ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/44642166

Molecular dynamics simulations of Zn(2+) coordination in protein binding sites

ARTICLE *in* THE JOURNAL OF CHEMICAL PHYSICS · MAY 2010

Impact Factor: 2.95 · DOI: 10.1063/1.3428381 · Source: PubMed

CITATIONS
9

READS
178

2 AUTHORS:

Richard Tjörnhammar
KTH Royal Institute of Technology
10 PUBLICATIONS **32** CITATIONS
SEE PROFILE

Olle Edholm
KTH Royal Institute of Technology
78 PUBLICATIONS **4,052** CITATIONS
SEE PROFILE

Available from: Olle Edholm
Retrieved on: 16 January 2016

Molecular dynamics simulations of Zn^{2+} coordination in protein binding sites

Richard Tjörnhammar and Olle Edholm
Department of Theoretical Physics, Theoretical Biological Physics, Royal Institute of Technology (KTH), AlbaNova University Center, SE-106 91 Stockholm, Sweden

(Received 22 January 2010; accepted 19 April 2010; published online 24 May 2010)

A systematic molecular dynamics (MD) study of zinc binding to a peptide that mimics the structural binding site of horse liver alcohol dehydrogenase (HLADH) has been conducted. The four zinc binding cysteines were successively mutated into alanines to study the stability, zinc coordination, and free energy of binding. The zinc ion is coordinated to four sulfurs in the native peptide as in x-ray structures of HLADH. When the cysteines are replaced by alanines, the zinc coordinating sulfurs are replaced by waters and/or polypeptide backbone carbonyl oxygens. With two or fewer cysteines, the coordination number increases from four to six, while the coordination number varies between four and six with three cysteines depending on which of the cysteines that is replaced by an alanine. The binding free energies of zinc to the proteins were calculated from MD free energy integration runs to which corrections from quantum mechanical cluster calculations were added. There is a reasonable correlation with experimental binding free energies [T. Bergman *et al.*, Cell. Mol. Life Sci. **65**, 4019 (2008)]. For the chains with the lowest structural fluctuations and highest free energies lower coordination numbers for zinc are obtained. Finally, x-ray absorption fine structure spectra were calculated from the MD structures. © *2010 American Institute of Physics*. [doi:10.1063/1.3428381]

I. INTRODUCTION

The zinc metalloprotein[1,2] alcohol dehydrogenase (ADH) was one of the first proteins in which zinc binding sites were recognized. A correct depiction of the coordination around zinc in such sites is vital for the understanding of the role of the binding site. This picture is complex in which the dynamics of different atomic species as well as polarization of nearby molecules all contribute to the binding energy. In horse liver alcohol dehydrogenase (HLADH) two binding sites are present where the first (catalytic) one coordinates the zinc with two histidines and two cysteines or with two cysteines, one histidine, and a water.[3] Other conformations for the catalytic site include zinc either with two histidines, one carboxyl oxygen, and a water or three histidines and a water.[1,4,5] The second (structural) binding site coordinates a zinc with four tetrahedrally arranged cysteine sulfurs.[1,4–8] A similar binding motif, the zinc finger motif, occurs in several proteins and is essential for binding of nucleic acids. Overall the zinc ions are then coordinated in cysteine/histidine repeats (e.g., Cys_2His_2, Cys_4, and Cys_6).[4]

Experimental studies of the zinc binding sites in ADH have shown some detail of the immediate environment[9] and also assessed the importance of carbonyl oxygen coordination as an active bridge in the catalytic reactions that the zinc ion is involved in.[10–13] Zinc binding in HLADH can be categorized into three different types depending on whether the site is catalytic, cocatalytic, or purely structural. The present study is also designed to test computational methods against experimental data,[9] such as the binding free energies and predictions for the zinc coordination in these sites.

The zinc ion ligands are generally thought to influence enzyme function indirectly where the structural features corresponding to selective reactivity are not well understood. It is claimed that the structural zinc site in HLADH is energetically strained compared to the configuration found in truncated 23-residue peptide replicas.[9] This is considered a condition for reactivity of the active sites. As such the zinc ligand coordination and energetics of these truncated peptide replicas can help shed light on the perceived structural to catalytic transition thought to occur for these peptides.

Computational efforts have previously been set forth in order to simulate the chemically active site, while the structural site has drawn less attention. This has still been described at different levels of theory in a few publications.[14,15] Since zinc is a transition metal, it is not obvious that the binding can be described as a purely classical electrostatic binding of an ion with two fixed positive charges to a surrounding with fixed net charge. There is some degree of charge transfer as well as elements of covalent binding. Still, the relatively weak binding affinity of the structural site in HLADH suggests that covalent binding plays a smaller part[16] and classical electrostatic binding can explain the main part.[14]

X-ray diffraction plays a fundamental role in obtaining the configuration of any biological structure and is often employed to yield static information of proteins. The structure of the HLADHs has been revealed in previous crystallographic studies.[17] X-ray absorption fine structure (EXAFS) spectroscopy and x-ray near edge absorption spectroscopy[18] are also well exercised tools for experimentally obtaining detailed information of the binding site coordination in metalloproteins.[19] EXAFS is a very useful technique to

 © 2010 American Institute of Physics

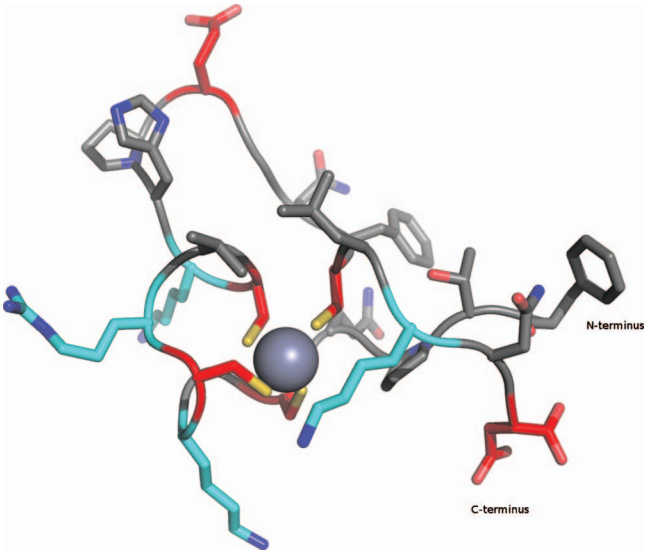

FIG. 1. Depiction of the native zinc binding in HLADH. The amino acids with polar side chains are marked with cyan/red for $\pm 1e$, respectively. The N-terminus corresponds to Phe and the C-terminus to Asp.

probe fine details of the coordination shell. Still it has to be complemented by other experimental techniques and/or theoretical modeling to get an initial structure of the coordination shell.[20] EXAFS does not unambiguously decide which atom belongs to the coordination shell. Oxygen and nitrogen atoms, for instance, have almost indistinguishable scattering properties with nearly identical absorption coefficients.

The ultimate understanding of polarization effects in water and protein as well as the correct sampling of the zinc-protein bonding[21,22] might require calculations at the electron structure level. Such calculations would, however, be very difficult to perform at long enough time scales to obtain the necessary sampling. We therefore use classical force field molecular dynamics (MD) as a main tool in this study, with added quantum mechanical corrections (QMCs) calculated from small clusters. It offers a wealth of fairly reliable information regarding possible candidates for chemical specificity and solvation of the protein backbone. Alternatively one can make use of a classical, polarizable force field with explicit charge transfer between the zinc and coordinating ligands.[21]

II. METHODS

All MD simulations were performed using the MD simulation package GROMACS.[23] The structure of the ADH peptide were taken from the structure 1n8k (Ref. 17) as found in the Research Collaboration for Structural Bioinformatics protein database.[24] One of the binding sites in the ADH sequence was isolated and cropped to correspond to experimentally studied 23-residue sequences[9] (see Fig. 1). The ADH mutants were created by point mutation of the appropriate cysteines to alanines. The full sequences of the binding site of regular HLADH, comprised of four cysteines (C) and four of the mutants containing various amounts of alanine are shown in Table I. Snapshots of the MD structure of the zinc sites of these sequences are shown in Fig. 2. The protonation state of the peptide was chosen with the N-terminus, the arginine, and the three lysines protonated and positively

TABLE I. Residue 93-115 in ADH and mutants. The charge of the side chain is marked with *italic* for $-e$, plain for 0 and underline for $+e$. The N-terminus is positively charged while the C-terminus is negative. The net charge of the peptides ranges from $-2e$ for CCCC to $+2e$ for AAAA. The numbering of the residues corresponds to that of the protein structure 1n8k.

LABEL	Sequence
CCCC	$F_{93}TPQ\mathit{C}_{97}G\underline{K}\mathit{C}_{100}\underline{R}V\mathit{C}_{103}\underline{K}HP\mathit{E}GNF\mathit{C}_{111}L\underline{K}NQ\mathit{D}_{115}$
CCCA	$F_{93}TPQ\mathit{C}_{97}G\underline{K}\mathit{C}_{100}\underline{R}V\mathit{C}_{103}\underline{K}HP\mathit{E}GNFA_{111}L\underline{K}NQ\mathit{D}_{115}$
CCAA	$F_{93}TPQ\mathit{C}_{97}G\underline{K}\mathit{C}_{100}\underline{R}VA_{103}\underline{K}HP\mathit{E}GNFA_{111}L\underline{K}NQ\mathit{D}_{115}$
CAAA	$F_{93}TPQ\mathit{C}_{97}G\underline{K}A_{100}\underline{R}VA_{103}\underline{K}HP\mathit{E}GNFA_{111}L\underline{K}NQ\mathit{D}_{115}$
AAAA	$F_{93}TPQA_{97}G\underline{K}A_{100}\underline{R}VA_{103}\underline{K}HP\mathit{E}GNFA_{111}L\underline{K}NQ\mathit{D}_{115}$

charged while the C-terminus, the aspartic and glutamic acids, and the cysteines were deprotonated and negatively charged. This corresponds to the expected protonation of the free amino acids at pH=7.5 (based on the pK_A values of the free amino acids) except for the cysteines where we assume that the presence of the close zinc ion will shift the pK_A and leads to deprotonation. Monte Carlo simulations to predict the protonation state was also conducted in order to test the charge state of the peptide.[25] The only amino acid except for cysteine with a prominent shift in protonation state was histidine that had acquired a higher (pK_A=7.34) than its free solvated counterpart (pK_A=6.3). Since this pK_A is close to the simulated pH the native peptide was simulated at different His protonation states. The net charge of the peptide with neutral His will then range from $-2e$ to $+2e$ depending on the number of cysteine to alanine mutations.

A. Simulation details

The peptides were solvated in cubic 5 nm boxes containing approximately 4000 water molecules. This system was treated with periodic boundary conditions. The volume of the periodic box corresponds to a protein concentration of 13 mM, which is much larger than the experimental ones. We utilized the optimized potentials for liquid simulations all

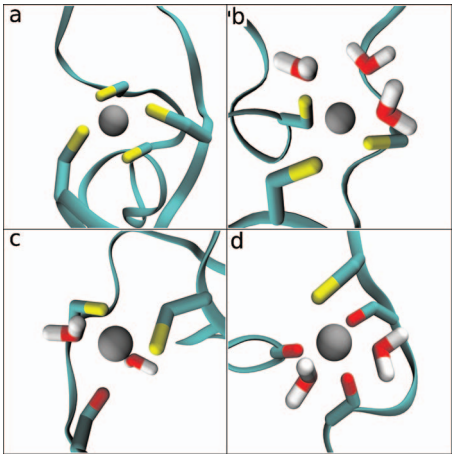

FIG. 2. Representative structures of the zinc binding site of four peptides. (a) The native peptide (CCCC) with four sulfurs coordinating the zinc. (b) The peptide CCCA with three sulfurs and three waters coordinating the zinc. (c) The peptide CAAC with two sulfurs, one carbonyl oxygen and two waters coordinating the zinc. (d) The peptide AAAC with one sulfur, three carbonyl oxygens, and two waters coordinating the zinc.

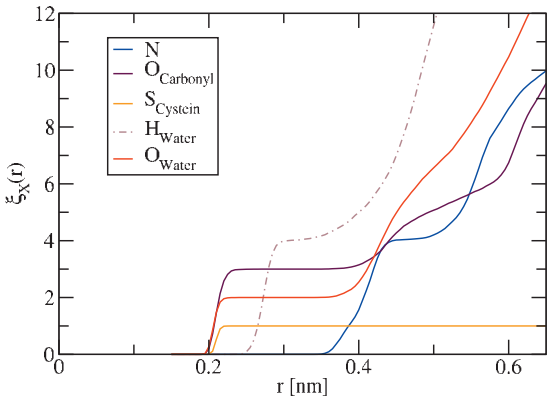

FIG. 3. $\xi_X(r)$ coordination function for the AAAC peptide, where $X = O_{carbonyl}$ denotes carbonyl oxygen.

atoms force field[26] (OPLSAA) for the protein atoms together with a fixed charge of $+2e$ and widely adapted Lennard-Jones parameters from Stote and Karplus[27] for zinc.[28] For the water the TIP4P model[29] was used for consistency with the OPLSAA force field. The bond lengths in the protein were constrained using the LINCS algorithm[30] while the water molecules were constrained using the analytic SETTLE algorithm.[31]

The equations of motion were integrated using a leap frog algorithm with a time step of 2 fs. Simulations to equilibrate the system were performed during 2 ns using Berendsen coupling algorithms to obtain a temperature of 300 K and the isotropic pressure of 1 atm. The electrostatics was treated using fast particle mesh Ewald (PME) summation in the Fourier space.[32] The Lennard-Jones interactions were truncated at 1.0 nm. A neighbor list that was updated every tenth step was used to speed up calculation of the nonbonded interactions in real space. In the (production) MD runs the V-rescale thermostat[33] was employed instead of the nonstochastic thermostat used during equilibration. The free energy integration was done in the NVT ensemble and simulations were run for at least 20 ns for each intermediate state.

B. Structural properties

The coordination was determined by calculating the function

$$\xi_X(r) \equiv 4\pi\rho_X \int_0^r g_{IX}(r')r'^2 dr' = \sum_{i\in X}^{N_X} \theta(r - r_i),$$
$$\theta(r) = \begin{cases} 0 & \text{if } r < 0 \\ 1 & \text{if } r \geq 0, \end{cases} \tag{1}$$

where ρ_X is the number density of atom type X and $g_{IX}(r)$ is the pair correlation function between the zinc ion and atom type X. For a uniform density (pair correlation function one) ξ_X would grow monotonously as r^3. In our case (see Fig. 3), a pronounced plateau between 0.2 and 0.4 nm is prevalent with an integer value for $\xi_X(r)$ showing the number of atoms of type X coordinating the zinc ion in a first neighbor shell.

To study the nonelectronic polarization of water and neutral protein residues, the polarization was calculated as a function of distance from the zinc ion as

$$P(r) = \frac{1}{4\pi r^2 \Delta r} \sum_i q_i r_i, \tag{2}$$

where the sum goes over all fractional charges in a neutral spherical shell of thickness Δr at a radial distance r from the zinc ion. The q_i's are the fractional charges and the r_i's are their distance to the zinc ion.

The root mean square distance (RMSD) of the peptide backbone to the backbone 1n8k x-ray structure[17] was calculated after translational and rotational fitting. The root mean square fluctuations (RMSFs) of the C_α's of the peptide mutants were calculated with respect to the average structure after translational and rotational fitting. Both RMSDs and RMSFs were calculated utilizing the implementation found within the GROMACS (Ref. 23) suite. They were calculated from the last 15 ns of the 20 ns simulations.

C. MD-EXAFS and comparison

The x-ray absorption fine structure spectroscopy code FEFF6 (Refs. 18 and 34) was employed to calculate EXAFS spectra from the coordination shells of the different peptides. The clusters used for the calculation were taken from the final equilibrium structure of a specific peptide with solute included and cropped to include 500 atoms, corresponding to those within a sphere of radius about 1.5 nm around the zinc ion. The cluster was configured to contain one hole at the Zn K edge, $E_K = 9.659$ keV, and only the first scattering paths closest to the zinc by setting parameters NLEG to 10, RMAX to 4.5 Å, and $E_0 = 3.0$ eV. Muffin tin potentials and multiple scattering Debye–Waller factors included in and calculated by the small FEFF6 program was used for all the simulations.

A Fourier transform with an applied Sine filter[35] was performed on the generated EXAFS spectra in the window $k = 0.8-12$ Å^{-1} to yield the real space signature of the cluster.

Ten different snapshots of each peptide were used to generate EXAFS spectra which then were averaged. Since EXAFS spectra are very sensitive to details in the closest coordination shell and our classical simulations tend to give slightly too small zinc-sulfur distances (but correct zinc-oxygen distances), a simple procedure to adjust for this (by moving the sulfurs) was implemented.

D. The free energy of binding

The free energy of the binding was calculated via integration over intermediate states. The Hamiltonian of the system was made functionally dependent on a coupling parameter λ with $H(\lambda = 0)$ corresponding to zinc bound in the site and $H(\lambda = 1)$ to a state with zinc deleted. It can then be shown that the free energy becomes

TABLE II. Quantum and classical binding energies for Zn^{2+} clusters (kJ mol^{-1}). R denotes mean Zn–S or Zn–O distances (Å).

Cluster	E_{QM}	E_{MM}	E_{QM}-E_{MM}	R
$Zn^{2+}(SCH_3)_4$	−2722.9	−2620.4	−102.5	2.41
$Zn^{2+}(H_2O)_6$	−1388.8	−1188.0	−200.8	2.12

$$\Delta G = G_{\text{bound}} - G_{\text{free}} = \int_1^0 \left\langle \frac{d\mathrm{H}}{d\lambda} \right\rangle_{\text{NVT}} d\lambda. \quad (3)$$

The integral was discretized using a step of 0.05 in λ. For each λ point, the system was equilibrated for 2 ns after which the integrand was sampled for 20 ns. Similarly, the zinc ion was deleted in water solution and the free energy of binding was estimated from difference in free energy cost to delete the zinc ions in the two different surroundings. For the nonbonded interactions between zinc and the rest of the system soft core interactions were used. Thus smoothly interpolating with soft core parameters $\alpha=0.5$, $\sigma=0.3$, and $p=1.0$ in order to get rid of the singularity in the Lennard-Jones and Coulomb potentials, as described for GROMACS 4.0.[23,36] The error in each intermediate step was estimated using block averaging over the final trajectory, the details of which were covered in Ref. 37.

The net charge of the peptide with zinc varies between 0 and $+4e$ depending on which mutant was considered. When the zinc ion is integrated away this charge decreased with $2e$ to end up in the interval $-2e$–$+2e$. For the zinc ion in water there is a corresponding decrease in the net charge from $+2e$ to 0. The Ewald methods (PME) require a net charge of zero, which was obtained by adding a uniform smooth counter charge density. The alternative way to achieve this by adding explicit counterions is less reliable unless one uses very long equilibration and sampling times to probe the positions of these counterions.[38] Three peptides—CCCC, CCAC, and AACA—were chosen to explore possible effects of different protonations. In all cases the protonation was changed to give a neutral peptide without the zinc ion, in the CCCC case by protonating Glu107 and Asp115, in the CCAC case by protonating Asp115, and in the AACA case by deprotonating Arg101. This stabilized the state with zinc bound by 20, 40, and 10 kJ mol^{-1}, respectively. This indicates that protonation effects, although not excessive, might have to be considered. The default protonation described previously was utilized for all other simulations.

1. Corrections to the free energy

Since the classical modeling does not include electronic configuration effects, these were estimated by performing cluster calculations on systems consisting of $Zn^{2+}(SCH_3)_4$ and $Zn^{2+}(H_2O)_6$, as in Ref. 14. A B3LYP exchange density functional was used with the 6-31++G(3d,p) basis set as implemented in the program GAUSSIAN.[39] The quantum mechanical (QM) and classical binding energies of these clusters are given in Table II, where R (Å) is the closest zinc-ligand distance. The origin of the difference is in the sulfur case that there is a small covalent contribution to the zinc/sulfur bonds due to the fact that electrons are shared. For the zinc ion in water solution, the two positive charges of the zinc ion cause a polarization of the neighboring waters that increases the classical electrostatic binding energy of each of these waters by about 33 kJ mol^{-1}. QM energy minimizations for all unique coordination spheres in the studied peptides can be found in Sec. III E.

Although there are reasonable evidences for that the cysteine sulfurs are deprotonated with the zinc ion bound to them, the situation is probably different when zinc is released. In the free energy integration scheme the sulfurs are left deprotonated (negatively charged) after the deletion of the zinc ion. The protein can then gain the additional free energy $2.3RT(\mathrm{pK}_A-\mathrm{pH})$ for each protonated one cysteine. Assuming a pH of 7.5 and using the pK_A value of the free amino acid (8.2) gives, as previously discussed,[14] a small correction of 4 kJ mol^{-1} per sulfur.

The experimental dissociation constants K_D (Ref. 9) were converted to standard state free energy differences using the equation

$$\Delta G^0 = -RT \ln(\mathrm{K}_D/C^0), \quad (4)$$

with C^0 being $1M$. In the simulations, the zinc ion probes a volume corresponding to about 20% of this concentration. Therefore, we get a standard state correction $\Delta G_{SS} = -k_B T \ln 5 = -4$ kJ mol^{-1} that further destabilizes the bound states compared to the calculated free energies.

III. RESULTS

A. Coordination

The simulations show a stable fourfold coordination of the sulfurs in the native peptide in agreement with the x-ray structures of LADH proteins. Contrary to this, the zinc ion in water solution coordinates six nearest neighbor water oxygens. In the peptide mutants with two or fewer cysteines, the coordination has switched into sixfold, while the mutants with one alanine and three cysteines show fourfold, fivefold, as well as sixfold coordination. The coordination can clearly be seen from figures such as Fig. 3 while the complete data for all the mutants are given in Table III. It is clear that water oxygens enter and take the coordination role of the lacking sulfurs in the mutants, but it is also seen that one or more backbone carbonyl oxygens may enter and coordinate the zinc instead of sulfurs or water oxygens. There is no clear preference for water or carbonyl oxygens although the average number of water oxygens coordinating the zinc ion is about twice as large as the number of carbonyl oxygens. Still there are some mutants with equally many or more carbonyl than water oxygens coordinating the zinc. Since the coordination number is calculated as an average over 20 ns of simulation, it could very well be a noninteger number if there are fluctuations between differently coordinated conformations during this time. The coordination numbers remain, however, very close to integers in most cases, indicating that such fluctuations are rare on this time scale. It is only for the CCAC peptide that a substantial deviation from an integer coordination number is observed. In this case we have sixfold coordination 1/3 of the time and fivefold coordination during 2/3 of the time. The x-ray structure as well as the

TABLE III. Number of different atoms in first Zn^{2+} coordination shell. Plateau values for the first coordination sphere with mean distances (in Å) to water and carbonyl oxygens and sulfurs. Underline and *italic* correspond to positive and negative residue charges, respectively. $\mathcal{R}_c^O$ is the one letter amino acid residue code for the coordinating carbonyls.

Solvent	O_w	O_c	S	$\mathcal{R}_c^O$	R_{O_w}	R_{O_c}	R_S
CCCC	0.00	0.00	4.00	⋯	3.39	3.99	2.06
ACCC	0.00	1.98	3.00	$\underline{K}_{99}$, F_{110}	3.52	2.08	2.07
CACC	1.05	0.00	3.00	⋯	2.06	4.86	2.08
CCAC	2.33	0.00	3.00	⋯	2.11	4.57	2.09
CCCA	2.98	0.00	3.00	⋯	2.12	5.06	2.08
AACC	3.00	1.00	2.00	F_{110}	2.08	2.09	2.10
CAAC	2.00	1.00	2.00	P_{95}	2.06	2.07	2.09
CCAA	2.00	1.99	2.00	$\underline{K}_{99}$, C_{100}	2.07	2.08	2.09
ACAC	3.00	1.00	2.00	F_{110}	2.07	2.10	2.10
CACA	2.99	1.00	2.00	C_{97}	2.07	2.11	2.08
ACCA	2.00	1.98	2.00	G_{98}, V_{102}	2.08	2.10	2.12
CAAA	4.00	1.00	1.00	N_{109}	2.03	2.10	2.13
ACAA	4.00	1.00	1.00	$\underline{K}_{99}$	2.03	2.09	2.12
AACA	3.00	2.00	1.00	$\underline{K}_{99}$, L_{112}	2.04	2.07	2.10
AAAC	2.00	3.00	1.00	V_{102}, G_{108}, F_{110}	2.03	2.04	2.10
AAAA	6.00	0.00	0.00	⋯	2.00	⋯	⋯
Water	6.00	0.00	0.00	⋯	2.00	⋯	⋯

native peptide has 11 hydrogen bonds contributed from the protein/peptide to the cysteine sulfurs, which corresponds to 2.75 hydrogen bonds per sulfur. The mutant peptides all exhibit stable hydrogen bond networks with an average of 2.5 ± 0.6 hydrogen bonds per sulfur. The mutants with a single cysteine have the largest fluctuations (± 1.0) around the same average of hydrogen bonds per sulfur.

The average minimum distances between the zinc ion and other atoms, such as water oxygens, backbone oxygens, or cysteine sulfurs, denoted R_{O_w}, R_{O_c}, and R_S, respectively, are also tabulated in Table III. They are in good agreement with similar distances in other simulations.[40] The sulfur to zinc distances are, however, about 10% too short compared to experimental x-ray distances and to distances obtained from QM calculations on smaller clusters of atoms. This is an artifact of the classical modeling that is present for sulfur but not for oxygen. The closest coordination shell was void of nitrogen in all the studied peptides. Nitrogen appears first at a distance of 3.82 Å from the zinc. Experimental EXAFS data[9] indicate the coordination of one of the nitrogens in His to zinc in the CAAC-peptide. The histidine was kept neutral (unprotonated) in the main simulations. To test whether the results were sensitive to the details of the histidine modeling, simulations were performed for the CAAC peptide with different histidine protonation (moving the proton between the two nitrogens in the neutral state and in a positively charged state with both nitrogens protonated). This did not alter any results. However, longer simulation times might be necessary in order to capture this possible coordination event.

The backbone carbonyl oxygens involved in coordinating the zinc usually belong to fairly close neighbors to the cysteines. One could have suspected that the negatively charged glutamic or aspartic acid would enter and take the coordinating role of the mutated cysteine. This was never observed in the present simulations. For the aspartic acid this might be due to its peripheral position in the sequence. In the case of the glutamic acid, we note the presence of a neighboring trans-proline that might play a role in hindering it to coordinate the zinc. The residues involved in the zinc coordination are summed up in Table III and denoted with one letter codes in the column $\mathcal{R}_c^O$. For the sequence numbering we refer to Table I.

It seems to be the position in the sequence rather than the side chain of the amino acid that determines which residues that may coordinate the zinc. Most common are Phe and Val which both are hydrophobic and Lys, which is positively charged. What these three amino acids have in common is that they are located just before a cysteine in the sequence and are thus close to the zinc.

One clear discrepancy between the present data and previous simulations can be found for the CCAC peptide for which[14] it reported a coordination shell with three waters. Here we have two or three with an average of 2.33 from a substantially longer simulation. This shows that we get some exchange of zinc ligands in 20 ns, but it is clear that the sampling of different possibilities may be insufficient.

B. Polarization

The radial part of the orientational molecular polarization (of water as well as polypeptide) was calculated around the zinc ion in water and around zinc-containing peptides with one to four cysteines using Eq. (2). Out of the peptides the ones with a given amount of cysteine with largest amount of bound water were chosen. The results are shown in Fig. 4 together with the polarization expected in a continuum environment with large relative dielectric constant,

$$P_{\text{cont.}}(r) = \frac{q_I}{4\pi r^2}\left[1 - \frac{4\pi}{3}\left(\frac{r}{L}\right)^3\right], \tag{5}$$

where $q_I = 2e$ is the ionic charge and the last factor accounts for the uniform background charge. For the zinc ion in water

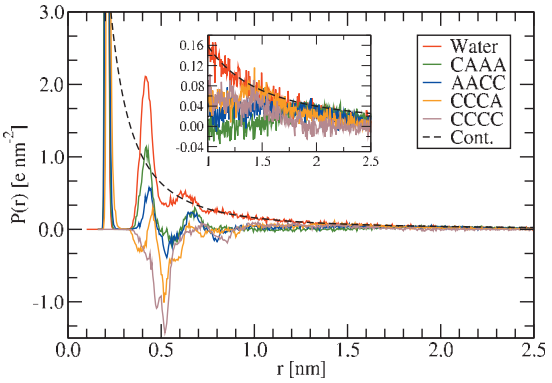

FIG. 4. Zn^{2+} polarization function for some representative polypeptides.

solution, we note a good fit to this expression beyond about 0.7 nm. It is apparent that the polarization decays faster around the polypeptides and that the CCCC peptide is better described by a $1/r^3$ function as would be expected for a dipole field. From the 0.5 nm region it is clear that the binding site structure screens the ionic potential. For two and less sulfurs the water is still polarized and becomes essentially zero going between two and three sulfurs in the binding site. The first coordination sphere becomes increasingly negative when enclosing larger amount of negative charge, thus making the polarization contribution of the second coordination sphere increasingly negative. As such protein carbonyl and water dipole moments are facing away from the zinc ion in this region of the binding site structure.

C. Structure

The root mean square distances of the backbone of the equilibrated peptides (with and without zinc) with respect to the protein x-ray structure 1n8k (Ref. 17) are shown in Table IV. We note that the native peptide with zinc stays close to the protein x-ray structure (0.19 nm), while the RMSD of the native peptide after release of zinc as well as the RMSDs of all mutants with and without zinc are substantially larger (0.43–0.93 nm). This indicates that the zinc release as well as the mutations induce substantial structural changes in the peptide backbone.

The RMSFs also shown in Table IV are typically 0.12 nm or four times smaller than the RMSDs for the zinc coordinating peptides indicating that the structures are stable with zinc present. Without zinc the fluctuations are a factor 2 larger indicating less rigid but still stable structures.

TABLE IV. rms deviations and fluctuations. rms (D) distances and (F) fluctuations for the peptides in nm.

Peptide	D(λ=0)	D(1)	F(λ=0)	F(1)
CCCC	0.19±0.02	0.43±0.04	0.06	0.24
AACA	0.56±0.05	0.51±0.03	0.06	0.25
ACCC	0.44±0.05	0.58±0.02	0.07	0.16
CACC	0.60±0.06	0.73±0.03	0.09	0.42
CCAC	0.64±0.06	0.62±0.05	0.09	0.29
ACAA	0.56±0.05	0.51±0.03	0.09	0.21
AAAC	0.45±0.02	0.43±0.03	0.10	0.18
CCCA	0.65±0.11	0.74±0.05	0.12	0.12
AACC	0.54±0.02	0.59±0.03	0.12	0.26
CAAC	0.62±0.03	0.70±0.09	0.12	0.26
ACCA	0.80±0.04	0.93±0.09	0.12	0.61
ACAC	0.70±0.08	0.61±0.07	0.15	0.42
CAAA	0.62±0.03	0.62±0.05	0.16	0.24
CCAA	0.71±0.05	0.61±0.04	0.18	0.18
CACA	0.73±0.17	0.49±0.03	0.35	0.17

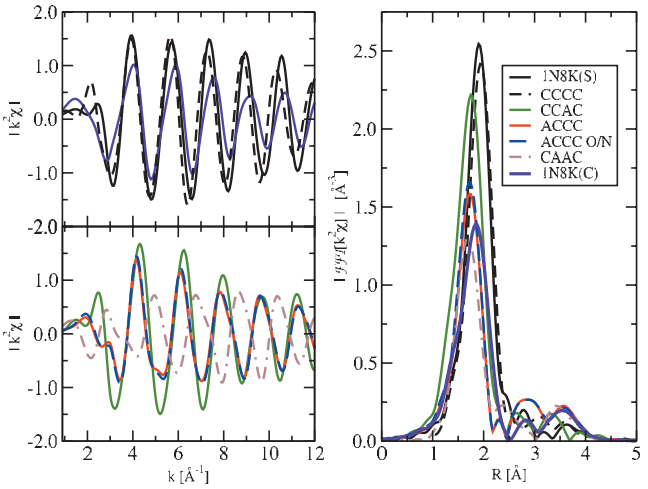

FIG. 5. Simulated EXAFS spectra (left side) and their Fourier transform (right side). ACCC O/N denotes the ACCC peptide with one of its carbonyl oxygens replaced with nitrogen. 1N8K(S) is the structural site. 1N8K(C) is the catalytic site.

D. EXAFS

EXAFS spectra calculated from the simulated structures of selected mutants are shown in Fig. 5, while some characteristic features of the spectra are tabulated in Table V and compared to the same quantities from an experimental study.[9] The classical simulations resulted in a 9.5% too short sulfur to zinc distance. Compensating for this by moving the sulfurs 0.02 nm further off from zinc resulted in structures that produce EXAFS spectra, which agree better to experimental ones. All present data are based on such adjusted structures. The EXAFS spectra are sensitive to details of the structure around zinc as well as to which type of atoms that is located there. Oxygen and nitrogen are, however, quite similar. This has been illustrated, replacing one of the carbonyl oxygens in the ACCC structure by a nitrogen. It is seen

TABLE V. Fourier transformed EXAFS data from Fig. 5. FWHM denotes full width half maximum and PP denotes peak position.

Peptide	PP (Å)	FWHM (Å)	$PP_{expt.}$ (Å)	$FWHM_{expt.}$ (Å)
1N8K(S)	1.90	0.52	⋯	⋯
1N8K(C)	1.85	0.63	⋯	⋯
CCCC	1.93	0.53	1.78	0.80
ACCC	1.73	0.53	⋯	⋯
ACCC[a]	1.73	0.53	⋯	⋯
CCAC	1.75	0.59	⋯	⋯
CAAC	1.70	0.54	1.71	0.87

[a]Innermost oxygen exchanged with nitrogen.

TABLE VI. Zinc-ligand QM and MD binding energy (kJ/mol) comparison for model clusters using B3LYP/6-31++G(3d,p) with zero point energies (ZPEs), where Nr specifies the amount of ligand molecule in the first coordination shell.

	Ligand/Nr	1	2	3	4	5	6
QM	SCH_3	−1821.87	−2693.87	−2916.74	−2722.9		
	H_2O	−321.73	−762.41	−993.75	−1174.49	−1291.72	−1388.8
	OCH_2	−432.48	−822.79	−1048.24	−1220.81	−1301.11	
MM	SCH_3	−1303.26	−2257.01	−2696.61	−2620.4		
	H_2O	−265.71	−523.05	−762.95	−982.93	−1084.81	−1188
	OCH_2	−261.18	−513.13	−750.56	−971.69	−1151.81	
QM-MM	SCH_3	−518.61	−436.86	−220.13	−102.5		
	H_2O	−56.02	−239.36	−230.8	−191.56	−206.91	−200.8
	OCH_2	−171.3	−309.66	−297.68	−249.12	−149.3	

that this has minor effect on the spectra and their Fourier transforms from Fig. 5. Sulfur and oxygen are on the other hand more different. Replacing one of the cysteines in the CCCC peptide by an oxygen results in a new spectra that more resemble that of the ACCC peptide than that of the original CCCC peptide (data not shown in the figure).

One important feature of Fig. 5 is that the Fourier transformed spectra have a weak signature at 3.5 Å, which is also experimentally present and is due to second shell scattering.[18] For the CAAC peptide, the spectrum from the simulated structure lacks the double peak that is observed experimentally[9] (their peptide 7) and has been attributed to the presence of histidine in the coordination sphere. This is explained by the substantial separation (1.35 nm) between the histidine imidazole ring and the zinc present in the simulated CAAC-peptide. It is also clear from the 1N8K catalytic site structure that such a histidine coordination would yield the experimentally present double peak. In 1N8K(C) the zinc is coordinated by two cysteines and two histidines. To explore this further, simulations were performed with the histidine in different protonation states. With a $pK_A=6$ it should be deprotonated (neutral) and neutral pH. A protonated (positively) charged state should not and did not result in attraction to the positively charged zinc, neither did a switch of protonation between the two nitrogens in the neutral imidazole changes the present results. It might be that the sampling of the peptide is still not sufficient in order to capture the binding site conformation as found in EXAFS experiments. It is clear from our coordination that switching between metastable conformations does occur during the 20 ns simulation.

Finally, we note that the somewhat large difference and strong signature of the 1n8k peptide compared to the rest of the mutants are due to the lack of water in that x-ray structure. However, the signature of the CCCC peptide is still in good correspondence with 1n8k reflecting the fact that the innermost coordination sphere is truly void of water in both cases.

E. Free energies of binding

The binding free energies were calculated from classical free energy integrations, with corrections added according to the description in Sec. II. The main correction was the QMC calculated from small atomic clusters. This is summarized in Table II and yields a correction of 24.6 kJ mol^{-1} per sulfur. The mean statistical error for the calculated binding energy averaged over all the peptides becomes ±13 kJ mol^{-1}, while the error in the experimental figures is substantially smaller, probably a few kJ mol^{-1}.

The quantum and classical energies of model clusters containing variable numbers of SCH_3, H_2O, or OCH_2 are shown in Table VI. The results for H_2O are in good agreement with previous theoretical predictions for zinc hydration.[41] From these QM calculations we see that the carbonyl bind even slightly stronger than water to the zinc. Polarization plays a minor role for the zinc/OCH_2 clusters especially for those with more carbonyls. The charge transfer between the oxygens (water or carbonyl) and the zinc is negligible in contrast to the considerable electron transfer/electron sharing between zinc and sulfur. The cluster calculations give energy minima that are deeper in the QM case compared to the molecular mechanics (MM). The sulfur QM as well as MM energy minimal structure have threefold planar symmetric sulfur coordination.

The correction for the QM effects was conducted for all unique coordinating clusters found in the peptides. These were simulated in order to derive the difference between the minimized QM and MM energies, see Table VII. From these data it is clear that the energy differences are not additive in the number of sulfurs, waters, or carbonyls. The classical free energies are also tabulated in Table VII. The application of the corrections yields qualitatively correct binding energies, but too strong binding for the peptides with two or three cysteines. The mean difference between the calculated and the experimental binding energies for the peptides with two or three cysteines is −36 kJ mol^{-1} ± 25 kJ mol^{-1}. The energy minimized QM structures are, in these cases, notably more dissimilar and generally slightly less dense than the equivalent energy minimized MD structures.

One should also note that only one of the four mutants with one cysteine did bind experimentally, while all four of them did bind in the classical simulations. This result is explained by the quantum corrections. The mutant AAAA,

TABLE VII. Zinc-ligand energies E and G in kJ/mol. Calculations using B3LYP/6-31++G(3d,p) with ZPE included and where R_X is the minimal mean distance (Å) to ligand X in the QM cluster. $QMC=\Delta E_{peptide}-\Delta E_{H_2O}$. ΔG_{prot} is the protonation state correction energy. ΔG_{cls} is the classical free energy integration energy. $\Delta\Delta G_{cls}=\Delta G_{cls,peptide}-\Delta G_{cls,H_2O}$. ΔG_{theor} and ΔG_{expt} are the theoretical and experimental binding energies, respectively.

Peptide	O_w	O_c	S	R_{O_w}	R_{O_c}	R_S	E_{QM}	E_{MM}	ΔE	QMC	ΔG_{prot}	ΔG_{cls}	$\Delta\Delta G_{cls}$	ΔG_{theor}	ΔG_{expt}
CCCC	0	0	4	⋯	⋯	2.41	−2722.9	−2620.4	−102.5	98.3	16.0	−1789.2	−172.9	−58.6	−56.5
ACCC	0	2	3	⋯	4.95	2.28	−2957.4	−2775.5	−181.9	18.9	12.0	−1762.4	−146.1	−115.2	−54.2
CACC	1	0	3	3.77	⋯	2.29	−2940.4	−2762.6	−177.8	23.0	12.0	−1737.7	−121.4	−86.4	−54.1
CCAC[a]	2	0	3	3.56	⋯	2.28	−2973.5	−2814.2	−159.2	41.6	12.0	−1746.8	−130.5	−61.6	−56.5
CCCA	3	0	3	3.67	⋯	2.28	−3006.0	−2892.8	−113.2	87.6	12.0	−1768.0	−151.7	−52.1	−57.7
AACC	3	1	2	2.14	2.23	2.39	−2815.8	−2622.1	−193.7	7.1	8.0	−1721.0	−104.7	−89.6	−50.5
ACAC	3	1	2	2.14	2.23	2.39	−2815.8	−2622.1	−193.7	7.1	8.0	−1746.0	−129.7	−114.6	−45.4
CACA	3	1	2	2.14	2.23	2.39	−2815.8	−2622.1	−193.7	7.1	8.0	−1723.1	−106.8	−91.7	−47.0
CAAC	2	1	2	2.17	4.08	2.25	−2777.6	−2580.2	−197.4	3.4	8.0	−1737.0	−120.7	−109.3	−45.3
CCAA	2	2	2	2.28	2.14	2.25	−2782.3	−2613.5	−168.8	32.0	8.0	−1729.6	−113.3	−73.3	−45.1
ACCA	2	2	2	2.28	2.14	2.25	−2782.3	−2613.5	−168.8	32.0	8.0	−1726.0	−109.7	−69.7	−45.4
CAAA	4	1	1	2.13	2.19	2.47	−2186.8	−2085.2	−101.6	99.2	4.0	−1688.9	−72.6	30.6	>0
ACAA	4	1	1	2.13	2.19	2.47	−2186.8	−2085.2	−101.6	99.2	4.0	−1672.3	−56.0	47.2	>0
AACA	3	2	1	2.13	2.30	2.25	−2260.8	−2101.2	−159.6	41.2	4.0	−1669.5	−53.2	−8.0	−37.7
AAAC	2	3	1	2.16	2.14	2.48	−2189.5	−2107.4	−82.1	118.7	4.0	−1674.1	−57.8	64.9	>0
H2O	6	0	0	2.12	⋯	⋯	−1388.8	−1188.0	−200.8	⋯	0.0	−1616.3	⋯	⋯	⋯

[a]ΔG_{theor} calculated from coordination of 2.33 waters.

without cysteines, did not bind zinc experimentally. In the simulations we also observed a fast release of the zinc in this case in agreement with experiment.

The fairly systematic overestimate of the binding strength in the mutants with two or three cysteines needs some discussion. First, the energy minimization with conjugate gradient methods used both for the QM and MM clusters does not guarantee that the minimum reached is a global minimum. The minimization problem is much easier for the symmetric cluster with zinc and four sulfurs. It is therefore reassuring that the results are good in that case. We note, however, that it is the minimum energies of the classical clusters that have to be reduced to improve the agreement with experiment. We therefore made some trials to search the phase space of the classical cluster better. This included among other things MD with simulated annealing. Still, we were not able to find minima with a lower energy. We are therefore convinced that we have found the global minima of the clusters. The second possibility is that the second neighbor coordination shell and/or exchange between the first and second coordination shells may be of importance in the more complex clusters. We may also get a situation in which hydrogen bonds between ligands replace hydrogen bonds to a second coordination shell and thus changes the coordination to the zinc. To explore this better, one would either have to go to larger clusters or to a QM/MM simulation.

IV. SUMMARY

The coordination of solvated zinc was studied in water and in different peptides. The general trend was the favored sixfold coordination just as in water. The most tightly bound peptide has only fourfold coordination while the coordination approaches sixfold for the loosely bound peptides. The structure of the ADH peptides with less than sixfold coordination had smaller backbone fluctuations than the average peptide. The zinc coordination in water is in good agreement with other theoretical predictions,[41,42] as well as experimental data.[43] The cysteine free AAAA peptide releases zinc in accordance with experimental data and favors water molecule solvation over backbone carbonyl solvation. The binding energies are also seen to qualitatively agree with those found experimentally. For the CCCC and CCCA peptides this correspondence is accurate while the binding energy is generally overestimated for the peptides with two or three cysteines and underestimated for the single cysteine peptide that binds. The QM clusters find energetically stable structures that are less densely packed compared to the initial molecular mechanical structures. We conclude that the correction scheme as suggested in Ref. 14 might require a large cluster size or cluster embedding to produce accurate results for more asymmetric clusters. In the classical simulations we see that rare switching of coordination do occur with lifetimes of 2 ns. Long simulation times become important both to decrease statistical errors and to fully sample all the different protein coordinations.

ACKNOWLEDGMENTS

This work has been supported by the Swedish National Infrastructure for Computing (SNIC) including PDC Center for High Performance Computing and High Performance Computing Center North (HPC2N), and by the Swedish Science Research Council (VR). The author would also like to thank T. Bergman and E. Brandt for valuable comments.

[1] B. L. Vallee and D. S. Auld, Biochemistry **29**, 5647 (1990).
[2] C. Andreini, L. Banci, I. Bertini, and A. Rosato, J. Proteome Res. **5**, 196 (2006).
[3] D. S. Auld and T. Bergman, Cell. Mol. Life Sci. **65**, 3961 (2008).
[4] D. S. Auld, BioMetals **14**, 271 (2001).
[5] B. L. Vallee and D. S. Auld, Proc. Natl. Acad. Sci. U.S.A. **87**, 220 (1990).
[6] C. I. Brändén, H. Jörnvall, H. Eklund, and B. Furugren, *The Enzymes* (Academic, New York, 1975), Vol. 11, pp. 103–190.
[7] T. Bergman, H. Jörnvall, B. Holmquist, and B. L. Vallee, Eur. J. Biochem. **205**, 467 (1992).

[8] J. Jelokova, C. Karlsson, M. Estonius, H. Jörnvall, and J. Höög, Eur. J. Biochem. **225**, 1015 (1994).
[9] T. Bergman, K. Zhang, C. Palmberg, H. Jörnvall, and D. Auld, Cell. Mol. Life Sci. **65**, 4019 (2008).
[10] B. L. Vallee and D. S. Auld, Acc. Chem. Res. **26**, 543 (1993).
[11] S. R. Billeter, S. P. Webb, T. Iordanov, P. K. Agarwal, and S. Hammes-Schiffer, J. Chem. Phys. **114**, 6925 (2001).
[12] T. Dudev and C. Lim, Acc. Chem. Res. **40**, 85 (2007).
[13] U. Ryde, Biophys. J. **77**, 2777 (1999).
[14] E. Brandt, M. Hellgren, T. Brinck, T. Bergman, and O. Edholm, Phys. Chem. Chem. Phys. **11**, 975 (2009).
[15] U. Ryde, Eur. Biophys. J. **24**, 213 (1996).
[16] U. Heinz, M. Kiefer, A. Tholey, and H.-W. Adolph, J. Biol. Chem. **280**, 3197 (2005).
[17] J. K. Rubach and B. V. Plapp, Biochemistry **42**, 2907 (2003).
[18] S. I. Zabinsky, J. J. Rehr, A. Ankudinov, R. C. Albers, and M. J. Eller, Phys. Rev. B **52**, 2995 (1995).
[19] A. V. Poiarkova and J. J. Rehr, Phys. Rev. B **59**, 948 (1999).
[20] N. Dimakis, M. J. Farooqi, E. S. Garza, and G. Bunker, J. Chem. Phys. **128**, 115104 (2008).
[21] D. V. Sakharov and C. Lim, J. Am. Chem. Soc. **127**, 4921 (2005).
[22] A. J. Sillanpää, R. Aksela, and K. Laasonen, Phys. Chem. Chem. Phys. **5**, 3382 (2003).
[23] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, J. Comput. Chem. **26**, 1701 (2005).
[24] See: http://www.rcsb.org for 1N8K.
[25] L. Sandberg and O. Edholm, Proteins **36**, 474 (1999).
[26] W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc. **110**, 1657 (2002).
[27] R. H. Stote and M. Karplus, Proteins **23**, 12 (1995).
[28] S. Obst and H. Bradaczek, J. Mol. Model. **3**, 224 (1997).
[29] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926935 (1983).
[30] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. **18**, 1463 (1997).
[31] S. Miyamoto and P. A. Kollman, J. Comput. Chem. **13**, 952 (1992).
[32] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[33] V. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).
[34] See: http://cars9.uchicago.edu/ifeffit/Ifeffit for FEFF6.
[35] M. Newville, J. Synchrotron Radiat. **8**, 322 (2001).
[36] See: ftp://ftp.gromacs.org/pub/manual/manual-4.0.pdf for the Gromacs4.0 manual.
[37] B. Hess, J. Chem. Phys. **116**, 209 (2002).
[38] S. Donnini, A. E. Mark, A. H. Juffer, and A. Villa, J. Comput. Chem. **26**, 115 (2005).
[39] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
[40] D. Toledo, A. Cordomi, M. G. Proietti, M. Benfatto, and L. J. del Valle J. J. Prez, P. Garriga, and F. Sepulcre, Photochem. Photobiol. **85**, 497 (2009).
[41] M. Hartmann and T. Clark, J. Am. Chem. Soc. **119**, 7843 (1997).
[42] G. Marini and N. Texler, J. Phys. Chem. **100**, 6808 (1996).
[43] T. Ohtaki, H. Yamaguchi, and M. Maeda, Bull. Chem. Soc. Jpn. **49**, 701 (1976).