

---

# Toward a BSSE-Free Description of Strongly Interacting Systems

---

G.J. HALÁSZ,<sup>1</sup> Á. VIBÓK,<sup>2</sup> S. SUHAI,<sup>3</sup> I. MAYER<sup>4</sup>

<sup>1</sup>*Institute of Mathematics and Informatics, University of Debrecen, P.O. Box 12, H-4010 Debrecen, Hungary*

<sup>2</sup>*Institute of Theoretical Physics, University of Debrecen, P.O. Box 5, H-4010 Debrecen, Hungary*

<sup>3</sup>*Department of Molecular Biophysics, German Cancer Research Centre, Im Neuenheimer Feld 280, D69120 Heidelberg, Germany*

<sup>4</sup>*Institute of Chemistry, Chemical Research Center, Hungarian Academy of Sciences, P.O. Box 17, H-1525 Budapest, Hungary*

Received 23 April 2001; accepted 17 September 2001

DOI 10.1002/qua.10286

---

**ABSTRACT:** The so-called “chemical Hamiltonian approach” (CHA) gives perfect *a priori* BSSE-free description of weak intermolecular interactions, but has been found inappropriate for describing strong interactions taking place within a molecule. Here, we propose a simple modification of the CHA/F BSSE-free SCF method, which retains all the good properties of the CHA/F method for the intermolecular case but can be used also to describe covalent and ionic interactions. This is vital for calculating whole potential surfaces of chemical reactions in a consistent manner, which was found impossible by using the *a posteriori* counterpoise correction method. Model calculations are presented for covalent and ionic chemical bonds and for a rare gas–proton system. © 2002 Wiley Periodicals, Inc. *Int J Quantum Chem* 89: 190–197, 2002

**Key words:** basis set superposition error (BSSE); covalent and ionic bonds; chemical Hamiltonian approach (CHA); BSSE-free SCF theory

---

## 1. Introduction

The notion of basis set superposition error (BSSE) is mostly considered in the framework of interatomic and intermolecular interactions, where it causes too low minima on the calculated potential surfaces because the internal energy of

each molecule is somewhat lowered by utilizing the basis orbitals of the partner. However, the BSSE problem is encountered also when considering a chemical reaction. In fact, similarly to the case of intermolecular complexes, the wave function corresponding to the transition structure (TS) of a chemical reaction also has more degrees of variational freedom than do those of the free reactant (or product) molecules together; thus, one can expect the TS energy to be also artificially lowered by BSSE. Fur-

Correspondence to: Á. Vibók; e-mail: vibok@cseles.atomki.hu

ther, many reactions exhibit weak prereaction and/or postreaction complexes, which essentially are regular intermolecular van der Waals systems. To be able to treat such reactions—especially if their dynamics is to be considered in detail—one needs a scheme in which all parts of the potential surface are treated using exactly the same technique. (Otherwise, one cannot avoid discontinuities of the calculated potential surfaces.) For the prereaction (postreaction) complexes, one has to introduce BSSE correction, without which even the existence of these minima cannot be established reliably. At TS of an “extraction” reaction in which an atom or a group is transferred from one molecule to another, one cannot assign the “moving” atom (group) to either of the reactant (product) molecules and, therefore, has to consider the system as consisting of at least three subunits. However, these subunits are not the same as those one should consider for the prereaction and postreaction complexes. This problem has recently been studied in detail in Ref. [1]. It was concluded that no consistent version of the *a posteriori* counterpoise correction (CP) method [2] can be constructed to deal with this problem, and the only remedy (in addition of using practically complete saturated basis sets) is to develop a version of the *a priori* chemical Hamiltonian approach (CHA) [3] in which *every atom* is treated as an independent subunit. It is not usual to consider BSSE *within* the molecule, but one has to conclude that such a treatment is inevitable looking at, for example, the different interactions the atom(s) transferred in a reaction exhibit in the reactants in the prereaction and post-reaction complex and at the TS.

According to this conclusion, we started to study the behavior of the CHA method for intramolecular interactions. For covalent and ionic bonds (as, in the simplest case, those in  $H_2$  or  $N_2$  molecules and in HF molecule, respectively), the results were discouraging: When the basis sets increased, the CHA scheme gave a slower convergence to the basis set limit as did the uncorrected or CP-corrected SCF methods. The respective MP2 scheme was also developed, and the results obtained for the correlation energy were more satisfactory. In such a situation, we met the challenge of understanding why the CHA-SCF procedures give a perfect *a priori* BSSE-free description of weak intermolecular interactions, but are inappropriate for describing the strong interactions taking place within a molecule. Reconsidering our earlier results [4] obtained for the rare gas-proton interaction, we concluded that

similar problems occurred for these ionic systems, at least in the case of some basis sets. Also, in few cases of intermolecular interactions some minor problems occurred at the (unphysically) short intermolecular distances and largest basis sets with diffuse functions.

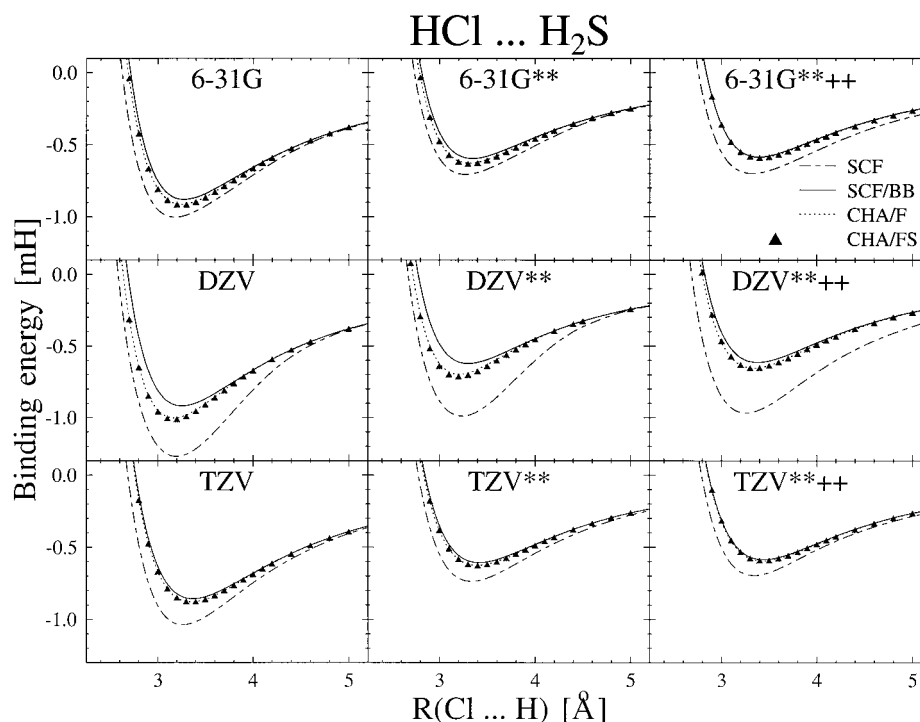
Here, we describe a new version of the CHA applied to the Fockian (CHA/F) scheme [5–7] that gives practically coinciding results with the original CHA/F method in the case of intermolecular interactions, retains all the good properties of it for treating weak complexes, but can be used also to describe covalent and ionic interactions. We shall denote it as CHA/FS, where “S” refers to the “strong” interactions. It is based on the distinction between the role the occupied and virtual monomer orbitals are playing in causing BSSE. In fact, in the case of weak interactions the interacting molecules retain their individuality to a great extent, which means that the virtual orbitals of the free monomers are weakly populated even within the complex. Therefore, only terms causing BSSE-type delocalizations from the occupied orbitals of one molecule to the basis orbitals of the partner molecule have any real significance in determining BSSE. For the case of strong intramolecular interactions, the virtual orbitals of the free monomers also become significantly populated, and it represents essentially a new problem: how exactly they should be treated to simultaneously conserve the BSSE-free properties of their CHA/F scheme and not cause unphysical restrictions on the interactions related to the virtual monomer orbitals.

Here, we describe the new CHA/FS procedure permitting us to solve the above problem. The results of sample calculations performed on five different systems ( $HCl \dots H_2S$ ,  $HF \dots HF$ ,  $H_2$ ,  $HF$ , and  $He \dots H^+$ ) are also presented and discussed.

---

## 2. CHA/FS Scheme

The basic idea of the CHA/F scheme was to construct a combined monomer Fockian by using a projection operator technique, which has the property that its solutions in the overlapping supermolecule basis recover exactly all the eigenvalues and eigenvectors of the free monomer SCF problems. Then, one adds the terms describing actual “true” intermolecular interactions, thus providing BSSE to be absent. This method keeps consistency of the supermolecule results with the original free mono-



**FIGURE 1.** Potential curves of the HCl . . . H<sub>2</sub>S dimer calculated in nine different basis sets as a function of the bond length. The curves illustrate four different approximations to the binding energy: SCF, standard Hartree–Fock; SCF/BB, standard Hartree–Fock energy corrected by the *a posteriori* Boys–Bernardi counterpoise method; CHA/F, the original *a priori* BSSE-free theory based upon the CHA; CHA/FS, a new version of the *a priori* BSSE-free theory based on the CHA.

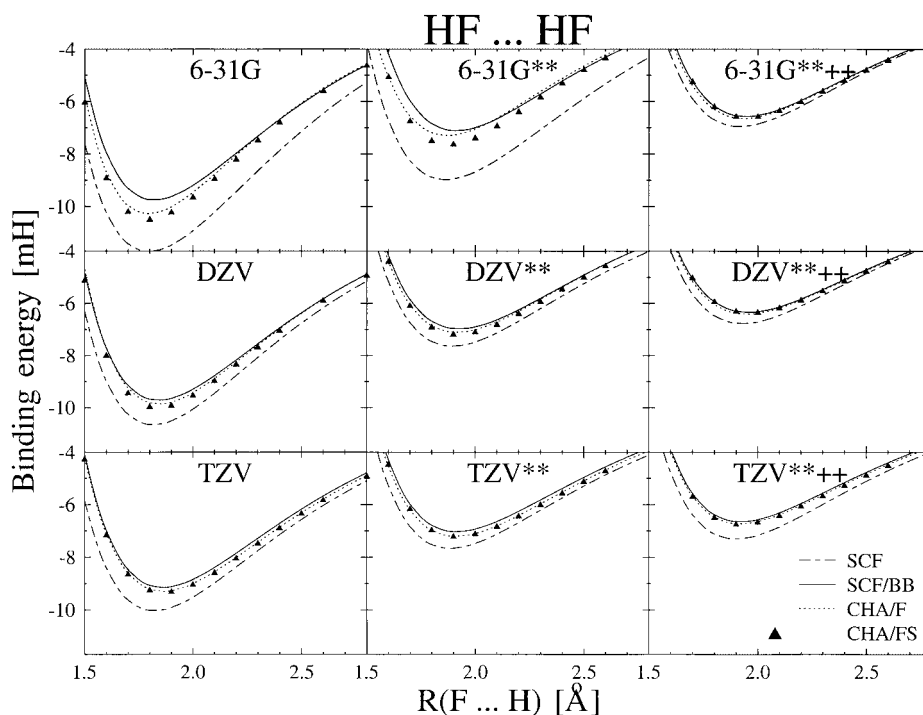
mer calculations performed in the monomer basis sets, as opposed to the CP scheme, in which the monomer calculations are adjusted to the supermolecule basis.

The way of constructing the CHA/F scheme was such that not only the occupied monomer orbitals but also the virtual ones represent its eigenvectors, that is, the projections used prevent them also from the “BSSE-type” delocalizations to the partner molecule(s). This was not really necessary but did not cause any trouble, as practically the virtual monomer orbitals remain almost empty in the weak intermolecular complexes. At the same time, the construction of the CHA Fockian was the simplest in this manner. However, as noted in Section 1, the stronger the interaction, the larger is the occupancy of the virtual monomer orbitals that are empty in the free monomers. If one takes care of the absence of any BSSE-type delocalization from the occupied monomer orbitals to the virtual ones, then the occupancy of the latter is a result of the “true” intermolecular interactions. Consequently, there is no meaning to consider any delocalization from the

virtual orbitals as being connected with BSSE. (These delocalizations obviously represent some higher-order effects that are negligible in the true intermolecular complexes.)

According to this discussion, there is no need to include the monomer virtual orbitals in the transformations devoted to exclude BSSE effects. This means that one cannot perform the analysis and programming directly in terms of the monomer atomic orbital basis, but at first all the quantities should be expressed in terms of the free monomer molecular orbitals (MOs), as done in Ref. [5]. Then, one can modify the Fockian in the monomer’s MO basis to exclude BSSE appropriately. Finally, one transforms back to the AO basis to solve the modified Hartree–Fock–Roothaan equations. In the following considerations, we consider the case of two interacting monomers; the generalization for an arbitrary number is straightforward (cf. [3, 8]).

Consider the “ghost orbitals” problem of monomer *A* in the basis set of the supersystem *AB*. In the special case when the monomer occupied orbitals



**FIGURE 2.** Potential curves of the HF . . . HF dimer calculated in nine different basis sets as a function of the bond length. The curves illustrate four different approximations to the binding energy: SCF, standard Hartree–Fock; SCF/BB, standard Hartree–Fock energy corrected by the *a posteriori* Boys–Bernardi counterpoise method; CHA/F, the original *a priori* BSSE-free theory based upon the CHA; CHA/FS, a new version of the *a priori* BSSE-free theory based on the CHA.

$\varphi_i^{A\text{occ}}$  are *exact* eigenvectors<sup>1</sup> of the respective monomer Fock operator  $\hat{F}_A$ , then one has the equality

$$\hat{F}_A \varphi_i^{A\text{occ}} = \varepsilon_i^A \varphi_i^{A\text{occ}}. \quad (1)$$

In this special case, the extension of the basis with the orbitals of monomer *B* does not change  $\varphi_i^{A\text{occ}}$ —the *exact* solution cannot be “improved” any more. In other words, BSSE is absent for an exact orbital. Multiplying Eq. (1) with an orbital of monomer *B*, one has

$$\langle \varphi_j^B | \hat{F}_A | \varphi_i^{A\text{occ}} \rangle = \varepsilon_i^A \langle \varphi_j^B | \varphi_i^{A\text{occ}} \rangle. \quad (2)$$

Taking into account the equality  $\varepsilon_i^A = \langle \varphi_i^{A\text{occ}} | \hat{F}_A | \varphi_i^{A\text{occ}} \rangle$ , Eq. (1) can also be written as

$$\langle \varphi_j^B | \hat{F}_A | \varphi_i^{A\text{occ}} \rangle = \langle \varphi_j^B | \varphi_i^{A\text{occ}} \rangle \langle \varphi_i^{A\text{occ}} | \hat{F}_A | \varphi_i^{A\text{occ}} \rangle. \quad (3)$$

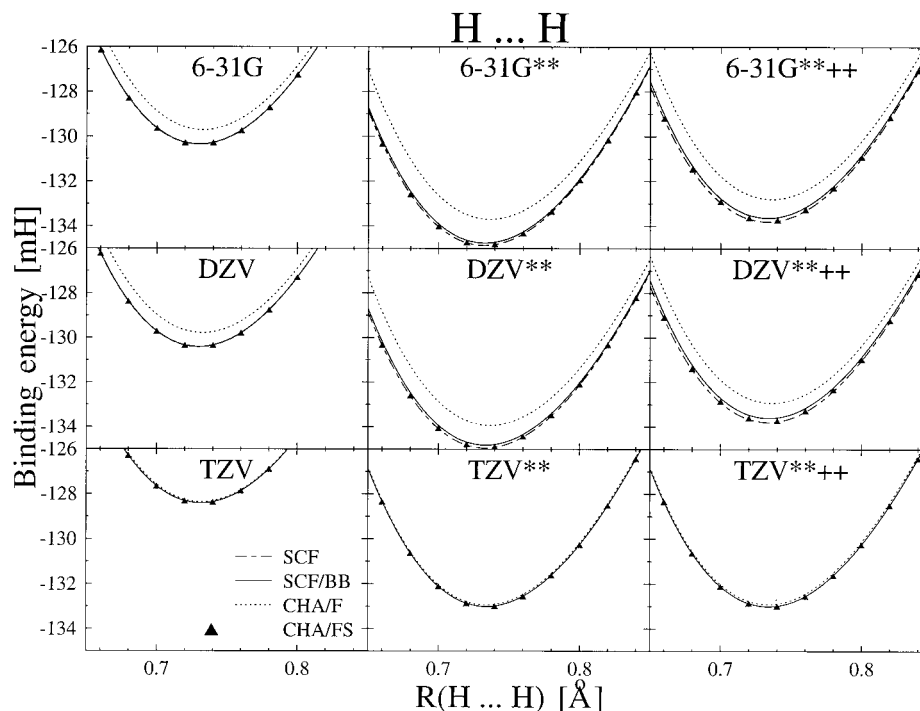
<sup>1</sup>Except for special cases (e.g., exact hydrogenic orbitals if *A* is a hydrogen atom), this is possible only in the limit of a complete basis set.

Further, the exact monomer molecular orbitals diagonalize the monomer Fock operator, that is,  $\langle \varphi_k^{A\text{occ}} | \hat{F}_A | \varphi_i^{A\text{occ}} \rangle = \varepsilon_i^A \delta_{ki}$ . Utilizing this, Eq. (3) can be extended as

$$\langle \varphi_j^B | \hat{F}_A | \varphi_i^{A\text{occ}} \rangle = \sum_{k \in A} \langle \varphi_j^B | \varphi_k^{A\text{occ}} \rangle \langle \varphi_k^{A\text{occ}} | \hat{F}_A | \varphi_i^{A\text{occ}} \rangle. \quad (4)$$

Equality (4) represents the most general condition that the monomer orbital  $\varphi_i^{A\text{occ}}$  remains unchanged when the basis set is extended. In practice, this is not fulfilled, and so causes energy lowering both in the ghost orbitals calculations and within an actual complex. In accord with the basic idea of the CHA scheme, one has to *enforce* the fulfillment of this equality, even if a finite basis is used, by replacing the left side of Eq. (4) with its right side during the correct supersystem calculations. This prevents any BSSE-caused delocalizations from the occupied orbitals to the basis orbitals of the partner monomer.

In actual calculations, this can be accomplished by computing a correction matrix that is to be



**FIGURE 3.** Potential curves of the  $H_2$  molecule calculated in nine different basis sets as a function of the bond length. The curves illustrate four different approximations to the binding energy: SCF, standard Hartree–Fock; SCF/BB, standard Hartree–Fock energy corrected by the *a posteriori* Boys–Bernardi counterpoise method; CHA/F, the original *a priori* BSSE-free theory based upon the CHA; CHA/FS, a new version of the *a priori* BSSE-free theory based on the CHA.

added to that block of the LCAO Fockian that corresponds to all the orbitals of monomer  $B$  from one side and the occupied orbitals of monomer  $A$  from another:

$$\mathbf{F}^{BA_{\text{occ}}} \Rightarrow \mathbf{F}^{BA_{\text{occ}}} + (\mathbf{S}^{BA} \mathbf{F}_A^{AA_{\text{occ}}} - \mathbf{F}_A^{BA_{\text{occ}}}). \quad (5)$$

Here,  $\mathbf{F}$  is the total supersystem Fock matrix,  $\mathbf{S}$  is the intermonomer overlap matrix in the monomer MO basis, and  $\mathbf{F}_A$  is the Fock matrix of monomer  $A$ , calculated by the actual intramonomer block of the supersystem density matrix. (An alternative could be to compute  $\mathbf{F}_A$  by using the density matrix of the free monomer.)

Analogous manipulations should be performed for monomer  $B$  as well. The difference with the previous versions of the CHA/F scheme is only in restricting the correction matrix to the blocks with occupied monomer orbitals. (Because of this distinction, the calculations of the present CHA/FS method must be explicitly performed in the basis of the free monomers' MOs, while in the usual CHA/F scheme all calculations could be done directly in the AO basis.) Similarly to the other

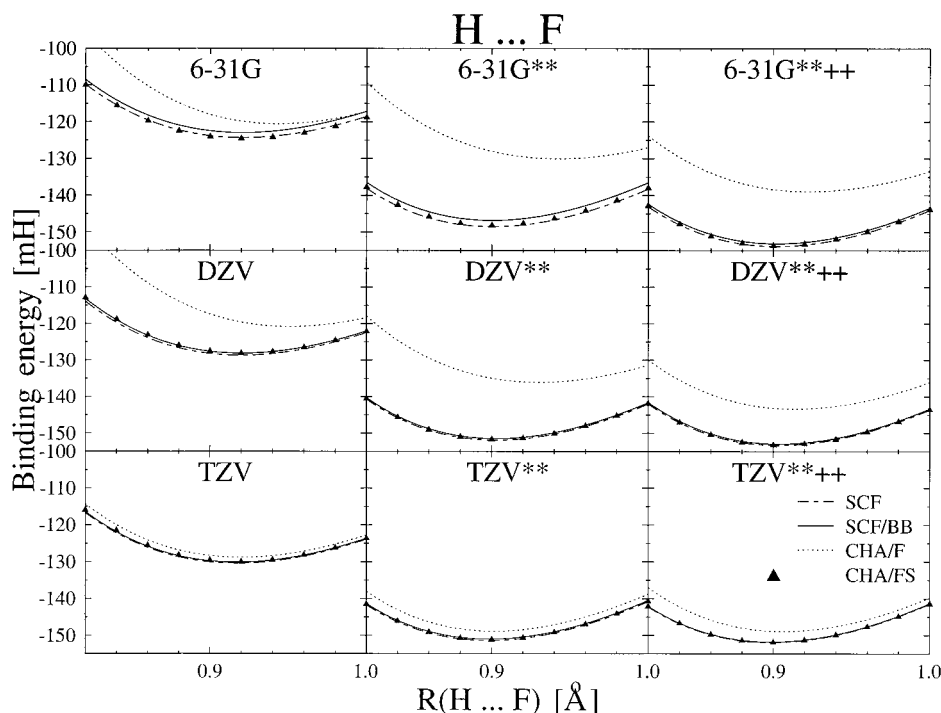
CHA/F schemes, the present Fockian is also not Hermitian,<sup>2</sup> and the energy, as usual, should be computed with the original (conventional) Fockian by using the customary equality

$$E = \frac{1}{2} \text{tr}[(\mathbf{F} + \mathbf{h})\mathbf{P}]. \quad (6)$$

### 3. Sample Calculations

A FORTRAN program has been written to realize the CHA/FS algorithm defined above. This code is based partly on our earlier work [6, 7] and also uses the same modified version of the Hondo-8 system [9] to get the necessary one- and two-electron integrals. Calculations have been carried out on a Pentium PC running Linux. In the calculations presented here, Pople's 6-31G and Dunning's valence double zeta (DZV) and valence triple zeta (TZV) basis sets were used with the standard parameters generated by the HONDO-8 system, ex-

<sup>2</sup>BSSE is not a physical phenomenon, so there exists no Hermitian operator that could be associated with it.



**FIGURE 4.** Potential curves of the HF molecule calculated in nine different basis sets as a function of the bond length. The curves illustrate four different approximations to the binding energy: SCF, standard Hartree–Fock; SCF/BB, standard Hartree–Fock energy corrected by the *a posteriori* Boys–Bernardi counterpoise method; CHA/F, the original *a priori* BSSE-free theory based upon the CHA; CHA/FS, a new version of the *a priori* BSSE-free theory based on the CHA.

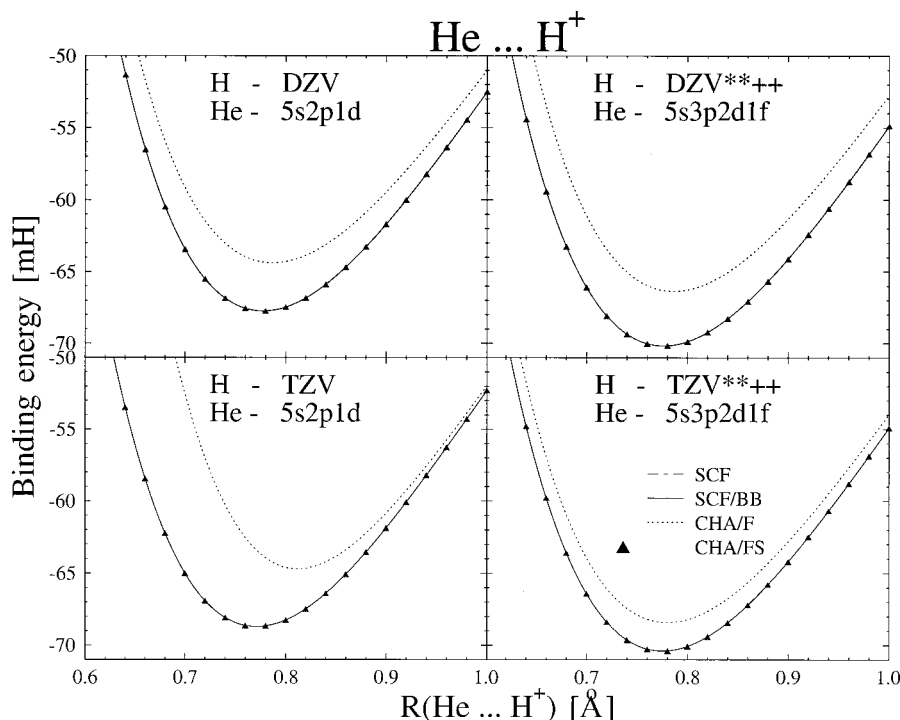
cept for helium. For helium, the basis sets applied ( $5s2p1d, 5s3p2d1f$ ) were taken from Ref. [10]. For the polarized basis sets (denoted\*\*), the exponents were taken as the default values of the Gaussian 98 system [11]. (There are no standard exponents given by HONDO-8 for this basis.) The actual values were: hydrogen ( $p$ -orbitals), 1.1; fluorine ( $6d$ -orbitals), 0.9. In the DZV\*\* and TZV\*\* basis sets, the hydrogen  $p$ -exponent was 1.0 while the fluorine  $d$ -orbital exponent was 1.62. In all cases, the diffuse functions (denoted ++) had the following exponents: hydrogen ( $s$ -orbitals), 0.036; fluorine ( $sp$ -orbitals), 0.1076. The geometries used for hydrogen-bonded complexes were the same as in our previous articles [7, 9].

In the figures, we present inter- and intramolecular potential curves at the SCF level of theory. Four different methods have been compared: the uncorrected SCF curve, the result obtained by the *a posteriori* counterpoise correction scheme of Boys and Bernardi (SCF/BB), and those given by the two different variants of the CHA/F theory. These are the straightforward original intermolecular CHA/F

method [7] and the present CHA/FS scheme, described in Section 2.

First, we checked the behavior of the new method on some typical intermolecular complexes. Figure 1 shows the potential curves of the complex  $\text{HCl} \dots \text{H}_2\text{S}$  in nine different basis sets. As shown, the original and new CHA/FS methods give practically undistinguishable results for this weak hydrogen bond. For the strongest hydrogen bond of the  $\text{HF} \dots \text{HF}$  system, shown in Figure 2, some minor discrepancies can be observed between the two CHA schemes, except in the cases of the largest basis sets of VTZ quality. For the smaller basis sets, it is difficult to decide what CHA curves could be preferred as the differences are small; in the case of the 6-31G\*\*, the new curve looks more balanced than the original CHA/F, judging on the comparison with the CP-corrected curve. These results indicate that the new scheme does not, indeed, destroy the good behavior of the CHA/F method as far as weak interactions are considered.

Concerning the strong interactions for which our new method has been designed, we present here



**FIGURE 5.** Potential curves of the  $\text{He} \dots \text{H}^+$  system calculated in four different basis sets as a function of the bond length. The curves illustrate four different approximations to the binding energy: SCF, standard Hartree–Fock; SCF/BB, standard Hartree–Fock energy corrected by the *a posteriori* Boys–Bernardi counterpoise method; CHA/F, the original *a priori* BSSE-free theory based upon the CHA; CHA/FS, a new version of the *a priori* BSSE-free theory based on the CHA.

three different covalent, ionic, and rare gas–proton systems as ( $\text{H}_2$ , HF, and  $\text{He} \dots \text{H}^+$ ). The aim of the present work consists of the comparison of the different schemes, facilitated by the use of such simple model systems. Figures 3–5 show the results obtained for these systems using different basis sets.

Inspecting the curves, one immediately sees that all methods considered give results closer to each other than the original CHA/F. This observation is in accord with the discussion given in Section 1: The original CHA/F method is not appropriate to describe interactions between atoms and ions that are connected by covalent or ionic bonding. At the same time, the new CHA/FS procedure always gives curves close to the CP-corrected and/or uncorrected ones, indicating that the basic problem encountered with the CHA/F method when applied to strong interactions has been appropriately solved by the new approach.

While all the BSSE-corrected curves are always above the uncorrected SCF one, it is difficult to predict whether the CHA/FS curve will be above or below the CP-corrected one in the given system and basis sets.

It may appear surprising at first that there is no BSSE correction for the case of the  $\text{H}_2$  molecule in the basis sets without polarization functions but there is one if polarization functions are also used. Adding polarization functions, however, does not improve the description of the  $1s$  state of the hydrogen atoms, but can serve for decreasing the energy in the ghost orbitals calculations. (The atomic energy of the free H atom calculated by the 6-31G and DZV basis sets is only slightly below  $-0.498$  Hartree, compared with the exact, so there is significant room for its improvement by utilizing the orbitals of the partner atom. However, the basis orbitals of the valence part of the basis are too compact for that purpose.) At the SCF level of theory, the BSSE in the  $\text{H}_2$  system practically disappears when turning to a TZV basis set, and the three methods (uncorrected SCF, SCF/BB, and CHA/FS) give undistinguishable results. (Even the original CHA/F gives curves close to the other methods.)

The inadequacy of the original CHA/F method for treating strong interactions is even more apparent for the HF molecule and the  $\text{He} \dots \text{H}^+$  systems, displayed in Figures 4 and 5. On the HF curves, the

new method gives results comparable to the SCF/BB ones; one hardly could give preference to either of them, as the differences are small. (However, it is interesting that in some cases the CHA/FS scheme indicates somewhat larger corrections than the SCF/BB one. But, this overcorrection, if any, can in no way be compared with the ill-behaving results of the original CHA/F.) There is practically no BSSE in the case of the He...H<sup>+</sup> system because proton bears no electrons and the helium basis used was good enough to describe the helium 1s orbital to a near-HF accuracy. Despite this, the original CHA/F indicates a large spurious BSSE correction.

#### 4. Conclusions

The present preliminary study indicates that we succeeded in identifying the main factor making the original CHA-SCF schemes inadequate for treating intramolecular effects. It may be expected that the present approach can be generalized without significant difficulties, at least to the MP2 level of theory, for treating electron correlation. Fortunately, the BSSE content of correlation contributions higher than second order is usually negligible, so a well-designed BSSE-free SCF+MP2 theory can be used to guess the BSSE content in the total uncorrected energy of a complex system. This indicates that with the present results we probably could get rid, at least in principle, of the major obstacle for calculating BSSE-corrected potential surfaces of chemical reactions by using the same technique at all configurations of the system: reactants, pre- and postreaction complexes, transition region, and products. To be able to perform actual calculations of that type, it will be necessary to introduce the modifications discussed here in the more demanding CHA/CE SCF formalism [5, 12] and the CHA-MP2 theory [13]. We hope this task can also be performed soon.

#### ACKNOWLEDGMENTS

This work was supported by OTKA Grants T25369, T34812, and FKFP 0498/2000. The authors thank the European Union for a fellowship under Contract BMH4-CT96-1618 that funded their study visit to Heidelberg.

#### References

1. Lendvay, G.; Mayer, I. *Chem Phys Lett* 1998, 297, 365.
2. Boys, S. B.; Bernardi, F. *Mol Phys* 1970, 19, 553.
3. Mayer, I. *Int J Quantum Chem* 1983, 23, 341; Mayer, I. *Int J Quantum Chem* 1998, 70, 41.
4. Mayer, I.; Vibók, Á. *Acta Phys Hung* 1991, 70, 403.
5. Mayer, I.; Surján, P. R.; Vibók, Á. *Int J Quantum Chem Quantum Chem Symp* 1989, 23, 281.
6. Mayer, I.; Vibók, Á. *Int J Quantum Chem* 1991, 40, 139; Vibók, Á.; Mayer, I. *Int J Quantum Chem* 1992, 43, 801; Mayer, I.; Vibók, Á.; Halász, G.; Valiron, P. *Int J Quantum Chem* 1996, 57, 1049; Halász, G.; Vibók, Á.; Suhai, S. *Int J Quantum Chem* 1998, 68, 151.
7. Valiron, P.; Vibók, Á.; Mayer, I. *J Comput Chem* 1993, 14, 401.
8. Halász, G.; Vibók, Á.; Valiron, P.; Mayer, I. *J Phys Chem* 1996, 100, 6332.
9. Dupuis, M.; Farazdel, A. HONDO-8, from MOTECC-91; IBM Corporation Center for Scientific & Engineering Computations: Kingston, NY, 1991.
10. Mayer, I.; Vibók, Á.; Valiron, P. *Chem Phys Lett* 1994, 224, 166.
11. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
12. Mayer, I.; Surján, P. R. *Int J Quantum Chem* 1989, 36, 225.
13. Mayer, I.; Valiron, P. *J Chem Phys* 1998, 109, 3360; Halász, G.; Vibók, Á.; Mayer, I. *J Comput Chem* 1999, 20, 274.