CrystEngComm

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/crystengcomm

Exploring the influence of steric hindrance and electronic nature of the substituents in the supramolecular arrangements of 5-(substituted phenyl)-2-formylpyrroles†

Cláudia A. Figueira, Patrícia S. Lopes, Clara S. B. Gomes, Luís F. Veiros and Pedro T. Gomes*

Centro de Química Estrutural, Departamento de Engenharia Química, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal

Phone: (+351)218419612; Fax: (+351)218417246; E-mail: pedro.t.gomes@tecnico.ulisboa.pt

Abstract

A family of 5-(substituted phenyl)-2-formylpyrrole compounds, exhibiting different electronic and steric features at the phenyl ring, was synthesised through the formylation reaction of the corresponding 2-(substituted phenyl)pyrrole precursors, using Vilsmeier-Haack acylation conditions. The products were obtained in moderate to high yields, being systematically characterised by NMR spectroscopy, elemental analysis and single crystal X-ray diffraction. The corresponding crystalline packings were discussed on the basis of three types of arrangements, leading to the formation of dimers, polymers or the newly observed tetramers, all of them essentially dictated by strong N-H···O hydrogen bonding interactions. Important C-H···O, C-H···π, π···π and, in the case of fluorinated compounds, C-H···F interactions also contributed to the growth of the three-dimensional crystalline network. DFT calculations helped to rationalise the relationship between the steric and electronic properties of the molecules and the basic units observed in the corresponding solid state structures.

Introduction

The 2-formylpyrrole is one of the simplest α-functionalised pyrrole derivatives, having received in the last ten years a considerable attention from several research areas. The presence of a reactive carbonyl group as a substituent of one of the α-carbon atoms of the pyrrole ring makes this compound an important intermediate in the formation of higher molecules. The 2-formylpyrrole has been mostly used in annulation reactions to

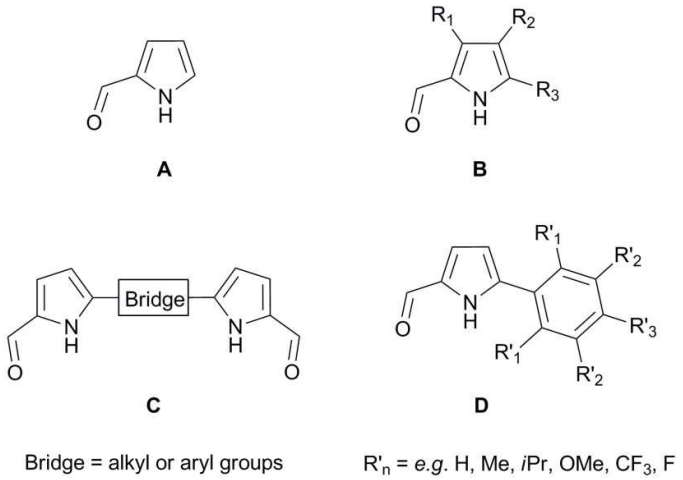
† Electronic supplementary information (ESI) available: Details of data collection and structure refinement for compounds **1**-**7**; selected bond lengths (Å) and angles (º) for compounds **1**-**7**; and atomic coordinates of the structures optimised by DFT. CCDC 1064026–1064032. For ESI and crystallographic data in CIF or other electronic format see DOI: …

obtain fused rings, giving rise to an extensive diversity of new nitrogen heterocyclic compounds.[1] It is also a precursor to the synthesis of boron-dipyrrin (BODIPY) complex dyes, with applications in biomedicine and molecular biology.[2] However, it is in the field of coordination chemistry that 2-formylpyrrole finds its major applicability, either as a ligand precursor[3–6] or as starting material for other ligand precursors,[7–9] such as the 2-iminopyrrole Schiff bases.[10–15]

In particular, the α-aryl substituted 2-formylpyrrole derivatives (5-aryl-2-formylpyrroles) have received scarce attention, so far being only employed in annulation reactions,[1d] in medicinal chemistry as precursors for the synthesis of HIV-1 entry inhibitors[16] and as intermediates for the synthesis of BODIPY derivatives[2a] and organic dyes with pyrrolic conjugated π spacers.[17] However, only few examples in the literature reported their synthesis, and often just for the simplest aryl derivative, the 5-phenyl-2-formylpyrrole.[18] The more general synthetic methods described for 5-aryl-2-formylpyrroles included two-step procedures,[19–21] with the most efficient method involving a one-step catalytic regioselective C5 direct arylation of *N*-substituted 2-formylpyrroles with aryl bromides.[22] Nevertheless, the final products are usually *N*-substituted 5-aryl-2-formylpyrroles, and not the corresponding *NH*-analogues.

In contrast with 2-iminopyrroles,[23] studies on the crystalline supramolecular arrangements of 2-formylpyrrole (Scheme 1, **A**) or its derivatives are very rare in the literature. Nevertheless, a search in the Cambridge Crystallographic Data Centre (CCDC) displayed a number of results involving the 2-formylpyrrole core. The crystalline structure of the simple 2-formylpyrrole is known,[4b] along with some other 2-formylpyrroles derivatives substituted in positions 3, 4 and/or 5 of the pyrrole ring (Scheme 1, **B**). In the majority of the cases, regardless of the type and number of substituents, the crystalline networks exhibit the formation of dimers, via complementary N-H···O hydrogen-bonding between the pyrrole N-H group and the formyl oxygen atom of a neighbouring molecule.[4b,24,25] Nevertheless, in some cases there are exceptions, such as in the crystalline network of the 2-formylpyrrole, formed by zigzagged polymer chains through intermolecular N-H···O interactions,[4b] a behaviour also observed for the 2,5-diformylpyrrole,[26] and for a few other cases when the 2-formylpyrrole is substituted in positions 3 and/or 4 of the pyrrole ring.[24a,27,28] Similar zigzagged polymer chains are formed, though through N-H···N hydrogen-bonding interactions, in the crystalline structure of 4-(4-pyridyl)-2-formylpyrrole.[29]

Bis(2-formylpyrrole) derivatives, with the two moieties linked by the remaining α-positions through alkyl- or aryl-bridges (Scheme 1, **C**), also self-assemble via complementary N-H···O hydrogen-bonding with one or two neighbouring molecules, leading respectively to dimerisation or polymerisation.[30]



Scheme 1 Schematic representations of 2-formylpyrrole (**A**) and of different examples of its derivatives (**B**, **C**), including the 5-(substituted phenyl)-2-formyl-1*H*-pyrroles of the present work (**D**).

While planning the preparation of a new set of 5-(substituted phenyl)-2-formyl-1*H*-pyrroles, in which the phenyl group has different electronic and steric features (Scheme 1, **D**), we performed a specific search for this type of compounds in the CCDC but no results were retrieved. The closest examples reported in the literature included the 3-(morpholin-4-yl)-5-phenyl-2-formylpyrrole[31] or, considering any kind of substituent at the position 5 of the pyrrole ring, the compounds 5-((2-aminophenyl)iminomethyl)-2-formylpyrrole[32] and the 5-(pyrrol-2-ylmethyl)-2-formylpyrrole,[33] all presenting the typical dimerisation pattern.

In the present work, we report the synthesis and characterisation of a family of seven 5-(substituted phenyl)-2-formyl-1*H*-pyrroles via the α-formylation of the corresponding 2-(substituted phenyl)-1*H*-pyrroles, among which five are new compounds. All of them were analysed by elemental analysis, solution NMR spectroscopy and single-crystal X-ray diffraction. A rationalisation of the different patterns of the supramolecular arrangements found in the X-ray crystalline structures is attempted on the basis of the influence of the steric and electronic nature of the phenyl

ring substituents upon the intermolecular interactions, this study being complemented by DFT calculations.[34] The present work is, to the best of our knowledge, the first report on the solid state structures of 5-(substituted phenyl)-2-formyl-1*H*-pyrroles.

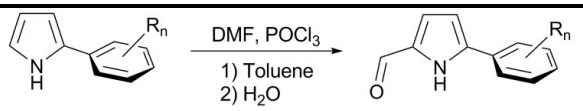
Results and Discussion

The 5-(substituted phenyl)-2-formyl-1*H*-pyrroles studied were prepared through the formylation reaction of the previously reported 2-(substituted phenyl)-1*H*-pyrroles[35,36] using the classical Vilsmeier-Haack electrophilic substitution method.[2,17,37] The *in situ* prepared Vilsmeier reagent reacted with the 2-(substituted phenyl)pyrroles *via* an electrophilic aromatic substitution reaction and the iminium salt obtained was then hydrolysed, compounds **1**-**7** (Table 1) being obtained as crystalline solids in moderate to high yields, from which **2**-**5** and **7** are new molecules.

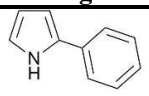
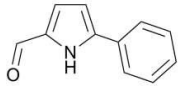
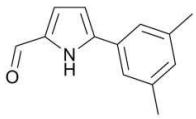
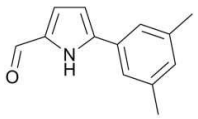
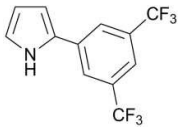
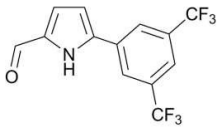
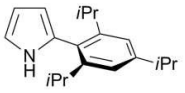
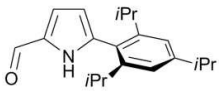
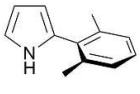
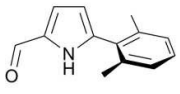
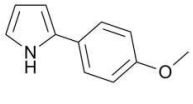
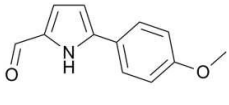
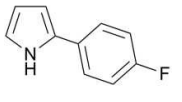
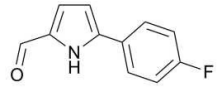
The crystalline and molecular structures of compounds **1**-**7** were determined by Single Crystal X-ray Diffraction, crystallographic and experimental details of crystal structure determinations being listed in Tables S1 and S2 in Supporting Information. The obtained molecular structures are depicted in Fig. 1.

The bond lengths and angles are similar for all compounds (see Table S3 in Supporting Information), with the C2-C6 bond slightly shorter than C5-C7 (≈1.42-1.43 *vs*. 1.46-1.49 Å) due to the higher extension of π-conjugation of the carbonyl group with the pyrrole ring. For compounds **4**-**6**, it was also observed that their molecular structures are very similar to those of the parent precursors.[35b] The plane containing the carbonyl bond is coplanar with that of the pyrrole ring, showing a lowest dihedral angle of 2.196(522)º for **2**, the highest being 5.797(228)º for $\mathbf{5_B}$ (the 2 molecules found in the asymmetric unit of **5** are designated as $\mathbf{5_A}$ and $\mathbf{5_B}$). With exception of **4** and **5**, all structures have planar geometries, with small dihedral angles between the phenyl and pyrrole rings of 11.15(6)º for **1**, 6.9(2)º for **2**, 7.51(13)º for **3**, 4.84(16)º for **6**, and 6.2(2)º for **7**. For compounds **4** and **5**, the phenyl rings are almost perpendicular to the pyrrole ones, with dihedral angles of 86.45(10)º for **4**, and 67.67(12)º and 74.63(12)º for **5** (molecules A and B), respectively.

Table 1. Formylation reactions of 2-(substituted phenyl)pyrroles **I-VII** to give 5-(substituted phenyl)-2-formylpyrroles **1-7**, using Vilsmeier-Haack acylation conditions.



I - VII 1 - 7

2-(substituted phenyl) pyrrole reagent	5-(substituted phenyl)-2-formylpyrrole product	Yield (%)
		96
I	**1**	
		81
II	**2**	
		65
III	**3**	
		75
IV	**4**	
		62
V	**5**	
		71
VI	**6**	
		64
VII	**7**	

Conditions: Vilsmeier reagent (excess); solvent: toluene; T = 0 ºC to r.t. (overnight)

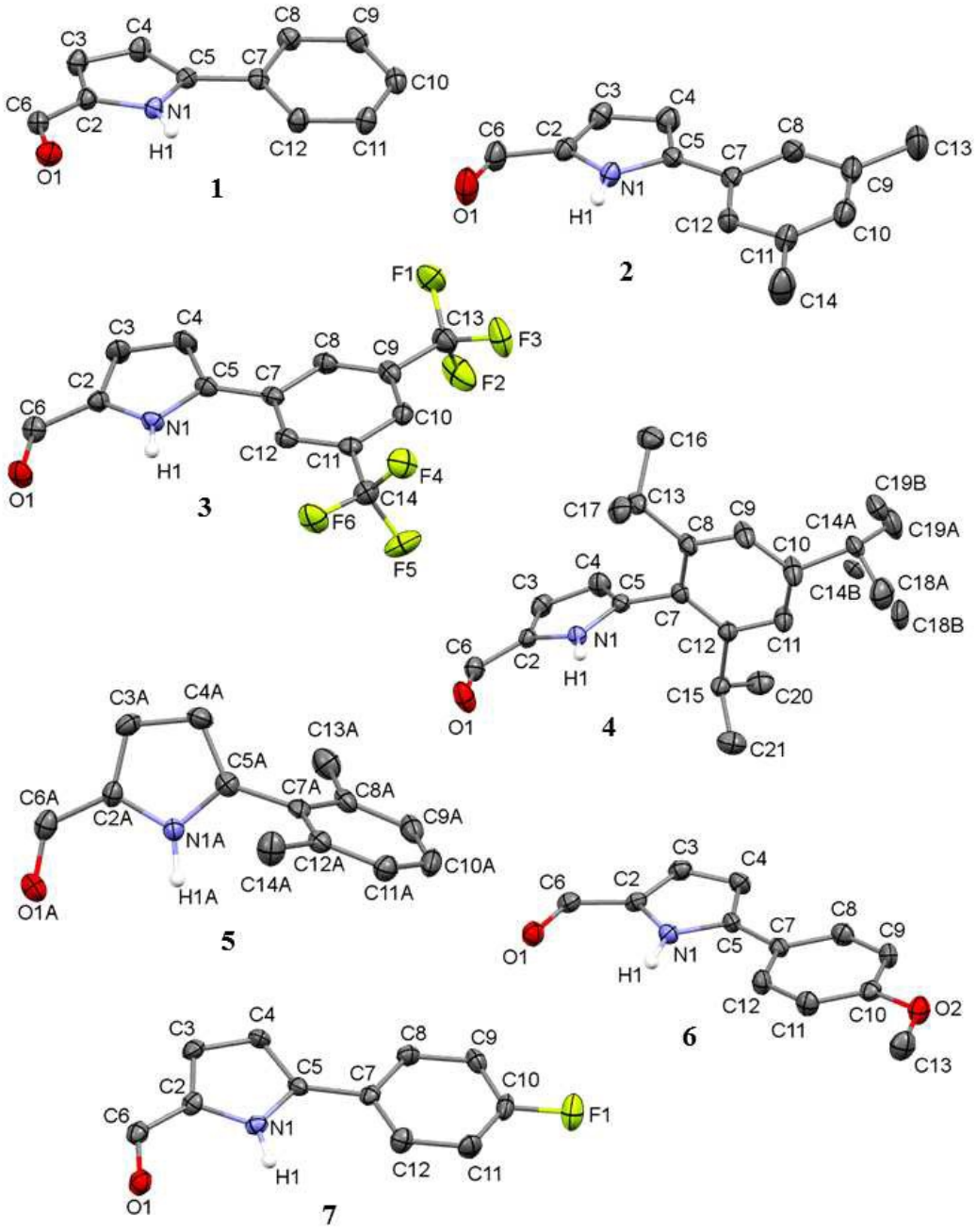


Fig. 1 Molecular structures of 5-(substituted phenyl)-2-formylpyrroles **1**-**7**, showing the atom-numbering scheme. H atoms, with exception of the N*H*, were omitted for clarity. Compound **4** has a disordered isopropyl substituent over two positions at the *para*-position of the phenyl ring, A and B representing both positions. Compound **5** contains two molecules in the asymmetric unit, the only difference between them being the value of the dihedral angle between the pyrrole and the 2,6-dimethylphenyl rings. Due to their high similarity, only molecule A is represented.

In the crystalline structures of all the compounds **1**-**7**, the most relevant intermolecular interaction is the N-H···O hydrogen-bond between the pyrrole N-H group (a strong H-bond donor) and the formyl oxygen atom (a strong H-bond acceptor) of a neighbouring molecule. In compound **1**, 5-phenyl-2-formylpyrrole, the N-H···O interaction occurs with a complementary counterpart of a neighbouring molecule generated by symmetry.[38] This leads to the formation of coplanar dimers (dihedral angle of 0.00º), which are obtained through a $R^2_2(10)$ synthon (Fig. 2, compound **1**).

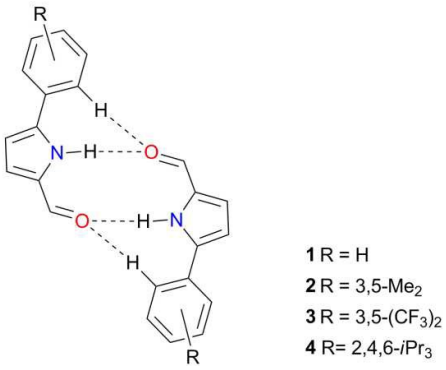


Fig. 2 Synthon $R^2_2(10)$ displayed by compounds **1**-**4**.

The assisted non-classical hydrogen-bond C-H···O, which shares the same oxygen atom of the previous interaction, reinforces the strength/cohesion between dimers. The network grows along the *ab* plane as an alternate zigzag dimer chain (angle of 79.74º) through a C_{10}-H_{10}···π $C_{g(Ar)}$ interaction between the phenyl rings of neighbouring dimers. This chain is repeatedly stacked through π···π interactions of 3.780 Å, between the phenyl and pyrrole moieties, due to the extended resonance of the molecules (Fig. 3).

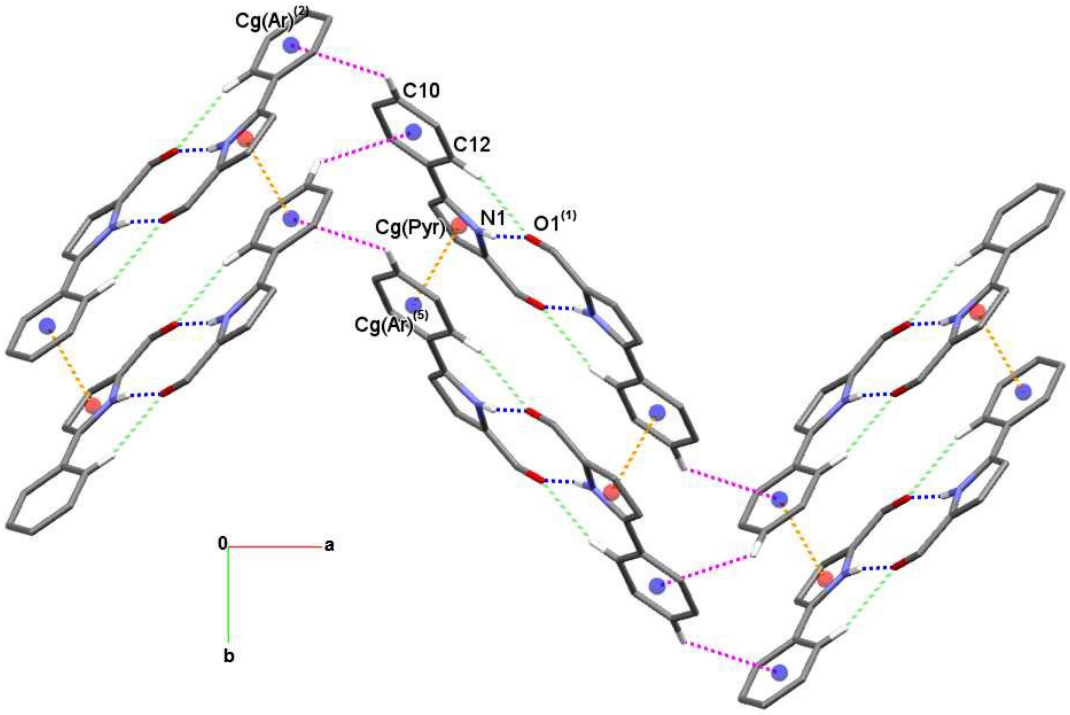


Fig. 3 View of the 3D packing of compound **1**, showing the one-dimensional zigzag chains in the (110) plane – *a* direction. Donor and acceptor atoms are identified. Blue dashed lines represent N_1-$H_1\cdots O_1$ hydrogen-bonds, light-green dashed lines the C_{12}-$H_{12}\cdots O_1$ assisted non-classical hydrogen-bonds, pink dashed lines the C_{10}-$H_{10}\cdots\pi$ $C_{g(Ar)}$ interactions and the orange dashed lines the $\pi\cdots\pi$ interactions between the pyrrole and phenyl centroids ($C_{g(Pyr)}$ and $C_{g(Ar)}$).

In compound **2** the dimerisation occurs through the formation of the same $R^2_2(10)$ synthon (Fig. 2, compound **2**) and the complementary C-H···O interaction, as observed for compound **1**, even though the packing is different (Fig. 4(a)). With the presence of methyl substituents in the *meta* positions of the phenyl ring, the C-H···π interaction no longer occurs between 3,5-dimethylphenyl rings due to the steric hindrance imposed by those substituents. However, the establishment of an additional interaction between the carbonyl CH group and the 3,5-dimethylphenyl ring leads to alternating dimer chains along the plane (011) – *b* direction with an angle of 61.18º (Fig. 4(b)).

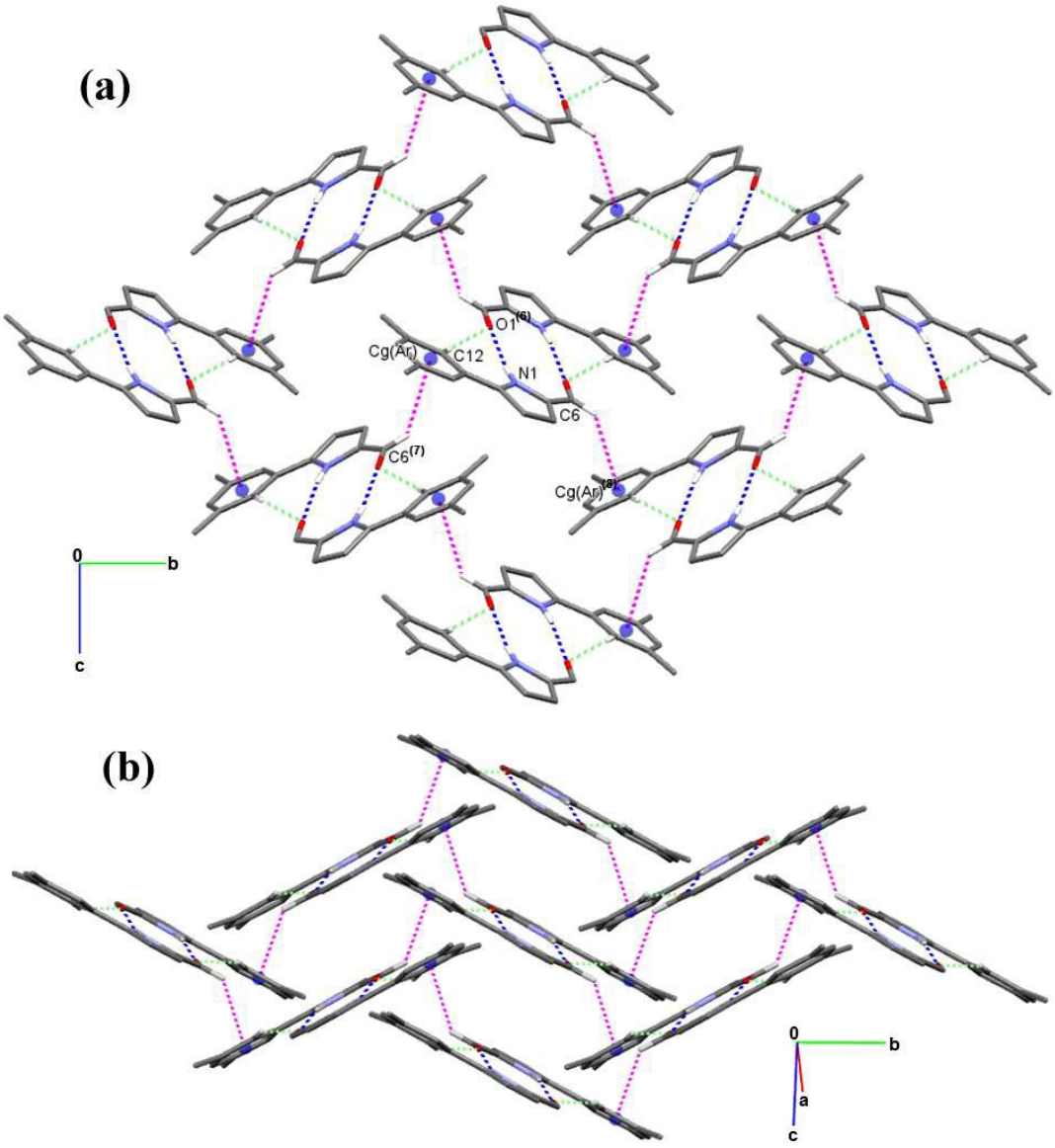


Fig. 4 Views of the 2D packing of compound **2** (a) in the plane (011) and (b) highlight of the alternate dimer chains. Donor and acceptor atoms are identified. Blue dashed lines represent the N_1-$H_1\cdots O_1$ hydrogen-bonds, light-green dashed lines the C_{12}-$H_{12}\cdots O_1$ assisted non-classical hydrogen-bonds and the pink dashed lines the C_6-$H_6\cdots\pi$ $C_{g(Ar)}$ interactions.

Furthermore, the occurrence of $\pi\cdots\pi$ interactions between 3,5-dimethylphenyl and pyrrole moieties, within a 4.055 Å distance, completes the 3D packing of compound **2** (Fig. 5).

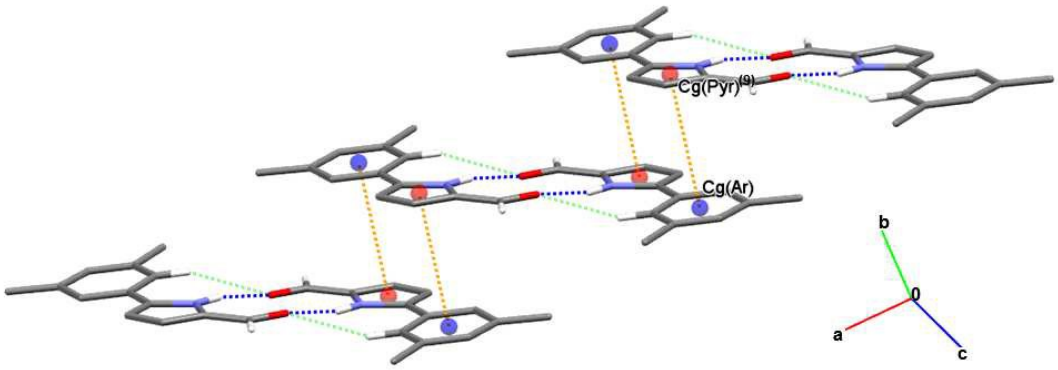


Fig. 5 Detail of the π-stacking displayed in each dimer of the crystalline structure of compound **2**. The orange dashed lines represent the π···π interactions between the pyrrole and 3,5-dimethylphenyl centroids ($C_{g(Pyr)}$ and $C_{g(Ar)}$).

In compound **3**, with the replacement of the methyl substituents by trifluoromethyl groups a new $R^2_2(16)$ synthon was also identified besides the previously observed coplanar dimer that forms a $R^2_2(10)$ synthon (Fig. 2, compound **3**). In the $R^2_2(10)$ synthon, the complementary N-H···O interaction is again reinforced with the C-H···O interaction. The new $R^2_2(16)$ synthon is formed due to a paired C-H···F interaction (Fig. 6).

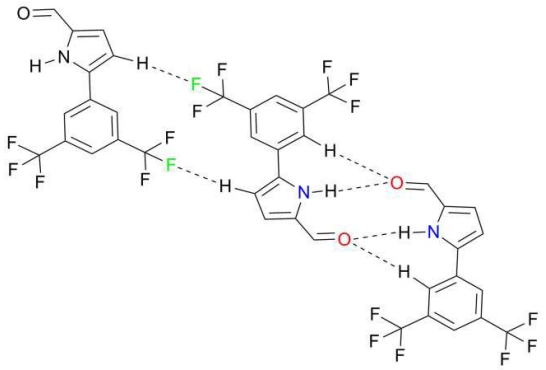


Fig. 6 Typical $R^2_2(10)$ synthon, formed by complementary N-H···O hydrogen-bond interactions, and $R^2_2(16)$ synthon, formed by paired C-H···F interactions, in compound **3**.

Since both synthons are consecutive, an infinite chain is formed along *a* (Fig. S1 in the Supporting Information). Another C-H···F interaction (C_3-H_3···F_5) occurs consecutively along *c*, generating a second infinite chain. The dihedral angle between the pyrrole and 3,5-bis(trifluoromethyl)phenyl ring planes of interrelated molecules is of 88.52º. This second chain crosses the previous one, which contains the $R^2_2(10)$ and

$R^2_2(16)$ synthons, leading to the 2D growth in the plane (101) (Fig. 7(a)). In addition, π-stacking, which arises owing to C-O···π $C_{g(Pyr)}$ interactions is observed when viewed along *b* (Fig. 7(b)).

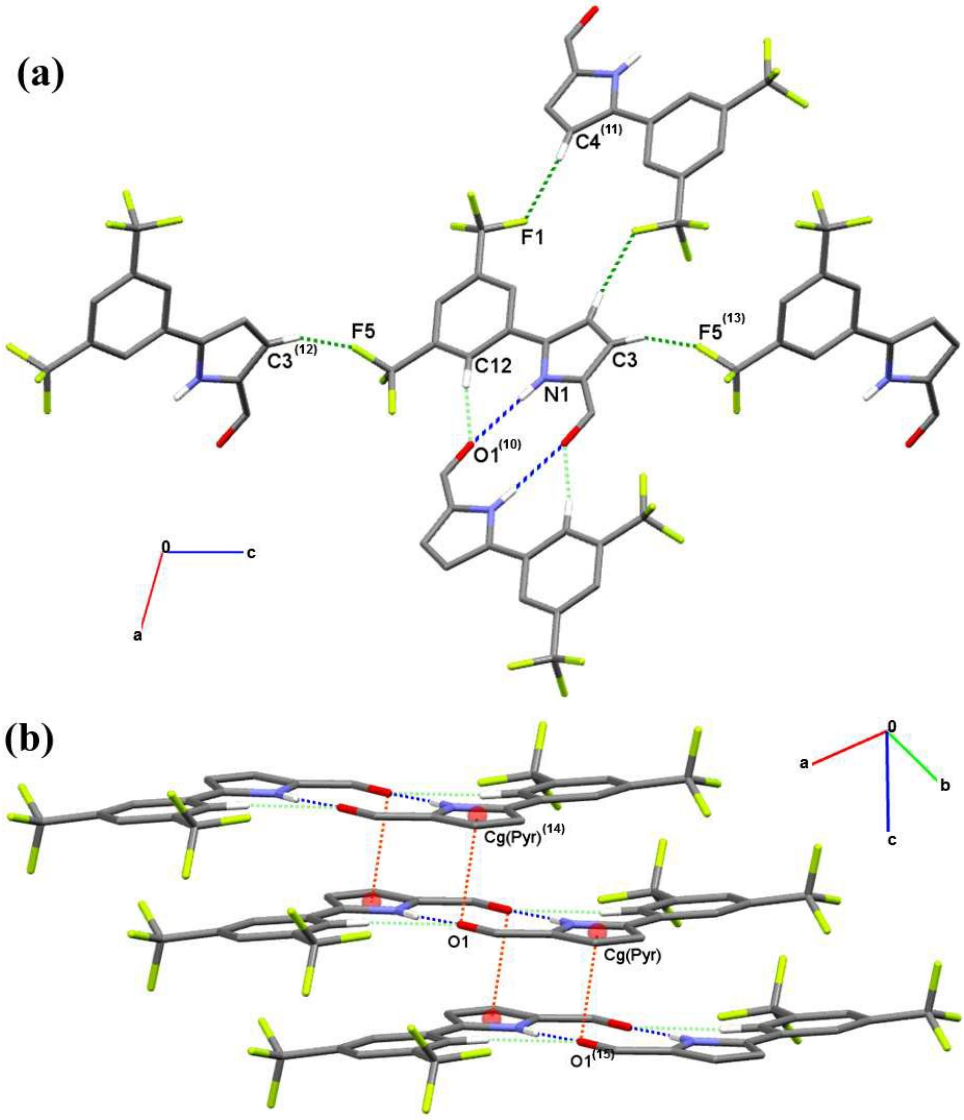


Fig. 7 (a) 2D packing in the crystalline structure of compound **3**, viewed along *b*, formed around the asymmetric unit; (b) π-stacking observed between the carbonyl group and the pyrrole moieties. Blue dashed lines represent the N_1-H_1···O_1 hydrogen-bonds, light-green and green dashed lines representing, respectively, the assisted non-classical hydrogen-bonds C_{12}-H_{12}···O_1 and both C_4-H_4···F_1 and C_3-H_3···F_5.

Conversely, compounds **4** and **5** bearing, respectively, isopropyl or methyl substituents in the *ortho*-positions of the phenyl ring, show the phenyl and pyrrole ring planes in an almost perpendicular fashion. In compound **4**, the $R^2_2(10)$ synthon is still identified

(Fig. 2, compound **4**). Nevertheless, in this case, the N_1-$H_1\cdots O_1$ is not accompanied by the C_{12}-$H_{12}\cdots O_1$ interaction due to the presence of the *ortho*-substituents in the phenyl ring. Even with the increased bulkiness introduced by the isopropyl substituents, the dimer is again coplanar, with both rings of the dimer perfectly parallel to each other. Both *ortho*-substituents show a C-H$\cdots\pi$ C_g interaction, one with the pyrrole ring of a neighbouring molecule and the other with the 2,4,6-tris(isopropyl)phenyl ring of another neighbouring one. On one hand, the C_{16}-$H_{16C}\cdots\pi$ $C_{g(Ar)}$ interaction, complementary with the counterpart of a neighbouring molecule, allows the stacking of dimers along the *c* axis. On the other hand, the also complementary interaction C_{21}-$H_{21B}\cdots\pi$ $C_{g(Pyr)}$, allows the stacking of dimers along the *b* axis. The combination of these three complementary interactions gave rise to the formation of small cavities in the 3D packing, owing to the interrelation between two dimers and two half dimers, in a total of 12 interactions (Fig. 8).

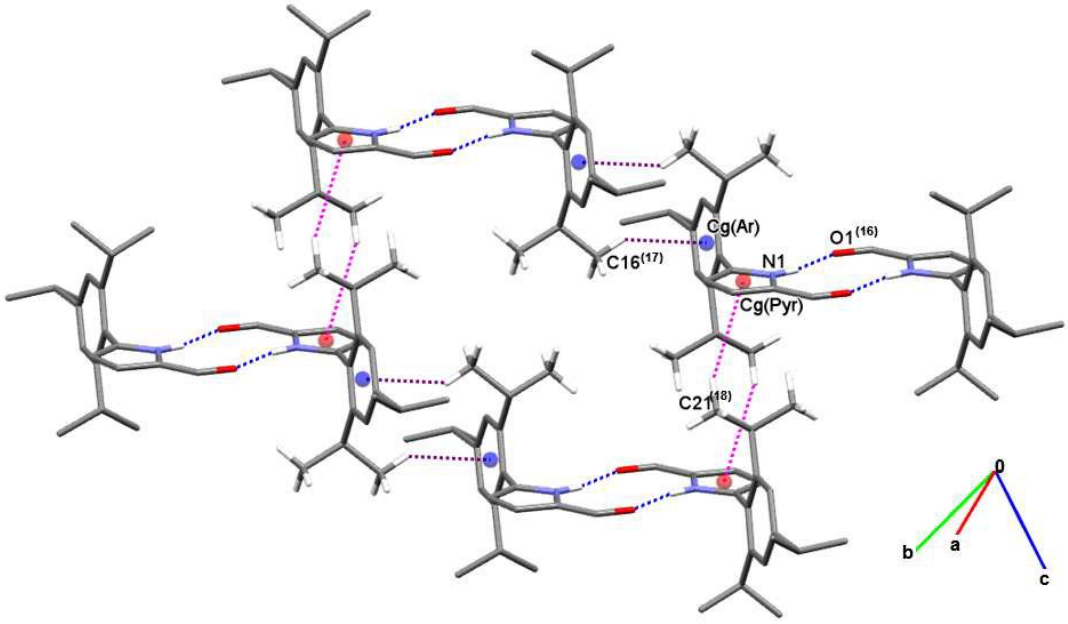


Fig. 8 Detail of a cavity in the 3D packing of compound **4**, due to the combination of the three different complementary interactions: Blue dashed lines represent the N_1-$H_1\cdots O_1$ hydrogen-bonds, purple dashed lines the C_{16}-$H_{16C}\cdots\pi$ $C_{g(Ar)}$ and pink dashed lines the C_{21}-$H_{21B}\cdots\pi$ $C_{g(Pyr)}$.

In turn, compound **5** showed a very distinctive supramolecular arrangement. In this case, the asymmetric unit is composed by two independent molecules, labelled as A and B, respectively. Each molecule A interacts with two neighbouring molecules B *via* two

N-H···O hydrogen-bonds, as donor and acceptor, and vice-versa. The self-organization of these molecules gives rise to nearly square planar tetramers, with dihedral angles between the planes of the pyrrole moieties of molecules A and B of 20.50º, corresponding to the formation of a $R^4_4(20)$ synthon (Fig. 9(a)). The high proximity and the perpendicularity between rings reinforces the formation of the tetramer by a C-H···π $C_{g(Ar)}$ interaction between the carbonyl carbon of one molecule and the centroid of the 2,6-dimethylphenyl ring of the neighbouring molecule (Fig. 9(b)).

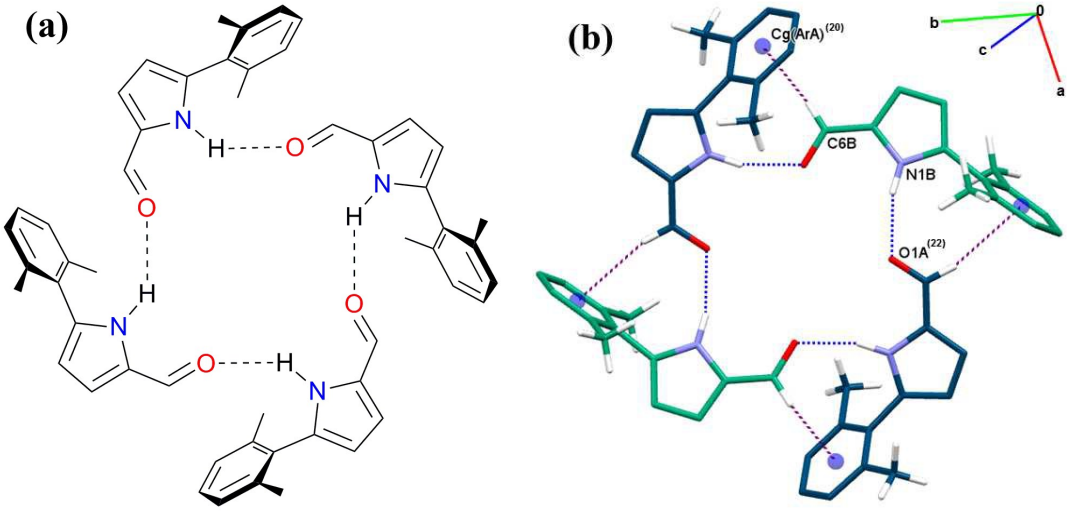


Fig. 9 (a) Synthon $R^4_4(20)$ characteristic of compound **5** and (b) evidence of the N-H···O (blue dashed lines) and the C-H···π $C_{g(Ar)}$ (purple dashed lines) interactions in the tetramer (Molecules A are represented in blue and B in green).

When analysed as separated entities, each molecule A relates with 6 different B molecules and vice-versa, including the ones of the tetramer (Fig. 10). Each tetramer relates with 8 other neighbouring tetramers (Fig. 11), involving a total of 14 molecules, directly interacting with the tetramer through different types of C-H···π and of C-H···O interactions. The C-H···π $C_{g(Pyr)}$ interactions between the *para*-carbon of the phenyl ring of molecules B and the pyrrole ring of molecules A and the C-H···π $C_{g(Ar)}$ between one methyl substituent of B and the phenyl ring of A enables the growing of the network along the plane (011) (Fig. 10(a)). In addition, two C-H···O interactions between one methyl substituent of molecules A and two oxygen atoms of two molecules B, belonging to the same neighbour tetramer, Fig. 10(b), gives rise to the extension of the network along the plane (101).

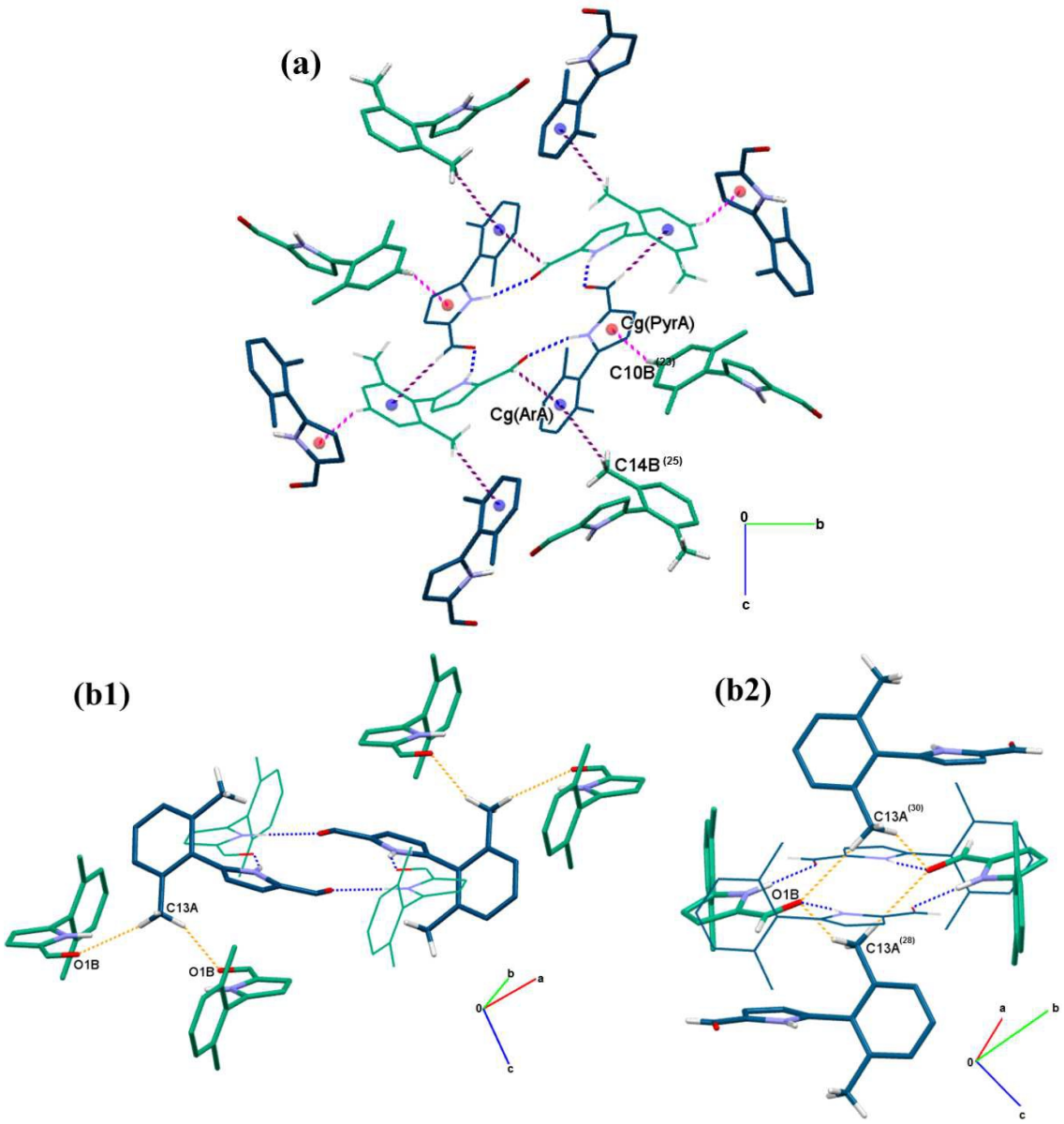


Fig. 10 Views of the overall interactions of each tetramer with the neighbouring molecules in the crystalline structure of **5**: (a) View of the C-H···π interactions of the tetramer along plane (011); the C-H···π $C_{g(Ar)}$ are represented as purple dashed lines and the C-H···π $C_{g(Pyr)}$ as pink dashed lines. (b) View of the C-H···O interactions (orange dashed lines) in (b1) molecules A and (b2) molecules B of the same tetramer.

With exception of the N-H···O and C-H···π $C_{g(Ar)}$ interactions inside the tetramer, which are differently oriented, the remaining interactions that form the intricate network among the tetramers tend to be oriented along one of the matrix axes: the C_{10B}-H_{10B}···π $C_{g(PyrA)}$ occurs along *b*, the C_{14B}-H_{14E}···π $C_{g(ArA)}$ along *c* and the C_{13A}-H_{13A}···O_{1B} and C_{13A}-H_{13B}···O_{1B} along *a* (Fig. 11).

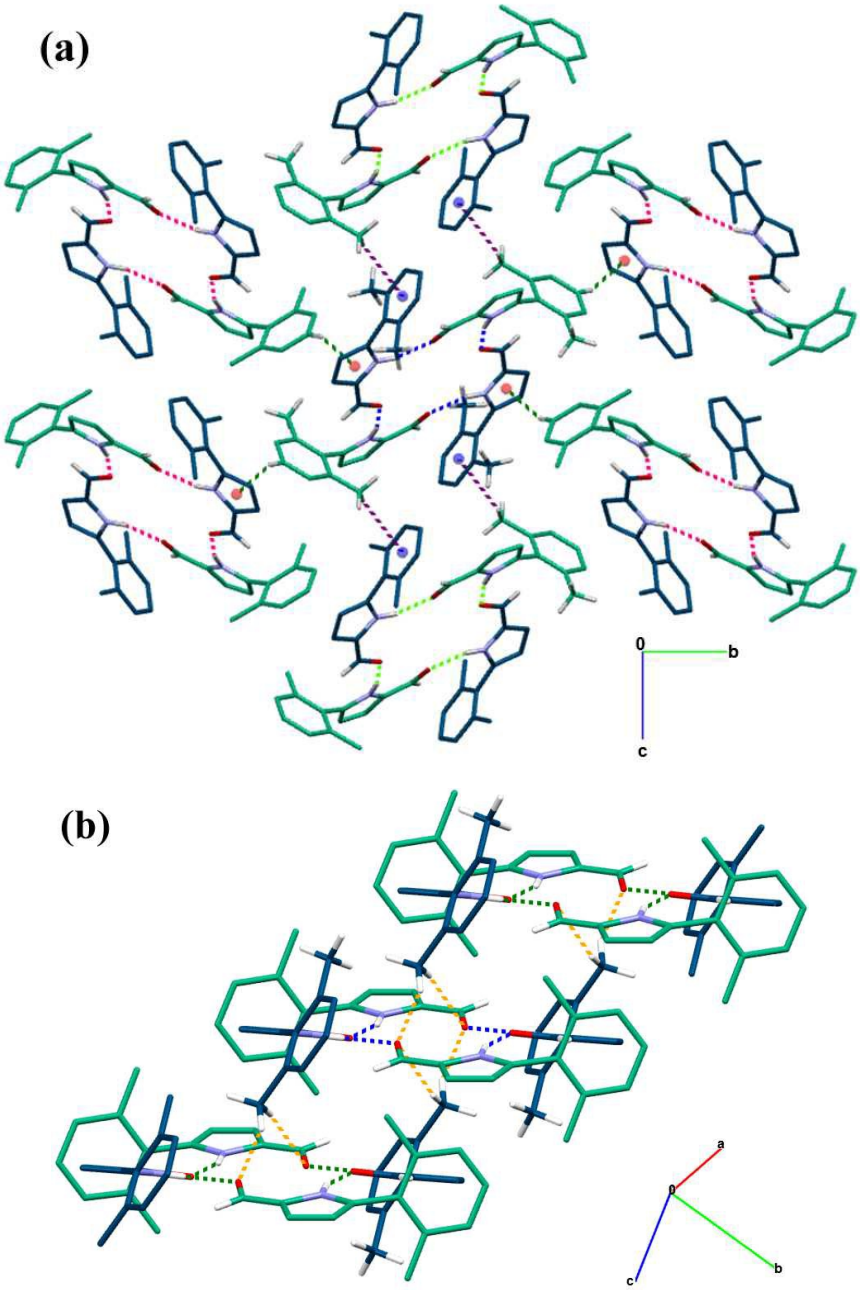


Fig. 11 Views of the overall interactions of each tetramer with the neighbouring tetramers in the crystalline structure of **5**: (a) View of the packing along plane (011); the C-H···π $C_{g(Ar)}$ interactions around the central tetramer are represented in purple dashed lines and the C-H···π $C_{g(Pyr)}$ in green dashed lines; the surrounding tetramers are represented in different colours (N-H···O interactions in pink and light green dashed lines), according to the representative interaction (C-H···π $C_{g(Pyr)}$ and C-H···π $C_{g(Ar)}$, respectively) to the central one (blue dashed lines). (b) View of the packing along *a* axis; the C-H···O interactions around the central tetramer are represented in orange dashed lines; the N-H···O interactions of the surrounding tetramers are presented in dark green dashed lines.

The introduction of *para*-phenyl substituents with strong electron donor or acceptor properties, such as in compounds **6** and **7**, respectively, do not lead to the typical dimerisation observed for compounds **1**-**4**. Instead, for **6**, the molecule present in the asymmetric unit interacts with eight different analogues (Fig. 12).

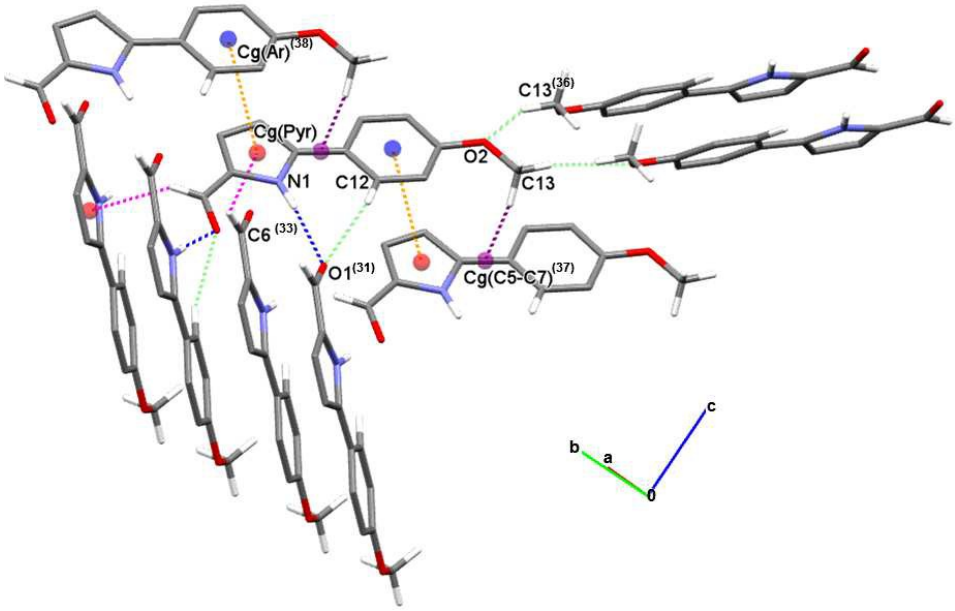


Fig. 12 Supramolecular arrangement of compound **6**, showing the N_1-H_1···O_1 interactions in blue dashed lines and both the C_{12}-H_{12}···O_1 and the C_{13}-H_{13B}···O_2 in light-green dashed lines. The C_6-H_6···π $C_{g(Pyr)}$ are represented in pink dashed lines. The orange and purple dashed lines represent, respectively, the π···π interaction between the centroids of the pyrrole and 4-methoxyphenyl ring and the C_{13}-H_{13C}···π $C_{g(C5\text{-}C7)}$ interaction, both responsible for part of the π-stacking along plane (011).

The two possible N-H···O_1 interactions (each molecule having a NH donor and an O_1 acceptor group), instead of being paired, occur now with two different neighbouring parallel molecules, giving rise to a zigzagged polymer chain arrangement. However, the complementary C_{12}-H_{12}···O_1 bond is maintained. The dihedral angle between the planes of those molecules is of 65.33º. Two interactions of the type C_6-H_6···π $C_{g(Pyr)}$ with two other molecules, having a very similar dihedral angle as that referred previously (65.18º), generate the second layer of the zigzag arrangement. The methoxy substituent in the *para* position of the phenyl ring gave rise to two new interactions of type C-H···O, between methoxy groups (C_{13}-H_{13B}···O_2) of adjacent molecules, displaying a dihedral angle of 62.01º and creating a linear arrangement. Finally, as already noticed in

the previous compounds **1**-**5**, it is possible to observe π···π interactions between the centroids of the pyrrole and 4-methoxyphenyl rings, which are reinforced by a C-H···π interaction, leading to a π-stacking arrangement (Figs. 12, 13 and S2 in Supporting Information).

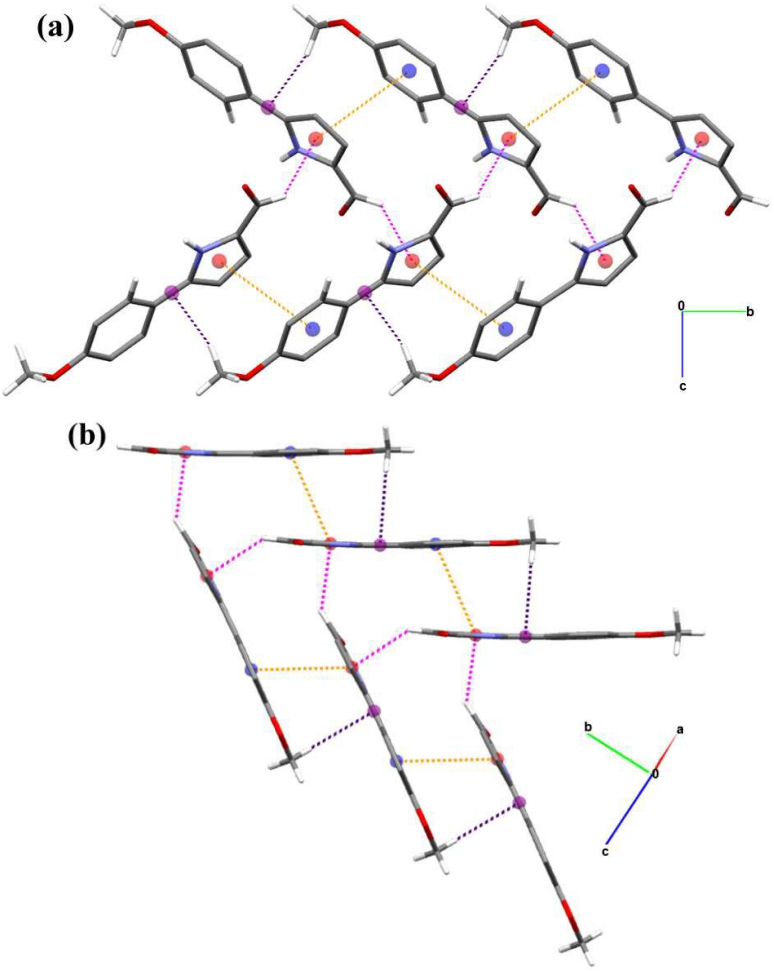


Fig. 13 Evidence of the π···π interactions between the centroids of the pyrrole and 4-methoxyphenyl rings (orange), C_{13}-H_{13C}···π $C_{g(C5-C7)}$ interactions (purple), and C_6-H_6···π $C_{g(Pyr)}$ interactions (pink), (a) along plane (011) and (b) in the 3D view of the crystalline structure of **6**. Hydrogen-bond interactions were omitted for clarity.

For compound **7**, containing a *p*-fluorophenyl substituent, the arrangement around the asymmetric unit is similar to that observed for the previous compound **6**. However, in this case, one molecule interacts with nine other ones (Fig. 14). The N-H···O interactions, complemented by the C-H···O ones, give rise to a zigzag arrangement displaying a dihedral angle of 65.35º. A second layer is then generated due to

interactions of two other molecules, through C_6-$H_6\cdots$ π $C_{g(Pyr)}$ interactions between the carbonyl CH group and the pyrrole ring, with a dihedral angle of 64.84º. On the other hand, C-H···F interactions allow not only a linear alignment with two other molecules (dihedral angle of 59.50º) but also an interaction with a third molecule, along with a C_9-$H_9\cdots$ π $C_{g(Ar)}$, with a dihedral angle of 60.29º. The usually observed π···π interactions of the 4-fluorophenyl and pyrrole ring centroids with two other molecules lead to the formation of the third type of layers, through π-stacking. Similarly to **6**, a 2D packing along plane (011) is observed as an intercalated zigzagged and linear network.

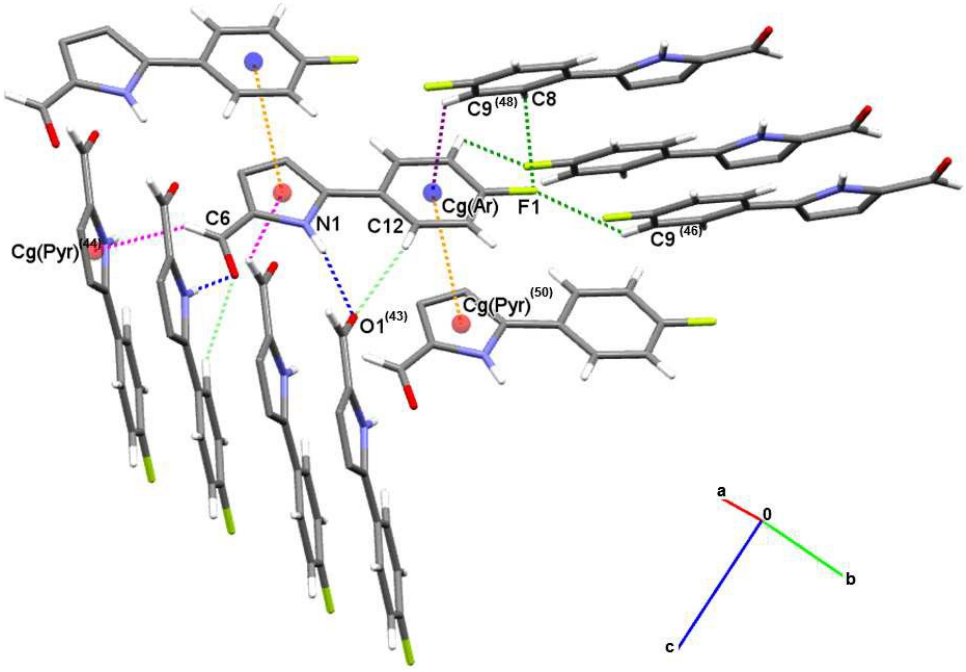


Fig. 14 View of the 3D supramolecular arrangement of compound **7**, showing all the interactions: N_1-$H_1\cdots O_1$ (blue), C_{12}-$H_{12}\cdots O_1$ (light-green), C_8-$H_8\cdots F_1$ and C_9-$H_9\cdots F_1$ (green), C_6-$H_6\cdots$ π $C_{g(Pyr)}$ and C_9-$H_9\cdots$ π $C_{g(Ar)}$ (pink and purple, respectively) and π···π (orange dashed lines).

A list of the hydrogen-bonds and other molecular interactions observed for compounds **1**-**7** and the corresponding symmetry operations in relation to their asymmetric units are summarised in Table 2.

Table 2 List of hydrogen-bonds and other intermolecular interactions for compounds **1**-**7**, and corresponding symmetry operations in relation to the asymmetric unit.

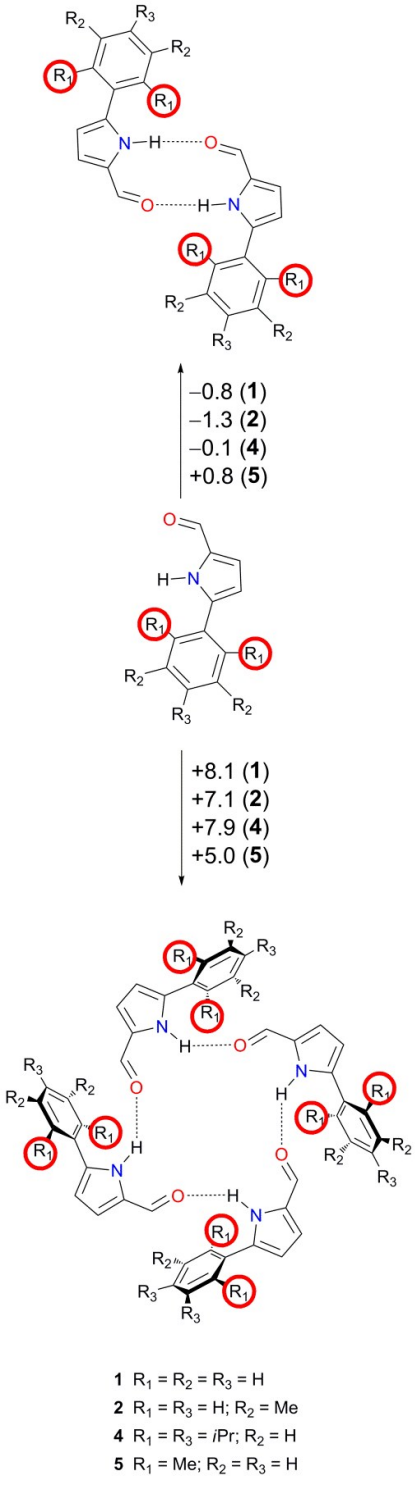
	D-H···A	d(H···A) (Å)	d(D···A) (Å)	(DĤA) (º)	Symmetry Operation
1	N_1-H_1···O_1	2.009(16)	2.9345(18)	170.8(14)	(1) 1-*x*,-*y*,1-*z*
	C_{12}-H_{12}···O_1	2.35	3.277(2)	163	(1)
	C_{10}-H_{10}··· π $C_{g(Ar)}$	3.223	3.977	138	(2) -*x*,-1/2+*y*,1/2-*z*
					(3) -*x*,1/2+*y*,1/2-*z*
	π $C_{g(Pyr)}$-π $C_{g(Ar)}$		3.7798(18)		(4) *x*,-1+*y*,*z*
					(5) *x*,1+*y*,*z*
2	N_1-H_1···O_1	2.07(4)	2.967(5)	170(4)	(6) 2-*x*,1-*y*,-*z*
	C_{12}-H_{12}···O_1	2.54	3.437(6)	157	(6)
	C_6-H_6···π $C_{g(Ar)}$	2.89	3.653(5)	138	(7) 1.5-*x*,-1/2+*y*,1/2-*z*
					(8) 1.5-*x*,1/2+*y*,1/2-*z*
	π $C_{g(Pyr)}$-π $C_{g(Ar)}$		4.055(3)		(9) 1-*x*,1-*y*,-*z*
3	N_1-H_1···O_1	2.01(3)	2.852(3)	177.1(15)	(10) 1-*x*,-2-*y*,1-*z*
	C_{12}-H_{12}···O_1	2.38	3.305(3)	164	(10)
	C_4-H_4···F_1	2.61		134	(11) -*x*,-*y*,1-*z*
	C_3-H_3···F_5	2.48	3.313(3)	147	(12) *x*,-1.5-*y*,-1/2+*z*
					(13) *x*,-1.5-*y*,1/2+*z*
	C_6-O_1···π $C_{g(Pyr)}$	3.394(3)	3.654(2)	68.11(14)	(14) *x*,-1+*y*,*z*
					(15) *x*,1+*y*,*z*
4	N_1-H_1···O_1	1.94(2)	2.834(2)	165.8(16)	(16) 1-*x*,-*y*,2-*z*
	C_{16}-H_{16C}··· π $C_{g(Ar)}$	2.79	3.637(2)	145	(17) 1-*x*,1-*y*,1-*z*
	C_{21}-H_{21B}··· π $C_{g(Pyr)}$	3.264	4.166	154	(18) 1-*x*,1-*y*,2-*z*
5[a]	N_{1A}-H_{1A}···O_{1B}	1.92(3)	2.845(2)	161(3)	(19) 1-*x*,-1/2+*y*,1/2-*z*
					(20) 1-*x*,1/2+*y*,1/2-*z*
	N_{1B}-H_{1B}···O_{1A}	2.02(2)	2.907(2)	164(2)	(21) -1+*x*,1/2-*y*,-1/2+*z*
					(22) 1+*x*,1/2-*y*,1/2+*z*
	C_{6A}-H_{6A}···π $C_{g(ArB)}$	2.70	3.628(3)	166	(21), (22)
	C_{6B}-H_{6B}··· π $C_{g(ArA)}$	2.64	3.405(2)	137	(19), (20)
	C_{10B}-H_{10B}··· π $C_{g(PyrA)}$	2.78	3.679(3)	158	(23) -1+*x*,*y*,*z*
					(24) 1+*x*,*y*,*z*
	C_{14B}-H_{14E}··· π $C_{g(ArA)}$	3.09	3.854	135	(25) *x*,1/2-*y*,1/2+*z*
					(26) *x*,1/2-*y*,-1/2+*z*
	C_{13A}-H_{13A}···O_{1B}	2.50	3.407(3)	153	(27) *x*,1/2-*y*,-1/2+*z*
					(28) *x*,1/2-*y*,1/2+*z*
	C_{13A}-H_{13B}···O_{1B}	2.64	3.573(3)	159	(29) 2-*x*,-1/2+*y*,1/2-*z*
					(30) 2-*x*,1/2+*y*,1/2-*z*
6	N_1-H_1···O_1	1.98(3)	2.873(3)	170(3)	(31) -x,-1/2+y,1/2-z

					(32) -x,1/2+y,1/2-z
	C_{12}-$H_{12}\cdots O_1$	2.40	3.294(3)	156	(31) (32)
	C_6-$H_6\cdots$ π $C_{g(Pyr)}$	2.59	3.367(2)	139	(33) 1-x,-1/2+y,1/2-z
					(34) 1-x,1/2+y,1/2-z
	C_{13}-$H_{13(B)}\cdots O_2$	2.58	3.352(3)	135	(35) -1/2+-x,-1/2-y,1-z
					(36) 1/2+x,-1/2-y,1-z
	π $C_{g(Pyrr)}$-π $C_{g(Ar)}$		3.8892(13)		(37) x,-1+y,z
					(38) x,1+y,z
	C_{13}-$H_{13(c)}\cdots$ π		2.829		(37) (38)
7	N_1-$H_1\cdots O_1$	1.99(3)	2.867(3)	166(3)	(42) -x,-1/2+y,1.5-z
					(43) -x,1/2+y,1.5-z
	C_{12}-$H_{12}\cdots O_1$	2.45	3.317(4)	152	(42) (43)
	C_6-$H_6\cdots$π $C_{g(Pyr)}$	2.60	3.370(3)	138	(44) 1-x,-1/2+y,1.5-z
					(45) 1-x,1/2+y,1.5-z
	C_9-$H_9\cdots F_1$	2.62	3.253(4)	125	(46) -1/2+x,1/2-y,1-z
					(47) 1/2+x,1/2-y,1-z
	C_8-$H_8\cdots F_1$	2.83	3.511	130	(48) -1/2+x,-1/2-y,1-z
	C_9-$H_9\cdots$π $C_{g(Ar)}$	2.94	3.590(4)	127	(48)
	π $C_{g(Pyr)}$-π $C_{g(Ar)}$		3.8501(19)		(49) x,-1+y,z
					(50) x,1+y,z

[a] Molecules A and B of the asymmetric unit do not belong to the same tetramer in the supramolecular arrangement.

DFT calculations[34] were used to study the structural units existing in the solid state structures discussed above and, in particular, the higher degree of aggregation observed in the case of compound **5**, when compared to its analogues. In fact, in the cases of compounds **6** and **7** the presence of heteroatoms on the *para*-position of the phenyl ring (X = O for **6**, and X = F for **7**) allows the establishment of X···H hydrogen interactions, and a resulting extended 3D pattern in the structure of the solid. However, for all other solids, except **5**, the basic unit of the structure is a dimer of molecules based on two complementary N-H···O hydrogen-bonds (the $R^2_2(10)$ synthon referred above). In the case of **5**, a tetramer is the basic unit of the structure.

The free energy balance for the formation of both a dimer and a tetramer was computed for compounds **1**, **2**, **4** and **5** and the results obtained are presented in Scheme 2. Compound **3** has the same aryl substitution pattern as **2** with *meta*-CF_3 groups instead of methyl, thus not being considered in the calculations.



Scheme 2 Free energy balance (kcal mol^{-1}) calculated for the formation of a dimer (top) or a tetramer (bottom) for compounds **1**, **2**, **4** and **5**.

The results in Scheme 2 indicate that the formation of dimers is clearly favoured, compared to the corresponding tetramers, for compounds **1**, **2** and **4**. In all those cases dimer formation is preferred over tetramer formation by 8 kcal mol^{-1} or more. In the case of **5** that difference drops to ca. 4 kcal mol^{-1} and the formation of the dimer is not exergonic. This is in accordance with the crystal structures, namely the presence of tetramers as basic units in the solid structure of compound **5**.

The calculated free energy balances (see Scheme 1) are directly related to the substituent in the *ortho*-position of the phenyl ring (R_1). In compounds **1** and **2** there is no substituent (R_1 = H) and the torsion angle between the aryl and pyrrole rings has an optimised value of 26° in the isolated molecules. Conversely, when the aryl rings present bulkier substituents in the *ortho*-positions, as in the cases of **4** (R_1 = *i*Pr) and **5** (R_1 = Me), steric repulsion between those groups and the pyrrole ring increases the torsion angle between rings in the corresponding molecules. The values obtained in the isolated molecules show that the two rings are almost perpendicular for **4** (85°) and approach that value for **5** (65°). Naturally, the formation of an extended π system justifies the relative orientation of the two rings in **1** and in **2**. For example, the energy required to twist the two rings in a molecule of **1**, from a torsion angle of 26° (the value that corresponds to the optimised structure) to 65° (the value optimised for **5**) is 2 kcal mol^{-1}, showing clearly that there is an electronic reason for the preferred conformation of the two rings in molecules **1** and **2**.

The degree of aggregation observed for the molecules in the solid state is a consequence of the relative geometry of the two rings. In the cases of compounds **1** and **2**, the rings are closer to co-planar and, thus, there is formation of dimers because in the geometry of the tetramers the substituted phenyl ring of one molecule is brought to close proximity to the pyrrole ring of its next neighbour, destabilising this arrangement. That does not occur in the case of compound **5**. Here, the two rings are almost perpendicular and the formation of tetramers becomes possible, yielding a greater degree of aggregation in the basic unit of the corresponding solid state structure. For compound **4**, the rings are practically perpendicular but the steric repulsion between the bulky *i*Pr groups has the same effect that the relative orientation of the ring planes has in compounds **1** and **2**. That is, for **4** the most favoured unit to build the solid state structure is the dimer again, not the tetramer.

Experimental

Materials and methods

The Vilsmeier-Haack reagents were acquired from commercial suppliers and used after a preliminary purification: *N,N*-dimethylformamide (DMF) and phosphoryl chloride ($POCl_3$) were distilled "trap-to-trap" and stored in glass ampoules equipped with J. Young-type taps. All the solvents were pre-dried with 4 Å molecular sieves and purified by refluxing over a suitable drying agent followed by distillation under nitrogen. Diethyl ether, toluene and THF were dried over sodium/benzophenone and dichloromethane (CH_2Cl_2) over calcium hydride. Deuterated chloroform was dried with 4 Å molecular sieves. The starting materials 2-arylpyrroles were synthesised according to a procedure previously reported and partially developed in our group,[35b] with exception of the starting material of compound **3**, which was synthesised by a different procedure.[36] Nuclear Magnetic Resonance (NMR) spectra of compounds **1**-**7**, dissolved in $CDCl_3$, were recorded on Bruker Avance III 300 (^{1}H, 300.130 MHz; ^{13}C, 75.468 MHz; ^{19}F, 282.404 MHz) or Bruker Avance III 400 (^{1}H, 400.130 MHz; ^{13}C, 100.613 MHz) spectrometers. Spectra were referenced internally using the residual protio-solvent (^{1}H) or solvent (^{13}C) resonances, and are reported relative to tetramethylsilane (δ=0). The ^{19}F NMR spectra were referenced using $CFCl_3$ (δ=0). All resonances were characterised by their chemical shifts (δ), quoted in ppm, and coupling constants (*J*), given in Hz. Multiplicities were abbreviated as follows: broad (br), singlet (s), doublet (d), doublet of doublets (dd), quartet (q), septet (sept) and multiplet (m). Elemental analyses were performed in a Fisons Instrument Mod EA-1108, at *Laboratório de Análises* (IST).

General procedure for the synthesis of 5-(substituted phenyl)-2-formylpyrroles

The reaction was adapted from a procedure described in the literature involving Vilsmeier-Haack acylation conditions.[37b] A solution of DMF in toluene was prepared and cooled in an ice-bath under nitrogen atmosphere. A toluene solution of $POCl_3$ was slowly added to the previous DMF solution and the ice-bath removed, allowing the mixture to warm up to room temperature. The resulting Vilsmeier reagent was cooled again in an ice-bath and a solution of the appropriate 2-arylpyrrole in toluene added dropwise. The reaction mixture was allowed to warm up slowly to room temperature with stirring overnight.

The reaction mixture was cooled to 0 ºC and cold water was added, followed by the addition of $NaHCO_3$ until reaching a pH=7. A solution of NaOH 40% (m/v) was then added in order to attain a final pH=12. The basified mixture was stirred for 3 h and the organic phase separated. The aqueous phase was then extracted with $CHCl_3$ until it became colourless, and the combined organic phases were dried over anhydrous Na_2SO_4. Unless otherwise stated, after filtration and removal of all the volatiles, the resulting solid was washed with hot *n*-hexane and dried under vacuum. The product was recrystallised in an appropriate solvent, at -20 ºC, giving crystals suitable for X-ray diffraction.

5-phenyl-2-formylpyrrole (1): The Vilsmeier reagent was prepared from the reaction between solutions of DMF (4.60 ml, 59 mmol) and $POCl_3$ (4.76 ml, 51 mmol), each in 20 mL of toluene. Subsequent addition of the solution of 2-phenylpyrrole **I** (7.16 g, 50 mmol) in 170 ml of toluene afforded a light-orange solid after work-up. Recrystallisation in diethyl ether at -20 ºC yielded **1** (8.19 g, 96%) as red crystals.
δ_H (300 MHz, $CDCl_3$) 9.81 (1H, br s, N*H*), 9.52 (1H, s, CO*H*), 7.65-7.61 (2H, m, *o*-Ph*H*, 7.47-7.41 (2H, m, *m*-Ph*H*), 7.38-7.33 (1H, m, *p*-Ph*H*), 7.03 (1H, dd, J_{HH} 4.0, J_{HH} 2.4, C(3)*H*), 6.65 (1H, dd, J_{HH} 4.0 , J_{HH} 2.6, C(4)*H*). δ_C (75 MHz) 178.9 (*C*OH), 140.3 (C5), 133.3 (C2), 130.7 (C_{ipso}), 129.1 (*m*-Ph*C*), 128.5 (*p*-Ph*C*), 125.4 (*o*-Ph*C*), 123.1 (C3), 109.0 (C4). Found: C 77.1; H 5.3; N 8.2. Calc. for $C_{11}H_9NO$: C 77.2; H 5.3; N 8.2%.

5-(3,5-dimethylphenyl)-2-formylpyrrole (2): The Vilsmeier reagent was prepared from the reaction between solutions of DMF (0.45 ml, 5.85 mmol) and $POCl_3$ (0.45 ml, 5.16 mmol), each in 1 mL of toluene. Subsequent addition of the solution of 2-(3,5-dimethylphenyl)pyrrole **II** (0.87 g, 5.08 mmol) in 10 ml of toluene afforded a light-orange solid after work-up. Recrystallisation in toluene at -20 ºC yielded **2** (0.82 g, 81%) as colourless crystals.
δ_H (300 MHz, $CDCl_3$) 10.01 (1H, br s, N*H*), 9.51 (1H, s, CO*H*), 7.28 (2H, s, *o*-Ph*H*), 7.02 (1H, dd, J_{HH} 3.9, J_{HH} 2.4, C(3)*H*), 7.00 (1H, s, *p*-Ph*H*), 6.62 (1H, dd, J_{HH} 3.9, J_{HH} 2.6, C(4)*H*), 2.36 (6H, s, CH_3). δ_C (75 MHz) 178.8 (*C*OH), 140.7 (C5), 138.8 (*m*-Ph*C*), 133.3 (C2), 130.6 (C_{ipso}), 130.5 (*p*-Ph*C*), 123.4 (*o*-Ph*C*), 123.0 (C3), 109.0 (C4), 21.5 (*C*H_3). Found: C 78.0; H 6.7; N 6.8. Calc. for $C_{13}H_{13}NO$: C 78.4; H 6.6; N 7.0%.

5-(3,5-bis(trifluoromethyl)phenyl)-2-formylpyrrole (3): The Vilsmeier reagent was prepared from the reaction between solutions of DMF (0.51 ml, 6.60 mmol) and $POCl_3$ (0.54 ml, 5.80 mmol), each in 1 mL of toluene. Subsequent addition of the solution of 2-(3,5-bis(trifluoromethyl)phenyl)pyrrole **III** (1.603 g, 5.70 mmol) in 10 ml of toluene afforded an orange solid after the general work-up, until the removal of all the volatiles. Subsequent washings with cold diethyl ether yielded **3** (1.15 g, 65%) as a beige solid, colourless crystals being obtained from the diethyl ether solution at -20ºC.

δ_H (400 MHz, DMSO) 12.8 (1H, br s, N*H*), 9.57 (1H, s, CO*H*), 8.63 (2H, s, *o*-Ph*H*), 7.99 (1H, s, *p*-Ph*H*), 7.19 (1H, d, J_{HH} 3.0, C(3)*H*), 7.14 (1H, d, J_{HH} 3.0, C(4)*H*). δ_C (101 MHz) 179.9 (*C*OH), 135.9 (C5), 134.9 (C2), 133.5 (C_{ipso}), 131.0 (q, J_{CF} 33, *m*-Ph*C*), 125.8 (*o*-Ph*C*), 123.4 (q, J_{CF} 271, *C*F_3), 121.8 (C3), 120.7 (*p*-Ph*C*,), 111.1 (C4). δ_F (282 MHz) -61.37 (s, C*F*$_3$). Found: C 50.9; H 2.2; N 4.4. Calc. for $C_{13}H_7F_6NO$: C 50.8; H 2.3; N 4.6%.

5-(2,4,6-tris(isopropyl)phenyl)-2-formylpyrrole (4): The Vilsmeier reagent was prepared from the reaction between solutions of DMF (0.49 ml, 6.3 mmol) and $POCl_3$ (0.52 ml, 5.6 mmol), each in 1 mL of toluene. Subsequent addition of the solution of 2-(2,4,6-tris(isopropyl)phenyl)pyrrole **IV** (1.48 g, 5.5 mmol) in 20 ml of toluene afforded a brown oil after the general work-up, until the removal of all the volatiles. Stirring in cold *n*-hexane and filtration yielded **4** (1.23 g, 75%) as a light-brown solid. Light-brown crystals were recovered from the *n*-hexane solution at -20ºC.

δ_H (300 MHz, $CDCl_3$) 9.50 (1H, s, CO*H*), 9.01 (1H, br s, N*H*), 7.06 (2H, s, *m*-Ph*H*), 7.04 (1H, dd, J_{HH} 3.8, J_{HH} 2.6, C(3)*H*), 6.24 (1H, dd, J_{HH} 3.6, J_{HH} 2.7, C(4)*H*), 2.94 (1H, sept, J_{HH} 6.8, *p*-C*H*(CH_3)$_2$), 2.59 (2H, sept, J_{HH} 6.9, *o*-C*H*(CH_3)$_2$), 1.29 (6H, d, J_{HH} 6.9, *p*-CH(C*H*$_3$)$_2$), 1.13 (12H, d, J_{HH} 6.9, *o*-CH(C*H*$_3$)$_2$). δ_C (75 MHz) 178.8 (*C*OH), 150.4 (*p*-Ph*C*), 148.9 (*o*-Ph*C*), 138.8 (C5), 132.5 (C2), 126.6 (C_{ipso}), 121.7 (C3), 120.9 (*m*-Ph*C*), 113.0 (C4), 34.6 (*p*-*C*H(CH_3)$_2$), 30.9 (*o*-*C*H(CH_3)$_2$), 24.6 (*o*-CH(*C*H$_3$)$_2$), 24.1 (*p*-CH(*C*H$_3$)$_2$). Found: C 80.7; H 9.1; N 4.6. Calc. for $C_{20}H_{27}NO$: C 80.8; H 9.2; N 4.7%.

5-(2,6-dimethylphenyl)-2-formylpyrrole (5): The Vilsmeier reagent was prepared from the reaction between solutions of DMF (0.88 ml, 11.5 mmol) and $POCl_3$ (0.95 ml, 10.2 mmol), each in 2 mL of toluene. Subsequent addition of the solution of 2-(2,6-dimethylphenyl)pyrrole **V** (1.71 g, 10.0 mmol) in 30 ml of toluene afforded a brown

solid after work-up. Recrystallisation in diethyl ether at -20 ºC yielded **5** (1.23 g, 62%) as light-brown crystals.

δ_H (300 MHz, $CDCl_3$) 9.47 (1H, s, CO*H*), 9.34 (1H, br s, N*H*), 7.26-7.18 (1H, m, *p*-Ph*H*), 7.10 (2H, d, J_{HH} 7.5, *m*-Ph*H*), 7.05 (1H, br s, C(3)*H*), 6.23 (1H, br s, C(4)*H*), 2.14 (6H, s, CH_3). δ_C (75 MHz) 179.0 (*C*OH), 138.9 (C5), 138.1 (*o*-Ph*C*), 132.7 (C2), 131.6 (C_{ipso}), 129.0 (*p*-Ph*C*), 127.6 (*m*-Ph*C*), 121.9 (C3), 112.0 (C4), 20.5 (*C*H_3). Found: C 77.8; H 6.85; N 7.0. Calc. for $C_{13}H_{13}NO$: C 78.4; H 6.6; N 7.0%.

5-(4-methoxyphenyl)-2-formylpyrrole (6): The Vilsmeier reagent was prepared from the reaction between solutions of DMF (0.79 ml, 10.2 mmol) and $POCl_3$ (0.84 ml, 8.98 mmol), each in 2 mL of toluene. Subsequent addition of the solution of 2-(4-methoxyphenyl)pyrrole **VI** (1.53 g, 8.85 mmol) in 40 ml of toluene and 10 ml of THF afforded a light-pink solid after work-up. Recrystallisation in THF at -20 ºC gave a low amount of crystals that were characterised by X-ray diffraction. The mother liquor was evaporated, affording the majority of the product, which was re-dissolved in CH_2Cl_2 and stored at -20 ºC, yielding **6** (1.26 g, 71%) as pink crystals.

δ_H (400 MHz, $CDCl_3$) 9.87 (1H, br s, N*H*), 9.48 (1H, s, CO*H*), 7.59 (2H, d, J_{HH} 8.7, *o*-Ph*H*), 7.01 (1H, dd, J_{HH} 3.7, J_{HH} 2.4, C(3)*H*), 6.96 (2H, d, J_{HH} 8.7, *m*-Ph*H*), 6.55 (1H, dd, J_{HH} 3.6, J_{HH} 2.4, C(4)*H*), 3.85 (3H, s, OCH_3). δ_C (101 MHz) 178.6 (*C*OH), 160.2 (*p*-Ph*C*), 140.5 (C5), 133.1 (C2), 126.9 (*o*-Ph*C*), 123.5 (C_{ipso}), 123.3 (C3), 114.7 (*m*-Ph*C*), 108.4 (C4), 55.5 (O*C*H_3). Found: C 71.6; H 5.5; N 7.0. Calc. for $C_{12}H_{11}NO_2$: C 71.6; H 5.5; N 7.0%.

5-(4-fluorophenyl)-2-formylpyrrole (7): The Vilsmeier reagent was prepared from the reaction between solutions of DMF (0.71 ml, 9.2 mmol) and $POCl_3$ (0.76 ml, 8.12 mmol), each in 2 mL of toluene. Subsequent addition of the solution of 2-(4-fluorophenyl)pyrrole **VII** (1.29 g, 8.0 mmol) in 25 ml of toluene afforded an orange solid after work-up. Recrystallisation in CH_2Cl_2 at -20 ºC yielded **7** (0.965 g, 64%) as yellow crystals.

δ_H (300 MHz, $CDCl_3$) 10.21 (1H, br s, N*H*), 9.51 (1H, s, CO*H*), 7.70-7.63 (2H, m, *o*-Ph*H*), 7.17-7.09 (2H, m, *m*-Ph*H*), 7.04 (1H, dd, J_{HH} 3.8, J_{HH} 1.8, C(3)*H*), 6.59 (1H, dd, J_{HH} 3.7, J_{HH} 2.1, C(4)*H*). δ_C (75 MHz) 179.1 (*C*OH), 163.0 (d, J_{CF} 249.1, *p*-Ph*C*), 139.6 (C5), 133.5 (C2), 127.4 (d, J_{CF} 8.2, *o*-Ph*C*), 127.2 (d, J_{CF} 3.4, C_{ipso}), 123.3 (C3), 116.4

(d, J_{CF} 21.9 Hz, *m*-Ph*C*), 109.1 (C4). δ_F (282 MHz) -112.33 (s, *p*-Ph*F*). Found. C 69.5; H 4.3; N 7.4. Calc. for $C_{11}H_8FNO$: C 69.8; H 4.3; N 7.4%.

Crystallographic data

Crystals of compounds **1**-**7** were covered with polyfluoroether oil and mounted on a nylon loop. Crystallographic data were collected using graphite monochromated Mo-Kα radiation (λ=0.71073Å) on a Bruker AXS-KAPPA APEX II diffractometer equipped with an Oxford Cryosystems open-flow nitrogen cryostat, at 150 K. Cell parameters were retrieved using Bruker SMART software and refined using Bruker SAINT on all observed reflections. Absorption corrections were applied using SADABS.[39] Structure solution and refinement were performed using direct methods with program SIR2004,[40] included in the software package WinGX-V2014.1[41] and SHELXL.[42] All non-hydrogen atoms were refined anisotropically and the hydrogen atoms, except the N*H* protons, were inserted in idealised positions and allowed to refine riding on the parent carbon atom. The N*H* protons were located in the electron density map. For compound **2**, a DFIX restraint was applied in the latter bond. Moreover, the crystals of **2** were of poor quality, though the structure refined to a perfect convergence. In compound **4** the isopropyl group *para*-positioned presented disordered carbon atoms. The disorder was modelled with a free occupancy in the corresponding carbon atoms (57:43), which were subsequently refined anisotropically. The crystals of compound **7** had poor diffracting power, reaching only 0.96 of completeness, though the structure also refined to a perfect convergence. The crystallographic and experimental details of crystal structure determinations are listed in Tables S1 and S2, and selected bond lengths and angles are listed in Table S3. Data was deposited in CCDC under the deposit numbers 1064026 for **1**, 1064027 for **2**, 1064028 for **3**, 1064029 for **4**, 1064030 for **5**, 1064031 for **6** and 1064032 for **7**. CCDC-1064026 to 1064032 contain the supplementary crystallographic data for this paper.

Computational Details

All calculations were performed using the GAUSSIAN 09 software package[43] and the M06-2X functional, without symmetry constraints. That is a hybrid meta-GGA functional developed by Truhlar and Zhao[44] and it was shown to perform very well for

main-group systems, providing a good description of long range effects such as van der Waals interactions or π-π stacking.[45] The optimised geometries were obtained using a standard 6-31G[46] basis set. Frequency calculations confirmed the nature of the stationary points as minima and allowed the conversion of the electronic energies (E_{b1}) to free energy at 298.15 K and 1 atm (G_{b1}) using zero point energy and thermal energy corrections calculated at the same level. Single point energy calculations were performed using a 6-311++G(d,p)[47] basis set and the same functional. The free energy values presented along the text (G_{b2}) were derived from the electronic energy values obtained at the M06-2X/6-311++G(d,p)//M06-2X/6-31G level (E_{b2}), according to the following expression:

$$G_{b2} = E_{b2} + G_{b1} - E_{b1}$$

The basis set superposition error (BSSE) associated with the higher level energy calculations of the aggregates (dimers and tetramers) was estimated by means of Counterpoise calculations[48] and the corrections obtained are included in the corresponding values (Eb2).

Conclusions

A family of 5- (substituted phenyl)-2-formylpyrroles with different electronic and steric features at the phenyl ring was synthesised and fully characterised. The employed synthetic route involved the straightforward formylation reaction of the corresponding 2-(substituted phenyl)pyrrole molecules using Vilsmeier-Haack acylation conditions, moderate to high yields being obtained. Solution NMR spectroscopy, elemental analysis and single-crystal X-ray diffraction were used to characterise all the compounds, and a detailed study of their crystalline arrangements in the solid state was performed. Three different core supramolecular arrangements, dictated by strong N-H···O hydrogen bonding interactions, were found: a) a typical dimerisation observed for compounds **1**-**4**, owing to complementary N-H···O hydrogen bonding, which also occurs in most of the substituted 2-formylpyrrole derivatives reported in the literature; b) the formation of zigzagged polymer chains for compounds **6** and **7**, a pattern observed also in some reported cases; and c) the formation of cyclic tetramers for compound **5**, which is the first example of this type of arrangement reported so far. Important non-classical hydrogen-bonding C-H···O and, in the case of fluorinated compounds, C-H···F

interactions are also found, along with recurrent C-H···π and π···π interactions, owing to the highly aromatised nature of the molecules, all contributing to the three-dimensional crystalline network. DFT calculations indicate that the degree of aggregation observed in the structural units of those compounds is related to the steric repulsion between the aryl ring of a molecule and the pyrrole ring of its closest neighbours. The formation of tetramers is only possible for compound **5**, where there are methyl substituents in the *ortho*-positions of the aryl ring. Those groups bring the two rings in the molecule to an almost perpendicular conformation, on one side, but, on the other, are small enough to allow the proximity between the phenyl ring of a molecule and the pyrrole ring of its closest neighbour existing in the tetramers structure.

Acknowledgements

The authors thank the Fundação para a Ciência e Tecnologia for financial support (Projects PTDC/EQU-EQU/110313/2009 and UID/QUI/00100/2013 and RECI/QEQ-QIN70189/2012) and for fellowships to C.A.F. (SFRH/BD/47730/2008), to P.S.L. (SFRH/BD/88639/2012) and to C.S.B.G. (SFRH/BPD/64423/2009). We thank Prof. M. Teresa Duarte for helpful discussions.

Notes and references

1 (a) J. Montgomery, M. V. Chevliakov, H. L. Brielmann, *Tetrahedron*, 1997, **53**, 16449; (b) L. D. Miranda, R. Cruz-Almanza, M. Pavón, E. Alva, J. M. Muchowski, *Tetrahedron Lett*., 1999, **40**, 7153; (c) Y. Wang, W. Zhang, V. J. Colandrea, L. S. Jimenez, *Tetrahedron*, 1999, **55**, 10659; (d) J. T. Reeves, D. R. Fandrick, Z. Tan, J. J. Song, H. Lee, N. K. Yee, C. H. Senanayake, *J. Org. Chem*., 2010, **75**, 992; (e) J. T. Reeves, Z. Tan, B. Z. Lu, C. H. Senanayake, *J. Heterocyclic Chem*., 2013, **50**, 680.

2 (a) K. Gieβler, H. Griesser, D. Göhringer, T. Sabirov, C. Richert, *Eur. J. Org. Chem*., 2010, 3611; (b) C. Yu, L. Jiao, H. Yin, J. Zhou, W. Pang, Y. Wu, Z. Wang, G. Yang, E. Hao, *Eur. J. Org. Chem*., 2011, 5460.

3 D. Reardon, J. Guan, S. Gambarotta, G. P. A. Yap, *Organometallics*, 2002, **21**, 4390.

4 (a) R. F. R. Jazzar, M. Varrone, A. D. Burrows, S. A. Macgregor, M. F. Mahon, M. K. Whittlesey, *Inorg. Chim. Acta*, 2006, **359**, 815; (b) T. Lundrigan, C. L. M. Jackson, Md. I. Uddin, L. A. Ticker, A. A.-S. Ali, A. Linden, T. S. Cameron, A. Thompson, *Can. J. Chem*., 2012, **90**, 693.

5 J. D. E. T. Wilton-Ely, P. J. Pogorzelec, S. J. Honarkhah, D. H. Reid, D. A. Tocher, *Organometallics*, 2005, **24**, 2862.

6 C. Jablonski, Z. Zhou, J. Bridson, *Inorg. Chim. Acta*, 1997, **254**, 315; S. A. Carabineiro, P. T. Gomes, L. F. Veiros, C. Freire, L. C. J. Pereira, R. T. Henriques, J. E. Warren, S. I. Pascu, *Dalton Trans*., 2007, 5460.

7 R. Alonso, E. Bermejo, R. Carballo, A. Castiñeiras, T. Pérez, *J. Mol. Struct*., 2002, **606**, 155.

8 A. Castiñeiras, R. Carballo, T. Pérez, *Polyhedron*, 2001, **20**, 441.

9 R. Li, B. Moubaraki, K. S. Murray, S. Brooker, *Dalton Trans*., 2008, 6014.

10 K. Mashima, H. Tsurugi, *J. Organomet. Chem*., 2005, **690**, 4414 and references herein cited.

11 C. S. B. Gomes, D. Suresh, P. T. Gomes, L. F. Veiros, M. T. Duarte, T. G. Nunes, M. C. Oliveira, *Dalton Trans*., 2010, **39**, 736.

12 (a) S. A. Carabineiro, L. C. Silva, P. T. Gomes, L. C. J. Pereira, L. F. Veiros, S. I. Pascu, M. T. Duarte, S. Namorado, R. T. Henriques, *Inorg. Chem*., 2007, **46**, 6880; (b) S. A. Carabineiro, R. M. Bellabarba, P. T. Gomes, S. I. Pascu, L. F. Veiros, C. Freire, L. C. J. Pereira, R. T. Henriques, M. C. Oliveira, J. E. Warren, *Inorg. Chem*., 2008, **47**, 8896; (c) C. S. B. Gomes, S. A. Carabineiro, P. T. Gomes, M. T. Duarte, M. A. N. D. A. Lemos, *Inorg. Chim. Acta*, 2011, **367**, 151.

13 R. M. Bellabarba, P. T. Gomes, S. I. Pascu, *Dalton Trans*., 2003, 4431.

14 C. S. B. Gomes, P. T. Gomes, M. T. Duarte, R. E. Di Paolo, A. L. Maçanita, M. J. Calhorda, *Inorg. Chem*., 2009, **48**, 11176.

15 (a) D. Suresh, C. S. B. Gomes, P. T. Gomes, R. E. Di Paolo, A. L. Maçanita, M. J. Calhorda, A. Charas, J. Morgado, M. T. Duarte, *Dalton Trans*., 2012, **41**, 8502; errata: *Dalton Trans.*, 2012, **41**, 14713; and *Dalton Trans.*, 2013, **42**, 16969; (b) D. Suresh, P. S. Lopes, B. Ferreira, C. A. Figueira, C. S. B. Gomes, P. T. Gomes, R. E.

Di Paolo, A. L. Maçanita, M. T. Duarte, A. Charas, J. Morgado, M. J. Calhorda, *Chem. Eur. J.*, 2014, **20**, 4126; (c) D. Suresh, C. S. B. Gomes, P. S. Lopes, C. A. Figueira, B. Ferreira, P. T. Gomes, R. E. Di Paolo, A. L. Maçanita, M. T. Duarte, A. Charas, J. Morgado, D. Vila-Viçosa, M. J. Calhorda, *Chem. Eur. J.*, 2015, **21**, 9133.

16 (a) S. Jiang, S. R. Tala, H. Lu, N. E. Abo-Dya, I. Avan, K. Gyanda, L. Lu, A. R. Katriyzky, A. K. Debnath, *J. Med. Chem*., 2011, **54**, 572; (b) S. Jiang, S. R. Tala, H. Lu, P. Zou, I. Avan, T. S. Ibrahim, N. E. Abo-Dya, A. Abdelmajeid, A. K. Debnath. A. R. Katritzky, *Bioorg. Med. Chem. Lett*., 2011, **21**, 6895.

17 M. Hara, D. Kondo, T. Shono, Y. Kunugi, H. Aihara, *Chem. Lett*., 2013, **42**, 140.

18 (a) H. Sukawa, O. Seshimoto, T. Tezuka, T. Mukai, *J. Chem. Soc., Chem. Commun*., 1974, 696; (b) A. Padwa, J. Smolanoff, A. Tremper, *J. Am. Chem. Soc*., 1975, **97**, 4682; (c) J.-P. Boukou-Poba, M. Farnier, R. Guilard, *Can. J. Chem.* 1981, **59**, 2962; (d) A. Padwa, T. Stengel, *Tetrahedron Lett*., 2004, **45**, 5991; (e) J. Zeng, Y. Bai, S. Cai, J. Ma, X.-W. Liu, *Chem. Commun*., 2011, **47**, 12855.

19 J. T. Reeves, J. J. Song, Z. Tan, H. Lee, N. K. Yee, C. H. Senanayake, *Org. Lett.*, 2007, **9**, 1875.

20 E. Y. Schmidt, A. I. Mikhaleva, A. M. Vasil'tsov, A. B. Zaitsev, N. V. Zorina, *Arkivoc*, 2005, **VII**, 11.

21 E. Y. Schmidt, A. M. Vasil'tov, N. V. Zorina, A. V. Ivanov, A. I. Mikhaleva, B. A. Trofimov, *Chem. Heterocycl. Compd*., 2012, **47**, 1300.

22 J. Roger, H. Doucet, *Adv. Synth. Catal*., 2009, **351**, 1977.

23 For example: (a) O. Q. Munro, S. D. Joubert, C. D. Grimmer, *Chem. Eur. J.*, 2006, **12**, 7987; (b) C. S. B. Gomes, C. A. Figueira, P.T. Gomes, M. T. *Acta Crystallogr. C*, 2008, **64**, o303; (c) C. S. B. Gomes, C. A. Figueira, P. S. Lopes, D. Suresh, P. T. Gomes, M. T. Duarte, *Acta Crystallogr. C* 2011, **67**, o315.

24 For some examples of crystal structures see: (a) K. M. Smith, F. W. Bobe, O. M. Minnetian, H. Hope, M. D. Yanuck, *J. Org. Chem.*, 1985, **50**, 790; (b) X. L. M. Despinoy, S. G. Harris, H. McNab, S. Parsons, K. Withell, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, 1998, **54**, 231; (c) J.-H. Liu, Q.-C. Yang, T. C. W. Mak, H. N. C. Wong, *J. Org. Chem.*, 2000, **65**, 3587; (d) R. A. Davis, A. R. Carroll, R. J. Quinn, P. C. Healy, E. R. T. Tiekink, *Acta Crystallogr., Sect. E: Struct. Rep.*

Online, 2005, **61**, o3401; (e) W. G. O'Neal, W. P. Roberts, I. Ghosh, H. Wang, P. A. Jacobi, *J. Org. Chem.*, 2006, **71**, 3472; (f) R. A. Davis, A. R. Carroll, R. J. Quinn, P. C. Healy, A. R. White, *Acta Crystallogr.*, *Sect. E: Struct. Rep. Online*, 2007, **63**, o4076; (g) D. M. Pinkerton, M. G. Banwell, A. C. Willis, *Org. Lett.*, 2007, **9**, 5127.

25 (a) S.-S. Kang, H.-L. Li, H.-S. Zeng, H.-B. Wang, *Acta Crystallogr., Sect. E: Struct. Rep. Online,* 2008*,* **64,** o1125; (b) K. Sidoryk, M. Malinska, K. Bankowski, M. Kubiszewski, M. Laszcz, M. Bodziachowska-Panfil, M. Kossykowska, T. Giller, A. Kutner, K. Wozniak, *J.Pharm.Sci.*, 2013, **102**, 706.

26 H. Adams, N. A. Bailey, D. E. Fenton, S. Moss, C. O. Rodriguez de Barbarin, G. Jones, *J. Chem. Soc., Dalton Trans.*, 1986, 693.

27 A. Rana, P. K. Panda, *Tetrahedron Lett.*, 2011, **52**, 2697.

28 (a) A. J. Blake, H. McNab, S. Parsons, C. Thornley, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, 1995, **51**, 2348; (b) T. Morimoto, H. Furuta, *Supramol. Chem.*, 2007, **19**, 493.

29 M. Krayer, M. Ptaszek, H.-J. Kim, K. R. Meneely, D. Fan, K. Secor, J. S. Lindsey, *J. Org. Chem.*, 2010, **75**, 1016.

30 For some examples of crystal structures see: (a) J. L. Sessler, G. Berthon-Gelloz, P. A. Gabe, S. Camiolo, E. V. Anslyn, P. Anzenbacher Jr., H. Furuta, G. J. Kirkovits, V. M. Lynch, H. Maeda, P. Morosini, M. Scherer, J. Shiver, R. S. Zimmerman, *Polyhedron*, 2003, **22**, 2963; (b) J. B. Love, A. J. Blake, C. Wilson, S. D. Reid, A. Novak, P. B. Hitchcock, *Chem. Commun.*, 2003, 1682; (c) M. K. Şener, D. Sanchez-Garcia, M. Akkurt, S. Ö. Yildirim, H.-K. Fun, J. L. Sessler, *Turk. J. Chem*., 2006, 243; (d) S. P. Mahanta, B. S. Kumar, S. Baskaran, C. Sivasankar, P. K. Panda, *Org. Lett.*, 2012, **14**, 548.

31 J. Zeng, Y. Bai, S. Cai, J. Ma, X.-W. Liu, *Chem. Commun.*, 2011, **47**, 12855.

32 A. Y. Chernyadyev, Y. A. Ustynyuk, O. V. Yazev, E. A. Kataev, M. D. Reshetova, A. A. Sidorov, G. G. Aleksandrov, V. N. Ikorskii, V. M. Novotortsev, S. E. Nefedov, I. L. Eremenko, I. I. Moiseev, *Russ. Chem. Bull., Int. Ed.*, 2001, **50**, 2445.

33 C. Devillers, D. Lucas, A. K. D. Dime, Y. Rousselin, Y. Mugnier, *Dalton Trans.*, 2010, **39**, 2404.

34 R. G. Parr and W. Yang, in *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.

35 (a) C. A. Figueira, P. T. Gomes, *Catal. Lett.*, 2015, **145**, 762; erratum: C. A. Figueira, P. T. Gomes, *Catal. Lett*. 2015, **145**, 1103; (b) C. A. Figueira, P. S. Lopes, P. T. Gomes, *Tetrahedron*, 2015, **71**, 4362.

36 R. D. Rieth, N. P. Mankad, E. Calimano, J. P. Sadighi, *Org. Lett.*, 2004, **6**, 22, 3981.

37 Some other examples include: (a) R. M. Silverstein, E. E. Ryskiewicz, S. W. Chaikin, *J. Org. Chem*., 1954, 4485; (b) D. O. A. Garrido, G. Buldain, B. Frydman, *J. Org. Chem*., 1984, **49**, 2692; (c) A. I. Mikhaleva, A. B. Zaitsev, A. V. Ivanov, E. Y. Schmidt, A. M. Vasil'tsov, B. A. Trofimov, *Tetrahedron Lett.*, 2006, **47**, 3693; (d) V. Leen, E. Braeken, K. Luckermans, C. Jackers, M. Van der Auweraer, N. Boens, W. Dehaen, *Chem. Commun*., 2009, 4515.

38 In order to simplify the text, the symmetry operations of each interaction are listed in Table 2.

39 G. M. Sheldrick, *SADABS - Program for Empirical Absorption Correction*, Univ. Göttingen, Germany, 1996.

40 SIR2004 - M. C. Burla, R. Caliandro, M. Camalli, B. Carrozzini, G. L. Cascarano, L. De Caro, C. Giacovazzo, G. Polidori, R. Spagna, *J. Appl. Cryst.*, 2005, **38**, 381.

41 L. J. Farrugia, *J. Appl. Cryst*., 1999, **32**, 837.

42 G. M. Sheldrick, *Acta Cryst*., 2008, **A64**, 112.

43 Gaussian 09, Revision A.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador,

J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

44 Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215.

45 (a) Y. Zhao, D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 157; (b) Y. Zhao, D. G. Truhlar, *Chem. Phys. Lett.*, 2011, **502**, 1.

46 (a) R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* 1971, **54**, 724; (b) W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* 1972, **56**, 2257; (c) P. C. Hariharan, J. A. Pople, *Mol. Phys.* 1974, **27**, 209; (d) M. S. Gordon, *Chem. Phys. Lett.* 1980, **76**, 163; (e) P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* 1973, **28**, 213.

47 (a) A. D. McClean, G. S. Chandler *J. Chem. Phys.* 1980, **72**, 5639; (b) R. Krishnan, J. S. Binkley, R. Seeger, J. A. Pople *J. Chem. Phys.* 1980, **72**, 650; (c) A. J. H.Wachters *J. Chem. Phys.* 1970, **52**, 1033; (d) P. J. Hay *J. Chem. Phys.* 1977, **66**, 4377; (e) K. Raghavachari, G. W. Trucks *J. Chem. Phys.* 1989, **91**, 1062; (f) R. C. Binning Jr., L. A. Curtiss *J. Comp. Chem.*, 1990, **11**, 1206; (g) M. P. McGrath, L. Radom *J. Chem. Phys.* 1991, **94**, 511; (h) T. Clark, J. Chandrasekhar, G. W. Spitznagel, P. v. R. Schleyer *J. Comp. Chem.* 1983, **4**, 294; (i) M. J. Frisch, J. A. Pople, J. S. Binkley *J. Chem. Phys.* 1984, **80**, 3265.

48 (a) S. F. Boys, F. Bernardi, *Mol. Phys.* 1970, **19**, 553; (b) S. Simon, M. Duran, J. J. Dannenberg, *J. Chem. Phys.*, 1996, **105**, 11024.

Graphical Abstract

The supramolecular arrangements of 5-(substituted phenyl)-2-formylpyrroles display the core formation of dimers, tetramers or polymers, depending on the phenyl substituents.