LA-SUB--95-103

Molecular Dynamics Simulation of Hydration in Myoglobin

WEI GU* AND BENNO P. SCHOENBORN

Life Science Division, Los Alamos National Lab., Los Alamos, N.M.
*Department of Biochemistry, University of New Mexico, Albuquerque N.M.

Telephone number: (505)-665-2033
Fax number: (505)-665-3024

Key words: computer simulation, molecular dynamics,
bound water, hydrogen bonds, neutron analysis

ABSTRACT This study was carried out to evaluate the stability of the 89 bound water molecules that were observed in the neutron diffraction study of CO myoglobin. The myoglobin structure derived from the neutron analysis was used as the starting point in the molecular dynamics simulation using the software package CHARMM. After solvation of the protein, energy minimization and equilibration of the system, 50 pico seconds of Newtonian dynamics was performed. This data showed that only 4 water molecules are continously bound during the length of this simulation while the other solvent molecules exhibit considerable mobility and are breaking and reforming hydrogen bonds with the protein. At any instant during the simulation, 73 of the hydration sites observed in the neutron structure are occupied by water.

INTRODUCTION

The structural role and catalytic effect that water plays in proteins and enzymes is still rather speculative and controversial and relies often on experimental data that is seemingly contradictory. Simple hydrogen bond energy calculations show that in systems at room temperature at least 2 or 3 hydrogen bonds are required to localize a water molecule. Nevertheless in many x-ray diffraction analyses, hundreds of water molecules have been located. This implies that at least in their crystalline form, proteins have well established long range water structures. Some NMR analyses of small proteins have, however, shown that only a few, if any, water molecules are firmly bound to the proteins surface. In order to determine if these observations are mainly due to crystal effects, we studied the results of a neutron investigation of myoglobin derivatives with a dynamics simulation. This study had the potential to explain the fundamental difference between NMR and the neutron data.

In x-ray diffraction, the scattering power of an atom is proportional to the number of its electrons. When applied to protein crystallography with the inherent limited resolution, hydrogen atoms are not observed. Therefore, the position of hydrogens can only be inferred by the oxygen (or nitrogen) atom position, and a direct knowledge of hydrogen bonding is not possible. By contrast, neutron radiation interacts with the nucleus and consequently isotopes, such as H and D, can be readily distinguished in a neutron density-map.[1] Therefore, neutron diffraction experiments can reliably and unambiguously locate the time averaged occupancy of H atoms. The NMR technique is also sensitive to hydrogen bonding but identifies water molecules only if they are bound for periods of time of the order of the protein tumbling time.[2] NMR studies of protein hydration rely on phenomena related to nuclear spin relaxation. A strong interaction of the nuclear Overhauser effects (NOE) is needed to detect the interaction between water protons and polypeptide protons. The strong distance-dependence of this interaction implies that NOEs can be observed only between spatially close protons; in practice this requires an approach distance (dij) <

4.0 Å, where dij is the distance between two protons i and j. The measured NOE intensities directly reflect the cross-relaxation rates, and the existence of hydration water is derived by the signs and values of these cross-relaxation rates.[2] Presently, with this method, at a Larmor frequency of 600 MHz, only hydration water strongly bound to protein atoms with a residence time ~ 300 ps can possibly be detected.[2] Therefore if a hydrogen bond of a "bound" water is broken and even reformed during this time, a NMR experiment would not see this so-called bound water.

This is demonstrated in x-ray diffraction and NMR experiments that have revealed identical locations for the interior water molecules of bovine pancreatic trypsin inhibitor (BPTI), but not for the hydration water on the surface of BPTI.[2]

A neutron diffraction analysis of myoglobin showed that eighty-nine water molecules were well localized on the surface of this protein molecule;[3] while only a few have been reported from NMR experiments.[4] A hydrogen bond energy distribution for the observed 89 water molecules are depicted in Fig. 1. A number of molecular dynamics simulations have been carried out to study hydration effects in myoglobin. [5-8] These studies[5] show that myoglobin is fully hydrated by 350 water molecules. The hydration shell of 350 water molecules is not a uniform monolayer but a patchwork of water clusters, with charged groups covered by two water layers while 37% of the protein surface is uncovered. The number of bound water molecules reported in neutron and x-ray diffraction studies are far fewer than 350 and do not cover the protein's surface. We refer to bound water as those water molecules observed in diffraction studies; in dynamic simulation studies these water molecules remain within a hydration shell of a specific atom for a long time at least for the time span of the simulation. Lounnas and Pettitt[7,8] showed that dynamic simulations are informative when addressing the general problem of protein hydration. They confirmed the general nature of hydration layers as demonstrated by Cheng and Schoenborn.[3] They further postulate that the differences in observed bound water sites between x-ray and neutron analyses are mainly due to the difference between the scattering power of Hydrogen and Oxygen atoms. The differences of observed water sites can be accounted for by the refinements involving different weight of hydrogen bonds and structural variations in different unit cells. In this study we focus on the behavior of the bound water molecules as depicted in neutron maps and start the simulation with the detailed atomic information including hydrogent atoms and bound water molecules as determined by the neutron diffraction analysis. This approach contrasts with the simulations based on stereochemical models. The hydration sites localized in the neutron study is then analyzed for the occupany of water as a function of the time, to explain the observed differences between NMR measurements and the neutron data.

METHODS

The molecular dynamics simulation of myoglobin surrounded with water was performed using the molecular dynamics program CHARMM.[9] The typical potential energy functions of CHARMM were applied, and all the hydrogens were explicitly included. The TIP3 water model[10] was used in these calculations. The starting structure for the simulation, including 89 hydration water, 5 ammonia ions, 1 sulfate ion and a CO ligand molecule was based on the coordinates determined by neutron diffraction.[3] The coordinates of the starting structure were augmented with coordinates of a pre-equilibrated water buffer at 300 K, both centers of mass being set to the laboratory frame origin. Water molecules with oxygen overlapping the protein were removed with an atom cutoff distance for non hydrogen atoms of 2.7 Å. To obtain a manageable data set, water molecules beyond a 9.0 Å distance from the protein's surface were removed. This 9 Å thick water shell contains 2242 water molecules. This protein water system was energy-minimized[11] resulting in a coordinate r.m.s. change for the whole protein of 0.37 Å with an r.m.s. change in backbone positions of 0.26 Å.

The simulation was performed by numerically integrating Newton's equations of motion for all the atoms, with a step size of 0.001 ps. A constant dielectric model was used to calculate electrostatic interactions. A long cut-off distance of 14.5 Å was applied in calculating all non-bonding interactions.

A Switching function was used to calculate the potential energy for van der Waals interactions with a range between 9.0 Å and 13.5 Å.[9] The potential energy function used to calculate electronic interactions between atoms was in the form of a Shifting function.[9] All covalent bonds involving hydrogen atoms were constrained by the SHAKE algorithm.[12] A model of stochastic boundary conditions was applied to the system.[13] Water molecules in the 2 Å shell at the edge of the system were constrained by harmonic forces ranging from 0.0 Kcal/mol to 7.0 Kcal/mol at the outmost region. The total simulation time was 80 ps calculated on a Silicon Graphics computer. The first step in this simulation used 15 ps to heat the molecules to room temperature and the next 15 ps were used to equilibrate the system at room temperature. These 30 ps are defined as the initial period and preceded the 50 ps used to analyze the dynamics of the proteins surface water molecules. The r.m.s. deviation for the averaged dynamic structure for the final 50 ps simulation compared to the energy minimized neutron structure is 0.76 Å for backbone atoms, 1.32 Å for sidechain atoms and 1.1 Å for the whole protein. The r.m.s. deviation to the original neutron coordinates are 1.34 Å for the whole protein; 0.87 Å for backbone atoms and 1.5 Å for side chains.

RESULTS AND DISCUSSION

The water radial distribution function calculated for the equilibrated myoglobin water

complex is shown in Fig. 2 and is typical for such a system.[11] The peak at 1.74 Å is caused by the hydrogen bonding distances between oxygen atoms of water molecules and hydrogen atoms belonging to the protein's polar groups. The second peak at 2.74 Å shows the interactions between water molecules and non H atoms of the protein. The second peak is a little higher than the normal water-water distribution function, this fact may be explained by the strong interactions of water molecules with hydrogen bond acceptor groups on the protein's surface.

To compare the results from the neutron diffraction analysis with the dynamic simulation, sites occupied by water in the neutron structure were tested for water occupancy during every time step in the simulation. To evaluate such water sites in the dynamically changing protein structure, water molecules were evaluated relative to their relevant protein acceptor/donor atoms and had to remain within hydrogen bonding distance ≤ 4.7 Å from these atoms. In this way the number of originally occupied hydration sites still showing water at any given time during the simulation were compiled as shown in Fig. 3. At any time during the simulation more than 80% of the hydration sites are occupied by water molecules. But for the whole time span of the simulation only 4 sites were occupied continously by the same water molecules. If we reduce the distance to 4.5 Å within which a water molecule has to be bound to protein atoms, we find only two such water sites. As expected, the tighter the constraints the fewer stable water molecules are observed.

An analysis of the time-averaged occupancy of water at all neutron-determined hydration sites showed that 72 sites have occupancies greater than 70%. This simulation shows that most bound water molecules exist in an equilibrium state. H bonds are beeing made and broken continuously in such a way that a hydrogen bonding loci is occupied for most of the time by a water molecule but not necessarily the same molecule. NMR measurement of protein hydration rely on permanent hydrogen bonds. In other words, the type of fluctuations observed for most of the so-called bound water molecules prevents their observation by NMR measurements. In contrast, diffraction measurement depict water sites with inherently high occupancies. From our simulation, we can see that the number of tightly bound water molecules is much less than the number of hydration waters observed for any instantaneous dynamic structure.[14]

The 4 water molecules bound during the whole time course of the simulation are also present in the neutron map. They all have multiple interactions with protein atoms and have high binding energy in both neutron and simulation structures.[15] The hydrogen bond energies, B factors and occupancies are listed in table 1. Comparisons between these parameters for the diffraction and simulation analysis show similar trends. The major exception is the relatively high B factor for one of the water molecules. This temperature factor of 34 is higher than the average value of 28 and much higher than the B's for the other bound waters. This seems to contradict the results from Loumas and Pettit.[8] It is however possible that the large B value of 34 is caused by two sites shared by the same water molecule. It should also be noticed that the simulations are

carried out in solution while the observed B's are for the crystal case. Observation of the structure itself shows that this particular water molecule can form 2 H bonds in a volume with little stereographic constraints. The average residence time of a given water molecule at a specific site is not simply correlated to the occupancy as postulated by Lounnas and Pettitt.[8] This is demonstrated by the interchange of two water molecules as shown in Fig 4. The neutron Fourier map shows that Water 19 is hydrogen-bonded (Fig 5) to aspartic acid 44, histidine 48 and to water molecule (water 35) with an energy of -8 Kcal/mol. The simulation shows that this hydration site is permanently occupied, but that after 15 ps, water 19 and water 35 exchange positions with each other(Fig 4). This region from residue 44 to residue 48 in the protein molecule is in the highly polar loop-region which has high B factors, suggesting large mobility. The bound water molecules balance the electrostatic interactions, but the large mobility of this loop permits the exchange of water molecules. Most hydration sites observed in the neutron maps that exhibit multiple H bonds to protein atoms are observed to have high occupancy in the simulation. In concurance with Lounnas and Pettit[7] we conclude that the occupancy of hydration sites is correlated to the number of H-bonds shared with the protein.

CONCLUSION

Although the difference between solution and crystal state will have some effects on hydration sites, we can conclude that most bound water molecules observed in the neutron analysis do not have residence times long enough to be detected by NMR experiments. The time average occupancies of these water sites in the simulation are high enough to be observed by neutron diffraction. This highlights the different information obtained by NMR and neutron diffraction experiments. Neutron diffraction gives an average picture of the protein-water configuration and shows the average pattern of hydrogen bonding between water and protein atoms. In other words, neutron diffraction gives a more complete picture of the average interactions between water and protein than NMR does. NMR experiments are sensitive to the dynamic stability of hydration and measure only water sites that remain bound to the protein longer than the simulation time; we call this permanently bound water.

The accuracy and length of the molecular dynamics simulation was limited by the level of our present computational chemistry capability. Thus, the major factors affecting the accuracy of this simulation are the lack of explicit calculations for atomic polarization, the cut-off of long-range non-bonding interactions, and lack of a possible quantum mechanical treatment for exchangeable proton. We expect our limitations to be overcome with the introduction of more powerful computers and better algorithms.

ACKNOWLEDGMENT

This research was done under the auspices of the Office of Health and Environmental Research of the U.S. Department of Energy. The authors acknowledge frequent fruitful discussion with Angel Garcia and editorial comments by Robert Knott. Some of the H bond energy calculations were performed by Fong Shu.

REFERENCES

1. Schoenborn, B. P. "Neutrons in Biology." Plenum Press, New York and London, 1982.
2. Otting, G., Liepinsh, E. and Wüthrich, K. Protein hydration in aqueous solution. Science, 254: 974-980, 1991.
3. Cheng, X. and Schoenborn, B. P. Neutron diffraction study of carbonmonoxymyoglobin. J. Mol. Biol., 220: 381-399, 1991.
4. Koenig, S. H., Brown III, R. D. and Ugolini, R. A unified view of relaxation in protein solutions an tissue, including hydration and magnetization transfer. Magnetic Resonance in Medicine. 193: 77-83, 1992.
5. Steinbach P. J. and Brooks B. R. Protein hydration elucidated by molecular dynamics simulation. Proc. Natl. Acad. Sci. USA, 90: 9135-9139, 1993.
6. Lounnas V., Pettitt B. M., and Phillips G. N. Jr. A global model of the protein-solvent interface. Biophys. J., 66: 601-614, 1994.
7. Lounnas V. and Pettitt B. M. A connected-cluster of hydration around myoglobin: correlation between molecular dynamics simulations and experiment. Proteins, 18: 133-147, 1994.
8. Lounnas V. and Pettitt B. M. Distribution function implied dynamics versus residence times and correlations: solvation shells of myoglobin. Proteins, 18: 148-160, 1994.
9. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S and Karplus, M. CHARMM: A program for macormolecular energy, minimization, and dynamics calculations. J. Comp. Chem., 4: 187-217, 1983.
10. Jorgensen, W. L., Chandrasekhar, J. and Madura, J. D. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79(2): 926-935, 1983.
11. Brooks III, C. L., Karplus, M. and Pettitt, B. M. "Proteins: A theoretical perspective of dynamics, structure and thermodynamics." John Wiley & Sons,

New York, 1988.

12. van Gunsteren, W. F. and Berendsen, H. J. C. Algorithms for macromolecular dynamics and constraint dynamics. Mol. Phys. 34: 1311-1327, 1977.
13. Brooks III, C. L., Br_nger, A. and Karplus, M. Active site dynamics in protein molecules: A Stochastic boundary molecular-dynamics approach. Biopolymers. 24: 843-865, 1985.
14. Garcia A. E., Stiller L. Computation of the mean residence time of water in the hydration shells of biomolecules. J. of Comp. Chem. 14: 1396-1406, 1993.
15. Schoenborn, B. P., Gu, W. and Shu, F. Hydrogen bonding and solvent in proteins. In: "Synchrotron Radiation in the Biosciences." Oxford Univ. Press, 1994.
16. Boobbyer, N.A., Goodford, J.P., McWhinnie, P.M. and Wade, R.C. New hydrogen-bond potentials for use in determining energetically favorable binding sites on molecules of known structure. J. Med. Chem. 32: 1083-1094, 1989.

Table 1

Hydrogen bond energies (Kcal/mol) are calculated according to Boobbyer for the 4 permanently bound water molecules as observed in the dynamic simulation in solution and in the neutron map. The energy calculated includes only water to protein terms and excludes any water to water interactions. The target atoms listed are the protein atoms with the strongest hydrogen bond energies with the given water molecule. Hydrogen bonds are only included with energies lower than -0.5 Kcal/Mol. Water molecules are identified according to the nomenclature used by Cheng & Schoenborn.[3]

		Dynamic Simulation		Neutron Structure			
Water #	Target atom	energy	#H-bonds	energy	# H-bonds	B-factor	occupancy
12	TYR103 OH	-5.08	2	-4.65	2	34.22	1.00
16	ILE75 O	-6.96	3	-6.10	3	24.23	1.00
23	GLU4 O	-5.29	2	-3.93	2	23.62	1.00
33	HEM154 OB31	-6.49	3	-5.47	3	23.19	0.79

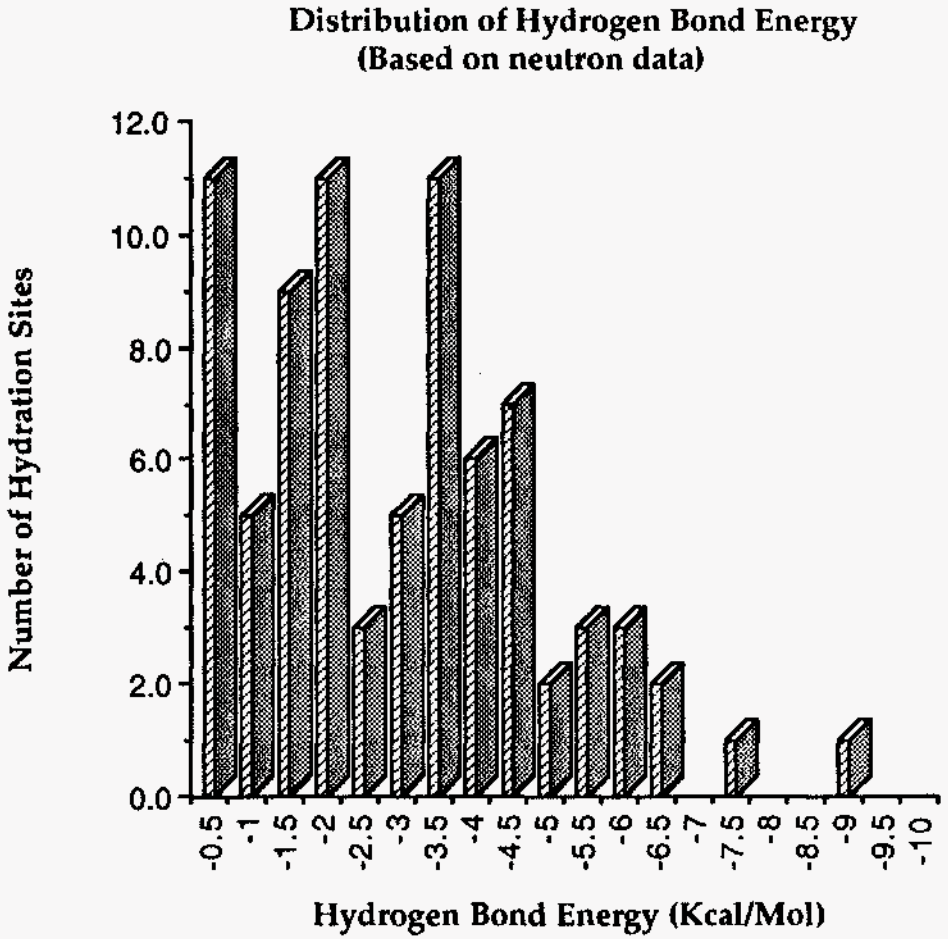
Fig. 1: The distribution of hydrogen bond energies for the 89 hydration waters determined by a neutron diffraction analysis. The hydrogen bond energies are calculated according to Boobbyer's.[16] Only contributions from protein atoms to each water site are considered.

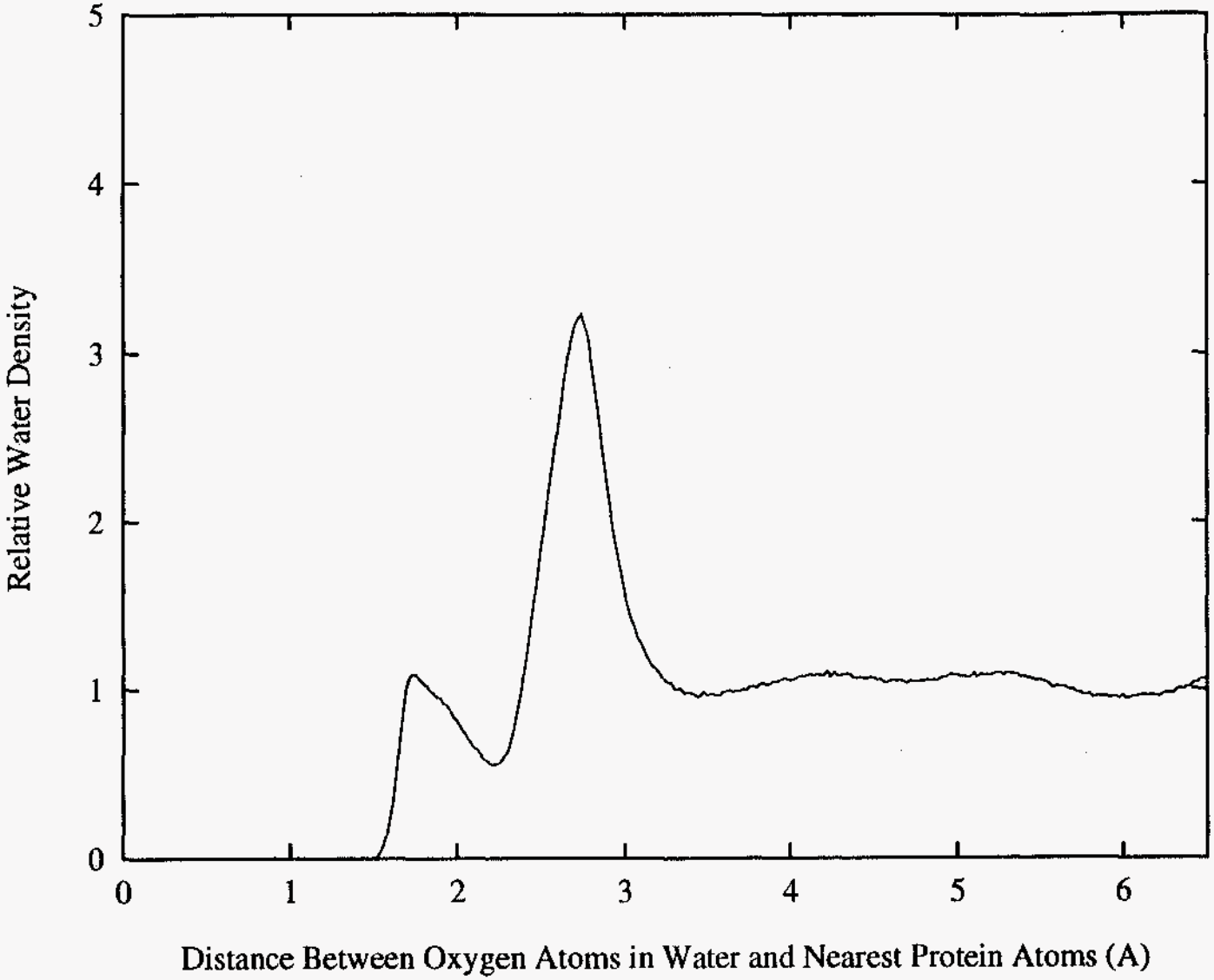
Fig. 2: Water-protein radial distributions calculated by using trajectory coordinates at an interval of 0.02 *ps.*

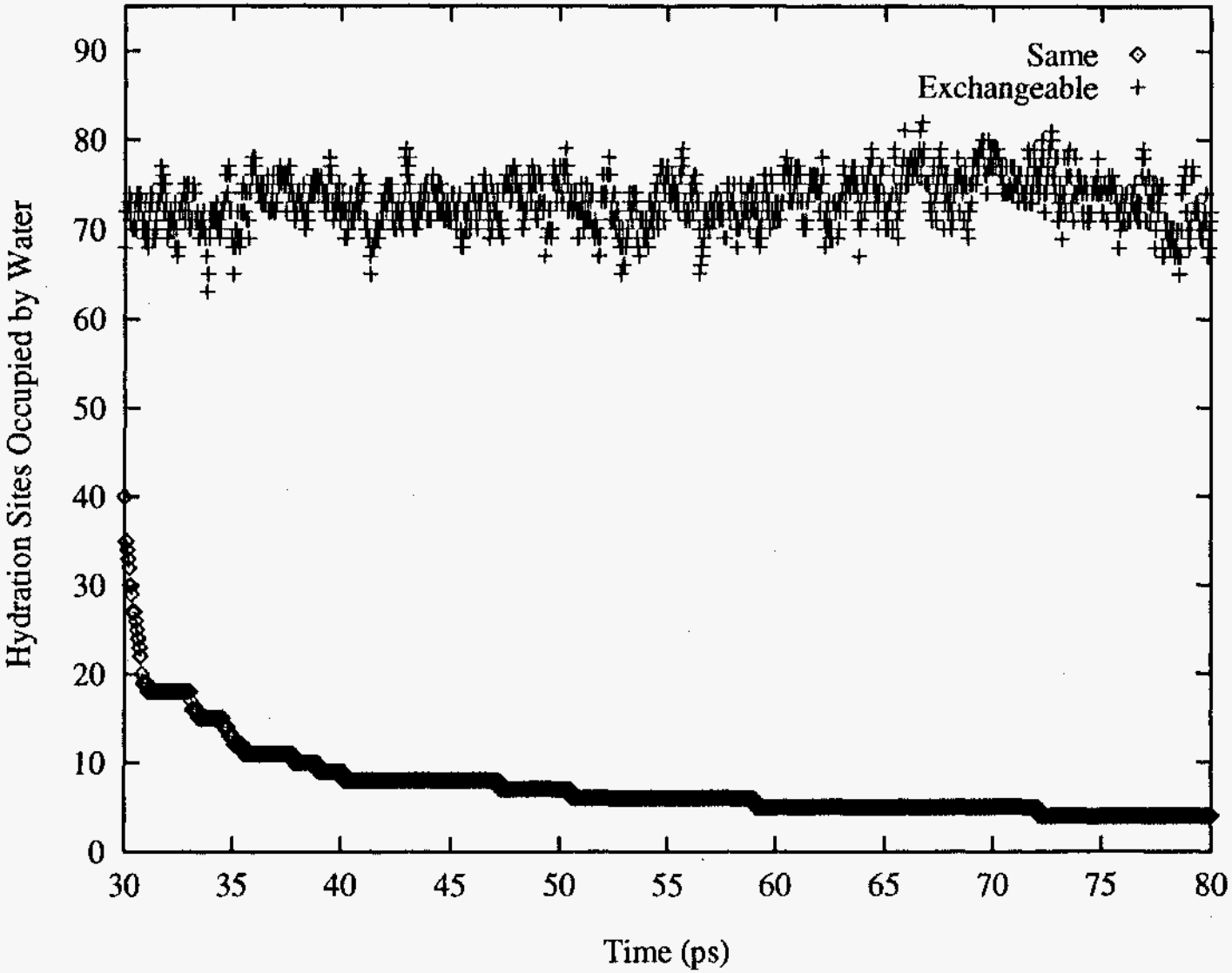
Fig. 3: The number of hydration sites observed in the simulation based on the 89 hydration waters determined in the neutron diffraction analysis . The hydration site is occupied by water which is allowed to shift from the original mean position no more than 2.0 Å, which requires hydrogen bond distance ≤ 4.7 Å. This calculation used original coordinates of hydration sites and the relevant bound protein atoms as observed in the neutron data as the reference. The bottom curve shows the number of hydration site occupied by the same water at each time step. The time interval is 0.04 *ps.* The top curve shows the number of hydration sites occupied by water at a given time step

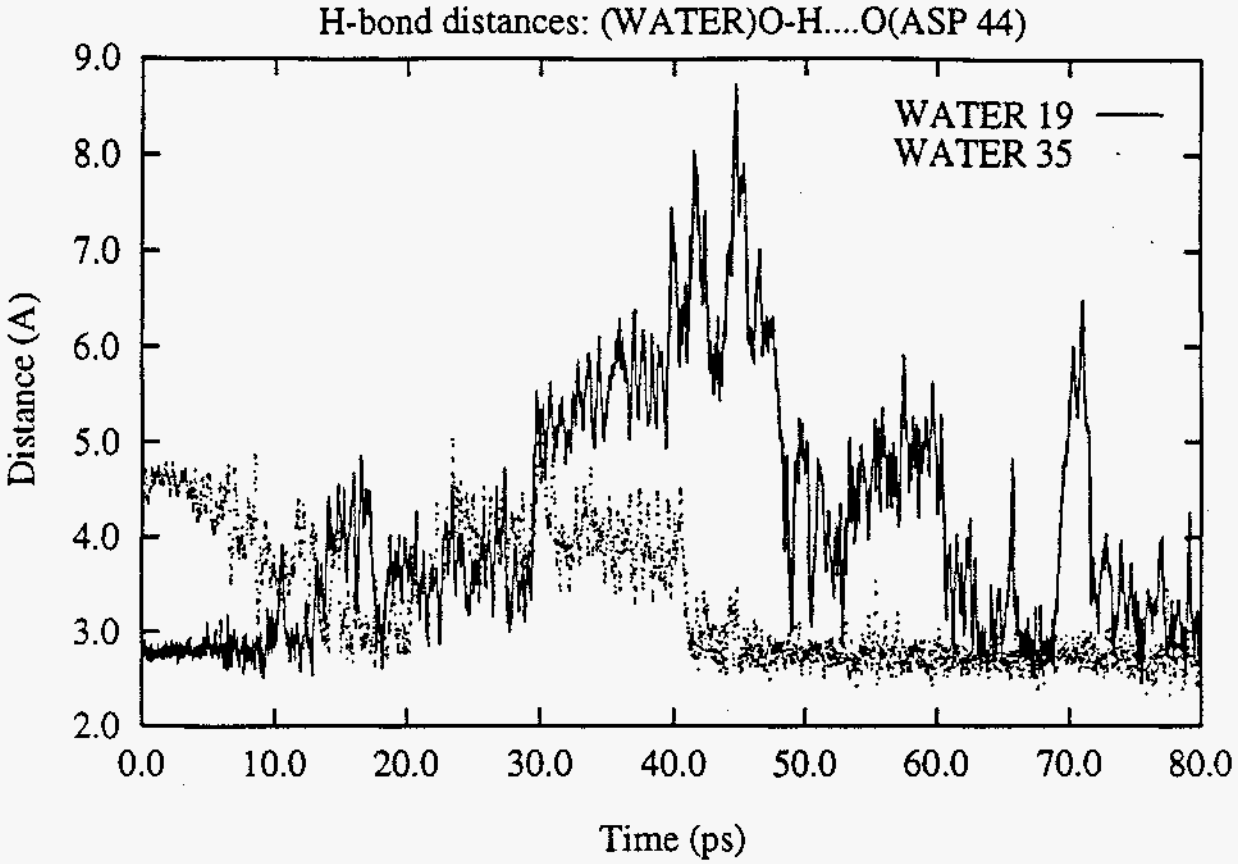
Fig. 4: The observed variation in hydrogen-bond length are depicted over the time course of the simulation for water molecules 19 and 35. The solid lines show variations in bond length between water 19 bound to aspartic acid 44 or histidine 48; while the dotted lines depict the same for water 35.

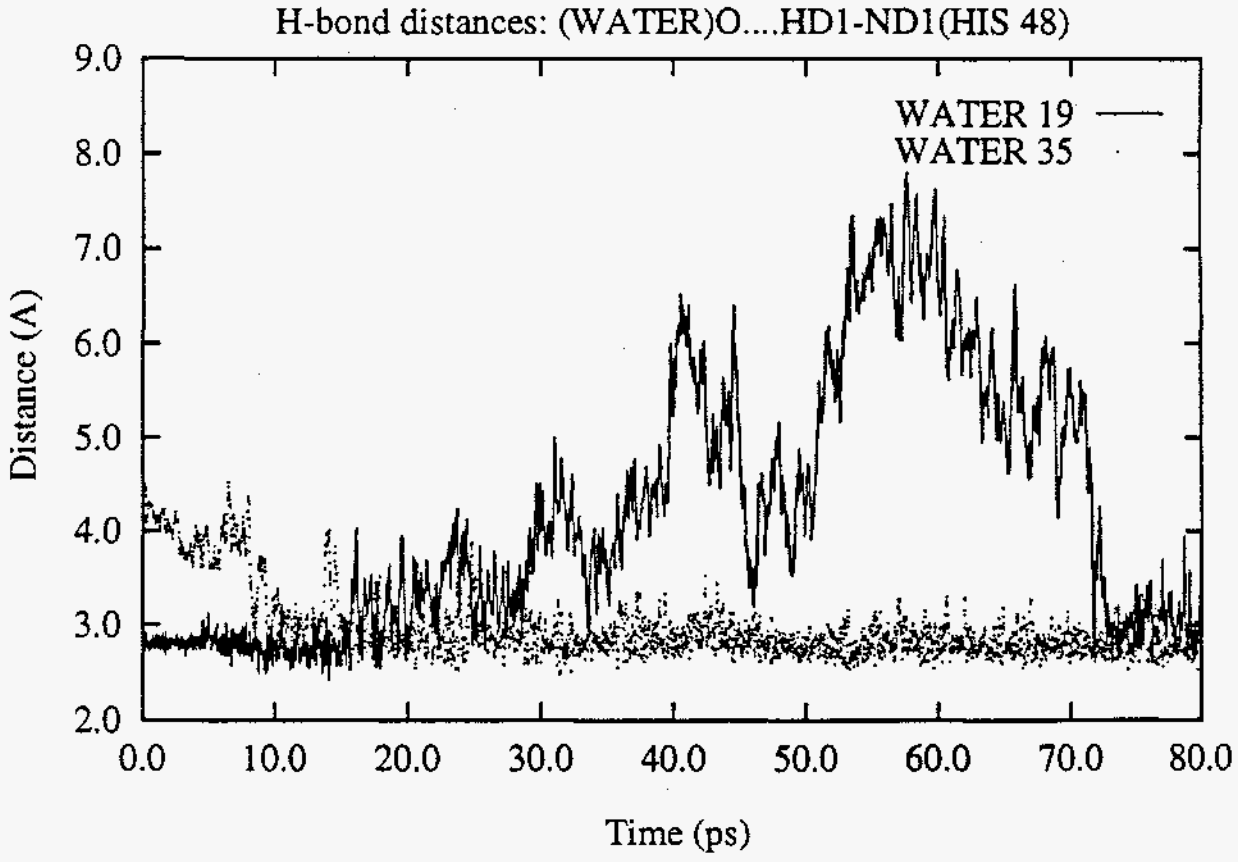
Fig. 5: The hydrogen-bonding geometry for the observed water molecule 19 according to the myoglobin structure derived from the neutron map. The hydrogen bonds are marked as purple dashed lines; oxygen atoms are displayed in the red; nitrogen atoms are shown in the blue; hydrogen atoms are displayed as white; carbon atoms and deuterium atoms are displayed in green.

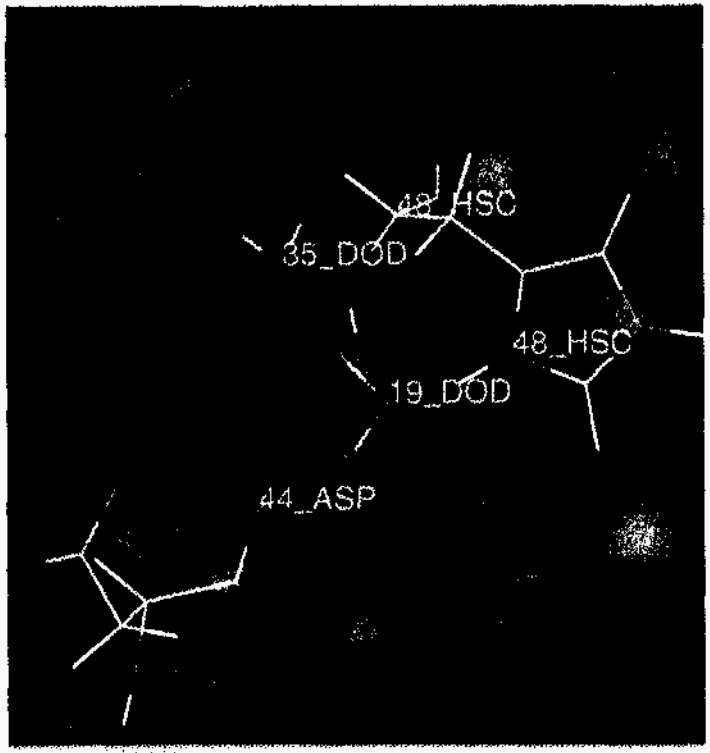

Distribution of Hydrogen Bond Energy
(Based on neutron data)
Number of Hydration Sites
12.0
10.0
8.0
6.0
4.0
2.0
0.0
-0.5
-1
-1.5
-2
-2.5
-3
-3.5
-4
-4.5
-5
-5.5
-6
-6.5
-7
-7.5
-8
-8.5
-9
-9.5
-10
Hydrogen Bond Energy (Kcal/Mol)


Relative Water Density
0
1
2
3
4
5
0
1
2
3
4
5
6
Distance Between Oxygen Atoms in Water and Nearest Protein Atoms (A)


Occupancies of Hydration Sites (Neutron Data)
Observed During the Simulation
Same
Exchangeable
Hydration Sites Occupied by Water
Time (ps)
0
10
20
30
40
50
60
70
80
90
30
35
40
45
50
55
60
65
70
75
80


H-bond distances: (WATER)O-H....O(ASP 44)
WATER 19
WATER 35
Distance (A)
Time (ps)
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0


H-bond distances: (WATER)O....HD1-ND1(HIS 48)
WATER 19
WATER 35
Distance (A)
Time (ps)
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0


48_HSC
35_DOD
48_HSC
19_DOD
44_ASP