

Computational Chemistry | Hot Paper |

Stephanie C. C. van der Lubbe^[a] and Célia Fonseca Guerra^{*[a, b]}

Dedicated to Professor Roald Hoffmann on the occasion of his 80th birthday

Abstract: Theoretical and experimental studies have elucidated the bonding mechanism in hydrogen bonds as an electrostatic interaction, which also exhibits considerable stabilization by charge transfer, polarization, and dispersion interactions. Therefore, these components have been used to rationalize the differences in strength of hydrogen-bonded systems. A completely new viewpoint is presented, in which the Pauli (steric) repulsion controls the mechanism of hydrogen bonding. Quantum chemical computations on the mismatched DNA base pairs CC and GG (C = cytosine, G = guanine) show that the enhanced stabilization and shorter distance of GG is determined entirely by the difference in the Pauli repulsion, which is significantly less repulsive for GG than for CC. This is the first time that evidence is presented for the Pauli repulsion as decisive factor in relative hydrogen-bond strengths and lengths.

The self-assembly characteristics of hydrogen bonds play an important role in many applicative fields, such as the development of self-healing materials,^[1–3] shape-memory polymers,^[4,5] hydrogels,^[6,7] drug delivery applications,^[8,9] and antisense technology.^[10] As the physical properties of materials can be controlled by tuning the strength of noncovalent interactions, a thorough understanding of the hydrogen-bonding mechanism is essential.^[11,12] Nevertheless, there is still an ongoing debate about the nature of this chemical interaction.^[13–16]

Most current undergraduate chemistry textbooks, as well as the International Union of Pure and Applied Chemistry (IUPAC) *Gold Book*, define hydrogen bonds as electrostatic interactions between an electronegative atom and a hydrogen atom attached to a second electronegative atom.^[17,18] As such, the hydrogen-bond strength is rationalized entirely by the size of the partial charges of the atoms that participate in the hydrogen bond.

However, experimental and theoretical studies have shown that hydrogen bonds are not purely electrostatic, but also partly covalent in nature.^[14,19] The idea of covalency was proposed in 1960 by Linus Pauling in his famous book *The Nature of the Chemical Bond*,^[20] in which he predicted that hydrogen bonds in ice have about 5% covalent character based on their relatively short bond lengths. This reasoning has been further developed since, and has been confirmed experimentally.^[14,21–24] As such, it is now generally believed that charge-transfer interactions enhance the hydrogen bonding via donor–acceptor interactions between the σ -lone pair orbital on the hydrogen-acceptor atom, and the anti-bonding σ^* empty orbital on the A–H group of the opposing monomer.^[25–27]

The importance of other stabilizing components has been investigated as well, including resonance-assistance by the π -electrons^[28–30] and dispersion interactions.^[31] However, an often-overlooked component that is used to tune hydrogen bond strength is the Pauli repulsion, which originates from the fact that electrons with the same spin are not allowed to be at the same position in space, and is a manifestation of the Pauli principle. It is the Pauli repulsion that is responsible for any steric repulsion, and thus prevents atoms from moving any closer toward each other. Herein, we will demonstrate that this repulsion can be a decisive factor for relative hydrogen-bonding strengths and lengths.

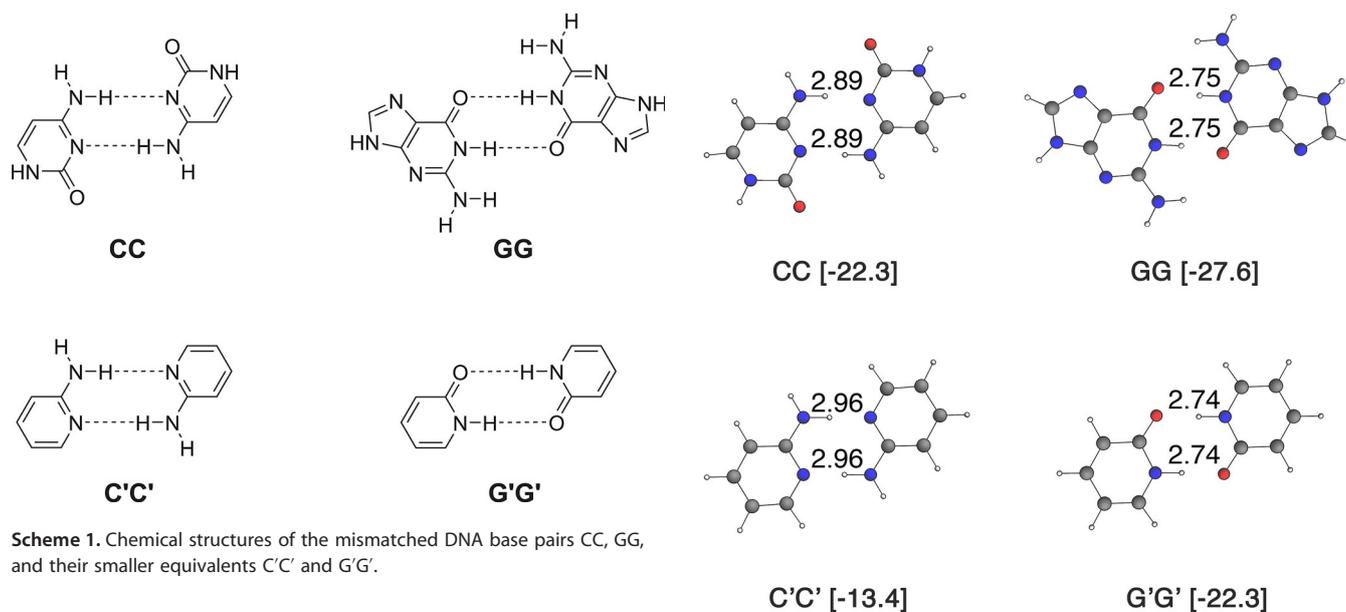
We investigated two symmetrical, mismatched DNA base pairs CC and GG (C = cytosine, G = guanine; Scheme 1). From previous theoretical work, it is known that the GG dimer is much stronger and has shorter hydrogen bonds than the CC pair.^[32,33] Both systems have resonance-assisted hydrogen bonds and the same amount of the well-known secondary electrostatic interactions.^[25,34] The pK_a values of the hydrogen-bond donor and acceptor groups do not explain the difference in binding strength between these dimers.^[35] Another possibility to rationalize the hydrogen-bond strength is by using the atomic charges of the frontier atoms. Different methods of computing atomic charges (see Supporting Information) point

[a] S. C. C. van der Lubbe, Prof. Dr. C. Fonseca Guerra
Department of Theoretical Chemistry and Amsterdam Center for Multiscale Modeling
Vrije Universiteit Amsterdam (The Netherlands)
E-mail: c.fonseca Guerra@vu.nl

[b] Prof. Dr. C. Fonseca Guerra
Leiden Institute of Chemistry, Gorlaeus Laboratories
Leiden University (The Netherlands)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/chem.201701821>.

© 2017 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited, and is not used for commercial purposes.



Scheme 1. Chemical structures of the mismatched DNA base pairs CC, GG, and their smaller equivalents C'C' and G'G'.

toward the greater charge on the oxygen atom in guanine as primary reason for the enhanced stability in GG. However, this seemingly logical reasoning does not hold, as our computational analyses will demonstrate.

We emphasize that our findings do not necessarily apply to DNA base pairs only, but rather aim to demonstrate that the Pauli repulsion can be a decisive factor in the strength of non-covalent interactions in general. Hydrogen-bond strengths and lengths are determined by an interplay of, among others, electrostatic, covalent, and steric interactions. The importance of each contribution is system-dependent and can be identified by performing quantum chemical computations.

We computed the hydrogen-bond strengths of the CC and GG pairs by using dispersion-corrected density functional theory at the BLYP-D3(BJ)/TZ2P level of theory, which is known to accurately reproduce the structural and energetic properties of biological hydrogen-bonded systems.^[36,37] The CC pair has a bond energy of $-22.3 \text{ kcal mol}^{-1}$, while the bond energy for GG is $-27.6 \text{ kcal mol}^{-1}$. The hydrogen bonds are 2.89 \AA for CC, and 2.75 \AA for GG (Figure 1).

To find the origin of these differences in hydrogen-bond strengths and lengths, we have included two simplified analogues of CC and GG in our analyses, namely C'C' and G'G' (Scheme 1). The advantage of these smaller analogues is that their number of π -electrons is identical, and they only incorporate functional groups that participate directly in the hydrogen bonds. A direct comparison between them is therefore more straightforward, and will help us to pinpoint the true cause of the difference in binding strength between the two mismatched DNA base pairs. The C'C' pair has a bond energy of $-13.4 \text{ kcal mol}^{-1}$, and the G'G' pair has a bond energy of $-22.3 \text{ kcal mol}^{-1}$. Thus, both bond energies are higher (that is less stable) than their GG and CC counterparts, but the trend remains unchanged, that is, $C'C' < G'G'$. The hydrogen bonds are 2.96 \AA for C'C', and 2.74 \AA for G'G' (Figure 1).

To understand the origin of the stronger interaction energy and shortened hydrogen bond lengths in GG and G'G', we ana-

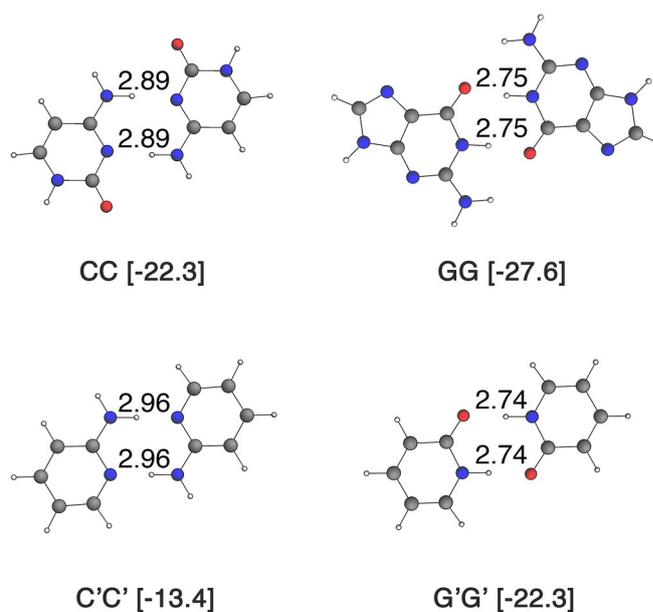


Figure 1. Hydrogen-bond distances [\AA] and bond energies between brackets [kcal mol^{-1}] for the fully optimized complexes CC, GG, C'C' and G'G' at the BLYP-D3(BJ)/TZ2P level of theory.

lyzed the bond energy of each dimer in terms of its original monomers. This is done by decomposing the bond energy ΔE into the preparation energy ΔE_{prep} and interaction energy ΔE_{int} as a function of the hydrogen-bond distance r :

$$\Delta E(r) = \Delta E_{\text{prep}}(r) + \Delta E_{\text{int}}(r) \quad (1)$$

The preparation energy ΔE_{prep} is the energy that is needed to deform the monomers from their optimal geometry into the geometry that they acquire in the dimer with hydrogen bond distance r . Since $\Delta E_{\text{prep}}(r)$ is very similar for all systems (see Supporting Information), we will focus on the interaction energy only. The interaction energy $\Delta E_{\text{int}}(r)$ accounts for the actual chemical interaction between the prepared monomers, and can be further decomposed into three physically meaningful terms:

$$\Delta E_{\text{int}}(r) = \Delta V_{\text{elstat}}(r) + \Delta E_{\text{Pauli}}(r) + \Delta E_{\text{oi}}(r) \quad (2)$$

The term ΔV_{elstat} corresponds to the classical electrostatic interactions between the prepared monomers, and is usually attractive in nature. The Pauli repulsion ΔE_{Pauli} comprises the destabilizing interactions between overlapping, occupied orbitals of the two monomers, and is responsible for any steric repulsion. The orbital interaction ΔE_{oi} accounts for charge transfer (namely donor-acceptor interactions between the hydrogen-bonded monomers) and polarization (empty-occupied orbital mixing on one monomer owing to the presence of another monomer). A theoretical overview of this energy decomposition energy (EDA) scheme is given in the Supporting Information and Ref. [38].

So, which energy component is responsible for the shape of the interaction profile, and thus for the position of equilibrium

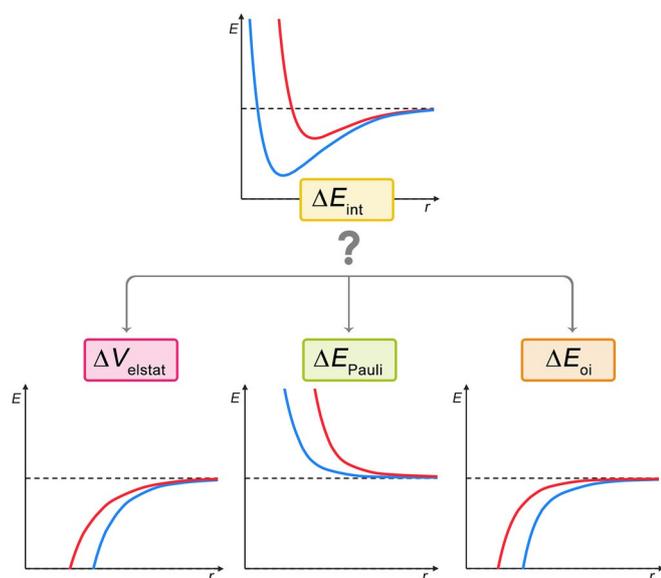


Figure 2. Overview of the interaction energy profile for GG (blue) and CC (red). Their positions of equilibrium could be determined by the electrostatic interaction ΔV_{elstat} , Pauli repulsion ΔE_{Pauli} , orbital interaction ΔE_{oi} or a combination of these terms.

(Figure 2)? In other words, are the relative hydrogen-bond strengths and lengths of GG, CC, and their smaller counterparts, determined by the electrostatic interaction, Pauli repulsion, orbital interaction, or a combination of these terms? We have addressed this question by computing the interaction energy profile of each dimer around its point of equilibrium. In

this approach, all O...H–N and N...H–N angles are kept linear, and the hydrogen-bond distances are varied over an interval from 2.70 Å to 3.00 Å with 0.01 Å per step. So, our analysis resulted in 31 new optimizations per complex, and reveals the driving force behind the stronger and shorter hydrogen bonds of GG and G'G'. The essential results are graphed in Figure 3; the complete dataset can be found in the Supporting Information.

The GG dimer is again more stable than the CC pair by up to 8.7 kcal mol⁻¹, even though their hydrogen-bond distances are the same. This confirms the interaction energy profile as schematically shown in Figure 2. Surprisingly, inspection of the attractive energy components ΔV_{elstat} and ΔE_{oi} reveals that these terms are more stable for CC, that is, the weaker bound pair, by up to 3.4 and 4.5 kcal mol⁻¹, respectively. In other words, the electrostatic and orbital interactions are more attractive for CC than for GG, but nevertheless, it is the GG pair that has the highest overall stability. As can be seen in Figure 3, this is entirely caused by the Pauli repulsion ΔE_{Pauli} , which is up to 17.5 kcal mol⁻¹ more repulsive for CC.

The smaller analogues G'G' and C'C' show similar trends; G'G' is up to 11.9 kcal mol⁻¹ more stable than C'C', even though the attractive energy components ΔV_{elstat} and ΔE_{oi} are more favorable for C'C' by up to 3.8 and 2.9 kcal mol⁻¹, respectively. Again, the higher relative stability for G'G' is entirely caused by the Pauli repulsion ΔE_{Pauli} , which is up to 20.0 kcal mol⁻¹ more repulsive for C'C' than for G'G'.

So, why then is the Pauli repulsion so much stronger for CC and C'C', than for GG and G'G'? We have addressed this question by studying the overlap between the filled orbitals of the

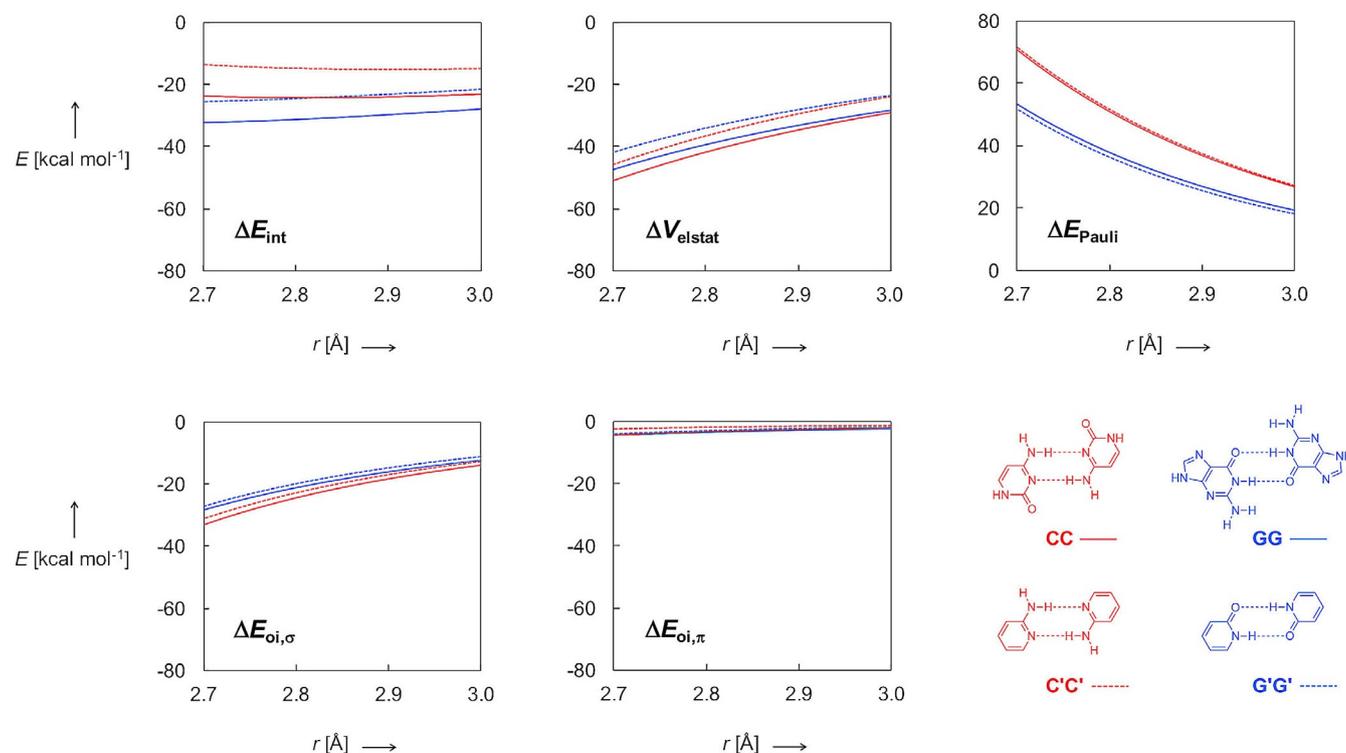


Figure 3. Decomposed energy terms [kcal mol⁻¹] as a function of the hydrogen-bond distance r [Å] for CC, GG, C'C', and G'G'. At each point, the dimers are optimized with constrained linear hydrogen bonds with distance r at the BLYP-D3(BJ)/TZ2P level of theory.

monomers, which provides us with an intuitive and chemically meaningful understanding of the difference in Pauli repulsion.^[38,39] Figure 4 presents the most important (i.e. largest) orbital overlap S^2 as a function of the hydrogen-bond distance, r , and shows the contour plots of the corresponding orbitals. (The contour plots of the second- and third-largest orbital overlaps are given in the Supporting Information.)

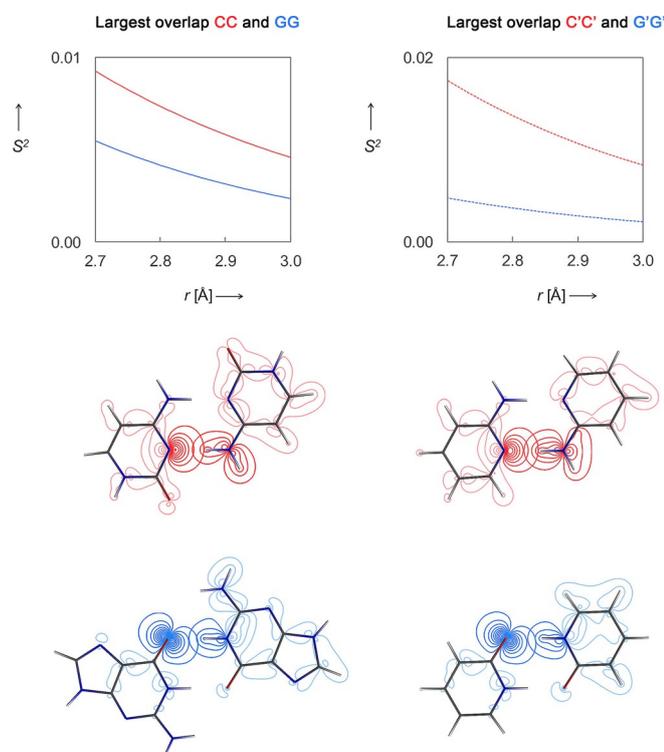


Figure 4. Largest orbital overlap S^2 as a function of the hydrogen-bond distance r [Å], and contour plots of the associated orbitals at 2.70 Å, computed at the BLYP-D3(BJ)/TZ2P level of theory. Contour plots contain 15 contours from 0.05–1.0 au.

The overlap between the filled orbitals is larger for CC and $C'C'$, than for GG and $G'G'$, which is in line with their stronger Pauli repulsion. Furthermore, the increase in overlap as a function of the hydrogen-bond distance r is larger for CC and $C'C'$ than for GG and $G'G'$. It is the steeper increase in Pauli repulsion that is responsible for the elongated hydrogen-bond lengths in the CC and $C'C'$ equilibrium structures.

As can be seen in the contour plots, the larger overlap and its steeper increase originates mainly from the alignment of the lone-pair orbital of the nitrogen atom with the H–N bond of the other monomer. So, it is the direction of the lone pair that plays an important role in the size of orbital overlap, and thus the size of the Pauli repulsion, and thus the overall stability of GG, CC, and their smaller equivalents. However, we have also found systems with the same lone pairs, in which the filled orbitals of the hydrogen-donating groups are responsible for the larger Pauli repulsion (see Supporting Information). Thus, the difference in Pauli repulsion can be caused by the direction of the lone pair only, but can also be determined by

the shape of the filled orbitals on the hydrogen-donating groups. These subtle effects can be captured by state-of-the-art quantum chemical software, which can assist supramolecular chemists in understanding, predicting, and tuning the interaction strength of self-assembled systems.

In conclusion, we have demonstrated that the determining factor for the enhanced stabilization and shortened hydrogen bonds of GG relative to CC is not (as we would predict from our chemical intuition) the electrostatic nor the donor–acceptor interaction, but the Pauli repulsion, which is up to 17.5 kcal mol⁻¹ larger for CC when both dimers have the same intermolecular distance. The larger Pauli repulsion originates mainly from the better alignment of the lone pair with the opposing H–N bond. Our results emphasize the complex nature of hydrogen bonds, which are an interplay between steric, electrostatic, covalent, cooperative, and π -resonance interactions, and highlight the importance of state-of-the-art quantum chemical analysis to shed light on their bonding mechanism.

Experimental Section

All calculations were performed using the Amsterdam Density Functional (ADF) program (2016.105) using dispersion-corrected density functional theory at the BLYP-D3(BJ)/TZ2P level of theory for geometry optimizations and energies. Full computational details are available in the Supporting Information.

Acknowledgements

We thank the Netherlands Organization for Scientific Research (NWO/CW) for financial support.

Conflict of interest

The authors declare no conflict of interest.

Keywords: bonding analysis • density functional calculations • DNA base pairs • hydrogen bonds • steric repulsion

- [1] N. Roy, B. Bruchmann, J.-M. Lehn, *Chem. Soc. Rev.* **2015**, *44*, 3786–3807.
- [2] P. Cordier, F. Tournilhac, C. Soulié-Ziakovic, L. Leibler, *Nature* **2008**, *451*, 977–980.
- [3] J.-L. Wietor, R. P. Sijbesma, *Angew. Chem. Int. Ed.* **2008**, *47*, 8161–8163; *Angew. Chem.* **2008**, *120*, 8282–8284.
- [4] L. Tan, J. Hu, K. Y. Rena, Y. Zhu, P. Liu, *J. Polym. Sci. Part A* **2017**, *55*, 767–775.
- [5] C. Liu, H. Qin, P. T. Mather, *J. Mater. Chem.* **2007**, *17*, 1543–1558.
- [6] X. Yan, F. Wang, B. Zheng, F. Huang, *Chem. Soc. Rev.* **2012**, *41*, 6042–6065.
- [7] J. Kopeček, J. Yang, *Angew. Chem. Int. Ed.* **2012**, *51*, 7396–7417; *Angew. Chem.* **2012**, *124*, 7512–7535.
- [8] B.-S. Kim, S. W. Park, P. T. Hammond, *ACS Nano* **2008**, *2*, 386–392.
- [9] L. J. De Cock, S. De Koker, B. G. De Geest, J. Grooten, C. Vervaet, J. P. Remon, G. B. Sukhorukov, M. N. Antipina, *Angew. Chem. Int. Ed.* **2010**, *49*, 6954–6973; *Angew. Chem.* **2010**, *122*, 7108–7127.
- [10] *Antisense Research and Application* (Ed: S. T. Crooke), Springer Verlag, Berlin, **1998**.
- [11] R. J. Wojtecki, A. Nelson, *J. Polym. Sci. Part A* **2016**, *54*, 457–472 and references therein.

- [12] S. I. S. Hendrikse, S. P. W. Wijnands, R. P. M. Lafleur, M. J. Pouderoijen, H. M. Janssen, P. Y. W. Dankers, E. W. Meijer, *Chem. Commun.* **2017**, 53, 2279–2282.
- [13] F. Weinhold, R. A. Klein, *Mol. Phys.* **2012**, 110, 565–579.
- [14] E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci, D. J. Nesbitt, *Pure Appl. Chem.* **2011**, 83, 1619–1636 and references therein.
- [15] G. Frenking, G. F. Caramori, *Angew. Chem. Int. Ed.* **2015**, 54, 2596–2599; *Angew. Chem.* **2015**, 127, 2632–2635.
- [16] F. Weinhold, R. A. Klein, *Angew. Chem. Int. Ed.* **2015**, 54, 2600–2602; *Angew. Chem.* **2015**, 127, 2636–2638.
- [17] IUPAC. *Compendium of Chemical Terminology 2nd ed.* (the “Gold Book”), Blackwell Scientific Publications, Oxford, **1997**.
- [18] C. E. Housecroft, E. C. Constable, *Chemistry 4th ed.* Person Education Limited, Harlow, **2010**.
- [19] H. Umeyama, K. Morokuma, *J. Am. Chem. Soc.* **1977**, 99, 1316–1332.
- [20] L. Pauling, *The Nature of the Chemical Bond 3rd ed.* Cornell University Press, Ithaca, **1960**.
- [21] J. Novotný, S. Bazzi, R. Marek, J. Kozelka, *Phys. Chem. Chem. Phys.* **2016**, 18, 19472–19481.
- [22] J. Poater, X. Fradera, M. Solà, M. Duran, S. Simon, *Chem. Phys. Lett.* **2003**, 369, 248–255.
- [23] M. Noguera, M. Sodupe, J. Bertrán, *Theor. Chem. Acc.* **2007**, 118, 113–121.
- [24] H. Elgabarty, R. Z. Khaliullin, T. D. Kühne, *Nat. Commun.* **2015**, 6, 8318.
- [25] C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *Chem. Eur. J.* **1999**, 5, 3581–3594.
- [26] C. Fonseca Guerra, H. Zijlstra, G. Paragi, F. M. Bickelhaupt, *Chem. Eur. J.* **2011**, 17, 12612–12622.
- [27] F. Weinhold, R. A. Klein, *Angew. Chem. Int. Ed.* **2014**, 53, 11214–11217; *Angew. Chem.* **2014**, 126, 11396–11399.
- [28] G. Gilli, F. Bulluci, V. Ferretti, V. Bertolasi, *J. Am. Chem. Soc.* **1989**, 111, 1023–1028.
- [29] P. Sanz, O. Mó, M. Yáñez, J. Elguero, *J. Phys. Chem. A* **2007**, 111, 3585–3591.
- [30] L. Guillaumes, S. Simon, C. Fonseca Guerra, *ChemistryOpen* **2015**, 4, 318–327.
- [31] W. Hujo, S. Grimme, *J. Chem. Theory Comput.* **2013**, 9, 308–315.
- [32] C. Fonseca Guerra, F. M. Bickelhaupt, E. J. Baerends, *Cryst. Growth Des.* **2002**, 2, 239–245.
- [33] J. Šponer, J. Leszczynski, P. Hobza, *Biopolymers* **2001**, 61, 3–31.
- [34] W. L. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, 112, 2008–2010.
- [35] B. Lippert in *Progress in Inorganic Chemistry, Vol. 54* (ed.: K. D. Karlin), John Wiley & Sons, Inc., New Jersey, **2005**, pp. 385–447.
- [36] C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *J. Am. Chem. Soc.* **2000**, 122, 4117–4128.
- [37] C. Fonseca Guerra, T. van der Wijst, M. Swart, J. Poater, F. M. Bickelhaupt, *Theor. Chem. Acc.* **2010**, 125, 245–252.
- [38] F. M. Bickelhaupt, E. J. Baerends in *Reviews in Computational Chemistry, Vol. 15* (Eds.: K. B. Lipkowitz, D. B. Boyd), Wiley-VCH, New York, **2000**.
- [39] T. A. Albright, J. K. Burdett, M.-H. Whangbo, *Orbital Interactions in Chemistry, 2nd ed.*, John Wiley & Sons, Inc., New Jersey, **2013**, pp. 15–24.

 Manuscript received: April 24, 2017

Accepted manuscript online: May 9, 2017

Version of record online: June 1, 2017