RESEARCH ARTICLE

Journal of COMPUTATIONAL CHEMISTRY   WILEY

Conformational energies of microsolvated Na$^+$ clusters with protic and aprotic solvents from GFN*n*-xTB methods

Arseniy A. Otlyotov[1] | Yury Minenkov[1,2]

[1]N.N. Semenov Federal Research Center for Chemical Physics RAS, Moscow, Russian Federation

[2]Joint Institute for High Temperatures, Russian Academy of Sciences, Moscow, Russian Federation

Correspondence
Yury Minenkov, N.N. Semenov Federal Research Center for Chemical Physics RAS, Kosygina Street 4, Moscow, 119991, Russian Federation.
Email: yury.minenkov@chph.ras.ru

Abstract

Performance of contemporary tight-binding semiempirical GFN*n*-xTB methods for the conformational energies of singly charged sodium clusters Na$^+$(S)$_n$ ($n = 4$–8) with 3 protic and 8 aprotic solvents is examined against the reference RI-MP2/CBS method. The median Pearson correlation coefficients of $\rho = 0.84$ (GFN2-xTB) and $\rho = 0.82$ (GFN1-xTB) do not give the clear preference to any tested approach. GFN1-xTB method demonstrates more stable performance than its GFN2-xTB successor with the average mean absolute errors (MAEs)/mean signed errors (MSEs) of 1.2/0.2 and 2.3/1.6 kcal mol^{-1}, respectively. Conformational energies produced by the computationally efficient DFT functional PBE and double-ζ basis set complemented with –D3(BJ) dispersion correction are suitable for the preliminary sampling (median $\rho = 0.93$), but should be used with a caution for the calculations of the average ensemble properties (MAE/MSE = 1.7/1.1 kcal mol^{-1}). Higher-ranking PBE0-D3 (BJ) and ωB97M-V with triple-ζ basis sets yield significantly lower MAEs/MSEs of 0.55/0.20 and 0.51/0.23 kcal mol^{-1}, respectively.

KEYWORDS
cluster-continuum theory, conformational analysis, microsolvation, semiempirical methods, sodium cation clusters

1 | INTRODUCTION

The evolution of observable physical and chemical properties when moving from an isolated molecule in the gas phase to solution is a central topic of chemical sciences. An avenue from gas to a solution crosses a so-called microsolvation region, where an isolated species is surrounded by a fixed number of solvent molecules.[1–11] As the number of explicit solvent molecules grows, the properties of such a system start mimicking the real solvent. The overall accuracy of the description of microsolvation appears in the proper identification of the low-energy spatial structures of such A$^{\pm}$S$_n$ clusters, where A represents a solute (either neutral, or ionic) and S denotes solvent molecules.

Obviously, the systematic conformational analysis is limited to the case of the small *n* values. Larger clusters possess many degrees of freedom associated with: (1) the number of solvent molecules in the solvation shells surrounding a solute; (2) relative orientation of

solvent molecules; (3) inherent conformational flexibility of solute and solvent molecules.

Such systems require automatic conformational sampling which is implemented into some quantum chemistry codes. One of the most contemporary and universal tools for the black-box determination of the thermally accessible ensembles of low energy structures is the CREST program[12,13] which uses metadynamic sampling (MTD) in conjunction with the GFN*n*-xTB/FF methods developed by Grimme and co-workers.[14–16] In their very recent work,[17] this procedure was incorporated into the new method, quantum cluster growth (QCG), aiming to describe the solvation process in an explicit fashion. In addition, the CREST/GFN*n*-xTB scheme is proposed for calculation of conformational entropy.[18] It should, however, be noted, that the applicability of the fast GFN*n*-xTB/FF methods for such challenging systems as A$^{\pm}$S$_n$ clusters should be carefully tested first.

Another recently developed tool for the conformational sampling of A$^{\pm}$S$_n$ clusters is the ABCluster program of Zhang.[19,20] It employs an

artificial bee colony (ABC) algorithm for the global optimization[21] and the CHARMM force field[22,23] for the potential energy calculations. Despite a number of encouraging results,[24–31] the general accuracy of such an algorithm for an arbitrary solvent requires further validation. Moreover, the applicability of this force field is limited, and some solvents require careful optimization of the CHARMM parameters.

Alternatively, a number of force field and semiempirical methods can be used in conjunction with the ABCluster, but only a few are of general applicability and require no parameterization for an individual system. These are the GFN-FF,[16] PM6*/7[32,33] and GFNn-xTB.[14,15] According to our internal tests, the GFN-FF demonstrated an unstable performance for some $A^{\pm}S_n$ systems and this method was ruled out in the present work. The semiempirical methods PM6*/7 were shown in general to be inferior to their GFNn-xTB competitors.[34–36]

For these reasons, in the present contribution we examine the performance of the emerging semiempirical GFNn-xTB methods in reproducing conformational energies of singly charged sodium clusters (Na$^+$S$_n$) with 3 protic (water, methanol, ethanol) and 8 aprotic (formamide, N-methylformamide, dimethylformamide, dimethylacetamide, diethylacetamide, dimethyl sulfoxide, 1,2-dichloroethane, propylene carbonate) solvents. Conformational sampling of these clusters is of crucial importance for detailed comprehension of the Na$^+$ solvation process both experimentally[37] and theoretically via the cluster-continuum approach.[38–40]

The diverse spatial conformations were generated with the use of the ABCluster and optimized with the pure GGA PBE functional in conjunction with a double-ζ basis. High-quality reference conformational energies were then obtained using the RI-MP2/CBS (TZ/QZ) method.

Along with the tight-binding GFNn-xTB, the two well-established dispersion-corrected DFT methods, hybrid PBE0-D3(BJ) and range-separated meta-hybrid ωB97M-V, in combination with a triple-ζ quality basis set were also examined against the RI-MP2 reference method. Besides, the performance of the computationally cheaper PBE functional complemented by –D3(BJ) dispersion correction, in conjunction with a double-ζ basis, that was used in this work for the geometry optimizations is discussed.

The following abbreviations of the solvents are used throughout the paper: MeOH, methanol; EtOH, ethanol; FA, formamide; NMF, N-methylformamide; DMF, dimethylformamide; DMA, dimethylacetamide; DEA, N,N-diethylacetamide; DMSO, dimethyl sulfoxide; 1,2-DCIE, 1,2-dichloroethane; PC, propylene carbonate.

2 | METHODS

The sets of spatially diverse conformations were obtained for each Na$^+$S$_n$ cluster (n = 4–8) using the following procedure. First, 50 structures were automatically generated with use of the *rigidmol* program included to the ABCluster suite.[19,20] To increase structural diversity the following settings were applied: the population size S$_n$ = 6000, the maximum of search generation g_{max} = 5, scout limit g_{limit} = 3. For the conformationally flexible solvents (EtOH, NMF, DEA, 1,2-DCIE)

two conformations were used to construct the corresponding clusters (see ESI). Second, all the structures were optimized in the Priroda 19 program.[41] The PBE functional[42,43] in combination with double-ζ all-electron λ1 basis sets[44] was used. The Dyall Hamiltonian was used to take scalar relativistic effects into account. Dispersion corrections –D3(BJ) were calculated using the standalone DFTD3 V3.0 program,[45,46] and added to the total electronic energies of the optimized conformations to yield the corresponding E(PBE-D3(BJ)) values further used for the comparison purposes.

Prior to the further statistical analyses, all the optimized structures were checked to exclude duplicates. The clusters with small numbers of the unique conformations were ruled out from the selection. Thus, at least 19 diverse spatial structures were obtained for each Na$^+$S$_n$ cluster. The final set of the clusters included to the benchmark set is listed below:

- for protic solvents: Na$^+$(H$_2$O)$_n$ (n = 6–8), Na$^+$(MeOH)$_n$ (n = 5–8), Na$^+$(EtOH)$_n$ (n = 4–8);
- for aprotic solvents: Na$^+$(FA)$_n$ (n = 6–8), Na$^+$(NMF)$_n$ (n = 5–8), Na$^+$(DMF)$_n$ (n = 5–8), Na$^+$(DMA)$_n$ (n = 5–8), Na$^+$(DEA)$_n$ (n = 4–8), Na$^+$(DMSO)$_n$ (n = 5–8), Na$^+$(1,2-DCIE)$_n$ (n = 4–8), Na$^+$(PC)$_n$ (n = 5–8).

Single-point energy calculations with use of contemporary semiempirical methods GFN1-xTB and GFN2-xTB were performed on the PBE/λ1 optimized geometries with use of the xtb 6.4.1 code.[47]

The reference conformational energies were obtained for all the optimized structures with use of the RI-MP2 method[48–51] as implemented in the ORCA 5.0.3 program.[52,53] Triple- and quadruple-ζ cc-pVnZ (n = 3, 4) basis sets were employed for (H, C, N, O),[54] S[55] and Cl[55] atoms. Na atom was described with the cc-pwCVnZ (n = 3,4)[56] and the sub-valence 2s^{2}2p^6 electrons were correlated according to the literature.[57–60] The corresponding correlation fitting basis sets cc-pVnZ/C[61] and cc-pwCVnZ/C (generated via the AutoAux routine)[62] requested for the resolution of identity (RI) approximation were employed. The auxiliary Coulomb and exchange fitting cc-pVnZ/JK and cc-pwCVnZ/JK (generated via AutoAux) basis sets were utilized in conjunction with the RIJCOSX approximation[63] to speed up the calculations. Complete basis set (CBS) extrapolation was performed with use of the Martin scheme[64]:

$$E^n_{RI-MP2} = E^{\infty}_{RI-MP2} + \frac{A}{\left(n+\frac{1}{2}\right)^4}$$

The choice of the reference level is motivated by its feasible computational cost, taking into account the large size of many structures (up to 169 atoms for Na$^+$(DEA)$_8$). Together with the large total number of structures (1874) it makes more accurate domain-based local pair natural orbital coupled cluster singles, doubles, and perturbative triples DLPNO-CCSD(T)/CBS[65–67] reference energies practically unaffordable.

Conformational energies were also obtained with the DFT methods accounting for the dispersion effects. Hybrid DFT functional

PBE0[68] complemented by Grimme's D3(BJ) correction and range-separated hybrid meta-GGA ωB97M-V[69] including VV10[70] nonlocal correlation to treat dispersion effects, in conjunction with triple-ζ all-electron def2-TZVP basis sets[71] were chosen. Tighter-than-default DFT integration grid *DEFGRID3* was used. The calculations were speeded up by RIJCOSX approximation.[63]

The quantitative comparison of the conformational energies was performed in terms of the Pearson correlation coefficients.

$$\rho(X,Y) = \frac{\sum_{i=1}^{n}\left(E_{x,i}-\overline{E_x}\right)\left(E_{y,i}-\overline{E_y}\right)}{\sqrt{\sum_{i=1}^{n}\left(E_{x,i}-\overline{E_x}\right)^2\left(E_{y,i}-\overline{E_y}\right)^2}},$$

where X is the examined electronic structure method (GFN1-xTB, GFN2-xTB or DFT), Y is the reference method (RI-MP2/CBS), E_i are the relative conformational energies, $\overline{E}$ are the average conformational energies for a given method. The positive ρ values close to 1 indicate a strong correlation between the tested methods, while negative ρ values close to -1 signify an anti-correlation.

Mean absolute error (MAE) and mean signed error (MSE) were also used to quantify the difference between relative conformational energies according to the formulae:

$$MAE(X,Y) = \frac{\sum_{i=1}^{n}|E_i(X)-E_i(Y)|}{n}$$

$$MSE(X,Y) = \frac{\sum_{i=1}^{n}(E_i(X)-E_i(Y))}{n}$$

where X and Y are the examined and reference electronic structure methods, n is the number of conformations for a Na$^+$S$_n$ cluster (see Tables S1–S3), E_i are the relative conformational energies.

3 | RESULTS AND DISCUSSION

This section is organized as follows. First, we describe the performance of the tight-binding GFNn-xTB methods for the group of protic solvents (H$_2$O, MeOH, EtOH) and two subgroups of aprotic solvents: (FA, NMF, DMF, DMA, DEA) and (1,2-DCIE, DMSO, PC). For each subgroup we examine (in)ability of the GFNn-xTB methods to predict the lowest or low-energy conformations. It is assumed that an examined method predicts low-energy conformation if its reference RI-MP2/CBS energy ($\Delta E_{ref.}$) is less than 0.1 kcal mol^{-1}. Besides, statistical measures (Pearson correlation coefficient, MAE and MSE) applied to particular Na$^+$S$_n$ clusters and the overall subgroups are discussed.

Second, the performance of the DFT methods (PBE-D3(BJ)/λ1, PBE0-D3(BJ)/def2-TZVP, ωB97M-V/def2-TZVP) is described in the similar fashion.

Finally, the overall performance of all tested methods applied to the complete set of 45 Na$^+$S$_n$ clusters is outlined in terms of the median and Q1/Q3 values of Pearson correlation coefficient and MAEs/MSEs averaged over the set.

3.1 | Water, methanol, and ethanol

Both GFNn-xTB methods correctly reproduce the low(est) energy structures for Na$^+$(H$_2$O)$_{6,8}$ Na$^+$(MeOH)$_{5,6,8}$ and Na$^+$(EtOH)$_{4,5}$ clusters. In the case of Na$^+$(H$_2$O)$_7$, Na$^+$(MeOH)$_7$ and Na$^+$(EtOH)$_7$ the most favorable conformations predicted by the examined semiempirical methods are 0.6–1 kcal mol^{-1} higher in their reference RI-MP2/CBS energies. Two clusters with ethanol, Na$^+$(EtOH)$_8$ and especially Na$^+$(EtOH)$_6$ are more challenging with the $\Delta E_{ref.}$ = 2.5/3.8 kcal mol^{-1} and $\Delta E_{ref.}$ = 6.8/2.4 kcal mol^{-1}, according to the GFN1-xTB/GFN2-xTB methods, respectively.

Statistical analysis in terms of the Pearson correlation coefficients reveals no superiority of the more recent GFN2-xTB method over its GFN1 predecessor. The respective mean correlation coefficients (Figure 1, left) averaged over 12 clusters are ρ = 0.67 and ρ = 0.70. The average MAEs of 1.3 kcal mol^{-1} (GFN1-xTB) and 1.2 kcal mol^{-1} (GFN2-xTB) are fairly small values. The hexamer ethanol cluster Na$^+$(EtOH)$_6$ is the most problematic with the lowest correlation coefficients for both GFN(1/2)-xTB methods (ρ = 0.33/0.43) and relatively large MAEs of 2.2/1.9 kcal mol^{-1}.

3.2 | Formamide, *N*-methyl formamide, dimethylformamide, dimethylacetamide, and diethylacetamide

The performance of the examined semiempirical methods for sodium clusters with FA, NMF, DMF, DMA, and DEA is nonhomogeneous: good results are interspersed with particular failures (Figure 2). There is a number of clusters, for which either both methods, or at least one of them successfully predict the low(est) energy conformations: Na$^+$(FA)$_6$, Na$^+$(FA)$_6$ (GFN1), Na$^+$(NMF)$_5$, Na$^+$(NMF)$_6$ (GFN1), Na$^+$(NMF)$_7$, Na$^+$(DMF)$_{7,8}$, Na$^+$(DMA)$_6$ (GFN2), Na$^+$(DMA)$_7$ (GFN1), and Na$^+$(DEA)$_{4,5,8}$. The largest discrepancies for the GFN1-xTB was yielded for Na$^+$(DEA)$_6$ ($\Delta E_{ref.}$ = 2.3 kcal mol^{-1}). Na$^+$(NMF)$_8$ is a puzzling system for the GFN2-xTB method with $\Delta E_{ref.}$ = 3.6 kcal mol^{-1}.

More insights were gained from statistical measures (ρ and MAE) applied to the considered subset of solvents. First, both GFNn-xTB conformational energies reveal anti-correlation with their reference RI-MP2/CBS counterparts for Na$^+$(DMF)$_{5,6}$ clusters. GFN1-xTB method additionally gives low (ρ = 0.19) or even absent (ρ = 0) correlation for Na$^+$(FA)$_8$ and Na$^+$(DMA)$_5$, respectively. An apparent reason for the problems in the conformational ranking for these systems lies in the narrow spans of the conformational energies (2.1–3.6 kcal mol^{-1} for Na$^+$(DMF)$_{5,6}$ and Na$^+$(DMA)$_{5,6}$) and corresponding low average relative conformational energies ($\Delta E_{ref.,avg.}$ = 1.0–1.8 kcal mol^{-1}). Engaging quite large MAEs (2.1/2.9 kcal mol^{-1} for Na$^+$(DMF)$_5$) in the analysis, we conclude that these clusters possessing different spatial structures with close energies are challenging for
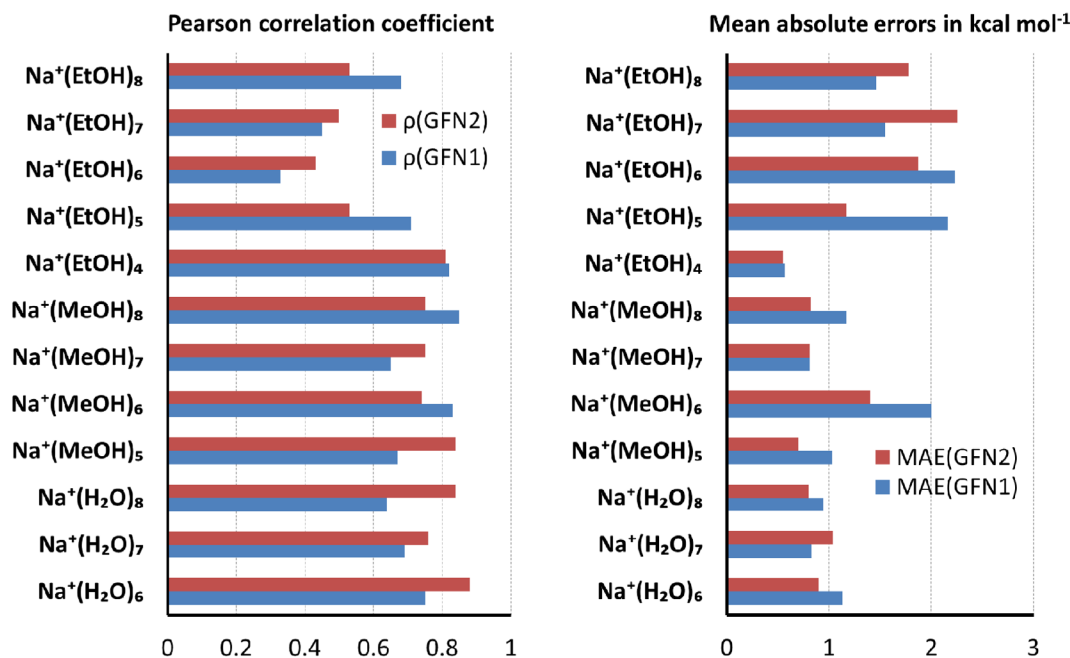
FIGURE 1 Pearson correlation coefficients (left) and mean absolute errors (right) between GFNn-xTB and reference (RI-MP2/CBS) conformational energies for sodium clusters with protic solvents (H_2O, MeOH, and EtOH)
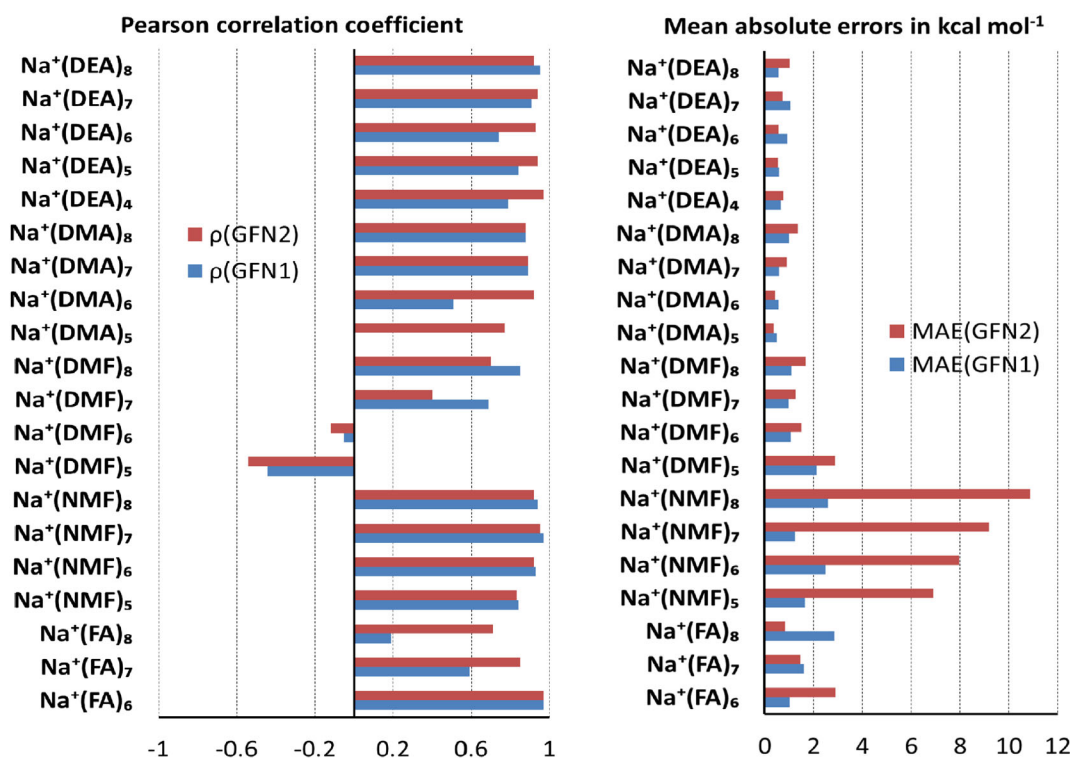
FIGURE 2 Pearson correlation coefficients (left) and mean absolute errors (right) between GFNn-xTB and reference (RI-MP2/CBS) conformational energies for sodium clusters with aprotic solvents (FA, NMF, DMF, DMA, and DEA)

the conformational sampling via the tight-binding semiempirical methods.

Second, inspecting the MAEs for the considered group of solvents, we find GFN2-xTB-predicted conformational energies of

$Na^+(NMF)_n$ clusters to deviate significantly from the reference values. The largest MAE of 10.9 kcal mol^{-1} for $Na(NMF)_8$ is especially alerting. The corresponding MSE of 10.6 kcal mol^{-1} reflects systematic overestimation of the conformational energies by GFN2-xTB for this
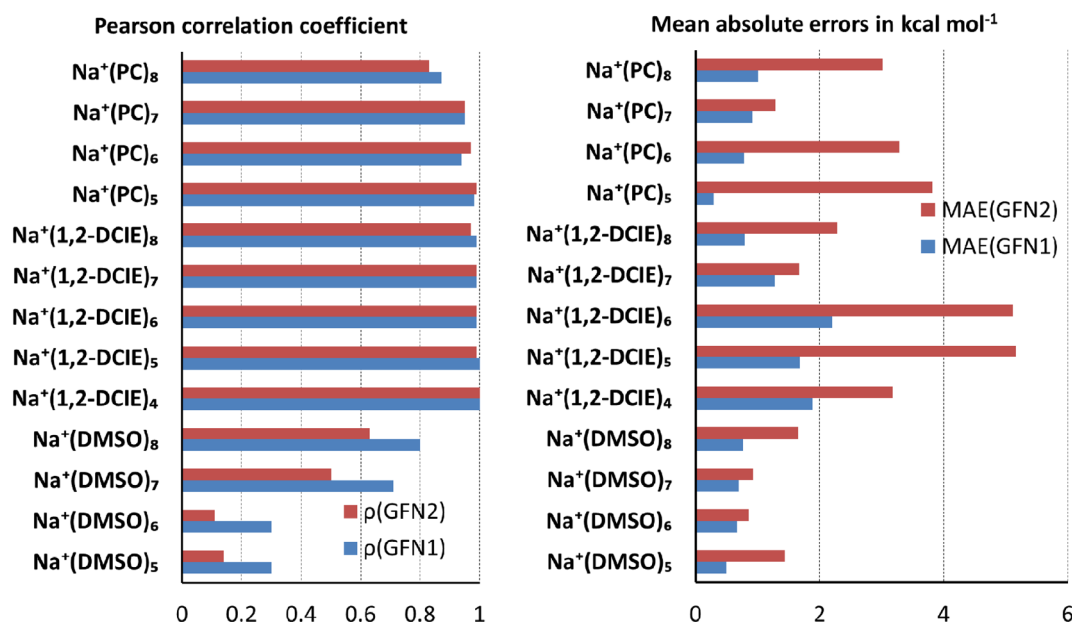
FIGURE 3 Pearson correlation coefficients (left) and mean absolute errors (right) between GFNn-xTB and reference (RI-MP2/CBS) conformational energies for sodium clusters with aprotic solvents (1,2-DCIE, DMSO, and PC)

cluster. It should however be noticed, that the conformational energies of the Na$^+$(NMF)$_n$ occupy the widest ranges of 13.2–23.4 kcal mol^{-1} for n = 5 and 8, respectively. At the same time, GFN1-xTB method demonstrates stable performance for Na$^+$(NMF)$_n$ systems with the average MAE of 2 kcal mol^{-1}. Additional internal test ascribes the difference in performance of the GFN1-xTB and GFN2-xTB approaches for this particular system to non-dispersion and non-repulsion components of the total energies. A number of changes were introduced when going from the GFN1-xTB to GFN2-xTB method, and not all of these can be considered as merely improvements. For example, in GFN2-xTB the basis set on hydrogen contains a single 1s function, while in its GFN1 predecessor hydrogen is described by two s-functions. Besides, instead of the classical force-field type corrections for hydrogen bonds used in GFN1-xTB, GFN2-xTB is proposed[15] to capture these interactions via multipole-extended electrostatics.

3.3 | 1,2-Dichloroethane, dimethyl sulfoxide, and propylene carbonate

For the second subset of aprotic solvents GFNn-xTB methods provide the lowest energy conformations for Na$^+$(PC)$_{5,6,7}$ and Na$^+$(1,2-DCIE)$_{4,6}$. For the other considered clusters with DMSO and PC, the most favorable structures predicted by the tight-binding methods are not more than 1.4 kcal mol^{-1} higher in the reference RI-MP2/CBS energies. The largest deviations were obtained for the GFN2-xTB method applied to the remaining (n = 5, 7, 8) clusters with 1,2-DCIE ($\Delta E_{\text{ref.}}$ = 3.8 kcal mol^{-1} for Na(1,2-DCIE)$_8$), while the GFN1-xTB-predicted lowest energy conformations are less than ca. 1 kcal mol^{-1} higher in their reference relative energies.

Pearson correlation coefficients for the considered subset indicate strong positive correlation for the conformational energies of the clusters with 1,2-DCIE and PC (Figure 3), but moderate ($\rho_{\text{avg.}}$ = 0.76/0.57) and low ($\rho_{\text{avg.}}$ = 0.30/0.13) for the larger (n = 7, 8) and smaller (n = 5, 6) clusters with DMSO, respectively. Again, low correlation for Na$^+$(DMSO)$_{5,6}$ can be attributed to the narrow conformational energy ranges (2.4 and 3.2 kcal mol^{-1}) for these systems. MAEs yielded for GFN2-xTB method and some clusters with PC and especially 1,2-DCIE reach quite large values, with the maximum of 5.1–5.2 kcal mol^{-1} in the case of Na(1,2-DCIE)$_{5,6}$. At the same time, MAEs calculated for the GFN1-xTB method do not exceed 1.9 kcal mol^{-1}.

3.4 | Performance of the DFT methods

Geometry optimizations of the conformations of Na$^+$S$_n$ clusters were performed with relatively cheap PBE/λ1 approximation. An appealing idea is to complement the energies of the optimized structures with the corresponding dispersion energy corrections –D3(BJ) calculated *a posteriori* and almost free of charge and use the resulting PBE-D3(BJ) energies for the conformational sampling. Examination of these energies against their reference RI-MP2/CBS counterparts leads to the following conclusions: (1) in most cases the low(est) energy conformations are successfully identified. However, several outliers, Na$^+$(DMSO)$_8$ and Na$^+$(PC)$_6$ with the reference relative energies of ca. 2 kcal mol^{-1} were also found; (2) MAEs in many cases are either comparable, or even larger than the corresponding values for the GFNn-xTB methods. Large (more than 4 kcal mol^{-1}) and systematically (as revealed by MSEs) overestimated conformational energies were yielded for Na$^+$(H$_2$O)$_{6,7}$, Na$^+$(MeOH)$_{7,8}$, Na$^+$(EtOH)$_8$, and all Na$^+$(NMF)$_{5-8}$ clusters.
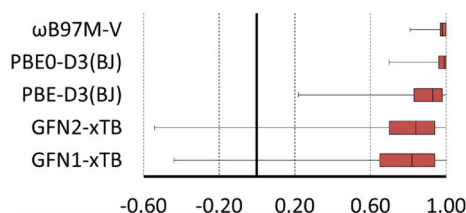
FIGURE 4 Pearson correlation coefficients (ρ) for the examined methods: The left and right sides of the boxes correspond to the first ($Q1$) and third ($Q3$) quartiles, respectively. The vertical solid line inside each box gives the median ρ value. The whiskers give the lowest and largest ρ values for each method

These findings suggest that in general relatively cheap DFT energies such as PBE-D3(BJ)/$\lambda 1$ can be used for the preliminary conformational sampling, that is, ranging of conformations of the Na^+S_n clusters by their energies, but the calculations of the ensemble properties, based on the absolute energy differences (e.g., Boltzmann distribution) should be performed at the higher levels of theory. For this reason, we extend the examination of the DFT conformational energies and involve higher-ranking functionals, hybrid GGA PBE0 complemented by Grimme's dispersion correction -D3(BJ) and range-separated hybrid meta-GGA ωB97M-V including VV10[70] nonlocal correlation to treat dispersion effects. Conformational energies from these methods were found to be much closer to the RI-MP2/CBS reference values. High median Pearson correlation coefficients of $\rho = 0.99$ (PBE0-D3(BJ)) and $\rho = 0.98$ (ωB97M-V) along with the very narrow interquartile ranges (Figure 4 and Table S4) confirm both DFT methods to provide an accurate conformational ranking. The mean MAE/MSE values averaged over all 45 considered clusters are 0.55/0.20 (PBE0-D3(BJ)) and 0.51/0.23 (ωB97M-V) kcal mol^{-1}. $Na^+(NMF)_n$ clusters still remain somewhat challenging for PBE0-D3 (BJ) method with the average MAE of 1.5 kcal mol^{-1}. At the same time, considerable average MAEs of ca. 1 kcal mol^{-1} were obtained for the ωB97M-V conformational energies of $Na^+(DMSO)_n$ and $Na^+(PC)_n$ systems.

3.5 | Recommendations

The overall results (Figure 4) obtained for the Na^+ clusters with protic and aprotic solvents show that fast GFNn-xTB methods are promising for the conformational sampling of such systems. Indeed, quite high median Pearson correlation coefficients ($\rho = 0.82/0.84$), acceptable average mean absolute errors (MAE = 1.2/2.3 kcal mol^{-1} and MSE = 0.2/1.6 kcal mol^{-1}) and correctly predicted low(est) energy structures for a number of clusters reflect a progress in the semiempirical methods.

As a fly in the ointment, there are some particular alerting aspects which should be taken into account before using these methods in a black-box fashion for the conformational sampling. First, we cannot claim the overall better performance of the more recent GFN2-xTB as compared to its GFN1-xTB predecessor for our benchmark set.

Second, particular systems ($Na^+(EtOH)_6$, $Na^+(PC)_n$, $Na^+(1,2-DCIE)_n$, and especially $Na^+(NMF)_n$) seem to be troublesome for at least one of the GFNn-xTB methods in either identification of the lowest energy structures, or large mean absolute errors against the reliable reference conformational energies. Third, low and even anti-correlations between the GFNn-xTB and reference conformational energies were observed for the clusters with different spatial structures, but close energies ($Na^+(DMF)_{5,6}$, $Na^+(DMA)_5$, and $Na^+(DMSO)_{5,6}$).

Conformational energies from the computationally much more demanding DFT methods examined in this work exhibit excellent correlation with their reference RI-MP2 counterparts and should be applied when higher accuracy is desirable.

4 | CONCLUSIONS

Contemporary semiempirical methods GFN1- and GFN2-xTB were examined to reproduce RI-MP2/CBS conformational energies of singly charged sodium clusters with 3 protic and 8 aprotic solvents. Comparing two versions of the GFNn-xTB method, we cannot claim the better performance of the more recent GFN2-xTB method (median Pearson correlation coefficient $\rho = 0.84$) over its GFN1-xTB predecessor ($\rho = 0.82$). Moreover, we find particular pronounced deviations of the GFN2-xTB conformational energies from the reference RI-MP2 values for $Na^+(PC)_n$, $Na^+(1,2-DCIE)_n$, and especially $Na^+(NMF)_n$ clusters, reaching, in terms of mean absolute error, 10.9 kcal mol^{-1} for $Na^+(NMF)_8$. On the optimistic side, tight-binding semiempirical methods correctly identify lowest energy structures for a number of the considered systems, including three clusters with propylene carbonate, the most common organic solvent used in lithium batteries. Fairly small mean absolute errors of 1.2/2.3 kcal mol^{-1} and mean signed errors of 0.2/1.6 kcal mol^{-1} were obtained for the GFN(1/2)-xTB methods applied to the whole set of 45 Na^+S_n clusters. Overall, these results suggest that the tight-binding GFNn-xTB methods are promising for the fast conformational sampling of the Na^+S_n clusters, but preliminary validation for a specific system is still a must.

Pure GGA PBE functional complemented by -D3(BJ) dispersion correction in conjunction with the double-ζ basis set ($\lambda 1$) yields conformational energies suitable for the preliminary sampling (median $\rho = 0.93$), but not for the reliable calculations of the ensemble properties such as Boltzmann distribution of the conformations due to considerably large mean absolute error of 1.7 kcal mol^{-1}. Higher-level single-point energy calculations involving hybrid PBE0-D3(BJ) or range-separated meta-hybrid ωB97M-V in conjunction with triple-ζ basis sets produce sufficiently more accurate conformational energies with the average MAE/MSE of 0.55/0.20 and 0.51/0.23 kcal mol^{-1}, respectively. The conformational ranking by these methods is in the excellent correlation with the reference RI-MP2 one, according to the high median Pearson correlation coefficients of $\rho = 0.99$ (PBE0-D3 (BJ)) and $\rho = 0.98$ (ωB97M-V).

The insights gained from this study may be of assistance to the further improvement of the contemporary semiempirical and

force-field methods and to the proper conformational sampling of the $A^{\pm}S_n$ clusters as an indispensible first step in the modeling of microsolvation. For the processes in the solution additional energy corrections introduced by implicit solvation models can also play a role in the determination of low-energy conformational ensembles.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Arseniy A. Otlyotov https://orcid.org/0000-0001-6974-6327
Yury Minenkov https://orcid.org/0000-0001-8993-056X

REFERENCES

[1] F. M. Bickelhaupt, E. J. Baerends, N. M. M. Nibbering, *Chem.—A Eur. J.* **1996**, *2*, 196.

[2] D. Feller, *J. Phys. Chem. A* **1997**, *101*, 2723. https://doi.org/10.1021/jp9700185

[3] S. M. Bachrach, *J. Phys. Chem. A* **2008**, *112*, 3722.

[4] A. Malloum, J. J. Fifen, J. Conradie, *Phys. Chem. Chem. Phys.* **2018**, *20*, 29184. https://doi.org/10.1039/c8cp05823g

[5] O. Boukar, J. J. Fifen, A. Malloum, Z. Dhaouadi, H. Ghalila, J. Conradie, *N. J. Chem.* **2019**, *43*, 9902. https://doi.org/10.1039/c9nj02462j

[6] T. E. Da-Yang, J. J. Fifen, A. Malloum, S. Lahmar, M. Nsangou, J. Conradie, *N. J. Chem.* **2020**, *44*, 3637. https://doi.org/10.1039/c9nj05169d

[7] A. Malloum, J. J. Fifen, J. Conradie, *J. Comput. Chem.* **2020**, *41*, 21. https://doi.org/10.1002/jcc.26071

[8] A. Malloum, J. Conradie, *Comput. Theor. Chem.* **2020**, *1191*, 113042.

[9] A. Malloum, J. Conradie, *Data Br.* **2021**, *37*, 107144. https://doi.org/10.1016/j.dib.2021.107144

[10] A. Malloum, J. Conradie, *J. Mol. Liq.* **2021**, *335*, 116032. https://doi.org/10.1016/j.molliq.2021.116032

[11] A. Malloum, J. Conradie, *Polyhedron* **2021**, *193*, 114856. https://doi.org/10.1016/j.poly.2020.114856

[12] S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 2847. https://doi.org/10.1021/acs.jctc.9b00143

[13] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169. https://doi.org/10.1039/c9cp06869d

[14] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989. https://doi.org/10.1021/acs.jctc.7b00118

[15] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652. https://doi.org/10.1021/acs.jctc.8b01176

[16] S. Spicher, S. Grimme, *Angew. Chem.—Int. Ed.* **2020**, *59*, 15665.

[17] S. Spicher, C. Plett, P. Pracht, A. Hansen, S. Grimme, *J. Chem. Theory Comput.* **2022**, *18*, 3174. https://doi.org/10.1021/acs.jctc.2c00239

[18] P. Pracht, S. Grimme, *Chem. Sci.* **2021**, *12*, 6551.

[19] J. Zhang, M. Dolg, *Phys. Chem. Chem. Phys.* **2015**, *17*, 24173. https://doi.org/10.1039/c5cp04060d

[20] J. Zhang, M. Dolg, *Phys. Chem. Chem. Phys.* **2016**, *18*, 3003. https://doi.org/10.1039/c5cp06313b

[21] J. Zhang, V. A. Glezakou, *Int. J. Quantum Chem.* **2021**, *121*, e26553. https://doi.org/10.1002/qua.26553

[22] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell, *J. Comput. Chem.* **2010**, *31*, 671.

[23] W. Yu, X. He, K. Vanommeslaeghe, A. D. MacKerell, *J. Comput. Chem.* **2012**, *33*, 2451.

[24] H. Li, J. Zhong, H. Vehkamäki, T. Kurtén, W. Wang, M. Ge, S. Zhang, Z. Li, X. Zhang, J. S. Francisco, X. C. Zeng, *J. Am. Chem. Soc.* **2018**, *140*, 11020. https://doi.org/10.1021/jacs.8b04928

[25] V. Besel, J. Kubečka, T. Kurtén, H. Vehkamäki, *J. Phys. Chem. A* **2020**, *124*, 5931. https://doi.org/10.1021/acs.jpca.0c03984

[26] J. Li, S. Zhou, J. Zhang, M. Schlangen, D. Usharani, S. Shaik, H. Schwarz, *J. Am. Chem. Soc.* **2016**, *138*, 11368. https://doi.org/10.1021/jacs.6b07246

[27] X. Cao, N. Heinz, J. Zhang, M. Dolg, *Phys. Chem. Chem. Phys.* **2017**, *19*, 20160. https://doi.org/10.1039/c7cp02861j

[28] Y. Liu, J. Yang, L. Cheng, *Inorg. Chem.* **2018**, *57*, 12934. https://doi.org/10.1021/acs.inorgchem.8b02159

[29] Y. Basdogan, J. A. Keith, *Chem. Sci.* **2018**, *9*, 5341.

[30] Y. Basdogan, M. C. Groenenboom, E. Henderson, S. De, S. B. Rempe, J. A. Keith, *J. Chem. Theory Comput.* **2020**, *16*, 633. https://doi.org/10.1021/acs.jctc.9b00605

[31] J. Zhang, E. T. Baxter, M. T. Nguyen, V. Prabhakaran, R. Rousseau, G. E. Johnson, V. A. Glezakou, *J. Phys. Chem. Lett.* **2020**, *11*, 6844. https://doi.org/10.1021/acs.jpclett.0c01671

[32] J. J. P. Stewart, *J. Mol. Model.* **2007**, *13*, 1173.

[33] J. J. P. Stewart, *J. Mol. Model.* **2013**, *19*, 1.

[34] D. I. Sharapa, A. Genaev, L. Cavallo, Y. Minenkov, *ChemPhysChem* **2019**, *20*, 92. https://doi.org/10.1002/cphc.201801063

[35] M. Bursch, A. Hansen, P. Pracht, J. T. Kohn, S. Grimme, *Phys. Chem. Chem. Phys.* **2021**, *23*, 287. https://doi.org/10.1039/d0cp04696e

[36] A. Otlyotov, A. Moshchenkov, L. Cavallo, Y. Minenkov, *Phys. Chem. Chem. Phys.* **2022**, *24*, 17314. https://doi.org/10.1039/d2cp01659a

[37] Y. Marcus, *Ion Solvation in Neat Solvents*, John Wiley & Sons, Hoboken, New Jersey **2015**, p. 107.

[38] J. R. Pliego, J. M. Riveros, *J. Phys. Chem. A* **2001**, *105*, 7241. https://doi.org/10.1021/jp004192w

[39] N. F. Carvalho, J. R. Pliego, *Phys. Chem. Chem. Phys.* **2015**, *17*, 26745. https://doi.org/10.1039/c5cp03798k

[40] J. R. Pliego, J. M. Riveros, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2020**, *10*, e1440. https://doi.org/10.1002/wcms.1440

[41] D. N. Laikov, Y. A. Ustynyuk, *Russ. Chem. Bull.* **2005**, *54*, 820. https://doi.org/10.1007/s11172-005-0329-x

[42] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865. https://doi.org/10.1103/PhysRevLett.77.3865

[43] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1997**, *78*, 1396. https://doi.org/10.1103/PhysRevLett.78.1396

[44] D. N. Laikov, *Chem. Phys. Lett.* **2005**, *416*, 116. https://doi.org/10.1016/j.cplett.2005.09.046

[45] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104. https://doi.org/10.1063/1.3382344

[46] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456. https://doi.org/10.1002/jcc.21759

[47] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2021**, *11*, e1493. https://doi.org/10.1002/wcms.1493

[48] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618. https://doi.org/10.1103/PhysRev.46.618

[49] M. Feyereisen, G. Fitzgerald, A. Komornicki, *Chem. Phys. Lett.* **1993**, *208*, 359. https://doi.org/10.1016/0009-2614(93)87156-W

[50] F. Weigend, M. Häser, *Theor. Chem. Acc.* **1997**, *97*, 331. https://doi.org/10.1007/s002140050269

[51] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Chem. Phys. Lett.* **1998**, *294*, 143. https://doi.org/10.1016/S0009-2614(98)00862-8

[52] F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73. https://doi.org/10.1002/wcms.81

[53] F. Neese, *WIREs Comput. Mol. Sci.* **2022**, e1606. https://doi.org/10.1002/wcms.1606

[54] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007. https://doi.org/10.1063/1.456153

[55] D. E. Woon, T. H. Dunning, *J. Chem. Phys.* **1993**, *98*, 1358. https://doi.org/10.1063/1.464303

[56] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, A. K. Wilson, *Theor. Chem. Acc.* **2011**, *128*, 69.

[57] A. Schulz, B. J. Smith, L. Radom, *J. Phys. Chem. A* **1999**, *103*, 7522. https://doi.org/10.1021/jp991577+

[58] M. B. Sullivan, M. A. Iron, P. C. Redfern, J. M. L. Martin, L. A. Curtiss, L. Radom, *J. Phys. Chem. A* **2003**, *107*, 5617. https://doi.org/10.1021/jp034851f

[59] M. A. Iron, M. Oren, J. M. L. Martin, *Mol. Phys.* **2003**, *101*, 1345. https://doi.org/10.1080/0026897031000094498

[60] Y. Minenkov, G. Bistoni, C. Riplinger, A. A. Auer, F. Neese, L. Cavallo, *Phys. Chem. Chem. Phys.* **2017**, *19*, 9374. https://doi.org/10.1039/c7cp00836h

[61] F. Weigend, A. Köhn, C. Hättig, *J. Chem. Phys.* **2002**, *116*, 3175. https://doi.org/10.1063/1.1445115

[62] G. L. Stoychev, A. A. Auer, F. Neese, *J. Chem. Theory Comput.* **2017**, *13*, 554. https://doi.org/10.1021/acs.jctc.6b01041

[63] F. Neese, F. Wennmohs, A. Hansen, U. Becker, *Chem. Phys.* **2009**, *356*, 98. https://doi.org/10.1016/j.chemphys.2008.10.036

[64] J. M. L. Martin, *Chem. Phys. Lett.* **1996**, *259*, 669. https://doi.org/10.1016/0009-2614(96)00898-6

[65] C. Riplinger, F. Neese, *J. Chem. Phys.* **2013**, *138*, 34106. https://doi.org/10.1063/1.4773581

[66] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101. https://doi.org/10.1063/1.4821834

[67] C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, F. Neese, *J. Chem. Phys.* **2016**, *144*, 24109. https://doi.org/10.1063/1.4939030

[68] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158. https://doi.org/10.1063/1.478522

[69] N. Mardirossian, M. Head-Gordon, *J. Chem. Phys.* **2016**, *144*, 214110. https://doi.org/10.1063/1.4952647

[70] O. A. Vydrov, T. Van Voorhis, *J. Chem. Phys.* **2010**, *133*, 244103. https://doi.org/10.1063/1.3521275

[71] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297. https://doi.org/10.1039/b508541a

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: A. A. Otlyotov, Y. Minenkov, *J. Comput. Chem.* **2022**, *43*(27), 1856. https://doi.org/10.1002/jcc.26988