Physical basis of self-assembly. Part 2.† A theoretical and experimental study of the self-assembly of a zinc *meso*-pyridyl porphyrin

Gianfranco Ercolani,* Marcella Ioele and Donato Monti

Dipartimento di Scienze e Tecnologie Chimiche, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 00133 Roma, Italy. E-mail: ercolani@uniroma2.it; Fax: +3906 72594328

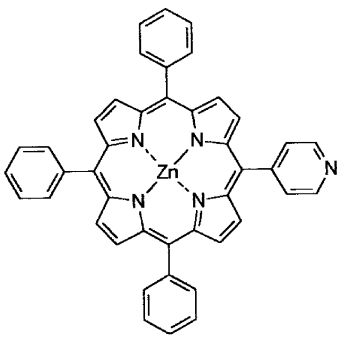
Received* (*in London, UK*) *29th January 2001, Accepted 29th March 2001
First published as an Advance Article on the web 16th May 2001

A previously reported theoretical treatment for self-assembly macrocyclisations occurring under thermodynamic control has been tested experimentally. The fundamental quantities on which the treatment is based are the effective molarity (EM) of the self-assembling cyclic *n*-mer and the equilibrium constant for the intermolecular model reaction between monofunctional reactants (K_{inter}). Provided that estimates of EM and K_{inter} are available, this treatment can be used to predict not only whether the self-assembly process is more or less favoured, but also the distribution of all the species present in solution. Since K_{inter} values are approximately known from the literature, we have proposed a method, based on molecular modelling techniques, to estimate the EM. The method has been applied to the self-assembly of $Zn(PyP_3P)$, where PyP_3P is 5-(4-pyridyl)-10,15,20-triphenylporphyrinato dianion. An EM greater than 0.1 mol L^{-1} has been estimated for its cyclotetramerisation by PM3 calculations, suggesting that self-assembly should be favoured in solvents like toluene and chloroform. Self-assembly of $Zn(PyP_3P)$ has been studied in these solvents by UV/visible spectroscopy. The data are consistent with the formation of the cyclotetramer, and at variance with the model of linear polymerisation. The experimental values of the EM were little affected by the nature of the solvent (EM values were 20 mol L^{-1} in toluene and 15 mol L^{-1} in chloroform), indicating that the solvent affects the process of self-assembly mainly through the value of K_{inter}.

Introduction

Self-assembly consists of the spontaneous generation of a well-defined, discrete supramolecular architecture from a given set of components under thermodynamic equilibration.[1] Application of this phenomenon to macrocyclisation is attractive, since cyclisation of a large molecule by the usual synthetic methods, employing kinetically controlled reactions, is frequently a tedious and low-yield process. Numerous examples of self-assembling systems can be found in the literature, however the physicochemical basis of self-assembly has not, with some exceptions,[2–5] received adequate attention.

We have recently reported a theoretical treatment for self-assembly macrocyclisations occurring under thermodynamic control.[5] The fundamental quantities on which the treatment is based are the effective molarity[6–8] (EM) of the self-assembling cyclic *n*-mer and the equilibrium constant for the intermolecular model reaction between monofunctional reactants (K_{inter}). Knowledge of these quantities would not only allow prediction of whether the self-assembly process is more or less favoured, but also the distribution of all the species present in solution. It appears, therefore, that it is crucial to have methods that allow an estimate of both K_{inter} and EM. In fact, an estimate of K_{inter} can generally be made on the basis of literature data, whereas an estimate of the EM is more difficult. The present paper has the twofold aim of both proposing a method, based on molecular modelling techniques, to obtain a rough estimate of the EM and testing our previously reported treatment experimentally. In search of a suitable system for such an investigation, we were attracted by $Zn(PyP_3P)$ [PyP_3P = 5-(4-pyridyl)-10,15,20-triphenylporphyrinato dianion] as a building block because its structure is relatively simple and, most stimulating for us, its self-assembly behaviour has given rise to some controversy which requires further investigation. In fact, Fleischer and Shachter reported a study on the aggregation of $Zn(PyP_3P)$ both in solution and in the solid state:[9] X-ray crystallography of $Zn(PyP_3P)$ showed that the pyridyl nitrogen is bound to the metal centre of an adjacent porphyrin to form a linear polymeric aggregate. Since a spectrophotometric study of $Zn(PyP_3P)$ in $CHCl_3$ solution showed a good fit to a model involving reversible linear polymerisation, it was concluded that the linear polymeric structure is present not only in the solid state, but also in solution. However, Hunter *et al.* pointed out that the stability constant for the formation of the polymeric assembly in solution obtained by Fleischer and Shachter is unusually large for pyridine coordination of a zinc porphyrin and suggested, by analogy with the behaviour of their macrocyclic porphyrin



$Zn(PyP_3P)$

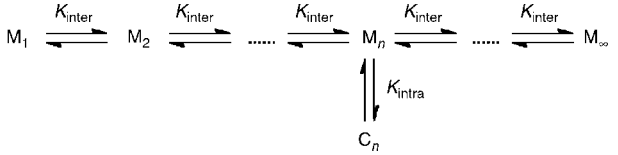
† For part 1, see: ref. 5.

oligomers, that $Zn(PyP_3P)$ probably forms a cyclic tetrameric structure in solution, *c*-$[Zn(PyP_3P)]_4$.[10] Imamura *et al.* reported a variable temperature ^1H NMR study of $Zn(PyP_3P)$ in $CDCl_3$ and found that at $-40\,^\circ C$ the spectrum is the same as that of *c*-$[Ru(PyP_3P)(CO)]_4$, with sharp pyridyl and β-pyrrole signals that broaden at $25\,^\circ C$.[11] They concluded that at low temperature the cyclic tetramer is the major component and that a fast exchange, relative to the NMR time scale, occurs at $25\,^\circ C$ between cyclic and linear structures. However, it is still not clear whether the cyclic tetramer is the prevalent species at room temperature.

Results and discussion

Theoretical considerations

Before illustrating our results, it is useful to briefly recall the principal features of the previously reported theoretical treatment for self-assembly macrocyclisations.[5] Consider a bifunctional monomer A–B (M_1) bearing two different functional groups, each capable of reacting with the other only in a reversible addition reaction. After equilibration, a system initially composed of monomer units M_1 contains, in principle, an infinite number of cyclic oligomers C_i as well as an infinite number of linear oligomers M_i, i being the degree of polymerisation. If the monomer, however, has a rigid structure predisposed[12] in such a way that formation of all the cyclic oligomers except one (the cyclic n-mer) is prevented by strain, the equilibration scheme can be simplified as shown in Scheme 1.



Scheme 1 Equilibria for the self-assembly of the cyclic oligomer C_n accompanied by linear polymerisation.

The mass balance equation relative to Scheme 1 is expressed by eqn. (1).[5]

$$[M_1]_0 = nEMx^n + \frac{x}{K_{inter}(1-x)^2} \quad (1)$$

where $[M_1]_0$ is the initial monomer concentration, K_{inter} is the equilibrium constant for the intermolecular model reaction between the functional groups A and B, assumed to be identical for all the associations between the linear oligomers, EM is the equilibrium effective molarity (defined as K_{intra}/K_{inter}, where K_{intra} refers to the equilibrium of cyclisation of the acyclic n-mer M_n to yield the cyclic n-mer C_n) and x, confined to the range $0 \leqslant x < 1$, is the fraction of reacted end groups in the acyclic part of the polymer. The first term in the right-hand side of eqn. (1) is the weighted concentration of C_n, whereas the second term is the weighted concentration of the acyclic fraction. Of course, in the absence of cyclic oligomers eqn. (1) reduces to eqn. (2).

$$[M_1]_0 = \frac{x}{K_{inter}(1-x)^2} \quad (2)$$

Note that since $x = K_{inter}[M_1]$, eqn. (2) is equivalent to the mass balance equation reported by Fleischer and Shachter [eqn. (10) of their paper].[9,13]

Although eqn. (1) cannot be solved analytically, knowing $[M_1]_0$, K_{inter} and EM, it is possible to obtain a numerical solution for x by the Newton–Raphson method.[14] In turn, knowledge of x allows the concentrations of all the species present in solution to be found, since eqn. (3) and eqn. (4) hold.

$$[M_i] = \frac{x^i}{K_{inter}} \quad (3)$$

$$[C_n] = EMx^n \quad (4)$$

One of the principal results of our treatment is that in order for self-assembly to be virtually complete $[C_n(\%)_{max} > 99\%]$ in a certain range of initial monomer concentration the following condition must be verified.

$$EMK_{inter} \geqslant 185n \quad (5)$$

If, in addition, $n \leqslant 4$, the acyclic fraction is essentially composed of the monomer only; in this case, Scheme 1 reduces to eqn. (6), and eqn. (1) to eqn. (7)

$$nM_1 \overset{EMK_{inter}^n}{\rightleftharpoons} C_n \quad (6)$$

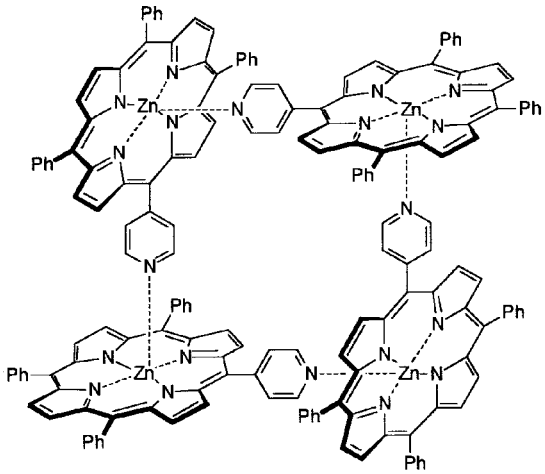
$$[M_1]_0 = nEMx^n + \frac{x}{K_{inter}} \quad (7)$$

An estimate of both EM and K_{inter} would allow a prediction using eqn. (5) of whether the process of self-assembly is more or less favoured. In fact, K_{inter} can generally be estimated from literature data. For example, the interaction between zinc tetraphenylporphyrin (ZnTPP) and pyridine has been studied by various authors in several solvents. A list of the available association constants is reported in Table 1. These values, which appear to be significantly affected by the nature of solvent, provide an estimate of the equilibrium constant K_{inter} for the process under consideration. An estimate of the EM is more difficult, depending on the structural features of both M_n and C_n. However, as a rule of thumb, a relatively high EM, say >0.1 mol L^{-1}, is a result of the absence of significant strain in the cyclic oligomer C_n (the main factor affecting the enthalpic component of the EM) and of a small number of rotatable bonds in the linear oligomer M_n (the main factor affecting the entropic component of the EM).[7] To ascertain whether significant strain energy is present in the self-assembled cyclic oligomer C_n, the enthalpy change ($\Delta_r H$) for the isodesmic equilibrium[23] reported in eqn. (8) has to be evaluated. The presence of strain would be signalled by a significantly positive $\Delta_r H$ value.

$$M_n + M_2 \overset{EM}{\rightleftharpoons} C_n + 2M_1 \quad (8)$$

In the case under consideration, eqn. (8) becomes:

$$[Zn(PyP_3P)]_4 + [Zn(PyP_3P)]_2 \rightleftharpoons c\text{-}[Zn(PyP_3P)]_4 + 2Zn(PyP_3P) \quad (9)$$



c-$[Zn(PyP_3P)]_4$

A way of estimating the reaction enthalpy is by using quantum mechanical calculations. To this end, we have

resorted to semi-empirical methods, since the structural complexity of the species involved prohibits the use of accurate *ab initio* calculations. The PM3 semi-empirical molecular orbital method[24] has been used to calculate the geometry and heat of formation of all the species involved in eqn. (9) (see Experimental for details). There is evidence in the literature[25] that both PM3 and AM1 semi-empirical methods predict geometrical parameters for zinc porphyrins which are close to those observed in X-ray studies, however, since PM3 performed better than AM1 as far as heats of formation are concerned, it was the method of choice. Earlier, Stewart reported a root mean square error of 7.9 kcal mol^{-1} for absolute values of heats of formation obtained by PM3 employing 276 compounds containing H, N, O and C.[24] Although there are indications that in the case of zinc porphyrins the error could be even larger (*ca.* 25 kcal mol^{-1}),[25] owing to the isodesmic nature of the equilibrium in eqn. (9), most of systematic errors should cancel out. The calculated reaction enthalpy turned out to be $\Delta_r H = -2.2$ kcal mol^{-1}. A negative $\Delta_r H$ value should be interpreted as due either to a self-template effect stabilising the cyclic tetramer or to a strain factor present in the linear tetramer that is released upon ring closure. However, since the $\Delta_r H$ value is rather small, and since there are no evident structural features that suggest either of the two possible effects, we are more inclined to believe that the negative $\Delta_r H$ value is trivially due to the approximate nature of the PM3 method. We assume therefore, as is more likely, that the reaction in eqn. (9) occurs without any appreciable real enthalpy change and, therefore, the EM should be dominated by entropy only.

The entropic component of the EM is more difficult to evaluate. Page and Jencks were the first to recognise that the entropic advantage of an intramolecular reaction with respect to the corresponding intermolecular one is due to the significant losses of translational plus rotational entropy that occur in the latter.[26] They estimated that such an advantage can be translated into a maximum EM value of about 10^8 mol L^{-1}. The presence of rotatable bonds in the chain undergoing ring closure involves a decrease in the EM because of restriction of torsional motion in the cyclic structure. In the case of short alkane chains, this decrease has been estimated to be about a factor of 10 per rotor.[7,8,26,27] The linear oligomer $[Zn(PyP_3P)]_4$ is characterised by the presence of three rotors, namely the three N(pyridyl)–Zn bonds, however the decrease of the EM per rotor is expected to be much greater than a factor of 10 because of the larger moments of inertia of the rotating groups with respect to the simple aliphatic chain. To estimate the reduced moment of inertia of each rotor (I_r) an approximate method, recommended for asymmetric top rotations, has been used.[28] The method is based on eqn. (10)

$$I_r = \frac{I_1 I_2}{I_1 + I_2} \tag{10}$$

where I_1 and I_2 are the moments of inertia of the two rotating groups about the axis of internal rotation. The moment of inertia of each top is computed, however, not about the axis containing the twisting bond as in the case of symmetric tops, but about the axis parallel to the bond and passing through the centre of mass of the top. A computer program has been written that calculates I_r for each rotor from the atomic coordinates of the PM3 optimised structure of the linear tetramer. The following I_r values (in amu Å^2) were obtained: 1.39×10^4; 2.39×10^4; 7.92×10^3. The geometrical mean of these values gives an average I_r value per rotor of 1.38×10^4 amu Å^2. The entropy of an internal rotation depends not only on the reduced moment of inertia of the rotating groups but also on the potential energy barrier to rotation (the entropy decreases on increasing the barrier).[29] Taking the dimer $[Zn(PyP_3P)]_2$ as a model compound, we have evaluated the adiabatic potential energy curve relative to the internal rotation about the N(pyridyl)–Zn bond by PM3 calculations (see Experimental for further details). The energy barrier turned out to be very low (0.59 kcal mol^{-1}), thus the internal rotation about N(pyridyl)–Zn can be considered virtually free at room temperature. According to statistical thermodynamics, the entropic contribution of a free internal rotation is given by eqn (11).[29]

$$S = R\left(0.5 \ln \frac{8\pi^3 I_r kT}{h^2} - \ln \sigma_{int} + 0.5\right) \tag{11}$$

where T is the absolute temperature, σ_{int} is the symmetry number for internal rotation, and R, k and h are the usual fundamental constants. Considering that for the three rotors under examination $\sigma_{int} = 1$, the average entropy per rotor at 25 °C turns out to be 14 cal mol^{-1} K^{-1}. If one assumes that upon ring closure internal rotations are frozen and this entropy is completely lost,[29] the EM will decrease by a factor $\exp(S/R)$, *i.e.* by a factor of about 10^3 per rotor. The resulting EM ($\approx$0.1 mol L^{-1}), however, should be considered as a lower limit because complete, or almost complete, freezing of torsional motion, although usual in small ring compounds, is unlikely in a macrocyclic structure such as *c*-$[Zn(PyP_3P)]_4$ because of new low frequency vibrations.[29,30] In spite of this uncertainty, from the estimated values of K_{inter} and EM, it can be concluded from eqn. (5) that in a solvent like toluene a large fraction of $Zn(PyP_3P)$ should be present in the cyclic tetrameric form. As a result, we initially chose this solvent to carry out the experimental study of self-assembly. We then extended the study to chloroform to gain information about possible solvent effects.

Table 1 Equilibrium constants for the addition of pyridine to ZnTPP at 298 K

Solvent	K/mol L^{-1}	Ref.
benzene	4800	15
benzene	6030	16
benzene	3900	17
benzene	6460	18
benzene	5300	19
toluene	6030	20
chloroform	610	18
cyclohexane	25100	21
methylene chloride	6900	22

Self-assembly in toluene

Binding studies were carried out at 25.0 °C using UV/visible spectroscopy. Our first goal was the evaluation of the equilibrium constant K_{inter}, and the very intense absorptions of $Zn(PyP_3P)$ in the 410–440 nm region (Soret band) were used to this end. We recorded the spectra of $Zn(PyP_3P)$ at a concentration of about 2 μM upon addition of increasing amounts of pyridine up to [Py] $\approx$ 10 mM. Analysis of the absorbance data at λ 422 and 430 nm *vs.* pyridine concentration led to well-behaved 1 : 1 complexation curves which gave a binding constant of the expected order of magnitude ($K_{inter} = 4.1 \pm 0.2 \times 10^3$ L mol^{-1}).

Dilution studies were carried out in the concentration range $[Zn(PyP_3P)] \approx 0.04$–7 μM. Plots of the apparent molar absorptivity in the 405–445 nm region, reported in Fig. 1, show a well-defined isosbestic point. There is evidence in the literature that cyclotetrameric complexes, such as *c*-[Ru$(PyP_3P)(CO)]_4$, show broadening of the Soret bands because of the excitonic interactions between the porphyrin subunits which are aligned parallel.[11] This implies that the molar absorptivity of a coordinated metalloporphyrin subunit in an ordered cyclic tetrameric structure is different from that in a random linear polymer. The isosbestic point in Fig. 1, is therefore not compatible with a model involving the simultaneous

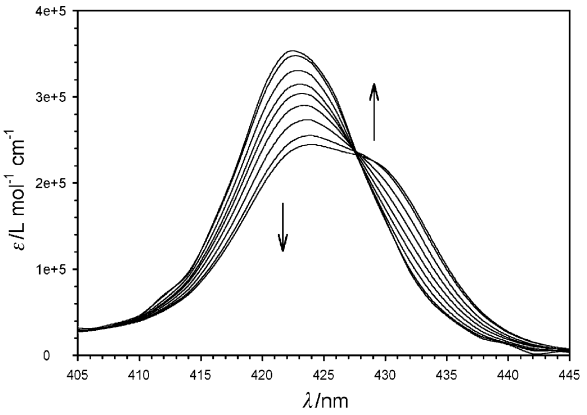


Fig. 1 Plots of the apparent molar absorptivity (ε) of $Zn(PyP_3P) \approx$ 0.04–7 μM in toluene. The arrows indicate how ε changes with increasing $Zn(PyP_3P)$ concentration.

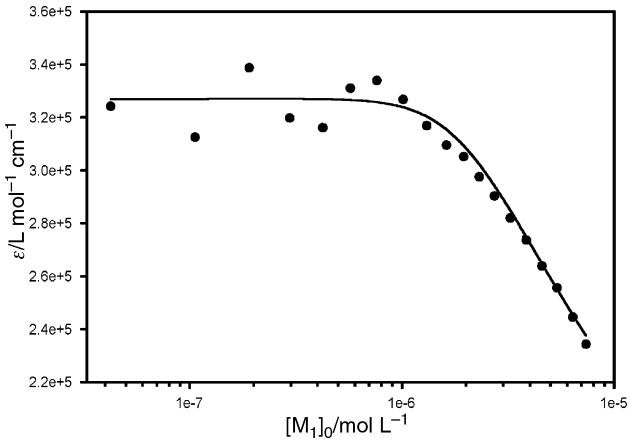


Fig. 2 Plot of the apparent molar absorptivity (ε) at λ 422 nm as a function of $Zn(PyP_3P)$ concentration in toluene. The points are experimental and the curve is calculated (see text).

presence of significant amounts of the monomer, the cyclic tetramer and the linear polymer, whereas it would be compatible either with a model involving an equilibrium between the monomer and the cyclic tetramer [eqn. (6)] or with a model involving linear polymerisation without formation of cyclic species, *i.e.* the model of Fleischer and Shachter.[9]

Let us consider how the absorbance depends on the analytical concentration of $Zn(PyP_3P)$ in the two model cases. In the first case, the absorbance (A) would be given by eqn. (12)

$$A = \varepsilon_0[M_1] + 4\varepsilon_\infty[C_4] \tag{12}$$

where ε_0 and ε_∞ are the molar absorptivities of the free monomer M_1 and of a coordinated porphyrin subunit in the cyclic tetramer C_4. Considering the mass balance equation, $[M_1]_0 = [M_1] + 4[C_4]$, and eqn. (3), eqn. (12) is transformed into eqn. (13).

$$A = \varepsilon_\infty[M_1]_0 + (\varepsilon_0 - \varepsilon_\infty)\frac{x}{K_{inter}} \tag{13}$$

To obtain A as a function of $[M_1]_0$ only, one should solve eqn. (7) with $n = 4$ for x, and introduce the resulting expression into eqn. (13). Although it is not possible to obtain a simple analytical solution of eqn. (7), it can be solved numerically by the Newton–Raphson method.[14] A nonlinear least-squares procedure based on eqn. (13) and on the Newton–Raphson solution of eqn. (7) with $n = 4$ was carried out. The input data were the absorbance data at λ 422 nm as a function of $[M_1]_0$, K_{inter} was held fixed at 4.1×10^3 L mol^{-1}, the EM, ε_0 and ε_∞ were treated as adjustable parameters to improve the fit of calculated to experimental absorbance data. Fig. 2 reports the experimental apparent molar absorptivities (ε) and the curve calculated on the basis of eqn. (7) and (13) with the following optimised parameters: EM $= 20 \pm 5$ mol L^{-1}; $\varepsilon_0 = 3.3 \pm 0.1 \times 10^5$ L mol^{-1} cm^{-1}; $\varepsilon_\infty = 1.5 \pm 0.1 \times 10^5$ L mol^{-1} cm^{-1}.

In the second case, considering that in each linear i-mer the zinc of one of the terminal subunit is not coordinated, the absorbance would be given by eqn. (14)

$$A = \sum_{i=1}^{\infty}(\varepsilon_0 + (i-1)\varepsilon_\infty) \cdot [M_i] \tag{14}$$

where ε_0 and ε_∞ are the molar absorptivities of an uncoordinated and a coordinated porphyrin subunit, respectively, in the linear oligomers. From eqn. (14), eqn. (15) is immediately obtained.

$$A = \varepsilon_\infty \sum_{i=1}^{\infty} i[M_i] + (\varepsilon_0 - \varepsilon_\infty)\sum_{i=1}^{\infty}[M_i] \tag{15}$$

Considering that the first summation is simply $[M_1]_0$, and substituting eqn. (3) for $[M_i]$ into the second summation, eqn. (15) becomes:

$$A = \varepsilon_\infty[M_1]_0 + \frac{(\varepsilon_0 - \varepsilon_\infty)}{K_{inter}}\sum_{i=1}^{\infty} x^i \tag{16}$$

Since the summation appearing in the right-hand side of eqn. (16) is equal to $x/(1-x)$, eqn. (16) is transformed into eqn. (17).

$$A = \varepsilon_\infty[M_1]_0 + \frac{(\varepsilon_0 - \varepsilon_\infty)}{K_{inter}}\frac{x}{(1-x)} \tag{17}$$

To obtain A as a function of $[M_1]_0$ only, one should solve eqn. (2) for x, and introduce the resulting expression into eqn. (17). Also, in the case of eqn. (2) a simple analytical solution is not available, and one must resort to the Newton–Raphson method.[14] A nonlinear least-squares procedure based on eqn. (17) and on the Newton–Raphson solution of eqn. (2) was carried out. The input data were the absorbance data at λ 422 nm as a function of $[M_1]_0$, whereas K_{inter}, ε_0 and ε_∞ were treated as adjustable parameters. The following optimised parameters were obtained: $K_{inter} = 8.5 \pm 1.2 \times 10^4$ L mol^{-1}; $\varepsilon_0 = 3.5 \pm 0.1 \times 10^5$ L mol^{-1} cm^{-1}; $\varepsilon_\infty = -4.0 \pm 2.4 \times 10^4$ L mol^{-1} cm^{-1}. It appears not only that the value of K_{inter} is more than one order of magnitude greater than the experimental value, but also that ε_∞, being negative, is completely unrealistic. By imposing a more realistic value of ε_∞, K_{inter} increases; for example, by assuming $\varepsilon_\infty = 1 \times 10^5$ L mol^{-1} cm^{-1}, $K_{inter} = 3.4 \pm 0.3 \times 10^5$ L mol^{-1}. In conclusion, although the quality of the fit is analogous for the two models, the model of linear polymerisation can be unequivocally ruled out because the required parameters for the best fit are not realistic. By contrast, the model involving a simple equilibrium between the monomer and the cyclic tetramer is supported by realistic values of the molar absorptivities and by the fact that the EM is greater than our estimated lower limit of 0.1 mol L^{-1}, as expected. Furthermore, the product EM K_{inter} ($=8.2 \times 10^4$) largely satisfies the condition shown in eqn. (5) with $n = 4$, thus justifying the validity of the simplified model in eqn. (6) instead of the more complex one shown in Scheme 1. Since the model in eqn. (6) holds, a lower self-assembly concentration (lsac) of 7.1×10^{-6} mol L^{-1} can be calculated from eqn. (18).[5]

$$\text{lsac} = \frac{2}{n^{1/(n-1)}\text{EM}^{1/(n-1)}K_{inter}^{n/(n-1)}} \tag{18}$$

The lsac has been defined as the initial monomer concentration at which the complex is half-assembled, *i.e.* when the yield

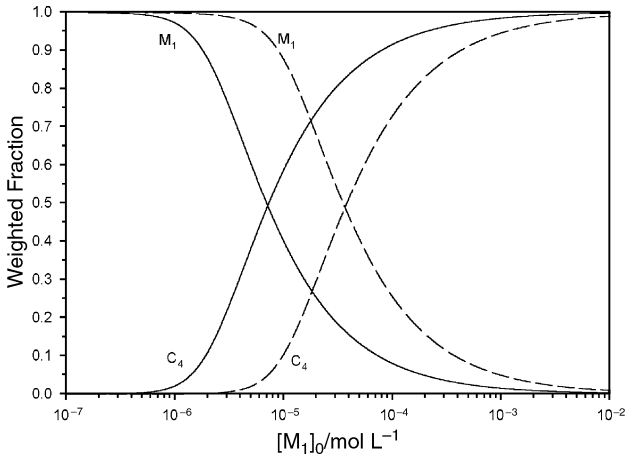


Fig. 3 Distribution curves of the monomer M_1 and of the cyclic tetramer C_4 as a function of the analytical concentration of $Zn(PyP_3P)$. Solid and dashed curves refer to toluene and chloroform solutions, respectively.

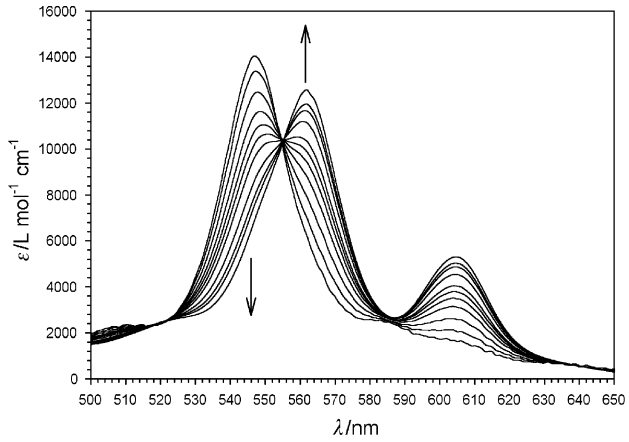


Fig. 4 Plots of the apparent molar absorptivity (ε) of $Zn(PyP_3P) \approx$ 7–200 μM in chloroform. The arrows indicate how ε changes with increasing $Zn(PyP_3P)$ concentration.

is 50%.[5,31] More elaborate calculations, involving the solution of eqn. (1), (3) and (4), allow the evaluation of the distribution curves of the various species present in solution. The solid curves in Fig. 3 are the distribution curves of the cyclic tetramer and the monomer in toluene. Since the weighted fraction of the linear polymer (excluding the monomer) is negligibly low (always lower than 0.016), the lsac coincides with the intersection of the two curves. Note that satisfaction of the condition in eqn. (5) implies that there is a range of initial monomer concentrations at which the process of self-assembly is more than 99% complete. This range begins at $[M_1]_0 = 2 \times 10^{-3}$ mol L^{-1}, as shown in Fig. 3.

It is worth noting that either formation of a cyclic trimer or pentamer fits the data almost as well as formation of the cyclic tetramer, thus, one cannot tell on the basis of absorbance data only which is the cyclic oligomer actually present in solution. However, in the present case, ^1H NMR evidence from Imamura *et al.*,[11] as well as the computational study reported here, strongly support the formation of the cyclic tetramer.

Self-assembly in chloroform

Binding studies were carried out at 25.0 °C with UV/visible spectroscopy using the same procedures illustrated for toluene. The equilibrium constant K_{inter} ($=1.3 \pm 0.2 \times 10^3$ L mol^{-1}) was obtained by analysing the changes in absorbance at λ 549 and 562 nm (Q bands) of $Zn(PyP_3P) \approx 20$ μM upon addition of increasing amounts of pyridine up to $[Py] \approx 10$ mM.

Since the EM represents the intramolecular reactivity (K_{intra}) corrected for the inherent reactivity of end-groups (K_{inter}), its value should be little affected by solvent effects. This has been largely verified in practice[7] and is the rationale for our believing that a theoretical gas phase estimate of the EM is also meaningful for intramolecular processes occurring in solution. By assuming for $CHCl_3$ the same EM value obtained in toluene ($=20$ mol L^{-1}), one can predict, from eqn. (5), that the process of self-assembly of $Zn(PyP_3P)$ should be strongly favoured in $CHCl_3$ as well, and that it should be described by the model in eqn. (6). Accordingly, one can estimate from eqn. (18) the value of the lsac in $CHCl_3$ that should be $\approx 3 \times 10^{-5}$ mol L^{-1}, that is, about five times greater than that in toluene. Dilution studies were therefore carried out in the concentration range $[Zn(PyP_3P)] \approx 7$–200 μM by following the absorbances corresponding to the Q bands (at these concentrations Soret bands display absorbance values which are too high for our purposes). Plots of the apparent molar absorptivity in the 500–650 nm region, reported in Fig. 4, show well-defined isosbestic points. The considerations made for the results in toluene also apply here. A nonlinear least-squares procedure based on eqn. (13) and on the Newton–Raphson solution of eqn. (7) with $n = 4$ and $K_{inter} = 1.3 \times 10^3$ L mol^{-1} was carried out. Absorbance data at λ 547 and 562 nm were taken into account as a function of $[M_1]_0$. Fig. 5 reports the experimental apparent molar absorptivities (ε) and the curves calculated on the basis of eqn. (7) and (13) with the following optimised parameters: at λ 547 nm, EM $= 11 \pm 1$ mol L^{-1}, $\varepsilon_0 = 1.41 \pm 0.02 \times 10^4$ L mol^{-1} cm^{-1}, $\varepsilon_\infty = 5.50 \pm 0.02 \times 10^3$ L mol^{-1} cm^{-1}; at λ 562 nm, EM $= 15 \pm 4$ mol L^{-1}, $\varepsilon_0 = 6.2 \pm 0.3 \times 10^3$ L mol^{-1} cm^{-1}, $\varepsilon_\infty = 1.37 \pm 0.01 \times 10^4$ L mol^{-1} cm^{-1}. From these results, an average EM value of 15 ± 5 mol L^{-1} can be obtained. This value can be considered identical within the experimental errors to that obtained in toluene, confirming, beyond our most optimistic expectations, the independence of the EM from solvent effects.

Best fits of the same absorbance data with the model of linear polymerisation described by eqn. (17) and (2) yielded the following optimised parameters: at λ 547 nm, $K_{inter} = 2.9 \pm 0.6 \times 10^5$ L mol^{-1}, $\varepsilon_0 = 2.6 \pm 0.2 \times 10^4$ L mol^{-1} cm^{-1}, $\varepsilon_\infty = 4.3 \pm 0.1 \times 10^3$ L mol^{-1} cm^{-1}; at λ 562 nm, $K_{inter} = 3.8 \pm 0.8 \times 10^5$ L mol^{-1}, $\varepsilon_0 = -5 \pm 2 \times 10^3$ L mol^{-1} cm^{-1}, $\varepsilon_\infty = 1.47 \pm 0.1 \times 10^4$ L mol^{-1} cm^{-1}. They also appear unrealistic in this case.

The dashed curves in Fig. 3, calculated from eqn. (1), (3) and (4), are the distribution curves of the cyclic tetramer and the monomer in chloroform; the weighted fraction of the linear polymer (excluding the monomer) is also negligibly low

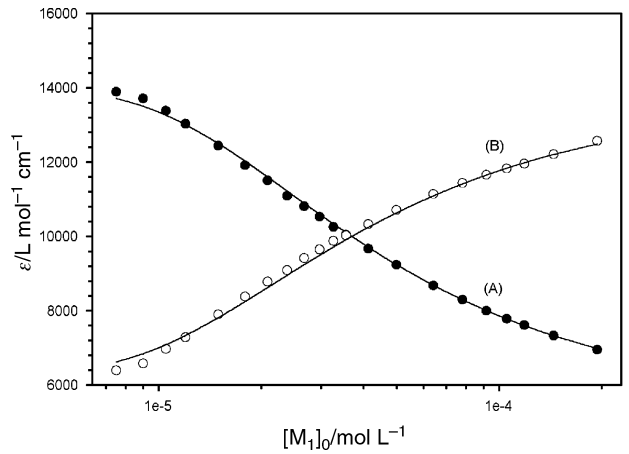


Fig. 5 Plots of the apparent molar absorptivity (ε) at λ 547 (A) and 562 nm (B) as a function of $Zn(PyP_3P)$ concentration in chloroform. The points are experimental and the curves are calculated (see text).

(<0.025) in this case. The calculated lsac ($=3.6 \times 10^{-5}$ mol L^{-1}) is in excellent agreement with the predicted value. The range of initial monomer concentrations at which the process of self-assembly is more than 99% complete begins in $CHCl_3$ at $[M_1]_0 = 0.12$ mol L^{-1}.

Conclusions

Values of EM can be roughly estimated by molecular modelling techniques in the case of rigid subunits undergoing self-assembly into fairly rigid cyclic oligomeric structures. Such estimates should be sufficient in most cases to establish the feasibility of the self-assembly process. Application of the method to $Zn(PyP_3P)$ gave indications that self-assembly should occur in solution to yield a cyclic tetramer. This has been found to be consistent with an UV/visible study of the self-assembly process both in toluene and in chloroform. The experimental values of the EM were, as expected, little affected by the nature of the solvent indicating that the solvent affects the process of self-assembly mainly through the value of K_{inter}. These results revise the conclusions of other authors[9] who suggested that $Zn(PyP_3P)$ yields a linear polymer in solution.

Experimental

Materials and methods

Reagents (Aldrich, Merck or Fluka) were of the highest grade available and were used without further purification. Solvents were dried, distilled and degassed prior to use, using standard procedures.[32] Spectrophotometric grade chloroform (Uvasol, Merck) was stored over activated molecular sieves, and passed through an aluminium oxide (activated, basic, Brockmann 1, Aldrich) column prior to use, in order to remove traces of acids. Silica gel 60 (70–230 mesh) was used for column chromatography. ^{1}H-NMR spectra were recorded with a Bruker AC 300 P (300 MHz) spectrometer. Routine UV/visible spectra were measured on a Varian Cary 1E Spectrophotometer, whereas more delicate measurements were performed on a Perkin Elmer λ18 spectrophotometer equipped with a thermostatted cell holder. Mass spectra (FAB) were recorded on a VG Quattro spectrometer using *m*-nitrobenzyl alcohol (NBA, Aldrich) as a matrix in positive ion mode. PM3 calculations were performed with the GAUSSIAN 98 software package.[33] Structure visualisation and manipulation was carried out with version 4.0 of the molecular modelling package HYPERCHEM (Hypercube Inc., Gainesville, FL).

PM3 calculations

The PM3 optimised structure of $Zn(PyP_3P)$ resulted in C_{2v} symmetry with the C_2 axis passing through the zinc and the nitrogen of the pyridyl group, the porphyrin ring is planar and the meso aromatic rings are orthogonal to the porphyrin plane ($\Delta_f H^\circ = 323.95$ kcal mol^{-1}). As to the dimer, $[Zn(PyP_3P)]_2$, since there is the possibility of internal rotation about the N(pyridyl)–Zn bond, the corresponding adiabatic potential energy curve was evaluated in the range $0^\circ \leqslant \theta \leqslant 180^\circ$. The lowest energy minimum ($\Delta_f H^\circ = 640.83$ kcal mol^{-1}), corresponds to the conformation in which the two pyridyl rings are coplanar ($\theta = 0, 180^\circ$); this conformation resulted in C_s symmetry, with the symmetry plane coplanar with pyridyl rings. Another relative energy minimum 0.02 kcal mol^{-1} was found at $\theta = 90^\circ$. Two maxima of practically the same relative energy (0.59 kcal mol^{-1}) were found at $\theta = 45$ and 135°. Since the barrier height is very low, the internal rotation can be considered virtually free at room temperature. As to the linear tetramer $[Zn(PyP_3P)]_4$, owing to the almost free rotation about the N(pyridyl)–Zn bonds, its potential energy surface is rather flat and this caused severe difficulties in localising the global minimum conformation. After some unsuccessful attempts, we decided to consider the structure with the three N(pyridyl)–Zn bonds in the *anti* conformation as a plausible representative of the lowest energy conformation. To constrain the structure in the *anti* conformation and to speed up the geometry optimisation we chose, on the basis of the results obtained with the linear dimer, to impose a plane of symmetry coplanar with all the pyridyl rings ($\Delta_f H^\circ = 1272.95$ kcal mol^{-1}). An attempt to optimise the geometry of the cyclic tetramer *c*-$[Zn(PyP_3P)]_4$ without symmetry constraints failed at an intermediate stage of the calculation because of the inability of the program to achieve self-consistency. However, by imposing C_4 symmetry, which should be a reasonable assumption, the geometry optimisation went on to completion. The optimised structure of *c*-$[Zn(PyP_3P)]_4$ resulted in C_{4h} symmetry with the zinc atoms forming a square with sides of 9.86 Å in length ($\Delta_f H^\circ =$ 1263.70 kcal mol^{-1}).

Synthesis of $Zn(PyP_3P)$

The porphyrin derivative $Zn(PyP_3P)$ was synthesised by a modification[34] of the earlier published procedure,[9] as follows: to a solution of 26 mL (0.256 mol) of benzaldehyde and 6.0 mL (0.063 mol) of 4-pyridinecarboxaldehyde in 250 mL of refluxing propionic acid, a solution of 17.5 mL (0.252 mol) of freshly distilled pyrrole in 50 mL of propionic acid was added dropwise over a period of 30 min. The reaction mixture was stirred at reflux temperature for additional 3 h, then cooled to room temperature and allowed to stand overnight. The crystalline residue, separated from the bulk of the reaction mixture by filtration, was thoroughly washed with methanol. 2.5 g of shiny dark-purple crystals, consisting of a mixture of two products, was chromatographed on SiO_2. Elution with chloroform gave a first moving band consisting of the tetraphenylporphyrin, H_2TPP, by-product. The desired porphyrin was subsequently eluted with a 1% methanol–chloroform mixture to give 0.9 g (1.5 mmol, 3% yield) of H_2PyP_3P as a purple crystalline solid, which was purified by dissolving in dichloromethane and adding hexane until complete precipitation. The Zn-metallation of the free base was accomplished according to the literature[5] to give, after work-up and dichloromethane–hexane crystallisation (as above), 870 mg (80% yield) of $Zn(PyP_3P)$. The identity of the porphyrin products was confirmed by UV/visible, ^{1}H-NMR and FAB mass spectroscopies.

Determination of K_{inter}

The intermolecular equilibrium constants K_{inter} were obtained *via* spectrophotometric titration by adding a pyridine solution to a solution of $Zn(PyP_3P)$ in a 1 cm path length quartz cuvette, using microliter syringes. The porphyrin was present in the ligand at the same concentration as that in the cuvette, to avoid dilution effects. Titrations were carried out at 25.0 °C up to a pyridine concentration of about 1×10^{-2} mol L^{-1}. Titrations in toluene were followed at λ 422 and 430 nm ($[Zn(PyP_3P)] \approx 2 \times 10^{-6}$ mol L^{-1}). Titrations in chloroform were followed at λ 549 and 562 nm ($[Zn(PyP_3P)] \approx 2 \times 10^{-5}$ mol L^{-1}). The association constants K_{inter} were evaluated by nonlinear least-squares curve fitting of the absorbance data to a standard equation for 1 : 1 complexation.[35]

Self-assembly studies

UV/visible self-assembly macrocyclisation studies of $Zn(PyP_3P)$ were performed at 25.0 °C by adding aliquots of a concentrated stock solution of the title porphyrin (2.5×10^{-4} mol L^{-1} in toluene; 1.80×10^{-3} mol L^{-1} in chloroform) to a 1 cm quartz cell containing 2.5 mL of solvent (toluene or chloroform). The UV/visible spectra after each addition were

recorded in digital format. Nonlinear least-squares fittings of the absorbance data to both the models of simple cyclisation and linear polymerisation were carried out by appositely written computer programs (see Results and discussion for further details). The programs, coded in BASIC, are based on the Gauss method and have been adapted from a published coding,[36] the principal variation being in the routine of function evaluation that calls for a subroutine performing the numerical solution of the appropriate algebraic equation by the Newton–Raphson method.

Moments of inertia

Reduced moments of inertia were evaluated by an approximate literature method[28] (see Results and discussion for further details) implemented into an appositely written MatLab (The MathWorks, Natick, MA) script. This script extracts the atomic coordinates, by dynamic data exchange, from the appropriate structure displayed in the HYPERCHEM program, and performs the relevant calculations.

References and notes

1 For recent reviews, see: J.-M. Lehn, *Supramolecular Chemistry: Concepts and Perspectives*, VCH, Weinheim, 1995; D. S. Lawrence, T. Jiang and M. Levett, *Chem. Rev.*, 1995, **95**, 2229; D. Philp and J. F. Stoddart, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 1154; *Comprehensive Supramolecular Chemistry*, ed. J. L. Atwood, J. E. D. Davies, D. D. MacNicol, F. Vögtle, J.-P. Sauvage and M. W. Hosseini, Pergamon, Oxford, 1996, vol. 9; F. W. Zeng and S. C. Zimmerman, *Chem. Rev.*, 1997, **97**, 1681; B. Linton and A. D. Hamilton, *Chem. Rev.*, 1997, **97**, 1669; M. M. Conn and J. Rebek, *Chem. Rev.*, 1997, **97**, 1647; M. Fujita, *Chem. Soc. Rev.*, 1998, **27**, 417; S. Leininger, B. Olenyuk and P. J. Stang, *Chem. Rev.*, 2000, **100**, 853; L. F. Lindoy and I. Atkinson, *Self-Assembly in Supramolecular Systems* (*Monographs in Supramolecular Chemistry*), ed. J. F. Stoddart, RSC, Cambridge, 2000; *Molecular Self-Assembly* (*Structure and Bonding, Vol. 96*), ed. M. Fujita, Springer, Heidelberg, 2000; J.-C. Chambron, V. Heitz and J.-P. Sauvage, *Noncovalent Multiporphyrin Assemblies*, in *The Porphyrin Handbook*, ed. K. M. Kadish, K. M. Smith and R. Guilard, Academic, London, 2000, vol. 6, pp. 1–42.

2 X. Chi, A. J. Guerin, R. A. Haycock, C. A. Hunter and L. D. Sarson, *J. Chem. Soc., Chem. Commun.*, 1995, 2563.

3 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.

4 H. L. Anderson, *Inorg. Chem.*, 1994, **33**, 972; P. N. Taylor and H. L. Anderson, *J. Am. Chem. Soc.*, 1999, **121**, 11538.

5 G. Ercolani, *J. Phys. Chem. B*, 1998, **102**, 5699.

6 A. J. Kirby, *Adv. Phys. Org. Chem.*, 1980, **17**, 183.

7 L. Mandolini, *Adv. Phys. Org. Chem.*, 1986, **22**, 1.

8 C. Galli and L. Mandolini, *Eur. J. Org. Chem.*, 2000, 3117.

9 E. B. Fleischer and A. M. Shachter, *Inorg. Chem.*, 1991, **30**, 3763.

10 X. Chi, A. J. Guerin, R. A. Haycock, C. A. Hunter and L. D. Sarson, *J. Chem. Soc., Chem. Commun.*, 1995, 2567.

11 K. Funatsu, T. Imamura, A. Ichimura and Y. Sasaki, *Inorg. Chem.*, 1998, **37**, 1798.

12 The term predisposition has been used in this context to indicate a strong conformational or structural preference expressed by the building block once incorporated into a larger structure, giving rise to a thermodynamic preference for a particular product: S. J. Rowan, D. G. Hamilton, P. A. Brady and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1997, **119**, 2578.

13 Owing to a misprint in eqn. (10) in this paper, however, this equivalence is not readily recognised (the power 2 was erroneously printed inside, instead of outside, the parenthesis).

14 H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry*, 2nd edn., Van Nostrand, Princeton, 1962, p. 492.

15 J. R. Miller and G. D. Dorough, *J. Am. Chem. Soc.*, 1952, **74**, 2977.

16 C. H. Kirksey, P. Hambright and C. B. Storm, *Inorg. Chem.*, 1969, **8**, 2141.

17 S. J. Cole, G. C. Curthoys, E. A. Magnusson and J. N. Phillips, *Inorg. Chem.*, 1972, **11**, 1024.

18 G. C. Vogel and L. A. Searby, *Inorg. Chem.*, 1973, **12**, 936.

19 G. C. Vogel and B. A. Beckmann, *Inorg. Chem.*, 1976, **15**, 483.

20 P. Bhyrappa, V. Krishnan and M. Nethaji, *J. Chem. Soc., Dalton Trans.*, 1993, 1901.

21 G. C. Vogel and J. R. Stahlbush, *Inorg. Chem.*, 1977, **16**, 950.

22 M. Nappa and J. S. Valentine, *J. Am. Chem. Soc.*, 1978, **100**, 5075.

23 An isodesmic reaction is one with the same number and types of bonds on both sides. Isodesmic reactions allow extensive cancellation of correlation energy, and systematic errors introduced in semi-empirical methods through parameterisation: W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.

24 J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 209; J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 221.

25 M. E. Zandler and F. D'Souza, *J. Mol. Struct. (THEOCHEM)*, 1997, **401**, 301.

26 M. I. Page and W. P. Jencks, *Proc. Natl. Acad. Sci. U. S. A.*, 1971, **68**, 1678; M. I. Page, *Chem. Soc. Rev.*, 1973, **2**, 295.

27 H. E. O'Neal and S. W. Benson, *J. Chem. Eng. Data*, 1970, **15**, 266.

28 D. R. Herschbach, H. S. Johnston, K. S. Pitzer and K. S. Powell, *J. Chem. Phys.*, 1956, **25**, 736; G. Ercolani, *J. Chem. Educ.*, 2000, **77**, 1495.

29 G. Ercolani, *J. Org. Chem.*, 1999, **64**, 3350, and references therein.

30 D. H. Williams and M. S. Westwell, *J. Chem. Soc. Rev.*, 1998, **27**, 57, and references therein.

31 This concept, originally introduced by Hunter *et al.* (ref. 2) as the csac (critical self-assembly concentration), has been successively commented upon, and its expression slightly modified by us (ref. 5).

32 D. D. Perrin, W. L. F. Armarego and D. E. Perrin, *Purification of Laboratory Chemicals*, 2nd edn., Pergamon, New York, 1980.

33 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle and J. A. Pople, GAUSSIAN 98, Rev. A.7, Gaussian, Inc., Pittsburgh, PA, 1998.

34 D. Monti, M. Venanzi, G. Mancini, F. Marotti, L. La Monica and T. Boschi, *Eur. J. Org. Chem.*, 1999, 1901.

35 K. A. Connors, *Binding Constants. The Measurement of Molecular Complex Stability*, Wiley, New York, 1987.

36 M. D. Johnston, Jr., *Computational Chemistry*, Elsevier, Amsterdam, 1988, pp. 498–507.