Peptide Self-Assembly Hot Paper

Peptide Backbone Directed Self-Assembly of Merocyanine Oligomers into Duplex Structures

Bin Liu, Yvonne Vonhausen, Alexander Schulz, Claudia Höbartner,* and Frank Würthner*

In memory of Prof. Dr. Ulf Diederichsen

Abstract: The pseudopeptide backbone provided by *N*-(2-aminoethyl)-glycine oligomers with attached nucleobases has been widely utilized in peptide nucleic acids (PNAs) as DNA mimics. Here we demonstrate the suitability of this backbone for the formation of structurally defined dye stacks. Toward this goal a series of peptide merocyanine (PMC) dye oligomers connected to a *N*-(2-aminoethyl)-glycine backbone were prepared through peptide synthesis. Our concentration-, temperature- and solvent-dependent UV/Vis absorption studies show that under the control of dipole–dipole interactions, smaller-sized oligomers consisting of one, two or three dyes self-assemble into defined duplex structures containing two up to six chromophores. In contrast, upon further extension of the oligomer, the chosen peptide backbone cannot direct the formation of a defined duplex architecture anymore due to intramolecular aggregation between the dyes. For all aggregate species a moderate aggregation-induced emission enhancement is observed.

Dye assembly has attracted considerable attention in the past two decades.[1] This is not only because well-ordered ensembles of multiple chromophores afford unique and desirable photophysical and photochemical properties different from those of the monomers but also because the functional properties of π-conjugated molecular solid-state materials are governed by the electronic coupling between the molecular building blocks.[2,3] To explore structural and

[*] Dr. B. Liu, Y. Vonhausen, A. Schulz, Prof. C. Höbartner, Prof. F. Würthner
Institut für Organische Chemie
Universität Würzburg
Am Hubland, 97074 Würzburg (Germany)
E-mail: claudia.hoebartner@uni-wuerzburg.de
wuerthner@uni-wuerzburg.de

Prof. C. Höbartner, Prof. F. Würthner
Center for Nanosystems Chemistry (CNC)
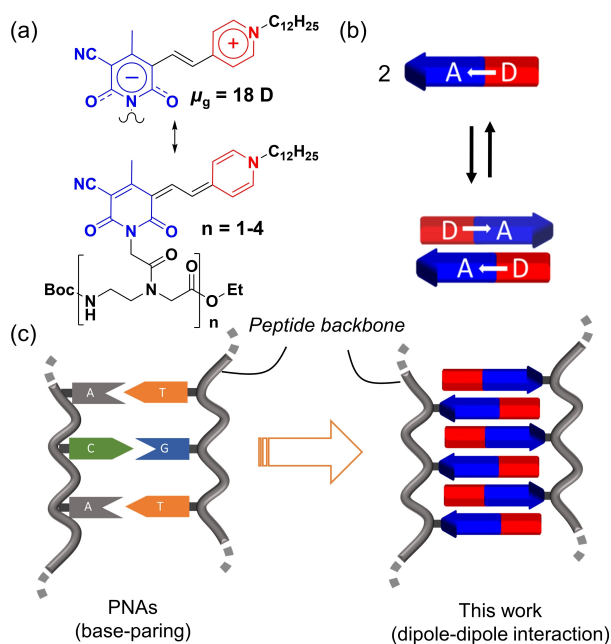Universität Würzburg
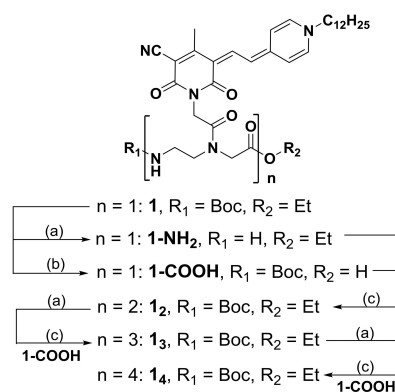Theodor-Boveri-Weg, 97074 Würzburg (Germany)

functional properties, multichromophore systems with well-defined size and structure are needed.[1,4] So far, chemists have developed a variety of strategies to construct multi-chromophore assemblies which include steric shielding,[5,6] templating,[7,8] and backbone-directed self-assembly.[9–11] These examples reveal general strategies for the design of dimers and tetramers but also highlight the lack of a general approach towards larger oligomers of defined structure and size. In contrast, the well-ordered double helix structure of DNA demonstrates a base pair guided backbone-directed self-assembly process that enables the formation of DNA double strands of any length. In principle, the sugar phosphate backbone found in DNA can also be utilized for the assembly of other π-scaffolds which, however, limits the research to aqueous environments and π-scaffolds often appended to natural nucleobases.[12,13] In another approach, dyes are embedded within DNA duplex structures by replacing entire nucleotides in the phosphodiester backbone, but here it is challenging to accommodate the dyes within the DNA backbone in a predictable supramolecular arrangement due to the spatial constraints of the DNA duplex.[14–17] Inspired by research on peptide nucleic acids (PNAs)[18,19] and research on the folding of peptide backbones driven by donor-acceptor interactions[20,21] we herein explore the suitability of *N*-(2-aminoethyl)-glycine-based pseudopeptide backbones for the formation of oligomeric dye duplex structures (Figure 1).

N-(2-aminoethyl)-glycine-based pseudopeptide backbones have indeed been used very successfully as a substitute for the (deoxy)ribose phosphate backbone in nucleic acids.[22,23] Because of the similarity of the structures, PNA can also form complementary duplexes of high stability with DNA and RNA through base-paring.[24] By introducing cyanine dyes into PNA, a series of PNA-FIT probes was successfully constructed and applied to the imaging of specific mRNA molecules.[25] Therefore, if a chromophore is used to replace the nucleobase on PNA, it should be possible to direct the formation of duplexes consisting of multiple chromophores. Advantageously, the *N*-(2-aminoethyl)-glycine-based peptide backbone is not charged and quite hydrophobic, thereby not limiting the studies to polar solvents and water. Whilst we anticipate that different chromophores can be easily attached to the PNA skeleton by stepwise deprotection-coupling protocols, thereby enabling a large variety of hetero dye stacks, in our initial study we focus on the formation of dye stacks consisting of only one dye, i.e. the dipolar merocyanine

Figure 1. Concept for a peptide backbone-directed self-assembly of peptide MC (PMC) oligomers into duplex structures. a) Neutral and zwitterionic resonance structures of the merocyanine (MC) chromophore, and the chemical structures of the oligomers investigated in this study. b) Self-assembly of dipolar MCs into antiparallel stacked dimers. c) Schematic representation of the self-assembly of the complementary duplex of PNAs and peptide MC oligomers.


Scheme 1. Synthesis of PMC oligomers. Reaction conditions: a) 4 M HCl in dioxane, DCM, 1 h, rt; b) NaOH aq, THF/MeOH, 2 h, rt; c) HBTU, DIPEA, DMF, 3 h, rt.

(MC) dye shown in Figure 1a.[26] As shown in our earlier work the self-assembly of this MC dye is strongly supported by electrostatic dipole-dipole interactions (Figure 1b) due to a high ground state dipole moment ($\mu_g = 18$ D) originating from a pronounced contribution of the zwitterionic resonance structure (Figure 1a).[26,27] Therefore, the interaction between these MCs is directional, similar to the directionality of hydrogen bonds.[28] Accordingly, similar to the directional base pairing that directs the self-assembly of nucleic acid sequences in DNA, directional interactions of MC dyes should drive peptide backbone-attached MC oligomers (Figure 1a) into ordered double-stranded architectures under controlled conditions (Figure 1c).[29]

Peptide merocyanine (PMC) oligomers 1_n were synthesized according to the route outlined in Scheme 1 (see the Supporting Information for details). Monomer **1** bearing protected amine and carboxylic acid functional groups was obtained via a convenient four-step reaction sequence with yields of 66–78% for the individual steps (Scheme S1). The synthesis of PMCs 1_n follows a classical peptide synthesis strategy (Scheme 1). First, the Boc protecting group of **1** is removed under acidic conditions to form the corresponding amine **1-NH$_2$**. Second, the ethyl protecting group of the ester is removed by an aqueous solution of sodium hydroxide to give **1-COOH**. Last, **1-NH$_2$** and **1-COOH** are coupled under activation with HBTU/DIPEA to generate dimer **1$_2$** in a yield of 45%. Similarly, by repetitive deprotection and coupling steps with **1-COOH**, **1$_3$** and **1$_4$** were obtained in 41% and 35% yield, respectively. The fact that the yields

remain in the same range despite of chain extension is promising with respect to future research on larger oligomers.

The self-assembly properties of **1**, **1$_2$**, **1$_3$**, and **1$_4$** were studied by concentration-dependent UV/Vis spectroscopy in CHCl$_3$ (Figures 2 and S1). Monomer **1** shows only one narrow absorption band at 556 nm at a low concentration of 1.4×10^{-7} M (Figure 2a). With increasing concentration a new hypsochromically shifted band appears at 475 nm that is characteristic for co-facially stacked MC dimers with H-type coupled transition dipole moments.[27,30] The concomitant decrease of the monomer band and an isosbestic point at 496 nm indicate a two-state equilibrium process between monomer and dimer species.

Unlike those simple two absorption bands displayed during the self-assembly process of monomer **1**, PMC oligomers **1$_2$**, **1$_3$**, and **1$_4$** exhibit already two absorption bands located at 560 nm (monomer-like) and 480 nm (dimer-like) even at very low concentrations (Figure 2b–d). The fact that these bands do not change in the lower concentration regime (Figure S2) suggests a partial folding of these oligomers (Figure 3), similar as observed for some previously investigated MC dyes that were interconnected by alkyl chains of variable length.[31] Due to the constraints imparted by the *N*-(2-aminoethyl)-glycine spacer units perfectly antiparallel stack arrangements cannot be achieved for the dipolar MC dyes and accordingly in addition to the most intense hypsochromically shifted band (located close to the dimer band observed for **1**) also J-type exciton states at lower energy may exhibit notable oscillator strength for such twisted geometries. Furthermore, within the conformational manifold, especially for the longer PMCs **1$_3$** and **1$_4$**, the presence of partially folded state is most likely that consist of strongly coupled dimeric and weakly coupled monomeric units in equilibrium, leading to the band structure observed in the absorption spectra at low concentrations.[32]

As the concentration gradually increases, however, self-assembly into duplex structures takes place as evidenced by hypsochromically shifted absorption bands at 447 nm and 441 nm for PMCs **1$_2$** and **1$_3$**, respectively. Thus, as expected for co-facially stacked dyes,[34] H-type coupling between
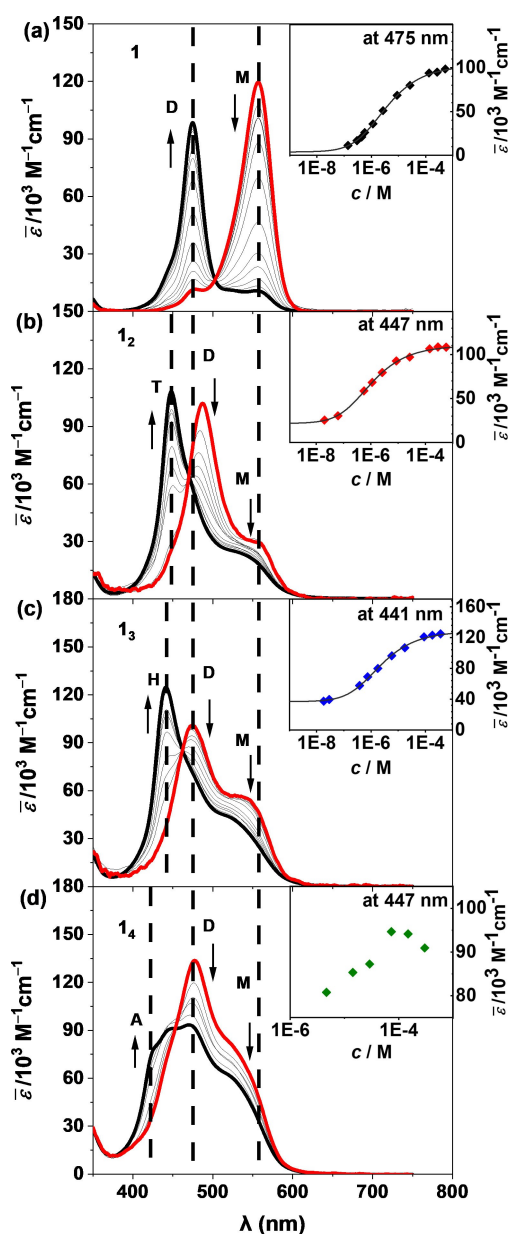
Figure 2. Concentration-dependent UV/Vis spectra of a) **1** ($c=1.40\times10^{-7}$–5.29×10^{-4} M), b) **1₂** ($c=2.04\times10^{-8}$–5.74×10^{-4} M), c) **1₃** ($c=1.81\times10^{-8}$–3.64×10^{-4} M), and d) **1₄** ($c=4.65\times10^{-6}$–2.99×10^{-4} M) in CHCl₃ at 298 K. Black arrows indicate the increase of the aggregate band (dimer band for **1** (D), tetramer band for **1₂** (T), hexamer band for **1₃** (H) and undefined aggregate band (A) for **1₄**) and the decrease in the intensity of the monomer (M) and dimer bands (D) upon increasing concentration. The dashed lines indicate the positions of the absorption maxima of the monomer, dimer and the various aggregate species. The inserts show concentration-dependent apparent extinction coefficients at 298 K for **1** (at 475 nm), **1₂** (at 447 nm) and **1₃** (at 441 nm) at the maximum of the aggregate band in comparison with the simulated curve (solid lines) according to the dimer model. Note: the data for **1₄** (at 447 nm) could not be properly fitted with dimer model due to the presence of more than two species in the investigated concentration range as indicated by the data points at 447 nm in the inset.
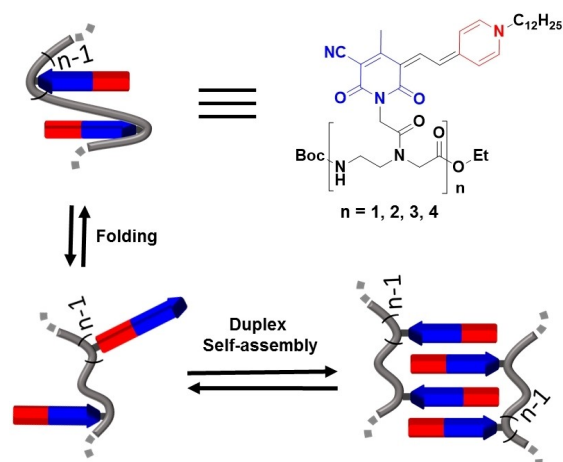
Figure 3. Schematic representation of folding and duplex self-assembly processes observed for PMC oligomers.[33]

neighboring dyes shifts the most intense highest energy exciton state progressively with increasing stack size by 0.38 eV (dimer stack **1**), 0.54 eV (tetramer stack of **1₂**) and 0.58 eV (hexamer stack of **1₃**) towards higher energy. For all concentration-dependent spectra of these oligomers (**1**, **1₂**, and **1₃**), we notice a well-defined isosbestic point in the studied concentration range, indicating a thermodynamic equilibrium between two different species, (folded) monomeric oligomers and duplex aggregates.

The concentration-dependent UV/Vis spectra of PMCs **1**, **1₂**, and **1₃** were fitted with a dimerization model through a global fit algorithm (Figure S1).[29,35] The inserts of Figure 2 show the experimental data points of the concentration-dependent absorbance at the absorption maximum of the duplex (475 nm for **1**, 447 nm for **1₂**, and 441 nm for **1₃**) as well as the calculated lines from the global fit analysis. The excellent match of experimental and calculated data throughout the whole concentration range supports our hypothesis that the oligomers self-assemble in duplex structures. This result is further corroborated by atomic force microscopy (AFM, Figure S10) and DOSY NMR (Figure S12) studies which both indicate the absence of larger aggregate species. For the DOSY experiments the calculated hydrodynamic radii (r_H) of the assemblies (**1**)₂, (**1₂**)₂, and (**1₃**)₂ are 10.9, 15.5, and 22.4 Å, respectively. This result further proves that as the length of the oligomer increases, the size of the assembled structures increases. The dimerization constants calculated from our UV/Vis data for the duplex formation of PMCs **1**, **1₂**, and **1₃** in CHCl₃ at 298 K are 3.0×10^{5}, 1.1×10^{6}, and 3.8×10^{5} M^{-1}, respectively. The similar magnitude of these values is at first glance surprising because the number of π–π-contacts and concomitant dipole-dipole interactions increases significantly within the series from one (for **1**) to three (**1₂**) and five (**1₃**). However, as discussed before, intramolecular dye-dye contacts originating from folding compete with intermolecular ones as illustrated in Figure 3, thereby reducing the Gibbs energy for duplex formation for the oligomers. Further insights from temperature-dependent studies in 5 % DMSO/

1,4-dioxane[36] (Figures S3, S4) indeed confirm the enthalpic gain upon duplex formation for the larger oligomers (Table S1) originating from the increased number of perfect antiparallel dye-dye contacts and the associated entropic cost due to the loss of motional freedom for the dyes and the backbone in the duplex. This leads to steeper melting curves and an increase of the melting temperatures in the series from $T_m = 296$ K for **1** to $T_m = 340$ K for **1$_3$** (Figure S4 and Table S1), similarly as observed for the duplex formation between smaller oligonucleotides.[37]

The presented results for the smaller oligomers **1**, **1$_2$** and **1$_3$** revealed supramolecular structures that result from a competition between intramolecular folding and bimolecular self-assembly that can be controlled either by the concentration or the temperature (Figure 3). At low concentrations or elevated temperature, these oligomers tend to fold intramolecularly while the self-assembled duplex structure is preferred at higher concentrations and lower temperatures. Because of the large contribution of dipole-dipole interactions, this competition is also affected by the polarity of the solvent. Thus, for the compared to CHCl$_3$ less polar solvent dioxane self-assembled duplex formation is even more favored whilst upon addition of the polar solvent DMSO the dis-assembly is observable at constant concentration (Figure S5). For this dissociated state the UV/Vis spectra also clearly indicate the presence of less folded structures with a more pronounced monomer-like absorption band due to the randomly oriented chromophores. In line with our earlier observations of an aggregation-induced emission enhancement for this class of dyes upon self-assembly into a more rigid aggregate state,[26b,32,35] solvent-dependent fluorescence studies (Figure S8 and S9) afforded fluorescence quantum yields of 1.4–1.8 % for the monomeric species of all four compounds in dilute DMSO which increased to 2.6–4.6 % in dilute DMSO/dioxane (1:99) where duplex structures (**1**, **1$_2$** and **1$_3$**) or folded aggregates (**1$_4$**, vide infra) prevail.

For **1$_4$** we note an obviously different behavior. This was already indicated by the concentration-dependent UV/Vis spectra shown in Figure 2d and is further corroborated by the solvent-dependent studies shown in Figure S7. Both data sets suggest a thermodynamically more favored monomer folded state, similar as known for protein folding, which might be due to the multiple options for the formation of antiparallelly stacked dye pairs that resists dissociation even in the most dipolar DMSO solvent (Figure S5). Upon decreasing the temperature (Figure S6) or reducing the solvent polarity by addition of methylcyclohexane into CHCl$_3$ solutions (Figure S7) the hypsochromically shifted band edge (H-aggregate) becomes more pronounced, thereby indicating the formation of extended dye stacks by self-assembly. However, due to a lack of isosbestic points we attribute these changes not to a structurally defined duplex but to polymeric aggregates which is also corroborated by AFM studies (Figure S11). Thus, for **1$_4$** the here utilized peptide backbone cannot direct the preferential formation of a MC duplex structure and depending on conditions (temperature, solvent, concentration) various folded and self-assembled species coexist.

In summary, in this study we reported a new approach toward structurally defined dye aggregates. A series of MC dye oligomers was synthesized by replacing the nucleobases of PNA with dipolar MCs. Driven by the dipole–dipole interactions **1**, **1$_2$**, and **1$_3$** oligomers self-assemble into duplex structures containing two, four, and six π-stacked chromophores under the guidance of the peptide backbone. The competing processes of intramolecular folding and intermolecular self-assembly for these oligomers can be controlled through appropriate conditions (solvent, temperature and concentration). We envision that duplex structures can also be accomplished for other dyes connected to the *N*-(2-aminoethyl)-glycine-based pseudopeptide backbones and that subtle modifications of this backbone may also enable the duplex formation for highly desired extended oligomer strands bearing eight, ten or more π-stacked chromophores. However, already the herein reported tetra- and hexamer stacks should be quite useful, for instance by incorporating them within DNA backbones where their highly structure-dependent exciton coupling can sense structural changes upon hybridization,[25b] hairpin formation or catalysis.[38]

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1] D. Bialas, E. Kirchner, M. I. Röhr, F. Würthner, *J. Am. Chem. Soc.* **2021**, *143*, 4500–4518.

[2] Z. B. Henson, K. Müllen, G. C. Bazan, *Nat. Chem.* **2012**, *4*, 699–704.

[3] F. C. Spano, C. Silva, *Annu. Rev. Phys. Chem.* **2014**, *65*, 477–500.

[4] M. R. Wasielewski, *Acc. Chem. Res.* **2009**, *42*, 1910–1921.

[5] M. Hutin, J. K. Sprafke, B. Odell, H. L. Anderson, T. D. W. Claridge, *J. Am. Chem. Soc.* **2013**, *135*, 12798–12807.

[6] M. M. Safont-Sempere, P. Osswald, K. Radacki, F. Würthner, *Chem. Eur. J.* **2010**, *16*, 7380–7384.

[7] D. González-Rodríguez, J. L. J. van Dongen, M. Lutz, A. L. Spek, A. P. H. J. Schenning, E. W. Meijer, *Nat. Chem.* **2009**, *1*, 151–155.

[8] Y. Yamauchi, M. Yoshizawa, M. Akita, M. Fujita, *J. Am. Chem. Soc.* **2010**, *132*, 960–966.

[9] S. Bhosale, A. L. Sisson, P. Talukdar, A. Fürstenberg, N. Banerji, E. Vauthey, G. Bollot, J. Mareda, C. Röger, F. Würthner, N. Sakai, S. Matile, *Science* **2006**, *313*, 84–86.

[10] C. Shao, M. Stolte, F. Würthner, *Angew. Chem. Int. Ed.* **2013**, *52*, 10463–10467; *Angew. Chem.* **2013**, *125*, 10657–10661.

[11] A. Das, S. Ghosh, *Angew. Chem. Int. Ed.* **2014**, *53*, 2038–2054; *Angew. Chem.* **2014**, *126*, 2068–2084.

[12] P. G. A. Janssen, J. Vandenbergh, J. L. J. van Dongen, E. W. Meijer, A. P. H. J. Schenning, *J. Am. Chem. Soc.* **2007**, *129*, 6078–6079.

[13] W. Xu, K. M. Chan, E. T. Kool, *Nat. Chem.* **2017**, *9*, 1043–1055.

[14] V. L. Malinovskii, D. Wenger, R. Häner, *Chem. Soc. Rev.* **2010**, *39*, 410–422.

[15] P. Ensslen, H.-A. Wagenknecht, *Acc. Chem. Res.* **2015**, *48*, 2724–2733.

[16] D. Baumstark, H.-A. Wagenknecht, *Chem. Eur. J.* **2008**, *14*, 6640–6645.

[17] C. B. Winiger, S. M. Langenegger, G. Calzaferri, R. Häner, *Angew. Chem. Int. Ed.* **2015**, *54*, 3643–3647; *Angew. Chem.* **2015**, *127*, 3714–3718.

[18] P. E. Nielsen, M. Engholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497–1500.

[19] M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566–568.

[20] R. Scott Lokey, B. L. Iverson, *Nature* **1995**, *375*, 303–305.

[21] X. Zhao, M. X. Jia, X. K. Jiang, L. Z. Wu, Z. T. Li, G. J. Chen, *J. Org. Chem.* **2004**, *69*, 270–279.

[22] B. Hyrup, P. E. Nielsen, *Bioorg. Med. Chem.* **1996**, *4*, 5–23.

[23] P. E. Nielsen, G. Haaima, *Chem. Soc. Rev.* **1997**, *26*, 73–78.

[24] P. E. Nielsen, M. Egholm, *Curr. Issues Mol. Biol.* **1999**, *1*, 89–104.

[25] a) S. Kummer, A. Knoll, E. Socher, L. Bethge, A. Herrmann, O. Seitz, *Bioconjugate Chem.* **2012**, *23*, 2051–2060; b) G. Fang, J. Chamiolo, S. Kankowksi, F. Hövelmann, D. Friedrich, A. Löwer, J. C. Meier, O. Seitz, *Chem. Sci.* **2018**, *9*, 4794–4800.

[26] a) F. Würthner, S. Yao, *Angew. Chem. Int. Ed.* **2000**, *39*, 1978–1981; *Angew. Chem.* **2000**, *112*, 2054–2057; b) U. Rösch, S. Yao, R. Wortmann, F. Würthner, *Angew. Chem. Int. Ed.* **2006**, *45*, 7026–7030; *Angew. Chem.* **2006**, *118*, 7184–7188.

[27] F. Würthner, S. Yao, T. Debaerdemaeker, R. Wortmann, *J. Am. Chem. Soc.* **2002**, *124*, 9431–9447.

[28] F. Würthner, *Acc. Chem. Res.* **2016**, *49*, 868–876.

[29] For other approaches to accomplish ordered MC dye stacks via dipolar aggregation within polymeric architectures, see: a) A. Rajak, C. K. Karan, P. Theato, A. Das, *Polym. Chem.* **2020**, *11*, 695–703; b) A. Rajak, A. Das, *ACS Polym. Au* **2022**, https://doi.org/10.1021/acspolymersau.1c00054.

[30] M. Kasha, H. Rawls, M. A. El-Bayoumi, *Pure Appl. Chem.* **1965**, *11*, 371–392.

[31] A. Zitzler-Kunkel, E. Kirchner, D. Bialas, C. Simon, F. Würthner, *Chem. Eur. J.* **2015**, *21*, 14851–14861.

[32] X. Hu, A. Schulz, J. O. Lindner, M. Grüne, D. Bialas, F. Würthner, *Chem. Sci.* **2021**, *12*, 8342–8352.

[33] The peptide backbone utilized in this study has a directionality and accordingly we may anticipate either the formation of parallel or antiparallel duplex structures. Due to the strong dipole–dipole interactions between dipolar merocyanine dyes (much stronger than hydrogen bonds between nucleobases in PNA) the here studied duplexes are assumed to be mixtures of parallel and antiparallel strands. This explains why we could not get insights into the duplex structures by 2D NMR spectroscopy.

[34] E. Kirchner, D. Bialas, F. Fennel, M. Grüne, F. Würthner, *J. Am. Chem. Soc.* **2019**, *141*, 7428–7438.

[35] Y. Vonhausen, A. Lohr, M. Stolte, F. Würthner, *Chem. Sci.* **2021**, *12*, 12302–12314.

[36] Due to the low boiling point of CHCl₃ we prepared a DMSO/dioxane solvent mixture that provides a similar polarity but allows for a larger temperature range for the temperature-dependent self-assembly studies.

[37] J. SantaLucia, H. T. Allawi, P. A. Seneviratne, *Biochemistry* **1996**, *35*, 3555–3562.

[38] R. Micura, C. Höbartner, *Chem. Soc. Rev.* **2020**, *49*, 7331–7353.