Theoret. Chim. Acta (Berl.) 33, 7—17 (1974)

# A Theory of Molecules in Molecules

## V. Study of the Relative Stability of Some Para-, Meta-, and Ortho-Disubstituted Benzene Molecules

Wolfgang von Niessen
Lehrstuhl für Theoretische Chemie, Technische Universität München, 8 München 2, Germany

Received September 13, 1973

SCF wave functions have been calculated using a minimal atomic basis set of Gaussian lobe functions for the para-, meta-, and ortho-forms of the molecules $C_6H_4XY$, where X, Y can be either of CN, OH, or F. It is found that in all cases the total energies increase in the sequence meta-, para-, ortho-compound. For the molecules containing the CN group the energy differences are extremely small (0.1–1 kcal/mole) for the other molecules they are one to two orders of magnitude larger. The reliability of these results is discussed. The theory of molecules in molecules is applied to these cases. The wave function of $C_6H_4XY$ is constructed from the fragments $C_6H_5X$ and HY by transferring some of the localized orbitals of the wave functions of the fragments and recalculating the orbitals in the region of interaction. For the molecules containing the CN group the energy differences are too small so that they are not correctly reproduced except by the most exact calculations, which involve no approximations other than the transfer of localized orbitals. For the other molecules satisfactory results are obtained.

*Key words:* Transferability of localized orbitals – Wave functions from fragment wave functions – $C_6H_4XY$, para-, meta-, and ortho-forms of ~

### 1. Introduction

The present investigations serve to examine the relative stability of some para-(*p*), meta-(*m*), and ortho-(*o*) disubstituted benzene molecules, $C_6H_4XY$. Three different substituents have been investigated: CN, OH, and F. *Ab initio* SCF wave functions have been calculated for these molecules. In addition the theory of molecules in molecules developed in previous articles [1, 2] has been applied to study these systems and their relative stability. This theory of molecules in molecules has been previously applied to a number of small molecules [1, 2] and to the hydrogen bonded systems $(FH)_2$ [3] and $NH_3 \cdot H_2O$ [4]. But it is equally well applicable to study benzene derivatives, because it can be regarded as a quantum chemical substitution method. The method will only be described in a few summarizing sentences which appear necessary for the understanding of the present investigations. For a detailed description and references to related approaches see the Refs. [1] and [2].

The wave function of the molecule to be calculated is constructed from the wave functions of fragment molecules by transferring some of the localized molecular orbitals (LMO's) of the fragments and recalculating the orbitals in the region of interaction, where the electronic rearrangement caused by the different chemical surrounding in the molecule to be formed must be allowed for. Chemical and physical intuition determines which molecular orbitals can be transferred unaltered from the fragments – they form the fixed core – and which are recalculated to describe optimally the new bonds formed and their immediate neighbourhood. The approximate transferability of LMO's, which is used here, has been the subject of a number of investigations [1–17]. A projection operator [18] is used to obtain molecular orbitals (MO's) in the region of interaction which are orthogonal to the core orbitals. In order to give the possibility to save computational time further approximations have been introduced: 1) The expansion of the MO's in the region of interaction is truncated to include only those basis functions which are regarded as essential for their expansion (this subset is denoted by $\Gamma$). Basis functions whose centers are distant from the region of interaction will contribute only insignificantly and will be excluded from the entire set of basis functions. 2) The LMO's in the projection operator for orthogonality to which orthogonality can be expected because of their spatial separation from the region of interaction are taken out of the projection operator and the expansion of the remaining LMO's is restricted to the same subset $\Gamma$ of basis functions mentioned above. 3) The nonorthogonality of the MO's is neglected. 4) The Coulomb integrals between a LMO transferred for one of the fragments and a LMO transferred for the other one are calculated by a point charge approximation for every orbital and the corresponding exchange integrals are neglected. These approximations lead to a reduction of the dimension of the matrices to be diagonalized and to the neglect of a part of the basic integrals. The theory of molecules in molecules (MIM) has been applied in a number of approximate forms. The investigations thus also serve the purpose to demonstrate the applicability of the method, to justify the approximations described above and examine the range of their validity in the case of the disubstituted benzenes.

## 2. Applications

The geometries used in the calculations for the *p*-, *m*-, and *o*-forms of the molecules $C_6H_4XY$ with $X, Y = CN$, OH, or F were partly taken from the literature [19], partly idealized geometries have been used. The geometry of the benzene molecule has been used for all substituted products. The bond distances of the CN, OH, and F substituents are taken from the monosubstituted compounds as given in the literature and were also used in the disubstituted molecules ($-CN: R(C–C) = 1.419$ Å, $R(C–N) = 1.158$ Å; $-F: R(C–F) = 1.30$ Å; $-OH: R(C–O) = 1.36$ Å, $R(O–H) = 0.956$ Å, $\sphericalangle COH = 120°$). The COH angle was chosen to be 120° with all atoms lying in a plane. The geometry of the *p*-dihydroxybenzene has the two H atoms in *cis*-position. The *m*- and *o*-dihydroxybenzene are derived from this structure by rotating one of the OH groups (the OH group on C atom $C_4$) counterclockwise into these positions keeping the OH group on C atom $C_1$ fixed. In this way no intramolecular hydrogen bonds can be formed in the *o*-compound, which would introduce a factor alien to the question to be investigated, and the H atoms of the OH groups have the least sterical interaction. The same geometries for the OH group have been used in the compounds $C_6H_4CNOH$ and $C_6H_4OHF$. (The numbering system for the atoms in the molecules is the same as in benzene. The H atoms carry the same indeces as the C atoms to which they are bound.) No bond distances or angles were varied in the calculations. The basis set used in the calculations is a minimal atomic Gaussian lobe function basis, whose parameters are taken from the literature [20, 21]. The contraction coefficients were determined from calculations on small molecules. On the atoms C, N, O, and F 3 functions of *s*-type [20] contracted to two functions and one function of *p*-type [20] and on the H atoms a single *s*-type function [21] has been used. The basis set is certainly of poor quality and it therefore does not allow to vary bond distances and angles. The computed total SCF energies are listed in Table 1. The energy differences $\Delta_1 = E(\text{para}) - E(\text{meta})$ and $\Delta_2 = E(\text{ortho}) - E(\text{para})$ are relatively small quantities ranging from 0.000248 a.u. = 0.16 kcal/mole to 0.035285 a.u. = 22.2 kcal/mole, but they are without any exception positive quantities. The *m*-disubstituted benzenes are always the most stable compounds in the cases considered and the *o*-disubstituted compounds are always the least stable ones having the highest total energies. An explanation for the latter fact could be the (repulsive) interaction between the two substituents in *o*-position, which are spatially close to each other, but there is no apparent explanation of why the *m*-compounds are always more stable than the *p*-compounds. The reliability of the wave functions computed with the present basis set is small and one cannot state with certainty that the observed results are true, i.e. will be observed as well if large basis sets are used for the calculations, especially because some of the energy differences are extremely small. It can only be argued that the fact that all wave functions have been calculated with the same basis set, which leads to an equal description of all molecules, speaks favourably for the reliability of the answer obtained. But this argumentation can only be regarded as valid if the energy differences are of the order of 10 kcal/mole. A variation of bond distances and angles could also lead to energy changes which could be larger than the differences observed and could result in other orders of stability as measured by the total SCF energies. This question can again not be investigated due to the quality of the basis set.

The energy differences acquire their smallest values if one of the substituents is the CN functional group and the largest value if one of the substituents is the F atom. (22.2 kcal/mole in the case of $C_6H_4F_2$, Table 1). No explanation will be attempted why the energy differences are so small for the cyano-compounds and

Table 1. Total SCF energies for the *p*-, *m*-, and *o*-forms of the disubstituted benzenes $C_6H_4XY$ with X, Y = CN, OH, F. $\Delta_1 = E$ (para – $E$ (meta), $\Delta_2 = E$ (ortho) – $E$ (para). (All values in atomic units)

| Molecule | $E$(para) | $E$(meta) | $E$(ortho) | $\Delta_1$ | $\Delta_2$ |
|---|---|---|---|---|---|
| $C_6H_4(CN)_2$ | −392.334033 | −392.334281 | −392.332991 | 0.000248 | 0.001042 |
| $C_6H_4CNOH$ | −375.817025 | −375.817745 | −375.813919 | 0.000720 | 0.003106 |
| $C_6H_4CNF$ | −397.892863 | −397.893164 | −397.891027 | 0.000301 | 0.001836 |
| $C_6H_4(OH)_2$ | −359.285606 | −359.302633 | −359.273415 | 0.017027 | 0.012191 |
| $C_6H_4OHF$ | −381.358415 | −381.378588 | −381.341240 | 0.020173 | 0.017175 |
| $C_6H_4F_2$ | −403.430215 | −403.454093 | −403.394930 | 0.023878 | 0.035285 |

two orders of magnitude larger for the molecules containing only the functional groups OH and F. The energy differences in the cyano compounds are so small that they are thermodynamically completely irrelevant, for the fluorine compounds on the other hand the energy differences could play a role in their formation. However, the unanimous trend in the stability precludes a thermodynamic explanation of the *m*- or *o*- and *p*-directing property of the different substituents.

The observations made above on the relative stability of the disubstituted benzene molecules will serve now as a laboratory to test the theory of molecules in molecules for its reliability to reproduce the SCF results. The theory is applied in the following manner. As fragments for the molecules $C_6H_4XY$ serve the molecules $C_6H_5X$ and HY both taken in the geometry appropriate for the molecule to be formed. For X = CN, Y can be either CN, OH, or F, for X = OH, Y can be either OH or F, and for X = F, Y can only be F. The other possible cases were not considered. These fragments $C_6H_5X$ and HY are combined to form the final molecule by transferring some of the LMO's and redetermining the MO's in the region of interaction. The two H atoms (one in each fragment) and their bond orbitals are taken out, they will be replaced by the C–Y bond orbital which is to be calculated. The MO's of the fragments were localized using the method of Boys [22] and the MO's of $\sigma$- and $\pi$-symmetry were always separately localized. The results of these calculations will be reported in another article. For every molecule three approximations have been investigated. They are denoted by MIM I, MIM II, and MIM III. In all cases the valence orbitals of the new substituent are recalculated as well as the three $\pi$-orbitals of the benzene ring. In MIM I the LMO's which are nearest and next nearest neighbour orbitals to the bond orbital to be calculated are modified, in MIM II and MIM III only nearest neighbour orbitals are modified. It is to be noted that in the molecules $C_6H_4(OH)_2$, $C_6H_4OHF$, and $C_6H_4F_2$ there is in the *o*-form one more next nearest neighbour orbital than in the *p*- and *m*-form because the C–X bond orbitals to the first substituent consist in fact of two bond orbitals, one of $\sigma$- one of $\pi$-type. This is in contrast to the cyanocompounds where an analysis of the LMO's shows that there is only one bond orbital. The $\pi$-orbitals localized on the first substituent are in general not recalculated because they are considered to be distant enough from the region of interaction not to be strongly influenced by the introduction of the new substituent. This amounts to assuming that the electronic rearrangement in the $\pi$-electron part of the wave function caused by the second substituent is sufficiently taken into account by modifying only the three $\pi$-orbitals in the benzene ring. Test calculations further on showed that the recalculation of the $\pi$-orbital on the first substituent is energetically in general not necessary. Approximations MIM II and MIM III are identical in the transfer of LMO's, but MIM III involves a truncation of the basis set. The following notation will be used to identify the different calculations: $N_1$, $N_2$, $N_3 \Gamma M$. $N_1$ is the number of LMO's transferred for the first fragment ($C_6H_5X$), $N_2$ is the number of LMO's transferred for the second one (HY), $N_3$ is the number of MO's calculated in the region of interaction, and $M$ is the number of basis functions included in the set $\Gamma$. The detailed listing of all calculations performed is given in Table 2. The basis functions taken out of the entire set of basis

Table 2. List of MIM approximations considered for the molecules $C_6H_4XY$

| Molecule | MIM I | MIM II | MIM III |
|---|---|---|---|
| $C_6H_4(CN)_2$ | 17, 2, 14 $\Gamma$ 54 | 21, 2, 10 $\Gamma$ 54 | 21, 2, 10 $\Gamma$ 40(47)[a] |
| $C_6H_4CNOH$ | 17, 1, 13 $\Gamma$ 50 | 21, 1, 9 $\Gamma$ 50 | 21, 1, 9 $\Gamma$ 36(43)[a] |
| $C_6H_4CNF$ | 17, 1, 13 $\Gamma$ 49 | 21, 1, 9 $\Gamma$ 49 | 21, 1, 9 $\Gamma$ 35(42)[a] |
| $C_6H_4(OH)_2$ | 15, 1, 13 $\Gamma$ 46<br>(14, 1, 14 $\Gamma$ 46)[a] | 19, 1, 9 $\Gamma$ 46 | 19, 1, 9 $\Gamma$ 35(39)[a] |
| $C_6H_4OHF$ | 15, 1, 13 $\Gamma$ 45<br>(14, 1, 14 $\Gamma$ 45)[a] | 19, 1, 9 $\Gamma$ 45 | 19, 1, 9 $\Gamma$ 34(38)[a] |
| $C_6H_4F_2$ | 15, 1, 13 $\Gamma$ 44<br>(14, 1, 14 $\Gamma$ 44)[a] | 19, 1, 9 $\Gamma$ 44 | 19, 1, 9 $\Gamma$ 33(37)[a] |

[a] The notation given in brackets refers to the calculations on the *o*-compounds if they differ from the other calculations.

functions in approximation MIM III are for the *p*-compound the functions on the first substituent and on the atoms $C_1$, $H_6$, $H_2$, for the *m*-compound the functions on the first substituent and on the atoms $C_6$, $H_6$, $H_5$, and for the *o*-compound the functions on the atoms $C_5$, $H_4$, $H_5$, and $H_6$. Only $\sigma$-type basis functions have been taken out. As a consequence of this truncation some LMO's have to be taken out of the projection operator for orthogonality, which are mainly constructed from the deleted basis functions. In addition to these approximations the effect of neglecting the nonorthogonality of the MO's and the effect of the point charge approximation for the calculation of the interaction energy between the two sets of LMO's has been investigated. The letter *a* denotes the exact calculation of the energy with the non-orthogonality of the MO's taken into account, *b* the exact calculation of the energy neglecting the non-orthogonality of the MO's, *c* and *d* correspond to *a* and *b*, but involve the point charge approximation in the calculation of the interaction energy. Thus for every calculation $N_1$, $N_2$, $N_3 \Gamma M$ four energy values have been computed: $E^{\mathrm{MIM}}(a)$, $E^{\mathrm{MIM}}(b)$, $E^{\mathrm{MIM}}(c)$, $E^{\mathrm{MIM}}(d)$. The MIM calculations made for the *p*- and *m*-compounds can always be done in a closely corresponding form, i.e. the number of transferred LMO's is the same for both and the number of functions in the set $\Gamma$ is the same, although there may be some different LMO's transferred and different members in the set $\Gamma$. This is not possible for the *o*-compound because the two substituents are close together. For these molecules the number of transferred LMO's and the number of basis functions in the set $\Gamma$ can differ from the calculations performed for the *p*- and *m*-disubstituted benzenes. It is difficult to find corresponding calculations which can be compared with each other. The calculations for the *o*-compounds tend to be either more or less exact than the ones made for the other molecules. Thus the difference $\Delta_1$ can be regarded as a good criterion for the quality of the MIM approximation, but $\Delta_2$ is subject to the uncertainties mentioned above. The greater spread in the values of $\Delta_2$ will exhibit this.

## $C_6H_4(CN)_2$, $C_6H_4CNOH$, $C_6H_4CNF$

The results for these molecules are contained in the Tables 3–5. The SCF energy differences $\Delta_1$ and $\Delta_2$ are extremely small quantities ranging from 0.16–1.95 kcal/mole and cannot be interpreted at all in particular because they are based on wave functions calculated with an insufficient basis set. Although the total energy values computed in the MIM approximation are quite close to the SCF energies (approximate error $10^{-3}$ to $10^{-1}$ a.u.) the energy differences are reproduced in a very poor way. Only the approximations MIM I (*a*), MIM II (*a*), and sometimes MIM III (*a*) give good values for $\Delta_1$ and $\Delta_2$. Even if the energy difference derived from total energies of 400 a.u. is as small as 0.1 kcal/mole the calculation involving only the transfer of LMO's reproduces this value closely

Table 3. Total energies for *p*-, *m*-, and *o*-$C_6H_4(CN)_2$ calculated by the SCF and MIM methods. $\Delta_1 = E\,(\text{para}) - E\,(\text{meta})$, $\Delta_2 = E\,(\text{ortho}) - E\,(\text{para})$. (For notation see text; all values in atomic units)

| Method | $E(p\text{-}C_6H_4(CN)_2)$ | $\Delta_1$ | $\Delta_2$ |
|---|---|---|---|
| SCF | −392.334033 | 0.000248 | 0.001042 |
| MIM I *a* | −392.333020 | 0.000432 | 0.003092 |
| MIM I *b* | −392.381689 | 0.000128 | −0.003534 |
| MIM I *c* | −392.351570 | −0.002887 | 0.028848 |
| MIM I *d* | −392.403672 | −0.003369 | 0.017369 |
| MIM II *a* | −392.329503 | 0.000366 | 0.003331 |
| MIM II *b* | −392.387683 | 0.000033 | −0.003653 |
| MIM II *c* | −392.357484 | −0.002808 | 0.028835 |
| MIM II *d* | −392.428616 | −0.003654 | 0.017706 |
| MIM III *a* | −392.315294 | 0.000005 | 0.003785 |
| MIM III *b* | −392.376497 | −0.000613 | −0.001918 |
| MIM III *c* | −392.343415 | −0.003009 | 0.028772 |
| MIM III *d* | −392.417429 | −0.004299 | 0.019441 |

Table 4. Total energies for *p*-, *m*-, and *o*-$C_6H_4CNOH$ calculated by the SCF and MIM methods. $\Delta_1 = E\,(\text{para}) - E\,(\text{meta})$, $\Delta_2 = E\,(\text{ortho}) - E\,(\text{para})$. (For notation see text; all values in atomic units)

| Method | $E(p\text{-}C_6H_4CNOH)$ | $\Delta_1$ | $\Delta_2$ |
|---|---|---|---|
| SCF | 375.817025 | 0.000720 | 0.003106 |
| MIM I *a* | −375.815932 | 0.000821 | 0.004037 |
| MIM I *b* | −375.868250 | 0.000515 | −0.002630 |
| MIM I *c* | −375.829018 | −0.000964 | 0.012900 |
| MIM I *d* | −375.884307 | −0.001339 | 0.003157 |
| MIM II *a* | −375.812052 | 0.000790 | 0.004072 |
| MIM II *b* | −375.873649 | 0.000445 | −0.002999 |
| MIM II *c* | −375.831898 | −0.000888 | 0.012438 |
| MIM II *d* | −375.902709 | −0.001509 | 0.002713 |
| MIM III *a* | −375.797528 | 0.000495 | 0.004526 |
| MIM III *b* | −375.861448 | −0.000333 | −0.001520 |
| MIM III *c* | −375.817669 | −0.001126 | 0.012550 |
| MIM III *d* | −375.890508 | −0.001509 | 0.004192 |

Table 5. Total energies for *p*-, *m*-, and *o*-$C_6H_4CNF$ calculated by the SCF and MIM methods. $\Delta_1 = E\,(\text{para}) - E\,(\text{meta})$, $\Delta_2 = E\,(\text{ortho}) - E\,(\text{para})$. (For notation see text; all values in atomic units)

| Method | $E(p\text{-}C_6H_4CNF)$ | $\Delta_1$ | $\Delta_2$ |
|---|---|---|---|
| SCF | −397.892863 | 0.000301 | 0.001836 |
| MIM I *a* | −397.891642 | 0.000428 | 0.002604 |
| MIM I *b* | −397.956477 | 0.000114 | −0.004356 |
| MIM I *c* | −397.905347 | −0.001371 | 0.010334 |
| MIM I *d* | −397.976966 | −0.001674 | −0.000183 |
| MIM II *a* | −397.887780 | 0.000375 | 0.002560 |
| MIM II *b* | −397.961907 | 0.000023 | −0.004937 |
| MIM II *c* | −397.908616 | −0.001306 | 0.009719 |
| MIM II *d* | −397.997090 | −0.001862 | −0.000927 |
| MIM III *a* | −397.870075 | 0.000344 | 0.003224 |
| MIM III *b* | −397.946009 | 0.000173 | −0.003279 |
| MIM III *c* | −397.891486 | −0.001297 | 0.010064 |
| MIM III *d* | −397.981191 | −0.001711 | 0.000731 |

if the energy value is calculated correctly. It does not matter very much whether the nearest and next nearest neighbour orbitals or only the nearest neighbour orbitals are modified, the approximation is a very good one. LMO's are thus seen to be transferable to a very great accuracy. The other approximations, however, are not capable of reproducing such small energy differences since the corresponding results have to be considered as useless. It will become apparent that it is the smallness of the energy differences of 0.1 kcal/mole which causes this.

### $C_6H_4(OH)_2, C_6H_4OHF, C_6H_4F_2$

Whereas in the case of the molecules containing a CN substituent the SCF energy differences are of the order of 0.1–1.0 kcal/mole they increase in the present cases by one to two orders of magnitude. For the molecule $C_6H_4(OH)_2$ $\Delta_1$ and $\Delta_2$ take the values: $\Delta_1 = 10.7$ kcal/mole and $\Delta_2 = 7.7$ kcal/mole (Table 6). The MIM approximation is in this case capable of reproducing correctly these energy differences not only in case *a* of the calculations but in all other cases as well. The MIM values of $\Delta_1$ are in all cases in satisfactory agreement with the SCF results, the maximum deviation of about 23% is encountered for approximation MIM III *c*, which involves the point charge approximation. For reasons previously mentioned the energy differences $\Delta_2$ cannot be expected to show the same agreement as the $\Delta_1$ values. The $\Delta_2$ values calculated for MIM approximation I, however, agree well with the SCF value. This is also true for the cases *c* and *d*. For MIM II and MIM III results *a* and *b* can be regarded as satisfactory or acceptable, but *c* and *d* are too large by a factor of about two.

The situation encountered for the molecule $C_6H_4OHF$ is very similar as in the preceding case. The SCF energy differences $\Delta_1$ and $\Delta_2$ have about the same value ($\Delta_1 = 12.7$ kcal/mole and $\Delta_2 = 10.8$ kcal/mole) and are satisfactorily reproduced by

Table 6. Total energies for *p*-, *m*-, and *o*-$C_6H_4(OH)_2$ calculated by the SCF and MIM methods. $\Delta_1 = E\,(\text{para}) - E\,(\text{meta})$, $\Delta_2 = E\,(\text{ortho}) - E\,(\text{para})$. (For notation see text; all values in atomic units)

| Method | $E(p\text{-}C_6H_4(OH)_2)$ | $\Delta_1$ | $\Delta_2$ |
|---|---|---|---|
| SCF | −359.285606 | 0.017027 | 0.012191 |
| MIM I *a* | −359.284464 | 0.017285 | 0.012409 |
| MIM I *b* | −359.336354 | 0.017565 | 0.010297 |
| MIM I *c* | −359.295186 | 0.014883 | 0.011642 |
| MIM I *d* | −359.349804 | 0.015273 | 0.008956 |
| MIM II *a* | −359.280817 | 0.017056 | 0.012796 |
| MIM II *b* | −359.341919 | 0.016913 | 0.005977 |
| MIM II *c* | −359.297711 | 0.015387 | 0.031684 |
| MIM II *d* | −359.367593 | 0.015255 | 0.022911 |
| MIM III *a* | −359.266532 | 0.015121 | 0.013883 |
| MIM III *b* | −359.329195 | 0.015718 | 0.008074 |
| MIM III *c* | −359.284029 | 0.013148 | 0.032604 |
| MIM III *d* | −359.354869 | 0.014060 | 0.025008 |

Table 7. Total energies for *p*-, *m*-, and *o*-$C_6H_4OHF$ calculated by the SCF and MIM methods. $\Delta_1 = E\,(\text{para}) - E\,(\text{meta})$, $\Delta_2 = E\,(\text{ortho}) - E\,(\text{para})$. (For notation see text; all values in atomic units)

| Method | $E(p\text{-}C_6H_4OHF)$ | $\Delta_1$ | $\Delta_2$ |
|---|---|---|---|
| SCF | −381.358415 | 0.020173 | 0.017175 |
| MIM I *a* | −381.357114 | 0.020489 | 0.017205 |
| MIM I *b* | −381.421482 | 0.020794 | 0.014982 |
| MIM I *c* | −381.368396 | 0.018021 | 0.016446 |
| MIM I *d* | −381.439159 | 0.018507 | 0.013454 |
| MIM II *a* | −381.353486 | 0.020255 | 0.017505 |
| MIM II *b* | −381.427071 | 0.020136 | 0.010367 |
| MIM II *c* | −381.371285 | 0.018543 | 0.026487 |
| MIM II *d* | −381.458629 | 0.018499 | 0.026891 |
| MIM III *a* | −381.335837 | 0.018426 | 0.018783 |
| MIM III *b* | −381.410495 | 0.019786 | 0.012564 |
| MIM III *c* | −381.354544 | 0.016369 | 0.037664 |
| MIM III *d* | −381.442053 | 0.018149 | 0.029087 |

the different MIM approximations (Table 7). The greatest deviation of $\Delta_1$ (19%) occurs for the approximation MIM III *c* as before. The MIM energy differences $\Delta_2$ agree well with the SCF result for MIM I (*a*, *b*, *c*, and *d*). In the two other approximations it is again the cases *c* and *d* which lead to somewhat too large energy differences.

The largest energy differences between the *p*-, *m*-, and *o*-compounds are encountered for the difluorobenzene molecules: $\Delta_1 = 15$ kcal/mole and $\Delta_2 = 22.2$ kcal/mole (Table 8). All MIM approximations reproduce the value of $\Delta_1$ in a very satisfactory way. The deviations from the SCF value are very small for MIM I and MIM II. When the basis set is truncated in approximation III the

Table 8. Total energies for *p*-, *m*-, and *o*-$C_6H_4F_2$ calculated by the SCF and MIM methods. $\Delta_1 = E\,(\text{para}) - E\,(\text{meta})$, $\Delta_2 = E\,(\text{ortho}) - E\,(\text{para})$. (For notation see text; all values in atomic units)

| Method | $E(p\text{-}C_6H_4F_2)$ | $\Delta_1$ | $\Delta_2$ |
|---|---|---|---|
| SCF | −403.430215 | 0.023878 | 0.035285 |
| MIM I *a* | −403.429029 | 0.024076 | 0.035178 |
| MIM I *b* | −403.493382 | 0.024752 | 0.021855 |
| MIM I *c* | −403.441602 | 0.023315 | 0.030581 |
| MIM I *d* | −403.512365 | 0.024432 | 0.010307 |
| MIM II *a* | −403.425474 | 0.023853 | 0.035354 |
| MIM II *b* | −403.499097 | 0.024328 | 0.015633 |
| MIM II *c* | −403.444626 | 0.024217 | 0.052882 |
| MIM II *d* | −403.532116 | 0.025543 | 0.024962 |
| MIM III *a* | −403.408017 | 0.021007 | 0.036707 |
| MIM III *b* | −403.482627 | 0.025969 | 0.018005 |
| MIM III *c* | −403.428101 | 0.021078 | 0.054230 |
| MIM III *d* | −403.515646 | 0.027184 | 0.024054 |

difference to the SCF value is 14% maximally. This can still be regarded as a good result. The MIM approximation reproduces the greater stability of the *p*- compared to the *o*-compound as well, but the quality of the agreement of the $\Delta_2$ values computed by the MIM and SCF methods is poorer than for the $\Delta_1$ values except for case *a*. The $\Delta_2$ values are in some cases somewhat too small in other cases somewhat too large. This is not unexpected.

## 3. Conclusions

The SCF calculations on the *p*-, *m*-, and *o*-forms of the molecules $C_6H_4(CN)_2$, $C_6H_4CNOH$, $C_6H_4CNF$, $C_6H_4(OH)_2$, $C_6H_4OHF$, and $C_6H_4F_2$ suggest that the stability as measured by the total energy always increases in the sequence *m*-, *p*-, *o*-compound. For the latter three molecules the energy differences were approximately 10–20 kcal/mole, for the former three only 0.1–1 kcal/mole. The wave functions have been calculated with a minimal atomic basis set of Gaussian lobe functions, which can only give a poor description of the electronic density. But since all three forms (*p*-, *m*-, and *o*-form) are described in an equal way it is expected that more exact calculations will corroborate the findings of this work concerning the stability at least for the compounds containing only the substituents OH and F. Nothing can be said about the cyano-compounds. Definite conclusions can certainly only be drawn from calculations with extended basis sets. From the calculations it can be concluded that the meta directing property of the CN substituent and the *o*- and *p*-directing property of the OH and F substituent cannot be explained by the thermodynamic stability of the final compounds, which are formed, but must be a kinetic effect as already indicated by the mesomeric structures drawn to explain the observed behaviour of the molecules.

The theory of molecules in molecules has been applied with success to the three molecules $C_6H_4(OH)_2$, $C_6H_4OHF$, and $C_6H_4F_2$, but it was found to be incapable of reproducing the energy differences in the CN substituted benzenes correctly. These latter energy differences are too small and the MIM results turn out to be unreliable. In other cases (Ne · $H_2O$) [23], however, the theory of molecules in molecules has proved to be capable of reproducing such small energy differences of 0.1 kcal/mole. In these cases two closed shell molecules interacted with each other at relatively large distances. In the present case the distances are shorter and the interactions in the benzene derivatives are stronger. This may be the main reason why the MIM approximation fails in the case of the CN substituted benzenes. In applications like the present one the energy differences must be about 10 kcal/mole for the MIM method to yield reliable results. It must be mentioned, however, that if the transfer of LMO's is the only approximation made, very good results are obtained in all cases. This demonstrates that LMO's are transferable to a great accuracy and describe each separate regions of a molecule.

An examination of the Tables 6–8 in particular leads to the following conclusions. The truncation of the basis set for the expansion of the MO's in the region of interaction and of the LMO's in the projection operator for orthogonality finds its justification in this investigation as well. Of all approximations considered (MIM I, II, and III) the greatest variations in the values of $\Delta_1$ and $\Delta_2$ are found for MIM III which involves the truncation of the basis set, but the values are acceptable and reproduce correctly the energy differences in the Tables 6–8. The results of this approximation must obviously be poorer than for the other cases, but it gives the possibility to save computational time. If the energy differences are of the order of 10 kcal/mole than the data calculated with the MIM approximation including the truncation of the basis set differ by about 20% from the SCF values if questions similar to the present one are examined. For the hydrogen bonding interactions considered in previous articles better results were obtained [3, 4, 23]. The approximation will henceforth be accepted as a standard approximation in this theory of molecules in molecules. The neglect of the nonorthogonality of the MO's introduces a relatively large error in the value of the total energy (approximately 0.1 a.u. for the molecules considered here), but the results *b* are not really worse than the results *a*, if the energy differences are not too small. The data in Table 6–8 can be accepted as a justification of this approximation. But if one wishes to make this theory of molecules in molecules applicable also to reproduce very small energy differences – which might be possible because the basic approximation, the transfer of LMO's, works so well – an improvement in the method dealing with the nonorthogonality problem should be made. Among the approximations of the MIM theory it is first of all the point charge approximation which has to be improved because it introduces the greatest unreliability of all approximations. This has been found in previous applications as well [3, 4, 23]. But this approximation does not fail as it did in some cases. A subsequent article will thus be devoted to an improvement of the method dealing with the nonorthogonality problem and in particular to an improvement of the point charge approximation [24].

It has been demonstrated in this series of articles that wave functions calculated by transferring LMO's from the wave functions of fragment molecules and by redetermining the MO's in the region of interaction are of a quality comparable to the one of SCF wave functions. This theory of molecules in molecules appears to be at least one possible path towards reliable and time saving calculations on larger molecules (subject to the limitation of the method in its present form in the case of very small energy differences), but it is expected that it can be improved in the two points mentioned above.

*Acknowledgement.* These investigations were started when the author was at the Max-Planck-Institut für Physik und Astrophysik in München. The calculations were performed on an IBM 360/91 computer. The author would like to thank Professor Dr. L. Biermann and Dr. G. Diercksen for the opportunity of working in the institute and using its facilities, and Professor Dr. G. L. Hofacker for giving the possibility to finish this work. Financial support of the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

## References

1. von Niessen, W.: J. Chem. Phys. **55**, 1948 (1971)
2. von Niessen, W.: Theoret. Chim. Acta (Berl.) **31**, 111 (1973)
3. von Niessen, W.: Theoret. Chim. Acta (Berl.) **31**, 297 (1973)
4. von Niessen, W.: Theoret. Chim. Acta (Berl.) **32**, 13 (1973)
5. Edmiston, C., Ruedenberg, K.: Rev. Mod. Phys. **35**, 457 (1963); J. Chem. Phys. **43**, S 97 (1965)
6. England, W., Salmon, L.S., Ruedenberg, K.: Topics in current chemistry, Vol. 23 p. 31. Berlin: Springer Verlag, 1971
7. Weinstein, H., Pauncz, R., Cohen, M.: Advan. At. Mol. Phys. Vol. 7, p. 97. New York: Academic Press, 1971
8. von Niessen, W.: J. Chem. Phys. **56**, 4290 (1972); Theoret. Chim. Acta (Berl.) **27**, 9 (1972), **29**, 29 (1973)
9. Rothenberg, S.: J. Chem. Phys. **51**, 3389 (1969); J. Am. Chem. Soc. **93**, 68 (1970)
10. Switkes, E., Lipscomb, W.N., Newton, M.D.: J. Am. Chem. Soc. **92**, 3837, 3847 (1970)
11. Newton, M.D., Switkes, E.: J. Chem. Phys. **54**, 3179 (1971)
12. Levy, M., Stevens, W.J., Shull, H., Hagstrom, S.: J. Chem. Phys. **52**, 5483 (1970)
13. Adams, W.H.: J. Am. Chem. Soc. **92**, 2198 (1969); J. Chem. Phys. **57**, 2340 (1972)
14. Peters, D.: J. Am. Chem. Soc. **94**, 707 (1972)
15. England, W., Gordon, M.S.: J. Am. Chem. Soc. **93**, 4649 (1971); **94**, 4818, 5168 (1972); Chem. Phys. Letters **15**, 59 (1972)
16. Epstein, I.R.: J. Chem. Phys. **53**, 4425 (1970); Roux, M., Epstein, I.R.: Chem. Phys. Letters **18**, 18 (1973)
17. Eisenberger, P., Marra, W.C.: Phys. Rev. Letters **27**, 1413 (1971)
18. Löwdin, P.O.: Phys. Rev. **139**, A 357 (1965)
19. Tables of interactomic distances and configurations in molecules and ions. The Chemical Society (London), Spec. Publ. No. 11 (1958) and No. 18 (1965)
20. Csizmadia, I.G., Harrison, M.C., Moskowitz, J.W., Sutcliffe, B.T.: Theoret. Chim. Acta (Berl.) **6**, 191 (1966)
21. Preuss, H.: Z. Naturforsch. **20**a, 1290 (1965)
22. Boys, S.F.: Rev. Mod. Phys. **32**, 296 (1960); Foster, J.M., Boys, S.F.: Rev. Mod. Phys. **32**, 300 (1960), – Boys, S.F.: Quantum theory of atoms, molecules, and the solid state, P. O. Löwdin. Ed., p. 253. New York: Academic Press 1966
23. von Niessen, W.: Unpublished results
24. von Niessen, W.: In preparation

Dr. W. von Niessen
Lehrstuhl für Theoretische Chemie
Technische Universität München
D-8000 München, Arcisstraße 21
Federal Republic of Germany