

# Parametrization of the SCC-DFTB Method for Halogens

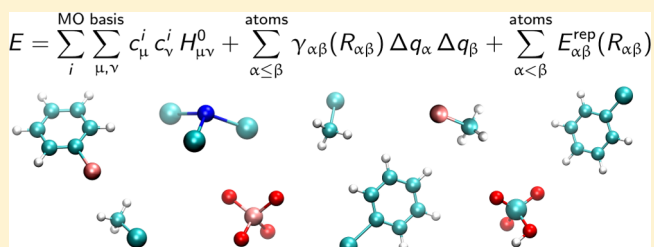
Tomáš Kubař,<sup>†</sup> Zoltán Bodrog,<sup>‡,§</sup> Michael Gaus,<sup>†</sup> Christof Köhler,<sup>‡</sup> Bálint Aradi,<sup>‡</sup> Thomas Frauenheim,<sup>‡</sup> and Marcus Elstner<sup>\*,†</sup>

<sup>†</sup>Institute of Physical Chemistry, Karlsruhe Institute of Technology, Kaiserstr. 12, 76131 Karlsruhe, Germany

<sup>‡</sup>Bremen Center for Computational Materials Science, University of Bremen, Am Fallturm 1, 28359 Bremen, Germany

## S Supporting Information

**ABSTRACT:** Parametrization of the approximative DFT method SCC-DFTB for halogen elements is presented. The new parameter set is intended to describe halogenated organic as well as inorganic molecules, and it is compatible with the established parametrization of SCC-DFTB for carbon, hydrogen, oxygen, and nitrogen. The performance of the parameter set is tested on a representative set of molecules and discussed.



## 1. INTRODUCTION

Semiempirical quantum chemical methods (SE)<sup>1–3</sup> represent a suitable compromise to describe complex molecular systems. Their favorable computational efficiency makes it possible to perform calculations of considerably larger systems than with ab initio wave function methods or density functional theory (DFT). On the other hand, once parametrized duly and benchmarked, they do not suffer from the limitations inherent to molecular mechanics methods, such as the inability to describe chemical reactions. SE methods found a wide range of applications in chemistry, from clusters of small molecules to synthetic polymers and biomolecular complexes.

Most of today's successful SE methods are based on the formalism of neglect of diatomic differential overlap approximation (NDDO) to the Hartree–Fock method<sup>4</sup> and on the early MNDO model involving a minimal basis set for the valence electron shells.<sup>5</sup> Several methods were developed and have been used frequently, including but not limited to the following. AM1<sup>6</sup> involves additional off-center attractive or repulsive functions, reducing the too high repulsion of closely approaching atoms; this makes a larger number of parameters necessary. AM1 was also reparametrized for organic molecules containing halogen atoms, giving the RM1 method.<sup>7</sup> The other most frequently used SE method is PM3,<sup>8</sup> with Hamiltonian very similar to that of AM1, the difference being the rather statistical approach to parametrization of PM3. Contrary to that, AM1 considers some of the parameters to possess physical meaning and estimates them on the basis of spectroscopy data. The more recent parametrization PM6<sup>9</sup> emphasizes biochemical systems somewhat. A slightly different class of methods are represented by OM2 and OM3,<sup>10</sup> which include orthogonalization corrections and approximated Pauli exchange–repulsion terms in the Fock matrix.

Unlike the SE framework, the self consistent charge density functional based tight-binding (SCC-DFTB) method,<sup>11</sup> which is the subject of this work, is derived from DFT. The

description of SCC-DFTB will follow in due detail in the Methodology section; here, we would like to mention several points that occurred in the discussion of SCC-DFTB and its comparison with SE methods by several authors.<sup>12–14</sup> There are undoubtedly formal similarities between SCC-DFTB and SE, for instance the DFTB Hamilton matrix may be likened to the Fock matrix. On the other hand, SCC-DFTB includes electron correlation explicitly by construction because of its DFT origin and involves an explicit overlap matrix while most popular SE methods are orthogonal (except OMx). Notably, two-electron integrals are a perturbation due to the differential electron density in SCC-DFTB, while they cover the complete electron–electron interaction in NDDO. So, a straightforward relation of DFTB to SE is not quite possible, as the different foundations provide the similar-looking terms with a different meaning.

As for the performance of SCC-DFTB and its relative merit with respect to SE methods, some general observations have been made. Thiel et al.<sup>13</sup> state that “errors in SE calculations are normally less uniform and thus harder to correct”, while the situation with SCC-DFTB is more favorable, meaning that it is rather predictable.<sup>13</sup> A further characteristic of SE methods is that “there are elements that are more ‘difficult’ than others”, and this is possibly similar in the case of SCC-DFTB. Also, while it is true in a certain sense about SCC-DFTB that “different parametrizations... may be required for different properties to obtain useful accuracy...”, it is perhaps not so difficult to find a single parametrization that works well for biomolecular complexes. (Still, a different parametrization would be required for solid state, for instance.) Jorgensen et al.<sup>12</sup> raised the issue of parametrization of SCC-DFTB: “Other issues that make SCC-DFTB less attractive at this time than PM3 are the lack of parameters for additional atom types

Received: March 10, 2013

including halogens, the larger number of parameters in the core repulsion formulas, and their quadratic rather than linear growth with addition of new atom types.<sup>12</sup> Surely, the parametrization effort is  $O(N^2)$  for  $N$  chemical elements as pairwise parameter sets are needed. The large freedom in the repulsive parametrization (expressed in terms of many parameters to determine) may be seen in a positive sense—there is a lot of freedom in the parametrization process that may be used to create good parameters: while it is indeed quite involved to determine the repulsive potentials carefully, these represent to a considerable extent the basis for the good performance of SCC-DFTB for molecular systems, which was established, for example, in ref 15. Also, automated procedures for the parametrization of repulsive potentials were developed.<sup>16,17</sup> Finally, the parameters for halogens were truly missing, and it is this work that aims to fill the gap. We note that a parametrization of fluorine is available already.<sup>18</sup>

Indeed, the considerable interest in interactions of halogenated molecules with biomolecular complexes<sup>19,20</sup> makes a parametrization of SE methods for halogen compounds topical. As mentioned, SCC-DFTB has missed such a parametrization so far, unlike most of the SE methods introduced above. Therefore, the aim of the current work is the parametrization of SCC-DFTB for compounds of chlorine, bromine, and iodine, considering organic molecules with halogen atoms bound to carbon atoms as well as inorganic molecules (dihalogen molecules, hydrogen halides, oxoacids, and oxoanions). This work will be compatible with the well established original parametrization of SCC-DFTB for organic molecules (called MIO),<sup>11</sup> rather than with the newer development of the extended variant DFTB3<sup>21</sup> and the corresponding parameters (called 3OB).<sup>22</sup> To assess the performance of the parametrization, the obtained parameter sets will be applied to a representative set of molecules containing the most frequently occurring binding motifs.

## 2. METHODOLOGY

**2.1. Self Consistent Charge Density Functional Tight-Binding.** The SCC-DFTB energy expression is derived from a second-order Taylor expansion of exact DFT energy with respect to the differential electron density  $\delta\rho$  relative to a given reference density  $\rho'_0 \equiv (\rho_0(\vec{r}'), \int' \equiv \int d\vec{r}')$ :<sup>11</sup>

$$E = \sum_i^{\text{occ}} \langle \Psi_i | \hat{H}^0 | \Psi_i \rangle + \frac{1}{2} \int \int' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{\text{XC}}}{\delta\rho\delta\rho'} \right) \Delta\rho\Delta\rho' - \frac{1}{2} \int \int' \frac{\rho'_0 \rho_0}{|\vec{r} - \vec{r}'|} + E_{\text{XC}}[\rho_0] - \int V_{\text{XC}}[\rho_0] \rho_0 + E_{\text{cc}} \quad (1)$$

A minimal basis set consisting of atomic-orbital-like functions  $\eta_\mu$  is introduced, and the LCAO wave function is  $\Psi_i = \sum c_\mu^i \eta_\mu$ . Then, the first term, representing the sum of the Kohn–Sham eigenvalues, becomes

$$\sum_i^{\text{occ}} \langle \Psi_i | \hat{H}^0 | \Psi_i \rangle = \sum_i^{\text{occ}} \sum_\mu \sum_\nu c_\mu^i c_\nu^i H_{\mu\nu}^0 \quad (2)$$

The DFTB Hamilton matrix elements for the state with reference electron density

$$H_{\mu\nu}^0 = \langle \eta_\mu | \hat{H}[\rho_0] | \eta_\nu \rangle \quad (3)$$

are subject to several approximations; in particular, a two-center approximation is applied, as discussed by Seifert.<sup>23</sup>  $H_{\mu\nu}^0$  and the

overlap matrix elements  $S_{\mu\nu} = \langle \eta_\mu | \eta_\nu \rangle$  are calculated once with a DFT program and are tabulated. This procedure will be described in some detail in a dedicated section.

The last four terms in eq 1 depend only on the reference density  $\rho_0$  and represent the DFT double-counting contributions and core–core repulsion energy contributions. They are approximated with a sum  $E_{\text{rep}}^{\text{rep}}$  of repulsive pairwise potentials (between atoms  $\alpha$  and  $\beta$ )  $E_{\alpha\beta}^{\text{rep}}$ .<sup>23</sup>

$$E_{\text{rep}} = \sum_{\alpha < \beta}^{\text{atoms}} E_{\alpha\beta}^{\text{rep}}(R_{\alpha\beta}) \quad (4)$$

The second-order term in the differential charge density  $\Delta\rho$  (second term in eq 1) is approximated by writing  $\Delta\rho$  as a superposition of atomic contributions

$$\Delta\rho = \sum_\alpha \Delta\rho_\alpha \quad (5)$$

and applying a monopole approximation for the differential density on atom  $\alpha$ , which is assumed to be spherical:  $\Delta\rho_\alpha \approx \Delta q_\alpha F_{00}^\alpha Y_{00}$ . Here,  $F_{00}^\alpha$  denotes the normalized radial dependence of the differential density, and  $Y_{00}$  is the  $l = 0$  real spherical harmonic. The second-order part becomes

$$E_{2\text{nd}} \approx \frac{1}{2} \sum_{\alpha\beta} \Delta q_\alpha \Delta q_\beta \int \int' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{\text{XC}}}{\delta\rho\delta\rho'} \right) F_{00}^\alpha F_{00}^\beta Y_{00}^2 \quad (6)$$

The quite simple shape of this function becomes apparent by noting the following: (i) For large distances, the XC terms vanish and the integral describes merely the Coulomb interaction of two spherical normalized charge densities, which reduces to  $\Delta q_\alpha \Delta q_\beta / R_{\alpha\beta}$ . (ii) For vanishing interatomic distance  $R_{\alpha\beta} \rightarrow 0$ , the integral describes the electron–electron interaction on a single atom  $\alpha$  and is approximated as  $1/2 U_\alpha \Delta q_\alpha^2$  with  $U_\alpha$  being the Hubbard parameter or the chemical hardness. For SCC-DFTB,  $U_\alpha$  of nontransition elements is calculated by taking the first derivative of the energy of the highest occupied molecular orbital of an isolated atom with respect to the charge. (iii) The intermediate region of chemical bonding (1–3 Å) is not covered by these limits. A function  $\gamma(R)$  is used to interpolate between the limits, similarly as in SE methods. There, the interpolating function has a simple form such as the Klopman–Ohno or Mataga–Nishimoto functions. SCC-DFTB uses rather an analytic expression derived by the integration of eq 6

$$\gamma_{\alpha\beta}(R) = \frac{1}{R} - \left( e^{-\tau_\alpha R} \left( \frac{\tau_\beta^4 \tau_\alpha}{2(\tau_\alpha^2 - \tau_\beta^2)^2} - \frac{\tau_\beta^6 - 3\tau_\beta^4 \tau_\alpha^2}{(\tau_\alpha^2 - \tau_\beta^2)^3 R} \right) + e^{-\tau_\beta R} \left( \frac{\tau_\alpha^4 \tau_\beta}{2(\tau_\alpha^2 - \tau_\beta^2)^2} - \frac{\tau_\alpha^6 - 3\tau_\alpha^4 \tau_\beta^2}{(\tau_\beta^2 - \tau_\alpha^2)^3 R} \right) \right) \quad (7)$$

with  $\tau_\alpha = 16/5 U_\alpha$ .<sup>11</sup> Then, the second-order term becomes a simple pairwise expression depending on atomic-like charges from the Mulliken population analysis.

With these definitions and approximations, the SCC-DFTB energy reads

$$E^{\text{SCC-DFTB}} = \sum_{i\mu\nu} c_{\mu}^i c_{\nu}^i H_{\mu\nu}^0 + \frac{1}{2} \sum_{\alpha\beta} \gamma[U_{\alpha}, U_{\beta}](R_{\alpha\beta}) \Delta q_{\alpha} \Delta q_{\beta} + \frac{1}{2} \sum_{\alpha\beta} E_{\alpha\beta}^{\text{rep}}(R_{\alpha\beta}) \quad (8)$$

The variational principle leads to a generalized eigenvalue problem, and the diagonalization of the Hamilton matrix is the main computational cost of SCC-DFTB calculation. The variational problem has to be solved iteratively for the wave function expansion coefficients  $c_{\mu}^i$  because the Hamilton matrix elements depend on  $c_{\mu}^i$  by way of the Mulliken charges (therefore, 'self consistent charge DFTB'; function  $\mu$  on atom  $\alpha$ , function  $\nu$  on atom  $\beta$ , the sum runs over all of the atoms  $\delta$  in the system):

$$H_{\mu\nu}^{\alpha\beta} = H_{\mu\nu}^{0\alpha\beta} + \frac{1}{2} S_{\mu\nu}^{\alpha\beta} \sum_{\delta} \Delta q_{\delta} (\gamma_{\alpha\delta} + \gamma_{\beta\delta}) \quad (9)$$

In principle, the two-body contributions  $E_{\alpha\beta}^{\text{rep}}(R)$  can be determined by comparison of the energy according to eq 8 with that from full DFT calculations, as function of the interatomic distance  $R_{\alpha\beta}$ .<sup>23</sup> Then, the resulting energy curve  $E_{\alpha\beta}^{\text{rep}}$  is represented by splines. The determination of repulsive potentials for organic molecules was reviewed briefly.<sup>24</sup>

Also, SCC-DFTB was extended to treat open-shell electronic systems,<sup>25–27</sup> with the following expression for total energy (analogue of eq 8 in standard SCC-DFTB):

$$E^{\text{spin-pol}} = \sum_{\sigma=\uparrow,\downarrow} \sum_i n_{i\sigma} \sum_{\mu\nu} c_{\mu}^{i\sigma} c_{\nu}^{i\sigma} H_{\mu\nu}^0 + \frac{1}{2} \sum_{\alpha\beta} \sum_{l \in \alpha} \sum_{l' \in \beta} \gamma[U_{\alpha l}, U_{\beta l'}](R_{\alpha\beta}) \Delta q_{\alpha l} \Delta q_{\beta l'} + \frac{1}{2} \sum_{\alpha} \sum_{l \in \alpha} \sum_{l' \in \alpha} W_{all'} p_{\alpha l} p_{\alpha l'} + \frac{1}{2} \sum_{\alpha\beta} E_{\alpha\beta}^{\text{rep}}(R_{\alpha\beta}) \quad (10)$$

The molecular orbitals  $i$  are different for spin-up and spin-down electrons, and the differential density is expressed in terms of contributions from the shells with angular momenta  $l, l'$  on atoms  $\alpha, \beta$  in the form of shell-resolved Mulliken charges  $\Delta q_{\alpha l}$ . The third term involves differences between spin-up and spin-down Mulliken populations of the individual shells  $l$  on atoms  $\alpha, p_{\alpha l}$ . The spin constants  $W_{all'}$  have to be determined for each chemical element.

The methodology as well as a range of applications of SCC-DFTB were reviewed comprehensively.<sup>28</sup> The DFTB+ software suite represents a modern implementation of SCC-DFTB.<sup>29</sup>

**2.2. SCC-DFTB Parameters.** The following parameters are necessary to perform SCC-DFTB calculations. The atomic parameters are the energies of  $s$  and  $p$  orbitals, the Hubbard parameter  $U$ , and, optionally, the spin-polarization parameters  $W$  or the spin-polarization energy. They are obtained from a DFT calculation on a single atom with the PBE exchange–correlation functional<sup>30</sup> solving the Schrödinger equation

$$(T + V_{\text{nuc}} + V_{\text{Har}} + V_{\text{XC}}) \varphi_i = \varepsilon_i \varphi_i \quad (11)$$

The Hamiltonian contains the kinetic energy  $T$ , the electron–nucleus attraction  $V_{\text{nuc}}$ , the Hartree energy  $V_{\text{Har}}$ , and the exchange–correlation contribution  $V_{\text{XC}}$ . The form of the Slater-type basis set is shown in the Supporting Information.

The calculation of a neutral atom in the (possibly hypothetical) singlet state yields the orbital energies directly.

To obtain the Hubbard parameter  $U$ , a series of calculations is performed on a single atom with the charge varying for instance from  $-0.2$  e to  $+0.2$  e, with a step of  $0.02$  e. The energy of the atom as a function of charge is then analyzed with quadratic regression, and  $U$  is obtained as twice the quadratic coefficient. The simpler procedure used to obtain  $U$  in the previous parametrizations<sup>11,18</sup> is consistent with the current procedure.

The spin-polarization constants  $W$  needed for calculations of open-shell systems are obtained from DFT calculations of individual atoms by means of one-side numerical differentiation of the single-particle energies with respect to the orbital occupation numbers; see ref 27 for details. The step width was  $0.01$  e, and the occupation numbers were reduced from  $2.5$  to  $2.49$  and from  $1$  to  $0.99$  for the  $p$  and  $s$  orbitals, respectively, in the calculation on each of the halogen atoms. Since the absolute values of  $W$  are small and thus the contribution to the DFTB total energy is small, too, the simple parametrization procedure (used in the previous work as well)<sup>27</sup> should be sufficiently accurate for most purposes.

When the atomization energies of molecules are of interest, it is desirable to correct the energies of atoms for the correct spin state. To this end, the spin-polarization energy is calculated. In a SCC-DFTB calculation, it is an a posteriori correction, which does not enter the Hamiltonian and does not affect any observables but the atomization energy. In the parametrization, it is evaluated as the difference of energy of two different states of the respective atom: the hypothetical, spin-unpolarized singlet state and the genuine, spin-polarized state. In the case of halogens, this is a doublet with 3 spin-up electrons and 2 spin-down electrons. For sake of consistency with the previous parametrization,<sup>11</sup> the spin-polarization energy is obtained with the LDA functional.<sup>31</sup>

For the parametrization of heavier chemical elements such as iodine and possibly bromine, an account for relativistic effects is desirable. In this work, the DFT calculations of bromine and iodine atoms are performed with a relativistic Hamiltonian for the kinetic energy by means of a new scaled scalar-relativistic ZORA<sup>32,33</sup> implementation with a point-charge nucleus, accounting for the relativistic kinetic energy and for the Darwin shift (which applies to  $s$  orbitals). We note that DFTB calculations involving a (different) relativistic treatment were performed already.<sup>34,35</sup>

Two-center electronic parameters are the Hamiltonian element and the overlap of two atom-centered basis functions. These parameters are obtained in accordance with the proposal for tight-binding method<sup>36</sup> as follows. First, the electron density  $\rho_0$  of individual atoms is evaluated by means of DFT calculations with the Schrödinger equation

$$\left( T + V_{\text{nuc}} + V_{\text{Har}} + V_{\text{XC}} + \left( \frac{r}{r_{\text{dens}}} \right)^2 \right) \varphi_i = \varepsilon_i \varphi_i \quad (12)$$

The Hamiltonian contains an additional term that compresses the density in space effectively, and  $r_{\text{dens}}$  represents the compression radius for the atomic density. The resulting density is discretized on a 3D grid. Then, the basis functions  $\eta_{\mu}$  centered on the atoms are obtained from a similar Schrödinger equation

$$\left( T + V_{\text{nuc}} + V_{\text{Har}} + V_{\text{XC}} + \left( \frac{r}{r_{\text{wf}}} \right)^2 \right) \eta_{\mu} = \varepsilon_{\mu} \eta_{\mu} \quad (13)$$



with a compression radius  $r_{wf}$  (different from  $r_{dens}$ ) and evaluated on the same 3D grid as the density. Finally, the two-center Slater–Koster parameters<sup>37</sup> for functions  $\mu$  and  $\nu$  located on atoms  $a$  and  $b$ , respectively, are computed numerically as

$$S_{\mu\nu} = \int \eta_{\mu} \eta_{\nu} dV \quad (14)$$

$$H_{\mu\nu} = \int \eta_{\mu} \hat{H} [\rho_a^a] + \rho_b^b [\eta_{\nu}] dV \quad (15)$$

where the Hamiltonian is constructed with the superposition of atomic densities.

The two-center repulsive potentials are to be generated in order to fit the structures, vibrational frequencies and atomization energies of selected simple molecules to a suitable reference. The philosophy of SCC-DFTB is to cut off each repulsive potential at a distance below the region of second-nearest-neighbor interactions. While this treatment prevents the repulsive potentials from affecting longer-distance interactions unexpectedly, thus improving numerical stability, it is not always possible to obtain repulsive functions that would yield satisfactory dissociation curves. To correct for this deficiency, the repulsive energy functions are often shifted down to some extent. This results in overestimated atomization energies of molecules containing the affected bond and will be referred to as overbinding in the following. We note that the previous parametrization of SCC-DFTB<sup>11</sup> features a certain degree of overbinding, too; this is reflected, for example, by not quite accurate prediction of heats of formation by SCC-DFTB.<sup>12</sup>

### 3. PARAMETRIZATION OF DFTB FOR HALOGENS

**3.1. Parameters for Fluorine.** DFTB parameters for fluorine and its interactions with other atomic species were developed already.<sup>18,38,39</sup> The original application was aimed at the mechanical and electronic properties of various fluorinated carbon allotropes (i.e., diamond, graphite, and nanotubes).<sup>38,39</sup> The non-SCC formulation of DFTB was used here. Later, while retaining the F–F interaction parameters, a new F–C interaction, together with additional heteronuclear interactions (F–O, F–N, F–H and F–Si), was generated for SCC-DFTB<sup>18,40</sup> and applied to fluorination reactions at SiN and SiO surface and concomitant gas-phase reactions. While special care had been taken to describe the relative reaction energetics between C–F and C–Si bond formation,<sup>18</sup> the description of the F<sub>2</sub> molecule in general, as well as an accurate description of the interactions F–O, F–N, and F–H, was not needed and therefore not in the focus of this previous work.

To retain the continuity of the SCC-DFTB parametrizations and avoid possible ambiguity as much as possible, the interactions of fluorine are not reparametrized in this work. Instead, the performance of the existing parametrization with respect to molecular systems of interest will be assessed, and any possible points of deficiency will be pointed out.

The one-center parameters for F are shown in Table 1. To illustrate the performance of the parameter sets, the structure, vibrational frequencies, and atomization energies of several fluorinated organic molecules were obtained and compared with experimental or DFT reference; see Table 2 and the Supporting Information. These DFT calculations as well as those following in this work were performed with *Gaussian* 03,<sup>41</sup> except where otherwise stated.

**Table 1. Electronic Parameters for Fluorine (in au) from Ref 18**

<i>s</i> -orbital energy	−1.095847
<i>p</i> -orbital energy	−0.408705
Hubbard parameter	0.5924
spin-pol. energy	−0.0150

When concentrating on halogenated organic molecules, the most important bond is the C–F bond, which was parametrized based on the fluoromethane molecule only. Consequently, the C–F bond length as well as vibration frequency is reproduced accurately with this parametrization, at the cost of a mild overbinding of 6 kcal/mol. A similar overbinding is present for the F–H bond of the HF molecule as well, while a large overbinding in the F<sub>2</sub> molecules is apparent, as compared to the other halogen dimer molecules with the current parametrization (see Table 8).

Actually, the parametrization of fluorine involves two slightly different values of Hubbard parameter  $U$ , which are suited for use in a modified variant of the SCC-DFTB method: 0.5924 and 0.5586 au for *s* and *p* orbitals, respectively. In standard SCC-DFTB calculations, as described in this work, the value for the *s* orbital is used; note that this behavior is undesired, as the occupation of *p* orbital varies in a fluorine atom upon charge transfer. We tested the dependence of the obtained results on the choice of  $U$  (that for *s* orbital or that for *p* orbital) in standard SCC-DFTB calculations. The maximum deviations were 0.004 Å for bond lengths (in (CH<sub>3</sub>)<sub>3</sub>CF), 0.2° for angles (in CH<sub>3</sub>F), 7 cm<sup>−1</sup> for vibrational frequencies (in HF), and 1.3 kcal/mol for atomization energies (in HF). From this we conclude that the exact choice of the Hubbard parameter  $U$  has a negligible overall effect.

**3.2. Electronic Parameters for Cl, Br, and I.** The atomic parameters were generated in a way that is consistent with the MIO parametrization of SCC-DFTB for the elements composing organic molecules.<sup>11</sup> The atomic parameters were obtained with a DFT calculation on a single atom of the respective halogen, which were performed with the PBE exchange–correlation functional.<sup>30</sup> The usual DFTB atomic reference state of a hypothetical spin-unpolarized spherical symmetric pseudoatom is assumed. Note that this state of a halogen atom requires the valence *p* orbitals to be filled with 2<sup>1/2</sup> spin-up electrons and 2<sup>1/2</sup> spin-down electrons. The atomic parameters are presented in Table 3, and the spin-polarization constants  $W$  are presented in Table 4 for all of the halogens. The exponents of the Slater-type basis set functions used in the DFT calculations are shown in the Supporting Information.

Calculations of bromine and iodine atoms were performed with a relativistic Hamiltonian. Since the current work is one of the first applications of this implementation, the results needed to be compared to suitable reference data. The valence orbital energies (which are prominently influenced by relativistic effects) obtained in this work and those from the Atomic Reference Data (ARD) database<sup>42</sup> are shown in Table 5. To make a comparison possible, these orbital energies were calculated with a Perdew–Wang LDA functional,<sup>31</sup> as no GGA data are available in the ARD. The orbital energies calculated in the current work agree well with the reference; the deviation is under 0.001 au for the *s*-orbital energy of Br atom and under 0.0002 au for all of the other orbital energies.

Table 2. Basic Characterization of the Parametrization of Fluorine<sup>a</sup>

molecule	bond	length		atomization energy			vib. frequency		
		ref.	DFTB	ref.	DFTB	overbinding	exp.	PBE	DFTB
CH <sub>3</sub> F	C–F	1.386	1.391	422.1	440.9	5.8	1049	1037	1056
F <sub>2</sub>	F–F	1.398	1.467	37.7	72.2	34.5	917	1000	936
HF	H–F	0.922	0.934	141.2	146.8	5.6	4138	3982	3867

<sup>a</sup>Presented are similar data as for the parametrization of other halogens. Bond length in angstrom (Å), atomization energies in kcal/mol, and vibrational frequencies in cm<sup>−1</sup>; the values of overbinding are related to the respective bonds.

Table 3. Atomic Electronic Parameters for Cl, Br, and I (in au)

element	s-orbital energy	p-orbital energy	Hubbard parameter	spin-pol. energy
Cl	−0.754424	−0.316395	0.3668	−0.0072
Br	−0.739422	−0.288802	0.3277	−0.0059
I	−0.638239	−0.260627	0.2842	−0.0047

In the computation of two-center Slater–Koster electronic parameters, the compression radii for the wave functions  $r_{\text{wf}}$  were chosen as twice the covalent radii. The compression radii for the density  $r_{\text{dens}}$  were expanded from the value of 9 taken for fluorine previously.<sup>18</sup> The radii are shown in Table 6. The two-center Slater–Koster parameters were evaluated for interatomic distances of up to 14–28 au with a step of 0.02 au. Note that while the wave functions entering the calculation as well as the density that the Hamiltonian depends on had been obtained with a relativistic treatment for Br and I, the Hamiltonian itself is nonrelativistic in the calculation of two-center parameters.

### 3.3. Two-Center Repulsive Potentials for Cl, Br, and I.

The repulsive potentials were obtained with cutoff distances kept below the typical 1–3 interactions; these are listed in the Supporting Information. The protocol used to obtain the repulsive functions for the distinct types of bonds is described in the following sections.

**3.3.1. C–Cl, C–Br, and C–I.**  $E_{\text{rep}}$  was fitted to the structure (C–X bond length), atomization energy, and stretching vibrational frequency of halogenmethane molecules. For these particular interactions, the function takes the form of a sequence of three splines of fourth order, starting at the distance of 0.1 au below the optimal C–X bond length and exceeding up to the cutoff distance, where the function as well as its first and second derivatives are bound to vanish. This treatment ensures a smooth function and unproblematic behavior of DFTB calculations. The recently developed procedure was used for fitting.<sup>16</sup>

The reference structure was obtained with gradient minimization on the B3LYP/cc-pVTZ level.<sup>43–46</sup> Atomization energies reported by Dixon et al.<sup>47</sup> from high-level ab initio calculations were corrected for the DFTB overbinding brought about by the MIO parametrization<sup>11</sup> of 4.3 kcal/mol per C–H bond, that is, of 13.0 kcal/mol for the molecule. The vibrational

Table 5. Scaled Scalar-Relativistic LDA Valence Orbital Energies of Heavy Halogen Atoms (in au)<sup>a</sup>

	s orbital	p orbital
Br		
this work	−0.746217	−0.294053
reference	−0.745247	−0.294158
deviation	−0.000970	+0.000105
I		
this work	−0.647127	−0.266126
reference	−0.646946	−0.266037
deviation	−0.000181	−0.000089

<sup>a</sup>Reference data taken from ref 42.

Table 6. Compression Radii for the Calculation of Two-Center Slater–Koster Parameters (in au)

element	$r_{\text{wf}}$	$r_{\text{dens}}$
F <sup>a</sup>	2.5	9
Cl	3.8	15
Br	4.4	17
I	5.1	20

<sup>a</sup>Values for fluorine taken from ref 18.

frequency was taken from experimental reports (CCCBDB databank).<sup>48</sup> An additional overbinding of 7, 10, and 12 kcal/mol for C–Cl, C–Br and C–I, respectively, was introduced in order to meet the remaining two criteria. The resulting description of CH<sub>3</sub>X molecules with DFTB is shown in Table 7. Evidently, with the introduction of such a mild overbinding, the structure as well as the vibration frequencies of the halogenmethane molecules is reproduced well, with the C–X bonds up to 0.02 Å longer and the frequencies within 20 cm<sup>−1</sup> of the reference, which may be considered a negligible deviation.

Thus, the strategy of the current parametrization is to use a minimal set of molecules for the fitting of parameters. The advantages of such an approach are that the obtained repulsive functions are perfectly smooth, preventing numerical problems in calculations that would be conceivable otherwise, and most likely perfect results for molecules with a similar bonding pattern as that in the fitting set. Obviously, the behavior of the parameter sets must be tested on a sufficiently large set of

Table 4. Parameters  $W$  for Spin-Polarized SCC-DFTB Calculations (in au)

	F		Cl		Br		I	
	LDA	PBE	LDA	PBE	LDA	PBE	LDA	PBE
$W_{ss}$	−0.034	−0.037	−0.020	−0.022	−0.016	−0.018	−0.012	−0.014
$W_{sp}$	−0.030	−0.031	−0.016	−0.018	−0.013	−0.014	−0.010	−0.011
$W_{ps}$	−0.030	−0.031	−0.016	−0.018	−0.013	−0.014	−0.010	−0.011
$W_{pp}$	−0.029	−0.030	−0.014	−0.016	−0.012	−0.013	−0.010	−0.011

Table 7. Data Used to Fit the C–X Repulsive Potentials<sup>a</sup>

molecule	C–X length		atomization energy			C–X frequency	
	ref.	fit	ref.	fit	overbinding	ref.	fit
CH <sub>3</sub> Cl	1.801	1.812	395.1	415.4	7.3	732	714
CH <sub>3</sub> Br	1.959	1.973	381.0	404.3	10.3	611	601
CH <sub>3</sub> I	2.161	2.178	367.8	393.1	12.3	533	515

<sup>a</sup>Bond lengths in angstrom (Å), atomization energies in kcal/mol and vibrational frequencies in cm<sup>−1</sup>. The amount of overbinding corresponding to 3 C–H bonds (13 kcal/mol) had been subtracted, and each value of overbinding is related to the single C–X bond. The experimental vibrational frequencies are taken from CCCBDB.<sup>48</sup>

Table 8. Data Used to Fit the Repulsive Potentials in Diatomic Molecules Containing Halogens<sup>a</sup>

molecule	bond length		atomization energy			vib. frequency		
	ref.	fit	ref.	fit	overbinding	exp.	PBE	fit
Cl <sub>2</sub>	2.024	2.024	58.1	85.1	27.0	560	537	528
Br <sub>2</sub>	2.315	2.315	45.9	73.9	28.0	325	316	306
I <sub>2</sub>	2.696	2.696	36.6	65.6	29.0	215	217	214
HCl	1.283	1.283	107.1	118.1	11.0	2991	2892	2923
HBr	1.425	1.425	90.3	101.3	11.0	2649	2558	2610
HI	1.618	1.618	74.5	85.5	11.0	2309	2264	2281
ClF	1.648	1.648	61.6	98.6	37.0	784	767	737
BrF	1.774	1.774	59.6	96.6	37.0	670	695	638
IF	1.932	1.932	64.4	96.4	32.0		600	590
BrCl	2.170	2.170	52.4	79.4	27.0	444	429	418
ICl	2.353	2.353	50.3	76.3	26.0		385	369
IBr	2.502	2.502		69.0			264	252

<sup>a</sup>Bond length in angstrom (Å), atomization energies in kcal/mol, and vibrational frequencies in cm<sup>−1</sup>; experimental frequencies taken from ref 48.

molecules, which would ideally span all relevant bonding patterns involving halogen atoms in organic molecules. Such testing will be carried out in the following, to assess the performance of the current parametrization.

**3.3.2. H–Cl, H–Br, H–I, Cl–F, Cl–Cl, Cl–Br, Cl–I, Br–F, Br–Br, Br–I, I–F, and I–I.**  $E_{\text{rep}}$  was fitted to the bond length in the respective molecules obtained with B3LYP/cc-pVTZ (calculations involving iodine were performed with an effective core potential and the cc-pVTZ-PP basis set),<sup>49</sup> as well as their atomization energies from ref 47 and experimental vibration frequencies (CCCBDB) where available. The same form of the functions was assumed as in the case of C–X repulsive potentials. The bonds Br–I and I–I were parametrized without fitting of atomization energies as these are missing in ref 47. The need to introduce an additional overbinding was particularly pronounced in the case of halogen–halogen bonds, as mentioned earlier in the assessment of the previous fluorine parametrization. So, the binding energy of these molecules is overestimated largely, for instance by 27 kcal/mol for Cl<sub>2</sub> or by 37 kcal/mol for ClF and BrF. Interestingly, the situation is the worst for bonds involving a fluorine atom. The description of the diatomic molecules with the current parametrization is summarized in Table 8.

**3.3.3. O–Cl, O–Br, O–I, N–Cl, N–Br, and N–I.** The fitting of oxygen–halogen repulsives has been carried out in a standalone fitting process, using the automaton of ref 47, fitting for O–X (X = Cl, Br, I) on the molecules HXO<sub>2</sub>, HXO<sub>4</sub>, and the ionic XO<sub>4</sub><sup>−</sup>. The fit objectives were those of energy and forces on atoms.

Overbinding energies put in by hand were determined first in an automatic way, fitted as pairwise inhomogeneous energy terms for each bond to adjust the dissociation energy limit (equivalent to ‘shifting’ the repulsive potential heuristically, as common in DFTB parametrization). The machinery was quite

similar to the fitting of one-body energy terms in ref 47, except that inhomogeneous terms were bound to bonds instead of atoms here. Because this ‘natural’ level of overbinding gave quite high values, the whole set of repulsives treated here was refitted in a second run, with all of the overbindings reduced to 20% of their natural values. In parallel, the fit was slightly modified (an appropriate term was added to the objectives based on energy derivatives) to discourage high derivatives beyond the first-neighbor region in order to obtain reasonable dissociation curves even with the systematically reduced overbindings. A reduction to 20% was the largest possible reduction available with the above methodology still retaining good dissociation curves. We present results with the latter, reduced overbindings; the respective values are presented in the Supporting Information. Note that the process outlined here corresponds to manually shifting the repulsive potential to avoid false minima of energy, finding the minimal necessary shift, and making a smooth interpolation to zero between the first- and second-neighbor distance; this procedure is implemented in a somewhat automatized way.

The fitting of N–X repulsives was done using energy and force targets of NH<sub>3–n</sub>X<sub>n</sub>,  $n = 1, 2, 3$ , and NOX, using the repulsives of the above O–X fit. Natural overbindings were fitted here using the ‘naturally overbound’ O–X set, and the resulting natural N–X overbindings were reduced to 20% to yield the fixed final ones in another fitting (using the parallel O–X repulsives with reduced overbinding in the second fit), similarly to the O–X fit.

The reference DFT calculator for O–X and N–X fitting was NWChem<sup>50</sup> with the B3LYP functional<sup>43</sup> and the cc-pVTZ basis set.<sup>44–46</sup> All electrons were considered on nonhalogen atoms, Cl, and Br, while relativistic pseudopotential treatment and the respective cc-pVTZ-PP basis set was applied for iodine.<sup>49</sup>

## 4. TESTING OF THE PARAMETRIZATION

**4.1. Halogenated Organic Molecules.** The parametrization was performed with a minimal possible set of halogenated organic molecules. Therefore, thorough testing of the parameter sets for all of the common bonding patterns involving halogens in organic molecules is necessary at this point. Structural properties and vibrational frequencies of 1–2 as well as 1–3 interactions will be monitored. As for reference data, B3LYP/cc-pVTZ calculations were used for the structure of the molecules, and PBE/cc-pVTZ calculations and experimental data available in the CCCBDB database were considered for vibrational frequencies. The detailed results of the testing are presented in the Supporting Information; a summary of the deviation of structural properties from the reference data is provided in Table 9.

**Table 9. Testing of the Parametrization for Halogenated Organic Molecules—Deviations from Reference DFT Data (B3LYP/cc-pVTZ for the Structures, PBE/cc-pVTZ for the Vibrational Frequencies)<sup>a</sup>**

	distances (Å)	stretching vib. (cm <sup>-1</sup> )	angles (deg.)	bending vib. (cm <sup>-1</sup> )
MSD	+0.027	+4	−1.1	+15
MAD	0.028	46	1.2	16
RMSD	0.033	60	1.6	22

<sup>a</sup>MSD, mean signed deviation; MAD, mean absolute deviation; RMSD, root-mean-squared deviation.

The description of halogenmethane molecules, which were used for parameter fitting, is good overall, considering also the H–C–X bonding angle and the corresponding vibrational frequency, with deviations below 1.5° and 50 cm<sup>-1</sup>. The correct behavior for these bonding angles shows the good performance of the H–X electronic parameter sets. Similarly, the C–X electronic parameters had to be checked for 1–3 interactions between carbon and halogens. The performance of the parametrization was good for halogenethane molecules, too, giving C–C–X angles within 0.5° and vibrational frequencies within 35 cm<sup>-1</sup> of the reference.

Three other groups of molecules contain a halogen atom connected to a pure carbon frame with different bonding character—a double bond C=C in halogenethenes, a triple bond C≡C in halogenethynes, and an aromatic ring in halogenbenzenes. Common to halogenethenes and halogenbenzenes is that the C–X bonds are slightly longer than in the reference, by up to 0.04 Å, while the vibrational frequencies as well as bonding angles match the reference. When the halogen atom is bound to a carbon participating in a triple bond, the C–X bonds assume a correct length; however, this seems to be a result of cancellation of errors to certain extent, as indicated by the vibrational frequencies that are ca. 120 cm<sup>-1</sup> too small except for fluorine. As an example of molecules with a more complex bonding aliphatic pattern, tert-butylhalides were included in the testing. The C–X bonds are up to 0.04 Å too long, and the corresponding vibrational frequencies are ca. 100 cm<sup>-1</sup> too large. Such errors are considered acceptable, taking into account the performance of the standard DFTB parametrization for a similar molecule, the tert-butylethyne, where the deviations are of a similar magnitude.

Another important binding pattern consists of several halogen atoms bound to a carbon atom. Here, particular attention shall be paid to the bonding angles, which are affected

by the electronic parameters for the concerned combination of halogens. Thus, dihalogenmethane molecules of types CH<sub>2</sub>X<sub>2</sub> and CH<sub>2</sub>XY were investigated. The X–C–X (X–C–Y) angles are consistently smaller than in reference calculations, with deviation below 1.5°, and the bending frequencies match the reference perfectly. Apparently, there are no issues regarding the halogen–halogen electronic parameters. To obtain an idea how the parametrization performs in more extreme situations, haloform molecules (CHX<sub>3</sub>) were tested, too. The C–X bonds are up to 0.05 Å too long with DFTB, and the corresponding vibrational frequencies are correct.

Similar potentially difficult cases are represented by molecules that feature a halogen atom and another electronegative atom bonded to a single carbon atom. Thus, the testing was concluded with three groups of molecules containing oxygen or nitrogen. (Note that some of these compounds are thermodynamically unstable, yet their study provides insight into the performance of the DFTB parametrization.) In CH<sub>2</sub>(OH)X, the C–X bonds are up to 0.05 Å too long and the O–C–X angles up to 3° too small, while the stretching frequencies are within 40 cm<sup>-1</sup> of the reference, with the bending ones even closer. In acetyl halides (CH<sub>3</sub>COX), the bond lengths are slightly worse (deviation up to 0.08 Å) and the angles are correct (within 1° of the reference), whereas the quality of vibrational frequencies is the same as for CH<sub>2</sub>(OH)X. The performance of the parametrization for the nitrogen-containing molecules CH<sub>2</sub>(NH<sub>2</sub>)X resembles that for CH<sub>2</sub>(OH)X somewhat: The C–X bonds are correct but the N–C–X angles are up to 4° too small, and the stretching frequencies are up to 80 cm<sup>-1</sup> too large. In this particular case, a nonhalogenated molecule CH<sub>2</sub>(NH<sub>2</sub>)C≡CH was included in the test set, to obtain an idea of the performance of the original DFTB parametrization for similar compounds. Actually, the respective bonding angle is 3° too small, and the C–C stretching frequency is 90 cm<sup>-1</sup> too large, which is roughly the same results as with the new parametrization for halogens.

Therefore, the binding situations among halogenated organic molecules that are the most difficult for the current parametrization of SCC-DFTB appear to be (i) cumulation of several halogen atoms on a single carbon atom and (ii) bonding of a halogen atom to a carbon that carries another electronegative atom. Additional calculations performed with the newer variant of DFTB involving the expansion of energy in density up to the third order (DFTB3)<sup>21</sup> showed no significant improvement of the results; see the Supporting Information for details. Also, these findings are in agreement with the previously observed performance of SCC-DFTB for similar binding motifs. The remaining molecules tested are described with a better accuracy.

**4.2. Halogenated Molecules with Oxygen and Nitrogen.** The performance of the O–X and N–X repulsive potentials was assessed for the molecules HXO<sub>n</sub>, X<sub>2</sub>O, anionic XO<sub>n</sub><sup>−</sup>, NH<sub>n</sub>X<sub>3−n</sub>, and NOX. The various bonding motifs will be discussed in the following sections. The observed deviations from reference DFT calculations are shown in Table 10.

First, the molecules of oxoacids of halogens HXO<sub>n</sub> are tested. The challenging character of halogen-oxygen bonding is illustrated by the fact that the atomization energies are not reproduced well; they are overestimated by 35 to 130 kcal/mol. This value increases with the increasing number of oxygen atoms in the molecule and thus the increasing number of O–X bonds, and it is the largest for molecules with chlorine. Still, the molecular geometries are fair: The single bonds O–X are 0.02–



**Table 10. Testing of the Parametrization for Halogenated Molecules with Oxygen and Nitrogen—Deviations from Reference DFT/B3LYP/cc-pVTZ Data<sup>a</sup>**

	molecules with O		molecules with N		HOX, HXO <sub>4</sub> , NX <sub>3</sub>
	distances (Å)	angles (deg.)	distances (Å)	angles (deg.)	vibrations (cm <sup>-1</sup> )
MSD	+0.010	−1.1	+0.009	+1.5	+40
MAD	0.043	3.6	0.030	1.9	57
RMSD	0.066	5.1	0.043	2.5	80

<sup>a</sup>MSD, mean signed deviation; MAD, mean absolute deviation; RMSD, root-mean-squared deviation.

0.08 Å too short except in HXO<sub>3</sub> where they are slightly too long, the double bonds O=X are up to 0.06 Å too long except in HXO<sub>4</sub> where they are slightly too short. The involved bonding angles O–X–O agree with the reference within 7° except in HXO<sub>2</sub>.

The molecules HXO<sub>2</sub> turn out to be truly difficult for SCC-DFTB to describe, and it is worth analyzing these difficulties in detail. Most distinctly, the method fails to describe the energy profile along the dihedral angle H–O–X=O: while the B3LYP calculation predicts the dihedral angle of about 80°, planar structures are obtained with SCC-DFTB. Additional calculations performed with PM6 and even with B3LYP involving a minimal basis set predicted planar structures as minima of energy, too. Therefore, the failure to describe the structure of HXO<sub>2</sub> correctly shall be ascribed rather to the SCC-DFTB method itself or, more specifically, to the fact that a minimal basis set is applied; it is not considered to be a problem of the parametrization.

The description of molecular geometry of HXO<sub>3</sub> and HXO<sub>4</sub> is slightly deficient in a related way. While the bond lengths and angles are obtained in satisfactory accordance with the reference, there is an issue with the energy profiles along dihedrals. So, for HXO<sub>3</sub>, geometries with C<sub>s</sub> symmetry are obtained with SCC-DFTB, while full DFT calculations give slightly distorted structures. On the other hand, for HClO<sub>4</sub> and HBrO<sub>4</sub>, the minimum of energy is found for staggered geometries with full DFT, which correspond to transition states between minima represented by eclipsed conformations with SCC-DFTB. This is also indicated by the negative vibrational frequency of the corresponding mode of motion, see below and the Supporting Information. We note again that this deficiency has most likely a deeper reason than a parametrization issue, as discussed in the previous paragraph.

Turning the attention to the ions XO<sub>n</sub><sup>−</sup>, a similar overestimation of atomization energies by 10 to 130 kcal/mol is found, increasing with the increasing number of oxygen atoms and being the largest for molecules with chlorine. The bonds are within 0.02 Å of the reference for XO<sub>4</sub><sup>−</sup> (which had been used in the parametrization) while they are up to 0.10 Å too long for the other molecules. The O–X–O angles agree with the reference within 1° except in IO<sub>2</sub><sup>−</sup>. Regarding the distribution of charge in these species, the magnitudes of partial atomic charges seem to be too large with SCC-DFTB; detailed results are presented in the Supporting Information. No significant improvement is observed upon passing to the third-order DFTB3 method.

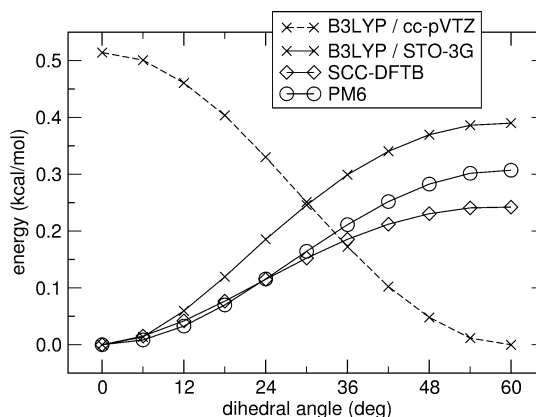
It should be noted that the molecules of oxoacids HXO<sub>n</sub> and oxoanions XO<sub>n</sub><sup>−</sup> are hypervalent, and as such, they are more challenging to describe with the minimal basis set employed in

the current parametrization. While the small basis set may be one of the factors that affect the distribution of charge in XO<sub>n</sub><sup>−</sup>, which was mentioned in the previous paragraph, the structure and vibrations of the tested molecules are described well with the current parametrization. Therefore, the parametrization is kept with the current basis set.

As for the molecules of halogenated compounds of nitrogen, these seem to be less problematic. The atomization energy is overestimated by at most 28 kcal/mol, and the geometries are described well. The N–X bonds are just up to 0.03 Å too short, and all of the angles are within 4° of the reference. The situation gets worse if an oxygen atom is bound to the nitrogen. Then, the deviation of atomization energy reaches 50 kcal/mol, and the N–X bond is up to 0.10 Å too long. Apparently, the O–X or O–N bonds are problematic in SCC-DFTB calculations to a certain extent.

Also tested was the performance of the current SCC-DFTB parametrization for the vibrational frequencies of several molecules with O–X and N–X bonds. For HOX molecules, the frequency of O–X stretching modes are up to 100 cm<sup>−1</sup> too high; notably, the deviation reaches nearly 300 cm<sup>−1</sup> for the stretching modes. Since the description of vibrational modes is not the primary objective of the current work, the parametrization will be kept in spite of the large deviation of vibrational frequencies. Another reason is that the frequencies obtained for the other groups of molecules, HXO<sub>4</sub> and NX<sub>3</sub>, are reproduced better: the deviation from the DFT reference values is below 100 cm<sup>−1</sup>, except one mode of NCl<sub>3</sub>, and is even below 50 cm<sup>−1</sup> for the vibrational modes weaker than 500 cm<sup>−1</sup>.

At this point, it is worth mentioning the energy profile of the rotation along the X–O bond in HXO<sub>4</sub> again. As reported, such a rotation is described poorly with SCC-DFTB. To analyze this in somewhat more detail, this energy profile was evaluated with the reference DFT method B3LYP/cc-pVTZ as well as B3LYP with a minimal basis set STO-3G, the current parametrization of SCC-DFTB, and the modern SE method PM6; the result is shown in Figure 1. Notably, the energy curves are quite shallow



**Figure 1.** Energy profile for the rotation along the Cl–O bond in the HClO<sub>4</sub> molecule evaluated with the various methods. 0°, eclipsed; 60°, staggered conformation.

with all of the methods, with barrier heights below 1  $k_B T$  at 300 K. All of the methods but B3LYP with the larger triple- $\zeta$  basis set favor the eclipsed conformation, failing to predict the correct minimum. While we are unable to identify the cause of this failure with certainty, the issue seems to be linked to the quality of the basis set. In correlation with this observation, the



structure of  $\text{HClO}_4$  that is a minimum of energy with B3LYP/cc-pVTZ is actually a saddle point with SCC-DFTB, as indicated by one imaginary vibrational frequency, see the Supporting Information. The situation is similar for  $\text{HBrO}_4$ , while the energy profile along the I–O rotation in  $\text{HIO}_4$  is shallower.

**4.3. On Reaction Energies.** The absolute value of binding energy of bonds involving halogen atoms had been resigned with the introduction of overbinding in the current parametrization. Obviously, the wrong binding energies will spoil the reaction enthalpies for processes where the halogen bonding pattern changes. While it is not the main aim of the current parametrization, the reaction energies for several simple chemical changes is presented in the following for the sake of completeness. Several kinds of reactions were studied: synthesis of hydrogen halides from elements, dissociation of ethane by halogens to form methyl halides, substitution of one hydrogen atom in methane by a halogen atom, halogen interchange in methyl halides, and substitution of OH and  $\text{NH}_2$  groups by a halogen atom. The reaction energies obtained as differences of relevant atomization energies with the current parametrization of DFTB are compared to reference data in Table 11.

The observations differ by the type of reactions. Whenever a halogen–halogen bond is broken (left column in Table 11), the strong overstabilization of the dihalogen molecule makes the reaction energy too endothermic. The deviation of DFTB reaction energy from the reference amounts to 27–32 kcal/mol whenever the fluorine molecule is involved, and 9–17 kcal/mol otherwise. Contrary to this, quite good reaction energies are obtained if no dihalogen molecules are involved in the reaction, as is the case for reactions  $\text{CH}_3\text{X} + \text{HY} \rightarrow \text{CH}_3\text{Y} + \text{HX}$ . The compensation of overbinding is nearly perfect in these reactions, making the reaction energies agree within 3 kcal/mol of the reference values.

Evidently, the energetics of chemical reactions is not described well but in cases with the same bonding patterns in the reactants and in the products. Thus, the application of the current DFTB parametrization is limited to cases where the reaction energetics is not the main object of interest. Structural studies of biomolecular complexes with organic ligands are an example.

## 5. CONCLUSION

The approximative DFT method SCC-DFTB was parametrized for molecules containing Cl, Br, and I. The obtained parameter sets are compatible with parameters for C, H, N, O, and for F from refs 11 and 18, respectively. The performance of current parametrization was assessed for a set of halogenated organic molecules with commonly occurring binding situations, as well as for a set of inorganic halogen molecules containing oxygen and nitrogen. The geometries and vibrational frequencies obtained with SCC-DFTB were compared to quantum chemical and experimental reference data and discussed critically. Several situations were identified where the performance is poorer than it is on average, and possible reasons were discussed. Finally, we note that it may be advantageous to complement the current parametrization of SCC-DFTB with empirical corrections for certain types of nonbonded interactions. These would be a correction for dispersion energy,<sup>53,54</sup> or one for halogen bonding (due to the presence of  $\sigma$ -hole on bromine or iodine atoms).<sup>55</sup>

**Table 11. Reaction Energies (kcal/mol) Obtained for Several Simple Chemical Reactions Involving Halogens with the Current Parametrization of DFTB<sup>a</sup>**

halogen X	ref.	DFTB
$\text{H}_2 + \text{X}_2 \rightarrow 2\text{HX}$		
F	−140.7	−108.2
Cl	−52.1	−37.9
Br	−30.7	−16.4
I	−8.4	+5.8
$\text{C}_2\text{H}_6 + \text{X}_2 \rightarrow 2\text{CH}_3\text{X}$		
F	−94.9	−64.7
Cl	−21.3	−4.6
Br	−4.5	+9.3
I	+12.6	+24.3
$\text{CH}_4 + \text{X}_2 \rightarrow \text{CH}_3\text{X} + \text{HX}$		
F	−105.8	−78.2
Cl	−24.7	−13.0
Br	−5.6	+4.7
I	+14.1	+23.3
halogen(s) X, (Y)	ref.	DFTB
$\text{CH}_3\text{X} + \text{HY} \rightarrow \text{CH}_3\text{Y} + \text{HX}$		
F, Cl	−7.5	−5.1
F, Br	−9.8	−8.9
F, I	−12.4	−12.5
Cl, Br	−2.3	−3.8
Cl, I	−4.9	−7.4
Br, I	−2.6	−3.6
$\text{CH}_3\text{OH} + \text{HX} \rightarrow \text{CH}_3\text{X} + \text{H}_2\text{O}$		
F	−0.7	−2.4
Cl	−8.2	−7.5
Br	−10.5	−11.3
I	−13.1	−14.9
$\text{CH}_3\text{NH}_2 + \text{HX} \rightarrow \text{CH}_3\text{X} + \text{NH}_3$		
F	+2.8	+4.8
Cl	−4.7	−0.3
Br	−7.0	−4.1
I	−9.6	−7.7

<sup>a</sup>The miomod:hh parameters as published at [www.dftb.org](http://www.dftb.org) were used to describe the  $\text{H}_2$  molecule properly. Reference data are taken from Dixon et al.<sup>47,51</sup> where available and from G3 calculations<sup>52</sup> otherwise.

The described SCC-DFTB parameter sets can be downloaded from [www.dftb.org](http://www.dftb.org) free of charge, after prior registration at the Web site.

## ■ ASSOCIATED CONTENT

### Supporting Information

Specification of the DFT basis set used in the parametrization; parameters used in the generation of two-center repulsive potentials; full data from the testing of the parametrization; discussion of the possible improvement of performance upon passing to the third-order approximation (DFTB3). This material is available free of charge via the Internet at <http://pubs.acs.org/>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Phone: +49 721 608 45705. E-mail: [marcus.elstner@kit.edu](mailto:marcus.elstner@kit.edu)

### Present Address

<sup>§</sup>Institute for Solid State Physics and Optics, Wigner Research Center for Physics, Hungarian Academy of Sciences, P.O. Box 49, 1525 Budapest, Hungary

## Notes

The authors declare no competing financial interest.

## REFERENCES

- (1) Levine, I. *Quantum Chemistry*, 4th ed.; Prentice Hall: Englewood Cliffs, NJ, 1991; pp 579–580.
- (2) Cramer, C. J. *Essentials of Computational Chemistry*; Wiley: Chichester, U.K., 2002; pp 126–131.
- (3) Thiel, W. In *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed.; NIC Series; John von Neumann Institute for Computing: Jülich, Germany, 2000; Vol. 3; pp 261–283.
- (4) Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*; McGraw-Hill Series in Advanced Chemistry; McGraw-Hill: New York, 1970; pp 57–84.
- (5) Dewar, M. J. S.; Thiel, W. Ground states of molecules. 38. The MNDO method. Approximations and parameters. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
- (6) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (7) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *J. Comput. Chem.* **2006**, *27*, 1101–1111.
- (8) Stewart, J. J. P. Optimization of parameters for semiempirical methods. I. Method. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (9) Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (10) Weber, W.; Thiel, W. Orthogonalization corrections for semiempirical methods. *Theor. Chem. Acc.* **2000**, *103*, 495–506.
- (11) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* **1998**, *58*, 7260–7268.
- (12) Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. L. Comparison of SCC-DFTB and NDDO-based semiempirical molecular orbital methods for organic molecules. *J. Phys. Chem. A* **2006**, *110*, 13551–13559.
- (13) Otte, N.; Scholten, M.; Thiel, W. Looking at self-consistent-charge density functional tight binding from a semiempirical perspective. *J. Phys. Chem. A* **2007**, *111*, 5751–5755.
- (14) Elstner, M. SCC-DFTB: What is the proper degree of self-consistency. *J. Phys. Chem. A* **2007**, *111*, 5614–5621.
- (15) Krüger, T.; Elstner, M.; Schiffels, P.; Frauenheim, T. Validation of the density-functional based tight-binding approximation method for the calculation of reaction energies and other data. *J. Chem. Phys.* **2005**, *122*, 114110.
- (16) Gaus, M.; Chou, C.-P.; Witek, H.; Elstner, M. Automated parametrization of SCC-DFTB repulsive potentials: Application to hydrocarbons. *J. Phys. Chem. A* **2009**, *113*, 11866–11881.
- (17) Bodrog, Z.; Aradi, B.; Frauenheim, T. Automated repulsive parametrization for the DFTB method. *J. Chem. Theory Comput.* **2011**, *7*, 2654–2664.
- (18) Köhler, C.; Frauenheim, T. Molecular dynamics simulations of  $\text{CF}_x$  ( $x = 2, 3$ ) molecules at  $\text{Si}_3\text{N}_4$  and  $\text{SiO}_2$  surfaces. *Surf. Sci.* **2006**, *600*, 453–460.
- (19) Lu, Y.; Wang, Y.; Zhu, W. Nonbonding interactions of organic halogens in biological systems: Implications for drug discovery and biomolecular design. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4543–4551.
- (20) Ibrahim, M. A. A. Performance assessment of semiempirical molecular orbital methods in describing halogen bonding: Quantum mechanical and quantum mechanical/molecular mechanical-molecular dynamics study. *J. Chem. Inf. Model.* **2011**, *51*, 2549–2559.
- (21) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.
- (22) Gaus, M.; Goez, A.; Elstner, M. Parametrization and benchmark of DFTB3 for organic molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.
- (23) Seifert, G. Tight-binding density functional theory: An approximate Kohn-Sham DFT scheme. *J. Phys. Chem. A* **2007**, *111*, 5609–5613.
- (24) Elstner, M. The SCC-DFTB method and its application to biological systems. *Theor. Chem. Acc.* **2006**, *116*, 316–325.
- (25) Köhler, C.; Seifert, G.; Gerstmann, U.; Elstner, M.; Overhof, H.; Frauenheim, T. Approximate density-functional calculations of spin densities in large molecular systems and complex solids. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5109–5114.
- (26) Köhler, C.; Seifert, G.; Frauenheim, T. Density functional based calculations for  $\text{Fe}_n$  ( $n \leq 32$ ). *Chem. Phys.* **2005**, *309*, 23–31.
- (27) Köhler, C.; Frauenheim, T.; Hourahine, B.; Seifert, G.; Sternberg, M. Treatment of collinear and noncollinear electron spin within an approximate density functional based method. *J. Phys. Chem. A* **2007**, *111*, 5622–5629.
- (28) Elstner, M.; Frauenheim, T.; McKelvey, J.; Seifert, G. Density functional tight binding: Contributions from the American Chemical Society Symposium. *J. Phys. Chem. A* **2007**, *111*, 5607–5608.
- (29) Aradi, B.; Hourahine, B.; Frauenheim, T. DFTB+, a sparse matrix-based implementation of the DFTB method. *J. Phys. Chem. A* **2007**, *111*, 5678–5684.
- (30) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (31) Perdew, J. P.; Wang, Y. Accurate and simple analytic representation of the electron–gas correlation energy. *Phys. Rev. B* **1992**, *45*, 13244–13249.
- (32) Koelling, D. D.; Harmon, B. N. Technique for relativistic spin-polarized calculations. *J. Phys. C* **1977**, *10*, 3107–3114.
- (33) Van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component Hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597–4610.
- (34) Heera, V.; Seifert, G.; Ziesche, P. A semi-relativistic variant of the scattered-wave  $X\alpha$  method. *J. Phys. B* **1984**, *17*, 519–530.
- (35) Koskinen, P.; Häkkinen, H.; Seifert, G.; Sanna, S.; Frauenheim, T.; Moseler, M. Density-functional based tight-binding study of small gold clusters. *New J. Phys.* **2006**, *8*, 9.
- (36) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B* **1995**, *51*, 12947–12957.
- (37) Slater, J. C.; Koster, G. F. Simplified LCAO method for the periodic potential problem. *Phys. Rev.* **1954**, *94*, 1498–1524.
- (38) Seifert, G.; Köhler, T.; Frauenheim, T. Molecular wires, solenoids, and capacitors by sidewall functionalization of carbon nanotubes. *Appl. Phys. Lett.* **2000**, *77*, 1313–1315.
- (39) Amkreutz, M.; Jungnickel, G.; Seifert, G.; Köhler, T.; Frauenheim, T. On the structural and electronic properties of fluorinated carbon allotropes. *New Diamond Front. Carbon Technol.* **2001**, *11*, 207–220.
- (40) Knaup, J. M.; Köhler, C.; Hoffmann, M.; Koenig, P. H.; Frauenheim, T. Ab initio simulation of interface reactions as a foundation of understanding polymorphism. *Eur. Phys. J.: Spec. Top.* **2007**, *149*, 127–144.
- (41) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.

Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C.02.; Gaussian, Inc.: Wallingford, CT, 2004.

(42) Kotochigova, S.; Levine, Z. H.; Shirley, E. L.; Stiles, M. D.; Clark, C. W. *Atomic Reference Data for Electronic Structure Calculations*; NIST: Gaithersburg, MD, 2003. <http://physics.nist.gov/DFTdata> (accessed Jan 15, 2013).

(43) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle–Salvetti correlation energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(44) Dunning, T. H. Gaussian-basis sets for use in correlated molecular calculations. 1. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(45) Woon, D. E.; Dunning, T. H. Gaussian-basis sets for use in correlated molecular calculations. 3. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

(46) Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. IX. The atoms gallium through krypton. *J. Chem. Phys.* **1999**, *110*, 7667–7676.

(47) Feller, D.; Peterson, K. A.; De Jong, W. A.; Dixon, D. A. Performance of coupled cluster theory in thermochemical calculations of small halogenated compounds. *J. Chem. Phys.* **2003**, *118*, 3510–3522.

(48) Johnson R. D., III, Ed.; *NIST Computational Chemistry Comparison and Benchmark Database*; NIST: Gaithersburg, MD, 2011. <http://cccbdb.nist.gov> (accessed Jan. 15, 2013).

(49) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16–18 elements. *J. Chem. Phys.* **2003**, *119*, 11113–11123.

(50) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **2010**, *181*, 1477–1489.

(51) Feller, D.; Dixon, D. A. Predicting the heats of formation of model hydrocarbons up to benzene. *J. Phys. Chem. A* **2000**, *104*, 3048–3056.

(52) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

(53) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density functional theory based treatment. *J. Chem. Phys.* **2001**, *114*, 5149–5155.

(54) Řezáč, J.; Hobza, P. Advanced corrections of hydrogen bonding and dispersion for semiempirical quantum mechanical methods. *J. Chem. Theory Comput.* **2012**, *8*, 141–151.

(55) Řezáč, J.; Hobza, P. A halogen-bonding correction for the semiempirical PM6 method. *Chem. Phys. Lett.* **2011**, *506*, 286–289.