The thermodynamics of guest complexation to octa-acid and tetra-*endo*-methyl octa-acid: reference data for the sixth statistical assessment of modeling of proteins and ligands (SAMPL6)

Matthew R. Sullivan[#], **Wei Yao**[#], **Bruce C. Gibb**
Department of Chemistry, Tulane University, New Orleans, LA, 70118, USA

[#] These authors contributed equally to this work.

Abstract

Although computer-aided drug design has greatly improved over time, its application in the pharmaceutical industry is still limited by the accuracy of association constant predictions. Towards improving this situation, the Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) is a series of community-wide blind challenges aimed to advance computational techniques as standard predictive tools in rational drug design (https://en.wikipedia.org/wiki/SAMPL_Challenge).

As an empirical contribution to the sixth assessment (SAMPL6), we report here the association constant (K_a) and thermodynamic parameters (G, H, $-T$ S) of eight guests (**G0**-**G7**) binding to two subtly different hosts (**OA** and **TEMOA**) using ITC. Both hosts contain a unique, well-defined binding pocket capable of storing guests with up to ten non-hydrogen atoms, whilst the selection of amphiphilic guests contain a range of saturated and unsaturated substituents from C6 to C10. The thermodynamic data from this study will allow the challenge participants of SAMPL6 to test the accuracy of their computational protocols for calculating host-guest affinities.

TOC Graphic
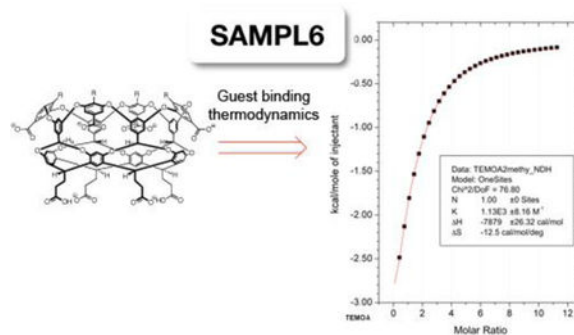
Keywords

Binding; cavitands; thermodynamics; Isothermal Titration Calorimetry

Introduction

The Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) is a series of community-wide blind challenges aimed to advance computational techniques as standard predictive tools in the rational design of drugs[1]. As part of the sixth statistical analysis (SAMPL6) we presented to the computational community two hosts (**OA** and **TEMOA**, Figure 1) and eight different amphiphilic guests (**G0**-**G7**) as part of a challenge to accurately determine the thermodynamics of host-guest complexation. The challenge segment now complete, we present here our ITC and NMR studies qualifying and quantifying the different binding events.

The hosts octa-acid (**OA**), and *tetra-endo*-methyl octa-acid **TEMOA**, are water-soluble deep-cavity cavitands that contain an exterior coating of eight carboxylic acid groups arranged in an anti-cube array, and a hydrophobic binding pocket well suited for association with a wide array of guests. Both hosts are obtained using similar synthetic approaches.[2, 3] The well-defined binding pocket of each host approximates to a truncated cone *ca.* 10 Å in depth, and with a width varying from 10 Å at the portal to 3 Å at its base. Within the walls of the pocket are four benzal protons (H_b; see label in host structures) that project into the cavity and act as weak hydrogen bond donors.[4] Four "upper" aromatic rings at the rim provide rigidity to the pocket, and have been shown to promote self-assembly of these hosts in the presence of non-polar guests.[5–7] The two hosts differ subtly at this upper rim. Specifically, **TEMOA** possesses four methyl groups at the rim of the pocket that project inwards and upwards. This array results in a slightly narrower cavity portal for **TEMOA**. Under the basic conditions of study, both hosts were assumed to be in the six-minus state, with two charged-hydrogen bonds between pairs of the 2-carboxyethyl pendent groups.

Under the basic aqueous conditions used for binding constant determinations, the guests chosen were all deprotonated and therefore strongly amphiphilic. This ensured that all of the guests bound with their alkyl or alkenyl groups filling the pocket of each host, and their (well solvated) carboxylate group at the portal of the complex. The strong solvation of the carboxylate ensured that neither host undergoes dimerization to form capsular complexes.

The alkyl or alkenyl groups of the guests ranged from C6 to C10 and contained a mix of fully saturated chains (**G0** and **G6**), or those with a single point of unsaturation (**G1**, **G3**-**G7**). Note that guests **G2**, **G4**, and **G5** are chiral. Guests **G2** and **G4** were used as pure enantiomers, whilst **G5** was used as a racemate mixture. However, as both hosts are themselves achiral and hence incapable of differentiating between enantiomers, there was no thermodynamic consequence using racemate **G5** for these studies.

Experimental

Full experimental details are given in the Supporting Information. Both hosts were synthesized following previously reported procedures.[2, 3] All reagents and guests were purchased from Aldrich and were used without purification. All NMR experiments were carried out in 50 mM phosphate buffer (pD = 12) prepared with D_2O (Cambridge Isotopes,

99.9%+). All NMR spectra were recorded on a Bruker 500 MHz spectrometer at 25 °C. Spectral processing was carried out using Mnova software (Mestrelab Research S.L).

Isothermal Titration Calorimetry (ITC) analyses of the binding events were performed in 50 mM phosphate buffer (pH = 11.5) at 25 °C using a VP-ITC MicroCalorimeter from Microcal, USA. Curve fitting of the binding isotherms were processed using ORIGIN 7.0 software included with the instrument.

Results and Discussion

Host-guest complexation was confirmed by ^{1}H NMR. The assignment of signals from the complexed guest was made possible by calculating the shift of each signal (δ) between the free (δ_{free}) and the bound (δ_{bound}) state, i.e., $\delta = \delta_{bound} - \delta_{free}$. This is of utility because the deeper a bound guest proton resides in the cavity, the more it is shielded and the greater its upfield shift relative to the free state. Combined with COSY NMR experiments revealing through-bond coupling of protons between adjacent carbons, the δ calculations reveal the orientation of the bound guest within the hosts.

Guest binding was either fast or slow on the NMR timescale. In the former, a solution of the host must be saturated with excess guest to ensure maximal signal shifts of the guest. Figure 2 shows the case of **G4**, where a two-fold excess of guest was used to ensure maximal upfield shifting of the bound guest signals. As can be seen, it is the terminal methyl groups that are shifted upfield the most. CPK models suggest that it is Me_2 can most easily be located at the base of the pocket, and therefore it is this signal that is shifted to –2.7 ppm. In contrast, Me_3 undergoes a much smaller shift because it is located nearer to the portal region of the pocket. Hence, the guest adopts an overall conformation with the carboxylate at the portal and the Me_2 anchoring the guest to the base of the pocket.

ITC was used to measure all principle thermodynamic parameters: G, H, and $-T S$. In these experiments the enthalpy of binding (H) is measured directly as the total amount of heat absorbed or released over the titration of excess guest into a solution of the host. Additionally, in each experiment the association constant (K_a) is determined by measuring the heat released or absorbed as a function of host-guest ratio and fitting the resulting data to a 1:1 binding model. These parameters obtained, the calculation of G and $-T S$ is straightforward.

For the complexation of **G5** and **G7** the resulting ITC data provided relatively low Wiseman parameters. As a result, the binding of these guests were analyzed as described by Turnbull[8] and Tellinghusen.[9] Furthermore, to ensure a measurable amount of heat from these systems, relatively high concentrations of guest (titrant) solutions were used. These high concentrations resulted in measurable heats of dilution, and as a result control experiments in which the guest was titrated into buffer solution alone were carried out. As an example, the binding data for guest **G5** complexing to **TEMOA** is shown in Figure 3. A standard ITC protocol was implemented for all of the remaining host-guest pairs. In all cases, the resulting data was successfully fitted to a 1:1 binding model.

The thermodynamic data obtained for each complexation event is shown in Table 1. Based on the current understanding of these systems, the greatest factors that affect binding affinity are the complementary relationship between host and guest as well as the changes that occur in solvation for both host and guest upon complexation.[10, 11]

We begin our discussion of guest binding by first considering the four C6 guests, before moving on to discuss the two C7 guests, and finally the remaining C10 guests. In all cases, at the temperature studied (25 ˚C) binding was exothermic, and with only two exceptions (**G0** and **G4** binding to **OA**), all complexations were entropically unfavorable. In both of these cases, the favorable contribution from entropy was only slight. An inspection of the data also reveals that with only one exception, the exothermicities of guest binding were stronger in the case of **TEMOA**. The one exception to this was **G2**. As we discuss below, we attribute this to the width of the cyclohexene moiety being incompatible with the narrow cavity rim of **TEMOA**.

C6 Guests: G1, G3, G5 and G6

The four C6 guests examined were **G1**, **G3**, **G5**, and **G6**. Two of these were straight-chain derivatives (**G1** and **G3**) that differed in the position of a double bond. A comparison of these two guests binding to **OA**, reveals that **G3** has a higher affinity: $K_a = 6.21 \times 10^3$ M^{-1} versus 2.57×10^3 M^{-1}. The affinities were reversed in the case of **TEMOA**, with **G1** binding more strongly. Indeed, the greatest difference between hosts in binding affinity amongst the hexanoate isomers was **G1**. The binding constant of **TEMOA-G1** was approximately an order of magnitude greater than for **OA**. This was due to a very sizable difference in the entropy of binding to **OA** versus **TEMOA**, namely: –5.52 and –8.25 kcal mol^{-1} respectively.

It is interesting to compare the affinity of **G3** for the previously reported data for the binding of *n*-hexanoate[12] and 5-hexynoate[13] to **OA**. The binding constants to **OA** for the alkyl, alkenyl, and alkynyl guests are $K_a = 6.75 \times 10^3$ M^{-1}, 6.21×10^3 M^{-1}, and 9.04×10^3 M^{-1} respectfully. Surprisingly, this does not provide a simple trend, but a close look at the individual contributions of H and –T S do. Thus, the respective enthalpic contributions are –5.13, –7.53, and –7.71 kcal mol^{-1}, from which we infer that with increasing p-character of the terminal CC bond there are more favorable interactions with the walls of the host (π-π stacking). Furthermore, the **OA**-hexanoate binding has a marginally favorable entropy of complexation (–T S = –0.1 kcal mol^{-1}), whilst the remaining two guests are both entropically penalized to the same degree (–T $S \approx 2.3$ kcal mol^{-1}). This indicates the more rigid the anchor of the guest, the greater the loss in degrees of freedom with packing the base of the cavity.

The second pair of C6 guests – **G5** and **G6** – are branched derivatives. Like **G3**, **G5** has a terminal double bond, but the main-chain is one carbon shorter and a methyl is present at the α-position. On the other hand, **G6** has a C5 main-chain with a geminal dimethyl terminus. Of all the C6 guests, **G5** was the only one to bind stronger to **OA** ($K_a = 2.33 \times 10^3$ M^{-1}) than to **TEMOA** ($K_a = 1.13 \times 10^3$ M^{-1}). Models suggest that α-substituents on a guest and

the rim methyls of **TEMOA** are co-localized in the complex, and we attribute this fact to the considerable entropic penalty in binding **G5** to this host.

The association of **TEMOA-G6** ($K_a = 9.12 \times 10^3$ M^{-1}) was twice as strong as **OA-G6** ($K_a = 4.37 \times 10^3$ M^{-1}). For both hosts, binding of **G6** is enthalpically dominated (**OA-G6**: –5.29 kcal mol^{-1} and **TEMOA-G6**: –6.19) with near-negligible entropic influence (**OA-G6**: 0.33 kcal mol^{-1} and **TEMOA-G6**: 0.79 kcal mol^{-1}). Compared to the methyl anchor of *n*-hexanoate ($K_a = 6.75 \times 10^3$ M^{-1}),[12] dimethyl terminated **G6** bound a little weaker, at $K_a = 4.37 \times 10^3$ M^{-1}. That noted, the enthalpy of complexation for these two guests were comparable: –5.29 to –5.13 kcal mol^{-1}, indicating that despite the very different termini the two guests fill the binding pocket with similar complementarity. The **OA-G6** was however slightly entropically penalized (0.33 kcal mol^{-1}) relative to the slightly favorable change in entropy for *n*-hexanoate (–0.1 kcal mol^{-1}). Hence, we view the two methyl groups of G6 – which can alternately fill the base of the host – being unable to compensate for the ensuing entropy cost of the greater width of the termini.

C7 Guests: G0 and G7

The binding of **G0** to **TEMOA** was twice as strong than to **OA**: $K_a = 2.81 \times 10^4$ and 1.47×10^4 M^{-1} respectively. The entropic contribution (–T S) to bind to **OA** was a slightly favorable –0.84 kcal mol^{-1}, compared to an unfavorable 1.77 kcal mol^{-1} for **TEMOA**. However, the H for **TEMOA** was a much stronger –7.85 kcal mol^{-1} versus 4.84 kcal mol^{-1}, which signifies a better complementarity between host and guest.

Conversely, **G7** formed a significantly stronger complex with **OA** ($K_a = 3.60 \times 10^4$ M^{-1}) than **TEMOA** ($K_a = 1.07 \times 10^3$ M^{-1}). Similar to **G0**, the enthalpic contribution for **TEMOA** was greater (–8.33 versus. –7.44 kcal mol^{-1}), but in this case, the value for –T S was almost four times greater (4.20 vs 1.23 kcal mol^{-1}). As with **G5**, the entropic penalty forming the **TEMOA-G7** complex is attributed to the co-localization of the rim methyl groups of the host and the *α*-substituents of the guest.

C10 Guests: G2 and G4

Both terpene guests, **G2** and **G4**, contain distinct anchor groups: methylethenyl and methylethylidene moieties respectively. The major difference between the guests lies however in their mid-section; cyclic **G2** is a rotund guest with two α-substituents, whilst **G4** is a more flexible guest with a β-methyl substituent. Guest **G2** is the strongest binder to OA ($K_a = 1.40 \times 10^6$ M^{-1}); the nine carbons of its chain being close to optimal for filling the cavity. It has the strongest enthalpy of complexation, but also the strongest entropy penalty of binding; presumably because the large chain can make multiple non-covalent contacts with the walls of the host, but in doing so guest packing is tight. **G2** is the second strongest binder to **TEMOA**, and again the enthalpy of complexation is strong. However, there is a considerable entropic penalty associated with binding. Conversely, complexation of more flexible **G4** to **TEMOA** is stronger than to **OA**: $K_a = 5.12 \times 10^5$ M^{-1} and 1.64×10^5 M^{-1} respectively. The complex between **G4** and **OA** exhibits favorable influence from both H and –T S (–6.92 and –0.19 kcal mol^{-1}), but the enthalpic contribution to the complexation

of **G4** to **TEMOA** (–8.87 kcal mol^{-1}) is sufficient to offset the corresponding entropic penalty (1.08 kcal mol^{-1}).

Conclusion

We have measured the thermodynamic data for the binding of eight guests to two hosts by Isothermal Titration Calorimetry. The association constants determined from this study covered a range of three orders of magnitude: from as low as $K_a = 1.07 \times 10^3$ M^{-1} (**G7-TEMOA**), up to 1.40×10^6 M^{-1} for **G2** binding to **OA**. Although the two hosts are very similar, they offer a unique thermodynamic profile for each guest within this series. In addition to the guests within this series, previously reported host-guest binding data was used to observe trends amid similar guests. This revealed that the hybridization state of the terminal atoms plays an important role in guest complexation. We look forward to inspecting the computational endeavors of the SAMPL6 participants to gain their insight into the binding of these hosts and guests.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Website https://en.wikipedia.org/wiki/SAMPL_Challenge

2. Liu S; Whisenhunt-Ioup SE; Gibb CLD; Gibb BC, An improved synthesis of 'octa-acid' deep-cavity cavitand. Supramol. Chem 2011, 23, 480–485. [PubMed: 21666831]

3. Gan H; Benjamin CJ; Gibb BC, Nonmonotonic Assembly of a Deep-Cavity Cavitand. J. Am. Chem. Soc 2011, 133, 4770–4773. [PubMed: 21401093]

4. Gibb CLD; Stevens ED; Gibb BC, C-H⋯X-R (X = Cl, Br, and I) Hydrogen Bonds Drive the Complexation Properties of a Nanoscale Molecular Basket. J. Am. Chem. Soc 2001, 123, 5849–5850. [PubMed: 11403638]

5. Gibb CLD; Gibb BC, Well-Defined, Organic Nanoenvironments in Water: The Hydrophobic Effect Drives a Capsular Assembly. J. Am. Chem. Soc 2004, 126, 11408–11409. [PubMed: 15366865]

6. Sullivan MR; Gibb BC, Differentiation of small alkane and alkyl halide constitutional isomers via encapsulation. Org. Biomol. Chem 2015, 13, 1869–1877. [PubMed: 25504155]

7. Jordan JH; Gibb BC, Molecular containers assembled through the hydrophobic effect. Chem. Soc. Rev 2015, 44, 547–585. [PubMed: 25088697]

8. Turnbull WB; Daranas AH, On the Value of c: Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? J. Am. Chem. Soc 2003, 125, 14859–14866. [PubMed: 14640663]

9. Tellinghuisen J, Isothermal titration calorimetry at very low c. Anal. Biochem 2008, 373, 395–397. [PubMed: 17920027]

10. Ben-Amotz D; Underwood R, Unraveling Water's Entropic Mysteries: A Unified View of Nonpolar, Polar, and Ionic Hydration. Acc. Chem. Res 2008, 41, 957–967. [PubMed: 18710198]

11. Ben-Amotz D, Water-Mediated Hydrophobic Interactions. Annu. Rev. Phys. Chem 2016, 67, 617–638. [PubMed: 27215821]

12. Wang K; Sokkalingam P; Gibb BC, ITC and NMR analysis of the encapsulation of fatty acids within a water-soluble cavitand and its dimeric capsule. Supramol. Chem 2016, 28, 84–90. [PubMed: 26997853]
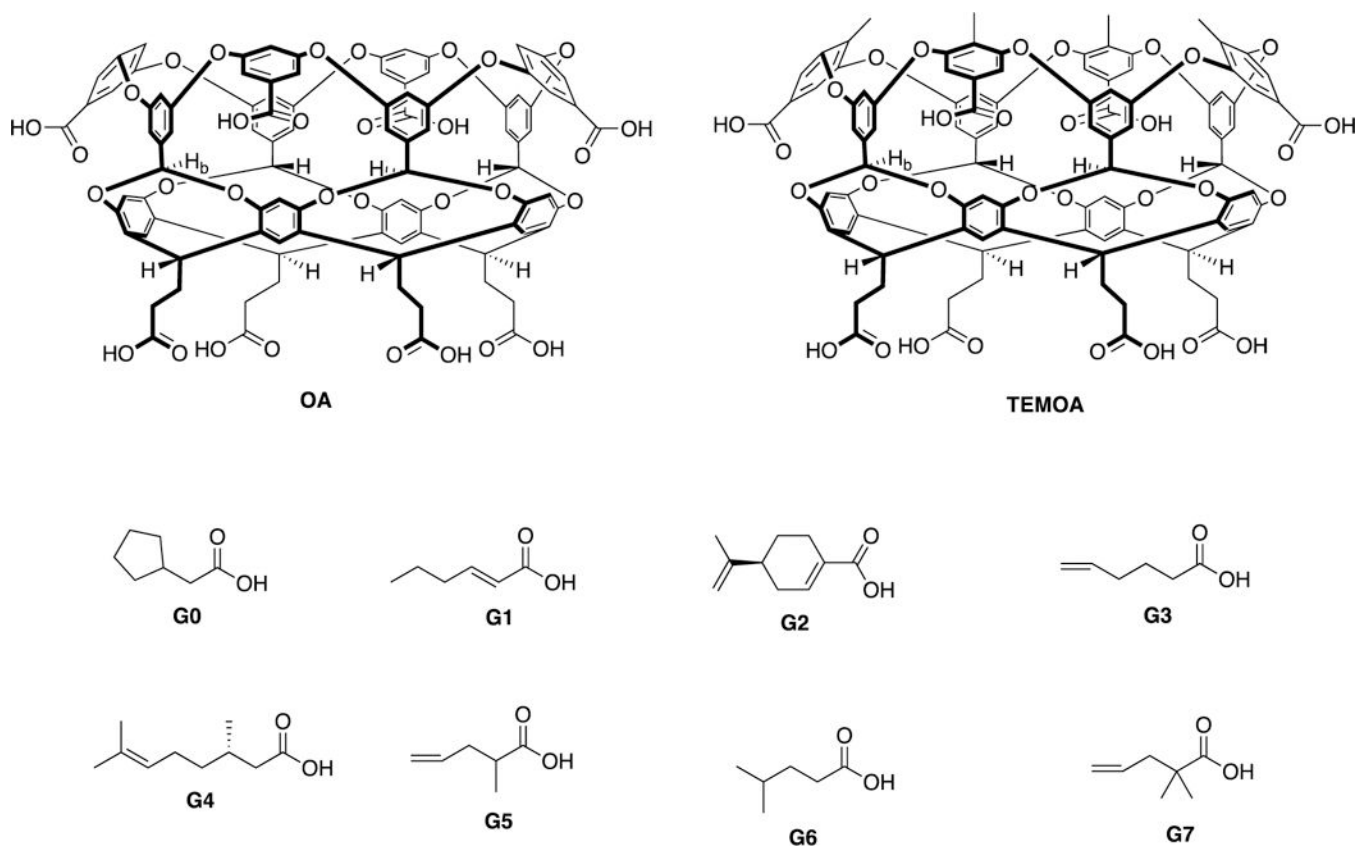
13. Sullivan MR, Sokkalingam P, Nguyen T, Donahue JP, Gibb BC Binding of carboxylate and trimethylammonium salts to octa-acid and TEMOA deep-cavity cavitands. J. Comput. Aided Mol. Des, 2016, 31 (SAMPL5 Special Issue), 21–8. [PubMed: 27432339]

Figure 1:
Chemical structures of hosts **OA** and **TEMOA** and guests **G0-G7**. Both hosts and guests are shown in their protonated, carboxylic acid state. Binding studies were performed under basic conditions in which the guests were fully deprotonated, and the hosts were assumed to be in the six-minus state (with two charged-hydrogen bonds between the pendent carboxylic acid/carboxylates).
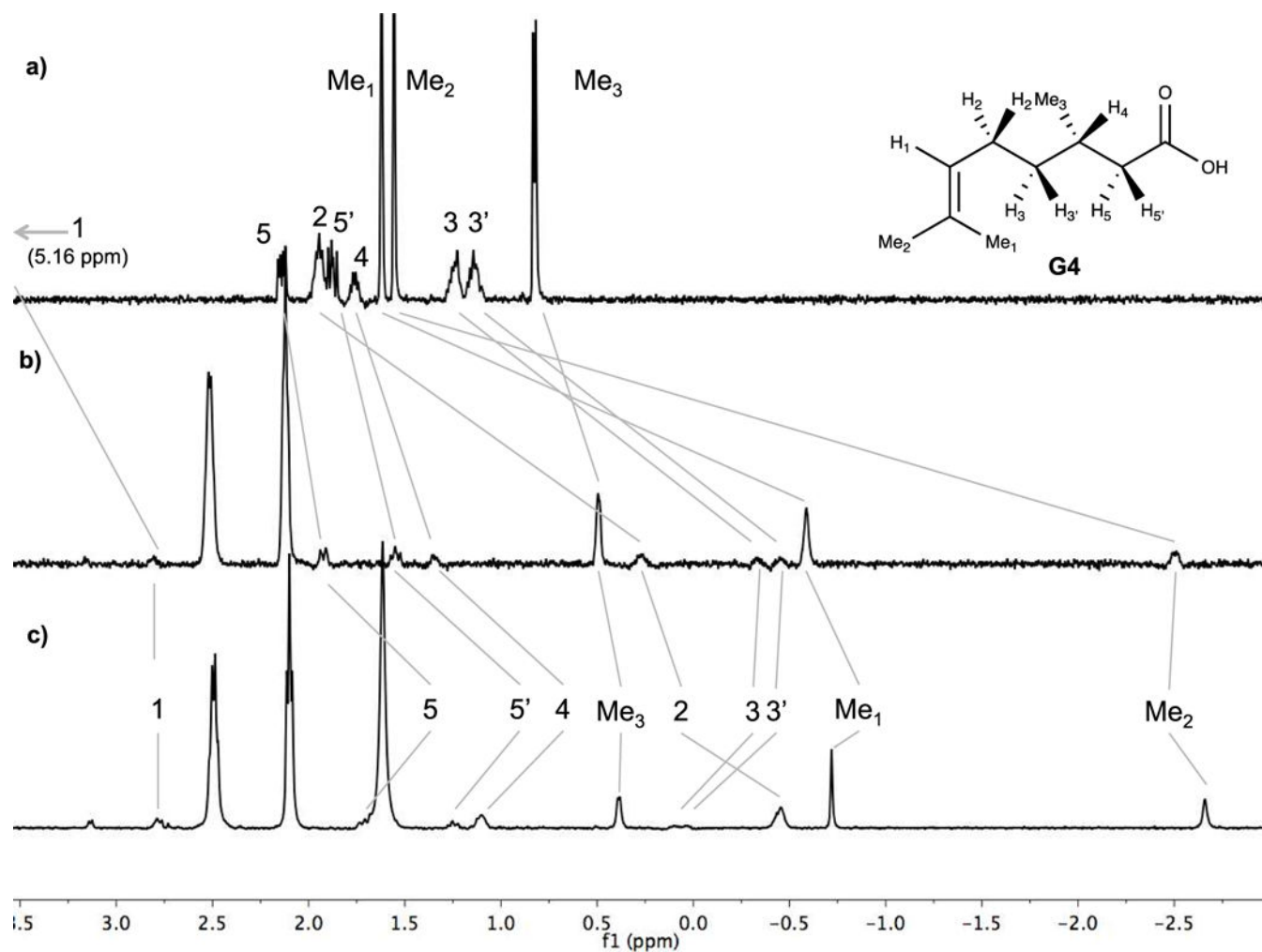
Figure 2:
^{1}H NMR of: a) free guest **G4**; b) a 2:1 host-guest ratio of **OA** and **G4**; c) a 2:1 host-guest ratio of **TEMOA** and **G4**. All solutions in 50 mM phosphate buffer in D_2O (pD = 12).
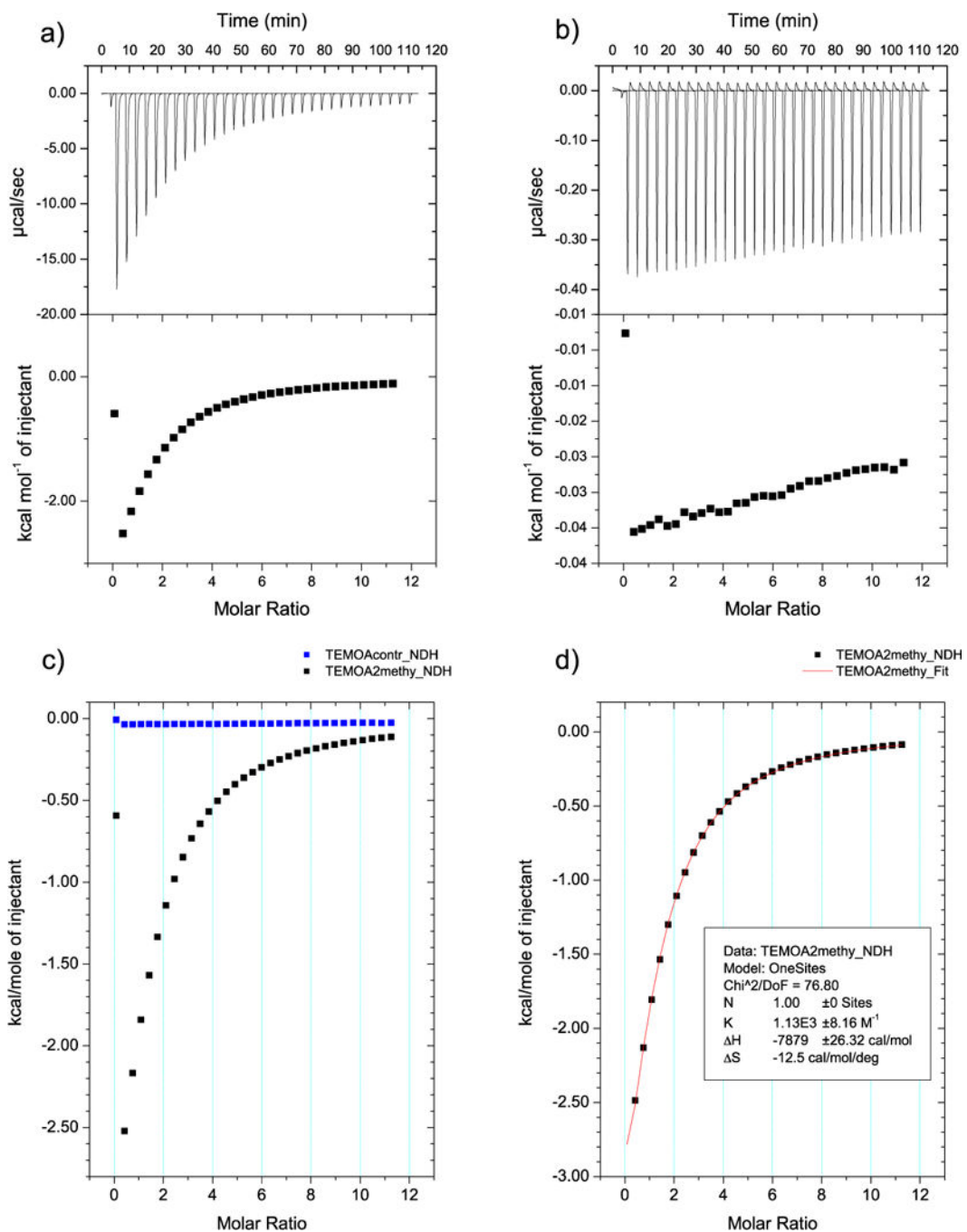
Figure 3:

Low affinity ITC data from **a)** complexation of **TEMOA-G5**; **b) G5** into buffer alone; **c)** an overlay of **a** and **b**; **d)** the final binding curve after subtraction of **b** from **a**. All solutions in 50 mM phosphate buffer in H_2O (pH = ~11.5).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1:

Thermodynamic data from ITC for the binding of guests **G0-G7** to hosts **OA** and **TEMOA**.[a]

Guests	OA				TEMOA			
	K_a^b M^{-1}	G^c kcal/mol	H^d kcal/mol	$-T S^e$ kcal/mol	K_a^b M^{-1}	G^c kcal/mol	H^d kcal/mol	$-T S^e$ kcal/mol
G0	1.47 E04	−5.68	−4.84	−0.84	2.81 E04	−6.06	−7.85	1.77
G1	2.57 E03	−4.65	−5.52	0.86	2.40 E04	−5.97	−8.25	2.27
G2	1.40 E06	−8.38	−12.07	3.69	9.82 E04	−6.81	−9.27	2.46
G3	6.21 E03	−5.18	−7.53	2.35	1.28 E04	−5.60	−8.86	3.25
G4	1.64 E05	−7.11	−6.92	−0.19	5.12 E05	−7.79	−8.87	1.08
G5	2.33 E03	−4.59	−5.31	0.71	1.13 E03	−4.16	−7.96	3.80
G6	4.37 E03	−4.97	−5.29	0.33	9.12 E03	−5.40	−6.19	0.79
G7	3.60 E04	−6.22	−7.44	1.23	1.07 E03	−4.13	−8.33	4.20

[a] 25 ˚C, 50 mM sodium phosphate buffer, pH = 11.5.

[b] Errors in K_a were < 4% in all cases (see SI for specific values).

[c] Calculated error in G is negligible due to the logarithmic relationship with K_a.

[d] Errors in H were < 4% in all cases (see SI for specific values).

[e] Errors in −T S were < 14% in all cases (see SI for specific values).