Proline-rich proteins—deriving a basis for residue-based selectivity in polyphenolic binding†

A. K. Croft* and M. K. Foley

Received 9th January 2008, Accepted 22nd February 2008
First published as an Advance Article on the web 19th March 2008
DOI: 10.1039/b800365c

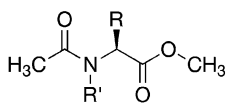
^{1}H NMR titration experiments have been used to establish that minimal proline-based models show enhanced binding selectivity towards phenol in $CDCl_3$, relative to other similarly protected amino acid residues. Cooperative binding effects appear to play a role, with sarcosine models affording binding constants to phenol intermediate to those obtained from proline models and other amino acid models. The mechanism for binding, based on DFT calculations and the application of Hunter's molecular recognition toolbox model, cannot be solely attributed to hydrogen bond strength, and appears to be mediated through C–H-π bonds and the rotational freedom of the amide substrate.

Introduction

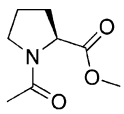
Phenolic compounds are extremely important components of plants, and are consumed in the diet through a number of sources, including in high levels in both tea and wine. These molecules have received much publicity due to their antioxidant nature, and the perceived health benefits that arise from this activity.[1] However, free in solution, phenolics may interact adversely with biomolecules because they are able to participate directly in oxidative processes, as well as possessing one or more nucleophilic sites in the reduced form and reactive electrophilic sites once oxidised.[2]

In nature, animals are able to compensate for the destructive properties of phenolics through the selective binding of these tannic components to proteins.[2,3] The phenolics are then able to impart their antioxidant ability, whilst being protected from reaction with the general cellular environment.[4–8] This is achieved in humans and other animals, when they are exposed to a tannin-rich diet, through the excretion of two classes of salivary proteins,[9] histidine-rich proteins (HRPs), which are present constitutively, and proline-rich proteins (PRPs) that are present in increased levels (up to 70% of the parotid salivary protein).[3,10–12]

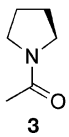
The binding of proteins, particularly PRPs, to polyphenolics has been extensively studied,[13–39] since protein–tannin complexes are key to the sensation of astringency.[40–46] Specificity for binding to the proline features of PRPs, which constitute approximately 35% of the protein content, was postulated by Hagerman and Butler from the results of competitive assays.[21] In this work, all polypeptides that showed good affinity for proanthocyanidin also contained high amounts of proline. It was noted that polyproline and polysarcosine showed similar affinities for proanthocyanidin. Small peptides were unable to compete satisfactorily in this assay, highlighting the requirement for several additive or synergistic interactions. Subsequent detailed ^{1}H NMR studies have since confirmed that the main interaction of the phenolic rings of tannins is primarily with the proline of PRPs.[35–39]



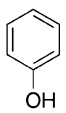
1a R = H, R' = H
1b R = H, R' = CH_3
1c R = CH_3, R' = H
1d R = CH_2Ph, R' = H
1e R = CH_2Ph-*p*-OPh, R' = H



2



3



4

The study of minimal models provides a framework by which to better understand residue-specific interactions and their contribution to supramolecular assemblies. This can help, with the aid of the studies already carried out on full protein–tannin models, to establish firstly what are the underlying residue-specific driving forces for interaction on an easily understood single-molecule scale, and secondly where the gaps in understanding are on going from intermolecular to supramolecular species. In this context, this work examines a selection of *N*-acetylated amino acid esters **1a–e** and **2**, as well as *N*-acetyl pyrrolidine **3**, and their interaction with phenol **4** by ^{1}H NMR titration methods. The thermodynamics of the interaction of *N*-acetyl proline methyl ester **2** and phenol **4** are examined in further detail. These results have been complemented by calculations of the structures of relevant amino derivatives **1–3**, with determination of their electrostatic potentials to elucidate key contributions to the binding mechanism, discussed in the context of Hunter's molecular recognition toolbox model.[47] This combination allows separate contributors to the binding to start to be understood.

Experimental methods

Synthesis of compounds

The L-amino acid derivatives **1a–e** and **2** were prepared from the corresponding L-amino acids as previously described in the literature,[48–50] as was the *N*-acetyl pyrrolidine **3**.[51] Extended experimental data is provided in the ESI.† Phenol **4** was purchased

School of Chemistry, University of Wales Bangor, Bangor, Gwynedd. E-mail: a.k.croft@bangor.ac.uk; Fax: +44 1248 370 528; Tel: +44 1248 382 375

† Electronic supplementary information (ESI) available: Synthetic procedures and characterisation, ^{1}H NMR details and titration curves. See DOI: 10.1039/b800365c

from Sigma-Aldrich, and used as received. The structures and purity of all compounds were verified by NMR prior to use.

^{1}H NMR titrations

^{1}H NMR titrations were performed on a 500 MHz Bruker Avance spectrometer with a 5 mm broad band observation (BBO) gradient probe with calibration from the position of residual non-deuterated solvent, cross referenced to trimethylsilane (TMS). Each measurement was taken on at least three freshly prepared samples for each concentration point. Solutions (10 ml) of the compounds **1–4** were prepared at known concentrations in $CDCl_3$. For the direct titration, aliquots of 1.0 mol dm^{-3} of phenol **4** were added to 45 μl each of 0.1 mol dm^{-3} of the amides **1–3**, with addition of extra $CDCl_3$ to make up the total volume to 500 μl. A ^{1}H NMR spectrum was recorded after each aliquot addition and the chemical shifts of the proton signals of the amides **1–3** were monitored with each addition. The association constants (K_a) and predicted change in chemical shift to produce saturation of association ($\Delta\delta_{pred}$) were derived by plotting non-linear titration curves using the computer graphing program, Mac-Curve-Fit™. The association constant, K_a, is obtained by application of eqn (a)–(c), which allow the observed chemical shift values of any given proton signal, δ_{obs}, to be fitted as a function of the initial concentration of phenol **4**. The association constant for each experiment was evaluated as the weighted average on the observed change in chemical shift for all signals monitored.

$$[HG] = \frac{1 + K_a([H]_i + [G]_i) - \left((1 + K_a([H]_i + [G]_i))^2 - 4K_a^2[H]_i[G]_i\right)^{0.5}}{2K_a} \quad \text{(a)}$$

$$[H] = [H]_i - [HG] \quad \text{(b)}$$

$$\delta_{obs} = \frac{[H].\delta_u + [HG].\delta_b}{[H]_i} \quad \text{(c)}$$

Where δ_u is the chemical shift of the unbound host and δ_b is the shift of the bound complex and $[H]_i$ and $[G]_i$ are the initial concentrations of host and guest, respectively.

The stoichiometry of the association between the proline derivative **2** and phenol **4** was ascertained by the continuous variation method, otherwise known as Job plot analysis.[52] ^{1}H NMR spectra were obtained using a range of different ratios of **2** and **4** and every proton signal of the proline derivative **2** was analysed for each rotamer. At the concentrations chosen to perform the Job plot analysis (0.1 mol dm^{-3} to 1.0 mol dm^{-3}), the proton signals of the proline derivative **2** were generally easier to resolve than when performing the ^{1}H NMR titration experiments, especially for the *trans*-rotamer.

Standard spectra were analysed at 20 °C, with probe temperature maintained by electronic heating and liquid nitrogen cooling. In a similar fashion, all temperature dependent plots were cooled or heated as required to the specific temperature with equilibration of samples prior to measurement.

Table 1 Association constants of selected amides **1–3** with phenol **4** in $CDCl_3$ based on ^{1}H NMR titration results

Compound	$K_{a\,1}$ (mol^{-1} dm^3)
N-Ac-Gly-OMe **1a**	2.4 ± 0.6
trans-*N*-Ac-Sar-OMe **1b**	6.0 ± 2.7
cis-*N*-Ac-Sar-OMe **1b**	6.9 ± 1.1
N-Ac-Ala-OMe **1c**	2.4 ± 0.8
N-Ac-Phe-OMe **1d**	2.2 ± 1.1
N-Ac-O-Bzl-Tyr-OMe **1e**	2.7 ± 0.7
trans-*N*-Ac-Pro-OMe **2**	18.5 ± 2.9
cis-*N*-Ac-Pro-OMe **2**	28.0 ± 8.1
N-Ac-Pyr **3**	14.0 ± 8.5

Electrostatic surface calculations

Density functional theory (DFT) calculations were performed using the Spartan '02 programme.[53,54] The minimum energies of both rotamer conformations of **2** were obtained by an initial conformer distribution calculation, at the B3LYP/6-31(d) level of theory. The generated conformations were then optimised to obtain equilibrium geometry structures. The equilibrium geometry structures of **1–4** were also calculated at B3LYP/6-31(d). All structures were confirmed as minima by the absence of negative frequencies. Electrostatic potential surfaces of molecules **1–4** were generated also using Spartan '02[53,54] at the B3LYP/6-31G(d) level.

Results and discussion

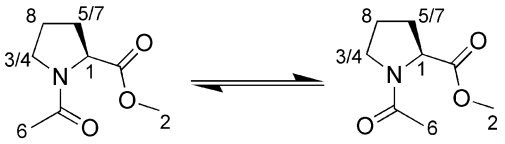
^{1}H NMR titrations

NMR titrations provide one method to identify binding interactions between species. For ^{1}H NMR, the change in chemical shift of protons of a molecule of interest can be monitored in response to increasing quantities of a complexing agent, and usually indicate a change in the chemical environment of species one initiated by species two. From these changes in chemical shift, binding constants, K_a, can be calculated. The K_a values obtained for the amide derivatives **1–3** with phenol **4** are presented in Table 1.

The K_a values for the glycyl, alanyl, phenylalanyl and tyrosyl derivatives **1a** and **1c–e**, respectively, indicate negligible binding within experimental error. A slightly more significant binding is observed for the sarcosine derivative **1b**, with a predicted K_a of around 6 mol^{-1} dm^3. For the proline and pyrrolidine derivatives **2** and **3**, the measured K_a increases to around 14 mol^{-1} dm^3 in the case of **3**, 18 mol^{-1} dm^3 for the *trans*-rotamer of **2** and 28 mol^{-1} dm^3 for the *cis*-rotamer. Whilst these are small values compared with typically reported binding constants, the trend and relative values are reproducible for these systems.

Two rotamer conformations of the tertiary amides **1b** and **2**, termed *trans*- and *cis*-rotamers (see *e.g.* Scheme 1), are observed directly on the NMR time scale. This phenomenon is seen with many tertiary amides.[55–59] The *trans*-rotamer is more prevalent than the *cis*-rotamer and the rotamer conformations of both sarcosine **1b** and proline **2** were estimated to be at approximately a 4 : 1 ratio in a $CDCl_3$ solution from analysis of the ^{1}H NMR integrals.

The two rotamer conformers of both sarcosine **1b** and proline **2** were analysed separately by ^{1}H NMR titration. The changes in chemical shift of the α-protons of proline **2** (proton 1, Scheme 1), with addition of phenol **4**, are illustrated in Fig. 1. The α-proton



Scheme 1 The two rotamer conformations of *N*-Ac-Pro-OMe **2**. The amide bond rotation equilibrium is observed directly on the NMR time scale and both conformers are observed on a ^{1}H NMR spectrum.

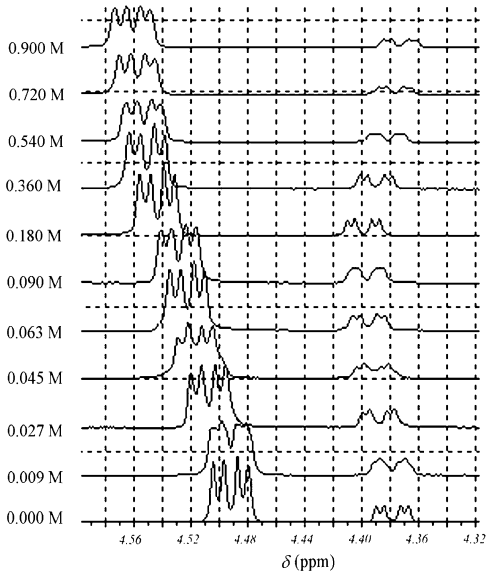


Fig. 1 A stack plot showing the α-proton signals of *N*-Ac-Pro-OMe **2** as a gradually increasing concentration of phenol **4** is added, shown in mol dm^{-3} (M) to the left. The solutions were made up in $CDCl_3$ solution, and the concentration of **2** is 0.009 mol dm^{-3}. When no phenol is present the signals correspond to (from left to right): δ^H 4.49, doublet of doublets (*trans*) α-*CH*; 4.38 ppm, doublet of doublets (*cis*) α-*CH*. Both α-proton signals appear to change in chemical shift with addition of phenol.

signal was monitored due to the relatively large change in chemical shift compared to the other signals, and also the ease of resolving both rotamer conformation signals.

The association of the *trans*-rotamer of **2** to phenol **4** resulted in one saturation point, but association to the *cis*-rotamer produced two saturation points (Fig. 2). This suggests that a secondary association occurs, unique to the *cis*-rotamer, with a quantifiable K_a value. However, the K_a values obtained for the secondary association are appreciably smaller than the initial association with values around 7 ± 1 calculated.

With association of the proline derivative **2** to phenol **4**, the largest changes in chemical shift observed for the signals of the *trans*-rotamer protons are those corresponding to the α-proton and one specific N*CH*$_2$ proton. Although it is not possible to assign the facial locality of the protons, the larger change in chemical shift of a specific N*CH*$_2$ proton does suggest specific facial association. The direction in chemical shift is also significant; the down-field shift observed for the α-proton signal is presumably due to deshielding by either an electron-withdrawing group, for example the phenolic oxygen, or due to the ring current of the aromatic ring. The up-field shift observed by the N*CH*$_2$ proton

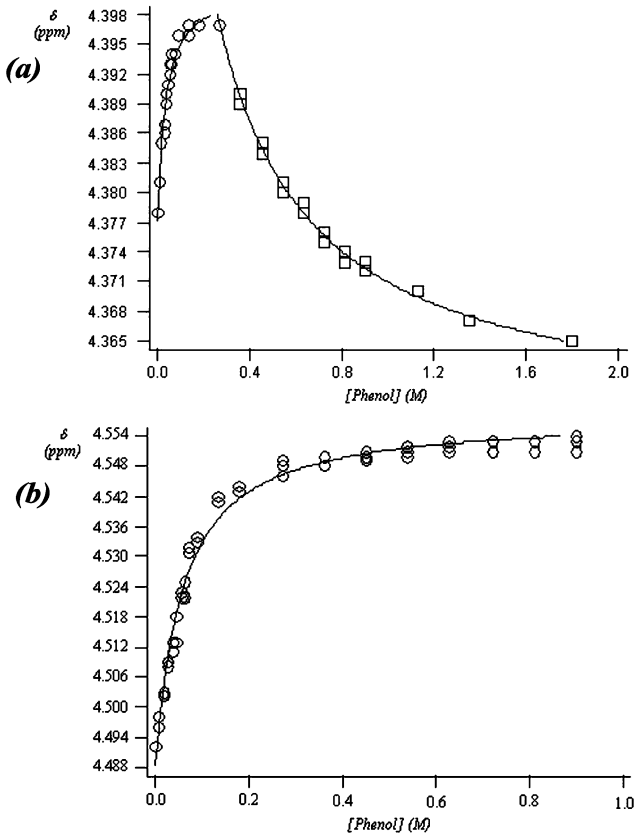


Fig. 2 *(a) trans*- and *(b) cis*-*N*-Ac-Pro-OMe **2** α-^{1}H NMR titration curves with addition of phenol **4**. The curves indicate that the *trans*-rotamer associates at a 1 : 1 ratio that results in a down-field change in chemical shift of the α-proton, but the *cis*-rotamer appears to produce a secondary association with **4**, and an up-field shift is observed after the initial down-field change in chemical shift.

signal is most probably due to shielding caused by the ring current of the phenolic aromatic ring. The observed chemical shift changes of the *cis*-rotamer proton signals are significantly different. The initial association results in the chemical shift of the protons moving down-field, apart from one of the ring protons, although this observation is confounded by overlapping proton signals. The secondary association causes the chemical shift of most of the protons to move up-field, with the most significant differences being observed for the ring protons. Although the low concentration of the *cis*-rotamer makes the observations less reliable, the results suggest that a very different mode of association occurs than is seen for the *trans*-rotamer of **2**.

The difference in association mechanism of the two rotamers is further highlighted when continuous variation (Job plot) analysis is carried out. This method allows the stoichiometry of binding to be determined through variation in the concentration of each analyte whilst maintaining a constant concentration of the combined analytes. The resulting Job plots are presented in Fig. 3.

Fig. 3 indicates that for both rotamers of the proline derivative **2**, the largest changes in shift are associated with the α-proton in each case. For the *trans*-rotamer the maxima for the α-proton, *O*CH$_3$ protons and *N*Ac protons are around 0.45 indicating that the stoichiometry is likely to be 1 : 1, although this maximum value could suggest that a second stoichiometry is possible.

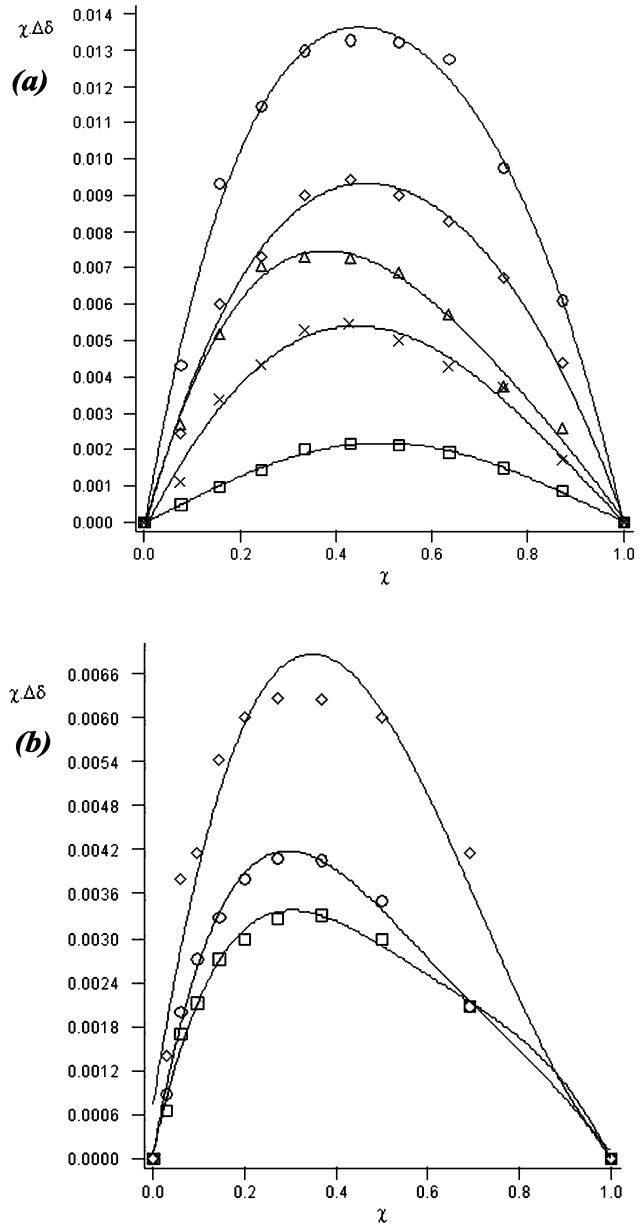


Fig. 3 Job plots monitoring the α-proton (○), the OCH_3 protons (□), the *N*-AcCH_3 protons (◊), and the NCH_2 protons (Δ and ×) of *trans*-*N*-Ac-Pro-OMe (a), and *cis*-*N*-Ac-Pro-OMe (b) at different ratios of phenol **4**.

However, when monitoring the NCH_2 proton signals, a maximum of around 0.40 was obtained. This can be attributed due to errors in measurement arising from signals corresponding to the *cis*-rotamer overlapping with the *trans*-rotamer proton signals. The maxima for the *cis*-rotamer is about 0.3 and indicates that a different stoichiometry is prevalent, probably 2 : 1. However, with such small changes in chemical shift for the *cis*-rotamer association, the results become difficult to interpret.[60]

The association of both rotamers **2** with phenol **4** was examined at different temperatures to extract thermodynamic information. When the temperature of the system is decreased, the initial association (K_{a1}) to both of the rotamers is strengthened (Table 2). A secondary association (K_{a2}) to the *trans*-rotamer was also seen,

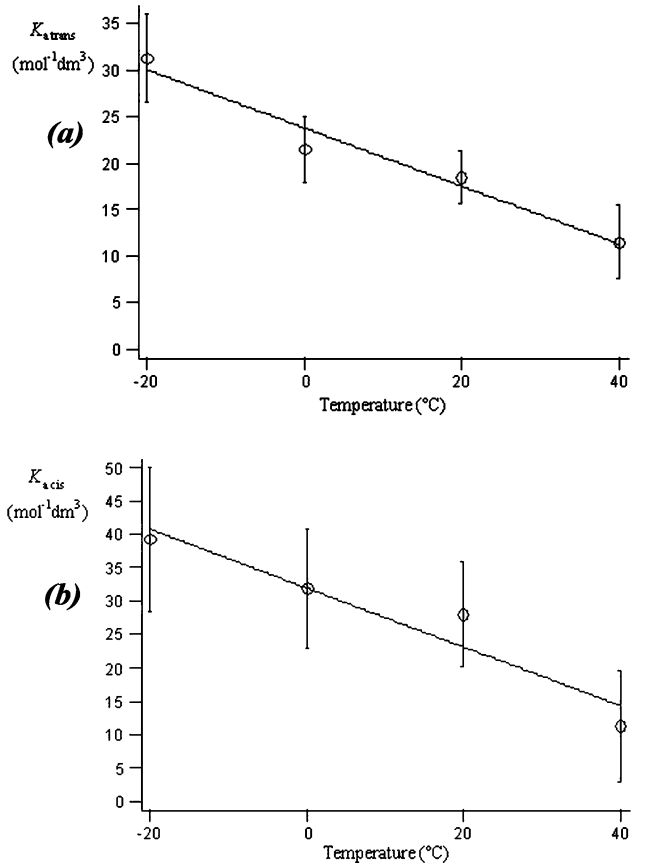


Fig. 4 (a) *trans*- and (b) *cis*-*N*-Ac-Pro-OMe **2** K_{a1} values for the association to phenol **4** *vs.* temperature. $R^2 = 0.96$ (*trans*) and 0.92 (*cis*).

in a similar manner as observed with the *cis*-rotamer at room temperature. As the temperature is increased, this secondary association to the *trans*-rotamer disappears. Fig. 4 illustrates how the K_{a1} values correlate with temperature. A linear relationship fits both sets of data relatively well, with R^2 values of 0.96 and 0.92 for the *trans*- and *cis*-rotamers respectively. The lower correlation of the *cis*-rotamer association can be accounted for by larger errors associated with the measurements than for the *trans*-rotamer. Both correlations indicate that lower temperatures enhance the initial association, as would be expected for a binding process and where competitive interactions may occur.

The value of the change in enthalpy when association occurs (ΔH) was calculated as -10.3 ± 3.9 kJ mol^{-1} for the *trans*-rotamer using the van't Hoff equation,[61] from the gradient of a plot of $\ln K_a$ *vs.* $1/T$. This value is consistent with a moderate hydrogen-bonding interaction.[62] d_6-DMSO typically competes for hydrogen-bonding sites[47] and addition (0.1 M) to the system commensurately reduced the K_{a1} to a very low level, namely from 18.5 ± 2.9 to 0.43 ± 0.06 for the *trans*-rotamer and from 28.0 ± 8.1 to 0.93 ± 0.30 for the *cis*-rotamer. This suggests that hydrogen bonding may be important in maintaining the complex.

The ΔG values of the association of the *trans*-rotamer of *N*-Ac-Pro-OMe **2** with phenol **4** were calculated from the experimental K_a values at each temperature and the value of ΔH derived above (Table 2). These values are all consistent, all being around 7 kJ mol^{-1}, reflective of a process that is favoured at the temperatures examined.

Table 2 α-^{1}H NMR titration results for *trans*- and *cis*-*N*-Ac-Pro-OMe **2**—phenol **4** association at different temperatures with corresponding thermodynamic data

	Temperature/K	K_{a1}/mol^{-1} dm^3	K_{a2}/mol^{-1} dm^3	ΔG/kJ mol^{-1}	ΔH/kJ mol^{-1}	ΔS/J K^{-1}
trans-Rotamer	253	31.3 ± 4.9	5.3 ± 1.7	−7.2 ± 0.7		
	273	21.4 ± 3.6	6.8 ± 3.4	−7.0 ± 0.8		
	293	18.5 ± 2.9	—	−7.1 ± 0.8	−10.3 ± 3.9	−10.8 ± 5.9
	313	11.5 ± 4.0	—	−6.4 ± 1.9		
cis-Rotamer	253	39.2 ± 10.9	3.4 ± 1.4			
	273	31.9 ± 9.1	2.2 ± 4.5			
	293	28.0 ± 8.1	7.0 ± 1.4			
	313	11.1 ± 8.5	—			

Electrostatic surface calculations

The calculated electrostatic surface properties of the molecules allow the analysis of the most likely modes of association. Hunter has recently postulated a simplified model for molecules whereby interactions between two species in solution can be predicted primarily by their electrostatic charges.[47] Using this model, the calculated electrostatic surface charges of the amide **2** and phenol **4** predict that the most likely interaction would be a hydrogen bonding interaction between the E_{max} of phenol, namely the phenolic hydrogen, and the E_{min} of the amide, *viz* the amide carbonyl oxygen (Fig. 5). This observation is consistent with previous literature on the binding of phenols with amides.[63–66]

Hydrogen-bond donor and acceptor constants (α and β,[67] Table 3) were calculated for the molecules **1–4** and the solvents chloroform and water from the corresponding E_{max} and E_{min} values. The free energy of hydrogen-bonding interaction ($\Delta\Delta G_{H\text{-bond}}$) can be crudely predicted from these numbers.[47] This gives a rough estimation of the K_a if the hydrogen bonding interaction were the overwhelming driving force for binding. Since the carbonyl of the ester could also act as an alternative hydrogen-bonding site, the E_{ester} value is also presented, along with the corresponding hydrogen bond acceptor constant, denoted ε here. The values of E_{ester} are all significantly smaller than E_{min}, confirming that the amide is the preferred binding site for the amino acid derivatives investigated in this study. Calculating $\Delta\Delta G$ and the corresponding K_a for water (Table 4), indicates that hydrogen bonding is negligible in aqueous solution, consistent with the assertion that hydrogen bonding is not the main driving association between tannins and proteins in solution.[22] The same calculation for association in CCl_4 affords K_a values highly consistent with the magnitude expected for phenol–amide complexation, based on reported values in the literature.[65,66] This correlates with the evidence supporting hydrogen bonding as the major interaction in this solvent.

The predicted K_a values for the derivatives **1a–e**, **2** and **3** in $CHCl_3$ are all significantly smaller than the measured K_a values. Furthermore, the predicted binding order contrasts with the trends seen for the measured K_a values for these species. To rationalise the observed binding, an interaction of the phenol ring with the hydrogen atoms of either the NCH_3 of the sarcosine **1b** or the NCH_2 of the pyrrolidine derivatives **2** and **3** can be proposed. Related CH–π interactions have been recognised in peptides in the form of phenylalanine–proline interactions,[68] and may contribute enough of either an additive or synergistic effect to improve the binding to a measurable level, relative to a derivative with only a secondary amide. This would be consistent with already well-established work that points to the influence of a strong 'hydrophobic effect' in the binding of larger tannins to polypeptides,[22] and particularly to a specific interaction with the CH groups of the proline.[37–39]

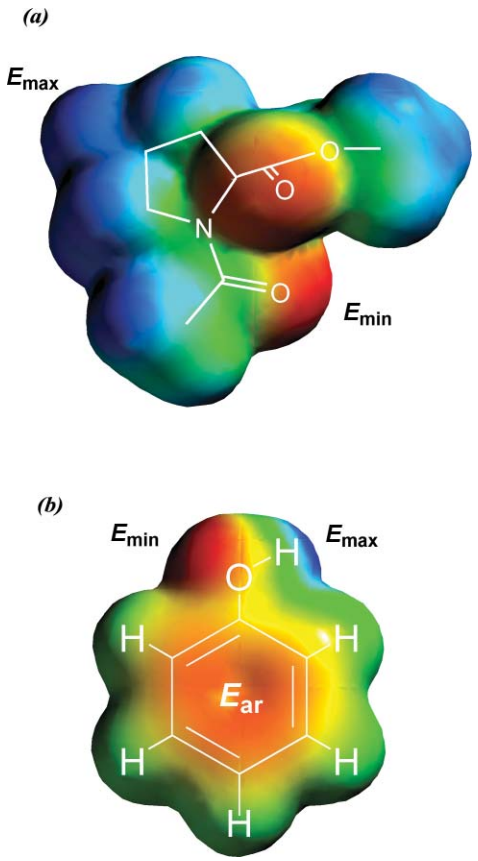


Fig. 5 Molecular electrostatic surfaces of: (a) *trans*-*N*-Ac-Pro-OMe **2**, and (b) phenol **4**, plotted on a van der Waals' surface calculated using B3LYP/6-31G(d). Positive regions are shown in blue, negative regions are shown in red, and neutral regions are shown in green. The locations of the maximum electrostatic potential, E_{max}, the minimum electrostatic potential, E_{min}, and the electrostatic potential of the aromatic ring (E_{ar}) are labelled.

A small increase in the $\Delta\Delta G$ of binding, to account for CH–π and related interactions, is able to account for the observed binding constants satisfactorily. The higher binding observed for the pyrrolidine derivatives **2** and **3**, relative to the sarcosine **1b** can be rationalised through an increased number of possible CH–π interactions.

Table 3 Electrostatic surface charges for compounds **1–4** obtained at the B3LYP/6-31G(d) level with Spartan and derived α, β and ε constants,[47] with calculated ΔΔ*G* for the hydrogen bond interaction and predicted and experimental K_as

Molecule	E_{min}/kJ mol^{-1}	E_{max}/kJ mol^{-1}	E_{ester}/kJ mol^{-1}				$\Delta\Delta G_{H\text{-}bond}$ [h]/ kJ mol^{-1}	Predicted K_a/mol^{-1} dm^3
N-Ac-Gly-OMe **1a**	−211.9[a]	171.6 [b]	−146.4	3.3	4.1	2.8	−5.7	0.74
trans-*N*-Ac-Sar-OMe **1b**	−209.5 [a]	123.6 [c]	−146.4	2.4	2.0	2.8	−5.6	0.68
cis-*N*-Ac-Sar-OMe **1b**	−211.5 [a]	124.2 [c]	−143.2	2.4	4.1	2.8	−5.7	0.72
N-Ac-Ala-OMe **1c**	−203.3 [a]	174.6 [b]	−133.9	3.4	3.9	2.6	−5.4	0.56
N-Ac-Phe-OMe **1d**	−211.0 [a]	184.0 [b]	−125.5	3.5	4.1	2.4	−5.6	0.71
N-Ac-O-Bzl-Tyr-OMe **1e**	−210.2 [a]	178.9 [b]	−131.2	3.4	4.0	2.5	−5.6	0.70
trans-*N*-Ac-Pro-OMe **2**	−209.0 [a]	100.1 [d]	−158.9	1.9	4.0	3.1	−5.6	0.67
cis-*N*-Ac-Pro-OMe **2**	−217.6 [a]	107.8 [e]	−152.3	2.1	4.2	2.9	−5.9	0.88
N-Ac-Pyr **3**	−219.2 [a]	86.6 [d]	—	1.7	4.2		−5.9	0.92
Phenol **4**	−133.4 [f]	258.0 [g]	—	5.0	2.6	—	−3.0	0.06
Chloroform	−43.6	167.7	—	3.2	0.8	—		
Water	−189.1	216.0	—	4.2	3.6	—		
Carbon tetrachloride	−22.6	96.5	—	1.9	0.4	—		

[a] Located over the amide carbonyl oxygen. [b] Located over the amide NH hydrogen. [c] Located on the CH_2 hydrogens. [d] Located over the ring hydrogens. [e] Located over the OMe protons. [f] Located over the OH oxygen. [g] Located over the OH hydrogen. [h] Calculated with chloroform as solvent at 298 K.

Table 4 Predicted ΔΔ*G* and association constants (K_a) of selected molecules with phenol **4** in H_2O and CCl_4, based on the B3LYP/6-31G(d) derived α and β values (Table 3)

Molecule	ΔΔ*G* (H_2O)	K_a (H_2O)	ΔΔ*G* (CCl_4)	K_a (CCl_4)
N-Ac-Gly-OMe **1a**	−0.4	0.01	−11.4	145.59
trans-*N*-Ac-Sar-OMe **1b**	−0.3	0.01	−11.2	127.02
cis-*N*-Ac-Sar-OMe **1b**	−0.4	0.01	−11.3	142.01
N-Ac-Ala-OMe **1c**	−0.2	0.00	−10.8	90.04
N-Ac-Phe-OMe **1d**	−0.3	0.01	−11.3	138.18
N-Ac-O-Bzl-Tyr-OMe **1e**	−0.3	0.01	−11.3	132.08
trans-*N*-Ac-Pro-OMe **2**	−0.3	0.01	−11.2	123.50
cis-*N*-Ac-Pro-OMe **2**	−0.5	0.01	−11.7	199.53
N-Ac-Pyr **3**	−0.5	0.01	−11.8	218.15
Phenol **4**	0.9	0.00	−6.6	1.82

The strength of these combined interactions is clearly mediated by the orientation of the amide, as can be ascertained from the differences seen between the *cis* and *trans* rotamers of the *N*-acetylproline methyl ester **2**. This is consistent with an underlying, directing hydrogen-bonding interaction. An additional contribution of around 3.6 kJ mol^{-1} from CH–π interactions for each rotamer is then sufficient to account for the observed K_as and is entirely consistent with the expected energetic contribution from this class of interaction. This estimation also adequately accounts for the effect of hydrogen-bond disruption seen in d_6-DMSO, as, even if these interactions are fully retained, the energetic contribution is insufficient to maintain a measurable binding.

The same magnitude of binding might be expected for the *N*-acetylpyrrolidine **3**, and, in concert with the predicted hydrogen bonding, a K_a slightly larger than that for the *cis*-*N*-acetylproline methyl ester **2** might be anticipated. The discrepancy in the experimental observation could be attributed to error within the experimental measurements, but might also be attributed either to differences in the degree of rotational freedom, or the electron withdrawing nature of the carboxylic side chain. Clearly, more detailed studies would be required to delineate these effects.

Conclusions

^{1}H NMR titrations of minimal peptide-based models with phenol **4** have established that, in $CHCl_3$ solution, there is a selective binding of proline-based peptides, relative to other similarly protected amino acid residues. This binding appears to be mediated by interactions with the tertiary amide structure primarily, with pyrrolidine-based structures **2** and **3** having the largest interactions, followed by the less-substituted sarcosine derivative **1b** and negligible binding for the secondary amide derivatives **1a** and **1c–e**.

Binding to the *trans*-rotamer of *N*-acetylproline methyl ester **2** appears to be in a roughly 1 : 1 ratio at 298 K, whereas titration of the *cis*-rotamer with phenol shows a secondary, smaller binding with different overall stoichiometry of around 2 : 1. This secondary association is also observed with the *trans*-rotamer **2** in cooled solutions. Thermodynamic analysis indicates that the association for both rotamers is an enthalpically favoured process. Interactions with d_6-DMSO indicate that hydrogen bonding is likely to be important in the association mechanism.

Electrostatic calculations, based on Hunter's molecular recognition toolbox methods, were used to help identify the contribution of hydrogen bonding. These calculations confirm that the preferred electrostatic interaction is that between the amide carbonyl and the phenolic hydroxyl, in line with previous studies on H-bonding carried out in CCl_4.[63–66] The predicted K_as for $CHCl_3$ as solvent illustrate that this hydrogen bond alone is insufficient to account for the binding observed in the experimental system. However, an additional, small contribution from proposed CH–π interactions is sufficient to explain the experimental observations. This is consistent with the work on full protein–tannin systems that indicate in polar solvents that hydrogen bonding cannot be the major driving force and that 'hydrophobic' interactions direct binding,[22,37–39] illustrating that these are reasonable models for the larger systems and the utility of combining experimental observations with theoretical calculations.

Whilst it is recognised that H-bonding is crucial to the overall association, it cannot work as the only factor in mediating the association. CH–π interactions offer a clear residue-level driving

force for the preferential binding of phenolics to proline residues in PRPs, but in our systems are also unlikely to be strong enough to act alone. Thus the two interactions must work together, and in doing so afford the selectivity observed in larger systems. An additive combination of interactions is sufficient to account for the binding in the reported systems. Further work is ongoing within our group to more clearly delineate the particular contributions of each interaction.

Acknowledgements

The authors would like to thank Dr J. B. Harper for extended useful discussions and assistance with Mac-Curve-Fit.

References

1 K.-T. Chung, T. Y. Wong, C.-I. Wei, Y.-W. Huang and Y. Lin, *Crit. Rev. Food Sci. Nutr.*, 1998, **38**, 421–464.
2 H. M. Appel, *J. Chem. Ecol.*, 1993, **19**, 1521–1551.
3 H. Mehansho, L. G. Butler and D. M. Carlson, *Annu. Rev. Nutr.*, 1987, 423–440.
4 Y. Chen and A. E. Hagerman, *J. Agric. Food Chem.*, 2004, **52**, 6061–6067.
5 M. J. T. J. Arts, G. R. M. M. Haeman, L. C. Wilms, S. A. J. N. Beetstra, C. G. M. Heijnem, H.-P. Voss and A. Bast, *J. Agric. Food Chem.*, 2002, **50**, 1184–1187.
6 K. M. Riedl and A. E. Hagerman, *J. Agric. Food Chem.*, 2001, **49**, 4917–4923.
7 S. Rohn, H. M. Rawel and J. Kroll, *J. Agric. Food Chem.*, 2004, **52**, 4725–4829.
8 W. Wang and M. T. Goodman, *Nutr. Res.*, 1999, **19**, 191–202.
9 A. Bennick, *Crit. Rev. Oral Biol. Med.*, 2002, **13**, 184–196.
10 H. Mehansho, D. K. Ann, L. G. Butler, J. C. Rogler and D. M. Carlson, *J. Biol. Chem.*, 1987, **262**, 12344–12350.
11 H. Mehansho, T. N. Asquith, L. G. Butler, J. C. Rogler and D. M. Carlson, *J. Agric. Food Chem.*, 1992, **40**, 93–97.
12 H. Mehansho, S. Clements, B. T. Sheares, S. Smith and D. M. Carlson, *J. Biol. Chem.*, 1985, **260**, 4418–4423.
13 E. Haslam, *Biochem. J.*, 1974, **139**, 285–288.
14 J. P. McManus, K. G. Davis, T. H. Lilley and E. Haslam, *J. Chem. Soc., Chem. Commun.*, 1981, 309–311.
15 J. P. McManus, K. G. Davis, J. E. Beart, S. H. Gaffney, T. H. Lilley and E. Haslam, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1429–1438.
16 C. M. Spencer, Y. Cai, R. Martin, S. H. Gaffney, P. N. Goulding, D. Magnolato, T. H. Lilley and E. Haslam, *Phytochemistry*, 1988, **27**, 2397–2409.
17 J. E. Beart, T. H. Lilley and E. Haslam, *Phytochemistry*, 1985, **24**, 33–38.
18 K. S. Feldman, A. Sambandam, S. T. Lemon, R. B. Nicewonger, G. S. Long, D. F. Battaglia, S. M. Ensel and M. A. Laci, *Phytochemistry*, 1999, **51**, 867–872.
19 R. A. Frazier, A. Papadopoulou, I. Mueller-Harvey, D. Kisoon and R. J. Green, *J. Agric. Food Chem.*, 2003, **51**, 5189–5195.
20 A. E. Hagerman and L. G. Butler, *J. Agric. Food Chem.*, 1980, **28**, 944–947.
21 A. E. Hagerman and L. G. Butler, *J. Biol. Chem.*, 1981, **256**, 4494–4497.
22 H. I. Oh, J. E. Hoff, G. S. Armstrong and L. A. Haff, *J. Agric. Food Chem.*, 1980, **28**, 394–398.
23 T. N. Asquith and L. G. Butler, *Phytochemistry*, 1986, **25**, 1591–1593.
24 M. Zhu, J. D. Phillipson, P. M. Greengrass, N. E. Bowery and Y. Cai, *Phytochemistry*, 1997, **44**, 441–447.
25 A. E. Hagerman, M. E. Rice and N. T. Ritchard, *J. Agric. Food Chem.*, 1998, **46**, 2590–2595.
26 K. J. Siebert, N. V. Troukhanova and P. Y. Lynn, *J. Agric. Food Chem.*, 1996, **44**, 80–85.
27 V. A. P. deFreitas and N. Mateus, *J. Sci. Food Agric.*, 2002, **82**, 113–119.
28 V. A. P. deFreitas and N. Mateus, *J. Agric. Food Chem.*, 2001, **49**, 940–945.
29 J. R. Bacon and M. J. C. Rhodes, *J. Agric. Food Chem.*, 1998, **46**, 5083–5088.
30 S. Verge, T. Richard, S. Moreau, S. Richelme-David, J. Vercauteren, J.-C. Prome and J.-P. Monti, *Tetrahedron Lett.*, 2002, **43**, 2363–2366.
31 Y. Chen and A. E. Hagerman, *J. Agric. Food Chem.*, 2004, **52**, 4008–4011.
32 A. J. Charlton, N. J. Baxter, M. L. Khan, A. J. G. Moir, E. Haslam, A. P. Davies and M. P. Williamson, *J. Agric. Food Chem.*, 2002, **50**, 1593–1601.
33 K. Wroblewski, R. Muhandiram, A. Chakrabartty and A. Bennick, *Eur. J. Biochem.*, 2001, **268**, 4384–4397.
34 N. Naurato, P. Wong, Y. Lu, K. Wroblewski and A. Bennick, *J. Agric. Food Chem.*, 1999, **47**, 2229–2234.
35 A. J. Charlton, N. J. Baxter, T. H. Lilley, E. Haslam, C. J. McDonald and M. P. Williamson, *FEBS Lett.*, 1996, **382**, 289–292.
36 C. Simon, K. Barathieu, M. Laguerre, J. M. Schmitter, E. Fouquet, I. Pianet and E. J. Dufourc, *Biochemistry*, 2003, **42**, 10385–10395.
37 A. J. Charlton, E. Haslam and M. P. Williamson, *J. Am. Chem. Soc.*, 2002, **124**, 9899–9905.
38 N. J. Murray, M. P. Williamson, T. H. Lilley and E. Haslam, *Eur. J. Biochem.*, 1994, **219**, 923–925.
39 N. J. Baxter, T. H. Lilley, E. Haslam and M. P. Williamson, *Biochemistry*, 1997, **36**, 5566–5577.
40 J. L. Goldstein and T. Swain, *Phytochemistry*, 1963, **2**, 371–383.
41 E. C. Bate-Smith, *Phytochemistry*, 1973, **12**, 907–912.
42 L. J. Porter and J. Woodruffe, *Phytochemistry*, 1984, **23**, 1255–1256.
43 T. Ozawa, T. H. Lilley and E. Haslam, *Phytochemistry*, 1987, **26**, 2937–2942.
44 G. Luck, H. Liao, N. J. Murray, H. R. Grimmer, E. E. Warminski, M. P. Williamson, T. H. Lilley and E. Haslam, *Phytochemistry*, 1994, **37**, 357–371.
45 E. Jobstl, J. O'Connell, J. P. A. Fairclough and M. P. Williamson, *Biomacromolecules*, 2004, **5**, 942–949.
46 I. Lesschaeve and A. C. Noble, *Am. J. Clin. Nutr.*, 2005, **81**, 330S–335S.
47 C. A. Hunter, *Angew. Chem.*, 2004, **43**, 5310–5324.
48 R. G. Webb, M. W. Haskell and C. H. Stammer, *J. Org. Chem.*, 1969, **34**, 576–580.
49 A. G. Griesbeck, S. Bondock and J. Lex, *J. Org. Chem.*, 2003, **68**, 9899–9906.
50 M. Prein, P. J. Munley and A. Padwa, *Tetrahedron*, 1997, **53**, 7777–7794.
51 S. Tsuboi, Y. Nooda and A. Takeda, *J. Org. Chem.*, 1984, **49**, 1204–1208.
52 J. W. Steed and J. L. Atwood, in *Supramolecular Chemistry*, John Wiley and Sons Ltd., Chichester, England, 2000, pp. 16–19.
53 C. A. Irvine, *Spartan '02*, Wavefunction, Inc., 2002.
54 J. Kong, C. A. White, A. I. Krylov, C. D. Sherrill, R. D. Adamson, T. R. Furlani, M. S. Lee, A. M. Lee, S. R. Gwaltney, T. R. Adams, C. Ochsenfeld, A. T. B. Gilbert, G. S. Kedziora, V. A. Rassolov, D. R. Maurice, N. Nair, Y. Shao, N. A. Besley, P. E. Maslen, J. P. Dombroski, H. Daschel, W. Zhang, P. P. Korambath, J. Baker, E. F. C. Byrd, T. V. Voorhis, M. Oumi, S. Hirata, C.-P. Hsu, N. Ishikawa, J. Florian, A. Warshel, B. G. Johnson, P. M. W. Gill, M. Head-Gordon and J. A. Pople, *J. Comput. Chem.*, 2000, **21**, 1532–1548.
55 D. A. Block, D. Yu, D. A. Armstrong and A. Rauk, *Can. J. Chem.*, 1998, **76**, 1042–1049.
56 F. X. Schmid, *Adv. Protein Chem.*, 2002, **59**, 243–282.
57 K. R. Williams, B. Adhyaru, I. German and E. Alvarez, *J. Chem. Educ.*, 2002, **79**, 372–373.
58 W. J. Wedemeyer, E. Welker and H. A. Scheraga, *Biochemistry*, 2002, **41**, 14637–14644.
59 C. Dugave and L. Demange, *Chem. Rev.*, 2003, **103**, 2375–2532.
60 J. M. Sanderson and E. J. Whelan, *Phys. Chem. Chem. Phys.*, 2004, **6**, 1012–1017.
61 P. W. Atkins, *The Elements of Physical Chemistry*, Oxford University Press, Oxford, 1996.
62 J. W. Steed and J. L. Atwood, *Supramolecular Chemistry*, John Wiley & Sons, Ltd, Chitchester, 2000.
63 J.-Y. Le Questel, C. Laurence, A. Lachkar, M. Helbert and M. Berthelot, *J. Chem. Soc., Perkin Trans. 2*, 1992, 2091–2094.
64 M. D. Joesten and R. S. Drago, *J. Am. Chem. Soc.*, 1962, **84**, 2696–2699.
65 T. Gramstad and W. J. Fuglevik, *Acta Chem. Scand.*, 1962, **16**, 1369–1377.
66 E. Kwiatkowski, K. Kozubek and Z. Pepliński, *Z. Naturforsch., B: Anorganische, organische und biologische Chemie, Botanik, Zoologie und verwandte Gebiete*, 1978, **33**, 230–236.
67 M. H. Abraham and J. A. Platts, *J. Org. Chem.*, 2001, **66**, 3484–3491.
68 Z. Ciunik, S. Berski, Z. Latajka and J. Leszczynski, *J. Mol. Struct.*, 1998, **442**, 125–134.

