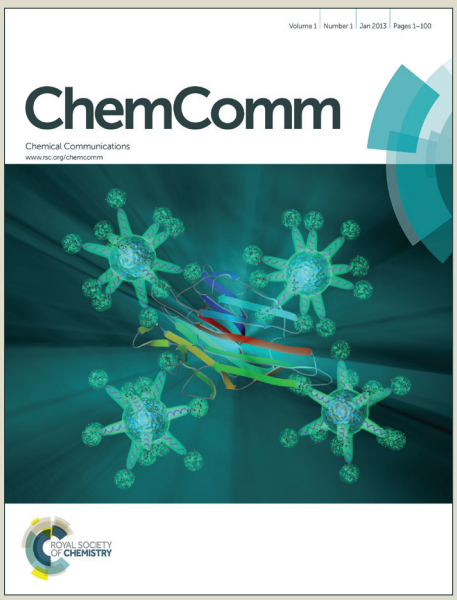
ChemComm

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.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

RSCPublishing

COMMUNICATION

A lipophilic "fully-*anti*" dodecamer from a (5'*S*)-5',8-cyclo-2'-deoxyguanosine†

Cite this: DOI: 10.1039/xoxxoooooox

Silvia Pieraccini,[a] Michael A. Terzidis,[b] Enrico J. Baldassarri,[c] Giovanna Fragneto,[d] Paolo Mariani,[c] Stefano Masiero,*[a] and Chryssostomos Chatgilialoglu*[b]

Received ooth January 2012,
Accepted ooth January 2012

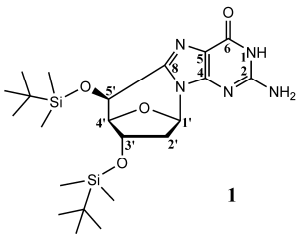
DOI: 10.1039/xoxxoooooox

www.rsc.org/

The self-assembly of a lipophilic derivative of (5'*S*)-5',8-cyclo-2'-deoxyguanosine, a mutagenic product formed by hydroxyl radical attack against DNA, has been investigated. This derivative forms, with high fidelity, a dodecameric complex composed of three stacked G-quartets in the presence of strontium picrate. This is the first example of a fully-*anti* lipophilic G-quadruplex.

Guanine (G)-rich sequences of DNA and RNA can spontaneously fold into G-quadruplexes (GQs).[1-3] Due to their potential role in multiple biological processes and genetic instabilities, GQs represent attractive targets for drug design, and much research is focused on their structural analysis and interacting properties.[4-6] 5',8-cyclo-2'-deoxyguanosines (cdGs), containing a covalent bond between the deoxyribose C5' and the purine C8 carbons, are tandem DNA lesions formed after the C5' methylene group insult by the hydroxyl radical.[7] Lesion cdGs have been detected both *in vitro* and *in vivo*.[7,8] (5'*S*)-cdG has been recently reported to be a strong block for replication and highly mutagenic in *Escherichia coli*, although it participates in Watson-Crick H-bonding in DNA.[9,10] Significant helicoidal and base stacking perturbations into the duplex have been evidenced.[11] In spite of the biological relevance of this modified nucleoside, the propensity of (5'*S*)-cdG to give G-quartet (G4) based assemblies has not been investigated so far, and could represent a point of interest for both biology and supramolecular chemistry. In fact, G4 based helical superstructures reminiscent of DNA-GQs can be obtained in non-aqueous media by lipophilic guanosine derivatives upon addition of ions.[12] These non-covalent assemblies, besides mimicking morphologies of biological GQs, can be exploited as versatile scaffolds for developing functional nanomaterials and devices.[13]

Recently, a short synthetic protocol for the synthesis of the (5'*S*)-cdG lipophilic analogue **1** (figure 1) has been reported.[14] This derivative, an 8-substituted guanosine, possesses a covalently blocked *anti* conformation around the glycosydic bond as its relevant structural feature. Here we describe the self-assembly behavior of **1** in chloroform in the presence of Sr^{2+} as a templating cation.[15,16]



1

Figure 1

The Circular Dichroism (CD) spectrum of a $CDCl_3$ solution of **1** shows a negative monosignate band in the 250 ÷ 300 nm region, corresponding to the π-π* transition of the G chromophore, in accordance with the presence of monomeric or oligomeric ribbon-like species (Figure 2a).[13,17] Upon addition of an excess of solid

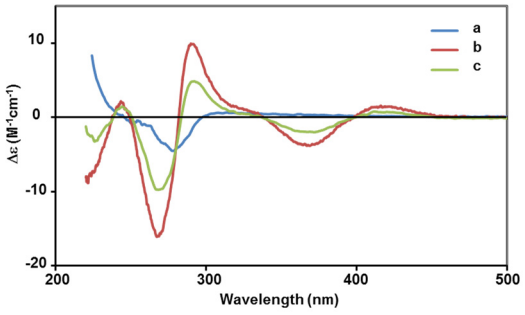


Figure 2. CD spectra recorded in $CDCl_3$ at 25°C. **a**, **1** (8 mM). **b**, **1**-SrP_2 solution obtained after extraction of 1/6 eq. of strontium picrate. **c**, **1**-SrP_2 solution added with ½ eq. of [2.2.2] relative to Sr^{2+}.

ChemComm Accepted Manuscript

strontium picrate (SrP_2), 1/6 eq. of the salt is extracted into the organic solution, as shown from ^{1}H-NMR analysis (*vide infra*).

The CD spectrum of the resulting solution shows drastic changes (Figure 2b). A positive CD band appears at 290 nm, accompanied by a negative non-conservative exciton coupling centered at 250 nm (λ_{max} **1** = 259 nm, see Figure S1). The latter spectral feature is diagnostic of a self-assembled species formed by the chiral stacking of at least two G_4s.[18] In addition, two opposite signed CD bands, not of the exciton type, are observed at 420 and 365 nm, corresponding to the two absorption maxima of the picrate anion. The emergence of this induced optical activity indicates that the intrinsically achiral counterion tightly interacts with the cation-templated chiral superstructure.

Variable temperature experiments show that CD band intensities decrease upon heating, possibly due to disaggregation. By increasing the temperature from 5 to 25°C only a slight decreasing is detected, while at 40°C signals reduce by *ca.* 15% (Figure S1). Analogously, CD bands decrease upon titration of the **1**-SrP_2 solution with cryptand [2.2.2], due to the gradual removal of Sr^{2+} and concomitant disaggregation (Figure 2c).[19] The CD band-shape of uncomplexed **1** reappears when one eq. of cryptand [2.2.2] relative to Sr^{2+} is added (Figure S2). Subsequent titration with trifluoromethanesulfonic acid (Htf) causes the gradual release of Sr^{2+} from cryptate [$Sr^{2+}\subset$ 2.2.2] and the consequent regeneration of the assembled species. The CD spectrum of the **1**-SrP_2 complex is almost quantitatively recovered upon addition of one eq. of Htf relative to cryptand [2.2.2] (Figure S2). Noteworthily, an isodichroic point is detectable at 280 nm (Figure 2 and S2), suggesting that only two species are present in solution during the titration experiments, *i.e.* uncomplexed **1** and the G_4 based complex. The self-assembly process of **1** in the presence of SrP_2 can then be considered in terms of a simple two-species equilibrium.

In-solution Small Angle Neutron Scattering (SANS) experiments have been performed on the **1**-SrP_2 complex in chloroform. Results obtained as a function of temperature (see Figure 3) clearly suggest that heating does not modify the structure of the scattering particles, although the reduction of the scattering intensity at zero angle ($I(0)$) indicates that some disaggregation process is taking place.

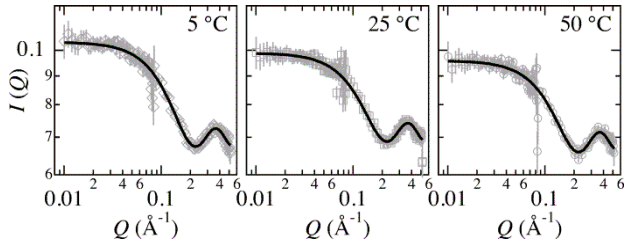


Figure 3. SANS profiles of the **1**-SrP_2 solution ([**1**] 50 mM), obtained upon extraction of 1/6 eq. of strontium picrate. Continuous lines are best fitting curves. See S.I. for fitting parameters.

In particular, the particle number density calculated from $I(0)$ reduces to 87% and 64% when the temperature is increased from 5 to 25 and 50°C, respectively. According to previous results,[20] a core-shell cylinder model is used to fit the curves: fitting parameters are the core radius (R_c), shell thickness (t_s) and length (L) of the cylinder model and the scattering length densities of the core (SLD_c) and shell (SLD_s) regions. Fitting results (Figure 3) show that **1** in the presence of SrP_2 forms short, monodisperse cylindrical particles, whose dimensions are compatible with the presence of discrete assembled dodecamers (**D**s), arising from the piling up of three G_4s. Indeed, and irrespective of temperature, the cylinders length results about 12 Å (*ca.* 3 times 3.4 Å, which is the stacking distance), and the total radius is *ca.* 12.3 Å (*i.e.* compatible with the G_4 radius).[20,21] Fitted dimensions and *SLDs* also confirm the supposed stoichiometry of the complex (**1**/Sr^{2+} = 6/1). In fact, the fitted R_c (8 Å) and SLD_c (4.3 10^{-6} Å^{-2})-indicate that the core region includes not only the G_4 central hole but also the guanine region, because of a similar *SLD*. Note that the calculated *SLD* for the guanine residues is 4.32 10^{-6} Å^{-2}, while the *SLD* of the **D** central cavity filled by two strontium ions only is just 4.31 10^{-6} Å^{-2}.

^{1}H-NMR experiments show that **1** in $CDCl_3$ in the absence of salts (Figure 4a) self-assembles into dimers or oligomeric ribbon-like structures.[13,17] Indeed, a deshielding of amino and imino protons is observed either by lowering temperature or increasing concentration, indicating their progressive engagement in H-bonded aggregates (Figures S3-S4).

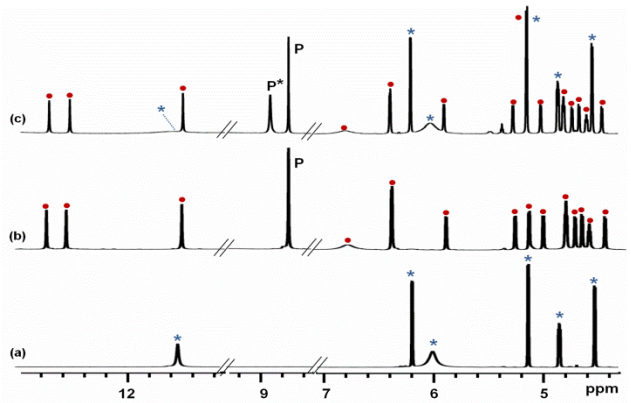


Figure 4. ^{1}H NMR spectra recorded in $CDCl_3$ at 0°C. **a**, **1** (8 mM). **b**, **1**-SrP_2 solution obtained after extraction of 1/6 eq. of SrP_2. **c**, **1**-SrP_2 solution added with ½ eq. of [2.2.2] relative to Sr^{2+}. Blue stars: **1** signals; red dots: **D** signals; "P": picrate resonance corresponding to the **D** counterion; "P*": picrate resonance corresponding to the [$Sr^{2+}\subset$ 2.2.2] counterion.

Upon solid-liquid extraction of increasing amounts of SrP_2, the original set of signals is progressively replaced by three new sets (Figure S5), that are sharp and of equal intensity at 0°C. These three sets of signals become exclusive when 1/6 of salt has been extracted (Figure 4b). Beyond this point, no further Sr^{2+} uptake or spectral change occurs. The **1**/Sr^{2+} stoichiometry of 6:1, inferred from the relative signal intensities of **1** and picrate (at 8.74 ppm), is hence in accordance with the formation of a discrete **D** species[22-24] whose molecularity is $[\mathbf{1}]_{12}Sr_2P_4$. ^{1}H NMR analysis confirms that complex **D** can be disaggregated and regenerated upon titration with cryptand [2.2.2] and Htf, respectively (Figures 4c and S6). As already evinced from CD data, no species other than uncomplexed **1** and **D** are observed in solution during the titration experiments. All results point out that the cation-templated self-assembly of **1** is a reversible and high cooperative process.[25] In the spectrum of **D** at 0°C the exocyclic amino protons are exchange-broadened into the baseline, and only a weak signal is visible at *ca.* 6.8 ppm (Figure 4b). However, the signature of at least two G_4s emerges by lowering temperature down to -30°C, when two sets of well separated resonances, corresponding to H-bonded and free amino protons, are observed (Figure S7).[21,26] By combined COSY, HSQC, HMBC, 1D and 2D-NOE experiments performed at -20°C, the three sets of signals observed in the ^{1}H-NMR spectrum of **D** can be assigned to the three G_4s (labeled α, β and γ, Figure S8 and table S1) and a full structural characterization of **D** in solution is obtained (see SI for details).

The two faces of a G_4 are diastereotopic and usually are referred to as *head* (*H*) (clockwise motion on going from H-bond donors to acceptors) or *tail* (*T*) (counterclockwise motion) (Figure S10).[26] From the stereochemical point of view, every possible combination

ChemComm Accepted Manuscript

This journal is © The Royal Society of Chemistry 2012

of three stacked G4s results in a dodecameric complex of C_4 symmetry, which is in agreement with the three sets of signals observed in both ^{1}H and ^{13}C-NMR spectra (Figure S9). The 2D-Noesy spectrum (Figure 5) shows several correlations between quartets α and γ: cross peaks (confirmed by Noesy1d experiments) relating αH1' and γH1', αH5' and γH5', αH4' and γH4', αH5' and γH4' are clearly visible.

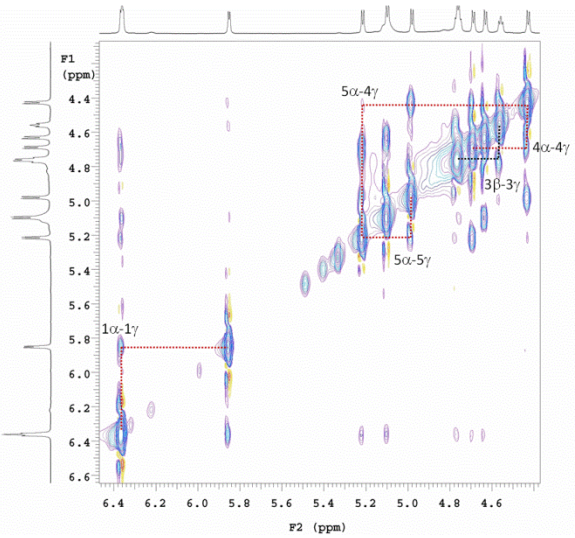


Figure 5. Sugars region of the 2D-NOESY spectrum ($CDCl_3$, -20°C) of the **1**-SrP_2 solution ([**1**] 30 mM), obtained after extraction of 1/6 eq. of strontium picrate (mixing time 300 ms). (Apices are omitted for clarity).

As **1** is covalently blocked in the *anti* conformation and possesses (*S*) configuration at C5', protons H1', H4' and H5' are all *syn* with respect to the molecular plane defined by the nucleobase within each **1** unit as well as within each G4. In particular, in a G4 they all point towards the *H* side, which, according to models, appears to be the less sterically crowded (Figure S10). The NOEs listed above thus suggest an *H* to *H* heteropolar stacking between quartets α and γ. As a consequence, quartet β must necessarily sit on the *T* face of either quartet α or γ. While no NOE relating quartets α and β could be observed, 2D-NOESY and Noesy1d spectra (Figure 5, S14) show proximity between γH3' and both βH3' and βH4', implying stacking between quartets γ and β: in particular, the *T* face of quartet γ is in contact with the *H* face of quartet β. Molecular modeling supports this conclusion, as the *T* side of the quartet turns out to be far more sterically crowded, thus ruling out a *T* to *T* stacking. A model summarizing the observed contacts is shown in Figure 6. It is interesting to notice that the picrate anion shows NOE contacts with both H1' and H5' of the γ middle quartet (Figure S14,15). Moreover, the anion self-diffuses exactly at the same rate as the **D** complex in DOSY experiments at 25°C (diffusion coefficient = $3.8 \cdot 10^{-10} \pm 0.2$ m^2/s, hydrodynamic radius = 10.6 ± 0.3 Å). This confirms that the anions are tightly bound and located at the periphery of the complex (Figure 6).

Many G4-based lipophilic complexes have been reported so far in the literature, ranging from isolated G4s up to *pseudo*-polymeric stacked assemblies.[12,27] Concerning complex stoichiometry, octamers are probably the most frequent examples, together with hexadecamers.[27,28] On the other hand, while isolated tetramers have been described in a few papers,[29,30] to the best of our knowledge only three examples of dodecameric complexes have been reported so far,[22-24] and only in one case the complex, although present in a mixture, has been fully characterized.[22] Remarkably, in all lipophilic GQs so far reported, guanosines adopt preferentially, if not exclusively, a *syn* conformation around the glycosydic bond. Thus,

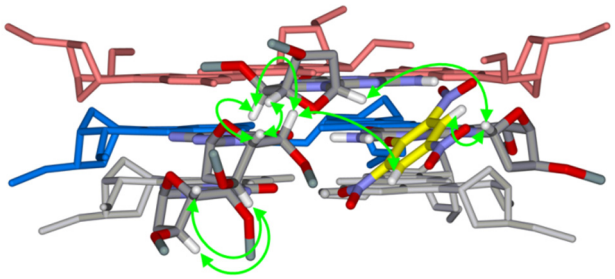


Figure 6. Proposed model for the $[1]_{12}Sr_2P_4$ dodecamer. The double-headed arrows indicate selected interquartet NOEs. Some atoms, including the two central Sr^{2+} ions are omitted for clarity. The three stacked quartets are shown in different colors (top: G4 α, pink; middle: G4 γ, blue; bottom: G4 β, gray). The picrate (only one shown) aromatic ring is highlighted in yellow.

isolated G4s have been observed in solution only in cases where a bulky substituent at the 8 position of the nucleobase forces molecules in the *syn* conformation. Octamers can be composed of either an all-*anti*[26] or an all-*syn*[31] quartet stacked on top of an all-*syn* quartet. Only *syn* conformers are present in at least one of the dodecamers so far reported.[22] In hexadecamers both conformations can be present, but *syn* is the preferred one,[12] and the same holds for the *pseudo*-polymer.[21]

Conclusions

The system described here is the first example of a lipophilic all-*anti* GQ and this finding rules out the presence of the *syn* conformation as a critical factor for G4 based self-assembly in non-aqueous medium. From the biological point of view, this modified nucleoside surely retains, at least qualitatively, the ability of pristine guanosine to participate to the GQ formation. However, its presence could introduce on GQ architecture a structural perturbation, which may play a role for the *in vivo* signaling and recognition processes connected with the repair machinery in cells. Remarkably, no dodecameric assemblies have been observed for the (5'R) diastereoisomer under the same or similar conditions. A comparative study of the self-assembly behavior of (5'S)-**1**, its (5'R) diastereoisomer and 3',5'-di-tert-butyldimethylsilyl-2'-deoxyguanosine in the presence of different salts is currently in progress.

This work was financially supported by the EC FP7 ICT-Molarnet project (318516) (S.M. and S.P.) and by the Marie Curie Action FP7-PEOPLE-2011-IEF-CYCLOGUO (298555) (M.A.T and C.C.). The support and sponsorship of the COST Action CM1201 on Biomimetic Radical Chemistry is also gratefully acknowledged. PM and EJB thank the Italian MIUR for financial support in the frame of the National Interest Research Program (PRIN 2010-2011, grant 2010LKE4CC).

Notes and references

[a] Dipartimento di Chimica 'Giacomo Ciamician' Alma Mater Studiorum - Università di Bologna, via San Giacomo 11, 40126 Bologna, Italy E-mail: stefano.masiero@unibo.it.

ChemComm Accepted Manuscript

[b] Istituto per la Sintesi Organica e la Fotoreattività, Consiglio Nazionale delle Ricerche, Via P. Gobetti 101, 40129 Bologna, Italy E-mail: chrys@isof.cnr.it.
[c] Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Via Ranieri 65, Ancona, Italy.
[d] Institut Laue-Langevin, BP 156 F-38042 Grenoble, France

† To the memory of Gian Piero Spada.
Electronic Supplementary Information (ESI) available: materials and methods, CD, NMR spectra and spectral data. See DOI: 10.1039/c000000x/

1 G. Biffi, M. Di Antonio, D. Tannahill and S. Balasubramanian, Nature Chem, 2014, 6, 75.
2 G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nature Chem.*, 2013, **5**, 182.
3 H. J. Lipps and D. Rhodes, *Trends Cell Biol.*, 2009, **19**, 414.
4 T. Biver, *Coord. Chem. Rev.*, 2013, **257**, 2765.
5 *Guanine quartets: structure and application*. (Eds. W. Fritzsche, L. Spindler), RSC, 2013.
6 S. Balasubramanian and S. Neidle, *Curr. Opin. Chem. Biol*, 2009, **13**, 345.
7 C. Chatgilialoglu, C. Ferreri and M. A. Terzidis, *Chem. Soc. Rev*., 2011, **40**, 1368.
8 D. Mitra, X. Luo, A. Morgan, J. Wang, M. P. Hoang, J. Lo, C. R. Guerrero, J. K. Lennerz, M. C. Mihm, J. A. Wargo, K. C. Robinson, S. P. Devi, J. C. Vanover, J. A. D'Orazio, M. McMahon, M. W. Bosenberg, K. M. Haigis, D. A. Haber, Y. Wang and D. E. Fisher, *Nature*, 2012, **491**, 449.
9 V. P. Jasti, R. S. Das, B. A. Hilton, S. Weerasooriya, Y. Zou and A. K. Basu, *Biochemistry*, 2011, **50**, 3862.
10 C. You, X. Dai, B. Yuan, J. Wang, J. Wang, P. J. Brooks, L. J. Niedernhofer and Y. Wang, *Nat. Chem. Biol.,* 2012, **8**, 817.
11 H. Huang, R. S. Das, A. K. Basu and M. P. Stone, *J. Am. Chem. Soc.*, 2011, **133**, 20357.
12 J. T. Davis, *Angew. Chem. Int. Ed.*, 2004, **43**, 668.
13 S. Masiero, S. Pieraccini and G. P. Spada, in *Molecular Self-Assembly: Advances and Applications*, (Ed. A. D. Li), Pan Stanford Publishing, Singapore, 2012, pp. 93-122.
14 M. A. Terzidis and C. Chatgilialoglu, *Aust. J. Chem.*, 2013, **66**, 330.
15 J. Deng, Y. Xiong and M. Sundaralingam, *PNAS* 2001, **98,** 13665.
16 B. I. Kankia and L. Marky, *J. Am. Chem. Soc.*, 2001, **123***,* 10799.
17 G. Gottarelli, S. Masiero, E. Mezzina, S. Pieraccini, J. Rabe, P. Samorì and G. P. Spada, *Chem. Eur. J.*, 2000, **6**, 3242.
18 A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo and M. W. da Silva, *Angew. Chem. Int. Ed.,* 2011, **50**, 10645.
19 A. Ciesielski, S. Lena, S. Masiero, G. P. Spada and P. Samorì, *Angew. Chem. Int. Ed.*, 2010, **49**, 1963.
20 A. Gonnelli, M. G. Ortore, E. J. Baldassarri, G. P. Spada, S. Pieraccini, R. C. Perone, S.S. Funari and P. Mariani, *J. Phys. Chem. B*, 2013, **117**, 1095.
21 E. Mezzina, P. Mariani, R. Itri, S. Masiero, S. Pieraccini, G. P. Spada, F. Spinozzi, J. T. Davis and G. Gottarelli, *Chem. Eur. J.*, 2001, **7**, 388.
22 M. del C. Rivera-Sánchez, I. Andújar-de-Sanctis, M. García-Arriaga, V. Gubala, G. Hobley and J. M. Rivera, *J. Am. Chem. Soc.*, 2009, **131,** 10403.
23 D. González-Rodríguez, J. L. J. van Dongen, M. Lutz, A. L. Spek, A. P. H. J. Schenning and E. W. Meijer, *Nat. Chem.*, 2009, **1**, 151.
24 I. C. M. Kwan, Y. M. She and G. Wu, *Chem. Commun*., 2007, 4286.
25 L. J. Prins, D. N. Reinhoudt and P. Timmerman, *Angew. Chem. Int. Ed.*, 2001, **40**, 2382.
26 A. L. Marlow, E. Mezzina, G. P. Spada, S. Masiero, J. T. Davis and G. Gottarelli, *J. Org. Chem.*, 1999, **64**, 5116.
27 J. T. Davis and G. P. Spada, *Chem. Soc. Rev.*, 2007, **36**, 296.
28 M. Martin-Hidalgo and J. M. Rivera, *Chem. Commun.* 2011, **47**, 12485.
29 J. L. Sessler, M. Sathiosatham, K. Doerr, V. Lynch and K.A. Abboud, *Angew. Chem. Int. Ed*., 2000, **39**, 1300.
30 M. Nikan and J. C. Sherman, *J. Org. Chem*., 2009, **74**, 5211.
31 X. Liu, I. C. M. Kwan, S. Wang and G. Wu, *Org. Lett*., 2006, **8**, 3685.

ChemComm Accepted Manuscript

This journal is © The Royal Society of Chemistry 2012