



A thermodynamic model of auto-regulated protein assembly by a supramolecular scaffold

| | |
|------------------|---|
| Title | A thermodynamic model of auto-regulated protein assembly by a supramolecular scaffold |
| Author(s) | Rennie, Martin L.;Crowley, Peter B. |
| Publication Date | 2019-03-12 |
| Publisher | Wiley |
| Repository DOI | 10.1002/cphc.201900153 |

A Thermodynamic Model of Auto-regulated Protein Assembly by a Supramolecular Scaffold

Dr. Martin L. Rennie*^[a,b], and Prof. Dr. Peter B. Crowley*^[a]

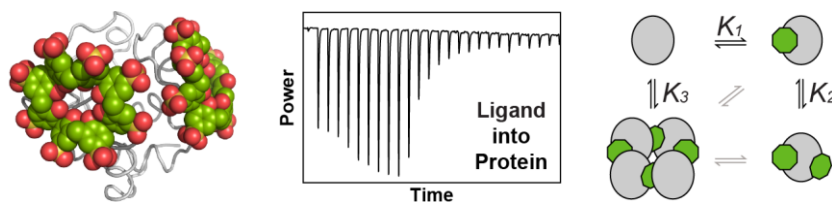
^[a]School of Chemistry, National University of Ireland Galway, University Road, Galway, Ireland

^[b]Present address: Institute of Molecular Cell and System Biology, University of Glasgow, University Avenue, Glasgow, UK

Correspondence: martin.rennie@glasgow.ac.uk +44 141 330 8119
peter.crowley@nuigalway.ie +353 91 49 24 80

Keywords: Bayesian, calix[8]arene, ITC, multi-body, oligomerization

Table of Contents



Auto-regulated Assembly of cytochrome *c* by sulfonato-calix[8]arene results in biphasic ITC thermograms. The interpretation of such data is non-trivial. We quantified the thermodynamics of ligand binding and assembly by testing several models with Bayesian fitting. Simultaneous data fitting to standard and reverse titrations was performed. Variations of the model may aid the analysis of auto-regulated assembly in general.

Abstract

Ligand-mediated regulation of protein assembly occurs frequently in different cellular contexts. Auto-regulated assembly, where a ligand acts as its own competitive inhibitor, provides a mechanism for exquisite control of assembly. Unlike simple protein-ligand systems a quantification of the binding thermodynamics is not straightforward. Here, we characterize the interactions of a recently identified model system in which the oligomerization of cytochrome *c* is controlled by sulfonato-calix[8]arene, an anionic supramolecular scaffold. Isothermal titration calorimetry and thermodynamic modelling, in combination with Bayesian fitting, were used to quantify the ligand binding and assembly equilibria for this system. The approach and variations of this model may prove useful for the analysis of auto-regulated protein assembly in general.

Introduction

Regulated protein assembly and disassembly is central to biological function.^[1,2] There are numerous examples of small molecule binding to proteins that controls the oligomeric state. This mechanism has been referred to as polysteric linkage,^[3,4] contrasting with allosteric linkage in which protein conformation and dynamics are regulated by ligands. Recently, protein dimerization and higher order assembly has been engineered *via* synthetic molecules, including bivalent ligands^[5-7] and supramolecular scaffolds.^[8-11] Competitive inhibitors of these molecules can be used to control disassembly.^[6,8]

In nature, multivalent ligands can act as their own competitive inhibitors, the exploitation of which may yield more sophisticated methods to control protein assembly. These systems are characterized by bell-shaped dose-response curves, where there is a maximum concentration of ligand required to achieve peak response, above which the response is reduced.^[12] This effect was reported more than a century ago for antibodies (“prozone effect” or “hook effect”),^[13] and has since been observed in DNA transcription (“auto-regulation”),^[14] heparin-regulated enzyme activity (“template mechanism”),^[15,16] scaffold proteins (“combinatorial inhibition”),^[17] phytate-mediated protein precipitation,^[18,19] and supramolecular control of protein assembly.^[5,11] Here, we refer to this property as auto-regulated assembly. A quantitative description of the binding thermodynamics is crucial for understanding auto-regulated assembly and comparing different systems. The simplest case of two protein molecules assembled *via* a bidentate ligand has been well-studied.^[12] This three-body system is equivalent to a two-site model^[20] (the ligand has two protein binding sites). However, there are examples where assembly proceeds beyond a three-body supramolecule.

Recently, we uncovered such an example studying the interaction between a cationic model protein, cytochrome *c*, and a flexible anionic calixarene, sulfonato-calix[8]arene (**sclx₈**).^[11] Cytochrome *c* is a monomeric electron transport protein, while calixarenes are synthetic macrocycles that have been shown to modulate protein assembly.^[9-11,21,22] Auto-regulated assembly was demonstrated by using a variety of techniques including, light scattering, NMR measurements and X-ray crystallography. It was found that tetramerization of cytochrome *c* can be triggered by ~1 equivalent of **sclx₈**, while disassembly to the (ligand-coated) monomer occurs at >2 equivalents.^[11] Structural characterization revealed that cytochrome *c* has at least two **sclx₈** binding sites and that **sclx₈** can link two cytochrome *c* molecules (Figure 1). These observations indicate that the protein and ligand are reciprocally bidentate *i.e.* a molecule of cytochrome *c* can bind two **sclx₈** molecules *and* a molecule of **sclx₈** can bind two cytochrome *c* molecules. This arrangement provides a mechanism for the formation of oligomers greater than a dimer,^[23] which is not possible when only the ligand is bidentate. We have shown that **sclx₈** can trigger also the assembly of a *Penicillium* anti-

fungal protein (PAF).^[22] However, PAF only has a single **sclx₈** binding site (Figure 1) and the binding thermodynamics were described adequately by the bidentate ligand model. In an extreme case, a bell-shaped precipitation response has been observed for lysozyme and phytate, another polycationic protein-polyanionic ligand system.^[19] Precipitation was observed at ~4 mM phytate, while lysozyme remained soluble at 30 mM phytate. The thermodynamics of this system have been described partially.^[24] However, a more complete quantitative description, modelling both ligand binding and assembly, is lacking.

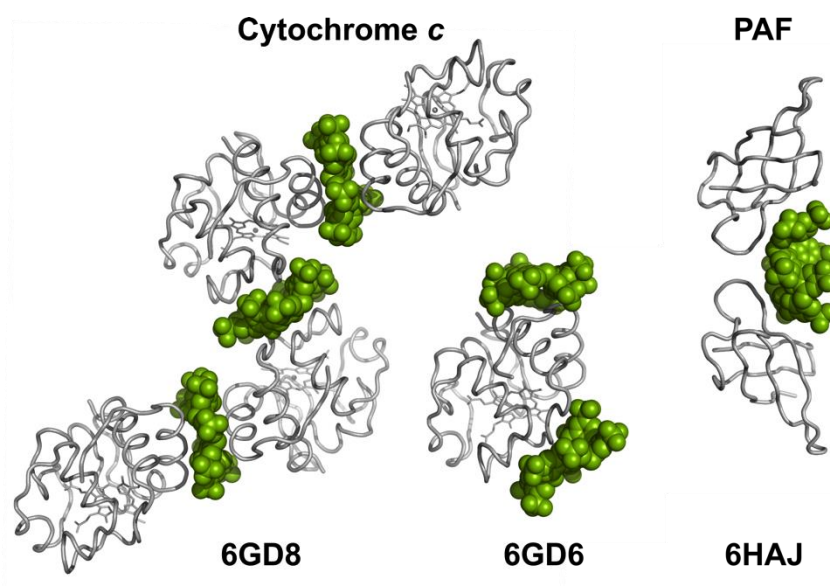


Figure 1. Crystal structures of protein – **sclx₈** assemblies. A **sclx₈**-mediated tetramer of cytochrome c is evident in PDB 6GD8. Two molecules of **sclx₈** mask the surface of cytochrome c in PDB 6GD6.^[11] An **sclx₈**-mediated dimer of *Penicillium* antifungal protein (PAF) in PDB 6HAJ.^[22] Protein and **sclx₈** are represented as grey ribbons and green spheres, respectively. The heme group of cytochrome c is shown as sticks.

Isothermal titration calorimetry (ITC) is the method of choice for thermodynamic characterization of molecular interactions.^[25,26] Importantly for auto-regulated assembly, both ligand binding and assembly processes yield heat changes that can be detected by ITC. For standard ITC analysis a single set of independent binding sites model is fit to the experimental data with the affinity and enthalpy parameters extracted simultaneously by non-linear least squares methods. Fitting models with multiple affinities and enthalpies is often non-trivial, as has been highlighted even for the relatively simple two-site model.^[4,20] These models have many adjustable parameters which can complicate the determination of parameter uncertainties.^[20,27-29] Bayesian methods provide a robust approach to this problem and have been deployed for ITC analysis in the form of an

open source software called *pytc*.^[30,31] Recently, we used Bayesian fitting in *pytc* to determine the thermodynamic parameters of the bidentate ligand (two-site) model for the PAF – **sclx₈** complex.^[22]

Here, we have characterized the cytochrome *c* – **sclx₈** system in detail by using ITC and thermodynamic modelling that incorporates ligand binding and oligomerization. The ITC data are atypical and could be described adequately by two different models. Only one of these models accurately reproduces the experimentally determined auto-regulated assembly. Simultaneous fitting of standard and reverse titrations using Bayesian methods is shown to determine the affinities and enthalpies, and the associated parameter uncertainties. Finally, an extension of the model is proposed which may facilitate the description of related systems.

Results and Discussion

Complexation between cytochrome *c* and **sclx₈** is an auto-regulated process that shifts between monomers and tetramers, depending on the protein:ligand ratio.^[11] In an effort to quantify these interactions, ITC measurements were carried out in standard (**sclx₈** into cytochrome *c*) and reverse (cytochrome *c* into **sclx₈**) mode (Figure 2). Both isotherms were atypical with reproducible extrema at near-equimolar ratios, consistent with a complex process. The bidentate ligand model^[12,22] (Figure S1a) was able to reproduce the main trends. However, it failed to account for the sharpness of the transitions (Figure S1b). Furthermore, this model does not account for tetramer assembly / disassembly, which is a key aspect of the system. Therefore, we sought alternative binding models that could provide a more accurate description of the data. Models involving oligomers up to a tetramer were explored.

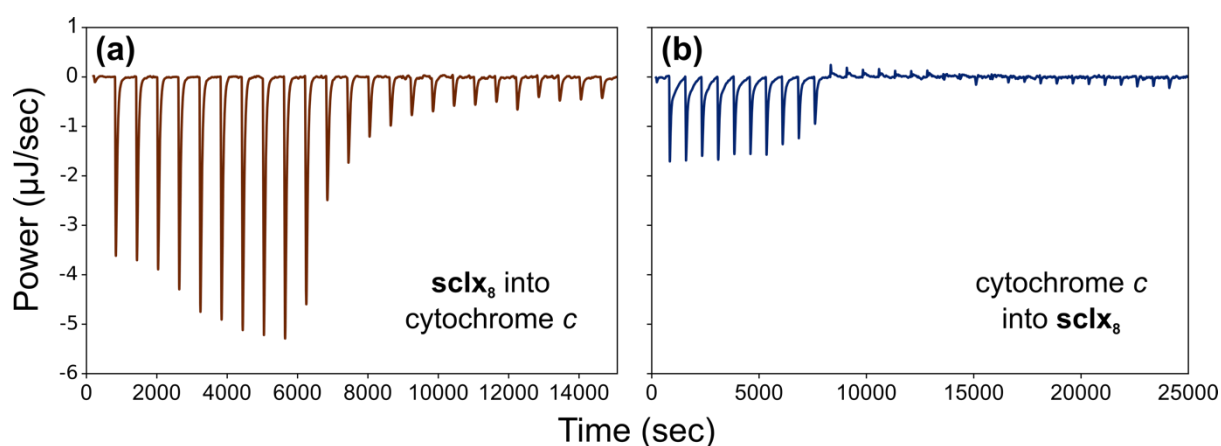


Figure 2. Thermograms for (a) Standard titration of **sclx₈** (1200 μM in the syringe) into cytochrome *c* (74 μM in the cell) and (b) Reverse titration of cytochrome *c* (640 μM in the syringe) into **sclx₈** (73 μM in the cell). See methods for details and baseline correction.

Two candidate models were identified that provided good agreement with the ITC data (Figures 3 and 4). Both models posit two monomeric ligand-bound states, with one or two ligands. Model 1 was relatively simple with a single oligomeric state; a tetramer with flexible ligand stoichiometry (n_L ; Figure 3a). This model has a total of 7 thermodynamic parameters (three equilibrium constants, three enthalpies, and one stoichiometry). The best fit to the ITC data for Model 1 was obtained with a tetramer that involved five **sclx₈** molecules ($n_L = 5$, Figure S2). Model 2, is more complex, positing two dimeric intermediates and a tetramer (Figure 3b), and has a total of 10 thermodynamic parameters (five equilibrium constants and five enthalpies).

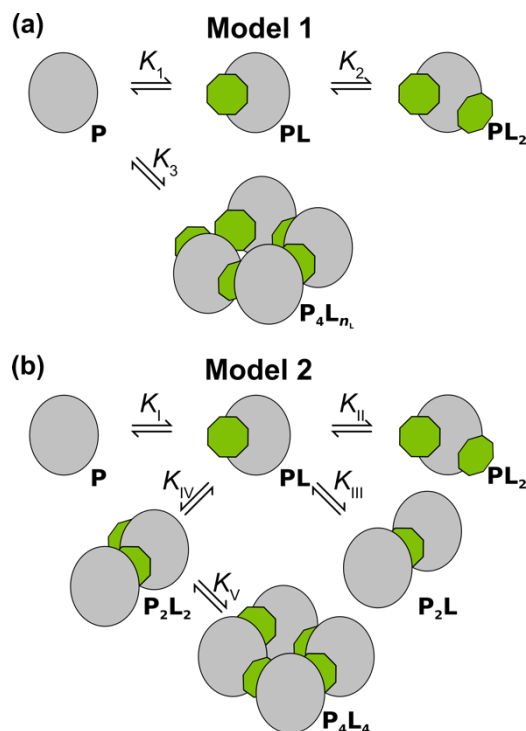


Figure 3. Proposed models for auto-regulated assembly of cytochrome *c* (grey; P) by sclx₈ (green; L). **(a)** Model 1 consists of four protein states governed by three equilibrium constants and one stoichiometry defining the number of ligands (n_L) in the assembled state. **(b)** Model 2 consists of six protein states governed by five equilibrium constants. Schematic representations are shown for illustrative purposes only. The binding models posit only the stoichiometries of each state.

To distinguish between these models we compared the species predicted by each model to be present during titration of ligand into protein. Auto-regulated assembly of cytochrome *c* by sclx₈ proceeds with tetramer formation at approximately 1 equivalent of ligand.^[11] The concentrations of oligomeric species, based on the fit parameters, should mirror this trend. The concentration of tetramer for Model 1 peaked at approximately 1 equivalent of ligand. In contrast, Model 2 predicted a negligible population of tetramer (< 10%), and the dimeric species peaking at 0.5 equivalents of ligand (Figure 5). Despite its simplicity Model 1 provided the most accurate description of auto-regulated assembly, in agreement with other experimental data.^[11] The Bayesian Information Criterion (BIC) provides an alternative method to compare models, accounting for both the fit to data and the number of parameters (an increasing number of parameters is penalized).^[32] Model 1 has a BIC significantly lower than Model 2 (Figure 4), supporting rejection of the latter model. In summary, although Model 2 involves several extra parameters, the flexibility afforded to Model 1 by the variable ligand stoichiometry (in the oligomer) facilitates a better description of the experimental data.

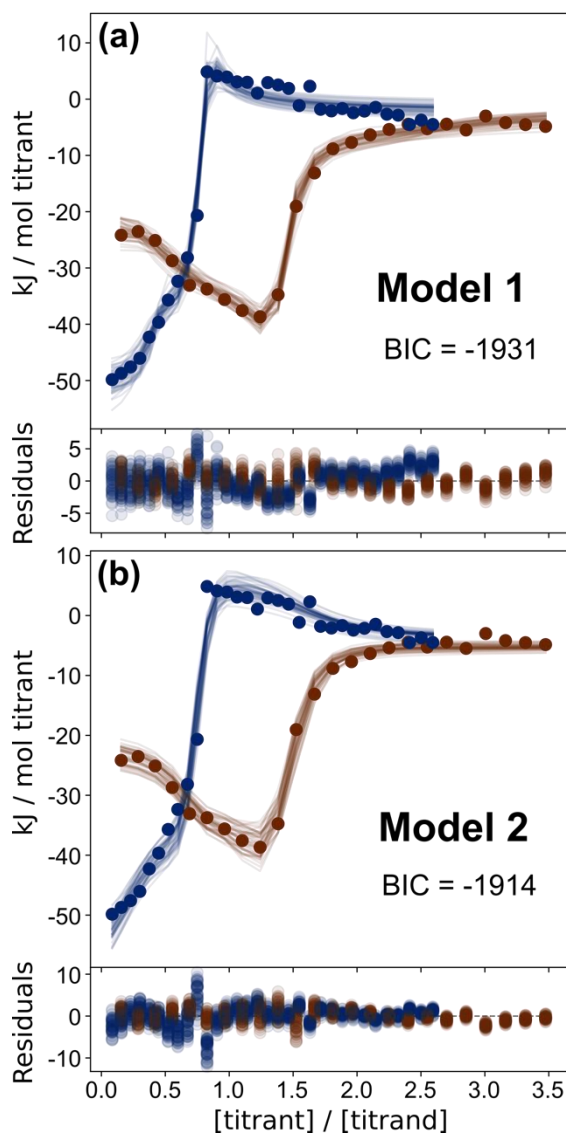


Figure 4. Comparison of fits to the ITC data for the models shown in Figure 3. Global fits of each model to the isotherms for standard (**sclx**₈ into cytochrome *c*; red) and reverse (cytochrome *c* into **sclx**₈; blue) titrations are shown. For each model, 50 randomly sampled fits (from 2500 Markov Chain Monte Carlo iterations) are shown.

The thermodynamic parameters of Model 1 for the cytochrome *c* – **sclx**₈ system were determined using Bayesian Markov Chain Monte Carlo (MCMC) as implemented in *pytc*.^[31] MCMC fitting to either the standard or reverse titrations individually resulted in several parameters that were poorly determined, as indicated by large uncertainties and strong inter-parameter correlations (Figures S3 and S4). In particular the enthalpies were poorly determined. In contrast global MCMC fitting resulted in all parameters (Table 1 and Figure 6) being well-defined (Figures S5). The dissociation constants are reported on a logarithmic scale to facilitate clarity of the associated uncertainty.

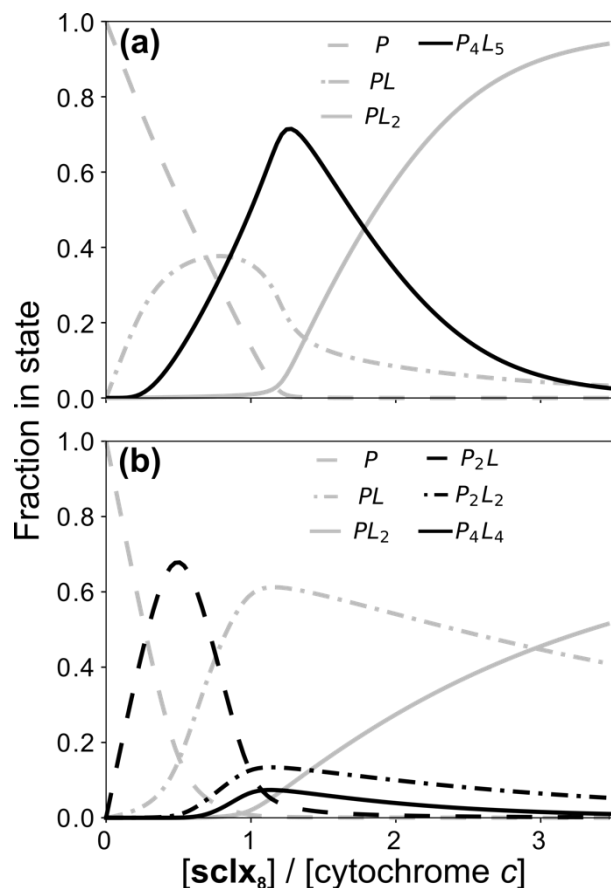


Figure 5. Proportion of cytochrome *c* predicted to be in each state. The concentrations of each state were simulated based on the fit equilibria for **(a)** Model 1 and **(b)** Model 2. The oligomer concentration (black line) in Model 1 mirrors the trends determined by light scattering (See Figure 1c in Rennie *et al*, 2018^[11]).

The first sclx_8 ligand binds to cytochrome *c* with $K_d = 0.01 \mu\text{M}$, with favourable enthalpic and entropic contributions. The second sclx_8 binds with $K_d = 2 \mu\text{M}$. This ~200-fold lower affinity is due primarily to a decreased entropic contribution (Figure 6). This reduction in entropy is greater than would be expected due to the statistical effects, *i.e.* the availability two sites for the first ligand binding versus one for the second ligand binding. Alternatively, the reduced entropy may be rationalized by considering that the first sclx_8 can ‘explore’ a larger protein surface,^[21] while the second ligand is restricted. The binding affinities of sclx_8 are comparable to the $K_d \sim 2 - 8 \mu\text{M}$ for binding of the related ligand, phosphonato-calix[6]arene (pclx_6),^[10] and are much tighter than the $K_d \sim 30 \mu\text{M}$ for the smaller sulfonato-calix[4]arene.^[32] Interestingly, pclx_6 triggers dimerization of cytochrome *c* but not in an auto-regulated fashion,^[10] suggesting a fundamentally different assembly mechanism compared to sclx_8 .

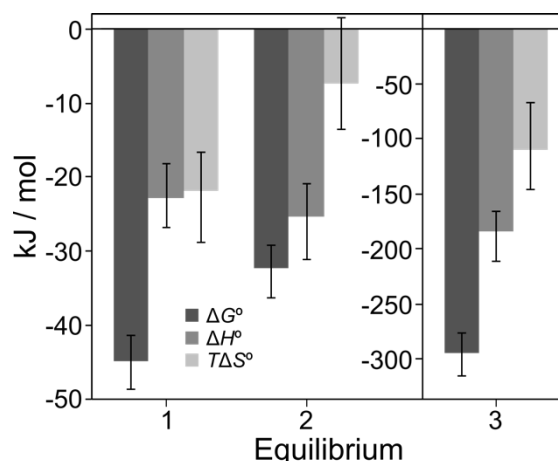


Figure 6. The fit thermodynamic parameters of Model 1. The equilibria are defined in Figure 2a and representative fits are shown in Figure 4a. Error bars are at 95 % credibility intervals, determined from Markov Chain Monte Carlo fitting.

Table 1. Dissociation constants and uncertainties for Model 1^[a]

| | Equilibrium ^[b] | | |
|------------------|----------------------------|----------------------|-------------------------|
| | 1 | 2 | 3 |
| $\log_{10}(K_d)$ | $-7.9^{+0.6}_{-0.7}$ | $-5.7^{+0.6}_{-0.7}$ | $-51.5^{+3.1}_{-3.8}$ |
| | | | $-6.4^{+0.4[c]}_{-0.5}$ |

^[a]Median values and, in parenthesis, 95 % credibility intervals.

^[b]See Figure 3a for the corresponding equilibria.

^[c]The average dissociation constants for the oligomerization equilibrium.

An interpretation of the thermodynamic parameters for oligomer formation (K_3 , Figure 3a and Equation 1b) is hampered by the high order and the dependence on ligand concentration. Previous quantitative analyses of homo-oligomerization have used half-points,^[34,35] critical transition concentrations,^[36] or critical micelle concentrations.^[35] However, in our case, oligomerization is mediated by the ligand and as such these quantities are functions of the ligand concentration.^[35] As an alternative we introduce a heuristic parameter, the *average dissociation constant*, defined as:

$$K_{d,ave} = \sqrt[N]{K_d}$$

where N is the order of the equilibrium (here, $N = 8$ as there are 9 bodies in the equilibrium). The average dissociation constant has units of Molar and is a reasonable approximation for the half-point of a homo-oligomerization equilibrium.^[34,35] Unlike the half-point it is straightforward to apply to any high order equilibrium. The calculated average dissociation constant for the cytochrome *c* – **sclx**₈ tetramer is $\sim 0.4 \mu\text{M}$, suggesting that the assembly persists at sub-micromolar concentrations. This result is consistent with SEC-MALS experiments, in which the **sclx**₈-mediated tetramer of cytochrome *c* was observed with an injected sample of $60 \mu\text{M}$ 1:1 protein:ligand.^[11] The enthalpy for

tetramerization is about 7-fold larger than the enthalpy of a single **sclx₈** binding (Figure 6). This result is approximately consistent with the formation of 8-9 cytochrome *c* – **sclx₈** interfaces in the assembly (**P₄L₅**, Figure 3a). Enthalpies of similarly large magnitude have been observed before for high order equilibria.^[34,37]

Although ITC does not provide direct information on the structure of the assembled state, previous SAXS measurements suggested that the cytochrome *c* molecules are arranged in a closed or semi-closed cyclic structure.^[11] The result here of 5 **sclx₈** molecules within the assembly is consistent with a semi-closed ring in which either end of the ring is capped with an **sclx₈** molecule. It is also plausible that a closed ring, mediated by 4 **sclx₈** molecules may bind an additional **sclx₈** molecule on one side of the ring.

As a further validation of the fitting results the apparent affinity between cytochrome *c* and **sclx₈** was determined by using an orthogonal experimental approach. The rate of reduction of cytochrome *c* by ascorbic acid was found to be dependent on the **sclx₈** concentration (Figure S6a). This observation is consistent with X-ray crystal structures that show the exposed heme edge is partially masked by **sclx₈**.^[11] The kinetic data were adequately described by a simple single-site binding model yielding an apparent dissociation constant of < 1 μM (Figure S6b) consistent with the ITC data.

Model 1 provided a reasonable description of the cytochrome *c* – **sclx₈** system. Variations of this model may prove useful to describe other auto-regulated assembly systems. An obvious extension is that the assembled state may involve *n* protein molecules. Another possibility is that the protein monomer may have more than two ligand binding sites. These features are straightforward to incorporate using the binding polynomial, which provides a concise mathematical description of a thermodynamic model.^[4] The binding polynomial (*Z*) for the generalized model is given by:

$$Z = 1 + K_1[L] + K_1K_2^{m-1}[L]^m + K_3^{n_P+n_L-1}[P]^{n_P-1}[L]^{n_L}$$

where *m* is the number of binding sites on the monomer, *K*₁ is the association constant for binding the first ligand, *K*₂ is the average association constant for binding *m* ligands, and *K*₃ is the average association constant for oligomerization. [*P*] and [*L*] are the concentrations of the free protein monomer and the free ligand. For the cytochrome *c* – **sclx₈** system, *m* = 2, *n_P* = 4, and *n_L* = 5. The model is available *via* https://pytc.readthedocs.io/en/latest/indiv_models/assembly-auto-inhibition.html

Conclusions

Auto-regulation provides a means to control protein assembly. Unlike simple protein-ligand systems a thermodynamic quantification of auto-regulation is not straightforward. We addressed this challenge in the cytochrome *c* – **sclx**₈ system. Isothermal titration calorimetry data collected in both standard and reverse mode provided a rich source of quantitative data to test different models. A relatively simple model, involving 4 protein states, provided the best agreement with the experimental data. The use of Bayesian MCMC and simultaneous fitting of the model to both isotherms facilitated robust determination of the multiple affinities and enthalpies of the model. There is increasing interest in thermodynamic binding models that treat multiple bodies.^[4,20,23,25,28,29,34,36,38–41] By generalizing the model developed here we have provided an approach that may be useful to describe other auto-regulated systems.

The anionic **sclx**₈ prompts the assembly of two cationic proteins, cytochrome *c*^[11] and PAF,^[22] suggesting its utility as a generic mediator of auto-regulated assembly for cationic proteins. For both proteins the binding affinity increases with the size of the calixarene macrocycle. The eight unit **sclx**₈ has a low nanomolar affinity for cytochrome *c*. Larger calixarenes^[42] may have even tighter affinity for cationic proteins. Similarly, protein size may correlate with the number of calixarene binding sites. PAF (6 kDa) possesses a single calix[8]arene binding site,^[22] while cytochrome *c* (13 kDa) possesses at least 3 binding sites under high ionic strength.^[11] Larger cationic proteins may have even more binding sites or different inter-site geometry, resulting in different oligomeric forms. The generalized model may facilitate quantification of such systems.

Experimental Section

Materials. *Saccharomyces cerevisiae* cytochrome *c* C102T was produced as described previously.^[21] 100 mM stock solutions of **sclx₈** (TCI S0471) were prepared in 20 mM potassium phosphate, 50 mM NaCl and adjusted to pH 6.0.

Isothermal titration calorimetry. Oxidized cytochrome *c* samples were dialyzed (3 kDa MWCO membrane) overnight at 4 °C against 20 mM potassium phosphate, 50 mM NaCl, pH 6.0. This buffer (post dialysis) was used to dilute 100 mM stocks of **sclx₈** to the required concentration. Samples were degassed and the protein concentration was determined immediately prior to the titration. Measurements were made at 25 °C using a Standard Volume NanoITC equipped with a Hastelloy cell (TA Instruments). Separate titrations of cytochrome *c* into buffer and **sclx₈** into buffer confirmed that the heats of dilution were small, exothermic and approximately constant. Standard and reverse titrations were performed by titrating **sclx₈** (530 - 2400 μM) into cytochrome *c* (40 - 170 μM) or cytochrome *c* (260 - 640 μM) into **sclx₈** (20 - 70 μM), respectively. The cell concentrations after each injection were calculated assuming that the solution exiting the active volume of the cell was expelled before any mixing occurred with the injected solution.^[10,41] The baseline around each pulse was approximated using a low order polynomial^[10,41] and the heat pulses were numerically integrated following baseline correction.

ITC data fitting. *pytc*^[31] was used to perform model fitting and parameter estimation. Two thermodynamic models were devised, informed by previous structural and biophysical data.^[11] The system of equations relating the independent variables of the models (total concentrations) to the experimental observable (heat generated during injections) is:

Model 1

$$\begin{aligned} [P_T]_i &= [P]_i + [PL]_i + [PL_2]_i + 4[P_4L_{n_L}]_i \\ [L_T]_i &= [L]_i + [PL]_i + 2[PL_2]_i + n_L[P_4L_{n_L}]_i \end{aligned} \tag{1a}$$

$$\begin{aligned} [PL]_i &= K_1[P]_i[L]_i \\ [PL_2]_i &= K_1K_2[P]_i[L]_i^2 \\ [P_4L_{n_L}]_i &= K_3[P]_i^4[L]_i^{n_L} \end{aligned} \tag{1b}$$

$$\begin{aligned}
q_i = V_{\text{cell}} & \left(\Delta H^\circ_1 ([PL]_i - [PL]_{i-1}(1 - v_i/V_{\text{cell}})) \right. \\
& + (\Delta H^\circ_1 + \Delta H^\circ_2) ([PL_2]_i - [PL_2]_{i-1}(1 - v_i/V_{\text{cell}})) \\
& \left. + \Delta H^\circ_3 \left([P_4L_{n_L}]_i - [P_4L_{n_L}]_{i-1}(1 - v_i/V_{\text{cell}}) \right) \right) + q_{\text{dil}}
\end{aligned} \tag{1c}$$

Model 2

$$\begin{aligned}
[P_T]_i &= [P]_i + [PL]_i + [PL_2]_i + 2[P_2L]_i + 2[P_2L_2]_i + 4[P_4L_4]_i \\
[L_T]_i &= [L]_i + [PL]_i + 2[PL_2]_i + [P_2L]_i + 2[P_2L_2]_i + 4[P_4L_4]_i
\end{aligned} \tag{2a}$$

$$\begin{aligned}
[PL]_i &= K_I [P]_i [L]_i \\
[PL_2]_i &= K_I K_{II} [P]_i [L]_i^2 \\
[P_2L]_i &= K_I K_{III} [P]_i^2 [L]_i \\
[P_2L_2]_i &= K_I^2 K_{IV} [P]_i^2 [L]_i^2 \\
[P_4L_4]_i &= K_I^4 K_{IV}^2 K_V [P]_i^4 [L]_i^4
\end{aligned} \tag{2b}$$

$$\begin{aligned}
q_i = V_{\text{cell}} & \left(\Delta H^\circ_I ([PL]_i - [PL]_{i-1}(1 - v_i/V_{\text{cell}})) \right. \\
& + (\Delta H^\circ_I + \Delta H^\circ_{II}) ([PL_2]_i - [PL_2]_{i-1}(1 - v_i/V_{\text{cell}})) \\
& + (\Delta H^\circ_I + \Delta H^\circ_{III}) ([P_2L]_i - [P_2L]_{i-1}(1 - v_i/V_{\text{cell}})) \\
& + (2\Delta H^\circ_I + \Delta H^\circ_{IV}) ([P_2L_2]_i - [P_2L_2]_{i-1}(1 - v_i/V_{\text{cell}})) \\
& \left. + (4\Delta H^\circ_I + 2\Delta H^\circ_{IV} + \Delta H^\circ_V) ([P_4L_4]_i - [P_4L_4]_{i-1}(1 - v_i/V_{\text{cell}})) \right) + q_{\text{dil}}
\end{aligned} \tag{2c}$$

Where,

$[P_T]_i$ is the total cell concentration of protein at the i^{th} injection (independent variable)

$[L_T]_i$ is the total cell concentration of ligand at the i^{th} injection (independent variable)

K_1, K_2, K_3 and $K_I, K_{II}, K_{III}, K_{IV}, K_V$ are the macroscopic association constants (fit parameters)

n_L is the ligand stoichiometry in the tetramer (fixed at 3, 4, 5, or 6)

$\Delta H^\circ_1 - \Delta H^\circ_3$ and $\Delta H^\circ_I - \Delta H^\circ_V$ are the enthalpies (fit parameters) associated with $K_1 - K_3$ and $K_I - K_V$, respectively

q_i is the heat generated from the i^{th} injection (dependent variable)

q_{dil} is the heat of dilution per mole of injectant (fit parameter)

V_{cell} is the volume of the cell

v_i is the volume of the i^{th} injection

Equations 1a and 2a are the mass conservation equations for Model 1 and 2, respectively, while

Equations 1b and 2b are the equilibrium equations. Numerical solution of these equations for each

model and each injection was performed using the Levenberg-Marquardt algorithm (implemented in SciPy) and estimates of the equilibrium constants to yield the free protein ($[P]_i$) and free ligand concentrations ($[L]_i$). The free concentrations were then used to compute the concentrations of the other states via the equilibrium equations. The calculated heat generated for each injection was then determined by Equation 3. Parameters were constrained to physically reasonable bounds ($K_1, K_2, {}^{(nL+nP-1)}\sqrt{K_3}$ between 10^2 and 10^{10} M^{-1} ; $\Delta H^\circ_1, \Delta H^\circ_2, \Delta H^\circ_3$ between -500 and 500 kJ/mol; q_{dil} between -1000 and 0 μJ) and best-fits were obtained by maximum likelihood starting from a range of initial estimates. Parameter uncertainties and correlations were estimated using the Bayesian MCMC (Figure S5). A uniform uncertainty in the measured heat of 20 μJ was applied to both titrations. A concentration correction parameter^[31,43] (fx) was used to model uncertainty in the **sclx₈** concentration. This parameter was allowed to float between 0.5 and 1.5 to account for an unknown residual water content in the **sclx₈** sample used to prepare the ligand stocks. The value of fx was 0.9 for the combined fit (Figure S5).

References

1. D. S. Goodsell, A. J. Olson, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105–153.
2. J. Janin, R. P. Bahadur, P. Chakrabarti, *Q. Rev. Biophys.* **2008**, *41*, 133–180.
3. A. Colosimo, M. Brunori, J. Wyman, *J. Mol. Biol.* **1976**, *100*, 47–57.
4. S. Vega, O. Abian, A. Velazquez-Campoy, *Methods* **2015**, *76*, 99–115.
5. N. Dotan, D. Arad, F. Frolow, A. Freeman, *Angew. Chem. Int. Ed.* **1999**, *38*, 2363–2366.
6. Q. Li, C. R. So, A. Fegan, V. Cody, M. Sarikaya, D. A. Vallera, C. R. Wagner, *J. Am. Chem. Soc.* **2010**, *132*, 17247–17257.
7. F. Sakai, G. Yang, M. S. Weiss, Y. Liu, G. Chen, M. Jiang, *Nat. Commun.* **2014**, *5*, 4634.
8. H. D. Nguyen, D. T. Dang, J. L. J. Van Dongen, L. Brunsveld, *Angew. Chem. Int. Ed.* **2010**, *49*, 895–898.
9. R. E. McGovern, A. A. McCarthy, P. B. Crowley, *Chem. Commun.* **2014**, *50*, 10412–10415.
10. M. L. Rennie, A. M. Doolan, C. L. Raston, P. B. Crowley, *Angew. Chem. Int. Ed.* **2017**, *56*, 5517–5521.
11. M. L. Rennie, G. C. Fox, J. Pérez, P. B. Crowley, *Angew. Chem. Int. Ed.* **2018**, *57*, 13764–13769.
12. E. F. Douglass, C. J. Miller, G. Sparer, H. Shapiro, D. A. Spiegel, *J. Am. Chem. Soc.* **2013**, *135*, 6092–6099.
13. B. Buxton, *J. Med. Res.* **1905**, *13*, 431–459.
14. A. Garcia-Pino, S. Balasubramanian, L. Wyns, E. Gazit, H. De Greve, R. D. Magnuson, D. Charlier, N. A. J. van Nuland, R. Loris, *Cell* **2010**, *142*, 101–111.
15. V. J. Streusand, I. Bjork, P. G. Gettins, M. Petitou, S. T. Olson, *J. Biol. Chem.* **1995**, *270*, 9043–9051.
16. A. R. Rezaie, *J. Biol. Chem.* **1998**, *273*, 16824–16827.
17. A. Levchenko, J. Bruck, P. W. Sternberg, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *91*, 5818–5823.
18. S. S. Deshpande, S. Damodaran, *J. Food Sci.* **1989**, *54*, 695–699.
19. J. W. Bye, N. P. Cowieson, A. J. Cowieson, P. H. Selle, R. J. Falconer, *J. Agric. Food Chem.* **2013**, *61*, 290–295.
20. C. A. Brautigam, *Methods* **2015**, *76*, 124–136.
21. R. E. McGovern, H. Fernandes, A. R. Khan, N. P. Power, P. B. Crowley, *Nat. Chem.* **2012**, *4*, 527–533.
22. J. M. Alex, M. L. Rennie, S. Engilberge, G. Lehoczki, H. Dorottya, A. Fizil, G. Batta, P. B. Crowley, *IUCrJ* **2019**, *6*, In press.
23. C. A. Hunter, H. L. Anderson, *Angew. Chem. Int. Ed.* **2009**, *48*, 7488–7499.
24. S. J. Darby, L. Platts, M. S. Daniel, A. J. Cowieson, R. J. Falconer, *J. Therm. Anal. Calorim.* **2017**,

127, 1201–1208.

25. A. Brown, *Int. J. Mol. Sci.* **2009**, *10*, 3457–3477.
26. R. J. Falconer, *J. Mol. Recognit.* **2016**, *29*, 504–515.
27. M. L. Johnson, *Methods Cell Biol.* **2008**, *84*, 781–805.
28. M. L. Rennie, R. L. Kingston, *bioRxiv* **2016**, 068577.
29. E. C. Ihms, I. R. Kleckner, P. Gollnick, M. P. Foster, *Biophys. J.* **2017**, *112*, 1328–1338.
30. K. E. Hines, *Biophys. J.* **2015**, *108*, 2103–2113.
31. H. Duvvuri, L. C. Wheeler, M. J. Harms, *Biochemistry* **2018**, *57*, 2578–2583.
32. R. E. Kass, A. E. Raftery, *J. Am. Stat. Assoc.* **1995**, *90*, 773–795.
33. A. M. Doolan, M. L. Rennie, P. B. Crowley, *Chem. Eur. J.* **2018**, *24*, 984–991.
34. K. Luke, D. Apiyo, P. Wittung-Stafshede, *Biophys. J.* **2005**, *89*, 3332–3336.
35. N. E. Olesen, P. Westh, R. Holm, *J. Colloid Interface Sci.* **2015**, *453*, 79–89.
36. S. Barranco-Medina, S. Kakorin, J. J. Lázaro, K. J. Dietz, *Biochemistry* **2008**, *47*, 7196–7204.
37. A. G. Kozlov, T. M. Lohman, *Methods Mol. Biol.* **2012**, *922*, 37–54.
38. A. Velazquez-Campoy, *Anal. Biochem.* **2006**, *348*, 94–104.
39. A. Brown, C. J. Robinson, J. T. Gallagher, T. L. Blundell, *Biophys. J.* **2013**, *104*, 1720–1730.
40. G. Gutiérrez-Magdaleno, M. Bello, M. C. Portillo-Téllez, A. Rodríguez-Romero, E. García-Hernández, *J. Mol. Recognit.* **2013**, *26*, 67–75.
41. G. D. Bailey, J. K. Hyun, A. K. Mitra, R. L. Kingston, *Structure* **2009**, *17*, 737–748.
42. V. Guérineau, M. Rollet, S. Viel, B. Lepoittevin, L. Costa, P. Saint-Aguet, R. Laurent, P. Roger, D. Gigmès, C. Martini, V. Huc, *Nat. Commun.* **2019**, *10*, 113.
43. C. A. Brautigam, H. Zhao, C. Vargas, S. Keller, P. Schuck, *Nat. Protoc.* **2016**, *11*, 882–894.

Acknowledgements

This research was supported by NUI Galway and Science Foundation Ireland (grants 13/ERC/B2912 and 13/CDA/2168 to PBC).