# The Thermodynamics of Self-assembly

**Xianglan Chi, Andrea J. Guerin, Richard A. Haycock, Christopher A. Hunter* and Luke D. Sarson**

*Krebs Institute for Biomolecular Science, Department of Chemistry, University of Sheffield, Sheffield, UK S3 7HF*

The concentration range over which a self-assembled macrocyclic structure is stable is given by the critical self-assembly concentration (lower limit) and the effective molarity for cyclisation (upper limit); the relationships between these parameters, the size of the assembly and the properties of the component interactions are used to describe the efficiency of self-assembly processes.

Self-assembly is an attractive and viable alternative to covalent synthesis for the construction of large multicomponent architectures.[1–6] However, abandoning the covalent bond as a construction tool is not without drawbacks. Structures held together by weak non-covalent interactions can not be isolated and characterised in the same way as single compounds. Self-assembled complexes are in dynamic equilibrium with their monomeric components and alternative aggregates or complexes. The structure of the assembly is critically dependent on solvent, concentration and other external influences such as temperature. Thus we must develop both different construction and characterisation strategies if we are to realise the potential of non-covalent synthesis.[2] In particular, the issue which we address here is 'how can we prevent the formation of open polymeric structures and generate discrete closed structures which are stable over a wide concentration range?'

We have studied the self-assembly of cyclic structures using zinc porphyrin–pyridine coordination interactions[7,8] This system is attractive because we can determine the stability of the macrocyclic assemblies and characterise the constituent interactions by using simple reference compounds. To date, we have reported the self-assembly of a macrocyclic dimer $(Zn\mathbf{1})_2$, trimer $(Zn\mathbf{2})_3$ and tetramer $(Zn\mathbf{3})_4$ (Fig. 1). Although the individual interactions used are almost identical, the behaviour of these three systems is quite different, and there is an interesting trend along the series. Table 1 shows the relationship between the size of the oligomer and the concentration at which self-assembly takes place (the critical self-assembly concentration as defined below): the larger the oligomer, the less stable the cyclic structure. Fig. 2 shows a computer simulation of how the mole fractions of the self-assembled structures vary with concentration.

The results can be rationalised using Scheme 1. The effective molarity (EM) is a useful parameter, because it is a measure of the concentration at which open polymeric structures start to compete with the closed cyclic structure.[9] In other words, EM defines the upper limit of the concentration range over which a self-assembled structure is stable. In our analysis, we assume that $K_{open}$, and hence EM, can be estimated by measuring $K_{ref}$ for the structurally related complexes formed between porphyrins Zn**4a** and Zn**4b** and ligands, **5**, **6** and **7** (Table 1). Thus for a cyclic assembly containing $n$ monomer units,

$$EM = \frac{K_{closed}}{K_{open}{}^{n}} = \frac{K_{closed}}{K_{ref}{}^{n}} \quad (1)$$

where $K_{closed}$ is the monomer–cyclic $n$-mer association constant (Scheme 1). $K_{open}$ is the monomer–open dimer association constant which we assume to be identical to the open dimer–open trimer association constant and so on for the formation of longer open oligomers (Scheme 1) and $K_{ref}$ is the corresponding reference porphyrin–reference ligand association constant.

In the absence of entropic solvent effects, the upper limit of the value of EM is given by[10]

$$EM_{max} = \exp(-\Delta S_{ref}/R) \quad (2)$$

Thus the entropy change for the formation of any given interaction may be used to assess its utility in a self-assembly process. For the formation of zinc porphyrin–pyridine complexes in toluene,[11] $\Delta S$ = *ca.* 50 $JK^{-1}$ $mol^{-1}$ which would give an upper limit on the value of EM of 400 mol $dm^{-3}$. Although we used a different solvent, dichloromethane, the measured value of EM for the cyclic trimer $(Zn\mathbf{2})_3$ is 100 mol $dm^{-3}$ which is similar in magnitude and is likely to be close to the upper limit for this system. The EM for the dimer $(Zn\mathbf{1})_2$ is somewhat lower which may be due to conformational ring strain since the geometry of the monomer is not quite optimal. The EM for the tetramer $(Zn\mathbf{3})_4$ is significantly lower which reflects an increase

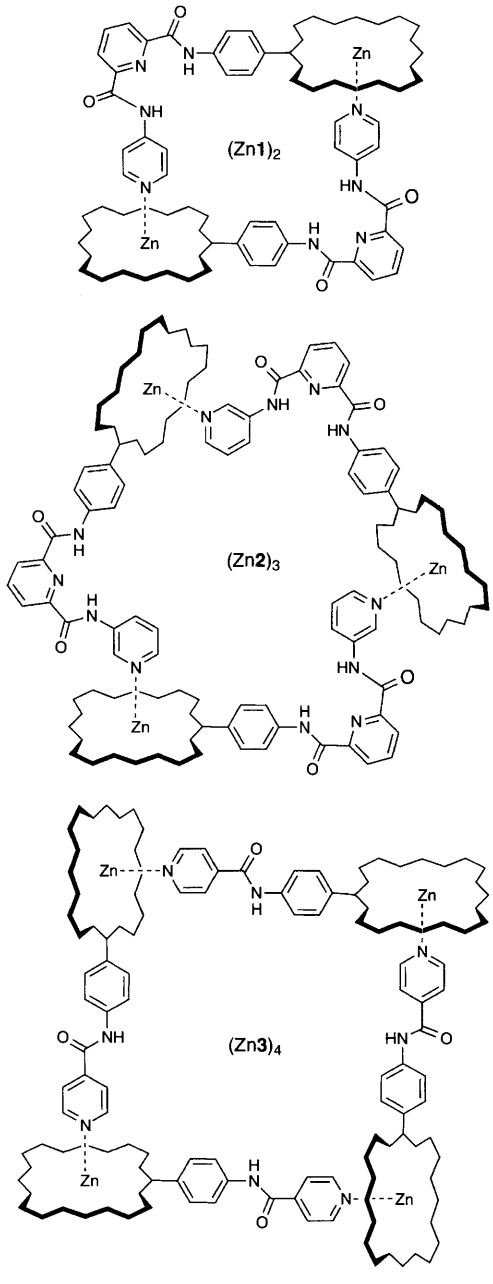



**Fig. 1** Self-assembled macrocyclic porphyrin oligomers. For clarity, the solubilising groups R at the porphyrin *meso* positions and the porphyrin ring bonds are not shown.

in the number of internal degrees of freedom which are frozen out on cyclisation of the larger macrocycle (Table 1). This would clearly become a serious problem if we were to try to construct much larger assemblies using our current approach. Thus for very large multicomponent structures, a strategy for reducing the conformational freedom or preorganising the open oligomeric assemblies must be found: solutions are to use steric constraints or multipoint binding.[2,3]

Table 1 shows that the lower limit of the concentration range over which a self-assembled complex is stable is dramatically dependent on the size of the oligomer ($n$). We term this concentration the critical self-assembly concentration (csac) and define it as the concentration at which the complex is half assembled, *i.e.* the mole fraction of monomer units present in the form of the fully-assembled complex is 0.5. The relationship

**Table 1** Self assembly properties of macrocyclic porphyrin oligomers in dichloromethane at room temperature

| Monomer[a] | $n$ | $K_{ref}$ dm$^3$ mol$^{-1}$ | csac mol dm$^{-3}$ | EM/mol dm$^{-3}$ | ε |
|---|---|---|---|---|---|
| Zn**1a** | 2 | $5.6 \times 10^3$ | $3 \times 10^{-9}$ | 6 | 9.3 |
| Zn**2a** | 3 | $3.8 \times 10^3$ | $2 \times 10^{-7}$ | 100 | 8.7 |
| Zn**3a** | 4 | $1.9 \times 10^3$ | $3 \times 10^{-5}$ | 0.6 | 4.5 |
| Zn**3b** | 4 | $1.4 \times 10^3$ | $4 \times 10^{-5}$ | 0.9 | 4.4 |

[a] **a** and **b** refer to the solubilising groups $R^1$ and $R^2$ respectively.

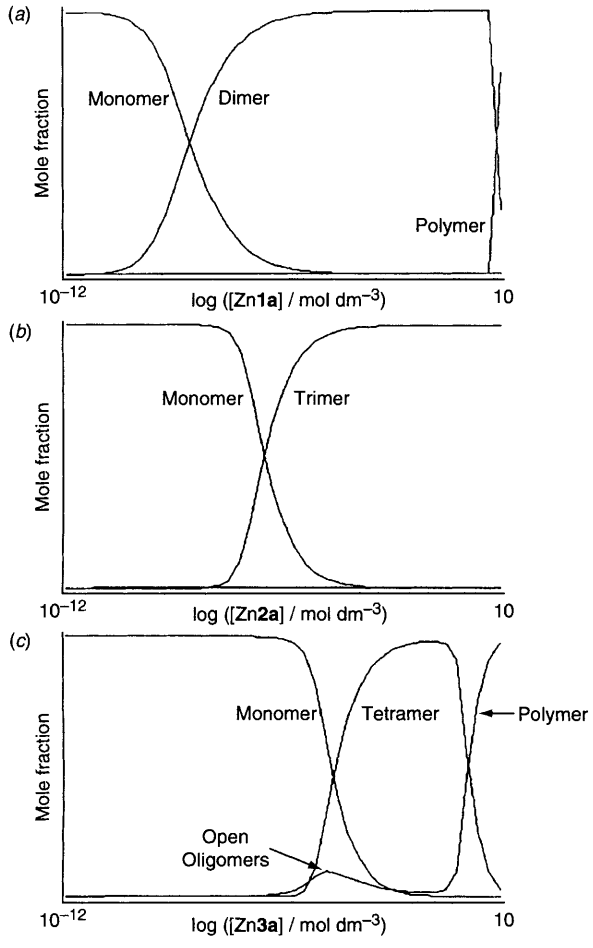



**Fig. 2** Mole fraction of self-assembled porphyrin macrocycles as a function of concentration (log scale) for (*a*) Zn**1a** (*b*) Zn**2a** (*c*) Zn**3a**. The populations of the various species were determined by solving the simultaneous equations for the multiple equilibria in Scheme 1 using the experimental association constants in Table 1 (assuming that $K_{open} = K_{ref}$).

of the csac to $n$, $K_{ref}$ and EM can be derived as follows for a simple two state equilibrium

$$\text{csac} = [\text{Monomer}], \text{ when} \frac{n[\text{Cycle}]}{[\text{Monomer}]} = 1$$

$$[\text{Cycle}] = K_{closed}\,[\text{Monomer}]^n \quad (3)$$

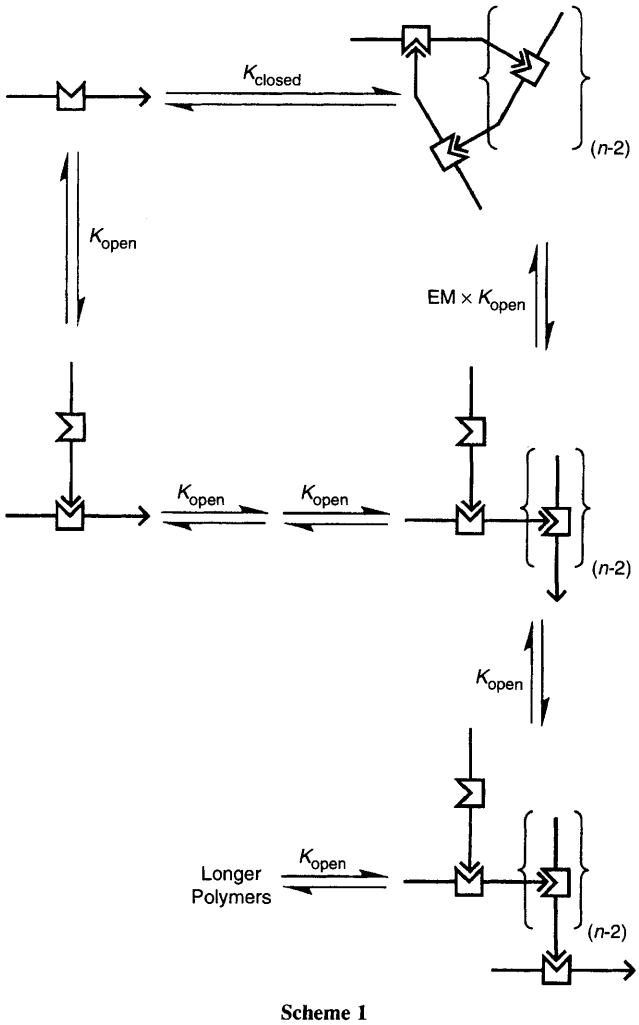
Thus

$$\text{csac} = \frac{1}{n^{1/(n-1)}\,K_{closed}{}^{1/(n-1)}} \quad (4)$$

substituting from eqn. (1) for $K_{closed}$ gives

$$\text{csac} = \frac{1}{n^{1/(n-1)}\,\text{EM}^{1/(n-1)}\,K_{ref}{}^{n/(n-1)}} \quad (5)$$

It is clear from this equation why csac rises rapidly as $n$ increases in Table 1. In the limit as $n$ tends to infinity, $n^{1/(n-1)} \rightarrow 1$, $\text{EM}^{1/(n-1)} \rightarrow 1$ and $K_{ref}{}^{n/(n-1)} \rightarrow K_{ref}$. The entropic advantage of the intramolecular cyclisation (EM) is dissipated as it is shared out among more and more components. Thus as $n$ becomes large, csac tends towards a limiting value of $1/K_{ref}$. The effective molarity affects the value of csac when $n$ is small, but has no influence on the value of csac for large $n$. Thus very large multicomponent systems will self-assemble to give stable structures at concentrations between $1/K_{ref}$ and EM: provided EM is maintained as a reasonably high value by using preorganised monomers, there is no limit on the size of the assembly. These closed assemblies will not show any enhanced stability relative to the individual interactions used to hold them




**Scheme 1**

together, but they will be stable relative to open polymeric species.

Substituting from eqn. (2) for $EM_{max}$ gives a lower limit for csac for a given type of interaction.

$$csac_{min} = \frac{1}{n^{1/(n-1)}} \exp\left[\frac{\Delta H_{ref}}{(n-1)\,RT}\right]\frac{1}{K_{ref}} \quad (6)$$

This expression is related to Whitesides $HB/(N-1)$ parameter which defines the stability of his hydrogen-bonded assemblies ($HB$ is the number of hydrogen-bonds and $N$ is the number of molecules in the assembly).[2] Eqn. (6) can be rewritten as eqn. (7):

$$csac_{min} = \frac{1}{n^{1/(n-1)}} \exp\left[\frac{n\,\Delta H_{ref}}{(n-1)\,RT} - \frac{\Delta S_{ref}}{R}\right] \quad (7)$$

In general, for an assembly of $n$ molecules which make $m$ identical interactions, the lower limit for csac will be given by

$$csac_{min} = \frac{1}{n^{1/(n-1)}} \exp\left[\frac{(m-n+1)\,\Delta H_{ref}}{(n-1)\,RT}\right]\frac{1}{K_{ref}} \quad (8)$$

or

$$csac_{min} = \frac{1}{n^{1/(n-1)}} \exp\left[\frac{m\,\Delta H_{ref}}{(n-1)\,RT} - \frac{\Delta S_{ref}}{R}\right] \quad (9)$$

In practice, the value of csac will always be higher than this due to conformational mobility and strain which will reduce EM for cyclisation from its maximum possible value in eqn. (2) (as observed in our porphyrin oligomers). This analysis implies that polymacrocyclic three-dimensional assemblies which require multiple interaction sites on the monomer units are likely to be significantly more stable than their simple macrocyclic two-dimensional counterparts which need only two interaction sites per monomer.[2,3]

A measure of the efficiency of a self-assembly process (ε) can be defined as the concentration range over which the assembly is stable. Thus

$$\varepsilon = \mathrm{Log}\left(\frac{EM}{csac}\right) \quad (10)$$

$$\varepsilon = \left[\frac{n}{n-1}\,\mathrm{Log}(EM\,K_{ref})\right] + \frac{\mathrm{Log}(n)}{n-1} \quad (11)$$

Self-assembly will only be observed if ε is significantly greater than zero. We can obtain a value for the upper limit of ε in any given system by substituting from eqn. (2) for $EM_{max}$:

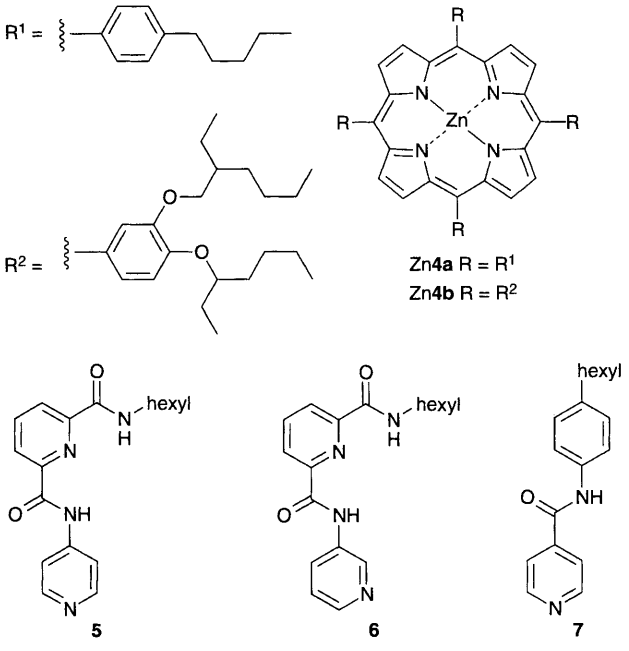
$$\varepsilon_{max} = \left(\frac{n}{n-1}\,\frac{\Delta H_{ref}}{2.303\,RT}\right) + \frac{\mathrm{Log}(n)}{n-1} \quad (12)$$

Thus while the entropy of the component interactions is important in determining the upper limit of the stability of an assembly, the enthalpy of the component interactions is important in determining the maximum concentration range over which assembly can take place. The values of ε given in Table 1 for the three porphyrin oligomers reflect the stability range of the self-assembled structures as illustrated in Fig. 2. The tetramer $(Zn\mathbf{3})_4$ is close to the lower limit of ε for which self-assembly may realistically be characterised (ε = 4), because the macrocyclic tetramer constitutes greater than 90% of the species present over a concentration range of less than 100-fold and open oligomers are always present at the level of a few percent.

These expressions provide guidelines for estimating the efficiency of a self-assembly process from the properties of the individual interactions to be used and should prove useful in the design of new self-assembling systems. In particular, they highlight the need to maintain a high effective molarity if large multicomponent assemblies are to be generated. Such large assemblies will have *apparent* stability constants which are the same as the *anticipated* stability constant for the formation of open polymeric assemblies ($K_{ref}$), but they will be stable with respect to polymerisation over a reasonable concentration range.

This research was supported by Glaxo, the CVCP and the University of Sheffield (L. D. S.), the EPRSC (R. A. H.), Pfizer Ltd and the University of Sheffield (A. J. G.) and the Lister Institute (C. A. H.).

*Received, 22nd June 1995; Com. 5/04153H*




5 6 7

## References

1 J. S. Lindsey, *New J. Chem.*, 1991, **15**, 153.
2 G. M. Whitesides, E. E. Simanek, J. P. Mathias, C. T. Seto, D. N. Chin, M. Mammen and D. M. Gordon, *Acc. Chem. Res.*, 1995, **28**, 37.
3 P. Baxter, J. M. Lehn, A. DeCian and J. Fischer, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 69; P. Baxter, J. M. Lehn, J. Fischer and M. T. Youinou, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 2284.
4 M. Fujita, J. Yazaki and K. Ogura, *J. Am. Chem. Soc.*, 1990, **112**, 5645; M. Fujita, J. Yazaki and K. Ogura, *J. Am. Chem. Soc.*, 1995, **117**, 1649.
5 S. C. Zimmerman and B. F. Duerr, *J. Org. Chem.*, 1992, **57**, 2215.
6 J. Yang, J. L. Marendaz, S. J. Geib and A. D. Hamilton, *Tetrahedron Lett.*, 1994, **35**, 3665.
7 C. A. Hunter and L. D. Sarson, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 2313.
8 X. Chi, A. J. Guerin, R. A. Haycock, C. A. Hunter and L. D. Sarson, *J. Chem. Soc., Chem. Commun.*, 1995, (preceding paper).
9 H. L. Anderson, *Inorg. Chem.*, 1994, **33**, 972.
10 H. L. Anderson, S. Anderson and J. K. M. Sanders, *J. Chem. Soc., Perkin 1*, in the press.
11 F. A. Walker and M. Benson, *J. Am. Chem. Soc.*, 1980, **102**, 5530; H. Imai and E. Kyuno, *Inorg. Chem.*, 1990, **29**, 2416.