nature chemistry

ARTICLES

PUBLISHED ONLINE: 27 JUNE 2010 | DOI: 10.1038/NCHEM.699

Sequence-selective assembly of tweezer molecules on linear templates enables frameshift-reading of sequence information

Zhixue Zhu, Christine J. Cardin, Yu Gan and Howard M. Colquhoun*

Information storage and processing is carried out at the level of individual macromolecules in biological systems, but there is no reason, in principle, why synthetic copolymers should not be used for the same purpose. Previous work has suggested that monomer sequence information in chain-folding synthetic copolyimides can be recognized by tweezer-type molecules binding to adjacent triplet sequences, and we show here that different tweezer molecules can show different sequence selectivities. This work, based on ^{1}H NMR spectroscopy in solution and on single-crystal X-ray analysis of tweezer–oligomer complexes in the solid state, provides the first clear-cut demonstration of polyimide chain-folding and adjacent-tweezer binding. It also reveals a new and entirely unexpected mechanism for sequence recognition, which, by analogy with a related process in biomolecular information processing, may be termed 'frameshift-reading'. The ability of one particular tweezer molecule to detect, with exceptionally high sensitivity, long-range sequence information in chain-folding aromatic copolyimides is readily explained by this novel process.

The idea that digital information might be encoded at the molecular level as a linear sequence of monomer residues in a copolymer chain first crystallized some fifty years ago with the discovery of the structure[1–3] and function[4] of DNA, a linear, high-molecular-weight organopolyphosphate in which the information represented by its nucleotide monomer sequences was found to ultimately specify the amino-acid sequences of proteins[5]. Viewed purely as digital information, however, DNA—with its four co-monomers—is far more complex than strictly necessary, because even the simplest two-monomer copolymer is the logical equivalent of a string of binary numbers[6]. It is significant that, in biology, no specific information is ever written to DNA. Random copying errors may be captured by evolution, but the key problem—solved in life by the operation of a dauntingly intricate set of molecular machinery—is how co-monomer sequences are to be read[7].

We have recently begun, as have others[8,9], to explore the possibility that synthetic copolymers might, in principle, be used to store and process digital information. Several groups have described the development of small, π-electron-rich, aromatic tweezer molecules[10–18], and we have recently shown that certain molecules of this type can recognize and bind to specific triplet sequences in high-molecular-weight aromatic copolyimides[19–22]. Binding occurs by means of sterically and electronically complementary π–π stacking[23–31] and hydrogen bonding between the tweezer molecules and monomer residues (Fig. 1). As in many supramolecular systems where aromatic π–π stacking plays a key role[32–40], this interaction results in very large, ring-current-induced complexation shifts of

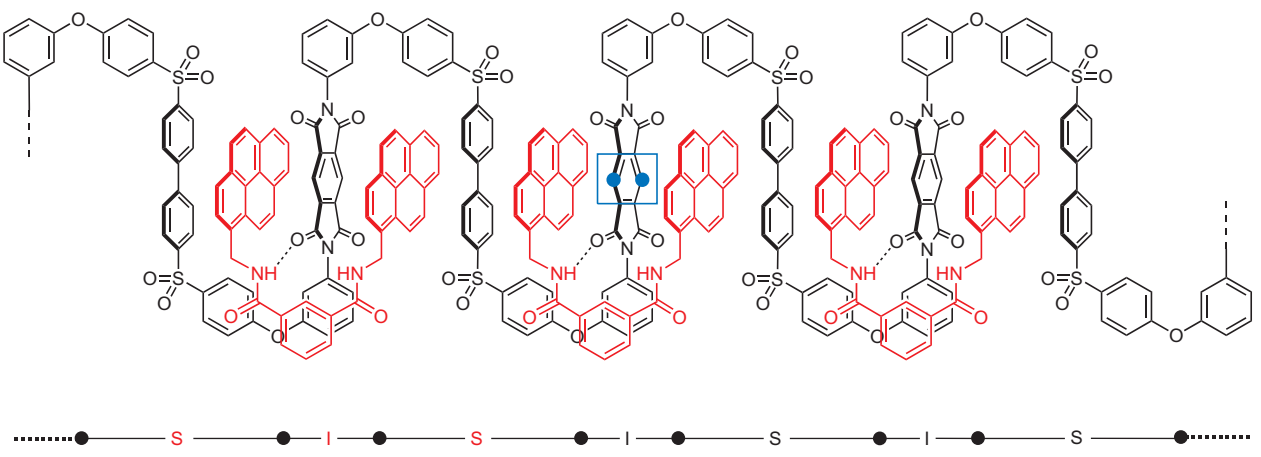
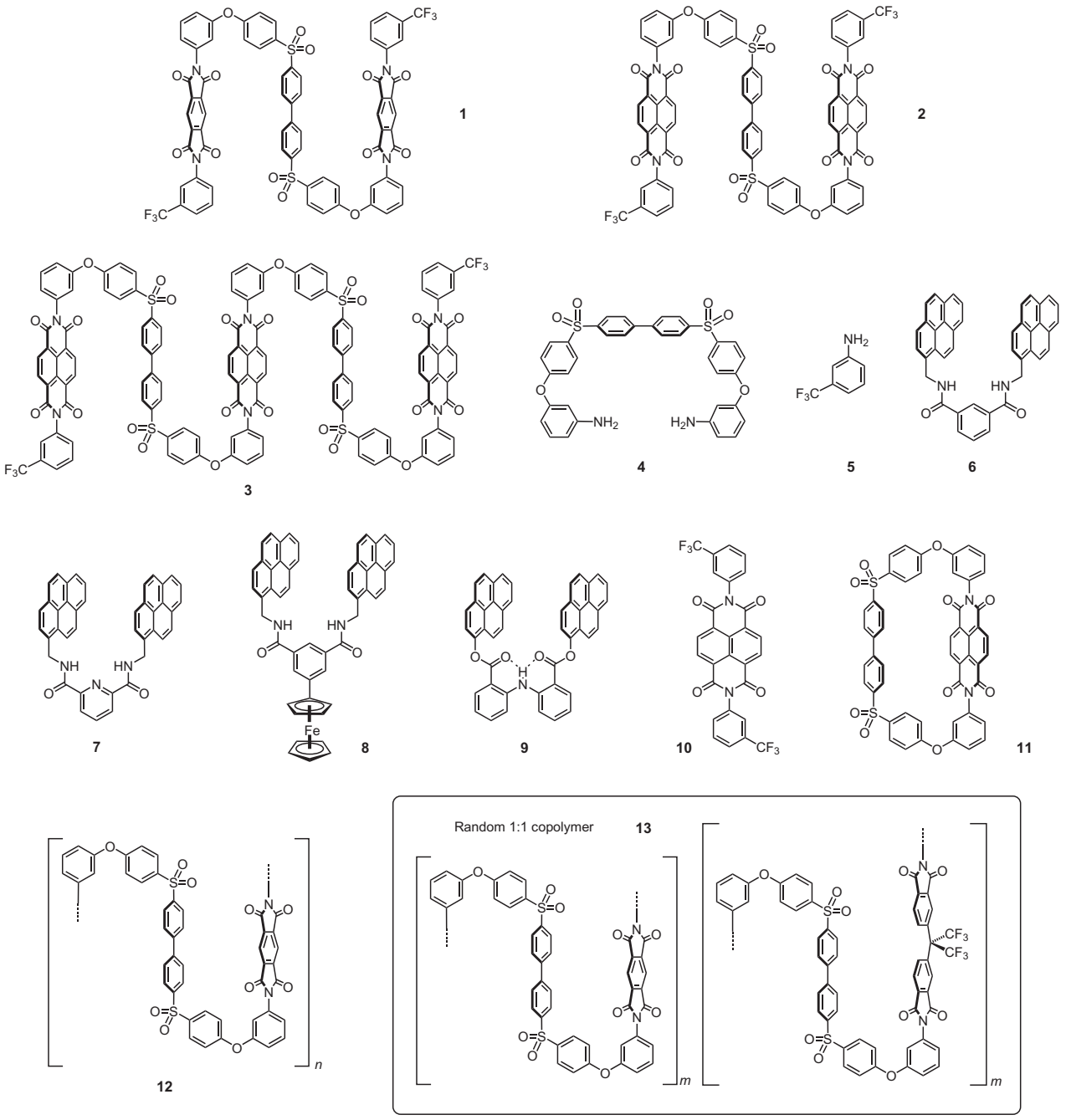


Figure 1 | Proposed[19–21] chain folding of a polyimide-sulfone and multiple tweezer binding at adjacent triplet sequences. Defining the diarylene-pyromellitimide unit as '**I**' and the *bis*(arylsulfonyl)-4,4′-biphenylene residue as '**S**', each tweezer molecule binds to the triplet monomer sequence '**SIS**' through a combination of complementary π–π stacking and NH···O=C hydrogen bonding. (Note that adjacent triplets are overlapping, because each '**S**' residue interacts with two tweezer molecules.).

Department of Chemistry, University of Reading, Whiteknights, Reading RG6 6AD, UK. *e-mail: h.m.colquhoun@rdg.ac.uk



^{1}H NMR resonances arising from the associating components, here involving triplet binding sequences[19,20]. Moreover, NMR data for a range of copolymers show that the tweezer molecules are able to 'read' sequences that are more extended than simple triplets, by a mechanism that we have suggested may involve polymer chain-folding and multiple adjacent-tweezer binding[21]. In the presence of the tweezer, co-monomer sequences in which two, one or no tweezer molecules can be bound at sequences adjacent to the central diimide binding site are found to give three separate diimide resonances[21]. Adjacent-tweezer binding to a chain-folded polyimide (Fig. 1) would produce additional ring-current shielding of the 'observed' diimide protons (highlighted in blue in the figure), so providing a possible mechanism for the detection of extended sequence information.

Here, we report the synthesis and characterization (both in solution and, crystallographically, in the solid state) of discrete complexes between tweezer molecules and fully defined diimide oligomers, both linear and macrocyclic. This work finally validates the hypothesis of polyimide chain-folding and adjacent-tweezer binding, and also reveals a novel mechanism for sequence recognition, which, by analogy with a related process in biomolecular information processing, we term 'frameshift-reading'. This wholly unexpected result explains, for the first time, the ability of one particular tweezer molecule to detect, with extraordinarily high sensitivity (see Supplementary Information), long-range sequence information in chain-folding aromatic copolyimides[22].

Results and discussion

Diimide oligomers **1**, **2** and **3** were synthesized by chemical imidization[41] of mixtures of diamine **4** and the end-capping monoamine **5** with either pyromellitic dianhydride (for **1**) or 1,4,5,8-naphthalenetetracarboxylic dianhydride (for **2** and **3**), in the presence of acetic anhydride. The products were fractionated by column chromatography to afford pure, single oligomers. Attempts to grow

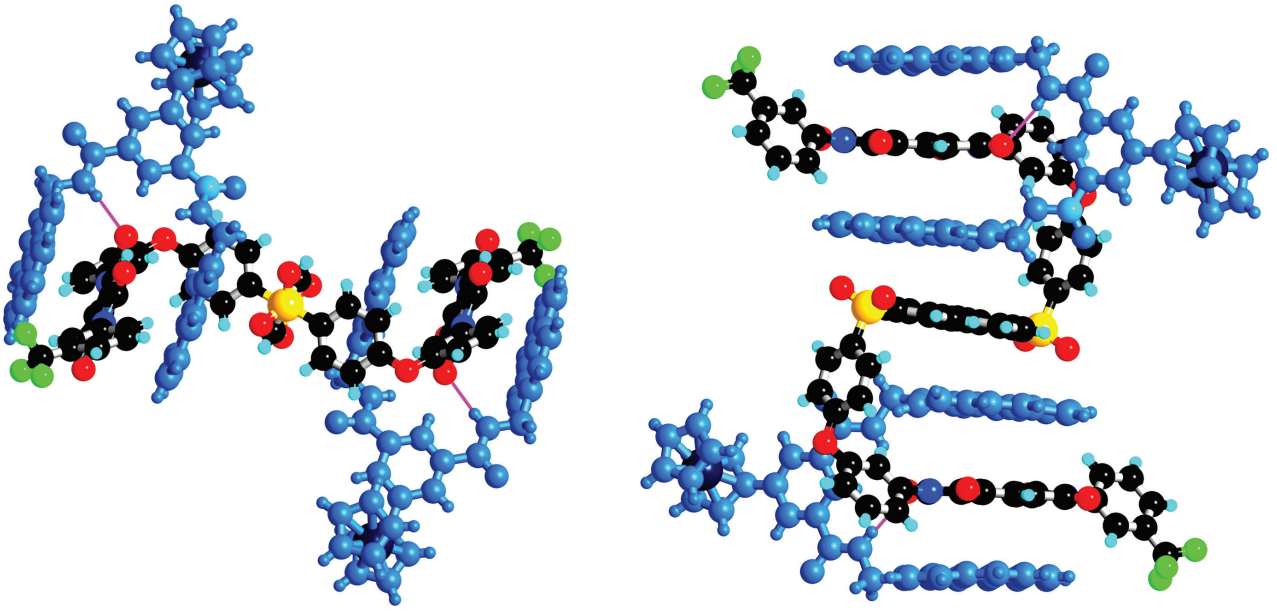


Figure 2 | X-ray structure (two perpendicular views) of a 2:1 complex [8_2+1] between ferrocenyl tweezer molecule 8 and *bis*-pyromellitimide oligomer 1. The two tweezer molecules are shown in blue to distinguish the different components of the assembly. The electron-rich pyrenyl 'arms' of adjacent tweezers interact with electron-poor pyromellitimide and 4,4′-biphenylenedisulfone units in the oligoimide to generate discrete π stacks containing seven alternating donor and acceptor aromatic π systems. Tweezer–oligomer binding also involves an essentially linear hydrogen bond (shown in magenta) between an amide NH group of each tweezer molecule and a pyromellitimide carbonyl-oxygen atom (H···O = 2.26 Å).

diffraction-quality single crystals of complexes between these oligomers and the previously reported tweezer molecules **6** and **7** were unsuccessful, but a new type of tweezer (**8**) containing a ferrocenyl substituent (but with an otherwise similar structure to tweezer molecule **6**) ultimately gave well-formed crystals of a 2:1 complex with oligoimide **1**. The X-ray structure of [$\mathbf{8_2}$**+1**] is shown in Fig. 2, from which it is evident that the previously hypothesized[19–21] adjacent binding of tweezer molecules does indeed occur. In [$\mathbf{8_2}$**+1**], the two tweezer molecules are related by a crystallographic inversion centre, and bind to the terminal diimide residues of the oligomer. Chain folding of the oligomer leads to the formation of a multiple π stack comprising seven electronically complementary components (four electron-rich tweezer arms, two electron-deficient pyromellitimide residues and one electron-deficient 4,4′-biphenylene-disulfone unit). A strong hydrogen bond (N–H···O=C=2.26 Å, ∠N–H···O=173°) is found between each tweezer molecule and a carbonyl group of its bound pyromellitimide residue.

The perpendicular distance in the crystal between each proton of a pyromellitimide residue and the mean plane of a non-adjacent pyrenyl ring (see Fig. 2) is ~10.5 Å, very close to the value of 10.8 Å predicted by an earlier computational modelling study of the corresponding tweezer–polymer complex[21]. This brings the diimide protons well within the '0.1 ppm shielding radius' of the adjacently bound pyrenyl residue, calculated at ~12.0 Å (ref. 42). Consequently, in addition to the large complexation shift of the diimide resonance resulting from tweezer binding at the diimide unit, additional upfield shifts of the diimide resonance are predicted (and observed) when tweezer binding is possible at one or both adjacent sequences in a chain-folded polymer[21]. The crystal structure of [$\mathbf{8_2}$**+1**] thus provides a clear-cut demonstration of the chain folding and adjacent binding proposed for the detection of long-range sequence information by tweezer molecules **6**, **7** and **8**.

Confirmation that the adjacent-tweezer binding observed in the crystal can also occur in solution was provided by Job plots for complexation of tweezer molecule **8** with oligoimides **1** and **2** in chloroform/hexafluoropropan-2-ol (6:1 v/v), using the complexation shifts ($\Delta\delta$) of the singlet diimide resonance in each case. These plots provide very good evidence for 2:1 (tweezer:oligomer) stoichiometry of binding, with peaks found at $x = 0.39$ and 0.32, respectively (cf. theoretical values of 0.33 for 2:1 and 0.50 for 1:1 binding). Remarkably however, Job analyses of complexation between the internally hydrogen-bonded tweezer **9** and oligomers **1** and **2** showed a clear-cut preference for 1:1 binding, with the latter plots peaking at $x = 0.50$ and 0.48, respectively (Fig. 3). The 2:1 tweezer–oligomer binding model established from the structure of complex [$\mathbf{8_2}$**+1**] (Fig. 2) thus fails completely for tweezer molecule **9**. However, computational modelling (molecular mechanics with charge equilibration) quickly showed that 1:1 complexation

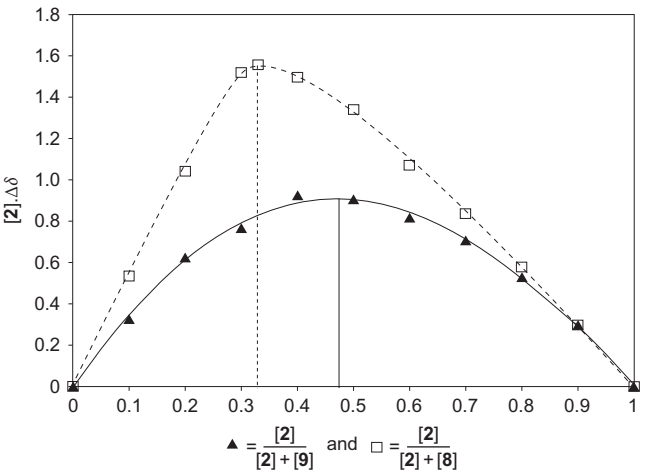


Figure 3 | Job plots for complexation of 8 (dashed line) and 9 (solid line) with *bis*(diimide) oligomer 2, based on ^{1}H NMR complexation shifts of the aromatic diimide resonances in chloroform/hexafluoropropan-2-ol (6:1 v/v). Samples had 6 mM total concentration of tweezer and oligomer. The observed peak abscissa values of 0.32 and 0.48 indicate 2:1 and 1:1 stoichiometries of binding (tweezer/oligomer) for tweezers **8** and **9**, respectively (theoretical values are 0.33 and 0.50).

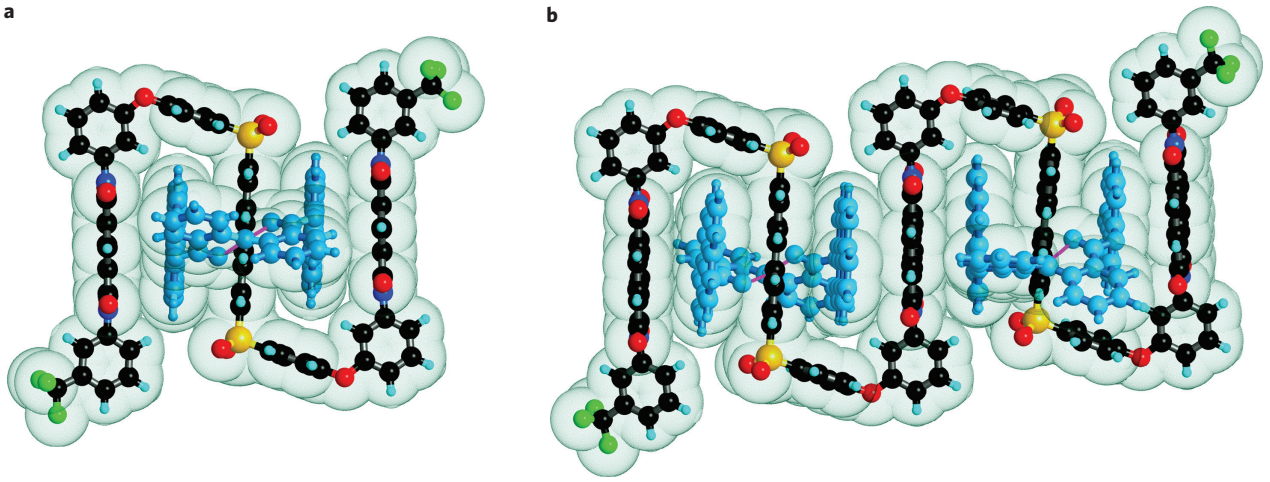


Figure 4 | Minimized computational models (molecular mechanics with charge equilibration) of the 1:1 complex between tweezer molecule 9 and the *bis*-pyromellitimide oligomer 1 (a), and the 2:1 complex between tweezer molecule 9 and the *tris*-diimide oligomer 3 (b). Intramolecular hydrogen bonds are shown in magenta. In complex [$\mathbf{9}_2$+**3**] (**b**) the two tweezer molecules are shown approaching from opposite sides of the oligomer chain, although the system is dynamic, and simultaneously complexed tweezer molecules could in fact be bound *syn* or *anti* to one another. Van der Waals surfaces (2.4 × covalent radii) demonstrate the high degree of shape-complementarity between tweezer molecule **9** and the chain-folded oligoimides. The outer surfaces of the tweezer arms interact with the pyromellitimide units in both models.

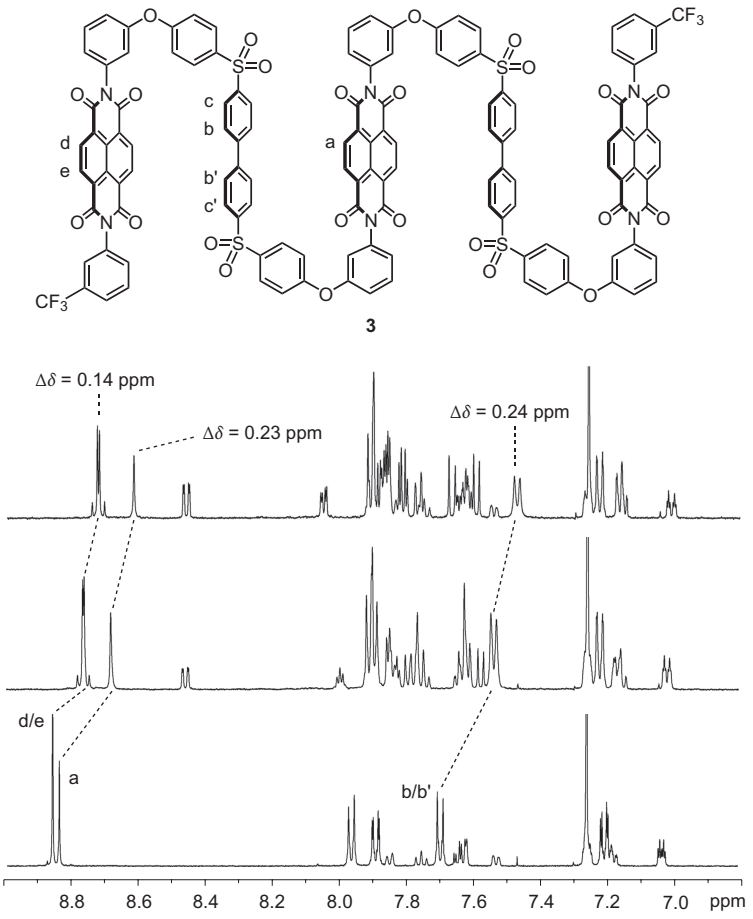


Figure 5 | ^{1}H NMR spectra showing complexation shifts of different resonances for *tris*-diimide oligomer 3 (lower trace) in the presence of tweezer molecule 9, at 1:1 (centre trace) and 1:2 (upper trace) molar ratios, respectively. The relative complexation shifts are entirely consistent with the model shown in Fig. 4b, in that the protons of the central diimide unit (a: $\Delta\delta = 0.23$ ppm) and 4,4′-biphenylene groups (b/b′: $\Delta\delta = 0.24$ ppm) are shielded in the model by the aromatic ring currents of two pyrenyl groups, whereas the protons of the outer diimide units (d/e: $\Delta\delta = 0.14$ ppm) are shielded by only one such group. Note that the chemical inequivalence of protons d and e (formally an AB system) is only revealed on tweezer complexation. Spectra were run at 500 MHz in $CDCl_3$/hexafluoropropan-2-ol (6:1 v/v) at 50 °C and 2 mM total concentration.

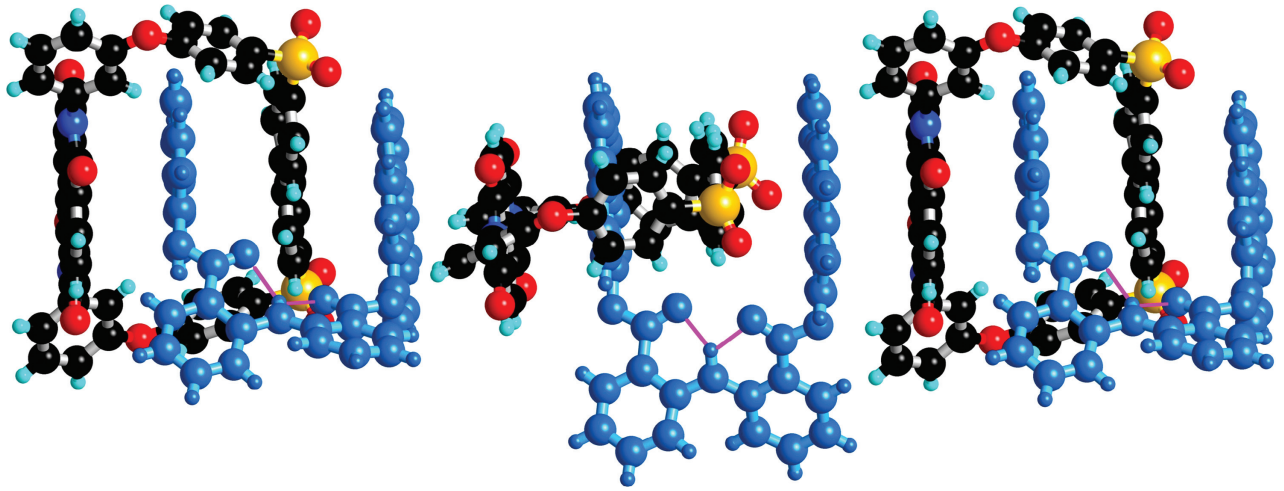


Figure 6 | X-ray structure of the complex between macrocycle 11 and tweezer molecule 9, showing molecular stacking along the crystallographic *a*-direction. Tweezer molecules are shown in blue, emphasizing the alternation of the complementary donor (pyrenyl) and acceptor (diimide and biphenylenedisulfone) subunits in the π stack. Intramolecular hydrogen bonds (N–H···O=C) are shown in magenta. The tweezer molecule clearly binds at the 4,4′-biphenylenedisulfone site, but the outer faces of its arms are both able to π-stack, in the crystal, with strongly electron-accepting diimide units. There is thus a very close analogy between this structure and the model (Fig. 4b) for the binding of tweezer molecule **9** to chain-folding oligomer **3**, and by extension to polymers **12** and **13**.

between tweezer molecule **9** and diimide oligomers **1** and **2** could be accounted for in terms of tweezer binding at the 4,4′-biphenylenedisulfone unit of the oligomer (Fig. 4a). This gives rise to a five-component, complementary π stack in which the diimide residues interact with the outer faces of the tweezer arms rather than with their inner faces as in complex [$\mathbf{8}_2$+**1**] (Fig. 2).

A closely related structure was found, computationally, for the complex between the *tris*-diimide oligomer **3** and tweezer molecule **9** (Fig. 4b), where two molecules of the tweezer bind at consecutive, biphenylene-centred '**ISI**' sequences (Fig. 1). This model leads to the prediction that the ^{1}H NMR resonances of oligomer **3** associated with the biphenylene units and the central diimide residue (all of which are shielded by two tweezer–pyrenyl groups) should show significantly greater complexation shifts ($\Delta\delta$) with tweezer **9** than those arising from the terminal diimide groups, which are shielded by only a single pyrene unit. As shown in Fig. 5, this prediction is borne out in practice, with $\Delta\delta$ values (2:1 molar ratio of tweezer molecule **9** to the *tris*-diimide oligomer **3**) of 0.23 and 0.24 ppm for protons H_a and $H_{b/b'}$ contrasting with a $\Delta\delta$ value of only 0.14 ppm for the AB system representing the protons ($H_{d/e}$) of the terminal diimide units. The spectra shown in Fig. 5 were obtained at 50 °C (fast exchange conditions), because room-temperature spectra were significantly broadened, suggesting an approach to slow exchange between bound and unbound species.

The different sequence selectivities of tweezer molecules **8** (**SIS**) and **9** (**ISI**) may be accounted for on the basis of their different hydrogen-bonding characteristics. Historically, tweezer-type molecules have invariably been found to bind small aromatic guests through a mechanism analogous to chelation in coordination chemistry, that is, by using the inner surfaces of their tweezer arms[10–16]. However, when the guest molecule is a chain-folding oligomer or polymer, it is entirely feasible that—depending on the detailed geometrical parameters and conformational flexibility of the tweezer—π stacking of the diimide residues to the outer faces of the tweezer arms could be preferred. Energetically, it makes no difference which face of the pyrenyl residue interacts with the diimide. Indeed, in the absence of other types of interaction, outer-face complexation might even be the norm, and it is clearly favoured for the intramolecularly hydrogen-bonded diester tweezer **9** (Figs 4– 6), leading to complexation at the sequence '**ISI**'. In contrast, the diamide tweezer molecule **8** and analogues such as **6** and **7** have two convergent N–H groups, each of which can form an intermolecular hydrogen bond to the carbonyl group of a diimide residue in a poly- or oligoimide chain (Fig. 2). This would obviously promote tweezer binding at the diimide unit, now leading to complementary π–π stacking between the diimide ring and the inner surfaces of the tweezer arms. As a result, the diamide-type tweezers invariably show selective binding at the sequence '**SIS**' (refs 19,20,21).

We were unable to isolate diffraction-quality crystals of the proposed 1:1 oligoimide complex [**1**+**9**], but the modelled biphenylene-centred binding mode was indeed identified for tweezer molecule **9** by a single-crystal X-ray study of the 1:1 complex

Table 1 | Graphical representation of frameshift-reading resulting from the different sequence-recognition characteristics of tweezer molecules 6, 7, 8 and 9.

Tweezer molecule	Monomer sequence	Pyrene units	
		Proximal	Distal
6, 7, 8	S I S I S I S	2	2
9	S I S I S I S	2	2
6, 7, 8	S I S I S F S	2	1
9	S I S I S F S	1	1
6, 7, 8	S F S I S F S	2	0
9	S F S I S F S	0	0

The 'observed' protons at the central diimide unit (in green) in the first sequence are shielded to a similar extent by the pyrenyl residues of both types of tweezer (two proximal and two distal pyrenes). However, with tweezer molecule **9**, the presence of a non-binding hexafluoroisopropylidene-diimide (**F**) unit at one or both adjacent sites has a more drastic effect on the number of pyrenes shielding the diimide protons. As a result, the chemical shifts of the three sequences are much more strongly differentiated by aromatic ring-current shielding in the presence of **9** than **6**, **7** or **8**.

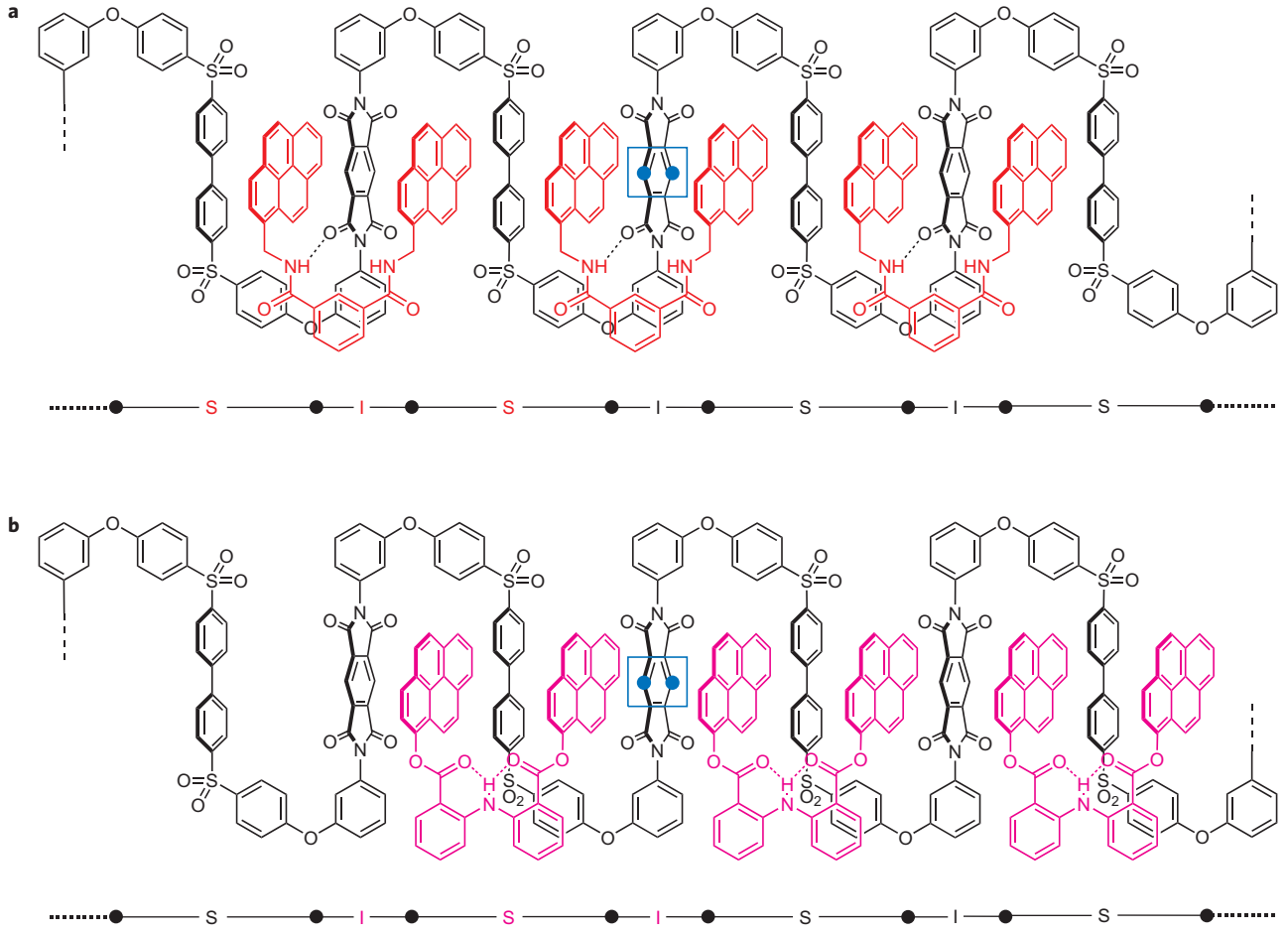


Figure 7 | Chain folding and multiple binding to different polyimide triplet sequences by different tweezer molecules. Defining the diarylpyromellitimide unit as '**I**' and the *bis*(arylsulfonyl)biphenylene residue as '**S**', each tweezer molecule binds either to monomer sequence '**SIS**' (tweezer **6**, Fig. 7a) or to **ISI** (tweezer **9**, Fig. 7b). The tweezer–polymer hydrogen-bonding interaction (NH···O=C) found with diamide tweezers such as **6** and **8**, but not with the diester tweezer **9**, is possible only for binding at sequence **SIS**. It should be noted that, following the frameshift, the right-hand tweezer molecule in the lower diagram is too distant from the 'observed' protons (shown in blue) to play any role in ring-current shielding.

between **9** and an analogous macrocyclic oligoimide **11**. Here, in contrast to all previous tweezer–diimide assemblies for which structural data are available[19–21,43], the imide unit of **11** does not represent the central binding site for the tweezer molecule, which instead binds at the 4,4′-biphenylene–disulfone unit. However, as with tweezer molecule **6**, where binding to macrocycle **11** occurs at the diimide residue[43], the new macrocycle–tweezer complex [**9**+**11**] packs to generate an 'infinite', complementary stack in which π-electron-deficient diimide residues alternate with π-electron-rich pyrenyl units (Fig. 6).

Conceptually, ring-opening polymerization of the macrocycles in such an infinite stack would generate a tweezer-bound, chain-folded homopolymer, with the tweezer molecules located on the 4,4′-biphenylenedisulfone units ('**S**'). The diimide residues ('**I**') would again stack onto the outer surfaces of the tweezer arms. The structure of [**9**+**11**] thus confirms that, in a chain-folded polymer, tweezer **9** favours binding to the sequence '**ISI**' rather than to '**SIS**'. Moreover, further analysis shows that this particular sequence selectivity provides an immediate explanation for the exceptional power of **9** to reveal long-range sequence information in the ^{1}H NMR spectra of high-molecular-weight copolyimides[22].

Consider the 1:1 copolymer **13**, in which the hexafluoroisopropylidene-diphthalimide groups generate entirely non-tweezer-binding sequences. Here, the steric bulk of the hexafluoroisopropylidene unit and the twisted orientation of the two connected phthalimide groups combine to ensure that direct tweezer binding cannot take place at this residue. Specifying the pyromellitimide unit as '**I**', the hexafluoroisopropylidene-diphthalimide group as '**F**', and the biphenylene-disulfonyl-diamine residue as '**S**', the allowed septet sequences in this copolymer (centring on the pyromellitimide unit observed by ^{1}H NMR) are then [–**SISISIS**–], [–**SFSISIS**– or –**SISISFS**–] and [–**SFSISFS**–]. The structure of complex [$\mathbf{8}_2$+**1**] reported here, together with previous spectroscopic studies[21], establishes that these sequences can bind three, two and one tweezer molecule of type **8**, respectively, at their **SIS** triplet sequences (note that the '**S**' units are shared or 'overlapping' under this terminology). Tweezer binding at **SIS** results in large upfield shifts (>2.5 ppm) for the central diimide resonance in each case, but additional shielding from the adjacently bound tweezer(s) present produces further (small) complexation shifts, so resolving a different diimide resonance for each different septet sequence. The key point is that the central diimide residue can bind a molecule of type **8** in all three sequences of this type (Table 1).

Conversely, tweezer **9** binds predominantly at the sequence '**ISI**' and only very weakly, if at all, at the directionally degenerate sequences **FSI** and **ISF**. As a result, the central diimide protons of the septet sequences are now much more sharply differentiated in terms of their susceptibility to ring-current shielding by complexing tweezer molecules. As shown for **9** in Table 1, two tweezer molecules bind to [–**SISISIS**–], one to [–**SFSISIS**– or –**SISISFS**–] and none to [–**SFSISFS**–]. Consequently, interaction with **9** means that the

central diimide residues in these septet sequences are shielded directly by two, one and no pyrenyl units, respectively. This is a very different situation from that described for complexation of **8** to the same sequences, where the central, 'observed' diimide protons are shielded directly by two pyrenyl residues in all cases. The novel sequence selectivity discovered in this work for tweezer molecule **9** (that is, '**ISI**' rather than '**SIS**') thus provides a simple explanation for the extraordinarily high sensitivity to this tweezer of ^{1}H NMR resonances associated with long-range sequence information[22]. Conversely, the 'anomalous' behaviour of copolyimide ^{1}H NMR resonances in the presence of tweezer molecule **9** can now be seen, in conjunction with Table 1, to provide strong evidence that this molecule does indeed display a general preference for binding to the triplet sequence '**ISI**' rather than '**SIS**'.

In reality the situation is not quite so clear-cut, because even the diimide resonance arising from the 'non-**9**-binding' sequence [–**SFSISFS**–] does undergo a small upfield shift in the presence of this tweezer. It is already known that **9** can bind weakly to a diimide residue[22], and because the NMR experiments described here all show fast exchange between bound and unbound tweezer molecules, it seems that the strong '**ISI**' complexation proposed for tweezer **9** will be dynamically superimposed on much weaker '**SIS**' binding, with both mechanisms operating simultaneously on the NMR timescale. Consistent with this, the 1:1 association constant K_a, measured (using the UV–Vis dilution method, based on the charge-transfer band at 551 nm) for the binding of tweezer molecule **9** to the chain-folding oligomer **2**, was found in this work to be 780 M^{-1}, more than five times higher than the value (140 M^{-1})[22] of K_a for the binding of **9** to the simple diimide **10**. In the context of this 500% increase in binding constant, it should be noted that the experimental error in binding constants determined by the UV–Vis dilution method is normally of the order of 15% (ref. 44).

We have thus established, through studies of fully defined oligomers and their tweezer complexes, that two different designs of tweezer molecules (exemplified by **8** and **9**) have structures that are complementary to two different triplet sequences in chain-folding polyimide sulfones. This situation is faintly reminiscent of the relationship between the different tRNAs (each having a different 'binding codon') and mRNA (ref. 45). Correspondingly, homopolyimide **12** can be represented as the sequence . . .**SISISISISIS**. . ., with tweezer molecules **6**, **7** and **8** binding to the triplets '**SIS**', while the present work shows that tweezer molecule **9** binds preferentially to the sequence '**ISI**', so that the 'reading frame' for **9** can be thought of as being shifted to the left or right by a single monomer residue (Fig. 7).

Methods

Materials and instrumentation. Synthetic procedures were carried out under an atmosphere of dry nitrogen, unless otherwise specified. Commercial solvents and reagents were used without purification, unless otherwise stated. *N*,*N*-Dimethylacetamide (DMAc) was distilled over calcium hydride before use. Diamine **4** (ref. 46), tweezer molecules **6** (ref. 43), **7** (ref. 43) and **9** (ref. 22), model imide **10** (ref. 21), macrocycle **11** (ref. 46), homopolymer **12** (ref. 19) and copolymer **13** (ref. 21) were prepared according to literature procedures. Proton and ^{13}C NMR spectra were recorded on Bruker AV-700 and DPX 250 MHz spectrometers, respectively. Computational modelling (molecular mechanics with charge equilibration, *Cerius*2, Accelrys, San Diego) was carried out on an SGI-O2 workstation using a reparametrized version of the Dreiding-II force-field. Single-crystal X-ray data for complexes [$\mathbf{8}_2$+**1**] and [**9**+**11**] were measured on an Oxford Diffraction X-Calibur CCD diffractometer using Cu-Kα radiation. Structure solution and refinement were carried out using the SHELXS-97 suite of programmes.

Synthesis of oligoimides. The synthesis and isolation of oligomer **1** is given as an example. A mixture of pyromellitic dianhydride (0.436 g, 20.0 mmol), diamine **4** (0.648 g, 10.0 mmol) and 3-(trifluoromethyl)aniline **5** (0.646 g, 40.0 mmol) in DMAc (120 ml) was stirred at room temperature for 0.5 h to give a clear solution. To this was added a mixture of acetic anhydride (1 ml) and pyridine (0.5 ml), and the reaction mixture was heated at 100 °C under dry nitrogen for 14 h, before cooling to room temperature and pouring into methanol. The precipitate was filtered, washed with methanol and dried at 80 °C for 3 h. The crude product was purified by column chromatography (dichloromethane:methanol 99:1, v/v), giving **1** as a pale yellow, crystalline solid (0.36 g, 27%). M.p. 309 °C; ^{1}H NMR (700 MHz, $CDCl_3$:hexafluoropropan-2-ol 6:1, v/v) δ (ppm) = 8.49 (s, 4H), 7.97 (d, J = 8.6 Hz, 4H), 7.90 (d, J = 9.0 Hz, 4H), 7.77 (d, J = 7.9 Hz, 2H), 7.75 (s, 2H), 7.71 (d, J = 8.5 Hz, 4H), 7.70 (t, J = 7.9 Hz, 2H), 7.66 (d, J = 7.9 Hz, 2H), 7.60 (t, J = 8.1 Hz, 2H), 7.33 (dm, 2H), 7.19 (dm, 2H), 7.18 (t, J = 2.1 Hz, 2H), 7.17 (d, J = 9.0 Hz, 4H); ^{13}C NMR (62.5 MHz, $CDCl_3$:hexafluoropropan-2-ol 6:1, v/v) δ (ppm) = 165.7, 165.5, 162.4, 155.8, 144.9, 140.9, 137.4, 137.3, 134.3, 132.8, 132.3, 131.5, 131.3, 130.5, 130.4, 129.9, 128.9, 128.4, 126.3, 123.7, 123.6, 123.4, 121.3, 120.1, 118.9, 118.8; IR (Nujol): 1,774, 1,713 (imide νC = O), 1,369 (νC–N), 1,241 (νC–O–C), 1,107 cm^{-1}; MS (MALDI-TOF): m/z = 1,357. Calcd. for $[C_{70}H_{36}F_6N_4O_{14}S_2 + Na]^+$ 1,357. Analysis (calcd., found for $C_{70}H_{36}F_6N_4O_{14}S_2$): C (62.97, 62.64), H (2.72, 2.88), N (4.19, 4.12%). Oligomers **2** and **3** were synthesized and isolated by an analogous method, replacing pyromellitic dianhydride with 1,4,5,8-naphthalenetetracarboxylic dianhydride.

Synthesis of the ferrocenyl tweezer molecule 8. A suspension of 5-bromoisophthalic acid (0.245 g, 1.00 mmol) in thionyl chloride (10 ml) was stirred at reflux under dry nitrogen for 2 h. Evaporation of excess thionyl chloride under reduced pressure gave an off-white solid. This residue was dissolved in dichloromethane (50 ml), and 1-pyrenemethylamine hydrochloride (0.536 g, 2.00 mmol) and triethylamine (1 ml) were added to the solution. The mixture was stirred at room temperature overnight, and the resulting solid was filtered, washed with water (3 × 30 ml) and methanol (2 × 40 ml), and dried to yield a cream crystalline solid (0.659 g, 95%). To a suspension of this bromo-substituted tweezer molecule (0.336 g, 0.500 mmol) in toluene (500 ml) was added tetrakis(triphenylphosphine)palladium(0) (0.029 g, 0.025 mmol), ferroceneboronic acid (0.120 g, 0.520 mmol) and sodium carbonate (0.106 g, 1.00 mmol) in water (5 ml). The mixture was heated under reflux for 2 days, cooled and diluted with toluene (200 ml) and water (200 ml). The organic phase was washed with 5% HCl, then with water, dried over $MgSO_4$, and evaporated to dryness. The crude product was purified by column chromatography (dichloromethane as eluent) to give **8** as a yellow, crystalline solid (0.284 g, 73% yield), m.p. 243 °C; ^{1}H NMR (250 MHz, DMSO-d_6) δ (ppm) = 9.41 (t, J = 5.4 Hz, 2H), 8.54 (d, J = 9.3 Hz, 2H), 8.33–8.25 (m, 9H), 8.21 (d, J = 1.5 Hz, 2H), 8.16 (s, 4H), 8.15–8.04 (m, 4H), 5.28 (d, J = 5.4 Hz, 4H), 4.88 (s, 2H), 4.38 (s, 2H), 4.01 (s, 5H); ^{13}C NMR (62.5 MHz, DMSO-d_6) δ (ppm) = 166.2, 140.3, 135.0, 133.2, 131.1, 130.7, 130.5, 128.5, 128.0, 127.8, 127.4, 127.3, 127.1, 126.6, 125.6, 125.5, 125.1, 124.6, 124.4, 124.3, 123.6, 83.7, 69.8, 69.7, 67.0, 43.5; IR (Nujol): 1,620 (νC = O), 1,376 (νC-N), 845 cm^{-1}; MS (MALDI-TOF): m/z = 777. Calcd. for $[C_{52}H_{36}FeN_2O_2 + H]^+$, 777; for $[C_{52}H_{36}FeN_2O_2 + Na]^+$, calcd. 799; found 799; Analysis (calcd., found for $C_{52}H_{36}FeN_2O_2$): C (80.41, 79.95), H (4.67, 4.66), N (3.61, 3.59%).

Received 13 January 2010; accepted 10 May 2010; published online 27 June 2010

References

1. Watson, J. D. & Crick, F. H. C. A structure for deoxyribose nucleic acid. *Nature* **171,** 737–738 (1953).
2. Wilkins, M. H. F., Stokes, A. R. & Wilson, H. R. Molecular structure of deoxypentose nucleic acids. *Nature* **171,** 738–740 (1953).
3. Franklin R. E. & Gosling, R. G. Molecular configuration in sodium thymonucleate. *Nature* **171,** 740–741 (1953).
4. Crick, F. H. C. On protein synthesis. *Symp. Soc. Exptl Biol.* **12,** 138–163 (1958).
5. Nirenberg, M. Historical review: deciphering the genetic code—a personal account. *Trends Biochem. Sci.* **29,** 46–54 (2004).
6. Dawkins, C. R. *The Blind Watchmaker* (Longmans, 1986).
7. Lodish, H. *et al. Molecular Cell Biology* 6th edn (Freeman, 2007).
8. Burd, C. & Weck, M. Self-sorting in polymers. *Macromolecules* **38,** 7225–7230 (2005).
9. Weck, M. Side-chain functionalized supramolecular polymers. *Polym. Int.* **56,** 453–460 (2007).
10. Harmata, M. Chiral molecular tweezers. *Acc. Chem. Res.* **37,** 862–873 (2004).
11. Klärner, F.-G. & Kahlert, B. Molecular tweezers and clips as synthetic receptors. Molecular recognition and dynamics in receptor−substrate complexes. *Acc. Chem. Res.* **36,** 919–932 (2003).
12. Chen, C. W. & Whitlock, H. W. Molecular tweezers—a simple model of bifunctional intercalation. *J. Am. Chem. Soc.* **100,** 4921–4922 (1978).
13. Zimmerman, S. C. & VanZyl, C. M. Rigid molecular tweezers—synthesis, characterization and complexation chemistry of a diacridine. *J. Am. Chem. Soc.* **109,** 7894–7896 (1987).
14. Kurebayashi, H., Haino, T., Usui, S. & Fukazawa, Y. Structure of supramolecular complex of flexible molecular tweezers and planar guest in solution. *Tetrahedron* **57,** 8667–8674 (2001).
15. Balzani, V. *et al.* Host−guest complexes between an aromatic molecular tweezer and symmetric and unsymmetric dendrimers with a 4,4′-bipyridinium core. *J. Am. Chem. Soc.* **128,** 637–648 (2006).

16. Klärner, F.-G. *et al.* Molecular tweezer and clip in aqueous solution: unexpected self-assembly, powerful host−guest complex formation, quantum chemical H-1 NMR shift calculation. *J. Am. Chem. Soc.* **128,** 4831–4841 (2006).
17. Goshe, A. J., Steele, I. M., Ceccarelli, C., Rheingold, A. L. & Bosnich, B. Supramolecular recognition: on the kinetic lability of thermodynamically stable host−guest association complexes. *Proc. Natl Acad. Sci. USA* **99,** 4823–4829 (2002).
18. Peng, X.-X., Lu, H.-Y., Han, T. & Chen, C.-F. Synthesis of a novel triptycene-based molecular tweezer and its complexation with paraquat derivatives. *Org. Lett.* **9,** 895–898 (2007).
19. Colquhoun, H. M. & Zhu, Z. Recognition of polyimide sequence information by a molecular tweezer. *Angew. Chem. Int. Ed.* **43,** 5040–5045 (2004).
20. Colquhoun, H. M., Zhu, Z., Cardin, C. J. & Gan, Y. Principles of sequence-recognition in aromatic polyimides. *Chem. Commun.* 2650–2652 (2004).
21. Colquhoun, H. M., Zhu, Z., Cardin, C. J., Gan Y. & Drew, M. G. B. Sterically controlled recognition of macromolecular sequence information by molecular tweezers. *J. Am. Chem. Soc.* **129,** 16163–16174 (2007).
22. Colquhoun, H. M., Zhu, Z., Cardin, C. J., Drew, M. G. B. & Gan, Y. Recognition of sequence-information in synthetic copolymer chains by a conformationally-constrained tweezer molecule. *Faraday Discuss.* **143,** 205–220 (2009).
23. Hamilton, D. G., Sanders, J. K. M., Davies, J. E., Clegg, W. & Teat, S. J. Neutral [2]catenanes from oxidative coupling of π-stacked components. *Chem. Commun.* 897–898 (1997).
24. Hamilton, D. G., Lynch, D. E., Byriel, K. A., Kennard C. H. L. & Sanders, J. K. M. *N*-alkylation of pyromellitic diimide: solid-state structure of the 1:1 *N*,*N*′-diethylpyromellitic diimide 2,6-dimethoxynaphthalene cocrystal. *Aust. J. Chem.* **51,** 441–444 (1998).
25. Hamilton, D. G., Davies, J. E., Prodi, L. & Sanders, J. K. M. Synthesis, structure and photophysics of neutral π-associated [2]catenanes. *Chem. Eur. J.* **4,** 608–620 (1998).
26. Hamilton, D. G., Feeder, N., Teat, S. J. & Sanders, J. K. M. Reversible synthesis of π-associated [2]catenanes by ring-closing metathesis: towards dynamic combinatorial libraries of catenanes. *New J. Chem.* **22,** 1019–1021 (1998).
27. Kaiser, G. *et al.* Lithium-templated synthesis of a donor−acceptor pseudorotaxane and catenane. *Angew. Chem. Int. Ed.* **43,** 1959–1962 (2004).
28. Vignon, S. A. *et al.* Switchable neutral bistable rotaxanes. *J. Am. Chem. Soc.* **126,** 9884–9885 (2004).
29. Iijima, T. *et al.* Controllable donor−acceptor neutral [2]rotaxanes. *Chem. Eur. J.* **10,** 6375–6392 (2004).
30. Pascu, S. I. *et al.* Cation-reinforced donor−acceptor pseudorotaxanes. *New J. Chem.* 80–89 (2005).
31. Hansen, J. G., Bang, K. S., Thorup N. & Becher, J. Donor−acceptor macrocycles incorporating tetrathiafulvalene and pyromellitic diimide: syntheses and crystal structures. *Eur. J. Org. Chem.* 2135–2144 (2000).
32. Colquhoun, H. M., Stoddart, J. F., Williams, D. J., Wolstenholme, J. B. & Zarzycki, R. Second-sphere coordination of cationic platinum complexes by crown ethers—the X-ray crystal-structure of $[Pt(bpy)(NH_3)_2$.dibenzo[30]crown-$10]_2^+[PF_6]_2^-.xH_2O$ (x=0.6). *Angew. Chem. Int. Ed. Engl.* **20,** 1051–1053 (1981).
33. Colquhoun, H. M. *et al.* Complex-formation between dibenzo-3*n*-crown-*n* ethers and the diquat dication. *J. Chem. Soc., Chem. Commun. (pre-1996)* 1140–1142 (1983).
34. Allwood, B. L., Spencer, N., Shahriari-Zavareh, H., Stoddart, J. F. & Williams, D. J. Complexation of diquat by a bisparaphenylene-34-crown-10 derivative. *J. Chem. Soc., Chem. Commun. (pre-1996)* 1061–1064 (1987).
35. Ashton, P. R., Slawin, A. M. Z., Spencer, N., Stoddart, J. F. & Williams, D. J. Complex formation between bisparaphenylene-(3*n*+4)-crown-*n* ethers and the paraquat and diquat dications. *J. Chem. Soc., Chem. Commun. (pre-1996)* 1066–1069 (1987).
36. Krebs, F. C. & Jørgensen, M. A new versatile receptor platform. *J. Org. Chem.* **66,** 6169–6173 (2001).
37. Koshkakaryan, G. *et al.* Alternative donor−acceptor stacks from crown ethers and naphthalene diimide derivatives: rapid, selective formation from solution and solid state grinding. *J. Am. Chem. Soc.* **131,** 2078–2079 (2009).
38. Peinador, C., Blanco V. & Quintela, J. A new doubly interlocked [2]catenane. *J. Am. Chem. Soc.* **131,** 920–921 (2009).
39. Ulrich, S. & Lehn, J.-M. Adaptation and optical signal generation in a constitutional dynamic network. *Chem. Eur. J.* **15,** 5640–5645 (2009).
40. Lohr, A., Grüne, M. & Würthner, F. Self-assembly of bis(merocyanine) tweezers into discrete bimolecular π-stacks. *Chem. Eur. J.* **15,** 3691–3705 (2009).
41. Harris, F. W. in *Polyimides* (eds Wilson, D., Stenzenberger, H. D. & Hergenrother, P. M.) Ch. 1 (Chapman and Hall, 1990).
42. Klod, S. & Kleinpeter, E. *Ab initio* calculation of the anisotropy effect of multiple bonds and the ring current effect of arenes—application in conformational and configurational analysis. *Perkin Trans.* 1893–1898 (2001).
43. Colquhoun, H. M., Zhu, Z. & Williams, D. J. Extreme complementarity in a macrocycle−tweezer complex. *Org. Lett.* **5,** 4353–4356 (2003).
44. Nielsen, M. B. , *et al.* Binding studies between tetrathiafulvalene derivatives and cyclobis(paraquat-*p*-phenylene). *J. Org. Chem.* **66,** 3559–3563 (2001).
45. Johnson, Z. I. & Chisholm, S. W. Properties of overlapping genes are conserved across microbial genomes. *Genome Res.* **14,** 2268–2272 (2004).
46. Colquhoun, H. M., Williams, D. J. & Zhu, Z. Macrocyclic aromatic ether-imide-sulfones: versatile supramolecular receptors with extreme thermochemical and oxidative stability. *J. Am. Chem. Soc.* **124,** 13346–13347 (2002).

Acknowledgements

This research was supported by EPSRC (grant nos EP/C533526/1, EP/E00413X/1, EP/F013663/1 and EP/G026203/1) and by the Royal Society (a Senior Research Fellowship to H.M.C).

Author contributions

Z.Z. and H.M.C. designed the synthetic and spectroscopic experiments and interpreted the resulting data. Z.Z. carried out the experiments and co-wrote the paper. C.J.C. and Y.G. undertook the crystallographic work, including structure solution, refinement and data analysis. H.M.C. conceived and supervised the project, generated the graphics and wrote the paper.

Additional information

The authors declare no competing financial interests. Supplementary information and chemical compound information accompany this paper at www.nature.com/naturechemistry. Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/. Correspondence and requests for materials should be addressed to H.M.C.