Orbital Interactions in Strong and Weak Hydrogen Bonds are Essential for DNA Replication**

Célia Fonseca Guerra and F. Matthias Bickelhaupt*

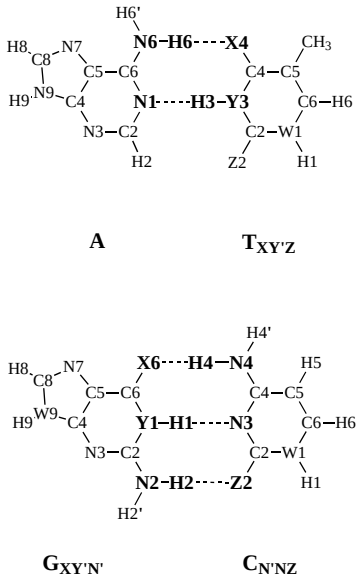
Dedicated to Professor Friedrich Bickelhaupt on the occasion of his 70th birthday

Based on a series of elegant experiments, Kool and coworkers[1] concluded that *not* Watson–Crick hydrogen bonding but steric effects, that is, the shape of DNA bases is mainly responsible for the high fidelity of DNA replication. They showed, amongst others, that 2,4-difluorotoluene (F), an isoster of thymine (T), encodes in a template strand the DNA-polymerase-catalyzed insertion of deoxyadenosine triphosphate (dATP) and that adenine (A) encodes the insertion of the deoxynucleoside triphosphate of difluorotoluene (dFTP), in spite of the supposed apolarity and the absence of hydrogen-bonding ability in F.

The above conclusion has led to a controversy in which experimental[1, 2] and theoretical[3] arguments are raised both against and in favor of the standard model[4] that reserves a key role for Watson–Crick hydrogen bonds in DNA replication. Evans and Seddon,[3a] in particular, pointed out that, according to ab initio calculations, 2,4-difluorotoluene is not apolar and thus capable of acting as a hydrogen-bond donor and acceptor. Consequently, they concluded that Kool's experiments confirm the standard model.

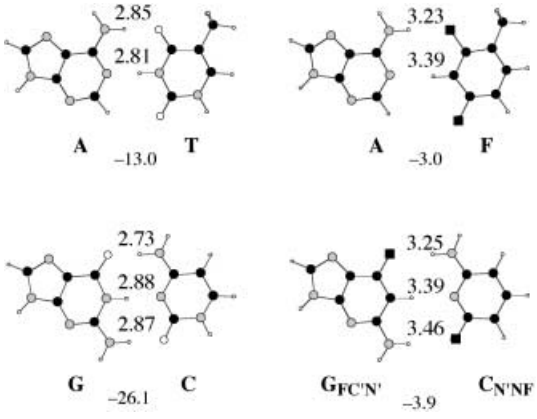
Herein, we present arguments for a third model in which steric effects and hydrogen bonding are both essential for and go hand-in-hand in high-fidelity DNA replication. An important point is that the net bond strength is insufficient for characterizing a hydrogen bond and that understanding and predicting its role and behavior requires knowledge of the different components of the bonding mechanism. In particular, it appears that covalent bonding (i.e., charge-transfer or donor–acceptor-orbital interactions) makes a significant contribution to the stability not only of the natural base pairs,[5a,b] but even to that of pairs held together by weak hydrogen bonds, such as AF.

Our evidence is based on a thorough Kohn–Sham density functional theory (DFT) investigation at BP86/TZ2P with the ADF program[6] of AT and GC (G = guanin, C = cytosin) Watson–Crick pairs and mimics thereof in which C=O and N–H bonds have been replaced by C–F and C–H (see Scheme 1). We wish to emphasize that, although seemingly a one-electron or molecular orbital (MO) model, Kohn–Sham DFT yields, in principle, exact and, in practice, with the available functionals, rather accurate energies that do take into account electron correlation and related effects, such as dispersion.[7] It is known that the BP86/TZ2P level of theory adequately describes hydrogen bonds in gas-phase and microsolvated AT and GC pairs, but also the weaker hydrogen bond between two water molecules.[5a,c]



Scheme 1. Model systems of this investigation: AT and GC Watson–Crick pairs and mimics thereof in which C=O and N–H bonds in T, G, and C have been replaced by C–F and C–H bonds (X, Z = O, F, H; W, Y = N, C).

The base-pairing energy ΔE (at 0 K) for AF (-3.0 kcal mol^{-1}) is about four-times smaller than for AT (-13.0 kcal mol^{-1}) and that the hydrogen bonds in AF are 0.4–0.6 Å longer than in AT (Scheme 2), in good agreement with other theoretical studies.[3e,g] These results raise a number of questions: 1) is F really an isoster of T if it leads to such enormous changes in the geometry relative to the Watson–Crick pair? 2) Is the A–F interaction predominantly electrostatic as one would expect for weak hydrogen bonds[8] or does it contain a sizable covalent component as has been found for the stronger hydrogen bonds in AT?[5a,b] 3) How important are hydrogen bonds for accurate DNA replication?



Scheme 2. N6–X4 and N1–Y3 distances [Å] in AT and AF and X6–N4, Y1–N3, and N2–Z2 distances in GC and $G_{FC'N'}C_{N'NF}$ obtained at the BP86/TZ2P level (see also Scheme 1).

[*] Dr. F. M. Bickelhaupt, Dr. C. Fonseca Guerra
Afdeling Theoretische Chemie
Scheikundig Laboratorium der Vrije Universiteit
De Boelelaan 1083, 1081 HV Amsterdam (The Netherlands)
Fax: (+31) 20-4447-629
E-mail: bickel@chem.vu.nl

[**] We thank the National Research School Combination—Catalysis (NRSCC) for a postdoctoral fellowship for C.F.G. and the National Computer Facilities (NCF) foundation of the Netherlands Organization for Scientific Research (NWO) for financial support.

  1433-7851/02/4112-2092 $ 20.00+.50/0

For answering the first of these questions, one must realize that steric shape or, in other words, the forces that cause steric repulsion have their origin in the quantum-mechanical effect of Pauli (and *not* electrostatic!) repulsion between the charge distributions of the two fragments.[7] Thus, making use of the conceptual framework provided by Kohn–Sham MO theory, we have decomposed the actual interaction energy (ΔE_{int}) into three physically well-defined terms: the Pauli repulsive orbital interactions between closed shells (ΔE_{Pauli}), the classical electrostatic interaction (ΔV_{elst}), and the attractive orbital interactions (ΔE_{oi}), that is, $\Delta E_{int} = \Delta E_{Pauli} + \Delta V_{elst} + \Delta E_{oi}$.[7] The results are summarized in Table 1. The orbital interactions ΔE_{oi} comprise the often distinguished attractive contributions, charge transfer, polarization (induction), and dispersion. Whether or with what precision these contributions can be quantified remains a controversial subject and we refrain from further decomposition, except by symmetry. However, we have observed that the orbital interactions in the σ-electron systems of Watson–Crick base pairs and mimics thereof are mostly of the donor–acceptor type (see below and refs. [5a,b]), and we feel it is therefore justified to denote the attractive orbital-interaction term as a covalent (or charge-transfer) contribution, as opposed to the contributions of electrostatic attraction and Pauli repulsion.

Table 1. Decomposition of base-pairing interaction energies ΔE_{int} [kcal mol^{-1}] from Kohn–Sham DFT analyses at the BP86/TZ2P level.[a]

	AF	AF*	AT	AB*
ΔE_{Pauli}	7.8	41.5	38.7	36.7
ΔV_{elst}	−7.0	−22.6	−31.8	−15.5
ΔE_{oi}	−4.0	−16.0	−22.1	−12.0
ΔE_{int}	−3.2	2.9	−15.2	9.2

[a] $\Delta E_{int} = \Delta E_{Pauli} + \Delta V_{elst} + \Delta E_{oi}$ (see text and ref. [7]). The data for AT and AF comes from equilibrium structures. In AF* and AB*, hydrogen bonds have been compressed to the equilibrium distances of AT.

The bond analyses nicely confirm that F is indeed an isoster of T. This is, however, not immediately clear if one examines only the equilibrium structure of AF. Because of the reduced bonding capabilities of F (see above), AF achieves a geometry optimum at much longer hydrogen-bond lengths than AT (Scheme 2). This situation causes all bonding components, including ΔE_{Pauli}, to be much smaller in AF: ΔE_{Pauli}, ΔV_{elst}, and ΔE_{oi} are 7.8, −7.0, and −4.0 kcal mol^{-1} in AF and 38.7, −31.8, and −22.1 kcal mol^{-1} in AT (Table 1). Kool and coworkers have provided strong evidence that the polymerase-catalyzed DNA replication involves a relatively tight active site that requires the new pair formed between template base and incoming base to conform to the Watson–Crick geometry.[1, 9] Thus, proceeding from Kool's model,[1] we have compressed AF to a geometry AF* in which the N6-H6···F4 and N1···H3-C3 hydrogen-bond distances adopt the values of the corresponding bonds in the equilibrium structure of AT, namely, 2.85 and 2.81 Å. The A–F interaction ΔE_{int} in AF* is repulsive by 2.9 kcal mol^{-1}, and ΔE_{Pauli}, ΔV_{elst}, and ΔE_{oi} are 41.5, −22.6, and −16.0 kcal mol^{-1} (Table 1). Note that the Pauli repulsion in AF* is only 2.8 kcal mol^{-1} higher than in AT and that the large difference in base-pairing energy ΔE (where $\Delta E = \Delta E_{prep} + \Delta E_{int}$; ΔE_{prep} = deformation energy) is because of a combined reduction of electrostatic and attractive orbital interactions of 15.3 kcal mol^{-1}. Thus, the forces causing the repulsive wall on the potential-energy surface are similar for AT and AF, and 2,4-difluorotoluene is a true isoster of thymine.

Next, we address the question about the nature of the A–F interaction. Unexpectedly, we find that the bonding orbital interactions in AF are of the same order of magnitude as the electrostatic attraction, that is, 36% and 64% of the total bonding interaction $\Delta E_{oi} + \Delta V_{elst}$ (Table 1). Figure 1 shows the relevant frontier-orbital interactions for AT and AF that emerge from our Kohn–Sham MO analyses. It appears that, analogous to the situation for AT,[5a,b] the ΔE_{oi} term in the A–F interaction has a significant amount of covalent character which is provided by charge transfer, that is, donor–acceptor orbital interactions in the σ-electron system between a lone pair on the fluorine atom in F and the N–H σ*-acceptor orbitals in A and between a lone pair on the nitrogen atom in A and the C–H σ*-acceptor orbitals in F. The hydrogen bonds

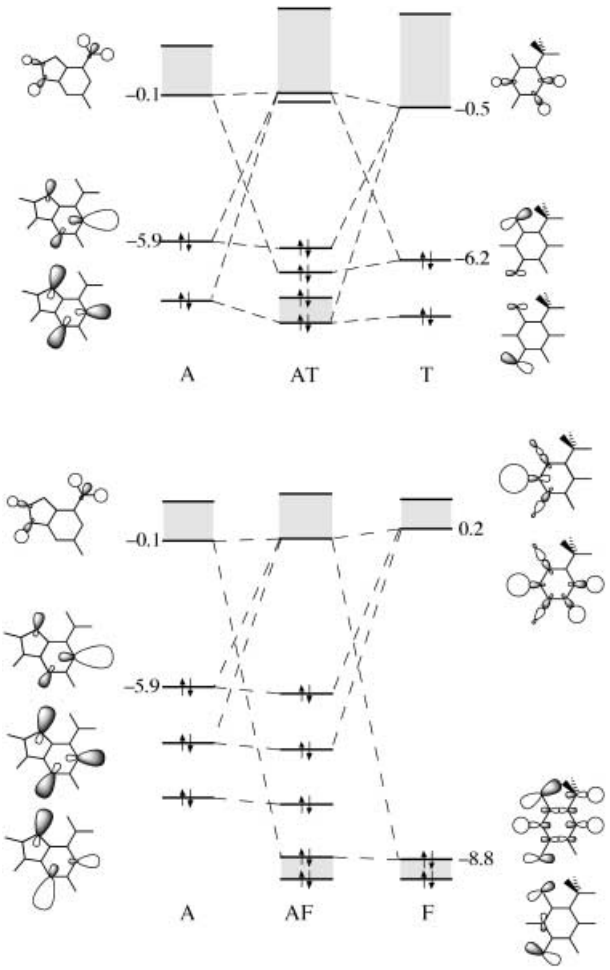


Figure 1. Frontier-orbital interactions (in the σ-electron system) between adenine and thymine in AT and between adenine and difluorotoluene in AF from Kohn–Sham DFT analyses at the BP86/TZ2P level, with HOMO and LUMO energies [eV] of the individual bases. The group of lowest unoccupied orbitals involved is represented as a gray block.

  1433-7851/02/4112-2093 $ 20.00+.50/0

in AF are weaker than in AT not only because of less electrostatic attraction ΔV_{elst} but also because of reduced orbital interactions ΔE_{oi}. The latter are weaker in AF because, compared to T, the electron-donor and -acceptor orbitals of F are lower and higher in energy, respectively, which leads to a larger orbital energy gap with the partner orbitals on A (Figure 1).

Additional support for charge transfer comes from an analysis of the redistribution of charge density caused by the interaction between two DNA bases by using the Voronoi deformation density (VDD) method.[5b, 10] In AF, the donor–acceptor interactions associated with charge transfer from A→F through the hydrogen bond N1···H-C3 are stronger than those back from F→A through the hydrogen bond N6-H···F4 leading overall to a small, negative charge of −0.004 electrons on F. All the charge transfer occurs in the σ-electron system. The π-electron systems of A and F polarize such that the accumulation of positive or negative charge around the donating or accepting atoms, caused by the charge transfer in the σ-electron system, is counteracted and partly relieved.

The base-pairing interaction in AF is not an isolated example of the importance of orbital interactions in weak hydrogen bonds, there are more cases. We have analyzed a variety of Watson–Crick base pairs, mimics, mismatches, and hydrogen-bonded molecular crystals. Figure 2, shows the

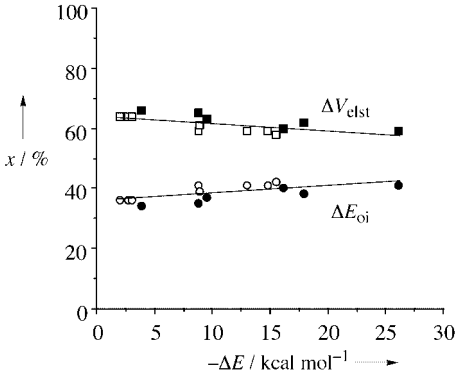


Figure 2. Relative contribution x (in % of $\Delta V_{elst} + \Delta E_{oi}$) of orbital interactions (ΔE_{oi}, circles) and electrostatic attraction (ΔV_{elst}, squares) as a function of the overall bond energy ΔE of AT (white) and GC (black) and the corresponding mimics thereof.

relative contribution of the orbital-interaction term ΔE_{oi} in the hydrogen bonds of AT, GC, and artificial mimics thereof as a function of the base-pairing energy ΔE. For the stronger hydrogen bonds ($\Delta E = -26$ kcal mol^{-1} Figure 2, right), the orbital interaction term ΔE_{oi}, which is responsible for any covalent character, contributes around 40% of the total attractive interactions ($\Delta V_{elst} + \Delta E_{oi}$). As hydrogen bonding becomes weaker (ΔE decreases from −26 to −2 kcal mol^{-1} in Figure 2), the orbital-interaction term decreases only slightly, and it is never less than one third of the total attractive interaction.

Finally, from the importance of steric factors and our analyses, it follows that hydrogen bonding also plays a key role in DNA replication. Steric factors require that the DNA bases fit into the active site of DNA polymerase, causing steric repulsion and thus a high overall activation energy for the insertion if an incoming nucleotide is too large. Base pairs of the right shape do fit into the active site without much repulsion and, therefore, with a low barrier for insertion. However, the transition state is also significantly stabilized by the Watson–Crick interaction in AT and GC, the binding energies ΔE of which are −13.0 and −26.1 kcal mol^{-1}, respectively, at 0 K (not shown in Table 1). We have shown previously[5a,c] that this intrinsic base-pairing interaction ΔE_{int} is basically affected neither by microsolvation of the major and minor grooves nor by formation of a DNA oligomer consisting of two stacked base pairs. Thus, if, for what ever reason, the bases cannot adopt the Watson–Crick geometry in the active site of the polymerase, this stabilizing interaction is lost, causing a sizable increase of the overall barrier for insertion.

But a similar argument holds true also for the formation of an AF pair: without hydrogen bonding, the barrier for this process would be restrictively high. This situation is not clear at first sight, if we just consider the *net* hydrogen-bond interaction ΔE_{int}, in AF* which is slightly repulsive (2.9 kcal mol^{-1}) and thus not stabilizing. The importance of hydrogen bonding becomes clear if one realizes that the net interaction would increase much more strongly in AF* in the absence of the underlying electrostatic attraction (−22.6 kcal mol^{-1}) and covalent orbital interactions (−16.0 kcal mol^{-1}, Table 1). Loss of the latter, for example, would raise the overall barrier for the formation of a new base pair by 16.0 kcal mol^{-1}. Note at this point that we compute the effect of the intrinsic base-pairing interaction (and its components ΔE_{oi}, ΔV_{elst}, ΔE_{Pauli}) on the barrier but not the barrier height as such. The latter depends, of course, on various other factors, such as solvent effects, stacking interactions, and the barrier associated with the nucleophilic substitution involved in the process of elongating the primer strand.[4]

In practice, it is of course difficult to switch off the orbital interactions completely, but they can be further reduced if 2,4-difluorotoluene is replaced, for example, by toluene (B). The planar, C_s-symmetric AB pair (Scheme 1, left: X, Z = H and Y, W = C) is practically unbound. To simulate again a tight DNA-polymerase active site as proposed by Kool,[1] we have compressed the AB pair to a geometry AB* in which the N6-H6···H4 and N1···H3-C3 hydrogen-bond lengths are of the same length as the corresponding bonds in the equilibrium structure of AT. The net base-pairing interaction ΔE in AB* (9.2 kcal mol^{-1}) becomes substantially more repulsive than in AF* (2.9 kcal mol^{-1}) because of a significant decrease of both the orbital interactions ΔE_{oi} and electrostatic attraction ΔV_{elst} (by a factor of 2 if compared to AT!).

Our results support aspects of both Kool's steric model of DNA replication[1] and the standard model that is based on hydrogen bonding[3a, 4] and lead to a coupling between these two models. They show that electrostatic and orbital interactions between DNA bases can also reduce the overall barrier for insertion of a nucleotide if the net hydrogen-bond strength is weak or even moderately repulsive. But our results do not rule out the possibility of this barrier being low for other reasons, for example, favorable solvent effects or

  1433-7851/02/4112-2094 $ 20.00+.50/0

stacking interactions.[11] In fact, it is plausible that one or more of these factors may be active in achieving efficient and selective replication. Full details will be published elsewhere.

Received: October 12, 2001
Revised: February 21, 2002 [Z18052]

[1] a) E. T. Kool, J. C. Morales, K. M. Guckian, *Angew. Chem.* **2000**, *112*, 1046; *Angew. Chem. Int. Ed.* **2000**, *39*, 990; b) K. M. Guckian, T. R. Krugh, E. T. Kool, *Nat. Struct. Biol.* **1998**, *5*, 954; c) J. C. Morales, E. T. Kool, *Nat. Struct. Biol.* **1998**, *5*, 950; d) K. M. Guckian, E. T. Kool, *Angew. Chem.* **1997**, *109*, 2942; *Angew. Chem. Int. Ed.* **1997**, *36*, 2825; e) S. Moran, R. X.-F. Ren, S. Rumney IV, E. T. Kool, *J. Am. Chem. Soc.* **1997**, *119*, 2056; f) B. A. Schweitzer, E. T. Kool, *J. Am. Chem. Soc.* **1995**, *117*, 1863; see also: g) U. Diederichsen, *Angew. Chem.* **1998**, *110*, 1745; *Angew. Chem. Int. Ed.* **1998**, *37*, 1655.

[2] K. S. Schmidt, R. K. O. Sigel, D. V. Filippov, G. A. van der Marel, B. Lippert, J. Reedijk, *New J. Chem.* **2000**, *24*, 195.

[3] a) T. A. Evans, K. R. Seddon, *Chem. Commun.* **1997**, 2023; b) X. Wang, K. N. Houk, *Chem. Commun.* **1999**, 2631; c) D. Barsky, E. T. Kool, M. E. Colvin, *J. Biomol. Struct. Dyn.* **1999**, *16*, 1119; d) B. D. Silverman, M. C. Pitman, D. E. Platt, *J. Biomol. Struct. Dyn.* **1999**, *16*, 1169; e) F. Ryjácek, M. Kratochvíl, P. Hobza, *Chem. Phys. Lett.* **1999**, *313*, 393; f) C. Santhosh, P. C. Mishra, *Int. J. Quantum Chem.* **1998**, *68*, 351; g) M. Meyer, J. Sühnel, *J. Biomol. Struct. Dyn.* **1997**, *15*, 619.

[4] L. Stryer, *Biochemistry*, W. H. Freeman and Company, New York, **1988**.

[5] a) C. Fonseca Guerra, F. M. Bickelhaupt, *Angew. Chem.* **1999**, *111*, 3120; *Angew. Chem. Int. Ed.* **1999**, *38*, 2942; b) C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *Chem. Eur. J.* **1999**, *5*, 3581; c) C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *J. Am. Chem. Soc.* **2000**, *122*, 4117.

[6] a) G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. Fonseca Guerra, E. J. Baerends, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931; b) C. Fonseca Guerra, J. G. Snijders, G. te Velde, E. J. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391.

[7] 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**, pp. 1–86.

[8] a) G. R. Desiraju, T. Steiner, *The Weak Hydrogen Bond*, Oxford University Press, Oxford, **1999**; b) G. A. Jeffrey, *An Introduction to Hydrogen bonding*, Oxford University Press, New York, Oxford, **1997**; c) see also T. Steiner, *Angew. Chem.* **2002**, *114*, 50–80; *Angew. Chem. Int. Ed.* **2002**, *41*, 48–76.

[9] The steric demands of the DNA-polymerase active site are less stringent for the sugar moieties, as shown by D. Summerer, A. Marx, *Angew. Chem.* **2001**, *113*, 3806; *Angew. Chem. Int. Ed.* **2001**, *40*, 3693.

[10] F. M. Bickelhaupt, N. J. R. van Eikema Hommes, C. Fonseca Guerra, E. J. Baerends, *Organometallics* **1996**, *15*, 2923.

[11] Experiments suggest indeed that hydrophobicity, for example, may also be a sufficient driving force for selective replication: a) D. L. McMinn, A. K. Ogawa, Y. Wu, J. Liu, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **1999**, *121*, 11585; b) M. Berger, A. K. Ogawa, D. L. McMinn, Y. Wu, P. G. Schultz, F. E. Romesberg, *Angew. Chem.* **2000**, *112*, 3069; *Angew. Chem. Int. Ed.* **2000**, *39*, 2940.

$[\{Li(thf)_3\}_2Ga_2\{As(Si\mathit{i}Pr_3)\}_4]$—A Compound with Gallium–Arsenic Double Bonds**

Carsten von Hänisch* and Oliver Hampe

The synthesis of molecular compounds with multiple bonds between the heavier main-group elements is still a preparative challenge for the inorganic chemist. While various symmetric compounds with multiple bonds between two heavier atoms of Group 13 or 15 were synthesized in the last years,[1–3] most of the asymmetric compounds with multiple bonds between Group 13 and 15 elements contain one atom of the 2nd period of the periodic table.[3] Power and co-workers recently published compounds with aluminum– or gallium–nitrogen multiple bonds,[4] and compounds with boron–phosphorus and boron–arsenic double bonds $[Mes_2B{=}ER\{Li(thf)_3\}]$ (**1a**: E = P, R = $SiMe_3$; **1b**: E = As, R = Ph; Mes = mesityl) have also been described.[5] Partial multiple-bond character between the heavier elements of Group 13 and 15, however, can be observed in the ternary phases $Cs_6M_2E_4$ (M = Al, Ga; E = P, As) or in the molecular compounds $[t\mathrm{Bu}_2\mathrm{GaPMes}^*(\mathrm{SiPh}_3)]$ and $[\mathrm{MesP}\{\mathrm{Ga(Trip)}_2\}_2]$ (Trip = 2,4,6-$i\mathrm{Pr}_3\mathrm{C}_6\mathrm{H}_2$; Mes* = 2,4,6-$t\mathrm{Bu}_3\mathrm{C}_6\mathrm{H}_2$).[6, 7] In contrast, multiple bonding between the heavier Group 13 and 15 elements can be excluded in the cyclic compounds $[RMER']_n$ (M = Al, Ga; E = P, As; R, R′ = alkyl, aryl; $n = 2$, 3), because of the ring folding and the pyramidal coordination of the phosphorus and arsenic atoms.[8] Molecular compounds with double bonds between gallium and arsenic are unknown to date.

We could synthesize the compound $[\{Li(thf)_3\}_2Ga_2\{As(Si\mathit{i}Pr_3)\}_4]$ (**2**) containing two As–Ga double bonds from the reaction of $GaCl_3$ with $Li_2AsSi\mathit{i}Pr_3$ [Eq. (1)].

$$2\,GaCl_3 + 4\,Li_2AsSi\mathit{i}Pr_3 \xrightarrow{\text{THF/heptane}} [\{Li(thf)_3\}_2Ga_2\{As(Si\mathit{i}Pr_3)\}_4]\ (\mathbf{2}) + 6\,LiCl \quad (1)$$

Compound **2** crystallizes in the monoclinic space group $P2_1/n$ with two formula units per unit cell (Figure 1).[9] The central structural motive of **2** is a planar Ga_2As_2 four-membered ring with two further exocyclic As atoms bound to the Ga atoms. Each As atom is bonded to a tri(isopropyl)silyl group and each of the exocyclic As atoms is additionally bound to one $Li(thf)_3$ fragment.

The angular sum at the Ga atoms comes to 359.8(1)° thus, the coordination of these atoms is essentially trigonally planar. However, the As-Ga-As angles differ, the endocyclic As(1)-Ga-As(1′) angle is 98.36(3)°, and the As(1)-Ga-As(2) and As(1′)-Ga-As(2) angles are 121.51(4) and 139.90(4)°, respectively. The considerable difference between the exocyclic angles can be explained by the mutual repulsion of the neighboring $i\mathrm{Pr}_3\mathrm{Si}$ groups on As(1′) and As(2). The Li–As

[*] Dr. C. von Hänisch, Dr. O. Hampe
Institut für Nanotechnologie
Forschungszentrum Karlsruhe
Postfach 3640, 76021 Karlsruhe (Germany)
Fax: (+49) 7247-82-6368
E-mail: carsten.vonhaenisch@int.fzk.de

[**] This work was supported by the Strategiefond of the Helmholtz Foundation. We thank Dr. P. Kramkowski for his valuable help with the translation.