Trends in Strong Chemical Bonding in C2, CN, CN-, CO, N2, NO, NO+, and O2

Kepp, Kasper Planeta

Published in:

Link to article, DOI:
10.1021/acs.jpca.7b08201

Publication date:
2017

Document Version
Peer reviewed version

Citation (APA):
Trends in Strong Chemical Bonding in C₂, CN, CN⁻, CO, N₂, NO, NO⁺, and O₂

Kasper P. Kepp*

Technical University of Denmark, DTU Chemistry, Building 206, 2800 Kgs. Lyngby, DK – Denmark. *Phone: +045 45 25 24 09. E-mail: kpj@kemi.dtu.dk

ABSTRACT: The strong chemical bonds between C, N, and O play a central role in chemistry, and their formation and cleavage are critical steps in very many catalytic processes. The close-lying molecular orbital energies and large correlation effects pose a challenge to electronic structure calculations and have led to different bonding interpretations, most notably for C₂. One way to approach this problem is by strict benchmark comparison of related systems. This work reports reference electronic structures and computed bond dissociation enthalpies D₀ for C₂, CN, CN⁻, CO, N₂, NO, NO⁺, O₂ and related systems C₂⁺ and C₂⁻ at chemical accuracy (~1 kcal/mol or 4 kJ/mol) using CCSD(T)/aug-cc-pV5Z, with additional benchmarks of HF, MP2, CCSD, explicitly correlated F12 methods, and four density functionals. Very large correlation and basis set effects are responsible for up to 93% of total D₀. The order of the molecular orbitals 1π_u and 3σ_g changes, as seen in text books, depending on total and effective nuclear charge. Linear trends are observed in 2σ_u – 2σ_g orbital splitting (R² = 0.91) and in D₀ of C₂, C₂⁻, and C₂⁺ (R² = 0.99). The correlation component of D₀ of C₂ is by far the largest (~93%) due to a poor HF description. Importantly, density functional theory fails massively in describing this series consistently in both limits of effective nuclear charge, and Hartree-Fock exchange or meta functionals do not remedy this 100 kJ/mol-error, which should thus be addressed in future density functional developments as it affects very many studies involving cleavage or formation of these bonds.
INTRODUCTION

The diatomic molecules of second-period elements C, N, and O are fundamental to our understanding of chemistry and are the basis of teaching chemical bonding and molecular orbital (MO) theory in chemistry classes. N₂ is notoriously inert and present at 78% in the atmosphere, whereas O₂, with its paramagnetic state and role as main electron acceptor of higher life forms, makes up most of the remaining part. CO and NO are important biological messenger molecules and are central to a wide range of catalytic processes, and the molecules play a major role as ligands in coordination chemistry.

Accordingly, these simple molecules have been widely studied for a long time. N₂ has revealed the dramatic failure of Møller-Plesset perturbation theory for such systems, which require accurate treatment of electron correlation. The 1\(\pi_u\) and 3\(\sigma_g\) molecular orbitals are conspicuously close in energy and differ in their occupation as noted in typical text books. As discussed further below, even the 2\(\sigma_u\) MO can be close to 1\(\pi_u\) and 3\(\sigma_g\), providing in some cases space for eight electrons within 0.2-0.3 eV. Some of the strongest bonds known in chemistry include the triple bonds of CN⁻, CO, and N₂. The molecules routinely enter density functional calculations via the description of catalytic processes, even though these tightly bound systems can produce large errors that could affect overall reliability of such studies. As an example, any model that estimates how O₂ or N₂ is cleaved by a catalyst is sensitive to the chosen method’s error in describing the respective bond dissociation enthalpy, D₀. Because of this complexity, our understanding of the chemical bonding in these species is regularly challenged, both during interaction with transition metal complexes such as O₂ binding to heme, and for the bonds themselves, as is the case for C₂. In both examples, the different
interpretations from MO theory and valence bond theory play a notable role: Thus, C$_2$ can yield valence bond structures resembling a quadruple bond$^{24,27,28}$. Single-determinant MO theory traditionally describes C$_2$ as a double bond$^{22,25,29,30}$, but different orbital definitions in the configuration state functions from multi-configurational self-consistent field (MC-SCF) methods can help reconcile MO and valence bond theory$^{31,32}$. The main requirement of any such model is of course 1) agreement with all available experimental data, notably D$_0$, and 2) predictive value, e.g. in relation to catalytic cleavage of the chemical bond in question.

Particular questions of interest include: 1) What is required to accurately (within 1 kcal/mol) compute D$_0$ of these molecules? 2) How large are the correlation effects in these strongly bonded molecules? 3) What is the order of MO energies in the various species and how does it affect bonding? 4) To which extent do density functionals have problems with these systems? 5) Do any of the molecules behave irregularly in the overall comparison, and if so, how? To address these questions, the electronic structures of a range of diatomic molecules and molecular ions were computed with the aim to reproduce experimental D$_0$ to within chemical accuracy (defined as errors smaller than 1 kcal/mol or ~4 kJ/mol) and subsequently, at this level of accuracy, analyze trends in electronic structure.

COMPUTATIONAL DETAILS

All computations were performed using the Turbomole software, version 7.0$^{33}$. All densities and energies were converged to 10$^{-7}$ a.u., and the resolution of identify approximation was used to speed up all HF, MP2, CCSD, and CCSD(T) calculations$^{34,35}$. In some cases also the explicitly correlated F12-CCSD and F12-CCSD(T) results were computed$^{36}$. Experimental bond lengths from NIST were used as follows: NO$^+$ = 1.0657 Å, NO = 1.1538 Å, O$_2$ = 1.2075 Å, N$_2$ = 1.0977 Å.
Å, CO = 1.1282 Å, CN = 1.1718 Å, CC = 1.2425 Å; CN\(^-\) = 1.177 Å. The energies of all molecules and atoms were computed using the aug-cc-pV5Z basis set\(^{37}\) (numerical data in Table S1 with errors vs. experiment in Table S2). The bond dissociation enthalpies were computed as:

\[
D_0(XY) = E_{el}(X) + E_{el}(Y) - E_{el}(XY) - ZPE(XY)
\]

(1)

where \(E_{el}(X)\) is the electronic energy of species X computed and ZPE(XY) is the vibrational zero-point energy of XY (numerical details in Table S3). For NO\(^+\), the fragments are N and O\(^+\), because any double occupation of p-orbitals is avoided, as seen also by comparison of the energies \(E(N) + E(O^+)\) vs. \(E(N^+) + E(O)\) (Supporting Information, Table S1). Similar reasoning applies to CN\(^-\), which forms C\(^-\) and N\(^38\). To investigate the role of the basis set in describing accurately the bonding, the following smaller basis sets were further tested for all eight molecules C\(_2\), CN, CN\(^-\), CO, N\(_2\), NO, NO\(^+\), and O\(_2\): \text{def2-SVP, def2-TZVPP, def2-QZVPP}^{39}, \text{aug-cc-pVTZ, and aug-cc-pVQZ}^{37}\) (numerical data in Tables S4-S10). Scalar-relativistic corrections (Cowan-Griffin approximation\(^{40}\)) for the energies of X, Y and XY were also computed and increase the \(D_0\) by 0.3-1.3 kJ/mol, the least for C\(_2\) and the most for NO\(^+\), in line with the effective nuclear charge and associated compactness of the 1s orbitals (Table S11). Hence, relativistic corrections were not considered further in this work. As the experimental data are in gas phase, the enthalpy corrections \(\Delta H(X)\) were further computed as \(3/2 \text{RT}\) for each \(D_0\).

Four density functionals were investigated to understand how such methods perform in the “tight-bonding” regime studied here: PBE0, PBE, B3LYP, and TPSSH. Among these, PBE0\(^{41}\) and B3LYP\(^{42-44}\) represent hybrid GGA functionals with 25% and 20% HF exchange, whereas PBE is a GGA non-hybrid\(^{41}\), and TPSSH is a meta hybrid functional\(^{45,46}\). To understand C\(_2\) in more detail, C\(_2^+\), C\(_2^-\), C\(^+\), and C\(^-\) were also studied as specified in Table S12, in order to compute \(D_0\) of C\(_2^+\) and C\(_2^-\) for comparison with C\(_2\) (numerical details in Table S13). The geometries of C\(_2^+\), and C\(_2^-\) were
obtained by geometry optimization at the B3LYP/def2-TZVPP level, which gives an error vs. the experimental bond lengths of C₂ of 0.005 Å (B3LYP: 1.247 Å; BLYP: 1.256 Å; TPSSh: 1.251 Å).

RESULTS AND DISCUSSION

Correlation and Basis Set Effects on Computed D₀ Values. The D₀ values were computed according to Equation (1), using HF, MP2, CCSD, CCSD(T), and the four density functionals PBE0, PBE, TPSSh, and B3LYP, and are shown in Table 1. The computed zero-point energies of the molecules (Table S3 in Supporting Information) range from 10-15 kJ/mol (9.7 for O₂; 14.6 for NO⁺) and are thus required to accurately reproduce the experimental D₀. In contrast, relativistic corrections affect D₀ by ~1 kJ/mol for the systems (Supporting Information, Table S11) and thus do not affect accuracy, as expected for second-period atoms.

Table 1. Computed D₀ (kJ/mol) using different methods corrected for zero-point energy and enthalpy (+3/2 RT); basis set: aug-cc-pV5Z.

<table>
<thead>
<tr>
<th></th>
<th>HF</th>
<th>MP2</th>
<th>CCSD</th>
<th>CCSD(T)</th>
<th>PBE0</th>
<th>PBE</th>
<th>TPSSh</th>
<th>B3LYP</th>
<th>EXP.</th>
<th>E_corr a</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO⁺</td>
<td>502.0</td>
<td>1105.7</td>
<td>1000.2</td>
<td>1046.1</td>
<td>1048.2</td>
<td>1150.8</td>
<td>1042.1</td>
<td>1063.8</td>
<td>1046.9</td>
<td>0.52</td>
</tr>
<tr>
<td>O₂</td>
<td>134.0</td>
<td>545.7</td>
<td>461.9</td>
<td>497.0</td>
<td>516.5</td>
<td>595.5</td>
<td>497.7</td>
<td>509.0</td>
<td>498.5</td>
<td>0.73</td>
</tr>
<tr>
<td>NO</td>
<td>215.7</td>
<td>659.4</td>
<td>589.3</td>
<td>627.9</td>
<td>635.7</td>
<td>712.6</td>
<td>619.7</td>
<td>637.4</td>
<td>630.6</td>
<td>0.66</td>
</tr>
<tr>
<td>N₂</td>
<td>472.5</td>
<td>995.6</td>
<td>902.2</td>
<td>924.4</td>
<td>934.2</td>
<td>1008.6</td>
<td>917.2</td>
<td>944.9</td>
<td>944.9</td>
<td>0.50</td>
</tr>
<tr>
<td>CO</td>
<td>721.0</td>
<td>1136.4</td>
<td>1045.1</td>
<td>1079.0</td>
<td>1060.9</td>
<td>1115.4</td>
<td>1035.9</td>
<td>1055.3</td>
<td>1076.6</td>
<td>0.33</td>
</tr>
<tr>
<td>CN</td>
<td>369.6</td>
<td>697.6</td>
<td>701.6</td>
<td>742.7</td>
<td>740.4</td>
<td>816.8</td>
<td>729.3</td>
<td>737.8</td>
<td>749.3</td>
<td>0.50</td>
</tr>
<tr>
<td>CC</td>
<td>42.9</td>
<td>658.0</td>
<td>525.0</td>
<td>608.5</td>
<td>497.4</td>
<td>594.3</td>
<td>495.1</td>
<td>489.2</td>
<td>605.0</td>
<td>0.93</td>
</tr>
<tr>
<td>CN⁻</td>
<td>604.0</td>
<td>1050.3</td>
<td>964.9</td>
<td>1000.7</td>
<td>978.2</td>
<td>1025.3</td>
<td>958.7</td>
<td>993.7</td>
<td>N/A</td>
<td>0.40</td>
</tr>
</tbody>
</table>

MSE -442.0 35.2 -46.6 -1.2 -16.9 64.4 -29.4 -15.1
MAE 442.0 50.0 46.6 2.8 23.9 67.1 29.6 26.3

a Correlation energy contribution to D₀, estimated as D₀(CCSD(T)) – D₀(HF) / D₀(CCSD(T)).
Table 1 shows that CCSD(T)/aug-cc-pV5Z achieves chemical accuracy (MAE ~2.8 kJ/mol) for these systems. Thus, CCSD(T) at this basis set level accurately describes the bonding and correlation energy of the systems. From Table 1, the correlation energy, defined as $E_{\text{CCSD(T)}} - E_{\text{HF}}$, constitutes a very large part and often more than half of the total $D_0$. $C_2$ stands out by having a very poor HF description of bonding, as pointed out previously\(^3\), and a corresponding very large (~ 93%) contribution of correlation energy to $D_0$. This makes the C-C bond in $C_2$ the most correlated of those in the series.

Table 2. $D_0$ computed for $C_2$ using different wave-function methods and basis sets.

<table>
<thead>
<tr>
<th>Basis set</th>
<th>HF</th>
<th>MP2</th>
<th>CCSD</th>
<th>CCSD(T)</th>
<th>F12-CCSD</th>
<th>F12-CCSD(T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>aug-cc-pV5Z</td>
<td>41.7</td>
<td>656.7</td>
<td>523.8</td>
<td>607.2</td>
<td>523.9</td>
<td>606.1</td>
</tr>
<tr>
<td>aug-cc-pVQZ</td>
<td>41.6</td>
<td>650.2</td>
<td>519.2</td>
<td>601.4</td>
<td>524.6</td>
<td>604.7</td>
</tr>
<tr>
<td>aug-cc-pVTZ</td>
<td>39.0</td>
<td>648.6</td>
<td>522.7</td>
<td>601.9</td>
<td>525.8</td>
<td>601.2</td>
</tr>
<tr>
<td>def2-QZVPP</td>
<td>41.5</td>
<td>647.0</td>
<td>517.1</td>
<td>598.8</td>
<td>525.9</td>
<td>604.9</td>
</tr>
<tr>
<td>def2-TZVPP</td>
<td>40.2</td>
<td>628.0</td>
<td>502.6</td>
<td>580.9</td>
<td>524.9</td>
<td>598.8</td>
</tr>
<tr>
<td>def2-SVP</td>
<td>40.7</td>
<td>612.1</td>
<td>513.0</td>
<td>585.1</td>
<td>552.9</td>
<td>616.7</td>
</tr>
</tbody>
</table>

As seen from a basis set effect study of $C_2$ in Table 2, high accuracy is only achieved when using very large basis sets, and chemical accuracy (1 kcal/mol) requires aug-cc-pV5Z. Basis set effects have been studied before for some of these molecules, reaching similar conclusions\(^12,47\). Even a polarized triple-zeta basis set underestimates bonding considerably. This can be traced to the large correlation energy and substantial electron-electron cusp of the tightly bound electrons as previously discussed for $N_2$\(^48\). Basis set effects for the other seven molecules are summarized in Supporting Information, Tables S4-S10. The basis set effects estimated from CCSD(T) are larger for the triple-bonded $N_2$ (~77.8 kJ/mol from def2-SVP to aug-cc-pV5Z, Table S4). For $O_2$ the largest difference is 47.6 kJ/mol, whereas it is 58.5 kJ/mol for NO. The basis set errors for all eight molecules scale somewhat with $D_0$ (correlation coefficient of linear regression $R \sim 0.33$) as estimated by plotting the standard deviation of computed $D_0$ for the six basis sets against $D_0$ at the
CCSD(T)/aug-cc-pV5Z level (Figure S1). C₂ which has an average D₀ in the series thus has a less pronounced basis set effect of 22.1 kJ/mol (Table 2).

Although Table 1 already shows that the basis set is saturated from the point of view of D₀ at aug-cc-pV5Z (as seen from the uniform high accuracy for all systems), a further validation of the basis set saturation was carried out using explicitly correlated F12 methods⁴⁹ for CCSD, and CCSD(T)³⁶, as shown in the last two columns of Table 2; they show faster convergence with smaller basis sets towards the value of the aug-cc-pV5Z basis set. Chemical accuracy is achieved at the polarized quadruple-zeta or even triple-zeta level using F12 methods because the electron-electron cusp is recovered at lower basis set costs. Polarization functions are essential to polarize electrons within the bonding region. Again, MP2 consistently overbinds and CCSD consistently underbinds relative to CCSD(T) for a given basis set. Importantly, D₀ increases monotonically with basis set size, because the tightly bound diatomic molecules benefit the most from additional freedom to distribute the electrons. Since small basis sets underbind, MP2 provides an example of commonly encountered error cancellation, with MP2/def2-TZVPP being within 10 kJ/mol of experiment due to cancellation of errors in the one-electron basis and correlation treatment.

As seen from Table 2, any model based on a triple-zeta basis set or smaller will miss some part of D₀, as the electrons are not optimally distributed in the bonding region. An indication of this is seen in Supporting Information, Figure S2, showing that differences in the electron density along the bond axis are very basis-set dependent. More electron density is recruited to the bonding region as correlation is recovered accurately.

**Trends in Accuracy vs. Effective Nuclear Charge.** The individual errors for each method and system, defined as D₀ (computed) - D₀(experimental), are shown in Figure 1 (numerical data in Table S2). The largest error for CCSD(T) (6.6 kJ/mol) is seen for the open-shell molecule CN, and uniform high accuracy is otherwise evident. Figure 1 shows that both MP2 and CCSD fail in
producing accurate bonding, and they miss by approximately the same amount in most cases, but in different directions such that CCSD underbinds and MP2 overbinds; accordingly, \( E = \frac{1}{2} (E_{\text{MP2}} + E_{\text{CCSD}}) \) gives a conspicuously good estimate of \( D_0 \) in most cases (except for CN).

![Figure 1](image_url)

**Figure 1.** Errors in computed \( D_0 \), corrected for vibrational zero-point energies and enthalpy.

The molecules in Figure 1 have been ordered according to effective nuclear charge, with the highest effective nuclear charge listed to the left (this order is also the same as for the general trend in MO energies, as discussed below). From this choice of ordering, some interesting observations can be made: While the *ab initio* methods MP2, CCSD, and CCSD(T) show no general trend in \( D_0 \) (computed) - \( D_0 \) (experimental) with this ordering, the density functionals all show a trend, in particular if CN is placed with the other open-shell molecule NO. For PBE, the trend is very clear because the GGA functional overbinds by > 100 kJ/mol for \( \text{NO}^+ \), with a trend towards good accuracy but somewhat underbinding in \( \text{C}_2 \). However, the three hybrid functionals confirm this tendency but are generally less strongly binding, and are thus accurate for \( \text{NO}^+, \text{O}_2, \) and \( \text{NO} \), but increasingly inaccurate due to underbinding in \( \text{N}_2, \text{CN}, \text{CO}, \) and in particular \( \text{C}_2 \). \( \text{C}_2 \) stands notably out in the series with the smallest effective nuclear charge and the most diffuse...
electrons; the hybrid functionals and CCSD fail to bind strongly enough and any hybrid functional, regardless of the amount of HF exchange, underbinds by 100 kJ/mol.

This comparison shows that none of the density functionals provide a balanced description of bonding through the series, and the imbalance amounts to ~100 kJ/mol for all four density functionals: *In other words, no density functional is capable of describing this series even remotely correctly in both limits of high and low effective nuclear charge.* Because the aug-cc-pV5Z basis set is saturated as shown above, this discrepancy is due to the functional design, but neither HF exchange at any amount, nor the meta functional character (TPSSh), nor the quantum mechanical bounds on the density imposed by PBE and TPSSh can remedy this error, which relates to the balance between tight and weak bonding regimes enforced by the effective nuclear charge. This makes the series studied here an outstanding problem to DFT that begs further scrutiny also in the development of new functionals. This is particularly true because of the importance of these molecules in catalysis, and because the 100-kJ/mol errors will strongly affect the reliability of DFT studies in catalysis where O-O, C-O, and N-O bonds are formed or cleaved.

**Ordering of Molecular Orbital Energies.** Figure 2 shows the HF molecular orbital energies of the largest basis set, aug-cc-pV5Z, ordered according to decreased effective nuclear charge. For the unrestricted calculations the energies are averages of the α- and β-orbitals. For smaller basis sets, the more diffuse virtual MOs would be less well described than occupied MOs but at this basis set level, the general trends are quite smooth and thus worthy of discussion. For the neutral species the 1π_u orbitals change regularly and only by about 0.2 eV due to effective nuclear charge from O_2, over NO, N_2, CO, and CN, to C_2, whereas all the σ-orbitals 2σ_g, 2σ_u and 3σ_g change by ~0.5 eV along the neutral series, because their more compact nature renders them more sensitive to nuclear charge; this difference is responsible for the change in MO ordering. N_2 and O_2 have the same order of MOs, with 3σ_g slightly lower than 1π_u, as confirmed also for the
intermediate case NO. However, in CO, more so in CN, and even more so in C₂ (in order of reduced effective nuclear charge) 3σ₈ increases substantially in energy compared to 1πᵤ because 3σ₈ (which is formed by the pₓ orbitals) is more sensitive to the reduced effective nuclear charge than 1πᵤ. In CN⁻ the total charge of an additional electron causes destabilization of all MOs, compared to CN.

**Figure 2.** HF/aug-cc-pV5Z orbital energies (average of alpha and beta for unrestricted calculations).

**Figure 3.** D1 diagnostic for the CCSD plotted against A) the fraction of the total bond dissociation enthalpy due to correlation, \((D₀^{CCSD(T)} - D₀^{HF}) / D₀^{CCSD(T)}\), B) total computed \(D₀^{CCSD(T)}\).
Correlation Complexity (Measured by D1) Grows Linearly with Bond Strength. To further understand the challenge of describing the bonding in these molecules, Figure 3 shows the D1 diagnostic obtained for the CCSD calculations\textsuperscript{50} at the aug-cc-pV5Z level, plotted against the fraction of \( D_0 \) that is due to correlation (Figure 3A) and the total \( D_0 \), computed at the CCSD(T) level (Figure 3B). The first plot indicates how the quality of the CCSD calculation depends on how important electron correlation is for the total bond energy, whereas the second plot simply shows the correlation with total \( D_0 \). Except for the radicals CN and NO, and \( C_2 \) all D1 diagnostics are smaller than 0.05, implying that CCSD treats these systems well without any complications of multi-configuration mixing. However, \( C_2 \) is notable by having the clearly largest D1 value among the closed-shell systems. The same order for three of these molecules CN > \( C_2 \) > O\(_2\) was obtained previously\textsuperscript{50}. All the triple-bonded closed-shell systems display low D1 values and also cluster tightly in the plot against the correlation component of \( D_0 \) (Figure 3A). For the two other groups of formal bond orders (BO) 2½ and 2, the molecule with the smallest effective nuclear charge displays the highest D1 diagnostic, consistent with the challenge of computational chemistry discussed above relating to the low effective nuclear charge. When D1 is plotted directly against \( D_0 \) as shown in Figure 3B, however, strongly linear relationships are obtained for the two groups of \( \text{BO} = 2, 2\frac{1}{2}, \) and \( \text{BO} = 3 \), showing that the quality of the CCSD treatment decreases essentially linearly (\( R^2 = 0.98-0.99 \)) with the strength of bonding, i.e. the tight bonding regime, which again relates to the electron cusp as discussed previously for \( N_2 \)\textsuperscript{48}, but the sensitivity is much larger for the \( \text{BO} = 2,2\frac{1}{2} \) group where the open-shell molecules and extreme \( C_2 \) case occur.

**Correlation Effects and Bonding in \( C_2 \).** While the comparison above provides insight into the series of molecules as a whole, in the following the special case of \( C_2 \) is discussed a bit more due to its prominent featuring in recent discussions\textsuperscript{22–25,27}. The most significant debate is on the suitability of describing \( C_2 \) as a quadruply bound molecule using valence bond theory or simply
as a double bond as obtained from MO diagrams. As discussed previously\textsuperscript{24,27}, valence bond theory can lead to construction of a fourth loose bond formed by the two anti-bonding (in MO theory) $2\sigma_u$ electrons, whereas in single-determinant MO theory this orbital is considered strictly anti-bonding. One argument against the double bonded C\textsubscript{2} is that $2\sigma_u$ is ”fuzzy”, i.e. it is only weakly anti-bonding or perhaps even non-bonding. In valence bond theory these two electrons can be described as being in an ”inverted” bond\textsuperscript{28}.

Without going into discussions on this debate, it is relevant to discuss C\textsubscript{2} within the present benchmark CCSD(T)/aug-cc-pV5Z computations because they provide $D_0$ at chemical accuracy (Table 1) as not seen in these previous discussions on C\textsubscript{2}. One way to do this is by bonding analysis, which has been done excessively in the past\textsuperscript{24–27,29}. Another is to compare directly the electronic correlation energies at the CCSD(T)/aug-cc-pV5Z level, where experimental $D_0$ is recovered, for a series of related systems that includes C\textsubscript{2}.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{A) The trends in $D_0$ computed using MP2, CCSD, and CCSD(T) for C\textsubscript{2}$^+$, C\textsubscript{2}, and C\textsubscript{2}$^-$; all three methods show almost perfectly linear trends in $D_0$. B) Plot of the energy of the bonding $2\sigma_b$ vs. that of the anti-bonding $2\sigma_a$ for NO$^+$, NO, O\textsubscript{2}, N\textsubscript{2}, CO, CN, C\textsubscript{2}, and CN$^-$.}
\end{figure}
Figure 4A shows a comparison of \( D_0 \) obtained using MP2, CCSD, and CCSD(T) for \( C_2 \) and its one-electron-enriched or depleted counterparts, \( C_2^- \) and \( C_2^+ \) (numerical data in Tables S12 and S13). According to MO theory, the \( C_2 \) bond should be between \( C_2^- \) and \( C_2^+ \) in strength because the additional electron is placed in the \( 3\sigma_g \) orbital which is bonding, and formally the BO increases from 2 to 2\( \frac{1}{2} \), whereas in \( C_2^+ \) a bonding electron is lost from \( 1\pi_u \) (see Figure 1) to produce a BO of 1\( \frac{1}{2} \). It is not clear what the relationship will be if the true bonding in \( C_2 \) is distinct from what simple single-determinant MO theory predicts (BO = 2). Possibly \( D_0 \) of \( C_2 \) would still be smaller than \( D_0 \) of \( C_2^- \), but if the bond in \( C_2 \) is unique, one would expect some deviation from an exact linear trend, if \( C_2^- \) and \( C_2^+ \) are normal and disrupt the unique quadruple bond character, even if \( C_2^- \) is stronger as measured by \( D_0 \). Figure 4A shows almost perfectly linear relationships for \( D_0 \) computed with all three methods MP2, CCSD, and CCSD(T) with \( R^2 \) values of 0.99, i.e. \( C_2 \) falls exactly in between \( C_2^- \) and \( C_2^+ \) in terms of its \( D_0 \). This observation is fully in agreement with the expectations from basic MO theory of a double bonded \( C_2 \), but it is possible that valence bond theory can also explain this linearity with \( C_2 \) still being qualitatively different from the two other bonds.

Figure 4B shows another comparison of the energies of the bonding \( 2\sigma_g \) vs. the energy of the anti-bonding \( 2\sigma_u \) for \( NO^+, NO, O_2, N_2, CO, CN, C_2, \) and \( CN^- \). In single-determinant MO theory these two orbitals form a tightly correlated bonding-antibonding set, and the antibonding orbital \( 2\sigma_u \) plays a major role in the left-right electron correlation description of the bonds. Both these two orbitals become more negative with higher effective nuclear charge, as expected. The linear regression suggests that the two orbitals are strongly correlated with \( R^2 = 0.91 \) and \( C_2 \) falls very close to the regression line. These electrons are mainly located along the (left-right correlated) axis of the molecules and thus do not interfere with the \( \pi \)-system.
One picture that summarizes these various observations is that C$_2$ represents very strong left-right correlation due to the low effective nuclear charge combined with the tight bonding, an effect that forces electrons that are formally anti- or non-bonding into the bonding region and requires very large augmented basis sets (Table 2). Such strong left-right correlation in C$_2$, CN, and CN$^-$ could be interpreted as additional bonding in various bonding models, yet it does not significantly affect D$_0$ which is perfectly normal, probably because this electron movement is compensated by alternative movements of other electrons as witnessed by the very large role of basis sets and triple corrections to CCSD. Thus, putting eight electrons in the bonding region is too costly even though it would complete the octet rule, but it takes a very large basis set and CCSD(T) to see the compensation that makes C$_2$ behave normally in Figure 4.

CONCLUSIONS

In summary, the experimental D$_0$ is well reproduced for all studied molecules where data are available using CCSD(T)/aug-cc-pV5Z; less correlated methods such as MP2 and CCSD are not sufficient, and smaller basis sets even at the quadruple-zeta level are not sufficient either. Thus, models based on limited basis sets may wrongly estimate D$_0$ and the amount of electrons formally residing in the bonding region. It is shown that electron correlation constitutes most of the bond energy in these tightly bound systems, and C$_2$ is extreme in this regard as the HF picture is very poor. The correlation energy amounts to an amazing 93% of the D$_0$ of C$_2$. Using the chemical accuracy CCSD(T) calculations, the D$_0$ of C$_2$ is shown to fall perfectly on a line with those of C$_2^-$ and C$_2^+$, consistent with a regular bonding behavior in C$_2$. It is then shown that effective nuclear charge provides a logical basis for comparing the systems both in terms of D$_0$ and in terms of errors associated with various methods. Notably, all four tested density functionals fail in describing bonding in one of the two limits of effective nuclear charge, and meta functionals or any amount
of HF exchange does not change this very large error of 100 kJ/mol across the series, making this a prominent problem to density functional theory. The order of MO energies plays a critical role in understanding these challenges; it depends on the effective nuclear charge, because more diffuse $1\pi_u$ orbitals are less stabilized by higher nuclear charge than the $\sigma$ orbitals. An analysis shows that the most difficult cases are CN and NO, which are the two open-shell systems.

ASSOCIATED CONTENT

Supporting Information

The supporting information is available free of charge on the ACS homepage. The file contains electronic energies at the aug-cc-pV5Z basis set level for molecules and atoms and ions (Table S1); errors in various methods vs. experimental $D_0$ (Table S2); computed zero point energies in a.u. and in kJ/mol (Table S3); basis set sensitivity study of $N_2$ (Table S4), $O_2$ (Table S5), NO (Table S6), CN (Table S7), CN$^-$ (Table S8), NO$^+$ (Table S9), and CO (Table S10); scalar-relativistic corrections (Table S11); electronic energies of carbon species (Table S12); computed $D_e$ of $C_2$, $C_2^-$, and $C_2^+$ (Table S13); basis set effects vs. bond strength (Figure S1); dependency of the HF electron density along the bond projection on basis set (Figure S2).

ACKNOWLEDGEMENTS

Computer time at the Aarhus supercomputer cluster "Grendel" (formerly the Danish Center for Scientific Computing) is gratefully acknowledged.
REFERENCES

(1) Mulliken, R. S. Electronic population analysis on LCAO–MO molecular wave functions. I. 

(2) Bader, R. F. W.; Bandrauk, A. D. Molecular charge distributions and chemical binding. III. 


(4) Murad, F. Discovery of some of the biological effects of nitric oxide and its role in cell 

(5) Dawson, T. M.; Snyder, S. H. Gases as biological messengers: Nitric oxide and carbon 


(7) Harcourt, R. D. Transition metal complexes with CO, N₂, NO and O₂ ligands. In Bonding 

(8) Frenking, G.; Fröhlich, N. The nature of the bonding in transition-metal compounds. Chem. 
Rev. 2000, 100 (2), 717–774.

(9) Ford, P. C.; Lorkovic, I. M. Mechanistic aspects of the reactions of nitric oxide with 

1967, 46 (9), 3320–3324.


Figure 1

204x93mm (300 x 300 DPI)
Figure 2

194x115mm (300 x 300 DPI)
Figure 3

186x60mm (300 x 300 DPI)
Figure 4

102x37mm (300 x 300 DPI)
TOC graphic

37x17mm (600 x 600 DPI)