JCTC Journal of Chemical Theory and Computation

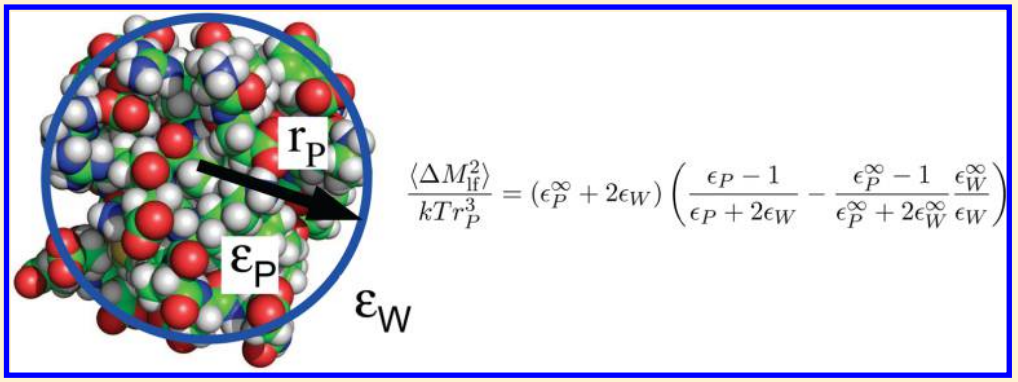
Article

pubs.acs.org/JCTC

What Is the Dielectric Constant of a Protein When Its Backbone Is Fixed?

Thomas Simonson*

Laboratoire de Biochimie (CNRS UMR7654), Department of Biology, Ecole Polytechnique, 91128 Palaiseau, France



ABSTRACT: Monte Carlo (MC) simulations with a fixed protein backbone but mobile sidechains are common for acid/base constants and protein design. To characterize the fluctuations in these models, estimating the Fröhlich−Kirkwood dielectric constant can give physical insight and allow comparison both with models that are more rigorous (fully flexible) and ones that are simpler (Poisson−Boltmann without any explicit protein flexibility). MC simulations of two small proteins yield protein dielectric constants of 12 and 14, about 70% of the result from MD (16 and 22). Thus, the consistency between the fully explicit MD and partly explicit MC is only fair.

1. INTRODUCTION

What is the dielectric constant of a protein when its backbone is fixed, and should we care? Simulations with a fixed protein backbone and flexible side chains were introduced in the 1990s and have become common. In particular, acid/base calculations are often performed with the protein backbone fixed but with the side chains free to explore different rotameric states.[1−9] This idea is even more prevalent in computational protein design.[10−13] With these models, although the backbone is held fixed during the simulations, its degrees of freedom and their fluctuations are not neglected. Rather, they are represented through a dielectric continuum model and absorbed into a protein dielectric constant.[14] Such models represent an intermediate level of theory, between fully explicit ones, where all the protein's degrees of freedom are explicitly modeled, and fully implicit ones.[15]

Since the continuum model can be applied to different sets of degrees of freedom, it is worth defining some vocabulary. A simulation model where the protein's degrees of freedom (and their fluctuations) are explicitly modeled (through MD or MC), we refer to as "explicit" (even though the solvent may be implicit). Models where some of the protein degrees of freedom are implicit and others are explicit we refer to as "hybrid". The first hybrid model was introduced by Fröhlich (for liquids):[14] the nuclear degrees of freedom were treated explicitly and embedded in a dielectric continuum whose dielectric constant ε^∞ modeled the electronic degrees of freedom. The exponent ∞ indicates the ability of the electrons to respond ∼instantly to structural rearrangements of the nuclei. Fröhlich then went on to compute the total dielectric constant of the system, ε ($>\varepsilon^\infty$), which arises from both the nuclear *and* the electronic degrees of freedom. Following this example, we define two dielectric constants for the protein: ε_P is the total protein dielectric constant; ε_P^∞ is the dielectric constant used during the hybrid simulations to model any protein degrees of freedom that are not explicit. Thus, in simulations where the backbone is held fixed but the side chains are mobile,[1−11] ε_P^∞ should account for both the protein's electronic polarizability and its backbone motions; another, larger dielectric constant is used for the aqueous solvent.

We can now reformulate our starting question differently: if we do a simulation with a fixed backbone using a reasonable ε_P^∞ value, measuring the protein fluctuations, how well do they compare with an ordinary MD simulation, where the protein is fully flexible? We obviously cannot answer this by comparing the atomic fluctuations directly, since different atom sets are free to move in each model. We can, however, ask how well the *total* protein dielectric constant ε_P from the hybrid simulation matches the MD result. Not only is ε_P an essential thermodynamic parameter, directly related to flexibility but it can also be defined and computed in a consistent way from both explicit and hybrid models, following the method of Fröhlich.[14] Thus, it provides a good observable to test the

Received: May 15, 2013
Published: September 3, 2013

   dx.doi.org/10.1021/ct400398e | *J. Chem. Theory Comput.* 2013, 9, 4603−4608

consistency between the explicit and fixed backbone approaches.

The calculation of protein dielectric properties from atomistic simulations has a long history.[16−36] Most of these studies have used macroscopic continuum theory to interpret atomistic simulations, by comparing the dipolar fluctuations to the theory, often framed as the Fröhlich−Kirkwood theory of dielectrics.[14,16,37−39] Such comparisons are useful, even though the continuum theory is only approximate for an object as small as a protein.[40,41] In fact, surprisingly good qualitative agreement is seen, with the interior of several proteins displaying polar fluctuations that are similar to a simple dielectric medium with a plausible dielectric constant, comparable to that of dry protein powders or acetamide crystals;[42,43] see Simonson[16] for a review. More sophisticated analyses have also been done, comparing reorganization energies or electrostatic potentials to continuum theory.[26,44−48] These comparisons give physical intuition and lend support to the simple continuum models.

Here, we consider two small proteins: the SH3 domain of Crk (57 amino acids) and Staphylococcal nuclease (SNase; 136 amino acids). For each one, we have performed 10−20 ns MD simulations with the protein fully flexible, in explicit solvent. We have measured the protein's dipolar fluctuations and analyzed them through the Fröhlich−Kirkwood theory to obtain an estimate of the total dielectric constant, ε_P(MD). The values obtained are similar to other small proteins.[23,25,31] We compare the results to MC simulations with a fixed backbone, flexible side chains, and a generalized Born implicit solvent. The MC simulations use a dielectric constant ε_P^∞ of either 4 or 8. These values gave good results for pK_a calculations in several proteins, including SNase.[7,9,49] The lower value ($\varepsilon_P^\infty = 4$) is comparable to the dielectric constant of dry protein powders.[42,43] We show that the total dielectric constants ε_P from MD and MC are in rough but not close agreement, with the hybrid model underestimating the Kirkwood G-factor and dielectric constant by about 30%. The MC results with $\varepsilon_P^\infty = 4$ and 8 are similar. Thus, for the overall protein fluctuations and dipole variance, consistency between the explicit and hybrid levels of theory is only fair.

2. THEORETICAL METHOD

By combining linear response theory and dielectric continuum theory, Kirkwood and Fröhlich showed how to obtain relations between the dipolar fluctuations of a simple medium and the dielectric constant(s) of it and its surroundings. Their work was expanded by later authors.[37−39] Here, we are interested in the simplest case: the inner region (medium "P") is a sphere, and it and its surroundings (medium "W") are homogeneous and isotropic, with total and high frequency dielectric constants ε_P, ε_P^∞, ε_W, and ε_W^∞. We then have[14]

$$\frac{f(\varepsilon_P, \varepsilon_W)(\varepsilon_P - 1) - f(\varepsilon_P^\infty, \varepsilon_W^\infty)(\varepsilon_P^\infty - 1)}{f(\varepsilon_P^\infty, \varepsilon_W)} = \frac{\langle \Delta M_{lf}^2 \rangle}{kTr_P^3} \stackrel{\text{def}}{=} \langle g \rangle \stackrel{\text{def}}{=} G \tag{1}$$

Here, ΔM_{lf} is the deviation of the dipole moment of medium P from its mean, excluding the contributions of the highest-frequency degrees of freedom (usually, the electronic polarizability); the subscript stands for "low frequency". The brackets represent an ensemble average (in the absence of any applied external field), r_P is the radius of medium P, and $f(\varepsilon_P,\varepsilon_W)$ has the form

$$f(\varepsilon_P, \varepsilon_W) = \frac{3\varepsilon_W}{\varepsilon_P + 2\varepsilon_W} \tag{2}$$

Equation 1 can be seen as a fluctuation−dissipation relation.[50] Its derivation involves treating the high-frequency degrees of freedom as a dielectric continuum, in which the lower-frequency degrees of freedom are embedded (following Frohlich[14]). Notice that it is valid even if medium P has a nonzero net charge.[51] To interpret the Monte Carlo simulations, we will view not only the electronic polarizability of the protein but also its backbone degrees of freedom as "high-frequency" degrees of freedom, to be absorbed into ε_P^∞. We repeat that viewing the backbone motions as part of ε_P^∞ is not the same as neglecting them.

The dimensionless quantity G on the right of eq 1 will be referred to as the Kirkwood G-factor;[14] its instantaneous value is denoted g. The probability distribution of g can be derived from continuum electrostatics, as pointed out long ago by Kusalik:[14,52−54]

$$P(g) \propto \sqrt{g}\, \exp\left(-\frac{3g}{2\langle g \rangle}\right) \tag{3}$$

Deviations from this distribution will arise if a system does not behave as a macroscopic continuum; if it is too small, for example. Deviations can also arise from insufficient sampling, especially of the largest, rarest, fluctations. Sampling the entire probability distribution of g is more demanding than simply estimating its expectation value G. Comparing the simulations to eq 3 will thus provide a test of the simulation quality and the use of continuum theory.

3. COMPUTATIONAL DETAILS

Monte Carlo Simulations. We started from the crystal structures, respectively, of the Crk SH3 domain (PDB code 1CKA; 57 amino acids, discarding a peptide ligand), and Staphylococcal nuclease (SNase; PDB code 1STN; 136 amino acids). Standard side chain protonation states were used and supported by pK_a predictions with the PROPKA program.[55,56] The system was modeled with the Amber ff99SB all-atom force field and an implicit solvent model that combines the GB/HCT generalized Born variant[57,58] with a solvent accessible surface area contribution.[58,59] The atomic surface energy coefficients (in kcal/mol/Å^2) were as follows: polar = −0.008; ionic = −0.009; aromatic = −0.012; alkane = −0.005; hydrogens = 0. The protein dielectric constant during the MC simulations is denoted ε_P^{MC}; it was set to 4 or 8. This value can be viewed as the "high-frequency" protein dielectric constant, ε_P^∞ in eq 1. For the side chain degrees of freedom, we used a rotamer library from 1995 by Tuffery et al.,[60] which has given good results for other applications,[9,49,61,62] including constant-pH Monte Carlo. To make the simulations efficient, we borrow a procedure from computational protein design, where all the pairwise interactions between protein side chains, or a side chain and the backbone, are computed ahead of time and stored in an energy matrix, allowing for all possible rotamer combinations for each pair.[3,49,63] For each matrix element (pairwise energy), N_{min} = 15 steps of conjugate gradient energy minimization are performed, to allow rotamer pairs to adjust slightly, with only the side chain pair allowed to move and only the intrapair and pair-backbone interactions included. With this procedure, we do not use a reduced or reparameterized van der Waals potential, as in some related methods.[3,64] There is no on-the-fly

dx.doi.org/10.1021/ct400398e | J. Chem. Theory Comput. 2013, 9, 4603−4608

minimization during the MC simulation. MC moves are accepted or rejected with the standard Metropolis algorithm.[65]

The energy matrix precalculation raises a question, since the GB solvent model is not decomposable into a sum of pairwise contributions.[57] However, it can be made pairwise-additive by applying a "Native Environment" approximation to each residue pair, where the pair's environment is assumed to be in its experimental, crystal conformation. More precisely, for each pair of residues *i,j*, the GB interaction is computed using atomic solvation radii that were estimated ahead of time for each residue, assuming the rest of the system is in its crystal conformation.[49] This procedure was shown to give good accuracy.[3,9,49] The Monte Carlo move set included rotamer changes at single sites or pairs of sites where the unsigned interaction energy was above a 3 kcal/mol threshold. Twenty million steps were run at a high temperature, kT = 0.8 kcal/mol, followed by 100 or 200 million steps at room temperature, kT = 0.6 kcal/mol, with the Metropolis MC algorithm.[66] Simulations were done with our in-house Proteus software package, which combines a modified version of the XPLOR program with in-house Monte Carlo code.[49,67]

Molecular Dynamics Simulations. Each protein was also simulated by molecular dynamics, immersed in a cubic box of explicit, TIP3P water,[68] using Particle Mesh Ewald, tinfoil boundary conditions,[69] and the CHARMM program.[70] Simulations were done with the same, Amber ff99SB protein force field used for the MC and run for 10 ns (Crk) or 20 ns (SNase). For comparison, we also did MD simulations with the Charmm27 force field.[71]

We also ran simulations for isolated amino acid analogues, with their backbones truncated before the leading NH and after the trailing CO, with the side chains Asp, Glu, Arg, or Lys. Each one was simulated at 300 K in GB solvent with its backbone held fixed, by MD and MC. The MD runs lasted 100 ns and used Langevin dynamics for all non-hydrogen atoms with a friction coefficiant of 60 ps^{-1}. The MC runs lasted 20 million steps. The solute dielectric constant was =4 in both cases.

4. RESULTS

G-factors and Dielectric Constants. The G-factors and dielectric constants estimated from MD and MC for Crk and SNase are given in Table 1. For Crk, the MD simulations lead to an overall G-factor of 18.3 (with protein radii estimated as $(5/3)^{1/2}$ times the radius of gyration[23]). This leads to a total dielectric constant of 21.6, similar to other small proteins.[16,23,25,31] Results with the Charmm27 force field are very similar (G = 19.3, ε_P = 22.9). Computing G from the backbone groups alone leads to a much lower dielectric constant of 6.8, not too different from dry protein powders.[42,43] Computing without the ionized side chains gives an even lower dielectric constant of 2.5. Notice that the G-factor is not a simple sum over protein groups and charges but depends on them quadratically. The result with the backbone alone can be larger than the result with all nonionized groups, as here, due to cross terms that couple the nonionized side chains and the backbone. The observation that the ionized side chains dominate the total dielectric constant is consistent with earlier work on several other proteins.[16,23,25,31]

The MD simulations do not include any explicit electronic polarizability, but polarizability is represented implicitly, through the atomic partial charges. Setting ε_P^∞ = 2 must then represent an overestimate and an upper limit on the contribution of the electronic polarizability. With $\varepsilon_P^\infty = \varepsilon_W^\infty$ = 2, we obtain slightly different results: ε_P = 7.5 for the backbone and 22.3 overall.

Table 1. Computed G-factors and Dielectric Constants

	groups	radius r_P	G	ε_P	ε_P^∞	ε_W^∞
		Crk				
MD	backbone	12.9	5.64	7.5	2	2
MD	backbone	12.9	5.64	6.8	1	1
MD	nonionized	12.9	1.52	2.5	1	1
MD	all	13.4	18.3	**21.6**[a]	1	1
MC(ε_P^{MC} = 8)	all	13.4	9.7	**13.6**	8	2
MC(ε_P^{MC} = 4)	all	13.4	10.3	**13.6**	4	2
MC(ε_P^{MC} = 4)	nonionized	12.9	0.87	3.4	4	2
		SNase				
MD	backbone	18.4	0.74	2.4	2	2
MD	backbone	18.4	0.74	1.7	1	1
MD	nonionized	18.4	1.31	2.3	1	1
MD	all	18.9	13.8	**16.0**	1	1
MC(ε_P^{MC} = 4)	all	18.9	8.76	**11.8**	4	2
MC(ε_P^{MC} = 4)	nonionized	18.4	0.62	3.1	4	2

[a]Values in bold correspond to the protein dielectric constant with the most plausible ε_i^∞ choices.

The MC simulations using $\varepsilon_P^{MC} = \varepsilon_P^\infty$ = 4 lead to a much smaller G value of 10.3 for Crk. This is mainly because the protein backbone degrees of freedom do not contribute (either directly or through cross terms involving side chains). Rather, they are represented as an underlying dielectric continuum, which also includes the protein's electronic degrees of freedom. The Fröhlich–Kirkwood eq 1 then converts the G factor into a total dielectric constant of 13.6. This is much smaller (63%) than the MD value of 21.6. Notice that the difference between the MD and MC values is larger than the "backbone" dielectric constant of 6.8 estimated from the MD. Similarly, the difference between the MD and MC G-factors is much larger than the backbone G-factor from the MD (5.6). To obtain the MD dielectric constant, in combination with ε_P^∞ = 4, the MC G-factor would have to be 17.1; the observed result is 60% of this.

Since the Crk MD simulations yield an ε_P of 6.8 when the G-factor is computed with the backbone charges only, we repeated the MC simulations with a larger ε_P^{MC} of 8. The total MC G-factor was 9.7 (compared to 10.3 above when ε_P^{MC} = 4); eq 1 then converts this into a total dielectric constant of ε_P = 13.6. This is the same as the result above with ε_P^{MC} = 4. Thus, increasing the value ε_P^{MC} used in the MC simulations does not improve the agreement with MD.

Results for SNase are similar. The G-factor from MD is 13.8 and the total protein dielectric constant is ε_P = 16.0. The backbone contribution alone leads to a very small G-factor of 0.7 and a dielectric constant of 1.7; nonionized groups alone lead to a G-factor of 1.3 and a dielectric constant of 2.3. As with Crk, the protein dielectric constant arises almost entirely from the ionized side chains. MD simulations with the Charmm27 force field led to a slightly larger ε_P of 18.6. Earlier 10 ns Charmm27 simulations[34] led to an even larger value of 24, with an estimated uncertainty of ±6. The Amber/Charmm difference found here is less than this uncertainty and less than the difference between the two Charmm27 runs.

The SNase MC simulations using ε_P^{MC} = 4 give a G-factor of 8.8 and an overall dielectric constant of 11.8. This is about 74% of the overall MD result. The G-factor is 71% of the value (12.4) that would lead (in combination with ε_P^∞ = 4) to the same total dielectric constant as MD. With only ionized side

chains, the dielectric constant from MC is just 3.1, about 1/4 of the total. Since the "backbone only" G-factor from MD was so small (0.7), we did not do MC simulations with a larger ϵ_P^{MC}.

Given the smaller G-factors from MC, compared to MD, we checked whether the rotamer approximation for the side chains might limit their fluctuations. We considered four amino acid analogues in isolation, with their backbones held fixed, simulated by MD and MC, using the GB solvent. The fluctuations were characterized by the variance of the side chain dipole moment, $\langle \Delta M^2 \rangle$. For Asp and Glu, the MD and MC fluctuations were very similar (Table 2). For the large Arg side chain, despite two 100 ns MD simulations, the results were noisier. The Arg MD fluctuations are somewhat larger than with MC; for Lys, they are somewhat smaller. High-temperature MC results (where all rotamers are equally populated) are given for comparison.

Table 2. Dipole Variance for Isolated Amino Acids[a]

	$\langle \Delta M^2 \rangle$			no. per protein[c]	
side chain	MD	MC	MC($T \to \infty$)[b]	SNase	Crk
Asp	3.7(4)	3.7	4.1	6	6
Glu	5.3(1)	5.1	6.6	11	8
Arg	9.9(5)	7.9	10.6	5	4
Lys	4.6(2)	6.8	9.7	22	5

[a]Variances of the side chain dipole moment over an MD or MC simulation (in eÅ^2). The Arg MD result is an average over two 100 ns runs. Uncertainties (in parenethesis, in significant digits) are the difference between two runs (Arg) or 1/2 runs (50 ns each). [b]High temperature limit: all rotamers are equally occupied. [c]The number of each residue type in SNase and Crk.

To put the isolated amino acid results on a scale more comparable to the protein G-factors, we can sum up the contributions of each side chain type, considering the number of instances in each protein, SNase or Crk. Dividing this sum by kTR^3 (where R is the protein radius), we obtain a dimensionless quantity that can be thought of as a "diagonal" G-factor: the side chains contribute independently, with the magnitude they would have separately, in solution. For SNase, we obtain 11.4 with the MD values and 13.1 with the MC values; for Crk, we obtain 17.7 and 17.9. In both cases, the MC values lead to slighly larger (diagonal) G-factors. Overall, it seems that using rotamers in the MC simulations does not significantly limit the side chain fluctuations.

Convergence. Convergence of the MD data is only fair (Figure 1), because we deliberately use rather short simulations (10 ns for Crk and 20 ns for SNase). Indeed, comparison to the MC runs makes more sense if we focus on MD simulations where the backbone does not make large excursions, in the flexible loops for example. MC with a fixed backbone is not designed to model such situations. With longer MD runs, the fluctuations would probably increase, but we expect that the overall dielectric constants would not change dramatically, since MD simulations on this time scale have been found to yield a reasonable qualitative picture of protein fluctuations. Indeed, another recent MD study of SNase (with the Charmm27 force field)[34] used comparable, 10 ns simulations and gave a somewhat larger value of 24, with a statistical error of 6, so that our estimate and the earlier one are consistent. Convergence of the MC runs is good; for example, when we replicate the SNase MC run, the dielectric constant changes by just 3%.

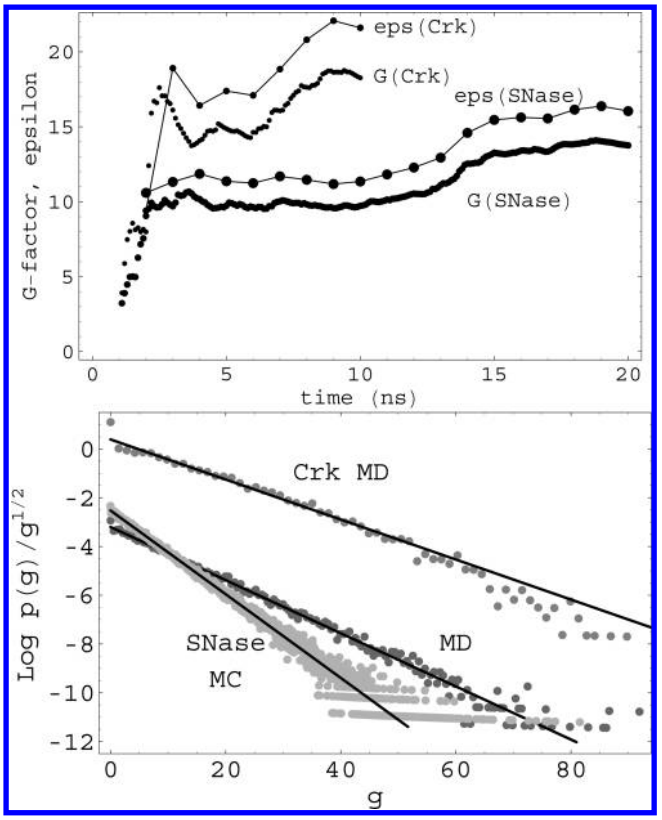


Figure 1. Top: Time convergence of the G-factor and dielectric constant from MD. Bottom: Probability distribution of the instantaneous G-factor. MD results and SNase MC results (dots) compared to continuum theory (lines). The Crk values are offset vertically by +4 for clarity.

The instantaneous value g of the G-factor from the MD and MC simulations provides another indication of convergence. It follows a probability distribution that is in good agreement with the form predicted by continuum electrostatics,[14,52−54] for both Crk (MC not shown) and SNase (Figure 1). Although the upper tails are noisy, as expected, the simulated log distributions follow the expected straight lines. Notice that the lines in the Figure are not fitted to the distributions but are computed directly from eq 3 and have the slopes $-3/2G$. Thus, despite the limited simulation lengths, the fluctuations of g are rather well-described.

5. DISCUSSION

MC simulations with a fixed backbone but mobile sidechains represent an intermediate level of theory, between models where all the protein flexibility is treated explicitly and models where all of it is treated implicitly, through a dielectric constant.[15] By comparing the total protein dielectric constant from the hybrid and fully explicit models, we can check that the intermediate model does not alter the overall dynamics too drastically, and that by combining the implicit degrees of freedom (backbone, electronic polarizability) and the explicit ones (side chain rotamers), we get back an overall picture that is at least roughly comparable to the fully explicit model. More generally, the dielectric constant is an essential thermodynamic parameter and is a good criterion to measure the consistency between the two classes of models. The theory of dielectrics, with the idea of a hybrid explicit/implicit model first proposed by Fröhlich, allows us to combine the implicit and explicit

degrees of freedom consistently and make this comparison in a meaningful way. The comparison is expected to be useful, even though dielectric theory is not rigorous for an object as small as a protein.

For the two proteins considered here, using 10−20 ns MD simulations, the hybrid model is only roughly consistent with the explicit model, giving total dielectric constants that are about 70% of the MD result: 13.6 and 11.8 from MC vs 21.6 and 16.0 from MD. We noted that the difference between the results with a fixed backbone (MC) and a flexible one (MD) is actually greater than the direct contribution of the protein backbone itself, as measured by its G-factor and the corresponding dielectric constant (6.8 and 1.7 for the two proteins). Evidently, although the direct contribution of the protein backbone to the G-factor is not very large, the backbone fluctuations lead, indirectly, to a large enhancement of the overall G-factor, probably for two reasons. First, the moving backbone carries the side chains along with it, directly increasing their fluctuations; second, the backbone motion helps activate side chain librations and rotamer transitions that would not otherwise occur.

Hybrid models are commonly used for computational protein design and for acid/base properties, where they represent a distinct improvement over the fully implicit, PB models introduced earlier.[72−74] To compare the intermediate and the fully implicit models, the total protein dielectric constant is an essential parameter. Here, we saw that when the MC is done with a protein dielectric of $\varepsilon_p^{MC} = 4$, the hybrid model leads to an overall protein dielectric constant ε_p of about 12−14. Thus, the total dielectric ε_p is much larger than the underlying ε_p^{∞}. Notice that the range 12−14 is larger than the dielectric value that was needed to interpret dielectric relaxation in the active site of aspartyl-tRNA synthetase[46] and much larger than that needed for the active site of chymotrypsin (around 2).[45]

Overall, the hybrid, MC simulations, performed with a GB model and a protein dielectric constant of 4 or 8, appear to reproduce most of the protein fluctuations, as measured by the overall G-factor. They lead to a computed dielectric constant that is 30% lower than the fully flexible MD model but is still in a range commonly used in simpler, fully implicit PB models.[15,73] To go further, it would be interesting to include some explicit backbone flexibility in the hybrid model by performing the MC simulations using a small ensemble of backbone models, but determining the statistical weights of the different backbones would not be straightforward.[75] Such extensions are left for future work.

AUTHOR INFORMATION

Corresponding Author

*E-mail: thomas.simonson@polytechnique.fr.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Savvas Polydorides, David Mignon, and Thomas Gaillard for discussions.

REFERENCES

(1) You, T.; Bashford, D. *Biophys. J.* **1995**, *69*, 1721−1733.
(2) Beroza, P.; Case, D. A. *J. Phys. Chem.* **1996**, *100*, 20156−20163.
(3) Georgescu, E. R.; Alexov, E.; Gunner, M. *Biophys. J.* **2002**, *83*, 1731−1748.
(4) Barth, P.; Alber, T.; Harbury, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4898−4903.
(5) Kieseritzky, G.; Knapp, E. W. *Proteins* **2008**, *71*, 1335−1348.
(6) Song, Y.; Mao, J.; Gunner, M. R. *J. Comput. Chem.* **2009**, *30*, 2231−2247.
(7) Aleksandrov, A.; Polydorides, S.; Archontis, G.; Simonson, T. *J. Phys. Chem. B* **2010**, *114*, 10634−10648.
(8) Gunner, M.; Zhu, X.; Klein, M. C. *Proteins* **2011**, *79*, 3306−3319.
(9) Polydorides, S.; Simonson, T. *J. Comput. Chem.* **2013**, In Press.
(10) Vizcarra, C. L.; Mayo, S. L. *Curr. Opin. Chem. Biol.* **2005**, *9*, 622−626.
(11) Vizcarra, C. L.; Zhang, N. G.; Marshall, S. A.; Wingreen, N. S.; Zeng, C.; Mayo, S. L. *J. Comput. Chem.* **2008**, *29*, 1153−1162.
(12) Pantazes, R. J.; Greenwood, M. J.; Maranas, C. D. *Curr. Opin. Struct. Biol.* **2011**, *21*, 467−472.
(13) Samish, I.; MacDermaid, C. M.; Perez-Aguilar, J. M.; Saven, J. G. *Annu. Rev. Phys. Chem.* **2011**, *62*, 129−149.
(14) Fröhlich, H. *Theory of Dielectrics*; Clarendon Press: Oxford, 1949.
(15) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144−1149.
(16) Simonson, T. *Rep. Prog. Phys.* **2003**, *66*, 737−787.
(17) Simonson, T. *Photosynth. Res.* **2008**, *97*, 21−32.
(18) Gilson, M.; Honig, B. *Biopolymers* **1986**, *25*, 2097−2119.
(19) Nakamura, H.; Sakamoto, T.; Wada, A. *Protein Eng.* **1988**, *2*, 177−183.
(20) King, G.; Lee, F.; Warshel, A. *J. Chem. Phys.* **1991**, *95*, 4366−4377.
(21) Simonson, T.; Perahia, D.; Brünger, A. T. *Biophys. J.* **1991**, *59*, 670−90.
(22) Smith, P.; Brunne, R.; Mark, A.; van Gunsteren, W. *J. Phys. Chem.* **1993**, *97*, 2009−2014.
(23) Simonson, T.; Perahia, D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1082−1086.
(24) Simonson, T.; Perahia, D. *J. Am. Chem. Soc.* **1995**, *117*, 7987−8000.
(25) Simonson, T.; Brooks, C. L. *J. Am. Chem. Soc.* **1996**, *118*, 8452−8458.
(26) Simonson, T.; Archontis, G.; Karplus, M. *J. Phys. Chem. B* **1999**, *103*, 6142−6156.
(27) Krishtalik, L.; Kuznetsov, A. M.; Mertz, E. L. *Biophys. J.* **1996**, *70*, A225.
(28) Song, X.; Chandler, D.; Marcus, R. A. *J. Phys. Chem.* **1996**, *100*, 11954−11959.
(29) Voges, D.; Karshikoff, A. *J. Chem. Phys.* **1998**, *108*, 2219−2227.
(30) Boresch, S.; Höchtl, P.; Steinhauser, O. *J. Phys. Chem. B* **2000**, *104*, 8743−8752.
(31) Pitera, J.; Falta, M.; van Gunsteren, W. *Biophys. J.* **2001**, *80*, 2546−2555.
(32) Song, X. *J. Chem. Phys.* **2002**, *116*, 9359−9363.
(33) Park, H.; Jeon, Y. H. *Phys. Rev. E* **2007**, *75*, DOI: 10.1103/PhysRevE.75.021916.
(34) Goh, G. B.; Garcia-Moreno, B.; Brooks, Charles L., I. *J. Am. Chem. Soc.* **2011**, *133*, 20072−20075.
(35) Guest, W. C.; Cashman, N. R.; Plotkin, S. S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6286−6295.
(36) Martin, D. R.; Matyushov, D. V. *J. Chem. Phys.* **2012**, *137*, 165101.
(37) Neumann, M. *Mol. Phys.* **1983**, *50*, 841−858.
(38) De Leeuw, S.; Perram, J.; Smith, E. *Annu. Rev. Phys. Chem.* **1986**, *37*, 245−270.
(39) Perram, J.; Smith, E. *J. Stat. Phys.* **1986**, *46*, 179−190.
(40) Jackson, J. *Classical Electrodynamics*; Wiley: New York, 1975.
(41) Netz, R. R.; Orland, H. *Eur. Phys. J. E* **2000**, *1*, 203−214.
(42) Gascoyne, P.; Pethig, R. *J. Chem. Soc., Faraday Trans.* **1981**, *77*, 1733−1735.
(43) Bone, S.; Pethig, R. *J. Mol. Biol.* **1982**, *157*, 571−575.
(44) Zhou, H. *J. Chem. Phys.* **1996**, *105*, 3726−3733.
(45) Mertz, E. L.; Krishtalik, L. I. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2081−2086.

(46) Archontis, G.; Simonson, T. *J. Am. Chem. Soc.* **2001**, *123*, 11047−11056.
(47) Simonson, T. *Curr. Opin. Struct. Biol.* **2001**, *11*, 243−252.
(48) Simonson, T. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6544−6549.
(49) Simonson, T.; Gaillard, T.; Mignon, D.; Schmidt am Busch, M.; Lopes, A.; Amara, N.; Polydorides, S.; Sedano, A.; Druart, K.; Archontis, G. *J. Comput. Chem.* **2013**, DOI: 10.1002/jcc.23418.
(50) Landau, L.; Lifschitz, E. *Statistical Mechanics*; Pergamon Press: New York, 1980.
(51) Simonson, T. *Int. J. Quantum Chem.* **1999**, *73*, 45−57.
(52) Kusalik, P. *Mol. Phys.* **1993**, *80*, 225−231.
(53) Kusalik, P.; Mandy, M.; Svishchev, I. *J. Chem. Phys.* **1994**, *100*, 7654−7664.
(54) Simonson, T.; Perahia, D. *Farad. Disc* **1996**, *103*, 71−90.
(55) Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins* **2008**, *73*, 765−783.
(56) Olsson, M. H. M.; Sondergaard, C. R.; Rostowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(57) Hawkins, G. D.; Cramer, C.; Truhlar, D. *Chem. Phys. Lett.* **1995**, *246*, 122−129.
(58) Lopes, A.; Aleksandrov, A.; Bathelt, C.; Archontis, G.; Simonson, T. *Proteins* **2007**, *67*, 853−867.
(59) Schmidt am Busch, M.; Lopes, A.; Amara, N.; Bathelt, C.; Simonson, T. *BMC Bioinformatics* **2008**, *9*, 148−163.
(60) Tuffery, P.; Etchebest, C.; Hazout, S.; Lavery, R. *J. Biomol. Struct. Dyn.* **1991**, *8*, 1267.
(61) Aleksandrov, A.; Thompson, D.; Simonson, T. *J. Mol. Recog.* **2010**, *23*, 117−127.
(62) Polydorides, S.; Amara, N.; Simonson, T.; Archontis, G. *Proteins* **2011**, *79*, 3448−3468.
(63) Dahiyat, B. I.; Mayo, S. L. *Science* **1997**, *278*, 82−87.
(64) Pokala, N.; Handel, T. M. *J. Mol. Biol.* **2005**, *347*, 203−227.
(65) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.
(66) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: New York, 1996.
(67) Brünger, A. T. *X-plor version 3.1, A System for X-ray Crystallography and NMR*; Yale University Press: New Haven, CT, 1992.
(68) Jorgensen, W.; Chandrasekar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926−935.
(69) Bogusz, S.; Cheatham, T. E.; Brooks, B. R. *J. Chem. Phys.* **1998**, *108*, 7070−7084.
(70) Brooks, B.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(71) Mackerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(72) Alexov, E.; Mehler, E. L.; Baker, N.; Baptista, A. M.; Huang, Y.; Milletti, F.; Nielsen, J. E.; Farrell, D.; Carstensen, T.; Olsson, M. H. M.; Shen, J. K.; Warwicker, J.; Williams, S.; Word, J. M. *Proteins* **2011**, *79*, 3260−3275.
(73) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219−10225.
(74) Krishtalik, L. *Russ. J. Electrochem.* **2006**, *42*, 1006−1016.
(75) Mandell; Kortemme, T. *Curr. Opin. Biotechnol.* **2009**, *20*, 420−428.