Supplementary Information (SI) for Nanoscale.

Supplementary Information for "The Nature of Trehalose–Protein Interactions in Aqueous Solution Revealed by Neutron Scattering"

Kajsa Ahlgren, [a,*] Christoffer Olsson, [b] Tristan Youngs, [c] Jan Swenson[a]

[a] Department of Physics, Chalmers University of Technology, 412 96 Gothenburg, Sweden.
[b] Department of Biomedical Engineering and Health Systems, KTH Royal Institute of Technology, Stockholm SE-114 28, Sweden.
[c] ISIS Pulsed Neutron and Muon Source, STFC Rutherford Appleton Laboratory, Didcot, UK.
[*] Corresponding author: kajsa.ahlgren@chalmers.se

Contents

1 Structural modelling: Species composition

Table S1 Species composition of the molecules included in the model composition.

Element in myoglobin	Amount in 1 molecule	Amount in the M10T20W70 configuration	Amount in the M10T20W70C configuration
C	801	3204	801
N	213	852	213
O	221	884	221
H	1239	4956	1239
S	2	8	2
Fe	1	4	1
Element in trehalose	Amount in 1 molecule	Amount in the M10T20W70 configuration	Amount in the M10T20W70C configuration
C	12	4944	1236
O	11	4532	1133
H	22	9064	2266
Element in water	Amount in 1 molecule	Amount in the M10T20W70 configuration	Amount in the M10T20W70C configuration
O	1	27356	6839
H	2	54712	13678

2 Structural modelling: Lennard-Jones and Coulomb parameters

Table S2 shows the Lennard-Jones and Coulomb parameters used in the Dissolve reference potential. The Lennard-Jones parameters σ and ε correspond to self-interaction parameters for each atom type and define interactions between identical atom types ($i-i$). Lennard-Jones parameters between different atom types ($i-j$) were generated using Lorentz-Berthelot combination rules:

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \qquad \text{and} \qquad \varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}. \tag{1}$$

Table S2 Lennard-Jones and Coulomb parameters for atoms in the Dissolve model.

Element in myoglobin	σ (Å)	ε (kJ/mol)	q (e)
H	0	0	0.0897
O	2.96	0.87864	-0.5
C	3.5	0.27614	0.2
N	3.25	0.71128	-0.76
S	3.55	1.046	-0.47
Fe	2.594	0.0539	2
Element in trehalose	σ (Å)	ε (kJ/mol)	q (e)
Ot	3.1	0.71128	-0.5
O1	2.9	0.58576	-0.5
O2	2.9	0.58576	-0.5
O3	3.1	0.71128	-0.5
Ht	1.7	0.05	0.3005
M	1.7	0.12552	0.0
Ct	3.5	0.27614	0.258
Element in water	σ (Å)	ε (kJ/mol)	q (e)
Hw	0	0	0.4238
Ow	3.166	0.65	-0.8476

3 QENS: Relative scattering contribution

Table S3 Theoretical relative scattering amplitudes for myoglobin, trehalose, and water in the different samples. Values without parenthesis are for the M10T20W70 system and the values within parenthesis are for the M25T25W50 system, they include both coherent and incoherent scattering contributions.

Sample	Relative Scattering Contribution (A^i_C)		
	A^i_{Prot}	A^i_{Tre}	A^i_{Wat}
Mb in D-Tre and H_2O	0.0827 (0.23)	0.0664 (0.09)	0.8509 (0.68)
Mb in H-Tre and D_2O	0.2569 (0.46)	0.3745 (0.34)	0.3686 (0.19)
Mb in H-Tre and H_2O	0.0767 (0.20)	0.1347 (0.18)	0.7886 (0.61)

4 Structural modelling: Magnified images for each data set

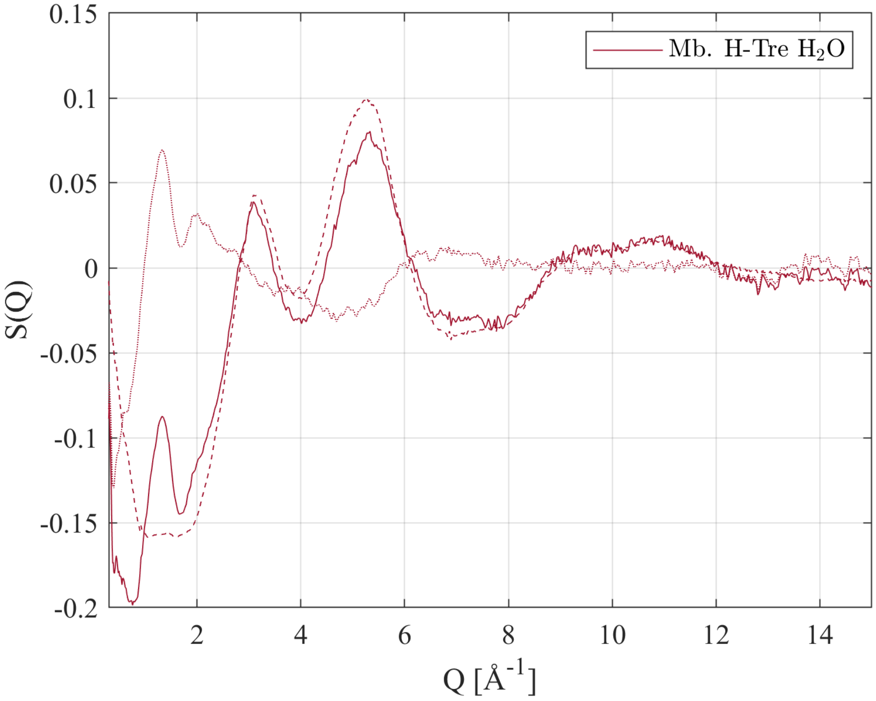

Figure S1 Structure factor for the Mb. H-Tre H_2O system, where the solid line is the experimental diffraction data, the dashed line the Dissolve fit, and the dotted line the Δ between the two.

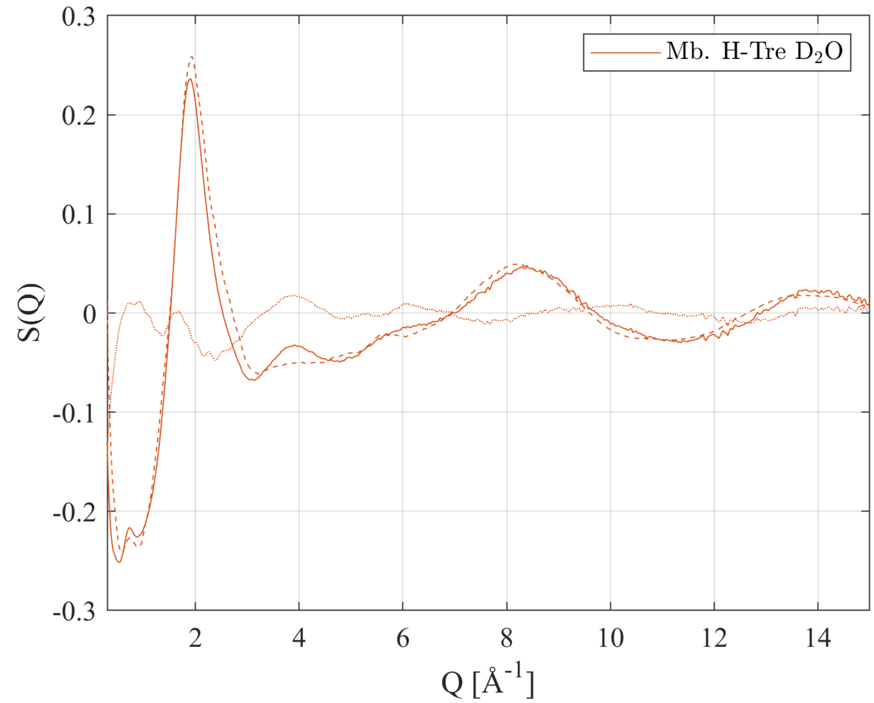

Figure S2 Structure factor for the Mb. H-Tre D_2O system, where the solid line is the experimental diffraction data, the dashed line the Dissolve fit, and the dotted line the Δ between the two.

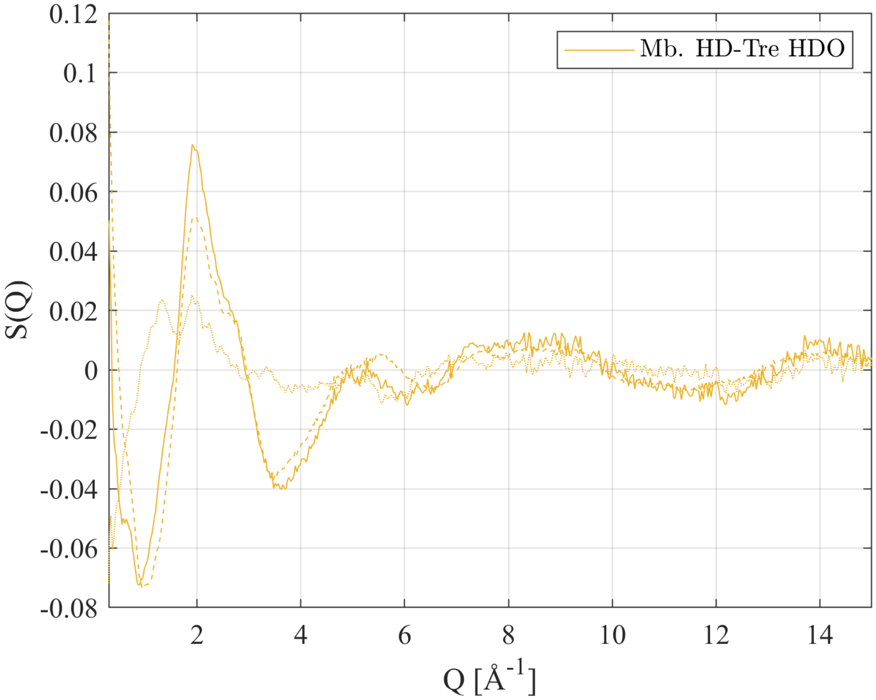

Figure S3 Structure factor for the Mb. HD-Tre HDO system, where the solid line is the experimental diffraction data, the dashed line the Dissolve fit, and the dotted line the Δ between the two.

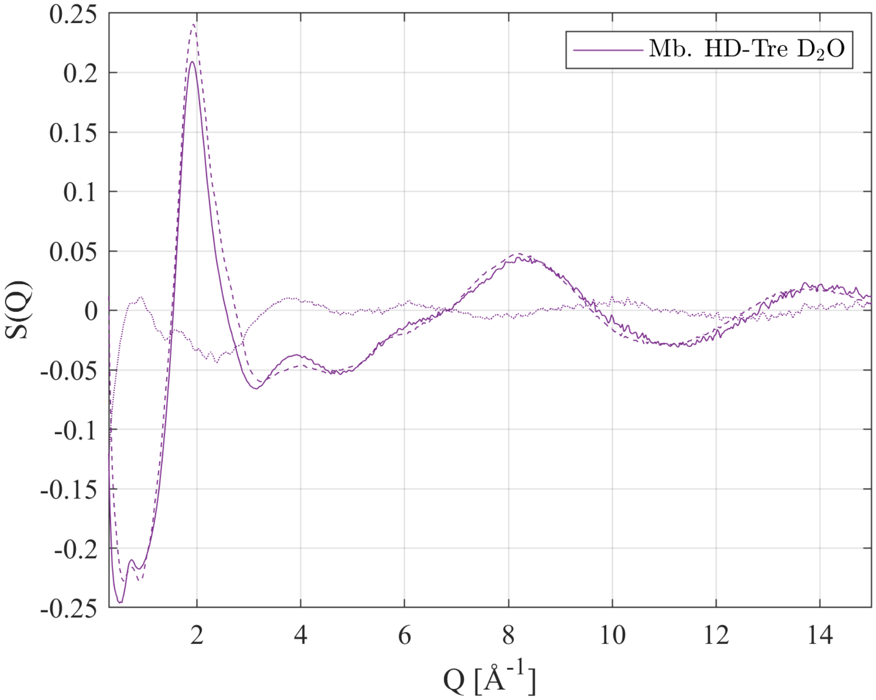

Figure S4 Structure factor for the Mb. HD-Tre D_2O system, where the solid line is the experimental diffraction data, the dashed line the Dissolve fit, and the dotted line the Δ between the two.

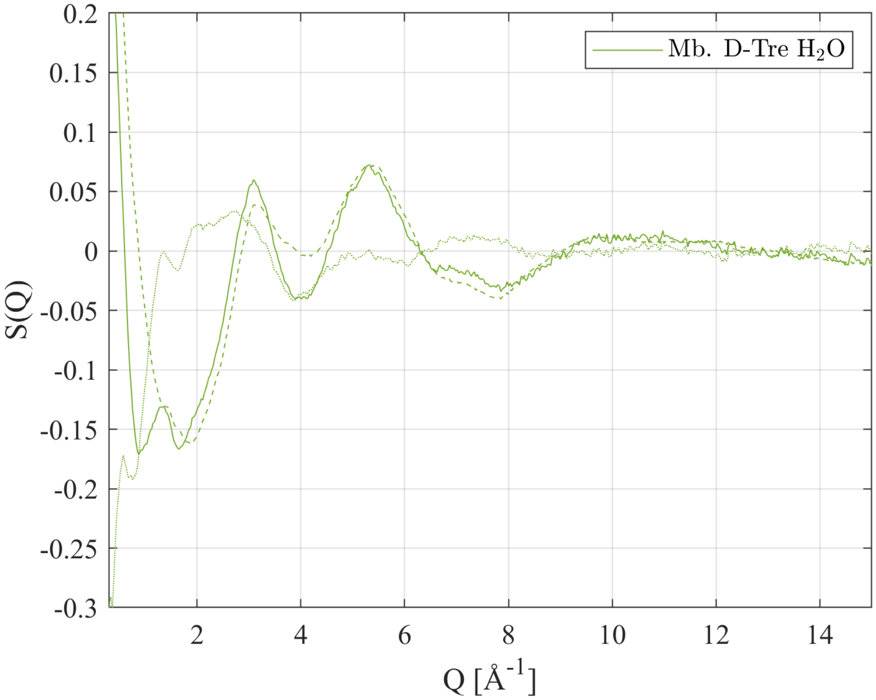

Figure S5 Structure factor for the Mb. D-Tre H_2O system, where the solid line is the experimental diffraction data, the dashed line the Dissolve fit, and the dotted line the Δ between the two.

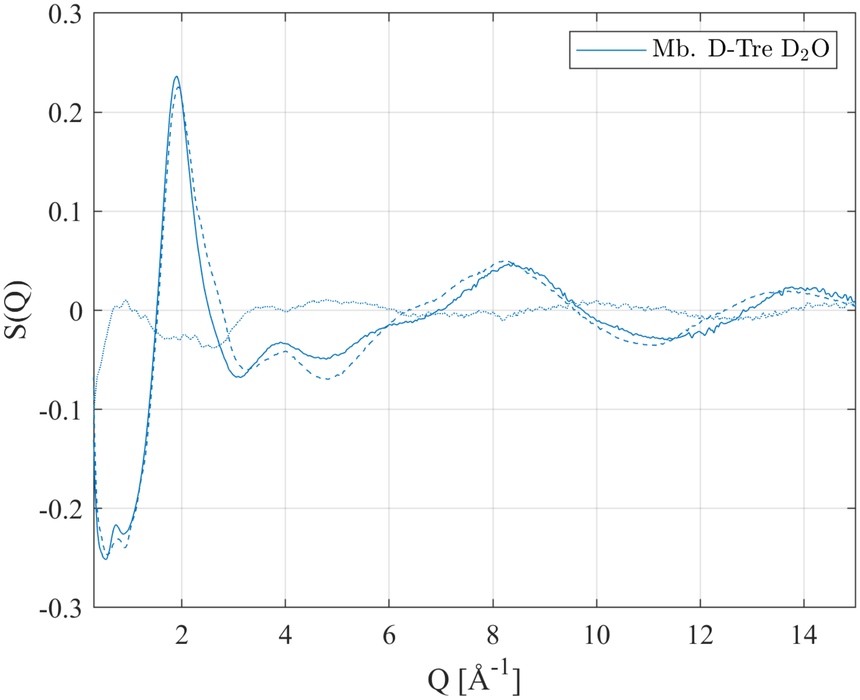

Figure S6 Structure factor for the Mb. D-Tre D_2O system, where the solid line is the experimental diffraction data, the dashed line the Dissolve fit, and the dotted line the Δ between the two.

5 Structural modelling: Comparison between M10T20W70 and M10T20W70C

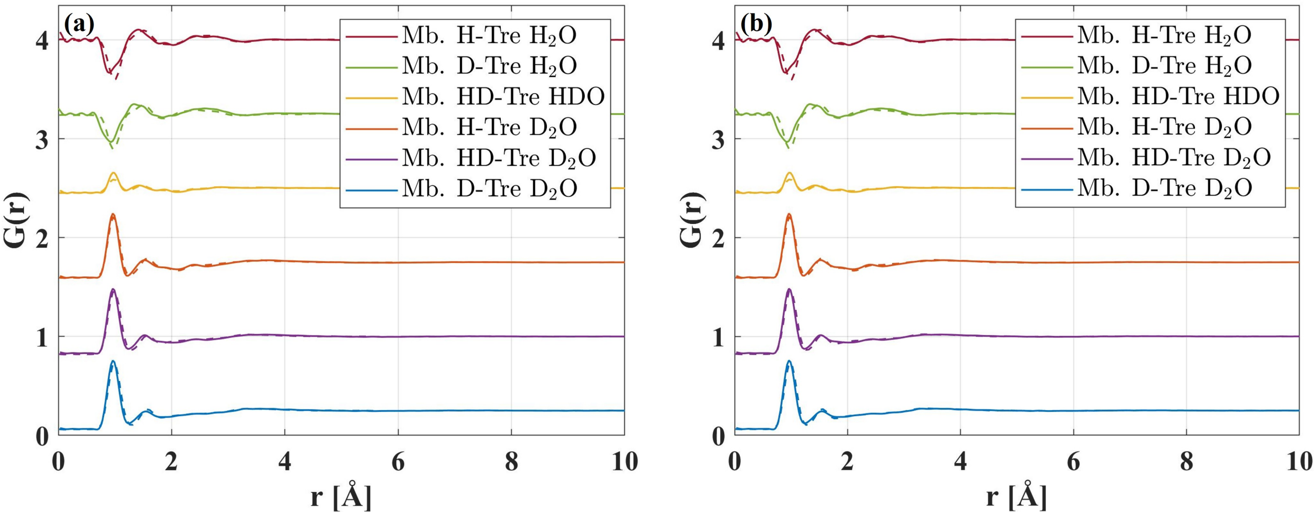

Figure S7 Real-space pair correlation functions, G(r), of the experimental and modelled neutron diffraction data of the (a) M10T20W70 and (b) M10T20W70C system.

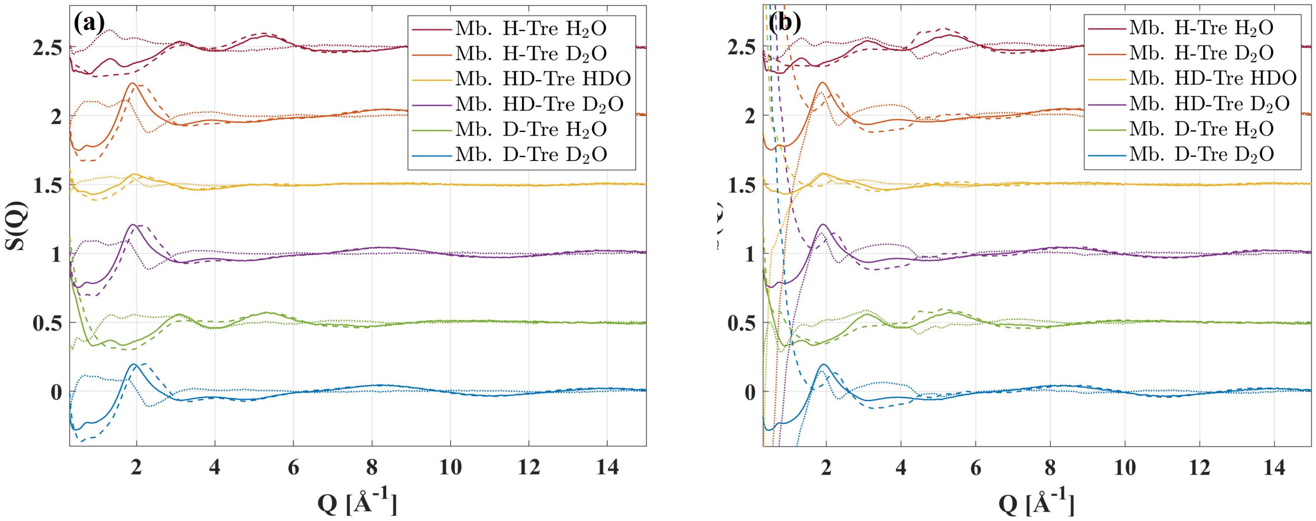

Figure S8 Structure factor of the six isotopically different samples for the (a) M10T20W70 and (b) M10T20W70C system before potential refinement but after energy minimization.

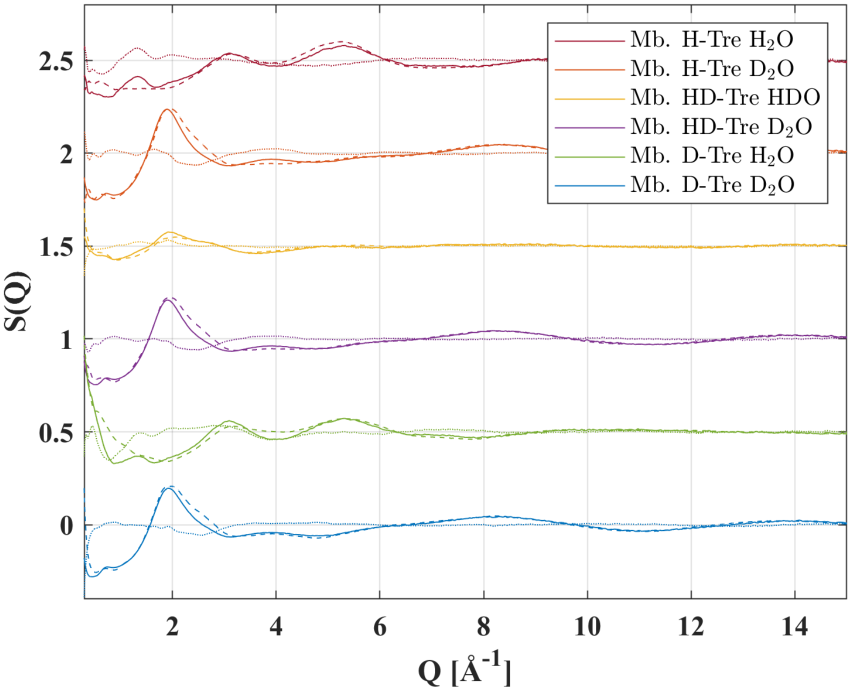

Figure S9 Structure factor of the six isotopically different samples for the M10T20W70C system. The solid lines represent the experimental neutron diffraction data, the dashed lines the Dissolve fits, and the dotted lines the Δ between the experimental and simulated curves.

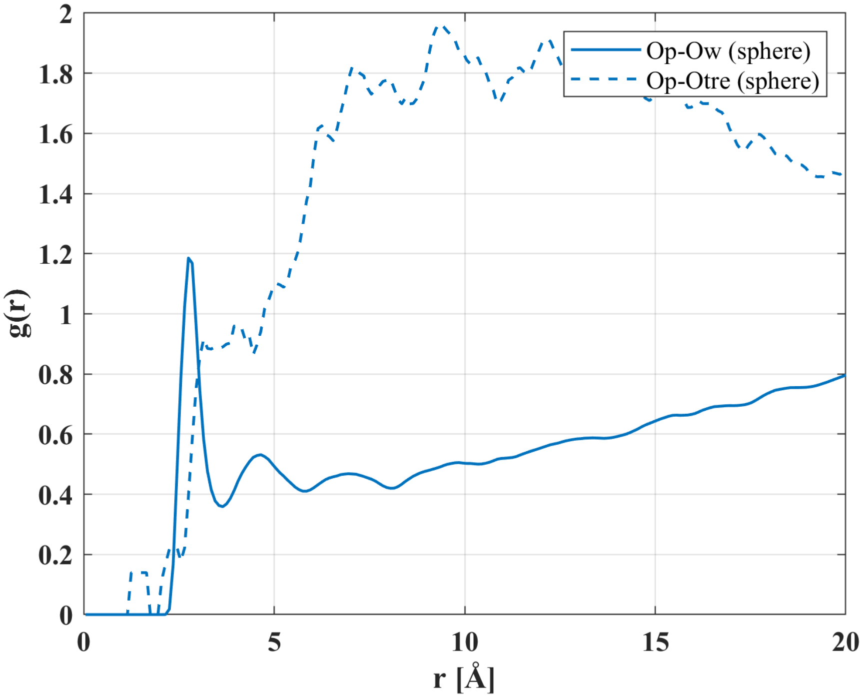

Figure S10 Partial pair correlation functions between the oxygen atoms on the surface of the protein (O_{MB}) and the oxygen atoms in water (O_W) or trehalose (O_{Tre}) for the M10T20W70C model. The blue solid line represents the correlation between O_{MB} and O_W and the blue dashed line the correlation between O_{MB} and O_{Tre}. The oxygens on the surface of the protein are defined as all O_{MB} within 2.5 Å of any atom of the solvent.

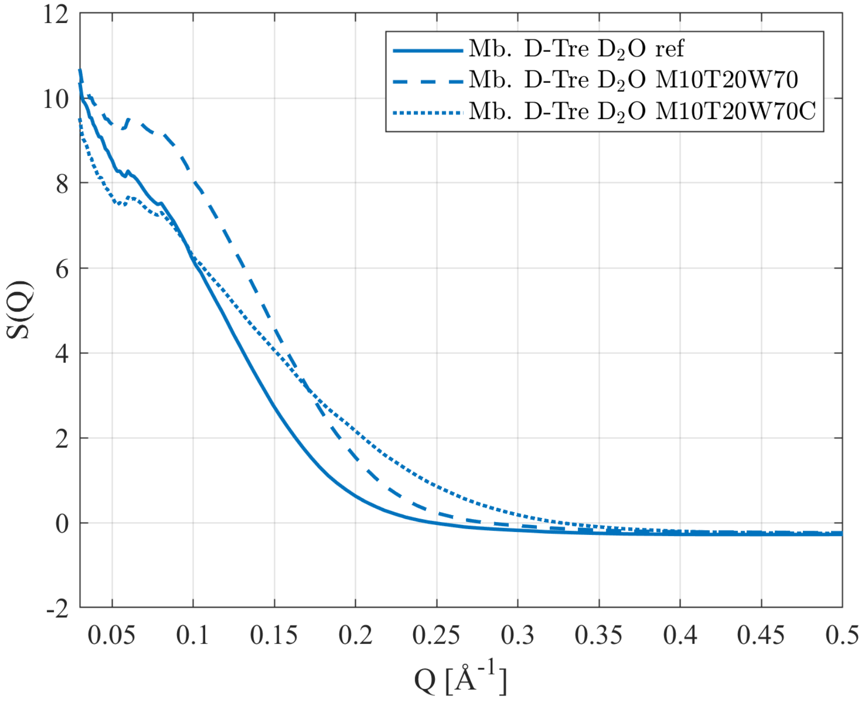

Figure S11 The structure factor in the low Q-range for the Mb D-Tre D2O system. The solid line is the experimental data (ref), the dashed line the M10T20W70 system and the dotted line the M10T20W70C system.

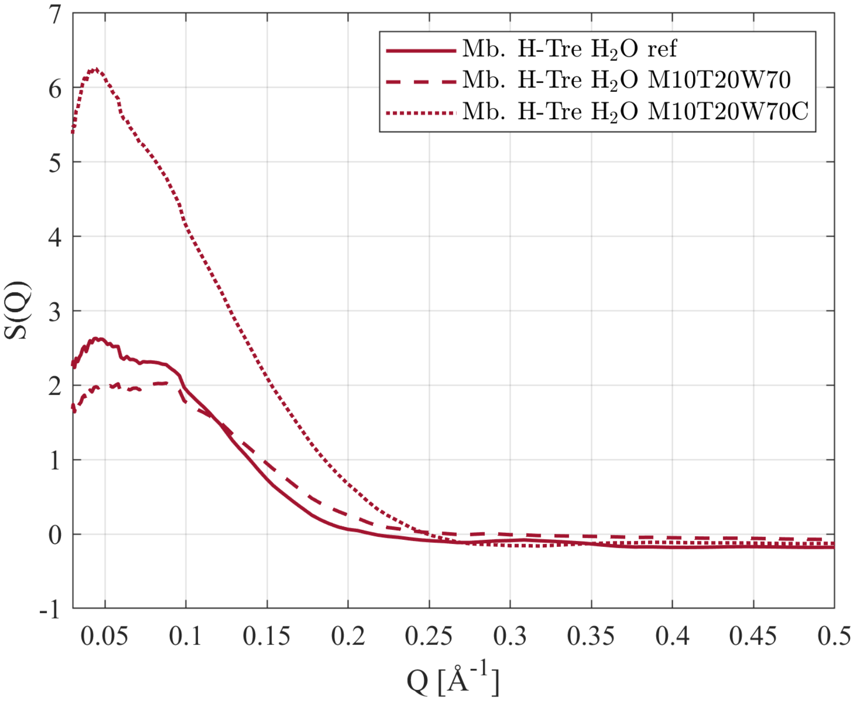

Figure S12 The structure factor in the low Q-range for the Mb H-Tre H2O system. The solid line is the experimental data (ref), the dashed line the M10T20W70 system and the dotted line the M10T20W70C system.

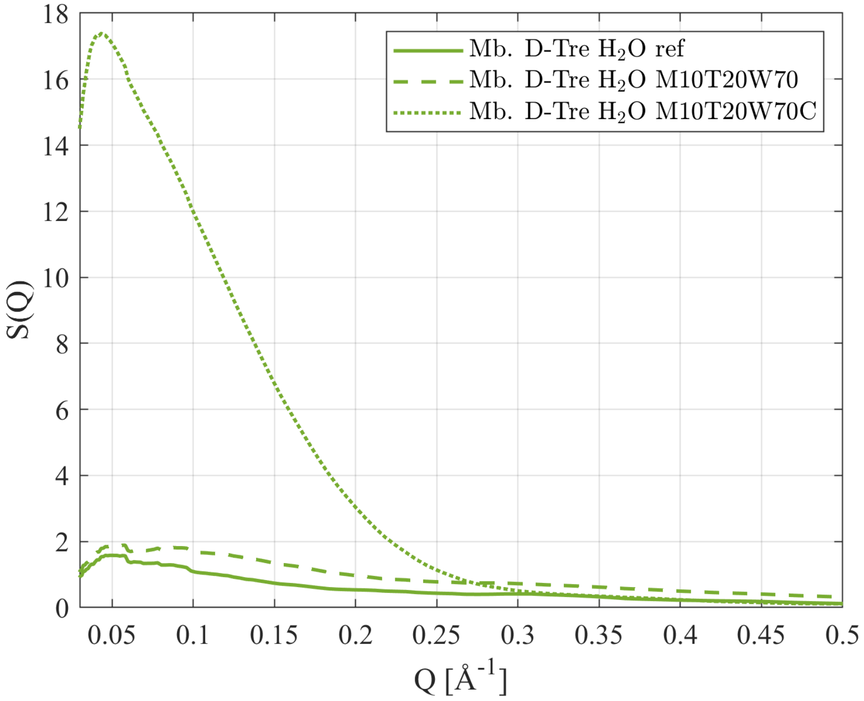

Figure S13 The structure factor in the low Q-range for the Mb D-Tre H2O system. The solid line is the experimental data (ref), the dashed line the M10T20W70 system and the dotted line the M10T20W70C system.

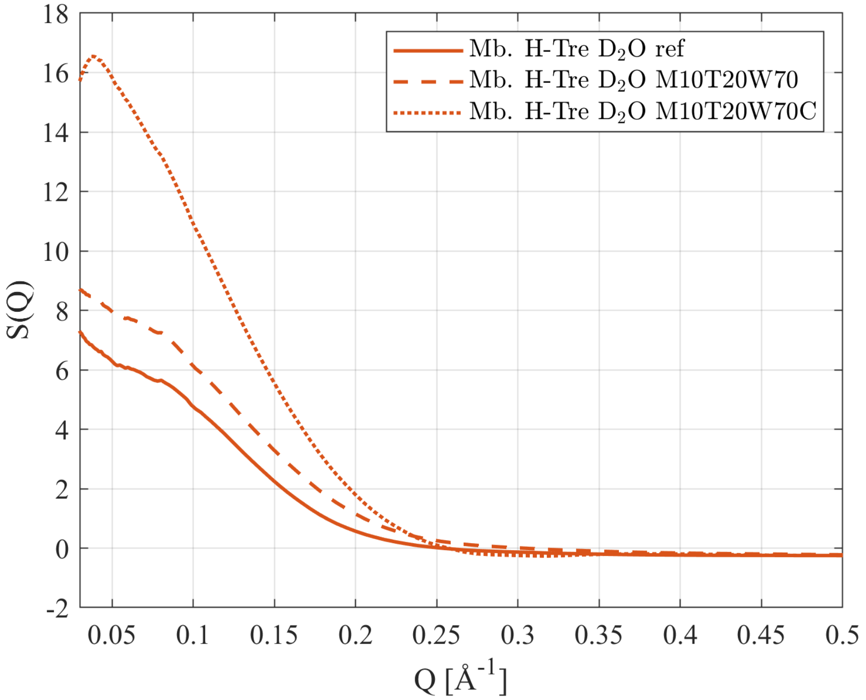

Figure S14 The structure factor in the low Q-range for the Mb H-Tre D2O system. The solid line is the experimental data (ref), the dashed line the M10T20W70 system and the dotted line the M10T20W70C system.

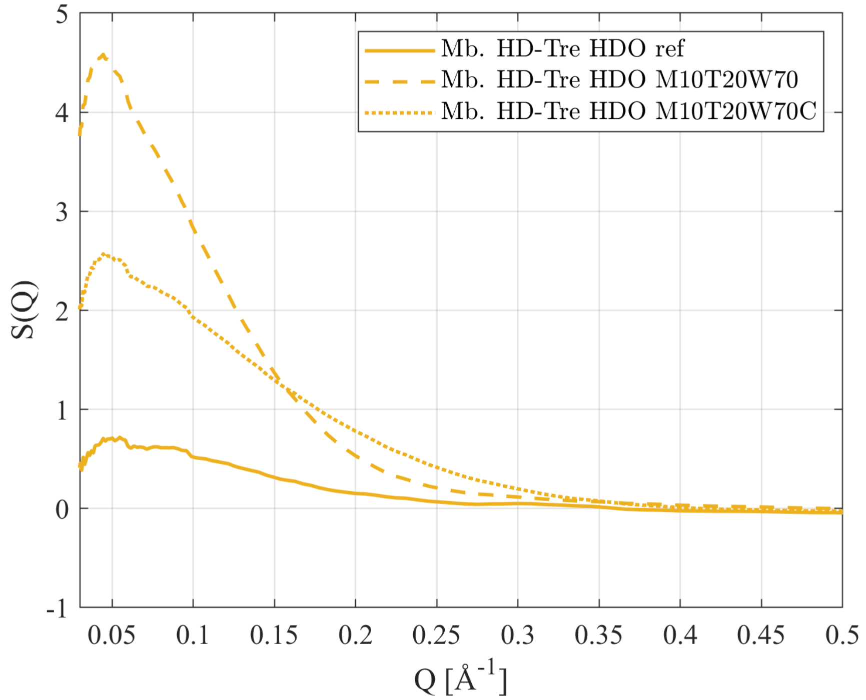

Figure S15 The structure factor in the low Q-range for the Mb HD-Tre HDO system. The solid line is the experimental data (ref), the dashed line the M10T20W70 system and the dotted line the M10T20W70C system.

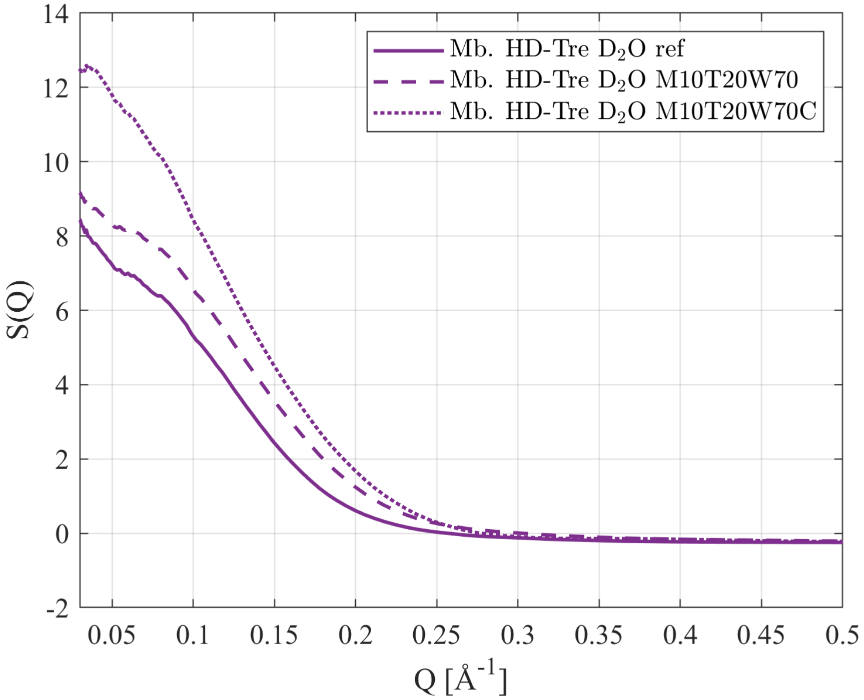

Figure S16 The structure factor in the low Q-range for the Mb HD-Tre D2O system. The solid line is the experimental data (ref), the dashed line the M10T20W70 system and the dotted line the M10T20W70C system.

6 Structural modelling: Comparison between amino acids

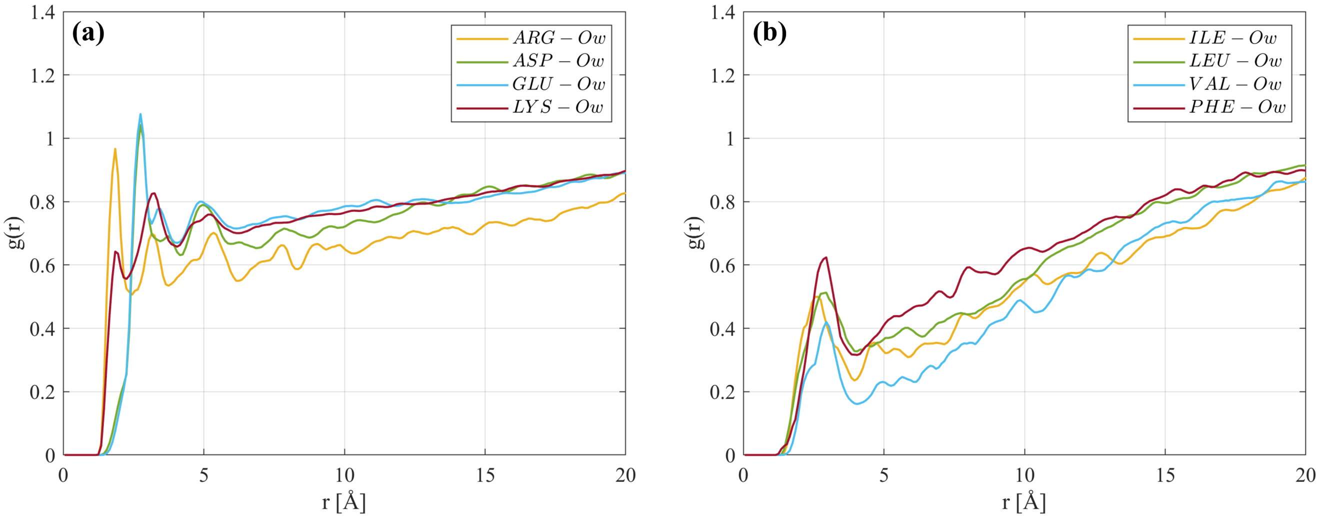

Figure S17 Partial pair correlation functions between (a) any atom in the hydrophilic amino acids on the surface of the protein and Ow and (b) any atom in the hydrophobic amino acids on the surface of the protein and Ow.

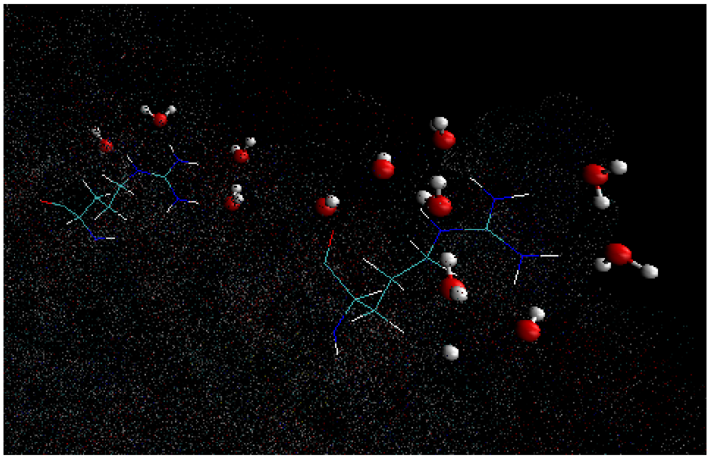

Figure S18 Three-dimensional image from the M10T20W70 model showing Arginine and water molecules within 3.4 Å.

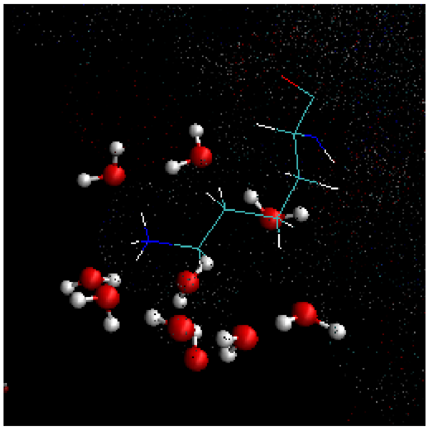

Figure S19 Three-dimensional image from the M10T20W70 model showing Lysine and water molecules within 3.4 Å.

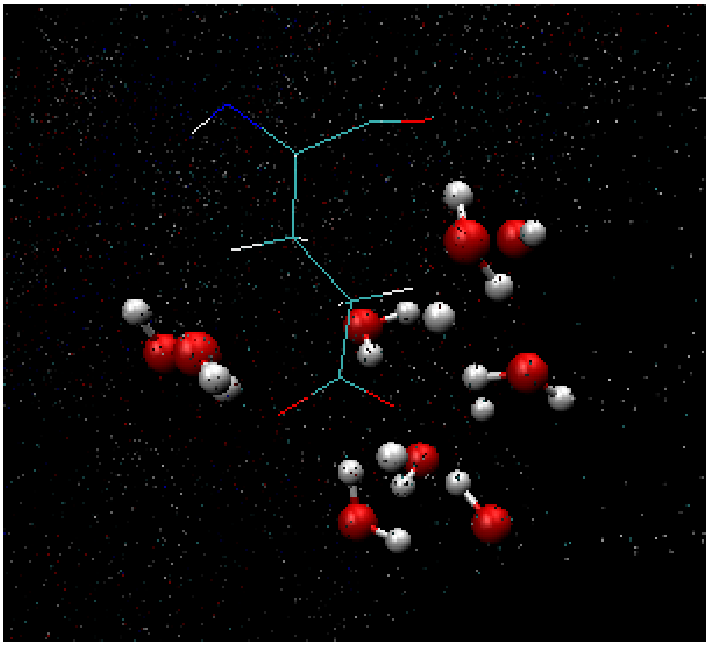

Figure S20 Three-dimensional image from the M10T20W70 model showing Glutamic acid and water molecules within 3.4 Å.

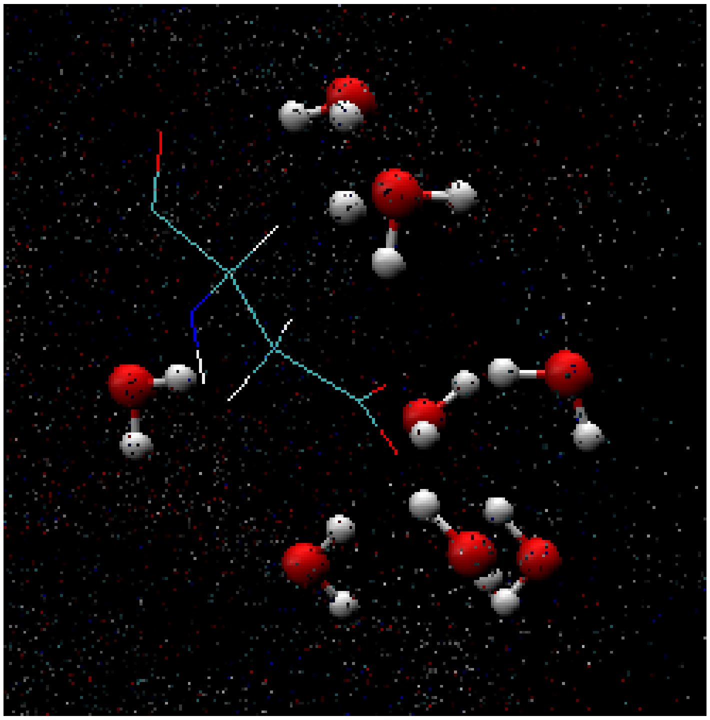

Figure S21 Three-dimensional image from the M10T20W70 model showing Aspartic acid and water molecules within 3.4 Å.

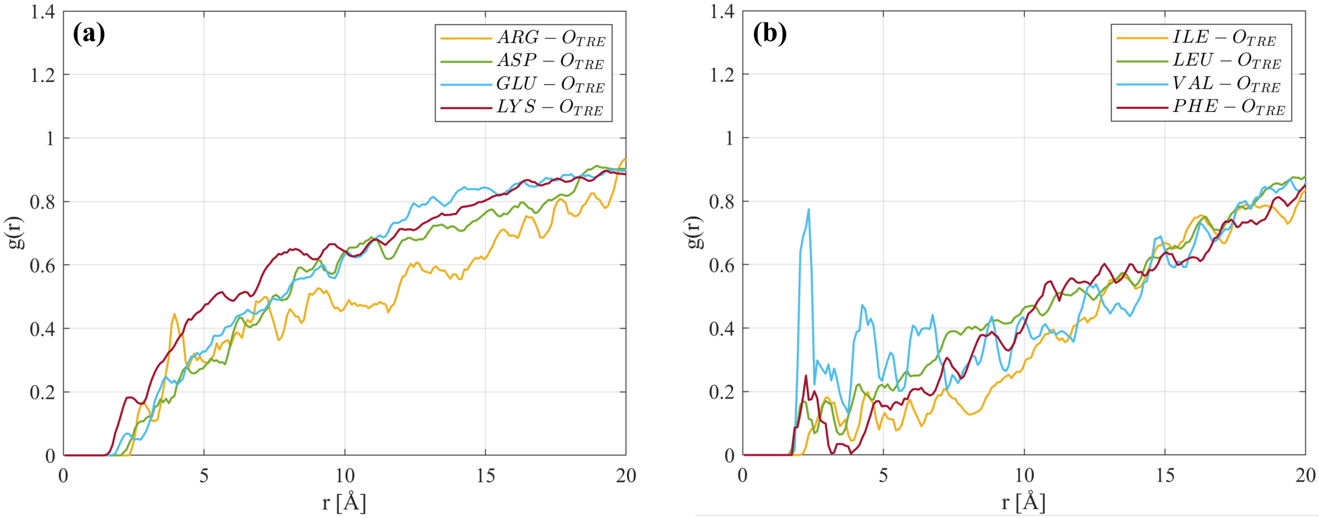

Figure S22 Partial pair correlation functions between (a) any atom in the hydrophilic amino acids on the surface of the protein and Otre and (b) any atom in the hydrophobic amino acids on the surface of the protein and Otre.