THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

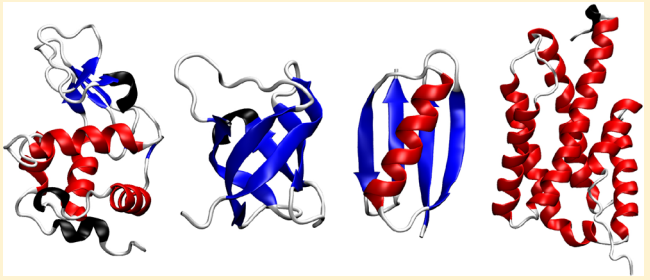
Structural Effects of an Atomic-Level Layer of Water Molecules around Proteins Solvated in Supra-Molecular Coarse-Grained Water

Sereina Riniker, Andreas P. Eichenberger, and Wilfred F. van Gunsteren*

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zürich, Switzerland

Supporting Information

ABSTRACT: Atomistic molecular dynamics simulations of proteins in aqueous solution are still limited to the multinanosecond time scale and multinanometer range by computational cost. Combining atomic solutes with a supra-molecular solvent model in hybrid fine-grained/coarse-grained (FG/CG) simulations allows atomic detail in the region of interest while being computationally more efficient. A recent comparison of the properties of four proteins in CG water versus FG water showed the preservation of the secondary and tertiary structure with a computational speed-up of at least an order of magnitude. However, an increased occurrence of hydrogen bonds between side chains was observed due to a lack of hydrogen-bonding partners in the supra-molecular solvent. Here, the introduction of a FG water layer around the protein to recover the hydrogen-bonding pattern of the atomistic simulations is studied. Three layer thicknesses of 0.2, 0.4, and 0.8 nm are considered. A layer thickness of 0.8 nm is found sufficient to recover the behavior of the proteins in the atomistic simulations, whereas the hybrid simulation is still three times more efficient than the atomistic one and the cutoff radius for nonbonded interactions could be increased from 1.4 to 2.0 nm.



INTRODUCTION

So-called hybrid systems where atomic, fine-grained (FG) and supra-molecular, coarse-grained (CG) particles are combined in a single molecular dynamics (MD) simulation analogously to hybrid quantum/classical (QM/MM) models[1] are a promising approach to access longer time scales and larger system sizes compared to fully atomistic systems while retaining atomic detail for the region of interest.[2−7] The proposed models differ mainly in the way the interactions between the FG and CG particles are handled and parametrized.[8] Recently, we introduced a hybrid model which requires only two parameters to calibrate the interactions between the FG and CG particles.[9,10] The CG water model[11] used represents five FG water molecules by a CG bead with two interaction sites. The two oppositely charged sites connected by an unconstrained bond form a polarizable dipole, thus allowing the explicit treatment of the electrostatic interactions. The CG water model reproduces the experimental density, surface tension and static relative dielectric permittivity.[11] The two parameters for the FG/CG interactions were calibrated based on mixtures of FG and CG water as the properties of water should ideally be independent of the ratio of FG molecules and CG beads.[9] The proposed hybrid model reproduces the before mentioned properties as well as the free energy of solvation of a FG water molecule and FG alkanes.[9]

The hybrid FG/CG model was further tested for four proteins, hen egg-white lysozyme (HEWL), major cold shock protein (CspA), G-protein (GP), and chorismate mutase (CM), solvated in CG water. The four proteins represent different secondary structure motifs, i.e., all α-helical, all β-strand, and combinations thereof. Analysis of the 20 ns of simulation showed the preservation of the secondary structure elements and similar backbone atom-positional root-mean-square deviation (rmsd) from the crystal structure and root-mean-square fluctuations (RMSF) of C_α atoms and radius of gyration compared to the fully atomistic simulations. Although a much larger cutoff radius was used for the long-ranged nonbonded interactions, i.e., 2.0 nm instead of 1.4 nm, the hybrid FG/CG simulations were computationally 1 order of magnitude more efficient than the fully atomistic simulations. The main difference observed between the use of a CG instead of a FG solvent was an increased occurrence of hydrogen bonds formed between protein side-chain atoms due to the lack of hydrogen-bonding partners in the solvent. This resulted in a more negative intraprotein electrostatic energy and a less negative intraprotein van der Waals energy compared to simulation in FG water. Further analysis of the simulations of HEWL in FG water, in CG water, and in vacuo showed that the extent of agreement with the available experimental data is not much decreased in the mixed-grained simulation compared to the fully atomistic one. Nevertheless, the alteration in the protein surface due to the formation of side chain - side chain hydrogen bonds may limit the applicability of the hybrid FG/CG simulation model.

Received: May 1, 2012
Revised: July 4, 2012
Published: July 20, 2012

dx.doi.org/10.1021/jp304188z | J. Phys. Chem. B 2012, 116, 8873−8879

Here, we study the effect of introducing a layer of FG water of varying size around the protein in hybrid FG/CG protein/solvent simulations. Three thicknesses of the FG water layer, 0.2, 0.4, and 0.8 nm, and their influence on the structural properties of the proteins and the computational efficiency are investigated. The simulated averages obtained from the simulations of HEWL are further compared to experimental NMR data.

METHODS

Fine-Grained and Coarse-Grained Models. The proteins were simulated with the atomistic GROMOS force field[12] 54A7, and the simple-point-charge water model[13] (SPC) was used as atomistic fine-grained (FG) solvent. The coarse-grained (CG) solvent model represents five FG water molecules and is described in detail in ref 11. The calibration of the parameters for the FG-CG interactions is described in ref 9.

The four proteins were chosen to cover the diverse secondary and tertiary structure elements commonly found, i.e., purely α-helical, purely β-strand, and combinations thereof. The proteins were hen egg-white lysozyme (HEWL, RCSB protein data bank (PDB) entry[15] 1AKI), major cold shock protein (CspA, PDB entry[16] 1MJC), protein G (GP, PDB entry[17] 1PGB), and chorismate mutase (CM, PDB entry[18] 2FP2) (Figure 1). The protonation states were chosen to match a pH of 7, and hydrogen atoms were added to the X-ray structures according to standard geometric criteria. The histidine side chains were protonated at N_δ or N_ε depending on their hydrogen-bonding environment. In CM, the first three residues were not considered.

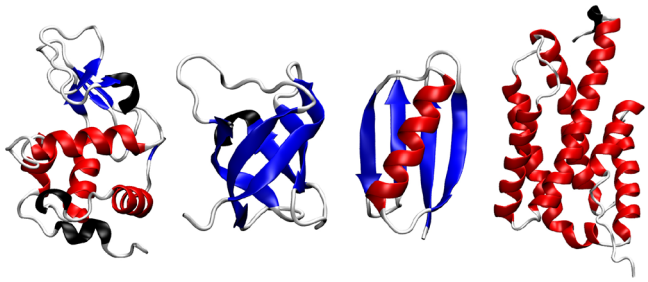


Figure 1. From left to right, crystal structures of hen egg-white lysozyme (HEWL, 129 residues, PDB entry[15] 1AKI), major cold shock protein (CspA, 69 residues, PDB entry[16] 1MJC), protein G (GP, 56 residues, PDB entry[17] 1PGB), and chorismate mutase (CM, 162 residues, PDB entry[18] 2FP2). α-helices are shown in red, 3_{10}-helices in black, and β-strands in blue.

Simulation Details. All simulations were performed for 20 ns under NPT conditions using the GROMOS package of programs.[19−21] The temperature was maintained close to its reference value T = 298 K by weak coupling to a temperature bath with a relaxation time of 0.1 ps.[22] The protein and solvent were coupled separately to the heat bath. The pressure was maintained close to its reference value P = 1.013 bar (1 atm) by weak coupling to a pressure bath with a relaxation time of 0.5 ps and using the isothermal compressibility $\kappa_T = 4.575 \times 10^{-4}$ (kJ mol^{-1} nm^{-3})$^{-1}$. Newton's equations of motion were integrated using the leapfrog scheme[23] with a time step of 2 fs. The bond lengths of the protein and the O−H bond lengths and H−H distance in the FG solvent were constrained to the ideal values applying the SHAKE algorithm.[24] A reaction field force[25] was applied using the experimental relative dielectric permittivity[26] ε_{rf} = 78.5. The contributions to the reaction field by the covalently bound atoms that are excluded from the nonbonded van der Waals and Coulomb interaction were considered in all simulations.

In the fully atomistic simulations, a twin cutoff method was used for the nonbonded interactions with a short-range cutoff radius of 0.8 nm, an intermediate-range cutoff radius of 1.4 nm, and an update frequency of 5 time steps for the short-range pairlist and intermediate-range interactions. In the mixed-grained simulations, a short-range cutoff radius of 1.4 nm and an intermediate-range cutoff radius of 2.0 nm was used. For the electrostatic interactions within the cutoff sphere, a relative dielectric permittivity ε_{cs}^{mix} = 2.3 was used for the FG−CG interactions,[9] ε_{cs}^{CG} = 2.5 for the CG−CG interactions,[11] and ε_{cs}^{FG} = 1 for the FG−FG interactions. Further simulation details of the simulations of the proteins in pure atomistic water and in pure coarse-grained water are given in ref 10.

Setup of the Simulations with an Atomistic Water Layer and Coarse-Grained Water. Three thicknesses of the layer of FG water molecules with respect to the protein surface were considered: 0.2, 0.4, and 0.8 nm. The equilibrated configurations from the fully atomistic simulations were used to select the molecules of the FG water layer, see Table 1. CG water beads were added until a similar volume as in the pure protein in CG water simulations was reached. Although the FG−CG water interaction was parametrized to reproduce the thermodynamic and dielectric properties of liquid water for different mole fractions of FG water in CG water,[9] this does not inhibit spatial mixing of FG water molecules and CG water beads. In order to keep the FG water molecules around the protein, attractive harmonic distance restraints beyond a distance r_0 were applied between the oxygen atoms of the FG molecules and the center of mass (COM) of the protein approximated by the COM of the C_α atoms of four selected

Table 1. Number of Fine-Grained (FG) Water Molecules and Coarse-Grained (CG) Water Beads, Distance r_0 in nm from the Center of Mass of the Protein at Which Attractive Harmonic Distance Restraints Are Applied to the FG Water Molecules in the FG/CG Solvent Simulations, and the Residues Defining the Centre of Mass (COM) of the Protein Used in the Simulations of Hen Egg-White Lysozyme (HEWL), Major Cold Shock Protein (CspA), G-Protein (GP), and Chorismate Mutase (CM)

	FG layer thickness										
	fully FG		0.8 nm		0.4 nm		0.2 nm		fully CG		
protein	FG/CG	r_0	FG/CG	r_0	FG/CG	r_0	FG/CG	r_0	FG/CG	r_0	residues defining COM
HEWL	14332/0		1616/2636	2.9	592/2913	2.5	135/2931	2.2	0/3217		26/35/63/90
CspA	8128/0		1159/1308	2.5	403/1528	2.1	89/1589	1.8	0/1659		7/35/45/56
GP	7444/0		995/1351	2.5	336/1523	2.1	88/1542	1.8	0/1669		14/24/38/50
CM	20643/0		2082/3272	3.6	783/3616	3.2	185/3726	3.0	0/3939		25/37/114/133

dx.doi.org/10.1021/jp304188z | J. Phys. Chem. B 2012, 116, 8873−8879

Table 2. Average Total Distance Restraint Energy V_{dr} for the Half-Harmonic Restraining Function Applied to N_{dr} FG Water Molecules in the 20 ns MD Simulations of Hen Egg-White Lysozyme (HEWL), Major Cold Shock Protein (CspA), G-Protein (GP), and Chorismate Mutase (CM)[a]

	FG layer thickness								
	0.8 nm			0.4 nm			0.2 nm		
protein	V_{dr} kJ mol^{-1}	N_{dr}	V_{dr}/N_{dr} kJ mol^{-1}	V_{dr} kJ mol^{-1}	N_{dr}	V_{dr}/N_{dr} kJ mol^{-1}	V_{dr} kJ mol^{-1}	N_{dr}	V_{dr}/N_{dr} kJ mol^{-1}
HEWL	115	1616	0.07	78	592	0.13	26	135	0.19
CspA	93	1159	0.08	57	403	0.14	19	89	0.22
GP	87	995	0.09	48	336	0.14	17	88	0.20
CM	157	2082	0.08	98	783	0.12	22	185	0.12

[a]The force constant was 300 kJ mol^{-1} nm^{-2}, and the values for r_0 are given in Table 1.

residues, see Table 1. The force constant for the distance restraints was in all cases 300 kJ mol^{-1} nm^{-2}. The FG and CG solvents were energy-minimized and subsequently equilibrated in three steps. In the first 100 ps, the protein atoms were positionally restrained with a force constant of 2.5×10^4 kJ mol^{-1} nm^{-2}, and the volume of the computational box was held constant. In the second step, a NPT simulation of 250 ps length was carried out with a reduced position-restraining force constant of 2.5×10^2 kJ mol^{-1} nm^{-2}, and subsequently a 250 ps simulation was performed without these position restraints.

Effect of the Protein−FG Water Distance Restraints. The distance r_0 between the oxygen atom of a FG water molecule and the approximate center of mass of the protein beyond which the attractive harmonic force inhibits a drifting of the FG water molecules away from the protein was chosen such as to influence the motions of the FG water molecules as little as possible, while keeping them around the protein. Table 2 shows that the average restraint energy per water molecule is very small for the four proteins, less than 0.1 kJ mol^{-1} for the 0.8 nm layer simulations, less than 0.15 kJ mol^{-1} for the 0.4 nm layer one, and less than 0.23 kJ mol^{-1} for the 0.2 nm layer simulation with the least FG water molecules. These numbers are 1 order of magnitude smaller than the thermal energy per degree of freedom at the simulated temperature. The applied distance restraints will thus have a negligible effect upon the structural and motional properties of the proteins.

Analysis. The GROMOS analysis programs to calculate energies *ene_ana*, atom-positional root-mean-square deviations *rmsd*, atom-positional root-mean-square fluctuations *rmsf*, secondary structure elements *dssp*, hydrogen bonds *hbond*, radii of gyration *rgyr*, torsional-angle transitions *ditrans*, NOE distance-bound violations *prep_noe*, *noe* and *post_noe*, ^3J-coupling values *jval*, residual dipolar couplings *svd_fit*, and ^{1}H−^{15}N order parameters *nhoparam* are described in ref 27. Details and formulas of the methods are given in ref 10.

RESULTS AND DISCUSSION

First, the influence of the size of the FG water layer on various nonmeasurable properties is analyzed by a comparison of the configurational ensembles generated in the five different environments, i.e., fully CG water, FG layer size 0.2, 0.4, and 0.8 nm, and fully FG water. Second, a comparison of simulated with measured values for quantities that can be measured or derived from experimental values is presented for HEWL, for which ample NMR data are available.

Structural Properties of the Proteins. The atom-positional root-mean-square deviations (rmsd) of the backbone atoms with respect to the crystal structure are shown in Figure 2. For HEWL and CspA, the rmsd of all simulations are similar, independent of the amount of FG water around the protein.

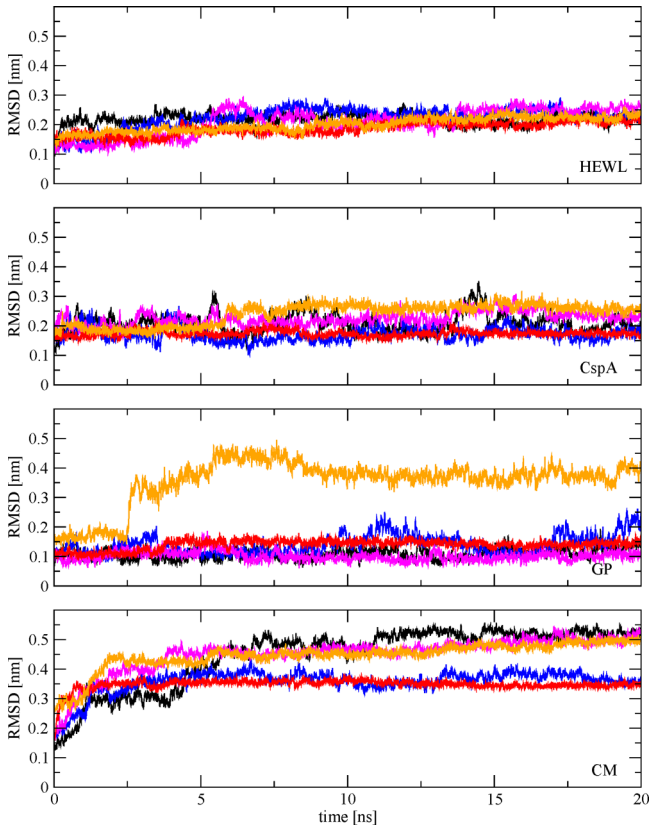


Figure 2. Atom-positional root-mean-square deviations (rmsd) of the backbone with respect to the crystal structures of hen egg-white lysozyme (HEWL), major cold shock protein (CspA), G-protein (GP), and chorismate mutase (CM). Simulations of the proteins in fully FG water (black) are compared to those in CG water with a FG water layer of size 0.8 nm (blue), 0.4 nm (magenta), 0.2 nm (red), and no layer, i.e., only CG water (orange).

The rmsd of GP in pure CG water is increased due to partial unfolding of the helix,[10] which is no longer the case when a FG water layer is present around the protein. In the case of CM, no clear trend is observable. The simulations in pure CG water and with a FG water layer of 0.4 nm give a similar rmsd as the fully atomistic simulation, whereas the rmsd of the simulations with a FG water layer of 0.2 and 0.8 nm is smaller. The radius of gyration is similar in all simulations (Figure S1 in the Supporting Information). The average atom-positional root-

dx.doi.org/10.1021/jp304188z | J. Phys. Chem. B 2012, 116, 8873−8879

mean-square fluctuations (RMSF) over 20 ns of simulation of the backbone C_α atoms, and of the last non-hydrogen atom of each side chain are shown in Figure S2 in the Supporting Information. In most cases, the RMSF values of the simulations with a FG water layer of 0.8 or 0.4 nm are closest to those in pure FG water. However, the large standard deviation of such averages makes a clear interpretation difficult. A clearer picture may be obtained from the RMSF values with respect to the residue number. This is shown for the simulation with a FG water layer of size 0.8 nm in Figure 3 in comparison with the

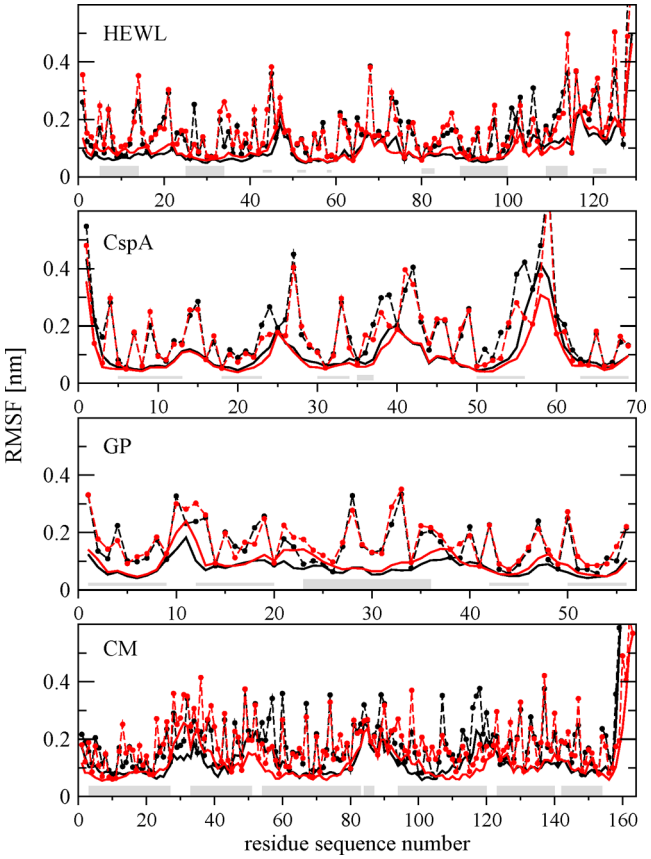


Figure 3. Atom-positional root-mean-square fluctuations (RMSF) of the backbone C_α atom (thick, solid lines) and the last atom of the side chains (dashed lines with filled circles) of hen egg-white lysozyme (HEWL), major cold shock protein (CspA), G-protein (GP), and chorismate mutase (CM) over 20 ns simulation time as function of residue sequence number. Simulations of the proteins in fully FG water (black) are compared to those in CG water with a FG water layer of size 0.8 nm (red). The gray bars at the bottom of each graph indicate the secondary structure elements present in the crystal structure of the protein, i.e., α-helix (thick bars), 3_{10}-helix (medium bars), and β-strand (thin bars).

fully atomistic simulation, and for the other layer sizes and pure CG water in Figure S3–S5 in the Supporting Information. As can be seen in Figure 3, the RMSF of the backbone C_α atoms is largely similar in both simulations. In the case of CM, the fluctuations in some of the loop regions are sometimes lower in the fully atomistic simulation (residues 28–34) and sometimes larger (residues 48–56 and 102–122) compared to the simulation with a FG water layer of size 0.8 nm. For the last atoms of the side chains, the RMSF values are also largely similar in both simulations. For the proteins in pure CG water on the other hand (Figure S5 in the Supporting Information), the RMSF values of the last atoms of the side chains are in most cases smaller than in the fully atomistic simulation, most probably due to the increased number of side-chain hydrogen bonds. The time evolution of the secondary structure elements is shown in Figure S6 in the Supporting Information for the simulations in fully atomic, FG water, in CG water, with a FG water layer of size 0.8, 0.4, and 0.2 nm. The corresponding evolution for the proteins in pure CG water is shown in Figure 4 in ref 10 and revealed the overall preservation of the secondary structure elements under these conditions. The introduction of a FG water layer of varying size does not change this observation. The partial loss of the α-helix in GP in CG water does not occur if a FG water layer is present.

The intraprotein hydrogen bonds (H-bonds) were split into backbone–backbone (bb–bb), backbone–side chain (bb–sc), and side chain–side chain (sc–sc) H-bonds and are shown in Figure 4. While the number and occurrence of the bb–bb

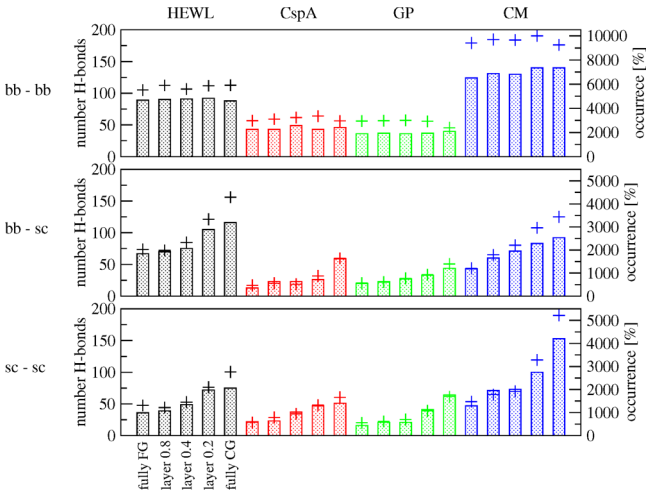


Figure 4. Total number (bars, scale left) and total occurrence (crosses, scale right) over 20 ns of hydrogen bonds in the protein of hen egg-white lysozyme (HEWL, black), major cold shock protein (CspA, red), G-protein (GP, green), and chorismate mutase (CM, blue), split into backbone–backbone (bb–bb), backbone–side chain (bb–sc), and side chain–side chain (sc–sc) hydrogen bonds. The total occurrence is given as the sum of the occurrences of the individual hydrogen bonds obtained using eq 6 in ref 10. For each protein, the simulations are ordered from left to right: fully FG water, FG water layer of size 0.8 nm, of size 0.4 nm, of size 0.2 nm, and no layer, i.e., only CG water.

hydrogen bonds is similar in all simulations independent of the size of the FG water layer, the number and occurrence of the bb–sc and sc–sc hydrogen bonds decreases with increasing thickness of the FG water layer. At a layer thickness of 0.8 nm, a similar number of bb–sc and sc–sc hydrogen bonds with similar occurrence is present as in the fully atomistic simulation, indicating that this layer thickness is sufficient to recover the hydrogen-bonding behavior of the fully atomic simulation while still being computationally more efficient (see Computational Efficiency section). As discussed in ref 10, the observation of an increased number of sc–sc hydrogen bonds in the proteins in CG water is reflected by a more negative intraprotein potential energy of the protein and a less negative intraprotein Lennard-Jones energy, due to hydrogen bonds being electrostatically highly favorable but repulsive in terms of van der Waals interactions. When the number of side-chain hydrogen bonds

dx.doi.org/10.1021/jp304188z | J. Phys. Chem. B 2012, 116, 8873–8879

decreases with increasing thickness of the FG water layer, the intraprotein electrostatic energy V_{CRF} of the protein becomes less negative and the intraprotein Lennard-Jones energy V_{LJ} becomes more negative (Figure 5). At a layer thickness of 0.8 nm, the intraprotein potential energy and its components are comparable to those of the protein in fully FG water.

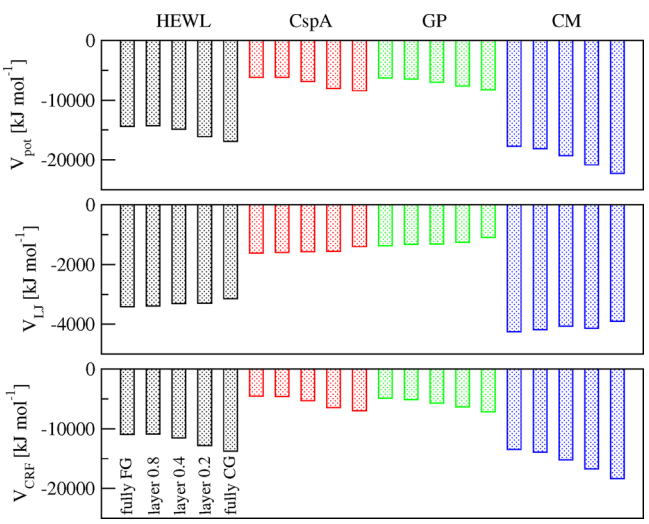


Figure 5. Intraprotein potential energy V_{pot} and its components, the Lennard-Jones energy V_{LJ}, and the electrostatic energy V_{CRF} of hen egg-white lysozyme (HEWL, black), major cold shock protein (CspA, red), G-protein (GP, green), and chorismate mutase (CM, blue). For each protein, the simulations are ordered from left to right: fully FG water, FG water layer of size 0.8 nm, of size 0.4 nm, of size 0.2 nm, and no layer, i.e., only CG water.

The total number of transitions of the backbone torsional angles ϕ and ψ and of the side-chain torsional angles χ_1, χ_2, and χ_3 are listed together with the respective average root-mean-square fluctuation in Table S1, Supporting Information. For the backbone angles, the total number of torsional-angle transitions observed for a protein with a FG water layer of 0.8 or 0.4 nm comes generally closest to the value of the protein in fully FG water. The corresponding average RMSF values are similar in all simulations. The number of transitions and the average RMSF of the side-chain torsional angles generally increase with increasing thickness of the FG water layer, most pronounced for χ_3 where the average RMSF of the proteins in pure CG water was significantly lower compared to the proteins in fully FG water due to the increased number of sc-sc hydrogen bonds.

Comparison to Experimental Data for Lysozyme. The five simulations of hen egg-white lysozyme (HEWL) were compared to experimental data, i.e., NOE distance bounds, ^{1}H−^{15}N order parameters, $^3J_{H_NH_\alpha}$-coupling constants and residual dipolar couplings (RDC).

The distributions of the NOE distance bound violations observed in the five simulations are shown in Figure S7 in the Supporting Information. There were six violations larger than 0.3 nm found in the protein in FG water, which are, however, not crucial as all of them involve the carboxy terminal residue (LEU129) which is very mobile. The same six violations larger than 0.3 nm were also found in the mixed-grained simulations independent of the thickness of the FG water layer. In the range between 0.1 and 0.3 nm, 32 violations were found for the protein in pure CG water, 25 for layer size 0.2 nm, 21 for layer size 0.4 nm, 21 for layer size 0.8 nm, and 22 for the fully atomistic simulation. The ^{1}H−^{15}N order parameters of the protein in the five environments are compared to order parameters derived from experiment in Figure S8 in the Supporting Information. The order parameters from the simulations are generally smaller than those derived from experiment. The largest order parameters are found for residues in helices. This pattern is reproduced roughly by all five simulations. Experimental proton−proton $^3J_{H_NH_\alpha}$-coupling constants of HEWL[28] were compared to calculated values from the simulations. The root-mean-square deviations (rmsd) of the calculated coupling constants from the experimental values are 1.6 (fully FG water and layer size 0.8 nm), 1.7 (layer sizes 0.4 and 0.2 nm), and 1.8 Hz (no layer, i.e., pure CG water). However, the quality of the calculated ^3J-couplings is low because of the inaccuracy due to the empirical nature of the parameters of the Karplus relation which is used to obtain a ^3J-coupling from a configuration.[29,30] Additionally, the sampling in a simulation is limited compared to the experimental time scale.

The quality of fitting the backbone ^{15}N−^{1}H, $^{13}C_\alpha$−^{13}C, and ^{13}C−^{15}N residual dipolar couplings (RDC) calculated from simulation to the experimental RDCs of HEWL[31] is assessed by a so-called Q value (eq 11 in ref 10). The closer a Q value is to zero, the better is the quality of the fit. The distributions of the Q values observed in the five simulations are shown in Figure 6. The values of the simulations are all much higher than the value calculated from the crystal structure.[15] The protein in fully FG water results in Q values closest to zero compared to the mixed-grained simulations, whereas the Q values resulting from the protein in pure CG water are deviating most. With increasing thickness of the FG water layer, the distributions of the RDCs are shifted to smaller Q values. At a layer thickness of 0.8 nm, the distributions from the mixed-grained simulation overlay fully with the distributions from the fully atomistic simulation for all RDCs except the ^{13}C−^{15}N ones. Interestingly, for these RDCs the results from the simulation with a FG water layer of size 0.4 nm performs best compared to those of the fully atomistic simulation.

Figure S6 in the Supporting Information shows that the secondary structure of HEWL is well preserved in the simulation with a 0.8 nm layer of FG water around the protein. Structural differences may though appear when analyzing the parts of the protein not involved in secondary structure. Table S2 in the Supporting Information shows the occurrence of hydrogen bonds involving residues not part of secondary structure elements for the simulation in pure FG water, in a 0.8 nm layer of FG water in CG water, and in pure CG water. All three simulations show similar bb−bb hydrogen-bond occurrences. The bb−sc and sc−sc occurrences in the 0.8 nm FG water layer simulation are significantly closer to the ones observed in the pure FG simulations than the occurrences in the pure CG water simulation. This leads to the conclusion that inclusion of a 0.8 nm FG water layer around the protein largely removes the artifacts of increased intrasolute hydrogen bonding observed for a protein solvated in pure CG water because of the limited hydrogen-bonding capacity of the latter.

Computational Efficiency. The effect of the thickness of the FG water layer on the computational efficiency of the mixed-grained simulations was assessed using a simulation of 5000 steps of HEWL in fully FG water, with a FG water layer of size 0.8, 0.4, and 0.2 nm, and in pure CG water on four CPU's using MPI parallelization. The number of CG beads in the

dx.doi.org/10.1021/jp304188z | J. Phys. Chem. B 2012, 116, 8873−8879

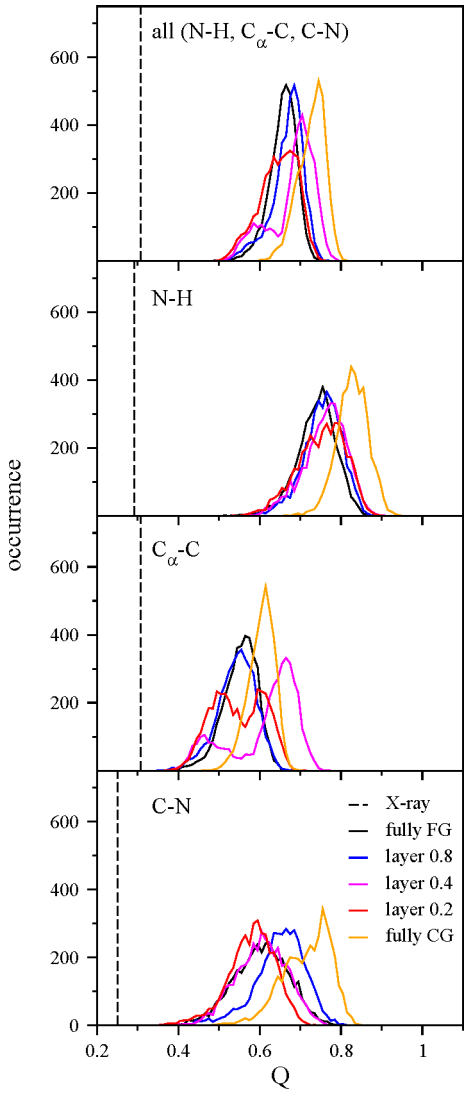


Figure 6. Q-value distributions of the calculated backbone N−H, C_α−C, and C−N RDCs of hen egg-white lysozyme (HEWL) in fully FG water (black), with a FG water layer of size 0.8 nm (blue), of size 0.4 nm (magenta), of size 0.2 nm (red), and with no layer, i.e., fully CG water (orange) when fitting the calculated residual dipolar couplings (RDC) of each configuration to the experimentally derived RDCs of HEWL.[31] The Q value of the fits from the X-ray crystal structure[15] is shown by a vertical black, dashed line.

mixed-grained simulations was chosen to match the number of FG water molecules in the fully atomistic simulation to allow a fair comparison (Table 3). Thus, the computational boxes had approximately the same volume. In addition, a fully atomistic simulation with the larger cutoff radii used in the mixed-grained simulations was performed for comparison.

In Table 3, the speed-up factors are given for the different layer thicknesses using either the cutoff radii 0.8/1.4 nm or 1.4/2.0 nm with respect to the fully atomistic simulation using the cutoff radii 0.8/1.4 nm. If the cutoff radii 1.4/2.0 nm were used in the mixed-grained simulations, the speed-up was reduced from a factor seven for the protein in pure CG water to a factor three in the simulation with a FG water layer of size 0.8 nm. However, if the cutoff radii 0.8/1.4 nm were also used in the mixed-grained simulations, the simulation with a FG water layer of 0.8 nm is a factor 6 faster than the fully atomistic simulation. If the larger cutoff radii 1.4/2.0 nm were used in both the atomistic and mixed-grained simulations, the computational speed-up is even a factor ten using the FG layer thickness 0.8 nm.

Table 3. Number of Fine-Grained (FG) Water Molecules and Coarse-Grained (CG) Water Beads Used in the Benchmarking Simulations (5000 Steps) of Hen Egg-White Lysozyme (HEWL) and the Corresponding Speed-up Factors for Two Combinations of the Short-Range and Intermediate-Range Cut-off Radii of the Nonbonded Interactions

layer thickness	FG/CG	cut-offs 0.8/1.4 nm	cut-offs 1.4/2.0 nm
fully FG	14332/0	1	0.3
0.8 nm	1616/2543	6	3
0.4 nm	592/2748	11	5
0.2 nm	135/2839	14	6
fully CG	0/2866	16	7

SUMMARY AND CONCLUSIONS

In a previous study, mixed-grained molecular dynamics (MD) simulations of the four proteins hen egg-white lysozyme (HEWL), major cold shock protein (CspA), G-protein (GP), and chorismate mutase (CM) in CG water were compared to fully atomistic simulations of the proteins. While the secondary and tertiary structure of the proteins was preserved in the hybrid FG/CG solute/solvent simulations, an increased occurrence of hydrogen bonds between the side chains was observed due to a lack of hydrogen-bonding partners in the solvent. Here, the effect of a fine-grained (FG) water layer around FG proteins in coarse-grained (CG) water on their structural properties was investigated. The thickness of the FG water layer around the protein needed to recover the hydrogen-bonding pattern observed in fully atomistic simulations and the subsequent reduction of the computational efficiency was assessed by considering three thicknesses of 0.2, 0.4, and 0.8 nm. In addition, the simulations of HEWL were compared to experimental NOE distance bounds, ^3J-coupling constants, N−H order parameters, and backbone residual dipolar couplings.

The structural properties such as atom-positional root-mean-square deviation (rmsd) of the backbone, root-mean-square fluctuations (RMSF) of the backbone and the side chains, radius of gyration, and the secondary structure elements, for which properties the simulations in pure CG water already showed agreement with the fully atomistic simulations, were not affected by the introduction of a FG water layer. The partial unfolding of the α-helix in the simulation of GP in CG water was not observed when a FG water layer was present. As expected the introduction of a FG water layer led to a decrease of side chain−side chain hydrogen bonds and subsequently to a more negative internal Lennard-Jones energy V_{LJ} and a less negative electrostatic energy V_{CRF} of the protein. A layer thickness of 0.8 nm was found to be sufficient to recover the properties of the fully atomistic simulation. Figure 7 illustrates that the CG waters do not replace the FG waters around the protein. This comes at the cost of losing a factor 2−3 in computational efficiency. Yet, the simulation of a protein with a 0.8 nm layer of FG water solvated in CG water using a nonbonded interaction cutoff radius of 2.0 nm is still three times faster than the simulation of the protein solvated in FG water using a 1.4 nm cutoff radius. Using a 2.0 nm cutoff radius in the simulation of the fully FG solvated protein makes it 10 times slower than the one using the 0.8 nm FG water layer.

dx.doi.org/10.1021/jp304188z | J. Phys. Chem. B 2012, 116, 8873−8879

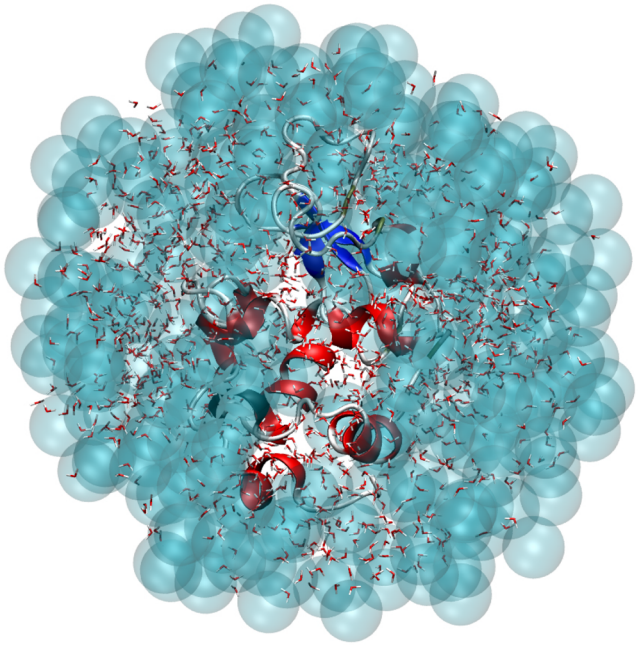


Figure 7. Configuration of HEWL solvated in a 0.8 nm layer of 1616 FG water molecules and 2636 CG water beads after 20 ns of MD simulation. The CG water beads within 0.35 nm distance of the FG water are shown as transparent, cyan spheres, the FG water in licorice mode with the oxygen atoms in red and the hydrogen atoms in white. The secondary structure elements of the proteins are colored as follows: α-helices are shown in red, 3_{10}-helices in black, and β-strands in blue.

Thus, the use of the CG solvent allows 1 order of magnitude faster simulation of proteins in aqueous solution without disturbing the structural properties of the protein significantly.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional figures depicting RGYR, RMSF, secondary structure elements, distribution of 1630 experimental NOE bound violation, and $^{1}H-^{15}N$ order parameter data. Tables of additional expreimental data and additional references. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: wfvgn@igc.phys.chem.ethz.ch.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was financially supported by the National Center of Competence in Research (NCCR) in Structural Biology and by Grant No. 200020-137827 of the Swiss National Science Foundation, and by Grant No. 228076 of the European Research Council, which is gratefully acknowledged.

■ REFERENCES

(1) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(2) Neri, M.; Anselmi, C.; Cascella, M.; Maritan, A.; Carloni, P. *Phys. Rev. Lett.* **2005**, *95*, 218102.
(3) Shi, Q.; Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2006**, *110*, 15045−15048.
(4) Michel, J.; Orsi, M.; Essex, J. W. *J. Phys. Chem. B* **2008**, *112*, 657−660.
(5) Masella, M.; Borgis, D.; Cuniasse, P. *J. Comput. Chem.* **2008**, *29*, 1707−1724.
(6) Rzepiela, A. J.; Louhivuori, M.; Peter, C.; Marrink, S. J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10437−10448.
(7) Masella, M.; Borgis, D.; Cuniasse, P. *J. Comput. Chem.* **2011**, *32*, 2664−2678.
(8) Riniker, S.; Allison, J. R.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2012**; DOI: 10.1039/C2CP40934H.
(9) Riniker, S.; van Gunsteren, W. F. *J. Chem. Phys.* **2012**, accepted.
(10) Riniker, S.; Eichenberger, A. P.; van Gunsteren, W. F. *Eur. Biophys. J.* **2012**; DOI: 10.1007/s00249-012-0837-1.
(11) Riniker, S.; van Gunsteren, W. F. *J. Chem. Phys.* **2011**, *134*, 084110.
(12) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. *Eur. Biophys. J.* **2011**, *40*, 843−856.
(13) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. in *Intermolecular Forces*; Pullmann, B., Ed.; Reidel: Dordrecht, 1981; pp 331−342.
(14) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zürich: Zürich, Groningen, 1996.
(15) Artymiuk, P. J.; Blake, C. C. F.; Rice, D. W.; Wilson, K. S. *Acta Crystallogr., Sec. B* **1982**, 778−783.
(16) Schindelin, H.; Jiang, W.; Inouye, M.; Heinemann, U. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5119−5123.
(17) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721−4729.
(18) Okvist, M.; Dey, R.; Sasso, S.; Grahn, E.; Kast, P.; Krengel, U. *J. Mol. Biol.* **2006**, *357*, 1483−1499.
(19) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; van Gunsteren, W. F. *Comput. Phys. Commun.* **2012**, *183*, 890−903.
(20) Kunz, A.-P. E.; Allison, J. R.; Geerke, D. P.; Horta, B. A. C.; Hünenberger, P. H.; Riniker, S.; Schmid, N.; van Gunsteren, W. F. *J. Comput. Chem.* **2012**, *33*, 340−353.
(21) The GROMOS package of programs can be obtained from www.gromos.net.
(22) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(23) Hockney, R. W. *Methods Comput. Phys.* **1970**, *9*, 136−211.
(24) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(25) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(26) Lide, D. R. *Handbook of Chemistry and Physics*, 88th ed.; CRC Press/Taylor and Francis: Boca Raton, FL, 2007−2008.
(27) Eichenberger, A. P.; Allison, J. R.; Dolenc, J.; Geerke, D. P.; Horta, B. A .C.; Meier, K.; Oostenbrink, C.; Schmid, N.; Steiner, D.; Wang, D.; et al. *J. Chem. Theory. Comp* **2011**, *7*, 3379−3390.
(28) Smith, L.; Sutcliffe, M.; Redfield, C.; Dobson, C. *Biochemistry* **1991**, *30*, 986−996.
(29) Allison, J.; van Gunsteren, W. F. *ChemPhysChem* **2009**, *10*, 3213−3228.
(30) Steiner, D.; Allison, J. R.; van Gunsteren, W. F. *J. Biomol. NMR* **2012**; DOI: 10.1007/s10858-012-9634-5.
(31) Higman, V. A. *The use of bicelles and other ordered media to study protein structure and dynamics*; Ph.D. thesis; University of Oxford: Oxford, U.K., 2004.