



<http://www.diva-portal.org>

## Postprint

This is the accepted version of a paper published in *International Journal of Quantum Chemistry*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Lundberg, M., Nishimoto, Y., Irle, S. (2012)

Delocalization errors in a hubbard-like model: Consequences for density-functional tight-binding calculations of molecular systems.

*International Journal of Quantum Chemistry*, 112(6): 1701-1711

<http://dx.doi.org/10.1002/qua.23178>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-181243>

# Delocalization Errors in a Hubbard-like Model: Consequences for Density-Functional Tight-Binding Calculations of Molecular Systems

Marcus Lundberg,<sup>1\*‡</sup> Yoshio Nishimoto<sup>2</sup>, Stephan Irle<sup>2</sup>

<sup>1</sup> *Fukui Institute for Fundamental Chemistry, Kyoto University, 34-4 Takano Nishihiraki-cho,  
Sakyo-ku, Kyoto 606-8103, Japan*

<sup>2</sup> *Department of Chemistry, Graduate School of Science, Nagoya University, Furo-cho,  
Chikusa-ku, Nagoya 464-8601, Japan*

## Abstract

It has previously been shown that self-consistent-charge density-functional tight-binding (SCC-DFTB) suffers from a self-interaction error that leads to artificial stabilization of delocalized states. The effects of the error are similar to those appearing for many density functionals. In SCC-DFTB the delocalization error is inherently related to the use of a Hubbard-like term to describe on-site charge interactions. The mathematical simplicity of this Hubbard-like term makes it easy to estimate if a complex system is subject to artificial stabilization of delocalized states, and to quantitatively predict the delocalization error in the system energy at large fragment separation. The error is directly proportional to the on-site charge interaction term, but decreases as the fragments become more asymmetric. The difference in orbital energies required to eliminate the delocalization error becomes equal to the Hubbard-like parameter of the fragment with the highest electron affinity. However, in most cases the localized state will be favored by spin polarization, fragment repulsion, solvent effects and large reorganization energies, in analogy to density functional theory (DFT), from which SCC-DFTB is derived. The presented analysis gives an early indication whether the standard approach is suitable, or if a different method is required to correct the delocalization error. In addition to cationic dimers, we discuss the effects of the delocalization error for asymmetric systems, bond dissociation of neutral molecules, and the description of mixed valence transition metal systems, exemplified by the enzyme cytochrome oxidase.

\* To whom correspondence should be addressed: E-mail: [marcus.lundberg@kvac.uu.se](mailto:marcus.lundberg@kvac.uu.se), Telephone: +46-18-4713208. Fax: +46-18-4715830.

‡ Present address: Department of Physical and Analytical Chemistry, Ångström laboratory, Uppsala University, Box 518, SE-751 20 Uppsala, Sweden.

## I. Introduction

Simulation of molecular processes requires a variety of methods to cover different time scales, system sizes, and accuracy demands. Density-functional tight-binding (DFTB) is a computationally efficient molecular orbital method that occupies a position between the elaborate wavefunction or density functional methods, and the very fast molecular mechanics methods.<sup>1-4</sup> One- and two-electron integrals are approximated by interatomic, distance-dependent Hamiltonian matrix elements and repulsive potentials within the framework of a two-center approximation, respectively, and a minimum basis set of Slater-type pseudoatomic orbitals. Both Hamiltonian matrix elements and repulsive potentials are obtained from first-principles density functional theory (DFT) calculations of atoms, diatomic systems, and small model compounds. DFTB calculations are several orders of magnitude faster than DFT, and similar in cost to semiempirical methods like the semiempirical AM1 and PM3 methods.<sup>5-7</sup> The efficiency of DFTB allows for the treatment of large systems over significant time scales with applications in solid-state physics, biology, and nanoscale simulations.<sup>4,8-10</sup>

As discussed in previous works, the self-consistent-charge version of DFTB (SCC-DFTB) has a self-interaction error that leads to artificial stabilization of delocalized states.<sup>11-13</sup> For charge resonance situations in molecular complexes, the behavior is similar to the effects of the self-interaction error in many density functionals. A classic illustration of the delocalization error is the dissociation of  $\text{H}_2^+$ .<sup>14</sup> As this one-electron system lacks correlation, Hartree–Fock (HF) gives the correct dissociation curve, see Figure 1. As the internuclear distance increases, the interaction should decrease and at infinite distance the energy should converge to the sum of energies of a hydrogen atom  $E(\text{H})$  and a proton  $E(\text{H}^+)$ . This is also true for non-charge-consistent DFTB (NCC-DFTB), which does not contain a Hubbard-like charge term. However, as shown in reference <sup>13</sup>, the SCC-DFTB energy decreases with increasing internuclear distance and does not converge toward the correct limit, see Figure 1. At infinite distance the energy is underestimated by 54.5 kcal/mol. This behavior is very similar to that of standard generalized gradient approximation (GGA) functionals,<sup>14</sup> and analogous calculations using the density functional PBE (with 6-31G(d) basis set) underestimate the energy of the delocalized solution by 67 kcal/mol.

Figure 1.

The artificial stabilization of delocalized states in standard DFT has several effects: incorrect dissociation limits of odd-electron systems (e.g.  $\text{H}_2^+$ ); underestimation of transition state barriers; exaggerated delocalization in conjugated systems; too low energy of charge-transfer excitations; and exaggerated delocalization in transition metal dimers.<sup>14-20</sup> There exists self-interaction corrected functionals that address this issue, but they typically perform worse in more general areas, e.g., thermochemistry.<sup>18</sup> To choose an adequate DFT treatment it is therefore important to understand the effects of the delocalization error for a specific model. By analogy, it is equally important to understand the effects of the delocalization error in SCC-DFTB simulations of complex molecular systems.

We have previously used the NCC- and SCC-DFTB methods to perform molecular dynamics (MD) simulations of fullerene and transition metal catalyzed nanotube formation.<sup>21,22</sup> SCC-DFTB has also been used in ONIOM multi-scale models to describe polarization and charge transfer in large transition-metal systems.<sup>9,23,24</sup> These reactions often involve open-shell species and intermittent radicals, for which the delocalization error of SCC-DFTB is of considerable relevance.

Here we show that an analysis of the main contribution to the delocalization error SCC-DFTB, the Hubbard-like term describing charge-charge interactions, makes it easy to estimate the effect of the delocalization error, even in very complex systems. Using this analysis, we investigate the effects of the delocalization error for asymmetric radical systems, bond dissociation of neutral molecules, and the description of mixed valence transition metal systems, exemplified by the enzyme cytochrome oxidase. The aim is to provide an educational illustration of different aspects of the delocalization error and help in designing stable SCC-DFTB models. The presented analysis should be seen as a complement to more fundamental theoretical work on the self-interaction in SCC-DFTB, e.g. by Hourahine *et al.*<sup>11</sup>

## II. Spin-Polarized Self-Consistent-Charge DFTB Approach

Although the DFTB method has been described extensively in other contributions,<sup>1-4</sup> we repeat the details that are important for an understanding of the delocalization error. DFTB can be seen as a parameterized approximation to DFT.<sup>2</sup> In the basic non-charge-consistent (NCC)-DFTB formulation the energy is expressed as:<sup>3</sup>

$$E_{\text{tot}}^{\text{NCC-DFTB}} = \sum_{\sigma=\uparrow,\downarrow} \sum_i^{\text{MO}} n_i^\sigma \langle \psi_i^\sigma | \hat{H}^0[\rho_0] | \psi_i^\sigma \rangle + \sum_{A<B}^{\text{atom}} E_{AB}^{\text{rep}}(R_{AB}) \quad (1)$$

The molecular orbitals ( $\psi_i^\sigma$ ) are expanded in a minimum valence basis set of localized Slater pseudoatomic orbitals using the LCAO-MO approach. The Hamiltonian ( $\hat{H}^0[\rho_0]$ ) depends on the initial atomic densities  $\rho_0$  and only contains two-center contributions (characteristic for all tight-binding methods). In this approximation the matrix elements are simple functions of pair-wise interatomic distances. The diagonalization of the Hamiltonian leads to a new electron density, but as the Hamiltonian itself is a function of the input density only, the NCC procedure is not iterative. Finally, as an approximation to the terms neglected in DFTB, a corrective potential as a function of the pair-wise interatomic distance, often called ‘‘repulsive potential’’  $E_{AB}^{\text{rep}}(R_{AB})$  is introduced. It is worth mentioning that in NCC-DFTB no spin-polarization is included, and that alpha and beta spin orbitals are identical in shape and energies.

To improve the description of heteroatomic systems containing charge density fluctuations relative to the atomic densities, the SCC-DFTB method has been developed.<sup>1</sup> In addition, fluctuations in spin densities can be considered by using a spin-polarized variation.<sup>2</sup> The spin-polarized SCC-DFTB energy can be written as:

$$E_{\text{tot}}^{\text{SCC-DFTB}} = \sum_{\sigma=\uparrow,\downarrow} \sum_i^{\text{MO}} n_i^\sigma \langle \psi_i^\sigma | \hat{H}^0[\rho_0] | \psi_i^\sigma \rangle + \frac{1}{2} \sum_{A,B}^{\text{atom}} \gamma_{AB}(R_{AB}) \Delta q_A \Delta q_B + \frac{1}{2} \sum_A^{\text{atom}} \sum_l \sum_{l'} p_{Al} p_{Al'} W_{All'} + \sum_{A<B}^{\text{atom}} E_{AB}^{\text{rep}}(R_{AB}) \quad (2)$$

where  $n_i^\sigma$  is the occupation number of the spin orbital  $\psi_i^\sigma$  and  $\uparrow$  and  $\downarrow$  denote  $\alpha$  and  $\beta$  spin orientation. The second order correction ( $\gamma_{AB}(R_{AB}) \Delta q_A \Delta q_B$ ) describes charge-charge interactions between induced Mulliken charges  $\Delta q_A$  and  $\Delta q_B$ , scaled by a distance-dependent interaction function  $\gamma_{AB}(R_{AB})$ .<sup>1</sup> The on-site terms (e.g.,  $\gamma_{AA}$ ) are Hubbard-like parameters, related to the chemical hardness. As will be shown later, the on-site terms are responsible for the delocalization error in SCC-DFTB. Spin-polarization is expressed as interactions between Mulliken spin populations  $p_{Al}$  and  $p_{Al'}$ , where  $W_{All'}$  is an interaction parameter that couples orbitals in shells with different  $l$  quantum numbers on the same atom.<sup>2</sup> In this work the term SCC-DFTB refers to the spin-polarized version, unless explicitly mentioned otherwise. Atomic parameters are usually derived using the PBE density functional.<sup>1</sup>

Calculations have been performed using a private development version of Gaussian,<sup>25</sup> as well as the publically available DFTB+ code, which can be downloaded at <http://www.dftb-plus.info>.

### III. Results

**Dissociation of  $H_2^+$ .** Although the effect of the delocalization error for  $H_2^+$  is similar in SCC-DFTB and DFT GGA functionals, the origin of the error is much easier to analyze. For large internuclear distances, SCC-DFTB gives a solution with fractional occupation numbers, where the electron is shared ( $H^{0.5+}$ ) between the two centers at all distances. This is the only possible solution for a single-reference method, but the artificial stabilization originates in a difference in energy between the delocalized ( $H^{0.5+} + H^{0.5+}$ ) and the localized solution ( $H + H^+$ ). Using equation 2:

$$\begin{aligned}
 E^{SCC-DFTB}(H) &= \varepsilon_i + \frac{1}{2}W_{Hss} \\
 E^{SCC-DFTB}(H^+) &= \frac{1}{2}\gamma_{HH} \\
 E^{SCC-DFTB}(H^{0.5+}) &= \frac{1}{2}\varepsilon_i + \frac{1}{2}0.5*0.5\gamma_{HH} + \frac{1}{2}0.5*0.5W_{Hss} = \frac{1}{2}\varepsilon_i + \frac{1}{8}\gamma_{HH} + \frac{1}{8}W_{Hss}
 \end{aligned} \tag{3,4,5}$$

The delocalization energy ( $E_{del}$ ), i.e., the energy difference between the localized and the delocalized solution is:

$$E_{del} = 2E^{SCC-DFTB}(H^{0.5+}) - E^{SCC-DFTB}(H + H^+) = -\frac{1}{4}\gamma_{HH} - \frac{1}{4}W_{Hss} \tag{6}$$

Using the Hubbard-like parameter for hydrogen ( $\gamma_{HH}=0.4195$  h) and the spin polarization parameter for the hydrogen 1s orbital ( $W_{Hss}=-0.072$  h), gives a delocalization energy of  $-0.0869$  h ( $-54.5$  kcal/mol). In NCC-DFTB, the energy of the localized and the delocalized solution is the same, and the dissociation limit is correct, see Figure 1. It is thus the Hubbard-like on-site interaction that is the origin of the delocalization error in this system, while spin-polarization favors a localized solution.

**Estimating the amount of delocalization.** Although it is relatively difficult to analyze the electron self-interaction, the problem can be significantly simplified by focusing only on the *effects* of the self-interaction. One useful probe is the delocalization error, i.e., the energy difference between the delocalized and the localized state of a given system, as done for  $H_2^+$  above.

In a system with one electron and two fragments, A and B, with negligible orbital overlap, the electron should localize on the fragment with the highest electron affinity. Assuming this is fragment A, the degree of delocalization ( $x$ ) can vary from 0 (electron on A) to 0.5 (fully delocalized). To outline the general strategy, we initially ignore spin polarization, and assume infinite distance between the two fragments. The delocalization energy  $E_{del(\infty)}(x)$  is then:

$$\begin{aligned} E_{del(\infty)}(x) &= E(x) - E(x=0) = (1-x)\varepsilon_{A^-} + x\varepsilon_{B^-} + \frac{(1-x)^2}{2}\gamma_{AA} + \frac{x^2}{2}\gamma_{BB} - (\varepsilon_{A^-} + \frac{1}{2}\gamma_{AA}) \\ &= \frac{(1-x)^2}{2}\gamma_{AA} + \frac{x^2}{2}\gamma_{BB} - x(\varepsilon_{A^-} - \varepsilon_{B^-}) - \frac{1}{2}\gamma_{AA} \end{aligned} \quad (7)$$

where  $\varepsilon_{A^-}$  and  $\varepsilon_{B^-}$  are the orbital energies of the extra electron on fragments A and B,  $\gamma_{AA}$  and  $\gamma_{BB}$  are the Hubbard-like on-site charge-interaction parameters. For a given system, the degree of delocalization ( $x$ ) is given by a minimum in the function  $E_{del(\infty)}(x)$ . The search for a point with zero gradient with respect to  $x$  yields:

$$\frac{\partial E_{del(\infty)}}{\partial x}(x) = x\gamma_{AA} + x\gamma_{BB} - (\varepsilon_{A^-} - \varepsilon_{B^-}) - \gamma_{AA} = 0 \quad (0 \leq x \leq 0.5) \quad (8)$$

$$x = \frac{\gamma_{AA} + (\varepsilon_{A^-} - \varepsilon_{B^-})}{\gamma_{AA} + \gamma_{BB}} \quad (0 \leq x \leq 0.5) \quad (9)$$

For a symmetric system (A=B),  $x=0.5$  (full delocalization) and  $E_{del(\infty)}(0.5) = -\gamma_{AA}/4$ , in agreement with the result for  $H_2^+$ . The system is fully localized ( $x=0$ ) only if  $|\varepsilon_{A^-} - \varepsilon_{B^-}| \geq \gamma_{AA}$ , i.e., if the energy penalty for moving the electron from A to B is greater or equal to the Hubbard-like parameter of the fragment with the highest electron affinity. Spin-polarization should also be included, but in this case the effects depend on the orbital occupations. The effects can easily be calculated for individual cases, as done for  $H_2^+$  above.

In cases where charge and/or spin appear on several atoms in a fragment, equation 9 needs to be modified by adding several centers. However, as the energy depends on the square of the charge/spin, only the largest term (or terms) needs to be included.

To illustrate the effect of increasing asymmetry on the delocalization error, the breaking of the C-H bond in the methane and ethane cation radicals have been compared with dissociation of  $H_2^+$ , see Figure 2. There is a clear decrease in the delocalization error as the asymmetry of the fragments increase (ionization potential (IP) of the H,  $CH_3$ , and  $C_2H_5$  fragments are 302, 234, and 197 kcal/mol respectively). Comparing the behavior to the PBE functional, see Figure 3, it can be seen that at large distances, the two methods give very similar trends.

Figure 2.

Figure 3.

At finite distances, the analysis has to take into account the interaction between the two fragments. As  $E_{AB}^{\text{rep}}(R_{AB})$  does not depend on the electronic structure and cancels in all calculations of the delocalization error. For relatively long distances, the orbital overlap is small, and the dominating distance-dependent term is the charge-charge interaction which behaves in this region as a Coulomb potential. In the localized state, the charges of the two fragments are  $(q_A-1)$  and  $q_B$ , because the extra electron resides on fragment A. In the delocalized state, the value of  $q_A$  increases by  $x$  (losing  $x$  electrons) and  $q_B$  increases by the same amount. The contribution to the delocalization energy is then:

$$E_{del\gamma_{AB}}(x) = \frac{1}{2} \sum_{A,B}^{\text{atom}} \gamma_{AB}(R_{AB}) [(q_A - (1-x))(q_B - x) - (q_A - 1)q_B] = \sum_{A,B}^{\text{atom}} \gamma_{AB}(R_{AB}) [x(1 + q_B - q_A) - x^2] \quad (10)$$

The effect thus depends on the charges of the fragments, but in symmetric models,  $q_A=q_B=q$ , we obtain the following contribution:

$$E_{del\gamma_{AB}}(x) = \frac{1}{2} \sum_{A,B}^{\text{atom}} \gamma_{AB}(R_{AB}) [x - x^2] \quad (11)$$

This value is positive for all  $x$  ( $0 \leq x \leq 0.5$ ), and largest for  $x=0.5$  (fully delocalized). The interatomic charge-charge interactions thus favor the localized solution in most systems. The delocalization error is therefore largest at infinite distance, as seen in Figure 1. Systems that do not artificially delocalize at infinite distance are correctly described at all distances. Systems that delocalize at infinite distance can still localize if the contributions from  $E_{del\gamma_{AB}}(x)$  are larger than the delocalization energy at infinite distance. Approximating this value by the Coulomb effect for a system with  $x=0.5$  and  $R_{AB}=5 \text{ \AA}$  gives a contribution of 26.5 kcal/mol, and if at a given distance  $E_{del(\infty)}$  is smaller than  $E_{del\gamma_{AB}}(x)$  the system localizes at that particular distance despite the delocalization error.

In the region of small orbital overlap, the effects of the self-interaction can mainly be described by classical electrostatics. The argument above should therefore be equally valid for both DFT and SCC-DFTB calculations. Many other effects are also largely independent of the specific electronic structure method, and we therefore reiterate some general observations of the appearance of delocalized states from reference<sup>20</sup>.

For systems with more than two atoms, the localized state is favored by an extra degree of freedom since the two fragments can adopt different optimal geometries. The term vibronic coupling has also been used to describe the geometric effects associated with geometry relaxation.

Finally, the solvent effect further works to stabilize the localized state in most cases. For a spherical cavity, the solvation energy is:

$$E_{solv} = -\frac{\epsilon - 1}{2\epsilon} * \frac{q^2}{R} \quad (12)$$

where  $\epsilon$  is the dielectric constant of the solvent, and  $q$  and  $R$  are charge and radius of the solute respectively. The solvent contribution to the delocalization energy is:

$$E_{solv}(x) = -\frac{\epsilon - 1}{2\epsilon} * \frac{1}{R} [(q_A - (1 - x))^2 + (q_B - x)^2 - (q_A - 1)^2 - q_B^2] = -\frac{\epsilon - 1}{2\epsilon} * \frac{1}{R} [2x^2 - 2x + 2q_A x - 2q_B x] = \{q_A = q_B\} = -\frac{\epsilon - 1}{\epsilon} * \frac{x^2 - x}{R} \quad (13)$$

which is positive for all values of  $x$  ( $0 \leq x \leq 0.5$ ) assuming  $q_A = q_B$ . For  $R = 10 \text{ \AA}$  and  $\epsilon = 4$ , the localized state is favored by 12.4 kcal/mol. This is balanced against the decrease in Coulomb repulsion between the delocalized fragments. Assuming an 8- $\text{\AA}$  fragment separation, the Coulomb repulsion decreases from 10.4 to 2.6 kcal / mol, an effect of -7.8 kcal/mol. The solvent effect on the balance between localized and delocalized states thus depends on system size and fragment separation.

To summarize, it is relatively straightforward to analyze the factors that favor or disfavor localization. First, a basic analysis can be made to see if the system delocalizes at infinite distance between the fragments. At a second stage, the effects of fragment-fragment repulsion, geometry reorganization, and solvent effects can be taken into account to estimate if the system still delocalizes at more realistic modeling conditions. In regions with significant orbital overlap, the analysis becomes much more complicated, and the present analysis is not applicable.

To distinguish between a correctly and an artificially delocalized system it is possible to look at the distance dependence. Exchange interactions, that can also lead to delocalized states, decrease in strength with increasing distance, while artificial delocalization increases with increasing distance. A system that becomes more delocalized as the fragments move apart must therefore be affected by artificial stabilization of the delocalized state.

**Dissociation of closed-shell molecules.** Although there are many simulations where the treatment of an odd number of electrons is important, most processes involve closed-shell molecules. In our research program we have studied the formation of fullerenes and carbon nanotubes, using DFTB molecular dynamics simulations.<sup>21,22</sup> In these simulations we have often used C<sub>2</sub> molecules as feedstock, and we therefore take a deeper look at the activation of neutral diatomics.

The simplest example is the dissociation of H<sub>2</sub>. The correct dissociation limit corresponds to two hydrogen atoms with one electron each. A closed shell singlet calculation using spin-polarized SCC-DFTB gives an electronic structure with half an electron in both  $\alpha$  and  $\beta$ -orbitals on each atom. This leads to an incorrect dissociation limit, 45.2 kcal above the correct value, see Figure 4. In SCC-DFTB, the correct solution includes spin-polarization ( $W_{Hss}$ ), while the closed-shell singlet lacks all spin polarization. The energy difference is equal to  $2 \times 1/2 \times W_{Hss} = -0.072$  h (-45.2 kcal/mol). The correct dissociation energy can be reached by using a broken-symmetry (BS) open-shell singlet calculation, see Figure 4, similar to the case for HF and DFT.

Figure 4.

An interesting consequence of the incorrect dissociation behavior of H<sub>2</sub><sup>+</sup> and closed-shell singlet H<sub>2</sub> is the large dependence of the vertical IP on the H–H distance. As the distance increases, SCC-DFTB underestimates the energy of H<sub>2</sub><sup>+</sup> and overestimates the energy of H<sub>2</sub>, which leads to a large drop in the vertical ionization potential of H<sub>2</sub>, see Figure 5. At long distances, CASSCF(2,2) dissociates correctly, while at 4 Å the error for SCC-DFTB is ~90 kcal/mol. PBE again gives very similar results to SCC-DFTB at larger distances.

Figure 5.

Dissociation of C<sub>2</sub> is qualitatively similar to H<sub>2</sub>, see Figure 6. The electronic ground state is a singlet and dissociation on the closed-shell singlet surface leads to a singlet instability. The correct dissociation limit is again obtained by allowing for spin-contamination in a BS open-

shell singlet solution. As expected, dissociation of  $C_2^+$  is affected by the delocalization error, as shown by the decrease in energy as the C-C distance increases, see Figure 6.

Figure 6.

The calculated vertical IP of  $C_2$  shows a similar behavior to  $H_2$ , see Figure 7. The vertical IP at the equilibrium distance is 290 kcal/mol (from ground state to ground state), which is a good match to the value calculated by CCSD(T)/PVTZ (288 kcal/mol).<sup>26</sup> However, as the C-C distance increases, the IP drops significantly.

Figure 7.

In high-temperature molecular dynamics simulations, thermal fluctuations of the C-C distance artificially decreases the ionization potential and may lead to electron donation from  $C_2$  to an electron acceptor. This would shift the dissociation curve to the  $C_2^+$  potential, with its lower barrier and repulsive potential at large C-C distances, see Figure 6. It is thus possible that the delocalization error in SCC-DFTB can lead to artificial dissociation of  $C_2$  molecules in a system with a good electron acceptor. However, even if spontaneous dissociation of  $C_2^+$  can occur, it only happens at C-C distances  $>3.7 \text{ \AA}$ , at which distance the bond is already broken. In addition, the uncatalyzed activation energy of  $C_2^+$  is 140 kcal/mol, so the rate is very low even at high temperatures (e.g., 2000 K) and is not expected to significantly affect the mechanism of  $C_2$  activation.

**Mixed-valence transition metal systems.** To show how the analysis can be used in more complex systems, we investigated parts of the electron transfer chain in cytochrome oxidase. This system includes multiple transition-metal centers, and is a possible target for multi-scale models with DFTB as a fast molecular orbital method to describe environmental effects.<sup>14</sup> In an ONIOM DFT:DFTB model, DFTB is applied to the entire system and must therefore give a reasonable description of the electronic structure of the reactive region.<sup>27</sup>

The final step in the electron-transfer chain in cytochrome oxidase is a transfer from heme *a* to heme *a3*, that together with  $Cu_B$  makes up the binuclear center.<sup>28,29</sup> Dioxygen activation is believed to take place in a mixed-valence state, where heme *a* is oxidized (+3)

while the binuclear center is reduced (heme  $a_3$ /Cu<sub>B</sub>, +2/+1). To design the computational model, heme  $a$ , heme  $a_3$  and the amino acids ligating the iron centers were taken from the crystal structure (1OCR),<sup>30</sup> see Figure 8. The Cu<sub>B</sub> center is not included as DFTB parameters for the Fe-Cu pair have not yet been developed. The propionates were removed from the heme groups and a water molecule was added to the open coordination site of heme  $a_3$  (Fe-O=2.25 Å). In the present model with two six-coordinated sites, both hemes are assigned as low-spin. This is correct for heme  $a$ , while the spin state of heme  $a_3$  is sensitive to the coordination environment and may vary during the reaction cycle.

Figure 8.

The final model consists of two heme groups in a mixed valent (+3/+2) state with an Fe–Fe distance of 13.4 Å. The total system charge is +1. Heme  $a_3$  has the highest electron affinity, and as the coupling between the iron centers is very small, the unpaired electron should reside almost fully on heme  $a_3$ . To analyze whether the present model is prone to artificial delocalization, the degree of delocalization ( $x$ ) can be estimated using equation 9. The energy difference ( $\epsilon_A - \epsilon_B$ ) is calculated as the difference in electron affinity (EA) of the oxidized heme groups in the X-ray structure ( $EA_{\text{heme}a_3} - EA_{\text{heme}a} = 0.00477$  h), and  $\gamma_{\text{FeFe}} = 0.2005$  h. Using these values in equation 9 gives  $x = 0.488$ , an almost fully delocalized solution. A single point calculation of the X-ray structure gives  $x = 0.465$  (from the Mulliken spin population analysis) In this model, the difference in electronic structure between the two heme groups is too small to give a correct description of the electronic structure. The smaller value of  $x$  in the real calculation can be attributed to a finite Fe–Fe distance (13.4 Å), spin polarization, geometry relaxation and delocalization of charge from iron to the porphyrins.

The degree of delocalization would of course change if the Cu<sub>B</sub> center was included, and with different heme  $a_3$  coordination or spin state. However, the major point is that the same analysis can be used to evaluate all these cases.

#### IV. Discussion

As the delocalization error appears in similar situations for both DFT and SCC-DFTB, it is interesting to discuss the similarities between the two descriptions. The

delocalization error in DFT is caused by the lack of cancellation between exchange and Coulomb terms. The self-interaction decreases in spatially diffuse orbitals, which, among other things, leads to a preference for delocalized states. Spatially constrained orbitals in the localized state thus leads to relatively larger delocalization errors in DFT. In DFTB the delocalization error is directly proportional to the Hubbard-like parameter. This is calculated as the second derivative of the total energy of a single atom with respect to the occupation number of the highest occupied atomic orbital. This value is higher for tightly bound orbitals, which are also spatially more localized.

At large distances, the distance dependence of the delocalization energy is dominated by the electrostatic interaction between the fragments, and this factor is similar for both DFT and SCC-DFTB. It has previously been shown that a transition metal dimer that is localized at short metal-metal distance becomes increasingly delocalized as the distance decreases.<sup>20</sup> This distance effect does not appear very often in DFT calculations, as the size of the system is still limited. However, DFTB can be applied to very large systems, and it becomes more likely that the method is applied to systems with two or more transition-metal centers at significant distance from each other, as exemplified by the calculation of the two heme groups in cytochrome oxidase above. The problem of delocalization in open-shell systems is therefore expected to be more frequent in DFTB than in standard DFT calculations.

The Hubbard model is used extensively in solid-state physics and the Hubbard repulsion term typically drives systems to insulating behavior. As previously known, the Hubbard model does not show a delocalization error for systems like  $H_2^+$ . The standard Hubbard Hamiltonian contains two terms, an inter-site hopping term and a Hubbard on-site repulsion term. If different site (orbital) energies are also included, the Hamiltonian can be written as:

$$H = -t \sum_{i,j,\sigma} c_{i\sigma}^\dagger c_{j\sigma} + \sum_{i,\sigma} \epsilon_i c_{i\sigma}^\dagger c_{i\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \quad (14)$$

where  $t$  is the hopping parameter,  $c^\dagger$  and  $c$  are the creation and annihilation operators respectively,  $\epsilon_i$  the orbital energy,  $U$  the Hubbard parameter and  $n_{i\uparrow}$ ,  $n_{i\downarrow}$  are the number of spin-up and spin-down electrons respectively. Notice that the Hubbard term only appears if there are both spin-up and spin-down electrons on a specific site, which is logical for a site (orbital) that follows the Pauli principle. This is different from the Hubbard-like term in SCC-DFTB that depends on the charge of the site.

In the Hubbard model  $H_2^+$  is a two-center system with a single electron. Including the site energies, the Hamiltonian matrix becomes:

$$\begin{bmatrix} \varepsilon_i & -t \\ -t & \varepsilon_i \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = E \begin{bmatrix} a \\ b \end{bmatrix} \quad (15)$$

for which the energies are  $E_1 = \varepsilon_i - t$  and  $E_2 = \varepsilon_i + t$  for the symmetric and asymmetric solutions respectively. As the distance increases, the size of the hopping parameter ( $t$ ) goes to zero and the energy goes to the correct dissociation limit of a proton and a hydrogen atom ( $\varepsilon_i$ ). There is also no delocalization error for multiple electron systems like  $\text{H}_2^-$ . The delocalization effect of the Hubbard term has received attention in solid-state physics,<sup>31-34</sup> but the origin of the effect is not as straightforward to analyze as in SCC-DFTB.

Semiempirical molecular orbital methods like AM1 or PM3 do not show delocalization errors for systems like  $\text{H}_2^+$  because they are based on HF. In these methods the approximations are made in the evaluation of the integrals, and no self-interaction appears as long as the same approximations are made both for Coulomb and exchange integrals.

For the SCC-DFTB calculation of  $\text{H}_2^+$ , the only term that does not lead to a delocalization (or localization) error has to be linear in charge fluctuations. However, this would lead to many other problems in SCC-DFTB, among them the lack of a defined ground state because charge separation between two atoms gives a quadratic gain in Coulomb attraction but only a linear increase in the on-site interactions.<sup>11</sup>

If an early analysis of the system indicates that it can be prone to effects of the delocalization error, it is possible to use alternative methods that in different ways handle this error. Treatments of the self-interaction in DFTB have also been outlined in the paper by Hourahine et.al.,<sup>11</sup> and by Kubar et. al.,<sup>12</sup> but these have not been included in the present investigation. Another alternative is the DFTB-VBCI approach recently proposed by Rapacioli et.al., that constrains the charge of the fragments using Lagrange multipliers, but for systems with significant exchange interactions, e.g., metal dimers, it can be difficult to set the proper charge constraints.

## V. Summary

In many systems the delocalization error in SCC-DFTB is dominated by the Hubbard-like term used to describe on-site interactions. For symmetric systems, the size of the error is directly proportional to the value of the Hubbard-like parameter. In asymmetric systems, the difference in orbital energies required to fully eliminate the delocalization error at infinite

fragment distance is equal the Hubbard-like parameter of the fragment with the highest electron affinity. In most cases, the localized state is also favored by spin polarization, a large reorganization energy, and solvent effects. These relations can be used to test if complex systems are subject to artificial delocalization.

For many systems, the delocalization error is similar in GGA DFT and SCC-DFTB. However, the delocalization error increases as a function of the distance between fragments. As DFTB is designed to handle very large systems, the problem of delocalization in open-shell systems, e.g., those with multiple transition metal centers, is expected to be more frequent compared to the smaller systems that can be treated by standard DFT.

In spin-polarized DFTB, singlet systems are subject to singlet-triplet instability, and the error in dissociation energy is proportional to the spin interaction parameter. As for other single-determinant methods, correct dissociation limits can be achieved by orbital relaxation to a broken-symmetry open-shell singlet. The delocalization error together with the singlet-triplet instability leads to an artificial dissociation path of closed shell-molecules in presence of an electron donor or acceptor, for both SCC-DFTB and DFT. In the systems investigated here, the barrier for this dissociation path is relatively high and should not be an important factor even in high-temperature molecular dynamics simulations.

## **Acknowledgments**

The authors would like to thank Keiji Morokuma, Yasuhito Ohta and Marcus Elstner for helpful discussions. Two authors (ML and SI) acknowledge a Fukui Institute for Fundamental Chemistry Fellowship. The work was in part supported by a CREST (Core Research for Evolutional Science and Technology) grant in the Area of High Performance Computing for Multi-scale and Multi-physics Phenomena from the Japan Science and Technology Agency (JST).

## Figure legends

Figure 1. Dissociation of  $\text{H}_2^+$  at Hartree-Fock (no delocalization error), NCC-DFTB (no delocalization error), and self-consistent charge SCC-DFTB (with delocalization error) levels of theory.

Figure 2. Dissociation of the C-H bond in  $\text{CH}_4^+$  and  $\text{C}_2\text{H}_6^+$  compared with dissociation of  $\text{H}_2^+$  calculated using SCC-DFTB. Convergence problems at isolated points distort the smooth potential energy surface of  $\text{C}_2\text{H}_6^+$ .

Figure 3. Dissociation of the C-H bond in  $\text{CH}_4^+$  and  $\text{C}_2\text{H}_6^+$  compared with dissociation of  $\text{H}_2^+$  calculated using PBE/6-31G(d). Convergence problems at isolated points distort the smooth potential energy surface of  $\text{C}_2\text{H}_6^+$ .

Figure 4. Dissociation curves for  $\text{H}_2$  with broken-symmetry open-shell and closed-shell singlet in spin-polarized SCC-DFTB.

Figure 5. Ionization potential of  $\text{H}_2$  calculated using CASSCF(2,2), as well as SCC-DFTB and PBE/6-31G(d) (closed-shell singlet for  $\text{H}_2$  for the two latter).

Figure 6. Dissociation profiles for  $\text{C}_2$  (BS open and closed-shell singlet) and  $\text{C}_2^+$  using SCC-DFTB.

Figure 7. Ionization potential and electron affinity (EA) of  $\text{C}_2$  calculated using SCC-DFTB (closed-shell singlet for  $\text{C}_2$ ).

Figure 8. Computational model of the heme groups in the electron transfer chain in cytochrome oxidase.

## References

1. Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys Rev B* 1998, 58, 7260-7268.
2. Kohler, C.; Seifert, G.; Gerstmann, U.; Elstner, M.; Overhof, H.; Frauenheim, T. *Phys Chem Chem Phys* 2001, 3, 5109-5114.
3. Porezag, D.; Frauenheim, T.; Kohler, T.; Seifert, G.; Kaschner, R. *Phys Rev B* 1995, 51, 12947-12957.
4. Elstner, M. *Theor Chem Acc* 2006, 116, 316-325.
5. Zheng, G. S.; Irle, S.; Morokuma, K. *Chem Phys Lett* 2005, 412, 210-216.
6. Otte, N.; Scholten, M.; Thiel, W. *J Phys Chem A* 2007, 111, 5751-5755
7. Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. L. *J Phys Chem A* 2006, 110, 13551-13559.
8. Frauenheim, T.; Seifert, G.; Elstner, M.; Hajnal, Z.; Jungnickel, G.; Porezag, D.; Suhai, S.; Scholz, R. *Phys Status Solidi B* 2000, 217, 41-62.
9. Morokuma, K. *Bull Chem Soc Jpn* 2007, 80, 2247-2261.
10. Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem Phys* 2001, 263, 203-219.
11. Hourahine, B.; Sanna, S.; Aradi, B.; Kohler, C.; Niehaus, T.; Frauenheim, T. *J Phys Chem A* 2007, 111, 5671-5677.
12. Kubar, T.; Elstner, M. *J Phys Chem B* 2010, 114, 11221-11240.
13. Rapacioli M, S. F., Scemama A, Mirtschink A. *J Chem Theory Comput* 2011, 7, 44-55.
14. Bally, T.; Sastry, G. N. *J Phys Chem A* 1997, 101, 7923-7925.
15. Grafenstein, J.; Kraka, E.; Cremer, D. *J Chem Phys* 2004, 120, 524-539.
16. Mori-Sanchez, P.; Cohen, A. J.; Yang, W. T. *J Chem Phys* 2006, 125, 201102.
17. Mori-Sanchez, P.; Cohen, A. J.; Yang, W. T. *Phys Rev Lett* 2008, 100, 146401.
18. Vydrov, O. A.; Scuseria, G. E. *J Chem Phys* 2004, 121, 8187-8193.
19. Johansson, A. J.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J Chem Phys* 2008, 129, 154301.
20. Lundberg, M.; Siegbahn, P. E. M. *J Chem Phys* 2005, 122, 224103.
21. Irle, S.; Zheng, G. S.; Wang, Z.; Morokuma, K. *J Phys Chem B* 2006, 110, 14531-14545.
22. Ohta, Y.; Okamoto, Y.; Irle, S.; Morokuma, K. *Phys Rev B* 2009, 79, 195415.
23. Zheng, G. S.; Lundberg, M.; Jakowski, J.; Vreven, T.; Frisch, M. J.; Morokuma, K. *Int J Quantum Chem* 2009, 1841-1854.
24. Zheng, G. S.; Witek, H. A.; Bobadova-Parvanova, P.; Irle, S.; Musaev, D. G.; Prabhakar, R.; Morokuma, K.; Lundberg, M.; Elstner, M.; Köhler, C.; Frauenheim, T. *J Chem Theory Comput* 2007, 3, 1349-1367.
25. Gaussian Development Version, M. J. F., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J.

- W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
26. Watts, J. D.; Bartlett, R. J. *J Chem Phys* 1992, 96, 6073-6084.
  27. Lundberg, M.; Sasakura, Y.; Zheng, G. S.; Morokuma, K. *J Chem Theory Comput* 2010, 6, 1413-1427.
  28. Blomberg, M. R. A.; Siegbahn, P. E. M. *Mol Phys* 2010, 108, 2733-2743.
  29. Kaila, V. R. I.; Verkhovski, M. I.; Wikström, M. *Chem Rev* 2010, 110, 7062-7081.
  30. Yoshikawa, S.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yamashita, E.; Inoue, N.; Yao, M.; Fei, M. J.; Libeu, C. P.; Mizushima, T.; Yamaguchi, H.; Tomizaki, T.; Tsukihara, T. *Science* 1998, 280, 1723-1729.
  31. Srinivasan, B.; Benenti, G.; Shepelyansky, D. L. *Phys Rev B* 2003, 67, 205112.
  32. Garg, A.; Krishnamurthy, H. R.; Randeria, M. *Phys Rev Lett* 2006, 97, 046403.
  33. Gulacsi, Z. *Phys Rev B* 2008, 77, 245113.
  34. Henseler, P.; Kroha, J.; Shapiro, B. *Phys Rev B* 2008, 77, 075101.