SUPPORTING INFORMATION

Structural Characterization of Modular Supramolecular Architectures in Solution

David M. Tiede[*, 1], *Ruitian Zhang*[1], *Lin X. Chen*[1], *Lianhe Yu*[2], *and Jonathan S. Lindsey*[2]

[1]Chemistry Division, Argonne National Laboratory, Argonne, IL 60439 and

[2]Department of Chemistry, North Carolina State University, Raleigh, NC 27695

*Corresponding author: Tel. 630-252-3539, fax: 630-252-9289

Email: tiede@anl.gov

Experimental Section

Synthesis of the cyclic porphyrin assemblies. The shape-persistent cyclic hexameric architecture (*cyclo*-Zn_3Fb_3U-*p/m*), comprised of three free base (Fb) porphyrins and three zinc porphyrins linked at the *meso*-positions via diphenylethyne units, was prepared as described previously.[1,2] The synthesis involved the templated-assisted, Pd-mediated coupling of a *p/p*-substituted diethynyl Zn porphyrin and a *m/m*-substituted diiodo Fb porphyrin, forming *p/m*-substituted diphenylethyne linkages.[1,2] The tripyridyl template (1,3,5-tris{4-[2-(4-pyridyl)ethynyl]phenyl}benzene) was removed from the completed cyclic assembly upon chromatography and added back as a guest molecule from toluene solution in the x-ray scattering experiments.

X-ray scattering measurements. X-ray scattering measurements were carried out using the undulator beam line 12-ID at the Advanced Photon Source (APS), Argonne National Laboratory. The X-ray scattering instrument utilized a double-crystal Si(111) monochromator and a two-dimensional mosaic CCD detector for detection[3]. The X-ray wavelength was set at λ = 1.0 Å and the sample to detector distances were adjusted to achieve scattering measured across the range of momentum transfer 0.02 $Å^{-1}$ < q < 0.8 $Å^{-1}$, where $q = (4\pi / \lambda) \sin\theta$, and λ is the X-ray wavelength and 2θ is the scattering angle. Individual 2-D scattering pattern images were measured for flowed macrocycle solution and solvent background samples (5 x 10^{-5} l/min flow rate) with image exposure times between 2 to 10 seconds [4]. A data acquisition cycle included recording of 10 to 30 solution and solvent images.

Measurement of molecular high-angle scattering requires accurate subtraction of solvent and instrument background scattering. 2-D scattering pattern images were radially averaged about the transmitted beam position, and the radial distances were converted to the scattering vector q, calibrated from a silver behenate standard [5]. Figure S1A shows examples of raw experimental scattering for a 1

mM solution of macrocycle **1** and the toluene solvent background. The scattered intensities are plotted as detector counts as a function of scattering angle, q. The illustrated traces are the average of ten ten-second exposures. Standard deviations for the ten exposures are also plotted for the solution and solvent

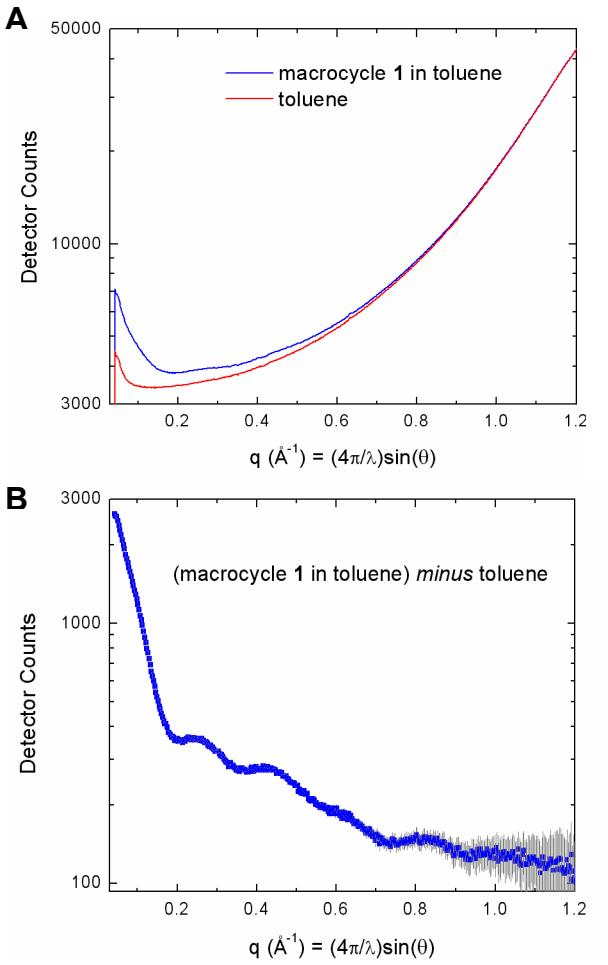

Figure S1. Experimental scattering patterns for macrocycle **1**. Part A shows radially averaged scattering patterns measured for the solution (1 mM macrocycle **1** in toluene) and the solvent (toluene). Part B shows the scattering pattern measured for macrocycle **1** by subtraction the solution and solvent scattering patterns.

curves in figure S1, but they can not be seen in the broad gain scale of figure S1A. In the scattering region below $q = 0.5$ Å^{-1} the extra scattered intensity due to the presence of macrocycle **1** in the solvent is apparent. In the region above $q = 0.5$ Å^{-1} the scattering for the solution sample is seen to be dominated by large background scattering of the solvent. The reproducibility in this set of the scattering

measurements is indicated by the small standard deviations between the ten successive measurements that correspond to measurement uncertainties of less than 0.15% compared to the scattered intensity incident at each q position. This precision is necessary for detecting molecular scattering on top of the solvent scattering background. The scattering pattern obtained for macrocycle **1** by subtraction of the solvent background from the solution sample is shown in figure S1B. At high-angle scattered intensities approach the level of 100 counts on top of the solution background of approximately 40,000 counts. While solution and solvent scattering patterns were found to be highly reproducible, the largest source of error in these scattering measurements arose from variations in the instrument background scattering that reflect the stability of the synchrotron X-ray beam and parasitic scattering from each of the optical components. Alternate sets of solvent and solution scattering patterns were collected and the statistical uncertainties within each set were evaluated. If beam line instabilities caused the measurement uncertainties to exceed 0.15%, these data were rejected, and the measurement sequences were repeated.

X-ray scattering analyses. Coordinate-based simulation of X-ray scattering curves from model structures were performed using computational approaches described previously.[6,7] Briefly, molecular scattering is calculated as the Fourier transform of the average positions of N atoms of the molecular assembly:

$$I(\boldsymbol{q}) = \sum_{i}^{N}\sum_{j}^{N} A_i A_j e^{i\boldsymbol{q}\cdot\boldsymbol{r}_{i,j}} , \tag{1}$$

where A_i is the atomic scattering amplitude for the ith atom. In solution the molecular scattering is orientationally averaged, yielding the following form:

$$I(q) = \left\langle I(\boldsymbol{q}) \right\rangle_{\Omega} = \sum_{i}^{N}\sum_{j}^{N} A_i A_j \frac{\sin qr_{i,j}}{qr_{i,j}}. \tag{2}$$

The atomic scattering amplitudes are of the form:

$$A_i = f_i(q)e^{-B_i q^2/16\pi^2} - g_i(q) \ , \tag{3}$$

where $f_i(q)$ are the atomic X-ray scattering form factors,[8] $g_i(q)$ are the atom solvent excluded volumes,[6] and B_i are the atomic thermal factors.[9] Note that B_i are assumed to be orientation independent. The structure of the macromolecule-solvent interface is a crucial component of solution scattering that is absent from gas measurements. The interface structure is approximated here by $g_i(q)$ that describes the packing of solvent next to the macromolecule through "dummy atoms" consisting of Gaussian spheres of zero electron density centered on the macromolecular atomic coordinates and having volumes determined by the atomic solvent excluded volumes.[6]. The atomic solvent excluded volumes can be expected to vary with the physical properties of the solvent molecule steric constraints determined by the positions of the atoms in the supramolecular architecture. The atomic solvent excluded volumes used here were adapted from the library atomic volumes measured in proteins.[10]

An analytical approach was developed to account for rigid-body motions in supramolecular architectures by dividing the molecular structure into composite groups, and treating group motion as a perturbation that affects the amplitude of the scattering cross-product between atom pairs in different groups. The motion of each atom during a rigid-body movement was characterized by $\boldsymbol{u}_i$, the displacement of atoms from their average positions. Each $\boldsymbol{u}_i$ is assumed to have a harmonic (Gaussian) distribution. Then, the configurationally averaged molecular scattering was calculated as following:

$$I(\boldsymbol{q}) = \sum_i^N \sum_j^N A_i A_j e^{i\boldsymbol{q}\cdot\boldsymbol{r}_{i,j}} e^{-\left\langle (\boldsymbol{q}\cdot\boldsymbol{u}_{i,j})^2 \right\rangle / 2} \tag{4}$$

where the amplitude $\boldsymbol{u}_{i,j} = (\boldsymbol{u}_i - \boldsymbol{u}_j)$ reflects the spatial coherence of the atom pair during the rigid-body movement. Equation 4 is the most accurate for the cases when $|\boldsymbol{u}_{i,j}| << r_{i,j}$. Equation 4 can be extended to include the orientational average to yield an expression for isotropic solution scattering:

$$I(q) = \left\langle I(\boldsymbol{q}) \right\rangle_{\Omega} = \sum_{i}^{N} \sum_{j}^{N} A_i A_j \left\langle e^{i\boldsymbol{q}\cdot\boldsymbol{r}_{i,j}} e^{-\left\langle (\boldsymbol{q}\cdot\boldsymbol{u}_{i,j})^2 \right\rangle / 2} \right\rangle_{\Omega} \quad . \tag{5}$$

This expression allows the effects of rigid body motion on scattering to be calculated by mapping group motion onto configurationally averaged inter-atomic displacements between groups, $\boldsymbol{u}_{i,j}$. Note that for atom pairs in the same rigid-body group, $\boldsymbol{u}_{i,j} = 0$. In the limit of a strictly rigid supramolecular structure with no dispersion in atomic group positions, *i.e.*, for all atom pairs $\boldsymbol{u}_{i,j} = 0$, equation 5 reduces to equation 2 and the calculated scattering corresponds to that of the complete array. In the limit of complete disorder of atomic groups, $\boldsymbol{u}_{i,j} \rightarrow \infty$ for all atomic pairs between groups but remains $\boldsymbol{u}_{i,j} = 0$ within the group, and the scattering calculated from equation 5 reduces to that of the sum of the individual atomic groups. In the region of group disorder between these two limits, equation 5 treats rigid-body group motion in a manner analogous to that of the atomic thermal factors.

Equation 5 can be simplified for special cases: For molecular architectures with planar groups experiencing in-plane disorders, such as porphyrin group translational disorder along the radial supramolecular direction, the $\boldsymbol{u}_{i,j}$ can be decomposed into parallel and perpendicular components with respect to $\boldsymbol{r}_{i,\mathrm{j}}$, and equation 5 is well approximated by the following:

$$I(q) \approx \sum_{i}^{N} \sum_{j}^{N} A_i A_j \int_{0}^{1} \cos(q r_{i,j} x) e^{-q^2 \{ <\boldsymbol{r}_{\mathrm{ij}} \bullet \boldsymbol{u}_{\mathrm{ij}} / r_{\mathrm{ij}}>^2 x^2 + <u_{ij}^2 - (\boldsymbol{r}_{\mathrm{ij}} \bullet \boldsymbol{u}_{\mathrm{ij}} / r_{\mathrm{ij}})^2 > (1-x^2)/2 \} / 2} dx \quad . \tag{6}$$

In the case of fully isotropic rigid body motion the inter-atomic displacements make a constant projection on $\boldsymbol{q}$ and equation 5 reduces to:

$$I(q) = \sum_{i}^{N} \sum_{j}^{N} A_i A_j \frac{\sin q r_{i,j}}{q r_{i,j}} . e^{-q^2 \left\langle u_{i,j}^2 \right\rangle / 2} \quad . \tag{7}$$

Equation 7 illustrates the analogy between isotropic rigid body motions and the Debye-Waller treatment of atomic thermal motions. This analogy can be extended by defining a term describing the atom-pair spatial coherence:

$$G_{i,j} = 8\pi^2 \left\langle u_{i,j}^2 \right\rangle \tag{8}$$

which allows equation 7 to be written:

$$I(q) = \sum_i^N \sum_j^N A_i A_j \frac{\sin qr_{i,j}}{qr_{i,j}} . e^{-G_{i,j} q^2 / 16\pi^2} \quad . \tag{9}$$

Rigid body rotations of the porphyrin groups within the cyclic hexamer architecture introduce analytically complex atom-pair disorder. At one limit, the atomic distribution can be considered for the case in which the angular orientation of each porphyrin group is distributed equally around the *meso*-porphyrin position-aryl linker axes. There is no relative movement (and no disorder) between atoms within a single porphyrin group, but the rotation creates a distribution of inter-group atom-pair correlations. With the assumption of full rotation, the equilibrium positions, $\boldsymbol{r}_i$(p), of the atoms fall onto their rotation centers, which are on their rotation axes, and a rotationally averaged displacement vector $\boldsymbol{u}_i$ is determined by the vectors, $\boldsymbol{r}_{i\perp} = \boldsymbol{r}_i - \boldsymbol{r}_i(p)$. Plugging these relations into equation 5, and after several steps of mathematical calculations, we obtain the approximate formula for scattering intensity as:

$$I(q) \approx \sum_i^N \sum_j^N A_i A_j \int_0^1 \cos(qr_{i,j}(p)x) e^{-q^2 (r_{i\perp}^2 + r_{j\perp}^2)\{\cos^2(\omega_{ij})x^2 + (1+\sin^2(\omega_{ij}))(1-x^2)/2\}/4} dx, \tag{10}$$

where $r_{i,j}(p) = |\boldsymbol{r}_i(p) - \boldsymbol{r}_j(p)|$, and ω_{ij} is the angle between the in-plane normal direction to the *meso*-porphyrin-linker axes and the vector $\boldsymbol{r}_{i,j}(p)$.

Pair Distance Distribution Function. The atomic pair distribution function (PDF), *P(r)*, was calculated directly from atomic coordinates for the model structures. For simplicity, the electron distribution on each atom was assumed to be uniformly distributed within a spherical volume V_i of radii R_i determined

by the atomic solvent excluded volume, and the electron density weighted pair correlation was calculated from the excess number of electrons with respect to the solvent electron density. *P(r)* is related to scattering intensity *I(q)* by the following:

$$P(r) = \frac{1}{2\pi^2} \int_0^\infty dq(qr)\sin(qr)I(q) \quad . \tag{11}$$

In an alternative method, the PDF for model structures was determined from the calculated scattering patterns using indirect Fourier transform methods.[11-13] This procedure allowed both model and experimental PDF to be determined in identical fashion using the program GNOM.[12,13]

Molecular Modeling. Energy minimized structures for **1** and **1+2**, were generated from energy minimized structures of individual building blocks, consisting of zinc or free base porphyrins connected with diphenylethyne linkers, calculated using the MM+ force field in the Hyperchem programs (Hypercube). Both steepest descent and conjugate gradient methods of energy minimization were employed. These energy minimized building block structures were then connected to form the macrocycle structures **1** and **1+2**, which consequently are energy minimized using the same program. Because of the size of the assemblies **1** and **1+2**, many isoenergetic structures were found that varied depending on the starting structures. The dihedral angle defining the relative orientation between the porphyrin macrocycle and the assembled macrocycle planes was found to be a critical parameter determining the conformation of the minimized structure. These multiple nearly isoenergetic conformations are presumably examples of the dynamic ensemble in solution. Structural dynamics of these macrocycles were investigated using molecular dynamics (MD) calculations carried out with Hyperchem. The calculations started from the energy minimized structures that were assigned to be the unthermalized, 0 K structures. The heating time from the assemblies to reach 290K was 100 ps with 0.5 fs/step. The MD simulations continue thereafter to 1 ns with the same step size.

References

(1) Yu, L.; Lindsey, J. S. *J. Org. Chem.* **2001**, *66*, 7402-7419.

(2) Li, J.; Ambroise, A.; Yang, S. I.; Diers, J. R.; Seth, J.; Wack, C. R.; Bocian, D. F.; Holten, D.; Lindsey, J. S. *J. Am. Chem. Soc.* **1999**, *121*, 8927-8940.

(3) Seifert, S.; Winans, R. E.; Tiede, D. M.; Thiyagarajan, P. *J. Appl. Cryst.* **2000**, *33*, 782-784.

(4) Tiede, D. M.; Zhang, R.; Seifert, S. *Biochemistry* **2002**, *41*, 6605-6614.

(5) Huang, T. C.; Toraya, H.; Blanton, T. N.; Wu, Y. *J. Appl. Cryst.* **1993**, *26*, 180-184.

(6) Svergun, D.; Barberato, C.; Koch, M. H. J. *J. Appl. Cryst.* **1995**, *28*, 768-773.

(7) Zhang, R.; Seifert, S.; Thiyagarajan, P.; Tiede, D. M. *J. Appl. Cryst.* **2000**, *33*, 565-568.

(8) Ibers, J. A.; Hamilton, W. C., Eds. *International Tables for X-ray Crystallography*, 1973; Vol. IV.

(9) Warren, B. E. *X-Ray Diffraction*; Dover Publications, Inc: New York, 1990.

(10) Pontius, J.; Richelle, J.; Wodak, J. *J. Mol. Biol.* **1996**, *264*, 121-136.

(11) Glatter, O. In *Neutron, X-ray and Light Scattering*; Lindner, P., Zemb, T., Eds.; Elsevier Science Publishers B.V.: Amsterdam, 1991, pp 33-82.

(12) Svergun, D. I.; Semenyuk, A. V.; Feigin, L. A. *Acta. Crystallogr. Section A* **1988**, *44*, 244-250.

(13) Svergun, D. I. *J. Appl. Cryst.* **1992**, *25*, 495-503.