Testing and optimizing the Drude polarizable force field for blocked amino acids based on high-level quantum-mechanical energy surfaces

Jinfeng Chen

Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China

Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China

Institute of Biology, Westlake Institute for Advanced Study, Hangzhou, Zhejiang , China

Gerhard König

Centre for Enzyme Innovation, University of Portsmouth St. Michael's Building, Portsmouth PO1 2DT, UK gerhard.koenig@port.ac.uk

Received Day Month Year
Revised Day Month Year
Accepted Day Month Year

The correct reproduction of conformational substates of amino acids was tested for the CHARMM Drude polarizable force field. This was achieved by evaluating the reorganization energies for all low lying energy minima occurring in all 15 neutral blocked amino acids on a quantum-mechanical (QM) energy surface at the MP2/cc-pVDZ level. The results indicate that the bonded parameters of the N-acetyl (ACE) and N-Methylamide (CT3) blocking groups lead to significant discrepancies. A reparametrization of five bond angles significantly improved the agreement with the QM energy surface. The corrected Drude force field exhibits almost the same average reorganization energies relative to the MP2 energy surface as the AM1 and PM3 semiempirical methods.

Keywords: Polarizable force field; validation; QM; amino acids

1. Introduction

Proteins are the most important biomolecules both biologically and for industrial applications. The structure and dynamics of proteins can be understood in terms of two factors:[1–9] First, the interactions of the protein with itself, which correspond to the conformational equilibria in the gas phase. Those equilibria are governed by the potential energies of the different conformational states. The second driving force for protein dynamics are the interactions of the protein with its environment, particularly through the hydrophobic effect that is caused by the surrounding water. Those solute-solvent interactions modulate the well depths and positions of the

conformational energy minima according to their solvation free energy.[10,11] The solvation free energy is defined as the free energy of transfer from the gas phase to solution.[12–14] To simulate a protein, both the potential energies of the conformational substates in the gas phase and the solute-solvent interactions have to be correctly represented in the model.[15]

Polarizable force fields are particularly suited to gauge the effect of solvation.[16–23] In contrast to fixed charge force fields, which employ the same charge distribution in the gas phase and in solution, the charge distributions of polarizable force fields can respond to the changes of the environment. The hydration free energies of simple solutes with the Drude polarizable force field[17,24] were previously compared to the ones from the fixed charge CHARMM36 force field,[25] exhibiting a 40% reduction of the root mean square error from the experimental reference results.[26] Quantum-mechanical (QM) and hybrid quantum-mechanics/molecular mechanics (QM/MM) methods are also particularly suited to determine solvation free energies if they are combined with a suitable representation of the solvent environment.[27,28,28–33] However, the computational expenses of simulations with QM methods are significantly higher than with polarizable force fields, which makes the use of force fields much more attractive for most practical applications.

This leaves the question in how far polarizable force fields are able to reproduce conformational equilibria in the gas phase. To this end, high level quantum-mechanical calculations can be used as accurate reference results to test the performance of force fields. To evaluate the correct reproduction of the conformational substates of proteins, it is possible to resort to the YMPJ database.[34,35] This database contains the low-energy confomers of all amino acids at the MP2/cc-pVDZ[36,37] level of theory. It includes 11 low-energy structures for Ala, 12 for Asn, 21 for Gln, 9 for Gly, 28 or Leu, 25 for Ile, 30 for Phe, 5 for Pro, 26 for Ser, 17 for Thr, 26 for Trp, 17 for Tyr, and 15 structures for Val. This makes in total 242 different structures of the energy minima of conformational substates. For these structures the differences between the Drude polarizable force field and QM at the MP2/cc-pVDZ level of theory are evaluated.

The thermodynamically relevant measure for the agreement between two energy surfaces is the phase space overlap.[38] The phase space overlap is proportional to the required sampling on one energy surface to cover all relevant low energy regions of the other energy surface, which also drives the convergence of free energy calculations.[39] For harmonic systems, like a single energy well, the phase space overlap is proportional to the reorganization energy (ΔU_{reorg}).[39] To determine the reorganization energy for an MM force field relative to a QM energy surface in one conformational substate, one can simply perform an MM energy minimization starting from the structure of the QM energy minimum and measure how much the potential energy drops as the MM energy minimum is reached (Fig. 1). Thus, the reorganization energy measures how much energy has to be invested in an MM simulation to visit the correct QM energy minimum. This principle was recently employed to evaluate the quality of several fixed charged force fields.[40]

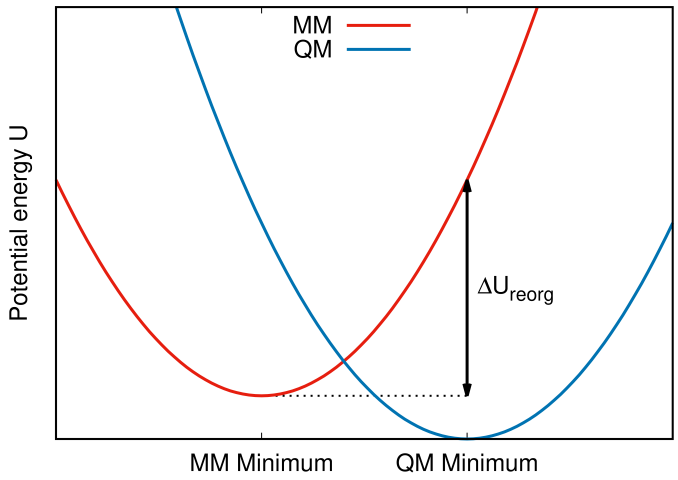

Fig. 1. The reorganization energy ΔU_{reorg} is a measure for the energy that has to be invested to sample the QM energy minimum within the MM energy surface when starting from the MM energy minimum. The reaction coordinate on the x-axis and the potential energy on the y-axis are in arbitrary units.

Preliminary calculations of the reorganization energies for blocked serine using the Drude polarizable force field indicated high deviations between the MM and QM energy surfaces.[40] Those discrepancies could be addressed by automatically adapting the reference bond lengths and reference bond angles of the force field, which indicates that at least one of those types of energy terms is responsible for the observed deviations from the QM energy surface. Here, the discrepancies are systematically assessed for all 15 neutral amino acids and then corrected by reparametrizing the degrees of freedom with the highest deviations between the QM and the MM optimal structures. The average reorganization energies for all 15 amino acids are then evaluated with the new parameters and compared to the outcome with the original parameters. Finally the performance of the new parameters is compared with the reorganization energies from the AM1[41] and PM3[42] semiempirical methods. Semiempirical approaches represent the next higher level of theory with increased computational demands[43–46] and, therefore, serve as a useful reference for the relative merits of polarizable force fields.

2. Theory

The main criterion for the similarity of two energy surfaces is the phase space overlap.[38] Formally, the phase space overlap between an MM energy surface and a QM energy surface can be determined by exhaustively sampling every point of both phase spaces, and evaluating the potential energies and probabilities of each confor-

mation for both MM and QM. However, this procedure is extremely computationally expensive. If the phase space overlap only has to be evaluated for one particular conformational substate, it is possible to resort to linear response theory[47–55] and Marcus theory[56] to estimate the expected phase space overlap via the reorganization energy (ΔU_{reorg}).[39,40] As shown previously,[39] the expected variance of a free energy estimate between the MM and the QM energy surface with the Zwanzig equation[57] (also known as free energy perturbation or exponential formula) is then given by

$$\sigma^2_{\Delta G} \approx \frac{e^{\frac{2}{kT}\Delta U_{reorg}} - 1}{n}, \tag{1}$$

where k is the Boltzmann constant, T is the temperature, and n represents the number of independent potential energy difference samples.[39,58,59] If multiple low-lying energy mimima are present in the system, Equation 1 has to be evaluated for all relevant conformational substates. If all conformational substates are equally important for the comparison, i.e., if one is interested in the overall similarity of the two phase spaces, then the relevant metric for the similarity is the average reorganization energy over all conformational substates ($|\Delta U_{reorg}|$). To estimate the similarity between the MM and QM energy surfaces in proteins we resort to the overall performance for all amino acid types, which is given by the average over all $|\Delta U_{reorg}|$ values over all considered amino acids (denoted as $||\Delta U_{reorg}||_{aa}$).

3. Methods

3.1. *Determination of the reorganization energies*

All molecular mechanics calculations were carried out with the CHARMM simulation package,[60,61] using the Drude polarizable force field.[17,24] For the determination of the average reorganization energies $|\Delta U_{reorg}|$, all energy minima of the N-acetyl-X-methylamide versions of the 15 neutral amino acids alanine (Ala), asparagine (Asn), cysteine (Cys), glutamine (Gln), glycine (Gly), isoleucine (Ile), leucine (Leu), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tryptophan (Trp), tyrosine (Tyr), and valine (Val) were evaluated. The structures of the energy minima at the MP2/cc-pVDZ level of theory[36,37] were obtained from the work of Kesharwani et al.[34] The same protocol as in previous calculations[40] of ΔU_{reorg} was used for the Drude polarizable force field. The only exception is the addition of the amino acid Trp to the benchmark, which was omitted in the previous calculations. All protonizable amino acids were excluded from the benchmark to avoid the introduction of net charges in the system, which might bias the results. The gas phase energy minimizations used 100 steps of steepest descent, followed by up to 10,000 steps of Adopted Basis Newton-Raphson (ABNR) energy minimization. To verify that the reorganization energies are converged, the minimizations were repeated with the conjugate gradient method. This only lead to insignificant changes of the results (on average about 0.05%). The nonbonded

interactions employed a switching function between 10 and 12 Å. For the MM calculations, the contributions of the individual energy terms to the reorganization energy were determined by evaluating the potential energy difference of each energy term between the MM and the QM optimized structure. The energy minimizations with the AM1[41] and PM3[42] semiempirical methods were performed with the MNDO package, using the default settings.[62]

3.2. *Reparametrization of the bond angles*

As will be discussed in the next section, the deviations of $||\Delta U_{reorg}||_{aa}$ mainly arise from the bond angle terms of the common backbone. Therefore, the simple alanine dipeptide (ALAD) system was used to perform the reparametrization. The optimized geometry of ALAD at the molecular mechanics level was generated with 100 steps of steepest descent, followed by 10000 steps of ABNR energy minimization using the lowest energy QM geometry at the MP2/cc-pVTZ level of theory as the initial structure. Thus, different basis sets are used for the reparametrization (cc-pVTZ) and the reorganization energies from the reference YMPJ database (cc-pVDZ), but this is necessary for consistency with the CHARMM parametrization strategy. To conduct the potential energy surface scan, the geometries of every QM scan were extracted with a fixed scanning angle, followed by the same MM optimization procedure as above. The vibrational frequency analysis was performed using MOLVIB module of CHARMM. All quantum-mechanical scan calculations were carried out using the Gaussian16 program.[63] The relaxed potential energy scan and the vibrational frequency analysis were performed at the MP2/cc-pVTZ level of theory.[36,37]

4. Results

4.1. *Reorganization energies with the original parameters*

The left side of Table 1 lists the total average of the average reorganization energies $||\Delta U_{reorg}||_{aa}$for the original parameters of the Drude polarizable force field for proteins,[17,24] considering all energy minima of the 15 neutral blocked amino acids. The new parameters are listed in Tables S1 and S2 of the SI. In preliminary calculations,[40] this force field exhibited high reorganization energies (above $100\,\mathrm{kJ\,mol^{-1}}$) for blocked serine relative to the energy surfaces of both MP2 and the OM2[64,65] semiempirical method. The high reorganization energies relative to OM2 was interpreted as an indicator that semiempirical calculations of ΔU_{reorg} can help to detect errors of the force field. A decomposition of the reorganization energies in regard to the individual force field terms revealed that the main contribution of this error arises from the bond angle terms. An automatic procedure[26] to adjust the bond length and bond angle parameters based on the quantum-mechanical reference structure with the lowest energy was able to reduce the average reorganization energy by 69%. This further indicates that the main discrepancies between the po-

Table 1. Total average of the average reorganization energies for the 15 neutral blocked amino acids ($||\Delta U_{reorg}||_{aa}$) with the original (old) and the new parameters for the Drude polarizable force field, as well as the average contributions from different MM energy terms.

Old Parameters		**New Parameters**	
$||\Delta U_{reorg}||_{aa}$	31.5 kcal mol^{-1}	$||\Delta U_{reorg}||_{aa}$	13.1 kcal mol^{-1}
Bond	5 %	Bond	33 %
Angle	66 %	Angle	7 %
Dihedral	4 %	Dihedral	20 %
Improper	5 %	Improper	1 %
CMAP	1 %	CMAP	2 %
vdW	15 %	vdW	31 %
Elec	4 %	Elec	8 %

larizable force field and the reference QM energy surface arise from the bonded terms.

To verify that this problem is not restricted to serine alone, the average reorganization energies and the individual contributions from the force field terms were determined. While a decomposition of the potential energy into contributions is only rigorous if there is no coupling between the energy terms,[66] this analysis can still help to guide the interpretation of the reorganization energies.[67,68] The results for the original parametrization are listed on the left side of Table 1. The total average reorganization energy of 31.5 kcal mol^{-1} (131.8 kJ mol^{-1}) shows that the discrepancies are high for all amino acids, which indicates that the main source of error lies in the common protein backbone or the N-acetyl (ACE) and N-Methylamide (CT3) blocking groups.

Table 1 also provides the average contributions of the individual force field energy terms, such as bond stretching (Bond), angle bending (Angle), dihedral angles (Dihedral), improper dihedrals (Improper), the cross term correction potential for the backbone dihedrals (CMAP), the van der Waals interactions (vdW), and electrostatic interactions (Elec). Strikingly, 66% of the observed discrepancies can be attributed to bond angle terms. Therefore, the reparametrization strategy aimed mainly at the bond angle terms of the backbone and the blocking groups. Because the reparametrization does not affect specific residues but the common backbone, the simple alanine dipeptide system (ALAD) is used for reference QM calculations at the MP2/cc-pVTZ level to derive the new parameters. The different basis set is necessary for consistency with the CHARMM parameterization strategy.

4.2. *Reparametrization*

The main targets of the reparametrization were identified based on the differences of the bond angle values between the QM global energy minimum at the MP2/cc-

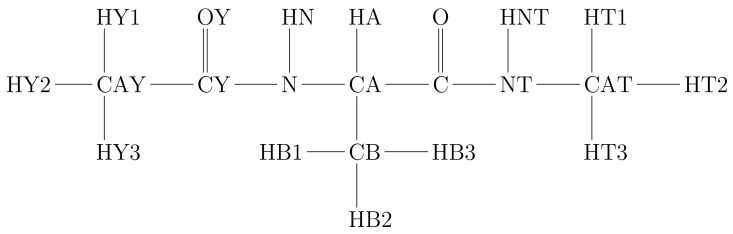

Fig. 2. Atom names of alanine dipeptide (ALAD).

Table 2. Comparison of the optimal angle values for the lowest energy structures at the MP2/cc-pVTZ level of theory and with the original and new force field parameters.

Angle	Optimal angle [°]			ΔU_{reorg} [kcal mol^{-1}]	
	MP2	Old MM	New MM	ΔU_{reorg}^{old}	ΔU_{reorg}^{new}
OY-CY-N	122.5	124.0	121.5	0.08	0.03
OY-CY-CAY	122.5	114.0	121.5	3.53	0.08
C-NT-CAT	119.5	123.0	120.5	0.52	0.04
N-CY-CAY	115.0	120.0	116.5	2.38	0.18
NT-CAT-HT1	109.0	120.0	110.0	3.89	0.02

pVTZ level and the polarizable force field optimal geometry of alanine dipeptide. The naming scheme for the atoms in alanine dipeptide can be found in Fig. 2. The interface between the protein backbone and the terminal blocking groups exhibited the largest discrepancies, particularly the angles formed by CAY, CY, OY, N. The optimal values of these angles at the MP2/cc-pVTZ level of theory and the Drude polarizable force field are listed in Table 2. Importantly, these discrepancies only arise when using blocking groups together with amino acids, so the present reparametrization does not affect the polypeptide chains in proteins.

The differences between the QM lowest energy structure at the MP2/cc-pVTZ level of theory and the original parameters of the Drude polarizable force field are 1.5°, 8.5°, 3.5°, 5°, 11° for the angles OY-CY-N, OY-CY-CAY, C-NT-CAT, N-CY-CAY, NT-CAY-HT1, respectively. Combining those discrepancies with the corresponding force constants of the bond angles leads to reorganization energies for the individual terms between 0.08 and 3.89 kcal mol^{-1}, as listed on the right side of Table 2. In particular, the angles N-CY-CAY and NT-CAT-HT1 exhibit high deviations from the QM energy minimum.

The new force field parameters are generated by fitting to a scan of the corresponding degrees of freedom. The corresponding results of the new parameters are shown in Table 2. The differences between the new reference angle values and the optimal geometry at the MP2 level of theory are all below 1.5°. The corresponding reorganization energies of the new parameters for those degrees of freedom range

between 0.02 and 0.18 kcal mol^{-1}. The low reorganization energies signify that the phase space overlap between the new parameters and the QM energy surface of the MP2/cc-pVTZ level of theory is significantly higher. The underlying potential energy scans can be found in the Supporting Information.

4.3. *Comparison of original and new parameters*

The total average of the average reorganization energies over all amino acids and the contributions of the individual force field terms for the new parameters are listed on the left side of Table 1. By adjusting five bond angle parameters of the Drude polarizable force field, the average reorganization energy drops by 58% from 31.5 to 13.1 kcal mol^{-1}. The contribution from the bond angle energy terms to the total average of the reorganization energies is merely 7% with the new parameters. Instead, the dominant sources of discrepancies between the polarizable force field and the quantum mechanical energy surface are the bond lengths (33%), van der Waals parameters (31%) and dihedral angles (33%). Those contributions are similar to the ones that were observed for the CHARMM22[69] and CHARMM36[25] force fields in a previous publication.[40] However, due to the extra degrees of freedom from the Drude particles to incorporate the effects of polarization, the reorganization energies and the contributions are not directly comparable.

Notably, all deviations of the original parameters are related to the peptide bonds. As shown in Table 2, the largest deviations are observed for the angles of NT-CAT-HT1, OY-CY-CAY, and N-CY-CAY, respectively. The atoms CY and CAT are bonded to the C_α atom (denoted as CA in the CHARMM force field) of the neighboring residues in proteins. The original parameters mostly reflect the situation in a protein environment. However, in the present work, these atoms are bonded to the terminal blocking groups, rather than to other amino acids. The extrapolation of the parameters that were developed for a protein backbone to terminal blocking groups leads to the observed discrepancies. Because there are only two termini in proteins, but many more normal peptide bonds between amino acids, the original parameters are optimized for the latter type. Vice versa, the new parameters for the termini will not significantly affect protein dynamics. However, when studying small peptides, like the ones in the YMPJ database, the original parameters might lead to biased results, especially in long molecular dynamics simulations. Thus special parameters for the terminal peptide bonds are necessary.

4.4. *Comparison with semiempirical methods*

To provide a comparison of the new parameters of the Drude polarizable force field to other methods that include polarization effects, we resort to the semiempirical quantum-mechanical methods AM1[41] and PM3.[42] Since the potential energies of the semiempirical methods are not composed from individual energy terms like a force field, it is only possible to compare the average reorganization energies for each amino acid type, as listed in Table 3. The last column shows the average

Table 3. Average reorganization energies $|\Delta U_{reorg}|$ for the 15 neutral blocked amino acids with the semiempirical methods AM1 and PM3, as well as the original and new force field parameters (MM). The last row gives the total average of the reorganization energies over all amino acids. All values are in kcal mol^{-1}.

	AM1	PM3	Old MM	New MM
Ala	10.6	11.1	28.2	10.4
Asn	13.3	14.7	33.2	13.6
Cys	11.7	12.3	30.6	12.0
Gln	14.5	16.6	36.6	16.3
Gly	9.4	10.3	27.5	10.9
Ile	13.3	12.4	29.9	11.9
Leu	13.8	12.8	29.6	12.0
Met	14.3	13.3	31.0	12.2
Phe	12.4	12.6	31.7	13.1
Pro	16.1	14.2	29.3	15.5
Ser	12.4	11.6	30.1	12.1
Thr	13.1	13.4	33.9	15.3
Trp	15.0	15.5	33.4	15.4
Tyr	13.1	12.8	32.9	14.3
Val	12.4	11.6	30.5	12.2
$\|\|\Delta U_{reorg}\|\|_{aa}$	**13.0**	**13.0**	**31.5**	**13.1**

reorganization energies of the Drude polarizable force field with the original (Old MM) and new parameters (New MM). The last row gives the total average of the reorganization energies over all amino acids. Surprisingly, although the average reorganization energies for the individual amino acids vary among the methods, the resulting total averages of the semiempirical methods and the new Drude force field parameters are very similar with 13.0, 13.0, and 13.1 kcal mol^{-1}. This indicates that those three methods reproduce the QM energy surface at the MP2/cc-pVDZ level of theory equally well.

5. Conclusions

Polarizable force fields are known to be superior to conventional fixed charge force fields when it comes to determining solute-solvent interactions. However, to evaluate whether polarizable force fields are also able to correctly capture protein dynamics, it also needs to be shown that they correctly reproduce the conformational equilibria in the gas phase. Here, this is achieved by determining the phase space overlap of the Drude polarizable force field with QM-generated conformational substates from the YMPJ database.[34,35] The previously reported weaknesses of the bonded terms of N-

acetyl-X-methylamide amino acids in the CHARMM polarizable force field[40] were corrected by adapting the bond angles at the interface between the protein backbone and the terminal blocking groups. While the original parameters exhibited a total average reorganization energy $||\Delta U_{reorg}||_{aa}$ of 31.5 kcal mol^{-1} relative to the QM reference structures at the MP2/cc-pVDZ level, the new parameters only exhibit a total average reorganization energy of 13.1 kcal mol^{-1} over all neutral amino acids. Thus, the agreement between the QM energy surface and the Drude polarizable force field has been greatly improved.

Some challenges remain in terms of the bond length parameters (which contribute 33% to the average reorganization energy), as well as the van der Waals parameters (31%) and dihedral angles (20%). However, similar discrepancies were observed previously for the CHARMM22 and CHARMM36 force fields,[40] which indicates that those deviations arise from the underlying parametrization strategy of the CHARMM force field. Most of the deviations of the bond length parameters are on the order of picometers,[40] but due to the stiffness of the bonded terms such small deviations can lead to sizeable potential energy differences. However, small deviations of the bond length can be easily addressed in multiscale free energy simulations by automatically updating the bond parameters to create a tailored force field for the QM target energy surface,[26,70,71] or by calculating corrections during post-processing.[72]

The average reorganization energies with respect to quantum-mechanical calculations serve as a useful indicator for the quality of the force field. To put the average reorganization energies of the Drude polarizable force field into perspective, the same test was also performed for the AM1 and PM3 semiempirical methods. The semiempirical methods represent the next higher level of theory in terms of rigour. In contrast to conventional fixed-charge force fields, quantum-mechanical semiempirical methods are able to incorporate polarization effects, which makes them a suitable reference for polarizable force fields in terms of performance. The total average reorganization energies over all energy minima of all considered 15 amino acids are 13.0 kcal mol^{-1} for both the AM1 and PM3 method. Thus the accuracy of the new parameters of the Drude polarizable force field (13.1 kcal mol^{-1}) is comparable to the semiempirical methods. Notably, the polarizable force field based on Drude particles exhibits lower computational demands and a better scalability than the semiempirical methods. This makes the Drude polarizable force field an interesting option for biomolecular simulations, particularly for proteins.

Acknowledgments

GK gratefully acknowledges financial support by Research England for E3 funding. The authors also thank Jing Huang for his help with the reparametrization of the polarizable force field.

References

1. Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1.
2. Makhtadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1993**, *232*, 639.
3. Privalov, P. L.; Makhtadze, G. I. *J. Mol. Biol.* **1993**, *232*, 660.
4. Dyson, H. J.; Wright, P. E.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13057–13061.
5. Langhorst, U.; Backmann, J.; Loris, R.; Steyaert, J. *Biochemistry* **2000**, *39*, 6586.
6. Kokubo, H.; Hu, C. Y.; Pettitt, B. M. *J. Am. Chem. Soc.* **2011**, *133*, 1849–1858.
7. Tarek, M.; Tobias, D. *J. Am. Chem. Soc.* **1999**, *121*, 9740.
8. Makarov, V.; Pettitt, B.; Feig, M. *Acc. Chem. Res.* **2002**, *35*, 376–384.
9. Fulle, S.; Gohlke, H. *J. Mol. Recognit.* **2010**, *23*, 220–231.
10. König, G.; Boresch, S. *J. Phys. Chem. B* **2009**, *113*, 8967–8974.
11. König, G.; Bruckner, S.; Boresch, S. *Biophys. J.* **2013**, *104*, 453–462.
12. Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792–803.
13. Ben-Naim, A.; Marcus, Y. J. *J. Chem. Phys.* **1984**, *81*, 2016–2027.
14. Wilhelm, E. *J. Solution Chem.* **2015**, *44*, 1004–1061.
15. van Gunsteren, W. F.; Daura, X.; Hansen, N.; Mark, A. E.; Oostenbrink, C.; Riniker, S.; Smith, L. J. *Angewandte Chemie International Edition* **2018**, *57*, 884–902.
16. Inakollu, V. S.; Geerke, D. P.; Rowley, C. N.; Yu, H. *Curr. Opin. Struct. Biol.* **2020**, *61*, 182–190.
17. Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell Jr, A. D. *Chem. Rev.* **2016**, *116*, 4983–5013.
18. Bradshaw, R. T.; Essex, J. W. *J. Chem. Theory Comput.* **2016**, *12*, 3871–3883.
19. Visscher, K. M.; Vosmeer, C. R.; Luirink, R. A.; Geerke, D. P. *jcc* **2017**, *38*, 508–517.
20. Huang, J.; Simmonett, A. C.; Pickard, F. C.; MacKerell, A. D., Jr.; Brooks, B. R. *J. Chem. Phys.* **2017**, *147*.
21. Belsare, S.; Esser, A.; Marx, D.; Head-Gordon, T. *Biophys. J.* **2017**, *112*, 497a.
22. Bedrov, D.; Piquemal, J.-P.; Borodin, O.; MacKerell Jr, A. D.; Roux, B.; Schröder, C. *Chem. Rev.* **2019**, *119*, 7940–7995.
23. Lin, F.-Y.; Lopes, P. E.; Harder, E.; Roux, B.; MacKerell Jr, A. D. *J. Chem. Inf. Model.* **2018**, *58*, 993–1004.
24. Lopes, P. E.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell Jr, A. D. *J. Chem. Theory Comput.* **2013**, *9*, 5430–5449.
25. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
26. König, G.; Pickard, F.; Huang, J.; Thiel, W.; MacKerell, A.; Brooks, B.; York, D. *Molecules* **2018**, *23*, 2695.
27. Pickard, F. C.; König, G.; Simmonett, A. C.; Shao, Y.; Brooks, B. R. *Biorg. Med. Chem.* **2016**, *24*, 4988–4997.

28. König, G.; Pickard, F. C.; Mei, Y.; Brooks, B. R. *J. Comput.-Aid. Mol. Design* **2014**, *28*, 245–257.
29. Huang, J.; Mei, Y.; König, G.; Simmonett, A. C.; Pickard IV, F. C.; Wu, Q.; Wang, L.-P.; MacKerell Jr, A. D.; Brooks, B. R.; Shao, Y. *J. Chem. Theory Comput.* **2017**, *13*, 679–695.
30. Jia, X.; Wang, M.; Shao, Y.; König, G.; Brooks, B. R.; Zhang, J. Z. H.; Mei, Y. *J. Chem. Theory Comput.* **2016**, *12*, 499–511.
31. König, G.; Mei, Y.; Pickard, F. C.; Simmonett, A. C.; Miller, B. T.; Herbert, J. M.; Woodcock, H. L.; Brooks, B. R.; Shao, Y. *J. Chem. Theory Comput.* **2016**, *12*, 332–344.
32. Pickard, F. C.; König, G.; Tofoleanu, F.; Lee, J.; Simmonett, A. C.; Shao, Y.; Ponder, J. W.; Brooks, B. R. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 1087–1100.
33. König, G.; Pickard, F. C.; Huang, J.; Simmonett, A. C.; Tofoleanu, F.; Lee, J.; Dral, P. O.; Prasad, S.; Jones, M.; Shao, Y.; Thiel, W.; Brooks, B. R. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 989–1006.
34. Kesharwani, M. K.; Karton, A.; Martin, J. M. *J. Chem. Theory Comput.* **2015**, *12*, 444–454.
35. Yuan, Y.; Mills, M. J.; Popelier, P. L.; Jensen, F. *J. Phys. Chem. A* **2014**, *118*, 7876–7891.
36. Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.
37. Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
38. Bennett, C. H. *J. Comp. Phys.* **1976**, *22*, 245–268.
39. König, G.; Brooks, B. R.; Thiel, W.; York, D. M. *Mol. Simulat.* **2018**, *44*, 1062–1081.
40. König, G.; Riniker, S. *Interface Focus* **2020**, *10*, 20190121.
41. Dewar, M.; Zoebisch, E.; Healy, E.; Stewart, J. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
42. Stewart, J. *J. Comput. Chem.* **1989**, *10*, 209–220.
43. Thiel, W. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 145–157.
44. Korth, M.; Thiel, W. *J. Chem. Theory Comput.* **2011**, *7*, 2929–2936.
45. Dral, P. O.; Wu, X.; Spörkel, L.; Koslowski, A.; Thiel, W. *J. Chem. Theory Comput.* **2016**, *12*, 1097–1120.
46. Margraf, J. T.; Dral, P. O. *J. Mol. Model.* **2019**, *25*, 1–12.
47. Åqvist, J.; Medina, C.; Samuelsson, J. E. *Prot. Eng.* **1994**, *7*, 385.
48. Hansson, T.; Marelius, J.; Åqvist, J. *J. Comp. Aid. Mol. Des.* **1998**, *12*, 27.
49. Almlof, M.; Carlsson, J.; Åqvist, J. *J. Chem. Theory Comput.* **2007**, *3*, 2162–2175.
50. Stjernschantz, E.; Marelius, J.; Medina, C.; Jacobsson, M.; Vermeulen, N. P. E.; Oostenbrink, C. *J. Chem. Inf. Model.* **2006**, *46*, 1972–1983.
51. Kolb, P.; Huang, D.; Dey, F.; Caflisch, A. *J. Med. Chem.* **2008**, *51*, 1179–1188.
52. Guitiérrez-de Terán, H.; Åqvist, J. *Computational Drug Discovery and Design*; Methods Mol. Biol.; 2012; Vol. 819; pp 305–323.
53. Hummer, G.; Pratt, L. R.; García, A. E. *J. Phys. Chem.* **1996**, *100*, 1206–1215.

54. Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512–9521.
55. Heid, E.; Moser, W.; Schröder, C. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10940–10950.
56. Marcus, R. *Ang. Chem. Int. Ed.* **1993**, *32*, 1111–1121.
57. Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
58. Zuckerman, D. M.; Woolf, T. B. *Phys. Rev. Lett.* **2002**, *89*, 180602.
59. Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 144107–1–144107–16.
60. Brooks, B. R. et al. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
61. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
62. Thiel, W. MNDO2005, version 7.1. Max-Planck-Institut für Kohlenforschung: Mülheim an der Ruhr, Germany, 2006.
63. Frisch, M. J. et al. Gaussian 16 Revision B.01. 2016; Gaussian Inc. Wallingford CT.
64. Weber, W.; Thiel, W. *Theor. Chem. Acc.* **2000**, *103*, 495–506.
65. Dral, P. O.; Wu, X.; Spörkel, L.; Koslowski, A.; Weber, W.; Steiger, R.; Scholten, M.; Thiel, W. *J. Chem. Theory Comput.* **2016**, *12*, 1082–1096.
66. Smith, P. E.; van Gunsteren, W. *J. Phys. Chem.* **1994**, *98*, 13735–13740.
67. Boresch, S.; Archontis, G.; Karplus, M. *Proteins: Struct. Funct. Genet.* **1994**, *20*, 25–33.
68. Boresch, S.; Karplus, M. *J. Mol. Biol.* **1995**, *254*, 801–807.
69. MacKerell, A. D. et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
70. Giese, T. J.; York, D. M. *J. Chem. Theory Comput.* **2019**, *15*, 5543–5562.
71. Hudson, P. S.; Boresch, S.; Rogers, D. M.; Woodcock, H. L. *J. Chem. Theory Comput.* **2018**, *14*, 6327–6335.
72. König, G.; Brooks, B. R. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850*, 932–943.