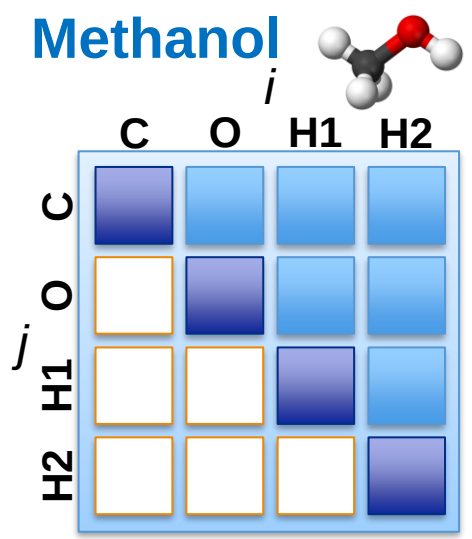
Also benzene, silica…

Formaldehyde



***N* = 6**

Methanol



***N* = 10**

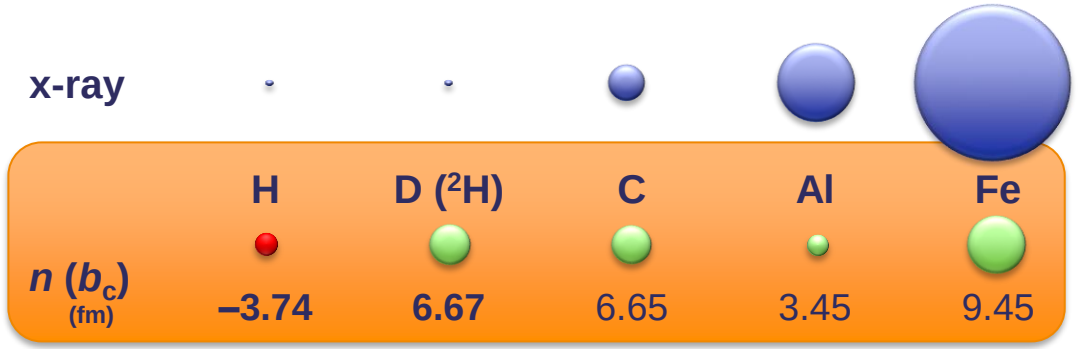
Science and Technology Facilities Council

ISIS Neutron and Muon Source

$$F(Q) = \sum_{i,j} (2 - \delta_{ij}) c_i c_j b_i b_j S_{ij}(Q)$$

Isotopic Substitution

$$F(Q) = \sum_{i,j} (2 - \delta_{ij}) c_i c_j \underbrace{\boldsymbol{b_i b_j}}_{\text{Scattering Lengths}} S_{ij}(Q)$$

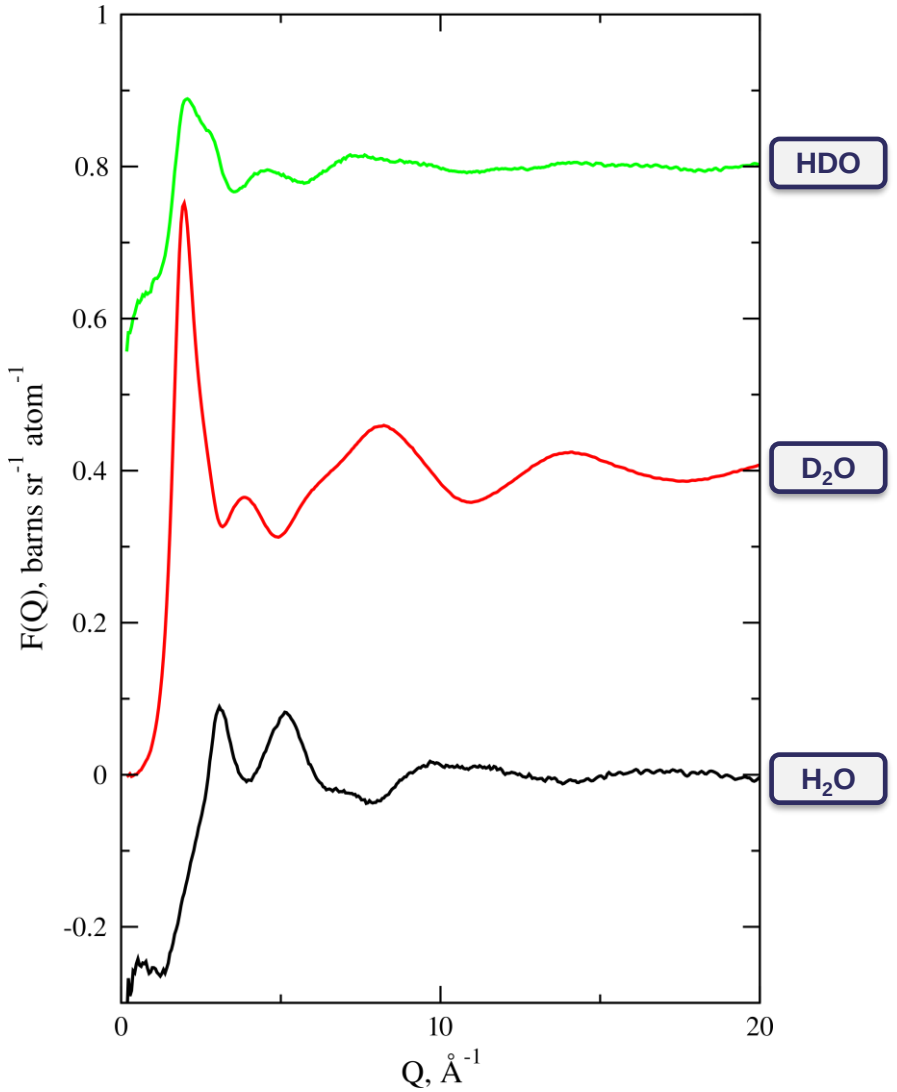


- Partial $S(Q)$ weighted by coherent scattering length, $\boldsymbol{b}$
- For neutrons, $\boldsymbol{b}$ is dependent on isotope
  - e.g. ^{2}H for H, ^{15}N for N, ^{6}Li for Li
- Key assumption: structure is independent of isotopes used
- Perform multiple measurements on the same system, with different isotopic substitutions

Isotopic Water

- Swapping H for ^2H (D) we can get three distinct datasets focussing on one atom type
  - H_2O
  - D_2O
  - 50:50 mix of H_2O and D_2O

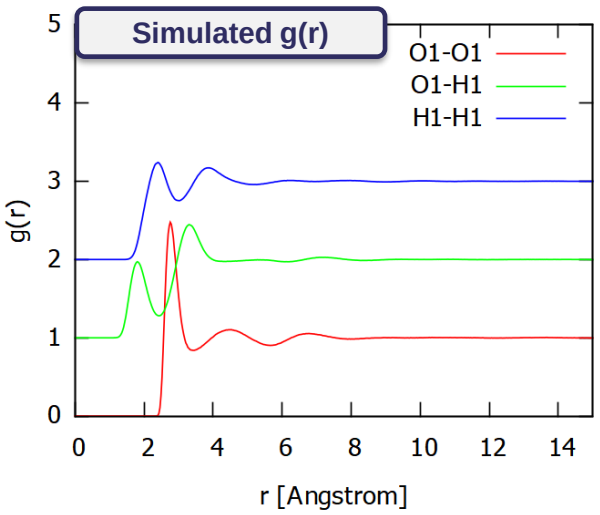
Swapping isotopes always needs to involve enough atoms to produce a noticeable change in the F(Q)



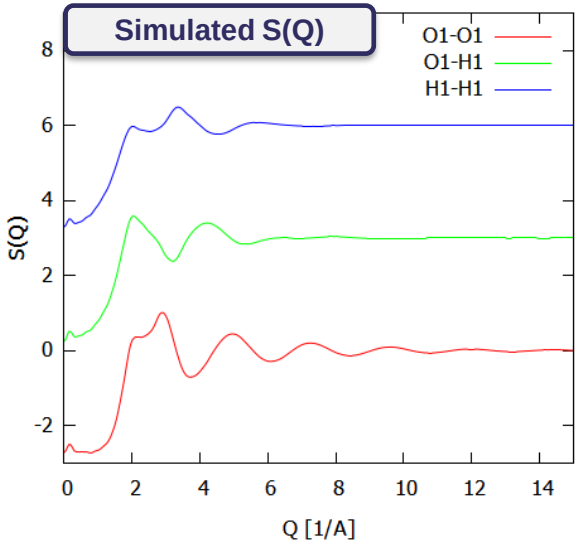
Science and Technology Facilities Council

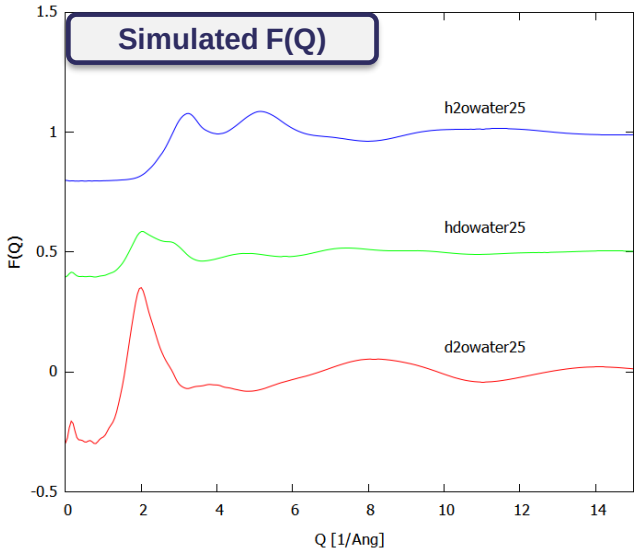
ISIS Neutron and Muon Source

g(r) to S(Q) to F(Q) and back again?



FT





$$w_{ij} = c_i c_j \boldsymbol{b_i} \boldsymbol{b_j}$$

$$\begin{bmatrix} 0.0374 & -0.0964 & 0.0622 \\ 0.0374 & 0.1722 & 0.1980 \\ 0.0374 & 0.0378 & 0.0096 \end{bmatrix} \begin{bmatrix} S_{OO}(Q) \\ S_{OH}(Q) \\ S_{HH}(Q) \end{bmatrix} = \begin{bmatrix} F_{H_2O}(Q) \\ F_{D_2O}(Q) \\ F_{HDO}(Q) \end{bmatrix}$$

Invert

$$\begin{bmatrix} 4.8238 & -2.7037 & 24.6061 \\ -5.8227 & 1.6265 & 4.1962 \\ 4.1525 & 4.1525 & -8.3050 \end{bmatrix} \begin{bmatrix} F_{H_2O}(Q) \\ F_{D_2O}(Q) \\ F_{HDO}(Q) \end{bmatrix} = \begin{bmatrix} S_{OO}(Q) \\ S_{OH}(Q) \\ S_{HH}(Q) \end{bmatrix}$$

Science and Technology Facilities Council

ISIS Neutron and Muon Source

The General Case

- For an “ideal” system such as H_2O where enough isotopic substitutions can be made, direct matrix inversion is possible
- What about cases where one or more partials only contribute weakly?
- What about cases where not enough isotopic substitutions can be made?

For example, silica:

$$\begin{bmatrix} F_{SiO_2}(Q) \end{bmatrix} = \begin{bmatrix} 0.0191 & 0.1070 & 0.1497 \end{bmatrix} \begin{bmatrix} S_{SiSi}(Q) \\ S_{SiO}(Q) \\ S_{OO}(Q) \end{bmatrix}$$

No inversion possible, so no route from $F(Q)$ to $S(Q)$ (and $g(r)$)

The Augmented Scattering Matrix

- We have a simulation which we assume gives us a good "guess" of the $F(Q)$ and hence a good guess for the partial $S(Q)$
- Define a feedback factor, $0 < f < 1$, and write new weighting factors

$$w_{ij} = f c_i c_j b_i b_j$$

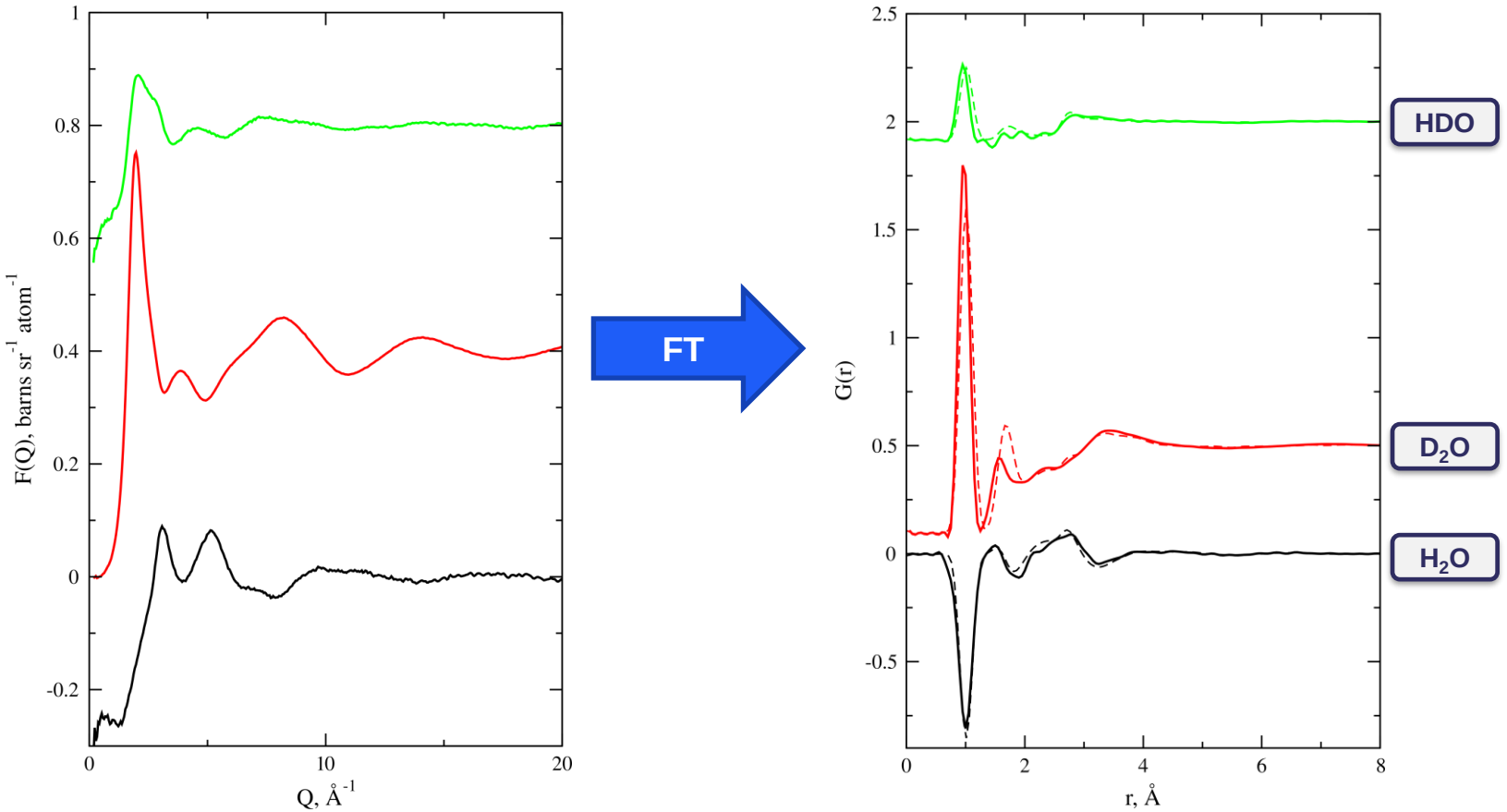
$$\begin{bmatrix} F_{SiO_2}(Q) \\ S^{sim}_{SiSi}(Q) \\ S^{sim}_{SiO}(Q) \\ S^{sim}_{OO}(Q) \end{bmatrix} = \begin{bmatrix} 0.0172 & 0.0963 & 0.1347 \\ 0.1 & 0.0 & 0.0 \\ 0.0 & 0.1 & 0.0 \\ 0.0 & 0.0 & 0.1 \end{bmatrix} \begin{bmatrix} S_{SiSi}(Q) \\ S_{SiO}(Q) \\ S_{OO}(Q) \end{bmatrix}$$

$$w_{ij} = (1 - f)$$

- Now we have an overdetermined matrix for which we can find a pseudoinverse

What About Bound Terms?

- Typically dealing with molecular species, so we have chemical bonds (and angles, torsions etc.)
- Assumed / forcefield-obtained values may not reflect reality…



Science and Technology Facilities Council
ISIS Neutron and Muon Source

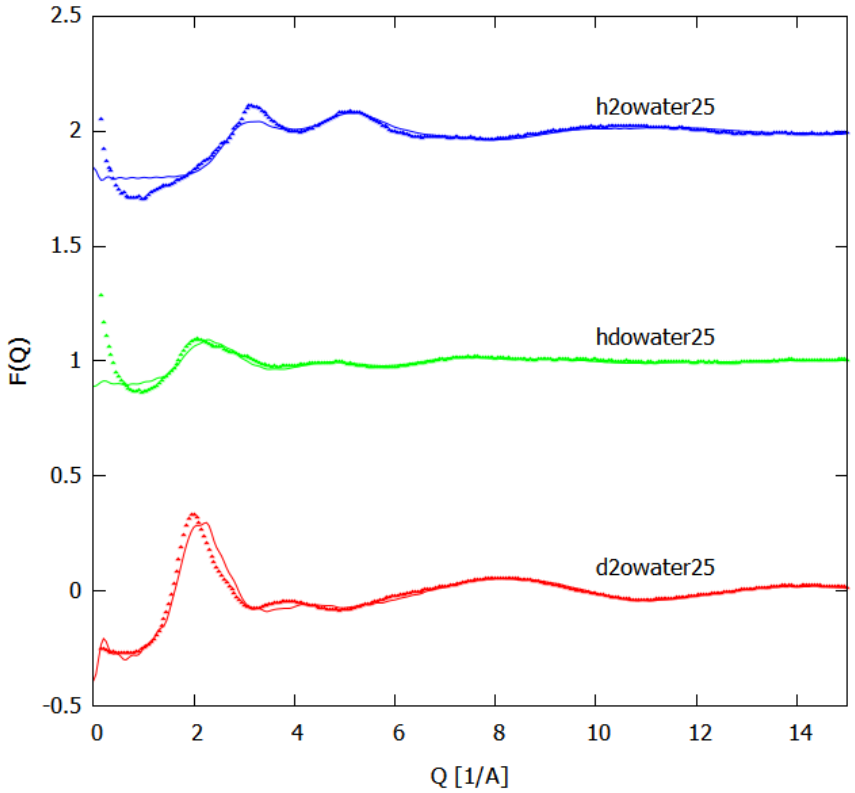
Science and Technology Facilities Council

ISIS Neutron and Muon Source

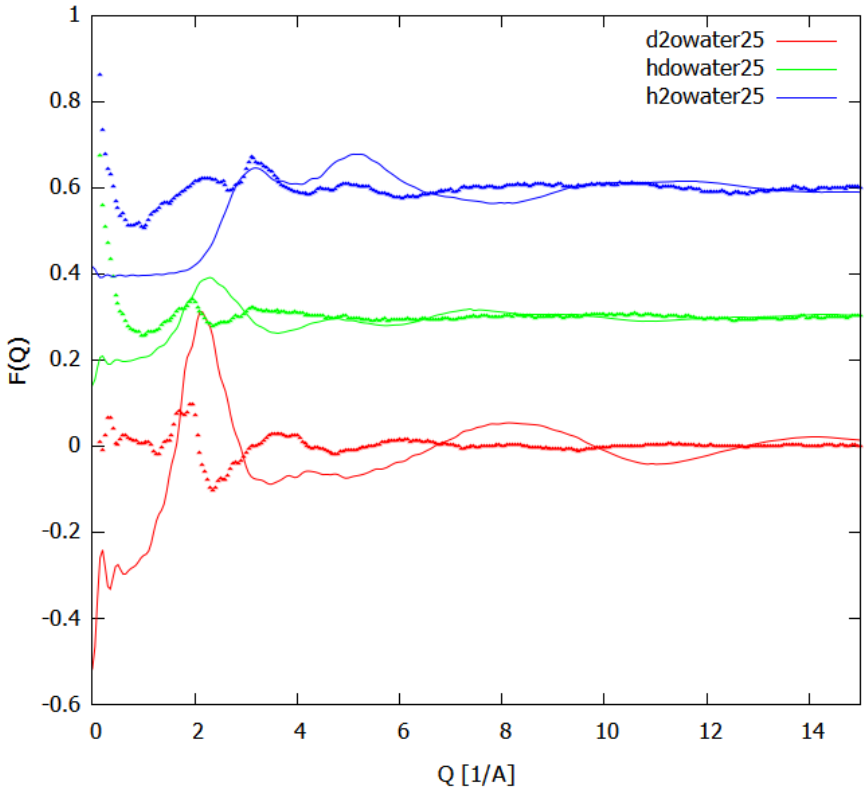
Empirical Potential Structure Refinement

The EPSR Method

1) Take differences between experimental and simulated $F(Q)$ to get $\Delta F(Q)$



F(Q) from simulation and experiment



ΔF(Q) between simulation and experiment

Science and Technology Facilities Council

ISIS Neutron and Muon Source

The EPSR Method

1) Take differences between experimental and simulated $F(Q)$ to get $\Delta F(Q)$
2) Enter these $\Delta F(Q)$ into the inverse scattering matrix to generate $\Delta S(Q)$

$$\begin{bmatrix} \Delta S_{OO}(Q) \\ \Delta S_{OH}(Q) \\ \Delta S_{HH}(Q) \end{bmatrix} = \begin{bmatrix} 4.8238 & -2.7037 & 24.6061 \\ -5.8227 & 1.6265 & 4.1962 \\ 4.1525 & 4.1525 & -8.3050 \end{bmatrix} \begin{bmatrix} \Delta F_{H_2O}(Q) \\ \Delta F_{D_2O}(Q) \\ \Delta F_{HDO}(Q) \end{bmatrix}$$

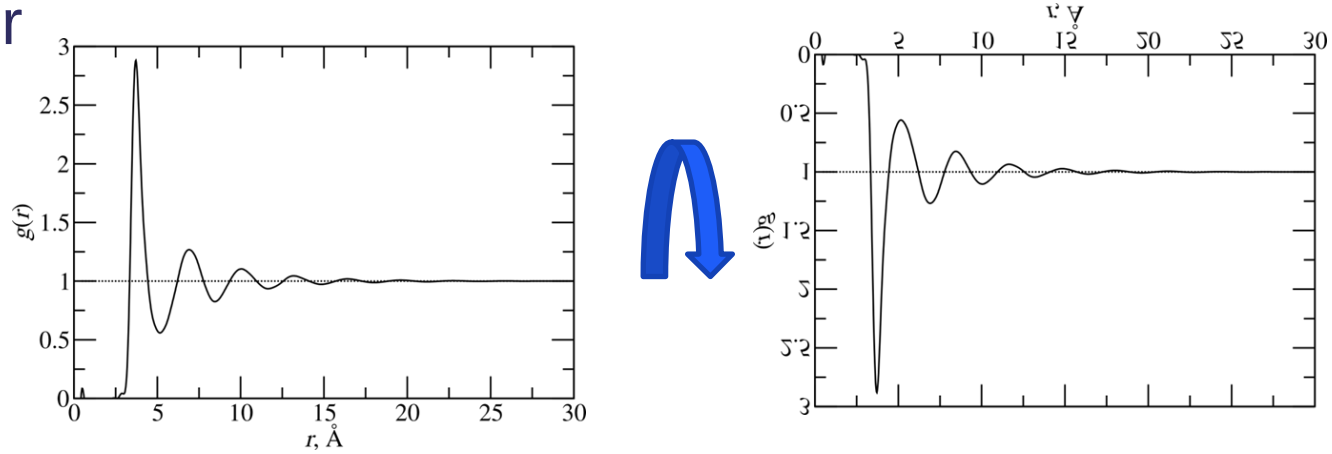
$$\begin{bmatrix} S_{OO}(Q) \\ S_{OH}(Q) \\ S_{HH}(Q) \end{bmatrix} = \begin{bmatrix} 4.8238 & -2.7037 & 24.6061 \\ -5.8227 & 1.6265 & 4.1962 \\ 4.1525 & 4.1525 & -8.3050 \end{bmatrix} \begin{bmatrix} F_{H_2O}(Q) \\ F_{D_2O}(Q) \\ F_{HDO}(Q) \end{bmatrix}$$

The EPSR Method

1) Take differences between experimental and simulated $F(Q)$ to get $\Delta F(Q)$
2) Enter these $\Delta F(Q)$ into the inverse scattering matrix to generate $\Delta S(Q)$

$$\begin{bmatrix} \Delta S_{OO}(Q) \\ \Delta S_{OH}(Q) \\ \Delta S_{HH}(Q) \end{bmatrix} = \begin{bmatrix} 4.8238 & -2.7037 & 24.6061 \\ -5.8227 & 1.6265 & 4.1962 \\ 4.1525 & 4.1525 & -8.3050 \end{bmatrix} \begin{bmatrix} \Delta F_{H_2O}(Q) \\ \Delta F_{D_2O}(Q) \\ \Delta F_{HDO}(Q) \end{bmatrix}$$

3) Transform the $\Delta S(Q)$ into $\Delta g(r)$ and use these to form an additional, empirical potential for each atom type pair



Science and Technology Facilities Council

ISIS Neutron and Muon Source

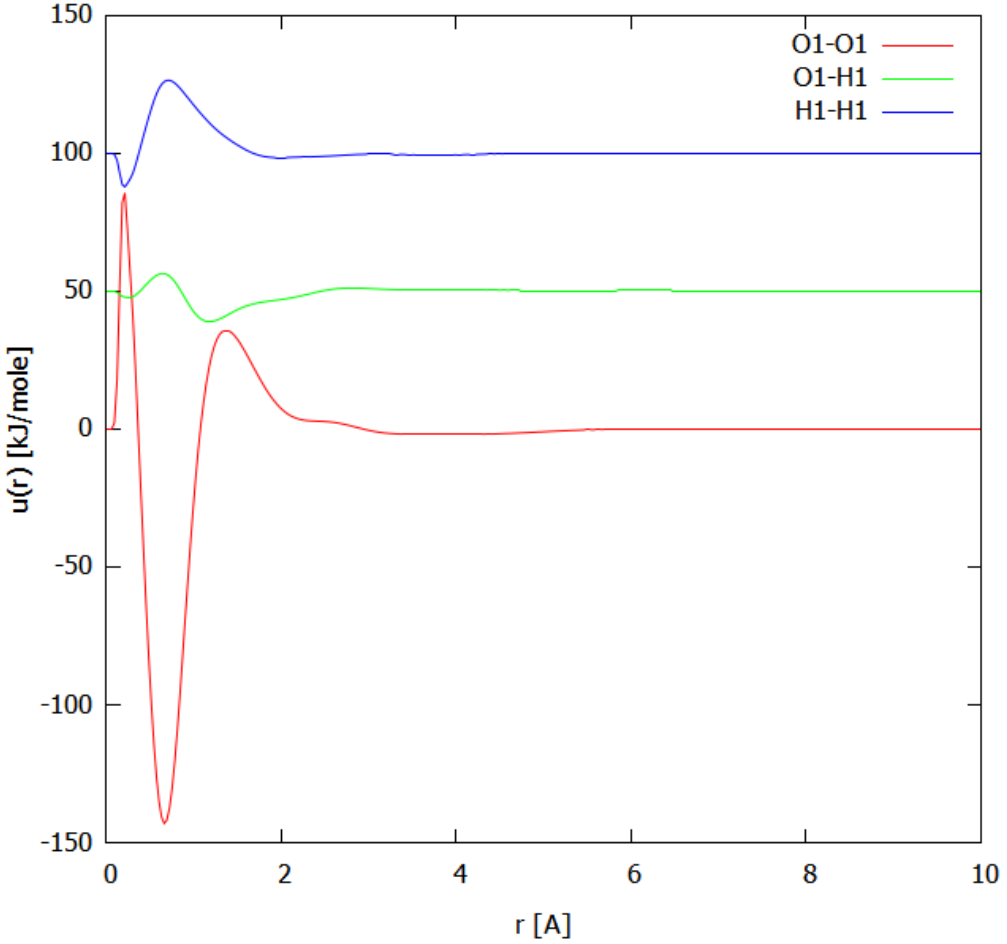
The EPSR Method

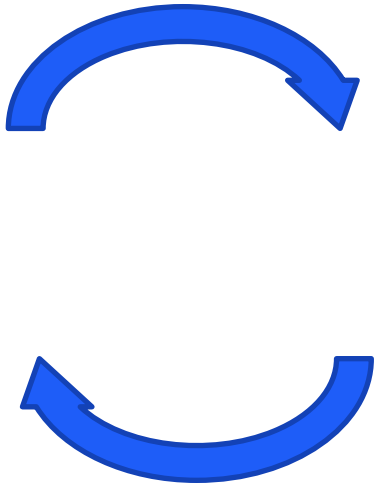
1) Take differences between experimental and simulated $F(Q)$ to get $\Delta F(Q)$
2) Enter these $\Delta F(Q)$ into the inverse scattering matrix to generate $\Delta S(Q)$

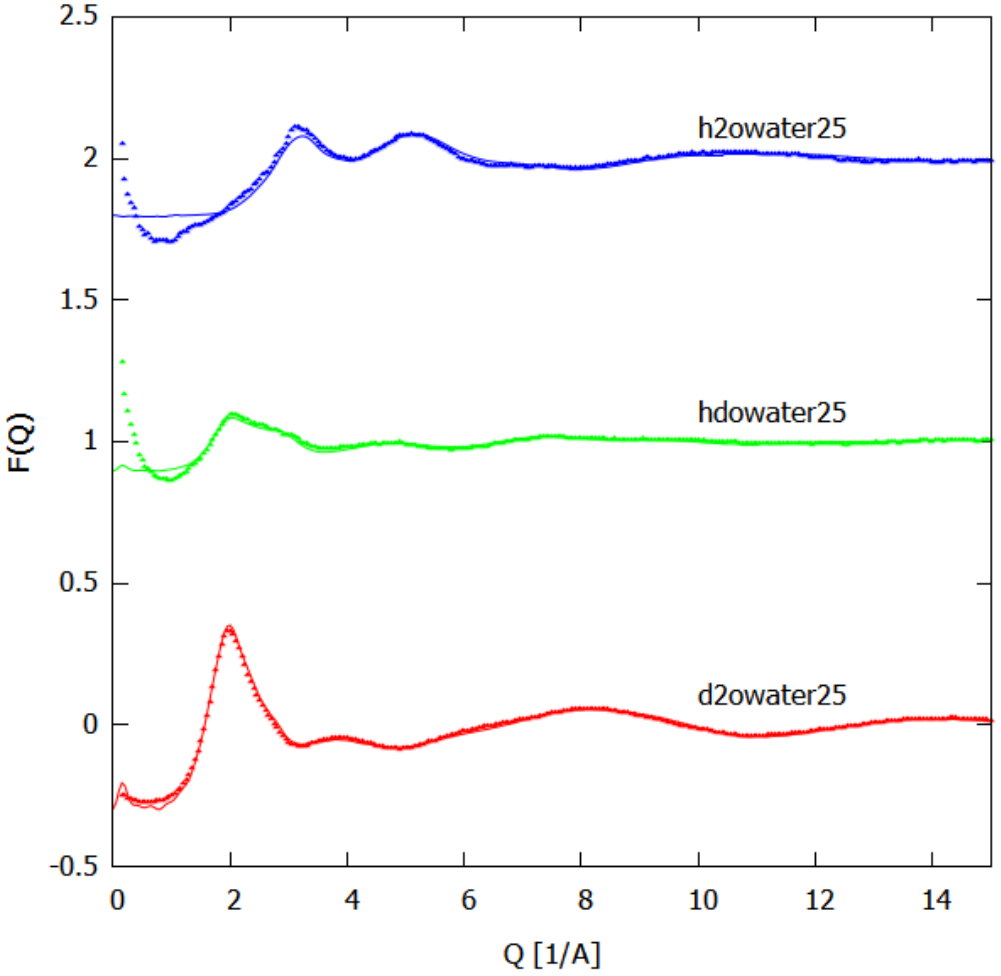
$$\begin{bmatrix} \Delta S_{OO}(Q) \\ \Delta S_{OH}(Q) \\ \Delta S_{HH}(Q) \end{bmatrix} = \begin{bmatrix} 4.8238 & -2.7037 & 24.6061 \\ -5.8227 & 1.6265 & 4.1962 \\ 4.1525 & 4.1525 & -8.3050 \end{bmatrix} \begin{bmatrix} \Delta F_{H_2O}(Q) \\ \Delta F_{D_2O}(Q) \\ \Delta F_{HDO}(Q) \end{bmatrix}$$

3) Transform the $\Delta S(Q)$ into $\Delta g(r)$ and use these to form an additional, empirical potential for each atom type pair
4) Repeatedly run the simulation and refine **additional potentials** until the experimental and simulated $F(Q)$ 'match'

Potential Refinement: EPSR results



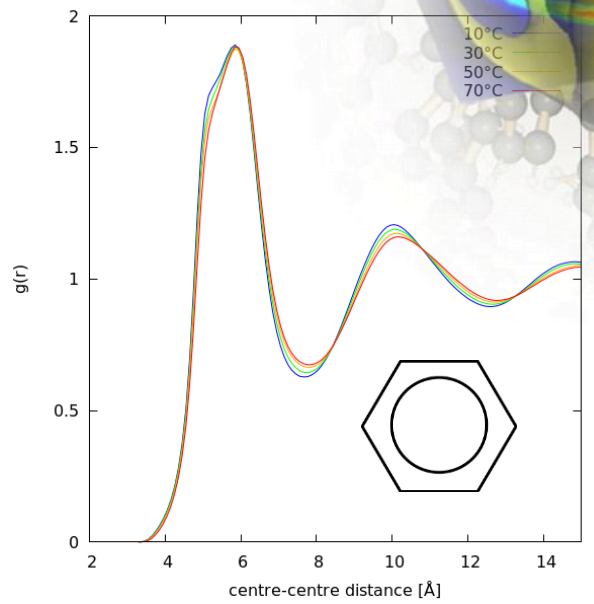


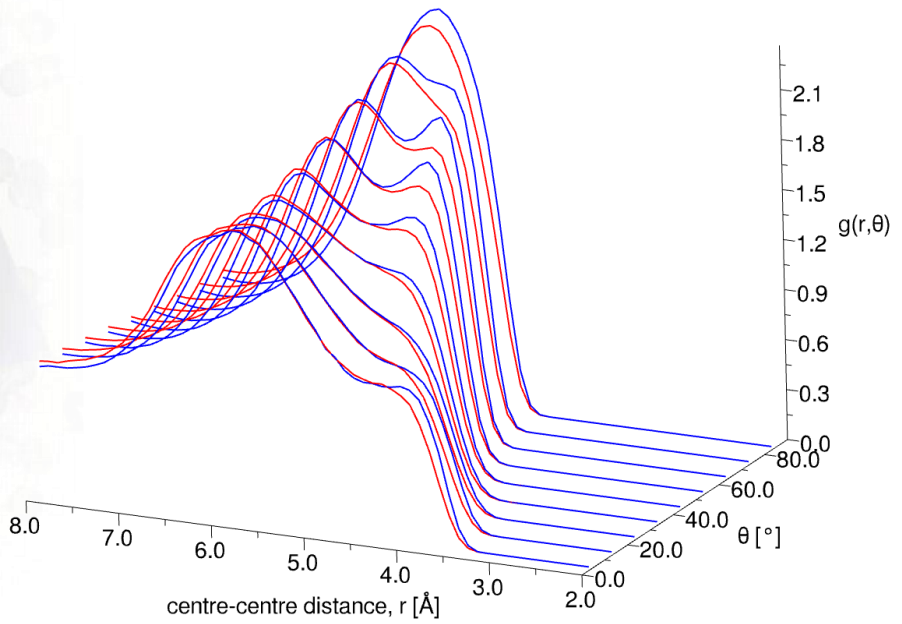


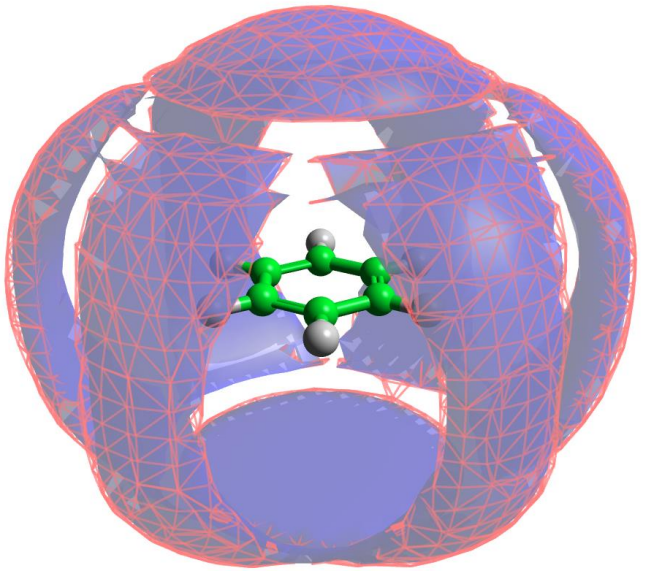
Science and Technology Facilities Council

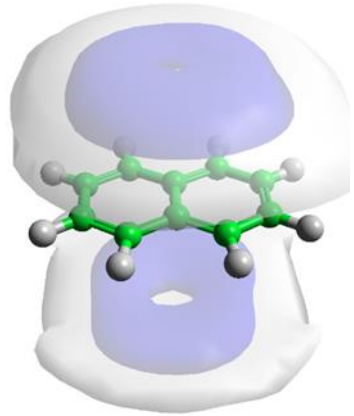
ISIS Neutron and Muon Source

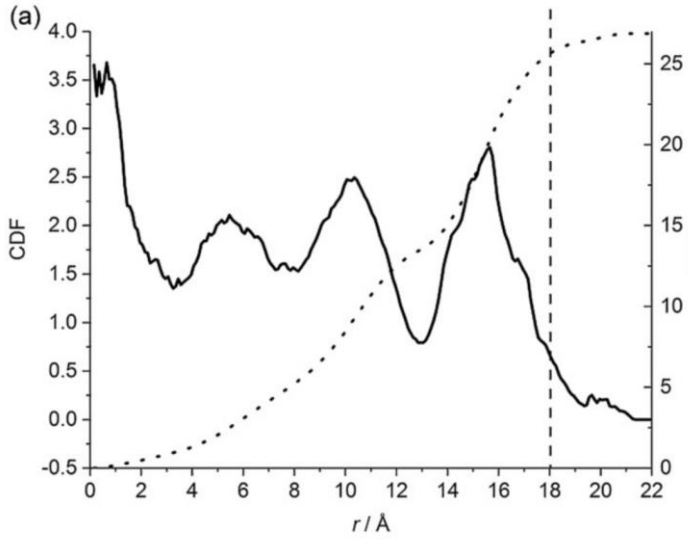
Analyse the Refined Simulation

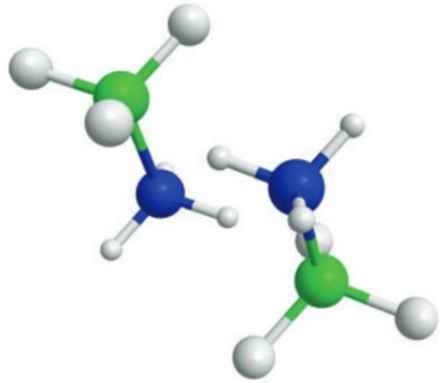


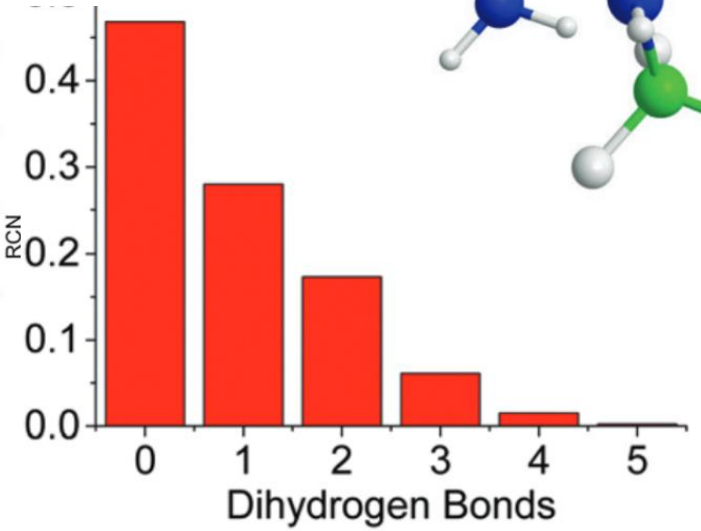












Science and Technology Facilities Council

ISIS Neutron and Muon Source

Science and Technology Facilities Council

ISIS Neutron and Muon Source

The Tools & Their Limits

EPSR & Dissolve

Empirical Potential Structure Refinement (EPSR)

- Developed by Prof. A. K. Soper
- Fortran / OpenMP
- Monte Carlo, custom forcefield, 140k atoms

https://www.isis.stfc.ac.uk/Pages/Empirical-Potential-Structure-Refinement.aspx

Dissolve

- Developed by Team Dissolve
- C++ / TBB / (OpenMPI)
- Monte Carlo / MD, standard forcefield, 3M+ atoms

https://www.projectdissolve.com

†With higher confidence than a simulation alone

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Data Quality

Both EPSR and Dissolve use (currently) the same procedure of matrix & potential inversion.

Theory can link measured scattering intensities to real-space simulation data – but is **idealised**

Data must be well processed:

- Sample environment and instrument backgrounds & containers removed
- Remove multiple scattering and beam attenuation effects
- Remove self-scattering – interference scattering only
- Remove inelastic scattering: important (and difficult) for light H.
- Have a normalised intensity on an absolute scale (barns sr^{-1} $atom^{-1}$)

It’s Just a Simulation!

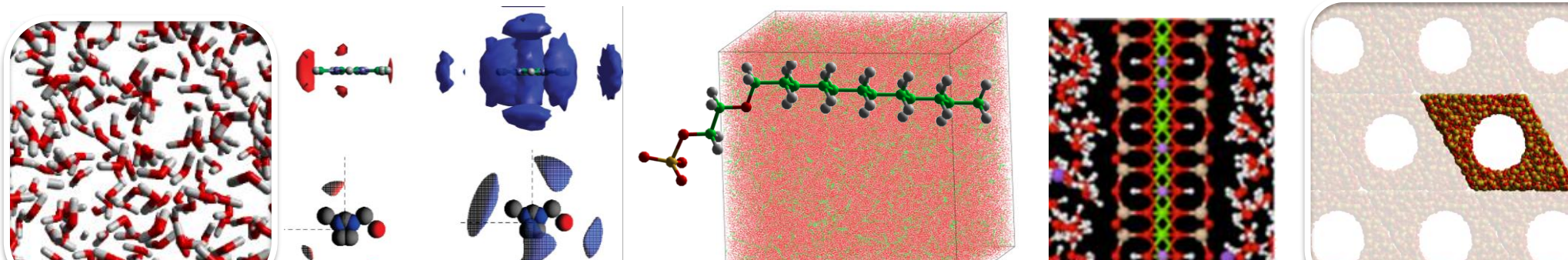
Perfect reproduction of data is not possible

1. Unknown (or not completely removed) instrument effects
2. Limitations in forcefield functional forms / potentials
3. Finite simulation size
4. Real-world systems are not 100% pure or 100% isotopically substituted

Can arguably show improvement over pure simulation, even if underdetermined, and has been successfully applied to a huge variety of systems…

Science and Technology Facilities Council

ISIS Neutron and Muon Source



Confined Fluid Mixtures

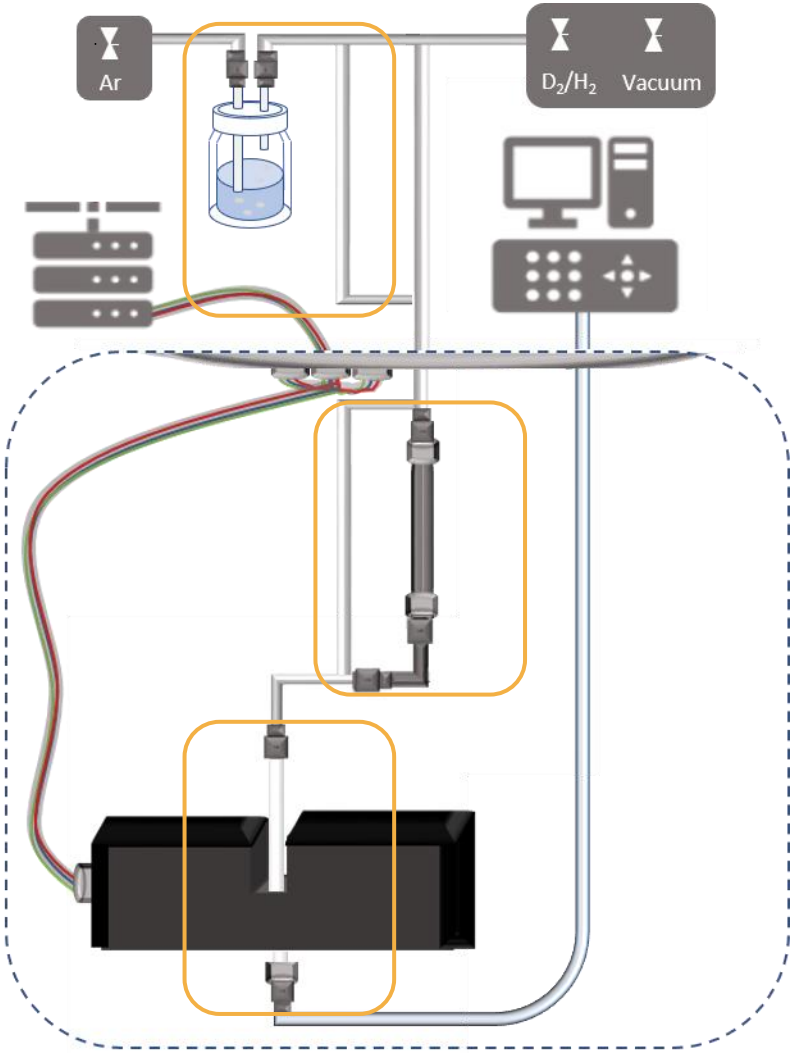
- Complete hydrogenation of benzene to cyclohexane using a 3wt% Pt catalyst supported on MCM-41

- Large contrast between reactant and product – adding six hydrogen atoms
- Mild conditions and no intermediates, good reaction to study prior to more complex more industrially relevant reactions.
- MCM-41 displays periodically arranged pores with minimal pore size disruption so could be modelled with less complexity than some less structurally ordered catalysts or catalyst supports.

Neutrons + NMR in Flow

- Use vanadium can (quartz would be good, but overlaps structurally with our substrate)
- Use Ar gas flow to (continually) deliver reactant / product mixture to neutron and NMR cells
- Flow H_2 or D_2 gas simultaneously
- Allows longer neutron data acquisition times
- NMR tells us the speciation

What is the liquid structure at various composition points within the catalyst support?

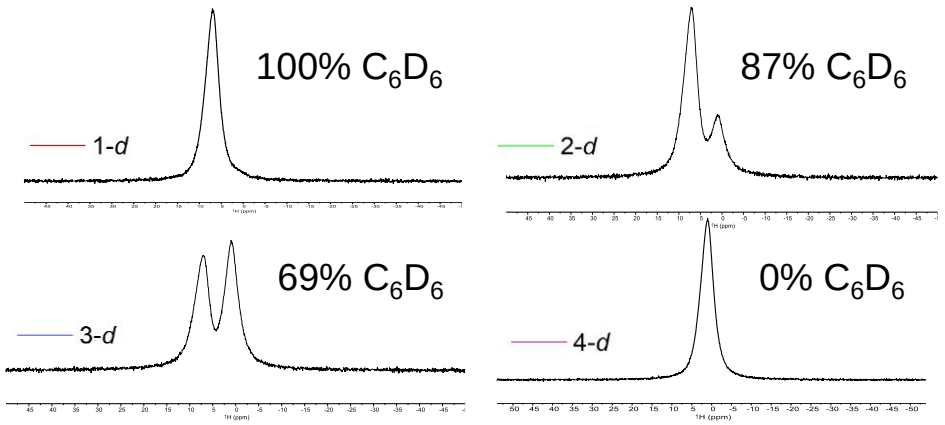


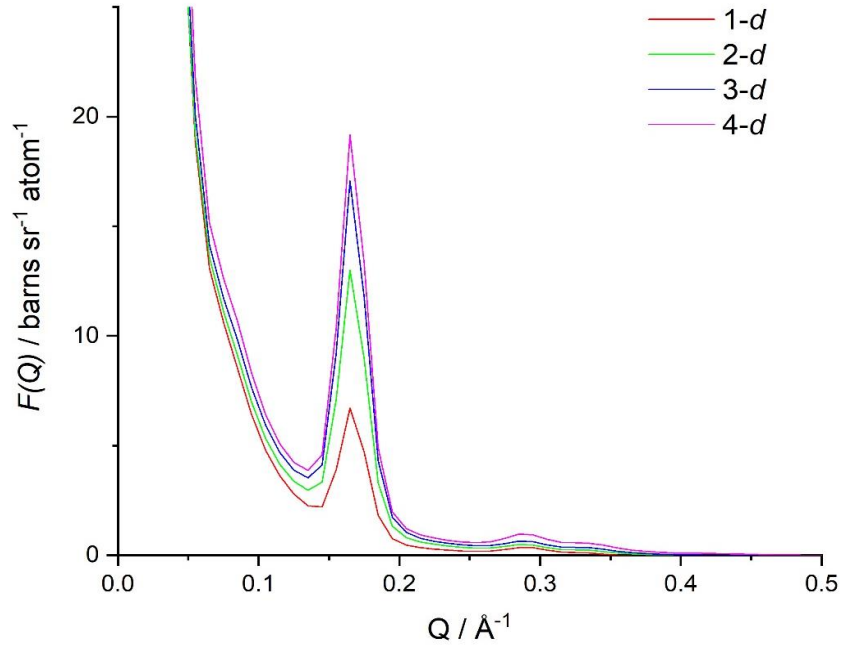
Science and Technology Facilities Council

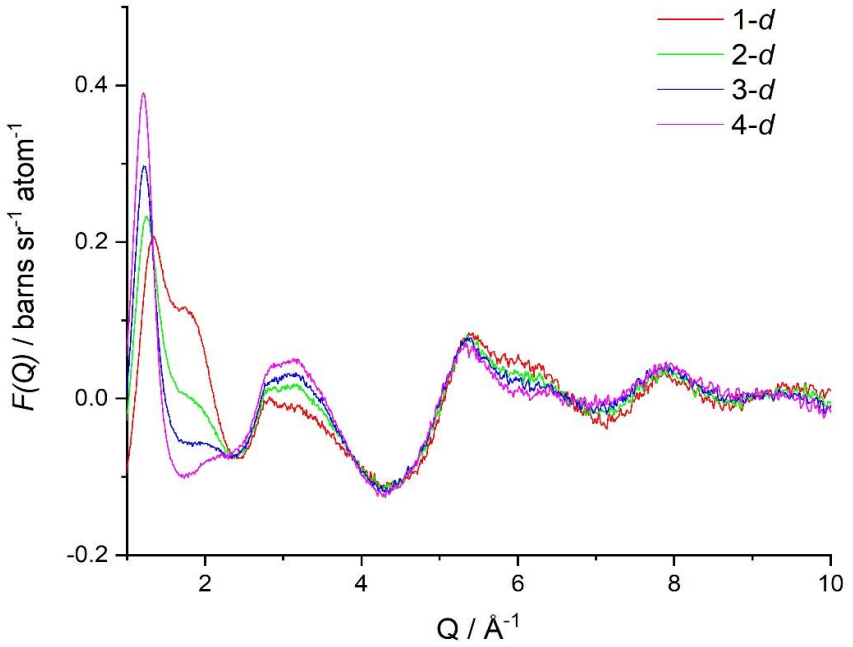
ISIS Neutron and Muon Source

Experimental Data

- Data collected at four compositions
- <100> Bragg peak
  - Position reflects MCM-41 pore array spacing
  - Intensity dependent on mixture composition
  - Increases as C_6D_6 reacts to C_6D_{12}
- Higher *Q* regime indicates changes in liquid arrangement and intramolecular structure







Science and Technology Facilities Council

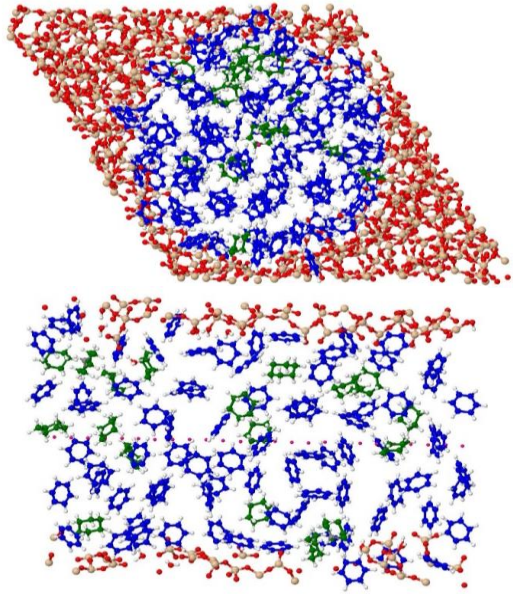
ISIS Neutron and Muon Source

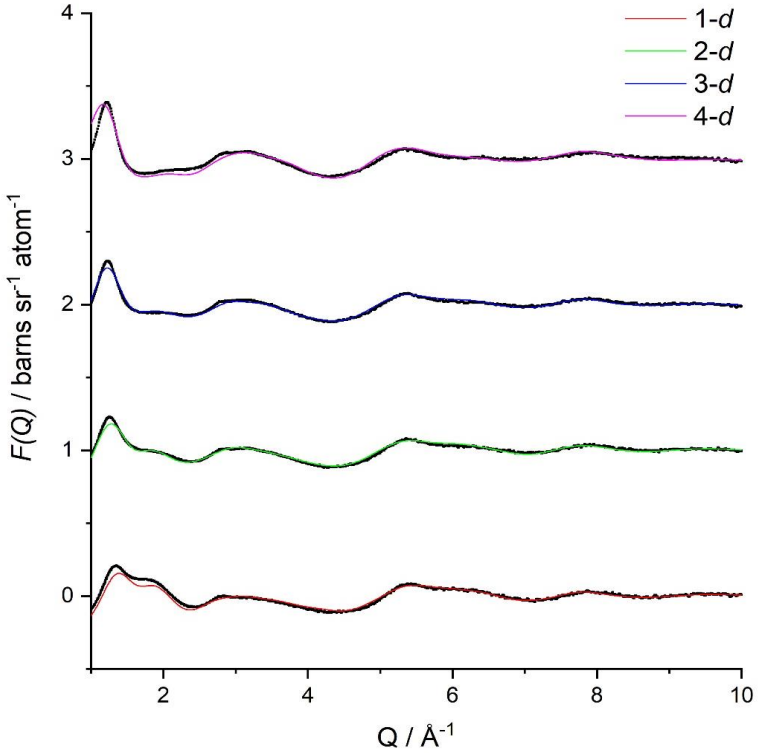
EPSR Simulation

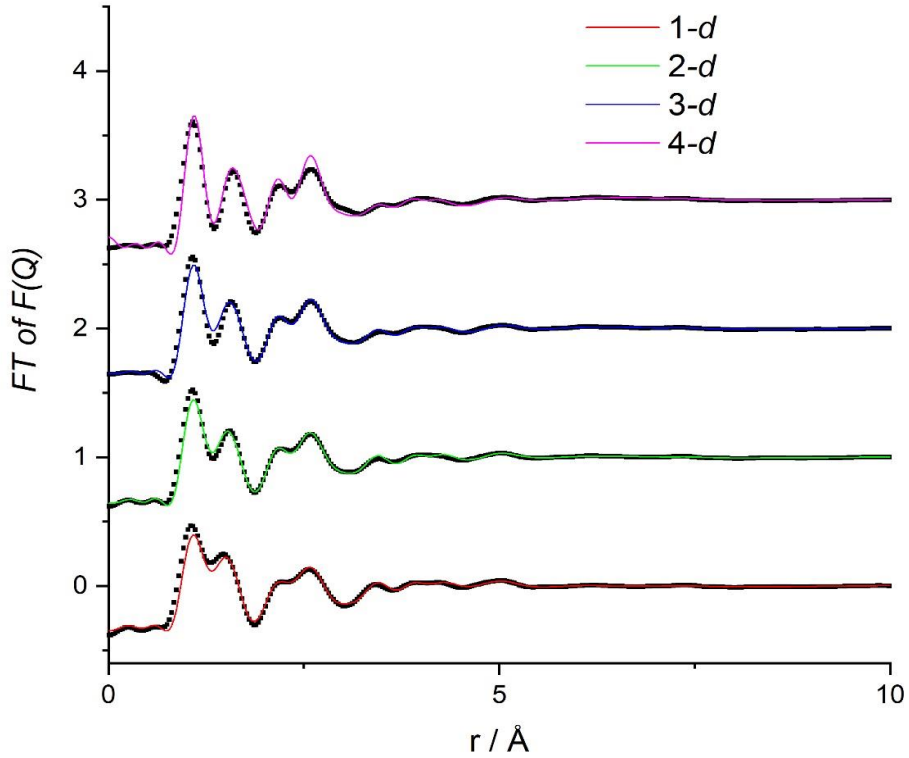
Two-stage process

1. Construct empty MCM-41 using <100> peak to inform unit cell size and pore radius

2. Add in NMR-determined ratio of benzene / cyclohexane at a sensible density and refine again



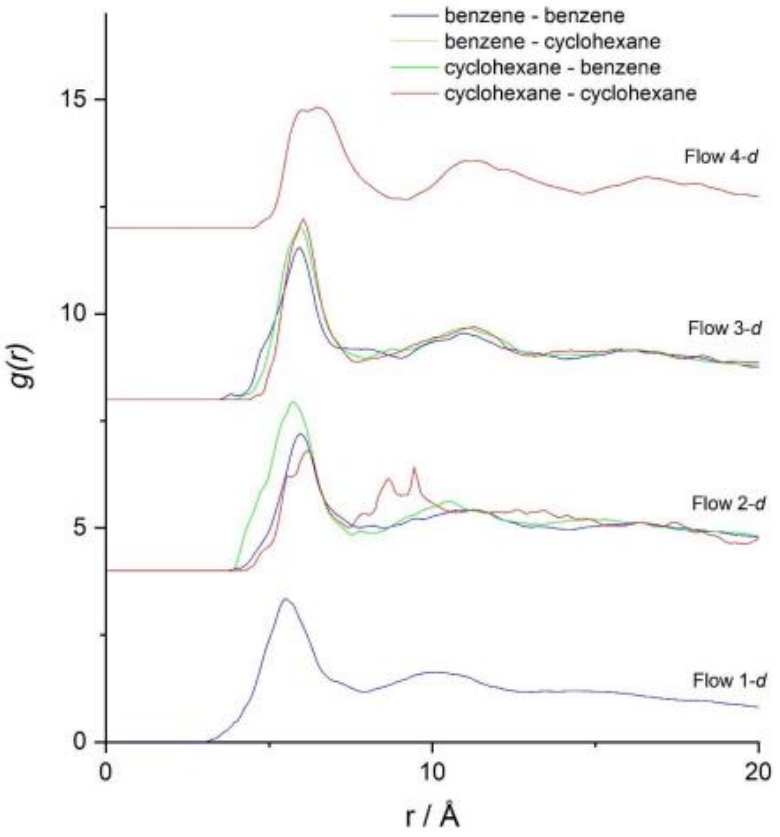


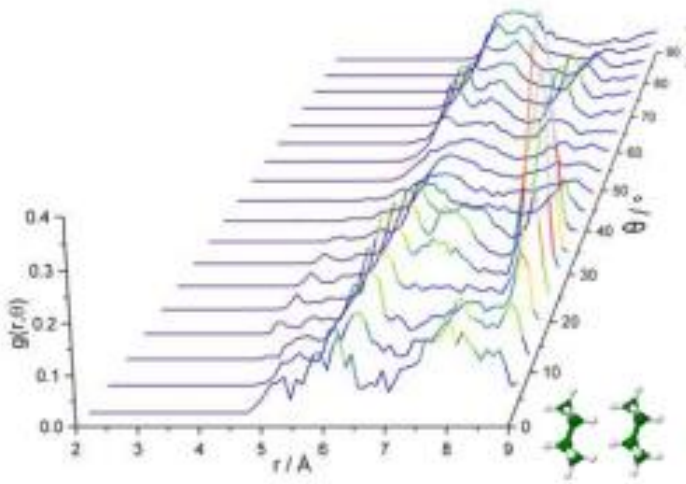


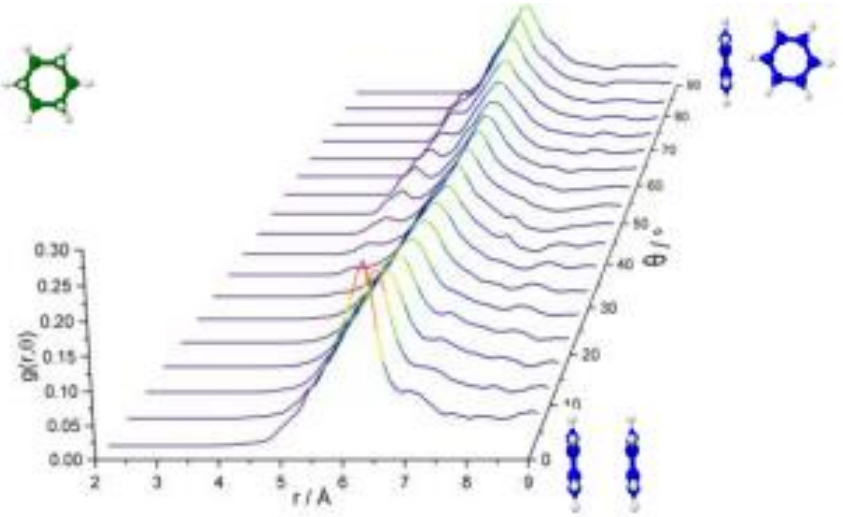
Science and Technology Facilities Council

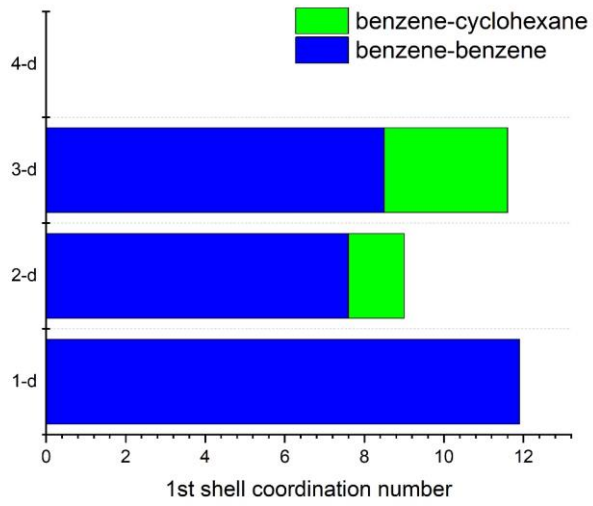
ISIS Neutron and Muon Source

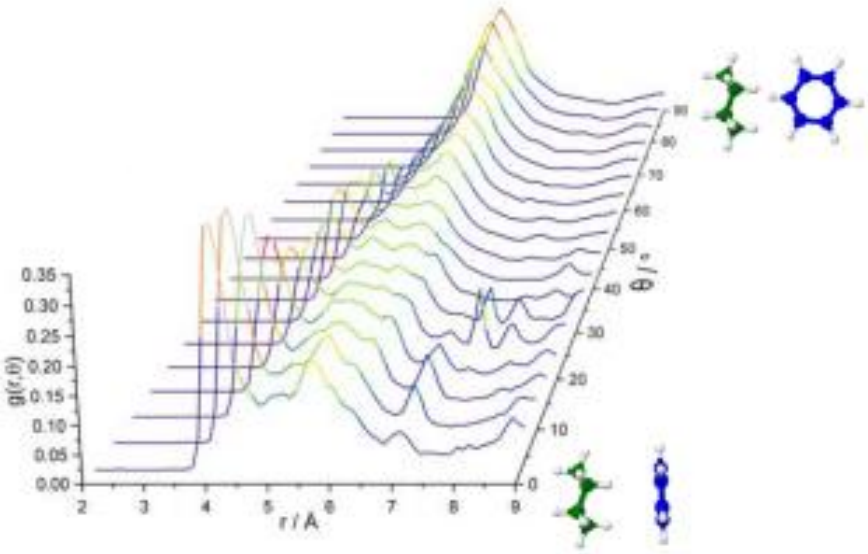
Basic Liquid Structure









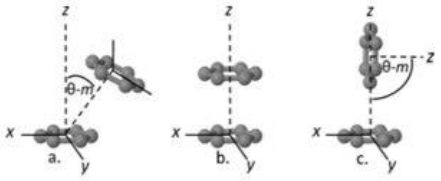


Science and Technology Facilities Council

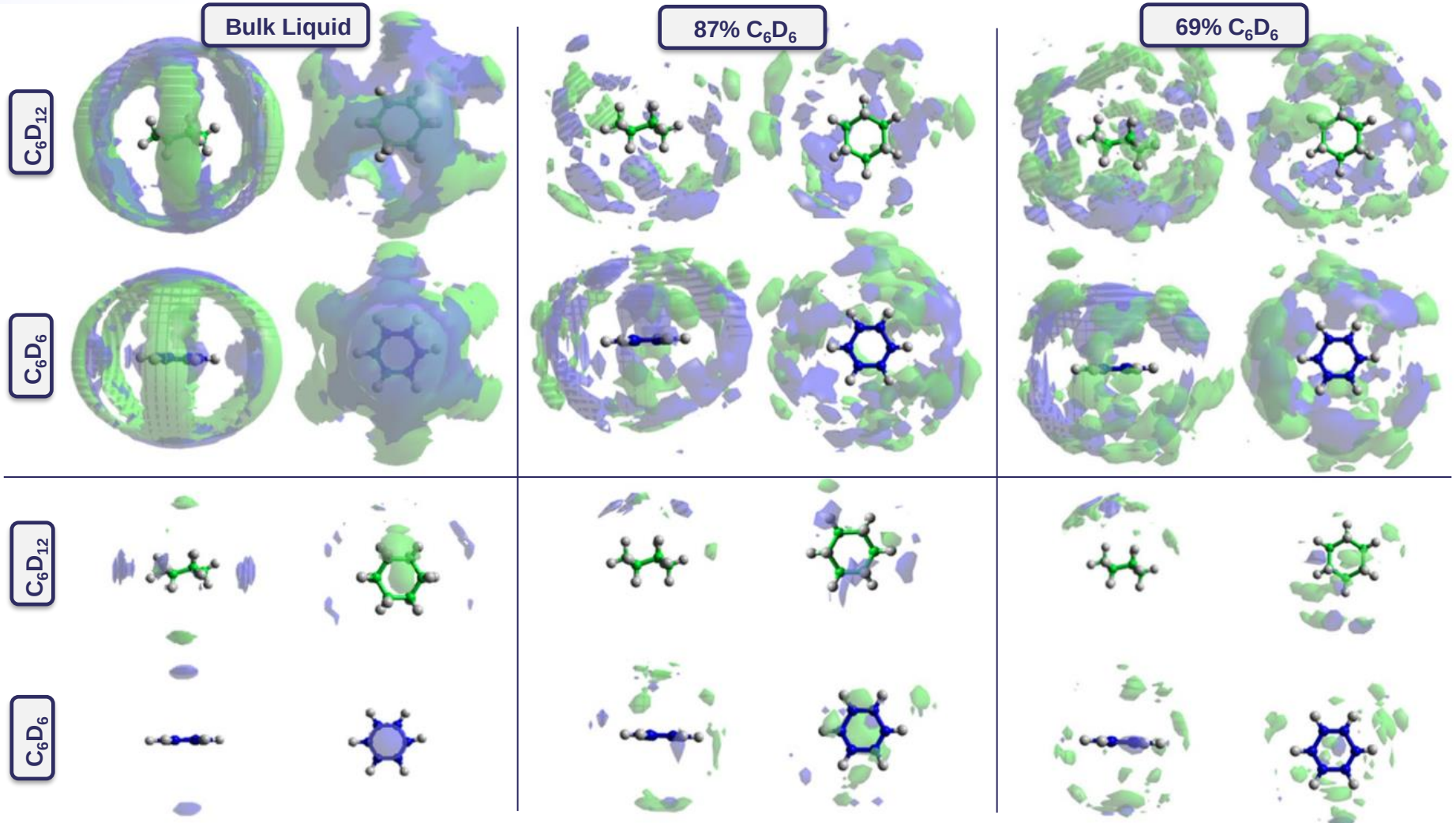
ISIS Neutron and Muon Source

3D Structure

Parallel / r > 4.95 Å



Parallel / r < 4.95 Å



Science and Technology Facilities Council

ISIS Neutron and Muon Source

Current / Emerging Challenges

- Accounting for heterogeneity of samples
  - Partial pore filling
  - Pore size distribution
- System size – coarse graining & refinement
- Soft link between data reduction and simulation
- Extending (improving?) the original EPSR methodology

Dissolve provides a modular framework on which all of this can be tested / built.

Acknowledgements

Disordered Materials @ ISIS

D. T. Bowron, T. F. Headen, O. Alderman, T.-L. Hughes, and A. K. Soper

Team Dissolve

A. Washington, D. Nixon, and all of our contributors

EPSR – Soper, *Mol. Phys.* **99**, 1503 (2001)
https://www.isis.stfc.ac.uk/Pages/Empirical-Potential-Structure-Refinement.aspx

Dissolve – Youngs, *Mol. Phys.* **117**, 3464 (2019)
https://www.projectdissolve.com

Science and Technology Facilities Council
ISIS Neutron and Muon Source

Topics in Catalysis (2021) 64:722–734
https://doi.org/10.1007/s11244-021-01437-w

ORIGINAL PAPER

Bulk and Confined Benzene-Cyclohexane Mixtures Studied by an Integrated Total Neutron Scattering and NMR Method

Terri-Louise Hughes[1] · Marta Falkowska[1] · Markus Leutzsch[2,4] · Andrew J. Sederman[2] · Mick D. Mantle[2] · Thomas F. Headen[3] · Tristan G. A. Youngs[3] · Daniel T. Bowron[3] · Christopher Hardacre[1]

Accepted: 10 April 2021 / Published online: 23 April 2021