# Uracil grafted imine-based covalent organic framework for nucleobase recognition

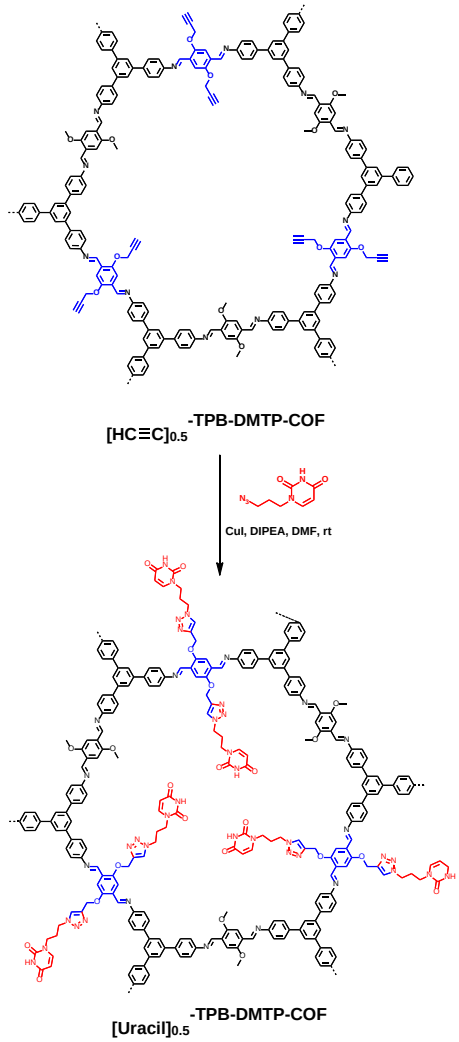
Sergio Royuela,[a,b] Eduardo Garcia-Garrido,[c] Miguel Martín Arroyo,[d] María J. Mancheño,[a] María M. Ramos,[b] David Gonzalez-Rodríguez,[d,f] Álvaro Somoza,*[c] Félix Zamora*[c,e,f,g] and José L. Segura*[a]

**An imine-based covalent-organic framework (COF) decorated in its cavities with uracil groups has shown selective recognition towards adenine in water. These results show how the confinement of the base-pair inside the COF's pores allows a remarkable selective recognition in aqueous media.**

The development of porous materials has accompanied the demands of modern society and a large variety of porous crystalline materials (e.g., zeolites,[1] carbon-based materials[2] and metal-organic frameworks (MOFs)[3]) have been actively investigated. Their distinctive properties allow them to find potential applications in many different areas such as catalysis, energy, nano-medicine, sensing, water treatment and electronics, among others. However, their use in biodisciplines, especially in biomolecule recognition (e.g., nucelobases), is still very limited[4-7] due to aspects related with inherent cytotoxicity, poor solubility-processability and stability in biological media. In this scenario, an emerging class of new porous materials, namely Covalent Organic Frameworks (COFs), seems to provide a plausible solution to many of these problems.[8] Indeed, some COFs have already shown affinity to host biomolecules such as enzymes,[9, 10] or sensing DNA.[11] In this respect, it is worth pointing out that nucleobases are among the most targeted biomolecules since they are the basic building blocks of nucleic acids and play key roles in biology.[12] However, while excellent recognition of nucleobases has been demonstrated in organic media, it is still a challenging task to use them in aqueous media because of significant interference from hydration of the binding site.[13]

[a.] *Departamento de Química Orgánica I, Facultad de CC. Químicas, Universidad Complutense de Madrid, Madrid 28040, Spain.*
[b.] *Departamento de Tecnología Química y Ambiental, Universidad Rey Juan Carlos, Madrid 28933, Spain*
[c.] *Instituto Madrileño de Estudios Avanzados en Nanociencia (IMDEA-Nanociencia), Cantoblanco, Madrid E-28049, Spain.*
[d.] *Departamento de Orgánica, Facultad de Ciencias, Universidad Autónoma de Madrid, Madrid 28049, Spain.*
[e.] *Departamento de Inorgánica, Facultad de Ciencias, Universidad Autónoma de Madrid, Madrid 28049, Spain.*
[f.] *Institute for Advanced Research in Chemical Sciences (IAdChem). Universidad Autónoma de Madrid. 28049 Madrid, Spain.*
[g.] *Condensed Matter Physics Center (IFIMAC). Universidad Autónoma de Madrid. 28049 Madrid, Spain.*

† Electronic Supplementary Information (ESI) is available providing experimental procedures and characterization data for all compounds, NMR and FT-IR spectra, TGA, powder XRD profiles and BET data. See DOI: 10.1039/x0xx00000x




**Scheme 1.** Representation of the synthesis of **[Uracil]$_{0.5}$-TPB-DMTP-COF**.

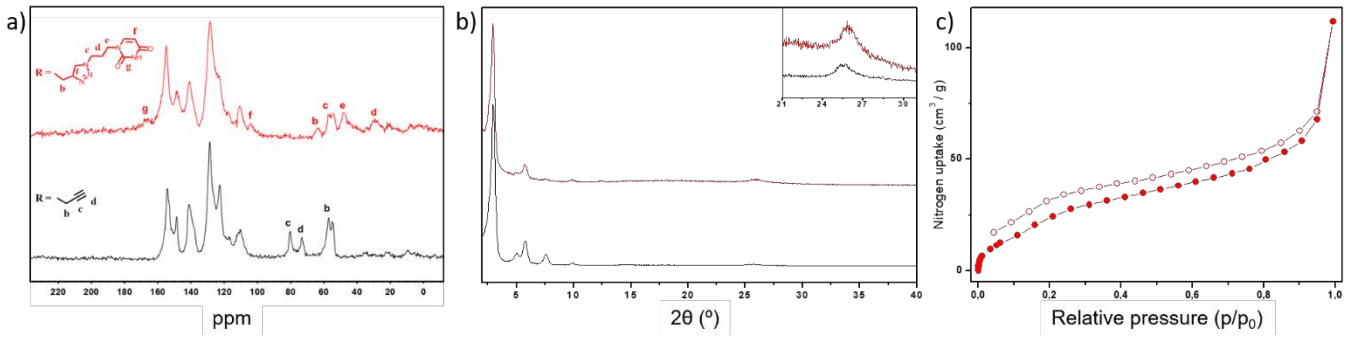



**Figure 1.** a) $^{13}$C CP/MAS NMR comparative spectra of **[HC≡C]$_{0.5}$-TPB-DMTP-COF** (black line) and **[Uracil]$_{0.5}$-TPB-DMTP-COF** (red line). b) Comparative PXRD patterns of **[HC≡C]$_{0.5}$-TPB-DMTP-COF** (black line) and **[Uracil]$_{0.5}$-TPB-DMTP-COF** (red line). c) Nitrogen adsorption isotherm of **[Uracil]$_{0.5}$-TPB-DMTP-COF**.

On the other hand, it is well-known that incorporation of desired molecules into the pore walls of porous materials may endow them with new characteristics thereby opening new avenues for applications. Indeed, solid-state porous materials have shown to be suitable platforms for DNA analysis.[14]

Thus, among porous materials, COFs show special advantageous features such as metal-free constitution, lower density and higher thermal and chemical stability, including water stability, in comparison with their closest congeners MOFs. Therefore, they seem to be excellent candidates for nucleobase recognition in aqueous media.

Hence, the integration of COFs with biomolecules, especially nucleobases, shows a highly promising value. In this regard, the use of "click chemistry", and more specifically the copper-catalyzed azide-alkyne cycloaddition reaction, has demonstrated to be a very efficient strategy to functionalize COFs and decorate their cavities with selected molecules.[15] Moreover, the additional control of the composition and density of the functional groups makes this strategy an attractive option for attaching nucleobases to COFs. In this work, we have prepared a nucleobase-functionalized COF, named as **[Uracil]$_{0.5}$-TPB-DMTP-COF** (Scheme 1), and evaluated its capability for selective nucleobase recognition.

The synthesis of **[HC≡C]$_{0.5}$-TPB-DMTP-COF** was accomplished by following the procedure previously described by Jiang and coworkers.[16] The subsequent characterization was performed by Fourier transform infrared spectroscopy (FTIR) and $^{13}$C cross-polarized magic angle spinning solid-state NMR ($^{13}$C CP/MAS NMR) spectroscopies. All the data are in accordance with those already reported (see ESI). **[Uracil]$_{0.5}$-TPB-DMTP-COF** was obtained by post-synthetic functionalization of **[HC≡C]$_{0.5}$-TPB-DMTP-COF** with 1-(3-azidopropyl)uracil (Scheme 1) by using the copper-catalysed Huisgen's 1,3-dipolar cycloaddition reaction between azides and terminal alkynes. FTIR (Figures S3 and S5) and $^{13}$C CP/MAS NMR (Figure 1a) confirm the successful incorporation of the nucleobase moiety in quantitative yield. FTIR clearly displays the carbonyl stretches associated to the CO groups of the uracil derivative as a broad signal at 1682 cm$^{-1}$. $^{13}$C CP/MAS NMR spectrum shows the complete disappearance of *C*sp signals at 80 and 73 ppm of the alkynyl fragment of **[HC≡C]$_{0.5}$-TPB-DMTP-COF**, together with the appearance of signals around 167 ppm assignable to the carbonyl and imine groups of the framework. New signals corresponding to the aliphatic core of the nucleobase are also encountered upfield.

Powder X-ray diffraction (PXRD) of all COFs exhibits a similar diffraction pattern to those previously reported adopting the AA stacking mode of a space group P6. (Figures S2 and S4 and Tables S1-S2). PXRD of **[Uracil]$_{0.5}$-TPB-DMTP-COF** also exhibits a similar profile to that of its precursor COF, indicating that the lattice remains almost unaltered after post-functionalization (Figure 1b).

The BET surface area, pore volume and pore size derived from nitrogen sorption isotherms are also in good agreement with the reported data for all of them (Table S6). The porosity of new **[Uracil]$_{0.5}$-TPB-DMTP-COF** was evaluated by its nitrogen sorption isotherm at 77 K revealing a type IV isotherm (Figure 1c), typical of mesoporous materials. The BET surface area and pore volume were estimated to be 105 m$^2$ g$^{-1}$ and 0.088 cm$^3$ g$^{-1}$, respectively. The pore size distribution calculated by using the non-local density functional theory (NLDFT) method resulted in a pore size of 1.3 nm.

Finally, the thermal stability of all the frameworks was determined by thermogravimetric analysis (TGA, Figures S10-S13). Although alkynyl functionalized **[HC≡C]$_{0.5}$-TPB-DMTP-COF** is stable up to 400 ºC stability decreases for its derivative, showing a 16 % weight loss at 360 ºC.

To evaluate the molecular recognition capability of the uracil group located at the cavity of the **[Uracil]$_{0.5}$-TPB-DMTP-COF**,

the material was incubated with adenine to study their interaction, since uracil and adenine form a canonical base pair in RNA duplexes. Such structure is promoted by the formation of two hydrogen bonds. During this process, the adenine was removed from the water solution by **[Uracil]$_{0.5}$-TPB-DMTP-COF**, leading to a significant reduction of this nucleobase compared to cytosine, uracil and thymine. (Figure 2). At the three concentrations tested (50, 100 and 150 μM) the same pattern was observed, but the result was much clearer at the highest concentration of nucleobase employed. Interestingly, when cytosine was used (Figure 2b), the nucleobase was not removed from the solution as in the previous case. This result suggests that the favourable interaction between uracil and adenine might be the main driving force in this process, which is quite remarkable since the stabilisation obtained by two hydrogen bonds in water is very weak.

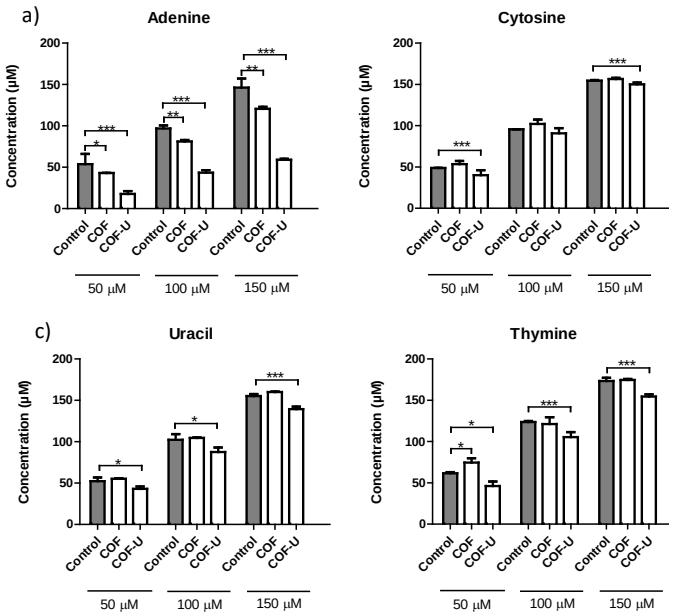



**Figure 2.** Concentration of nucleobases in the media after incubation at three different concentrations with **[Uracil]$_{0.5}$-TPB-DMTP-COF (COF-U)** and **TPB-DMTP-COF (COF)** for 22 hours. In the case of Adenine (a) significant changes in the concentration were observed, probably due to the favourable interaction uracil-adenine. In the case of cytosine (b) these changes were not observed. All the experiments were done in triplicates. Statistical analysis was performed using one-way ANOVA Tukey's test (each group vs Control). *P<0.01, **P<0.001, and ***P<0.0001.

When Uracil and Thymine were used similar results were obtained (Figure 2c,d). In this case, the slight reduction in the concentration observed might be due to the formation of non-canonical base pairs (U:U; U:T), as described in RNA structures.[17]

We also studied an additional modified COF since we wondered whether the chemical process employed in the modification of the COF was making the structure more susceptible to interact with adenine. Thus, we prepared a COF modified with a propyl group (**TPB-DMTP-COF-TAZ,** Scheme S4, Figures S8-S9)[18] instead of an uracil nucleobase (**[Uracil]$_{0.5}$-TPB-DMTP-COF**) using the same procedure. Therefore, the contribution of impurities or defects in the structure in the interaction with the nucleobases could be better assessed. The results obtained in this case (Figure S17) were similar to those obtained in the previous experiments (Figure 2) This outcome highlights the formation of the canonical base-pair inside the COF as the major driving force for the selective removal of adenine from the solution. This Watson-Crick pair shows a very weak interaction in water,[19] which might be enhanced by the confinement of the base-pair inside the COF. Such process resembles the one taking place inside polymerases[20, 21] or RNA complexes,[22] where the effect of the size and shape of the base pair inside the active pocket seems to be the main factor in the selectivity. In addition, the presence of multiple, closely spaced uracil units may induce the formation of H-bonded triplets[19] and/or stacking interactions between A-U pairs.

## Conclusions

The post-synthetic functionalization of COFs *via* the copper-catalyzed azide-alkyne cycloaddition reaction is an efficient method for the inclusion of biomolecules into the cavities of imine-based COFs. For complementary nucleobase recognition, the incorporation of the uracil nucleobase within the pores of the COF promotes the confinement of the base-pair inside the COF which allows a remarkable selective recognition in aqueous media. These results pave the way for the further investigation of appropriately functionalized COFs for recognition or delivering of oligonucleotides.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgments

This work was financially supported by MINECO (MAT2016-77608-C3-1-P and 2-P, SAF2017-87305-R). IMDEA Nanociencia acknowledges support from the 'Severo Ochoa' Programme for Centres of Excellence in R&D (MINECO, Grant SEV-2016-0686). Funding from the European Research Council (ERC-StG 279548) and MINECO (CTQ2014-27729-P and CTQ2017-84727-P) is gratefully acknowledged (DGR).

## Notes and references

‡ ESI include experimental procedures and characterization data for all compounds. Synthesis, NMR and FT-IR spectra, UV, TGA, BET isotherms and powder XRD profiles.

1. T. Ennaert, J. Van Aelst, J. Dijkmans, R. De Clercq, W. Schutyser, M. Dusselier, D. Verboekend and B. F. Sels, *Chem. Soc. Rev.*, 2016, **45**, 584-611.
2. L. Borchardt, Q.-L. Zhu, M. E. Casco, R. Berger, X. Zhuang, S. Kaskel, X. Feng and Q. Xu, *Mater. Today*, 2017, **20**, 592-610.
3. T. Faust, *Nat. Chem.*, 2016, **8**, 990-991.

4. L. Qin, L.-X. Lin, Z.-P. Fang, S.-P. Yang, G.-H. Qiu, J.-X. Chen and W.-H. Chen, *Chem. Commun.*, 2016, **52**, 132-135.
5. Y. Wu, J. Han, P. Xue, R. Xu and Y. Kang, *Nanoscale*, 2015, **7**, 1753-1759.
6. X. Luo, T. Shen, L. Ding, W. Zhong, J. Luo and S. Luo, *J. Hazard. Mater.*, 2016, **306**, 313-322.
7. M. Sarker, J. Y. Song, A. R. Jeong, K. S. Min and S. H. Jhung, *J. Hazard. Mater.*, 2018, **344**, 593-601.
8. P. J. Waller, F. Gándara and O. M. Yaghi, *Acc. Chem. Res.*, 2015, **48**, 3053-3063.
9. Q. Sun, C.-W. Fu, B. Aguila, J. Perman, S. Wang, H.-Y. Huang, F.-S. Xiao and S. Ma, *J. Am. Chem. Soc.*, 2018, **140**, 984-992.
10. S. Kandambeth, V. Venkatesh, D. B. Shinde, S. Kumari, A. Halder, S. Verma and R. Banerjee, *Nat. Commun.*, 2015, **6**, 6786.
11. W. Li, C.-X. Yang and X.-P. Yan, *Chem. Commun.*, 2017, **53**, 11469-11471.
12. C. Schmuck and H. Wennemers, *Highlights in Bioorganic Chemistry: Methods and Applications*, Wiley-VCH, Weinheim, 2002.
13. K. Yoshimoto, S. Nishizawa, M. Minagawa and N. Teramae, *J. Am. Chem. Soc.*, 2003, **125**, 8982-8983.
14. C. Chen, Y. Li, S. Kerman, P. Neutens, K. Willems, S. Cornelissen, L. Lagae, T. Stakenborg and P. Van Dorpe, *Nat. Commun.*, 2018, **9**, 1733.
15. A. Nagai, Z. Guo, X. Feng, S. Jin, X. Chen, X. Ding and D. Jiang, *Nat. Commun.*, 2011, **2**, 536.
16. H. Xu, J. Gao and D. Jiang, *Nat. Chem.*, 2015, **7**, 905.
17. K. J. Baeyens, H. L. De Bondt and S. R. Holbrook, *Nat. Struct. Biol.*, 1995, **2**, 56.
18. L. Meri-Bofi, S. Royuela, F. Zamora, M. L. Ruiz-Gonzalez, J. L. Segura, R. Munoz-Olivas and M. J. Mancheno, *J. Mater. Chem. A*, 2017, **5**, 17973-17981.
19. M. J. Mayoral, C. Montoro-García and D. González-Rodríguez, in *Comprehensive Supramolecular Chemistry II*, ed. J. Atwood, Elsevier, Oxford, 2017, vol. 4, pp. 191-257.
20. E. T. Kool, *Ann. Rev. Biophysics Biomol. Struct.*, 2001, **30**, 1-22.
21. J. C. Delaney, P. T. Henderson, S. A. Helquist, J. C. Morales, J. M. Essigmann and E. T. Kool, *Proc. Nat. Ac. Sci.*, 2003, **100**, 4469.
22. A. Somoza, J. Chelliserrykattil and T. Kool Eric, *Angew. Chem. Int. Ed.*, 2006, **45**, 4994-4997.