Organic & Biomolecular Chemistry

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 6885

www.rsc.org/obc

PAPER

Binding trimethyllysine and other cationic guests in water with a series of indole-derived hosts: large differences in affinity from subtle changes in structure†

Amanda L. Whiting and Fraser Hof*

Received 8th May 2012, Accepted 23rd June 2012
DOI: 10.1039/c2ob25882j

The binding of a series of indole-derived hosts to various ammonium cations in pure, buffered water is investigated using both solution phase ^{1}H NMR studies and computational modeling. These hosts can engage their targets *via* the cation–pi interaction, electrostatic attraction, and the hydrophobic effect. The hydrophobic effect is shown to be a dominant influence in the strength of the binding interactions, both in terms of the hydrophobicity of the host and of the guest. Our findings show that small changes that reduce the host hydrophobic surface area without reducing either the number of negative charges or amount of aromatic surface area are found to significantly decrease binding. Additionally, the position of solubilizing charges is also shown to influence the preferred host geometry and resulting binding constants.

Introduction

Molecular recognition arises from the contributions of many kinds of weak, intermolecular forces. In pure water, additional considerations such as solvation and the hydrophobic effect become important. Synthetic host–guest systems that function in pure water remain relatively underrepresented in the supramolecular literature.[1–4] While nature has perfected the art of encoding strong and selective binding in water, chemists still find it difficult to generate simple synthetic receptors that can reproduce this level of mastery. This is largely because the structure-function relations that have become relatively easy to predict in organic solvents still present a challenge to the designer who wishes to create a system that functions in pure water.

Host **1**, decorated with three indole rings appended with carboxylates, is a mimic of tryptophan-rich protein binding pockets that have evolved to engage quaternary and tertiary ammonium ions such as acetylcholine (ACh) and trimethyllysine (Kme3).[5–13] Preliminary studies of host **1** show that it binds organic ammonium cations in water, and that the hydrophobic effect plays a dominant role that far exceeds the energetic influence of, for example, cation–pi interactions, in determining its binding affinities and selectivities.[14] We sought to explore the extent of this effect in a series of related hosts – each containing three indole rings and three carboxylic acids attached to a central benzene ring but with variation in the exact positioning of functionality from one host to another (Fig. 1). All four hosts provide the same amount of aromatic surface area and the same number of negative charges to engage a cation.

In host **2**, we chose to retain the 3-propionic carboxylate tails on the indoles and append methoxy groups to the core benzene ring. It was thought by us and others[15] that the electron-donating ability of the oxygen atoms could contribute to an increase in the electron density on the core. This could potentially lead to an increase in binding due to an enhanced cation–pi effect at the central ring. Hosts **3** and **4** were obtained by shortening the propionate side chain of **1** to a simple carboxylate at the 3- and 2-position of the indoles, respectively, and were envisioned as less hydrophobic versions of **1**. The effect of removing the $-CH_2CH_2-$ linker could be determined directly by comparing **1** and **3**, while altering the position of the carboxylate (host **4**) would allow us to probe a different type of geometric variation within this family.

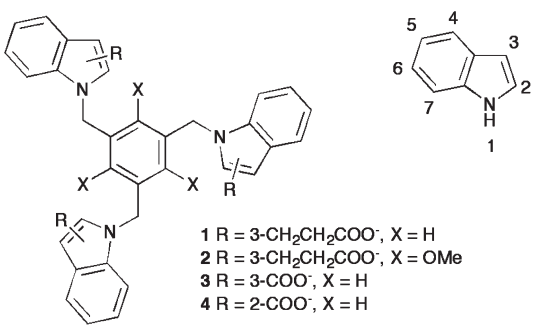


Fig. 1 Host studied in this work (indole numbering guide for reference).

University of Victoria, Department of Chemistry, PO Box 3065 STN CSC, Victoria, Canada, V8W3V6. E-mail: fhof@uvic.ca

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c2ob25882j

Results and discussion

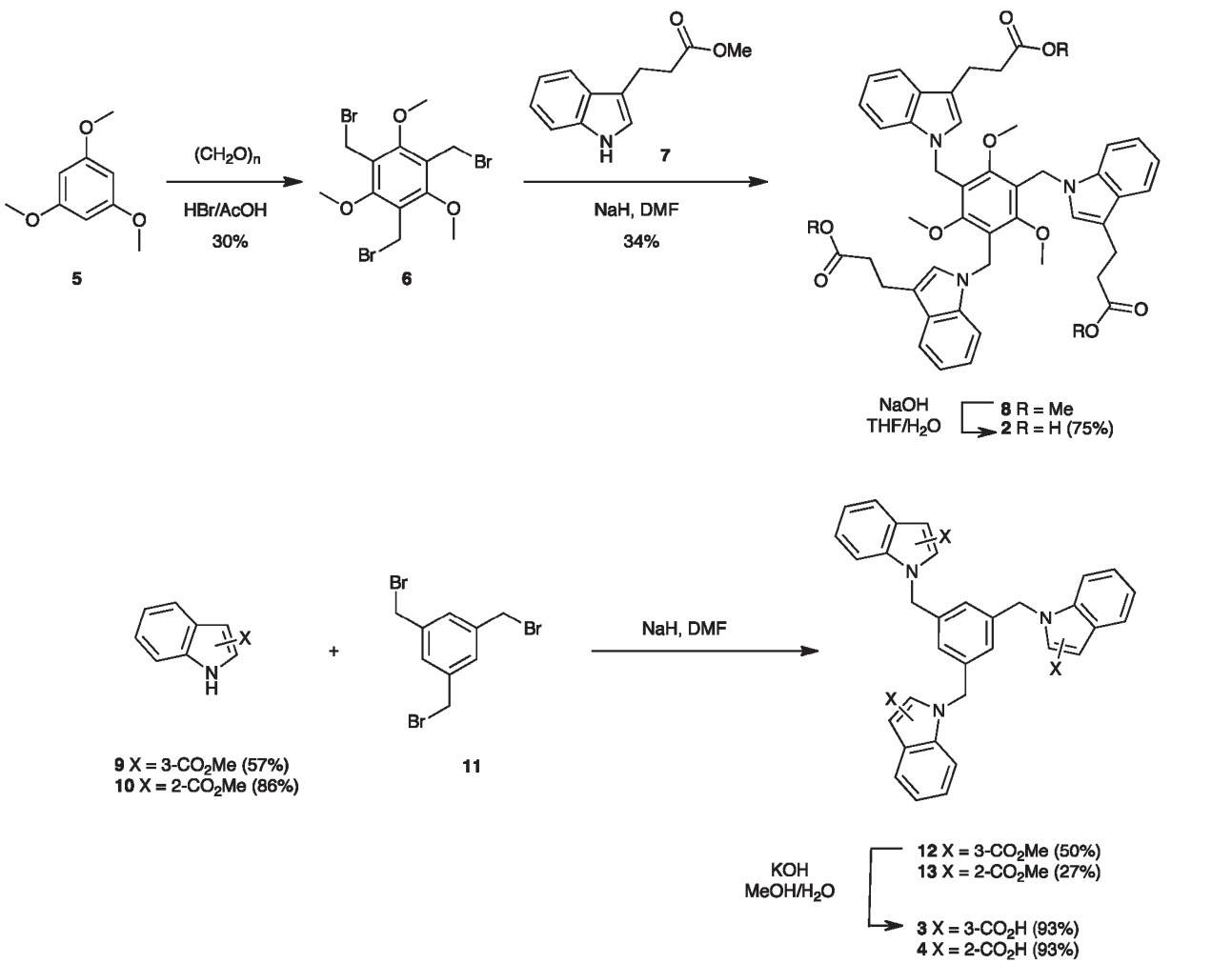
1. Synthesis

The synthesis of host **1** has been previously reported.[14] Host **2** is prepared starting with 1,3,5-trimethoxybenzene **5**, and treating with paraformaldehyde and HBr to provide the 1,3,5-tris(bromomethyl)-2,4,6-trimethoxybenzene core **6** in 30% yield (Scheme 1).[16] This methoxy core is then reacted with three equivalents of methyl indole-3-propionate **7** that has been pre-treated with NaH in DMF to give **8** in 34% yield. Deprotection of the methyl esters in basic aqueous solution gives the final host **2**. Hosts **3** and **4** are similarly synthesized starting with the protection of commercial 3- and 2-indole carboxylic acids to give methyl esters **9** and **10**. Coupling to 1,3,5-tris(bromomethyl)benzene **11** and deprotection with KOH in MeOH gives hosts **3** and **4**. All four hosts were isolated as carboxylic acids, and converted to their tri-sodium salts by treatment with stoichiometric NaOMe prior to their dissolution in buffered water for use in NMR studies. Our studies included both a detailed analysis of host conformations based on experimental and computational data, and NMR-based determinations of association constants for various ammonium ions of interest.

2. Host geometries in organic solution and water

Hosts **1–4** are inherently flexible due to the nature of the single bonds connecting the indole arms to the central benzene. As a result, they are likely able to adopt a number of conformations separated by low energetic barriers. This is especially true in organic solutions like dimethyl sulfoxide (DMSO) where the greasy aromatic elements of the hosts are well solvated. In water however, the geometries adopted will depend on the inherent bond rotational preferences of the host structures (as in DMSO) as well as aromatic clustering driven by the hydrophobic effect.[17,18] The preferred geometry prior to binding will not necessarily resemble an ideal open and bowl-like conformation.

We examined the aqueous host conformations without guests by comparing the chemical shifts of each host in DMSO to that in buffered D_2O (1 mM solutions) to see which protons on the hosts were most affected by the solvent change. The resulting spectra of the aromatic region protons of host **1** are shown in Fig. 2. Protons which were found to have ≥0.2 ppm change in chemical shift are highlighted in Fig. 3 (see Fig. S1–S5 in ESI† for additional full spectra and chemical shift data for all protons). Protons that significantly shift upfield upon the move to D_2O indicate a closer association with aromatic surfaces.



Scheme 1 Synthesis of hosts.

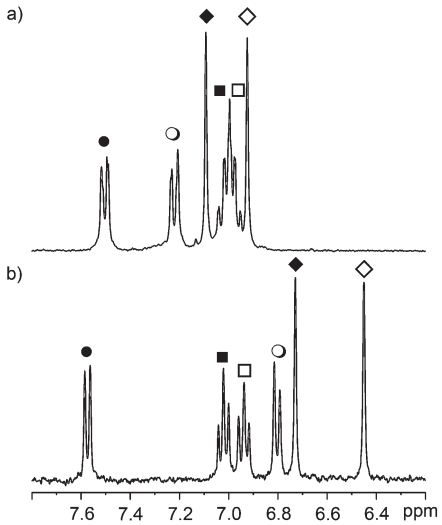


Fig. 2 Aromatic protons for host **1** (7.8–6.2 ppm) showing chemical shift upon solvent change in (a) DMSO and (b) H_2O. Proton identities as follows: 4-H (●), 7-H (○), 2-H (♦), 6-H (■), 5-H (□), and Ar–H (of central benzene ring, ◊).

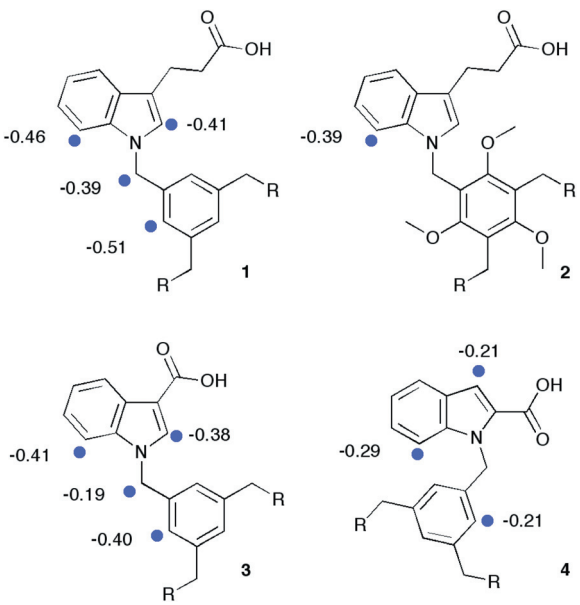


Fig. 3 Significant chemical shift changes from DMSO to D_2O. Magnitude of shift is indicated (ppm).

Hosts **1** and **3** each have several protons that move upfield to the same degrees, indicating a similar conformation is adopted for these two hosts in D_2O. Energy minimizations in implicit water (HF/6-31G* as implemented in Spartan '10)[19] suggest that the three indole rings are collapsed into a closed, propeller-like aromatic cluster, with each of the protons that are observed to be upfield-shifted in the NMR data located in the shielding cone of a neighboring aromatic ring (Fig. 4). The methoxy-containing host **2** has only one proton that experiences significant solvent-induced shifts (7-H indole) with all other shifts being negligible. Energy minimizations again suggest a propeller-like

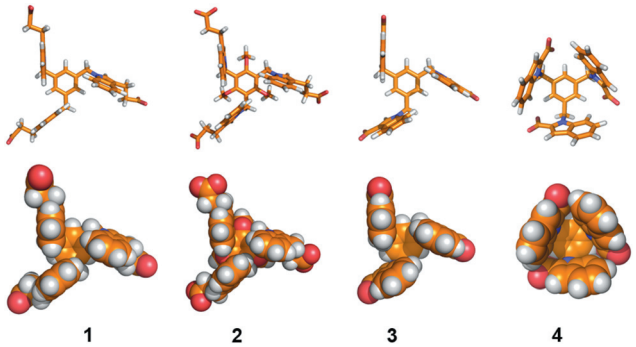


Fig. 4 Equilibrium geometry in implicit water for hosts **1**, **2**, **3** and **4** (HF/6-31G* as implemented in Spartan '10).[19]

conformation for **2** that is consistent with the (more limited) experimental data available for this host.

Host **4**, with a 2-carboxy substituent, shows overall smaller changes in chemical shift than the other hosts, but follows a similar pattern to the shifts observed in hosts **1** and **3** again indicating a closer association with aromatic surfaces. Interestingly, the energy minimized structure for **4** in implicit water suggests an open conformation, unlike the other hosts in this series. If accurate, the open conformation of **4** (Fig. 4) would be expected to have downfield chemical shifts for the 6-H and 7-H indole protons arising from CH···O contacts with the carboxylates. Experimentally, this was observed for proton 6-H on host **4** which at 7.08 ppm in D_2O and 7.14 ppm in DMSO-d_6, is located furthest downfield of all hosts in both water and also DMSO (see Fig. S1–S4† for NMR spectra).

Further NMR evidence that helps us to draw conclusions about the conformational changes experienced by all four hosts in DMSO and in buffered water comes from 1D ROESY experiments carried out in each solvent. In DMSO, key ROEs indicating through-space interactions are noted between the N–CH_2 methylene protons and indole H-7 for all hosts ("indole-out," Fig. 5a), and between N–CH_2 and indole H-2 for hosts **1**, **2** and **3** ("indole-in," Fig. 5b; host **4** has no H-2 proton). These two close contacts are mutually exclusive and occur in two different conformations that differ only by rotation about the indole N–CH_2 single bond. We conclude that the ROE data in DMSO are best explained by the rapid exchange between both conformations. In buffered water, hosts **1**, **2**, and **3** continue to show N–CH_2 to H-2 contacts that arise from the "indole-in" conformation shown in Fig. 5b, but show no evidence of contacts between H-7 and N–CH_2 that would arise from the "indole-out" conformation shown in Fig. 5a. Of note is that the "indole-in" rotamer, with H-7 closely packed over the central ring, is consistent with the calculated closed, propeller-like structures for hosts **1–3** shown in Fig. 4. Host **4** maintains H-7 to N–CH_2 contacts in D_2O, as well as showing H-7 – ArH contacts that are consistent with the calculated open conformation for **4** shown in Fig. 4.

Overall, the chemical shift, computational, and ROE data suggest geometries for hosts **1–3** in which the aromatic rings are further compressed onto each other and the charged carboxylate arms are mostly directed outwards into solution. This is consistent with the expected influence of the hydrophobic effect on these radially amphiphilic molecules. The collapsed form in

water also allows for stronger edge-to-face aromatic interactions between the rings than would be present in DMSO, where the hosts maintain significant mobility.

3. Solution-phase binding studies

The binding properties of hosts **1–4** were probed by NMR studies in phosphate-buffered D_2O (Table 1). Dilution titrations conducted on each host alone gave concentration-dependent chemical shifts that were fit to determine a self-association constant for each host ($K_{dimerization}$). Subsequent titrations with various ammonium ions provided data that could be fit to a 1 : 1 host–guest binding isotherm to determine K_{assoc} while taking into account the effects of host homodimerization using HypNMR[20] (see ESI† for experimental and calculation details).

Similar trends in chemical shifts upon binding were observed across the hosts irrespective of guest. Host signals identified in Fig. 3 as being most shielded (upfield) in water (Ar–H core, N–CH_2, 7-H indole and 2-H indole) were generally found to have downfield chemical shifts when a cationic guest was introduced. This suggested an opening of the aromatic rings of the indoles upon binding. The greatest downfield change in chemical shift ($\Delta\delta$) was observed for the protons closest to the benzene core: Ar–*H* (for hosts **1**, **3** and **4**) and Ar–O*CH*$_3$ (for host **2**). These shifts are consistent with a cationic guest positioned directly above the central ring, interacting both with it and the now face-on indole rings in an ideal bowl or "cage-like" conformation.

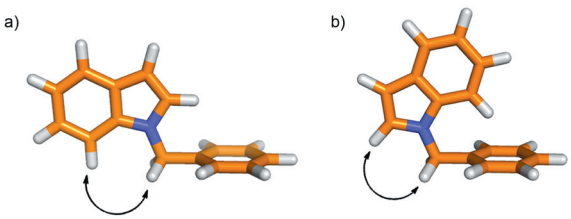


Fig. 5 Key ROESY interactions arising from the (a) "indole-out" rotamer (H-7 to CH_2 contact) and (b) "indole-in" rotamer (H-2 to CH_2 contact). The "closed" indole rotamer in 5b is the same as that adopted by all three indoles of hosts **1–3**, as shown in Fig. 4.

Table 1 Binding affinities for host **1–4** in phosphate-buffered D_2O[a]

Entry		K_{assoc} (M^{-1})			
		1	2	3	4
1	$K_{dimerization}$	330 ± 50	22 ± 6	22 ± 2	<1
2	NMe_4Cl	40 ± 15	27 ± 12	32 ± 14	22 ± 1
3	$BnNMe_3Cl$	47 ± 3	44 ± 1	41 ± 2	55 ± 1
4	nBuNMe$_3$I[b]	100 ± 20	35 ± 10	40 ± 10	22 ± 0
5	AChCl	120 ± 8	44 ± 1	30 ± 6	24 ± 0
6	Kme_3Cl	250 ± 9	22 ± 19	55 ± 27	24 ± 1
7	NEt_4Cl	180 ± 10	55 ± 15	80 ± 85	28 ± 1
8	NPr_4Cl	1100 ± 210	60 ± 20	70 ± 30	28 ± 1
9	NBu_4Cl	7060 ± 2100	90 ± 10	145 ± 55	23 ± 1
10	$HNMe_3Cl$	96 ± 10	8 ± 8	20 ± 10	80 ± 9
11	H_2NMe_2Cl	50 ± 20	n.b.[c]	26 ± 33	50 ± 20
12	H_3NMeCl	16 ± 27	n.b.	n.b.	n.b.

[a] Phosphate buffered D_2O (50 mM Na_2HPO_4/NaH_2PO_4) at pH 7.0 (pD 7.4). $K_{dimerization}$ is defined as the association constant (β_{HH}) for host dimerization while K_{assoc} is defined as β_{HG}, the association constant for the 1 : 1 host–guest complex. In all cases, $K_{dimerization}$ was first determined *via* dilution titrations of host alone, and then used as a constant in the fitting of host–guest titration data to determine the true value of K_{assoc}. See ESI for details.† [b] The iodide salt of this guest was used. The identity of counter ion for NMR studies in water has been shown to be negligible.[21] [c] n.b. = no binding.

3.1 Comparison of 1 *vs.* 2 – effect of methoxy substituents.

Host **1** exhibits a significant range in its association constants. For those guests containing a trimethyllysine-like R-NMe_3^+ group (entries 2–6) there is variation ranging from 40 M^{-1} to 250 M^{-1}. The simplest binding of a quaternary ammonium ion would be tetramethylammonium (entry 2); this could be considered a baseline of interaction with these types of hosts. For guests containing a single, simple alkyl substitution (nBuNMe$_3$I, entry 4), binding to host **1** in water becomes stronger as the length of the non-polar alkyl chain increases (from one carbon in NMe_4^+ to four carbons in nBuNMe$_3^+$). Given that the overall charge between these guests has not changed, this reflects an increased contribution from the hydrophobic effect. When the single alkyl substitution contains heteroatoms, such as the ester in AChCl (entry 5) and the amino acid head group in Kme3 (entry 6), the binding affinity to host **1** is again increased. While it is likely that there is an increased hydrophobic force associated with these larger guests, we cannot rule out additional polar interactions between host and guest. Binding to Kme3, for example, is significantly stronger than the other guests of this type, leading us to suggest additional favourable electrostatic interactions between the ammonium head group of Kme3 and the host. Interestingly, the aromatic quaternary ammonium $BnNMe_3^+$ (entry 3), does not result in a significant increase in the binding constant compared to NMe_4^+.

To examine the binding of larger quaternary ammonium ions, we increased the length of all four alkyl substituents across the series from methyl to *n*-butyl (entries 7–9). We observed a significant increase in binding, this time exclusively due to an increased hydrophobic contribution. The largest guest, NBu_4^+, had the strongest effect on the host propionate chains, shifting the methylene protons 0.1 ppm downfield and indicating the strongest involvement of the alkyl chain in binding. These small shifts are most consistent with the small deshielding influence of a cationic alkylammonium binding partner. Binding to primary, secondary and tertiary ammonium ions, (entries 10–12) showed weaker binding constants that diminished as methyl groups were removed. This was attributed to increased competition with the aqueous environment that could better solvate these guests as the number of hydrogen bonding NH's increases.

The addition of methoxy substituents to the benzene core in host **2** was intended to increase the electron-density available in the pi system and strengthen the resulting cation–pi interaction. However, the methoxy substituents were found to have significantly diminished binding relative to host **1**, especially with hydrophobic guests. A comparison of dimerization constants revealed **2** to be an order of magnitude less than **1** (entry 1). Presumably, the methoxy groups contribute to steric gearing[22] and can bias the indole arms of **2** to favor the closed "propeller" conformation (Fig. 4) regardless of solvation effects or the presence

of guests. This is consistent with the observation that host **2** has the fewest protons changing chemical shift upon moving from DMSO to water – its conformation in both solvents is fixed into the propeller shape. This also decreases the ability of **2** to associate with larger, hydrophobic guests (entries 7–9) that require an open and flexible host conformation. Similar diminished affinities (relative to **1**) are seen with almost all of the guests studied. It is clear that the added methoxy groups do not favor a conformation suitable for guests binding.

3.2 Comparison of 1 *vs.* 3 – effect of methylenes in propionate chains. Host **3**, with the water-solubilizing carboxylates attached directly to the indole 3-position, was studied in order to observe the effect that the $–CH_2CH_2–$ units in the indole arms of **1** would have on the strength of its binding interactions. Guest binding to host **3** in all cases is diminished relative to **1** (Table 1), and the effects are largest for the most hydrophobic guests. This is consistent with the assertion that the hydrophobic effect plays a dominant role in driving cation binding to **1**.

Interestingly, host **3** binds all quaternary R-NMe_3^+ guests with approximately the same affinity (entries 2–6), regardless of their degree of hydrophobic character. Given the similarity in structure, the aromatic skeletons of hosts **1** and **3** should have very similar conformational preferences and abilities to bind guests through their identical aromatic rings. Thus, for guests that do not cause a large hydrophobic effect, the magnitude of binding should be similar. This is clearly evident for guests with little hydrophobic contribution such as NMe_4^+ (entry 2; 40 M^{-1} *vs.* 32 M^{-1}, for **1** and **3**, respectively). For greasier guests where the hydrophobic effect is likely to be larger, the disparity in binding between **1** and **3** increases to a 50-fold difference for the largest cation, NBu_4^+ (entry 9; 7060 M^{-1} *vs.* 145 M^{-1}). The difference in surface area per $–CH_2CH_2–$ group between **1** and **3** is calculated to be 41 $Å^2$. Assuming that *all* surface area of the CH_2CH_2 in **1** is buried upon complexation (which must be an over-estimate of the surface area actually involved), this would lead to a prediction of 10 kcal mol^{-1} difference in binding energy (3.3 kcal mol^{-1} × 3 $–CH_2CH_2–$) between the two,[23] which is larger than that observed for NBu_4^+ (2.3 kcal mol^{-1}). The small downfield shifts of $–CH_2CH_2–$ protons upon binding NBu_4^+ are consistent with the formation of close contacts with the large cationic guest. Thus, although the deletion of the CH_2CH_2 groups seems intuitively to be a modest change of a peripheral group, the observed large changes in affinities are best interpreted as arising from the participation of those arms as important hydrophobic binding elements.

3.3 Comparison of 3 *vs.* 4 – effect of carboxylate position. We were also interested in what effect geometry would have on tripodal hosts of this nature. In all cases, the three indole rings need to adopt an open conformation to allow a cation to interact in a "face-on" manner with all four aromatic surfaces. Hosts **3** and **4** are structural isomers; they have the same number and identity of atoms and therefore the same degree of hydrophobicity. Their difference lies only in the position of the carboxylate and thus, the preferred geometry of the molecule. Experimentally, host **4** has no tendency to self-associate in water and binds all quaternary cations with approximately the same value regardless of size, shape or hydrophobic character. All binding constants are approximately the same as for NMe_4^+, the baseline guest molecule. This differs from host **3**, which has a small dimerization constant and a weak but measurable preference for hydrophobic guests. For host **4**, binding with $BnNMe_3^+$ stands out as having the highest binding constant for guests with the R-NMe_3^+ motif.

In looking at Fig. 4, carboxylates at the 2-positions of host **4** have the effect of drastically changing the geometric preferences of the indole arms compared to host **3**. Host **4** also has the smallest change in proton chemical shifts of all the similar, unrestricted hosts (**1**, **3** and **4**) indicating that it likely adopts a geometry that is highly similar in aqueous and organic solution. By moving the polar carboxylate groups to the indole-2 position, we have effectively excluded them from the binding pocket making them less able to participate in any electrostatic interaction with an incoming cation. For smaller R-NMe_3^+ guests that require little adjustment of the binding pocket, this translates into slightly smaller binding constants than for host **3** that could use its carboxylates in a favourable manner on the upper rim of the binding pocket. As we branch into larger quaternary ammonium cations, there is little change in binding as now host **4** must disrupt its preferred conformation and adjust to accommodate a larger guest. So while there is an increase in the hydrophobic association between host and the larger guests, the entropic penalty associated with opening the host binding pockets could negate any gain in binding affinity.

4. Calculated binding geometries

Energy minimizations in water were also used to evaluate the geometry of each host upon guest binding. For the typical guest NMe_4^+, the collapsed conformations of the empty hosts in each case convert to a more open conformation capable of making favourable contacts with guests by rotating the indole arms about the N–C single bond (Fig. 6). The process of creating a pocket for the guest changes the chemical environment experienced by the host protons as shown in the solution phase data. Protons which were most shielded in the collapsed form are now deshielded as they are moved out of shielding regions. In addition, the aromatic rings are now involved in interacting with a cation. While this would shield any protons attached to the cation, the aromatic rings themselves (and thus their protons) experience some deshielding (or a return to the more normal deshielded state).

The models of hosts **1** and **2** indicate that the methylene region of the propionate arms come into contact with cationic R-NMe_3^+ region of the guests. While the resulting change in NMR chemical shift might not be great for these protons, this

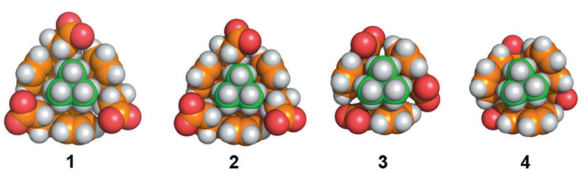


Fig. 6 Equilibrium geometry in implicit water for **1**, **2**, **3** and **4** with NMe_4^+ (HF/6-31G* as implemented in Spartan '10).[19]

would be an additional weak interaction that would give these hosts an advantage over the others. Indeed, the largest observed shift for these protons was only 0.1 ppm downfield and only for host **1** with NBu_4^+, the most hydrophobic guest. Still, the increased hydrophobic surface area was shown to play a significant role in the overall binding strength of these hosts given the differences in binding affinity, especially between hosts **1** and **3**. As well, these models show that the carboxylate groups of these hosts are associated with the upper rim of the host pocket, in position to interact favourably with cationic guests.

The models of hosts **3** and **4** demonstrate the physical difference of shifting the carboxylate from position 3 to 2. The carboxylates in **3** are more exposed, similar to those in hosts **1** and **2** and likely able to weakly participate in electrostatic interactions. In host **4**, those polar regions are buried in the base of the host cavity and less able to make contact with a cationic guest.

Conclusions

Within this series of highly similar indole-based hosts, two factors stand out as having the biggest influences on the strength of the host–guest interaction: (1) the amount of hydrophobic surface area available for binding (particularly for hydrophobic guests) and (2) the effect of small changes in functional group position on both geometric preference and binding affinity. The complementary and beneficial result of having additional hydrophobic regions on both host and guest is understandable as a consequence of a strong hydrophobic contribution within the aqueous environment. The more subtle effect of a small change in polar group position leading to a weaker electrostatic interaction was less expected and demonstrates our lack of ability to predict the outcome of such changes.

Though these hosts exhibit weak-to-moderately strong attraction to cations in water, there are important lessons here for future host designs targeting biologically relevant cations in water. Though we normally think of trying to build hosts that take advantage of as many attractive interactions as possible (hydrogen bonding and cation–pi interactions, for example), the complications that can be introduced by carrying out studies in pure water should not be underestimated. The current examples show that, in pure water, even subtle changes in pendant solubilizing groups can have dramatic effects on host–guest affinities. While past studies have been concerned with the possible role of, for example, an anionic water-solubilizing group *vs.* a neutral water-solubilizing group,[24] the current results show that even apparently innocent changes in a small number of linker CH_2 groups can have a strong and determinant role in guest binding. Related structural elements in proteins, such as the methylene stretches of lysine, arginine, and glutamate side chains, are increasingly appreciated to be binding elements that play a strong role in determining the structures and interactions of folded proteins.[25–29] Designs for aqueous host–guest systems will need to continue to consider the hydrophobic aspects of small functional group changes as a key driver of their molecular recognition performance.

Experimental section

General considerations

Proton (^{1}H) NMR and carbon (^{13}C) NMR spectra were recorded on a Brüker AC300 (300 MHz) spectrometer, a Brüker AVANCE360 (360 MHz) and a Brüker AVANCE500 (500 MHz). Proton (^{1}H) NMR spectra for NMR titration studies were recorded on a Brüker AVANCE360 (360 MHz) spectrometer. Chemical shifts (δ) are given in parts per million (ppm) and referenced to residual protonated solvent ($CHCl_3$: δH 7.26 ppm, δC 77.16 ppm; $(CH_3)_2SO$: δH 2.50 ppm, δC 39.52 ppm). J values are given in Hz. Abbreviations used are s (singlet), d (doublet), t (triplet), q (quartet), and m (multiplet). Infrared spectra were measured on a Thermo-Nicolet Nexus 670 FT-IR spectrometer with a resolution of 2 cm^{-1} using a Pike MIRacle attenuated total reflection (ATR) sampling accessory. Melting points were obtained using a Gallenkamp Melting Point Apparatus and are uncorrected. Accurate mass determination (HR-ESI-MS) was performed at the UVIC Genome BC Proteomics Center on a Thermo Scientific LTQ Velos Orbitrap. The synthesis of compounds **1** and **7** was performed as previously reported.[14] All final tri-acid compounds (**1**, **2**, **3** and **4**) were converted to their tri-sodium salts by reaction with stoichiometric NaOMe in MeOH before their dissolution in buffered water for use in NMR titration studies.

1,3,5-Tris(bromomethyl)-2,4,6-trimethoxybenzene (6). Synthesized by reported procedure for exact compound.[16] ^{1}H NMR ($CDCl_3$, 300 MHz): δ 4.14 (s, 9H), 4.60 (s, 6H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 22.7, 62.9, 123.5, 160.3.

Compound 8. 60% NaH (110 mg, 2.7 mmol) and indole ester **7** (550 mg, 2.7 mmol) were suspended in anhydrous DMF (5 mL) and stirred under N_2. After 40 minutes, 1,3,5-tris(bromomethyl)-2,4,6-trimethoxybenzene **6** (200 mg, 0.46 mmol) in DMF (2 mL) was added drop-wise and the reaction left to stir at room temperature for 17 hours. After complete reaction of compound **6**, most solvent was removed *via* vacuum and the reaction quenched with hexanes and THF. After salt removal *via* filtration, column chromatography (15 to 50% EtOAc/hexanes) yielded product **8** (127 mg, 34%) as a tan oil. IR (neat), ν (cm^{-1}): 3728, 3698, 3625, 3013, 2951, 2369, 2349, 2327, 1737, 1728, 1673, 1585 1467, 1193, 1107, 1093, 747. ^{1}H NMR ($CDCl_3$, 300 MHz): δ 2.73 (t, J = 7.7, 6H), 3.13 (t, J = 7.7, 6H), 3.45 (s, 9H), 3.68 (s, 9H), 5.28 (s, 6H), 6.97 (s, 3H), 7.16 (dt, J = 3.5, J = 0.6, 3H), 7.29 (dt, J = 4.0, J = 0.3, 3H), 7.61 (t, J = 8.9, 6H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 20.7, 35.1, 39.4, 51.5, 62.9, 109.8, 113.8, 118.9, 119.1, 121.4, 121.8, 125.1, 127.7, 136.6, 160.2, 173.7. HR-ESI-MS: 836.3519 ($[M + Na]^+$; $C_{48}H_{51}N_3O_9Na$; calc: 836.3523).

Compound 2. Tri-substituted methyl ester indole **8** (127 mg, 0.16 mmol) was dissolved in THF (4 mL) and distilled H_2O (2 mL). Excess $NaOH_{(s)}$ (91.5 mg) was added and the reaction was stirred at room temperature for 18 hours under N_2. Reaction was diluted with an equal volume of 1 M $HCl_{(aq)}$, extracted with EtOAc (3 × 30 mL), and dried over $MgSO_4$ before concentrating under vacuum. Precipitation from a minimum of DCM with hexanes and sonication gave triacid **2** (90 mg, 75%) as an

off-white solid. Mp: 117–120 °C (dec.) IR (neat), ν (cm^{-1}): 3024, 2957, 1708, 1581, 1463, 1249, 740. ^1H NMR ($CDCl_3$, 300 MHz): δ 2.71 (t, J = 6.5, 6H), 3.08 (t, J = 6.5, 6H), 3.20 (s, 9H), 5.24 (s, 6H), 6.90 (s, 3H), 7.11 (t, J = 7.2, 3H), 7.24 (t, J = 7.2, 3H), 7.48 (d, J = 8.2, 3H), 7.55 (d, J = 7.8, 3H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 20.4, 34.7, 39.7, 62.7, 109.8, 113.4, 118.8, 119.3, 120.9, 122.0, 124.8, 127.8, 136.8, 160.8, 179.3. HR-ESI-MS: 774.3051 ($[M + Na]^+$; $C_{45}H_{45}N_3O_9Na$; calc: 774.3054).

Methyl indole-3-carboxylate (9). Indole-3-carboxylic acid (2.00 g, 12.5 mmol) was dissolved in MeOH (23 mL) and concentrated H_2SO_4 (0.3 mL, 5.6 mmol) and heated to reflux for 3.5 hours. The reaction was cooled, poured into 75 mL ice, and extracted with DCM (3 × 60 mL). The combined organics were washed with saturated brine and saturated $NaHCO_{3(aq)}$, dried over $MgSO_4$, filtered and condensed. Column chromatography (20 to 25% ethyl acetate/hexanes) gave indole **9** (1.25 g, 57%) as a tan powder. Spectra matched known reference.[30] ^1H NMR (DMSO, 500 MHz): δ 3.80 (s, 3H), 7.19 (m, 2H), 7.47 (m, 1H), 7.99 (m, 1H), 8.07 (d, J = 3.0, 1H), 11.92 (br s, 1H). ^{13}C NMR (DMSO, 125 MHz): δ 50.6, 106.3, 112.3, 120.4, 121.2, 122.4, 125.6, 132.4, 136.4, 164.8.

Methyl indole-2-carboxylate (10). Indole-2-carboxylic acid (3.00 g, 18.7 mmol) was dissolved in MeOH (30 mL) and concentrated H_2SO_4 (0.3 mL, 5.6 mmol) and heated to reflux for 15 hours. The reaction was cooled, diluted with saturated $NaHCO_{3(aq)}$ (50 mL), and extracted with EtOAc (2 × 40 mL). The combined organics were dried over Na_2SO_4, filtered and condensed. The crude solid was suspended in hexanes, sonicated, and then filtered to give indole **10** (2.80 g, 86% yield) as a tan powder. Spectra matched known reference.[30] ^1H NMR ($CDCl_3$, 300 MHz): δ 3.96 (s, 3H), 7.16 (t, J = 7.5, 1H), 7.23 (s, 1H), 7.33 (t, J = 7.6, 1H), 7.43 (d, J = 8.3, 1H), 7.70 (d, J = 8.0, 1H), 8.98 (br s, 1H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 52.2, 109.0, 112.1, 121.1, 122.9, 125.7, 127.3, 127.7, 137.1, 162.7.

Compound 12. 60% NaH (245 mg, 6.1 mmol) and indole ester **9** (1.03 g, 5.9 mmol) were suspended in anhydrous DMF (8 mL) and stirred under N_2. After 25 minutes, 1,3,5-tris(bromomethyl)benzene **11** (359 mg, 1.0 mmol) in DMF (2 mL) was added drop-wise and the reaction left to stir at room temperature for 18 hours. Most solvent was removed *via* vacuum and the reaction quenched with hexanes and THF. After salt removal *via* filtration, column chromatography (15 to 50% EtOAc/hexanes) yielded **12** (320 mg, 50% yield) as a light yellow solid. Mp: 165–167 °C (dec.) IR (neat), ν (cm^{-1}): 2925, 1703, 1692, 1530, 1536, 1243, 1180, 1092, 757, 748. ^1H NMR ($CDCl_3$, 300 MHz): δ 3.93 (s, 9H), 5.18 (s, 6H), 6.74 (s, 3H), 7.05 (d, J = 8.2, 3H), 7.18 (td, J = 4.1, J = 1.1, 3H), 7.27 (td, J = 7.5, J = 0.9, 3H), 7.73 (s, 3H), 8.18 (d, J = 7.9, 3H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 50.2, 51.0, 108.0, 110.0, 121.9, 122.2, 123.2, 124.9, 126.7, 134.2, 136.4, 138.1, 165.2. HR-ESI-MS: 662.2264 ($[M + Na]^+$; $C_{39}H_{33}N_3O_6Na$; calc: 662.2267).

Compound 13. 60% NaH (245 mg, 6.2 mmol) and indole **10** (1.05 g, 6.0 mmol) were suspended in anhydrous DMF (8 mL) and stirred under N_2. After 30 minutes, 1,3,5-tris(bromomethyl)-benzene **11** (356 mg, 1.0 mmol) in DMF (2 mL) was added drop-wise and the reaction left to stir for 18 hours. Most solvent was removed *via* vacuum and the reaction quenched with THF. After salt removal *via* filtration, column chromatography (50 to 80% DCM/hexanes) yielded **13** (172 mg, 27% yield) as a beige solid. Mp: 174–176 °C (dec.) IR (neat), ν (cm^{-1}): 3050, 2948, 1712, 1709, 1249, 1197, 742. ^1H NMR ($CDCl_3$, 300 MHz): δ 3.65 (s, 9H), 5.52 (s, 6H), 6.44 (s, 3H), 6.96–7.14 (m, 9H), 7.17 (s, 3H), 7.57 (dd, J = 7.0, J = 1.1, 3H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 47.5, 51.5, 110.7, 111.2, 120.7, 122.6, 123.4, 125.2, 126.0, 127.1, 138.9, 139.3, 162.1. HR-ESI-MS: 662.2263 ($[M + Na]^+$; $C_{39}H_{33}N_3O_6Na$; calc: 662.2267).

Compound 3. Tri-substituted methyl ester indole **12** (122 mg, 0.19 mmol) was suspended in MeOH (12 mL). $KOH_{(s)}$ (221 mg) was dissolved in distilled H_2O (8 mL) and added drop-wise to the suspension. The reaction was heated to reflux for 16 hours until complete conversion observed by TLC (1 : 1 EtOAc/hexanes). Reaction was diluted with 1 M $HCl_{(aq)}$, extracted with EtOAc (3 × 30 mL), and dried over $MgSO_4$ before concentrating under vacuum. Solid was suspended in $CHCl_3$, filtered and air-dried to give triacid **3** (106 mg, 93% yield) as a tan solid. Mp: 261–263 °C (dec.) IR (neat), ν (cm^{-1}): 3055, 2938, 1665, 1659, 1536, 1531, 1278, 1253, 1190, 751. ^1H NMR (DMSO, 360 MHz): δ 5.38 (s, 6H), 7.08 (t, 3H, J = 7.6), 7.16 (t, 3H, J = 7.4), 7.23 (s, 3H), 7.32 (d, 3H, J = 8.1), 8.00 (d, 3H, J = 7.9), 8.16 (s, 3H), 12.01 (br s, 3H). ^{13}C NMR (DMSO, 90 MHz): δ 49.4, 106.8, 111.0, 120.8, 121.3, 122.2, 126.4, 126.6, 135.4, 136.1, 138.1, 165.6. HR-ESI-MS: 620.1797 ($[M + Na]^+$; $C_{36}H_{27}N_3O_6Na$; calc: 620.1797).

Compound 4. Tri-substituted methyl ester indole **13** (26 mg, 0.04 mmol) was suspended in MeOH (5 mL). $KOH_{(s)}$ (108 mg, 1.9 mmol) was dissolved in distilled H_2O (4 mL) and added drop-wise to the suspension. The reaction was heated to reflux for 17 hours until complete conversion observed by TLC (DCM). Reaction was diluted with 1 M $HCl_{(aq)}$, extracted with EtOAc (3 × 30 mL), and dried over $MgSO_4$ before concentrating under vacuum. Solid was suspended in $CHCl_3$, filtered and air-dried to give triacid **4** (23 mg, 93% yield) as a white solid. Mp: 262–264 °C (dec.) IR (neat), ν (cm^{-1}): 3035, 1721, 1687, 1678, 1673, 1519, 1267, 1200, 1171, 1136, 744. ^1H NMR (DMSO, 500 MHz): δ 5.65 (s, 6H), 6.64 (s, 3H), 7.10 (td, J = 3.9, J = 0.87, 3H), 7.15 (td, J = 3.5, J = 1.1, 3H), 7.19–7.23 (m, 6H), 7.65 (d, J = 7.9, 3H), 12.81 (br s, 3H). ^{13}C NMR (DMSO, 125 MHz): δ 46.8, 110.4, 111.0, 120.5, 122.3, 123.6, 124.7, 125.5, 127.7, 138.7, 139.0, 162.8. HR-ESI-MS: 620.1795 ($[M + Na]^+$; $C_{36}H_{27}N_3O_6Na$; calc: 620.1797).

Acknowledgements

This work was funded by the Michael Smith Foundation for Health Research (MSFHR) and NSERC. FH is a MSFHR Career Scholar and Canada Research Chair. ALW is a MSFHR Junior Graduate Research Trainee.

Notes and references

1 F. Diederich and K. Dick, *J. Am. Chem. Soc.*, 1984, **106**, 8024–8036.
2 C. S. Beshara, C. E. Jones, K. D. Daze, B. J. Lilgert and F. Hof, *ChemBioChem*, 2010, **11**, 63–66.

3 Y. Ma, X. Ji, F. Xiang, X. Chi, C. Han, J. He, Z. Abliz, W. Chen and F. Huang, *Chem. Commun.*, 2011, **47**, 12340–12342.
4 L. A. Ingerman, M. E. Cuellar and M. L. Waters, *Chem Commun*, 2010, **46**, 1839–1841.
5 D. L. Beene, G. S. Brandt, W. Zhong, N. M. Zacharias, H. A. Lester and D. A. Dougherty, *Biochemistry*, 2002, **41**, 10262–10269.
6 J. D. Schmitt, C. G. V. Sharples and W. S. Caldwell, *J. Med. Chem.*, 1999, **42**, 3066–3074.
7 X. Shi, I. Kachirskaia, K. L. Walter, J.-H. A. Kuo, A. Lake, F. Davrazou, S. M. Chan, D. G. E. Martin, I. M. Fingerman, S. D. Briggs, L. Howe, P. J. Utz, T. G. Kutateladze, A. A. Lugovskoy, M. T. Bedford and O. Gozani, *J. Biol. Chem.*, 2007, **282**, 2450–2455.
8 S. A. Jacobs and S. Khorasanizadeh, *Science*, 2002, **295**, 2080–2083.
9 J. F. Flanagan, L.-Z. Mi, M. Chruszcz, M. Cymborowski, K. L. Clines, Y. Kim, W. Minor, F. Rastinejad and S. Khorasanizadeh, *Nature*, 2005, **438**, 1181–1185.
10 Y. Huang, J. Fang, M. T. Bedford, Y. Zhang and R.-M. Xu, *Science*, 2006, **312**, 748–751.
11 K. D. Daze, T. Pinter, C. S. Beshara, A. Ibraheem, S. A. Minaker, M. C. F. Ma, R. Courtemanche, R. E. Campbell and F. Hof, *Chem. Sci.*, 2012, DOI: 10.1039/c2sc20583a.
12 S. A. Minaker, K. D. Daze, M. C. F. Ma and F. Hof, *J. Am. Chem. Soc.*, 2012, DOI: 10.1021/ja303465x.
13 K. D. Daze and F. Hof, *Acc. Chem. Res.*, 2012, DOI: 10.1021/ar300072g.
14 A. L. Whiting, N. M. Neufeld and F. Hof, *Tetrahedron Lett.*, 2009, **50**, 7035–7037.
15 P. C. Kearney, L. S. Mizoue, R. A. Kumpf, J. E. Forman, A. McCurdy and D. A. Dougherty, *J. Am. Chem. Soc.*, 1993, **115**, 9907–9919.
16 H. Li, E. A. Homan, A. J. Lampkins, I. Ghiviriga and R. K. Castellano, *Org. Lett.*, 2005, **7**, 443–446.
17 P. Martín-Gago, M. Gomez-Caminals, R. Ramón, X. Verdaguer, P. Martin-Malpartida, E. Aragón, J. Fernández-Carneado, B. Ponsati, P. López-Ruiz, M. A. Cortes, B. Colás, M. J. Macias and A. Riera, *Angew. Chem., Int. Ed.*, 2011, 1820–1825.
18 A. G. Cochran, N. J. Skelton and M. A. Starovasnik, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 5578–5583.
19 *Spartan '10*, Wavefunction, Inc, Irvine, CA, 2010.
20 P. Gans, A. Sabatini and A. Vacca, HypNMR – http://www.hyperquad.co.uk/hypnmr.htm 2006.
21 T. B. Gasa, C. Valente and J. F. Stoddart, *Chem. Soc. Rev.*, 2011, **40**, 57–78.
22 G. Hennrich, V. M. Lynch and E. V. Anslyn, *Chem.–Eur. J.*, 2002, **8**, 2274–2278.
23 D. Mendel, J. Ellman, Z. Chang, D. Veenstra, P. Kollman and P. Schultz, *Science*, 1992, **256**, 1798–1802.
24 S. M. Ngola, P. C. Kearney, S. Mecozzi, K. Russell and D. A. Dougherty, *J. Am. Chem. Soc.*, 1999, **121**, 1192–1201.
25 R. M. Hughes and M. L. Waters, *J. Am. Chem. Soc.*, 2005, **127**, 6518–6519.
26 H. J. Dyson, P. E. Wright and H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 13057–13061.
27 G. Némethy, I. Z. Steinberg and H. A. Scheraga, *Biopolymers*, 1963, **1**, 43–69.
28 A. Fernandez and H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 113–118.
29 C. D. Tatko and M. L. Waters, *Protein Sci.*, 2004, **13**, 2515–2522.
30 AIST, Spectral Database for Organic Compounds (SDBS) maintained by the National Institute of Advanced Industrial Science and Technology (Japan).

