



Universiteit  
Leiden  
The Netherlands

## Supramolecular Self-Sorting Networks Using Hydrogen-Bonded Motifs

Coubrough, H.; Lubbe, S.C.C. van der; Parashiv, K.; Minard, A.; Pask, C.; Howard, M.; ... ; Wilson, A.J.

### Citation

Coubrough, H., Lubbe, S. C. C. van der, Parashiv, K., Minard, A., Pask, C., Howard, M., ... Wilson, A. J. (2018). Supramolecular Self-Sorting Networks Using Hydrogen-Bonded Motifs. *Chemistry-A European Journal*, 25(3), 785-795. doi:10.1002/chem.201804791

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/74905>

**Note:** To cite this publication please use the final published version (if applicable).

## Supramolecular Chemistry



## Supramolecular Self-Sorting Networks using Hydrogen-Bonding Motifs

Heather M. Coubrough,<sup>[a, b]</sup> Stephanie C. C. van der Lubbe,<sup>[c]</sup> Kristina Hetherington,<sup>[a, b]</sup> Aisling Minard,<sup>[a, b]</sup> Christopher Pask,<sup>[a]</sup> Mark J. Howard,<sup>[a]</sup> Célia Fonseca Guerra,<sup>\*, [c, d]</sup> and Andrew J. Wilson<sup>\*, [a, b]</sup>

**Abstract:** A current objective in supramolecular chemistry is to mimic the transitions between complex self-sorted systems that represent a hallmark of regulatory function in nature. In this work, a self-sorting network, comprising linear hydrogen motifs, was created. Selecting six hydrogen-bonding motifs capable of both high-fidelity and promiscuous molecular recognition gave rise to a complex self-sorting system, which included motifs capable of both narcissistic and social self-sorting. Examination of the interactions be-

tween individual components, experimentally and computationally, provided a rationale for the product distribution during each phase of a cascade. This reasoning holds through up to five sequential additions of six building blocks, resulting in the construction of a biomimetic network in which the presence or absence of different components provides multiple unique pathways to distinct self-sorted configurations.

## Introduction

Nature has the ability to assemble multiple functional assemblies simultaneously in defined locations and with temporal precision.<sup>[1]</sup> For example, in a cellular signalling cascade, individual proteins interact selectively with certain proteins during one stage of a cascade and then with different proteins at another stage, driven by protein expression levels and enzymatically manipulated post-translational modifications (PTMs).<sup>[1,2]</sup> Thus, regulatory control in a cellular context requires molecular recognition motifs capable of selective but adaptive recogni-

tion behaviour.<sup>[1,3]</sup> In terms of biomimetic systems that recapitulate these features, defined supramolecular assemblies employing multiple consecutive narcissistic (self-loving) and/or social (self-loathing) integrative self-sorting<sup>[4]</sup> events have been used to transition between well-defined—and in some instances functional<sup>[5]</sup>—complexes.<sup>[6]</sup> The parallel assembly of different complexes and transition to express different architectural complexity has also been demonstrated.<sup>[5b]</sup> However, these multicomponent systems have typically relied on shape and geometrical complementarity together with dative coordination bonds. In contrast, non-integrative systems capable of transition between different self-sorted configurations are less established as is the exploitation of weaker interactions, for example, hydrogen-bonding set within the context of a recognition pattern of donors (D) and acceptors (A).<sup>[7]</sup> Achieving this objective is a fundamental challenge and more accurately mimics many of the bimolecular associations that control cellular processes including those that occur within the context of multicomponent protein assemblies (e.g., the binding of co-activators to transcription factor complexes).<sup>[8]</sup>

Previously, we have described a supramolecular system which we termed a self-sorting cascade;<sup>[9]</sup> it exploited both orthogonal and promiscuous recognition behaviour of linear arrays of hydrogen bonds<sup>[10]</sup> to achieve sequential self-sorting depending on which components were present. Comprising four different hydrogen-bonding motifs, four rounds of self-sorting could be achieved with different complexes formed depending on the sequence in which the components were added. Furthermore, the transitions could be triggered (albeit irreversibly) using a light-mediated reaction to unmask one of the H-bonding motifs, thus, mimicking the process of post-translational regulation. In this work, we studied in greater

[a] H. M. Coubrough, K. Hetherington, A. Minard, Dr. C. Pask, Dr. M. J. Howard, Prof. Dr. A. J. Wilson  
School of Chemistry, University of Leeds  
Woodhouse Lane, Leeds LS2 9JT (UK)  
E-mail: a.j.wilson@leeds.ac.uk

[b] H. M. Coubrough, K. Hetherington, A. Minard, Prof. Dr. A. J. Wilson  
Astbury Centre for Structural Molecular Biology, University of Leeds  
Woodhouse Lane, Leeds LS2 9JT (UK)

[c] S. C. C. van der Lubbe, Prof. Dr. C. Fonseca Guerra  
Department of Theoretical Chemistry and  
Amsterdam Centre for Multiscale Modelling, Vrije Universiteit  
De Boelelaan 1081, Amsterdam 1081 HV (The Netherlands)  
E-mail: c.fonseca Guerra@vu.nl

[d] Prof. Dr. C. Fonseca Guerra  
Leiden Institute of Chemistry, Gorlaeus Laboratories, Leiden University  
Wassenaarseweg 76, Leiden 2333 AL (The Netherlands)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:  
<https://doi.org/10.1002/chem.201804791>.

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited, and is not used for commercial purposes.

detail the ability of linear arrays of hydrogen bonds to self-sort using a combination of experimental (single-crystal X-ray diffraction and  $^1\text{H}$  NMR spectroscopy) and computational (DFT) methods. We illustrate that thermodynamically less preferential complexes can be formed during self-sorting, driven by the stability of the entire system, and that this can be influenced by the configuration and conformation of the hydrogen-bonding motif. This allowed us to extend the scope of the self-sorting cascade to six components and construct cascades that operate in the presence of each other but with cross-talk, recapitulating a new aspect of biological signalling and resulting in a self-sorting network.

## Results and Discussion

In this work, we examined the literature to identify additional H-bonding motifs 1–2 to add to those employed 3–6 in our previous study (Figure 1).<sup>[9]</sup> Prior work by Sijbesma and co-workers established that self-complementary dialkylaminoureidopyrimidinone (AUPy) motifs, for example, 1, and amidonaphthyridone (NAPyO) motifs, for example, 2, could be used to form alternating supramolecular copolymers through preferential heterocomplexation, with tuneable polymer composition.<sup>[11]</sup> These building blocks were thus known to elicit low fidelity recognition behaviour in the same way as ureidopyrimidinone (UPy) 4, an essential requirement for the assembly of a cascade.<sup>[12]</sup> Such variable recognition behaviour arises through the different arrays of donors and acceptors (e.g., DADA, DDAA) presented by different accessible tautomers/conformers of each motif. The use of diamidonaphthyridine (DAN) 3, ureidoimidazole (UIM) 5 and amidoisocytosine (AIC) 6, with more restricted molecular recognition properties enforced by a solitary arrangement of donors and acceptors in each case, was also considered necessary for the effective assembly of a cascade. In the prior study, motifs 3–6 were shown to form a pair of tightly bound heterodimers 3·4 and 5·6 in chloroform, resulting from effective self-sorting. Therefore, our study com-

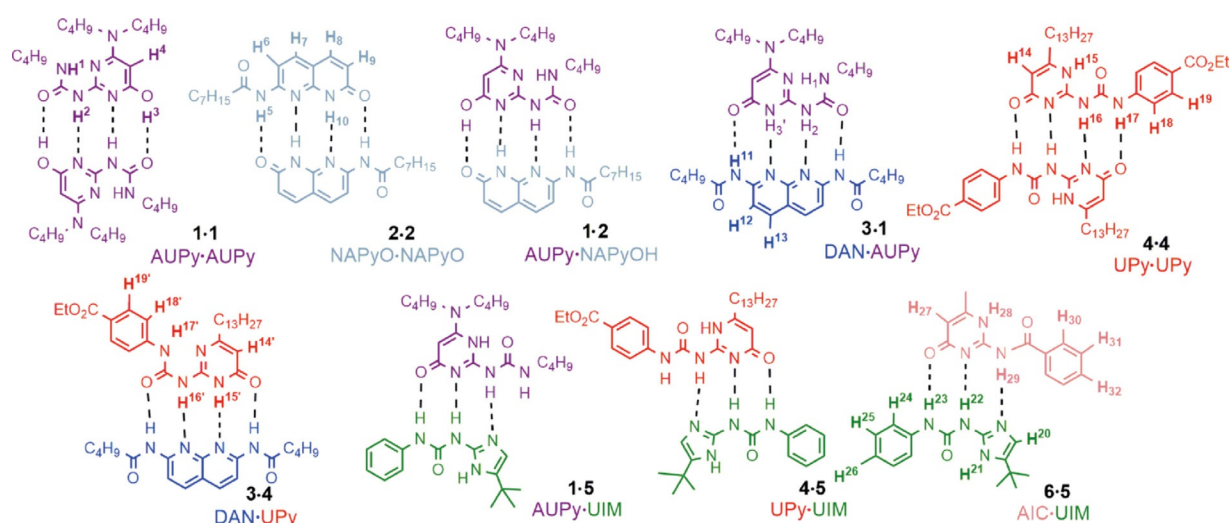
menced with a deeper analysis of the recognition properties of these compounds with an emphasis on 1 and 2.

## Synthesis of hydrogen bonding motifs

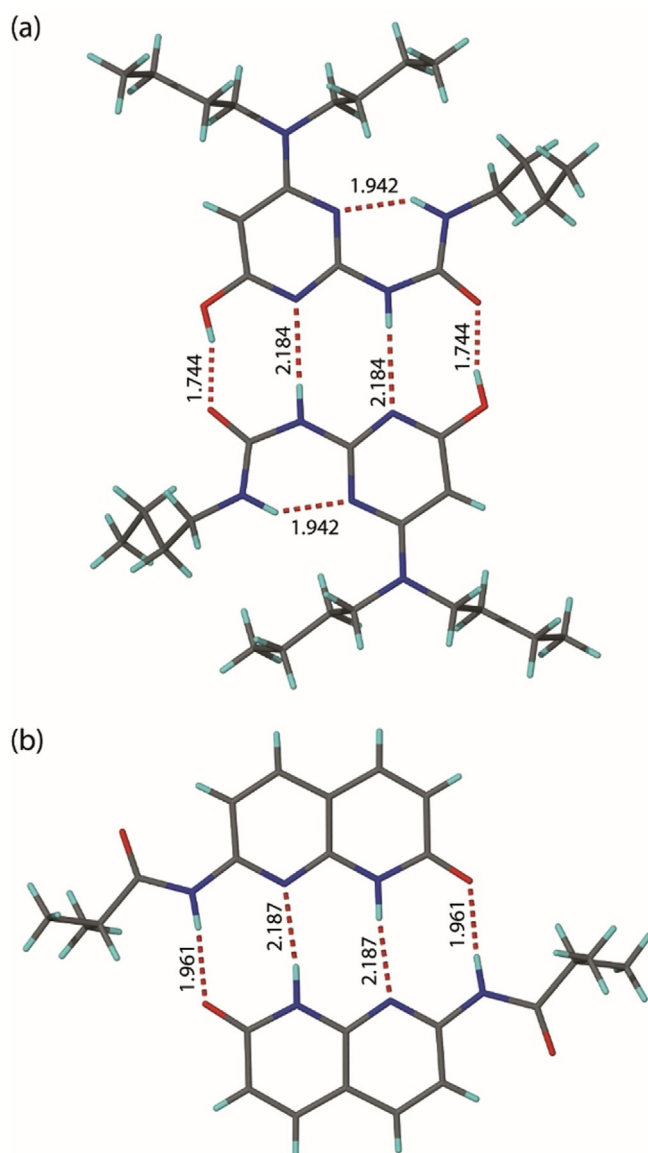
The synthesis of AUPy 1 and NAPyO 2 was achieved by adaptation of previously described methods (see Experimental Section and Supporting Information).<sup>[11b,c]</sup> Like UPy 4, AUPy 1 has the potential to adopt four different tautomers/conformers (Supporting Information, Figure S1a), three of which present a quadruple hydrogen-bond array. A single-crystal analysis confirmed that AUPy 1 forms a quadruple hydrogen-bonded homodimer 1·1 in the solid state; in this case with the monomer adopting the ADAD presenting tautomeric conformer (Figure 2a). This is entirely consistent with results observed previously.<sup>[11b]</sup> Similarly, NAPyO 2 may feasibly access two tautomeric configurations (see Supporting Information, Figure S1b). Single-crystal X-ray analysis on an isopropyl analogue of NAPyO 7 revealed a quadruple hydrogen-bonded homodimer 7·7 in which the monomer was present as the pyrimidinone tautomer, resulting in a DADA array (Figure 2b).

## Pairwise interactions

The interactions of these two additional components (AUPy 1 and NAPyO 2) with partners were then further investigated using  $^1\text{H}$  NMR spectroscopy in  $\text{CDCl}_3$ . The  $^1\text{H}$  NMR resonances from separate samples of AUPy 1·1 and NAPyO 2·2 homodimers were compared to a 1:1 mixture of AUPy 1 and NAPyO 2 (Supporting Information, Figure S2). The small, yet notable, changes in the chemical shift of the NH signals relative to the homodimers were indicative of complex formation and consistent with reported interactions.<sup>[11c]</sup> Attempted 2D  $^1\text{H}$ – $^1\text{H}$  NOESY analyses were confounded by the presence of exchange peaks for NH/OH resonances, preventing a more in-depth structural interpretation, although such exchange peaks are entirely consistent with an interaction between AUPy 1 and NAPyO 2.



**Figure 1.** The lowest energy dimerisation interactions between the six linear hydrogen bonding motifs 1–6, with dashed lines showing intermolecular hydrogen bonds.



**Figure 2.** X-ray crystal structures of a) AUPy 1 and b) NAPyO 7 homodimers (carbon is shown in grey, nitrogen in blue, oxygen in red, hydrogen in light blue and hydrogen bonds as dashed lines, with distances between the donor hydrogen and the acceptor expressed in Å).

To further confirm the presence of a heterodimer, variable-temperature (VT) NMR analyses were performed (Supporting Information, Figure S3). As the temperature was reduced from 277 to 244 K, the broad NH/OH signals in the region 9–15 ppm sharpened and split to reveal additional resonances consistent with the presence of homodimers of AUPy 1·1 and NAPyO 2·2 as well as AUPy·NAPyO 1·2 heterodimer as the dominant species; at 253 K the ratio homodimer/heterodimer was observed to be 1:2. This indicates fast exchange between homodimers 1·1/2·2 and heterodimer 1·2 at room temperature. This is a property that is useful in the context of the cascades/network discussed later, in that recognition simply needs to translate into a response, for example, a signal change/distinct spectrum and not necessarily diagnostic resonances associated with particular complexes. To further understand the observations

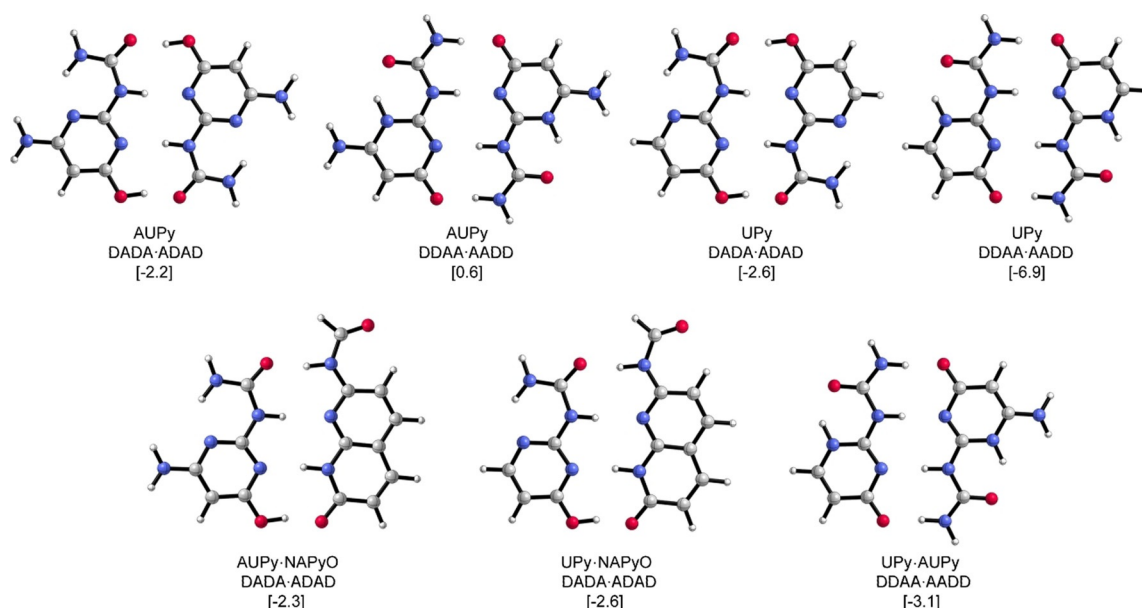
gleaned from  $^1\text{H}$  NMR analyses, we computed the Gibbs free energies of dimerisation ( $\Delta G$ ) at the BLYP-D3(BJ)/TZ2P level of theory with implicit chloroform solvation at 298 K. The Gibbs free energies for the AUPy 1·1 and NAPyO 2·2 homodimers were  $-2.2$  and  $-3.5$  kcal mol $^{-1}$ , respectively (Figure 3, Table 1 and Supporting Information for Cartesian coordinates). The lowest energy AUPy·NAPyO 1·2 heterodimer was seen for the ADAD·DADA interacting array with a Gibbs free energy of  $-2.3$  kcal mol $^{-1}$ . Hence, the dimerisation energy of the AUPy 1·1 homodimer is close to that of the AUPy·NAPyO 1·2 heterodimer, which is consistent with experimental observations of a mixture of homodimers and heterodimer at room temperature.

Next, the interactions of AUPy 1 and NAPyO 2 with UPy 4 were examined. Previous studies have highlighted the promiscuous nature of UPy 4 to interact through several hydrogen-bonding arrays.<sup>[9,12]</sup> Surprisingly,  $^1\text{H}$  NMR spectroscopy suggested that there was no significant interaction of UPy 4 with AUPy 1 or NAPyO 2 in 1:1 mixtures (Supporting Information, Figures S4 and S5). A variable-temperature  $^1\text{H}$  NMR analysis of a 1:1 mixture of NAPyO 2 and UPy 4 further confirmed narcissistic self-sorting (supporting information Figure S6). In such mixtures, UPy 4 remained as homodimer 4·4 and did not show heterodimer interactions with AUPy (1·4) or NAPyO (2·4).

Taken together, these observations raised the following interesting points: 1) AUPy 1 has a preference to adopt a DADA pyrimidinol tautomer in the homodimer, in contrast to the extensively observed DDAA pyrimidinone tautomer in UPy 4, and 2) AUPy 1 has the ability to form a heterodimer with NAPyO 2, whereas UPy 4 does not. To resolve these queries, we turned to DFT calculations. These showed the most stable UPy 4·4 and AUPy 1·1 homodimers to have  $\Delta G$  values of  $-6.9$  and  $-2.2$  kcal mol $^{-1}$ , respectively, whereas the most stable UPy·AUPy 1·4, UPy·NAPyO 2·4 and AUPy·NAPyO 1·2 heterodimers had  $\Delta G$  values of  $-3.1$ ,  $-2.6$  and  $-2.3$  kcal mol $^{-1}$ , respectively (Figure 3). Hence, the dimerisation energy of UPy (i.e., to form 4·4) contrasts with that of AUPy 1·1, being considerably stronger than the potential heterodimers (1·4 and 2·4), which explained why AUPy forms heterodimers, whereas UPy does not.

Next, the molecular basis for the tautomeric difference between AUPy and UPy was reconciled. Naturally, the secondary electrostatic model proposed by Jorgensen,<sup>[13]</sup> and verified by Zimmerman,<sup>[14]</sup> dictates that the DDAA array should be a higher affinity interaction than the DADA array; nevertheless, the DADA tautomer is preferred by AUPy 1·1. Our DFT calculations revealed that this is caused by the differences in tautomerization energy  $\Delta G_{\text{taut}}$  (Supporting Information, Figure S7). When considering the Gibbs free energies of dimerisation with respect to the actual tautomer in the dimer (i.e., excluding  $\Delta G_{\text{taut}}$ ) the AUPy 1·1 is more stable as DDAA than as DADA tautomer, and of similar energy to the calculated UPy 4·4 dimer. Thus, the higher energetic penalty of switching from the pyrimidinol to the pyrimidinone tautomer for AUPy is responsible for its preference to be in DADA tautomeric form. This may be attributed to the intramolecular hydrogen bond between the [6]N and ureido NH (Supporting Information, Figure S8).





**Figure 3.** Optimised dimer interactions and their Gibbs free energies  $\Delta G$  between brackets [in kcal mol<sup>-1</sup>] with respect to their most stable tautomer ADDA. Computed at the BLYP-D3(BJ)/TZ2P level of theory (without alkyl chains).

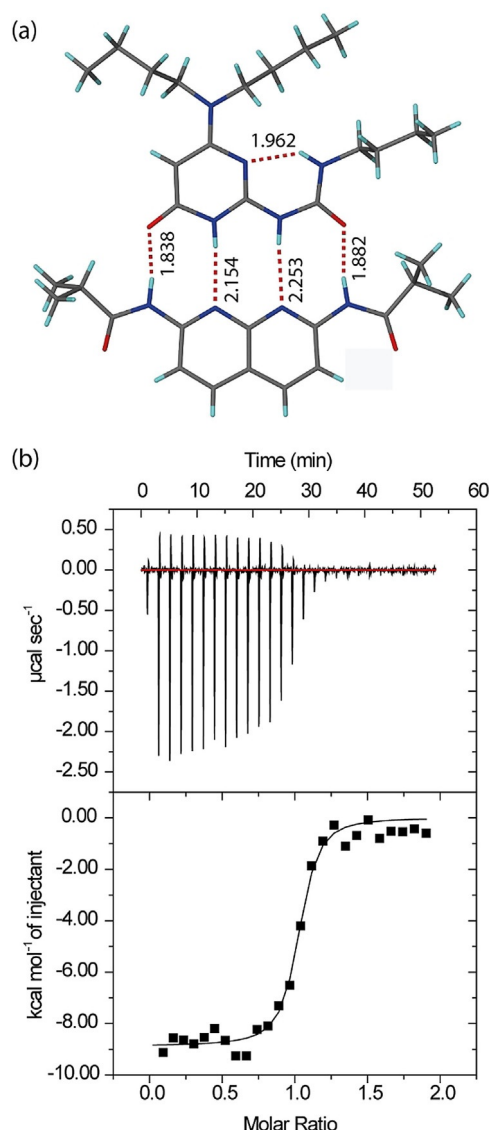
**Table 1.** The Gibbs free energies computed for all dimerisation interactions with respect to the most stable tautomer ADDA in implicit chloroform solvation at the BLYP-D3(BJ)/TZ2P level of theory (without alkyl chains).

Hydrogen-bonding motif	Interacting array	$\Delta G$ [kcal mol <sup>-1</sup> ]
<i>Homodimers</i>		
AUPy	ADAD <sup>[a]</sup>	-2.2
AUPy	AADD	0.6
UPy	ADAD <sup>[a]</sup>	-2.6
UPy	AADD <sup>[a]</sup>	-6.9
NApYO	ADAD <sup>[a]</sup>	-3.5
NApYO	AADD	7.4
<i>Heterodimers</i>		
AUPy-DAN	ADDA-DAAD <sup>[a]</sup>	-5.0
UPy-DAN	ADDA-DAAD <sup>[a]</sup>	-4.8
AUPy-NApYO	ADAD-DADA <sup>[a]</sup>	-2.3
AUPy-NApYO	AADD-DDAA	4.1
AUPy-NApYO	ADDA-DAAD	0.0
UPy-NApYO	ADAD-DADA	-2.6
UPy-NApYO	AADD-DDAA	0.8
UPy-NApYO	ADDA-DAAD	0.5
AUPy-UPy	ADAD-DADA	-2.4
AUPy-UPy	AADD-DDAA	-3.1

[a] Interacting array of dimers seen experimentally.

We then characterised the interaction of the additional components AUPy **1** and NApYO **2** with DAN **3**, presenting only a DAAD arrangement of hydrogen-bonding functionalities. <sup>1</sup>H NMR analyses on 1:1 mixtures of NApYO **2** and DAN **3** demonstrated no interaction, as expected given the incompatible linear arrays. In contrast, <sup>1</sup>H NMR analyses on a 1:1 mixture of AUPy **1** and DAN **3** indicated heterodimer **1·3** formation (Supporting Information Figure S9), in a similar manner to the observation of heterodimer **3·4** formation for DAN **3** and UPy **4** (Supporting Information, Figure S10). 2D <sup>1</sup>H-<sup>1</sup>H NOESY analyses also supported the formation of a heterodimer with NOE cross

peaks between NH signals of both AUPy **1** and DAN **3** (Supporting Information, Figure S11). Additionally, single-crystal X-ray diffraction studies on a co-crystal of AUPy **1** and DAN **8** (isopropyl analogue) revealed the anticipated ADDA-DAAD **1·8** heterodimer interaction, as expected (Figure 4a). Isothermal titration calorimetry (ITC) confirmed a strong interaction with an association constant of  $K_a = 10^6$  M<sup>-1</sup> in chloroform (Figure 4b). Next, we tested which dimers would “win-out” when AUPy **1**, DAN **3** and UPy **4** were combined in a 1:1:1 ratio. The <sup>1</sup>H NMR data were indicative of a mixture of UPy **4·4** homodimer and AUPy-DAN **1·3** heterodimer (Figure 5), consistent with prior observations.<sup>[11c]</sup> Our DFT analyses indicated no significant Gibbs free energy difference  $\Delta G$  between UPy-DAN **3·4** (-4.8 kcal mol<sup>-1</sup>) and AUPy-DAN **1·3** (-5.0 kcal mol<sup>-1</sup>). Thus, the energetic difference between the two heterodimers fails to explain preferential AUPy-DAN (**1·3**) heterodimer formation. Instead, the behaviour is explained by considering the free energy change in this system and the additive effects of each dimer formed. Again, the high stability of the UPy **4·4** homodimer drove the behaviour of the system; considering that the  $\Delta G$  for the UPy **4·4** homodimer is lower (i.e., more negative) than both heterodimers by 1.9 (AUPy-DAN **1·3**) and 2.1 (UPy-DAN **3·4**) kcal mol<sup>-1</sup>, UPy dimerises with itself instead of with DAN, and the AUPy **1·1** homodimer consumes two equivalents on DAN to form the AUPy-DAN **1·3** heterodimer. Although calculations revealed that AUPy-DAN **1·3** is 2.8 kcal mol<sup>-1</sup> more stable than the AUPy **1·1** homodimer, the requirement to maximise hydrogen-bonding interactions dictates this behaviour. To provide further evidence that the overall number of hydrogen bonds in the system must be maximised before considering the strength of the dimerisation interaction and lowest energy conformers, a further <sup>1</sup>H NMR experiment was performed with DAN **3**, AUPy **1** and UPy **4** in 2:1:1 ratio (Supporting Information, Figure S12). This predictably resulted in an equal amount of AUPy-DAN **1·3**



**Figure 4.** Structural and thermodynamic characterisation of AUPy-DAN 1-8 heterodimer. a) Single-crystal structure of AUPy-DAN 1-8 heterodimer (carbon is shown in grey, nitrogen in blue, oxygen in red, hydrogen in light blue and hydrogen bonds as dashed lines, with distances between the donor hydrogen and the acceptor expressed in Å). b) ITC thermograms (top) and binding isotherm for the addition of a solution of AUPy 1 (1 mM) into DAN 3 (0.1 mM) in  $\text{CHCl}_3$  (bottom).

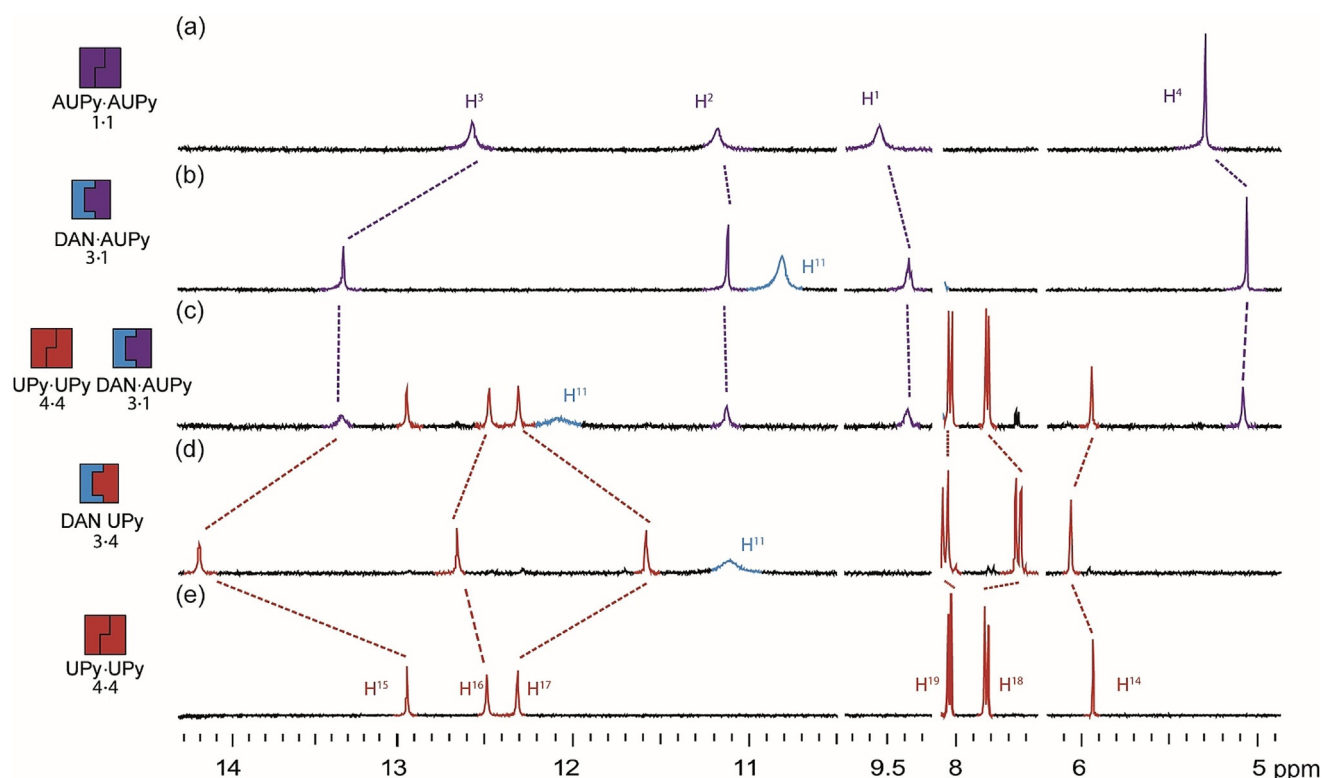
and UPy-DAN 4-3 heterodimers with no homodimers present. Here, the excess of DAN 3 drives the formation of UPy-DAN 4-3 heterodimer over the UPy 4-4 homodimer, despite an energetic bias towards the latter.

Finally, the AUPy-DAN 1-3 heterodimer was also shown to “win-out” in the presence of NAPyO 2. Upon addition of DAN 3 to the weak AUPy-NAPyO 1-2 heterodimer, NAPyO 2 was displaced, resulting in the formation of AUPy-DAN 1-3 heterodimer and NAPyO 2-2 homodimer (Supporting Information, Figure S13). This can be readily explained in terms of both maximising the number of hydrogen bonds in a system and association strength. Unsurprisingly, DFT computations mirror these experimental observations; AUPy-NAPyO 1-2 and AUPy-DAN 1-3 heterodimers have a  $\Delta G$  of  $-2.3$  and  $-5.0 \text{ kcal mol}^{-1}$ .

### Self-sorting cascades

Having characterised the pairwise dimerisation behaviour of the individual components, this allowed for the construction of a number of self-sorting cascades. The simplest of which was conceived though the addition of NAPyO 2 to AUPy 1 followed by the addition of DAN 3, as described above (Supporting Information, Figure S13). More complex self-sorting cascades exploiting five hydrogen-bonding motifs were also investigated (Figure 6a). Beginning with AUPy-AUPy 1-1 homodimer (Figure 6e), the addition of UIM 5 (Figure 6d) resulted in a small shift in the  $\text{H}^4$  resonance of AUPy 1 and broadening of the NH signals, consistent with AUPy-UIM 1-5 heterodimer formation (Supporting Information, Figure S14). The AIC 6 motif was then added to the mixture (Figure 6c), resulting in the formation of UIM-AIC 5-6 indicated by the downfield shift in the  $\text{H}^{24}$  resonance and broadening of the  $\text{H}^{20}$  resonance of UIM 5 (Supporting Information, Figure S15). Concomitant recovery of the AUPy homodimer 1-1 was observed on the basis of the restoration of its NH resonances towards the expected frequency. Upon the addition of NAPyO 2 (Figure 6b) the interaction between AIC-UIM 6-5 was unchanged, but small changes in chemical shifts of the NH resonances of AUPy 1 were observed, consistent with the formation of AUPy-NAPyO 1-2 as the dominant species comprising either components (as previously discussed). The cascade was completed by the addition of DAN 3 (Figure 6a). The  $^1\text{H}$  NMR spectrum of AIC-UIM 6-5 heterodimer showed that no resonances indicative of this complex were affected. However, a significant upfield shift for proton  $\text{H}^4$  and sharpening of NH resonances for AUPy 1 indicated formation of AUPy-DAN 1-3 and the displacement of NAPyO 2 from complexation with AUPy 1 (Supporting Information, Figure S12). Based on this reasoning, the final mixture was comprised of AUPy-DAN 1-3 and AIC-UIM 6-5 heterodimers as well as NAPyO-NAPyO 2-2 homodimer (as predicted) and a clear pathway of sequential self-sorting established for five components.

A series of  $^1\text{H}$  NMR experiments were then performed to exemplify a six-component cascade (Figure 7). The first six-component cascade (pathway E, see Figure 9) initiated with UPy 4-4 homodimer; addition of AUPy 1 led to no change in chemical shifts consistent with narcissistic self-sorting of UPy 4-4 and AUPy 1-1 homodimers (Figure 7e). Upon addition of UIM 5, UPy-UIM 4-5 and AUPy-UIM 1-5 heterodimers formed as major products (Figure 7d). This was indicated by the upfield shift and broadening of  $\text{H}^{18}$  of UPy 4 as well as by the upfield shift of  $\text{H}^4$  and broadening of NH signals of AUPy 1, in line with heterodimerisation observed in the two-component mixture (Supporting Information, Figures S14 and S16). Notably, distinct spectral changes are sufficient to distinguish this phase of the cascade from the preceding and subsequent phases, demonstrating that the absence of clear speciation may be tolerated in cascades (see earlier discussion on AUPy-NAPyO 1-2). AIC 6 was then added (Figure 7c), resulting in a strong heterodimer formation with UIM 5 (illustrated by the downfield shift of  $\text{H}^2$ , Supporting Information, Figure S14), and simultaneous reformation of the UPy 4-4 and AUPy 1-1 homodimers. Formation of the AUPy-NAPyO 1-2 heterodimer occurred on addition



**Figure 5.** Analysis of pairwise interactions and preferences for AUPy 1, DAN 3 and UPy 4 by  $^1\text{H}$  NMR spectroscopy (10 mm,  $\text{CDCl}_3$ ). a) AUPy 1; b) 1:1 AUPy 1: DAN 3; c) 1:1:1 AUPy 1: DAN 3: UPy 4; d) 1:1 DAN 3: UPy 4; (e) UPy 4.

of NAPyO 2 to the cascade in the presence of UPy 4-4 homodimer and UIM-AIC 5-6 heterodimers (Figure 7b). Addition of DAN 3, as expected, disrupted the 1-2 complexation resulting in AUPy-DAN 1-3 heterodimer and NAPyO 2-2 homodimers (Figure 7a). Thus, a final product distribution with four major components AUPy-DAN 1-3, UIM-AIC 5-6 heterodimers and NAPyO 2-2, UPy 4-4 homodimers was observed.

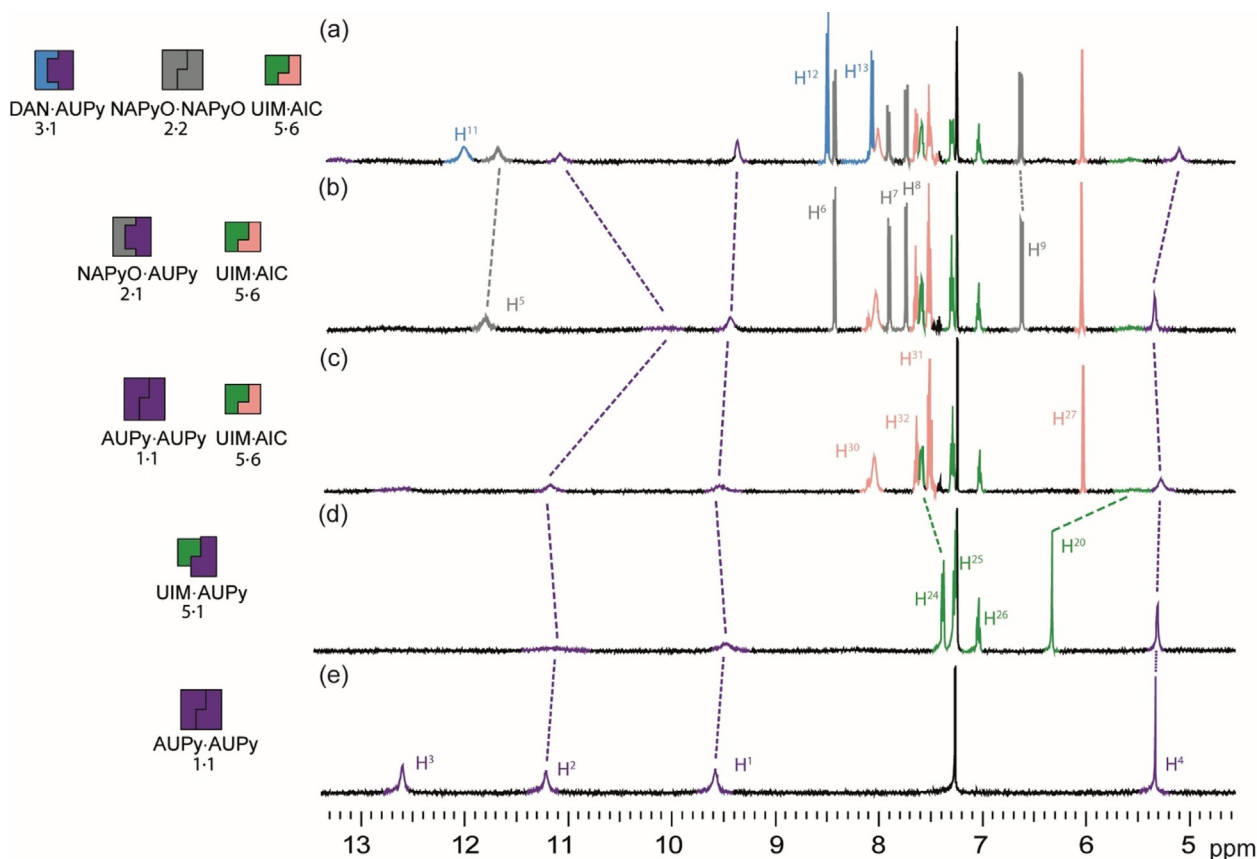
### Orthogonal self-sorting cascade

To develop cascades occurring in parallel (but comprising components capable of cross-talk), an approach to perform two orthogonal self-sorting cascades (pathway G, Figure 8) was developed. The aim was to create two parallel cascades that can operate in sequence, in spite of the ability of complexes from both cascades to interact with one another in different scenarios. Starting with a mixture of UPy 4-4 and AUPy 1-1 homodimers (Figure 8e), the cascade could be split in two with UPy 4-4 at the head of one channel and AUPy 1-1 at the head of the other. When NAPyO 2 was added (Figure 8d), the "UPy channel" was unchanged, whereas AUPy 1 interacted with NAPyO 2 to form AUPy-NAPyO 1-2 heterodimer. Upon addition of UIM 5 to this three-component mixture (Figure 8c), the "AUPy channel" was unchanged and UPy-UIM 4-5 heterodimer formed in the UPy channel. With the addition of AIC 6 (Figure 8b), the "AUPy channel" was unchanged through a further stage but, in the "UPy channel", UPy-UIM 4-5 heterodimer was disrupted to form UPy 4-4 homodimer and UIM-AIC 5-6 heterodimer. The

cascade was completed with the addition of DAN 3; in this case, the "UPy channel" was unchanged but DAN 3 interacted with AUPy 1 in the "AUPy channel" to form AUPy-DAN 1-3 heterodimer liberating NAPyO 2-2 homodimer. Overall, four social self-sorting transitions occurred, two in one channel and two in the other; however, the transitions did not cross over resulting in orthogonal cascades.

### Self-sorting network

To elaborate a self-sorting network, several different additional cascades were exemplified; understanding potential hydrogen-bonding interactions between motifs 1-6 was key in creating a self-sorting network with intersecting pathways (Figure 9). By comparing the  $^1\text{H}$  NMR spectra of each mixture with the distinct chemical shifts of heterodimers and homodimers, the major components in each pathway were identified. Hence, several different self-sorting pathways could be plotted together based on their product distribution, resulting in a network of diverging and converging paths with the same end point. For example, pathway B (red, Figure 9) initiated with UPy 4-4 homodimer and, upon addition of DAN 3, led to production of DAN-UPy 3-4 heterodimer (Supporting Information, Figure S18e). Pathway B continued with the addition of AUPy 1, resulting in disassembly of DAN-UPy 3-4 heterodimer at the expense of DAN-AUPy 3-1, together with regeneration of UPy 4-4 homodimer (Supporting Information, Figure S18d). The hydrogen-bonding interactions were unchanged on the addition of



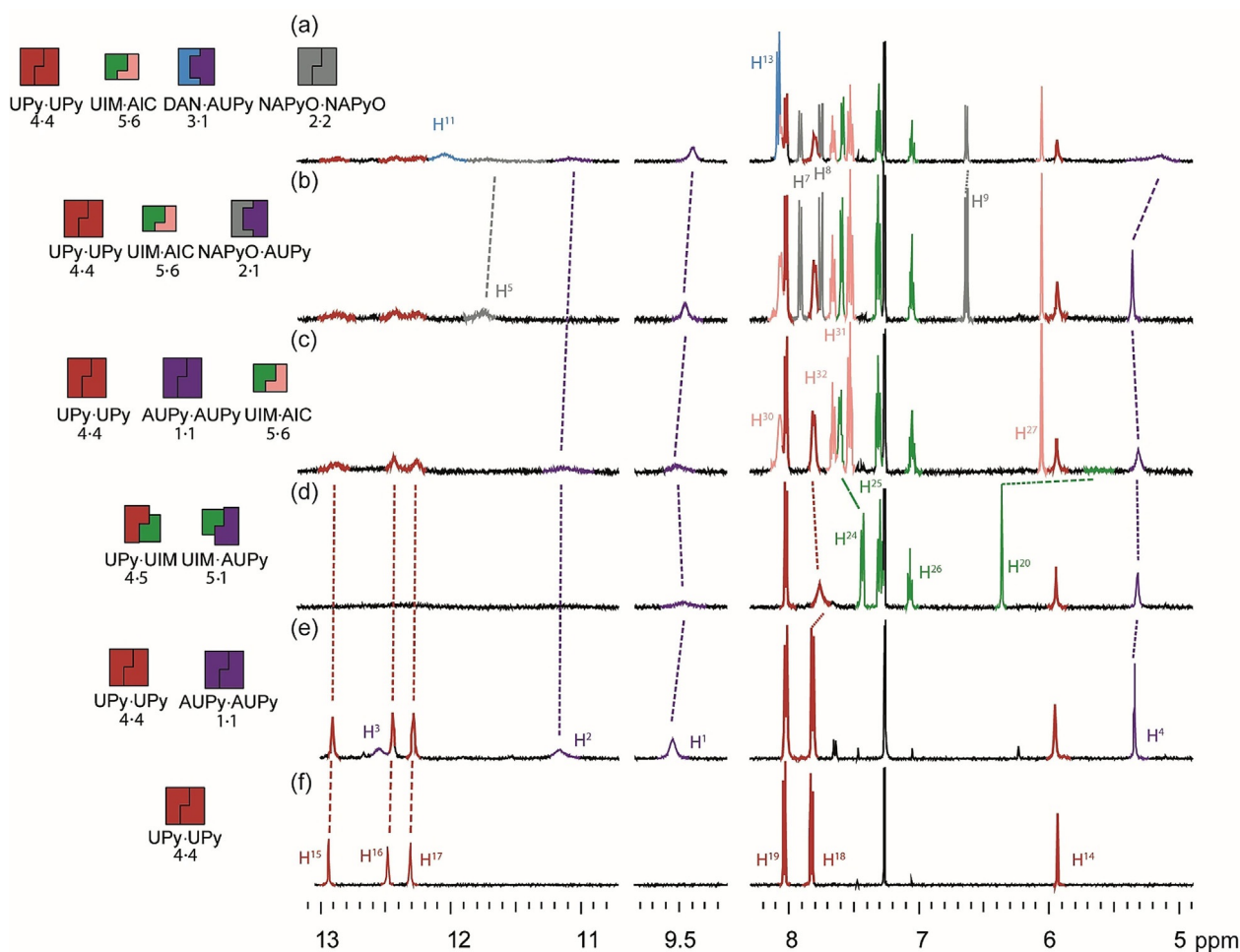
**Figure 6.** Five-component signalling cascade studied by  $^1\text{H}$  NMR (500 MHz, 10 mm,  $\text{CDCl}_3$ ) (a) AUPy 1, NAPyO 2, DAN 3, UIM 5 and AIC 6; (b) AUPy 1, NAPyO 2, UIM 5 and AIC 6; (c) AUPy 1, UIM 5 and AIC 6; (d) AUPy 1, UIM 5; (e) AUPy 1.

NAPyO 2, which was present as the 2:2 homodimer (Supporting Information, Figure S18c). UIM 5 was able to interfere with the UPy 4:4 homodimer interactions to form UPy-UIM 4:5 heterodimer (Supporting Information, Figure S18b), which was in turn disassembled to form UIM-AIC 5:6 heterodimer with concomitant UPy 4:4 homodimer formation on addition of AIC 6 (Supporting Information, Figure S16a). Pathway C (yellow, Figure 9) took a different route from pathway B, in that UIM 5 and AIC 6 are successively added to the DAN-UPy 3:4 heterodimer (considered here as an interchange) to create UPy-UIM 4:5 and UIM-AIC 5:6 heterodimers, respectively (Supporting Information, Figure S19c,d). The ability of UIM 5 to disrupt the DAN-UPy 3:4 heterodimer is moderate and a series of low-fidelity complexes formed. Nonetheless, the resultant  $^1\text{H}$  NMR spectrum was diagnostic of a distinct state within the network, and, the addition of AIC 6 restored a well-resolved spectrum indicative of a well-defined self-sorted configuration. Finally, in pathway C, addition of AUPy 1 switched the configuration from DAN-UPy 3:4 and UIM-AIC 5:6 to AUPy-DAN 1:3, UPy 4:4 and UIM-AIC 5:6 with subsequent addition of NAPyO 2:2 promoting no change in the distribution of the other components (Supporting Information, Figure S19a,b). Pathways A, D, E and F further exemplify the different network configurations that can be obtained depending on which components are present and, therefore, expressed in the system (see Supporting Information, Figures S19–S21 for experimental data).

## Conclusion

Detailed analysis of molecular recognition behaviour of hydrogen-bonding motifs by experiments and computations allowed for a complex self-sorting network made up of several signalling cascades to be created. The sequential addition of six linear hydrogen motifs led to cascades capable of both narcissistic and social self-sorting phases as well as a mixture of both. Through examination of the interactions between the individual components of each cascade, the product distribution of the overall system can be understood in terms of orthogonal recognition. The varying degrees of fidelity and promiscuity of the hydrogen-bonding interactions of these motifs and critically an understanding of their behaviour was essential in developing the cascades comprising the network. These cascades include for the first-time parallel cascades that operate in the presence of each other, but which are capable of cross-talk. Future studies will centre on exploiting these motifs as components of self-sorting and reconfigurable materials, and on developing reversible transitions between the different phases of the cascade. Beyond this, developing approaches to temporarily perturb the system may provide access to out-of-equilibrium networks with emergent behaviour.





**Figure 7.** Pathway E.  $^1\text{H}$  NMR (500 MHz, 10 mm,  $\text{CDCl}_3$ ) six-component signalling cascade a) AUPy 1, NAPyO 2, DAN 3, UPy 4, UIM 5, and AIC 6; b) AUPy 1, NAPyO 2, UPy 4, UIM 5, and AIC 6; c) AUPy 1, UPy 4, UIM 5, and AIC 6; d) AUPy 1, UPy 4, and UIM 5; e) AUPy 1 and UPy 4; f) UPy 4.

## Experimental Section

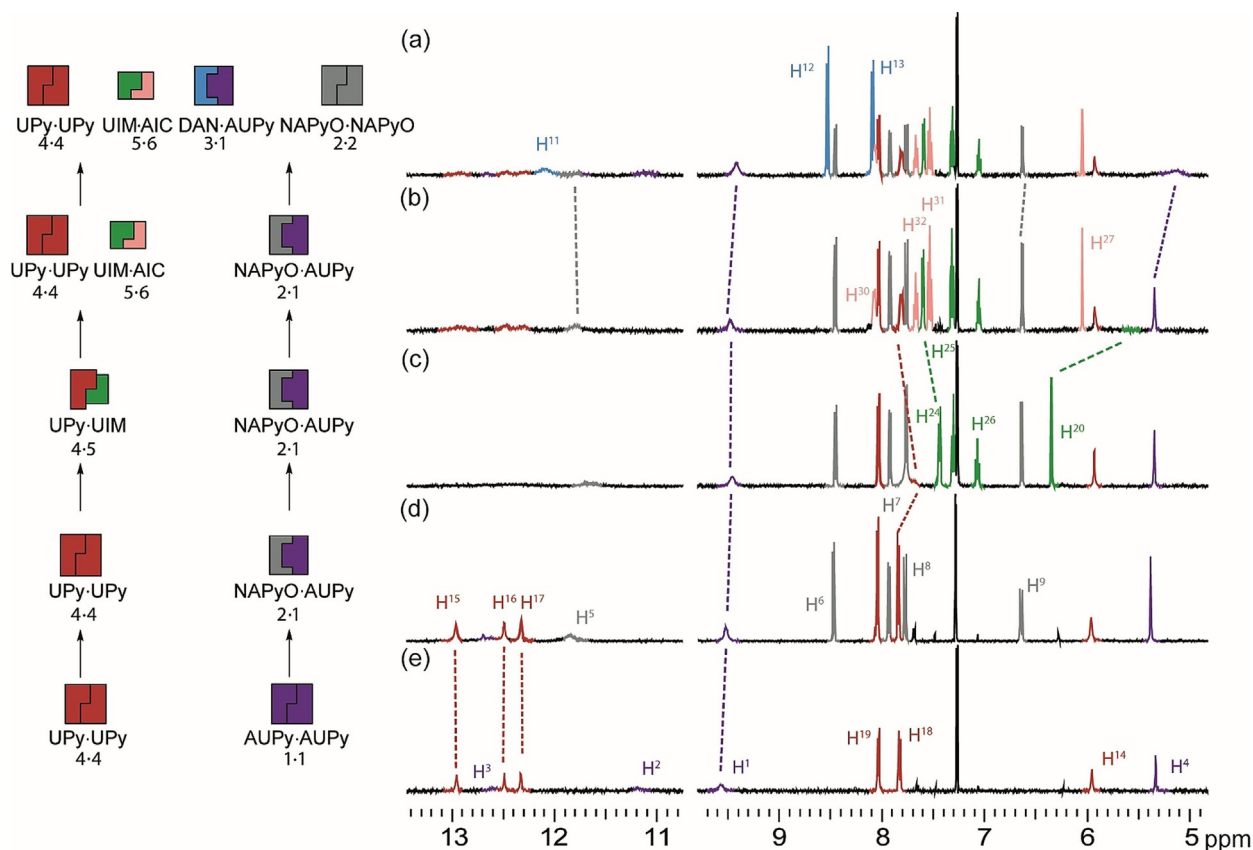
**General considerations:** Solvents and reagents were purchased from Sigma Aldrich or Fisher Scientific and used without further purification unless otherwise stated. Where anhydrous solvents were required, dichloromethane, chloroform, tetrahydrofuran and acetonitrile were obtained from the in-house solvent purification system Innovative Inc. PureSolv®. Anhydrous pyridine was placed over KOH for 24 hours before being refluxed for 2 hours and distilled over Linde 5 Å molecular sieves and solid KOH before use. All non-aqueous reactions were carried out under a nitrogen atmosphere. Chloroform- $d$  was placed on  $\text{CaCl}_2$  before being distilled over Linde 5A molecular sieves before use in  $^1\text{H}$  NMR cascade experiments. Analytical thin layer chromatography was performed on Merck Kieselgel 60  $F_{254}$  0.25 mm pre-coated aluminium plates. Product spots were visualised under UV light ( $\lambda_{\text{max}}=254\text{ nm}$ ). Flash chromatography was carried out using Merck Kieselgel 60 silica gel. Nuclear magnetic resonance spectra were obtained at 298 K (unless stated) using a Bruker AV500 spectrometer operating at 11.4 T (500 MHz for  $^1\text{H}$ ) and JEOL ECA600ii operating at 14.1 T (150 MHz for  $^{13}\text{C}$ ) and NOESY spectra as stated. Infra-red spectra were obtained using a PerkinElmer FTIR spectrometer in which absorption maxima ( $\tilde{\nu}_{\text{max}}$ ) are expressed in wavenumbers ( $\text{cm}^{-1}$ ) and only structurally relevant absorptions have been included. High-Resolution mass spectra were recorded with a BrukerDaltonicsmi-

croTOF using electrospray ionisation (ESI). Detailed synthetic procedures and characterization are given in the Supporting Information.

**Crystal structure determination for 1:** Single crystals were grown by the slow evaporation of **1** in acetonitrile. X-Ray diffraction data were collected at the University of Leeds. Crystal data.  $\text{C}_{17}\text{H}_{31}\text{N}_5\text{O}_2$ ;  $M=337.47$ ; crystal size  $0.21\times0.07\times0.04\text{ mm}$ ; triclinic; space group  $P\bar{1}$ ;  $a=5.1924(6)$ ,  $b=11.8545(15)$ ,  $c=15.3976(10)\text{ Å}$ ,  $\alpha=97.911(8)$ ,  $\beta=93.859(7)$ ,  $\gamma=102.021(10)^\circ$ ;  $V=913.65(17)\text{ Å}^3$ ;  $T=120.3(7)\text{ K}$ ;  $Z=2$ ;  $I=0.661\text{ mm}^{-1}$ ;  $\lambda=1.54184\text{ Å}$  [Cu-K $\alpha$ ]; 6466 reflections measured; 3410 unique reflections ( $R_{\text{int}}=0.0462$ ); observed  $I>2\sigma(I)$ . The final  $R_1$  was 0.0532 (observed reflections 0.0774) and  $wR(F^2)$  was 0.1306 (all data 0.1468) for 232 parameters.

**Crystal structure determination for 7:** Single crystals were grown by the slow evaporation of **7** in acetonitrile. X-Ray diffraction data were collected at the University of Leeds. Crystal data.  $\text{C}_{12}\text{H}_{13}\text{N}_3\text{O}_2$ ,  $M=231.25$ ; crystal size  $0.15\times0.06\times0.03\text{ mm}$ ; monoclinic; space group  $P2_1/n$ ;  $a=4.8837(11)$ ,  $b=22.962(4)$ ,  $c=9.9321(18)\text{ Å}$ ;  $\alpha=90$ ,  $\beta=101.26(2)$ ,  $\gamma=90^\circ$ ;  $V=1092.4(4)\text{ Å}^3$ ;  $T=120.00(10)\text{ K}$ ;  $Z=4$ ;  $I=0.811\text{ mm}^{-1}$ ;  $\lambda=1.54184\text{ Å}$  [Cu-K $\alpha$ ]; 4018 reflections measured; 2121 unique reflections ( $R_{\text{int}}=0.0500$ ); observed  $I>2\sigma(I)$ . The final  $R_1$  was 0.0832 (observed reflections 0.1048) and  $wR(F^2)$  was 0.2057 (all data 0.2223) for 164 parameters.

**Crystal structure determination for 1-8:** Single crystals were grown by slow evaporation of a 1:1 mixture of **1** and **8** in acetonitrile.



**Figure 8.** Pathway G.  $^1\text{H}$  NMR (500 MHz, 10 mm,  $\text{CDCl}_3$ ) orthogonal signalling cascade a) AUPy 1, NAPyO 2, DAN 3, UPy 4, UIM 5, and AIC 6; b) AUPy 1, NAPyO 2, UPy 4, UIM 5, and AIC 6; c) AUPy 1, NAPyO 2, UPy 4 and UIM 5; d) AUPy 1, NAPyO 2 and UPy 4; e) AUPy 1 and UPy 4.

trile. X-Ray diffraction data were collected at the University of Leeds. Crystal data.  $\text{C}_{33}\text{H}_{51}\text{N}_9\text{O}_4$ ,  $M = 637.82$ ; crystal size  $0.29 \times 0.11 \times 0.08$  mm; monoclinic; space group  $P2_1/n$ ;  $a = 9.52157(15)$ ,  $b = 15.8558(3)$ ,  $c = 22.3751(3)$  Å;  $\alpha = 90$ ,  $\beta = 92.0659(13)$ ,  $\gamma = 90^\circ$ ;  $V = 3375.81(9)$  Å<sup>3</sup>;  $T = 119.99(13)$  K;  $Z = 4$ ;  $I = 0.684$  mm<sup>-1</sup>;  $\lambda = 1.54184$  Å [Cu-K $\alpha$ ]; 13422 reflections measured; 6628 unique reflections ( $R_{\text{int}} = 0.0348$ ); observed  $I > 2\sigma(I)$ . The final  $R_1$  was 0.0412 (observed reflections 0.0541) and  $wR(F^2)$  was 0.0996 (all data 0.1084) for 442 parameters.

**Crystallographic data:** CCDC 1868220 (1), 1868221 (7), and 1868219 (1-8) contain the supplementary crystallographic data for this paper. These data are provided free of charge by The Cambridge Crystallographic Data Centre.

**NOESY data acquisition and processing:** Phase sensitive  $^1\text{H}$ - $^1\text{H}$  NOESY experiments were performed on a 1:1 mixture of components at 50 mM concentration with respect to each component in  $\text{CDCl}_3$ . Spectra were recorded using a 750 ms mixing time with 256 increments, 2048 data points on a JEOL ECA600ii spectrometer at 298 K operating at 14.1 T (600 MHz for  $^1\text{H}$ ).

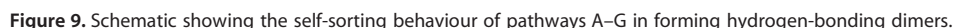
**ITC experiments:** ITC experiments were carried out using Microcal ITC200i instrument (Malvern) at 25 °C in chloroform. 100  $\mu\text{M}$  DAN 3 was present in the cell and titrated with 1 mM AUPy 1 loaded into the syringe using 26 1.4  $\mu\text{L}$  injections with 120 s spacing between the injections. Heats of chloroform dilution was subtracted from each measurement raw data. Data was analysed using Microcal Origin 8 and fitted to a one-binding site model.

**Molecular structure calculations:** The computations were carried out by using the density functional theory-based program Amsterdam Density Functional (ADF) 2017.208<sup>[15]</sup> at the BLYP-D3(BJ)/

TZ2P<sup>[16]</sup> level of theory, which is known to accurately reproduce hydrogen-bond strengths and lengths.<sup>[17]</sup> Solvent effects were accounted for by using the implicit conductor-like screening model (COSMO), in which the solute molecule is surrounded by a dielectric medium.<sup>[18]</sup> All optimised structures have been verified to be true minima (zero imaginary frequencies). The molecular figures were illustrated using CYLview.<sup>[19]</sup> Full computational details are given in the Supporting Information.

**$^1\text{H}$  NMR experiments pairwise analysis:** The one-component and two-component  $^1\text{H}$  NMR experiments were performed on separate samples. The mass of each component was calculated to make a final concentration of 10 mM in 0.6 mL of  $\text{CDCl}_3$  (dried as described previously<sup>[10f]</sup>). The required mass of each component was dissolved in 0.6 mL of  $\text{CDCl}_3$ . The sample was allowed to equilibrate for a minimum of ten minutes before acquisition. The acquired spectra of the one-component and two-component mixtures were compared for changes in the chemical shifts of specific resonances, indicative of hydrogen bonding.

**$^1\text{H}$  NMR experiments cascades:** The mass of each component was calculated to make a final concentration of 10 mM in 0.6 mL of  $\text{CDCl}_3$ . The required mass of each component for the first step of the cascade was dissolved in 0.6 mL of  $\text{CDCl}_3$ . The sample was allowed to equilibrate for a minimum of ten minutes before acquisition. The required mass of the next component in the cascade was added to the same sample. The sample was allowed to equilibrate for a minimum of ten minutes before acquisition. The remaining components were added in the same sequential manner after each acquisition until all the components were added and the  $^1\text{H}$  NMR spectra was acquired for each stage of the cascade.



This work was supported by the EPSRC (EP/KO39292/1). H.C. thanks the University of Leeds for a Leeds Anniversary Research Scholarship. C.F.G. and S.L. thank the Netherlands Organization for Scientific Research (NWO/CW) for financial support.

The authors declare no conflict of interest.

- [1] J. Nelson, *Structure and function in cell signalling*, Wiley, Chichester, **2009**.
- [2] R. Nussinov, C.-J. Tsai, *Cell* **2013**, 153, 293.
- [3] P. Csermely, R. Palotai, R. Nussinov, *Trends Biochem. Sci.* **2010**, 35, 539.

- Chem. Eur. J. 2019, 25, 785–795

- Nitschke, *Angew. Chem. Int. Ed.* **2015**, *54*, 3988; *Angew. Chem.* **2015**, *127*, 4060.
- [7] A. J. Wilson, *Soft Matter* **2007**, *3*, 409.
- [8] D. Wu, N. Potluri, J. Lu, Y. Kim, F. Rastinejad, *Nature* **2015**, *524*, 303.
- [9] M. L. Pellizzaro, K. A. Houton, A. J. Wilson, *Chem. Sci.* **2013**, *4*, 1825.
- [10] a) P. K. Baruah, S. Khan, *RSC Adv.* **2013**, *3*, 21202; b) P. Otte, J. Taubitz, U. Lüning, *Eur. J. Org. Chem.* **2013**, 2130; c) U. Lüning, C. Kuhl, *Tetrahedron Lett.* **1998**, *39*, 5735; d) X. Zhao, X.-Z. Wang, X.-K. Jiang, Y.-Q. Chen, Z.-T. Li, G.-J. Chen, *J. Am. Chem. Soc.* **2003**, *125*, 15128; e) M. L. Pellizzaro, A. M. McGhee, L. C. Renton, M. G. Nix, J. Fisher, W. B. Turnbull, A. J. Wilson, *Chem. Eur. J.* **2011**, *17*, 14508; f) A. Gooch, A. M. McGhee, M. L. Pellizzaro, C. I. Lindsay, A. J. Wilson, *Org. Lett.* **2011**, *13*, 240; g) S. Rayavarapu, S. Kheria, D. R. Shinde, R. G. Gonnade, G. J. Sanjayan, *Org. Biomol. Chem.* **2017**, *15*, 10087; h) S. Kheria, S. Rayavarapu, A. S. Kotmale, G. J. Sanjayan, *Chem. Commun.* **2017**, *53*, 2689; i) S. Kheria, S. Rayavarapu, A. S. Kotmale, R. G. Gonnade, G. J. Sanjayan, *Chem. Eur. J.* **2017**, *23*, 783; j) Y. Zhao, A. L. Connor, T. A. Sobiech, B. Gong, *Org. Lett.* **2018**, *20*, 5486; k) Y. Zhang, R. Cao, J. Shen, C. S. F. Detchou, Y. Zhong, H. Wang, S. Zou, Q. Huang, C. Lian, Q. Wang, J. Zhu, B. Gong, *Org. Lett.* **2018**, *20*, 1555.
- [11] a) T. F. A. de Greef, G. Ercolani, G. B. W. L. Ligthart, E. W. Meijer, R. P. Sijbesma, *J. Am. Chem. Soc.* **2008**, *130*, 13755; b) T. F. A. de Greef, G. B. W. L. Ligthart, M. Lutz, A. L. Spek, E. W. Meijer, R. P. Sijbesma, *J. Am. Chem. Soc.* **2008**, *130*, 5479; c) T. Felder, T. F. A. de Greef, M. M. L. Nieuwenhuizen, R. P. Sijbesma, *Chem. Commun.* **2014**, *50*, 2455.
- [12] F. H. Beijer, R. P. Sijbesma, H. Kooijman, A. L. Spek, E. W. Meijer, *J. Am. Chem. Soc.* **1998**, *120*, 6761.
- [13] W. L. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, *112*, 2008.
- [14] T. J. Murray, S. C. Zimmerman, *J. Am. Chem. Soc.* **1992**, *114*, 4010.
- [15] a) ADF2017, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, <http://www.scm.com>; b) G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931; c) C. Fonseca Guerra, J. G. Snijders, G. te Velde, E. J. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391.
- [16] a) A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098; b) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785; c) E. R. Johnson, A. D. Becke, *J. Chem. Phys.* **2005**, *123*, 024101; d) S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [17] a) C. Fonseca Guerra, T. van der Wijst, J. Poater, M. Swart, F. M. Bickelhaupt, *Theor. Chem. Acc.* **2010**, *125*, 245; b) B. Brauer, M. K. Kesharwani, S. Kozuch, J. M. L. Martin, *Phys. Chem. Chem. Phys.* **2016**, *18*, 20905.
- [18] C. C. Pye, T. Ziegler, *Theor. Chem. Acc.* **1999**, *101*, 396.
- [19] CYLview, 1.0b; C. Y. Legault, Université de Sherbrooke, **2009** (<http://www.cylview.org>).

Manuscript received: September 19, 2018

Accepted manuscript online: October 31, 2018

Version of record online: December 13, 2018