Article

Arrangement of Hydrogen Bonds in Aqueous Solutions of Different Globular Proteins

Amber R. Titus[1], Pedro P. Madeira[2], Luisa A. Ferreira[1], Alexander I. Belgovskiy[1], Elizabeth K. Mann[3], J. Adin Mann Jr.[4], William V. Meyer[5], Anthony E. Smart[6], Vladimir N. Uversky[7,*], and Boris Y. Zaslavsky[1,*]

[1] Cleveland Diagnostics, 3615 Superior Ave., Cleveland, OH 44114, USA; amber.titus@clevelanddx.com (A.R.T.); luisa.ferreira@clevelanddx.com (L.A.F.); alexander.belgovskiy@clevelanddx.com (A.I.B.); boris.zaslavsky@clevelanddx.com (B.Y.Z.)
[2] Centro de Investigacao em Materiais Ceramicos e Compositos, Department of Chemistry, Aveiro, Portugal; p.madeira@ua.pt
[3] Department of Physics, Kent State University, Kent, OH, USA; emann@kent.edu
[4] Department of Chemical and Biomolecular Engineering, Case Western Reserve University, Cleveland, OH, USA; j.mann@case.edu
[5] Scattering Solutions, Inc., Cleveland, OH, USA; william.v.meyer@sbcglobal.net
[6] Scattering Solutions, Inc., Costa Mesa, CA, USA; anthonyericsmart@gmail.com
[7] Department of Molecular Medicine and Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, USA; vuversky@usf.edu
* Correspondence: vuversky@usf.edu; Phone: +1-813-974-5816; 12901 Bruce B. Downs Blvd., MDC07; Tampa, FL 33612, USA (V.N.U.); boris.zaslavsky@clevelanddx.com; Phone: +1-216-432-2700, Cleveland Diagnostics, 3615 Superior Ave., Cleveland, OH 44114, USA (B.Y.Z.)

Abstract: This work presents the first evidence that dissolved globular proteins change the arrangement of hydrogen bonds in water, with different proteins showing quantitatively different effects. Using ATR-FTIR (Attenuated Total Reflection – Fourier Transform Infrared) spectroscopic analysis of OH-stretch bands, we obtain quantitative estimates of the relative amounts of the previously reported four subpopulations of water structures coexisting in a variety of aqueous solutions. Where solvatochromic dyes can measure the properties of solutions of non-ionic polymers, the results correlate well with ATR-FTIR measurements. In protein solutions to which solvatochromic dye probes cannot be applied, NMR (Nuclear Magnetic Resonance) spectroscopy was used for the first time to estimate the hydrogen bond donor acidity of water. We found strong correlations between the solvent acidity and arrangement of hydrogen bonds in aqueous solutions for several globular proteins. Even quite similar proteins are found to change water properties in dramatically different ways.

Keywords: Fourier Transform Infrared spectroscopy; water structure; hydrogen bonds; protein solution; solvent properties

1. Introduction

All biological processes, specifically protein folding, biomolecular recognition, hydrophobic effects, protein-ligand binding, and liquid-liquid phase separation *in vivo* only occur with the active participation of water [1-3]. The role of water in many of these biological processes has been mostly ignored [4]. Recently, the active role of water in biological processes has become a subject of focused studies [5-8]. For example, human heat shock protein HSP6 [9] and dehydrins (dehydration proteins) of various molecular weights [10] alter the solvent features of water. These solvent features in solutions were characterized with solvatochromic dyes [9, 10] and included the solvent dipolarity/polarizability (π*) representing dipole-dipole and dipole-induced dipole interactions, solvent hydrogen bond donor acidity (α), and solvent hydrogen bond acceptor basicity (β). We established [11] that one or two of these properties strongly correlate with various physicochemical properties of aqueous solutions of various solutes, such as activity coefficient,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2022 doi:10.20944/preprints202208.0327.v1

osmotic pressure, viscosity, surface tension, and relative permittivity. The effects of nonionic polymers on the solvent features of water are important in macromolecular crowding [12, 13]. The excluded volume introduced by macromolecular crowding was earlier thought to dominate the functional and structural properties of proteins and nucleic acids both in vivo and in model conditions *in vitro*. Phenomena such as folding mechanisms, conformational stability, aggregation propensity, and interactions with other molecules earlier ascribed to space restrictions are now shown to be dependent upon the structure of the water in which they are dissolved. Nonionic polymers are typically used as model crowders, but analysis of their influence on the solvent features of aqueous media [12, 13] indicates that they act by changing the solvent properties of water.

In systems showing phase separation, the solvent properties of water differ between phases [14]. These differences govern partitioning between two phases of various solutes ranging from small organic compounds to proteins [15] and are strongly correlated with the interfacial tension between the phases [16].

Solvatochromic dyes permit the estimation of changes in the properties of water in various solutions [11-15] but may be used only for proteins that do not bind to aromatic compounds. This method may not therefore be used for many proteins such as serum albumin or lysozyme. We recently used ATR-FTIR spectroscopy to show that the solvent features of water are strongly correlated with hydrogen bond arrangement in solutions of various compounds [17-22]; here we use it to explore the rearrangement of hydrogen bonds in aqueous solutions of several globular proteins.

2. Results and Discussion

Figure 1 shows an example of decomposition of the OH-stretch band into four different Gaussian components in a protein solution. The ratio of the water subpopulations/clusters changes differently in the presence of different proteins.

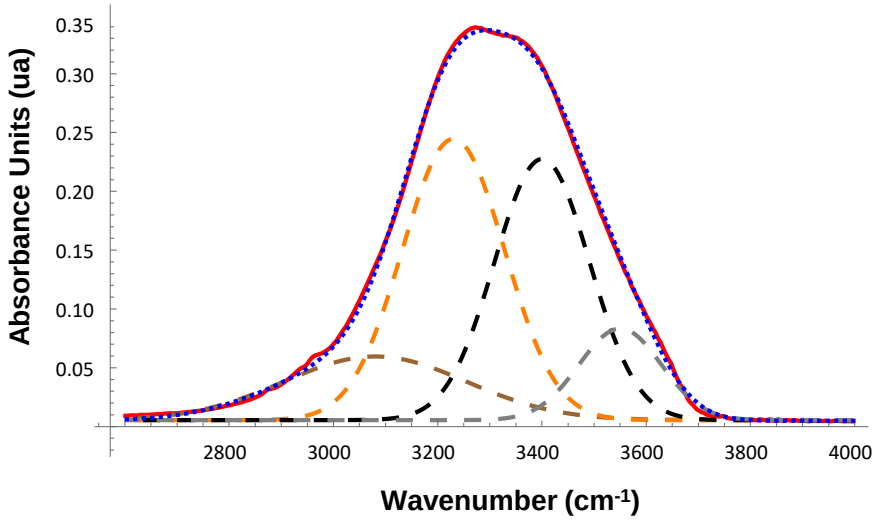

Figure 1. Examples of the ATR-FTIR spectra of OH-stretch band in the solvent for 250 mg/mL β-lactoglobulin A in 0.15 M NaCl in 0.01 M Na-phosphate buffer, pH 7.4 (PBS). The blue dotted line is the measured absorption spectrum, and the red envelope is the best fit of the sum of our four Gaussian components (dashed lines at positions (3080 cm^{-1}, 3230 cm^{-1}, 3400 cm^{-1}, and 3550 cm^{-1}). Experimental data and fit are visually almost indistinguishable.

Estimates of the relative contributions of the Gaussian components I (3080 cm^{-1}), component II (3230 cm^{-1}), component III (3400 cm^{-1}), and component IV (3550 cm^{-1}) for all the proteins examined here at various concentrations are listed in Supplementary Information (**Table S1**). Concentration dependences of the relative contributions of the Gaussian

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2022 doi:10.20944/preprints202208.0327.v1

components I-IV for human and bovine albumins, ovalbumin, and human γ-globulin are plotted in **Figures 2a-d**. Analogous dependences for lysozyme, β-lactoglobulins A and B, and trypsinogen are shown in **Figures 3a-d**. We display the eight proteins, measured separately, in two groups for clarity.

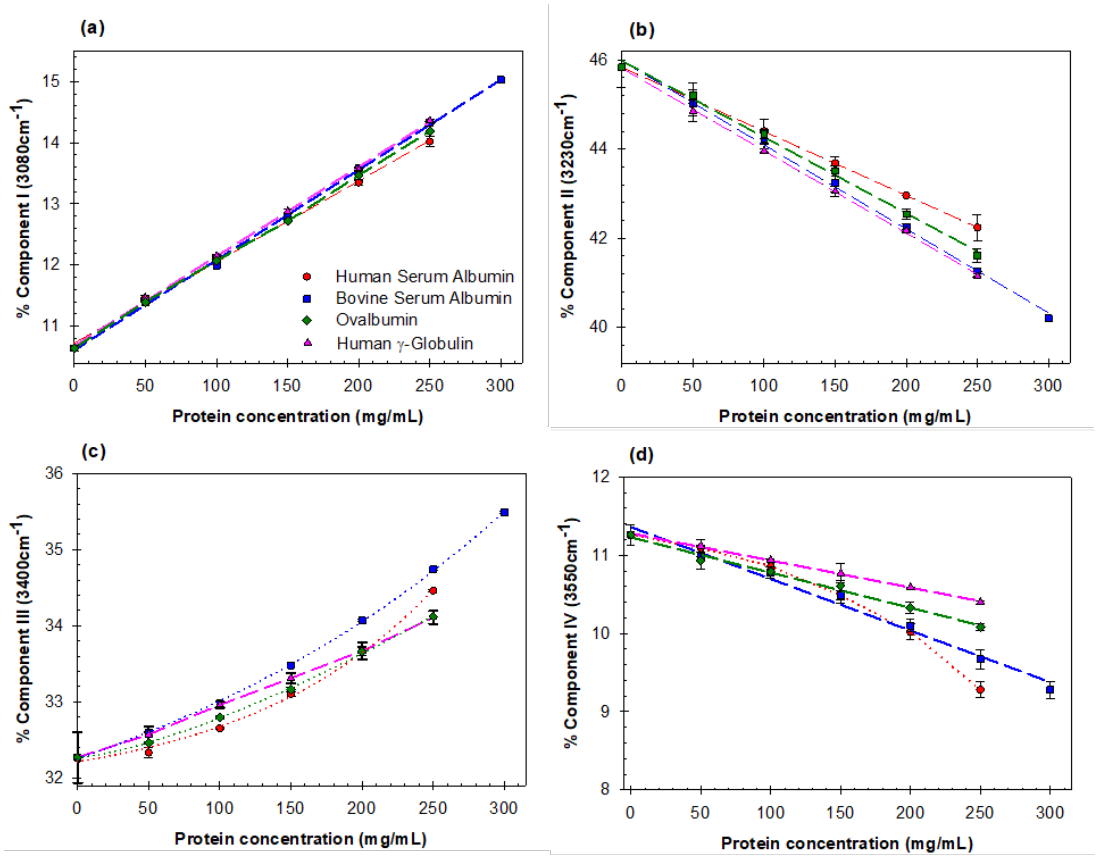

Figure 2. Concentration dependence of the relative contributions of the Gaussian components (a) I, (b) II, (c) III, and (d) IV for human serum albumin (red circles), bovine serum albumin (blue squares), ovalbumin (green diamonds), and human γ-globulin (magenta triangles).

The contribution of Gaussian component I (3080 cm^{-1}), assigned to water molecules with four tetrahedrally-arranged hydrogen bond subpopulations, increases linearly with concentration for all globular proteins explored here, whereas that for Gaussian component II (3230 cm^{-1}), assigned to water molecules with four distorted hydrogen bonds, decreases linearly with the same materials and constraints. There is no systematic apparent relationship between the slope of Gaussian component I and protein molecular weight.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2022 doi:10.20944/preprints202208.0327.v1

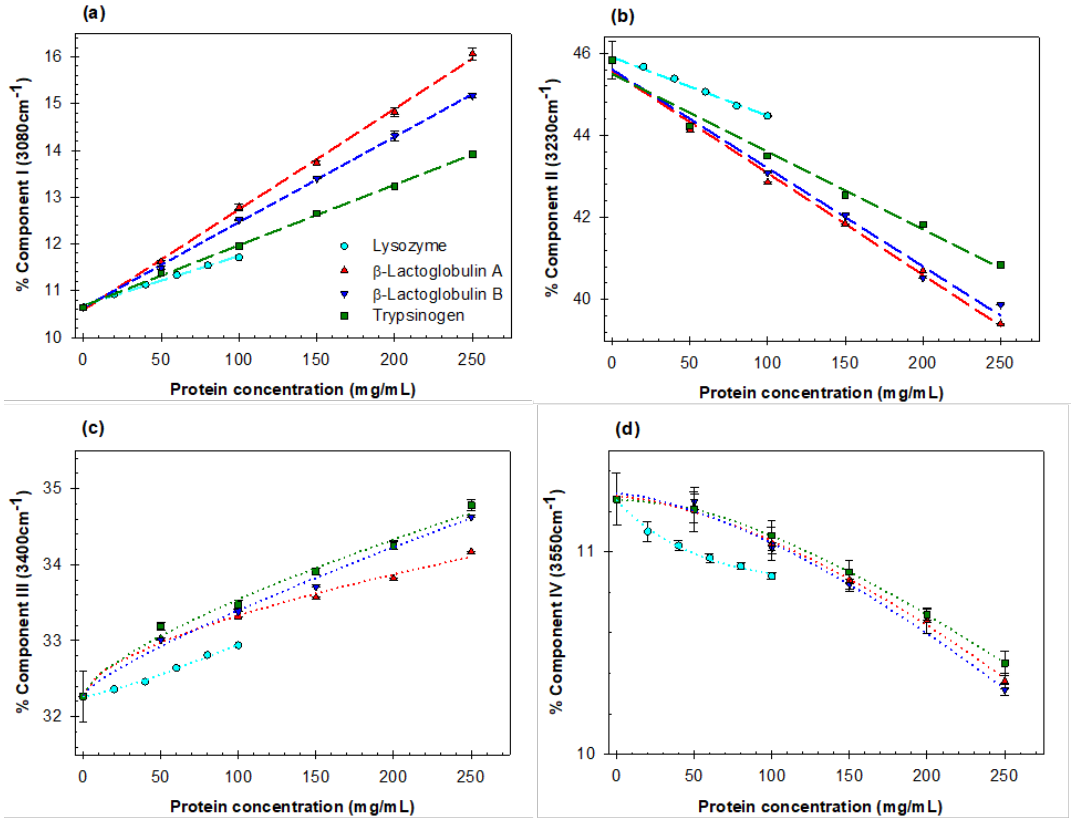

Figure 3. Concentration dependences of the relative contributions of the Gaussian components I (a), II (b), III (c), and IV (d) for chicken egg lysozyme (cyan circles), β-lactoglobulins A (red triangles) and B (blue inverted triangles), and trypsinogen (green squares).

For each examined protein, the fraction of ice-like water structure (component I) and the fraction of distorted ice-like water structure (component II) both change linearly with increasing protein concentration, but in opposite directions (**Figures 2a**, **2b**, **3a**, and **3b**). The relative change of component III increases linearly only for γ-globulin. Component IV decreases linearly for bovine serum albumin, ovalbumin, and γ-globulin, and becomes non-linear for human serum albumin, lysozyme, β-lactoglobulins A and B, and trypsinogen. This suggests that the effects of different proteins on the structure of water depend strongly on the nature and spatial arrangement of the solvent-exposed protein groups.

The rather unexpected finding is how significantly different are the effects of two β-lactoglobulins A and B on hydrogen bond arrangement in aqueous solution (**Figure 3a** and **3c)**. These two proteins differ at only two positions: Gly64 in β-lactoglobulin B substitutes for Asp64 in β-lactoglobulin A, and Ala118 in β-lactoglobulin B substitutes for Vall118 in β-lactoglobulin A. Both proteins form dimers at pH 7.4. All the observed differences in the effects of various proteins on rearrangements of hydrogen bonds in aqueous solutions may be due to differences in the nature and steric arrangements of the solvent-exposed residues in these proteins.

Analysis of the data in **Table S1** shows a strong linear correlation between the α_{24} parameter and the relative contributions of Gaussian components II (3230 cm^{-1}) and III (3400 cm^{-1}) for a given protein. The correlations observed are presented graphically in **Figures 4a-d**. For (a) and (b) we observe dependence only on the value of Gaussian component III, whereas for (c) and (d) the dependence is upon both Gaussian components II and III.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2022 doi:10.20944/preprints202208.0327.v1

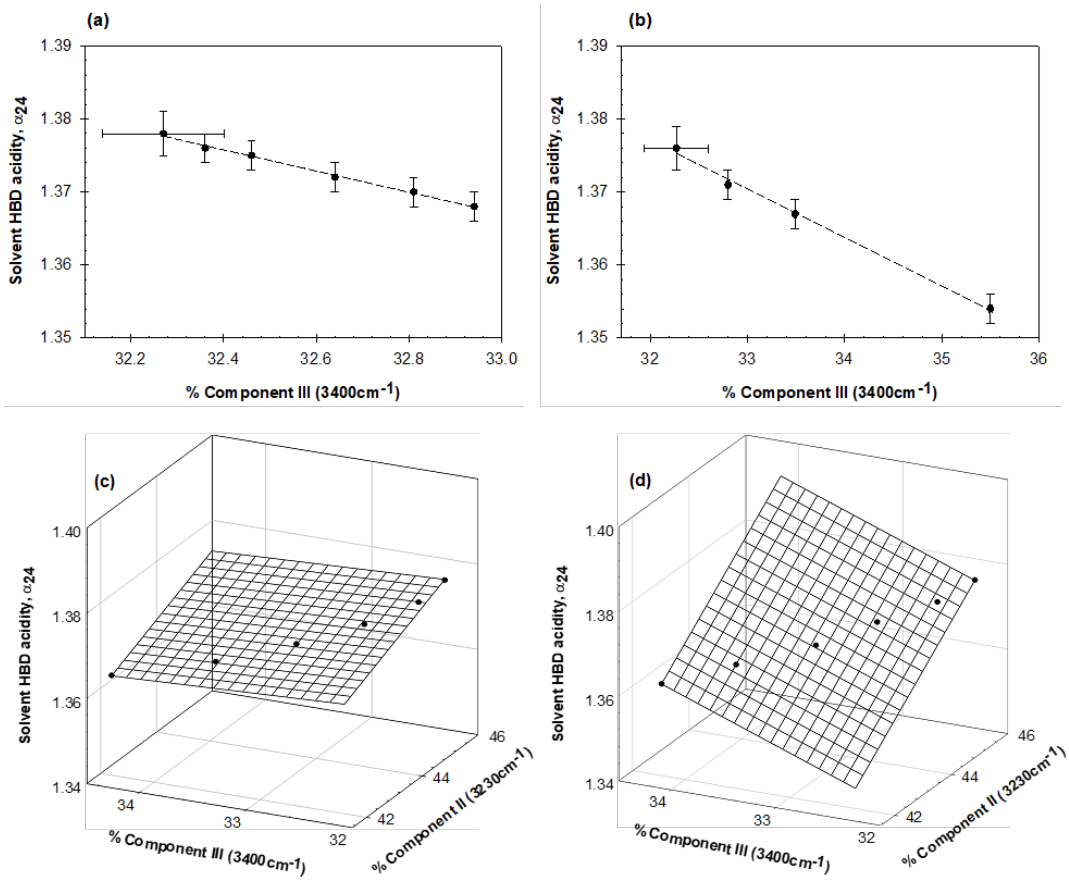

Figure 4. Relationship between the solvent hydrogen bond donor acidity, α_{24}, and the relative contributions of Gaussian component III for (a) lysozyme and (b) bovine serum albumin, and the relative contributions of Gaussian components II and III for (c) human serum albumin and (d) ovalbumin.

As above, solvatochromic dyes typically interact with most proteins and therefore cannot be used to estimate the solvent properties of water in protein solutions. In aqueous solutions of ionic liquids, however, hydrogen bond donor acidity, α_{24}, was evaluated with the pyridine-N-oxide probe using the ^{13}C NMR chemical shift [23]. Here we used this NMR technique for the first time to evaluate α_{24} in solutions of globular proteins with the results presented in **Figure 4** and **Table S2**.

These data show that the relative contributions of Gaussian components II (3230 cm^{-1}) and/or III (3400 cm^{-1}) are strongly correlated with solvent hydrogen bond donor acidity in protein solutions. Similar relationships were previously reported [20] for solutions of nonionic polymers, inorganic salts, and several small organic compounds. Data presented in **Figure 4** show that changes in the arrangement of hydrogen bonds of water in protein solutions cause changes in the solvent properties of water.

4. Materials and Methods

2.1. Materials

The following specimens were purchased from Sigma-Aldrich: (1) human serum albumin (fatty acid and human globulin free (~99%)), (2) albumin from bovine serum, (3) human γ-globulin, (4) β-lactoglobulin A from bovine milk (>90%), (5) β-lactoglobulin B from bovine milk (>90%), (6) lysozyme from chicken egg white, (7) trypsinogen from bovine pancreas, (8) ovalbumin, and (9) pyridine-N-oxide (PyO) with a purity of 95% for the NMR measurements and analysis.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 August 2022 doi:10.20944/preprints202208.0327.v1

2.2. Methods

2.2.1. ATR-FTIR measurements

ATR-FTIR spectra for each sample were measured in two separately prepared solutions using an Alpha II FT-IR spectrometer equipped with Platinum single reflection ATR single reflection diamond ATR module (Bruker Scientific, LLC, Billerica, MA, USA). All measurements were performed at ambient temperature (approximately 23 °C) using 24 scans for each sample and 24 scans for background in the spectral range of 4000-1000 cm^{-1} with resolution of 4 cm^{-1}. The spectra were reproducible to better than 1 cm^{-1}.

2.2.2. Analysis of ATR-FTIR Spectra

ATR-FTIR spectra were analyzed using custom software written in Wolfram Mathematica and run under version 12. The software analyzed the OH-stretch band by fitting the data using the 'NonlinearModelFit' function, with a model using the sum of three, four, or five Gaussian distributions with floated peak frequencies, amplitudes, and widths. Using this model, fitting a baseline is necessary where real values may extend beyond the available measured dataset or the data may have a residual offset. We found that the best and most reliable fits gave peak frequencies within the error band of values from literature [22], which were therefore fixed for all later fitting routines. These values, which we confirmed earlier [17, 20], are 3080 cm^{-1}, 3230 cm^{-1}, 3400 cm^{-1}, and 3550 cm^{-1} , represented by μ_{1-4} respectively. The fit includes the four Gaussian expressions identified by the subscript, with baseline B and normalization by A to the measured total amplitude, where σ_i is the standard deviation of each Gaussian component, and a_i is the amplitude contribution of each:

$$B + A\sum_{i=1}^{4}\frac{a_i}{2\sigma_i\sqrt{\pi}}e^{-\left[\frac{1}{2}\left(\frac{x-\mu_i}{\sigma_i}\right)^2\right]}.$$

The program displays graphically the raw data, model function fit, and individual Gaussian distributions. Also available are all fitted parameter values together with a suite of boundary metrics quantifying confidence for each fit.

In FTIR spectra, vibrational bands such as the C-O-H stretch mode of alcohols and biomolecules together with their N-H stretch (Amide A) modes overlap the broad OH-stretch band of water. We suggest that this OH-stretch band in solutions of the proteins examined here represents water sufficiently closely to support the results presented in this paper. This is justified by two observations. First, the molar concentration of water exceeds those of the protein functional groups by several orders of magnitude, even at the highest concentrations used in this study. Second, in the region of interest (2600 to 3800 cm^{-1}), water absorbs more than ten times more strongly than any protein functional group.

ATR-FTIR spectroscopic analysis of the OH-stretch band showed that the hydrogen bonding of water depends on specific solutes, such as inorganic salts, trimethylamine N-oxide, urea, the polymers PEG, PVP, and a copolymer of ethylene glycol and propylene glycol (Ucon), and their concentrations [20]. For every compound examined, the minimal unstructured residuals were obtained by fitting the measured OH-stretch band with four Gaussian components peaking at 3080, 3230, 3400, and 3550 cm^{-1}. We confirmed the quality of this fit by calculating the correlation matrix, assuring the independence of the variables.

Our results show that this simple model of exactly four Gaussian components suggests the simultaneous coexistence of four subpopulations of water with different H-bond arrangements; water with four tetrahedrally arranged hydrogen bonds (3080 cm^{-1}), water with four distorted hydrogen bonds (3230 cm^{-1}), water with loosely arranged three or four hydrogen bonds (3400 cm^{-1}), and water with three, two, or one hydrogen bond(s) (3550

 doi:10.20944/preprints202208.0327.v1

cm^{-1}) [20]. The proportion of each subpopulation depends on the properties and concentration of the solute. However, the physical distribution, geometry, structure, molecular arrangement, and scale of each subpopulation are currently unknown. Previously reported solvent properties of water [20], such as solvent dipolarity/polarizability, π*, solvent H-bond donor acidity, α, and solvent H-bond acceptor basicity, β, correlate strongly with the fractional contributions of subpopulations of water.

This approach was used successfully to analyze the coexisting phases of several aqueous two-phase systems [17]. We apply the same model here for the analysis of water in solutions of different globular proteins.

2.2.3. NMR measurements.

NMR chemical shifts (d) in ppm were determined using a Bruker Avance 300 spectrometer (operating at 300.13 MHz for ^{1}H and 75.47 MHz for ^{13}C NMR). Solutions containing 0.25 mol dm^{-3} of pyridine-N-oxide in each sample aimed to be characterized, and a solution of tetramethylsilane (TMS) in pure deuterated water (99.9% D) as an internal standard, were used in NMR tubes adapted with coaxial inserts. The TMS/D_2O solution was always used as the inner part of the concentric tubes, while each sample was used in the outer part of the NMR tube. Using this approach, it is possible to guarantee that the TMS standard and D_2O are not in direct contact with the sample, avoiding possible interferences or deviations in the ^{13}C NMR chemical shifts. The Mnova software (Santiago de Compostela, Spain) was used for data processing. The parameter α_{24} is determined as the ^{13}C NMR chemical shift of carbon 2, relative to that of carbon 4 of pyridine-*N*-oxide (PyO) probe and represents the solvent hydrogen bond donor acidity of aqueous media [23] in protein solution.

5. Conclusions

- The earlier model of four different structures for water represented by exactly four Gaussian spectral components and describing the OH-stretch band in the ATR-FTIR spectra of solutions of non-ionic polymers and inorganic salts [20], is also applicable to protein solutions.
- The effects of different proteins on the structure of water depend strongly on the nature and spatial arrangement of the solvent-exposed protein groups.
- The solvent hydrogen bond donor acidity in aqueous protein solutions (α_{24}) may be estimated by the NMR technique described in [23], which has not previously been applied to protein solutions.
- For the globular proteins in the aqueous solutions examined here, the contributions of Gaussian components II and III are strongly correlated with the hydrogen bond donor acidity of water.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1: Figure S.1 showing further examples of Figure 1 in the text. Table S.1. Relative contributions of Gaussian components I-IV at indicated fixed wavelengths obtained from the literature for all proteins and concentrations measured; Table S.2. Solvent hydrogen bond donor (HBD) acidity, α_{24}, in solutions of proteins.

Author Contributions: Conceptualization, B.Y.Z. and V.N.U.; methodology, P.P.M. and A.R.T.; software, A.I.B. and W.V.M.; validation, J.A.M.J., E.K.M. and A.E.S.; investigation, P.P.M., A.R.T. and L.A.F.; data curation, P.P.M., A.R.T. and L.A.F.; writing—original draft preparation, B.Y.Z. and V.N.U.; writing—review and editing, A.R.T., E.K.M., J.A.M.J., W.V.M., A.E.S., V.N.U. and B.Y.Z.; visualization, L.A.F., B.Y.Z. and V.N.U.; supervision, B.Y.Z.; project administration, B.Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: P.P.M. acknowledges University of Aveiro, CICECO-Aveiro Institute of Materials for funding in the framework of the project UIDB/5011/2020 and UIDP/50011/2020, financed by national funds through the FCT/MEC contract foreseen in the numbers 4, 5, and 6 of the article 23, of the

Decree-Law 57/2016, of August 29, changed by Law 57/2017, of July 19. This work was performed while E.K.M. served at the National Science Foundation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All raw ATR-FTIR data along with the Wolfram Mathematica code used are available upon request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Ball P. Water as an active constituent in cell biology. Chem Rev. 2008;108(1):74-108.
2. Ball P. Water is an active matrix of life for cell and molecular biology. Proc Natl Acad Sci U S A. 2017;114(51):13327-35.
3. Ball P. Biophysics: More than a bystander. Nature. 2011;478(7370):467-8.
4. Ball P. Concluding remarks: Cum grano salis. Faraday Discuss. 2013;160:405-14.
5. Templeton C, Elber R. Why Does RNA Collapse? The Importance of Water in a Simulation Study of Helix-Junction-Helix Systems. J Am Chem Soc. 2018;140(49):16948-51.
6. Zaslavsky BY, Uversky VN. In Aqua Veritas: The Indispensable yet Mostly Ignored Role of Water in Phase Separation and Membrane-less Organelles. Biochemistry. 2018;57(17):2437-51.
7. Bellissent-Funel MC, Hassanali A, Havenith M, Henchman R, Pohl P, Sterpone F, et al. Water Determines the Structure and Dynamics of Proteins. Chem Rev. 2016;116(13):7673-97.
8. You X, Shirley JC, Lee E, Baiz CR. Short-and long-range crowding effects on water's hydrogen bond networks. Cell Reports Physical Science. 2021;2(5):100419.
9. Ferreira LA, Gusev NB, Uversky VN, Zaslavsky BY. Effect of human heat shock protein HspB6 on the solvent features of water in aqueous solutions. J Biomol Struct Dyn. 2018;36(6):1520-8.
10. Ferreira LA, Walczyk Mooradally A, Zaslavsky B, Uversky VN, Graether SP. Effect of an Intrinsically Disordered Plant Stress Protein on the Properties of Water. Biophys J. 2018;115(9):1696-706.
11. Ferreira LA, Loureiro JA, Gomes J, Uversky VN, Madeira PP, Zaslavsky BY. Why physicochemical properties of aqueous solutions of various compounds are linearly interrelated. Journal of Molecular Liquids. 2016;221:116-23.
12. Ferreira LA, Uversky VN, Zaslavsky BY. Role of solvent properties of water in crowding effects induced by macromolecular agents and osmolytes. Mol Biosyst. 2017;13(12):2551-63.
13. Ferreira LA, Madeira PP, Breydo L, Reichardt C, Uversky VN, Zaslavsky BY. Role of solvent properties of aqueous media in macromolecular crowding effects. J Biomol Struct Dyn. 2016;34(1):92-103.
14. Madeira PP, Reis CA, Rodrigues AE, Mikheeva LM, Zaslavsky BY. Solvent properties governing solute partitioning in polymer/polymer aqueous two-phase systems: nonionic compounds. J Phys Chem B. 2010;114(1):457-62.
15. da Silva NR, Ferreira LA, Teixeira JA, Uversky VN, Zaslavsky BY. Effects of sodium chloride and sodium perchlorate on properties and partition behavior of solutes in aqueous dextran-polyethylene glycol and polyethylene glycol-sodium sulfate two-phase systems. J Chromatogr A. 2019;1583:28-38.
16. Titus AR, Ferreira LA, Belgovskiy AI, Kooijman EE, Mann EK, Mann JA, Jr., et al. Interfacial tension and mechanism of liquid-liquid phase separation in aqueous media. Phys Chem Chem Phys. 2020;22(8):4574-80.
17. Madeira PP, Titus AR, Ferreira LA, Belgovskiy AI, Mann EK, Mann JA, Jr., et al. Hydrogen Bond Arrangement Is Shown to Differ in Coexisting Phases of Aqueous Two-Phase Systems. Biomolecules. 2021;11(12).
18. Alexandrova AN, Jorgensen WL. Why urea eliminates ammonia rather than hydrolyzes in aqueous solution. J Phys Chem B. 2007;111(4):720-30.

 doi:10.20944/preprints202208.0327.v1

19. Masuda K, Haramaki T, Nakashima S, Habert B, Martinez I, Kashiwabara S. Structural change of water with solutes and temperature up to 100 degrees C in aqueous solutions as revealed by attenuated total reflectance infrared spectroscopy. Appl Spectrosc. 2003;57(3):274-81.
20. da Silva N, Ferreira LA, Belgovskiy AI, Madeira PP, Teixeira JA, Mann EK, et al. Effects of different solutes on the physical chemical properties of aqueous solutions via rearrangement of hydrogen bonds in water. Journal of Molecular Liquids. 2021;335:116288.
21. Chen Y, Zhang Y-H, Zhao L-J. ATR-FTIR spectroscopic studies on aqueous LiClO4, NaClO 4, and Mg (ClO 4) 2 solutions. Physical Chemistry Chemical Physics. 2004;6(3):537-42.
22. Kitadai N, Sawai T, Tonoue R, Nakashima S, Katsura M, Fukushi K. Effects of ions on the OH stretching band of water as revealed by ATR-IR spectroscopy. Journal of solution chemistry. 2014;43(6):1055-77.
23. Madeira PP, Passos H, Gomes J, Coutinho JA, Freire MG. Alternative probe for the determination of the hydrogen-bond acidity of ionic liquids and their aqueous solutions. Physical Chemistry Chemical Physics. 2017;19(18):11011-6.