# DFT Calculations of Core−Electron Binding Energies of the Peptide Bond

**Delano P. Chong,[†] Philippe Aplincourt,*[,‡] and Christophe Bureau[‡]**

*Department of Chemistry, 2036 Main Mall, University of British Columbia, Vancouver, BC, Canada V6T 1Z1, and CEA-Saclay, DSM-DRECAM-SRSIM, bât. 466, F-91191 Gif-sur-Yvette Cedex, France*

*Received: August 1, 2001; In Final Form: October 23, 2001*

Although an efficient DFT method using the generalized transition-state model to calculate core−electron binding energies had been successfully applied to over 200 cases, with an average absolute deviation of only 0.21 eV from experiment, a new $\Delta E_{KS}$(PW86−PW91)/cc-pCVTZ model based on total Kohn−Sham energy difference was recently developed. Not only was the model error-free, but also the average absolute deviation for 32 cases studied was found to be 0.15 eV. In this study, we first confirm the excellent performance of such a $\Delta E_{KS}$ approach with 46 new cases, with the result that the average absolute deviation from experiment for the 78 cases remains at 0.15 eV. With such consistent accuracy, this new method is applied to the peptide bond. The model molecules studied in this work include formamide, *N*-methylformamide, *N*,*N*-dimethylformamide, acetamide, *N*-methylacetamide, *N*,*N*-dimethylacetamide, and two model dipeptides, one cyclic and one acyclic. The difference in the computed nitrogen core−electron binding energy between the two model dipeptides is found to be 0.85 eV, several times our average absolute deviation. This may be of interest to other workers studying other aspects of the peptide bond.

## 1. Introduction

In the past few years, an efficient method of computing accurate core−electron binding energies (CEBEs) based on density-functional theory (DFT) has been developed and thoroughly tested.[1−7] The method used the unrestricted generalized transition-state (uGTS) model of Williams et al.,[8] together with Becke's 1988 exchange functional[9] and Perdew's 1986 correlation functional.[10] Small relativistic corrections, based on Pekeris' study of two-electron ions,[11] were added to the nonrelativistic values. Also, an efficient scaled basis set was proposed and developed.[1,2] In short, the method may be labeled: uGTS(B88-P86)/scaled-pVTZ. Over 200 cases, including some unpublished results, were compared with experiment, and the average absolute deviation (AAD) of the calculated CEBEs was only 0.21 eV. Although the accuracy of the uGTS(B88−P86) method is due to fortuitous cancellation of errors, the fact that this cancellation persists throughout the large set of 200 molecules previously studied lends some credibility to this method.

Very recently, Triguero et al.[12,13] and Chong et al.[14] showed that such an accuracy was the result of fortuitous cancellation of the two main sources of error, namely, from the uGTS model (positive error) and from the B88−P86 functional (negative error). First, the model error can be eliminated by going after the total Kohn−Sham energy difference. $\Delta E_{KS}$ calculations by using various local and nonlocal potentials have been first performed by Pedocchi et al.[15] with an AAD close to 0.5 eV. In a recent study, Triguero et al. [13] have tested the reliability of the $\Delta E_{KS}$ method for two functionals BP86 and PW86P86. They observed strong functional dependency and obtained quite large AADs (−0.43 and 0.45 eV respectively). Thus, the $\Delta E_{KS}$ approach was not reliable until the investigation of Chong et al.[14] By testing 10 different functional combinations, they reduced the functional error by a large extent. The Perdew−Wang 1986 exchange functional[16] and the Perdew-Wang 1991 correlation functional[17] were found to be the best combination, giving CEBEs closest to experiment. Finally, various basis sets were tested, and the best one was found to be the correlation-consistent polarized core-valence triple-$\zeta$ basis set.[18] In short, this new and more reliable method may be called: $\Delta E_{KS}$(PW86−PW91)/cc-pCVTZ. The AAD for the 32 cases studied[14] was 0.15 eV.

On the other hand, two recent articles reported the calculated electronic spectra of dipeptides, which are the basic elements describing larger polypeptides since they have many properties (flexible backbone, dihedral angles) encountered in real peptides. Serrano-Andrés and Fülscher[19] examined a "linear" model dipeptide (LMD), 2-(acetylamino)-*N*-methylacetamide and two other model polypeptides by CASPT2 method at MP2/6-31G* optimized geometry, while Hirst and Persson[20] used CASPT2//MP2/cc-pVDZ to study the cyclic diketopiperazine (DKP). The calculated electronic transitions were found to be similar. What is interesting about DKP is that it can be regarded as two linked acetamide units or as two linked *N*-methylformamide molecules. On the basis of the results, Hirst and Persson[20] concluded that DKP is better considered as two linked *N*-methylformamides. In this work, we wish to apply our accurate method to investigate the CEBEs of simple amides as well as these two model dipeptides.

Before we proceed with that task, we first make calculations on the CEBEs of more molecules with our new reliable and accurate method so that more confidence may be placed in the results on LMD and DKP.

## 2. Method

The CEBEs have been computed with the deMon[21] software according to the method that we have developed recently. All

* Corresponding author. P.A.: e-mail, aplincourt@drecam.cea.fr; fax, 33(0)169086462.
[†] University of British Columbia. D.P.C.: e-mail, chong@chem.ubc.ca; fax, (604)822-2847.
[‡] CEA-Saclay. C.B.: e-mail, christophe.bureau@cea.fr.

10.1021/jp0129737 CCC: $22.00 
Published on Web 12/11/2001

**TABLE 1: Δ*E*-KS/P86−P91 Calculations of Core−Electron Binding Energies (in eV) with cc-pCVTZ Basis Set**

| molecule | calculated | experiment | deviation |
|---|---|---|---|
| **C**$H_2$=C($CH_3$)$_2$ | 289.86 | 289.83[a] | +0.03 |
| **C**$H_2$=CHCH=$CH_2$ | 290.27 | 290.23[a] | +0.04 |
| **C**$H_2$=C$HCH_3$ | 290.30 | 290.25[a] | +0.05 |
| $CH_2$=**C**($CH_3$)$_2$ | 290.73 | 290.65[a] | +0.08 |
| $CH_2$=C(**C**$H_3$)$_2$ | 290.51 | 290.69[a] | −0.18 |
| $CH_2$=**C**$HCH_3$ | 290.80 | 290.73[a] | +0.07 |
| $CH_2$=CH**C**$H_3$ | 290.92 | 290.81[a] | +0.11 |
| **C**$_2H_4$ | 290.93 | 290.82[a] | +0.11 |
| $CH_2$=**C**HCH=$CH_2$ | 290.90 | 290.87[a] | +0.03 |
| **C**$H_3$COO$CH_3$ | 291.51 | 291.30 | +0.21 |
| **C**$H_3$COOH | 291.59 | 291.55 | +0.04 |
| HCON(**C**$H_3$)$_2$b | 292.21 | 292.03 | +0.18 |
| **C**$H_3$O$CH_3$ | 292.20 | 292.34 | −0.14 |
| **C**$H_3$OH | 292.42 | 292.42 | 0.00 |
| **C**$H_3$Cl | 292.50 | 292.43[a] | +0.07 |
| $CH_3$COO**C**$H_3$ | 292.39 | 292.55 | −0.16 |
| **C**·$H_2$OC.$H_2$[c] | 292.55 | 292.91 | −0.36 |
| **C**$H_3$NC | 293.40 | 293.35 | +0.05 |
| H**C**ON($CH_3$)$_2$[b] | 293.32 | 293.45 | −0.13 |
| **C**$H_2Cl_2$ | 293.86 | 293.81[a] | +0.05 |
| H**C**O$NH_2$ | 294.11 | 294.45 | −0.34 |
| $H_2$**C**O | 294.55 | 294.47 | +0.08 |
| $CH_3$**C**OO$CH_3$ | 294.47 | 294.85 | −0.38 |
| $CH_3$**C**OOH | 295.13 | 295.38 | −0.25 |
| **C**$H_2F_2$ | 296.02 | 296.40 | −0.38 |
| **C**$HCl_3$ | 295.25 | 295.16[a] | +0.09 |
| **N**($CH_3$)$_3$ | 404.98 | 404.81 | +0.17 |
| ($CH_3$)$_2$**N**H | 405.07 | 404.92 | +0.15 |
| **N**$H_3$ | 405.71 | 405.56 | +0.15 |
| HCO**N**($CH_3$)$_2$[b] | 406.19 | 405.90 | +0.29 |
| HC**N** | 406.98 | 406.14 | +0.84[d] |
| HCO**N**$H_2$ | 406.70 | 406.39 | +0.31 |
| $CH_3$**N**C | 406.91 | 406.67 | +0.24 |
| HC**O**N($CH_3$)$_2$[b] | 536.98 | 536.95 | +0.03 |
| HC**O**$NH_2$ | 537.64 | 537.74 | −0.10 |
| $CH_3$C**O**O$CH_3$ | 537.92 | 537.92 | 0.00 |
| $CH_3$C**O**OH | 538.27 | 538.33 | −0.06 |
| $CH_3$**O**$CH_3$ | 538.71 | 538.74 | −0.03 |
| $CH_3CH_2$**O**H | 538.89 | 538.82 | +0.07 |
| $CH_3$CO**O**$CH_3$ | 539.41 | 539.46 | −0.05 |
| $H_2$C**O** | 539.61 | 539.48 | +0.13 |
| $CH_3$CO**O**H | 540.27 | 540.12 | +0.15 |
| $CH_3$**F** | 692.84 | 692.92 | −0.08 |
| $CH_2$**F$_2$** | 693.69 | 693.65 | +0.04 |
| CH**F**$_3$ | 694.47 | 694.62 | −0.15 |
| $C_2$**F$_6$** | 694.96 | 695.07 | −0.11 |
| average absolute deviation | | (0.00) | 0.15 |

[a] Determined by synchrotron studies. [b] Geometry optimized by AM1 semiempirical method. [c] Ethylene epoxide. [d] See text for possible reason of this large deviation.

the details of this method are given in ref 14. In shorthand notation, it is called $\Delta E_{KS}$(PW86−PW91)/cc-pCVTZ.

When there are two or more atoms of the same element in the molecule, some difficulties were encountered at the beginning. A typical example is 2-methylpropene. We have developed two different methods of localizing the core hole at any desired atom. Both approaches rely on first creating a desirable restart density. The first procedure, applied in most of our earlier studies,[1−7] replaces the atom of interest by an isoelectronic ion, such as C, by $B^-$ or $N^+$. The second method is based on the concept of blocking. If a core-hole is desired at atom x but ends up at atom y, then we artificially scale the first two s-type orbital basis functions of atom y by a factor of 0.2 (for example) so that the basis set becomes relatively unfavorable for a core-hole. Both methods work quite efficiently.

What remains to be specified is the geometry used. For the set of 46 molecules employed to complete our database, we have used experimental geometries given in convenient compilations[22−24] when available. If not, we optimized them at the MP2/6-31+G* level of theory. The simple amides as well as the dipeptides have been optimized at the MP2/6-31+G* level. The Gaussian98[25] package has been employed for geometry optimizations.

**TABLE 2: Summary of Deviations (in eV) of Calculated Core−Electron Binding Energies from Experiment**

| | Σ abs dev | no. of cases | ave abs dev |
|---|---|---|---|
| this work | 6.76 | 46 | 0.15 |
| previous study[a] | 4.57[b] | 32 | 0.14 |
| all cases to date | 11.33 | 78 | 0.15 |
| CEBEs for boron | 1.11 | 6 | 0.19 |
| CEBEs for carbon | 5.29 | 35 | 0.15 |
| CEBEs for nitrogen | 2.41 | 12 | 0.20 |
| CEBEs for oxygen | 1.45 | 15 | 0.10 |
| CEBEs for fluorine | 1.07 | 10 | 0.11 |
| ref = synchrotron only | 1.05 | 14 | 0.08 |
| other observed values as refs | 10.28 | 64 | 0.16 |

[a] See ref 12. [b] Synchrotron results now used as reference for **C**$H_4$ and **C**$Cl_4$.

### 3. Forty-Six New Cases

The new results of calculated CEBEs are presented in Table 1. Whenever CEBEs are available from synchrotron studies,[26−28] we assume that those values are most reliable. Otherwise, we rely on the compilation of Jolly et al.[29] For example, three values were listed for $NH_3$: 405.52, 405.6, and 405.60 eV. We discarded the "inaccurate" one 405.6 and took the average of the other two. When this procedure was followed for HCN, we discarded 406.8 eV and took the average of 406.13 and 406.15. With the computed values at 406.98 eV, we are inclined to believe that the discarded value of 406.8 eV (though imprecise) may well be more accurate.

The error statistics of Table 1 combined with the earlier results[14] are summarized in Table 2. The relatively high AAD of CEBEs for nitrogen is mainly caused by the large deviation of 0.84 eV for HCN. Without that entry, the AAD of CEBEs(N) would drop to 0.14 eV. On the other hand, when only synchrotron results are compared, our AAD for the 14 cases is reduced to 0.08 eV. Unless both theory and experiment suffer systematic error in the same direction, such a small AAD is excellent amazing and suggests that both our method of calculation and the synchrotron measurements (and subsequent analysis) are accurate. In any case, the overall reliability of our new procedure $\Delta E_{KS}$(PW86−PW91)/cc-pCVTZ is now firmly established, with the AAD of 0.15 eV for a total of 78 cases.

### 4. Results for Simple Amides

First of all, we have studied a series of simple amides that come from two parent amides, i.e., the formamide (formamide, *N*-methylformamide, *N*,*N*-dimethylformamide) and the acetamide (acetamide, *N*-methylacetamide, *N*,*N*-dimethylacetamide) molecules. In the case of the *N*-methylformamide, we have considered the two cis and trans isomers. Experimentally, the trans conformer is dominant. The presence of a small fraction of the cis conformer is suggested, but the experimental evidence is not conclusive. We have to precise that the three main atoms in amides or peptide chains are those of the framework, i.e., oxygen O, nitrogen N, and carbonyl carbon $C_{carb}$. Geometries of these simple amides have been optimized at the MP2/6-31+G* level and are reported in Figure 1. The agreement between theoretical parameters and experimental data available in the literature[30] is good. The AAD for bond lengths is 0.012

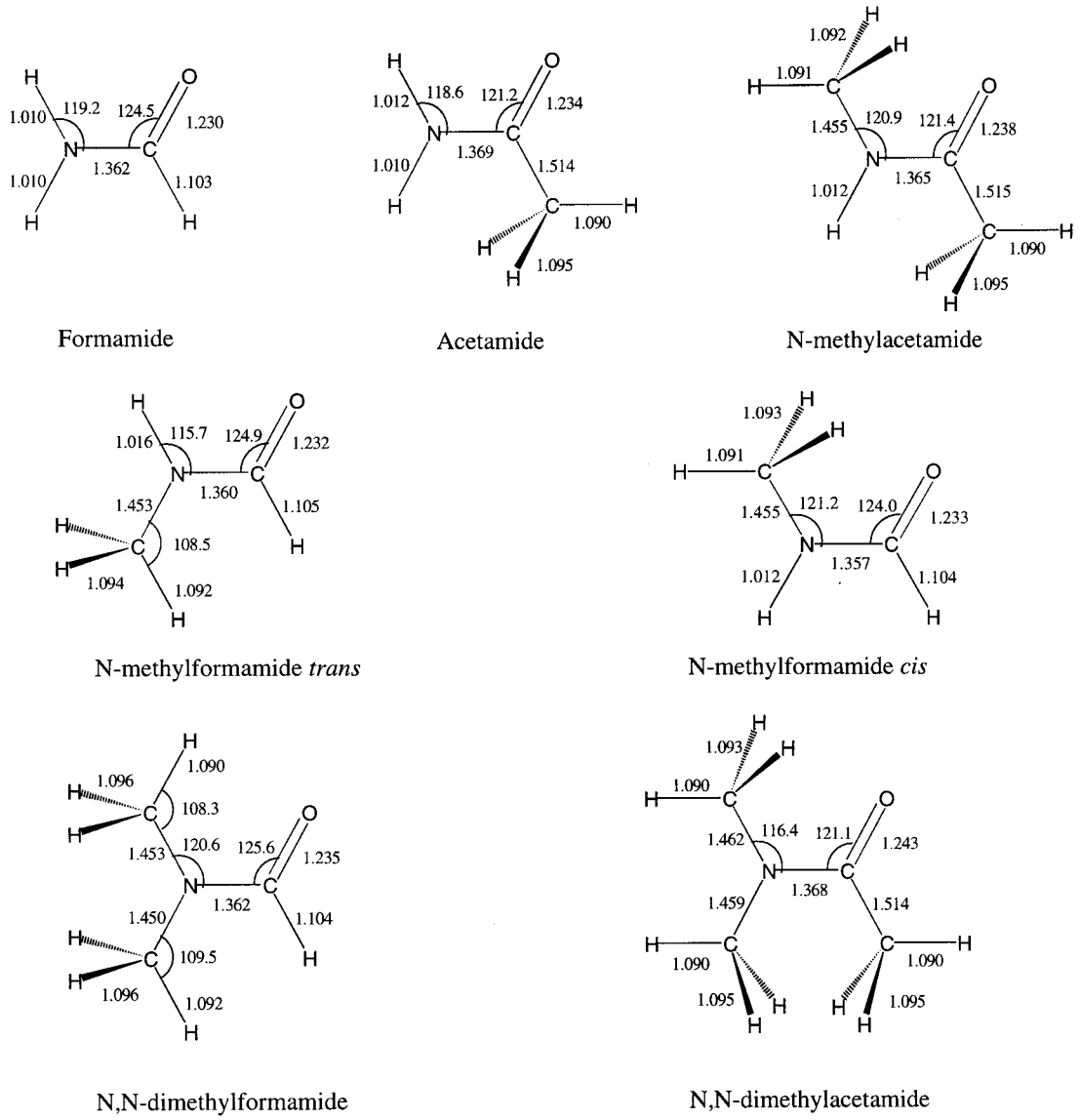
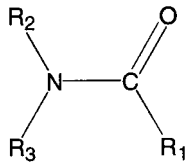



**Figure 1.** MP2/6-31+G* geometries of all the simple amides studied in this work. Distances are given in Å and angles in deg.

**TABLE 3: Calculated Core−Electron Binding Energies (in eV) of Simple Amides**

| molecule | formula | $R_1$[a] | $C_{carb}$[a] | O | N | $R_2$[a] | $R_3$[a] |
|---|---|---|---|---|---|---|---|
| formamide | $CH_3NO$ | | 294.16 | 537.65 | 406.70 | | |
| *trans*-*N*-methylformamide | $C_2H_5NO$ | | 293.64 | 537.23 | 406.37 | | 292.50 |
| *cis*-*N*-methylformamide | $C_2H_5NO$ | | 293.58 | 537.26 | 406.38 | 292.18 | |
| *N*,*N*-dimethylformamide | $C_3H_7NO$ | | 293.25 | 536.97 | 406.18 | 291.92 | 292.27 |
| Acetamide | $C_2H_5NO$ | 291.20 | 293.87 | 537.17 | 406.17 | | |
| *N*-methylacetamide | $C_3H_7NO$ | 291.12 | 293.37 | 536.86 | 405.97 | 292.03 | |
| *N*,*N*-dimethylacetamide | $C_4H_9NO$ | 290.90 | 292.99 | 536.54 | 405.79 | 291.69 | 292.11 |

[a] $C_{carb}$ is the carbon of the carbonyl group, and $R_1$ is on the carbonyl carbon; $R_2$ is cis to the C=O bond, and $R_3$ is trans to the C=O bond:



Å with a maximum of 0.021 Å. For the valence angles, the AAD is 0.6° with a maximum of 1.2°.

CEBEs for simple amides are gathered in Table 3 and are also plotted in Figure 2 as a function of the size of the molecules. There are many data entries to compare. No experimental XPS spectra are available on simple amides, and then it is necessary to check the reliability of our theoretical results. Only one previous study[31] has reported some core ionization energy calculations for the formamide and the *N*,*N*-dimethylacetamide compounds. The data are obtained with the Koopman's theorem and are corrected by a factor suited to each heavy atom. The results have been reported in Table 4 and compared with our

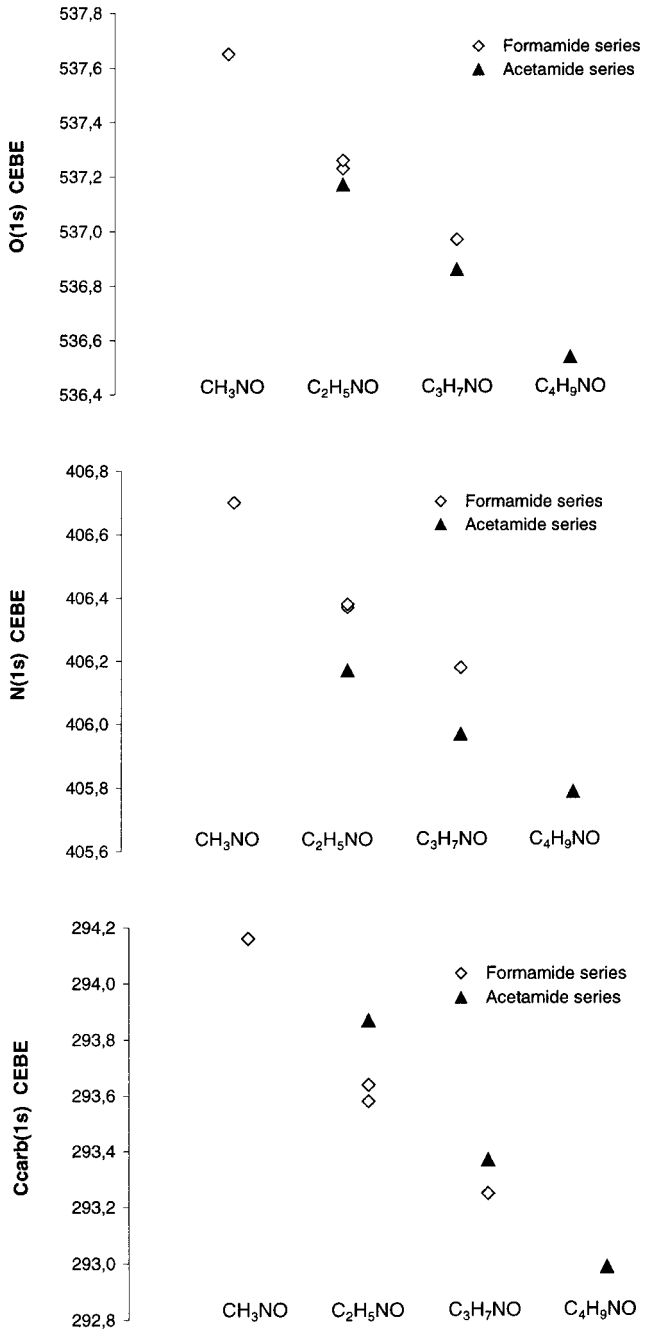



**Figure 2.** From top to bottom: O(1s), N(1s), and $C_{carb}$(1s) CEBEs (in eV) of all the simple amides studied in this work as function of the molecular size.

**TABLE 4: Comparison of Calculated Core−Electron Binding Energies (in eV) for the Formamide and the N−N-dimethylacetamide**

| molecule | | O | N | $C_{carb}$ |
|---|---|---|---|---|
| formamide | this work | 537.65 | 406.70 | 294.16 |
| | ref 28 | 537.71 | 406.33 | 294.56 |
| *N*,*N*-dimethylacetamide | this work | 536.54 | 405.79 | 292.99 |
| | ref 28 | 536.60 | 405.55 | 291.67 |

$\Delta E_{KS}$ values. We noticed large discrepancies between the two series of XPS results especially for the carbon atom in the *N*,*N*-dimethylacetamide (difference of 1.32 eV). However, as our method has been validated on a large series of compounds, we assume that our results are closer to reality.

Results in Table 3 and Figure 2 show that for molecules coming from the same parent compound (e.g., formamide and *N*-methylformamide), the CEBEs of the O, N and $C_{carb}$ atoms are downshifted with the successive addition of methyl groups because this electron-donating functional group contributes to the stabilization of the core-hole cation. Binding energies also decrease from the formamide-like compounds to the corresponding acetamide-like ones because of the presence of the electrodonating methyl group in $R_1$ position in the molecules of the acetamide series (e.g., *N*-methylformamide and *N*-methylacetamide).

**TABLE 5: Calculated CHelpG Charges for the $C_2H_5No$ and $C_3H_7NO$ Compounds and DKP**

| formula | molecule | O | N | $C_{carb}$ |
|---|---|---|---|---|
| $C_2H_5NO$ | *trans*-*N*-methylformamide | −0.56 | −0.50 | 0.48 |
| | *cis*-*N*-methylformamide | −0.53 | −0.51 | 0.54 |
| | Acetamide | −0.62 | −1.11 | 0.91 |
| $C_3H_7NO$ | *N*,*N*-dimethylformamide | −0.53 | −0.06 | 0.36 |
| | *N*-methylacetamide | −0.58 | −0.59 | 0.74 |
| | DKP [a] | −0.59 | −0.59 | 0.55 |

[a] Atomic charges for DKP have been computed on $C_2$ (or $C_6$), $N_4$ (or $N_8$), and $O_3$ (or $O_7$) atoms.

CEBEs for conformers having the same $C_xH_yNO$ formula have also been compared. Table 3 shown that compounds that belong to the same family (e.g., *cis*- and *trans*-*N*-methylformamide) give approximately the same O(1s), N(1s), and $C_{carb}$(1s) CEBEs. On the other hand, isomers of two different series, i.e., *N*-methylformamide and acetamide ($C_2H_5NO$) as well as *N*,*N*-dimethylformamide and *N*-methyacetamide ($C_3H_7NO$), yield to CEBEs that are clearly separated by few electrovolts. Thus, experimental X-ray photoelectron spectra coupled with theoretical calculations are a powerful tool to distinguish such molecules.

In addition, Figure 2 shows that members of the formamide series (white rhombus) yield higher O(1s) and the N(1s) CEBEs than the equivalent members of the acetamide series (black triangles). In the case of the $C_{carb}$ atom, the opposite trend is observed. This result can be interpreted in terms of charges brought by the three atoms of interest in neutral compound. We have then computed electrostatic potential derived charges using the CHelpG method[32] developed in Gaussian98. As the comparison can be made only on the $C_2H_5NO$ and $C_3H_7NO$ members, we have reported in Table 5 atomic charges of the compounds corresponding to one of these formula. For a given formula, Table 5 shows that the atomic charges brought by the oxygen atom are relatively similar from one conformer to another, but a little more negative for molecules of the acetamide series (average difference of −0.06 au). The computed O(1s) CEBE are relatively close between the two series, the values found for the acetamide being a little bit more downshifted because of the extra stabilization of the core-hole cation due to the more negative atomic charge (average difference of −0.08 eV). For the N-atom, the charges computed in the case of the acetamide-like compound are much more negative than the ones for the formamide-like compound (average difference of −0.58 au). As for the oxygen atom, the CEBEs computed for compounds of the acetamide series are therefore more downshifted than those from compounds of the formamide series, but the difference in computing CEBEs in this case is larger (average difference of −0.20 eV). For the $C_{carb}$ atom, the situation is reversed because the atomic charge is more positive for molecules of the acetamide series than for those of the formamide series (average difference of +0.39 au). The core-hole cation is therefore less stabilized in the case of acetamide compounds. Then these compounds present a less important downshift (average difference +0.21 eV).

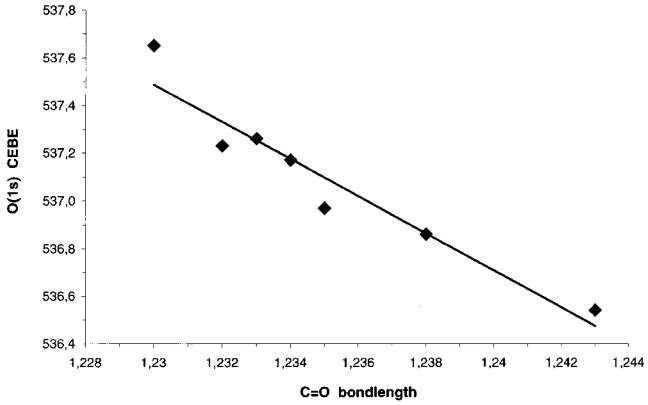



**Figure 3.** O(1s) CEBEs (in eV) of all the simple amides studied in this work as function of the C═O bond lengths (in Å).

Finally, we noticed that there is a correlation between the C═O bond length and the O(1s) binding energy shift. As shown in Figure 3, the O(1s) CEBE decreases with the increasing of the C═O bond length. This type of relation between geometrical parameters and calculated CEBEs has been also mentioned in our previous paper on hydrogen-bonded systems.[33] We have established a correlation between hydrogen bond lengths in the clusters studied and binding energy shifts.

## 5. Results for Polypeptides

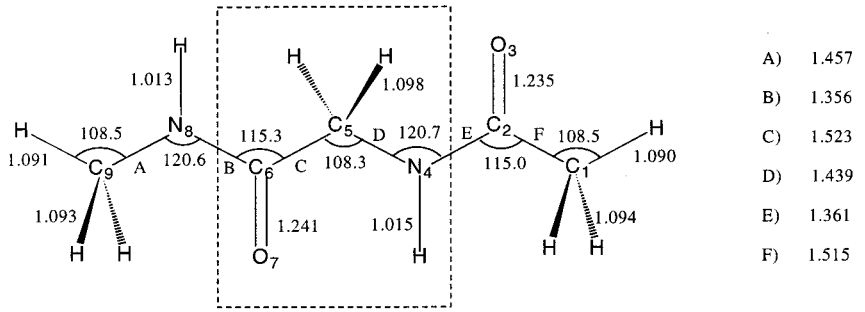
MP2/6-31+G* geometries of the two model dipeptides envisages are reported in Figure 4. Previous theoretical studies have investigated the structure of DKP by means of semi-empirical[34] and ab initio[20,35,36] methods. The boat form of $C_2$ symmetry is reported to be the only equilibrium structure. Recently, Bettens et al.[36] have explored the ring-puckering potential energy surface. They have shown that the minimum energy pathway linking the two boat enantiomeric conformers passes over a very small barrier of about 470 $cm^{-1}$ (1.34 kcal/mol). The chair conformer of $C_i$ symmetry is involved at the summit of the barrier. Another form, a planar one of $C_{2h}$ symmetry, has also been reported[20] as a saddle point higher on the potential energy surface. Distance parameters of LMD and DKP are very close. The AAD for bond lengths is 0.05 Å. The largest discrepancies are found for bonds around the central $C_5$ atom. Deviations for the $N_4-C_5$ and the $C_5-C_6$ bond lengths, respectively, reach 0.20 and 0.12 Å. Without these two values, the AAD is reduced to 0.03 Å.

**TABLE 6: Comparison of Calculated Core−Electron Binding Energies (in eV) of the Two Model Dipeptides LMD and DKP**

| atom[a] | LMD[b] | DKP[c] |
|---|---|---|
| $C_1$ | 290.86 | 292.57 |
| $C_2$ | 293.16 | 293.94 |
| $O_3$ | 536.67 | 537.25 |
| $N_4$ | 405.60 | 406.45 |
| $C_5$ | 291.98 | 292.57 |
| $C_6$ | 293.84 | 293.94 |
| $O_7$ | 537.35 | 537.25 |
| $N_8$ | 406.36 | 406.45 |
| $C_9$ | 292.25 | |

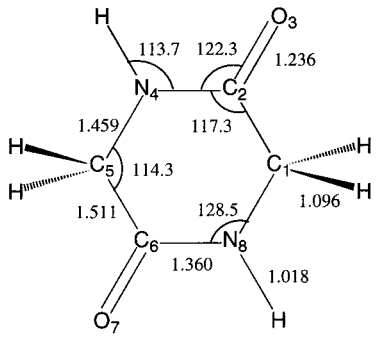
[a] See Figure 4 for numbering. [b] Linear model dipeptide: 2-(acetylamino)-*N*-methylacetamide. [c] Cyclic model dipeptide: diketopiperazine.

CEBEs for the two model dipeptides are presented in Table 6. On Figure 4, the part least affected by the methyl terminations




"linear" model dipeptide (LMD)

(2-(acetylamino)-N-methylacetamide)




Diketopiperazine (DKP)

**Figure 4.** MP2/6-31+G* geometries of the linear 2-(acetylamino)-*N*-methylacetamide (LMD) and diketopiperazine (DKP). The part in the former molecule least affected by the terminating methyl groups is indicated by the dashed rectangle. Distances are given in Å and angles in deg.

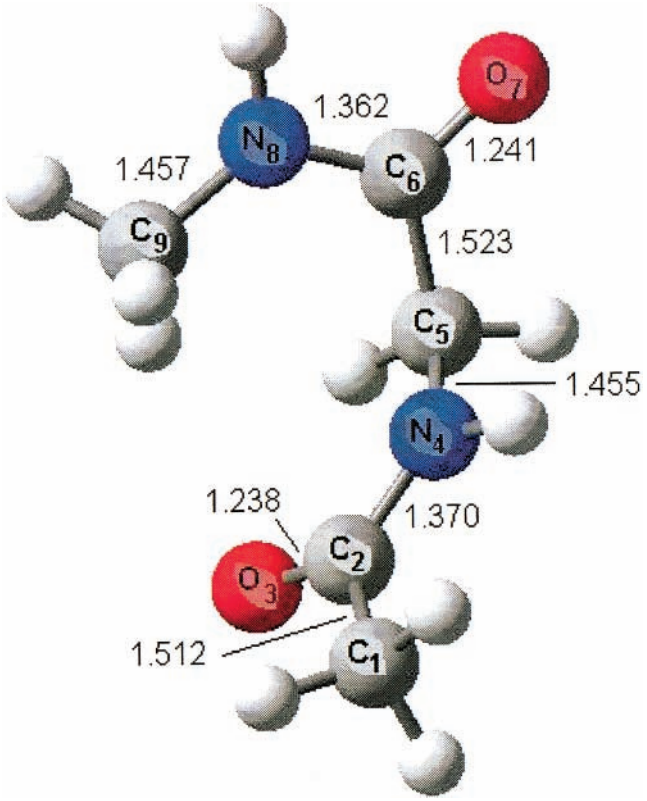



**Figure 5.** MP2/6-31+G* geometries of the nonlinear 2-(acetylamino)-*N*-methylacetamide (NCNLD). Distances are given in Å.

in LMD is indicated by the dashed rectangle. In the DKP, this problem does not exist because this compound can be seen as an infinite dipeptide. We find that the carbonyl $C_6$═$O_7$ has essentially the same CEBEs in LMD and DKP, indicating that they are in almost identical chemical environment. However, the CEBEs of $N_4$, and of $C_5$ to a smaller extent, are quite different, despite the similar environment surrounding them in the two model dipeptides. Besides the terminating methyl groups in LMD, which are two bonds away from the part marked off by the dashed rectangle in Figure 4, the only difference between the two model dipeptides is the conformation about the rigid $C_2-N_4$ and $C_6-N_8$ bonds: *Z*-conformation in LMD but *E*-conformation in DKP. This is a surprising finding, since the $\Delta E_{KS}$ predictions for *cis*- and *trans*-*N*-methylformamide are very similar for the nitrogen atom of interest.

Thus, we have examined if the discrepancies found in the case of $N_4$ and $C_5$ atoms between LMD and DKP were due principally to the presence of the terminating methyl groups in LMD or to the orientation of bonds around the two atoms of interest. We therefore looked for other noncyclic-nonlinear dipeptides having the same number of atoms as that in the LMD but a blackbone closer to that in the DKP. We optimized at the MP2/6-31+G* level only one dipeptide almost filling those terms. This compound, hereafter abbreviated NCNLD, is represented in Figure 5. Although the structure of NCNLD is still quite different from the structure of DKP, the atomic chain $C_5-C_6-O_7-N_8-C_9$ has a conformation similar to that encountered in the cyclic dipeptide, as shown in Table 7. Thus, this intermediate dipeptide presents two distinct atomic chains, i.e., one that resemble to atomic chain in LMD ($C_1-C_2-O_3-N_4-C_5$) and the other that resembles the atomic chain in DKP ($C_5-C_6-O_7-N_8-C_9$).

**TABLE 7: Comparison of MP2/6-31+G* Structures of LMD, DKP, and NCNLD**

| | DKP | LMD | NCNLD |
|---|---|---|---|
| $C_2-N_4$ | 1.360 | 1.361 | 1.370 |
| $C_2-O_3$ | 1.236 | 1.235 | 1.238 |
| $C_1-C_2$ | 1.511 | 1.515 | 1.512 |
| $N_4-H$ | 1.018 | 1.015 | 1.014 |
| $N_4-C_5$ | 1.459 | 1.439 | 1.455 |
| $(N_4C_2O_3)$ | 120.4 | 121.5 | 122.3 |
| $(N_4C_2C_1)$ | 122.3 | 115.0 | 115.7 |
| $(O_3C_2C_1)$ | 120.3 | 123.2 | 122.0 |
| $(C_2N_4H)$ | 113.7 | 122.8 | 118.7 |
| $(C_2N_4C_5)$ | 128.5 | 120.7 | 122.7 |
| $(HN_4C_5)$ | 117.8 | 116.4 | 117.4 |
| $[O_3C_2N_4H]$ | 0.0 | 180.0 | −174.5 |
| $[O_3C_2N_4C_5]$ | 180.0 | 0.0 | −7.0 |
| $[C_1C_2N_4H]$ | 180.0 | 0.0 | 5.7 |
| $[C_1C_2N_4C_5]$ | 0.0 | 180.0 | 173.1 |
| $[C_5C_6N_8H]$ | 180.0 | 0.0 | 174.9 |
| $[O_7C_6N_8H]$ | 0.0 | 180.0 | −5.2 |
| $[C_5C_6N_8C_9]$ | | 180.0 | 9.5 |
| $[C_5C_6N_8C_1]$ | 0.0 | | |
| $[O_7C_6N_8C_9]$ | | 0.0 | −170.6 |
| $[O_7C_6N_8C_1]$ | 180.0 | | |

**TABLE 8: Calculated Core−Electron Binding Energies (in eV) of NCNLD**

| atom[a] | NCNLD |
|---|---|
| $C_1$ | 291.28 |
| $C_2$ | 293.53 |
| $O_3$ | 537.09 |
| $N_4$ | 406.17 |
| $C_5$ | 291.90 |
| $C_6$ | 293.30 |
| $O_7$ | 536.85 |
| $N_8$ | 405.97 |
| $C_9$ | 291.95 |

[a] See Figure 5 for numbering.

CEBEs computed for the NCNLD are presented in Table 8. For the DKP-like moiety, we expected calculated binding energies intermediate between those obtained for the two model dipeptides. This is clearly not the case since CEBEs for this part of the molecule are lower than those found for both model dipeptides. Therefore, we suggest that the large difference between CEBEs in DKP and in LMD is not principally due to the orientation of bonds in each dipeptides. The difference between the two sets of CEBEs of DKP and LMD rather comes from the intrinsic nature of these two dipeptides. The structure of LMD may be considered as two linked acetamide molecules. Besides, comparison of CEBEs of these two compounds shows similarities. For example, CEBEs of the $C_6$, $N_7$, and $O_8$ atoms in LMD and $C_{carb}$, O, and N in acetamide are within 0.2 eV. On the other hand, the structure of DKP can be viewed as two linked *N*-methylformamide molecules or two linked acetamide molecules. Although the CEBE of $C_2$ (and $C_6$) in DKP is closer to that of the carbonyl carbon in acetamide than that in *N*-methylformamide, the other three CEBEs of DKP are within 0.1 eV of those of *N*-methylformamide. This result suggests that DKP is better considered as two linked *N*-methylformamide units as suggested by Hirst and Persson.[20] To verify this assumption, we have computed the atomic charges for the DKP. They are reported in Table 5. Atomic charges computed for DKP are clearly closer to those obtained for *N*-methylformamide (AAD = 0.06) than those calculated for acetamide (AAD = 0.30). These new data support the conclusions above. Therefore, the two distinct sets of CEBEs encountered for the two model dipeptides are principally due the nature of each of them, i.e.,

LMD as two linked acetamide units and DKP as two linked *N*-methylformamide units.

## Conclusion

In this work, CEBE calculations have been performed on simple amides and dipeptides with our $\Delta E_{KS}$(PW86−PW91)/cc-pCVTZ method. First, we have definitively checked the validity of our method to fit XPS spectra on a new set of 46 compounds that have been added to the other 32 tested in a previous study.[14] For this complete set of 78 molecules, the deviation between theoretical and experimental CEBEs is excellent since the AAD is only 0.15 eV. We have then applied our method to the study of several simple amides that come from two parent compounds, formamide, and acetamide. As far as the amides are concerned, our results suggest that electron-donating methyl groups tend to lower the CEBEs of the three main atoms O, N, and $C_{carb}$. The downshift in the case of O(1s) and N(1s) CEBEs is more important in the acetamide series than in the formamide one, whereas the opposite trend is observed for $C_{carb}$(1s) CEBEs. This can be explained in terms of charges brought by these three atoms in the respective amides. We pointed out a strong correlation between the charge of the atom considered in the neutral compound and its corresponding CEBEs. Finally, two model dipeptides, one cyclic (DKP) and one noncyclic (LMD), have been investigated. They differ essentially by the orientation of peptide chain (bonds in cis in one case and trans in the other). Unlike in the case of the *N*-methylformamide, CEBEs of the nitrogen atom in the part least affected by the methyl terminations, i.e., $N_4$ in Figure 4, are very different from a dipeptide to another. CEBE calculations for a dipeptide with a structure intermediate between LMD and DKP (NCNLD) shown that this difference is not due to the sole orientation of bonds around the nitrogen atom. Comparison of the CEBEs of DKP and LMD with those computed for amides revealed that each dipeptide can be viewed as the linkage of two different subunits, i.e., two acetamide for LMD and two *N*-methylformamide for DKP. This result can explain therefore the discrepancies observed for the nitrogen atom in the two model dipeptides. Calculations of atomic charges in both dipeptides support this conclusion.

**Acknowledgment.** We thank Drs. Serrano-Andrés and Hirst for providing us with the coordinates of their model dipeptides. D.P.C. is grateful to the Natural Sciences and Engineering Research Council (NSERC) of Canada for financial support.

## References and Notes

(1) Chong, D. P.; Hu, C.-H.; Duffy, P. *Chem. Phys. Lett.* **1996**, *249*, 491.
(2) Pulfer, M.; Hu, C.-H.; Chong, D. P. *Chem. Phys.* **1997**, *216*, 91. For errata, see ref 4.
(3) Hu, C.-H.; Chong, D. P. *Chem. Phys.***1997**, *216*, 99.
(4) Chong, D. P.; Hu, C.-H. *J. Chem. Phys.* **1998**, *108*, 8950.
(5) Bureau, C.; Chong, D. P.; Lécayon, G.; Delhalle, J. *J. Electron Spectrosc. Relat. Phenom.* **1997**, *83*, 227.
(6) Cavigliasso, G.; Chong, D. P. *Can. J. Chem.* **1999**, *77*, 24.
(7) Chong, D. P.; Bureau, C. *J. Electron Spectrosc. Relat. Phenom.* **2000**, *106*, 1.
(8) Williams, A. R.; deGroot, R. A.; Sommers, C. B. *J. Chem. Phys.* **1975**, *63*, 628.
(9) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(10) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
(11) Pekeris, C. K. *Phys. Rev.* **1958**, *112*, 1649.
(12) Triguero, L.; Pettersson, L. G. M.; Ågren, H. *Phys. Rev. B* **1998**, *58*, 8097.
(13) Triguero, L.; Plashkevych, O.; Pettersson, L. G. M.; Ågren, H. *J. Electron Spectrosc. Relat. Phenom.* **1999**, *104*, 195.
(14) Cavigliasso, G.; Chong, D. P. *J. Chem. Phys.* **1999**, *111*, 9485.
(15) Pedocchi, L.; Russo, N.; Salahub, D. R. *Phys. Rev. B* **1993**, *47*, 12992.
(16) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1986**, *33*, 8800.
(17) Perdew, J. P.; Kurth, S. In *Density Functionals: Theory and Applications*; Joubert, D., Ed.; Springer, Berlin, 1998.
(18) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1995**, *103*, 4572.
(19) Serrano-Andrés, L.; Fülscher, M. P. *J. Am. Chem. Soc.* **1998**, *120*, 10912.
(20) Hirst, J. D.; Persson, B. J. *J. Phys. Chem. A* **1998**, *102*, 7519.
(21) Casida, M. E.; Daul, C. D.; Goursot, A.; Koester, A.; Pettersson, L.; Proynov, E.; St-Amant, A.; Salahub, D. R.; Duarte, H.; Dodbout, N.; Guan, J.; Jamorski, C.; Leboeuf, M.; Malkin, V.; Malkina, O.; Sim, F.; Vela, A. *deMon-KS*, Version 3.4; deMon Software: University of Montreal, Montreal, Canada, 1997.
(22) In *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*; Hellwege, K.-L., Ed.; Springer: Berlin, 1976; Vol. 7, New Series, Group II.
(23) In *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*; Madelung, O., Ed.; Springer: Berlin, 1987; Vol. 15, New Series, Group II.
(24) In *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*; Madelung, O., Ed.; Springer: Berlin, 1992; Vol. 21, New Series, Group II.
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; 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.; 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. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.5; Gaussian, Inc.: Pittsburgh, PA, 1998.
(26) Sorensen, S. L.; Osborne, S. J.; Ausmees, A.; Kikas, A.; Correia, N.; Svensson, S.; Naves de Brito, A.; Persson, P.; Lunell, S. *J. Chem. Phys.* **1996**, *105*, 10719.
(27) Sæthre, L. J.; Sværen, O.; Svensson, S. L.; Osborne, S.; Thomas, T. D.; Jauhiainen, J.; Aksela, S. *Phys. Rev. A* **1997**, *55*, 2748.
(28) Sundin, S.; Sæthre, L. J.; Sorensen, S. L.; Ausmees, A.; Svensson, S. *J. Chem. Phys.* **1999**, *110*, 5806.
(29) Jolly, W. L.; Bomben, K. D.; Eyermann, C. J. *Atomic Data Nucl. Data Tables* **1984**, *31*, 433.
(30) (a) Kitano, M.; Fukuyama, T.; Kuchitsu, K. *Bull. Chem. Soc. Jpn.* **1973**, *46*, 384. (b) Kitano, M.; Kuchitsu, K. *Bull. Chem. Soc. Jpn.* **1973**, *46*, 3048. (c) Kitano, M.; Kuchitsu, K. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 67. (d) Kitano, M.; Kuchitsu, K. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 631.
(31) Greenberg, A.; Moore, D. T. *J. Mol. Struct.* **1997**, *413*−*414*, 477.
(32) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.
(33) Aplincourt, P.; Bureau, C.; Anthoine, J. L.; Chong, D. P. *J. Phys. Chem. A* **2001**, *105*, 7364.
(34) Ciarkowski, J. *Biopolymers* **1984**, *23*, 397.
(35) Jonsson, M.; Wayner, D. D. M.; Armstrong, D. A.; Yu, D.; Rauk A. *J. Chem. Soc., Perkin Trans. 2* **1998**, 1967.
(36) Bettens, F. L.; Bettens, R. P. A.; Brown, R. D.; Godfrey, P. D. *J. Am. Chem. Soc.* **2000**, *122*, 5856.