The Hong Kong Polytechnic University

Theoretical Studies on Potassium Cation–Ligand Interaction

Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

by

Wong Hoi Shan

18th August 2004

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text. The experimental results quoted in this thesis were carried out by Dr. Y. Tsang and Dr. C.L. Wong of our research group. Parts of the theoretical calculations presented in Chapter 3 of the thesis were carried out by Dr. F.M. Siu, Mr. J.K.C. Lau, and Ms. P.S. Ng of our research group. This thesis has not been submitted to this or any other institution for academic qualification purposes.

Wong Hoi Shan

Date: 18th August, 2004

Abstract of thesis entitled

"Theoretical Studies on Potassium Cation-Ligand Interaction"

submitted by Wong Hoi Shan

for the degree of Doctor of Philosophy in Chemistry

at The Hong Kong Polytechnic University in August, 2004.

Abstract

Potassium cation is one of the most abundant metal ions in biological systems, where it is involved in numerous biochemical functions, such as stabilization of protein structures and osmotic equilibrium of cells. For a better understanding of such interactions, information about the *intrinsic* binding modes and energies of K^+ to simple model systems ligands is essential.

The present study addresses this subject by determining the geometries and energetics of K^+ binding to: (a) small organic ligands, (b) aliphatic amino acids, (c) linear dipeptides (glycylglycine, **GG**, and alanylalanine, **AA**), (d) dipeptides containing phenylalanine and (e) the basic amino acid, histidine, by quantum chemical molecular modeling methods. The theoretical studies were carried out by high-level density functional theory calculations at the B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d)) level of theory (abbreviated as the '$EP(K^+)$ protocol'). Our study shows that theoretical protocol yields absolute affinities for 65 model organic ligands in excellent agreement with experimental affinities (mean absolute deviation of 4.5 kJ mol^{-1}), in support of the reliability of the B3-LYP protocol adopted in our study.

Using this EP(K^+) protocol, ten stable isomers of K^+-glycine on the potential energy surface have been located and the most stable mode of K^+ binding involves a bidentate interaction between the cation and the O=C and –OH sites in the charge-solvated (CS) form. We also found that the stabilization energies of these complexes can be well approximated by a linear function of the 'dipole interaction parameter (DIP)' and 'polarizability interaction parameter (PIP)'. For the larger aliphatic amino acids, we found that the most stable modes of K^+ binding is still the same as K^+-glycine, even though the zwitterionic (ZW) form is stabilized more than the CS mode. The effect of alkyl chain length of the relative K^+ affinities of aliphatic amino acids are also discussed.

The most stable K^+-GG/AA/FG/GF complexes are found to be in the charge-solvated (CS) form, with K^+ bound to the two carbonyl oxygen atoms of the peptide backbone. For the **FG/GF** dipeptides, K^+ additionally bound to the phenyl π-ring of phenylalanine are comparably stable. The K^+ is found to be in close alignment with the molecular dipole moment vector of the bound ligand, indicating that electrostatic ion-dipole interaction is the key stabilizing factor in these complexes. Furthermore, the strong ion-dipole interaction between K^+ and the amide carbonyl oxygen of the peptide bond is important in determining the relative stabilities of different CS binding modes. When compared to amino acids, the most stable ZW complex is found to be much less stable than the CS form for dipeptides. The usefulness of proton affinity as a criterion for estimating the relative stability of ZW versus CS binding modes is examined. Based on the results of this study, the interaction of K^+ with longer peptides consisting of aliphatic amino acids is also rationalized.

Theoretical studies on protonated and potassiated histidine ($[His + H/K]^+$) and their mass spectrometric fragmentations in the gas phase have been conducted. Fragmentation of $[His + H]^+$ occurs when the proton attaching to N^{π} is transferred to the hydroxyl oxygen (yielding $[CO + H_2O]$ or H_2O), or when the backbone amino nitrogen is protonated (yielding $[CO_2 + NH_3]$). The critical energy required for the formation of $[CO + H_2O]$ is found to be much lower than that for $[CO_2 + NH_3]$, in agreement with the dominant loss of $[CO + H_2O]$ observed in the low-energy (eV, laboratory scale) collision-induced dissociation (CID) mass spectrum. However, the fragmentation pattern of potassiated histidine is entirely different. While the dominant fragment in $[His + H]^+$ corresponds to the loss of $[CO + H_2O]$, such fragment is absent in the $[His + K]^+$ CID mass spectrum. Instead, loss of K^+ is dominant in potassiated histidine, with minor fragment ions arising from the losses of CO_2 and NH_3 separately. The difference in fragmentation behavior between protonated and potassiated histidine are discussed.

Acknowledgement

I would like to thank my co-supervisor Dr. N. L. Ma, the Institute of High Performance Computing and National University of Singapore, who helped me to understand thoroughly on the theoretical calculations of my study, and her generous guidance for providing critical comments on this thesis and related publications.

I would like to express my deepest gratitude to my chief supervisor, Prof. C. W. Tsang, for his generous guidance, supervision and encouragement throughout the conduct of this project, especially his valuable comments on the manuscript of this thesis and related publications.

The award of a Hong Kong Polytechnic University graduate studentship to me, and the funding support of the Research Grant Council of Hong Kong (CERG Project No. GW-021) are gratefully acknowledged.

I would also like to thank all of my research group members, who have helped me in various ways during the years of my study. In particular, I would like to thank Dr. Y. Tsang of our research group who kindly provided me the experimental results, Miss H. M. Lee and Dr. F. M. Siu who often encouraged me and helped me when I encountered difficulties during my project, and also Mr. K. F. Tam, Mr. K. C. Lau and Mr. Y. K. Au who kindly helped me a lot on many technical problems. I also wish to express my gratitude to all academic and technical staff from the Department of Applied Biology and Chemical Technology for their support.

Finally, I would like to thank all members of my family for their continuous encouragement during the course of my postgraduate study.

Table of Contents

List of Figures

for the loss of HCOOH), postulated by [Rogalewicz and Hoppilliard, 2000b]. The relative enthalpies at 0K (ΔH_0), with respect to species 1, are shown in parentheses, square bracket and large bracket for loss of $[CO_2 + NH_3]$, $[CO_2 + H_2]$, and HCOOH, respectively, in kJ mol^{-1}.

Figure S-7.2. The loss of $C(OH)_2$ (46 u) from $[His + H]^+$, initiated by proton transfer from imidazole N^{π} to carboxyl O_C). The relative enthalpies at 0K (ΔH_0), with respect to species 1, are shown in parentheses (in kJ mol^{-1}).

Figure S-7.3. The loss of $[CO + H_2O]$ from $[His + H]^+$, via the mechanism analogous to that for protonated aliphatic amino acids (i.e. initiated by proton transfer from amino N^{α} to hydroxyl O_H), postulated by [Rogalewicz and Hoppilliard, 2000b]. The relative enthalpies at 0K (ΔH_0), with respect to species 1, are shown in parentheses (kJ mol^{-1}).

List of Tables

List of Schemes

protonated histidine. Pathway (a) corresponds to the equivalent mechanism proposed in ref. [Rogalewicz et al., 2000c] for $[His + H]^+$; pathway (b) corresponds to the equivalent mechanism reported in ref. [Lioe et al., 2004] for protonated tryptophan; pathway (c) and corresponds to the equivalent mechanism reported in ref. [Aribi et al., 2004] for protonated phenylalanine.

Scheme 7.3. Scrambling that initiated by the ring N^{τ} exchanges its deuterium with the hydrogen on the imidazole C_2 in protonated histidine, followed by the formation of carbon dioxide and ammonia mechanism presented in Fig. 7.7 (1 → 1-12 → 12 etc.) that the ND_3 exchanges the deuterium with N^{τ} hydrogen.

Scheme 7.4. Proposed fragmentation mechanism for the loss of K^+ from $[His + K]^+$

Scheme 7.5. Proposed pathway and mechanism for the loss of NH_3 from $[His + K]^+$, analogous to that for $[Ala + K]^+$ [Abirami et al., 2004].

Scheme 7.6. Proposed pathway and mechanism for the loss of CO_2 from $[His + K]^+$, analogous to that for free glycine ([Li and Brill, 2003], Scheme 2).

Scheme 7.7. Proposed pathway and mechanism for the loss of H_2O from $[His + K]^+$, analogous to that for $[Ala + K]^+$ [Abirami et al., 2004].

Scheme 7.8. Proposed pathway and fragmentation mechanism for the loss of H_2O from $[His + K]^+$, analogous to that for $[His + H]^+$.

Scheme 7.9. Resonance form of (a) species 11 in $[His + H]^+$ (b) the analogous species in $[His + K]^+$, identical to species KH8 in Scheme 7.8.

Scheme 7.10. Key structures for the loss of $[CO_2 + NH_3]$ in $[His + K]^+$. Species KH9 corresponds to the analogous species of 14 reported in $[His + H]^+$, stabilized by hydrogen bond $N^{\tau}H...NH_3$; species KH10 is the alternative structure of species 14, which stabilized by electrostatic interaction of $K^+...NH_3$.

List of Abbreviations

B3LYP	Beck's three parameters exchange function (B3-) and the correlation functional Lee, Yang and Parr (-LYP)
BSSE	Basis Set Superposition Error
CGTOs	Contracted gaussian type orbitals
CID	Collisional induced dissociation
CS	Charge-solvated form
DFT	Density Functional Theory
DIP	Dipole interaction parameter
E_{def}	Deformation energy
$E_{destabilization}$	Destabilization energy
E_{elec}	Electronic energy
$EP(K^+)$	Energetic protocol for K^+
$E_{stabilization}$	Stabilization energy
FT-ICR	Fourier transform ion cyclotron resonance
GB	Gas phase basicity
GCA	Gradient Corrected Approximation
GTO	Gaussian-type orbital
HF	Hartree-Fock
HPMS	High pressure mass spectrometric
LDA	Local Density Approximation
MAD	Mean absolute deviation

MCMM	Monte Carlo Multiple Minimum
M O	Molecular Orbital
MP2	Second-order for Møller-Plesset
MS	Mass spectrometry
PA	Proton affinities
PCA	Potassium cation affinity
PES	Potential energy surface
PIP	Polarization interaction parameter
PGTOs	Primitive gaussian type orbitals
RF	Radiofrequency
SCF	Self-Consistent Field
STO	Slater type orbitals
ZPE	Zero-point energy
ZW	Zwitterionic form
ΔH_0	Binding affinity at 0K/ Enthalpy change at 0K
ΔH_{298}	Binding affinity at 298K
ΔG_{298}	Gibbs free energy of binding at 298K or gas basicity at 298K
ΔG_{600}	Gibbs free energy of binding at 600K or gas basicity at 600K
ΔS	Entropy change
3-W	Three–dimensional Wiener index
$3\text{-}W_{\text{side chain}}$	Side chain analogue of the three-dimensional Wiener index

Abbreviations of Compounds

Me = $-CH_3$

Et = $-C_2H_5$

i-Bu = $-(CH_3)_2CHCH_2$

i-Pr = $-(CH_3)_2CH$

n-Bu= $-C_4H_{11}$

n-Pr = $-C_3H_8$

Ph= $-C_6H_5$

t-Bu= $-(CH_3)_3C$

RNAs/DNAs, Amino acids and peptides

A	adenine
AA	alanylalanine
Ala	alanine
Arg	arginine
C	cytosine
FG	phenylalanylglycine
G	guanine
GF	glycylphenylalanine
GG	glycylglycine
Gly	glycine

His	histidine
Ile	isoleucine
Leu	leucine
Phe	phenylalanine
T	thymine
U	uracil
Val	valine

Chapter 1 Introduction

1.1 The Potassium Cation: Its Biological and Analytical Importance

Potassium cation, K^+, is one of the most abundant metal cations found in biological systems. Studies have shown that K^+ can induce conformational changes when it binds to proteins. [Kohtani et al., 2000; Taraszka et al., 2001] As a result, the interaction between potassium cation and proteins/carbohydrates structures underlies many fundamental biological processes and enzyme functions [Hughes, 1972; Lippard and Berg, 1994; Stryer, 1995], including maintaining the osmotic equilibrium of cells, [Lippard and Berg, 1994; Stryer, 1995] stabilization of protein structures, [Lippard and Berg, 1994] activation of enzyme functions, [Hughes, 1972] signaling in the nervous system, generation of rhythmic signals by the heart, selective transport of K^+ across cell membranes [Jiang et al., 2003; Zhou and MacKinnon, 2003] and unceasing sifting of toxic solutes in the blood by the kidney. [Miller, 1993; Kaim and Schwederski, 1994; Hille, 2001] Thus, knowledge of the K^+ binding modes and *intrinsic* binding energies (affinities) of smaller model ligand/amino acid systems are basic to a full understanding of the interaction of K^+ in the more complex and larger peptide/protein system.

Furthermore, understanding the interaction between K^+ (and other alkali cations) and model amino acids/peptides is important in the interpretation of the mass spectra of M^+-peptide/protein complexes, from which sequencing information of peptides can be obtained. [Teesch et al., 1991a; 1991b; Loo and Muenster, 1999; Hoyes and Gaskell, 2001; Huang et al., 2002b]

A related issue that has attracted a lot of attention is the stability of the neutral versus zwitterionic (ZW) form of amino acids/peptides in the gas phase. In aqueous media, all common α-amino acids are found in the zwitterionic form, with a deprotonated carboxyl and a protonated amine (or basic side chain) termini. [Zubay, 1988] In contrast, gas-phase α-amino acids are believed to exist in the 'neutral' free acid form because of the absence of stabilizing intermolecular interactions with solvent molecules. Williams et al. has suggested that arginine, the most basic amino acid, could be zwitterionic in the gas phase. [Williams et al., 1997] However, later experimental study by Chapo et al. [Chapo et al., 1998] suggested that it should occur in the 'neutral' form; i.e., the carboxyl group is not sufficiently acidic to drive spontaneous intramolecular deprotonation even if the fairly basic guanidine group is present. [Hunter and Lias, 1998]

However, the stability of zwitterionic amino acids can be enhanced by complexation with alkali metal ions. For example, the energy difference between zwitterionic (ZW) and charge-solvated (CS) Na^+-glycine is estimated to be only 8-12 kJ mol^{-1}, primarily because of the formation of a salt bridge between the carboxylate terminus and the metal cation. [Jensen, 1992; Hoyau and Ohanessian, 1998; Wyttenbach et al., 1998; 1999a] Based on theoretical calculations, the K^+-arginine complex (and larger alkali cations like Cs^+ and Rb^+) is found to be zwitterionic; the enhanced stability of the zwitterionic form is proposed to have important implication for the observed mass spectrometric fragmentation pattern by facilitating the loss of NH_3 from K^+-arginine. [Jockusch et al., 1999]

1.2 Quantifying the Interaction between Potassium Cation and Ligands

A parameter describing the interaction between K^+ and a ligand (L) is the binding energy/affinity, defined as the enthalpy change (ΔH) for the reaction [1.1]:

$$K^+ + L \rightarrow K^+\text{-}L \qquad [1.1]$$

and can be obtained from Eqn. [1.2] as:

$$\Delta H_0 = [(E_{K^+} + E_L) - E_{K^+\text{-}L}] \qquad [1.2]$$

- where E_{K^+}, E_L and $E_{K^+\text{-}L}$ are the electronic energies of the potassium cation, the ligand and the K^+-ligand complex, respectively.

Affinity is clearly an important parameter as it determines the stability, structure, and reactivity of a K^+-ligand complex. Despite its importance, experimental determination of this parameter *in vivo* is difficult. In a typical biological environment, the stability of different forms of complexes may be modified by the presence of solvent molecules. Thus, a variety of techniques has been applied to obtain the *intrinsic* affinity between the potassium cation and ligands, in the absence of solvent. On the one hand, *absolute* affinities were obtained experimentally by threshold collisional induced dissociation (threshold-CID), [Klassen et al., 1996; More et al., 1997; Walter et al., 1998; Rodgers and Armentrout, 1999; Amunugama and Rodgers, 2000; Rodgers and Armentrout, 2000a; 2000b; Amicangelo and Armentrout, 2000; Armentrout and Rodgers, 2000; Rodgers, 2001; Huang and Rodgers, 2002; Amunugama and Rodgers, 2002] radiative associative kinetics measurements, [Ryzhov and Dunbar, 1999] and high pressure mass spectrometric

(HPMS) equilibrium measurements, [Dzidic and Kebarle, 1970; Davidson and Kebarle, 1976a,b,c; Castleman, 1978a; Sunner et al., 1981; Sunner and Kebarle, 1984; Keesee and Castleman, 1986; Guo and Castleman, 1991; Kebarle, 1997; Hoyau et al., 1999; Ryzhov and Dunbar, 1999; Amunugama and Rodgers, 2002] while *relative* affinities have been obtained by Fourier transform ion cyclotron resonance (FT-ICR) ligand exchange equilibrium measurements, [Alcami et al., 1990; Burk et al., 2000; McMahon and Ohanessian, 2000; Gapeev and Dunbar, 2001] and the mass spectrometric kinetic method. [Cerda and Wesdemiotis, 1996; Cooks and Wong, 1998; Cooks et al., 1999; Ryzhov et al., 2000]

Complementing the progress made in experimental measurements, quantum chemical methods have advanced to a stage that both *relative* and *absolute* alkali cation affinities can be obtained in excellent agreement (with an accuracy of ± 15 kJ mol^{-1}) with experimental values. [Hoyau et al., 1999; Armentrout and Rodgers, 2000; Rodgers and Armentrout, 2000a; Burk et al., 2000; McMahon and Ohanessian, 2000; Ma et al., 2000; Siu et al., 2001a] In many instances, reliable theoretical results have been shown to provide a complementary/alternative route for obtaining and confirming alkali cation affinities. [Siu et al., 1998; Hoyau and Ohanessian, 1998; Russo et al., 2001a; 2001b] Also, theoretical findings on the most stable and low-lying binding modes/structures often provide new insights in the interpretation of experimental data. For example, the recent theoretical study on Li^+, Na^+ and K^+ affinities of DNA/RNA nucleobases highlighted the problem of assigning correct metal cation binding structures to measured alkali metal affinities when the free ligand exhibits tautomerism. [Russo et al., 2001a; 2001b]

While there exists good compilations of the *intrinsic* interaction energies between Li^+/Na^+ with model organic ligands, [Hoyau et al., 1999; Alcami et al. 1990; Burk et al., 2000; McMahon and Ohanessian, 2000; Armentrout and Rodgers, 2000] much fewer potassium cation affinities (both experimentally and theoretically) have been reported in the literature.

1.3 Factors Affecting Potassium Cation Affinities (PCAs)

The binding affinity of potassiated ligand complexes (ΔH_0, Eqn. [1.2]) is a result of the difference between stabilizing ($E_{stabilization}$) and destabilizing ($E_{destabilization}$) interactions, and is given by:

$$\Delta H_0 = E_{stabilization} - E_{destabilization} \qquad [1.3]$$

The stabilizing forces between potassium cation and ligands are believed to be non-covalent in nature. Non-covalent interactions are significant in many areas of science and technological applications, including nanoelectronics, biochemical sensors, and the design of new drugs. [Stenesh, 1998] Comparing to covalent interactions, non-covalent interactions are relatively weak. In the case of alkali metal cations and model organic ligands, the binding affinity is in the range of 20 – 300 kJ mol^{-1}, [Burk et al., 2000; Petrie, 2001; Russo et al., 2001b] which is less than half of the strength of a carbon-carbon single bond. The non-covalent interaction between K^+ and ligand is largely electrostatic in nature, arising from ion-dipole, ion-induced dipole, and in some cases, cation-π interaction. [Armentrout, 2001]

1.3.1 Ion-Dipole Interactions ($E_{ion\text{-}dipole}$)

For molecules, even though electrically neutral, their internal electronic charge distribution generally is not uniform. This separation of charges within a molecule gives rise to the permanent molecular dipole. The interaction energy ($E_{ion\text{-}dipole}$) between an ion and a molecular dipole moment, μ, of the ligand is given by Eqn. [1.4] based on the Coulomb's Law: [Isracelachvili, 1992]

$$E_{ion-dipole} = \frac{-Q\mu\cos\Phi}{4\pi\varepsilon_0 r^2} \qquad [1.4]$$

where $E_{ion\text{-}dipole}$ = ion-dipole interaction (binding) energy (in Joules, J)

Q = electric charge of the ion (in Coulomb, C)

μ = dipole moment of the ligand (in Coulomb-meter, C-m)

Φ = angle of deviation between the ion from the dipole moment vector (in degrees, °)

r = distance between ion and the center of the dipole axis (in meter, m)

ε_0 = dielectric permittivity of free space (8.854×10^{-12} $C^2J^{-1}m^{-1}$)

Thus, the strength of ion-dipole interactions is strongest when the metal ion is in perfect alignment ($\Phi = 0°$) with the dipole moment vector and has been suggested to be one of the major contributions to electrostatic binding in metal cationized-ligand complexes. [Davidson and Kebarle, 1976a]

1.3.2 Ion-induced Dipole Interactions ($E_{ion\text{-}induced\ dipole}$)

The presence of a point charge around a molecule will induce a dipole moment in the molecule. In a uniform field, **E**, this induced dipole moment, μ, can be written as

$$\alpha = \mu / \mathbf{E} \qquad [1.5]$$

in which α is the molecular polarizability. Thus, polarizability of a molecule indicates the response of its electrons to an externally applied electric field. A metal cation can induce an electric dipole moment in a neighboring ligand, no matter the molecule is polar or not. The classical interaction energies between the charge and the center of polarizability is given by Eqn. [1.6]: [Isracelacvili, 1992]

$$E_{ion\text{-}induced\ dipole} = \frac{-Q^2\alpha}{2(4\pi\varepsilon_0)^2 r^4} \qquad [1.6]$$

where $E_{ion\text{-}induced\ dipole}$ = ion-induced dipole interaction energy (in Joules, J)

Q = electric charge of the ion (in Coulombs, C)

α = polarizability of the ligand (in Coulombs-meter, C-m)

r = distance between the interacting ion and ligand (in meter, m)

ε_0 = dielectric permittivity of free space (8.854×10^{-12} $C^2J^{-1}m^{-1}$)

Note that no angle Φ appears since the induced dipole is taken to be parallel to the field. As shown in Eqn. [1.6], the ion-induced dipole interaction energy depends on the polarizability, α, of the ligand, and is inversely proportional to r^4. Thus, compared to ion-dipole interaction, the relative importance of ion-induced dipole is expected to diminish more rapidly, with increasing cation-ligand distance.

1.3.3 Dipole Interaction Parameter (DIP) and Polarization Interaction Parameter (PIP)

To aid our discussion on cation-ligand interaction, and based on the classical electrostatic theory of Eqns. [1.4] and [1.6], [Isracelachvili, 1992] we define two properties of metal cation-ligand interaction, the 'dipole interaction parameter (DIP, Eqn. [1.7])' and 'polarization interaction parameter (PIP, Eqn. [1.8])' here:

$$DIP = \mu * \cos(\Phi) / r_{\mu}^{2} \qquad [1.7]$$

$$PIP = \alpha / r_{\alpha}^{4} \qquad [1.8]$$

where μ = permanent molecular dipole moment of the deformed ligand in the complexed state (from Mulliken population analysis, in Debye)

Φ = angles of deviation between the K^+ and the molecular dipole moment vector (in degrees)

r_{μ} = the distance between the K^+ to the center of the dipole (in Å)

Thus, DIP is in the unit of Debye $Å^{-2}$.

α = molecular polarizability of the deformed ligand (from Mulliken population analysis, in $Å^3$)

r_{α} = the distance between the K^+ to the center of charge (the origin of the molecule) of the deformed ligand (in Å)

Thus, PIP is in the unit of $Å^{-1}$.

Graphical representation of these two parameters are depicted in Fig. 1.1, using K^+-glycine complex as the example:

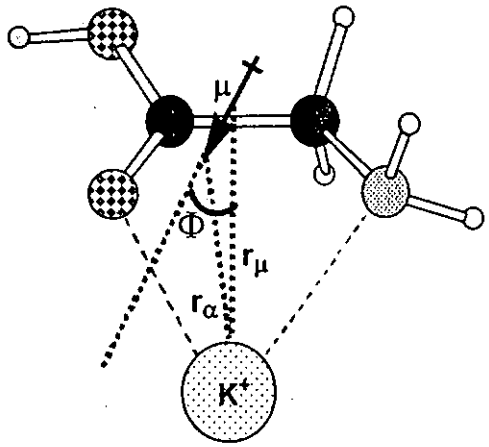
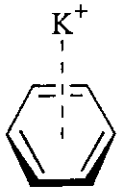


Figure 1.1. Definition of parameters used in the calculation of DIP (Eqn. [1.7]) and PIP (Eqn. [1.8]). Here, μ is the permanent molecular dipole moment of the deformed ligand; Φ is the angle of deviation between the K^+ and the dipole moment vector; r_μ is the distance between the K^+ to the center of the dipole moment vector; and r_α (in Å) is the distance between the K^+ to the center of charge for the deformed ligand.

These electrostatic properties of the deformed ligand (DIP and PIP) take into account of the orientation of the molecular dipole moment and polarizability with respect to the K^+ in the deformed ligand via the Φ, r_μ, and r_α terms, and are expected to provide a better correlation with the stabilization energies than μ and α. (The idea of ligand deformation will be discussed further in Section 1.3.5)

1.3.4 Cation-π Interaction

Cation-π interaction is a relatively new type of non-covalent interaction, involving the electrostatic attraction between a cation and the negative electrostatic potential associated with the π face of a simple aromatic system.



Scheme 1.1. Cation-π interaction.

Various theoretical studies of the interaction between alkali, [Caldwell and Kollman, 1995; Cubero, 1998; Amicangelo and Armentrout, 2000; Feller et al., 2000a; Ikuta, 2000; Tsuzuki et al., 2001; Dunbar, 2002b; Kim et al., 2003b; Lau et al., 2003] alkaline earth, [Zhu et al., 2003] transition metal, [Dunbar, 2002b; Caraiman et al., 2004] and ammonium cation [Zhu et al., 2000; Kim et al., 2003b] and the simple aromatic system benzene and its derivatives have been conducted. These studies suggested that the cation-π interaction is essentially electrostatic in nature, [Ma and Dougherty, 1997; Cubero, 1998; Tsuzuki et al., 2001; Kim et al., 2003b] with induction force playing an important role. [Cubero, 1998; Tsuzuki et al., 2001; Kim et al., 2003b] For K^+, it has been found that dispersion forces are also significant. [Kim et al., 2003] In the context of potassium cation and biological ligands, K^+-π interaction can occur between the metal cation and the aromatic side chains of phenylalanine/tyrosine/tryptophan on the peptide/protein chain, [Hirota et al., 2000; De Wall et al., 2000; Shoeib et al., 2002; Hu et al., 2002a; 2002b; Williams et al., 2003; Meyer, 2003] and the DNA/RNA nucleobases [Wintjens et al., 2000; Mao et al., 2004].

1.3.5 Deformation Energies (E_{def})

In order to accommodate the K^+ at the binding sites, the preferred intramolecular hydrogen bonding pattern of the free amino acids/peptides might be disturbed. Furthermore, comparing to the free ligand, intramolecular electrostatic repulsion between electron-rich functional groups are expected to increase. To understand how metal-cation binding affects the stability of the *free* ligand, we define and calculated the deformation energy, E_{def}, where E_{def} is given by Eqn. [1.9]:

$$E_{def} = E(\text{L in the } K^+\text{-L complex}) - E(\text{L in the } \textit{free} \text{ uncomplexed form}) \qquad [1.9]$$

where E is the electronic energy of the ligand, L, in the two different forms. As the deformed ligand in the complexed form is always less stable than the free ligand, E_{def} is always positive.

Rearranging Eqn. [1.3], the raw interaction energy, $E_{stabilization}$, which only accounts for the total favorable (stabilizing) interaction energy of the metal-ligand interaction, is then given by:

$$E_{stabilization} = \Delta H_0 + E_{def} \qquad [1.10]$$

1.4 Research Objectives

The objective of this research work is to apply theoretical modeling to:

(a) quantify the interaction between K^+ and model ligands, i.e., obtain the relative and absolute potassium cation affinities;

(b) understand the relative importance of various factors affecting the K^+ affinities;

(c) elucidate the relation between strength/preferred mode of binding and the mass spectrometric fragmentation of H^+/K^+-histidine.

The objectives are achieved by a series of studies as described in various chapters below:

Chapter 3: Benchmarking theoretical K^+ affinity (calculated with high level *ab initio* and density functional theory (DFT) theoretical protocols) against the binding affinities obtained experimentally, and constructing quantitative-structure-parameter-relations between binding affinity and ligand properties.

Chapter 4: Using the validated theoretical protocol obtained in Chapter 3 to study the alkyl substituent effect on the K^+ binding affinities of aliphatic amino acids (**Gly, Ala, Val, Leu** and **Ile**). The determinant factors governing the relative stabilities of different CS and ZW modes of binding of K^+-aliphatic amino acids were also investigated.

Chapter 5: Extending the theoretical study, i.e., the determination of absolute and relative affinities of different binding mode, etc., from K^+-Gly/Ala to K^+-model aliphatic dipeptides (glycylglyine, **GG** and alanylalanine, **AA**). The effect of alkali

metal cation on metal cation-dipeptide interactions is elucidated by comparing K^+-GG with previous studies on Ag^+/Na^+-GG.

Chapter 6: Further extending our studies form model aliphatic dipeptides **GG** to dipeptides containing phenylalanine (**FG** and **GF**), in order to understand the role of cation-π interaction in: (a) the relative stability of charge-solvated versus zwitterionic forms, and (b) the effect of the position of the aromatic ring of phenylalanine (as N-terminal versus C-terminal substituent) on K^+ binding geometry and affinity.

Chapter 7: Explaining the difference in the observed mass spectrometric fragmentation pattern of H^+-histidine and K^+-histidine by obtaining the detailed potential energy surface of H^+-histidine, and propose plausible mechanisms for the fragmentation of K^+-histidine based on a key stable structures obtained by theoretical calculations.

Chapter 8: Conclusions arising from this series of studies will be summarized.

Chapter 2 Theoretical Methodology

In this chapter, the theoretical modeling methodologies used in our study will be introduced. All *ab initio* and density functional theory calculations have been carried out using the Gaussian98 [Frisch et al., 1998] package of programs Compaq XP900, Intel Pentium 4 and SGI Supercomputers workstations.

2.1 *Ab Initio* Molecular Orbital (MO) Methods

Molecular orbital theory uses one-electron functions to approximate the full molecular wavefunction and *ab initio* means from the beginning. Thus, *ab initio* molecular orbital theory aims to predict the properties of atomic and molecular systems based on the laws of quantum mechanics, and using only fundamental constant, without experimental parameterizations.

2.1.1 The Schrödinger Equation

The energy of a stationary state of a molecule can be obtained by solving the Schrödinger equation: [Hehre et al., 1986]

$$\hat{H}\Psi = E\Psi \qquad [2.1]$$

where $\hat{H}$ is the Hamiltonian (a differential operator representing the total energy, consists of kinetic and potential parts); Ψ is the wavefunction (a function of the Cartesian coordinates of all particles and spin coordinates in the atomic/molecular system) and E is the energy of state of a system (relative to a state in which the constituent particles, nuclei and electrons, are infinitely separated and at rest).

2.1.2 The Born-Oppenheimer Approximation

The Hamiltonian for a molecular system, $\hat{H}$, can be written as the kinetic and potential energy operator of the nuclei and electrons as shown below: [Hehre et al., 1986]

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \qquad [2.2]$$

where $\hat{T}_n$ *and* $\hat{T}_e$ are the kinetic energy operator for nuclei and electrons; $\hat{V}_{ne}, \hat{V}_{ee}$ *and* $\hat{V}_{nn}$ are the potential energy operator for nuclei-electron attraction, electron-electron repulsion and nuclei-nuclei repulsion, respectively.

The Born-Oppenheimer approximation simplifies the molecular problem by separating nuclear and electronic motions. This approximation is generally reasonable since the mass of nucleus is thousands of times greater than that of an electron. The nuclei move very slowly with respect to the electrons, and the electrons react essentially instantaneously to the changes in nuclear positions. Thus, we can construct an electronic Hamiltonian ($\hat{H}_e$), which neglects the kinetic energy term for the nuclei:

$$\hat{H} \approx \hat{H}_e = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \qquad [2.3]$$

2.1.3 The Slater Determinant

A molecular orbital (ψ) is a function of the Cartesian coordinates (x, y, z) of a single electron. Its square (ψ^2) is the probability distribution of the electron in space. To

describe an electron completely, the spin coordinates (ξ = +1/2 or −1/2 for α, β respectively) also have to be included. Thus, a complete wavefunction for a single electron is the product of a molecular orbital (ψ) and a spin function (ξ), named as the spin orbital.

For an n-electron system, the full wavefunction (Ψ), often referred to as a Slater determinant, Ψ_{Slater}, is represented by Eqn. [2.4].

$$\Psi_{Slater} = (n!)^{-\frac{1}{2}} \begin{vmatrix} \psi_1(1)\xi_\alpha(1) & \psi_2(1)\xi_\beta(1) & \bullet\bullet\bullet & \psi_n(1)\xi_\beta(1) \\ \psi_1(2)\xi_\alpha(2) & \psi_2(2)\xi_\beta(2) & \bullet\bullet\bullet & \psi_n(2)\xi_\beta(2) \\ & & \bullet & \\ & & \bullet & \\ & & \bullet & \\ \psi_1(n)\xi_\alpha(n) & \psi_2(n)\xi_\beta(n) & \bullet\bullet\bullet & \psi_n(n)\xi_\beta(n) \end{vmatrix} \qquad [2.4]$$

The first row of the Ψ_{Slater} included all the possible spin orbital assignments for electron 1, and similarly for the other electrons in other rows. On expansion, the Ψ_{Slater} becomes a sum of products of spin orbitals. In order to ensure the probability of finding the electron anywhere in space is unity, a normalization constant of $(n!)^{-1/2}$ is needed. The Ψ_{Slater} fulfills two important properties of molecular wavefunction:

(i) It satisfies the antisymmetric requirement of a wavefunction, i.e., when the coordinates of any two electrons are interchanged, it is equivalent to the interchange of the corresponding rows of these two electrons in the determinant, which does have the effect of changing the +/- sign. [Jensen, 1999]

(ii) It also satisfies the requirement arises from the Pauli exclusion principle, i.e., no two electrons of the same spin may occupy the same molecular orbital. It

is because the product terms vanish when two columns in Ψ_{Slater} are identical. [Jensen, 1999]

2.1.4 Basis Set Approximation

The molecular orbitals (ψ) can be expressed as a linear combinations of a finite set of N prescribed one-electron functions represented by *known* basis functions (ϕ_x). For example, Eqn. [2.5], the i^{th} molecular orbital is represented by a linear combination of a series (x=1, 2, 3, ...N) basis functions (ϕ_x) with coefficient (c_{xi}) as the molecular orbital expansion coefficients.

$$\psi_i = \sum_{x=1}^{N} c_{xi}\phi_x \qquad [2.5]$$

where i represents a series of molecular orbitals, x represents a series of basis functions, and N is the dimension of the basis function. The expansion coefficients in Eqn. [2.5] are determined by applying the variational principle in a self-consistent fashion (described in Section 2.1.6 below).

2.1.5 Basis Sets

Slater and Gaussian Type Orbitals

There are two types of basis functions, the Slater and Gaussian type of orbital functions. The Slater Type Orbitals (STO) have exponential radial parts. They are

labeled like hydrogen atomic orbitals, 1s, 2s, 2px,..., and have the normalized form. The one with STO_{1s} is as follows: [Hehre et al., 1986]

$$STO_{1s} = \left(\frac{\zeta_1^3}{\pi}\right)^{1/2} \exp(\zeta_1 r) \qquad [2.6]$$

where ζ_1 is constant determining the size of the orbitals. STOs provide reasonable representations of atomic orbitals with ζ values recommended by Slater. However, they are not suitable for computational calculations. Thus, a second type of basis functions, the Gaussian-type basis function, GTO, is introduced. The Gaussian basis function for the s-type atomic orbitals, GTO_s, which have angular symmetries, is represented as follows: [Hehre et al., 1986]

$$GTO_s = \left(\frac{2\alpha}{\pi}\right)^{3/4} \exp(-\alpha r^2) \qquad [2.7]$$

Pople Style Basis Sets

Pople style basis sets are the basis sets that we used extensively in this work, and thus more details are introduced here. Pople style basis sets are one type of contracted basis sets. Their basis functions, also known as the primitive gaussian type orbitals (PGTO), are combined together by fixed linear combinations, to form the contracted gaussian type orbitals (CGTOs). Here are some of the examples:

(i) STO-nG basis set: n = number of PGTO, that are contracted to form a fit to the Slater type orbital (STO). Usually, STO-3G is the widely used minimum basis.

(ii) k-nlmG basis set: split valence basis sets. The first part 'k' means k PGTO are used for representing the core orbitals, while the second part

'nlm' describes how many functions the valence orbitals (CGTO) are split into (now is three) with different sizes, and how many PGTO are used to represent the valence orbitals (denoted as 'nlm' here, which represented n-, l- and m- PGTOs, respectively).

(iii) Sometimes, more parameters are added before the G (for Gaussian) and after the G, to indicate diffuse and polarization functions added. The diffuse functions (denoted as '+') are large-size version of s- and p-type functions. They allow orbitals to occupy a larger region of space and directional characters. They are found to be important for systems where electrons are relatively far away from the nucleus, for molecules with lone pairs, anions and other systems with significant negative charge, and also for systems in their excited states. The polarization functions are higher angular momentum functions, i.e., a p-orbital introduces polarization of the s-orbitals, while d-orbitals can be used for polarizing p-orbitals, and f-orbitals for d-orbitals, etc. They are particularly useful in describing chemical bonds.

Taking the highest basis set we used in this thesis as an example, the 6-311+G(3df,2p) basis set indicates a triple split valence basis, where the core orbital is made up of a contraction of six PGTOs, and the valence orbitals is split into three functions, represented by three-, one- and one PGTOs, respectively. The '+' indicates the application of one set of diffuse s-p-functions on heavy atoms. The polarization functions (3df,2p) indicate three d- and one f-functions are added on heavy atoms, as well as two p-functions on the hydrogens.

2.1.6 Hartree-Fock Self-Consistent Field (SCF) Theory

To generate approximate solutions for the Schrödinger equation, the variational principle, which states that any approximate wavefunction has an energy above or equal to the exact energy, is employed (i.e., $E_{trial} \geq E_{exact}$). In other words, the variational principle is to minimize the energy of a single Slater determinant by choosing suitable values for the MO coefficient, under the constraint that the MOs retain orthonormal. By making a trial wavefunction containing a number of parameters, we can generate the "best" trial function of the given form by minimizing the energy as a function of these parameters.

Hartree-Fock (HF) self-consistent field method is based on this variational principle. In simple terms, if a basis set is selected for orbital expansion, the coefficients (c_{xi} in Eqn. [2.5]) may then be adjusted iteratively to minimize the expectation value of energy, E_{trial}. Hence, the best single-determinant wavefunction, Ψ_{HF}, is found by minimizing E_{trial} with respect to the coefficients c_{xi}, by the following relationship: [Hehre et al., 1986]

$$\partial E_{trial} / \partial c_{xi} = 0 \qquad [2.8]$$

2.1.7 Roothaan-Hall Equations

By applying the variational principle, a set of algebraic equations for c_{xi} can be formed. They can be solved independently by Roothaan-Hall equation for a closed shell system (Eqn. [2.9]). In other words, the Roothaan-Hall equations allow the solution of HF equations using matrix techniques, since these equations can be considered as a matrix representation of the HF equation in an infinite basis. In

practice, for computational efficiency, the number of basis functions is limited. Thus, these calculations cannot reach the HF limit. The Roothaan-Hall equation is given as below: [Hehre et al., 1986]

$$\sum_{\nu=1}^{N}(\bar{F}_{\chi\nu} - \varepsilon_i \bar{S}_{\chi\nu})c_{\nu i} = 0 \qquad [2.9]$$

where $\bar{F}_{\chi\nu}$ and $\bar{S}_{\chi\nu}$ are the Fock matrix and overlap matrix, respectively, between χ and ν basis functions, while ε_i is the one electron energy of the molecular orbital, ψ_i. For normalization conditions, the MO coefficients are given as

$$\sum_{\chi=1}^{N}\sum_{\nu=1}^{N} c_{\chi i}^{*} \bar{S}_{\chi\nu} c_{\nu i} = 1 \qquad [2.10]$$

To determine the unknown MO coefficients, $c_{\nu i}$, the Fock matrix, $\bar{F}_{\chi\nu}$, must be known. On the other hand, the Fock matrix is only known if all the coefficients, $c_{\nu i}$, are known. Thus, an iterative process is needed to test for convergence within a certain threshold (energy and $c_{\nu i}$). If the iteration fails, then the next iteration will be initiated. If the iteration is successful (convergence achieved), the energy and $c_{\nu i}$ needed can be solved. The whole iteration process is illustrated in Fig. 2.1.

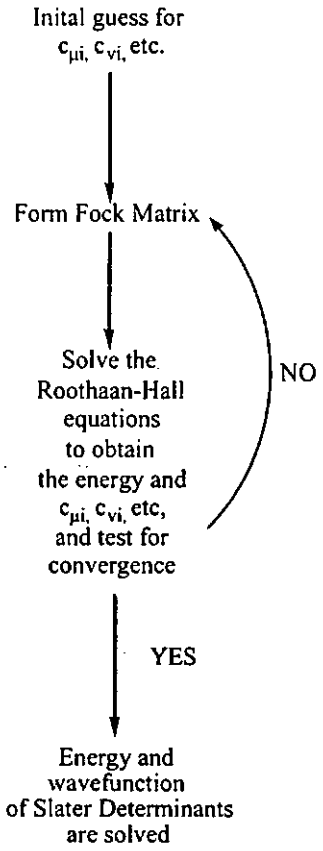


Figure 2.1. Iterative process of SCF procedure.

2.2 Density Functional Theory (DFT)

In contrast to *ab initio* theory where all analyses are based on the wavefunction, electron density (ρ) is used in all DFT methods. In DFT, the electronic energy is expressed as a functional of density, $E[\rho]$, i.e., the goal of DFT methods is to design functionals connecting the electron density with energy. [Parr and Yang, 1989; Bartolotti and Flurchick, 1996; Baerends and Gritsenko, 1997]

Under DFT formalism, the total energy is decomposed into three contributions, a kinetic energy, a Coulomb energy (due to classical electrostatic interactions among all

charged particles in the system), and a term called the exchange-correlation energy which captures all many-body interactions. [Koch and Holthausen, 2000]

According to Kohn and Sham, the kinetic energy functional can be split up further into two parts, one of which can be calculated exactly and a small correction term:

$$\begin{aligned} E[\rho] &= T[\rho] + T_s[\rho] + E_{ne}[\rho] + E_{ee}[\rho] \\ &= T[\rho] + T_s[\rho] + E_{ne}[\rho] + (J[\rho] + K[\rho]) \\ &= T_s[\rho] + E_{ne}[\rho] + J[\rho] + (K[\rho] + T[\rho]) \\ &= T_s[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \end{aligned} \qquad [2.11]$$

The $E[\rho]$ is the energy functional, $T_s[\rho]$ is the part of the kinetic energy contribution, which can be solve exactly; $T[\rho]$ is the part of the kinetic energy with a small correction, which cannot be solved exactly; $E_{ne}[\rho]$ is the attraction between the nuclei and electrons, while the coulomb and an exchange contribution to the electron-electron repulsion ($E_{ee}[\rho]$), is given by $J[\rho]$ and $K[\rho]$, respectively. Note that part of the kinetic energy which cannot be solved by $T_s[\rho]$, i.e., $T[\rho]$, is also absorbed into an exchange contribution, $K[\rho]$, to form the exchange-correlation term of electron-electron repulsion, $E_{xc}[\rho]$.

It is noted that $T_s[\rho]$, $E_{ne}[\rho]$ and $J[\rho]$ can also be solved exactly. Thus, the first task for DFT is deriving suitable formulae for the exchange-correlation term, $E_{xc}[\rho]$. Then, the second task is to determine a set of orthogonal (to enforce the Pauli principle) orbitals which minimize the energy. Since both $J[\rho]$ and $E_{xc}[\rho]$ functional depends on the total density, an iterative procedure is needed. Usually, the $E_{xc}[\rho]$ is

further partitioned into two parts, a pure exchange (E_x) and a pure correlation part (E_c), i.e., $E_{xc}[\rho] = E_x + E_c$. The difference between various DFT methods relies precisely in the choice of the functional form of the exchange and the correlation function (i.e. $E_{xc}[\rho]$).

There are three types of DFT methods: Local Density Approximation (LDA) methods, Gradient Corrected Approximation (GCA) Method, and Hybrid Methods. In LDA, The density is treated locally as a uniform electron gas. For GCA, the exchange and correlation energies are dependent not only on the electron density, but also on derivatives of the density. The density functional used in this work, B3-LYP, is one of the most commonly used DFT methods and it is a hybrid functional method.

2.2.1 Hybrid Methods

In hybrid method, the exchange energy of a Slater determinant is computed from the HF method. The B3-LYP functional employed in this work makes use of Becke's three parameters exchange function (B3-) and the correlation functional of Lee, Yang and Parr (-LYP). [Lee, et al. 1988] The E_{xc} is given as below:

$$E_{xc}^{B3-LYP} = aE_x^{HF} + (1-a)E_x^{LSDA} + b\Delta E_x^{B} + cE_c^{LYP} + (1-c)E_c^{LSD} \qquad [2.12]$$

The constants a, b and c are those determined by Becke by fitting the G1 molecule set. Becke determined the values of the three parameters by fitting to the 56 atomization energies, 42 ionization potentials and 8 proton affinities and 10 first-row atomic energies in the G1 molecule set, and the computed values of a=0.80, b=0.72 and c=0.81 were used. [Frisch et al., 1998]

Because of the parameterization, B3-LYP functionals perform almost as well as the more elaborate G1/G2 model, with the unsigned error for energetics at about 8 kJ mol^{-1}. [Koch and Holtausen, 2000] The performance of the B3-LYP level of theory in term of geometries is also impressive. [Koch and Holthausen, 2000] For a set of 20 organic molecules, geometries optimized at the B3-LYP/6-31G(d) level were found to be in error by less than 0.005Å on average for bond lengths, and bond angels were accurate to within a few tenths of a degree, which is of the same order as the uncertainties in the experimental equilibrium structures for most polyatomic molecules. Because of this, several authors suggested [Frisch et al., 1998] that the B3-LYP functional could be used instead of MP2 geometries for the highly accurate extrapolation schemes like G2 method.

2.3 Zero-point Corrections and Frequencies

All molecules have some vibrational energy even at 0 Kelvin, arising from the Heisenburg uncertainty principle. Thus, a zero-point energy term, ZPE, has to be added to the calculated total energy (i.e., the electronic energies, E_{elec}, calculated above, which assume the system is completely at rest) to obtain the 'true' energy of the molecule, that is:

$$E_0 = E_{elec} + ZPE \qquad [2.13]$$

Molecular frequencies, ν, depend on the second derivative of the energy with respect to the nuclear positions, r, that is:

$$\nu = \partial^2 E / \partial r^2 \qquad [2.14]$$

Raw frequency values computed with the Hartree-Fock method (and other quantum mechanical methods) are known to contain systematic errors, generally overestimates the frequency for the Hartree-Fock method by 10%-12%. Therefore, it is a common practice to scale down frequencies predicted at the Hartree-Fock level, with a scale factor of 0.8929. [Foresman, 1998]

2.4 Thermochemical Corrections

In order to obtain the energy of a system at higher temperature, a thermal energy correction must be added into the zero-point corrected electronic energy, E_0 (Eqn. [2.15]). The thermal correction includes the effects of molecular translation, E_{trans}, rotation, E_{rot}, and vibration, E_{vib}, at the specified temperature, T, and pressure, that is:

$$E_{tot} = E_0 + E_{vib} + E_{rot} + E_{trans} \qquad [2.15]$$

The E_{trans}, E_{rot} and E_{vib} can be estimated by statistical mechanics equation, assuming that the molecule is a rigid rotor and a harmonic oscillator. [McQuarrie and Simon, 1997] From the total energy, enthalpy can be obtained:

$$H = E_{tot} + RT \qquad [2.16]$$

, where R is gas constant (8.314 J K^{-1} mol^{-1}). In order to obtain the Gibbs free energy (G):

$$G = H - TS \qquad [2.17]$$

Entropy (S) is estimated from harmonic vibrational frequencies, via statistical mechanics equations. [McQuarrie and Simon, 1997]

2.5 Molecular Dipole Moment

The molecular dipole moment reported in our study is calculated at the B3-LYP levels, and is obtained according to the following equation:

$$\mu = -\left(\frac{\partial E}{\partial F}\right)_0 \qquad [2.18]$$

where E and F is the energy and external electric field. The dipole moment calculated by hybrid functionals are found to be very reliable, for example, by using B3-LYP protocol, the mean absolute deviation is only 0.04 Debye in average. [Koch, Holthausen, 2000]

2.6 Molecular Polarizability

The polarizabilities reported in this project are calculated at the B3-LYP level according to Eqn. [2.19]:

$$\alpha_{ij} = -\left(\frac{\partial^2 E}{\partial F_i \partial F_j}\right)_0 \qquad [2.19]$$

where E and F are the energy and external electric field, respectively. Compare to experimental results, the polarizabilities calculated at B3-LYP functionals have been shown to be reliable. [Koch and Holthausen, 2000]

2.7 The 'Energetic Protocol for K^+'

In this work, the interaction between potassium cation and the ligand in the gas phase was modeled with the 'Energetic Protocol for K^+', abbreviated as EP(K^+). The procedure is based on the energetics and structures determined with the hybrid Becke3-Lee-Yang-Parr (B3-LYP) [Becke, 1993] density functionals.

The computational procedures of the EP(K^+) protocol is as follows:

(1) Geometry optimization at the HF level using the standard 6-31G(d), followed by frequency calculations in order to confirm to be real minima and obtain the zero-point-energy (ZPE) correction at this level;

(2) Geometries of these stable minima were further refined by the B3-LYP functional with the 6-31G(d) basis set in order to obtain the effect of electron correlation on structures of ligands and K^+-ligand complexes; [Becke, 1993]

(3) Energetics are obtained by using the B3-LYP functional with the large and flexible 6-311+G(3df,2p) basis set based on geometry determined in step (2), i.e., the energetic calculations were carried out at the B3–LYP/6–311+G(3df,2p)//B3–LYP/6-31G(d) level.

Potassium cation affinities (PCAs) at 0K, ΔH_0, are obtained using Eqn. [2.20]:

$$\Delta H_0 = [(E_{K^+} + E_L) - E_{K^+\text{-}L}] + [ZPE_L - ZPE_{K^+\text{-}L}] \times 0.8929 \qquad [2.20]$$

, where E_{K^+}, E_L and $E_{K^+\text{-}L}$ are the electronic energies of the potassium metal cation, the ligand and the K^+-ligand complex, respectively, obtained from step (3); ZPE_L and $ZPE_{K^+\text{-}L}$ are the zero-point-energy corrections for the ligand and the K^+-L complex obtained from step (1), respectively, with a scaling factor of 0.8929 for the Hartree-Fock frequencies. [Frisch et al., 1998]

For the purpose of comparison with experimental values, the EP(K^+) theoretical values at 0K (ΔH_0) are converted to affinities at 298K (ΔH_{298}) and free energy of binding (ΔG_{298}) of different conformers/isomers by applying standard statistical thermodynamics relations [DelBene, 1983] calculated from the scaled HF/6-31G(d) vibrational frequencies.

In summary, all studies reported in this work employed is modeled by the following level of theory, unless otherwise noted:

(1) Geometries: B3-LYP/6-31G(d);

(2) Single point energies: B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d);

(3) Zero-point energy (ZPE) and thermal correction: HF/6-31G(d);

(4) Deformation energy (E_{def}, defined in Section 1.3.5): B3–LYP/6–31G(d)//B3–LYP/6–31G(d);

(5) Dipole moment (μ), polarizabilities (α), dipole interaction parameter (DIP, defined in Section 1.3.3) and polarization interaction parameter (PIP, defined in Section

1.3.3): B3-LYP/6-31G(d)//B3-LYP/6-31G(d) by Mulliken population analysis, using the geometries of the deformed ligand.

The B3-LYP hybrid method has been found to be the best overall performance for problems related to gas-phase ion chemistry, [Alcami et al., 2001] with deviations from the best experimental values typically by no more than 2-3 kJ mol^{-1}. [Wang et al., 1999; Rodriquez et al., 2000; Addario, 2000] We found this observation to be valid in the calculation of K^+ affinities, even though a relatively large basis set has to be used. The agreement between the experimental and calculated ΔH_0 values lends credence to the optimized geometries and energetics obtained in this study. In addition, the basis set we used, 6-31G(d), is a medium size basis set which has been demonstrated to be adequate for geometry optimizations and excitation energy calculations on 'co-valent' molecules. [Nguyen and Pachter, 2001] The 6-311+G(3df,2p) basis set were used because calculations without diffuse and polarization functions for the heavy and hydrogen atoms may be inadequate for calculations on ion-molecule complexes. [Bruyneel, 2003]

2.8 Conformation Search for Dipeptides

The majority of the studies reported in this work is modeled by *ab initio* molecular orbital and density functional theory methods, which are based on quantum mechanics (described in Sections 2.1. and 2.2). These methods are expected to be reasonably reliable and accurate, but at the same time computational expensive, especially for the larger dipeptides and their complexes described in Chapter 5 and 6. To reduce the number of conformers/isomers of dipeptides systems that we have to calculate quantum mechanically, we employed Monte Carlo conformation search, using force field calculated energies. For this purpose, the Macromodel 7.0 package of programs (running from Intel Pentium III workstation) based on the AMBER* force field and Monte Carlo Multiple Minimum (MCMM) search method was used in our study. Macromodel is an integrated software system for modeling organic and bioorganic molecules using molecular mechanics written by Mohamadi et al. [Mohamadi et al., 1990]

2.8.1 Force Field

In the force field method, a molecular system is treated as a collection of hard spheres, connected to each other by springs (depicted in Fig. 2.2):

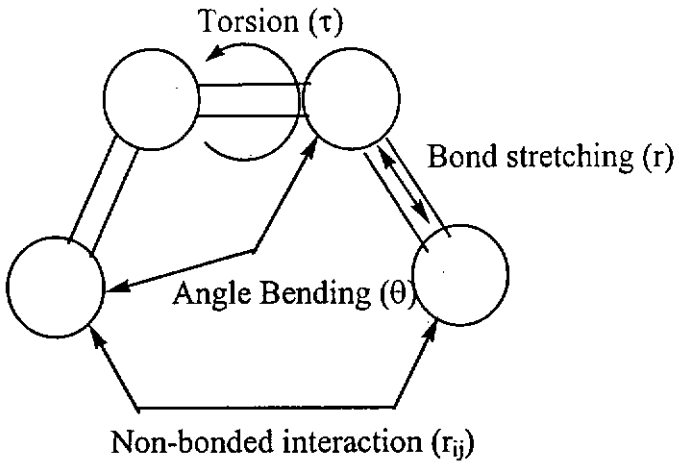


Figure 2.2. The definition of bond stretching, angle bending, torsional and non-bonded interaction in molecular mechanics modeling.

The total energy (relative to the strained free system) of a molecule, E_{FF}, is modeled as the sums of the following terms:

$$E_{FF} = E_{str} + E_{bend} + E_{tors} + E_{non\text{-}bonded} \qquad [2.21]$$

, where E_{str} represents the energy required for stretching a bond between two atoms from its equilibrium position, r; E_{bend} is energy associated with bending an angle, θ; E_{tors} is torsional energy for rotation around a bond, τ; and $E_{non\text{-}bonded}$ is the non-bonded atom-atom interactions, as functions of different conformations.

The main difference between different force fields lies in the functional form of each energy term and the type of information used for fitting the parameters. In the current study, AMBER* force field [Weiner et al., 1984] is employed. This force field is chosen as it is developed especially for the simulation of nucleic acids and proteins.

2.8.2 Monte Carlo Multiple Minimum (MCMM) Search

In Monte Carlo conformational search, a starting structure is first chosen. From this structure, random variations to selected coordinates are applied in order to generate a series of displaced structures. These displaced structure is then energy minimized, in this case, with the AMBER* force field calculated energy (described in Section 2.8.1). The energy of these structures is compared with minima found during previous conformational search steps. This new minimized structure is accepted if it is within the cutoff energy window from the instant minimum.

In the MCMM search methodology, the K^+ was not connected to any atoms of the dipeptides so that the ion can move freely to interact with any part of the ligand. Point charges and distant-dependent dielectric constant were used in the electrostatic interaction treatment. Extended cutoff bond lengths of 20Å were employed for plausible van der Waals and electrostatic interactions, and hydrogen bonding distances for K^+ complexes. For K^+–dipeptides in the CS and ZW forms, n_{conf} Monte Carlo steps (where n_{conf} = 1500 x number of rotatable torsional angles of the dipeptide ligand in CS or ZW form) were carried out to locate the low energy structures. Using this method, a limited number of isomers/conformers were obtained within an energy window of 50 kJ mol^{-1}. Any complexes generated from the MCMM step without maximum number of intramolecular hydrogen bondings were discarded. The candidates chosen for further exploration are modeled quantum mechanically.

Chapter 3 Nature of K^+-Small Model Ligand Interactions

3.1 Background

Since K^+ is so important in biological systems, we are interested to study the K^+ binding modes and *intrinsic* binding energies (affinities) of smaller model ligands to serve as the basis to fully understand the binding interaction of K^+ in more complex and larger biological systems. Thus, to start with, we have carried out a systematic theoretical study on 136 organic ligands (20 classes) in the gas phase by using hybrid density functional theory (B3–LYP) calculations with the 6-311+G(3df,2p) basis set.

The present work represents the most comprehensive theoretical studies on the potassium cation affinity (PCA) scale reported up-to-date. For the 136 ligands studied here, the K^+ interacts with different atoms (rare gas, carbon, oxygen, nitrogen, sulfur, and phosphorus, etc.) in the ligands, and with a wide range of functional groups (alcohol, sulfide, sulfoxide, amine, amide, ether, aldehyde, ketone, nitrile, carboxylic acid, aromatic, heterocyclic, etc.). With such a broad spectrum of ligands, we believe that the findings presented here are of general chemical interests and useful in revealing the *intrinsic* nature of K^+ binding to organic and biological ligands in the gas phase without the complicating effect of solvent molecules.

Out of these 136 ligands, 70 experimental values and 64 theoretical values reported in the literature are available for comparison in order to establish the accuracy of our current theoretical *e*nergetic *p*rotocol for K^+, *EP* (K^+) (refer to Chapter 2 for the details). The effect of substituents on the modes of binding and the PCAs of unsubstituted parent ligands are discussed. The relations between Li^+/Na^+ and K^+

affinities for the wide range of ligands and their nature of binding are examined. In addition, in order to test the quality of quantitative results (geometries and binding affinities) obtained by density functional calculations, test calculations were first carried out on smaller organic ligands. The aim is to show that the accuracy of the hybrid density functional theory calculations (B3-LYP with the 6-311+G(3df,2p) basis set, i.e., the PCA obtained by the EP(K^+) protocol as compared with those obtained by the B3-P86, G2(MP2,SVP) protocol, and previously reported theoretical values. The effect of B3-LYP/6-31G(d) correction on zero point vibrational energies, temperature corrections, and the effect of full counterpoise correction on EP(K^+) PCAs are also discussed.

Because of the broad scope of the work, calculations on the ligands hydrogen sulfide, thiols, dioxanes, hydrogen cyanide, alkyl nitriles, sulfoxides, and some members of the alcohols, aldehydes, ketones, carboxylic acids, esters, amide, azoles, amino acids were performed by Mr. J.K.C. Lau, Ms. P.S. Ng, and Dr. F.M. Siu of Prof. Tsang's research group. For the total of 136 ligands, calculations on 66 ligands were carried out by the author.

3.2 Results and Discussion

3.2.1 Overview of Theoretical and Experimental Alkali Cation Affinities

In this section, we would like to highlight some recent advances and issues related to the theoretical and experimental determination of alkali cation affinities reported in the literature. Using density functional based method, Burk et al. recently reported the Li^+ affinity for 63 ligands, calculated at the B3–LYP/6–311+G(d,p) level. [Burk et al., 2000] Calibration against experimental data suggested that this level of theory carried an average unsigned error of 15 kJ mol^{-1}, and the accuracy could be improved if systematic errors are taken into account. [Burk et al., 2000]

Several theoretical studies of Na^+ affinities appeared recently. [Hoyau et al., 1999; McMahon and Ohanessian, 2000; Armentrout and Rodgers, 2000; Petrie, 2001] Most of these studies were conducted at the MP2(full)/6-311+G(2d,2p)//MP2/6-31G(d) level, with basis set superposition error (BSSE) corrections. [Hoyau et al., 1999; McMahon and Ohanessian, 2000] Armentrout and Rodgers further compared the performance of this level of theory and other models (DFT, CBS, G2, G3, etc.) with experimental data. [Armentrout and Rodgers, 2000] Theoretical and experimental sodium affinities were found to be in good general agreement. However, while BSSE corrected B3-LYP/6-311+G(2d,2p)//B3LYP/6-31G(d) sodium affinities are consistently too high (mean-absolute-deviation (MAD) of 8.5 kJ mol^{-1} in comparison with experiment), the BSSE corrected B3–P86/6-311+G(2d,2p)//B3-P86/6-31G(d) values appear to far better (MAD of 5.5 kJ mol^{-1} in comparison with experiment). [Armentrout and Rodgers, 2000] Very recently, in order to improve the quality of the basis set for the sodium inner–valence 2s/2p orbitals, Petrie decontracted the standard 6-311+G(3df) basis for the Na atom. With this more flexible basis set, Na^+ affinities

of 38 ligands at the geometry–corrected counterpoise "CPd-G2thaw" level were obtained, [Petrie, 2001] and it was suggested that the currently established experimental sodium cation affinities were systematically too low by 3–5 kJ mol^{-1}.

For K^+, much fewer theoretical and experimental values are available. Without going into details, we compiled the EP(K^+) theoretical PCAs at 298K obtained in this study and the available experimental PCAs as reported in the original sources in Table 3.1 (also graphically presented in Fig. 3.1), while the geometries of K^+ binding modes for representative class of ligands are shown in Fig. 3.2. Here, we wish to make a note on the temperature reported in the experimental literature. The experimental PCA values were largely obtained by threshold-CID, high pressure (equilibrium) mass spectrometry (HPMS) and kinetic method measurements. The threshold-CID measurements yielded PCAs at 0K, and corrections to 298K (generally less than 2 kJ mol^{-1}) are carried out by theoretical calculations. In HPMS measurements, it is generally assumed that the standard enthalpy of cation binding, ΔH, is approximately equal to the ion-ligand bond dissociation energy, and is independent of temperature effects. Similarly, in kinetic method measurements, the relative affinity measured, $\Delta(\Delta H)$, is also implicitly assumed to be independent of the 'effective temperature', even though the $\Delta(\Delta H)$ term may be anchored to a reference ΔH value at a known temperature. Consequently, experimental values obtained by HPMS and the kinetic methods are usually not reported at any specified temperature. However, temperature effects on the ΔH or $\Delta(\Delta H)$ term are expected to be small. For the 136 ligands studied, we found general agreement (within ±15 kJ mol^{-1}) between our EP(K^+) PCA values and available experimental values reported in the literature. Only for a few sets of PCA values (phenylalanine, cytosine, guanine, adenine obtained by kinetic method measurements, and Me_2SO obtained by high pressure mass spectrometric

equilibrium measurement, as indicated by open circles in Fig. 3.1), the difference between EP(K^+) values and reported experimental values exceeds 15 kJ mol^{-1}. These cases of discrepancy will be further discussed in the individual sections below.

Table 3.1. Theoretical EP(K^+) values at 298K and experimental potassium cation affinities (kJ mol^{-1}).

Molecule [a]	Theoretical [b]	Experimental [c]	Molecule [a]	Theoretical [b]	Experimental [c]
He	4.1		1,2-dioxane	84.1	
Ne	4.1		1,3-dioxane	82.9	
Ar	9.4	15.4(7)[d]	1,4-dioxane	71.0	
CO	24.8	19.0(5.0)[d]	HCHO	77.9	
HF	51.3		MeCHO	93.7	
HCl	27.6		EtCHO	95.8	
P_4	34.4		n-PrCHO	97.3	
PH_3	42.2		n-BuCHO	98.5	
C_2H_2	37.4		CF_3CHO	59.5	
C_2H_4	34.7		CCl_3CHO	74.0	
H_2S	39.6		Me_2CO	104.6	102.1,[k] 108.8[l]
MeSH	52.7		MeCOEt	105.9	
EtSH	57.5		NH_3	77.2	74.9,[e] 84.1,[n] 82(8)[m]
n-PrSH	59.3		$MeNH_2$	79.2	79.9[e]
i-PrSH	60.2		Me_2NH	77.5	81.6[e]
n-BuSH	60.6		Me_3N	74.2	83.7[e]
i-BuSH	63.2		$EtNH_2$	81.8	
t-BuSH	62.0		n-$PrNH_2$	81.3	91.2[e]
Me_2S	61.8		HCN	78.4	
H_2O	70.4	70.7,[e] 74.9[f, g]	MeCN	101.9	102.1(1.7)[o]
MeOH	75.7	79.5,[h] 83.7[h]	EtCN	105.5	
EtOH	81.4		n-PrCN	107.0	
n-PrOH	82.2		i-PrCN	108.2	
i-PrOH	85.2		t-BuCN	110.3	
n-BuOH	86.2		$PhCH_2CN$	110.3	
i-BuOH	80.5		CF_3CN	58.6	
s-BuOH	87.6		CCl_3CN	76.4	
t-BuOH	88.1		HCO_2H	80.4	
1,2-Propanediol	116.2		HCO_2Me	90.2	
1,3-Propanediol	122.5		$MeCO_2H$	91.8	
Ethylene glycol	119.3		HCO_2Et	95.3	
Glycerol	133.9		HCO_2n-Pr	97.0	
CF_3CH_2OH	71.4		$MeCO_2Me$	99.5	
CCl_3CH_2OH	76.1		$MeCO_2Et$	105.7	
Me_2O	74.9	87.0,[e] 74.0(4.0)[i]	$EtCO_2Me$	100.9	
Et_2O	85.0	93.3[e]	CF_3CO_2Me	82.2	
$(MeOCH_2)_2$	123.5	120.0(4.0),[i] 129.7[j]	$ClCO_2Me$	80.9	

Table 3.1. (cont.).

Molecule [a]	Theoretical [b]	Experimental [c]	Molecule [a]	Theoretical [b]	Experiment [c]
SO_2	51.9		1-Me-pyrazole	94.5	94.8(3.6)[w]
Me_2SO	128.1	130.1,[k] 146.4(12.6)[l]	3-Me-pyrazole	92.8	
PhSOMe	132.2		4-Me-pyrazole	96.4	
$HCONH_2$	114.1		1,4-Me_2-pyrazole	100.0	
HCONHMe	120.7	117.7[p]	1,5-Me_2-pyrazole	101.3	
$HCONMe_2$	126.2	123.4,[k] 129.7[l]	1,3,5-Me_3-pyrazole	103.4	
$MeCONH_2$	123.2	124.3,[k] 118.9[p]	3,4,5-Me_3-pyrazole	105.4	
MeCONHMe	128.4	127.2[k]	1,3,4,5-Me_4-pyrazole	106.2	
$MeCONMe_2$	130.6	121.3,[k] 129.7,[l] 131.2[p]	1-Me-imidazole	118.8	117.7(2.7)[v]
Benzene	67.6	76.6,[g] 74.2(4.1),[q] 80.3[r]	1,2-Me_2-imidazole	120.5	
Borazine	46.8		2,4,5-Me_3-imidazole	122.0	
Phenol	70.0	77.9(12.6),[s] 74.6(3.8)[t]	2H-1,2,3-triazole	64.5[y]	55.6(5.5)[x]
Pyridine	93.3	86.6,[e] 90.6(3.9)[u]	1H-1,2,4-triazole	87.0[y]	87.5(4.5)[x]
2-Me-pyridine	94.6	98.5(3.5)[v]	2H-tetrazole	88.5[y]	89.6(4.6)[x]
3-Me-pyridine	99.1	100.1(3.5)[v]	1H-tetrazole	109.7[z]	
4-Me-pyridine	101.1	99.0(4.0)[v]	4H-1,2,4-triazole	140.1[z]	
2-F-pyridine	100.3		1H-1,2,3-triazole	118.6[z]	
3-Cl-pyridine	81.9		Glycine	118.4	125.5,[k] 119.3[a']
1,8-naphthyridine	155.6		Alanine	124.0	123.6[a']
Pyridazine	130.0	130.9(2.6)[u]	Valine	128.2	128.0[a']
Pyrimidine	75.7	69.7(4.3)[u]	Leucine	128.8	129.3[a']
Pyrazine	69.8	67.6(3.6)[u]	Iso-Leucine	129.8	129.9[a']
1,3,5-triazine	56.5	55.6(3.0)[u]	Proline	143.0	
Pyrrole	77.1	85.1(3.6)[w]	Serine	137.5	
Indole	89.1	99.0(12.6)[s]	Cysteine	123.5	
Pyrazole	90.5	84.2(3.3)[w]	Phenylalanine	145.6	104.2 (20.9)
Imidazole	111.1	109.6(5.6)[x]	Adenine	87.1[c']	106,[d'] 97.4(3.2)[e']
Thiazole	87.1		Thymine	112.0[c']	102,[d'] 104.6(3.8)[e']
Isothiazole	86.1		Uracil	113.1[c']	101,[d'] 105.0(2.8)[e']
Oxazole	83.5		Cytosine	166.3[c']	110[d']
Isoxazole	94.4		Guanine	143.3[c']	117[d']

[a]: Abbreviations: Me = $-CH_3$, Et = $-C_2H_5$, n-Pr = $-C_3H_8$, i-Pr = $-(CH_3)_2CH$, n-Bu= $-C_4H_{11}$, i-Bu = $-(CH_3)_2CHCH_2$, t-Bu= $-(CH_3)_3C$, Ph= $-C_6H_5$. [b]: This work , theoretical EP(K^+) affinities at 298K

(ΔH_{298}). [c]: Experimental affinities are tabulated at 298K (ΔH_{298}) or at unspecified temperatures (see text for further discussions). Numbers in parenthesis represent reported experimental uncertainties, if available. [d]: Threshold-CID, ΔH_{298}, adjusted from ΔH_0 reported in ref. [Walter et al., 1998] in this work. [e]: HPMS, ref. [Davidson and Kebarle, 1976a]. [f]: HPMS, ref. [Dzidic and Kebarle, 1970]. [g]: HPMS, ref. [Sunner et al., 1981]. [h]: estimated from ligand exchange-HPMS data, ref. [Nielsen et al., 1999]; first value anchored to experimental PCA of water reported in ref. [Davidson and Kebarle, 1976a] and the second value anchored to experimental PCA of water reported in ref. [Dzidic and Kebarle, 1970] and ref. [Sunner et al., 1981]. [i]: Threshold-CID, ΔH_{298}, ref. [More et al., 1997]. [j]: HPMS, ref. [Davidson and Kebarle, 1976c]. [k]: Threshold-CID, ΔH_{298}, ref. [Klassen et al., 1996]. [l]: HPMS, ref. [Sunner and Kebarle, 1984]. [m]: Threshold-CID, ΔH_{298}, ref. [Iceman and Armentrout, 2003]. [n]: HPMS, ref. [Castleman, 1978a]. [o]: HPMS, ref. [Davidson and Kebarle, 1976b]. [p]: ref. [Tsang et al., 2001a], kinetic method measurements using theoretical G2(MP2,SVP)-ASC(GCP) K^+ affinity values (ΔH_0) values at 0K of formamide (109.2 kJ mol^{-1}) / N,N'- dimethylformamide (123.9 kJ mol^{-1}) / N-methylacetamide (125.6 kJ mol^{-1}) as reference values as reported in ref. [Siu et al., 2001a], and corrected to ΔH_{298}. [q]: Threshold-CID, ΔH_{298}, ref. [Amicangelo and Armentrout, 2000]. [r]: HPMS, ref. [Keesee and Castleman, 1986]. [s]: Radiative associative kinetics measurement, ref. [Ryzhov and Dunbar, 1999], reported value reduced by 6.2 kJ mol^{-1} according to text, and with thermal correction to 298K. [t]: Threshold-CID, ΔH_{298}, ref. [Amunugama and Rodgers, 2002]. [u]: Threshold-CID, ΔH_{298}, ref. [Amunugama and Rodgers, 2000]. [v]: Threshold-CID, ΔH_{298}, ref. [Rodgers, 2001]. [w]: Threshold-CID, ΔH_{298}, ref. [Huang and Rodgers, 2002]. [x]: Threshold-CID, ΔH_{298}, ref. [Rodgers and Armentrout, 1999]. [y]: EP(K^+) PCA of the *most stable* tautomer of the *free* ligands: 1H-1,2,4-triazole, 2H-1,2,3-triazole and 2H-tetrazole *kinetically (but not energetically) favored* to bind to K^+ in the gas phase, resulting in the formation of K^+-(1H-1,2,4-triazole), K^+–(2H–1,2,3-triazole) and K^+-(2H-tetrazole) complexes proposed to be likely observed in threshold-CID experiments as explained in ref. [Rodgers and Armentrout, 1999]. [z]: EP(K^+) PCA of the less stable tautomer of the free ligands : 4H-1,2,4-triazole, 1H-1,2,3-triazole and 1H-tetrazole but energetically (but not kinetically) favored to bind to K^+ in the gas phase, leading to formation of the most stable K^+–(4H–1,2,4–triazole), K^+-(1H-1,2,3-triazole) and K^+-(1H-tetrazole) complexes with the largest PCA, but not likely observed in threshold-CID experiments as explained in ref. [Rodgers and Armentrout, 1999]. [a']: kinetic method measurements using theoretical G2(MP2,SVP)–ASC(GCP) K^+ affinity values at 0K (ΔH_0) of acetamide (118.7 kJ mol^{-1}) / N-methylacetamide (125.6 kJ mol^{-1}) / N,N'-dimethylacetamide (129.2 kJ mol^{-1}) as reference values in ref. [Tsang et al., 2004], and the experimentally found values were corrected to ΔH_{298}. [b']: Kinetic method, ref. [Ryzhov et al., 2000]. [c']: the PCA is estimated for the most stable tautomer of the free ligands, corresponding to species A1, T1, U1, C1 and G1 in ref. [Russo et al., 2001b]. [d']: Kinetic method, ref. [Cerdà and Wesdemiotis, 1996]. [e']: Threshold-CID, ΔH_{298}, ref. [Rodgers and Armentrout, 2000b].

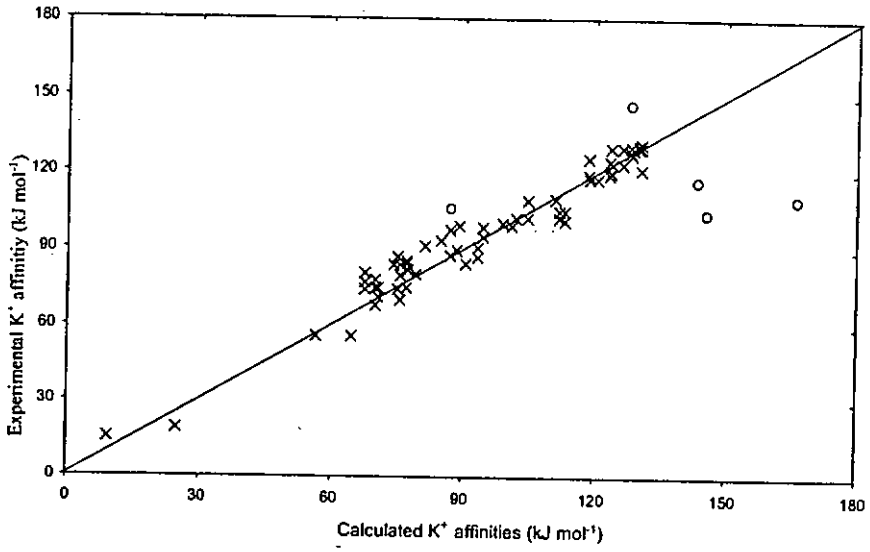


Figure 3.1. Plot of experimental potassium cation affinities versus EP(K^+) calculated potassium cation affinities: the diagonal line with a slope of 1.0 is drawn for reference purposes. Large differences exceeding 15 kJ mol^{-1} are depicted as 'o' in the figure.

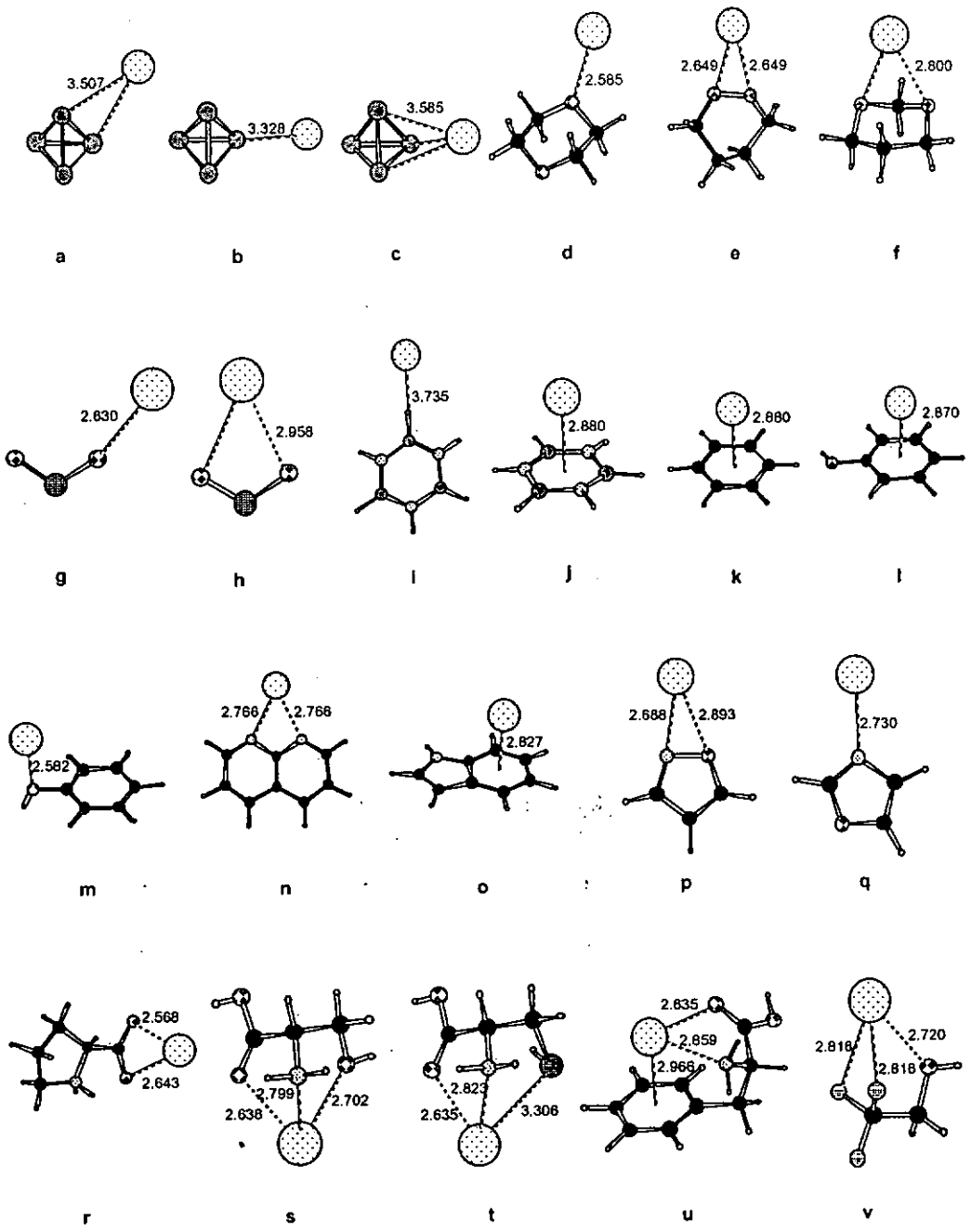


Figure 3.2. The geometries of selected K^+-ligand complexes, optimized at the B3–LYP/6–31G(d) level of theory. Ligand in a-c = P_4; d = 1,4-dioxane; e = 1,2–dioxane; f = 1,3-dioxane; g-h = SO_2; i-j = borazine; k = benzene; l–m = phenol; n = 1,8-naphthyridine; o = indole; p = isoxazole; q = oxazole; r = proline; s = serine; t = cysteine; u = phenylalanine; v = CF_3CH_2OH. Selected non-bond distances are shown in Angstroms.

In addition, we also compiled the existing theoretical PCAs in the literature (Table 3.2), corrected to 298K if necessary, to facilitate comparison. Different theoretical models were employed in various studies, in which different geometries, electron correlation methods, basis sets, core sizes, zero-point energies are employed and BSSE corrections may or may not be included. [Ikuta, 1984; Magnusson, 1994; More et al., 1997; Hoyau and Ohanessian, 1998; Ryzhov and Dunbar, 1999; Rodgers and Armentrout, 1999; Amicangelo and Armentrout, 2000; Amunugama and Rodgers, 2000; Dunbar, 2000; Rodgers and Armentrout, 2000b; Russo et al., 2001b; Siu et al., 2001a; Rodgers, 2001; Abirami et al., 2002; Amunugama and Rodgers, 2002; Huang and Rodgers, 2002; Iceman and Armentrout, 2003] Even though direct comparison is difficult, it appears that the PCAs estimated by the EP(K^+) protocol tend to be slightly larger than the affinities estimated by other methods, except when the K^+ binds to the ligand via aromatic π binding (as in the case of benzene, phenol and pyrrole). Nevertheless, the general agreement is good (within ± 5 kJ mol^{-1}) except for a few species (CO, H_2S, NH_3, imidazole, glycine and the five DNA/RNA nucleobases). Detailed discussions of the newly found discrepancies will be given in individual sections below.

Table 3.2. Theoretical potassium cation affinities (PCAs) (kJ mol^{-1}).

Molecule [a]	Theoretical [b]	Literature [c]	Molecule [a]	Theoretical [b]	Literature [c]
CO	24.8	37.1[d]	Phenol	70.0	74.1[l]
PH_3	42.2	41.4[e]	Pyridine	93.3	91.3[m], 91.4[n]
H_2S	39.6	31.6[e]	2-Me-pyridine	94.6	94.5[m]
H_2O	70.4	74.7[e], 70.7[f], 68.4[g]	3-Me-pyridine	99.1	97.0[m]
MeOH	75.7	74.9[g]	4-Me-pyridine	101.1	97.9[m]
EtOH	81.4	81.4[g]	Pyridazine	130.0	130.5[n]
n-PrOH	82.2	81.2[g]	Pyrimidine	75.7	73.2[n]
i-PrOH	85.2	85.2[g]	Pyrazine	69.8	68.7[n]
n-BuOH	86.2	85.8[g]	1,3,5-triazine	56.5	53.5[n]
i-BuOH	80.5	83.6[g]	Pyrrole	77.1	82.1[o]
s-BuOH	87.6	87.4[g]	Indole	89.1	87.2[p]
t-BuOH	88.1	88.7[g]	Pyrazole	90.5	86.7[o]
1,2-Propanediol	116.2	117.9[h]	Imidazole	111.1	109.3[o], 116.8[q]
1,3-Propanediol	122.5	122.2[h]	1-Me-pyrazole	94.5	91.3[o]
Ethylene glycol	119.3	118.2[h]	1-Me-imidazole	118.8	117.2[o]
Glycerol	133.9	134.3[h]	2H-1,2,3-triazole	64.5	65.5[q]
Me_2O	74.9	79.0[i]	1H-1,2,4-triazole	87.0	91.7[q]
NH_3	77.2	79.5[e], 77.6[f], 75.1[g], 75.1,[j] 77.1,[j] 76.1,[j] 64.1,[j] 66.1[j]	2H-tetrazole	88.5	90.7[q]
$HCONH_2$	114.1	111.2[g]	Glycine	118.4	110.9[r]
HCONHMe	120.7	118.9[g]	Phenylalanine	145.6	146.4[s]
$HCONMe_2$	126.2	124.5[g]	Adenine	87.1[t]	78.7[f], 85.2[u]
$MeCONH_2$	123.2	120.2[g]	Thymine	112.0[t]	107.1[f], 104.0[u]
MeCONHMe	128.4	126.9[g]	Uracil	113.1[t]	108.4[f], 104.5[u]
$MeCONMe_2$	130.6	129.5[g]	Cytosine	166.3[t]	159.0[f]
Benzene	67.6	72.0[k]	Guanine	143.3[t]	139.7[f]

[a]: Abbreviations: Me = $-CH_3$, Et = $-C_2H_5$, n-Pr = $-C_3H_8$, i-Pr = $-(CH_3)_2CH$, n-Bu= $-C_4H_{11}$, i-Bu = $-(CH_3)_2CHCH_2$, t-Bu= $-(CH_3)_3C$, Ph= $-C_6H_5$. [b]: This work , theoretical EP(K^+) affinities at 298K (ΔH_{298}). [c]: Previously reported PCA at 298K from literature. For cases where only 0K values are reported, thermal corrections to 298K using HF/6-31G(d) geometries and frequencies are applied. [d]: ref. [Ikuta, 1984]. [e]: ref. [Magnusson, 1994]. [f]: ref. [Russo et al., 2001b]. [g]: ref. [Siu et al., 2001a]. [h]: ref. [Abirami et al., 2002]. [i]: ref. [More et al., 1997]. [j]: ref. [Iceman and Armentrout, 2003]. [k]: ref. [Amicangelo and Armentrout, 2000]. [l]: ref. [Amunugama and Rodgers, 2002]. [m]: ref. [Rodgers, 2001]. [n]: ref. [Amunugama and Rodgers, 2000]. [o]: ref. [Huang and Rodgers, 2002]. [p]: ref. [Ryzhov and Dunbar, 1999]. [q]: ref. [Rodgers and Armentrout, 1999]. [r]: ref. [Hoyau and Ohanessian, 1998]. [s]: ref. [Dunbar, 2000]. [t]: the PCA is estimated for the most stable tautomer of the free ligands, corresponding to species A1, T1, U1, C1 and G1 in ref. [Russo et al., 2001b]. [u]: ref. [Rodgers and Armentrout, 2000b].

3.2.2 Accuracy of EP(K^+) Protocol for K^+ Affinities

Before we discuss the PCAs of individual class of ligands, we would like to comment on our choice of B3-LYP over B3-P86 density functional. In order to obtain the energetics at the B3-P86/6-311+G(3df,2p)//B3-LYP/6-31G(d) level for comparison, we replaced the single point energy calculations (step (3) of the EP(K^+) protocol in Chapter 2) with the B3–P86 functionals, and the results are shown in the Appendix I, Table S-3.1. We found that the affinities obtained by EP(K^+) and B3-P86 functional are both in good agreement with existing experimental affinities, with MAD of 4.5 and 4.3 kJ mol^{-1}, respectively. The EP(K^+) affinities tend to be slightly larger in most cases, except when the K^+ interacts with π-type ligands, e.g., C_2H_2, C_2H_4, benzene and borazine, etc. The largest difference is found in glycerol where the EP(K^+) affinity is larger than the B3-P86 affinities by 6 kJ mol^{-1}. For species which show relatively large difference in EP(K^+) and B3-P86 affinities, we carried out additional calculations at the G2(MP2,SVP) level for benchmarking (Table S-3.1). We found that the EP(K^+) affinities are marginally more comparable to this benchmark level with a MAD of 3.2 kJ mol^{-1} (as opposed to 4.1 kJ mol^{-1} for B3-P86 values). As the performance of both functionals does not differ significantly, the B3–LYP functional was chosen for obtaining PCAs as it is more widely used, thus possibly allowing a more direct comparison with other theoretical studies.

For a representative subset of 14 ligands shown in the Appendix I, Table S-3.2, we further explored the effects of employing B3-LYP geometries and vibrational frequencies on zero-point vibrational energy and thermal correction to 298K, and the magnitude of basis set superposition error. Using the B3-LYP frequencies to correct for zero-point energies tend to decrease the theoretical PCA by ~0.5 kJ mol^{-1}, as

compared to using the HF frequencies. The thermal correction at 298K with B3-LYP parameters leads to a further decrease of ~0.3 kJ mol^{-1}. We found that the estimated BSSE obtained by the density functional-based protocol is small (average of 0.7 kJ mol^{-1}), but can be as large as 1.8 kJ mol^{-1} when the cation binds to ligand via aromatic-π interactions. For this subset of ligands, seven experimental PCAs (CO, H_2O, NH_3, benzene, pyridine, pyrrole and uracil) are available for comparison. For these seven species, inclusion of the above refinement in fact leads to a slight increase of MAD (6.4 kJ mol^{-1}, as opposed to 6.0 kJ mol^{-1} with our proposed EP(K^+) protocol). Therefore, the corrections may not lead to better agreement with experimental data. Given that the magnitude of these corrections (~1-2 kJ mol^{-1}) has minimal effect on calculated PCAs, it appears that they could be omitted for computational efficiency.

3.2.3 Geometries and Potassium Cation Affinities of 20 Classes of Ligands

Rare Gas Atoms

The strength of K^+ binding to the rare gas atoms is at the lowest end of the PCA scale. While the experimental K^+ affinities for He and Ne are not known, the Ar affinity [Walter et al., 1998] has been determined to be 14(7) kJ mol^{-1} at 0K (experimental uncertainty in parenthesis). Our calculated K^+ affinity for Ar (8 kJ mol^{-1} at 0K) is within the error limit of this experimental value.

As the potassium cation is isoelectronic with the argon atom, it is of interest to compare the PCAs of the rare gas atoms with the bond dissociation energy of the corresponding isoelectronic rare gas dimer of HeAr, NeAr and Ar_2. The bond dissociation energies for the dimers are very small (less than 1 kJ mol^{-1}) as the rare

gas atoms are held together by weak dispersion forces. [Ma et al., 1993] Comparatively, the interaction between K^+ and rare gas atoms are much stronger, ranging from 4 to 9 kJ mol^{-1}, reflecting the strength of the induction forces when a cation interacts with the polarizable rare gas atoms.

Carbon Monoxide

In agreement with previous findings on Li^+, [Ikuta, 1984; DelBene, 1996], Na^+ [Ikuta, 1984; Hoyau et al., 1999] and K^+, [Ikuta, 1984] the potassium cation prefers to bind to the carbon in CO (estimated K^+...C distance of 3.04 Å), forming a linear cation-ligand complex. The preferential binding of K^+ to carbon can be explained by the fact that even though oxygen is more electronegative than carbon, the negative end of the dipole moment of CO in fact resides on the carbon atom. [Ernzerhof et al., 1993; Schautz and Flad, 1999] Thus, our result highlights the importance of ion-dipole interaction in this complex.

The EP(K^+) calculated PCA is approximately 12 kJ mol^{-1} smaller than that reported previously by Ikuta, [Ikuta, 1984] and the discrepancy arises presumably from the small basis sets employed in the previous study. We note in passing that although the experimental PCA of CO at 298K (19(5) kJ mol^{-1}) [Walter et al., 1998] is closer to our theoretical estimate for the K^+...OC (18 kJ mol^{-1}) than for the K^+...CO (25 kJ mol^{-1}) binding mode, both theoretical values are within the error bar of the experimental measurement.

Hydrogen Fluoride and Hydrogen Chloride

K^+ prefers to bind to the halogen atoms in hydrogen halides. While K^+-HF is linear, the K^+-HCl complex is bent (with $\angle K^+$...Cl–H of 119°). Sodiated hydrogen chloride

(Na^+-HCl) is in a similar shape as the K^+-HCl complex, in which the $\angle Na^+...Cl–H$ is estimated to be 114°. [McMahon and Ohanessian, 2000] Hence, for both Na^+ and K^+, the cation does not bind along the dipole moment vector of the HCl, reflecting a certain degree of covalency when the alkali metal cation binds to second row elements in the periodic table. [McMahon and Ohanessian, 2000]

Phosphorus and Phosphine

Theoretical study on Li^+-P_4 suggested that Li^+ binds to P_4 in a monodentate fashion, along a 3-fold axis of the tetrahedron (species **2**, in ref. [Abboud et al., 2000]). We found that the larger K^+ prefers to bind bidentately along the edge of the P_4 tetrahedron, with a $K^+...P$ distance of 3.51Å (C_{2v}, species **a**, Fig. 3.2) and a PCA of 34 kJ mol^{-1}. The monodentate complex (C_{3v}, species **b**, Fig. 3.2) and the tridentate face-bound K^+-P_4 complex (C_{3v}, species **c**, Fig. 3.2) are approximately 4 kJ mol^{-1} less stable.

The PCA of phosphine (PH_3) is slightly larger than that of P_4. The cation binds to phosphine along the C_{3v} axis of the ligand, with a $K^+...P$ distance of 3.37Å. Our estimate of K^+-PH_3 affinity at 0K (41 kJ mol^{-1}) is in excellent agreement with a previous theoretical value of 40 kJ mol^{-1}. [Magnusson, 1994]

Ethene and Ethyne

Hoyau et al. [Hoyau et al., 1999] showed that Na^+ binds to the π bond of C_2H_2 and C_2H_4 at an average distance of 2.65Å above the plane of the ligand. For K^+, we found that cation-π binding distance is longer at 3.13Å. The increase in bonding distance leads to a decrease in cation affinities by approximately 16 kJ mol^{-1}.

Hydrogen Sulfide and Thiols

For these sulfur-containing ligands, the cation binds to the sulfur at approximately 3.20 Å. The cation has a tendency to bind to one of the lone pairs of the sulfur, and hence, align rather poorly (~40°) with the molecular dipole moment of the ligand.

The estimated PCAs for this class of ligand studied here range from approximately 40 to 60 kJ mol^{-1}. Our theoretical K^+ affinity of H_2S at 0K (38 kJ mol^{-1}) is 8 kJ mol^{-1} too large compared to a previously reported value by Magnusson. [Magnusson, 1994] This discrepancy may be attributed to two factors: no zero point energy correction was made and the affinity was calculated with rather small basis sets in that paper. However, no experimental PCAs for this class of species is available for comparison

Water and Alcohols

For this class of ligands, the K^+ binds to the oxygen atom at a distance of approximately 2.58Å. For K^+-H_2O, the cation binds along the 2-fold axis of the ligand, in perfect alignment with the dipole moment of the ligand. This is in contrast with the preferred mode of K^+ binding in H_2S discussed in the previous section, again reflecting the presence and influence of covalency when the cation binds to second row atoms.

Two sets of experimental PCA for H_2O have been reported by Kebarle and co-workers using the HPMS technique. Our present estimate of 70 kJ mol^{-1} is within ±5 kJ mol^{-1} of both experimental values: 70.7 kJ mol^{-1} [Davidson and Kebarle, 1976a] and 74.9 kJ mol^{-1} [Dzidic and Kebarle, 1970; Sunner et al., 1981]. Three theoretical PCAs for H_2O are available for comparison. [Magnusson, 1994; Russo et al., 2001b; Siu et al., 2001a] The EP(K^+) estimated PCA is in good agreement with all three

values, and it is virtually identical to that obtained by Russo et al. [Russo et al., 2001b] at the B3-LYP/6-311+G(2df,2p)//B3-LYP/6-311+G(2df,2p) level, corrected with BSSE.

For simple alcohols, the cation is in reasonable alignment (~17°) with the dipole moment of the ligands. Previously studies [Rodgers and Armentrout, 1997; Siu et al., 1998] found that when n–BuOH binds to Li^+, the alkyl chain wraps around so that the terminal carbon atom of the ligand forms a secondary interaction with the cation. In the case of K^+, apparently the additional favorable interaction cannot compensate for the unfavorable ligand deformation so that this cyclic form is slightly less stable than the extended open form (by 5 kJ mol^{-1}). [Siu et al., 2001a]

Direct experimental determination of the absolute PCA of simple alcohols has not been reported. However, the relative enthalpy change when water is exchanged by methanol has been determined to be 8.8 kJ mol^{-1}, [Nielsen et al., 1999] which is in reasonable agreement with the corresponding relative affinity values estimated by various theoretical protocols (i.e., EP(K^+) in Table 3.1; B3-P86, and G2(MP2,SVP) in the Appendix I, Table S-3.1), in the range of 4.3 to 6.4 kJ mol^{-1}. We note in passing that using the two experimental PCAs of water and the experimental relative enthalpy change value of 8.8 kJ mol^{-1}, [Nielsen et al., 1999] an absolute PCA for methanol can be estimated to be either 79.5 [Davidson and Kebarle, 1976a] or 83.7 kJ mol^{-1} [Dzidic and Kebarle, 1970; Sunner et al., 1981]. All the theoretical protocols employed here (EP(K^+), B3-P86, G2(MP2,SVP)) yield PCA values that are more consistent with the lower estimate.

Compared with simple aliphatic alcohols, K^+ interactions with polyhydroxyl ligands are less studied. Our results suggested that the K^+ binds bidentately to ethylene glycol

(1,2-ethanediol), 1,2-propanediol and 1,3-propanediol, while interacting in a tridentate fashion with glycerol (1,2,3-propanetriol). Comparing with the monodentate bindings in simple alcohols, bidentate K^+ interactions with the diols increase the PCA by ~30 kJ mol^{-1}. However, the transition from bidentate (in diols) to tridentate (in glycerol) binding only leads to a further enhancement of PCA by ~15 kJ mol^{-1}. This suggests that the destabilizing effect of ligand deformation plays an important role in determining the PCA of multidentate complexes. [More et al., 1997; Rodgers and Armentrout, 2000a; Abirami et al., 2002]

Ethers and Dioxane

For this class of ligands, K^+ binds to the oxygen atom with a typical K^+...O distance of 2.65Å, slightly longer than that found in K^+ bound complexes of aliphatic alcohols, and hence a smaller PCA as compared to the corresponding alcohol analogue.

Dioxanes ($C_4O_2H_8$), which are well-known carcinogens, can be considered as cyclic ethers. For 1,4-dioxane, the cation binds monodentately to one of the oxygen atoms in the ligand (C_s, species **d**, Fig. 3.2). In K^+-1,2-dioxane (C_1, species **e**, Fig. 3.2) and K^+-1,3-dioxane (C_s, species **f**, Fig. 3.2), the cation is bidentately co–ordinated to the two closely situated oxygen atoms. As a result, the PCA of 1,4–dioxane is lower than that of 1,2- and 1,3–dioxane.

Experimental affinities are available for some ethers. The 298K PCA for Me_2O and $(MeOCH_2)_2$ have been determined by Armentrout and co-workers using the threshold-CID method to be 74(4) and 120(4) kJ mol^{-1}, respectively. [More et al., 1997] Both values are in good agreement (within ±4 kJ mol^{-1}) with our theoretical estimates. In

contrast, an older reported PCA of Me_2O [Davidson and Kebarle, 1976a] determined by HPMS (87 kJ mol^{-1}) differs more from our theoretical estimate of 74.9 kJ mol^{-1}.

Aldehydes and Ketones

For this class of ligand, K^+ binds to the carbonyl oxygen (O=C). Compared to the alcohols and ethers, the K^+...O distance is shorter at 2.56Å, with a typical $\angle K^+$...O=C angle of 165°. In general, the PCA of a ketone is larger than that of the corresponding aldehydes and alcohol, for the same number of carbon. Two experimental values for the PCA of Me_2CO were reported, [Sunner and Kebarle, 1984; Klassen et al., 1996] and both are within ±4.5 kJ mol^{-1} of our calculated value (105 kJ mol^{-1}).

Ammonia and Amines

K^+ binds to the electronegative nitrogen atom in these ligands, with a typical K^+...N distance of 2.78Å. A few theoretical PCAs are available for comparison for NH_3. [Magnusson, 1994; Russo et al., 2001b; Siu et al., 2001a; Iceman and Armentrout, 2003] The EP(K^+) estimated PCA is in good agreement with all the reported values based on all electron basis sets calculations. In comparison, the two values reported in ref. [Iceman and Armentrout, 2003] using pseudopotential are too low by over 10 kJ mol^{-1}. Our EP(K^+) PCA is virtually identical to that obtained at the B3-LYP/6-311+G(2d,2p)//B3-LYP/6-311+G(2d,2p) [Russo et al., 2001b] and B3-LYP/6-311+G(2d,2p)//B3-LYP/6-31G(d) levels, [Iceman and Armentrout, 2003] indicating that if sufficiently large basis set are used, the effects of geometry and zero point corrections on K^+ binding affinities would be minimal. It is also interesting to note that for K^+-$(NH_3)_n$ complexes (where n = 1-5), the BSSE at the B3-LYP level is small (within ±1 kJ mol^{-1}) when compared to the BSSE obtained by MP2 calculations (3-4

kJ mol^{-1}) using the same basis set. [Iceman and Armentrout, 2003] The rather small BSSE corrections found for these systems at the B3-LYP level is in agreement with our general findings presented in the Appendix I, Table S-3.2.

Experimental affinities are available for ammonia and four alkylamines. [Davidson and Kebarle, 1976a; Castleman, 1978; Iceman and Armentrout, 2003] While the experimental PCAs are in good general agreement with our theoretical estimates (within ±10 kJ mol^{-1}), qualitative differences are found on the order of relative affinities upon successive methyl substitution of ammonia (refer to section 3.2.5 on "Effect of Substituents" below for further discussion).

Hydrogen Cyanide and Alkyl Nitriles

For HCN and the six alkyl nitriles studied here (including $PhCH_2CN$), K^+ prefers to bind to the nitrogen atom of the ligand. The average K^+...N distance is 2.68Å, slightly shorter than that found in potassiated amines complexes. This reflects that the K^+ interaction with the *sp* hybridized nitrogen in nitriles is stronger than the sp^3 nitrogen in amines. The 298K experimental PCA of MeCN [Davidson and Kebarle, 1976b] at 102 kJ mol^{-1} is in excellent agreement with our calculated value.

Carboxylic Acids and Esters

Two potential K^+ sites of binding are available for this class of ligands (two carboxylic acids and eight esters), namely, carbonyl oxygen and hydroxyl/alkoxyl oxygen. We found that the K^+ prefers to bind to the carbonyl oxygen (O=C), with an average K^+...O distance of 2.51Å, $\angle K^+$...O=C of 165°, and adopting a "trans" conformation for the K^+...O=C–O. In this class of ligand, the K^+ is in fairly good alignment with the molecular dipole moment of the ligand (angle of deviation ~10°).

In general, the PCA of a carboxylic acid is larger than that of the corresponding alcohol, e.g., the PCA of methanoic acid is 5 kJ mol^{-1} above that of methanol. As we are not aware of any experimental and theoretical values in the literature, the PCAs reported here is the first set of estimates available for this class of ligands.

Sulfoxides

For the three sulfoxides studied here, including PhSOMe, K^+ binds exclusively to the oxygen atom in the ligand, with an average K^+...O distance of 2.50Å. It is interesting to compare the structure of the K^+-SO_2 complex with the Na^+–SO_2 complex reported previously. [Hoyau et al., 1999] Ohanessian and co-workers found that Na^+ binds to the SO_2 in a bidentate fashion (with C_{2v} symmetry). For the larger K^+, we found that the cation prefers to bind monodentately to one of the oxygen atoms (C_s, species **g**, Fig. 3.2), and this mode of binding is about 4 kJ mol^{-1} more stable than the bidentate C_{2v} mode (species **h**, Fig. 3.2).

Two experimental values for the PCA of Me_2SO: 146 (HPMS) [Sunner and Kebarle, 1984] and 130 kJ mol^{-1} (threshold-CID) [Klassen et al., 1996] have been reported by Kebarle and co-workers. However, it was pointed out that the HPMS value was based on the slope of a van't Hoff plot covering only a narrow temperature range, and hence subjected to greater experimental error. Thus, the threshold-CID value is considered to be more reliable. [Klassen et al., 1996] Our calculated PCA (128 kJ mol^{-1}) at 298K is in very good agreement with the threshold–CID value of 130 kJ mol^{-1}, and provides further evidence in support of the more recent value.

Amides

Amides are perceived as the entry point for understanding the peptide linkage in protein structures, and thus the K^+-amides interactions are of special biological interest. For the amide complexes studied here, the K^+ binds monodentately to the oxygen atom (K^+...O of 2.45Å, $\angle K^+$...O=C of 166°), in good alignment with the molecular dipole moment of the amides. [Klassen et al., 1996]

Kebarle and co-workers reported the PCAs of four (out of the six) amides studied here, using two different experimental techniques (HPMS [Sunner and Kebarle, 1984] and threshold-CID [Klassen et al., 1996]). The reported 298K PCA of N,N'-dimethylacetamide ($MeCONMe_2$) deserves special attention. For this species, the PCA determined by threshold-CID [Klassen et al., 1996] (121 kJ mol^{-1}) is noticeably lower than that determined by HPMS method (130 kJ mol^{-1}) [Sunner and Kebarle, 1984] More importantly, the affinity determined by threshold-CID is even lower than that of N,N'-dimethylformamide ($HCONMe_2$, 123 kJ mol^{-1}) and acetamide ($MeCONH_2$, 124 kJ mol^{-1}). [Klassen et al., 1996] This is quite surprising as amongst these three amides, the more polarizable $MeCONMe_2$ is expected to have the highest K^+ affinity. In order to resolve this discrepancy, we recently re–measured the PCA of $MeCONMe_2$ using the kinetic method. Anchoring to the *ab initio* G2(MP2,SVP)-ASC(GCP) theoretical value of $HCONMe_2$ corresponding to 125 kJ mol^{-1} at 298K [Siu, 2001a] (which is consistent with the threshold-CID value [Klassen et al., 1996] of 123 kJ mol^{-1}), we obtained a PCA of $MeCONMe_2$ at 131 kJ mol^{-1}. [Tsang et al., 2001a] This indicates that, in contrast to what is found in the case of K^+–Me_2SO (discussed in the 'sulfoxides' section above), the K^+-$MeCONMe_2$ affinity determined by the HPMS method [Sunner and Kebarle, 1984] is more consistent with available theoretical and experimental values. Excluding the threshold-CID value of

$MeCONMe_2$, it is pleasing to note that our protocol yields affinities to within ±3 kJ mol^{-1} for all the six amides studied here.

Benzene, Borazine and Phenol

The interaction between cations and aromatic π electrons is a relatively new type of electrostatic interaction. This so-called cation-π interaction is implicated in many important biological functions, [Mecozzi et al., 1996; Dougherty, 1996; Ma and Dougherty, 1997; Nakamura et al., 1997; Dougherty and Lester, 1998; Zacharias and Dougherty, 2002] and benzene is often used as the prototype ligand for understanding this interaction. The most recent experimental dissociation energies of K^+-benzene (73 kJ mol^{-1} at 0K), determined by Amicangelo and Armentrout, [Amicangelo and Armentrout, 2000] is in good agreement with our present estimated value of 67 kJ mol^{-1} at 0K. We note in passing that our 298K PCA reported in Table 3.1 (67.6 kJ mol^{-1}) is obtained from the EP(K^+) PCA value at 0K (67.1 kJ mol^{-1}), with thermal corrections using HF/6-31G(d) geometry and vibrational frequencies. The experimental PCA (74.2 kJ mol^{-1}, in Table 3.1) is taken directly from ref. [Amicangelo and Armentrout, 2000], with thermal correction at the MP2(full)/6-31G(d) level. The apparent inconsistency (0.7 kJ mol^{-1}) between the two sets of PCA values at 0K and 298K probably arises from the level of theory employed for the thermochemical correction.

Even though borazine ($B_3N_3H_6$) is isoelectronic with benzene, its electronic nature has been controversial. Criteria based on magnetic properties, energetic and geometric indices have suggested that borazine may not be aromatic. [Schleyer et al., 1995; Katritzky et al., 1998] However, recent findings showed that borazine should be aromatic, though its aromaticity is about half of that of benzene. [Kiran et al., 2001]

The most stable structure we obtained for borazine possesses the D_{3h} symmetry, and is in agreement with previous experimental and theoretical studies. [Harshbarger et al., 1969; Parker and Davis, 1997; Chiavarino et al., 1999]

We located two stable minima on the K^+-borazine potential energy surface. The less stable complex is planar, with C_{2v} symmetry (species **i**, Fig.3.2). The more stable complex possesses C_{3v} symmetry (species **j**, Fig.3.2), with the cation situated at 2.88Å above the ring centroid and K^+...N and K^+...B distances of 3.22 and 3.33Å, respectively.

While the structural features of the most stable K^+-borazine complex (species **j**, Fig. 3.2) is quite comparable to that obtained for K^+-benzene (C_{6v}, species **k**, Fig. 3.2), the estimated PCA of borazine is 21 kJ mol^{-1} less than that for benzene (68 kJ mol^{-1}).

Phenol represents a prototypical case for the competition of π and non-π hydroxyl oxygen binding site for K^+. For the Na^+-phenol complexes, [Hoyau et al., 1999; Dunbar, 2002b] the binding affinities at 298K for the aromatic-π and non-π complexes are comparable, and differ by ~1–4 kJ mol^{-1} depending on the computational protocol used to obtain the Na^+ affinities. However, the free energy of binding, ΔG_{298}, indicates that the non–π complex is the favored form of Na^+-phenol complex in the gas phase. [Hoyau et al., 1999]

For K^+-phenol, two stable minima are again found. In the more stable form (in terms of ΔH_{298}), the K^+ is bound to the aromatic-π ring (species **l**, Fig. 3.2), while the less stable (by 3 kJ mol^{-1}) non-π mode involves a K^+...O interaction of 2.58Å (species **m**, Fig. 3.2). Similar to Na^+–phenol system, the difference in EP(K^+) free energy of binding (ΔG_{298}) suggests that the non–π complex is favored over the π mode (by about 4 kJ mol^{-1}) in K^+-phenol.

The EP(K^+) PCA we obtained for phenol is 70 kJ mol^{-1}. Two experimental values are available for comparison: a recent threshold-CID value of 75 kJ mol^{-1}, [Amunugama and Rodgers, 2002] and an earlier value of 84 kJ mol^{-1} based on results of radiative association kinetics measurements and density functional calculations. [Ryzhov and Dunbar, 1999] Here, we note that the latter value is anchored to the K^+–benzene affinity (at 298K) of 80 kJ mol^{-1}. [Keesee and Castleman, 1986] However, recent experiment [Amicangelo and Armentrout, 2000] suggested that the PCA at 298K for benzene might need to be lowered to 74 kJ mol^{-1}. Adopting this lower PCA value of benzene as the anchoring point, a revised PCA value of 78 kJ mol^{-1} is obtained (corrected to 298K). With this downward revision of the PCA of K^+-phenol from ref. [Ryzhov and Dunbar, 1999], the two experimental values and our theoretical EP(K^+) affinity are now much more consistent.

Azines

Pyridine (C_5NH_5), pyridazine, pyrimidine, pyrazine (isomers of $C_4N_2H_4$) and 1,3,5-triazine ($C_3N_3H_3$) are six-membered ring nitrogen heterocycles. The presence of nitrogen atom(s) disturbs the symmetry of the π–electrons distribution, creating a localization of charge on the nitrogen atom, thus decreasing the resonance stabilization and aromatic character of the molecule. [Amunugama and Rodgers, 2000] Hence, for this class of ligands, the K^+ prefers to bind to the nitrogen lone pair, with an average K^+...N distance of 2.75Å, rather than to the π cloud. Our present relative and absolute PCAs are in very good agreement with a combined experimental and theoretical study: [Amunugama and Rodgers, 2000] the MAD with the reported experimental and BSSE corrected MP2(full)/6-311+G(2d,2p)//MP2(full)/6-31G(d) values are 2.5 (Table 3.1) and 1.8 kJ mol^{-1} (Table 3.2), respectively.

The modes of cation binding in 1,8-naphthyridine ($C_8N_2H_6$) and its PCA have not been reported previously. Our model (C_{2v}, species **n**, Fig. 3.2) suggests that K^+ binds bidentately to the two nitrogen atoms of the ligand. Because of the combined effect of polarizability and multidentate interaction, the PCA of 1,8-naphthyridine (at 156 kJ mol^{-1}) is much higher than the other pyridines, and it is in fact at the top end of the PCA scale presented here.

Pyrrole and Indole

Pyrrole (C_4NH_5) and indole (C_8NH_7) are important models for understanding of the cation-π interactions in tryptophan containing proteins. Unlike pyridine, the nitrogen atom in pyrrole is electron-deficient, [Fessenden and Fessenden, 1998] thus, the pyrrole nitrogen is not a favorable site for cation binding. As in the case of Na^+-pyrrole, the potassium cation prefers to bind to the π ring (K^+-π distance of approximately 2.85Å) of the ligand.

The experimental PCA for pyrrole has recently been determined to be 85 kJ mol^{-1}, [Huang and Rodgers, 2002] in good agreement with our theoretical PCA (77 kJ mol^{-1}). This is approximately 10 kJ mol^{-1} larger than the theoretical estimate for benzene, and can be attributed to the larger quadrupole moment and dipole moment of pyrrole. [Doerksen and Thakkar, 1999]

In the Na^+-indole complex, the cation may bind to either the benzo-π or pyrrolo–π face of the ligand. [Dunbar, 1998] For the larger potassium cation, only one type of K^+-indole complex is found where the cation binds to the benzo-π face, with an estimated PCA of 89 kJ mol^{-1} (species **o**, Fig. 3.2). Thus, even though the PCA of pyrrole is larger than that of benzene, it appears that the K^+ binds exclusively to the benzo-π face of indole. This suggests that the distribution of π electrons is

sufficiently altered in the polycyclic aromatic ligands so that indole should not be considered as a "fused" ring of benzene and pyrrole. [Mecozzi et al., 1996; Ryzhov and Dunbar, 1999]

The experimental PCA of indole at 0K was reported in the same study of K^+–phenol by Ryzhov and Dunbar at 105 kJ mol^{-1}. [Ryzhov and Dunbar, 1999] In view of our discussions in the "benzene, borazine and phenol" section, this experimental PCA of indole might need to be lowered (by 6 kJ mol^{-1}) because of the anchoring value of benzene. The revised value is 99 kJ mol^{-1} (at 298K), which brings it in closer agreement with our estimated PCA at 298K for indole of 89 kJ mol^{-1}.

Azole

Azoles are building blocks for many antibiotics, anticancer agent and drugs. [Cartledge et al., 1997; 1998; Hopewell, 1997] Even though these five-membered ring heterocycles may be perceived as derivatives of pyrrole, the mode of potassium cation binding is different from that found in pyrrole. For the study of azoles here (except isoxazole), the K^+ binds exclusively to the nitrogen lone pair, with an average K^+...N distance of 2.71Å, comparable to that found in azines. No cation-π complexes were located, i.e., all azoles favor σ binding interaction (except pyrrole). [Rodgers and Armentrout, 1999; Huang and Rodgers, 2002] Two theoretical PCA for imidazole ($C_3N_2H_4$) were reported previously. [Rodgers and Armentrout, 1999; Huang and Rodgers, 2002] Our value is in good agreement with the more recent value, calculated at the MP2(full)/6-311+G(2d,2p)//MP2(full)/6-31G(d) level, corrected with BSSE. [Huang and Rodgers, 2002] The earlier value, [Rodgers and Armentrout, 1999] calculated at HF/6–31G(d,p), was about 6 kJ mol^{-1} too large when compared to the EP(K^+) estimated PCA.

In general, the PCA of azoles increases with increasing number of methyl substituents, [Huang and Rodgers, 2002] and with increasing number of ring nitrogen atoms present, except for tetrazole. [Rodgers and Armentrout, 1999] We have also estimated the PCA of the oxygen- and sulfur-containing azoles (isomers of C_3ONH_3: oxazole, isoxazole; isomers of C_3SNH_3: thiazole and isothiazole), which has not been reported previously. Isoxazole has a higher PCA (by ~10 kJ mol^{-1}) than oxazole as K^+ is bound bidentately to both N and O atoms in isoxazole (species **p**, Fig. 3.2), but solely to the N atom in oxazole (species **q**, Fig. 3.2). On the other hand, thiazole and isothiazole have similar PCA as K^+ binds to the N atom exclusively in both ligands.

Our EP(K^+) PCA values at 0K for 2H-1,2,3-triazole, 1H-1,2,4-triazole and 2H–tetrazole are in good agreement (within ±9 kJ mol^{-1}) with the previously reported experimental threshold-CID PCA values (MAD of 4 kJ mol^{-1}). However, our theoretical PCA for the corresponding 1H-1,2,3-triazole, 4H-1,2,4-triazole and 1H–tetrazole tautomers are significantly greater (and hence more stable K^+ bound complexes) by 20–63 kJ mol^{-1}. [Rodgers and Armentrout, 1999] Our findings here are consistent with the rationalization put forward by Rodgers and Armentrout: in their threshold-CID experiments, the K^+ are *kinetically* favored to bind to the most stable 2H–1,2,3–triazoles, 1H-1,2,4-triazoles and 2H–tetrazoles azole tautomers, even though the resulting K^+ bound complexes are not the most stable K^+ bound species *thermodynamically* (refer to footnote [y] and [z] of Table 3.1). [Rodgers and Armentrout, 1999]

Amino Acids

Because of the presence of the acidic carboxylic group and basic amino group, amino acids can exist in two forms: the charge-solvated (CS) and zwitterionic (ZW) form. [Hoyau and Ohanessian, 1998; Marino et al., 2000; Pulkkine et al., 2000; Strittmatter et al., 2000; Wyttenbach et al., 2000] The ZW form of amino acids is dominant in solution, [Wada et al., 1982] and can be stabilized in the gas phase by binding to cations. [Strittmatter et al., 2000] Similar to the case of glycine and alanine reported earlier, [Wyttenbach et al., 2000] our calculations show that for the other aliphatic amino acids like valine, leucine, and isoleucine, the potassium cation prefers to bind bidentately to the carbonyl and hydroxyl oxygen atoms, with the ligands in the charge-solvated CS form.

Comparing our EP(K^+) PCA of glycine with that estimated by Hoyau and Ohanessian, [Hoyau and Ohanessian, 1998] their value is too small by almost 8 kJ mol^{-1}; the difference most likely is due to the different basis sets employed. We are not aware of theoretical and experimental PCAs for the larger aliphatic amino acids in the literature. Using our theoretical G2(MP2,SVP)-ASC(G-CP) PCA at 0K of acetamide (118.7 kJ mol^{-1}), N–methylacetamide (125.6 kJ mol^{-1}) and N,N'-dimethylacetamide (129.2 kJ mol^{-1}) as anchoring reference values, [Tsang et al., 2004] we have obtained the experimental PCA of aliphatic amino acids by the kinetic method as shown in Table 3.1 (refer also to footnote [a'] of Table 3.1). The relative and absolute EP(K^+) PCAs of all five aliphatic amino acids were found to be in very good agreement with the quantitative values determined by the mass spectrometric kinetic method in the order: [Tsang et al, 2001b]

glycine < alanine < valine ~ leucine ~ isoleucine

, the absolute PCAs are also within ±1 kJ mol^{-1} of that determined experimentally. [Tsang et al., 2001b] We note in passing the anchoring value of 129.2 kJ mol^{-1} we used for N,N'–dimethylacetamide is very close to the experimental HPMS value of 129.5 kJ mol^{-1} reported by Kebarle and co-workers. [Sunner and Kebarle, 1984] Thus, the experimental set of PCAs for aliphatic amino acids is not only consistent with our EP(K^+) estimates, but it is also consistent with the reported experimental affinity of N,N'-dimethylacetamide.

We have also obtained the PCA of proline, serine, cysteine and phenylalanine (species **r, s, t, u**, Fig. 3.2, respectively). The mode of K^+ binding for these ligands are similar to that found in the corresponding Na^+ complexes. [Hoyau et al., 1999; Siu et al., 2001b] No experimental or theoretical PCA is available in the literature, except for K^+–phenylalanine. The theoretical PCA reported by Ryzhov et al. (Table 3.2, 146.4 kJ mol^{-1}) is very close to our EP(K^+) PCA of 145.6 kJ mol^{-1}. [Ryzhov et al., 2000] Using the experimental PCA of adenine (106 kJ mol^{-1}), cytosine (110 kJ mol^{-1}) and guanine (117 kJ mol^{-1}) as reference values (Table 3.1), [Cerda and Wesdemiotis, 1996] Ryzhov et al. reported a kinetic method value of 104.2 kJ mol^{-1} for the PCA of phenylalanine; [Ryzhov et al., 2000] which is even lower than the threshold-CID value of K^+–glycine (126 kJ mol^{-1} at 298K) reported by Kebarle and co-workers. [Klassen et al., 1996] This is counter-intuitive because phenylalanine is expected to have greater PCA than glycine due to its greater polarizability and ion-induced dipole interactions. Hence, it is very likely that the kinetic method value for phenylalanine might involve a relatively large margin of experimental uncertainty greater than 20.9 kJ mol^{-1} (5 kcal mol^{-1}) as estimated by Ryzhov et al. (refer to the section on "nucleobases" below for further discussions on the reference values). We have

conducted detailed theoretical and experimental studies on the K^+-phenylalanine system, and our findings will be reported elsewhere.

Nucleobases

Given the biological importance of these ligands as models for cation–RNA/DNA interactions, several experimental [Cerda and Wesdemiotis, 1996; Rodgers and Armentrout, 2000b] and theoretical works [Burda et al., 1996; Rodgers and Armentrout, 2000b; Russo et al., 2001a; 2001b] have been reported on the gas-phase alkali cations affinities of the five nucleobases adenine (A), thymine (T), uracil (U), cytosine (C) and guanine (G).

Experimental PCAs for all five nucleobases (A, T, U, C, G) have been obtained by Cerda and Wesdemiotis [Cerda and Wesdemiotis, 1996] using the extended mass spectrometric kinetic methods, and the PCA of adenine/thymine/uracil (A, T, U) has been determined by Rodgers and Armentrout [Rodgers and Armentrout, 2000b] using the threshold-CID method. There have been some concerns regarding the accuracy of the PCA reported by Cerda and Wesdemiotis. [Cerda and Wesdemiotis, 1996] Firstly, Rodgers and Armentrout, [Rodgers and Armentrout, 2000b; Armentrout and Rodgers, 2000] noted that the kinetic method measurements were carried out at only two excitation conditions or effective temperatures. As a result, relative enthalpy (affinity) and entropy changes may not be extractable in a statistically meaningful way from the experiment. Secondly, under the experimental condition of the kinetic method measurement, [Cerda and Wesdemiotis, 1996] several tautomeric forms of the free ligand *may* co-exist, leading to different K^+ bound structures and affinities. [Russo et al., 2001b] Depending on the energy barrier of interconversion between these tautomeric forms in the free ligand and the K^+ bound form, the kinetic method

values determined may or may not correspond to the PCA of a particular alkali metal cation-nucleobase complex. [Russo et al., 2001b] Such complication is suggested to be the case for cytosine and guanine. [Russo et al., 2001b] Moreover, it has also been noted that the kinetic method values for adenine could not be easily ascribed to the alkali cation affinity of any one of the tautomers. [Russo et al., 2001b] Nevertheless, on careful examination, we consider the PCA of U and T determined by the kinetic method should be reliable. In the experiment, pyridine, aniline and n-propylamine were used as the reference compounds. As noted by Rodgers and Armentrout, [Rodgers and Armentrout, 2000b] these three reference compounds and U/T all bind to K^+ in a monodentate fashion; thus, entropic effect in the kinetic method measurement would be negligible. Moreover, for these two species, the PCA determined by kinetic method is comparable to that determined by threshold-CID. Taken all the above factors into consideration, we have omitted the kinetic method values for A/C/G, but retaining the ones for U/T in the evaluation of our EP(K^+) protocol. We have estimated the EP(K^+) affinities of the most stable K^+ bound structure of these ligands in the *most stable free tautomeric forms*. When compared to the experimental threshold-CID and selected kinetic method values, [Rodgers and Armentrout, 2000b] our EP(K^+) affinity is on average about 9.6 kJ mol^{-1} too large.

Comparing the EP(K^+) PCA with previously reported *ab initio* MP2 and density functional theory values, [Rodgers and Armentrout, 2000b; Russo et al., 2001b] our present estimate is approximately 7-9 kJ mol^{-1} too large. In the case of the MP2 values, [Rodgers and Armentrout, 2000b; Russo et al., 2001b] the difference can be at least attributed partly to BSSE corrections. As an example, for K^+-uracil, the reported BSSE for the MP2 based model is 4 kJ mol^{-1} [Rodgers and Armentrout, 2000b] which accounts for about 50% of the difference. We note here again that our EP(K^+) model

does not include the BSSE correction, but for this species, the EP(K^+)-BSSE correction is calculated to be small at 0.8 kJ mol^{-1} only (refer to Appendix I, Table S-3.2).

However, the differences between our values and the density function based estimates by Russo et al. [Russo et al., 2001b] could not be easily explained. The two DFT protocols, even though not identical, are expected to be comparable, and indeed we found evidence of support in the case of PCA of water and ammonia (see discussions in the respective Sections above). Using K^+-uracil as an example, we tried to identify the source of discrepancies. We found that the geometries are similar: for this complex, the calculated K^+...O distance at B3-LYP/6-31G(d) level (2.465Å) is only 0.002Å shorter than that obtained at the B3-LYP/6-311+G(2df,2p) level. The zero point corrections are also similar: the B3LYP/6-31G(d) level ZVPE is only 0.4 kJ mol^{-1} larger than that at B3LYP/6-311+G(2df,2p). In order to resolve the differences, we attempted to obtain the PCA of uracil using the model reported in Russo's work (B3-LYP/6-311+G(2df,2p)//B3-LYP/6-311+G(2df,2p) with BSSE correction). A PCA of 113.8 kJ mol^{-1} (without BSSE correction) was obtained, which is very close to our EP(K^+) estimated PCA of 113.1 kJ mol^{-1}. Applying the BSSE correction reduced the estimate from 113.8 to 113.0 kJ mol^{-1}, but still differs from the reported value of Russo (108.1 kJ mol^{-1}, corrected to 298K) by 4.9 kJ mol^{-1}. [Russo et al., 2001b] In conclusion, it appears that DFT based protocols could be over-estimating the PCAs for the nucleobases. Given these uncertainties, further theoretical studies and experimental measurements on guanine, cytosine and adenine are clearly required.

3.2.4 Comparison between PCA with Lithium and Sodium Cation Affinity Scales

Our theoretical PCA obtained using the EP(K^+) protocol were plotted against the reported theoretical Li^+ affinities [Burk et al., 2000] (calculated at the B3-LYP/6-311+G(d,p) level) and Na^+ affinities [Hoyau et al., 1999; Armentrout and Rodgers, 2000; McMahon and Ohanessian, 2000; Amunugama and Rodgers, 2000; Rodgers, 2001; Petrie, 2001] in Fig. 3.3. The theoretical methods used in ref. [Hoyau et al., 1999; Armentrout and Rodgers, 2000; McMahon and Ohanessian, 2000; Amunugama and Rodgers, 2000; Rodgers, 2001] and related studies were all carried out at the MP2(full)/6–311+G(2d,2p)//MP2/6–31G(d) level, and hence, the theoretical Na^+ affinity values from these studies were pooled together as one set and correlated with the EP(K^+) PCAs (denoted as Na^+-1 in Fig. 3.3). However, the theoretical Na^+ affinities reported by Petrie [Petrie, 2001] were calculated at a different level of theory (CPd-G2thaw), [Petrie, 2001] and hence was correlated separately with the EP(K^+) PCAs (denoted as Na^+-2 in Fig. 3.3).

The PCAs correlates linearly with Li^+ and Na^+ affinities:

$$\Delta H_0 (Li^+\text{-}L) = 1.55\ PCA + 40.93,\ R^2 = 0.96 \qquad [3.1]$$

$$\Delta H_0 (Na^+\text{-}L) = 1.16\ PCA + 12.84,\ R^2 = 0.97,\ \text{for } Na^+\text{-1 affinities} \qquad [3.2]$$

$$\Delta H_0 (Na^+\text{-}L) = 1.22\ PCA + 10.48,\ R^2 = 0.97,\ \text{for } Na^+\text{-2 affinities} \qquad [3.3]$$

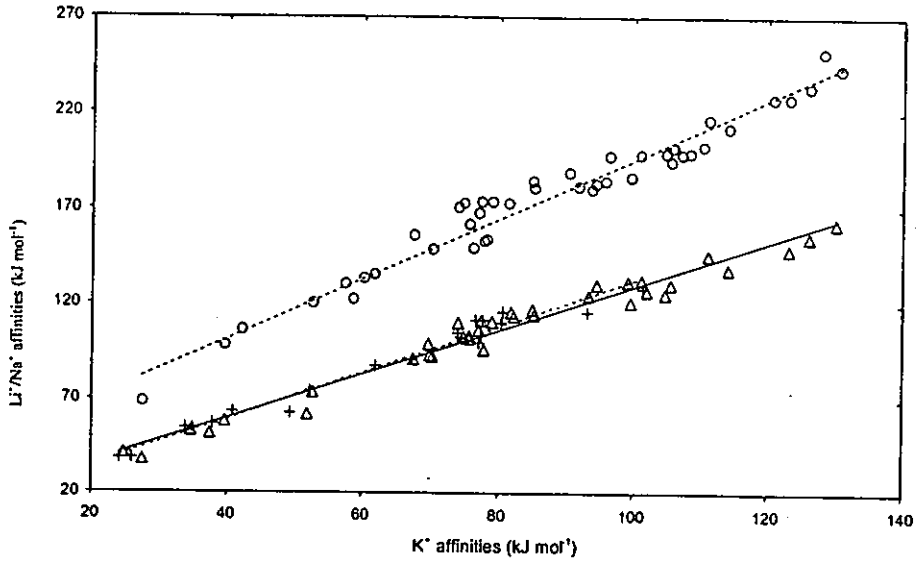


Figure 3.3. The correlation of potassium cation affinities, with (a) theoretical Li^+ affinities reported in ref. [Burk et al., 2000] (298K, indicated by 'o'); (b) theoretical Na^+ affinities reported in ref. [Hoyau et al., 1999; Armentrout and Rodgers, 2000; McMahon and Ohanessian, 2000; Amunugama and Rodgers, 2000; Rodgers, 2001] and related studies (298K, indicated by 'Δ', Na^+-1) and ref. [Petrie, 2001] (0K, indicated by '+', Na^+-2).

Such good linear correlation indicates that the nature of interaction between alkali cations (Li^+, Na^+ and K^+) and the wide range of ligands studied here are indeed very similar. Furthermore, we note that the correlation relation obtained using Na^+-1 (Eqn. [3.2]) and Na^+–2 (Eqn. [3.3]) differs slightly only in the intercept. In other words, different theoretical models are likely to yield the same *relative* affinity ladder, but the *absolute* affinities obtained may be different slightly. This highlights the importance of obtaining a set of *absolute* theoretical PCAs that is consistent with experimental values in order to minimize the presence of systematic errors.

3.2.5 Effect of Substituents

We would now like to comment on the effect of substituents, in terms of how they affect the binding mode and the PCA as compared to the unsubstituted parent ligands.

For substituents with no electronegative atoms (e.g., alkyl group), the cation is not bound to these groups. Hence, the presence of these substituents only affects the PCA, but not the site of binding. For aromatic ligands such as pyridine and azoles, methylation increases the PCA. Compared to *o*- and *m*-substitution, the effect of *p*-substitution on PCA is most significant in pyridine, and has been attributed to the larger dipole moment found in *p*-methylpyridine. [Rodgers, 2001]

Using classical electrostatic theory, Davidson and Kebarle suggested that alkyl substituent affects four properties of a ligand: its polarizability, dispersion, intra–molecular repulsion (between ligand and ion), and dipole moment. [Davidson and Kebarle, 1976c] While successive alkyl substituents increase the binding affinity of ligand by enhancing the polarizability and dispersion component of the cation-ligand interactions, it also increases the repulsion and decreases the dipole moment of the ligand, thus, leading to the decrease of binding affinities. Here, we studied the effect of successive methylation at the O/N/S binding sites on the PCA of water, ammonia, formamide/acetamide and hydrogen sulfide.

In the H_2O/MeOH/Me_2O series, first methylation increases the PCA by ~5 kJ mol^{-1}, while slight decrease in PCA is found in the second methylation. In the H_2S/MeSH/Me_2S series, the PCA is in the order of H_2S < MeSH < Me_2S. Similar observations are found for the corresponding theoretical sodium cation affinities for both series, [Hoyau et al., 1999; McMahon and Ohanessian, 2000] and have been rationalized in terms of opposing effects of changes in ligand polarizability (increase)

and dipole moment (decrease). [McMahon and Ohanessian, 2000] While similar rationale can be applied in explaining the observed trends in PCAs, we would like to point out that repulsive effect may also play a role here. For the smaller oxygen atom, the repulsive steric effect of successive methylation would be much strongly felt when compared against the larger sulfur atom. Thus, it may not be surprising that while PCA increases with increasing methylation in the H_2S series, it "tails-off" or decreases in the H_2O series.

For the ammonia series, our theoretical results show that the effect of multiple methyl substitutions on the PCA of NH_3 is small, spanning a range of only 5 kJ mol^{-1}. Similar to what is found in the case of the H_2O series, the theoretical PCA increases with first methylation (by 2 kJ mol^{-1} from NH_3 to $MeNH_2$), but decreases on second and third methylations. Interestingly, earlier experimental HPMS results suggest that successive methyl substitution increases the PCA, i.e., $NH_3 < MeNH_2 < Me_2NH < Me_3N$. [Davidson and Kebarle, 1976c] As the difference in PCA in successive methylation is small, and can be considered as within the expected error limits of theoretical models, we only note here that our EP(K^+) affinities are more in line with the recent reported trends for the experimental (FT-ICR) and theoretical (MP2) sodium cation affinities, [McMahon and Ohanessian, 2000] and our B3–P86 affinities (Appendix I, Table S-3.1), yet differ from the prediction at the G2(MP2,SVP) level (Appendix I, Table S-3.1). Thus, further calculations and experimental measurements may be needed to resolve the difference in qualitative trends found between experimental and theoretical PCAs of ammonia and its methyl substituted derivatives.

For electronegative (e.g., -fluoro, -chloro) or electron-rich substituents (e.g., aromatic π-rings), not only are the binding affinities affected, but the presence of these

substituents opens up new modes of binding that is not possible in the parent ligand. We found that the -F and -Cl substituents are in general not competitive against the O, N binding sites already present. Hence, for most classes of ligands (e.g., carboxylic acids, aldehydes, ketones, nitriles, etc.), the modes of binding remain the same as the parent ligand upon halogenation. In general, the PCA of halogenated ligands decreases as these electron withdrawing groups decrease the dipole moment of the ligand. However, in a few cases (e.g., CF_3CH_2OH, species **v**, Fig. 3.2) when the halogen atom is closed to the original site of binding of the parent ligand, it offers additional binding site for K^+. In these cases, the PCA are increased compared to the unsubstituted parent ligands.

For aromatic-π substituents, the PCA increase in all cases as polarizability of the ligand is greatly enhanced by the presence of the highly polarizable phenyl ring. In some cases, the aromatic-π substituents (e.g., from alanine to phenylalanine) are also involved in the binding, leading to further increase in PCA than what is expected solely from polarizability effect.

3.2.6 Relating PCAs to Properties of Ligands

In the previous section, the effect of substituents on PCA is discussed in a qualitative manner. Here, we would like to take the discussion one step further by establishing quantitative relations between the PCA of the 136 ligands and its properties. Our aim is to use molecular properties which are readily available and accessible, so that the PCAs can be predicted with relative ease.

As the K^+-ligand interaction is mainly electrostatic in nature, ligand properties such as dipole moment and polarizability are expected to be important. The number of interactions (co-ordinate number or dentate) and which type of atom(s) the K^+ binds to should also be important. Based on the goodness of fit (in terms of adjusted R^2) from multiple linear regression analysis, four parameters were identified to be important in governing the PCA of a ligand. They are: the dipole moment (μ, in Debye, which are determined at the B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d) level), the polarizability (α, in $Å^3$, which are determined at the HF/6-31G(d)//HF/6-31G(d) level) of the ligand, the number of first row atoms the K^+ interacts with (n_1), and the number of second row atoms the K^+ interacts with (n_2). The PCA is related to these four parameters via:

$$PCA = 10.7\mu + 3.6\alpha + 16.1n_1 - 11.8n_2 + 29.6 \quad (\text{adjusted } R^2 = 0.82) \qquad [3.4]$$

Fig. 3.4 shows the relation between the calculated PCA (Table 3.1) and the predicted PCA using Eqn. [3.4]. It is clear from Fig. 3.4 that several points show large deviations from the ideal correlation line with slope of 1. At the lowest end of the PCA scale, the error could be very large; in the case of Ne, the error could be as large as 12-fold (calculated PCA 4 kJ mol^{-1}, predicted PCA 47 kJ mol^{-1}). Such deviation partly arises from the simplicity of the correlation equation employed, but is also due to the small numerical PCA value for this ligand.

It is clear that our proposed model is quite crude and has neglected the other effects such as ion-local dipole interactions. [Amunugame and Rodgers, 2000] Despite the crudeness of the model, the equation could yield reasonable estimates of PCA. If the four lowest PCA values (He, Ne, Ar and CO) are ignored; the MAD for the remaining

132 ligands is then reduced to 12 kJ mol^{-1} (error of 10%), with a maximum of 38 kJ mol^{-1} for CF_3CH_2OH (error of 52%).

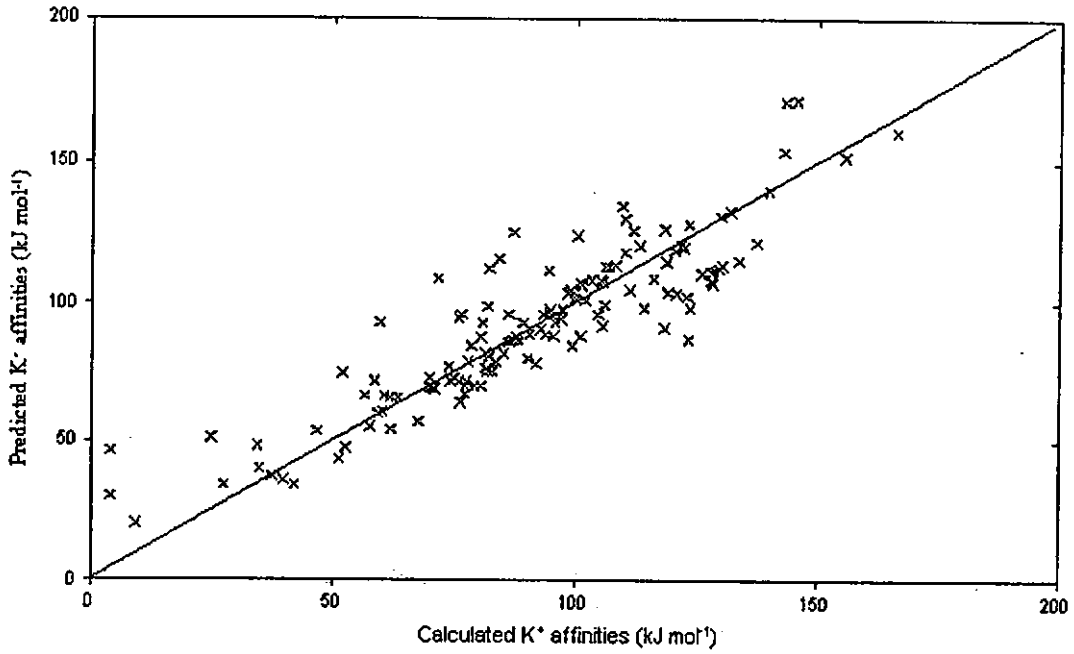


Figure 3.4. Plot of predicted PCAs (Eqn. [3.4]) against EP(K^+) PCAs: the diagonal line with a slope of 1.0 is drawn for reference proposes.

3.3 Conclusion

In this chapter, we reported the theoretical potassium cation affinities (PCA) of 136 ligands, spanning a range from 4 to 166 kJ mol^{-1}. Out of these 136 ligands, 70 experimental and 64 theoretical values reported in the literature are available for comparison. We found that our theoretical estimates and most of the experimental affinities are in very good general agreement (within ±10 kJ mol^{-1}). Based on our theoretical EP(K^+) values, we are able to conduct a critical evaluation of the reported values for Me_2SO, $MeCONMe_2$ and phenol obtained by different experimental techniques, for which PCA differences of over ±10 kJ mol^{-1} have been reported. Large discrepancies (exceeding 26 kJ mol^{-1}) has been found in the case of phenylalanine, [Ryzhov et al., 2000] cytosine, [Cerda and Wesdemiotis, 1996] guanine [Cerda and Wesdemiotis, 1996] and adenine. [Cerda and Wesdemiotis, 1996] However, in all these cases, the discrepancies are likely to arise from complications in the experimental measurements. Ignoring these four values and the PCA of Me_2SO determined by HPMS, the mean-absolute-deviation of our theoretical PCA with the remaining experimental values is 4.5 kJ mol^{-1}. Our EP(K^+) PCA is also consistent with most of the previously reported theoretical values to within ±5 kJ mol^{-1}. For species with larger differences, we are able to account for the difference in terms of different basis sets used in the theoretical calculations and/or basis set superposition errors. However, the rather large difference of 7-9 kJ mol^{-1} found between our values and those reported by Russo et al. [Russo et al., 2001b] for the DNA/RNA nucleobases remains unclear.

The effects of substituents on PCA of parent ligands are also discussed. For a halogenated ligand, the PCA decreases in general, except when the halogen is

proximal to the original binding site. For aromatic-π substituents, the PCA increases in all cases as polarizability of the ligand is greatly enhanced by the presence of the highly polarizable phenyl ring. First methylation tends to increase the PCA, while PCA may decrease upon further methylation.

We have also compared the PCA with previously reported lithium and sodium cation affinities in the literature. Excellent linear correlation is found, indicating that the nature of interaction between alkali cations (Li^+, Na^+ and K^+) and the wide range of ligands studied here are indeed very similar. Thus, such relations (Eqns. [3.1-3.3]) allow estimation of PCA for ligands with known Li^+ and/or Na^+ affinities, in particular in cases where the mode of binding for K^+ is not expected to differ from the smaller Li^+/Na^+.

Finally, we established one correlation equations (Eqn. [3.4]) relating PCAs of ligands with their properties (dipole moment, polarizability, and number of interactions). This equation offers relatively simple and efficient methods of estimating the PCA of ligands not reported here.

Chapter 4 Interaction between K^+ and Aliphatic Amino Acids

4.1 Background

In Chapter 3, the energetic protocol for obtaining potassium cation affinities, $EP(K^+)$ has been validated. While the mean-absolute-deviation (MAD) between the $EP(K^+)$ estimates and available experimental values is 4.5 kJ mol^{-1} for 65 ligands, the protocol performs particular well for the five aliphatic amino acids (within ±1 kJ mol^{-1}). Encouraged by this, we expanded the theoretical study to other modes of binding between K^+ and the aliphatic amino acids.

Glycine is the simplest of the aliphatic amino acids. The -N-C-C-O- motif, which is found in glycine, is the fundamental building block for all amino acids and proteins. [Lehninger, 1982; March, 1985] As an important model compound in biophysics and biochemistry, glycine has been the subject of many experimental and theoretical investigations. [Iijima et al., 1991; Csaszar, 1996; Shirazian and Gronert, 1997; Balta et al., 2000; Matta and Bader, 2000] A number of theoretical studies have been conducted to elucidate the fundamental nature of the interaction between alkali metal cations and simpler aliphatic amino acids (glycine/alanine) in the gas phase. [Jensen, 1992; Hoyau and Ohanessian, 1998; Wyttenbach et al., 1998; Marino et al., 2000; Pulkkinen et al., 2000; Strittmatter et al., 2000; Moision and Armentrout, 2002; Talley et al, 2002; Kish et al., 2003] In these works, the relative stabilities of the charge-solvated (CS) or zwitterionic (ZW) forms of the aliphatic amino acids with different metal cations were studied.

It was found that the smaller alkali metal cation M^+ (Li^+/Na^+) prefers to interact with glycine [Hoyau and Ohanessian, 1998; Marino et al., 2000; Talley et al., 2002; Moision and Armentrout, 2002; Kish, 2003] and alanine [Marino et al., 2000; Talley et al., 2002; Kish, 2003] via bidentate coordination to the amino nitrogen and the carbonyl oxygen of the amino acid in CS form, while the larger M^+ ($K^+/Rb^+/Cs^+$) prefers to bind with the two carboxylic oxygen atoms of glycine in the CS form. [Hoyau and Ohanessian, 1998; Talley et al., 2002] In contrast to the monovalent alkali metal cations, and with the exception of Be^{2+}, most divalent alkaline earth metal cations like $Mg^{2+}/Ca^{2+}/Sr^{2+}/Ba^{2+}$ [Strittmatter et al., 2000] stabilize the ZW form of glycine where the M^{2+} prefers to bind bidentately to the carboxylate oxygen atoms.

Recent theoretical studies also suggested the use of proton affinity (PA) could be a criterion to determine the relative stability of CS versus ZW forms of metal cationized amino acids and N-substituted amino acids, [Pulkkinen et al., 2000; Wyttenbach et al., 2000] because the CS and ZW forms of these complexes can be related by an intramolecular proton transfer process. Given the increase in PA from **Gly** to the larger aliphatic amino acids, would the ZW form gain more stability when compared to the CS form? Despite these studies, the underlying physico-chemical factors governing the relative stabilities of the CS versus ZW forms of the K^+-aliphatic amino acids remain unexplored. Thus, the objective of our present study is to investigate the nature of *intrinsic* interaction between K^+ and aliphatic amino acids, i.e., how ion-dipole and ion-induced dipole interactions affect the stability of different CS versus ZW binding modes.

4.2 Results and Discussion

4.2.1 Geometries of Ligands - the Aliphatic Amino Acids

In order to understand the effect of M^+ complexation on the geometry of aliphatic amino acids, stable conformers of the free amino acid of glycine (**Gly**), alanine (**Ala**), valine (**Val**), leucine (**Leu**) and isoleucine (**Ile**) have been explored. The conformations of **Gly, Ala, Val** and **Leu** have been the subject of several *ab initio* and DFT studies [Csaszar, 1996; Shirazian and Gronert, 1997; Balta et al., 2000; Matta and Bader, 2000], and these structures are in good agreement with these previously reported theoretical studies at various levels of theory. However, to our best knowledge, there are no similar studies for isoleucine previously. Isoleucine can be regarded as a methyl-substituted derivative of valine. Hence, we made use of our optimized geometry for **Val**, replacing one of the C4 hydrogen atoms with a methyl group, and rotate the C3-C4 bond to orient the C5 methyl group in order to minimize the steric repulsion. This trial structure of **Ile** was fully optimized at HF/6–31G(d), confirmed to be a minimum with frequency analysis, and refined at the B3-LYP/6-31G(d) level. The most stable conformers of the five aliphatic amino acids are shown in Fig. 4.1.

For **Gly, Ala** and **Val**, the rotational constants derived from our theoretical structures carry an error bar of less than 0.5% as compared to previously reported experimental [Iijima et al., 1991; Godfrey et al., 1993] and theoretical values [Stepanian et al., 1999] (Table 4.1). Hence, our calculated rotational constants of **Leu** and **Ile** are expected to serve as useful references for further experimental studies.

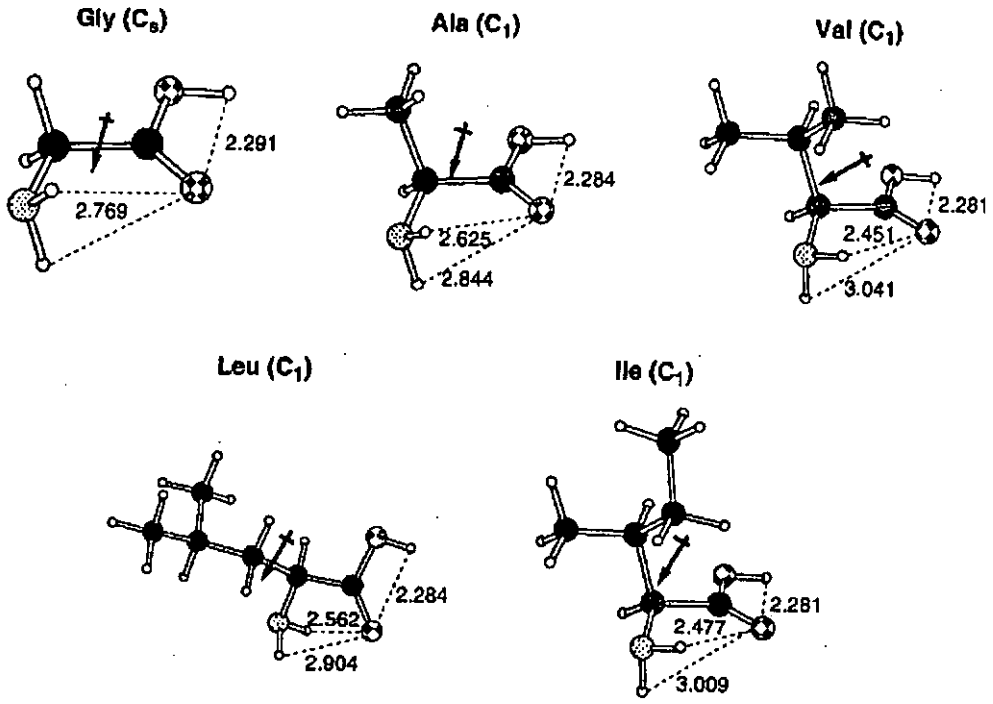


Figure 4.1. Geometries of the most stable conformer of five aliphatic amino acids obtained at the B3-LYP/6-31G(d) level of theory. The dipole moment vectors are drawn for reference only and are not to scale. Selected non-bond distances are shown in Angstroms.

Table 4.1. Theoretical rotational constants (in MHz) for the B3-LYP/6-31G(d) optimized structures of aliphatic amino acids.[a]

Species	Rotational Constants		
	A	B	C
Gly [b]	10297.6 (-0.3)	3858.9 (-8.6)	2898.9 (-12.1)
Ala [b]	5027.9 (-38.2)	3091.3 (-9.6)	2237.7 (-26.3)
Val [c]	2921.2 (11.2)	1415.0 (5.0)	1318.1 (9.1)
Leu	2745.2	831.1	782.3
Ile	2074.1	1081.6	965.2

[a]: Deviations from experimental values are given in parenthesis. [b]: The experimental rotational constant (A, B, C) are: 10297.9, 3867.5 and 2911.0 MHz, respectively, for **Gly** (ref. [Iijima et al., 1991]); and 5066.1, 3100.9 and 2264.0 MHz, respectively, for **Ala** (ref. [Godfrey et al., 1993]). [c]: In this work, no experimental rotational constant has been obtained. Holver, as the experimental infrared vibrational frequencies and intensity are in good agreement with calculated values for the most stable **Val** conformer 'Ia' (ref. [Stepanian et al., 1999]), thus suggesting that the theoretical structure obtained in this work is reliable. For this species, the reported theoretical rotational constants (A, B, C) are 2910, 1410 and 1309 MHz, respectively.

4.2.2 Binding Modes between K^+ and Glycine

There are three electron-rich binding sites (NH_2, O=C, and OH) in glycine, and hence, ten complexes with different modes of binding can be formed in principle. Out of these ten modes, seven of them are charge-solvated (CS) complexes (three monodentate, three bidentate and one tridentate modes) and the remaining three are zwitterionic (ZW) complexes (two monodentate and one bidentate modes). Our detailed exploration on the K^+-Gly potential energy surface reveals six complexes which differ in the mode of binding (**CS1** to **CS5**, and **ZW1**), and four other species,

CS1', **CS2'**, **CS5'** and **CS5''**, which are less stable conformers of **CS1**, **CS2** and **CS5**. The geometries of these ten K^+-Gly complexes are summarized in Fig. 4.2.

In agreement with previous findings, [Hoyau and Ohanessian, 1998] the tridentate CS mode (O=C, OH, NH_2) is not a stable minimum on the K^+-Gly potential energy surface. Starting with sensible candidates, these structures collapsed to the bidentate **CS2** species at the HF/6-31G(d) level. The monodentate CS mode of binding to hydroxyl group -OH is also found to be unstable, presumably due to the relatively weak interaction between K^+ and -OH as sensible candidates becomes either **CS1** or **CS4** upon geometry optimization. Similar result is observed in simpler model like carboxylic acids (Chapter 3), in which the K^+ is not preferable to bind with hydroxyl oxygen atom. Moreover, the ZW complexes where the K^+ binds only to *one* of the carboxylate (COO^-) oxygen atoms are not stable as these species optimize to the bidentate **ZW1** species, that involves the proton is transferred intramolecularly to the basic $-NH_2$ site. This suggests that the intrinsic attraction between K^+ and the formally negative charged carboxylate group is strong enough to overcome the strain of the bidentate binding mode.

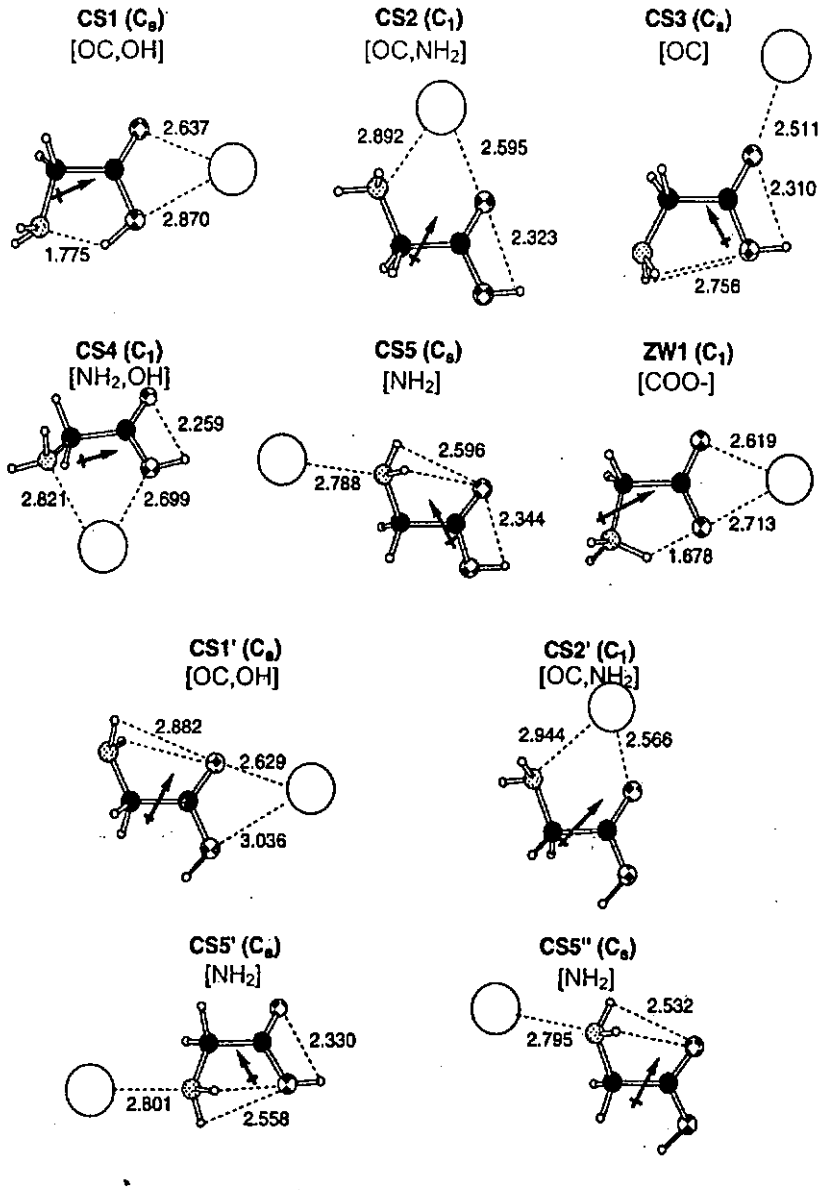


Figure 4.2. Geometries of ten stable K^+-Gly complexes, obtained at the B3–LYP/6–31G(d) level of theory. Their corresponding binding sites are shown in square bracket. The dipole moment vectors are drawn for reference only and are not to scale. Selected non-bond distances are shown in Angstroms.

Our structures obtained at the B3-LYP/6-31G(d) level for **CS1, CS2, CS4, CS5** and **ZW1** are comparable with that reported by Hoyau and Ohanessian (denoted as **3**, **1**, **2**, **5** and **6**, respectively, in ref. [Hoyau and Ohanessian, 1998] for K^+-Gly. One might argue that calculations without diffuse functions on heavy atoms are inadequate for ion-molecule complexes. In order to investigate the validity of our reported B3-LYP/6-31G(d) structures in Fig. 4.2, and its consequence on the reported energetics, the geometries of these K^+-Gly complexes at the B3-LYP/6-31+G(d) level of theory are also obtained. In terms of geometries, we found that both sets of geometries are similar with the distances between K^+ and various sites of binding shortened somewhat (average of 0.04Å) with the incorporation of the diffuse function, except for the geometries of species **CS1'**.

For species **CS1'** at the B3-LYP/6-31G(d) level of theory (Fig. 4.2), the interaction between K^+ and the two carboxylic oxygen atoms are quite asymmetric in **CS1'** where the binding distance of K^+...O=C is approximately 0.4Å shorter than that of K^+...OH. However, optimization of this species at the HF/6-31G(d) and B3LYP/6-31+G(d) level of theory yielded an essentially monodentate species, i.e., K^+ binding with C=O bond only. Hence, this species should be quite sensitive to the level of theory employed.

4.2.3 Binding Affinities for K^+-Glycine Complexes

From Table 4.2, the most stable binding mode is bidentate coordination to the two carboxylic oxygen atoms of Gly in CS form (**CS1**), which is even more stable than the ZW form (**ZW1**), with binding affinity obtained at the EP(K^+) level of theory.

Table 4.2. Binding affinities at 0K (ΔH_0, in kJ mol^{-1}), stabilization energies ($E_{stabilization}$, in kJ mol^{-1}), deformation energies (E_{def}, in kJ mol^{-1}), Dipole Interaction Parameter (DIP, in D Å^{-2}), Polarization Interaction Parameter (PIP, in Å^{-1}) for K^+-Gly conformers optimized at the B3-LYP/6-31G(d) level.

Species	Binding mode	ΔH_0 [a]	E_{def} [b]	$E_{stabilization}$ [c]	DIP [d]	PIP [e]
CS1	O=C, OH	117.7	7.7	125.4	0.3184	0.0256
CS2	O=C, NH_2	115.2	24.1	139.3	0.2141	0.0434
CS3	O=C	84.4	11.4	95.8	0.0869	0.0145
CS4	OH, NH_2	75.9	24.0	99.9	-0.0060	0.0429
CS5	NH_2	67.2	3.9	71.1	0.0697	0.0110
ZW1	COO^-	104.5	86.2	190.7	0.4594	0.0312
CS1'	O=C, OH	80.2	28.4	108.6	0.1507	0.0238
CS2'	O=C, NH_2	96.4	50.7	147.1	0.3194	0.0421
CS5'	NH_2	55.6	9.9	65.5	0.0339	0.0112
CS5''	NH_2	38.5	29.0	67.5	-0.0056	0.0105

[a]: Calculated by the EP(K^+) protocol. [b]: Refer to Eqn. [1.9]. [c]: Refer to Eqn. [1.10]. [d]: Refer to Eqn. [1.7]. [e]: Refer to Eqn. [1.8].

One might argue that calculations without diffuse functions on heavy atoms are inadequate for ion-molecule complexes. In order to study the effect of geometry on the calculated affinities, we also obtained the geometries of different modes of binding of K^+-Gly complexes (Fig 4.2) at the B3-LYP/6-31+G(d) level of theory and further compare the binding affinities in Appendix II, Table S-4.1. The binding

affinities obtained at the B3-LYP/6-311+G(3df,2p) level using the B3-LYP/6-31G(d) and the B3-LYP/6-31+G(d) geometries are virtually identical (within 0.7 kJ/mol^{-1} except **CS1'**). Even for the apparently problematic species **CS1'**, the difference in geometry only affects our reported energetics by 2 kJ mol^{-1}. Given the saving of computational time, and the geometry of the **CS1'** species might be even more "reliable" than that obtained at B3-LYP/6-31+G(d) level, the use of B3-LYP/6-31G(d) geometries for the larger K^+-aliphatic amino acids would appear to be justified. Besides, the geometries obtained with the Blaudeau 6-31G(d) basis set [Blaudeau et al., 1997] is comparable to that obtained here. With this alternative geometry, the binding affinity obtained at the B3-LYP/6-311+G(3df,2p) level again is virtually identical to our EP(K^+) protocol (within 0.7 kJ mol^{-1}). Hence, our results suggest that the choice of geometry is of secondary importance if the final energetics are reported at a sufficiently high level of theory. In other words, the binding affinities obtained at the same level of theory, with minor differences in geometry, could in fact be compared directly.

4.2.4 Factors Affecting the Relative Affinities of Different Binding Modes of K^+-Gly Complexes

Even though many theoretical studies on metal cationized amino acid complexes have been reported, the underlying factors affecting the relative affinities of different binding modes of K^+-Gly are seldom discussed. Thus, we would like to correlate the binding affinities of metal-ligand complexes with stabilization energies, $E_{stabilization}$, between the metal cation and the ligand arising from ion-dipole, ion-induced dipole and other electrostatic interactions, as well as the deformation energy (E_{def}) arising

from structural distortion (see Chapters 1 and 2 for details). The binding affinities at 0K, ΔH_0, deformation energy (E_{def}) and the stabilization energy ($E_{stabilization}$) for various K^+-Gly isomers are summarized in Table 4.2.

The stabilization energy of the M^+-Gly complexes arises primarily from the interactions of the cation with the permanent molecular dipole and the induced dipole of the amino acid. Thus, the defined molecular properties 'Dipole Interaction Parameter (DIP, Eqn. [1.7])' and 'Polarization Interaction Parameter (PIP, Eqn. [1.8])' will be applied to correlate with the K^+ affinities of different modes of binding. These molecular properties of the ligand (DIP and PIP) take into account of the orientation of the molecular dipole moment and polarizability with respect to the cation in the deformed **Gly** via the Φ, r_μ and r_α terms; and are expected to provide a better correlation with the stabilizing energies than μ and α. Both DIP and PIP, calculated at the B3-LYP/6-31G(d) level of theory, are tabulated in Table 4.2.

From Table 4.2, the $E_{stabilization}$ in general increases with increasing DIP of the M^+-Gly conformers, suggesting that ion-dipole interaction is a key factor in determining their relative stabilizing energy. Interestingly, for the case of **CS4** and **CS5''** (Fig. 4.2), the DIP is negative. This indicates that the positive end of the dipole moment vector is in fact closer to the positively charged K^+ than the negative end of the dipole vector. Hence, the ion-dipole interaction in these cases might in fact be destabilizing. In such cases, the ion-induced dipole interaction would be an important force binding the K^+ and the **Gly** together. Hence, a multiple linear regression analysis on $E_{stabilization}$ using DIP and PIP as the regression parameters were carried out, which can be described by Eqn. [4.1] for the ten K^+-Gly complexes:

$$E_{stabilization} = 194\ DIP + 1007\ PIP + 54 \qquad [4.1]$$

The correlation is very good, with an adjusted $R^2 = 0.92$. Fig. 4.3 shows the relation between the calculated $E_{stabilization}$ (from Table 4.2) and the predicted $E_{stabilization}$ using Eqn. [4.1]. With our simple model, the maximum deviation of the predicted $E_{stabilization}$ from the calculated value is 17 kJ mol^{-1}, with an average deviation of 8 kJ mol^{-1}. To test the applicability of this equation, test cases based on non-stationary points on the K^+-Gly potential surface were created (indicated by solid circles in Fig. 4.3). We found that for species with $E_{stabilization}$ within the range of 66 (**CS5'**) and 191 kJ mol^{-1} (**ZW1**), the average deviation between calculated and predicted $E_{stabilization}$ is 19 kJ mol^{-1}, suggesting that Eqn. [4.1] can be applied to estimate the raw binding energies within this range. Outside this range, Eqn. [4.1] is apparently less applicable as our test cases yielded an average error of 65 kJ mol^{-1}. No apparent pattern can be found between the magnitude of deviation and the structures of these test cases.

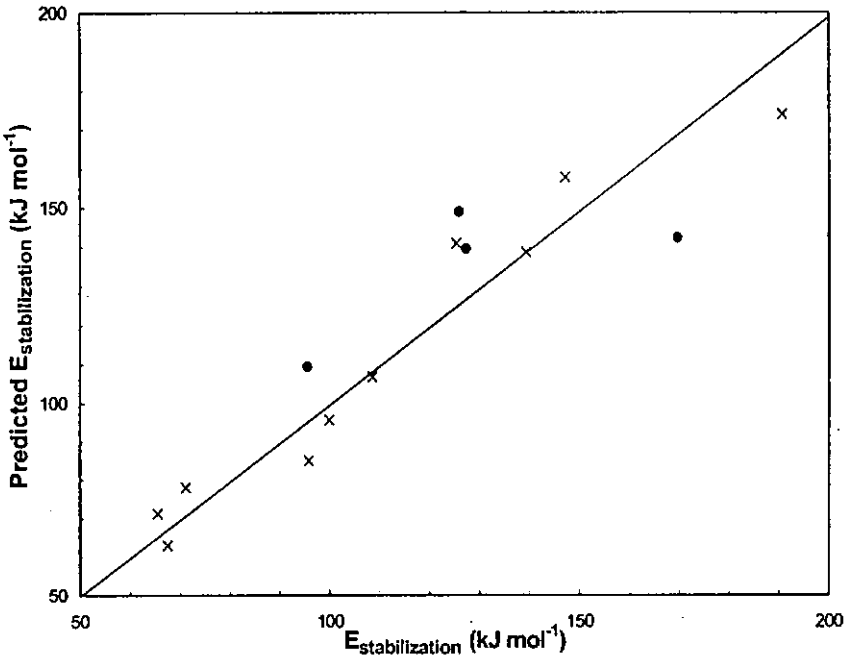


Figure 4.3. Plot of predicted stabilization energy (by Eqn. [4.2]) against calculated stabilization energy. The stationary points on the K^+-Gly potential energy surface are designated by 'x' while the non-stationary test points are labeled as '•'. The diagonal line with a slope of 1.0 is drawn for reference purposes.

Comparing our model with that of Gresh et al. as applied to Zn^{2+}-Gly complexes, [Rogalewicz et al., 2000a] the interaction energy of their model (equivalent to the $E_{stabilization}$ here) is partitioned into contributions from ion-dipole, ion-quadrupole, repulsion, polarization, charge-transfer and dispersion interactions. Since Eqn. [4.1] suggests that if the effect of ion-dipole and polarization is modeled classically via DIP and PIP, a constant term of 54 kJ mol^{-1} would represent the contribution from all the other kinds of interactions, that is ion-quadrupole, repulsion, charge-transfer and dispersion interactions, which are virtually constant compared with our model. Moreover, in all the K^+-Gly complexes, the natural charge of K^+ is very close to one (approximately 0.98 Coulombs). This suggests that the contribution due to covalent charge-transfer is minimal. Given that ion-quadrupole interaction is significant for cation-aromatic π interactions only, it appears that the sum of repulsive and dispersive interactions can be approximated by the constant term of 54 kJ mol^{-1}.

4.2.5 Types of Bonding Present in K^+-Aliphatic Amino Acid Complexes

For the larger aliphatic amino acids (**Ala/Val/Leu/Ile**), only the three most stable **CS1**, **CS2** and **ZW1** isomers are discussed, which also are the first three most stable complexes found in K^+-Gly case.

The geometries of **CS1**, **CS2** and **ZW1** for K^+ bound complexes of the larger aliphatic amino acids were found to be very similar to that of the corresponding K^+-Gly complexes (Fig. 4.4). For example, the average bond lengths for K^+...O=C and K^+...OH in the most stable **CS1** bidentate binding mode in K^+-Ala/Val/Leu/Ile

are 2.634Å and 2.827Å (Fig. 4.4), respectively, which are comparable to that obtained for the K^+-Gly complex (Fig. 4.2).

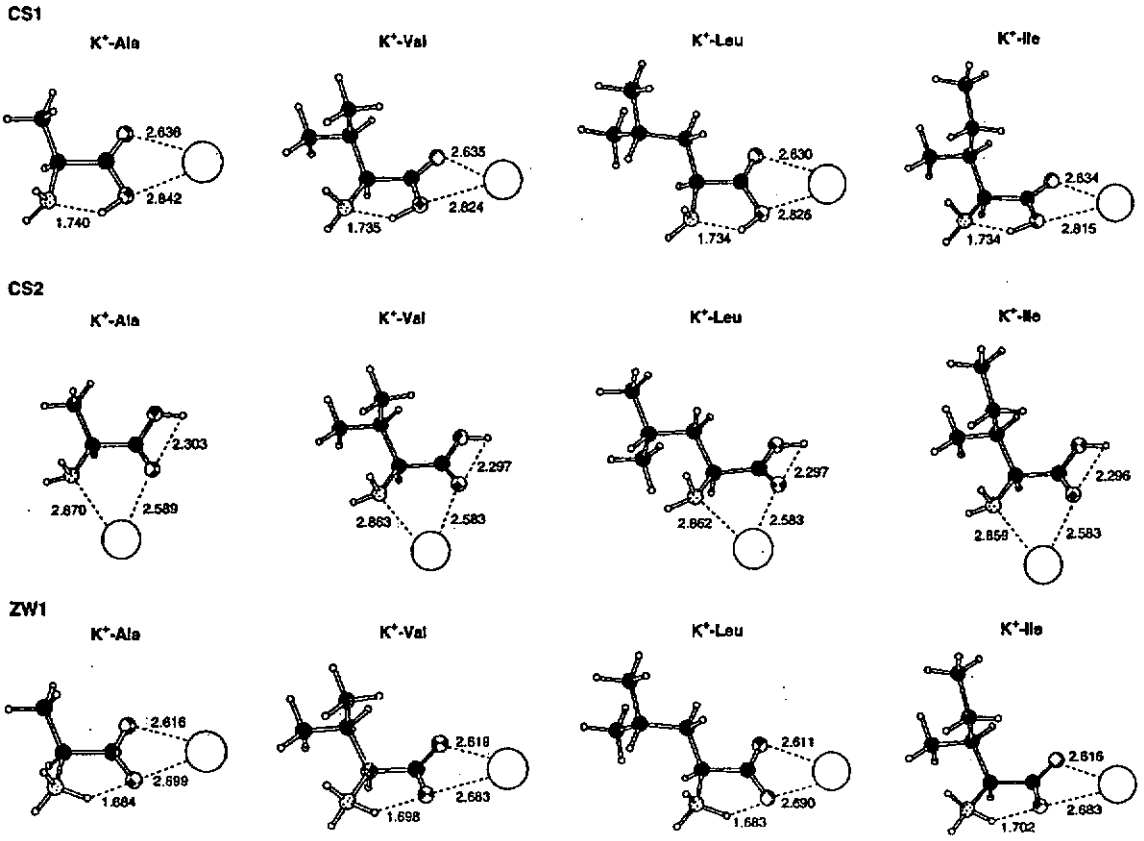


Figure 4.4 The optimized geometries of **CS1, CS2** and **ZW1** binding modes for different aliphatic amino acids, obtained at the B3-LYP/6-31G(d) level of theory. Selected non-bond distances are shown in Angstroms.

For **Val, Leu** and **Ile**, the potassiated complexes in which the K^+ is additionally stabilized by secondary interaction with a 'coiled' alkyl chain were also investigated, similar to that observed in the case of the Li^+/K^+ bound n–butanol complex [Ma et al., 2000]. Several plausible complexes were constructed where the K^+ might interact with the carbon on alkyl side chain (with K^+...C-alkyl initially at approximately 3.4Å), while maintaining the original **CS1/CS2/ZW1** modes of binding. Upon optimization at the HF/6-31G(d) level, these K^+...C-alkyl bonding distances

lengthened to at least 4.9Å, suggesting that these secondary interactions do not exist in the case of alkali metal cation binding to amino acids. Presumably, in order to establish these K^+...C-alkyl interactions, the amino acid ligand has to adopt a very unfavorable geometry where the potassium cation cannot interact efficiently with the oxygen/nitrogen binding sites. As a result, no stable **CS1/CS2/ZW1** isomers with additional K^+...C-alkyl ("coiled" form) interactions can be formed. Based on our study here, we could conclude that the presence of a long alkyl side chain does not alter the K^+ binding geometries at the O/N binding sites in the larger aliphatic amino acids, and the geometries of the CS and ZW binding modes would be very similar to that of K^+-Gly complexes shown in Fig. 4.2.

4.2.6 Theoretical Energetics K^+-Aliphatic Amino Acids Complexes

The theoretical K^+ binding affinities (ΔH_0), deformation energies (E_{def}) and stabilization energies ($E_{stabilization}$) determined at the EP(K^+) level for the **CS1, CS2,** and **ZW1** binding mode of all five aliphatic amino acids are summarized in Table 4.3.

Similar to the case of K^+-Gly complexes, the **CS1, CS2** and **ZW1** binding modes of larger aliphatic amino acids are energetically comparable, lying within a range of 8 kJ mol^{-1}, with the most stable mode of K^+ binding for all five aliphatic amino acids is the **CS1** mode.

Table 4.3. Theoretical EP(K^+) potassium cation affinities at 0K, ΔH_0, (in kJ mol^{-1}), stabilization energies ($E_{stabilization}$, in kJ mol^{-1}), deformation energies (E_{def}, in kJ mol^{-1}), three-dimensional Wiener Index (3-W, in Å), and three-dimensional alkyl side chain Wiener Index (3-$W_{side\ chain}$, in Å), for K^+-aliphatic amino acid (Gly, Ala, Val, Leu and Ile) conformers optimized at B3-LYP/6-31G(d) level.

Species	Modes	ΔH_0 [a]	E_{def} [b]	$E_{stabilization}$ [c]	3-W	3-$W_{side\ chain}$
K^+-Gly	CS1	117.7	7.7	125.4	45.64	0.00
	CS2	115.2	24.1	139.3		
	ZW1	104.5	86.2	190.7		
K^+-Ala	CS1	123.1	5.3	128.4	52.41	8.61
	CS2	117.9	24.7	142.6		
	ZW1	115.1	78.5	193.6		
K^+-Val	CS1	127.3	4.6	131.9	84.14	48.27
	CS2	122.1	25.4	147.5		
	ZW1	122.7	75.6	198.3		
K^+-Leu	CS1	127.9	6.1	134.0	112.69	43.50
	CS2	121.6	26.1	147.7		
	ZW1	122.2	78.4	200.6		
K^+-Ile	CS1	128.9	4.5	133.4	98.26	49.59
	CS2	123.7	25.3	149.0		
	ZW1	124.9	75.2	200.1		

[a]: Calculated by the EP(K^+) protocol. [b]: Refer to Eqn. [1.9]. [c]: Refer to Eqn. [1.10].

4.2.7 Comparison with Experimental and Literature Data

In order to compare with experimental and literature data, standard thermodynamic relations [DelBene et al., 1983] were applied to obtain the 298K affinities (ΔH_{298}) and free energy of binding (ΔG_{298}) of the most stable K^+ binding mode of the five aliphatic amino acid, i.e., the **CS1** mode (Table 4.4). As expected, theoretical ΔH_{298} is slightly larger than ΔH_0 by an average of 0.9 kJ mol^{-1}. Moreover, as entropy is expected to increase when the cation is released from complexes, ΔG_{298} is smaller than ΔH_{298} (by average of 27.8 kJ mol^{-1}). Nevertheless, the relative affinity and free energy of binding scales are essentially parallel, and this suggests that entropy effects are not important in determining the preferred interaction of K^+ with aliphatic amino acids.

For K^+–Gly, our theoretical value (118.6 kJ mol^{-1}) is in good agreement with the threshold-CID value (125.5 kJ mol^{-1}) reported by Kebarle and coworkers [Klassen et al., 1996], and with the MP2 estimate (120 kJ mol^{-1}) reported by Talley et al. [Talley et al., 2002]. While BSSE correction is small at the B3-LYP level (0.5 kJ mol^{-1}, n=5), it is much larger at the MP2 level (7 kJ mol^{-1}, n=2), leading to the consistently lower affinity (~6 kJ mol^{-1}) estimated by Hoyau et al. [Hoyau and Ohanessian, 1998; Talley et al., 2002] Previously, various research groups [Feller et al., 1995; Hoyau et al., 1999; McMahon and Ohanessian, 2000; Feller et al., 2000a; Moision and Armentrout, 2002] have commented that the full counterpoise approximation to BSSE can provide worse agreement with experiment than that of the theoretical values without BSSE corrections in systems such as those studied here. Given the additional computational cost and the small BSSE correction, we believed that full counterpoise correction is unnecessary for the DFT based protocol we employed here.

More importantly, the BSSE correction shows little variation amongst the five aliphatic amino acids. With or without BSSE, the order of relative K^+ binding affinity estimated theoretically for the **CS1** binding mode is:

Gly < Ala < Val ~ Leu < Ile

It is pleasing to note that the above theoretically established order is in exact agreement with the experimentally observed order or 'affinity ladder' obtained by the mass spectrometric kinetic method measurements [Tsang et al., 2001b]. Furthermore, as shown in Table 4.4, excellent agreement is achieved between our experimental and theoretical K^+ affinity values; the theoretical K^+ affinities are in excellent agreement with the absolute experimental affinities to within 1 kJ mol^{-1}, which is within the experimental uncertainty of ~8 kJ mol^{-1} for the kinetic method measurement [Tsang et al., 2001b], with the mean-absolute-deviation (MAD) is 0.4 kJ mol^{-1} only (n = 5). Such good quantitative agreement between theory and experiment provides confidence to the quality of our theoretical model employed and the conclusions drawn here.

Table 4.4. Potassium affinities (ΔH_{298}) and Gibbs free energies (ΔG_{298}) at 298K in kJ mol^{-1} for aliphatic amino acids.

Species	ΔH_{298}			ΔG_{298} [g]
	Experimental	Theoretical		
		Without BSSE	With BSSE	
Glycine (Gly)	119.3[a]	118.4[c]	118.0[e]	92.1
	125.5 [b]	120[d]	113[d]	
			110.9[f]	
Alanine (Ala)	123.6[a]	124.0[c]	124.9[e]	96.3
		126[d]	119[f]	
Valine (Val)	128.0[a]	128.2[c]	128.6[e]	99.8
Leucine (Leu)	129.3[a]	128.8[c]	129.6[e]	100.5
Isoleucine (Ile)	129.9[a]	129.8[c]	129.9[e]	101.5

[a]: Kinetic-method measurements using theoretical G2(MP2,SVP)-ASC(GCP) K^+ affinity values at 0K(ΔH_0) of acetamide (118.7 kJ mol^{-1})/ N-methyacetamide (125.6 kJ mol^{-1})/ N, N'-dimethylacetamide (129.2 kJ mol^{-1}) as reference values in ref. [Tsang et al., 2004]. The experimentally determined values approximated to ΔH_0 values at 0K and further corrected to ΔH_{298} using HF/6-31G(d) frequencies in this work. [b]: Experimental threshold CID values, ΔH_{298}, ref. [Klassen et al., 1996]. [c]: This work: theoretical EP(K^+) affinities at 298K (ΔH_{298}) for the most stable **CS1** binding mode (Figs. 4.2 and 4.4), with thermal corrections to 298 K using HF/6-31G(d) frequencies. [d]: Theoretical ΔH_{298} values at 298K reported by Talley et al. [Talley et al., 2002], calculated at the MP2/basis2//MP2/basis1 level. [e]: This work, EP(K^+) affinities at 298K with full counterpoise correction, including effect of geometry deformation calculated at the B3-LYP/6-31G(d) [Siu et al., 2001a]. [f]: Theoretical ΔH_{298} reported by Hoyau and Ohanessian [Hoyau and Ohanessian, 1998], calculated at the MP2/basis2//MP2/basis1 level including BSSE. [g]: Standard thermodynamic relations [DelBene et al., 1983] were applied to obtain the free energy of binidng (ΔG_{298}) at 298 K.

Lastly, we wish to comment on the difference between sodiated and potassiated aliphatic amino acid. Moision and Armentrout has recently proposed that volatilization of the ZW complex directly into the gas phase via electrospray ionization might be plausible, as the zwitterionic complexes enjoy better stability in the solution phase. [Moision and Armentrout, 2002] This proposal arises from both thermodynamics and kinetics consideration. Thermodynamically, the CS and ZW forms are of comparable stability: for sodiated glycine, the **ZW** form is only ~ 8 kJ mol^{-1} less stable than the most stable **CS2** mode. [Hoyau and Ohanessian, 1998] Furthermore, the rearrangement barrier between the two forms are high. In sodiated glycine, the barrier is estimated to be ~70 kJ mol^{-1}. [Hoyau and Ohanessian, 1998] Thus, if the ZW complex is formed initially, it may not be able to rearrange to a more stable CS form in the gas phase. [Moision and Armentrout, 2002]

Interestingly, our current theoretical K^+ affinities for the most stable mode of binding in aliphatic amino acid is in excellent agreement with experimental values. The difference in behavior in Na^+ and K^+ can be attributed to the different most stable binding modes. In the gas phase, the **CS2** mode (Na^+ binds to O=C and NH_2) is preferred in Na^+-Gly. If **ZW1** form were initially generated in the solution phase, it would take a few steps for this form to rearranged to the **CS2** form (estimated energy barrier of 70 kJ mol^{-1} [Hoyau and Ohanessian, 1998]) On the other hand, the most stable mode of K^+ binding is **CS1** (K^+ binds to O=C and OH). The formation of **CS1** from **ZW1** would be via a simple, one-step proton shift from the amino nitrogen to the carboxylate oxygen. Our calculation suggests that this process would be virtually barrierless. In other words, even if K^+–Gly generated in our experiment is in the **ZW1** form, it would be converted to the **CS1** form easily as the latter species is thermodynamically more stable.

4.2.8 The Effect of Alkyl Chain Length on the Relative Affinities of K^+-Aliphatic Amino Acids

While the effect of cationic charge and size on the relative stability of the most stable CS/ZW binding modes have been explored, the effect of alkyl chain length have not received much attention. Using the proton affinity argument, Wyttenbach et al. [Wyttenbach et al., 2000] suggested that while the K^+-**Gly/Ala** complexes should be in the CS mode, the K^+-Ile and possibly that of K^+-**Val/Leu** complexes, might be in the ZW mode. This is in partial disagreement with our theoretical results, which show that the **CS1** mode is *always* favored in the aliphatic amino acids. We further tested our conclusions by calculating the affinity with MP2(full) and B3-P86 protocols with the 6-311+G(3df,2p) basis set, using optimized B3-LYP/6-31G(d) geometry. While the absolute affinities differed somewhat (average deviation of less than 4 kJ mol^{-1}), the relative stabilities for the **CS1/CS2/ZW1** species are not affected. In other words, all three levels of theory (EP(K^+), MP2, and B3-P86) indicated that the most favorable K^+ binding sites in aliphatic amino acids are at the carboxylic oxygen atoms, designated as the "**CS1**" mode in this study.

The relative stabilities of the **CS1**, **CS2** and **ZW1** binding modes for the K^+-aliphatic amino acid complexes as a function of the alkyl side chain length are shown in Fig. 4.5. While the binding affinities of all three binding modes increase with increasing alkyl chain length, the rate of increase for **ZW1** is steeper than the two CS modes. In the case of **Val/Leu/Ile**, the **ZW1** mode is, in fact, more stable than the **CS2** mode of binding, even though it is still less stable than the **CS1** mode.

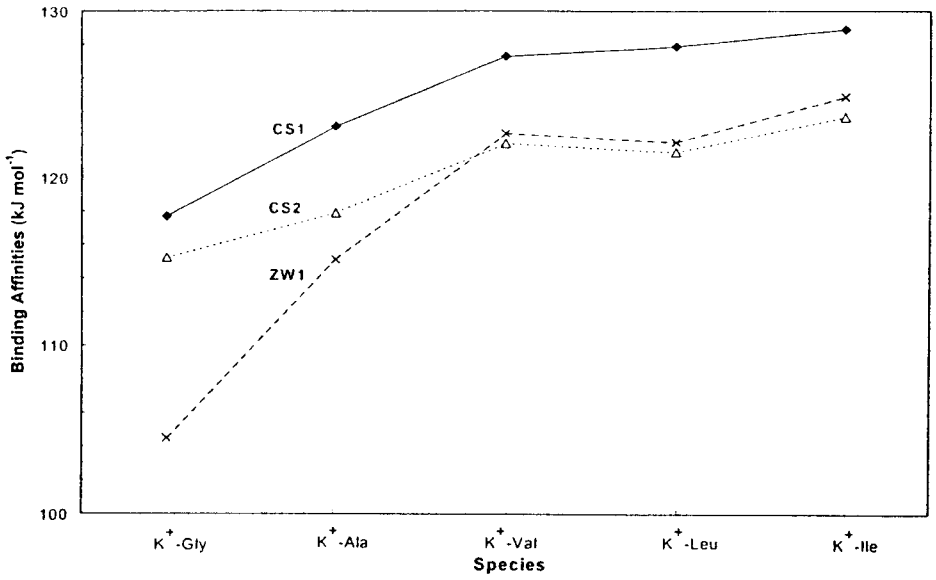


Figure 4.5. Variation of **CS1, CS2** and **ZW1** binding affinities for different aliphatic amino acids.

The E_{def} of the **ZW1** mode of K^+-Gly is particularly large (86 kJ mol^{-1}, Table 4.3) compared to the complexes of other K^+-aliphatic amino acids (75 to 79 kJ mol^{-1}, Table 4.3). This is in contrast to the relative insensitivity (less than 3 kJ mol^{-1}) of E_{def} for **CS1** and **CS2** modes of binding when the aliphatic amino acid is changed. Hence, the apparent increase in the stability of the **ZW1** mode, as compared to the other two CS binding modes for the larger aliphatic amino acids is, in fact, due to the greater deformation energy and hence instability of the **ZW1** binding mode observed especially for the smaller K^+-Gly complex. One can understand this by comparing the structurally similar **CS1** and **ZW1**, where the two species are related by an intramolecular proton transfer reaction. The larger E_{def} of the K^+-Gly complex suggests that the amino nitrogen of the **Gly**, as compared to the other aliphatic amino acids, is more reluctant to act as a proton acceptor. This could be related to the presence of an alkyl substituent at the C2 carbon of **Ala/Val/Leu/Ile**, but not for **Gly**. In the absence of an electron donating alkyl group at C2, protonation at the amino

nitrogen site might be more difficult. It is further supported by the proton affinity (PA) of the aliphatic amino acids. The PA of **Gly** (887 kJ mol^{-1}) [Hunter and Lias, 1998] is the smallest of all the aliphatic amino acids, and the difference of PA (in kJ mol^{-1}) between the **Gly/Ala** pair (15) is larger than the **Ala/Val** (9), **Val/Leu** (4) and **Leu/Ile** (3) pairs [Hunter and Lias, 1998]. Hence, in agreement with the predictions of Wyttenbach et al. [Wyttenbach et al., 2000], the preference for the **CS1** mode of binding over **ZW1** is reduced for the larger aliphatic amino acids, even though our calculations would suggest that the **CS1** mode is still slightly more stable than the **ZW1** mode for the larger aliphatic amino acids, **Leu** and **Ile**.

Previously, it has been suggested that a close to linear + - + arrangement for amino hydrogen, carboxylate oxygen, and metal cation (<H...O...M) is one of the factor stabilizing the formation a zwitterionic complex in Na^+-Pro. [Hoyau et al., 1999; Talley et al., 2002; Kish et al., 2003] Thus, we would like to explore whether there is any correlation between the stability of the **ZW1** mode in the aliphatic amino acids with the <H...O...M. We found that there is a slight increase of the angle from 0.5° (**Gly**) to 1.4°(**Ile**). Since the deviation from linearity is not significant (within 1.4° for five aliphatic amino acids), we believed that the only determinant factor governing the relative stabilities of the ZW form is the proton affinities of aliphatic amino acids.

Given the similarity of binding geometries for K^+-Gly and the larger K^+-aliphatic amino acid complexes, we attempted to predict the $E_{stabilization}$ of the K^+-Ala/Val/Leu/Ile complexes by the correlation equation (Eqn. [4.1]) obtained for K^+-Gly. Unfortunately, Eqn. [4.1] was found not applicable to K^+ bound complexes of other aliphatic amino acids. This suggests that the correlation coefficients are not constant when the amino acid is changed, so that Eqn. [4.1] is not transferable

between different amino acids. In fact, one may expect that the slightly increases in binding affinities with greater alkyl chain length from **Gly** to **Ile** should be mainly parallel to the increase in polarizability of the amino acids since the substitution of a CH_3 group for an H will not appreciably change the dipole moment of the ligand, but will increase the molecular polarizability. We further try to correlate the $E_{stabilization}$ with the three–dimensional Wiener index (3-W) of the amino acid [Trinajstic], which is a parameter correlated to molecular polarizability. The 3-W index is a topographical index which measures the compactness of a molecule, and is given by:

$$\text{Three-dimensional Wiener index}\,(3-\text{W}) = \frac{1}{2}\sum_{i}\sum_{j}(3-\text{D})_{ij} \qquad [4.2]$$

where $(3\text{-D})_{ij}$ is the matrix element representing the shortest Cartesian distance between all atom pairs *i* and *j* of the molecule [Trinajstic]. This index has been found previously to correlate well with shape-dependent physical properties like boiling points [Trinajstic], but to the best of our knowledge, this is the first time that it has been used to correlate with alkali cation affinities of neutral ligands. We found that the 3-W index (Table 4.3) correlates linearly with $E_{stabilization}$ (R^2 = 0.93), but only fairly with the absolute cation binding affinity, ΔH_0 (R^2=0.80).

In light of these observations, we like to introduce a modified form of the Wiener index, i.e., a side chain analogue of the 3-W index (denoted as $3\text{-W}_{side\ chain}$ in Table 4.3) which only restricts itself or models the compactness of the alkyl side chain while ignoring the conformation of the amino acid backbone. In other words, only the atoms in the side chain are included in the calculation of Eqn. [4.2]. As an example, in the case of **Ala**, $3\text{-W}_{side\ chain}$ was calculated from the shortest Cartesian distances of the methyl group attached to the C2 carbon. This $3\text{-W}_{side\ chain}$ index is found to

correlate very well linearly with both $E_{stabilization}$ and ΔH_0 with $R^2 = 0.90$. From this, we conclude that the variation of absolute binding affinity, ΔH_0, within the series of aliphatic amino acids might be more of a substituent effect, and less of a property of the whole ligand molecule. On the other hand, the overall stabilization interaction between the cation and the ligand might be better modeled as a linear function of the 3-W index of the whole ligand because the effect of ligand deformation is already taken into account in the $E_{stabilization}$ term.

4.2.9 Biological Implications

The relative stability of various charge solvation CS and zwitterionic ZW forms of metal cationized glycine has been studied previously [Hoyau and Ohanessian, 1998; Strittmatter et al., 2000]. From these two studies, one can draw the following conclusions: firstly, the smaller cations (Li^+, Na^+ and Be^{2+}) tend to favor the **CS2** modes of binding where the cation is bidentately coordinated to NH_2 and O=C groups. Secondly, for the larger alkali cations (K^+, Rb^+ and Cs^+), **CS2** is less stable relative to **CS1** where M^+ binds to the two oxygen atoms of the carboxylic acid group (OH and O=C). Thirdly, the larger divalent M^{2+} cations (Mg^{2+}, Ca^{2+}, Sr^{2+} and Ba^{2+}) are more capable of stabilizing the zwitterionic form of glycine, where the cation binds to the both carboxylate oxygen atoms (COO^-) than the monovalent M^+ in the metal cation-glycine complex. These results might have some interesting biological implications. In biological systems, due to the similarity in charge density, smaller cations of lower charges often behave chemically similar to larger cations of higher charges. Because of this so-called "diagonal relations", Li^+ shows similar physiological affinity for the phosphate ligand as compared to Mg^{2+} [Kaim and

Schwederski, 1994]. However, comparing the alkali cation [Hoyau and Ohanessian, 1998] to the alkaline earth cation bound complexes of glycine [Strittmatter et al., 2000], we found that such relation does not appear to hold in the gas phase. For the alkali cations, the charge solvation (CS) mode is always preferred. On the other hand, the zwitterionic (ZW) species tend to be more stable for the alkaline earth cation bound complexes. As different mode of binding is often associated with different binding energies, it would appear that the "diagonal relations" observed in biological systems could not be explained solely by the *intrinsic* binding energies of the metal cations, and solvent effects might be playing an important role in the establishment of the diagonal relations.

Based on a protein crystal structure at 3.2Å resolution for K^+ channels from *Streptomyces lividans*, MacKinnon and co-workers [Doyle et al., 1998] proposed a molecular mechanism for the selective transport of K^+ across the protein structures of K^+ channels. In this model, as the K^+ enters the pore region or selectivity filter of the ion channel, the water solvating the K^+ is removed. It is postulated that the amide carbonyl oxygen atoms of the peptide backbone act as surrogate co-ordination centers in place of water to stabilize the bare K^+ ion, even though the precise orientation of individual carbonyl oxygen atoms cannot be discerned at the resolution of the X-ray analysis. So it seems that the stability of **CS1** binding mode (binding to O=C and OH), which involves preferred binding to the amide carbonyl oxygen over other binding sites/modes, may be of direct relevance to the transmembrane transport of potassium ions in biological systems, and our reported theoretical geometries may serve as the reference on K^+-carbonyl oxygen bonding distances for further structural studies on K^+ channels.

4.3 Conclusion

This is the first time for the structures and binding affinities of potassium cation bound complexes of aliphatic amino acids (**Gly, Ala, Val, Leu, Ile**) in the gas phase are obtained at one uniform level of theory by the EP(K^+) protocol. The theoretical rotational constants of the free ligands and the calculated K^+ binding affinities are in excellent agreement with the available experimental data, in support of the reliability of the EP(K^+) protocol presented here.

In the case of K^+-Gly, all the stable stationary points (ten conformers) on the potential energy surface have been located. By analyzing the stabilization energies (raw interaction energies) in terms of classical molecular dipole moments and polarizabilities, we found that ion-dipole and ion-induced dipole interactions are the major factors in determining the relative stabilities of different modes of binding in K^+-Gly.

For the K^+-aliphatic amino acid complexes, even though the ZW form is stabilized more in the larger aliphatic amino acids than **Gly**, we found that the most stable mode of binding still involves the K^+ interacting bidentately to the carboxylic oxygen atoms in the CS mode. Important determinant for the formation of zwitterionic complexes is a higher proton affinity for aliphatic amino acids, but not the linearity for the <H...O...M). Moreover, we found good correlation of the absolute K^+ binding affinities with the side-chain three-dimensional Wiener index, but not with the three-dimensional Wiener index of the whole amino acids. This suggests that the increasing trend of binding affinities found within the series of aliphatic amino acids is more related to the polarizability of the side chain, rather than the property of the whole molecule. Lastly, our quantitative results may provide further insights on the

molecular basis of the "diagonal relations", and the transmembrane transport of K^+ in biological systems.

Chapter 5 Interaction between K^+ and Aliphatic Dipeptides: K^+-Glycylglycine (GG) and K^+-Alanylalanine (AA)

5.1 Background

In chapter 4, we have studied the interactions of alkali metal cations with simpler aliphatic amino acids. The lowest energy (most stable) structure of K^+-aliphatic amino acids complexes is found to be **CS1**, with K^+ binding to the [CO, OH] sites. (Chapter 4, Figs. 4.2 and 4.4) This may suggest that K^+ binding to the carboxylic oxygen atoms of peptides containing aliphatic amino acids residues should at least be a stable binding modes. Furthermore, since the species **CS2** with K^+ binding to [NH_2, CO] sites loses its competitiveness to **CS1** for the larger aliphatic amino acids, one may expect that the aliphatic peptides may not favor K^+ binding to the N-terminal – NH_2. On the other hand, the greater flexibility of the peptide backbone (as compared to the C-C backbone of an aliphatic amino acid) may allow alternative and more stable K^+ binding modes to occur. To extrapolate from the small amino acid model systems to metal cation-protein systems, the interaction between metal cation and the simplest dipeptide glycylglycine (**GG**) and alanylalanine (**AA**) is a vital bridge. With no functionalized side chain, the interaction between metal cations and the dipeptides, **GG/AA**, are restricted to O/N binding sites at the amide oxygen of the peptide bond, the amino nitrogen at the N-terminal, and the carboxylic oxygen atoms at the C–terminal, thus yielding valuable information on the individual metal cation binding sites on the peptide/protein backbone.

Because of the relatively large molecular sizes, only five high-level theoretical studies on Li^+-GG, [Feng et al., 2003] Na^+-GG [Cerda et al., 1998; Wyttenbach et al., 1998] and Cu^+/Ag^+-GG [Chu et al., 2001; Shoeib et al., 2001] peptide complexes have been reported. To the best of our knowledge, theoretical determination of K^+ affinities of dipeptides has not been reported in literature. The most stable mode of interaction between Na^+ and **GG** is a charge-solvated (CS) complex involving Na^+ binding bidentately to the two carbonyl O=C oxygen atoms of both amino acids residues, [Cerda et al., 1998] while the ZW forms of Li^+/Na^+ are found to be not as stable as CS forms. [Cerda et al., 1998; Feng et al., 2003] The Na^+-GG affinity was determined by extended mass spectrometric kinetic method measurements, [Cerda et al., 1998] and appears to be supportive of the theoretical findings. On the other hands, the more polarizable Ag^+ cation prefers to bound bidentately to the N-terminal amine nitrogen and amide carbonyl O=C oxygen in the charge-solvated (CS) form. [Shoeib et al., 2001]

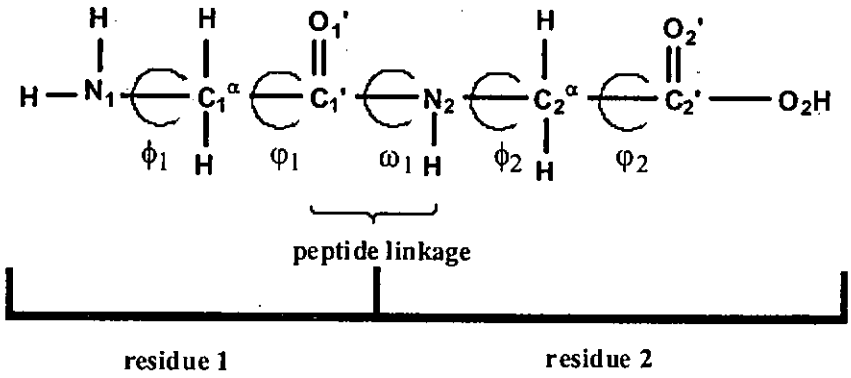
In order to gain a better understanding of how different metal ions interact with the peptide backbone and the nature of binding, we carried out density functional theory (DFT) studies of the potassiated–glycylglycine (K^+-GG) and the related potassiated-alanylalanine (K^+-AA) system. The Monte Carlo conformational search technique is used to generate plausible conformers for further *ab initio*/DFT calculations. Through this combined approach, we located several modes of metal cation binding on the **GG** and **AA** backbone that were not reported in previous studies. In the case of K^+-GG/K^+-AA complexes, some of these binding modes are fairly low-lying, and hence, may be energetically accessible under experimental conditions. The key stabilizing factor of 'local' ion-dipole electrostatic interactions is highlighted. Factors affecting the relative K^+ affinities of these different ZW versus CS binding modes will be

discussed. The usefulness of proton affinity (PA) as a criterion for estimating the relative stability of ZW versus CS binding mode is also examined. Based on the results of this study, the interaction of K^+ with longer peptides consisting of aliphatic amino acids are proposed and rationalized.

5.2 Computational Details

For the free ligand **GG**, the stability of its various conformers has been the subject of several theoretical publications. [Wu and Lebrilla, 1993; Zhang et al., 1993; Cassady et al., 1995; Cerda et al., 1998; Shoeib et al., 2001] We re-optimized the four most stable conformers presented in ref. [Cerda et al., 1998; Shoeib et al., 2001] at the HF/6-31G(d) level, and refined their geometries at the B3–LYP/6-31G(d) level. Single point energy calculations were performed at the B3-LYP/6-311+G(3df,2p) level based on the B3–LYP/6–31G(d) geometries.

To our best knowledge, there are no prior high level theoretical studies on the stabilities of K^+-GG complexes. The K^+ may interact with the dipeptide ligand in charge-solvated (CS) or zwitterionic (ZW) forms. The zwitterionic of the ligand can in turn be classified into three types in which the proton is attached to the terminal amino nitrogen atom (N_1), the nitrogen in the peptide linkage (N_2) or the amide carbonyl oxygen atom of the peptide bond (O_1') as depicted in Scheme 5.1.



Scheme 5.1. Systematic naming of atoms and dihedral angles of the glycylglycine ligands according to the IUPAC recommendation. [Anonymous, 1975] Atoms and their positions in the peptide backbone are indicated by letters/symbol/numbers in bold fonts. [Anonymous,1975]

We adopted a Monte Carlo Multiple Minimum (MCMM) conformational searching technique, [Saunders et al., 1990] outlined in details in Chapter 2, to obtain the K^+-GG conformations, with AMBER* force field [Weiner, 1984] implemented in the Macromodel 7.0 package. [Mohamadi et al., 1990] While the n_{conf} is 9000 for CS and ZW(O1'), and 7500 for ZW(N_1) and ZW(N_2) were carried out to locate the low energy structures for K^+-GG/AA. With an energy window of 50 kJ mol^{-1}, more than 130 of K^+-GG isomers/conformers were obtained. After discarding conformers with no intramolecular hydrogen bonds, ~60 K^+-GG conformers were retained. These species are re-optimized and confirmed to be a real minimum by using EP(K^+) protocol. (refer to Section 2.7)

The affinities at 0K (ΔH_0) for all K^+-GG complexes calculated using the EP(K^+) protocol are summarized in Table 5.1. Standard thermodynamic relations [DelBene et al., 1983] were applied to obtain the affinities (ΔH_{298}) and free energy of binding (ΔG_{298}) of various modes of binding at 298K (Table 5.1). As expected, for a given mode of binding, ΔH_{298} is larger than ΔH_0 (average of 1.4 kJ mol^{-1}). Moreover, as entropy is expected to increase when the cation is released from the complexes, ΔG_{298} is smaller than ΔH_{298} (by an average of 31.5 kJ mol^{-1}). Nevertheless, the relative affinity and free energy scale is essentially parallel, suggesting that entropy effect is not important in determining the preference of how K^+ interacts with the **GG** ligand.

Table 5.1. The theoretical energetics of potassiated glycylglycine (K^+-GG) complexes, in kJ mol^{-1}.

Species	Site of binding	ΔH_0 [a]	ΔH_{298} [b]	ΔG_{298} [b]	E_{def} [c]	$E_{stabilization}$ [d]
CS1	**O_1', O_2'**	151.8	153.0	120.7	31.1	182.9
CS2	**O_1', O_2', N_1**	145.3	147.2	110.6	45.4	190.7
CS3	**O_1', O_2**	122.5	123.4	92.6	21.5	144.0
CS4	**O_1', O_2, O_2'**	106.1	107.0	75.9	63.6	169.7
CS5	**O_2', O_2**	137.4	139.1	105.8	9.3	146.7
CS6	**O_1', N_1**	137.2	138.6	106.9	33.3	170.5
CS7	**O_1'**	127.3	127.8	100.1	6.7	134.0
CS8	**N_1**	80.2	80.7	54.9	14.0	94.2
ZW(O_1')	Carboxylate COO^-	103.8	105.5	74.6	149.5	253.3
ZW(N_1)	Carboxylate COO^-	90.4	93.5	55.9	88.8	179.2
ZW(N_2)	Carboxylate COO^-	46.1	47.8	18.6	156.1	202.2

[a]: Calculated by the EP(K^+) protocol. [b]: Standard thermodynamic relations [DelBene et al., 1983] were applied to obtain the affinities (ΔH_{298}) and free energy of binding (ΔG_{298}) at 298K. [c]: Refer to Eqn. [1.9]. [d]: Refer to Eqn. [1.10].

Furthermore, in order to understand how metal cation binding affects the structural and electronic energy of the ligand, we also calculated the deformation energy, E_{def}, and stabilization energy, $E_{stabilization}$, by using Eqns. [1.9] and [1.10] (Table 5.1).

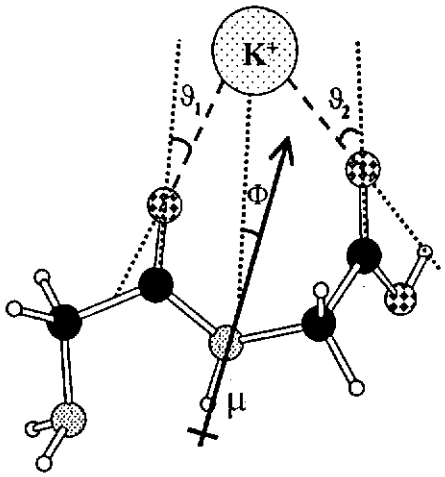
By replacing one hydrogen on each of C_1^{α} and C_2^{α} with a methyl group (without performing the MCMM conformation search again), we carried out EP(K^+) calculations to obtain ΔH_0 for the potassiated-alanylalanine (K^+-AA) system. Also, the same computational procedures were applied to obtain the ΔH_{298}, ΔG_{298}, E_{def} and $E_{stabilization}$ of the K^+-AA complexes. (Table 5.2)

As mentioned in Chapter 1, the strength of the ion-dipole interaction is directly proportional to the molecular dipole moment (μ) of the *deformed* ligand in the complexed state, the cosine of the angle of deviation between the cation and the dipole moment vector (Φ, in degree) (μ and Φ are shown in Scheme 5.2). The strength of ion-dipole interactions is strongest when the metal ion is in perfect alignment (Φ = 0°) with the dipole moment vector. The alignment of K^+ with the amide O=C, carboxylic O=C and –OH bonds at the individual O/N heteroatom binding sites are represented by the angles of deviation ϑ_1, ϑ_2 and ϑ_3, respectively (only ϑ_1 and ϑ_2 are shown in Scheme 5.2).

Table 5.2. The theoretical energetics of potassiated alanylalanine (K^+-AA) complexes, in kJ mol^{-1}.

Species	Site of binding	ΔH_0 [a]	ΔH_{298} [b]	ΔG_{298} [b]	E_{def} [c]	$E_{stabilization}$ [d]
CS1	**O_1', O_2'**	157.2	158.1	127.0	28.8	186.0
CS2	**O_1', O_2', N_1**	145.9	147.5	112.3	50.2	196.1
CS3	**O_1', O_2**	130.1	130.8	100.0	19.6	149.7
CS4	**O_1', O_2, O_2'**	113.1	113.6	83.6	64.9	178.0
CS5	**O_2', O_2**	140.6	142.2	109.7	14.0	154.6
CS6	**O_1', N_1**	135.8	137.1	106.0	40.3	176.1
CS7	**O_1'**	123.8	124.1	98.0	6.4	130.2
CS8	**N_1**	83.7	84.1	58.1	15.3	99.0
ZW(O_1')	Carboxylate COO^-	110.6	112.2	79.8	143.5	254.1
ZW(N_1)	Carboxylate COO^-	104.0	106.4	71.2	79.7	183.8
ZW(N_2)	Carboxylate COO^-	59.6	60.8	31.7	148.5	208.1

[a]: Calculated by the EP(K^+) protocol. [b]: Standard thermodynamic relations [DelBene et al., 1983] were applied to obtain the affinities (ΔH_{298}) and free energy of binding (ΔG_{298}) at 298K. [c]: Refer to Eqn. [1.9]. [d] Refer to Eqn. [1.9].



Scheme 5.2. Representation of the ion-dipole interactions in the K^+-Glycylglycine complex in which the angle of deviation between the cation and the dipole moment vector is Φ (in °), and the distance between the cation and the center of the dipole moment vector is r_μ (in Å, with origin at center of charge of the *deformed* ligand). The alignment of K^+ with the O=C bond axis is represented by the angles of deviation ϑ_1 and ϑ_2.

5.3 Results and Discussion

5.3.1 Glycylglycine Ligand

Four glycylglycine conformers, **GG1** to **GG4** (Fig. 5.1) were investigated. The most stable conformer of glycylglycine we obtained, **GG1**, is stabilized by three sets of intramolecular hydrogen bonds.

Our findings are in agreement with Cerda et al (species **X** in ref. [Cerda et al., 1998]) but in contrast with that reported by Cassady et al. [Cassady et al., 1995] and Siu et al. [Shoeib et al., 2001]. Without electron correlation (at the HF/6-31G(d) level), Cassady et al. [Cassady et al., 1995] suggested that **GG3** is most stable. On the other hand, Siu et al. [Shoeib et al., 2001] reported that conformer **GG2** (species **5N** in ref. [Shoeib et al., 2001]) is the global minimum. At our current level of theory, the energy difference between these three conformers is within 4.0 kJ mol^{-1} and such minor energy difference can easily be the result of using different theoretical treatment or protocols. Nevertheless, to obtain theoretical K^+ binding affinity of glycylglycine, it may not be important to locate the global minimum on the **GG** potential energy surface. [Shoeib et al., 2001]

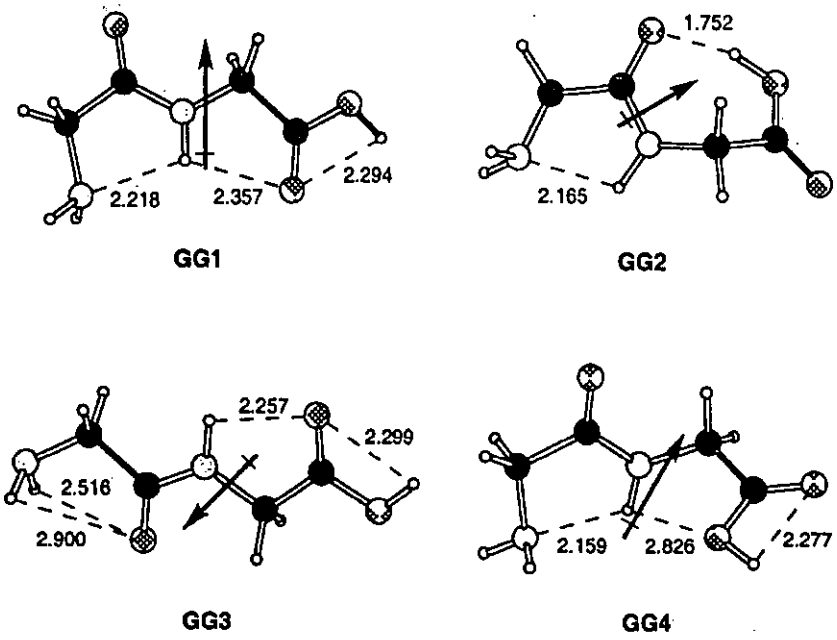


Figure 5.1. The geometries of four conformers of the glycylglycine (**GG**) ligand, optimized at the B3-LYP/6-31G(d) level. The dipole moment vectors are drawn for reference only and are not to scale. Selected non-bond distances are shown in Angstroms.

5.3.2 The Most Stable K^+-GG Complex

Using the MCMM and the EP(K^+) protocols, we have located 27 complexes (18 in CS forms and nine ZW forms) within 140 kJ mol^{-1} from the most stable K^+-GG complex. These 27 species can be further classified according to their modes of binding into eight CS (Fig. 5.2) and three ZW (Fig. 5.3) classes. The remaining 16 low-lying CS and ZW forms are shown in Appendix III, Fig. S-5.1. In our discussion below, we have reserved the term "form" to distinguish between charge-solvated (CS) and zwitterionic (ZW) species; "conformers" for species with same chemical formula and

mode of binding, but different conformations; and "isomers" for complexes with different mode of binding between the metal cation and the dipeptides.

We found that the most stable mode of interaction between K^+ and **GG** is one in which the ligand is in the charge-solvated CS form. This mode of binding, denoted as **CS1** here (Fig. 5.2), involves K^+ binding to the two carbonyl O=C oxygen atoms (one at the peptide amide bond, and one at the carboxylic acid functional group) on both amino acid residues. The K^+ is in very close alignment ($\Phi = 9°$) with the molecular dipole moment vector of *deformed* **GG** (Scheme 5.2), suggesting that ion–dipole interaction is important in stabilizing the **CS1** mode of binding. We note that **CS1** is also identified to be the most stable binding mode in the Na^+-GG complex. [Cerda et al., 1998]

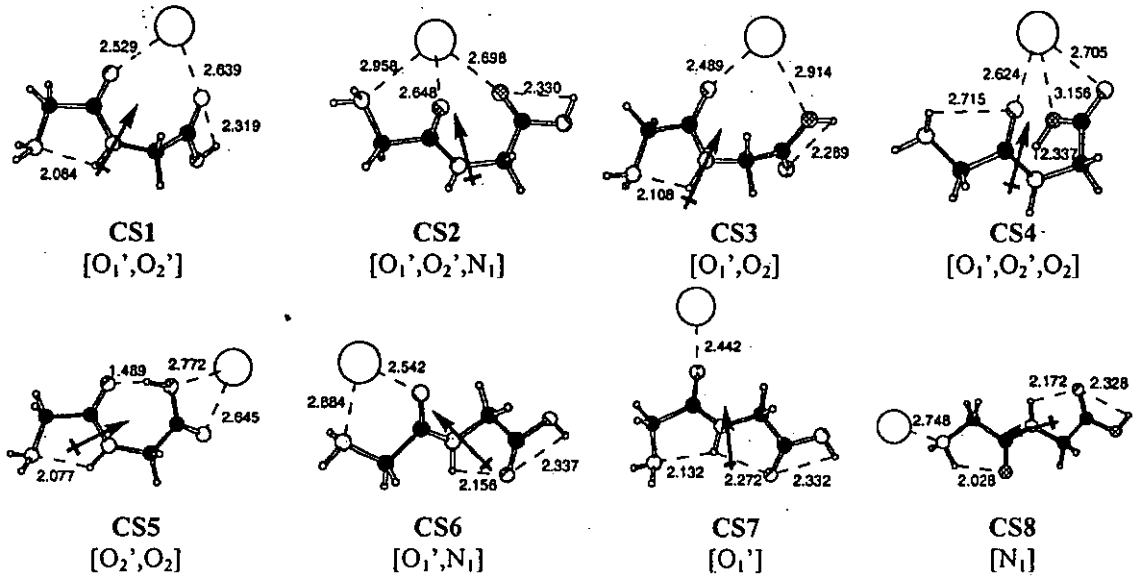


Figure 5.2. The geometries of eight CS binding modes of K^+-GG, optimized at the B3-LYP/6-31G(d) level. The molecular dipole moment vector of the *deformed* **GG** ligand is indicated by an arrow (not to scale). Selected non-bond distances are shown in Angstroms.

5.3.3 Other Charge-solvated (CS) Complexes

We found eight complexes (Fig. 5.2) in which the K^+ binds to **GG** in the CS form; the least stable CS isomer (**CS8**) has binding affinity 71.6 kJ mol^{-1} above the **CS1** complex (Table 5.1). These CS complexes (**CS1** to **CS8**) can be classified into two groups in terms of how the K^+ interacts with the **GG** ligand: those that involve binding to O/N sites of *both* amino acid residues and those that involve binding to *only one* residue. As in the case of **CS1**, we found good general alignment of K^+ with the dipole moment vector (Φ from 7° to 24°) in the **CS2** to **CS8** modes of binding.

For **CS2**, **CS3** and **CS4** complexes, the K^+ is attached to O/N heteroatoms from both residues. One may view these complexes as "derivatives" of the **CS1** mode of binding. Comparing **CS2** and **CS1**, the tridentate **CS2** complex is stabilized by an additional interaction between K^+ and the N-terminal amino nitrogen (N_t). While this additional interaction stabilizes the complex, it also leads to greater deformation of the **GG** ligand (E_{def} is 14.3 kJ mol^{-1} greater than that of **CS1**), hence decreasing the overall stability of the **CS2** mode of binding.

The binding affinity of **CS3** is quite comparable to that of **CS1**: in **CS1**, the K^+ interacts with two carbonyl oxygen atoms (O=C), while in **CS3**, the cation interacts with one O=C and one –OH. The interaction of K^+ with the hydroxyl (–OH) functional group is known to be noticeably weaker than binding to O=C, (Chapters 3 and 4) hence, accounting for the decrease in stability of the **CS3** mode of binding relative to **CS1**.

The **CS4** mode differs from the **CS1** mode in two aspects. Firstly, the N-terminal amino group adopted a "*cis*" (φ_1 = 29°), rather than a "*trans*" conformation (φ_1 = -

173°) in **CS1**. Secondly, the K^+ interacts with an additional -OH in the **CS4** mode of binding. Even though these changes might be considered minor, it is interesting to note that the E_{def} of **CS4** is more than twice that estimated for **CS1**. We attribute this large E_{def} to intra-ligand repulsion. In the **CS4** mode of binding, in order for the three negative sites to bind to K^+ simultaneously, the three oxygen atoms have to be very close to each other. The distance between $\mathbf{O_1}$' and $\mathbf{O_2}$ is only 3.04Å in **CS4** mode, which is 1.41Å shorter than what is found in **CS1**. Given that the $K^+\cdots$OH interaction is weaker, hence, the **CS4** mode of binding is less stable than **CS1** by 46 kJ mol^{-1}.

Complexes **CS5, CS6, CS7** and **CS8** are charge-solvated complexes in which the K^+ interacts with the O/N heteroatoms sites of one amino acid residue only. Hence, it is of interest to compare these species with the corresponding modes of binding in K^+-Gly (Chapter 4) so that the effect of the additional glycyl residue can be elucidated. Firstly, the ΔH_0 of K^+-GG is at least 13 kJ mol^{-1} higher than the corresponding modes of binding in K^+-Gly, with the largest difference (43 kJ mol^{-1}) found in **CS7**. (Chapter 4) Secondly, the order of relative affinity for these modes for binding (**CS5 > CS6 > CS7 > CS8**) of K^+-GG is identical to that of K^+-Gly. (Chapter 4) This implies that for the same mode of binding, the role of the additional spectator glycyl group in **GG** is simply to enhance the K^+ affinity, presumably due to the increase of permanent dipole moment and polarizability in the presence of the glycyl group. In other words, when one further extends the peptide backbone from **GG** to **GGG** (and other longer aliphatic peptides as well), and as long as the K^+ only binds to one glycine residue, the relative stabilities of the binding modes would not be expected to differ in the following order:

$$O{=}C + OH \sim O{=}C + NH_2 > O{=}C > NH_2$$

It is also interesting to compare the stability of **CS1** relative to that of **CS5** and **CS6**. When K^+ binds to small ligands, the raw interaction energy (in kJ mol^{-1}) $E_{stabilization}$, using EP(K^+) protocol, for formamide is found to be especially large: (Chapter 3)

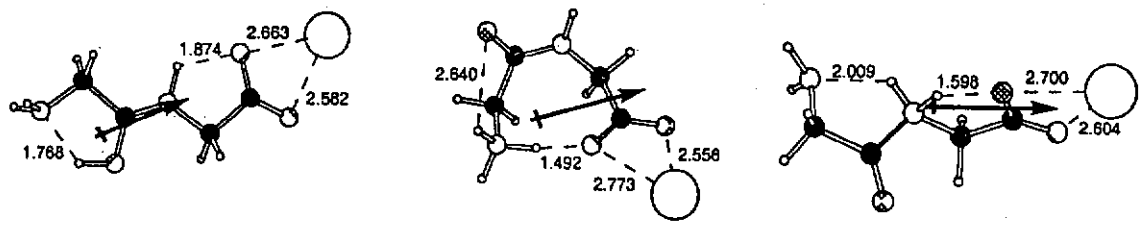
formamide (115) >> formic acid (83) ~ ammonia (74) ~ water (68)

, and is in line with the much larger theoretical dipole moment of *deformed* formamide (~ 4 Debye) in the complexed state when compared to the other three ligands (~2 Debye). The K^+ binding to these small organic ligands can be viewed as model interactions between K^+ and the individual O/N heteroatom binding site in peptides, i.e., K^+ binding to the amide C=O oxygen atoms (formamide) versus carboxylic C=O oxygen atoms (formic acid), N-terminal $-NH_2$ (ammonia) and the C-terminal –OH (water) functional groups. Here, for the K^+-GG complexes, we found that the $E_{stabilization}$ term (Table 5.1) for these bidentate modes of binding is in the order of: **CS1** (amide C=O + carboxylic C=O) > **CS6** (amide C=O + N-terminal $-NH_2$) > **CS3** (amide C=O + carboxyl OH), in line with the greater $E_{stabilization}$ derived from K^+ binding to the amide C=O oxygen. The only exception is **CS5** (carboxyl C=O and OH), which shows comparable $E_{stabilization}$ to **CS3**, presumably because **CS5** is stabilized by a particularly strong intramolecular hydrogen bond (~ 1.5Å, Fig. 5.2). Furthermore, the K^+...O=C interaction is enhanced by better alignment of K^+ with the bond axis of the binding sites (and presumably the 'local' dipole moment vector of the C=O binding sites) in **CS1** (with angles of deviation ϑ_1= 33° and ϑ_2=47° for the amide and carboxylic C=O bonds, respectively) than that in K^+-Gly, in which K^+ binds to the carboxylic C=O and –OH oxygen atoms (with angles of deviation ϑ_2= 77° and ϑ_3=91°, respectively). (Chapter 4) Thus, the stability of **CS1** is also related to the strength of the 'local' ion–dipole interaction between K^+ and the carbonyl oxygen

atoms, and especially K^+ binding to the amide carbonyl oxygen of the peptide bond which is energetically much favored. In fact, the highest $E_{stabilization}$ values (Table 5.1) are found for the four CS forms (**CS1, CS2, CS4** and **CS6**) which involve K^+ binding to the amide carbonyl oxygen of the dipeptide (Fig. 5.2). Our findings are in line with previous postulates that $M^+\cdots O{=}C$ ion dipole interactions (M^+= Na^+ or K^+) are important sources of attractive interaction contributing to the stability of M^+-formamide/acetamide and Na^+-GG complexes. [Roux and Karplus, 1995; Tortajada et al., 1995; Klassen et al., 1996]

5.3.4 Zwitterionic (ZW) Modes of Binding

We have identified three types of complexes where the K^+ interacts with the zwitterionic (ZW) form of **GG** (Fig. 5.3). In all three types of ZW complexes, the K^+ interacts in a bidentate fashion with the two carboxylate oxygen atoms COO^-, but differs in the site of intramolecular protonation: (i) at the amide carbonyl oxygen ($\mathbf{O_1'}$) of the peptide bond, (ii) at the N-terminal amino nitrogen ($\mathbf{N_1}$), or (iii) at the amide nitrogen ($\mathbf{N_2}$) of the peptide linkage.



ZW(O_1') [COO^-] **ZW(N_1)** [COO^-] **ZW(N_2)** [COO^-]

Figure 5.3. The geometries of three ZW binding modes of K^+-GG, optimized at the B3-LYP/6-31G(d) level. The molecular dipole moment vector of the deformed **GG** ligand is indicated by an arrow (not to scale). Selected non-bond distances are shown in Angstroms.

The binding affinities of the ZW complexes in K^+-GG system are at least 48 kJ mol^{-1} lower than the **CS1** mode of binding. It is interesting to compare this difference with our previous study of the K^+-Gly system. (Chapter 4) In the case of K^+–Gly, the lowest energy ZW mode is only 13 kJ mol^{-1} less stable compared to the most stable CS complex. (Chapter 4) As the site of K^+ binding is identical (at the carboxylate COO^-) in both K^+-Gly and K^+-GG, the relative instability of ZW modes of binding in K^+-GG arises from the more unfavorable (greater) charge separation in the zwitterionic dipeptide backbone. Similar conclusion has been drawn in the corresponding Na^+ system. [Wyttenbach et al., 1998]

Previously studies [Wu and Lebrilla, 1993; Zhang et al., 1993; 1994; Paizs et al., 2001; Rodriquez et al., 2001] suggested that, when an external proton is attached to **GG** and tripeptide **GGG**, the relative stability and basicity of different sites is in the order of:

amino nitrogen N_1 ~ amide carbonyl O_1' >> peptide amide N_2

In the case of ZW K^+-GG complexes, the preference of intramolecular proton transfer from the carboxylic acid group to the three basic sites is in the order (Table 5.1):

amide carbonyl $\mathbf{O_1'}$ > amino nitrogen $\mathbf{N_1}$ >> peptide amide $\mathbf{N_2}$

Thus, it can be concluded that complexation with K^+ has not altered the amide nitrogen ($\mathbf{N_2}$) as the least favorable site of protonation. This could be attributed to the loss of resonance stabilization of the peptide bond after protonation at the amide nitrogen [Zhang et al., 1994] which destabilizes the **ZW(N_2)** structure. Hence, the **ZW(N_2)** complex is approximately 44 kJ mol^{-1} less stable than **ZW(N_1)**, which has the protonation site at the N-terminal amino nitrogen $\mathbf{N_1}$.

For the **GG** ligand, it has been estimated that in terms of proton affinity at 0K (ΔH_0), protonation at the N-terminal amino $\mathbf{N_1}$ site is only favored by 3.7 kJ mol^{-1} than the amide carbonyl oxygen $\mathbf{O_1'}$ site (Table IV of ref. [Zhang et al., 1993]). [Zhang et al., 1993] Recent high-level density functional calculations of the **GGG** ligand also indicate that the proton affinity (ΔH_0) of the N-terminal amide carbonyl oxygen is only marginally smaller (by 0.9 kJ mol^{-1}) than that of the amino site (supplementary information of ref. [Rodriquez et al., 2001]). [Rodriquez et al., 2001] In both cases, protonation at the N-terminal amide C=O is stabilized by internal hydrogen bonding between the additional proton and the amino nitrogen ($-NH_2$), leading to very similar proton affinities for these two proton binding sites at the N-terminal glycyl residue. However, in the case of the zwitterionic K^+-GG complexes, protonation at the N-terminal amide carbonyl $\mathbf{O_1'}$ site is noticeably preferred by 13.4 kJ mol^{-1} (in terms of ΔH_0, Table 5.1) over protonation at the N-terminal amino $\mathbf{N_1}$ site. We attribute the greater stability of the **ZW(O_1')** binding mode of K^+-GG to three factors: (i) the enhanced resonance stabilization (partial double bond character) in the peptide

linkage $\mathbf{O_1'—C_1'—N_2—H}$, (ii) the formation of a stable hydrogen bonding interaction $\mathbf{O_1'H^{+}\cdots N_1}$ (1.77Å), analogous to the 'internal proton solvation' found at the N-terminal glycyl residue of protonated **GGG**, and (iii) the better alignment of the molecular dipole moment with K^+ in the **ZW(O_1')** mode of binding (Φ of 15° in **ZW(O_1')** versus 35° in **ZW(N_1)**).

The ZW complex with protonation site at the N-terminal N_1, **ZW(N_1)**, deserves further attention. Its deformation energy is ~40% smaller than that of the **ZW(O_1')** and **ZW(N_2)**, and quite comparable to that found in some CS modes of binding (e.g., **CS4**). The relatively small E_{def} in **ZW(N_1)** probably arises from an extra-ordinarily strong electrostatic bond between the hydrogen of the positively charged amino group and the oxygen of the negatively charged carboxylate group, i.e., $N_1H^{+}\cdots{}^{-}OOC$ interaction of 1.49Å. Such strong hydrogen bond could even compensate the two destabilizing factors arising from N_1 protonation: the preference for a more stable '*cis*' conformation in the peptide bond, [Luque and Orozco, 1993] and the charge separation in a ZW structure. This interaction is so stabilizing that upon metal complexation, the dipeptide is driven into a compact 'cyclic' configuration in **ZW(N_1)** as opposed to an 'extended' conformation found in the other zwitterionic K^+-GG complexes (Fig. 5.3).

5.3.5 Nature of Cation on Metal Cation-glycylglycine (M^+-GG) Interactions

Two high level theoretical studies on Na^+-GG have been reported by Bowers et al. [Wyttenbach et al., 1998] and Cerda et al. [Cerda et al., 1998] independently. The study conducted by Bowers et al. [Cerda et al., 1998; Wyttenbach et al., 1998] reported the relative energies of three Na^+-GG isomers. This sub-set of isomers are also included in a more comprehensive study on Na^+-GG interaction by Cerda et al., with qualitatively the same results. [Cerda et al., 1998] Thus, our comparison will only be made against the reported results of Cerda et al. More recently, the interaction between **GG** and Cu^+/Ag^+ has also been reported. [Chu et al., 2001; Shoeib et al., 2001] As Cu^+ is a substantially smaller cation than K^+, and the similarity/difference between Cu^+-GG and Ag^+-GG interactions have been extensively discussed in ref. [Shoeib et al., 2001], we will only focus on the interaction of **GG** with K^+, as opposed to Na^+ and Ag^+ here. Since Feng et al. have only located two geometries in Li^+-GG: one sample ground state of CS form and one ZW form, we may not compare their findings here. [Feng et al., 2003]

By carrying out geometry search described in this work, we have located some fairly stable modes of binding (e.g., **CS3, CS4, CS5, CS7, CS8, ZW(O_1') and ZW(N_2))** previously not reported in the Na^+, Cu^+ and Ag^+-GG systems. [Cerda et al., 1998; Wyttenbach et al., 1998; Shoeib et al., 2001; Chu et al., 2001] Certain modes of binding previously reported (e.g., species **V** in ref. [Cerda et al., 1998]) are less stable conformers on the K^+-GG potential energy surface that share the same mode of cation binding in Figs. 5.2 and 5.3 here, and the structure and energetics of these low-lying conformers are summarized in Appendix III, Fig. S-5.1. We note in passing that the relative stabilities of **GG** conformers sharing the same mode of binding are in fact

governed by the hydrogen-bonding patterns adopted by the **GG** ligand. The result of this aspect of K^+-GG binding will be reported in Appendix IV.

Here, we focus our attention on how the nature of metal cations affects the relative affinity of various modes of binding (Fig. 5.4). When compared to Na^+-GG system, [Cerda et al., 1998] the relative affinity scale of K^+-GG is compressed. Almost all the analogous Na^+-GG isomers are much less stable relative to the **CS1** mode of binding (species **II** in ref. [Cerda et al., 1998]), except for **CS2** (species **I** in ref. [Cerda et al., 1998]) which is of comparable stability (0.4 kJ mol^{-1}) to that of **CS1.**

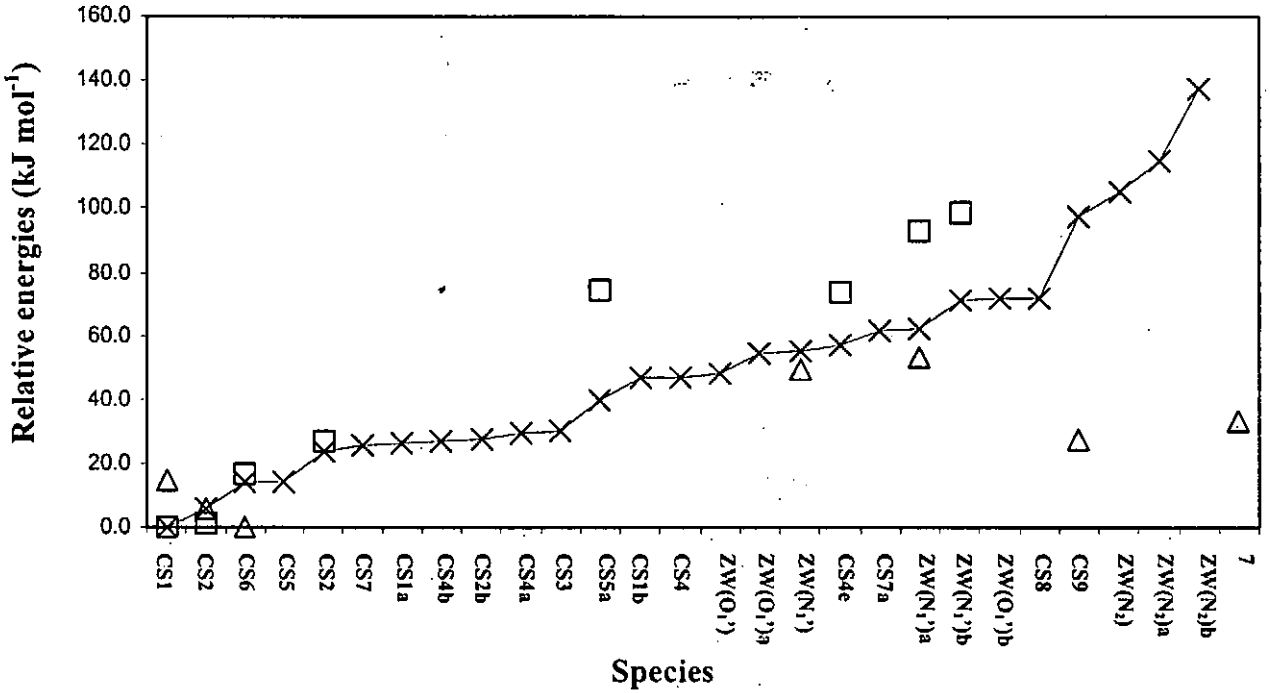


Figure 5.4. The relative energies (at 0K) of different modes of binding of glycylglycine for various cations: K^+ (this work at the EP(K^+) level, 5 with connecting lines for ease of visualization), Na^+ (ref. [Cerda et al., 1998] at the HF/6-31G(d) level, indicated by □) and Ag^+ (ref. [Shoeib et al., 2001] at the B3LYP/DZVP, indicated by Δ). The effect of zero–point energies are not included in the comparison as it was not reported in ref. [Cerda et al., 1998]. Species **7** in ref. [Shoeib et al., 2001] (identical to the structure **4** denoted in ref. [Chu et al., 2001]) for the Ag^+-GG could not be located on the K^+-GG potential energy surface here.

On the other hand, the K^+-GG relative affinity scale is expanded compared to that of the Ag^+-GG system (Fig. 5.4). Moreover, the most stable mode of binding also differs. In the most stable charge-solvated Ag^+-GG conformer (species **3** in ref. [Chu et al., 2001; Shoeib et al., 2001], corresponding to our **CS6**), the Ag^+ interacts with $\mathbf{O_1}$' and $\mathbf{N_1}$, while K^+ prefers to bind to $\mathbf{O_1}$' and $\mathbf{O_2}$' of the GG ligand (**CS1** mode of binding). One can rationalize this difference in terms of the hard–soft-acid-base (HSAB) principle. [Huheey et al., 1993] The silver cation is a softer acid (hence more polarizable) than Na^+/K^+. This means that Ag^+ is a better electron acceptor, and would prefer to bind to the softer nitrogen binding sites than the harder oxygen donor site in a ligand. The preference of Ag^+ binding to nitrogen binding site is also exemplified in the **CS2** mode of binding. As the **CS2** binding mode is already of comparable stability to the **CS1** binding mode in the case of Na^+, the preference of Ag^+ for the softer nitrogen site further stabilizes the **CS2** mode of binding. A further example can be found in species **6** in ref. [Shoeib et al., 2001] where the Ag^+ binding to carboxylic C=O and N-terminal $-NH_2$. With reference to the corresponding **CS1** mode, this mode of binding is relatively stable in Ag^+-GG (~15 kJ mol^{-1}), but very unstable (the **CS9** binding mode shown in Appendix III, Fig. S-5.1, ~100 kJ mol^{-1}) in K^+–GG. (This mode of binding cannot be found within the 50 kJ mol^{-1} energy window utilized in our MCMM calculations. The present **CS9** complex shown in Appendix III, Fig. S-5.1 is obtained from fully optimized K^+-GG complexes based on the geometries of species **6** in ref. [Shoeib et al., 2001])

We have failed to locate the corresponding zwitterionic species **7** in ref. [Shoeib et al., 2001] (identical to species **5** in ref. [Chu et al., 2001] on the K^+-GG potential energy surface. Starting from sensible trial K^+-GG structures, these complexes invariably optimized to **CS5** with K^+ binding to carboxylic oxygen atoms (Fig. 5.2) at both

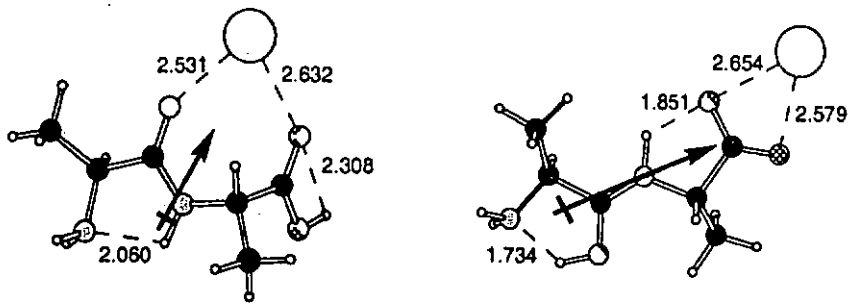
HF/6-31G(d) and B3-LYP/6-31G(d) levels. Species **7** (ref. [Shoeib et al., 2001]) and the **CS5** complex presented here are isomers that differs in the location of protonation: **CS5** is charge–solvated in nature while species **7** is zwitterionic in which the carboxylic proton has been transferred from the C–terminal to the N-terminal amide oxygen $\mathbf{O_1}$'. We have carried out additional geometry optimization with larger basis sets (i.e., 6-31+G(d), 6-31G(d,p) and 6–31+G(d,p)) with the B3-LYP function. As these more flexible basis sets also failed to yield stable complex similar to that of species **7** (ref. [Shoeib et al., 2001]) found in the Ag^+–GG system, it appears that such mode of binding is in fact unstable on the K^+-GG potential energy surface.

Examination of species **7** (ref. [Shoeib et al., 2001]) reveals that the proton bridges between the carboxylate oxygen and amide oxygen $\mathbf{O_1}$' are short (at 1.40 and 1.07Å, respectively), [Shoeib et al., 2001] and hence the proton would be expected to be quite mobile between these two alternative protonation sites. While cation binding could stabilize a ZW complex (through strong interactions between positively charged metal cations and the negatively charged carboxylate COO^- group), it also brings along instability to the ligand because of the charge separation effect. It appears that as the smaller Ag^+ binds more strongly and closely to the ligand, the stabilization factor outweighs the charge separation destabilization effect. As the interaction between the larger K^+ and **GG** is generally weaker, the stabilizing effect of the $K^+\cdots COO^-$ interaction is not strong enough to overcome the instability arising from the charge separation effect.

5.3.6 Effect of Alkyl Side Chain and Proton Affinity

The energy difference between the first two most stable conformers (**AA1** and **AA2**) is within 1.2 kJ mol^{-1}, with reverse in relative order that found in glycylglycine ligand. As we have discussed in Section 5.3.1, such minor energy differences can easily be the result of using a different theoretical protocol. Thus, the final theoretical K^+ binding affinity of alanylalanine will not be significantly affected even if **AA1** is not found to be the most stable conformers here. Thus, the theoretical energetics of K^+-AA that found in Table 5.2 are using the energy of conformers **AA1** for the calculation.

The modes of binding in K^+-AA is very similar to that of K^+-GG, with the more stable CS and ZW modes of binding depicted in Fig. 5.5. Generally speaking, the relative affinities of analogous binding modes in K^+-AA is parallel to that of K^+-GG. Presumably, because of the increase in polarizability, the raw interaction energy ($E_{stabilization}$) is increased, leading to a general increase in K^+ binding affinities of approximately 5 kJ mol^{-1} from **GG** to **AA**, which is identical to the change observed in going from K^+-Gly to K^+-Ala. Hence, it seems that the effect of increasing the alkyl chain length (polarizability) on relative stability of CS versus ZW in K^+ binding modes is the same for both aliphatic amino acids and dipeptides.



CS1 [O_1',O_2'] **ZW(O_1')** [COO^-]

Figure 5.5. The geometries of the most stable conformer of alanylalanine ligand (**AA**), charge-solvated complex (**CS1**), zwitterionic forms (protonated at **O_1'**, **ZW(O_1')**, optimized at the B3–LYP/6-31G(d) level of theory. The molecular dipole moment vector of the *deformed* **AA** ligand is indicated by an arrow (not to scale). Selected non-bond distances are shown in Angstroms.

A greater proton affinity (PA) would favor intramolecular proton transfer from the C–terminal carboxylic –OH to the amino group (-NH_2) group at the N-terminal, and is expected to confer greater stability to ZW forms of the M^+-amino acid complex. [Wyttenbach et al., 2000; Pulkkinen et al., 2000; Chapter 4] Here, we would like to investigate how one can extend this criterion to metal cationized dipeptides **GG** and **AA**. The proton affinities of glycine, glycylglycine, alanine and alanylalanine are summarized in Table 5.3, along with the relative stability of the CS/ZW forms in these K^+–ligand systems.

Table 5.3. A comparison of proton affinities (PA, in kJ mol^{-1}) of glycine (Gly), glycylglycine (GG), alanine (Ala), alanylalanine (AA) and the relative stabilities (E_{zw-cs}, in kJ mol^{-1}) of K^+ bound CS/ZW forms.

Species	Gly	GG	Ala	AA
PA [a]	902.5	934.7	912.5	946.8
E_{zw-cs} [b]	13.2	48.0	8.0	46.6

[a]: Experimental data at 298K and 1 atm taken from ref. [Cassady et al., 1995]. [b]: The ΔH_0 of the most stable K^+-ligand complex in the ZW form, relative to the most stable CS complex of the same system.

The PA of **Ala** is greater than that of **Gly** by 10 kJ mol^{-1}, suggesting that the more basic N-terminal site in **Ala** is more prone to proton attachment. Accordingly, formation of the ZW complex is found to be more favorable for **Ala** than **Gly**. A similar trend is observed in the dipeptide case when **AA** is compared against **GG**: the energy difference between the most stable CS and ZW forms becomes smaller for K^+-AA (at 46.6 and 48.0 kJ mol^{-1}, respectively). Hence, we could conclude that proton affinity affects the relative stabilities of CS and ZW complexes for both aliphatic amino acids and dipeptides in the same way.

However, the same could not be said for the K^+-Gly/K^+-GG or K^+-Ala/K^+-AA pairs. While **GG** and **AA** have greater proton affinities than **Gly** and **Ala**, the most stable ZW complex for the dipeptide ligand is much less stable than the most stable **CS1** complexes by 47 – 48 kJ mol^{-1} (Table 5.3), a difference significantly greater than the corresponding 8 – 13 kJ mol^{-1} found in the case of K^+-Gly and K^+-Ala complexes. (Chapter 4) It is because the most stable CS (K^+ binding to both carboxylic oxygen atoms) and ZW (protonation at the most basic N–terminal -NH_2 site) binding modes

in the aliphatic amino acids (refer to Chapter 4) are different from that of the dipeptides. Thus, the PA criterion can only be applied to systems that have the same CS or ZW modes of metal-cation binding, and their corresponding ZW forms should have the same protonation sites.

5.3.7 Comparison between Experimental and Theoretical K^+ Affinities

Since this is the first theoretical study on K^+ binding to dipeptides **GG** and **AA**, no literature values are available for comparison. To provide a better understanding on the K^+ binding to the dipeptides investigated, the experimental K^+ affinities obtained by Dr. Y. Tsang of our research group [Tsang, 2003] using the mass spectrometric kinetic-method and our theoretical results for the most stable CS and ZW binding modes are summarized in Table 5.4. The reference K^+ affinity compounds/values used in the experimental measurements are the K^+ affinities at 0K of amino acids with functionalized side chains and their methylated -OMe/ ethylated -OEt derivatives (in kJ mol^{-1}), i.e., Ser-OMe (137.7), Ser-Oet (140.4), Phe-OMe (143.9), Tyr-OMe (144.3), Phe-OEt (145.5).

Table 5.4. Theoretical and experimental K^+ affinities at 0K (kJ mol^{-1}) of K^+-GG and K^+-AA dipeptides.

Dipeptides	K^+ affinities				$\Delta H_{0\,(Expt.)} - \Delta H_{0\,(Theory)}$	
	Theoretical [a]			Experimental [b]		
	CS		ZW			
	With BSSE	Without BSSE	Without BSSE		With BSSE	Without BSSE
GG	151.4	**151.8**	103.8	139.0	12.4	12.8
AA	150.6	**157.2**	110.6	142.0	8.6	15.2

[a] Theoretical K^+ affinities, $\Delta H_{0\,(Theory)}$, using EP (K^+) protocol with or without BSSE correction, based on the B3-LYP/6-31G(d) optimized geometries. The affinities of the most stable binding mode are indicated by bold fonts. [b] Experimental K^+ affinities, $\Delta H_{0\,(expt.)}$, determined by Dr. Y. Tsang using extended kinetic-method measurements, with K^+ affinity values at 0K of Ser-OMe (137.7), Ser-OEt (140.4), Phe-OMe (143.9), Tyr-OMe (144.3), Phe-OEt (145.5) as reference values.

The order of theoretical K^+ affinities for **GG** and **AA** dipeptides, without BSSE correction, is in exact agreement with the experimentally observed order of the mass spectrometric kinetic method measurements. Furthermore, better agreement is achieved between the experimental and theoretical K^+ affinity values with charge-solvated (CS) binding mode, with the theoretical K^+ affinity values systematically higher than the experimental values by 13-15 kJ mol^{-1}, and a mean-absolute-deviation (MAD, without BSSE correction) of 14 kJ mol^{-1} (n=2). The deviation between experimental and theoretical absolute K^+ affinities is in the range of the experimental limits of error of 12-14 kJ mol^{-1}. One possible cause for the relatively large discrepancy (but still very acceptable) between experimental and theoretical values is that because of the relatively large size of the dipeptides, there are noticeable inter-

ligand repulsion in the K^+-bound heterodimer ion between the dipeptide and the reference compound, causing the measured K^+ affinity to be lower than the actual K^+ binding affinities. Another possible cause is that the error of the theoretical protocol increases with the molecular size of the calculated systems. Both of these causes may contribute to the observed increasing deviations between experimental and theoretical values, which are still close to the limits of experimental uncertainty and thus, the affinities can be regarded as still acceptable. Even though the inclusion of BSSE calculation are found to be in fair better agreement with experimental data (MAD =10.5 kJ mol^{-1}), the order of K^+ affinities for **GG** and **AA** dipeptides is reversed relative to experimental affinities. Therefore, the corrections with BSSE may not lead to better agreement in the order of experimental affinities and has minimal effect on calculated K^+ affinities. Therefore, it appears that BSSE correction could be omitted for computational efficiency. In addition, the experimentally measured K^+ affinities should correspond to the binding affinities of the K^+-dipeptide complexes in the most stable CS forms in gas phase. Theoretically, the most stable ZW form of K^+ bound dipeptides are generally ~ 50 kJ mol^{-1} less stable than the CS form. Because of the large difference in binding affinities and stability between the most stable CS and ZW forms, the K^+ bound dipeptides complexes are expected to be found predominantly in the CS forms in the gas phase under the experimental conditions of the mass spectrometer system adapted in the measurements.

5.3.8 Interaction of K^+ with Peptide Backbones

In the previous sections, we have discussed the factors governing the relative stability of the CS and ZW modes of binding in K^+-GG/K^+-AA systems. By comparing our present results with the intrinsic gas phase K^+-peptide backbone interactions reported in the literature, [Wyttenbach et al., 1998; Kohtani et al., 2000; Taraszka et al., 2001; Sudha et al., 2002] as well as that found in biological systems, [Larsen et al., 1994; Doyle et al., 1998; Isupov et al., 1998; Wyttenbach et al., 1999b] we could gain new insight into the interactions of these and related systems.

Results from the present study illustrate the importance of 'local' $K^+\cdots$O=C ion–dipole interactions between K^+ and **GG/AA** so that the cation prefers to bind to two carbonyl oxygens of the dipeptide backbone in the charge-solvated (CS) mode. Given this, one expects that for the longer peptides, the increase in flexibility of the backbone would allow the ligand to orient its various carbonyl O=C groups to align more closely with the cation, so that the number of $K^+\cdots$O=C interactions could be maximized. At the same time, the zwitterionic (ZW) K^+ binding modes become much less stable (by at least about 47 kJ mol^{-1}) in the aliphatic dipeptides **GG** and **AA** than in the amino acids glycine and alanine. This is due to (i) the enhanced stability conferred by the strong $K^+\cdots$O=C interaction associated with the peptide bond in the most stable **CS1** structure, and (ii) the instability of the ZW structures arising from the greater charge-separation effect in the dipeptide backbone.

Both factors suggests that the K^+ is likely to be encapsulated inside the peptide chain in a macrocyclic CS conformation with multi-dentate binding to mostly amide O=C sites of the peptide backbone. We found first support in the potassiated valinomycin (a dodecadepsipeptide) complexes where the K^+ binds to the six valine carbonyl

oxygens in a near octahedral arrangement in the gas phase. [Wyttenbach, 1999b] The second support is in the metalliated CHO-Gly_3-OCH_3 complexes where the alkali metal cations bind to the three C-terminus CO groups is the energetically most favorable for Li^+, Na^+ and K^+. [Sudha et al., 2002] Moreover, given the similarity between the mode of binding between Na^+ and K^+ illustrated in this work for the **GG/AA**, and previous works for smaller ligands, [Rodgers and Armentrout, 2000a] our results can be used to explain the occurrence of macrocyclic CS modes of Na^+ binding to backbone carbonyl oxygens in sodiated oligoglycines (Na^+-Gly_n, n=2-6) and oligoalanines (Na^+-Ala_n, n=10, 15, 20, and $[Ala_n + 3Na]^{3+}$, n=18-36). [Cerda et al., 1998; Wyttenbach et al., 1998; Kohtani et al., 2000; Taraszka et al., 2001] In biological systems, K^+ binding to backbone carbonyl oxygens in a macrocyclic pattern are found in the X-ray protein structures of K^+ channels [Doyle et al., 1998], tryptophanase [Isupov et al., 1998] and pyruvate kinase [Larsen et al., 1994], suggesting that the importance of local $K^+\cdots O{=}C$ ion–dipole interaction may be extendible from the gas phase to solution phase.

Finally, we found that the most stable ZW binding mode **ZW(O_1')** for K^+-GG and K^+–AA has the protonation site at the amide carbonyl oxygen **O_1'** of the N-terminal glycyl residue (Scheme 5.1), which has proton affinity very similar to that of the amino nitrogen **N_1** site. Despite stabilization by very strong hydrogen bonding (as indicated by the very short bonding distance of 1.77Å) between the **O_1'H** and **N_1** within the N-terminal glycyl residue, this potassiated **ZW(O_1')** structure of **GG** and **AA** remains much less stable than the most stable conformer **CS1**. For longer aliphatic peptides, the ZW conformers are expected to become even less stable due to the greater charge separation effect. Drawing on the similarity in binding modes between K^+ and Na^+ again, our results on the relative instability of ZW structures

ZW(O_1') and **ZW(N_1)** for the dipeptides **GG** and **AA** are in line with a previous report that a longer helical ZW Na^+-Ala_{15} structure collapsed to a random globular structure (presumably in the CS form) in numerical simulations. [Kohtani et al., 2000] Furthermore, a helix with a salt bridge from the deprotonated C-terminus (zwitterionic $COO^{-}\cdots Na^{+}$) and protonation of the backbone C=O near to the C-terminus (to minimize charge separation effect) appears to be a stable structure in the numerical simulation. Hence, the larger aliphatic peptides are most likely to remain in the CS form when sodiated or potassiated.

5.4 Conclusion

To our knowledge, this is the first high-level *ab initio*/density functional study on K^+ interaction with dipeptides. In this study, we have located eighteen charge-solvated (CS) and nine zwitterionic (ZW) stable isomers/conformers for the K^+-GG/AA dipeptide complexes at the B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d) (abbreviated at EP(K^+)) level of calculations, which can be classified into eight CS and three ZW K^+ binding modes to different O/N heteroatom sites of the peptide. Several of these binding modes are not found in previous studies of Na^+, Cu^+ and Ag^+ binding to the dipeptide **GG**.

The most stable K^+-GG and K^+-AA complexes involve a bidentate interaction where the K^+ coordinates to two C-terminus carbonyl oxygen atoms of two amino acid residues in the CS form. We found good general alignment of K^+ with the dipole moment vector of the *complexed (deformed)* aliphatic dipeptides in all of the CS modes of binding, suggesting that ion dipole interaction is the key electrostatic interaction contributing to the stability of the K^+-GG/AA complexes. Among the different O/N heteroatom binding sites on the peptide backbone, K^+ binding to the amide carbonyl oxygen is energetically very much preferred, and this is attributed to the very strong *local* ion-dipole interaction between K^+ and the peptide amide C=O bond. Consequently, the more stable CS forms are those that involve K^+ binding to and in close alignment with the amide C=O.

By definition, zwitterionic K^+-dipeptide structures have the potassium cation bound to the two carboxylate oxygen atoms at the C-terminus, thereby losing the significant stability that can be gained from the $K^+\cdots$C=O (amide) interaction. While K^+ binding to backbone amide carbonyl oxygen atoms in the Na^+/K^+ bound ZW structures of

larger helical or globular peptide chains are still possible, the CS form will always have the advantage of being able to bind to more amide carbonyl oxygen atoms than the ZW form, thus gaining extra stability for the CS form. We also found that protonation and internal proton solvation at the basic amide carbonyl oxygen of the N-terminal glycyl/alanyl residue in **GG** and **AA** cannot compensate for the destabilizing charge separation effect (between the proton positive charge and the carboxylate negative charge) even for the lowest energy (most stable) **ZW(O_1')** structure involving protonation and strong hydrogen bonding at the more basic N-terminal glycyl/alanyl residue. Since the charge separation effect is expected to be amplified in ZW structures of longer peptides, and coupling with the lesser probability of K^+ binding to amide C=O of the peptide backbone for ZW structures, it is very likely that the most stable K^+ stabilized aliphatic peptide complexes have charge-solvated conformations in the gas phase.

The stability of the lowest energy ZW binding mode of **GG/AA** (relative to their respective most stable CS forms) is found to increase slightly with the proton affinity of the dipeptide. However, the proton affinity criterion fails when cross comparisons are made between the dipeptides and the aliphatic amino acids. Hence, we would suggest caution in extending the proton affinity criterion to compare or predict the relative stability of different ZW structures of amino acids or peptides having dissimilar metal cation binding modes.

While the CS and ZW binding modes and the trend of relative stabilities are similar for Na^+/K^+-dipeptide complexes, we found the most stable CS form of K^+-GG and Ag^+-GG complex to have different O/N heteroatom binding sites, and other stable CS conformers also show significant differences in their relative stabilities. The origin of

these differences may reflect the different nature of bonding and ionic sizes of Ag^+ and K^+. Hence, differences in the mass spectral fragmentation patterns between Ag^+-peptides and K^+/Na^+-peptides are expected. Mass spectra of Na^+, [Grese et al., 1989; Teesch and Adam, 1991a; Teesch et al., 1991b], K^+ [Cody et al., 1985; Grese et al., 1989; Teesch and Adam, 1991a] and Ag^+ [Chu et al., 1999] cationized peptides have been usefully applied to provide identification and sequence information of peptides. It might be of practical interest to examine whether such differences can be exploited to provide complementary mass spectral information on peptide sequence, which has become an important issue in proteomic analysis today.

Chapter 6 K^+-Glycylphenylalanine (GF) and K^+-Phenylalanylglycine (FG): The Importance of Cation-π Interactions

6.1 Background

The attractive force between positively charged species/centers and an electron-rich aromatic ring, the "cation-π" interaction, is a relatively new kind of non-covalent force. [Gokel et al., 2000; Tsuzuki, et al., 2001; Chen, 2002; Zacharias and Dougherty, 2002] Such force has been postulated to play important biological roles, like protein-DNA binding [Wintjens et al., 2000], antigen-antibody interactions, [Hsieh-Wilson, 1996; Bracci et al., 2003] and selective ion transport across cell membranes channels. [Silverman et al., 1998; Gokel et al., 2002; Nino et al., 2003] Cation-π interaction is found to be implicated in the stability of secondary/tertiary structures of peptides and proteins [Gallivan and Dougherty, 1999; Fernandez-Recio et al., 1999; Toth et al., 2001b], and is being exploited in the design of drugs. [Wouters and Ooms, 2001; Mao et al., 2004]

In biological systems, the cation-π interaction can occur between the metal cations, [Hirota et al., 2000; De Wall et al., 2000; Shoeib et al., 2002; Hu et al., 2002a; 2002b; Williams et al., 2003; Meyer, 2003] cationic side chains of lysine/arginine/histidine, [Gallivan and Dougherty, 1999; Chen 2000; Nurizzo et al., 2001; Okada et al., 2001; Zacharias and Dougherty, 2002] and organic cations like acetylcholine and its derivatives [Zhong et al., 1998; Smit et al., 2001; Koellner et al., 2002], with the aromatic side chains of phenylalanine/tyrosine/tryptophan on the peptide/protein chain. Given the importance of such interactions, insight into the fundamental nature

of cation-π interactions is necessary. This prompted many theoretical studies of the interaction between alkali, [Amicangelo, and Armentrout, 2000; Feller et al., 2000a; Ikuta, 2000; Tsuzuki et al., 2001; Dunbar, 2002b; Kim, et al. 2003b; Chapter 3] alkaline, [Zhu et al., 2003] transition metal, [Dunbar, 2002b; Caraiman et al., 2004] and ammonium cation [Zhu et al., 2000; Kim, et al. 2003b] with simple aromatic systems benzene and its derivatives. Because of the size of the system, high level theoretical study of the interaction between cations and aromatic amino acids/peptides remained few in number. [Ryzhov et al., 2000; Dunbar, 2000; Siu et al., 2001b; 2004] Density functional calculation by Dunbar and co-workers indicated that the phenyl-π face was one of the most preferable sites involved in the Na^+/K^+ bound complexes of phenylalanine/tyrosine/tryptophan. By comparing against the corresponding non-π mode, the additional π interaction is found to increase the binding affinity by ~20 kJ mol^{-1}. [Ryzhov et al., 2000; Dunbar, 2000] Siu et al. [Siu et al., 2001b; 2004] further examine the role of the two forms of phenylalanine (**Phe**), charge-solvated (CS) versus zwitterionic (ZW), plays in the competition between π and non-π binding of $Li^+/Na^+/K^+$.

Here we extend the study on cation-π interaction into the binding of potassium cation (K^+) with prototypical dipeptides containing an aromatic amino acid, phenylalanylglycine (**FG**) and glycylphenylalanine (**GF**). In this work, the interaction between K^+ and **FG/GF** dipeptides was modeled using high-level density function theory. Theoretical geometries of free **FG/GF** and their K^+ complexes will be presented, and the potassium cation affinity (relative and absolute) for different binding modes will be compared. Our study aims to address the questions: what is the preferred site(s) of K^+ binding on dipeptides containing aromatic amino acids? How does the K^+ binding to aromatic dipeptides differ from that of aliphatic dipeptide (like

glycylglycine), and to aromatic amino acid (like phenylalanine)? What is the effect of the position of the aromatic ring (at N-terminal versus C-terminal) on the geometry and affinity of K^+ binding?

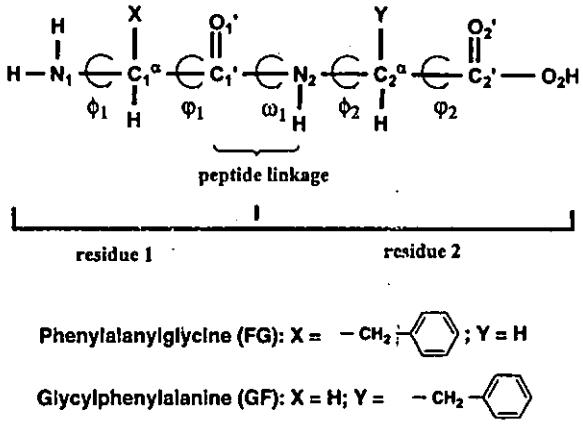
6.2 Computational Details

There are no prior high level theoretical studies on the interactions of K^+ and **FG/GF** ligands. Similar to K^+-GG/AA, (Chapter 5) the K^+ may interact with the dipeptide ligand in charge-solvated (CS) or zwitterionic (ZW) forms, and their naming system are presented in Scheme 6.1.

We adopted a Monte Carlo Multiple Minimum (MCMM) conformational searching technique [Saunders et al., 1990], outlined in details in Chapter 2, to obtain the conformation of the free ligand and their K^+ complexes, with AMBER* force field [Weiner et al., 1984] implemented in the Macromodel 7.0 package. [Mohamadi et al., 1990] While the default extended non-bonded cutoff (8, 12 and 4 Å for Van der Waals, electrostatic and hydrogen bonds, respectively) is employed for **FG/GF** in the MCMM search, the cutoff was increased to 20Å for K^+-FG/GF, identical to our previous studies on K^+-GG complexes. (Chapter 5) With an energy window of 50 kJ mol^{-1}, 251/242 **FG/GF** conformers, and 619/607 K^+-FG/GF complexes, respectively, were obtained. After discarding conformers with no intramolecular hydrogen bonds, ~40 **FG/GF** and ~100 K^+-FG/GF conformers were retained. These species were re-optimized and confirmed to be real minima by using EP(K^+) protocol (refer to Section 2.7).

The affinities at 0K (ΔH_0) for all K^+-FG/GF complexes calculated using the EP(K^+) protocol are summarized in Tables 6.1 and 6.2. In order to convert ΔH_0 to affinities at 298K (ΔH_{298}) and free energy of binding (ΔG_{298}), standard statistical thermodynamics relations [DelBene et al., 1983] (Tables 6.1 and 6.2) were employed. The deformation energy (E_{def}) and the stabilization energy ($E_{stabilization}$) for all the K^+-FG/GF complexes are also summarized in Tables 6.1 and 6.2. Polarization interaction

parameter (PIP) and dipole interaction parameter (DIP) (refer to Section 1.3.3) of the deformed ligands in the complexed states are also considered. The PIP and DIP take into account the effect of ligand framework deformation, binding distance/geometry effect along with the molecular electrostatic properties (polarizability and dipole moment), which have been found to provide better quantitative description of the interaction energy between K^+ and glycine. (Chapter 4) Finally, we wish to point out that the IUPAC naming scheme [Anonymous, 1975] is adopted here, and illustrated in Scheme 6.1.



Scheme 6.1. Schematic naming of atoms and dihedral angles of the phenylalanylglycine (**FG**) and glycylphenylalanine (**GF**) ligands according to the IUPAC recommendation. [Anonymous, 1975] Atoms and their positions in the peptide backbone are indicated by letters/ symbols/numbers in bold fonts.

6.3 Results and Discussion

6.3.1 Conformation Preference of Free Phenylalanylglycine (FG) and Glycylphenylalanine (GF) Ligands

The four most stable phenylalanylglycine (**FG**) and glycylphenylalanine (**GF**) conformers are depicted in Figs. 6.1(a) and (b), respectively.

There are no previous high level studies on the free **FG/GF** ligand. For phenylalanylglycine (Fig. 6.1(a)), the hydrogen bonding patterns exhibited in the four most stable conformers are identical to what we have found for **GG** previously, spanning a similar energy range (refer to Section 5.3.1). On the other hand, for glycylphenylalanine (Fig. 6.1(b)), the most stable conformer we obtained is **GF2**, not the analogous species of **GF1** as might have been expected. The unexpected stability of **GF2** could be due to the presence of NH...π interaction (~2.77Å in this species), a stabilizing non-covalent force implicated in biological recognition and receptor binding. [Jeffrey and Saenger, 1991; Fong et al., 1993; Koehl and Levitt, 1999; Toth et al., 2001b] More importantly, the NH...π moiety in **GF2** is part of a non-covalent OH...NH...π "daisy-chain" network of interactions (illustrated in Scheme 6.2(c)), which is believed to confer extra stability in aromatic amino acids. [Snoek et al., 2000; Robertson and Simons, 2001]

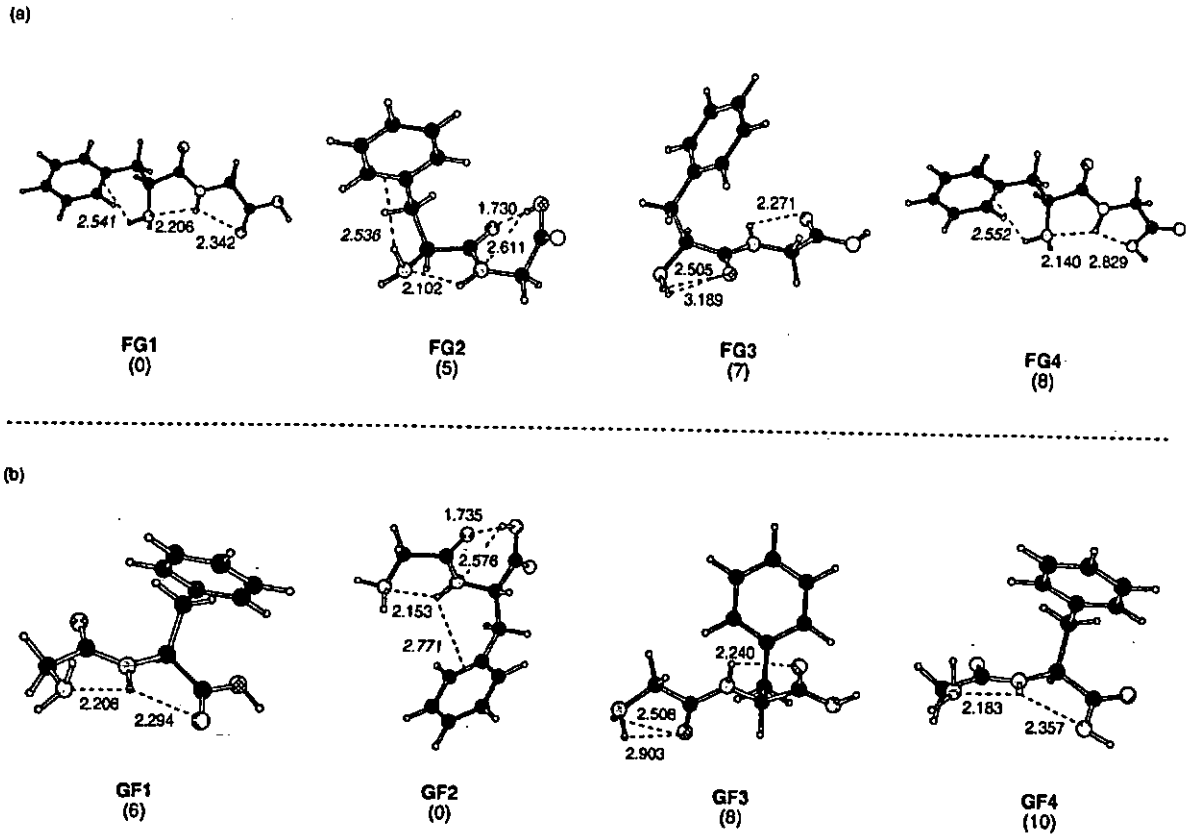
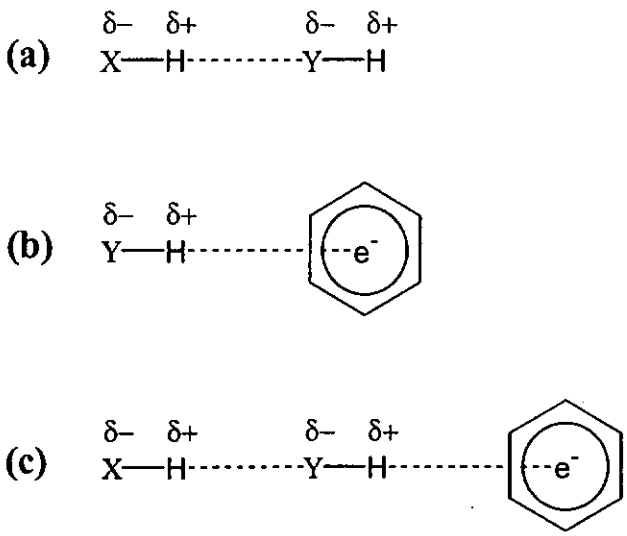


Figure 6.1. The geometries of four conformers of (a) phenylalanylglycine (**FG**) and (b) glycylphenylalanine (**GF**) ligands, optimized at the B3-LYP/6-31G(d) level. Selected non-bond distances are shown in Angstroms. The values in bracket is the relative stability (in kJ mol^{-1}) with reference to the most stable form in the respective ligands: **FG1** and **GF2**.



where X or Y = heteratom O or N of the dipeptides FG/GF

Scheme 6.2. Illustration of the "daisy-chain" effect: (a) hydrogen bonding, with atom "X" being the hydrogen bond donor and atom "Y" being the hydrogen bond acceptor, (b) YH-π interaction, (c) "daisy-chain" effect: because of the X-H...Y-H hydrogen bonding interaction, the Y-H bond would become more polarized, leading to stronger electrostatic interaction between the H and the π face.

6.3.2 Potassium Cation Binding Modes of Phenylalanylglycine and Glycylphenylalanine Complexes

Non-π Modes of Binding (Geometries)

The optimized structures of K^+-FG and K^+-GF complexes, in which the π ring is not involved in K^+ binding, are depicted in Figs. 6.2 and 6.3, respectively. All these non-π modes of binding in K^+-FG/GF are analogous to those reported for K^+-GG (Chapter 5). The average K^+ binding distance to carbonyl oxygen atoms ($\mathbf{O_1'}$, $\mathbf{O_2'}$), hydroxyl oxygen atoms ($\mathbf{O_2}$), amino nitrogen ($\mathbf{N_1}$) and the carboxylic COO^- of K^+-FG/GF ranges between 2.53 to 2.98 Å, slightly shorter (with difference less than 0.03Å) than what is found in K^+-GG previously (Chapter 5).

Similar to what is found in K^+-GG, the nitrogen of the peptide bond, $\mathbf{N_2}$, remains a unfavorable site for K^+ binding. For formamide (which is often used as the model of the peptide linkage), the most favorable binding site for $Li^+/Na^+/K^+$ is the amide oxygen atom. [Tortajada et al., 1995; Remko, 1997; Chapter 3] Furthermore, the common view is that alkali metal ions bind to peptides at one or more of the following sites: amino nitrogen, carbonyl oxygen, carboxylate oxygen atoms (of a zwitterionic peptide), but not at the amide $\mathbf{N_2}$. [Mallis and Russell, 1986; Russell et al., 1988; Grese et al., 1989; Teesch and Adams, 1991a; Kenny et al., 1997; Cerda et al., 1998; Shoeib et al., 2001] Thus, even though the amide $\mathbf{N_2}$ holds a key role in the peptide conformation, [Meyer et al., 2003; Duan et al., 1999] it may not be a favorable K^+ binding site.

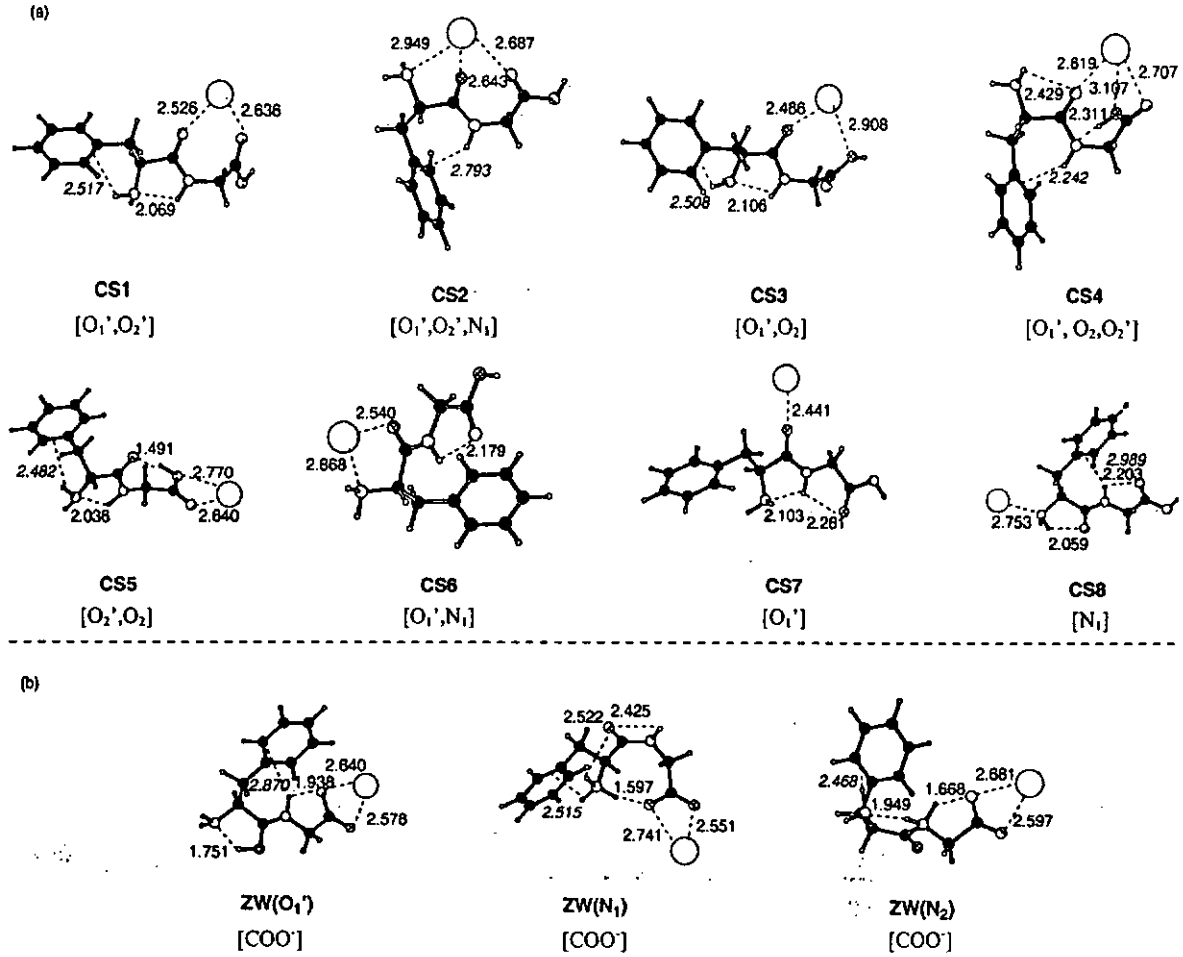


Figure 6.2. The geometries of the non-π binding modes of potassiated phenylalanylglycine (K^+-FG) complexes in the (a) charge-solvated (CS) and (b) zwitterionic (ZW) form. Selected non-bond distances are shown in Angstroms.

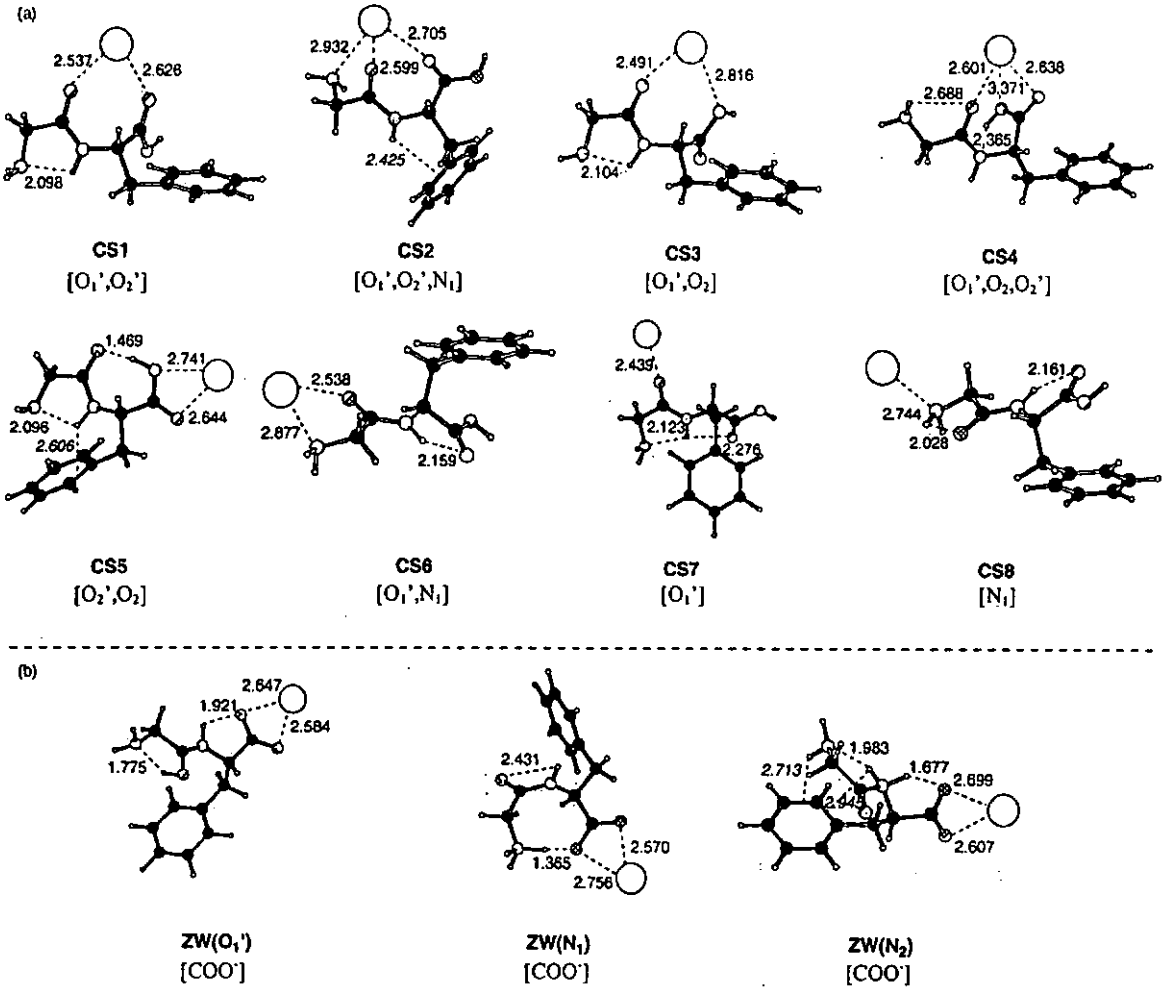
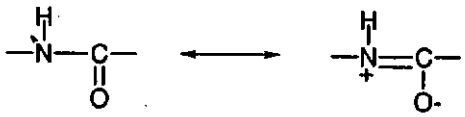


Figure 6.3. The geometries of the non-π binding modes of potassiated glycylphenylalanine (K^+-GF) complexes in the (a) charge-solvated (CS) and (b) zwitterionic (ZW) form. Selected non-bond distances are shown in Angstroms.

We would like to propose two rationales for this observation here. Firstly, because of resonance (Scheme 6.3), the amide N_2 would be expected to be less negative as compared to the amino nitrogen N_1.

Indeed, natural charge estimated at the B3-LYP/6-31G(d) level for N_2 and N_1 are –0.66 and –0.93, respectively, thus making the amide N_2 a less favorable site of K^+ binding. Secondly, binding of metal cation to the amide nitrogen would perturb the resonance stabilization (Scheme 6.3). One finds supporting evidence in the Cu^+-formamide complex [Luna, et al., 1998]. For Cu^+, binding to the amide nitrogen (species '1c' in ref. [Luna, et al., 1998]) is ~80 kJ mol^{-1} less favorable than binding to the amide oxygen (species '1a' in ref. [Luna, et al., 1998]). This is partly due to the loss of resonance stabilization, as reflected by the substantially longer CN bond (by ~0.16Å) in species '1c' than '1a'. [Luna, et al., 1998] With a larger ionic radii, the interaction of K^+ with the amide nitrogen would be even weaker, thus, unlikely to recover the energetic penalty due to the loss of resonance stabilization arising from metal cation binding.



Scheme 6.3. The resonance hybrid forms of the amide bond. Because of resonance, the amide nitrogen is expected to carry less negative charge, compared to the amino nitrogen.

Non-π Modes of Binding (Energetics)

The binding affinity at 0 and 298K (ΔH_0 and ΔH_{298}), free energy of binding (ΔG_{298}), deformation energy (E_{def}) and stabilization energy ($E_{stabilization}$) are summarized in Table 6.1.

Given the very similar binding characteristics (discussed in the previous section), and the fact that the π ring is not involved in K^+ binding, the variation of ΔH as a function of these non-π modes is expected to be parallel to that of K^+-GG. (Chapter 5) This is, in fact, what is found in general (Fig. 6.4(a)), with an average increase in absolute K^+ affinity of 4-5 kJ mol^{-1} from **GG** to **FG/GF**, with the only notable exception noted for the **CS7** mode of binding. For this mode (K^+ binds monodentate to O=C), we found that the K^+ affinity has *decreased* by 5 and 3 kJ mol^{-1}, for **FG** and **GF**, respectively, when compared to **GG** (Fig. 6.4(a)). We wish to note here that, for the **CS7** mode, the K^+ affinities for alanylalanine (Chapter 5) and glycylproline [Abirami et al.2004b] are also *smaller* than that of **GG**. Thus it appears that that K^+ binding to O=C in glycylglycine is unexpectedly strong (when compared to other aliphatic dipeptides), even though the origin of this abnormality is not clear.

Table 6.1. The theoretical energetics (in kJ mol^{-1}) of non-π binding modes in potassiated phenylalanylglycine (K^+-FG) and glycylphenylalanine (K^+-GF) complexes.

		ΔH_0 [a]		ΔH_{298} [b]		ΔG_{298} [b]		E_{def} [c]		$E_{stabilization}$ [d]	
Species	Site of binding	K^+-FG	K^+-GF	K^+-FG	K^+-GF	K^+-FG	K^+-GF	K^+-FG	K^+-GF	K^+-FG	K^+-GF
CS1	O_1', O_2'	152	160	154	161	120	127	30	38	182	197
CS2	O_1', O_2', N_1	154	149	155	151	119	113	47	63	201	212
CS3	O_1', O_2	123	133	124	134	92	101	20	29	142	162
CS4	O_1', O_2, O_2'	109	115	111	116	74	85	66	77	175	192
CS5	O_2', O_2	140	143	143	145	106	111	8	13	148	156
CS6	O_1', N_1	139	136	140	138	108	105	32	40	171	177
CS7	O_1'	122	124	123	124	96	96	6	14	129	138
CS8	N_1	80	79	81	80	53	53	17	21	98	100
ZW(O_1')	Carboxylate COO^-	105	107	107	109	73	75	160	153	264	260
ZW(N_1)	Carboxylate COO^-	111	94	114	96	77	60	86	83	197	177
ZW(N_2)	Carboxylate COO^-	52	63	54	65	22	31	160	151	212	214

[a]: Calculated by the EP(K^+) protocol. [b]: Standard thermodynamic relations [DelBene et al., 1983] were applied to obtain the affinities (ΔH_{298}) and free energy of binding (ΔG_{298}) at 298K. [c]: Refer to Eqn. [1.9]. [d]: Refer to Eqn. [1.10].

The molecular polarizability of the **FG/GF** dipeptides (~17Å^3) are estimated to be more than twice of that of **GG** (~8Å^3). Thus, based on this factor alone, one would expect the K^+ affinity of **FG/GF** to be substantially larger than that of **GG**. Thus, the rather small increase in affinity (~4-5 kJ mol^{-1}) may be due to the distance factor (the increased distance between K^+ and the center of charge) and/or repulsion between the nucleus of K^+ and the various atoms on the ligand. Here we tried to correlate the strength of interaction between K^+ and **GG/FG/GF** dipeptides with polarization interaction parameter (PIP) and dipole interaction parameter (DIP). Despite the success for K^+ and glycine, (Section 4.2.4) one finds no correlation between the PIP and DIP with the variation in K^+ affinity of the three dipeptides (Fig. 6.4(b) and (c)). In terms of effect of molecular dipole moment on K^+ affinity, one finds that the DIP (Fig. 6.4(c)) of **GG** is consistently larger than that for **FG/GF**. In other words, based on DIP alone, one may expect that the K^+ affinity of **GG** to be higher than that of **FG/GF**, which is opposite to what is in generally observed here. The failure of DIP and PIP in describing the variation of K^+ affinity of **GG/FG/GF** may not be too surprising as the electron distribution in these species are more complex than in glycine, so that simple parameters like PIP and DIP are not sufficient to accurately describe the underlying physico-chemical basis of the interaction.

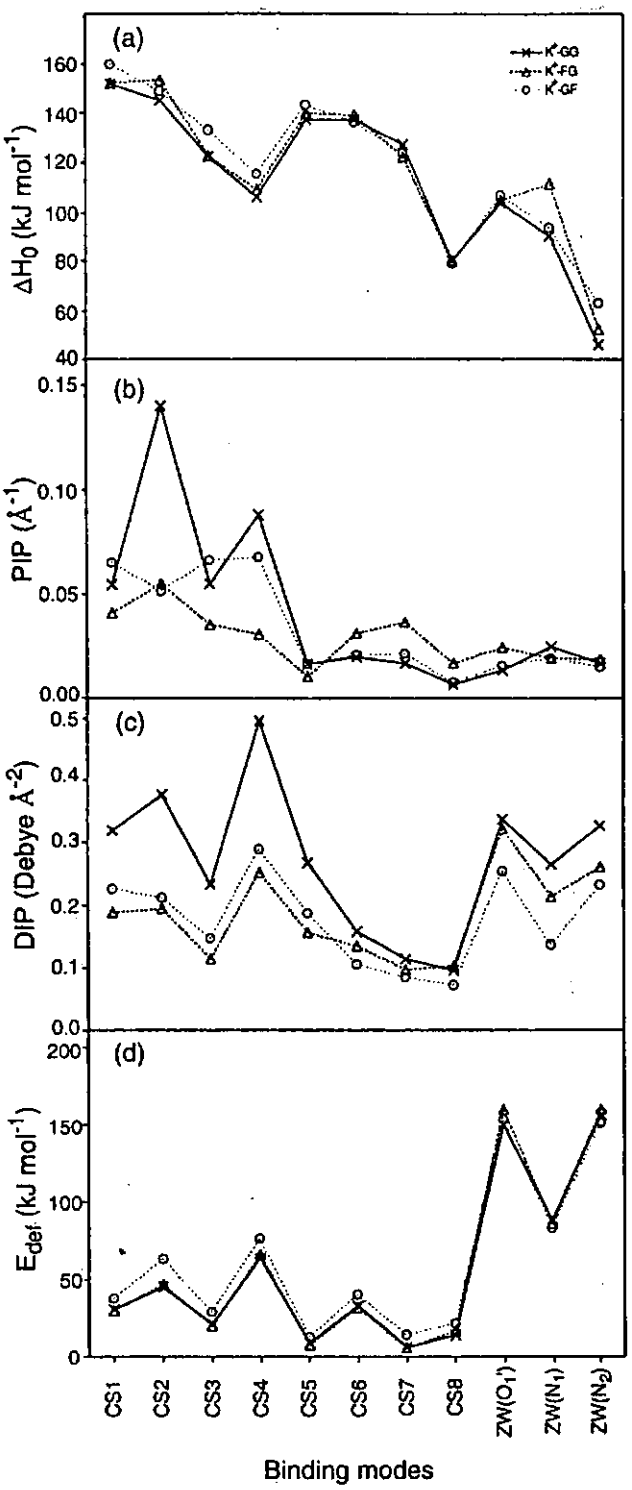


Figure 6.4. The variation of (a) binding affinities, ΔH_0, (b) polarization interaction parameter, PIP, (c) dipole interaction parameter, DIP and (d) deformation energy, E_{def}, as a function of the non-π mode of binding for K^+-GG (cross, with solid line), K^+-FG (triangle, with dotted line), and K^+-GF (circle, with dotted line).

Interestingly, the absolute affinity of **ZW(N_1)** in K^+-FG and **ZW(N_2)** in K^+-GF are noticeably higher (~ 20 kJ mol^{-1}) when compared to the analogous species in K^+-GG. (Section 5.3.4) The species **ZW(N_1)** is so stable that it becomes the most stable zwitterionic species in K^+-FG, surpassing the stability of **ZW(O_1')**. This could be attributed to the relative change in E_{def} for these two species (Table 6.1). In fact, when the variation of E_{def} of these non-π modes in K^+-FG/GF are compared (Fig. 6.4(d)), two quite distinct trends can be observed:

- for the non-π CS modes, the E_{def} of K^+-FG is *smaller* when compared to K^+-GF (by 4-16 kJ mol^{-1})
- for the non-π ZW modes, the E_{def} of K^+-FG is *greater* when compared to K^+-GF (by 3-9 kJ mol^{-1})

In other words, replacing glycine with phenylalanine at the C-terminal (i.e., in **GF**) tends to disfavor K^+ binding in the CS mode but to favor the ZW mode, even though the phenyl π ring is not involved in the cation binding directly.

This observation may be attributed to the interaction between the aromatic π ring and other binding sites. In the CS mode, as the K^+ prefers to bind to the oxygen atoms, the cation would prefer to bind to the oxygen rich C-terminal side. When glycine is substituted with phenylalanine, the electron-rich π ring is in the same residue as the carboxylic oxygen atoms, and the electrostatic repulsion between the sites would increase E_{def}.

On the other hand, the electron-rich π ring may stabilize the positively charged part of the dipeptide in the zwitterionic form, i.e., via cation-π interaction. It is enlightening to note that in the case of **ZW(O_1')**, the proton attached to **O_1'** is interacting with the

N-terminal amino group in both K^+-GG/FG/GF. Without cation-π interaction (in this case, $OH^+...\pi$), the E_{def} of **ZW(O_1')** in K^+-FG/GF is larger than that for K^+-GG. In species where $NH^+...\pi$ interaction are present (2.52Å in **ZW(N_1)** of K^+-FG in Fig. 6.2, 2.95Å in **ZW(N_2)** of K^+-GF in Fig. 6.3), the E_{def} is in fact smaller than what is found in K^+-GG.

In conclusion, possibly due to geometrical constraints, the π ring of phenylalanine is involved in more repulsive interactions in the CS mode, but more attractive interactions in the ZW mode, when it is situated at the C-terminal (**GF**) than when it is at the N-terminal (**FG**).

π-Modes of Binding (Geometries)

Most of the non-π binding modes depicted in Figs. 6.2 and 6.3 have corresponding π binding modes, in which the K^+ forms *additional* binding/interaction with the aromatic-π ring (Figs. 6.5 and 6.6 for K^+-FG and K^+-GF, respectively, with energetics summarized in Table 6.2). Compared to the corresponding non–π modes, the average K^+ binding distance to **O_1'**, **O_2'**, **O_2**, **N_1** and the carboxylic COO^- in the π modes are slightly longer (maximum of 0.12 Å in K^+-FG/GF). The K^+-π distance ranges from 2.92-3.84 and 2.96-3.73 Å in **FG** and **GF**, respectively, again longer than that reported for K^+-Phe previously (2.83-3.15Å) [Siu et al., 2004]. Thus, in order to attain larger number of interactions, the interaction between K^+ and individual binding sites is weakened from K^+-Phe to K^+-FG/GF.

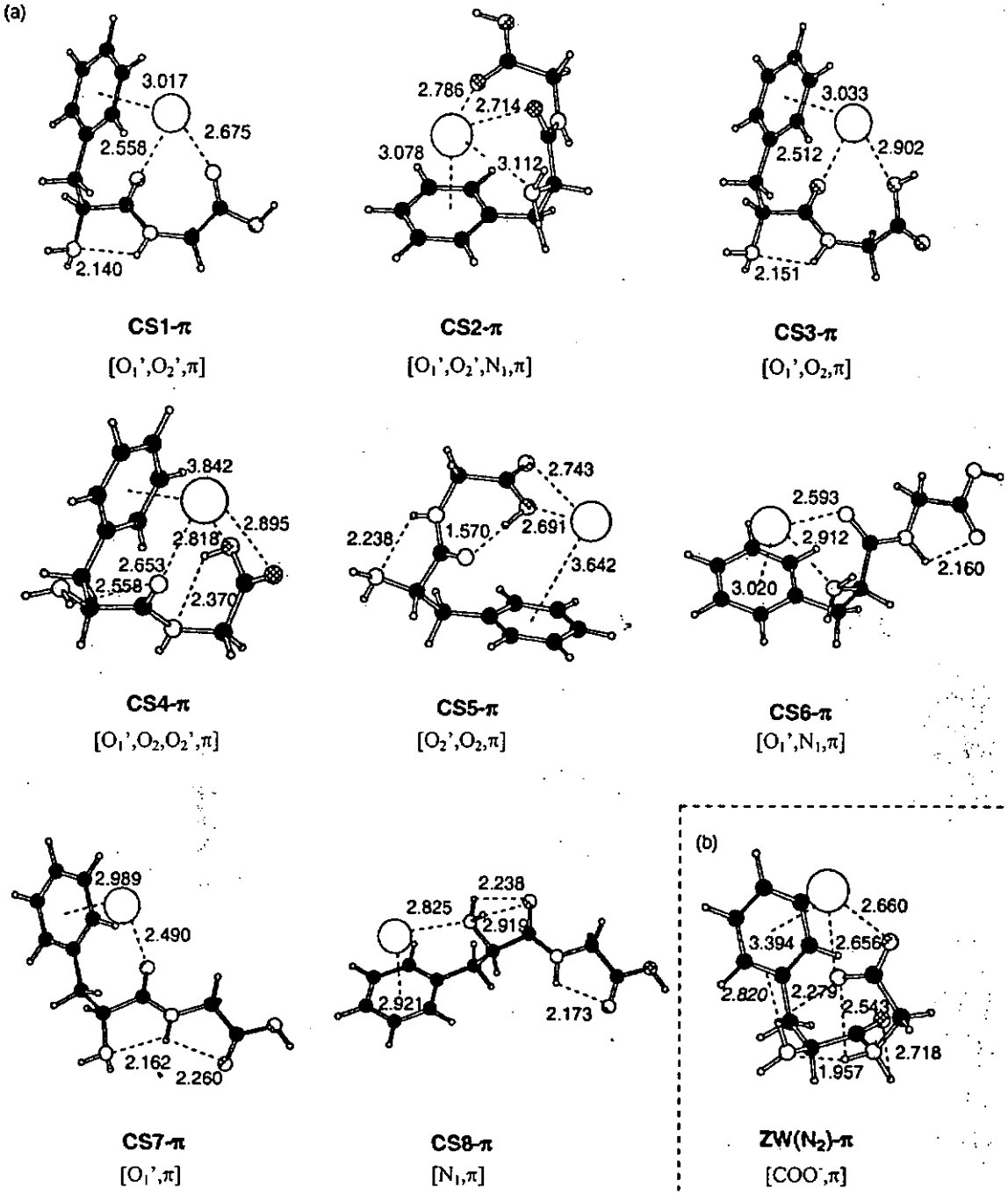


Figure 6.5. The geometries of the π binding modes of potassiated phenylalanylglycine (K^+-FG) complexes in the (a) charge-solvated (CS) and (b) zwitterionic (ZW) form. Selected non-bond distances are shown in Angstroms.

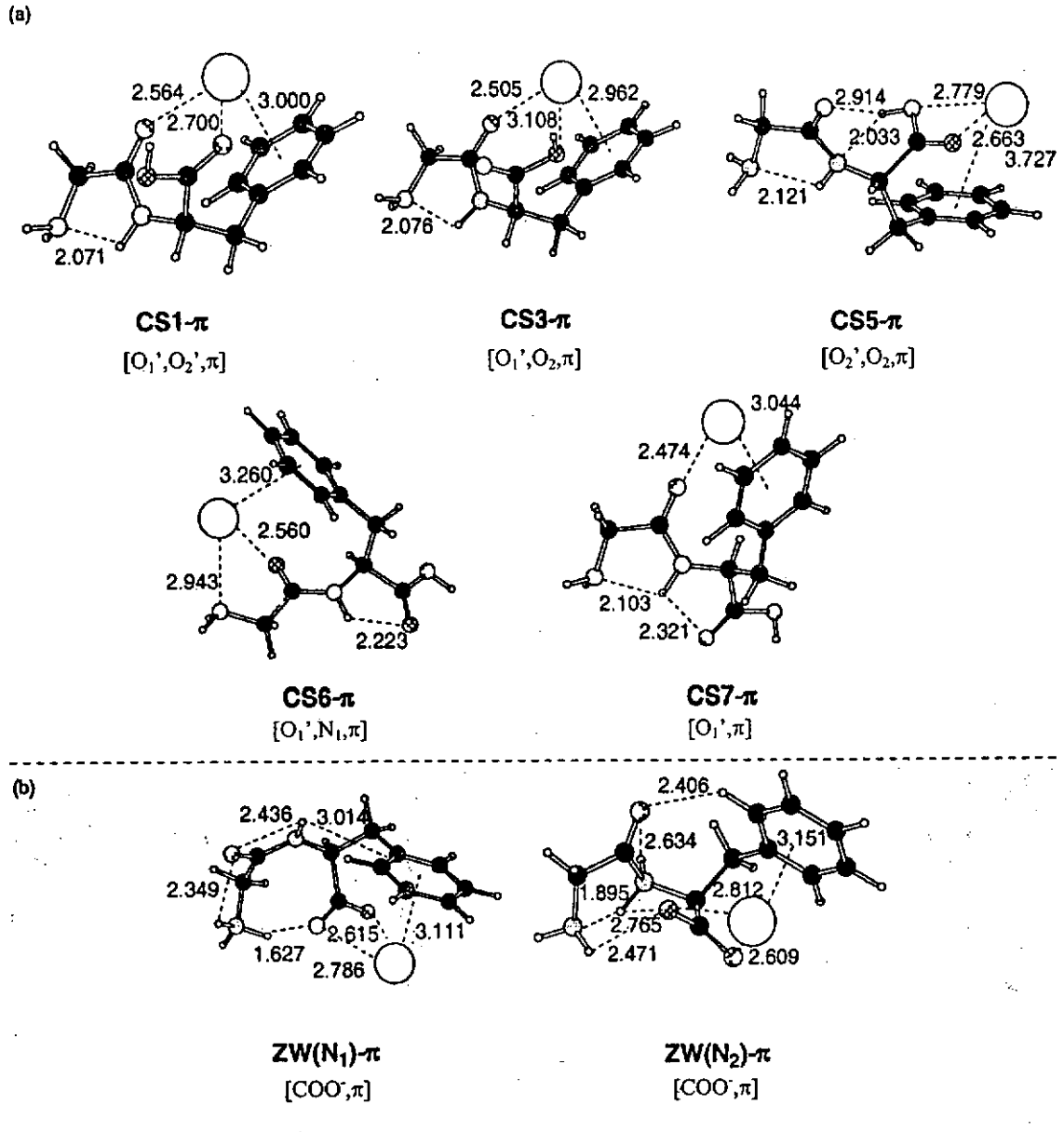


Figure 6.6. The geometries of the π binding modes of potassiated glycylphenylalanine (K^+-GF) complexes in the (a) charge-solvated (CS) and (b) zwitterionic (ZW) form. Selected non-bond distances are shown in Angstroms.

Table 6.2. The theoretical energetics (in kJ mol^{-1}) of the π binding modes in potassiated phenylalanylglycine (K^+-FG) and glycylphenylalanine (K^+-GF) complexes.

		ΔH_0 [a]		ΔH_{298} [b]		ΔG_{298} [b]		E_{def} [c]		$E_{stabilization}$ [d]	
Species	Site of binding	K^+-FG	K^+-GF	K^+-FG	K^+-GF	K^+-FG	K^+-GF	K^+-FG	K^+-GF	K^+-FG	K^+-GF
CS1-π	O_1', O_2', π	167	158	169	159	129	121	42	53	209	210
CS2-π	O_2', O_1', N_1, π	147	—	149	—	110	—	62	—	210	—
CS3-π	O_1', O_2, π	136	129	137	131	100	92	38	53	174	183
CS4-π	O_1', O_2, O_2',π	108	—	110	—	71	—	72	—	180	—
CS5-π	O_2, O_2', π	127	102	129	103	90	68	33	49	160	151
CS6-π	O_1', N_1, π	156	150	158	152	121	114	49	46	205	196
CS7-π	O_1', π	151	152	152	153	118	118	20	24	171	176
CS8-π	N_1, π	117	—	118	—	85	—	24	—	141	—
ZW(O_1')-π	Carboxylate COO^-, π	—	—	—	—	—	—	—	—	—	—
ZW(N_1)-π	Carboxylate COO^-, π	—	85	—	89	—	45	—	123	—	208
ZW(N_2)-π	Carboxylate COO^-, π	43	56	46	60	2	15	165	176	207	233

[a]: Calculated by the EP(K^+) protocol. [b]: Standard thermodynamic relations [DelBene et al., 1983] were applied to obtain the affinities (ΔH_{298}) and free energy of binding (ΔG_{298}) at 298K. [c]: Refer to Eqn. [1.9]. [d]: Refer to Eqn. [1.10].

We note that not all non-π mode of binding can form additional π interaction with the K^+. As an example, in **GF**, the corresponding charge-solvated π mode for **CS2**, **CS4**, and **CS8** (i.e., **CS2-π**, **CS4-π** and **CS8-π**) cannot be located. It appears that geometry of binding plays an important role in governing the formability/stability of the π binding modes:

- the π ring could be too far from the other binding site, e.g., for **CS2**, the π-ring is too far away from the other sites ($\mathbf{O_1'}$, $\mathbf{O_2'}$, $\mathbf{N_1}$) for the K^+ to bind to all four sites simultaneously; or

- the π ring could be too close from the other binding sites, e.g., sensible trial structures for the **CS8-π** binding mode (with the K^+ interacting bidentately with $\mathbf{N_1}$ and π) invariably optimized to the **CS6-π** mode (with the K^+ binds to $\mathbf{O_1'}$, $\mathbf{N_1}$, and π) as the carbonyl oxygen $\mathbf{O_1'}$ site is close by.

It also appears that the K^+ has a low preference to bind to carbonyl oxygen, hydroxyl oxygen and aromatic π on the *same* residue/amino acid. In the case of K^+-Phe [Siu et al., 2004], the K^+...π distance for such mode of binding is ~ 0.2Å longer (**CS6** in ref. [Siu et al., 2004]) than the others. Here, for the **GF** dipeptide, the **CS4-π** mode of binding could not be located. This probably arises from geometrical constraints as when the O=C/OH and π ring are located on *different* residues, as in the case of **FG**, such modes are found to be fairly stable, making additional binding to the π-ring less preferred.

As one can see from Table 6.2, the ZW forms of these K^+-dipeptides are energetically not competitive against the CS modes, even in the presence of cation-π interaction. For the ZW modes to engage in K^+–π interaction while maintaining the metal cation

binding to the COO^- group means that the π ring and the carboxylate group has to be in close proximity. Such conformation is expected to be unfavorable in general because of the strong electrostatic repulsion between the binding sites. While this hypothesis is supported by the fact that such mode of binding could not be located in K^+-Phe, [Siu et al., 2004] we found cases that K^+ can simultaneously bind to both carboxylate oxygen atoms and the π ring: **ZW(N_1)-π** in K^+-GF; and **ZW(N_2)-π** in K^+-FG/GF. In these ZW complexes, we notice that the positively charged nitrogen (N_1 for **ZW(N_1)-π** or N_2 for **ZW(N_2)-π**) is at close vicinity to the carboxylate group (with NH_2^+...OCO^- distance ranges from 1.63 to 2.77 Å). Such electrostatic interaction not only stabilize the zwitterionic mode in general, but may also help to disperse the negative charges on the COO^-, thus, reducing the destabilizing electrostatic repulsion between COO- and the π ring to allow the formation of these ZW-π modes.

π-Modes of Binding (Energetics)

In this section, we would like to discuss the effect of additional π interaction on the K^+ binding energetics. As not all corresponding π modes can be located, the number of species that we can compare is limited. However, some trends/patterns can be observed.

The effect of additional π interaction on the binding affinity for CS modes and ZW modes are quite different (Figs. 6.7(a) and 6.7(c)) for K^+-FG and K^+-GF complexes, respectively). For the ZW modes, additional π interaction reduces the binding affinity, while for the CS modes, such interaction *tends to* enhance the affinity, with the monodentate modes (**CS7** and **CS8**) benefit most from the additional π interaction (by 30 to 40 kJ mol^{-1}). Reduction of affinity is observed in some CS modes, but they are generally small (1 to 6 kJ mol^{-1}), with the only exception being **CS5** in which the

decrease in affinity due to additional π interaction can be up to ~40 kJ mol^{-1} for K^+-GF.

Thus, additional π interaction in **FG/GF** dipeptide does not *always* increase the K^+ binding affinity when compared to the corresponding non-π mode. This could be largely attributed to the effect of deformation energy. Comparing the π and non-π mode, the E_{def} is always larger (by 6 to 40 kJ mol^{-1}) for the π complex (Figs. 6.7(b) and 6.7(d)). Thus, the unfavorable factor would partially outweigh, and in some cases offset, the increase in stabilizing forces from K^+-π interaction, leading to the observed difference in ΔH between π and non-π modes.

Moreover, the effect is rather dependent on the mode of binding, which is similar to what is found in K^+-Phe: when K^+ binds to NH_2 and O=C, additional π interaction increase the binding affinity by ~20 kJ mol^{-1}; [Dunbar, 2000] however, when K^+ binds to OH and O=C, additional π interaction is in fact destabilizing. [Siu et al., 2004]

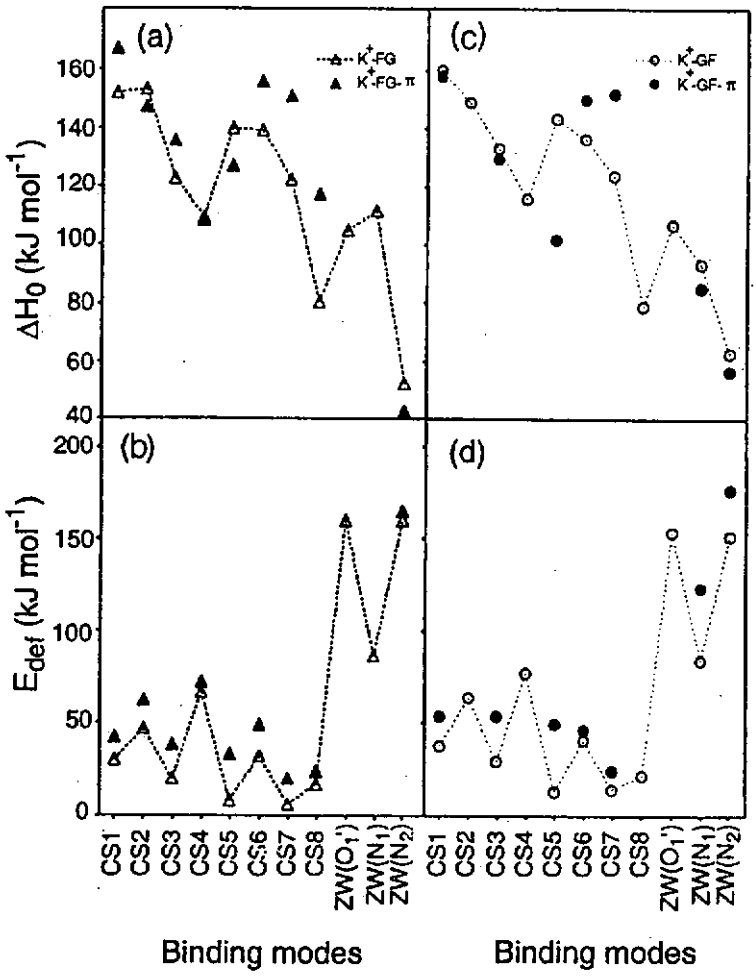


Figure 6.7. The variation of (a) binding affinities, ΔH_0 (b) deformation energy, E_{def}, as a function of binding mode: non-π mode with open triangle, dotted line and π mode with solid triangle in K^+-FG complexes. The corresponding quantities (ΔH and E_{def}) for K^+-GF are depicted in (c) and (d), respectively, with the non-π mode denoted in open circle, dotted line and π mode in solid circle.

6.3.3 Potassium Cation Affinity of Phenylalanylglycine (FG) and Glycylphenylalanine (GF)

In the previous section, we noted that the K^+ binding affinity of a few π modes are found to be less stable than the corresponding non-π modes. Most importantly, the most stable mode of binding is found to be the **CS1** mode in K^+-GF (with the **CS1-π** mode less stable by about 2 kJ mol^{-1}). Thus, it remains unclear whether cation-π interaction is present in the most stable mode of binding in K^+-GF; and whether such interaction is stabilizing or destabilizing for the **CS1** complex.

In order to shed some light into these questions, we conducted additional single point energies calculation with several other theoretical models (Table 6.3). All these single points calculations are conducted with the 6-311+G(3df,2p) basis set, using the B3-LYP/6-31G(d) geometries, and corrected with zero-point-vibrational energy effect at the HF/6-31G(d) level. The effect of BSSE correction is also investigated.

Table 6.3. Estimated K^+ affinities (ΔH, kJ mol^{-1}) for GF and FG systems obtained by different exchange-correlation functionals.

Level of theory	FG		GF	
	$\Delta H(\pi)$ [a]	$\Delta(\Delta H)$ [b]	$\Delta H(\pi)$ [a]	$\Delta(\Delta H)$ [b]
B-LYP	159	11	149	-5
B3-LYP	167	15	158	-2
B3-P86	169	18	158	3
B1-LYP	168	15	159	-2
MPW1-PW91	169	19	159	3
PW91-PW91	174	19	161	6
BHandH-LYP	176	19	169	1
MP2(full)	192	34	186	23
B-LYP [c]	157	10	147	-6
B3-LYP [c]	165	14	156	-3
B3-P86 [c]	167	16	156	2
B1-LYP [c]	165	14	157	-3
MPW1-PW91 [c]	167	17	157	2
PW91-PW91 [c]	171	18	159	5
BHandH-LYP [c]	174	18	167	0
MP2(full) [c]	181	29	175	18
Experiment [d]	155		152	

[a]: K^+ affinities of **CS1-π** at different level of theory, with basis set 6-311+G(3df,2p) for single point energy calculation, with geometry optimized at B3-LYP/6-31G(d), corrected with zero-point vibrational energies calculated at HF/6-31G(d). [b]: The difference in affinity between **CS1-π** and **CS1** modes (a positive $\Delta(\Delta H)$ indicates that the **CS1-π** is more stable than **CS1**). [c]: ΔH and $\Delta(\Delta H)$ at various levels of theory, with BSSE correction. [d]: Experimental K^+ affinities determined by kinetic method performed by Dr. Y. Tsang.

Because of the size of the system, the calculation we conducted is neither systematic nor exhaustive. However, some general conclusions about the *absolute* affinity of these species can be drawn:

(1) All models agree that the K^+ affinity of **FG** is larger than that of **GF**, which is in general agreement with the experimental values of 155 and 152 kJ mol^{-1}, respectively, but tend to be too large.

(2) The K^+ affinity estimated by B3-LYP, B3-P86, B1-LYP and MPW1-PW91 functionals are very similar while that estimated by PW91-PW91 and BHandH-LYP tends to be higher. This is in agreement with what is found previously for K^+ affinity of proline [Marino et al., 2003].

(3) In agreement with previous works [Rodgers and Armentrout, 2000b; Iceman and Armentrout, 2003; Chapter 4], BSSE is large for the MP2 method (6-11 kJ mol^{-1}) but small (0-3 kJ mol^{-1}) for DFT. Even with BSSE correction, the K^+ affinity estimated at the MP2 level is still larger than the DFT estimates.

Interestingly, the *relative* stability of the **CS1** versus **CS1-π** mode in K^+-GF is very sensitive to the level of theory. At some level of theory (BLYP, B1-LYP and B3-LYP), the non-π mode is marginally more stable, while at other DFT levels (PW91PW91, B3-P86 etc.), the π mode could be marginally more stable. In fact, similar trends are also observed in K^+-FG but because the energy difference of the two modes, $\Delta(\Delta H)$, is larger (> 10 kJ mol^{-1}), the **CS1-π** mode is found to be *consistently* more stable than the **CS1** mode. Based on the above, we propose an error

bar of ±3 kJ mol^{-1} for the *relative* affinity estimated by our B3-LYP based model (obtained by averaging all the DFT Δ(ΔH) values for K^+-GF).

The energy difference between **CS1** and **CS1-π**, Δ(ΔH), estimated by the MP2 model, is large, strongly favoring the complex with cation-π interaction. However, the trend is the same: the **CS1-π** mode is dominant in K^+-FG, while comparatively, the **CS1** and **CS1-π** modes are energetically more competitive in K^+-GF. Comparing the K^+...π distance of **CS1-π** in K^+-FG (~3.00Å) with that in K^+-GF (~3.02Å), the latter one is marginally longer so that the stability of the **CS1-π** in K^+-GF is unlikely to arise from stronger cation-π interaction. We note that for K^+-FG, the binding distance from K^+ to the C-terminal carbonyl oxygen (O_2') has only increased by ~0.04Å from **CS1** (2.64Å) to **CS1-π** (2.68Å) complex; while the corresponding distance in K^+-GF has lengthened by ~0.07Å. Furthermore, the *difference* in deformation energy (E_{def}) for **CS1-π** (Table 6.2) versus **CS1** (Table 6.1) for the two species for K^+-FG is also slightly smaller than that of K^+-GF (15 and 12 kJ mol^{-1}, respectively). Thus, for the **CS1-π** mode in K^+-GF, the gain in stability arising from cation-π interaction is offset by the loss of stabilization in K^+...O_2' and the larger E_{def}, so that the **CS1** and **CS1-π** modes in this species are of similar stability.

6.4 Conclusion

A density functional study of phenylalanylglycine (**FG**) and glycylphenylalanine (**GF**) dipeptides and their potassiated complexes has been conducted. It is found that, while the relative stability of the **FG** conformers closely resembles that of glycylglycine (**GG**), a different conformer, stabilized by a "daisy-chain" OH...NH...π network, is preferred in **GF**.

All basic sites on the **FG/GF** dipeptides ($\mathbf{O_1'}$, $\mathbf{O_2'}$, $\mathbf{O_2}$, $\mathbf{N_1}$, the carboxylic COO^-, and the aromatic-π ring) are accessible to K^+ binding, but the binding to amide nitrogen ($\mathbf{N_2}$) is in general not favored. Some plausible reasons have been proposed.

For the non-π modes of binding, the K^+ binding characteristics are similar to that reported previously in K^+-GG (Chapter 5), with binding affinity increased by ~4 kJ mol^{-1}. Unlike K^+-glycine (Chapter 4), we found that the polarization interaction parameter and dipole interaction parameter has no correlation with the K^+ affinities of **GG/FG/GF**. From the analysis of deformation energy, we found that replacing the glycine with phenylalanine at the C-terminal (i.e., **GG** to **GF**) tends to disfavor K^+ binding in the CS mode but to favor the ZW mode, even though the phenyl π ring is not involved in the cation binding directly.

Because of geometrical constraints and electronic factors, not all of the non-π modes may engage in additional K^+-π interactions. While the formation of ZW-π modes are not favorable in K^+-phenylalanine due to electrostatic repulsion, [Siu et al., 2004] for the **FG/GF** dipeptides, some of the negative charges on the COO^- may be dispersed via the $NH_2^+...OCO^-$ interaction. This reduces the destabilizing electrostatic

repulsion between COO- and the π ring, hence facilitating the formation of ZW-π modes.

Comparing the π and non-π mode for the **FG/GF** dipeptides, the deformation energy is *always* larger for the π complex. Thus, the unfavorable factor would partially outweigh the increase in stabilizing forces from K^+-π interaction so that the additional π interaction does not necessarily enhance the binding affinity. As a result, we found that while the **CS1-π** mode (with K^+ binds the aromatic-π ring and the two carbonyl oxygen atoms) is the most favorable mode of binding for K^+-FG, for K^+-GF, the **CS1** mode (with K^+ binds the two carbonyl oxygen atoms) is of comparable stability with the **CS1-π** mode. Finally, we note that the absolute K^+ affinity for the most stable mode of binding in **FG (CS1-π)** and **GF (CS1)** are 167 and 160 kJ mol^{-1}, respectively, in fair comparison with the experimental affinities recently determined by the mass spectrometry kinetics method (155 and 152 kJ mol^{-1}, respectively).

Chapter 7 Fragmentation of Cationized Histidine in the Gas Phase: Protonation versus Potassiation

7.1 Background

Mass spectrometric studies of protonated/alkaliated peptides/proteins yield important information about these biomolecules. Under favorable conditions, the fragmentation of these cationized species allows the determination of the sequence of amino acids in a given peptide. [Loo and Muenster, 1999; Hoyes and Gaskell, 2001; Huang et al., 2002b] Furthermore, H/D exchange studies provide information on the number of active (labile) hydrogen atoms on the surface of the macromolecule, thus yielding important structural information on the proteins/peptides. [Roepstorff, 1997; Dongre et al., 1997; Yates, 1998; Figueroa and Russell, 1999]

In a typical mass spectrometric experiment on determination of peptide sequence, the sample of interest is first cationized with a proton or a metal cation. Under metastable and low-energy collision-induced dissociation (CID) conditions, characteristic fragment ions are formed along with loss of small stable neutrals, such as H_2O, CO and NH_3. [Rogalewicz et al., 2000c; Farrugia et al., 2001a; 2001b; Aribi et al., 2003] As the fragment ions formed are dependent on the constituent amino acids, [Grese and Gross, 1990; Yalcin et al., 1995], it is of both fundamental and practical interest to elucidate the fragmentation mechanisms of cationized amino acids.

Our research group has previously conducted a combined experimental and theoretical study on the fragmentation of a prototypical cationized aliphatic amino acid, i.e., protonated and alkaliated alanine. [Abirami et al., 2004] Here, we extend the study to an amino acid with a functionalized side-chain, histidine (**His**).

Experimental studies on protonated histidine $[His + H]^+$ suggested that, like many protonated amino acids, the loss of a neutral of 46 u, $[CO + H_2O]$, is dominant. [Kulik and Heerma, 1988; Dookeran et al., 1996; Rogalewicz et al., 2000c] With the exception of histidine, deuterium-labeling experiments indicated that all α-amino acids lost $[CO + D_2O]$ exclusively from $[ND_2RHCOOD + D]^+$. [Rogalewicz et al., 2000c] On the other hand, protonated histidine also showed other minor fragment ions due to the loss 17 u (NH_3), 18 u (H_2O) and 61 u ($[CO_2 + NH_3]$), depending on the experimental conditions. [Kulik and Heerma, 1988; Dookeran et al., 1996; Rogalewicz et al., 2000c] This shows that the fragmentation of protonated histidine is different from the other amino acids. Theoretical studies on protonated histidine derivatives (methylated histidine [Turecek et al., 1998], histidine methyl esters [Farrugia et al., 2001a] and N-acyl methyl ester histidine [Farrugia et al., 2001b], etc.) have been reported. However, the scopes of these studies are limited, and the complete fragmentation pathways and mechanisms have not been elucidated.

On the other hand, experimental studies in our group revealed that the formation of m/z 39 (K^+) from $[His + K]^+$ was dominant, with other less intense fragment ions arising from the loss of NH_3 (17 u) and CO_2 (44 u) were also observed, depending on the experimental conditions [Wong, 2003; Abirami et al., 2004]. These results clearly illustrate that the products formed from $[His + K]^+$ and $[His + H]^+$ are very different. Thus far, there are only limited studies on the fragmentation of metal cationized basic amino acids like histidine in the literature; the previous focus lies mainly on transition metal cations likes Cu^+ [Lavanant and Hoppilliard, 1997; Caraiman et al., 2003], Ag^+ [Shoeib et al., 2001b; Talaty et al., 2001; Shoeib et al., 2002; Caraiman et al., 2003] and Ni^+ [Rodriguez-Santiago et al., 2001], etc. Hoppilliard and co-workers studied the dissociation of $[(His - H) + M]^{2+}$ (M = Fe, Co, Ni, Cu and Zn) [Lavanant et al.,

1999] and found the characteristic $[(His - H) + M - CO_2]^+$ fragment ion derived from decarboxylation of the amino acid. More recently, Hoppillard analyzed the pathways for the losses of CO_2, $[CO + H_2O]$ and CO from $[(Gly - H) + Zn]^{2+}$. [Lavanant et al., 1999] However, the fragmentation of mono-cationized and di-cationized amino acids [Lavanant et al., 1999; Rogalewicz et al., 2001; Hoppilliard et al., 2001] are expected to be quite different. Firstly, di-cationized species is expected to be more reactive in the gas phase when compared to their mono-cationized counterpart as di-cationic species are intrinsically unstable in the gas phase. Secondly, the interaction between a di-cation and a ligand is much stronger than that between a mono-cation and the ligand. With such strong interaction, a cation "insertion" mechanism, in which the di-cation may be inserted between carbon-carbon bonds, is viable. [Eller and Schwarz, 1991] Such "insertion" mechanism would be unlikely to occur in mono-cation systems as the interaction is too weak to compensate for the energy penalty for breaking up the covalent bonds in the ligand.

In this Chapter, detailed theoretical study on the fragmentation mechanism of $[His + H]^+$ for the loss of $[CO + H_2O]$, H_2O, and $[CO_2 + NH_3]$ will be presented. With the time available and the complexity of the system, only a partial theoretical study on $[His + K]^+$ was conducted. Even though the scope of study on $[His + K]^+$ is limited only to the relative stability of a few potassiated histidine complexes, we will show that useful insight can be gained in terms of understanding the fragmentation of the $[His + H]^+$ versus $[His + K]^+$ systems.

7.2 Computational Details

Density functional calculations were performed using the GAUSSIAN98 package [Frisch et al., 1998]. Structures for both the protonated and potassiated histidine systems were fully optimized at the B3-LYP/6-31G(d) level, and the nature of stationary points were confirmed by frequency calculation. For protonated histidine in which a full potential energy surface study is conducted, transition structures are verified to be connecting the specific minimum by Intrinsic Reaction Coordinate (IRC) calculations.

With these structures, single point energy calculations were performed at the B3-LYP/6-311+G(3df,2p) level. The electronic energies at 0K of all species were corrected with B3-LYP/6-31G(d) zero–point-vibrational-energies, scaled by 0.9806. [Scott and Radom, 1996] All discussions below are based on energetics at 0K unless otherwise noted. To obtain the enthalpy (ΔH) and Gibbs free energy (ΔG) of the various species at finite temperatures, standard thermodynamic relations [DelBene et al., 1983] were applied, calculated from the scaled B3-LYP/6-31G(d) vibrational frequencies.

7.3 Results and Discussion

7.3.1 Low-Energy Collision-induced Dissociation (CID) of Protonated Histidine

The experimental results presented in this Chapter were obtained by Dr C. L. Wong. The original spectrum can be found in ref. [Wong, 2003], but are reproduced here for easy reference.

Fig. 7.1 presents the MS/MS spectra of protonated histidine, $[His + H]^+$ (m/z 156), obtained under low-energy CID (MS/MS) conditions with an ion trap mass spectrometer system. The identity of the fragment ion arising from the loss of the three neutrals, i.e., H_2O (m/z 138), $[CO + H_2O]$ (m/z 110) and $[CO_2 + NH_3]$ (m/z 95), were further confirmed by deuterium labeling experiments (will be discussed hereafter).

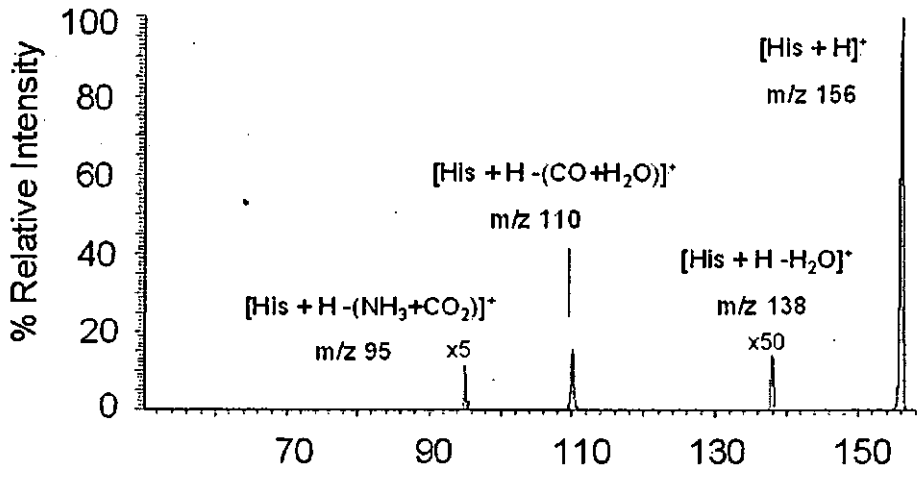


Figure 7.1. Ion trap MS/MS spectra of protonated histidine, $[His + H]^+$ (m/z 156). (ion trap mass analyzer condition: normal scan mode; ion activation time, 10 ms; trap offset at -10 V, q_z at 0.25 and activation RF voltage (V_{p-p}) at 0.4 V).

Previous studies on $[Gly + H]^+$ [Abirami et al., 2004] suggested that ion-trap-CID shows preference for dissociation reactions with low critical energies. As shown in the breakdown curve obtained under such conditions (Fig. 7.2), loss of $[CO + H_2O]$ is the pathway with the lowest threshold voltage (threshold RF voltage 0.3 volts), followed by loss of H_2O (0.35 volts) and loss of $[CO_2 + NH_3]$ (0.4 volts) (Fig. 7.2). Thus, we tend to conclude that the order of critical energies for the three dissociation channels is: $-[CO + H_2O] < -H_2O < -[CO_2 + NH_3]$

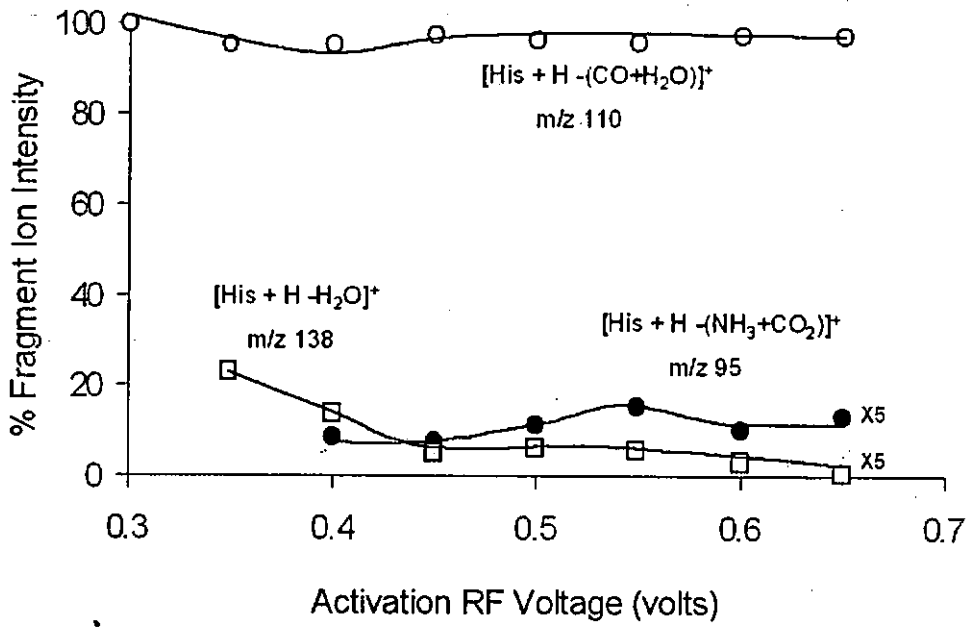
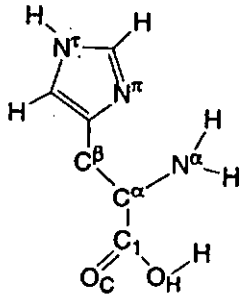


Figure 7.2. Ion trap energy resolved MS/MS breakdown graph of protonated histidine, $[His + H]^+$ (m/z 156). (ion trap mass analyzer conditions: 'low mass' scan mode; ion activation time, 5 ms; main RF set for optimum trapping and detection of m/z 30; trap offset at –5 V).

7.3.2 Proton Affinity and Basicity of Histidine (His)

We first attempt to establish the reliability of our protocol by comparing the theoretical values against the experimental proton affinity (PA) and gas phase basicity (GB) of histidine.

In order to estimate PA and GB, energetic information for both free histidine (**His**) and protonated histidine $[His + H]^+$ is required. We are not aware of detailed *ab initio* conformation studies for free histidine in the literature. Because of tautomerism, the 'external' hydrogen may be located on either the N^{π} or N^{τ} of the imidazole ring. Based on previous theoretical studies of aliphatic amino acids, [Csaszar, 1996; Stepanian et al., 1999; Balta et al., 2000] analogous **His** conformers were constructed with hydrogen bonds between the imidazole N^{τ}/N^{π} and hydrogen bond donor sites (i.e., H-N^{α}, and H-O_H); or between N^{π}-H/N^{τ}-H and hydrogen bond acceptor sites (i.e., N^{α}, O_H, and O_C) (Scheme 7.1) on the backbone. (IUPAC Nomenclature, ref. [Anonymous, 1975])



Scheme 7.1. The schematic labeling of the atoms of histidine that follows the IUPAC instruction. (IUPAC Nomenclature, ref. [Anonymous, 1975])

As in the case of glycine [Jensen and Gordon, 1995; Ding and KroghJespersen, 1996] we have not been able to obtain free histidine in the zwitterionic form. For the four most stable conformers we found (Fig. 7.3(a)), the relative energy is in the order of (in kJ mol^{-1}):

His1 (0) < **His2** (2) < **His3** (5) < **His4** (6)

i.e., the two N^{τ} tautomers (**His1** and **His2**) are marginally more stable than the two N^{π} tautomers (**His3** and **His4**). Our finding differs from that reported by Blinzyuk et al., [Bliznyuk et al., 1993] who suggested that, based on semi-empirical calculations **His3** is the most stable form. Here, we choose **His1** as the reference form of free histidine in the calculation of proton affinity, but note that the choice would have no consequence to the potential energy surface presented in this work.

In the two theoretical studies of histidine protonation, [Bliznyuk et al., 1993; Maksic and Kovacevic, 1999] it appears that only protonation at N^{π} is considered. In fact, four potential protonation sites are present on **His1**: (i) N^{π}, nitrogen on the side chain imidazole ring, (ii) N^{α}, the amino nitrogen, (iii) O_H, the C-terminal hydroxyl oxygen, or (iv) O_C, the C-terminal carbonyl oxygen (Scheme 7.1). All these possibilities were explored and the optimized H^+-His isomer obtained is depicted in (Fig. 7.3(b)). The theoretical proton affinity, PA, at 0K and 298K (ΔH_0 and ΔH_{298}, respectively) and gas basicity, GB, at 298K (ΔG_{298}) for various basic sites, are summarized in Table 7.1.

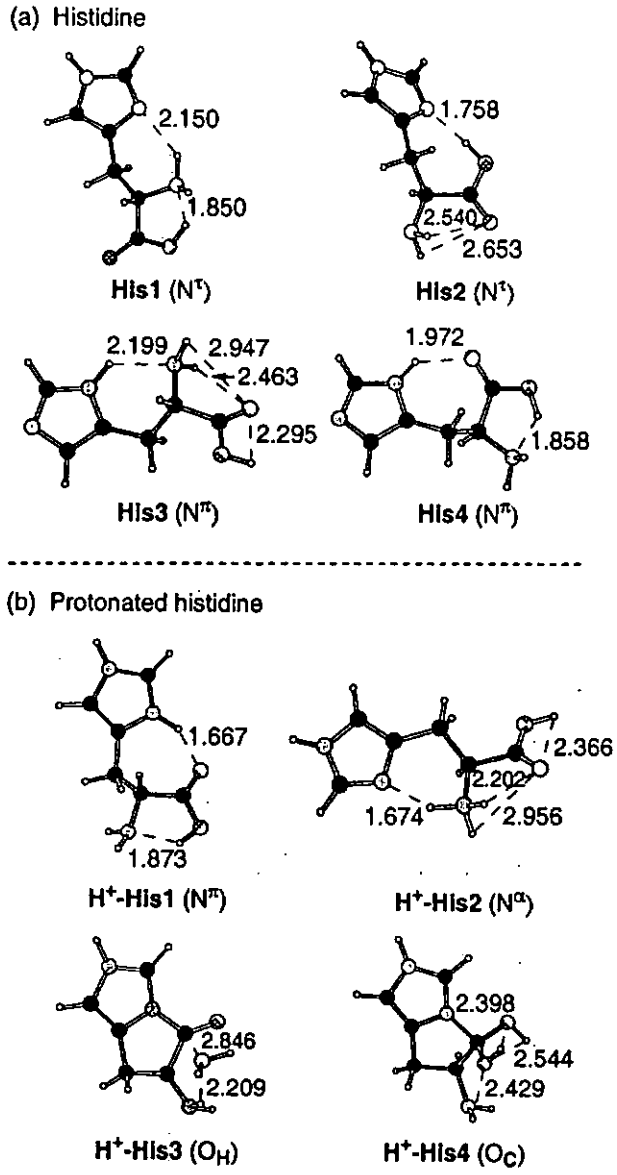


Figure 7.3. Optimized geometry (at B3-LYP/6-31G(d) level) of (a) histidine and (b) protonated histidine. The weak non-covalent interactions are highlighted with dotted lines, with distance given in Å.

Table 7.1. Gas phase proton affinity, PA and basicity, GB (in kJ mol^{-1}) of histidine.

Sites	This work			Literature [a]	
	ΔH_0	ΔH_{298}	ΔG_{298}	ΔH_{298}	ΔG_{298}
Imidazole N^{π}	981	987	952	971 [b], 952 [c], 988 [d] 956 [e] - 988 [f]	936 [b], 915 [c], 950 [d] 892 [e] - 950 [f]
amino N^{α}	974	984	952		
hydroxyl O_H	909	915	891		
carbonyl O_C	893	904	866		

[a]: Thermal corrections with B3-LYP/6-31G(d) geometries and frequencies were applied, if necessary, in order to adjust the reported values to 298K. [b]: Theoretical ΔH_0 (M(II)) from ref. [Maksic and Kovacevic, 1999]. [c]: Theoretical ΔH_0 (AM1) from ref. [Bliznyuk et al., 1993]. [d]: Experimental value from NIST webbook, ref. [Hunter and Lias, 1998]. [e]: Minimum of experimental values from ref. [Locke and McIver, 1983; Lias et al., 1984; 1988; Decouzon et al., 1991; Zhang et al., 1993; Wu and Fenselau, 1994; Bojesen, and Breindahl, 1994; Cassady et al., 1995; Carr and Cassady, 1996; Harrison, 1997; Hunter and Lias, 1998]. [f]: Maximum of experimental values from ref. [Locke and McIver, 1983; Lias et al., 1984; 1988; Decouzon et al., 1991; Zhang et al., 1993; Wu and Fenselau, 1994; Bojesen, and Breindahl, 1994; Cassady et al., 1995; Carr and Cassady, 1996; Harrison, 1997; Hunter and Lias, 1998].

Protonation at N^{π} (**H^+-His1** in Fig. 7.3(b)) is at least 7 kJ mol^{-1} more stable than the other three basic sites. The relative stability of the four basic sites is:

H^+-His1 (imidazole N^{π}) > **H^+-His2** (amino N^{α})

> **H^+-His3** (hydroxyl O_H) > **H^+-His4** (carbonyl O_C)

Thus, protonation at nitrogen is preferred over oxygen. Similar to the case of glycine [Rogalewicz and Hoppilliard, 2000b], protonation at hydroxyl O_H of histidine leads to

the formation of an ion–water complex (**H^+-His3**, Fig. 7.3(b)). Interestingly, while protonation at the carbonyl oxygen (O_C) is more stable than O_H-protonation in glycine (by ca. +10 kJ mol^{-1}), the opposite is observed in histidine (by –16 kJ mol^{-1}). We attribute the difference between glycine and histidine to the presence of the aromatic imidazole ring in histidine. In the presence of the ring, a bicyclic species is produced when a covalent bond is formed between C_1 and N^{π}. For this bicyclic ion, conjugation of the sp^2 C_1 in **H^+-His3** (O_H) with the imidazole ring conjures extra stability. Such stabilizing factor is absent in **H^+–His4** (O_C), nor can be found in the corresponding protonated glycine complexes.

We also explored the possibility of the formation of H^+-His in the zwitterionic form. Previously, H^+-Arg has been suggested to be in zwitterionic form in the gas phase [Price et al., 1997] because of the presence of the highly basic guanidine side chain. As the proton affinities of H^+–Arg and H^+-His are the closest among the 20 amino acids [Harrison, 1997] (with PA at 1051 and 988 kJ mol^{-1}, respectively) [Hunter and Lias, 1998], one may expect H^+-His to be zwitterionic as well. However, we have not been able to locate any stable H^+-His in zwitterionic form, and thus, we conclude that such protonated histidine form(s) may not be stable in the gas phase.

Based on the energetic difference of **His1** and **H^+-His1**, our estimated PA and GB for histidine is 987 and 952 kJ mol^{-1}, respectively. These values are in excellent agreement (within 2 kJ mol^{-1}) with the recommended values in the NIST Standard Reference Database [Hunter and Lias, 1998], at 988 and 950 kJ mol^{-1}, respectively. Previously estimated values [Bliznyuk et al., 1993; Maksic and Kovacevic, 1999] are at least 10 kJ mol^{-1} too low.

7.3.3 Dissociation of Protonated Histidine

Experimentally, $[His + H]^+$ yields predominantly fragment ions due to the loss of 46 u, $[CO + H_2O]$, regardless of the ionization techniques used. [Kulik and Heerma, 1988; Dookeran et al., 1996; Rogalewicz and Hoppilliard, 2000b] Furthermore, minor fragment ions due to the loss of 18 u [Kulik and Heerma, 1988] and 61 u [Dookeran et al., 1996] are sometimes observed. This is in contrast to protonated aliphatic amino acids in which $[CO + H_2O]$ is lost exclusively. In the discussion below, we will first focus on the loss of 46 u for $[CO + H_2O]$, followed by the loss of 18 u for H_2O, and finally the loss of 61 u for $[CO_2 + NH_3]$ in protonated histidine. The potential energy surface of all three channels is summarized in Fig. 7.4, while the structures of the stationary points are depicted in Figs. 7.5, 7.6 and 7.7 (Cartesian coordinates can be found in Appendix V, Table S-7.1). In this work, the labels **1**, **2**, etc. are used to denote the various minima on the PES, while labels **1-2**, **2-3**, etc. denote the transition structures between minima **1** and **2**; **2** and **3**, respectively.

Loss of $[CO + H_2O]$

Previously, Rogalwicz and Hoppilliard [Rogalwicz and Hoppilliard, 2000b] have compared the energetic requirement for the loss of 46 u from $[Gly + H]^+$: (i) formic acid (HCOOH), (ii) carbon dioxide and hydrogen ($CO_2 + H_2$), (iii) dihydroxycarbene ($C(OH)_2$), and (iv) water and carbon monoxide $[CO + H_2O]$, and found that the last pathway (i.e., loss of $[CO + H_2O]$) is energetically most favorable. In this project, we have postulated the mechanism for the loss of all these four sets of neutrals, and our findings are summarized in Appendix VI and VII. Details aside, we note here that similar to $[Gly + H]^+$, the formation of $[CO + H_2O]$ from protonated histidine is energetically more favored. Thus, in the discussions below, we will assume that the

loss of 46 u corresponds to that of [CO + H_2O]. Furthermore, several pathways leading to the loss of [CO + H_2O] from protonated histidine has been investigated. The lowest energy pathway is depicted in Fig. 7.4, with the relevant structures and energetics displayed in Fig. 7.5 and Table 7.2, respectively.

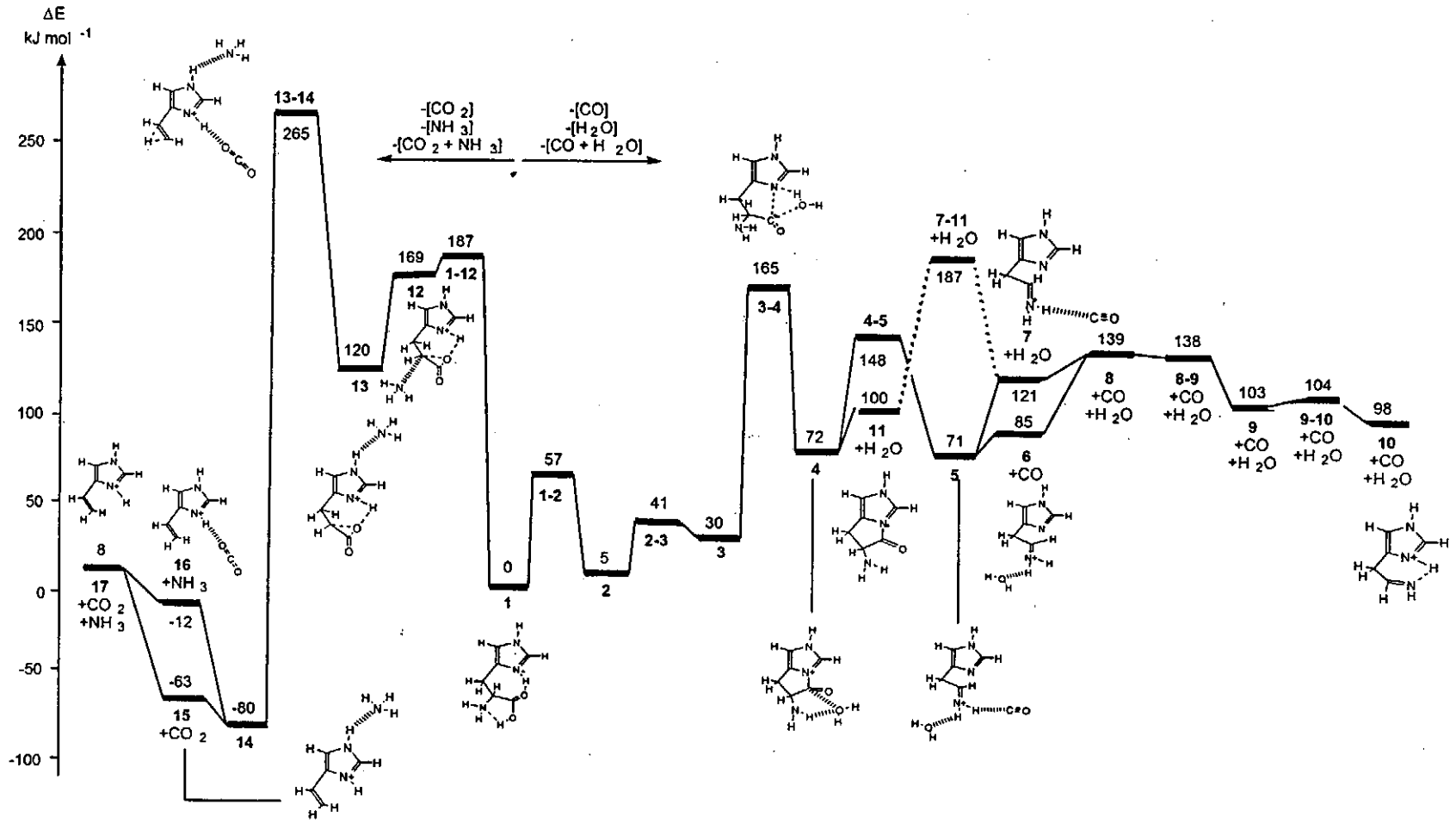


Figure 7.4. The potential energy surface (PES) at 0K for the fragmentation of $[His + H]^+$ in the gas phase, calculated at the B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d).

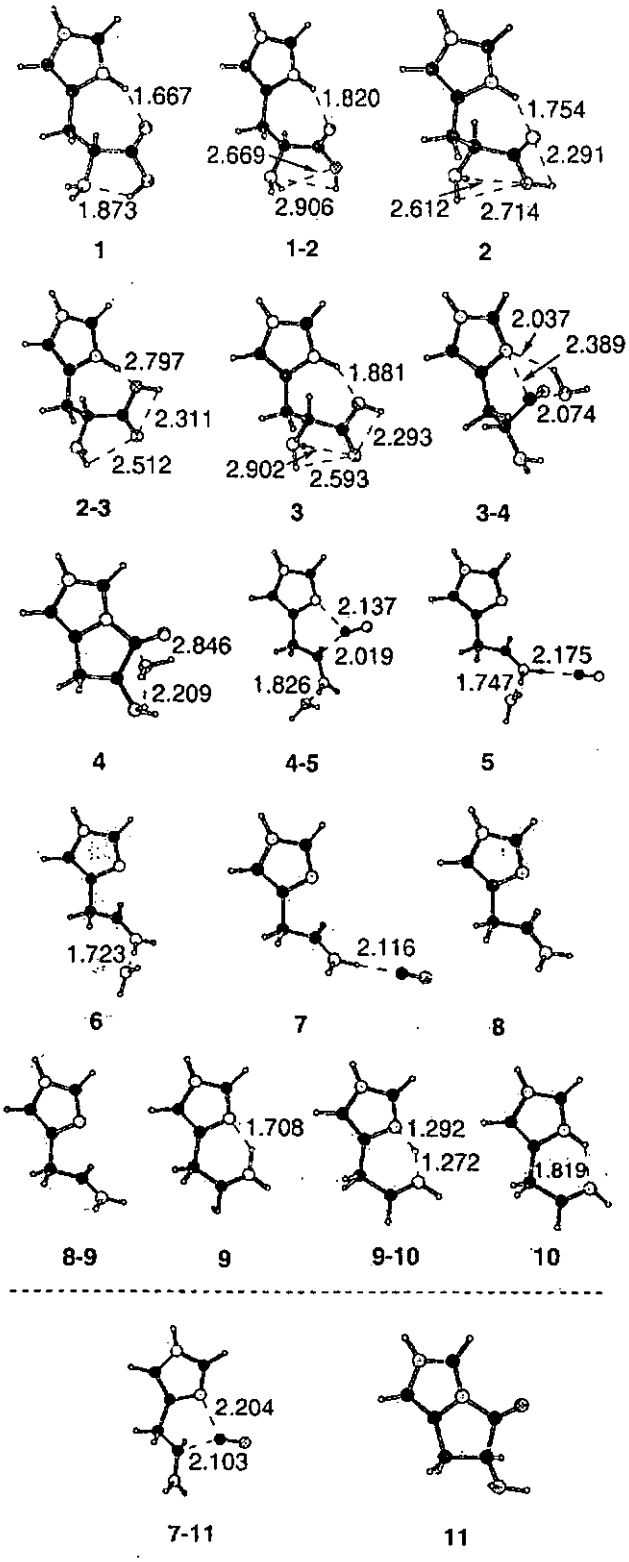


Figure 7.5: Optimized geometry (at B3-LYP/6-31G(d) level) of the species related to the loss of 46 u ([CO + H_2O]) and 18 u (H_2O) from $[His + H]^+$.

From the most stable form of protonated histidine (**1** in Fig. 7.5, identical to **H^+-His1** in Fig. 7.3 (b)), the species goes through several low energy conformational changes (**1-2**, **2**, **2-3**) to yield **3**, in which the O_H is positioned to accept the proton from N^{π}-H. Proton transfer occurs via TS **3-4** to form a 6-amino-5-oxo-2,5,6,7-tetrahydro-pyrrolo[1,2-c]imidazol-4-ylium(bicyclic ion)...H_2O complex (**4** in Fig. 7.5, identical to **H^+-His3** in Fig. 7.3 (b)). Through TS **4-5**, species **5** is formed, which can be described as a H_2O...immonium ion...CO ion-molecule complex. Species **5** could fragment to yield (**6** + CO) or (**7** + H_2O) directly. Immonium ion **8** can be formed from either **6** or **7**, via further loss of H_2O and CO, respectively. Through **8-9**, a more stable form (stabilized by N^{α}-H...N^{π} hydrogen bond) of the immonium ion **9** is formed. Via a facile proton transfer from N^{α}-H to N^{π} (**9-10**), an imine cation **10** is produced. The 0K enthalpy of reaction of this process (energy difference between **1** and **10**) is 98 kJ mol^{-1}, with an estimated energy barrier of 165 kJ mol^{-1}.

Table 7.2. The relative enthalpies at 0K (ΔH_0), relative Gibbs free energies at 298K (ΔG_{298}) and 600K (ΔG_{600}), in kJ mol^{-1}, with respect to species 1.

Species	Energetics		
	ΔH_0	ΔG_{298}	ΔG_{600}
1	0	0	0
1-2	57	55	52
2	5	3	-1
2-3	41	38	34
3	30	27	21
3-4	165	164	156
4	72	66	52
4-5	148	137	116
5	71	47	7
6 + CO	85	30	-40
7 + H_2O	121	65	-8
8 + CO + H_2O	139	51	-52
8-9 + CO + H_2O	138	53	-43
9 + CO + H_2O	103	16	-85
9-10 + CO + H_2O	104	19	-77
10 + CO + H_2O	98	11	-88
11 + H_2O	100	60	13
7-11 + H_2O	187	142	87
1-12	187	182	172
12	169	157	136
13	120	108	89
13-14	265	240	202
14	-80	-107	-148
15 + CO_2	-63	-118	-185
16 + NH_3	-12	-73	-149
17 + CO_2 + NH_3	8	-82	-185

While loss of [CO + H_2O] is the dominant fragmentation pathway for both protonated aliphatic amino acids and $[His + H]^+$, H/D exchange studies suggested that the mechanisms involved are different. [Rogalewicz et al., 2000c] Our mechanism presented above is consistent with experimental H/D exchange results and that proposed previously, but it does not explain *why* there is a change of fragmentation mechanism from protonated aliphatic amino acids to $[His + H]^+$ even though the loss of neutrals is the same. In order to address this, we examined the loss of [CO + H_2O] from $[His + H]^+$, in a manner analogous to that reported in $[Gly + H]^+$and $[Ala + H]^+$ [Uggerud, 1997; O'Hair, 2000; Rogalewicz and Hoppilliard, 2000b; Abirami et al., 2004]. This analogous mechanism (initiated by proton transfer from amino N^{α} to hydroxyl O_H, summarized in Appendix VII, Fig. S-7.3), is found to be at least 30 kJ mol^{-1} (via **TSa** and **TSb,** Fig. 7.6) higher in energy as compared to TS **3-4** presented above. In other words, the presence of an imidazole ring in protonated histidine allows the formation and occurrence of a "bicyclic-like" transition structure **3-4**, resulting in the switch of fragmentation pathway for loss of [CO + H_2O] different from that of protonated aliphatic amino acids.

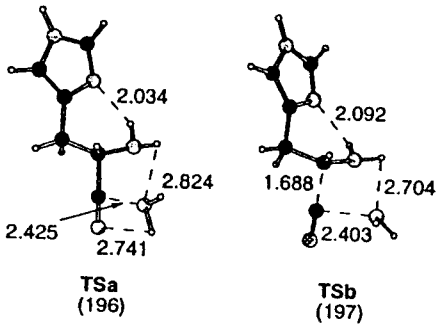


Figure 7.6. Higher energy transition structures for the loss of 46 u ([CO + H_2O]) in $[His + H]^+$, optimized at the B3-LYP/6-31G(d) level of theory. The bracket values are relative enthalpies at 0K (ΔH_0), in kJ mol^{-1}, with respect to the ground state of species **1**.

The imidazole ring plays important role here. As a very stable conjugated species **4** is formed from transition structure **3-4**, and according to the Hammond postulate such TS would be expected to be energetically low. [Hammond, 1995] In other words, it appears that the change of mechanism from $[Ala + H]^+$ to $[His + H]^+$ is due to the *stability* of transition state **3-4**.

Loss of H_2O

Our experimental results showed that there was a minor peak at m/z 138 in the metastable and low energy CID spectra (Fig. 7.1), corresponding to loss of H_2O from protonated histidine. In the mechanism presented above, water can be produced along with species **7** from species **1** via **5** (Fig. 7.4). Species **5** is a weak ion-molecule complex of immonium ion with H_2O and CO (Fig. 7.4). Given that the immonium ion...CO interaction would be weaker than the immonium ion...H_2O interaction, one expect the loss of CO (28 u) would be observed if loss of H_2O is observed from **5**. As this is not the case, it suggests that the observed loss of H_2O peak at m/z 138 most likely is formed from an alternative pathway.

Here we note that species **4** is an ion-water ion-molecule complex. Removal of water from this complex only requires ca. 30 kJ mol^{-1} to yield (**11** + H_2O). The 6-amino-5-oxo-2,5,6,7-tetrahydro-pyrrolo[1,2-c]imidazol-4-ylium **11** (the bicyclic ion) (Fig. 7.4) is particularly stable because of the conjugation between carbonyl C=O and the imidazole ring. For species **11** to react further (via transition structure **7-11**), an additional 22 kJ mol^{-1} above the transition structure **3-4** is required. In other words, once (**11** + H_2O) is formed, it may be kinetically stable enough to withstand further fragmentation. As the imidazole ring is not present in aliphatic amino acids, formation of stable intermediates like species **11** is not possible. Thus, the loss of

H_2O peak could be observed in the mass spectra of $[His + H]^+$ (Fig. 7.1), but not in the spectra of protonated aliphatic amino acids.

Furthermore, we note that species **4** (Fig. 7.5) is identical to species **H^+-His3** (hydroxyl O_H protonated histidine) depicted in Fig. 7.3 (b). As protonation at N^{π} is ca. 70 kJ mol^{-1} more stable than at O_H (Table 7.1), initial O_H protonation is thermodynamically unlikely. Thus, the reaction mechanism proposed here is additional evidence that species **4** is unlikely to be formed initially: if O_H protonation did occur, then, the dominant neutral loss would be H_2O, not $[CO + H_2O]$ as observed experimentally; as loss of $[CO + H_2O]$ via **4-5** requires ca. 70 kJ mol^{-1}, but loss of H_2O via **11** only requires ca. 30 kJ mol^{-1} from species **4**.

Loss of $[CO_2 + NH_3]$

In agreement with Harrison and co-workers [Dookeran et al., 1996], a minor fragment ion due to the loss of 61 u (attributed to the loss of $[CO_2 + NH_3]$) was also observed in the metastable and low energy CID mass spectra of protonated histidine (Fig. 7.1), but it could not be found in the corresponding mass spectra of protonated aliphatic amino acids. [Dookèran et al., 1996; Rogalewicz et al., 2000c] Therefore, the loss of the $[CO_2 + NH_3]$ neutral must be due to the presence of the imidazole ring in the side chain of histidine.

Two pathways leading to the loss of $[CO_2 + NH_3]$ have been investigated in this work. The higher energy pathway (with a barrier of 318 kJ mol^{-1} above species **1**, summarized in Appendix VII, Fig. S-7.1), is derived from the loss of CO_2 pathway postulated for $[Gly + H]^+$ in ref. [Rogalewicz and Hoppilliard, 2000b]. The lower energy pathway (Fig. 7.7) has a barrier of 265 kJ mol^{-1} above species **1**, which is still significantly higher (~100 kJ mol^{-1}) than the energy barrier associated with the

transition structure **3-4** for the loss of [CO + H_2O] pathway at 165 kJ mol^{-1} (Fig. 7.4). Thus, our calculated energetics is consistent with the experimental observation that loss of 61 u ([CO_2 + NH_3]) would be only a minor channel when compared to the loss of 46 u [CO + H_2O] channel.

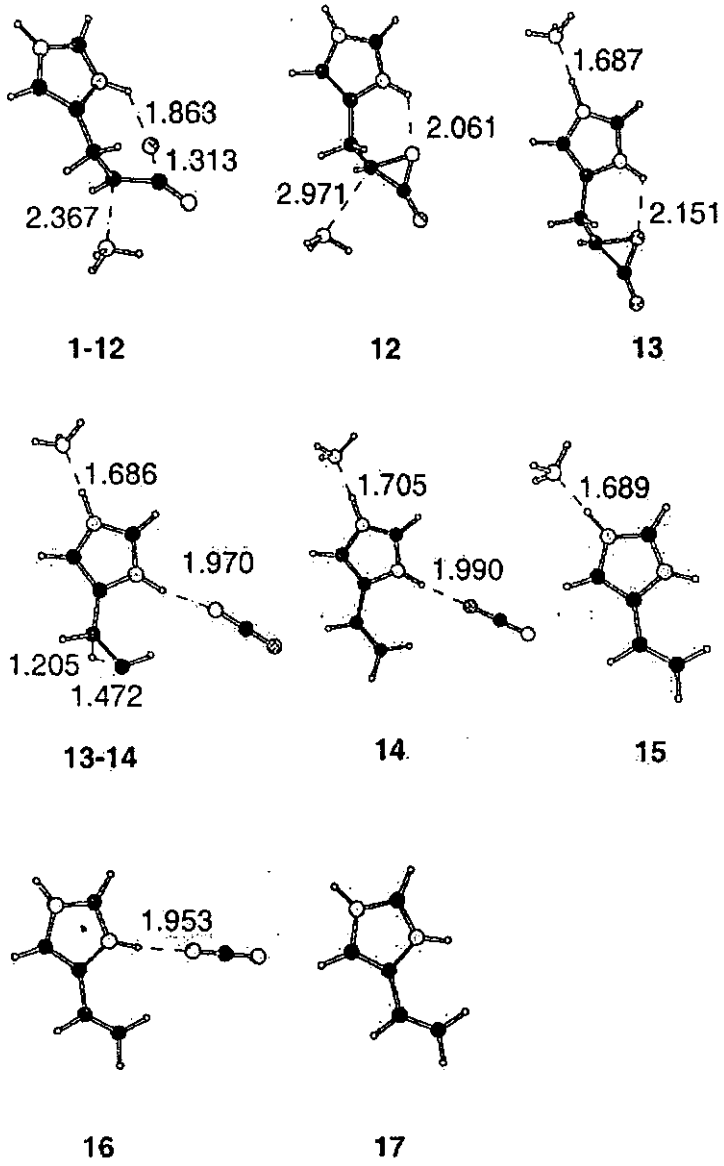


Figure 7.7. Optimized geometry (at B3-LYP/6-31G(d) level) of the species related to the loss of 61 u ([CO_2 + NH_3]) from $[His + H]^+$.

To lose [CO_2 + NH_3] from the most stable form of protonated histidine (**1**), the species goes through a transition structure **1-12** (Fig. 7.7, 187 kJ mol^{-1} above **1**), with the reaction coordinate corresponding to the bending of angle O_H-C^1-C^α. From **1-12**, an ion-molecule complex **12** is formed, in which NH_3 is weakly coordinated to the C^α. As the C^α...NH_3 distance is almost 3Å, and with slight excess excitation energy, it is very likely that the two constituents parts of the ion-molecule complex can "freely" rotate with respect to each other. With this relative 'free' rotation, complex **12** could rearrange to an energetically more stable isomer in which the NH_3 binds to the imidazole ring via a hydrogen bond (complex **13**, with H_3N...HN^τ of ca. 1.7Å). When a 1,2-hydrogen shift occurred between C^β and C^α, via a barrier (**13-14**) of 265 kJ mol^{-1}, a very stable NH_3...5-vinyl-3*H*-imidazol-1-ium...CO_2 ion-molecule complex is formed (species **14**, ca. 80 kJ mol^{-1} below species **1**), which could fragment to yield (**15** + CO_2) or (**16** + NH_3). The stable species **17**, the 5-vinyl-3*H*-imidazol-1-ium, can be formed from either **15** or **16** via further loss of NH_3 and CO_2, respectively.

Based on our potential energy surface, we concluded that the imidazole ring serves two functions in the 'unique' loss of [CO_2 + NH_3] from protonated histidine. Firstly, the formation of the 5-vinyl-3*H*-imidazol-1-ium (**17**) is thermodynamically favorable because of the conjugation between the imidazole ring and the C^α=C^β bond. Secondly, the imidazole ring provides an "anchor" for the small neutral fragments NH_3 and CO_2. Without this anchoring effect of the ring, via the strong H_3N...HN^τ interaction (and a weaker CO_2...HN^π interaction), the loss of [CO_2 + NH_3] pathway will have to proceed via some higher energy pathway, like that proposed in Appendix VII, Fig. S-7.1. It is also noteworthy that such "anchoring" of neutrals to an ion is not

present in protonated aliphatic amino acids. Thus, the loss of $[CO_2 + NH_3]$ pathway is not found in protonated aliphatic amino acids.

Experimental H/D Exchange Studies

The low-energy CID (eV, laboratory scale) spectrum of deuterated d_4-histidine, $[d_4\text{-His} + D]^+$ (m/z 161), in which all the labile acidic/basic hydrogen atoms have been replaced by deuterium atoms, is shown in Fig. 7.8 (taken from ref. [Wong, 2003]).

As shown in Fig. 7.8, loss of $[CO + D_2O]$ and [CO + HDO] in the ratio ~2:1 (m/z 113 : m/z 114) and $[CO_2 + ND_3]$ and $[CO_2 + ND_2H]$ in the ratio of 4:5 (m/z 97 : m/z 98) were observed. With all labile protons in histidine replaced by deuterium in $[d_4\text{-His+D}]^+$, elimination of D_2O and ND_3 is expected.

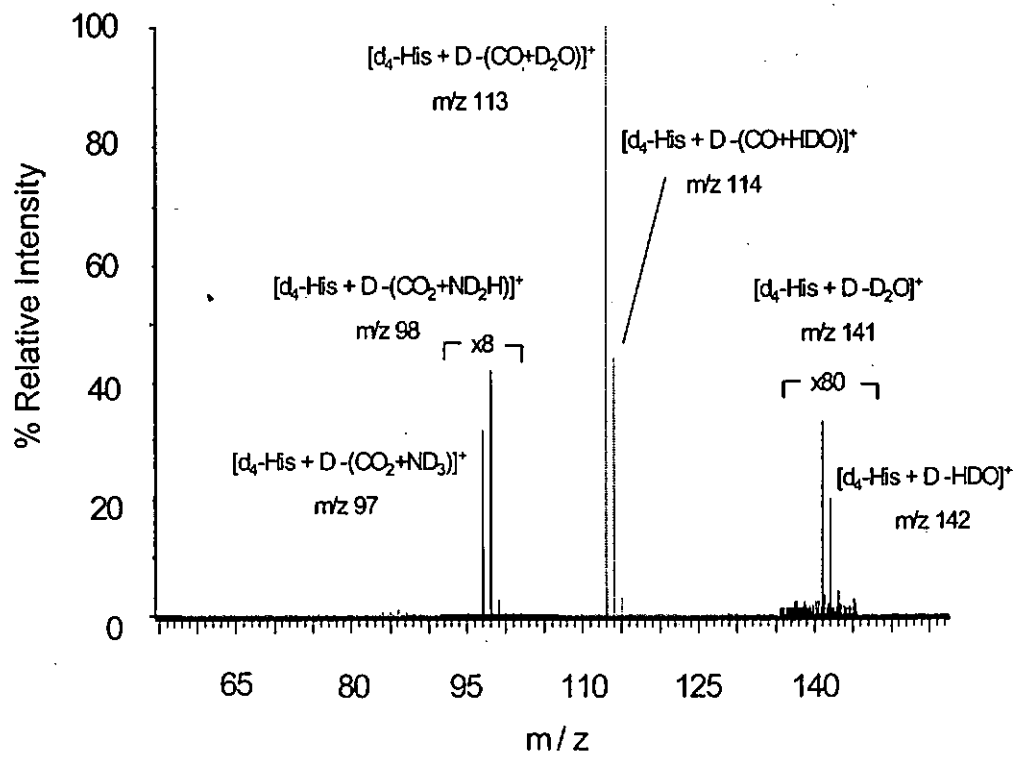


Figure 7.8. Q-TOF MS/MS spectra of deuterated d_4-histidine, $[d_4\text{-His} + D]^+$ (m/z 161) generated by electrospray ionization.

Here we like to comment on the experimental H/D exchange results. We found that full H/D exchange and dissociation of deuterated histidine yielded mixture of peaks due to the loss of D_2O/HDO; CO + D_2O/HDO; and CO_2 + ND_3/ND_2H. While our mechanism proposed above is consistent with the H/D experiment, the mixture of product fragment ions suggests that scrambling of the exchanged deuterium has occurred, i.e., the presence of the imidazole ring apparently allows the exchange of labile deuterium with the ring carbon hydrogen atoms.

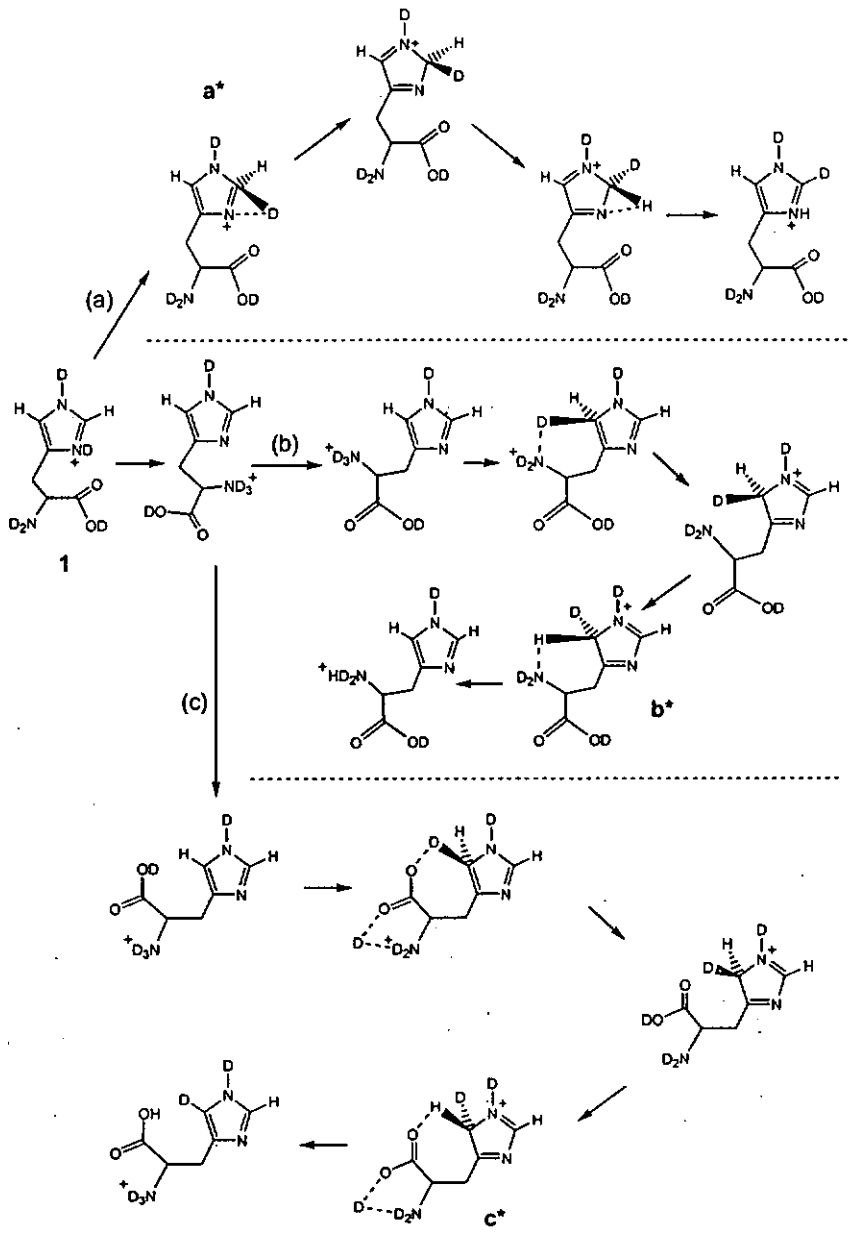
Three mechanisms [Rogalewicz et al., 2000c; Aribi et al., 2004; Lioe et al., 2004] have been proposed on how such scrambling may occur in protonated aromatic amino acids: histidine [Rogalewicz et al., 2000c], tryptophan [Lioe et al., 2004] and phenylalanine [Aribi et al., 2004]. Given the similarities between the three aromatic amino acids, we believe that the proposed scrambling mechanism for protonated tryptophan/phenylalanine may also be plausible for protonated histidine. In Scheme 7.2, we summarize the key steps of these mechanisms for protonated histidine (species **1**), with necessary modifications on the proposed mechanism of ref. [Lioe et al., 2004 and Aribi et al., 2004]. Scrambling may occur via the exchange of the deuterium on:

(a) ring N^{π} with imidazole C_2 hydrogen [Rogalewicz et al., 2000c];

(b) amino nitrogen with imidazole C_5 hydrogen [Lioe et al., 2004];

(c) hydroxyl oxygen with imidazole C_5 hydrogen [Aribi et al., 2004].

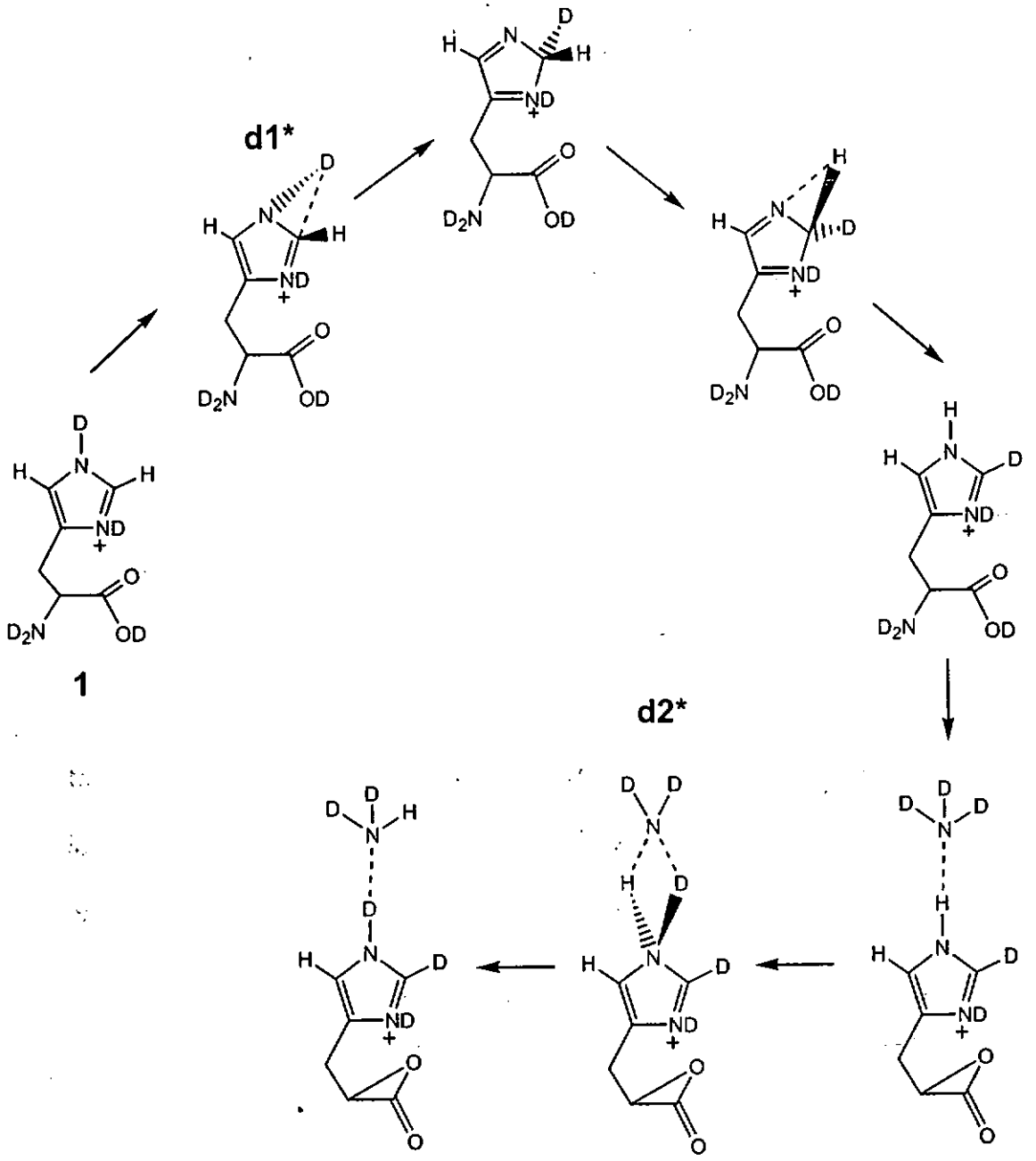
We have obtained the key transition structures of the three mechanisms, labeled as **a***, **b*** and **c*** in Scheme 7.2. The energy barriers relative to protonated histidine (species **1**) are 280, 138 and 145 kJ mol^{-1} for **a***, **b*** and **c***, respectively.



Scheme 7.2. Three mechanisms [Rogalewicz et al., 2000c; Aribi et al., 2004; Lioe et al., 2004] have been proposed on H/D scrambling occurring in protonated histidine. Pathway (a) corresponds to the equivalent mechanism proposed reported in ref. [Rogalewicz et al., 2000c] for [His + H]$^+$; pathway (b) corresponds to the equivalent mechanism reported in ref. [Lioe et al., 2004] for protonated tryptophan; and pathway (c) corresponds to the equivalent mechanism reported in ref. [Aribi et al., 2004] for protonated phenylalanine. Only key steps are shown and some modifications (if necessary) are made for these mechanisms.

The key barrier for the loss of D_2O/HDO; [CO + D_2O]/HDO from species **1**, (species **3-4**, Fig. 7.4) is estimated to be approximately 165 kJ mol^{-1}. Given the low energetic requirement for pathways (b) and (c), these two scrambling mechanisms would certainly be accessible under experimental CID conditions. Pathway (a), proposed by Rogalewicz et al., [Rogalewicz et al., 2000c] is considerably higher in energy. However, as the barrier would be expected to be much reduced in the presence of solvating water or methanol (in solution or vapor phase) present under electrospray ionization conditions [Rodriquez et al., 2000b; Aribi et al., 2004], the mechanism cannot be ruled out either.

The loss of [CO_2 + ND_3/ND_2H] via species **13-14** requires 265 kJ mol^{-1}. Given such high energy requirement, there could be a fourth mechanism, i.e., via ion-molecule complex **13** (presented in Scheme 7.3). In this mechanism, the ring N^{τ} exchanges its deuterium with the hydrogen on the imidazole C_2. This is related to pathway (a) presented above but the deuterium of different ring nitrogen has been exchanged, with the barrier for the exchange (**d1***) estimated to be 241 kJ mol^{-1}. With the N^{τ} deuterium replaced by hydrogen, following the formation of carbon dioxide and ammonia mechanism presented above (**1** → **1-12** → **12**, etc., Figs. 7.4 and 7.7), the ND_3 has another opportunity to exchange the deuterium with the N^{τ} hydrogen. The transition structure involved, **d2***, is estimated to be 195 kJ mol^{-1}, which is ~ 70 kJ mol^{-1} lower than the barrier **13-14** that leads to the eventual formation of carbon dioxide and ammonia.



Scheme 7.3. Scrambling that initiated by the ring N^{τ} exchanges its deuterium with the hydrogen on the imidazole C_2 in protonated histidine, followed by the formation of carbon dioxide and ammonia mechanism presented in Fig. 7.7 (1 → **1-12** → **12**, etc.) that the ND_3 exchanges the deuterium with N^{τ} hydrogen. Only key steps are shown here.

Effect of Temperature

In the previous sections, all energetics and discussions pertained to the thermodynamic information at 0K. We have recently found that the ion–trap condition is more akin to reactions conditions at ~600K [Abirami et al., 2004]. Thus, in this section, we would like to investigate how temperature affects the ΔG of various fragmentation processes discussed above. The theoretically calculated ΔG_{298} and ΔG_{600} values are shown in Table 7.2.

In Fig. 7.9, the variation of the relative Gibbs free energy is plotted against temperature for: (a) the competitive loss of CO, H_2O and [CO + H_2O] from the highest (determining) energy barrier TS **3-4** (Fig. 7.9 (a)) and (b) the competitive loss of CO_2, NH_3 and [CO_2 + NH_3] from the highest (determining) energy barrier TS **13-14** (Fig. 7.9 (b)). As the negative of ΔG is plotted in these figures, a *more* negative value (lower on the y-axis) indicates a thermodynamically *less* favorable process. As an example, the Gibbs free energies of reactions (open symbols) are above the free energy of the associated barrier (solid symbols), suggesting that the fragmentation are kinetically controlled at all temperatures.

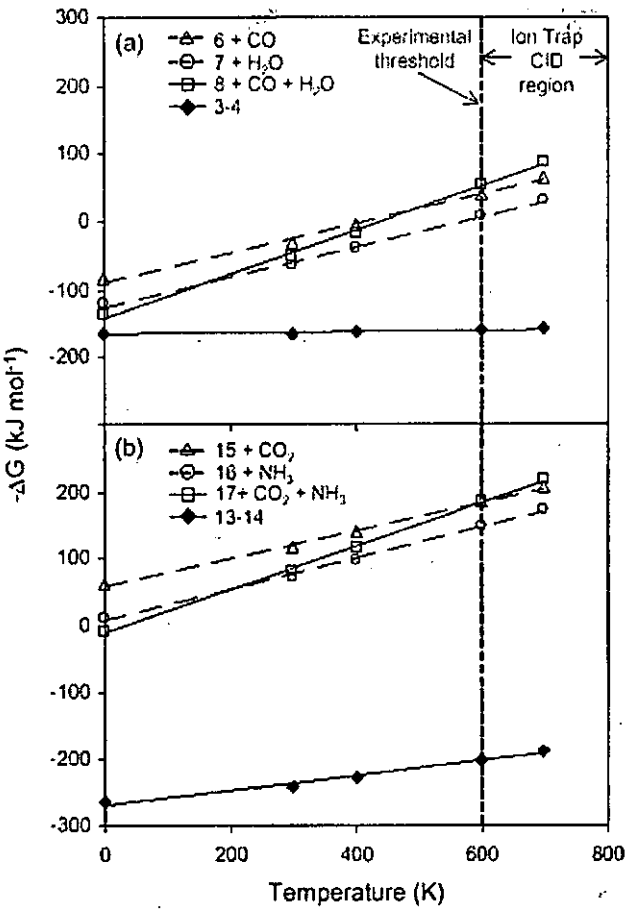


Figure 7.9. The variation of relative Gibbs free energy (kJ mol^{-1}) as a function of temperature of (a) competitive loss of CO (triangle), H_2O (open circle), and [CO + H_2O] (open square) from barrier **3-4** and (b) competitive loss of CO_2 (triangle), NH_3 (open circle) and [CO_2+ NH_3] (open square) from barrier **13-14** from $[His + H]^+$.

The variation of ΔG with temperature is essentially linear, indicating that enthalpy (ΔH) and entropy (ΔS) are fairly constant over the range of temperature depicted in Fig. 7.9. On comparing the variation of ΔG for **3-4** (Fig. 7.9 (a)) and **13-14** (Fig. 7.9 (b)), the latter barrier remained less competitive even at higher temperature. Thus, as far as the competitive loss of [CO + H_2O] vs [CO_2 + NH_3] is concerned, the results found at 0K (discussed in the previous sections) remain unchanged. This is consistent

with the trend observed in lower threshold voltage/ critical energies for the loss of [CO + H_2O] than the loss of [CO_2 + NH_3] pathway (Fig. 7.2).

Moreover, the variation of ΔG for [**7** + H_2O] and [**6** + CO] with temperature are essentially parallel (Fig. 7.9 (a)). At 0K, the formation of [**8** + CO + H_2O] is less favorable than the two competing products, loss of CO and loss of H_2O with the same energy barrier at **3-4**. However, as the ΔG of [**8** + CO + H_2O] is a more steeply variation function of temperature (arising from the entropic factor), this fragmentation channel becomes thermodynamically preferred as temperature increases so that two "cross-over" points are found: the first point at ~170K suggesting that loss of [CO + H_2O] is favored over the loss of H_2O, the second point at ~500K indicating that the loss of [CO + H_2O] become the thermodynamically preferred channel, over *both* loss of CO and loss of H_2O. Our theoretical findings on the loss of H_2O pathway are consistent with experimental observations (Fig. 7.2).

Similarly, the competitive dissociation of [**15** + CO_2], [**16** + NH_3] and [**17** + CO_2 + NH_3], which also share the same energy barrier **13-14**, show the same trend in Fig. 7.9 (b). The pathway for loss of two neutrals, [**17** + CO_2 + NH_3] becomes dominant over the loss of one neutral, i.e., the loss of CO_2 or NH_3 pathway at a higher temperature (~600K). As a result, the peaks due to the loss of CO_2 (44 u) or NH_3 (17 u) are not found in the metastable or CID spectra of $[His + H]^+$.

In summary, our potential energy surface provides theoretical rationale for the experimental results. Firstly, the lost of 46 u [CO + H_2O] is favored over 61 u [CO_2 + NH_3] because the barrier to the formation of the latter product is higher at all temperatures. Secondly, for competing products that shared the same barrier, entropic factor ($T\Delta S$) would favor the product channel that forms more number of neutrals and

fragment ions. However, whether the entropy factor is large enough to change the thermodynamic preference (ΔG) of one channel over another (from the same barrier) will depend largely on the difference in enthalpy at 0K of these channels.

7.3.4 Low-Energy Collision-induced Dissociation (CID) of Potassiated Histidine

It is clear from the ion trap-CID spectrum presented in Fig. 7.10(a) (taken from ref. [Wong, 2003]) that loss of K^+ (m/z 39) is dominant from potassiated histidine, $[His + K]^+$ (m/z 194). Under CID conditions as shown in Fig. 7.10 (b), losses of CO_2 and NH_3 are also observed. According to the breakdown graph shown in Fig. 7.11 (b), the order of threshold voltages (critical energies) for the three dissociation channels is: $K^+ < -CO_2 < -NH_3$.

Thus, comparing to $[His + H]^+$, in which the MS/MS spectrum is dominated by the loss of $[CO + H_2O]$ peak, with loss of H_2O and $[CO_2 + NH_3]$ as the minor peaks (Fig. 7.1), it is clear that the nature of cation (H^+ versus K^+) plays an important role in the fragmentation of cationized histidine. In the remaining sections in this Chapter, three issues will be discussed:

(a) where is the likely site(s) for the potassium cation to bind to histidine

(b) how does K^+ formation and loss of NH_3 and CO_2 occur from $[His + K]^+$

(c) why the preferred fragmentation pathways differs in $[His + H]^+$ and $[His + K]^+$.

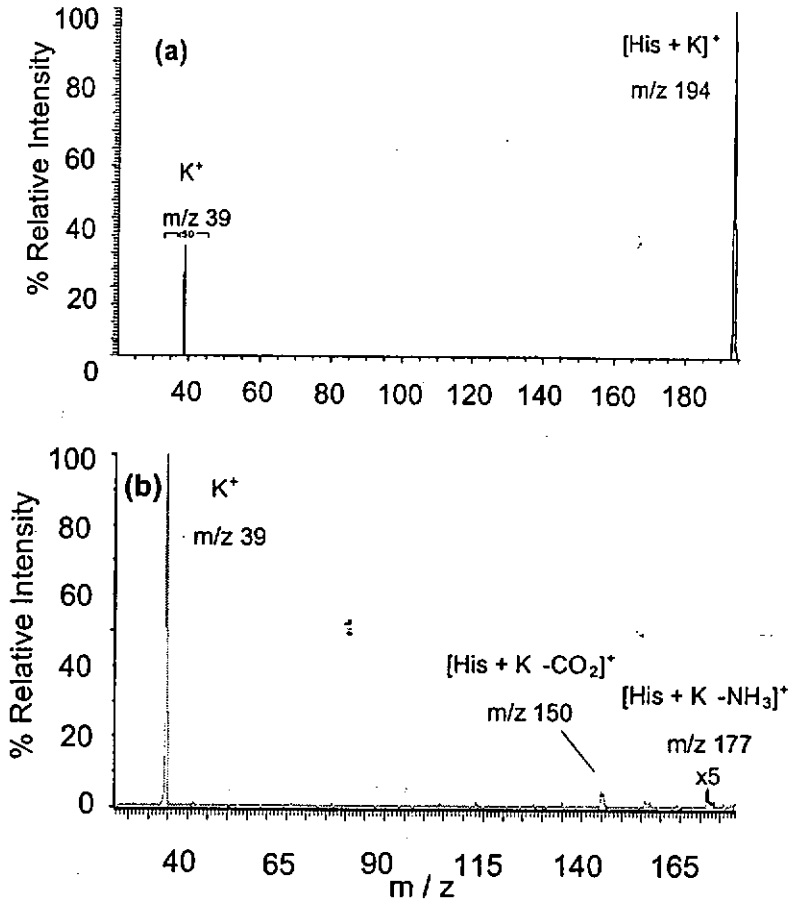


Figure 7.10. Low-energy CID mass spectra of $[His + K]^+$: (a) ion trap-CID: RF activation voltage: 0.66V, ion activation time: 5ms; trap offset: -5V and qz:0.2); b) triple quadrupole CID at collision energy of 14V (laboratory scale).

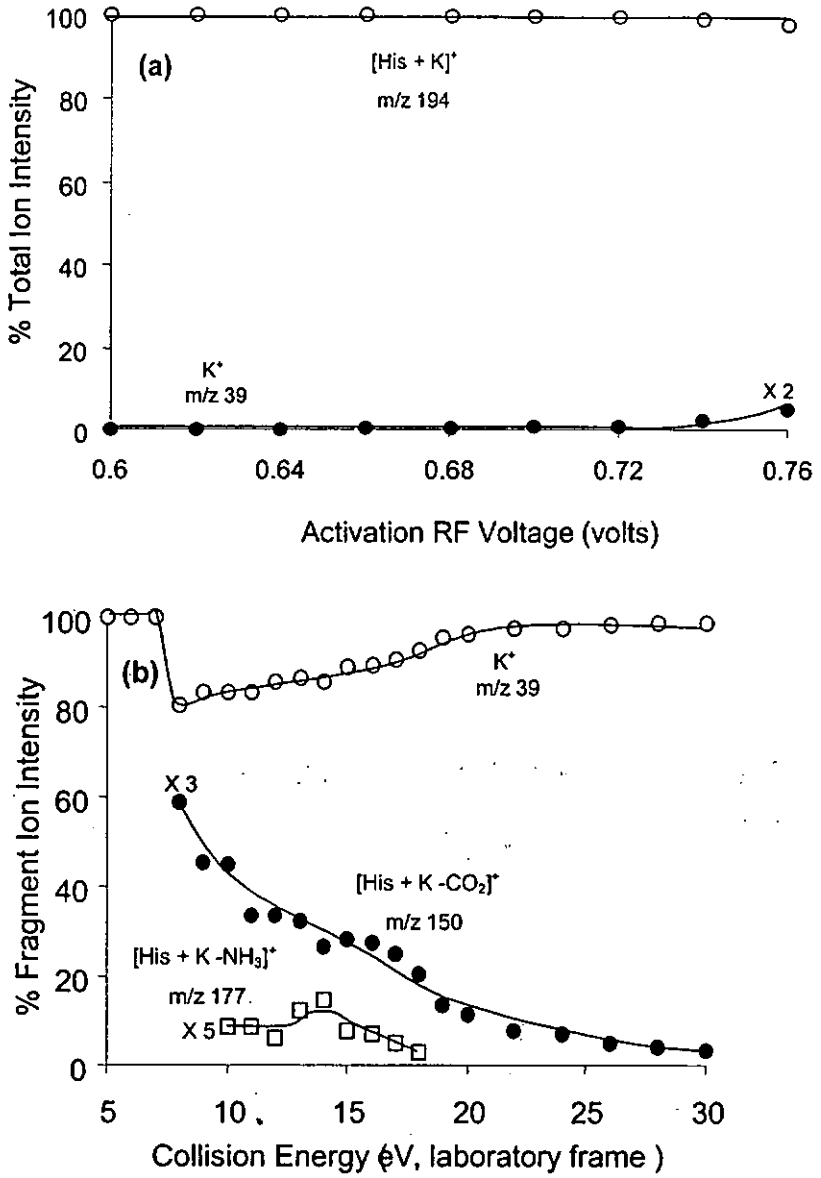


Figure 7.11. (a) Ion trap and (b) triple quadrupole energy resolved MS/MS breakdown graph of potassiated histidine, $[His + K]^+$ (m/z 194).

7.3.5 Stable Binding Modes of Potassiated Histidine

For histidine, because of the presence of functionalized side chain, the K^+ may interact with the ring nitrogen atoms (N^{π} and N^{τ}) or π-cloud of the imidazole ring, in addition to the binding sites available in aliphatic amino acids (NH_2, O=C and OH). However, based on the results presented in Section 3.2.3 on "imidazole" for K^+-imidazole, we concluded that the cation-π mode of binding is not likely in histidine. Furthermore, as N^{τ} is far away from the C-C backbone binding sites, multidentate interaction may not be attainable if the K^+ binds to this nitrogen. Based on all the above, we assume that the K^+ would prefer to bind to more than one of the following sites of binding: NH_2, O=C, OH and N^{π}. Plausible structures of these K^+-His complexes are constructed, and allowed to be fully optimized at the B3-LYP/6-31G(d) level. The four most stable forms of potassiated histidine complex we found theoretically are summarized in Fig. 7.12.

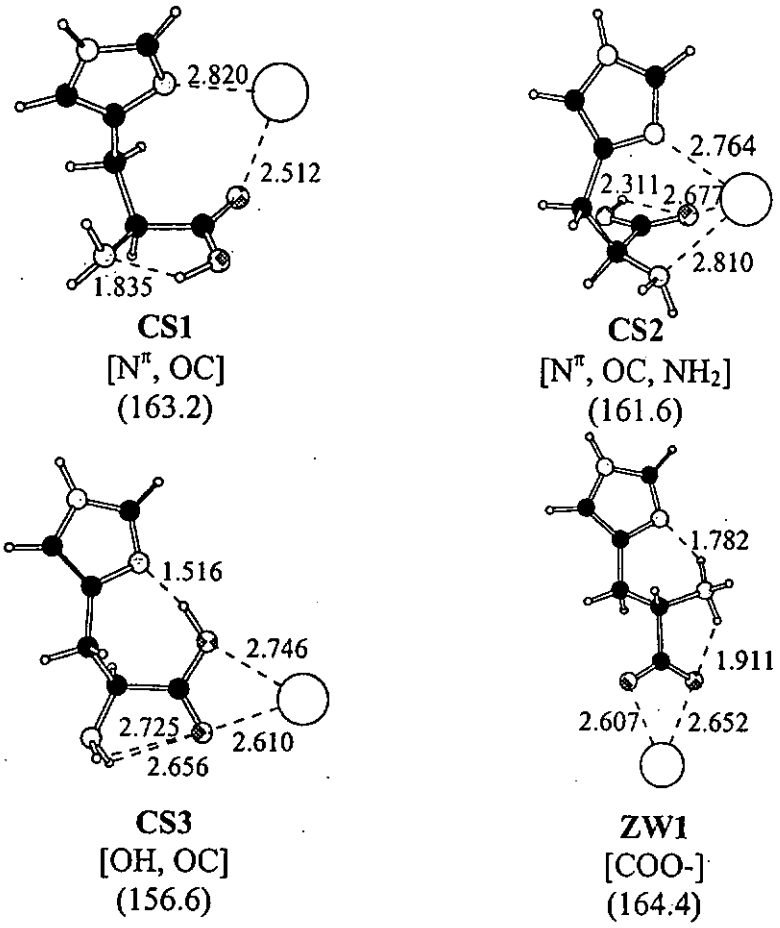


Figure 7.12. The geometries of the four most stable modes of bindings of K^+-His at the B3-LYP/6-31G(d) level. Binding affinities obtained at the B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d) level, with zero-point energy correction at the B3-LYP/6-31G(d) are enclosed in parentheses. Selected non-bond distances are in Angstroms.

We have located four K^+-His complexes: three in the charge-solvated (CS) from, and one in the zwitterionic (ZW) form. The most stable form of charge-solvated complex, **CS1**, involves K^+ bidentately coordinated to N^{π} and O=C, has an estimated affinity of 163 kJ mol^{-1}. The tridentate **CS2** complex, with the K^+ binding to N^{π}, O=C, and NH_2, is approximately 2 kJ mol^{-1} less stable than **CS1**. Thus, comparing the **CS2** to the

CS1 mode, the additional $K^+...NH_2$ binding in **CS2** has apparently destabilized the K^+-His complex.

Interestingly, we note that the most stable mode of binding in potassiated aliphatic amino acids (Chapter 4), i.e., K^+ binds to C=O and OH (denoted **CS3** in Fig. 7.12), is at least 5 kJ mol^{-1} less stable than **CS1/CS2**. Thus, it appears that K^+ binding to N^{π} of imidazole is particularly stabilizing, which is in agreement to our results presented in Chapter 3. For K^+ binding to simply ligands in the monodentate fashion, the binding affinities (in kJ mol^{-1}) is in the order:

Imidazole (111)>> formic acid (80)~ ammonia (77) > water (70)

The K^+ affinity of imidazole is particularly high when the cation binds to the N^{π} of the imidazole ring.

The calculated affinity of the zwitterionic mode (**ZW1**) is 164 kJ mol^{-1}, which is marginally higher (1 kJ mol^{-1}) than that of the **CS1** mode. Thus, it suggests that the K^+-His complex could be in zwitterionic form in the gas phase. Several factors facilitating the formation of zwitterionic complexes have been discussed previously, which include: a high proton affinity (PA) of amino acids, [Wyttenbach et al., 2000] and the formation of strong intramolecular hydrogen bond in the amino acid [Wyttenbach et al., 1999a]. Both criteria are apparently satisfied in potassiated histidine. Histidine is the second most basic amino acid (after **Arg**) in the gas phase, with a PA of 988 kJ mol^{-1} [Hunter and Lias, 1998]. Furthermore, because of the presence of the imidazole side chain, a strong $N^{\pi}...H\text{-}N^{\alpha}$ hydrogen bond (1.782Å) is present in **ZW1** so that the formation of a zwitterionic K^+-His may be favorable. However, this is in contrast to previous findings by Talley et al. [Talley et al., 2002],

who suggested that histidine is coordinated to all of the alkali metal cations (including K^+) in the charge-solvated mode. As the very small difference in relative stability of the **ZW1** over the **CS1** mode is within the error bar of our theoretical protocol (3 kJ mol^{-1}, Section 6.3.3), further investigation is clearly needed. Nevertheless, the key conclusion is: for K^+-His, the zwitterionic **ZW1** mode is of comparable stability with the charge-solvated **CS1** mode. This similarity in stability is apparently reflected in the fragmentation of $[His + K]^+$, as discussed in details below.

7.3.6 Proposed Fragmentation Pathways and Mechanisms for K^+-His

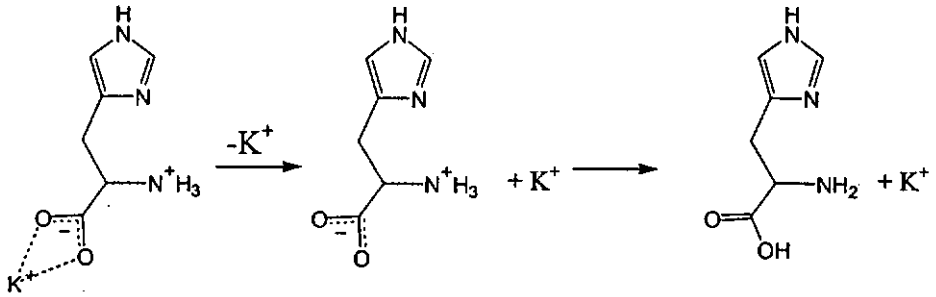
Armed with the very limited theoretical studies conducted and the experimental results presented in Section 7.3.5, we will attempt to rationalize the mass spectrometric fragmentation of potassiated histidine. As our theoretical results suggest that the zwitterionic mode of binding (**ZW1**) is the most favored mode of K^+ binding for **His**, we will assume that the reaction is initiated from this structure for all the mechanisms we postulated below.

Loss of K^+

Previously, it has been noted that the formation of K^+ is common in the MS/MS spectra of potassiated dipeptides. [Grese and Gross, 1990] Thus, given the possibly lower potassium cation affinity of histidine as compared to the dipeptides, it is not surprising that formation of K^+ is predominant in $[His + K]^+$ (Fig. 7.10).

The pathway for formation of K^+ should be trivial as the K^+ could be released from the stable K^+-His complex by simple bond fission. However, as the zwitterionic form of the free histidine ligand is not stable (Section 7.3.2), rearrangement of the free

ligand to the neutral form (**His1**, Fig. 7.3) would be required if the formation of K^+ is to be derived from **ZW1** (Scheme 7.4).



ZW1 **His1**

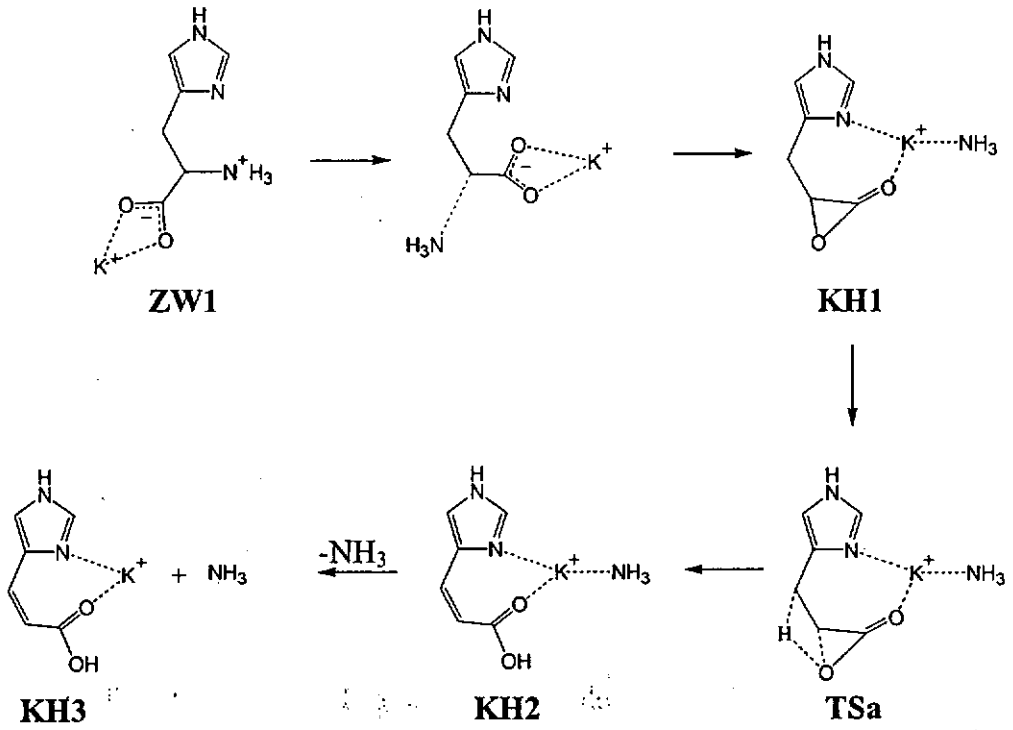
Scheme 7.4. Proposed pathways and mechanism for the loss of K^+ from $[His + K]^+$.

Loss of NH_3

Loss of NH_3 (17 u) is found in the low-energy CID mass spectrum of potassiated histidine. In the literature, loss of NH_3 is often perceived as an indicator of the involvement of ZW complexes in the fragmentation, whereas the loss of H_2O is often associated to a precursor in the charge-solvated form. [Jockusch et al., 1999] This gives a strong indication that under our experimental conditions, the $[His + K]^+$ complex is predominantly in the zwitterionic (ZW) form. This is in agreement with our theoretical predictions presented in Section 7.3.5.

Loss of NH_3 is also found in the MS/MS spectra of $[Ala + K]^+$. [Abirami, 2004] Thus, it is plausible that $[His + K]^+$ yields fragment ions due to loss of this neutral via analogous mechanisms as in $[Ala + K]^+$. It is also possible that the presence of the imidazole ring in the histidine side-chain opens up new pathways that is not available for **Ala**. Based on the assumption that the fragmentation of $[His + K]^+$ leading to the

loss of NH_3 is similar to that of $[Ala + K]^+$, we present our proposed fragmentation schemes below (Scheme 7.5)

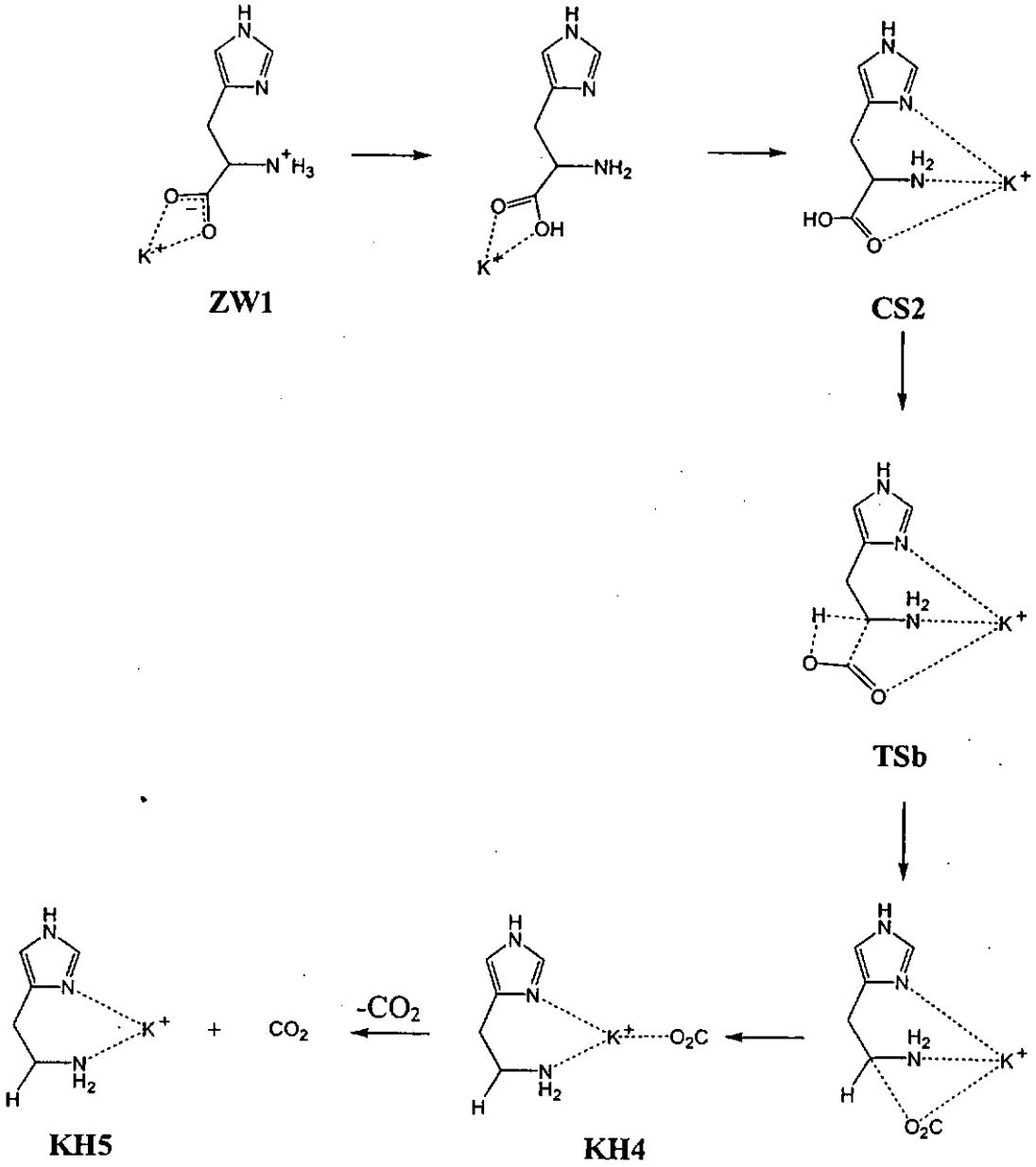


Scheme 7.5. Proposed pathway and mechanism for the loss of NH_3 from $[His + K]^+$, analogous to that for $[Ala + K]^+$ [Abirami et al., 2004].

Essentially, loss of NH_3 occurs via a bond fission between C^{α} and NH_3^+ from **ZW1**. However, in order to stabilize the intermediate, ion-molecule complexes like **KH1** and **KH2** are expected to be formed. The key barrier **TSa,** involving a 1,3 hydrogen shift from C^{β} to the epoxy oxygen, is expected to be more stable than the corresponding transition structure in $[Ala + K]^+$ because of the highly conjugated product formed (**KH3**), as well as via multidentate binding with K^+ strongly coordinated to the N^{π}.

Loss of CO_2

The minor fragment ion due to the loss of CO_2 (44 u) is found in the low-energy CID spectrum of $[His + K]^+$ (Fig. 7.10). Here, we propose a loss of CO_2 pathway and mechanism (Scheme 7.6) based on the decarboxylation of free glycine [Li and Brill, 2003, Scheme 2].



Scheme 7.6. Proposed pathway and mechanism for the loss of CO_2 from $[His + K]^+$, analogous to that for free glycine ([Li and Brill, 2003], Scheme 2).

In this proposed mechanism, decarboxylation is initiated by proton shift from NH_3 in **ZW1**. Through this proton shift, the **ZW1** species is rearranged to the less stable form of potassiated histidine isomer, **CS2**.

A 1,3 hydroxyl hydrogen shift occurs from O_H to C^{α} (via **TSb**), and through the ion-molecule complex **KH4**, leading to the formation of an ion-molecule complex between potassiated histamine (**KH5**) and CO_2. Interestingly, loss of CO_2 is not observed in $[Ala + K]^+$. Thus, the imidazole ring is playing an important role here, presumably by reducing the energetic requirement for the fragmentation via multidentate binding occurring throughout the course of the reaction with K^+ strongly coordinated to the N^{π}.

7.3.7 Fragmentation of Protonated Histidine versus Potassiated Histidine

In Section 7.3.3, we have summarized our theoretical potential energy surface study for the loss of $[CO + H_2O]$, H_2O and $[CO_2 + NH_3]$ from $[His + H]^+$. In Section 7.3.6, we have *proposed* some pathways and mechanisms for the experimentally observed for the formation of K^+, and losses of CO_2 and NH_3 from $[His + K]^+$. Even though we have not obtained the potential energy surface for $[His + K]^+$, a few issues/conclusions can be highlighted with regards to the role of cation in the fragmentation of $[His + K]^+$ versus $[His + H]^+$:

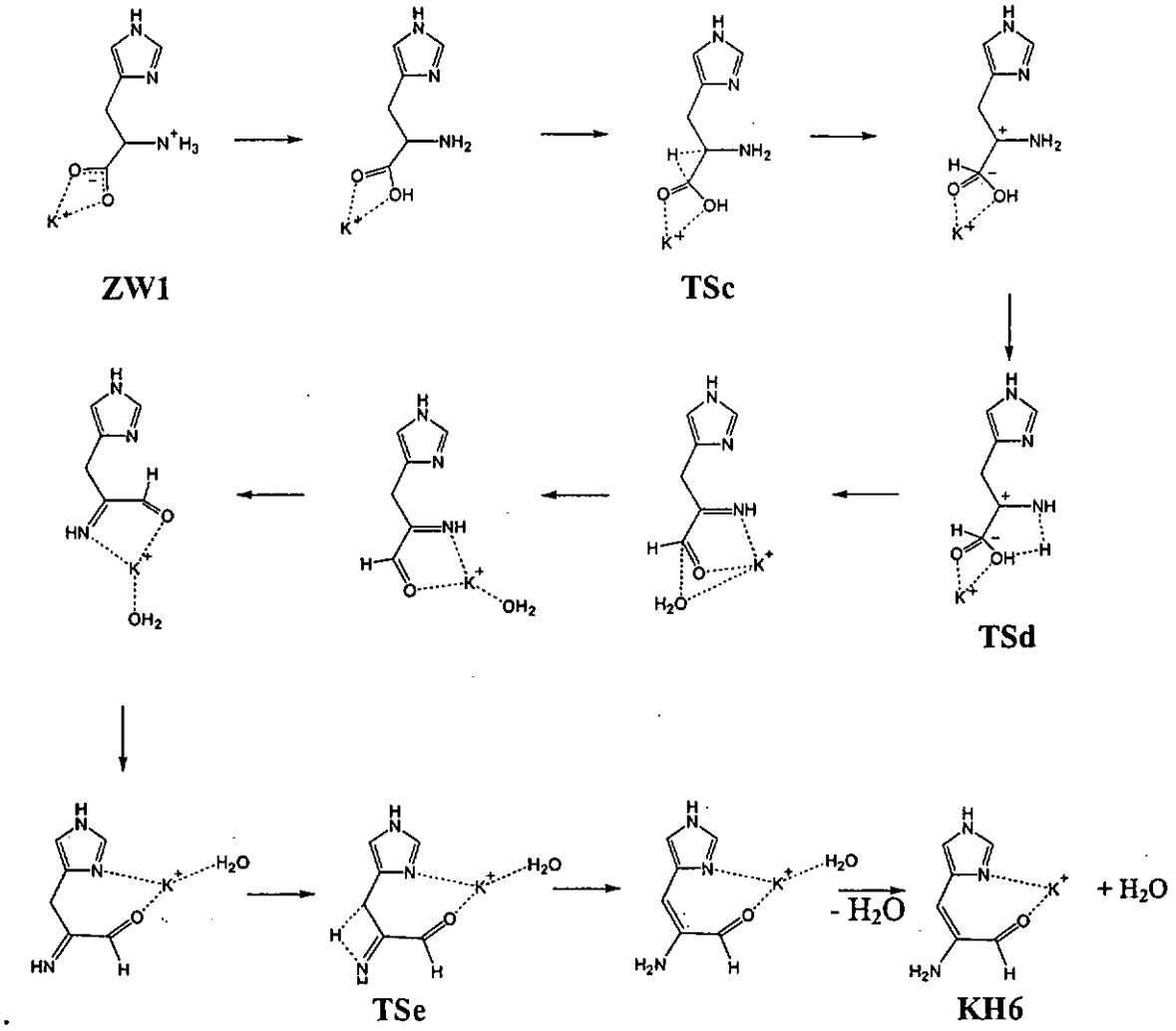
(1) The absence of atomic cation fragment in $[His + H]^+$

Loss of atomic cation (K^+) is dominant in $[His + K]^+$, yet the proton fragment ion, H^+, was not observed in $[His + H]^+$. This difference arises from the nature of cation binding to histidine: while the binding of K^+ to **His** is mainly electrostatic,

the binding of H^+ binds to the amino acid is purely covalent. Because of the difference in nature of binding, the strength of binding differs tremendously. With the proton being more tightly bound than the K^+ (the estimated H^+ and K^+ affinity of **His** is 981 and 164 kJ mol^{-1}, respectively), the loss of H^+ would not be energetically competitive to the loss of other neutrals from protonated histidine.

(2) The absence of loss of [CO + H_2O] neutral from [His + K^+]

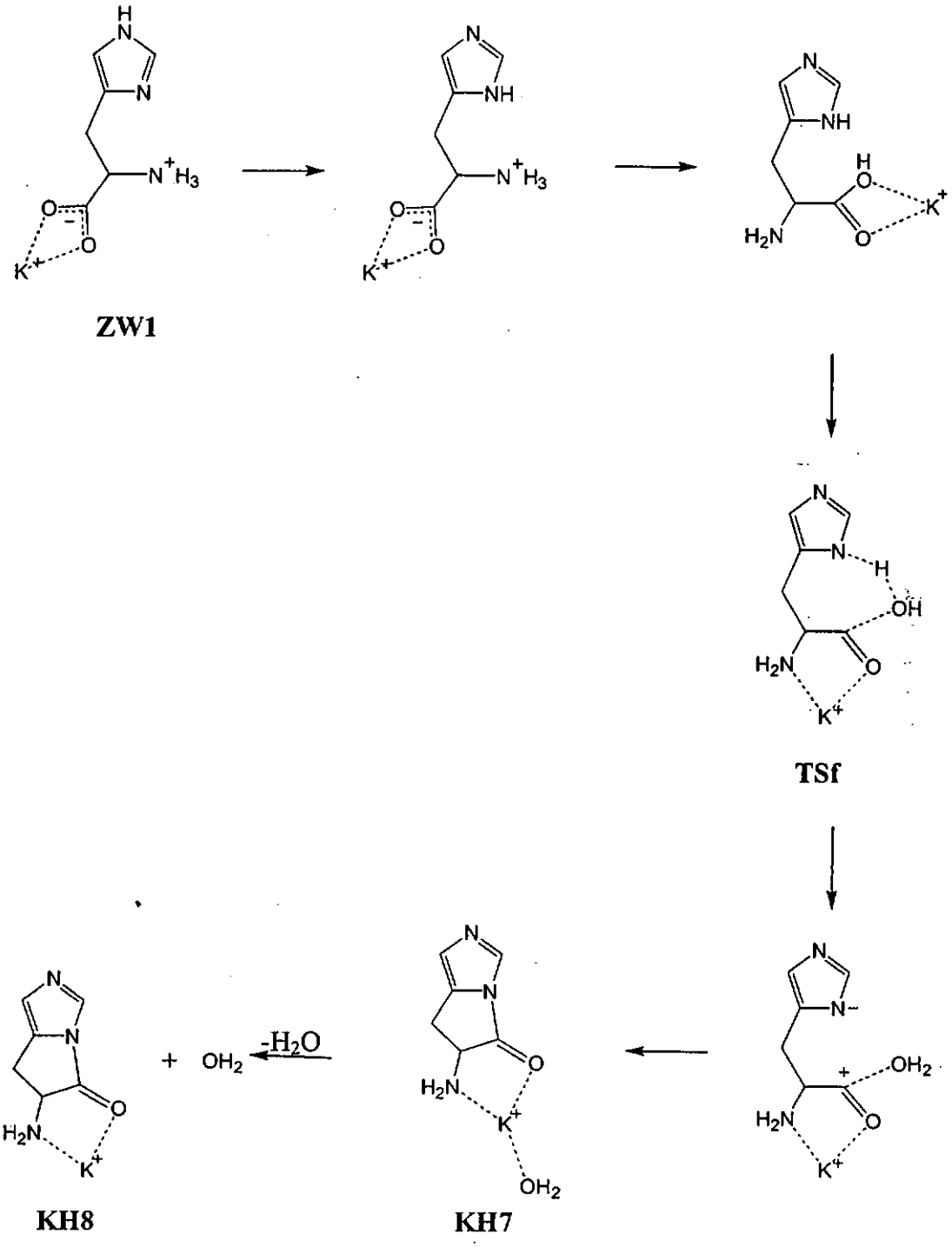
For [His + H^+], the loss of [CO + H_2O] is dominant (together with a minor fragment ion due to the loss of H_2O), while the loss of [CO + H_2O]/H_2O was not found in [His + K]$^+$. In [Ala + K]$^+$, loss of H_2O is found as a minor fragmentation pathway. [Abirami et al., 2004] Based on that study [Abirami et al., 2004], we propose that on analogous mechanism for [His + K]$^+$ (Scheme 7.7). In this mechanism, the loss of H_2O is initiated by a 1,2-hydrogen shift from C^{α} to C_1 (**TSc**), followed by a 1,4-hydrogen shift from the amino nitrogen to hydroxyl oxygen (**TSd**). In the case of [Ala + K]$^+$, the energy barrier for the loss of NH_3 is lower than that for the loss of H_2O by 7 kJ mol^{-1}. [Abirami et al., 2004] It is possible that the energy difference between the corresponding barriers (**TSc/TSd**, Scheme 7.7 versus **TSa**, Scheme 7.5) are even larger in [His + K]$^+$, leading to the absence of loss of H_2O pathway in this case.



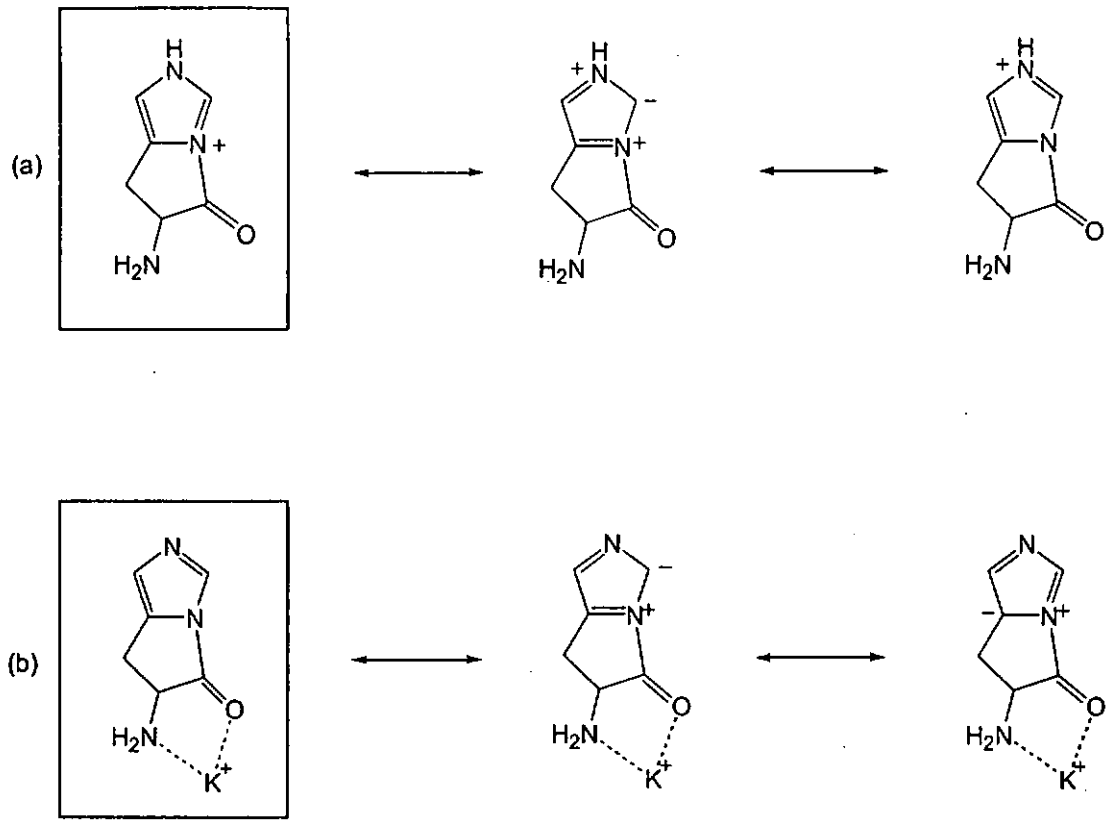
Scheme 7.7. Proposed pathway and mechanism for the loss of H_2O from $[His + K]^+$, analogous to that for $[Ala + K]^+$ [Abirami et al., 2004].

Apart from the above scheme, we have also considered whether $[His + K]^+$ may loss $[CO + H_2O]/H_2O$ in a fashion akin to $[His + H]^+$ (discussed in Section 7.3.3). Such scheme is presented in Scheme 7.8, in which the loss of H_2O is initiated by a proton shift from N^{π}-H to the hydroxyl oxygen (O_H). In Section 7.3.3, we mentioned that because of the conjugation between carbonyl C=O and the imidazole ring in species **11** (Fig. 7.5), loss of H_2O is a favorable process in $[His + H]^+$. Compared to $[His + H]^+$, $[His + K]^+$ is *proton-deficient*. With one proton less, the C=O group of the final product for the loss of H_2O is in fact not in conjugation with the imidazole ring in the

analogous K^+ complex (Scheme 7.9). With the loss of stability, this mechanism becomes less viable, thus, explaining the absence of the loss of H_2O pathway (and related fragment ions) in $[His + K]^+$.



Scheme 7.8. Proposed pathway and mechanism for the loss of H_2O from $[His + K]^+$, analogous to that for $[His + H]^+$.

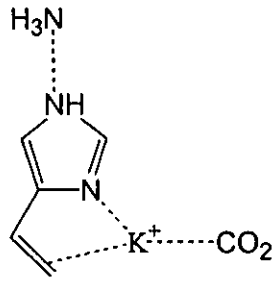


Scheme 7.9. Resonance form of (a) species **11** in $[His + H]^+$, (b) the analogous species in $[His + K]^+$, i.e., species **KH8** in Scheme 7.8. The dominant form is enclosed in a box.

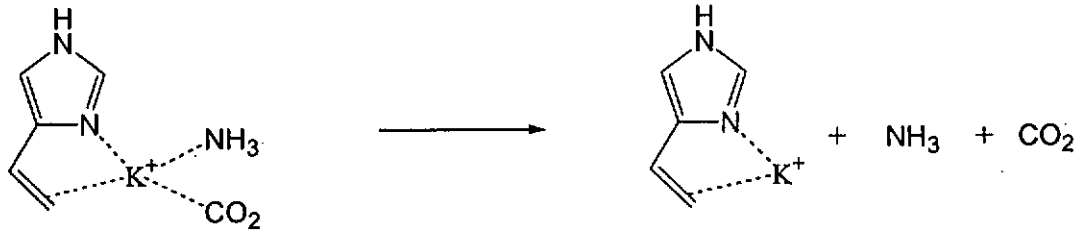
(3) Carbon dioxide and ammonia are lost together in $[His + H]^+$, but by two separate neutrals in $[His + K]^+$.

CO_2 and NH_3 are lost in the fragmentation of *both* protonated (Fig. 7.1) and potassiated histidine (Fig. 7.10). However, in the former case, the two neutrals are lost together, i.e., loss of 61 u, while in the latter case, they are observed as leading to individual fragment ions. Thus, the loss of CO_2 and of NH_3 in $[His + K]^+$ is likely to go through different mechanisms as compared to the simultaneous loss of $[CO_2 + NH_3]$ pathway we formed for $[His + H]^+$ (Fig. 7.7), leading to our proposed Schemes 7.5 and 7.6, respectively.

Without carrying out detailed theoretical studies on potassiated histidine, it is difficult to explain this difference. However, we wish to note that one of the key structures we proposed for the loss of $[CO_2 + NH_3]$ in $[His + H]^+$, species **14** (Fig. 7.7), in which CO_2 and NH_3 are coordinated to the imidazole ring via the hydrogen attached to N^{π} and N^{τ}, respectively. In the case of $[His + K]^+$, with the H^+ on N^{π} replaced by K^+, the analogous species of **14** would be **KH9** (Scheme 7.10). However, as the $K^+...NH_3$ interaction is strong, the formation of ion-molecule complex **KH10** (Scheme 7.10) is more likely. From this, the simultaneous loss of $[CO_2 + NH_3]$ is certainly possible, yet it is not observed experimentally. However, without further investigation, it is impossible to make further comment on the experimental observations.



KH9



KH10 **KH11**

Scheme 7.10. Key structures for the loss of $[CO_2 + NH_3]$ from $[His + K]^+$. Species **KH9** corresponds to the analogous species of **14** reported for $[His + H]^+$, stabilized by hydrogen bonding of $N^1H...NH_3$; species **KH10** is the alternative structure of species **14**, which is stabilized by electrostatic interaction of $K^+...NH_3$.

7.4 Conclusion

Detailed theoretical studies on $[His + H]^+$ and its fragmentation in the gas phase have been conducted. Our estimated proton affinity and gas phase basicity are in excellent agreement (2 kJ mol^{-1}) with experimental values; the most favorable protonation site for histidine is the N^{π} on the imidazole side chain. The mechanism for the loss of [CO + H_2O] (46 u), H_2O (18 u), and [CO_2 + NH_3] (61 u) from $[His + H]^+$ were also studied. The potential energy surface obtained for these reactions is found to be valuable in explaining the experimentally observed mass spectrometric fragmentation patterns.

When compared to protonated histidine, the fragmentation pattern of potassiated histidine is entirely different. While the dominant fragment ion in $[His + H]^+$ corresponds to the loss of [CO + H_2O], such fragment ion is absent in the CID MS/MS spectrum of $[His + K]^+$. Instead, formation of K^+ is dominant in potassiated histidine, with minor fragment ions arising from the losses of CO_2 and NH_3. Based on our limited theoretical studies on $[His + K]^+$, the difference in fragmentation behavior between the protonated and the potassiated histidine can be traced back to the difference in the nature of H^+ versus K^+ binding to the amino acid, and the ability of the K^+ in stabilizing the zwitterionic form of the histidine ligand.

Furthermore, by comparing with protonated/potassiated alanine/glycine, the role of the imidazole ring in the gas phase fragmentation of $[His + H]^+$ and $[His + K]^+$ has been elucidated. Firstly, the presence of the imidazole ring allows the formation of stable transition structures and intermediates via hyperconjugation, thus changing the mechanism for the losses of neutrals (when compared to those of protonated/potassiated aliphatic amino acids). Secondly, low energy pathways for the

loss of $[CO_2 + NH_3]$ in $[His + H]^+$ are accessible, with stable ion-molecule complexes as key reaction intermediates when small neutrals like CO_2 and NH_3 are "anchored" onto the imidazole ring. Thirdly, for the loss of CO_2 and of NH_3 from $[His + K]^+$, the imidazole ring plays only a minor role in the fragmentation, apart from providing an additional site (via N^{π}) for K^+ binding.

Chapter 8 Conclusion:

In this thesis, we reported our theoretical studies, at one uniform level of theory (Chapter 2), on the structures and energies (affinities) of K^+ binding to: small organic ligands (Chapter 3), aliphatic amino acids (Chapter 4), aliphatic dipeptides (Chapter 5), aromatic dipeptides (Chapter 6) and basic functionalized amino acids (Chapter 7). The following key conclusions can be drawn:

(1) Performance of the proposed energetic protocol EP(K^+):

The protocol, B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d), abbreviated as EP$(K)^+$ here, yields *absolute* K^+ affinities for model organic ligands to within ±4.5 kJ mol^{-1} for 65 model ligands when compared to available experimental affinities in the literature. When compared to experimental affinities of dipeptides obtained by the mass spectrometric kinetic method, the *absolute* affinities are systematically higher by 8-15 kJ mol^{-1}, with an estimated error bar of 3 kJ mol^{-1} in *relative* affinities.

In general, the discrepancy between theoretical and experimental K^+ affinities increases with the size of the molecular systems. While the difference could arise from the approximations employed in the theoretical modeling method, it could also be attributed to experimental errors. For the larger ligands, increase in inter-ligand repulsion in the K^+-bound heterodimer ion are more severe that the mode of binding formed in the heterodimer may not be the same as that in the K^+ monomer complex, thus, leading to a reduction in measured experimental K^+ affinities.

(2) Factors affecting the stability of K^+ binding to ligands

The interactions between K^+ and organic/amino acids/dipeptides are mainly electrostatic in nature, and in general, more similar to that of Li^+/Na^+, than Ag^+. The overall interaction arises from the contribution of both stabilizing and destabilizing factors.

Polarizability, dipole moment and co-ordination number between K^+ and binding sites (dentate) are the three key stabilizing factors in K^+ binding to ligands. Linear correlation between K^+ binding affinity and these three factors has been obtained for small model organic ligands. For a series of related ligands with little variation in molecular dipole moment, and if the mode of binding is identical with changes mainly in alkyl side chain length (e.g., K^+ binding to O=C, OH in **Gly/Ala/Val/Leu/Ile**; K^+ binding to amide O=C, C-terminal O=C in **GG/AA** and **GG/FG/GF**), then the increase of binding affinities is found to correlate to the polarizability of the side chain.

We also looked at the underlying factors that govern the stability of different modes of binding in the same ligands. In order to include the effect of binding geometries, the effect of molecular polarizability and dipole moment can be described by the polarization interaction parameter (PIP) and dipole interaction parameter (DIP) terms, respectively. While the correlation between PIP and DIP to the overall stabilization energy of different modes of binding in K^+-Gly is almost linear, the same correlation cannot be extended to the dipeptides. We propose that the failure arises from the more complex electronic distribution in the dipeptides than in glycine, so that simple parameters like DIP and PIP cannot describe accurately the underlying physico-chemical basis of the interactions. As an example, *local* interaction, as opposed to

global, between K^+ and the amide carbonyl oxygen is energetically very much preferred in the dipeptides, which has been attributed to the very strong *local* ion-dipole interaction between K^+ and the peptide amide C=O bond.

The increasing stabilizing factor is often associated with increases in destabilizing forces. To bind to K^+, the hydrogen bonding network in free amino acids/dipeptides are usually perturbed. To achieve multidentate interaction, the various negatively charged binding sites are located closer to each other than otherwise would have been in the case of free ligand, which would in turn increase the intramolecular electrostatic repulsion. Furthermore, to increase the electrostatic interaction, the free amino acids/dipeptides may rearrange to the zwitterionic form, in which the carboxylate end of these species bears the formal negative charge in order to facilitate the binding of the positively charged potassium cation. The disruption of the hydrogen bonding pattern, increase in intramolecular electrostatic repulsion, and the increase in charge-separation in the ligand in the ZW form, are the main sources of destabilization, which is represented quantitatively and collectively as the deformation energy, E_{def}, in this work. Thus, it is not surprising that the tridentate **CS2** (O_1', O_2', N_1) modes is not more stable than the bidentate **CS1** (O_1', O_2') modes for dipeptides **GG/AA/FG/GF**, as the deformation energies for the tridentate modes are always higher, and may even outweigh the stabilization factor bought along by multidentate binding. Similarly in K^+-FG/GF, it is not surprising that not all π modes are more stable than the corresponding non-π modes, as the larger deformation energy for the π modes may again outweigh the stabilization factor of K^+-π interaction.

(3) Factors affecting the relative stability of charge-solvated (CS) and zwitterionic (ZW) modes of K^+ binding

For a given amino acid/peptides, adopting the zwitterionic form is intrinsically destabilizing in the gas phase. For a given cation (K^+ in this case), the key factor favoring the formation of the ZW form is the proton affinity (PA) of the ligand. Our studies in K^+ binding to amino acid supports this view as the PA criterion can be used to explain the stability of ZW species in K^+-His, and the increase in the stability of the ZW form in the aliphatic amino acid series, **Gly/Ala/Val/Leu/Ile**. However, we wish to caution against extending the PA criterion to compare/predict the relative stability of different ZW structures of amino acids or peptides having dissimilar metal cation binding modes. This is illustrated by cross comparisons between the aliphatic dipeptides and aliphatic amino acids (i.e. compare **GG/Gly** pair or **AA/Ala** pair).

Apart from PA, it has been suggested previously that a close to linear + - + arrangement for amino hydrogen, carboxylate oxygen, and metal cation (<H...O...M) is one of the factors stabilizing the zwitterionic complex in Na^+-Pro. In the case of aliphatic amino acids, the deviation from linearity is not significant, nor have any correlation to the variation of stability of the ZW forms in **Gly/Ala/Val/Leu/Ile**. It has been suggested previously that the formation of a strong hydrogen bond from the functionalized side chain is another stabilizing factor for the zwitterionic complex in metal cationized **Arg**. There is a strong hydrogen bond between the imidazole side chain and the amino group in K^+-His. Based on this, we concluded that the latter factor is only of secondary importance in the stabilization of zwitterionic complexes.

For all the dipeptides we studied here (**GG/AA/FG/GF**), the CS binding modes are in generally more stable than the ZW binding modes. Even though the ZW modes can

be stabilized somewhat if the dipeptide adopts a more *compact*, rather than an *extended* form, it appears that the instability arises from the charge-separation effect in the larger/longer dipeptides would outweigh the gain of electrostatic interaction between K^+ and ^-OOC in general.

(4) Preferred mode of K^+ binding to amino acids and dipeptides

Because of the interplay of the various stabilizing and destabilizing factors outlined above, and for the few amino acids/dipeptide we studied here (aliphatic amino acids, histidine, **GG/AA/FG/GF**), the following *generalization* can be made:

(i) The most favorable single site of K^+ binding to the amino acid and dipeptide is the carbonyl oxygen, O=C, with the ligand in the CS mode.

(ii) The optimal number of interaction between K^+ and ligands is two, i.e., bidentate interaction.

(iii) For aliphatic amino acid with no functional side chain, K^+ prefers to bind to O=C and OH. For aliphatic dipeptides, binding to the two O=C is most favorable.

(iv) For amino acids with functional side chain, e.g. aromatic-π in **Phe** and imidazole in His, K^+ may additionally bind to the functional group site on the side chain. For the dipeptides **FG** and **GF**, binding to the two O=C (**CS1**) and two O=C, π (**CS1-π**) are very favorable, with comparable K^+ affinities for K^+-GF.

(v) The ZW modes are in generally not competitive against CS modes; the only exceptional case is K^+-His due to the basicity of the imidazole ring in

histidine. Out of the three possible forms of ZW dipeptides (protonation at peptide bond nitrogen, peptide bond oxygen, and terminal amino nitrogen), the **ZW(N_2)** form (protonation at peptide bond nitrogen) is least favored because of the loss of resonance stabilization.

(5) Relation between cation (H^+/K^+) binding in amino acids to mass spectrometric fragmentation

The mechanisms for the loss of [CO + H_2O] (46 u), H_2O (18 u), and [CO_2 + NH_3] (61 u) from $[His + H]^+$ are studied. The potential energy surface obtained for these reactions is found to be consistent with the experimentally observed fragmentation patterns. Furthermore, by comparing with $[Ala + H]^+$, the role of the imidazole ring in the gas phase fragmentation of $[His + H]^+$ has been elucidated. When compared to $[His + H]^+$, the fragmentation pattern of $[His + K]^+$ is entirely different. Based on our limited theoretical studies on $[His + K]^+$, the difference in fragmentation behavior between the protonated and potassiated histidine can be traced back to the difference in the nature of H^+ versus K^+ binding to the amino acid, and the ability of the K^+ in stabilizing the zwitterionic form of the histidine ligand.

In this thesis, we have comprehensively studied the interaction between K^+ and small model amino acids/peptides. As ionic species are usually stabilized by interaction with polar solvent molecules like water, better knowledge of the interaction of K^+ with 'hydrated' amino acids/peptides is a vital bridge to understanding and prediction of biomolecular properties in vivo. However, direct comparison, in particular the quantitative comparison of K^+ affinities, with such water solvated systems in the gas phase is experimentally difficult, and the situation could be complicated by

competitive binding of the alkali metal cations to the O/N heteratom binding sites of amino acid/peptides at the same time.

In biological systems, K^+ binding to backbone carbonyl oxygens in a macrocyclic pattern are found in the X-ray protein structures of K^+ channel [Doyle et al., 1998], tryptophanase [Isupov et al., 1998] and pyruvate kinase [Larrsen et al., 1994], suggesting that our current study of identifying the importance of local $K^+\cdots O{=}C$ ion-dipole interaction may be extendible from the gas phase to the solution phase. Apart from these, gas-phase studies of model systems, such as those comprised of a single amino acid, a metal ion, and individual water molecules may provide a means for investigating the solvent effects. An example is the work of Williams and co-workers, [Jockusch et al., 2001] who investigated the structures of gas-phase amino acid-K^+-water complexes theoretically and experimentally. Their calculations indicated that K^+ is coordinated to both oxygens (CO, OH) of nonzwitterionic valine, which is in line with our findings. The addition of a single water molecule does not significantly affect the relative K^+ binding energies calculated for potassiated valine clusters. Experimentally, relative dissociation kinetics of Val-$K^+(H_2O)_{1\rightarrow0}$ also suggested that potassium was coordinated to both oxygens, but could not distinguish whether the valine was zwitterionic or nonzwitterionic from experimental observations alone. [Jockusch et al., 2001]

In summary, much more experimental studies, along with development of new experimental techniques for observing the interaction of metal ions with amino acid/peptide-water cluster complexes, are desired and necessary before the effect of solvent molecules on metal cation binding affinities to bio-molecules can be elucidated.

Appendix I: Supporting information for Chapter 3

Table S-3.1. The B3-P86 and G2(MP2,SVP) potassium cation binding affinities (kJ mol^{-1}) at 298K.

Molecule [a]	B3-P86	G2(MP2,SVP)	Molecule [a]	B3-P86	G2(MP2,SVP)
He	2.8		n-PrCHO	96.1	
Ne	2.4		n-BuCHO	97.3	
Ar	8.4		CF_3CHO	58.4	
CO	24.6		CCl_3CHO	73.1	
HF	49.8		Me_2CO	103.7	
HCl	26.9		MeCOEt	105.0	
P_4	33.6		NH_3	77.8	76.7
PH_3	42.4		$MeNH_2$	79.3	80.5
C_2H_2	37.6		Me_2NH	77.6	81.2
C_2H_4	35.0		Me_3N	74.8	80.3
H_2S	39.3		$EtNH_2$	81.7	84.1
MeSH	52.5		n-$PrNH_2$	81.6	85.6
EtSH	57.2		HCN	77.6	
n-PrSH	59.0		MeCN	101.4	
i-PrSH	60.0		EtCN	105.0	
n-BuSH	60.3		n-PrCN	106.4	
i-BuSH	63.1		i-PrCN	107.7	
t-BuSH	61.7		t-BuCN	109.9	
Me_2S	61.8		$PhCH_2CN$	109.7	
H_2O	69.9	70.3	CF_3CN	58.4	
MeOH	74.2	76.7	CCl_3CN	76.0	
EtOH	80.2	83.6	HCO_2H	79.6	
n-PrOH	81.1	83.5	HCO_2Me	88.6	
i-PrOH	84.1	87.5	$MeCO_2H$	91.0	
n-BuOH	84.7	88.2	HCO_2Et	93.7	
i-BuOH	82.1	86.0	HCO_2n-Pr	95.3	
s-BuOH	86.4	89.8	$MeCO_2Me$	98.2	
t-BuOH	87.1	91.3	$MeCO_2Et$	104.4	
1,2-Propanediol	114.0	119.1	$EtCO_2Me$	99.7	
1,3-Propanediol	118.7	123.8	CF_3CO_2Me	80.7	
Ethylene glycol	114.7	119.3	$ClCO_2Me$	79.8	
Glycerol	127.9	136.4	SO_2	49.8	
CF_3CH_2OH	69.7		Me_2SO	126.7	
CCl_3CH_2OH	74.7		PhSOMe	130.0	
Me_2O	73.2		$HCONH_2$	113.3	113.2
Et_2O	83.9		HCONHMe	118.8	120.9
$(MeOCH_2)_2$	122.1		$HCONMe_2$	124.6	126.6
1,2-dioxane	82.3		$MeCONH_2$	122.5	122.2
1,3-dioxane	82.3		MeCONHMe	126.6	129.0
1,4-dioxane	69.3		$MeCONMe_2$	128.7	131.7
HCHO	76.3		Benzene	70.8	78.6
MeCHO	92.6		Borazine	48.8 (48.7) [b]	57.6
EtCHO	94.6		Phenol	73.7 (73.6) [b]	80.7

Pyridine	92.6		3,4,5-Me_3-pyrazole	104.8	
2-Me-pyridine	94.2		1,3,4,5-Me_4-pyrazole	105.9	
3-Me-pyridine	98.6		1-Me-imidazole	118.4	
4-Me-pyridine	100.7		1,2-Me_2-imidazole	120.2	
2-F-pyridine	98.6		2,4,5-Me_3-imidazole	122.2	
3-Cl-pyridine	81.5		2H-1,2,3-triazole	63.7	
1,8-naphthyridine	155.4		1H-1,2,4-triazole	86.8	
Pyridazine	129.6		2H-tetrazole	88.3	
Pyrimidine	75.0		1H-tetrazole	109.8	
Pyrazine	69.1		4H-1,2,4-triazole	140.3	
1,3,5-triazine	55.8		1H-1,2,3-triazole	118.7	
Pyrrole	79.9 (79.3) [b]	87.0	Glycine	122.1 (121.4) [b]	118.3
Indole	93.2 (93.0) [b]	102.8	Alanine	128.0 (127.4) [b]	122.5
Pyrazole	89.8		Valine	131.9 (131.5) [b]	126.5
Imidazole	111.0		Leucine	132.8 (132.6) [b]	127.3
Thiazole	86.8		Iso-Leucine	133.5 (133.0) [b]	128.5
Isothiazole	85.2		Proline	142.3	
Oxazole	83.3		Serine	136.3	
Isoxazole	92.6		Cysteine	121.4	
1-Me-pyrazole	93.9		Phenylalanine	144.0	
3-Me-pyrazole	92.2		Adenine	86.6 [b]	
4-Me-pyrazole	95.8		Thymine	110.9 [c]	
1,4-Me_2-pyrazole	99.5		Uracil	111.9 [c]	
1,5-Me_2-pyrazole	100.9		Cytosine	166.9 [c]	
1,3,5-Me_3-pyrazole	103.1		Guanine	143.7 [c]	

[a]: Abbreviations: Me = $-CH_3$, Et = $-C_2H_5$, n-Pr = $-C_3H_8$, i-Pr = $-(CH_3)_2CH$, n-Bu= $-C_4H_{11}$, i-Bu = $-(CH_3)_2CHCH_2$, t-Bu= $-(CH_3)_3C$, Ph= $-C_6H_5$. [b]: The value in bracket is the PCA calculated using B3-P86/6-31G(d) geometry, i.e., at the B3-P86/6-311+G(3df,2p)//B3-P86/6-31G(d) level. The similarity between the two sets of PCAs suggests that the effect of geometry difference: (B3-LYP/6-31G(d) versus B3-P86/6-31G(d)), on binding affinity is negligible. [c]: The PCA is estimated for the most stable tautomer of the free ligands, corresponding to species A1, T1, U1, C1 and G1 in ref. [Russo et al., 2001b].

Table S-3.2. The effect of B3-LYP/6-31G(d) correction on zero point vibrational energies, temperature corrections; and the effect of full counterpoise correction.

Species	$\Delta(ZVPE)^{[a]}$	$\Delta(H_{298})^{[b]}$	$\Delta(BSSE)^{[c]}$
CO	-0.39	-0.08	-0.64
C_2H_2	-0.77	-0.45	-0.39
MeSH	-0.42	-0.14	-0.61
H_2O	-0.58	-0.42	-0.53
HCHO	-0.50	-0.41	-0.63
NH_3	-0.83	-0.56	-0.30
HCN	-0.26	-0.12	-0.41
HCO_2H	-0.45	-0.36	-0.65
SO_2	-0.15	-0.11	-0.81
Benzene	-0.30	-0.14	-1.77
Pyridine	-0.49	-0.32	-0.64
Pyrrole	-0.47	-0.21	-1.18
Glycine	-0.52	0.47	-0.75
Uracil[d]	-0.61	-0.49	-0.76

[a]: Difference in zero-point vibrational energy corrections (ZVPE), $\Delta(ZVPE)$ = ZVPE calculated at the B3-LYP/6-31G(d) level - ZVPE calculated at the HF/6-31G(d) level, in kJ mol^{-1}. [b]: Difference in thermal corrections, $\Delta(H_{298})$ = temperature correction calculated at the B3-LYP/6-31G(d) level - temperature correction at the HF/6-31G(d) level, in kJ mol^{-1}. [c]: $\Delta(BSSE)$ = full counterpoise correction, including effect of geometry deformation, ref. [Siu, et al., 2001a], at the B3-LYP/6-311+G(3df,2p) level of theory, in kJ mol^{-1}, based on B3-LYP/6-31G(d) geometries. [d]: The PCA is estimated for the most stable tautomer of the free ligands U1 in ref. [Russo et al., 2001b].

Appendix II: Supporting information for Chapter 4.

Table S-4.1. Calculated binding affinities at 0K (ΔH_0, in kJ mol^{-1}) of various K^+-Gly complexes at the B3-LYP/6-311+G(3df,2p) level using geometry obtained at different levels of theory.

Species	ΔH_0 [a]	ΔH_0 [b]	ΔH_0 [c]
CS1	117.7	117.5	118.2
CS2	115.2	114.8	115.7
CS3	84.4	84.0	85.0
CS4	75.9	75.4	75.6
CS5	67.2	66.9	67.0
ZW1	104.5	103.8	104.2
CS1'	80.2	82.5	80.0
CS2'	96.4	96.0	96.0
CS5'	55.6	55.2	56.3
CS5"	38.5	38.1	39.2

[a]: Using B3-LYP/6-31G(d) optimized geometry (with potassium basis set in ref. [Frisch et al., 1998]), abbreviated as "EP(K^+)" in this work. [b]: Using B3-LYP/6-31+G(d) optimized geometry (with potassium basis set in ref. [Frisch et al., 1998]). [c]: Using B3-LYP/6-31G(d) optimized geometry (with potassium basis set in ref. [Blaudeau et al., 1997]).

Appendix III: Supporting information for Chapter 5.

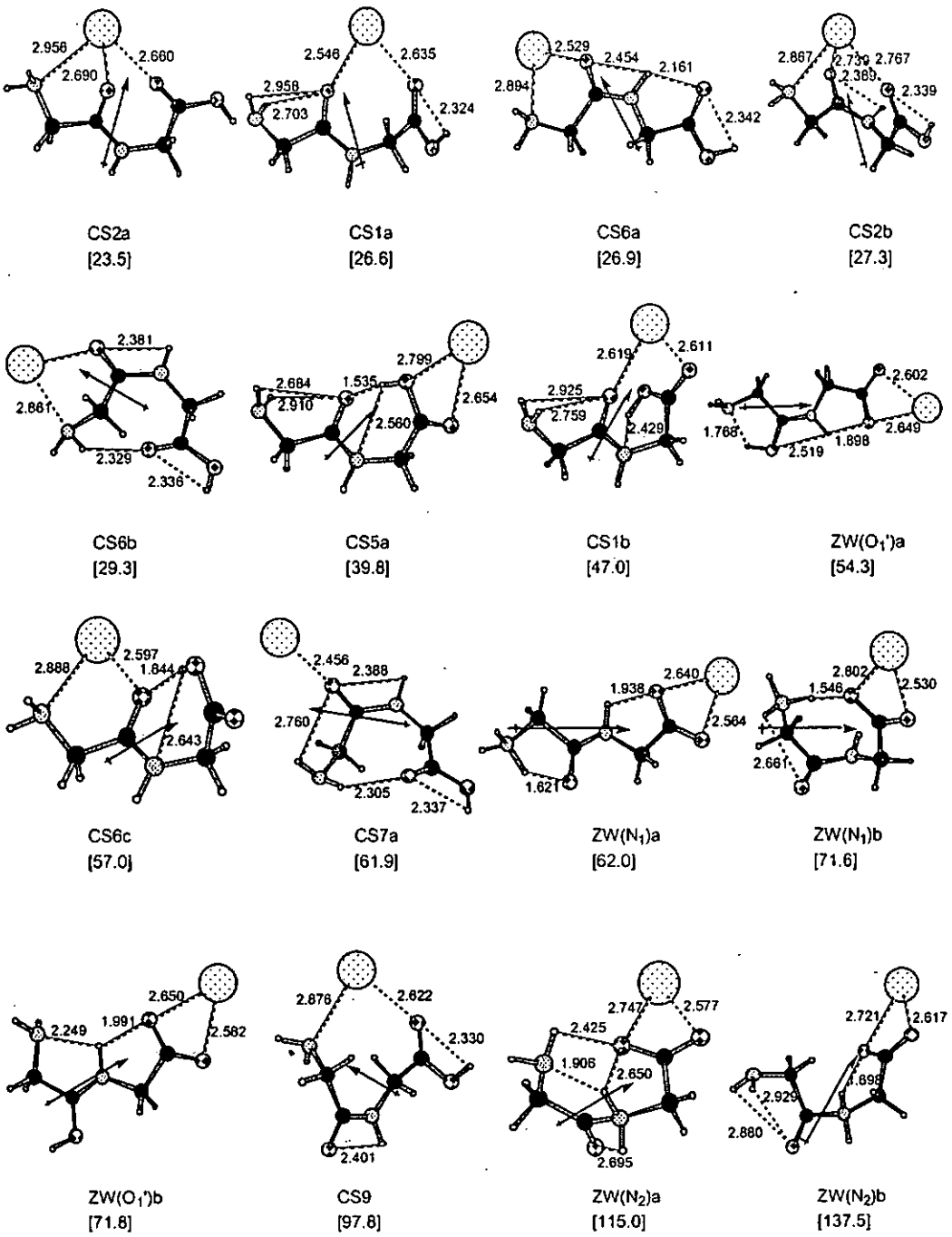


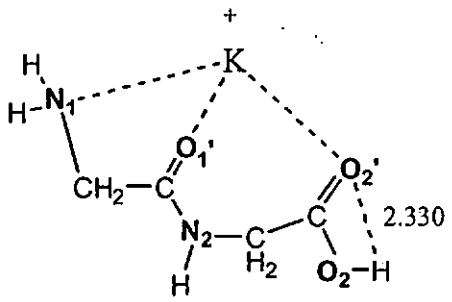
Figure S-5.1. Summary of the structures and energetics of low-lying conformers on the K^+–GG potential energy surfaces other than those shown in Figs. 5.2 and 5.3. Selected non-bond distances are shown in Angstroms. The molecular dipole moment vectors of the *deformed* AA ligand (not to scale) are included for reference. The K^+ binding energies at 0K of these complexes, without zero-point energy correction, relative to that of CS1, are included in brackets.

Appendix IV: Supporting information for Chapter 5.

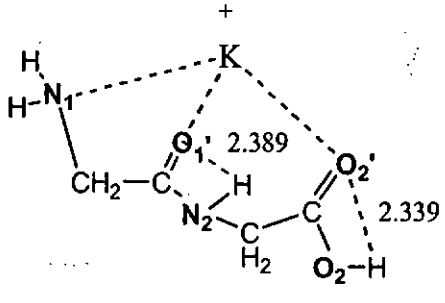
Relative Stabilities of Conformers Sharing the Same Modes of Binding in K^+-GG:

The relative stabilities of conformers sharing the same modes of K^+ binding can be attribute to two main factors:

(1) conformation of peptide bonds: By comparing the conformer pair: **CS2/CS2b**; **ZW(O_1')/ZW(O_1'a)** and **CS6/CS6a** (see Figs. 5.2, 5.3 and S-5.1), we found that the *trans* conformation is more stable than the *cis* conformation (by 6-22 kJ mol^{-1}). This is in line with the general expectation that peptide bonds prefers to be in the *trans* configuration. [Luque and Orozoc, 1993]



CS2



CS2b

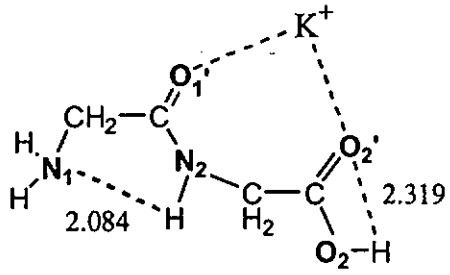
(2) Hydrogen bonds and other non-bonded interactions: Hydrogen bonds play a dominant role in the structure and function of proteins including features such as protein folding, protein-legend recognition, enzymatic activity, protein hydration, and material design. [Desiraju and Steiner, 1999; Sarkhel and Desiraju, 2004] The common hydrogen bonds found in aliphatic peptides includes N-H...O, O-H...O, [Jeffrey and Saenger, 1991; Sarkhel and Desiraju, 2004], with bond distances ranges from 2.03 to 2.55Å [Sarkhel and Desiraju, 2004]. Previously,

Zhang et al. [Zhang et al., 1993] suggested that the strength of non-bonded interaction (including hydrogen bond) between atom **A** and **B** (**A**-H...**B**, in the case of hydrogen bond) for free and protonated **Gly/GG**, can be correlated to the distance between **A** (or **A**-H) and **B**, and the Mulliken population on the atoms. We would like to carry out similar analysis here to understand the relative stabilities of K^+-GG complexes sharing the same mode of binding. In K^+-GG, the non-bonded interactions (mainly hydrogen bonds) can be classified into six types as below:

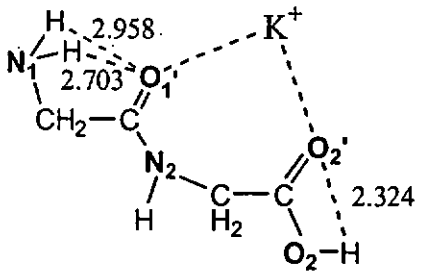
Non-bonded interaction	Atoms involved	Nature	Found in species	Typical distance (Å)
Type 1	$N_2H...N_1$	Hydrogen bond	CS1	2.084
Type 2	$N_2^+H...N_1$	Hydrogen bond	ZW(N_2)	2.009
Type 3	$N_1H_2...O_1'$	Hydrogen bond	CS1a	2.958; 2.703
Type 4	$O_2H...O_2'$	Charge-dipole	CS1	2.319
Type 5	$O_2H...N_2$	Hydrogen bond	CS1b	2.429
Type 6	$O_1'^+H...N_1$	Hydrogen bond	ZW(O_1')	1.768

Under this classification, we found that:

(1) Hydrogen bond type 1 versus type 3: **CS1** is more stable than **CS1a** by ~25 kJ mol^{-1}, implying that hydrogen bond type 1 found in **CS1** is more stabilizing than the hydrogen bond type 3 in **CS1a**. Similarly conclusion can be found when **CS5** is compared against **CS5a**. This suggests that the pair of longer bifurcated hydrogen bond of type 3 is not as strong as one single shorter hydrogen bond of type 1.

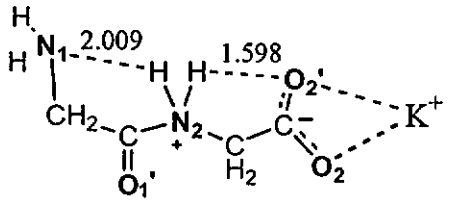


CS1

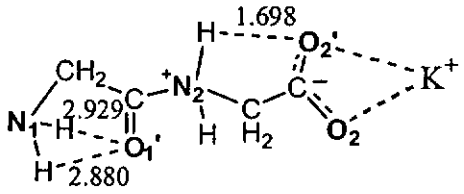


CS1a

(2) Hydrogen bond type 2 versus type 3: The energy difference between **ZW(N_2)** and **ZW(N_2)b** is ~30 kJ mol^{-1}, indicating that hydrogen bond type 2 is more stabilizing than type 3. Again, it appears that a pair of longer bifurcated hydrogen bond (type 3) is not as strong as one single shorter hydrogen bond (type 2).

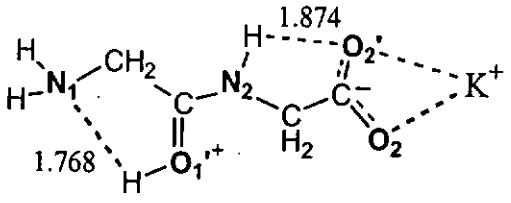


ZW(N_2)

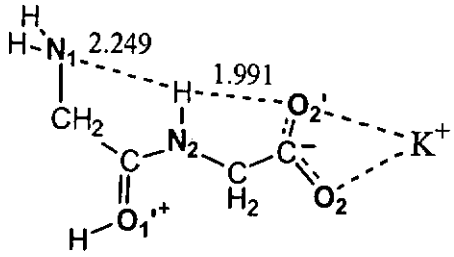


ZW(N_2)b

(3) Hydrogen bond type 6 versus type 1: **ZW(O_1')** is more stable (by ~20 kJ mol^{-1}) than **ZW(O_1')b** because Type 6 in **ZW(O_1')** is much shorter than Type 1 in **ZW(O_1')b**.

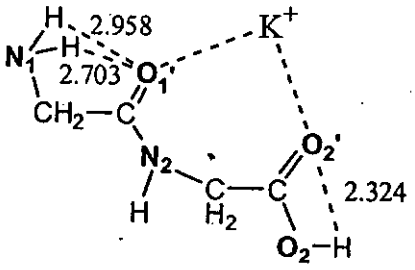


ZW(O_1')

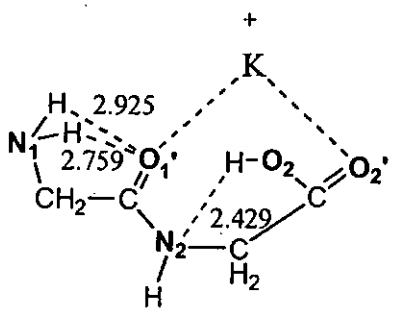


ZW(O_1')b

(4) Charge-dipole interaction type 4 [Zhang et al., 1993] versus hydrogen bond type 5: **CS1a** is more stable than **CS1b** by ~20 kJ mol^{-1}, as type 4 (in **CS1a)** is much shorter than Type 5 (in **CS1b)**. Here we note that the strength of the charge-dipole interaction type 4 would be ~18 kJ mol^{-1}. This value is obtained by comparing **CS2** with **CS2a**, in which the two complexes differs mainly in the presence/absence of the charge-dipole interaction type 4.



CS1a



CS1b

Appendix V: Supporting information for Chapter 7.

Table S-7.1. Protonated histidine (B3-LYP/6-31G(d) geometries).

1		E = -549.3654865 Hartress		1-2		E = -549.3415214 Hartress	
C	1.457019	-0.515407	-1.254689	C	1.478940	-0.556834	-1.148984
C	1.424604	-0.507089	0.299798	C	1.463724	-0.411671	0.384627
N	2.810562	-0.535498	0.769350	N	2.780838	-0.236069	0.980961
H	2.924386	-1.053018	1.636308	H	3.362156	-1.060437	0.859815
H	3.197920	0.397600	0.900905	H	3.274948	0.571468	0.606721
C	0.579674	0.662897	0.884453	C	0.549436	0.738378	0.881267
C	-0.912909	0.532720	0.782901	C	-0.929678	0.546701	0.737633
N	-1.562139	0.269513	-0.411890	N	-1.565967	0.415477	-0.488394
C	-1.897228	0.674473	1.724730	C	-1.926579	0.514807	1.676131
C	-2.877569	0.245100	-0.225946	C	-2.881626	0.302190	-0.322498
N	-3.103602	0.494612	1.071033	N	-3.120812	0.364456	0.992677
H	-0.992465	0.076874	-1.268825	H	-1.017228	0.355970	-1.356795
H	0.963503	-1.459951	0.589641	H	1.071778	-1.366938	0.761766
H	0.826343	0.749711	1.946992	H	0.770526	0.867335	1.943401
H	0.893194	1.602732	0.410270	H	0.838677	1.671824	0.380716
H	-1.846430	0.893555	2.779841	H	-1.888532	0.594492	2.751308
H	-3.628826	0.054826	-0.977522	H	-3.619753	0.178360	-1.100375
H	-4.020948	0.541991	1.499326	H	-4.040647	0.310741	1.414959
O	2.606310	-0.889275	-1.768425	O	2.425777	-1.415957	-1.614163
O	0.495090	-0.227907	-1.955446	O	0.671363	-0.050631	-1.902109
H	3.216849	-1.023417	-0.992901	H	3.287223	-0.998577	-1.777112
2		E = -549.3632358 Hartress		2-3		E = -549.3490397 Hartress	
C	-1.727558	0.978668	0.053154	C	-1.771680	0.764972	0.124332
C	-1.392107	-0.456732	0.510405	C	-1.449789	-0.711392	0.391669
N	-2.514833	-1.363894	0.649024	N	-2.583760	-1.583940	0.145410
H	-3.057228	-1.416527	-0.211567	H	-3.067817	-1.289921	-0.702902
H	-3.152563	-1.046273	1.374980	H	-3.257541	-1.519802	0.905598
C	-0.333871	-1.113775	-0.425716	C	-0.291576	-1.176440	-0.545138
C	1.090268	-0.685198	-0.242501	C	1.054385	-0.587682	-0.268291
N	1.497775	0.639420	-0.289428	N	1.418691	0.718354	-0.586056
C	2.229978	-1.419133	-0.047577	C	2.167396	-1.138505	0.308664
C	2.815910	0.725946	-0.129927	C	2.681273	0.958856	-0.222504
N	3.281154	-0.520888	0.015034	N	3.150235	-0.165028	0.323329
H	0.795050	1.394647	-0.385468	H	0.806134	1.406409	-1.009708
H	-0.943309	-0.341653	1.507244	H	-1.129642	-0.823035	1.431326
H	-0.389946	-2.187224	-0.230327	H	-0.225226	-2.261580	-0.435045
H	-0.624789	-0.966201	-1.474211	H	-0.580500	-0.979687	-1.585568
H	2.380015	-2.483508	0.043718	H	2.337600	-2.130643	0.696755
H	3.403811	1.631067	-0.114645	H	3.221002	1.885328	-0.347306
H	4.256223	-0.761207	0.151431	H	4.090566	-0.280355	0.685002
O	-3.020148	1.269375	0.173320	O	-1.142969	1.605163	0.985807
O	-0.912200	1.795268	-0.359263	O	-2.476966	1.149965	-0.779108
H	-3.152877	2.198610	-0.103041	H	-1.471304	2.507693	0.797087

3			E = -549.3535532 Hartress	3-4			E = -549.2977441 Hartress
C	-1.334659	0.677346	-0.394444	C	2.176071	-1.119003	-0.363647
C	-2.047139	-0.637116	-0.032387	N	3.054219	-0.111363	-0.156606
O	-1.211208	-1.732774	-0.080451	C	2.359026	0.985453	0.31406
H	0.662525	-1.567427	-0.047231	C	1.051882	0.583156	0.38392
H	-1.752896	-2.512948	0.154467	C	-0.20812	1.284004	0.784844
C	-0.192628	1.003344	0.62503	C	-1.318586	1.024427	-0.24545
N	-2.240653	1.806238	-0.460647	N	-2.614076	1.359788	0.224907
H	-2.870095	1.706225	-1.254775	O	-1.49867	-1.407023	1.060974
H	-2.839991	1.824005	0.364382	H	-0.538296	-1.620445	1.030223
C	1.162928	0.463551	0.290322	H	-1.972691	-2.246852	0.91823
N	1.413967	-0.872528	0.004172	H	-2.895061	0.873523	1.070767
C	2.378005	1.090142	0.21146	H	-3.351956	1.37194	-0.471179
C	2.710668	-1.062428	-0.241269	H	-0.556494	0.939438	1.766213
N	3.313444	0.123708	-0.114274	H	-0.071396	2.365963	0.854701
H	-0.887301	0.544473	-1.387315	H	2.847234	1.91893	0.548667
H	-0.119786	2.093521	0.648839	H	4.05282	-0.154503	-0.318032
H	-0.478922	0.683231	1.635586	H	2.460773	-2.09129	-0.739511
H	2.649299	2.123617	0.360238	N	0.950816	-0.732184	-0.037968
H	3.186683	-1.996404	-0.498689	C	-1.306496	-0.442531	-0.764775
H	4.306503	0.282565	-0.242072	O	-1.642385	-1.147589	-1.605109
O	-3.20459	-0.741202	0.269144	H	-1.105858	1.573472	-1.173038
4			E = -549.3341022 Hartress	4-5			E = -549.3017759 Hartress
C	-0.717521	-0.488470	-0.801700	C	1.096417	-0.612097	0.027631
C	-1.325772	0.912500	-0.553738	C	-0.280926	-1.091211	-0.330484
N	-2.639037	0.933644	0.028756	C	-1.349994	-0.582124	0.617503
H	-2.672338	0.304790	0.830401	C	-0.717141	1.325265	0.806030
H	-3.352660	0.656791	-0.639823	O	-0.904988	2.440828	0.968716
C	-0.266993	1.679290	0.300319	N	-2.619482	-0.610210	0.216668
C	0.988130	0.878136	0.132140	H	-2.853008	-0.534450	-0.783351
N	0.707637	-0.343102	-0.457798	H	-3.365741	-0.464044	0.884994
C	2.329510	0.881999	0.383737	N	1.286343	0.696992	0.409620
C	1.812771	-1.086715	-0.563946	C	2.305710	-1.258919	0.020247
N	2.807367	-0.349818	-0.058463	C	2.578559	0.854000	0.638617
H	-1.376263	1.375351	-1.549429	N	3.230914	-0.310638	0.406553
H	-0.156475	2.719158	-0.014782	H	-1.187001	-0.777508	1.675833
H	-0.583983	1.678970	1.348172	H	-0.321340	-2.186563	-0.282989
H	2.981090	1.616998	0.829696	H	-0.557766	-0.808483	-1.354134
H	1.886262	-2.078007	-0.985539	H	2.587828	-2.272054	-0.222554
H	3.775954	-0.649065	-0.012808	H	3.064440	1.762552	0.963393
O	-1.185035	-1.512672	-1.193472	H	4.227328	-0.460614	0.503138
O	-1.357027	-1.110655	1.901056	O	-2.880513	-0.411132	-2.604496
H	-1.772294	-1.965083	1.701990	H	-3.146667	0.397814	-3.069257
H	-1.445313	-1.018790	2.863118	H	-3.244249	-1.136711	-3.135983

5		E = -549.3297848 Hartress		6		E = -435.9664899 Hartress	
C	1.354900	-1.261561	-0.136994	C	0.870464	-0.269300	-0.358039
C	-0.038071	-1.440274	0.400358	C	-0.225088	0.564377	-0.962599
C	-0.523670	-0.099896	0.832800	C	-1.488505	-0.208699	-0.810970
C	-2.409524	3.354224	1.456999	N	-2.546236	0.261485	-0.247079
O	-2.672891	4.398214	1.806563	H	-2.560810	1.217743	0.173767
N	-1.628710	0.423829	0.435582	H	-3.381153	-0.312676	-0.155599
H	-2.235307	-0.067601	-0.254286	N	0.546125	-1.071760	0.717166
H	-1.908853	1.356948	0.757273	C	2.196989	-0.406601	-0.687423
N	1.656457	-0.074631	-0.773611	C	1.665673	-1.683714	1.045485
C	2.448236	-2.092946	-0.111705	N	2.689624	-1.308401	0.227727
C	2.916909	-0.186228	-1.139983	H	-1.529620	-1.226975	-1.191974
N	3.434516	-1.390481	-0.765776	H	-0.054948	0.716964	-2.038637
H	0.086189	0.478693	1.523983	H	-0.328249	1.551397	-0.496104
H	-0.034472	-2.094197	1.284946	H	2.816330	0.038083	-1.451538
H	-0.731681	-1.872575	-0.330796	H	1.790208	-2.395301	1.849639
H	2.612032	-3.082215	0.287991	H	3.648273	-1.626211	0.287682
H	3.494854	0.559713	-1.667405	O	-2.302875	2.795773	0.814266
H	4.374942	-1.717656	-0.944286	H	-2.154768	2.935819	1.763101
O	-3.019486	-1.085957	-1.437489	H	-2.766053	3.589317	0.502114
H	-2.970141	-0.890014	-2.386507				
H	-3.877069	-1.520077	-1.305722				

7		E = -472.8447569 Hartress		8		E = -359.4799957 Hartress	
C	1.637697	0.512639	-0.268740	N	-0.367268	0.064728	-0.933366
C	0.537129	1.336114	-0.876383	C	-0.367908	0.102711	0.446197
C	-0.715287	0.548871	-0.713807	C	0.979378	0.124463	1.111208
C	-4.345133	-0.805328	0.155662	C	1.909208	-0.528910	0.151130
O	-5.264103	-1.451564	0.288485	N	2.995398	0.016691	-0.284698
N	-1.787550	0.998238	-0.158003	H	3.258224	0.965335	-0.022043
H	-1.815460	1.934268	0.243302	H	3.605891	-0.456531	-0.947748
H	-2.633779	0.420042	-0.056932	C	-1.650591	0.099030	0.935822
N	1.281153	-0.357906	0.740834	C	-1.637648	0.046582	-1.284160
C	2.984277	0.437500	-0.527328	N	-2.445097	0.071534	-0.187126
C	2.400691	-0.952396	1.101148	H	1.686259	-1.532170	-0.202761
N	3.453781	-0.498601	0.364971	H	0.986085	-0.468694	2.038231
H	-0.749700	-0.470121	-1.091719	H	1.325511	1.135424	1.363897
H	0.696887	1.489084	-1.954376	H	-2.060398	0.118769	1.934149
H	0.432951	2.327763	-0.416202	H	-2.020315	0.016979	-2.294502
H	3.630724	0.944050	-1.227754	H	-3.457129	0.086965	-0.194080
H	2.504581	-1.702447	1.872436				
H	4.418690	-0.786639	0.465438				

8-9			E = -359.4801001 Hartress	9			E = -359.4929896 Hartress
N	0.172128	-0.913175	0.527862	N	-0.033265	-0.899583	0.300564
C	0.468414	0.295148	-0.045787	C	0.409238	0.323830	-0.169911
C	-0.552186	1.397087	0.025303	C	-0.535351	1.473270	-0.400652
C	-1.853060	0.787220	-0.370987	C	-1.973196	1.241125	-0.101072
N	-2.816180	0.591226	0.435543	N	-2.431133	0.123168	0.337145
H	-2.742119	0.830824	1.406881	H	-1.713619	-0.648716	0.479171
H	-3.660867	0.132882	0.150994	H	-3.421359	-0.007451	0.533837
C	1.703455	0.274630	-0.594192	C	1.767946	0.278042	-0.365892
C	1.219469	-1.645733	0.337122	C	1.033114	-1.671655	0.388479
N	2.177162	-0.978783	-0.333069	N	2.140767	-0.996158	-0.005366
H	-1.992591	0.461943	-1.387005	H	-2.671949	2.061684	-0.261983
H	-0.319631	2.194751	-0.671654	H	-0.501073	1.827497	-1.444492
H	-0.622554	1.816628	1.023118	H	-0.247960	2.362573	0.184833
H	2.275736	1.010129	-1.117237	H	2.480385	1.009237	-0.716653
H	1.339825	-2.659447	0.657433	H	1.047870	-2.698598	0.723421
H	3.073873	-1.332700	-0.583641	H	3.082618	-1.365890	-0.030255
9-10			E = -359.4890866 Hartress	10			E = -359.4956589 Hartress
N	-0.587507	-0.665426	0.000000	N	-0.870887	-0.298880	0.000000
C	0.000000	0.588453	0.000000	C	0.512641	-0.245821	0.000000
C	1.501083	0.716854	0.000000	C	1.275493	1.050578	0.000000
C	2.246351	-0.588579	0.000000	C	0.456979	2.323127	0.000000
N	1.630005	-1.709522	0.000000	N	-0.816244	2.327492	0.000000
H	0.370353	-1.532289	0.000000	H	-1.393571	0.603097	0.000000
H	2.181526	-2.566817	0.000000	H	-1.210782	3.269973	0.000000
C	-0.997742	1.528121	0.000000	C	0.941152	-1.547529	0.000000
C	-1.900616	-0.504540	0.000000	C	-1.285817	-1.562019	0.000000
N	-2.179398	0.815398	0.000000	N	-0.193547	-2.337680	0.000000
H	3.337165	-0.541547	0.000000	H	1.058964	3.238457	0.000000
H	1.852254	1.290347	0.870870	H	1.944608	1.084080	-0.871649
H	1.852254	1.290347	-0.870870	H	1.944608	1.084080	0.871649
H	-0.978153	2.607055	0.000000	H	1.933030	-1.971975	0.000000
H	-2.642419	-1.288997	0.000000	H	-2.309718	-1.903179	0.000000
H	-3.109136	1.216892	0.000000	H	-0.205081	-3.351076	0.000000
7-11			E = -472.8205131 Hartress	11			E = -472.8571563 Hartress
C	0.755562	-0.664074	-0.142788	C	-1.035198	-0.419299	-0.251384
C	-0.566236	-1.340037	-0.370338	C	-1.000374	-0.537355	1.243453
C	-1.700351	-0.726877	0.416864	C	0.504875	-0.324088	1.603917
C	-1.274212	1.296491	0.032471	C	1.143235	0.326751	0.351168
O	-1.514051	2.401855	-0.112112	O	2.170569	0.904788	0.182280
N	-2.953382	-0.993217	0.071939	N	0.703203	0.360389	2.851190
H	-3.191123	-1.261060	-0.877091	H	0.472545	1.349960	2.799848
H	-3.729826	-0.749570	0.675907	H	1.657874	0.273557	3.188543
N	0.842200	0.709100	-0.153469	N	0.160339	0.101504	-0.725055
C	2.003885	-1.200156	0.048837	C	-1.832770	-0.599695	-1.343602
C	2.113593	1.014219	0.032681	C	0.119973	0.253345	-2.054080
N	2.850774	-0.117826	0.153969	N	-1.086239	-0.173027	-2.439947
H	-1.535635	-0.558413	1.477924	H	0.986422	-1.309160	1.678057
H	-0.508336	-2.388646	-0.049208	H	-1.376030	-1.496814	1.605402
H	-0.834188	-1.351985	-1.435953	H	-1.607409	0.247514	1.709461
H	2.359081	-2.217781	0.109094	H	-2.838863	-0.973987	-1.452079
H	2.529796	2.009879	0.085408	H	0.909461	0.634949	-2.684716
H	3.852045	-0.161058	0.297386	H	-1.408113	-0.184340	-3.402903

TSa		E= -549.2833225 Hartress		TSb		E = -549.2823548 Hartress	
C	2.306321	-0.636388	-0.233686	C	0.975045	0.006209	0.304848
C	0.991299	-0.036882	0.157666	C	2.409665	-0.741289	-0.177176
N	0.974866	1.071354	-0.753806	N	0.923131	1.131623	-0.482075
H	4.466162	0.813024	0.705716	H	4.263102	1.441321	0.289656
H	0.048332	1.190691	-1.180025	H	0.126929	1.194463	-1.121322
H	1.348314	1.925953	-0.356656	H	1.269800	1.996638	-0.088815
C	-0.120517	-1.099385	-0.141217	C	-0.101383	-1.067513	0.012737
C	-1.472355	-0.475962	0.016285	C	-1.472567	-0.461649	0.068643
N	-1.851615	0.568446	-0.806636	N	-1.829106	0.503245	-0.854025
C	-2.496433	-0.804108	0.872716	C	-2.530632	-0.742651	0.898658
C	-3.085240	0.865444	-0.457216	C	-3.083883	0.800225	-0.589481
N	-3.518195	0.058668	0.552246	N	-3.552094	0.069121	0.460850
H	1.029439	0.189280	1.228879	H	1.154768	0.176273	1.365158
H	0.007703	-1.457536	-1.169374	H	0.063198	-1.478947	-0.990488
H	0.005629	-1.948404	0.536366	H	0.006688	-1.884818	0.731320
H	-2.595588	-1.557036	1.640086	H	-2.654110	-1.433788	1.718761
H	-3.700946	1.638221	-0.895725	H	-3.690813	1.522134	-1.117820
H	-4.437181	0.073665	0.975027	H	-4.488984	0.102492	0.840729
O	3.580105	0.957845	1.076958	O	3.605841	1.048996	0.888626
H	3.716443	1.679396	1.712634	H	4.102198	0.854160	1.701086
O	3.187126	-1.121951	-0.755314	O	3.128689	-1.196227	-0.927488
1-12		E = -549.2898215 Hartress		12		E = -549.296530 Hartress	
C	-2.338600	1.245865	-0.161796	C	-2.325815	1.265031	-0.126181
N	-1.611525	-0.829138	-0.125195	N	-1.712092	-0.845777	-0.168282
C	0.239691	0.876191	-0.447398	C	0.235852	0.745349	-0.398012
C	1.215343	0.184230	0.484492	C	1.088985	-0.080467	0.542320
C	1.993742	-0.974503	0.052061	C	2.074000	-1.023080	0.058724
C	-2.934037	-0.892188	0.005557	C	-3.038955	-0.842992	-0.040263
N	-3.395624	0.363478	-0.020238	N	-3.429002	0.435004	-0.017964
C	-1.198362	0.492179	-0.227316	C	-1.226178	0.457500	-0.222232
O	0.877083	-1.587337	0.369352	O	0.892628	-1.616713	0.352701
O	3.075328	-1.269895	-0.372852	O	3.143105	-1.327374	-0.358556
N	2.980410	1.753834	0.325243	N	3.122465	2.068343	0.275292
H	-0.931988	-1.596319	-0.075154	H	-1.101232	-1.662876	-0.165221
H	-3.525800	-1.787962	0.119564	H	-3.679214	-1.708934	0.036044
H	-4.373388	0.620331	0.054982	H	-4.391808	0.741603	0.066098
H	-2.487331	2.313320	-0.209544	H	-2.416113	2.340035	-0.129695
H	0.333085	1.954589	-0.302232	H	0.451212	1.798735	-0.203619
H	0.516692	0.666298	-1.487678	H	0.524709	0.548583	-1.437751
H	1.166934	0.402159	1.545009	H	1.086019	0.194689	1.593868
H	3.027823	2.548700	-0.312448	H	3.279757	2.831835	-0.383275
H	3.337976	2.077201	1.224089	H	3.476930	2.409148	1.169444
H	3.637224	1.051676	-0.020853	H	3.766949	1.328831	-0.008504

13		E = -549.3152857 Hartress	
C	-1.505387	0.969738	-0.190618
C	1.105411	1.110206	-0.471183
N	-0.374901	-0.910617	-0.140416
N	-2.373208	-0.092341	-0.027092
C	2.130087	0.612819	0.530931
C	3.369713	-0.028541	0.114024
C	-1.669226	-1.221018	0.005366
H	0.409932	-1.557028	-0.108561
H	-2.066369	-2.216912	0.130901
C	-0.234543	0.466523	-0.263519
H	-3.447811	-0.025588	0.059383
H	-1.855471	1.988872	-0.246776
O	2.384164	-0.900375	0.436675
O	4.487734	-0.054189	-0.276770
H	2.013942	0.936444	1.564055
H	1.005345	2.196030	-0.371206
H	1.457783	0.916050	-1.491781
N	-5.125667	0.087082	0.191763
H	-5.617670	-0.498251	-0.484840
H	-5.463954	1.038067	0.037488
H	-5.470813	-0.188399	1.112311

13-14		E = -549.2518598 Hartress	
C	2.280423	-0.344787	0.107837
N	0.125349	0.066041	0.035744
N	2.070743	1.017044	0.105748
C	0.703355	-2.410138	0.014950
C	-0.559927	-3.017684	-0.211782
C	0.758068	1.243798	0.062396
H	2.837587	1.777960	0.131426
C	1.056212	-0.961554	0.062225
H	-0.887150	-0.046148	-0.012995
H	0.238288	-2.888530	1.017858
H	0.280014	2.211608	0.049081
O	-2.847845	0.134678	-0.084309
C	-4.012180	-0.000705	-0.183010
O	-5.160948	-0.129404	-0.279057
H	3.271359	-0.771572	0.137534
H	1.592104	-3.035861	-0.118448
H	-1.341210	-2.234904	-0.114315
N	4.049155	2.949873	0.168694
H	4.127029	3.411072	1.076237
H	3.908769	3.692240	-0.518067
H	4.971112	2.557658	-0.028392

14		E = -549.3886522 Hartress	
C	-2.101250	0.978963	0.030524
N	-0.149856	-0.026733	-0.000703
C	-0.038927	2.497672	0.002407
C	1.289711	2.655984	-0.018357
C	-1.092665	-0.977879	0.011579
N	-2.285028	-0.387103	0.030575
H	0.849726	-0.213927	-0.015899
C	-0.748150	1.229222	0.010686
H	-0.909198	-2.041312	0.006919
H	1.727337	3.648168	-0.022993
H	-3.232708	-0.893564	0.044014
H	-0.687157	3.369896	0.014687
H	-2.931604	1.667521	0.044470
H	1.989488	1.824359	-0.031534
O	4.979734	-1.500316	-0.060670
O	2.742403	-0.828407	-0.044772
C	3.868569	-1.166209	-0.052727
N	-4.754027	-1.662390	0.067420
H	-5.232209	-1.548070	0.962100
H	-4.707281	-2.668109	-0.101251
H	-5.384839	-1.288111	-0.642688

15		E = -360.721286 Hartress	
C	0.710691	0.557707	-0.026274
N	-1.238673	-0.444347	0.030373
C	-1.354878	2.077654	-0.041681
C	-2.684132	2.231400	-0.027635
C	-0.297131	-1.399581	0.045148
N	0.893217	-0.808400	0.011169
H	-2.235556	-0.619964	0.049479
C	-0.641538	0.811966	-0.014744
H	-0.481895	-2.462438	0.078728
H	-3.127532	3.220677	-0.050237
H	1.845022	-1.317490	0.012776
H	-0.710245	2.951749	-0.076100
H	1.541911	1.244682	-0.058091
H	-3.385982	1.400913	0.006300
N	3.351343	-2.080440	0.013691
H	3.866462	-1.921005	-0.853286
H	3.296225	-3.093640	0.126589
H	3.952307	-1.746043	0.768335

16			E = -492.7734797 Hartress	17			E = -304.1049043 Hartress
C	2.761421	0.748323	-0.034519	C	2.183494	1.583298	0.000000
N	0.805262	-0.259384	0.000459	H	2.733237	2.518011	0.000000
C	0.703175	2.266848	-0.051060	H	2.790470	0.680467	0.000000
C	-0.625748	2.423427	-0.044407	C	0.845348	1.579658	0.000000
C	1.731338	-1.219844	0.013661	C	0.000000	0.397649	0.000000
N	2.926006	-0.623095	-0.007315	N	0.460267	-0.918464	0.000000
H	-0.197402	-0.442146	0.011435	C	-1.371217	0.296388	0.000000
C	1.408835	0.996453	-0.029886	C	-0.564573	-1.776049	0.000000
H	1.548271	-2.283287	0.036994	N	-1.683838	-1.049747	0.000000
H	-1.063160	3.415583	-0.061781	H	1.434652	-1.197382	0.000000
H	3.814918	-1.110865	-0.003626	H	0.299341	2.518890	0.000000
H	1.353130	3.137273	-0.073990	H	-2.130857	1.062246	0.000000
H	3.599320	1.427221	-0.054609	H	-0.498821	-2.853633	0.000000
H	-1.325997	1.592530	-0.021881	H	-2.621339	-1.436781	0.000000
O	-4.280759	-1.752993	0.062718				
O	-2.053098	-1.050201	0.036548				
C	-3.175203	-1.404109	0.049765				

Appendix VI: Supporting information for Chapter 7.

Loss of 46 u Other Than the Loss of [CO + H_2O]

Our theoretical findings is similar to previous findings on $[Gly + H]^+$ on the fragmentation pathways for the loss of 46 u for $[His + H]^+$. There are four different neutral species that may correspond to 46 u; they are: (i) formic acid (HCOOH) from the amino N^{α}-protonated histidine, (ii) carbon dioxide and hydrogen from the amino N^{α}-protonated histidine, (iii) dihydroxycarbene ($C(OH)_2$) from the O_HC-protonated isomer, (iv) water and carbon monoxide [CO + H_2O] from the O_HH-protonated form. [Rogalwicz and Hoppilliard, 2000b] We have calculated the complete pathway of each of the fragmentation product due to the loss of 46 u. In order to simplify the problem, we firstly reported the energy of the neutral species (final product after fragmentation) and their highest energy barrier to the formation of various 46 u fragments as summarized in Table S-7.2. In fact, the mechanism for the loss of other 46 u we reported here for $[His + H]^+$ is identical to $[Gly + H]^+$ [Rogalwicz and Hoppilliard, 2000b], except the mechanism postulated for the loss of $C(OH)_2$. The complete mechanism of $[Gly + H]^+$ has already been discussed by Rogalwicz and Hoppilliard, thus, in the following discussion, we will only highlight some possible differences of the fragmentation mechanism in $[His + H]^+$; the analogous structures and energetics are summarized in Appendix VII, Figs. S-7.1 and S-7.2.

Table S-7.2. The relative enthalpies at 0K (ΔH_0) of the final state and their barrier to formation of various 46 u fragments of $[His + H]^+$, in kJ mol^{-1}.

	Final State			Barrier		
Species	ΔH_0	ΔG_{298}	ΔG_{600}	ΔH_0	ΔG_{298}	ΔG_{600}
HCOOH	62	9	-47	387	382	375
$CO_2 + H_2$	32	-42	-129	318	315	309
$C(OH)_2$	240	187	131	219	213	202
$CO + H_2O$	98	11	-88	165	164	156

According to Table S-7.2, our results show that among the four pathways for the loss of 46 u, the loss of $[CO + H_2O]$ requires the least energy barrier; that is 165 kJ mol^{-1} in energy. It seems that the loss of HCOOH or $[CO_2 + H_2]$ is not likely since their energy barrier is substantially higher, at 387 or 318 kJ mol^{-1} less stable than species **1**, respectively, and similar observation is found at 600K. However, the most kinetically favorable loss of $[CO + H_2O]$ is not so energetically favorable among 46 u neutrals, which is 98 kJ mol^{-1} higher in energy relative to the species **1**; that is at least 36 kJ mol^{-1} less stable than the neutral species of $[CO_2 + H_2]$ and HCOOH. Though the energy barrier required for loss of $C(OH)_2$ is quite competitive to the loss of $[CO + H_2O]$, the loss of $C(OH)_2$ is two times less stable than that of $[CO + H_2O]$. In other words, the loss of $[CO + H_2O]$ is very likely because its mechanism among all loss of 46 u pathways is demonstrated to be energetically and kinetically very favorable, and is identical to the simplest amino acids, $[Gly + H]^+$. [Rogalwicz and Hoppilliard, 2000b]

For the loss of other 46 u from $[His + H]^+$, the process is assumed to arise from the most stable $[His + H]^+$ complex, species **1**, in which the protonation site is found at the imidazole N^{π}, different from that of $[Gly + H]^+$ at the amino N^{α} nitrogen [Rogalwicz and Hoppilliard, 2000b]. This means that there would be a variation in the early part of the fragmentation pathway of $[His + H]^+$. However, this is not likely to play an important role in the fragmentation as the barrier of conversion from protonation at N^{π} to N^{α} is small (within 40 kJ mol^{-1} relative to species **1** (Table 7.2)). Since the loss of 46 u has been proven to be an imine rather than an immonium structure found in $[Gly + H]^+$ [Rogalwicz and Hoppilliard, 2000b] in Section 7.3.3, the final products for the loss of HCOOH, $[CO_2 + H_2]$ and $C(OH)_2$ that discussed here are also imine (Species **10** in Fig. 7.5).

Loss of HCOOH: (Appendix VII, Fig. S-7.1) Similar to Hoppilliard et al., [Rogalwicz and Hoppilliard, 2000b] the key step for the loss of HCOOH is the proton migration of one hydrogen from the protonated amino N^{α} to the carboxyl carbon C_t and the estimated barrier is 387 kJ mol^{-1} higher in energy than **1** (the analogous species in $[Gly + H]^+$ is 'TS 1-4' in ref. [Rogalwicz and Hoppilliard, 2000b]). The intermediate **A** (analogous species '4' in ref. [Rogalwicz and Hoppilliard, 2000b]) is a very stable ion that is an immonium ion weakly coordinated to the HCOOH, and its energy is 34 kJ mol^{-1} above **1**. However, in $[Gly + H]^+$, such species is much higher in energy than that found in $[His + H]^+$, (38 kJ mol^{-1} differs in energy), suggesting that the imidazole ring of histidine plays an important role for stabilizing the whole species as it could form an extraordinary stable hydrogen bond between imidazole N^{π} and one hydrogen of amino N^{α} (~1.7Å), which is not present in aliphatic amino acids.

Loss of [CO_2 + H_2]: (Appendix VII, Fig. S-7.1) Similar to Rogalwicz and Hoppilliard et al., [Rogalwicz and Hoppilliard, 2000b] there are two key steps for the loss of [CO_2 + H_2]. The first key step is the proton transfer from hydroxyl O_H to the C^{α} atom and facilitate the formation of a four center transition state with a relatively high energy barrier at 318 kJ mol^{-1} (the analogous species 'TS 1-5' in ref. [Rogalwicz and Hoppilliard, 2000b]). This is followed by the stretching of the C-C bond facilitate the loss of CO_2 in the second step. The product for the loss of CO_2 is –48 kJ mol^{-1} more stable. The next key step is the loss of H_2 (or the loss of NH_3 as discussed in Section 7.3.3) involving a four center elimination of H_2 at the protonated amino N^{α} hydrogen to the leaving H from the C^{α} atom (analogous species 'TSB_1-B_2' in ref. [Rogalwicz and Hoppilliard, 2000b]. This requires a greater amount of energy; the transition state lies 370 kJ mol^{-1} higher in energy than the product for the loss of CO_2, i.e., 304 kJ mol^{-1} above **1**. In other words, at 0 K, the loss of CO_2 should be followed by the loss of NH_3 rather than the loss of H_2 since the ions may not have enough internal energy to overcome such a high energy barrier.

Loss of $C(OH)_2$: (Appendix VII, Fig. S-7.2) For this dissociation reaction, the key step is not similar to that found by Rogalwicz and Hoppilliard for the $[Gly + H]^+$. [Rogalwicz and Hoppilliard, 2000b] Thus, we postulated the dissociation reaction in Appendix VII, Fig. S-7.2. The key steps for such dissociation reaction of $[His + H]^+$ is the hydrogen migration from imidazole N^{π} to carbonyl O_c via transition state **B** located 134 kJ mol^{-1} above **1** to yield the *cyclized* species **C** (Appendix VII, Fig. S-7.2). The cyclized species **C** is particularly stable when compared to the analogous non-cyclic species, '2' in ref. [Rogalwicz and Hoppilliard, 2000b]. Further elongation of the C-C bond to 2.95Å via a transition state **D** leads to the loss of $C(OH)_2$ (with 219.0 kJ mol^{-1} above **1**). The product (immonium ion, **8**) for the loss of $C(OH)_2$

required 63 kJ mol^{-1} higher in energy than the transition state. In other words, the loss of $C(OH)_2$ has no barrier for the reverse reaction. This is contrary to the loss of other neutrals of 46 u; they all have a substantial energy barrier for the reverse reaction. This is inline with Rogalwicz and Hoppiliard findings on $[Gly + H]^+$. [Rogalwicz and Hoppilliard, 2000b]

Appendix VII: Supporting information for Chapter 7.

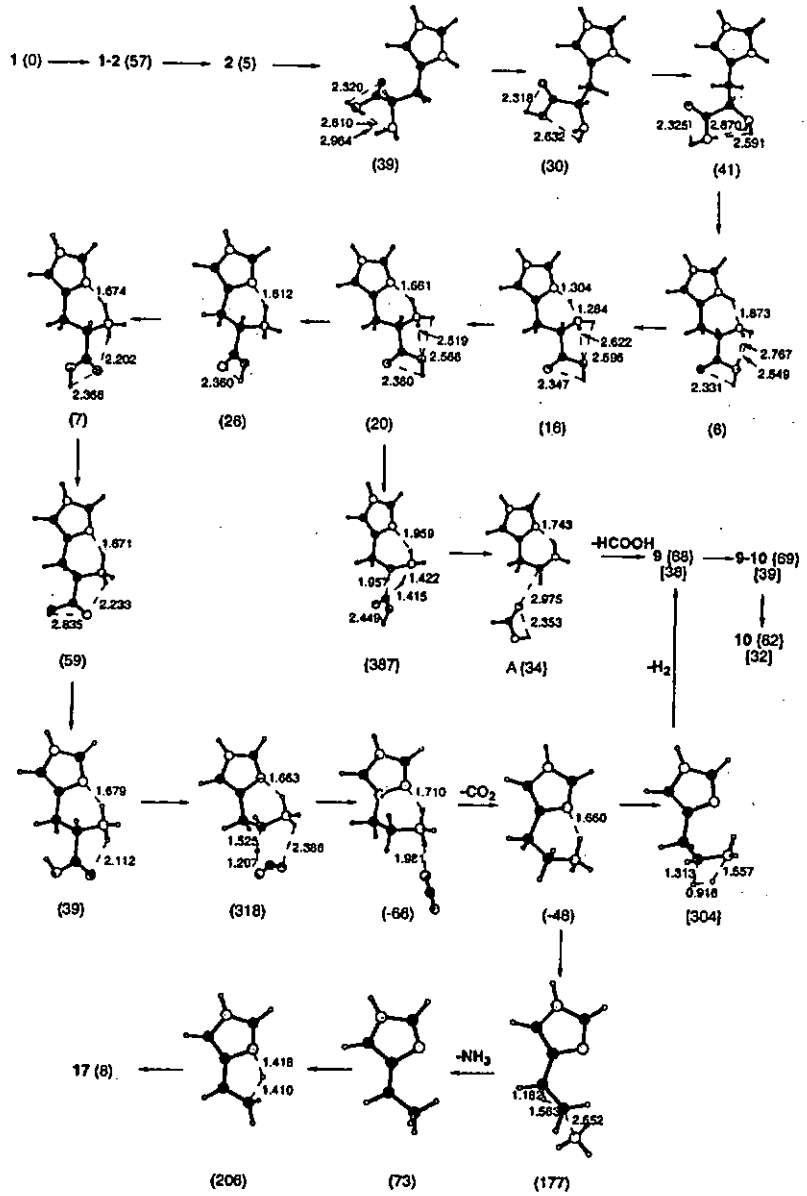


Figure S-7.1. The loss of $[CO_2 + NH_3]$ (61 u), $[CO_2 + H_2]$ (46 u) and HCOOH (46u) from $[His + H]^+$, analogous to that for protonated glycine (i.e. initiated by proton transfer from hydroxyl O_H to carbon C^{α}) for the loss of $[CO_2 + NH_3]$ and $[CO_2 + H_2]$; proton transfer from amino N^{α} to carboxyl C_1 for the loss of HCOOH), postulated by [Rogalewicz and Hoppilliard, 2000b]. The relative enthalpies at 0K (ΔH_0), with respect to species **1**, are shown in parentheses, square bracket and large bracket for loss of $[CO_2 + NH_3]$, $[CO_2 + H_2]$, and HCOOH, respectively, in kJ mol^{-1}.

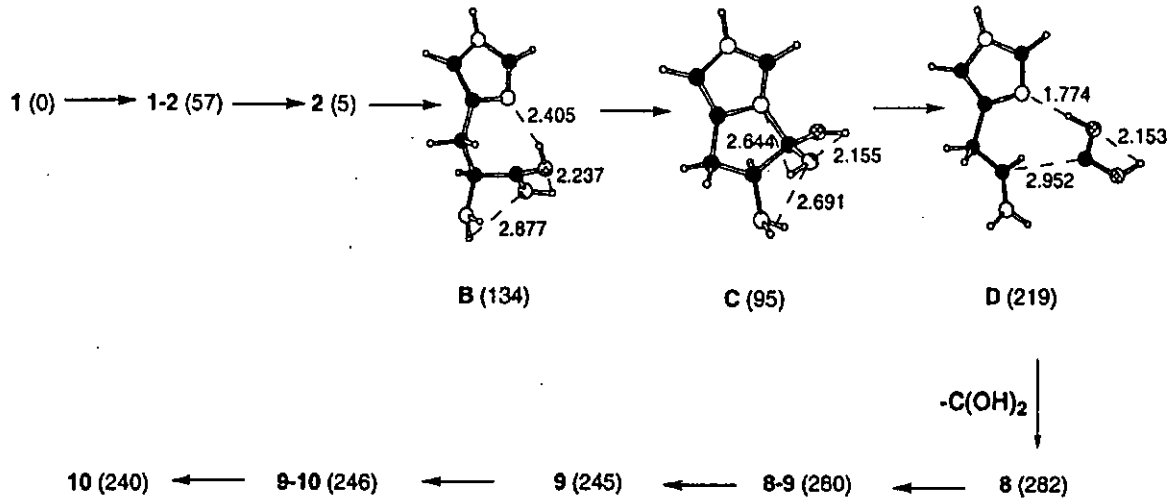


Figure S-7.2. The loss of $C(OH)_2$ (46 u) from $[His + H]^+$, initiated by proton transfer from imidazole N^{π} to carboxyl O_C). The relative enthalpies at 0K (ΔH_0), with respect to species **1**, are shown in parentheses (in kJ mol^{-1}).

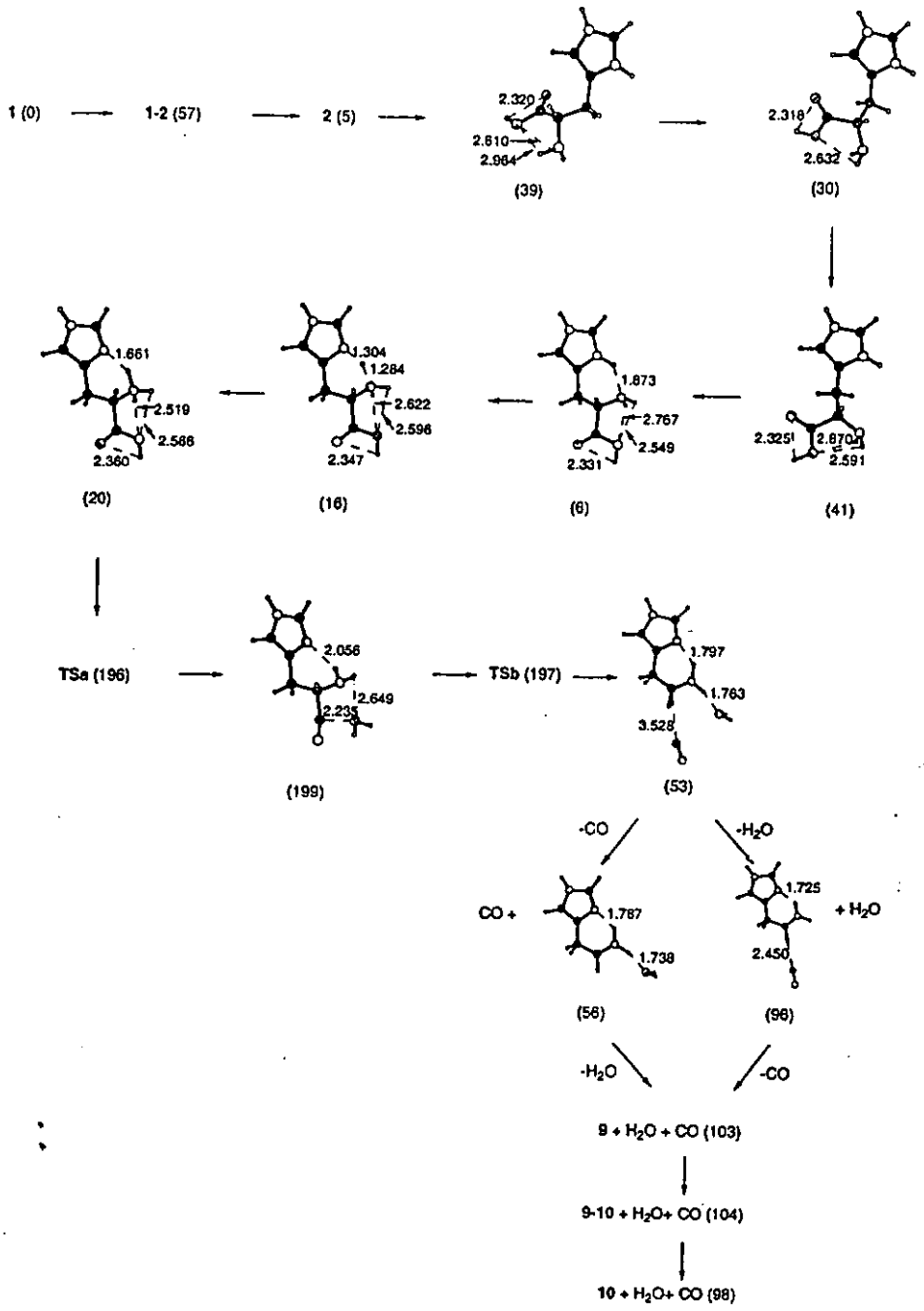


Figure S-7.3. The loss of [CO + H_2O] from [His + H]$^+$, analogous to that for protonated aliphatic amino acids (i.e. initiated by proton transfer from amino N^{α} to hydroxyl O_H), postulated by [Rogalewicz and Hoppilliard, 2000b]. The relative enthalpies at 0K (ΔH_0), with respect to species **1**, are shown in parentheses (kJ mol^{-1}).

References

Abboud, J.-L.M., Alkorta, I., Davalos, J.Z., Gal, J.-F., Herreros, M., Maria, P.-C., Mo, O., Molina, M.T., Notario, R. and Yanez, M. "The P-4...Li^+ ion in the gas phase: A planetary system". *Journal of the American Chemical Society*, Vol. 122, pp. 4451-4454 (2000)

Abirami, S., Ma, N.L. and Goh, N.K. "A Gaussian-3 (G3) theoretical study of the interactions between alkali metal cations and polyhydroxyl ligands". *Chemical Physics Letters*, Vol. 359, pp. 500-506 (2002)

Abirami, S., Wong, C.H.S., Ma, N.L. and Tsang, C.W. "K^+-glycylproline (GP) and K^+-prolylglycine (PG)". *Manuscript in preparation* (2004b)

Abirami, S., Wong, C.L., Ma, N.L. and Tsang, C.W. "Fragmentation of prototypical cationized aliphatic amino acids: protonated and alkali cationized alanine". *Unpublished work* (2004)

Adamo, C. and Barone, V. "Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models". *Journal of Chemical Physics,* Vol. 108, pp. 664 (1998)

Addario, V., Guo, Y., Chu, I.K., Ling, Y., Ruggerio, G., Rodriquez, C.F., Hopkinson, A.C. and Siu, K.W.M. "Proton affinities of methyl esters of N-acetylated amino acids". *International Journal of Mass Spectrometry*, Vol. 219, pp. 101-114 (2002)

Alcami, M., Mo, O., Yanez, M., Anvia, F. and Taft, R.W. "Experimental and theoretical study of Li^+ affinities of methyldiazoles". *Journal of Physical Chemistry,* Vol. 94, pp. 4796-4804 (1990)

Alcami, M., Mo, O. and Yanez, M., "Computational chemistry: A useful (sometimes mandatory) tool in mass spectrometry studies". *Mass Spectrometry Reviews*, Vol. 20, pp. 195-245 (2001)

Armentrout, P.B. and Rodgers, M.T. "An absolute sodium scales: Threshold collision-induced dissociation experiments and ab initio theory". *Journal of Physical Chemistry A*, Vol. 104, pp. 2238-2247 (2000)

Amicangelo, J.C. and Armentrout, P.B. "Absolute binding energies of alkali-metal cation complexes with benzene determined by threshold collision-induced dissociation experiments and ab initio theory". *Journal of Physical Chemistry A*, Vol. 104, pp. 11420-11432 (2000)

Amunugama, R. and Rodgers, M.T. "Absolute alkali metal ion binding affinities of several azines determined by threshold collision-induced dissociation and *ab initio* theory". *International Journal of Mass Spectrometry*, Vol. 195/196, pp. 439-457 (2000)

Amunugama, R. and Rodgers, M.T. "The influence of substituents on cation-pi interactions. 4. Absolute binding energies of alkali metal cation - Phenol complexes determined by threshold collision-induced dissociation and theoretical studies". *Journal of Physical Chemistry A*, Vol. 106, pp. 9718-9728 (2002)

Anonymous, "Nomenclature of Alpha-amino-acids – Recommendations, 1974 - IUPAC Commission on nomenclature of organic – chemistry (CNOC) and IUPAC-IUB Commission on biochemical nomenclature (CBN)". *European Journal of Biochemistry*, Vol. 53, pp.1-14 (1975)

Aribi, H.E., Rodriquez, C.F., Almeida, D.R.P., Ling, Y., Mak, W.W.-N., Hopkinson, A.C. and Siu K.W.M., "Elucidation of fragmentation mechanisms of protonated peptide ions and their products: a case study on glycylglycylglycine using density functional theory and threshold collision-induced dissociation". *Journal of the American Chemical Society*, Vol. 125, pp. 9229-9236 (2003)

Aribi, H.E., Orlova, G., Hopkinson, A.C. and Siu, K.W.M. "Gas-phase fragmentation reactions of protonated aromataic amino acids: Concomitant and consecutive neutral eliminations and radical cation formations". *Journal of Physical Chemistry A*, Vol. 108, pp. 3844-3853 (2004)

Baerends, E.J. and Gritsenko, O.V., "A quantum chemical view of density functional theory". *Journal of Physical Chemistry*, Vol. 101, pp. 5383–5403 (1997)

Balta, B., Basma, M., Aviyente, V., Zhu, C.B. and Lifshitz, C. "Structures and reactivity of gaseous glycine and its derivatives". *International Journal of Mass Spectrometry*, Vol. 201, pp. 69-85 (2000)

Barrow, G., eds., *Physical Chemistry*, 6th edition, McGraw-Hill Companies, Inc., Chapter 14 (1996).

Bartolotti, L.J. and Flurchick, K. *Rev. Comput. Chem.*, Vol. 7, pp. 197 (1996)

Becke, A.D. "Density-functional exchange-energy approximation with correct asymptotic-behaviour". *Physical Review A,* Vol. 38, pp. 3098-3100 (1988).

Becke, A.D. "Density – functional thermochemistry .3. The role of exact exchange". *Journal of Chemical Physics*, Vol. 98, pp. 5648-5652 (1993)

Blaudeau, J.–P., McGrath, M.P., Curtiss, L.A. and Radom, L. "Extension of Gaussian-2 (G2) theory to molecules containing third-row atoms K and Ca". *Journal of Chemical Physics*, Vol., 107, pp. 5016-5021 (1997)

Bliznyuk, A.A., Scheafer III, H.F. and Amster, I.J. "Proton affinities of lysine and histidine – a theoretical consideration of the discrepancy between experimental results from the kinetic and bracketing methods". *Journal of the American Chemical Society*, Vol. 115, pp. 5149-5154 (1993)

Bojesen, G. and Breindahl, T. "On the proton affinity of some alpha-amino acids and the theory of the kinetic method". *Journal of the Chemical Society - Perkin Transactions*, Vol. 2, pp. 1029-1037 (1994)

Bracci, L. Pini, A., Bernini, A. Lelli, B., Ricci, C., Scarselli, M. Niccolai, N. and Neri, P. " Biochemical filtering of a protein-protein docking simulation identifies the structure of a complex between a recombinant antibody fragment and alpha-bungarotoxin". *Biochemical Journal*, Vol. 371, pp. 423-427 (2003)

Bruyneel, C., Pham-Tran, N.N., Nguyen, M.T. and Zeegers-Huyskens, T. "Theoretical and experimental study of the conformation and vibrational frequencies of N-acetyl-L-alanine and N-acetyl-L-alaninate". *Spectroscopy Letters*, Vol. 36, pp. 537-550 (2003)

Burda, J.V., Sponer, J. and Hobza, P. "*Ab Initio* study of the interaction of guanine and adenine with various mono- and bivalent metal cations (Li^+, Na^+, K^+, Rb^+, Cs^+; Cu^+, Ag^+, Au^+; Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}; Zn^{2+}, Cd^{2+}, and Hg^{2+})". *Journal of Physical Chemistry*, Vol. 100, pp. 7250-7255 (1996)

Burk, P., Koppel, I.A., Koppel, I., Kurg, R., Gal, J.-F., Maria, P.-C., Herreros, M., Notario, R., Abboud, J.-L.M., Anvia, F. and Taft, R.W. "Revised and expanded scale of gas-phase lithium cation basicities. An experimental and theoretical study". *Journal of Physical Chemistry A*, Vol. 104, pp. 2824-2833 (2000)

Caldwell, W. and Kollman, P.A. "Cation-π interactions: nonadditive effects are critical in their accurate representation". *Journal of the American Chemical Society*, Vol. 117, pp.4177-4178 (1995)

Caraiman, D., Koyanagi, G.K. and Bohme, D.K. "Gas-phase reactions of transition-metal ions with hexafluorobenzene: Room-temperature kinetics and periodicities in reactivity". *Journal of Physical Chemistry A*, Vol. 108, pp. 978-986 (2004)

Caraiman, D., Shoeib, T. Siu, K.W.M., Hopkinson, A.C. and Bohme, D.K. " Investigations of the gas-phase reactivity of Cu^+ and Ag^+ glycine complexes towards CO, D_2O and NH_3". *International Journal of Mass Spectrometry*, Vol. 228, pp. 629-646 (2003)

Carr, S.R. and Cassady, C.J. "Gas-phase basicities of histidine and lysine and their selected di- and tripeptides". *Journal of the American Society for Mass Spectrometry*, Vol. 7, pp. 1203-1210 (1996)

Cartledge, J.D., Denning, P.W., Dupont, B., Clumeck, N., Dewit, S., Midgley, J., Hawkins, D.A. and Gazzard, B.G. "Treatment of HIV-related fluconazole-resistant oral candidosis with D0870, a new triazole antifungal". *AIDS*, Vol. 12, pp. 411-416 (1998)

Cartledge, J.D., Midgley, J., Petrout, M., Shanson, D. and Gazzard, B.G. "Unresponsive HIV-related oro-oesophageal candidosis - an evaluation of two new in-vitro azole susceptibility tests". *Journal of Antimicrobial Chemotherapy*, Vol. 40, pp. 517-523 (1997)

Cassady, C.J., Carr, S.R., Zhang, K. and Chung-Phillips, A. "Experimental and ab-Initio studies on protonations of alanine and small peptides of alanine and glycine". *Journal of Organic Chemistry*, Vol. 60, pp. 1704-1712 (1995)

Castleman, A.W. "Properties of clusters in gas-phase – ammonia about Bi^+, Rb^+, and K^+". *Chemical Physics Letters*, Vol. 53, pp. 560-564 (1978a)

Cerda, B. A., Hoyau, S., Ohanessian, G. and Wesdemiotis, C., "Na^+ binding to cyclic and linear dipeptides. bond energies, entropies of Na^+ complexation, and attachment sites from the dissociation of Na^+-bound heterodimers and ab Initio calculations". *Journal of the American Chemical Society*, Vol. 120, pp. 2437-2448 (1998)

Cerda, B.A. and Wesdemiotis, C. "Li^+, Na^+, and K^+ binding to the DNA and RNA nucleobases. Bond energies and attachment sites from the dissociation of metal ion-bound heterodimers". *Journal of the American Chemical Society*, Vol. 118, pp. 11884-11892 (1996)

Chapo, C.J., Paul, J.B., Provencal, R.A., Roth, K. and Saykally, R.J., "Is arginine zwitterionic or neutral in the gas phase? Results from IR cavity ringdown spectroscopy". *Journal of the American Chemical Society*, Vol. 120, pp.12956-12957 (1998)

Chen, R., Cheng, Y.H., Liu, L., Li, X.S. and Guo, Q.X., "pi-type and sigma-type cation-pi complexes of atomic cations". *Research on Chemical Intermediates*, Vol. 28, pp.41-48 (2002)

Chen, X.Z., Steel, A. and Hediger, M.A., "Functional roles of histidine and tyrosine residues in H^+-peptide transporter pepT1". *Biochemical and Biophysical Research Communications*, Vol. 272, pp.726-730 (2000)

Chiavarino, B., Crestoni, M.E., Marzio, A.D., Fornarini, S. and Rosi, M. "Gas-phase ion chemistry of borazine, an inorganic analogue of benzene". *Journal of the American Chemical Society*, Vol. 121, pp. 11204-11210 (1999)

Chu, I.K., Guo, X., Lau, T.C. and Siu, K.W.M. "Sequencing of argentinated peptides by means of electrospray tandem mass spectrometry". *Analytical Chemistry*, Vol. 71, pp. 2364-2372 (1999)

Chu, I.K., Shoeib, T, Guo, X., Rodriquez, C.F., Lau, T.-C., Hopkinson, A.C. and Siu, K.W.M. "Characterization of the product ions form the collision-induced dissociation of argentinated peptides". *Journal of the American Society Mass Spectrometry*, Vol. 12, pp.163-175 (2001)

Cody, R.B., Amster, I.J. and Mclafferty, F.W. "Peptide mixture sequencing by tandem fourier – transform mass-spectrometry". *Proceedings of the National Academy fo Sciences of the United States of America*, Vol. 82, pp. 6367-6370 (1985)

Cooks, R.G. and Wong, P.S.H. "Kinetic method of making thermochemical determinations: Advances and applications". *Accounts of Chemical Research*, Vol. 31, pp. 379-386 (1998)

Cooks, R.G., Koskinen, J.T. and Thomas, P.D. "Special feature: Commentary-the kinetic method of making thermochemical determinants". *Journal of Mass Spectrometry*, Vol. 34, pp. 85-92 (1999)

Csaszar, A.G. "Conformers of gaseous α-alanine". *Journal of Physical Chemistry*, Vol. 100, pp. 3541-3551 (1996)

Cubero, E, Luque, F.J. and Orozco, M. "Is polarization important in cation-pi interactions?". *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 95, pp. 5976-5980 (1998)

Davidson, W.R. and Kebarle, P. "Binding-energies of stabilities of potassium-ion complexes from studies of gas-phase ion equilibria $K^+ + M = K^+M$". *Journal of the American Chemical Society*, Vol. 98, pp. 6133-6138 (1976a)

Davidson, W.R. and Kebarle, P. "Ionic solvation by aprotic-solvents-gas-phase solvation of alkali ions by acetonitrile". *Journal of the American Chemical Society*, Vol. 98, pp. 6125-6133 (1976b)

Davidson, W.R. and Kebarle, P. "Binding-energies and stabilities of potassium-ion complexes with ethylene diamine and dimethoxyethane (GlyMe) from measurements of complexing equilibria in gas-phase". *Canadian Journal of Chemistry-Revue Canadienne De Chimie*, Vol. 54, pp. 2594-2599 (1976c)

Decouzon, M., Gal, J.F., Maria, P.-C. and Raczynska, E.D. "Gas-phase basicity of N_1, N_1-dimethyl-N_2-alkylformamidines-substituent polarizability effects". *Journal of Organic Chemistry*, Vol. 5, pp. 3669-3673 (1991)

De Wall, S.L., Meadows, E.S., Barbour, L.J. and Gokel, G.W., "Synthetic receptors as models for alkali metal cation-π binding sites in proteins". *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 97, pp.6271-6276 (2000)

DelBene, J.E. "Basis set and correlation effects on computed lithium ion affinities". *Journal of Physical Chemistry*, Vol. 100, pp. 6284-6287 (1996)

DelBene, J.E., Mettee, H.D., Frisch, M.J., Luke, B.T. and Pople, J.A. "*Ab initio* computation of the enthalpies of some gas-phase hydration reactions". *Journal of Physical Chemistry*, Vol. 87, pp. 3279-3282 (1983)

Desiraju, G.R. and Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*, New York, Oxford University Press (1999)

Ding, Y.B. and KroghJespersen, K. "The 1:1 glycine zwitterion-water complex: An ab initio electronic structure study". *Journal of Computational Chemistry*, Vol. 17, pp. 338-349 (1996)

Doerksen, R.J. and Thakkar, A.J. "Quadrupole and octopole moments of heteroaromatic rings". *Journal of Physical Chemistry A*, Vol. 103, pp. 10009-10014 (1999)

Dongre, A.R. "Emerging tandem-mass-spectrometry techniques for the rapid identification of proteins". *Trends in Biotechnology*, Vol. 15, pp. 418-425 (1997)

Dookeran, N.N., Yalcin, T. and Harrison, A.G. "Fragmentation reactions of protonated α-amino acids ". *Journal of Mass Spectrometry*, Vol. 31, pp. 500-508 (1996)

Dougherty, D.A. "Cation-π interactions in chemistry and biology: A new view of benzene, Phe, Tyr, and Trp". *Science*, Vol. 271, pp. 163-168 (1996)

Dougherty, D.A. and Lester, H.A., "The crystal structure of a potassium channel - A new era in the chemistry of biological signaling". *Angewante Chemie-International Edition*, Vol. 37, pp. 2329-2331 (1998)

Doyle, D.A., Cabral, J.M., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T. and Mackinnon, R., "The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity". *Science*, Vol. 280, pp. 69-77 (1998)

Duan, G., Smith, V.H. and Weaver, D.F. " An *Ab initio* and data mining study on aromatic-amide interactions". *Chemical Physics Letters*, Vol. 310, pp. 323-332 (1999)

Dunbar, R.C. "Binding of Na^+, Mg^+, and Al^+ to the pi faces of naphthalene and indole: *Ab initio* mapping study". *Journal of Physical Chemistry A.*, Vol. 102, pp. 8946-8952 (1998)

Dunbar, R.C. "Complexation of Na^+ and K^+ to aromatic amino acids: A density functional computational study of cation-π interactions". *Journal of Physical Chemistry A*, Vol. 104, pp. 8067-8074 (2000)

Dunbar, R.C. "Metal cation binding to phenol: DFT comparison of the competing sites". *Journal of Physical Chemistry A*, Vol. 106, pp.7328-7337 (2002b)

Dzidic, I. and Kebarle, P. "Hydration of alkali ions in gas phase – enthalpies and entropies of reactions $M^+(H_2O)_{n-1}+H2O = M^+(H_2O)_n$". *Journal of Physical Chemistry*, Vol. 74, pp. 1466-1474 (1970)

Eller, K. and Schwarz, H. "Organometallic chemistry in the gas-phase". *Chemical Reviews*, Vol. 91, pp. 1121-1177 (1991)

Ernzerhof, M., Marian, C.M. and Peyerimhoff, S.D. "Energy derivative versus expectation value approach – the dipole-moment of CO". *Chemical Physics Letters*, Vol. 204, pp. 59-64 (1993)

Farrugia, J.M., Taverner, T. and O'Hair, R.A.J. "Side-chain involvement in the fragmentation reactions of the protonated methyl esters of histidine and its peptides ". *International Journal of Mass Spectrometry*, Vol. 209, pp. 99-112 (2001a)

Farrugia, J.M., O'Hair, R.A.J. and Reid, G.E. "Do all b(2) ions have oxazolone structures? Multistage mass spectrometry and ab initio studies on protonated N-acyl amino acid methyl ester model systems ". *International Journal of Mass Spectrometry*, Vol. 210/211, pp. 71-87 (2001b)

Feller, D., Dixon, D.A. and Nicholas, J.B. "Binding enthalpies for alkali cation-benzene complexes revisited". *Journal of Physical Chemistry A*, Vol. 104, pp.11414-11419 (2000a)

Feller, D., Glendening, E.D., Woon, D.E. and Feyereisen, M.W. 'An extended basis-set ab-initio study of alkali-metal cation-water clusters". *Journal of Chemical Physics*, Vol. 103, pp. 3526-3542 (1995)

Feng, W.Y., Gronert, C., Fletcher, K.A., Warres, A. and Lebrilla, C.B. "The mechanism of C-terminal fragmentations in alkali-metal ion complexes of peptides". *International Journal of Mass Spectrometry*, Vol. 222, pp.117-134 (2003)

Fernandez-Recio, J., Romero, A. and Sancho, J. "Energetics of a hydrogen bond (charged and neutral) and of a cation-π interaction in apoflavodoxin". *Journal of Molecular Biology*, Vol. 290, pp. 319-330 (1999)

Fessenden, R.J. and Fessenden, J.S., *Organic Chemistry*, Pacific Grove, Calif.: Brooks/Cole; London: International Thomson Pub, W. Grant Press, 6th edition (1998)

Figueroa, I.D. and Russell, D.H., "Matrix-assisted laser-desorption ionization hydrogen/deuterium exchange studies to probe peptide conformational changes". *Journal of the American Society for Mass Spectrometry*, Vol. 10, pp. 719-731 (1999)

Flocco, M.M. and Mowbray, S.L. "Planar stacking interactions of arginine and aromatic side-chains in proteins". *Journal of Molecular Biology*, Vol. 235, pp. 709-717 (1994)

Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Zakrzewski, V.G., Montgometry, J.A., Stratmann, R.E., Burant, J.C., Dapprich, S., Millam, J.M., Daniels, A.D., Kudin, K.N., Strain, M.C., Farkas, Oh., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D.

K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J.V., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M.A., Peng, C.., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P.M.W., Johnson, B.G., Chen, W., Wong, M.W., Andres, J.L., Gonzalez, C., Head-Gordon, M., Replogle, E.S. and Pople, J.A. *Gaussian 98*, Revision A.6, Gaussian, Inc., Pittsburgh, PA (1998)

Fong, T.M., Cascieri, M.A., Yu, H., Bansal, A., Seain, C. and Strader, C.D. "NH/pi interaction is important for the selective binding of receptor protein with neurotransmitter". *Nature*, Vol. 362, pp. 350-353 (1993)

Foresman, J.B. *Exploring Chemistry with Electronic Structure Methods*, Pittsburgh, P.A.; Gaussian, Inc., 2nd. Ed. (1996)

Gallivan, J.P. and Dougherty, D.A. "Cation-pi interactions in structural biology". *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 96, pp. 9459-9464 (1999)

Gallivan, J.P. and Dougherty, D.A. "A computational study of cation-π interactions vs salt bridges in aqueous media: implications for protein engineering". *Journal of the American Chemical Society*, Vol. 122, pp.870-874 (2000)

Gapeev, A. and Dunbar, R.C. "Cation-pi interactions and the gas-phase thermochemistry of the Na^+/phenylalanine complex". *Journal of the American Chemical Society*, Vol. *123*, 8360- 8365 (2001)

Grese, R.P., Cerny, R.L. and Gross, M.L., "Metal-ion peptide interactions in the gas-phase – A tandem mass spectrometry study of alkali-metal cationized peptides". *Journal of the American Chemical Society*, Vol. 111, pp. 2835-2842 (1989)

Grese R.P. and Gross M.L. "Gas-phase interactions of lithium Ions and dipeptides'. *Journal of the American Chemical Society*, Vol. 112, pp. 5098-5104 (1990)

Godfrey, P.D., Firth, S., Hatherley. L.D., Brown, R.D. and Pierlot, A.P. 'Millimeter-wave spectroscopy of biomolecules: Alanine". *Journal of the American Chemical Society*, Vol. 115, pp.9687-9691 (1993)

Gokel, G.W., Barbour, L.J., Ferdani, R. and Hu, J.X., "Lariat ether receptor systems show experimental evidence for alkali metal cation-π interactions". *Accounts of Chemical Research*, Vol. 35, pp. 878-886 (2002)

Gokel, G.W., De Wall, S.L. and Meadows, E.S., "Experimental evidence for alkali metal cation-π interactions". *European Journal of Organic Chemistry*, Vol. 17, pp. 2967-2978 (2000)

Guo, B.C. and Castleman, A.W. "The bonding strength of Ag^+ (C_2H_4) and Ag^+ $(C_2H_4)_2$ complexes". *Chemical Physics Letters,* Vol. 81, 16-20 (1991)

Hammond, G.S. "A Correlation of xn Rate: Hammond Postulate". *Journal of the American Chemical Society*, Vol. 77, pp. 334-338 (1995)

Harrison, A.G. "The gas-phase basicities and proton affinities of amino acids and peptides". *Mass Spectrometry Reviews*, Vol. 16, pp. 201-217 (1997)

Harshbarger, W., Lee, G., Porter, R.F. and Bauer, S.H. *Inorganic Chemistry*, Vol. 8, pp. 1683 (1969)

Hehre, W.J., Radom, L., Scleyer, P.V.R. and Pople, J., eds., *Ab initio Molecular Orbital Theory,* Wiley, New York (1986)

Hille, B. *Ion Channels of Excitable Membranes*, Sinauer, Sunderland, Massachusetts, 3rd. Edition (2001)

Hirota, S. Hayamizu, K., Okuno, T., Kishi, M., Iwasaki, H., Kondo, T., Hibino, T. Takabe, T. Kohzurna, T. and Yamauchif, O. "Spectroscopic and electrochemical studies n structural change of plastocyanin and its tyrosine 83 mutants induced by interaction with lysine peptides". *Biochemistry*, Vol. 39, pp. 6357-6364 (2000)

Hoppilliard Y, Rogalewicz F. and Ohanessian G., "Structures and fragmentations of zinc(II) complexes of amino acids in the gas phase. II. Decompositions of glycine-Zn(II) complexes". *International Journal of Mass Spectrometry*, Vol. 204, pp. 267-280 (2001)

Hoyau, S. and Ohanessian, G., " Interaction of alkali metal cations (Li^+-Cs^+) with glycine in the gas phase: A theoretical study". *Chemistry-A European Journal*, Vol. 4, pp. 1561-1569 (1998)

Hoyau, S., Norrman, K., McMahon, T.B. and Ohanessian, G. "A quantitative basis for a scale of Na^+ affinities of organic and small biological molecules in the gas phase". *Journal of the American Chemical Society*, Vol. 121, pp. 8864-8875 (1999)

Hoyes, E. and Gaskell, S.J. "Automatic function switching and its usefulness in peptide and protein analysis using direct infusion microspray quadrupole time-of-flight mass spectrometry". *Rapid Communications in Mass Spectrometry*, Vol. 15, pp. 1802-1806, (2001)

Hpewell, P.C. "Tuberculosis in persons with human immunodeficiency virus nfection: Clinical and public health aspects". *Seminars in Respiratory and Critical Care Medicine*, Vol. 18, pp. 471-484 (1997)

Hsieh-Wilson, L.C., Schultz, P.G. and Stevens, R.C. "Insights into antibody catalysis: structure of an oxygeonation catalyst at 1.9-angstrom resolution". *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 93, pp.5363-5367 (1996)

Hu, J.X., Barbour, L.J. and Gokel, G.W. "Probing alkali metal-pi interactions with the side chain residue of tryptophan". *Proceedings of the National academy of Sciences of the United States of America*, Vol. 99, pp.5121-5126 (2002a)

Hu, J.X., Barbour, L.J. and Gokel, G.W. "The indole side chain of tryptophan as a versatile pi-donor". *Journal of the American Chemical Society*, Vol. 124, pp. 10940-10941 (2002b)

Huang, H. and Rodgers, M.T. "Sigma versus pi interactions in alkali metal ion binding to azoles: Threshold collision-induced dissociation and ab initio theory studies". *Journal of Physical Chemistry A*, Vol. 106, pp. 4277-4289 (2002)

Huang, Y., Wysocki, V.H., Tabb, D.L. and Yates, J.R., III, "The influence of histidine on cleavage C-terminal to acidic residues in doubly protonated tryptic peptides". *International Journal of Mass Spectrometry*, Vol. 219, pp. 233-244 (2002b)

Huisken, F. and Werhahn, O. "The O-H stretching vibrations of glycine trapped in rare gas matrices and helium clusters". *Journal of Chemical Physics*, Vol. 111, pp. 2978-2984 (1999)

Hughes, M.N., *The Inorganic Chemistry of Biological Processes*, 2nd Ed., John Wiley; Chichester, New York, pp.89-124, 257-295 (1972)

Huheey, J.E., Keiter, E.A. and Keiter, R.L. *Inorganic Chemistry: Principles of Structure and Reactivity*. In 4th, , Harper College Publishers, pp.187 (1993)

Hunter, E.P. and Lias, S.G., *Journal of Physical and Chemical Reference Data*, Vol. 27, pp. 413-656 (1998)

Iceman, C. and Armentrout, P.B. "Collision-induced dissociation and theoretical studies of K^+ complexes with ammonia: a test of theory for potassium ions". *International Journal of Mass Spectrometry*, Vol. 222, pp. 329-349 (2003)

Iijima, K., Tanaka, K. and Onuma, S. "Main conformer of gaseous glycine: molecular structure and rotational barrier from electron diffraction data and rotational constants". *Journal of Molecular structure (Theochem)*, Vol. 246, pp. 257-266 (1991)

Ikuta, S. "*Ab initio* MO calculations on the stable structures and binding-energies of H^+CO and H^+N_2, Li^+CO and Li^+N_2, Na^+CO and Na^+N_2, K^+CO and K^+N_2 ions". *Chemical Physics Letters*, Vol. 109, pp. 550-553 (1984)

Ikuta, S. "A theoretical study on the conformations and energetics on the cation-pi interaction between monovalent ions ($M^+ = Li^+$, Na^+ and K^+) and anthracene and phenanthrene molecules". *Journal of Molecular structure (Theochem)*, Vol. 530, pp. 201-207 (2000)

Israelachvili, J. N. *Intermolecular and Surface Force*; Academic Press: London (1992)

Isupov, M.N., Antson, A.A., Dodson, E.J., Dodson, G.C., Dementieva, I.S., Zakomirdina, L.N., Wilson, K.S., Dauter, Z.J., Lebedev, A.A. and Harutyunyan, E.H. "Crystal structure of tryptophanase". *Journal of Molecular Biology*, Vol. 276, pp. 603-623 (1998)

Jeffrey G. A. and Saenger, W. *Hydrogen bonding in biological structures*; Springer: New York (1991)

Jensen, F. "Structure and stability of complexes of glycine and glycine methyl analogous with H^+, Li^+, and Na^+". *Journal of the American Chemical Society*, Vol. 114, pp. 9533-9537 (1992)

Jensen, F. *An Introduction to computational Chemistry*, Chichester: New York, N.Y.; Wiley (1999)

Jensen, J.H. and Gordon, M.S. "On the number of water-molecules necessary to stabilize the glycine zwitterion". *Journal of the American Chemical Society*, Vol. 117, pp. 8159-8170 (1995)

Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chalt, B.T. and Mackinnon, R. "X-ray structure of a voltage-dependent K^+ Channel". *Nature*, Vol. 423, pp. 33-41 (2003)

Jockusch, R. A., Lemoff, A. S., and Williams, E. R. "Effect of metal ion and water coordination on the structure of a gas-phase amino acid". *Journal of the American Chemical Society*, Vol. 123, pp. 12255-12265 (2001)

Jockusch, R.A., Price, W.D. and Williams, E.R. "Structure of cationized arginine (Arg center dot M^+, M = H, Li, Na, K, Rb, and Cs) in the gas phase: Further evidence for zwitterionic arginine". *Journal of Physical Chemistry A*, Vol. 103, pp. 9266-9274 (1999)

Kabsh, W. *Acta Crystallogr.*, Vol. A32, pp. 944 (1976)

Kabsh, W. *Acta Crystallogr*, Vol. A34, pp. 827 (1978)

Kaim, W. and Schwederski, B., eds., *Bioinorganic Chemistry: Inorganic Elements in the Chemistry of Life: An Introduction and Guide,* John Wiley and Sons, New York, pp. 267-286 (1994)

Katritzky, A.R., Karelson, M., Sild, S., Krygowski, T.M. and Jug, K. "Aromaticity as a quantitative concept. 7. Aromaticity reaffirmed as a multidimensional characteristic". *Journal of Organic Chemistry*, Vol. 63, pp. 5228-5231 (1998)

Kebarle, P. "Equilibrium studies of the solvated proton by high pressure mass spectrometry. Thermodynamic determinations and implications for the electrospray ionization process". *Journal of Mass Spectrometry*, Vol. 32, pp. 922-929 (1997)

Keesee, R.G. and Castleman, A.W. "Thermochemical data on gas-phase ion-molecule association and clustering reactions". *Journal of Physical and Chemical Reference Data*, Vol. 15, pp. 1011-1071 (1986)

Kenny, P.T.M., Nomoto, K. and Orlando, R. "Fragmentation of gas-phase complexes between alkali metal ions and a biologically active peptide, Achatin-I". *Rapid Communication in Mass Spectrometry*, Vol. 11, pp.224-227 (1997)

Kim, D., Hu, S., Tarakeshwar, P. and Kim, K.S. "Cation-π Interactions: A theoretical investigation of interaction of metallic and organic cations with alkenes, arenes, and heteroarenes". *Journal of Physical Chemistry A*, Vol. 107, pp.1228-1238 (2003b)

Kiran, B., Phukan, A.K. and Jemmis, E.D. "Is borazine aromatic? Unusual parallel behavior between hydrocarbons and corresponding B-N analogues". *Inorganic Chemistry*, Vol. 40, pp. 3615-3618 (2001)

Kish, M. M., Ohanessian, G. and Wesdemiotis, C. " The Na^+ affinities of alpha-amino acids: side-chain substituent effects". *International Journal of Mass Spectrometry*, Vol. 227, pp. 509-524 (2003)

Klassen, J.S., Anderson, S.G., Blades, A.T. and Kebarle, P. "Reaction Enthaplies for $M^+L = M^+ + L$, where $M^+ = Na^+$ and K^+ and L = Acetamide, N-methylacetamide, N,N-dimethylacetamide, glycine, and glycylglycine, from determinations of collision-induced dissociation thresholds". *Journal of Physical Chemistry*, Vol. 100, pp. 14218-14227 (1996)

Koch, W. and Holthausen, M.C., eds., *A Chemist's Guide to Density Functional Theory*, Wiley-VCH (2000)

Koehl, P. and Levitt, M. "De Novo Protein Design. I. In Search of Stability and Specificity". *Journal of Molecular biology*, Vol. 293, pp. 1161-1181 (1999)

Koellner, G., Steiner, T., Millard, C.B., Silman, I. and Sussman, J.L. "A neutral molecule in a cation-binding site: Specific binding of a PEG-SH to acetylcholnesterase from Torpedo californica". *Journal of Molecular Biology*, Vol. 320, pp. 721-725 (2002).

Kohtani, M., Kinnear, B.S., and Jarrold, M.F., "Metal-ion enhanced helicity in the gas phase'. *Journal of the American Chemical Society*, Vol., 122, pp.12377-12378 (2000)

Kulik, W. and Heerma, W. "A study of the positive and negative-ion fast atom bombardment mass-spectra of alpha-amino acids". *Biomedical and Environmental Mass Spectrometry*, Vol. 15, pp. 419-427 (1988)

Larsen, T.M., Laughlin, L.T., Holden, H.M., Rayment, I. and Reed, G.H. "Structure of rabbit muscle pyruvate-kinase complexed with Mn^{2+}, K^{+} and pyruvate". *Biochemistry*, Vol. 33, pp. 6301-6309 (1994)

Lau, J.K.-C., Wong, C.H.S., Siu, F.M., Ma, N.L. and Tsang, C.W. "Absolute potassium cation affinities (PCAs) in the gas phase". *Chemistry-A European Journal*, Vol.9, pp. 3383-3396 (2003)

Lavanant, H. and Hoppilliard, Y. "Formation and fragmentation of alpha-amino acids complexed by Cu^{+}". *Journal of Mass Spectrometry*, Vol. 32, pp. 1037-1049 (1997)

Lavanant H., Hecquet, E. and Hoppilliard, Y. "Complexes of L-histidine with Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+} studied by electrospray ionization mass spectrometry". *International Journal of Mass Spectrometry*, Vol. 187, pp. 11-23 (1999)

Lee, C., Yang, W. and Parr, R.G. "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density". *Physical Review*, Vol. 37, pp. 785-789 (1988)

Lehniger, A.L. *Principles of biochemistry*; Worth Publishers Inc.: New York (1982)

Li, J. and Brill, T.B. "Decarboxylation mechanism of amino acids by density functional theory". *Journal of Physical Chemistry A*, Vol. 107, pp.5993-5997 (2003)

Lias, S.G., Bartmass, J.E., Liebman, J.F., Homes, J.L., Levin, R.D. and Mallard, W.G. "Gas phase ion and neutral thermochemistry". *Journal of Physical and Chemical Reference Data*, Vol. 17, pp. 1-861, Suppl. 1 (1988)

Lias, S.G., Liebman, J.F. and Levin, R.D. "Evaluated gas phase basicities and proton affinities of molecules – heats of formation of protonated molecules". *Journal of Physical and Chemical Reference Data*, Vol. 13, pp. 696-808 (1984)

Lioe, H., O'Hair, R.A. and Reid, G.E. "Gas-phase reactions of protonated tryptophan'. *Journal of the American Society in Mass Spectrometry*, Vol. 15, pp. 65-76 (2004)

Lippard, S.J. and Berg, J.M., *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, California (1994)

Locke, M.J. and Mclver, J.R.T. "Effect of solvation on the acid-base properties of glycine". *Journal of the American Chemical Society*, Vol. 105, pp. 4226-4232 (1983)

Loo, J.A. and Muenster, H., "Magnetic sector-ion trap mass spectrometry with electrospray ionization for high sensitivity peptide sequencing". *Rapid Communications in Mass Spectrometry*, Vol. 13, pp. 54-60 (1999)

Luna, A., Amekraz, B., Tortajada, J., Morizur, J. P.,Alcami, M., Mo, O. and Yanez, M. "Modeling the interactions between peptide functions and Cu(I): Formamide – CU+ reactions in the gas phase". *Journal of the American Chemical Society*, Vol. 120, pp. 5411-5426 (1998)

Luque, F.J. and Orozco, M. "Reactivity of planar and twisted amides in vacuum and aqueous environments and ab initio mep study". *Journal of the Chemical Society-Perkin Transactions 2*, Vol. 4, pp. 683-690 (1993)

Ma, J.C. and Dougherty, D.A. "The cation-π interaction". *Chemical Reviews*, Vol. 97, pp. 1303-1324 (1997)

Ma, N.L., Li, W.K. and Ng, C.Y. "A Gaussian-2 *Ab initio* study of van der Waals dimers R_1R_2 and their cations $R_1R_2^+$ (R1, R2 = He, Ne, Ar, and Kr)". *Journal of Chemical Physics*, Vol. 99, pp. 3617-3621 (1993)

Ma, N.L., Siu, F.M. and Tsang, C.W. "Interaction of alkali metal cations and short chain alcohols: effect of core size on theoretical affinities". *Chemical Physics Letters*, Vol. 322, pp. 65-72 (2000)

Magnusson, E. "Binding of polar-molecules to Li^+, Na^+, K^+, Mg^{2+}, and Ca^{2+} in single-ligand adducts M^+L and $M^{2+}L$ (L=H_2O, NH_3, H_2S, PH_3)". *Journal of Physical Chemistry*, Vol. 98, pp. 12558-12569 (1994)

Mallis, L.M. and Russell, D.H. "Fast atom bombardment – tandem mass-spectrometry studies of organo-alkyli metal ions of small peptides". *Analytical Chemistry*, Vol. 58, pp. 1076-1080 (1986).

Maksic, Z.B. and Kovacevic, B. "Towards the absolute proton affinities of 20 alpha-amino acids". *Chemical Physics Letters*, Vol. 307, pp. 497-504 (1999)

Mao, L., Wang, Y., Liu, Y. and Hu, X. "Molecular determinants for ATP-binding in Proteins: A data mining and quantum chemical analysis". *Journal of Molecular Biology*, Vol. 336, pp. 787-807 (2004)

March, J. Advanced Organic Chemistry: Reactions, Mechanisms, and Structure, 3rd. Ed., J. Wiley and Sons: McGraw-Hill, New York (1985).

Marino, T., Russo, N. and Toscano, M. "Gas-phase metal ion (Li^+, Na^+, Cu^+) affinities of glycine and alanine". *Journal of Inorganic Biochemistry*, Vol. 79, pp. 179–185 (2000)

Marino, T., Russo, N. and Toscano, M. "Internation of Li^+, Na^+, and K^+ with the proline amino acid. Complexation modes, potential energy profiles, and metal ion affinities". *Journal of Physical Chemistry B*, Vol. 107, pp. 2588-2594 (2003)

Matta, C.F. and Bader, R.F.W. "An atoms-in-molecules study of the genetically-encoded amino acids: I. Effects of conformation and of tautomerization on geometric, atomic, and bond properties". *Proteins: Structure, Function, and Genetics*, Vol. 40 pp.310-329 (2000)

McMahon, T.B. and Ohanessian, G. "An experimental and ab initio study of the nature of the binding in gas-phase complexes of sodium ions". *Chemistry-A European Journal*, Vol. 6, pp. 2931-2941 (2000)

McQuarrie, D. A. and Simon, J. D. *Physical Chemistry: a Molecular Approach*, Sausalito, Calif. : University Science Books (1997)

Mecozzi, S., West, A.P., Jr. and Dougherty, D.A. "Cation-pi interactions in aromatics of biological and medicinal interest: Electrostatic potential surfaces as a useful qualitative guide". *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 93, pp. 10566-10571 (1996)

Melo, A., Ramos, M.J., Floriano, W. B., Gomes, J.A.N.F., Leao, J.F.R., Magalhaes, A.L., Maigret, B., Nascimento, M.C. and Reuter, N. "Theoretical study of arginine-carboxylate interactions". *Journal of Molecular structure (Theochem)*, Vol. 463, pp.81-90 (1999)

Meyer, E.A., Castellano, R.K. and Diederich, F. "Interactions with aromatic rings in chemical and biological recognition". *Angewante Chemie-international Edition*, Vol. 42, pp. 1210-1250 (2003)

Miehlich, B., Savin, A., Stoll, H. and Preuss, H. "Results obtained with the correlation-energy density functionals of Becke and Lee, Yang and Parr". *Chemical Physics Letters*, Vol. 157, pp. 200-206 (1989)

Miller, C., "Potassium selectivity in proteins-oxygen cage or in the F-ace". *Science*, Vol. 261, pp. 1692-1693 (1993)

Mohamadi, F., Richards, N.G.J., Guida, W.C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T. and Still, W.C. "Macromodel – an integrated software system for modling organic and bioorganic molecules using molecular mechanics". *Journal of. Computational Chemistry*, Vol. 11, pp. 440-467 (1990). Macromodel is an integrated software system for modeling organic and bioorganic molecules using molecular mechanics written by them

Moision, R.M. and Armentrout, P.B., "Experimental and theoretical dissection of sodium cation/glycine interactions". *Journal of Physical Chemistry A*, Vol. 106, pp. 10350-10362 (2002)

More, M.B., Ray, D. and Armentrout, P.B. "Cation-ether complexes in the gas phase: Bond dissociation energies of K^+ (dimethyl ether)$_x$, x = 1-4; K^+(1,2-dimethoxyethane)$_x$, x = 1 and 2; and K^+(12-crown-4)". *Journal of Physical Chemistry A*, Vol. 101, pp. 4254 –4262 (1997)

Nakamura, R.L., Anderson, J.A. and Gaber, R.F. "Determination of key structural requirements of a K^+ channel pore". *Journal of Biological Chemistry*, Vol. 272, pp. 1011-1018 (1997)

Ng, K.M., Ma, N.L. and Tsang, C.W. "Cation-aromatic pi interaction in the gas phase: an experimental study on relative silver (I) ion affinities of polyaromatic hydrocarbons". *Rapid Communications in Mass Spectrometry*, Vol. 12, pp. 1679-1684 (1998)

Nguyen, K.A. and Pachter, R. "Ground state electronic structures and spectra of zinc complexes of porphyrin, tetraazaprophyrin, tetrabenzoporphyrin, and phthalocyanine: A density functional theory study". *Journal of Chemical Physics*, Vol. 114, pp. 10757-10767 (2001)

Nielsen, S.B., Masella, M. and Kebarle, P. "Competitive gas-phase solvation of alkali metal ions by water and metal". *Journal of Physical Chemistry A*, Vol. 103, pp. 9891-9898 (1999)

Nino, A., Munoz-Caro, C., Carbo-Dorca, R. and Girones, X. "Rational modeling of the voltage-dependent K^+ Channel inactivation by amino pyridines". *Biophysical Chemistry*, Vol. 104, pp. 417-427 (2003)

Nurizzo, D., Baker, H.M., He, Q.Y., MacGillivray, R.T.A., Mason, A.B., Woodworth, R.C. and Baker, E.N., 'Crystal structures and iron release properties of mutants (K206A and K96A) that abolish the dilysine interaction in N-lobe of human transferring". *Biochemistry*, Vol. 40, pp. 1616-1623 (2001)

Oh, K.S., Lee, C.W., Choi H. S., Lee, S.J. and Kim, K.S. "Origin of the high affinity and selectivity of novel receptors for NH_4^+ over K^+: charged hydrogen bonds versus cation-π interaction". *Organic Letters*, Vol. 2, pp. 2679-2681 (2000)

O'Hair, R.A.J., Broughton, P.S., Styles, M.L., Frink, B.T. and Hadad, C.M. "the fragmentation pathways of protonated glycine: A computational study". *Journal of the American Society for Mass Spectrometry*, Vol. 11, pp. 687-696 (2000)

Okada, A., Miura, T. and Takeuchi, H. "Protonation of histidine and histidine-tryptophan interaction in the activation of the M ion channel from influenza A". *Biochemistry*, Vol.40 pp. 6053-6060 (2001)

Paizs, B., Csonka, I.P., Lendvay, G. and Suhai, S. "Proton mobility in protonated glycylglycine and N-formylglycylglycinamide: a combined quantum chemical and RKKM study". *Rapid Communications in Mass Spectrometry*, Vol. 15, pp. 637-650 (2001)

Parker, J.K. and Davis, S.R. "Ab initio study of the relative energies and properties of fluoroborazines". *Journal of Physical Chemistry A*, Vol. 101, pp. 9410-9414 (1997)

Parr, R.G. and Yang, W. *Density Functional Theory*, Oxford University Press (1989)

Perdew, J.P. "Density-functional approximation for the correlation energy of the inhomogeneous electron gas". *Physical Review B*, Vol. 33, pp. 8822-8824 (1986).

Petrie, S. "An improved theoretical sodium cation affinity scale?". *Journal of Physical Chemistry A*, Vol. 105, pp. 9931-9938 (2001)

Price, W.D., Jockusch, R.A. and Williams, E.R. " Is arginine a zwitterion in the gas phase?". *Journal of the American Chemical Society*, Vol. 119, pp. 11988-11989 (1997)

Pulkkinen, S., Noguera, M., Rodriguez-Santiago, L., Sodupe, M. and Bertran, J. "Gas phase intramolecular proton transfer in cationized glycine and chlorine substituted derivatives (M-Gly, M = Na^+, Mg^{2+}, Cu^+, Ni^+, and Cu^{2+}): Existence of zwitterionic structures?". *Chemistry - A European Journal*, Vol. 6, pp. 4393-4399 (2000)

Rappe, A.K. and Casewit, C.J. *Molecular Mechanics Across Chemistry*, University Science Books (1997)

Reed, A.E., Curtiss, L.A. and Weinhold, F. "Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint". *Chemical Reviews*, Vol. 88, pp. 899-926 (1988)

Remko, M. "Structure and gas phase stability of complexes L-M, where M=Li^+, Na^+, Mg^{2+} and L is formaldehyde, formic acid, formate anion, formamide and their sila derivatives". *Molecular Physics*, Vol. 91, pp. 929-926 (1997)

Robertson, E.G. and Simons, J.P. "Getting into shape: Conformational and supramolecular landscapes in small biomolecules and their hydrated clusters". *Physical Chemistry Chemical Physics*, Vol. 3, pp.1-18 (2001)

Rodgers, M.T., "Substituent Effects in the binding of alkali metal ions to pyridines, studied by threshold collision-induced dissociation and *ab initio* theory: the methylpyridines". *Journal of Physical Chemistry A*, Vol. 105, pp. 2374-2383 (2001)

Rodgers, M.T. and Armentrout, P.B. "Absolute binding energies of lithium ions to short chain alcohols, $C_nH_{2n+2}O$, n = 1-4, determined by threshold collision-induced dissociation". *Journal of Physical Chemistry A*, Vol. 101, pp. 2614-2625 (1997)

Rodgers, M.T. and Armentrout, P.B. "Absolute alkali metal ion binding affinities of several azoles,determined by threshold collision-induced dissociation". *International Journal of Mass Spectrometry*, Vol. 185/186/187, pp. 359-380 (1999)

Rodgers, M.T., and Armentrout, P.B., "Noncovalent metal-ligand bond energies as studied by threshold collision-induced dissociation". *Mass Spectrometry Reviews*, Vol. 19, pp. 215-247 (2000a)

Rodgers, M.T. and Armentrout, P.B. "Noncovalent interactions of nucleic acid bases (uracil, thymine, and adenine) with alkali metal ions. Threshold collision-induced dissociation and theoretical studies". *Journal of the American Chemical Society*, Vol. 122, pp. 8548-8558 (2000b)

Rodriquez, C.F., Cunje, A., Shoeib, T., Chu, I.K., Hopkinson, A.C. and Siu, K.W.M. "Proton migration and tautomerism in protonated triglycine". *Journal of the American Chemical Society*, Vol. 123, pp. 3006-3012 (2001)

Rodriquez, C.F.; Shoeib, T., Chu, I.K., Siu, K.W.M. and Hopkinson, A.C. "Comparison between protonation, lithiations, and argentination of 5-oxazolones: A study of a key intermediate in gas-phase peptide sequencing". *Journal of Physical Chemistry A*, Vol. 104, pp. 5335-5342 (2000)

Rodriquez, C.F., Cunje, A., Shoeib, T., Chu, I.K., Hopkinson, A. and Siu, K.W.M. "Solvent-assisted rearrangements between tautomers of protonated peptides". *Journal of Physical Chemistry A*, Vol. 104, pp. 5023-5028 (2000b)

Rodriguez-Santiago, L., Sodupe, M. and Tortajada, J. "Gas-phase reactivity of Ni^+ with glycine". *Journal of Physical Chemistry A*, Vol. 105, pp. 5340-5347 (2001)

Roepstorff, P. "Mass Spectrometry in protein studies from genome to function". *Current Opinion in Biotechnology*, Vol. 8, pp. 6-13 (1997)

Rogalewicz, F., Ohanessian, G. and Gresh, N. "Interaction of neutral and zwitterionic glycine with Zn^{2+} in gas phase: Ab Initio and SIBFA molecular mechanics calculations". *Journal of Computational Chemistry*, Vol. 21, pp. 963-973 (2000a)

Rogalewicz, F. and Hoppilliard, Y. "Low energy fragmentation of protonated glycine. An ab initio theoretical study". *International Journal of Mass Spectrometry*, Vol. 199, pp. 235-252 (2000b)

Rogalewicz, F., Hoppilliard, Y. and Ohanessian, G. "Fragmentation mechanisms of α-amino acids protonated under electrospray ionization: a collisional activation and ab initio theoretical study". *International Journal of Mass Spectrometry*, Vol. 195/196, pp. 565-590 (2000c)

Rogalewicz, F., Hoppilliard, Y. and Ohanessian, G. "Structures and fragmentations of zinc(II) complexes of amino acids in the gas phase. III. Rearrangement versus desolvation in the electrospray formation of the glycine-zinc complex". *International Journal of Mass Spectrometry*, Vol. 206, pp. 45-52 (2001)

Roux, B. and Karplus, M. "Potential energy function for cation-peptide interactions – An ab –initio study". *Journal of Computational Chemistry*, Vol. 16, pp. 690-704 (1995)

Russell, D.H., McGlohon, E.S. and Mallis, L.M. "Fast-atom bombardment tandem mass spectrometry studies of organo-alkali-metal ions of small peptides. Competitive interaction of sodium with basic amino acid substituents". *Analytical Chemistry*, Vol. 60, pp. 1818-1824 (1988).

Russo, N., Toscano, M. and Grand, A. "Lithium affinity for DNA and RNA nucleobases. The role of theoretical information in the elucidation of the mass spectrometry data". *Journal of Physical Chemistry B*, Vol. 105, pp. 4735-4741 (2001a)

Russo, N., Toscano, M. and Grand, A. "Bond energies and attachments sites of sodium and potassium cations to DNA and RNA nucleic acid bases in the gas phase ". *Journal of the American Chemical Society*, Vol. 123, pp. 10272-10279 (2001b)

Ryzhov, V. and Dunbar, R.C. "Interactions of phenol and indole with metal ions in the gas phase: models for Tyr and Trp side-chain binding". *Journal of the American Chemical Society*, Vol. 121, pp. 2259-2268 (1999)

Ryzhov, V., Dunbar, R.C., Cerda, B. and Wesdemiotis, C. "Cation-π effects in the complexation of Na^+ and K^+ with Phe, Tyr, and Trp in the gas phase ". *Journal of the American Society for Mass Spectrometry*, Vol. 11, pp. 1037-1046 (2000)

Sarkhel, S. and Desiraju, G.R. " N-H...O, O-H...O, and C-H...O hydrogen bonds in protein-ligamd complexes: Strong and weak interactions in molecular recognition". *Proteins: Structure, Function, and Bioinformatics*, Vol. 54, pp. 247-259 (2004)

Saunders, M., Houk, K.N., Wu, Y.D., Still, W.C., Lipton, M., Chang, G. and Guida, W.C. "Conformations of cyclohept adecane – A comparison of methods for conformational searching". *Journal of the American Chemical Society*, Vol. 112, pp. 1419-1427 (1990)

Schleyer, P.v.R., Freeman, P.K., Jiao, H.J. and Goldfuss, B. "City and antiaromaticity in 5-membered C4H4X ring systems – classical and magnetic concepts may not be

orthogonal". *Angewante Chemi-International Edition in English*, Vol. 34, pp. 337-340 (1995)

Scott, A.P. and Radom, L. "Harmonic vibrational frequencies: An evaluation of Hartree-Fock, Moller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors". *Journal of Physical Chemistry*, Vol. 100, pp. 16502-16513 (1996)

Shirazian, S. and Gronert, S. "The gas-phase conformations of valine: an ab initio study" *Journal of Molecular Structure (Theochem)*, Vol. 397, pp. 107-112 (1997)

Shoeib, T., Cunje, A., Hopkinson, A.C. and Siu, K.W.M. "Gas-phase fragmentation of the Ag^+ phenylalanine complex: Cation-π interactions and radical cation formation". *Journal of the American Society for Mass Spectrometry*, Vol. 13, pp.408-416 (2002)

Shoeib, T., Hopkinson, A.C and Siu, K.W.M. "Collision-induced dissociation of the Ag^+ proline complex: Fragmentation pathways and reaction mechanisms – A synergy between experiment and theory". *Journal of Physical Chemistry B*, Vol. 105, pp.12399-12409 (2001b)

Shoeib, T., Rodriquez, C.F., Siu, K.W.M. and Hopkinson, A.C., "A Comparison of copper (I) and silver (I) complexes of glycine, diglycine and triglycine". *Physical Chemistry Chemical Physics*, Vol. 3, pp.853-861 (2001)

Silverman, S.K., Lester, H.A. and Dougherty, D.A. "Asymmetrical contributions of subunit pore regions to ion selectivity in an inward rectifier K^+ channel". *Biophysical Journal*, Vol. 75, pp. 1330-1339 (1998)

Siu, F.M., "Studies on non-covalent metal cation-molecule complexes". *PhD thesis,* (2001d)

Siu, F.M., Ma, N.L. and Tsang, C.W. "Theoretical binding energies of lithium ions to short-chain alcohols". *Chemical Physics Letters*, Vol. 288, pp. 408-412 (1998)

Siu, F.M., Ma, N.L. and Tsang, C.W. "Alkali metal cation-ligand affinities: Basis set superposition correction for Gaussian protocols". *Journal of Chemical Physics*, Vol. 114, pp. 7045-7051 (2001a)

Siu, F.M., Ma, N.L. and Tsang, C.W. "Cation-pi interactions in sodiated phenylalanine complexes: Is phenylalaine in the charge-solvated or zwitterionic form". *Journal of the American Chemical Society*, Vol. 123, pp. 3397-3398 (2001b)

Siu F.M., Ma, N.L. and Tsang, C.W., "The competition between π and non-π cation binding sites in aromatic amino acid: a theoretical study of the alkali metal cation (Li^+, Na^+, K^+)-phenylalanine complexes". *Chemistry – A European Journal*, (2004), accepted

Smit, A.B., 'A glia-derived acetylcholne-binding protein that modules synaptic transmission". *Nature*, Vol. 411, pp. 261-268 (2001)

Snoek, L.C., Robertson, E.G., Kroemer, R.T. and Simons, "Conformational landscapes in amino acids: infrared and ltraviolet ion-dip spectroscopy of phenylalanine in the gas phase". *Chemical Physics Letters*, Vol. 321, pp. 49-56 (2000)

Spears, K.G. "Ion-neutral bonding". *Journal of Chemical Physics*, Vol. 57, pp. 1850-1972 (1972)

Stenesh, J., *Biochemistry*, New York: Plenum, Press, pp. 117-139 (1998)

Stepanian, S.G., Reva, I.D., Radchenko, E.D. and Adamowicz, L. "Combined matrix-isolation infrared and theoretical DFT and ab initio study of nonionized valine conformers". *Journal of Physical Chemistry A*, Vol. 103, pp. 4404-4412 (1999)

Strittmatter, E.F., Lemoff, A.S. and Williams, E.R. "Structure of cationized glycine, Gly M^{2+} (M = Be, Mg, Ca, Sr, Ba), in the gas phase: Intrinsic effect of cation size on zwitterion stability". *Journal of Physical Chemistry A*, Vol. 104, pp. 9793-9796 (2000)

Stryer, L., *Biochemistry*, New York, W. H. Freeman, pp.291-322, Chapter 12, 4th edition (1995)

Sudha, R.; Panda, M., Chandrasekhar, J. and Balaram, P., "Structural effects on the formation of proton and alkali metal ion adducts of apolar, neutral peptides: electrospray ionization mass spectrometry and ab initio theoretical studies". *Chemisty –A European Journal*, Vol. 8, pp. 4980–4991 (2002)

Sunner, J. and Kebarle, P. "Ion solvent molecule interactions in the gas-phase- the potassium-ion and Me_2SO, DMA, DMF, and acetone". *Journal of the American Chemical Society*, Vol. 106, pp. 6135-6139 (1984)

Sunner, J., Nishizawa, K. and Kebarle, P. "Ion-solvent molecule interactions in the gas-phase – The potassium-ion and benzene". *Journal of Physical Chemistry*, Vol. 85, pp. 1814-1820 (1981)

Sussman, J.L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L. and Silman, I. "Atomic – structure of acetylcholinesterase from torpedo-californica – a prototypoc acetylcholine-binding protein". *Science*, Vol. 253, pp. 872-879 (1991)

Talaty, E.R., Perera, B.A., Gallardo, A.L., Barr, J.M. and Van Stipdonk, M.J. "Elucidation of fragmentation pathways for the collision-induced dissociation of the binary Ag(I) complex with phenylalanine". *Journal of Physical Chemistry A*, Vol. 105, pp. 8059-8068 (2001)

Talley, J.M., Cerda, B.A., Ohanessian, G. and Wesdemiotis, C. "Alkali metal ion binding to amino acids versus their methyl esters: Affinity trends and structural changes in the gas phase". *Chemistry - A European Journal*, Vol. 8, pp.1377-1388 (2002)

Taraszka, J.A., Counterman, A.E. and Clemmer, D.E. "Large anhydrous polyalanine ions: substitution of Na^+ for H^+ destabilizes folded states". *International Journal Mass Spectrometry*, Vol. 204, pp.87-100 (2001)

Teesch, L.M. and Adams, J. "Fragmentations of gas-phase complexes between alkali-metal ions and peptides – metal-ion binding to carbonyl oxygens and other neutral functional-groups". *Journal of the American Chemical Society*, Vol. 113, pp. 812-820, (1991a)

Teesch, L.M., Orlando, R.C. and Adams, J. "Location of the alkali metal ion in gas-phase peptide complexes". *Journal of the American Chemical Society*, Vol. 113, pp. 3668 –3675 (1991b)

Tortajada, A.J., Leon, E., Morizur, J.-P., Luna, A., Mo, O. and Yanez, M. "Potential-energy surface of protonated formamide and of formamide –X(+) (X= Li, Na, Mg and Al) complexes". *Journal of Physical Chemistry*, Vol. 99, pp. 13890-13898 (1995)

Toth, G., Watts, C. R., Murphy, R. F. and Lovas, S. "Significance of aromatic-backbone amide interactions in protein structure". *Proteins: Structure, Function, and Genetics*, Vol. 43, pp.373-381 (2001b).

Trinajstic, N. "Chemical Graphic Theory". In *Chemical Graphic Theory*, CRC Press: Flordia; Vol. I, II. A program for calculating the 3-D Wiener index from Gaussian 98 output can be obtained from one of the author NLM.

Tsang, C.W., Tsang, Y., Wong, C.H.S. and Ma, N.L. *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics (No. TPB 045), May 27-31,Chicago, USA, and manuscript under preparation* (2001b).

Tsang, Y., Siu, F.M., Ma, N.L. and Tsang, C.W. *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics (No. TPB 044), May 27-31, Chicago, USA* (2001a)

Tsang, Y., "Mass Spectrometric Studies on Alkali Metal Cationized Amino Acids and Peptides". *PhD thesis,* (2003)

Tsang, Y., Siu, F.M., Ho, C.S., and Ma, N.L. and Tsang, C.W. "Experimental validation of theoretical potassium and sodium cation affinities of amides by mass spectrometric kinetic method measurements". *Rapid Communications in Mass Spectrometry*, Vol. 18, pp. 345-355 (2004)

Tsuzuki, S., Yoshida, M. Uchimaru, T. and Mikami, M. "The origin of the cation/π interaction: The significant importance of the induction in Li^+ and Na^+ complexes". *Journal of Physical Chemistry A*, Vol. 105, pp.769-773 (2001)

Turecek, F., Kerwin, J.L., Xu, R. and Kramer, K.J. "Distinction of N-substituted histidines by electrospray ionization mass spectrometry". *Journal of Mass Spectrometry*, Vol. 33, pp. 392-396 (1998)

Uggerud, E. "The unimolecular chemistry of protonated glycine and the proton affinity of glycine: a computational model". *Theoretical Chemistry Accounts*, Vol. 97, pp. 313-316 (1997)

Wada, G., Tamura, E., Okina, M. and Nakamura, M. "On the ratio of zwitterion form to uncharged form of glycine at equilibrium in various aqueous-media". *Bulletin of the Chemical Society of Japan*, Vol. 55, pp. 3064-3067 (1982)

Walter, D., Sievers, M.R. and Armentrout, P.B. "Alkali ion carbonyls: sequential bond energies of $Li^+(CO)_x$ (x = 1-3), $Na^+(CO)_x$ (x = 1,2) and $K^+(CO)$". *International Journal of Mass Spectrometry and Ion Processes*, Vol. 175, pp. 93-106 (1998)

Wang, Z., Chu, I.K., Rodruquez, C.F., Hopinson, A.C., and Siu. K.W.M. "α,ω-Diaminoalkanes as models for bases that dicoordinate the proton: an evaluation of the kinetic method for estimating their proton affinities". *Journal of Physical Chemistry A*, Vol. 103, pp. 8700-8705 (1999)

Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., Jr. and Weiner, P. "A new force field for molecular mechanical simulation of nucleic-acids and proteins". *Journal of the American Chemical Society*, Vol. 106, pp. 765-784 (1984)

Williams, P.D., Jockusch, R.A. and Williams, E.R., "Is Arginine a zwitterion in the gas phase?". *Journal of the American Chemical Society*, Vol. 119, pp. 11988-11989 (1997)

Williams, S.M., Brodbelt, J.S., Huang, Z.L., Lai, H.G. and Marchand, A.P. "Complexation of silver and co-recovered metals with novel aza-crown ether macrocycles by electrospray ionization mass spectrometry". *Analyst*, Vol .128 pp.1352-1359 (2003)

Wintjens, R., Lievin, J., Rooman, M. and Buisine, E. "Contribution of caton-pi interactions to the stability of protein-DNA complexes". *Journal of Molecular Biology*, Vol. 302, pp. 395-410 (2000)

Wong, C.H.S., Ma, N.L. and Tsang, C.W. "A theoretical study of potassium cation binding to glycylglycine (GG) and alanylalanine (AA) dipeptides". *Chemistry – A European Journal*, Vol. 8, pp. 4909-4918 (2002a)

Wong, C.H.S., Siu, F.M., Ma, N.L. and Tsang, C.W. "A theoretical study of potassium cation-glycine (K^+-Gly) interactions". *Journal of Molecular Structure (Theochem)*, Vol. 588, pp. 9-16 (2002b)

Wong, C.L. "Mass Spectrometric studies on protonated and alkalimetal cationized amino acids and peptides". *PhD thesis* (2003)

Wouters, J. and Ooms, F. " Small molecule crystallography in drug design". *Current Pharmaceuticla design*, Vol. 7, pp. 529-545 (2001)

Wu, J.Y. and Lebrilla, C.B. "Gas phase basicities and sites of protonation of glycine oligomers (Gly_n, n=1-5) ". *Journal of the American Chemical Society*, Vol. 115, pp. 3270-3275 (1993)

Wu, Z.C. and Fenselau, C. "Gas-phase basicities and proton affinities of lysine and histidine measured from the dissociation of proton-bond dimers". *Rapid Communications in Mass Spectrometry*, Vol. 8, pp. 777-780 (1994)

Wyttenbach, T., Batka Jr., J.J., Gidden, J. and Bowers, M.T. "Host/guest conformations of biological systems: valinomycin/alkali ions". *International Journal of Mass Spectrometry*, Vol. 193, pp. 143-152 (1999b)

Wyttenbach, T., Bushnell, J.E. and Bowers, M.T., "Salt bridge structures in the absence of solvent? The case for oligoglycine". *Journal of the American Chemical Society*, Vol. 120, pp. 5098-5103 (1998)

Wyttenbach, T., Witt, M. and Bowers, M.T. "On the question of salt bridges of cationized amino acids in the gas phase: glycine and arginine". *International Journal of Mass Spectrometry*, Vol. 182/183, pp. 243-252 (1999a)

Wyttenbach, T., Witt, M. and Bowers, M.T. " On the stability of amino acids zwitterions in the gas phase: The influence of derivatization, proton affinity, and alkali ion addition". *Journal of the American Chemical Society*, Vol. 122, pp. 3458-3464 (2000)

Yalcin, T., Khouw, C., Csizmadia, I. G., Peterson, M.R., Harrison, A.G. "Why are B ions stable species in peptide spectra?". *Journal of the American Society for Mass Spectrometry*, Vol. 6, pp. 1164-1174 (1995)

Yates, J.R., "Mass spectrometry and the age of the proteome". *Journal of Mass Spectrometry*, Vol. 33, pp. 1-19 (1998)

Zacharias, N. and Dougherty, D.A. "Cation-π interactions in ligand recognition and catalysis". *Trends in Pharmacological Sciences*, Vol. 23, pp. 281-287 (2002).

Zhang, K., Cassady, C.J. and Chung-Phillips, A. "Ab initio studies of neutral and protonated triglycines – comparison of calculated and experimental gas-phase basicity". *Journal of the American Chemical Society*, Vol. 116, pp. 11512-11521 (1994)

Zhang, K., Zimmerman, D.M., Chung-Phillips, A. and Cassady, C.J. "Experimental and ab-initio studies of the gas-phase basicities of polyglycines". *Journal of the American Chemical Society*, Vol. 115, pp. 10812-10822 (1993)

Zhong, W., Gallivan, J.P., Zhang, Y., Li, L., Lester, H.A. and Dougherty, D.A. "From Ab initio quantum mechanics to molecular neurobiology: a cation-π binding site in the nicotinic receptor". *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 95, pp. 12088-12093 (1998)

Zhou, Y. and MacKinnon, R. " The occupancy of ions in the K^+ selectivity filter: Charge balance and coupling of ion binding to a protein conformational change underlie high conduction rates". *Journal of Molecular Biology*, Vol. 333, pp. 965-975 (2003)

Zhu, W.L., Tan, X.J., Puah, C.M., Gu, J.D., Jiang, H.L., Chen, K.X., Felder, C.E., Silman, I. and Sussman, J.L. "How does ammonium interact with aromatic groups? A

density functional theory (DFT/B3LYP) investigation". *Journal of Physical Chemistry A*, Vol. 104, pp. 9573-9580 (2000)

Zhu, W.L., Tan X.J., Shen, J.H., Luo, X.M., Cheng, F., Mok, P.C., Ji, R.Y., Chen, K.X. and Jiang, H.L., "Differentiation of cation-π bonding from cation-π intermolecular interactions: A quantum chemistry study using density-functional theory and Morokuma decomposition methods". *Journal of Physical Chemistry A*, Vol., 107, pp. 2296-2303, (2003)

Zubay, G. L. *Biochemistry*,2nd.Ed., Macmillan, New York, Dubuque, Iowa: Wm.C. Brown Publishers (1993)

Research Publications

A. Journal Papers

"A Theoretical Study of Potassium Cation-Glycine (K^+-Gly) Interactions". C.H.S. Wong, F.M. Siu, N.L. Ma and C.W. Tsang. Journal of Molecular Structure (Theochem), Vol. 588, pp. 9-16 (2002).

"A Theoretical Study of Potassium Cation Binding to Glycylglycine (GG) and Alanylalanine (AA) Dipeptides". C.H.S. Wong, N.L. Ma and C.W. Tsang. Chemistry-A European Journal, Vol. 8, pp.4909-4918 (2002).

"Absolute Potassium Cation Affinities (PCAs) in the Gas Phase". J.K.-C. Lau, C.H.S. Wong, F.M. Siu, N.L. Ma and C.W. Tsang. Chemistry-A European Journal, Vol. 9, pp. 3383-3396 (2003).

"Potassium Cation Affinity of Aliphatic Amino Acids: a Combined Experimental and Theoretical Study". Y.Tsang, C.H.S. Wong, F.M. Siu, N.L. Ma and C.W. Tsang. Journal of the American Society for Mass Spectrometry. Manuscripts under preparation.

"Fragmentation of the Protonated Histidine in the Gas Phase" C.H.S. Wong, C.L.C. Wong, N.L. Ma and C.W. Tsang. Journal of Physical Chemistry B. Manuscripts under preparation.

B. Conference Proceedings

C.H.S. Wong, N.L. Ma and C.W. Tsang, "Interaction between Potassium Cation (K^+) with Glycylglycine (glygly)". Proceedings of the 223rd ACS Conference, Orlando, USA, April 7 (2002).

C.H.S. Wong, N.L. Ma and C.W. Tsang, "Interaction between Potassium Cation (K^+) with Phenylalanylglycine Dipeptides". The 1st Asian Pacific Conference on Theoretical Computational Chemistry, Institute for Molecular Science, Okazaki, Japan, May 12-15 (2004)

C. Conference Presentation (Abstract)

C.H.S. Wong, N.L. Ma and C.W. Tsang, 'Interaction between Potassium Cation (K^+) with Glycylglycine (GlyGly) and Alanylalanine (AlaAla) Dipeptides". The 9th Symposium on Chemistry Postgraduate Research in Hong Kong, The Chinese University of Hong Kong, pp. AE-16, 23rd March (2002).

C.H.S. Wong, N.L. Ma and C.W. Tsang, 'Cation-π Interactions and the Gas Phase Thermochemistry of Potassium Cation - Glycylphenylalanine (K^+-GF) Interactions". The 10th Symposium on Chemistry Postgraduate Research in Hong Kong, The Hong Kong Baptist University, pp. P-4, 22nd March (2003).

ELSEVIER

Journal of Molecular Structure (Theochem) 585 (2002) 9–16

THEO CHEM

www.elsevier.com/locate/theochem

A theoretical study of potassium cation-glycine (K^+-Gly) interactions

C.H.S. Wong[a,b], F.M. Siu[a,b], N.L. Ma[c,1], C.W. Tsang[a,b,*]

[a]*Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China*

[b]*Institute of Molecular Technology for Drug Discovery and Synthesis (An Area of Excellence of the University Grants Committee), The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China*

[c]*Institute of High Performance Computing, 89C Science Park Drive, #02-11/13, The Rutherford, Singapore, Singapore 118261*

Received 3 December 2001; accepted 22 February 2002

Abstract

The structures and binding affinities of potassium cation (K^+) bound complexes of glycine (Gly) are established using a B3-LYP density functional based energetic protocol 'EP(K^+)'. Ten stable isomers on the potential energy surface have been located and the most stable mode of binding involves a bidentate interaction between the cation with O=C and −OH. The dipole moment of the glycine ligand plays a dominant role in governing the relative stability of binding modes, while the effect of ligand polarizability plays a less important role. We found that the stabilization energies (raw interaction energies) of these complexes can be well approximated by a linear function of the 'dipole interaction parameter' and 'polarizability interaction parameter'. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Metal–ligand interaction; Potassium ion; Glycine; Density functional theory; Electrostatic interaction

1. Introduction

As one of the most abundant alkali metal cations found in living systems, K^+ has numerous and important biochemical functions. It participates in osmotic equilibrium of cells [1,2], stabilization of DNA structures [1], and activation of enzyme functions [3]. The interaction between potassium cation and protein is the molecular basis accounting for the selective transport of potassium cations across cell membranes [2]. The trans-membrane movement of K^+ underlies many fundamental biological processes, including biological (electrical) signaling in the nervous system, generation of rhythmic signals by the heart, and unceasing sifting of toxic solutes in the blood by the kidney [4]. Hence, detailed knowledge of the structural and energetic aspects of the local interaction between K^+ and prototypical amino acid residues is essential for the understanding of these processes.

In order to elucidate the fundamental nature of the interaction between alkali metal cations and various amino acids in the gas phase, a number of experimental and theoretical studies have been conducted [5–10]. Hoyau and Ohanessian performed ab initio theoretical studies on the stability of sodium

* Corresponding author. Address: Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China. Tel.: +852-27665610; fax: +852-23309372.

E-mail address: bctsang@polyu.edu.hk (C.W. Tsang).

[1] Corresponding co-author.

0166-1280/02/$ - see front matter
PII: S0166-1280(02)00083-0

complexes of serine, cysteine and proline [6]. Dunbar and co-workers [7,9] studied the cation–π interactions in sodiated and potassiated phenylalanine, tyrosine, and tryptophan, while Siu et al. further elucidated the possible roles of cation–π interaction in determining the relative stability of charge-solvation (CS) versus zwitterionic (ZW) forms of Li^+/Na^+/K^+-phenylalanine in the gas phase [8,10].

For alkali metal cation bound complexes of aliphatic amino acids, the first comprehensive ab initio molecular orbital study was conducted by Jensen [11] on Li^+/Na^+-glycine (Gly) complexes. This work was later extended by other groups to other metal cations: Li^+/Na^+ [12–14]; K^+/Rb^+/Cs^+ [12]; Cu^+ [13–15]; Ni^+, Be^{2+}/Mg^{2+}/Ca^{2+}/Sr^{2+}/Ba^{2+} [14,16]; Cu^{2+} [14,17] and Zn^{2+} [18]. In these works, the relative affinities of different modes of binding of a given cation were studied, but none of them has attempted to elucidate the role of different types of electrostatic interaction on the relative stabilities of different modes of binding. The objective of our present study is to investigate how ion-dipole and ion-induced dipole interactions govern the stability of a particular mode of binding, which hopefully would provide new insights into the development of force field for metal–protein simulations.

2. Computational details

The interaction between potassium cation and glycine in the gas phase was modeled with the '$EP(K^+)$' method outlined below. This procedure is based on the energetics and structures determined with the hybrid Becke3-Lee-Yang-Parr (B3-LYP) [19] density functionals. The geometries of glycine and K^+–glycine complexes were first optimized at the HF level using the standard 6-31G(d) basis in GAUSSIAN 98 [20], and confirmed to be real minima at this level of theory. Geometries of these stable minima were further refined by the B3-LYP functional with the 6-31G(d) basis set. Using these structures, single point energy calculations were performed at the B3-LYP/6-311+G(3df,2p) level to yield the theoretical affinities (ΔH_0) given in Eq. (1):

$$\Delta H_0 = [(E_{K^+} + E_{Gly}) - E_{K^+\text{–}Gly}] + [ZPE_{Gly} - ZPE_{K^+\text{–}Gly}] \times 0.8929 \quad (1)$$

where E_{K^+}, E_{Gly}, $E_{K^+\text{–}Gly}$ are the electronic energies of the potassium cation, the glycine ligand and the K^+–Gly complex, respectively, calculated at the B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d) level and the 'ZPE' is the zero point energies of the various species calculated at the HF/6-31G(d) level, scaled by 0.8929 [21]. For simplicity, we have abbreviated this protocol as EP(K^+) for 'Energetic Protocol for estimating K^+ binding affinity'.

The EP(K^+) method shares the same target basis level and zero-point correction factor with the sophisticated ab initio G2-type procedure [23]. On comparing the K^+ affinities determined using EP(K^+) and that calculated with the G2(MP2,SVP)-ASC protocol [22,23], we found that the mean-absolute-deviation for the affinities of 13 model organic ligands is only 4 kJ mol^{-1} [24]. Hence, even though our methodology are not directly related to the G2-type methods, it can yield K^+ affinities comparable to that derived from the computationally more expensive ab initio G2(MP2,SVP) method. Furthermore, we have applied EP(K^+) to obtain the theoretical K^+ affinities for the other amino acids [25–27], and found that they are in excellent agreement with the absolute experimental affinities to within 2 kJ mol^{-1}, which is within the estimated experimental uncertainty of ~10 kJ mol^{-1} for the kinetic method measurements [25–26]. Such good quantitative agreement between theory and experiment provides confidence to the quality of our theoretical model employed and the conclusions drawn here.

3. Results and discussion

3.1. Mode of interaction between K^+ and glycine

There are three electron-rich binding sites (NH_2, O=C, and OH) in glycine, and hence, ten complexes with different modes of binding can be formed in principle. Out of these ten modes, seven of them are CS complexes (three monodentate, three bidentate and one tridentate modes) and the remaining three are ZW complexes (two monodentate and one bidentate modes). Our detailed exploration on the K^+–Gly potential energy surface reveals six complexes which differ in the mode of binding (CS1–5, and ZW1), and four other species, CS1′, CS2′, CS5′ and CS5″, which are less stable conformers of CS1, CS2 and CS5. The

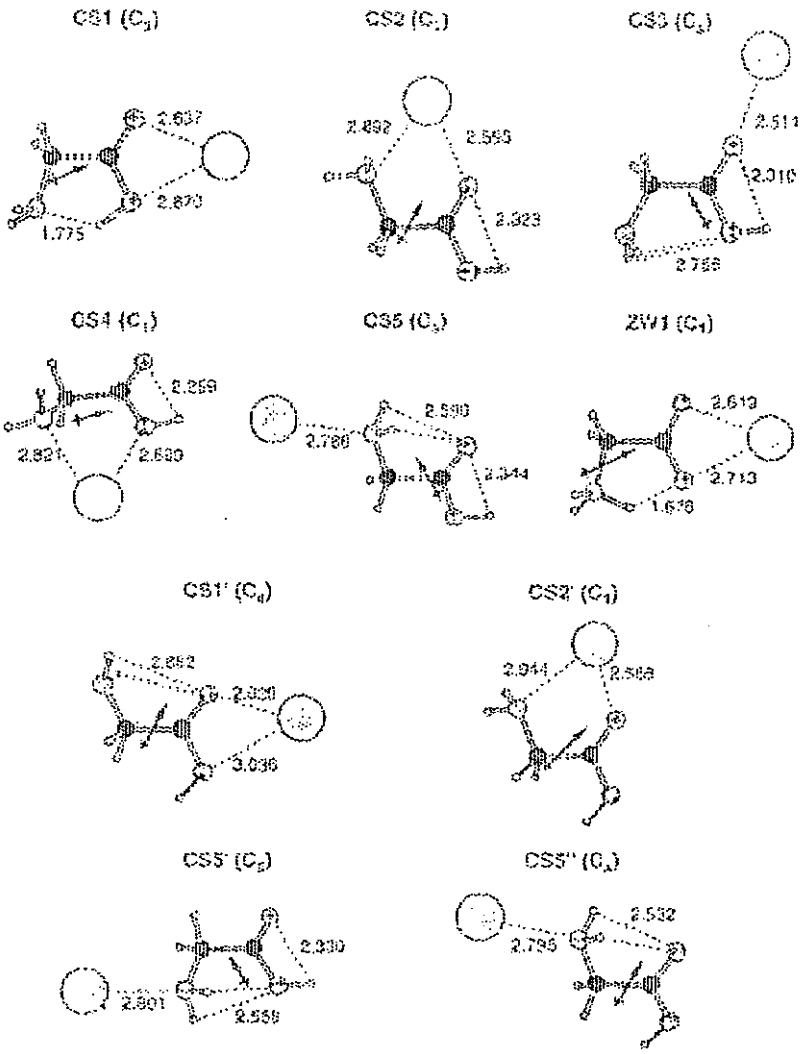


Fig. 1. Geometries of ten stable K^+–Gly complexes, obtained at the B3-LYP/6-31G(d) level of theory. The dipole moment vectors are drawn for reference only and are not to scale.

geometries of these ten K^+–Gly complexes are summarized in Fig. 1.

In agreement with previous findings [12], the tridentate CS mode (O=C, OH, NH_2) is not a stable minimum on the K^+–Gly potential energy surface. Starting with sensible candidates, these structures collapsed to the bidentate CS2 species at the HF/6-31G(d) level. The monodentate CS mode of binding to –OH is also found to be unstable, presumably due to the relatively weak interaction between K^+ and –OH as sensible candidates becomes either CS1 or CS4 upon geometry optimization. Moreover, the ZW complexes where the K^+ binds only to one of the carboxylate (COO^-) oxygens are not stable as these species optimize to the bidentate ZW1 species. This suggests that the intrinsic attraction between K^+ and the formally negative charged carboxylate group is strong enough to overcome the strain of the bidentate binding mode.

Our structures obtained at the B3-LYP/6-31G(d) level for CS1, CS2, CS4, CS5 and ZW1 are comparable with that reported by Rovau and Ohanessian (denoted as 3, 1, 2, 5 and 6, respectively, in Ref. [12] for K^+–Gly at the MP2(full) level of theory

Table 1
Variation of geometrical parameters around the K^+ binding site at different levels of theory for complex CS1′

Level	r_1 (Å)[a]	r_2 (Å)[b]	θ_1 (°)[c]	θ_2 (°)[d]
HF/6-31G(d)	2.591	3.267	136.8	70.8
HF/6-31+G(d)	2.591	4.031	147.3	64.6
HF/6-31G(d,p)	2.586	3.798	138.0	70.0
HF/6-31+G(d,p)	2.588	4.040	148.3	64.2
HF/6-31++G(d,p)	2.588	4.039	148.2	64.2
HF/6-311G(d)	2.568	3.858	141.2	68.0
HF/6-311+G(d)	2.566	4.038	149.3	63.4
HF/6-311+G(d,p)	2.562	4.035	151.7	62.0
HF/6-311++G(d,p)	2.563	4.083	151.5	62.2
B3-LYP/6-31G(d)	2.629	3.036	110.7	87.1
B3-LYP/6-31+G(d)	2.565	4.057	148.9	63.8
B3-LYP/6-31G(d,p)	2.659	3.033	110.3	88.0
B3-LYP/6-31+G(d,p)	2.564	4.043	148.4	64.0
B3-LYP/6-31++G(d,p)	2.659	3.033	110.3	88.0
B3-LYP/6-311G(d)	2.622	2.969	109.0	88.2
B3-LYP/6-311+G(d)	2.624	3.008	110.2	87.8
B3-LYP/6-311+G(d,p)	2.624	3.032	112.5	86.0
B3-LYP/6-311++G(d,p)	3.625	3.038	111.8	86.7
MP2/6-31G(d)[e]	2.680	2.872	104.6	91.5
MP2/6-31+G(d)	2.698	3.888	105.0	91.8
MP2/6-31G(d,p)	2.680	2.883	104.9	91.2
MP2/6-31+G(d,p)	2.697	2.901	105.4	91.5
MP2/6-31++G(d,p)	2.696	2.903	105.4	91.4
MP2/6-311G(d)	2.691	2.839	104.2	91.5
MP2/6-311+G(d)	2.674	2.873	105.1	91.3
MP2/6-311+G(d,p)	2.671	2.924	106.6	90.2
MP2/6-311++G(d,p)	2.672	2.924	106.6	90.2

[a] r_1 is defined as the distance between the K^+ and the carbonyl oxygen (O=C) in glycine.
[b] r_2 is defined as the distance between the K^+ and the hydroxyl oxygen (OH) in glycine.
[c] θ_1 is defined as the angle between K^+, the carbonyl oxygen, and Cl of glycine.
[d] θ_2 is defined as the angle between K^+, the hydroxyl oxygen, and Cl of glycine.
[e] Full core correlation is employed in all MP2 optimizations.

using the 6-31G(d) basis for O, N, and H, and a modified double-zeta basis set for K from Schäfer et al. [28]). One might argue that calculations without diffuse functions on heavy atoms are inadequate for ion–molecule complexes. In order to investigate the validity of our reported B3-LYP/6-31G(d) structures in Fig. 1, and its consequence on the reported energetics, we also obtained the geometries of these K^+–Gly complexes at the B3-LYP/6-31+G(d) level of theory. In terms of geometries, we found that both sets of geometries are similar with the distances between K^+ and various sites of binding shortened somewhat (average of 0.04 Å) with the incorporation of the diffuse function. The only notable exception is observed in the species CS1′.

The species CS1′, a higher energy conformer of CS1, is an interesting species. At the B3-LYP/6-31G(d) level of theory, as shown in Fig. 1, the interaction between K^+ and the two oxygens are quite asymmetric in CS1′ where the binding distance of K^+···O=C is approximately 0.4 Å shorter than that of K^+···OH. However, optimization of this species at the HF/6-31G(d) level of theory yielded an essentially monodentate species, similar to the potassium analogue of species M9 in Ref. [11]. We found that whether CS1′ is predicted to be a mono- or bi-dentate binding mode is, in fact, quite sensitive to the level of theory employed (Table 1). Without electron correlation at the Hartree–Fock (HF) level, K^+ is suggested to be monodentately bound to O=C, independent of the basis set. The result obtained with the hybrid density functional, B3-LYP, shows some interesting basis set dependence. As noted above, the B3-LYP/6-31G(d) level favors a bidentate complex (Fig. 1). However, addition of a diffuse function on heavy atoms (6-31+G(d)) and further addition of a polarization function on the hydrogen (6-31+G(d,p)) favor monodentate complexes. As shown in Table 1, one finds that density functional calculations with the more flexible split-valence triple zeta type basis would lead to a bidentate species for CS1′. This is further supported by more conventional ab initio modeling at the MP2 level: the structure we determined with the MP2 Hamiltonian unanimously pointed towards the bidentate complex, independent of the basis sets.

We attribute the peculiarity of CS1′ to the very soft K^+···O(H)···C bending potential. At the HF/6-31G(d) level, the harmonic frequency of this mode is only 24 cm^{-1}, which is the lowest vibration mode calculated for the species CS1′. With electron correlation at the B3-LYP level, the stiffness of this mode is increased somewhat to 48 cm^{-1}. At the MP2 level, this mode becomes the second lowest vibrational model, with the frequency increases to 82 cm^{-1}. With this generally small bending potential, it is reasonable to expect that only a theoretical model with sufficient electron correlation and flexible basis set could describe this species correctly.

Table 2
Calculated binding affinities at 0 K (ΔH_0, in kJ mol^{-1}) of various K^+-Gly complexes at the B3-LYP/6-311+G(3df,2p) level using geometry obtained at different levels of theory

Species	ΔH_0[a]	ΔH_0[b]	ΔH_0[c]
CS1	117.7	117.5	118.2
CS2	115.2	114.8	115.7
CS3	84.4	84.0	85.0
CS4	75.9	75.4	76.6
CS5	67.2	66.9	67.0
ZW1	104.5	103.8	104.3
CS1′	80.2	82.9	80.0
CS3′	96.4	96.0	96.0
CS5′	55.6	53.2	56.3
CS2″	38.5	38.1	39.2

[a] Using B3-LYP/6-31G(d) optimized geometry (with potassium basis set in Ref. [20]), abbreviated as EP(K^+) in this work.
[b] Using B3-LYP/6-31+G(d) optimized geometry (with potassium basis set in Ref. [20]).
[c] Using B3-LYP/6-31G(d) optimized geometry (with potassium basis set in Ref. [29]).

3.2. Relative binding affinities of various K^+-glycine complexes

In order to study the effect of geometry on the calculated affinities, we further compare the binding affinities obtained at the B3-LYP/6-311+G(3df,2p) level using the B3-LYP/6-31G(d) and B3-LYP/6-31+G(d) geometries, and found them to be virtually identical (Table 2). Even for the apparently problematic species CS1′, the difference in geometry only affects our reported energetics by 2 kJ mol^{-1}. Given the saving of computational time, and the geometry of the CS1′ species might be even more 'reliable' than that obtained at B3-LYP/6-31+G(d) level, the use of B3-LYP/6-31G(d) geometries for the larger K^+-aliphatic amino acids would appear to be justified. We also note in passing that the geometries obtained with the Blaudeau 6-31G(d) basis set [29] is comparable to those obtained here. With this alternative geometry, the binding affinity obtained at the B3-LYP/6-311+G(3df,2p) level again is virtually identical to that calculated with EP(K^+) (within 0.7 kJ mol^{-1} except CS1′). Hence, our results suggest that the choice of geometry is of secondary importance if the final energetics is reported at a sufficiently high level of theory. In other words, the binding affinities obtained at the same level of theory, with minor differences in geometry, could in fact be compared directly.

3.3. Nature of interactions between K^+ and glycine

Even though many theoretical studies on metal cationized glycine complexes have been reported [11–18], the underlying factors affecting the relative affinities of different binding modes are seldom discussed. The binding affinity of metal–ligand complexes is a result of the balance between stabilizing and destabilizing interactions. Qualitatively speaking, the stabilizing interaction is the attractive interaction between the metal cation and the ligand glycine arising from ion-dipole, ion-induced dipole and other electrostatic interactions. In our discussion below, the stabilization energy resulting from attractive electrostatic interactions between the metal cation and the ligand is termed $E_{stabilization}$. On the other hand, in order to accommodate the K^+ at the binding sites, the preferred hydrogen bonding pattern in the free glycine might be disturbed. Such perturbations might lead to less stable conformations of the ligand, which is an important factor in determining the overall strength of the metal–ligand binding affinity, ΔH_0. Here, we called this energy, which destabilizes the ligand in the complexed form relative to its free ligand form, the deformation energy, E_{def}

$$E_{def} = E(\text{glycine in the complex}) - E(\text{glycine in the uncomplexed form}) \quad (2)$$

where E is the electronic energy of the glycine in the two different forms. As the deformed ligand in the complexed state is always less stable than the free ligand, E_{def} is always positive. We found that the E_{def} determined at B3-LYP/6-311+G(3df,2p) and B3-LYP/6-31G(d) levels only differs by less than 1 kJ mol^{-1} in a few test cases. Hence, given the saving in computational time, we estimated the destabilization effect at the B3-LYP/6-31G(d) level.

The raw interaction energy $E_{stabilization}$, which only accounts for the favorable, stabilizing part of the metal–ligand interaction, is thus related to the deformation energy (E_{def}) and binding affinity (ΔH_0) by Eq. (3):

$$E_{stabilization} = \Delta H_0 + E_{def} \quad (3)$$

The analysis on ΔH_0, E_{def} and $E_{stabilization}$ for various K^+–Gly isomers is summarized in Table 3.

Table 3
Binding affinities at 0 K (ΔH_0, in kJ mol^{-1}), stabilization energies ($E_{stabilization}$, in kJ mol^{-1}), deformation energies (E_{def}, in kJ mol^{-1}), DIP, in Debye Å^{-2}, PIP, in Å^{-3}) for K$^+$–Gly conformers optimized at the B3-LYP/6-31G(d) level

Species	Binding mode	ΔH_0[a]	E_{def}[b]	$E_{stabilization}$[c]	DIP	PIP
CS1	O=C, OH	117.7	7.7	125.4	0.3184	0.0256
CS2	O=C, NH_2	115.2	24.1	139.3	0.2141	0.0434
CS3	O=C	84.4	11.4	95.8	0.0869	0.0143
CS4	OH, NH_2	75.9	24.0	99.9	−0.0060	0.0329
CS5	NH_2	62.2	3.9	71.1	0.0697	0.0110
ZW1	COO^-	184.5	86.2	199.7	0.4944	0.0313
CS1′	O=C, OH	80.2	28.4	108.6	0.1907	0.0238
CS2′	O=C, NH_2	96.4	50.7	147.1	0.3104	0.0421
CS3′	NH_2	55.6	9.9	65.5	0.0339	0.0113
CS5″	NH_2	33.5	29.0	67.5	−0.0056	0.0105

[a] Calculated with the EPI(K$^+$) method.
[b] Calculated at the B3-LYP/6-31G(d) level of theory, using Eq. (3). We found that the deformation energy calculated at the B3-LYP/6-311+G(3df,2p) and B3-LYP/6-31G(d) levels using the B3-LYP/6-31G(d) geometries differs by no more than 1 kJ mol^{-1} for a few test cases. Hence, the deformation energies calculated at the B3-LYP/6-31G(d) level are considered to be sufficient.
[c] Calculated using Eq. (5).

The stabilization of the M$^+$–Gly complexes arises primarily from the interactions of the cation with the permanent molecular dipole and the induced dipole of the amino acid. Based on classical electrostatic theory [30], we defined two properties of the amino acid, the dipole interaction parameter (DIP, E_{dip}, (4a)) and polarization interaction parameter (PIP, Eq. (4b)) as:

$$\mathrm{DIP} \propto \mu \cos(\Phi)/r_\mu^2 \tag{4a}$$

$$\mathrm{PIP} \propto \alpha/r_\alpha^4 \tag{4b}$$

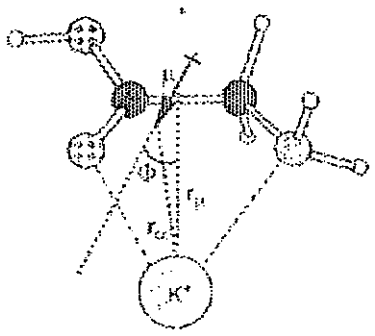


Fig. 2. Definition of parameters used in the calculation of DIP, E_{dip}, (4a) and PIP, Eq. (4b). Here, μ is the permanent molecular dipole moment of the deformed ligand; Φ is the angle of deviation between the metal cation and the dipole moment vector; r_μ is the distance between K$^+$ to the center of the dipole moment vector; and r_α (in Å) is the distance between K$^+$ to the center of charge for the deformed ligand.

where μ (from Mulliken population analysis, in Debye) is the permanent molecular dipole moment of the deformed ligand; Φ is the angle of deviation (in degrees) between the metal cation and the molecular dipole; r_μ (in Å) is the distance between the K$^+$ to the center of the dipole moment vector; α (from Mulliken population analysis, in Å^3) is the molecular polarizability of the deformed ligand, and r_α (in Å) is the distance between K$^+$ to the center of charge of the deformed ligand (Fig. 2). These effective electrostatic properties of the deformed ligand (DIP and PIP) take into account of the orientation of the molecular dipole moment and polarizability with respect to the cation in the deformed glycine via the Φ, r_μ and r_α terms; and are expected to provide a better correlation with the stabilizing energies than μ and α. Both DIP and PIP, calculated at the B3-LYP/6-31G(d) level of theory, are tabulated in Table 3. We note in passing that our analysis is based on classical electrostatic theory with molecular properties determined from Mulliken population analysis (MPA). Even though MPA is known to be sensitive to basis-set effect and often overestimate the absolute magnitude of molecular dipole moments, the relative magnitudes of dipole moment for related species are quite acceptable [31]. Hence, despite these potential pitfalls, our simple approach adopted here should provide a reasonable quantitative understanding for the role of different electrostatics forces.

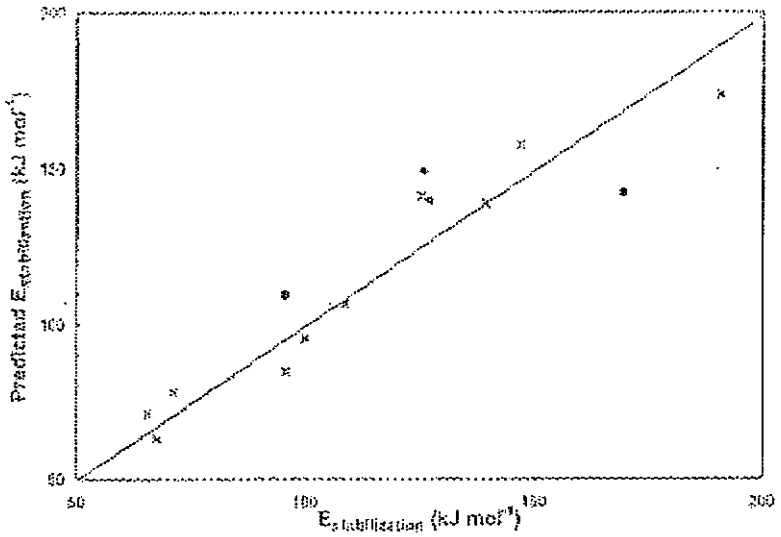


Fig. 3. Plot of predicted stabilization energy (by Eq. (5)) against calculated stabilization energy. The stationary points on the K^+–Gly potential energy surface are designated by × while the non-stationary test points are labeled as ●. The diagonal line with a slope of 1.0 is drawn for reference purposes.

From Table 3, we find that $E_{stabilization}$ in general increases with increasing DIP of the K^+–Gly conformers, suggesting that ion-dipole interaction is a key factor in determining their relative stabilizing energy. Interestingly, in the case of CS4 and CS5[6] (Fig. 1), the DIP is negative (Table 3). This indicates that the positive end of the dipole moment vector is in fact closer to the positively charged K^+ than the negative end of the dipole vector. Hence, the ion-dipole interaction in these cases might be destabilizing. In such cases, the ion-induced-dipole interaction would be an important force holding the K^+ and Gly together. We carried out a multiple linear regression analysis on $E_{stabilization}$ using DIP and PIP as the regression parameters and found that for the ten K^+–Gly complexes, the $E_{stabilization}$ can be described by Eq. (5):

$$E_{stabilization} = 194 \times DIP + 1007 \times PIP + 54 \quad (5)$$

The correlation is very good, with an adjusted $R^2 = 0.92$. Fig. 3 shows the relation between the calculated $E_{stabilization}$ (from Table 3) and the predicted $E_{stabilization}$ using Eq. (5). With our simple model, the maximum deviation of the predicted $E_{stabilization}$ from the calculated value is 17 kJ mol^{-1}, with an average deviation of 8 kJ mol^{-1}. To test the applicability of this equation, we created test cases based on non-stationary points on the K^+–Gly potential surface (indicated by solid circles in Fig. 3). We found that for species with $E_{stabilization}$ within the range of 66 (CS5′) and 191 kJ mol^{-1} (ZW1), the average deviation between calculated and predicted $E_{stabilization}$ is 19 kJ mol^{-1}, suggesting that Eq. (5) can be applied to estimate the raw binding energies within this range. Outside this range, Eq. (5) is apparently less applicable as our test cases yielded an average error of 65 kJ mol^{-1}. No apparent pattern can be found between the magnitude of deviation and the structures of these test cases.

It is interesting to compare our model with that of Gresh et al. as applied to Zn^{2+}–Gly complexes [18]. In their model, the interaction energy (equivalent to the $E_{stabilization}$ here) is partitioned into contributions from ion-dipole, ion-quadrupole, repulsion, polarization, charge-transfer and dispersion interactions. On the other hand, Eq. (5) suggests that if the effect of ion-dipole and polarization is modeled classically via DIP and PIP, a constant term of 54 kJ mol^{-1} would represent the contribution from all the other kinds of interactions. In other words, comparison between the two models would suggest that in the K^+–Gly system, the total contribution from ion-quadrupole, repulsion, charge-transfer and dispersion interactions to the relative stability might be constant, and is independent of the mode of binding. Moreover, in all the K^+–Gly complexes, the natural charge of K^+ is very close

to one (approximately 0.98 C). This suggests that the contribution due to covalent charge-transfer is minimal. Given that ion-quadrupole interaction is significant for cation-aromatic π interactions only [32], it appears that the sum of repulsive and dispersive interactions can be approximated by the constant term of 54 kJ mol^{-1}. From this, we estimated that the contribution of ion-dipole and ion-induced dipole interaction to the relative stability of the various K^+–Gly conformers varies from ~1 to 51% and 16 to 45%, respectively, depending on the geometry of the K^+–Gly complex (Table 3).

4. Conclusions

We have located all the stable stationary points (ten conformers) on the K^+–Gly potential energy surface. By analyzing the stabilization energies (raw interaction energies) in terms of classical molecular dipole moments and polarizabilities, we found that ion-dipole and ion-induced dipole interactions are the key factors in determining the relative stability of different CS and ZW modes of binding in K^+–Gly.

Acknowledgements

NLM thanks the Institute of High Performance Computing and National University of Singapore for generous allocation of supercomputer time. The award of Hong Kong Polytechnic University research studentships GW-021, and GV-540 to CHSW and PMS, respectively, and an Area of Strategic Development Fund (acc. no. A024) to CWT are gratefully acknowledged.

References

[1] S.J. Lippard, J.M. Berg, Principles of Bioinorganic Chemistry, University Science Books, Mill Valley, California, 1994.
[2] L. Stryer (Ed.), Biochemistry W.H. Freeman, New York, 1995 Chapter 12.
[3] M.N. Hughes (Ed.), The Inorganic Chemistry of Biological Processes Wiley, New York, 1972.
[4] C. Miller, Science 261 (1993) 1692.
[5] R.A. Jockusch, W.D. Price, E.R. Williams, J. Phys. Chem. A 103 (1999) 9266.
[6] S. Hoyau, K. Norrman, T.B. McMahon, G. Ohanessian, J. Am. Chem. Soc. 121 (1999) 8864.
[7] V. Ryzhov, R.C. Dunbar, B. Cerda, C. Wesdemiotis, J. Am. Soc. Mass Spectrom. 11 (2000) 1037.
[8] P.M. Siu, N.L. Ma, C.W. Tsang, J. Am. Chem. Soc. 123 (2001) 3393.
[9] R.C. Dunbar, J. Phys. Chem. A 104 (2000) 8067.
[10] P.M. Siu, N.L. Ma, C.W. Tsang, unpublished results.
[11] F. Jensen, J. Am. Chem. Soc. 111 (1992) 9533.
[12] S. Hoyau, G. Ohanessian, Chem. Eur. J. 4 (1998) 1561.
[13] T. Marino, N. Russo, M. Toscano, J. Inorg. Biochem. 79 (2000) 179.
[14] S. Pelikkinen, M. Noguera, L. Rodriguez-Santiago, M. Sodupe, J. Bertran, Chem. Eur. J. 6 (2000) 4393.
[15] S. Hoyau, G. Ohanessian, J. Am. Chem. Soc. 119 (1997) 2016.
[16] E.F. Strittmatter, A.S. Lemoff, E.R. Williams, J. Phys. Chem. A 104 (2000) 9793.
[17] J. Bertrán, L. Rodriguez-Santiago, M. Sodupe, J. Phys. Chem. B 103 (1999) 2310.
[18] F. Rogalewicz, G. Ohanessian, N. Gresh, J. Comput. Chem. 21 (2000) 965.
[19] A.D. Becke, J. Chem. Phys. 98 (1993) 5648.
[20] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.J. Zakrzewski, J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.A. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, GAUSSIAN 98, Revision A.6, Gaussian, Inc., Pittsburgh, PA, 1998.
[21] J.B. Foresman, A.E. Frisch (Eds.), Exploring Chemistry with Electronic Structure Methods Gaussian, Pittsburgh, PA, 1996.
[22] N.L. Ma, P.M. Siu, C.W. Tsang, Chem. Phys. Lett. 322 (2000) 65.
[23] P.M. Siu, N.L. Ma, C.W. Tsang, J. Chem. Phys. 113 (2001) 7045.
[24] P.M. Siu, N.L. Ma, C.W. Tsang, unpublished work.
[25] C.W. Tsang, Y. Tsang, C.H.S. Wong, N.L. Ma, Proceedings of the 48th ASMS Conference on Mass Spectrometry and Allied Topics, May 23–31, Chicago, USA, 2001.
[26] Y. Tsang, C.H.S. Wong, N.L. Ma, C.W. Tsang, unpublished work.
[27] R.G. Cooks, P.S.H. Wong, Acc. Chem. Res. 31 (1998) 379.
[28] A. Schäfer, H. Horn, R. Ahlrichs, J. Chem. Phys. 97 (1992) 2571.
[29] J.-P. Blaudeau, M.P. McGrath, L.A. Curtiss, L. Radom, J. Chem. Phys. 107 (1997) 5016.
[30] J.N. Israelachvili, Intermolecular and Surface Force, Academic Press, London, 1992.
[31] S. Tsuzuki, T. Uchimaru, K. Tanabe, A. Yliniemela, J. Mol. Struct. (Theochem) 565 (1999) 81.
[32] J.H. Williams, Acc. Chem. Res. 26 (1993) 593.

A Theoretical Study of Potassium Cation Binding to Glycylglycine (GG) and Alanylalanine (AA) Dipeptides

Carrie Hoi Shan Wong,[a] Ngai Ling Ma,*[b] and Chun Wai Tsang*[a]

Abstract: By combining Monte Carlo conformational search technique with high-level density functional calculations, the geometry and energetics of K^+ interaction with glycylglycine (GG) and alanylalanine (AA) were obtained for the first time. The most stable K^+-GG and K^+-AA complexes are in the charge-solvated (CS) form with K^+ bound to the carbonyl oxygens of the peptide backbone, and the estimated 0 K binding affinities (ΔH_0) are 152 and 157 kJ mol^{-1}, respectively. The K^+ ion is in close alignment with the molecular dipole moment vector of the bound ligand, that is, electrostatic ion–dipole interaction is the key stabilizing factor in these complexes. Furthermore, the strong ion–dipole interaction between K^+ and the amide carbonyl oxygen atom of the peptide bond is important in determining the relative stabilities of different CS binding modes. The most stable zwitterionic (ZW) complex involves protonation at the amide carbonyl oxygen atom and is approximately 48 kJ mol^{-1} less stable than the most stable CS form. The usefulness of proton affinity (PA) as a criterion for estimating the relative stability of ZW versus CS binding modes is examined. The effect of chain length and the nature of metal cations on cation–dipeptide interactions are discussed. Based on results of this study, the interaction of K^+ with longer peptides consisting of aliphatic amino acids are rationalized.

Keywords: cations · density functional calculations · peptides · potassium · zwitterions

Introduction

As one of the most abundant alkali metal cations in living systems, K^+ has numerous biochemical functions such as osmotic equilibrium of cells,[1–3] stabilization of protein structures,[1] and activation of enzyme functions.[1] Recent studies have shown that K^+ and the other alkali metal cations can induce conformational changes when they bind to proteins.[4–5] The trans-membrane movement of potassium ions underlies many fundamental biological processes, including biological (electrical) signaling in the nervous system, generation of rhythmic signals by the heart, and unceasing sifting of toxic solutes in the blood by the kidney.[6] X-ray crystallographic data suggest that the interaction between K^+ and the main chain carbonyl oxygen atoms of the potassium ion channel (protein) is the molecular basis that accounts for the selective transport of K^+ across cell membranes.[7] Hence, detailed knowledge of the structural and energetic aspects of the local interaction between K^+ and prototypical amino acid residues and peptides is essential for understanding these processes. Such knowledge is of practical importance in interpreting the mass spectra of K^+-peptide/protein complexes, from which sequencing information can be obtained.[8–10]

The interactions of alkali metal cations with simple amino acids are reasonably well-studied theoretically.[11–14] In the gas phase, experimental evidence indicates that binding of K^+ to arginine stabilizes the zwitterionic (ZW) form of the amino acid,[12–14] and the stability of the metal-cationized ZW structure relative to the charge-solvated (CS) structure is postulated to be enhanced by increase in the proton affinity of the amino acid.[12–14] To extrapolate from these small model amino acid systems to metal cation–protein systems, the interaction between metal cation and the simplest dipeptide glycylglycine (GG) and alanylalanine (AA) is a vital bridge. With no functionalized side chain, the interaction between metal cations and GG/AA are restricted to O/N binding sites

[a] Dr. C. W. Tsang, C. H. S. Wong
Department of Applied Biology and Chemical Technology
and Institute of Molecular Technology
for Drug Discovery and Synthesis**
The Hong Kong Polytechnic University, Hung Hom (Hong Kong)
Fax: (+852) 2364-9932
E-mail: bectsan@polyu.edu.hk

[b] Dr. N. L. Ma
Institute of High Performance Computing
1 Science Park Road, No. 01-01, The Capricorn
Singapore Science Park II, 117 528 (Singapore)
Fax: (+65) 6778-0553
E-mail: ida@ihpc.a-star.edu.sg

[**] An Area of Excellence of the University Grants Committee (Hong Kong)

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

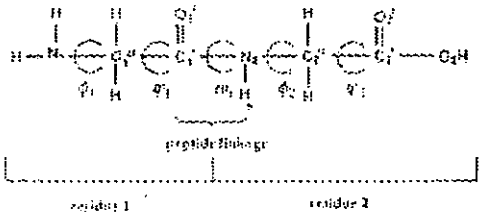
at the amide oxygen of the peptide bond, the amino nitrogen at the N-terminal, and the carboxyl oxygen atoms at the C-terminal, and this provides valuable information on the individual metal-cation-binding sites on the peptide/protein backbone. Because of the large molecular sizes, only three high-level theoretical studies on peptide complexes have been reported, Na^+-GG[20, 21] and Cu^+/Ag^+-GG[22].

To gain a better understanding of how different metal ions interact with the peptide backbone, we carried out DFT studies on K^+-GG and K^+-AA. The Monte Carlo conformational search technique is used to generate plausible conformers for further ab initio/DFT studies. Using this combined approach, we located several modes of metal-cation binding on the GG and AA backbones that were not reported in previous studies. In the case of K^+-GG/K^+-AA complexes, some of these binding modes are fairly low lying and hence may be energetically accessible under laboratory conditions. Factors affecting the relative K^+ affinities of these different ZW versus CS binding modes will be discussed.

Computational Methods

For the free ligand GG, the stability of its various conformers has been the subject of several theoretical publications[7, 8, 20] We re-optimized the four most stable conformers presented in refs. [21, 23] at the HF/6-31G(d) level, and refined their geometries at the B3-LYP/6-31G(d) level. Single-point energy calculations were performed at the B3-LYP/6-311+G(3df,2p) level based on the B3-LYP/6-31G(d) geometries.

There are no prior high level theoretical studies on the interaction of K^+ and GG ligands. The K^+ may interact with the dipeptide ligand in charge-solvated (CS) or zwitterionic (ZW) forms. The zwitterionic form of the ligand can in turn be classified into three types in which the proton is attached to the terminal amino nitrogen atom (N_t), the nitrogen in the peptide linkage (N_l) or the amide carbonyl oxygen atom of the peptide bond (O_l') as depicted in Scheme 1.[24]



Scheme 1. Systematic naming of atoms and dihedral angles of the glycylglycine ligands according to the IUPAC recommendation.[24] Atoms and their positions in the peptide backbone are indicated by lateral symbols/numbers in bold fonts).

Because of the flexibility of the peptide backbone by single bond rotation, the K^+-GG complex can exist as many isomers/conformers. Given the complexity, we first obtained plausible geometries of these species using the Monte Carlo multiple minimum (MCMM) conformational searching technique[25] with AMBER* force field[26] implemented in the Macromodel 5.0 package.[27]

In the MCMM search, the K^+ was not covalently bound to any atoms of glycylglycine so that the ion could move freely to interact with any part of the ligand. Point charges and distant-dependent dielectric constant were used in the electrostatic interaction treatment.[28] Extended cutoff bond lengths of 30 Å were employed for physical van der Waals and electrostatic interactions, and hydrogen-bonding distances. For the interaction of K^+-GG in CS and ZW forms, n_{tot} Monte Carlo steps (where $n_{tot} = 1500 \times$ number of rotatable torsional angles of the GG ligand in CS or ZW form, that is $n_{tot} = 9000$ for CS and ZW(O_l'), and 7500 for ZW(N_t) and ZW(N_l)) were carried out to locate the low energy structures for K^+-GG. Thus more than 100 conformers/isomers were obtained within an energy window of 50 kJ mol^{-1}. Any complexes generated from the MCMM step without intermolecular hydrogen bondings were discarded.

The remaining 60 lowest-energy conformers were re-optimized at the HF/6-31G(d) level by using the Gaussian 98 package.[29] Within each CS/ZW series, only stable isomers/conformers (<80 kJ mol^{-1} above the most stable K^+-GG complex calculated at the HF/6-31G(d) level) were retained for further calculations in which the geometries were refined at the B3-LYP/6-31G(d) level.[30] These structures were used for single-point energy calculations at the B3-LYP/6-311+G(3df,2p) level to yield the theoretical affinities at 0 K (ΔH_0) given in Equation (1):

$$\Delta H_0 = [(E_{K^+} + E_{GG}) - E_{K^+\cdot GG}] + [ZPE_{K^+\cdot GG} - ZPE_{GG}] \times 0.9806^{[31]} \quad (1)$$

where E_{K^+}, E_{GG}, $E_{K^+\cdot GG}$ are the electronic energies (calculated at the B3-LYP/6-311+G(3df,2p)//B3-LYP/6-31G(d) level) of the potassium cation, the glycylglycine ligand (in this case, the energy of conformer GG1; see below) and the K^+-glycylglycine complex, respectively; and ZPE is the zero-point energy of the various species calculated at the HF/6-31G(d) level and scaled by 0.8939.[31] For simplicity of expression, we abbreviate this protocol as "EP(K$^+$)" for "Energetic Protocol for estimating K^+ binding affinity". We calibrated this protocol against the computationally more expensive ab initio G2(MP2,SVP) method for 13 small organic ligands, and found that the K^+ affinities (ΔH_0) are comparable[32] and the mean absolute deviation (MAD) is only 4 kJ mol^{-1}. Furthermore, we have applied EP(K$^+$) to obtain the theoretical K^+ affinities for all five aliphatic amino acids,[33] and found that they are in excellent agreement with the absolute affinities determined by the mass spectrometric kinetic method to within 2 kJ mol^{-1}, which is well within the estimated experimental uncertainty of about 10 kJ mol^{-1}.[32, 34]

The affinities at 0 K (ΔH_0) for all K^+-GG complexes calculated with the EP(K$^+$) protocol are summarized in Table 1. Standard thermodynamic relations[35] were applied to obtain the affinities ΔH_{298} and basicity ΔG_{298} of various modes of binding at 298 K (Table 1). As expected, for a given mode of binding, ΔH_{298} is larger than ΔH_0 (by ca. 1.5 kJ mol^{-1}). Moreover, as entropy is expected to increase when the cation is released from the complexes, ΔG_{298} is smaller than ΔH_{298} (by ca. 31.5 kJ mol^{-1}). Nevertheless, the relative affinity and basicity scales are essentially parallel, and this suggests that entropy effects are not important in determining the preferred interaction of K^+ with the GG ligand.

To understand how metal-cation binding affects the structural and electronic energy of the ligand, we also calculated the deformation energy

Table 1. The theoretical energetics of potassiated glycylglycine (K^+-GG) complexes [kJ mol^{-1}].

Species	Binding site	ΔH_0[a]	ΔH_{298}[b]	ΔG_{298}[b]	E_{def}[c]	$E_{coulombic}$[d]
CS1	O_t', O_l'	151.8	155.0	120.7	31.3	182.9
CS2	O_t', O_t', N_t	145.4	148.2	110.6	43.4	190.7
CS3	O_t', O_t	122.5	123.4	92.6	21.5	144.0
CS4	O_t', O_t, O_l'	106.1	107.9	75.9	60.6	169.7
CS5	O_t', O_t	103.4	104.1	106.5	9.3	146.7
CS6	O_t', N_t	157.2	158.6	116.9	33.3	170.5
CS7	O_t'	127.3	127.8	109.1	6.7	134.0
CS8	N_t	80.2	80.7	54.9	11.0	94.2
ZW(O_t')	carboxylate COO^-	103.8	105.5	77.6	139.5	253.3
ZW(N_t)	carboxylate COO^-	90.4	93.5	53.9	78.5	179.2
ZW(N_l)	carboxylate COO^-	46.1	47.6	15.6	138.1	203.3

[a] Calculated by the EP(K$^+$) protocol. [b] Standard thermodynamic relations[35] were applied to obtain the affinities (ΔH_{298}) and basicity (ΔG_{298}) at 298 K. [c] Calculated at the B3-LYP/6-31G(d) level of theory. We found that the deformation energy calculated at the B3-LYP/6-311+G(3df,2p) and B3-LYP/6-31G(d) levels using the B3-LYP/6-31G(d) geometries differs by no more than 1 kJ mol^{-1} for a few test cases. Hence, the deformation energies calculated at the B3-LYP/6-31G(d) level are considered to be sufficient. [d] Sum of ΔH_0 and E_{def}, representing the raw interaction energy between the cation and the ligand.

0947-6539/03/0921-4910 $ 20.00+.50/0 Chem. Eur. J. 2003, 9, No. 21

E_{int} at the B3-LYP/6-31G(d) level of theory, where E_{int} (listed in Table 1) is given by Equation (2)[9a, 9b]

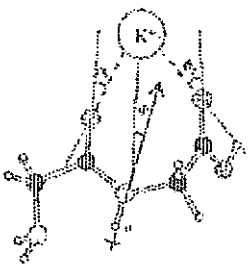
$$E_{int} = E(\text{GG in the K}^+\text{-GG complex}) - E(\text{GG in the uncomplexed form}) \quad (2)$$

Physically, E_{int} represents the destabilization energy arising from structural distortion, disruption of intramolecular hydrogen bonding, and intramolecular electrostatic repulsion among electron-rich functional groups of the ligand when the ligand deforms itself to accommodate the metal cation. As the deformed ligand in the complexed form is always less stable than the free ligand, E_{int} is always positive. The total favorable (enthalpic) interaction energy at 0 K is then given by $E_{complexation}$, which is the sum of ΔH_0 and E_{int}. By replacing one hydrogen atom on each of C_1^a and C_1^b with a methyl group (without performing the MCMM conformation search again), we carried out EP(K$^+$) calculations to obtain ΔH_0 for the K$^+$-AA system at the HF/LYP/6-31[1 + G(3df,2p)//B3LYP/6-31G(d) level. Also, the same computational procedures were applied to obtain ΔH_{298}, ΔG_{298}, E_{int} and $E_{complexation}$ of the K$^+$-AA complexes (Table 2).

Table 2. The theoretical energetics of K$^+$-AA complexes [kJ mol^{-1}].

Species	Binding site	ΔH_0[a]	ΔH_{298}[a]	ΔG_{298}[a]	E_{int}[a]	$E_{complexation}$[a]
CS1	O_1^a, O_1^b	157.2	153.1	127.0	28.8	186.0
CS2	O_1^a, O_1^b, N_1	145.9	147.5	112.5	50.2	196.1
CS3	O_1^a, O_2	138.1	130.8	100.0	19.6	149.7
CS4	O_1^a, O_2, O_1^b	113.1	113.6	83.6	64.9	178.0
CS5	O_2^a, O_2	138.6	142.2	109.7	14.0	154.6
CS6	O_2^a, N_1	135.8	137.1	106.0	40.3	176.1
CS7	O_1^a	123.8	123.1	93.0	6.4	130.3
CS8	N_1	83.7	84.1	58.1	15.3	99.0
ZW(O_1^a)	carboxylate COO^-	119.6	113.2	79.8	140.5	254.1
ZW(N_1)	carboxylate COO^-	104.9	106.4	71.3	79.7	183.8
ZW(N_2)	carboxylate COO^-	59.6	60.5	31.7	148.5	208.1

In the present study, the dipole moment (μ in Debye) of the deformed GG and AA in the complexed states were calculated by standard Mulliken population analysis in the Gaussian 98 package.[9c] Classically, the strength of the ion–dipole interaction is directly proportional to the molecular dipole moment of the *deformed* ligand in the complexed state, the cosine of the angle of deviation between the cation and the dipole moment vector (Φ in ° degrees), and inversely proportional to the square of the distance between the cation and the centre of the dipole moment vector (r_c in Å), with the origin at centre of charge of the deformed ligand (Scheme 2).[9d] The ion–dipole interactions are strongest when the metal ion is in perfect alignment ($\Phi = 0°$) with the dipole-moment vector. The alignment of K$^+$ with the amide O=C, carboxyl O=C and OH bonds at the individual O/N heteroatom binding sites are represented by the angles of deviation θ_1, θ_2, and θ_3, respectively (only θ_1 and θ_2 are shown in Scheme 2).



Scheme 2. Representation of the ion–dipole interactions in the K$^+$-glycylglycine complex in which the angle of deviation between the cation and the dipole moment vector is Φ (in °), and the distance between the cation and the centre of the dipole moment vector is r_c (in Å, with origin at centre of charge of the *deformed* ligand). The alignment of K$^+$ with the O=C bond axis is represented by the angles of deviation θ_1 and θ_2.

Results and Discussion

Glycylglycine ligand: Four glycylglycine conformers, GG1 to GG4 (Figure 1) were investigated. The most stable conformer

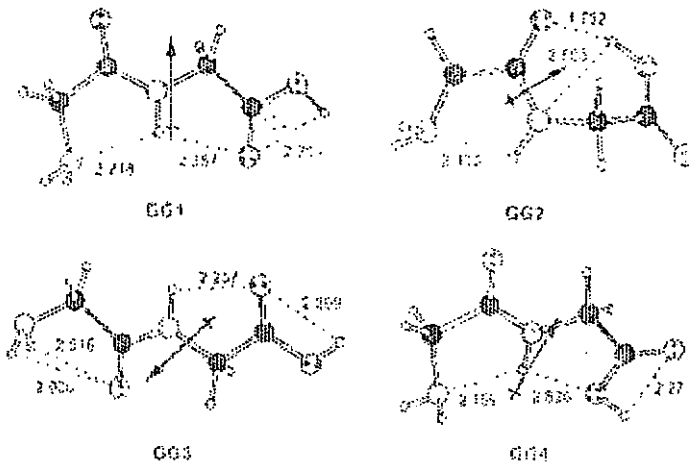


Figure 1. The geometries of four conformers of the glycylglycine (GG) ligand, optimized at the B3-LYP/6-31G(d) level. The intramolecular hydrogen bonds are indicated by dotted lines.

we obtained, GG1, is stabilized by three intramolecular hydrogen bonds.

Our findings are in agreement with Cerda et al. (species X in ref. [21]) but in contrast to those of Cassady et al.[9a] and Siu et al.[9b] Without electron correlation (at the HF/6-31G(d) level), Cassady et al.[9a] suggested that GG3 is most stable, while Siu et al.[9b] reported that conformer GG2 (species 5N in ref. [23]) is the global minimum. At our current level of theory, the energy difference between these three conformers is within 4.0 kJ mol^{-1} and such minor energy differences can easily be the result of using a different theoretical treatment or protocol. Given such small energy differences among these "low-lying" GG conformers, the final theoretical K$^+$ binding affinity of glycylglycine will not be significantly affected even if GG1 is not found to be the most stable conformer at another level of theory.

The most stable K$^+$-GG complex: Using the MCMM and the EP(K$^+$) protocols, we have located 27 complexes (18 in CS forms and nine in ZW forms) within 140 kJ mol^{-1} from the most stable K$^+$-GG complex. These 27 species can be further classified according to their modes of binding into eight CS (Figure 2) and three ZW (Figure 3) forms; the remaining 16 low-lying CS and ZW forms are shown in the Supporting Information, Figure S1. The structures shown in Figure S-1 are less stable although they have the same K$^+$ binding modes as the structures shown in Figures 2 and 3.

We found that the most stable mode of interaction between K$^+$ and GG is one in which the ligand is in the charge-solvated CS form. This mode of binding, denoted CS1 here (Figure 2), involves binding of K$^+$ to the two carbonyl O=C oxygen atoms (one at the peptide amide bond, and one at the carboxyl group) on both amino acid residues. The K$^+$ ion is in very close alignment ($\Phi = 9°$) with the molecular dipole moment vector of *deformed* GG (Scheme 2), and this suggests that the ion–dipole interaction is important in stabilizing the CS1 mode of binding. We note that CS1 was also identified as the most stable binding mode in the Na$^+$-GG complex.[9d]

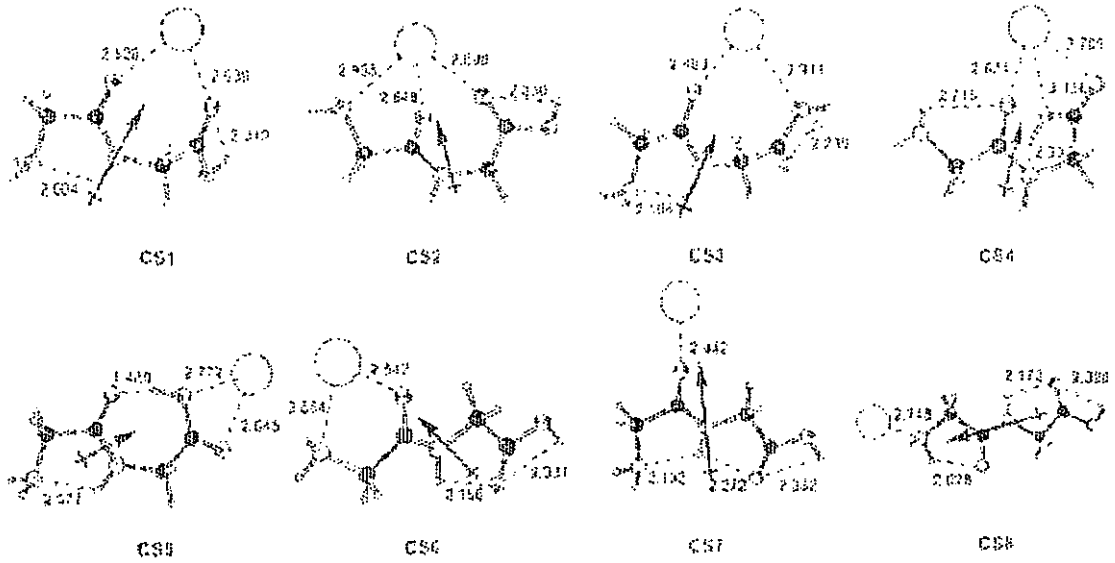


Figure 2. The geometries of eight CS binding modes of K^+-GG, optimized at the B3-LYP/6-31G(d) level. The intramolecular hydrogen bond and the interaction between K^+ and the binding site of the ligand are indicated by dotted lines. The molecular dipole moment vector of the *deformed* GG ligand is indicated by an arrow (not to scale).

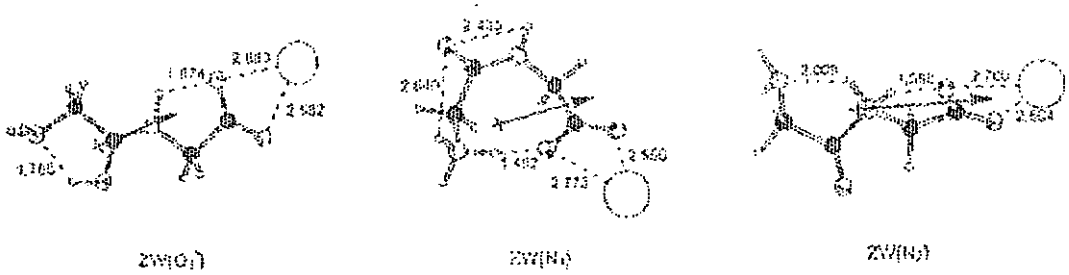


Figure 3. The geometries of three ZW binding modes of K^+-GG, optimized at the B3-LYP/6-31G(d) level. The intramolecular hydrogen bonds and the interaction between K^+ and the binding site of the ligand are indicated by dotted lines. The molecular dipole moment vector of the *deformed* GG ligand is indicated by an arrow (not to scale).

Other charge-solvated (CS) complexes. We found eight complexes (Figure 2) in which K^+ binds to GG in the CS form; the least stable CS isomer CS8 has a binding affinity 71.6 kJ mol^{-1} above the CS1 complex (Table 1). Complexes CS1 – CS8 can be classified into two groups in terms of how K^+ interacts with the GG ligand: those that involve binding to O/N sites of *both* amino acid residues and those that involve binding to only *one* residue. As in the case of CS1, we generally found good alignment of K^+ with the dipole moment vector ($\Phi \approx 7$ to 24°) in the CS2 to CS8 modes of binding.

In CS2, CS3 and CS4 complexes, K^+ is attached to O/N heteroatoms of both residues. One can view these complexes as "derivatives" of the CS1 mode of binding. Relative to CS1, the tridentate CS2 complex is stabilized by an additional interaction between K^+ and the N-terminal amino nitrogen (N_1). While this additional interaction stabilizes the complex, it also leads to greater deformation of the GG ligand (E_{def} is 14.3 kJ mol^{-1} greater than that of CS1), hence decreases the overall stability of the CS2 mode of binding.

The binding affinity of CS3 is comparable to that of CS1: in CS1, K^+ interacts with two carbonyl oxygen atoms, while in CS3, the cation interacts with one O=C and one OH. The interaction of K^+ with the hydroxyl group is known to be weaker than with O=C,[24] and this accounts for the decrease in stability of the CS3 mode relative to CS1.

The CS4 mode differs from the CS1 mode in two aspects. Firstly, the N-terminal amino group adopts a "*cis*" ($\varphi_1 = 29°$), rather than a "*trans*" conformation ($\varphi_1 = -173°$) in CS1. Secondly, K^+ interacts with an additional OH group in the CS4 mode. Although these changes might be considered minor, it is interesting to note that the E_{def} of CS4 is more than twice that estimated for CS1. We attribute this large E_{def} to intraligand repulsion. In the CS4 mode, simultaneous binding of the three negative sites to K^+ requires that the three oxygen atoms be very close to each other. The distance between O_1' and O_2 is only 3.08 Å in CS4 mode, which is 1.41 Å shorter than that in CS1. Since the $K^+ \cdots$OH interaction is weaker, the CS4 mode of binding is less stable than CS1 by 46 kJ mol^{-1}.

Complexes CS5 – CS8 are charge-solvated complexes in which the K^+ interacts with the O/N heteroatom sites of one amino acid residue only. Hence, it is of interest to compare these species with the corresponding modes of binding in K^+-Gly[18] so that the effect of the additional glycyl residue can be

  0947-6539/02/0821-4912 $ 20.00+.50/0 Chem. Eur. J. 2002, 8, No. 21

elucidated. Firstly, the ΔH_0 of K^+-GG is at least 13 kJ mol^{-1} higher than the corresponding modes of binding in K^+-Gly, and the largest difference (43 kJ mol^{-1}) found in CS7.[24] Secondly, the order of relative affinity for these modes for binding (CS5 > CS6 > CS7 > CS8) of K^+-GG is identical to that of K^+-Gly.[25] This implies that for the same mode of binding, the role of the additional spectator glycyl group in GG is simply to enhance the affinity of K^+, presumably due to the increase of permanent dipole moment and polarizability in the presence of the glycyl group. In other words, when the peptide backbone is extended from GG to GGG and longer, and K^+ only binds to one glycine residue, the relative stabilities of the binding modes should have the following order: O=C + OH ≈ O=C + NH_2 > O=C > NH_2.

It is also interesting to compare the stability of CS1 to that of CS5 and CS6. When K^+ binds to small ligands,[24] the raw interaction energy $E_{\text{stabilization}}$ for formamide is especially large: formamide (135) ≫ formic acid (83) ≈ ammonia (74) ≈ water (68 kJ mol^{-1}). This is in line with the much larger theoretical dipole moment of *deformed* formamide (≈4 D) in the complexed state compared with the other three ligands (≈2 D). The binding of K^+ to these small organic ligands can be viewed as model interactions between K^+ and the individual O/N heteroatom binding site in peptides: K^+ binding to the amide C=O oxygen atoms (formamide), carboxyl C=O oxygen atoms (formic acid), N-terminal NH_2 (ammonia) and the C-terminal OH groups (water).[24] Here, we found that the $E_{\text{stabilization}}$ term (Table 1) for these bidentate modes of binding is in the order of: CS1 (amide C=O + carboxyl C=O) > CS6 (amide C=O + N-terminal NH_2) > CS5 (amide C=O + carboxyl OH), in line with the greater $E_{\text{stabilization}}$ derived from binding of K^+ to the amide C=O oxygen atom. The only exception is CS8 (carboxyl C=O and OH), which shows comparable $E_{\text{stabilization}}$ to CS3, presumably because CS8 is stabilized by a particularly strong intramolecular hydrogen bond (≈1.5 Å, Figure 3). Furthermore, the $K^+ \cdots$ O=C interaction is enhanced by better alignment of K^+ with the bond axis of the binding sites (and presumably the "local" dipole moment vector of the C=O binding sites) in CS1 (with angles of deviation $\theta_1 \approx 3°$ and $\theta_2 \approx 47°$ for the amide and carboxyl C=O bonds, respectively) than that in K^+-Gly, in which K^+ binds to the carboxyl C=O and OH oxygen atoms (with angles of deviation $\theta_1 \approx 77$ and $\theta_2 \approx 91°$, respectively).[24] Thus, the stability of CS1 is also related to the strength of the local ion–dipole interaction between K^+ and the carbonyl oxygens, especially binding of K^+ to the amide carbonyl oxygen atom of the peptide bond, which is energetically favored. In fact, the highest $E_{\text{stabilization}}$ values (Table 1) are found for the four CS forms (CS1, CS2, CS4 and CS6) which involve binding of K^+ to the amide carbonyl oxygen atom of the dipeptide (Figure 2). Our findings are in line with previous postulates that $M^+ \cdots$ O=C ion–dipole interactions (M = Na or K) are important sources of attractive interaction that contribute to the stability of M^+-formamide/acetamide and Na^+-GG complexes.[5, 6]

Zwitterionic (ZW) modes of binding: We have identified three zwitterionic (ZW) forms of the K^+-GG complex in which the K^+ interacts in a bidentate fashion with the two carboxylate oxygen atoms COO^-, but differ in the site of attachment of the carboxyl proton (Figure 3): i) at the amide carbonyl oxygen atom (O_1') of the peptide bond, ii) at the N-terminal amino nitrogen (N_1), or iii) at the amide nitrogen (N_2) of the peptide linkage.

The binding affinities of the ZW complexes in K^+-GG system are at least 48 kJ mol^{-1} lower than the CS1 mode. It is interesting to compare this difference with our previous study of the K^+-Gly system.[24] In the case of K^+-Gly, the lowest energy ZW mode is only 13 kJ mol^{-1} less stable compared to the most stable CS complex.[24] As the site of K^+ binding is identical (at the carboxylate COO^-) in both K^+-Gly and K^+-GG, the relative instability of ZW modes of binding in K^+-GG arises from the more unfavorable (greater) charge separation (additional coulombic energy required to maintain the separation of the *positive* proton charge and the *negative* carboxylate charge) in the zwitterionic dipeptide backbone. A similar conclusion has been drawn in the corresponding Na^+ system.[22]

Previously studies[24, 25, 43–45] suggested that, when an external proton is attached to GG and tripeptide GGG, the relative stability and basicity of different sites is in the order of: amino nitrogen N_1 ≈ amide carbonyl O_1' ≫ peptide amide N_2.

In the case of ZW K^+-GG complexes, the preference for intramolecular proton transfer from the carboxyl acid group to the three basic sites is in the order (Table 1): amide carbonyl O_1' > amino nitrogen N_1 ≫ peptide amide N_2. Thus, it can be concluded that on complexation with K^+ the amide nitrogen (N_2) remains the least favorable site of protonation. This could be attributed to the loss of resonance stabilization of the peptide bond after protonation at the amide nitrogen atom[43] which destabilizes the ZW(N_2) structure. Hence, the ZW(N_2) complex is approximately 44 kJ mol^{-1} less stable than ZW(N_1), which has the protonation site at the N-terminal amino nitrogen N_1.

For the GG ligand, it has been estimated that in terms of proton affinity at 0 K (ΔH_0), protonation at the N-terminal amino site N_1 is only favored by 3.7 kJ mol^{-1} over that at the amide carbonyl oxygen site O_1' (Table IV of Ref. [25]). Recent high-level DFT calculations on the GGG ligand also indicate that the proton affinity (ΔH_0) of the N-terminal amide carbonyl oxygen atom is only marginally smaller (by 0.9 kJ mol^{-1}) than that of the amino site (Supplementary information of ref. [45]). In both cases, protonation at the N-terminal amide C=O group is stabilized by internal hydrogen bonding between the additional proton and the amino nitrogen NH_2 atom; this leads to very similar proton affinities for these two proton-binding sites at the N-terminal glycyl residue. However, in the case of the zwitterionic K^+-GG complexes, protonation at the N-terminal amide carbonyl oxygen site O_1' is preferred by 13.4 kJ mol^{-1} (in terms of ΔH_0, Table 1) over protonation at the N-terminal amino site N_1. We attribute the greater stability of the ZW(O_1') binding mode of K^+-GG to three factors: i) the enhanced resonance stabilization (partial double bond character) in the peptide linkage O_1'-C_1'-N_2-H, ii) the formation of a stable hydrogen-bonding interaction $O_1'H^+ \cdots N_1$ (1.77 Å), analogous to the "internal proton solvation" found at the N-terminal glycyl residue of protonated GGG; and iii) the better alignment of the

molecular dipole moment with K^+ in the $ZW(O_1^-)$ mode of binding (θ of 15° in $ZW(O_1^-)$ versus 39° in $ZW(N_1)$).

The ZW complex with protonation site at the N-terminal N_1, $ZW(N_1)$, deserves further attention. Its deformation energy is ≈40% smaller than that of the $ZW(O_1^-)$ and $ZW(N_2)$, and quite comparable to that found in some CS modes of binding (e.g. CS4). The relatively small E_{def} in $ZW(N_1)$ probably arises from an extraordinarily strong hydrogen bond between the hydrogen of the positively charged amino group and the oxygen of the negatively charged carboxylate group ($N_1H^+\cdots{}^-OOC$ 1.49 Å). Such strong hydrogen bond could even compensate the two destabilizing factors arising from N_1 protonation: the preference for a more stable "*cis*" conformation in the peptide bond,[29] and the charge separation in a ZW structure. This interaction is so stabilizing that upon metal complexation, the dipeptide is driven into a compact "cyclic" configuration in $ZW(N_1)$ as opposed to an "extended" conformation found in the other zwitterionic K^+-GG complexes (Figure 2).

Nature of cation on metal cation-glycylglycine (M^+-GG) interactions: Two high-level theoretical studies on Na^+-GG were independently reported by Bowers et al.[24] and Cerda et al.[25] independently. Bowers et al.[24] reported the relative energies of three Na^+-GG isomers, which are also included in a more comprehensive study on Na^+-GG interaction by Cerda et al., with qualitatively the same results.[25] Thus, our comparison is only against the results of Cerda et al. More recently, the interaction between GG and Cu^+/Ag^+ was also reported.[26] As Cu^+ is a substantially smaller cation than K^+, and the similarities and differences between Cu^+-GG and Ag^+-GG interactions have been extensively discussed,[26] we focus only on the interaction of GG with K^+, as opposed to Na^+ and Ag^+ here.

By geometry search we have located some fairly stable modes of binding (e.g. CS3, CS4, CS5, CS7, CS8, $ZW(O_1^-)$ and $ZW(N_2)$) previously not reported in the Na^+-, Cu^+-, and Ag^+-GG systems.[24–26] Certain modes of binding previously reported (e.g. species V in ref. [21]) are less stable conformers on the K^+-GG potential energy surface that share the same mode of cation binding in Figures 2 and 3 here, and the structure and energetics of these low-lying conformers are summarized in the Supporting Information (Figure S-1). We note in passing that the relative stabilities of GG conformers sharing the same mode of binding are in fact governed by the hydrogen-bonding patterns adopted by the GG ligand. The results on this aspect of K^+-GG binding will be reported elsewhere.[27]

Here, we focus on how the nature of the metal cation affects the relative affinity of various modes of binding (Figure 4). Relative to Na^+-GG[24] the relative affinity scale of K^+-GG is compressed. Most of the analogues of Na^+-GG isomers are much less stable relative to the CS1 mode of binding (species I in ref. [21]), except for CS2 (species I in ref. [21]) which is of comparable stability (0.4 kJ mol^{-1}) to that of CS1.

On the other hand, the K^+-GG relative affinity scale is expanded compared to that of the Ag^+-GG system (Figure 4). Moreover, the most stable mode of binding also differs. In the most stable charge-solvated Ag^+-GG conformer (species 3 in ref. [23], corresponding to our CS6), Ag^+ interacts with O_1' and N_1, while K^+ prefers to bind to O_1' and O_2' of the GG ligand (CS1 mode of binding). One can rationalize this difference in terms of the hard-soft-acid-base (HSAB) principle.[28] The silver cation is a softer acid (hence more

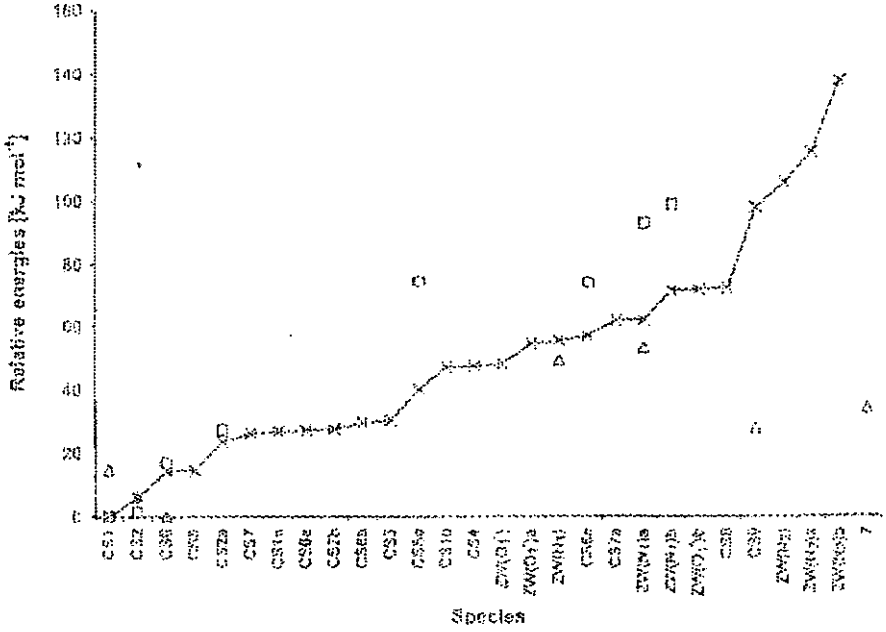


Figure 4. The relative energies (at 0 K) of different modes of binding of glycylglycine for various cations: K^+ (this work at the ED/K$^+$) level, ×, with connecting lines for ease of visualization), Na^+ (ref. [21] at the HF/6-31G(d) level, indicated by ○) and Ag^+ (ref. [23] at the B3LYP/D/ZVP, indicated by △). The effect of zero-point energies are not included in the comparison as it was not reported in ref. [21]. Species 7 in ref. [20] for the Ag^+-GG could not be located on the K^+-GG potential energy surface here.

  0947-6539/02/0821-4914 $ 20.00+.50/0 Chem. Eur. J. 2002, 8, No. 21

polarizable) than Na^+/K^+. Therefore Ag^+ is a better electron acceptor and prefers to bind to the softer nitrogen binding sites than the harder oxygen donor sites of a ligand. The preferred binding of Ag^+ to nitrogen sites is also exemplified in the CS3 mode of binding. As the CS2 binding mode is already of comparable stability to the CS1 binding mode in the case of Na^+, the preference of Ag^+ for the softer nitrogen site further stabilizes the CS2 mode of binding. A further example can be found in species 6 in ref. [23] in which the Ag^+ binds to carboxyl C=O and N-terminal NH_2. Relative to the corresponding CS1 mode, this mode of binding is relatively stable in Ag^+-GG (≈15 kJ mol^{-1}), but very unstable (the CS9 binding mode, ≈100 kJ mol^{-1}, see Supporting Information Figure S-1) in K^+-GG.[24]

We have failed to locate the corresponding zwitterionic species 7 from ref. [23] on the K^+-GG potential energy surface. Starting from sensible trial K^+-GG structures, these complexes invariably optimized to CS5 with K^+ bound to carboxyl oxygen atoms (Figure 2) at both HF/6-31G(d) and B3-LYP/6-31G(d) levels. Species 7 (ref. [23]) and the CS5 complex presented here are isomers that differ in the site of protonation: CS5 is charge-solvated in nature, while species 7 is a zwitterion in which the carboxyl proton has been transferred from the C-terminal to the N-terminal amide oxygen atom O_1'. We carried out additional geometry optimization with larger 6-31+G(d), 6-31G(d,p) and 6-31+G(d,p) basis sets with the B3-LYP function. As these more flexible basis sets also failed to yield stable complex similar to that of species 7[23] found in the Ag^+-GG system, it appears that such mode of binding is in fact unstable on the K^+-GG potential energy surface.

Examination of species 7[23] reveals that the proton bridges between the carboxylate oxygen atom and amide oxygen atom O_1' are short (at 1.40 and 1.07 Å, respectively),[23] and hence the proton would be expected to be quite mobile between these two alternative protonation sites. While cation binding could stabilize a ZW complex (through strong interactions between positively charged metal cations and the negatively charged carboxylate COO^- group), it also introduces instability into the ligand because of the charge-separation effect. It appears that as the smaller Ag^+ binds more strongly and closely to the ligand, the stabilization factor outweighs the charge-separation destabilization effect. As the interaction between the larger K^+ and GG is generally weaker, the stabilizing effect of the $K^+ \cdots CO(O^-)$ interaction is not strong enough to overcome the instability arising from the charge separation effect.

Effect of alkyl side chain and proton affinity: The modes of binding in K^+-AA is very similar to that of K^+-GG, with the more stable CS and ZW modes of binding depicted in Figure 5. Generally speaking, the relative affinities of analogous binding modes in K^+-AA is parallel to that of K^+-GG.

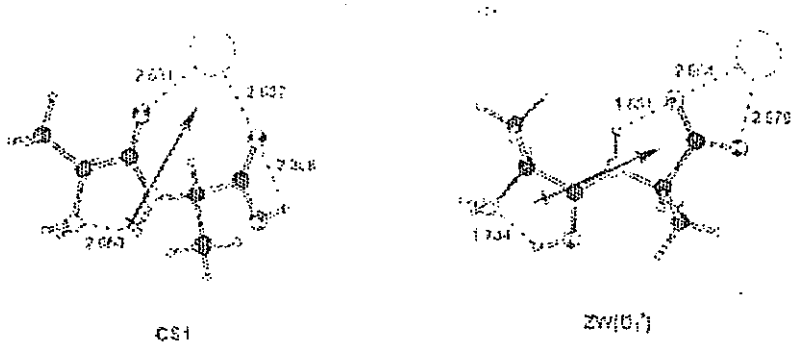


Figure 5. The geometries of the most stable conformer of alanylalanine ligand (AA), charge-solvated complex (CS1), zwitterionic forms (protonated at O_1', ZW(O_1')), optimized at B3-LYP/6-31G(d) level of theory. The intramolecular hydrogen bonds and the interaction between K^+ and the binding sites of the ligand are indicated by dotted lines. The molecular dipole moment vector of the *deformed* AA ligand is indicated by an arrow (not to scale).

Presumably, because of the increase in polarizability, the raw interaction energy ($E_{\text{raw interaction}}$) is increased; this leads to a general increase in K^+ binding affinities of approximately 5 kJ mol^{-1} from GG to AA, which is identical to the change observed in going from K^+-Gly to K^+-Ala.[16] Hence, it seems that the effect of increasing the alkyl chain length (polarizability) on relative stability of CS versus ZW in K^+ binding modes is the same for both aliphatic amino acids and dipeptides.

A greater proton affinity (PA) would favor intramolecular proton transfer from the C-terminal carboxyl OH group to the N-terminal NH_2 group, and is expected to confer greater stability to ZW forms of the M^+ · amino acid complex.[15, 20, 21] Here, we would like to investigate how this criterion can be extended to metal complexes of the dipeptides GG and AA. The proton affinities of glycine, glycylglycine, alanine and alanylalanine are summarized in Table 3, along with the relative stability of the CS/ZW forms in these K^+–ligand systems.

Table 3. A comparison of proton affinities [kJ mol^{-1}] of glycine (Gly), glycylglycine (GG), alanine (Ala), alanylalanine (AA) and the relative stabilities (E_{rel}, in kJ mol^{-1}) of K^+ bound CS/ZW forms

Species	Gly	GG	Ala	AA
PA[a]	902.5	934.7	912.5	946.8
E_{rel}[b]	13.2	48.0	8.0	46.6

[a] Experimental data at 298 K and 1 atm from ref. [20]. [b] The ΔE_0 of the most stable K^+–ligand complex in the ZW form, relative to the most stable CS complex of the same system.

The PA of Ala is greater than that of Gly by 10 kJ mol^{-1}, and this suggests that the more basic N-terminal site in Ala is more prone to proton attachment. Accordingly, formation of the ZW complex is more favorable for Ala than Gly. A similar trend is observed when AA is compared against GG: the

energy difference between the most stable CS and ZW forms is smaller for K^+-AA. Hence, the proton affinity affects the relative stabilities of CS and ZW complexes of aliphatic amino acids and dipeptides in the same way.

However, this does not hold for K^+-Gly/K^+-GG or K^+-Ala/K^+-AA pairs. While GG and AA have greater proton affinities than Gly and Ala, the most stable ZW complex of the dipeptide ligand is much less stable than the most stable CS1 complexes (by 47–48 kJ mol^{-1}, Table 3), a difference significantly greater than the corresponding 8–15 kJ mol^{-1} for K^+-Gly and K^+-Ala complexes.[16] This is because the most stable CS binding mode in the aliphatic amino acids[16] differs from that of the dipeptides. Moreover, the most stable ZW complex for aliphatic amino acids[16] also has a different protonation site to that of the dipeptides. Thus, the PA criterion can only be applied to systems that have the same CS or ZW modes of metal-cation binding, and their corresponding ZW forms should have the same protonation sites.

Interaction of K^+ with peptide backbones: Comparing our present results on the factors governing the relative stability of the CS and ZW modes of binding in K^+-GG/K^+-AA with the intrinsic K^+–peptide backbone interactions in the gas phase[1,4,23] and biological systems[1,5,6,24] provided new insights into the interactions of these and related systems.

Results from the present study illustrate the importance of "local" $K^+ \cdots O{=}C$ ion–dipole interactions between K^+ and GG/AA, so that the cation prefers to bind to two carbonyl oxygen atoms of the dipeptide backbone in the CS mode. Therefore, in longer peptides, the increased flexibility of the backbone should allow the ligand to align its various O=C groups more closely with the cation and this maximize the number of $K^+ \cdots O{=}C$ interactions. At the same time, the ZW binding modes of K^+ become much less stable (by at least about 47 kJ mol^{-1}) in the aliphatic dipeptides GG and AA than in the amino acids glycine and alanine. This is due to i) the enhanced stability conferred by the strong $K^+ \cdots O{=}C$ interaction associated with the peptide bond in the most stable CS1 structure, and ii) the instability of the ZW structures that arise from the greater charge-separation effect in the dipeptide backbone.

Both factors suggest that the K^+ is likely to be encapsulated inside the peptide chain in a macrocyclic CS conformation with multidentate binding to mostly amide C=O sites of the peptide backbone. Support for this is provided by the potassium complex of valinomycin, a dodecadepsipeptide in which the K^+ binds to the six valine carbonyl oxygen atoms in a near octahedral arrangement in the gas phase.[25] Moreover, given the similarity between the mode of binding between Na^+ and K^+ shown here for GG and AA, and previous works for smaller ligands,[16] our results are consistent with previous reports that macrocyclic CS modes of Na^+ binding to backbone carbonyl oxygen atoms in sodium oligoglycine complexes (Na^+-Gly_n, $n = 2$–6) and oligoalanines (Na^+-Ala_n, $n \approx 10$, 15, 20, and $[Ala_n + 3Na]^{3+}$, $n \approx 18$–36).[2, 3, 23] In biological systems, K^+ binding to backbone carbonyl oxygen atoms is a macrocyclic pattern are found in the X-ray protein structures of K^+ channels,[5] tryptophanase[6] and pyruvate kinase,[5] and suggest that the importance of local $K^+ \cdots O{=}C$ ion–dipole interaction can be extrapolated from the gas phase to solution phase.

Finally, we found that the most stable ZW binding mode ZW(O_1') for K^+-GG and K^+-AA is protonated at the amide carbonyl oxygen atom O_1' of the N-terminal glycyl residue (Scheme 1), which has proton affinity very similar to that of the amino N_1 site. Despite stabilization by very strong hydrogen bonding (as indicated by the very short bond length of 1.77 Å) between the $O_1'H$ and N_1 within the N-terminal glycyl residue, this protonated ZW(O_1') structure of GG and AA remains much less stable than the most stable conformer CS1. For longer aliphatic peptides, the ZW conformers are expected to become even less stable due to the greater charge-separation effect. Drawing on the similarity in binding modes between K^+ and Na^+ again, our results on the relative instability of ZW structures ZW(O_1') and ZW(N_1) for the dipeptides GG and AA are in line with a previous report that a longer helical ZW Na^+-Ala_{15} structure collapsed to a random globular structure (presumably in the CS form) in numerical simulations.[11] Furthermore, a helix with a salt bridge from the deprotonated C-terminus (zwitterionic $CO_2^- \cdots Na^+$) and protonation of the backbone C=O near to the C-terminus (to minimize charge-separation effect) appears to be a stable structure in the numerical simulation. Hence, the larger aliphatic peptides are most likely to remain in the CS form when sodiated or potassiated.

Conclusion

To our knowledge, this is the first high-level ab initio/density functional study on K^+ interaction with dipeptides. In this study, we have located 18 charge-solvated (CS) and nine zwitterionic (ZW) stable isomers/conformers for the K^+-GG/AA dipeptide complexes at the B3-LYP/6-311+G(3*df*,2*p*)//B3-LYP/6-31G(d) (abbreviated as EP(K^+)) level of calculations, which can be classified into eight CS and three ZW K^+ binding modes to different O/N heteroatom sites of the peptide. Several of these binding modes are not found in previous studies of Na^+, Cu^+ and Ag^+ binding to the dipeptide GG.

The most stable K^+-GG and K^+-AA complex involve a bidentate interaction in which the K^+ coordinates to two carbonyl oxygen atoms of two amino acid residues in the CS form. We found good general alignment of K^+ with the dipole moment vector of the *unperturbed (deformed)* aliphatic dipeptides in all of the CS modes of binding, and this suggests that ion–dipole interaction is the key electrostatic interaction contributing to the stability of the K^+-GG/AA complexes. Among the different O/N heteroatom binding sites on the peptide backbone, K^+ binding to the amide carbonyl oxygen is energetically very much preferred, and this is attributed to the very strong *local* ion-dipole interaction between K^+ and the peptide amide C=O bond. Consequently, the more stable CS forms are those that involve K^+ binding to and in close alignment with the amide C=O.

Since the most stable ZW form (with K^+ binding to two carboxylate oxygens) is 48 kJ mol^{-1} less stable than the most stable CS form (with K^+ binding to two amide carbonyl

0947-6539/02/0821-4916 $ 20.00+.50/0 Chem. Eur. J. 2002, 8, No. 21

oxygens), it appears that the $K^+ \cdots C{=}O$ (amide) interaction can confer greater stability to K^+-GG/AA complex than $K^+ \cdots COO^-$ (carboxylate) interaction. By definition, in ZW K^+–peptide structures K^+ is bound to the two carboxylate oxygen atoms at the C-terminus. While K^+ binding to backbone amide carbonyl oxygen atoms in the ZW structures of larger helical or globular peptide chains is still possible, the CS form always has the advantage of being able to bind to more amide C=O binding sites than the ZW form, and the former thus gains additional stability. Protonation and internal proton solvation at the basic amide carbonyl oxygen atom of the N-terminal glycyl/alanyl residue in GG and AA cannot compensate for the destabilizing charge-separation effect even for the lowest energy (most stable) $ZW(O_t^-)$ structure involving protonation and strong hydrogen bonding at the more basic N-terminal glycyl/alanyl residue. Since the charge-separation effect is expected to be amplified in ZW structures of longer peptides (due to greater separation distances between the postive proton charge and the negative carboxylate charge), and coupling with the lesser probability of K^+ binding to amide C=O of the peptide backbone for ZW structures, it is very likely that the most stable K^+ stabilized aliphatic peptide complexes have charge-solvated conformations in the gas phase.

The stability of the lowest energy ZW binding mode of GG/AA (relative to their respective most stable CS forms) increases slightly with the proton affinity of the dipeptide. However, the proton affinity criterion fails when cross comparisons are made between the dipeptides and the aliphatic amino acids (i.e., compare GG/Gly pair or AA/Ala pair). Hence, we would suggest caution in extending the proton-affinity criterion to compare or predict the relative stability of different ZW structures of amino acids or peptides having different metal cation binding modes or sites of proton attachment.

While the CS and ZW binding modes and the trend of relative stabilities are similar for Na^+/K^+–dipeptide complexes, we found the most stable CS form of K^+-GG and Ag^+-GG complex to have different O/N heteroatom binding sites, and other stable CS conformers also show significant differences in their relative stabilities. The origin of these differences may reflect the different nature of bonding and ionic sizes of Ag^+ and K^+. Hence, differences in the mass spectral fragmentation patterns between Ag^+–peptides and K^+/Na^+–peptides are expected. Mass spectra of Na^+-,[9,10,54] K^+-[9,10] and Ag^+-[10] cationized peptides have been used to identify peptides and provide sequence information. It may be of practical interest to examine whether such differences can be exploited to provide complementary mass spectral information on peptide sequence, which has become an important issue in proteomic analysis today.

Acknowledgement

N.L.M. thanks the Institute of High Performance Computing and National University of Singapore for generous allocation of supercomputer time. The funding support of the Hong Kong Polytechnic University (Project No. G-W021 to C.H.S.W., and Area of Strategic Development Fund Project No. A024 to C.W.T.), and the Research Grant Council of Hong Kong (Area of Excellence Project No. P-10/2001 and CERG Project No. 5103/01P to C.W.T.) are gratefully acknowledged.

[1] S. J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, California, 1994.
[2] L. Stryer, *Biochemistry*, W. H. Freeman, New York, 1995.
[3] M. N. Hughes, *The Inorganic Chemistry of Biological Processes*, Wiley, New York, 1972.
[4] M. Kohtani, B. S. Kinnear, M. F. Jarrold, *J. Am. Chem. Soc.* 2000, *122*, 12377.
[5] J. A. Tanaka, A. B. Countermen, D. E. Clemmer, *Int. J. Mass Spectrom.* 2001, *204*, 87.
[6] C. Miller, *Science* 1993, *261*, 1692.
[7] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, R. MacKinnon, *Science* 1998, *280*, 69.
[8] R. D. Cody, J. J. Amster, E. W. McLafferty, *Proc. Natl. Acad. Sci. USA* 1985, *82*, 6367.
[9] X. Tang, W. Werner Ens, K. G. Standing, J. B. Westmore, *Anal. Chem.* 1988, *60*, 1791.
[10] R. P. Grese, R. L. Cerny, M. L. Gross, *J. Am. Chem. Soc.* 1989, *111*, 2835.
[11] L. M. Teesch, J. Adams, *J. Am. Chem. Soc.* 1991, *113*, 832.
[12] F. Jensen, *J. Am. Chem. Soc.* 1992, *114*, 9533.
[13] S. Hoyau, G. Ohanessian, *Chem. Eur. J.* 1998, *4*, 1561.
[14] T. Marino, N. Russo, M. Toscano, *J. Inorg. Biochem.* 2000, *79*, 179.
[15] S. Pelikkinen, M. Noguera, L. Rodriguez-Santiago, M. Sodupe, J. Bertran, *Chem. Eur. J.* 2000, *6*, 4393.
[16] C. H. S. Wong, P. M. Siu, S. L. Ma, C. W. Tsang, *J. Mol. Struct. (THEOCHEM)* 2002, *588*, 9.
[17] T. Wyttenbach, M. Witt, M. T. Bowers, *Int. J. Mass Spectrom.* 1999, *182/183*, 243.
[18] R. A. Jockusch, W. D. Price, E. R. Williams, *J. Phys. Chem. A* 1999, *103*, 9266.
[19] B. A. Cerda, C. Wesdemiotis, *Analyst* 2000, *125*, 637.
[20] T. Wyttenbach, M. Witt, M. T. Bowers, *J. Am. Chem. Soc.* 2000, *122*, 3458.
[21] B. A. Cerda, S. Hoyau, G. Ohanessian, C. Wesdemiotis, *J. Am. Chem. Soc.* 1998, *120*, 2437.
[22] T. Wyttenbach, J. E. Bushnell, M. T. Bowers, *J. Am. Chem. Soc.* 1998, *120*, 5098.
[23] T. Shoeib, C. F. Rodriquez, K. W. M. Siu, A. C. Hopkinson, *Phys. Chem. Chem. Phys.* 2001, *3*, 853.
[24] J. Wu, C. B. Lebrilla, *J. Am. Chem. Soc.* 1993, *115*, 3270.
[25] K. Zhang, D. M. Zimmerman, A. Chung-Phillips, C. J. Cassady, *J. Am. Chem. Soc.* 1993, *115*, 10812.
[26] C. J. Cassady, S. R. Carr, K. Zhang, A. Chung-Phillips, *J. Org. Chem.* 1995, *60*, 1704.
[27] C. Liebecq, *Biochemical Nomenclature and Related Documents*, Portland Press, London, 1992.
[28] M. Saunders, K. N. Houk, Y. D. Wu, W. C. Still, M. J. Lipton, *J. Am. Chem. Soc.* 1990, *112*, 1419.
[29] This AMBER* force field is enhanced and/or modified versions of the corresponding native force-fields by Macromodel in ref. [10], and the native force-fields is from Kollman's united atom and all atom fields with additional parameters for organic functionality. Key references: S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta Jr., P. Weiner, *J. Am. Chem. Soc.* 1984, *106*, 765; S. J. Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case, *J. Comput. Chem.* 1986, *7*, 230.
[30] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* 1990, *11*, 440.
[31] *Gaussian 98*, Revision A.6, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A.

Lindenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc., Pittsburgh, PA, 1998.
[32] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.
[33] J. B. Foresman, A. Frisch, *Exploring Chemistry with Electronic Structure Methods*, 2nd ed., Gaussian, Inc. Pittsburgh, PA, 1996.
[34] C. W. Tsang, N. L. Ma, unpublished results.
[35] C. W. Tsang, Y. Tsang, C. H. S. Wong, N. L. Ma, in *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics*, Chicago, USA, 2001, and unpublished results.
[36] R. G. Cooks, P. S. H. Wong, *Acc. Chem. Res.* **1998**, *31*, 379.
[37] J. E. Del Bene, H. D. Mettee, M. J. Frisch, B. T. Luke, J. A. Pople, *J. Phys. Chem.* **1983**, *87*, 3279.
[38] F. M. Siu, N. L. Ma, C. W. Tsang, *J. Am. Chem. Soc.* **2001**, *123*, 5397.
[39] J. N. Israelachvili, *Intermolecular and Surface Force*, Academic Press, London, 1992.
[40] J. S. Klassen, S. G. Anderson, A. T. Blades, P. Kebarle, *J. Phys. Chem.* **1996**, *100*, 14218.
[41] B. Boys, M. Karplus, *J. Comput. Chem.* **1995**, *16*, 698.
[42] J. Tortajada, E. Leon, J.-P. Morizur, A. Luna, O. Mo, M. Yanez, *J. Phys. Chem.* **1995**, *99*, 13890.
[43] K. Zhang, C. J. Cassady, A. Chung-Phillips, *J. Am. Chem. Soc.* **1994**, *116*, 11512.
[44] B. Paizs, I. P. Csonka, O. Lendvay, S. Suhai, *Rapid Commun. Mass Spectrom.* **2001**, *15*, 637.
[45] C. F. Rodriquez, A. Cunje, T. Shoeib, I. K. Chu, A. C. Hopkinson, K. W. M. Siu, *J. Am. Chem. Soc.* **2001**, *123*, 3086.
[46] E. L. Lopez, M. Orozco, *J. Chem. Soc. Perkin Trans. 2* **1993**, *4*, 663.
[47] C. H. S. Wong, N. L. Ma, C. W. Tsang, unpublished results.
[48] J. E. Huheey, E. A. Keiter, R. L. Keiter, *Inorganic Chemistry: Principles of Structure and Reactivity*, 4th ed., Harper College Publishers, 1993.
[49] This mode of binding cannot be found within the 50 kJ mol energy window utilized in our MCMM calculations. The present C89 complex shown in Figure S-3 (Supporting Information) is obtained from fully optimized K⁺-CO complexes based on the geometry of species 6 in ref. [33].
[50] T. Wyttenbach, J. J. Batka Jr., J. Gidden, M. T. Bowers, *Int. J. Mass Spectrom.* **1999**, *185*, 143.
[51] M. N. Jasper, A. A. Aarset, R. J. Dodson, G. C. Dodson, I. S. Demestichas, L. N. Zakharkine, K. S. Wilson, Z. J. Davies, A. A. Lebedev, H. Harutyunyan, *J. Mol. Biol.* **1998**, *276*, 491.
[52] T. M. Larsen, L. T. Laughlin, H. M. Holden, I. Rayment, G. H. Reed, *Biochemistry* **1994**, *33*, 6301.
[53] M. T. Rodgers, P. B. Armentrout, *Mass Spectrom. Rev.* **2000**, *19*, 215.
[54] L. M. Theisen, R. C. Orlando, J. Adams, *J. Am. Chem. Soc.* **1991**, *113*, 3668.
[55] I. K. Chu, K. Guo, T. C. Lau, K. W. M. Siu, *Anal. Chem.* **1999**, *71*, 1364.

Received: April 26, 2002 [F4052]

0947-6539/02/0821-4918 $ 20.00+.50/0

Chem. Eur. J. 2002

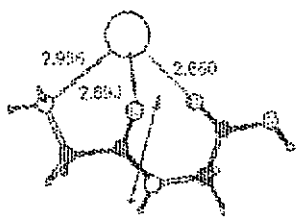
Supporting Information

for

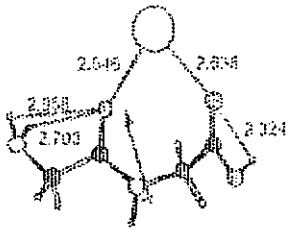
A Theoretical Study of Potassium Cation Binding to Glycylglycine (GG) and Alanylalanine (AA) Dipeptide

Carrie H. S. Wong,[a] N. L. Ma,*[b] and C. W. Tsang*[a]

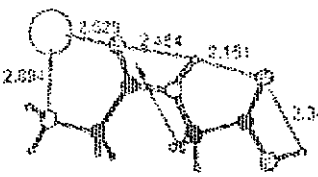
Figure S1. Summary of the structures and energetics (in kJ mol^{-1}) of low-lying conformers on the K^+-GG potential energy surfaces other than those shown in Figures 2 and 3. The intramolecular hydrogen bonds are indicated by dotted lines. The molecular dipole moment vectors of the *deformed* AA ligand (not to scale) are included for reference. The K^+ binding energies at 0K of these complexes, relative to that of **CS1**, are included in brackets.



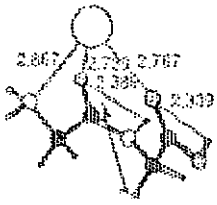
CS2a
[23.5]



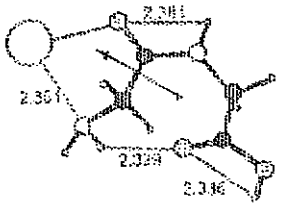
CS1a
[26.6]



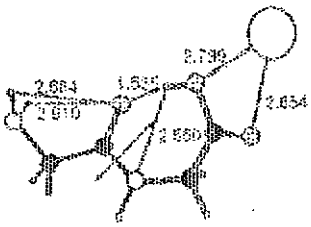
CS5a
[26.9]



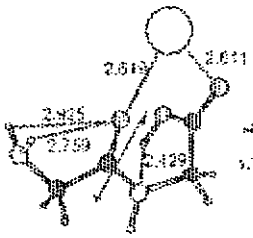
CS2b
[27.3]



CS0b
[29.3]

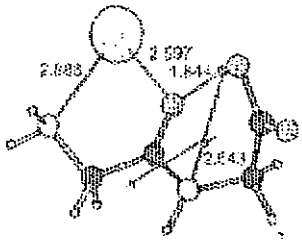


CS5b
[30.8]

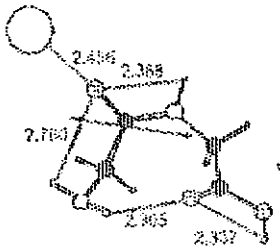


CS1b
[47.0]

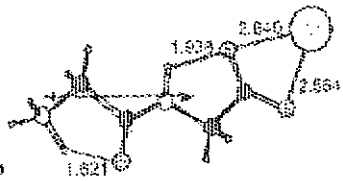
ZW(O_1)a
[54.3]



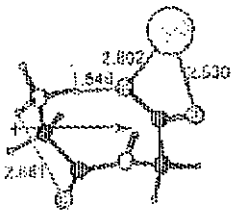
CS6c
[57.0]



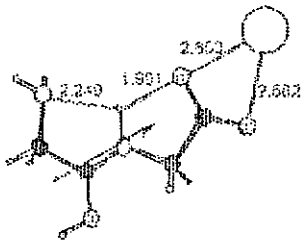
CS7a
[61.9]



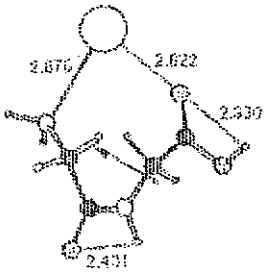
ZW(N_1)a
[62.0]



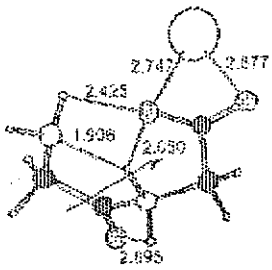
ZW(N_1)b
[71.6]



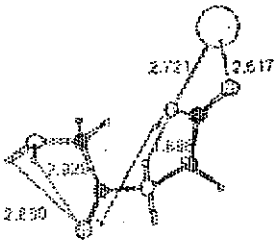
ZW(O_1')b
[71.8]



CS9
[97.8]



ZW(N_2)a
[115.0]



ZW(N_2)b
[137.5]

Absolute Potassium Cation Affinities (PCAs) in the Gas Phase

Justin Kai-Chi Lau,[a] Carrie Hoi Shan Wong,[a] Po Shan Ng,[a] Fung Ming Siu,[a] Ngai Ling Ma,*[b] and Chun Wai Tsang*[a]

Abstract: The potassium cation affinities (PCAs) of 136 ligands (20 classes) in the gas phase were established by hybrid density functional theory calculations (B3-LYP with the 6-311+G(3df,2p) basis set). For these 136 ligands, 70 experimental values are available for comparison. Except for five specific PCA values—those of phenylalanine, cytosine, guanine, adenine (kinetic-method measurement), and Me_2SO (by high-pressure mass spectrometric equilibrium measurement)—our theoretical estimates and the experimental affinities are in excellent agreement (mean absolute deviation (MAD) of 4.5 kJ mol^{-1}). Comparisons with previously reported theoretical PCAs are also made. The effect of substituents on the modes of binding and the PCAs of unsubstituted parent ligands are discussed. Linear relations between Li^+/Na^+ and K^+ affinities suggest that for the wide range of ligands studied here, the nature of binding between the cations and a given ligand is similar, and this allows the estimation of PCAs from known Li^+ and/or Na^+ affinities. Furthermore, empirical equations relating the PCAs of ligands with their dipole moments, polarizabilities (or molecular weights), and the number of binding sites were established. Such equations offer a simple method for estimating the PCAs of ligands not included in the present study.

Keywords: alkali metals · binding affinities · cations · density functional calculations · potassium

Introduction

The potassium cation is one of the most abundant metal cations in biological systems. The binding between potassium cation and protein/DNA–RNA/carbohydrate structures underlies many fundamental biological processes and enzymatic functions.[1] Knowledge of the K^+ binding modes and intrinsic binding energies (affinities) of smaller model ligand systems are fundamental to a full understanding of the interaction of K^+ in the more complex and larger biological systems.

A variety of experimental techniques has been employed to determine the alkali metal cation affinities of small model ligands. Absolute affinities were obtained by threshold collisional induced dissociation (threshold CID),[2-13] radiative associative kinetics measurements,[14] and high-pressure mass spectrometric (HPMS) equilibrium measurements,[15-24] while relative affinities were obtained by Fourier transform ion cyclotron resonance (FT-ICR) ligand exchange equilibrium measurements[25-28] and the mass spectrometric kinetic method.[29-31] Complementing the progress made in experimental measurements, quantum chemical methods have advanced to a stage that not only relative but also absolute alkali metal cation affinities can be obtained in excellent agreement (accuracy: ≈15 kJ mol^{-1}) with experimental values.[7,11,21,28,32-34] In many instances, reliable theoretical results have been shown to provide a complementary/alternative route for obtaining and confirming alkali metal cation affinities.[5,34-37] Also, theoretical findings on the most stable and low-lying binding modes/structures often provide new insights into the interpretation of experimental data. For example, a recent theoretical study on Li^+, Na^+, and K^+ affinity of DNA/RNA nucleobases highlighted the problem of assigning correct binding structures to measured alkali metal affinities when the free ligand exhibits tautomerism.[36,37] On the other hand, measured experimental affinities are essential for the calibration, validation, and establishment of reliable theoretical protocols.

[a] Dr. C. W. Tsang, J. K.-C. Lau, C. H. S. Wong, P. S. Ng, Dr. F. M. Siu
Department of Applied Biology and Chemical Technology
and Central Laboratory of the Institute of Molecular Technology
for Drug Discovery and Synthesis**
The Hong Kong Polytechnic University, Hung Hom, Hong Kong
Fax: (+852)23649932
E-mail: bcctsang@polyu.edu.hk

[b] Dr. N. L. Ma
Institute of High Performance Computing,
1 Science Park Road, #01-01,
The Capricorn, Singapore Science Park II, Singapore 117528
Fax: (+65)67780522
E-mail: ida@ihpc.a-star.edu.sg

[**] An Area of Excellence of the University Grants Committee (Hong Kong)

Supporting information for this article is available on the WWW under http://www.chemeurj.org or from the author.

While good compilations of intrinsic interaction energies between Li^+/Na^+ and model organic ligands are available,[26, 27, 28] far fewer potassium cation affinities (both experimental and theoretical) have been reported in the literature. The present work represents the most comprehensive theoretical study on the potassium cation affinity (PCA) scale reported to date. By comparison with the existing experimental data, the accuracy of the theoretical protocol was established. For the 136 ligands studied here, the K^+ interacts with different atoms (noble gases, carbon, oxygen, nitrogen, sulfur, phosphorus, etc.) in the ligands, and with a wide range of functional groups (alcohol, sulfide, sulfoxide, amine, amide, ether, aldehyde, ketone, nitrile, carboxylic acid, aromatic, heterocyclic, etc.). With such a broad spectrum of ligands, we believe the findings presented here are of general chemical interest and useful in revealing the nature of K^+ binding to organic and biological ligands in the gas phase.

Methods of Calculation

Based on our previous studies on a smaller set of ligands,[26, 28] the interactions between potassium cations and neutral ligands in the gas phase were modeled using the following protocol:

1) Geometry optimization at the HF/6-31G(d) level, followed by frequency calculations to obtain the zero-point energy (ZPE) correction.
2) The effect of electron correlation on structures of ligands and K^+–ligand complexes was obtained by full geometry optimization at the B3-LYP/6-31G(d) level.[29]
3) Energetics were obtained by using the B3-LYP functional with the large and flexible 6-311 + G(3df,2p) basis set, based on geometries determined in step (2), that is, the energetic calculations were carried out at the B3-LYP/6-311 + G(3df,2p)//B3-LYP/6-31G(d) level.

Potassium cation affinities (PCAs) at 0 K, ΔH_0, were obtained by using Equation (1) where E_K, E_L, and E_{K^+L} are the electronic energies of the potassium cation, the ligand, and the K^+–ligand complex, respectively, obtained from step (3); ZPE_L and ZPE_{K^+L} are the zero-point energy corrections for the ligand and the K^+–ligand complex, respectively, obtained from step (1) with a scaling factor of 0.8929 for the Hartree–Fock frequencies.[30, 31] For ease of description, this protocol is abbreviated as energetic protocol for K^+, EP(K^+), in the following. For comparison with experimental values, the EP(K^+) theoretical values at 0 K (ΔH_0) were converted to affinities at 298 K (ΔH_{298}) by standard statistical thermodynamics relations[28] calculated from the scaled HF/6-31G(d) vibrational frequencies.

$$\Delta H_0 = [(E_K + E_L) - E_{K^+L}] + [ZPE_L - ZPE_{K^+L}] + 0.5929 \quad (1)$$

Results and Discussion

Overview of theoretical and experimental alkali metal cation affinities: In this section, we highlight some recent advances in and issues related to the theoretical and experimental determination of alkali metal cation affinities reported in the literature. Using a density functional based method, Bauschlicher et al. recently reported the Li^+ affinity for 63 ligands, calculated at the B3-LYP/6-311 + G(d,p) level.[28] Calibration against experimental data suggested that this level of theory carried an average unsigned error of 15 kJ mol^{-1}, and the accuracy could be improved if systematic errors were taken into account.[28] Several theoretical studies of Na^+ affinities appeared recently.[26, 27, 32, 33] Most of these were conducted at the MP2(full)/6-311 + G(2d,2p)//MP2/6-31G(d) level, with corrections for basis set superposition error (BSSE).[34, 35] Armentrout and Rodgers further compared the performance of this level of theory and other models (DFT, CBS, G2, G3, etc.) with experimental data.[36] Theoretical and experimental sodium affinities were found to be in good general agreement. However, while BSSE-corrected B3-LYP/6-311 + G(2d,2p)//B3-LYP/6-31G(d) sodium affinities are consistently too high (MAD of 8.5 kJ mol^{-1} relative to experiment), the BSSE-corrected B3-P86/6-311 + G(2d,2p)//B3-P86/6-31G(d) values appear to fare better (MAD of 5.5 kJ mol^{-1} in comparison with experiment).[26] Recently, to improve the quality of the basis set for the sodium inner-valence 2s/2p orbitals, Petrie deconstructed the standard 6-311 + G(3df) basis for the Na atom. With this more flexible basis set, Na^+ affinities of 38 ligands at the geometry-corrected counterpoise "CPd-G2thaw" level were obtained,[26] and it was suggested that the currently established experimental sodium cation affinities were systematically too low by 3 - 5 kJ mol^{-1}.

For K^+, far fewer theoretical and experimental values are available. The EP(K^+) theoretical PCAs at 298 K obtained in this study and the available experimental PCAs are compiled in Table 1 and graphically presented in Figure 1, while the

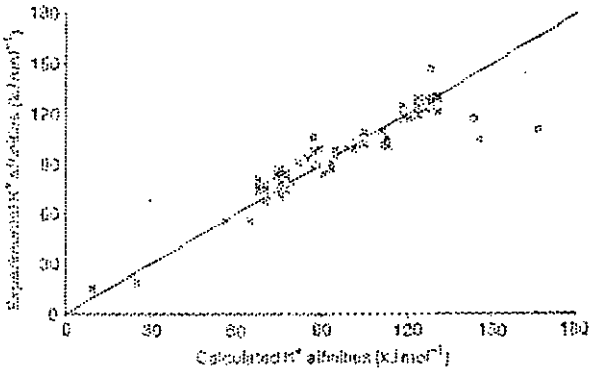


Figure 1. Plot of experimental versus EP(K^+) PCAs; the diagonal line with a slope of 1.0 is drawn for reference purposes. Large differences (> 15 kJ mol^{-1}) are depicted as o.

geometries of K^+ binding modes for representative class of ligands are shown in Figure 2. Here, we wish to comment on the temperatures reported in the experimental studies. The experimental PCA values were largely obtained by threshold CID, HPMS, and kinetic methods. The threshold CID measurements yielded PCAs at 0 K, and corrections to 298 K (generally less than 2 kJ mol^{-1}) were carried out by theoretical calculations. In HPMS measurements, it is generally assumed that the standard enthalpy of cation binding ΔH is approximately equal to the ion–ligand bond dissociation energy, and is independent of temperature effects. Similarly, in measurements by the kinetic method, the relative affinity measured, $\Delta(\Delta H)$, is also implicitly assumed to be independent of the "effective temperature", even though the $\Delta(\Delta H)$ term may be anchored to a reference ΔH value at a

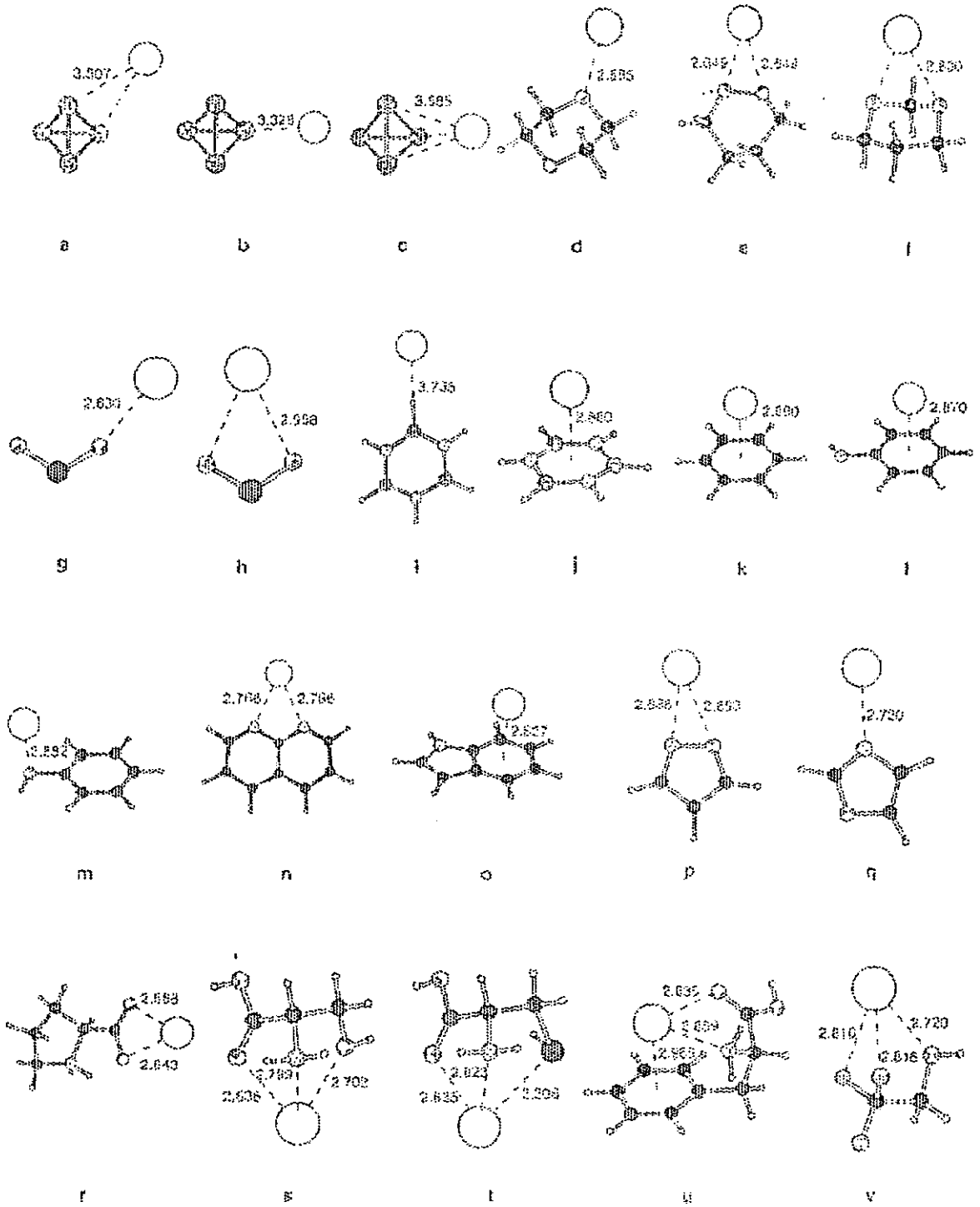


Figure 2. Geometries of selected K^+–ligand complexes, optimized at the B3-LYP/6–31G(d) level of theory. Ligand is a = c = H_2; d = 1,4-dioxane; e = 1,2-dioxane; f = 1,3-dioxane; g, b = SO_2; i, j = furazine; k = benzene; l, m = phenol; n = 1,8-naphthyridine; o = indole; p = isoxazole; q = oxazole; r = proline; s = serine; t = cysteine; u = phenylalanine; v = CF_3CH_2OH

known temperature. Consequently, experimental values obtained by HPMS and the kinetic method are usually not reported at any specific temperature. However, temperature effects on the ΔH or $\Delta(\Delta H)$ term are expected to be small. For the 136 ligands studied, we found general agreement (within ±15 kJ mol^{-1}) between our EP(K^+) PCA values and experimental values reported in the literature. Only for a few sets of PCA values (phenylalanine, cytosine, guanine, adenine obtained by kinetic method measurements, and Me_2SO obtained by high-pressure mass spectrometric equilibrium measure-

Table 1. Theoretical EP(K+) PCAs at 298 K and experimental PCAs [kJ mol−1].

Molecule[a]	Theoretical[b]	Experimental[c]	Molecule[a]	Theoretical[b]	Experimental[c]
He	4.1		Ne	4.1	
Ar	9.4	15.4(2)[d]	CO	24.3	19.0(5.0)[d]
HF	51.3		HCl	27.6	
P_4	31.1		PH_3	42.1	
C_2H_2	37.4		C_2H_4	34.7	
H_2S	39.6		MeSH	52.3	
EtSH	57.3		nPrSH	59.1	
iPrSH	60.2		nBuSH	60.6	
tBuSH	63.2		iBuSH	62.6	
Me_2S	61.8		H_2O	70.4	70.7,[e] 74.9[f]
MeOH	75.7	79.5,[e] 83.7[g]	EtOH	81.4	
nPrOH	82.2		iPrOH	85.2	
nBuOH	86.2		iBuOH	80.5	
sBuOH	87.6		tBuOH	85.1	
1,2-propanediol	116.2		1,3-propanediol	122.5	
ethylene glycol	119.3		glycerol	133.9	
CF_3CH_2OH	71.4		CCl_3CH_2OH	76.1	
Me_2O	74.9	87.0,[e] 71.0(4.0)[i]	Et_2O	85.0	93.3[e]
$(MeOCH_2)_2$	123.5	120.0(4.0),[i] 129.7[j]	1,2-dioxane	84.1	
1,3-dioxane	83.9		1,4-dioxane	71.8	
HCHO	77.9		MeCHO	93.7	
EtCHO	95.8		nPrCHO	97.3	
nBuCHO	98.5		CF_3CHO	59.5	
CCl_3CHO	74.0		Me_2CO	101.6	102.1,[e] 108.8[f]
MeCOEt	104.9		NH_3	77.2	74.9,[k] 84.1,[f] 80(5)[m]
$MeNH_2$	79.2	75.9[e]	Me_2NH	77.5	81.6[e]
Me_3N	74.2	83.7[e]	$EtNH_2$	81.8	
$nPrNH_2$	81.3	91.3[e]	HCN	78.4	
MeCN	101.9	102.1(1.7)[d]	EtCN	105.5	
nPrCN	107.0		iPrCN	108.2	
tBuCN	110.5		$PhCH_2CN$	110.3	
CF_3CN	58.6		CCl_3CN	76.4	
HCO_2H	88.4		HCO_2Me	90.2	

[a] Abbreviations: Me = CH_3, Et = C_2H_5, nPr = C_3H_7, iPr = $(CH_3)_2CH$, nBu = C_4H_9, iBu = $(CH_3)_2CHCH_2$, sBu = $C_2H_5CH_3CH$, tBu = $(CH_3)_3C$, Ph = C_6H_5. [b] This work, theoretical EP(K+) affinities at 298 K (ΔH_{298}). [c] Experimental affinities are tabulated at 298 K (ΔH_{298}) or at unspecified temperatures (see text for further discussion). Numbers in parentheses represent reported experimental uncertainties where available. [d] Threshold CID, ΔH_{298}, adjusted from ΔH_0 reported in ref. [1]. [e] HPMS, ref. [17]. [f] HPMS, ref. [15]. [g] HPMS, ref. [20]. [h] Estimated from ligand-exchange HPMS data, ref. [53]; first value anchored to experimental PCA of water reported in ref. [17], and the second to experimental PCA of water reported in refs. [15] and [20]. [i] Threshold CID, ΔH_{298}, ref. [3]. [j] HPMS, ref. [16]. [k] Threshold CID, ΔH_{298}, ref. [2]. [l] HPMS, ref. [21]. [m] Threshold CID, ΔH_{298}, ref. [17]. [n] HPMS, ref. [19]. [o] HPMS, ref. [18].

ment, as indicated by open circles in Figure 1) does the difference between EP(K+) values and reported experimental values exceed 15 kJ mol−1. These cases of discrepancy are further discussed in the individual sections below.

In addition, we also compiled the theoretical PCAs from the literature (Table 2), corrected to 298 K if required, to facilitate comparison. Different theoretical models were employed in various studies, in which different geometries, electron correlation methods, basis sets, core sizes, and zero-point energies were employed, and BSSE corrections may or may not be included.[3–5, 6, 8, 9, 12–13, 15, 33, 35, 45–47] Even though direct comparison is difficult, it appears that the PCAs estimated by the EP(K+) protocol tend to be slightly larger than the affinities estimated by other methods, except when the K+ binds to the ligand by aromatic π binding (as in the case of benzene, phenol, and pyrrole). Nevertheless, the general agreement is good (within ±5 kJ mol−1) except for a few species (CO, H_2S, NH_3, imidazole, glycine, and the five DNA/RNA nucleobases). Detailed discussions of the discrepancies are given in individual sections below.

Before we discuss the PCAs of individual class of ligands, we would like to comment on our choice of B3-LYP over B3-P86 density functional. To obtain the energetics at the B3-P86/6-311+G(3df,2p)//B3-LYP/6-31G(d) level for comparison, we replaced the single-point energy calculations (step (3) of the EP(K+) protocol) with the B3-P86 functionals, and the results are shown in the Supporting Information (Table S1). We found that the affinities obtained by EP(K+) and B3-P86 functional are both in good agreement with existing experimental affinities, with MADs of 4.3 and 4.3 kJ mol−1, respectively. The EP(K+) affinities tend to be slightly larger in most cases, except when K+ interacts with π-type ligands such as C_2H_2, C_2H_4, benzene, and borazine. The largest difference is found for glycerol, for which the EP(K+) affinity is larger than the B3-P86 affinities by 6 kJ mol−1. For species which show a relatively large difference in EP(K+) and B3-P86 affinities, we carried out additional calculations at the G2(MP2,SVP) level for benchmarking. We found that the EP(K+) affinities are marginally more comparable to this benchmark level with a MAD of 3.2 kJ mol−1 (as opposed to

Table 1 (cont.).

Molecule[a]	Theoretical[b]	Experimental[b]	Molecule[a]	Theoretical[c]	Experimental[c]
$MeCO_2H$	91.8		HCO_2Et	95.5	
HCO_2nPr	97.0		$MeCO_2Me$	99.5	
$MeCO_2Et$	105.7		$EtCO_2Me$	106.9	
CF_3CO_2Me	82.2		$ClCO_2Me$	86.9	
SO_2	81.9		Me_2SO	128.1	130.1[d] 146.4(12.6)[g]
PhSOMe	135.7		$HCONH_2$	134.1	
HCONHMe	120.7	117.2[d]	$HCONMe_2$	136.2	123.4[d] 129.2[e]
$MeCONH_2$	125.3	134.3[d] 118.9[f]	MeCONHMe	126.4	127.5[e]
$MeCONMe_2$	130.6	131.3[d] 129.7[d] 131.2[e]	benzene	67.6	70.6[d] 74.2(6.1)[d] 80.3[e]
benzene	46.8		phenol	70.0	77.9(12.6)[d] 75.6(3.8)[g]
pyridine	93.3	86.6[d] 90.0(5.9)[d]	2-methylpyridine	94.6	98.5(3.5)[d]
3-methylpyridine	99.1	100.1(3.5)[d]	4-methylpyridine	101.1	99.0(4.0)[d]
3-fluoropyridine	180.3		3-chloropyridine	81.9	
1,8-naphthyridine	151.6		pyridazine	136.0	130.9(2.6)[d]
pyrimidine	75.7	69.7(4.3)[d]	pyrazine	69.8	67.6(3.6)[d]
1,3,5-triazine	56.5	51.6(3.6)[d]	pyrrole	77.1	85.1(3.6)[d]
indole	89.1	99.8(12.6)[d]	pyrazole	90.5	84.2(3.3)[d]
imidazole	111.1	109.6(5.6)[d]	thiazole	87.1	
isothiazole	86.1		oxazole	84.5	
isoxazole	94.4		1-methylpyrazole	94.5	94.6(3.6)[d]
3-methylpyrazole	92.8		4-methylpyrazole	96.4	
1,4-dimethylpyrazole	100.0		1,5-dimethylpyrazole	101.3	
1,3,5-trimethylpyrazole	103.4		3,4,5-trimethylpyrazole	105.4	
1,2,4,5-tetramethylpyrrole	106.2		1-methylimidazole	116.8	117.7(2.2)[d]
1,3-dimethylimidazole	120.5		2,4,5-trimethylimidazole	122.0	
2*H*-1,2,3-triazole	64.5[h]	55.6(5.5)[d]	1*H*-1,2,4-triazole	87.0[d]	87.5(4.5)[d]
2*H*-tetrazole	85.5[h]	89.6(4.6)[d]	1*H*-tetrazole	109.7[d]	
4*H*-1,2,4-triazole	140.1[d]		1*H*-1,2,3-triazole	118.6[h]	
glycine	118.4	115.5[d] 119.3[e]	alanine	124.0	123.6[d]
valine	128.2	128.8[d]	leucine	128.8	129.3[d]
isoleucine	128.8	129.9[d]	proline	140.0	
serine	137.3		cysteine	123.5	
phenylalanine	145.6	104.2 (20.9)[d]	adenine	87.1[d]	106.3[d] 97.4(3.2)[d]
thymine	112.8[d]	102.3[d] 104.6(3.8)[d]	uracil	113.1[d]	101.3[d] 105.8(3.8)[d]
cytosine	146.3[d]	110[d]	guanine	143.3[d]	112[d]

[a] Ref. [3–4], kinetic-method measurements using theoretical G2(MP2,SVP)-ASCI(GCP) K^+ affinity (ΔH_0) values at 0 K of formamide (109.2 kJ mol^{-1})/*N*,*N*-dimethylformamide (123.9 kJ mol^{-1})/*N*-methylacetamide (129.6 kJ mol^{-1}) as reference values, as reported in ref. [33] and corrected to ΔH_{298}. [b] Threshold CID, ΔH_{298}, ref. [8]. [c] RPNS, ref. [32]. [d] Radiative association kinetics measurement, ref. [14], reported value reduced by 6.2 kJ mol^{-1}, as described in text, and with thermal correction to 298 K. [e] Threshold CID, ΔH_{298}, ref. [11]. [f] Threshold CID, ΔH_{298}, ref. [6]. [g] Threshold CID, ΔH_{298}, ref. [11]. [h] Threshold CID, ΔH_{298}, ref. [12]. [i] Threshold CID, ΔH_{298}, ref. [5]. [j] EP(K^+) PCA of the most stable tautomer of the free ligands 1*H*-1,2,4-triazole, 2*H*-1,2,3-triazole, and 2*H*-tetrazole kinetically (but not energetically) favored to bind to K^+ in the gas phase, resulting in the formation of K^+–(1*H*-1,2,4-triazole), K^+–(2*H*-1,2,3-triazole), and K^+–(2*H*-tetrazole) complexes proposed as likely to be observed in threshold-CID experiments as explained in ref. [5]. [k] EP(K^+) PCA of the less stable tautomer of the free ligands 4*H*-1,2,4-triazole, 1*H*-1,2,3-triazole, and 1*H*-tetrazole, energetically (but not kinetically) favored to bind to K^+ in the gas phase, leading to formation of the most stable K^+–(4*H*-1,2,4-triazole), K^+–(1*H*-1,2,3-triazole), and K^+–(1*H*-tetrazole) complexes with the largest PCA, but not likely observed in threshold-CID experiments, as explained in ref. [5]. [l] Ref. [70], kinetic-method measurements using theoretical G3(MP2,SVP)-ASCI(GCP) (see also footnote [p]) K^+ affinity values at 0 K (ΔH_0) of acetamide (118.7 kJ mol^{-1})/*N*-methylacetamide (129.6 kJ mol^{-1})/*N*,*N*-dimethylacetamide (139.2 kJ mol^{-1}) as reference values, as reported in ref. [33]; the experimentally determined values were corrected to ΔH_{298}. [m] Kinetic method, ref. [31]. [n] The PCA is estimated for the most stable tautomer of the free ligands, corresponding to species A1, T1, U1, C1, and G1 in ref. [37]. [o] Kinetic method, ref. [30]. [p] Threshold CID, ΔH_{298}, ref. [9].

4.1 kJ mol^{-1} for B3-P86 values). As the performance of both functionals does not differ significantly, the B3-LYP functional was chosen for obtaining PCAs as it is more widely used, and this will possibly allow more direct comparison with other theoretical studies.

For a representative subset of 14 ligands shown in the Supporting Information (Table S2), we further explored the effects of employing B3-LYP geometries and vibrational frequencies on zero-point vibrational energy and thermal correction to 298 K, and the magnitude of the BSSE. Using the B3-LYP frequencies to correct for zero-point energies tends to decrease the theoretical PCA by about 0.5 kJ mol^{-1}, as compared to using the HF frequencies. The thermal correction at 298 K with B3-LYP parameters leads to a further decrease of about 0.3 kJ mol^{-1}. We found that the estimated BSSE obtained by the density functional-based protocol is small (average of 0.7 kJ mol^{-1}), but can be as large as 1.8 kJ mol^{-1} when the cation binds to the ligand by aromatic π interactions. For this subset of ligands, seven experimental PCAs (CO, H_2O, NH_3, benzene, pyridine, pyrrole, and uracil) are available for comparison. For these seven species, inclusion of the above refinement in fact leads to a slight increase in MAD (6.4 kJ mol^{-1}, as opposed to 6.0 kJ mol^{-1} with our EP(K^+) protocol). Therefore, the corrections may not lead to better agreement with experimental data. Given that the magnitude of these corrections (ca. 1–2 kJ mol^{-1}) has

Table 3. Theoretical PCAs (kJ mol^{-1})

Molecule[a]	Theoretical[b]	Literature[c]	Molecule[a]	Theoretical[b]	Literature[c]
CO	24.8	37.1[e]	PH_3	42.5	41.1[d]
H_2S	17.6	31.0[e]	H_2O	78.4	71.7[e], 70.7[f], 66.9[g]
MeOH	75.7	74.9[g]	EtOH	81.4	82.4[d]
nPrOH	83.2	81.2[g]	iPrOH	86.2	85.9[d]
nBuOH	86.3	85.8[g]	sBuOH	86.5	83.6[d]
tBuOH	87.6	87.4[g]	iBuOH	86.1	88.7[g]
1,2-propanediol	116.2	117.2[g]	1,3-propanediol	122.5	123.2[g]
ethylene glycol	119.1	118.2[g]	glycerol	133.5	134.3[g]
Me_2O	74.9	79.0[g]	NH_3	79.2	79.5[e], 77.6[g] 76.1[g] 75.1[g] 77.1[g] 76.1[g] 64.1[g] 66.1[g]
$HCONH_2$	114.1	111.7[d]	HCONHMe	130.7	118.9[g]
$HCONMe_2$	136.2	134.5[g]	$MeCONH_2$	123.2	120.2[g]
MeCONHMe	128.4	126.9[g]	$MeCONMe_2$	130.6	129.5[g]
benzene	67.6	72.0[g]	phenol	79.0	74.1[g]
pyridine	93.5	91.3[g], 91.4[g]	2-methylpyridine	94.6	94.2[g]
3-methylpyridine	99.1	97.0[g]	4-methylpyridine	101.1	97.6[g]
pyridazine	130.0	130.9[g]	pyrimidine	75.7	73.2[g]
pyrazine	69.8	68.7[g]	1,3,5-triazine	56.5	53.5[g]
pyrrole	72.1	83.1[g]	indole	89.1	87.2[g]
pyrazole	90.5	86.7[g]	imidazole	111.1	109.2[g], 116.5[g]
1-methylpyrazole	94.5	91.3[g]	1-methylimidazole	118.8	117.2[g]
2*H*-1,2,3-triazole	61.5	65.5[g]	1*H*-1,2,4-triazole	87.0	91.7[g]
2*H*-tetrazole	53.5	50.7[g]	glycine	138.4	119.5[g]
phenylalanine	145.6	146.4[g]	adenine	83.1[g]	78.7[g], 85.2[g]
thymine	112.0[g]	107.1[g], 104.0[g]	uracil	113.1[g]	105.4[g], 104.5[g]
cytosine	106.3[g]	109.0[g]	guanine	145.3[g]	139.2[g]

[a] Abbreviations: Me = CH_3, Et = $-C_2H_5$, *n*Pr = C_3H_7, *i*Pr = $(CH_3)_2CH$, *n*Bu = C_4H_{9-}, *i*Bu = $(CH_3)_2CHCH_2$, *s*Bu = $C_2H_5CH(CH_3)$, *t*Bu = $(CH_3)_3C$, Ph = C_6H_5. [b] This work, theoretical EP(K$^+$) affinities at 298 K (ΔH_{298}). [c] Previously reported PCA at 298 K from literature. For cases where only values at 0 K are reported, thermal corrections to 298 K using HF/6-31G(d) geometries and frequencies were applied. [d] Ref. [43]. [e] Ref. [44]. [f] Ref. [33]. [g] Ref. [35]. [h] Ref. [46]. [i] Ref. [3]. [j] Ref. [47]. [k] Ref. [8]. [l] Ref. [15]. [m] Ref. [11]. [n] Ref. [6]. [o] Ref. [12]. [p] Ref. [14]. [q] Ref. [5]. [r] Ref. [45]. [s] Ref. [48]. [t] The PCA is estimated for the most stable tautomer of the free ligands, corresponding to species A1, T1, U1, C1, and G1 in ref. [37]. [u] Ref. [9].

minimal effect on calculated PCAs, it appears that they could be omitted for computational efficiency.

Noble gas atoms: The strength of K$^+$ binding to the noble gas atoms is at the lowest end of the PCA scale. While the experimental K$^+$ affinities for He and Ne are not known, the Ar affinity[d] was determined to be 14(7) kJ mol^{-1} at 0 K (experimental uncertainty in parenthesis). Our calculated K$^+$ affinity for Ar (8 kJ mol^{-1} at 0 K) is within the error limit of this experimental value.

As the potassium cation is isoelectronic with the argon atom, it is of interest to compare the PCAs of the noble gas atoms with the bond dissociation energy of the corresponding isoelectronic HeAr, NeAr, and Ar_2. The bond dissociation energies for the species are very small (< 1 kJ mol^{-1}), as the rare gas atoms are held together by weak dispersion forces.[49] In comparison, the interactions between K$^+$ and noble gas atoms are much stronger, ranging from 4 to 9 kJ mol^{-1}, and reflect the strength of the inductive forces when a cation interacts with the polarizable noble gas atoms.

Carbon monoxide: In agreement with previous findings on Li^+,[40,41] Na^+[34,41] and K$^+$,[39] the potassium cation prefers to bind to the carbon atom of CO (estimated K$^+\cdots$C distance of 3.04 Å) with formation of a linear cation–ligand complex. The preferential binding of K$^+$ to the carbon atom can be explained by the fact that even though oxygen is more electronegative than carbon, the negative end of the CO dipole in fact resides on the carbon atom.[39] Thus, our result highlights the importance of ion–dipole interaction in this complex.

The EP(K$^+$) PCA is approximately 12 kJ mol^{-1} smaller than that reported previously by Ikuta,[39] and the discrepancy arises presumably from the small basis sets employed in the previous study. We note in passing that although the experimental PCA of CO at 298 K (19(3) kJ mol^{-1})[d] is closer to our theoretical estimate for the K$^+\cdots$OC (18 kJ mol^{-1}) than for the K$^+\cdots$CO (25 kJ mol^{-1}) binding mode, both theoretical values are within the error bar of the experimental measurement.

Hydrogen fluoride and hydrogen chloride: K$^+$ prefers to bind to the halogen atoms in hydrogen halides. While K$^+$–HF is linear, the K$^+$–HCl complex is bent (K$^+\cdots$Cl-H angle of 119°). Sodiated hydrogen chloride (Na$^+$–HCl) has a similar shape to the K$^+$–HCl complex, with an estimated Na$^+\cdots$Cl-H angle of 114°.[39] Hence, for both Na$^+$ and K$^+$, the cation does not bind along the dipole moment vector of HCl; this reflects

a certain degree of covalency when the alkali metal cation binds to second-row elements.[21]

Phosphorus and phosphine: A theoretical study on Li^+–P_4 suggested that Li^+ binds to P_4 in a monodentate fashion, along a threefold axis of the tetrahedron (species 2 in ref. [51]). We found that the larger K^+ prefers to bind in a bidentate manner on an edge of the P_4 tetrahedron, with a $K^+ \cdots P$ distance of 3.51 Å (C_{2v}, species **a**, Figure 2) and a PCA of 34 kJ mol^{-1}. The monodentate complex (C_{3v}, species **b**, Figure 2) and the tridentate face-bound K^+–P_4 complex (C_{3v}, species **c**, Figure 2) are approximately 4 kJ mol^{-1} less stable.

The PCA of phosphine (PH_3) is slightly higher than that of P_4. The cation binds to phosphine along the C_3 axis of the ligand, with a $K^+ \cdots P$ distance of 3.37 Å. Our estimate of K^+–PH_3 affinity at 0 K (41 kJ mol^{-1}) is in excellent agreement with a previous theoretical value of 40 kJ mol^{-1}).[21]

Ethene and ethyne: Boyan et al.[52] showed that Na^+ binds to the π bonds of C_2H_2 and C_2H_4 at an average distance of 2.65 Å above the plane of the ligand. For K^+, we found that the cation–π binding distance is longer (3.13 Å). The increased distance leads to a decrease in cation affinities by approximately 10 kJ mol^{-1}.

Hydrogen sulfide and thiols: For these sulfur-containing ligands, the cation binds to the sulfur atom at a distance of approximately 3.20 Å. The cation has a tendency to bind to one of the lone pairs of the sulfur atom and is hence rather poorly aligned (ca. 40°) with the molecular dipole moment of the ligand.

The estimated PCAs for ligands of this class studied here range from approximately 40 to 60 kJ mol^{-1}. Our theoretical K^+ affinity of H_2S at 0 K (38 kJ mol^{-1}) is 8 kJ mol^{-1} larger than the value previously reported by Magnusson.[19] This discrepancy may be attributed to two factors: no zero-point energy correction was made and the affinity was calculated with rather small basis sets in that paper. However, no experimental PCAs for this class of compounds are available for comparison.

Water and alcohols: For this class of ligands, K^+ binds to the oxygen atom at a distance of approximately 2.58 Å. For K^+–H_2O, the cation binds along the twofold axis of the ligand, in perfect alignment with its dipole moment. This is in contrast with the preferred mode of K^+ binding in H_2S discussed above, and again reflects the presence and influence of covalency when the cation binds to second-row atoms.

Two sets of experimental PCA for H_2O were determined by Kebarle et al. using the HPMS technique. Our present estimate of 70 kJ mol^{-1} is within ±5 kJ mol^{-1} of both experimental values: 70.7 kJ mol^{-1}[19] and 74.9 kJ mol^{-1}.[18,20] Three theoretical PCAs for H_2O are available for comparison.[20,37,41] The EP(K^+) PCA is in good agreement with all three values, and it is virtually identical to that obtained by Russo et al.[37] at the B3-LYP/6-311 + G(2df,2p)//B3-LYP/6-311 + G(2df,2p) level with BSSE correction.

For simple alcohols, the cation is in reasonable alignment (ca. 12°) with the dipole moment of the ligands. Previous studies[14,53] found that when nBuOH binds to Li^+, the alkyl chain wraps around so that the terminal carbon atom of the ligand undergoes a secondary interaction with the cation. In the case of K^+, this additional favorable interaction apparently cannot compensate for the unfavorable ligand deformation, so that this cyclic form is slightly less stable than the extended open form (by 5 kJ mol^{-1}).[20]

Direct experimental determination of the absolute PCA of simple alcohols has not been reported. However, the relative enthalpy change when water is exchanged by methanol was determined to be 8.8 kJ mol^{-1},[20] which is in reasonable agreement with the corresponding relative affinity values estimated by various theoretical protocols (i.e., EP(K^+) in Table 1; B3-PS6 and G2(MP2,SVP) in the Supporting Information, Table S1) to be in the range of 4.3–6.4 kJ mol^{-1}. We note that using the two experimental PCAs of water and the experimental relative enthalpy change of 8.8 kJ mol^{-1},[20] the absolute PCA for methanol can be estimated to be 79.5[19] or 83.7 kJ mol^{-1}.[18,20] All theoretical protocols employed here [EP(K^+), B3-PS6, G2(MP2,SVP)] yield PCA values that are more consistent with the lower estimate.

Compared with simple aliphatic alcohols, K^+ interactions with polyhydroxyl ligands have been less widely studied. Our results suggested that K^+ bind in a bidentate manner to ethylene glycol (1,2-ethanediol), 1,2-propanediol, and 1,3-propanediol, and in a tridentate fashion with glycerol (1,2,3-propanetriol). Compared with the monodentate binding in simple alcohols, bidentate K^+ interactions with the diols increase the PCA by about 30 kJ mol^{-1}. However, the transition from bidentate (in diols) to tridentate (in glycerol) binding leads to a further enhancement of the PCA by only about 15 kJ mol^{-1}. This suggests that the destabilizing effect of ligand deformation plays an important role in determining the PCA of multidentate complexes.[6,7,49]

Ethers and dioxanes: For this class of ligands, K^+ binds to the oxygen atom with a typical $K^+ \cdots O$ distance of 2.65 Å, slightly longer than that found in K^+ complexes of aliphatic alcohols, and hence the PCA is smaller than that of the corresponding alcohol analogue.

Dioxanes ($C_4O_2H_8$), which are well-known carcinogens, can be considered as cyclic ethers. For 1,4-dioxane, the cation binds in a monodentate fashion to one of the oxygen atoms in the ligand (C_s, species **d**, Figure 2). In K^+–1,2-dioxane (C_1, species **e**, Figure 2) and K^+–1,3-dioxane (C_s, species **f**, Figure 2), the cation is coordinated in a bidentate manner to the two closely situated oxygen atoms. As a result, the PCA of 1,4-dioxane is lower than those of 1,2- and 1,3-dioxane.

Experimental affinities are available for some ethers. The PCAs at 298 K of Me_2O and $(MeOCH_2)_2$ were determined by Armentrout et al. using the threshold CID method to be 74(4) and 120(6) kJ mol^{-1}, respectively.[8] Both values are in good agreement (within ±4 kJ mol^{-1}) with our theoretical estimates. In contrast, an earlier reported PCA of Me_2O[19] determined by HPMS (87 kJ mol^{-1}) differs more widely from our theoretical estimate of 74.9 kJ mol^{-1}.

Aldehydes and ketones: For this class of ligand, K^+ binds to the carbonyl oxygen atom. Compared to the alcohols and

ethers, the $K^+ \cdots O$ distance is shorter (2.56 Å), with a typical $K^+ \cdots O{=}C$ angle of 165°. Two experimental PCAs of Me_2CO were reported[5,21] and both are within ±4.5 kJ mol^{-1} of our calculated value (103 kJ mol^{-1}).

Ammonia and amines: K^+ binds to the electronegative nitrogen atom in these ligands with a typical $K^+ \cdots N$ distance of 2.78 Å. A few theoretical PCAs are available for comparison for NH_3:[5,21,42,43] The EP(K^+) PCA is in good agreement with all the reported values based on calculations with all-electron basis sets. In comparison, the two values reported in ref. [42] using pseudopotential are too low by over 10 kJ mol^{-1}. Our EP(K^+) PCA is virtually identical to that obtained at the B3-LYP/6-311+G(3d,2p)//B3-LYP/6-311+G(3d,2p)[21] and B3-LYP/6-311+G(3d,2p)//B3-LYP/6-31G(d) levels[43] and this indicates that if sufficiently large basis sets are used, the effects of geometry and zero-point corrections on K^+ binding affinities would be minimal. Interestingly, for $K^+(NH_3)_n$ complexes ($n = 1-3$), the BSSE at the B3-LYP level is small (within ±1 kJ mol^{-1}) when compared to the BSSE obtained by MP2 calculations (3–4 kJ mol^{-1}) using the same basis set.[21] The rather small BSSE corrections found for these systems at the B3-LYP level are in agreement with our general findings presented in the Supporting Information (Table S2).

Experimental affinities are available for ammonia and four alkylamines.[13,18,44] While the experimental PCAs are in good general agreement with our theoretical estimates (within ±10 kJ mol^{-1}), qualitative differences are found in the order of relative affinities upon successive methyl substitution of ammonia (see section "Effect of substituents" below for further discussion).

Hydrogen cyanide and alkyl nitriles: For HCN and the six alkyl nitriles studied here (including $PhCH_2CN$), K^+ prefers to bind to the nitrogen atom of the ligand. The average $K^+ \cdots N$ distance is 2.68 Å, slightly shorter than that found in potassiated amine complexes. This reflects the fact that the interaction of K^+ with the sp-hybridized nitrogen atom in nitriles is stronger than that with the sp^3-hybridized nitrogen atom in amines. The experimental PCA of MeCN[18] of 102 kJ mol^{-1} at 298 K is in excellent agreement with our calculated value.

Carboxylic acids and esters: Two potential K^+ binding sites are available for this class of ligands (two carboxylic acids and eight esters): the carbonyl and hydroxyl/alkoxyl oxygen atoms. We found that the K^+ prefers to bind to the carbonyl oxygen atom (average $K^+ \cdots O$ distance of 2.51 Å, average $K^+ \cdots O{=}C$ angle of 165°), with a "*trans*" conformation of the $K^+ \cdots O{=}C{-}O$ moiety. In this class of ligand, K^+ is in fairly good alignment with the molecular dipole moment of the ligand (angle of deviation ca. 10°). In general, the PCA of a carboxylic acid is larger than that of the corresponding alcohol: for example, the PCA of acetic acid is 5 kJ mol^{-1} higher than that of methanol. As we are not aware of any experimental and theoretical values in the literature, the PCAs reported here are the first set of estimates available for this class of ligands.

Sulfoxides: For the three sulfoxides studied here, including PhSOMe, K^+ binds exclusively to the oxygen atom, with an average $K^+ \cdots O$ distance of 2.60 Å. It is interesting to compare the structure of the $K^+{-}SO_2$ complex with that of the $Na^+{-}SO_2$ complex reported previously.[14] Ohanessian et al. found that Na^+ binds to the SO_2 group in a bidentate fashion (with C_s symmetry). We found that the larger K^+ prefers to bind in a monodentate fashion to one of the oxygen atoms (C_1 species g, Figure 2), and this mode of binding is about 4 kJ mol^{-1} more stable than the bidentate C_s mode (species h, Figure 2). Two experimental values for the PCA of Me_2SO [146 (HPMS)[20] and 130 kJ mol^{-1} (threshold CID)[18] were reported by Keheyle et al. However, it was pointed out that the HPMS value was based on the slope of a van't Hoff plot covering only a narrow temperature range, and hence was subject to a greater experimental error. Thus, the threshold-CID value is considered to be more reliable.[21] Our calculated PCA (126 kJ mol^{-1} at 298 K) is in very good agreement with the threshold-CID value of 130 kJ mol^{-1}, and provides further support for this more recently determined value.

Amides: Amides are perceived as an entry point for understanding the peptide linkage in protein structures, and hence K^+–amide interactions are of special biological interest. For the amide complexes studied here, K^+ binds in a monodentate fashion to the oxygen atom ($K^+ \cdots O$ distance of 2.48 Å, $K^+ \cdots O{=}C$ angle of 166°), in good alignment with the molecular dipole moment.[21] Keheyle et al. determined PCAs for four of the six amides studied here by HPMS[20] and threshold CID.[18] The reported 298 K PCA of *N,N*-dimethylacetamide ($MeCONMe_2$) deserves special attention. For this species, the PCA determined by threshold CID[18] (121 kJ mol^{-1}) is noticeably lower than that determined by HPMS (139 kJ mol^{-1}).[20] More importantly, the affinity determined by threshold CID is even lower than those of *N,N*-dimethylformamide ($HCONMe_2$, 123 kJ mol^{-1}) and acetamide ($MeCONH_2$, 124 kJ mol^{-1}).[18] This is quite surprising, as the more polarizable $MeCONMe_2$ is expected to have the highest K^+ affinity of these three amides. To resolve this discrepancy, we recently re-measured the PCA of $MeCONMe_2$ using the kinetic method. By anchoring to the ab initio G2(MP2,SVP)-ASC(GCP) theoretical value of $HCONMe_2$ corresponding to 125 kJ mol^{-1} at 298 K[21] (which is consistent with the threshold-CID value[18] of 123 kJ mol^{-1}), we obtained a PCA of $MeCONMe_2$ of 131 kJ mol^{-1}.[24] This indicates that, in contrast to K^+–Me_2SO (see above), the $K^+ \cdots MeCONMe_2$ affinity determined by the HPMS method[20] is more consistent with available theoretical and experimental values. If the threshold-CID value for $MeCONMe_2$ is excluded, our protocol yields affinities to within ±3 kJ mol^{-1} for all the six amides studied here.

Benzene, borazine, and phenol: The interaction between cations and aromatic π electrons is a relatively newly discovered type of electrostatic interaction. Such cation···π interaction are implicated in many important biological functions,[45–48] and benzene is often used as the prototype ligand for understanding them. The most recent experimental dissociation energy of K^+–benzene (71 kJ mol^{-1} at 0 K),

determined by Amicangelo and Armentrout,[8] is in good agreement with our estimated value of 69 kJ mol^{-1} at 0 K. We note that our 298 K PCA reported in Table 1 (67.6 kJ mol^{-1}) is obtained from the EP(K^+) PCA value at 0 K (62.1 kJ mol^{-1}), with thermal corrections using the HF/6-31G(d) geometry and vibrational frequencies. The experimental PCA (74.2 kJ mol^{-1}, Table 1) is taken directly from ref. [8], with thermal correction at the MP2(full)/6-31G(d) level. The apparent inconsistency (0.7 kJ mol^{-1}) between the two sets of PCA values at 0 K and 298 K probably arises from the level of theory employed for the thermochemical correction.

Even though borazine ($B_3N_3H_6$) is isoelectronic with benzene, its electronic nature has been controversial. Criteria based on magnetic properties and energetic and geometric indices suggested that borazine may not be aromatic.[37] However, recent findings showed that borazine should be aromatic, although its aromaticity is about half of that of benzene.[38] The most stable structure we obtained for borazine has D_{3h} symmetry and is in agreement with previous experimental and theoretical studies.[39] We located two stable minima on the K^+–borazine potential energy surface. The less stable complex is planar and has C_{3v} symmetry (species i, Figure 2). The more stable complex has C_s symmetry (species j, Figure 2), with the cation 2.58 Å above the ring centroid and $K^+ \cdots N$ and $K^+ \cdots B$ distances of 3.23 and 3.33 Å, respectively.

While the structural features of the most stable K^+–borazine complex (species j, Figure 2) are quite comparable to those obtained for K^+–benzene (C_{6v}, species k, Figure 2), the estimated PCA of borazine is 21 kJ mol^{-1} lower than that of benzene (68 kJ mol^{-1}). As the quadrupole moment of a ligand is expected to play a key role in determining the strength of the cation–π interaction,[40] we calculated the quadrupole moment of these two ligands. Not only are the calculated values (−3.62 and −7.88 Buckingham for borazine and benzene, respectively) in excellent agreement with experiment (−4.18 and −7.99 Buckingham, respectively),[41] but the increase in quadrupole moment of the ligand also correlates with the increase in PCA from borazine to benzene. Thus, the important contribution of ion–quadrupole interactions in this class of aromatic ligands appears to be confirmed.

Phenol is a prototypical case for competition between π and non-π hydroxyl oxygen binding sites for K^+. For the Na^+–phenol complexes,[24, 42] the binding affinities at 298 K for the aromatic π and non-π complexes are comparable, and differ by about 1–4 kJ mol^{-1}, depending on the computational protocol used to obtain the Na^+ affinities. However, the free energy of binding ΔG_{298} indicates that the non-π complex is the favored form of Na^+–phenol in the gas phase.[24]

For K^+–phenol, two stable minima were again found. In the more stable form (in terms of ΔH_{298}), K^+ is bound to the aromatic π ring (species l, Figure 2), while the less stable (by 3 kJ mol^{-1}) non-π mode has a $K^+ \cdots O$ interaction of 2.58 Å (species m, Figure 2). Similar to the Na^+–phenol system, the difference in EP(K^+) free energy of binding (ΔG_{298}) suggests that the non-π complex is favored over the π mode (by about 4 kJ mol^{-1}) in K^+–phenol.

We obtained an EP(K^+) PCA for phenol of 70 kJ mol^{-1}. Two experimental values are available for comparison: a recent threshold-CID value of 75 kJ mol^{-1},[8] and an earlier value of 84 kJ mol^{-1} based on results of radiative association kinetics measurements and density functional calculations.[43] We note that the latter value is anchored to the K^+–benzene affinity (at 298 K) of 80 kJ mol^{-1}.[26] However, recent experiments[8] suggested that the PCA at 298 K for benzene might need to be lowered to 74 kJ mol^{-1}. Adopting this lower PCA value of benzene as the anchoring point leads to a revised PCA value of 78 kJ mol^{-1} (corrected to 298 K). With this downward revision of the PCA of K^+–phenol from ref. [14], the two experimental values and our theoretical EP(K^+) affinity are now much more consistent.

Azines: Pyridine (C_5NH_5), pyridazine, pyrimidine, pyrazine (isomers of $C_4N_2H_4$), and 1,3,5-triazine ($C_3N_3H_3$) are six-membered nitrogen heterocycles. The presence of nitrogen atom(s) disturbs the symmetry of the π-electron distribution: charge is localized on the nitrogen atoms, and the resonance stabilization and aromatic character of the molecule are thus decreased.[5] Hence, for this class of ligands, K^+ prefers to bind to the nitrogen lone pair, with an average $K^+ \cdots N$ distance of 2.75 Å, rather than to the π cloud. Our present relative and absolute PCAs are in very good agreement with a combined experimental and theoretical study:[26] the MADs with respect to the reported experimental and BSSE-corrected MP2(full)/6-311 + G(2d,2p)//MP2(full)/6-31G(d) values are 2.5 (Table 1) and 1.5 kJ mol^{-1} (Table 2), respectively.

The modes of cation binding in 1,8-naphthyridine ($C_8N_2H_6$) and its PCA have not been reported previously. Our model (C_{2v}, species n, Figure 2) suggests that K^+ binds in a bidentate fashion to the two nitrogen atoms of the ligand. Because of the combined effect of polarizability and multidentate interaction, the PCA of 1,8-naphthyridine (136 kJ mol^{-1}) is much higher than those of the other pyridines, and it is at the top end of the PCA scale presented here.

Pyrrole and indole: Pyrrole (C_4NH_5) and indole (C_8NH_7) are important models for understanding the cation–π interactions in tryptophan-containing proteins. Unlike pyridine, the nitrogen atom in pyrrole is electron-deficient,[5] so the pyrrole nitrogen atom is not a favorable site for cation binding. As in the case of Na^+–pyrrole, the potassium cation prefers to bind to the π ring ($K^+ \cdots \pi$ distance of ca. 2.85 Å) of the ligand.

The experimental PCA for pyrrole was recently determined to be 85 kJ mol^{-1},[8] in good agreement with our theoretical PCA (77 kJ mol^{-1}). This is approximately 10 kJ mol^{-1} larger than the theoretical estimate for benzene, and can be attributed to the larger quadrupole and dipole moments of pyrrole.[40]

In the Na^+–indole complex, the cation can bind to either the benzo-π or pyrrole-π face of the ligand.[44] For the larger potassium cation, only one type of K^+–indole complex is found, in which the cation binds to the benzo-π face, with an estimated PCA of 89 kJ mol^{-1} (species o, Figure 2). Thus, even though the PCA of pyrrole is larger than that of benzene, it appears that K^+ binds exclusively to the benzo-π face of indole. This suggests that the distribution of π electrons is sufficiently altered in the polycyclic aromatic ligands that indole should not be regarded as fused benzene and pyrrole rings.[44, 45]

The experimental PCA of indole at 0 K was reported to be 910 kJ mol^{-1} in the study on K$^+$–phenol by Ryzhov and Dunbar.[29] In view of our discussions in the above section "Benzene, toluene, and phenol", this experimental PCA of indole may need to be lowered (by 6 kJ mol^{-1}) because of the anchoring value of benzene. The revised value is 99 kJ mol^{-1} (at 298 K), which brings it into closer agreement with our estimated PCA at 298 K for indole of 89 kJ mol^{-1}.

Azole: Azoles are building blocks for many antibiotics, anticancer agents, and other drugs.[30] Even though these five-membered heterocycles may be perceived as derivatives of pyrrole, the mode of potassium cation binding is different from that found in pyrrole. For the azoles studied here (except isoxazole), K$^+$ binds exclusively to the nitrogen lone pair, with an average K$^+$···N distance of 2.71 Å, comparable to that found for azines. No cation–π complexes were located, that is, all azoles favor σ binding interaction (except pyrrole).[6, 13] Two theoretical PCAs for imidazole ($C_3N_2H_4$) were reported previously.[6, 13] Our value is in good agreement with the more recent value, calculated at the MP2(full)/6-311+G(2d,2p)//MP2(full)/6-31G(d) level with BSSE correction.[13] The earlier value,[6] calculated at the HF/6-31G(d,p) level, is about 6 kJ mol^{-1} too large when compared to the EP(K$^+$) PCA.

In general, the PCA of azoles increases with increasing number of methyl substituent,[31] and with increasing number of ring nitrogen atoms, except for tetrazole.[6] We also estimated the PCAs of oxygen- and sulfur-containing azoles (isomers of C_3ONH_3: oxazole, isoxazole; isomers of C_3SNH_3: thiazole and isothiazole), which have not been reported previously. Isoxazole has a higher PCA (by about 10 kJ mol^{-1}) than oxazole, as K$^+$ is found in a bidentate fashion to both N and O atoms in isoxazole (species p, Figure 2), but solely to the N atom in oxazole (species q, Figure 2). On the other hand, thiazole and isothiazole have similar PCAs, as K$^+$ binds exclusively to the N atom in both ligands.

Our EP(K$^+$) PCA values at 0 K for 2*H*-1,2,3-triazole, 1*H*-1,2,4-triazole, and 2*H*-tetrazole are in good agreement (within ± 9 kJ mol^{-1}) with the previously reported experimental threshold-CID PCA values (MAD of 4 kJ mol^{-1}). However, our theoretical PCAs for the corresponding 1*H*-1,2,3-triazole, 4*H*-1,2,4-triazole, and 1*H*-tetrazole tautomers are significantly higher by 20–63 kJ mol^{-1}.[6] Our findings here are consistent with the rationalization put forward by Rodgers and Armentrout: in their threshold-CID experiments, binding of K$^+$ to the most stable tautomers of 2*H*-1,2,3-triazole, 1*H*-1,2,4-triazole, and 2*H*-tetrazole is kinetically favored, even though the resulting K$^+$ complexes are not the thermodynamically most stable K$^+$-bound species (see footnotes [y] and [z] to Table 1).[6]

Amino acids: Because of the presence of the acidic carboxyl group and basic amino group, amino acids can exist in two forms: charge-solvated (CS) and zwitterionic (ZW).[26, 32, 33, 34] The ZW form of amino acids is dominant in solution[35] and can be stabilized in the gas phase by binding to cations.[36] Similar to the case of glycine and alanine reported earlier,[26] our calculations show that for other aliphatic amino acids such as valine, leucine, and isoleucine, the potassium cation prefers to bind in a bidentate fashion to the carbonyl and hydroxyl oxygen atoms of the CS form.

Comparing our EP(K$^+$) PCA of glycine with that estimated by Hoyau and Ohanessian[37] shows their value to be too small by almost 8 kJ mol^{-1}; the difference most likely is due to the different basis sets employed. We are not aware of theoretical and experimental PCAs for the larger aliphatic amino acids in the literature. Using our theoretical G2(MP2,SVP)-ASC-(G-CP) PCAs at 0 K of acetamide (118.7 kJ mol^{-1}), *N*-methylacetamide (125.6 kJ mol^{-1}), and *N*,*N*′-dimethylacetamide (129.2 kJ mol^{-1}) as anchoring reference values,[31] we obtained the experimental PCA of aliphatic amino acids by the kinetic method, as shown in Table 1 (see also footnote [a′] to Table 1). The relative and absolute EP(K$^+$) PCAs of all five aliphatic amino acids were found to be in very good agreement with the quantitative values determined by the mass spectrometric kinetic method in the order[38] glycine < alanine < valine ≈ leucine ≈ isoleucine; the absolute PCAs are also within ± 1 kJ mol^{-1} of those determined experimentally.[38] We note that the anchoring value of 129.2 kJ mol^{-1} we used for *N*,*N*′-dimethylacetamide is very close to the experimental HPMS value of 129.3 kJ mol^{-1} reported by Kebarle et al.[23] Thus, the experimental set of PCAs for aliphatic amino acids is not only consistent with our EP(K$^+$) estimates, but it is also consistent with the reported experimental affinity of *N*,*N*′-dimethylacetamide.

We also obtained the PCAs of proline, serine, cysteine, and phenylalanine (species r, s, t, u, respectively, in Figure 2). The modes of K$^+$ binding for these ligands are similar to those of the corresponding Na$^+$ complexes.[34, 39] No experimental or theoretical PCAs are available in the literature, except for K$^+$–phenylalanine. The theoretical PCA reported by Ryzhov et al. (Table 2, 146.4 kJ mol^{-1}) is very close to our EP(K$^+$) PCA of 145.6 kJ mol^{-1}.[29, 40] Using the experimental PCAs of adenine (106 kJ mol^{-1}), cytosine (110 kJ mol^{-1}), and guanine (117 kJ mol^{-1}) as reference values (Table 1),[28] Ryzhov et al. reported a value of 104.2 kJ mol^{-1} (kinetic method) for the PCA of phenylalanine,[29] which is even lower than the threshold-CID value of K$^+$–glycine (126 kJ mol^{-1} at 298 K) reported by Kebarle et al.[9] This is counterintuitive, because phenylalanine is expected to have a higher PCA than glycine due to its greater polarizability and ion-induced dipole interactions. Hence, it is likely that the kinetic-method value for phenylalanine involves a relatively large margin of experimental uncertainty (> 20.9 kJ mol^{-1} /5 kcal mol^{-1}), as estimated by Ryzhov et al. (see section "Nucleobases" below for further discussions on the reference values). We have conducted detailed theoretical and experimental studies on the K$^+$–phenylalanine system, and our findings will be reported elsewhere.

Nucleobases: Given the biological importance of these ligands as models for cation–RNA/DNA interactions, several experimental[28, 29] and theoretical studies[9, 26, 27, 31] have been reported on the gas-phase alkali metal cation affinities of the five nucleobases adenine (A), thymine (T), uracil (U), cytosine (C), and guanine (G).

Experimental PCAs for all five nucleobases were obtained by Cerda and Wesdemiotis[28] using the extended mass

spectrometric kinetic method, and the PCAs of A, T, and U were determined by Rodgers and Armentrout[5] using the threshold CID method. There have been some concerns regarding the accuracy of the PCAs reported by Cerda and Wesdemiotis.[24] Firstly, Rodgers and Armentrout[6, 7] noted that the kinetic-method measurements were carried out under only two excitation conditions or effective temperatures. As a result, relative enthalpy (affinity) and entropy changes may not be extractable from the experiment in a statistically meaningful way. Secondly, under the experimental condition of the kinetic-method measurement,[25] several tautomeric forms of the free ligand may co-exist, and this would lead to different K^+-bound structures and affinities.[27] Depending on the energy barrier for interconversion between these tautomeric forms in the free ligand and the K^+-bound form, values determined by the kinetic method may or may not correspond to the PCA of a particular alkali metal cation complex of a nucleobase.[27] Such a complication is suggested to be the case for cytosine and guanine.[27] Moreover, it has also been noted that the values determined by the kinetic method for adenine could not be easily ascribed to the alkali metal cation affinity of any one of the tautomers.[27] Nevertheless, on careful examination, we consider that the PCAs of U and T determined by the kinetic method should be reliable. In the experiment, pyridine, aniline, and *n*-propylamine were used as reference compounds. As noted by Rodgers and Armentrout,[5] these three reference compounds and U/T all bind to K^+ in a monodentate fashion; thus, entropic effects in the kinetic-method measurement would be negligible. Moreover, for these two species, the PCA determined by the kinetic method is comparable to that determined by threshold CID. Taking all the above factors into consideration, we omitted the kinetic-method values for A/C/G, but retained those for U/T in the evaluation of our EP(K^+) protocol. We estimated the EP(K^+) affinities of the most stable K^+-bound structures of these ligands in the most stable free tautomeric forms. When compared to the experimental threshold-CID and selected kinetic-method values,[5] our EP(K^+) affinity is on average about 9.6 kJ mol^{-1} too large.

Comparing the EP(K^+) PCA with previously reported values calculated by ab initio MP2 and DFT methods,[8, 26] our present estimate is approximately 7–9 kJ mol^{-1} too large. In the case of the MP2 values,[8] the difference can be at least partly attributed to BSSE corrections. As an example, for K^+–uracil, the reported BSSE for the MP2-based model is 4 kJ mol^{-1},[8] which accounts for about 50% of the difference. We note here again that our EP(K^+) model does not include the BSSE correction, but for this species, the EP(K^+) BSSE correction is calculated to be small (only 0.8 kJ mol^{-1}; see Supporting Information, Table S2).

However, the differences between our values and the DFT-based estimates by Russo et al.[10] cannot be easily explained. The two DFT protocols, even though not identical, are expected to be comparable, and indeed we found evidence for this in the PCAs of water and ammonia (see discussions in the respective sections above). Using K^+–uracil as an example, we tried to identify the source of the discrepancies. We found that the geometries are similar: for this complex, the calculated $K^+\cdots O$ distance at the B3-LYP/6-31G(d) level (2.465 Å) is only 0.030 Å shorter than that obtained at the B3-LYP/6-311+G(2df,2p) level. The zero-point corrections are also similar: the B3-LYP/6-31G(d) level ZVPE is only 0.4 kJ mol^{-1} larger than that at the B3-LYP/6-311+G(3df,3p) level. To resolve the differences, we attempted to obtain the PCA of uracil using the model reported by Russo et al. (B3-LYP/6-311+G(2df,2p)//B3-LYP/6-311+G(2df,2p) with BSSE correction). A PCA of 113.8 kJ mol^{-1} (without BSSE correction) was obtained, which is very close to our EP(K^+) PCA of 113.1 kJ mol^{-1}. Applying the BSSE correction reduced the estimate from 113.8 to 113.0 kJ mol^{-1}, which still differs from the value reported by Russo et al. (108.1 kJ mol^{-1}, corrected to 298 K) by 4.9 kJ mol^{-1}.[26] In conclusion, it appears that DFT-based protocols could be overestimating the PCAs for the nucleobases. Given these uncertainties, further theoretical studies and experimental measurements on guanine, cytosine, and adenine are clearly required.

Comparison of PCAs with lithium and sodium cation affinity scales: Our theoretical PCA obtained by using the EP(K^+) protocol are plotted against the reported theoretical Li^+ affinities[28] (calculated at the B3-LYP/6-311+G(d,p) level) and Na^+ affinities[6, 10, 11, 24, 27] in Figure 3. The calculations in refs. [6, 10, 11, 24, 27] and related studies were all carried out at the MP2(full)/6-311+G(2d,2p)//MP2/6-31G(d) level, and hence the theoretical Na^+ affinity values from these studies were pooled together as one set and correlated with the EP(K^+) PCAs (denoted Na^+-1 in Figure 3). However, the theoretical Na^+ affinities reported by Petrie[24] were calculated at a different level of theory (CPd-G3theory),[29] and hence were correlated separately with the EP(K^+) PCAs (denoted Na^+-2 in Figure 3).

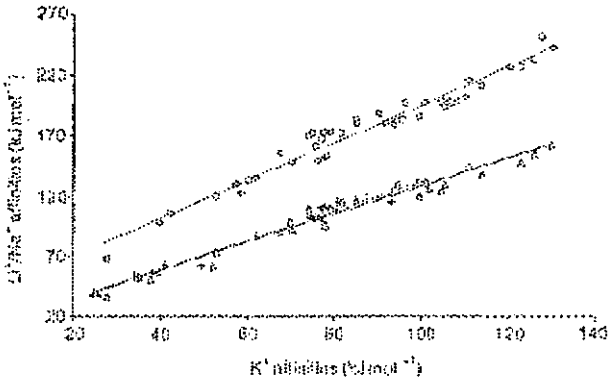


Figure 3. Correlation of PCAs with theoretical Li^+ affinities reported in ref. [28] (298 K, ○) and theoretical Na^+ affinities reported in refs. [6, 10, 11, 24, 27] and related studies (298 K, △, Na^+-1) and ref. [12] (0 K, +, Na^+-2).

The PCAs correlate linearly with Li^+ and Na^+ affinities [Eqs. (2)–(4)].

$$\Delta H_r(Li^+\text{-L}) = 1.55\,PCA + 40.83 \quad (R^2 = 0.96) \tag{2}$$

$$\Delta H_r(Na^+\text{-L}) = 1.16\,PCA + 12.84 \quad (R^2 = 0.93) \tag{3}$$

$\Delta H_{rxn}(Na^+-L) = 1.22PCA + 19.48$ ($R^2 = 0.97$) (4)

Such good linear correlation indicates that the nature of interactions between Li^+, Na^+, and K^+ and the wide range of ligands studied here are indeed very similar. Furthermore, we note that the correlation relations obtained for Na^+-1 [Eq. (3)] and Na^+-2 [Eq. (4)] differ only in the intercept. In other words, different theoretical models are likely to yield the same relative affinity scale, but the absolute affinities obtained may be different. This highlights the importance of obtaining a set of absolute theoretical PCAs that is consistent with experimental values in order to minimize the presence of systematic errors.

Effect of substituents: We now comment on the effect of substituents on the binding mode and the PCA relative to the unsubstituted parent ligands.

In the case of substituents without electronegative atoms (e.g., alkyl groups), the cation is not bound to these groups. Hence, the presence of these substituents affects the PCA but not the binding site. For aromatic ligands such as pyridine and azoles, methylation increases the PCA. Compared to *o* and *m* substitution, the effect of *p* substitution on PCA is most significant in pyridine, and has been attributed to the larger dipole moment of *p*-methylpyridine.[31]

Using classical electrostatic theory, Davidson and Kebarle suggested that alkyl substitution affects four properties of a ligand: its polarizability, dispersion, intramolecular repulsion (between ligand and ion), and dipole moment.[16] While successive alkyl substituents increase the binding affinity of a ligand by enhancing the polarizability and dispersion component of the cation - ligand interactions, it also increases the repulsion and decreases the dipole moment of the ligand and thus leads to a decrease in binding affinities. Here, we studied the effect of successive methylation at the O/N/S binding sites on the PCAs of water, ammonia, formamide/acetamide, and hydrogen sulfide.

In the H_2O/MeOH/Me_2O series, the first methylation increases the PCA by about 5 kJ mol^{-1}, while slight decrease in PCA is found for the second methylation. In the H_2S/MeSH/Me_2S series, the PCAs are in the order of H_2S < MeSH < Me_2S. Similar observations were made for the corresponding theoretical sodium cation affinities for both series[24, 25] and were rationalized in terms of opposing effects of changes in ligand polarizability (increase) and dipole moment (decrease).[25] While similar rationales can be applied to explain the observed trends in PCAs, we would like to point out that repulsive effects may also play a role here. For the smaller oxygen atom, the repulsive steric effect of successive methylation would be much more strongly felt than for the larger sulfur atom. Thus, it may not be surprising that while the PCA increases with increasing methylation in the H_2S series, it tails off or decreases in the H_2O series.

For the ammonia series, our theoretical results show that the effect of multiple methyl substitutions on the PCA of NH_3 is small, spanning a range of only 5 kJ mol^{-1}. Similar to the H_2O series, the theoretical PCA increases with the first methylation (by 2 kJ mol^{-1} from NH_3 to $MeNH_2$), but decreases on second and third methylations. Interestingly, earlier experimental FTMS results suggest that successive methyl substitution increases the PCA, that is, NH_3 < $MeNH_2$ < Me_2NH < Me_3N.[26] As the difference in PCA for successive methylations is small and can be considered to lie within the expected error limits of theoretical models, we only note here that our EP(K^+) affinities are more in line with the recent reported trends for the experimental (FT-ICR) and theoretical (MP2) sodium cation affinities[27] and our B3-P86 affinities (Supporting Information, Table S1) differ from the prediction at the G2(MP2,SVP) level (Supporting Information, Table S1). Thus, further calculations and experimental measurements may be needed to resolve the difference in qualitative trends found between experimental and theoretical PCAs of ammonia and its methyl-substituted derivatives.

For electronegative (e.g., fluoro, chloro) or electron-rich substituents (e.g., aromatic π rings), not only are the binding affinities affected, but the presence of these substituents also opens up new modes of binding that are not possible in the parent ligand. We found that the F and Cl substituents are in general not competitive with O or N binding sites already present. Hence, for most classes of ligands (e.g., carboxylic acids, aldehydes, ketones, nitriles), the binding modes remain the same as in the parent ligand on halogenation. In general, the PCA of halogenated ligands decreases, as these electron-withdrawing groups decrease the dipole moment of the ligand. However, in a few cases (e.g., CF_3CH_2OH, species v, Figure 2) when the halogen atom is close to the original binding site of of the parent ligand, it offers an additional binding site for K^+. In these cases, the PCA is increased relative to the unsubstituted parent ligands.

For aromatic π substituents, the PCA increases in all cases, as polarizability of the ligand is greatly enhanced by the presence of the highly polarizable phenyl ring. In some cases, the aromatic π substituents (e.g., from alanine to phenylalanine) are also involved in binding, and this leads to a further increase in PCA over that expected solely from the polarizability effect.

Relating PCAs to properties of ligands: In the previous section, the effect of substituents on PCA is discussed in a qualitative manner. Here, we take the discussion one step further by establishing quantitative relations between the PCA of the 136 ligands and their properties. Our aim is to use molecular properties which are readily available and accessible, so that the PCAs can be predicted with relative ease.

As the K^+–ligand interaction is mainly electrostatic in nature, ligand properties such as dipole moment and polarizability are expected to be important. The number of interactions (coordination number or denticity) and to which type of atom(s) K^+ binds should also be important. Based on the goodness of fit (in terms of adjusted R^2) from multiple linear-regression analyses, four parameters were found to be important in governing the PCA of a ligand: the dipole moment μ in Debye and the polarizability α in Å^3 of the ligand, the number n_1 of first-row atoms the K^+ ion interacts with, and the number n_2 of second-row atoms the K^+ interacts with. The PCA is related to these four parameters by Equation (5).

$$PCA = 10.2\mu + 3.6\alpha + 14.1n_l - 11.8n_r + 29.6 \quad \text{adjusted } R^2 = 0.82 \quad (5)$$

Figure 4 shows the relation between the calculated PCA (Table 1) and the PCA predicted by Equation (5). Clearly, several points show large deviations from the ideal correlation

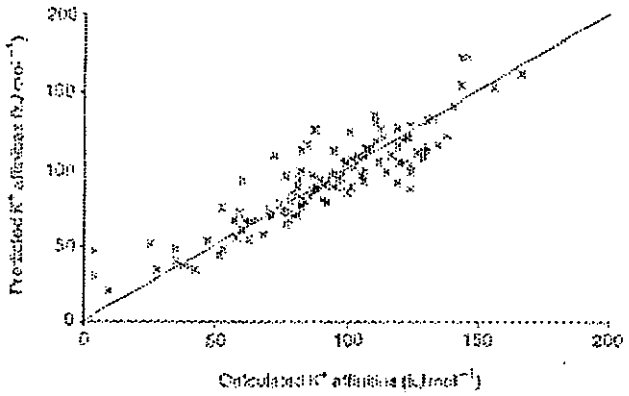


Figure 4. Plot of predicted PCAs [Eq. (5)] against BP(K^+) PCAs; the diagonal line with a slope of 1.0 is drawn for reference purposes.

line with a slope of unity. At the lowest end of the PCA scale, the error can be very large: in the case of Ne, as large as 12-fold (calculated PCA 4 kJ mol^{-1}, predicted PCA 47 kJ mol^{-1}). Such a deviation partly arises from the simplicity of the correlation equation employed, but is also due to the small numerical PCA value for this ligand.

It is clear that our proposed model is quite crude and has neglected other effects such as ion–local dipole interactions.[9] Despite the crudeness of the model, the equation can yield reasonable estimates of PCA. If the four lowest PCA values (He, Ne, Ar, and CO) are ignored, the MAD for the remaining 132 ligands is then reduced to 12 kJ mol^{-1} (error of 10%), with a maximum of 38 kJ mol^{-1} for CF_3CH_2OH (error of 52%). We note that substituting the polarizability by the molecular weight (MW, in g mol^{-1}) of the ligand also yields a reasonably good correlation [Eq. (6)].

$$PCA = 12.5\mu + 0.1MW + 15.4n_l - 11.6n_r + 40.0 \quad (6)$$

Compared to Equation (5), the adjusted R^2 is somewhat lower (0.74). However, as the molecular weight of a ligand is a parameter that can be more readily obtained than the polarizability, Equation (6) in fact provides a simpler alternative for estimating the PCA of a ligand.

Conclusion

We have reported the theoretical potassium cation affinities (PCA) of 156 ligands, spanning a range from 4 to 166 kJ mol^{-1}. Of these 156 ligands, 70 experimental and 64 theoretical values reported in the literature are available for comparison. We found that our theoretical estimates and most of the experimental affinities are in good general agreement (within ± 10 kJ mol^{-1}). Based on our theoretical BP(K^+) values, we were able to conduct a critical evaluation of the reported values for Me_2SO, $MeCONMe_2$, and phenol obtained by different experimental techniques, for which PCA differences of more than ± 10 kJ mol^{-1} have been reported. Large discrepancies (>20 kJ mol^{-1}) were found in the case of phenylalanine,[11] cytosine,[24] guanine,[24] and adenine.[24] However, in all these cases, the discrepancies likely arise from complications in the experimental measurements. Ignoring these four values and the PCA of Me_2SO determined by HPMS, the mean absolute deviation of our theoretical PCA from the remaining experimental values is 4.5 kJ mol^{-1}. Our BP(K^+) PCA is also consistent with most of the previously reported theoretical values to within ± 5 kJ mol^{-1}. For species with larger differences, we are able to account for the difference in terms of the different basis sets used in the theoretical calculations and/or basis set superposition errors. However, the origin of the rather large difference of 7–9 kJ mol^{-1} found between our values and those reported by Russo et al.[24] for the DNA/RNA nucleobases remains unclear.

The effects of substitution on PCAs of parent ligands are also discussed. For a halogenated ligand, the PCA decreases in general, except when the halogen is close to the original binding site. For aromatic substituents, the PCA increases in all cases, as the polarizability of the ligand is greatly enhanced by the presence of the highly polarizable phenyl ring. First methylation tends to increase the PCA, while the PCA may decrease upon further methylation.

We have also compared the PCAs with lithium and sodium cation affinities previously reported in the literature. The excellent linear correlation that was found indicates that the nature of the interactions between alkali metal cations (Li^+, Na^+, and K^+) and the wide range of ligands studied here are indeed very similar. Thus, such relations [Eqs. (2)–(4)] allow estimation of PCA for ligands with known Li^+ and/or Na^+ affinities, in particular where the mode of binding for K^+ is not expected to differ from those of the smaller Li^+/Na^+.

Finally, we established two correlation equations [Eqs. (5) and (6)] relating PCAs of ligands with their properties (dipole moment, polarizability, molecular weight, and number of interactions). These two equations offer relatively simple and efficient methods of estimating the PCA of ligands not reported here.

Acknowledgement

P.L.M. thanks the Institute of High Performance Computing for generous allocation of supercomputer time. The funding support of the Hong Kong Polytechnic University (Project No. GV-548 to P.M.S., GW-021 to C.H.S.W., and Area of Strategic Development Fund Project No. A024 to C.W.T.), and the Research Grant Council of Hong Kong (Area of Excellence Project No. P-10/2001 and CE-RG Project No. 5303/01P to C.W.T.) are gratefully acknowledged.

[1] a) M. N. Hughes, *The Inorganic Chemistry of Biological Processes*, 2nd ed., Wiley, New York, 1972; b) S. J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, California, 1994; c) L. Stryer, *Biochemistry*, W. H. Freeman, New York, 1995

[1] J. S. Klassen, S. G. Anderson, A. T. Blades, P. Kebarle, *J. Phys. Chem.* **1996**, *100*, 14218.
[2] M. B. Meuwly, D. Bojc, P. B. Armentrout, *J. Phys. Chem. A* **1997**, *101*, 4354.
[3] D. Walker, M. R. Nevers, P. B. Armentrout, *Int. J. Mass Spectrom. Ion Processes* **1998**, *175*, 93.
[4] M. T. Rodgers, P. B. Armentrout, *Int. J. Mass Spectrom.* **1999**, *185/186/187*, 359.
[5] K. Ariyasinghe, M. T. Rodgers, *Int. J. Mass Spectrom.* **2006**, *195/196*, 459.
[6] M. T. Rodgers, P. B. Armentrout, *Mass Spectrom. Rev.* **2000**, *19*, 215.
[7] J. C. Amicangelo, P. B. Armentrout, *J. Phys. Chem. A* **2000**, *104*, 11420.
[8] M. T. Rodgers, P. B. Armentrout, *J. Am. Chem. Soc.* **2000**, *122*, 8548.
[9] P. B. Armentrout, M. T. Rodgers, *J. Phys. Chem. A* **2000**, *104*, 2238.
[10] M. T. Rodgers, *J. Phys. Chem. A* **2001**, *105*, 2374.
[11] H. Huang, M. T. Rodgers, *J. Phys. Chem. A* **2002**, *106*, 4272.
[12] K. Ariyasinghe, M. T. Rodgers, *J. Phys. Chem. A* **2002**, *106*, 9715.
[13] V. Ryzhov, R. C. Dunbar, *J. Am. Chem. Soc.* **1999**, *121*, 2259.
[14] I. Dzidic, P. Kebarle, *J. Phys. Chem.* **1970**, *74*, 1466.
[15] W. R. Davidson, P. Kebarle, *Can. J. Chem.* **1976**, *54*, 2384.
[16] W. R. Davidson, P. Kebarle, *J. Am. Chem. Soc.* **1976**, *98*, 6133.
[17] W. R. Davidson, P. Kebarle, *J. Am. Chem. Soc.* **1976**, *98*, 6125.
[18] A. W. Castleman, Jr., *Chem. Phys. Lett.* **1978**, *53*, 560.
[19] J. Sunner, K. Nishizawa, P. Kebarle, *J. Phys. Chem.* **1981**, *85*, 1814.
[20] J. Sunner, P. Kebarle, *J. Am. Chem. Soc.* **1984**, *106*, 6135.
[21] R. G. Keesee, A. W. Castleman, Jr., *J. Phys. Chem. Ref. Data* **1986**, *15*, 1011.
[22] a) S. C. Guo, A. W. Castleman, Jr., *Chem. Phys. Lett.* **1991**, *181*, 16; b) P. Kebarle, *J. Mass Spectrom.* **1997**, *32*, 922.
[23] S. Hoyau, K. Norrman, T. B. McMahon, G. Ohanessian, *J. Am. Chem. Soc.* **1999**, *121*, 8864.
[24] M. Alcami, M. Y. O. Mo, F. Anvia, R. W. Taft, *J. Phys. Chem.* **1990**, *94*, 4196.
[25] P. Burk, I. A. Koppel, I. Koppel, R. Kurg, J.-F. Gal, P.-C. Maria, M. Herreros, R. Notario, J.-L. M. Abboud, F. Anvia, R. W. Taft, *J. Phys. Chem. A* **2000**, *104*, 2824.
[27] T. B. McMahon, G. Ohanessian, *Chem. Eur. J.* **2000**, *6*, 2931.
[28] A. Gapeev, R. C. Dunbar, *J. Am. Chem. Soc.* **2001**, *123*, 8560.
[29] B. A. Cerda, C. Wesdemiotis, *J. Am. Chem. Soc.* **1996**, *118*, 11884.
[30] a) R. G. Cooks, P. S. H. Wong, *Acc. Chem. Res.* **1998**, *31*, 379; b) R. G. Cooks, J. T. Koskinen, P. D. Thomas, *J. Mass Spectrom.* **1999**, *34*, 85.
[31] V. Ryzhov, R. C. Dunbar, B. Cerda, C. Wesdemiotis, *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1037.
[32] N. L. Ma, F. M. Siu, C. W. Tsang, *Chem. Phys. Lett.* **2000**, *322*, 65.
[33] F. M. Siu, N. L. Ma, C. W. Tsang, *J. Chem. Phys.* **2001**, *114*, 7365.
[34] F. M. Siu, N. L. Ma, C. W. Tsang, *Chem. Phys. Lett.* **1998**, *285*, 408.
[35] S. Hoyau, G. Ohanessian, *Chem. Eur. J.* **1998**, *4*, 1561.
[36] N. Russo, M. Toscano, A. Grand, *J. Phys. Chem. B* **2001**, *105*, 4735.
[37] N. Russo, M. Toscano, A. Grand, *J. Am. Chem. Soc.* **2001**, *123*, 10272.
[38] F. M. Siu, N. L. Ma, C. W. Tsang, *J. Am. Chem. Soc.* **2001**, *123*, 2317.
[39] a) C. H. S. Wong, N. L. Ma, C. W. Tsang, *Chem. Eur. J.* **2002**, *8*, 4099; b) C. H. S. Wong, F. M. Siu, N. L. Ma, C. W. Tsang, *J. Mol. Struct. (Theochem)* **2002**, *588*, 9.
[40] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.
[41] J. E. Del Bene, H. D. Mettee, M. J. Frisch, B. T. Luke, J. A. Pople, *J. Phys. Chem.* **1983**, *87*, 3279.
[42] S. Feller, *J. Phys. Chem. A* **2001**, *105*, 9851.
[43] S. Feller, *Chem. Phys. Lett.* **1984**, *109*, 550.
[44] E. Magnusson, *J. Phys. Chem.* **1984**, *88*, 12558.
[45] R. C. Dunbar, *J. Phys. Chem. A* **2000**, *104*, 8067.
[46] S. Ahlrichs, P. L. Ma, N. K. Goh, *Chem. Phys. Lett.* **2002**, *359*, 500.
[47] C. Lifshitz, P. B. Armentrout, *Int. J. Mass Spectrom.* **2003**, *222*, 329.
[48] N. L. Ma, W. K. Li, C. Y. Ng, *J. Chem. Phys.* **1993**, *99*, 3617.
[49] J. E. Del Bene, *J. Phys. Chem.* **1968**, *100*, 6384.
[50] a) M. Feierabend, C. M. Marian, S. D. Peyerimhoff, *Chem. Phys. Lett.* **1983**, *304*, 50; b) F. Schauer, H. J. Plöd, *J. Chem. Phys.* **1998**, *110*, 11201.
[51] J.-L. M. Abboud, I. Alkorta, I. Z. Davalos, J.-F. Gal, M. Herreros, P.-C. Maria, O. Mó, M. T. Molina, R. Notario, M. Yáñez, *J. Am. Chem. Soc.* **2000**, *122*, 4457.
[52] M. T. Rodgers, P. B. Armentrout, *J. Phys. Chem. A* **1999**, *103*, 2614.
[53] S. B. Nielsen, N. Masella, P. Kebarle, *J. Phys. Chem. A* **1999**, *103*, 9891.
[54] Y. Tang, F. M. Siu, N. L. Ma, C. W. Tsang, Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics (Poster TPB 044), May 27–31, 2001, Chicago, USA, and unpublished results.
[55] S. Meot-ner, A. F. Wexler, Jr., D. A. Dougherty, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10586.
[56] a) D. A. Dougherty, *Science* **1996**, *271*, 163; b) J. C. Ma, D. A. Dougherty, *Chem. Rev.* **1997**, *97*, 1303.
[57] a) P. von R. Schleyer, P. K. Freeman, H. Jiao, B. Goldfuss, *Angew. Chem.* **1995**, *107*, 332; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 337; b) A. R. Katritzky, M. Karelson, S. Sild, T. M. Krygowski, K. Jug, *J. Org. Chem.* **1998**, *63*, 5228.
[58] B. Kirsch, A. K. Pleshan, E. D. Jemmis, *Inorg. Chem.* **2001**, *40*, 3615.
[59] a) W. Hartshorger, G. Lee, R. E. Porter, S. H. Bauer, *Inorg. Chem.* **1968**, *8*, 1063; b) J. K. Parkes, S. R. Davis, *J. Phys. Chem. A* **1997**, *101*, 9410; c) B. Chiavarino, M. E. Crestoni, A. D. Marchis, S. Fornarini, M. Rosi, *J. Am. Chem. Soc.* **1999**, *121*, 11244.
[60] K. M. Pi, N. L. Ma, C. W. Tsang, *Rapid Comm. Mass Spectrom.* **1998**, *12*, 1679.
[61] R. J. Datellion, A. J. Thakkar, *J. Phys. Chem. A* **1999**, *103*, 10009.
[62] R. C. Dunbar, *J. Phys. Chem. A* **2002**, *106*, 7328.
[63] R. J. Fessenden, J. S. Fessenden, *Organic Chemistry*, 6th ed., W. Grant Press, Boston, MA, **1998**.
[64] R. C. Dunbar, *J. Phys. Chem.* **1998**, *102*, 8946.
[65] a) J. D. Cartledge, J. Midgley, M. Perreau, D. Shanson, B. G. Gazzard, *J. Antimicrob. Chemother.* **1987**, *40*, 517; b) P. C. Hogaweil, *Sem. Resp. Crit. Care Med.* **1987**, *38*, 271; c) J. D. Cartledge, P. W. Denning, B. Dupont, N. Clumeck, S. Dewit, J. Midgley, D. A. Hawkins, B. G. Gazzard, *AIDS* **1998**, *12*, 411.
[66] a) T. Marino, N. Russo, M. Toscano, *J. Inorg. Biochem.* **2000**, *79*, 179; b) N. S. Pelkkine, M. Noguera, L. Rodriguez-Santiago, M. Sodupe, I. Bertran, *Chem. Eur. J.* **2000**, *6*, 4305.
[67] E. P. Sirilumatter, A. S. Lascoff, E. R. Williams, *J. Phys. Chem. A* **2006**, *104*, 9793.
[68] T. Wyttenbach, M. Witt, M. T. Bowers, *J. Am. Chem. Soc.* **2000**, *122*, 3458.
[69] O. Wada, E. Tamura, M. Okina, M. Nakamura, *Bull. Chem. Soc. Jpn.* **1982**, *55*, 3064.
[70] C. W. Tsang, Y. Tang, C. H. S. Wong, N. L. Ma, Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics (Poster TPB 045), May 27–31, 2001, Chicago, USA, and unpublished results.
[71] J. V. Burda, J. Sponer, P. Hobza, *J. Phys. Chem.* **1996**, *100*, 7250.
[72] P. B. Armentrout, *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 371.

Received: December 18, 2002
Revised: March 24, 2004 [F4673]

Chem. Eur. J. 2003

Supporting Information

for

Absolute Potassium Cation Affinities (PCAs) in the Gas Phase

Justin Kai-Chi Lau, Carrie Hoi Shan Wong, Po Shan Ng, Fung Ming Siu, Ngai Ling Ma and Chun Wai Tsang

Table S1. The B3-P86 and G2(MP2,SVP) Potassium Cation Binding Affinities (kJ mol $^{-1}$) at 298K.

Molecule[a]	B3-P86	G2(MP2,SVP)
He	2.8	
Ne	2.4	
Ar	8.4	
CO	24.6	
HF	49.8	
HCl	26.9	
P_4	33.6	
PH_3	42.4	
C_2H_2	37.6	
C_2H_4	35.0	
H_2S	39.3	
MeSH	52.5	
EtSH	57.2	

n-PrSH	59.0	
i-PrSH	60.0	
n-BuSH	60.3	
i-BuSH	63.1	
t-BuSH	61.7	
Me_2S	61.8	
Et_2O	69.9	70.3
MeOH	74.2	76.7
EtOH	80.2	83.6
n-PrOH	81.1	83.5
i-PrOH	84.1	87.5
n-BuOH	84.7	88.2
i-BuOH	82.1	86.0
s-BuOH	86.4	89.8
t-BuOH	87.1	91.3
1,2-Propanediol	114.0	119.1
1,3-Propanediol	118.7	123.8
Ethylene glycol	114.7	119.3
Glycerol	127.9	136.4
CF_3CH_2OH	69.7	
CCl_3CH_2OH	74.7	
Me_2O	73.2	
Et_2O	83.9	
$(MeOCH_2)_2$	122.1	
1,2-dioxane	82.3	

1,3-dioxane	82.3	
1,4-dioxane	69.3	
HCHO	76.3	
MeCHO	92.6	
EtCHO	94.6	
n-PrCHO	96.1	
n-BuCHO	97.3	
CF_3CHO	58.4	
CCl_3CHO	73.1	
Me_2CO	103.7	
MeCOEt	105.0	
NH_3	77.8	76.7
$MeNH_2$	79.3	80.5
Me_2NH	77.6	81.2
Me_3N	74.8	80.3
$EtNH_2$	81.7	84.1
$n\text{-}PrNH_2$	81.6	85.6
HCN	77.6	
MeCN	101.4	
EtCN	103.0	
n-PrCN	106.4	
i-PrCN	107.7	
t-BuCN	109.9	
$PhCH_2CN$	109.7	
CF_3CN	58.4	

CCl_3CN	76.0	
HCO_2H	79.6	
HCO_2Me	88.6	
$MeCO_2H$	91.0	
HCO_2Et	93.7	
HCO_2n-Pr	95.3	
$MeCO_2Me$	98.2	
$MeCO_2Et$	104.4	
$EtCO_2Me$	99.7	
CF_3CO_2Me	80.7	
$ClCO_2Me$	79.8	
SO_2	49.8	
Me_2SO	126.7	
PhSOMe	130.0	
$HCONH_2$	113.3	113.2
HCONHMe	118.8	120.9
$HCONMe_2$	124.6	126.6
$MeCONH_2$	122.5	122.2
MeCONHMe	126.6	129.0
$MeCONMe_2$	128.7	131.7
Benzene	70.8	78.6
Borazine	48.8 (48.7)[b]	57.6
Phenol	73.7 (73.6)[b]	80.7
Pyridine	92.6	
2-Me-pyridine	94.2	

3-Me-pyridine	98.6	
4-Me-pyridine	100.7	
2-F-pyridine	98.6	
3-Cl-pyridine	81.5	
1,8-naphthyridine	155.4	
Pyridazine	129.6	
Pyrimidine	75.0	
Pyrazine	69.1	
1,3,5-trinzine	55.8	
Pyrrole	79.9 (79.3) [b]	87.0
Indole	93.2 (93.0) [b]	102.8
Pyrazole	89.8	
Imidazole	111.0	
Thiazole	86.8	
Isothiazole	85.2	
Oxazole	83.3	
Isoxazole	92.6	
1-Me-pyrazole	93.9	
3-Me-pyruzole	92.2	
4-Me-pyrazole	95.8	
1,4-Me_2-pyrazole	99.5	
1,5-Me_2-pyrazole	100.9	
1,3,5-Me_3-pyrazole	103.1	
3,4,5-Me_3-pyrazole	104.8	
1,3,4,5-Me_4-pyrazole	105.9	

1-Me-imidazole	118.4	
1,2-Me$_2$-imidazole	120.2	
2,4,5-Me$_3$-imidazole	122.2	
2H-1,2,3-triazole	63.7	
1H-1,2,4-triazole	86.8	
2H-tetrazole	88.3	
1H-tetrazole	109.8	
4H-1,2,4-triazole	140.3	
1H-1,2,3-triazole	118.7	
Glycine	122.1 (121.4)[b]	118.3
Alanine	128.0 (127.4)[b]	122.5
Valine	131.9 (131.5)[b]	126.5
Leucine	132.8 (132.6)[b]	127.3
Iso-Leucine	133.5 (133.0)[b]	128.5
Proline	142.3	
Serine	136.3	
Cysteine	121.4	
Phenylalanine	144.0	
Adenine	86.6[b]	
Thymine	110.9[c]	
Uracil	111.9[c]	
Cytosine	166.9[c]	
Guanine	143.7[c]	

[a] Abbreviations: Me = $-CH_3$, Et = $-C_2H_5$, n-Pr = $-C_3H_8$, i-Pr = $-(CH_3)_2CH$, n-Bu= $-C_4H_{11}$, i-Bu = $-(CH_3)_2CHCH_2$, t-Bu= $-(CH_3)_3C$, Ph= $-C_6H_5$. [b] The value in bracket is the PCA calculated using B3-P86/6-31G(d) geometry, i.e., at the B3-P86/6-311+G(3df,2p)//B3-P86/6-31G(d) level. The similarity between the two sets of PCAs suggests that the effect of geometry difference: (B3-LYP/6-31G(d) versus B3-P86/6-31G(d)), on binding affinity is negligible. [c] The PCA is estimated for the most stable tautomer of the free ligands, corresponding to species A1, T1, U1, C1 and G1 in ref. [37].

Table S2. The effect of B3-LYP/6-31G(d) correction on zero point vibrational energies, temperature corrections, and the effect of full counterpoise correction.

Species	$\Delta(ZVPE)^{[a]}$	$\Delta(H_{298})^{[b]}$	$\Delta(BSSE)^{[c]}$
CO	-0.39	-0.08	-0.64
C_2H_2	-0.77	-0.45	-0.39
MeSH	-0.42	-0.14	-0.61
H_2O	-0.58	-0.42	-0.53
HCHO	-0.50	-0.41	-0.63
NH_3	-0.83	-0.56	-0.30
HCN	-0.26	-0.12	-0.41
HCO_2H	-0.45	-0.36	-0.65
SO_2	-0.15	-0.11	-0.81
Benzene	-0.30	-0.14	-1.77
Pyridine	-0.49	-0.32	-0.64
Pyrrole	-0.47	-0.21	-1.18
Glycine	-0.52	0.47	-0.75
Uracil[d]	-0.61	-0.49	-0.76

[a] Difference in zero-point vibrational energy corrections (ZVPE), Δ(ZVPE) = ZVPE calculated at the B3-LYP/6-31G(d) level — ZVPE calculated at the HF/6-31G(d) level, in kJ mol^{-1}. [b] Difference in thermal corrections, $\Delta(H_{298})$ = temperature correction calculated at the B3-LYP/6-31G(d) level — temperature correction at the HF/6-31G(d) level, in kJ mol^{-1}. [c] Δ(BSSE) = full counterpoise correction, including effect of geometry deformation, ref. [33], at the B3-LYP/6-311+G(3df,2p) level of theory, in kJ mol^{-1}, based on B3-LYP/6-31G(d) geometries. [d] The PCA is estimated for the most stable tautomer of the free ligands U1 in ref. [37].