ARTICLE

OBC
www.rsc.org/obc

Molecular recognition in the gas phase ligand switching reactions of the proton bound dimer of sarcosine and glycylglycine.†‡

Richard A. J. O'Hair* and Ana K. Vrkic
School of Chemistry, University of Melbourne, Victoria 3010, Australia

Received 9th October 2002, Accepted 20th December 2002
First published as an Advance Article on the web 4th February 2003

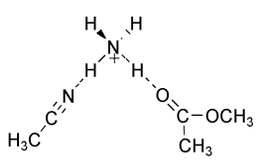
Ion–molecule reactions of the proton bound dimer of $[Sar+H+GlyGly]^+$ (where Sar = sarcosine and GlyGly = glycylglycine) proceed via two main reaction channels, *i.e.* association and ligand switching. The association reaction, which involves formation of an adduct between the protonated dimer and neutral base, occurs more readily for the oxygen containing bases and those with a lower gas phase basicity. Molecular recognition was demonstrated for the ligand switching reactions in which nitrogen containing bases preferred to switch out sarcosine while the oxygen containing bases preferred to switch out glycylglycine. Molecular dynamics followed by semiempirical PM3 calculations for the ligand switching reactions of $[Sar+H+GlyGly]^+$ with methylamine directly correlated with the experimental findings by predicting that the most stable product ion arises from switching out sarcosine. These calculations reveal that the most stable adduct structure and the most stable ligand switched structure arise from proton transfer to methylamine to yield ions of the type $[(Sar)(GlyGly)(LigH^+)]$ and $[(GlyGly)(LigH^+)]$.

Introduction

The four "S" advantages of mass spectrometry (specificity, sensitivity, speed and stoichiometry) coupled with soft ionization modes (such as electrospray ionization, ESI, and matrix assisted laser desorption ionization, MALDI) have made it an important complimentary analytical method for the examination of non-covalent interactions of biological interest.[1] Chemists have also been intrigued by the possibility of examining the fundamental properties (structure and reactivity) of these complexes, giving rise to the new area of gas phase molecular recognition and supramolecular chemistry.[2] Some of the "tools" available to probe the gas phase chemistry of non-covalent complexes include: (i) various tandem mass spectrometry (MS/MS) techniques such as collision induced dissociation (CID), blackbody infrared radiative dissociation (BIRD); (ii) bimolecular reactions including ion–molecule and ion–ion reactions; (iii) ion-mobility measurements; (iv) and spectroscopy. We note some recent highlights using each of these tools: (i) the modified kinetic method of Cooks[3] has been used to determine the chirality of a range of analytes,[4] BIRD has been used to show that Watson–Crick binding of double stranded DNA is preserved in the gas phase[5] and sustained off-resonance irradiated (SORI) CID has been used to determine the relative gas phase binding energies of inhibitors of carbonic anhydrase;[6] (ii) Lebrilla has used ion–molecule reactions of host–guest complexes to determine chirality of a range of analytes[7] and we have shown that trimethyl borate can been used as a gas phase "crosslinking reagent" for non-covalent anionic complexes containing phosphate moieties;[8] (iii) ion-mobility has been used to determine the gas phase structure of novel serine octamers;[9] (iv) REMPI has been used to provide exquisite structural and spectroscopic detail on hydrogen bonding in isolated neutral guanine–cytosine (G-C) and guanine–guanine (G-G) base pairs, which demonstrates gas phase Watson–Crick binding.[10]

In a landmark paper, Feng, Ling and Lifshitz used a selected ion flow tube apparatus to demonstrate molecular recognition in the ligand switching reactions of the proton bound heterodimer of acetonitrile and methylacetate.[11a] They found that both the rates of reaction as well as the branching ratios (*i.e.* relative product yields) of each of the three reaction channels (adduct formation in eqn. (1), ligand switching of methylacetate in eqn. (2) and ligand switching of acetonitrile in eqn. (3)) were highly dependent on the properties of the neutral ligand (Lig). Thus they found that nitrogen containing ligands preferentially switch out methylacetate (eqn. (2)) while oxygen-containing molecules preferentially switch out acetonitrile (eqn. (3)). These results were rationalized in terms of the energetics of these ligand switching reactions with the $[–O–H^+ \cdots O–]$ and $[–N–H^+ \cdots N–]$ binding energies being higher than those for $[–O–H^+ \cdots N–]$ and $[–N–H^+ \cdots O–]$.[11] In addition, the observation of adduct ions was ascribed to the formation of insertion complexes in which a protonated ligand is doubly hydrogen bonded to two other ligands as shown for ammonia in **(A)**. Evidence for these insertion reactions came from the observation that tertiary amines do not undergo these reactions,[11a] from multiply hydrogen bonded complexes in ether systems[11b] and also from *ab initio* calculations.[11b,c]

$$[CH_3CN+CH_3CO_2CH_3+H]^+ + Lig \rightarrow \begin{cases} [CH_3CN+CH_3CO_2CH_3+Lig+H]^+ & (1) \\ [CH_3CN+Lig+H]^+ + CH_3CO_2CH_3 & (2) \\ [CH_3CO_2CH_3+Lig+H]^+ + CH_3CN & (3) \end{cases}$$



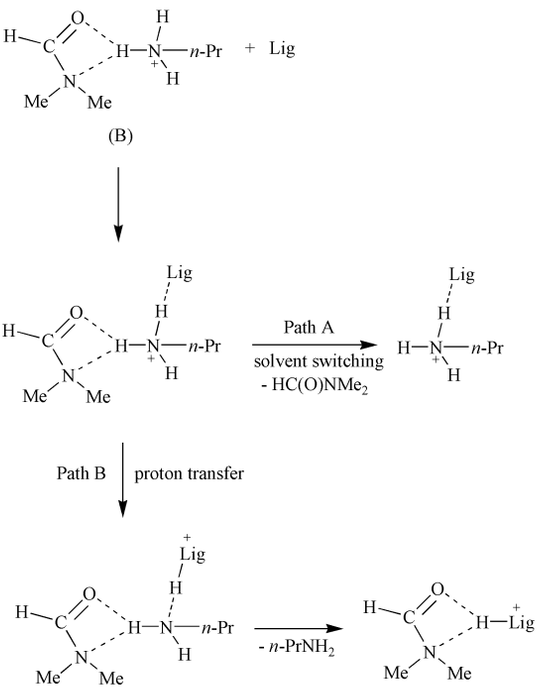
(A)

Very recently, Witt and Grützmacher have used FT-ICR mass spectrometry to examine the ligand exchange reactions of the proton bound heterodimer of *N*,*N*-dimethylformamide (DMF) and *n*-propyl amine with 14 reagents.[12] They also observed striking selectivity in these ligand switching reactions, which depended on two main properties of the neutral ligand:

† Part 35 of the series "Gas Phase Ion Chemistry of Biomolecules". Part 34: A. K. Vrkic and R. A. J. O'Hair, Gas phase reactions of trimethylborate with the $[M-H]^-$ ions of nucleotides and their non-covalent homo and heterodimer complexes, *Aust. J. Chem.*, in press.
‡ Electronic supplementary information (ESI) available: heats of formation of adducts and PM3-optimized adduct structures. See http://www.rsc.org/suppdata/ob/b2/b209848b/

DOI: 10.1039/b209848b

(i) its polarity; and (ii) its proton affinity. Thus the DMF ligand is selectively exchanged by polar reactants while the *n*-propylamine is selectively exchanged by a more basic amine. Furthermore, the rate of reactivity is also dependent on the ligand. They ascribed this difference in reactivity using a "solvation model" for the structure of the heterodimers, shown in Scheme 1. They suggest that the reactant dimer has the structure **(B)**,



Scheme 1

consisting of an ammonium ion solvated by the amide. Thus exchange of the amide by another polar reactant corresponds to a simple switching of solvent molecules (Path A of Scheme 1), while the exchange of the amine ligand by a more basic amine is essentially an exothermic proton transfer, with the amide acting as the solvent molecule of the complex which follows the migrating charge (Path B of Scheme 1).

Since discovering that H/D exchange of the proton bound dimer of sarcosine (Sar) and glycylglycine (GlyGly), $[Sar+H+GlyGly]^+$ proceeds more readily than that of the monomer ions,[13] we have been interested in examining other gas phase reactions of this ion. Unfortunately, CID of $[Sar+H+GlyGly]^+$ provides little information (both protonated monomers are formed in relative abundances which reflect their relative proton affinities), and so herein we evaluate the role of molecular recognition in the ligand switching reactions of this non-covalent complex. In particular, we are interested in whether the dramatic differences in ligand switching observed for simple proton bound heterodimers[11,12] occurs for a more complex heterodimer capable of multiple hydrogen bonding.

All of the ion–molecule reactions of proton bound dimers discussed above involve the intermediacy of a trimeric complex. Recently, related trimer complexes of biological significance have been formed *via* ESI and their CID reactions have been used to gain insights into relative binding affinities and/or chiral recognition. Although the general applicability of such an approach remains to be demonstrated, recent papers have shown chiral recognition in amino acids[14a,b] and non-covalent complexes between vancomycin antibiotics and peptide ligand stereoisomers.[14c] While the study of gas phase ionic trimeric complexes is still in its infancy, it is interesting to note their relationship to ion–molecule complexes, which have been shown to play important roles in both ion–molecule reactions and in fragmentation reactions.[15]

Experimental

(a) Mass spectrometry

All experiments were performed using a commercially available quadrupole ion trap mass spectrometer (Finnigan-MAT model LCQ, San Jose, CA) equipped with electrospray ionization (ESI) and recently modified to allow the introduction of neutral reagents via the helium background gas inlet line.[8,13,16] All reagents were commercially obtained and used without further purification.

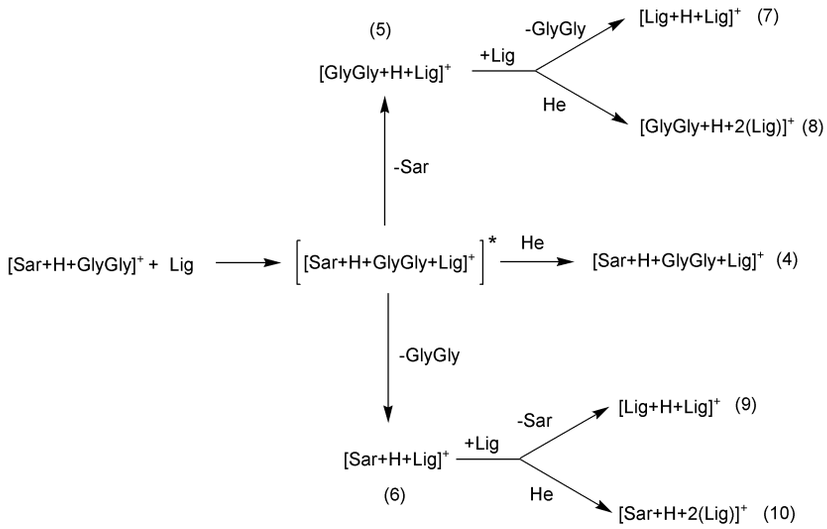
A 2 : 1 mixture of sarcosine and glycylglycine was dissolved in a 50 : 50 mixture of H_2O and CH_3OH (1% acetic acid) (0.1 mg mL^{-1}) and introduced to the mass spectrometer at 3.0 μL min^{-1} *via* electrospray ionization. Typical ESI conditions used were: spray voltage, 5.0 kV, capillary temperature, 200 °C, nitrogen sheath pressure, 40 psi, and capillary voltage/tube lens offset, 0–10V. Ion–molecule reactions were carried out as previously described.[8] Briefly, once a stable ESI ion signal was obtained, the neutral reagents were introduced into the trap as a part of the helium bath gas. A constant flow of the reagent (5–200 μL h^{-1}) was established using a syringe pump with the syringe needle directed into a measured flow of helium (850–4601 mL min^{-1}). The majority of the gas exits through a flowmeter, whereas a small amount (~1 mL min^{-1}) is drawn into the trap. The LCQ uses a constriction capillary to control the helium flow and is designed to maintain 1.75 mtorr in the trap when 3 psi of He pressure is applied to the capillary. In the stock system, the 3 psi is maintained by an internal regulator that steps down the 40 psi of He that is delivered at the external port. To avoid the dead volume in the internal regulator, we bypass it and deliver the He mixture (3 psi) directly to the capillary. This greatly decreases the lag time after changes in reagent concentration. Once an appropriate flow of the neutral reagent was established, the reagent pressure was allowed to equilibrate to a steady state over several minutes. Reported rates are the average of at least three kinetic runs using two to three (in the majority of cases) reagent flow rates. Standard deviations in absolute rate constants were typically <10%, except for the very slow reactions, which were <20%. A conservative estimate of error is ±25%, but relative rates are expected to be more accurate due to cancellation of errors. Branching ratios were extrapolated to zero reaction time in order to account for the effects of secondary reactions, and are estimated to have error limits of ±20%.

Previous studies by Gronert on a nearly identically modified LCQ, have shown that these systems give ion–molecule reaction rate constants comparable to those from flowing afterglow instruments[17a] and that the ions in the trap are essentially at ambient temperature (~300 K).[17b] Thus products from ion–molecule reactions can be collisionally cooled by the helium bath gas.

(b) Molecular modeling procedures

Molecular dynamics calculations were used to generate candidate structures for subsequent structural optimizations via molecular orbital calculations. The Insight II suite of programs (Molecular Simulations Inc. San Diego, CA, USA) was used to build and run the molecular dynamics simulations using the Discover program in conjunction with the cff91 force field. Semi-empirical molecular orbital calculations were performed using the Insight II molecular modeling package at the PM3 level of theory.

We note that we use two different ways to describe the proton-bound dimer and trimer ions. In those instances where we do not specify which species is protonated, we show each species (including the proton) in a square bracket. For example, the proton bound dimer of glycylglycine and sarcosine is designated as $[Sar+H+GlyGly]^+$. In contrast, when calculations have been carried out, the ligand which is protonated needs to



Scheme 2

be assigned. Thus we have adopted a different nomenclature, whereby each species in a complex is surrounded by a bracket, with the protonated species having an additional H. For example, the dimer consisting of protonated glycylglycine and sarcosine is designated as [(Sar)(GlyGlyH$^+$)].

Structural conformations of the trimer, dimer and monomeric species were explored as follows. Candidate structures of the proton bound trimers consisting of either the protonated sarcosine, neutral glycylglycine and neutral methylamine (Lig) [(SarH$^+$)(GlyGly)(Lig)], or protonated glycylglycine, neutral sarcosine and neutral methylamine [(GlyGlyH$^+$)(Sar)(Lig)], or protonated methylamine, neutral sarcosine and neutral glycylglycine [(LigH$^+$)(Sar)(GlyGly)] were generated by subjecting an initial conformer to 10 ps of dynamics at 300K. The resultant structure was minimized using MM and stored. This process was repeated 300 times to generate 300 candidate structures. The most stable structures were then optimized at the PM3 semi-empirical level of theory. Each of these optimized trimeric structures were then separated into the appropriate dimers (e.g. [(GlyGlyH$^+$)(Sar)+(Lig)] and [(GlyGlyH$^+$)(Lig)+(Sar)] and again optimized at the PM3 semi-empirical level of theory. These low energy structures were then subjected to the same molecular dynamics simulations as described above to explore the conformational space of each of the dimeric ions as well as the neutrals. Once again, the most stable structures were optimized at the PM3 level of theory. The same process was repeated to explore the conformational space of the monomers.

Results and discussion

Scheme 2 illustrates the various primary and secondary reaction channels available for neutral ligands reacting with [Sar+H+GlyGly]$^+$. Thus the initially formed [Sar+H+GlyGly+Lig]$^+$ trimeric complex can undergo collisional stabilization with the helium bath gas (eqn. (4) of Scheme 2) or it can switch out either sarcosine (eqn. (5) of Scheme 2) or glycylglycine (eqn. (6) of Scheme 2). The ligand substitution products [GlyGly+H+Lig]$^+$ and [Sar+H+Lig]$^+$ can undergo secondary reactions with another neutral ligand (eqns. (7)–(10) of Scheme 2). Table 1 lists the rates of reaction, reaction efficiencies, reaction channels and branching ratios of the ligand switching ion–molecule reactions of [Sar+H+GlyGly]$^+$ with 37 neutral ligands of two main classes (oxygen versus nitrogen nucleophiles/bases) in the order of increasing gas phase basicity of the ligand. The gas phase basicities of glycylglycine (GlyGly) and sarcosine (Sar), which are 882 and 888.7 kJ mol^{-1}, are highlighted on Table 1. Since Witt and Grützmacher have noted that ligand polarity also plays an important role in ligand exchange reactions, we have included the dipole moments of each of the ligands in Table 1.[12]

An examination of Table 1 reveals that the branching ratios are highly dependent on the type of ligand. In fact they have been extrapolated to zero reaction time in order to account for the effects of secondary reactions. This is illustrated in Fig. 1A and 1B for *n*-butylamine and *N*,*N*-diethylacetamide, two ligands of similar basicity but different structure/reactivity modes (nitrogen versus oxygen nucleophile/base). Thus butylamine (gas phase basicity (GB) = 886.6 kJ mol^{-1}), a nitrogen nucleophile/base, undergoes preferential ligand switching of sarcosine

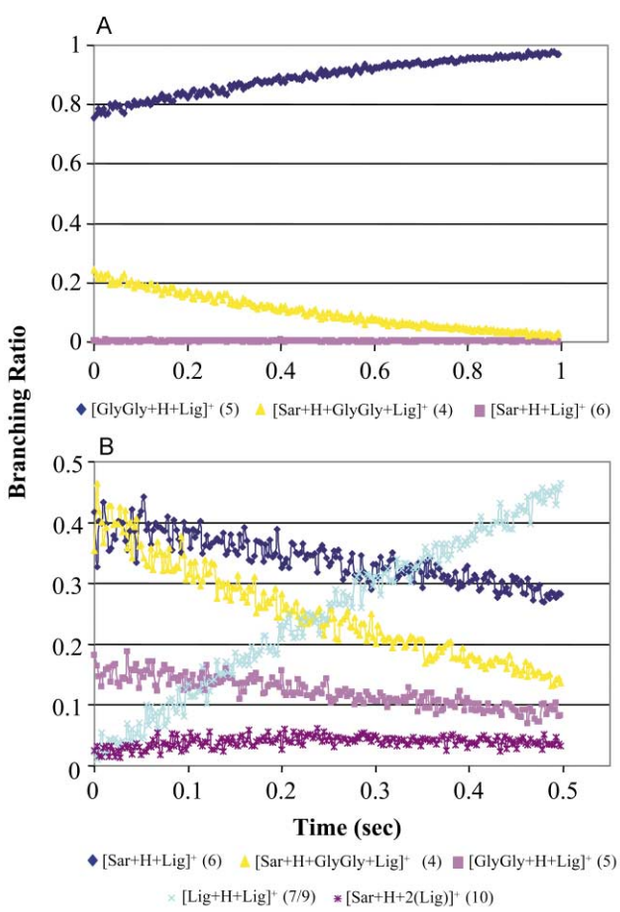


Fig. 1 Normalized branching ratios for the ligand substitution reactions of [Sar+H+GlyGly]$^+$ ions reacting with: (A) butylamine; (B) *N*,*N*-diethylacetamide.

Table 1 Rate constants and branching ratios for reactions of $[Sar+H^{+}+GlyGly]^{+}$ with various neutral ligands (Lig)

Neutral ligand (Lig)	Gas phase basicity/ kJ mol^{-1} [a]	μ/D [b]	Base	k_{exp} [c]	Reaction efficiency [d]	Branching ratios for primary products [e]		
						[Sar+H+Gly Gly+Lig]$^{+}$	[Sar+H+ Lig]$^{+}$	[GlyGly+ H+Lig]$^{+}$
(1) Methanol	724.5	1.70	O	No reaction				
(2) Ethanol	746	1.69	O					
(3) Acetonitrile	748	3.92	N					
(4) 2,3-Butanedione	770.1	1.03	O					
(5) Acetone	782.1	2.88	O					
(6) Methyl acetate	790.7	1.72	O					
(7) 2-Butanone	795.5	2.78	O					
(8) Methyl methacrylate	800.5	1.60	O					
(9) Diethylether	801	1.15	O					
(10) 3-Pentanone	807	2.75	O					
(11) *tert*-Butyl methyl ether	812.4	1.2	O					
(12) 4-Methylcyclohexanone	813	3.07	O					
(13) Hexamethyldisiloxane	816.2	≈ 0	O					
(14) 1,2-Dimethoxyethane	820.2	1.8	O	0.5	0.004	0.46	0.38	0.16
(15) Methylformamide [f]	820.3	3.83	O	1.0	0.005	0.10	0.60	0.30
(16) *tert*-Butyl ethyl ether	826.9	1.2	O	No reaction				
(17) Isopropyl ether	828.1	1.13	O					
(18) 2,4-Pentanedione	836.8	3.03	O					
(19) Pyrrole	843.8	1.74	N					
(20) 2,5-Hexanedione [f]	851.8	3.05	O	3.5	0.022	0.55	0.30	0.15
(21) Dimethyl sulfoxide	853.7	3.96	O	59	0.290	0.70	0.20	0.10
(22) Trimethyl phosphate	860.8	2.99	O	313	2.15	0.40	0.45	0.15
(23) Methylamine	864.5	1.31	N	19	0.136	0.20	0 [g]	0.80
(24) *N*-Ethylacetamide [f]	867	3.87	O	24	0.122	0.65	0.25	0.10
(25) 2-Chloropyridine	869	3.25	N	No reaction				
(26) 3-Chloropyridine	871.5	2.02	N					
(27) *N*,*N*-Dimethylacetamide [f]	877	3.81	O	124	0.632	0.50	0.35	0.15
Glycylglycine	**882**	**2.88**						
(28) *n*-Butylamine	886.6	1.23	N	226	1.69	0.15	0 [g]	0.85
Sarcosine	**888.7**	**1.2**						
(29) *N*,*N*-Diethylacetamide [f]	894.4	3.88	O	278	1.48	0.35	0.50	0.15
(30) Pyridine	898.1	2.22	N	146	1.01	0.06	0.02	0.92
(31) 2-Methoxypyridine	902.8	1.09	N	8.6	0.079	0 [g]	0.05	0.95
(32) Ethylenediamine	912.5	1.83	N	343	2.52	0 [g]	0.02	0.98
(33) 4-Methylpyridine	915.3	2.70	N	351	2.22	0 [g]	0.03	0.97
(34) 3-Ethylpyridine	915.5	2.4	N	367	2.51	0 [g]	0.04	0.96
(35) Diethylamine	919.4	0.92	N	274	2.45	0 [g]	0.01	0.99
(36) 2,6-Dimethylpyridine	931.1	1.78	N	306	2.34	0 [g]	0.06	0.94
(37) Triethylamine	951.0	0.66	N	198	1.88	0 [g]	0.07	0.93

[a] Gas phase basicities are from ref. 24. [b] Dipole moments are either taken from a review[25] or are sourced from the primary literature. Complete references are available from the authors upon request. [c] Rate is in units of 10^{-11}cm^{3} molecule^{-1} sec^{-1}. A conservative estimate of error is ±25%. [d] Reaction efficiency = $k_{(exp)}/k_{ADO}$, where k_{ADO} is the theoretical prediction of the collision rate between the $[Sar+H+GlyGly]^{+}$ ion and a neutral base. These collision rates are calculated by the method of Su and Bowers.[18] [e] Branching ratios are extrapolated to zero reaction time and are estimated to have error limits of ±20%. [f] Branching ratios complicated by secondary reactions including decay of adduct at longer reaction times. [g] Branching ratio is less than 0.01.

(Figure 1A). In contrast, *N*,*N*-diethylacetamide (an oxygen nucleophile/base with a GB = 894.4 kJ mol^{-1}), shows a broader reactivity pattern with both primary and secondary reactions being observed. $[Lig+H+Lig]^{+}$ and $[Sar+H+2Lig]^{+}$ are both secondary reaction products since their branching ratios extrapolate to zero at zero reaction time. They arise from the processes shown in eqs. 7, 9 and 10, as confirmed in MS^{3} experiments in which the primary ions $[GlyGly+H+Lig]^{+}$ and $[Sar+H+Lig]^{+}$ are mass selected and allowed to further react with *N*,*N*-diethylacetamide. More importantly, *N*,*N*-diethylacetamide shows a different selectivity, preferring to switch glycylglycine (Figure 1B). Interestingly, the same contrasting behaviour has been observed by Witt and Grützmacher, who noted that the more polar amides tend to undergo secondary reactions more readily.[12]

(a) Trends in ligand substitution

There are no simple trends observed for the rates of reaction of $[Sar+H+GlyGly]^{+}$ with ligands. Not all ligands are reactive, with the first ligand to show reactivity being the bidentate oxygen nucleophile/base 1,2-dimethoxyethane. The gas phase basicity of this ligand is 820.2 kJ mol^{-1}, over 60 kJ mol^{-1} less than that of either of the original partners of the $[Sar+H+GlyGly]^{+}$ complex. In contrast, the monodentate oxygen ligands *tert*-butyl ether and isopropyl ether, which have higher gas phase basicities, are unreactive. Thus gas phase basicity appears to be less important than how the ligand can interact with either glycylglycine or sarcosine in the substitution products (eqns. (5) and (6) of Scheme 1). This manifests itself even more dramatically in the relative yields for these ligand substitution reactions (eqns. (5) and (6)), which are plotted in Figure 2 as a function of gas phase basicity. In all cases, nitrogen nucleophiles/bases (amines and pyridines) prefer to switch out sarcosine (eqn. (5) of Scheme 1). In contrast all oxygen nucleophiles/bases (ketones, ethers, sulfoxides, phosphates and amides) prefer to switch out glycylglycine (eqn. (5) of Scheme 1). This selectivity is similar to that observed by Witt and Grützmacher (Scheme 1), who noted that ligand exchange reactions depend on two main properties of the neutral ligand: (i) its polarity; and (ii) its proton affinity.[12] Thus the more polar oxygen nucleophiles/bases prefer to switch out the more polar component of the proton bound dimer (glycylglycine). These ligand substitution reactions represent an example of gas phase molecular recognition in a proton bound, non-covalent complex of an amino acid and a dipeptide.

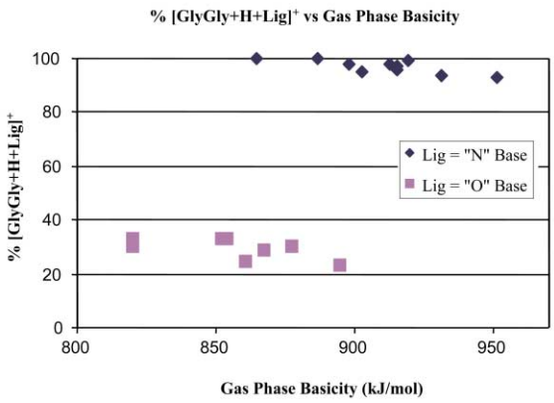


Fig. 2 Plot of the percentage of the $[GlyGly+H+Lig]^+$ product as a function of the gas phase basicity of the ligand. Note that $\%[GlyGly+H+Lig]^+ = 100\% \times \{[GlyGly+H+Lig]^+/([Sar+H+Lig]^+ + [GlyGly+H+Lig]^+)\}$.

Interestingly, the reactivity of $[GlyGly+H]^+$ changes dramatically as it becomes part of the $[Sar+H+GlyGly]^+$ complex. Just as our previous study demonstrated a difference in H/D exchange behaviour of this complex,[13] its mode of reactivity changes for acetone and 2,5-hexanedione. While $[GlyGly+H]^+$ reacts with both acetone and 2,5-hexanedione *via* a Schiff base formation reaction,[19,20] no such reaction is observed for the $[Sar+H+GlyGly]^+$ ion, which instead undergoes no reaction with acetone or ligand substitution with 2,5-hexanedione.

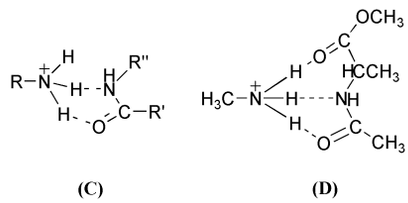
(b) Trends in adduct formation

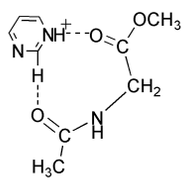
An examination of Table 1 reveals that unlike Lifshitz's study,[11a] there is no simple correlation between $[Sar+H+GlyGly+Lig]^+$ adduct formation and the structure of the ligand (Lig). Thus, while the amines methylamine and *n*-butylamine do form adducts, diethylamine does not. Based on Lifshitz's insertion model,[11] all the pyridines and triethylamine would not be expected to form an adduct, yet adduct formation is observed for pyridine. Adduct formation also appears to be more prevalent for the oxygen atom nucleophiles/bases, being observed in the majority of cases.

It is of considerable interest as to how these trimer ions are related to the overall ligand substitution channels observed. In particular does CID fragmentation of the collisionally stabilized $[Sar+H+GlyGly+Lig]^+$ adduct (formed via eqn. (4) in Scheme 2) yield the same branching ratios as that observed for the "direct" ligand substitution (*i.e.* eqns. (5) and (6) of Scheme 2)? Unfortunately our attempts to answer this question were unsuccessful. Briefly, we tried to examine the CID reactions of $[Sar+H+GlyGly+Lig]^+$ trimer ions in two ways: (i) via mass selection the collisionally stabilized trimers ions (formed via eqn. (4) of Scheme 2); (ii) by mixing the ligand with GlyGly and Sar in solution and examining the ESI/MS/MS spectra of the resultant trimer *via* CID. In both cases, insufficient signal of the trimers coupled with signal loss during mass selection of the trimer resulted in spectra which were both noisy and not reproducible. Note that signal loss during mass selection of weakly bound ions is a common phenomenon in quadrupole ion traps.[21]

(c) Insights from modeling studies

How can these molecular recognition reactions be rationalized in terms of the structures of the $[Sar+H+GlyGly]^+$ complex, the ligand (Lig) and the products $[Sar+H+Lig]^+$ and $[GlyGly+H+Lig]^+$? Mautner has studied the binding energies of a wide range of small proton bound dimers of biological significance.[22] It is interesting to speculate that some of the hydrogen bonding structural motifs that he has described may play an important role in the ligand substitution reactions observed. Thus, the preference for amides to switch out glycylglycine may be explained by Mautner's binding mode for ammonium ions to amides **(C)**, while the preference for amines and pyridines to switch out sarcosine could be rationalized by Mautner's binding modes for ammonium and pyridines binding to *N*-acetylated amino acid esters as shown in structures **(D)** and **(E)**.





While detailed molecular modeling for these systems are beyond the scope of this paper, we decided to focus on the $[Sar+H+GlyGly+Lig]^+$ system (where Lig = methylamine) since: (i) amines exhibit the greatest selectivity and (ii) this amine can potentially form insertion products of the type **(A)**. The molecular protocol is described in the experimental section and initially involves performing molecular dynamics calculations on all possible trimer structures in which one component is protonated (the "core" ion) and the other two are neutral. Thus we have considered protonated sarcosine surrounded by neutral glycylglycine and the ligand (designated as $[(SarH^+)(GlyGly)\ (Lig)]$) along with $[(Sar)(GlyGlyH^+)(Lig)]$ and $[(Sar)(GlyGly)(LigH^+)]$. Following optimization at the PM3 level of theory these trimers were separated into the appropriate dimers and subjected to the same molecular dynamics simulations (and optimizations) described above. The same process was repeated to explore the conformational space of the monomers. It should be noted that the PM3 level of theory was chosen to optimize the selected structures since it is generally accepted to be the best semiempirical method for calculating hydrogen bonded structures.[23]

The molecular dynamics (MD) calculations yielded a host of structures with different numbers and types of hydrogen bonds, consistent with non-covalent dimers (and trimers) in dynamic equilibrium. Since the molecular dynamics cannot be used to determine the overall energetics for the competing reactions (eqns. (4)–(6) of Scheme 2), we have taken the most stable structures from the MD calculations and have reoptimized them at the PM3 level of theory. We do not imply that these are the only unique structures, but have used them to gain insights into the possible types of hydrogen bonding in the trimers and dimers and the relative energies for adduct formation (eqn. (4)) and ligand switching (eqns. (5) and (6)). We recognize that the various protonated forms of the dimers and trimers are likely to be in dynamic equilibrium, but have not examined transition states for processes such as proton transfer within these cluster ions. All the ionic structures are shown in Fig. S1 of the electronic supplementary information (ESI,‡), while the energies of all species relevant to calculating the energetics of ligand substitution can be found in Table S1 (ESI). As expected, the PM3 calculations reveal the formation of several types of hydrogen bonds in the dimer and trimer ions. Thus $[(SarH^+)(GlyGly)]$ contains three hydrogen bonds (of the type $[–N–H^+ \cdots O–]$ and $[–O–H \cdots N–]$) while $[(Sar)(GlyGlyH^+)]$ contains two hydrogen bonds (of the type $[–N–H^+ \cdots O–]$ and $[–O–H \cdots O–]$). Similarly, each of the three possible trimers, *i.e.* $[(SarH^+)(GlyGly)(Lig)]$, $[(Sar)(GlyGlyH^+)(Lig)]$ and

[(Sar)(GlyGly)(LigH$^+$)] contains three hydrogen bonds and the types vary from [–N–H$^+$ ··· O–], [–N–H$^+$ ··· N–], [–O–H ··· N–] to [–O–H ··· O–]. Thus multiple hydrogen bonds play an active role in stabilizing these ion–molecule complexes.

To gain further insight into the relative stabilities of these complexes, we have used the PM3 heats of formation (Table S1, ESI) to calculate the energy diagram for ligand switching reactions of [Sar+H+GlyGly]$^+$ with methylamine (Fig. 3). Note

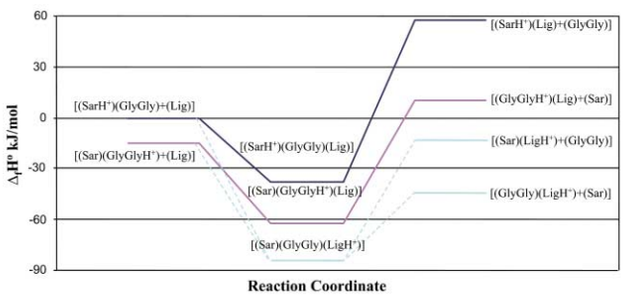


Fig. 3 Energy diagram for ligand switching reactions of [Sar+H+GlyGly]$^+$ with methylamine, using the PM3 optimized structures. All energies are derived from Table S1 (ESI,‡) and are relative to the total energy of the separated reactants ([(SarH)$^+$(GlyGly)] and methylamine), which are at 0 kcal mol^{-1}.

that this is not meant to represent the potential energy surfaces for these reactions, but rather to reveal the relative energies for: (i) formation of the three types of isomeric trimer adduct ions; (ii) ligand substitution. Of the two initial dimer complexes, [(Sar)(GlyGlyH$^+$)] is observed to be energetically favored over its protonated sarcosine counterpart. While this trend continues to hold for their respective trimer structures, the most stable trimer is actually [(Sar)(GlyGly)(LigH$^+$)], in which proton transfer to the ligand has occurred. An examination of the traditional product ions, *i.e.* [(SarH$^+$)(Lig)] (eqn. (6)) and [(GlyGlyH$^+$)(Lig)] (eqn. (5)) reveals both these complexes are endothermic and each possesses only one hydrogen bond (of the type [–N–H$^+$ ··· N–]) between the protonated ion and methylamine. Similar analysis of the proton transfer product ions (in which $MeNH_3^+$ is the core ion), *i.e.* [(Sar)(LigH$^+$)] and [(GlyGly)(LigH$^+$)] reveals that both reactions are exothermic and each contains two hydrogen bonds (of the type [–N–H$^+$ ··· N–] and [–N–H$^+$ ··· O–]). Overall, the most energetically stable product ion is observed to be [(GlyGly)(LigH$^+$)]. These calculations mirror the experimental results, which show that methylamine prefers to switch out sarcosine.

Conclusions

This study represents one of the first examples of molecular recognition in the ion–molecule ligand substitution reactions of a proton bound dimer of simple biomolecules ([Sar+H+GlyGly]$^+$). This dimer undergoes two main types of reaction, namely, association (resulting in adduct formation) and ligand switching. Association appears to be more prevalent for oxygen containing bases and those with a lower gas phase basicity. The ligand switching reactions show a remarkable selectivity in which nitrogen containing nucleophiles/bases prefer to switch out sarcosine, while oxygen containing nucleophiles/bases prefer to switch out glycylglycine. This selectivity is similar to that observed by Witt and Grützmacher (Scheme 1), who noted that the ligand exchange reactions of the proton bound heterodimer of*N*,*N*-dimethylformamide (DMF) and *n*-propyl amine depended on two main properties of the neutral ligand: (i) its polarity; and (ii) its proton affinity. The dimer [Sar+H+GlyGly]$^+$) appears to exhibit broader reactivity (e.g. Sar is switched out by the *less* basic amine, methylamine). Thus additional factors such as hydrogen bonding in the reactant and product ions may also play a role in the selectivity observed.

Acknowledgements

R.A.J.O. thanks the Australian Research Council for financial support (grant #A29930202) and the University of Melbourne for funds to purchase the LCQ. A.K.V acknowledges the award of a Science Faculty Scholarship (Studentship). We thank Dr Herbert Treutlein for help running and access to the Insight II program. We thank Dr Gavin Reid for carrying out some preliminary experiments.

References

1 (*a*) R. D. Smith, J. E. Bruce, Q. Y. Wu and Q. P. Lei, *Chem. Soc. Rev.*, 1997, **26**, 191; (*b*) J. M. Daniel, S. D. Friess, S. Rajagopalan, S. Wendt and R. Zenobi, *Int. J. Mass Spectrom.*, 2002, **216**, 1; (*c*) J. L. Beck, M. L. Colgrave, S. F. Ralph and M. M. Sheil, *Mass Spectrom. Rev.*, 2001, **20**, 61.

2 (*a*) C. A. Schalley, *Mass Spectrom. Rev.*, 2001, **20**, 253; (*b*) C. A. Schalley, *Int. J. Mass Spectrom.*, 2000, **194**, 11.

3 For a review of Cooks' kinetic method, which has been widely used to determine thermochemistry see: R. G. Cooks, J. S. Patrick, T. Kotiaho and S. A. Mcluckey, *Mass Spec. Rev.*, 1994, **13**, 287.

4 (*a*) W. A. Tao, L. M. Wu and R. G. Cooks, *Chem. Commun.*, 2000, 2023; (*b*) W. A. Tao, D. X. Zhang, E. N. Nikolaev and R. G. Cooks, *J. Am. Chem. Soc.*, 2000, **122**, 10598; (*c*) W. A. Tao and R. G. Cooks, *Angew. Chem., Int. Ed. Engl.*, 2001, **40**, 757.

5 (*a*) E. F. Strittmatter, P. D. Schnier, J. S. Klassen and E. R. Williams, *J. Am. Soc. Mass Spectrom.*, 1999, **10**, 1095; (*b*) P. D. Schnier, J. S. Klassen, E. F. Strittmatter and E. R. Williams, *J. Am. Chem. Soc.*, 1998, **120**, 9605.

6 Q. Wu, J. Gao, D. Joseph-McCarthy, G.B. Sigal, J. E. Bruce, G.M. Whitesides and R. D. Smith, *J. Am. Chem. Soc.*, 1997, **119**, 1157.

7 C. B. Lebrilla, *Acc. Chem. Res.*, 2001, **34**, 653.

8 S. Gronert and R. A. J. O'Hair, *J. Am. Soc. Mass Spectrom.*, 2002, **13**, 1088.

9 A. E. Counterman and D. E. Clemmer, *J. Phys. Chem. B*, 2001, **105**, 8092.

10 E. Nir, K. Kleinermanns and M. S. de Vries, *Nature*, 2000, **408**, 949.

11 (*a*) W. Y. Feng, Y. Ling and C. Lifshitz, *J. Phys. Chem.*, 1996, **100**, 35; (*b*) C. Zhu, Y. Ling, W. Y. Feng and C. Lifshitz, *Int. J. Mass Spectrom.*, 2000, **194**, 93; (*c*) J. M. L. Martin, V. Aviyente and C. Lifshitz, *J. Phys. Chem. A*, 1997, **101**, 2597.

12 M. Witt and H.-F. Grützmacher, *Int. J. Mass Spectrom.*, 2003, **222**, 27.

13 G. E. Reid, R. A. J. O'Hair, M. L. Styles, W. D. McFadyen and R. J. Simpson, *Rapid Commun. Mass Spectrom.*, 1998, **12**, 1701.

14 (*a*) Z. P. Yao, T. S. M. Wan, K. P. Kwong and C. T. Che, *Anal. Chem.*, 2000, **72**, 5383; (*b*) G. Fago, A. Filippi, A. Giardini, A. Lagana, A. Paladini and M. Speranza, *Angew. Chem., Int. Ed. Engl.*, 2001, **40**, 4051; (*c*) T. J. D. Jorgensen, D. Delforge, J. Remacle, G. Bojesen and P. Roepstorff, *Int. J. Mass Spectrom. Ion Processes*, 1997, **167/168**, 135.

15 For excellent reviews on the role of ion–molecule complexes see: (*a*) R. D. Bowen, *Acc. Chem. Res.*, 1991, **24**, 364; (*b*) T. H. Morton, *Tetrahedron*, 1982, **38**, 3195.

16 A. K. Vrkic and R. A. J. O'Hair, *Int. J. Mass Spectrom.*, 2002, **218**, 131.

17 (*a*) S. Gronert, L. M. Pratt and S. Mogali, *J. Am. Chem. Soc.*, 2001, **123**, 3081; (*b*) S. Gronert, *J. Am. Soc. Mass Spectrom.*, 1998, **9**, 845.

18 T. Su, M. T. Bowers, *Gas Phase Ion Chemistry*, Academic, New York, 1979, vol. 1, pp. 83–118.

19 R. A. J. O'Hair and G. E. Reid, *J. Am. Soc. Mass Spectrom.*, 2000, **11**, 244.

20 E. H. Gur, L. J. de Koning and N. M. M. Nibbering, *Int. J. Mass Spectrom. Ion Processes*, 1997, **167/168**, 135.

21 J. E. McClellan, J. P. Murphy, J. J. Mulholland and R. A. Yost, *Anal. Chem.*, 2002, **74**, 402.

22 (*a*) M. M-N. Mautner, *J. Am. Chem. Soc.*, 1984, **106**, 278; (*b*) M. M-N. Mautner, *Acc. Chem. Res.*, 1984, **17**, 186; (*c*) M. M-N. Mautner, *Molecular Structure and Energetics*, VCH Publishers, Inc., 1987, vol 4, ch. 3, p. 71.

23 (*a*) B. Kallies and R. Mitzner, *J. Mol. Model.*, 1995, **1**, 68; (*b*) E.V. Denisov, V. Shustryakov, E.N. Nikolaev, F.J. Winkler and R. Medina, *Int. J. Mass Spectrom. Ion Processes*, 1997, **167/168**, 259.

24 *NIST Standard Reference Database 69*, National Institute of Standards and Technology, Gaithersburg, MD, July 2001, http://webbook.nist.gov/chemistry/CRC Handbook of Chemistry and Physics.

25 *CRC Handbook of Chemistry and Physics*, ed. R. C. Weast, Chemical Rubber Company, Boca Raton, FL, 69th edn., 1988–1989.