Supplemental material: Adaptive resolution simulation of an atomistic protein in MARTINI water

Julija Zavadlav,[1] Manuel Nuno Melo,[2] Siewert J. Marrink,[2, a)] and Matej Praprotnik[1, 3, b)]

[1)]*Laboratory for Molecular Modeling, National Institute of Chemistry, Hajdrihova 19, SI-1001 Ljubljana, Slovenia*

[2)]*Groningen Biomolecular Sciences and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands*

[3)]*Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia*

In the following we present the remainder of results of our adaptive resolution simulations. The simulation protocol is given in the original paper[1]. Figs. 1 and 2 depict the normalized density profiles of water bundles and oxygen atoms, respectively, around the protein's center of mass. The water density profiles are the same for all simulations. This is important because the stability of the protein structure depends critically on the density of water molecules around the protein.
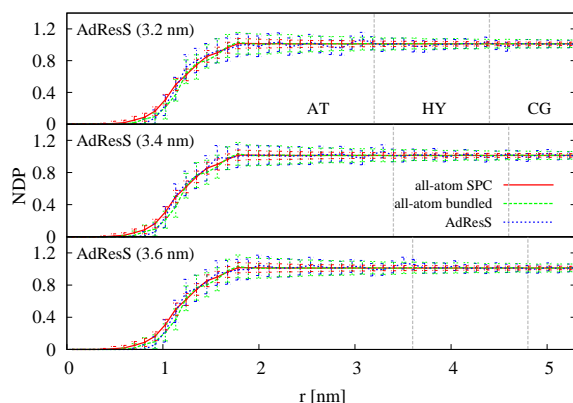

FIG. 1. Solvent density around the center of mass of protein G for AdResS simulations with atomistic region radius sizes of 3.2nm (top), 3.4nm (middle) and 3.6nm (bottom) using the bundled water model 2[2] (The results for the model 1[2] are presented in the original paper[1]). The plots include error bars[3]. The results are compared to the fully atomistic bundled and SPC solvations. The vertical lines denote the boundaries between resolution domains, i.e., the atomistic (AT), hybrid (HY), and the coarse-grained (CG) domains.

To achieve a flat density profile throughout the hybrid region, the adaptive resolution simulations (AdResS) require the
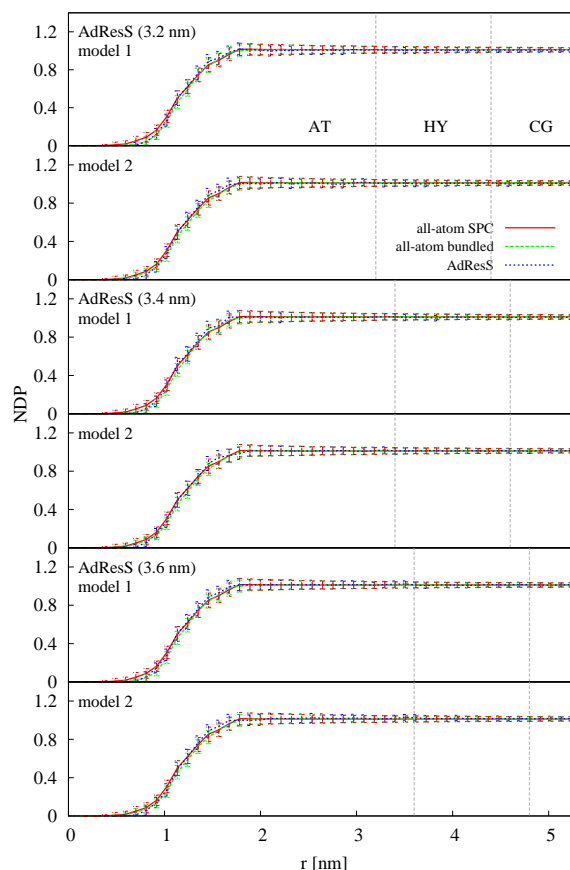
[a)]s.j.marrink@rug.nl
[b)]praprot@cmm.ki.si

FIG. 2. Oxygen density profiles around the center of mass of protein G. The presentation is the same as in Fig. 1 except that the results for the both models of bundled water are presented.

use of thermodynamic (TD) force[4,5]. The latter compensates for the difference in the chemical potential at different levels of resolution. We calculate the TD force with an iterative procedure as described in the literature on a water system without the presence of protein. The TD force that acts on bundle's center of mass in the hybrid region is shown for both models in Fig. 3.
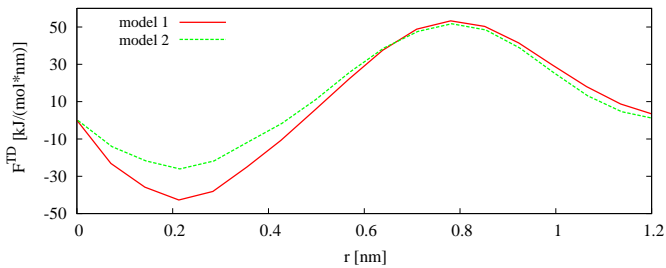
FIG. 3. The thermodynamics force that acts on molecule's center of mass in the hybrid region.
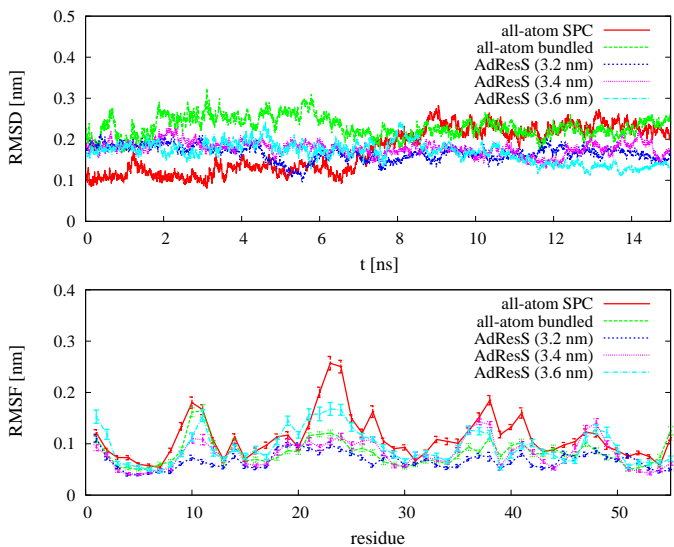

FIG. 4. RMSD (top) and RMSF with error bars (bottom) of the backbone atoms with respect to the crystal structure as a function of time. We compare the results obtained from the fully atomistic simulation using SPC and bundled water (model 2) to AdResS simulations with three atomistic region sizes: spheres of radii 3.2, 3.4 and 3.6nm, respectively. The corresponding results using the model 1 of the bundled water are presented in the original paper[1].

The protein's structural properties are verified with the root-mean-square deviation (RMSD) and the root-mean-square fluctuations (RMSF) of the backbone atoms with respect to the crystal structure (see Fig. 4). Using the model 2 of the bundled water the average RMSD values (in nm) are 0.23 ± 0.02, 0.17 ± 0.02, 0.18 ± 0.02, 0.17 ± 0.03 for the all-atom bundled and AdResS bundled (with the three atomistic region sizes) solvation, respectively. These values confirm that the protein structure is stable during our simulations.

We have also computed the protein's radius of gyration (R_g) shown in Fig. 5. The results from the adaptive reso-
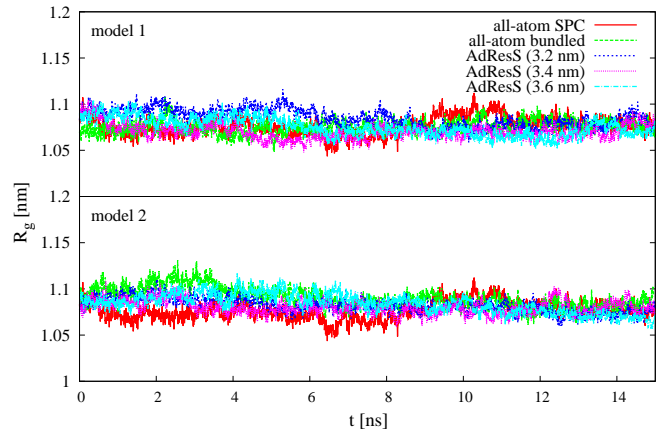

FIG. 5. The radius of gyration as a function of time. The presentation is the same as in Fig. 4 except that the results for the both models of bundled water are presented.

lution simulations closely match the atomistic counterparts. Even for the adaptive resolution simulation with the smallest atomistic region the protein remains compact during the total simulation run. With atomistic domain size almost three times larger than R_g we have ensured that the whole protein is always located within the atomistic region.
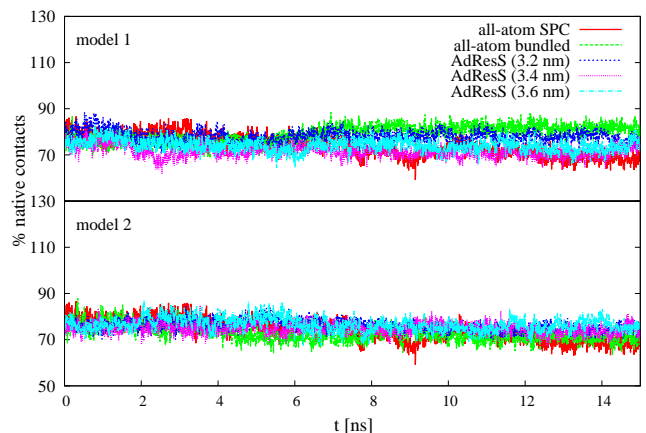

FIG. 6. Percentage of native contacts as a function of time. The presentation is the same as in Fig. 5.

Furthermore, we have determined the stability of native contacts, i.e., contacts between the side chains of two amino acids that are more than five residues apart in the protein but are spatially closer than 0.75nm, by the MaxCluster tool[6]. Figure 6 shows time evolution of the percentage of native contacts, i.e., the ratio between numbers of native contacts in the current and the crystal structure. The percentage remains con-

stant during the course of all simulations, further confirming the stable structure of the protein.

The solvent accessible surface area and the secondary structure were obtained using the DSSP program[7] and are shown on Figs. 7 and 8.
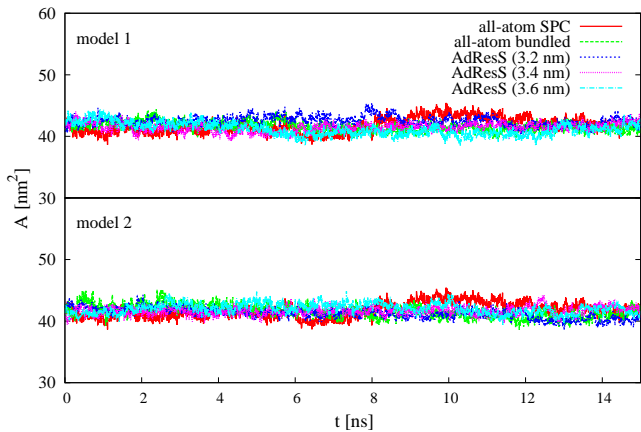

FIG. 7. The solvent accessible surface area as a function of time. The presentation is the same as in Fig. 5.
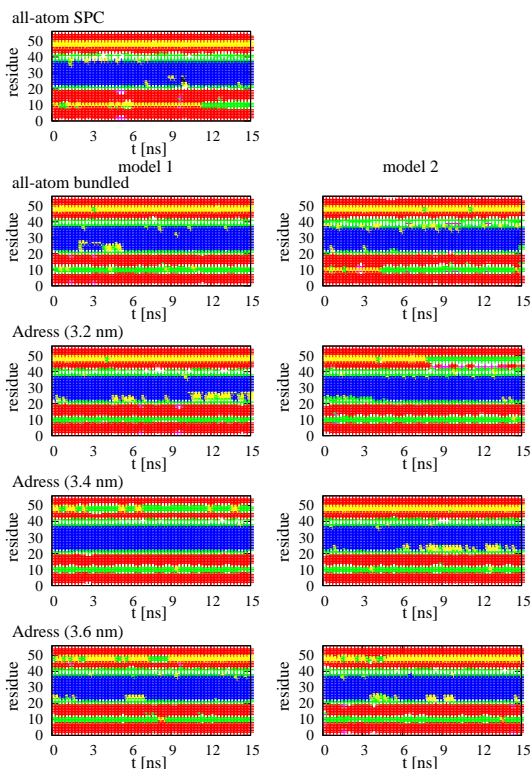

FIG. 8. The secondary structure elements of protein G for fully atomistic and AdResS simulations. The a-helices are shown in blue, the 3_{10}-helices in black, the b-strands in red, the hydrogen bonded turns in yellow and the bends in green.

The solvent accessible area results are in accordance with the relation[8] between solvent accessible area A and the protein's molar mass M_r, i.e. $A = \alpha M_r^{0.73}$, where $\alpha = 0.063 \mathrm{nm}^2/\mathrm{mol}$. For protein G the relation yields approximately $37 \mathrm{nm}^2$. We observe minor differences in the time evolution of secondary structure elements between different simulations. However, these differences appear to be uncorrelated with a specific simulation approach.
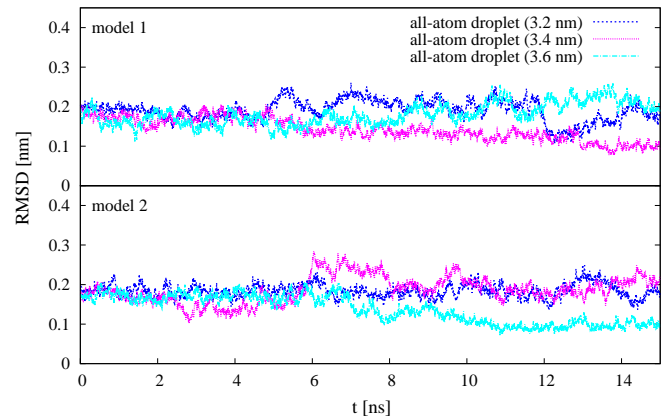

FIG. 9. RMSD of the backbone atoms with respect to the crystal structure as a function of time. We present the results obtained from the fully atomistic simulation of protein in bundled water droplet for three sizes: spheres of radii 3.2, 3.4 and 3.6nm. The average RMSD values (in nm) are 0.19 ± 0.03, 0.14 ± 0.03, 0.18 ± 0.03 (model 1 of the bundled water) and 0.18 ± 0.02, 0.19 ± 0.03, 0.14 ± 0.04 (model 2) for the three droplet sizes, respectively.
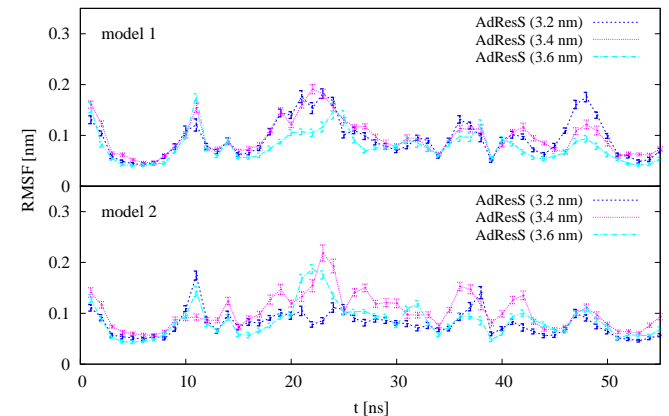

FIG. 10. RMSF with error bars of the backbone atoms as a function of the residue sequence number. The presentation is the same as in Fig. 9.

Finally, we have carried out additional simulations of protein enclosed in atomistic water droplet in vacuum. The results are shown in Figs. 9, 10 and 11. Simulations of pro-
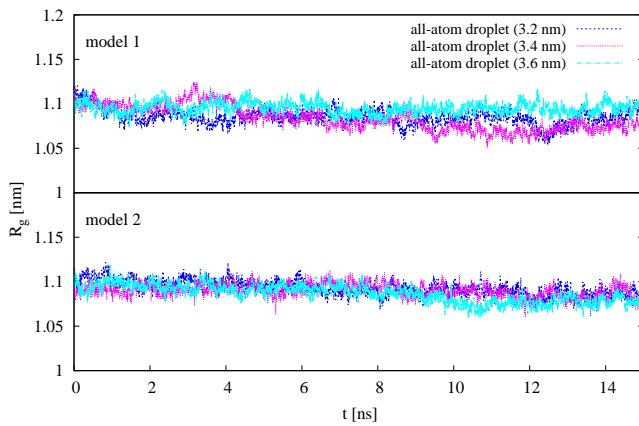
FIG. 11. The radius of gyration as a function of time. The presentation is the same as in Figs. 9 and 10.

tein within water droplet in vacuum give similar results as the bulk AdResS and fully atomistic simulations for all considered structural properties of the protein. Such results indicate that protein's effect on the water structure does not exceed the chosen sizes of atomistic regions and water droplets. The bulk is modeled in the AdResS simulations by CG water that is not able to form hydrogen bonds. Therefore, the stability of the protein does not seem to be affected strongly by the bulk hydrogen bond network. Note, however, that even in the smallest water droplet there are still several hydration shells around the protein.

[1] J. Zavadlav, M. N. Melo, S. J. Marrink, and M. Praprotnik, Original paper.

[2] M. Fuhrmans, B. P. Sanders, S. J. Marrink, and A. H. de Vries, Theor. Chem. Acc. **125**, 335 (2010).

[3] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987).

[4] S. Poblete, M. Praprotnik, K. Kremer, and L. Delle Site, J. Chem. Phys **132**, 114101 (2010).

[5] S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, and K. Kremer, Phys. Rev. Lett. **108**, 170602 (2012).

[6] A. Herbert, MaxCluster: A tool for Protein Structure Comparison and Clustering, http://www.sbg.bio.ic.ac.uk/ maxcluster/index.html.

[7] R. P. Joosten, T. A. H. te Beek, E. Krieger, M. L. Hekkelman, R. W. W. Hooft, R. Schneider, C. Sander, and G. Vriend, NAR **39**, 411 (2010).

[8] V. Z. Spassov, A. D. Karshikoff, and R. Ladenstein, Protein Science **4**, 1516 (1995).