This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Insight into the self-assembly of water-soluble perylene bisimide derivatives through a combined computational and experimental approach

Emily R. Draper,1# Liam Wilbraham,2 Dave J. Adams,1 Matthew Wallace,3 Ralf Schweins,4 Martijn A. Zwijnenburg2*

1 School of Chemistry, University of Glasgow, Glasgow, G12 8QQ (UK).
2 Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ (UK).
3 School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ (UK).
4 Large Scale Structures Group, Institut Laue-Langevin, 71 Avenue des Martyrs, CS 20156, F-38042 Grenoble, CEDEX 9, France

# corresponding author: emily.draper@glasgow.ac.uk
* corresponding author: m.zwijnenburg@ucl.ac.uk

We use a combination of computational and experimental techniques to study the self-assembly and gelation of water-soluble perylene bisimides derivatised at the imide position with an amino acid. Specifically, we study the likely structure of self-assembled aggregates of the alanine-functionalised perylene bisimide (PBI-A) and the thermodynamics of their formation using density functional theory and predict the UV-vis spectra of such aggregates using time-dependent density functional theory. We compare these predictions to experiments in which we study the evolution of the UV-Vis and NMR spectra and rheology of alkaline PBI-A solutions when gradually decreasing the pH. Based on the combined computational and experimental results, we show that PBI-A self-assembles at all pH values but that aggregates grow in size upon protonation. Hydrogel formation is driven not by aggregate growth but reduction of the aggregation surface-charge and a decrease in the colloidal stability of the aggregation with respect to agglomeration.
Introduction

Perylene bisimides (PBIs, or perylene diimides, PDIs) and related molecules, such as naphthalene bisimides, form a fascinating class of compounds. When functionalised with suitable substituents, they can form gels, liquid crystals or be dried down to form amorphous thin films. The resulting materials can act as photoconductors,\textsuperscript{1-6} with very long charge carrier lifetimes even in the presence of air,\textsuperscript{3} as hydrogen evolution photocatalysts,\textsuperscript{7-9} and battery electrolytes,\textsuperscript{10} as well as materials that structural respond to electrochemical reduction\textsuperscript{11} or illumination.\textsuperscript{12} All these properties arise from the interplay between their propensity to self-assemble and their rich redox and photochemistry.

PBIs can be functionalised both at the imide nitrogen and at the bay carbon atoms, where the role of the functional groups is three-fold; to (i) improve the solubility of the PBI, (ii) directly modify the optical and redox properties of the PBI, and/or (iii) indirectly modify the optical and redox properties of the PBI by changing the structure in which it self-assembles. Typically, long and branched alkyl chains are used as functional groups to improve the solubility in organic solvents and ionisable groups or oligo(ethylene oxide) chains to make the PBIs more water-soluble. The alkyl-chain functionalised PBIs can form organogels\textsuperscript{13-16} upon lowering the solution temperature or increasing the PBI concentration. PBIs functionalised with oligo(ethylene glycol) groups can form hydrogels upon mixing with water\textsuperscript{17-18} and PBIs functionalised with ionisable groups can form hydrogels\textsuperscript{2-6, 12, 19-22} upon neutralisation of the charge. Both organic and water-soluble functionalised PBIs have also reported to be able to form columnar liquid crystalline systems.\textsuperscript{23-29}

PBIs derivatised with amino acids at the imide nitrogen positions,\textsuperscript{2-6, 12, 21} can be easily prepared by reacting perylene-3,4,9,10-tetracarboxylic dianhydride with an amino acid at high temperature with imidazole, and are examples of PBIs functionalised with ionisable groups and the focus of this paper. Such amino acid functionalised PBIs are soluble in water at high pH, when the terminal carboxylic groups are deprotonated, and can form hydrogels upon lowering the pH and hence neutralisation of the charge. Previous work has shown that the UV-Vis spectra and physical properties of the formed gel or dried down solution change with the exact amino acid group used.\textsuperscript{5-6}

The amino acid functionalised PBIs display relatively subtle changes in the UV-Vis spectrum upon a reduction in pH.\textsuperscript{3, 5-6} The alkyl-chain functionalised PBIs in contrast show a quite dramatic change in the UV-Vis absorption and fluorescence spectrum upon decreasing the temperature of a hot solution.\textsuperscript{13-16} These changes in the optical properties, subtle or large, and the gelation or liquid crystal phase formation itself are linked to the growth of supramolecular aggregates and hence information about the structure of the aggregates formed is crucial to understand the properties of the formed material. As organogels and hydrogels – as well as the thin films formed from drying down the solutions and the xerogels formed by drying the gels – are typically amorphous, and the liquid crystalline system by definition has limited long-range order, theory must play an important role in extracting this information by linking (changes in) the spectral features to aggregate structure and vice versa.
The simplest approach to linking structures and spectra is based on a naïve use of Kasha's exciton model. For example, the formation of J- or H-aggregates is often suggested based on the red shift or blue shift of the UV-Vis absorption spectrum. However, we and others have argued that this interpretation is problematic as PBIs are unlikely to assemble in structures resembling the textbook J- or H-aggregate and the electronic coupling between the molecules in the aggregates is likely more complicated than assumed by Kasha and co-workers. A more sophisticated approach is based on building actual structural models of possible aggregates, optimising their structures using, for example, density functional theory (DFT) and calculating their optical properties using time-dependent DFT (TD-DFT) or other excited state methods. There are a number of papers in which this approach is applied to derivatised PBIs that form organogels and/or liquid crystalline phases, in which the aggregates are typically modelled as a (face-to-face) stacked dimer, often of unsubstituted PBI rather than its functionalised counterpart present in the experiment. In our previous work on PBIs derivatised with amino acids, we also modelled dimers but including the substituents. Recently Segaline and co-workers computationally studied aggregates of up to seven units for a PBI functionalised with quaternary nitrogen groups.

Scheme 1 Structure of PBI-A.

Here, we employ the computational approach sketched out above in combination with experimental spectroscopy and rheology to understand the changes upon gelation of amino acid functionalised PBIs using PBI-A, PBI functionalised with alanine at the imide nitrogen position (see Scheme 1), as an example. In contrast to most previous work, as discussed above, we explicitly model PBI-A aggregates and consider aggregates beyond a stacked dimer. Specifically, we explore different packing motifs of PBI-A and study the evolution of the optical properties upon self-assembly by considering aggregate models containing up to four PBI-A molecules. We address the question of why the changes in the UV-Vis spectrum of PBI-A and other amino-acid derivatised PBIs is so much subtler than is the case for their alkyl-chain substituted equivalents. We compare the prediction from theory to time-dependent UV-Vis, NMR and rheology experiments. Using both theory and experiment, we propose a microscopic model for what happens during gelation. Using the case of PBI-A as an illustration, we also show the importance of explicitly modelling the functionalised PBIs rather than unsubstituted PBI, as well as aggregates beyond a dimer.

Methodology

Ground-state optimisation

The structures of the different aggregate models were optimised using the PBEh-3c and B97-3c DFT methods by Grimme and co-workers, as implemented in Turbomole 7.01 and 7.3 respectively. PBEh-3c and B97-3c were especially
developed to give accurate predictions of the geometries and thermochemistry of supramolecular systems for a minimal computational cost. Use of PBEh-3c and B97-3c allows us to study much larger systems than otherwise possible.

For selected structures, the energies were also calculated using single-point B3LYP+D3\textsuperscript{39-43} and \omega B97XD\textsuperscript{44} DFT calculations, as well as (SCS/SOS-)RI-MP2\textsuperscript{45} and RI-RPA\textsuperscript{46} calculations, on the structures optimised using PBEh-3C. The additional DFT calculations used the triple-zeta def2-TZVP basis-set and were performed using Turbomole 7.01 (B3LYP+D3) and Gaussian16 (\omega B97XD) respectively. The (SCS/SOS-)RI-MP2 and RI-RPA calculations used the def2-TZVPP\textsuperscript{47} basis-set, involved the RI approximation and in the case of RI-RPA also the RI-JK approximation, and were performed using Turbomole 7.3.

For all aggregate models the harmonic frequencies were calculated using PBEh-3c and for dimers and monomers also with B97-3c. These frequency calculations allowed us to verify that the stack models corresponded to minima on the PBEh-3c/B97-3c potential energy surfaces, as well calculate the non-electronic, i.e. rotational, translational and vibrational; contributions to the free energies of the models. The free-energy calculations typically made the quasi-RRHO approximation by Grimme\textsuperscript{48} in which low frequency modes (\omega < 100 cm\textsuperscript{-1}) are treated as rotations rather than vibrations when calculating the entropy (using \omega_0 = 100 cm\textsuperscript{-1} and \alpha = 4), and no frequency rescaling. Finally, all calculations, except stated otherwise, involve a implicit solvation model, COSMO\textsuperscript{49} (\varepsilon_r 80) in the case of Turbomole 7.01/7.3 and SMD\textsuperscript{50} in the case of Gaussian16, to describe the aqueous environment of the aggregates.

**Optical property calculations**

UV-Vis spectra of the different aggregate models were predicted using time-dependent DFT (TD-DFT) single-point calculations on the structures optimised with PBEh-3c. These TD-DFT calculations, unless stated otherwise, were performed using the \omega B97X\textsuperscript{51} functional, the 6-31G**\textsuperscript{52-53} basis-set and the SMD solvation model in Gaussian09. By necessity these calculations approximate the UV-Vis spectra by the vertical excitation spectrum, non-vertical vibronic effects have been argued to appear in the UV-Vis spectra of PBI aggregates\textsuperscript{31, 54} and can in principle be modelled,\textsuperscript{55} however such calculations would be computationally intractable for the size of aggregates studied here. Adding explicit water molecules, hydrogen-bonded to the carboxylic acid groups and/or imide carbonyl oxygen, on top of the implicit solvation model results in no appreciable changes in the predicted spectra.

**Structural analysis**

To interpret the results and to compare them to what we would expect based on Kasha’s exciton model, we need a set of parameters to describe the geometries of the different aggregates. As the transition dipole moment of the lowest excitation of an isolated PBI-A molecule is located along a vector that connects the two nitrogen atoms, i.e. the long axis of the molecule, a natural description of the geometry of PBI aggregates is in terms of four parameters per pair of molecules. There is (i) the angle between these vectors in two PBI molecules in the stack, (ii) the length of the vector connecting the centroids of the same two PBI molecules, and (iii) for each of the molecules, the angle between the vector connecting the two nitrogen atoms in one of the molecules etc.
and the vector connecting the centroids (see Fig. 1a). The angle (i) will be referred to as the “twist” angle and the angles in (iii) as the “slide” angles. For reference, the classical textbook structure of an H-aggregate has a twist angle of zero and slide angles of ninety degrees, while a similar textbook structure of the J-aggregate has a twist angle of zero and slide angles of zero.

**Thermodynamic analysis**

The stability of the PBI-A aggregates relative to isolated PBI-A molecules in solution, as well as their thermodynamic propensity to form, are analysed in term of their binding free energy $\Delta G_{b,n}$, defined as:

$$\Delta G_{b,n} = G_{agg,n} - n \times G_{\text{free}}$$  \hspace{1cm} (1)

where $G_{agg,n}$ is the free energy of an aggregate consisting of $n$ PBI-A molecules and $G_{\text{free}}$ that of a PBI-A molecule in solution. To compare the binding free energies between aggregates of different size, the binding energy can be normalised with respect to the number of molecules in an aggregate:

$$\Delta G_{b,norm} = \Delta G_{b,n} / n$$ \hspace{1cm} (2)

To understand the propensity of molecules to form larger aggregates, it is instructive to not only consider the binding free energy but also the free-energy change associated with the formation of an aggregate of $n$ PBI-A molecules from an aggregate of $n-1$ molecules and a free PBI-A molecule in solution $\Delta G_{\text{step,n}}$ defined as:

$$\Delta G_{\text{step,n}} = G_{agg,n} - G_{agg,n-1} - G_{\text{free}}$$ \hspace{1cm} (3)

$\Delta G_b$ is associated with the overall stability constant $\beta_n$ and $\Delta G_{\text{step,n}}$ with the step-wise equilibrium constant $K_n$. Obviously for the dimer $\beta_n$ and $K_n$ are the same but for larger aggregates they will be different. The step-wise equilibrium constants are important parameters for models that allow one to predict the degree of aggregation and the average length of aggregates in solution.

**Preparation of solutions**

PBI-A was prepared as described previously. Solutions were prepared in water by suspending the PBI-A in water and adding two molar equivalents of NaOH (using a 1M solution) to give a final solution concentration of PBI-A of 5 mg/mL. Gels were prepared by the addition of GdL (8 mg/mL).

**Spectroscopy**

UV–Vis absorption spectra were collected on a Cary 60 UV–Vis spectrophotometer (Agilent Technologies). The samples were measured in a 0.1 mm pathlength quartz cuvette (Hellma Analytics). Spectra were collected at a scan rate of 2 nm/s.

**Rheology**

Rheological measurements were performed using an Anton Paar Physica 301 rheometer. The time sweeps were performed using a 50 mm sandblasted plate, with a gap distance of 0.8 mm. The samples were prepared as discussed above and 2 mL transferred onto the plate immediately after the addition of GdL and the plate lowered on top. A strain of 0.5% and a frequency of 10 rad/s were maintained whilst $G'$ and $G''$ were recorded every 30 seconds until they had...
reached a plateau. The plate was flooded with oil to ensure the sample did not dry out. A zero force of 0 N was maintained throughout the experiments. All measurements were recorded in triplicate. Viscosity measurements were measured using a 75 mm cone at a set shear rate of 10 s$^{-1}$. The samples were loaded onto the plate as described above and viscosity measured every 30 seconds until the sample had started gelling (when data became very noisy). All measurements were conducted at 25°C.

**pH measurements**

A FC200 pH probe (HANNA instruments) with a 6 mm x 10 mm conical tip was used along with a pH turtle data logger for the pH measurements. The stated accuracy of the pH measurements is ±0.1. The pH was recorded every 30 seconds until a gel was formed. The samples were prepared in a 7 mL Sterlin vial as described above and the temperature maintained at 25 °C during the titration by using a circulating water bath.

**NMR**

5 mg/mL solutions of PBI-A were prepared in D$_2$O with two equivalents NaOD added. NMR analyses were performed at high pH prior to the addition of GdL ($t = 0$). An aliquot of this solution was then added to GdL as described above and transferred to a 5 mm NMR tube for analysis. A 75 μL glass capillary (New Era Enterprises) containing 30 mM 3-(trimethylsilyl)propionic-2,2,3,3-d$_4$ acid sodium salt (TSP) in D$_2$O was inserted into the sample to act as a chemical shift and integral reference. The time quoted below in the discussion of the changes in the NMR data with time corresponds to the time elapsed since the PBI-A solution was first added to the GdL.

All NMR experiments were performed on a Bruker Avance II 400 MHz wide bore spectrometer operating at 400.20 MHz for $^1$H. The temperature of the samples was maintained at 298±0.5 K, the variation in the temperature being less than 0.1 K. $^1$H integrals were recorded in 65536 points with a 30 degree pulse in two scans with a relaxation delay of 40 s and signal acquisition time of 4.1 s. $^{23}$Na spectra were recorded with a 33 μs pulse (π/2) and a signal acquisition time of 0.3 s. 6144 data points were collected in 512 scans giving a total acquisition time of 4 minutes. See Figs. S1 and S2 for example $^1$H and $^{23}$Na NMR spectra.

The $^1$H integrals discussed are normalised to the value measured with two equivalents NaOD in the absence of GdL. $^{23}$Na spectra were processed with 16384 points and an exponential line broadening factor of 3.0. RQCs were extracted by Lorentzian deconvolution of the spectra. Example $^1$H and $^{23}$Na spectra are provided in the Supporting Information. See refs. 56-57 for an in-depth discussion of the use of $^1$H integrals and $^{23}$Na RQC to probe gelation.

**Small Angle Neutron Scattering**

The solutions were prepared as described above in D$_2$O. The samples were measured in UV spectrophotometer grade, quartz cuvettes (Hellma) with a 2 mm path length. These were placed in a temperature-controlled sample rack during the measurements. Small angle neutron scattering (SANS) measurements were performed using the D11 instrument (Institut Laue Langevin, Grenoble, France). The neutron beam, with a fixed wavelength of 6 Å and divergence of $\Delta \lambda / \lambda =$
9%, allowing measurements over a Q [Q = 4πsin(θ)/λ] range of 0.001 to 0.4 Å⁻¹ using three sample-detector distances of 1.5 m, 8 m, and 39 m. After adding GdL to the solution, the sample was loaded onto the sample rack and the data collection started after 90 seconds (due the experimental requirements). Data was collected and averaged over 3 minutes. Repeat sample collection on the same sample was carried out over 10 hours to provide information as to the evolution of the structures.

The data were reduced to 1D scattering curves of intensity vs. Q using the facility provided software. The electronic background was subtracted, the full detector images for all data were normalized and scattering from the empty cell was subtracted. The scattering from D₂O was also measured and subtracted from the data, and put on absolute scale via the measurement of the secondary calibration standard H₂O with 1 mm path-length. The data were then fitted using the SasView software package version 4.2.0. The SLD of PBI-A was calculated assuming a density of 1.55 g/cm³ to be 3.445x10⁻⁶ Å⁻².

**Results and Discussion**

**Aggregate structures**

During our manual exploration of the potential energy surface of neutral protonated PBI-A aggregates, we found three main classes of structures, all face-to-face stacked helices (see Fig. 1b-h). Two, I and II, are based on a PBI-A conformer where both alanine groups are orientated in the same direction and one, III, on a conformer in which the two alanine groups are rotated by ~180° relative to one another. In class I aggregates, there are hydrogen bonds between the carboxylic acid protons and imide carbonyl oxygen atoms of adjacent PBI-A molecules in the aggregate. Class II and III aggregates lack these intra-aggregate hydrogen bonds. The different aggregates also display other differences in how they are stacked (see Table S1). The class I aggregates have twist angle of ~30° and 90° slides angles, while the class II and III aggregates have twist and slide angles of ~50° and ~110° respectively. Finally, on average the distance between the centroids of adjacent molecules is slightly shorter in the class II and III than in the class I structures. All structures we encountered in our exploration, finally, were right-handed helices but we have no reason to believe that left-handed helices would be less likely and experimental samples are likely to be a mixture.

Interestingly, the core structure of the class I aggregates is similar to that reported in the literature for unsubstituted PBI. We hypothesize that the reason that the class II aggregates have such a different core structure than aggregates of unsubstituted PBI, while naively both have similar intermolecular interactions (i.e. no hydrogen bonds) is related to the fact that PBI-A has a dipole while unsubstituted PBI has none, as well as an increased quadrupole.

In the case of aggregates where all of the carboxylic acid groups of the PBI-A molecule are deprotonated, i.e. the likely (de)protonation state at high pH, we only found two classes of structures. These doubly-deprotonated aggregate structures, IV and V, are very similar to the type II and III structures, respectively, observed for the neutral aggregate, with twist and slide angles of ~50° and ~110° respectively. The similarity between II and IV and III and V is, perhaps, not surprising as all lack the hydrogen bonds present in I.
**Fig. 1** a) definition of the centroid distance (R), twist angle (TA) and slide angles (SA); b) class I dimer; c) class II dimer; d) class III dimer; e) class IV dimer; f) class V dimer; g) class I tetramer; f) class II tetramer.

**Predicted binding (free) energies**

Focussing first on the neutral aggregates, the PBEh-3c calculations predict that class I-III aggregates all have a normalised binding energy of ~-90 kJ/mol with respect to isolated PBI-A molecules, which reduces to ~-40 kJ/mol for the binding free energy (see Table 1 and Table S2-S9 for the underlying data). Calculation of the binding (free) energy of the class I dimer (see Table S10), suggest that these number do not substantially change in the presence of explicit water molecules. The B97-3c calculations predict a similar but slightly more negative normalised binding energies of ~-100 kJ/mol, which again reduces to ~-45 kJ/mol when including free-energy effects. PBEh-3c predicts that class II aggregates are consistently more stable than class I and III aggregates by 5-10 kJ/mol in terms of binding (free) energy, while B97-3c predicts that class I aggregates are more stable than their class II and III counterparts by a similar amount. As a test, we ran ωB97XD, B3LYP+D3, MP2 and RI-RPA single-point calculations on the PBEh-3c optimised geometries of dimers of I and II. ωB97XD and B3LYP+D3 calculations gave a similar energy ranking as B97-3c, as did (SCS/SOS-)MP2/def2-TZVPP, while RI-RPA/def2-TZVPP predicted a similar ranking as PBEh-3c (see Table S11). Irrespective of what is the correct picture, all methods agree that the differences in free energy between the class I-III aggregate are very small and thus that in experiment very likely a mixture of aggregates is present.
Both PBEh-3c and B97-3c predict that doubly-deprotonated aggregate structures IV and V are less stable than (most of) their neutral counterparts, with normalised binding (free) energies relative to isolated doubly deprotonated PBI-A molecules of $\sim -75$ kJ/mol (PBEh-3c, $\sim -30$ kJ/mol, when including free energy effects) and $\sim -80$ kJ/mol (B97-3c, $\sim -30$ kJ/mol, when including free energy effects) respectively. Class V aggregates are predicted to be more stable than class IV aggregates by both PBEh-3c and B97-3c. However, the differences are small and thus again in experiment likely a mixture of aggregate types is present. The models for the doubly-deprotonated aggregate structures do not contain the solvated (Na$^+$) counter-ions present in experiment. Inclusion of solvated ions in static calculations is inherently fraught with problems. We suspect that their absence most likely does not significantly influence the predicted binding (free) energies.

### Predicted aggregate size and degree of aggregation

The fact that the doubly-deprotonated class IV and V aggregates have less negative normalised binding free energies than the neutral class I-III aggregates suggest that aggregates at high pH will be likely shorter than at low/neutral pH and that perhaps more PBI-A is present as free molecules. Use of a simple isodesmic model,\textsuperscript{59-60} which assumes that the step-wise equilibrium constants are the same for all aggregates sizes including the dimer (i.e. $\beta_2 = \beta_3 = \beta_n$), PBI-A concentration of 5 mg/mL (0.001 mol/L) and dimer equilibrium constants based on the binding energies mentioned above, suggest number average aggregate sizes of 100s to 1000s for neutral PBI-A and in between 10 to 200 for deprotonated PBI-A (see Table 2). Even at high pH, more than 99.9% of all PBI-A is predicted to be part of an aggregate using this isodesmic model.

### Table 1 Normalised binding (free) energies, step-wise addition free energy changes and step-wise equilibrium constants for aggregates of different size of the different aggregates classes, as predicted using PBEh-3c (P) and B97-3c (B). See table S2-S9 for the underlying data.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\Delta U_{b,norm}$</th>
<th>$\Delta G_{b,norm}$</th>
<th>$\Delta G_{step,n}$</th>
<th>$\beta_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>B</td>
<td>P</td>
<td>B</td>
</tr>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-87</td>
<td>-110</td>
<td>-36</td>
<td>-49</td>
</tr>
<tr>
<td>3</td>
<td>-84</td>
<td>-109</td>
<td>-29</td>
<td>-22</td>
</tr>
<tr>
<td>II</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-90</td>
<td>-101</td>
<td>-41</td>
<td>-43</td>
</tr>
<tr>
<td>3</td>
<td>-91</td>
<td>-103</td>
<td>-40</td>
<td>-38</td>
</tr>
<tr>
<td>III</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-85</td>
<td>-96</td>
<td>-37</td>
<td>-40</td>
</tr>
<tr>
<td>3</td>
<td>-86</td>
<td>-99</td>
<td>-35</td>
<td>-32</td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-79</td>
<td>-85</td>
<td>-33</td>
<td>-30</td>
</tr>
<tr>
<td>3</td>
<td>-73</td>
<td>-80</td>
<td>-25</td>
<td>-16</td>
</tr>
<tr>
<td>V</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-83</td>
<td>-85</td>
<td>-38</td>
<td>-31</td>
</tr>
<tr>
<td>3</td>
<td>-72</td>
<td>-24</td>
<td>-24</td>
<td>-11</td>
</tr>
</tbody>
</table>
Table 2 Number average aggregate size (N) and degree of aggregation values ($\alpha_{agg}$) for the different aggregates classes, as predicted using PBEh-3c and B97-3c.

<table>
<thead>
<tr>
<th></th>
<th>PBEh-3c</th>
<th>B97-3c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>$\alpha_{agg}$</td>
</tr>
<tr>
<td>I</td>
<td>138</td>
<td>0.9999</td>
</tr>
<tr>
<td>II</td>
<td>419</td>
<td>1.0000</td>
</tr>
<tr>
<td>III</td>
<td>157</td>
<td>1.0000</td>
</tr>
<tr>
<td>IV</td>
<td>72</td>
<td>0.9998</td>
</tr>
<tr>
<td>V</td>
<td>204</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Comparing the stepwise free energies for the dimer and trimer, however, suggests that the isodesmic model is probably too simple to describe PBI-A aggregation. While the normalised binding free energies of the dimer and trimer are very similar in all cases, the stepwise free energy of the trimer ($\Delta G_{step,3}$) is typically significantly smaller than that of the dimer ($\Delta G_{step,2}$). The approximation that the stepwise equilibrium constants are all similar, therefore, seems too crude and, as also previous observed for alkyl-chain substituted PBIs,

Predicted optical spectra

Figs. 2a-c show the TD-ωB97X predicted evolution of the spectra of neutral class I, II and III aggregates with the number of molecules in aggregates, as calculated on the PBEh-3c optimised structures. In all cases, the intensities of an aggregate of $n$ molecules are multiplied by $1/n$ to correct for the presence of more molecules. Using this correction, perhaps the most prominent observation is that the absorption intensity relative to isolated molecules in solution is predicted to decrease with increasing number of molecules in either aggregate. A similar decrease in intensity with increasing aggregate size was previously observed by Nogueira et al. in calculations on polynucleotides.

When considering the shape of the spectra of the aggregates it is instructive to start with the dimer. In the case of the class I aggregates, the long wavelength ($\lambda > 300$ nm) spectrum of the dimer displays essentially three peaks; a weak peak at $\lambda \sim 475$ nm, to the red of the absorption peak of the isolated molecule ($\lambda \sim 435$ nm), and to the blue, a strong peak at $\lambda \sim 423$ nm and a weaker peak at $\lambda \sim 375$ nm. Increasing the number of molecules in the aggregate, results in the slow
merger of all three peaks, where the “475 nm” peak shifts progressively to the blue and the “375 nm” peak to the red, suggesting that in the large aggregate limit the class I aggregates will display one significantly broadened peak slightly blue-shifted relative to that of an isolated molecule.

**Fig. 2** TD-ωB97X predicted spectra of a) class I; b) class II; c) class III; d) class IV; and class V aggregates of different size.

In the case of the class II and III dimer, the long wavelength spectrum displays one main peak at essentially the same wavelength as the isolated molecule but with a red-shifted shoulder and a weak blue-shifted peak at $\lambda \approx 370$ nm. Increasing the number of molecules in the aggregate, results, in contrast to the case of I, not in a further merger of peaks. The $\lambda \approx 370$ nm peak stays fixed, the main peak shifts slightly around, to red for odd-membered aggregates and to the blue for the even-membered case, while the red-shifted shoulder becomes more pronounced. Extrapolation the spectra to the large aggregate limit for class II and III aggregates is complicated due to the circular shifts of the main peak but most likely they will look quite similar to the dimer but with a more pronounced red-shifted shoulder. Fig. 2d-e shows the equivalent spectra for the class IV and V doubly-deprotonated aggregates. Clearly, also for these aggregates a reduction in absorption intensity with aggregate size is predicted. The shape of the predicted spectra is very similar to that of the structurally related class II and III neutral aggregates.

Overall, the calculations suggest that the main effect to be expected as a result of aggregation, be it for neutral or deprotonated PBI-A molecules, is a reduction in the absorption intensity, possibly together with the formation of a red-shifted shoulder.

**Comparison with experimental spectra**

Next, we compare the results of the calculations with time-resolved experiments. Here, starting from a solution of PBI-A and two equivalents of sodium hydroxide
(Na$_2$PBI-A), the pH is lowered by addition of glucono-δ-lactone (GdL), which hydrolyses to gluconic acid, and we study the change in the optical and mechanical properties as well as NMR spectra. The change in the pH is smooth bar a plateau from 60-150 minutes, see Fig. 3a, which can be linked to the half-way and equivalence points for the first protonation of PBI-A and thus the formation of significant amounts of singly-deprotonated PBI-A.

Fig. 3 a) Change in pH, b) UV-Vis spectra, c) G’ and G”, d) viscosity, e) $^{23}$Na RQC and $^1$H methyl integral, and f) SANS intensity at $Q = 0.020$ Å$^{-1}$ after addition of GdL. In the case of the UV-Vis spectra, the initial spectrum before GdL addition is shown in black, the first spectrum after GdL addition at 2 minutes in dark blue, and all the subsequent spectra in increasing lighter shades of blue.

We will start with the change in the absorption spectrum. Fig. 3b shows the time evolution of the experimental UV-Vis spectrum of a PBI-A solution after addition of GdL. Initially, minutes after addition of the GdL, the spectrum sharpens considerably and the shoulder at ~550 nm becomes more pronounced before the spectrum appears to dim and broaden. Overall, the main peak does not shift relative to before the addition of the GdL but a distinct new shoulder at ~600 nm appears. Another distinct change upon the addition of GdL is the clear gradual reduction in peak intensity with time.

Comparing the experimental spectra with the predictions discussed above, the first thing one notices is that the latter is blue-shifted by ~50 nm relative to the former. Such a blue shift is a well-known artefact of the use of range-separated functionals such as ωB97X, which avoid issues with the description of charge-transfer states, but at the expense of slightly blue shifting the overall spectrum. Indeed, in previous work we have often applied a rigid ad-hoc shift to correct for this, though we do not do that here. The second and more interesting observation is that, if we assume that the evolution of the experimental UV-Vis spectrum after addition of GdL is due to aggregation, experiment and our calculations agree on both the absence of large shifts in the spectrum upon aggregation and the decrease in absorption intensity. The initial sharpening of the spectrum then can be explained by the fact that addition of GdL temporarily
drives the system away from equilibrium, resulting in dissolution of the aggregates before they reform, and a momentarily increased concentration of monomeric species.

**Rheology**

Figs. 3c and 3d shows the time evolution of the sample’s storage and loss moduli ($G'$ and $G''$ respectively) and viscosity $\eta$. Based on these data, the sample behaves like a non-viscous Newtonian liquid up to 50 minutes after addition of the GdL ($\eta < 1$ Pa.S), a viscous non-Newtonian liquid ($\eta > 1$ Pa.S, $G'$ increase from $\sim 0.1$ to $\sim 10$) from 50-250 minutes before gelling ($G'$ jumps from $\sim 10$ to $\sim 1000$, and $G'$ dominates significantly over $G''$). The gelation thus appears to occur on a slower time-scale than the changes in the sample’s UV-Vis spectrum. The largest changes in the UV-Vis spectrum occur when the samples is still a liquid.

**NMR**

Fig. 3e shows the time evolution of the integral of the methyl peak in the $^1$H NMR spectrum and the $^{23}$Na residual quadrupolar coupling (RQC) splitting. Focussing first on the former, the methyl integral decreases slowly from 1 to $\sim 0.9$ in the first 150 minutes, before dropping at an increased rate and becoming negligible at approximately 450 minutes. The decrease of the $^1$H methyl integral with time is the result of an increasing fraction of the PBI-A being concentrated in NMR silent aggregates that have ceased rotating on the NMR time-scale because they are too large and/or become part of a solid phase, be it a gel or be it a precipitate. The $^{23}$Na RQC splitting is unobservable until 180 minutes, afterwards it rises to a maximum at $\sim 400$ minutes before slowly decreasing again. For $^{23}$Na RQC splitting to occur, aggregates should become sufficiently large to align with the magnetic field. The absence of observable $^{23}$Na RQC splitting before 180 minutes suggest that at that time the aggregates are too small for this to occur. Similarly the increase of the $^{23}$Na RQC splitting from 180 to 400 minutes suggest that with time an increasing amount of PBI-A is tied up in aggregates large enough for $^{23}$Na RQC splitting to be observable. Finally, the decrease in $^{23}$Na RQC splitting after 400 minutes can be explained by the increased protonation of PBI-A and reduction of the PBI-A net charge with time. Both NMR experiments thus suggest that assembly into aggregates that are sufficiently large for significant changes to the NMR spectra to occur happens on a slower time-scale than the changes in the sample’s UV-vis spectrum.

**Neutron scattering**

Initially, there is very little scattering. The first dataset can be fit to a power law only (see table S12), indicative of poorly scattering structures, perhaps existing as non-persistent aggregates. Over time, the scattering increased and the final data set after 10 hours was fitted to a flexible elliptical cylinder (see table S11) with a power law to take into account the scattering at low Q as we have described previously for other systems. To fit the data, the data were initially fitted over a reduced Q range from 0.01 to 0.4 Å$^{-1}$ using a flexible elliptical cylinder model only (this gave a better fit than a cylinder or flexible cylinder model). The full data set was then fitted using this model, and it was found that a power law was needed to take into account the excess scattering at low Q. This can be ascribed to scattering from the network.
To probe the evolution of the cylindrical structures forming, the intensity at a 0.020 Å\(^{-1}\) was followed with time. The intensity at this value of Q is indicative of the cylindrical structures. The data in Fig. 3f show that the scattering intensity increases after an induction period of around 22 minutes and reaches a plateau at around 300 minutes. The neutron scattering data thus suggests, similarly to the UV-Vis spectroscopy data, that the formation of fibrous aggregates is complete before gelation occurs.

**Discussion**

The combination of experimental UV-vis spectroscopy and (TD-)DFT calculations suggests that already in aqueous solution of Na\(_2\)PBI-A, PBI-A self-assembles in the form of face-to-face stacked helices and that a lowering of the pH merely results in an increase of the aggregate size. This combined with fact that gelation occurs at a longer time-scale than the largest changes in the UV-vis spectrum and neutron scattering data suggests that the observed gelation is not simply due to the formation of aggregation and perhaps not even linked to them growing especially large, although depending on the exact computational set-up neutral aggregates are predicted to be 4-40x larger than their doubly-deprotonated counterparts. An alternative explanation for the gelation, akin to that previously proposed by a number of us for the case of the gelation of a functionalised dipeptide,\(^{56}\) is that reduction in the surface-charge of the aggregates upon protonation makes them increasingly colloidally unstable with respect to agglomeration into a gel (see Fig. 4). Here, the reduction in surface-charge of the aggregates reduces the electrical double-layer that stabilises aggregates in the doubly-deprotonated state against agglomeration. This explanation is supported by the fact that the point in time where the \(^1\)H methyl integral starts to significantly decrease and the \(^23\)Na RQC splitting becomes observable, both of which require (a fraction of) aggregates to not display molecular rotation on the NMR time-scale, roughly coincides with what we assume to be the half-way and equivalence points for the first protonation of PBI-A and thus a significant reduction in the surface-charge relative to aggregates formed doubly-deprotonated PBI.

**Fig. 4** Cartoon description of the proposed gelation process, showing (a), the short aggregates present at high pH before GdL addition, (b), the initial dissolution of aggregates shortly after GdL addition, and, (c) to (e), the increase of aggregate size, reduction of aggregate surface charge and agglomeration of aggregate with increasing time and lower pH.

The more subtle changes in the UV-Vis spectra of amino acid functionalised PBIs upon lowering the pH relative to those observed for alkyl-chain functionalised PBIs upon lowering the temperature is probably due to a combination of two factors. Firstly, as discussed above, even at high pH the amino acid functionalised
PBIs self assemble into aggregates and the effect of lowering the pH is only in a transition from short to longer aggregates. In contrast, if alkyl-chain functionalised PBIs at high-temperature are truly monomeric, the self assembly triggered by temperature lowering for these systems is a true molecule to aggregate transition. It would not be surprising if such a more drastic change in the molecular environment would also result in a more drastic change in the UV-Vis spectrum. Secondly, our TD-DFT calculations suggest that for type of aggregate geometries predicted to occur for PBI-A, the change in spectra upon self-assembly are inherently relatively modest.

The calculations also provide insight into the likely structure of thin films formed by drying aqueous solution of Na\textsubscript{2}PBI-A. As the PBI-A in the initial dilute solutions of Na\textsubscript{2}PBI-A is predicted to already be aggregated, and as concentrating the solution during the drying will drive the equilibrium to longer aggregates, the structure of the thin films on the molecular scale will be very similar to that of the gels but denser. This prediction of a similar structure but different density is inline with the fact that experimentally thin films have similar optical properties as dried gels but much better photoconductivity.\textsuperscript{12}

From a computational point of view, finding the many different classes of aggregates that lie low in (free) energy is especially challenging. Here we found such structures by hand but for more complicated molecules, e.g. PBI functionalised with L-DOPA at the imide position,\textsuperscript{5} which has two additional hydroxyl groups that can hydrogen bond per L-DOPA substituent, a more systematic computational approach might be beneficial. An alternative might be sampling structures from molecular dynamics (MD), a strategy recently employed by Segalina \textit{et al.},\textsuperscript{36} but this might require many more TD-DFT calculations. The barriers between the different classes of aggregates also should be low enough for switching between them to be observable during a MD run.

\textbf{Conclusions}

Using a combination of DFT and TD-DFT calculations, experimental UV-vis and NMR spectroscopy and experimental rheology measurements, we studied the self-assembly and gelation of perylene bisimide functionalised with the amino acid alanine at the imide position (PBI-A). The DFT calculations predict that both fully-deprotonated and protonated PBI-A molecules self-assemble in helical aggregates but that the expected size of the aggregates is 4 – 40 times larger in the latter case. TD-DFT calculations predict that because aggregates are predicted to be formed in both cases and because their predicted structures are very similar, lowering of the pH should result in only minor changes in the shape of the UV-Vis spectrum of PBI-A solutions but a significant reduction in intensity. Experiments where the pH was gradually lowered through the hydrolysis of glucono-\textgreek{D}-lactone agree with that prediction. Gelation is found to occur on a much longer time-scale than the most significant changes in the UV-vis intensities, suggesting that gelation upon pH reduction is not primarily the result of aggregate growth. Using pH and NMR data we propose that instead gelation occurs because protonation reduces the surface charge on the aggregates, decreasing their colloidal stability, resulting in their agglomeration into a hydrogel. Thin films formed by drying fully-deprotonated PBI-A solutions, finally, are predicted to be similar in structure to the hydrogels but denser.
Acknowledgements

The authors kindly acknowledge Dr. Gerit-Jan Brandenburg, Dr. Tim Gould, Dr. Mike Porter and Dr. Felix Plasser for useful discussion. D.J.A. thanks the UK Engineering and Physical Sciences Research Council (EPSRC) for a Fellowship (EP/L021978/1), E.R.D. the Leverhulme Trust for funding (ECF-2017-223) and the University of Glasgow for a LKAS leadership Fellowship, and M.A.Z and L.W. the EPSRC for funding (EP/N004884/1). MW thanks The Royal Commission for the Exhibition of 1851 for a Research Fellowship. Dr Jonathan Iggo (University of Liverpool) is thanked for the use of his NMR spectrometer. The NMR spectrometer was funded by the EPSRC (EP/C005643/1 and EP/K039687/1). Computer time on ARCHER, the UK supercomputer, provided by the UK Materials Chemistry Consortium (EP/L000202/1) is kindly acknowledged. The neutron scattering experiment at the Institut Laue-Langevin was allocated beam time under experiment number 9-11-1905 (DOI: 10.5291/ILL-DATA.9-11-1905). This work benefitted from the SasView software, originally developed by the DANSE project under NSF award DMR-0520547.

Supporting information

Calculated total (free) energies; predicted twist angle, slide angle and centroid distances, example NMR spectra, DFT optimised structures, TD-DFT predicted excitation energies and oscillator values underlying the spectra in Fig. 2 and the experimental data underlying Fig. 3 are available online.

References


25. van Herrikhuizen, J.; Syamakumari, A.; Schenning, A.; Meijer, E. W., Synthesis of N-Type Perylene Bisimide Derivatives and Their Orthogonal Self-Assembly with P-Type Oligo(P-Phenylene Vinylene)S. *Journal of the American Chemical Society* 2004, 126, 10021-10027.


We use a combination of computational and experimental techniques to study the self-assembly and gelation of amino-acid functionalised water-soluble perylene bisimides.