ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2021, **57**, 1659

Received 24th November 2020,
Accepted 4th January 2021

DOI: 10.1039/d0cc07707k

rsc.li/chemcomm

Crystalline supramolecular organic frameworks *via* hydrogen-bonding between nucleobases†

Miguel Martín-Arroyo, [a] Javier Castells-Gil, [b] Nerea Bilbao, [a] Neyvis Almora-Barrios, [b] Carlos Martí-Gastaldo *[b] and David González-Rodríguez *[ac]

We report a crystalline supramolecular framework assembled by H-bonding interactions between covalently fused monomers equipped with two guanine–cytosine nucleobase pairs.

Supramolecular organic frameworks (SOFs) have reached an important position in the field of crystalline molecular materials.[1–3] As their covalent and metal–organic counterparts, COFs[4,5] and MOFs,[6,7] they are also formed by the interaction between diverse building blocks through dynamic and reversible bonds, and present similar features in terms of predictable structure,[8] high surface area and tailored pore size/shape.[9] One of the most interesting supramolecular interactions to construct SOF materials is the hydrogen (H)-bond, which gives rise to the so called H-bonded organic frameworks (HOFs).[10–12] Many chemical groups are susceptible to be used as H-bond motifs, and some of them, such as the diaminotriazine,[13,14] urea,[15] benzoic acid units[16,17] or amidinium-carboxylate pairs,[18] have been employed in the literature to produce HOFs. The first reports of porous materials made from these motifs were described decades ago, as part of the structural study of other systems, but their consideration and analysis as ordered frameworks for structure-to-function correlations is relatively recent.[19] It is now clear that the challenges imposed by the high directionality of H-bonds and the isomerism and multiple binding possibilities of many H-bonding fragments, which often results in crystal polymorphs,[20] hampers the assembly of targeted frameworks. Hence, the design of HOFs remains nowadays quite unpredictable, with most HOF materials made serendipitously or by trial-and-error approaches.

A prototypical group of H-bonding modules is constituted by the nucleobases. It is clear that mastering the use of these bioinspired motifs in HOFs might bring unprecedented structural designs due to their synthetic versatility and the availability of two complementary DNA base pairs. However, despite their popular use in functional supramolecular systems,[21] there are not many examples of the utilization of nucleobases in crystalline HOF materials. Wasielewski and co-workers recently reported different crystalline materials by the formation of G-quartets between building blocks formed by a large π-conjugated systems functionalized with two guanines at the edges.[22,23] Their structural determination was realized by comparing the recorded PDRX data and different simulated diffractograms, in which they evaluated various stacking modes between 2D layers. On the other hand, Richert and co-workers described the use of a tetrahedral core substituted with dinucleotide arms.[24] This molecule formed aggregates at 95 °C in water, supposedly by the complementary interaction between nucleobases, but their small size prevented to characterize them by X-ray diffraction. In any case, these scarce examples reveal the difficulties in producing crystalline frameworks by controlling H-bonding between nucleobases.

Here, we describe our first advances in the preparation of a new family of HOFs built from supramolecular interactions between nucleobases. This work is supported on our previous studies on the cyclization of self-complementary dinucleoside monomers in solution[25] and onto graphite substrates,[26] where highly ordered 2D networks formed by cyclic units could be imaged by STM at the solid–liquid interface (Fig. 1a). These results clearly invited to generate related 3D crystalline materials in which the bidimensional cavities could be extended along the *z*-axis and materialized into actual pores. Unfortunately, the formation of these 2D networks relied on weak and ill-defined secondary interactions between cyclic entities, and therefore the preparation of bulk crystalline materials from the same dinucleoside monomers could not be realized. In order to solve this problem, we considered a novel design that involved the covalent fusion of two monomers (Fig. 1b). The idea was to rely exclusively on Watson–Crick interactions between nucleobases

[a] *Departamento de Química Orgánica, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain. E-mail: david.gonzalez.rodriguez@uam.es*
[b] *Instituto de Ciencia Molecular, Universitat de Valencia, Catedrático José Beltrán 2, 46980 Paterna, Spain*
[c] *Institute for Advanced Research in Chemical Sciences (IAdChem), Madrid, Spain*

† Electronic supplementary information (ESI) available: Synthetic details, ^{1}H NMR spectra, UV-vis absorption, fluorescence, PDRX diffractograms, theoretical study and gas absorption. See DOI: 10.1039/d0cc07707k

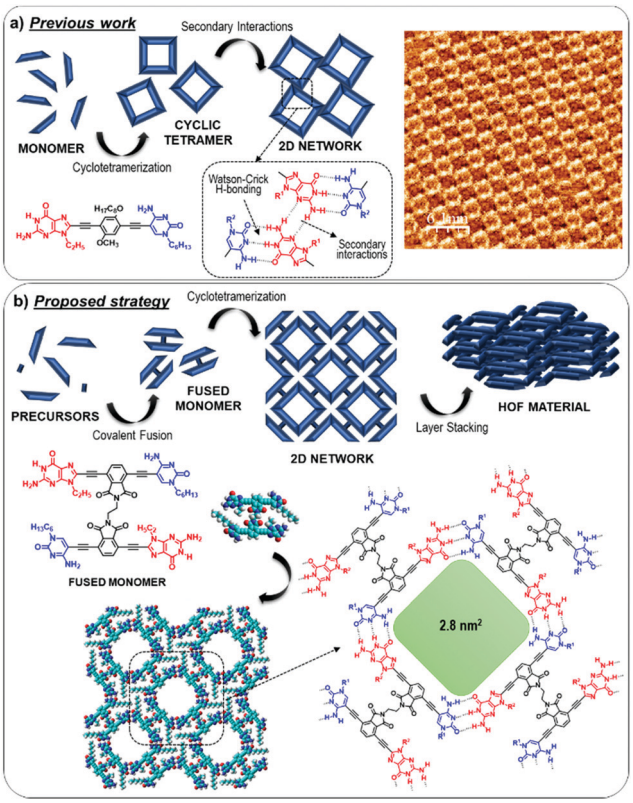


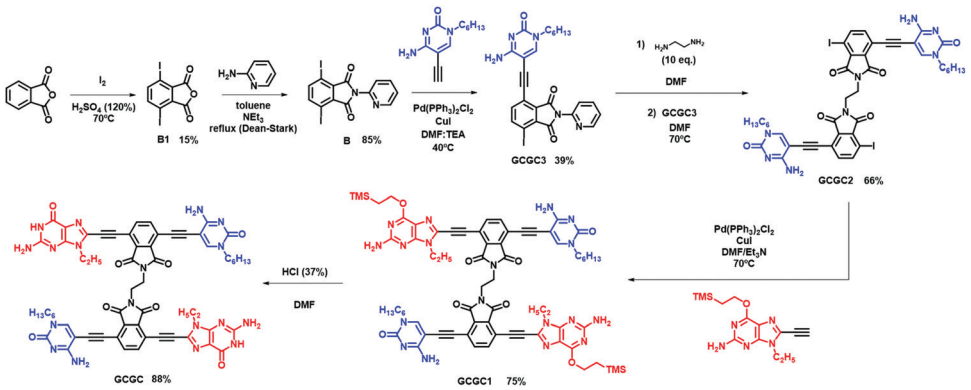
Fig. 1 Comparison between (a) previous work in the group and (b) the proposed process for the preparation of HOFs by covalent fusion, H-bonding association in the plane, and consecutive layer staking.

to direct network formation. This would minimize the contribution of weaker and less defined intermolecular interactions, thus facilitating the formation of crystalline materials with a targeted structure. In the designed fused monomer **GCGC** (Fig. 1b), two guanine(G)–cytosine (C) linear fragments are connected by a bis-phthalimide linker through the central block. The selected nucleobases were substituted at specific positions by alkyl chains in order to occupy the (presumably formed) secondary pores and to avoid as much as possible other relative conformations between nucleobases. With this design, we intended to induce the formation of a 2D network with pores of *ca.* 2.8 nm^2 (Fig. 1b).

The synthetic route to **GCGC** is shown in Scheme 1. As a preliminary step, 5-/8-ethynylated C and G derivatives (**C**, **G**) were obtained following reported routes.[27] The synthesis starts with the iodination of phthalic anhydride. This anhydride was too labile to stand subsequent couplings and purification steps, so its protection was mandatory. Inspired by strategies used in the field of polymeric materials,[28,29] we decided to protect it as a phthalimide through a condensation reaction with 2-aminopyridine, to obtain the dihalogenated central block **B**. Due to the electron-poor nature of the resulting phthalimide we were able to functionalize later this position through a transimidation reaction with another amine nucleophile. The synthetic route continues with a Sonogosahira reaction to couple the central block **B** and nucleobase **C**. Then, the resulting compound (**GCGC3**) was fused first by a transimidation reaction with excess of ethylenediamine, and, after the evaporation of the unreacted diamine, a second condensation with additional 1.1 equivalents of **GCGC3** was carried out, to yield **GCGC2**. Finally, a second Sonogashira reaction was used to couple guanosine **G**, producing **GCGC1**. At this point, the protecting group at the carbonyl position of the guanosine fragment, which supplies solubility to the fused monomer, can be cleaved in acid media (or, alternatively, in the presence of fluoride salts), leading to the final fused compound **GCGC**.

GCGC turned out to be very insoluble even in polar solvents that strongly compete for H-bonding, such as DMSO or DMF. The aggregation process could only be monitored by ^{1}H NMR in DMSO-d_6 (Fig. 2a and Fig. S1, ESI†). At room temperature, the spectrum shows broad peaks, characteristic of strongly aggregated species, that became sharper when the sample was heated up to 100 °C, so they could be assigned to each proton. No NMR peaks related to the formation of G:C H-bonded species could be detected along this temperature range: either the monomer is detected, with a characteristic DMSO-bound G-amide proton at *ca.* 10.5 ppm, or large aggregates that produce too broad H-bonded signals and that precipitate out of solution are formed within the 10^{-2}–10^{-3} M range.

These DMSO suspensions were also analysed within a lower concentration range with absorption and emission spectroscopies at lower concentrations (Fig. S2, ESI†). Room temperature **GCGC** suspensions revealed a broad, low-intensity fluorescence emission



Scheme 1 Synthetic route to the fused monomer **GCGC**.

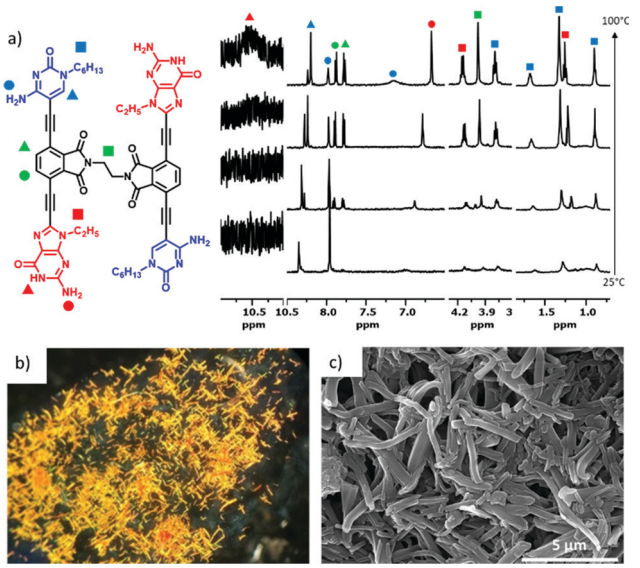


Fig. 2 (a) ^{1}H RMN spectra of GCGC in DMSO-d_6 at different temperatures; (b) needle like crystals at the optic microscope with polarized light; and (c) SEM images of **GCGC** crystals.

with maxima at 538 nm. When the temperature is increased or the concentration is decreased, a significant increase in emission intensity and a blue-shift to 458 nm was immediately noted, a behaviour that is similar to other G–C monomers studied in the group and that can be attributed to monomer dissociation.[25] However, the fully disaggregated samples at high temperatures required several days to aggregate again and recover the spectroscopic features of the original suspensions, thus hampering us to analyze the polymerization process as a function of temperature.

Next, we concentrated our efforts on optimizing the conditions required for producing crystals. However, because of the extremely low solubility of **GCGC**, and thus our limitation to the use of DMSO at high temperatures and low concentrations to fully dissolve it, the crystals obtained were of low quality. We therefore decided to work with the more soluble **GCGC1** precursor and carry out the deprotection of the carbonyl group in a controlled way. This strategy, also used in MOFs,[30] allows to gain control over the crystallization process due to the gradual release of **GCGC** in the reaction medium. After testing different concentrations, solvent compositions and deprotection conditions, the best results were obtained when a solution of **GCGC1** in DMF was treated with a slight excess of HCl, and MeOH was added. Additionally, a thermal treatment was realised in order to improve crystallinity. Then, after washing with MeOH, orange needle-shaped crystals were obtained.

Optical microscopy images of the solid obtained from MeOH (Fig. 2b and Fig. S3, ESI†) suggested the formation of needle-shaped crystals of 0.5 μm, as confirmed by SEM (Fig. 2c and Fig. S4, ESI†). On the other hand, quantitative deprotection of the molecules was confirmed by ^{1}H NMR analysis. In addition, FT-IR (Fig. S5, ESI†) analysis in the C=O stretching range reveals a new couple of peaks at 1640 and 1650 cm^{-1} that were not observed in the protected fused monomer nor in each of the nucleobases separately. However, the mixture of both nucleobases also reveals the presence of those peaks, so we can attribute them to H-bonded carbonyl groups in **GCGC**.[31]

We preliminary checked the crystallinity of the solid in the diffractometer available in our laboratory. As shown in Fig. S6a (ESI†), the powder X-ray diffraction (PXRD) pattern displays broad diffraction lines with low intensity. Similar patterns were obtained under different acid deprotection conditions, which suggests that the structure formed is not very sensitive to slight pH variations. In order to obtain data of sufficient quality to guide the refinement of the PXRD pattern for building a structural model, we collected Synchrotron diffraction at the beamline BL04-MSPD in ALBA. These measurements are consistent with the PXRD pattern collected in our facilities and suggest the formation of a polycrystalline solid featuring diffraction at low-angles, with the most intense diffraction line at 2θ = 2.4°, as well as a peak of high relative intensity at 13.5°, which was attributed to a π–π stacking distance of *ca.* 3.5 Å (λ = 0.82629 Å, Fig. S6, ESI†).

We computationally designed several 2D H-bonded networks by using Materials Studio (Fig. S7A, ESI†) to analyse similarities with the experimental pattern. This set of potential candidates included the targeted Watson–Crick-bonded network shown in Fig. 1b. These 2D layers were then stacked in diverse relative arrangements and compared to the experimental PXRD. In a first approach, we simulated networks in an eclipsed conformation, *i.e.* ···AA··· stacking, and with a shift between layers, *i.e.* ···AB··· stacking with 5 Å shifts (see ESI† for details). Among the different potential networks, the best fit was obtained with the one with cell parameters a = 22.6 Å, b = 7.2 Å, c = 19.4 Å and β = 97°, in which alternated C:C and G:G interactions direct the assembly of a layered framework (Fig. 3b and Fig. S7Ab, ESI†).

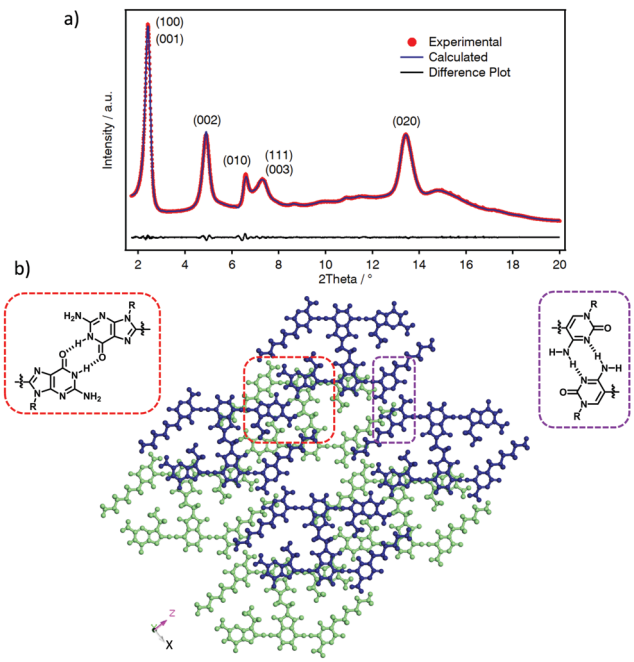


Fig. 3 (a) LeBail profile fitting of the experimental PXRD pattern (λ = 0.82629 Å) and (b) proposed structure with C:C and G:G interactions.

This network was then studied in more detail by changing the relative disposition of the layers with a higher number of possibilities. We also considered different numbers of layers in the unit cell by varying the *b* axis without substantial changes in the simulated PXRD pattern (Fig. S7C, ESI†). Eventually, the best match was obtained by assuming an ···AB··· packing with a shift of 12 Å along the *a* axis y 18 Å along the *c* axis between layers (Fig. 3). For this case, the LeBail refinement converged with excellent residual values (R_{wp} = 2.55%, R_{exp} = 1.77%) to a monoclinic crystal system. Furthermore, the absence of large pores in the simulated structure is consistent with the low adsorption capacity of the material, which is negligible for N_2 and displays a small uptake of CO_2 at 273 K of 85 $cm^3\ g^{-1}$ (Fig. S8, ESI†). The final **GCGC** material is not very sensitive to temperature changes and no significant weight loss is detected up to 350 °C in thermogravimmetric analyses (Fig. S9, ESI†).

In short, novel crystalline framework materials derived from H-bonding interactions between nucleobases have been obtained in this work from a fused monomer substituted with two complementary G and C base pairs. However, the resulting structure obtained under thermodynamic conditions is not the one expected in the original design. Since nucleobases present multiple possibilities to form H-bonds aside from the Watson–Crick pair,[32] in this particular case the bases prefer to interact through weaker doubly H-bonded C:C and G:G interactions, thus leading to a different, more compact network with smaller voids. These interaction modes, however, do not bring the versatility and potential that complementary nucleobases could supply if they are made to bind *via* Watson–Crick interactions, as shown in our previous work.[25,26,33] This first work suggests that monomer structure needs to be redesigned to reach such goal.

Funding from the European Research Council (ERC-StG 279548) and MINECO (CTQ2014-57729-P and CTQ2017-84727-P) is gratefully acknowledged. M. M.-A. is grateful to MINECO for an FPI Grant (BES-2015-071795). We also acknowledge ALBA Synchrotron and Catalin Popescu (BL04-MSPD) for provision of synchrotron radiation facilities.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 J. Tian, H. Wang, D.-W. Zhang, Y. Liu and Z.-T. Li, *Natl. Sci. Rev.*, 2017, **4**, 426–436.
2 J. Tian, T. Y. Zhou, S. C. Zhang, S. Aloni, M. V. Altoe, S. H. Xie, H. Wang, D. W. Zhang, X. Zhao, Y. Liu and Z. T. Li, *Nat. Commun.*, 2014, **5**, 5574.
3 M. Pfeffermann, R. Dong, R. Graf, W. Zajaczkowski, T. Gorelik, W. Pisula, A. Narita, K. Müllen and X. Feng, *J. Am. Chem. Soc.*, 2015, **137**, 14525–14532.
4 S. Y. Ding and W. Wang, *Chem. Soc. Rev.*, 2013, **42**, 548–568.
5 M. S. Lohse and T. Bein, *Adv. Funct. Mater.*, 2018, **28**, 1705553.
6 J. R. Li, R. J. Kuppler and H. C. Zhou, *Chem. Soc. Rev.*, 2009, **38**, 1477–1504.
7 A. Corma, H. García and F. X. Llabrés I Xamena, *Chem. Rev.*, 2010, **110**, 4606–4655.
8 A. Pulido, L. Chen, T. Kaczorowski, D. Holden, M. A. Little, S. Y. Chong, B. J. Slater, D. P. McMahon, B. Bonillo, C. J. Stackhouse, A. Stephenson, C. M. Kane, R. Clowes, T. Hasell, A. I. Cooper and G. M. Day, *Nature*, 2017, **543**, 657–664.
9 H. Deng, S. Grunder, K. E. Cordova, C. Valente, H. Furukawa, M. Hmadeh, F. Gándara, A. C. Whalley, Z. Liu, S. Asahina, H. Kazumori, M. O'Keeffe, O. Terasaki, J. F. Stoddart and O. M. Yaghi, *Science*, 2012, **336**, 1018–1023.
10 I. Hisaki, C. Xin, K. Takahashi and T. Nakamura, *Angew. Chem., Int. Ed.*, 2019, **58**, 11160–11170.
11 R. B. Lin, Y. He, P. Li, H. Wang, W. Zhou and B. Chen, *Chem. Soc. Rev.*, 2019, **48**, 1362–1389.
12 B. Wang, R. B. Lin, Z. Zhang, S. Xiang and B. Chen, *J. Am. Chem. Soc.*, 2020, **142**, 14399–14416.
13 F. Helzy, T. Maris and J. D. Wuest, *J. Org. Chem.*, 2016, **81**, 3076–3086.
14 P. Li, Y. He, J. Guang, L. Weng, J. C.-G. Zhao, S. Xiang and B. Chen, *J. Am. Chem. Soc.*, 2014, **136**, 547–549.
15 M. Mastalerz and I. M. Oppel, *Angew. Chem., Int. Ed.*, 2012, **51**, 5252–5255.
16 I. Hisaki, S. Nakagawa, N. Ikenaka, Y. Imamura, M. Katouda, M. Tashiro, H. Tsuchida, T. Ogoshi, H. Sato, N. Tohnai and M. Miyata, *J. Am. Chem. Soc.*, 2016, **138**, 6617–6628.
17 K. Ma, P. Li, J. H. Xin, Y. Chen, Z. Chen, S. Goswami, X. Liu, S. Kato, H. Chen, X. Zhang, J. Bai, M. C. Wasson, R. R. Maldonado, R. Q. Snurr and O. K. Farha, *Cell Reports Phys. Sci.*, 2020, **1**, 100024.
18 S. A. Boer, M. Morshedi, A. Tarzia, C. J. Doonan and N. G. White, *Chem. – Eur. J.*, 2019, **25**, 10006–10012.
19 Y. He, S. Xiang and B. Chen, *J. Am. Chem. Soc.*, 2011, **133**, 14570–14573.
20 B. Moulton and M. J. Zaworotko, *Chem. Rev.*, 2001, **101**, 1629–1658.
21 A. Prado, D. González-Rodríguez and Y. Wu, *ChemistryOpen*, 2020, **9**, 409–430.
22 Y. L. Wu, N. E. Horwitz, K. S. Chen, D. A. Gomez-Gualdron, N. S. Luu, L. Ma, T. C. Wang, M. C. Hersam, J. T. Hupp, O. K. Farha, R. Q. Snurr and M. R. Wasielewski, *Nat. Chem.*, 2017, **9**, 466–472.
23 Y.-L. Wu, N. S. Bobbitt, J. L. Logsdon, N. E. Powers-Riggs, J. N. Nelson, X. Liu, T. C. Wang, R. Q. Snurr, J. T. Hupp, O. K. Farha, M. C. Hersam and M. R. Wasielewski, *Angew. Chem., Int. Ed.*, 2018, **57**, 3985–3989.
24 A. Singh, M. Tolev, M. Meng, K. Klenin, O. Plietzsch, C. I. Schilling, T. Muller, M. Nieger, S. Bräse, W. Wenzel and C. Richert, *A ngew. Chem., Int. Ed.*, 2011, **50**, 3227–3231.
25 C. Montoro-García, J. Camacho-García, A. M. López-Pérez, M. J. Mayoral, N. Bilbao and D. González-Rodríguez, *Angew. Chem., Int. Ed.*, 2016, **55**, 223–227.
26 N. Bilbao, I. Destoop, S. De Feyter and D. González-Rodríguez, *Angew. Chem., Int. Ed.*, 2016, **55**, 659–663.
27 N. Bilbao, V. Vázquez-González, M. T. Aranda and D. González-Rodríguez, *Eur. J. Org. Chem.*, 2015, 7160–7175.
28 C. G. Herbert, H. Ghassemi and A. S. Hay, *J. Polym. Sci., Part A: Polym. Chem.*, 1997, **35**, 1095–1108.
29 T. P. Bender and Z. Y. Wang, *J. Polym. Sci., Part A: Polym. Chem.*, 2000, **38**, 3991–3996.
30 N. M. Padial, E. Quartapelle Procopio, C. Montoro, E. López, J. E. Oltra, V. Colombo, A. Maspero, N. Masciocchi, S. Galli, I. Senkovska, S. Kaskel, E. Barea and J. A. R. Navarro, *Angew. Chem., Int. Ed.*, 2013, **52**, 8290–8294.
31 Y. Kyogoku, R. C. Lord and A. Rich, *Science*, 1966, **154**, 518–520.
32 M. J. Mayoral, N. Bilbao and D. González-Rodríguez, *ChemistryOpen*, 2016, **5**, 10–32.
33 S. Romero-Pérez, J. Camacho-García, C. Montoro-García, A. M. López-Pérez, A. Sanz, M. J. Mayoral and D. González-Rodríguez, *Org. Lett.*, 2015, **17**, 2664–2667.

