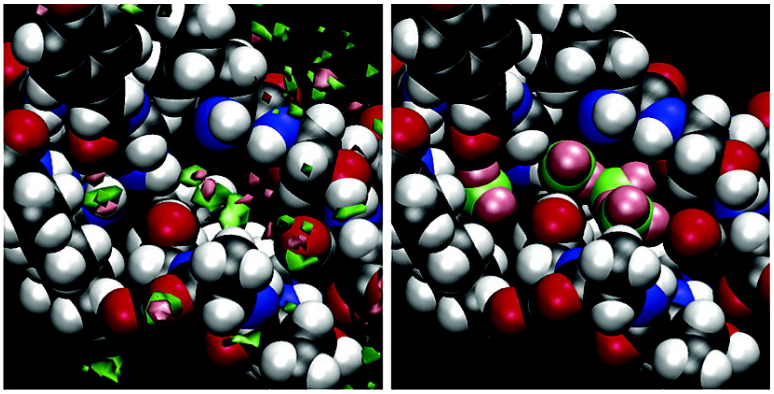
Communication

Water Molecules in a Protein Cavity Detected by a Statistical–Mechanical Theory

Takashi Imai, Ryusuke Hiraoka, Andriy Kovalenko, and Fumio Hirata

J. Am. Chem. Soc., **2005**, 127 (44), 15334-15335• DOI: 10.1021/ja054434b • Publication Date (Web): 18 October 2005



J|A|C|S
COMMUNICATIONS

Published on Web 10/18/2005

Water Molecules in a Protein Cavity Detected by a Statistical−Mechanical Theory

Takashi Imai,[†] Ryusuke Hiraoka,[†] Andriy Kovalenko,[‡] and Fumio Hirata*[,§]

Department of Bioscience and Bioinformatics, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan, National Institute for Nanotechnology, National Research Council of Canada, and Department of Mechanical Engineering, University of Alberta, Edmonton, AB, T6G 2V4 Canada, and Department of Theoretical Studies, Institute for Molecular Science, Okazaki Aichi 444-8585, Japan

Received July 5, 2005; E-mail: hirata@ims.ac.jp

Water molecules confined inside cavities in a protein are of great importance in understanding the structure, stability, and functions of the biomolecule.[1−4] Considerable efforts have been devoted to observe such water molecules by experiments,[5] but it is still a nontrivial task. It is virtually impossible to "find" water molecules in a protein cavity by the ordinary molecular simulation because they are most likely trapped in the biomolecule through a process of large conformational fluctuation or "folding". The simulation of such water molecules is as difficult as the protein folding itself. The three-dimensional reference interaction site model (3D-RISM) theory employed in the present study is a recently developed statistical−mechanical theory of molecular solutions.[6−8] Starting from the solute−solvent molecular interactions, it yields the 3D distribution of solvent around a solute, say, protein, as well as its solvation thermodynamics, such as the free energy and the partial molar volume. The theory has proven itself to be capable of predicting the solvation structure and thermodynamics at least qualitatively.[9,10] It was not entirely sure, however, that the theory can be applied to water molecules confined in a small cavity of protein because there has been a common understanding or "prejudice" that statistical−mechanical theories are not designed for such a heterogeneous system. The present study proves unambiguously that the common understanding was really a prejudice.

In this study, we have carried out the 3D-RISM calculation for a hen egg-white lysozyme immersed in water and obtained the 3D distribution function of oxygen and hydrogen of water molecules around and inside the protein. The native 3D structure of the protein is taken from the protein data bank (PDB code 1hel).[11] The protein is known to have a cavity composed of the residues from Y53 to I58 and from A82 to S91, in which four water molecules have been determined by means of the X-ray diffraction measurement.[11] In our calculation, those water molecules are not included explicitly.

Figure 1 gives isosurface representations of the 3D distribution function $g(\mathbf{r})$ ($g(x,y,z)$) of water oxygen around lysozyme. The green surfaces show the areas where $g(\mathbf{r})$ is larger than 2, 4, and 8 in the left, center, and right pictures, respectively. For example, $g(\mathbf{r}) = 2$ means water molecules are distributed twice as probable as in the bulk phase. The green areas in the right picture indicate the positions of the water molecules tightly bound to protein sites. It is found from the isosurface map that water molecules form rather strong hydrogen bonds with almost each of the exposed polar sites.

More significantly, we have observed some peaks of the 3D distribution function in a cavity within the protein. This is highly nontrivial because the cavity was not filled with water molecules explicitly. The peaks are located in the largest cavity surrounded by the residues from Y53 to I58 and from A82 to S91. The left picture in Figure 2 shows the isosurfaces of $g(\mathbf{r}) > 8$ for water oxygen (green) and hydrogen (pink) in the cavity. In the figure, only the surrounding residues are displayed, except for A82 and L83, which are located in the front side. There are four distinct peaks of water oxygen and seven distinct peaks of water hydrogen in the cavity. The spots colored by green and pink indicate water oxygen and hydrogen, respectively. From the isosurface plot, we have reconstructed the most probable model of the hydration structure. It is shown in the center of Figure 2, where the four water molecules are numbered in the order from the left. Water 1 is hydrogen-bonding to the main-chain oxygen of Y53 and the main-chain nitrogen of L56. Water 2 forms hydrogen bonds to the main-chain nitrogen of I56 and the main-chain oxygen of L83, which is not drawn in the figure. Waters 3 and 4 also form hydrogen bonds with protein sites, the former to the main-chain oxygen of S85 and the latter to the main-chain oxygens of A82 (not displayed) and of D87. There is also a hydrogen bond network among Waters 2, 3, and 4. The peak of the hydrogen between Waters 3 and 4 does not appear in the figure because it is slightly less than 8, which means the hydrogen bond is weaker or looser than the other hydrogen-bonding interactions. Although the hydroxyl group of S91 is located in the center of the four water molecules, it makes only weak interactions with them.

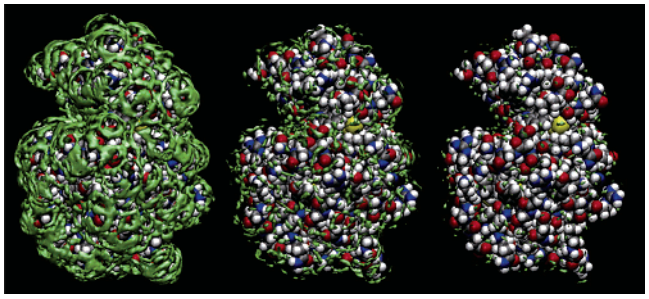


Figure 1. Isosurface representations of the three-dimensional distribution function of water oxygen around lysozyme. The green surfaces show the areas where the distribution function, $g(\mathbf{r})$, is larger than 2 (left), 4 (center), and 8 (right). The lysozyme molecule is represented by the standard space-filling model. VMD software[12] was used for molecular visualization.

It is interesting to compare the hydration structure obtained by the 3D-RISM theory with crystallographic water sites of X-ray structure. The crystallographic water molecules in the cavity are depicted in the right of Figure 2, showing four water sites in the cavity, much as the 3D-RISM theory has detected. Moreover, the water distributions obtained from the theory and experiment are quite similar to each other. Thus the 3D-RISM theory can predict the water-binding sites with great success.

[†] Ritsumeikan University.
[‡] University of Alberta.
[§] Institute for Molecular Science.

10.1021/ja054434b CCC: $30.25

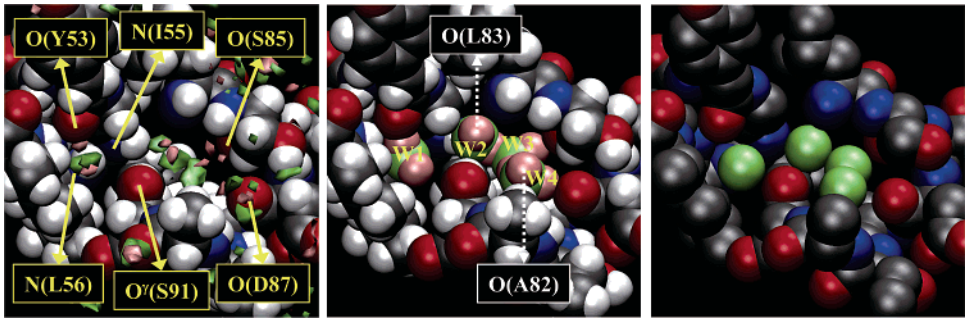


Figure 2. Water molecules in the cavity surrounded by the residues from Y53 to I58 and from A82 to S91. Residues A82 and L83 located in the front side are not displayed. The isosurfaces of water oxygen (green) and hydrogen (pink) for the three-dimensional distributions larger than 8 (left), the most probable model of the hydration structure reconstructed from the isosurface plots (center), and the crystallographic water sites (right). VMD software[12] was used for molecular visualization.

It should be noted that one peak of the 3D distribution function does not necessarily correspond to one molecule. If a water molecule transfers back and forth between two sites in the equilibrium state, two peaks correspondingly appear in the 3D distribution function. In fact, the number of the water molecules within the cavity calculated from the 3D distribution function is 3.8. It is less than the number of the water-binding sites and includes decimal fractions. To explain that, we carried out 1 ns molecular dynamics (MD) simulation using the same parameters and under the same thermodynamic conditions as the 3D-RISM calculation. Only one exception was that the four crystallographic water molecules in the cavity as well as the other crystallographic water molecules were initially put at their own sites in the MD simulation. The result of MD simulation also shows the hydration number less than 4, that is, 3.1. From the MD trajectory, it is found that two inner water molecules, Waters 1 and 2, stay at their own sites for all of the simulation time and make only small fluctuation around the sites. On the other hand, two outer water molecules, Waters 3 and 4, sometimes enter and leave the sites and, by chance, exchange with other water molecules from the bulk phase. As a result, the number of water molecules at the outer sites is 1.1 on average. It is noted here that the value might be underestimated because water molecules entering into a cavity are sampled less than those leaving, and there is a possibility that the simulation does not reach an equilibrium. Thus the 3D-RISM theory can provide the reasonable hydration number including decimal fractions through statistical−mechanical relations, even though the theory takes no explicit account of the dynamics of molecules.

In conclusion, we have presented a rather surprising result demonstrating the unusual ability of the statistical−mechanical theory of molecular solutions for detecting water molecules confined in a small cavity of a protein. The new result will have great impact on the fields of biochemistry and biophysics, including recognition of a drug molecule by a receptor or an ion channel, accommodation of a guest molecule by a host molecule, enzymatic reaction, etc. The result strongly suggests that the theory can access configurations peculiar to molecules confined in a small space if the configuration is stable by any reason. One might be skeptical about the method because the result is presented only for water. So, let us explain how one can readily extend the method to find drug or guest molecules in a cavity of a host (macro)molecule. We replace water solvent in the present calculation by aqueous solution of guest molecules and analyze the 3D distribution of cosolvent (guest) and solvent molecules in a cavity of a host molecule, say, protein. Such extensions are well-established and straightforward in this theory. One may find peaks of either the guest molecule or water molecules, depending on the ratio of their affinities to the host cavity. The higher the peak, the greater the binding affinity, just as we inspected above for water molecules in the protein cavity. The consequence is nothing but the recognition of guest molecules by a host molecule. The study along this line is in progress in our group.

Acknowledgment. The authors thank Dr. Takeshi Yamazaki for his helpful information on computer graphics. This work was supported in part by Grant-in-Aids for the NAREGI Nanoscience Project and for Scientific Research on Priority Area of "Water and Biomolecules" from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MONBUKAGAKUSHO). T.I. is also grateful for the support by a Grant-in-Aid for Young Scientists from the ministry.

Supporting Information Available: Detailed description of our calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Branden, C.; Tooze, J. *Introduction to Protein Structure*, 2nd ed.; Garland: New York, 1999.
(2) Morais-Cabral, J. H.; Zhou, Y.; MacKinnon, R. *Nature* **2001**, *414*, 37.
(3) Zhou, Y.; Morais-Cabral, J. H.; Kaufman, A.; MacKinnon, R. *Nature* **2001**, *414*, 43.
(4) Tanimoto, T.; Furutani, Y.; Kandori, H. *Biochemistry* **2003**, *42*, 2300.
(5) Nakasako, M. *Philos. Trans. Biol. Sci.* **2004**, *359*, 1191.
(6) Kovalenko, A. In *Molecular Theory of Solvation*; Hirata, F., Ed.; Kluwer: Dordrecht, The Netherlands, 2003; pp 169−275.
(7) Kovalenko, A.; Hirata, F. *Chem. Phys. Lett.* **1998**, *290*, 237.
(8) Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 7821.
(9) Imai, T.; Kovalenko, A.; Hirata, F. *Chem. Phys. Lett.* **2004**, *395*, 1.
(10) Imai, T.; Kovalenko, A.; Hirata, F. *J. Phys. Chem. B* **2005**, *109*, 6658.
(11) Wilson, K. P.; Malcolm, B. A.; Matthews, B. W. *J. Biol. Chem.* **1992**, *267*, 10842.
(12) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

JA054434B