

ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

COMMUNICATION

Guanosine-based hydrogen-bonded 2D Scaffolds: metal-free formation of G-quartets and G-ribbons architectures at the solid/liquid interface

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

We report on the synthesis and self-assembly of three novel lipophilic guanosine derivatives exposing a ferrocene moiety on the C(5') position of the sugar unit. Their self-association in solution, and at the solid/liquid interface, can be tuned by varying the size and nature of the C(8)-substituent, leading to the generation of either G-ribbons, lamellar G-dimer based arrays or the G₄ cation-free architectures.

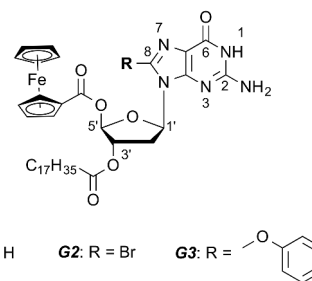
The controlled self-assembly of suitably designed molecular building blocks is a viable approach towards the construction of highly sophisticated nanostructured materials.¹ Among various molecular components, supramolecular architectures with *ad hoc* structural motifs can be obtained through the non-covalent self-association of natural² and unnatural³ nucleobases on flat surfaces. Such structures, when decorated with appropriate electrically/optically active units, can be used as scaffolds to locate such units in pre-determined positions in 2D,⁴ thereby paving the way towards a wide range of applications, e.g. in opto-electronics.⁵ Among the four nucleobases of DNA, guanine (G)⁶ exhibits a very rich self-assembly behaviour: depending on the environmental conditions it can undergo different self-assembly pathways resulting in various well-distinct architectures including dimers,⁷ tetramers,⁸ ribbons,⁹ and helical structures.¹⁰ In the presence of certain metal ions, G can form cyclic tetrameric architectures, known also as G-quartets (hereafter G₄), which further pile up into octamers or higher order columnar aggregates. It is commonly believed that templating alkali metal cations such as Na⁺ and K⁺ as well as alkaline earth and lanthanide cations are needed to stabilize the G₄ formation.¹⁰ However, suitably designed guanosines, e.g. derivatives exposing a sterically demanding N,N-dimethylaniline moiety in C(8) position of the guanine core, were found to form cation-free G₄ structures both in solution and in the solid state of the bulk.¹¹

On solid surfaces, G-based H-bonded supramolecular architectures were self-assembled into highly ordered motifs and studied with scanning tunnelling microscopy (STM) under ultra-high vacuum.¹²

However, the STM explorations at the solid/liquid interfaces have shown numerous advantages, e.g. they provide an excellent environment for *in situ* chemical modifications of adsorbed molecules.¹³ When guanosine derivatives are physisorbed at the solution/graphite interface, thermodynamically stable supramolecular ribbons, characterized by N(2)-H⁺O(6) and N(1)-H⁺N(7) H-bonds, were observed.

Given the possibility to functionalize the guanosines in the sugar moiety, they represent ideal building blocks for the fabrication of conformationally rigid and structurally complex architectures based on ribbons or G₄ motifs. Yet, the formation of G₄ at the solid/liquid interface was observed only upon use of a templating metal center.^{2c} Ferrocenes are organometallic compounds possessing unique opto-electronic properties, which made them important active components for applications in medicine and materials science.¹⁷ In this context, the control over the self-assembly of ferrocene-based architectures through molecular engineering is crucial in order to control and improve their opto-electronic properties.

Here we have designed and synthesized three novel lipophilic guanosine derivatives **G1-G3** (see Scheme 1, and ESI for synthetic details), exposing a ferrocene moiety on the C(5') position of the sugar unit.



Scheme 1. Chemical formulae of investigated guanosine derivatives.

In order to tune the molecular self-assembly process at the graphite-solution interface we substituted the nucleobase C(8) position with different sterically demanding groups. The presence of a long stearate side chain in C(3') position of the sugar is expected to promote the molecular physisorption on HOPG.

In line with previous studies on other guanosines,^{2a, 2b} in the absence of metal ions, **G1** in solution forms an H-bonded ribbon-like structure that involves the pairing N(2)-H \cdots O(6) and N(1)-H \cdots N(7). ¹H NMR spectra in CDCl₃ (Fig. S1 in ESI) show a progressive downfield shift for both N(1)-H and N(2)-H signals upon cooling, while considerable line broadening occurs (e.g. see the C(8)-H signal at $\delta = 7.8$ in Fig. S1). **G1** can complex alkali metal ions to form a C₄ symmetric octamer consisting of two stacked **G1₄**, as evidenced from the characteristic changes both in the ¹H NMR and in the CD spectrum (Fig. S2).¹⁹ The self-assembly of **G1** at the solid-liquid interface has been explored by applying a 4 μ L drop of a (100 \pm 2) μ M **G1** solution in 1-phenyloctane on HOPG. The STM image showed a crystalline structure consisting of ribbon-like architectures forming a lamellar motif (Fig. 1a). In this 2D crystal, the stearate side chains are physisorbed flat on the surface and they are interdigitated between adjacent lamellae. The unit cell parameters amount to $a = (7.4 \pm 0.1)$ nm, $b = (1.0 \pm 0.1)$ nm, and $\alpha = (88 \pm 2)^\circ$, leading to an area $A = (7.4 \pm 0.2)$ nm², where each unit cell contains four molecules. Thus, the area occupied by a single molecule **G1** corresponds to (1.85 \pm 0.10) nm². Given the size of the unit cell there is not enough space to accommodate the ferrocene units on the basal plane of the HOPG surface, thus it is most likely that they are either back-folded into supernatant solution or physisorbed as second layer on the top of the guanine first layer. Unfortunately, despite the high spatial resolution attained by STM imaging, we are unable to rule out any of these two scenarios. The monitored supramolecular motif can be well-described by the formation of a 1D hydrogen-bonded ribbon that involves the pairing N(2)-H \cdots O(6) and N(1)-H \cdots N(7) (see model in Fig. 1b and S12). This self-assembly behaviour is in good agreement with NMR solution data.

In order to steer **G** self-assembly towards different supramolecular motifs, we explored the effect of the functionalization of the C(8) position of the guanine core, by substituting the proton with a Br atom (**G2**). Monolayers of **G2** have been generated by applying on the HOPG surface a 4 μ L drop of a (100 \pm 2) μ M solution of **G2** in 1-phenyloctane. The STM imaging (Fig. 1c) displays a crystalline structure consisting of lamellar architectures. In **G2**-based 2D crystal, the stearate side chains are physisorbed flat on the surface and are interdigitated between adjacent lamellae. The unit cell parameters, $a = (4.1 \pm 0.1)$ nm, $b = (0.9 \pm 0.1)$ nm, and $\alpha = (90 \pm 2)^\circ$, lead to an area $A = (3.7 \pm 0.1)$ nm², where each unit cell contains two molecules. Thus, the area occupied by a single molecule **G2** amounts to (1.85 \pm 0.10) nm². While the area occupied by single molecule **G2** is identical with the one of **G1**, their self-assembled patterns are markedly different (see Figures 1a vs. 1c). In particular, the appearance of hollow features within **G2** ribbon core (indicated with arrows in the inset of Fig. 1c) as well as different orientation of stearate side chains vs. the main lamellar axis (60° and 90° for **G1** and **G2** patterns, respectively) provides unambiguous evidence for a different self-assembly motif. In fact, the **G2** supramolecular motif can be well-described by the formation of H-bonded dimers that involves the pairing N(1)-H \cdots O(6) (see models in Figs. 1d and 2). Each dimer interacts laterally with

neighbouring dimers via N(2)-H \cdots Br(8) bonding, resulting in the formation of 1D lamellar arrays. Similarly to the case of **G1** ribbons, ferrocene units are likely back-folded into supernatant solution or adsorbed as a second layer. Formation of such structures highlights the role played by bulky bromine atoms in C(8) position of **G** core, which introduced N(2)-H \cdots Br(8) bonding, thus forcing self-assembly towards an unprecedented ribbon structure.

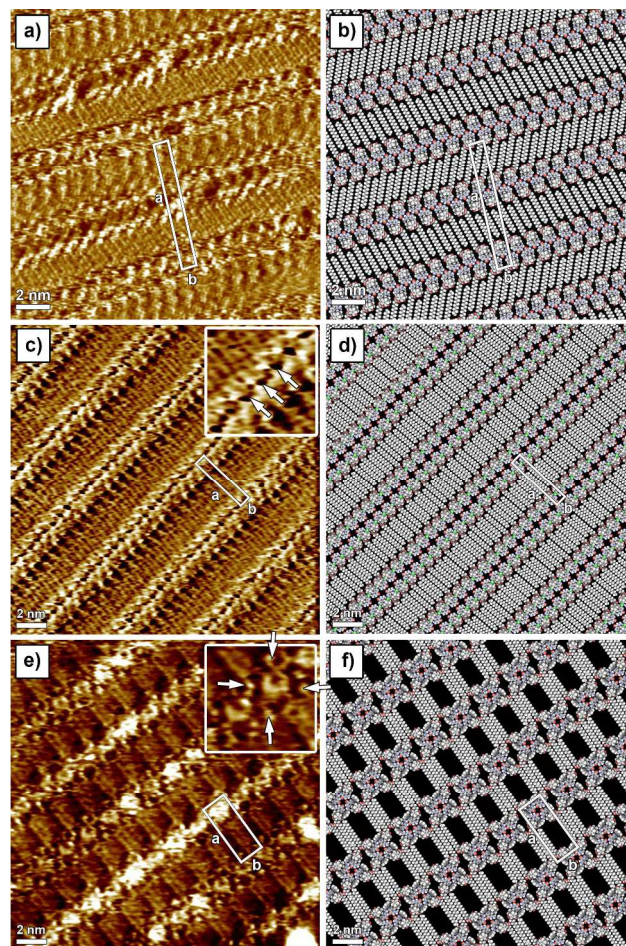


Figure 1. STM images of physisorbed monolayers of the investigated guanosine derivatives self-assembled on HOPG from solutions in 1-phenyloctane. Ribbon-like structure of (a) **G1**, and (c) **G2**. (e) **G3**-based structure of **G3**. The packing motifs are shown in (b), (d) and (f), respectively. Tunneling parameters in (a), (c) and (e): average tunnelling current (I_t)=(35 \pm 2) pA, bias voltage (V_t)=(400 \pm 25) mV.

¹H-NMR spectra of **G2** recorded upon cooling a solution in CDCl₃ (Fig. S3) show a progressive splitting of the broad N(2)-H singlet at $\delta = 6.1$ into two signals (bonded and free N(2)-H, at $\delta = 8.7$ and $\delta = 5.7$, respectively). The chemical shifts for the N(2)-H protons are close to those reported for a similar compound ($\delta = 8.50$ and $\delta = 5.44$),^{10b} but differ from those of the two stacked **G4** formed by **G1** in the presence of metal ions (marked signals in Fig. S1) as well as from those of an isolated **G4** ($\delta = 9.81$ and $\delta = 5.15$).¹¹ Previously,^{10b} some of us have studied the self-assembly of a similar 8-bromo lipophilic guanosine derivative in solution by NMR spectroscopy. It was concluded that this signal splitting can be attributed to the existence of isolated **G4** on the basis of the well-known preference of 8-bromo guanosine to adopt a *syn* conformation around the glycosidic bond and the lack of any

liquid crystalline behavior. Although **G2** behaves similarly and the lack of substantial line broadening points to the existence of small size aggregates, the present results suggest to reconsider the supramolecular behavior of **G2** in solution, taking into account the existence of small, possibly dimer-like, aggregates.

We then decided to replace the Br atom with a more neutral group, which is also more sterically demanding, i.e. phenol (**G3**). The behaviour of **G3** in solution is very peculiar. In analogy to **G2**, the **G3** molecule is unable to complex metal ions to form G_4 stacked structures, as no changes can be detected both on CD (Fig. S4, S8) and on ^1H NMR spectra after the addition of K^+ . Furthermore, in the absence of added ions, both N(1)-H and N(2)-H signals split upon cooling. In particular, the N(2)-H signal splits into two couples of new signals in a 2:1 ratio (Fig. S5). A couple of signals at ca. $\delta = 8$, can be attributed to H-bonded N(2)-Hs, while the other couple of signals appearing at ca. $\delta = 3$ ppm can be ascribed to free N(2)-Hs. The existence of two sets of resonances for both imino and amino protons in a 2:1 ratio points to the existence of two different supramolecular species. On the basis of NOE experiments (Fig. S6, S7), the major specie can be ascribed to the formation of all-*syn* isolated **G3**. Although no direct and conclusive evidence could be gathered from the spectra, in depth inspection of the models suggests *anti* **G3**, or *anti* **G3**-dimers, analogous to those found for **G2** on surfaces, as the possible structure for the minor specie. STM investigation of sub-monolayer-thick films obtained from (100 ± 2) μM solution of **G3**, revealed the formation of a new type of pattern. In this 2D crystal, because of steric hindrance brought into play by the phenol unit, only three out of four stearate side chains are physisorbed flat on the surface. The unit cell parameters amount to $a = (4.5 \pm 0.1)$ nm, $b = (1.8 \pm 0.1)$ nm, and $\alpha = (90 \pm 2)^\circ$, leading to an area $A = (8.1 \pm 0.1)$ nm 2 , where each unit cell contains two molecules. Thus, the area occupied by a single molecule **G3** corresponds to (4.1 ± 0.1) nm 2 . The packing of **G3** molecules is very loose as evidenced by the large discrepancy between the areas occupied by single molecules **G1**, **G2** and **G3**. The STM inset in Figure 1e clearly shows the presence of macrocyclic bright features decorated with four small protrusions (indicated with white arrows in the inset), which can be assigned to **G3** and ferrocene groups (backfolded into the supernatant solution), respectively. Because of the presence of sterically demanding phenol substituents in the C(8) position of **G3**, the formation of H-bonded ribbon-like structures is hindered,³¹ leading to the generation of cyclic tetrameric H-bonded structures characterized by the N(1)-H \cdots O(6) and N(2)-H \cdots N(7) motif, whose existence was also indicated by NOE data in solution. While NOE analysis suggests the presence of all-*syn* isolated **G3**, as the main specie, once adsorbed on the surface both all-*syn* and all-*anti* **G3**, will occupy the same areas, therefore we cannot unambiguously exclude the existence of the former over the latter. Noteworthy, some of **G3**, appear brighter in the STM image, which can be explained by the interference of the supramolecular lattice and the underlying HOPG surface.

To provide a molecular understanding of the three **G** derivatives self-assembly in 2D and gain insight into the formation and stability of supramolecular structures, we have performed density functional theory (DFT) calculations using the hybrid Gaussian and plane-wave method (GPW), implemented in the QUICKSTEP module of the CP2K package.³⁴ We used the B₃LYP hybrid exchange-correlation

potential,³⁵ whereas the Grimme's DFT-D2 method³⁶ was employed for taking into account the dispersion forces. The additional details on the computational methodology, as well as of the results for the structural and electronic properties of the different assembly motifs are detailed in ESI. To bestow information onto the intermolecular binding mechanisms, we have focused our attention on unravelling the interplay between H-bonds, holding together the guanine cores, and the effective metallic repulsion coming from the four iron cations present in the ferrocenes.

Table 1. Calculated bonding energies [a]

System	E_{tot}	$E_{\text{H-bond}}$	$E_{\text{Br-HN}}$
G1	-52.1	-13.0	-
G2	-50.8	-11.9	-13.9
G3	-40.1	-10.0	-

[a] E_{tot} total intermolecular interaction energy, $E_{\text{H-bond}}$ average hydrogen-bond energy, $E_{\text{Br-HN}}$ energy of bromine-NH interactions, all energies in kcal mol $^{-1}$

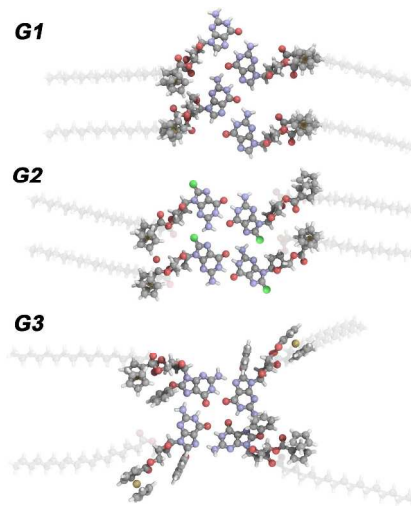


Figure 2. Calculated structure of **G1** H-bonded ribbon, **G2** lamella and **G3** cation-free quartet structure.

Noteworthy, as can be seen in the suggested monolayer packing motifs, two types of intramolecular interactions can be distinguished, i.e., the hydrogen-bonding (or N(2)-H \cdots Br(8) interactions in the case of the **G2** structure) between **G** cores, and the van der Waals interaction, resulting from the interdigitation of the stearate chains. In order to determine their contribution in the total cohesive energy, we calculated the intermolecular dissociation energy, for each of the different four-molecule-based configurations (see Fig. 2) exhibited in three **G**-based complexes and the results are reported in Table 1. The presented first-principle calculations have not only elucidated the different mechanically stable molecular arrangements, but more importantly, have shed light on the energetics of intra-molecular (within unit cells) as well as inter-molecular (between neighbouring cells) interactions determining the stability of the molecular networks. According to E_{tot} values, the ribbon structure of **G1** is greatly stabilized by four strong H-bonds. In the **G2** lamellar structure molecules are

held together by two H-bonds to form dimers, which further self-assemble via two strong N(2)-H \cdots Br bonds to form 1D array. As expected the **G**₃ macrocycle is energetically unfavored, since the H-bonds involved in pairing are of the weakest nature. Our findings indicate that the formation of intermolecular H-bonds guides the self-assembly, since the interactions between the stearate chains are much weaker. In gas phase, the calculated electronic structure of the dimers and ribbons exhibits hybridization between the states stemming from the organic complex, namely the guanine backbone, and the metallic states associated to the ferrocene groups. The information given by the electronic structure of calculated complexes confirms that they are primarily held together by H-bonds even in presence of the metallic repulsion coming from the occupied molecular orbitals with clear *d* symmetries.

In summary, we have designed and synthesised novel organic soluble ferrocene-exposing guanines. Their self-association in solutions, occurring via H-bonding, depends on the steric hindrance and H-bonding ability of the substituent attached to the nucleobase C(8)-position. When physisorbed at the solid/liquid interface the diversity of self-assembly behaviour upon chemical design is reflected in the generation of either different *G*-ribbon structures or the *G*₄ cation-free architectures. These structures have been monitored on the sub-nm scale by *in situ* STM imaging. Our approach demonstrates that a careful molecular design of the guanosine starting building block makes it possible to steer the self-assembly towards the formation of different supramolecular architectures, even in the absence of templating ions. Such motifs are different, genuine supramolecular 2D scaffolds dictating, in the present case, the spatial localization of ferrocenes, ultimately forming 1D arrays that may be of interest in opto-electronics.

This work was supported by the European Community through the project EC FP7 ICT-MOLARNET (318516) and the European Research Council project SUPRAFUNCTION (GA-257305), the Agence Nationale de la Recherche through the LabEx project Chemistry of Complex Systems (ANR-10-LABX-0026_CSC) and the International Center for Frontier Research in Chemistry (icFRC).

Notes and references

^a *ISIS & icFRC, Université de Strasbourg & CNRS, 8 allée Gaspard Monge, 67000 Strasbourg, France. E-mail: samori@unistra.fr, ciesielski@unistra.fr*

^b *Alma Mater Studiorum - Università di Bologna, Dipartimento di Chimica "G. Ciamician", Via S. Giacomo 11, 40126 Bologna, Italy. E-mail: stefano.masiero@unibo.it*

^c *Institute for Materials Science and Max Bergmann Center of Biomaterials, Dresden University of Technology, 01062 Dresden, Germany.*

^d *Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany.*

^e *Center for Advancing Electronics Dresden, Dresden Center for Computational Materials Science, Dresden University of Technology, 01062 Dresden, Germany. E-mail: g.cuniberti@tu-dresden.de*

|| *These authors contributed equally to the work.*

† *Electronic Supplementary Information (ESI) available: experimental details, synthesis, full characterisation of new compounds, NMR spectra, DFT.*

1. a) J. S. Lindsey, *New J. Chem.*, 1991, **15**, 153-180; b) J.-M. Lehn, *Supramolecular chemistry: concepts and perspectives*, VCH New

- York, 1995; c) G. M. Whitesides, E. E. Simanek, J. P. Mathias, C. T. Seto, D. N. Chin, M. Mammen and D. M. Gordon, *Acc. Chem. Res.*, 1995, **28**, 37-44; d) D. Philp and J. F. Stoddart, *Angew. Chem. Int. Ed.*, 1996, **35**, 1155-1196; e) J. H. van Esch and B. L. Feringa, *Angew. Chem. Int. Ed.*, 2000, **39**, 2263-2266; f) D. N. Reinhoudt and M. Crego-Calama, *Science*, 2002, **295**, 2403-2407; g) J. F. Stoddart and H. R. Tseng, *Proc. Nat. Acad. Sci. U.S.A.*, 2002, **99**, 4797-4800.
2. a) G. Gottarelli, S. Masiero, E. Mezzina, S. Pieraccini, J. P. Rabe, P. Samori and G. P. Spada, *Chem. Eur. J.*, 2000, **6**, 3242-3248; b) S. Lena, G. Brancolini, G. Gottarelli, P. Mariani, S. Masiero, A. Venturini, V. Palermo, O. Pandoli, S. Pieraccini, P. Samori and G. P. Spada, *Chem. Eur. J.*, 2007, **13**, 3757-3764; c) A. Ciesielski, S. Lena, S. Masiero, G. P. Spada and P. Samori, *Angew. Chem. Int. Ed.*, 2010, **49**, 1963-1966; d) A. Ciesielski, R. Perone, S. Pieraccini, G. P. Spada and P. Samori, *Chem. Commun.*, 2010, **46**, 4493-4495.
3. a) M. Yu, J. G. Wang, M. Mura, Q. Q. Meng, W. Xu, H. Gersen, E. Laegsgaard, I. Stensgaard, R. E. A. Kelly, J. Kjems, T. R. Linderoth, L. N. Kantorovich and F. Besenbacher, *ACS Nano*, 2011, **5**, 6651-6660; b) A. Ciesielski, S. Haar, A. Benyei, G. Paragi, C. F. Guerra, F. M. Bickelhaupt, S. Masiero, J. Szolomajer, P. Samori, G. P. Spada and L. Kovacs, *Langmuir*, 2013, **29**, 7283-7290; c) A. Ciesielski, S. Haar, G. Paragi, Z. Kupihar, Z. Kele, S. Masiero, C. F. Guerra, F. M. Bickelhaupt, G. P. Spada, L. Kovacs and P. Samori, *Phys. Chem. Chem. Phys.*, 2013, **15**, 12442-12446.
4. a) G. P. Spada, S. Lena, S. Masiero, S. Pieraccini, M. Surin and P. Samori, *Adv. Mater.*, 2008, **20**, 2433-2438; b) T. F. A. Greef and E. W. Meijer, *Nature*, 2008, **453**, 171-173.
5. Y. L. Wu, K. E. Brown and M. R. Wasielewski, *J. Am. Chem. Soc.*, 2013, **135**, 13322-13325.
6. a) J. T. Davis, *Angew. Chem. Int. Ed.*, 2004, **43**, 668-698; b) J. T. Davis and G. P. Spada, *Chem. Soc. Rev.*, 2007, **36**, 296-313.
7. J. L. Sessler and R. Z. Wang, *Angew. Chem. Int. Ed.*, 1998, **37**, 1726-1729.
8. V. Andrisano, G. Gottarelli, S. Masiero, E. H. Heijne, S. Pieraccini and G. P. Spada, *Angew. Chem. Int. Ed.*, 1999, **38**, 2386-2388.
9. G. Gottarelli, S. Masiero, E. Mezzina, G. P. Spada, P. Mariani and M. Recanatini, *Helv. Chim. Acta*, 1998, **81**, 2078-2092.
10. a) J. A. Walmsley and J. F. Burnett, *Biochemistry*, 1999, **38**, 14063-14068; b) T. Giorgi, S. Lena, P. Mariani, M. A. Cremonini, S. Masiero, S. Pieraccini, J. P. Rabe, P. Samori, G. P. Spada and G. Gottarelli, *J. Am. Chem. Soc.*, 2003, **125**, 14741-14749.
11. J. L. Sessler, M. Sathiosatham, K. Doerr, V. Lynch and K. A. Abboud, *Angew. Chem. Int. Ed.*, 2000, **39**, 1300-1303.
12. a) R. Otero, M. Schock, L. M. Molina, E. Laegsgaard, I. Stensgaard, B. Hammer and F. Besenbacher, *Angew. Chem. Int. Ed.*, 2005, **44**, 2270-2275; b) R. Otero, W. Xu, M. Lukas, R. E. A. Kelly, E. Laegsgaard, I. Stensgaard, J. Kjems, L. N. Kantorovich and F. Besenbacher, *Angew. Chem. Int. Ed.*, 2008, **47**, 9673-9676.
13. A. Ciesielski and P. Samori, *Nanoscale*, 2011, **3**, 1397-1410.
14. a) M. Krack and M. Parrinello, *High Performance Computing in Chemistry*, 2004, **25**, 29-51; b) J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comput. Phys. Commun.*, 2005, **167**, 103-128.
15. a) A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372-1377; b) A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098-3100.
16. S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787-1799.