ARTICLE TYPE www.rsc.org/xxxxxx | XXXXXXXX

Free-Energy Relationships for the Interactions of Tryptophan with Phosphocholines

Georg Blaser,[a] John M. Sanderson*[a] and Mark R. Wilson[a]

Received (in XXX, XXX) Xth XXXXXXXXX 200X, Accepted Xth XXXXXXXXX 200X
First published on the web Xth XXXXXXXXX 200X
DOI: 10.1039/b000000x

In membrane proteins and peptides, tryptophan exhibits a marked tendency to occur in locations that correspond to the interfacial region of the lipid bilayer. The relative contributions of electrostatic, dipolar, hydrophobic and conformational effects on the interactions of tryptophan with lipids have been the subject of much speculation. In order to elucidate the fundamental properties of tryptophan-phosphocholine interactions in the absence of competing factors such as protein conformation and membrane perturbation, we have determined the binding characteristics of a homologous series of tryptophan analogues to 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) in deuterochloroform using NMR titrimetric approaches. The data are analysed using a binding model that includes lipid aggregation and the explicit association of water with the lipid. For a series of substituents (OMe, Me, H, F, Cl, Br, I, NO_2) at the 5-position of the indole ring, the trends in the free energy of association for the formation of 1:1 and 1:2 lipid:tryptophan adducts both follow an inverted-U relationship as a function of the corresponding para-Hammett parameter, with tryptophan (R = H) exhibiting the weakest binding. These trends are shown to be consistent with participation of the indole side chain in both hydrogen bonds and cation-π interactions. Molecular dynamics simulations of tryptophan and DMPC in an explicit chloroform solvent model demonstrate that for the formation of lipid-tryptophan adducts, binding is driven predominantly by carbonyl-cation and cation-π interactions with the choline ammonium group, alongside hydrogen bonding interactions with the lipid phosphate. Some of these interactions operate co-operatively, which may account for the observed trends in free energy.

Introduction

In membrane proteins and peptides, tryptophan, along with phenylalanine and tyrosine, is found statistically more often at the lipid-water interface than other amino acids[1-4] and forms favorable interactions with lipids that are essential for biological function.[5-9] There is still some debate concerning the reasons for this preferred localisation, for which a number of factors may be influential. Important enthalpic factors are likely to include electrostatic interactions involving the indole side-chain, such as cation-π interactions and hydrogen bonds, alongside contributions arising from peptide conformational changes and dipolar interactions with membrane lipids.[10,11] Entropic factors include perturbation of lipid and bilayer structure, as well as the hydrophobic effect.[12,13] NMR studies on indole derivatives embedded in lipid bilayers have demonstrated that the dipole moment of the indole group, cation-π interactions and the ability of the indole group to participate in hydrogen bonding interactions with lipids all contribute to localisation in the interfacial region.[11,14-16] More recently however, steric effects and dipole-dipole contributions have been suggested to be more important than hydrogen bonding or cation-π interactions for determining the membrane partitioning and alignment of these and related indole derivatives.[10] Molecular dynamics simulations of mixed indole-lipid and tryptophan-lipid systems indicate a preference for indole to localise at specific locations in relation to membranes, including the middle of the bilayer, proximal to the choline headgroup and in the region of the lipid carbonyl groups, with the latter being the preferred site.[17,18] This is supported by experimental work on model helical transmembrane peptides bearing interfacial tryptophan residues, for which the presence of hydrogen bonding interactions with lipid carbonyl groups is notable.[19,20] However, an interfacial preference of tryptophan residues is also observed for similar peptides with ether-linked lipids,[21] which complements observations on the membrane partitioning behaviour of simple indole derivatives that similarly localise into the headgroup region of these lipids.[14,15] This implies that hydrogen bonding interactions with acyl carbonyls are not solely responsible for interfacial localisation. The ability of cation-π interactions to influence the membrane activity of synthetic molecules with pendant aromatic groups, such as ion channels, has been demonstrated, suggesting that these interactions may be a general feature of interactions with lipids.[22] As a further point of note, the surface sites available for interaction of a peripherally bound peptide or protein are likely to be different from those of an integral species, as, for example, lipid carbonyl groups will not be accessible, with significant consequences for our understanding of peripheral membrane activity. It is of interest therefore, to probe the interactions of tryptophan with phospholipids in the absence of competing restraints in order to understand the fundamental properties of the interactions

between them and deduce the electrostatic interactions that contribute to binding.

Recently, association constants for the interaction of a selection of *N*-acetyl amino acid ethyl amides (analogous to non-terminal protein residues) with 1,2-diacetyl-*sn*-glycero-3-phosphocholine (DAPC) in $CDCl_3$ were determined by NMR titration.[23] Although the dielectric constant of chloroform is lower than that expected of the membrane interfacial region, the lack of competitive interactions with this solvent enabled fundamental aspects of amino acid and lipid binding to be probed. These experiments demonstrated that tryptophan and tyrosine exhibit significantly more favorable free energies of association with phospholipids than other amino acids and permitted hypotheses to be generated concerning the nature of the intermolecular interactions, particularly for tryptophan.[24] Hydrogen bonding and cation-π interactions were proposed to contribute significantly to adduct formation. However, in order to make a proper assessment of the interactions that contribute to association of tryptophan with phospholipids, it is necessary to perturb the electron density on the indole ring and the hydrogen bonding potential of the indole-NH group and assess the changes in association free energy that result. The most convenient method for achieving this objective is to include a range of substituents on the indole ring at a position, which has the potential to modify both types of interactions. Experiments conducted along these lines to determine the free energies of association of DMPC (**2**) with tryptophan homologues bearing a range of substituents at the 5-position of the indole group (Fig. 1, **1a-h**), are reported in this paper, together with the development of a model for analyzing the data that accounts for the interactions of water with the lipid. In addition, molecular dynamics simulations have been carried out to determine the nature of the binding interactions in chloroform.

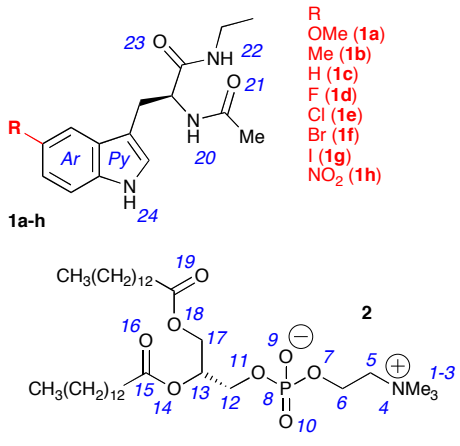
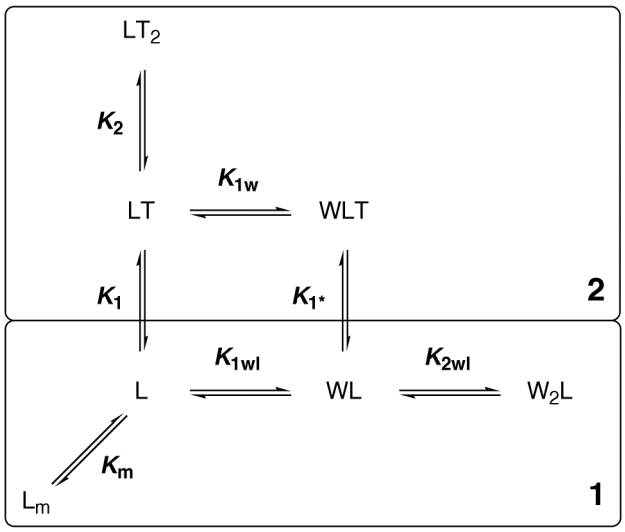


Fig. 1 Compounds used in this study: *N*-acetyl *L*-tryptophan ethyl amides (**1a-h**) and 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC, **2**). Atom numbers shown in blue correspond to the numbers used in the main text to describe contacts and torsions. Ar and Py correspond to the carbon and nitrogen atoms of the aryl and pyrrole parts of the indole ring respectively.

DMPC was selected as guest for these binding studies as it is significantly less hygroscopic than lipids such as DAPC. This leads to significant improvements in the reproducibility of determined association constants, which are sensitive to the water content of the sample. Moreover, the water signal in ^{1}H NMR spectra of DMPC is sharp and readily identified, in contrast to DAPC where it tends to be broad and frequently masked by other signals. This permits separate NMR titrations to be performed to account for interactions of DMPC with water and lipid self-aggregation, which has been reported to occur for phosphocholines in chloroform.[25,26] A model that included simple isodesmic aggregation of the lipid and 1:1 and 2:1 adducts with water was found to model the experimental data extremely well (Scheme 1) and was used for analysis of binding data for all adducts (see the Experimental section for a full account of this model).



Scheme 1 Equilibria used to model the interactions of tryptophan derivatives (**1a-h**) with DMPC (**2**) in $CDCl_3$. Box 1 describes DMPC self-association and water binding equilibria, where K_m is the association constant for the formation of lipid aggregates (L_m) and K_{1wl} and K_{2wl} are the association constants for the formation of 1:1 and 2:1 complexes of water (W) with the lipid (L). Box 2 describes equilibria associated with the formation of tryptophan (T) complexes, where K_1 and K_2 are the equilibrium constants for the formation of 1:1 and 2:1 complexes of tryptophan with lipid, K_{1*} is the equilibrium constant for the association of tryptophan with the 1:1 complex between water and lipid and K_{1w} is the equilibrium constant for the association of water with the 1:1 complex between tryptophan and lipid.

Materials and Methods

Materials and General Methods

DMPC was obtained from Bachem (St. Helens, UK) and used as purchased. Tryptophan derivatives **1a-h** were synthesised according to literature methods.[27] All ^{1}H NMR titrations were performed in duplicate in $CDCl_3$ on a Varian Unity spectrometer operating at 200 MHz, using tetramethylsilane as internal standard (0 ppm). Signals that exhibited ^{1}H chemical shift changes ≥0.05 ppm were used for association constant determination. The concentrations of the various species present were calculated by standard protocols as the root of a quadratic equation derived by combining mass balance equations with those of equilibrium constants.[28] Non-linear curve fitting procedures were performed using the Solver function in Excel (Microsoft Corp.).

Development of the Binding Model

Two factors needed to be considered in developing a binding model: the inclusion of sufficient equilibria to account for the formation of water:lipid complexes; and whether the aggregated form of the lipid was hydrated or not. Previous work on the aggregation of DPPC in chloroform has estimated the number of molecules per adduct to be 3,[25] which would not constitute adducts of a sufficient size to entrap water. This is supported by the sharpness of the water signal observed in DMPC ^{1}H NMR spectra in deuterochloroform. The data therefore support a model which excludes water from the aggregated form of the lipid[25,29] and uses an isodesmic model[30] for lipid micellisation. A consideration when selecting the number of water:lipid complexes to include was the desire to maintain the number of independent variables for regression analysis at the minimum required for accurate modelling of the data, in order to maximise the number of degrees of freedom and minimise the standard error of the fit. The number of water:lipid complexes was therefore restricted to two (Scheme 1, Box 1). Lipid:tryptophan adducts with 1:1 and 1:2 stoichiometries were included in the model, in accordance with the literature.[23]

The low solubility of tryptophan derivatives **1a-h** in chloroform prevented their use as guests in reverse NMR titrations with DMPC as host and the determination of binding constants for self-association. Nevertheless, due to the low concentrations used throughout titrations with **1a-h** as hosts, self-association and association with water could reasonably be excluded from the binding model. Due to the explicit inclusion of a 1:1 water:lipid complex, it was necessary to include the corresponding complex formed either by association of tryptophan with the 1:1 water:lipid complex or association of water with the 1:1 lipid:tryptophan complex, requiring two additional association constants (Scheme 1, Box 2). However, due to the nature of the thermodynamic cycle involving 1:1 adducts, one of these could be calculated with knowledge of the others. Considering all of the association constants and chemical shift parameters, this model required a total of seven independent variables per signal.

Water-DMPC Binding Experiments

A solution of DMPC in $CDCl_3$ (0.11 M) was added in aliquots to $CDCl_3$ (0.5 ml) at 20 °C and a ^{1}H NMR spectrum acquired after each addition, following mixing by agitation. The concentration of water in each sample ($[W]_t$) was determined by comparison of the water integral (where this could be distinguished) with that of a known lipid signal according to eqn. (1).

$$[W]_t = [L]_t n_I I_w / 2I_L \qquad (1)$$

where: $[L]_t$ is the total lipid concentration, I_L is the integral of the lipid signal, n_I is the number of lipid protons responsible for the signal and I_W is the integral of the water signal. The number of water molecules per lipid (n_w) and the concentration of free water in the $CDCl_3$ stock ($[W]_C$) were calculated by linear regression of eqn. (2) using the water concentrations ($[W]_t$) determined by eqn. (1), over a range of lipid ($[L]_t$) concentrations.

$$[W]_t = [L]_t n_w + [W]_C \qquad (2)$$

Chemical shift changes for water (δ_w) and lipid (δ_l) were fitted to the corresponding experimental data using eqns. (3) and (4) respectively:

$$\delta_w = \frac{[WL](\delta_{1wl} - \delta_{fw})}{[W]_t} + \frac{2[W_2L](\delta_{2wl} - \delta_{fw})}{[W]_t} + \delta_{fw} \qquad (3)$$

$$\delta_l = \frac{[WL](\delta_{1wl*} - \delta_{fl})}{[L]_t} + \frac{[W_2L](\delta_{2wl*} - \delta_{fl})}{[L]_t} + \frac{[L_m](\delta_{lm} - \delta_{fl})}{[L]_t} + \delta_{fl} \qquad (4)$$

where: $[WL]$ is the concentration of 1:1 water:lipid complex, $[W_2L]$ is the concentration of 2:1 water:lipid complex, $[L_m]$ is the concentration of lipid in micellar form, δ_{fw} is the chemical shift of free water, δ_{1wl} is the water chemical shift in the 1:1 water:lipid complex, δ_{2wl} is the water chemical shift in the 2:1 water:lipid complex, δ_{fl} is the chemical shift of free DMPC, δ_{1wl*} is the DMPC chemical shift in the 1:1 water:lipid complex, δ_{2wl*} is the DMPC chemical shift in the 2:1 water:lipid complex and δ_{lm} is the DMPC chemical shift in micellar aggregates. The concentration of lipid in the micellar form ($[L_m]$) was determined from the mass balance. In each titration, equations for all signals were solved globally to shared association constant values.

Tryptophan-DMPC Binding Experiments

A solution of each of the *N*-acetyl tryptophan ethylamides (**1a-h**) in $CDCl_3$ was prepared at a typical concentration of 5 mM. Of this, 0.5 mL was used as the initial host solution and 1 mL used to prepare a guest solution of DMPC (2) at ~100 mM. The guest solution was added to the host solution at 20 °C in aliquots and a ^{1}H NMR spectrum acquired after each addition, following mixing by agitation. Chemical shift changes for the tryptophan analogue (δ) were fitted to the experimental data using eqn. (5):

$$\delta = \frac{K_1[L](\delta_l - \delta_f) + K_{1*}[WL](\delta_{l*} - \delta_f) + 2K_2[LT](\delta_2 - \delta_f)}{1 + K_1[L] + K_{1*}[WL] + 2K_2[LT]} + \delta_f \qquad (5)$$

where: $[L]$ is the concentration of free lipid, $[LT]$ is the concentration of 1:1 lipid:tryptophan complex, δ_f is the chemical shift of free tryptophan, δ_1 is the tryptophan chemical shift in the 1:1 lipid:tryptophan complex, δ_2 is the tryptophan chemical shift in the 1:2 lipid:tryptophan complex and δ_{1*} is the tryptophan chemical shift in the 1:1:1 water:lipid:tryptophan complex. Values for K_{1wl} and K_{2wl} were fixed at those determined in separate experiments (see above). For each titration, data for all chemical shifts were initially subjected to global non-linear least squares regression to shared association constant values. Subsequently, each signal was subjected to regression analysis individually using the output from the global regression to obtain starting values for the independent variables. Association constants were evaluated as the weighted mean based on the observed chemical shift changes for the signals used. Chemical shift changes for water (δ_w) were calculated according to eqn. (6):

$$\delta_w = \frac{[WLT](\delta_{1w} - \delta_{fw})}{[W]_t} + \frac{[WL](\delta_{1wl} - \delta_{fw})}{[W]_t} + \frac{2[W_2L](\delta_{2wl} - \delta_{fw})}{[W]_t} + \delta_{fw} \qquad (6)$$

where δ_{1w} is the chemical shift in the 1:1:1 water:lipid:tryptohan complex. Values for δ_{fw}, δ_{1wl} and δ_{2wl} were fixed at those determined in separate experiments (see

above). Regression analysis was not performed using water chemical shift data.

Molecular simulations

Molecular mechanics calculations were performed using GROMACS (version 3.3.1).[31,32] Chloroform molecules were represented with a 4-site united atom model available from the literature, which provides excellent agreement with the experimental density.[33] DMPC was represented using the model of Tieleman *et al*,[34-36] with partial charges, binding and non-binding parameters of the lipid adapted from previous examples.[37] Water was represented using the SPC model. The OPLS-AA forcefield was used in the simulations.[38] Simulations at 300 K were performed in the constant-NpT ensemble with a time step of 0.002 fs. The temperature was controlled using the weak coupling algorithm of a Berendsen thermostat[39] with a coupling constant of 0.1 ps, while the pressure was set to 1 bar using a Berendsen barostat with a time constant of 5 ps. DMPC, tryptophan. water and the solvent were treated and coupled independently. Smooth particle mesh Ewald (PME) summation[40,41] was used to evaluate the electrostatic interactions with a real space cutoff of 1.2 nm, fourth-order spline interpolation, and a grid spacing of 0.12 nm. Van der Waals interactions were calculated using a Lennard-Jones potential with a truncation at 1.2 nm. Bond lengths of all molecules were constrained using the SHAKE algorithm[42] with a relative tolerance of 0.0001. Multiple starting configurations were used, with the amino acid derivative placed at 6 octahedral positions on a sphere with a 2 nm radius from the center of a DMPC molecule. All starting configurations were equilibrated for 10 ns before data were collected. Production runs were then performed over 200 ns. Simulations at elevated temperatures, 600 K, were conducted over 600 ns in the constant-*NVT* ensemble at a density of 1.46 ± 0.01 g/cm^3, with coordinates saved to the trajectory each 150 ps. Other parameters were the same as those used at 300 K.

Data analysis

Data from trajectories were subjected to analysis to determine the proportion of the simulation time during which selected atoms or groups of atoms were within the cut-off for non-covalent interactions. The cut-off for hydrogen bonding interactions was taken as 0.35 nm. For other interactions involving groups of atoms, the group centre of mass was used for calculating distances. Cut-off distances in these cases were calculated from the radial distribution function, $g(r)$, for the atoms involved. The $g(r)$ data were smoothed as an interpolated spline and the distance at which the 1st derivative of the smoothed spline was zero at a distance longer than the first peak in the $g(r)$ was taken as the cut-off distance. Smoothing and derivative calculation were performed using the R statistical graphics package (version 2.7.1).[43]

Cluster analysis was conducted on a subset of the coordinates comprising the amino acid and lipid headgroup only, using the GROMOS clustering method,[32,44] with rms cut-offs in the range 0.10-0.14 nm, chosen such that the average rms deviation per atom was <0.01 nm.

Results and Discussion

DMPC-water binding

Titration of hydrated DMPC into deuterochloroform produced significant chemical shift changes for the water signal and a number of the lipid signals (Fig. 2). From analysis of the water and lipid integrals in ^{1}H NMR spectra, the water concentration in the deuterochloroform was determined to be 4.6 mM, with 2.4 water molecules per lipid in the stock DMPC sample. Regression analysis using the model described in Scheme 1 (Box 1) produced good fits to the experimental chemical shift data and enabled the binding constants for the association of water with the lipid, K_{1wl} and K_{2wl}, to be determined as 30 ± 3 M^{-1} and 120 ± 12 M^{-1} respectively, with the association constant for isodesmic self-association of lipid molecules (K_m) evaluated as 134 ± 13 M^{-1}. These values were used as constants for calculations of tryptophan:lipid binding isotherms. The good fit to the experimental data for both lipid and water signals was indicative that this model described the system well.

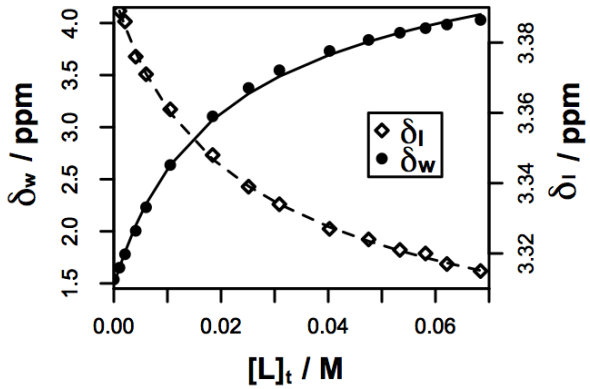


Fig. 2 Binding isotherms for the association of water with DMPC (**2**) in $CDCl_3$ at 20 °C. Experimental data for water (δ_w) and the DMPC ammonium group (δ_l) are shown as points (●) and open diamonds (◇) respectively. Values calculated by non-linear least squares fitting of these data using eqns. (3) and (4) (Experimental section) are shown as solid lines (——) for water and dashed lines (- - - -) for DMPC.

Fig. 3 shows the concentrations of each complex during the course of the titration using the experimentally-determined regression parameters and indicates that the major species present was the 2:1 water:lipid complex at all total lipid concentrations ($[L]_t$) above 10 mM. The concentration of free lipid ($[L]$) attained a steady state level of ~3.6 mM in this region, broadly consistent with concentrations described elsewhere for 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC)[25] in $CHCl_3$ and egg phosphatidylcholine in CCl_4.[29]

DMPC-tryptophan binding

In order to implement the binding model described in Scheme 1, two scenarios for analyzing the chemical shift data could be envisaged according to the means by which the 1:1:1 water:lipid:tryptophan complex was formed: 1) for non-competitive (independent) association of water and tryptophan with the lipid, $K_1 = K_{1*}$ and $K_{1w} = K_{1wl}$ enabling the number of degrees of freedom to be increased by one; 2) for co-operative

binding, $K_1 \neq K_{1*}$ and $K_{1w} \neq K_{1wl}$ enabling all of these variables to be adjusted independently. In practice, both approaches were examined for all experiments, although the quality of the fits to the experimental data were not significantly different between the two. The data presented in this paper are those that represent the former, non-competitive binding approach. A third option of competitive binding between water and tryptophan, *i.e.* with $K_{1*} = K_{1w} = 0$, can be ruled out as this model gives very poor fits to the experimental data.

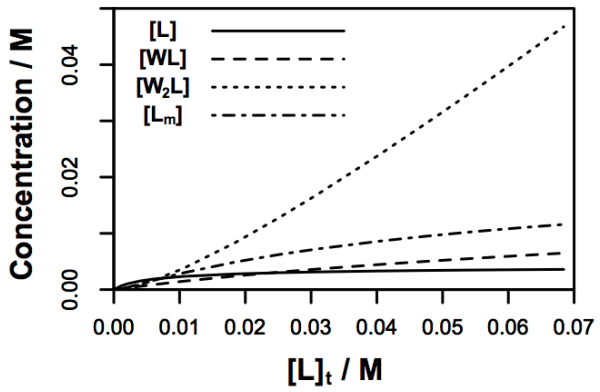


Fig. 3 Speciation of the lipid components during a DMPC (**2**)/water titration in $CDCl_3$ at 20 °C according to the model in Scheme 1 (Box 1). Data were calculated by least squares fitting of the chemical shift data using eqns. (3) and (4) (Experimental section and supporting information).

Titrations of DMPC into solutions of tryptophan derivatives **1a-h** at fixed concentrations in $CDCl_3$ produced marked downfield shifts of key heteroatom-associated protons in the ^{1}H NMR spectrum of the amino acids. In particular, even the weakest binding complex, **1c**, produced an observed chemical shift change of ~2 ppm for the indole-NH during the course of the titration and for others, such as **1d**, the observed shift was >2 ppm, enabling good fits to the experimental data to be obtained by nonlinear regression (Fig. 4A). Although the water signal could be followed throughout the titrations, these chemical shift data were not useful for regression procedures. Nevertheless, calculated water isotherms could be obtained (Fig. 4B) if the complexation-induced chemical shift of water in the 1:1:1 complex (δ_{1w}) was assumed to be the same as that in the 1:1 water:lipid complex (δ_{1wl}). These isotherms generally matched the experimental data very well, lending credence to the models used to analyze the binding data and providing further evidence for the non-competitive binding of water and tryptophan in the 1:1:1 complex.

The magnitudes of the complexation-induced chemical shift changes for the indole-NH groups of **1a-h** in both of the complexes with a 1:1 lipid:trytophan stoichiometry ($\Delta\delta_1$ and $\Delta\delta_{1*}$) were consistent across all of the titrations (Table 1), indicating that the structure of the these adducts did not change significantly in response to changes in electron density on the indole ring. A linear trend in complexation-induced chemical shift changes was observed for the 1:2 lipid:tryptophan complex, from $\Delta\delta_2$ = -0.01 ppm for **1a** to +3.5 ppm for **1h**, reflecting either the existence of secondary interactions between the two tryptophan molecules or a change in the structure of the 1:2 adduct. Significant downfield chemical shift changes of between +1.5 ppm and +4.5 ppm were observed for both of the amide-NH groups in the 1:1 lipid:tryptophan complexes, with those of the ethyl amide-NH groups being ~1 ppm greater than those of the acetyl-NH groups, in contrast to the 1:2 complexes, where $\Delta\delta_2$ values for the ethyl amide-NH protons were significantly smaller than for the 1:1 complexes (in the range +0.3 to +0.7 ppm) and the corresponding values for the acetyl-NH protons were negative (in the range –0.2 ppm to –1.1 ppm). As a whole, the complexation-induced shift changes, particularly for $\Delta\delta_1$, are in agreement with those observed previously for the interaction of **1c** with DAPC.[23]

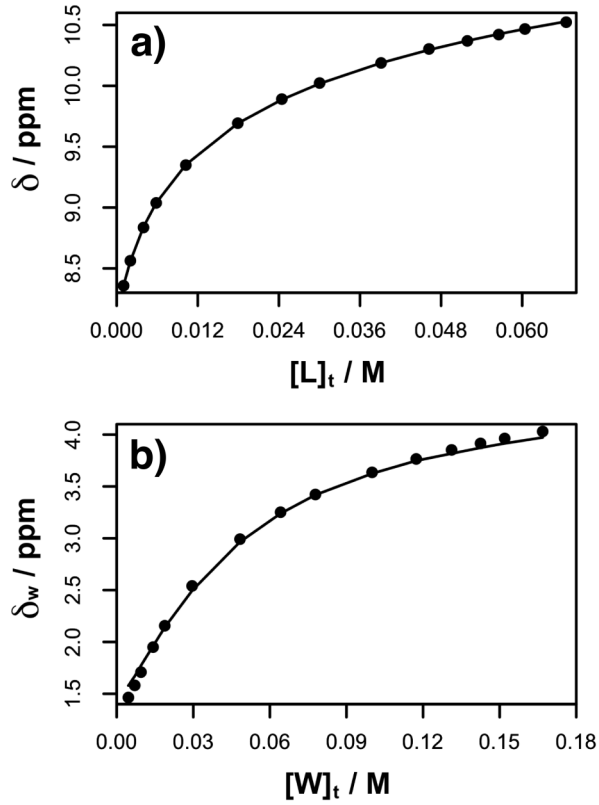


Fig. 4 Binding isotherms for the association of *N*-acetyl 5-fluoro-*L*-tryptophan ethyl amide (**1d**) with DMPC (**2**) in $CDCl_3$ at 20 °C. Experimental data are shown as points and calculated data as solid lines. (a) Chemical shift changes (δ) for the indole-NH of **1d**. The calculated isotherm was obtained by non-linear least squares fitting of experimental data to eqn. (5) (Experimental section) using K_{1wl} = 30 M^{-1}, K_{2wl} = 120 M^{-1} and K_m = 134 M^{-1}. (b) Chemical shift changes (δ_w) of water in the sample. The calculated isotherm was obtained using eqn. (6) without least squares regression, with δ_{1wl} = 3.01 and δ_{2wl} = 5.85, assuming $\delta_{1w} = \delta_{1wl}$.

Free energy trends for tryptophan-lipid interactions

In order to interpret the binding data, the free energy of association of each of the complexes was analyzed in relation to the corresponding *para*-Hammett substituent parameter (σ_p). Similar approaches have been used for the analysis of aromatic interactions[46,47] and were deemed pertinent in this case due to the 1,4-relationship between the 5-substituent and the indole-NH group. For the formation of 1:1 (*LT*) and 1:2

(LT_2) lipid:tryptophan complexes, the free energy of association becomes more favorable in comparison to tryptophan (**1c**) as the 5-substituent becomes more electron withdrawing (Fig. 5). An inverse linear relationship between σ_p and ΔG_1 (the free energy of formation of the 1:1 complex from the free species in solution) exists in the range $\sigma_p = 0$ to $\sigma_p = 0.28$, *i.e.* between tryptophan (**1c**) and 5-iodo-tryptophan (**1g**). Although ΔG_1 is more favorable for the 5-nitro derivative (**1h**) than the 5-iodo derivative (**1g**), the deviation from the linear trend suggests that ΔG_1 is approaching its limiting value. This is not the case for ΔG_2 (the free energy of formation of the 1:2 complex from the 1:1 complex), where the inverse relationship extends to $\sigma_p = 0.78$ for **1h**. Similarly, more favorable free energies of association relative to tryptophan (**1c**) are produced by electron releasing substitutents, with linear relationships between σ_p and both ΔG_1 and ΔG_2 in the range $\sigma_p = 0$ to $\sigma_p = -0.27$.

Table 1 Association constants and limiting complexation-induced chemical shift changes for the non-competitive association of 5-substituted *L*-tryptophan derivatives (**1a-h**) with DMPC (**2**)[a]

R	σ_p [b]	K_1 / M^{-1}	K_2 / M^{-1}	$\Delta\delta_1$ [c]	$\Delta\delta_{1*}$ [c]	$\Delta\delta_2$ [c]
OMe (**1a**)	-0.27	95 ± 9	319 ± 32	4.65	5.95	-0.01
Me (**1b**)	-0.14	69 ± 7	138 ± 14	5.02	5.41	0.16
H (**1c**)	0	57 ± 6	47 ± 5	4.94	5.73	0.27
F (**1d**)	0.15	104 ± 10	62 ± 6	5.22	5.22	0.53
Cl (**1e**)	0.24	190 ± 19	85 ± 8	5.01	4.29	1.65
Br (**1f**)	0.26	196 ± 20	88 ± 9	4.70	4.58	0.68
I (**1g**)	0.28	194 ± 19	89 ± 9	4.39	4.43	0.63
NO_2 (**1h**)	0.78	342 ± 34	470 ± 47	4.65	4.65	3.50

[a] Experiments were conducted in $CDCl_3$ at 20 °C. Association constants and limiting complexation-induced chemical shift changes ($\Delta\delta$) were calculated using non-linear least squares fitting of eqn. (5) (Experimental section) to the experimental data assuming a non-competitive binding model (*i.e.* $K_{1w} = K_{1wl}$ and $K_1 = K_{1*}$ was enforced). Errors in K are ± 10% or the value calculated by curve fitting if the latter was greater than the former. [b] *p*-Hammett substituent constants were obtained from reference 45. [c] Complexation induced chemical shift changes (in ppm) are reported for the indole-NH of **1a-h**. $\Delta\delta_1 = \delta_1 - \delta_f$; $\Delta\delta_2 = \delta_2 - \delta_f$; $\Delta\delta_{1*} = \delta_{1*} - \delta_f$. Standard errors in $\Delta\delta$ are ± 10%.

The break in the trend either side of tryptophan (**1c**, R = H) is consistent with two or more factors influencing the free energy of association. As contributions from both hydrogen bonding and cation-π interactions of the indole group potentially contribute to binding, it is to be expected that substituents with a strong electron withdrawing effect that reduce the π-electron density of the indole ring system and produce a higher partial positive charge on the indole-NH, should favour hydrogen bonding; conversely, those that are electron releasing should favour cation-π interactions.[48,49] Either side of tryptophan (**1c**), the free energy trends match these predictions, suggesting that association that is dominated by cation-π interactions for $\sigma_p < 0$, and hydrogen bonding interactions for $\sigma_p > 0$. The position of tryptophan (**1c**) in the trend is such that neither cation-π nor hydrogen bonding effects will be predominant and both will contribute. In a number of supramolecular systems, free energies for the formation of specific electrostatic interactions have been found to scale linearly with σ_p.[46,47] Consequently, the inverted-U trends observed here would suggest that either the formation of hydrogen bonds and cation-π interactions is co-operative in this system, or that other interactions influenced by the 5-substituent of the indole contribute to the free energy of binding.

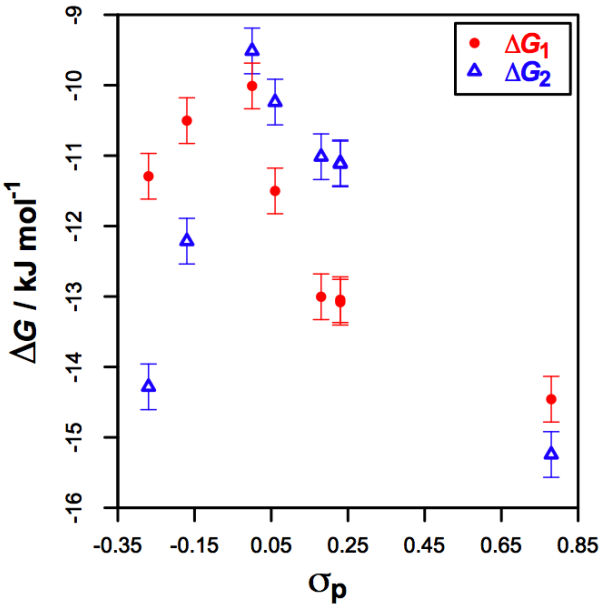


Fig. 5 Free energy trends for the association of 5-substituted *L*-tryptophan derivatives (**1a-h**) with DMPC (**2**) in $CDCl_3$ at 20 °C as a function of the *p*-Hammett substituent parameter (σ_p) of the 5-substituent. The points are calculated from the data presented in Table 1, which assume non-competitive binding of **1a-h** and water to **2**.

Ab initio calculations on ternary indole/alkali metal/water systems have concluded that polarisation and electrostatic factors make approximately equal contributions to the binding energy.[50] However, the cation-π and hydrogen bond interaction energies in these systems are non-additive; the formation of one leads to an interaction energy for the other that is more favorable by up to 5 kJ mol^{-1}, with a concomitant reduction in intermolecular distances. This free energy gain mostly comprises electrostatic rather than polarisation effects and is likely to be smaller for interactions with the ammonium group of choline, which has a considerably more disperse charge than alkali metal cations. Other studies have concluded that polarisation effects on aromatic systems with a common core structure remain relatively unchanged in response to substitution of the ring.[49,51] It is therefore reasonable to assess the trends in free energy of association in terms of changes to the electrostatic potential of the indole ring system, with synergistic behaviour between cation-π interactions and hydrogen bonding leading to co-operative behaviour as an explaination for the inverted-U trend.

We do not anticipate significant contributions to the observed trends as a consequence of secondary interactions in the adducts. Potential secondary effects may arise from intramolecular contacts that are disrupted upon adduct formation. The involvement of both amide groups of **1a-h** in binding interactions with **2** is implied by the observed chemical shift changes. However, the remote distance of these groups from the 5-position of the indole would suggest that substituent effects on these interactions should be minimal

and their contribution to binding free energies constant across all of the adducts, provided that there are no significant intramolecular contacts between the amide groups and the aromatic system. While dispersion and exchange repulsion are expected to make significant contributions to the overall stability of adducts, only the latter is likely to vary to any significant degree with changes in the 5-substituent on the indole. Given the r^{12} dependency of exchange repulsion, changes to this term would be reflected by changes to the geometry of contacts with the indole in this unconstrained system, with consequent effects on complexation-induced chemical shift changes. As the latter do not vary significantly as a function of σ_p for the formation of 1:1 adducts, significant changes in the exchange repulsion contribution can be ruled out. It is also reassuring to note that others have observed free energy trends for cation-π interactions that may be modelled in terms of changes to the electrostatic contribution to ΔG as a function of σ_p.[48,49]

Molecular Dynamics Simulations

The chloroform model used for this work gives a good reproduction of the bulk properties of the solvent for constant-*NpT* simulations at 300 K, with a calculated density of 1.46 ± 0.01 g/cm^3 (literature value of 1.48 g/cm^3 at 298 K[52]). *NpT* simulations of single molecules of either tryptophan derivative **1c** or DMPC (**2**) solvated in this chloroform model at 300 K produced acceptable torsional behaviour, with torsional distributions resembling those observed elsewhere[23] (see supporting information). In the case of tryptophan, the major conformer exhibited an intramolecular hydrogen bond between the NH of the ethyl amide group and the C=O of the *N*-acetyl group (Fig. 6A). The two other predominant conformers of tryptophan **1c** placed the *N*-acetyl-NH and ethyl amide-NH protons in close proximity to the pyrrole ring of the indole (Figs. 6B and 6C respectively). In the case of DMPC, the major conformer placed the choline ammonium group in close proximity to the carbonyl oxygen at the *sn*-2 position of glycerol as the result of the formation of an intramolecular carbonyl-cation interaction (Fig. 6D).

Six *NpT* simulations were conducted at 300 K with one molecule of tryptophan derivative **1c** and DMPC (**2**) per box, corresponding to a concentration for each of 42 mM. Unsurprisingly, given these concentrations and the experimental association constants for this interaction, adducts between DMPC and tryptophan formed rapidly (typically within 10 ps) and the molecules remained in a variety of binding configurations for the remainder of the simulation. While the extended lifetimes of the adducts at 300 K prevented a statistical analysis of association constants or an assessment of the most favorable binding configuration at this temperature, they did allow for identification of a series of key binding configurations arising from cluster analysis of the independent trajectories. The intramolecular π-facial (NH-Py) and hydrogen bonding interactions for tryptophan and cation-carbonyl interactions for DMPC, observed in the single molecule simulations (Fig. 6), were absent from these binding configurations.

Due to the limitations of the room temperature simulations for observing multiple association-dissociation events, the simulations were repeated under *NVT* conditions at 600 K at the same density. Both DMPC (**2**) and tryptophan derivative **1c** diplayed torsional distributions similar to those observed under *NpT* conditions when simulated individually in chloroform (see supporting information) and cluster analysis produced the same predominant conformations as the *NpT* simulations (Fig. 6). However, at this elevated temperature, multiple binding events could now be observed (Fig. 7).

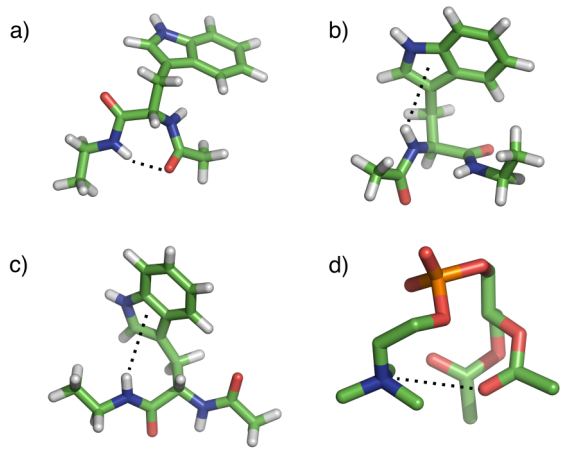


Fig. 6 The most abundant conformations of *N*-acetyl tryptophan *N'*-ethyl amide (**1c**) and DMPC (**2**) from constant-*NpT* simulations of single molecules of each in a chloroform box at 300 K. (a) tryptophan derivative **1c**. The median structure from a cluster of 153 structures is shown, representing 15% of the structures in the trajectory. ϕ, ψ, χ_1 and χ_2 are -84 ± 29°, 56 ± 37°, 43 ± 15° and 99 ± 17° respectively. The NH---O=C separation is 0.21 ± 0.07 nm. (b) **1c**. The median structure from a cluster of 107 structures is shown, representing 11% of the structures in the trajectory. ϕ, ψ, χ_1 and χ_2 are 13 ± 32°, -90 ± 33°, 38 ± 15° and -104 ± 15° respectively. The NH---Py centroid separation is 0.30 ± 0.08 nm. (c) **1c**. The median structure from a cluster of 89 structures is shown, representing 9% of the structures in the trajectory. ϕ, ψ, χ_1 and χ_2 are 158 ± 21°, -149 ± 22°, -160 ± 15° and -111 ± 17° respectively. The NH---Py centroid separation is 0.29 ± 0.06 nm. (d) DMPC (**2**). The median structure from a cluster of 308 structures is shown, representing 31% of the structures in the trajectory. The N---O=C(*sn*-2) separation is 0.39 ± 0.06 nm. Some atoms of the acyl chains have been omitted for clarity. Superimposed structures are presented in the supporting information.

Key groups of atoms potentially involved in hydrogen bonding, cation-π or carbonyl-cation interactions were analyzed in terms of the proportion of the trajectory that they spent within a specified cut-off distance (beyond which there could be said to be no interaction). In order to determine the appropriate cut-off distances for interactions involving groups of atoms, radial distribution functions were calculated for each (Fig. 8), with the cut-off obtained from the position of the first minimum in the $g(r)$ curve (see experimental section). While some caution must be exercised in interpreting the results from non-ambient temperature simulations, the radial distribution functions (Fig. 8), none-the-less, display some interesting features, particularly with respect to the presence of short and long range distributions. The short-range distributions were attributed to direct binding interactions between the groups concerned, whereas those at longer-range could be attributed to secondary effects from binding

interactions between other regions of the molecules. Analysis of the dynamics trajectory using these cut-off values enabled the relative abundance of each interaction to be assessed (Table 2).

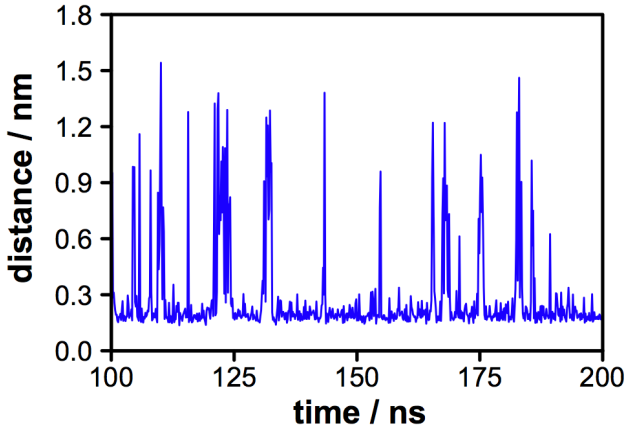


Fig. 7 A 100 ns sample of the 600 ns trajectory from an *NVT* simulation at 600 K of tryptophan derivative **1c** and DMPC (**2**) showing the minimum distance between the two molecules as a function of simulation time.

From these data it is evident that the most significant interactions observed are cation-π interactions involving the choline ammonium group and the indole ring, carbonyl-cation interactions involving the choline ammonium group and amide carbonyls, and hydrogen bonding interactions between the phosphate and the indole-NH or amide-NH groups. It is notable that there is a strong bias towards carbonyl-cation interactions involving the *N*-acetyl group (Table 2, entry **12**) and phosphate-amide interactions involving the ethyl amide group (Table 2, entry **6**). Analysis of these two interactions reveals that both are present in ~70% of cases where either one is less than the cut-off, suggesting a level of co-operativity between them. This is consistent with the observed chemical shift changes for the amide-NH groups, with the ethyl amide-NH exhibiting large positive changes and the *N*-acetyl-NH smaller positive or negative changes. The most significant hydrogen bonding interaction of the indole-NH group is with the lipid phosphate. For 80% of the time in which the latter interaction is less than the cut-off distance, cation-π interactions are observed with the phenyl ring of the indole, suggesting that these interactions are co-operative. Of the remaining 20% of the time that the indole-NH and phosphate groups are within the cut-off distance, approximately half is spent in the absence of other interactions. This reflects the steric environment of the tryptophan side chain, with the indole-NH being significantly less hindered than the other hydrogen bond donors of the tryptophan molecule. By contrast, in >99% of cases where cation-π interactions occur, with either the pyrrole or aryl rings of the indole, other interactions are present. This is presumably a reflection of the conformational properties of both molecules and the accessibility of functional groups, but nevertheless highlights the importance of this interaction for promoting binding.

Cluster analysis of structures from the trajectory produced binding configurations that were consistent with these observations (Fig. 9). As with configurations from *NpT* simulations at 300 K, the main clusters contained structures in which the predominant intramolecular contacts, observed in individual simulations of DMPC (**2**) and tryptophan **1c** (Fig. 6), were absent.

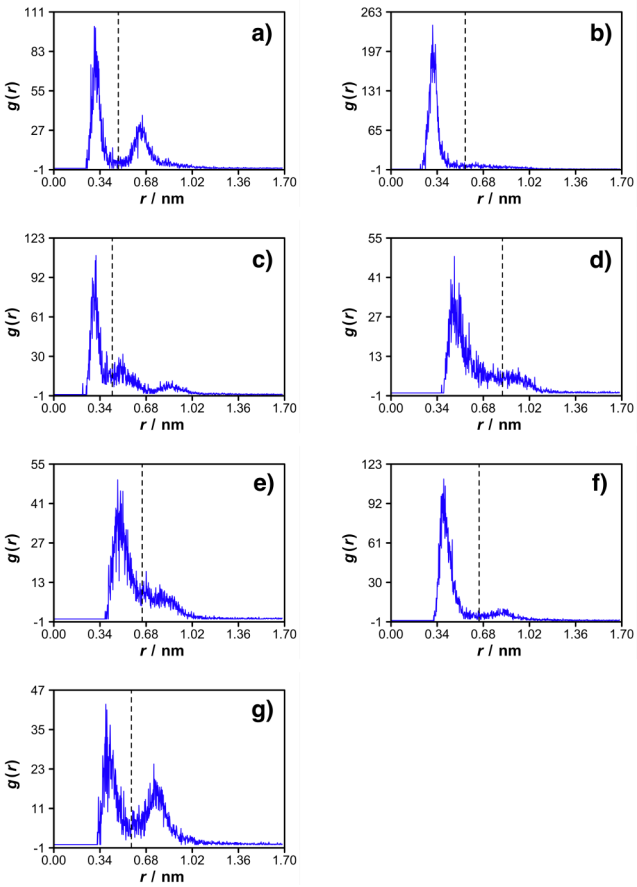


Fig. 8 Radial distribution functions for key groups of atoms during 600 ns *NVT* simulations of tryptophan **1c** and DMPC (**2**). Numbers correspond to the labels in Fig. 1. (a) 20-(7,9,10,11). (b) 22-(7,9,10,11). (c) 24-(7,9,10,11). (d) 4-(Ar). (e) 4-(Py). (f) 21-(1,2,3,5); (g) 23-(1,2,3,5). The centre of mass of the atoms in parentheses was used for $g(r)$ calculations. Vertical dashed lines represent the cut-off used for binding calculations and were determined from the first derivative of the smoothed curve.

The formation of the π-facial intramolecular contacts, in particular, would be expected to follow a similar free energy trend to cation-π interactions, becoming more favourable as the pyrrole ring of the indole becomes more electron rich. As the NH-Py distances observed (0.29-0.3 nm) are longer than the optimal distance of 0.19 nm calculated for pyrrole dimers,[53] their significance will be reduced somewhat. Nevertheless, their loss can potentially contribute unfavorably to the free energy of adduct formation. The experimental trend for $\sigma_p < 0$ would therefore indicate that these interactions are less significant, in terms of free energy, than contributions from competing intermolecular cation-π interactions. As π-facial hydrogen bonds tend to be weaker than hydrogen bonds,[54] free energy gains from hydrogen bond formation will also be more than sufficient to compensate for their loss. It nevertheless remains a possibility that for $\sigma_p > 0$, a reduction

in the population of states that involve these intramolecular contacts offsets the reduction in free energy contribution from cation-π interactions. The overall picture, therefore, is of a system of interacting molecules that contains a significant number of partially bound states, with enthalpic co-operativity[55] operating to populate specific tightly-bound configurations.

Table 2 Analysis of binding interactions during *NVT* simulations of tryptophan derivative **1c** and DMPC (**2**) in chloroform at 600 K.

Entry	Interacting Pair[a]	Cutoff / nm[b]	% time ≤ cut-off	% time as sole interaction[c]
1	20-19	0.35	1 ± 2	-
2	20-16	0.35	2 ± 2	-
3	20-(7,9,10,11)	0.475	20 ± 4	1
4	22-19	0.35	2 ± 1	-
5	22-16	0.35	5 ± 2	-
6	22-(7,9,10,11)	0.547	52 ± 11	2
7	24-19	0.35	6 ± 2	1
8	24-16	0.35	11 ± 2	6
9	24-(7,9,10,11)	0.431	22 ± 5	10
10	4-(Ar)	0.821	53 ± 5	1
11	4-(Py)	0.651	43 ± 5	0
12	21-(1,2,3,5)	0.649	52 ± 13	4
13	23-(1,2,3,5)	0.571	21 ± 6	3

[a] atom numbers correspond to those in Fig. 1. The centre of mass of the atoms grouped in parentheses was used for calculations. [b] cutoff set to 0.35 nm for hydrogen bonding interactions, or determined from point of zero gradient in the radial distribution function of the corresponding interaction (Fig. 8). [c] the proportion of the time spent at a distance less than the cut-off during which this was the only intermolecular interaction.

Separate simulations with two tryptophan molecules and one DMPC molecule per box allowed visualisation of 1:2 lipid:tryptophan adducts and analysis of the behaviour of each tryptophan independently. These were characterised by the same predominant contacts (see supporting information) as the 1:1 adduct. Significant changes were notable to the proportion of time during which the indole-NH was the sole point of contact with the lipid, which increased to 20%, and the nature of interactions with the indole, which now favoured the pyrrole ring over the aryl ring. Both of these changes reflect the more demanding steric environment around the lipid.

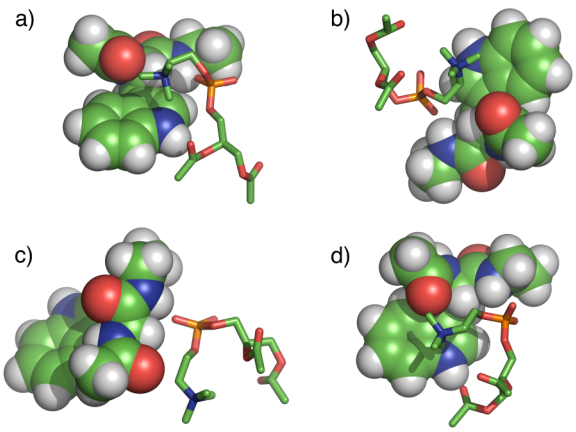


Fig. 9 The four most abundant binding interactions from cluster analysis of 4000 structures taken at 0.15 ns intervals during a 600 ns *NVT* simulaton of tryptophan derivative **1c** and DMPC (**2**) in a chloroform box at 600 K. The median structure is shown in each case. The numbers in each cluster are 48, 45, 41 and 31 for (a)-(d) respectively.

Direct interactions between the two tryptophan molecules in these adducts were insignificant, which further supports the model used to fit the experimental data. In order to investigate whether disruption to the intramolecular carbonyl-cation contacts of DMPC is also a feature of water binding, simulations of two water molecules and a DMPC molecule in a chloroform box were conducted under *NpT* conditions at 300 K. As with the case for the binding of tryptophan to DMPC, dissociation events were relatively rare, preventing a complete analysis of binding conformations. As with tryptophan/DMPC however, no conformations were observed in which the *sn*-2 acyl carbonyl was undergoing interaction with the choline ammonium group (Fig. 6D), suggesting that disruption of this interaction by co-ordination of the first water is the reason for the lower value of K_{1wl} in comparison to K_{2wl}.

Conclusions

Association constants for the binding of DMPC to a homologous series of tryptophan analogs have been determined experimentally by NMR titration in chloroform and augmented by computational studies in order to determine binding configurations. The model used to analyze the titration data describes the system very well, enabling accurate chemical shift predictions to be made for all tryptophan and water protons throughout each titration and a full account of the speciation in each sample to be made.

Trends in the experimental free energy of association suggest that cation-π interactions dominate for $\sigma_p < 0$, which is supported by the computational studies, where this is one of the predominant interactions. The large chemical shifts of the amide and indole amine protons observed during the course of NMR titrations is consistent with the involvement of these groups in hydrogen bonding interactions, and free energy trends indicate that these dominate for $\sigma_p > 0$. The significantly larger experimental chemical shift change for the indole-NH proton would suggest that this is involved in tighter interactions than the amide-NH protons, although this is not supported by the simulation data, which suggest a greater role for the ethyl amide-NH. However, a key feature of the indole-NH proton is its ability to form hydrogen bonds regardless of the conformational state of the tryptophan molecule, which will favour interactions in states where other groups are sterically masked or participating in other intramolecular interactions.

The observed chemical shift changes for the amide groups are consistent with the simulation data, where interactions between the ethyl amide-NH atom and the phosphate, and the carbonyl of the *N*-acetyl group and the choline ammonium group, are also predominant. These interactions would be expected to contribute to adduct formation irrespective of the nature of the remote indole substituent, unless there is competition from intramolecular interactions between the amide groups and the indole side chain. In this respect, interactions between the amide-NH groups and the pyrrole ring of the indole are observed in simulations of tryptophan (**1c**) in the absence of DMPC. However, analysis of the experimental free energy trends indicates that they are less significant than cation-π interactions, which compete for π-

facial association with indole.

Cation-π interactions contribute significantly to binding, especially as they appear to be able to form in co-operation with other electrostatic interactions, such as the indole-NH interaction with phosphate. The observed trend in the free energy of association may therefore reflect a predominant effect of cation-π interactions for electron-releasing R groups, which will favor electron density in the aromatic ring. Conversely, hydrogen bonding effects predominate for electron-withdrawing groups, which increase the delta-positive charge of the indole-NH, making it a better hydrogen bond donor. Unsubstituted tryptophan sits in a position where these two effects are balanced, rendering this residue maximally susceptible to perturbations of its interactions by the local electronic environment.

Notes and references

[a] *Centre for Bioactive Chemistry, Department of Chemistry, University Science Laboratories, South Road, Durham, DH1 3LE, UK. Fax: +44 (0)191 3844737; Tel: +44 (0)191 3342107; E-mail: j.m.sanderson@durham.ac.uk*

† Electronic Supplementary Information (ESI) available: full equations used for curve fitting, other models used for data analysis, chemical shift changes, σ_R and σ_I analysis, dynamics trajectories, torsional distributions, radial distribution functions, cluster analysis. See DOI: 10.1039/b000000x/

1 W. Liu and M. Caffrey, *Biochemistry*, 2006, **45**, 11713.
2 P. Raman, V. Cherezov and M. Caffrey, *Cell Mol. Life Sci.*, 2005, **63**, 36.
3 C. Landolt-Marticorena, K. A. Williams, C. M. Deber and R. A. Reithmeier, *J. Mol. Biol.*, 1993, **229**, 602.
4 M. B. Ulmschneider and M. S. Sansom, *Biochim. Biophys. Acta*, 2001, **1512**, 1.
5 D. I. Chan, E. J. Prenner and H. J. Vogel, *Biochim. Biophys. Acta*, 2007 **1758**, 1184.
6 W. M. Yau, W. C. Wimley, K. Gawrisch and S. H. White, *Biochemistry*, 1998, **37**, 14713.
7 E. K. Esbjörner, K. Ogleçka, P. Lincoln, A. Gräslund and B. Nordén, *Biochemistry*, 2007, **46**, 13490.
8 A. N. J. A. Ridder, S. Morein, J. G. Stam, A. Kuhn, B. de Kruijff and J. A. Killian, *Biochemistry*, 2000, **39**, 6521.
9 A. E. Daily, D. V. Greathouse, P. C. A. van der Wel and R. E. Koeppe II, *Biophys. J.*, 2008, **94**, 480.
10 E. K. Esbjörner, C. E. B. Caesar, B. Albinsson, P. Lincoln and B. Nordén, *Biochem. Biophys. Res. Commun.*, 2007, **361**, 645.
11 H. Sun, D. V. Greathouse, O. S. Andersen and R. E. Koeppe, *J. Biol. Chem.*, 2008, **283**, 22233.
12 T. J. McIntosh, A. Vidal and S. A. Simon, *Biochemical Society Transactions*, 2001, **29**, 594.
13 S. H. White and W. C. Wimley, *Annu. Rev. Biophys. Biomol. Struct.*, 1999, **28**, 319.
14 S. Persson, J. A. Killian and G. Lindblom, *Biophys. J.*, 1998, **75**, 1365.
15 H. C. Gaede, W.-M. Yau and K. Gawrisch, *J. Phys. Chem. B.*, 2005, **109**, 13014.
16 W.-M. Yau, W. C. Wimley, K. Gawrisch and S. H. White, *Biochemistry*, 1998, **37**, 14713.
17 K.E. Norman and H. Nymeyer, *Biophys. J.*, 2006, **91**, 2046.
18 J. L. MacCallum, W. F. Drew Bennett and D. P. Tieleman, *Biophys. J.*, 2008, **94**, 3393.
19 M. R. R. De Planque, B. B. Bonev, J. A. A. Demmers, D. V. Greathouse, R. E. Koeppe, F. Separovic, A. Watts and J. A. Killian, *Biochemistry*, 2003, **42**, 5341.
20 M. R. R. De Planque, J. A. W. Kruijtzer, R. M. J. Liskamp, D. Marsh, D. V. Greathouse, R. E. Koeppe II, B. De Kruijff and J. A. Killian, *J. Biol. Chem.*, 1999, **274**, 20839.
21 P. C. A. van der Wel, N. D. Reed, D. V. Greathouse and R. E. Koeppe II, *Biochemistry*, 2007, **46**, 7514.
22 M. E. Weber, E. K. Elliott and G. W. Gokel, *Org. Biomol. Chem.*, 2006, **4**, 83.
23 J. M. Sanderson and E. J. Whelan, *Phys. Chem. Chem. Phys.*, 2004, **6**, 1012.
24 J. M. Sanderson, *Org. Biomol. Chem.*, 2007, **5**, 3276.
25 R. Haque, I. J. Tinsley and D. Schmedding, *J. Biol. Chem.*, 1972, **247**, 157.
26 M. del C. Luzardo, F. Amalfa, A. M. Nuñez, S. Diaz, A. C. Biondi de Lopez and E. A. Disalvo, *Biophys. J.*, 2000, **78**, 2452.
27 G. Blaser, J. M. Sanderson, A. S. Batsanov and J. A. K. Howard, *Tetrahedron Lett.*, 2008, **49**, 2795.
28 K. A. Connors, *Binding Constants: The Measurement of Molecular Complex Stability*, John Wiley & Sons, New York, 1987.
29 T. M. Bray, J. A. Magnuson and J. R. Carlson, *J. Biol. Chem.*, 1974, **249**, 914.
30 T. H. Lilley, H. Linsdell and A. Maestre, *J. Chem. Soc. Faraday Trans.*, 1992, **88**, 2865.
31 H. J. C. Berendsen, D. van der Spoel and R. van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43.
32 E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Model.*, 2001, **7**, 306.
33 W. L. Jorgensen and D. L. Severance, *J. Am. Chem. Soc.*, 1990, 112, 4768.
34 C. Anézo, A. H. de Vries, H.-D. Höltje, D. P. Tieleman and S.-J. Marrink, *J. Phys. Chem. B*, 2003, **107**, 9424.
35 D. P. Tieleman, J. L. MacCallum, W. L. Ash, C. Kandt, Z. Xu and L. Monticelli, *J. Phys.: Condens. Matter.*, 2006, **18**, S1221.
36 J. L. MacCallum and D. P. Tieleman, *J. Comput. Chem.*, 2003, **24**, 1930.
37 O. Berger, O. Edholm and F. Jähnig, *Biophys. J.*, 1997, **72**, 2002.
38 W. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657.
39 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.
40 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.* 1993, **98**, 10089.
41 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577.
42 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463.
43 R Development Core Team R: A language and environment for statistical computing; R Foundation for Statistical Computing, Vienna, Austria, 2008. (ISBN 3-900051-07-0, URL http://www.R-project.org).
44 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chem., Int. Ed.*, 1999, **38**, 236.
45 N. Isaacs, *Physical Organic Chemistry*, 2nd Ed. Longman, 1995, pp. 152.
46 C. A. Hunter, K. R. Lawson, J. Perkins and C. J. Urch, *J. Chem. Soc. Perkin Trans 2*, 2001, 651.
47 S. L. Cockroft and C. A. Hunter, *Chem. Soc. Rev.*, 2007, **36**, 172.
48 B. W. Gung and J. C. Amicangelo, *J. Org. Chem.*, 2006, **71**, 9261.
49 J. C. Ma and D. A. Dougherty, *Chem. Rev.*, 1997, **97**, 1303.
50 D. Escudero, A. Frontera and D. Quinonero, P. M. Deya, *Chem. Phys. Lett.*, 2008, **456**, 257.
51 E. Cubero, F. J. Luque and M. Orozco, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 5976.
52 D. R. Lide (Ed.), *CRC handbbok of Chemistry and Physics*, 86th Ed., CRC press, 2005.
53 H. Park and S. Lee, Chem. *Phys. Lett.*, 1999, **301**, 487.
54 H. Adams, K. D. M. Harris, G. A. Hembury, C. A. Hunter, D. Livingstone and J. F. McCabe, *Chem. Commun.*, 1996, 2531.
55 C. A. Hunter and S. Tomas, *Chem. Biol.*, 2003, **10**, 1023.