

Apparent non-statistical binding in a ditopic receptor for guanosine.

Asawin Likhitsup,[†] Robert J Deeth,[†] Sijbren Otto,[‡] Andrew Marsh^{†*}

[†]*Department of Chemistry, University of Warwick, Coventry, CV4 7AL, United Kingdom.*

[‡]*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2*

1EW, United Kingdom.

a.marsh@warwick.ac.uk

Supporting information

Part 1

Curve-fitting procedure

The curve-fitting routine was implemented using Microsoft Office Excel 2002 SP3 on a PC. The ‘Solver’ add-in option was used to minimize the chi-squared S value as discussed below. For best results, the solver option was set as: Max time: 100 sec; Iterations: 100; Precision: 0.0000000000001; Tolerance: 1%; Convergence: 0.0000000000001; Estimates: Quadratic; Derivatives: Central. For solving the mass balance, the ‘assume non-negative’ option was also used. The curve-fitting procedure was written separately for each experiment: (i) dimerisation of the receptors, (ii) self-association of guanosine **3**, (iii) benchmark against previously published data and (iv) binding between the receptors and guanosine taking into account the multiple equilibrium. Error limits expressed in the resulting association constants represent the standard deviation from propagated fitting errors as described in the main text.^{1, 2} The spreadsheets are shown below and also made available at <http://go.warwick.ac.uk/marshgroup>.

1. Dimerisation of 1 and 2.

The NMR titration data was fitted to the dimer model. The three diagnostic ¹H-NMR signals H-6, NH-1 and NH-2 were used for curve fitting simultaneously. For the Saunder-Hynes dimer model,³ a quadratic solution for [H] (eq. 9) was used. Given the equilibrium:



and an initial concentration of the receptor $[H]_0$, the following relationship between [H] and $[H]_0$ can be derived:

$$\text{Dimerisation constant: } K_{HH} = \frac{[HH]}{[H]^2} \quad (6)$$

$$\text{Mass balance: } [H]_0 = 2[HH] + [H] \quad (7)$$

$$\text{Substitute (7) in (6), rearrange: } 2K_{HH}[H]^2 + [H] - [H]_0 = 0 \quad (8)$$

Quadratic solution:
$$[H] = \frac{-1 \pm \sqrt{1 + 8K_{HH}[H]_0}}{4K_{HH}} \quad (9)$$

Using the ‘limiting chemical shift’ of the monomer δ_H and dimer δ_{HH} , the mole ratio of the monomer χ_H can be used to calculate the predicted chemical shift δ_{cal} for each signal

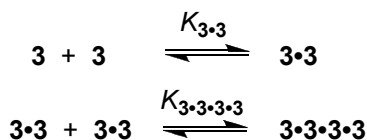
at each concentration:
$$\delta_{cal} = \delta_H \chi + \delta_{HH} (1 - \chi) \quad (10)$$

Curve fitting was done by varying K_{HH} , δ_H and δ_{HH} to minimize the sum of chi squared

(S) value which is defined as:
$$S = \sum_{i=1}^n (\delta_{cal} - \delta_{obs})^2 \quad (11)$$

2. Self-association of guanosine 3.

As described in the main text, the NMR titration data was fitted to the ‘dimer of dimers’ model. The three diagnostic $^1\text{H-NMR}$ signals N3-H, H-8 and NH_2 were used for curve-fitting simultaneously.



Mass balance:
$$[\mathbf{3}]_0 = [\mathbf{3}] + 2[\mathbf{3\cdot 3}] + 4[\mathbf{3\cdot 3\cdot 3\cdot 3}] \quad (3)$$

Chemical shift:
$$\delta_{cal} = \chi_3 \delta_3 + \chi_{3\cdot 3} \delta_{3\cdot 3} + \chi_{3\cdot 3\cdot 3\cdot 3} \delta_{3\cdot 3\cdot 3\cdot 3} \quad (4)$$

The curve-fitting routine starts by using an approximate value of $K_{3\cdot 3}$ and $K_{3\cdot 3\cdot 3\cdot 3}$ to generate a virtual equilibrium using an input concentration $[\mathbf{3}]_0$. The Excel solver function was then used to find the free guanosine concentration $[\mathbf{3}]$ that satisfies the mass balance [equation (3)] at each data point. Using an estimate value for limiting chemical shifts of free guanosine δ_3 , guanosine dimer $\delta_{3\cdot 3}$ and guanosine tetramer $\delta_{3\cdot 3\cdot 3\cdot 3}$, the calculated chemical shifts δ_{cal} for N3-H, H-8 and NH_2 were found from equation (4). All the limiting chemical shifts were then varied to minimize the S value (eq. 11) for each

pair of input $K_{3,3}$ and $K_{3,3,3,3}$. These variables were then manually iterated and the minimisation procedure was repeated to find the best fit which gave the smallest S value.

3. Benchmarking of Excel spreadsheets using Roelens *et al.* data.⁴

In order to validate the new Excel curve fitting procedure, a test set of data was selected from the literature that posed similar challenges to the data analyzed in our work in order to provide a benchmark. In developing a tripodal receptor for glycosides Roelens *et al.* observed self-association of the receptor, which when included in their subsequent analysis of guest glycoside binding using the HypNMR2004 package gave good fits of data to calculated titration curves. We thus took their NMR titration data from their supporting information⁴ and re-analysed it using our own Excel spreadsheet, obtaining firstly a receptor dimerisation constant K_{dim} of 56 M^{-1} (*c.f.* Roelens *et al.* $54 \pm 1 \text{ M}^{-1}$) and then equilibrium constants K_1 ($= 660 \text{ M}^{-1}$ (*c.f.* Roelens' $\beta_{11} = 663 \text{ M}^{-1}$) and $K_2 = 65\,000 \text{ M}^{-1}$ (*c.f.* Roelens' $\beta_{21} = 65\,614 \text{ M}^{-1}$) for binding of α -Gal to Roelens' dimerised receptor **1a**. With this data, alongside our own results we consider the use of Excel Solver function valid in this type of spreadsheet curve fitting routine.

4. Binding between the receptors and guanosine taking into account the multiple equilibrium.

The multiple equilibria considered are shown in **Figure 12** and **13**. Unfortunately, due to signal overlaps, the only diagnostic signal from NMR titration experiment was NH^b on the receptor. The dimerisation constants for the hosts $K_{1,1}$ and $K_{2,2}$ and the corresponding limiting chemical shifts were taken from NMR dilution experiment as described above. The self-association constants for guanosine $K_{3,3}$ and $K_{3,3,3,3}$ were taken

from NMR dilution experiment as described above. The curve fitting procedure was carried out by first estimating the two association constants $K_{2\cdot3}$ and $K_{2\cdot3\cdot3}$ as inputs. The mass balance was solved by non-linear least squares curve fitting, varying [2] and [3]. We then calculated δ_{cal} from the input δ_{H} , δ_{HG} and δ_{HGG} and these were varied in another non-linear least squares fitting routine to give the best possible fit for this pair of K_{HG} and K_{HGG} . The K_{HG} and K_{HGG} values were then iterated manually and the pair that gave the best fit (with reasonable values for δ_{H} , δ_{HG} and δ_{HGG}) was chosen.

Preliminary Isothermal Titration Calorimetry measurements.

Equilibrium constants, enthalpies and entropies of binding of **3** to **2** in chloroform were determined using isothermal titration calorimetry (MCS-ITC, Microcal LLC, Northampton, MA, USA) at 298 K. A solution of **3** (19.4 mM) in chloroform was titrated in 10 μL aliquots into solvent (control) or a solution of **2** (1 mM) made up using the same batch of chloroform. A solution of dialkyne **2** (10.5 mM) was likewise diluted into chloroform (**Figure S1.1**) revealing an association constant for dimerisation of 90 M^{-1} , somewhat different with that obtained from NMR dilution ($K_{2\cdot2}$ as $340 \pm 7 \text{ M}^{-1}$ in deuteriochloroform). Binding constants and enthalpies of binding were obtained by curve fitting of the titration data using the one-site binding model available in the Origin 2.9 software.

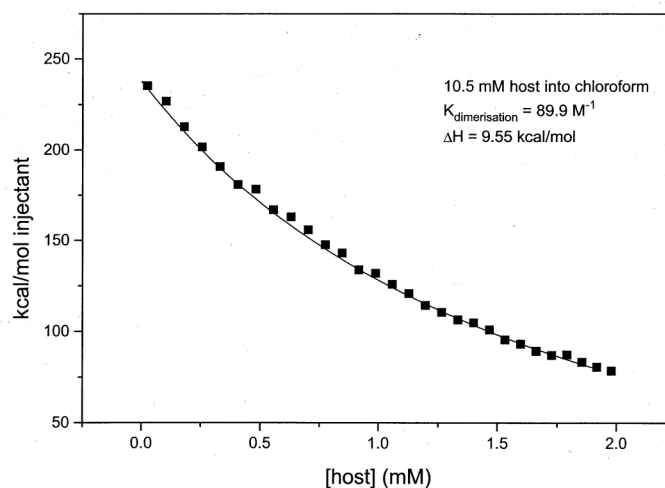


Figure S1.1: Dilution of dialkyne **2** (10.5 mM) into chloroform monitored by ITC. $K_{2,2} = 89.9 \text{ M}^{-1}$, $\Delta G^\circ = -11.2 \text{ kJ/mol}$, $\Delta H^\circ = -40 \text{ kJ/mol}$, $T\Delta S^\circ = -28.8 \text{ kJ/mol}$.

Isothermal titration calorimetry (ITC) was also carried out for the **2•3•3** system (**Figures S1.2 and S1.3**). During data analysis the data for the dilution of **3** into chloroform (carried out to account for guanosine self-association; **Figure S1.2**) was subtracted from the data for the titration of **3** into **2**. Subsequent curve fitting was performed assuming completely independent binding of **3** to the two binding sites of **2**; *i.e.* using an effective concentration of binding sites of 2 mM. The following results were obtained:

$$K = 1.80 \times 10^4 \text{ M}^{-1}$$

$$\Delta G^\circ = -24.3 \text{ kJ/mol}$$

$$\Delta H^\circ = -18.6 \text{ kJ/mol}$$

$$T\Delta S^\circ = 5.7 \text{ kJ/mol}$$

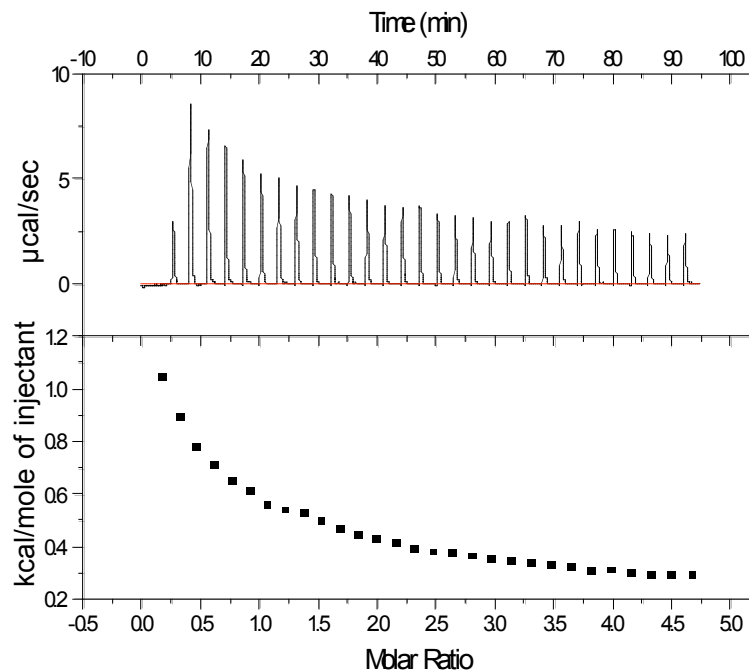


Figure S1.2: Titration of 29.4 mM **3** into chloroform at 298 K.

Comparing these values with K_1 and K_2 obtained from the NMR titrations requires a simple statistical correction, multiplying by a factor of 2. Thus $K_{1,ITC} = 3.60 \times 10^4 \text{ M}^{-1}$ and $K_{2,ITC} = 9.0 \times 10^3 \text{ M}^{-1}$. These values are approximately 6 times higher than those obtained by NMR. This discrepancy might be caused by the fact that subtracting the dilution experiment data, *while common practice*, is not strictly correct. During the dilution experiment a much higher concentration of unbound **3** is reached towards the end of the titration than during the binding experiment, where about half of the total amount of **3** that is added ends up in the complex with **2**. Moreover, subtracting the dilution data is not a correct way of filtering out a competing equilibrium and is more properly treated by simultaneous curve fitting, as for the NMR titration. This is apparent from a simple thought experiment: suppose the enthalpy change associated with the self-association of **3**

were zero; in this case the subtraction would not change anything, but that certainly does not mean that the competitive association of **3** does not take place. The ITC data analysis also ignores the self-association of receptor, although at a concentration of 1 mM only around 20% of **1** or **2** will self-associate.

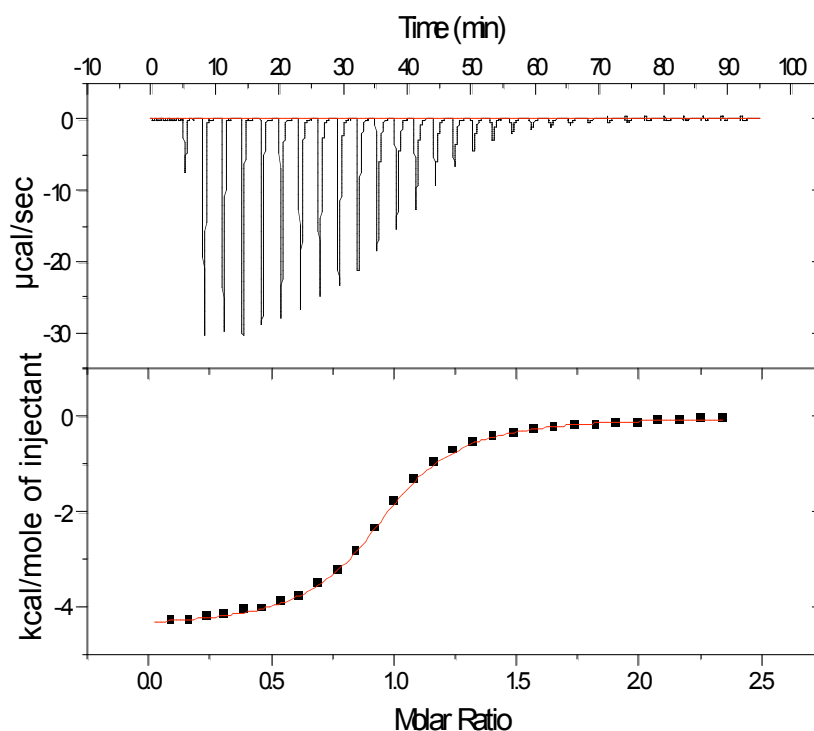


Figure S1.3: Titration of 29.4 mM **3** into a solution of **2** (1mM) in chloroform at 298 K. The molar ratio refers to the number of moles of **3** per number of binding sites.

The differing results probably stem from the exclusion of self-association in the ITC curve fitting itself. Thus, a correct description of all equilibria is crucial for the determination of stepwise association constants. It has been reported that the association constants from ITC are usually less than those obtained by spectroscopic techniques.⁵

References.

1. D. C. Harris, *J Chem Educ*, 1998, **75**, 119-121.
2. S. Walsh and D. Diamond, *Talanta*, 1995, **42**, 561-572.
3. M. Saunders and J. B. Hyne, *J. Chem. Phys.*, 1958, **29**, 1319-1323.
4. A. Vacca, C. Nativi, M. Cacciarini, R. Pergoli and S. Roelens, *J. Am. Chem. Soc.*, 2004, **126**, 16456-16465.
5. L. D. Williams, B. Chawla and B. R. Shaw, *Biopolymers*, 1987, **26**, 591-603.