Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2017

Ship in a bottle: confinement-promoted self-assembly

E. Lopez-Fontal, A. Grochmal, T. Foran, L. Milanesi, S. Tomas

Supplementary Information

Materials and Methods

General. Chemicals and solvents where obtained from commercial sources and used without further purification. Cobalt metalloporphyrin **C** was obtained by metalation of the corresponding free base porphyrin.[29] UV data was recorded using a Cary 3 UV spectrophotometer. The buffer used in this work was sodium phosphate 100 mM, pH 7.2, prepared by dissolving NaH_2PO_4 in water and adjusting the pH with a concentrated solution (e.g., 5M) of NaOH.

Lipid vesicle preparation. The appropriate amount of a stock solution of egg-yolk phosphatidylcholine (EYPC) in ethanol was evaporated under reduced pressure. The resulting lipid film was suspended, by vortexing the sample for a minimum of 2 minutes in 1.2 mL of buffer containing the appropriate concentration of cobalt porphyrin **C**. The suspension was then extruded through polycarbonate filters (pore size 100 nm) a minimum of 21 times. The suspension was then subjected to 6 cycles of freeze-thawing, after which it was extruded another 21 times through the same filter. For experiments of titration of non-confined **C** with bipiryridine **B** there was no **C** on the buffer used to generate the vesicles and the concentration of EYPC on the vesicle suspension was 10 mM. This suspension was then used to generate the corresponding samples containing 0.5 mM EYPC (see UV experiments below). For experiments with confined **C**, the concentration of C on the buffer use to generate the vesicle suspension was 1.0 mM, with the concentration of EYPC being also 10 mM. Once the vesicle suspension was generated, non-confined **C** was removed by size exclusion chromatography, using Sepharose 4B as the stationary phase. The fractions containing confined **C** where then used for the preparation of the corresponding dilution or titration experiments. The concentration of EYPC in the samples after SEC was estimated by comparing the volume of the fractions containing vesicles with the volume applied to the column. The concentration of **C** in relation to the bulk solution was estimated by UV, using the known extinction coefficient of the UV spectrum of **C** at the maximum of the Soret band (320,000 M^{-1} cm^{-1} at 423 nm)[29]

UV titration experiments. In a typical experiment, 12 samples, all containing a constant concentration of **C**, and 11 of them with increasing concentration of the appropriate ligand (with a concentration increase of 1.5 fold between samples) where prepared in buffer. To evaluate the binding of 3-pyridinesulfonic acid, **Ps**, the concentration of **C** was 2 μM and that of ligand ranged from 4 to 4000 μM (Supplementary Figure S1, Supplementary Table

S1). To evaluate the binding of pyridine, **P**, the concentration of **C** was 2 μM and that of ligand ranged from 0.35 to 20 μM (Figure 2, Supplementary Figures S4 and S5, Table 1). For the evaluation of the binding of non-confined **C** the concentration of **C** was either 1.8, 18 or 1000 μM, that of **B** ranged from 0.43 to 3700 μM and were carried out in the presence of 0.5 mM of EYPC (Figure 3, Figure 4a, Supplementary Figure S6 and S7, Table 1). All samples where incubated overnight before recording the UV spectrum. For the evaluation of binding of confined **C** to **B**, vesicle suspensions containing confined **C** where used to generate up to 14 samples with concentration of **B** ranging from 0.125 to 2000 μM. The samples where let to incubate overnight, after which the UV was recorded. Further analysis of the UV of these samples at later times (2 days and 1 week) revealed no further change in the spectra, showing that in 10-12 hours the equilibrium had been reached (Figure 4, Supplementary Figure S8, Supplementary Table S2, Table 1).

UV dilution experiments. In a typical experiment a sample containing confined **C**, with local concentration in the lipid cavity of 1mM and apparent concentration 4 μM, and 3 mM of **B** was used to generate 4 samples by serial dilution with buffer, halving the concentration at each dilution. The samples where let to equilibrate overnight and the UV spectrum recorded after this time (Figure 5a, Supplementary Figure S9, Supplementary Table 2).

Vesicle bursting experiment. A sample of confined **C** with a bulk concentration of 5 μM, and a local concentration of 1.0 mM, containing EYPC 0.5 mM was added **B** up to 100 μM and let equilibrate overnight. After this period, it was diluted 2 fold with a solution of triton X 10% in volume and the UV spectra were recorded at fixed time intervals (Figure 5b).

Cryo-TEM experiments.

Samples containing vesicle-confined **C** were prepared as described above. The local concentration of C was 1 mM and the apparent concentration 1.8 μM in all cases. The concentration of EYPC was 500 μM. The concentration of **B** ranged from 0 to 2 mM. Samples (3.5 μL) were loaded on freshly discharged (60 s) holey carbon Lacey grids (S166-3; Agar Scientific). The grids were blotted and plunged into liquid nitrogen cooled liquid ethane to embed the samples in vitreous ice. Images were collected using minimal electron dose at a nominal magnification of 26000x in a Tecnai 12 microscope (FEI, Eindhoven, NL) with a tungsten filament operating at 120 kv. Images were recorded with a Gatan 1K Ultrascan camera (Gatan, USA) between 1.2-2.0 μm underfocus (Supplementary Fig. S7 and Fig. 8).

UV Data processing. All the UV spectra where exported as ASCII data for processing with Microsoft Excel. In order to remove random baseline variations and the lipid vesicle scattering, a double derivative was applied to the spectral data. The first derivative was carried out graphically, by calculating the average of 5 contiguous data and subtracting the average of 5 contiguous data 5 nanometers apart (for example, derivative = average (A_{365}, A_{366}, A_{367}, A_{368}, A_{369}) – average (A_{360}, A_{361}, A_{362}, A_{363}, A_{364}). The second derivative of the absorbance, *A''*, was calculated by repeating the process over the results from the first

derivative. This procedure smooths out noise and ensures that only variations in the spectral bands relating to assembly processes are reflected in the data.

^{1}H-NMR of the CP and CP_2 complexes. Samples of **C** with increasing amounts of pyridine **P** where prepared in buffer containing 10% of D_2O. The concentration of **C** was 0.75 M while the concentration of **P** ranged from 0 to 1.5 mM. The samples were let to equilibrate for 2 hours and the spectrum recorded on a Bruker AMX 600 NMR Spectrometer (Supplementary Figure S3).

Simulations of polymerization vs. concentration of C and B. The data for the graphs shown in Fig. 7 was generated using equations (S31) and (S47) respectively as implemented in the program Micromath Scientist 3.0, using the parameters in Table 1. The data was imported into Excel to construct the plots.

Testing the membrane permeability of C

3-pyridinesulfonic acid (**Ps**) is an anionic molecule at pH 7, and is therefore expected to be little or no permeable to the lipid membrane. This feature make **Ps** ideally suited to test the membrane permeability of **C**. To carry out the permeability experiments it was necessary to establish beforehand to what extent **Ps** binds to non-confined **C** in the presence of lipid vesicles. UV titration experiments show changes in the UV spectrum that are consistent with the formation of complexes **CPs** and **CPs_2** (Fig. S1 A).

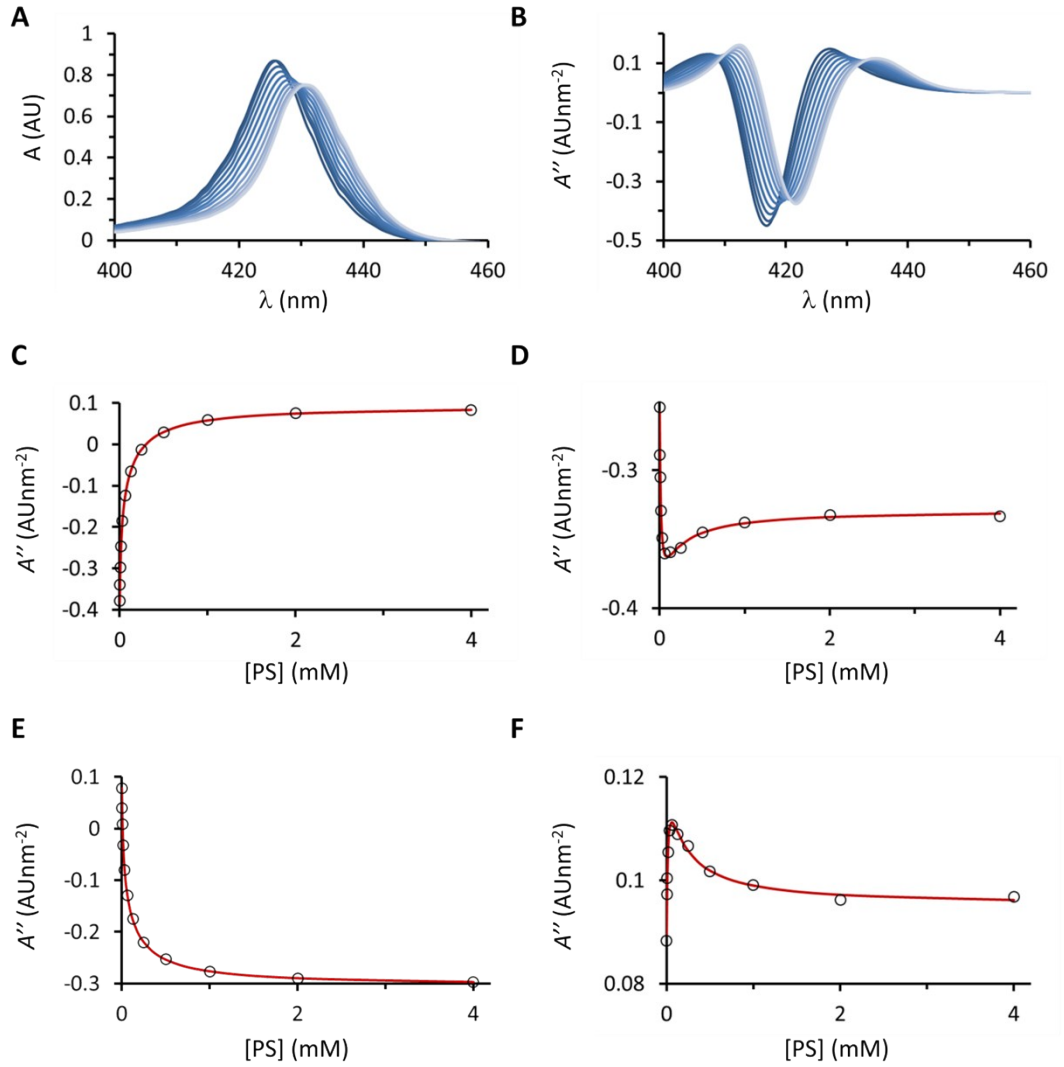

Figure S1. A. Changes in the Soret band region of the UV spectrum of **C** in buffer upon addition of increasing amounts of 3-pyridinesulfonic acid (**Ps**). The concentration of **C** was 2.0 μM and that of EYPC vesicles 500 μM. B. Changes in the second derivative of the spectra displayed in panel A. C. Changes in the intensity of the second derivative of the absorbance at 415 nm. D. Idem at 420 nm. E. Idem at 424 nm. F. Idem at 432 nm.

The distribution of species in equilibrium depends on the binding constants K_1 and K_2 that were determined by fitting the UV data to the appropriate system of equations. The system includes the equations of the binding for the binding constant (S1 and S2), the mass balances (S3 and S4) and the relationship between the second derivative of the absorbance, *A''*, to the concentration of the colored species present (S5):

$$K_1 = \frac{[CPs]}{[C][Ps]} \qquad \text{(S1)}$$

$$K_2 = \frac{[CPs_2]}{[CPs][Ps]} \qquad \text{(S2)}$$

$$[C]_0 = [C] + [CPs] + [CPs_2] \quad \text{(S3)}$$

$$[Ps]_0 = [Ps] + [CPs] + 2[CPs_2] \quad \text{(S4)}$$

$$A'' = E_C[C] + E_{CPs}[CPs] + E_{CPs2}[CPs_2] \quad \text{(S5)}$$

E_C, E_{CP} and E_{CP2} are the proportionality factors between the second derivative of the absorbance and the concentration of the species. To calculate the binding constants we fitted to the model changes at the two wavelength where the change in intensity is the greatest, together with two wavelengths where the presence of the two binding events where more clear (i.e., those wavelengths at which the trend of variation in signal changes direction as the formation of the dominant complexes switches from 1:1 to 1:2). For this purpose we used the program Scientist 3.0, which allows entering the model as a system of implicit equations (Supplementary Fig. S1, Table S1).

In order to establish the lipid membrane permeation properties of **C** samples of confined **C** where treated with **Ps** and the changes in the UV spectrum recorded. From the values of binding constants it was estimated that in the presence of 4 mM **Ps** more than 95% **C** should bind to the ligand. Addition of **Ps** up to this concentration leads to a rapid change of the UV spectrum that is completed within the first two minutes (Supplementary Fig S2). This change is consistent with the binding of **Ps** to a relatively minor fraction of non-confined **C**. Thereafter, a much slower change follows that is consistent with either the binding of **C** to **Ps** upon permeation of **Ps** into the cavity, **C** out of the cavity, or both. Changes in the second derivative of the spectra where fitted to a first order kinetic process for the second stage. The fitting of the data allow us to calculate the percentage change of signal on the first and second stages (Supplementary Fig. S2). In the first, fast, stage the signal experiences a 20 % of the total change. We can therefore conclude that around 80% of **C** is confined in the lipid vesicle at the start of the experiment. In order to establish whether it is **C** or **Ps** the molecule that permeates during the second stage a vesicle sample containing confined **C** was let to equilibrate for 24 hours and then treated with **Ps**. Analysis of the data reveals that both the first, nearly instantaneous, change of the signal and the slow change thereafter are the same than for a sample made out of vesicles freshly prepared. This result shows that, in a 24 hours period, the percentage of non-confined **C** in the sample has not increased. Therefore **C** does not permeate out of the vesicle during this period to a measurable extent, and that the slow change phase is result of the permeation of the ligand **Ps** into the lipid vesicle.

Form these results we can safely assume that, in our experiments, **C** does not leak out when confined. These experiments also show that a minor fraction of unconfined **C** is always present in samples containing confined **C**. In the conditions of our experiments, however, non-confined **C** does not contribute to the UV band attributed to the double stranded polymer and can therefore be disregarded when fitting the experimental data to our assembly model.

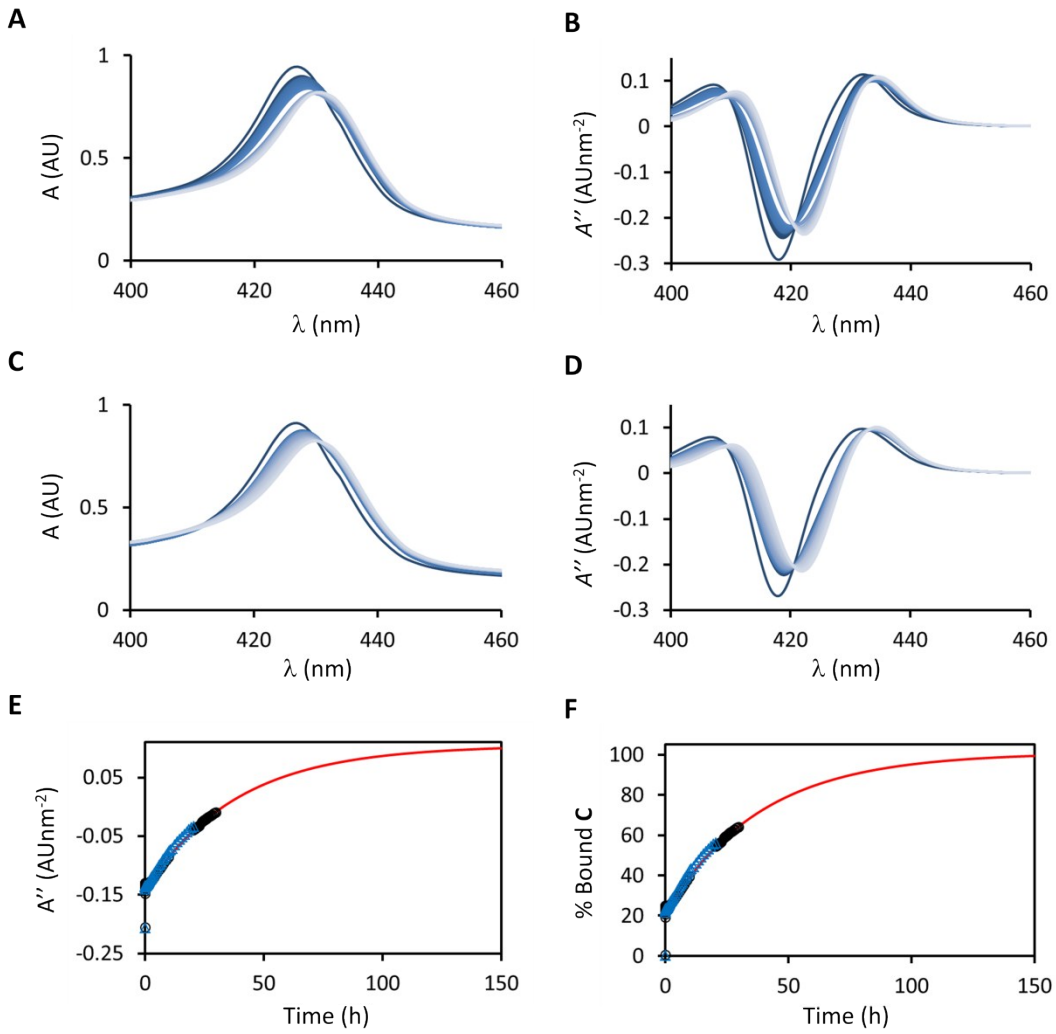

Figure S2. A. Changes in the Soret band region of the UV spectrum of lipid vesicle confined **C** upon addition of 3-pyridinesulfonic acid (**Ps**) in buffer. The first spectrum was recorded prior to the addition of **Ps**, and the intensity was corrected for the factor of dilution introduced by the addition of **Ps** stock. The concentration of **C** was 2.4 μM, the local concentration of **C** (i.e., in the lipid vesicle cavity) was 1 mM, that of EYPC 500 μM and that of **Ps** 4.0 mM. B. Changes in the second derivative of the absorbance for the data shown in panel A. C. Same than in panel A, for a sample that has been let to equilibrate for 24 hours at room temperature prior to the addition of **Ps**. D. Changes in the second derivative of the absorbance for the data shown in panel C. E. Changes in the second derivative of the absorbance at 415 nm for the spectra shown in panel B (black circles) and for the spectra shown in panel D (blue triangles). The red line is the fitting to a first order kinetics of the data excluding the stage. F. Plot of the percentage of complex formed obtained from the data shown in panel E.

Table S1. Binding constants of C with Ps

Ligand	K_1	K_2
Ps	$7.4 \times 10^4 \pm 3.5 \times 10^3$	$5.8 \times 10^3 \pm 2.8 \times 10^3$

The units are M^{-1} and are the average of a minimum of 3 experiments. The error reported is twice the standard deviation of the mean.

Local concentration of C

The local concentration of **C** in the cavity can be assumed to be equal to the concentration of **C** in the buffer used to prepare the sample. The local concentration could be independently estimated by measuring the apparent concentration of **C** in vesicle samples used for titration and dilution experiments and by determining the confined volume. To determine the confined volume the concentration of lipid is estimated by recording the dilution factor of the lipid sample before SEC, of known concentration. For example, for the titration experiment featured in Fig. 4c, we have that vesicle sample 9 mM in EYPC, generated by suspending the lipids in buffer containing 1 mM **C**, was applied to the SEC column. 3 mL where collected containing vesicles with confined **C**. and were diluted down 6-fold to generate the samples used in the titration experiment. The concentration of lipids in the sample is therefore 500 μM in EYPC. The apparent concentration of **C** was determined from the first point of the titration (i.e, 1.8 μM using the ε_{420} = 320000 $M^{-1}cm^{-1}$). Knowing the concentration of EYPC and the average area per lipid molecule (60 to 80 A^2) and the average size of the lipid vesicles (100 nm diameter), we calculated the fraction of volume in the cavity in relation to the bulk (i.e., 0.2 %). Since **C** is largely inside the cavity, the local concentration can be determined by dividing the apparent bulk concentration by the fraction of solvent in the cavity. Using this procedure $[\mathbf{C}]_{i0}$ was calculated to be between 0.9 and 1.2 mM, consistent with the initial assumption of 1 mM.

Binding of pyridine P to non-confined and confined C

The identity of the complexes **CP** and $\mathbf{CP}_2$ was corroborated by means of ^{1}H NMR spectroscopy (Supplementary Fig. S3).

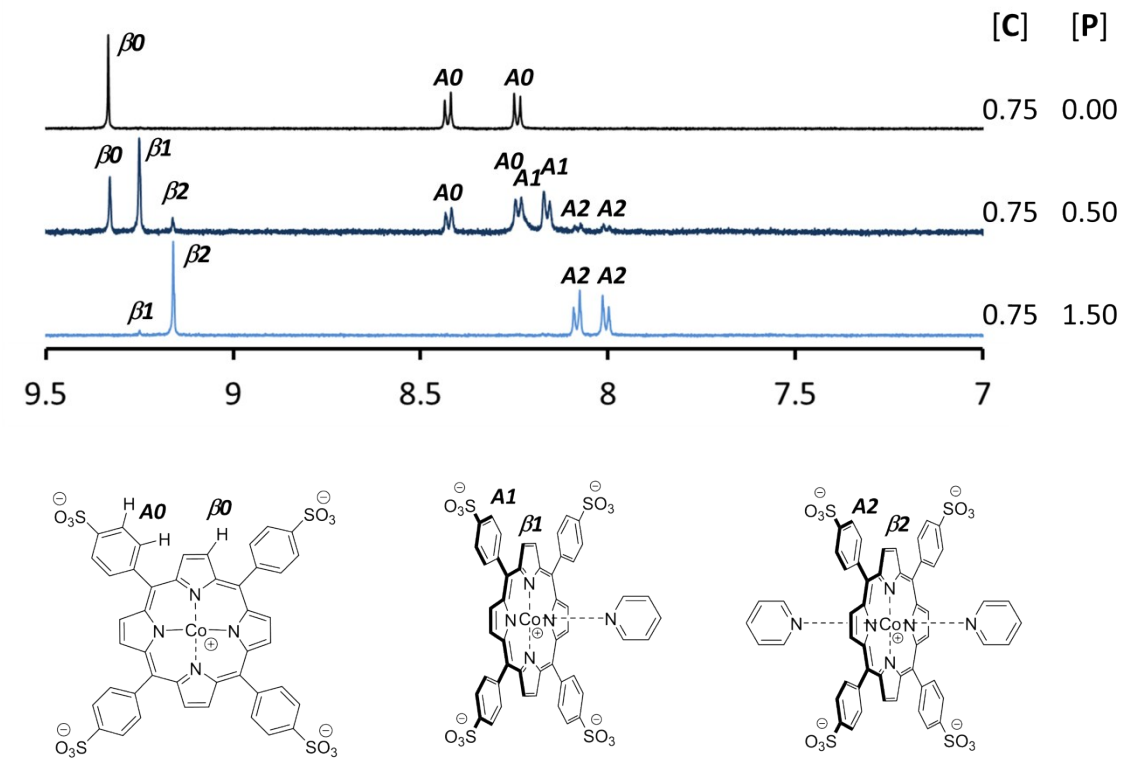

Figure S3. Changes in the aromatic region of the ^{1}H-NMR spectrum of **C** upon addition of increasing amounts of **P**. The peak assignment to the different hydrogen atoms is shown. The concentration of **C** and **P** is shown to the right and is quoted in mM.

The effect of confinement on the binding affinity of **C'**s metal center for the ligand was estimated by carrying out UV titration experiments of confined and non-confined **C** with pyridine **P**. The UV data was fit to the model described for **Ps** (Supplementary Fig. S4 and S5 and Table 1)

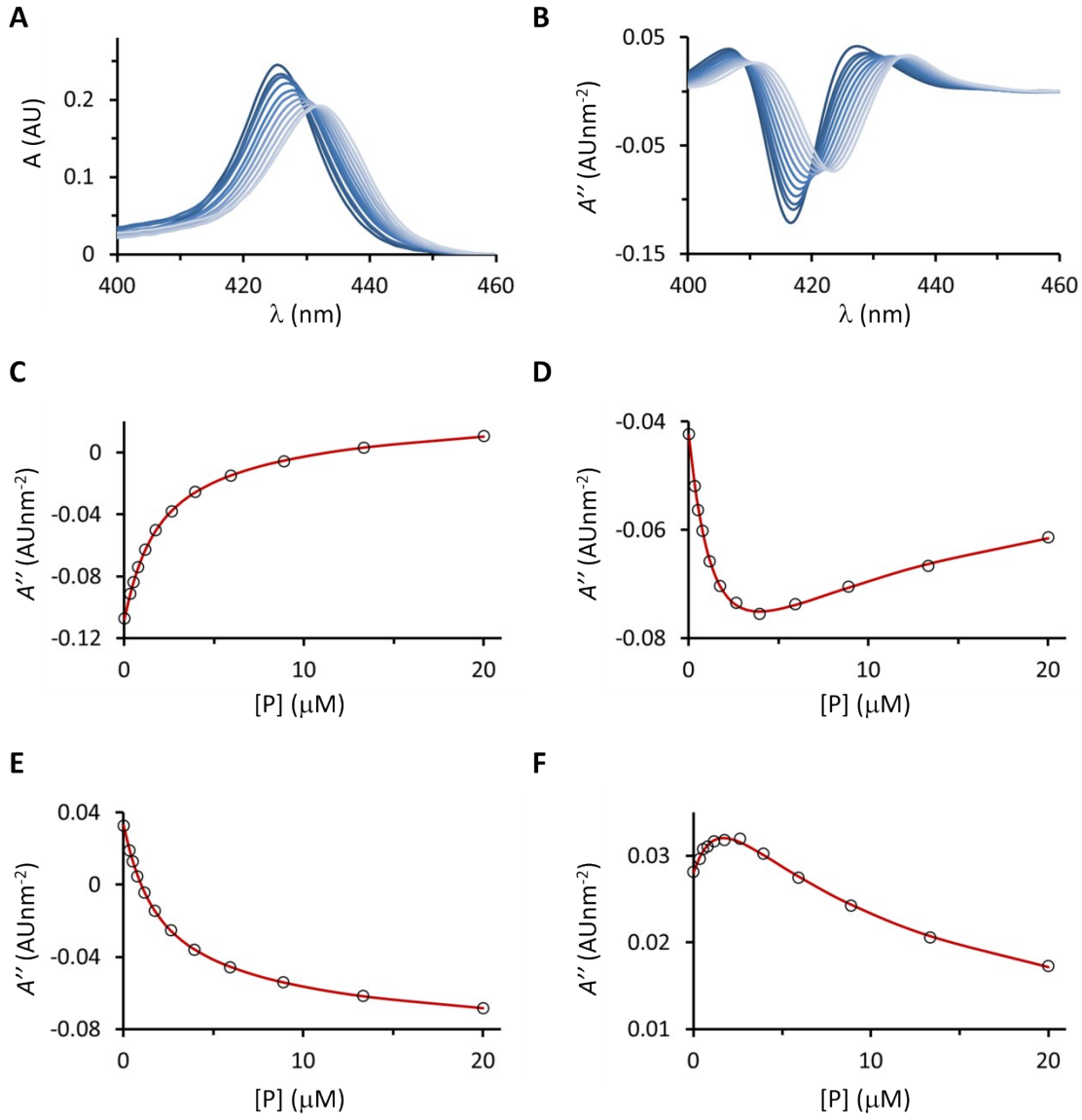

Figure S4. A. Changes in the Soret band region of the UV spectrum of **C** in buffer upon addition of increasing amounts of pyridine (**P**). The concentration of **C** was 1.0 μM. B. Changes in the second derivative of the spectra displayed in panel A. C. Changes in the intensity of the second derivative of the absorbance at 415 nm. D. Idem at 421 nm. E. Idem at 425 nm. F. Idem at 432 nm.

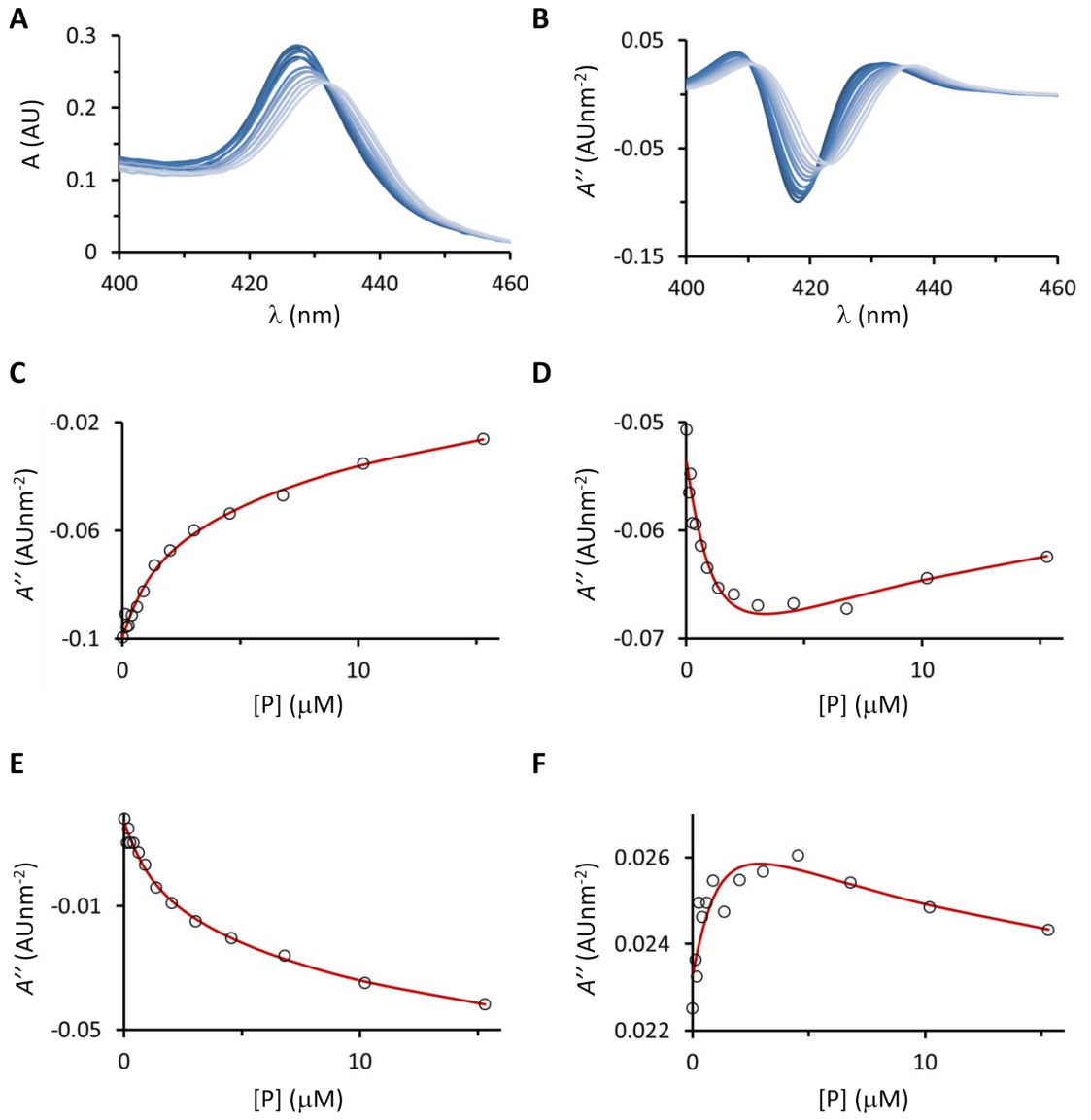

Figure S5. A. Changes in the Soret band region of the UV spectrum of lipid vesicle confined **C** in buffer upon addition of increasing amounts of pyridine (**P**). The concentration of **C** was 0.65 μM, the local concentration (i.e., within the cavity of the lipid vesicle) was 1 mM and that of EYPC was 500 μM. B. Changes in the second derivative of the spectra displayed in panel A. C. Changes in the intensity of the second derivative of the absorbance at 418 nm. D. Idem at 422 nm. E. Idem at 427 nm. F. Idem at 435 nm.

Assembly model for non-confined C and B and UV data fitting

As described elsewhere[29,30] we assume that, within long oligomers, the number of **C** and **B** building blocks are the same. We can therefore describe the formation of single stranded oligomers as the isodesmic assembly of building blocks **CB**, with oligomerization constant K_o, i.e.:

$$K_o = \frac{[(CB)_n]}{[(CB)_{n-1}][CB]} \quad \text{(S6)}$$

The formation of the double stranded polymer with n repeats, **D**, can be written as a function of the lateral association constant per unit repeat, K_l, as follows:

$$K_l^{\,n} EM_D^{\,n-1} = \frac{[D]}{[(CB)_n]^2} \quad \text{(S7)}$$

where EM_D is the effective molarity for the assembly of **D**, a measure of the local concentration of binding sites within the polymer **D**. For simplicity of notation, we have assumed throughout that EM_D is 1 M in our system.[29]

The formation of the building block **CB** is depends on the binding constant K_1

$$K_1 = \frac{[CB]}{[C][B]} \quad \text{(S8)}$$

On the other hand, the presence of an excess of either **B** or **C** leads to the formation of complexes $\mathbf{CB_2}$ and $\mathbf{C_2B}$, with binding constants K_2 and K_3 respectively, i.e.:

$$K_2 = \frac{[CB_2]}{[CB][B]} \quad \text{(S9)}$$

$$K_3 = \frac{[C_2B]}{[CB][C]} \quad \text{(S10)}$$

The mass balances of **C** and **B** can be written as:

$$[C]_0 = [C] + [CB_2] + 2[C_2B] + \sum_{n=1}^{\infty} n[(CB)_n] + \sum_{n=1}^{\infty} 2n[D] \quad \text{(S11)}$$

$$[B]_0 = [B] + 2[CB_2] + [C_2B] + \sum_{n=1}^{\infty} n[(CB)_n] + \sum_{n=1}^{\infty} 2n[D] \quad \text{(S12)}$$

which can be written as a function of the binding constants and the equilibrium concentration of **C** and **B**:

$$[C]_0 = [C] + K_1K_2[C][B]^2 + 2K_1K_3[C]^2[B] + \frac{K_1[C][B]}{\left(1 - K_oK_1[C][B]\right)^2} + \frac{}{(1 - K}$$

(S13)

$$[B]_0 = [B] + 2K_1K_2[C][B]^2 + K_1K_3[C]^2[B] + \frac{K_1[C][B]}{(1 - K_oK_1[C][B])^2} + \frac{}{(1 - K}$$

(S14)

Solving the system of implicit equations (S13)-(S14) is not an easy task. It is nonetheless possible to reduce the model to a single implicit equation if the appropriate assumptions are made. For example, given the relatively large value of K_1 (Table 1), for concentrations of **C** above 10 μM, it can be assumed that, of **C** and **B**, the component that is in the lowest concentration is fully bound to yield the complex **CB**, which is the monomer repeat in the polymerization process. The sum of all **CB** present we call $[\mathbf{CB}]_0$. We therefore have that

if $[C]_0 > [B]_0$ then $[CB]_0 = [B]_0$ (S15)

If $[C]_0 = [B]_0$ then $[CB]_0 = [B]_0 = [C]_0$ (S16)

If $[C]_0 < [B]_0$ then $[CB]_0 = [C]_0$ (S17)

We call $[\mathbf{B}]_x$ the excess concentration of **B** , which is equal to zero when the total concentration of **C** is larger or equal than that of **B**, but that when **B** is larger than **C**. Therefore:

$$[B]_x = [B]_o - [CB]_0 \quad \text{(S18)}$$

Likewise, $[\mathbf{C}]_x$ is the excess concentration of **C** , which is equal to zero when the total concentration of **B** is larger or equal than that of **C**, but that when **C** is larger than **B**, is:

$$[C]_x = [C]_o - [CB]_0 \quad \text{(S19)}$$

The total concentration of monomer $[\mathbf{CB}]_0$ is also the sum of all the species that contain **CB**, that is:

$$[CB]_0 = [CB_2] + [C_2B] + \sum_{n=1}^{\infty} n[(CB)_n] + \sum_{n=1}^{\infty} 2n[D] \quad \text{(S20)}$$

Using the corresponding mass balances:

$$[B]_x = [B] + [CB_2] \quad \text{(S21)}$$

$$[C]_x = [C] + [C_2B] \quad \text{(S22)}$$

and equations (S3) and (S4), the concentrations of species **CB_2** and **C_2B** can be written as a function of the corresponding constant, the concentration of free monomer **CB** and the total excess **B** or **C**, $[\mathbf{B}]_x$ or $[\mathbf{C}]_x$:

$$[CB_2] = \frac{K_2[CB][B]_x}{(1 + K_2[CB])} \quad \text{(S23)}$$

$$[C_2B] = \frac{K_3[CB][C]_x}{(1 + K_3[CB])} \quad \text{(S24)}$$

On the other hand re-arranging (S6) we have that:

$$[(CB)_n] = (K_o[CB])^{n-1} \quad \text{(S25)}$$

And combining and combining with (S7) we have:

$$[D] = K_l(K_l EM_D K_o^2 [CB]^2)^{n-1} \quad \text{(S26)}$$

Combining equations (S20) with equations (S23)-(S26) we have:

$$[CB]_0 = \frac{K_2[CB][B]_x}{(1 + K_2[CB])} + \frac{K_3[CB][C]_x}{(1 + K_3[CB])} + [CB]\sum_{n=1}^{\infty} n(K_o[CB])^{n-1} + 2K_l[CB$$

(S27)

and applying in equation (S27) the relevant convergent series[29] we have:

$$[CB]_0 = \frac{K_2[CB][B]_x}{(1 + K_2[CB])} + \frac{K_3[CB][C]_x}{(1 + K_3[CB])} + \frac{[CB]}{(1 - K_o[CB])^2} + \frac{2K_l[CB]^2}{(1 - K_l EM_D K_o^2 [CB]^2)^2} \quad \text{(S28)}$$

Equation (S28) allows calculating the concentration of building block **CB** in equilibrium. This equation however is only strictly applicable in those conditions in which it can be assumed that **CB** forms quantitatively. Equation (S28) cannot therefore be applied for cases where the concentration of **C** is very low (i.e., in the sub-micromolar range). It is possible to obtain a closer reflection of the binding scenario by replacing [**CB**] by the expression of the binding constant that leads to this complex, that is:

$$[CB] = K_1[C][B] \quad \text{(S29)}$$

Since the excess of **B** or **C** (depending in the particular titration point) are accounted already by the quantities $[\mathbf{B}]_x$ and $[\mathbf{C}]_x$ it follows that in our approximation [**C**] = [**B**], therefore:

$$[CB] = K_1[C]^2 \quad \text{(S30)}$$

which substituted in equation (S28) results in:

$$[CB]_0 = \frac{K_2K_1[C]^2[B]_x}{(1+K_2K_1[C]^2)} + \frac{K_3K_1[C]^2[C]_x}{(1+K_3K_1[C]^2)} + \frac{K_1[C]^2}{(1-K_oK_1[C]^2)^2} + \frac{2K_l(K_1[C]^2)^2}{(1-K_lEM_DK_o^2(K_1[C]^2)^2)^2} \quad \text{(S31)}$$

Changes in the second derivative of the absorbance at 458 nm, *A''*, are attributed to changes in concentration of building block **CB** present in the double stranded polymer **D** can can be written as:

$$A'' = E_D\sum_{n=1}^{\infty} 2n[D] \quad \text{(S32)}$$

where E_D is the proportionality factor between the concentration of **CB** units within the double stranded polymer and the second derivative of the absorbance at the wavelength under study. The concentration of **CB** within the double stranded polymer can be written as a function of the concentration of free, non-excess, **C** in equilibrium as follows:

$$\sum_{n=1}^{\infty} 2n[D] = \frac{2K_l(K_1[C]^2)^2}{(1-K_lEM_DK_o^2(K_1[C]^2)^2)^2} \quad \text{(S33)}$$

Therefore, *A''* can be written as

$$A'' = E_D\frac{2K_l(K_1[C]^2)^2}{(1-K_lEM_DK_o^2(K_1[C]^2)^2)^2} \quad \text{(S34)}$$

A convenient way to represent UV data where the spectra are recorded at different total concentration of **C** (such as, for example, the dilution data, see below) is to divide the observed absorbance, or its second derivative, by the total concentration of **C**. The UV data thus normalized can be then fitted to a modified version of equation (S34), i.e.:

$$\frac{A''}{[C]_0} = E_D\frac{2K_l(K_1[C]^2)^2}{[C]_0(1-K_lEM_DK_o^2(K_1[C]^2)^2)^2} \quad \text{(S35)}$$

To fit the experimental data we first obtained the values of $[\mathbf{CB}]_0$, $[\mathbf{C}]_x$ and $[\mathbf{B}]_x$ from $[\mathbf{C}]_0$ and $[\mathbf{B}]_0$. $[\mathbf{C}]_0$, $[\mathbf{CB}]_0$, $[\mathbf{C}]_x$ and $[\mathbf{B}]_x$ were then used as the independent variable input in the model described by equations (S30) and (S35). We used Micromath Scientist 3.0 to fit the data. In the fitting of the data all parameter were fixed except for K_I and E_D (Table 1 and S2). The value of E_D calculated from the fitting procedure was then used to transform the UV data into the percentage of C in the double stranded polymer as shown in Fig. 3, i.e.:

$$\%\ of\ C\ at\ D = 100 \times \frac{A''}{E_D} \qquad \text{(S36)}$$

See Supplementary Fig. S6 for the second derivative of the spectra of non-confined **C** upon addition of **B** and for the fitting of the data, and Table 1 and Table S2 for a summary of the fitting results.

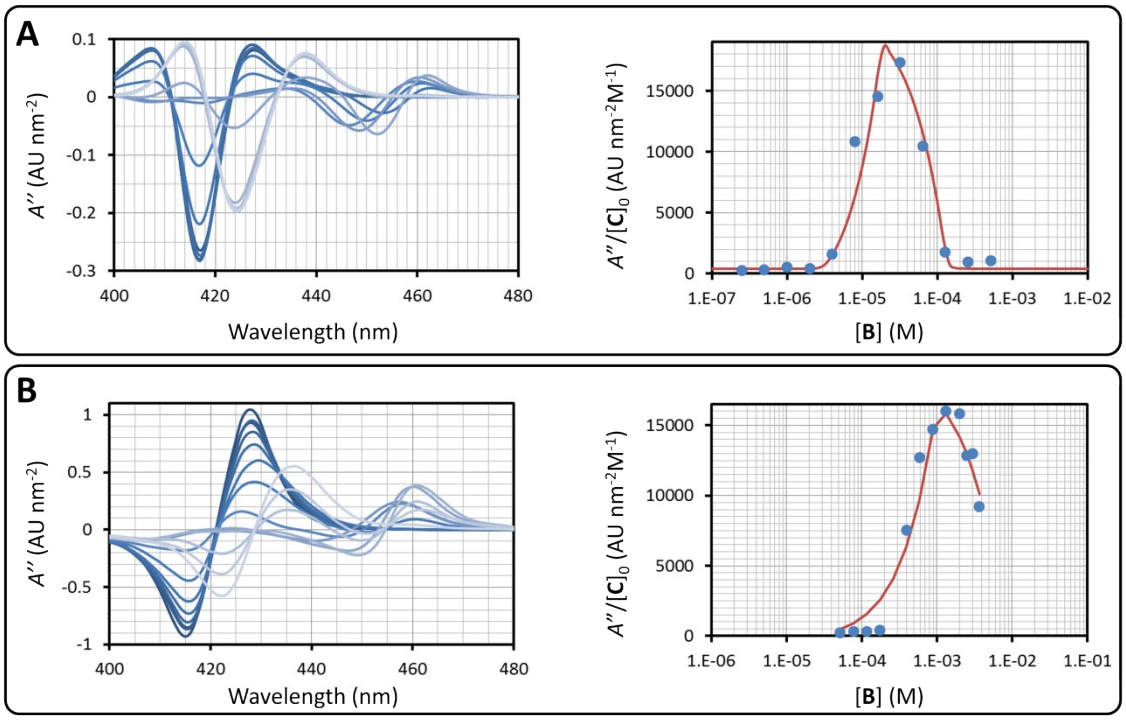

Figure S6. A. Second derivative of the Soret band region of the UV spectrum of **C** in the presence of increasing concentrations of **B** (left panel). The panel in the right shows the experimental values of the second derivative of the absorbance at 458 nm (blue circles) and the best fitting to the binding model defined by equations (S31) and (S35). The concentration of **C** was 18 μM and the concentration of egg-yolk phosphatidylcholine (EYPC) was 0.5 mM.B. Idem, with a concetration of **C** 1 mM.

Table S2. **Fitting parameters obtained from the data used in the displayed figures.**

	Fig. 3, Fig. S6A (Titration, non-confined)	Fig. 4b, Fig. S6B (Titration, non-confined)	Fig. 4c, Fig. S8A (Titration, confined)	Fig. 4d, Fig. S8B (Titration, confined)	Fig. 5a, Fig.S10 (Dilution, confined)
K_I	5800	5500	8700	9100	11300
E_D	22000	16500	15800	13800	16400

The units for K_I are in M^{-1} and for E_D are AU nm^{-2} M^{-1}

Nucleation concentration (NC)

Our assembly model allow us to determine the NC from the pairwise constants as follows:[29]

$$NC = \frac{(K_l EM_D)^{1/m}}{K_l EM_D K_o (1 - \frac{(K_l EM_D)^{1/m}}{K_l EM_D})^2} \quad \text{(S37)}$$

where EM_D has been assumed to be 1 in our calculations. The value of NC is therefore

$NC = 1.93 \mu M$

Integrity of the vesicles at different concentrations of B by cryo-EM

Cryo-EM images show that the presence of increasing amounts of **B** does not have a major effect on the appearance of the lipid vesicles within the samples in our experimental conditions (Supplementary Fig. S7 and Fig. 8)

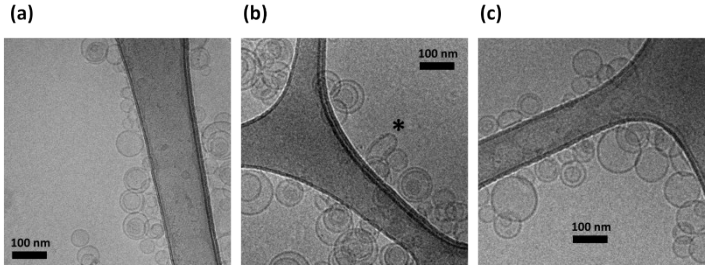

Figure S7. Cryo-TEM images generated for samples of vesicle-confined **C** with concentration of **B** 0 μM (a), 32 μM (b) and 2000 μM (c). The symbol (*) higlights a non-physically constrained flattened vesicle.

Assembly model for binding of B to confined C.

In a strict notation, when **C** is confined the expressions for the binding constants should reflect the relevant concentration of specie involved in the binding event, that is:

$$K_o = \frac{[(CB)_n]_i}{[(CB)_{n-1}]_i[CB]_i} \quad \text{(S38)}$$

$$K_l{}^{n}{EM_D}^{n-1} = \frac{[D]_i}{[(CB)_n]_i^2} \quad \text{(S39)}$$

$$K_1 = \frac{[CB]_i}{[C]_i[B]} \quad \text{(S40)}$$

$$K_2 = \frac{[CB_2]_i}{[CB]_i[B]} \quad \text{(S41)}$$

$$K_3 = \frac{[C_2B]_i}{[CB]_i[C]_i} \quad \text{(1)}$$

where the suffix *i* denotes concentration in relation to the confined volume and the absence of suffix concentration in relation to the bulk volume.

The ratio between concentrations of confined species in relation to the confined volume equals the ratio of these same species in relation to the bulk solution. For example:

$$\frac{[CB]}{[C]} = \frac{[CB]_i}{[C]_i} \quad \text{(S42)}$$

Therefore, equations S38-S41, and 1 can be written as

$$K_o = \frac{[(CB)_n]}{[(CB)_{n-1}][CB]_i} \quad \text{(S43)}$$

$$K_l{}^{n}{EM_D}^{n-1} = \frac{[D]}{[(CB)_n][(CB)_n]_i} \quad \text{(S44)}$$

$$K_1 = \frac{[CB]}{[C][B]} \quad \text{(S8)}$$

$$K_2 = \frac{[CB_2]}{[CB][B]} \quad \text{(S9)}$$

$$K_3 = \frac{[C_2B]}{[C][CB]_i} \quad \text{(S10)}$$

We define *Z* as the ratio between the bulk (*V*) and confined (V_i) volumes, or the ratio between local and bulk concentration of any confined species, i.e.:

$$Z = \frac{V}{V_i} \quad \text{(S45)}$$

Therefore *Z* can be written as the ratio of the concentration of the local concentration of any confined species in relation to the concentration of this species in relation to the bulk volume, that is:

$$Z = \frac{[C]_i}{[C]} = \frac{[CB]_i}{[CB]} \quad \text{(S46)}$$

Substituting *Z* in equations (S43) , (S44) and (10) we have that:

$$K_o Z = \frac{[(CB)_n]}{[(CB)_{n-1}][CB]} \quad (4)$$

$$K_l{}^n EM_D{}^{n-1} Z = \frac{[D]}{[(CB)_n]^2} \quad (5)$$

$$K_3 Z = \frac{[CB_2]}{[CB][C]} \quad (3)$$

which allows referring all the equilibria to bulk solution concentrations. Equations (3)-(5), (S8) and (S46) are used to obtain a modified version of the polymerization equation (S31):

$$[CB]_0 = \frac{K_2 K_1 [C]^2 [B]_x}{(1 + K_2 K_1 [C]^2)} + \frac{K_3 Z K_1 [C]^2 [C]_x}{(1 + K_3 Z K_1 [C]^2)} + \frac{K_1 [C]^2}{(1 - K_o Z K_1 [C]^2)^2} + \frac{2K_l Z (K_1 [C]^2}{(1 - K_l EM_D K_o^2 Z^2 (K_1} \quad \text{(S47)}$$

Like for non-confined **C**, changes in the second derivative of the absorbance at 458 nm, *A''*, are attributed to changes in concentration of building block **CB** in the of double stranded polymer **D** and can be written as:

$$\frac{A''}{[C]_0} = E_D \frac{2K_l Z (K_1[C]^2)^2}{[C]_0(1 - K_l E M_D K_o^2 Z^2 (K_1[C]^2)^2)^2} \qquad \text{(S48)}$$

To fit the experimental data we first obtained the values of $[\mathbf{CB}]_0$, $[\mathbf{C}]_x$ and $[\mathbf{B}]_x$ from $[\mathbf{C}]_0$ and $[\mathbf{B}]_0$. $[\mathbf{C}]_0$, $[\mathbf{CB}]_0$, $[\mathbf{C}]_x$ and $[\mathbf{B}]_x$ and Z were then used as the independent variable input in the model described by equations (S47) and (S48) (Fig. 4c and d and Supplementary Fig. S8) . We used Micromath Scientist 3.0 to implement de model. In the fitting of the data, all parameter are fixed except for K_l and E_D (Table 1 and S2).

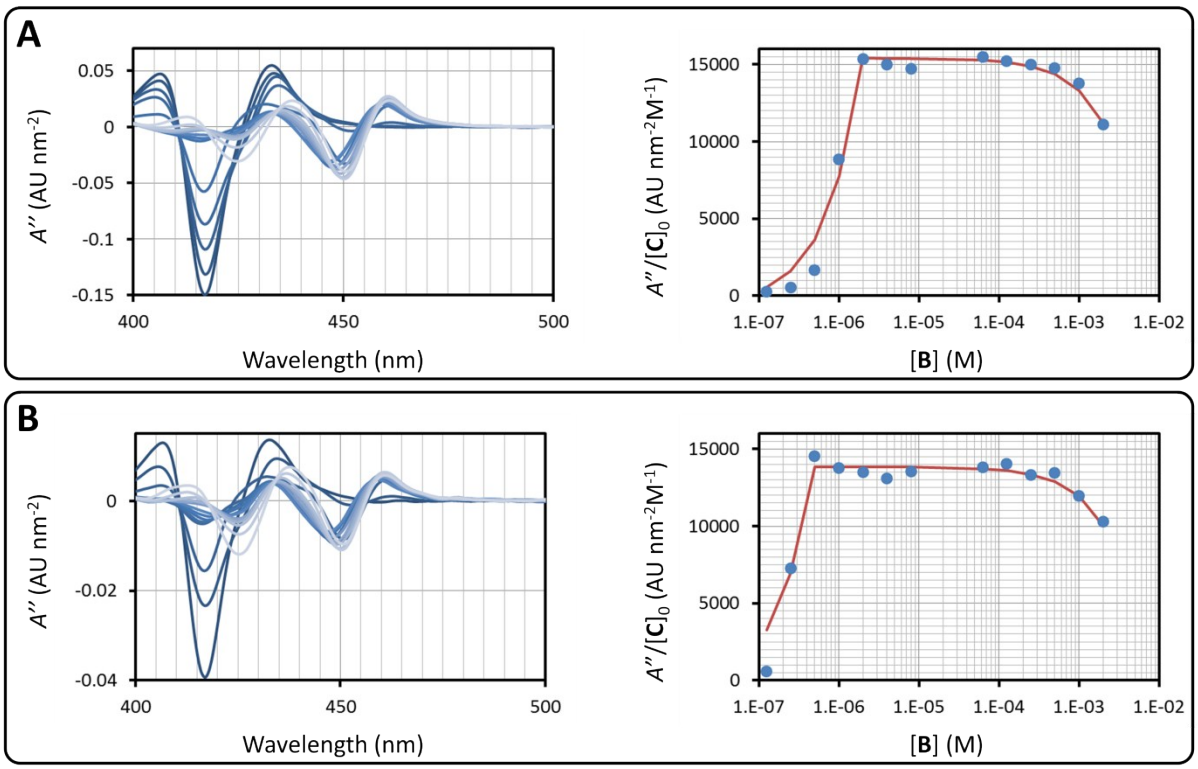

Figure S8. A. Second derivative of the Soret band region of the UV spectrum of **C** in the presence of increasing concentrations of **B** (left panel). The panel in the right shows the experimental values of the second derivative of the absorbance at 458 nm (blue circles) and the best fitting to the binding model defined by equations (S47) and (S48). The apparent concentration of **C** was 1.9 μM, the local concentration in the vesicle cavity was 1 mM and that of EYPC was 500 μM. C. Idem, for a sample with apparent concentration of **C** 0.45 μM and that of EYPC 125 μM.

In order to remove effects of baseline drift and the scattering of the liposomes the UV spectra were converted into the second derivatives, as explained in the Methods section. In the second derivative, the band attributed to the double stranded polymer appears between 450 and 480 nm (Fig. S8). Upon addition of **B** this band first grows, as the amount of double stranded polymer increases. After an equimolar amount of **B** has been added the intensity of the band does not increases, but it experiences a shift to slightly longer

wavelength as the concentration of **B** is increased. Either **C** or **B** can bind to the ends of a double stranded polymer formed of **CB** building blocks (Fig. S9). This process can be accounted for approximately introducing the appropriate modifications in equation (S48). In conditions where there is an excess of **C**, the binding of **C** to the double stranded polymer **D** can be written as

$$K_3 Z = \frac{[CD]}{[D][C]} \tag{S49}$$

$$(K_3 Z)^2 = \frac{[C_2 D]}{[D][C]^2} \tag{S50}$$

We can safely assume that the concentration of double stranded polymer ends is very small in relation with the excess of **C**. Therefore, equation (S49) and (S50) can be written as

$$K_3 Z = \frac{[CD]}{[D][C]_x} \tag{S51}$$

$$(K_3 Z)^2 = \frac{[C_2 D]}{[D][C]_x{}^2} \tag{S52}$$

Similarly, for the binding of **B** we have:

$$K_2 = \frac{[BD]}{[D][B]_x} \tag{S53}$$

$$(K_2)^2 = \frac{[B_2 D]}{[D][B]_x{}^2} \tag{S54}$$

The total concentration of **C**-caps in the double stranded polymer can be written as

$$[Ccaps] = \sum_{n=1}^{\infty} [BD] + \sum_{n=1}^{\infty} 2[D] + \sum_{n=1}^{\infty} 3[CD] + \sum_{n=1}^{\infty} 4[C_2 D] \tag{S55}$$

And that of **B**-caps is

$$[Bcaps] = \sum_{n=1}^{\infty} [CD] + \sum_{n=1}^{\infty} 2[D] + \sum_{n=1}^{\infty} 3[BD] + \sum_{n=1}^{\infty} 4[B_2 D] \tag{S56}$$

Substituting equations (S45), (5), (S51)-(S54) in (S55) we have:

$$[Ccaps] = K_2[B]_x[CB]^2K_l\sum_{n=1}^{\infty}(K_lEM_D(K_o)^2[CB]^2)^{n-1} + 2[CB]^2K_l\sum_{n=1}^{\infty}(K_lEM_D$$
$$+ 3K_3Z[C]_x[CB]^2K_l\sum_{n=1}^{\infty}(K_lEM_D(K_o)^2[CB]^2)^{n-1} + 4(K_3Z[C]_x)^2[CB]^2$$
$$\sum_{n=1}^{\infty}(K_lEM_D(K_o)^2[CB]^2)^{n-1}$$

(S57)

Applying the corresponding Taylor formula for convergent series and re-arranging we have that

$$[Ccaps] = \frac{(2 + K_2[B]_x + 3K_3Z[C]_x + 4(K_3Z[C]_x)^2)2K_lZ(K_1[C]^2)^2}{1 - K_lEM_DK_o^2Z^2(K_1[C]^2)^2}$$

(S58)

Substituting equations (S45), (5), (S51)-(S54) in (S56) and repeating the same procedure for **B** caps we have:

$$[Bcaps] = \frac{(2 + K_3Z[C]_x + 3K_2[B]_x + 4(K_2[B]_x)^2)2K_lZ(K_1[C]^2)^2}{1 - K_lEM_DK_o^2Z^2(K_1[C]^2)^2}$$

(S59)

We define the relative concentration of **C**-caps x_C as:

$$x_C = \frac{[Ccaps]}{[Ccaps] + [Bcaps]} \qquad \text{(S60)}$$

Substituting in equations (S58) and (S59) in (S60) we have that:

$$x_C = \frac{2 + K_2[B]_x + 3K_3Z[C]_x + 4(K_3Z[C]_x)^2}{4 + 4K_2[B]_x + 4K_3Z[C]_x + 4(K_3Z[C]_x)^2) + 4(K_2[B]_x)^2} \qquad \text{(S61)}$$

Similarly, the relative concentration of **B**-caps is:

$$x_B = \frac{2 + K_3Z[C]_x + 3K_2[B]_x + 4(K_2[B]_x)^2}{4 + 4K_2[B]_x + 4K_3Z[C]_x + 4(K_3Z[C]_x)^2) + 4(K_2[B]_x)^2} \qquad \text{(S62)}$$

Changes in x_C and x_B calculated using equation (S61) and (S62) follow a similar trend than changes in the second derivative of the absorbance observed at concentrations of **B** where the double stranded polymer is the main species (Fig. S9), which is consistent with the shift in this band being due to changes in the relative amount of **C** and **B** capped polymers.

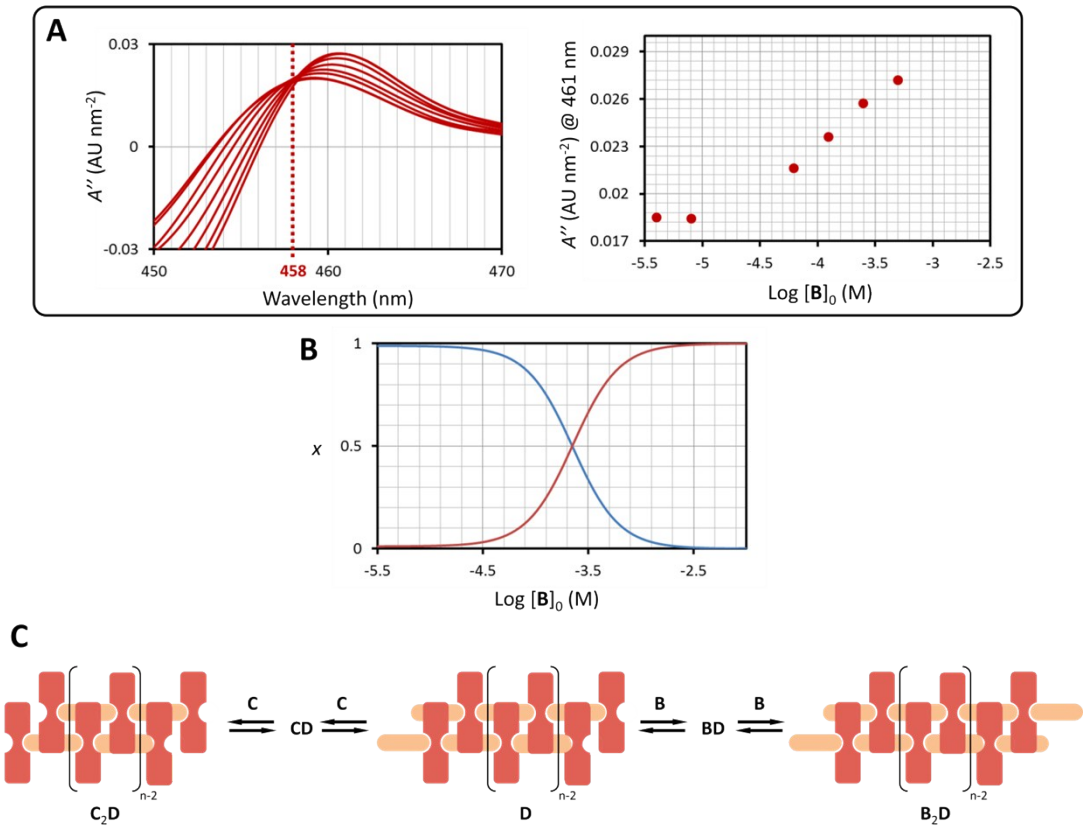

Figure S9. A. Detail of the band around 460 nm seen in the second derivative of the spectrum of **C** upon addition of increasing amounts of **B**. The isosbestic point at 458 nanometers is highlighted (left panel). The image is a zoom-in of the spectra displayed in Fig. S8A. The right panel shows the corresponding changes in absorbance at 461 nm. B. Change in the relative amounts of **C**-capping (blue trace) and **B**-capping (red trace) in the double stranded polymers, calculated using equations (S60) and (S61) in the experimental conditions used in A. C. Cartoon representation of the binding of **B** and **C** to double stranded polymer **D**.

During the fitting of titration data all the parameters where fixed except for K_I and E_D. In all experiments, the value of K_I was similar and, for the confined **C** so was the value of E_D (Table S2). The value of E_D for the non-confined titration is somewhat different, due to the fact that the Soret band has a slightly different shape when **C** is confined. This difference in shape is attributed to the interaction of confined **C** with the membrane, which concentration in the

cavity is up to 0.15 M.[17] Nonetheless, this interaction does not appear to interfere significantly in the assembly of **C** and **B**. The value of E_D for each corresponding experiment was used to generate the percentage value of **C** in the double stranded polymer for each experimental point as well as for the fitted curve, as seen in Fig. 4, using equation (S35)

For the dilution experiment the average value of E_D obtained from the fitting of the titration data of confined **C** was used to convert the experimental data (from the normalized second derivative of the spectra at 458 nm) to the percentage of **C** in the double stranded polymer. The data was fitted to the model described by equations (S47) and (S48). Form the fitting values for E_D and K_I were obtained and were the similar to those obtained from the fitting of the titration data (Table S2, Fig. 5a, Supplementary Fig. S10). The spectra shown in Fig. 5a have been normalized by dividing the values of absorbance by the total concentration of **C**, $[\mathbf{C}]_0$. The corresponding raw spectra and the normalized second derivative are shown in Fig. S10.

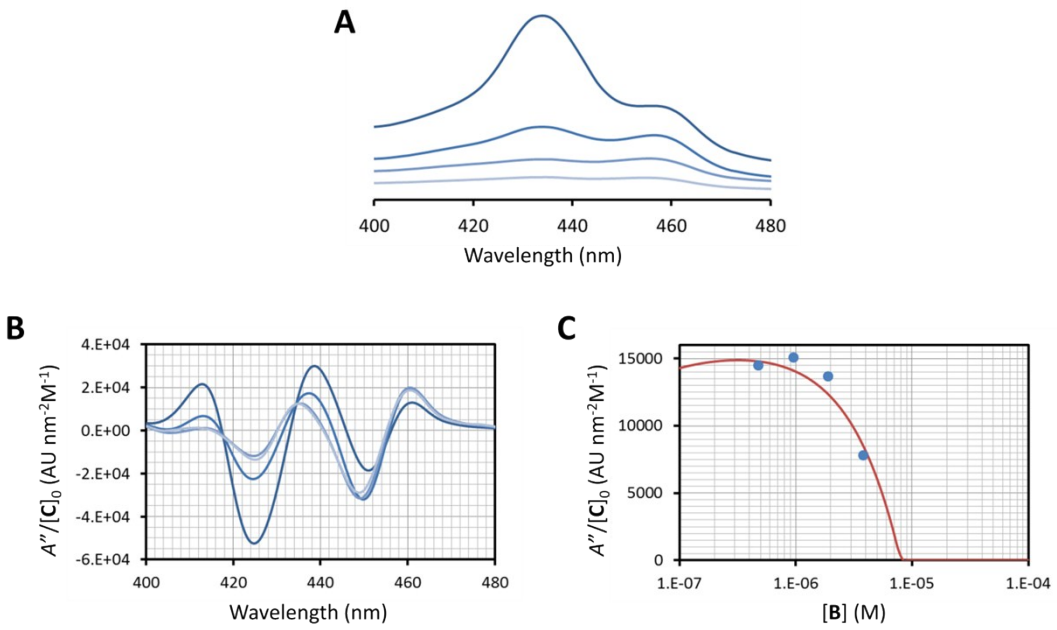

Figure S10. A. Soret band region of the UV spectrum of lipid vesicle confined **C** in the presence of an excess of **B**. In the most concentrated sample, corresponding to the darkest trace, the concentrations where $[\mathbf{C}]_0$ = 4 μM, $[\mathbf{C}]_i$ = 1 mM, $[\mathbf{B}]_0$ = 3 mM and [EYPC] = 1 mM. Each of the spectra is half the concentration of the preceding one. B. Second derivative of the spectra shown in A, normalized by the total concentration of **C**. C. Changes in the normalized second derivative of the absorbance at 458 nm corresponding to the spectra shown in panel A and B (blue circles). The red trace is the best fit to the model described by equations (S47) and (S48). See Table S2.

Average number of repeats for the free and confined double stranded polymer

The total number of repeats of a double stranded polymer <*N*> can be written as a function of free **C** as follows:[29]

$$< N >= \frac{1}{1 - K_l EMK_o^2 Z^2 (K_1 [C]^2)^2} \qquad \text{(S63)}$$

Using equation (S31) for un-confined **C** and **B**, or (S47) for confined **C**, and the pairwise binding constants obtained from the fitting (Table 1) it is possible to simulate the changes in <*N*> with changes in the concentration of both building blocks (Fig 7).