Supporting Information

A Discrete Three-Layer Stack Aggregate of a Linear Porphyrin Tetramer: Solution-Phase Structure Elucidation by NMR and X-Ray Scattering

*Marie Hutin, Johannes K. Sprafke, Barbara Odell, Harry L. Anderson * and Tim D. W. Claridge**

The Chemistry Research Laboratory, 12 Mansfield Road, OX1 3TA Oxford, UK

harry.anderson@chem.ox.ac.uk, tim.claridge@chem.ox.ac.uk

Table of Contents

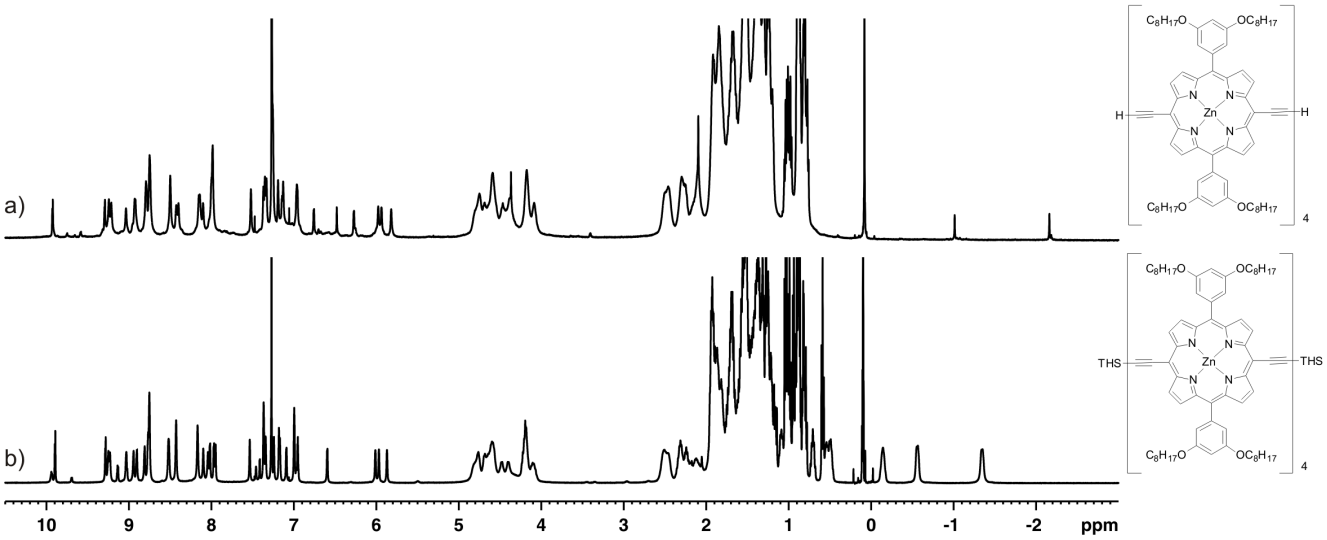

Figure S1. ^{1}H NMR spectra of a) the aggregate of the deprotected tetramer and b) the aggregate of the THS protected tetramer in $CDCl_3$, 500 MHz, 298 K.

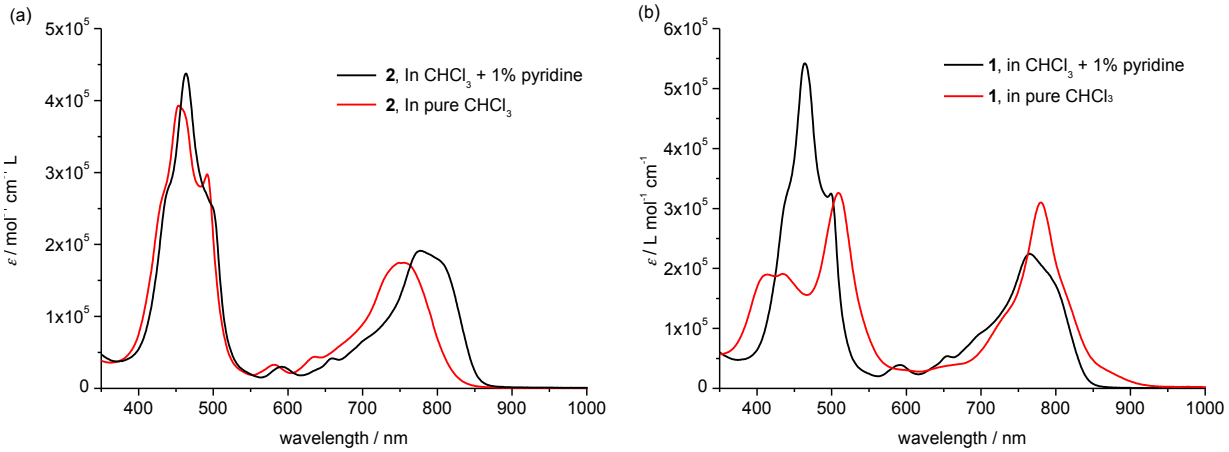

Figure S2. UV/vis absorption spectra of a) **2** and b) **1** with (black lines) and without (red lines) pyridine in $CHCl_3$ at 298 K.

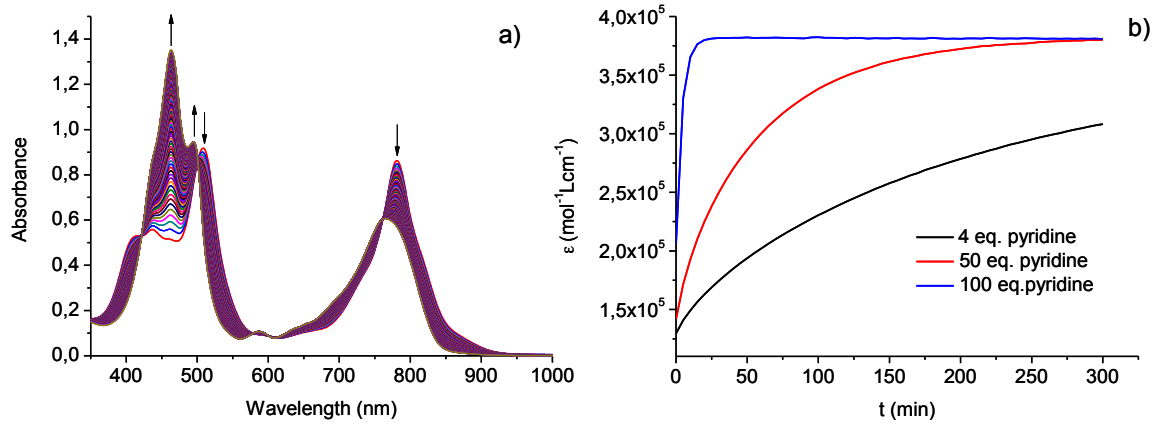

Figure S3. a) Evolution of the UV-vis spectrum of the aggregate of tetramer **1** after addition of 4 equivalent of pyridine, $CHCl_3$, 298 K and b) Kinetic of disaggregation of tetramer **1** with different amounts of pyridine in $CHCl_3$, 298 K.

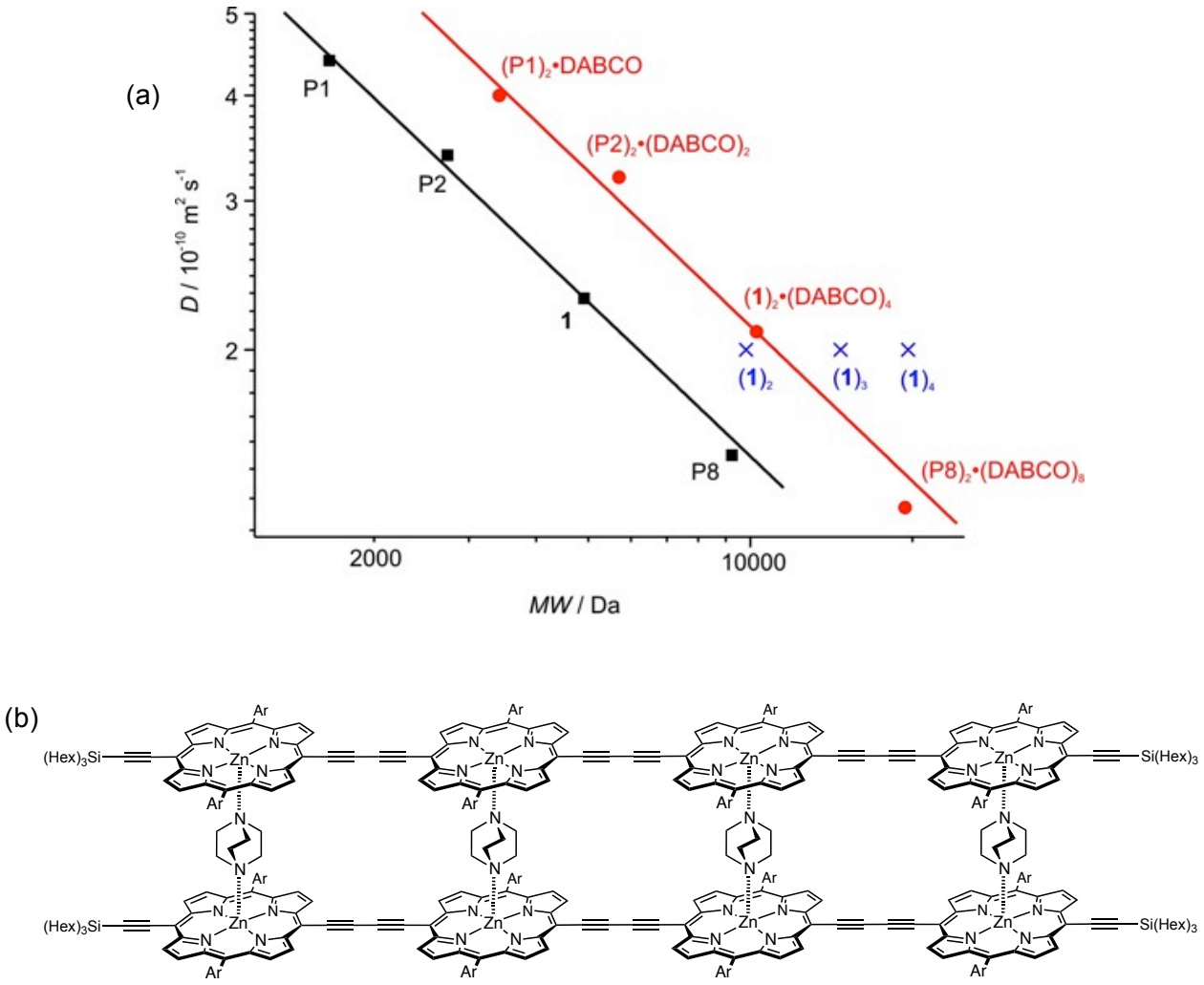

Figure S4. (a) Double logarithmic plot of the diffusion coefficients against the molecular weight of a series of disaggregated porphyrin oligomers (as pyridine adducts; black squares, P*N* where *N* is number of porphyrin units; **1** ≡ P4) and the corresponding oligomer-DABCO ladder complexes (red circles) fitted to exponential decays ($D = 400 \times 10^{-10}\ MW^{-0.61}$, black line and $D = 584 \times 10^{-10}\ MW^{-0.61}$, red line). Both series of data give individually good fits but the greater compactness of the ladder complexes results in reduced apparent molecular weights compared to the disaggregated oligomers. The tetramer aggregate (dimeric, trimeric and tetrameric, blue crosses) is expected to be even more compact than the ladder complexes and thus it is difficult to accurately predict its size. However, the formation of large, polymeric aggregates can be excluded from this plot and 2–4 porphyrin tetramers in the aggregate seems to be the most plausible number. NMR diffusion data were measured at 500 MHz, 298 K in $CDCl_3$ for the ladders and the aggregate and in $CDCl_3$/1% d_5-pyridine for the disaggregated oligomers. (b) Structure of the $\mathbf{1_2}$**•(DABCO)$_4$** complex, which has a similar diffusion coefficient to the $\mathbf{1}_3$ aggregate.

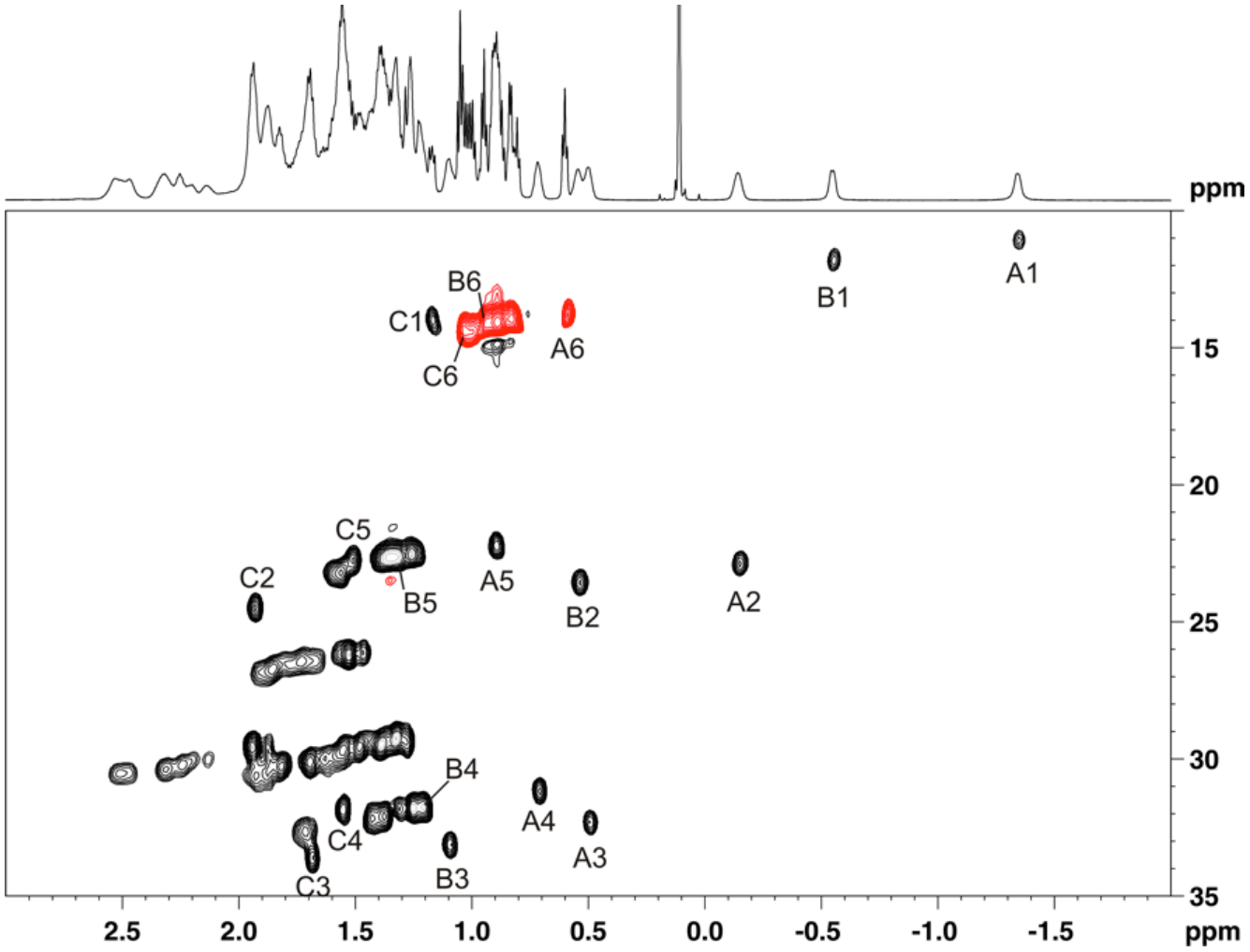

Figure S5. ^{1}H-^{13}C multiplicity edited HSQC ($CDCl_3$, 700 MHz, 298 K) expansion showing the assigned correlations of the shielded hexylsilyl sidechains of aggregate $\mathbf{1}_3$. Red crosspeaks arise from methyl CH_3 groups, black from methylene CH_2 groups. The positions along the hexyl chain are labeled 1–6 as Figures 5 and S9.

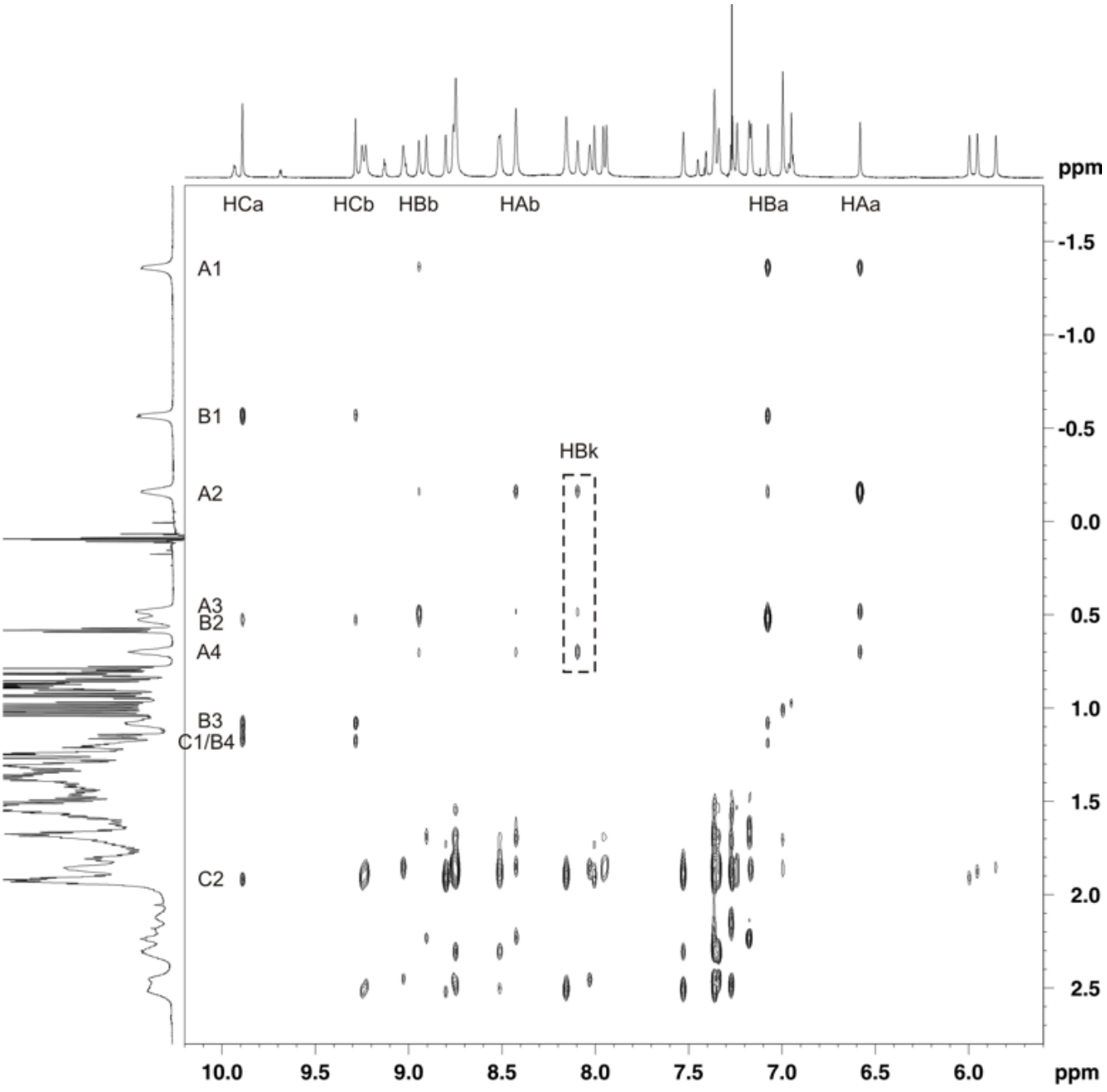

Figure S6. Expansion of the ^{1}H NOESY spectrum of the tetramer aggregate of **1** ($CDCl_3$, 700 MHz, 298 K, 250 ms mixing time) showing correlations between terminal β-pyrrole protons and the hexylsilyl sidechain protons of strands A, B and C. The boxed region shows correlations between the HBk *ortho*-aryl proton and the hexylsilyl chain protons of strand A. These correlations are consistent with the ends of the three tetramers sitting in an overlayed but offset geometry. The positions along the hexyl chain are labeled 1–6 as Figures 5 and S9.

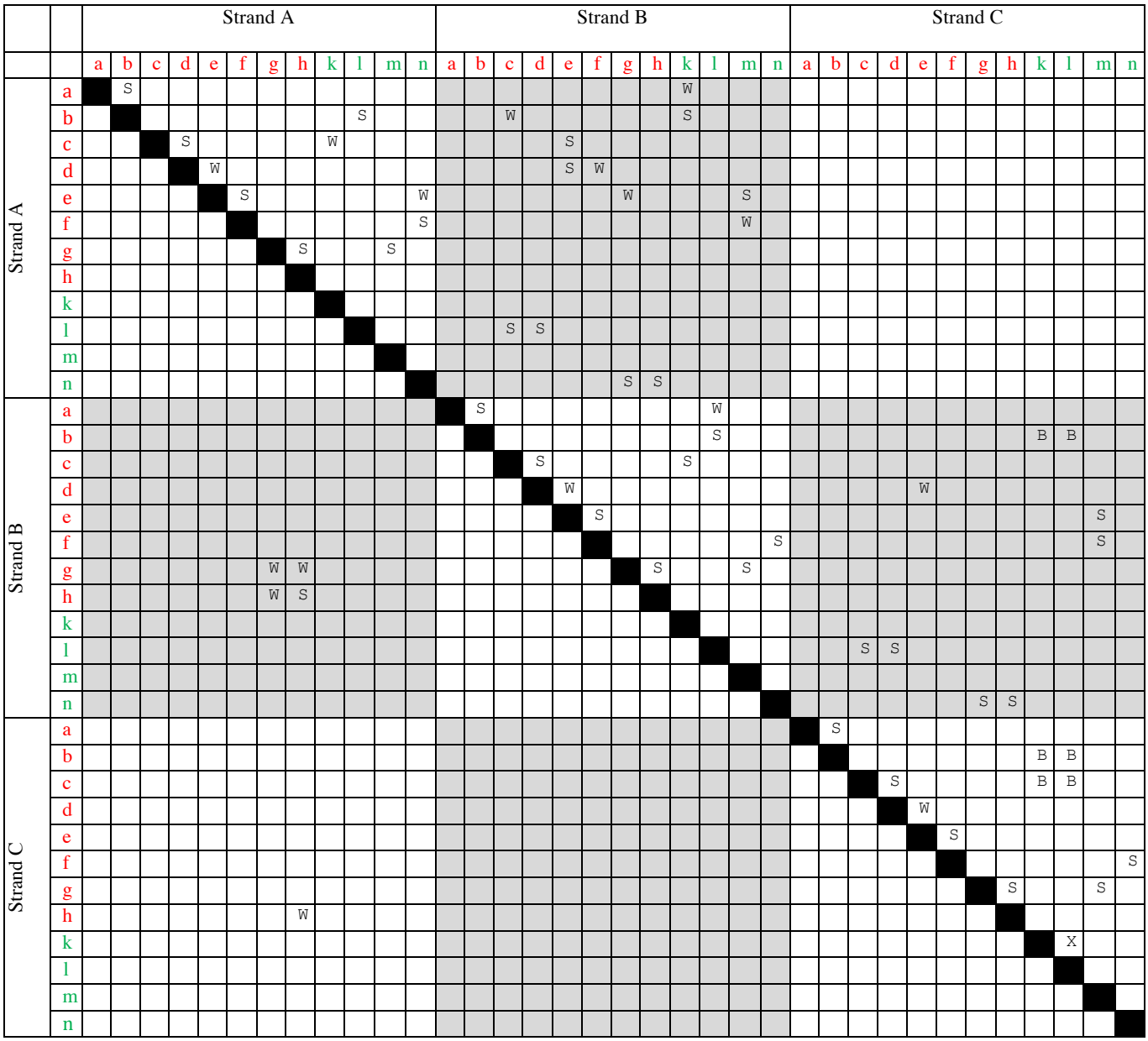

Figure S7. Graphical summary of the NOEs observed in 2D NOESY spectra amongst β-pyrrole (red) and *ortho*-aryl (green) protons in the tetramer aggregate $\mathbf{1}_3$ (no direct NOEs are observed involving the *para*-aryl protons i and j except to the adjacent octyloxy chains). Intensities are qualitatively assigned as being strong (**S**) or weak (**W**). Label **X** indicates the observed crosspeak arises from exchange between the broad *ortho*-aryl protons toward the terminus of strand C and labels **B** indicate NOEs involving these protons which are likely to be exchange averaged. NOEs tabulated in the lower half of the grid occur across the symmetry interface at the center of the aggregate. Grid shading is included to differentiate strands and guide the eye.

Details of NMR Resonance Assignments

In order to demonstrate consistency of the available NMR data with the proposed molecular model, we wished to determine as complete a set of resonance assignments as possible. This would also enable assignments of residual ^{1}H-^{13}C dipolar couplings to specific positions in the modeled structure, allowing computation of the magnetic field alignment tensor for the tetramer aggregate $\mathbf{1}_3$ and hence further evidence for the proposed parallel-stacked structure.

Initial ^{1}H assignments were made by comparison with chemical shift data of the disaggregated tetramer $\mathbf{1{\cdot}(py)_4}$ for which assignments were known, and from logical chemical shift arguments based on likely shielding effects. During the course of the study, data from the trimer aggregate also became available (Figure S8) and provided evidence for which β-pyrrole resonances in the tetramer aggregate must arise from those protons at the symmetry interface HAg/h, HBg/h and HCg/h (Figure 4a) as these were clearly missing in the trimer spectra (see discussions below). Site specific assignment of the porphyrin β-pyrrole and aryl protons was ultimately possible from thorough analyses of 2D COSY and NOESY spectra recorded at 500 and 700 MHz.

From 2D COSY data it was possible to identify the 12 unique pairs of pyrrole protons expected for the triple-stranded aggregate. Of these, only one pair exhibited proton chemical shifts similar to those observed for the disaggregated pyridine adduct and originated from the exposed proton pairs HCa-HCb at the end of strand C. Each of the remaining 11 pairs could be classified as having one proton resonating above 8.3 ppm (later assigned as that adjacent to the *meso*-aryl substituent) with the partner β-proton resonating below this shift (later assigned as that adjacent to the butadiyne linker) and experiencing significant ring-current shielding effects leading to shifts as low as ~ 6 ppm in some cases (Figure 4b).

In an attempt to assign resonances to specific strands and specific locations within these strands, we investigated NOE interactions within the complex via 2D NOESY spectra. Negative NOEs were observed in all cases, consistent with large rotational correlations times. An informative region was that of the shielded hexylsilyl groups below 0 ppm which gave rise to enhancements with six β-pyrrolic protons which were thus assigned as the terminal β-proton pairs (Ha and Hb) of the three strands. The two most deshielded protons (9.94 and 9.33 ppm) gave NOEs to the hexyl chain that experienced minimal ring current shielding (as identified in the 1D TOCSY experiments; Figure 5) and were thus assigned as the terminal β-pyrrole pair of strand C; HCa and HCb respectively (Figure S6). As noted above, these β-pyrrole proton shifts most closely resemble those of the disaggregated pyridine adduct $\mathbf{1{\cdot}(py)_4}$ (Ha 9.65 and Hb 8.95 ppm), consistent with assignment to the relatively exposed end of strand C. Protons HCa and HCb displayed further NOEs to protons of a second hexylsilyl chain that experienced significant ring current shielding, which were therefore assigned to the terminal chain of strand B. Notably, protons for methylene groups B1 and B2 resonated below 1 ppm (Figure 5) with the shielding being ascribed to these groups sitting directly above the terminal porphyrin ring of strand C (Figure 4a). Two further β-

pyrrole protons also displayed NOEs to the hexyl chain of strand B and were thus assigned as the terminal β-pyrrole pairs of strand A. Further NOEs from these pyrrole protons correlated with the third and most highly shielded hexyl chain (displaying four methylene groups below 1 ppm; Figure 5) that was thus assigned to strand A, here shielded by its location above the terminal porphyrin of strand B. For each β-pyrrole pair, that giving rise to the stronger NOEs to the hexylsilyl chain was assigned as proton a and the weaker as proton b. (It is likely that the NOEs involving the protons labeled b arise, at least in part, through spin diffusion of hexyl NOEs via neighboring a protons.) For both strands A and B, the most shielded pyrrole protons corresponded to proton a (HAa 6.65 and HBa 7.14 ppm). However, for strand C the most shielded pyrrole proton was HCb (9.33 ppm) and it is noteworthy that in the disaggregated porphyrin the protons adjacent to the *meso*-aryl substituents (Hb) are likewise the most shielded of the pair (Hb 8.9–9.1 ppm *vs*. Ha 9.6–9.9 ppm). Hence, the observed shifts for the terminal protons of strand C are consistent with these being remote from the core of the complex and not subject to substantial ring current effects. Only a single *ortho*-aryl proton resonance demonstrated an obvious NOE to any hexylsilyl chain; the A2, A3 and A4 methylene protons, suggesting this to be HBk that sits immediately adjacent to the hexylsilyl group of strand A (Figure S6). As noted in the main text, the aryl group on the terminal porphyrin of strand C gave rise to broad *ortho*-aryl proton resonances yet it was possible to discern NOEs from the *ortho*-aryl proton at 8.28 ppm to the B2, B3 and B4 methylene protons, confirming its assignment as HCk.

Further assignments could be secured only through consideration of NOEs involving the *ortho*-aryl protons. Chemical shift arguments, together with the recognition that aryl ring rotations were blocked in the aggregate, initially differentiated those on the outside of the aggregate from those buried within it, which would sit adjacent to the face of at least one neighboring aryl ring and would thus be substantially deshielded. Thus, the 7 aryl protons with shifts > 8.0 pm were classified as belonging to the "inner" protons HAl, HAn, HBk, HBl, HBm, HBn, and HCm (NB: HCk at ~ 8.3 ppm was broadened by ring dynamics and was not readily observed, as described in the main text). The three *ortho*-aryl resonances below 8 ppm were ascribed to the protons on the outer face of the aggregate (HAk, HAm and HCn; again HCl was not easily observed due to dynamic broadening). Their chemical shifts (7.22-7.32 ppm) were similar to those of the *ortho*-aryl protons of the disaggregated tetramer **1·(py)$_4$** (~7.4 ppm) consistent with their solvent exposed location.

Attempts to assign resonances in more detail were initially frustrated by the complexity of the NOE data and the possibility of *inter* as well as *intra*-strand NOEs. The breakthrough in analysis was made through the realization that for each β-pyrrole proton pair, that adjacent to the butadiyne linker may experience NOEs only with *ortho*-aryl protons that sat immediately below it in the stacked aggregate (HAa to HBk, for example) whereas the partner pyrrole proton (that adjacent to the aryl substituent) may show NOEs to the same *ortho*-aryl protons in the strand below *in addition* to the adjacent *ortho*-aryl protons of the same porphyrin unit (HAb to HBk *and* HAl for example). A systematic analysis of the NOE data yielded further assignments when supported by other

information. This included knowledge of the assignments of the three terminal β-pyrrole pairs for each strand and for the *ortho*-aryl protons HBk from NOEs with the hexylsilyl end groups (described above); chemical shift arguments differentiating the inner and outer *ortho*-aryl protons and, as also noted above, the identification of the β-pyrrole pairs at the symmetry interface (Hg/h) through comparison with the trimer aggregate data (although strand-specific assignments were not possible from this comparison alone). During the assignment process, it also became possible to identify site-specific *inter*-strand NOEs between β-pyrrole protons HAd-HBe, HBd-HCe and HBh-HAh (which occurs across the symmetry interface). These correspond to ~ 3.6 Å separations in the model structure and were again wholly consistent with the offset strand stacking. Finally, weaker *intra*-strand NOEs were also observed between protons in neighboring porphyrin units HAd-HAe, HCd-HCe and HAh-HCh (which also exists across the symmetry interface and hence occurs within the same tetramer strand); it is likely that a HBd-HBe NOE also exists but could not be observed owing to the similar chemical shifts of these protons. These inter-porphyrin distances correspond to ~ 4.3 Å in the model. The final NOE assignments are summarized in Figure S7.

Determining assignments for the 6 *para*-aryl protons was complicated by their rather small peak dispersion, consistent with their environments on the outer rim of the stacked strands, yielding clustered ^{1}H and ^{13}C correlations in the HSQC. Due to the isolated environments of these protons, assignments were derived from direct long-range COSY correlations with assigned *ortho*- protons of the same aryl group and by matching NOE correlations from these to the adjacent octyloxy OCH_2 protons. These analyses identified HAi, HBi and HBj. HAj and HCj could not be unambiguously differentiated from each other due to coincidence of their corresponding *ortho*-aryl proton shifts of HAn and HCm so their assignments remain interchangeable. HCi was assigned as it showed no correlations due to the previously described broadening of the *ortho* protons on the aryl ring.

The final ^{1}H and ^{13}C resonance assignments for the pyrrole and aryl protons are tabulated in Table 1 of the main text.

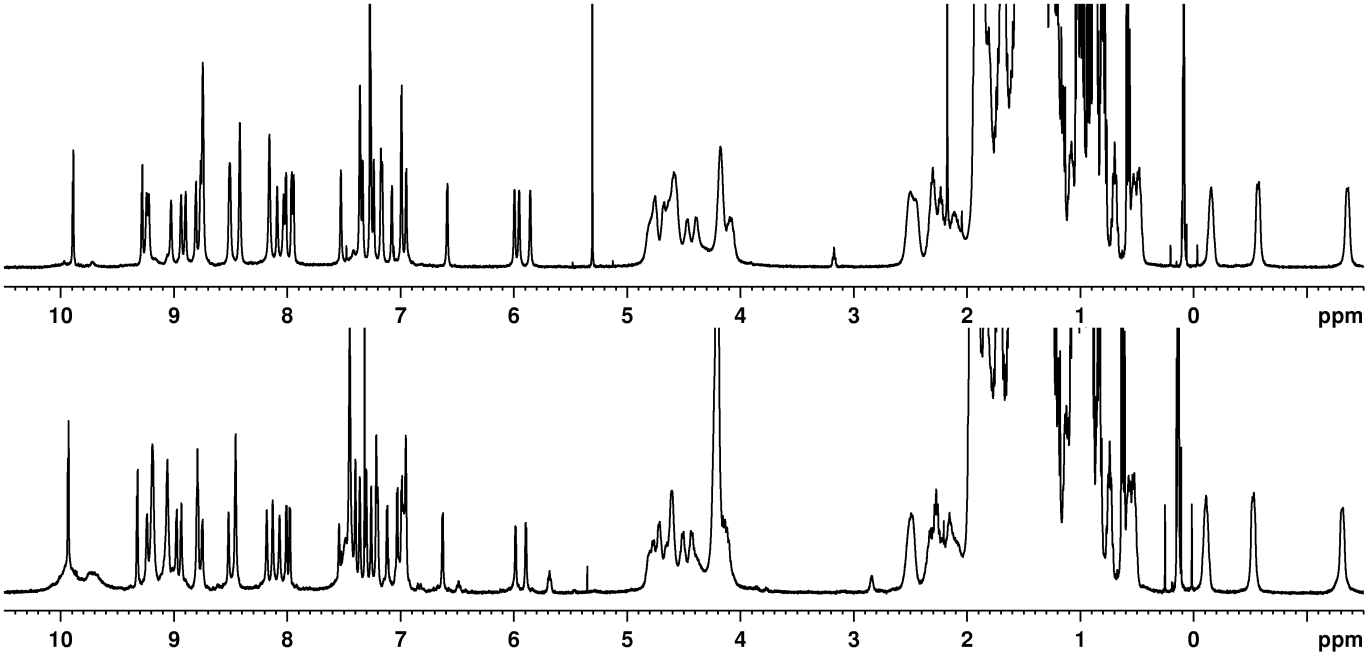

Figure S8. Overlay of the ^{1}H NMR spectra of the tetramer aggregate (upper trace) and the trimer aggregate (lower trace), $CDCl_3$, 500 MHz, 298 K. The spectra demonstrate remarkable similarity and suggest similar triple-stranded aggregates are formed for both oligomers, with specific resonances missing in the trimer spectrum due to its reduced length; this is most apparent for the shielded β-pyrrole resonances at 5.8–6.1 ppm, for example.

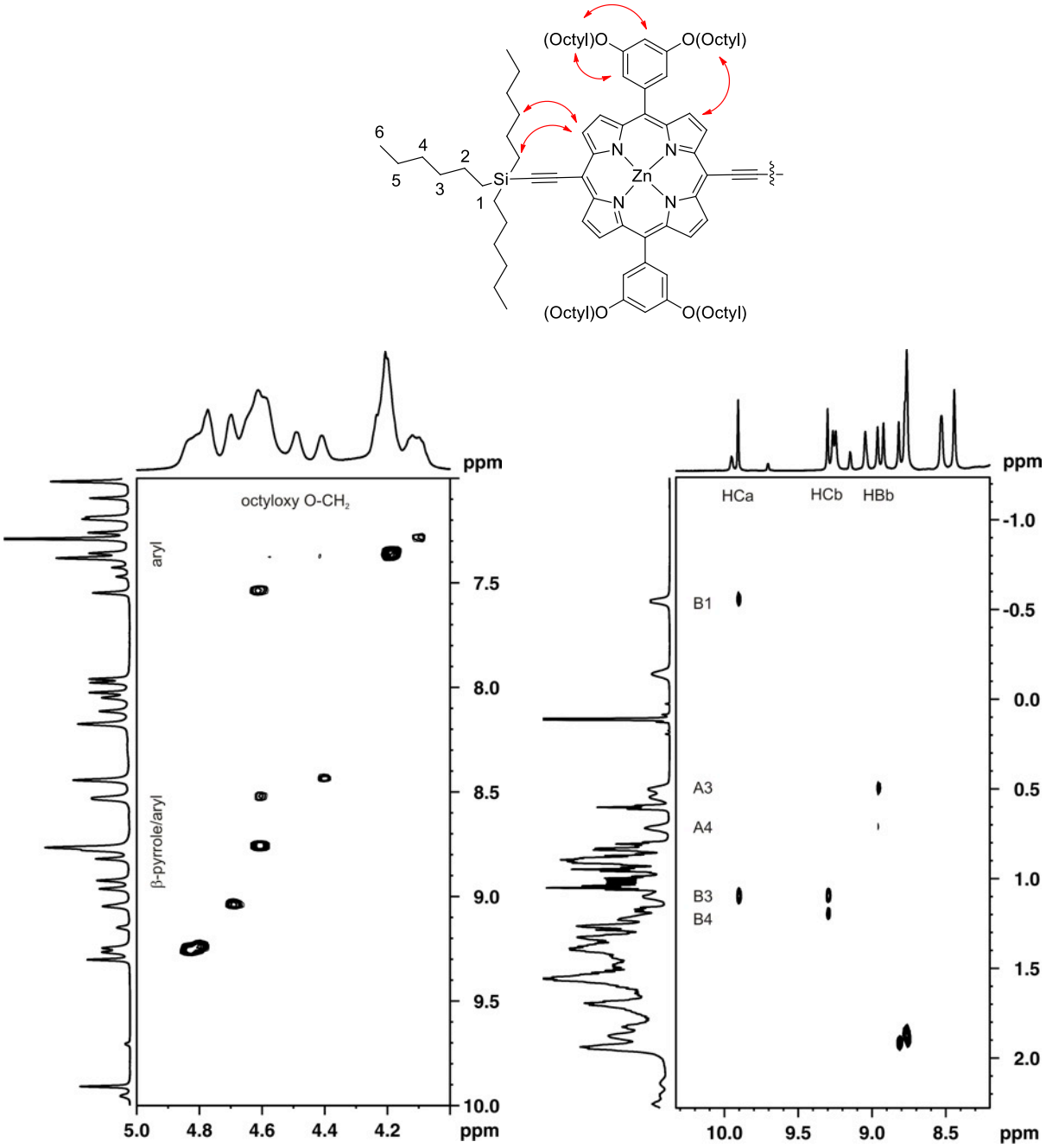

Figure S9. Selected regions of a ^{1}H COSY spectrum demonstrating unexpected correlations between (a) β-pyrrole or aryl and octyloxy sidechain protons and (b) β-pyrrole and hexylsilyl sidechain protons for the tetramer aggregate ($CDCl_3$, 700 MHz, 298 K). These correlations originate from ^{1}H-^{1}H dipolar couplings, not classical scalar (*J*) couplings, as illustrated schematically in the porphyrin structure (note: some correlations actually arise between protons in neighboring strands), and all show matching correlations in 700 MHz NOESY spectra.

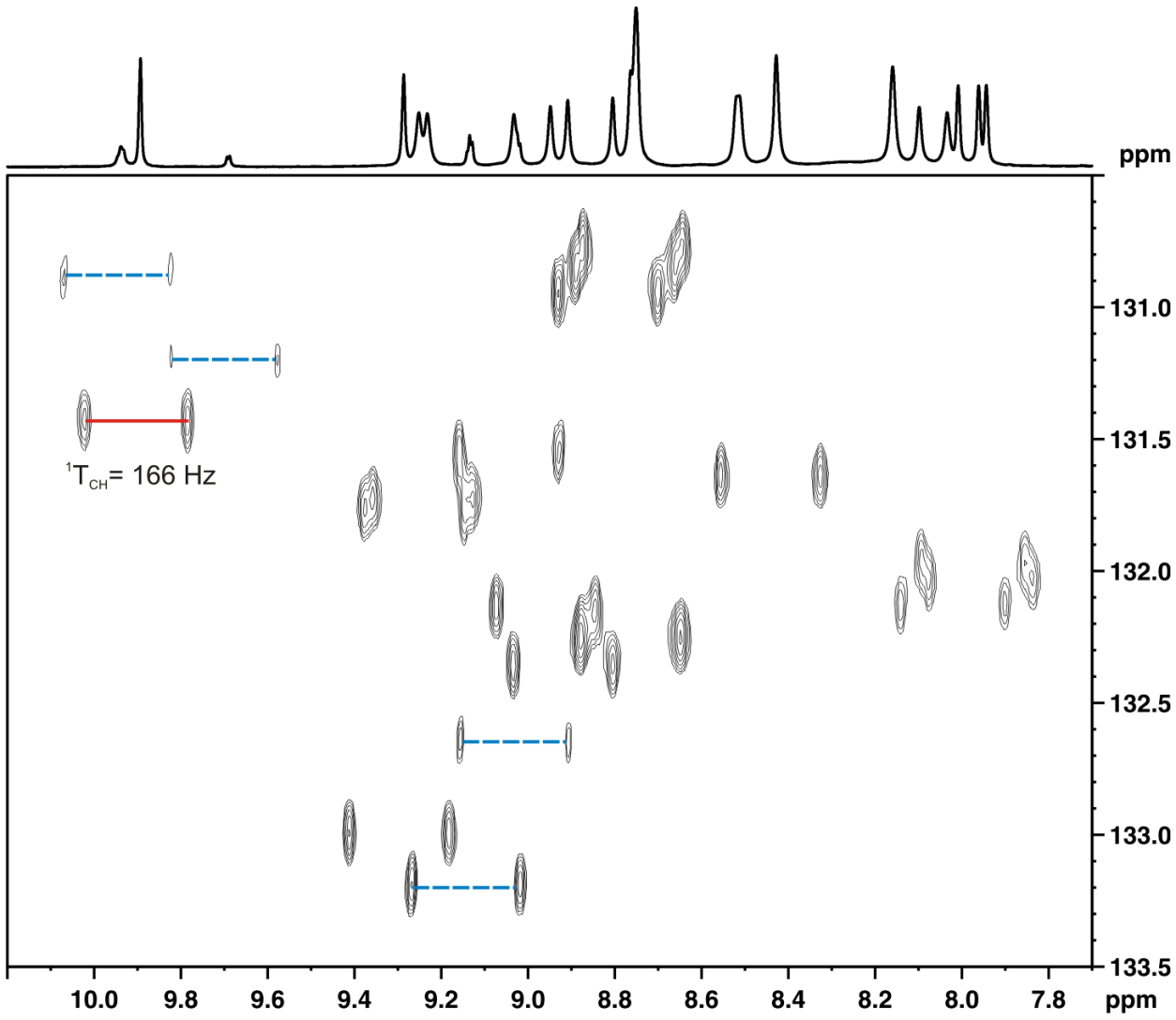

Figure S10. Region of the ^{13}C-coupled ^{1}H-^{13}C HSQC spectrum of the tetramer aggregate showing the ^{13}C satellite doublet structure from which the total splitting ($^1T_{CH}$= $^1J_{CH}$ + $^1D_{CH}$; red line) was measured in the ^{1}H dimension ($CDCl_3$, 700 MHz, 298 K). The doublet structures indicated with dashed blue lines arise from the disaggregated tetramer **1·(py)$_4$** (present due to residual d_5-pyridine). Exchange between the aggregate and the disaggregated tetramer-pyridine adduct is slow under these conditions.

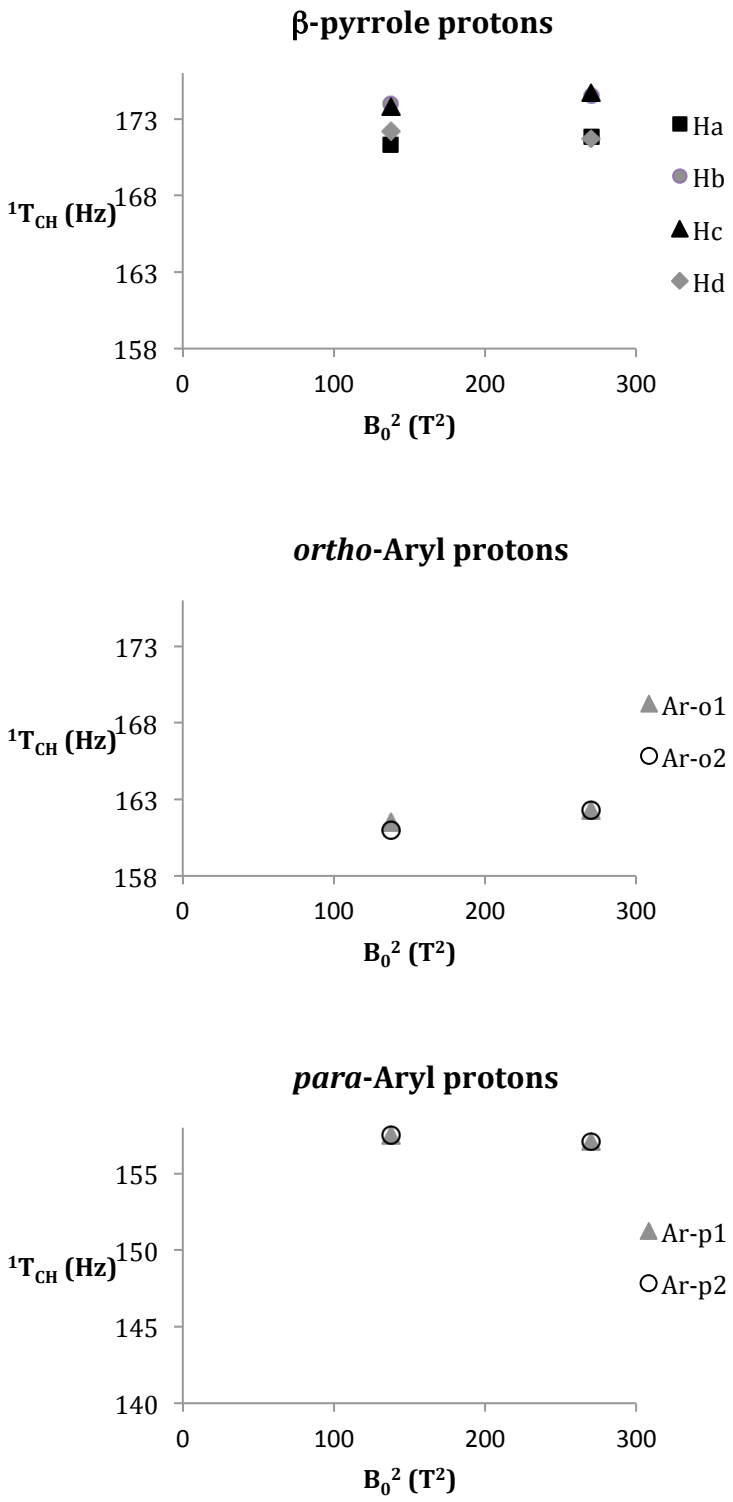

Figure S11. Observed total ^{1}H-^{13}C doublet splitting ($^1T_{CH}$) plotted against the square of the magnetic field of the disaggregated tetramer **1**·**(py)$_4$**. Plot ranges match those in Figure 6 of the main text to aid comparison.

Details of SAXS analysis

Both the disaggregated porphyrin tetramer in toluene/1% pyridine (3.0 mg/mL) and the tetramer aggregate in toluene (3.0 mg/mL) were analyzed by solution-phase small-angle X-ray scattering (SAXS) using synchrotron radiation at the Diamond Light Source at beamline I22. Linearity of the scattering data in Guinier plots of both samples confirmed their monodispersity. The experimental scattering data of the disaggregated tetramer is perfectly reproduced by a simulated scattering profile created from a molecular model (optimized with the semiempirical AM1 method) using the program Crysol. The Guinier analysis of the raw scattering data at small angles gives a radius of gyration R_g (the root mean square distance of the electrons from the center of gravity of the molecule) of 16.0 Å which is in excellent agreement with the value determined from the model (16.3 Å). The pair-distribution function (PDF) was obtained by indirect Fourier transformation of experimental and simulated scattering data (using the program Gnom). Pair distribution functions of porphyrin oligomers and assemblies are highly informative: the local point symmetry of the porphyrin macrocycle and the presence of a high atomic number metal in its center often result in distinct peaks for certain metal-metal distances in the PDF. The three peaks in the PDF of disaggregated tetramer correspond to the three Zn-Zn distances in the molecule. The good agreement between experimental and simulated PDF confirms the validity of the calculated model. The longest Zn-Zn distance in the experimental PDF occurs at slightly shorter length than in the model due to the intrinsic flexibility of the tetramer chain in solution, resulting in an averaged shorter distance.

In order to elucidate the structure of the porphyrin tetramer aggregate, the experimental data were compared to three different aggregates: a dimeric, a trimeric and a tetrameric aggregate. Models for each structure were calculated using the MM+ force field in HyperChem and the corresponding scattering profiles simulated using Crysol. PDFs for all simulated and the experimental scattering data were calculated with Gnom and R_g values determined from Guinier fits. The experimental scattering and PDF show poor agreement with the data simulated based on the tetrameric aggregate and the R_g value from this model (17.4 Å) is significantly higher than the experimentally determined value (15.1 Å). The dimeric and the trimeric model give both satisfactory fits to the experimental data and it is difficult to judge which of the two models reproduces the experimental data better. The PDF based on the trimeric model appears to fit the experimental PDF best, but the R_g and the experimental scattering give better agreement with a dimeric model. Experimental SAXS data allows the molecular weight determination of unknown assemblies by comparison of the extrapolated scattering intensity at zero angle $I(0)$ with a calibration standard of known concentration. The samples of the aggregated and the disaggregated tetramer were measured at nearly identical concentrations (3.0 mg/mL). This means the relative molecular weights can be directly obtained from the ratio of the values for $I(0)$. The ratio of 2.05/1 between aggregated and disaggregated tetramer indicates the formation of a dimeric aggregate in toluene.

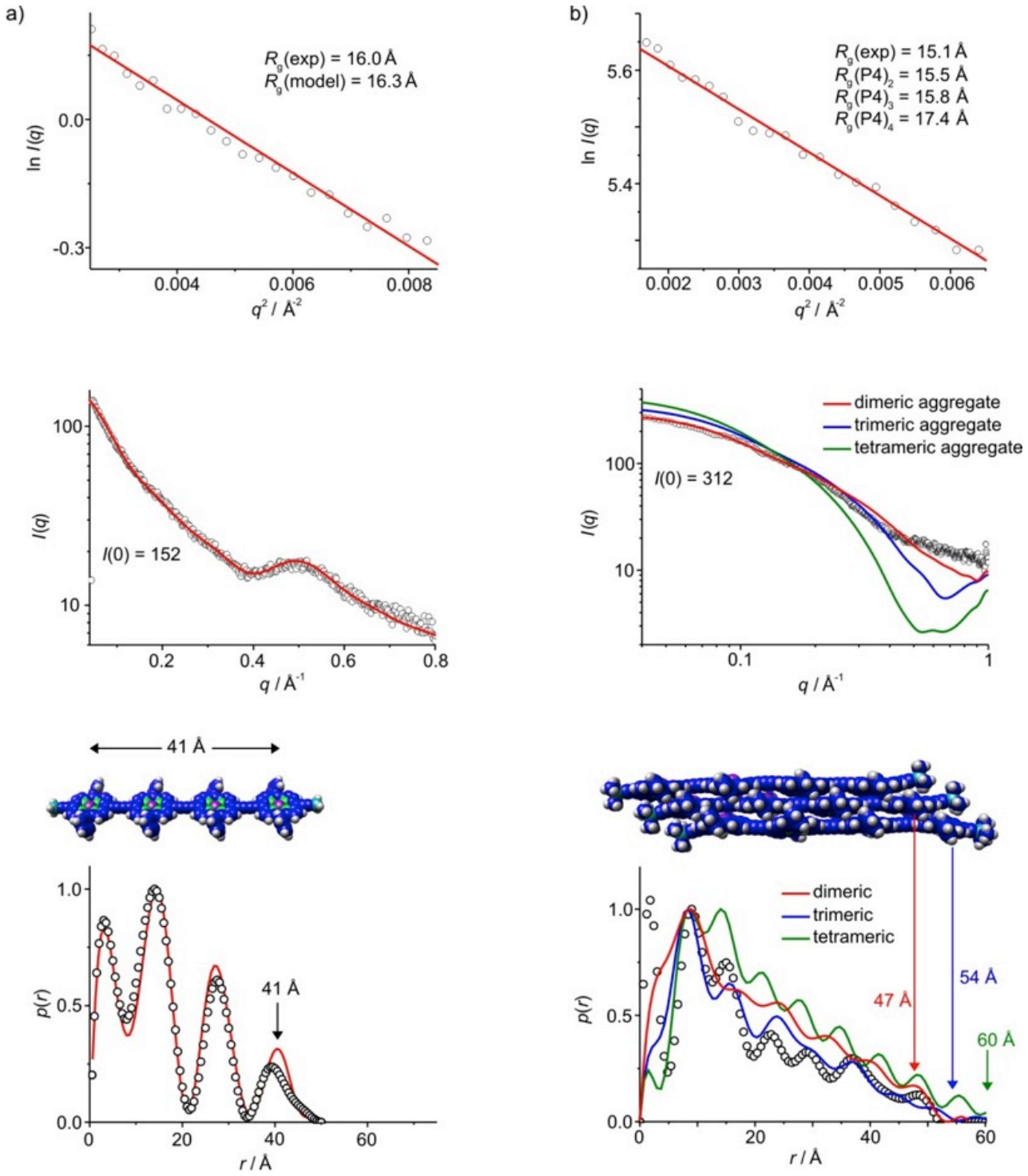

Figure S12. (a) SAXS data for **1** in toluene containing 1% pyridine, and (b) for the solution in pure toluene. Concentration: 3.0 mg/L. Experimental scattering (black circles) and Guinier fit (red line). MM+ minimized structure of the tetramer aggregate used for data fitting. Pair distribution function of the experimental scattering data (black circles) and the model (red line).