Journal of Molecular Liquids 238 (2017) 462–469

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Hydration sphere structure of proteins: A theoretical study

Anikó Lábas [a], Imre Bakó [b,*], Julianna Oláh [a,*]

[a] Department of Inorganic and Analytical Chemistry, Budapest University of Technology and Economics, Szent Gellért tér 4, H-1111 Budapest, Hungary
[b] Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar tudósok körútja 2, H-1519 Budapest, P.O. Box 286, Hungary

ARTICLE INFO

Article history:
Received 27 February 2017
Accepted 9 May 2017
Available online 10 May 2017

Keywords:
Hydrogen-bond network
Solvation layers
Insulin
MD simulations

ABSTRACT

Hydration is essential for the proper biological activity of biomolecules. We studied the water network around insulin (as a model protein) in aqueous NaCl solutions using molecular dynamics simulations and statistical analysis of the topological properties (hydrogen bond neighbor number and the interaction energy between hydrogen-bonded water molecules) of the water network. We propose a simple method to define the hydration layers around proteins. Water molecules in the first and second layers form significantly less, but stronger hydrogen bonds with each other than in the bulk phase. Furthermore, water molecules over the hydrophilic and hydrophobic surface of the protein possess slightly different H-bonding properties, supporting the hypothesis of structural and dynamical heterogeneity of the water molecules over protein surface. The protein molecule perturbs the solvent structure at least up to the fourth-fifth hydration layer. Our data suggest the peculiar role of the second hydration shell.

1. Introduction

The appropriate spatial structure is essential for the activity of proteins. It is affected by both intramolecular interactions in the proteins and the intermolecular interactions formed with the solvent molecules, which is water in living cells. It has been well-established that the dominant conformational motions of proteins are profoundly affected by their hydration shell [1,2]. As a consequence, structural changes of the solvent should inevitably affect protein structure and function as well. Indeed, addition of compounds such as inorganic salts, organic molecules, acids or bases to the solution can perturb the structure of liquid water leading to the denaturation of the biomolecule. Among these, the denaturing effect of salts has been the most extensively studied, and more than a century ago Franz Hofmeister ordered the ions according to their ability to precipitate egg-white proteins [3,4]. Kosmotropic ions (e.g. sulfate ion) or water structure makers strengthen the hydrogen-bonding network of bulk water and at the same time decrease the solubility of biomolecules. In contrast chaotropic ions (e.g. nitrate ion) supposedly break the hydrogen-bonding network of bulk water and increase the solubility of biomolecules. Recently, the interfacial tension at the protein–water interface was shown to play a central role in the Hofmeister phenomena [5]. Not only salts, but other chemical agents can also denature proteins. Bennion and Daggett simulated the urea-induced denaturation of chymotrypsin and suggested that the solvent plays various roles in the process. Most importantly, the structure and dynamics of the solvent changed in the solution, and intrusion of the solvent molecules into the hydrophobic core of chymotrypsin was responsible for diminishing the hydrophobic effect and encouraging solvation of the core and thereby changing the intramolecular hydrogen bond network in the protein [6].

Generally, it is accepted that different levels of hydration occur at a biomolecule. In the first hydration layer water interacts with the external surface of the protein through directional hydrogen-bonding (H-bonding) interaction especially on the hydrophilic surface of the protein, while on the hydrophobic surface of the protein the topology, roughness and spatial constrains of the surface orient the water molecules. As a consequence, the hydrogen-bonded properties of water is influenced significantly by the surface properties of macromolecules resulting in increased mean residence time [7–12] and 10–20% increase of the density of water molecules [13,14] compared to the bulk phase. However, a molecular dynamics study on myoglobin, also showed that only those water molecules have very long residence times that are found in cavities and clefts of the protein; other hydration sites of the protein are characterized by residence times similar to the bulk phase [15]. The water-water hydrogen bonding in the first layer forms a spanning, peptide homogeneously enveloped, percolated network, while lack of biological functions is always connected to the broken (not percolated) H-bonded network in the first layer [16]. Recent terahertz spectroscopic measurements, a method sensitive to the collective motion of water molecules, indicate that protein disturb the water structure beyond the 1–2 water layers as previously thought [17]. The radius of the dynamic hydration shell was greater than 10 Å for the studied proteins and correlated well with the dipole moment of the protein [18].

* Corresponding authors.
E-mail addresses: bako.imre@ttk.mta.hu (I. Bakó), julianna.olah@mail.bme.hu (J. Oláh).

http://dx.doi.org/10.1016/j.molliq.2017.05.038
0167-7322/

It is obvious from the above overview that gaining a better understanding of the hydration layer structure around proteins could contribute to our understanding of various processes involving protein-solvent interactions such as protein folding and unfolding. A possible way to study the structure of water is graph theory, which has recently been applied to the hydrogen bond network in various solutions and mixtures, e.g. of water, methanol and ethanol solutions adsorbed in microporous silicalite-1 [19], of ion aggregates in different high salt solutions [20], and of highly concentrated renal osmolyte solutions [21]. Recently, we studied the mixtures of water and formamide, the simplest model of the peptide bond, and showed that these two compounds form microhomogeneous mixtures, in which the number of hydrogen bonds formed by water and formamide are very similar [22].

When graph theory is used to study the structure (i.e. hydrogen bond network) of water, the network of interactions is mapped into a graph [23]. The vertices of the graph correspond to the water molecules and the edges to the hydrogen bonds formed by the water molecules. Once the hydrogen bond network is mapped into a graph, a thorough statistical analysis can be carried out in order to get insight into the water structure.

In the present work we extend this theory to explore the hydrogen bond network around a protein to obtain a better understanding of its structure and how it changes from the surface of the protein towards the bulk phase, and investigate how molecular dynamics simulations can give more insight into recent findings by terahertz spectroscopic measurements that proteins disturb the water structure beyond 1–2 water layers. Furthermore, we explore the effect of salt concentration on the properties of the hydrogen bond network around a protein. We chose insulin as a model protein because of its (1) small size (2) importance in human health (3) and as it has a balanced distribution of hydrophobic and hydrophilic patches on its surface. NaCl was selected as a co-solute to the protein, because Na^+ has a relatively high concentration in the cytosole compared to divalent cations and it has a weak hydrate sphere ordering capacity and Cl^- is by far the most common anion in living organisms.

First we carried out a series of molecular dynamics simulations at different salt concentrations on solvated systems (with and without the protein), then for each snapshot taken from the trajectory of the simulations, we determined the network of hydrogen bonds and transformed it to a graph. Finally, a thorough statistical analysis of the properties of the obtained graphs (hydrogen bond neighbor number and the interaction energy between hydrogen-bonded water molecules) were carried out. Importantly, we present here a simpler approach to define solvent layers around the protein compared to those that have been described in the literature [24–26], and investigate their hydrogen bond properties layer by layer, which enables us to compare the structure and hydrogen bond properties of these layers to those of reference solutions, which do not include the protein molecule. This methodology enables us to characterize the effect of the protein molecule on the hydrogen bond network and to study the structure of its hydration sphere in a statistical way.

2. Methods

2.1. Molecular dynamics simulations (MD)

The crystal structure of the monomer, which is the active form of human insulin (PDB code 3I40 [27]) was used as a starting structure for the MD simulations. Disulfide bonds were created between $Cys_{6,chainA}$-$Cys_{11,chainA}$, $Cys_{7,chain\ A}$-$Cys_{7,chainB}$ and $Cys_{20,chainA}$-$Cys_{19,chain\ B}$. The protonation state of the titratable amino acid residues were determined using the H++ webserver version 3.2 [28–30]. Based on the estimated pK_a values the His_5 residue was doubly protonated in chain B and after visual analysis of their surroundings all other histidine residues were protonated on the ε nitrogen atom. The CHARMM-GUI webserver was used for the system setups and generation of the input files [31], the NAMD software package [32] with the CHARMM27 force field [33] for the minimization of the structure and dynamics simulations. Hydrogen atoms were added using the standard CHARMM protocol [34]. We chose the TIP3P water model, which is a simple 3-point rigid water model, to simulate water as the non-bonded parameters of protein atom types in the CHARMM27 force field were determined to be in line with the TIP3P water model. As a consequence, when the CHARMM program package is used for modelling proteins the TIP3P water model is by far the most frequently used water model, yielding a reliable description of proteins. Therefore, the protein was solvated by TIP3P water molecules arranged in an octahedral shape with 15 Å edge distances. Three differently solvated protein systems were prepared. One contained only one sodium ion in order to generate a neutral system, while two other systems contained sodium and chloride ions in 0.5 and 1.5 molar concentrations. Although the cytosolic concentration of these ions is much lower, we have chosen these relatively high concentrations to obtain improved statistics for the effect of the desalting in the simulation. The ions were placed by a Monte Carlo approach. Reference systems of the salt solutions with 0 M, 0.5 M and 1.5 M concentrations were also prepared; these did not include the protein molecule. Afterwards, each system was minimized for 10,000 steps to eliminate bad initial contacts, followed by a 50 ps long NVT equilibration simulations at 303.15 K temperature. Then 5 ns long NPT Langevin MD simulations were carried out applying 2 fs step size with collecting configurations from every ps. All bonds in the molecules involving hydrogen atoms were kept fixed with the SHAKE [35] algorithm. Periodic boundary conditions (PBC) were used to handle boundary effects. The temperature was set to 303.15 °C in all simulations. The equilibration of the systems was reached by means of temperature reassignment. All of the velocities of the atoms in the systems were periodically reassigned in order to set the desired temperature. Therefore, in every 500 steps the temperature was rescaled during equilibration. The Constant Temperature Control making use of Langevin dynamics was applied together with the Nose-Hoover Langevin piston pressure control with the target pressure set to 1.01325 bar. As the random initial velocity distribution used in the MD simulation could influence the obtained results, all MD simulations were carried out with three different initial velocity distributions yielding 3-3 parallel trajectories for all systems studied.

3. Analysis

3.1. Hydrogen bonds

Two water molecules were regarded hydrogen bonded if the H···O distance between the two molecules was smaller than 2.5 Å and O—H···O angle was larger than 120°. This criteria are reasonable for protein simulations where correlation between the distance and the angle criteria have already been shown [36]. However, we have checked the dependency of our results on this definition by analyzing the trajectories obtained for the protein solvated in 0.5 M NaCl solution by setting the O—H···O angle criterion to 130° and 145° as well.

The average hydrogen bond number, N_{HB}, was calculated by averaging the number of hydrogen bonds over the trajectory and over all molecules (Eq. (1)):

$$N_{HB} = \frac{<\sum_{i=1}^{N} N_{HB,i}>}{2N} \quad (1)$$

where $N_{HB,i}$ is the number of hydrogen bonds around water molecule i, N is the number of water molecules in the simulated box and < > denotes averaging over all snapshots.

We calculated the hydrogen bond energy between water molecules i and j ($E_{HB,ij}$) for each hydrogen bond using the TIP3P force field following the common practice [37] to identify the hydrogen-bond energy with the interaction energy of the H-bonded molecular pair, even if