JACS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

Cite This: J. Am. Chem. Soc. 2018, 140, 11526–11536

pubs.acs.org/JACS

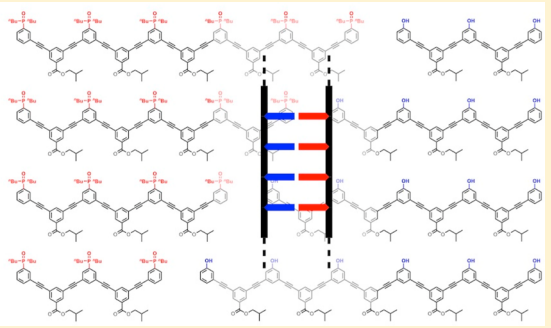
H-Bonded Duplexes based on a Phenylacetylene Backbone

Jonathan A. Swain, Giulia Iadevaia, and Christopher A. Hunter*

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Supporting Information

ABSTRACT: Complementary phenylacetylene oligomers equipped with phenol and phosphine oxide recognition sites form stable multiply H-bonded duplexes in toluene solution. Oligomers were prepared by Sonogashira coupling of diiodobenzene and bis-acetylene building blocks in the presence of monoacetylene chain terminators. The product mixtures were separated by reverse phase preparative high-pressure liquid chromatography to give a series of pure oligomers up to seven recognition units in length. Duplex formation between length complementary homo-oligomers was demonstrated by ^{31}P NMR denaturation experiments using dimethyl sulfoxide as a competing H-bond acceptor. The denaturation experiments were used to determine the association constants for duplex formation, which increase by nearly 2 orders of magnitude for every phenol-phosphine oxide base-pair added. These experiments show that the phenylacetylene backbone supports formation of extended duplexes with multiple cooperative intermolecular H-bonding interactions, and together with previous studies on the mixed sequence phenylacetylene 2-mer, suggest that this supramolecular architecture is a promising candidate for the development of synthetic information molecules that parallel the properties of nucleic acids.



INTRODUCTION

Nucleic acids encode information in the sequence of monomer units, and this structure provided the molecular basis for the evolution of biological life. The information is read through sequence-selective duplex formation and copied through template synthesis.[1] The sequence of monomer units also defines the three-dimensional structures and functional properties of single-stranded nucleic acids.[2] Modified analogues of nucleic acids have been reported,[3] where the sugar,[4] the phosphate linker,[5] or the base pairing system[6] have been replaced and the ability of forming duplexes was not affected, suggesting that fully synthetic information molecules may be able to have some or all the functions of nucleic acids. There are some examples of synthetic oligomers that form duplexes using metal coordination, salt bridges, or H-bonding as the base-pairing interactions.[7−11] Figure 1a shows a blueprint for duplex forming oligomers that is currently being investigated in our laboratory. The modular nature of this design allows the three different components to be explored independently, providing insights into the key requirements for architectures that will lead to sequence-selective duplex formation.

We have shown that oligomers equipped with single H-bond recognition modules form stable duplexes in nonpolar solvents.[12a−g] It is possible to interchange different H-bond donor−acceptor recognition motifs (D·A) on the same backbone and maintain the duplex forming properties of the oligomers.[12c] Oligomers have been prepared using reductive amination chemistry and using thiol−ene coupling, and both types of backbone lead to the formation of stable duplexes. Experiments on three isomeric reductive amination backbones indicate that an important requirement for duplex formation is a backbone with sufficient flexibility to adopt to a compatible conformation.[12b,g]

However, when mixed donor−acceptor sequences were prepared, it became clear that the conformational properties of the backbone also play a key role in discriminating between intramolecular H-bonding interactions that lead to folding and intermolecular interactions that lead to duplex formation. Highly flexible backbones lead to folding that precludes effective sequence-selective duplex formation in mixed sequence oligomers.[12e] These studies have identified two backbones that have promising conformational properties, and we recently showed that one of these systems supports sequence-selective duplex formation for mixed sequence 3-mers.[12f] Figure 1b shows the other backbone that is sufficiently rigid to preclude intramolecular folding in the mixed sequence AD 2-mer.[12e] This phenylacetylene backbone is particularly attractive for a number of reasons. Phenylacetylene oligomers have been extensively studied by Moore,[13] who developed solid phase synthesis methods that could be used to make oligomers with controlled length and sequence.[14] Oligomers longer than about eight repeat units fold into helical structures in acetonitrile solution, due to stacking interactions,[15] so the single-stranded forms may have interesting properties. Formation of phenylacetylene duplexes based on dynamic covalent base-pairs was investigated previously, but slow kinetics limited the fidelity.[16] Here we describe the synthesis of families of recognition-encoded oligomers based on the phenylacetylene backbone and show that this molecular

Received: July 30, 2018
Published: September 4, 2018

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

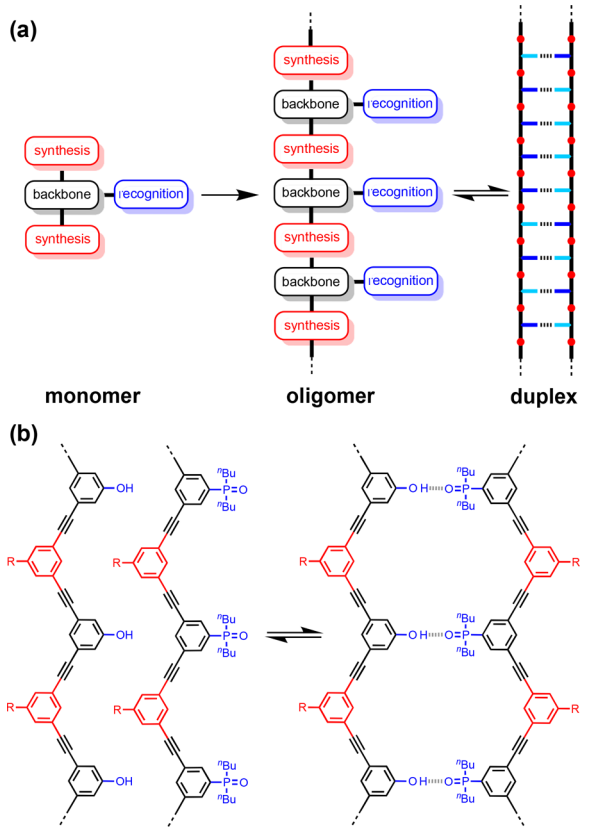


Figure 1. (a) A blueprint for duplex forming molecules. There are three key design elements: the coupling chemistry used for the synthesis of oligomers (red), the recognition module, which controls intermolecular binding (blue), and the backbone module, which links these components together (black). (b) Proposed duplex formed by phenylacetylene oligomers equipped with phenol and phosphine oxide recognition modules (R is a solubilizing group).

architecture leads to the formation of very stable H-bonded duplexes in toluene.

RESULTS AND DISCUSSION

Stepwise synthesis of oligomeric molecules requires multiple deprotection-coupling cycles and is labor intensive. Here we investigate a one-step process, statistically controlled oligomerization, to rapidly access families of oligomers of varying length (Figure 2). The average oligomer length can be controlled by adding monofunctional chain stoppers to the reaction mixture, and the resulting product mixture can be separated by preparative HPLC. Sonogashira coupling of bis-acetylene recognition modules with diiodo-linkers in the presence of the corresponding monoacetylene recognition modules was therefore investigated as a simple route to homo-oligomers.

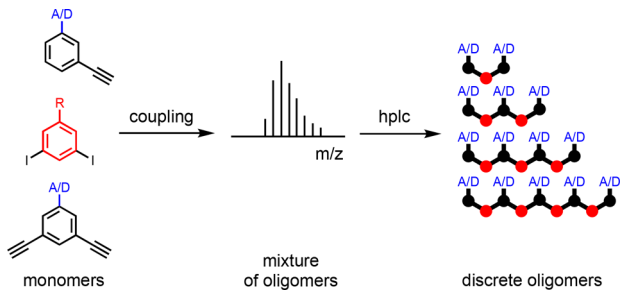
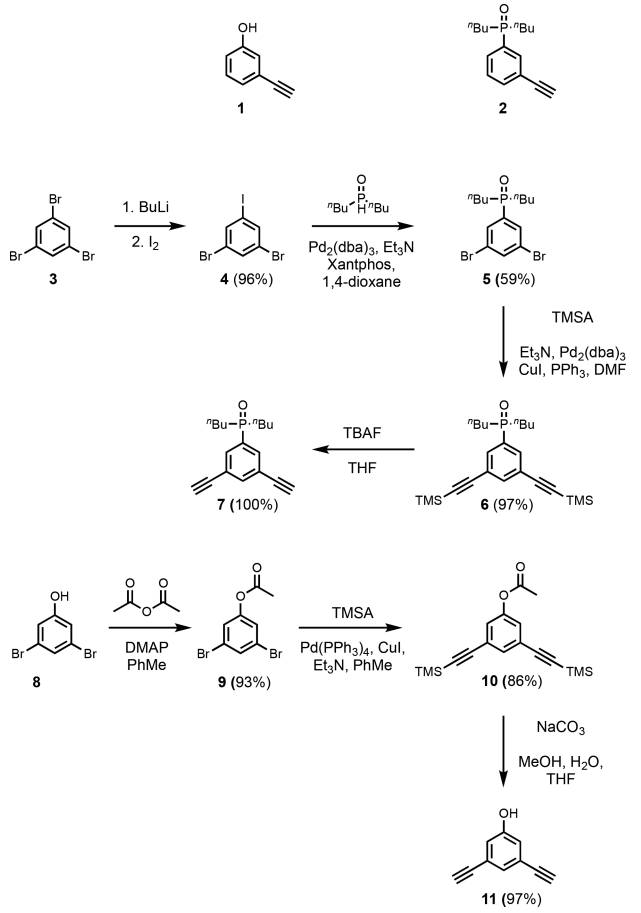


Figure 2. Oligomer synthesis using statistical coupling of mono- and bifunctional monomers to produce mixtures that can be separated by chromatography. H-bond donor (D) or acceptor (A) recognition groups can be used, and R is a solubilizing group.

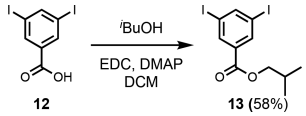
Synthesis of Building Blocks. The monoacetylene donor recognition module (**1**) was commercially available, and the monoacetylene acceptor recognition module (**2**) was synthesized as previously described (Scheme 1).[12e] Recognition modules bearing two acetylene moieties were synthesized using the route shown in Scheme 1. Treatment of tribromobenzene (**3**) with ^{n}BuLi followed by reaction with iodine gave dibromoiodobenzene (**4**). The difference in reactivity of halides allowed for selective P-arylation with di-*n*-butylphosphine oxide to give **5**. Sonogashira coupling with TMSA yielded **6**, and TBAF mediated deprotection yielded the bis-acetylene acceptor recognition module (**7**). Acetylation of 3,5-dibromophenol (**8**) with acetic anhydride yielded the acetate derivative (**9**). Acetylation was required to increase the reactivity in the subsequent Sonogashira coupling with TMSA to obtain **10**. Simultaneous base-mediated deprotection of the phenol and acetylenes yielded the bis-acetylene donor recognition module (**11**). The diiodo-linker (**13**) was synthesized by an EDC-mediated esterification of 3,5-diiodobenzoic acid (**12**) with ^{i}BuOH (Scheme 2).

Scheme 1. Mono-Acetylene Recognition Modules, and Synthesis of Bis-Acetylene Recognition Modules



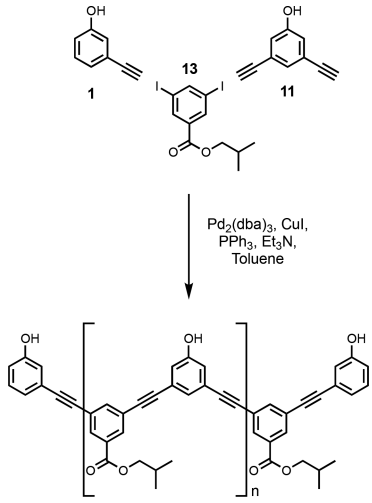
DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

Scheme 2. Synthesis of Diiodo-Linker



Oligomer Synthesis. Donor homo-oligomers were synthesized by oligomerization of the phenol recognition modules (**1** and **11**) with linker **13** under Sonogashira coupling conditions in toluene (Scheme 3).[12e,17] The ratio of starting materials was chosen to target the DDD 3-mer as the major product (**1**:**11**:**13** = 2:1:2). Although the crude reaction mixture contained a large amount of insoluble material, it could be dissolved in DMSO. LCMS analysis of the DMSO solution showed the presence of oligomers up to the 6-mer (Figure 3a). Preparative HPLC was used to separate the products (Figure 2b), and samples of donor homo-oligomers from the 3-mer to 6-mer were isolated by this method (Table 1).[18] The overall yield with respect to the diiodo-linker **13** was 15%, due to the low solubility of the oligomeric products. The most abundant oligomer isolated was the DDD 3-mer, as expected from the stoichiometry of the two recognition modules used in the reaction.

Scheme 3. Oligomerization of Phenol Recognition Modules to Yield Donor Homo-Oligomers (*n* = 1−4 Corresponding to the 3-mer to 6-mer Donor Homo-Oligomers)



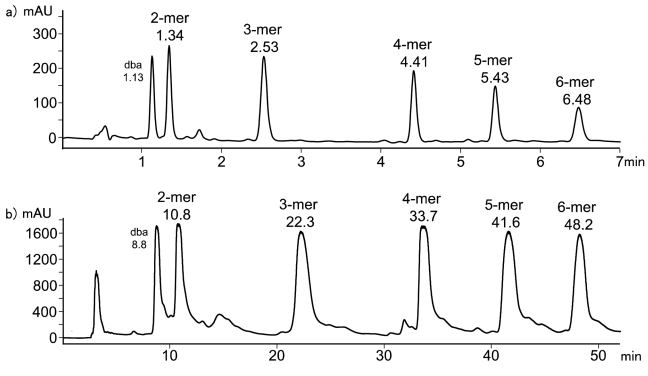


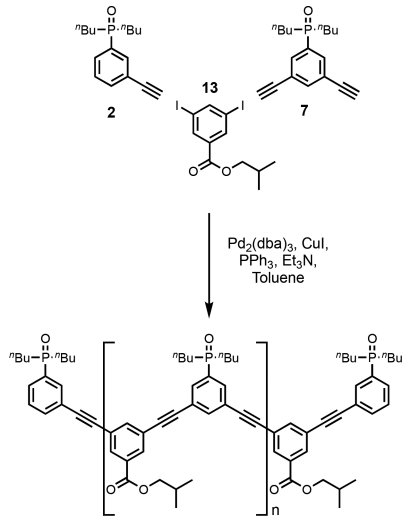
Figure 3. Product distribution from the donor oligomerization reaction shown in Scheme 3. (a) LCMS analysis of the crude reaction mixture using a Hichrom C_8C_{18} column with a water/THF (0.1% formic acid) solvent system. (b) Preparative HPLC separation of donor homo-oligomers using a HIRBP-6988 prep column with a water/THF solvent system. Samples were prepared in DMSO, UV/vis absorption was measured at 290 nm, and the peaks identified by MS are labeled with retention time in minutes. dba = dibenzylideneacetone.

Table 1. Donor homo-oligomers isolated from the oligomerization reaction in Scheme 3

product	sequence	% mol fraction	% yield
14 (3-mer)	DDD	40	6
15 (4-mer)	DDDD	27	4
16 (5-mer)	DDDDD	21	3
17 (6-mer)	DDDDDD	12	2

Acceptor homo-oligomers were synthesized by oligomerization of the phosphine oxide recognition modules (**2** and **7**) with linker **13** under Sonogashira conditions in toluene (Scheme 4). The ratio of starting materials was chosen to target the AAA 3-mer as the major product (**2**:**7**:**13** = 2:1:2). The reaction mixture was dissolved in ethanol and analyzed by LCMS (Figure 3a). Oligomers up to the 7-mer were observed. Preparative HPLC was used to separate the products (Figure 4b), and samples of the acceptor homo-oligomers from the 3-mer to 7-mer were isolated (Table 2). The overall yield with respect to the diiodo-linker **13** was 33%. The most abundant oligomer in the mixture was the AAA 3-mer, as expected from the stoichiometry of the two recognition modules used in the reaction.

Scheme 4. Oligomerization of Phosphine Oxide Recognition Modules to Yield Acceptor Homo-Oligomers (*n* = 1−5 Corresponding to the 3-mer to 7-mer Acceptor Homo-Oligomers)



Oligomer Characterization. Mass spectrometry and ^{1}H NMR spectroscopy were used to confirm the structures of the oligomers. The terminal recognition modules, the internal recognition modules, and the linkers all give rise to distinct signals in the ^{1}H NMR spectra, and the ratios of the integrals of these signals were used to confirm oligomer length. For the donor oligomers, the ratio of the integrals of the ^{1}H NMR signals due to the terminal linkers (blue and orange) compared with the signals due to the internal linkers (red and green)

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

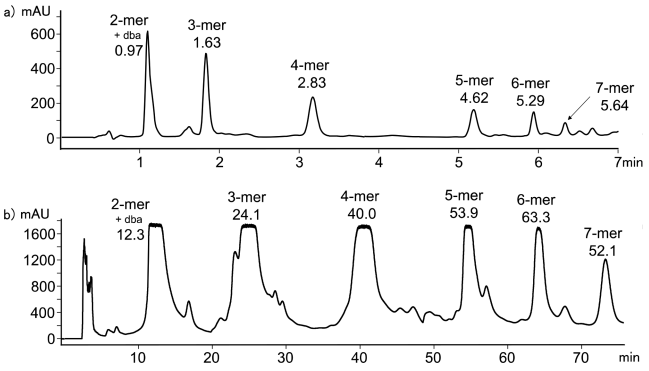


Figure 4. Product distribution from the acceptor oligomerization reaction shown in Scheme 4. (a) LCMS analysis of the crude reaction mixture using a Hichrom C_8C_{18} column with a water/THF (0.1% formic acid) solvent system. (b) Preparative HPLC separation of acceptor homo-oligomers using a HIRBP-6988 prep column with a water/THF solvent system. Samples were prepared in EtOH, UV/vis absorption was measured at 290 nm, and the peaks identified by MS are labeled with retention time in minutes. dba = dibenzylideneacetone.

Table 2. Acceptor Homo-Oligomers Isolated from the Oligomerization Reaction in Scheme 4

product	sequence	% mol fraction	% yield
18 (3-mer)	AAA	58	19
19 (4-mer)	AAAA	27	9
20 (5-mer)	AAAAA	10	3
21 (6-mer)	AAAAAA	4	1
22 (7-mer)	AAAAAAA	1	<1

allowed quantification of number of linkers and therefore oligomer length (Figure 5). Similarly, for the acceptor oligomers, the ratio of the integrals of the ^{1}H NMR signals due to the terminal linkers (blue) compared with the signals due to the internal linkers (red) allowed quantification of the number of linkers and therefore oligomer length. These results were confirmed using the ratio of integrals of the ^{31}P NMR signals due to the terminal recognition modules (blue) compared with the signals due to the internal recognition modules (red) (Figure 6).

For both the donor homo-oligomers in THF and the acceptor homo-oligomers in chloroform, the chemical shifts of the signals due to equivalent protons from equivalent building blocks do not change as the length of the oligomers increases. Moore and co-workers have studied the properties of a related family of phenylacetylene oligomers in acetonitrile, and they observed substantial upfield changes in chemical shift (0.5 ppm) for the signals due to the aromatic protons as the length of the oligomers increased.[14] These chemical shift changes are indicative of folding into a helical conformation with stacking of the aromatic rings. The ^{1}H NMR data in Figures 5 and 6 show that the oligomers described in this work do not fold in chloroform or THF, but the solubility is not sufficient to record NMR spectra in acetonitrile.

Binding Studies. Previous studies of duplex assembly used toluene as the solvent, because toluene is a very weak H-bond donor and acceptor, which maximizes the strength of the phenol-phosphine oxide interactions. The intrinsic strength of this base-pairing interaction is important for maximizing *K* EM, the parameter that governs chelate cooperativity in zipping up the duplex.[12,19] The poor solubility of the donor homo-oligomers in toluene described here made it impossible to measure association constants for duplex formation by titration experiments. However, in the presence of the complementary acceptor homo-oligomer, both DDD and DDDD are soluble at millimolar concentrations in toluene. The longer donor homo-oligomers do not dissolve in toluene even in the presence of the complementary acceptor homo-oligomer.

The ^{31}P NMR spectrum of a 1:1 mixture of AAA and DDD in toluene-d_8 shows one broad signal at 40−41 ppm. In

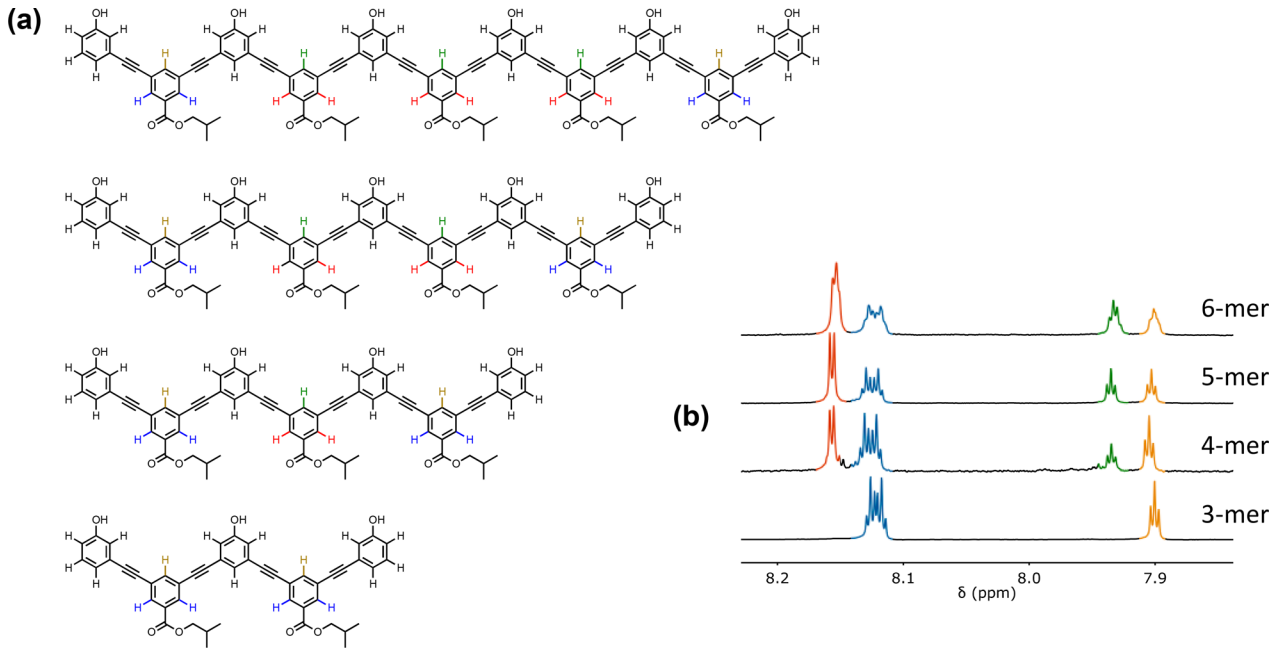


Figure 5. (a) Chemical structures of the donor homo-oligomers. (b) Partial 500 MHz ^{1}H NMR spectra in THF-d_6. The signals in the ^{1}H NMR spectra are assigned to the chemical structures using color coding. The 4-mer contains traces of an impurity that could not be removed by HPLC.

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

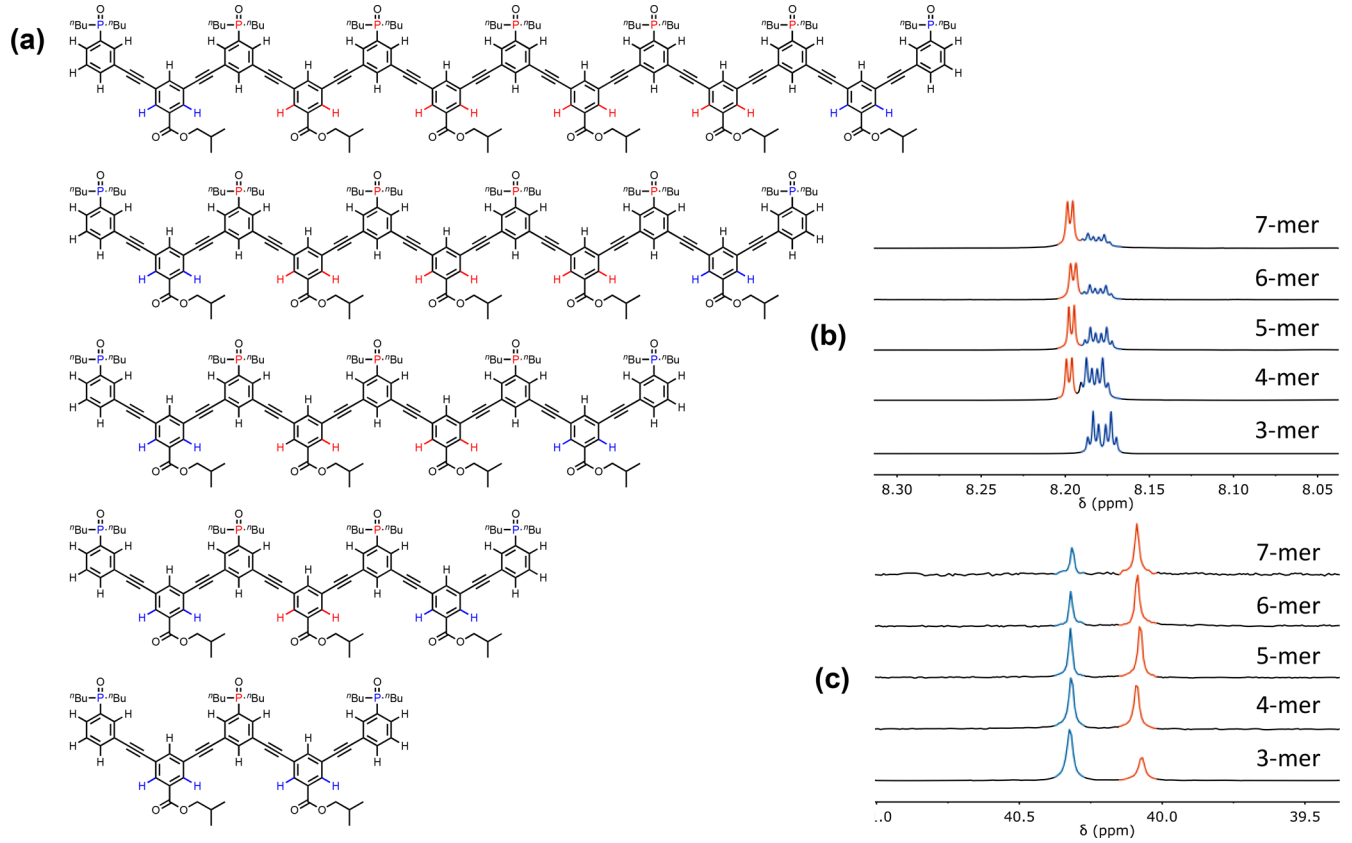


Figure 6. (a) Chemical structures of the acceptor homo-oligomers. (b) Partial 500 MHz ^{1}H NMR spectra in $CDCl_3$. (c) 162 MHz ^{31}P NMR spectra in $CDCl_3$. The signals in the ^{1}H and ^{31}P NMR spectra are assigned to the chemical structures using color coding.

contrast in the ^{31}P NMR spectrum of AAA in toluene-d_8, there are two signals at 35.9 ppm. These differences in ^{31}P chemical shift are characteristic of phosphine oxide groups that are fully H-bonded to phenol in the mixture of AAA and DDD. In the ^{1}H NMR spectrum of this mixture, the two signals due to the phenol OH groups appear at 11 ppm, which indicates that the three phenol groups of DDD are fully H-bonded to phosphine oxide groups on AAA. These data provide good evidence that at a concentration of 1 mM in toluene-d_8 the AAA·DDD duplex is fully assembled with three intermolecular phenol-phosphine oxide H-bonds. Similarly for a 1:1 mixture of AAAA and DDDD in toluene-d_8, the ^{31}P NMR spectrum has two signals at 40 ppm, and the ^{1}H NMR spectrum has two signals due to the phenol OH groups at 11 ppm. Thus, the AAAA·DDDD duplex is also fully assembled with four intermolecular phenol-phosphine oxide H-bonds at a concentration of 1 mM in toluene-d_8. In all cases, equilibration was rapid, and no slow exchange processes were observed.

The stabilities of the AAA·DDD and AAAA·DDDD duplexes were characterized by denaturation experiments with DMSO. DMSO can H-bond to the phenol groups on the donor homo-oligomers and thus dissociate the duplexes. If the association constant for the phenol-DMSO interaction is known (K_d), it is possible to use the denaturation isotherm to deduce the association constant for duplex formation between two complementary N-mers (K_N). To test the validity of this approach, the 1-mers and 2-mers, which were reported previously,[12e] were used to directly compare the results from a titration and a denaturation experiment (Figure 7).

Association constants for formation of the A·D and AA·DD complexes in toluene-d_8 were determined using ^{31}P NMR

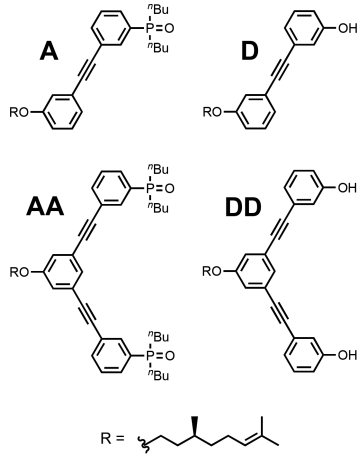


Figure 7. 1-mers and 2-mers used to validate the DMSO denaturation experiment.

titration experiments. The data fit well to 1:1 binding isotherms (see the Supporting Information, SI), and the results are summarized in Table 3. The large downfield changes in ^{31}P NMR chemical shift ($^{+}$6−7 ppm) are indicative of H-bond formation, and the bound chemical shifts are very similar to those observed for the AAA·DDD and AAAA·DDDD duplexes described above. The association constant for the AA·DD complex (K_2 = 44 000 M^{-1}) is 2 orders of magnitude higher than the association constant for the A·D complex (K_1 = 760 M^{-1}) indicating that the duplex is fully assembled with two cooperative intermolecular phenol-phosphine oxide H-bonds.

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

Table 3. Association Constants for Duplex Formation (K_N), Effective Molarities (EM), and ^{31}P NMR Chemical Shifts (δ) Measured by NMR Titration and DMSO Denaturation Experiments in Toluene-d_8 at 298 K[a]

method	complex	K_N (M^{-1})	EM (mM)	δ_f (ppm)	δ_b (ppm)
titration	A·D	760		35.3	42.1
titration	AA·DD	44 000	38	35.3	41.3
denaturation	AA·DD	83 000	72	34.8	42.1
denaturation	AAA·DDD	2 300 000	52	34.8	40.5
denaturation	AAAA·DDDD	130 000 000	59	35.3	40.1

[a]Errors are estimated at ±30% based on two repeats of the experiment.

The value of K_d was measured by ^{1}H NMR titration of DMSO into the D 1-mer in toluene (K_d = 150 M^{-1}, see SI). For denaturation of the AA·DD duplex, an equimolar 1 mM solution of AA and DD was prepared in toluene-d_8. At this concentration, AA is 90% bound to DD, and the ^{31}P signal was accordingly observed close to the bound chemical shift at 41 ppm. Addition of DMSO-d_6 lead to a decrease in the chemical shift of the ^{31}P NMR signal to 36 ppm, due to disruption of the H-bonding interactions (Figure 8b). However, the final chemical shift is 1 ppm higher than the free chemical shift from the titration experiment (35 ppm), and Figure 8b shows that the ^{31}P chemical shift starts to increase at very high concentrations of DMSO at the end of the denaturation experiment. This observation is presumably due to a change in the nature of the solvent for DMSO concentrations of the order 1 M. DMSO-d_6 was therefore titrated into a 1 mM sample of the A 1-mer in toluene-d_8 where no H-bonding interactions are present (Figure 8a). An increase of about 1 ppm was observed in the ^{31}P chemical shift at very high DMSO concentrations, confirming that the unusual shape of the AA·DD denaturation data is due to nonspecific effects of the change in solvent. The ^{31}P chemical shift data for the A 1-mer were therefore used to correct the ^{31}P chemical shifts from the AA·DD denaturation experiment for the effects of DMSO on the chemical shift of the unbound phosphine oxide signal. The proportion of duplex present at any point in the denaturation experiment is therefore given by eq 1.

$$\%\text{duplex} = (\delta_{obs} - \delta_A - \delta_f)/(\delta_b - \delta_f) \quad (1)$$

where δ_{obs} is the observed chemical shift for the 1:1 mixture at a specific concentration of denaturant, δ_A is the observed chemical shift of the A 1-mer at the same concentration of denaturant, δ_f is the free chemical shift of the acceptor homo-oligomer in pure toluene-d_8, and δ_b is the bound chemical shift of the duplex.

The data from the AA·DD denaturation experiment did not fit to a simple two-state, all-or-nothing denaturation isotherm (see SI), which suggests that partially denatured species must also be considered. Figure 9 shows all possible complexes present in the AA·DD denaturation experiment. Both the fully bound duplex and unbound species are present at the start of the experiment. Binding of one molecule of DMSO to the duplex leads to the partially denatured AA·DD·DMSO complex. It is possible to estimate the equilibrium constant for the formation of this termolecular complex as the product of K_d (the D·DMSO association constant) and K_1 (the A·D association constant), multiplied by a statistical factor of 4. The unbound DD 2-mer can bind one or two molecules of DMSO, and the equilibrium constants for formation of the DD·DMSO and DD·$(DMSO)_2$ complexes can be estimated using the D·DMSO association constant as $2K_d$ and K_d^2 respectively.

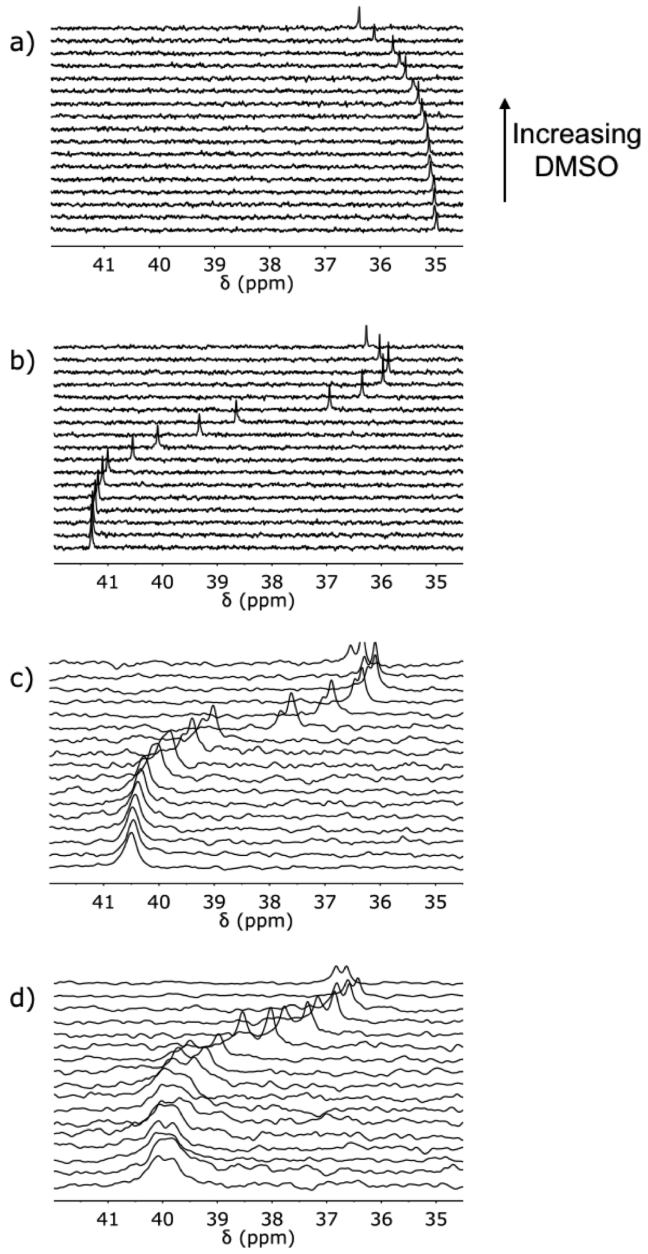


Figure 8. 202 MHz ^{31}P NMR spectra for titration of DMSO-d_6 into a) A (1 mM), and DMSO-d_6 denaturation of equimolar 1 mM solutions of b) DD and AA, c) DDD and AAA, and d) DDDD and AAAA in toluene-d_8 at 298 K.

The data for denaturation of the AA·DD duplex fit well to a denaturation isotherm that allows for all of these species, and this allowed determination of K_2 as the only unknown equilibrium constant. The results are shown in Table 3. The

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

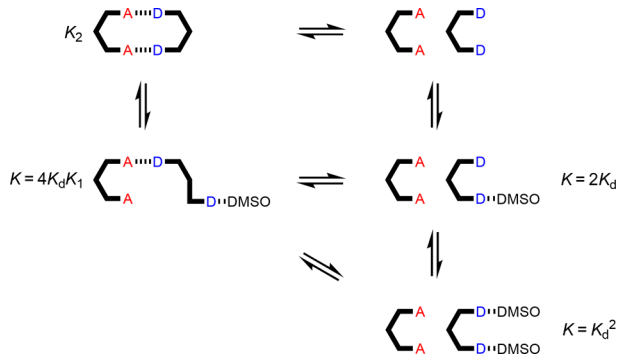


Figure 9. Equilibria involved in denaturation of AA·DD. K_d is the D·DMSO association constant, and K_N is the association constant for duplex formation between homo-oligomers of length N in toluene-d_8.

value of K_2 determined in the denaturation experiment is comparable to the value determined in the corresponding titration experiment in toluene, which indicates that the denaturation methodology provides a robust method for determining duplex stability.

The ^{31}P NMR DMSO denaturation experiments for the AAA·DDD and AAAA·DDDD duplexes are shown in Figure 8c and d, respectively, and the results are similar to those obtained for AA·DD. Fitting the denaturation data for the longer oligomers requires consideration of a larger number of partially denatured species (Figures 10 and 11). The equilibrium

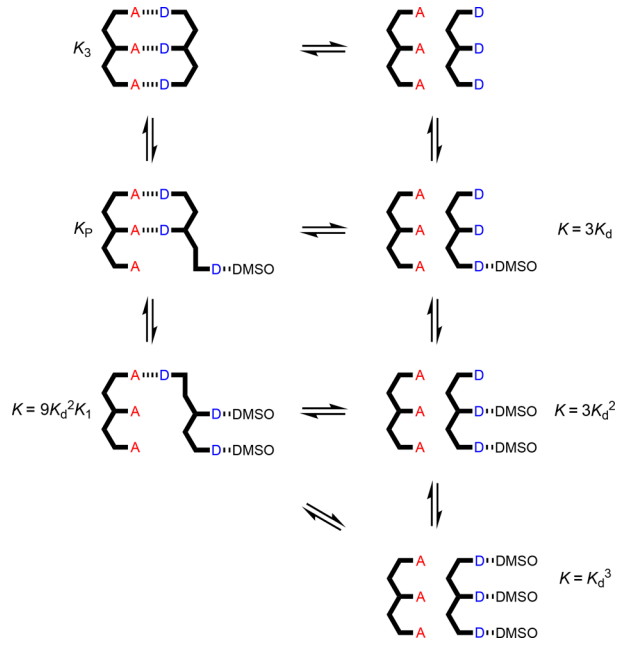


Figure 10. Equilibria involved in denaturation of AAA·DDD. K_d is the D·DMSO association constant, and K_N is the association constant for duplex formation between homo-oligomers of length N in toluene-d_8. For some of the complexes, isomeric arrangements are possible (not shown).

constants for formation of most of these species can also be estimated from the values of K_d and K_1, but some partially denatured duplexes have no direct analogues. For example, a number of different structures are possible for the AAA·DDD·DMSO complex, and we do not know whether complexes with DMSO bound to one of the terminal recognition units have

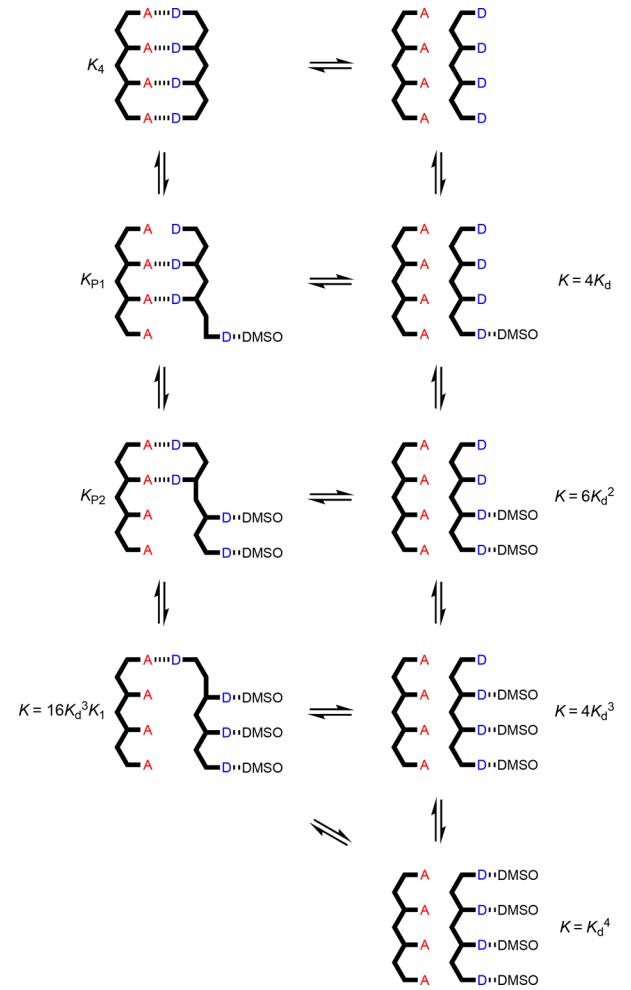


Figure 11. Equilibria involved in denaturation of AAAA·DDDD. K_d is the D·DMSO association constant, and K_N is the association constant for duplex formation between homo-oligomers of length N in toluene-d_8. For some of the complexes, isomeric arrangements are possible (not shown).

similar stability to the complex with DMSO bound to the central recognition unit. By analogy with the AA·DD·DMSO complex, one might assume that the association constant for AAA·DDD·DMSO complex, K_P, will be the product of K_d and K_2, multiplied by a statistical factor, but to avoid any bias in the fitting, K_P was optimized as a variable along with K_3 in analysis of the AAA·DDD denaturation data. The results of fitting the proportion of duplex present determined using eq 1 to the denaturation isotherm are shown in Figure 12, and the resulting association constant for duplex formation is reported in Table 3. The optimized value of K_P from the fitting is equal to 3.5 K_d K_2. If the duplex preferentially denatures from the ends, so that the central H-bond is intact in the partially denatured AAA·DDD·DMSO complex, the statistical factor would be 4, which is close to the optimized value of 3.5. The AAA·DDD duplex is more than an order of magnitude more stable than the AA·DD duplex, which confirms that the fully assembled triply H-bonded duplex is formed in toluene.

For denaturation of the AAAA·DDDD duplex, there are two partially denatured complexes with unknown stability, AAAA·DDDD·DMSO and AAAA·DDDD·$(DMSO)_2$ (Figure 11). Therefore, fitting of the denaturation data for this system involved optimization of three equilibrium constants, K_4, K_{P1},

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

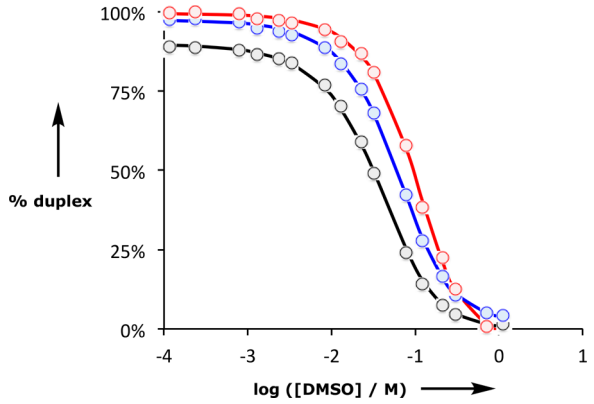


Figure 12. Duplex denaturation data plotted as a function of DMSO-d_6 concentration in toluene-d_8 at 298 K for AA·DD (black), AAA·DDD (blue), and AAAA·DDDD (red). The dots represent the experimental values obtained using eq 1, and the lines are the calculated denaturation isotherms.

and K_{P2}. The results are shown in Figure 12 and Table 3. The AAAA·DDDD is more than an order of magnitude more stable than the AAA·DDD duplex, which confirms that the fully assembled duplex with all four H-bonds is formed in toluene. The value of K_{P1} is equal to 7.5 K_d K_3. If the duplex preferentially denatures from the ends, so that the central H-bonds are both intact in the partially denatured AAAA·DDDD·DMSO complex, then the statistical factor would be 4, which is smaller than the optimized value of 7.5. This result suggests that there are some additional partially denatured states where the central H-bonds are broken. The value of K_{P2} is equal to 8.0 K_d^2 K_2. If the duplex preferentially denatures from the ends, then the statistical factor would be 9, which is close to the optimized value.

Figure 13 shows the speciation of different complexes calculated from the denaturation isotherms. In the denaturation of all three duplexes, the first step is breaking of one duplex H-bonds to form a phenol-DMSO H-bond, and this complex is the major partially denatured state in all cases (blue lines in Figure 13a–c). For the AA·DD duplex, this partially denatured complex is not very stable, so the DD·DMSO complex is also significantly populated (pink line in Figure 13a). At high concentrations of DMSO, the fully denatured state dominates in all cases (red lines in Figure 13a–c).

Denaturation of the duplexes can also be monitored by ^{1}H NMR spectroscopy (Figure 14). The signals due to the aromatic protons of the oligomers move by 0.1–0.6 ppm on denaturation. For the AA·DD duplex, the signal due to the phenol OH proton starts at 6.2 ppm, suggesting that the duplex is not fully bound, and moves to 10 ppm when bound to DMSO (Figure 14a). For both the AAA·DDD and AAAA·DDDD duplexes, the signals due to the phenol OH protons start at about 11 ppm, suggesting that the duplexes are fully bound, and move to about 10 ppm when bound to DMSO (Figure 14b and c). As observed for the ^{31}P NMR spectra, there are differences at very high concentrations of DMSO-d_6 due to the change in the nature of the solvent.

The association constants for duplex formation, the average effective molarities (EM) for intramolecular H-bonding interactions in the duplexes, and the free and bound chemical shifts are shown in Table 3. The consistency of the values suggests that the results obtained from the denaturation experiments are accurate. The association constants increase

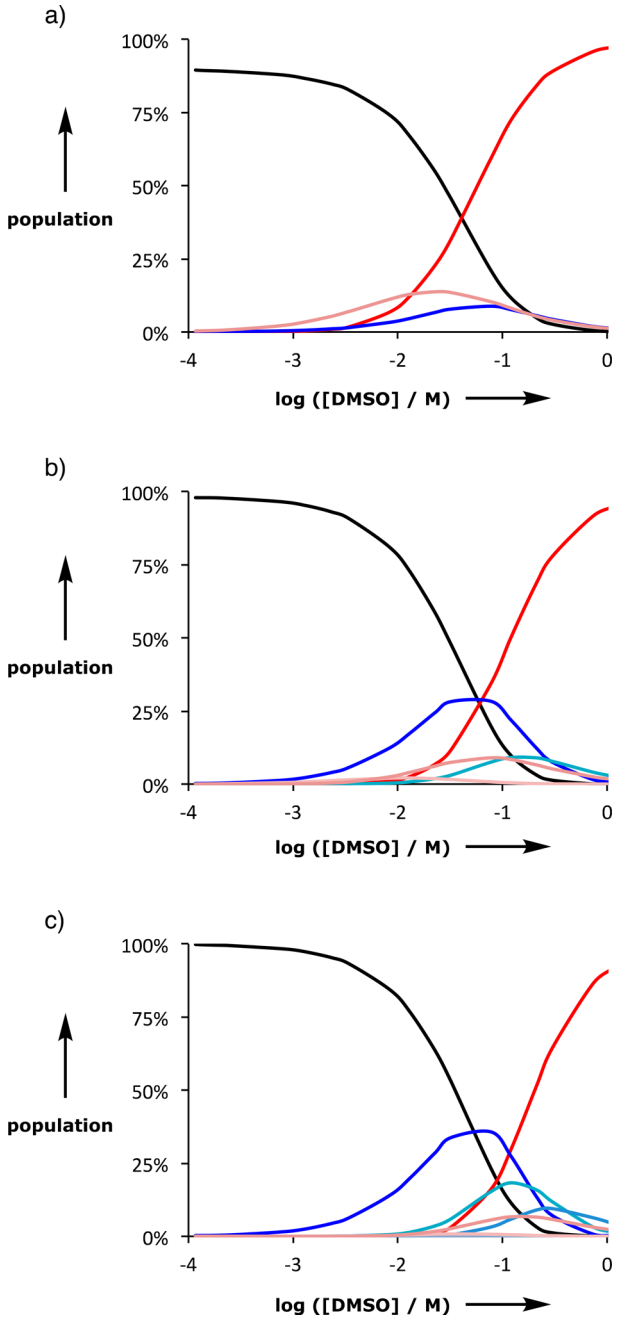


Figure 13. Speciation from DMSO denaturation experiments on length complementary homo-oligomer duplexes. a) The 2-mer duplex: AA·DD (black), DD·DMSO (pink), AA·DD·DMSO (blue), and DD·$(DMSO)_2$ (red). b) The 3-mer duplex: AAA·DDD (black), AAA·DDD·DMSO (blue), DDD·$(DMSO)_2$ (pink), AAA·DDD·$(DMSO)_2$ (green), and DDD·$(DMSO)_3$ (red). c) The 4-mer duplex: AAAA·DDDD (black), AAAA·DDDD·DMSO (blue), AAAA·DDDD·$(DMSO)_2$ (cyan), DDDD·$(DMSO)_3$ (pink), and DDDD·$(DMSO)_4$ (red). Free acceptor homo-oligomer populations are not shown, and all other possible complexes are not populated to a significant extent.

with each additional H-bond, and the value of EM does not vary significantly between duplexes.

Figure 15 shows the value of log K_N plotted as a function of the number of recognition modules in the oligomers (N). There is a uniform increase of 1.7 log units for every additional H-bond in the duplex. The results presented here can be compared with previous experiments on different backbone

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526–11536

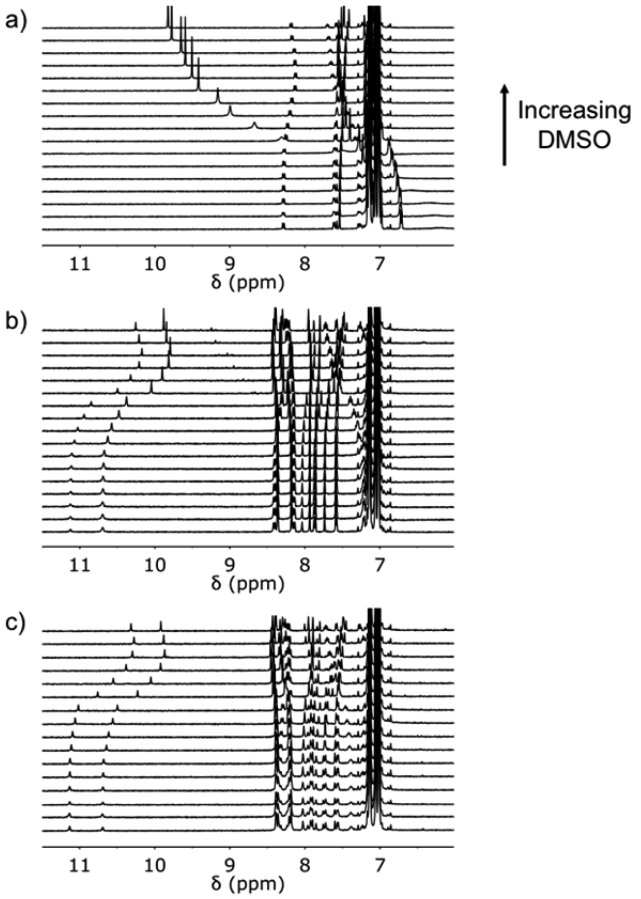


Figure 14. 500 MHz ^{1}H NMR spectra for DMSO-d_6 denaturation of equimolar 1 mM solutions of a) A, b) DD·AA, c) DDD·AAA, and d) DDDD·AAAA in toluene-d_8 at 298 K.

stuctures equipped with the same phenol-phosphine oxide recognition modules. The data in Figure 15 show that the phenylacetylene backbone (red) gives the most stable duplexes, presumably because these oligomers are more rigid than the aniline (green) and thioether (blue) backbones. Although the increased preorganization and geometric complementarity of the phenylacetylene backbone leads to higher values of EM for formation of the intramolecular H-bonds that zip up the duplex, the increase is actually quite small compared with the more flexible backbones (50 mM compared with 20 mM and 30 mM for the green and blue backbones, respectively, in Figure 15). However, small differences in H-bond strength and EM are cumulative along an oligomer, so that the phenylacetylene 4-mer duplex is 2 orders of magnitude more stable than the corresponding duplex with the aniline backbone.

CONCLUSIONS

Oligomerization of bifunctional building blocks in the presence of monofunctional chain terminators provides a straightforward approach to access families of oligomeric molecules with minimal synthetic effort. Here we describe the use of this strategy to prepare phenylacetylene oligomers equipped with phenol and phosphine oxide recognition sites. Pure samples of homo-oligomers up to seven recognition units in length were obtained by preparative hplc of the crude reaction mixtures. The phenol oligomers form stable duplexes with length complementary phosphine oxide oligomers via intermolecular

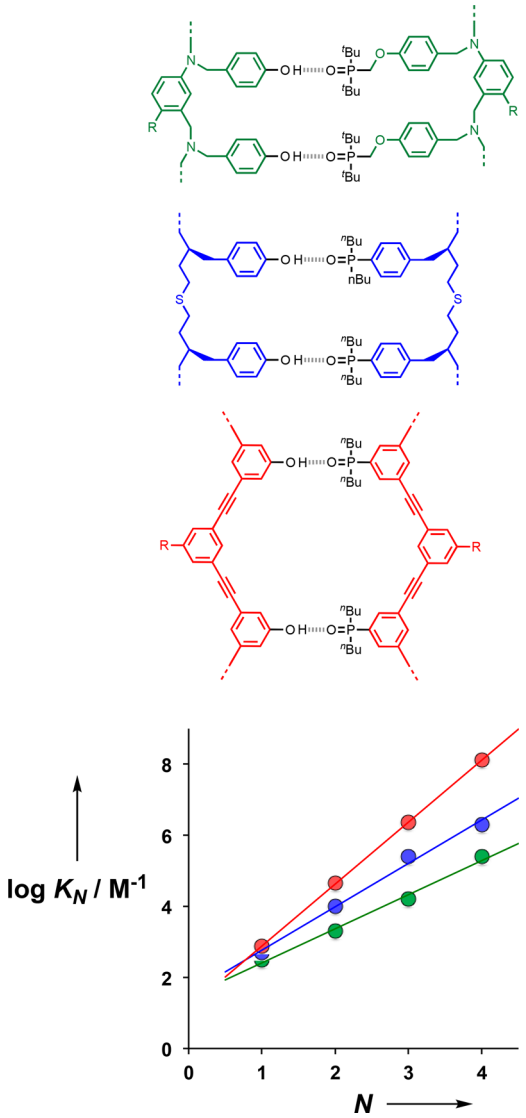


Figure 15. Relationship between the association constant for duplex formation between length-complementary oligomers in toluene-d_8 at 298 K (K_N) and the number of intermolecular H-bonds formed (N). Data are shown for three different backbones (color coded), each equipped with phenol-phosphine oxide recognition modules. The lines of best fit are shown: red log K_N = 1.7N + 1.3; blue log K_N = 1.2N + 1.6; green log K_N = 1.0N + 1.5. R represents a solubilizing group.

phenol-phosphine oxide H-bonding interactions in toluene solution. NMR spectroscopy indicates that all possible base-pairing interactions between recognition sites are fully bound in the duplexes.

Addition of a competing H-bond acceptor, DMSO, dissociated the duplexes, and DMSO denaturation experiments were used to measure the association constants for duplex formation. The association constants increase by almost 2 orders of magnitude for every additional H-bonded base-pair added to the duplex, reaching 10^8 M^{-1} for AAAA·DDDD. The average effective molarity for the intramolecular H-bonding interactions responsible for zipping up the duplex is EM = 50 mM, and the corresponding value of the intramolecular equilibrium constant that quantifies chelate cooperativity in these systems is K EM = 40 (the product of the association constant for formation of one intermolecular phenol-

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

phosphine oxide H-bond and the effective molarity). A value of 40 for *K* EM means that in the duplex the recognition sites are 98% bound, which is consistent with the NMR data. The uniform increase in duplex stability with oligomer length suggests that this supramolecular architecture is likely to be compatible with the formation extended H-bonded duplexes in longer oligomers. Given that the mixed sequence AD 2-mer does not exhibit any intramolecular folding, the recognition-encoded phenylacetylene oligomer system described here appears to be a promising candidate for the development of synthetic information molecules with sequence-selective duplex-forming properties that resemble nucleic acids.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b08087.

Detailed experimental procedures and ^{1}H and ^{13}C of all compounds, NMR titration and dilution spectra, details of fitting the binding isotherm, and detailed HPLC methods used for separation of oligomers (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*herchelsmith.orgchem@ch.cam.ac.uk

ORCID

Christopher A. Hunter: 0000-0002-5182-1859

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Engineering and Physical Sciences Research Council (EP/J008044/2) and European Research Council (ERC-2012-AdG 320539-duplex) for funding.

■ REFERENCES

(1) (a) Crick, F. H. C.; Watson, J. D. *Nature* **1953**, *171*, 964. (b) Kyogoku, Y.; Lord, R. C.; Rich, A. *Science* **1966**, *154*, 518. (c) Kyogoku, Y.; Lord, R. C.; Rich, A. *J. Am. Chem. Soc.* **1967**, *89*, 496. (d) Newmark, R. A.; Cantor, C. R. *J. Am. Chem. Soc.* **1968**, *90*, 5010. (e) Porschke, D. *Biopolymers* **1971**, *10*, 1989. (f) Crick, F. *Nature* **1970**, *227*, 561.

(2) (a) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818. (b) Rothemund, P. W. K. *Nature* **2006**, *440*, 297. (c) Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F. *Science* **2008**, *321*, 1795. (d) Seeman, N. C. *Nano Lett.* **2010**, *10*, 1971. (e) Ke, Y.; Ong, L. L.; Shih, W. M.; Yin, P. *Science* **2012**, *338*, 1177−1183. (f) Pinheiro, A. V.; Han, D.; Shih, W. M.; Yan, H. *Nat. Nanotechnol.* **2011**, *6*, 763. (g) Zhang, F.; Nangreave, J.; Liu, Y.; Yan, H. *J. Am. Chem. Soc.* **2014**, *136*, 11198.

(3) (a) Eschenmoser, A. *Science* **1999**, *284*, 2118. (b) Benner, S. A. *Acc. Chem. Res.* **2004**, *37*, 784. (c) Benner, S. A.; Chen, F.; Yang, Z. *Chemical Synthetic Biology*; Luisi, P. L., Chiarabelli, C., Eds.; John Wiley & Sons, Ltd: Chichester, UK, 2011; pp 69−106. (d) Benner, S. A. *Biological Theory* **2013**, *8*, 357. (e) Wilson, C.; Keefe, A. D. *Curr. Opin. Chem. Biol.* **2006**, *10*, 607. (f) Appella, D. H. *Curr. Opin. Chem. Biol.* **2009**, *13*, 687. (g) Kool, E. T. *Curr. Opin. Chem. Biol.* **2000**, *4*, 602.

(4) Sugar modifications of nucleic acids (a) Schöning, K.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347. (b) van Aerschot, A.; Verheggen, I.; Hendrix, C; Herdewijn, P. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1338. (c) Renneberg, D.; Leumann, C. J. *J. Am. Chem. Soc.* **2002**, *124*, 5993. (d) Braasch, D. A.; Corey, D. R. *Chem. Biol.* **2001**, *8*, 1. (e) Singh, S. K.; Koshkin, A. A.; Wengel, J.; Nielsen, P. *Chem. Commun.* **1998**, 455. (f) Nguyen, H. V.; Zhao, Z. Y.; Sallustrau, A.; Horswell, S. L.; Male, L.; Mulas, A.; Tucker, J. H. R. *Chem. Commun.* **2012**, *48*, 12165. (g) Zhang, L.; Peritz, A.; Meggers, E. *J. Am. Chem. Soc.* **2005**, *127*, 4174. (h) Schlegel, M. K.; Peritz, A. E.; Kittigowittana, K.; Zhang, L.; Meggers, E. *ChemBioChem* **2007**, *8*, 927. (i) Karri, P.; Punna, V.; Kim, K.; Krishnamurthy, R. *Angew. Chem., Int. Ed.* **2013**, *52*, 5840.

(5) Phosphate modifications of nucleic acids: (a) Benner, S. A.; Hutter, D. *Bioorg. Chem.* **2002**, *30*, 62. (b) Huang, Z.; Schneider, K. C.; Benner, S. A. *J. Org. Chem.* **1991**, *56*, 3869. (c) Huang, Z.; Benner, S. A. *J. Org. Chem.* **2002**, *67*, 3996. (d) Richert, C.; Roughton, A. L.; Benner, S. A. *J. Am. Chem. Soc.* **1996**, *118*, 4518. (e) Li, P.; Sergueeva, Z. A.; Dobrikov, M.; Shaw, B. R. *Chem. Rev.* **2007**, *107*, 4746. (f) Isobe, H.; Fujino, T.; Yamazaki, N.; Guillot-Nieckowski, M.; Nakamura, E. *Org. Lett.* **2008**, *10*, 3729. (g) Eriksson, M.; Nielsen, P. E. Q. *Q. Rev. Biophys.* **1996**, *29*, 369. (h) Nielsen, P. E. *Chem. Biodiversity* **2010**, *7*, 786. (i) Nielsen, P. E.; Egholm, M. *Curr. Issues Mol. Biol.* **1999**, *1*, 89. (j) Nielsen, P. E.; Haaima, G. *Chem. Soc. Rev.* **1997**, *26*, 73. (k) Ura, Y.; Beierle, J. M.; Leman, L. J.; Orgel, L. E.; Ghadiri, M. R. *Science* **2009**, *325*, 73.

(6) Bases modification of nucleic acids: (a) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33. (b) Yang, Z.; Hutter, D.; Sheng, P.; Sismour, A. M.; Benner, S. A. *Nucleic Acids Res.* **2006**, *34*, 6095. (c) Wojciechowski, F.; Leumann, C. J. *Chem. Soc. Rev.* **2011**, *40*, 5669. (d) Benner, S. A. *Curr. Opin. Chem. Biol.* **2012**, *16*, 581. (e) Liu, H.; Gao, J.; Lynch, S. R.; Saito, Y. D.; Maynard, L.; Kool, E. T. *Science* **2003**, *302*, 868. (f) Kool, E. T. *Acc. Chem. Res.* **2002**, *35*, 936. (g) Kool, E. T.; Lu, H.; Kim, S. J.; Tan, S.; Wilson, J. N.; Gao, J.; Liu, H. *Nucleic Acids Symp. Ser.* **2006**, *50*, 15. (h) Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006**, *4*, 4265.

(7) (a) Kramer, R.; Lehn, J.-M.; Marquis-Rigault, A. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5394. (b) Marquis, A.; Smith, V.; Harrowfield, J.; Lehn, J.-M.; Herschbach, H.; Sanvito, R.; Leize-Wagner, E.; van Dorsselaer, A. *Chem. - Eur. J.* **2006**, *12*, 5632.

(8) (a) Gong, B.; Yan, Y.; Zeng, H.; Skrzypczak-Jankunn, E.; Kim, Y. W.; Zhu, J.; Ickes, H. *J. Am. Chem. Soc.* **1999**, *121*, 5607. (b) Gong, B. *Synlett* **2001**, *2001*, 0582. (c) Zeng, H.; Miller, R. S.; Flowers, R. A.; Gong, B. *J. Am. Chem. Soc.* **2000**, *122*, 2635. (d) Zeng, H.; Ickes, H.; Flowers, R. A.; Gong, B. *J. Org. Chem.* **2001**, *66*, 3574. (e) Gong, B. *Polym. Int.* **2007**, *56*, 436. (f) Gong, B. *Acc. Chem. Res.* **2012**, *45*, 2077. (g) Yang, Y.; Yang, Z. Y.; Yi, Y. P.; Xiang, J. F.; Chen, C. F.; Wan, L. J.; Shuai, Z. G. *J. Org. Chem.* **2007**, *72*, 4936.

(9) (a) Tanaka, Y.; Katagiri, H.; Furusho, Y.; Yashima, E. *Angew. Chem.* **2005**, *117*, 3935. (b) Ito, H.; Furusho, Y.; Hasegawa, T.; Yashima, E. *J. Am. Chem. Soc.* **2008**, *130*, 14008. (c) Yamada, H.; Furusho, Y.; Ito, H.; Yashima, E. *Chem. Commun.* **2010**, *46*, 3487. (d) Yashima, E.; Ousaka, N.; Taura, D.; Shimomura, K.; Ikai, T.; Maeda, K. *Chem. Rev.* **2016**, *116*, 13752.

(10) (a) Anderson, H. L. *Inorg. Chem.* **1994**, *33*, 972. (b) Taylor, P. N.; Anderson, H. L. *J. Am. Chem. Soc.* **1999**, *121*, 11538. (c) Berl, V.; Huc, I.; Khoury, R. G.; Krische, M. J.; Lehn, J.-M. *Nature* **2000**, *407*, 720. (d) Berl, V.; Huc, I.; Khoury, R. G.; Lehn, J.-M. *Chem. - Eur. J.* **2001**, *7*, 2810. (e) Sánchez-Quesada, J.; Seel, C.; Prados, P.; de Mendoza, J.; Dalcol, I.; Giralt, E. *J. Am. Chem. Soc.* **1996**, *118*, 277. (f) Bisson, A. P.; Carver, F. J.; Eggleston, D. S.; Haltiwanger, R. C.; Hunter, C. A.; Livingstone, D. L.; McCabe, J. F.; Rotger, C.; Rowan, A. E. *J. Am. Chem. Soc.* **2000**, *122*, 8856. (g) Bisson, A. P.; Hunter, C. A. *Chem. Commun.* **1996**, 1723. (h) Yang, Y.; Yang, Z.-Y.; Yi, Y.-P.; Xiang, J.-F.; Chen, C.-F.; Wan, L.-J.; Shuai, Z.-G. *J. Org. Chem.* **2007**, *72*, 4936. (i) Chu, W.-J.; Yang, Y.; Chen, C.-F. *Org. Lett.* **2010**, *12*, 3156. (j) Chu, W.-J.; Chen, J.; Chen, C.-F.; Yang, Y.; Shuai, Z. *J. Org. Chem.* **2012**, *77*, 7815. (k) Archer, E. A.; Krische, M. J. *J. Am. Chem. Soc.* **2002**, *124*, 5074. (l) Gong, H.; Krische, M. J. *J. Am. Chem. Soc.* **2005**, *127*, 1719.

(11) (a) Dervan, P. B.; Burli, R. W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688. (b) Renneberg, D.; Dervan, P. B. J. *J. Am. Chem. Soc.* **2003**, *125*, 5707. (c) Doss, R. M.; Marques, M. A.; Foister, S.; Chenoweth, D. M.; Dervan, P. B. *J. Am. Chem. Soc.* **2006**, *128*, 9074. (d) Meier, J. L.;

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536

Yu, A. S.; Korf, I.; Segal, D. J.; Dervan, P. B. *J. Am. Chem. Soc.* **2012**, *134*, 17814.
(12) (a) Stross, A. E.; Iadevaia, G.; Hunter, C. A. *Chem. Sci.* **2016**, *7*, 94. (b) Iadevaia, G.; Stross, A. E.; Neumann, A.; Hunter, C. A. *Chem. Sci.* **2016**, *7*, 1760. (c) Stross, A. E.; Iadevaia, G.; Hunter, C. A. *Chem. Sci.* **2016**, *7*, 5686. (d) Núñez-Villanueva, D.; Hunter, C. A. *Chem. Sci.* **2017**, *8*, 206. (e) Núñez-Villanueva, D.; Iadevaia, G.; Stross, A. E.; Jinks, M. A.; Swain, J. A.; Hunter, C. A. *J. Am. Chem. Soc.* **2017**, *139*, 6654. (f) Stross, A. E.; Iadevaia, G.; Núñez-Villanueva, D.; Hunter, C. A. *J. Am. Chem. Soc.* **2017**, *139*, 12655. (g) Iadevaia, G.; Núñez-Villanueva, D.; Stross, A. E.; Hunter, C. A. *Org. Biomol. Chem.* **2018**, *16*, 4183.
(13) (a) Zhang, J.; Moore, J. S.; Xu, Z.; Aguirre, R. A. *J. Am. Chem. Soc.* **1992**, *114*, 2273−2274. (b) Wu, Z.; Lee, S.; Moore, J. S. *J. Am. Chem. Soc.* **1992**, *114*, 8730−8732. (c) Xu, Z.; Kahr, M.; Walker, K. L.; Wilkins, C. L.; Moore, J. S. *J. Am. Chem. Soc.* **1994**, *116*, 4537−4550. (d) Lahiri, S.; Thompson, J. L.; Moore, J. S. *J. Am. Chem. Soc.* **2000**, *122*, 11315−11319. (e) Prince, R. B.; Brunsveld, L.; Meijer, E. W.; Moore, J. S. *Angew. Chem., Int. Ed.* **2000**, *39*, 228−230. (f) Prince, R. B.; Moore, J. S.; Brunsveld, L.; Meijer, E. W. *Chem. - Eur. J.* **2001**, *7*, 4150−4154. (g) Tanatani, A.; Mio, M. J.; Moore, J. S. *J. Am. Chem. Soc.* **2001**, *123*, 1792−1793. (h) Brunsveld, L.; Meijer, E. W.; Prince, R. B.; Moore, J. S. *J. Am. Chem. Soc.* **2001**, *123*, 7978−7984. (i) Cary, J. M.; Moore, J. S. *Org. Lett.* **2002**, *4*, 4663−4666. (j) Zhao, D.; Moore, J. S. *Chem. Commun.* **2003**, 807−818. (k) Stone, M. T.; Moore, J. S. *Org. Lett.* **2004**, *6*, 469.
(14) (a) Young, J. K.; Nelson, J. C.; Moore, J. S. *J. Am. Chem. Soc.* **1994**, *116*, 10841−10842. (b) Nelson, J. C.; Young, J. K.; Moore, J. S. *J. Org. Chem.* **1996**, *61*, 8160−8168.
(15) (a) Prince, R. B.; Saven, J. G.; Wolynes, P. G.; Moore, J. S. *J. Am. Chem. Soc.* **1999**, *121*, 3114. (b) Prince, R. B.; Barnes, S. A.; Moore, J. S. *J. Am. Chem. Soc.* **2000**, *122*, 2758. (c) Prince, R. B.; Okada, T.; Moore, J. S. *Angew. Chem., Int. Ed.* **1999**, *38*, 233−236. (d) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chem. Rev.* **2001**, *101*, 3893.
(16) (a) Hartley, C. S.; Elliott, E. L.; Moore, J. S. *J. Am. Chem. Soc.* **2007**, *129*, 4512. (b) Elliott, E. L.; Hartley, C. S.; Moore, J. S. *Chem. Commun.* **2011**, *47*, 5028−5030.
(17) Sonogashira, K. J. *J. Organomet. Chem.* **2002**, *653*, 46−49.
(18) (a) Lawrence, J.; Lee, S.-H.; Abdilla, A.; Nothling, M. D.; Ren, J. M.; Knight, A. S.; Fleischmann, C.; Li, Y.; Abrams, A. S.; Schmidt, B. V. K. J.; Hawker, M. C.; Connal, L. A.; McGrath, A. J.; Clark, P. G.; Gutekunst, W. R.; Hawker, C. J. *J. Am. Chem. Soc.* **2016**, *138*, 6306. (b) Lawrence, J.; Goto, E.; Ren, J. M.; McDearmon, B.; Kim, D. S.; Ochiai, Y.; Clark, P. G.; Laitar, D. S.; Higashihara, T.; Hawker, C. J. *J. Am. Chem. Soc.* **2017**, *139*, 13735.
(19) Hunter, C. A.; Anderson, H. L. *Angew. Chem., Int. Ed.* **2009**, *48*, 7488−7499.

DOI: 10.1021/jacs.8b08087
J. Am. Chem. Soc. 2018, 140, 11526−11536