COMPUTATIONAL THERMODYNAMICS

OF GAS AND SOLUTION PHASE ANIONS

by

ASHLEY SHARI MCNEILL

DAVID A. DIXON, COMMITTEE CHAIR
MICHAEL K. BOWMAN
CAROLYN J. CASSADY
VINCENT F. SCALFANI
SHANE C. STREET

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Chemistry and Biochemistry
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2020

Copyright Ashley Shari McNeill 2020
ALL RIGHTS RESERVED

ABSTRACT

The work in this dissertation focuses on the computational analysis of the thermodynamics of anions in the gas phase and in aqueous solution to provide unique insights into the chemistry of a range of biologically and geochemically relevant chemical species. This often involves calculating properties for these species such as electron affinities and hydration free energies of the anions, which can be difficult or impossible to obtain experimentally. Systems of interest in this work include small peptides, enzyme-catalyzed biological reactions, and the gas phase and solvation energetics of a variety of anionic species including CO_2^-, H^-, X^- (halides), OX^- (hypohalites), and YH^- (chalcogen hydrides). The peptide work, performed largely with the composite correlated molecular orbital theory G3(MP2) method, is compared directly to experiments conducted with low-energy collision-induced dissociation negative ion mode mass spectrometry. Isotope fractionation studies, of significant use in many geochemical applications, are conducted on the overall reaction by the alanine transaminase enzyme ($^+H_3NCH(CH_3)COO^-$ + $^-OOCCH_2CH_2C(O)COO^-$ ↔ $CH_3C(O)COO^-$ + $^+H_3NCH(CH_2CH_2COO^-)COO^-$) in order to predict that ^{13}C preferentially collects in the C2 site of pyruvate over alanine by 9‰ at equilibrium. This prediction, calculated from gas phase- and aqueous-optimized clusters with explicit H_2O molecules at the MP2/aug-cc-pVDZ with and without the COSMO self-consistent reaction field for implicit solvation, is reflected in simpler models: without explicit solvation, with simpler analogues formaldehyde and methylamine, and from canonical functional group frequencies and reduced masses for $R_2C{=}O$ and $R_2CH\text{-}NH_2$. Solvation studies of the CO_2^-, H^-,

X^-, OX^-, and YH^- anions and corresponding neutrals gave adiabatic electron affinities, reduction potentials, and gas phase and aqueous acidities that are generally in excellent agreement with experiment. These studies used a variety of computational methods, including heavy application of coupled cluster calculations with the Feller-Peterson-Dixon method to obtain high accuracy thermodynamic values. Absolute hydration free energies are determined for neutral and anionic species clustered with 4 to 8 explicit H_2O molecules using a supermolecule-continuum approach.

DEDICATION

This dissertation is dedicated to my family, my friends, and the many, many wonderful educators I have had the good fortune to know and work with over the years. I am inspired by all of you to be the best version of me.

LIST OF ABBREVIATIONS AND SYMBOLS

α	isotope fractionation factor
A	alanine
AEA	adiabatic electron affinity
AIMD	ab initio molecular dynamics
AIMD/MM	combined ab initio molecular dynamics and classical molecular dynamics
ALT	alanine transaminase enzyme
AO	atomic orbital
ATcT	Active Thermochemical Tables
aug-cc-pVNZ	augmented, correlation consistent, polarized valence N zeta basis sets, where n = double (D), triple (T), quadruple (Q), 5, or 6
aug-cc-pVNZ-DK	aug-cc-pVNZ basis sets with all electron DK basis sets
aug-cc-pVNZ-PP	aug-cc-pVNZ basis sets with pseudopotentials for heavy atoms
aug-cc-pwCVNZ	augmented, correlation consistent, polarized weighted core-valence N (D, T, Q, 5, or 6) zeta basis sets
aug-cc-pwCVNZ-DK	aug-cc-pwCVNZ basis sets with all electron DK basis sets
aug-cc-pwCVNZ-PP	aug-cc-pwCVNZ basis sets with pseudopotentials for heavy atoms
aN	aug-cc-pVNZ basis sets
aN-DK	aug-cc-pVNZ-DK basis sets
aN-PP	aug-cc-pVNZ-PP basis sets

awCVNZ	aug-cc-pwCVNZ basis sets
awCNZ-DK	aug-cc-pwCVNZ-DK basis sets
awCNZ-PP	aug-cc-pwCVNZ-PP basis sets
β	reduced partition function ratio
B3LYP	Becke 93 (exchange), Lee-Yang-Parr (correlation) DFT functional
BLYP	Becke 88 (exchange), Lee-Yang-Parr (correlation) DFT functional
BDE	bond dissociation energy
BDFE	bond dissociation free energy
CBS	complete basis set
CC	coupled cluster
CCSD	coupled cluster singles and doubles
CCSD(T)	coupled cluster singles, doubles, and perturbative triples
CI	configuration interaction
CID	collision-induced dissociation
CISD	configuration interaction singles and doubles
COSMO	conductor-like screening model
CV	core-valence
ΣD_0 ($\Sigma D_{0,0K}$)	total atomization energy
D	aspartic acid
DFT	density functional theory
DK	Douglas-Kroll-Hess
DZVP2	DFT optimized double zeta valence basis set with polarization functions
ϵ	isotope fractionation factor

ΔE_{CBS}	complete basis set energy change
ΔE_{CV}	core valence set energy change
ΔE_{SR}	scalar relativistic energy change
ΔE_{SO}	spin orbit energy change
ΔE_{ZPE}	zero point energy change
E	glutamic acid
E°	reduction potential
EA	electron affinity
ECP	effective core potential
ESI	electrospray ionization
FCI	full configuration interaction
FPD	Feller-Peterson-Dixon
FT-ICR	Fourier transform ion cyclotron resonance
ΔG_{aq}	aqueous Gibbs free energy
ΔG_{gas}	gas phase Gibbs free energy
ΔG_{solv}	aqueous Gibbs free energy of solvation
G	glycine
G3(MP2)	Gaussian-3 theory calculation using 2nd order Møller-Plesset perturbation theory (MP2)
GA	gas phase acidity
GGA	generalized gradient approximation
GTO	Gaussian-type orbital
η	hardness
$\Delta H_{f,0K}$	enthalpy of formation at 0 K

$\Delta H_{f,298K}$	enthalpy of formation at 298 K
ΔH_{rxn}	reaction change in enthalpy
$\hat{H}$	molecular Hamiltonian operator
HF	Hartree-Fock
IFF	isotope fractionation factor
K	equilibrium constant
k_B	Boltzmann constant
KS	Kohn-Sham
LCAO	linear combination of atomic orbitals
LDA	local density approximation
m/z	mass-to-charge ratio
MBPT	many-body perturbation theory
MD	molecular dynamics
MO	molecular orbital
MP*n*	n^{th} order Møller-Plesset perturbation theory
NCS	non-charge separated
QCISD(T)	quadratic configuration interaction singles, doubles, and perturbative triples
PCM	polarizable continuum model
PCO_2	partial pressure of CO_2
pK_a	negative logarithm of the acid dissociation constant
PP	pseudopotential
R	gas constant
RPFR	reduced partition function ratio

RE	reaction efficiency
SCF	self-consistent field
SCRF	self-consistent reaction field
SHE	standard hydrogen electrode
SMD	solvation model based on density
SO	spin-orbit
SORI	sustained off-resonance irradiation
SR	scalar relativistic
STO	Slater-type orbital
T	temperature
TAE	total atomization energy
TC	temperature correction
TISE	time-independent Schrödinger equation
TS	transition state
VAE	vertical attachment energy
VDE	vertical detachment energy
V_m	molar volume
ZPE	zero point energy
>	greater than
<	less than
=	equal to
~	approximately
Å	Angstrom

°	degrees
±	plus or minus

ACKNOWLEDGMENTS

First and absolutely foremost, I would not be where I am or who I am today without the tireless dedication of my mother, Terri Dianne McNeill. She has shown me over the course of our more than 3 decades together what it means to be generous, determined, and strong. She is my best friend and a bottomless source of support and love. I am also grateful for other close family members who have supported me through this long academic journey. You have all loved me and cheered me on through many years and degrees, and I am filled with gratitude for each of you. To my nieces and nephews, current and future – you can do anything in this world that you put your minds to!

Next, I would like to express my gratitude to my friends. Some of you have been in my life for a very long time, and others are more recent additions. Regardless, you add the sunshine to my days and make it easier to keep pushing forward. I am excited to see what life has in store for us all.

A special class of friends that deserve special attention are my groupmates and fellow/former graduate students. The Dixon group has been a family to me for many years. Both current and former members have carried me when I couldn't take another step on my own, shared successes and failures without bitterness or competition, and shown me what it means to be stronger together. Beyond our tight-knit group, many other members of the department, including students, professors (especially my committee), and staff, have encouraged me daily

and reminded me why I signed up for this in the first place. Shelby Hall will always feel like home in my heart.

Finally, I want to take a moment to thank my advisor, Dr. David A. Dixon. You invited me to this world when I was an undergraduate student way back in the summer of 2010. I hadn't even had physical chemistry yet, but you welcomed me to your lab and gave me the opportunity to find a life in computational chemistry. When you saw me at ACS in New Orleans in the spring of 2013, just over a month before I would complete my MS in chemistry from URI, you told me that you would have a space for me if I ever wanted to come back and get my Ph.D. In 2016, I took you up on that offer, and it is the greatest career decision I have ever made. You push me to be better. You support me as a true mentor. I am overcome with gratitude that you believed in me and gave me a chance to thrive.

CONTENTS

ABSTRACT ii

DEDICATION iv

LIST OF ABBREVIATIONS AND SYMBOLS v

ACKNOWLEDGMENTS xi

LIST OF TABLES xvi

LIST OF FIGURES xxii

CHAPTER 1 INTRODUCTION 1

CHAPTER 2 BOND DISSOCIATION ENERGIES IN GLYCINE, ALANINE, AND DIPEPTIDE DEPROTONATED ANIONS FOR USE IN ANALYZING COLLISION-INDUCED DISSOCIATION PROCESSES 20

a. Introduction 20

b. Experimental Methods 21

c. Computational Methods 22

d. Results 23

e. Conclusions 33

f. Reference List 63

g. Appendix A2 – Supporting Information 67

CHAPTER 3 EXPERIMENTAL AND COMPUTATIONAL STUDY OF THE GAS-PHASE ACIDITIES OF ACIDIC DI- AND TRIPEPTIDES 79

a. Introduction 79

b. Experimental Methods 80

c. Computational Methods 82

d. Results 82

e. Conclusions 89

f. Reference List 94

g. Appendix A3 – Supporting Information 99

CHAPTER 4 ENERGETICS OF CO_2^- IN AQUEOUS SOLUTION 105

a. Introduction 105

b. Computational Methods 107

c. Results 110

d. Conclusions 129

e. Reference List 146

f. Appendix A4 – Supporting Information 153

CHAPTER 5 ^{13}C FRACTIONATION FOR THE AQUEOUS ALANINE TRANSAMINASE REACTION 184

a. Introduction 184

b. Computational Methods 185

c. Results 191

d. Conclusions 205

e. Reference List 222

f. Appendix A5 – Supporting Information 226

CHAPTER 6 THE H•/H^- REDOX COUPLE AND ABSOLUTE HYDRATION ENERGY OF H^- 247

a. Introduction 247

b. Computational Methods 250

c. Results 253

d. Conclusions 264

e. Reference List 271

f. Appendix A6.1 – Supporting Information 280

g. Appendix A6.2 – Polarizabilities 297

CHAPTER 7 ABSOLUTE HYDRATION FREE ENERGY OF SMALL ANIONS 308

a. Introduction 308

b. Computational Methods 309

c. Results 313

d. Conclusions 327

e. Reference List 341

f. Appendix A7.1 – Supporting Information 348

g. Appendix A7.2 – Diatomic Electron Affinities 361

CHAPTER 8 CONCLUSIONS 375

REFERENCES 381

LIST OF TABLES

2.1 Fragmentation Energies in Glycine $[M - H]^-$ in kcal/mol 37

2.2 Fragmentation Energies in Glycine Amide $[M - H]^-$ in kcal/mol 39

2.3 Selected Transition States and Energy Barriers at the B3LYP/DZVP2 Level in kcal/mol ... 41

2.4 Fragmentation Energies in Alanine $[M - H]^-$ in kcal/mol 43

2.5 Fragmentation Energies in Alanine Amide $[M - H]^-$ in kcal/mol 45

2.6 Fragmentation Energies in Diglycine $[M - H]^-$ in kcal/mol 48

2.7 Fragmentation Energies in Diglycine Amide $[M - H]^-$ in kcal/mol 51

2.8 Product Production Energies in Dialanine $[M - H]^-$ in kcal/mol 54

2.9 Fragmentation Energies in Dialanine Amide $[M - H]^-$ in kcal/mol 57

A2.1 Additional Transition State and Barrier Energy Figures at the B3LYP/DZVP2 Level in kcal/mol 68

A2.2 Comparison of Computational and Experimental Heats of Formation of Neutral Fragments in kcal/mol 69

3.1 Reaction Efficiencies of Ion/Molecule Reactions Between Acidic Di- and Tripeptides and Reference Compounds 91

3.2 Experimental and G3(MP2) Theoretical GAs in kcal/mol for Di- and Tripeptides 92

3.3 GAs for Relevant Amino Acids, Their Amides, and Peptides in kcal/mol 92

A3.1 Total energies of all lowest-energy peptides and anions in the gas phase at the G3(MP2) level (kcal/mol) 100

4.1 Relative Energies (ΔH_{298}) of Neutral and Anionic Clusters (kcal/mol) in the Gas Phase ... 131

4.2 Relative Energies (ΔG_{298}) of Neutral and Anionic Clusters (kcal/mol) in Aqueous Phase with COSMO and SMD ... 133

4.3 Adiabatic Electron Affinities Gas and Aqueous Phases (kcal/mol) ... 135

4.4 Vertical Detachment Energies in Aqueous Phase (kcal/mol) ... 137

4.5 Vertical Attachment Energies in Aqueous Phase (kcal/mol) ... 137

4.6 CO_2 Stretch Frequencies and Infrared Intensities at the MP2/aT and G3MP2(B3) Levels . 138

4.7 Decomposition Reaction Energies in the Gas and Aqueous Phases at 298 K (kcal/mol) ... 140

A4.1 Total Energies of Neutral and Anionic Species in the Gas Phase in Hartrees: MP2 and G3MP2(B3) ... 155

A4.2 Total Energies of Neutral and Anionic Species in the Gas Phase in Hartrees: CCSD(T) and CCSD(T)-F12b ... 157

A4.3 Total Energies of Neutral and Anionic Species from Single Point Calculations in Aqueous Solution at the B3LYP/aT Level Using MP2/aT or aD Geometries in Hartrees ... 159

A4.4 Relative Energies (ΔG_{298}) of Neutral and Anionic Clusters in the Gas Phase in kcal/mol ... 161

A4.5 OCO Bond Angles and C=O Bond Lengths for Gas Phase Optimized Structures and Aqueous Phase Optimized Structures ... 162

A4.6 Relative Aqueous Free Energy for DFT-Optimized Clusters at 298 K in kcal/mol ... 163

A4.7 Adiabatic Electron Affinities for Gas Phase and Aqueous Phase Optimized Clusters at 298 K in kcal/mol ... 164

A4.8 Vertical Detachment Energies for $CO_2^- \rightarrow CO_2 + e^-$ in Gas Phase in kcal/mol ... 165

A4.9 Vertical Attachment Energies for $CO_2 + e^- \rightarrow CO_2^-$ in Gas Phase kcal/mol ... 165

A4.10 Raw Values and CBS Extrapolation for FPD Approximations in Hartrees ... 165

A4.11 HOMO and LUMO Energies for CO_2^- and Oxalate Dianion (C2O42-) with and without Explicit H_2O Molecules in Hartrees ... 166

5.1 Gas Phase and Aqueous Free Energy Values for Various Reactions at the MP2/aD and MP2/aD/COSMO Levels at 298 K in kcal/mol ... 207

5.2 Reaction Energies with Explicit H_2O Molecules in the Gas Phase and in Aqueous Solution in kcal/mol with All Light Isotopes 208

5.3 Reduced Partition Function Ratios at Various Carbon Atoms at 298 K 208

5.4 K and ϵ Values for the Reactions of Interest With and Without Heavy Isotopes for Deprotonated Alanine and Pyruvate in the Gas Phase and in Aqueous Solution at 298 K 209

5.5 Gas Phase Isotope Exchange Values at 298 K at Various Computational Levels in kcal/mol 210

5.6 Site Specific β Values for Non-Solvated H_2CO, CH_3NH_2, $CH_3NH_3^+$, Alanine, Deprotonated Alanine, Protonated Alanine, Pyruvic Acid, and Pyruvate in the Gas Phase at 298 K at the MP2/aT Level 211

5.7 Estimated Reduced Masses and Frequencies for ^{12}C and ^{13}C Bonds with N, O, and H 211

5.8 Estimated Free Energy (A), ΔA, and ϵ_{ZPE} in kcal/mol for ^{12}C and ^{13}C Bonds with N, O, and H in kcal/mol at 298 K 211

A5.1 Parameters for the QM/MM Potential and the Toukan-Rahman Water Potential 229

A5.2 Site Specific β Values for H_2CO, CH_3NH_2, $CH_3NH_3^+$, Alanine, Deprotonated Alanine, Pyruvic Acid, and Pyruvate in the Gas Phase at 298 K at the MP2/aD Level 230

A5.3 $^{13}C/^{12}C$ Exchange Thermodynamics for Explicitly Solvated Pyruvate and Alanine in kcal/mol 231

A5.4 Q Values for All Species Optimized in the Gas Phase or in Aqueous Solution at the MP2/aD or MP2/aD/COSMO Level at 298 K 232

A5.5 Reduced Partition Function Ratios at Various Carbon Atoms at 298 K for Glutamic Acid and α-Ketoglutaric Acid 234

A5.6 Estimated Free Energy (A), ΔA, and ϵ_{ZPE} in kcal/mol for ^{12}C and ^{13}C Bonds with N, O, and H in kcal/mol at 310 K 235

6.1 Reduction Potential of the H•/H^- Couple vs SHE in V 265

6.2 Absolute Hydration Free Energy of Radical Hydrogen ($\Delta G_{hyd}[H\bullet,n]$) Calculated as the Free Energy Change from H•(gas) + $(H_2O)_n$(aq) → H•$(H_2O)_n$(aq) at 298 K in kcal/mol 266

6.3 Absolute Hydration Free Energy of Hydride ($\Delta G_{hyd}[H^-,n]$) Calculated as the Free Energy Change from H^-(gas) + $(H_2O)_n$(aq) → $H^-(H_2O)_n$(aq) at 298 K in kcal/mol 267

A6.1.1 Gas Phase MP2/aug-cc-pV*n*Z ("a*n*"; *n* = D, T, Q) Total Energies in Hartrees 283

A6.1.2 Total Electronic Energies Including Solvation at the MP2/aD Level in Hartrees 285

A6.1.3 Solvation Free Energies at the MP2/aD Level in kcal/mol 286

A6.1.4 Total Electronic Energies Including Solvation for H• and H^- at Various Levels of Theory in Hartrees 287

A6.1.5 Solvation Free Energies for H• and H^- at Various Levels of Theory in kcal/mol 287

A6.1.6 Aqueous Electron Affinities (kcal/mol) and $E°$ (V) for H•/H^- at 298 K 288

A6.1.7 Anion…H-OH Bond Distances (Å) for F^- and H^- Clustered With 4 and 8 H_2O Molecules 289

A6.1.8 Natural Atomic Charges from Natural Bond Orbitals (NBO) for $H^-(H_2O)_n$ and $F^-(H_2O)_n$ for n = 4, 8 at the B3LYP/aug-cc-pVDZ Level 290

A6.1.9 Energy Values for reaction $OH^- + H_2 \rightarrow H_2O + H^-$ in kJ/mol 291

A6.1.10 Reaction energetics for $OH^- + H_2 \rightarrow H_2O + H^-$ 291

A6.1.11 Polarizabilities of H^- and F^- in $Bohr^3$ 291

A6.2.1 Polarizabilities of X^- (X = H and Halogens) in $Å^3$ with Molpro 2018 300

A6.2.2 Polarizabilities of Rare Gases in $Å^3$ with Molpro 2018 302

A6.2.3 Polarizabilities of Alkali Metal Cations in $Å^3$ (CC = CCSD(T)) 304

7.1 FPD and Literature Results for Small Triatomic Molecules and Diatomic Anions in kcal/mol 328

7.2 Gas Phase and Aqueous Acidities at the Best Available Level at 298 K in kcal/mol 329

7.3 Relative Gas Phase and Aqueous Free Energies for Anionic Species with Multiple Conformations at 298 K in kcal/mol 330

7.4 Relative Gas Phase and Aqueous Free Energies for Neutral Species with Multiple Conformations at 298 K in kcal/mol 331

7.5 Absolute Hydration Free Energy of Anions Calculated as the Converged Free Energy Change from $X^-(gas) + (H_2O)_n(aq) \rightarrow X^-(H_2O)_n(aq)$ at 298 K in kcal/mol 332

7.6 Absolute Hydration Free Energy of Acids Calculated as the Converged Free Energy Change from $HX(gas) + (H_2O)_n(aq) \rightarrow HX(H_2O)_n(aq)$ at 298 K in kcal/mol 334

7.7 ΔG_{solv} for Single Molecules at 298 K in kcal/mol at MP2/aD/COSMO and MP2/aD/SMD 335

A7.1.1 Bulk Solvent Shift ($\Delta\Delta G_{sol}[X^-,n]$) of Anions at 298 K in kcal/mol 350

A7.1.2 Geometry Parameters for CCSD(T)/awCQ-DK Optimized Triatomic Molecules 351

A7.1.3 Polynomial Fit-Optimized Parameters for Diatomic Molecules at CCSD(T) Level with both aug-cc-pwCVQZ-DK and aug-cc-pwCV5Z-DK Basis Sets 352

A7.1.4 FPD Data for Small Molecules (kcal/mol) 353

A7.1.5 Spin-Orbit Corrections and Respective $\Sigma D_{0,0K}$ Values in kcal/mol 354

A7.2.1 Polynomial Fit-Optimized Parameters for Diatomic Molecules at CCSD(T) Level with both aug-cc-pwCVQZ-DK and aug-cc-pwCV5Z-DK Basis Sets 364

A7.2.2 FPD Data for Small Molecules in kcal/mol 366

A7.2.3 Spin-Orbit Corrections and Respective $\Sigma D_{0,0K}$ Values in kcal/mol 367

A7.2.4 FPD Results for Small Molecules Using Best Spin Orbit Corrections with Comparable Literature Values in kcal/mol 368

A7.2.5 Gas Phase Electron Affinities of Small Molecules Calculated at the CCSD(T) Level with aug-cc-pwCVnZ-DK or CBS Extrapolated TAE + ZPE With and Without Spin-Orbit Corrections in eV 369

LIST OF FIGURES

2.1 Bond dissociation site numbering for amino acids 60

2.2 Bond dissociation site numbering for dipeptides 60

2.3 Low-energy CID spectrum of $[M - H]^-$ from diglycine 61

2.4 Low-energy CID spectrum of $[M - H]^-$ from diglycine amide 62

A2.1 Low-energy CID spectrum of $[M - H]^-$ from glycine (G) 70

A2.2 Low-energy CID spectrum of $[M - H]^-$ from glycine amide (G-NH_2) 71

A2.3 Low-energy CID spectrum of $[M - H]^-$ from alanine (A) 72

A2.4 Low-energy CID spectrum of $[M - H]^-$ from alanine amide (A-NH_2) 73

A2.5 Low-energy CID spectrum of $[M - H]^-$ from diglycine (GG) 74

A2.6 Low-energy CID spectrum of $[M - H]^-$ from diglycine amide (GG-NH_2) 75

A2.7 Low-energy CID spectrum of $[M - H]^-$ from dialanine (AA) 76

A2.8 Low-energy CID spectrum of $[M - H]^-$ from dialanine amide (AA-NH_2) 77

A2.9 Relative energies (kcal/mol) for the dissociation of glycine amide to H_2NCH_3 and $N{=}C{=}O^-$ 78

A2.10 Relative energies (kcal/mol) for the dissociation of glycine amide to form H_2O and $H_2NCH\text{-}C{\equiv}N$ or $HN^-CH{=}C{=}NH$ 78

A2.11 Relative energies (kcal/mol) for the dissociation of glycine amide to form H_2O and either $H_2NC(CH_3)\text{-}C{\equiv}N$ or $HN^-C(CH_3){=}C{=}NH$ 78

3.1 Gas-phase deprotonation reaction pathways for AD (A) and AE (B) 93

3.2 Gas-phase deprotonation reaction pathways for AAE 94

A3.1 Alanine-Alanine, Dialanine, AA 101

A3.2 Alanine-Aspartic Acid, AD 101

A3.3 Aspartic Acid-Alanine, DA 101

A3.4 Alanine-Glutamic Acid, AE 102

A3.5 Glutamic Acid-Alanine, EA 102

A3.6 Alanine-Alanine-Aspartic Acid, AAD 102

A3.7 Alanine-Aspartic Acid-Alanine, ADA 103

A3.8 Aspartic Acid-Alanine-Alanine, DAA 103

A3.9 Alanine-Alanine-Glutamic Acid, AAE 103

A3.10 Alanine-Glutamic Acid-Alanine, AEA 104

A3.11 Glutamic Acid-Alanine-Alanine, EAA 104

4.1 Lowest energy $CO_2 \cdot (H_2O)_3$ and $CO_2 \cdot (H_2O)_4$ clusters in the gas phase and in aqueous solution 141

4.2 $CO_2^{-} \cdot (H_2O)_3$ clusters in the gas phase at the MP2/aT level with electron density images at the B3LYP/aT level. Relative gas-phase enthalpies at the CCSD(T)-F12b/CBS level and relative aqueous free energies CCSD(T)-F12b/CBS/COSMO level 142

4.3 Lowest energy $CO_2^{-} \cdot (H_2O)_4$ clusters in the gas phase optimized at the MP2/aT level. Relative gas-phase enthalpies at the CCSD(T)-F12b/aD level and relative aqueous free energies at the CCSD(T)-F12b/aD/COSMO level 143

4.4 Lowest energy $CO_2 \cdot (H_2O)_8$ clusters in the gas phase and in aqueous solution 144

4.5 Lowest energy $CO_2^{-} \cdot (H_2O)_8$ clusters in the gas phase and in aqueous solution at the CCSD(T)/aD and CCSD(T)/aD/COSMO levels, respectively, with electron density images at the B3LYP/aT level 145

A4.1 $CO_2 \cdot (H_2O)_3$ structures at the MP2/aT and G3MP2(B3) levels 167

A4.2 $CO_2^{-} \cdot (H_2O)_3$ structures at the MP2/aT and G3MP2(B3) levels 168

A4.3 $CO_2 \cdot (H_2O)_4$ structures at the MP2/aT and G3MP2(B3) levels 169

A4.4 $CO_2^{-} \cdot (H_2O)_4$ structures at the MP2/aT and G3MP2(B3) levels 170

A4.5 $CO_2^{-\cdot}(H_2O)_4$ (B3LYP/aT) electron density 172

A4.6 $CO_2^{-\cdot}(H_2O)_8$ structures at the MP2/aD level 174

A4.7 $CO_2^{-\cdot}(H_2O)_8$ structures at the MP2/aD level with B3LYP/aT electron densities 175

A4.8 Selected $CO_2^{-\cdot}(H_2O)_3$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries 177

A4.9 Selected $CO_2^{-\cdot}(H_2O)_3$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries 177

A4.10 Selected $CO_2^{-\cdot}(H_2O)_4$ cluster for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries 178

A4.11 Selected $CO_2^{-\cdot}(H_2O)_4$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries 178

A4.12 Selected $CO_2^{-\cdot}(H_2O)_8$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries 179

A4.13 Selected $CO_2^{-\cdot}(H_2O)_8$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries 180

A4.14 Electron spin density (B3LYP/aT) for the VAE products of **N3-3**, **N4-4**, and **N8-4** 181

A4.15 Oxalate-related compounds at the MP2/aT level with relative energies in kcal/mol 182

5.1 Reactions (5.1) through (5.4) 212

5.2 Carboxylic acid deprotonation sites for α-ketoglutaric acid 213

5.3 Cumulative reduced partition function ratio (RPFR) vs. vibrational frequency (cm^{-1}) at 298 K for ^{13}C at either the methyl, C2 site, or the C-terminus or equivalent 214

5.4 Isotope fractionation factors ϵ and ϵ_{ZPE} at each $^{13}C/^{12}C$ substitution site of alanine and pyruvate for reactions (5.1) through (5.4) with aqueous optimized, explicitly hydrated clusters 218

5.5 ^{13}C exchange between (5.15) deprotonated alanine and pyruvate, (5.16) neutral non-charge separated (NCS) alanine and pyruvate, and (5.17) formaldehyde (H_2CO) and methylamine (CH_3NH_2) 220

5.6 Cumulative reduced partition function ratio (RPFR) vs frequency in cm^{-1} for formaldehyde, methylamine, protonated methylamine, pyruvic acid, pyruvate, alanine, and deprotonated alanine in the gas phase at 298 K 221

A5.1 Starting explicitly solvated (single solvation shell) structures from MD simulations for deprotonated alanine, pyruvate, α-ketoglutarate, and glutamate 235

A5.2 Alanine and related structures with 7 explicit H_2O molecules optimized (A) in the gas phase and (B) in aqueous solution 236

A5.3 Glutamic acid and related structures with 8 explicit H_2O molecules optimized (A, C, E) in the gas phase and (B, D, F) in aqueous solution 237

A5.4 Zwitterionic glutamic acid deprotonation reactions and energetics compared to experimental pKa values for both (A) MP2/aD gas phase optimized and (B) MP2/aD/COSMO aqueous optimized clusters of glutamic acid or glutamate with 8 explicit H_2O molecules 240

A5.5 Pyruvic acid and pyruvate with 6 explicit H_2O molecules optimized (A) in the gas phase and (B) in aqueous solution 241

A5.6 α-Ketoglutaric acid and related structures with 9 explicit H_2O molecules optimized (A) in the gas phase and (B) in aqueous solution 242

A5.7 Cumulative reduced partition function ratio (RPFR) vs. vibrational frequency (cm^{-1}) at 298 K for ^{13}C at either the C2 site or the C-terminus 243

6.1 $H^-(H_2O)_n$ clusters with n = 4, 8, 12, and 16 269

6.2 $H•(H_2O)_n$ clusters with n = 1 – 4 and 8 270

6.3 Energy diagram for the creation of aqueous H^- from gas-phase H• with free energy values in kcal/mol at 298 K 271

A6.1.1 Optimized $H^-(H_2O)_n$ clusters with n = 4, 8, 12, and 16 292

A6.1.2 Optimized $H•(H_2O)_n$ clusters with n = 1 – 4 and 8 294

A6.1.3 Optimized $(H_2O)_n$ clusters with n = 2 – 4, 8, 12, and 16 295

7.1 Thermodynamic cycle used for calculating the aqueous acidity of all species from gas phase acidity and hydration free energy of HA, H^+, and A^- 336

7.2 Conformers for $Cl^-(H_2O)_4$ and $Cl^-(H_2O)_8$ with relative gas phase and aqueous free energies at 298 K in kcal/mol 337

7.3 MP2/aD optimized geometries of diatomic anions with 4 H_2O molecules 338

7.4 MP2/aD optimized geometries of diatomic anions with 8 H_2O molecules 339

7.5 Conformers for (A) $HF(H_2O)_4$ and (B) $HCl(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol 340

7.6 MP2/aD optimized clusters for H_2S, H_2Se, H_2Te, HOCl, HOBr, and HOI with 4 H_2O molecules 341

A7.1.1 Conformers for $Cl^-(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol 355

A7.1.2 Conformers for $HBr(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol 356

A7.1.3 Conformers for $Br^-(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol 357

A7.1.4 Conformers for $HI(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol 358

A7.1.5 Conformers for $I^-(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol 359

CHAPTER 1

INTRODUCTION

1.1 Background

Computational chemistry is a particularly powerful tool when used in conjunction with experimental methods. Based on approximate solutions to complex quantum mechanical equations for chemical systems, a wide variety of methods exist, many of which are accessible in convenient software packages, that allow calculation of fundamental properties of these species. Particularly with access to modern computing resources, high level approximations deliver results, many of which can be compared directly to experimental work, with accuracy for thermodynamic values within 1 kcal/mol, also known as chemical accuracy. Gas-phase properties can be determined most conveniently, as the single molecule or cluster in space obeys by definition the common ideal gas behavior assumed in laboratory settings. For example, calculations of gas-phase acidities and other bond dissociation energies can be compared to experimental mass spectrometry studies in order to better understand decomposition patterns and to explore the conformations of species within the instrument, which cannot be determined experimentally with this technique. Studies of the energetics of these dissociation reactions and their transition states can provide insight about the mechanism of these decomposition pathways and justify the child ions detected in mass spectrometry under various experimental conditions. Additionally, solution phase properties can be calculated using self-consistent reaction fields (SCRFs) to simulate the molecule-solution electrostatic interactions. This inclusion of implicit

solvation, especially with additional inclusion of explicit solvent molecules in the gas-phase portion of the simulated cluster, is critical for accurately modeling many aqueous systems.

Computational chemistry is necessary to our understanding of chemical properties that cannot be or are difficult to observe experimentally. For example, the ability to model transition states, which are difficult if not impossible to study in detail experimentally, yields important reaction information that often cannot be derived by laboratory exercises.

This dissertation will focus on the computational analysis of many anionic systems, both in the gas phase and in aqueous solution, which provide unique insight on problems in chemistry that are important to investigation of a range of biologically and geochemically relevant chemical species. This often involves calculating properties for these species such as electron affinities and hydration free energies of the anions, which can be difficult or impossible to obtain experimentally.[1,2,3] Historically, atomic and molecular electron affinities, defined as negative of the difference between the total energy of the ground state of A and its negative ion A^-,[4] have been measured using laser photodetachment, electron impact detachment, and electron detachment in heavy particle collisions among a variety of approaches.[2] The lifetime of anions with negative electron affinities is fleeting (as brief as 10^{-14} to 10^{-10} s), but this lifetime is enough for the anionic species to be involved in chemical processes such as bond breaking or electron transfer.[5] The adiabatic electron affinity (AEA) is defined as the total energy difference between the ground state of the neutral species and the ground state of the anionic species, reported as $E_{neutral} - E_{anion}$ so that a positive AEA indicates a bound electron on the anion. The vertical attachment energy (VAE) is defined as the energy difference between the ground state neutral and the anion at the neutral geometry. The vertical detachment energy (VDE) is similarly defined as the energy difference between the ground state of the anion and the neutral at the anionic

geometry.[1,5] The VDE is more important in terms of photodetachment spectroscopy and is now the one most often measured. Once known, these EAs can be used in a wide variety of applications, from ion-molecule reactions in the upper atmosphere to inductive effects in organic chemical bonding to oxidative-reduction reactions and charge transfer reactions.[1] In this work, high accuracy electronic structure calculations are performed to determine EAs and other thermodynamic properties for a variety of anionic systems that are otherwise challenging to observe.

1.2 Computational Chemistry

Computational chemistry is a powerful tool that provides means for predicting thermodynamic properties of molecules and clusters of molecules that, in some cases, might otherwise not be accessible to experimental observation. Even when experimental values are available, computational chemistry can provide insight and verification that is invaluable for interpreting and then guiding experimental design. The field of computational chemistry has undergone rapid growth with improvements in computational resources, increased access to powerful supercomputing resources, and constant improvement in methods and available software.

Computational chemistry can be broadly divided into two categories, force field and electronic structure theory, which reflect predictions based either upon classical or quantum mechanical physics. Molecular mechanics and the corresponding use of it in molecular dynamics (MD) simulations, is a force field computational method that is useful for making quick structural predictions. When incorporated in MD simulations, it is possible to predict the bulk properties of typically larger systems as they evolve over time. MD simulations are based on

Newton's equations of motion; classical force fields are used to describe the potential energy surface of the system as it evolves in time.[7]

Electronic structure theory, the fundamental approach behind the work in this dissertation, instead relies upon solutions to the time-independent Schrödinger (TISE) equation for the motion of electrons in molecules. There are generally three types of calculations that fall under this umbrella – semi-empirical, density functional theory, and ab initio molecular orbital theory – in order of increasing computational complexity. Semi-empirical calculations, while computationally the least expensive, are wholly dependent upon fitting parameters obtained from experimental or high level computational data. Ab initio methods, conversely, seek solutions from first principles using only physical constants to solve the TISE.[6] The TISE has the form:

$$\hat{H}\Psi = E\Psi \tag{1.1}$$

where $\hat{H}$ is the molecular Hamiltonian, the total energy operator representing all kinetic and electrostatic interactions for the molecule system, Ψ is the wavefunction, and E is the total energy of the molecular system. The Born-Oppenheimer approximation[7] is the assumption that the nuclei are fixed in space compared to electronic motion due to the nuclei's much greater mass and much slower motion. Therefore, the motion of the nuclei and the electrons can be separated and the electronic Hamiltonian ($\hat{H}_{elec}$) can be defined as

$$\hat{H}_{elec} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{|R_A - r_i|} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{|r_j - r_i|} \tag{1.2}$$

where ∇ is the gradient operator, N is the number of electrons in the molecular system, M is the number of nuclei, A represents the nuclei, Z is the atomic number, i and j represent individual electrons, and R and r represent coordinates of the nuclei and electrons respectively. The terms within $\hat{H}_{elec}$ represent the kinetic energy of the electrons (repulsive), the nuclear-electron potential (attractive), and the electron-electron potential (repulsive).[6] Taking the first derivative

of the energy with respect to the nuclear coordinates $dE/\{dx_i\}$ allows for geometry optimization. When $dE/\{dx_i\} = 0$, the second derivative of the energy with respect to the nuclear coordinates can be used to determine if the molecular species is at a stationary point, for example, an energy minimum, a transition state, or a more complex saddle-point in the potential energy surface as well as providing the vibrational frequencies. Both density functional theory (DFT) and molecular orbital (MO) theory are approaches to solving the TISE.

1.2.1 Molecular Orbital (MO) Theory

The simplest approach to mathematically describe the electronic structure of molecules is Hartree-Fock (HF) theory. This theory approaches solving the TISE by calculating the components of the total energy using a basis set often based on hydrogen-like orbitals using Slater determinants.[6] These results depend upon the independent-particle model, in which the interaction of every electron in the system with the electrons around it are actually represented as an average of all electron-electron interactions of each electron. The HF energy accounts for approximately 99% of the total energy of the system.[6] However, the HF model only accounts for the average electron-electron interactions, thereby neglecting the correlation between electrons, which is much smaller in magnitude than the other components of the total energy, but of high importance in terms of chemical accuracy.

In order to solve the TISE, one must first have a wavefunction for the system composed of molecular orbitals. One way of creating the molecular orbitals is as a linear combination of atomic orbitals (the LCAO approximation).[7] These are typically constructed from individual mathematical functions that resemble atomic orbitals (AOs) called basis functions, the collection of which is called a basis set. When these basis sets are assigned for each atom, the best MO coefficients are found by variational minimization. The LCAO approach to solving the electron-

electron interactions can be computationally expensive, especially when Slater-type orbitals (STOs) are used to construct the basis sets due to the fact that they are exponential functions that are computationally taxing to integrate. Thus, most computational electronic structure codes use Gaussian-type orbitals (GTOs) instead of STOs. Two electron integrals involving STOs do not have simple analytical solutions, but those composed of GTOs do. Thus, a linear combination of primitive GTOs is used to model each STO-type basis function or contracted basis function. Split valence basis sets in which each valence atomic orbital is represented by multiple basis functions are called double zeta, triple zeta, etc. Because these orbitals representing each valence atomic orbital have different spatial properties, a linear combination allows the electron density to adjust spatially appropriate to the molecular environment. As the number of basis functions increase, the LCAO approximation error should decrease and eventually converge to the Hartree-Fock limit. The difference between the HF limit and the exact energy is known as the correlation energy

Minimum basis sets do not typically produce reasonable results due to the fact that the orbitals cannot adjust in size to accommodate directional changes such as those in forming a chemical bond. To achieve better accuracy for interactions, one can double the number of valence orbitals and use a linear combination of these valence orbitals to provide a weighted average of the two valence levels.[7] Additional improvement to basis set behavior comes from the addition of polarization functions and diffuse functions. Polarization functions allow for the orbitals to distort more and diffuse functions are important for allowing further extension of the spatial extent, which is of critical importance for the anions studied in this dissertation. The work in this dissertation will largely use polarized valence multiple zeta (pVNZ) correlation consistent (cc) basis functions augmented (aug) with diffuse functions, written as aug-cc-pVNZ where N

can be D for double, T for triple, Q for quadruple, 5 for quintuple, 6 for sextuple, etc.[8] For some DFT calculations presented in this work, the polarized valence double zeta DGauss basis set DZVP2,[9,10] optimized for use with DFT, is used.

HF yields a self-consistent field (SCF) solution, but neglects correlation between electrons. In the configuration interaction approach, the addition of more determinants to HF theory improves the model, allowing for calculations that converge toward the exact solution of the TISE. According to the variational principle, any calculated ground state energy will always be higher in energy than the true minimum for the system. Thus, we can include correlation energy components to our model to better approximate the true minimum without fear of producing a ground state energy that is lower than the true minimum of the system. There are several methods that attempt to better approximate this energy minimum by adding correlation to the HF solution.

The configuration interaction (CI)[11,12,13] method is performed by including additional Slater determinants that include excited state configurations. Computing the HF wave function generates virtual orbitals in addition to the occupied orbitals, and these virtual orbitals can be combined with the occupied orbitals to make additional Slater determinants describing the excitation of an electron into these virtual orbitals. Exciting a single electron is known as CIS (singles) and the excitation of two electrons is known as CISD (singles and doubles). The CISD level is required for correlation energy, as the CIS solution gives us the estimates of the energy of the excited states that can be generated by single excitations. For a molecule with more than 2 electrons, higher order excitations are also possible. Including more excitations increases the accuracy of the expansion and, therefore, the energy. If all possible excitations were included (full CI or FCI), we would have a numerically exact wave function, but the computational

expense of that is extremely prohibitive due to the factorial scaling. Due to the fact that the CI method is not size consistent or size extensive, the CI method is now rarely used to calculate the correlation energy.[14] A method is considered to be size consistent if the model for a group of atoms at infinite separation is equal to the sum of the properties of the individual molecules using that method. If a system of two individual molecules, A and B, are calculated together at infinity to determine E(A + B), then E(A + B) = E(A) + E(B) for a size consistent method. A size extensive method is one where the energy of a system of n noninteracting particles is equal to the energy of n times the energy of a single particle and the energy scales appropriately with the size of the system.[14,15]

One popular method for finding the coefficients for the expansion yielding the correlation energy is by many-body perturbation theory (MBPT), which allows for corrections to be calculated with respect to a reference state. Perturbation theory with respect to the HF energy is called Møller-Plesset perturbation theory.[16] These methods increase in expense as additional terms are considered and are named accordingly: MP2 (second order), MP3 (third order), MP4 (fourth order).[7] However, the computational cost significantly increases with increasing order, so MP2 calculations are most commonly employed as they scale as N^5 with N the number of basis functions due to the need to sort and transform the 2-electron integrals from the atomic orbital basis in which they are calculated to the molecular orbital basis needed for MP2.

MBPT CC calculations are based on many-body calculations that improve on the HF use of single-electron wave functions. This is formalized as follows,

$$\hat{H} = \hat{H}_0 + \lambda V \tag{1.3}$$

where $\hat{H}$ is the Hamiltonian operator for the system, $\hat{H}_0$ is the unperturbed part of the system, and λV is the perturbed part (λ is the formal order parameter). If Ψ_0 is the ground state eigenfunction

of $\hat{H}$ with energy E_0 and Φ_0 is the ground state eigenfunction for $\hat{H}_0$ with energy E_0^0, the exact energy E_0 can be expressed as a power series in the perturbation λV.[14,17] The energy of the system can then be expressed as an expansion of orders in the perturbation $V = (\hat{H} - \hat{H}_0)$,[15]

$$E = \sum_{k=0}^{\infty} < \Phi_0 | \hat{H}[(E_0 - \hat{H}_0)^{-1}\hat{H}]^k | \Phi_0 > \quad (1.4)$$

in which Φ_0 is a SCF function, the terms with $k > 0$ represent the correlation energy, and $\hat{H}_0$ is a separable n-particle Hamiltonian whose eigenfunction is Φ_0.

Coupled cluster (CC) methods[18,] use a wavefunction written using an exponential ansatz

$$|\Phi_0 > = e^T |\Phi_0 > \quad (1.5)$$

where the reference wavefunction (Φ_0) is usually a Slater determinant constructed of HF molecular orbitals. The choice of this exponential ansatz is convenient because it guaranties that the solution is size extensive. CC methods are size consistent. The exponential ansatz can be expanded in a Taylor series

$$e^T = 1 + T + \frac{1}{2}T^2 + \frac{1}{6}T^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!} T^k \quad (1.6)$$

where T is the cluster operator. T is given by

$$T = T_1 + T_2 + T_3 + \ldots + T_n \quad (1.7)$$

where n describes a specific degree of excitation as well as the number of interacting electrons involved in the excitation – $n = 1$ are singles, $n = 2$ are doubles, $n = 3$ are triples, etc. This cluster operator must be truncated at some level, as n approaching infinity would scale as Full CI, and the choice of truncation is what yields the various CC methods, which are named similarly to the MP methods described above– CCSD (singles and doubles), CCSD(T) (CCSD plus an estimate of the contribution from triples from MP4), etc. While MP4 is somewhat less computationally expensive than CCSD(T) due to the MP4 calculation being non-iterative, CCSD(T) is much more accurate. MP2 is used to generate the starting guess for CCSD which is iterative. The (T)

stands for perturbative triples from an MP4 calculation using the CCSD amplitudes. CCSD scales as iterative N^6 and CCSD(T) scales as N^7 but is not iterative. A full CCSDT calculation scales as iterative N^8 and is only used in very high accuracy benchmarks of very small systems with modest basis sets. CCSD(T)[14,19,20,21,22,23,24] calculations are regularly employed in this work, as these calculations are considered to be the "gold standard" in computational electronic structure calculations.

The CCSD(T)-F12[25,26,27] approach includes terms in the wavefunction that explicitly depend on the interelectronic distances, which enables greatly improved performance with a smaller basis set. The F12 correction improves the basis set convergence of correlation and reaction energies and removes errors from the Hartree-Fock contributions by including MP2 single excitations directly in the auxiliary basis set. There are two approximations for the CCSD(T)-F12 energy, CCSD(T)-F12a and CCSD(T)-F12b. Calculations at the CCSD(T)-F12a level produce highly accurate relative energies, but only up to the aVTZ basis set. Beyond this basis set size, CCSD(T)-F12a convergence in terms of the basis set becomes slow, whereas CCSD(T)-F12b convergence remains smooth to the basis set limit.[26] Thus, this work will exclusively use the CCSD(T)-F12b method rather than CCSD(T)-F12a. An advantage of the F12 methods is the quick convergence to the CBS limit for properties such as geometries, vibrational frequencies, and energies.

When correlation consistent basis sets are used, as they are in all of the work presented here at the CCSD(T) level, the calculated values at a variety of basis set levels can be extrapolated to the complete basis set (CBS) limit according to the following equation. This provides the most accurate possible values at this level.

$$E_n = E_{CBS} + A(n + 1/2)^{-4} \tag{1.8}$$

1.2.2 Density Functional Theory (DFT)

Density functional theory (DFT) is based on the proof by Hohenberg and Kohn[28] that the ground state electronic energy of a molecular system can be determined completely from the electron density of that system. E. B. Wilson[29] offered an intuitive proof of why this is true by arguing that the integral of the density defines the number of electrons, the cusps in the density define the positions of the nuclei, and the heights of those cusps define the nuclear charges for each nucleus. Unlike the MO theory-based methods discussed above, Hohenberg-Kohn DFT methods, in principle, may not scale exponentially with the number of electrons in the system depending on the choice of functional. Calculations of the energy are based on three dimensions only – those of the density of the electron cloud – and, therefore, Hohenberg-Kohn DFT calculations may be much less computationally expensive as the size of the system grows. The components of the total energy for DFT calculations are

$$E[\rho] = T_s[\rho] + V_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \tag{1.9}$$

where T_s is the kinetic energy for the non-interacting system, ρ is the electron density, V_{ne} is the nuclear-electronic potential energy, J is the electron-electron repulsive energy, and E_{xc} is the exchange-correlation energy.[30] The exchange-correlation energy term is where DFT functionals differ, and it takes the form

$$E_{xc}[\rho] = (T[\rho] - T_s[\rho]) + (V_{ee}[\rho] - J[\rho]) \tag{1.10}$$

where T and V_{ee} are the kinetic energy and electron-electron repulsive energy for the interacting system respectively. The kinetic and potential terms from equation 1.10 are not well described in terms of electron density; Kohn and Sham[30] proposed that the kinetic, electron-nuclear, and Coulomb electron-electron energies should be calculated using orbitals, similar to calculations using the HF method, and this Kohn-Sham (KS) approach is what is currently used in

computational chemistry for DFT calculations. This increases the complexity of the calculation from 3 to 3N,[6] which may be significantly less computationally expensive than MO theory models, but is dependent on the choice of the functional. The wide variety of DFT approaches available arise from the variety of exchange-correlation functionals created to to treat the exchange and correlation portions of the electron interactions. Since the exact form of E_{xc} is unknown, the accuracy of DFT is directly dependent upon the quality of the functional that is used. There are a variety of functionals available, including local density approximation (LDA) functionals,[31,32] generalized gradient approximation (GGA) functionals,[33,34,35] and hybrid functionals.[36] LDA functionals depend only on the electronic density whereas GGA functionals depend on the electronic density and its gradient. Hybrid functionals are comprised of a portion of HF exchange in addition to exchange and correlation from a variety of other sources. The exchange-correlation term is roughly 10 times smaller in magnitude than the kinetic energy component of the system, therefore KS DFT methods are generally much less sensitive to inaccuracies in the functional than calculations that are orbital-free.[6]

The DFT method used in this work is B3LYP, which is a hybrid functional composed of the Becke 3-parameter exchange functional[37] and the Lee, Yang, Parr correlation functional.[38] This DFT functional is a workhorse in the computational chemistry field, particularly for geometry optimizations, but suffers from the classic DFT limitation of lacking accurate through-space interaction energy.[39] Thus, we use other approaches to arrive at our reported thermodynamic values, including the previously mentioned MP2 method.

1.2.3 Composite Approaches

In order to approach the accuracy of CCSD(T) but avoid the high computational expense, several composite approaches have been proposed. In this work, particularly for the work with

organic molecules, the Gaussian-*n* (G*n*) methods[40] were employed, as prior work[41,42,43,44,45,46] in our group has shown that G3(MP2)[47] has comparable accuracy with CCSD(T) calculations extrapolated to the complete basis set (CBS) limit for peptides, ± 1 kcal/mol or "chemical accuracy" when compared with experimentally derived values. The individual steps for the G3(MP2) calculation are

(1) Optimization and frequency calculation at the HF/6-31G(*d*) level of theory.

(2) Optimization at the MP2(FULL)/6-31(*d*) level of theory.

(3) Single point calculation at the QCISD(T,FC) /6-31G(*d*)//MP2(FULL)/6-31G(*d*) level of theory

(4) Single point calculation at the MP2(FC)/G3MP2large//MP2(FULL)/6-31G(*d*) level of theory.

The total energy for G3(MP2) at 0 K is then calculated as

$$E_0[G3(MP2)] = QCISD(T)/6-31G(d) + \Delta E(MP2) + \Delta E(SO) + E(HLC) + E(ZPE) \quad (1.11)$$

where the total energy is based on a higher level calculation at the quadratic configuration interaction (QCISD(T)) level of theory.[48] This level of theory is similar to CCSD(T) with the exception of a few small terms that are missing. A series of other corrections are added to the QCISD(T) energy, including a correction for the basis set extension at the MP2 level, ΔE(MP2), calculated as

$$\Delta E(MP2) = [E(MP2/G3MP2large)] - [E(MP2/6-31G(d))] \quad (1.12)$$

where the G3MP2large basis set[49] is a modification of the 6-311+G(3*df*,2*p*) basis set from G2(MP2) theory. It is necessary to include a spin orbit correction $\Delta E(SO)$ to account for this relativistic effect, which lowers the sum of the atomic energies decreasing the total atomization

energies. These corrections are included by replacing energies that correspond to an average over the available spin multiplets with energies for the lowest multiplets, as most electronic structure codes are only capable of producing spin multiplet averaged wavefunctions. In this dissertation, this correction is only needed for calculating heats of formation. The zero-point correction, $E(ZPE)$, is obtained by scaling the SCF frequencies from the HF/6-31G(d) level by 0.8929. The calculated harmonic frequencies need to be scaled to account for anharmonic effects because SCF harmonic frequencies overestimate experimental frequencies by up to ~ 12%.[50] The higher level correction, $E(HLC)$, is added to account for any other remaining deficiencies in the calculation such as correlation energy between spin-paired electrons present in single bonds. For molecules, the $E(HLC)$ is given as:

$$E(HLC) = -An_B - B(n_\alpha - n_\beta) \quad (1.13)$$

where A and B are 9.279 and 4.471 mhartrees respectively and are numbers chosen to give the smallest error to the energy after fitting to experiment, and n_α and n_β are the number of alpha and beta valence electrons respectively. Again, the HLC is only relevant to the calculation of heats of formation.

There are molecular systems for which the initial HF optimization in a G3(MP2) calculation is problematic, particularly for small anions. For these systems, especially when the initial geometry optimization is performed at the B3LYP level, G3MP2(B3)[51] can be used to preserve the original optimized structure. G3MP2(B3) is very similar to G3(MP2) except for the initial optimization and calculation of the zero point energy (steps 1 and 2). These calculations for G3MP2(B3) are performed instead at the B3LYP/6-31G(d) level, which avoids sometimes poor HF optimization.

A final composite method employed in some of the work in this dissertation is the Feller-Peterson-Dixon method (FPD),[52,53,54,55,56,57,58,59,60] developed in collaboration between the Dixon group and Peterson and Feller at Washington State University. The FPD approach improves the accuracy of the 3-point CCSD(T)[19,20,21] extrapolation to CBS by including correction factors. These correction factors account for spin-orbit coupling (ΔE_{SO}), core-valence (ΔE_{CV}), scalar relativistic effects (ΔE_{SR}), and a ZPE correction, which are combined to give a much more accurate calculation of the enthalpy of formation for the molecule of interest. The total atomization energy (TAE, $\Sigma D_{0,0K}$) is calculated as

$$TAE = \Delta E_{CBS} + \Delta E_{ZPE} + \Delta E_{CV} + \Delta E_{SR} + \Delta E_{SO} \quad (1.14)$$

ΔE_{ZPE} is taken from the MP2/aug-cc-pVTZ optimization of the molecule. ΔE_{CV} is calculated at the CCSD(T)/aug-cc-pwCVTZ level. It is found by taking the difference between the core and valence energies at that level. ΔE_{SR} is found by calculating the second order Douglas-Kroll-Hess (DK)[61,62,63] with the all-electron aug-cc-pwCVTZ-DK basis set or by CISD/aug-cc-pwCVTZ.[64] ΔE_{SO} is found by taking the atomic experimental values for the individual atoms from the Moore tables.[65] Heats of formation at 0 K and at 298 K can be calculated from the TAE using experimental $\Delta H_{f,0K}$ for the atoms[66] and temperature corrections (TCs) from the MP2/aug-cc-pVTZ optimization according to the Curtiss et al. procedure[67] as follows.

$$\Delta H_{f,0K} = \Sigma \Delta H_{0K,atoms} - TAE \quad (1.15)$$

$$\Delta H_{f,298K} = \Delta H_{f,0K} + TC_{molecule} - \Sigma TC_{atoms} \quad (1.16)$$

1.3 Computational Thermodynamics

Thermodynamics is used to describe energy changes associated with chemical and physical processes. According to the first law of thermodynamics, the change in internal energy (ΔU) is defined as the sum of heat (q) and work (w). If the system is at constant pressure, then the heat content, enthalpy (H), can be defined as

$$H = U + PV \tag{1.17}$$

The internal energy of a molecule at 0 K is defined as the sum of the electronic energy of the molecule and its zero point energy

$$U_0 = E_{elec} + \sum_i \frac{1}{2} h\omega_i \tag{1.18}$$

The molecular energy for an ideal gas over a temperature range T can be decomposed into energy contributions from electronic, vibrational, translational, and rotational motions:

$$U_T = U_0 + \Delta U_{0\rightarrow T(elec)} + \Delta U_{0\rightarrow T(vib)} + \Delta U_{0\rightarrow T(rot)} + \Delta U_{0\rightarrow T(trans)} \tag{1.19}$$

The enthalpy at some T for an ideal gas can be derived from equation 1.17 to be $H_T = U_T + RT$. The total entropy of the system, or the measure of the number of states a system has access to, is also the sum of contributions due to the electronic, vibrational, translation, and rotational motions as a function of temperature. The entropies are calculated using the statistical mechanical expressions for a harmonic oscillator, rigid rotor description of a molecule.[68] The Gibb's free energy (G) combines enthalpy and entropy into one value to predict whether a process is spontaneous or not at constant pressure and temperature.

$$G = H - TS \tag{1.20}$$

The change in free energy can be used to calculate the equilibrium constant (K)

$$K = e^{-\frac{\Delta G}{RT}} \tag{1.21}$$

where R is the gas constant and T is the temperature in Kelvin.

1.4 Chapter Descriptions

The study of peptide gas phase bond dissociation energies (BDEs) and gas phase acidities, Chapters 2 and 3 respectively, are motivated by partnership with experimental mass spectrometry to further the goals of the Human Proteome Project.[69] The Human Proteome Project was started by the international Human Proteome Organization in 2011 after the

successful completion of the Human Genome Project, a joint effort by the National Institute of Health and the Department of Energy of the United States.[70,71,72] A critical component of studying the human genome is understanding the sequence of the genomic proteins, which is studied largely by mass spectrometry due to being cost effective, highly reproducible, and capable of sensitive and rapid identification of peptides and proteins.[73] Chapter 2 is a detailed study of the bond dissociation energetics of glycine, alanine, and their dipeptides in the gas phase for comparison with mass spectrometry results. Select transition states are also studied to investigate why some favorable dissociations are not observed in low-energy collision dissociation experiments. Chapter 3 is a study of the gas phase acidity of di- and tripeptides composed of alanine with one acidic residue, either aspartic or glutamic acid. This work is compared with experimental mass spectrometry bracketing experiments that produce experimental gas phase acidities in excellent agreement with computational models.

An isotope fractionation factor (IFF) is the quantification of the preference between two isotopes of the same element in either chemical or physical equilibrium. These values are commonly used by geochemists to answer a variety of questions about modern and ancient Earth, such as $^{11}B/^{10}B$ fractionation for determination of the pH of ancient oceans and the PCO_2 of the ancient atmosphere,[74] $^{17}O/^{16}O$ and $^{18}O/^{16}O$ fractionation to indicate the temperature of ancient seas, water movement, and paleoclimate,[75] $^{13}C/^{12}C$ fractionation between gaseous CO_2, aqueous carbonate species, and major carbonate minerals as well as $^{13}C/^{12}C$ fractionation to allow differentiation between C_3 vs C_4 photosynthetic pathways for plant growth.[75] Chapter 4 describes the $^{13}C/^{12}C$ isotope fractionation between deprotonated alanine and pyruvate in the alanine transaminase overall reaction. This includes a discussion of the accuracy of a variety of

models, from high accuracy, explicitly and implicitly solvated species to simplistic gas phase analogues representing only the fundamental frequencies affected by the isotope exchange.

Computational tools are critically important to determining accurate thermodynamic values for anionic species in solution. In particular, chapters 5 and 6 of this work explore EAs and reduction potentials for CO_2^- and H^-, which are difficult to impossible to observe experimentally and require careful treatment with an explicit solvent continuum in addition to implicit solvation to accurately capture the behavior of these anions in solution. Hydration of ions plays a vital role in aqueous chemical and biological systems.[76,77,78,79,80,81] It is of interest to better understand the nature of CO_2^- in aqueous systems, as this is potentially important for a variety of geochemical processes in which the presence of CO_2^-, in addition to carbonate CO_3^{2-} and bicarbonate HCO_3^-, could play a role in the aqueous chemistry of CO_2. Furthermore, there is the possibility that CO_2^- can dimerize to produce value-added organic compounds, such as the oxalate dianion. Chapter 5 describes the aqueous CO_2^- anion and the CO_2/CO_2^- reduction energy in terms of electron affinity.

The solvation of the H^- (hydride) anion is important in developing an absolute hydricity scale for developing catalytic processes, especially those involving hydride transfer.[82,83,84,85,86,87,88] Chapter 6 similarly describes the solvation of hydride (H^-) and the reduction potential of H•. This study of the hydration and reduction potential of hydride is very important as it is the highest accuracy study of its kind. Former estimates of the hydration free energy of hydride depend on assumptions about the polarizability or atomic radius of hydride, which our calculations show are poor assumptions. Included in chapter 6 is an appendix with a discussion on the calculated polarizabilities of group 1 cations Li^+, Na^+, K^+, Rb^+, Cs^+, and Fr^+ and the halide anions F^-, Cl^-, Br^-, I^-, and At^- compared with the polarizabilities for the rare gases

with which they are isoelectronic, He, Ne, Ar, Kr, Xe, and Rn. These calculations are performed with doubly and triply augmented basis sets in order to most accurately capture the polarizabilities of the anionic species, which are not converged otherwise.

Finally, chapter 7 is a detailed study on the gas phase and aqueous thermodynamic properties of the halides (F^-, Cl^-, Br^-, I^-), hydride chalcogens (HO^-, SH^-, SeH^-, TeH^-), and the hypohalite anions (OCl^-, OBr^-, and OI^-). This work includes high accuracy calculations of their heats of formation and electron affinity using the FPD method. The hydration free energies, gas phase and aqueous acidities are also determined with clusters of up to eight explicit H_2O molecules.

CHAPTER 2

BOND DISSOCIATION ENERGIES IN GLYCINE, ALANINE, AND DIPEPTIDE DEPROTONATED ANIONS FOR USE IN ANALYZING COLLISION-INDUCED DISSOCIATION PROCESSES

From McNeill, A. S.; Cui, C.; Miller, S. R.; Stover, M. L.; Burns, M.; Cassady, C. J.; Dixon, D. A. *Int. J. Mass Spectrom.* **2018**, *429*, 212 – 226.

Introduction

Substantial information is required to continue the development of negative ion mode (anionic) mass spectrometry for use in proteomics. An improved understanding of the fragmentation patterns from collision-induced dissociation (CID) is an important component of this continued development. In the CID experiment, there is enough energy for the molecule to undergo structural rearrangements in order to reach alternative configurations with potentially lower energies.[1,2] Due to the nature of mass spectrometric measurements, experimental data can only determine whether ions of a particular mass-to-charge ratio (m/z) are observed, but it cannot directly give detailed information about the observed ion's actual molecular structure. Computational modeling can be used to elucidate the likely identity and conformation of the observed molecular fragments.

Although CID fragmentation mechanisms for deprotonated peptide ions have been proposed in numerous experimental studies,[3,4,5,6,7,8,9,10] relatively few reports that have employed a concerted approach of mass spectrometry and high level molecular orbital calculations to provide a more detailed understanding of the processes involved.[11,12,13] However,

in recent years, we have shown the utility of such a combined approach by experimentally and computationally determining the gas-phase acidities (GAs) and deprotonation sites of amino acids and amino acid amides.[14,15,16,17,18] A knowledge of peptide deprotonation sites can be very useful in interpreting negative ion CID spectra[11,12,19] because fragmentation is usually charge-directed.[20,21] Our GA studies involved composite correlated molecular orbital theory at the G3(MP2) level.[22] The validity of the G3(MP2) approach has been demonstrated by agreement between the G3(MP2) GA values and the experimental GA values to within about ±2 kcal/mol, as well as agreement of G3(MP2) values with higher computational level values obtained using the Feller-Peterson-Dixon composite approach based on extrapolating CCSD(T) calculations to the complete basis set limit with additional corrections.[2324,25,26,27]

In the current study, we have used the G3(MP2) approach to calculate key bond energies and the energies of important fragments to explain the observed product ions from low-energy CID experiments. We chose the simplest amino acids glycine and alanine to provide an initial understanding of the energetics of the backbone bonds. The side chains of glycine and alanine are hydrogen and a methyl group, respectively, which should not actively participate in the fragmentation mechanisms. In addition, amino acid amides were investigated because the replacement of a terminal carboxylic acid group by an amide allows modeling of residues that are not located at the C-terminus of a larger peptide chain.

Experimental and Computational Methods

Experimental Methods

All experiments were performed using a Bruker (Billerica, MA, USA) BioApex 7e Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer equipped with a 7.0 T superconducting magnet. Deprotonated analyte ions were generated in the negative ion mode by

electrospray ionization (ESI) on solutions with a concentration of ~6×10^{-5} M in a solvent of 50:50:1 (by volume) methanol:water:ammonium hydroxide. The solution was introduced into the ESI source with a syringe pump using a flow rate ~120 μl/h. For low-energy CID experiments, precursor ions were isolated by correlated resonance frequency ejection techniques[28] and excited using sustained off-resonance irradiation (SORI).[29] Ions were activated 500 Hz off-resonance (lower frequency, higher m/z) with an excitation pulse in the range of 1.8 to 3.6 V_{pp} for 170 ms in the presence of a pulsed pressure of argon that reached a maximum of 10^{-5} mb. CID product ions were detected following a 1.0 s delay to allow for collisions and for the argon to pump away. The maximum center-of-mass (cm) collision energy was in the range of 5 to 25 eV. The high resolution of FT-ICR made it possible to accurately assign chemical formulas for the CID product ions; for example, the instrument was capable of distinguishing between ammonia (NH_3) and hydroxyl (OH) elimination even though both have a nominal mass of 17 Da.

All biomolecules were obtained from Sigma-Aldrich Corp. (St. Louis, MO, USA) except glycine (J. T. Baker Inc., Phillipsburg, NJ, USA), alanine amide (TCI America, Portland, OR, USA), and dialanine amide (custom synthesized by Biomatik, Wilmington, DE, USA). All CID mass spectra can be found in the Appendix with peak labelling for dipeptides using both organic nomenclature and peptide fragmentation nomenclature.

Computational Methods

Initial geometries were optimized at the DFT level using the B3LYP exchange-correlation functional[30,31] and the DZVP2 basis set.[32] The vibrational frequencies were calculated to show that the optimized structures are energetic minima. The conformational space for each structure was searched by extensive manual sampling at the DFT level (see Appendix). The correlated molecular orbital method G3(MP2) performs better in the prediction of bond

energies as well as steric, non-bonded interactions than do the most widely used DFT exchange-correlation functionals for these types of compounds. The heats of formation were calculated from atomization energies at the G3(MP2) level. The program Gaussian09 was used to perform calculations at the density functional theory (DFT) and correlated molecular orbital (MO) theory levels.[33]

Results

Amino acids and amides

We have previously reported the structures and energetics of neutral glycine and glycine amide and the corresponding deprotonated anions.[16] The overall agreement with the available experimental data[34] for the heats of formation of neutral product molecules is excellent as shown in the Appendix. We first consider the bond energies in the backbone of the glycine anion formed by loss of a proton from the CO_2H group, $[M - H]^-$, and in the glycine amide anion generated by loss of a proton from the $C(O)NH_2$ group (Figure 2.1 and Tables 2.1 and 2.2). The bonds can undergo a homolytic cleavage to generate two radicals, one of which is an anion radical, which can lose an electron to form a closed shell molecule via dissociative electron detachment (DED). This will happen when the anion radical is not stable to electron loss. The most favorable adiabatic bond cleavage process for the glycine anion involving no rearrangement of either fragment is to break the C-C bond (bond 2) between the $-CO_2^-$ group and the NH_2CH_2- group to generate the CH_2NH_2 radical in a DED process. Slightly lower in energy is the process where the electron transfers to the NH_2CH_2, which then rearranges to form the more energetically favorable CH_3NH^- fragment and CO_2. Much higher in energy is the process to break the NH_2-CH_2R bond leading to the formation of the $CH_2CO_2^-$ radical and NH_2 radical. Even though NH_2^- is 17.2 kcal/mol more stable than $\cdot NH_2$, the charged carboxylate species is 34

kcal/mol more stable than the 3-member strained α-lactone ring, the lowest energy form of the neutral carboxylate species. The lower bond energy for bond 2 (54.9 kcal/mol) as compared to bond 1 (86.3 kcal/mol) is driven by the very favorable formation of CO_2.

The only product ions in the low-energy CID mass spectrum of the glycine $[M - H]^-$ anion (Figure A2.1) involve loss of the neutrals CO and $HN{=}CH_2$ or its isomeric carbene H_2N-C:H. These are low-energy processes with elimination of CO requiring 50.5 kcal/mol (2.19 eV) and elimination of $HN{=}CH_2$ requiring 26.2 kcal/mol (1.14 eV) whereas formation of the carbene requires 62.1 kcal/mol (2.69 eV). These energies are readily accessible experimentally because the maximum CID collision energy in the center of mass was ~20 eV. Formation of $HN{=}CH_2$/H_2N-C:H is accompanied by formation of the stable formate ion and loss of CO generates the $H_2NCH_2O^-$ anion. Two reactions with low reaction energies that are not observed experimentally for glycine under low-energy CID are loss of H_2 and loss of $H_2C(O)$ with reaction energies of 14.6 kcal/mol (0.63 eV) and 39.4 kcal/mol (1.71 eV) respectively. The direct bond breaking path involving CO_2 loss discussed above is also not observed in the CID experiments.

The loss of H_2 from the precursor ion is a low-energy process for all of the studied single amino acid structures but it is never observed in our experiments. However, Bowie and coworkers[3] observed elimination of H_2, H radical, and H_2O, as well as CO loss (but not loss of $HN{=}CH_2$) in high-energy CID experiments on deprotonated glycine. These experiments employed a double focusing sector mass analyzer where precursor ions generated by chemical ionization were accelerated with kilovolt (kV) potentials into helium. These high-energy CID experiments are both more energetic than our low-energy SORI CID experiments and pump energy into the precursor ions faster (μs timescale as opposed to our ms timescale).

In order to further understand potential processes, transition states starting from the deprotonated glycine anion were studied for selected reactions. These results are shown in Table 2.3 with additional results in the Appendix (Table A2.1). The barrier for H_2 elimination from deprotonated glycine is predicted to be 94 kcal/mol at the B3LYP/DZVP2 level. This is in the same range as elimination of H_2 across C-C bonds.[35] This barrier is high enough that the C-H or N-H bonds in the system may directly break before the energy is localized in this region preventing loss of H_2. If the energy is input quickly at high collision energy as in the experiments of Bowie and coworkers,[3] then both H and H_2 loss can be observed. The barrier for loss of CO is also high, 104 kcal/mol, but in this case, the O^- transfer is a consequence of the direct cleavage of the $–H_2C\text{-}CO_2^-$ bond and the transfer of the O^- to the CH_2 group so the energy can be localized in this region. The formation of HCO_2^- arises by a proton transfer step to form the carbene $NH_2C{:}H$ with essentially no barrier as compared to the endothermicity of the process (the barrier is slightly below the endothermicity at the G3MP2 level due to B3LYP predicting barriers that are too low as well as the formation of a product channel complex that is weakly bound). The transfer of the NH hydrogen to from the imine $HN{=}CH_2$ has an energy barrier $\Delta H^{\dagger}(298\ K)$ = 45.3 kcal/mol. The unobserved $H_2C(O)$ + $HNCH(O)^-$ path is a multistep process and was not further studied.

The bond energies for the deprotonated glycine amide anion are analogous to those of the deprotonated glycine anion (Table 2.2). Cleavage of bond 2 is preferred, leading to formation of HN=C=O rather than CO_2 as for the glycine anion. This fragmentation requires approximately 5 kcal/mol more energy than that for deprotonated glycine because the formation of HN=C=O is not as energetically favorable as the formation of CO_2. Both glycine and its amide could generate

a three-member α-lactam structure in the CID process although this process is not observed; the α-lactam derived from the glycine amide is more stable than the α-lactone derived from glycine.

The CID spectrum for glycine amide (Figure A2.2) shows loss of the carbene NH_2C:H, which requires 62.5 kcal/mol (2.71 ev) and might convert to the lower energy isomer $HN=CH_2$ which requires only 26.6 kcal/mol (1.15 eV) overall. As in the case of the deprotonated glycine anion, this process occurs with a barrier slightly less than the endothermicity of the reaction (Table 2.3). The other major CID product is elimination of CH_3NH_2 with formation of the $N=C=O^-$ anion. This is predicted to be an exothermic process with an energy of -2.2 kcal/mol (-0.10 eV), showing that the deprotonated glycine amide anion is metastable (although it does not dissociate in the FT-ICR in the absence of collisional activation). This reaction is a multistep process that involves the initial formation of $H_2NCH_2^-$ and HN=C=O (74.9 kcal/mol) by cleavage of the C-C bond in the amide (Figure A2.9). The hydrogen is subsequently transferred from the nitrogen to the negatively charged carbon to form methylamine; the final products are 77.1 kcal/mol lower in energy than $H_2NCH_2^-$ + HN=C=O so there is effectively no barrier for the proton transfer process. In contrast to the deprotonated glycine anion, the loss of CO is not observed even though this process is of comparable energy, 53.2 kcal/mol (2.31 eV). As for glycine above, the loss of H_2 (13.8 kcal/mol, 0.60 eV) and $H_2C(O)$ (41.7 kcal/mol, 1.81 eV) are not observed experimentally under low-energy CID despite energetically favorable reaction energies due to large energy barriers for the transition state in the case of H_2 or a multi-step mechanism that requires significant dissociation and/or rearrangement, to proceed to the predicted products. Loss of NH_3 is also not observed and this has a reasonably high barrier (Table 2.3). The production of H_2O from the glycine amide anion requires 63.2 kcal/mol (2.86 eV). The loss of H_2O is the result of a multi-step mechanism that involves a proton transfer from

the -NH^- group at the C-terminus to the carboxylic oxygen (Figure A2.10). As this newly formed hydroxide group separates from the rest of the molecule, it collects a hydrogen from the N-terminus. As in the case of loss of H_2CO or H_2O, there is no single step mechanism that can produce $H_2C{=}CH_2$.

The deprotonated alanine (Table 2.4) and alanine amide (Table 2.5) anions have backbone fragmentation energetics similar to glycine and glycine amide. The cleavage of each bond in alanine and alanine amide without molecular rearrangement is within 3 kcal/mol of the corresponding bond energy for glycine and glycine amide, consistent with a small effect of substituting a methyl group for a hydrogen. As for glycine and glycine amide, the most stable bond dissociation products for alanine and alanine amide are formed by breaking bond 2, the backbone C-C bond. The energetically favorable structures resulting from this dissociation for alanine and alanine amide are nearly identical to those formed by the same dissociation for glycine and glycine amide, except for the backbone substitution of a methyl group for hydrogen. The α-lactone that could be derived from deprotonated alanine was predicted to be more stable than the α-lactone that could be derived from deprotonated glycine because the methyl group further stabilizes the 3-member ring; processes leading to these species are not observed in our experiments.

For the alanine anion, CID results in the loss of the neutrals NH_3, $HN{=}CHCH_3/H_2N{=}C{:}CH_3$ and HCO_2H with energies of 16.0 (0.69 eV), 56.6 (2.45 eV), and 59.0 (2.56 eV) kcal/mol, respectively (Figure A2.3). These peaks also correspond to the most intense high-energy CID product ions in the work of Bowie and coworkers.[3] Lower energy processes to form $HN{=}CH_2/H_2NC{:}H$, CO_2, CO, C_2H_4, and H_2CO with energies of 25.5 (1.11 eV), 52.6 (2.28 eV), 47.1 (2.04 eV), 16.3 (0.71 eV), and 38.2 (1.66 eV) kcal/mol are not observed in our low-

energy CID experiments. The barrier to transfer a CH_3 group to form $CH_3CO_2^-$ is much higher than the barrier to transfer a proton to form HCO_2^- so only the latter path is observed (Table 2.3). The elimination of the H radical, H_2, H_2O, CO, and CO_2 are, however, observed in the high-energy CID spectrum.[3] The respective energies for these CID processes are predicted to be 87.2 (3.78 eV), 20.8 (0.90 eV), 68.9 (2.99 eV), 47.1 (2.04 eV), and 52.6 (2.28 eV) kcal/mol.

For the deprotonated alanine amide anion, loss of NH_3 and H_2O (a minor product) are observed in the low-energy CID spectrum and are low-energy processes of 19.4 (0.84 eV) and 61.3 (2.78 eV) kcal/mol, respectively (Figure A2.4). Loss of $HN{=}CHCH_3$ is a higher energy process requiring 57.6 kcal/mol (2.50 eV) that also occurs. The highest abundance in the low-energy CID spectrum is loss of $CH_3CH_2NH_2$ to form $N{=}C{=}O^-$, which is the most favorable process energetically. This process is slightly exothermic, -0.2 kcal/mol, again showing the alanine amide anion to be metastable. Low-energy processes that are not observed include loss of HCN, C_2H_4, and $HN{=}CH_2$ with energies of 44.5 (1.93 eV), 18.7 (0.81 eV), and 26.0 (1.13 eV) kcal/mol, respectively. It is possible that H_2O loss is observed for alanine amide and not for glycine amide due to the stabilizing nature of the methyl on alanine amide's central carbon (Figure A2.11). The elimination barriers studied are similar to those discussed above (Table 2.3).

Dipeptides and Corresponding Amides

We first consider the bond energies in the backbone of the deprotonated diglycine anion formed by loss of a proton from the CO_2H group and in the diglycine amide anion generated by loss of a proton from the $C(O)NH_2$ group (Figure 2.2 and Tables 2.6 and 2.7). The corresponding results for deprotonated dialanine anion and dialanine amide are in Tables 2.8 and 2.9. The homolytic cleavage of bond 5 to produce a CO_2 or $HN{=}C{=}O$ is the most favored reaction pathway for the adiabatic fragmentation of diglycine, dialanine, and the corresponding amides,

analogous to the favored bond breaking for the glycine and alanine amino acids. This is the bond cleavage that would produce members of a_2^- series using standard peptide fragmentation nomenclature. As with the deprotonated anion acids, breaking bond 5 of the deprotonated diglycine and dialanine dianions requires 5 kcal/mol less energy than breaking the corresponding bond of the deprotonated amide because CO_2 is a more favorable product than HN=C=O. The lowest energy ion of this homolytic cleavage undergoes a rearrangement whereby a proton is transferred from the N of the amide to form CH_3.

The adiabatic cleavage of the peptide bond, bond 3, for deprotonated diglycine and diglycine amide requires more than 100 kcal/mol. Rearrangement of the product neutrals or anions leads to lower energies but these are still high at greater than 70 kcal/mol. In the CID spectrum, which is shown in Figure 2.3, the observed y_1- anion involves cleavage of the central amide bond accompanied by hydrogen transfer from the neutral to the anion to generate $NH_2CHRCO_2^-$. The resulting product ion has lost the N-terminal residue from the dipeptide. This is a very common low-energy CID pathway for peptides. Our calculations show that in formation of y_1^-, which involves loss of the mass and chemical formula of a glycine residue, the eliminated neutral is most likely a 3-membered ring with a C=O group (cyclopropanamide) at an energy of 64.1 kcal/mol (2.78 eV). In contrast, a substantially rearranged neutral with a cyano group is predicted to require 41.6 kcal/mol (1.80 eV). Very similar results are predicted for the formation of the y_1^- anion from deprotonated diglycine amide, dialanine, and dialanine amide anions.

For the dipeptide species and their amides, the loss of a single amino acid residue (rearranged, as discussed above) is the most intense product ion in the CID spectra. Among the amide species studied, the $N{\equiv}C\text{-}O^-$ ion was universally observed. Fragmentation pathways found in the low-energy CID spectrum of the diglycine anion (Figure 2.3) correspond to loss of CO_2,

formic acid (HCOOH), and the 3-member ring with a C=O group (which might further rearrange to N≡CCH_2OH) with reaction energies of 11.9 (0.52 eV), 30.4 (1.32 eV), and 64.1 (2.78 eV) kcal/mol, respectively. The energy for the loss of CO_2 without rearrangement following CID is only 20.5 kcal/mol, which is favorable energetically but is nearly twice as much energy as the pathway that allows for rearrangement. Two reaction pathways in the CID spectrum suggest loss of two neutral species: CO_2 + HN=CH_2 (37.5 kcal/mol, 1.63 eV) and CO_2 + H_2O (41.4 kcal/mol, 1.80 eV). The path to produce CO_2 + H_2O is 7 kcal/mol (0.30 eV) less endothermic than the path to produce HOC(O)OH. Low-energy pathways that are not observed in the CID spectrum include the loss of NH_3, H_2O, and $C_4H_6NO_2$ with energies of 2.9 (0.13 eV), 11.9 (0.52 eV), and 48.2 (2.09 eV) kcal/mol, respectively.

The major product ions in the CID spectrum for diglycine amide anion (Figure 2.4) suggest the loss of H_2O, $H_2NCH(O)$, and c-$H_2CNHC(O)$ (cyclopropanamide) with reaction energies of 31.6 (1.37 eV), 19.7 (0.85 eV), and 63.8 (2.77 eV) kcal/mol. Elimination of H_2O has also been observed as a prominent CID product for larger deprotonated peptide amides (e.g., penta-alanine amide) even if the molecule lacks a hydroxyl group.11 Additionally, as predicted for both the glycine and alanine amides, the formation of $N=C=O^-$ by loss of $H_2NC(O)NHCH_2CH_3$ shows that the diglycine amide anion is metastable with an exothermic reaction energy value (-4.9 kcal/mol, -0.21 eV). There are several fragmentation pathways that are energetically favorable yet are not found in the CID spectrum for deprotonated diglycine amide, including the exothermic loss of NH_3 (-2.8 kcal/mol, -0.12 eV); the endothermic losses of HNCO (25.0 kcal/mol, 1.08 eV), CO_2 (22.7 kcal/mol, 0.98 eV), and $H_2NC(O)NHCH_2CH_2C{\equiv}N$ (48.5 kcal/mol, 2.10 eV); and the bimolecular loss of CO_2 + HN=CH_2 (45.0 kcal/mol, 1.95 eV).

For the dialanine anion, loss of CO_2, NH_2CO_2H, and c-$C(CH_3)HNHC(O)$, which can rearrange to $H_3CCH{=}NCH(O)$, $H_2C{=}CHC(O)OH$, and $H_2NC(O)C(O)CH_3$, are observed in the CID spectrum (Figure A2.7) with respective energies of 20.5 (0.89 eV), 43.4 (1.88 eV), 51.6 (2.24 eV), and 41.9 (1.82 eV) kcal/mol. Though the bimolecular loss of CO_2 + NH_3 could be suggested for the observed mass loss of 61 Da, it was found that the unimolecular loss of $H_2NC(O)OH$ (20.9 kcal/mol, 0.91 eV) is a significantly more favorable pathway by over 50 kcal/mol. Despite having favorable energetics, the exothermic loss of NH_3 (-0.6 kcal/mol, -0.03 eV) and the endothermic loss of H_2O (13.7 kcal/mol, 0.59 eV) are not observed in the CID spectrum. Again, the NH_3 loss energetics show that the deprotonated dialanine anion may be a metastable species.

The observed dialanine amide anion fragmentation pathways (Figure A2.8) include the loss of H_2O, $H_2NCH(O)$, $HC(NH_2)_3$, c-$C(CH_3)HNHC(O)$, c-$C_5H_{10}N_2O$, and $H_2NC(O)NHCH_2CH(CH_3)_2$, with respective energies of 34.2 (1.48 eV), 26.4 (1.14 eV), 45.5 (1.97 eV), 62.9 (2.72 eV), 17.4 (0.75 eV), and 0.3 (0.01 eV) kcal/mol. A loss of 61 Da suggests a possible bimolecular loss of $H_2NCH(O)$ + NH_2, but energetically this loss and the loss of the unimolecular $(H_2N)_2COH$ are over 50 kcal/mol higher in energy than the loss of $HC(NH_3)_3$. There are several fragmentation pathways that are energetically favorable, yet not observed, including the loss of NH_3 (-0.6 kcal/mol, -0.03 eV), $O{=}C{=}NH$ (25.4 kcal/mol, 1.10 eV), CO_2 (13.9 kcal/mol, 0.6 eV), and $HN{=}C(CH_3)C(O)NH_2$ (28.8 kcal/mol, 1.25 eV). Again, the deprotonated alanine amide anion is a metastable species with respect to the loss of NH_3.

Discussion

The loss of mass 18 corresponding to the H_2O molecule is observed for the amides of alanine, diglycine, and dialanine. Elimination of H_2O is also a common process in the CID spectra of larger deprotonated peptide amides, including neutral species that have no hydroxyl

group.[5] Even though the energy for loss of H_2O is lower for the non-amide dipeptides, this process is not observed. The loss of H_2O arises because the terminal anionic site in the amides is $R(R')C(H)\text{-}C(=O)NH^-$, where R' is the substituent side chain of the amino acid (in this case H or CH_3) and R is the remainder of the peptide. In the first step of this process (reaction (2.1) for the glycine amide anion), a hydrogen from the NH^- group is transferred to the oxygen on the carbonyl group at the C terminus to generate the $H_2NCHC(OH)NH^-$ intermediate (Energies at the

$$H_2NCH_2C(O)NH^- \xrightarrow{H^+\text{ transfer}} TS \rightarrow H_2NCHC(OH)NH^- \xrightarrow{H^+\text{ transfer}} H_2NCH^-C{\equiv}N + H_2O \quad (2.1a)$$
51.9 39.6 42.0

$$\rightarrow (HN^-)HC{=}C{=}NH + H_2O \quad (2.1b)$$
65.9

B3LYP/DZVP2 level). As an H is transferred from the amide NH (reaction (2.1a)) or from the NH_2 terminus (reaction (2.1b)) to the OH group, the H_2O is formed with essentially no barrier and dissociates from the anion. This process cannot occur at the acid terminus of the normal $[M - H]^-$ amino acid species due to the missing NH^- group, so we do not observe it for the non-amide precursor anions. For glycine amide, alanine amide, diglycine amide, and dialanine amide, the overall enthalpies for reaction (2.1a) are 38.3 (1.66 eV), 39.4 (1.70 eV), 31.6 (1.37 eV), and 34.2 (1.48 eV) kcal/mol, respectively, and those for reaction (2.1b) are 63.2 (2.74 eV), 61.3 (2.66 eV), 44.9 (1.95 eV), and 49.4 kcal/mol (2.14 eV), respectively. We note that the elimination of H_2O for glycine amide is not observed. There are lower energy isomers than the cyano carbanions but these involve substantially more rearrangement. The lowest energy anion isomers for alanine amide, diglycine amide, and dialanine amide are lower than the cyano-substituted carbanion by 9.5, 16.0, and 14.7 kcal/mol, respectively (see Appendix for the structures).

The loss of H_2 is a low-energy pathway for all anions in this work, It is never observed in our low-energy CID mass spectra although it is in high-energy CID spectra.[3] Values range from less than 10 kcal/mol (dialanine and diglycine) to approximately 24 kcal/mol (dialanine amide).

The amides have lower enthalpies for the H_2 elimination reaction (6-14 kcal/mol) than for the amino acids and dipeptides (14-24 kcal/mol).

When the C-CO_2^- bond is cleaved, the electron can either remain with the acid fragment or be ejected leaving CO_2 and a radical. The difference between CO_2 loss and $CO_2 + e^-$ loss for glycine and alanine are 54-60 kcal/mol because the electron affinity of the radical is small. A much larger difference is predicted for the dipeptides so that the loss of CO_2 is much more favorable, 12-23 kcal/mol, due to the much higher electron affinity of the radical. The dialanine and diglycine anions have CID spectra with a product ion corresponding to loss of CO_2 (loss of 44 Da). Notably, the loss of CO_2 from dialanine is energetically very similar with and without rearrangement of the remaining anionic product species (20.8 kcal/mol for rearranged species and 20.5 kcal/mol for the adiabatic bond dissociation energy). The adiabatic energy for loss of CO_2 from diglycine is energetically similar (20.5 kcal/mol), but the rearrangement of the remaining anionic species can be more energetically favorable (11.9 kcal/mol).

The calculations suggest that the loss of HN=C=O could be present in the CID spectra for the amides, but peaks corresponding to this loss are never observed. This is despite the fact that the energetics for losing this neutral species for diglycine amide and dialanine amide are similar to those for the loss of CO_2 for those species; it is approximately 5 kcal/mol higher in energy to lose HN=C=O as compared to loss of CO_2. The energetics are also similar when comparing the loss of HN=C=O and CO_2 for glycine amide and alanine amide (approximately 2-8 kcal/mol higher to lose HN=C=O as compared to loss of CO_2). The loss of 29 Da, most likely due to elimination of the imine $HN=CH_2$, is another low-energy pathway (less than 30 kcal/mol for glycine, alanine, and their amides) that is observed experimentally for glycine, but not for alanine.

Conclusions

Composite correlated molecular orbital theory calculations at the G3(MP2) level suggest that the dissociation resulting from low-energy CID mass spectrometry produced fragments that can undergo significant rearrangement in order to obtain low-energy isomers. A wide range of possible conformations were explored using density functional theory at the B3LYP/DZVP2 level. The G3MP2 heats of formation for the neutral gas-phase species were in excellent agreement with the available experimental values. The loss of H_2O was observed experimentally only for alanine amide, diglycine amide, and dialanine amide. The $R(R')C(H)\text{-}C(=O)NH^-$ moiety is only found in the deprotonated amide structures and is apparently necessary for the formation of water from the carboxyl group and the adjacent hydrogens on both the amide group and the neighboring carbon leading to a cyano carbanion. Glycine amide is the only exception as no H_2O loss peak was observed for this amide precursor ion. Lower energy pathways that required significantly more structural rearrangements were found, but are not likely as there is no evidence of such species for the non-amide anions. Each observed deprotonated amino acid and dipeptide anion with or without the amide modification were predicted to have energetically favorable pathways that were not observed in their mass spectra. The mass spectra of the amino acids and dipeptides and their respective amides are fairly similar, but the amide spectra give additional insight into the behavior of a fragment that does not originate from the terminus of a longer peptide chain. The loss of the C-terminus (CO_2 for regular amino acids and dipeptides, HN=C=O for respective amides) is always energetically favorable, even though it is only observed for the deprotonated dipeptide anions. The bond dissociation energies for the dipeptides and their amides were found to be very similar when the N-terminus did not retain the negative charge upon dissociation with differences always less than 5 kcal/mol and often less than 1

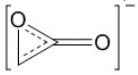
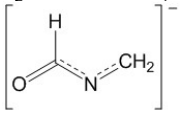
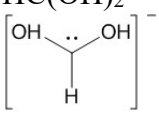
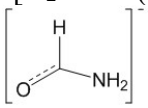
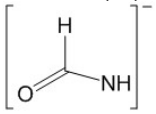
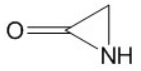
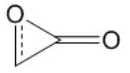
kcal/mol. However, the bond dissociation energies for these species for which the N-terminus is the site of negative charge did not display this trend. When the negative charge is not retained by the C-terminal fragment, the dipeptide amides showed lower bond dissociation energies than their non-amide counterparts, with a difference of 5-10 kcal/mol. Breaking the peptide bond in the dipeptides followed by hydrogen transfer to the anion yields the y_1^- anion, $NH_2CHRCO_2^-$, commonly observed in CID experiments. For the dipeptides this usually generated a cyclopropanamide neutral and requires about 2.7-2.8 eV. A lower energy pathway to generate a cyano compound with substantial rearrangement required 1.8-1.9 eV accompanying the formation of the y_1^- anion.

The computational results, in combination with our low-energy CID and other high-energy CID[3] experiments, show that complex dynamical processes are playing a role in determining the fragmentation patterns. The energy of the collision has a clear impact on the observed fragmentation patterns. The observed fragments may require rearrangements to reach low-energy structures and are not single event processes. This is clear from the thermodynamic predictions, the calculated transition states and energy barriers, and the observed peaks in the mass spectra. In addition, it can be extremely difficult to locate transition states for highly endothermic reactions where the transition state is in the exit channel very near to the products as found under CID conditions, as a consequence of Hammond's Postulate[36] and the Bell-Evans-Polanyi Principle.[37,38] Furthermore, the randomization of the collision energy which has been converted into vibrational energy will depend on the collision time and the complexity of the molecule.[39,40] Thus, not only does one have to map out the appropriate potential energy surfaces, one also requires detailed dynamical studies of each dynamical event at different collision energies, especially to understand how the collision energy is redistributed into the

rearrangement and decomposition modes. Such studies even for a simple species such as the glycine anion are not feasible today. However, our computational results do provide reliable adiabatic bond dissociation energies and the structures including specific conformations of a broad range of potential structures of the decomposition products, and in combination with experiments, lead to an improved understanding of the CID processes.

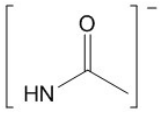
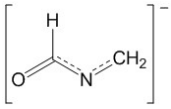
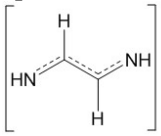
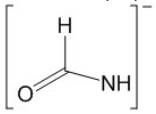
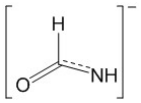
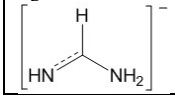
Tables

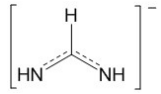
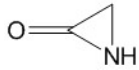
Table 2.1. Fragmentation Energies in Glycine $[M - H]^-$ in kcal/mol[a]

Anion	Neutral	Mass Loss [c]	ΔH_f A	ΔH_f B	ΔH_{rxn}
$^2[H_2N{\bullet}CHCO_2^-]$ [b]	$^2[H{\bullet}]$	1	-80.5	52.1	88.8
$^2[HN{\bullet}CH_2CO_2^-]$	$^2[H{\bullet}]$	1	-83.0	52.1	86.3
$HN{=}CHCO_2^-$	H_2	2	-100.5	-1.1	15.6
$^2[{\bullet}CH_2CO_2^-]$ [b] (1)	$^2[{\bullet}NH_2]$	16	-75.4	44.5	86.3
c-$HCOC(O)^-$ 	NH_3	17	-37.9	-10.0	69.3
$^2[H_2C{\bullet}NCH(O)]^-$ 	$^2[HO{\bullet}]$	17	-16.3	8.3	109.2
$HN{=}CHC(O)^-$	H_2O	18	12.3	-57.4	72.1
$HC(OH)_2^-$ 	HCN	27	-44.8	31.2	103.6
$H_2NCH_2(O)^-$	CO	28 (obs)	-39.3	-27.4	50.5
HCO_2^-	H_2N-C:H	29(obs)	-112.3	57.2	62.1
HCO_2^-	$HN{=}CH_2$	29 (obs)	-112.3	21.3	26.2
$^2[H_2N{\bullet}CH(O)]^-$ 	$^2[HCO{\bullet}]$	29	-9.0	9.5	117.7
$HN^-HC(O)$ 	$H_2C(O)$	30	-50.1	-27.7	39.4
$HNCH_3^-$ [b] (2)	CO_2	44	32.6	-94.9	54.9
$^2[NH_2CH_2{\bullet}]$ [b] (2) + e^-	CO_2		36.9	-94.9	59.2
HO^-	c-$H_2CNHC(O)$ 	57	-33.0	10.6	94.8
HO^-	$N{\equiv}CCH_2OH$	57	-33.0	-11.9	72.3
NH_2^- [b] (1)	c-$H_2COC(O)$ 	58	27.3	-41.4	103.1

[a] $H_2NCH_2CO_2^-$. ΔH_f(298 K) = -117.2 kcal/mol. Mass = 74 Da. [b] 'Adiabatic' bond cleavage. Bond labels in parentheses. See Figure 2.1. [c] "obs" indicates that this fragmentation pathway is observed in our CID spectra.

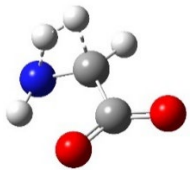
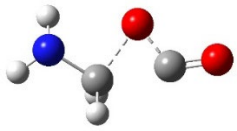
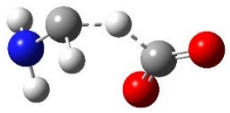
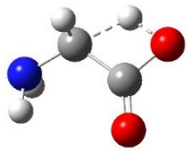
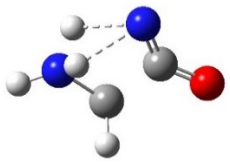
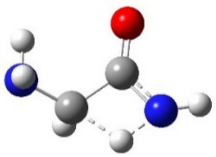
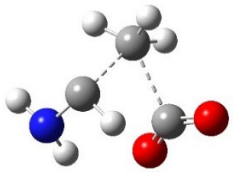
Table 2.2. Fragmentation Energies in Glycine Amide $[M - H]^-$ in kcal/mol[a]

Anion	Neutral	Mass Loss [c]	ΔH_f A	ΔH_f B	ΔH_{rxn}
$^2[H_2N{\bullet}CHC(O)NH]^-$ [b]	$^2[H{\bullet}]$	1	-19.5	*52.1*	88.0
$HN{=}CHC(O)NH^-$	H_2	2	-40.5	-1.1	13.8
$H_2NCH_2C(O)^-$ [b]	$^3[:NH]$	15	-6.4	84.0	133.0
$H_3CC(O)NH^-$ 	$^3[:NH]$	15	-59.6	84.0	79.8
$^2[{\bullet}CH_2C(O)NH]^-$ [b] (1)	$^2[{\bullet}NH_2]$	16	-15.3	44.5	84.6
$^2[H_2C{\bullet}NCH(O)]^-$ 	$^2[{\bullet}NH_2]$	16	-16.3	44.5	83.6
$HN{=}CHC(O)^-$	NH_3	17	12.3	-10.0	57.7
$^2[HNCH{\bullet}CHNH]^-$ 	$^2[HO{\bullet}]$	17	39.2	8.3	102.9
$(HN^-)CH{=}C{=}NH$	H_2O	18	65.2	-57.4	63.2
$H_2NCH^-C{\equiv}N$	H_2O	18	40.3	-57.4	38.3
$H_2NCH_2NH^-$	CO	28	25.2	-27.4	53.2
$HN^-HC(O)$ 	H_2N-C:H	29(obs)	-50.1	57.2	62.5
$HN{=}CH{-}O^-$ 	$HN{=}CH_2$	29 (obs)	-50.1	21.3	26.6
$^2[HN{\bullet}CHNH_2]^-$ 	$^2[HC{\bullet}(O)]$	29	53.1	9.5	118.0

HNCHNH$^-$ 	$H_2C(O)$	30	14.0	-27.7	41.7
^{2}[HN=•C=O]$^-$ [b] (2)	^{2}[H_2NC•H_2]	30	-12.8	36.9	79.5
N=C=O$^-$	CH_3NH_2	31 (obs)	-53.5	-4.1	-2.2
$H_2NCH_2^-$ [b] (2)	HN=C=O	43	47.8	-28.3	74.9
$HNCH_3^-$	HN=C=O	43	32.6	-28.3	59.7
^{2}[H_2N•CH_2] [b] (2) + e$^-$	HN=C=O		36.9	-28.3	64.0
HO$^-$	HN=CHCH=NH	56	-33.0	41.7	64.1
NH_2^- [b] (1)	c-H_2CNHC(O) 	57	27.3	10.6	93.3
NH_2^-	N≡CCH_2OH	57	27.3	-11.9	70.8

[a] $H_2NCH_2C(O)NH^-$. ΔH_f(298 K) = -55.4 kcal/mol. Mass = 73 Da. [b] 'Adiabatic' bond cleavage. Bond labels in parentheses. See Figure 2.1. [c] "obs" indicates that this fragmentation pathway is observed in our CID spectra.

Table 2.3. Selected Transition States and Energy Barriers at the B3LYP/DZVP2 Level in kcal/mol

Products	**ΔH_{298}**	**ΔG_{298}**	**Image of TS**
Glycine			
$H_2 + HN{=}CHCO_2^-$	94.0	94.3	
$CO + H_2NCH_2(O)^-$	103.7	103.5	
$H_2NC{:}H + HCO_2^-$	56.8	56.4	
$H_2NC{:}H + HOCO^-$	52.7	53.2	
Glycine Amide			
$NH_3 + HN{=}CHC(O)^-$	81.6	82.4	
$H_2NC{:}H + CONH_2^-$	52.2	52.4	
Alanine			
$H_2NC{:}H + H_3CCO_2^-$	93.8	93.9	
Alanine Amide			

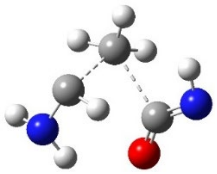
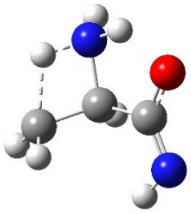
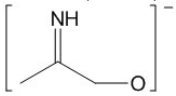
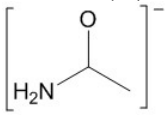
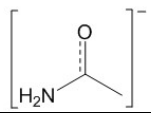
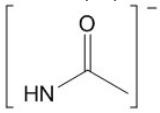
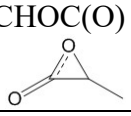
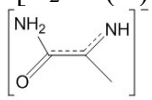
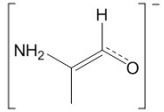
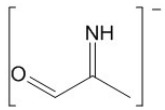
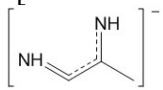
$H_2NC{:}H + H_3CC(O)NH^-$	92.8	92.4	
$NH_3 + HNC(O)CH{=}CH_2^-$	57.5	58.2	

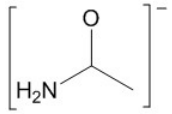
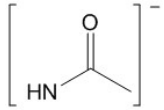
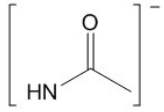
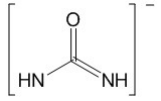
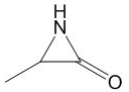
Table 2.4. Fragmentation Energies in Alanine $[M - H]^-$ in kcal/mol[a]

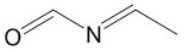
Anion	Neutral	Mass Loss [c]	ΔH_f A	ΔH_f B	ΔH_{rxn}
$^2[H_2N{\bullet}C(CH_3)CO_2]^-$ [b]	$^2[H{\bullet}]$	1	-90.4	*52.1*	87.2
$HN{=}C(CH_3)CO_2^-$	H_2	2	-102.6	-1.1	21.8
$H_2NC(=CH_2)CO_2^-$	H_2	2	-103.6	-1.1	20.8
$^2[H_2N{\bullet}CHCO_2]^-$ [b]	$^2[{\bullet}CH_3]$	15	-46.9	34.2	112.8
$^2[HN{\bullet}CH_2CO_2]^-$	$^2[{\bullet}CH_3]$	15	-83.0	34.2	76.7
$^2[H_3C{\bullet}CHCO_2]^-$ [b] (1)	$^2[{\bullet}NH_2]$	16	-83.9	44.5	86.1
$H_2C{=}CHCO_2^-$	NH_3	17 (obs)	-99.5	-10.0	16.0
$^2[H_2CC(O){\bullet}NCH_3]^-$	$^2[HO{\bullet}]$	17	-16.0	8.3	101.2
$H_3CC(=NH)C(O)^-$ 	H_2O	18	0.8	-57.4	68.9
$H_3CC(OH)_2^-$	HCN	27	-65.6	31.2	91.1
$^2[H_2N{\bullet}C(O)OH]^-$	$^2[H_2C{=}C{\bullet}H]$	27	-66.2	70.3	129.6
$H_2NCH(O)CH_3^-$ 	CO	28	-51.0	-27.4	47.1
$H_2NCO_2^-$	$H_2C{=}CH_2$	28	-121.0	11.9	16.4
$H_3CCO_2^-$	$H_2NC{:}H$	29	-121.6	57.2	61.1
$H_3CCO_2^-$	$HN{=}CH_2$	29	-121.6	21.3	25.2
$^2[H_2N{\bullet}C(O)CH_3]^-$ 	$^2[HC{\bullet}(O)]$	29	-21.8	9.5	113.2
$^2[HN{\bullet}CO_2]^-$	$^2[H_3CC{\bullet}H_2]$	29	-73.0	28.9	81.4
$H_3CC(O)NH^-$ 	$H_2C(O)$	30	-59.6	-27.7	38.2
HCO_2^-	$HN{=}CHCH_3$	43 (obs)	-112.3	43.4	56.6
$^2[H_3CC{\bullet}HO]^-$	$^2[O{=}C{\bullet}NH_2]$	44	-16.1	-2.2	107.2
$HNCH_2CH_3^-$ [b] (2)	CO_2	44	20.8	-94.9	52.6
$^2[H_2N{\bullet}CHCH_3]$ [b] (2) + e^-	CO_2		29.2	-94.9	59.8
$HNCH{=}CH_2^-$	HCOOH	46 (obs)	23.5	-90.0	59.0
OH^-	$H_3CCH{=}NCH(O)$	71	-33.0	-16.4	76.1
NH_2^- [b1]	CH_3-c-CHOC(O) 	72	27.3	-52.2	100.6

[a] $H_2NCH(CH_3)CO_2^-$. ΔH_f(298 K) = -125.5 kcal/mol. Mass = 88 Da. [b] 'Adiabatic' bond cleavage. Bond labels in parentheses. see Figure 2.1. [c] "obs" indicates that this fragmentation pathway is observed in our CID spectra.

Table 2.5. Fragmentation Energies in Alanine Amide [M - H]$^-$ in kcal/mol[a]

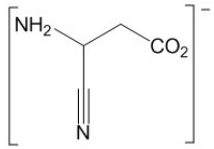
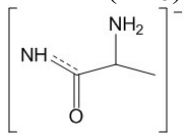
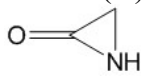
Anion	Neutral	Mass Loss [c]	ΔH_f A	ΔH_f B	ΔH_{rxn}
$^2[H_2N{\bullet}C(CH_3)C(O)NH^-$ [b]	$^2[H{\bullet}]$	1	-29.4	*52.1*	87.0
$^2[H_2NC(O){\bullet}C(NH)CH_3^-$ 	$^2[H{\bullet}]$	1	-36.6	*52.1*	79.8
$HN{=}C(CH_3)C(O)NH^-$	H_2	2	-50.3	-1.1	12.9
$^2[H_2N{\bullet}CHC(O)NH]^-$ [b]	$^2[{\bullet}CH_3]$	15	-19.5	34.2	79.0
$H_2NCH(CH_3)C(O)^-$ [b]	$^3[NH]$	15	-15.2	84.0	133.1
$H_2NC(CH_3){=}CH(O)^-$ 	$^3[NH]$	15	-41.2	84.0	107.1
$^2[H_3C{\bullet}CHC(O)NH]^-$ [b] (1)	$^2[{\bullet}NH_2]$	16	-23.7	44.5	85.1
$^2[HC(O){\bullet}C(=NH)CH_3]^-$ 	$^2[{\bullet}NH_2]$	16	-30.0	44.5	78.8
$HNC(O)CH{=}CH_2$	NH_3	17 (obs)	-34.9	-10.0	19.4
$^2[HN{=}CH{\bullet}C(NH)CH_3]^-$ 	$^2[HO{\bullet}]$	17	30.0	8.3	102.6
$(HN^-)(CH_3)C{=}C{=}NH$	H_2O	18 (obs)	54.4	-57.4	61.3
$H_2NC^-(CH_3)(C{\equiv}N)$	H_2O	18(obs)	32.5	-57.4	39.4

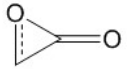
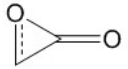
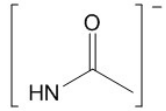
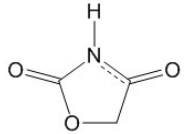
$H_2NCH(O)CH_3^-$ 	HCN	27	-51.0	31.2	44.5
$H_2NCH_2NCH_3^-$	CO	28	23.4	-27.4	60.3
$H_2NC(O)NH^-$	$H_2C{=}CH_2$	28	-57.5	11.9	18.7
$H_3CC(O)NH^-$ 	$H_2NC{:}H$	29	-59.6	57.2	61.9
$H_3CC(O)NH^-$ 	$HN{=}CH_2$	29	-59.6	21.3	26.0
$^2[H_2NCH_2{\bullet}CHNH]^-$	$^2[HC{\bullet}(O)]$	29	44.9	9.5	118.7
$^2[HN{\bullet}C(O){=}NH]^-$ 	$^2[H_3CC{\bullet}H_2]$	29	-9.2	28.9	84.1
$HNCH_2CH_3^-$	$H_2C(O)$	30	20.8	-27.7	57.4
$H_3CCH_2NH^-$ [b] (2)	$HN{=}C{=}O$	43	20.8	-28.3	56.8
$^2[H_2N{\bullet}CHCH_3]$ [b] (2) + e^-	$HN{=}C{=}O$		29.2	-28.3	65.2
$HN{=}CH{-}O^-$	$HN{=}CHCH_3$	43 (obs)	-50.1	43.4	57.6
$^2[HN{=}{\bullet}C{=}O]^-$ [b] (2)	$^2[H_2NC{\bullet}HCH_3]$	44	-12.8	29.2	80.7
$N{=}C{=}O^-$	$H_3CCH_2NH_2$	45 (obs)	-53.5	-11.0	-0.2
NH_2^- [b] (1)	c-$C(CH_3)HNHC(O)$ 	71	27.3	1.5	93.1

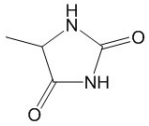
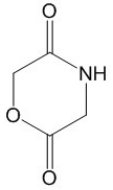
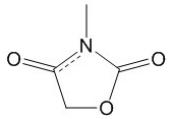
NH_2^-	$HC(O)N{=}CHCH_3$ 	71	27.3	-16.4	75.2

[a] $H_2NCH(CH_3)C(O)NH^-$. ΔH_f(298 K) = -64.3 kcal/mol. Mass = 87 Da. [b] 'Adiabatic' bond cleavage. Bond labels in parentheses. See Figure 2.1. [c] "obs" indicates that this fragmentation pathway is observed in our CID spectra.

Table 2.6. Fragmentation Energies in Diglycine $[M - H]^-$ in kcal/mol[a]

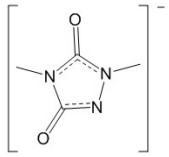
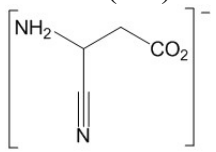
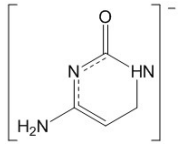
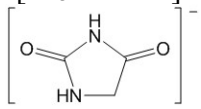
Anion	Neutral	Mass Loss [c]	ΔH_f A	ΔH_f B	ΔH_{rxn}
$H_2NC(O)NHCH=CHCO_2^-$	H_2	2	-161.1	-1.1	8.5
$^2[H_2•CC(O)NHCH_2CONH]^-$ [b] (1)	$^2[•NH_2]$	16	-132.1	44.5	83.1
$^2[H_2NC(O)CH•CHC(O)OH]^-$	$^2[•NH_2]$	16	-142.5	44.5	72.7
$H_2NC(O)CH=CHCO_2^-$	NH_3	17	-157.8	-10.0	2.9
$H_2NCH(C≡N)CH_2CO_2^-$ 	H_2O	18	-101.4	-57.4	11.9
$^2[•C(O)NHCH_2CO_2]^-$ [b] (2)	$^2[H_2NC•H_2]$	30	-127.6	36.9	80.0
$H_2NCH_2C(O)=NCH_3^-$ [b] (5)	CO_2	44(obs)	-55.3	-94.9	20.5
$H_2NCH(CH_3)C(O)NH^-$ 	CO_2	44 (obs)	-63.9	-94.9	11.9
$^2[H_2NCH_2C(O)NH•CH_2]$ [b] + e^-	CO_2		-2.8	-94.9	73.0
$^2[H_2NC(O)CH(CH_3)•NH] + e^-$	CO_2		-7.4	-94.9	68.4
$HNC(O)C(=NH)CH_3^-$	HCOOH	46 (obs)	-50.3	-90.0	30.4
$H_2NCH_2CO_2^-$ [b]	c-H_2CNHC(O) 	57 (obs)	-117.2	10.6	64.1
$NHCH_2C(O)OH^-$	N≡CCH_2OH	57	-68.4	-11.9	90.4
$H_3CNHCO_2^-$	N≡CCH_2OH	57	-119.6	-11.9	41.6
$^2[•NHCH_2CO_2]^-$ [b] (3)	$^2[H_2NCH_2•C(O)]$	58	-73.8	6.6	103.5
$^2[•NHCH_2CO_2]^-$ [b] (3)	$^2[H_2NCHC•CH(O)]$	58	-73.8	-7.3	89.6
$^2[H_2C•NHCO_2]^-$	$^2[H_2N•CHCH(O)]$	58	-83.0	-7.3	80.4

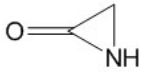
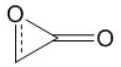
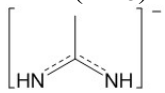
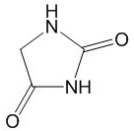
$H_2NCH_2C(O)NH^-$ [b] (2)	c-$H_2COC(O)$ 	58	-55.4	-41.4	73.9
$H_2NC(O)NCH_3^-$	c-$H_2COC(O)$ 	58	-57.1	-41.4	72.2
$HN{=}CHCHCH{=}NH^-$	$CO_2 + H_2O$	62 (obs)	23.0	-94.9 + -57.4	41.4
$HN{=}CHCHCH{=}NH^-$	HOC(O)OH	62(obs)	23.0	-145.3	48.4
$H_2NCH_2C(O)^-$ [b] (3)	$CO_2 + HN{=}CH_2$	73	-6.4	-94.9 + 21.3	90.7
$H_2NCH{=}CHO^-$	$CO_2 + HN{=}CH_2$	73	-33.3	-94.9 + 21.3	63.8
$H_3CC(O)NH^-$ 	$CO_2 + HN{=}CH_2$	73 (obs)	-59.6	-94.9 + 21.3	37.5
$^2[{\bullet}CH_2CO_2]^-$ [b] (4)	$^2[{\bullet}NHCH_2C(O)NH_2]$	73	-75.4	2.0	97.3
$^2[{\bullet}CH_2CO_2]^-$	$^2[H_2NC(O)NH{\bullet}CH_2]$	73	-75.4	-10.6	84.7
$HNCH_3^-$ [b] (2)	$C_3H_3NO_3$ 	101	32.6	-119.0	84.3
$HNCH_3^-$	$HOC(O)CH_2$-N=C=O	101	32.6	-120.3	83.0

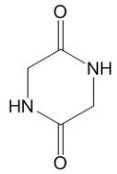
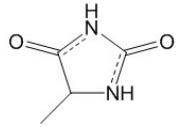
HO^-	$C_4H_6N_2O_2$ 	114	-33.0	-89.5	48.2
NH_2^- [b] (1)	$C_4H_5NO_3$ 	115	27.3	-111.0	87.0
NH_2^-	$C_4H_5NO_3$ 	115	27.3	-125.1	72.9

[a] $H_2NCH_2C(O)NHCH_2CO_2^-$. ΔH_f(298 K) = -170.7 kcal/mol. Mass = 131 Da. [b] 'Adiabatic' bond cleavage. Bond labels in parentheses. See Figure 2.2. [c] "obs" indicates that this fragmentation pathway is observed in our CID spectra.

Table 2.7. Fragmentation Energies in Diglycine Amide $[M - H]^-$ in kcal/mol[a]

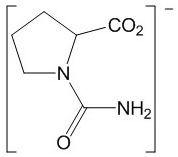
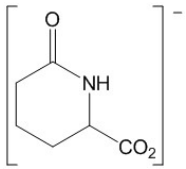
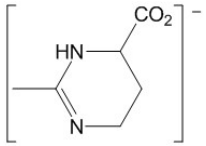
Anion	Neutral	Mass Loss [c]	ΔH_f A	ΔH_f B	ΔH_{rxn}
$C_4H_6N_3O_2^-$ 	H_2	2	-91.1	-1.1	16.4
$^2[H_2\bullet CC(O)NHCH_2CO_2]^-$ [b] (1)	$^2[\bullet NH_2]$	16	-70.9	44.5	82.2
$^2[H_2NC(O)CH\bullet CHC(O)NH_2]^-$	$^2[\bullet NH_2]$	16	-89.5	44.5	63.6
$H_2NCH(CN)CH_2CO_2^-$ 	NH_3	17	-101.4	-10.0	-2.8
$C_4H_6N_3O^-$ 	H_2O	18(obs)	-35.6	-57.4	15.6
$H_2NCH_2C(O)NHCHC{\equiv}N^-$	H_2O	18 (obs)	-19.6	-57.4	31.6
$(H_2N)CH_2C(O)N^-\text{-}C{=}C{=}N$	H_2O	18 (obs)	-6.3	-57.4	44.9
$^2[\bullet C(O)NHCH_2C(O)NH]^-$ [b] (2)	$^2[H_2NC\bullet H_2]$	30	-66.3	36.9	79.2
$^2[\bullet C_3H_4N_2O_2]^-$ 	$^2[H_2NC\bullet H_2]$	30	-64.8	36.9	80.7
$^2[H_2NC\bullet(O)CH(CH_3)NH_2]^-$	$^2[\bullet N{=}C{=}O]$	42	-25.9	29.4	112.1
$H_2NCH_2C(O){=}NCH_3^-$ [b] (5)	O=C=NH	43	-55.3	-28.3	25.0

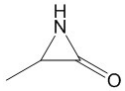
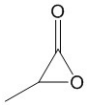
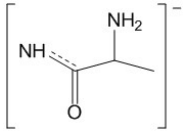
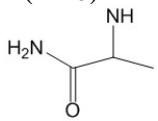
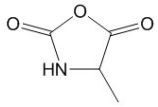
$H_3CNHCH_2C(NH)_2^-$	CO_2	44	8.1	-94.9	22.7
$^2[H_3CC(O)N•C(O)]^-$	$^2[H_2NC•HNH_2]$	45	-74.2	41.1	75.5
$NH_2C(O)NCH=CH_2^-$	$H_2NCH(O)$	45 (obs)	-44.8	-44.1	19.7
$H_2NCH_2C(O)NH^-$ [b] (4)	c-$H_2CNHC(O)$ 	57 (obs)	-55.4	10.6	63.8
$H_2NC(O)NCH_3^-$	$N≡CCH_2OH$	57(obs)	-57.1	-11.9	39.6
$^2[•NHCH_2C(O)NH]^-$ [b] (3)	$^2[H_2NCH_2•C(O)]$	58	-12.1	6.6	103.1
$^2[•NHCHC(O)NH_2]^-$	$^2[H_2N•CHCH(O)]$	58	-26.6	-7.3	74.7
$H_2NC(NH)NCH_3^-$	c-$H_2COC(O)$ 	58	17.3	-41.4	84.5
$^2[H_2NC(O)C•CHNH]^-$	$^2[H_2N•C(O)CH_2]$	58	-26.6	-8.6	73.4
$H_2NCH_2C(O)^-$ [b] (3)	$HN=CH_2 + O=C=NH$	72	-6.4	21.3 + -28.3	95.2
$H_2NCHCH(O)^-$ [b] (3)	$HN=CHC(O)NH_2$	72	-33.3	-23.6	51.7
$^2[•CH_2C(O)NH]^-$ [b] (4)	$^2[•NHCH_2C(O)NH_2]$	73	-15.3	2.0	95.3
$^2[H_2CN•CHO]^-$	$^2[H_2NC(O)NHC•H_2]$	73	-15.6	-10.6	82.4
$HNC(CH_3)NH^-$ 	$CO_2 + HN=CH_2$	73	10.0	-94.9 + 21.3	45.0
$HNCH_2NH_2^-$	$H_3CC(O)N=C=O$	85	27.4	-66.1	69.9
$N=C=O^-$	$H_2NC(O)NHCH_2CH_3$	88 (obs)	-53.5	-60.0	-4.9
$HNCH_3^-$ [b] (2)	$C_3H_4N_2O_2$ 	100	32.6	-80.4	60.8

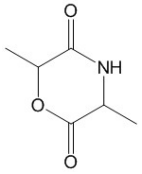
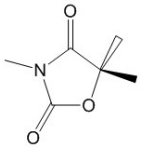
HO^-	$H_2NC(O)NHCH_2CH_2C{\equiv}N$	113	-33.0	-27.1	48.5
NH_2^- [b] (1)	$C_4H_6N_2O_2$ 	114	27.3	-74.2	61.7
NH_2^-	$C_4H_6N_2O_2$ 	114	27.3	-89.5	46.4

[a] $H_2NCH_2C(O)NHCH_2C(O)NH^-$. ΔH_f(298 K) = -108.6 kcal/mol. Mass = 130 Da. [b] 'Adiabatic' bond cleavage. Bond labels in parentheses. See Figure 2.2. [c] "obs" indicates that this fragmentation pathway is observed in our CID spectra.

Table 2.8. Product Production Energies in Dialanine $[M - H]^-$ in kcal/mol[a]

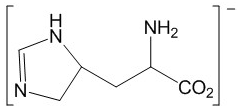
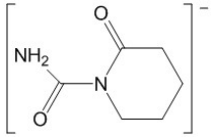
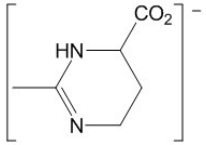
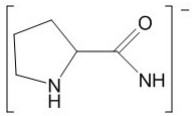
Anion	Neutral	Mass Loss [c]	ΔH_f A	ΔH_f B	ΔH_{rxn}
$C_6H_9N_2O_3^-$ 	H_2	2	-180.3	-1.1	6.0
$^2[H_2N•CHC(O)NHCH(CH_3)CO_2]^-$ [b]	$^2[•CH_3]$	15	-145.3	34.2	76.3
$^2[H_3C•CHC(O)NHCH(CH_3)CO_2]^-$ [b] (1)	$^2[•NH_2]$	16	-149.1	44.5	82.8
$^2[H_3CCH_2NHC(O)CH•CHC(O)OH]^-$	$^2[•NH_2]$	16	-147.7	44.5	84.2
$C_6H_8NO_3^-$ 	NH_3	17	-178.0	-10.0	-0.6
$C_6H_9N_2O_2^-$ 	H_2O	18	-116.3	-57.4	13.7
$^2[•C(O)NHCH(CH_3)CO_2]^-$ [b] (2)	$^2[H_2NC•HCH_3]$	44	-135.9	29.2	80.7
$^2[H_2NC(O)CH•CHC(O)OH]^-$	$^2[H_2NC•HCH_3]$	44	-142.5	29.2	74.1
$H_2NCH(CH_3)C(O)NCH_2CH_3^-$ [b] (5)	CO_2	44 (obs)	-72.0	-94.9	20.5
$CH_3CH_2NHC(O)NHCH_2CH_2^-$	CO_2	44 (obs)	-71.7	-94.9	20.8
$HN=C(CH_3)CHC(O)CH_3^-$	$CO_2 + NH_3$	61	-64.4	-94.9 + 44.5	72.6
$HN=C(CH_3)CHC(O)CH_3^-$	$H_2NC(O)OH$	61 (obs)	-64.4	-102.1	20.9

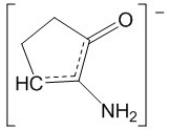
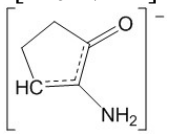
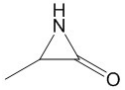
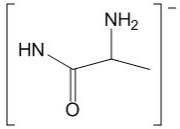
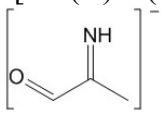
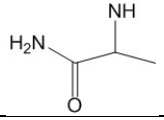
$H_2NCH(CH_3)CO_2^{-b}$	c-$C(CH_3)HNHC(O)$ 	71 (obs)	-125.5	1.5	63.4
$H_2NCH(CH_3)CO_2^-$	$H_3CCH{=}NCH(O)$	71	-125.5	-16.4	45.5
$H_3CCH_2NHCO_2^-$	$H_3CCH{=}NCH(O)$	71	-127.6	-16.4	43.4
$^2[H_2N{\bullet}C(CH_3)CO_2]^{-\,b}$ (2)	$^2[H_2N{\bullet}CH(CH_3)C(O)]$	72	-90.4	0.5	97.9
$^2[H_2NC(O){\bullet}C(O)CH_3]^-$	$^2[CH_3CH_2\text{-}NH{\bullet}C(O)]$	72	-94.0	-8.8	84.6
$H_2NCH(CH_3)C(O)NH^{-\,b}$ (4)	c-$CH(CH_3)OC(O)$ 	72	-64.3	-52.2	70.9
$HNC(O)CH(CH_3)NH_2^-$ 	$H_2C{=}CHC(O)OH$	72 (obs)	-63.9	-71.9	51.6
$H_2NCH(CH_3)C(O)^{-\,b}$ (3)	$CH_3CH{=}NH + CO_2$	87	-15.2	10.4 + -94.9	87.7
$H_2NC(CH_3)CH(O)^{-\,b}$ (3)	$CH_3CH{=}NH + CO_2$	87	-41.2	10.4 + -94.9	61.7
$^2[CH_3{\bullet}CHC(O)_2]^{-\,b}$ (4)	$^2[{\bullet}NHCH(CH_3)C(O)NH_2]$	87	-83.9	-5.3	98.2
$^2[CH_3C(O){\bullet}CH(O)]^-$	$^2[H_2NC(O)CH(CH_3){\bullet}NH$ 	87	-85.8	-7.4	94.2
$H_3CCH_2C(O)NH^-$	$H_2NC(O)C(O)CH_3$	87 (obs)	-61.2	-84.3	41.9
$HNCH_2CH_3^-$	$C_4H_5NO_3$ 	115	20.8	-128.2	80.0

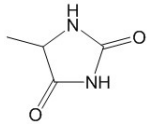
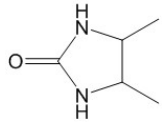
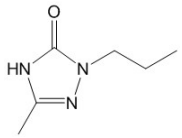
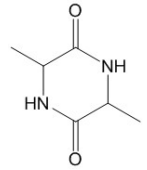
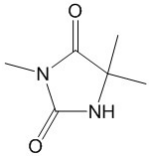
NH_2^- [b] (1)	$C_6H_9NO_3$ 	143	27.3	-128.7	86.0
NH_2^-	$C_6H_9NO_3$ 	143	27.3	-147.1	67.6

[a] $H_2NCH(CH_3)C(O)NHCH(CH_3)CO_2^-$. ΔH_f(298 K) = -187.4 kcal/mol. Mass=159 Da. [b] 'Adiabatic' bond cleavage. Bond labels in parentheses. See Figure 2.2. [c] "obs" indicates that this fragmentation pathway is observed in our CID spectra.

Table 2.9. Fragmentation Energies in Dialanine Amide $[M - H]^-$ in kcal/mol[a]

Anion	Neutral	Mass Loss [c]	ΔH_f A	ΔH_f B	ΔH_{rxn}
$C_6H_{10}N_3O_2^-$ 	H_2	2	-100.8	-1.1	23.8
$^2[H_2N{\bullet}CHC(O)NHCH(CH_3)C(O)NH]^-$ [b]	$^2[{\bullet}CH_3]$	15	-83.9	34.2	76.0
$^2[H_3C{\bullet}CHC(O)NHCH(CH_3)C(O)NH]^-$ [b1]	$^2[{\bullet}NH_2]$	16	-87.3	44.5	82.9
$^2[{\bullet}C_6H_{10}N_2O_2]^-$ 	$^2[{\bullet}NH_2]$	16	-90.4	44.5	79.8
$C_6H_9N_2O_2^-$ 	NH_3	17	-116.3	-10.0	-0.6
$H_2NC(CH_3)C(O)C^-(CH_3)C{\equiv}N$	H_2O	18 (obs)	-34.1	-57.4	34.2
$(H_2N)CH(CH_3)C(O)N^-C(CH_3){=}C{=}N$	H_2O	18 (obs)	-18.9	-57.4	49.4
$H_2NCH(CH_3)C(O)NCH_2CH_3^-$ [b] (5)	O=C=NH	43	-72.0	-28.3	25.4
$CH_3CH_2NHC(O)NHCH_2CH_2^-$	O=C=NH	43	-71.7	-28.3	25.7
$HNC(NH_2)NC(CH_3)_3^-$	CO_2	44	-16.9	-94.9	13.9
$^2[{\bullet}C(O)NHCH(CH_3)C(O)NH]^-$ [b] (2)	$^2[H_2NC{\bullet}HCH_3]$	44	-74.9	29.2	80.0
$^2[H_2NC(O)CH{\bullet}CHC(O)NH_2]^-$	$^2[H_2NC{\bullet}HCH_3]$	44	-89.5	29.2	65.4
$C_5H_9N_2O^-$ 	$H_2NCH(O)$	45 (obs)	-55.2	-44.1	26.4

$^2[\bullet C_5H_7NO]^-$ 	$H_2NCH(O) + {}^2[\bullet NH_2]$	61	-17.5	-44.1 + 44.5	108.6
$^2[\bullet C_5H_7NO]^-$ 	$^2[(H_2N)_2\bullet COH]$	61	-17.5	-5.6	102.6
$H_2C{=}CHC(O)CH{=}CH(O)^-$	$HC(NH_2)_3$	61 (obs)	-80.2	0.0	45.5
$H_2NCH(CH_3)C(O)NH^-$ [b] (4)	c-$C(CH_3)HNHC(O)$ 	71(obs)	-64.3	1.5	62.9
$HN^-C(O)CH(CH_3)NH_2$ 	$H_3CCH{=}NCH(O)$	71(obs)	-63.9	-16.4	45.4
$^2[\bullet NHCH(CH_3)C(O)NH]^-$ [b] (3)	$^2[H_2NCH(CH_3)\bullet C(O)]$	72	-19.7	0.5	106.5
$^2[H_2N\bullet C(CH_3)C(O)NH]^-$ [b] (3)	$^2[H_2N\bullet C(CH_3)CH(O)]$	72	-36.6	-17.3	71.8
$^2[H_2N\bullet C(CH_3)C(O)NH]^-$	$^2[CH_3CH_2NH\bullet C(O)]$	72	-36.6	-8.8	80.3
$H_2NCH(CH_3)C(O)^-$ [b] (3)	$CH_3CH{=}NH + O{=}C{=}NH$	86	-15.2	10.4 + -28.3	92.6
$H_2NC(CH_3)CH(O)^-$ [b] (3)	$HN{=}C(CH_3)C(O)NH_2$	86	-41.2	-35.7	48.8
$H_3CCH_2C(O)NH^-$	$HN{=}C(CH_3)C(O)NH_2$	86	-61.2	-35.7	28.8
$^2[H_3C\bullet CHC(O)NH]^-$ [b] (4)	$^2[\bullet NHCH(CH_3)C(O)NH_2]$	87	-23.7	-5.3	96.7
$^2[HC(O)\bullet C(=NH)CH_3]^-$ 	$^2[H_2NC(O)CH(CH_3)\bullet NH]$ 	87	-30.0	-7.4	88.3

$HNCH_2CH_3^-$ [b] (2)	$C_4H_6N_2O_2$ 	114(obs)	20.8	-89.5	57.0
$HNCH(O)^-$	$C_5H_{10}N_2O$ 	114 (obs)	-50.1	-58.2	17.4
$N=C=O^-$	$H_2NC(O)NHCH_2CH(CH_3)_2$	116 (obs)	-53.5	-71.9	0.3
HO^-	$C_6H_{11}N_3O$ 	141	-33.0	-31.9	60.8
NH_2^- [b] (1)	$C_6H_{10}N_2O_2$ 	142	27.3	-91.6	61.4
NH_2^-	$C_6H_{10}N_2O_2$ 	142	27.3	-103.7	49.3

[a] $H_2NCH(CH_3)C(O)NHCH(CH_3)\ C(O)NH^-$. ΔH_f(298 K) = -125.7 kcal/mol. Mass = 158 Da. [b] 'Adiabatic' bond cleavage. Bond labels in parentheses. See Figure 2.2. [c] "obs" indicates that this fragmentation pathway is observed in our CID spectra.

Figures

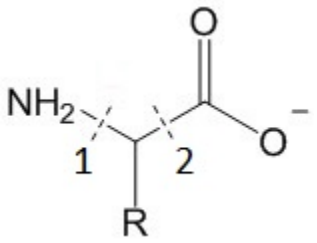


Figure 2.1. Bond dissociation site numbering for amino acids.

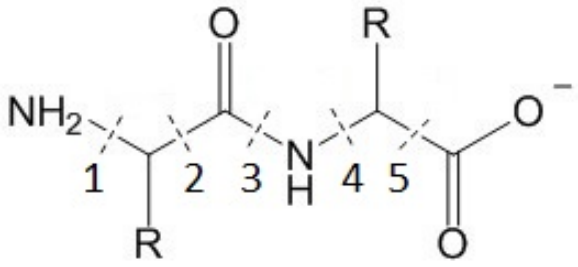


Figure 2.2. Bond dissociation site numbering for dipeptides.

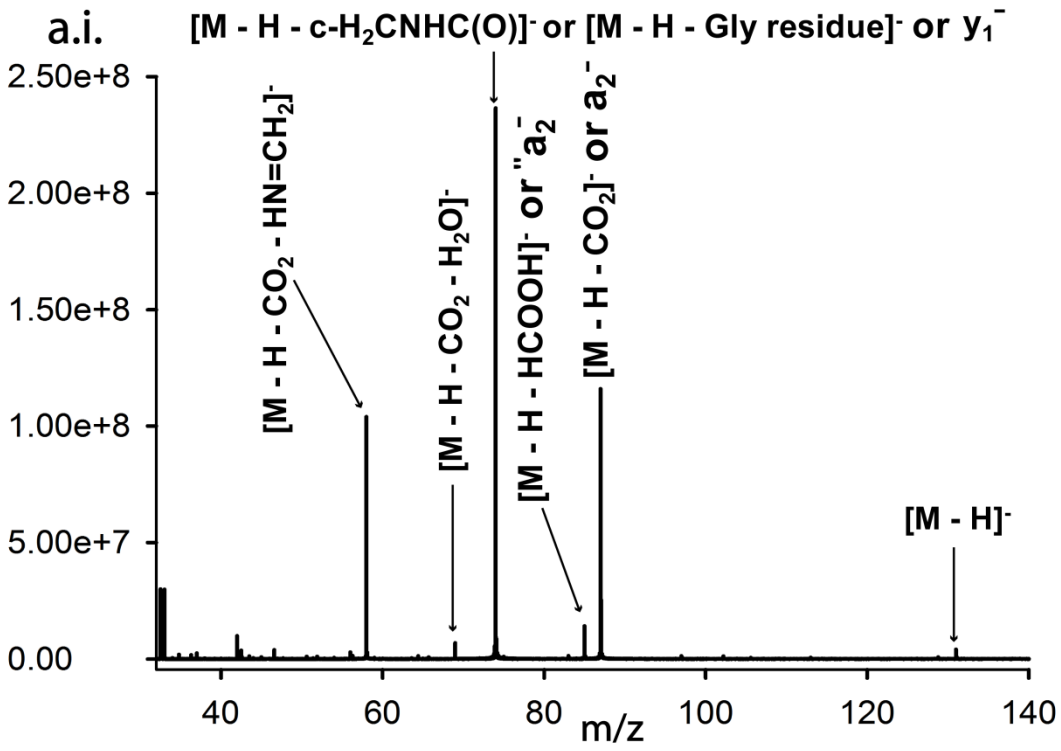


Figure 2.3. Low-energy CID spectrum of $[M - H]^-$ from diglycine.

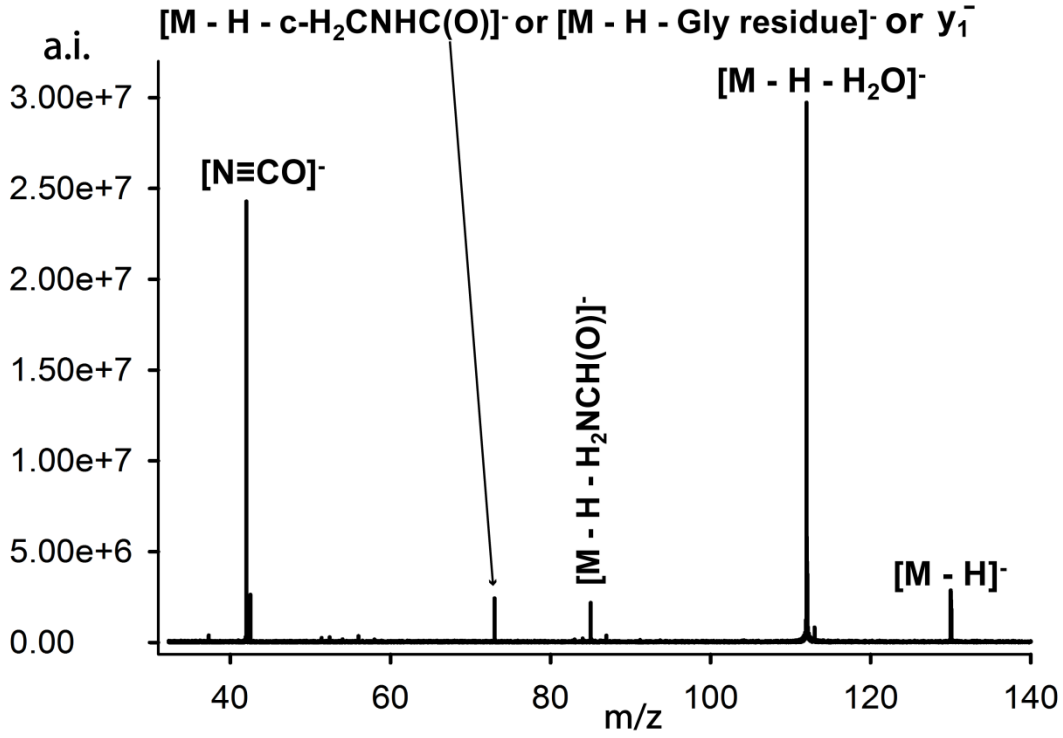


Figure 2.4. Low-energy CID spectrum of $[M - H]^-$ from diglycine amide.

References

[1] Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Anal. Chem.* **2016**, *88*, 30 – 50.

[2] McLuckey, S. A. Principles of Collisional Activation in Analytical Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599 – 614.

[3] Eckersley, M.; Bowie, J. H.; Hayes, R. N. Collision-induced Dissociations of Deprotonated α-Amino Acids. The Occurrence of Specific Proton Transfers Preceding Fragmentation. *Int. J. Mass Spectrom. Ion Proc.* **1989**, *93*, 199 – 213.

[4] Waugh, R. J.; Bowie, J. H. A Review of the Collision Induced Dissociations of Deprotonated Dipeptides and Tripeptides. An Aid to Structure Determination. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 169 – 173.

[5] Bowie, J. H.; Brinkworth, C. S.; Dua, S. Collision-Induced Fragmentations of the $[M - H]^-$ Parent Anions of Underivatized Peptides: An Aid to Structure Determination and Some Unusual Negative Ion Cleavages. *Mass Spectrom. Rev.* **2002**, *21*, 87 – 107.

[6] Bilusich, D.; Bowie, J. H. Fragmentation of $(M - H)^-$ Anions of Underivatised Peptides. Part 2. Characteristic Cleavages of Ser and Cys, and Disulfides and Other Post-Translational Modifications, Together with Some Unusual Internal Processes. *Mass Spectrom. Rev.* **2009**, *28*, 20 – 34.

[7] Wang, T.; Tran, Nha T. T.; Andreazza, H. J.; Bilusih, D.; Brinkworth, C. S.; Bowie, J. H. Negative Ion Cleavages of $(M - H)^-$ Anions of Peptides. Part 3. Post-Translational Modifications. *Mass Spectrom. Rev.* **2016**, E-issn: 1098 – 2787.

[8] Li, Z.; Yalcin, T.; Cassady, C. J. C-Terminal Amino Acid Residue Loss for Deprotonated Peptide Ions Containing Glutamic Acid, Aspartic Acid, or Serine Residues at the C-Terminus. *J. Mass Spectrom.* **2006**, *41*, 939 – 949.

[9] Clipston, N. L.; Jai-nhuknan, J.; Cassady, C. J. A Comparison of Negative and Positive Ion Time-of-Flight Post-Source Decay Mass Spectrometry for Peptides Containing Basic Residues. *Int. J. Mass Spectrom.* **2003**, *222*, 363 – 381.

[10] Harrison, A.G. Sequence-specific Fragmentation of Deprotonated Peptides Containing H or Alkyl Side Chains. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1 – 13.

[11] Bokatzian-Johnson, S. S.; Stover, M. L.; Dixon, D. A.; Cassady, C. J. Comparison of the Effects of Amide and Acid Groups at the C-Terminus on the Collision-Induced Dissociation of Deprotonated Peptides. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1544 – 1557.

[12] Nha Tran, T. T.; Wang T.;Hack, S.; Bowie, J. H. Fragmentations of $[M–H]^-$ Anions of Peptides containing Ser Sulfate. A Joint Experimental and Theoretical Study. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 2287 – 2296.

[13] Wang, T.; Nha Tran, T.T.; Calabrese, A. N.; Bowie, J. H. Backbone Fragmentations of [M – H]$^-$ Anions from Peptides. Reinvestigation of the Mechanism of the Beta Prime Cleavage. *Rapid Commun. Mass Spectrom*. **2012**, *26*, 1832 – 1840.

[14] Bokatzian, S. S.; Stover, M. L.; Plummer, C. E.; Dixon, D. A.; Cassady, C. J. An Experimental and Computational Investigation into the Gas-Phase Acidities of Tyrosine and Phenylalanine: Three Structures for Deprotonated Tyrosine. *J. Phys. Chem. B* **2014**, *118*, 12630 – 12643.

[15] Li, Z.; Matus, M. H.; Velazquez, H. A.; Dixon, D. A.; Cassady, C. J. Gas-Phase Acidities of Aspartic Acid, Glutamic Acid, and their Amino Acid Amides. *Int. J. Mass Spectrom.* **2007**, *265*, 213 – 223.

[16] Stover, M. L.; Jackson, V.; Matus, M.; Adams, M.; Cassady, C. J.; Dixon, D. A. Fundamental Thermochemical Properties of Amino Acids: Gas-Phase and Aqueous Acidities and Gas-Phase Heats of Formation. *J. Phys. Chem. B* **2012**, *116*, 2905 – 2916.

[17]. Stover, M. L.; Plummer, C. E.; Miller, S. R.; Cassady, C. J.; Dixon, D. A. Gas-Phase Acidities of Phosphorylated Amino Acids. *J. Phys. Chem. B* **2015**, *119*, 14604 − 14621.

[18] Plummer, C. E.; Stover, M. L.; Bokatzian, S. S.; Davis, J.; Dixon, D. A.; Cassady, C. J. An Experimental and Computational Study of the Gas-Phase Acidities of the Common Amino Acid Amides. *J. Phys. Chem. B* **2015**, 119, 9661 − 9669.

[19] Grzetic, J.; Oomens, J. Spectroscopic Evidence for an Oxazolone Structure in Anionic b-Type Peptide Fragments. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 290 – 300.

[20] Wysocki, V.; Tsaprailis, G.; Smith, L. L.: Breci, L. A. Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *J. Mass Spectrom.* **2000**, *35*, 1399 – 1406.

[21] Wright, P.; Alex, A.; Harvey, S.; Parsons, T.; Pullen, F. Understanding Collision-induced Dissociation of Dofetilide: A Case Study in the Application of Dennsity Functional Theory as an Aid to Mass Spectral Interpretation. *Analyst* **2013**, *138*, 6869 – 6880.

[22] Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V.; Pople, J. A. Gaussian-3 Theory Using Reduced Møller-Plesset Order. *J. Chem. Phys.* **1999**, *110*, 4703 – 4709.

[23] Dixon, D. A.; Feller, D.; Peterson, K. A. In *A Practical Guide to Reliable First Principles Computational Thermochemistry Predictions Across the Periodic Table;* Wheeler, R. A., Section Ed. Tschumper, G. S., Eds.; Annual Reports in Computational Chemistry; Elsevier: Amsterdam, 2012; Vol. 8, pp 1 – 28.

[24] Feller, D.; Peterson, K. A.; Dixon, D. A. Further Benchmarks of a Composite, Convergent, Statistically Calibrated Coupled-Cluster-Based Approach for Thermochemical and Spectroscopic Studies. *Mol. Phys.* **2012**, *110*, 2381 – 2399.

[25] Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy: Current Strategies and Future Challenges. *Theor. Chem. Acc.* **2012**, *131*, 1 – 20.

[26] Feller, D.; Peterson, K. A.; Dixon, D. A. A Survey of Factors Contributing to Accurate Theoretical Predictions of Atomization Energies and Molecular Structures. *J. Chem. Phys.* **2008**, *129*, 204105-1 – 204105-32.

[27] Feller, D.; Dixon, D. A. Extended Benchmark Studies of Coupled Cluster Theory through Triple Excitations. *J. Chem. Phys.* **2001**, *115*, 3484 – 3496.

[28] Heck, A. J. R.; de Koning, L. J.; Pinkse, F. A.; Nibbering, N. M. Mass-specific Selection of Ions in Fourier-transform Ion Cyclotron Resonance Mass Spectrometry: Unintentional Off-resonance Cyclotron Excitation of Selected Ions. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 406 – 414.

[29] Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. Sustained Off-resonance Irradiation for Collision-actived Dissociation Involving Fourier Transform Mas Spectrometry. Collision-activated Dissociation Technique that Emulates Infrared Multiphoton Dissociation. *Anal. Chim. Acta* **1991**, *246*, 211 – 225.

[30] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.

[31] Lee, C.; Yang, W.; Parr, R. G. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Phys. Rev. B* **1988**, *37*, 785 – 789.

[32] Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. Optimization of Gaussian-Type Basis-Sets for Local Spin-Density Functionals Calculations. 1. Boron Through Neon, Optimization Technique and Validation. *Can. J. Chem.* **1992**, *70*, 560 – 571.

[33] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision B.01; Gaussian, Inc.: Wallingford, CT, 2009.

[34] Sander, S. P.; Friedl, R. R.; Ravishankara, A. R.; Golden, D. M.; Kolb, C. E.; Kurylo, M. J.; Huie, R. E.; Orkin, V. L.; Molina, M. J.; Moortgat, G. K.; Finlayson-Pitts, B. J. Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies: Evaluation Number 14, JPL Publication 02-25, National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 2003. http://jpldataeval.jpl.nasa.gov/pdf/JPL_02-25_rev02.pdf

[35] Nguyen, M. T.; Nguyen, V. S.; Matus, M. H.; Gopakumar, G.; Dixon, D. A. Molecular Mechanism for H_2 Release from Amine Borane, BH_3NH_3, Including the Catalytic Role of the Lewis Acid BH_3. *J. Phys. Chem. A*, **2007**, *111*, 679 – 690.

[36] Hammond, G. S. A Correlation of Reaction Rates. *J. Am. Chem. Soc.* **1955**, *77*, 334 – 338.

[37] Bell, R. P. The Theory of Reactions Involving Proton Transfers. *Proc. R. Soc. London, Ser. A*, **1936**, *154*, 414 – 429.

[38] Evans, M. G.; Polanyi, M., Further Considerations on the Thermodynamics of Chemical Equilibria and Reaction Rates. *J. Chem. Soc., Faraday Trans.*, **1936**, *32*, 1333 – 1360.

[39] Holbrook, K. A.; Pilling, M. J.; Robertson, S. H. *Unimolecular Reactions,* 2nd Ed., Wiley: Chichester, UK, 1996.

[40] Gilbert R. G.; Smith, S. C. *Theory of Unimolecular and Recombination Reactions*, Blackwell Scientific: Oxford, 1990.

Appendix: Bond Dissociation Energies in Amino Acid Monomer and Dimer Anions for Use in Analyzing Collision-Induced Dissociation Processes

Full Reference for Reference 33 (Gaussian 09): Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone,V.; Petersson,G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J. ; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe,D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT, 2009.

Table A2.1. Additional Transition State and Barrier Energy Figures at the B3LYP/DZVP2 Level in kcal/mol

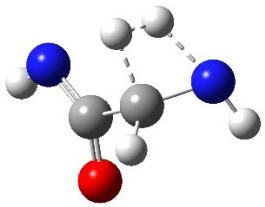
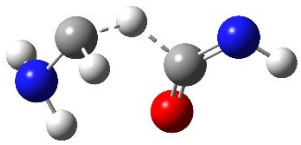
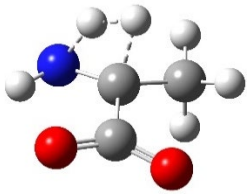
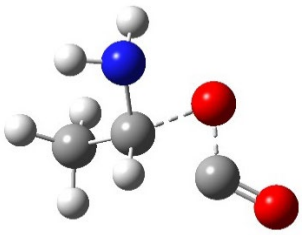
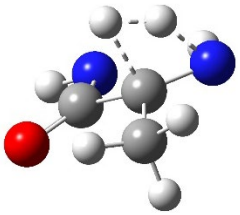
Products	**ΔH_{298}**	**ΔG_{298}**	**Image of TS**
Glycine Amide			
H_2 + HN=CHC(O)NH^-	107.9	107.6	
H_2NC:H + $HCONH^-$	60.8	60.4	
Alanine			
H_2 + HN=$C(CH_3)CO_2^-$	89.4	89.7	
CO + $H_2NCH(O)CH_3^-$	103.9	103.7	
Alanine Amide			
H_2 + HN=$C(CH_3)C(O)NH^-$	85.0	85.3	

Table A2.2. Comparison of Computational and Experimental Heats of Formation of Neutral Fragments in kcal/mol

Neutral	ΔH_f calc	ΔH_f expt
H_2	-1.1	0.0
$\bullet CH_3$	34.2	35.1
NH	84.0	85.3
$\bullet NH_2$	44.5	44.5
NH_3	-10.0	-11.0
HO•	8.3	8.9
H_2O	-57.4	-57.8
HCN	31.2	, 31.5
$H_2C{=}C\bullet H$	70.3	70.6
$H_2C{=}CH_2$	11.9	12.5
CO	-27.4	-26.4
$HN{=}CH_2$	21.3	21.7
HCO•	9.5	10.6
$H_3CC\bullet H_2$	28.9	28.9
CH_3NH_2	-4.1	-5.6±0.24
$\bullet CH_2NH_2$	36.9,	35.6
$H_2C(O)$	-27.7,	-26.0
HN=C=O	-28.3	-24.8±3
CO_2	-94.9	-94.0
HCOOH	-90.0	-90.5
$H_3CCH_2NH_2$	-11.0	-11.4±.14

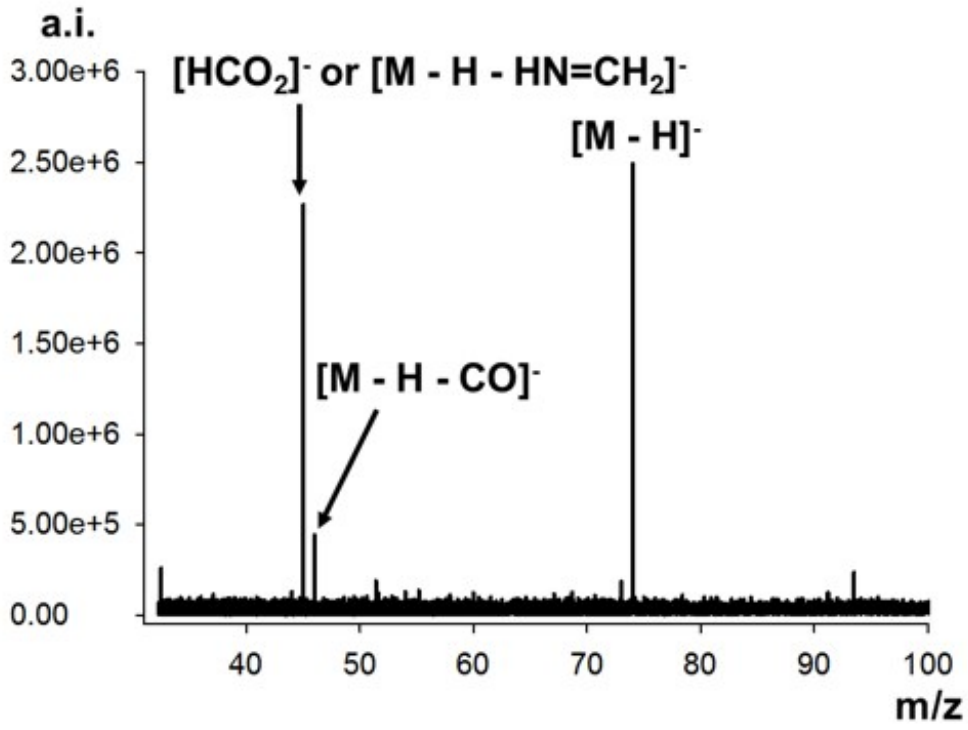


Figure A2.1. Low-energy CID spectrum of $[M - H]^-$ from glycine (G).

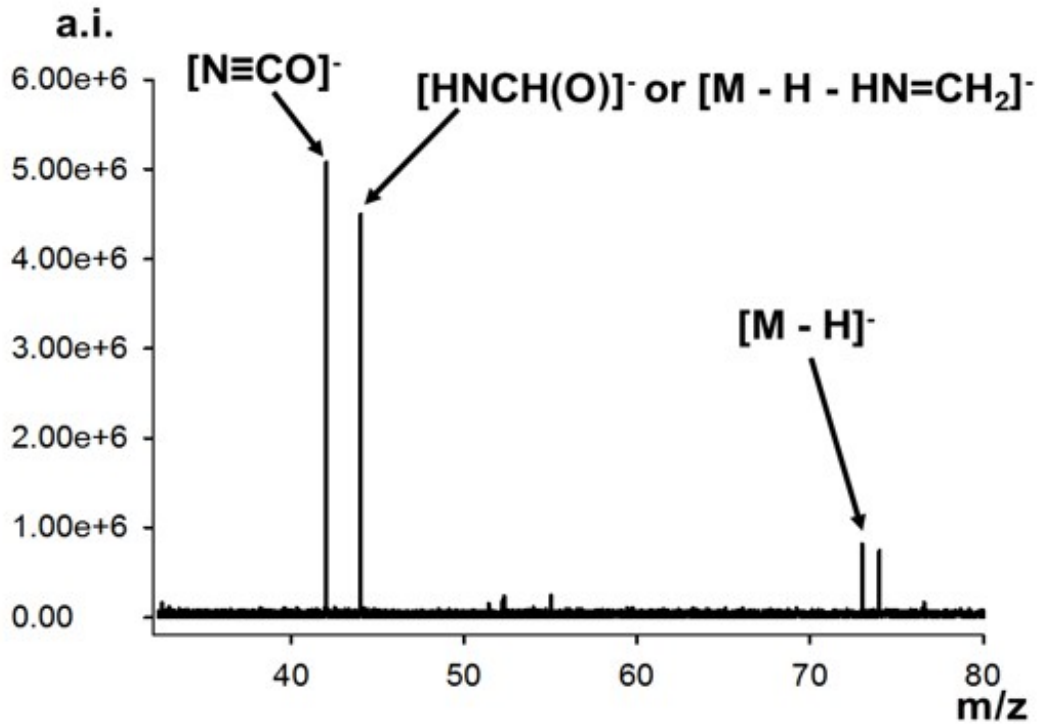


Figure A2.2. Low-energy CID spectrum of $[M - H]^-$ from glycine amide (G-NH_2). Due to its low acidity, glycine amide did not deprotonate readily by ESI. Multiple attempts over a period of months were made to reproduce this spectra but we only achieved sufficiently intense $[M - H]^-$ signal to study by CID once, for the spectra shown above. This spectrum, however, is consistent with our reproducible spectrum for alanine amide (Figure S4). All other CID spectra discussed in this manuscript were readily reproduced.

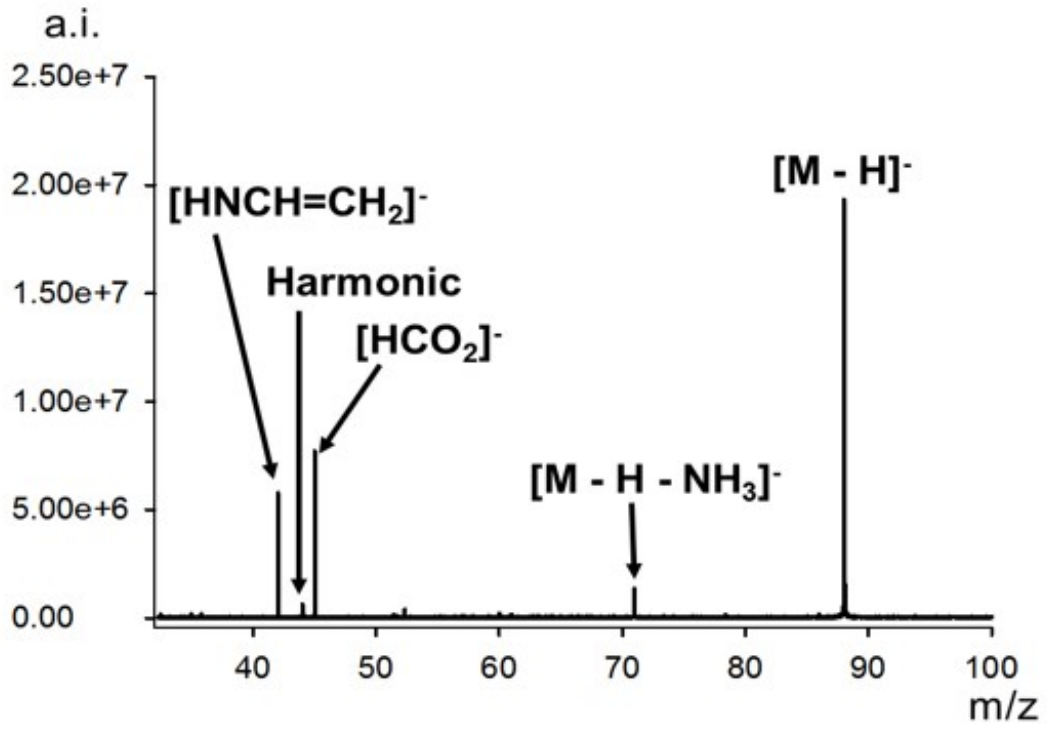


Figure A2.3. Low-energy CID spectrum of $[M - H]^-$ from alanine (A). "Harmonic" refers to a frequency harmonic of the precursor ion; this is not a CID product ion.

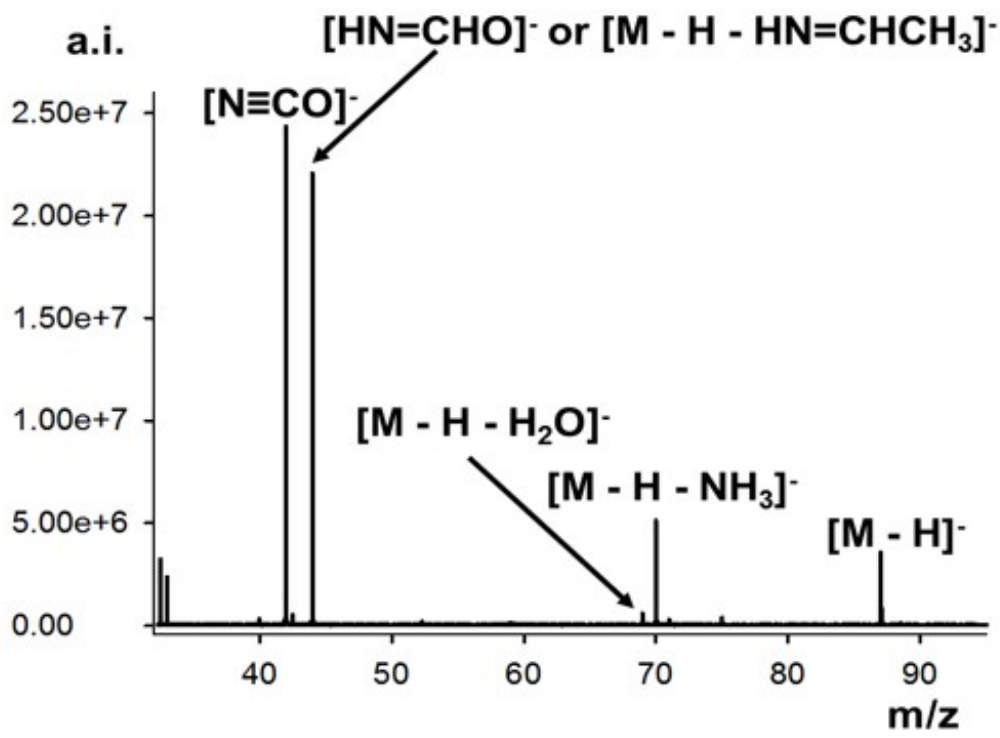


Figure A2.4. Low-energy CID spectrum of $[M - H]^-$ from alanine amide (A-NH_2).

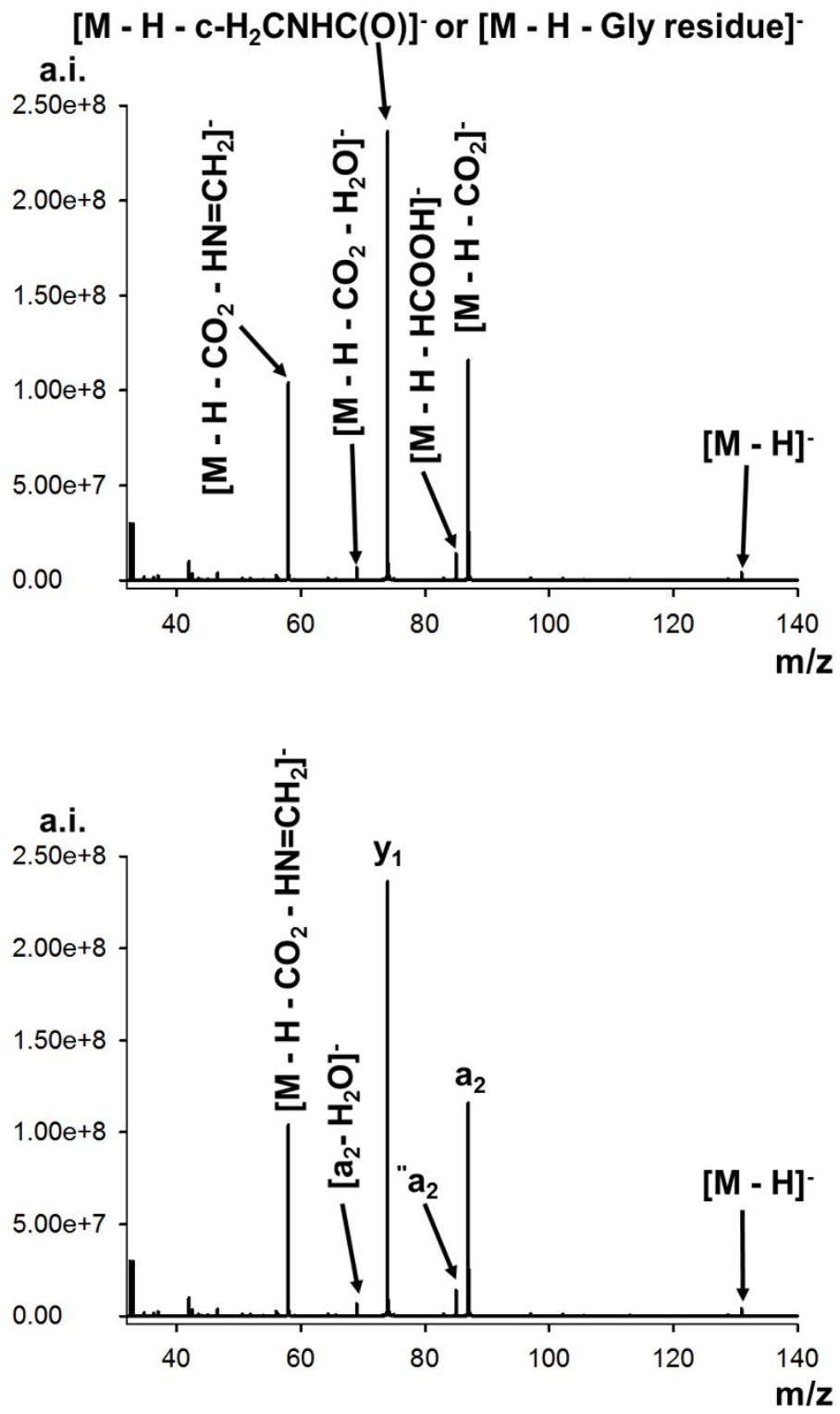


Figure A2.5. Low-energy CID spectrum of $[M - H]^-$ from diglycine (GG). The same spectrum is shown labelled with organic nomenclature (top) and peptide fragmentation nomenclature (bottom).

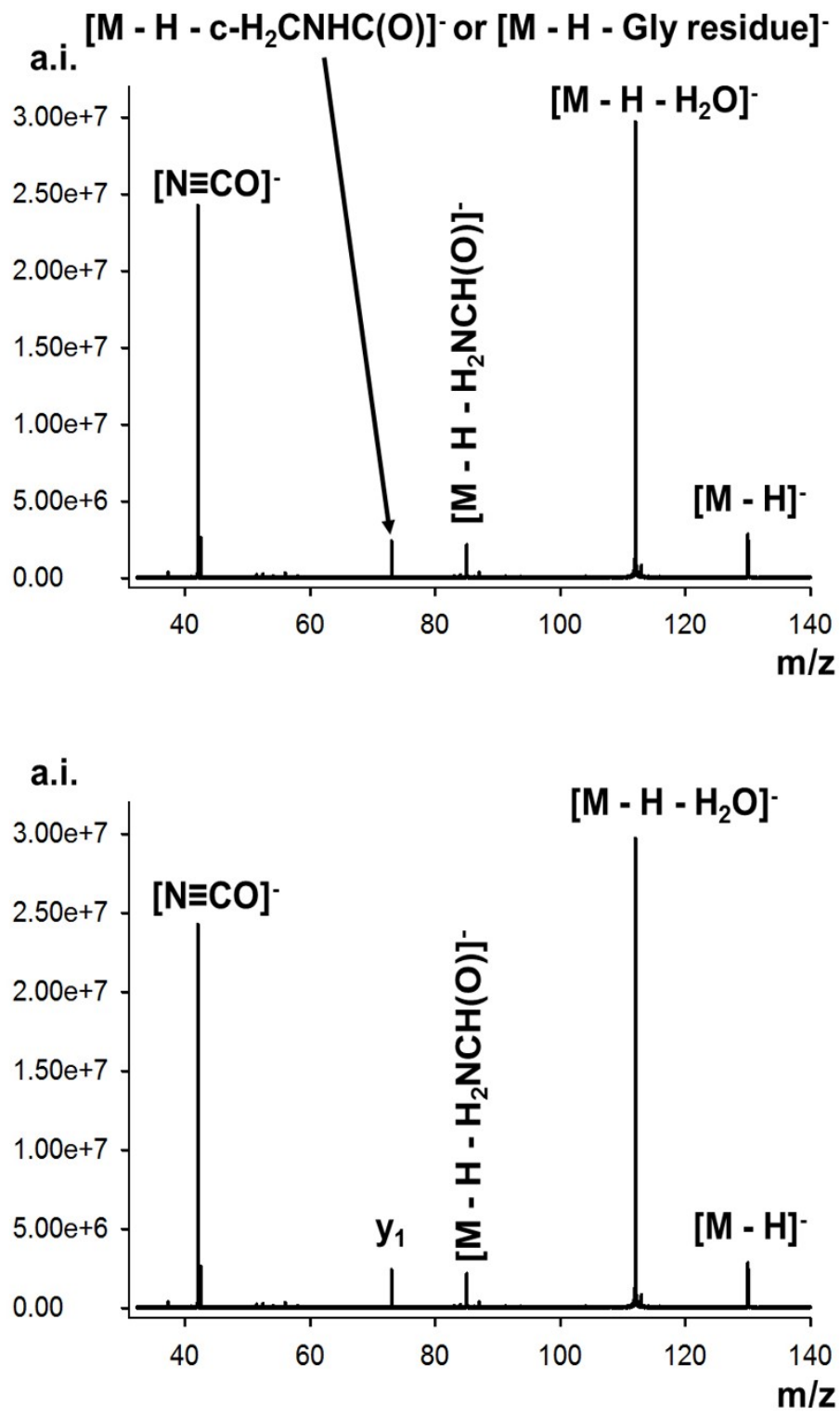


Figure A2.6. Low-energy CID spectrum of $[M - H]^-$ from diglycine amide ($GG\text{-}NH_2$). The same spectrum is shown labelled with organic nomenclature (top) and peptide fragmentation nomenclature (bottom).

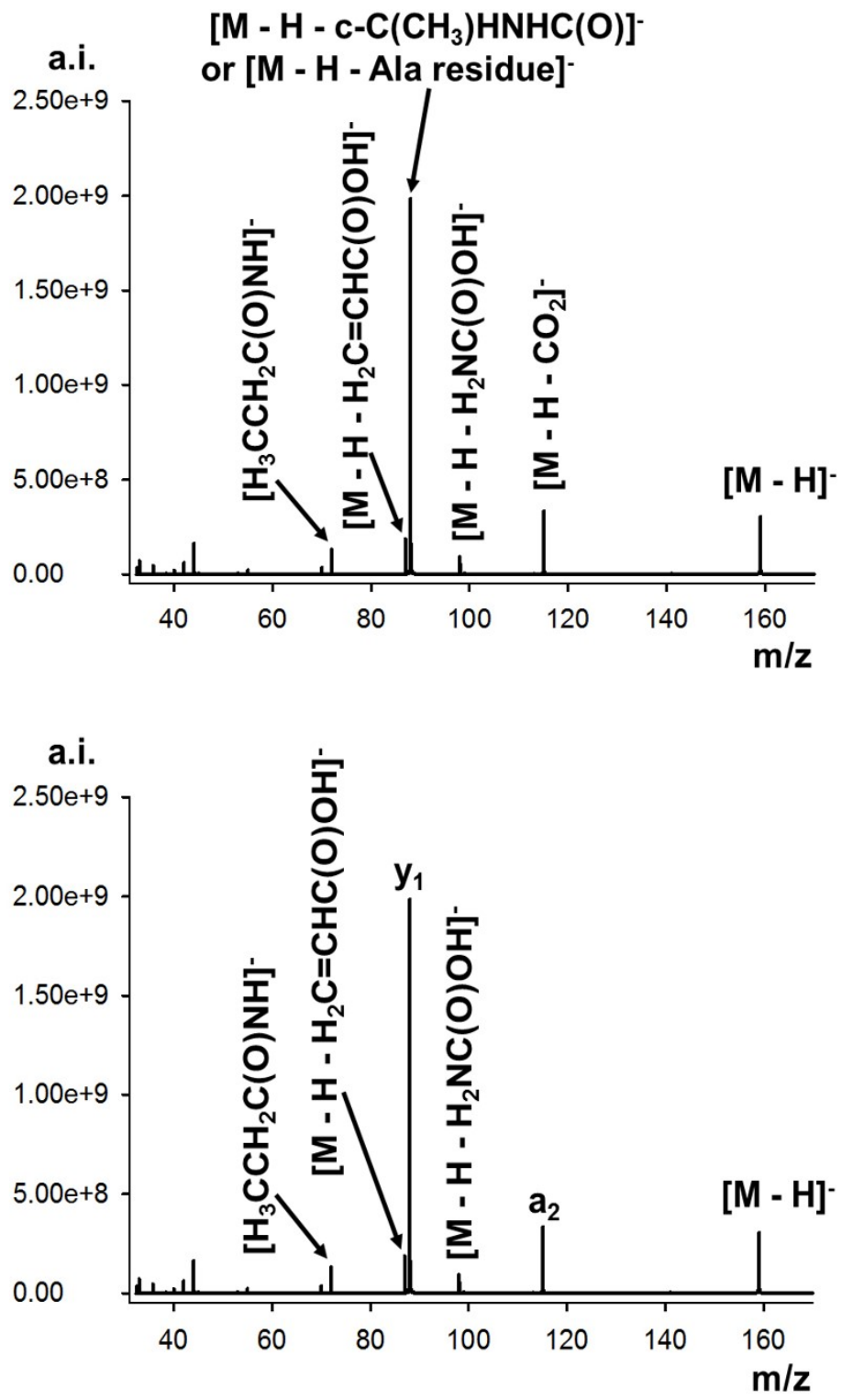


Figure A2.7. Low-energy CID spectrum of $[M - H]^-$ from dialanine (AA). The same spectrum is shown labelled with organic nomenclature (top) and peptide fragmentation nomenclature (bottom).

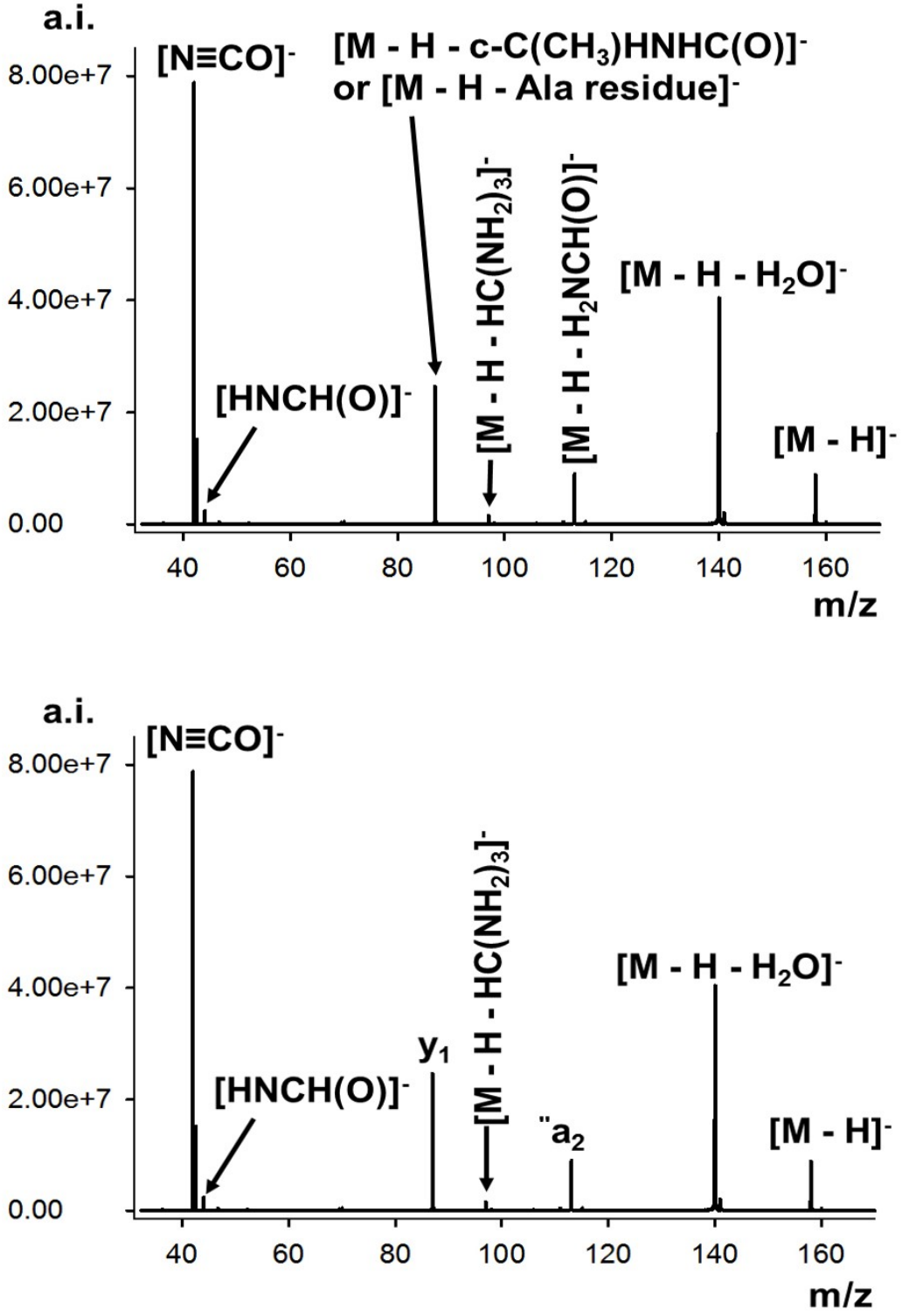


Figure A2.8. Low-energy CID spectrum of $[M - H]^-$ from dialanine amide ($AA\text{-}NH_2$). The same spectrum is shown labelled with organic nomenclature (top) and peptide fragmentation nomenclature (bottom).

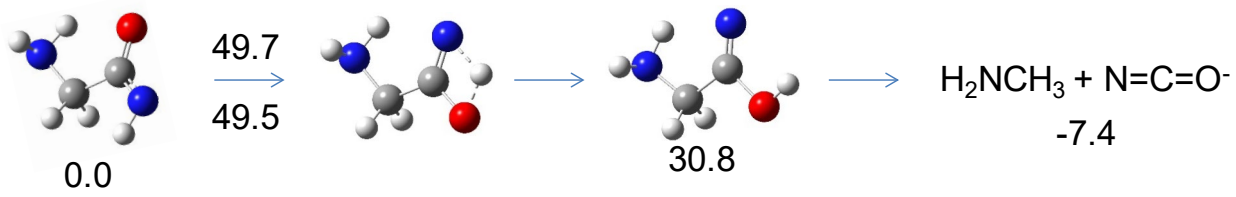


Figure A2.9. Relative energies (kcal/mol) for the dissociation of glycine amide to H_2NCH_3 and N=C=O$^-$. Energies shown are relative enthalpies (and Gibbs free energy, when shown) calculated at the B3LYP/DZVP2 level.

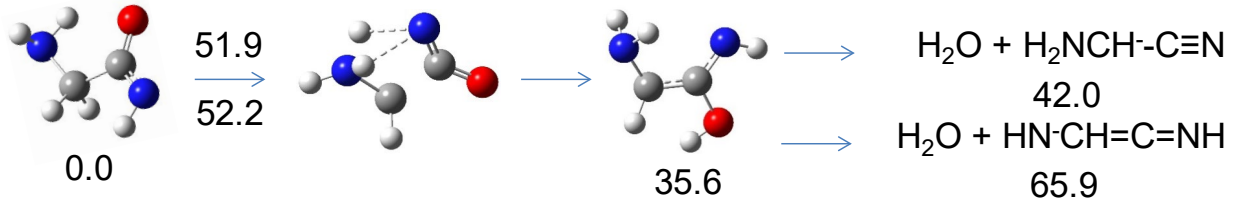


Figure A2.10. Relative energies (kcal/mol) for the dissociation of glycine amide to form H_2O and H_2NCH^--C≡N or HN^-CH=C=NH. Energies shown are relative enthalpies (and Gibbs free energy, when shown) calculated at the B3LYP/DZVP2 level.

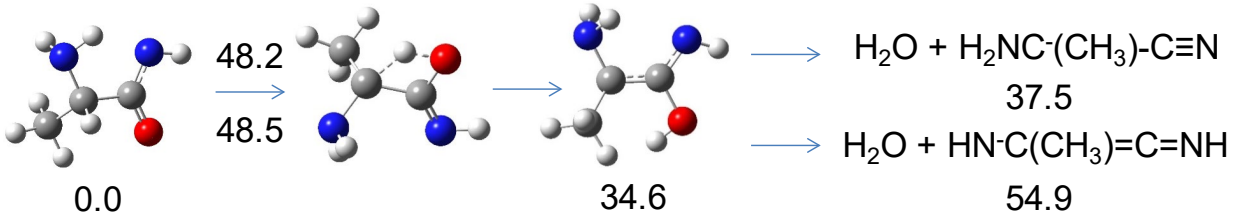


Figure A2.11. Relative energies (kcal/mol) for the dissociation of glycine amide to form H_2O and either $H_2NC(CH_3)$-C≡N or $HN^-C(CH_3)$=C=NH. Energies shown are relative enthalpies (and Gibbs free energy, when shown) calculated at the B3LYP/DZVP2 level.

CHAPTER 3

GAS-PHASE ACIDITIES OF ACIDIC DI- AND TRIPEPTIDES

From Cui, C.; McNeill, A. S.; Jackson, W. C.; Raddatz, M. A.; Dixon, D. A.; Cassady, C. J. *J. Phys. Chem. B* **2019**, *123*, 606 – 613.

Introduction

Acidic peptides are of importance because of their roles in biological activities such as blood clotting,[1] metabolism,[2] and neurotransmission.[3,4] Because acidic peptides readily deprotonate in the gas phase by electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), negative ion mode mass spectrometry (MS) can be a useful analysis tool.[5,6,7,8] In addition, tandem mass spectrometry (MS/MS) on deprotonated negative ions can provide sequence information that complements the more commonly employed protonated positive ion mode.[9],[10],[11],[12] Peptide fragmentation by MS/MS is often charge directed;[13,14] therefore, a knowledge of the location of deprotonation and the energetics of the process can be useful in elucidating peptide fragmentation mechanisms for proteomics applications.

Gas-phase acidity (GA or ΔG_{acid}) is defined as the free energy change (ΔG) of deprotonation reaction (3.1). The GAs of the twenty common amino acids[15,16,17,18,19,20,21,22,23,24,25,26,27,28,29] and three common

$$M \rightarrow [M - H]^- + H^+ \quad (3.1)$$

phosphorylated amino acids[30] have been determined, as well as the GAs of their amides,[30,31] which are representative of residues in peptides. However, there have been few studies on the GAs and sites of deprotonation in peptides. Using a combined approach employing

gas-phase proton transfer reactions and G3(MP2) composite correlated molecular orbital (MO) theory calculations, we have previously determined the GAs of six tripeptides containing neutral amino acid residues.[32] The tripeptides are about 10 kcal/mol more acidic (i.e., have a lower numerical GA value) than their constituent amino acids (glycine and alanine) and, as expected, their site of deprotonation is the C-terminal carboxylic acid group. The methyl esters of these tripeptides were also studied; the lack of a C-terminal acidic site resulted in preferred deprotonation at central amide nitrogens (-NH-) along the peptide backbone. In addition, Ren and coworkers have investigated the GAs of cysteine-containing peptide amides (with 2-5 residues) using the extended Cooks kinetic method of collision-induced dissociation (CID) on a proton-bound dimer composed of the analyte and a reference acid.[33,34,35,36] The thiol side chain, -SH, of cysteine can serve as a gas-phase deprotonation site. Ren and coworkers found that these peptides are more acidic when the cysteine residue is located at the N-terminus rather than the C-terminus.[33,34,35] In addition, increasing the length of the peptide chain made the peptides more acidic.[33,35,36] In previous studies on the gas-phase basicities (GBs) of polyglycines, peptides also become more basic as the length of the peptide chain increases.[37,38,39]

The goal of the current work is to explore the effects of acidic residue position on the acidity and deprotonation sites of peptides. Experimental and computational studies were performed on ten model di- and tripeptides composed of inert alanine (A) residues and one acidic X residue of either aspartic acid (D) or glutamic acid (E): AX, XA, AAX, AXA, and XAA.

Experimental and Computational Methods

Experimental Methods. All mass spectrometry experiments employed a Bruker (Billerica, MA, USA) BioApex 7T Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer. Peptide solutions were at concentrations of 60 μM in a solvent of 49.5:49.5:1 (v:v:v)

methanol:water:ammonium hydroxide. Solutions were introduced into an Apollo ESI source using a syringe pump at the rate of 120 μL/h. Air was the ESI drying and nebulizing gas at a temperature of ~220ºC. The needle of the ESI source was grounded and the capillary entrance and end plate were held at potentials of 3.5-4.0 kV for optimizing negative ion formation. Ions were allowed to accumulate in a hexapole for 500-1000 ms before being transferred into the FT-ICR cell by electrostatic focusing. Deprotonated peptides, $[M - H]^-$, were isolated by correlated frequency resonance ejection techniques[40] and were allowed to react with a series of reference compounds of known GAs.[41] Neutral reference compounds were introduced into the FT-ICR cell through a leak valve at a constant pressure in the range of $(1.0 - 10) \times 10^{-8}$ mbar. Pressures were measured with a calibrated ion gauge.[15,39]

Experimental GAs were assigned with the thermokinetic method, which we have previously used to obtain excellent agreement between experimental and computational GA values.[15,25,30,31,32] Experimental rate constants (k_{exp}) were obtained by observing the pseudo-first order decay of precursor ion intensity as a function of reaction time. Reaction times were varied in the range of 0-30 s. Under near thermoneutral conditions where deprotonation was in competition with proton-bound dimer formation, k_{exp} was determined by fitting the experimental reaction data as discussed previously.[39] The reaction efficiency (RE) was calculated as the ratio of k_{exp} to the theoretical thermal capture rate constant.[42,43] An RE value of 0.269 was used to assign the "break point" (based on the work of Bouchoux and co-workers[44,45,46,47]), where the ion/molecule reaction becomes exoergic. The GA value where RE is 0.269 was determined from a straight line fit of a plot of experimental REs versus GAs between the two bracketing reference compounds for each reaction. Reported experimental uncertainities in GA were obtained using

the closeness of the experimental GA to the bracketing reference compound GAs; an additional 2 kcal/mol was added to this value to account for uncertainities in the reference GAs.

Computational Methods. Initial geometries for all structures were optimized at the density functional theory (DFT) level using the B3LYP exchange-correlation functional[48,49] and the DZVP2 basis set.[50] The vibrational frequencies were calculated to show that the optimized structures are energetic minima. The calculated frequencies also provide zero point and thermal corrections to the enthalpy and the entropy for each species so that free energies could be calculated and compared directly to experimental GA values. The conformational space for each structure was searched by manual sampling at the DFT level for neutral di- and tripeptides as well as the anions that result from deprotonation at the C-terminus and at the carboxylic acid group on the side chain of the acidic residue. The composite correlated molecular orbital (MO) theory method G3(MP2)[51] performs better in the prediction of bond energies as well as steric, non-bonded interactions than do the most widely used DFT exchange-correlation functionals for these types of compounds. In previous work,[15,25,26,30,31,32] GAs predicted at the G3(MP2) level are generally in excellent agreement (±1 kcal/mol) with experimentally derived values and with higher level CCSD(T) calculations extrapolated to the complete basis set limit with additional corrections.[52,53,54,55,56] The reported GAs were calculated from free energy ($\Delta G_{gas,\ 298\ K}$) at the G3(MP2) level. The program Gaussian09 was used to perform calculations at the DFT and correlated MO theory levels.[57]

Results

In the mass spectrometry experiments, deprotonated peptides, $[M - H]^-$, were reacted with neutral reference compounds of known GAs. The reference compounds, their GAs, and the resulting reaction efficiencies (REs) are listed in Table 3.1. Note that a lower numerical GA

value indicates a more acidic species. To insure the validity of the reference compound GAs from the literature,[41] the GAs for these compounds were also calculated using the computational procedures outlined above. In all cases, excellent agreement was found between our calculated GAs and the literature values for the reference compounds.

For the ten acidic di- and tripeptides, all experimental pseudo-first order kinetic plots were linear (as illustrated in Figure 3.1), suggesting that there is only one major deprotonated ion structure produced under the present ESI conditions or that, if multiple structures are formed, their GAs are too close to be distinguished experimentally. In past work, we have used ion/molecule reactions to experimentally distinguish between the GAs of two deprotonated tyrosine structures that differ by ~1.2 kcal/mol,[15] although in this case both structures were of significant abundance.

The ten acidic peptides have two potential deprotonation sites, the C-terminal carboxylic acid group and the side chain carboxylic acid group of the acidic residue. Table 3.2 provides calculated deprotonation energies for both sites, as well as experimental GAs. C-terminal deprotonation is predicted to be more acidic than side chain deprotonation for DA, AE, AAD, DAA, and AAE, whereas side chain deprotonation is favored for AD, ADA, AEA, and EAA. The two values for EA are within 1 kcal/mol with side chain deprotonation being slightly favored. Overall, the experimental GA values are in excellent agreement with the calculated values within experimental error.

Dipeptides. Due to the uncertainties in the experimentally derived GAs (± 3 to 4 kcal/mol), it is not straightforward to assign deprotonation sites based on the calculated GA values. In addition, it is possible that two (or more) structures of similar energies are present experimentally because ESI does not always exclusively form the lowest energy structure. For example, under very

similar ESI conditions in our laboratory, twelve of the twenty common amino acid amides deprotonate to generate two structures even though the GAs associated with their formation differ by 4-7 kcal/mol.[31] For AD, the predicted deprotonation site is the side chain (306.4 kcal/mol), which is more favorable (lower in energy) by 4 kcal/mol than C-terminal deprotonation. The experimental value (308.2 kcal/mol) falls intermediate between the two calculated values. For DA, the more acidic site is predicted to be the C-terminus (313.5 kcal/mol) with the side chain site 5.5 kcal/mol less acidic; however, the experimental value (317.8 kcal/mol) is closer to the less acidic side chain value. The experimental GA for AE (310.6 kcal/mol) is only 0.8 kcal/mol lower than the calculated side chain deprotonation GA (311.4 kcal/mol), which is the less acidic site. Although it is tempting to assign the experimental deprotonation site as the side chain, the C-terminus is only 2.2 kcal/mol more acidic, so either site could be deprotonated given the experimental error bars and the fact that ESI does not always exclusively produce the lowest energy structure. For EA, the experimental value (317.3 kcal/mol) is less acidic than either calculated value (313.7 and 314.7 kcal/mol), which are very close to each other with the side chain being the more acidic by only 1 kcal/mol.

Experimental GA values indicate that AD and AE (Figure 3.2) are more acidic than DA and EA, respectively (9.6 kcal/mol lower in energy for AD and 6.7 kcal/mol lower in energy for AE). The calculated structures suggest that this residue position effect on GA is related to the ability of two carboxylic acid groups (at the C-terminus and the side chain) to share a lone proton when they are in close proximity, which increases the stability of the deprotonated ion and consequently makes its neutral more acidic. Similar deprotonation schemes are available for all other dipeptides and tripeptides in the Appendix.

When considering the effect of different acidic residues on GAs, DA has very similar acidity to EA and AD is more acidic than AE. To confirm this trend, ion/molecule reactions were performed using peptide mixtures. One mixture contained DA and EA, while another mixture contained AD and AE; all peptides were at equal molar concentration. The deprotonated ions from two peptides were reacted simultaneously with 1,1,1,5,5,5-hexafluoro-2,4-pentanedione (HFPD). For the mixture of DA and EA, the two ions reached the 100% completion at almost the same reaction time, which indicates that the GAs of DA and EA are very close to each other. For the other mixture, AE reached the 100% completion before AD, which confirms that AD is more acidic than AE. Thus, the results of both mixture experiments are consistent with the experimental GA values in Table 3.2.

Tripeptides. C-terminal deprotonation (306.7 kcal/mol) is probable for AAD because side chain deprotonation (310.3 kcal/mol) is predicted to be 3.6 kcal/mol less acidic. There is excellent agreement between the predicted value for C-terminal deprotonation and experiment (306.3 kcal/mol). The experimental GA value for ADA (312.8 kcal/mol) falls between the calculated values (310.0 and 316.2 kcal/mol) but is closer to the more acidic site, side chain deprotonation. For DAA, the C-terminus (305.3 kcal/mol) is the more acidic site by 5.1 kcal/mol. Again, the experimental value (308.1 kcal/mol) falls between the two calculated values.

The deprotonation sites can be more clearly assigned for the tripeptides with E residues. The experimental GA of AAE (307.8 kcal/mol) is very close to the calculated lowest energy (308.9 kcal/mol), which has deprotonation at the C-terminus. In addition, C-terminal deprotonation for AAE is evident because side chain deprotonation is 16.9 kcal/mol less acidic than its experimental GA. The significant difference in the GAs for the two deprotonation sites of AAE is due to the unique ability of the glutamic acid side chain near the C-terminus to orient

itself very close to the C-terminus. This is a consequence of the longer E side chain, which has one more methylene (-CH_2-) than a D side chain. If the two carboxylic acid groups are spatially close, the proton is shared between the C-terminus and side chain with the negative charge at the C-terminus. Therefore, as shown in Figure 3.3 for AAE, deprotonation of the side chain can only be found when the carboxylate group at the side chain is oriented far from the C-terminus, substantially reducing the possibility of hydrogen bonding to stabilize the side chain. We do not observe this large difference in GAs for the two deprotonation sites of AE because the N-terminus is close enough to the deprotonated side chain to stabilize that structure (Figure 3.2), which is not the case for AAE.

Side chain deprotonation for AEA (310.2 kcal/mol) is predicted to be more acidic by 5.8 kcal/mol. The experimental value (311.0 kcal/mol) is within 1.0 kcal/mol of the calculated value for the more acidic side chain site and, thus, is most consistent with side chain deprotonation. For EAA, side chain deprotonation (309.3 kcal/mol) is predicted to be 6.8 kcal/mol more acidic and there is very good agreement with experiment (308.3 kcal/mol) for side chain deprotonation.

Among tripeptides, the lowest acidity occurs when the acidic residue is located at the center (ADA or AEA). Again, tripeptides are slightly more acidic when the acidic residues are at the C-terminus (AAD or AAE). The presence of D either results in slightly more acidity (AAD at 306.3 kcal/mol versus AAE at 307.8 kcal/mol) or has negligible effect on their GAs (EAA and DAA). To confirm the GA trends between D and E residues in tripeptides, ion/molecule reactions involving mixtures (equal molar for each peptide) were performed using the reference compound bistrifluoroacetamide (BTFA). For AAD and AAE, $[M - H]^-$ from AAE reacted with BTFA faster and reached 100% completion first, indicating that AAD is more acidic. For DAA and EAA, the two ions reached 100% completion at almost the same time, which confirms near

identical GAs. Deprotonated ions from ADA and AEA reacted at a very similar rate with ADA reaching 100% completion slightly faster than AEA. The results for all three mixture experiments are in agreement with our experimental GAs in Table 3.2.

Comparison with Relevant Amino Acids, Amides, and Peptides. The acidities of the di- and tripeptides (Table 3.2) can be compared with those of the amino acids alanine (A),[17,26] aspartic acid (D),[25] and glutamic acid (E),[25] their corresponding amides,[25,31] dialanine (AA), and trialanine (AAA) (Table 3.3).[32] Our previous evaluation of experimental and computational GAs indicates that the amino acids forms of D and E are deprotonated at the C-terminus, which is the more acidic site.[25] All studied acidic di- and tripeptides are more acidic than their relevant amino acids and amides. Di- and tripeptides are larger than single amino acids, so there are greater opportunities for the formation of stabilizing hydrogen bonds, which reduces the energy required for deprotonation. This is consistent with the GA trend from alanine to trialanine (increasing the number of alanine residues results in more acidity). The acidic tripeptides are more acidic than dipeptides when the same acidic residue at the same location (C-terminus or N-terminus).

The GAs for the more acidic sites of AD (side chain deprotonation) and DA (C-terminal deprotonation) are more acidic than aspartic acid by 9 and 1.9 kcal/mol, respectively. The structures for the products of the two deprotonation reaction pathways of AD have the same number of hydrogen bonding interactions with nearly identical hydrogen bond lengths (≤ 0.03 Å differences) (Figure 3.2). However, the more acidic product of deprotonation at the side chain (by 4.0 kcal/mol) has two of its three hydrogen bonds stabilizing the site of the negative charge at the carboxylate, reducing its energy compared to the product of the C-terminal deprotonation. The more acidic deprotonation product for DA is from the deprotonation of the C-terminus (by 5.5 kcal/mol) due to the sterically hindered nature of the side chain deprotonation product,

despite the fact that this less acidic product has one more hydrogen bond than the more acidic C-terminal product.

The GAs for the more acidic sites of AE (C-terminal deprotonation) and EA (side chain deprotonation) are more acidic by 7.2 and 2.7 kcal/mol, respectively, as compared to that of glutamic acid. The structures resulting from the two deprotonation pathways for AE are similar in that both have the same number of hydrogen bonding sites, though the lengths of the hydrogen bonds for the more acidic C-terminal deprotonation pathway are surprisingly 0.07 to 0.27 Å longer than comparable hydrogen bonds in the side chain deprotonation product (see Appendix). The C-terminal deprotonation of AE results in a structure that is more linear, as opposed to the tightly curled structure for the side chain deprotonation, which may lead to slightly lower energy (2.2 kcal/mol difference) due to reduced steric hindrance. The GAs for the two deprotonation sites of EA are within 1 kcal/mol of each other, so it is difficult to state which site is truly more acidic, although the deprotonation of the side chain was found to be slightly more acidic and the structure resulting from side chain deprotonation has an additional hydrogen bond.

For the most acidic sites of tripeptides AAD (C-terminus), ADA (side chain), and AAD (C-terminus), the GAs are more acidic by 5 to 10 kcal/mol as compared to aspartic acid. In the case of AAD, the more acidic pathway by 3.6 kcal/mol involves deprotonation at the C-terminus and results in a more linear product than the deprotonation at the side chain (see Appendix). Both products have nearly identical hydrogen bonds with lengths of less than 2 Å (0.02 to 0.04 Å differences), but the structure resulting in lower steric hindrance is energetically favored. The favored side chain deprotonation pathway for AAD also results in a somewhat more linear structure, though it also has an increased number of hydrogen bonds with lengths less than 2 Å. Side chain deprotonation is favored for ADA by 6.2 kcal/mol as it results in an increased number

of hydrogen bonds with lengths less than 2 Å and a more linear, less sterically-hindered product structure. DAA favors the C-terminal deprotonation by 5.1 kcal/mol. This is likely due to the D being located two residues away from the C-terminus, which allows for the transfer of the C-terminal hydrogen to the deprotonated side chain if they are near one another. The lowest energy conformation for side chain deprotonation of DAA has the carboxylate group of the side chain spatially separated from the C-terminus by a significant distance.

Side chain deprotonation for AAE is within 0.4 kcal/mol of the side chain deprotonation value for glutamic acid. When the E residue is located at either the N-terminus (EAA) or the central residue (AEA), the GA for the less acidic pathway is within 0.4 kcal/mol of the value of the more acidic C-terminal deprotonation of the amino acid form of glutamic acid. Similar trends are not observed for the analogous tripeptides with D replacing E residues. The mixed di- and tripeptides are more acidic than the pure amino acids. The tripeptides with glutamic acid, AAE (C-terminus by 15.8 kcal/mol), AEA (side chain by 5.8 kcal/mol), and EAA (side chain by 6.8 kcal/mol), are 6 to 8 kcal/mol more acidic than glutamic acid. Each of these preferred deprotonation pathways results in a structure with a greater number of hydrogen bonds with distances less than 2 Å. Unlike the tripeptides involving D, the tripeptides with E have deprotonation products that are conformationally very similar to one another regardless of the deprotonation site. Side chain deprotonation requires that the side chain carboxylate be spatially distant from the C-terminus for all tripeptides with the E residue, as discussed above for AAE.

Conclusions

The GA values of ten acidic peptides have been assigned based on both experiments and electronic structure calculations, and these values are in excellent agreement. C-terminal deprotonation is predicted to be more acidic than side chain deprotonation for DA, AE, AAD,

DAA, and AAE. Side chain deprotonation is favored for AD, ADA, AEA, and EAA. The two values for EA are within 1 kcal/mol with the side chain deprotonation being slightly favored. Linear pseudo-first order kinetic plots of all ion/molecule reactions indicate either only one major ion structure is formed or multiple ion structures with very similar GAs are formed by ESI mass spectrometry. The studied di- and tripeptides are more acidic than the relevant amino acids and their amides, which is consistent with stronger hydrogen bonds formed in larger peptides. Dipeptides are more acidic when acidic residues are at the C-terminus. Tripeptides are only slightly more acidic with acidic residues at the C-terminus. Replacing the E by D at the termini of a peptide will cause the peptide to be more acidic (when E is at the C-terminus) or have negligible effect on its acidity (when E is at the N-terminus).

Tables

Table 3.1. Reaction Efficiencies of Ion/Molecule Reactions Between Acidic Di- and Tripeptides and Reference Compounds

Reference Compound	GA[a] Lit	GA[b] Calc	Reaction Efficiency (Mean ± Standard Deviation)									
			AD	DA	AE	EA	AAD	ADA	DAA	AAE	AEA	EAA
Trifluoropropionic acid	326.9	327.1	0.00	0.00	0.00	0.01 ± 0.01	0.00	0.00	0.00	0.00	0.01 ± 0.00	0.01 ± 0.01
Difluoroacetic acid	323.8	323.0	0.00	0.03 ± 0.00	0.00	0.00	0.00	0.02 ± 0.00	0.00	0.00	0.01 ± 0.00	0.01 ± 0.00
Pentafluorophenol	320.8	322.3	0.01 ± 0.01	0.07 ± 0.02	0.01 ± 0.00	0.08 ± 0.03	0.01 ± 0.00	0.03 ± 0.01	0.01 ± 0.00	0.01 ± 0.00	0.02 ± 0.00	0.02 ± 0.00
				Break[c]		Break						
Trifluoroacetic acid	317.4	316.9	0.02 ± 0.01	0.29 ± 0.07	0.06 ± 0.00	0.27 ± 0.04	0.02 ± 0.01	0.07 ± 0.01	0.03 ± 0.01	0.04 ± 0.01	0.08 ± 0.02	0.06 ± 0.01
Heptafluorobutyric acid	314.9	313.2	0.04 ± 0.01	0.35 ± 0.02	0.09 ± 0.03	0.32 ± 0.01	0.03 ± 0.02	0.21 ± 0.01	0.09 ± 0.02	0.08 ± 0.00	0.14 ± 0.01	0.11 ± 0.02
					Break			Break			Break	
1,1,1,5,5,5-Hexafluoro-2,4,pentanedione	310.3	309.8	0.08 ± 0.01	0.48 ± 0.00	0.29 ± 0.00	0.47 ± 0.09	0.02 ± 0.00	0.31 ± 0.03	0.13 ± 0.04	0.08 ± 0.01	0.29 ± 0.02	0.14 ± 0.02
			Break						Break	Break		Break
Bistrifluoroacetamide	307.5	308.5	0.33 ± 0.04		0.39 ± 0.02		0.19 ± 0.02	0.46 ± 0.02	0.29 ± 0.01	0.30 ± 0.02	0.35 ± 0.04	0.31 ± 0.04

[a] "GA Lit" values for the reference compounds are from reference 41 and are given in kcal/mol with error of ± 2 kcal/mol.

[b] "GA Calc" values were obtained in the present study using the same computational methods employed to obtain peptide GAs and are given in kcal/mol.

[c] "Break" indicates the point where the GA was assigned.

Table 3.2. Experimental and G3(MP2) Theoretical GAs in kcal/mol for Di- and Tripeptides

Peptide	Calc C-terminus (kcal/mol)	Calc Side Chain (kcal/mol)	Expt (kcal/mol)
AD	310.4	306.4	308.2 ± 3.4
DA	313.5	319.0	317.8 ± 3.7
AE	309.2	311.4	310.6 ± 4.3
EA	314.7	313.7	317.3 ± 3.7
AAD	306.7	310.3	306.3 ± 4.0
ADA	316.2	310.0	312.8 ± 4.3
DAA	305.3	310.4	308.1 ± 3.4
AAE	308.9	324.7	307.8 ± 3.4
AEA	316.0	310.2	311.0 ± 4.3
EAA	316.1	309.3	308.3 ± 3.4

Table 3.3. GAs for Relevant Amino Acids, Their Amides, and Peptides in kcal/mol

Peptide	Calc C-terminus	Calc Side Chain	Expt
A amino acid	334.6[a]		334.8[b]
A-NH_2	351.1[c]		no expt
AA[d]	324.1		no expt
AAA[e]	321.6		322.2
D amino acid[f]	315.4	325.9	315.3
D-NH_2[f]	325.9		326.5
E amino acid[f]	316.4	324.3	318.2
E-NH_2[f]	326.4		328.7

[a] Calculated GA from reference 26.
[b] Experimental GA from reference 17.
[c] Calculated GA from reference 31.
[d] GA calculated for the present work.
[e] Experimental and calculated GAs from reference 32.
[f] Experimental and calculated GAs from reference 25.

Figures

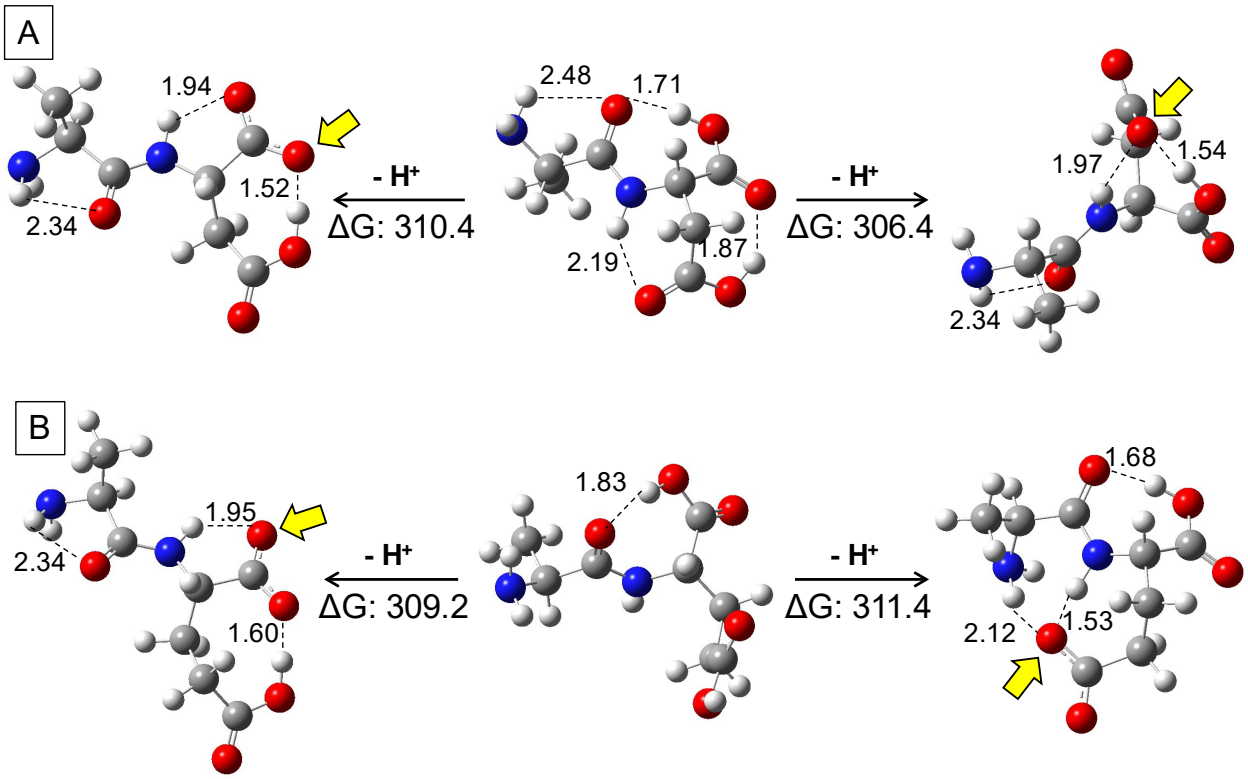


Figure 3.1. Gas-phase deprotonation reaction pathways for AD (A) and AE (B). GAs are given in kcal/mol, hydrogen bonding lengths are given in Å, and yellow arrows indicate the site of deprotonation for product anions.

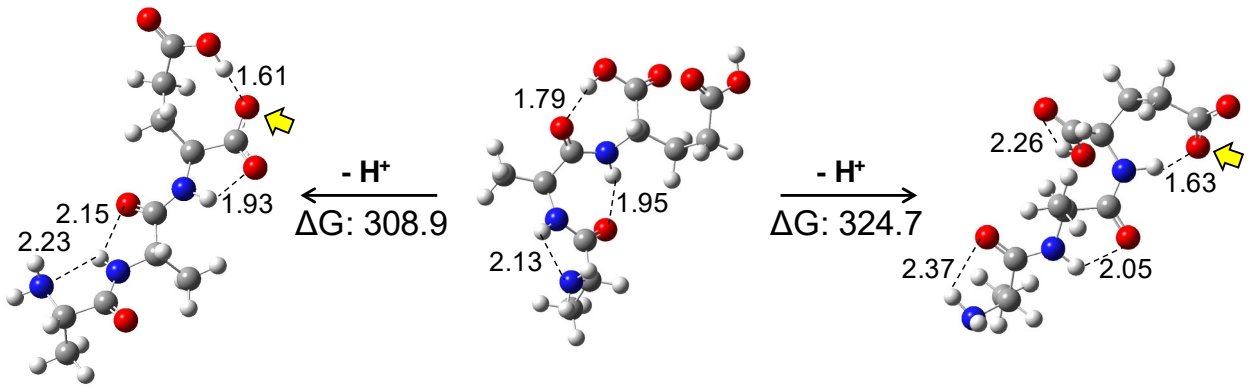


Figure 3.2. Gas-phase deprotonation reaction pathways for AAE. GAs are given in kcal/mol, hydrogen bonding lengths are given in Å, and yellow arrows indicate the site of deprotonation for product anions.

References

[1] Voet, D.; Voet, J. G.; Pratt, C. W. In *Fundamentals of Biochemistry: Life at the Molecular Level;* John Wiley & Sons: 2013; pp 1200.

[2] Shulkes, A.; Baldwin, G. S. In *Gastrin;* Handbook of Biologically Active Peptides (Second Edition); Elsevier: 2013; pp 1219 – 1226.

[3] Tatemoto, K. Neuropeptide Y: Complete Amino Acid Sequence of the Brain Peptide. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5485 – 5489.

[4] Pierotti, A. R.; Prat, A.; Chesneau, V.; Gaudoux, F.; Leseney, A. M.; Foulon, T.; Cohen, P. N-Arginine Dibasic Convertase, a Metalloendopeptidase as a Prototype of a Class of Processing Enzymes. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6078 – 6082.

[5] Jai-nhuknan, J.; Cassady, C. J. Negative Ion Postsource Decay Time-of-Flight Mass Spectrometry of Peptides Containing Acidic Amino Acid Residues. *Anal. Chem.* **1998**, *70*, 5122 – 5128.

[6] Gao, J.; Cassady, C. J. Negative Ion Production from Peptides and Proteins by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 4066 – 4072.

[7] Chen, Y.; Watson, H. M.; Gao, J.; Sinha, S. H.; Cassady, C. J.; Vincent, J. B. Characterization of the Organic Component of Low-Molecular-Weight Chromium-Binding Substance and its Binding of Chromium. *J. Nutr.* **2011**, *141*, 1225 – 1232.

[8] Ewing, N. P.; Cassady, C. J. Dissociation of Multiply-Charged Negative Ions for Hirudin (54-65), Fibrinopeptide B, and Insulin A (Oxidized). *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 105 – 116.

[9] Andreazza, H. J.; Wang, T.; Bagley, C. J.; Hoffmann, P.; Bowie, J. H. Negative Ion Fragmentations of Deprotonated Peptides. The Unusual Case of isoAsp: A Joint Experimental and Theoretical Study. Comparison with Positive Ion Cleavages. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1993 – 2002.

[10] Jai-nhuknan, J.; Cassady, C. J. Anion and Cation Post-Source Decay Time-of-Flight Mass Spectrometry of Small Peptides: Substance P, Angiotensin II, and Renin Substrate. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1678 – 1682.

[11] Yagami, T.; Kitagawa, K.; Futaki, S. Liquid Secondary-Ion Mass Spectrometry of Peptides Containing Multiple Tyrosine-O-Sulfates. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1335 – 1341.

[12] Pu, D.; Clipston, N. L.; Cassady, C. J. A Comparison of Positive and Negative Ion Collision-Induced Dissociation for Model Heptapeptides with One Basic Residue. *J. Mass Spectrom.* **2010**, *45*, 297 – 305.

[13] Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model. *J. Am. Chem. Soc.* **1996**, *118*, 8365 – 8374.

[14] Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *J. Mass Spectrom.* **2000**, *35*, 1399 – 1406.

[15] Bokatzian, S. S.; Stover, M. L.; Plummer, C. E.; Dixon, D. A.; Cassady, C. J. An Experimental and Computational Investigation into the Gas-Phase Acidities of Tyrosine and Phenylalanine: Three Structures for Deprotonated Tyrosine. *J. Phys. Chem. B* **2014**, *118*, 12630 – 12643.

[16] Fournier, F.; Afonso, C.; Fagin, A. E.; Gronert, S.; Tabet, J. C. Can Cluster Structure Affect Kinetic Method Measurements? The Curious Case of Glutamic Acid's Gas-Phase Acidity. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1887 – 1896.

[17] Jones, C. M.; Bernier, M.; Carson, E.; Colyer, K. E.; Metz, R.; Pawlow, A.; Wischow, E. D.; Webb, I.; Andriole, E. J.; Poutsma, J. C. Gas-Phase Acidities of the 20 Protein Amino Acids. *Int. J. Mass Spectrom.* **2007**, *267*, 54 – 62.

[18] Locke, M. J.; Hunt, R. L.; McIver Jr., R. T. Experimental Determinations of the Acidity and Basicity of Glycine in the Gas Phase. *J. Am. Chem. Soc.* **1979**, *101*, 272 – 273.

[19] Locke, M. J.; McIver Jr., R. T. Effect of Solvation on the Acid/Base Properties of Glycine. *J. Am. Chem. Soc.* **1983**, *105*, 4226 – 4232.

[20] O'Hair, R. A. J.; Bowie, J. H.; Gronert, S. Gas-Phase Acidities of the Alpha-Amino Acids. *Int. J. Mass Spectrom. Ion Proc.* **1992**, *117*, 23 – 36.

[21] Riffet, V.; Bourcier, S.; Bouchoux, G. Gas-Phase Basicity and Acidity of Tryptophan. *Int. J. Mass Spectrom.* **2012**, *316–318*, 47 – 56.

[22] Topol, I. A.; Burt, S. K.; Russo, N.; Toscano, M. Theoretical Calculations of Glycine and Alanine Gas-Phase Acidities. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 318 – 322.

[23] Vyas, N.; Ojha, A. K. Calculation of Dissociation Constants and Chemical Hardness of some Biologically Important Molecules: A Theoretical Study. *J. Quantum Chem.* **2011**, *111*, 3961 – 3970.

[24] Webb, I. K.; Muetterties, C. E.; Platner, C. B.; Poutsma, J. C. Gas-Phase Acidities of Lysine Homologues and Proline Analogs from the Extended Kinetic Method. *Int. J. Mass Spectrom.* **2012**, *316–318*, 126 – 132.

[25] Li, Z.; Matus, M. H.; Velazquez, H. A.; Dixon, D. A.; Cassady, C. J. Gas-Phase Acidities of Aspartic Acid, Glutamic Acid, and their Amino Acid Amides. *Int. J. Mass Spectrom.* **2007**, *265*, 213 – 223.

[26] Stover, M. L.; Jackson, V.; Matus, M.; Adams, M.; Cassady, C. J.; Dixon, D. A. Fundamental Thermochemical Properties of Amino Acids: Gas-Phase and Aqueous Acidities and Gas-Phase Heats of Formation. *J. Phys. Chem. B* **2012**, *116*, 2905 – 2916.

[27] Tian, Z.; Wang, X. B.; Wang, L. S.; Kass, S. R. Are Carboxyl Groups the most Acidic Sites in Amino Acids? Gas-Phase Acidities, Photoelectron Spectra, and Computations on Tyrosine, p-Hydroxybenzoic Acid, and their Conjugate Bases. *J. Am. Chem. Soc.* **2009**, *131*, 1174 – 1181.

[28] Tian, Z.; Pawlow, A.; Poutsma, J. C.; Kass, S. R. Are Carboxyl Groups the most Acidic Sites in Amino Acids? Gas-Phase Acidity, H/D Exchange Experiments, and Computations on Cysteine and its Conjugate Base. *J. Am. Chem. Soc.* **2007**, *129*, 5403 – 5407.

[29] Caldwell, G.; Renneboog, R.; Kebarle, P. Gas-Phase Acidities of Aliphatic Carboxylic Acids, Based on Measurements of Proton Transfer Equilibria. *Can. J. Chem.* **1989**, *67*, 611 – 618.

[30] Stover, M. L.; Plummer, C. E.; Miller, S. R.; Cassady, C. J.; Dixon, D. A. Gas-Phase Acidities of Phosphorylated Amino Acids. *J. Phys. Chem. B* **2015**, *119*, 14604 – 14621.

[31] Plummer, C. E.; Stover, M. L.; Bokatzian, S. S.; Davis, J. T. M.; Dixon, D. A.; Cassady, C. J. An Experimental and Computational Study of the Gas-Phase Acidities of the Common Amino Acid Amides. *J. Phys. Chem. B* **2015**, *119*, 9661 – 9669.

[32] Bokatzian-Johnson, S. S.; Stover, M. L.; Dixon, D. A.; Cassady, C. J. Gas-Phase Deprotonation of the Peptide Backbone for Tripeptides and their Methyl Esters with Hydrogen and Methyl Side Chains. *J. Phys. Chem. B* **2012**, *116*, 14844 – 14858.

[33] Morishetti, K. K.; De Suan Huang, B.; Yates, J. M.; Ren, J. Gas-Phase Acidities of Cysteine-Polyglycine Peptides: The Effect of the Cysteine Position. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 603 – 614.

[34] Shen, J.; Ren, J. Gas Phase Acidity of a Cysteine Residue in Small Oligopeptides. *Int. J. Mass Spectrom.* **2012**, *316-318*, 147 – 156.

[35] Ren, J.; Tan, J. P.; Harper, R. T. Gas-Phase Acidities of Cysteine-Polyalanine Peptides I: A(3,4)CSH and HSCA(3,4). *J. Phys. Chem. A* **2009**, *113*, 10903 – 10912.

[36] Tan, J.; Ren, J. Determination of the Gas-Phase Acidities of Cysteine-Polyalanine Peptides using the Extended Kinetic Method. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 188 – 194.

[37] Wu, Z.; Fenselau, C. Proton Affinities of Polyglycines Assessed by using the Kinetic Method. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 863 – 866.

[38] Wu, J.; Lebrilla, C. B. Gas-Phase Basicities and Sites of Protonation of Glycine Oligomers (GLY_N; N = 1-5). *J. Am. Chem. Soc.* **1993**, *115*, 3270 – 3275.

[39] Zhang, K.; Zimmerman, D. M.; Chung-Phillips, A.; Cassady, C. J. Experimental and Ab Initio Studies of the Gas-Phase Basicities of Polyglycines. *J. Am. Chem. Soc.* **1993**, *115*, 10812 – 10822.

[40] de Koning, L. J.; Nibbering, N. M. M.; van Orden, S. L.; Laukien, F. H. Mass Selection of Ions in a Fourier Transform Ion Cyclotron Resonance Trap using Correlated Harmonic Excitation Fields (CHEF). *Int. J. Mass Spectrom. Ion Proc.* **1997**, *165/166*, 209 – 219.

[41] Bartmess, J. E. In *Negative Ion Energetics Data;* Linstrom, P. J., Ed.; NIST Chemistry WebBook, NIST Standard Reference Database Number 69; National Institute of Standards and Technology: Gaithersburg MD, 20899, http://webbook.nist.gov, (retrieved November 5, 2018).

[42] Su, T.; Chesnavich, W. J. Parametrization of the Ion-Polar Molecule Collision Rate-Constant by Trajectory Calculations. *J. Chem. Phys.* **1982**, *76*, 5183 – 5185.

[43] Su, T. Erratum: Trajectory Calculations of Ion-Polar Molecule Capture Rate Constants at Low Temperatures. *J. Chem. Phys.* **1988**, *89*, 5355.

[44] Bouchoux, G.; Salpin, J. Y.; Leblanc, D. A Relationship between the Kinetics and Thermochemistry of Proton Transfer Reactions in the Gas Phase. *Int. J. Mass Spectrom. Ion Proc.* **1996**, *153*, 37 – 48.

[45] Bouchoux, G.; Salpin, J. Y. Re-Evaluated Gas Phase Basicity and Proton Affinity Data from the Thermokinetic Method. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 932 – 936.

[46] Bouchoux, G.; Salpin, J. Y. Gas-Phase Basicity of Glycine, Alanine, Proline, Serine, Lysine, Histidine and Some of their Peptides by the Thermokinetic Method. *Eur. J. Mass Spectrom.* **2003**, *9*, 391 – 402.

[47] Bouchoux, G.; Buisson, D.; Bourcier, S.; Sablier, M. Application of the Kinetic Method to Bifunctional Bases ESI Tandem Quadrupole Experiments. *Int. J. Mass Spectrom.* **2003**, *228*, 1035 – 1054.

[48] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.

[49] Lee, C.; Yang W.; Parr, R. G. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Phys. Rev. B* **1988**, *37*, 785 – 789.

[50] Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. Optimization of Gaussian-Type Basis Sets for Local Spin Density Functional Calculations. Part I. Boron through Neon, Optimization Technique and Validation. *Can. J. Chem.* **1992**, *70*, 560 – 571.

[51] Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V.; Pople, J. A. Gaussian-3 Theory Using Reduced Moller-Plesset Order. *J. Chem. Phys.* **1999**, *110*, 4703 – 4709.

[52] Dixon, D. A.; Feller, D.; Peterson, K. A. In *A Practical Guide to Reliable First Principles Computational Thermochemistry Predictions Across the Periodic Table;* Wheeler, R. A., Section Ed. Tschumper, G S, Eds.; Annual Reports in Computational Chemistry; Elsevier: Amsterdam, 2012; Vol. 8, pp 1 – 28.

[53] Feller, D.; Peterson, K. A.; Dixon, D. A. Further Benchmarks of a Composite, Convergent, Statistically Calibrated Coupled-Cluster-Based Approach for Thermochemical and Spectroscopic Studies. *Mol. Phys.* **2012**, *110*, 2381 – 2399.

[54] Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy: Current Strategies and Future Challenges. *Theor. Chem. Acc.* **2012**, *131*, 1 – 20.

[55] Feller, D.; Peterson, K. A.; Dixon, D. A. A Survey of Factors Contributing to Accurate Theoretical Predictions of Atomization Energies and Molecular Structures. *J. Chem. Phys.* **2008**, *129*, 204105-1 – 204105-32.

[56] Feller, D.; Dixon, D. A. Extended Benchmark Studies of Coupled Cluster Theory through Triple Excitations. *J. Chem. Phys.* **2001**, *115*, 3484 – 3496.

[57] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheesman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. *Gaussian 09, Revision B.01*; Gaussian, Inc.: Wallington, CT, **2009**.

Full Reference for Reference 57 (Gaussian 09): Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone,V.; Petersson,G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J. ; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe,D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT, 2009.

Table A3.1. Total energies of all lowest-energy peptides and anions in the gas phase at the G3(MP2) level (kcal/mol).

Peptide	Deprotonation Site	ΔH(0 K)	ΔH(298 K)	ΔG(298 K)
AA	No Deprotonation	-570.258075	-570.244720	-570.296892
	C-terminus	-569.730646	-569.717254	-569.770376
AD	No Deprotonation	-758.621031	-758.605175	-758.664174
	C-terminus	-758.122200	-758.106495	-758.165937
	Side-Chain	-758.115946	-758.100270	-758.159802
DA	No Deprotonation	-758.624470	-758.608128	-758.668728
	C-terminus	-758.108522	-758.093063	-758.150310
	Side-Chain	-758.114430	-758.098477	-758.159128
AE	No Deprotonation	-797.855103	-797.837412	-797.901787
	C-terminus	-797.353482	-797.336493	-797.399006
	Side-Chain	-797.349919	-797.333223	-797.395466
EA	No Deprotonation	-797.854408	-797.836465	-797.902639
	C-terminus	-797.345418	-797.328211	-797.391130
	Side-Chain	-797.348598	-797.331967	-797.392697
AAD	No Deprotonation	-1005.589936	-1005.568025	-1005.641896
	C-terminus	-1005.090681	-1005.069125	-1005.143127
	Side-Chain	-1005.088685	-1005.067848	-1005.137319
ADA	No Deprotonation	-1005.589037	-1005.567441	-1005.639934
	C-terminus	-1005.075693	-1005.054485	-1005.126048
	Side-Chain	-1005.083988	-1005.062691	-1005.135967
DAA	No Deprotonation	-1005.584485	-1005.562326	-1005.637584
	C-terminus	-1005.090923	-1005.069579	-1005.141119
	Side-Chain	-1005.081693	-1005.060504	-1005.132990
AAE	No Deprotonation	-1044.824462	-1044.801261	-1044.878352
	C-terminus	-1044.321803	-1044.298954	-1044.376031
	Side-Chain	-1044.295735	-1044.272473	-1044.350849
AEA	No Deprotonation	-1044.821803	-1044.798050	-1044.878522
	C-terminus	-1044.312315	-1044.289712	-1044.364936
	Side-Chain	-1044.321405	-1044.298954	-1044.374221
EAA	No Deprotonation	-1044.824945	-1044.801787	-1044.878800
	C-terminus	-1044.313389	-1044.290871	-1044.365119
	Side-Chain	-1044.323303	-1044.300848	-1044.37585

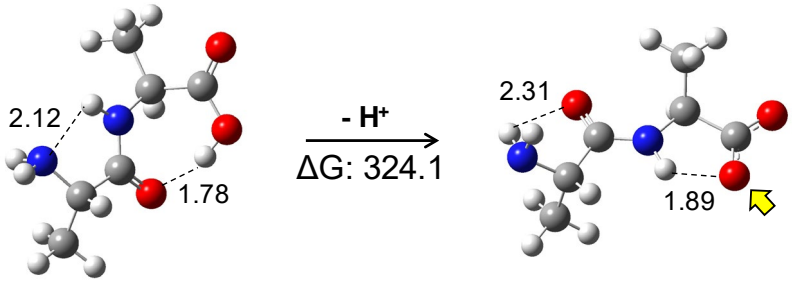


Figure A3.1. Alanine-Alanine, Dialanine, AA

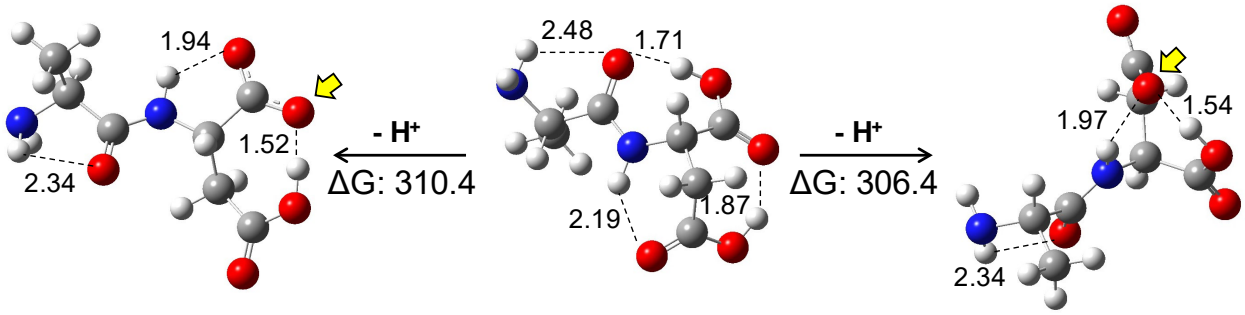


Figure A3.2. Alanine-Aspartic Acid, AD

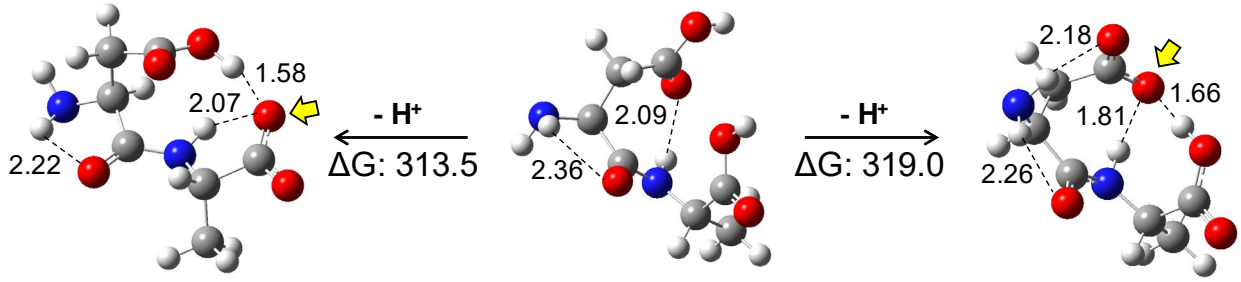


Figure A3.3. Aspartic Acid-Alanine, DA

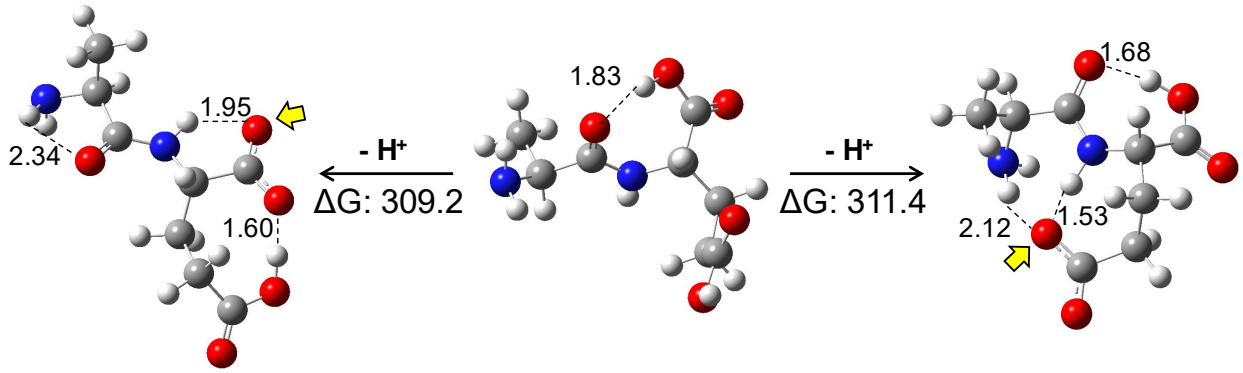


Figure A3.4. Alanine-Glutamic Acid, AE

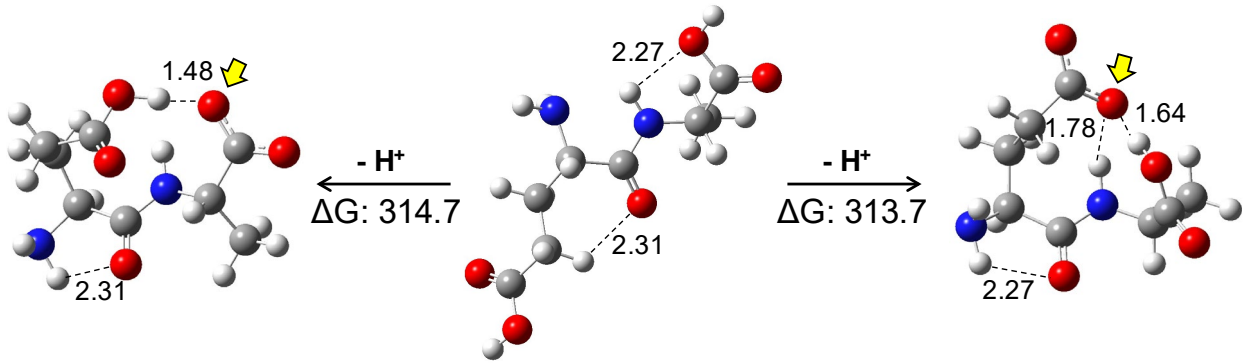


Figure A3.5. Glutamic Acid-Alanine, EA

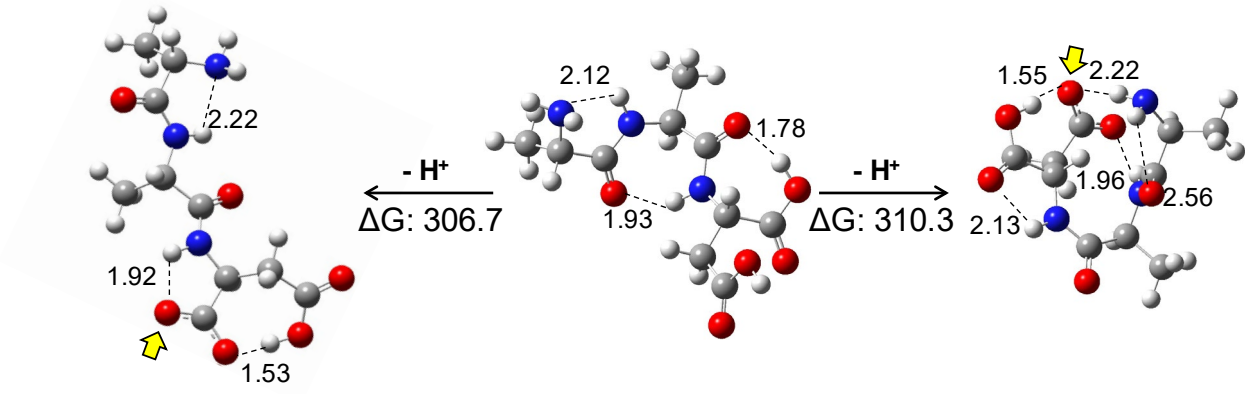


Figure A3.6. Alanine-Alanine-Aspartic Acid, AAD

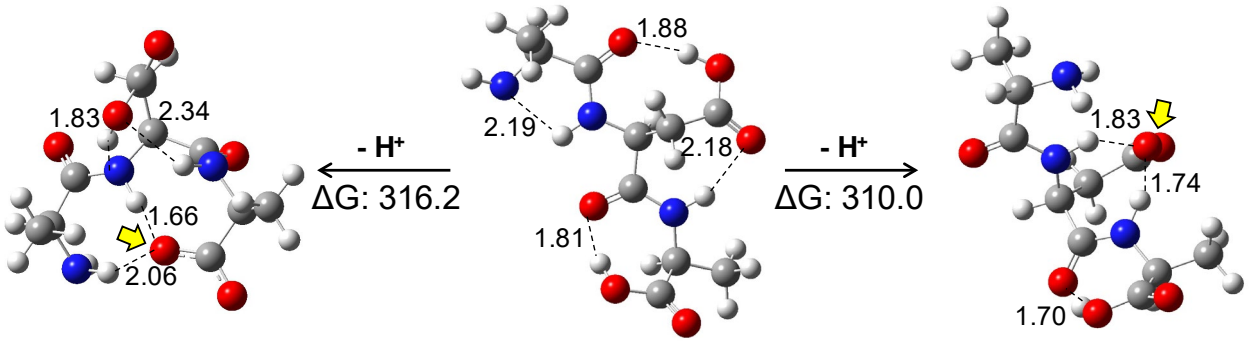


Figure A3.7. Alanine-Aspartic Acid-Alanine, ADA

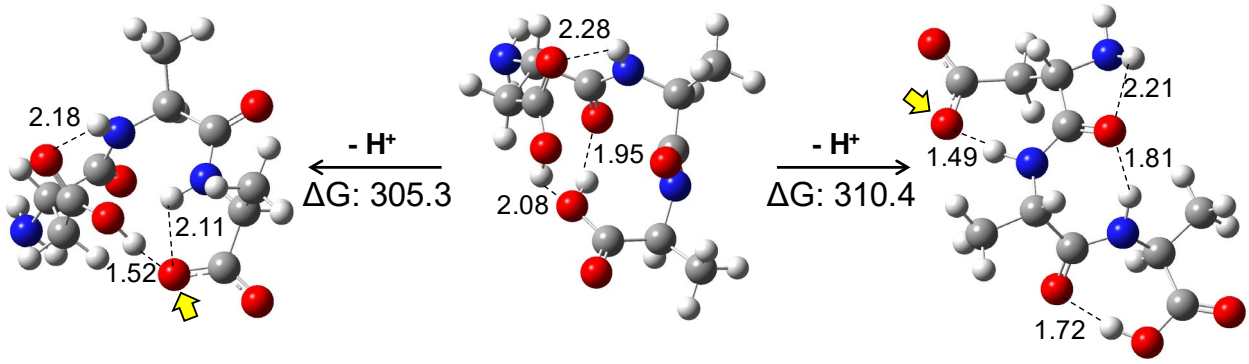


Figure A3.8. Aspartic Acid-Alanine-Alanine, DAA

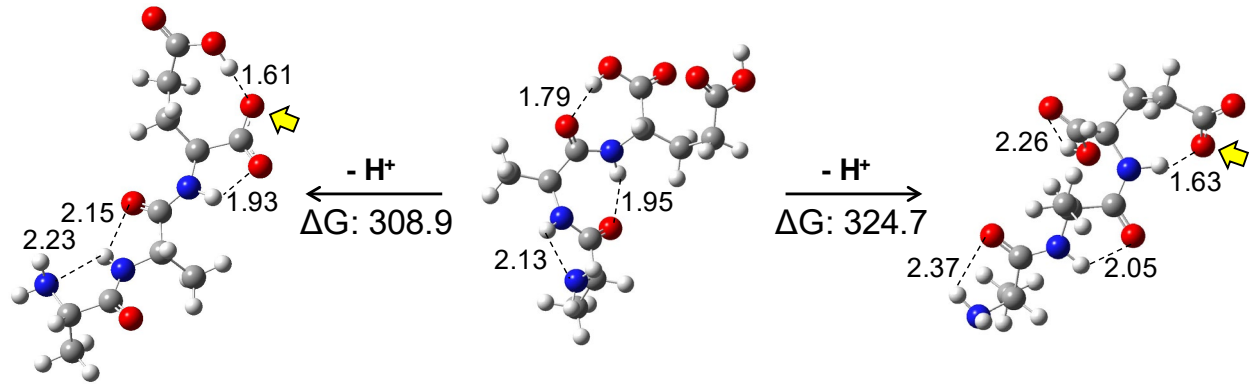


Figure A3.9. Alanine-Alanine-Glutamic Acid, AAE

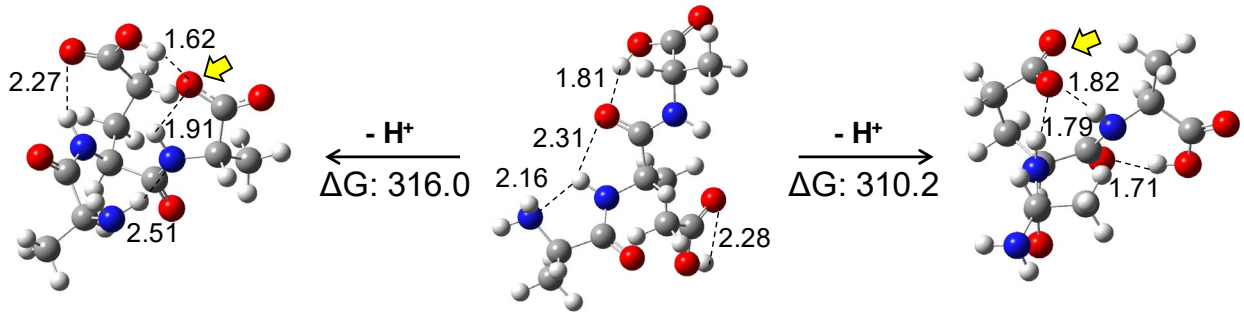


Figure A3.10. Alanine-Glutamic Acid-Alanine, AEA

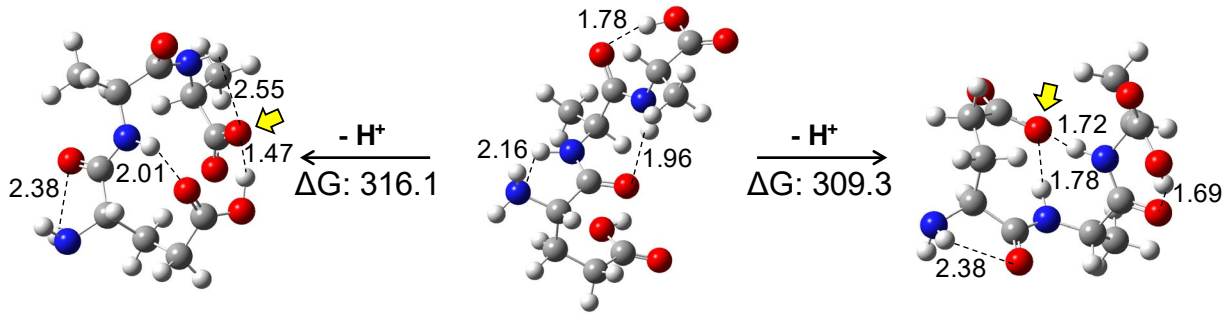


Figure A3.11. Glutamic Acid-Alanine-Alanine, EAA

CHAPTER 4

THE ENERGETICS OF CO_2^- IN AQUEOUS SOLUTION

From McNeill, A. S.; Dixon, D. A. Energetics of CO_2^- in Aqueous Solution. *J. Phys. Chem. A* **2019**, *123*, 1243 – 1259.

Introduction

There is significant interest in the properties of CO_2 in aqueous media. It is well-established that free molecular CO_2 cannot bind an electron as CO_2^- is predicted to be metastable with a potential energy minimum 0.4 eV above the ground state of the neutral CO_2 molecule.[1,2,3,4] In order to spectroscopically measure the properties of this unstable anionic species, early CO_2^- studies were conducted by trapping the anion in various salt matrices, such as alkali halides.[5] Time-of-flight (TOF) mass spectrometry (MS) was used to observe direct attachment of an electron to $(CO_2)_n$ clusters with $2 \leq n \leq 6$ generated from a supersonic nozzle;[6] however, the presence of CO_2^- was not observed in this TOF MS study. This suggests that the CO_2^- anion is only stable on the observable timescale when it is interacting with at least one additional CO_2 molecule. In contrast, the lifetime of the bent radical CO_2^- anion in the gas phase against autodetachment has been reported to be $9.0 \pm 2.0 \times 10^{-5}$ s based on results from a colliding alkali atom beam experiment.[7] When produced by collision ionization of cesium with maleic and succinic anhydrides, the observed lifetime of the dissociation product CO_2^- was found to be $6.0 \pm 1.0 \times 10^{-5}$ s.[8] The presence of even a single H_2O molecule stabilizes the CO_2^- radical anion, increasing its observed lifetime to the ms domain in a comparable TOF MS experiment with an absorption spectrum identical to that of CO_2^- in a crystalline matrix.[9] Recent

experiments and theoretical studies confirm that the CO_2 retains the additional negative charge in clusters of CO_2 with varying numbers of H_2O molecules.[10,11,12,13] The CO_2^- molecule has a bond angle of approximately 134-138° with C-O bond distances of ~1.25 Å, determined both experimentally and theoretically.[1,5,14,15] The one electron reduction potential of aqueous CO_2 has been reported to be -1.90 eV (SHE) from measurements of the equilibrium with the Tl^+/Tl couple with the CO_2 anion generated by pulse radiolysis.[16]

Motivated by the predicted improvement in stability for CO_2^- in the presence of H_2O molecules, particularly in an aqueous environment rather than in the gas phase clustered with explicit H_2O molecules,[12] pulse radiolysis was used in combination with time-resolved resonance Raman spectroscopy to prepare and observe aqueous CO_2^-.[13] The use of this spectroscopic technique was necessary for the aqueous study due to the fact that the CO_2^- radical is reactive and does not exist in sufficient concentration for steady state observation at room temperature. A transient Raman band at 1298 cm^{-1} for the C-O stretch and a 742 cm^{-1} for the O-C-O bend that is Raman forbidden in a linear molecule but expected for a bent CO_2^- with a bond angle of approximately 130° upon the capture of the free electron in solution by the linear neutral CO_2 were observed. The formation of CO_2^- in an aqueous environment is certainly favorable, as shown in a molecular dynamics study in which a neutral CO_2 molecule, upon binding with a solvated electron with six explicit H_2O molecules, accepts the free electron and bends to form a CO_2^- radical anion on the surface of the $(H_2O)_6$ cluster within 20 ps of contact.[11]

Thus, it is of interest to better understand the nature of CO_2^- in aqueous systems, as this is potentially important for a variety of geochemical processes in which the presence of CO_2^-, in addition to carbonate CO_3^{2-} and bicarbonate HCO_3^-, could play a role in the aqueous chemistry of CO_2. Furthermore, there is the possibility that CO_2^- can dimerize to produce value-added organic

compounds, such as the oxalate dianion. The creation of oxalate from CO_2 has been observed experimentally, though only via electrochemical or electrocatalytic methods. For example, CO_2 could be reduced to oxalate selectively at a PVA/[$Ni(dppm)_2Cl_2$]-coated platinum electrode with at least 5 turnovers before the CO_2 in the thin layer was consumed.[17] Binuclear copper complexes have been used to catalytically convert CO_2 to oxalate under mild conditions.[18,19] However, the oxalate product must be chemically removed from the catalytic complex. Also, despite being highly selective for CO_2,[18,19] the reaction proceeds extremely slowly (~0.019 hr^{-1}),[19] which prevents this from being a commercially viable way for converting CO_2 to oxalate.

We have used high level correlated molecular orbital theory in combination with self-consistent reaction field approaches to treat long range solvent effects in our study of the CO_2^- anion in aqueous systems. This allows us to predict the electron affinity of CO_2 in aqueous solution and to draw conclusions about water's role in the energetics related to oxalate formation in an aqueous environment.

Computational Methods

Initial geometries for the majority of the neutral CO_2 clusters with 3 or 4 H_2O molecules were taken from our previous work.[20] The anionic clusters related to these neutral structures were created by adding an electron to the neutral geometries. Additional geometries for neutral and anionic CO_2 clusters with 3 or 4 H_2O molecules were taken from Balaj et al.[10] Other CO_2 clusters with 4 H_2O molecules were created by adding CO_2 either above (in the region of the free electron density) or below the cluster of H_2O molecules from our prior work[21] with a free electron added to 4 H_2O to form $e^-\cdot(H_2O)_4$. CO_2^- clusters with 8 H_2O molecules were created by adding an additional layer of 4 H_2O molecules to **A4-1**, **A4-2**, and **A4-3** or by beginning with $e^-\cdot(H_2O)_8$ cluster geometries from our prior work[21] with a CO_2 molecule added either above (in the

region of free electron density), in the center of, or below the 8 H_2O cluster. All clusters of CO_2 with 3, 4, or 8 H_2O molecules are labelled as **N** for neutral or **A** for anion, followed by a number indicating the number of explicit molecules, and finally labeled in order of increasing gas-phase enthalpy at 298 K using the best CCSD(T) electronic energies.

All reported neutral and anionic species in this work were initially optimized at the density functional theory (DFT) level using the B3LYP[22,23] hybrid exchange-correlation functional and the aug-cc-pVTZ basis set.[24,25] The B3LYP optimized geometries were reoptimized at the MP2[26,27] level with the aug-cc-pVDZ basis set for all species, and the aug-cc-pVTZ basis set[24,25] for all species except those with 8 H_2O molecules. These basis sets are hereafter labelled as aD and aT, respectively. Harmonic vibrational frequencies were calculated to ensure that the optimized structures were energetic minima. The B3LYP/aT geometries were also used as a starting point for composite correlated level G3MP2(B3)[28] calculations for all species. All above calculations were performed using the Gaussian-16 software package.[29]

Coupled cluster R/UCCSD(T)[30,31,32,33,34,35,36] single point calculations were performed using the MP2/aT optimized structures with the aD and aT basis sets for $CO_2 \cdot (H_2O)_y^{x-}$ for x = 0 and 1 with y = 3. For clusters with y = 4 and 8, single point calculations were performed with only the aD basis set due to computational resource restraints. In addition, the explicitly correlated CCSD(T)-F12b method[37,38,39] was used with the equivalent aD and aT basis sets where possible.[40,41] The CCSD(T)-F12 approach uses functions for which the inter-electronic coordinates are explicitly incorporated into the wavefunction, which enables improved performance with a smaller basis set. These are the highest level calculations presented for all species, except for the anionic CO_2 clusters with 8 explicit H_2O molecules, for which the CCSD(T)-F12b calculations were too computationally expensive. For clusters with y = 3, both

aD and aT basis sets were used. For clusters with y = 4, only the aD basis set was used. The zero point energies (ZPEs) and thermal corrections for the enthalpies and entropies were taken from the MP2/aT calculations. All CCSD(T) calculations were performed using the MOLPRO 2015.1 program package.[42,43] For species calculated with both the aD (n = 2) and aT (n = 3) basis sets, the values for the R/UCCSD(T)-F12b calculations were extrapolated to the complete basis set (CBS) limit using Equation (4.1).[44,45]

$$E_n = E_{CBS} + A(n + 1/2)^{-4} \quad (4.1)$$

Free energies of solvation in water at 298 K were calculated for MP2/aT optimized geometries from single point calculations using the self-consistent reaction field (SCRF) approach[46] with the COSMO[47,48] parameters as implemented in Gaussian 16 at the B3LYP/aT level. The aqueous reaction Gibbs free energies were calculated from Equation (4.2)

$$\Delta G_{aq} = \Delta G_{gas} + \Delta G_{solv} \quad (4.2)$$

where ΔG_{gas} is the appropriate gas phase free energy at different computational levels and ΔG_{solv} is the aqueous solvation free energy calculated at the B3LYP/aT/COSMO level. A dielectric constant of 78.39 corresponding to that of bulk water was used in the COSMO calculations. The solvation energy only includes the electrostatic energy component. For comparison, solvation free energies were also calculated with the SMD[49] parameters using Gaussian 16 at the B3LYP/aT level.

The extrapolation to the complete basis set limit and additional small corrections to the oxalic acid gas-phase bond dissociation energy (BDE) were calculated following the Feller-Peterson-Dixon (FPD) approach.[50,51,52,53] We performed additional CCSD(T)-F12b calculations with the aQ basis set to include in the basis set extrapolation using Equation (4.1). Additional corrections include core-valence corrections and scalar relativistic corrections. The core-valence

corrections (ΔE_{CV}) were calculated at the CCSD(T) level with the aug-cc-pwCVTZ basis set[54,55] for C and O atoms as well as both oxalic acid and the carbonyl radical. The scalar relativistic corrections (ΔE_{Rel}) including the pseudopotential corrections were calculated at the second order Douglas-Kroll-Hess (DK) [56,57,58] level with the all-electron aug-cc-pwCVTZ-DK basis sets.[59] The atomic spin-orbit corrections (ΔE_{SO}) for the heats of formation were taken from the experimental values (C = -0.08 kcal/mol and O = -0.22 kcal/mol).[60] The total atomization energies for the heats of formation for oxalic acid and the carbonyl radical are calculated using Equation (4.3).

$$TAE = \Delta E_{CBS} + \Delta E_{SR} + \Delta E_{CV} + \Delta E_{ZPE} + \Delta E_{SO} \quad (4.3)$$

Heats of formation at 0 K were calculated from these TAEs and experimental heats of formation of the H, C, and O atoms at 0 K from the JANAF tables (51.63 ± 0.001 kcal/mol for H, 169.98 ± 0.1 kcal/mol for C, and 58.98 ± 0.02 kcal/mol for O).[61] Heats of formation at 298 K were then calculated by following the Curtiss et al. procedure.[62]

All calculations were performed on local parallel high-performance Xeon and Opteron based Penguin Computing clusters at The University of Alabama and the Xeon and Opteron based SGI UV (Ultraviolet) at the Alabama Supercomputer Center. Molecular visualization was done using the AGUI graphics program from the AMPAC program package.[63]

Results

We used a microsolvation approach in which a select number of explicit water molecules (n = 3, 4, or 8) are included together with the SCRF implicit calculations to predict the energetics of binding the electron to CO_2 in aqueous solution, building on our prior work on the hydrolysis of CO_2 to carbonic acid.[20] The same approach was applied to the lowest energy structures with 3

and 4 explicit water molecules from Balaj et al.,[10] and to the clusters derived from our prior work[21] on the solvated electron involving 4 and 8 explicit H_2O molecules, described above.

Gas-Phase Microsolvated Geometries and Energies of Neutral and Anionic Isomers. The relative gas-phase enthalpies (ΔH_{gas}(298 K)) of the isomers with 3, 4, and 8 H_2O molecules explicitly solvating CO_2 are given in Table 4.1. The neutral cluster with 3 H_2O molecules has three low energy structures. The lowest energy cluster in the gas phase (**N3-1**) was taken from Ref. 20 and has two H_2O molecules solvating the carbon and an oxygen with the third H_2O bridging the first two H_2O molecules so that the CO_2 is essentially bound to the edge of a water trimer (Figure 4.1). One H_2O molecule has a Lewis acid-base interaction with the carbon and the other H_2O molecule hydrogen bonds to one CO_2 oxygen. The next lowest energy neutral isomer of the CO_2 with 3 H_2O (**N3-2**) involves a plane of three H_2O molecules hydrogen bonded to one another with the linear CO_2 bound to the surface of the water cluster (Figure 4.1). The enthalpy of this neutral structure is only 0.6 kcal/mol higher at the CCSD(T)-F12b/CBS level. The highest energy structure (**N3-3**, Figure 4.1) has one H_2O forming a Lewis acid-base adduct with the carbon and the other two H_2O molecules hydrogen bonded to the two oxygen atoms on the CO_2 and to the central H_2O molecule and is 3.5 kcal/mol higher in gas-phase enthalpy than **N3-1**.

The three anionic CO_2 clusters with 3 H_2O molecules (Figure 4.2) have an O-C-O bond angle of 133 to 136°. The structure for the anion with 3 H_2O molecules (**A3-1**) with the lowest enthalpy (ΔH_{gas}(298 K)) in the gas phase is that of Balaj et al.[10] **A3-1** has three H_2O molecules hydrogen bonded to one another with the dangling hydrogens from each H_2O molecule bonding to the oxygens of CO_2. **A3-1** is the lowest energy by 1.2 kcal/mol at the CCSD(T)-F12b/aT level. **A3-1** is similar to what is predicted for anions with other small H_2O clusters with the anion on the surface of the cluster.[64,65] The anionic cluster with the second lowest energy (**A3-2**) has two

water molecules each with a hydrogen bond to each oxygen on the CO_2 for a total of 4 hydrogen bonds to the CO_2. The third H_2O molecule bridges the other two hydrogen bonding to the oxygen atoms on the two H_2O molecules. The electron spin density plot shows that the excess spin for both **A3-1** and **A3-2** is mostly localized on the CO_2 carbon atom and oriented away from the rest of the cluster. Another low energy structure, **A3-3**, has 3 H_2O molecules bonded to the CO_2 and is similar to the **N3-2** isomer of the neutral cluster with one CO_2 oxygen not involved in the hydrogen bonding. In this case, there is no Lewis acid interaction and the CO_2 is also bent with the excess electron spin density on the carbon oriented towards the H atoms in the hydrogen bonds. The energy difference between **A3-1** and **A3-3** is 2.0 kcal/mol at the CCSD(T)-F12b/CBS level.

There are three low enthalpy (ΔH_{gas}(298 K)) structures for the neutral cluster with 4 H_2O molecules, two of which involve essentially adding a water molecule to the lowest energy structure with 3 H_2O molecules. The lowest energy structures (**N4-1** and **N4-2**) were taken from Ref. 20 and both have a dangling oxygen atom on the CO_2 (Figure 4.1). The two lowest energy structures are essentially isoenergetic in the gas phase. **N4-3** has 4 H_2O molecules hydrogen bonded to one another in a plane with the CO_2 bound to the surface of this water cluster and is higher in energy than **N4-1** in the gas phase by 2.7 kcal/mol (see Appendix). The higher energy structures (**N4-4** and **N4-5**) are essentially built by adding a H_2O molecule to the highest energy trimer structure **A3-3** (see Appendix). These two cluster isomers are 4.6 and 4.9 kcal/mol higher in energy respectively than **N4-1**.

The O-C-O bond angles for all anionic clusters with 4 H_2O molecules were 132-137°, with the notable exceptions of **A4-7**, **A4-8**, and **A4-9**, which did not retain the radical electron density on the CO_2 but shared it among the H_2O molecules instead, leaving the CO_2 molecule

approximately linear. The two structures with lowest energy in the gas phase are shown in Figure 4.3, as well as the structure with the lowest energy in aqueous solution, with the remaining structures in the Appendix. The structure for the anion with 4 H_2O molecules (**A4-1**) with the lowest ΔH_{gas}(298 K) has the four H_2O molecules forming a plane in which they are each hydrogen bonded to one another with all four dangling hydrogens forming hydrogen bonds with the oxygens from CO_2. This structure is typical of how anions bind to water clusters.[64,65] **A4-2** is higher in energy than **A4-1** by 2.5 kcal/mol at the CCSD(T)-F12b/aD level. **A4-2** is structurally similar to **A3-3** (least favorable cluster with 3 H_2O molecules) except that the two O atoms on the CO_2 have inequivalent hydrogen bonding. The G3MP2(B3) calculations predicted structures with a dangling CO_2 anion, but these structures are about 20 kcal/mol higher in energy than **A4-1**. The next higher energy structure at the CCSD(T)-F12b/aD level is **A4-3**, which is 5.5 kcal/mol higher in energy than **A4-1**. It is built in part by adding an additional H_2O molecule to **A3-3**. **A4-4** retains the radical electron density on the CO_2 molecule (O-C-O angle of 134.4°) with the H_2O molecules arranged in a plane with the CO_2 molecule in the center of the cluster; the gas-phase enthalpy is 5.8 kcal/mol higher than that of **A4-1**. Each H_2O molecule has one hydrogen bond with an oxygen from the CO_2 molecule and one to a nearby H_2O molecule. Clusters **A4-5** and **A4-6** are 6.6 and 10.4 kcal/mol higher in energy than **A4-1**, respectively. The water molecules of **A4-5** are distributed more sparsely about the CO_2, resulting in a cluster that does not resemble the CO_2 binding to the surface of the water cluster, but being incorporated into the water cluster instead. The geometry of **A4-6** is like that of **A4-2**, though all hydrogen bonding from three of the H_2O molecules occurs with one of the oxygens of the CO_2 while its other oxygen dangles away from the cluster of H_2O molecules. Three higher energy structures, **A4-7**, **A4-8**, and **A4-9**, have the radical electron density localized on the water molecules rather than on the CO_2 and are

18 to 22 kcal/mol higher in energy than the ones with the electron localized on the CO_2 at the CCSD(T)-F12b/aD level of theory.

The three lowest energy structures of neutral clusters of CO_2 with 8 H_2O molecules have gas-phase enthalpies within 2 kcal/mol of one another at the CCSD(T)-F12b/aD level (Figure 4.4). The two lowest in energy, **N8-1** and **N8-2**, are virtually indistinguishable with **N8-2** only 0.8 kcal/mol higher in gas-phase enthalpy than **N8-1**. The geometries for these clusters are extremely similar to one another, with the linear CO_2 molecule bound to one vertex of the cube formed by the 8 H_2O molecules, which maximizes the number of hydrogen bonds formed between the explicit H_2O molecules. There are an equivalent number of hydrogen bonds in **N8-1** and **N8-2**, although the hydrogen bonding distances between the H_2O molecules and the linear CO_2 molecule for each cluster are slightly shorter in **N8-1**. The next lowest energy neutral structure in the gas phase, **N8-3**, (Figure 4.4) is based on the addition of linear CO_2 to the most favorable 8 H_2O cluster from our previous work.[21] **N8-3** is 3.4 kcal/mol higher in energy than **N8-1**. The positions of the explicit water molecules in the **N8-3** structure are much less ordered than those in **N8-1** and **N8-2**, and there are only two hydrogen bonds between the H_2O molecules and the oxygen atoms of the linear CO_2 rather than three hydrogen bonds as in **N8-1** and **N8-2**. **N8-4** (Figure 4.4), **N8-5**, and **N8-6** are approximately 5.4, 8.2, and 12.1 kcal/mol higher in energy than **N8-1**, respectively. **N8-4** is structurally similar to **N8-2** and **N8-1** in that the H_2O molecules form a cube to maximize H_2O-H_2O hydrogen bonding, but the CO_2 molecule is much farther from the water molecules and only shares one hydrogen bond with the nearest H_2O. **N8-5** and **N8-6** are much less ordered, like **N8-3**, but the H_2O-H_2O hydrogen bonds are more interrupted by the insertion of the linear CO_2 molecule in these clusters. Images of **N8-5** and **N8-6** are given in the Appendix.

The three lowest energy gas-phase anionic clusters with 8 H_2O also have very similar gas phase enthalpies at the CCSD(T)/aD level (Figure 4.5). For anionic clusters of CO_2 with 8 H_2O molecules, all clusters for which the CO_2 retains the radical electron density have an O-C-O bond angle of approximately 133 to 136°. **A8-1** is the lowest energy structure in the gas phase with 8 H_2O molecules arranged in a cube that maximizes the hydrogen bonds between the H_2O molecules. The four H_2O molecules on the face of the cube bound to the bent CO_2^- each have a hydrogen bond with one of the oxygen atoms from the CO_2^-. The geometry of **A8-2** (Figure 4.5) is very similar to that of **A8-1** except that the CO_2^- is bound to a vertex of the 8 H_2O cube and therefore only forms two hydrogen bonds between the oxygen atoms of CO_2^- and the explicit H_2O molecules; **A8-2** is 1.1 kcal/mol higher in enthalpy than **A8-1**. The **A8-3** geometry (Figure 4.5) is markedly less ordered than those of **A8-1** and **A8-2** as it is based on the geometry of the lowest energy $(H_2O)_8^-$ structure previously reported[21] with the carbon dioxide added to the center of that original cluster between the two layers of H_2O molecules. Despite this less ordered cluster geometry, **A8-3** is only 2.6 kcal/mol higher in energy than **A8-1**. **A8-4** (Appendix) and **A8-5** (Figure 4.5) are only 3.4 and 4.3 kcal/mol higher in energy than **A8-1**. The structures of **A8-4** and **A8-5** are similar to those described above except that the bent CO_2 somewhat penetrates the clustered H_2O molecules, preventing some H_2O-H_2O hydrogen bonds that exist for the lower energy clusters. **A8-6** (Appendix) is the only anionic cluster of CO_2 with 8 H_2O molecules that does not have the radical electron density on the CO_2. Instead the electron is localized on the H_2O molecules and the O-C-O bond angle is approximately linear at 177.2°. **A8-6** is 23.1 kcal/mol higher in energy than **A8-1**, which is similar to the relative energies for CO_2 clusters with 4 H_2O molecules, **A4-7**, **A4-8**, and **A4-9**, that retain the radical electron density among the water molecules.

Aqueous Energies of Neutral and Anionic Clusters. The relative energetics (ΔG_{aq}(298 K)) of the CO_2 with 3, 4, and 8 explicit H_2O molecules vary in the aqueous phase as compared to the gas phase (Table 4.2). The role of geometry optimization in the presence of an implicit solvent at the DFT level on the relative and overall energetics is small and discussed in the Appendix. Among the neutral clusters with 3 explicit H_2O molecules, **N3-1** is the lowest energy isomer in both the gas and aqueous phase. **N3-3** in solution is higher in free energy (ΔG_{aq}(298 K)) by 0.9 kcal/mol than **N3-1** at the CCSD(T)-F12b/CBS/COSMO level. In contrast, **N3-3** is lower in free energy than **N3-1** by 0.6 kcal/mol at the CCSD(T)-F12b/CBS/SMD level. Clearly both of these clusters can serve as model for hydrated CO_2 in aqueous solution. **N3-2** is higher in energy by approximately 4 kcal/mol than **N3-1** and **N3-3** in aqueous solution at the CCSD(T)-F12b/CBS/COSMO and SMD levels, whereas this structure has a nearly identical energy to that of **N3-1** in the gas phase.

A3-2 is the lowest energy anionic structure with 3 explicit H_2O molecules in aqueous solution at the CCSD(T)-F12b/CBS/COSMO and SMD levels. **A3-3** is higher in energy than **A3-2** by 1.6 kcal/mol at the CCSD(T)-F12b/CBS/COSMO level and by 1.0 kcal/mol at the CCSD(T)-F12b/CBS/SMD level. Unlike in the gas phase, **A3-1** is the highest energy structure of the three anionic conformers with 3 explicit H_2O molecules modeled in aqueous solution by 2.4 kcal/mol at the CCSD(T)-F12b/CBS/COSMO level and by 2.5 kcal/mol at the CCSD(T)-F12b/CBS/SMD level. The use of other electronic structure methods does not greatly affect the energy differences between the neutral and anionic clusters with 3 explicit H_2O molecules or their relative ordering.

The two lowest energy (ΔG_{aq}(298 K)) neutral clusters with 4 H_2O molecules in aqueous solution (**N4-1** (lowest energy) and **N4-2**) at the CCSD(T)-F12b/CBS level with both differ by

0.8 kcal/mol at the CCSD(T)-F12b/CBS/COSMO level and by 0.2 kcal/mol at the CCSD(T)-F12b/CBS/SMD level. **N4-5** and **N4-4** are 2.7 and 3.3 kcal/mol higher in energy at the CCSD(T)-F12b/CBS/COSMO, respectively, relative to **N4-1**, but the energy differences are smaller at the CCSD(T)-F12b/CBS/SMD level, 1.5 and 1.6 kcal/mol, respectively. **N4-3** is the highest energy neutral cluster with 4 explicit H_2O molecules in the aqueous phase by 6.7 kcal/mol at the CCSD(T)-F12b/CBS/COSMO level and 4.2 kcal/mol at the CCSD(T)-F12b/CBS/SMD level.

In aqueous solution at the CCSD(T)-F12b/aD/COSMO level, **A4-4** is the anionic cluster with 4 explicit H_2O molecules with the lowest aqueous free energy (ΔG_{aq}(298 K)) value. **A4-3** and **A4-5** are 3.6 and 3.7 kcal/mol higher in energy than **A4-4**, respectively, at the CCSD(T)-F12b/aD/COSMO level. **A4-2**, **A4-1**, and **A4-6** have relative aqueous free energies of 4.4, 7.9, and 12.4 kcal/mol, respectively, at the CCSD(T)-F12b/aD/COSMO level with respect to **A4-4**. However, at the CCSD(T)-F12b/aD/SMD level, **A4-3** is the lowest energy structure. **A4-4** and **A4-5** are higher in energy by 1.5 and 1.0 kcal/mol, respectively, compared to that of **A4-3** at the CCSD(T)-F12b/aD/SMD level. **A4-2**, **A4-1**, and **A4-6**, with respect to **A4-3**, are higher in energy by 2.5, 5.2, and 10.4 kcal/mol, respectively, at the CCSD(T)-F12b/aD/SMD level. **A4-7**, **A4-8**, and **A4-9**, for which the additional electron density is localized on the H_2O molecules rather than on the CO_2, are significantly higher in energy as compared to **A4-4** (COSMO) or **A4-3** (SMD) by 24 to 37 kcal/mol in the aqueous phase at the CCSD(T)-F12b/aD/COSMO and SMD levels.

In aqueous solution, all neutral clusters of CO_2 with 8 H_2O molecules are within 2.5 kcal/mol (ΔG_{aq}(298 K)) of the lowest energy structure with the exception of **N8-5** at the CCSD(T)-F12b/aD/COSMO level and are within 5 kcal/mol of the lowest energy cluster at the CCSD(T)/aD/SMD level. Due to the similarity in energetics for these clusters, the identity of the

lowest energy neutral cluster with 8 explicit H_2O molecules is dependent on the both the electronic structure method and the solvation model. The lowest energy neutral cluster with 8 explicit H_2O molecules in aqueous solution is **N8-4** at the CCSD(T)-F12b/aD/COSMO level and is **N8-6** at the CCSD(T)-F12b/aD/SMD level. **N8-4** has a very similar geometry to that of **N8-1** and **N8-2**, except that the CO_2 in **N8-4** has only one hydrogen bond with one of the explicit H_2O molecules. Despite the similarities in geometry, **N8-4** is lower in energy than **N8-1** and **N8-2** by 1.8 and 0.7 kcal/mol, respectively, in aqueous solution at the CCSD(T)-F12b/aD/COSMO level. **N8-6** is the most ordered cluster resulting from adding linear CO_2 to the 8 H_2O geometry from our previous work[21] and has a relative aqueous free energy of 1.7 kcal/mol with respect to **N8-4** at the CCSD(T)-F12b/aD/COSMO level. The less ordered **N8-3** geometry is 2.5 kcal/mol higher than **N8-4** in aqueous solution at the CCSD(T)-F12b/aD/COSMO level. **N8-5** is the highest energy neutral cluster with 8 H_2O at 4.1 kcal/mol higher than **N8-4** at the CCSD(T)-F12b/aD/COSMO level. At the CCSD(T)-F12b/aD/SMD level, all six neutral clusters with 8 explicit H_2O molecules had aqueous free energies within 2 kcal/mol of the lowest energy cluster (**N8-6**), showing that all may be reasonable models for solvated CO_2.

A8-3 is the lowest energy (ΔG_{aq}(298 K)) anionic CO_2 cluster with 8 explicit H_2O molecules in aqueous solution at the CCSD(T)/aD/COSMO and SMD levels. The geometry of this cluster is very similar to that of **A8-5** with the CO_2 positioned on the surface of the cluster rather than incorporated into the cluster. **A8-5** has the same energy as **A8-3** at the CCSD(T)/aD/COSMO level and is 0.5 kcal/mol higher in energy at the CCSD(T)/aD/SMD level. Other low energy clusters are **A8-2** (1.4 kcal/mol at the CCSD(T)/aD/COSMO level, 3.1 kcal/mol at the CCSD(T)/aD/SMD level), **A8-4** (1.7 kcal/mol with at the CCSD(T)/aD/COSMO level, 1.1 kcal/mol at the CCSD(T)/aD/SMD level), and **A8-1** (2.6 kcal/mol at both the

CCSD(T)/aD/COSMO and SMD levels). The energy for **A8-6** is significantly higher (25 to 30 kcal/mol) due to the radical electron density being shared among H_2O molecules rather than being localized on the CO_2 molecule, as found for **A4-7**, **A4-8**, and **A4-9**.

Electron Affinities. The adiabatic electron affinities (EAs) are defined as the energy difference between the lowest energy isomer of the neutral and the lowest energy isomer of the anion. We provide these values as free energies at 298 K because of the solvation energy term (Table 4.3). The adiabatic gas phase EA for CO_2 with 3 H_2O molecules (**N3-1** + e^- → **A3-1**) at the CCSD(T)-F12b/CBS level is ΔH_{gas}(298 K) = 11.7 kcal/mol (0.51 eV). The EA increases substantially when the effects of aqueous solvation are included. In the aqueous phase, the adiabatic EA (ΔG_{aq}(298 K)) for **N3-1** + e^- → **A3-2** is 52.5 kcal/mol (2.28 eV) at the CCSD(T)-F12b/CBS/COSMO and 54.0 kcal/mol (2.34 eV) at the CCSD(T)-F12b/CBS/SMD level, respectively. Although **N3-1** is the lowest energy neutral cluster with 3 H_2O molecules at the CCSD(T)-F12b/CBS/COSMO level, **N3-3** is lowest in energy at the CCSD(T)-F12b/CBS/SMD level. Thus, we additionally report that the adiabatic EA (ΔG_{aq}(298 K)) for **N3-3** + e^- → **A3-2** is 53.4 kcal/mol (2.32 eV) at the CCSD(T)-F12b/CBS/COSMO and 54.0 kcal/mol (2.34 eV) at the CCSD(T)-F12b/CBS/SMD level, respectively.

The addition of a fourth explicit H_2O molecule to the clusters does not make a large difference in either the gas phase or aqueous solution phase EAs. These adiabatic EA values increase relative to those of the clusters with 3 H_2O molecules by 1.9 kcal/mol in the gas phase (**N4-1** + e^- → **A4-1**) at the CCSD(T)-F12b/aD level. This results in a gas-phase adiabatic EA (ΔH_{gas}(298 K)) of 12.4 kcal/mol (0.54 eV). In aqueous solution, the lowest energy anionic cluster with 4 H_2O molecules is **A4-4** at the CCSD(T)-F12b/aD/COSMO level and is **A4-3** at the CCSD(T)-F12b/aD/SMD level. The adiabatic EA for **N4-1** + e^- → **A4-4** is 54.6 kcal/mol (2.37

eV) at the CCSD(T)-F12b/aD/COSMO level and 52.6 kcal/mol (2.28 eV) in the aqueous phase at the CCSD(T)-F12b/aD/SMD level. For **N4-1** + e^- → **A4-3**, the aqueous adiabatic EA is 51.0 kcal/mol (2.21 eV) at the CCSD(T)-F12b/aD/COSMO level and 54.1 kcal/mol (2.35 eV) at the CCSD(T)-F12b/aD/SMD level. These aqueous adiabatic EA values, which are between 51 and 55 kcal/mol, are similar to the values for clusters with 3 H_2O, which have aqueous adiabatic values between 52 and 54 kcal/mol.

Increasing the number of explicit H_2O molecules to 8 significantly increases the adiabatic electron affinity of the cluster in the gas phase to 19.7 kcal/mol (0.85 eV) at the CCSD(T)/aD level for **N8-1** + e^- → **A8-1**. In aqueous solution at the CCSD(T)/aD/COSMO level, **N8-4** is the lowest energy neutral cluster with 8 H_2O molecules. **A8-3** and **A8-5** are both energetically identical at the CCSD(T)/aD/COSMO level and are the lowest energy clusters of all anionic clusters with 8 H_2O molecules. For **N8-4** + e^- → **A8-3**, the aqueous adiabatic EA is 52.8 kcal/mol (2.29 eV) at the CCSD(T)/aD/COSMO level and 55.3 kcal/mol (2.40 eV) at the CCSD(T)/aD/SMD level. **N8-4** + e^- → **A8-5** has an adiabatic EA of 52.9 kcal/mol (2.30 eV) at the CCSD(T)/aD/COSMO level and 54.9 kcal/mol (2.38 eV) at the CCSD(T)/aD/SMD level. However, in aqueous solution at the CCSD(T)/aD/SMD level, **N8-6** is the lowest energy neutral cluster with 8 H_2O molecules. For **N8-6** + e^- → **A8-3** in aqueous solution, the adiabatic EA is 54.3 kcal/mol (2.35 eV) at the CCSD(T)/aD/COSMO level and 54.8 kcal/mol (2.38 eV) at the CCSD(T)/aD/SMD level. In aqueous solution, the adiabatic EA of **N8-6** + e^- → **A8-5** is 54.3 kcal/mol (2.35 eV) at the CCSD(T)/aD/COSMO level and 54.4 kcal/mol (2.36 eV) at the CCSD(T)/aD/SMD level. All of these aqueous adiabatic EA values fall within the 52 to 55 kcal/mol range. Thus, the adiabatic EA results in aqueous solution are essentially converged

within 1 kcal/mol even with only 3 H_2O molecules with each solvation model. Therefore, we can conclude that the adiabatic electron affinity of CO_2 in aqueous solution is 2.35 ± 0.08 eV.

Vertical detachment (VDE) and attachment (VAE) energies for lowest-energy clusters of CO_2 with 0, 3, 4, or 8 explicit H_2O molecules (Tables 4.4 and 4.5, respectively) are given in aqueous solution for comparison to the adiabatic EA values. These energies represent the process of the anionic/neutral structure losing/gaining the electron without allowing the geometry to change in order to accommodate the loss/gain of the free electron. Thus, a VDE is positive (requires energy) whereas a VAE may be positive or negative. We use the electron affinity sign convention (negative of the energy to bind an electron) for the VAE where a negative VAE means that the electron will not bind in a vertical attachment process, and a positive VAE means that the electron will bind in the vertical process. We report the aqueous VAE and VDE values obtained using both the COSMO and SMD solvation models. For the VDEs, COSMO and SMD model values are in good agreement, differing by 0.7 to 2.6 kcal/mol. The COSMO model VAEs tend to be larger in magnitude than SMD model VAEs by 2.3 to 6.3 kcal/mol; the trends with both COSMO and SMD models are similar, so only COSMO values will be discussed.

The VDEs (ΔE_{elec}), given in Table 4.5, are significantly larger in magnitude than the adiabatic EAs due to the fact that the neutral CO_2 of the product structures are bent, which significantly destabilizes the neutral CO_2 molecule by 35 kcal/mol. The gas phase values for these VDE processes are given in the Appendix. The aqueous VDE for **A3-2** is 103.8 kcal/mol (4.50 eV) at the CCSD(T)-F12b/aT/COSMO level, and the aqueous VDEs for **A4-3** and **A4-4** are 105.5 and 100.7 kcal/mol (4.58 and 4.37 eV) at the CCSD(T)-F12b/aD/COSMO level, respectively. The aqueous VDEs for **A8-3** and **A8-5** are 110.5 and 103.5 kcal/mol (4.79 and 4.49 eV) at the CCSD(T)/aD/COSMO level, respectively. For **A3-1**, the VDE at the CCSD(T)-

F12b/aD/COSMO and SMD levels (the best calculations available for this cluster) are ~5 kcal/mol higher in energy than the VDE for this cluster at the CCSD(T)/aD/COSMO and SMD levels. We can then use 5 kcal/mol as a correction factor to the CCSD(T)/aD/COSMO and CCSD(T)/aD/SMD calculated VDEs for **A8-3** and **A8-5** to better estimate that the VDEs for **A8-3** and **A8-5** are approximately 105 to 115 kcal/mol. The final vertical state results in a CO_2 bent to 135°. Bending from 180° to 135° in the gas phase requires 35.0 kcal/mol at the CCSD(T)-F12b/CBS level. Subtraction of this value from the VDE gives 70 to 80 kcal/mol, showing that there is additional reorganization of the H_2O and CO_2 molecules on detachment of the additional electron on the anion resulting in an additional ~ 15 to 25 kcal/mol of stabilization.

In the gas phase, all of the vertical attachment energies (VAEs; $-\Delta E_{elec}$) for CO_2 with up to 8 explicit H_2O molecules are negative except for **N8-4** at the CCSD(T)/aT level where it is barely bound by 0.1 kcal/mol (Appendix). In aqueous solution, addition of an electron to the linear CO_2 molecule without allowing the CO_2 to bend out of its ~180° geometry has positive VAEs at all computed levels for CO_2 with 3, 4, and 8 explicit H_2O molecules, so the solution enables binding of an electron to the CO_2/H_2O cluster even without bending. The electron density images for the VAE products of **N3-3**, **N4-4**, and **N8-4** are given in the Appendix. Examination of the electron spin density for these clusters confirms that the radical electron density is not retained on the linear CO_2 but is found on the explicit H_2O molecules instead. The VAEs for CO_2 with 3 H_2O molecules in aqueous solution are 19.8 kcal/mol (0.86 eV) for **N3-1** and 18.6 kcal/mol (0.81 eV) for **N3-3** at the CCSD(T)-F12b/aT/COSMO level. The VAE decreases to 12.0 kcal/mol (0.52 eV) for **N4-4** at the CCSD(T)-F12b/aD/COSMO level and increases to 17.1 kcal/mol (0.74 eV) for **N8-4** and 14.4 kcal/mol (0.62 eV) for **N8-6** at the CCSD(T)/aD/COSMO level. As this computational level is likely to give VAEs that are too low,

we can estimate that the VAE for CO_2 in aqueous solution is about 21 ± 3 kcal/mol. Bending the CO_2 molecule to 135° requires 35 kcal/mol in the gas phase which is a destabilization term. Given an adiabatic electron affinity of 54 kcal/mol, we can estimate that the additional solvation effect is ~ 68 kcal/mol ($\Delta E_{solvation}$ ~ (54-21+35 kcal/mol) to move the electron from the H_2O to generate solvated CO_2^-.

Reduction Potential of CO_2. In order to convert the adiabatic electron affinity into a reduction potential, the absolute potential of the standard hydrogen electrode (SHE) is needed. Although IUPAC recommends a value of -4.44 eV,[66] we prefer to use a more recent value of -4.28 eV.[67,68] We can write reaction (4.4) to get the E^0 for CO_2/CO_2^- relative to SHE. This reaction energy then

$$H^+(aq) + CO_2^-(aq) \rightarrow CO_2(aq) + ½\ H_2(g) \qquad (4.4)$$

corresponds to E^0 = (-4.28 + 2.35 eV) = -1.93 ± 0.08 eV. Following Marenich et al.,[67] we note that we can use the adiabatic electron affinity as the conversion from a standard state of 1 bar to 1 atm is small (<0.04 kJ/mol) and change from 1 molar to 1 molal is also small and can be neglected. Our value for E^0 is in excellent agreement with the reported value of -1.90 eV.[16] Camaioni et al.[69] developed a protocol to predict reduction potentials by varying the radii of the solvation cavity in a self-consistent reaction field approach in an iterative fashion using COSMO and CHELPG charges.[70,71,72,73] They predicted a reduction potential for CO_2(g)/CO_2^-(aq) of 2.00 eV at the G3B3 level,[74] which corresponds to -1.93 eV for CO_2(aq)/CO_2^-(aq) and is in excellent agreement with our result. The corresponding B3LYP/aug-cc-pVTZ value for CO_2(aq)/CO_2^-(aq) is -1.67 eV with their approach.

Vibrational Frequencies. The C-O symmetric stretching frequencies are given in Table 4.6 as a potential diagnostic tool. The C-O stretching frequency for neutral gas-phase CO_2 is 1326 cm^{-1} at the MP2/aT level as compared to an experimental value of 1333 cm^{-1}.[75] These values can be

compared to the symmetric C-O stretches of the neutral $CO_2 \cdot (H_2O)_x$ (x = 3, 4, and 8) clusters. Most of the neutral CO_2 and H_2O clusters have C-O symmetric stretches slightly blue-shifted in the presence of 3 or 4 water molecules by 3 to 8 cm^{-1} compared to the calculated C-O symmetric stretch of gaseous CO_2. The neutral clusters based on the Balaj et al.[10] geometries, **N3-2** and **N4-3**, have C-O symmetric stretches that are red-shifted by ~40 cm^{-1} compared to the predicted symmetric C-O stretch of CO_2. **N8-1**, **N8-2**, **N8-3**, **N8-4**, **N8-5,** and **N8-6** have symmetric C-O stretch frequencies that are red-shifted by less than 20 cm^{-1} compared to gaseous CO_2.

When an extra electron is added to CO_2 to form the CO_2^- anion, the molecule becomes bent at an angle of approximately 135° and the calculated symmetric stretch (1203.7 cm^{-1} at the MP2/aT level) of the anion red-shifts by 122 cm^{-1} compared to that of the neutral CO_2, which is similar to a previously reported calculated value of 1218 cm^{-1} obtained via DFT[10] and in exact agreement with a previously reported calculated value at the same MP2/aT level of theory.[76] The experimental value for the C-O stretch of CO_2^- in a Ne matrix is reported to be 1254 cm^{-1}, but the matrix also contained other ions and trace amounts of H_2O,[77] which may have a notable effect on this frequency. The vibrational stretches of CO_2^- are similar to those of the CO_2^- moiety in HCO_3^-, which has a symmetric C-O stretching frequency of 1293.6 cm^{-1} at the MP2/aT level of theory in the gas phase. The corresponding experimental value for the symmetric C-O stretch of HCO_3^- is 1355 cm^{-1} in aqueous solution.[78] This aqueous symmetric stretch value is blue-shifted by ~ 60 cm^{-1} as compared to the calculated gas phase value of this molecule due to the effect of solvent on the stretching frequencies.[80] Most anionic CO_2-water clusters have calculated symmetric C-O stretch values below 1300 cm^{-1}, and are very similar to previous DFT-calculated gas-phase values of the C-O stretch for $CO_2^- \cdot (H_2O)_4$ of 1294 cm^{-1} and 1298 cm^{-1} for $CO_2^- \cdot (H_2O)_5$.[13] **A4-7**, **A4-8**, and **A4-9** are the exceptions to this because the radical electron density in these clusters is

shared between the water molecules rather than on the CO_2. Compared to the calculated symmetric C-O stretch frequency for neutral CO_2, the symmetric stretches in **A4-7**, **A4-8**, and **A4-9** are blue-shifted by only 3, 7, and 11 cm^{-1}, respectively, which is consistent with a neutral CO_2. All other anionic clusters have symmetric C-O stretches that closely resemble the calculated value of the CO_2^- moiety of HCO_3^- in the gas phase. The C-O symmetric stretches in the anionic clusters are highly similar to the calculated value for the C-O symmetric stretch of the CO_2^- moiety of HCO_3^-. The C-O symmetric stretches for **A3-1**, **A3-2**, **A3-3**, **A4-1**, **A4-2**, **A4-3**, **A4-4**, **A4-5** and **A4-6** vary from this value by less than 35 cm^{-1}. However, the symmetric C-O stretch of **A8-6** is red-shifted compared to that of neutral gaseous CO_2 by 16.9 cm^{-1}, supporting the presence of neutral CO_2 in this cluster with the radical electron density shared among the H_2O molecules.

The C-O asymmetric stretching frequency for neutral gas-phase CO_2 is 2401 cm^{-1} at the MP2/aT level as compared to an experimental value of 2349 cm^{-1}.[75,79] The calculated asymmetric C-O stretch values for all neutral CO_2 clusters with 3 or 4 H_2O molecules based on the Nguyen et al.[20] geometries are blue-shifted by less than 10 cm^{-1} compared to the predicted asymmetric C-O stretch of CO_2. The clusters based on Balaj et al.[10] geometries, **N3-2** and **N4-3**, have asymmetric

C-O stretches that are red-shifted by ~75 cm^{-1}. Gas phase anionic CO_2^- has an asymmetric C-O stretch (1746.7 cm^{-1}) that is red-shifted by ~650 cm^{-1} from the gas phase neutral at the MP2/aT level. The CO_2^- moiety of gas phase HCO_3^- has a calculated asymmetric C-O stretch (1763.0 cm^{-1}) that is red-shifted by ~640 cm^{-1} compared to that of neutral CO_2. This value is only 10 cm^{-1} blue-shifted compared to the value for the calculated asymmetric C-O stretch of the gaseous CO_2^- molecule and is therefore similar enough to be used for comparison to the analogous

asymmetric C-O stretches in our reported anionic clusters. The experimental value for the asymmetric C-O stretch of the HCO_3^- is 1620 cm^{-1} in aqueous solution,[78] which differs substantially from the calculated gas-phase values due to the effect of solvent on the stretches of this molecule. All the anionic clusters with 4 H_2O molecules except **A4-7**, **A4-8**, and **A4-9** have asymmetric C-O stretch values between 1600 cm^{-1} and 1800 cm^{-1} which are similar to the calculated value for the asymmetric stretch of the CO_2^- moiety of HCO_3^-. The anionic clusters **A4-7**, **A4-8**, and **A4-9**, which retain the radical electron density on the H_2O molecules in the anionic cluster rather than on the CO_2, are red-shifted in the asymmetric C-O stretch from 0 to 16 cm^{-1} compared to the neutral CO_2 gas phase molecule. Several of the CO_2 anionic clusters with 4 H_2O molecules have calculated asymmetric C-O stretches that are coupled with the H_2O bending modes. Previous work by Rustad et al. found that the coupling of the vibrational modes between the solute and solvent significantly affect the vibrational frequencies for HCO_3^-,[80] so it is not surprising that the modes for CO_2^- are similarly affected by the presence of H_2O molecules. Anionic clusters **A3-1**, **A3-2**, **A3-3**, **A4-1**, **A4-2**, **A4-3**, **A4-4**, and **A4-6** have asymmetric C-O stretch frequencies that are red-shifted by 4 to 154 cm^{-1} as compared to the C-O asymmetric stretch of the CO_2^- moiety of HCO_3^- in the gas phase. Anionic cluster **A4-5** has an asymmetric C-O stretch frequency that is blue-shifted by 6 cm^{-1}. The anionic clusters **A4-7**, **A4-8**, **A4-9**, and **A8-6**, which retain the radical electron density among the H_2O molecules rather than on the CO_2 molecule, have symmetric CO_2 stretches that are incredibly similar to those of neutral gaseous CO_2. The asymmetric C-O stretch is of **A8-6** is red-shifted by 20.2 cm^{-1} compared to that of neutral CO_2. This is similar to the differences between gaseous CO_2 and **N8-1**, **N8-2**, **N8-3**, **N8-4**, **N8-5,** and **N8-6**, for which the asymmetric stretch is red-shifted by 1 to 18 cm^{-1}. The asymmetric C-O stretches for **A8-1**, **A8-2**, **A8-3**, **A8-4**, and **A8-5** are also similar to the

analogous calculated C-O stretches for the CO_2^- moiety of HCO_3^-. With the exception of **A8-3**, all of the anionic clusters with 8 H_2O molecules have asymmetric C-O stretches that are coupled with bending frequencies of the H_2O molecules, which can result in multiple frequencies with a component of the asymmetric C-O stretch. The asymmetric C-O stretches for the five other anionic clusters with 8 H_2O molecules (**A8-1**, **A8-2**, **A8-3**, **A8-4**, and **A8-5**) are similar to those of the CO_2^- moiety of HCO_3^- (red-shifted by 26 to 124 cm^{-1}).

Formation of Oxalate $[O_2CCO_2]^{2-}$. The following C-C bond dissociation energy (BDE) reactions ((4.5) – (4.9)) were explored to examine whether the CO_2^- anion dimerize in aqueous solution to form the oxalate dianion (Table 4.7). If the BDE is positive, then the compound will be stable to dissociation.

$$HO(O)C\text{-}C(O)OH \rightarrow 2\ \cdot C(O)OH \quad (4.5)$$

$$HO(O)C\text{-}C(O)O^- \rightarrow \cdot C(O)OH + C(O)O^- \quad (4.6)$$

$$^-O(O)C\text{-}C(O)O^- \rightarrow 2\ C(O)O^- \quad (4.7)$$

$$(H_2O)_3\cdot^-O(O)C\text{-}C(O)O^-\cdot(H_2O)_3 \rightarrow 2\ C(O)O^-\cdot(H_2O)_3 \quad (4.8)$$

$$(H_2O)_4\cdot^-O(O)C\text{-}C(O)O^-\cdot(H_2O)_4 \rightarrow 2\ C(O)O^-\cdot(H_2O)_4 \quad (4.9)$$

Reaction (4.5) yields the C-C bond dissociation energy (BDE) for oxalic acid, the most favorable conformer of which maximizes the number of hydrogen bonds by having the two OH groups trans from one another with the hydrogen pointing toward the =O of the opposite carbon atom (the CCOH dihedral angle is 0º). Images of all oxalic conformers and their relative energetics are given in the Appendix. The calculated ΔH_f(298 K) of oxalic acid at the FPD level is -176.5 kcal/mol, which is in reasonable agreement with the reported experimental ΔH_f(298 K) value of -173.0 ± 1.1 kcal/mol.[81] The ΔH_f(298 K) of the trans carboxyl radical HOCO is -44.4 kcal/mol at the FPD level. This is consistent with the previous FPD value of -44.9 ± 0.5 kcal/mol obtained using a slightly different approach.[82] Both values are in good agreement with the

recently summarized experimental value of -44.1 ± 0.4 kcal/mol.[83] The calculated gas-phase BDE (ΔH_{gas}(298 K)) for oxalic acid at the Feller-Peterson-Dixon (FPD) level is 87.7 kcal/mol. (The additional corrections to the CCSD(T)-F12b/CBS BDE of 87.4 kcal/mol are only 0.3 kcal/mol.) This value is similar to the literature experimental value for the C-C BDE for C_2H_6 (90.2 ± 0.2 kcal/mol).[84,85] The gas-phase C-C bond dissociation free energy (BDFE) at 298 K for the C-C bond dissociation of oxalic acid is 73.6 kcal/mol at the CCSD(T)-F12b/CBS level. The aqueous BDFEs (ΔG_{aq}(298 K)) are 69.9 kcal/mol at the CCSD(T)-F12b/CBS/COSMO level and 75.5 kcal/mol at the CCSD(T)-F12b/CBS/SMD level. Thus, the COSMO approach predicts that the products are favored over the reactant in solution by -3.7 kcal/mol (ΔG_{solv}(298 K)), and the SMD approach predicts that the reactant is favored in solution over the products by 1.9 kcal/mol.

If a proton is removed from one of the C(OH) groups to create the hydrogen oxalate monoanion before C-C bond dissociation (reaction (4.6)), the gas-phase C-C BDE is reduced by approximately 0.6 kcal/mol at the CCSD(T)-F12b/CBS level compared to that of oxalic acid. The difference between the gas-phase BDFE of hydrogen oxalate, 70.3 kcal/mol at the CCSD(T)-F12b/CBS level, and the aqueous solution C-C BDFEs, 75.8 and 70.2 kcal/mol at the CCSD(T)-F12b/CBS level with COSMO and SMD respectively. The solvation effects for reaction (4.6) are opposite those for reaction (4.5), with the COSMO approach favoring the reactant by -5.5 kcal/mol), and the SMD approach showing essentially no solvation effect.

Removing the second C(OH) proton to create the oxalate dianion (reaction (4.7)) leads to a decrease in the BDFE in solution. The BDFE is positive due to the large effect of solvation on the dianion, which makes the oxalate dianion stable to dissociation to two CO_2^- anions. Although the gas phase dianion is not going to be stable to autoionization, the highest occupied molecular orbital (HOMO) is slightly bound (-0.20 eV) at the Hartree-Fock (HF) level for the closed shell

singlet. The geometry of $C_2O_4^{2-}$ is weakly dependent on the presence of solvent. The BDFE for $C_2O_4^{2-}$ is not strongly dependent on the electronic structure method but is more dependent the solvent model with SMD giving larger BDFE's than with COSMO.

In order to improve our model to determine if two CO_2^- will dimerize in aqueous solution, we added 3 or 4 waters of solvation to each CO_2^- site in the dianion. Microsolvation of the oxalate with either six or eight total H_2O molecules makes the dianion stable in the gas phase with respect to breaking the C-C bond with BDEs of 15 kcal/mol for reaction (4.8) and 25 kcal/mol for reaction (4.9). The HF HOMOs for the closed shell singlet are strongly bound with values of -3.63 eV for six water molecules and -4.30 eV with eight water molecules. The gas-phase BDFE of the oxalate dianion with 6 H_2O molecules is virtually 0, and the gas-phase BDFE of the oxalate dianion with 8 H_2O molecules is 11 kcal/mol. The aqueous microsolvated oxalate dianion with 6 H_2O water molecules has a C-C BDFE of 52 kcal/mol at the CCSD(T)-F12b/aD/COSMO level and 57 kcal/mol at the CCSD(T)-F12b/aD/SMD level. An additional water of solvation on each side lowers the C-C BDFE to 47 kcal/mol at the CCSD(T)-F12b/aD/SMD level and to 43 kcal/mol at the CCSD(T)-F12b/aD/COSMO level. Although the results for the C-C BDFE of the aqueous oxalate dianion are not converged in terms of the number of waters of solvation, it is clear, barring kinetic constraints, that the oxalate dianion will form in aqueous solution if CO_2^- is present.

Conclusions

High level correlated molecular orbital theory calculations predict that CO_2^- is a stable species in the presence of 3, 4, and 8 explicit H_2O molecules in the gas phase. The radical electron density is localized on the carbon in the most stable structures as shown by analysis of the electron density and by changes in the symmetric and asymmetric C-O stretches. There are a

large number of low-energy conformations for the neutral and anionic clusters of CO_2 with 3, 4, or 8 explicit H_2O molecules in both the gas phase and aqueous solution. In general, the most energetically favorable clusters are those with the most strong hydrogen bonds. The highest energy (least favorable) anionic clusters of CO_2 in the gas phase with 4 and 8 H_2O molecules were those for which the free electron density was localized on H_2O molecules rather than being localized on the CO_2 molecule. The clusters with the electron density localized on the H_2O molecules are 10 to 20 kcal/mol higher in energy in the gas phase (ΔH_{gas}(298 K)) and 20 to 30 kcal/mol higher in energy in aqueous solution (ΔG_{aq}(298 K)). The EA of CO_2 in aqueous solution quickly converges with the number of explicit H_2O molecules and 3 H_2O molecules plus a self-consistent reaction field (SCRF) give results comparable to that with 8 explicit H_2O molecules plus a SCRF. The adiabatic electron affinity of CO_2 in aqueous solution is predicted to be 2.35 ± 0.08 eV. This value can be used to generate a reduction potential for CO_2/CO_2^- of -1.93 ± 0.08 eV (SHE) in good agreement with the experimental value of -1.90 eV.[16] The EA for the neutral CO_2 molecule solvated by explicit water molecules in aqueous solution is ~20 kcal/mol larger that the solvated free electron[21] with an EA of ~35 kcal/mol (1.52 eV). Thus, addition of a single CO_2 to H_2O substantially increases the ability of the system to bind an electron. The energy required to bend the linear CO_2 to the favored 135° of the CO_2 anion is large, 35.0 kcal/mol. The calculated VDE and VAE show that there is substantial reorganization of the local solvent environment in both cases. The generation of CO_2^- in aqueous solution could lead to the dimerization of two CO_2^- molecules to form the value-added oxalate dianion in solution. This is supported by experiment[17,18,19] as the oxalate dianion can be generated in aqueous solution from CO_2^- produced from CO_2 by electrochemical or electrocatalytic means.

Tables

Table 4.1. Relative Energies (ΔH_{298}) of Neutral and Anionic Clusters (kcal/mol) in the Gas Phase

Cluster	G3MP2 (B3)	MP2/ aT	CCSD(T)/ aD	CCSD(T)/ aT	CCSD(T)-F12b/aD	CCSD(T)-F12b/aT	CCSD(T)-F12b/CBS
N3-1 [a]	0.2	0.0	0.4	0.0	0.0	0.0	0.0
N3-2 [b]	0.0	0.2	0.0	0.3	0.5	0.5	0.6
N3-3 [a]	3.9	3.7	3.8	3.4	3.5	3.5	3.5
A3-1 [b,c]	0.0	0.0	0.0	0.0	0.0	0.0	0.0
A3-2 [c]	0.4	1.8	2.0	1.7	1.2	1.2	1.2
A3-3 [c]	6.4	2.1	2.9	2.2	2.2	2.1	2.0
N4-1 [a]	0.1	0.0	0.5	0.0	0.0	0.0	0.0
N4-2 [a]	0.0	0.0	0.3	0.1	0.1	0.2	0.2
N4-3 [b]	0.1	0.2	0.0	0.4	0.6	2.2	2.7
N4-4 [a]	5.6	4.9	4.9	4.6	4.6	4.6	4.6
N4-5 [a]	6.2	5.3	5.4	5.0	4.9	4.9	4.9
A4-1 [b,d]	0.0	0.0	0.0		0.0		
A4-2 [a,d]	2.8	2.4	3.4		2.5		
A4-3 [a,d]	8.6	7.5	5.8		5.5		
A4-4 [c]	7.6	5.8	7.0		5.8		
A4-5 [a,d]	6.5	6.3	7.7		6.6		
A4-6 [a,d]	3.9	4.1	10.8		10.4		
A4-7 [a,e]	[f]	15.8	20.6		18.5		
A4-8 [a,e]	[f]	18.7	23.8		21.1		
A4-9 [c]	[f]	19.9	24.5		22.0		
N8-1	0.0	9.2 [g]	0.0		0.0		
N8-2	0.3	0.0 [g]	1.4		0.8		
N8-3	4.0	2.5 [g]	3.4		3.4		
N8-4	5.4	6.8 [g]	5.8		5.4		
N8-5	8.5	9.4 [g]	8.7		8.2		
N8-6	6.7	12.5 [g]	12.3		12.1		
A8-1 [d]	1.2	0.0 [g]	0.0				
A8-2 [d]	2.2	0.0 [g]	1.1				
A8-3 [d]	0.0	1.8 [g]	2.6				
A8-4 [d]	3.7	2.5 [g]	3.4				
A8-5 [d]	3.8	3.5 [g]	4.3				
A8-6 [e]	25.7	20.5 [g]	23.1				

[a] Structures from Ref. 20 clusters with respect to the original labeling: **N3-1** = R3-3; **N3-3** = R3-2; **N4-1** = R4-4-1; **N4-2** = R4-4-2; **N4-5** = R4-2; **N4-4** = R4-3. [b] Structures relabeled from the Ref. 10 lowest energy clusters with 3 or 4 H_2O molecules: **A3-1**, **A4-1**. [c] Structures derived from $e^-\cdot(H_2O)_4$ or $e^-\cdot(H_2O)_8$ from Ref. 21. [d] CO_2 spin. [e] H_2O spin. [f] Did not optimize at the G3MP2(B3)

level with the radical electron density on the CO_2 molecules. [g] MP2/aT values are calculated with MP2/aT single point energy values and MP2/aD enthalpy corrections at 298 K.

Table 4.2. Relative Energies (ΔG_{298}) of Neutral and Anionic Clusters (kcal/mol) in Aqueous Phase with COSMO and SMD

Cluster	COSMO						SMD					
	MP2	CCSD(T)		CCSD(T) -F12b			MP2	CCSD(T)		CCSD(T) -F12b		
	aT	aD	aT	aD	aT	CBS	aT	aD	aT	aD	aT	CBS
N3-1[a]	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.7	0.6	0.6	0.6	0.6
N3-2 [b]	3.9	3.4	4.1	4.2	4.4	4.4	3.2	2.9	3.6	3.6	3.8	3.8
N3-3 [a]	1.1	0.8	0.8	0.9	0.9	0.9	0.0	0.0	0.0	0.0	0.0	0.0
A3-1 [b,d]	1.8	1.5	1.9	2.3	2.3	2.4	2.0	1.6	2.0	2.4	2.4	2.5
A3-2 [d]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
A3-3 [d]	1.2	1.6	1.3	1.8	1.6	1.6	0.6	1.0	0.7	1.2	1.0	1.0
N4-1 [a]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
N4-2 [a]	0.7	0.4	0.7	0.8	0.8	0.8	0.1	0.0	0.1	0.2	0.2	0.2
N4-3 [b]	4.4	3.8	4.4	4.6	6.2	6.7	3.1	1.5	1.9	2.1	3.7	4.2
N4-4 [a]	3.6	3.1	3.3	3.3	3.3	3.3	1.9	1.6	1.6	1.6	1.6	1.6
N4-5 [a]	3.1	2.6	2.7	2.7	2.7	2.7	1.9	1.6	1.5	1.5	1.5	1.5
A4-1 [b,d]	8.0	6.6		7.9			5.3	4.9		5.2		
A4-2 [d]	4.5	4.1		4.4			2.6	3.2		2.5		
A4-3 [d]	3.6	2.6		3.6			0.0	0.0		0.0		
A4-4 [d]	0.0	0.0		0.0			1.5	2.5		1.5		
A4-5 [d]	3.4	3.5		3.7			0.7	1.8		1.0		
A4-6 [d]	4.8	11.6		12.4			2.8	10.6		10.4		
A4-7 [e]	32.5	34.1		35.1			34.0	36.6		36.6		
A4-8 [e]	32.5	34.4		34.9			34.7	37.6		37.1		
A4-9 [e]	22.1	25.5		24.4			24.2	28.6		26.5		
N8-1	11.1 [f]	1.4		1.8			11.1	1.7		1.9		
N8-2	0.0 [f]	0.9		0.7			0.0	1.2		0.8		
N8-3 [c]	1.7 [f]	2.1		2.5			0.3	1.3		1.5		
N8-4	1.5 [f]	0.0		0.0			1.7	0.5		0.3		
N8-5 [c]	5.5 [f]	4.2		4.1			2.3	1.3		1.0		
N8-6 [c]	2.2 [f]	1.5		1.7			0.4	0.0		0.0		
A8-1 [d]	3.4 [f]	2.6					3.4	2.6				

A8-2 [d]	1.0 [f]	1.4					2.7	3.1				
A8-3 [c,d]	0.0 [f]	0.0					0.0	0.0				
A8-4 [d]	1.5 [f]	1.7					0.9	1.1				
A8-5 [c,d]	0.0 [f]	0.0					0.5	0.5				
A8-6 [c,e]	23.2 [f]	25.0					28.5	30.3				

[a] Structures from Ref. 20 clusters with respect to the original labeling: **N3-1** = R3-3; **N3-3** = R3-2; **N4-1** = R4-4-1; **N4-2** = R4-4-2; **N4-5** = R4-2; **N4-4** = R4-3. [b] Structures relabeled from Ref. 10 lowest energy clusters with 3 or 4 H_2O molecules: **A3-1**, **A4-1**. [c] Structures derived from $e^{-}\cdot(H_2O)_8$ from Ref. 21. [d] CO_2 spin. [e] H_2O spin. [f] MP2/aT values are calculated with MP2/aT single point energy values and MP2/aD free energy corrections at 298 K.

Table 4.3. Adiabatic Electron Affinities Gas and Aqueous Phases (kcal/mol)[a]

Energy	Method	CO_2/ CO_2^-	N3-1/ A3-1	N3-1/ A3-2	N3-3/ A3-2	N4-1/ A4-1	N4-1/ A4-3	N4-1/ A4-4	N8-1/ A8-1	N8-4/ A8-3	N8-6/ A8-3	N8-4/ A8-5	N8-6/ A8-5
ΔH_{gas}	MP2/aT	-17.3	7.1	5.4	9.1	8.7	1.3	3.0	25.3	21.1	26.7	19.4	25.1
ΔG_{gas}	MP2/aT	-15.4	3.6	4.7	7.5	4.8	-0.2	6.1	24.5	20.3	23.6	18.8	22.1
ΔG_{aq}	MP2/aT/COSMO	47.0	45.6	47.4	48.5	43.0	47.5	51.0	57.3	51.1	51.6	51.2	51.7
ΔG_{aq}	MP2/aT/SMD	48.6	47.3	48.9	49.1	45.3	50.6	49.0	59.6	53.6	52.2	53.2	51.8
ΔH_{gas}	G3MP2(B3)	-14.3	11.0	10.6	14.3	12.7	4.1	5.1	18.7	25.3	26.6	21.5	22.8
ΔG_{gas}	G3MP2(B3)	-12.4	8.1	9.5	12.8	10.0	2.2	9.4	16.4	23.8	24.1	22.0	22.3
ΔG_{aq}	G3MP2(B3)/COSMO	50.0	50.1	52.2	53.8	48.2	49.9	54.3	49.2	54.6	52.1	54.8	51.9
ΔG_{aq}	G3MP2(B3)/SMD	51.6	51.8	53.7	54.4	50.5	53.0	52.3	51.6	57.1	52.7	56.7	52.0
ΔH_{gas}	CCSD(T)/aD	-13.2	11.8	9.8	13.2	14.0	8.2	7.0	19.7	22.8	29.4	21.1	27.7
ΔG_{gas}	CCSD(T)/aD	-11.3	8.3	9.1	11.6	10.1	4.7	10.1	18.9	22.0	26.2	20.6	24.7
ΔG_{aq}	CCSD(T)/aD/COSMO	51.1	50.3	51.8	52.6	48.3	52.3	55.0	50.3	52.8	54.3	52.9	54.3
ΔG_{aq}	CCSD(T)/aD/SMD	53.8	52.0	53.3	53.2	50.6	55.5	53.0	52.8	55.3	54.8	54.9	54.4
ΔH_{gas}	CCSD(T)/aT	-10.9	11.9	10.2	13.7								
ΔG_{gas}	CCSD(T)/aT	-10.9	8.4	9.6	12.1								
ΔG_{aq}	CCSD(T)/aT/COSMO	51.5	50.4	52.3	53.1								
ΔG_{aq}	CCSD(T)/aT/SMD	54.3	52.1	53.7	53.7								
ΔH_{gas}	CCSD(T)-F12b/aD	-13.6	10.5	9.3	12.8	12.4	6.9	6.6					
ΔG_{gas}	CCSD(T)-F12b/aD	-11.8	7.0	8.7	11.2	8.5	3.3	9.7					
ΔG_{aq}	CCSD(T)-F12b/aD/COSMO	50.7	49.0	51.4	52.2	46.6	51.0	54.6					
ΔG_{aq}	CCSD(T)-F12b/aD/SMD	53.4	50.7	52.8	52.8	48.9	54.1	52.6					
ΔH_{gas}	CCSD(T)-F12b/aT	-12.6	11.4	10.2	13.7								
ΔG_{gas}	CCSD(T)-F12b/aT	-10.8	7.9	9.5	12.1								
ΔG_{aq}	CCSD(T)-F12b/aT /COSMO	51.6	49.9	52.2	53.1								
ΔG_{aq}	CCSD(T)-F12b/aT/SMD	54.4	51.6	53.7	53.7								

ΔH_{gas}	CCSD(T)-F12b/CBS	-12.3	11.7	10.5	14.0								
ΔG_{gas}	CCSD(T)-F12b/CBS	-10.5	8.2	9.9	12.4								
ΔG_{aq}	CCSD(T)-F12b/CBS/COSMO	52.0	50.2	52.5	53.4								
ΔG_{aq}	CCSD(T)-F12b/CBS/SMD	54.7	51.9	54.0	54.0								

[a] Positive values indicate bound states for electron affinities. All radical electron spin density for the anions is on CO_2.

Table 4.4. Vertical Detachment Energies[a] in Aqueous Phase (kcal/mol)

Method	COSMO					SMD				
	A3-2	A4-3	A4-4	A8-3	A8-5	A3-2	A4-3	A4-4	A8-3	A8-5
MP2/aT	96.8	97.1	95.2	105.4	95.1	96.1	95.8	91.0	94.3	92.5
CCSD(T)/aD	98.8	102.5	97.2	110.5	100.3	98.2	101.2	93.0	99.5	97.7
CCSD(T)/aT	102.1					101.4				
CCSD(T)-F12b/aD	102.4	105.5	100.7			101.8	104.1	96.5		
CCSD(T)-F12b/aT	103.8					103.1				

[a] Positive VDE (ΔE_{elec}) value indicates a bound electron.

Table 4.5. Vertical Attachment Energies[a] in Aqueous Phase (kcal/mol)

Method	COSMO					SMD				
	N3-1	N3-3	N4-4	N8-4	N8-6	N3-1	N3-3	N4-4	N8-4	N8-6
MP2/aT	16.5	15.3	14.8	14.9	12.2	13.7	12.4	8.8	8.5	10.4
CCSD(T)/aD	16.2	14.5	10.7	17.1	14.4	13.4	11.6	4.7	10.8	12.7
CCSD(T)/aT	19.0	17.5				16.1	14.5			
CCSD(T)-F12b/aD	17.8	16.3	12.0			15.0	13.3	6.0		
CCSD(T)-F12b/aT	19.8	18.6				17.0	15.7			

[a] Positive VAE ($-\Delta E_{elec}$) value indicates a bound electron.

Table 4.6. CO_2 Stretch Frequencies and Infrared Intensities at the MP2/aT and G3MP2(B3) Levels

	MP2/aT			G3MP2(B3)		
	Frequency (cm^{-1})	Intensity (km/mol)	Assign	Frequency (cm^{-1})	Intensity (km/mol)	Assign
CO_2	1325.8	0.0	Sym	1372.0	0.0	Sym
	2401.4	577.4	Asym	2436.3	545.8	Asym
CO_2^-	1203.7	66.8	Sym	1344.0	14.0	Sym
	1746.7	875.1	Asym	1706.0	484.9	Asym
HCO_3^-	1293.6	362.6	Sym	1333.6	344.5	Sym
	1763.0	794.5	Asym	1846.8	523.2	Asym
3 H_2O						
N3-1	1329.3	0.5	Sym	1371.2	0.8	Sym
	2405.8	557.0	Asym	2436.3	537.5	Asym
N3-2	1285.4	0.5	Sym	1372.4	0.7	Sym
	2326.9	541.5	Asym	2437.8	533.9	Asym
N3-3	1331.4	0.4	Sym	1359.9	4.8	Sym
	2406.0	632.4	Asym	2410.3	609.1	Asym
A3-1	1259.2	57.6	Sym	1336.4	14.4	Sym
	1619.4	749.9	Asym	1711.4	566.3	Asym
A3-2	1295.2	43.7	Sym	1336.4	6.6	Sym
	1682.8	570.6	Asym	1688.8	387.3	Asym
A3-3	1266.9	59.7	Sym	1326.2	73.9	Sym
	1758.9	607.9	Asym	1756.5	468.0	Asym [a]
				1775.1	145.2	Asym [a]
4 H_2O						
N4-1	1329.2	0.5	Sym	1372.0	0.5	Sym
	2405.6	550.0	Asym	2436.8	531.3	Asym
N4-2	1330.0	0.4	Sym	1373.7	0.5	Sym
	2407.2	523.2	Asym	2439.9	513.7	Asym
N4-3	1284.9	0.5	Sym	1371.9	0.5	Sym
	2326.1	535.1	Asym	2436.8	531.2	Asym
N4-5	1331.5	0.8	Sym	1369.6	2.5	Sym
	2409.0	573.5	Asym	2433.4	556.3	Asym
N4-4	1331.2	0.9	Sym	1367.8	1.7	Sym
	2409.2	537.6	Asym	2428.5	559.1	Asym
A4-1	1264.8	47.1	Sym	1334.8	8.6	Sym
	1609.0	721.8	Asym	1695.5	561.0	Asym
A4-2	1297.5	64.0	Sym	1337.3	16.8	Sym
	1683.3	80.7	Asym [a]	1713.6	569.0	Asym
	1684.0	492.5	Asym [a]			
	1727.5	260.5	Asym [a]			
A4-3	1279.0	37.8	Sym	1328.7	74.1	Sym
	1757.8	784.4	Asym	1764.8	606.6	Asym [a]
				1780.2	67.5	Asym [a]

A4-4	1307.2	18.8	Sym	1370.3	8.1	Sym
	1743.5	729.5	Asym	1701.8	1489.2	Asym
A4-5	1276.5	62.3	Sym	1313.1	106.9	Sym
	1769.1	697.8	Asym	1765.2	1001.3	Asym
A4-6	1263.1	155.1	Sym	1303.4	154.1	Sym
	1755.7	719.7	Asym	1733.6	945.1	Asym [a]
				1761.1	51.0	Asym [a]
A4-7 [b]	1328.9	13.3	Sym			
	2401.5	54.6	Asym			
A4-8 [b]	1319.1	256.4	Sym			
	2386.9	467.8	Asym			
A4-9 [b]	1315.1	151.2	Sym			
	2385.3	124.8	Asym			
8 H_2O [c]						
N8-2	1311.5	0.3	Sym	1371.2	1.1	Sym
	2386.6	564.5	Asym	2436.2	537.3	Asym
N8-3	1314.1	0.2	Sym	1375.3	0.2	Sym
	2390.9	466.6	Asym	2440.1	508.9	Asym
N8-4	1310.2	3.2	Sym	1368.9	4.0	Sym
	2385.7	565.2	Asym	2433.1	588.6	Asym
N8-1	1308.7	0.1	Sym	1372.4	0.0	Sym
	2383.6	457.5	Asym	2435.4	446.5	Asym
N8-5	1315.2	0.2	Sym	1370.4	0.4	Sym
	2392.5	527.8	Asym	2433.3	543.4	Asym
N8-6	1320.1	0.0	Sym	1372.8	0.0	Sym
	2400.7	503.6	Asym	2437.1	456.7	Asym
A8-1	1286.8	42.6	Sym	1335.3	8.2	Sym
	1638.9	468.4	Asym [a]	1684.8	501.4	Asym
	1693.0	287.6	Asym [a]			
A8-2	1284.7	56.2	Sym	1345.1	6.0	Sym
	1671.8	705.8	Asym	1723.8	278.6	Asym [a]
				1731.8	233.3	Asym [a]
A8-3	1283.2	46.2	Sym	1316.2	68.3	Sym
	1728.1	487.2	Asym	1747.7	705.4	Asym
A8-4	1281.0	49.6	Sym	1328.5	53.2	Sym
	1724.7	442.1	Asym	1744.5	262.4	Asym [a]
				1757.6	352.2	Asym [a]
A8-5	1286.0	15.9	Sym	1337.5	64.2	Sym
	1737.1	735.7	Asym	1754.8	483.3	Asym [a]
				1789.4	215.0	Asym [a]
A8-6 [b]	1308.9	1.0	Sym	1371.0	1.4	Sym
	2381.2	415.3	Asym	2433.2	465.7	Asym

[a] Asymmetric CO_2 stretches are coupled with water molecule bending modes. [b] Radical electron density on the H_2O molecules. All other anions maintain the radical electron density on the CO_2. [c] MP2/aD vibrations are given for clusters of CO_2 with 8 H_2O molecules.

Table 4.7. Decomposition Reaction Energies in the Gas and Aqueous Phases at 298 K (kcal/mol)

Energy	Method	(4.5)	(4,6)	(4.7)[a]	(4.8)	(4.9)
ΔH_{gas}	MP2/aT	90.4	85.7		20.7	29.8
ΔG_{gas}	MP2/aT	76.6	73.5		8.0	15.2
ΔG_{aq}	MP2/aT/COSMO	72.9	79.0	55.1	60.6	47.6
ΔG_{aq}	MP2/aT/SMD	78.5	73.4	63.6	65.6	51.3
ΔH_{gas}	G3MP2(B3)	86.9	83.9		28.2	18.6
ΔG_{gas}	G3MP2(B3)	73.2	72.0		14.9	5.4
ΔG_{aq}	G3MP2(B3)/COSMO	69.5	77.5	53.6	67.5	37.8
ΔG_{aq}	G3MP2(B3)/SMD	75.1	71.9	62.1	72.5	41.5
ΔH_{gas}	CCSD(T)/aD	84.1	80.5		15.1	25.5
ΔG_{gas}	CCSD(T)/aD	70.3	68.4		-0.3	10.9
ΔG_{aq}	CCSD(T)/aD/COSMO	66.6	73.9	50.5	52.3	43.3
ΔG_{aq}	CCSD(T)/aD/SMD	72.2	68.3	59.0	57.3	47.0
ΔH_{gas}	CCSD(T)/aT	86.3	81.6			
ΔG_{gas}	CCSD(T)/aT	72.5	69.4			
ΔG_{aq}	CCSD(T)/aT/COSMO	68.8	74.9	50.6		
ΔG_{aq}	CCSD(T)/aT/SMD	74.4	69.3	59.1		
ΔH_{gas}	CCSD(T)-F12b/aD	86.6	82.0		15.0	25.4
ΔG_{gas}	CCSD(T)-F12b/aD	72.8	69.9		-0.4	10.7
ΔG_{aq}	CCSD(T)-F12b/aD/COSMO	69.1	75.4	50.9	52.2	43.1
ΔG_{aq}	CCSD(T)-F12b/aD/SMD	74.7	69.8	59.4	57.2	46.8
ΔH_{gas}	CCSD(T)-F12b/aT	87.2	82.3			
ΔG_{gas}	CCSD(T)-F12b/aT	73.4	70.2			
ΔG_{aq}	CCSD(T)-F12b/aT/COSMO	69.7	75.7	51.2		
ΔG_{aq}	CCSD(T)-F12b/aT/SMD	75.3	70.1	59.7		
ΔH_{gas}	CCSD(T)-F12b/CBS	87.4	82.5			
ΔG_{gas}	CCSD(T)-F12b/CBS	73.6	70.3			
ΔG_{aq}	CCSD(T)-F12b/CBS/COSMO	69.9	75.8	51.3		
ΔG_{aq}	CCSD(T)-F12b/CBS/SMD	75.5	70.2	59.8		

[a] The gas phase values for reaction (4.7) are not included as $C_2O_4^{2-}$ is likely to autoionize in the gas phase.

Figures

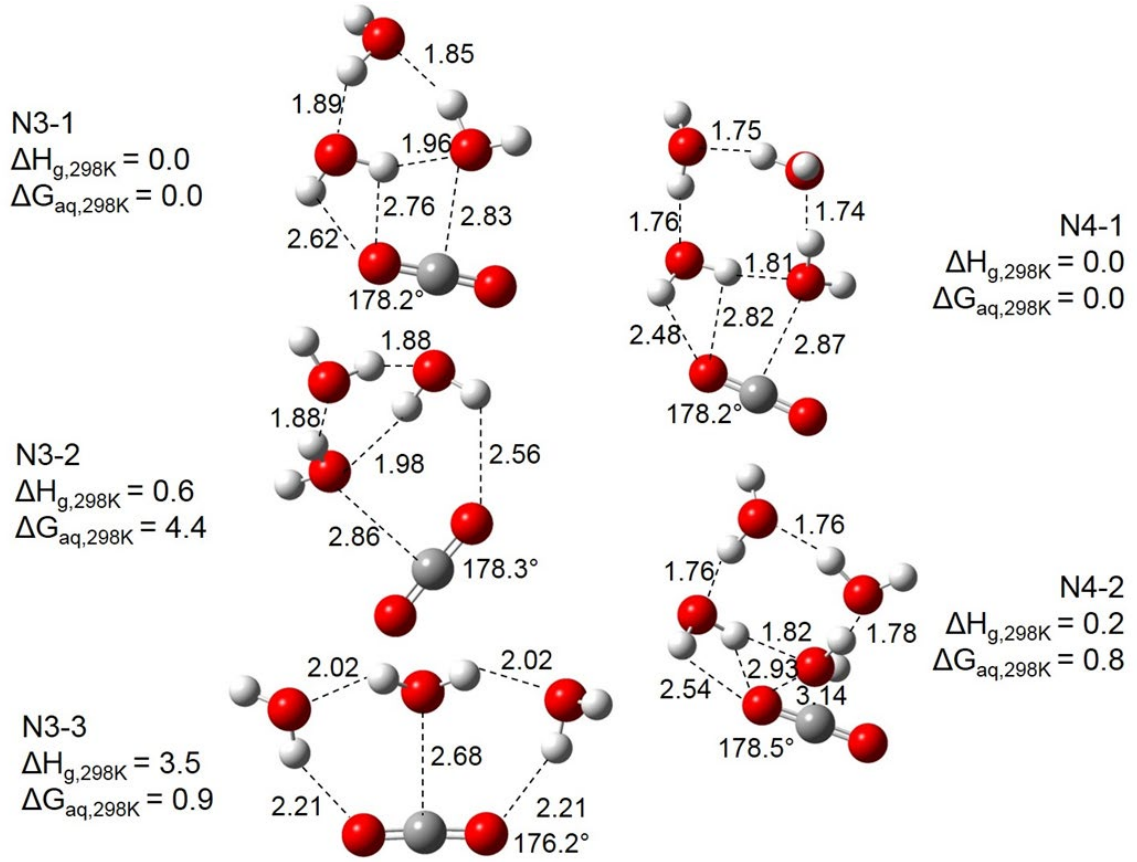


Figure 4.1. Lowest energy $CO_2 \cdot (H_2O)_3$ and $CO_2 \cdot (H_2O)_4$ clusters in the gas phase and in aqueous solution. Relative gas-phase enthalpies at the CCSD(T)-F12b/CBS level and relative aqueous free energies CCSD(T)-F12b/CBS/COSMO level. Bond distances in Å and energies in kcal/mol.

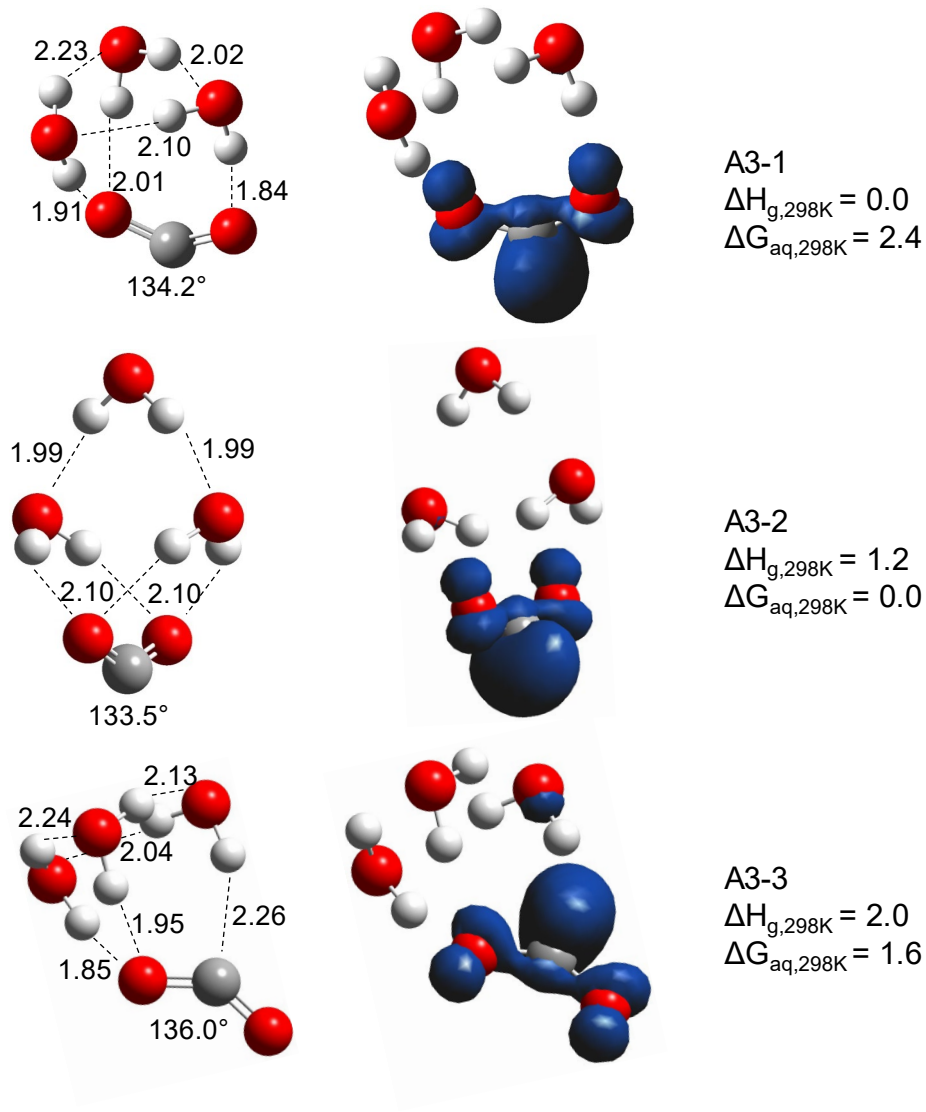


Figure 4.2. $CO_2^{-\cdot}(H_2O)_3$ clusters in the gas phase at the MP2/aT level with electron density images at the B3LYP/aT level. Relative gas-phase enthalpies at the CCSD(T)-F12b/CBS level and relative aqueous free energies CCSD(T)-F12b/CBS/COSMO level. Bond distances in Å and energies in kcal/mol. Spin populations for **A3-1** are 0.75 e^- on C and 0.12 e^- on each O in the CO_2, for **A3-2** are 0.77 e^- on C, and 0.12 e^- on each O in the CO_2, and for **A3-3** are 0.65 e^- on C and 0.15 and 0.18 e^- on the O atoms of the CO_2, with up to 0.03 e^- on the O atoms in H_2O molecules. The isovalue contour is 0.005000.

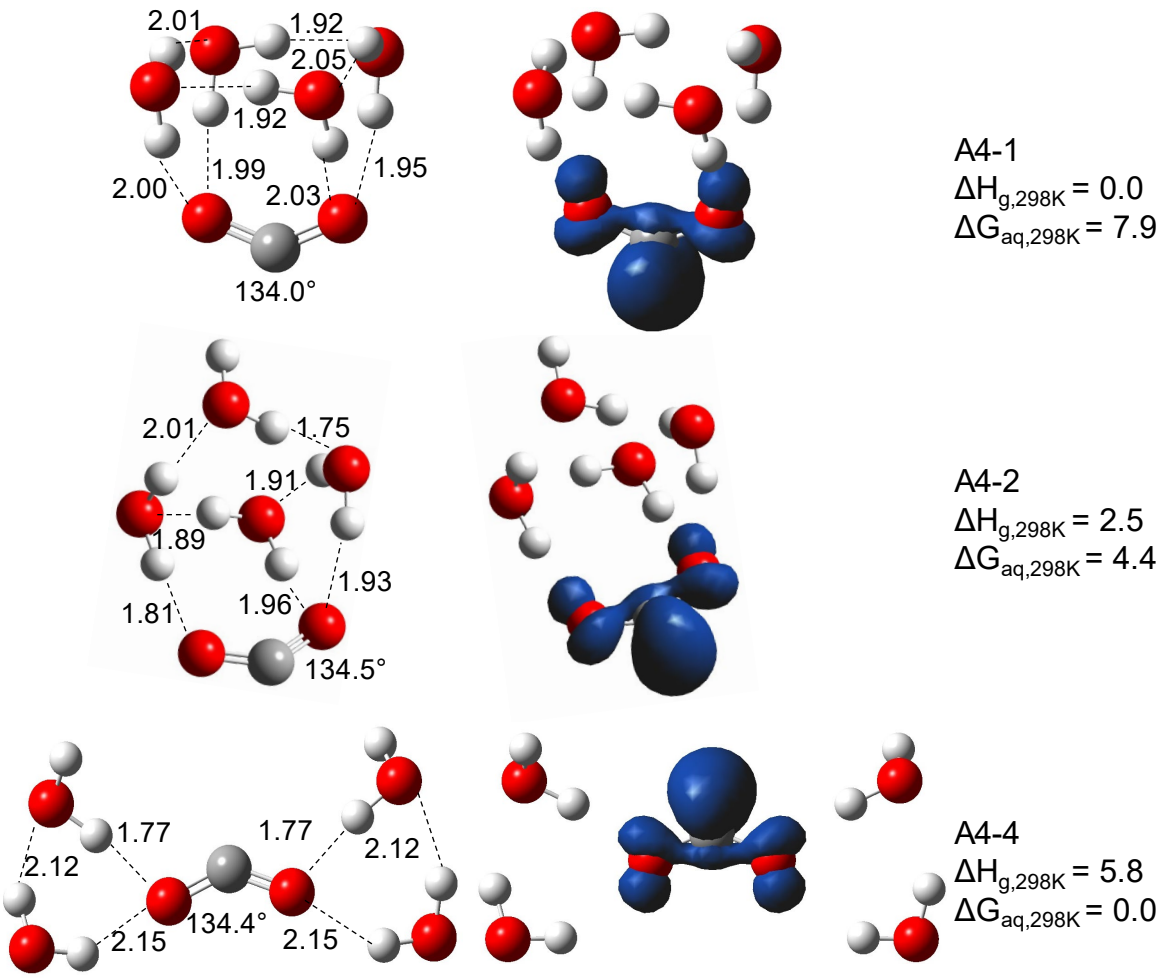


Figure 4.3. Lowest energy $CO_2^{-}\cdot(H_2O)_4$ clusters in the gas phase optimized at the MP2/aT level. Relative gas-phase enthalpies at the CCSD(T)-F12b/aD level and relative aqueous free energies at the CCSD(T)-F12b/aD/COSMO level. Bond distances in Å and energies in kcal/mol. Spin populations for **A4-1** and **A4-2** are 0.75 e^- on C and 0.12 e^- on each O in the CO_2 and for **A4-4** are 0.70 e^- on C and 0.16 e^- on both of the O atoms of the CO_2, with up to 0.03 e^- on the O atoms of the H_2O molecules. The isovalue contour is 0.005000.

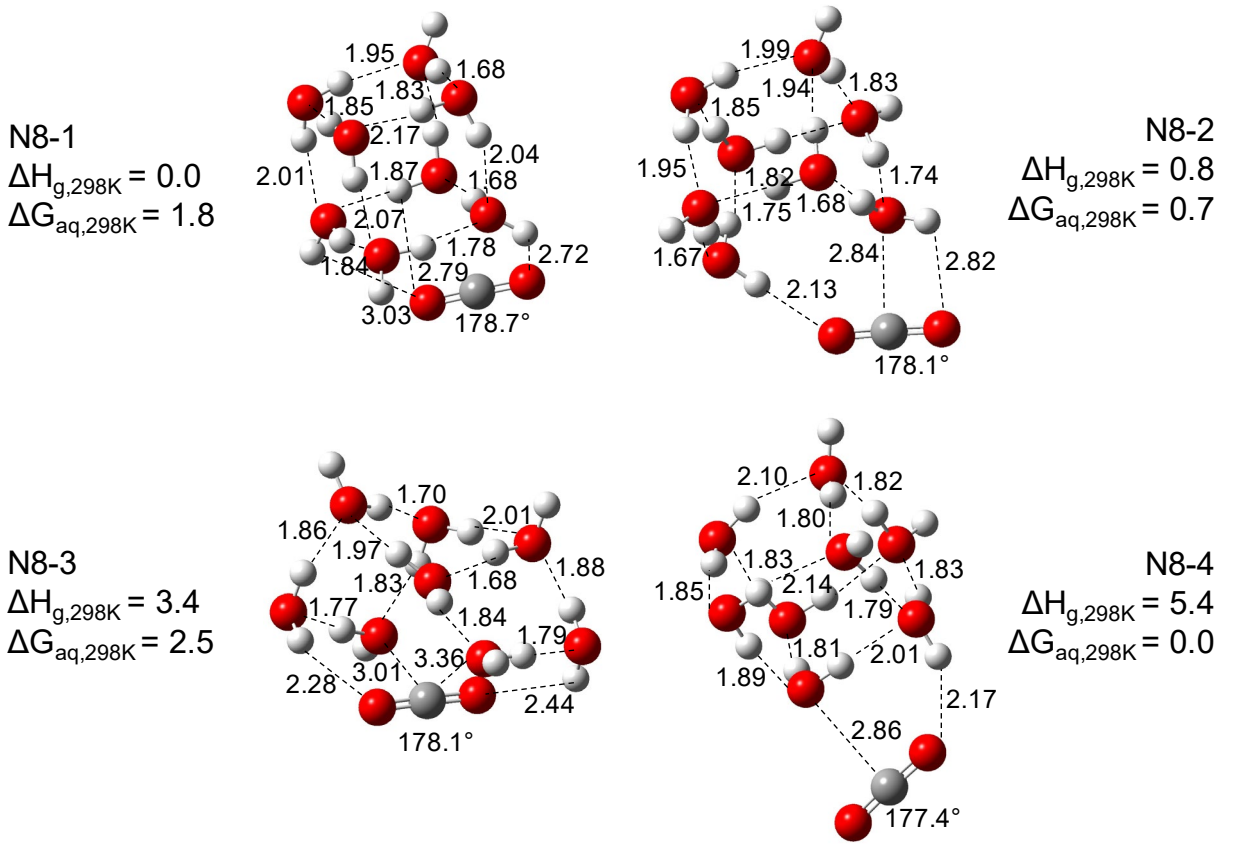


Figure 4.4. Lowest energy $CO_2 \cdot (H_2O)_8$ clusters in the gas phase and in aqueous solution. Relative gas-phase enthalpies at the CCSD(T)-F12b/aD level and relative aqueous free energies CCSD(T)-F12b/aD/COSMO level. Bond distances in Å and energies in kcal/mol.

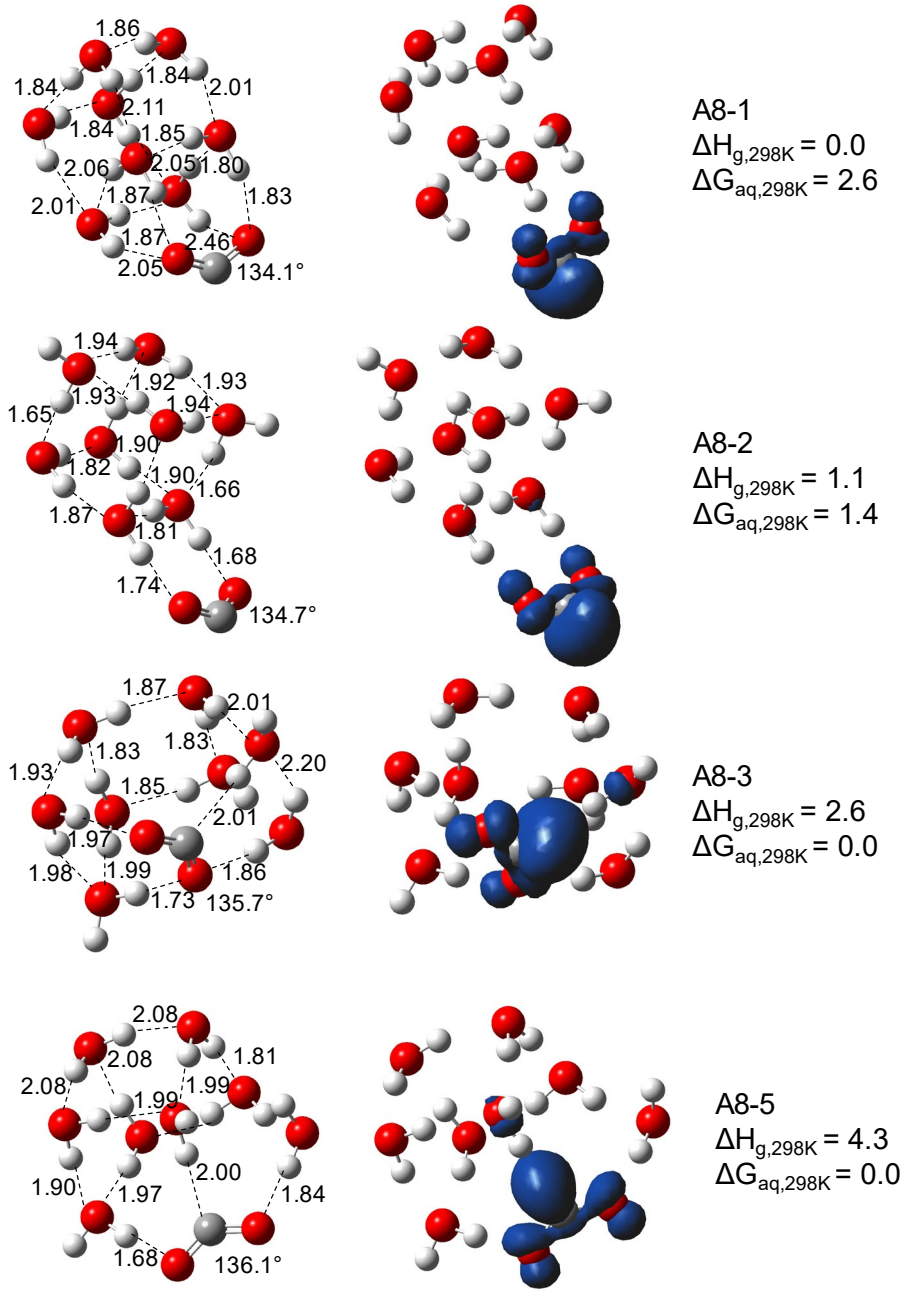


Figure 4.5. Lowest energy $CO_2^{-}\cdot(H_2O)_8$ clusters in the gas phase and in aqueous solution at the CCSD(T)/aD and CCSD(T)/aD/COSMO levels, respectively, with electron density images at the B3LYP/aT level. Bond distances in Å and energies in kcal/mol. Spin densities for **A8-1** are 0.83 e^- on the C and 0.07 e^- on each of the O atoms of the CO_2, for **A8-2** are 0.84 e^- on the C and 0.07 and 0.08 e^- on the O atoms of the CO_2, and for **A8-3** are 0.71 e^- on the C and 0.13 and 0.17 e^- on the O atoms of the CO_2. The lowest energy $CO_2^{-}(H_2O)_8$ cluster in the aqueous phase is **A8-5** with spin densities of 0.64 e^- on the C and 0.17 and 0.20 e^- on the O atoms of the CO_2. The isovalue contour is 0.005000.

References

[1] Krauss, M.; Neumann, D. Energy Curves of CO_2^-. *Chem. Phys. Lett.* **1972**, *14*, 26 – 28.

[2] Cooper, C. D.; Compton, R. N. Metastable Anions of CO_2. *Chem. Phys. Lett*. **1972**, *14*, 29 – 32.

[3] Pacansky, J.; Wahlgren, U.; Bagus, P. S. SCF Ab-Initio Ground State Energy Surfaces for CO_2 and CO_2^-. *J. Chem. Phys.* **1975**, *62*, 2740 – 2744.

[4] England, W. B.; Rosenberg, B. J.; Fortune, P. J.; Wahl, A. C. Ab Initio Vertical Spectra and Linear Bent Correlation Diagrams for the Valence States of CO_2 and Its Singly Charged Ions. *J. Chem. Phys.* **1976**, *65*, 684 – 691.

[5] Hartman, K. O.; Hisatsune, I. C. Infrared Spectrum of Carbon Dioxide Anion Radical. *J. Chem. Phys*. **1966**, *44*, 1913 – 1918.

[6] Klots, C. E.; Compton, R. N. Electron Attachment to Carbon Dioxide Clusters in a Supersonic Beam. *J. Chem. Phys.* **1977**, *67*, 1779 – 1780.

[7] Compton, R. N.; Reinhardt, P. W.; Cooper, C. D. Collision Ionization of Na, K, and Cs by CO_2, COS, and CS_2: Molecular Electron Affinities. *J. Chem. Phys.* **1975**, *63*, 3821 – 3827.

[8] Cooper, C. D.; Compton, R. N. Electronic Attachment to Cyclic Anhydrides and Related Compounds. *J. Chem. Phys.* **1973**, *59*, 3550 – 3565.

[9] Klots, C. E. Stable Gas Phase Hydrates of the CO_2^- Anion. *J. Chem. Phys.* **1979**, *71*, 4172.

[10] Balaj, O. P.; Siu, C.-K.; Balteanu, I.; Beyer, M. K.; Bondybey, V. E. Reactions of Hydrated Electrons $(H_2O)_n^-$ with Carbon Dioxide and Molecular Oxygen: Hydration of the CO_2^- and O_2^- Ions. *Chem. Eur. J.* **2004**, *10*, 4822 – 4830.

[11] Breen, K. J.; DeBlase, A. F.; Guasco, T. L.; Voora, V. K.; Jordan, K. D.; Nagata, T.; Johnson, M. A. Bottom-Up view of Water Network-Mediated CO_2 Reduction Using Cryogenic Cluster Ion Spectroscopy and Direct Dynamics Simulations. *J. Phys. Chem. A* **2012**, *116*, 903 – 912.

[12] Liu, P.; Zhao, J.; Liu, J.; Zhang, M.; Bu, Y. Ab Initio Molecular Dynamics Simulations Reveal Localization and Time Evolution Dynamics of an Excess Electron in Heterogeneous CO_2-H_2O Systems. *J. Chem. Phys.* **2014**, *140*, 044318-1 – 044318-11.

[13] Janik, I.; Tripathi, G. N. R. The Nature of the CO_2^- Radical Anion in Water. *J. Chem. Phys.* **2016**, *144*, 154307-1 – 154307-7.

[14] Rauk, A.; Armstrong, D. A.; Yu, D. The Lifetimes of Gas Phase $CO_2^{\cdot -}$ and $N_2O^{\cdot -}$ Calculated from the Transition Probability of the Autodetachment Process $A^- \rightarrow A + e^-$. *Int. J. Chem. Kinet.* **1994**, *26*, 7 – 24.

[15] Sommerfeld, T.; Meyer, H.-D.; Cederbaum, L. S. Potential Energy Surface of the CO_2^- Anion. *Phys. Chem. Chem. Phys.* **2004**, *6*, 42 – 45.

[16] Schwarz, H. A.; Dodson, R. W. Reduction potentials of CO_2^- and the Alcohol Radicals. *J. Phys. Chem.* **1989**, *93*, 409 – 414.

[17] Christensen, P. A.; Higgins, S. J. The Electrochemical Reduction of CO_2 to oxalate at a Pt Electrode Immersed in Acetonitrile and Coated with Polyvinylalcohol/$[Ni(dppm)_2Cl_2]$. *J. Electroanal. Chem.* **1995**, *387*, 127 – 132.

[18] Angamuthu, R.; Byers, P.; Lutz, M.; Spek, A. L.; Bouwman, E. Electrocatalytic CO_2 Conversion to Oxalate by a Copper Complex. *Science* **2010**, *327*, 313 – 315.

[19] Pokharel, U. R.; Fronczek, F. R.; Maverick, A. W. Reduction of Carbon Dioxide to Oxalate by a Binuclear Copper Complex. *Nat. Commun.* **2014**, *5*, 5883-1 – 5883-5.

[20] Nguyen, M. T.; Matus, M. H.; Jackson, V. E.; Ngan, V. T.; Rustad, J. R.; Dixon, D. A. Mechanism of the Hydration of Carbon Dioxide: Direct Participation of H_2O Versus Microsolvation. *J. Phys. Chem. A* **2008**, *112*, 10386 – 10398.

[21] Zhan, C.-G.; Dixon, D. A. The Nature and Absolute Hydration Free Energy of the Solvated Electron in Water. *J. Phys. Chem. B* **2003**, *107*, 4403 – 4417.

[22] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.

[23] Lee, C.; Yang, W.; Parr, R. G. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Phys. Rev. B* **1988**, *37*, 785 – 789.

[24] Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys*. **1989**, *90*, 1007 – 1023.

[25] Kendall, R. A.; Dunning, T. H. Jr.; Harrison, R. J., Electron Affinities of the First Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys*. **1992,** *96*, 6796 – 6806.

[26] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 0618 – 0622.

[27] Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical Models Incorporating Electron Correlation. *Int. J. Quantum. Chem.* **1976**, *10*, 1 – 19.

[28] Baboul, G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-3 Theory Using Density Functional Geometries and Zero-Point Energies. *J. Chem. Phys.* **1999**, *110*, 7650 – 7657.

[29] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16, Revision A.03, Gaussian, Inc., Wallingford CT, 2016.

[30] Purvis, G. D., III; Bartlett, R. J. A Full Coupled Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910 – 1918.

[31] Raghavachari, K.; Truck, G.W.; Pople, J.A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett*. **1989**, *157*, 479 – 483.

[32] Watts, J. D.; Gauss, J.; Bartlett, R. J. Coupled-Cluster Methods with Noniterative Triple Excitations for Restricted Open Shell Hartree-Fock and Other General Single Determinant Reference Functions. Energies and Analytical Gradients. *J. Chem. Phys.* **1993**, *98*, 8718 – 8733.

[33] Bartlett, R. J.; Musial, M. Coupled Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291 – 352.

[34] Knowles, P. J.; Hampel, C.; Werner, H.-J. Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **1993**, *99*, 5219 – 5227.

[35] Knowles, P. J.; Hampel, C.; Werner, H.-J. Erratum: Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **2000,** *112*, 3106 – 3107.

[36] Deegan, M. J. O.; Knowles, P. J. Perturbative Corrections to Account for Triple Excitations in Closed and Open Shell Coupled Cluster Theories. *Chem. Phys. Lett*. **1994**, *227*, 321 – 327.

[37] Adler, T. B.; Knizia, G.; Werner, H.-J. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Phys. Chem.* **2007,** *127,* 221106-1 – 221106-4.

[38] Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009,** *130,* 054104-1 – 054104-20.

[39] Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-ζ Quality Coupled-Cluster Correlation Energies with Triple- ζ Basis Sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921 – 1930.

[40] Yousaf, K. E.; Peterson, K. A. Optimized Auxiliary Basis Sets for Explicitly Correlated Methods. *J. Chem. Phys.* **2008**, *129*, 184108-1 – 184108-7.

[41] Hill, J. G.; Mazumder, S.; Peterson, K. A. Correlation Consistent Basis Sets for Molecular Core-Valence Effects with Explicitly Correlated Wave Functions: The Atoms B-Ne and Al-Ar. *J. Chem. Phys.* **2010**, *132*, 054108-1 – 054108-11.

[42] Werner H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, T.; Korona, T.; Lindh, R.; et al. MOLPRO, version 2015.1, a package of *ab initio* programs, See http://www.molpro.net. Accessed March 1, 2016.

[43] Werner H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242 – 253.

[44] Martin, J. M. L. Ab Initio Total Atomization Energies of Small Molecules - Towards the Basis Set Limit. *Chem. Phys. Lett.* **1996**, *259*, 669 – 678.

[45] Feller, D.; Peterson, K.A.; Hill, J. G. On the Effectiveness of CCSD(T) Complete Basis Set Extrapolations for Atomization Energies. *J. Chem. Phys.* **2011***, 135,* 044102-1 – 044102-18.

[46] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Contiuum Solvation Models. *Chem. Rev.* **2005**, 105, 2999 – 3093.

[47] Klamt, A. *Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam, 2005.

[48] Klamt, A.; Schüürmann, G. J. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *Chem. Soc. Perkin Trans.* **1993**, 2, 799 – 805.

[49] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378 – 6396.

[50] Dixon, D. A.; Feller, D.; Peterson, K. A. A Practical Guide to Reliable First Principles Computational Thermochemistry Predictions Across the Periodic Table. In *Annual Reports in Computational Chemistry*, Ralph, A. W., Ed.; Elsevier: 2012; Vol. 8, pp 1 – 28.

[51] Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy: Current Strategies and Future Challenges. *Theor. Chem. Acc.* **2012**, *131*, 1 – 20.

[52] Feller, D.; Peterson, K. A.; Dixon, D. A. Further Benchmarks of a Composite, Convergent, Statistically Calibrated Coupled Cluster Based Approach for Thermochemical and Spectroscopic Studies. *Mol. Phys.* **2012**, *110*, 2381 – 2399.

[53] Feller, D.; Peterson, K. A.; Dixon, D. A. The Impact of Larger Basis Sets and Explicitly Correlated Coupled Cluster Theory on the Feller-Peterson-Dixon Composite Method. In *Annual Reports in Computational Chemistry*, Vol. 12, Ed. D. A. Dixon, Elsevier, Amsterdam, 2016, pp. 47 – 78.

[54] Woon, D. E.; Dunning, T. H. Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-Valence Basis Sets for Boron Through Neon. *J. Chem. Phys.* **1995**, *103*, 4572 – 4585.

[55] Peterson, K. A.; Dunning, T. H. Accurate Correlation Consistent Basis Sets for Molecular Core-Valence Correlation Effects: The Second Row Atoms Al-Ar, and the First Row Atoms B-Ne Revisited. *J. Chem. Phys.* **2002**, *177*, 10548 – 10560.

[56] Douglas, M.; Kroll, N. M. Quantum Electrodynamical Corrections to the Fine Structure of Helium. *Ann. Phys.* **1974**, *82*, 89 – 155.

[57] Hess, B. A. Applicability of the No-Pair Equation with Free-Particle Projection Operators to Atomic and Molecular Structure Calculations. *Phys. Rev. A* **1985**, *32*, 756 – 763.

[58] Hess, B. A. Relativistic Electronic-Structure Calculations Employing a Two-Component No-Pair Formalism with External-Field Projection Operators. *Phys. Rev. A* **1986**, *33*, 3742 – 3748.

[59] De Jong, W. A.; Harrison, R. J.; Dixon, D. A. Parallel Douglas-Kroll Energy and Gradients in NWChem: Estimating Scalar Relativistic Effects Using Douglas-Kroll Contracted Basis Sets. *J. Chem. Phys*. **2001**, *114*, 48 – 53.

[60] Moore, C. E. *Atomic Energy Levels As Derived from the Analysis of Optical Spectra*; U.S. National Bureau of Standards Circular 467, COM-72-50282; U.S. Department of Commerce, National Technical Information Service: Washington, DC, 1949; Vol. *1*, H to V.

[61] Chase, M. W., Jr. *NIST-JANAF Thermochemical Tables*, 4th ed. *J. Phys. Chem. Ref. Data.*, *Mono. 9*; American Institute of Physics: Woodbury, NY, 1998; Suppl 1.

[62] Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies of Formation. *J. Chem. Phys.* **1997**, *106*, 1063 – 1079.

[63] AMPAC 10, 1992-2013 Semichem, Inc.: Shawnee, KS.

[64] Zhan, C.-G; Dixon, D. A. Hydration of the Fluoride Anion: Structures and Absolute Hydration Free Energy from First-Principles Electronic Structure Calculation. *J. Phys. Chem. A* **2004**, *108*, 2020 – 2029.

[65] Kemp, D. D.; Gordon, M. S. Theoretical Study of the Solvation of Fluorine and Chlorine Anions by Water. *J. Phys. Chem. A* **2005**, *109*, 7688 – 7699.

[66] Trasatti, S. The Absolute Electrode Potential: An Explanatory Note. *Pure Appl. Chem.*, **1986**, *58*, 955 – 966.

[67] Marenich, A. V.; Ho, J.; Coote, M. L.; Cramer, C. J.; Truhlar, D. G. Computational Electrochemistry: Prediction of Liquid-Phase Reduction Potentials. *Phys. Chem. Chem. Phys.,* **2014**, *16*, 15068 – 15106.

[68] Isse, A. A.; Gennaro, A. Absolute Potential of the Standard Hydrogen Electrode and the Problem of Interconversion of Potentials in Different Solvents. *J. Phys. Chem. B* **2010,** *114,* 7894 – 7899.

[69] Camaioni, D. M.; Dupuis, M.; Bentley, J. Theoretical Characterization of Oxoanion, XO_m^{n-}, Solvation. *J. Phys. Chem. A* **2003,** *107,* 5778 – 5788.

[70] Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5*, 129 – 145.

[71] Chirlian, L. E.; Francl, M. M. Atomic Charges Derived from Electrostatic Potentials: A Detailed Study. *J. Comput. Chem.* **1987**, *8*, 894 – 905.

[72] Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431 – 439.

[73] Breneman, C. M.; Wiberg, K. B. Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis. *J. Comput. Chem.* **1990**, *11*, 361 – 373.

[74] Curtiss, L. A.; Raghavachary, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-3 and Density Functional Theories for a Larger Experimental Test Set. *J. Chem. Phys.* **2000**, *112*, 7374 – 7383.

[75] Shimanouchi, T. *Tables of Molecular Vibrational Frequencies, Consolidated Volume I,* NSRDS NBS-39, National Institutes of Standards, Washington D.C., 1972.

[76] Knight, L. B.; Hill, D.; Berry, K.; Babb, R.; Feller, D. Electron Spin Resonance Rare Gas Matrix Studies of $^{12}CO_2^-$, $^{13}CO_2^-$, and $C^{17}O_2^-$: Comparison with Ab Initio Calculations. *J. Chem. Phys.* **1996**, *105*, 5672 – 5686.

[77] Thompson, W. E.; Jacox, M. E. The Vibration Spectra of CO_2^+, $(CO_2)_2^+$, CO_2^-, and $(CO_2)_2^-$ Trapped in Solid Neon. *J. Chem. Phys.* **1999**, *111*, 4487 – 4496.

[78] Davis, A. R.; Oliver, B. G. A Vibrational-Spectroscopic Study of the Species Present in the CO_2-H_2O System. *J. Solution Chem.* **1972**, *1*, 329 – 339.

[79] Rudolph, W. W.; Fischer, D.; Irmer, F. Vibrational Spectroscopic Studies and Density Functional Theory Calculations of Speciation in the CO_2^- Water System. *Appl. Spectrosc*. **2006**, *60*, 130 – 144.

[80] Rustad, J. R.; Nelmes, S. L.; Jackson, V. E.; Dixon, D. A. Quantum-Chemical Calculations of Carbon-Isotope Fractionation in CO_2(g), Aqueous Carbonate Species, and Carbonate Minerals. *J. Phys. Chem. A* **2008**, *112*, 542 – 555.

[81] Pedley, J. B. *Thermochemical Data and Structure of Organic Compounds*; Thermodynamics Research Center, Texas A&M University: College Station, TX, 1994.

[82] Feller, D.; Dixon, D. A.; Francisco, J. S. Coupled Cluster Theory Determination of the Heats of Formation of Combustion-related compounds: CO, HCO, CO_2, HCO_2, HOCO, HC(O)OH and HC(O)OOH. *J. Phys. Chem. A* **2003**, *107*, 1604 – 1617.

[83] Burkholder, J. B.; Abbatt, J. P. D.; Huie, R. E.; Kurylo, M. J.; Wilmouth, D. M.; Sander, S. P.; Barker, J. R.; Kolb, C. E.; Orkin, V. L.; Wine, P. H. *Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies: Evaluation Number 18*, JPL Publication 15-10, National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, Oct. 2015, https://jpldataeval.jpl.nasa.gov/pdf/JPL_Publication_15-10.pdf. accessed June 1, 2018.

[84] Luo, Y-R. *Handbook of Bond Dissociation Energies in Organic Compounds*; CRC Press: Boca Raton, FL, 2003.

[85] Pedley, J. B.; Naylor, R. D.; Kirby, S. P. *Thermochemical Data of Organic Compounds*, 2nd ed.; Chapman and Hall: New York, 1986.

Full Reference 29 (Gaussian 16): Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision A.03, Gaussian, Inc., Wallingford CT, 2016.

Full Reference 42 (Molpro 2015): Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, T.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.;Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. MOLPRO, version 2015.1, a package of *ab initio* programs, See http://www.molpro.net. Accessed March 1, 2016.

Table A4.1. Total Energies of Neutral and Anionic Species in the Gas Phase in Hartrees: MP2 and G3MP2(B3)

Species	MP2/aT (n = 3, 4, 6) or aD (n = 8)			G3MP2(B3)		
	ΔH(0 K)	ΔH(298 K)	ΔG(298 K)	ΔH(0 K)	ΔH(298 K)	ΔG(298 K)
CO_2	-188.310147	-188.306572	-188.330860	-188.385798	-188.382179	-188.406529
CO_2 (135°)	-188.253642	-188.249744	-188.276227			
CO_2^-	-188.282991	-188.279077	-188.306300	-188.363301	-188.359398	-188.386693
$^2HCO_2$	-188.801160	-188.797002	-188.825549	-188.887734	-188.883542	-188.912154
HCO_3^-	-264.069161	-264.064687	-264.094934	-264.179405	-264.174873	-264.205214
N3-1	-417.254988	-417.242306	-417.294175	-417.440003	-417.428033	-417.478108
N3-2	-417.253775	-417.242062	-417.290023	-417.440404	-417.428378	-417.478354
N3-3	-417.250164	-417.236480	-417.289805	-417.434933	-417.422096	-417.472889
A3-1	-417.264657	-417.253681	-417.299952	-417.456134	-417.445540	-417.491034
A3-2	-417.263276	-417.250917	-417.301732	-417.456654	-417.444945	-417.493248
A3-3	-417.262044	-417.250336	-417.298973	-417.446239	-417.435275	-417.482384
N4-1	-493.577725	-493.562790	-493.620958	-493.799847	-493.785474	-493.841682
N4-2	-493.577806	-493.562712	-493.619541	-493.800219	-493.785651	-493.841310
N4-3	-493.576265	-493.562354	-493.615847	-493.799864	-493.785486	-493.841703
N4-4	-493.570844	-493.555054	-493.613368	-493.791621	-493.776730	-493.833323
N4-5	-493.570358	-493.554432	-493.614111	-493.790776	-493.775821	-493.833783
A4-1	-493.590072	-493.576668	-493.628658	-493.818788	-493.805714	-493.857688
A4-2	-493.586964	-493.572878	-493.627417	-493.814276	-493.801179	-493.852884
A4-3	-493.579258	-493.564878	-493.620655	-493.805536	-493.792039	-493.845194
A4-4	-493.583828	-493.567622	-493.630678	-493.809426	-493.793663	-493.856621
A4-5	-493.581922	-493.566791	-493.625653	-493.809800	-493.795341	-493.851400
A4-6	-493.585326	-493.570191	-493.627889	-493.813978	-493.799451	-493.855564
A4-7	-493.566509	-493.551814	-493.608382			
A4-8	-493.562480	-493.547305	-493.606207			
A4-9	-493.563359	-493.545468	-493.615230			
N8-1	-798.166066	-798.141913	-798.219152	-799.228455	-799.205342	-799.280054
N8-2	-798.164363	-798.139943	-798.219109	-799.228057	-799.204883	-799.280449
N8-3	-798.161538	-798.136388	-798.217247	-799.222929	-799.199031	-799.275728
N8-4	-798.157144	-798.132224	-798.211825	-799.220208	-799.196730	-799.271812

N8-5	-798.153798	-798.127732	-798.211427	-799.216128	-799.191848	-799.269197
N8-6	-798.148418	-798.121456	-798.204793	-799.218898	-799.194595	-799.271357
A8-1	-798.188300	-798.164623	-798.240595	-799.257159	-799.235111	-799.306214
A8-2	-798.187137	-798.164146	-798.238940	-799.255176	-799.233530	-799.304364
A8-3	-798.185848	-798.161241	-798.239537	-799.259758	-799.237031	-799.309704
A8-4	-798.183823	-798.159937	-798.236418	-799.253597	-799.231106	-799.303659
A8-5	-798.183275	-798.158745	-798.237423	-799.254501	-799.230924	-799.306927
A8-6	-798.159575	-798.133224	-798.219079	-799.220786	-799.196067	-799.276370
HO(O)CC(O)OH (a)	-377.740081	-377.733751	-377.769296	-377.907811	-377.901354	-377.937092
HO(O)CC(O)OH (b)	-377.735422	-377.728959	-377.764993	-377.903046	-377.896436	-377.932690
HO(O)CC(O)OH (c)	-377.744290	-377.738098	-377.773209	-377.911935	-377.905613	-377.906557
HO(O)CC(O)OH (d)	-377.737984	-377.731469	-377.768256	-377.890837	-377.885119	-377.886063
$HO(O)CC(O)O^-$	-377.219098	-377.212608	-377.248985	-377.383262	-377.376593	-377.413570
$^-O(O)CC(O)O^-$	-376.527201	-376.521115	-376.555807	-376.685332	-376.679094	-376.714317
$^-O(O)CC(O)O^- + 6H_2O$	-834.556032	-834.533586	-834.610705	-834.937043	-834.915444	-834.988532
$^-O(O)CC(O)O^- + 8H_2O$	-986.369367	-986.342810	-986.428600	-987.657114	-987.632004	-987.714360

Table A4.2. Total Energies of Neutral and Anionic Species in the Gas Phase in Hartrees: CCSD(T) and CCSD(T)-F12b

	CCSD(T)/aD	CCSD(T)/aT	CCSD(T)-F12b/aD	CCSD(T)-F12b/aT
CO_2	-188.185856	-188.340554	-188.337594	-188.387126
CO_2 (135°)	-188.127849	-188.281597	-188.278330	-188.327589
CO_2^-	-188.161979	-188.317373	-188.313046	-188.364112
$^2HCO_2$	-188.695095	-188.851959	-188.847373	-188.898242
N3-1	-417.040097	-417.399646	-417.383881	-417.504827
N3-2	-417.040635	-417.399104	-417.383045	-417.503814
N3-3	-417.034173	-417.393624	-417.377771	-417.498738
A3-1	-417.057391	-417.417038	-417.399135	-417.521465
A3-2	-417.053828	-417.414077	-417.396874	-417.519224
A3-3	-417.052803	-417.413528	-417.395664	-417.518197
N4-1	-493.333529	-493.761366	-493.741151	-493.885851
N4-2	-493.333676	-493.761058	-493.740732	-493.885382
N4-3	-493.333875	-493.760647	-493.740134	-493.882352
N4-4	-493.326211	-493.753744	-493.733434	-493.878137
N4-5	-493.325377	-493.753046	-493.732914	-493.877626
A4-1	-493.354517		-493.759523	
A4-2	-493.348978		-493.755507	
A4-3	-493.344522		-493.749970	
A4-4	-493.342019		-493.749055	
A4-5	-493.341224		-493.747935	
A4-6	-493.337054		-493.742781	
A4-7	-493.321970		-493.730320	
A4-8	-493.316519		-493.725748	
A4-9	-493.313801		-493.722692	
N8-1	-798.501595		-799.160501	
N8-2	-798.499079		-799.158889	
N8-3	-798.495425		-799.154386	
N8-4	-798.492185		-799.151687	
N8-5	-798.486940		-799.146679	
N8-6	-798.480647		-799.139813	
A8-1	-798.531412			
A8-2	-798.529195			
A8-3	-798.526368			
A8-4	-798.525452			
A8-5	-798.523687			
A8-6	-798.492774			
HO(O)CC(O)OH (a)	-377.525642	-377.842477	-377.833809	-377.936466
HO(O)CC(O)OH (b)	-377.520827	-377.837300	-377.828549	-377.931179

HO(O)CC(O)OH (c)	-377.529676	-377.846846	-377.838161	-377.940910
HO(O)CC(O)OH (d)	-377.523566	-377.839959	-377.831260	-377.933821
$HO(O)CC(O)O^-$	-376.990082	-377.304017	-377.295823	-377.398264
$^-O(O)CC(O)O^-$	-376.284479	-376.595346	-376.587228	-376.689760
$^-O(O)CC(O)O^- + 6H_2O$	-834.136453		-834.822341	
$^-O(O)CC(O)O^- + 8H_2O$	-986.742712		-987.555474	

Table A4.3. Total Energies of Neutral and Anionic Species from Single Point Calculations in Aqueous Solution at the B3LYP/aT Level Using MP2/aT or aD Geometries in Hartrees

	COSMO (16)	SMD (16)
	HF_{aq}(298 K)	HF_{aq}(298 K)
CO_2	-188.666751	-188.659047
CO_2^-	-188.751432	-188.746333
$^2HCO_2$	-189.190224	-189.184114
HCO_3^-	-264.682510	-264.682265
N3-1	-418.103524	-418.098522
N3-2	-418.102253	-418.097368
N3-3	-418.099777	-418.096169
A3-1	-418.194154	-418.191844
A3-2	-418.191042	-418.188382
A3-3	-418.192512	-418.191129
N4-1	-494.591109	-494.585731
N4-2	-494.589301	-494.584666
N4-3	-494.590082	-494.584749
N4-4	-494.582113	-494.579286
N4-5	-494.583174	-494.579437
A4-1	-494.676377	-494.674661
A4-2	-494.679909	-494.676976
A4-3	-494.678942	-494.678588
A4-4	-494.679413	-494.670884
A4-5	-494.678896	-494.677097
A4-6	-494.673291	-494.670094
A4-7	-494.629957	-494.621504
A4-8	-494.631697	-494.622067
A4-9	-494.636241	-494.626877
N8-1	-800.526042	-800.520841
N8-2	-800.526215	-800.521153
N8-3	-800.519998	-800.516514
N8-4	-800.524553	-800.519158
N8-5	-800.514983	-800.514983
N8-6	-800.516619	-800.514335
A8-1	-800.610960	-800.609508
A8-2	-800.618554	-800.614601
A8-3	-800.616660	-800.615281
A8-4	-800.615326	-800.614892
A8-5	-800.617407	-800.615265
A8-6	-800.572797	-800.563004
HO(O)CC(O)OH (a)	-378.504189	-378.501206
HO(O)CC(O)OH (b)	-378.501536	-378.498716
HO(O)CC(O)OH (c)	-378.505329	-378.502106

HO(O)CC(O)OH (d)	-378.503656	-378.500682
$HO(O)CC(O)O^-$	-378.055209	-378.053029
$^-O(O)CC(O)O^-$	-377.594106	-377.597499
$^-O(O)CC(O)O^- + 6H_2O$	-836.480112	-836.483019
$^-O(O)CC(O)O^- + 8H_2O$	-989.452283	-989.457261

Table A4.4. Relative Energies (ΔG_{298}) of Neutral and Anionic Clusters in the Gas Phase in kcal/mol

Cluster	G3MP2 (B3)	MP2/ aT	CCSD(T)/ aD	CCSD(T)/ aT	CCSD(T)-F12b/aD	CCSD(T)-F12b/aT	CCSD(T)-F12b/CBS
N3-1[a]	0.2	0.0	0.0	0.0	0.0	0.0	0.0
N3-2 [b]	0.0	2.6	2.1	2.8	2.9	3.1	3.1
N3-3 [a]	3.4	2.8	2.5	2.5	2.6	2.6	2.6
A3-1 [b,d]	1.4	1.1	0.8	1.2	1.6	1.6	1.7
A3-2 [d]	0.0	0.0	0.0	0.0	0.0	0.0	0.0
A3-3 [d]	6.8	1.8	2.2	1.9	2.4	2.2	2.2
N4-1 [a]	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N4-2 [a]	0.2	0.9	0.6	0.9	1.0	1.0	1.0
N4-3 [b]	0.0	3.3	2.7	3.3	3.5	5.1	5.6
N4-4 [a]	5.3	4.8	4.3	4.5	4.5	4.5	4.5
N4-5 [a]	5.0	4.4	3.9	4.0	4.0	4.0	4.0
A4-1 [b,d]	0.0	1.3	0.0		1.2		
A4-2 [d]	3.0	2.1	1.8		2.0		
A4-3 [d]	7.8	6.4	5.5		6.4		
A4-4 [d]	0.7	0.0	0.1		0.0		
A4-5 [d]	3.9	3.2	3.4		3.5		
A4-6 [d]	1.3	1.8	8.7		9.4		
A4-7 [e]	[f]	14.2	15.9		16.8		
A4-8 [e]	[f]	15.6	17.6		18.0		
A4-9 [e]	[f]	9.8	13.3		12.1		
N8-1	0.2	10.4 [g]	0.0		0.4		
N8-2	0.0	0.0 [g]	0.2		0.0		
N8-3 [c]	3.0	1.4 [g]	1.1		1.5		
N8-4	5.4	6.5 [g]	4.3		4.3		
N8-5 [c]	7.1	6.6 [g]	4.6		4.5		
N8-6 [c]	5.7	9.9 [g]	8.5		8.7		
A8-1 [d]	2.2	0.0 [g]	0.0				
A8-2 [d]	3.4	0.7 [g]	1.9				
A8-3 [c,d]	0.0	0.3 [g]	1.1				
A8-4 [d]	3.8	2.1 [g]	3.1				
A8-5 [c,d]	1.7	1.8 [g]	2.6				
A8-6 [c,e]	20.9	14.3 [g]	16.9				

[a] Labeling compared to structures from Nguyen et al.: **N3-1** = R3-3; **N3-3** = R3-2; **N4-1** = R4-4-1; **N4-2** = R4-4-2; **N4-5** = R4-2; **N4-4** = R4-3. [b] Structures relabeled from Balaj et al. lowest energy clusters with 3 or 4 H_2O molecules: **A3-1**, **A4-1**. [c] Structures derived from $e^{-}\cdot(H_2O)_8$ as published by Zhan and Dixon. [d] CO_2 spin. [e] H_2O spin. [f] Will not optimize at G3MP2(B3) level with the radical electron density on the CO_2 molecules. [g] MP2/aT values are calculated with MP2/aT single point energy values and MP2/aD free energy corrections at 298 K.

Discussion of Implicit Solvation During Optimization

Table A4.5. OCO Bond Angles and C=O Bond Lengths for Gas Phase Optimized Structures and Aqueous Phase Optimized Structures

Species	B3LYP/aT		B3LYP/aT/COSMO		B3LYP/aT/SMD	
	OCO Angle (Degrees)	C=O Bond Length (Å)	OCO Angle (Degrees)	C=O Bond Length (Å)	OCO Angle (Degrees)	C=O Bond Length (Å)
CO_2^-	138.4	1.23	135.2	1.23	134.6	1.24
N3-1	178.2	1.17	179.0	1.16	179.5	1.16
N3-3	176.2	1.17	180.0	1.16	179.4	1.16
A3-1	135.0	1.23, 1.24	134.5	1.23, 1.24	134.5	1.23, 1.24
A3-2	132.3	1.25	134.3	1.22, 1.25	134.1	1.23, 1.24
N4-1	178.2	1.17	179.0	1.16	180.0	1.16
A4-1	134.8	1.24	134.9	1.24	134.6	1.24
A4-3	134.9	1.23, 1.24	134.5	1.23, 1.24	134.5	1.23, 1.24
A4-4	134.6	1.23	133.8	1.24	133.7	1.24
N8-1	178.4	1.16	178.9	1.16	179.6	1.16
N8-4	179.6	1.16	179.8	1.16	179.9	1.16
N8-6	178.2	1.16	180.0	1.16	179.5	1.16
A8-1	134.7	1.23, 1.24	135.0	1.24	134.4	1.23, 1.24
A8-3	136.2	1.22, 1.24	135.1	1.23, 1.24	135.0	1.23, 1.24
A8-5	136.3	1.23	135.0	1.23	137.4	1.23, 1.24
$C_2O_4^{2-}$	126.3	1.27	126.2	1.26	126.1	1.26
$[C_2O_4 \cdot 6H_2O]^{2-}$	125.5	1.27	127.0	1.26	126.7	1.26
$[C_2O_4 \cdot 8H_2O]^{2-}$	126.0	1.28	126.6	1.25, 1.27	126.4	1.25, 1.27

Table A4.6. Relative Aqueous Free Energy for DFT-Optimized Clusters at 298 K in kcal/mol

Species	Gas Phase Optimized		Aqueous Phase Optimized	
	B3LYP/aT/COSMO	B3LYP/aT/SMD	B3LYP/aT/COSMO	B3LYP/aT/SMD
N3-1	1.5	2.4	1.3	1.4
N3-3	0.0	0.0	0.0	0.0
A3-1	1.2	1.0	3.7	2.1
A3-2	0.0	0.0	0.0	0.0
N4-1	0.0	0.0	0.0	0.0
A4-1	8.5	4.2	3.8	3.3
A4-3	7.0	1.9	4.4	3.7
A4-4	0.0	0.0	0.0	0.0
N8-1	2.0	1.8	1.1	0.9
N8-4	0.0	0.0	3.0	0.0
N8-6	4.2	2.2	0.0	2.0
A8-1	6.4	6.0	0.3	1.6
A8-3	1.3	0.8	1.6	0.3
A8-5	0.0	0.0	0.0	0.0

The energetic results in the paper are based on the gas-phase optimized MP2/aT geometries with a solvation correction at the B3LYP/aT/COSMO and B3LYP/aT/SMD level with the same MP2/aT geometries. In order to examine the effect of optimizing in the presence of solvent, the optimized gas phase clusters with the lowest gaseous free energies and single point aqueous free energies were reoptimized at the B3LYP/aT/COSMO and B3LYP/aT/SMD levels. Frequencies were also calculated to confirm that these aqueous-optimized structures were energetic minima. The relative energies are given in Table 4.2. The lowest free energy (ΔG_{aq}) clusters for both neutral and anionic $CO_2 \cdot (H_2O)_x$ where x = 3 and 4 are the same for both gas-optimized and aqueous-optimized clusters (**N3-3**, **A3-2**, **N4-1**, and **A4-4**) at the B3LYP/aT level. The lowest aqueous free energy MP2/aT geometries used in the CCSD(T) calculations are very similar to the lowest energy ones optimized in the gas phase or in aqueous solution for three and four waters.

N8-4 has the lowest aqueous free energy for the neutral $CO_2 \cdot (H_2O)_8$ clusters at the reoptimized B3LYP/aT/SMD level as well as at the gas phase optimized B3LYP/aT and MP2/aT levels. **N8-6** is the lowest energy for aqueous-optimized B3LYP/aT/COSMO. The energy

differences between **N8-4** and **N8-6** are 4.2, 2.2, 2.0, 1.5, and 1.7 kcal/mol at the B3LYP/aT/COSMO (gas optimized), B3LYP/aT/SMD (gas optimized), B3LYP/aT/SMD(aq optimized), MP2/aT/COSMO(gas opt), and CCSD(T)/aD(gas opt) levels. At the B3LYP/aT/COSMO(aq opt) level, the energy difference is -3.0 kcal/mol, which should not significantly affect the results. **A8-5** is the anionic $CO_2^{-\cdot}(H_2O)_8$ cluster with the lowest free energy (ΔG_{aq}) for both the same for both gas phase optimized and aqueous optimized clusters at the B3LYP/aT level.

Table A4.7. Adiabatic Electron Affinities for Gas Phase and Aqueous Phase Optimized Clusters at 298 K in kcal/mol

Method	Opt	N3-3/A3-2	N4-1/A4-4	N8-4/A8-5	N8-6/A8-5
B3LYP/aT/COSMO	gas	57.3	61.7	57.2	61.4
B3LYP/aT/SMD	gas	57.9	59.7	59.3	61.5
B3LYP/aT/COSMO	aq	59.5	60.0	60.4	57.4
B3LYP/aT/SMD	aq	60.2	61.1	59.5	61.5

Gas phase and aqueous phase optimized ΔG_{aq} electron affinities are given in Table A4.7. The use of geometries from gas phase or aqueous phase optimization at the DFT level does not significantly the electron affinity calculations. We note that the DFT geometries were not used in the actual MP2 and CCSD(T) calculations of the electron affinities as the gas phase optimized MP2/aT geometries were used for the correlated molecular orbital theory calculations.

Table A4.8. Vertical Detachment Energies[a] for $CO_2^- \rightarrow CO_2 + e^-$ in Gas Phase in kcal/mol

	A3-1	A3-2	A4-1	A4-3	A4-4	A8-3	A8-5
MP2/aT	62.7	63.4	70.4	53.0	73.9	89.7	86.2
CCSD(T)/aD	64.8	65.5	72.4	58.5	75.9	94.8	91.4
CCSD(T)/aT	68.1	68.8					
CCSD(T)-F12b/aD	68.4	69.1	76.1	61.4	79.4		
CCSD(T)-F12b/aT	69.8	70.4					

[a] Negative VDE (ΔE_{elec}) value indicates a bound electron. **A3-1** and **A4-1** are included for comparison because they are the lowest energy clusters of their types in the gas phase.

Table A4.9. Vertical Attachment Energies[a] for $CO_2 + e^- \rightarrow CO_2^-$ in Gas Phase kcal/mol

Method	N3-1	N3-3	N4-1	N4-4	N8-4	N8-6
MP2/aT	-12.0	-12.7	-11.5	-6.8	-2.2	-6.5
CCSD(T)/aD	-12.3	-13.5	-12.1	-11.0	0.1	-4.2
CCSD(T)/aT	-9.6	-10.5				
CCSD(T)-F12b/aD	-10.7	-11.7	-10.5	-9.7		
CCSD(T)-F12b/aT	-8.7	-9.4				

[a] Positive VAE ($-\Delta E_{elec}$) value indicates a bound electron. **N3-1** and **N4-1** are included for comparison because they are the lowest energy clusters of their types in the gas phase.

Table A4.10. Raw Values and CBS Extrapolation for FPD Approximations in Hartrees

Method	Basis Set	Oxalic Acid (C)	Carbonyl Radical
CCSD(T)-F12b	aD	-377.838161	-188.847373
CCSD(T)-F12b	aT	-377.940910	-188.898242
CCSD(T)-F12b	aQ	-377.983404	-188.919348
CCSD(T)-F12b	CBS (DTQ)	-378.008579	-188.929023
CCSD(T)/Core	awT	-378.177791	-189.017351
CCSD(T)/Valence	awT	-377.869160	-188.863364
CCSD(T)-DK/Valence	awT-DK	-378.106214	-188.982009

Table A4.11. HOMO and LUMO Energies for CO_2^- and Oxalate Dianion ($C_2O_4^{2-}$) with and without Explicit H_2O Molecules in Hartrees

Species	B3LYP/aT		HF/aT		B3LYP/aT/COSMO		B3LYP/aT/SMD	
	HOMO	LUMO	HOMO	LUMO	HOMO	LUMO	HOMO	LUMO
CO_2^-	0.05238	0.13661	-0.09153	0.20784	-0.15532	-0.04317	-0.16503	-0.04812
$C_2O_4^{2-}$	0.13558	0.25063	-0.00722	0.28837	-0.20947	0.01107	-0.22320	0.01424
$[C_2O_4 \cdot 6H_2O]^{2-}$	0.01363	0.17312	-0.13342	0.20369	-0.23032	-0.01274	-0.23688	-0.00923
$[C_2O_4 \cdot 8H_2O]^{2-}$	-0.00974	0.16628	-0.15820	0.19789	-0.23494	-0.01564	-0.24021	-0.01152

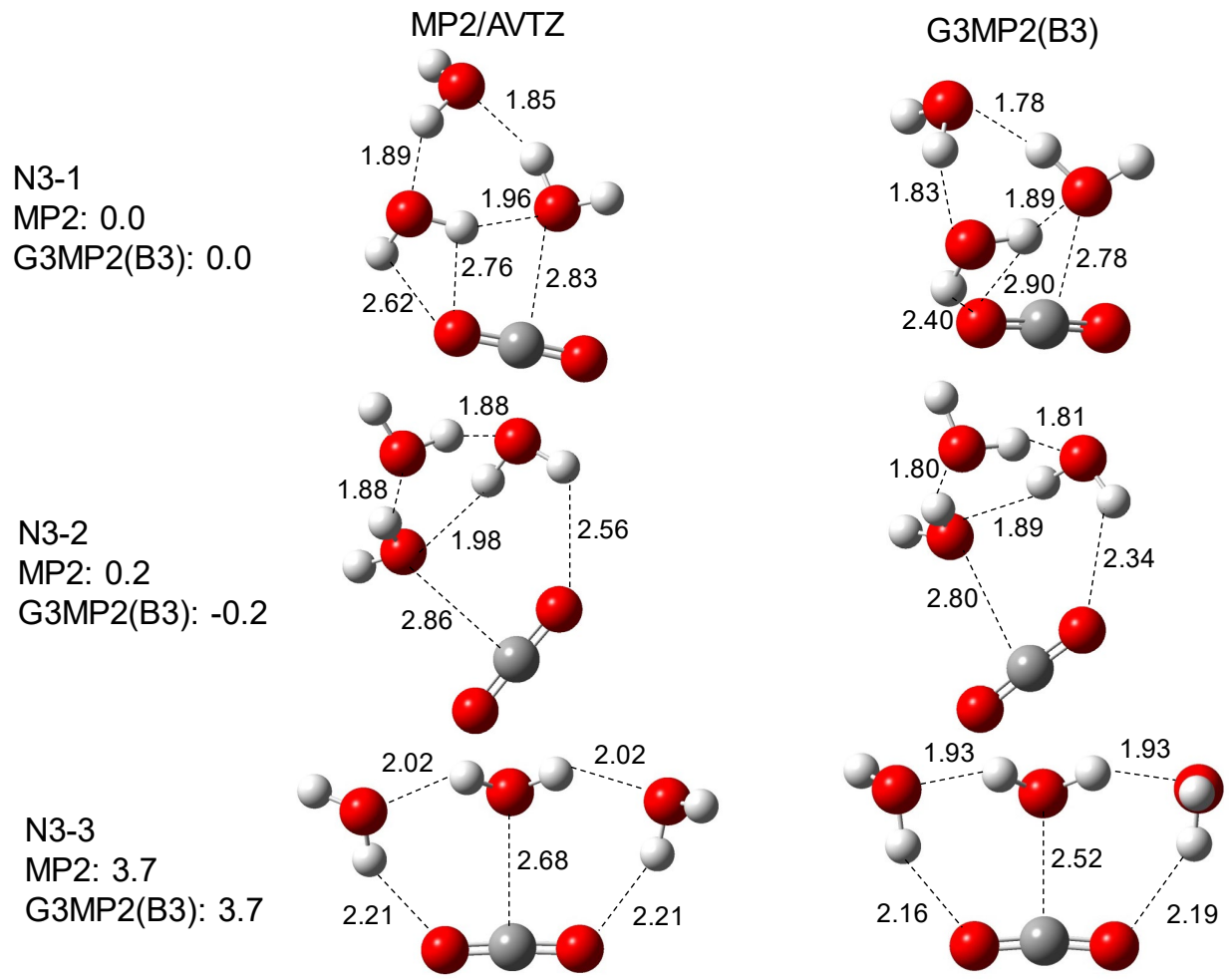


Figure A4.1. $CO_2 \cdot (H_2O)_3$ structures at the MP2/aT and G3MP2(B3) levels.

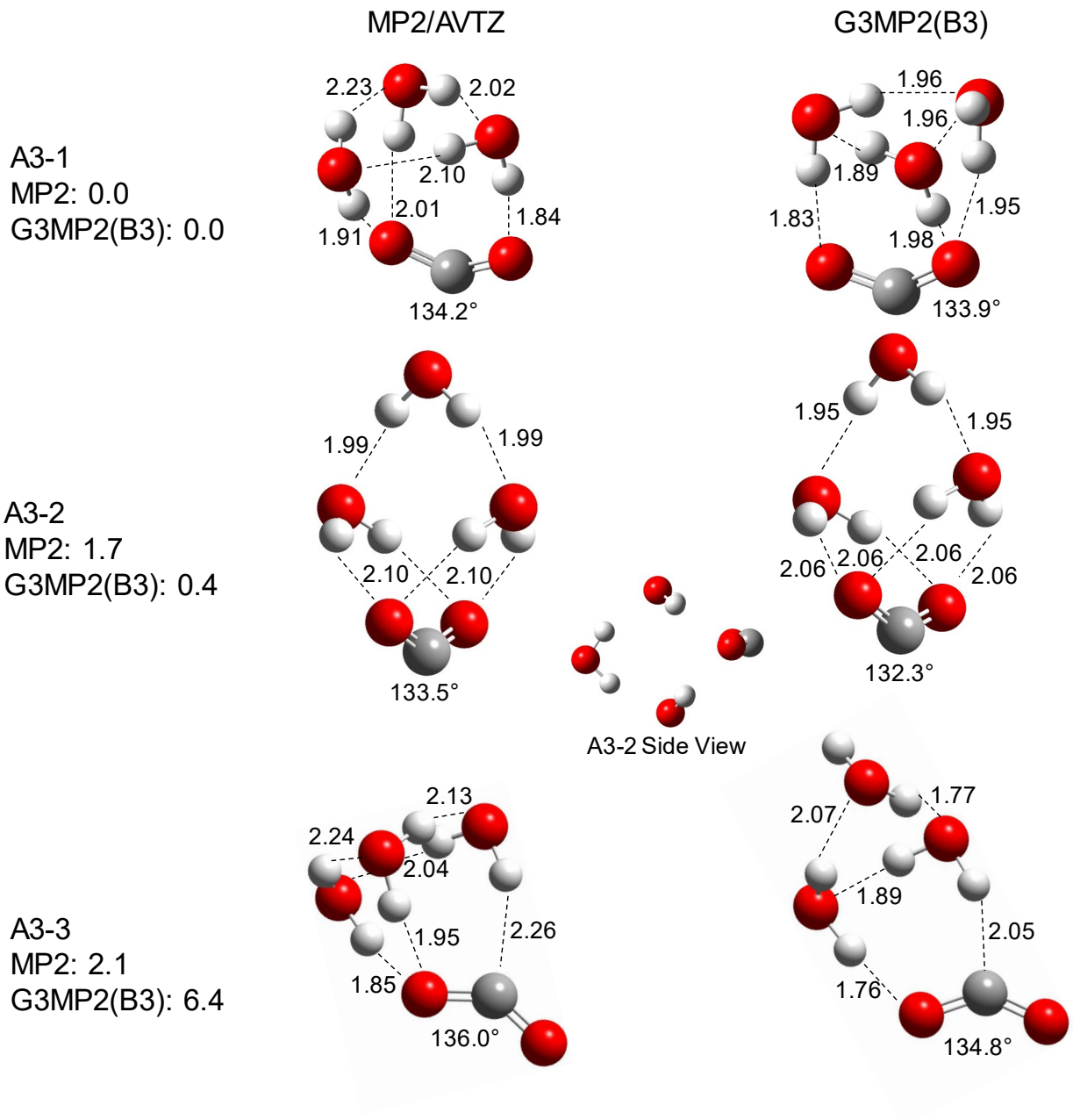


Figure A4.2. $CO_2^{-\cdot}(H_2O)_3$ structures at the MP2/aT and G3MP2(B3) levels.

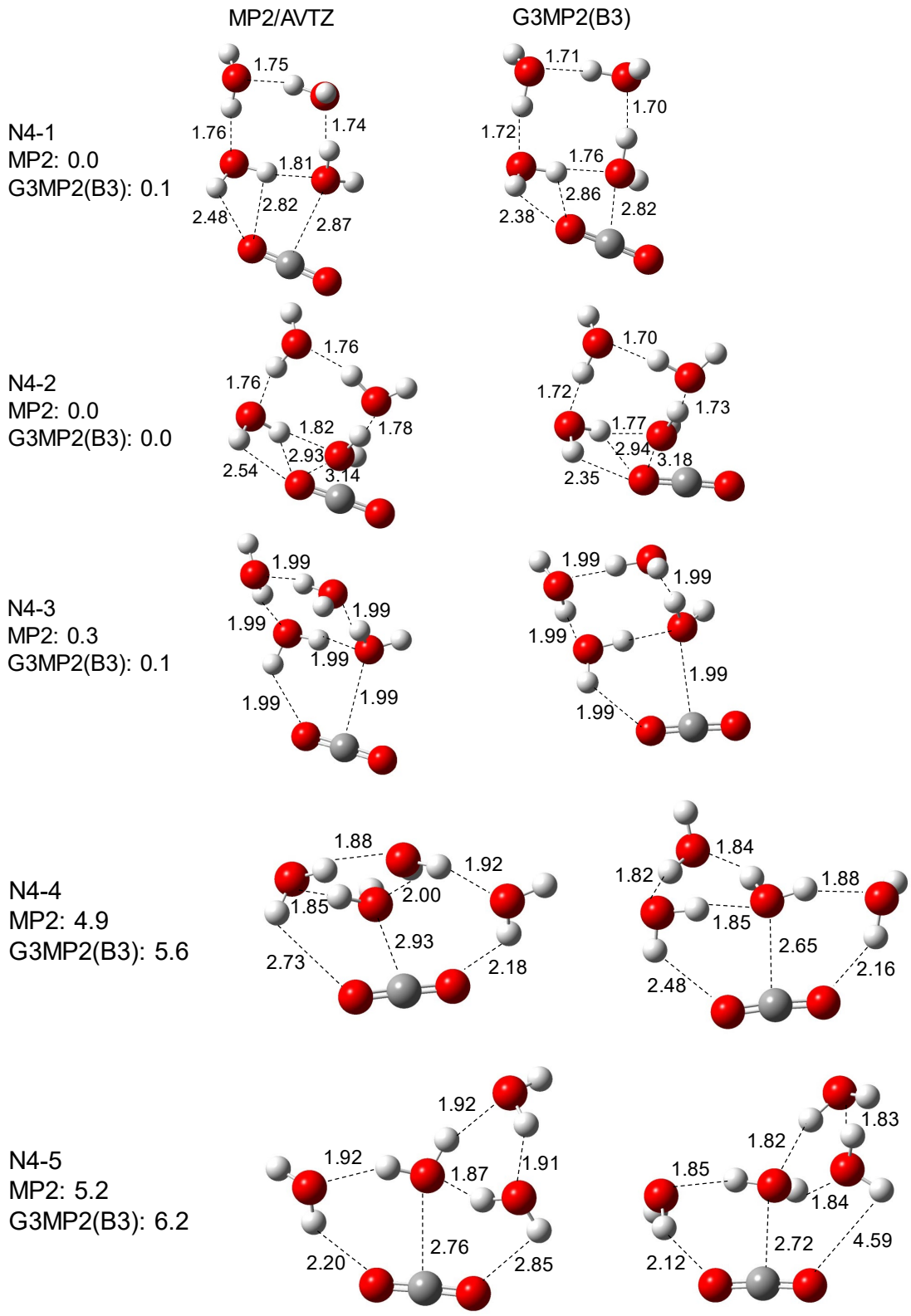
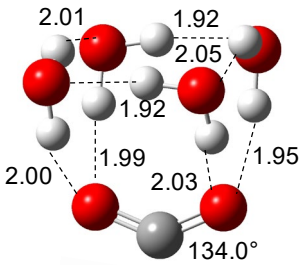
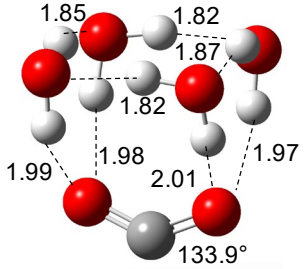
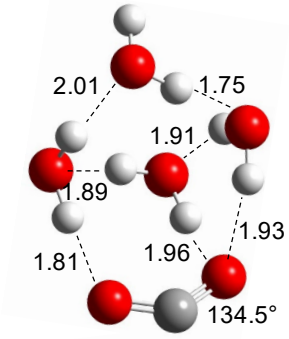
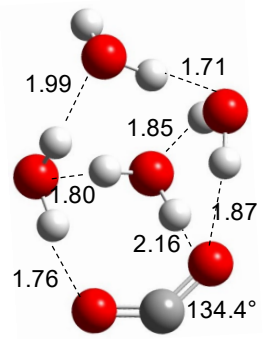
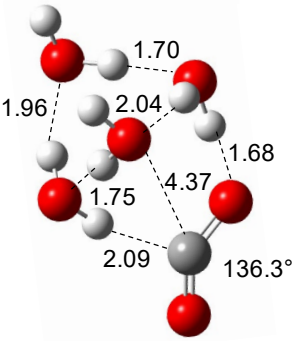
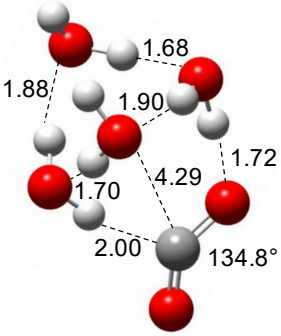
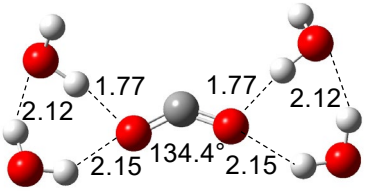
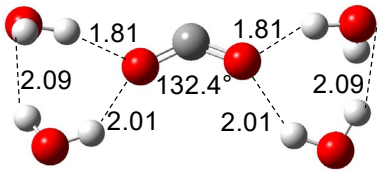
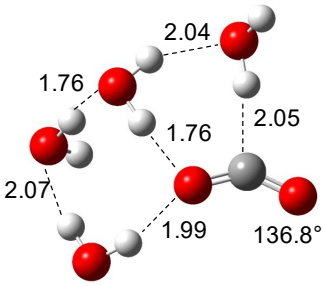
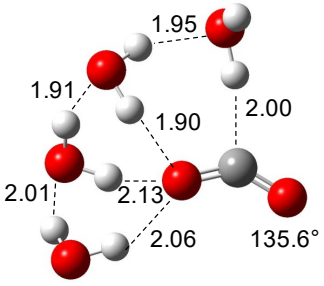
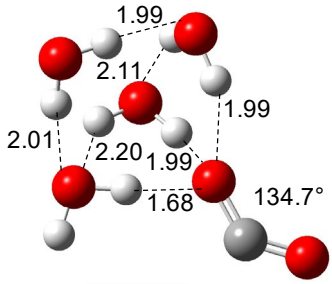
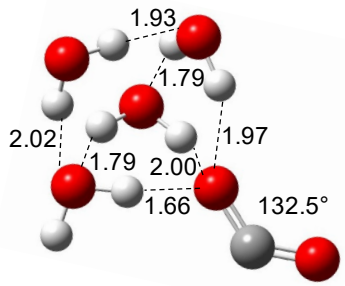


Figure A4.3. $CO_2 \cdot (H_2O)_4$ structures at the MP2/aT and G3MP2(B3) levels.

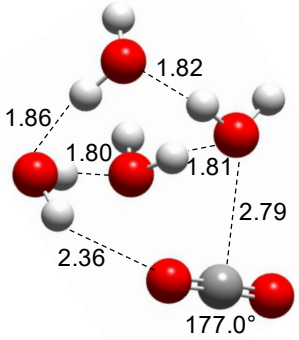
	MP2/AVTZ	G3MP2(B3)
A4-1 MP2: 0.0 G3MP2(B3): 0.0		
A4-2 MP2: 2.4 G3MP2(B3): 2.8		
A4-3 MP2: 7.5 G3MP2(B3): 8.6		
A4-4 MP2: 5.8 G3MP2(B3): 7.6		
A4-5 MP2: 6.3 G3MP2(B3): 6.5		

A4-6
MP2: 4.1
G3MP2(B3): 3.9

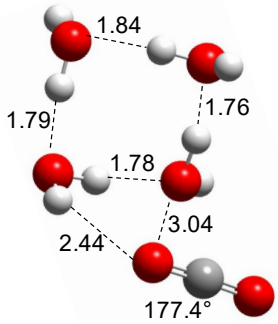




A4-7
MP2: 15.8



A4-8
MP2: 18.7



A4-9
MP2: 19.9

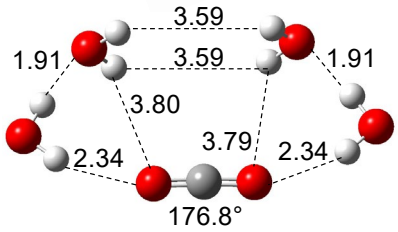
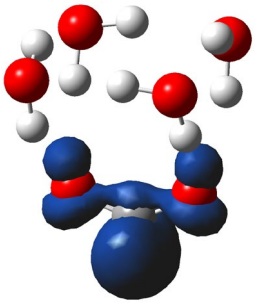


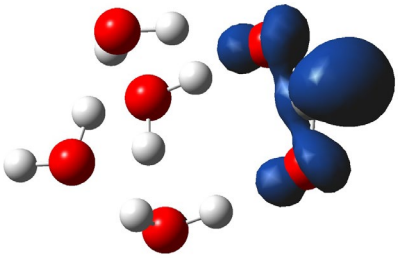
Figure A4.4. $CO_2^{-\cdot}(H_2O)_4$ structures at the MP2/aT and G3MP2(B3) levels.

B3LYP/AVTZ

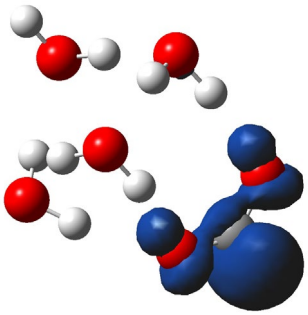
A4-1



A4-2



A4-3



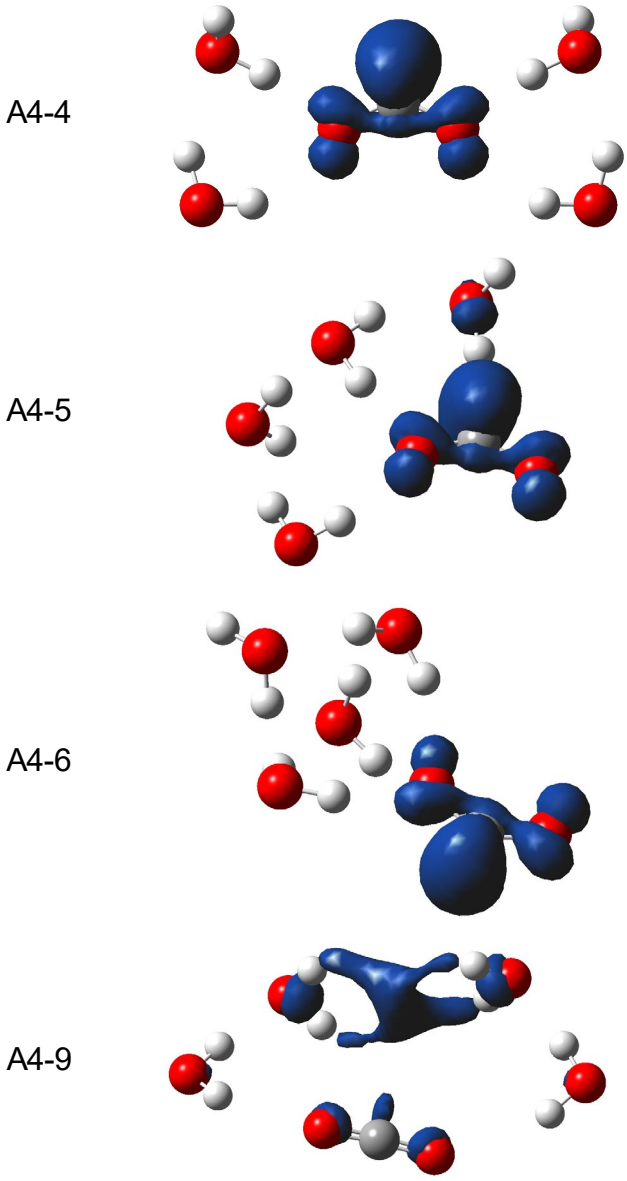


Figure A4.5. $CO_2^{-\cdot}(H_2O)_4$ (B3LYP/aT) electron density (Contour = 0.005000 for **A4-1** to **A4-6**. Contour level = 0.001700 for **A4-9**.) Note that **A4-7** and **A4-8** optimize to **A4-6** and **A4-3**, respectively, at the B3LYP/aT level.

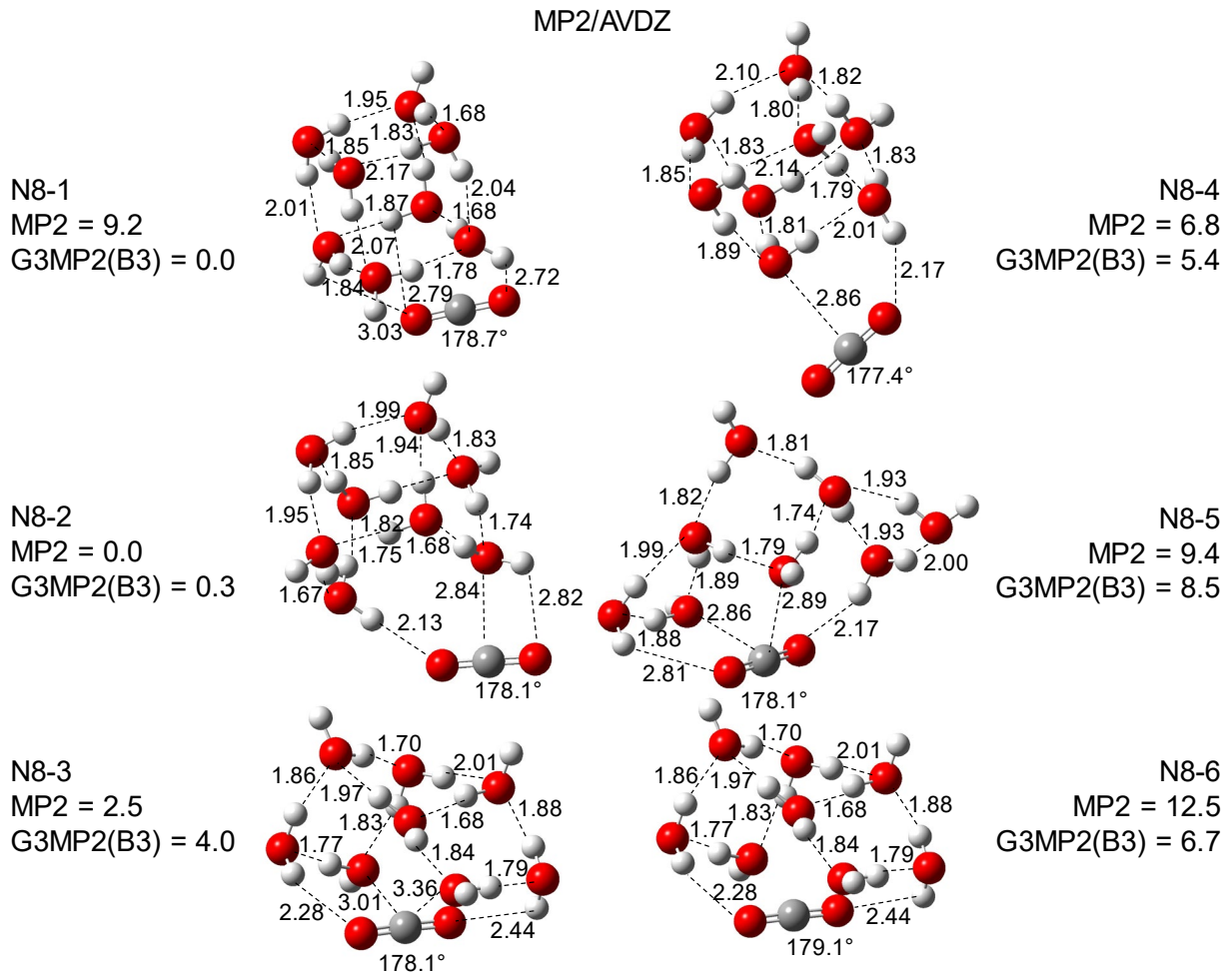
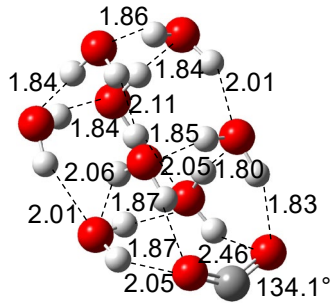
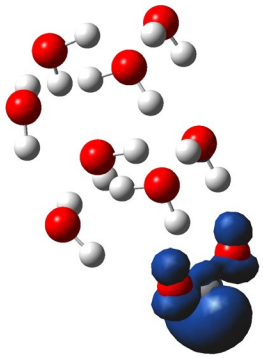


Figure A4.6. $CO_2 \cdot (H_2O)_8$ structures at the MP2/aD level.

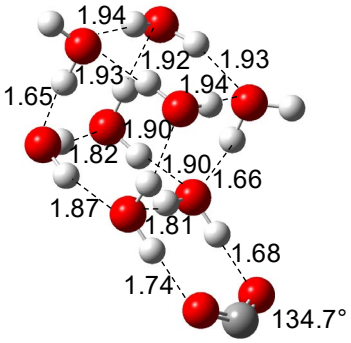
MP2/AVTZ B3LYP/AVTZ

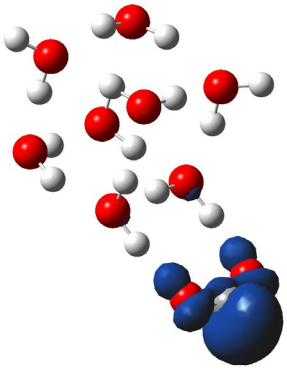
A8-1



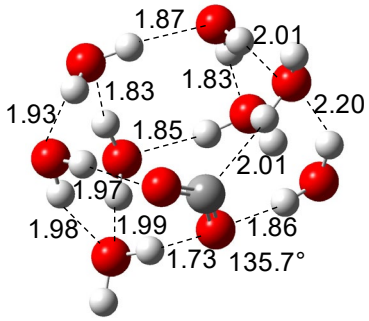


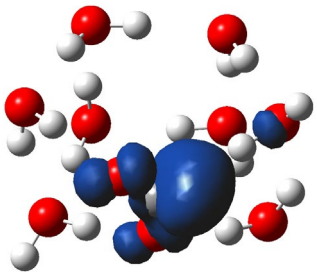
A8-2



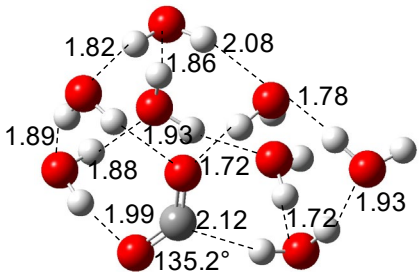


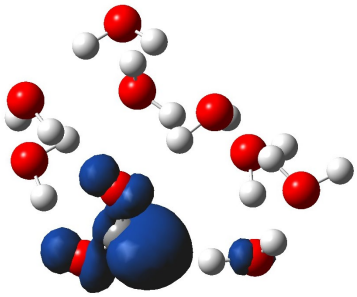
A8-3





A8-4





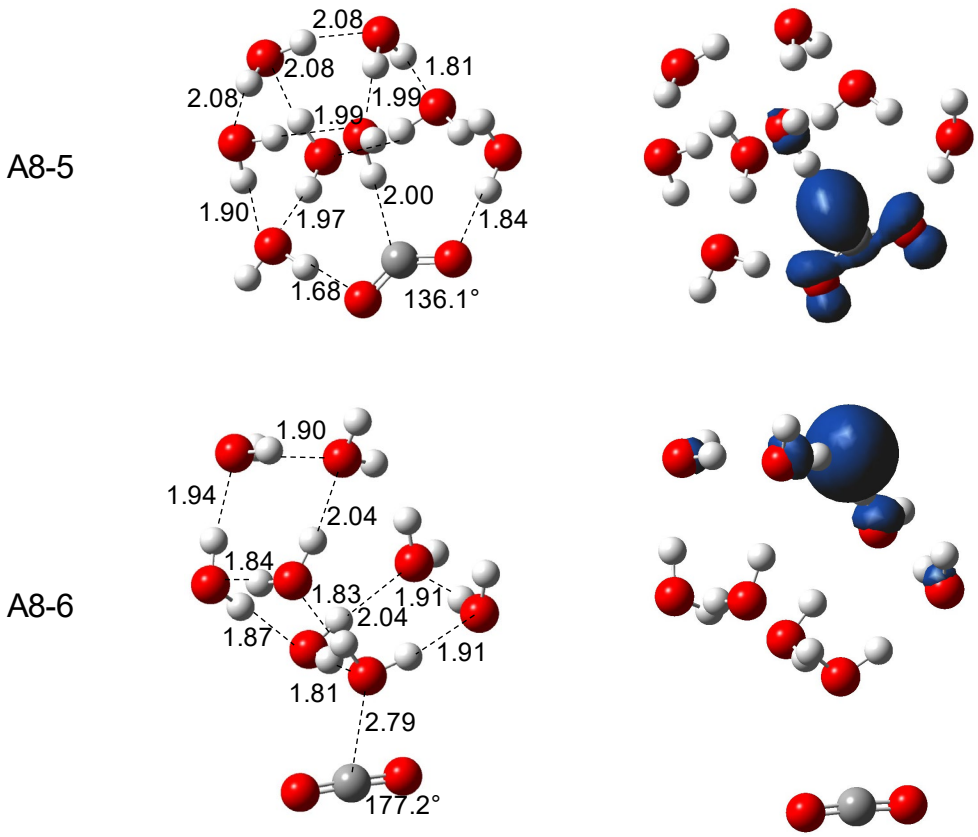


Figure A4.7. $CO_2^{-\cdot}(H_2O)_8$ structures at the MP2/aD level with B3LYP/aT electron densities. (Contour = 0.005000 for **A8-1** to **A8-5**. Contour = 0.002000 for **A8-6**.)

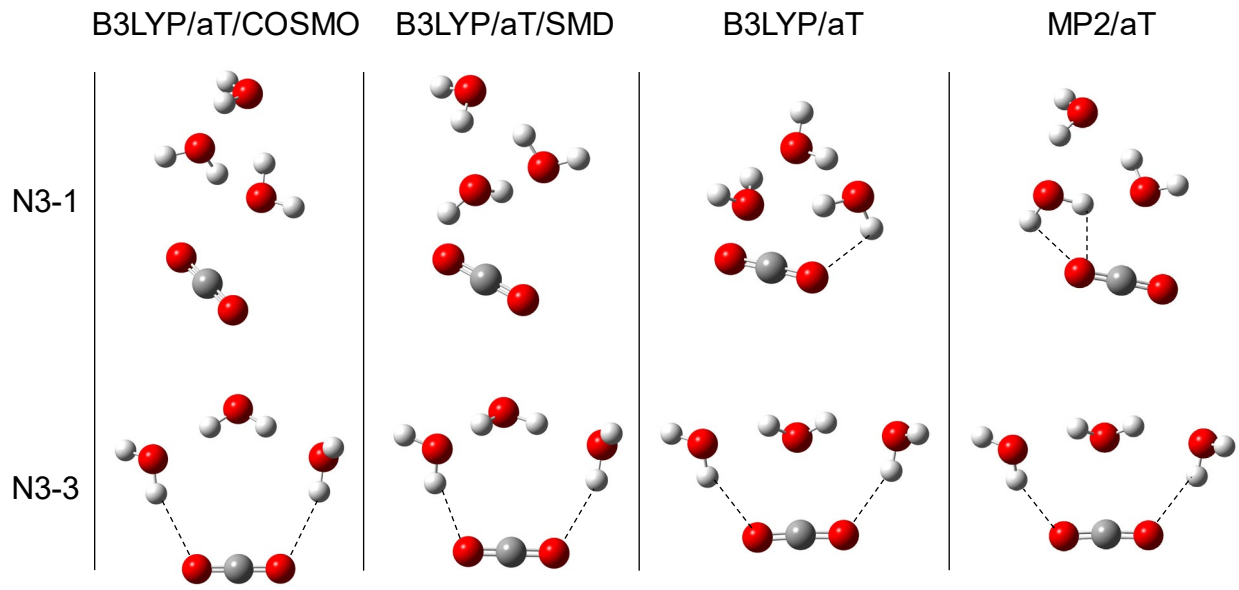


Figure A4.8. Selected $CO_2 \cdot (H_2O)_3$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries.

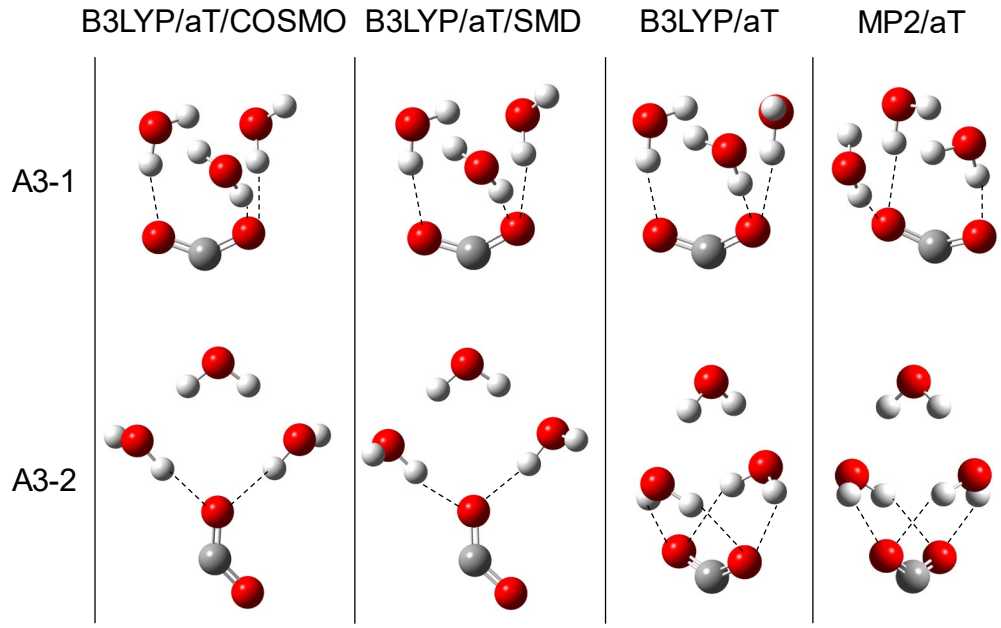


Figure A4.9. Selected $CO_2^{-\cdot} (H_2O)_3$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries.

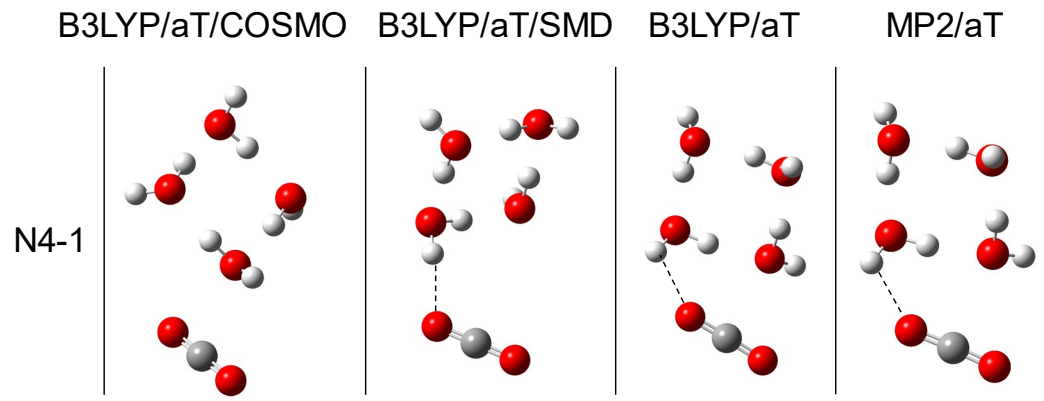


Figure A4.10. Selected $CO_2^{-}\cdot(H_2O)_4$ cluster for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries.

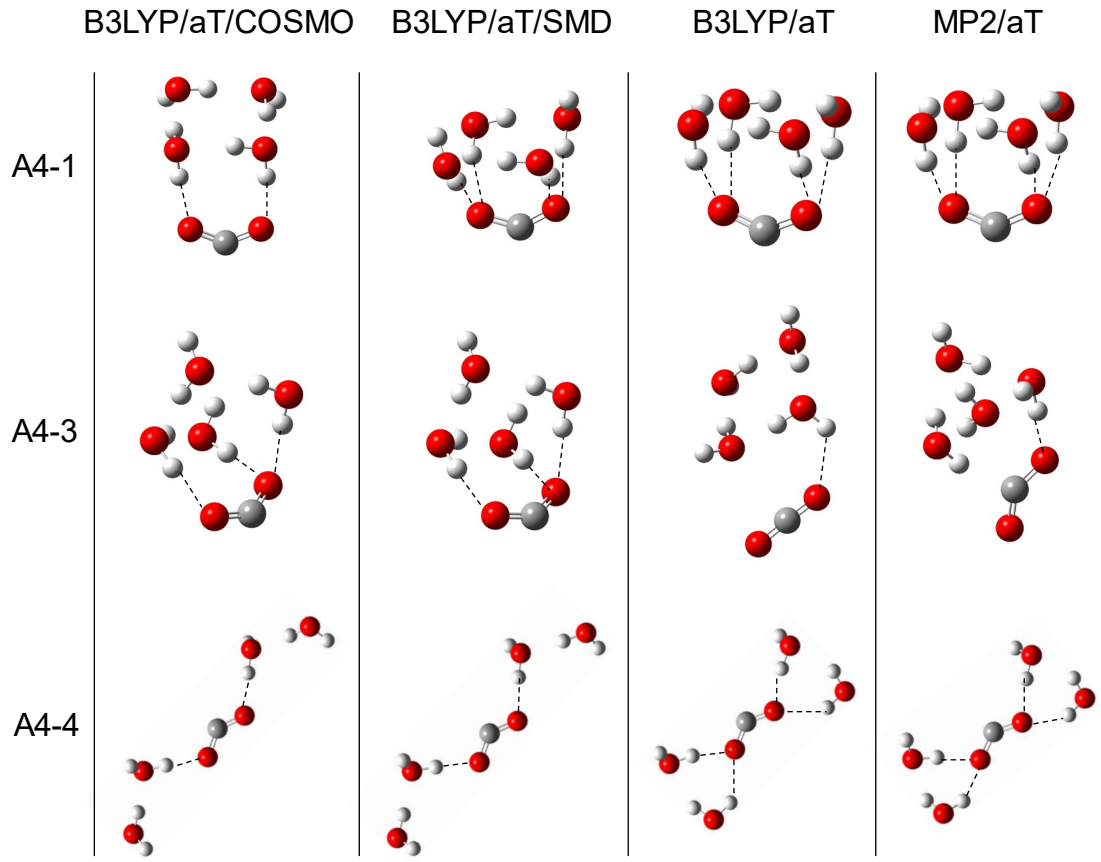


Figure A4.11. Selected $CO_2^{-}\cdot(H_2O)_4$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries.

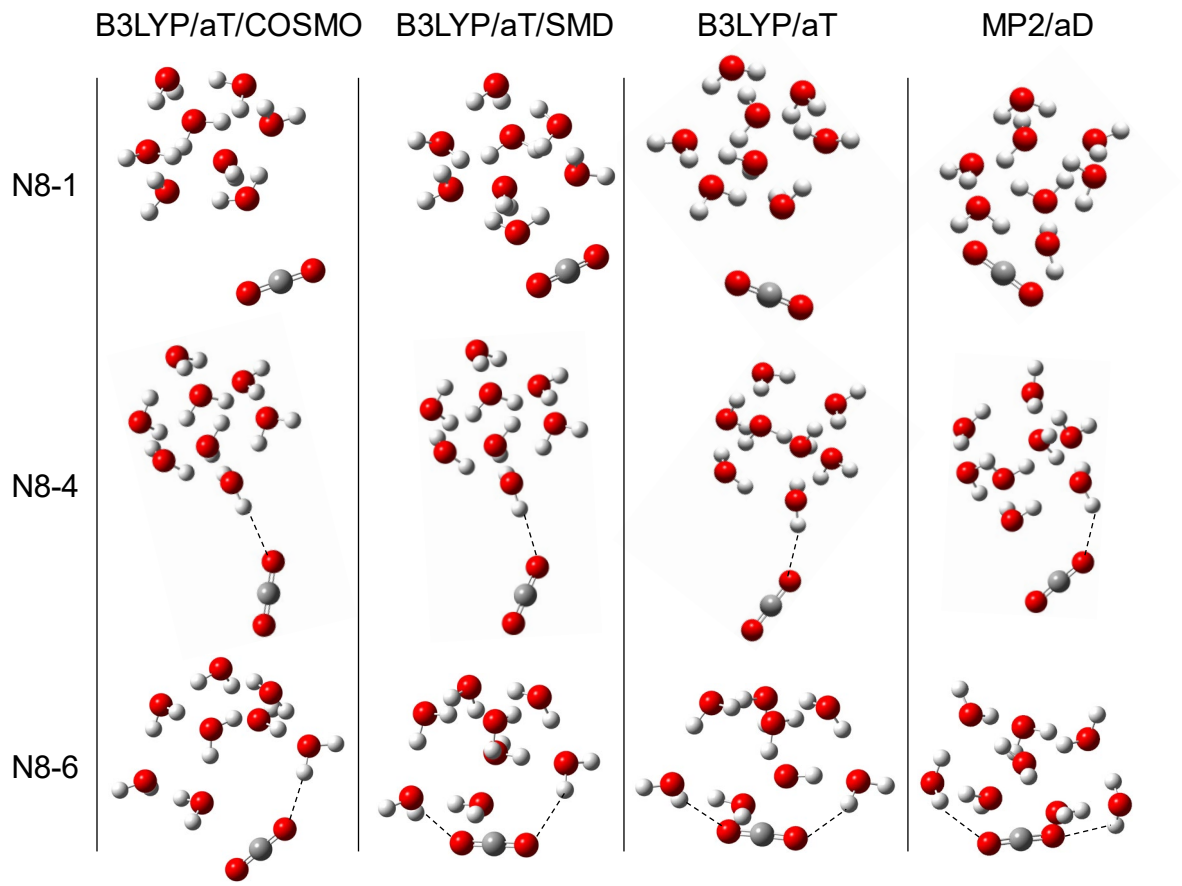


Figure A4.12. Selected $CO_2 \cdot (H_2O)_8$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries.

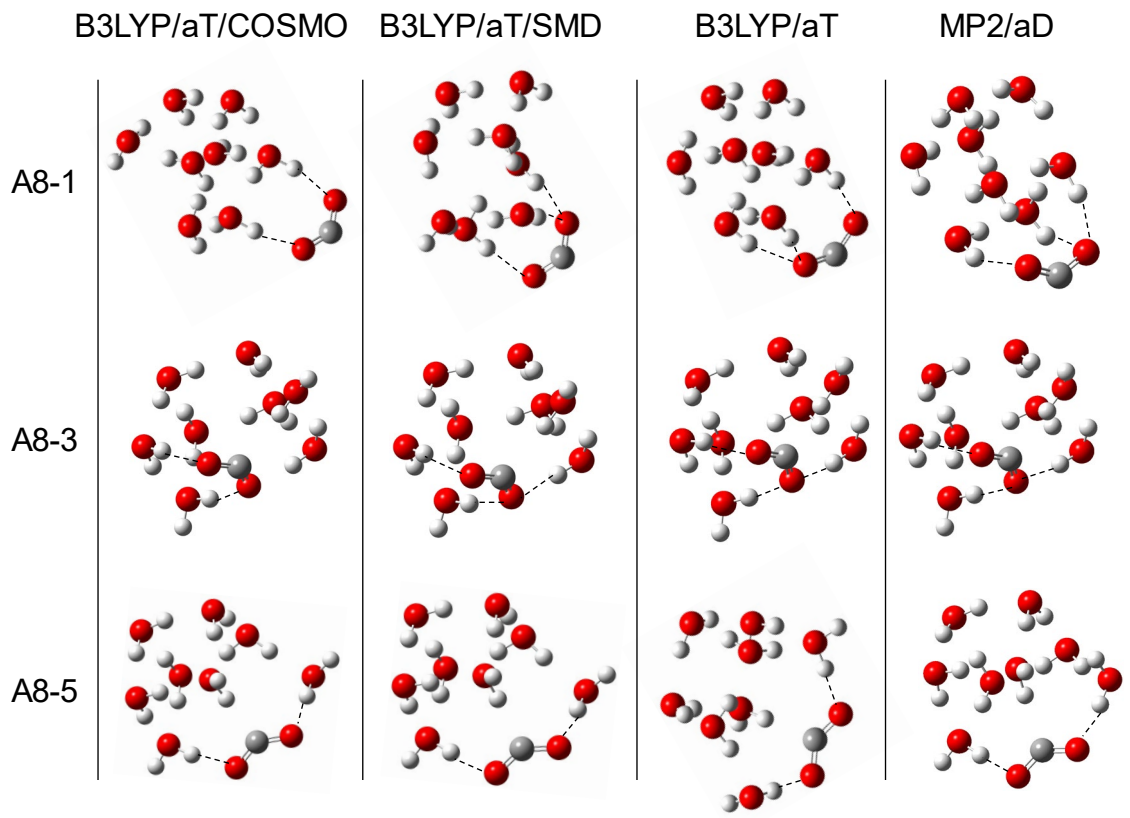


Figure A4.13. Selected $CO_2^{-\cdot}(H_2O)_8$ clusters for comparison of aqueous phase optimized (B3LYP/aT/COSMO, B3LYP/aT/SMD) and gas phase optimized (B3LYP/aT, MP2/aT) geometries.

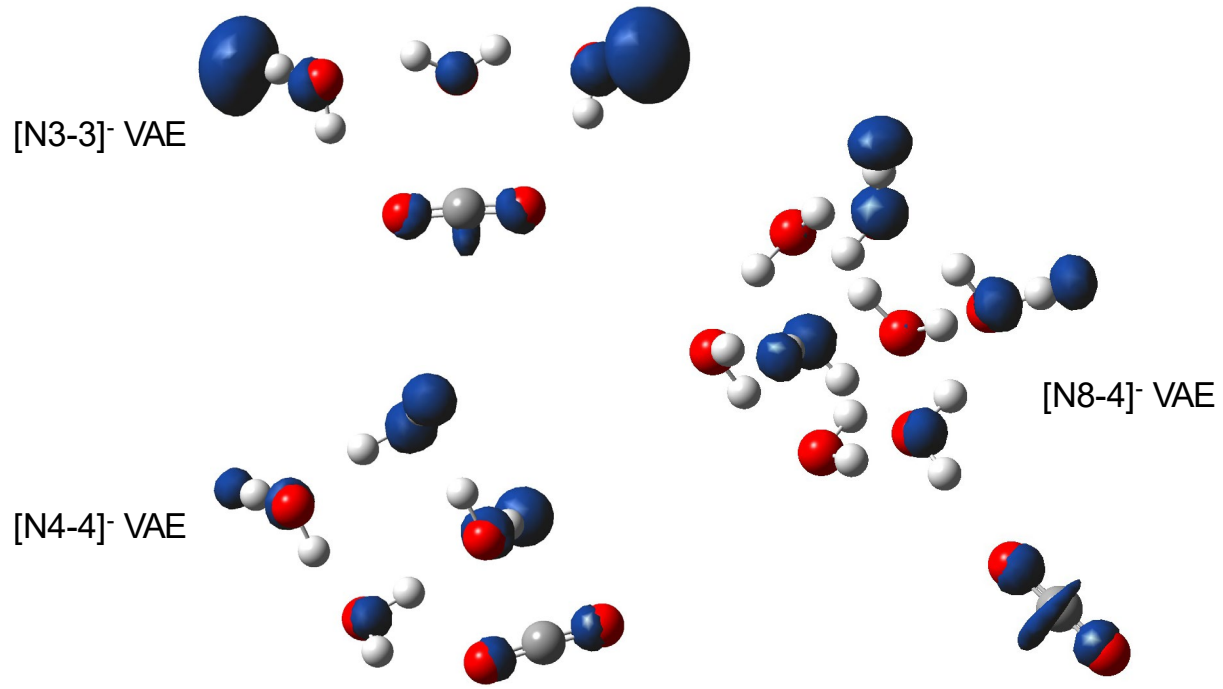
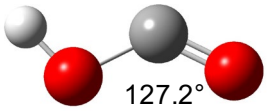
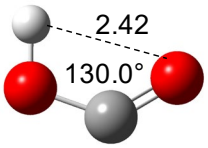


Figure A4.14. Electron spin density (B3LYP/aT) for the VAE products of **N3-3**, **N4-4**, and **N8-4** (contour = 0.0008).

·C(O)OH



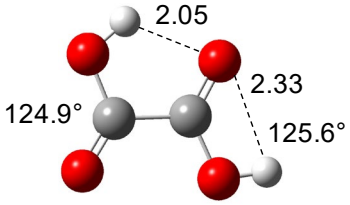
ΔH_{298} = 0.0
ΔG_{298} = 0.0

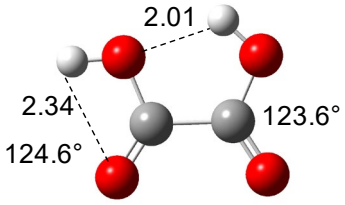


ΔH_{298} = 1.8
ΔG_{298} = 1.8

HO(O)C-C(O)OH

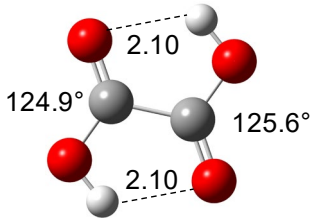
Oxalic Acid, a
ΔH_{298} = 2.7
ΔG_{298} = 2.4

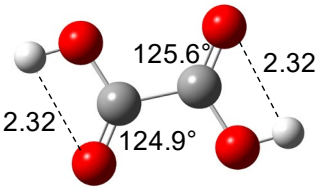




Oxalic Acid, b
ΔH_{298} = 5.8
ΔG_{298} = 5.2

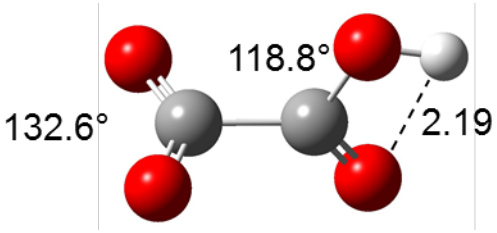
Oxalic Acid, c
ΔH_{298} = 0.0
ΔG_{298} = 0.0



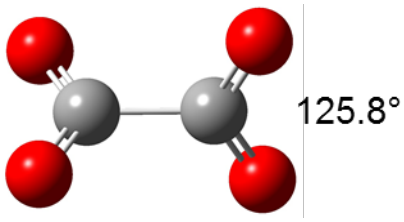


Oxalic Acid, d
ΔH_{298} = 12.9
ΔG_{298} = 13.7

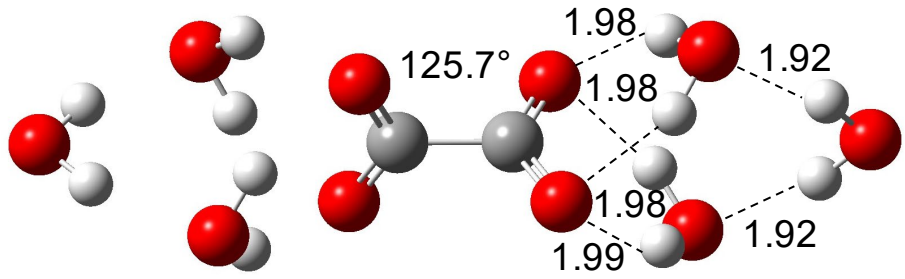
Hydrogen Oxalate, HO(O)C-C(O)O^-



Oxalate, $^{-}O(O)C\text{-}C(O)O^{-}$



Oxalate with 6 H_2O based on A3-2, $(H_2O)_3 \cdot {}^{-}O(O)C\text{-}C(O)O^{-} \cdot (H_2O)_3$



Oxalate with 8 H_2O based on A4-3, $(H_2O)_4 \cdot {}^{-}O(O)C\text{-}C(O)O^{-} \cdot (H_2O)_4$

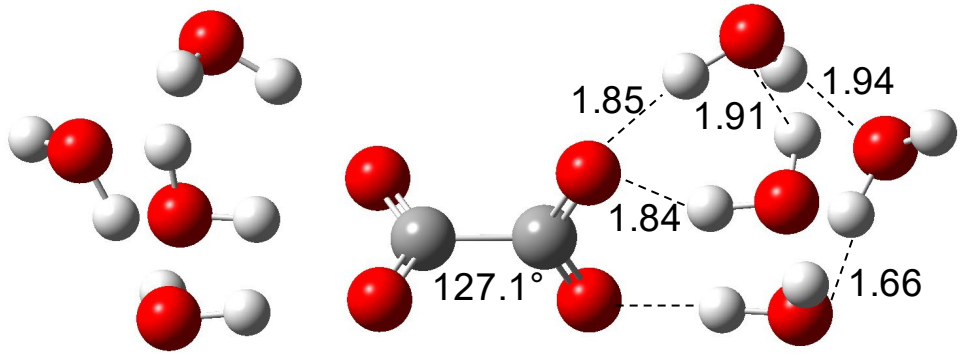


Figure A4.15. Oxalate-related compounds at the MP2/aT level with relative energies in kcal/mol.

CHAPTER 5

REACTION ENERGETICS AND ^{13}C FRACTIONATION OF ALANINE TRANSAMINATION IN THE AQUEOUS AND GAS PHASES

From McNeill, A. S.; Dallas, B.; Eiler, J.; Bylaska, E. J.; Dixon, D. A. *J. Phys. Chem. A* **2020**, *124*, 2077 – 2089.

Introduction

The alanine transaminase (ALT) enzyme, which can be found in plasma and other tissues but is most common in the liver, was first characterized in 1955.[1] ALT activity is commonly included in clinical liver function tests as a significantly increased level of ALT is an indicator of liver disease.[2] ALT activity is also influenced by other various factors including drug consumption, viral hepatitis, and metabolic abnormalities.[3] The ALT enzyme catalyzes the reversible transfer of an amino group from alanine to α-ketoglutarate to produce pyruvate and glutamate. An equilibrium isotopic fractionation is expected for the isotope exchange reaction between alanine ($H_2NCH(CH_3)COOH$) and pyruvate ($CH_3C(O)COOH$) with pyruvate preferentially enriched in ^{13}C because the force constant at the >C=O bond minimum is larger than that of $\equiv C\text{-}NH_2$.[4]

Experimental studies of the ALT reaction found that the equilibrium constant of the reaction is 0.66 (taken from the experimental value of 1.52 for the reverse reaction) at 25 °C and biological pH (7.4), with experimental variation of less than 9%.[5] Additional experiments with racemic mixtures of alanine and glutamate rather than pure L-alanine and L-glutamate resulted in

no change in the overall equilibrium of this reaction. Equilibrium constant values of 0.62 and 0.59 (again taken from the reverse reaction) have been reported for the reaction at biological temperature, 37 °C.[6,7,8] The free energy of the ALT reaction can be obtained from the K_{eq} values and is close to zero, 1.04 ± 0.13 kJ/mol (0.25 kcal/mol).[8]

In this work, we predict the equilibrium isotope fractionation factor (IFF) for the reversible ALT reaction at a variety of computational levels. Additionally, $^{13}C/^{12}C$ substitution of alanine and pyruvic acid at the methyl and carboxylic acid sites, unrelated to the ALT reaction, are also modeled.

Computational Methods

The Car-Parrinello simulations[9] in this study were performed using an *ab initio* molecular dynamics approach (AIMD) and a combined ab initio and classical molecular dynamics (AIMD/MM) approach that is part of the pseudopotential plane-wave program (NWPW module) contained in the NWChem quantum chemistry package.[10] This method couples the pseudopotential plane-wave solvers[11,12] of the Density Functional Theory (DFT) equations[13,14] for the quantum mechanics region to a molecular mechanics description of a larger region. In the AIMD/MM method the total energy for the system is given by a sum of E_{AIMD}, E_{MM}, and $E_{AIMD/MM}$, where E_{AIMD} and E_{MM} are energies for quantum mechanics and molecular mechanics regions, respectively. Details about the method can be found in Cauet et al.[15] The PBE96[16] gradient corrected exchange-correlation functional was used. The valence electron interactions with the atomic H, C, N and O cores were approximated using generalized norm-conserving Hamann pseudopotentials[17,18] modified into a separable form as suggested by Kleinman and Bylander.[19] The original pseudopotential parameterizations suggested by Hamann[18] were too 'hard,' and softer pseudopotentials were constructed by increasing the core radii giving: H: r_{cs} =

0.8 au, r_{cp} = 0.8 au; C: r_{cs} = 0.8 au, r_{cp} = 0.85 au, r_{cd} = 0.85 au; N: r_{cs} = 0.7 au, r_{cp} = 0.7 au, r_{cd} = 0.7 au; O: r_{cs} = 0.7 au, r_{cp} = 0.7 au, r_{cd} = 0.7 au. The electronic wavefunctions were expanded using a plane-wave basis set with periodic boundary conditions, sampled at the Γ point, with a wavefunction cutoff energy of 100 Ry and a density cutoff energy of 200 Ry. A homogeneous negative charge background neutralized the −1 charge used in these simulations. The Car-Parrinello equations of motion[9] were integrated in the presence of Nose-Hoover thermostats,[20,21,22] coupled to the electronic and ionic degrees of freedom, at T = 300 K with a time step of 0.17 fs and a fictitious mass of 750 au. The simulation cells were simple cubic. They were chosen to have a concentration of 1 M (i.e. solute + 56 H_2O) and 0.5 M (i.e. solute + 112 H_2O) with a density of water near 1. A Shrake-Rupley algorithm[23] was used to determine the volumes of the amino acid solute. The SPC/E MM potential for water was used in the QM/MM portion of this study.[24] The parameters for this potential and the QM/MM potential used in this work are given in the Appendix. EMSL Arrows[25] was used generate the input decks and perform these calculations.

In order to model the overall reactions of interest, shown in Figure 5.1, the structures of molecules with single solvation shells of explicit H_2O molecules for each of the four anions were isolated from the AIMD and AIMD/MM simulations; these starting structures are shown in Figure A5.1. This protocol is rationalized in terms of Inherent Structure Theory[26] in which a thermal ensemble of atomic configurations is mapped (by steepest descent) onto potential energy minima, to produce an ensemble averaged IFF. This type of sampling strategy allows for averaging over multiple solvent configurations, while at the same time only having to compute explicit IFFs on a small number of samples. Protons were added to the anions to form neutral species for all four molecules, which for alanine and glutamic acid includes non-charge separated

(NCS) and zwitterionic structures. There are two versions of the glutamic acid zwitterion, one with the negative charge on the C-terminus and one with the negative charge on the side chain carboxylate. A variation of the glutamate anion structure with a zwitterionic backbone was also modeled in which the N-terminus has an additional proton and both carboxylates are deprotonated. Protons were removed to generate the glutamate and α-ketoglutarate dianions. For glutamate, a second anionic structure was generated by moving the proton from the side chain carboxylic acid group to the C-terminus carboxylic acid group. For α-ketoglutarate, a second anion was generated by moving the proton from the carboxylic acid of glutarate moiety to the carboxylate of the α-keto moiety (Figure 5.2). These initial structures were optimized at the density functional theory (DFT) level using the hybrid exchange-correlation method B3LYP[27,28] with the DZVP2[29] basis set. Any explicit H_2O molecules that left the first solvation shell were removed before the next iteration of optimization for all neutral and ionic species. The number of H_2O molecules in the variations of a given chemical species with the largest first solvation shell was maintained for all other variations of that molecule. This results in species with additional H_2O molecules not contained in the first solvation shell for our simulations, but those additional H_2O molecules do not affect the energetics of the overall reaction.

The B3LYP/DZVP2-optimized geometries were reoptimized at the correlated molecular orbital (MO) theory MP2 level[30,31] with the aug-cc-pVDZ (aD) basis set[32,33] in the gas phase and in the presence of implicit aqueous solvation using a self-consistent reaction field (SCRF) approach[34] with COSMO[35,36] parameters. The MP2 level was chosen because we need to be able to calculate vibrational frequencies and due to issues with non-bonding interactions with most density functional theory exchange-correlation functionals. Such geometry optimizations and frequency calculations are not feasible computationally with higher levels of correlated

molecular orbital theory. In addition, the MP2 method provides excellent energetics for hydrogen bonding for water molecules[37] and has been shown to provide good isotope fractionation factors in aqueous and mineral carbonates.[38] For the gas phase optimized structures, aqueous corrections were also calculated at the MP2/aD/COSMO level and aqueous free energies (ΔG_{aq}) were calculated according to Equation (5.5).

$$\Delta G_{aq} = \Delta G_{gas} + \Delta G_{solv} \tag{5.5}$$

A dielectric constant of 78.39 corresponding to that of bulk water was used in the COSMO calculations. The solvation energy includes only the electrostatic energy component. Gas phase optimized geometries were also reoptimized in aqueous solution at the MP2/aD/COSMO level. All calculations were performed using the Gaussian 16 software package.[39] Harmonic vibrational frequencies were calculated for MP2/aD and MP2/aD/COSMO optimized clusters for use in the isotopic fractionation calculations.

The calculated vibrational frequencies for "light" (all isotopically light atoms: ^{1}H, ^{12}C, ^{14}N, ^{16}O) and "heavy" (one ^{13}C with all otherwise isotopically light atoms) isotopes were used to calculate isotopic fractionation ratios for $^{13}C/^{12}C$ at the C2 site and the carboxylic acid and methyl groups for alanine and pyruvic acid. The vibrational partition functions (relative to the minimum of the potential energy surface) for a given cluster, q_{vib}, were calculated using Equation (5.6)

$$q_{vib,i} = \frac{e^{-\theta_i/2T}}{1-e^{-\theta_i/T}} \tag{5.6}$$

where $\Theta_i = h\upsilon_i/k_B$ and υ_i are the individual fundamental modes for a given species as calculated at the MP2/aD or MP2/aD/COSMO level. The individual vibrational partition functions are then used to calculate Q for each cluster, (Equation (5.7)).

$$Q = \prod_i \frac{\theta_i}{T} q_{vib,i} = \prod_i \frac{\theta_i}{T} \frac{e^{-\theta_i/2T}}{1-e^{-\theta_i/T}} \tag{5.7}$$

We use the vibrational partition functions of each compound to evaluate the equilibrium constants for reactions shown in Figure 5.1, both for the versions of those reactions involving reactants and products containing ^{12}C in all carbon positions (indicated with the subscript 'l' to signify isotopically light) and for versions that contain a ^{13}C in one of the carbon positions of reactant alanine and product pyruvate (indicated with the subscript 'h' to signify isotopically heavy). This results in two related equations for the equilibrium constant for each reaction — K_l for the reaction involving only isotopically light reactants and K_h for the reaction involving one isotopically heavy reactant and product. Equations (5.8) and (5.9) illustrate this nomenclature.

$$K_l = \frac{Q_{Pyr,l}Q_{Glu,l}}{Q_{Ala,l}Q_{\alpha-Ket,l}} \tag{5.8}$$

$$K_h = \frac{Q_{Pyr,h}Q_{Glu,l}}{Q_{Ala,h}Q_{\alpha-Ket,l}} \tag{5.9}$$

The isotope fractionation factor (IFF; ϵ) associated with the reaction is determined by taking the ratio of the equilibrium constant for the version of the reaction having heavy isotopes to the equilibrium constant for the version of the reaction with all light isotopes. This value is calculated as shown in Equation (5.10).

$$\epsilon = \left[\left(\frac{K_h}{K_l}\right) - 1\right] * 1000 \tag{5.10}$$

We also use the vibrational frequencies of isotopically heavy and light forms of the studied compounds to calculate the reduced partition function ratios (RPFR; β), which provide a measure of the propensity of a compound of interest to concentrate a heavy isotope relative to a pool of free atoms of the relevant element. Comparison of β values for two compounds provides a measure of the tendency of heavy isotopes to partition into one vs. the other compound at equilibrium. Similarly, when β values are computed for two different isotopomers of the same compound, comparison of these values indicates whether the heavy isotope of interest

preferentially concentrates in one vs. the other atomic site. Values of β were calculated in the harmonic approximation for each individual species as shown in Equation (5.11)

$$\beta = \frac{Q_h}{Q_l} = \frac{\prod_i \theta_i q_{vib,h,i}}{\prod_j \theta_j q_{vib,l,j}} \tag{5.11}$$

where Q_h is the Q value of the isotope that contains a ^{13}C at a given site and Q_l is the Q value of the isotope with only ^{12}C atoms.

Equilibrium constants are also calculated from the free energies in the gas phase and in aqueous solution and are given as K′. These are determined using the standard thermodynamic relationship in Equation (5.12).

$$K' = e^{-\Delta G/RT} \tag{5.12}$$

Or, equivalently, lnK' = -ΔG/RT. This approach can also be used to estimate the IFF associated with a reaction by approximating ΔG values with zero-point energies of reaction (reaction at 0 K). Specifically, we can estimate the IFF as $\epsilon_{ZPE} = 1000*\ln K_{ZPE}$, where $\ln K_{ZPE}$ is computed as shown in Equation (5.13)

$$lnK_{ZPE} = \frac{-\Delta\Delta ZPE}{RT} \tag{5.13}$$

where ΔΔZPE is the difference between ΔZPE values for the heavy isotope-containing molecules in an isotope fractionation equilibrium, which are calculated as follows in Equation (5.14), in which ω_i^{12} and ω_i^{13} are fundamental frequencies with all isotopically light atoms or with one substituted ^{13}C atom, respectively.

$$\Delta ZPE = \sum_i \frac{1}{2} hc(\omega_i^{12} - \omega_i^{13}) \tag{5.14}$$

A positive ϵ or ϵ_{ZPE} value indicates that the reaction favors formation of the heavy isotope product over heavy isotope reactant in an equilibrium between two species.

Benchmarking was performed for all of the above calculation types by modeling equilibrium isotope fractionation between formaldehyde and methylamine as well as various deprotonated forms of pyruvic acid and alanine with no explicit solvation. Three DFT methods – B3LYP, LSDA,[13,14,40,41] and BP86[42,43] – using the DZVP2 basis set were compared with values derived from MP2 calculations with the aD and aT (aug-cc-pVTZ)[32,33] basis sets. These values are also compared to similar calculations performed using the NWChem software;[10] the B3LYP, LSDA, and BP86 DFT methods were used with the 6-311++G(2d,2p) basis set[44] in addition to MP2/aT calculations.

Results

Optimized Geometries and Deprotonation Energies with Explicit H_2O. Starting from the initial structures taken from the MD simulations (Figure A5.1), gas phase geometry optimization resulted in final structures for alanine with 7 explicit H_2O molecules, glutamic acid with 8 explicit H_2O molecules, pyruvic acid with 6 explicit H_2O molecules, and α-ketoglutaric acid with 9 explicit H_2O molecules (Figures A5.2 through A5.6). The inclusion of larger numbers of H_2O molecules resulted in the extra H_2O molecules moving into a second solvent shell. The number of explicit H_2O molecules were kept consistent despite the structure of or charge on the molecule of interest to compare the energies of the resulting optimized geometries directly. These gas phase optimized geometries were reoptimized in aqueous solution at the MP2/aD/COSMO level are also given in Figures A5.2 through A5.6. The reactions of interest involve the zwitterionic structures of alanine and glutamic acid as well as deprotonated forms of the amino acids, pyruvic acid, and α-ketoglutaric acid.

Gas phase and aqueous free energies for the formation of the various species for alanine, glutamic acid, pyruvic acid, and α-ketoglutaric acid and their deprotonated forms are

summarized in Table 5.1. When reoptimized in aqueous solution, the molecules of interest do not differ in geometry, but the explicit H_2O molecules in each cluster can reorganize in the single solvation shell and may leave the first solvent shell. This can result in differences between gas phase optimized ΔG_{aq} and aqueous optimized ΔG_{aq}. The calculated pK_a values for deprotonation of aqueous optimized clusters are systematically higher by 1 to 8 pK_a units with respect to deprotonation of gas optimized clusters. All the discussion is based on the organic molecules in their cluster of H_2O solvent molecules. These calculated pK_a values are compared to experimental values[45] in Table 5.1 and are discussed below.

For alanine (Figure A5.2), the formation of the zwitterion is exergonic and favorable in both the gas phase and aqueous optimized clusters by -15.3 and -10.5 kcal/mol respectively at the MP2/aD/COSMO level. This is expected for an amino acid in aqueous solution. The aqueous free energy of the deprotonation of alanine is -6.7 kcal/mol for the gas phase optimized reaction and 4.1 kcal/mol for the aqueous optimized structure. However, in aqueous solution alanine is present as a zwitterion rather than a structure without charge separation. Aqueous deprotonation of zwitterionic alanine is endergonic for both gas and aqueous optimized clusters by 8.6 and 14.6 kcal/mol respectively at the MP2/aD/COSMO level. The gas phase and aqueous optimized pK_a values for neutral zwitterionic alanine losing a proton to form anionic alanine are 6.3 and 10.7, as compared to an experimental value of 9.87.[45] In this case, the optimized aqueous phase structure yields a pK_a closer to experiment. The use of the gas phase structures, except in the case of formation of the glutamate dianion, give pK_a values that are too acidic.

For glutamic acid (Figure A5.3), the gas phase and aqueous optimized modeled reactions are endergonic. Deprotonation of glutamic acid was modeled at the C-terminus, the side chain carboxylic acid, or with double deprotonation at both carboxylic acid groups. Deprotonation at

the C-terminus of glutamic acid yields an aqueous reaction free energy of 6.7 and 14.2 kcal/mol for the gas and aqueous optimized reactions respectively at the MP2/aD/COSMO level. Deprotonation at the side chain of glutamic acid is more favorable than at the C-terminus with aqueous free energies of 0.8 kcal/mol for the gas phase optimized reaction and 12.1 kcal/mol for the aqueous phase optimized reaction. Thus, deprotonation at the C-terminus is 5.9 kcal/mol and 2.1 kcal/mol higher in aqueous free energy than deprotonation at the side chain of glutamic acid for the gas phase and aqueous phase optimized reactions, respectively. Aqueous deprotonation of both carboxylic acid sites requires 22.5 kcal/mol for the gas phase and aqueous optimized reactions. The deprotonation of the side chain following the first deprotonation at the C-terminus requires an aqueous free energy change of 9.3 and 10.4 kcal/mol for the gas phase and aqueous optimized reactions, respectively.

In aqueous solution, the zwitterionic glutamic acid structure is preferred. For the gas phase optimized clusters, the NCS glutamic acid with 8 explicit H_2O molecules is favored over the zwitterionic glutamic acid structures. Compared to the $\Delta G_{gas,298K}$ of NCS glutamic acid, zwitterionic glutamic acid with the C-terminus deprotonated is 7.3 kcal/mol higher and zwitterionic glutamic acid with the side chain deprotonated is 3.4 kcal/mol higher in free energy at 298 K. For the aqueous optimized clusters, the zwitterionic glutamic acid with deprotonated C-terminus is lowest in free energy at 298 K, with relative $\Delta G_{aq,298K}$ values for the zwitterionic glutamic acid with deprotonation at the side chain of 2.6 kcal/mol and for NCS glutamic acid of 6.4 kcal/mol. Aqueous glutamate anions also exists with a zwitterionic backbone. In the gas phase, the $\Delta G_{gas,298K}$ of the glutamate zwitterion is 4.0 kcal/mol higher than the glutamate structure with the side chain deprotonated and a NCS backbone, but the glutamate zwitterion is 5.7 kcal/mol more favorable than the side chain anion with the NCS backbone in $\Delta G_{aq,298K}$ for

the gas optimized cluster. Similarly, for the aqueous optimized cluster, $\Delta G_{aq,298K}$ is 6.0 kcal/mol higher for the NCS backbone glutamate side chain anion than for the zwitterionic glutamate anion. Zwitterionic glutamic acid deprotonation reactions compared to experimental pK_a's for aqueous glutamic acid are shown in Figure A5.4. The gas phase and aqueous optimized pK_a values for neutral zwitterionic glutamic acid losing a proton to form the zwitterionic glutamate anion are 1.7 and 9.2, which can be compared to the experimental pK_a of 4.31.[45] Again, the gas phase optimized structure pK_a is in better agreement with experiment. The gas phase and aqueous optimized pK_a values for zwitterionic glutamate anion losing a proton to form dianionic glutamate are 11.0 and 12.0, which can be compared to the experimental value of 9.67.[45] Both calculated values are not acidic enough. Aqueous deprotonation of pyruvic acid to form pyruvate is very slightly endergonic for the gas phase optimized reaction (0.5 kcal/mol) and endergonic for the aqueous optimized reaction (5.4 kcal/mol), as shown in Figure A5.5. The gas phase and aqueous optimized pK_a values for pyruvic acid losing a proton to form pyruvate are 0.4 and 4.0, which can be compared to an experimental pK_a of 2.39.[45] Both calculated values are in good agreement with experiment.

For α-ketoglutaric acid (Figure A5.6), reactions between gas phase optimized structures are exergonic and those between aqueous optimized structures are endergonic. Deprotonation of α-ketoglutaric acid was modeled at either the -C(O)C(O)OH (α-keto moiety), which is analogous to the C-terminus of an amino acid, the -CH_2C(O)OH (glutarate moiety), which we consider analogous to the side chain of glutamic acid, or with both carboxylic acid groups deprotonated. The gas phase optimized reaction free energies for formation of these respective deprotonation products are -4.0, -2.4, and -1.6 kcal/mol, respectively. Those for aqueous phase optimized reactions are 2.2, 6.1, and 10.6 kcal/mol, respectively. Regardless of optimization method,

deprotonation is preferred at the α-keto moiety (-C(O)C(O)O^-) over deprotonation at the glutarate moiety (-CH_2C(O)O^-) by 1.6 kcal/mol for the gas optimized reaction and by 3.9 kcal/mol for the aqueous optimized reaction at the MP2/aD/COSMO level. Deprotonation at both carboxylic acid groups on α-ketoglutaric acid is -1.6 kcal/mol for the gas phase optimized reaction and 10.6 kcal/mol for the aqueous phase optimized reaction, which requires an additional 3.4 kcal/mol for the gas phase optimized reaction and 8.4 kcal/mol for the aqueous phase optimized reaction beyond the deprotonation at the α-keto moiety site. The gas phase and aqueous optimized pK_a values for α-ketoglutaric acid losing a proton to form the lowest energy anion (deprotonated at the α-keto moiety) are -3.0 and 1.6, which can be compared to the experimental pK_a of 2.47.[45] Further deprotonation of this anion to form the α-ketoglutarate dianion yields gas phase and aqueous optimized pK_a's of 1.8 and 6.2, which can be compared to the experimental value of 4.68.[45] In this case, the solution optimized structures gives better agreement with experiment for both values.

Transamination Reaction Energetics. The calculated energetics of reactions (5.1) through (5.4), as shown in Figure 5.1, are given in Table 5.2 in terms of enthalpy, free energy, and equilibrium constant in the gas phase and the free energy and equilibrium constant in aqueous solution at 298 K for each species with explicit solvation. These values are given for the reactions with all light isotopes. In the gas phase, reaction (5.1), which involves α-ketoglutarate deprotonated at the α-keto moiety and glutamate deprotonated at the C-terminus, is both exothermic (ΔH_{gas}(298 K) = -8.8 kcal/mol) and exergonic (ΔG_{gas}(298 K) = -2.9 kcal/mol) at the MP2/aD level. In aqueous solution, both the gas phase and aqueous optimized reactions are endergonic at 5.8 and 8.0 kcal/mol respectively at the MP2/aD/COSMO level. Reaction (5.2), which involves α-ketoglutarate deprotonated at the glutarate moiety and glutamate deprotonated

at the side chain, is exothermic (ΔH_{gas}(298 K) = -2.4 kcal/mol) and slightly endergonic (ΔG_{gas}(298 K) = 0.9 kcal/mol) in the gas phase. Thus, this reaction is less favorable than reaction (5.1) by 3.8 kcal/mol (ΔG_{gas}(298 K)) at the MP2/aD level in the gas phase. In aqueous solution, the free energy of reaction (5.2) is more favorable than reaction (5.1) for the gas phase and aqueous optimized solutions. In aqueous solution, the gas phase optimized free energy is exergonic at -1.7 kcal/mol and is endergonic at 2.0 kcal/mol for the aqueous optimized reaction. These values are 7.5 and 6.0 kcal/mol more favorable than those of reaction (5.1) for the gas phase and aqueous optimized reactions, respectively. Reaction (5.3) contains each species as it would appear at biological pH (7.4) – zwitterionic alanine and doubly deprotonated α-ketoglutarate in equilibrium with deprotonated pyruvate and the zwitterionic glutamate anion structure. This reaction is much more highly exothermic in the gas phase (ΔH_{gas}(298 K) = -46.5 kcal/mol; ΔG_{gas}(298 K) = -47.9 kcal/mol) than any of the other equilibrium reactions between other variations of these species. However, in aqueous solution, both the gas phase and aqueous optimized reactions are comparable in energy to reaction (5.2) with gas phase optimized ΔG_{aq}(298 K) = -3.9 kcal/mol and aqueous optimized ΔG_{aq}(298 K) = 3.8 kcal/mol. These calculated values bracket the experimentally determined free energy for the ALT reaction, 0.25 kcal/mol.[8] Considering the difficulty of calculating the aqueous free energy of the dianion accurately with SCRF approaches, both of our ΔG_{aq}(298 K) values are consistent with experiment. Reaction (5.4), involving doubly deprotonated α-ketoglutarate and glutamate, is endothermic and endergonic in the gas phase and endergonic in aqueous solution. In the gas phase, reaction (5.4) has an enthalpy of 7.6 kcal/mol and a free energy of 6.8 kcal/mol at 298 K. In aqueous solution, both of the gas phase aqueous optimized reactions are nearly as endergonic as the aqueous optimized reaction (5.1) at 6.8 and 7.9 kcal/mol, respectively.

Position-Dependent RPFRs for Explicitly Solvated Species. Before describing the predicted IFFs, we describe the results for the Q values. The Q values calculated directly from the frequencies determined at the MP2/aD level for gas phase optimized clusters and the MP2/aD/COSMO level for aqueous phase optimized clusters for both light isotopes (^{12}C only; Q_l) and each heavy isotope (^{13}C at a single carbon site, Q_h) are given in Table A5.3 in the Appendix for both the gas phase and aqueous cluster geometries of alanine, glutamic acid, pyruvic acid, and α-ketoglutaric acid as different isomeric and charged (addition or loss of a proton) species with explicit H_2O molecules. These Q values were used to calculate the reduced partition function ratio (RPFR, β) values that are given in Table 5.3 for the vibrational component only with respect to the site of ^{13}C substitution. The β values indicate whether a given site in the molecule of interest will become enriched in ^{13}C. For all species modeled in both the gas phase and aqueous solution, $\beta > 1$, which indicates that enrichment in ^{13}C at all modeled sites is favorable to varying degrees. Each of the four species of interest and their various forms are modeled with all light isotopes as well as with one carbon replaced with a heavy ^{13}C. Values for β are reported for both gas phase and aqueous optimized clusters of the molecules of interest and their respective explicit H_2O molecules. The gas phase and aqueous optimized β values are very similar for all reported species, so only the aqueous optimized β_{aq} values will be discussed as these most accurately represent the environment of the aqueous species of interest during the ALT reaction. Additionally, we compare to the $^{13}C/^{12}C$ exchange at other C sites in the alanine and pyruvic acid species, despite the fact that the ALT reaction does not provide pathways for equilibrating some of these families of sites. For example, if you were to ^{13}C label the methyl group of pyruvate, that label would be transferred to the methyl group of alanine through the ALT chemistry, but it would never migrate into the terminus or C2 sites of either compound.

Figure 5.3 shows the cumulative β values with vibrational frequency (i.e., cumulative product of the contributions to β from each fundamental mode) for all ^{13}C-substituted forms of alanine (with 7 explicit H_2O molecules) and pyruvic acid (with 6 explicit H_2O molecules). The inclusion of 2 additional explicit H_2O molecules for pyruvate, to make a total of 8 explicit H_2O, did not affect the β values. Sharp increases in β indicate vibrational frequencies that contribute significantly to the reduced partition function ratio when either the methyl, the C2, or the C-terminus site contains a heavy carbon. The cumulative β vs. frequency plots for gas phase and aqueous optimized structures are very similar for alanine-based and pyruvic acid-based species, showing that different optimization procedures do not greatly affect the cumulative β. For alanine and pyruvic acid with $^{13}C/^{12}C$ substitution at the C2 and methyl sites, neutral and deprotonated species have extremely similar cumulative β plots, regardless of charge or conformation. For all alanine and pyruvic acid species with $^{13}C/^{12}C$ substitution at the methyl group, asymmetric C-H stretches at the methyl sites contribute most significantly to the cumulative β at ~ 3100 to 3200 cm^{-1}. $^{13}C/^{12}C$ substitution for alanine results in a significant contribution at ~ 3100 cm^{-1} from the C-H stretch at the C2 position. Substitution at the C2 site of pyruvic acid, on the other hand, has a significant contribution from the C=O stretch at the C2 site at ~1650 cm^{-1}. However, $^{13}C/^{12}C$ substitution at the carboxylic acid site, the neutral or deprotonated state of the alanine or pyruvic acid molecule affects the frequency that contributes most to the cumulative β. Because the protonated state of the molecule affects the atoms directly bonded to the carboxylic acid/carboxylate, neutral and anionic alanine-based and pyruvic acid-based species have significant contributions to the cumulative β that arise from unique fundamental modes. For NCS neutral alanine, $^{13}C/^{12}C$ substitution at the C-terminus yields a significant contribution to β at ~ 1690 cm^{-1} (C=O asymmetric stretch coupled with the H-O-C

bend). For deprotonated alanine, the loss of the H-O bond results in a red-shifted C=O asymmetric stretch coupled with the H-C-C bend that occurs at ~1500 cm^{-1}. For zwitterionic alanine, the NH_3 symmetric bend mode dominates with a small amount of coupling with the C=O asymmetric stretch at the C-terminus (~ 1520 cm^{-1}). For pyruvic acid, there are two significantly contributing fundamental modes, the C=O stretch at the C2 site, and the C=O asymmetric stretch at the carboxylic acid coupled with the H-O-C bend (1690 to 1720 cm^{-1}). $^{13}C/^{12}C$ substitution at the carboxylate for pyruvate, regardless of the number of explicit H_2O molecules, has a significant contribution at 1535 cm^{-1} from the C=O asymmetric stretch mode coupled with one or more H-O-H bends from the explicit H_2O molecules. Substitution of ^{13}C into carbon sites of glutamic acid and α-ketoglutaric acid species are not relevant to reactions (1) through (4), so discussion of these β values is in the Appendix.

Substitution of ^{13}C at the C-terminus site is the most favorable of the substitution sites modeled, as it yields the highest β_{aq} value for all alanine species with 7 explicit H_2O molecules and ^{13}C substitution at the methyl group is the least favorable (smallest β_{aq}). Analogous values for pyruvic acid species are comparable to those of alanine, as ^{13}C substitution is most favorable at the C-terminus and least favorable at the methyl group. The C2 position for pyruvic acid and pyruvate has a somewhat larger average β_{aq} value (1.1835) than the C2 position of alanine and its various species (average β_{aq} = 1.1742) by 9.3 ‰.

Isotope Fractionation During Aqueous Alanine Transamination. The ALT reaction is catalyzed by the alanine transaminase (ALT) enzyme and involves exchanging the =O at the C2 position of pyruvate with NH_3^+ to form zwitterionic alanine at biological pH in aqueous solution.[46] This reaction was modeled in terms of various components. The free energy of the reaction depends on the enthalpy and entropy which incorporates all of the terms including those

due to isotope substitution. This is more complicated than the often-studied isotopic substitution reaction which has a nominal $K_{eq} = 1$ for all of the same isotopes. Even if one were to just consider the vibrational contributions, the equilibrium constants K_l or K_h with all light or all heavy isotopes are not equal to one. To maintain an equal number of explicit H_2O molecules between reactants and products, pyruvic acid/pyruvate was modeled with 8 explicit H_2O molecules for the isotopic studies.

The aqueous organic reaction between alanine and α-ketoglutarate to form pyruvate and glutamate was modeled in a variety of ways. Equilibrium constants (K) given in Table 5.4 are calculated from the appropriate Q_l or Q_h values according to the reactions as they are defined in Figure 5.1. These reactions are also shown with optimized cluster geometries in Figure 5.4, for which the isotope fractionation factors (ϵ) can be directly compared with the simpler zero-point energy-based approximations (ϵ_{ZPE}), which are described below. A value of K greater than 1 indicates that the reaction equilibrium favors the production of pyruvate and glutamate. In the gas phase, reactions (5.1), (5.3), and (5.4) are predicted to favor the formation of products, pyruvate and glutamate. Aqueous reaction (5.3), which involves the chemical species as they would appear at biological pH, also slightly favors the formation of products. However, gas phase reaction (5.2) and aqueous reactions (5.1), (5.2), and (5.4) favor the reactants, alanine and α-ketoglutarate. Regardless of the individual equilibrium constant values, the ϵ values (i.e., the change in equilibrium constant on ^{13}C substitution) are in excellent agreement for all four modeled reactions during $^{13}C/^{12}C$ exchange at the C2 site and at the C-terminus. The calculated ϵ values for the C-terminus ^{13}C substitution indicate that the reaction is impeded by this substitution. A heavy carbon at the C-terminus favors the reactants, alanine and α-ketoglutarate, by ~ 1 ‰ in aqueous solution and ~ 3 ‰ in the gas phase. The ϵ for the C2 position ^{13}C

substitution predicts that this heavy carbon substitution favors the product, pyruvate, by ~ 9 ‰ in both the gas and aqueous phases. $^{13}C/^{12}C$ substitution at the methyl group is very consistent for reactions (5.1), (5.2), and (5.4) in both gas and aqueous optimized reactions and favors the reactants by 2.0 to 2.5 ‰. However, methyl substitution differs when the reaction consists of species present at biological pH (zwitterionic alanine rather than deprotonated alanine) and favors the reactants more strongly by a factor of 2 (4.9 to 5.0 ‰).

Zero Point Energy-Based Approach. A simpler approach is just to use zero-point energies to estimate ϵ. At T = 0 K, the ϵ values are determined only by changes in the ZPE using Equation 5.12. ϵ_{ZPE} values shown in Figure 5.4 for aqueous optimized clusters and are reported in Table A5.3 for both gas and aqueous optimized clusters for the equilibrium $^{13}C/^{12}C$ exchange reactions between zwitterionic or deprotonated alanine and pyruvate. In both the gas phase and in aqueous solution, ^{13}C substitution at the C2 position of both deprotonated alanine and pyruvate result in favored formation of the heavy product, pyruvate, by 9.2 and 9.5 ‰ in the gas and aqueous optimized reactions, respectively, which is in excellent agreement with the ϵ results of 8.9 to 9.0 ‰. However, $^{13}C/^{12}C$ exchange at the C-terminus favors the heavy reactant, deprotonated alanine, by 4.3 ‰ in the gas phase and 2.1 ‰ in aqueous solution, which is in good agreement with the predicted ϵ of ~ 1 ‰ in aqueous solution and ~ 3 ‰ in the gas phase. The ϵ_{ZPE} values for biological conditions (zwitterionic alanine) are in excellent agreement with the values calculated for $^{13}C/^{12}C$ substitution of deprotonated alanine and pyruvate at the C2 and C-terminus positions. The only exception is that the gas phase optimized ^{13}C substitution of zwitterionic alanine and pyruvate is predicted to favor the formation of heavy pyruvate slightly less ($\epsilon_{ZPE,gas}$ = 7.8 ‰) than the exchange when both species are deprotonated ($\epsilon_{ZPE,aq}$ = 9.2 ‰). $^{13}C/^{12}C$ exchange between the methyl groups of alanine and pyruvate depend much more on the form of species in

equilibrium. $^{13}C/^{12}C$ exchange at the methyl group favors zwitterionic alanine over pyruvate by 6.5 ‰ in the gas optimized equilibrium and 5.8 ‰ in the aqueous optimized equilibrium, which is in good agreement with ϵ = 4.9 to 5.0 ‰. Deprotonated alanine is favored over pyruvate, though to a much lesser extent, with a predicted $^{13}C/^{12}C$ substitution of 1.1 to 1.2 ‰ for both the gas and aqueous optimized reactions, which is also in good agreement with ϵ = 2.0 to 2.5 ‰.

An even simpler estimate to obtain ϵ is to exclude the explicit waters of solvation. ^{13}C exchange between pyruvate and alanine were studied as shown in reactions (5.15) and (5.16) (Figure 5.5). Vibrational frequencies at a variety of computational levels were used to calculate the IFFs of these reactions from both the zero-point energy-based ϵ_{ZPE} and the more exact ϵ approaches described above, shown in Table 5.5. For the isotope fractionation equilibria between deprotonated alanine and pyruvate, NCS neutral alanine and pyruvate, and protonated alanine and pyruvic acid in reactions (5.15) and (5.16) with ^{13}C at the C2 position, both ϵ_{ZPE} and ϵ are positive, so ^{13}C will be enriched in the product (pyruvate or pyruvic acid) preferentially over the reactant (NCS alanine or deprotonated alanine). This preference is largest for reaction (5.15) and is smallest for reaction (5.16).

Table 5.6 gives ^{13}C site specific β values for gas phase optimized NCS alanine, deprotonated alanine, pyruvic acid, and pyruvate with no explicit H_2O molecules at the MP2/aT level. Analogous values for these compounds at the MP2/aD level are in the Appendix. The carboxylic acid site is predicted to be the most favorable for ^{13}C substitution for alanine and pyruvic acid and their various species as these sites yield the largest β values. MP2/aT β values are very similar to those calculated at the MP2/aD level with each MP2/aT value being slightly larger in magnitude than the analogous MP2/aD value. These site specific β values for pyruvic acid and alanine species are in good agreement with the explicitly solvated results for the same

species as listed in Table A5.3, therefore the inclusion of solvent is not necessary for calculating accurate β values for this reaction. Our values for $^{13}C/^{12}C$ substitution in Table 5.6, when compared with the β value for CO_2 (1.1939) calculated at the MP2/aT level are -43.9 ‰ at the methyl, -14.1 ‰ at the C2 site, and 6.2 ‰ at the carboxylic acid. These values are in good agreement within 2 ‰ with previous literature values for the gas phase zwitterionic structure of alanine calculated at the same MP2/aT level.[47] Note that Rustad[47] has calculated isotopic enrichments in all of the gas phase zwitterionic amino acids relative to CO_2 at the MP2/aD level.

An even simpler model is to look at the isotope exchange between formaldehyde and methylamine as shown in reaction (5.17) (Figure 5.5). Formaldehyde and methylamine serve as a simplified version of the type of exchange that takes place at the C2 sites of pyruvate and alanine. The calculations for reaction (5.17) were done at the MP2/aT level given above. For the isotope fractionation equilibrium between formaldehyde and methylamine as written in reaction (5.17), both measures of preference for ^{13}C enrichment, ϵ_{ZPE} and ϵ, are positive, so the product containing heavy carbon ($H_2{}^{13}C(O)$) is preferred over the formation of the reactant containing heavy carbon ($H_3{}^{13}CNH_2$).

For each of the compounds in Table 5.6, the relationship between cumulative β and frequency is given in Figure 5.6. This plot demonstrates the frequencies that significantly contribute to the β for each ^{13}C substituted species and ultimately result in pyruvate and formaldehyde being favored for ^{13}C enrichment over alanine and methylamine, respectively. Pyruvate and alanine species shown have ^{13}C at their C2 positions so that formaldehyde and methylamine can be directly compared. Pyruvic acid, and pyruvate, and formaldehyde have C=O stretch frequencies of ~ 1700 cm^{-1} that contribute significantly to the overall β for these molecules. Additionally, pyruvate and, to a slightly lesser extent, pyruvic acid have a frequency

of ~ 1150 cm^{-1} (a coupled C-C with C-H or coupled C-C with C-H and O-H) that also increases their β values, which formaldehyde lacks. The cumulative β values of the other analogous molecules that lack the C=O stretch contribution never can catch up, despite having noteworthy increases in β due to significant C-H stretch contributions from frequencies of ~ 3000 to 3200 cm^{-1}. Alanine and methylamine have a set of frequencies from approximately 1000 cm^{-1} to 1600 cm^{-1} that coincide with C-N and the N-H stretches, which contribute significantly more to the β values of alanine compounds than those of methylamine.

Estimate of ϵ Using Canonical Functional Group Frequencies. As described above, the simple estimate of ϵ for the enrichment of ^{13}C in $R_2C{=}O$ relative to $R_2CH\text{-}NH_2$ gave results within 2 ‰ for the actual ALT reaction. An even simpler approach to predict the isotope fractionation factor ϵ_{ZPE} is to use canonical functional group frequencies rather than frequencies calculated or measured for particular molecules. To estimate how much ^{13}C enrichment one would expect in $R_2C{=}O$ vs $R_2CH\text{-}NH_2$, the system can be approximated in terms of three key frequencies,[48] the carbonyl C=O stretch, the amine C-N stretch, and the C-H stretch in the amine. The canonical group frequencies are given in Table 5.7. To simplify the analysis, we assume that these canonical group frequencies are associated with the most abundant isotopes; $^{12}C{=}^{16}O$, $^{12}C\text{-}^{14}N$, and $^{12}C\text{-}^{1}H$. Here we use cardinal masses instead of exact masses, due to the inexact nature of using canonical group, as opposed to measured or calculated, frequencies. The isotopically substituted frequency only depends only on the ratio of the reduced masses so we can readily estimate the frequency of the $^{13}C{=}^{16}O$, $^{13}C\text{-}^{14}N$, and $^{13}C\text{-}^{1}H$ vibrations as $\omega_{13}= \sqrt{\frac{\mu_{R12}}{\mu_{R13}}}\,\omega_{12}$. The heavy C frequencies estimated in this way are also given in Table 5.7.

The free energy of a quantum harmonic oscillator is given by[49]

$$A = \Sigma_i\{½hc\omega_i + k_bT\ln(1 - e^{-hc\omega_i/k_bT})\} \quad (5.18)$$

in which the first term is the zero-point energy contribution and the second term is the thermal energy contribution to the overall free energy. For the isotope exchange reaction,

$$^{13}C{=}O + {}^{12}CHNH_2 \leftrightarrow {}^{12}C{=}O + {}^{13}CHNH_2 \quad (5.19)$$

the vibrational free energy change ΔA can be calculated from Equation (5.18) summed over the three frequencies for the reactants and products; then K is calculated as $e^{-\Delta A/RT}$. Table 5.8 gives results at T = 298 K so that these values can be compared directly to other values in this work. Analogous values at T = 310 K (37 °C) are given in the Appendix, which is the temperature relevant for biological systems in warm-blooded mammals. This difference in temperature makes virtually no difference in the ϵ_{ZPE} reported in Table 5.8. Additionally, the thermal contributions beyond the ZPE are less than 1 ‰ and therefore are not significant to the overall energy of the reaction. Note that K is slightly less than 1 for the reaction as written in Equation 5.19 as the ^{13}C prefers the stronger bonding environment in the carbonyl bond. Although this very rough model is not in quantitative agreement with our more sophisticated models (too large by a factor of approximately 2), it is qualitatively correct in predicting ^{13}C enrichment in the carbonyl bond.

Conclusions

The alanine transaminase (ALT) enzyme catalyzes the reversible transfer of an amino group from alanine to α-ketoglutarate to produce pyruvate and glutamate. As one would expect based on the relative strength of the >C=O bond vs that of the ≡C-NH_2 bond,[4] all levels of calculated models represented in this work predict that ^{13}C will be enriched at the C2 position of pyruvate (>C=O) over the C2 position of alanine (≡C-NH_2). The calculated aqueous free energies for the ALT reaction with explicit and implicit solvation, -3.9 and 3.8 kcal/mol for gas phase and aqueous optimized reactions, respectively, are in decent agreement with the experimental value of 0.25 kcal/mol.[8] The deviation from experiment is due to the limitations of

implicit solvation approaches for dianions even with inclusion of explicit solvating water molecules. Isotope fractionation factors (IFFs) for this reaction were calculated from the partition functions of explicitly and/or implicitly solvated alanine, pyruvic acid, glutamic acid, α-ketoglutaric acid, and their various species, which predict a 9.0 ‰ enrichment of ^{13}C in pyruvate over alanine. The computational results are not dependent on whether gas phase or aqueous optimized clusters with and without explicit H_2O molecules were included with the implicit COSMO SCRF approach. Values derived solely from zero-point energy (ZPE) contributions are in excellent agreement with the results including all vibrational components, showing that thermal corrections have virtually no effect on ϵ for this reaction. The inclusion of explicit H_2O molecules does not significantly contribute to the magnitude of the calculated IFFs. In addition to the $^{13}C/^{12}C$ equilibrium between pyruvic acid and alanine or their various species, ZPE-based IFFs (ϵ_{ZPE}) were also calculated using simpler analogues of the molecules of interest, formaldehyde and methylamine. An enrichment in formaldehyde of 7 to 8 ‰ at the MP2/aug-cc-pVDZ and aug-cc-pVTZ levels was predicted to be consistent with the value for the more complex models of the reactants and products. Empirical estimates based solely on typical values for relevant vibrational modes and reduced masses for C=O, C-H, and C-N produce a qualitative prediction for C=O enrichment over C-N that agrees with the more sophisticated models. In combination with the prior work of Rustad,[47] the computational methods developed here can be used to explore the $^{13}C/^{12}C$ isotopic fractionation of the formation of terrestrial amino acids and extraterrestrial amino acids such as those discovered on meteorites[50,51,52] to further our understanding of the nature of how amino acids are formed in different environments.

Table 5.1. Gas Phase and Aqueous Free Energy Values for Various Reactions at the MP2/aD and MP2/aD/COSMO Levels at 298 K in kcal/mol.

Reaction	# of H_2O Molecules [a]	ΔG_{gas}	$\Delta G_{aq,\ gas\ opt}$	$\Delta G_{aq,\ aq\ opt}$	Calc. Gas Opt pK_a	Calc. Aq Opt pK_a	Expt pK_a [b]
Alanine → Alanine Zwitterion	7	-5.6	-15.3	-10.5			
Alanine → Deprotonated Alanine + H^+	7	299.2	-6.7	4.1	-4.9	3.0	
Protonated Alanine → Alanine + H^+	7	236.9	17.6	20.3	12.9	14.9	
Alanine Zwitterion → Deprotonated Alanine + H^+	7	304.8	8.6	14.6	6.3	10.7	9.87
Protonated Alanine → Alanine Zwitterion + H^+	7	231.3	2.3	9.8	1.7	7.2	2.34
Glutamic Acid → Glutamate (C-terminus Deprotonation) + H^+	8	297.3	6.7	14.2	4.9	10.4	
Glutamic Acid → Glutamate (Side Chain Deprotonation) + H^+	8	297.9	0.8	12.1	0.6	8.8	
Glutamic Acid → Glutamate (Doubly Deprotonated) + $2H^+$	8	661.3	10.1	22.5			
Glutamic Acid Zwitterion (C-terminus) → Zwitterionic Glutamate Anion + H^+	8	294.6	2.4	12.5	1.7	9.2	4.31
Zwitterionic Glutamate Anion → Glutamate (Doubly Deprotonated) + H^+	8	359.5	15.0	16.4	11.0	12.0	9.67
Pyruvic Acid → Pyruvate + H^+	6	295.7	0.5	5.4	0.4	4.0	2.39
α-Ketoglutaric Acid → α-Ketoglutarate ($-C(O)C(O)O^-$) + H^+	9	286.7	-4.0	2.2	-3.0	1.6	2.47
α-Ketoglutaric Acid → α-Ketoglutarate ($-CH_2C(O)O^-$) + H^+	9	283.5	-2.4	6.1	-1.8	4.5	2.47
α-Ketoglutarate ($-C(O)C(O)O^-$) → Doubly Deprotonated α-Ketoglutarate + $2H^+$	9	354.4	2.4	8.4	1.8	6.2	4.68

[a] Number of H_2O molecules indicates the total number of explicit H_2O molecules included in all gas phase and aqueous geometry optimizations for all species present in the given reactions. [b] Ref. 45.

Table 5.2. Reaction Energies with Explicit H_2O Molecules in the Gas Phase and in Aqueous Solution in kcal/mol with All Light Isotopes.

Energy	Reaction (5.1)	Reaction (5.2)	Reaction (5.3)	Reaction (5.4)
$\Delta H_{gas,298K}$	-8.8	-2.4	-46.5	7.6
$\Delta G_{gas,298K}$	-2.9	0.9	-47.9	6.8
$\Delta G_{aq,298K}$ Gas Opt	5.8	-1.7	-3.9	6.8
$\Delta G_{aq\ 298K}$ Aq Opt	8.0	2.0	3.8	7.9
$K'_{aq,298K}$ Gas Opt	5.9042×10^{-5}	1.6800×10^{1}	7.2695×10^{2}	1.0402×10^{-5}
$K'_{aq,298K}$ Aq Opt	1.3634×10^{-6}	3.4425×10^{-2}	3.3427×10^{-5}	1.6640×10^{-6}

[a] In each reaction, alanine anion has 7 explicit H_2O molecules, α-ketoglutarate has 9 explicit H_2O molecules, pyruvate has 8 explicit H_2O molecules, and glutamate has 8 explicit H_2O molecules.

Table 5.3. Reduced Partition Function Ratios at Various Carbon Atoms at 298 K.

Species	# of H_2O Molecules	^{13}C Site	$\beta_{gas\ opt}$	$\beta_{aq\ opt}$	$\Delta\beta$ (‰) [a]
Alanine	7	Methyl	1.1434	1.1439	-0.484
		C2	1.1727	1.1736	-0.857
		C-terminus	1.1986	1.1954	3.202
Alanine, Zwitterion	7	Methyl	1.1446	1.1448	-0.276
		C2	1.1731	1.1734	-0.311
		C-terminus	1.1986	1.1941	4.460
Alanine, Deprotonated	7	Methyl	1.1418	1.1413	0.500
		C2	1.1725	1.1733	-0.809
		C-terminus	1.1986	1.1941	4.533
Pyruvic Acid	6	Methyl	1.1416	1.1412	0.316
		C2	1.1836	1.1825	1.101
		Carboxylic Acid	1.1936	1.1904	3.160
Pyruvate	6	Methyl	1.1392	1.1391	0.151
		C2	1.1838	1.1841	-0.297
		Carboxylate	1.1954	1.1934	1.951
Pyruvate	8	Methyl	1.1390	1.1391	-0.129
		C2	1.1828	1.1839	-1.080
		Carboxylate	1.1948	1.1932	1.624

[a] $\Delta\beta$ is calculated as $(\beta_{gas} - \beta_{aq})*1000$.

Table 5.4. K and ϵ Values for the Reactions of Interest With and Without Heavy Isotopes for Deprotonated Alanine and Pyruvate in the Gas Phase and in Aqueous Solution at 298 K.[a]

Reaction	Opt	K_l	$K_{h,Me}$ [b]	$K_{h,C2}$ [b]	$K_{h,carboxylate}$ [d]	ϵ_{Me} [e]	ϵ_{C2} [e]	$\epsilon_{C\text{-}term}$ [e]
(5.1)	Gas	1.2149	1.2119	1.2255	1.2110	-2.4747	8.7729	-3.1521
	Aqueous	0.8217	0.8202	0.8291	0.8211	-1.9253	8.9977	-0.7278
(5.2)	Gas	0.1889	0.1884	0.1905	0.1883	-2.4747	8.7729	-3.1521
	Aqueous	0.3703	0.3696	0.3737	0.3701	-1.9253	8.9977	-0.7278
(5.3)	Gas	2.8771	2.8631	2.9009	2.8680	-4.8516	8.2957	-3.1348
	Aqueous	1.0220	1.0169	1.0312	1.0212	-4.9791	8.9485	-0.7718
(5.4)	Gas	1.5423	1.5385	1.5559	1.5375	-2.4747	8.7729	-3.1521
	Aqueous	0.5882	0.5871	0.5935	0.5878	-1.9253	8.9977	-0.7278

[a] Calculation includes Q values for deprotonated alanine with 7 explicit H_2O molecules and pyruvate with 8 explicit H_2O molecules. [b] ^{13}C at the methyl. [c] ^{13}C at the C2 site. [d] ^{13}C at the carboxylate. [e] ϵ values are given in ‰.

Table 5.5. Gas Phase Isotope Exchange Values at 298 K at Various Computational Levels in kcal/mol.

	B3LYP/ 6-311++ G(2d,2p) [a]	LSDA/ 6-311++ G(2d,2p) [a]	BP86/ 6-311++ G(2d,2p) [a]	B3LYP/ DZVP2 [b]	LSDA/ DZVP2 [b]	BP86/ DZVP2 [b]	MP2/aD [b]	MP2/aT [b]	MP2/aT [a]
Reaction (5.15)									
Pyruvate ΔZPE	0.1502	0.1492	0.1433	0.1513	0.1501	0.1445	0.1520	0.1513	
Deprotonated Alanine ΔZPE	0.1416	0.1422	0.1362	0.1435	0.1441	0.1380	0.1447	0.1452	
ΔΔZPE	-0.0086	-0.0070	-0.0071	-0.0078	-0.0060	-0.0064	-0.0074	-0.0062	
ϵ_{ZPE} (‰)	14.5	11.9	12.0	15.6	14.3	14.8	13.3	11.0	
ϵ (‰)	13.1	10.8	10.8	11.9	9.3	9.7	11.2	9.4	
Reaction (5.16)									
Pyruvate ΔZPE	0.1497	0.1480	0.1428	0.1513	0.1501	0.1445	0.1520	0.1513	
Alanine ΔZPE	0.1439	0.1436	0.1387	0.1477	0.1479	0.1424	0.1498	0.1501	
ΔΔZPE	-0.0058	-0.0044	-0.0041	-0.0036	-0.0022	-0.0021	-0.0022	-0.0012	
ϵ_{ZPE} (‰)	9.4	7.2	6.7	6.1	3.7	3.5	3.7	2.1	
ϵ (‰)		0.40		5.3	3.3	2.9	3.1	1.5	
Reaction (5.17)									
Formaldehyde ΔZPE	0.1119	0.1096	0.1082	0.1121	0.1103	0.1082	0.1099	0.1109	0.1110
Methylamine ΔZPE	0.1058	0.1044	0.1031	0.1068	0.1060	0.1037	0.1076	0.1082	0.1083
ΔΔZPE	-0.0061	-0.0052	-0.0051	-0.0053	-0.0043	-0.0045	-0.0023	-0.0027	-0.0027
ϵ_{ZPE} (‰)	9.8	8.4	8.4	10.4	7.0	8.2	3.9	4.6	4.4
ϵ (‰)	12.6	11.5	11.6	13.0	10.2	10.8	7.0	7.7	7.4

[a] Calculated using NWChem (ref 10). [b] Calculated using Gaussian 16 (ref 39).

Table 5.6. Site Specific β Values for Non-Solvated H_2CO, CH_3NH_2, $CH_3NH_3^+$, Alanine, Deprotonated Alanine, Protonated Alanine, Pyruvic Acid, and Pyruvate in the Gas Phase at 298 K at the MP2/aT Level.

	$\beta = Q_h/Q_l$		
H_2CO	1.1523		
CH_3NH_2	1.1435		
$CH_3NH_3^+$	1.1427		
	^{13}C Methyl	^{13}C Central Carbon	^{13}C COOH/COO^-
Alanine	1.1414	1.1771	1.2013
Deprotonated Alanine	1.1422	1.1677	1.1928
Pyruvic Acid	1.1416	1.1825	1.1929
Pyruvate	1.1363	1.1802	1.1922

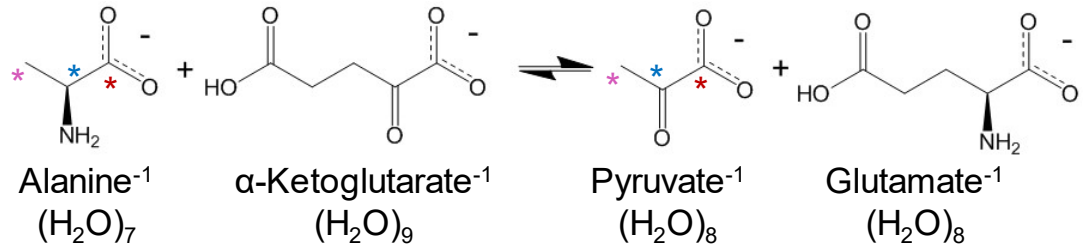
Table 5.7. Estimated Reduced Masses and Frequencies for ^{12}C and ^{13}C Bonds with N, O, and H.

	ω_{12} (cm^{-1})	ω_{13} (cm^{-1})
$^{12/13}C$-^{14}N	1100	1076.98
$^{12/13}C$=^{16}O	1750	1711.11
$^{12/13}C$-^{1}H	3000	2991.11

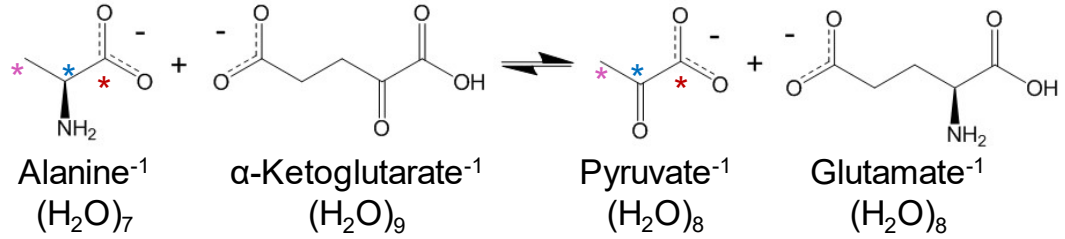
Table 5.8. Estimated Free Energy (A), ΔA, and ϵ_{ZPE} in kcal/mol for ^{12}C and ^{13}C Bonds with N, O, and H in kcal/mol at 298 K.

	A ^{12}C			A ^{13}C		
	ZPE	Thermal	Total	ZPE	Thermal	Total
C-N	1.57	-2.92×10^{-3}	1.57	1.54	-3.27×10^{-3}	1.54
C=O	2.50	-1.26×10^{-4}	2.50	2.45	-1.52×10^{-4}	2.45
C-H	4.29	-3.01×10^{-7}	4.29	4.28	-3.14×10^{-7}	4.28
	$^{12}C{=}O + {}^{13}CHNH_2 \leftrightarrows {}^{13}C{=}O + {}^{12}CHNH_2$					
	ΔA	K	ϵ_{ZPE} (‰)			
ZPE + Therm	-9.67×10^{-3}	-0.9844	15.7			
ZPE Only	-9.98×10^{-3}	-0.9839	16.2			

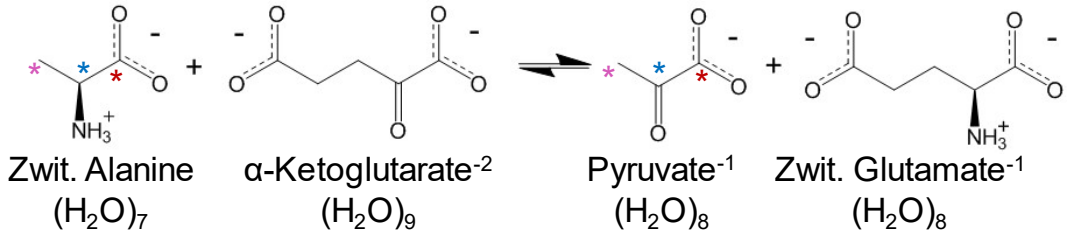
Figures



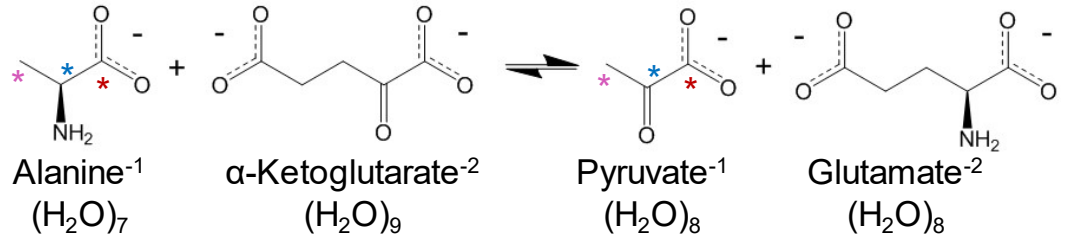
(5.1)



(5.2)



(5.3)



(5.4)

Figure 5.1. Reaction (5.1) is that for the alanine anion and α-ketoglutarate deprotonated at the α-keto moiety (-C(O)C(O)O$^-$) forming pyruvate and glutamate deprotonated at the C-terminus. Reaction (5.2) has α-ketoglutarate deprotonated at the glutarate moiety (-CH$_2$C(O)O$^-$) and glutamate deprotonated at the side chain. Reaction (5.3) represents each chemical species as it would be present at biological pH. Reaction (5.4) has all carboxylate groups deprotonated for all species. H$_2$O molecules not shown for clarity. Red, blue, and purple asterisks (*) indicate sites of ^{13}C substitution at the carboxylate group, the C2 position, and the methyl group, respectively.

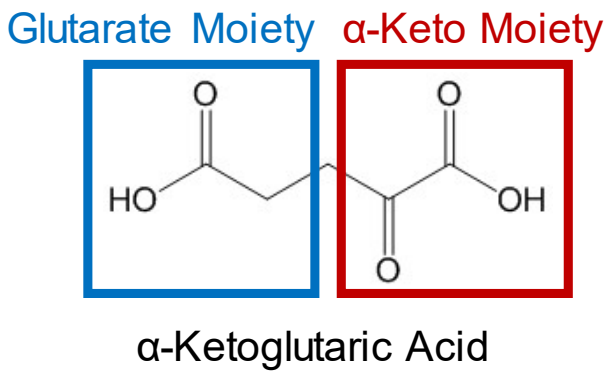
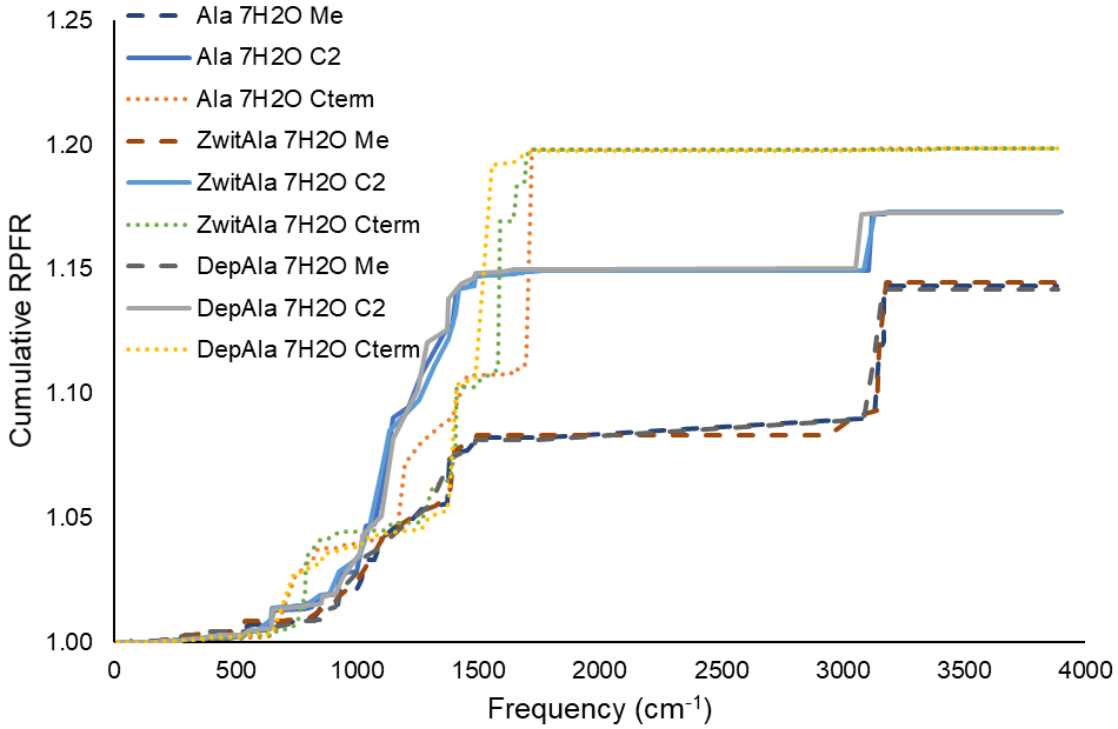
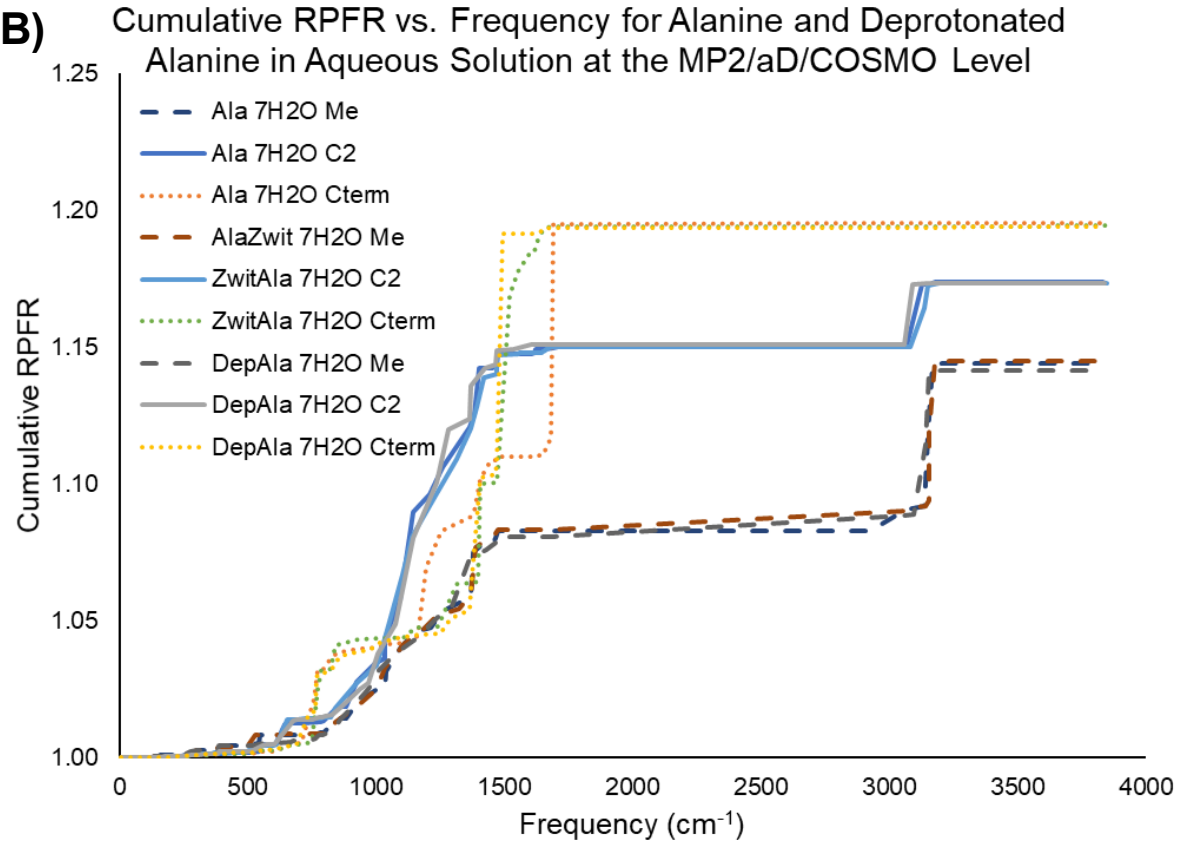


Figure 5.2. Carboxylic acid deprotonation sites for α-ketoglutaric acid. The glutarate moiety is analogous to the side chain of glutamic acid and the α-keto moiety is analogous to the C-terminus of glutamic acid.

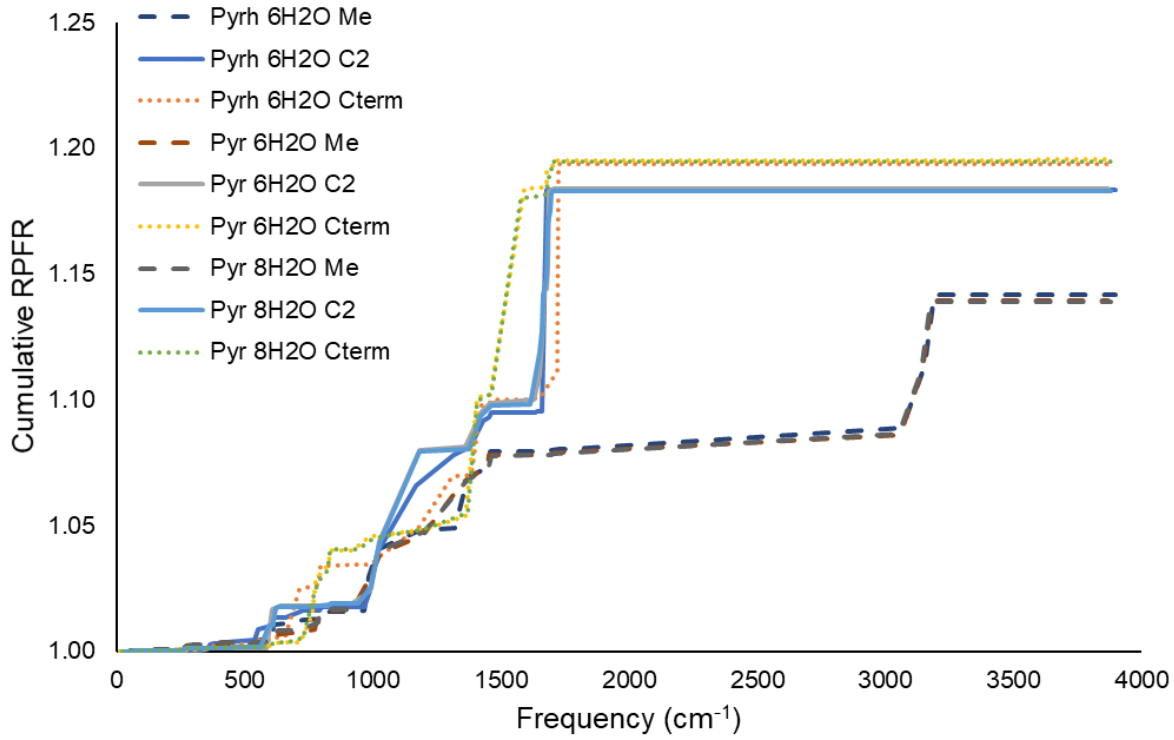
(A) Cumulative RPFR vs. Frequency for Alanine and Deprotonated Alanine in the Gas Phase at the MP2/aD Level



(B)


Cumulative RPFR vs. Frequency for Alanine and Deprotonated Alanine in Aqueous Solution at the MP2/aD/COSMO Level
Ala 7H2O Me
Ala 7H2O C2
Ala 7H2O Cterm
AlaZwit 7H2O Me
ZwitAla 7H2O C2
ZwitAla 7H2O Cterm
DepAla 7H2O Me
DepAla 7H2O C2
DepAla 7H2O Cterm
Cumulative RPFR
1.25
1.20
1.15
1.10
1.05
1.00
0
500
1000
1500
2000
2500
3000
3500
4000
Frequency (cm-1)

(C) Cumulative RPFR vs. Frequency for Pyruvic Acid and Pyruvate in the Gas Phase at the MP2/aD Level



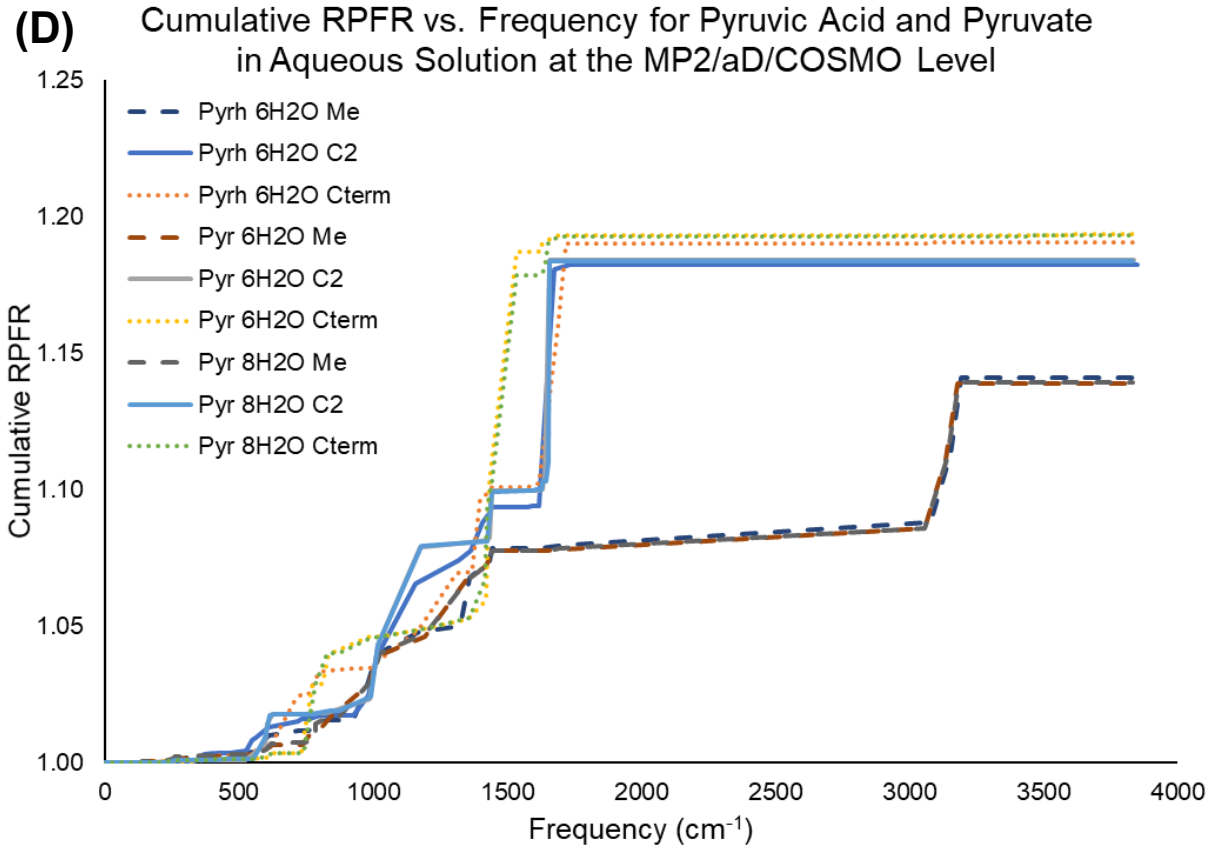
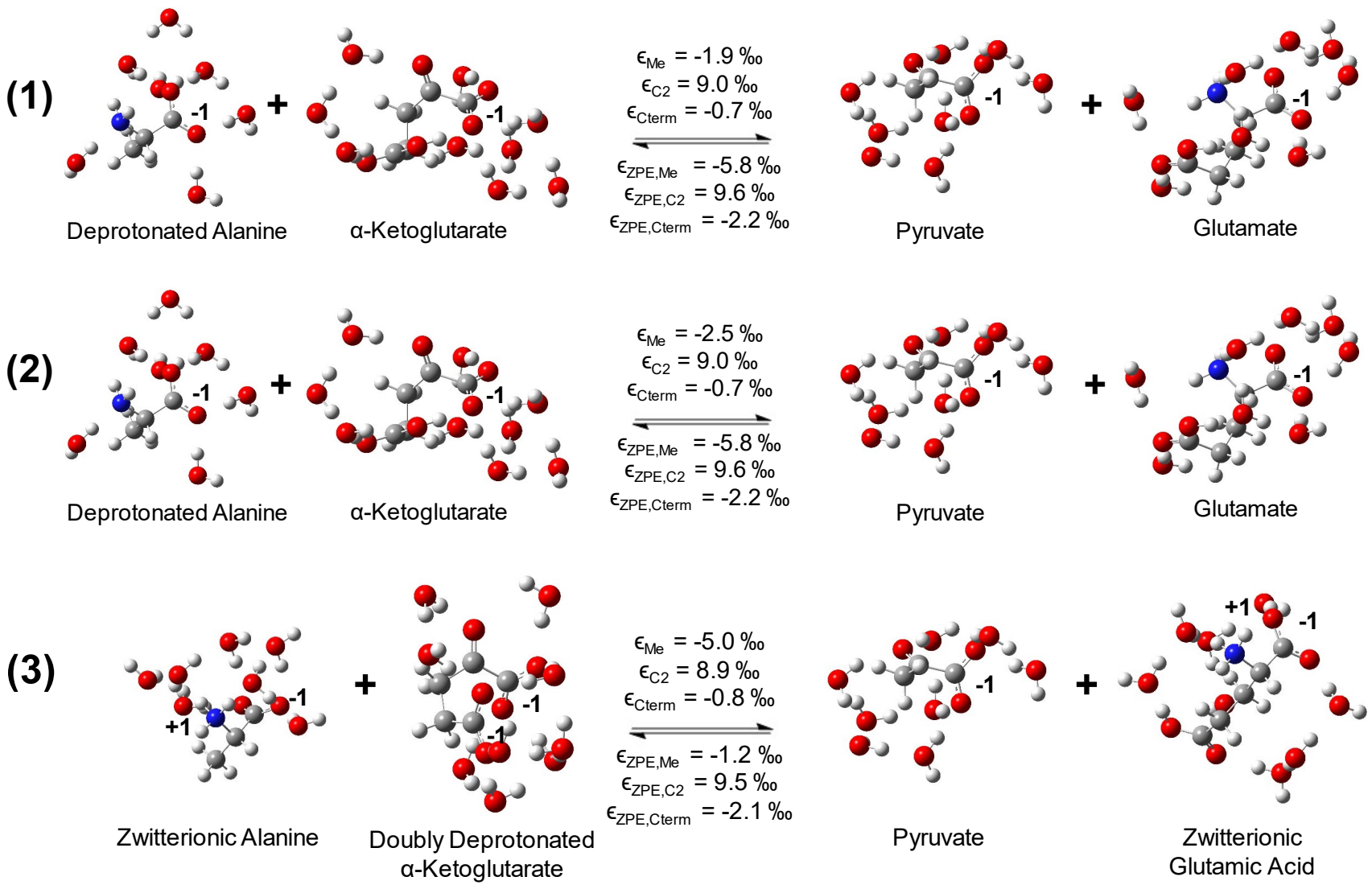


Figure 5.3. Cumulative reduced partition function ratio (RPFR) vs. vibrational frequency (cm^{-1}) at 298 K for ^{13}C at either the methyl (Me, dashed lines), C2 site (solid lines), or the C-terminus or equivalent (dotted lines). (A) and (B) are alanine in its NCS (Ala) and zwitterionic (ZwitAla) structures and deprotonated alanine (DepAla) with 7 explicit H_2O molecules in the gas phase and in aqueous solution, respectively. (C) and (D) are pyruvic acid (Pyrh) and pyruvate (Pyr) with 6 and 8 explicit H_2O molecules in the gas phase and in aqueous solution, respectively. Pyruvic acid-related results are labeled to correspond to analogous sites in alanine. All results are optimized at the MP2/aD level for the gas phase and MP2/aD/COSMO for the aqueous solution.


(1)
-1
-1
Deprotonated Alanine
α-Ketoglutarate
ϵ_{Me} = -1.9 ‰
ϵ_{C2} = 9.0 ‰
ϵ_{Cterm} = -0.7 ‰
$\epsilon_{ZPE,Me}$ = -5.8 ‰
$\epsilon_{ZPE,C2}$ = 9.6 ‰
$\epsilon_{ZPE,Cterm}$ = -2.2 ‰
-1
-1
Pyruvate
Glutamate
(2)
-1
-1
Deprotonated Alanine
α-Ketoglutarate
ϵ_{Me} = -2.5 ‰
ϵ_{C2} = 9.0 ‰
ϵ_{Cterm} = -0.7 ‰
$\epsilon_{ZPE,Me}$ = -5.8 ‰
$\epsilon_{ZPE,C2}$ = 9.6 ‰
$\epsilon_{ZPE,Cterm}$ = -2.2 ‰
-1
-1
Pyruvate
Glutamate
(3)
+1
-1
-1
-1
Zwitterionic Alanine
Doubly Deprotonated
α-Ketoglutarate
ϵ_{Me} = -5.0 ‰
ϵ_{C2} = 8.9 ‰
ϵ_{Cterm} = -0.8 ‰
$\epsilon_{ZPE,Me}$ = -1.2 ‰
$\epsilon_{ZPE,C2}$ = 9.5 ‰
$\epsilon_{ZPE,Cterm}$ = -2.1 ‰
-1
+1
-1
Pyruvate
Zwitterionic
Glutamic Acid

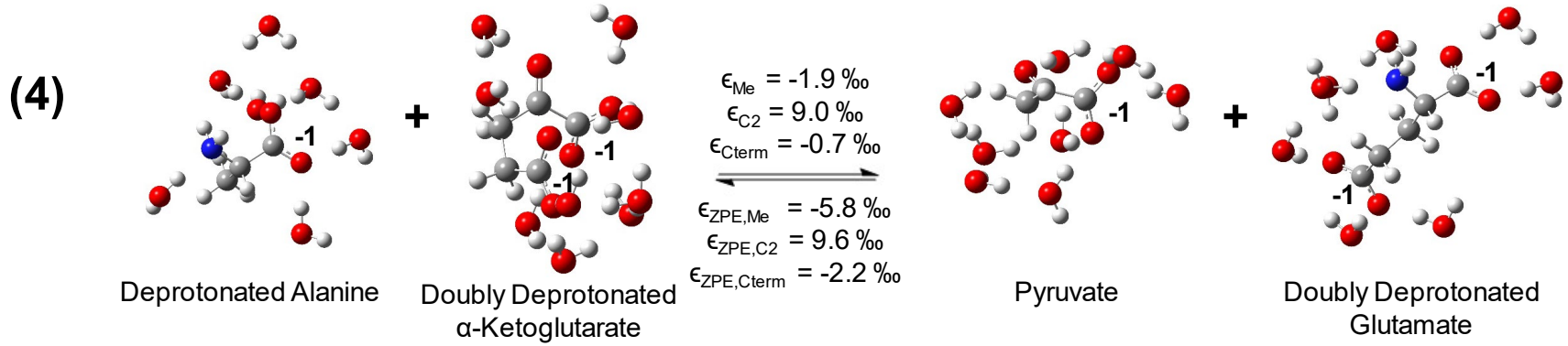
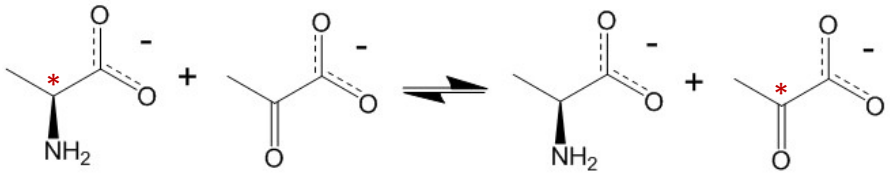
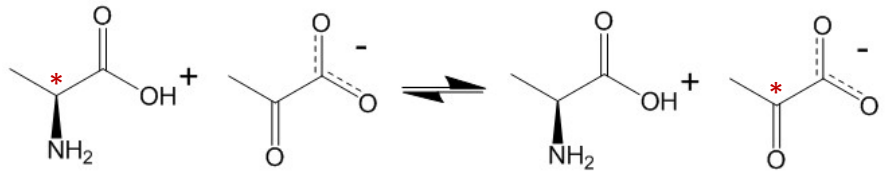


Figure 5.4. Isotope fractionation factors ϵ and ϵ_{ZPE} at each $^{13}C/^{12}C$ substitution site of alanine and pyruvate for reactions (5.1) through (5.4) with aqueous optimized, explicitly hydrated clusters. Calculated ϵ_{ZPE} values are solely based on the form of alanine and pyruvate in equilibrium with one another.



(5.15)



(5.16)

(5.17)

Figure 5.5. ^{13}C exchange between (5.15) deprotonated alanine and pyruvate, (5.16) neutral non-charge separated (NCS) alanine and pyruvate, and (5.17) formaldehyde (H_2CO) and methylamine (CH_3NH_2). The red asterisk indicates the location of the heavy carbon.

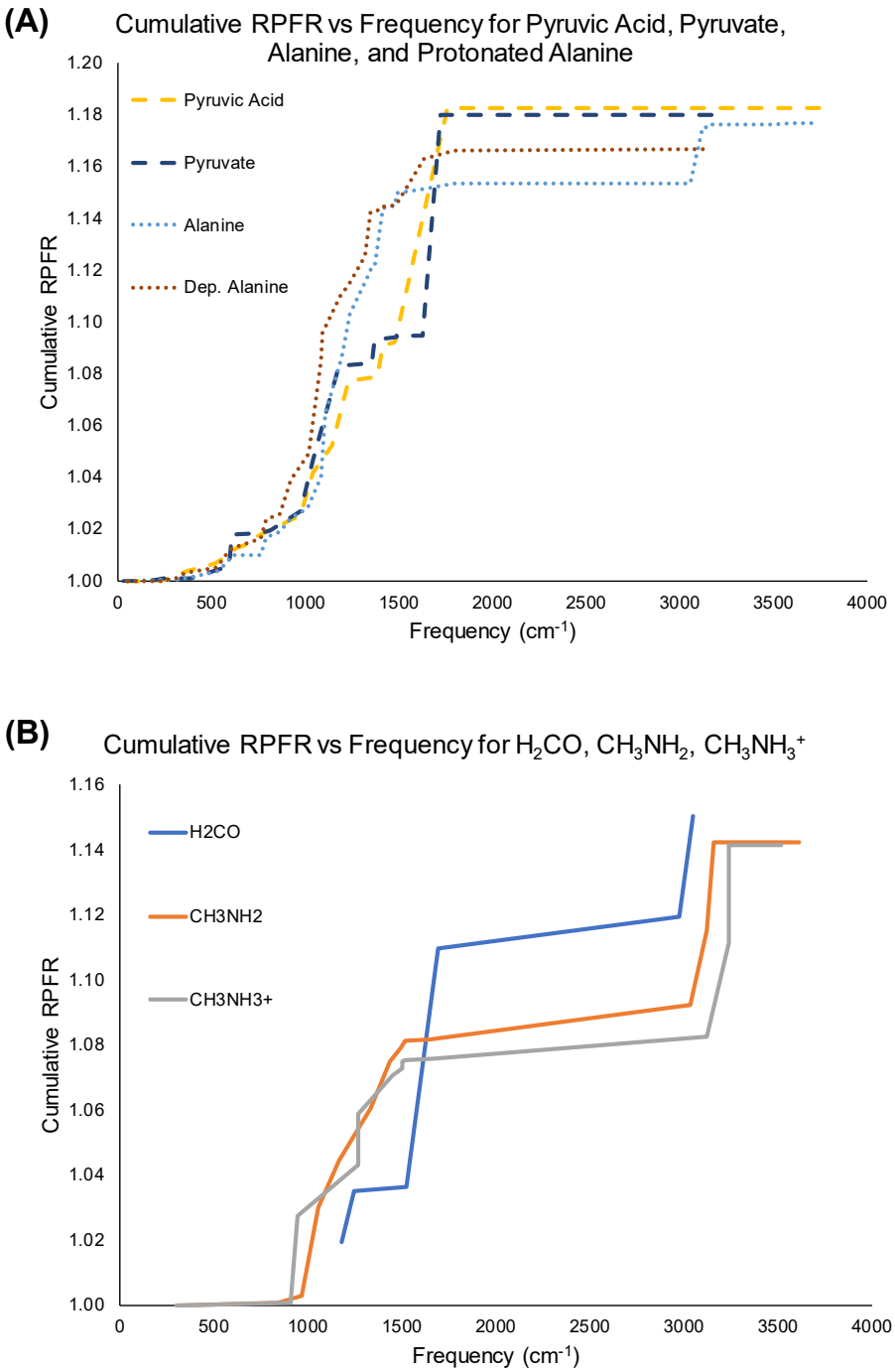


Figure 5.6. Cumulative reduced partition function ratio (RPFR) vs frequency in cm^{-1} for formaldehyde, methylamine, protonated methylamine, pyruvic acid, pyruvate, alanine, and deprotonated alanine in the gas phase at 298 K. All pyruvate and alanine species are ^{13}C-substituted at the C2 position.

References

[1] Karmen, A.; Wróblewski, F.; LaDue, J. S. Transaminase Activity in Human Blood. *J. Clin. Invest.* **1955**, *34*, 126 – 133.

[2] Lala, V.; Minter, D. A. *Liver Function Tests*. StatPearls Publishing: Treasure Island, FL, 2019.

[3] Liu, Z.; Que, S.; Zu, J.; Peng, T. Alanine Aminotransferase – Old Biomarker and New Concept: A Review. *Int. J. Med. Sci.* **2014**, *11*, 925 – 935.

[4] Luo, Y.-R. *Comprehensive Handbook of Chemical Bond Energies*, CRC Press, Taylor and Francis Group, 2007.

[5] Krebs, H. A. Equilibria in Transamination Systems. *Biochem. J.* **1953**, *54*, 82 – 86.

[6] Segal, H. L.; Beattie, D. S.; Hopper, S. Purification and Properties of Liver Glutamic-Alanine Transaminase from Normal and Corticoid-Treated Rats. *J. Biol. Chem.* **1962**, *237*, 1914 – 1920.

[7] Martinez-Carrion, M.; Jenkins, W. T. D-Alanine-D-Glutamate Transaminase. I. Purification and Characterization. *J. Biol. Chem.* **1965**, *240*, 3538 – 3551.

[8] Miller, S. L.; Smith-Magowan, D. The Thermodynamics of the Krebs Cycle and Related Compounds. *J. Phys. Chem. Ref. Data* **1990**, *19*, 1049 – 1073.

[9] Car, R.; Parrinello, M. Unified Approach to Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett*. **1985**, *55*, 2471 – 2474.

[10] Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; et al. NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Comput. Phys. Commun*. **2010**, *181*, 1477 – 1489.

[11] Bylaska, E. J.; Tsemekhman, K.; Govind, N.; Valiev, M. Large-Scale Plane-Wave-Based Density Functional Theory: Formalism, Parallelization, and Applications. In *Computational Methods for Large Systems: Electronic Structure Approaches for Biotechnology and Nanotechnology*. Reimers, J. R., Ed.; John Wiley and Sons: Hoboken, NJ, 2011; pp. 77 – 116.

[12] Bylaska, E. J. Plane-Wave DFT Methods for Chemistry. In *Annual Reports in Computational Chemistry*. Vol. 13, Elsevier, Amsterdam, 2017 Ch. 5, pp. 185 – 228.

[13] Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136,* B864 – B871.

[14] Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, **1965**, *140*, A1133 – A1138.

[15] Cauët, E.; Bogatko, S.; Weare, J. H.; Fulton, J. L.; Schenter, G. K.; Bylaska, E. J. Structure and Dynamics of the Hydration Shells of the Zn^{2+} Ion from Ab Initio Molecular Dynamics and

Combined Ab Initio and Classical Molecular Dynamics Simulations. *J. Chem. Phys*. **2010**, *132*, 194502-1 – 194502-13.

[16] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett*. **1996**, *77*, 3865 – 3868.

[17] Hamann, D. R.; Schlüter, M.; Chiang, C. Norm-Conserving Pseudopotentials. *Phys. Rev. Lett.* **1979**, *43*, 1494 – 1497.

[18] Hamann, D. R. Generalized Norm-Conserving Pseudopotentials. *Phys. Rev. B* **1989**, *40*, 2980 – 2987.

[19] Kleinman, L.; Bylander, D. M. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett*. **1982**, *48*, 1425 – 1428.

[20] Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys*. **1984**, *52*, 255 – 268.

[21] Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev*. **1985**, *31*, 1695 – 1697.

[22] Blöchl, P. E.; Parrinello, M. Adiabaticity in First-Principles Molecular Dynamics. *Phys. Rev. B* **1992**, *45*, 9413 – 9416.

[23] Shrake, A.; Rupley, J. A. Environment and Exposure to Solvent of Protein Atoms. Lysozyme and Insulin. *J. Mol. Biol*. **1973**, *79*, 351 – 371.

[24] Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269 – 6271.

[25] Environmental Molecular Sciences Laboratory. EMSL Arrows 2D Molecular and Reaction Editor. https://arrows.emsl.pnnl.gov/api/ (accessed Mar 11, 2018).

[26] Stillinger, F. H.; Weber, T. A. Inherent Structure Theory of Liquids in the Hard-Sphere Limit. *J. Chem. Phys.* **1985**, *83*, 4767 – 4775.

[27] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.

[28] Lee, C.; Yang, W.; Parr, R. G. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Phys. Rev. B* **1988**, *37*, 785 – 789.

[29] Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. Optimization of Gaussian-Type Basis-Sets for Local Spin-Density Functionals Calculations. 1. Boron Through Neon, Optimization Technique and Validation. *Can. J. Chem.* **1992**, *70*, 560 – 571.

[30] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 0618 – 0622.

[31] Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical Models Incorporating Electron Correlation. *Int. J. Quantum. Chem.* **1976**, *10*, 1 – 19.

[32] Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys*. **1989**, *90*, 1007 – 1023.

[33] Kendall, R. A.; Dunning, T. H. Jr.; Harrison, R. J., Electron Affinities of the First Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys*. **1992,** *96*, 6796 – 6806.

[34] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Contiuum Solvation Models. *Chem. Rev.* **2005**, 105, 2999 – 3093.

[35] Klamt, A. *Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*. Elsevier: Amsterdam, 2005.

[36] Klamt, A.; Schüürmann, G. J. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *Chem. Soc. Perkin Trans.* **1993**, 2, 799 – 805.

[37] Feyereisen, M. W.; Feller, D.; Dixon, D. A. The Hydrogen Bond Energy of the Water Dimer. *J. Phys. Chem*. **1996**, *100*, 2993 – 2997

[38] Rustad, J. R.; Nelmes, S. L.; Jackson, V. E.; Dixon, D. A. Quantum-Chemical Calculations of Carbon-Isotope Fractionation in CO_2(g), Aqueous Carbonate Species, and Carbonate Minerals. *J. Phys. Chem. A* **2008**, *112*, 542 – 555.

[39] Gaussian 16, Revision E.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; et al., Gaussian, Inc., Wallingford CT, 2004.

[40] Slater, J. C. *The Self-Consistent Field for Molecular and Solids*, *Quantum Theory of Molecular and Solids.* McGraw-Hill: New York, 1974; Vol. 4.

[41] Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200 – 1211.

[42] Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*, 8822 – 8824.

[43] Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *88*, 3098 – 3100.

[44] Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. IX. Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724 – 728.

[45] Lide, D. R. (E.) *CRC Handbook of Chemistry and Physics*, 86th Ed. CRC Press, Boca Raton, FL, USA, 2006.

[46] Barb, A. W.; Hekmatyar, S. K.; Glushka, J. N.; Prestegard, J. H. Probing Alanine Transaminase Catalysis with Hyperpolarized $^{13}CD_3$-Pyruvate. *J. Magn. Reson.* **2013**, *228*, 59 – 65.

[47] Rustad, J. R. Ab Initio Calculation of the Carbon Isotope Signatures of Amino Acids. *Org. Geochem.* **2009**, *40*, 720 – 723.

[48] Silverstein, R. M.; Bassler, G. C.; and Morrill, T. C. *Spectrometric Identification of Organic Compounds*, 4th ed. John Wiley and Sons: New York, 1981.

[49] McQuarrie, D. A. *Statistical Mechanics.* University Science Books: Sausalito, CA, 2000.

[50] Glavin, D. P.; Bada, J. L.; Brinton, K. L. F.; McDonald, G. D. Amino Acids in the Martian Meteorite Nakhla. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8835 – 8838.

[51] Koga, T.; Naraoka, H. A New Family of Extraterrestrial Amino Acids in the Murchison Meteorite. *Sci. Rep.* **2017**, *7*, 636-1 – 636-8.

[52] Chan, Q. H. S.; Zolensky, M. E.; Kebukawa, Y.; Fries, M.; Ito, M.; Steele, A.; Rahman, Z.; Nakato, A.; Kilcoyne, A. L. D.; Suga, H.; et al. Organic Matter in Extraterrestrial Water-Bearing Salt Crystals. *Sci. Adv.* **2018**, *4*, eaao3521-1 – eaao3521-10.

Full Reference 10 (NWChem): Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Comput. Phys. Commun*. **2010**, *181*, 1477 – 1489.

Full Reference 38 (Gaussian 16): Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision A.03, Gaussian, Inc., Wallingford CT, 2016.

Full Reference 51: Chan, Q. H. S.; Zolensky, M. E.; Kebukawa, Y.; Fries, M.; Ito, M.; Steele, A.; Rahman, Z.; Nakato, A.; Kilcoyne, A. L. D.; Suga, H.; Takahashi, Y.; Takeichi, Y.; Mase, K. Organic Matter in Extraterrestrial Water-Bearing Salt Crystals. *Sci. Adv.* **2018**, *4*, eaao3521-1 – eaao3521-10.

Discussion of β Values for Glutamic Acid and α-Ketoglutaric Acid. Glutamic acid and its related species are modeled with 8 explicit H_2O molecules. Neutral glutamic acid is modeled as non-charge separated (NCS) and as a zwitterion with the positive charge at the N-terminus and the negative charge at either the C-terminus or the carboxylate of the side chain. All forms and derivatives of glutamic acid are modeled with all isotopically normal atoms and with ^{13}C at either the C-terminus, the C2 position, or the carboxylic acid/carboxylate of the side chain. All glutamic acid-related compounds have C-terminus and side chain β_{aq} values that are very similar to those for the C-terminus of alanine with the largest β_{aq} value 1.1967 for NCS zwitterionic glutamic acid with the negative charge on the side chain and the smallest β_{aq} value 1.1929 for doubly deprotonated glutamate. Additionally, the C2 position of all glutamic acid-related compounds have β_{aq} values that are very similar to those of the C2 position of alanine and its derivatives with the largest β_{aq} value 1.1742 for zwitterionic glutamic acid with the negative charge on the side chain and the smallest β_{aq} value 1.1717 for zwitterionic glutamic acid with the negative charge at the C-terminus. ^{13}C substitution at the C-terminus is the most favorable site for non-charge separated glutamic acid, the zwitterionic glutamic acid with the negative charge at the side chain, glutamate deprotonated at the C-terminus, and doubly deprotonated glutamate by 0.7, 4.6, 0.7, and 0.8 ‰, respectively. ^{13}C substitution at the side chain is the most favorable site for the zwitterionic glutamic acid with the negative charge at the C-terminus and glutamate

deprotonated at the side chain by 2.5 and 1.8 ‰, respectively. The zwitterionic form of the glutamate anion has similar β values to the glutamate anions with NCS backbones. The differences between β_{aq} values for the zwitterionic glutamate anion vs the anions with NCS backbones are less than 0.5 ‰. The only exception is for ^{13}C substitution of the side chain carboxylic acid for the NCS glutamate anion deprotonated at the C-terminus, which is ~ 4 ‰ larger than the value for ^{13}C substitution in the zwitterionic glutamate anion. Thus, the zwitterionic backbone of glutamate makes only very minor differences in β. For gas phase optimized structures, the values for zwitterionic glutamate anion and glutamate anion deprotonated at the side chain with a NCS backbone are also very similar, with the largest difference less than 2 ‰ for ^{13}C substitution at the C2 position.

α-Ketoglutaric acid and its related species are modeled with 9 explicit H_2O molecules. The β_{aq} values of α-ketoglutaric acid and its derivatives for its ^{13}C substitution at the α-keto moiety carboxylic acid/carboxylate site and the glutarate moiety carboxylic acid/carboxylate site are in good agreement with similar ^{13}C sites in alanine, glutamic acid, and pyruvic acid. Substitution of ^{13}C at the carboxylic acid/carboxylate of the α-keto moiety yields a range of β_{aq} values from 1.1868 for doubly deprotonated α-ketoglutarate to 1.1929 for α-ketoglutarate deprotonated at the α-keto moiety. Substitution of ^{13}C at the carboxylic acid/carboxylate of the glutarate moiety yields a range of β_{aq} values from 1.1925 for α-ketoglutarate deprotonated at the glutarate moiety to 1.1950 for α-ketoglutarate deprotonated at the α-keto moiety. The β_{aq} values for C2 site ^{13}C substitution range from 1.1783 for doubly deprotonated α-ketoglutarate to 1.1826 for α-ketoglutarate deprotonated at the α-keto moiety. The average β_{aq} for the C2 site of α-ketoglutaric acid and its derivatives is 1.1808, which is 6.6 ‰ higher than the average β_{aq} of ^{13}C

at the C2 position of alanine, indicating that this site is much more similar to the C2 position of pyruvic acid than that of alanine or glutamic acid.

Table A5.1. Parameters for the QM/MM Potential and the Toukan-Rahman Water Potential.

QM/MM		**SPC/E MM inter-atomic**	
$A_{CO\text{^}}$	$4\sqrt{\varepsilon_C\varepsilon_O}\left(\frac{\sigma_C+\sigma_O}{2}\right)^{12}$	$q_{O\text{^}}$	-0.8476e
$B_{CO\text{^}}$	$4\sqrt{\varepsilon_C\varepsilon_O}\left(\frac{\sigma_C+\sigma_O}{2}\right)^{6}$	$q_{H\text{^}}$	0.4238e
$A_{NO\text{^}}$	$4\sqrt{\varepsilon_N\varepsilon_O}\left(\frac{\sigma_N+\sigma_O}{2}\right)^{12}$	r_{OH}	1Å
$N_{NO\text{^}}$	$4\sqrt{\varepsilon_N\varepsilon_O}\left(\frac{\sigma_N+\sigma_O}{2}\right)^{6}$	$\angle HOH$	109.47°
$A_{OO\text{^}}$	$4\varepsilon_O\sigma_O^{12}$	$A_{O\text{^}O\text{^}}$	$4\varepsilon_O\sigma_O^{12}$
$B_{OO\text{^}}$	$4\varepsilon_O\sigma_O^{6}$	$B_{O\text{^}O\text{^}}$	$4\varepsilon_O\sigma_O^{6}$
$A_{HO\text{^}}$	$4\sqrt{\varepsilon_H\varepsilon_O}\left(\frac{\sigma_H+\sigma_O}{2}\right)^{12}$	$A_{O\text{^}H\text{^}}, A_{H\text{^}H\text{^}}$	0.0
$B_{HO\text{^}}$	$4\sqrt{\varepsilon_H\varepsilon_O}\left(\frac{\sigma_H+\sigma_O}{2}\right)^{6}$	$B_{O\text{^}H\text{^}}, B_{H\text{^}H\text{^}}$	0.0
$A_{CH\text{^}}, A_{NH\text{^}}, A_{OH\text{^}}, A_{HH\text{^}}$	0.0		
$B_{CH\text{^}}, B_{NH\text{^}}, B_{OH\text{^}}, B_{HH\text{^}}$	0.0		
ε_C= 0.100 kcal/mol σ_C= 3.41Å ε_N= 0.069 kcal/mol σ_N= 3.26Å ε_O= 0.155 kcal/mol σ_O= 3.17Å ε_H= 0.044 kcal/mol σ_H= 0.70Å			

Table A5.2. Site Specific β Values for H_2CO, CH_3NH_2, $CH_3NH_3^+$, Alanine, Deprotonated Alanine, Pyruvic Acid, and Pyruvate in the Gas Phase at 298 K at the MP2/aD Level.

	$\beta = Q_h/Q_l$		
H_2CO	1.1504		
CH_3NH_2	1.1424		
$CH_3NH_3^+$	1.1415		
	^{13}C Methyl	^{13}C Central Carbon	^{13}C COOH/COO^-
Alanine	1.1405	1.1765	1.1983
Deprotonated Alanine	1.1414	1.1669	1.1990
Pyruvic Acid	1.1406	1.1810	1.1899
Pyruvate	1.1355	1.1788	1.1899

Table A5.3. $^{13}C/^{12}C$ Exchange Thermodynamics for Explicitly Solvated Pyruvate and Alanine in kcal/mol.

Gas Phase Optimized, ^{13}C Substituted		
Pyruvate $\Delta ZPE_{gas,Me}$	0.1182	
Alanine Zwitterion $\Delta ZPE_{gas,Me}$	0.1220	
Deprotonated Alanine $\Delta ZPE_{gas,Me}$		0.1188
$\Delta\Delta ZPE_{gas,Me}$	0.0039	0.0007
$1000*\ln K_{ZPE,gas,Me}$ (‰)	-6.5	-1.1
Pyruvate $\Delta ZPE_{gas,C2}$	0.15451	
Alanine Zwitterion $\Delta ZPE_{gas,C2}$	0.1498	
Deprotonated Alanine $\Delta ZPE_{gas,C2}$		0.1491
$\Delta\Delta ZPE_{gas,C2}$	-0.0047	-0.0054
$1000*\ln K_{ZPE,gas,C2}$ (‰)	7.8	9.2
Pyruvate $\Delta ZPE_{gas,Cterm}$	0.1621	
Alanine Zwitterion $\Delta ZPE_{gas,Cterm}$	0.1648	
Deprotonated Alanine $\Delta ZPE_{gas,Cterm}$		0.1646
$\Delta\Delta ZPE_{gas,Cterm}$	0.0027	0.0025
$1000*\ln K_{ZPE,gas,Cterm}$ (‰)	-4.6	-4.3
Aqueous Optimized, ^{13}C Substituted		
Pyruvate $\Delta ZPE_{aq,Me}$	0.1181	
Alanine Zwitterion $\Delta ZPE_{gas,Me}$	0.1215	
Deprotonated Alanine $\Delta ZPE_{aq,Me}$		0.1188
$\Delta\Delta ZPE_{aq,Me}$	0.0035	0.0007
$1000*\ln K_{ZPE,aq,Me}$ (‰)	-5.8	-1.2
Pyruvate $\Delta ZPE_{aq,C2}$	0.1553	
Alanine Zwitterion $\Delta ZPE_{gas,C2}$	0.1496	
Deprotonated Alanine $\Delta ZPE_{aq,C2}$		0.1496
$\Delta\Delta ZPE_{aq,C2}$	-0.0057	-0.0057
$1000*\ln K_{ZPE,aq,C2}$ (‰)	9.6	9.5
Pyruvate $\Delta ZPE_{aq,Cterm}$	0.1612	
Alanine Zwitterion $\Delta ZPE_{gas,Cterm}$	0.1625	
Deprotonated Alanine $\Delta ZPE_{aq,Cterm}$		0.1624
$\Delta\Delta ZPE_{aq,Cterm}$	0.0013	0.0013
$1000*\ln K_{ZPE,aq,Cterm}$ (‰)	-2.2	-2.1

Table A5.4. Q Values for All Species Optimized in the Gas Phase or in Aqueous Solution at the MP2/aD or MP2/aD/COSMO Level at 298 K.

Species	# of H_2O Molecules	$Q_{l,gas\ opt}$	$Q_{l,aq\ opt}$	^{13}C Site	$Q_{h,gas\ opt}$ [a]	$Q_{h,aq\ opt}$ [a]
Alanine	7	2.42680×10^{-71}	3.67591×10^{-70}	Methyl	2.77475×10^{-71}	4.20473×10^{-70}
				C2	2.84597×10^{-71}	4.31398×10^{-70}
				C-terminus	2.90870×10^{-71}	4.39408×10^{-70}
Alanine, Zwitterion	7	1.33845×10^{-71}	1.64825×10^{-70}	Methyl	1.53195×10^{-71}	1.88699×10^{-70}
				C2	1.57008×10^{-71}	1.93402×10^{-70}
				C-terminus	1.60421×10^{-71}	1.96818×10^{-70}
Alanine, Deprotonated	7	6.01801×10^{-68}	1.69660×10^{-66}	Methyl	6.87162×10^{-68}	1.93640×10^{-66}
				C2	7.05616×10^{-68}	1.99065×10^{-66}
				C-terminus	7.21309×10^{-68}	2.02582×10^{-66}
Glutamic Acid	8	5.67114×10^{-87}	2.50526×10^{-85}	Side Chain	6.81020×10^{-87}	2.99488×10^{-85}
				C2	6.65295×10^{-87}	2.93860×10^{-85}
				C-terminus	6.79896×10^{-87}	2.99665×10^{-85}
Glutamic Acid, Zwitterion (C-terminus)	8	7.25137×10^{-87}	6.52629×10^{-86}	Side Chain	8.70155×10^{-87}	7.80597×10^{-86}
				C2	8.46005×10^{-87}	7.64664×10^{-86}
				C-terminus	8.67348×10^{-87}	7.78964×10^{-86}
Glutamic Acid, Zwitterion (Side Chain)	8	1.16950×10^{-86}	1.64453×10^{-85}	Side Chain	1.40115×10^{-86}	1.96048×10^{-85}
				C2	1.37083×10^{-86}	1.93101×10^{-85}
				C-terminus	1.40321×10^{-86}	1.96801×10^{-85}
Glutamate, Zwitterionic Anion		2.13975×10^{-83}	1.71107×10^{-82}	Side Chain	2.50363×10^{-83}	2.00674×10^{-82}
				C2	2.56265×10^{-83}	2.03921×10^{-82}
				C-terminus	2.55844×10^{-83}	2.04140×10^{-82}
Glutamate, C-terminus Deprotonation	8	1.92245×10^{-83}	5.48760×10^{-82}	Side Chain	2.30675×10^{-83}	6.56171×10^{-82}
				C2	2.25362×10^{-83}	6.43344×10^{-82}
				C-terminus	2.30331×10^{-83}	6.54926×10^{-82}
Glutamate, Side Chain Deprotonation	8	2.30255×10^{-83}	6.00869×10^{-82}	Side Chain	2.75483×10^{-83}	7.15671×10^{-82}
				C2	2.69723×10^{-83}	7.04667×10^{-82}
				C-terminus	2.75644×10^{-83}	7.17956×10^{-82}

Glutamate, Doubly Deprotonated	8	$5.15756x10^{-80}$	$1.01366x10^{-78}$	Side Chain	$6.16202x10^{-80}$	$1.20839x10^{-78}$
				C2	$6.02672x10^{-80}$	$1.18856x10^{-78}$
				C-terminus	$6.16349x10^{-80}$	$1.20920x10^{-78}$
Pyruvic Acid	6	$1.36071x10^{-55}$	$2.10712x10^{-54}$	Methyl	$1.55333x10^{-55}$	$2.40473x10^{-54}$
				C2	$1.61052x10^{-55}$	$2.49164x10^{-54}$
				Carboxylic Acid	$1.62413x10^{-55}$	$2.50838x10^{-54}$
Pyruvate	6	$2.47092x10^{-52}$	$3.25271x10^{-51}$	Methyl	$2.81494x10^{-52}$	$3.70508x10^{-51}$
				C2	$2.92504x10^{-52}$	$3.85147x10^{-51}$
				Carboxylate	$2.95373x10^{-52}$	$3.88193x10^{-51}$
Pyruvate	8	$1.93272x10^{-65}$	$7.01755x10^{-64}$	Methyl	$2.20140x10^{-65}$	$7.99402x10^{-64}$
				C2	$2.28601x10^{-65}$	$8.30790x10^{-64}$
				Carboxylate	$2.30922x10^{-65}$	$8.37321x10^{-64}$
α-Ketoglutaric Acid	9	$7.74812x10^{-84}$	$2.60672x10^{-82}$	$-C(O)^{13}COOH$	$9.23382x10^{-84}$	$3.10272x10^{-82}$
				$-^{13}C(O)COOH$	$9.14008x10^{-84}$	$3.07668x10^{-82}$
				$-CH_2{}^{13}COOH$	$9.28701x10^{-84}$	$3.11433x10^{-82}$
α-Ketoglutarate, α-Keto Moiety Dep. ($-C(O)C(O)O^-$)	9	$5.08207x10^{-81}$	$2.76222x10^{-79}$	$-C(O)^{13}COO^-$	$6.07198x10^{-81}$	$3.29494x10^{-79}$
				$-^{13}C(O)COO^-$	$6.01137x10^{-81}$	$3.26658x10^{-79}$
				$-CH_2{}^{13}COOH$	$6.08189x10^{-81}$	$3.30082x10^{-79}$
α-Ketoglutarate, Glutarate Moiety Dep. ($-CH_2C(O)O^-$)	9	$3.91551x10^{-80}$	$6.71130x10^{-79}$	$-C(O)^{13}COOH$	$4.67645x10^{-80}$	$8.00102x10^{-79}$
				$-^{13}C(O)COOH$	$4.62461x10^{-80}$	$7.93247x10^{-79}$
				$-CH_2{}^{13}COO^-$	$4.68027x10^{-80}$	$8.00338x10^{-79}$
α-Ketoglutarate, Doubly Deprotonated	9	$1.07394x10^{-77}$	$7.12808x10^{-76}$	$-C(O)^{13}COO^-$	$1.27869x10^{-77}$	$8.45961x10^{-76}$
				$-^{13}C(O)COO^-$	$1.26541x10^{-77}$	$8.39895x10^{-76}$
				$-CH_2{}^{13}COO^-$	$1.28365x10^{-77}$	$8.50191x10^{-76}$

[a] "l" represents values for which all atoms are isotopically light and "h" denotes the presence of ^{13}C at the given carbon.

Table A5.5. Reduced Partition Function Ratios at Various Carbon Atoms at 298 K for Glutamic Acid and α-Ketoglutaric Acid.

Species	# of H_2O Molecules	^{13}C Site	$\beta_{gas\ opt}$	$\beta_{aq\ opt}$	$\Delta\beta$ (‰) [a]
Glutamic Acid	8	Side Chain	1.2009	1.1954	5.416
		C2	1.1731	1.1730	0.151
		C-terminus	1.1989	1.1961	2.726
Glutamic Acid, Zwitterion (C-terminus)	8	Side Chain	1.2000	1.1961	3.906
		C2	1.1667	1.1717	-4.984
		C-terminus	1.1961	1.1936	2.538
Glutamic Acid, Zwitterion (Side Chain)	8	Side Chain	1.1981	1.1921	5.949
		C2	1.1721	1.1742	-2.060
		C-terminus	1.1998	1.1967	3.136
Glutamate, Zwitterion	8	Side Chain	1.1976	1.1918	5.862
		C2	1.1701	1.1727	-2.693
		C-terminus	1.1957	1.1931	2.619
Glutamate, C-terminus Deprotonation	8	Side Chain	1.1999	1.1957	4.166
		C2	1.1723	1.1724	-0.094
		C-terminus	1.1981	1.1935	4.643
Glutamate, Side Chain Deprotonation	8	Side Chain	1.1964	1.1911	5.367
		C2	1.1714	1.1727	-1.337
		C-terminus	1.1971	1.1949	2.265
Glutamate, Doubly Deprotonated	8	Side Chain	1.1948	1.1921	2.651
		C2	1.1685	1.1725	-4.022
		C-terminus	1.1950	1.1929	2.136
α-Ketoglutaric Acid	9	$-C(O)^{13}COOH$	1.1917	1.1903	1.473
		$-^{13}C(O)COOH$	1.1797	1.1803	-0.638
		$-CH_2{}^{13}COOH$	1.1986	1.1947	3.884
α-Ketoglutarate, α-Keto Moiety Dep. ($-C(O)C(O)O^-$)	9	$-C(O)^{13}COO^-$	1.1948	1.1929	1.924
		$-^{13}C(O)COO^-$	1.1829	1.1826	0.265
		$-CH_2{}^{13}COOH$	1.1967	1.1950	1.743
α-Ketoglutarate, Glutarate Moiety Dep. ($-CH_2C(O)O^-$)	9	$-C(O)^{13}COOH$	1.1943	1.1922	2.167
		$-^{13}C(O)COOH$	1.1811	1.1820	-0.857
		$-CH_2{}^{13}COO^-$	1.1953	1.1925	2.792
α-Ketoglutarate, Doubly Deprotonated	9	$-C(O)^{13}COO^-$	1.1907	1.1868	3.852
		$-^{13}C(O)COO^-$	1.1783	1.1783	-0.003
		$-CH_2{}^{13}COO^-$	1.1953	1.1927	2.532

[a] $\Delta\beta$ is calculated as $(\beta_{gas} - \beta_{aq})*1000$.

Table A5.6. Estimated Free Energy (A), ΔA, and ϵ_{ZPE} in kcal/mol for ^{12}C and ^{13}C Bonds with N, O, and H in kcal/mol at 310 K.

	A ^{12}C			A ^{13}C		
	ZPE	Thermal	Total	ZPE	Thermal	Total
C-N	1.57	-3.76×10^{-3}	1.57	1.54	-4.18×10^{-3}	1.54
C=O	2.50	-1.84×10^{-4}	2.50	2.45	-2.20×10^{-4}	2.45
C-H	4.29	-5.57×10^{-7}	4.29	4.28	-5.80×10^{-7}	4.28
	$^{12}C{=}O + {}^{13}CHNH_2 \leftrightarrows {}^{13}C{=}O + {}^{12}CHNH_2$					
	ΔA	K	ϵ_{ZPE} (‰)			
ZPE + Therm	-9.59×10^{-3}	-0.9845	15.6			
ZPE Only	-9.98×10^{-3}	-0.9839	16.2			

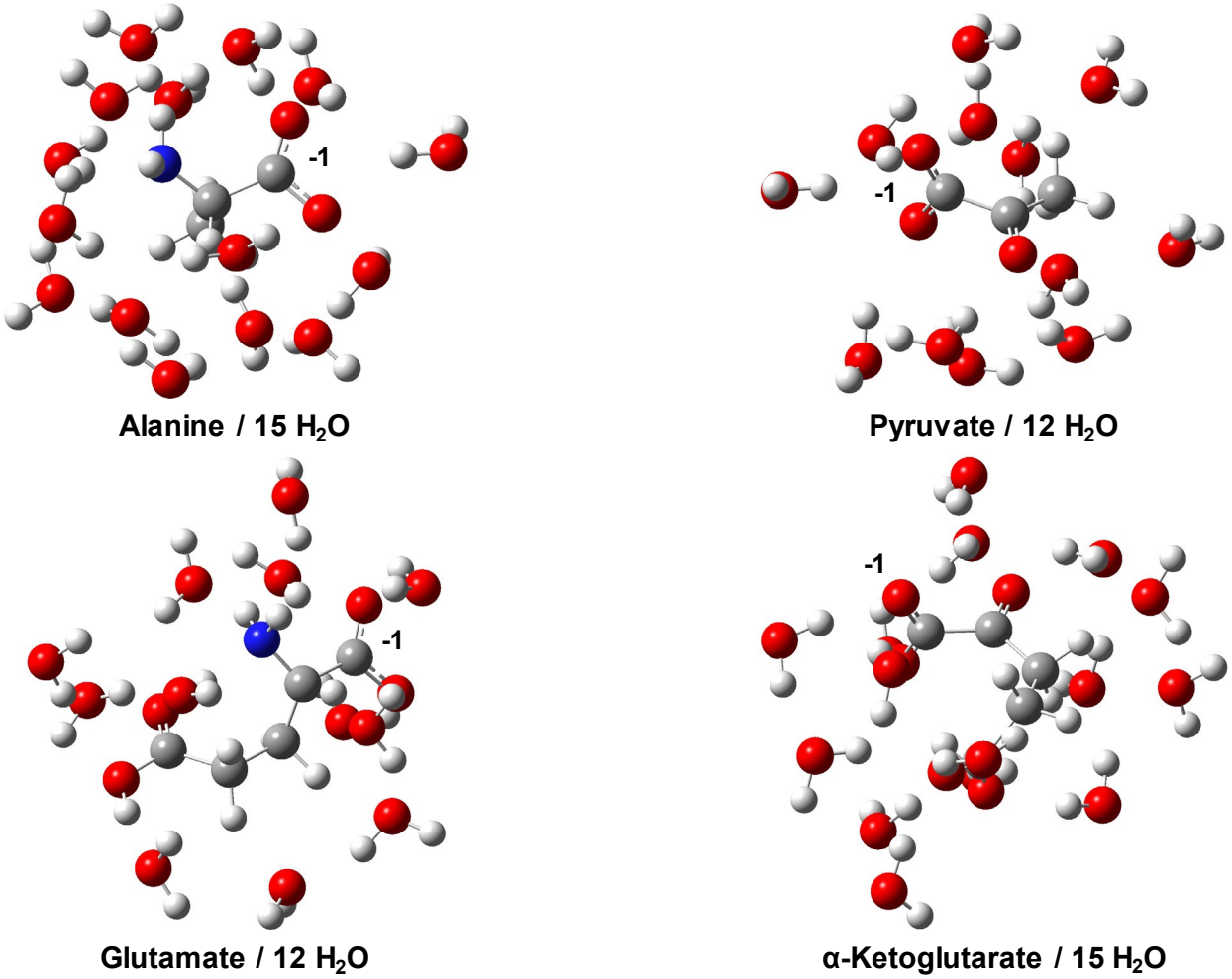


Figure A5.1. Starting explicitly solvated (single solvation shell) structures from MD simulations for deprotonated alanine, pyruvate, α-ketoglutarate, and glutamate.

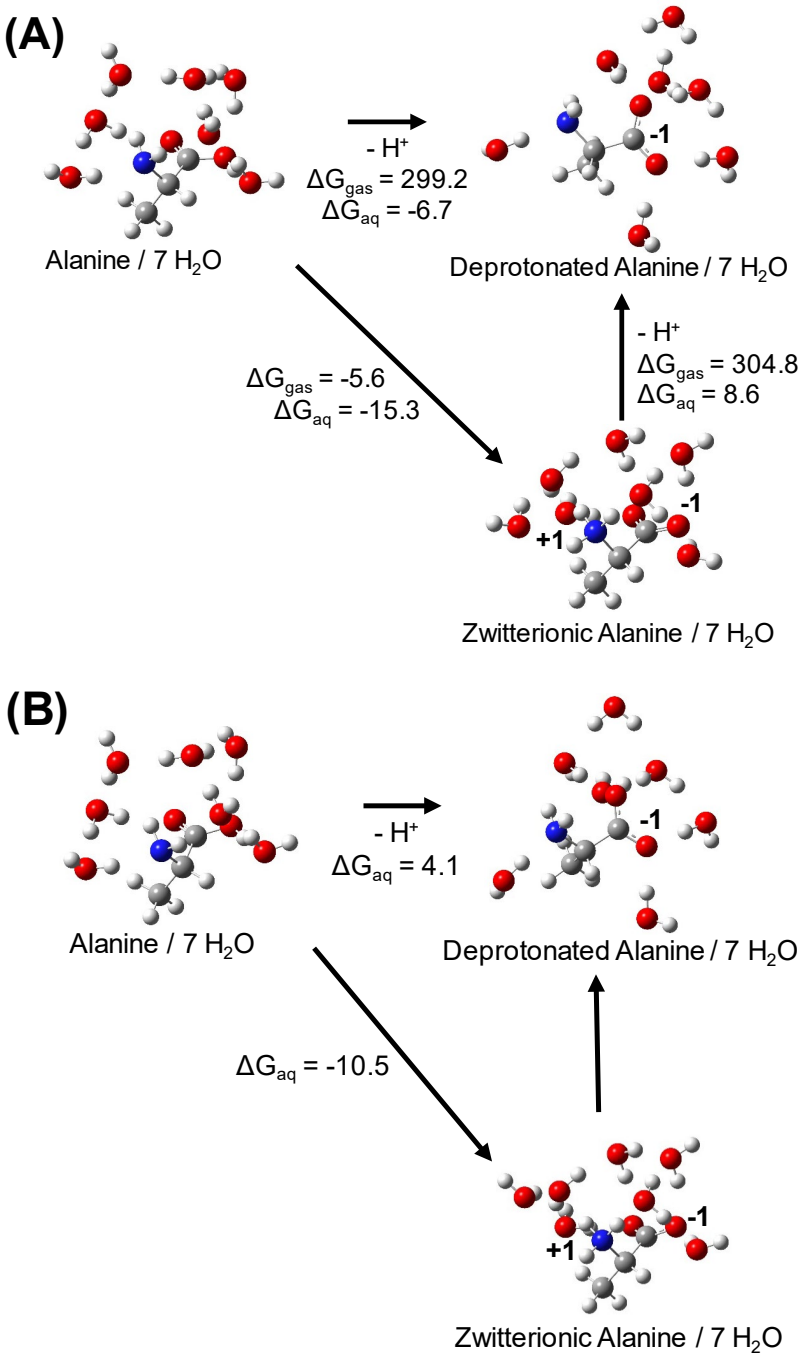
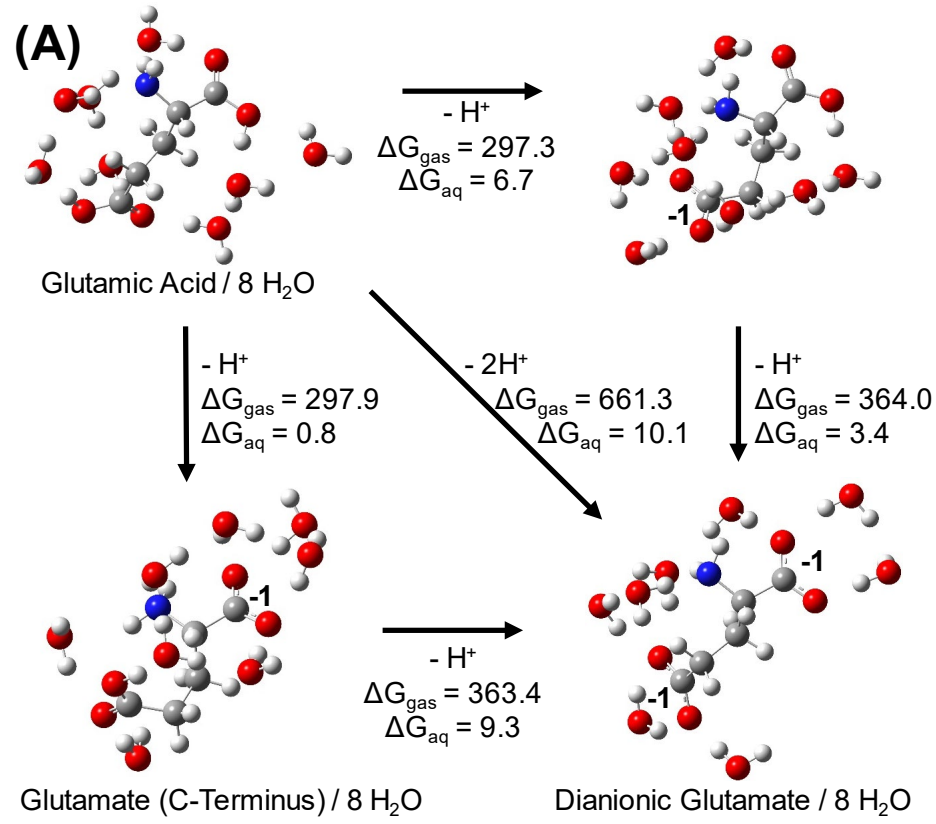
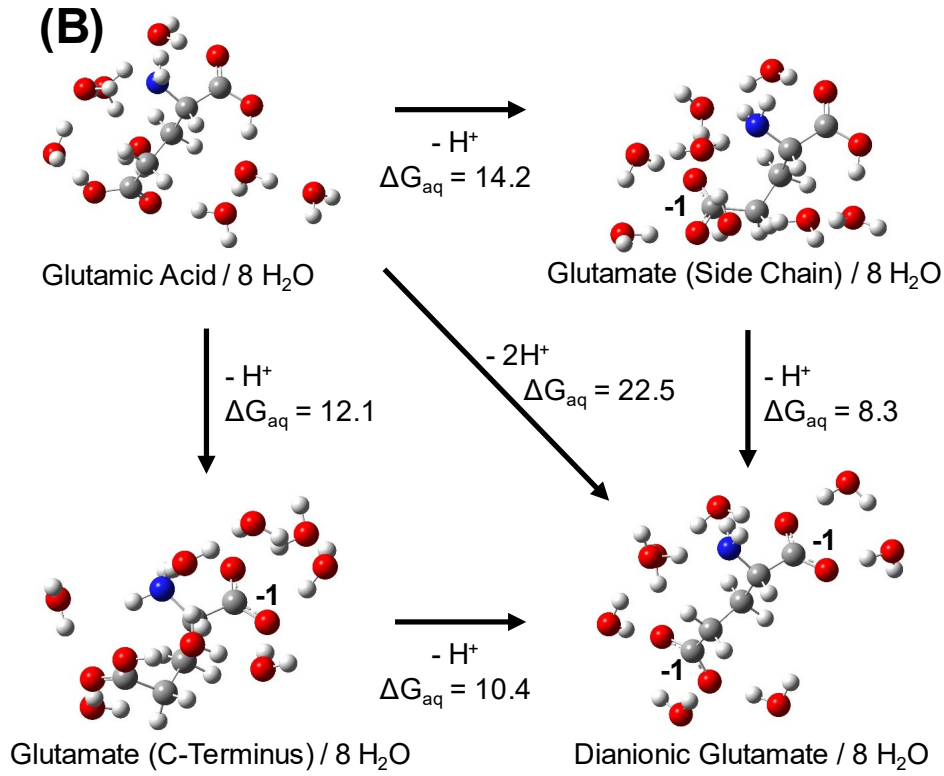
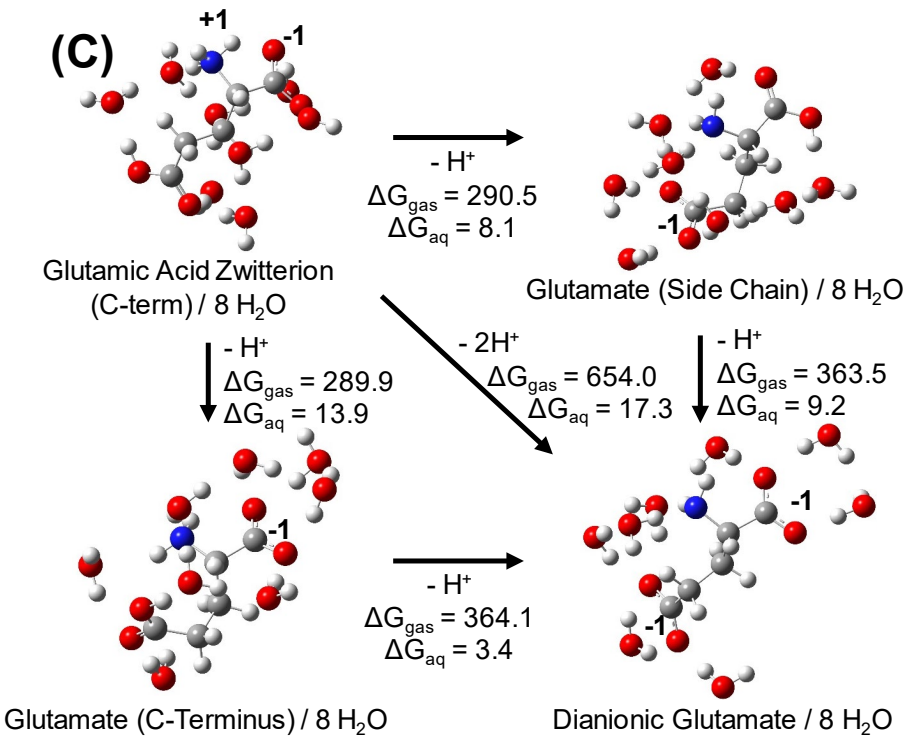
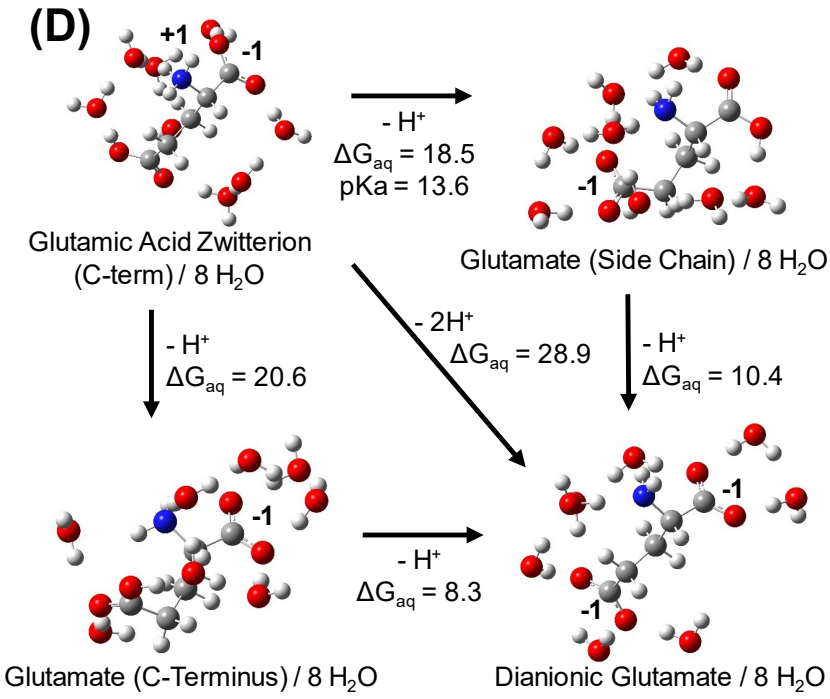


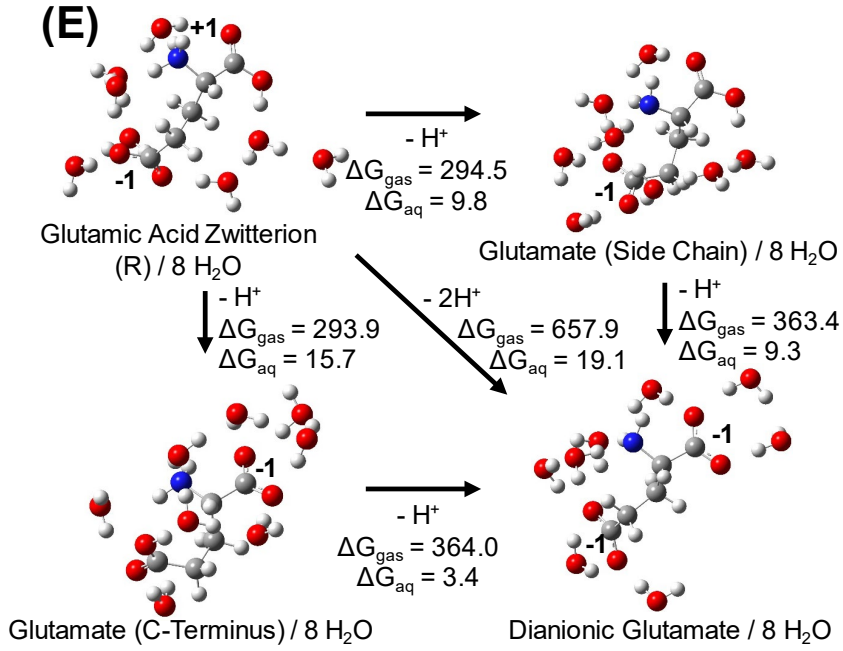
Figure A5.2. Alanine and related structures with 7 explicit H_2O molecules optimized (A) in the gas phase. Energies are given in kcal/mol at the MP2/aD level with and without SCRF via COSMO and (B) in aqueous solution in kcal/mol at the MP2/aD/COSMO level.


(A)
- H+
ΔGgas = 297.3
ΔGaq = 6.7
-1
Glutamic Acid / 8 H2O
- H+
ΔGgas = 297.9
ΔGaq = 0.8
- 2H+
ΔGgas = 661.3
ΔGaq = 10.1
- H+
ΔGgas = 364.0
ΔGaq = 3.4
-1
-1
-1
- H+
ΔGgas = 363.4
ΔGaq = 9.3
Glutamate (C-Terminus) / 8 H2O
Dianionic Glutamate / 8 H2O


(B)
- H+
ΔGaq = 14.2
-1
Glutamic Acid / 8 H2O
Glutamate (Side Chain) / 8 H2O
- H+
ΔGaq = 12.1
- 2H+
ΔGaq = 22.5
- H+
ΔGaq = 8.3
-1
-1
-1
- H+
ΔGaq = 10.4
Glutamate (C-Terminus) / 8 H2O
Dianionic Glutamate / 8 H2O


(C)
+1
-1
Glutamic Acid Zwitterion
(C-term) / 8 H2O
- H+
ΔGgas = 290.5
ΔGaq = 8.1
-1
Glutamate (Side Chain) / 8 H2O
- H+
ΔGgas = 289.9
ΔGaq = 13.9
- 2H+
ΔGgas = 654.0
ΔGaq = 17.3
- H+
ΔGgas = 363.5
ΔGaq = 9.2
-1
- H+
ΔGgas = 364.1
ΔGaq = 3.4
-1
-1
Glutamate (C-Terminus) / 8 H2O
Dianionic Glutamate / 8 H2O


(D)
+1
-1
Glutamic Acid Zwitterion
(C-term) / 8 H2O
- H+
ΔGaq = 18.5
pKa = 13.6
-1
Glutamate (Side Chain) / 8 H2O
- H+
ΔGaq = 20.6
- 2H+
ΔGaq = 28.9
- H+
ΔGaq = 10.4
-1
- H+
ΔGaq = 8.3
-1
-1
Glutamate (C-Terminus) / 8 H2O
Dianionic Glutamate / 8 H2O



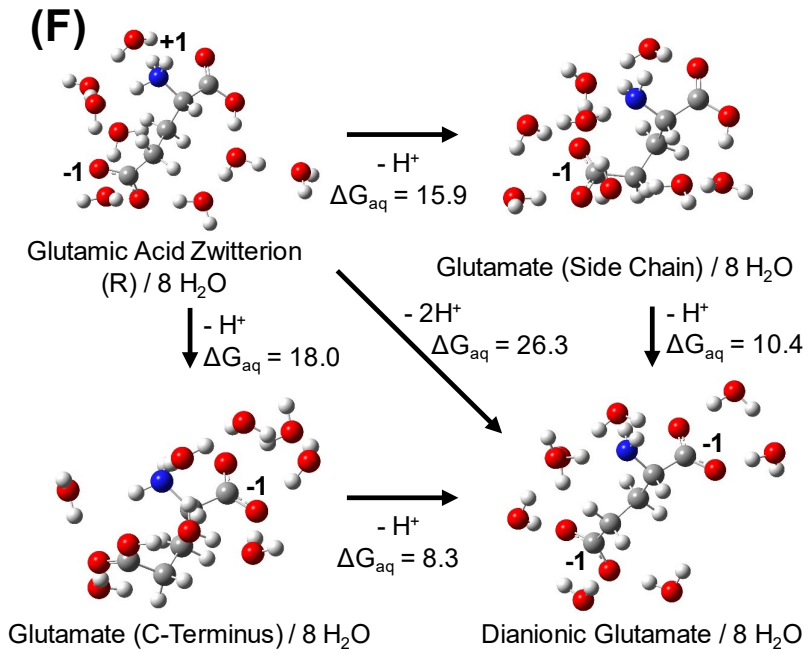


Figure A5.3. Glutamic acid and related structures with 8 explicit H_2O molecules optimized (A, C, E) in the gas phase and (B, D, F) in aqueous solution. Non-charge separated (NCS) glutamic acid deprotonation schemes are given in (A) and (B). Zwitterionic glutamic acid deprotonation reactions with the negative charge at the C-terminus are given in schemes (C) and (D). Zwitterionic glutamic acid deprotonation reactions with the negative charge at the side group are given in schemes (E) and (F). Energies are given in kcal/mol at the MP2/aD level with and without SCRF via COSMO and in aqueous solution in kcal/mol at the MP2/aD/COSMO level.

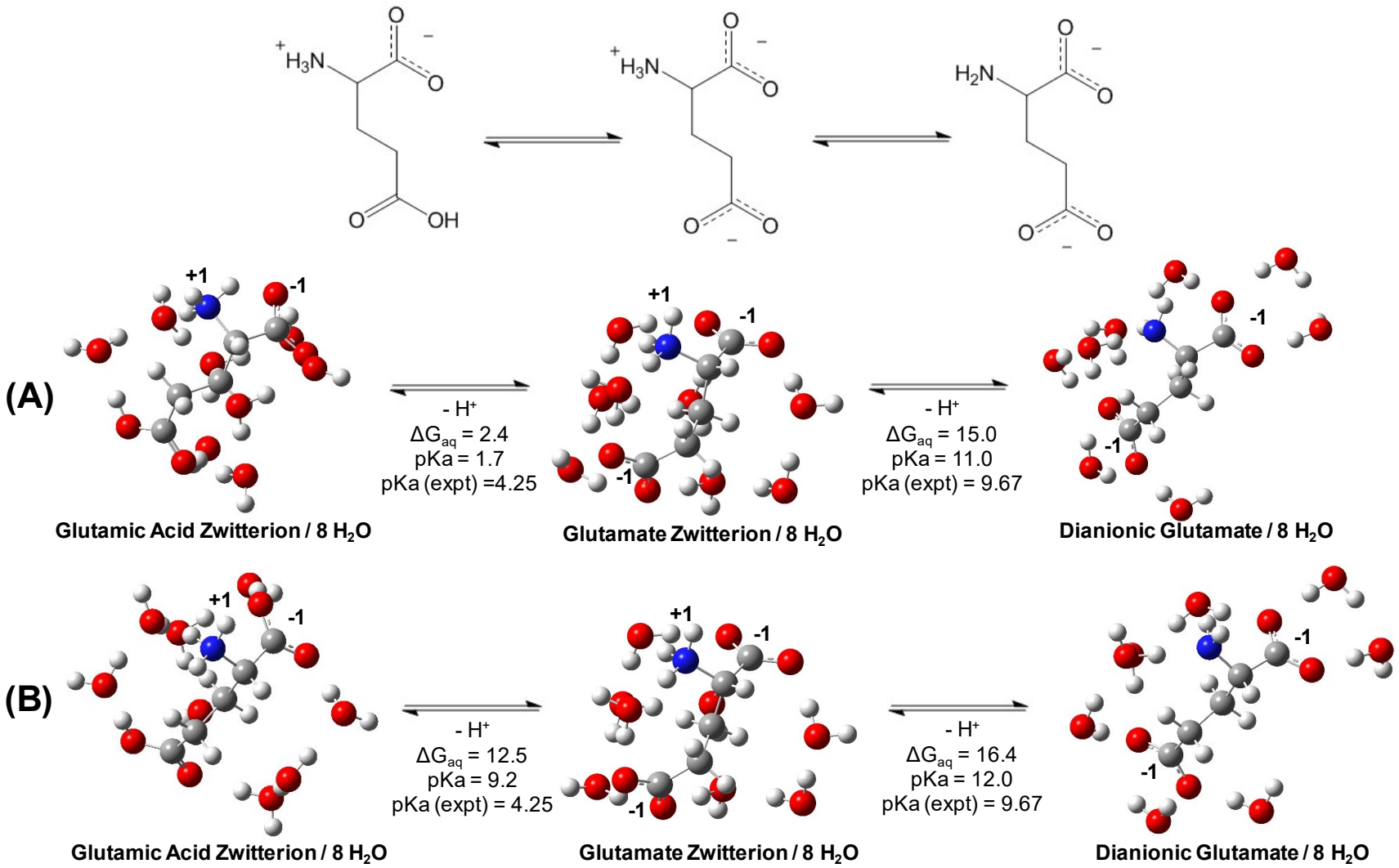


Figure A5.4. Zwitterionic glutamic acid deprotonation reactions and energetics compared to experimental pKa values for both (A) MP2/aD gas phase optimized and (B) MP2/aD/COSMO aqueous optimized clusters of glutamic acid or glutamate with 8 explicit H_2O molecules.

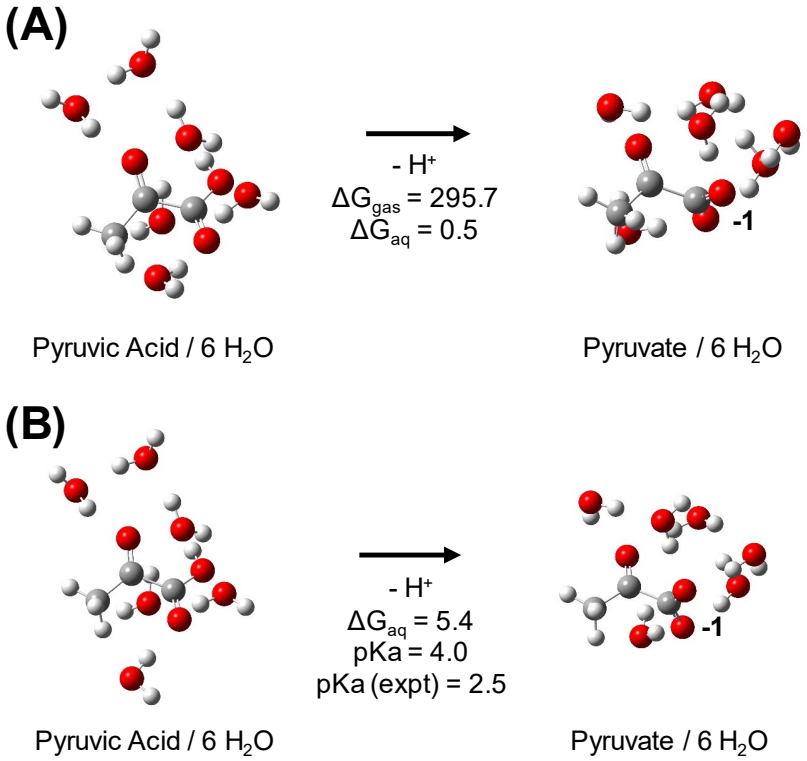


Figure A5.5. Pyruvic acid and pyruvate with 6 explicit H_2O molecules optimized (A) in the gas phase. Energies are given in kcal/mol at the MP2/aD level with and without SCRF via COSMO and (B) in aqueous solution in kcal/mol at the MP2/aD/COSMO level.

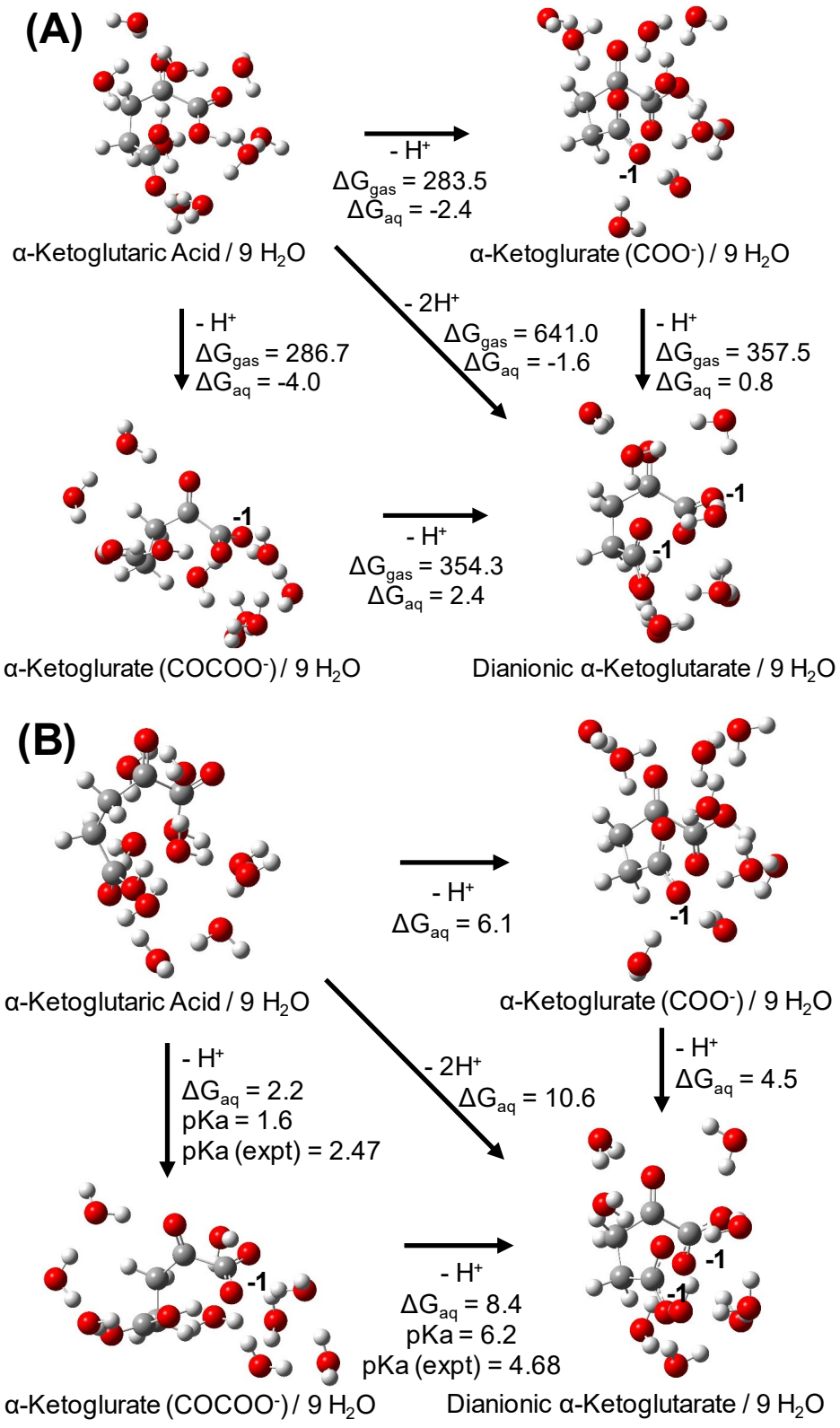
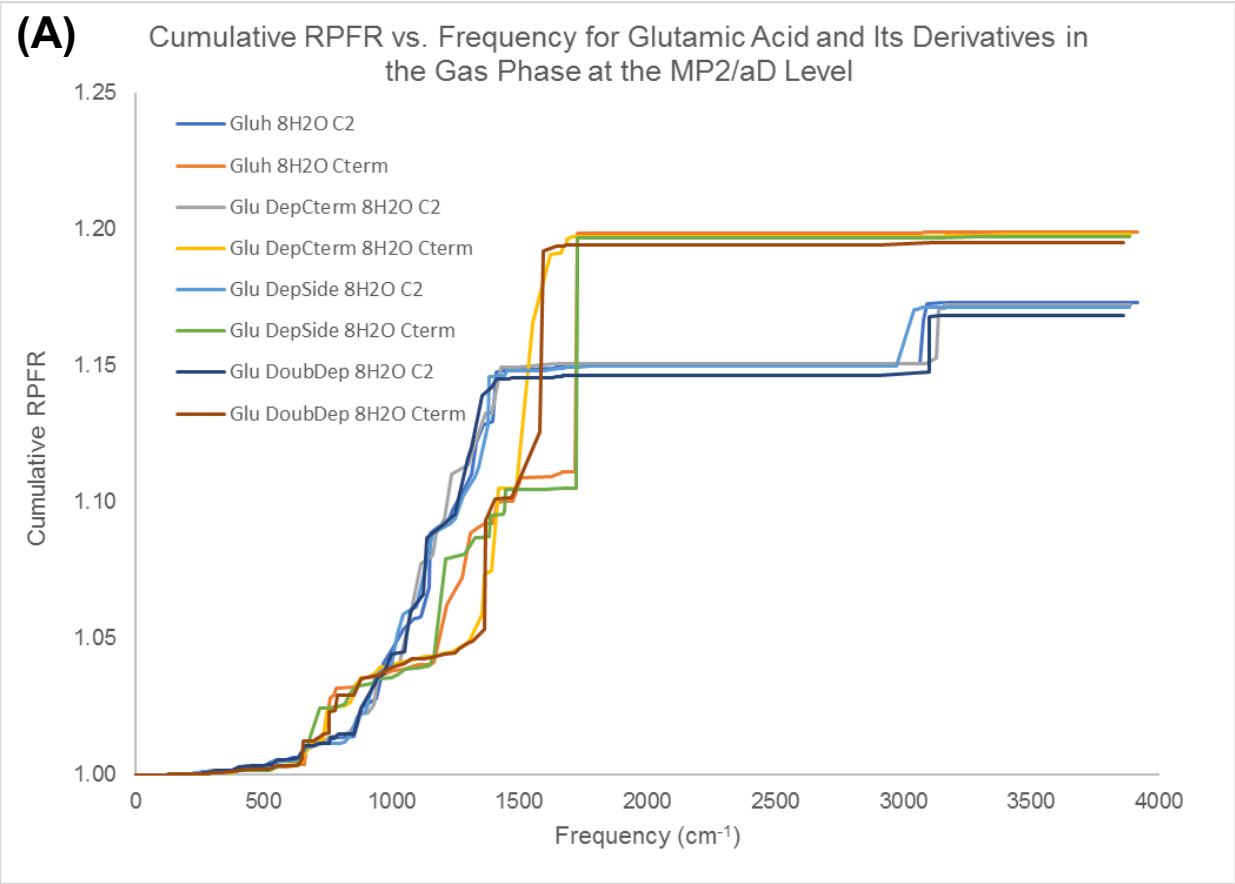
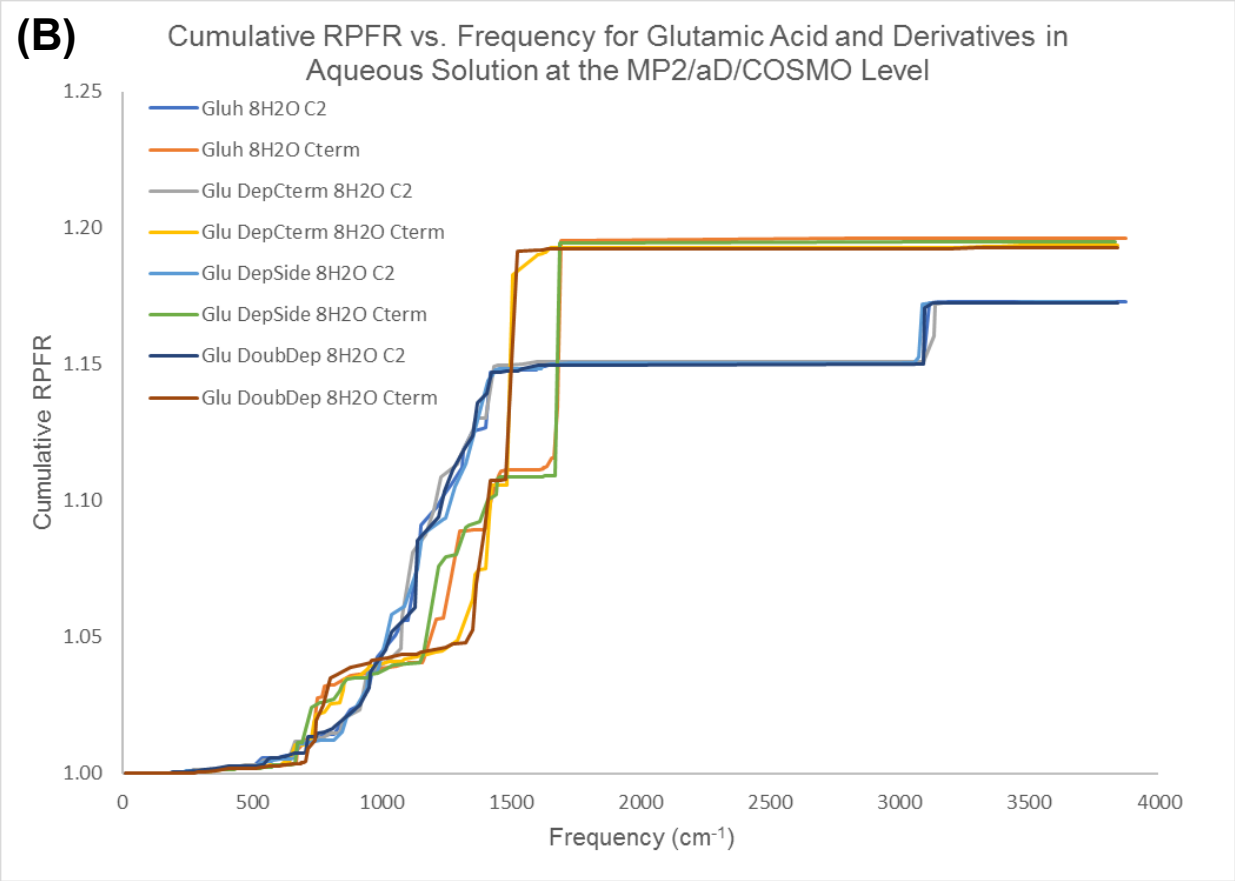
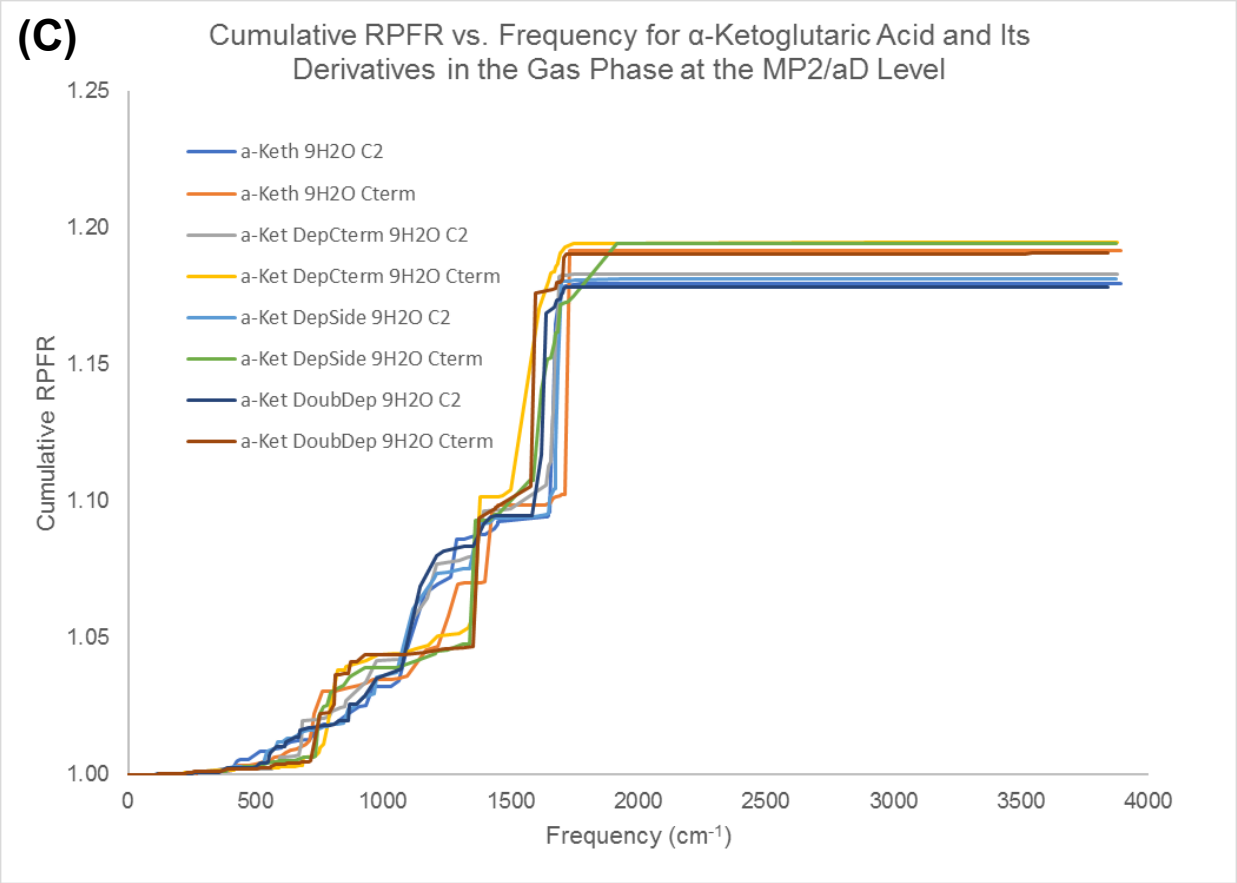


Figure A5.6. α-Ketoglutaric acid and related structures with 9 explicit H_2O molecules optimized (A) in the gas phase. Energies are given in kcal/mol at the MP2/aD level with and without SCRF via COSMO and (B) in aqueous solution in kcal/mol at the MP2/aD/COSMO level.


(A)
Cumulative RPFR vs. Frequency for Glutamic Acid and Its Derivatives in the Gas Phase at the MP2/aD Level
Cumulative RPFR
1.25
1.20
1.15
1.10
1.05
1.00
0
500
1000
1500
2000
2500
3000
3500
4000
Frequency (cm-1)
Gluh 8H2O C2
Gluh 8H2O Cterm
Glu DepCterm 8H2O C2
Glu DepCterm 8H2O Cterm
Glu DepSide 8H2O C2
Glu DepSide 8H2O Cterm
Glu DoubDep 8H2O C2
Glu DoubDep 8H2O Cterm


(B)
Cumulative RPFR vs. Frequency for Glutamic Acid and Derivatives in Aqueous Solution at the MP2/aD/COSMO Level
Gluh 8H2O C2
Gluh 8H2O Cterm
Glu DepCterm 8H2O C2
Glu DepCterm 8H2O Cterm
Glu DepSide 8H2O C2
Glu DepSide 8H2O Cterm
Glu DoubDep 8H2O C2
Glu DoubDep 8H2O Cterm
Cumulative RPFR
1.25
1.20
1.15
1.10
1.05
1.00
0
500
1000
1500
2000
2500
3000
3500
4000
Frequency (cm-1)


(C)
Cumulative RPFR vs. Frequency for α-Ketoglutaric Acid and Its Derivatives in the Gas Phase at the MP2/aD Level
Cumulative RPFR
1.25
1.20
1.15
1.10
1.05
1.00
0
500
1000
1500
2000
2500
3000
3500
4000
Frequency (cm-1)
a-Keth 9H2O C2
a-Keth 9H2O Cterm
a-Ket DepCterm 9H2O C2
a-Ket DepCterm 9H2O Cterm
a-Ket DepSide 9H2O C2
a-Ket DepSide 9H2O Cterm
a-Ket DoubDep 9H2O C2
a-Ket DoubDep 9H2O Cterm

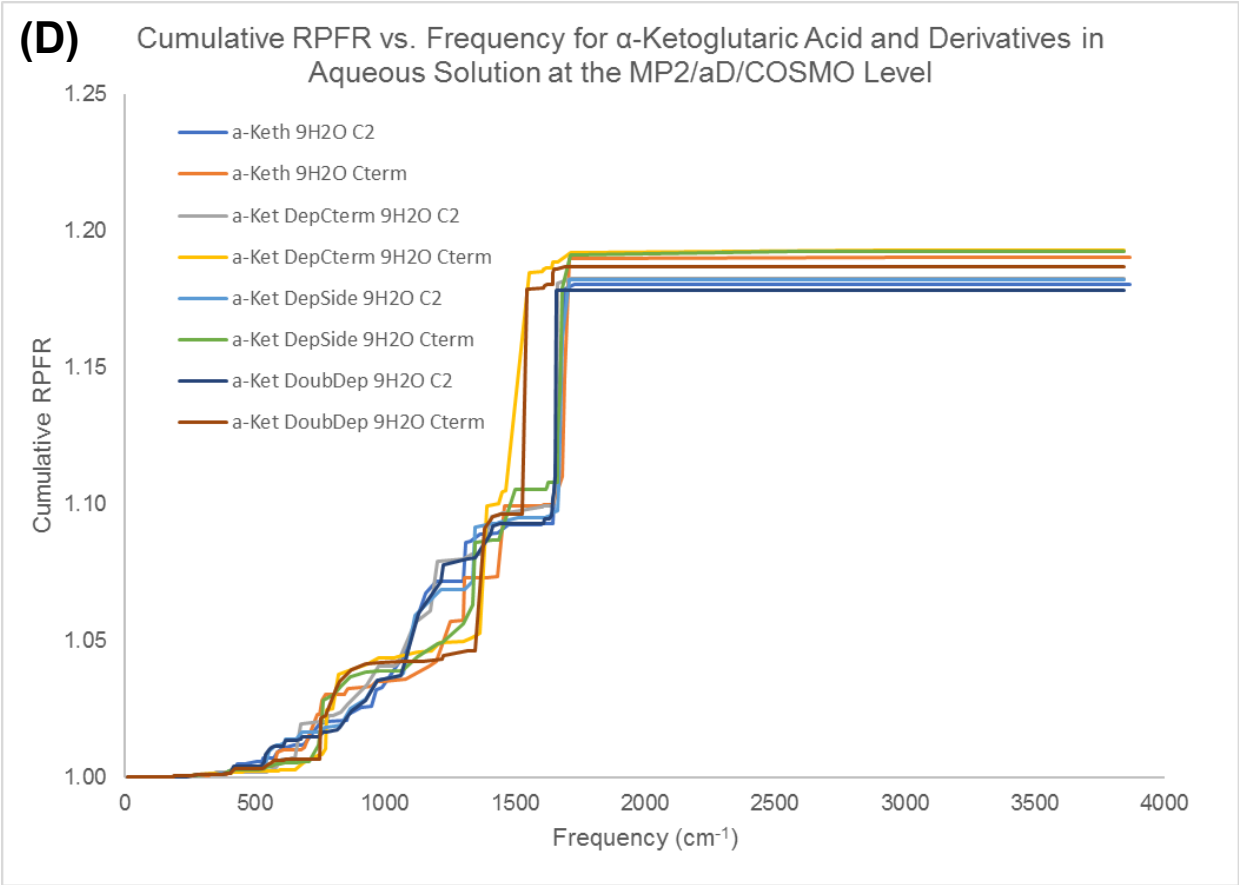


Figure A5.7. Cumulative reduced partition function ratio (RPFR) vs. vibrational frequency (cm^{-1}) at 298 K for ^{13}C at either the C2 site or the C-terminus. (A) and (B) are glutamic acid (Gluh) and its various deprotonated forms (Glu) with 8 explicit H_2O molecules in the gas phase and in aqueous solution, respectively. (C) and (D) are α-ketoglutaric (a-Keth) acid and its various deprotonated forms (a-Ket) with 9 explicit H_2O molecules in the gas phase and in aqueous solution, respectively. All results are optimized at the MP2/aD level for the gas phase and MP2/aD/COSMO for the aqueous solution.

CHAPTER 6

THE H•/H^- REDOX COUPLE AND ABSOLUTE HYDRATION ENERGY OF H^-

From McNeill, A. S.; Zhan, C.-G.; Appel, A. M.; Stanbury, D. M.; Dixon, D. A. Accepted to *J. Phys. Chem. A* **2020**.

Introduction

Hydration of ions plays a vital role in aqueous chemical and biological systems.[1,2,3,4,5,6] The solvation of the H^- (hydride) anion is important in developing an absolute hydricity scale for designing catalytic processes, especially those involving hydride transfer.[7,8,9,10,11,12,13,14,15,16,17,18,19,20] The gas phase fluoride and hydride affinity scales have been used to predict Lewis acidities and associated reactivity patterns,[21,22] following the gas phase proton affinity scale[23] that has been broadly developed and applied. It was possible using a range of techniques to get an absolute proton affinity scale, although much of the initial proton affinity scale was developed as sets of relative scales. The development of an absolute scale can be done in the gas phase for the addition of an atomic ion such as H^+, H^-, and F^- in a relatively straightforward fashion. However, it is far more complicated to get the absolute energy of the ion in solution. The hydricity scales used to date are in principle absolute hydricity scales, but these scales suffer from the limited accuracy for defining the free energy of formation of a solvated hydride. For the purpose of predicting reactivity between soluble species, such as in catalysis by transition metal hydride complexes, the relative values have been demonstrated to be very useful.[9]

It is important to note that the aqueous hydride anion has not yet been observed. In fact, the only report of the existence of H^- in a molecular solvent is from the solubility of KH in THF/crown ether, which was interpreted in terms of a solvated hydride anion.[24] The reaction that occurs between hydride and the solvent H_2O molecules, forming solvated OH^- (hydroxide) and molecular hydrogen (H_2) after proton transfer, has been computationally estimated to occur in 20 to 100 ps at 300 K.[25] The only reliable solution-phase H^- reports pertain to molten salts.[26] It has been suggested that H^- is produced as an intermediate in the photolysis of $Ta_6Br_{12}^{2-}$.[27]

If the free energies of solvation for H^- were available, then it would be possible to develop an absolute hydricity scale. One test of these values is $E°(H•/H^-_{aq})$ and the available results are summarized in Table 6.1. The free energy of solvation of H_2 can be calculated from the Henry's Law constant[28] K_{H2} = 0.00078 M/atm, giving ΔG_{solv} = 4.25 kcal/mol and this value is often used as an estimate for the free energy of solvation of H•. There have been several attempts to estimate the $H•/H^-$ standard reduction potential, many of which are dependent on the choice of H^- radius, due to the lack of bare hydride in solution. Latimer estimated $E°(1/2H_2/H^-_{aq})$ to be -2.25 V based on using an H^- radius of 1.50 Å derived from metal hydride crystals.[29] Eberson derived $E°(H•/H^-_{aq})$ = –0.14 V from Latimer's result by adjusting for the H-H bond energy in H_2.[30] Toffel and Henglein reported a reduction potential value of 0.05 V using an estimated hydration energy.[31] Using an ionic radius for hydride of 1.56 Å, Ross derived a reduction potential value for H_2/H^- of -2.25 V,[32] based on Latimer's method, but with updated data. Bratsch estimated $E°(1/2H_2/H^-_{aq})$ = –2.40 V by assuming that the dissolutions of alkali hydrides fall between those of alkali fluorides and alkali chlorides,[33] following the assumptions and estimates of Latimer[29] and Ross.[32] Pearson obtained $E°(H•/H^-_{aq})$ = 0.18 ± 0.15 V based on the known gas phase electron affinity of H• and estimated values for the free energies of

hydration for H• (assumed to be the same as H_2 at 4.0 kcal/mol) and H^-.[34] The estimated value of the hydration energy for H^- reported by Pearson is -87 ± 2 kcal/mol and is derived by subtracting 4 pK_a units from the pK_a value of H_2 in tetrahydrofuran from the work of Buncel and Menon.[24] Griller et al.[35] estimated potentials for the H^+/H• and H•/H^- couples using gas-phase data for H•, Toffel and Henglein's data for H^-,[31] and assuming $\Delta G_{hyd} = 0$ for H•, obtaining $E° = -0.36$ V. Parker[36] "corrected" the work of Griller et al.[35] on the basis that they neglected ΔG_{hyd} for H• giving a value of $E° = -0.12$ V. However, Parker's estimate (following the work of Pearson[34]) is based on a correlation of ΔG_{hyd} vs radius for the noble gases, and the data used in this correlation disagree substantially with the National Bureau of Standards (NBS) data. Subsequently, Parker published a correction[37] in which he noted that the data used in the first paper were referenced to a standard state of unit mole fraction. When corrected to a standard state of 1 atm, all of the data change. In particular, the potential for H^+/H• is now -2.29 V, which is in reasonable agreement with the value selected by Stanbury.[38] In addition, Parker notes that an even better estimate for the solvation energy of H• is available from the work of Roduner and Bartels.[39] These authors argue that solvation enthalpy is primarily determined by polarizability, and therefore H_2 is much better than He as a model for H•. Note that in Parker's correction,[37] he failed to correct his result for the H•/H^- couple. However, Parker cites a different value, 0.18 V, in a review.[40] This review cites another paper for this value,[41] and in this latter paper, Parker cites Pearson's result.[34] Kelly and Rosseinsky[42] estimated the ionic radius of H^- to be 1.38 Å by correlations with halides and thereby calculated $E° = 0.83$ V for H•/H^-, which would make H• a strong outer-sphere oxidant. Following this approach, Connelly et al.[43] used updated values to obtain a value of $E° = 0.81$ V. The resulting free energy of solvation of H^- is -96.1 kcal/mol. These two latter approaches

depend on the ionic radius of H^- and being able to extrapolate from the free energies of solvation of the halide anions based on their ionic radii.

A variety of computational approaches have been used to predict the structures and energies of hydride complexes with up to 8 H_2O.[44,45,46,47] Binding energies, electronic properties, and some spectroscopic properties have been predicted at levels of theory up to CCSD(T)/CBS.

Our goal is to predict the free energy of solvation of H• and H^- using a hybrid supermolecule-continuum approach. The cluster-continuum approach employed in this study has been extensively developed and used by the community for ~ 25 years in the successful prediction of a broad range of thermodynamic properties in aqueous and other solutions.[48,49,50,51,52,53,54,55,56,57,58,59,60,61,62] Our group has previously used this approach to predict the free energies of hydration at 298 K for H^+,[63] OH^-,[64] and F^-[65] and the results are within ± 1 kcal/mol of experiment. In addition, it has been used to predict the reduction potential ($E°$) for the CO_2/CO_2^- couple within 0.1 V of experiment.[66] We have used this approach to predict the free energies of solvation of H• and H^-, and these values are then used to estimate $E°$ for the H•/H^- couple. As a result, a more accurate anchoring point for the absolute hydricity scale in water is provided.

Computational Methods

A hybrid supermolecule-continuum approach was used to estimate the absolute hydration free energy of H• and H^-. The hydration free energy of species X with X = H• or H^- in water is the free energy of reaction (6.1), $\Delta G_{hyd}[X,n]$, converged to $n \rightarrow \infty$.

$$X(gas) + (H_2O)_n(aq) \rightarrow X(H_2O)_n(aq) \qquad (6.1)$$

To calculate this value, we must first determine the Gibbs free energy of the components of this reaction. Gas phase free energies are calculated directly and the aqueous free energies of

$(H_2O)_n$(aq) and $X(H_2O)_n$(aq) are calculated as the sum of their respective gas phase free energy and the bulk solvent shift according to equations 6.2 and 6.3.

$$G[(H_2O)_n(aq)] = G[(H_2O)_n(gas)] + \Delta G_{sol}[(H_2O)_n] \quad (6.2)$$

$$G[X(H_2O)_n(aq)] = G[X(H_2O)_n(gas)] + \Delta G_{sol}[X(H_2O)_n] \quad (6.3)$$

The bulk solvent shift from the gas phase cluster to the solution, $\Delta\Delta G_{sol}[X,n]$, and the hydration free energy of X in gas phase cluster $X(H_2O)_n$ are calculated according to equations 6.4 and 6.5, respectively.

$$\Delta\Delta G_{sol}[X,n] = \Delta G_{sol}[X(H_2O)_n] - \Delta G_{sol}[(H_2O)_n] \quad (6.4)$$

$$\Delta G_{gas}[X,n] = G[X(H_2O)_n(gas)] - G[X(gas)] - G[(H_2O)_n(gas)] \quad (6.5)$$

Then, we can evaluate the hydration free energy of X using equation (6.6).

$$\Delta G_{hyd}[X,n] = \Delta G_{gas}[X,n] + \Delta\Delta G_{sol}[X,n] \quad (6.6)$$

At 298 K, $\Delta G_{hyd}[X,n]$ is converged to $\Delta G_{hyd}^{298}(X)$ when $n \rightarrow \infty$. To determine $\Delta G_{hyd}[X,n]$, both $\Delta G_{gas}[X,n]$ and $\Delta\Delta G_{sol}[X,n]$ should be calculated at as high a level of theory as possible.

All geometry optimizations and frequency calculations were performed using the Gaussian 16[67] (G16) software package. Geometries were optimized at the MP2/aug-cc-pVDZ (aD, $n = 2$) level, [68,69,70,71] which includes molecular clusters $H•(H_2O)_n$ for $n = 1 – 4$ and 8, $H^-(H_2O)_n$ for $n = 1$, 4, 8, 12, and 16, and $(H_2O)_n$ for $n = 1 – 4$, 8, 12, and 16. All of these clusters with $n \leq 8$ were also optimized at the MP2/aug-cc-pVTZ (aT, $n = 3$) level. Clusters with $n = 12$ or 16 were not optimized at the MP2/aT level due to the computational expense. Single point energies were calculated at the MP2/aug-cc-pVQZ (aQ, $n = 4$) level for all species with $n \leq 4$, to extrapolate the electronic energy to the complete basis set (CBS) limit. This was calculated[72,73] using equation (6.7)

$$E_{l_{max}} = E_{CBS} + Ae^{-(l_{max}-1)} + Be^{-(l_{max}-1)^2} \quad (6.7)$$

where $l_{max} = n$ for D, T, and Q. The electronic energies were combined with free energy corrections at the MP2/aD level to determine the gas phase free energy.

MOLPRO 2015.1[74,75] was used to calculate coupled cluster R/UCCSD(T)[76,77,78,79,80,81,82] and R/UCCSD(T)-F12b[83,84,85] single point electronic energies using the aD and aT basis sets and the appropriate F12b equivalents[86,87] using the highest order optimized geometry available for clusters with 8 or less explicit H_2O molecules. The CCSD(T)-F12 approach uses functions for which the inter-electronic coordinates are explicitly incorporated into the wavefunction, which enables improved performance with a smaller basis set. Clusters with or without hydride and 12 or 16 H_2O molecules, all of which are closed shell, were calculated at the RCCSD(T)/aD and aD-F12b levels. CCSD(T)/aD-F12b and CCSD(T)/aT-F12b values were extrapolated to CBS(DT) using equation (6.8).

$$E(l_{max}) = E_{CBS} + B/(l_{max} + ½)^4 \quad (6.8)$$

Single point solvation calculations with optimized geometries were performed at the MP2/aD level using the self-consistent reaction field (SCRF) approach[88] with the conductor-like screening model (COSMO),[89,90] SMD,[91] and polarizable continuum model (PCM) using integral equation formalism variant[92,93,94] parameters as implemented in Gaussian 16.[67] SCRF calculations at the MP2/aD/COSMO level were also performed with Gaussian 03[95] (G03) for comparison with G16 values. Additionally, single point solvation calculations for H^- and H• were calculated at the CCSD[79] and QCISD[96] levels with the aD basis set to compare with the results of the general hybrid supermolecule-continuum approach. Natural atomic charges were determined for optimized clusters $X(H_2O)_n$ where X = H^- and F^-[65] and n = 4 and 8 by calculating single point natural bond orbital (NBO) analyses with Gaussian16 using the density functional theory (DFT) exchange-correlation functional B3LYP[97,98] with the aD basis set.

Results

Optimized geometries. There are a number of geometries with low-lying energies. All optimized $H^-(H_2O)_n$ for n = 4, 8, 12, 16 clusters and their relative $\Delta H_{g,298K}$, $\Delta G_{g,298K}$, and $\Delta G_{aq,298K}$ are shown in Figure A6.1; the lowest free energy structures are shown in Figure 6.1. The hydride in $H^-(H_2O)$ is bound to a single proton from the water molecule. For $H^-(H_2O)_4$ there are four optimized geometries, three of which (B, C, and D) have the hydride sitting on the surface of the water cluster. For $H^-(H_2O)_4$ A (C_2 symmetry), the hydride is positioned in the center of the cluster of H_2O molecules with each water arranged in a tetrahedral formation around the anionic center, and this is the most favorable geometry at the CCSD(T)/CBS(DT)-F12b/COSMO(G16) level. $H^-(H_2O)_4$ B (C_1 symmetry) is the only cluster in which only 3 of the 4 explicit H_2O molecules are interacting with the hydride. Clusters $H^-(H_2O)_4$ B and $H^-(H_2O)_4$ D (C_1 symmetry) are 0.3 and 0.6 kcal/mol higher in free energy than $H^-(H_2O)_4$ B in aqueous solution at the CCSD(T)/CBS(DT)-F12b/COSMO(G16) level, which are not significant energy differences at this level of calculation, whereas $H^-(H_2O)_4$ C (C_4 symmetry) is 3.5 kcal/mol higher in energy at the CCSD(T)/CBS(DT)-F12b/COSMO(G16) level. However, the most favorable structure in the gas phase is $H^-(H_2O)_4$ D, which is geometrically similar to $H^-(H_2O)_4$ B, but has all 4 H_2O molecules interacting with the hydride. $H^-(H_2O)_4$ A and $H^-(H_2O)_4$ B are slightly higher in gas phase free energy compared to $H^-(H_2O)_4$ D at 1.1 kcal/mol each with $H^-(H_2O)_4$ C highest in free energy relative to $H^-(H_2O)_4$ D by 2.7 kcal/mol at the CCSD(T)/CBS(DT)-F12b level. Because $H^-(H_2O)_4$ A is very similar energetically to the lowest energy structures in the gas and aqueous phases, it is clear that the rearrangement of the solvated hydride to form a tetrahedral arrangement of H_2O molecules around the anion is not a process that involves any significant energy.

Both $H^-(H_2O)_8$ A and $H^-(H_2O)_8$ B are very similar geometrically, although slight changes in the orientation of the cube of explicit H_2O molecules result in overall symmetries of C_2 and C_4, respectively. The hydride sits in the center of one face of the $(H_2O)_8$ cube for both geometries. Both structures are within 1 kcal/mol of one another in both the gas phase and in aqueous solution. Structure $H^-(H_2O)_8$ C (C_1 symmetry) is disordered with the hydride embedded in the surface of the H_2O molecules. This geometry is lowest in aqueous free energy by 0.6 kcal/mol at the CCSD(T)/CBS(DT)-F12b/COSMO(G16) level compared to $H^-(H_2O)_8$ A and B, respectively. $H^-(H_2O)_8$ A is favored by 0.8 and 1.1 kcal/mol in the gas phase at the CCSD(T)/CBS(DT)-F12b level compared to $H^-(H_2O)_8$ B and C. Our independently chosen structures of $H^-(H_2O)_4$ clusters are consistent with those previously reported by Prakash et al.[47] Our structures for $H^-(H_2O)_8$ clusters are completely different from those of Prakash et al.[47] The $H^-(H_2O)_8$ conformers of Prakash et al.[47] are 1.0 to 2.4 kcal/mol higher in energy in the gas phase than our best geometry at the CCSD(T)/aD-F12b level and are 4.5 to 7.2 kcal/mol higher in energy in aqueous solution than our best geometry at the CCSD(T)/aD-F12b/COSMO(G16) level and are not discussed further.

There are four $H^-(H_2O)_{12}$ optimized structures, three of which (A, B, and C) have C_4 symmetry and are very similar energetically in both the gas and aqueous phases. $H^-(H_2O)_{12}$ A has a $(H_2O)_8$ cube core with a H_2O on each vertex. The hydride sits on one face of this cube, and the additional 4 H_2O molecules also exist in the plane of this face. $H^-(H_2O)_{12}$ B and $H^-(H_2O)_{12}$ C are nearly identical in geometry. Each structure consists of three stacked $(H_2O)_4$ rings with the hydride just above the top-most ring. $H^-(H_2O)_{12}$ B is the lowest energy isomer in the gas phase and in aqueous solution. $H^-(H_2O)_{12}$ A is energetically equivalent to $H^-(H_2O)_{12}$ B in the gas phase but is 2.7 kcal/mol higher in aqueous free energy than $H^-(H_2O)_{12}$ B at the MP2/aD and

MP2/aD/COSMO levels, respectively. $H^-(H_2O)_{12}$ C is 0.5 and 0.6 kcal/mol higher in gas phase and aqueous free energy, respectively, than $H^-(H_2O)_{12}$ B at the MP2/aD and MP2/aD/COSMO levels of theory. Relative values for $H^-(H_2O)_{12}$ A and $H^-(H_2O)_{12}$ C are not calculated at the CCSD(T)/aD-F12b level of theory due to their extreme similarity in both geometry and energy to $H^-(H_2O)_{12}$ B. Structure $H^-(H_2O)_{12}$ D is very different geometrically from A, B, and C in that the hydride is embedded in the center of a more disordered group of 12 H_2O molecules (C1 symmetry) and is 5.7 and 9.0 kcal/mol higher in free energy in the gas phase and in aqueous solution at the MP2/aD and MP2/aD/COSMO levels and 5.1 and 3.8 kcal/mol higher in free energy in the gas phase and in aqueous solution at the CCSD(T)/aD-F12b and CCSD(T)/aD-F12b/COSMO(G16) levels of theory than the lowest energy isomer, $H^-(H_2O)_{12}$ B.

The MP2/aD optimized geometries for $H^-(H_2O)_{16}$ vary significantly in structure. $H^-(H_2O)_{16}$ A is a more spherically distributed structure with C_2 symmetry overall with the hydride at the center of the cluster surrounded by explicit H_2O molecules. $H^-(H_2O)_{16}$ B is a layered structure with four $(H_2O)_4$ rings topped by the hydride anion on one of the cluster's smaller faces with overall C_4 symmetry. In the gas phase, $H^-(H_2O)_{16}$ B is preferred by 2.2 kcal/mol at the CCSD(T)/aD-F12b level, whereas $H^-(H_2O)_{16}$ A is preferred in aqueous solution by 1.4 kcal/mol at the CCSD(T)/aD-F12b/COSMO(G16) level. The hydride structures for n = 8, 12, and 16 are similar to those for F^- interacting with the same water clusters.[65]

Structures for the MP2/aT optimized $H•(H_2O)_n$ with n = 1 – 4 and the MP2/aD optimized n = 8 are shown in Figure 6.2. Both $H•(H_2O)$ and $H•(H_2O)_2$ have two optimized geometries that are very similar to one another with both gas phase and aqueous free energies within 0.1 kcal/mol of one another. Both $H•(H_2O)$ A and $H•(H_2O)_2$ A involve a hydrogen from a single explicit H_2O molecule being closest to the radical hydrogen atom. $H•(H_2O)$ B and $H•(H_2O)_2$ B,

on the other hand, have the radical hydrogen atom closer to both a hydrogen and an oxygen from all available explicit H_2O molecules in the cluster, virtually isoenergetic with structure A for 1 H_2O and is more favorable than structure A for 2 H_2O by 0.4 and 0.2 kcal /mol in the gas phase and in aqueous solution at the CCSD(T)/CBS(DT)-F12b and CCSD(T)/CBS(DT)-F12b/COSMO(G16) levels, respectively.

Similarly, the lowest energy structures for $H•(H_2O)_3$ at the CCSD(T)/CBS(DT)-F12b and CCSD(T)/CBS(DT)-F12b/COSMO(G16) levels are $H•(H_2O)_3$ B and $H•(H_2O)_3$ C, which are virtually equivalent in energy in both the gas phase and in aqueous solution and are very similar in geometry. All $H•(H_2O)_3$ involve a trio of H_2O molecules hydrogen bonded to one another in a ring-like formation, but $H•(H_2O)_3$ B and $H•(H_2O)_3$ C have one hydrogen from a H_2O molecule interacting with the radical hydrogen. $H•(H_2O)_3$ A has the radical hydrogen centered over the $(H_2O)_3$ ring, which results in a free energy that is 1.0 and 1.1 kcal/mol higher in energy at the CCSD(T)/CBS(DT)-F12b and CCSD(T)/CBS(DT)-F12b/COSMO(G16) levels than that of $H•(H_2O)_3$ B or $H•(H_2O)_3$ C. This trend continues in $H•(H_2O)_4$ clusters, as the hydrogen radical in $H•(H_2O)_4$ A is centered on the face of the $(H_2O)_4$ ring, which is virtually isoenergetic in both gas phase and aqueous free energy at the CCSD(T)/CBS(DT)-F12b and CCSD(T)/CBS(DT)-F12b/COSMO(G16) levels with $H•(H_2O)_4$ B, in which the hydrogen radical only interacts with a single hydrogen from an explicit H_2O in the $(H_2O)_4$ ring. The optimized geometry of $H•(H_2O)_8$ is highly similar to that of $H•(H_2O)_4$ A with an additional tetramer ring of H_2O molecules beneath the first.

Important low free energy optimized structures for $(H_2O)_n$ with $n = 2 – 4$, 8, 12, and 16 are given in Figure A6.3. The optimized structures are very similar to the structures discussed above with the H• or H^- removed from the cluster. For $(H_2O)_2$, $(H_2O)_3$, and $(H_2O)_8$ there is only

one low free energy structure for the water cluster, all of which optimized with C_1 symmetry. For $(H_2O)_4$, there are three structures with similar aqueous free energies but different gas phase free energies and energy ordering. Structure $(H_2O)_{12}$ A (C_4 symmetry) consists of layers of H_2O tetramers hydrogen bonded to one another and is 7.2 kcal/mol more stable in free energy than $(H_2O)_{12}$ B (C_1 symmetry) in aqueous solution at the MP2/aD/COSMO(G16) level. $(H_2O)_{12}$ B resembles $(H_2O)_8$ with four additional H_2O molecules that create H_2O trimers at each corner of the cube of inner H_2O molecules and is more stable than $(H_2O)_{12}$ A in the gas phase by 2.0 kcal/mol at the MP2/aD level. There are two optimized geometries for $(H_2O)_{16}$, one of which involves a $(H_2O)_8$ core with 8 other waters distributed on opposing sides of the core structure (A, C_1 symmetry) and one that maintains a column of three stacked $(H_2O)_4$ rings with a fourth, more separated $(H_2O)_4$ ring hydrogen bonded to the lowest of the stacked tetramer rings (B, C_4 symmetry). $(H_2O)_{16}$ A is lowest in gas phase free energy by 3.5 kcal/mol and $(H_2O)_{16}$ B is lowest in aqueous free energy by 8.6 kcal/mol at the MP2/aD and MP2/aD/COSMO(G16) levels, respectively. Due to computational expense, only the lowest energy structures in aqueous solution, $(H_2O)_{12}$ A and $(H_2O)_{16}$ B, were used for higher accuracy calculations.

Free energy of hydration of the H atom. The bulk solvent shift and absolute hydration free energy for H• clustered with $(H_2O)_n$ for $n = 1 – 4$ are shown in Table 6.2. The best estimate of the absolute hydration free energy ($\Delta G_{hyd}[H•,n]$) for H•(H_2O), H•$(H_2O)_2$, H•$(H_2O)_3$, and H•$(H_2O)_4$ are given in Table 6.2 at the CCSD(T)/CBS(DT)-F12b level; the bulk solvent shifts are virtually 0 kcal/mol at the MP2/aD/COSMO level. The calculated value for $\Delta G_{hyd}[H•,8]$ yields a best calculated absolute hydration free energy of H• of 4.6 kcal/mol at the CCSD(T)/aD-F12b level. For all H•$(H_2O)_n$ clusters, the calculated values for the absolute hydration free energy are

not method dependent, as most MP2/aD values (lowest level of theory used) agree with the CCSD(T)/CBS(DT)-F12b values (highest level of theory used) within 1 kcal/mol.

Free energy of hydration of H^-. The arrangement of explicit H_2O molecules has very little effect on the free energy of a cluster of H^- with H_2O molecules in both the gas phase and in aqueous solution (< 5 kcal/mol). For H^-, the bulk solvent shift decreases as the number of explicit H_2O molecules in the cluster increases. Free energy corrections for higher order calculations were taken from the optimized MP2/aD structures and combined with single point calculations using the methods listed. Combining the bulk solvent shift calculated at MP2/aD/COSMO with the gas phase free energy of the hydration of H^- calculated at various computational levels yields the results in Table 6.3.

The calculations of the absolute hydration free energy are not strongly method dependent using correlated molecular orbital theory. The calculated electron affinity (EA; ΔH_{gas}(0 K) and ΔH_{gas}(298 K)) for H at our highest level of theory, CCSD(T)/CBS(DT)-F12b, is 17.3 kcal/mol, which is in excellent agreement with the experimental EA value of 17.41 kcal/mol (0.75497 eV).[99] In contrast, the corresponding DFT (B3LYP/aD) and MP2/aD EA values are 20.5 and 7.9 kcal/mol, respectively, which is a direct result of the difficulty of calculating gas phase H^- at lower levels of theory. Despite these difficulties, cancellation of errors at each level of theory represented in Table 6.3 leads to $\Delta G_{hyd}[H^-,4]$ values that are very consistent across all levels of theory used for this work.

The $\Delta\Delta G_{sol}[H^-,1]$ and $\Delta\Delta G_{sol}[H^-,4]$ calculated at our highest level of theory, CCSD(T)/CBS(DT)-F12b/COSMO(G16), are -77.8 and -77.2 kcal/mol respectively. All CCSD(T) values except for CCSD(T)/aD/COSMO(G16) are in exact agreement for $\Delta\Delta G_{sol}[H^-,1]$ at the highest level of theory, and the CCSD(T)/aD/COSMO(G16) value only differs by 0.4

kcal/mol. The CCSD(T)/aD-F12b/COSMO(G16) and CCSD(T)/aT-F12b/COSMO(G16) values that are used to extrapolate to the CBS level are in excellent agreement with the extrapolated value differing by -1.0 and 0.1 kcal/mol, respectively, for $\Delta\Delta G_{sol}[H^-,4]$. CCSD(T)/aD/COSMO(G16) and CCSD(T)/aT/COSMO(G16), without the additional explicit electron correlation of the F12b method, yield results that are 1.0 and 0.1 higher in energy than the best value. Thus, CCSD(T)/aT/COSMO(G16) and CCSD(T)/aT-F12b/COSMO(G16) are virtually identical to the extrapolated value. The G3MP2(B3)/COSMO(G16) absolute hydration value is *ca.* -80 kcal/mol, which is 3 kcal/mol more negative than our best value. The G3MP2(B3) value is in good agreement with values calculated at the MP2 level using various basis sets, including the MP2/CBS(DTQ)/COSMO(G16) value of *ca.* -79 kcal/mol. Therefore, even our lowest level of theory is within 2 kcal/mol of our best CCSD(T)/CBS(DT)-F12b/COSMO(G16) value.

For the absolute hydration free energies of $H^-(H_2O)_n$ with n = 8, 12, and 16, geometries with the hydride on the surface of the cluster of H_2O molecules and with the hydride embedded in the explicit H_2O molecules are both given in Table 6.3. For $\Delta\Delta G_{sol}[H^-,8]$, as with $\Delta\Delta G_{sol}[H^-,4]$, the best value at our highest level of theory is *ca.* -81 kcal/mol at the CCSD(T)/CBS(DT)-F12b/COSMO(G16) level. There is very little difference in the energetics at this level between the surface hydride (-80.6 kcal/mol) and the embedded hydride (-81.0 kcal/mol). As found for $\Delta\Delta G_{sol}[H^-,4]$, the values for $\Delta\Delta G_{sol}[H^-,8]$ calculated at the coupled cluster level are all in excellent agreement with the CBS value within 1.5 kcal/mol. Again, even our lowest level of theory is within 2 to 3 kcal/mol of the CCSD(T)/CBS(DT)-F12b/COSMO(G16) value.

The absolute hydration free energy for $H^-(H_2O)_{12}$ is -82.3 kcal/mol for the surface hydride at the CCSD(T)/aD-F12b/COSMO(G16) level, which is 3.8 kcal/mol lower in energy

than the embedded hydride with 12 H_2O molecules. The embedded hydride is very similar energetically to the $H^-(H_2O)_4$ cluster at less than 0.5 kcal/mol difference at the CCSD(T)/aD-F12b/COSMO(G16) level.

The absolute hydration free energy of H^- embedded in $(H_2O)_{16}$ is -77.7 kcal/mol at the CCSD(T)/aD-F12b/COSMO(G16) level, which is 1.4 kcal/mol more favorable than the surface hydride value. The bulk solvent shift for the surface geometry of hydride with 16 H_2O molecules is less negative than that for the embedded structure by 1.4 kcal/mol. Again, these values are very similar to that of $H^-(H_2O)_4$ at the CCSD(T)/aD-F12b/COSMO(G16) level of theory and within 5 kcal/mol of the values with 8 and 12 H_2O molecules.

It is useful to examine how the continuum contributions to the free energy of hydration are converging toward zero with cluster size. For the surface hydride, the continuum contribution is -46.7 kcal/mol (n = 4), -31.9 kcal/mol (n = 8), -23.0 kcal/mol (n =12), and -9.2 kcal/mol (n = 16). These contributions are clearly becoming smaller, as expected, as the cluster size grows. As noted above, the free energy of hydration is essentially converged at n = 4 and thus the gas phase and continuum values are appropriately adding. For the embedded hydride, the continuum contribution is -33.6 kcal/mol (n = 8), -24.4 kcal/mol (n =12), and -12.9 kcal/mol (n = 16). Again the continuum contributions are converging towards zero and the free energies of hydration are almost converged at n = 8. Thus, at n = 16, the continuum contribution is less than -10 kcal/mol for the surface hydride and less than -13 kcal/mol for the embedded hydride.

The best common level of hydration energies for all clusters is CCSD(T)/aD-F12b/COSMO(G16). As noted above, this is within 1 kcal/mol for 4 H_2O molecules and within 1.5 kcal/mol for 8 H_2O molecules. The variation from n = 4 to n =16 for the embedded cluster is less than 1.5 kcal/mol so this value is reasonably well converged even at n = 4 and we can derive

a value of -78.0 kcal/mol for the hydration energy of H^-. The surface hydration energies show more variation becoming more negative from $n = 4$ to $n = 8$ and from $n = 8$ to $n = 12$. The hydration energy of the surface species becomes less negative from $n = 12$ to $n = 16$ and is now less negative than the hydration energy of the embedded species.

We take -78 kcal/mol as our best estimate of the absolute hydration free energy of H^- from the embedded hydride. The surface hydride values are not as well converged but do become comparable to that of the embedded value for the largest cluster size. We estimate that the error bar for the hydration is ± 2 kcal/mol.

Aqueous electron affinity and reduction potential of H•. The electron affinity for H• in the aqueous phase can be calculated as shown in Figure 6.3. The gas-phase electron affinity in terms of free energy at 298 K (16.9 kcal/mol); the absolute hydration free energy for H• ($\Delta G_{hyd}[H\bullet]$ = 4.6 kcal/mol); and the absolute hydration free energy for H^- ($\Delta G_{hyd}[H^-]$ = -78 ± 2 kcal/mol) can be used to calculate the aqueous electron affinity of H•. This then yields a value of 99.5 ± 2 kcal/mol (4.31 ± 0.1 V) for the aqueous electron affinity of H•. The error bar of ± 2 kcal/mol encompasses both the surface and embedded values.

In order to convert the adiabatic electron affinity into a reduction potential, the absolute potential of the standard hydrogen electrode (SHE) is needed. Although IUPAC recommends a value of $E°$ = -4.44 V,[100] we prefer to use a more recent value of $E°$ = -4.28 V.[101,102] We write reaction (6.9) to obtain the $E°$ for H•/H^- relative to SHE. This reaction energy then corresponds to $E°$.

$$H^+(aq) + H^-(aq) \rightarrow H\bullet(aq) + \tfrac{1}{2}\, H_2(g) \qquad (6.9)$$

Following Marenich et al.,[101] we note that we can use the adiabatic electron affinity as the conversion from a standard state of 1 bar to 1 atm is small (< 0.04 kJ/mol) and change from 1

molar to 1 molal is also small and can be neglected, especially in comparison to other errors in the calculations. The energy for reaction (6.9) is 0.03 eV giving a reduction potential for the hydrogen atom of -0.03 ± 0.1 V. This value is in good agreement with several previous estimates of the H•/H^- reduction potential, shown in Table 6.1, which have potentials very close to that of the SHE electrode.[30,31,34,36,38,40] The notable exceptions are the results from Kelly and Rosseinsky and the analogous results from Connelly, et al.,[42,43] Both of these reports base the solvation free energy of hydride on an extrapolation from the halide anions using estimates of the hydride radius.

Why might the results based on extrapolating from the halide anions using the hydride radius not be appropriate? To address this question, we examined the same structures for H^- and F^- with 4 and 8 H_2O molecules. The bond distances are given in the Appendix and clearly show that the hydride is closer to the H atoms on the H_2O molecules involved in hydrogen bonding to the anion than is the fluoride. Thus, the H^- has an apparent smaller effective radius, not a larger one than the F^- in our solvent model. Thus, it is not appropriate to extrapolate based on the 'ionic radii.'

Why does the H^- then not have a larger solvation energy that does the F^-? In fact, the free energy of solvation of F^- is predicted[65] to be -104.3 ± 0.7 kcal/mol and our value for H^- is -78 ± 2 kcal/mol. Examination of the NPA charges from the NBO analysis (see Appendix) provides us some insights into this. The H^- charges for the 4 H_2O cluster range from -0.64 to -0.71 e and the corresponding F^- charge is -0.82 to -0.83 e. For the 8 H_2O cluster, the H^- charges range from -0.58 to -0.62 e and the corresponding F^- charge is -0.80 to -0.81 e. Clearly the F^- holds onto its charge more than does the H^-. This is totally consistent with the electron affinities with EA(F) = 3.40 eV[103] being much larger than EA(H) = 0.755 eV.[99] Thus, the additional electron density for

F^- remains mostly on the F, but for H^- the electron density delocalizes onto the adjacent H_2O molecules so the solvation energetics will be very different.

To provide additional insights into differences in H^- and F^-, we calculated the polarizabilities of these ions at the CCSD (full CI) level for H^- and at the CCSD(T) level for F^- using triply augmented correlation consistent basis sets[104] at the sextuple level. The value for H^- is 34.05 $Å^3$, whereas that for F^- is 2.46 $Å^3$. The value for F^- is more converged in terms of the basis set size than for H^-. As an additional comparison, the value of the polarizability for Cl^- is 5.36 $Å^3$ at the same computational level. The polarizability of H^- is thus more than ten times that of F^-, consistent with the delocalization of charge onto the associated waters for H^- being larger than that for F^-. In summary, these results suggest that one should not use an ionic radii extrapolation from the halides to get the solvation free energy for H^- as the two types of species are very different.

The absolute free energy of solvated H^- can be estimated from the above data. The heat of formation of H^- in the gas phase at 298 K from the Active Thermochemical Tables[105,106] is 34.7 kcal/mol. The conversion[107] from $\Delta H_{f,gas}$(298 K) to $\Delta G_{f,gas}$(298 K) is -1.6 kcal/mol, giving $\Delta G_{f,gas}(H^-,298\ K) = 33.1$ kcal/mol. Addition of our solvation energy gives $\Delta G_f(H^-_{aq}, 298\ K) = -44.9 \pm 2$ kcal/mol.

An application of this result is to the isotope exchange reaction of OH^- and H_2 in H_2O represented by reaction (6.10) which has been used as evidence of the existence of the aqueous

$$OH^- + H_2 = H_2O + H^- \qquad (6.10)$$

hydride anion. The kinetics of this reaction was reported by Wilmarth et al. in 1953.[108] The experiments were performed at rather high temperatures, so it is a long extrapolation to 25 °C. From the available data, rate constant of 1.7×10^{-4} M^{-1} min^{-1} can be estimated, which

corresponds to $\Delta G^{\pm}$ = 25 kcal/mol. Using the data given in the Appendix, reaction (6.10) has $\Delta H_{gas,298K}$ = 10.1 kcal/mol and ΔG_{gas}(298 K) = 10.5 kcal/mol. Inclusion of the solvation energy corrections, gives ΔG_{aq}(298 K) = 24.3 kcal/mol for reaction (6.10) so this could be the rate limiting step if the barrier height is just the endothermicity of the reaction.

Conclusions

Using the supermolecule-continuum approach, we have calculated the absolute hydration free energies for atomic hydrogen (H•) and hydride (H^-) to be 4.6 ± 1 kcal/mol and -78 ± 2 kcal/mol at the CCSD(T)/CBS(DT)-F12b/COSMO(G16) level, respectively. This value for the absolute hydration free energy of H^- encompasses the behavior of both surface and embedded H^- in clusters of up to 16 H_2O molecules. The electron affinity of H• (ΔH_{gas}(0 K)) in the gas phase at the CCSD(T)/CBS(DT)-F12b level was found to be 17.3 kcal/mol, which is in excellent agreement with the experimental EA value of 17.41 kcal/mol (0.75497 eV).[99] In terms of ΔG_{gas}(298 K), the electron affinity of H• is 16.9 kcal/mol at the CCSD(T)/CBS(DT)-F12b level. The combination of free energy values for gas phase electron affinity with the solvation free energies for H• and H^- allows for the calculation of the aqueous electron affinity of H•, which we find to be 99.5 ± 2 kcal/mol or 4.31 ± 0.1 eV. This aqueous electron affinity yields a predicted reduction potential for H• of -0.03 ± 0.1 V vs SHE, which is in good agreement with most previous estimates.[29,30,31,32,33,34,38] Errors in the absolute accuracy of thermodynamic parameters, such as the reduction potential for H•, do not affect the ability to predict chemical reactivity using established relative hydricity scales, as the differences in free energies between reacting species will remain unchanged. However, a more accurate determination of these parameters provides fundamental insight into how a putative hydride interacts with a variety of solvents and

enable the development of an absolute scale as well as the relative ones that play an important role in understanding reactivity.

Tables

Table 6.1. Reduction Potential of the H•/H^- Couple vs SHE in V

$E°(H•/H^-_{aq})$	Reference
-0.36	Griller, et al.[35]
-0.14	Eberson[30]
-0.12	Parker[36]
-0.03 ± 0.1	this work
0.05	Toffel & Henglein[31]
0.18	Pearson[34]
0.18	Parker[40]
0.81	Connelly, et al.[43]
0.83	Kelly & Rosseinsky[42]

Table 6.2. Absolute Hydration Free Energy of Radical Hydrogen (ΔG_{hyd}[H•,n]) Calculated as the Free Energy Change from H•(gas) + $(H_2O)_n$(aq) → H•$(H_2O)_n$(aq) at 298 K in kcal/mol

Method [a]	n = 1	n = 2	n = 3	n = 4	n = 8
Bulk Solvent Shift ($\Delta\Delta G_{sol}$[H•,n])					
MP2/aD/COSMO(G16)	-0.2	-0.1	-0.1	-0.1	0.0
MP2/aD/SMD(G16)	1.1	1.1	1.1	1.5	1.5
MP2/aD/PCM(G16)	0.0	0.1	0.0	0.0	0.0
MP2/aD/COSMO(G03-Electrostatic)	0.2	0.3	0.3	0.1	0.0
MP2/aD/COSMO(G03-all)	0.2	0.3	0.3	0.1	0.0
MP2/aD	2.1	2.4	2.4	1.8	4.4
MP2/aT	2.2	2.8	2.8	3.3	
MP2/aQ with MP2/aT corr.	2.3	2.9	3.9	3.7	
MP2/CBS(DTQ)	2.4	3.0	4.6	3.9	
G3MP2(B3)				3.1	
CCSD(T)/aD	2.2	2.9	3.1	3.5	4.1
CCSD(T)/aT	2.2	2.7	3.1	3.5	
CCSD(T)/aD-F12b	2.2	2.8	3.1	3.5	4.6
CCSD(T)/aT-F12b	2.2	2.7	3.1	3.5	
CCSD(T)/CBS(DT)-F12b	2.2	2.7	3.1	3.5	

[a] Method for the energy calculations in the gas phase. All solvation contributions were calculated at the MP2/aD/COSMO, MP2/aD/SMD, and MP2/aD/PCM levels with Gaussian 16 and at the MP2/aD/COSMO level with Gaussian 03 for the bulk solvent shift, either with electrostatic components only ("G03-Electrostatic") or with both electrostatic and non-electrostatic components ("G03-all"). Values for the absolute hydration energy are reported with MP2/aD/COSMO(G16) electrostatic only.

Table 6.3. Absolute Hydration Free Energy of Hydride ($\Delta G_{hyd}[H^-,n]$) Calculated as the Free Energy Change from $H^-(gas) + (H_2O)_n(aq) \rightarrow H^-(H_2O)_n(aq)$ at 298 K in kcal/mol

Method [a]	Energy	*n* = 1	*n* = 4	*n* = 8		*n* = 12		*n* = 16	
			Surface	Embed	Surface	Embed	Surface	Embed	Surface
Bulk Solvent Shift ($\Delta\Delta G_{sol}[H^-,n]$)									
MP2/aD/COSMO(G16)		-67.2	-46.7	-33.6	-31.9	-24.4	-23.0	-12.9	-9.3
MP2/aD/SMD(G16)		-65.8	-47.0	-33.0	-33.0	-25.6	-23.9	-12.6	-9.9
MP2/aD/PCM(G16)		-65.6	-46.8	-32.1	-32.1	-13.8	-12.3	-17.7	-12.9
MP2/aD/COSMO(G03-Electrostatic)		-67.4	-46.5	-31.7	-31.7	-24.4	-22.2	-13.0	-7.8
MP2/aD/COSMO(G03-all)		-74.4	-49.0	-31.1	-31.1	-23.7	-21.2	-12.1	-5.7
MP2/aD	$\Delta G_{gas}[H^-,n]$	-12.3	-32.2	-50.1	-50.7	-54.1	-59.7	-64.8	-67.0
MP2/aD	$\Delta G_{hyd}[H^-,n]$	-79.5	-78.9	-83.7	-82.6	-78.5	-82.7	-77.7	-76.3
MP2/aT	$\Delta G_{gas}[H^-,n]$	-12.6	-33.3						
MP2/aT	$\Delta G_{hyd}[H^-,n]$	-79.8	-80.0						
MP2/aQ	$\Delta G_{gas}[H^-,n]$	-12.6	-33.0						
MP2/aQ	$\Delta G_{hyd}[H^-,n]$	-79.8	-79.6						
MP2/CBS(DTQ)	$\Delta G_{gas}[H^-,n]$	-12.6	-32.7						
MP2/CBS(DTQ)	$\Delta G_{hyd}[H^-,n]$	-79.8	-79.3						
G3MP2(B3)	$\Delta G_{gas}[H^-,n]$	-11.7	-33.5						
G3MP2(B3)	$\Delta G_{hyd}[H^-,n]$	-78.9	-80.2						
CCSD(T)/aD	$\Delta G_{gas}[H^-,n]$	-10.2	-29.6	-46.9	-47.6	-52.9	-58.4	-64.7	-67.3
CCSD(T)/aD	$\Delta G_{hyd}[H^-,n]$	-77.4	-76.2	-80.5	-79.5	-77.2	-81.4	-77.6	-76.6
CCSD(T)/aT	$\Delta G_{gas}[H^-,n]$	-10.6	-30.4	-47.5	-48.6				
CCSD(T)/aT	$\Delta G_{hyd}[H^-,n]$	-77.8	-77.1	-81.1	-80.4				
CCSD(T)/aD-F12b	$\Delta G_{gas}[H^-,n]$	-10.6	-31.5	-47.6	-48.7	-54.1	-59.3	-65.1	-67.3
CCSD(T)/aD-F12b	$\Delta G_{hyd}[H^-,n]$	-77.8	-78.2	-81.2	-80.6	-78.5	-82.3	-78.0	-76.6
CCSD(T)/aT-F12b	$\Delta G_{gas}[H^-,n]$	-10.6	-30.5	-47.5	-48.7				
CCSD(T)/aT-F12b	$\Delta G_{hyd}[H^-,n]$	-77.8	-77.1	-81.1	-80.6				
CCSD(T)/CBS-F12b	$\Delta G_{gas}[H^-,n]$	-10.6	-30.5	-47.4	-48.8				
CCSD(T)/CBS-F12b	$\Delta G_{hyd}[H^-,n]$	-77.8	-77.2	-81.0	-80.6				

[a] Method for the energy calculations in the gas phase. All solvation contributions were calculated at the MP2/aD/COSMO, MP2/aD/SMD, and MP2/aD/PCM levels with Gaussian 16 and at the MP2/aD/COSMO level with Gaussian 03 for the bulk solvent shift, either with electrostatic components only ("G03-Electrostatic") or with both electrostatic and non-electrostatic components ("G03-all"). Values for the absolute hydration energy are reported with MP2/aD/COSMO(G16) electrostatic only.

Figures

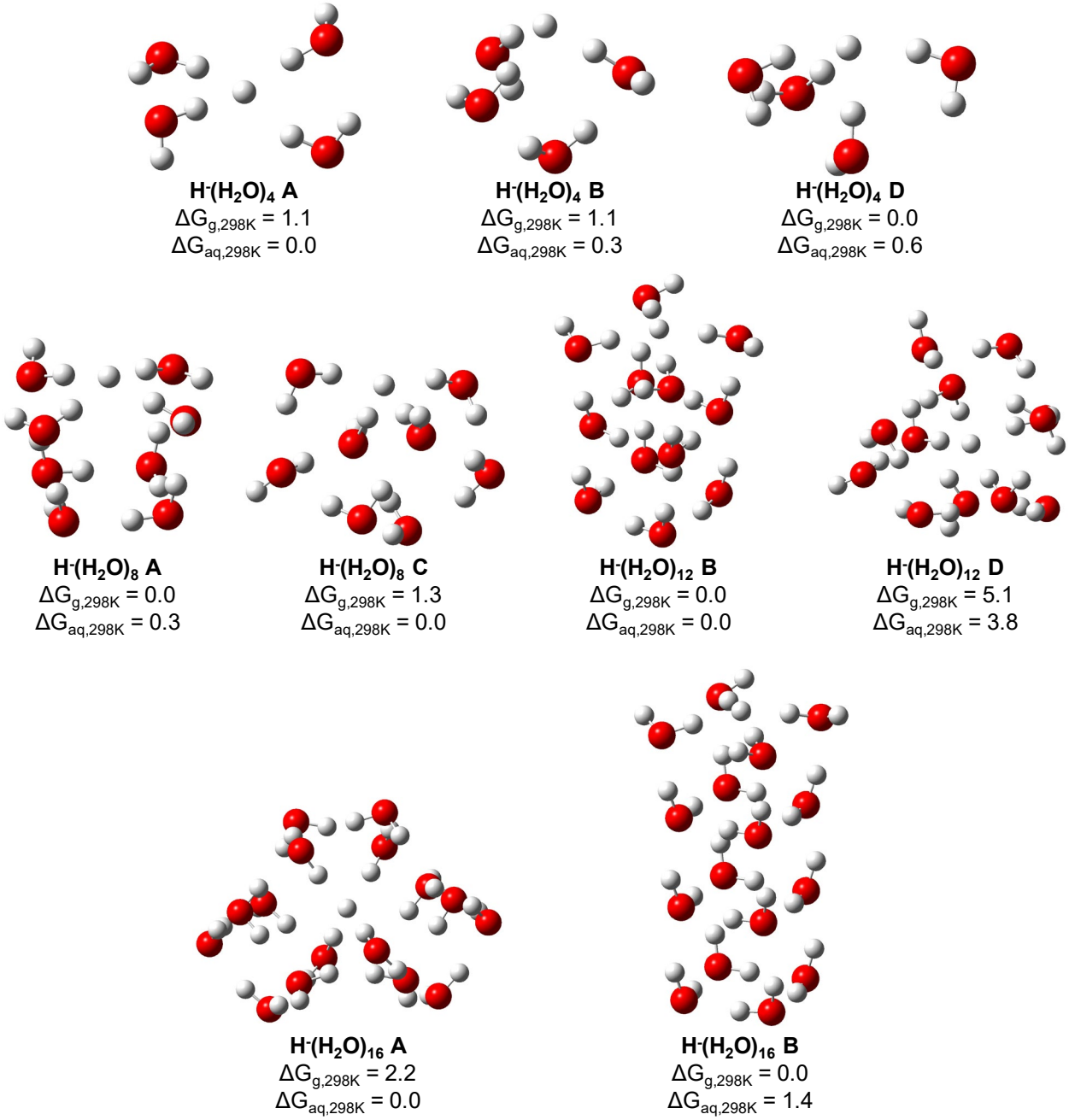


Figure 6.1. $H^-(H_2O)_n$ clusters with $n = 4$, 8, 12, and 16. Gas phase and aqueous free energy values relative to the lowest energy isomer for a given species are calculated at the CCSD(T)/CBS(DT)-F12b level for n = 4, 8 and at the CCSD(T)/aD-F12b level for n = 12, 16. All solvation contributions are calculated at the MP2/aD/COSMO(G16) level.

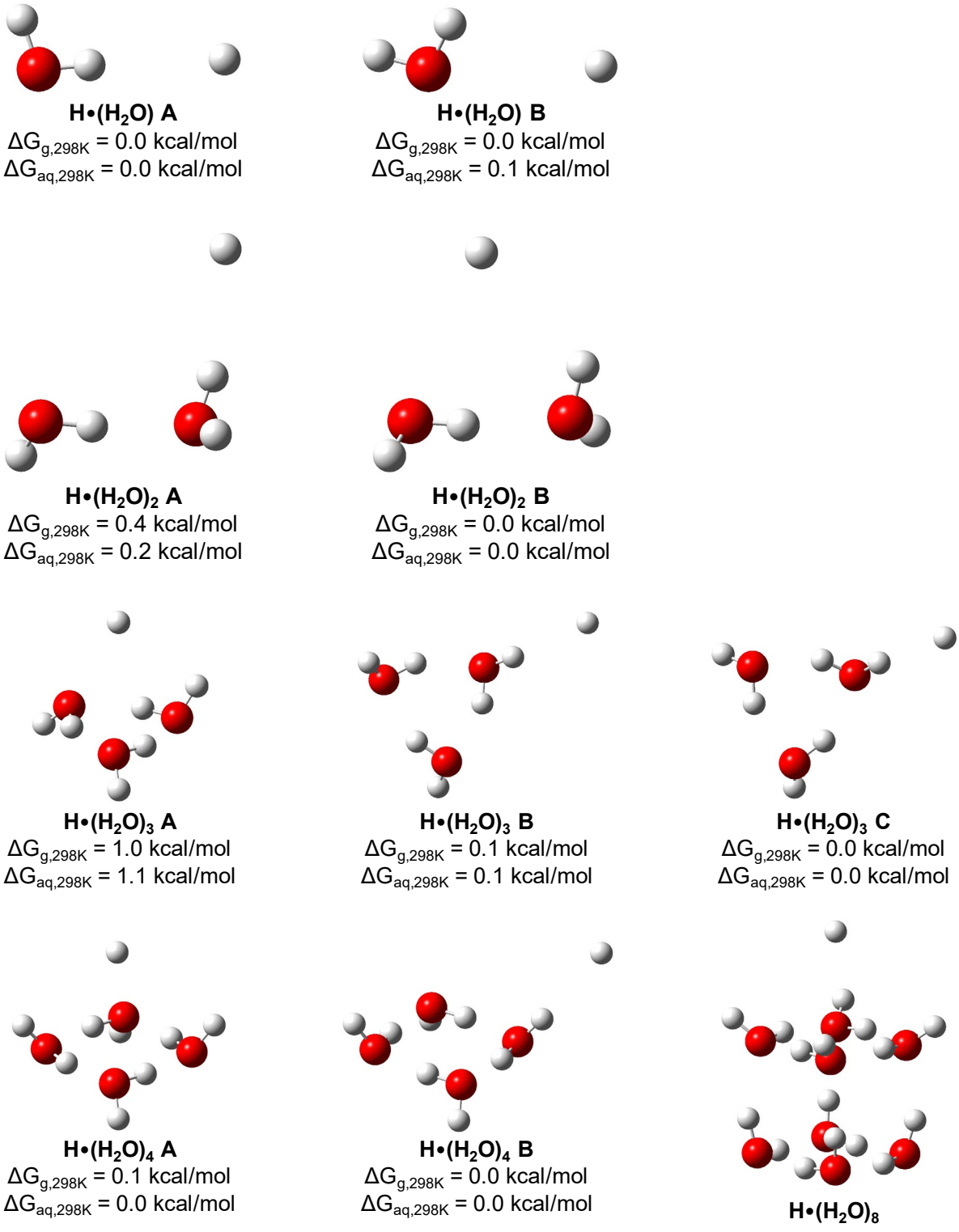


Figure 6.2. H•$(H_2O)_n$ clusters with n = 1 – 4 and 8. Gas phase and aqueous free energy values relative to the lowest energy isomer for a given species are calculated at the CCSD(T)/CBS(DT)-F12b level for n = 1 – 4 and at the CCSD(T)/aD-F12b level for n = 8. Solvation contributions are calculated at the MP2/aD/COSMO(G16) level.

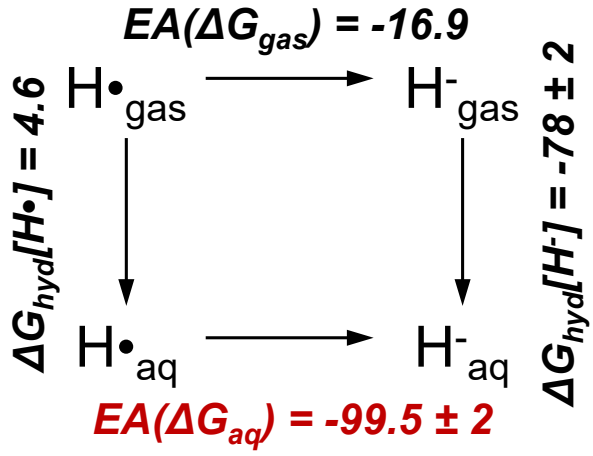


Figure 6.3. Energy diagram for the creation of aqueous H^- from gas-phase H• with free energy values in kcal/mol at 298 K.

References

[1] Conway, B. E. *Ionic Hydration in Chemistry and Biophysics*; Elsevier: New York, 1981.

[2] Marcus, Y. *Ion Solvation*; Wiley: New York, 1985.

[3] Weber, J. M.; Kelley, J. A.; Nielsen, S. B.; Ayotte, P.; Johnson, M. A. Isolating the Spectroscopic Signature of a Hydration Shell With the Use of Clusters: Superoxide Tetrahydrate. *Science* **2000**, *287*, 2461 – 2463.

[4] Pasquarello, A.; Petri, I.; Salmon, P. S.; Parisel, O.; Car, R.; Tóth, É.; Powell, D. H.; Fischer, H. E.; Helm, L.; Merbach, A. First Solvation Shell of the Cu(II) Aqua Ion: Evidence for Fivefold Coordination. *Science* **2001**, *291*, 856 – 859.

[5] Kropman, M. F.; Bakker, H. J. Dynamics of Water Molecules in Aqueous Solvation Shells. *Science* **2001**, *291*, 2118 – 2120.

[6] Robertson, W. H.; Diken, E. G.; Price, E. A.; Shin, J.-W.; Johnson, M. A. Spectroscopic Determination of the OH^- Solvation Shell in the $OH^-\cdot(H_2O)_n$ Clusters. *Science* **2003**, *299*, 1367 – 1372.

[7] Burgess S. A.; Kendall, A. J.; Tyler, D. R.; Linehan, J. C.; Appel. A. M. Hydrogenation of CO_2 in Water Using a Bis(Diphosphine) Ni-H Complex. *ACS Catal.* **2017**. *7*, 3089 – 3096.

[8] Matsubara, Y.; Fujita, E.; Doherty, M. D.; Muckerman, J. T.; Creutz, C. Thermodynamic and Kinetic Hydricity of Ruthenium(II) Hydride Complexes. *J. Am. Chem. Soc.* **2012**, *134*, 15743 – 15757.

[9] Wiedner, E. S.; Chambers, M. B.; Pitman, C. L.; Bullock, R. M.; Miller, A. J. M.; Appel, A. M. Thermodynamic Hydricity of Transition Metal Hydrides. *Chem. Rev.* **2016**, *116*, 15, 8655 – 8692.

[10] Wiedner, E. S.; Linehan, J. C. Making a Splash in Homogeneous CO_2 Hydrogenation: Elucidating the Impact of Solvent on Catalytic Mechanisms. *Chem. Eur. J.* **2018**, *24*, 16964 – 16971.

[11] Ceballos, B. M.; Yang, J. Y. Directing the Reactivity of Metal Hydrides for Selective CO_2 Reduction. *Proc. Nat. Acad. Sci. USA* **2018**, *115*, 12686 – 12691.

[12] Ceballos, B. M.; Tsay, C.; Yang, J. Y. CO_2 reduction or HCO_2^- oxidation? Solvent-Dependent Thermochemistry of a Nickel Hydride Complex. *Chem. Commun.* **2017**, *53*, 7405 – 7408.

[13] Tsay, C.; Livesay, B. N.; Ruelas, S.; Yang, J. Y. Solvation Effects on Transition Metal Hydricity. *J. Am. Chem. Soc.* **2015**, *137*, 14114 – 14121.

[14] Brereton, K. R.; Jadrich, C. N.; Stratakes, B. M.; Miller, A. J. M. Thermodynamic Hydricity Across Solvents: Subtle Electronic Effects and Striking Ligation Effects in Iridium Hydrides. *Organometallics* **2019**, *38*, 3104 – 3110.

[15] Brereton, K. R.; Pitman, C. L.; Cundari, T. R.; Miller, A. J. M. Solvent-Dependent Thermochemistry of an Iridium/Ruthenium H_2 Evolution Catalyst. *Inorg. Chem.* **2016**, *55*, 12042 – 12051.

[16] Pitman, C. L.; Brereton, K. R.; Miller, A. J. M. Aqueous Hydricity of Late Metal Catalysts as a Continuum Tuned by Ligands and the Medium. *J. Am. Chem. Soc*. **2016**, *138*, 2252 – 2260.

[17] Connelly Robinson, S. J.; Zall, C. M.; Miller, D. L.; Linehan, J. C.; Appel, A. M. Solvent Influence on the Thermodynamics for Hydride Transfer from Bis(diphosphine) Complexes of Nickel. *Dalton Trans.* **2016**, *45*, 10017 – 10023.

[18] Cammarota, R. C.; Xie, J.; Burgess, S. A.; Vollmer, M. V.; Vogiatzis, K.; Ye, J.; Linehan, J. C.; Appel, A. M.; Hoffmann, C.; Wang, X.; et al. Thermodynamic and Kinetic Studies of H_2 and N_2 Binding to Bimetallic Nickel-Group 13 Complexes and Neutron Structure of a Ni(η^2-H_2) Adduct. *Chem. Sci.* **2019**, *10*, 7029 – 7042.

[19] Burgess, S. A.; Appel, A. M.; Linehan, J. C.; Wiedner, E. S. Changing the Mechanism for CO_2 Hydrogenation Using Solvent-Dependent Thermodynamics. *Angew. Chem. Int. Ed*. **2017**, *56*, 15002 – 15005.

[20] Taheri, A.; Berben, L. A. Tailoring Electrocatalysts for Selective CO_2 or H^+ Reduction: Iron Carbonyl Clusters as a Case Study. *Inorg. Chem.* **2016**, *55*, 378 – 385.

[21] Christe, K. O.; Dixon, D. A.; McLemore, D. K.; Wilson, W. W.; Sheehy, J.; Boatz, J. A. On a Quantitative Scale for Lewis Acidity and Recent Progress in Polynitrogen Chemistry. *J. Fluor. Chem.*, **2000**, *101*, 151 – 153.

[22] Grant, D. J.; Dixon, D. A.; Potter, R. G.; Camaioni, D.; Christe, K. O. Lewis Acidities and Hydride, Fluoride, and X^- Affinities of the $H_{3-n}BX_n$ compounds for (X = F, Cl, Br, I, NH_2, OH, and SH) from Coupled Cluster Theory. *Inorg. Chem.* **2009**, *48,* 8811 – 8821.

[23] Hunter, E. P.; Lias, S. G. Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update. *J. Phys. Chem. Ref. Data*, **1998**, *27*, 3, 413 – 656.

[24] Buncel, E.; Menon, B. Carbanion Mechanisms. 6. Metalation of Arylmethanes by Potassium Hydride/18-Crown-6 Ether in Tetrahydrofuran and the Acidity of Hydrogen. *J. Am. Chem. Soc.* **1977**, *99*, 4457 – 4461.

[25] Anick, D. J.; Leung, K. Transition States for Hydride-Water $(H^-)(H_2O)_n$ Clusters, n = 2 – 6, 20. *J. Mol. Struc. – THEOCHEM* **2009**, *916*, 61 – 71.

[26] Kasajima, T.; Nishikiori, T.; Hohira, T.; Ito, Y. Electrochemical Behavior of Hydride Ion in a LiBr-KBr-CsBr Eutectic Melt. *J. Electrochem. Soc.* **2003**, *150*, E403 – E408.

[27] Vogler, A.; Kunkely, H. Photolysis of the Tantalum Cluster $Ta_6Br_{12}^{2+}$ in Aqueous Acidic Solution. *Inorg. Chem.* **1984**, *23*, 1360 – 1363.

[28] Sander, R. Compilation of Henry's Law Constants (Version 4.0) for Water as Solvent. *Atmos. Chem. Phys.*, **2015**, *15*, 4399 – 4981.

[29] Latimer, W. M. Hydrogen. In *The Oxidation States of the Elements and their Potentials in Aqueous Solution*, 2nd ed.; Prentice-Hall, Englewood Cliffs, NJ, 1952, pp. 26 – 33.

[30] Eberson, L. Electron Transfer Reactions in Organic Chemistry. V. Examination of Postulated Electron Transfer Processes Involving Anionic and Carbanionic Nucleophiles with the Aid of the Marcus Theory. *Acta Chem. Scand. B* **1984**, *38*, 439 – 459.

[31] Toffel, P.; Henglein, A. Polarogram of the Free Hydrogen Atom and of Some Simple Organic Radicals. *Faraday Discuss. Chem. Soc.* **1977**, *63*, 124 – 133.

[32] Ross, P. N. Hydrogen. In *Standard Potentials in Aqueous Solution*, 1st ed.; A. J. Bard, R. Parsons, and J. Jordan, Eds.; Marcel Dekker, New York, 1985, pp. 39 – 48.

[33] Bratsch, S. G. Standard Electrode Potentials and Temperature Coefficients in Water at 298.15 K. *J. Phys. Chem. Ref. Data* **1989**, *18*, 1 – 21.

[34] Pearson, R. G. Ionization Potentials and Electron Affinities in Aqueous Solution. *J. Am. Chem. Soc.* **1986**, *108*, 6109 – 6114.

[35] Griller, D.; Martinho Simoes, J. A.; Mulder, P.; Sim, B. A.; Wayner, D. D. M. Unifying the Solution Thermochemistry of Molecules, Radicals, and Ions. *J. Am. Chem. Soc.* **1989**, *111*, 7872 – 7876.

[36] Parker, V. D. Homolytic Bond (H-A) Dissociation Free Energies in Solution. Applications of the Standard Potential of the (H^+/H•) Couple. *J. Am. Chem. Soc.* **1992**, *114*, 7458 – 7462.

[37] Parker, V. D. Correction to Dissociation Free Energies in Solution. Applications of the Standard Potential of the (H^+/H•) Couple. *J. Am. Chem. Soc.* **1993**, *115*, 1201.

[38] Stanbury, D. M. Reduction Potentials Involving Inorganic Free Radicals in Aqueous Solution. *Adv. Inorg. Chem.* **1989**, *33*, 69 – 138.

[39] Roduner, E.; Bartels, D. M. Solvent and Isotope Effects on Addition of Atomic Hydrogen to Benzene in Aqueous Solution. *Ber. Bunsenges. Phys. Chem.* **1992**, *96*, 1037 – 1042.

[40] Wayner, D. D. M.; Parker, V. D. Bond Energies in Solution from Electrode Potentials and Thermochemical Cycles. A Simplified and General Approach. *Acc. Chem. Res.* **1993**, *26*, 287 – 294.

[41] Parker, V. D. Hydride Affinities of Organic Cations in Solution. *Acta Chem. Scand.* **1992**, *46*, 1133 – 1134.

[42] Kelly, C. A.; Rosseinsky, D. R. Estimates of Hydride Ion Stability in Condensed Systems: Energy of Formation and Solvation in Aqueous and Polar-Organic Solvents. *Phys. Chem. Chem. Phys.* **2001**, *3*, 2086 – 2090.

[43] Connelly, S. J.; Wiedner, E. S.; Appel, A. M. Predicting the Reactivity of Hydride Donors in Water: Thermodynamic Constants for Hydrogen. *Dalton Trans*. **2015**, *44*, 5933 – 5938.

[44] Chałasiński, G.; Kendall, R. A.; Simons, J. *Ab Initio* Studies of the Structures and Energies of the $H^-(H_2O)$ and $H^-(H_2O)_2$ Complexes. *J. Chem. Phys.* **1987**, *87*, 2965 – 2975.

[45] Xantheas, S. S.; Dunning, Jr., T. H. Structures of Anion-Water Clusters: $H^-(H_2O)_n$, $n = 1 – 3$. *J. Phys. Chem.* **1992**, *96*, 7505 – 7506.

[46] Lee, H. M.; Anick, D. J.; Kim, K. S. Structures of Tri-, Tetra-, and Hexahydrated Hydride Anion Clusters. *Int. J. Quantum Chem.* **2009**, *109*, 1820 – 1826.

[47] Prakash, M.; Gopalsamy, K.; Subramanian, V. Studies on the Structure, Stability, and Spectral Signatures of Hydride Ion-Water Clusters. *J. Chem. Phys.* **2011**, *135*, 214308-1 – 214308-18.

[48] Pliego Jr., J. R.; Riveros, J. M. Hybrid Discrete-Continuum Solvation Methods. *WIREs Comput. Mol. Sci.* **2020**, *10:e1440*, 1 – 25.

[49] Mennucci, B.; Cammi, R. (Eds). *Continuum Solvation Models in Chemical Physics: From Theory to Applications.* Wiley: Chichester, UK, 2007.

[50] Martin, R. L.; Hay, P. J.; Pratt, L. R. Hydrolysis of Ferric Ion in Water and Conformational Equilibrium. *J. Phys. Chem. A* **1998**, *102*, 3565 – 3573.

[51] Asthagiri, D.; Pratt, L. R.; Kress, J. D.; Gomez, M. A. The Hydration State of HO^-(aq). *Chem Phys Lett.* **2003**, *380*, 530 – 535.

[52] Asthagiri, D.; Pratt, L. R.; Ashbaugh, H. S. Absolute Hydration Free Energies of Ions, Ion-Water Clusters, and Quasichemical Theory. *J Chem Phys.* **2003**, *119*, 2702 – 2708.

[53] Pliego Jr., J. R.; Riveros, J. M. The Cluster-Continuum Model for the Calculation of the Solvation Free Energy of Ionic Species. *J. Phys. Chem. A* **2001**, *105*, 7241 – 724.

[54] Grabowski, P.; Riccardi, D.; Gomez, M. A; Asthagiri, D.; Pratt, L. R. Quasi-chemical Theory and the Standard Free Energy of H^+(aq). *J. Phys. Chem. A* **2002**, *106*, 9145 – 9148.

[55] Pliego Jr., J. R.; Riveros, J. M. Theoretical Calculation of pK_a Using the Cluster-Continuum Model. *J. Phys. Chem. A* **2002**, *106*, 7434 – 7439.

[56] Ho, J.; Coote, M. L. First-Principles Prediction of Acidities in the Gas and Solution Phase. *WIREs Comput. Mol. Sci.* **2011**, *1*, 649 – 660.

[57] Hummer, G.; Pratt, L. R.; Garcia, A. E. Free Energy of Ionic Hydration. *J. Phys. Chem.* **1996**, *100*, 1206 – 1215.

[58] Rempe, S. B.; Pratt, L. R.; Hummer, G.; Kress J. D.; Martin, R. L.; Redondo, A. The Hydration Number of Li^+ in Liquid Water. *J. Am. Chem. Soc*. **2000**, *122*, 966 – 967.

[59] Ho, J.; Coote, M. L. A Universal Approach for Continuum Solvent pK(a) Calculations: Are We There Yet? *Theor. Chem. Acc*. **2010**, *125*, 3 – 21.

[60] Seybold, P. G.; Shields, G. C. Computational Estimation of pK_a Values. *WIREs Comput. Mol. Sci.* **2015**, *5*, 290 – 297.

[61] Alongi, K. S.; Shields, G. C. Theoretical Calculations of Acid Dissociation Constants: A Review Article, In *Annual Reports in Computational Chemistry*; Wheeler, R. A., Ed. Elsevier: Amsterdam, 2010; Vol. 6, pp 113 – 138.

[62] Ho, J.; Ertem, M. Z. Calculating Free Energy Changes in Continuum Solvation Models. *J. Phys. Chem. B* **2016**, *120*, 1319 – 1329.

[63] Zhan, C.-G.; Dixon, D. A. Absolute Hydration Free Energy of the Proton from First-Principles Electronic Structure Calculations. *J. Phys. Chem. A* **2001**, *105*, 11534 – 11540.

[64] Zhan, C.-G.; Dixon, D. A. First-Principles Determination of the Absolute Hydration Free Energy of the Hydroxide Ion *J. Phys. Chem. A* **2002**, *106*, 9737 – 9744.

[65] Zhan, C.-G.; Dixon, D. A. Hydration of the Fluoride Anion: Structures and Absolute Hydration Free Energy from First-Principles Electronic Structure Calculations. *J. Phys. Chem. A* **2004**, *108*, 2020 – 2029.

[66] McNeill, A. S.; Dixon, D. A. Energetics of CO_2^- in Aqueous Solution. *J. Phys. Chem. A* **2019**, *123*, 1243 – 1259.

[67] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. *Gaussian 16, Revision A.03*, Gaussian, Inc., Wallingford CT, 2016.

[68] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 0618 – 0622.

[69] Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical Models Incorporating Electron Correlation. *Int. J. Quantum. Chem.* **1976**, *10*, 1 – 19.

[70] Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys*. **1989**, *90*, 1007 – 1023.

[71] Kendall, R. A.; Dunning, T. H. Jr.; Harrison, R. J., Electron Affinities of the First Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys*. **1992,** *96*, 6796 – 6806.

[72] Martin, J. M. L. Ab Initio Total Atomization Energies of Small Molecules - Towards the Basis Set Limit. *Chem. Phys. Lett.* **1996**, *259*, 669 – 678.

[73] Feller, D.; Peterson, K.A.; Hill, J. G. On the Effectiveness of CCSD(T) Complete Basis Set Extrapolations for Atomization Energies. *J. Chem. Phys.* **2011***, 135,* 044102-1 – 044102-18.

[74] Werner H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, T.; Korona, T.; Lindh, R.; et al. *MOLPRO, version 2015.1, a package of ab initio programs*, See http://www.molpro.net. Accessed March 1, 2016.

[75] Werner H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242 – 253.

[76] Purvis, G. D., III; Bartlett, R. J. A Full Coupled Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910 – 1918.

[77] Raghavachari, K.; Truck, G.W.; Pople, J.A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett*. **1989**, *157*, 479 – 483.

[78] Watts, J. D.; Gauss, J.; Bartlett, R. J. Coupled-Cluster Methods with Noniterative Triple Excitations for Restricted Open Shell Hartree-Fock and Other General Single Determinant Reference Functions. Energies and Analytical Gradients. *J. Chem. Phys.* **1993**, *98*, 8718 – 8733.

[79] Bartlett, R. J.; Musial, M. Coupled Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291 – 352.

[80] Knowles, P. J.; Hampel, C.; Werner, H.-J. Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **1993**, *99*, 5219 – 5227.

[81] Knowles, P. J.; Hampel, C.; Werner, H.-J. Erratum: Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **2000,** *112*, 3106 – 3107.

[82] Deegan, M. J. O.; Knowles, P. J. Perturbative Corrections to Account for Triple Excitations in Closed and Open Shell Coupled Cluster Theories. *Chem. Phys. Lett*. **1994**, *227*, 321 – 327.

[83] Adler, T. B.; Knizia, G.; Werner, H.-J. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Chem. Phys.* **2007,** *127,* 221106-1 – 221106-4.

[84] Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009,** *130,* 054104-1 – 054104-20.

[85] Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-ζ Quality Coupled-Cluster Correlation Energies with Triple- ζ Basis Sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921 – 1930.

[86] Yousaf, K. E.; Peterson, K. A. Optimized Auxiliary Basis Sets for Explicitly Correlated Methods. *J. Chem. Phys.* **2008**, *129*, 184108-1 – 184108-7.

[87] Hill, J. G.; Mazumder, S.; Peterson, K. A. Correlation Consistent Basis Sets for Molecular Core-Valence Effects with Explicitly Correlated Wave Functions: The Atoms B-Ne and Al-Ar. *J. Chem. Phys.* **2010**, *132*, 054108-1 – 054108-11.

[88] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, 105, 2999 – 3093.

[89] Klamt, A. *COSMO RS, Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam, 2005.

[90] Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc. Perkin Trans.* **1993**, 2, 799 – 805.

[91] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378 – 6396.

[92] Miertus, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum - a Direct Utilization of Ab Initio Molecular Potentials for the Prevision of Solvent Effects. *Chem. Phys.* **1981**, *55*, 117 – 129.

[93] Pascual-Ahuir, J. L.; Silla, E.; Tunon, I. GEPOL: An Improved Description of Molecular-Surfaces 3. A New Algorithm for the Computation of a Solvent-Excluding Surface, *J. Comput. Chem.* **1994**, *15*, 1127 – 1138.

[94] Scalmani, G.; Frisch, M. J. Continuous Surface Charge Polarizable Continuum Models of Solvation. I. General Formalism. *J. Chem. Phys.* **2010**, *132*, 114110-1 – 114110-15.

[95] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; et al. *Gaussian 03, Revision E.01*, Gaussian, Inc., Wallingford CT, 2004.

[96] Pople, J. A.; Head-Gordon, M.; Raghavachari, K. Quadratic Configuration Interaction – a General Technique for Determining Electron Correlation Energies. *J. Chem. Phys.* **1987**, *87*, 5968 – 5975.

[97] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648 − 5652.

[98] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys*. **1988**, *37*, 785 − 789.

[99] Shiell, R.C.; Hu, X.K.; Hu, Q.C.J.; Hepburn, J.W. Threshold Ion-Pair Production Spectroscopy (TIPPS) of H_2 and D_2. *Faraday Discuss. Chem. Soc.* **2000**, *115*, 331 – 343.

[100] Trasatti, S. The Absolute Electrode Potential: An Explanatory Note. *Pure Appl. Chem.* **1986**, *58*, 955 – 966.

[101] Marenich, A. V.; Ho, J.; Coote, M. L.; Cramer, C. J.; Truhlar, D. G. Computational Electrochemistry: Prediction of Liquid-Phase Reduction Potentials. *Phys. Chem. Chem. Phys.* **2014**, *16*, 15068 – 15106.

[102] Isse, A. A.; Gennaro, A. Absolute Potential of the Standard Hydrogen Electrode and the Problem of Interconversion of Potentials in Different Solvents. *J. Phys. Chem. B* **2010,** *114,* 7894 – 7899.

[103] Blondel, C.; Delsart, C.; Goldfarb, F. Electron Spectrometry at the mu eV Level and the Electron Affinities of Si and F. *J. Phys. B: Atom. Mol. Opt. Phys.* **2001**, *34*, L281 – L288.

[104] Woon, D. E.; Dunning, T.H., Jr, Gaussian Basis Sets for Use in Correlated Molecular Calculations. IV. Calculation of Static Electrical Response Properties. *J. Chem. Phys*. **1994**, *100*, 2975 – 2988.

[105] Ruscic, B.; Bross, D. H. Thermochemistry in *Mathematical Modelling of Gas-Phase Complex Reaction Systems: Pyrolysis and Combustion. Computer-Aided Chemical Engineering 45*, Eds. Faravelli, T.; Manenti, F.; Ranzi, E. Elsevier: Amsterdam, 2019, Chapt. 1, pp 3 – 114.

[106] Ruscic, B.; Bross, D. H. Active Thermochemical Tables (ATcT) values based on ver.1.122d of the thermochemical network, Available at https://ATcT.anl.gov/Thermochemical%20Data/version%201.122d, accessed March 31, 2020.

[107] Chase, M. W., Jr. *NIST-JANAF Thermochemical Tables*, 4th ed.; *J. Phys. Chem. Ref. Data, Mono. 9*; American Institute of Physics: Woodbury, NY, 1998; Suppl 1.

[108] Wilmarth, W. K.; Dayton, J. C.; Flournoy, J. M. The Mechanism of Exchange of Hydrogen Gas and Aqueous Alkali. *J. Am. Chem. Soc.* **1953**, *75*, 4549 – 4553.

Full Reference 67 (Gaussian 16): Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision A.03, Gaussian, Inc., Wallingford CT, 2016.

Full Reference 74 (Molpro 2015): Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, T.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.;Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.

MOLPRO, version 2015.1, a package of *ab initio* programs, See http://www.molpro.net. Accessed March 1, 2016.

Full Reference 95 (Gaussian 03): Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. Gaussian 03, Revision E.01, Gaussian, Inc., Wallingford CT, 2004.

Discussion of $(H_2O)_n$ Clusters. The ring-like structure $(H_2O)_4$ A (C_2 symmetry) has the lowest gas phase free energy and has a relative aqueous free energy of 0.4 kcal/mol higher than the lowest energy structure, $(H_2O)_4$ B (C_1 symmetry), which consists of the $(H_2O)_3$ ring with one additional H_2O molecule hydrogen bonded to one hydrogen from a H_2O in the ring. Despite being the lowest aqueous free energy structure, $(H_2O)_4$ B is 2.5 kcal/mol higher than $(H_2O)_4$ A in the gas phase. The geometry of $(H_2O)_4$ C (C_4 symmetry) is very similar to that of $(H_2O)_4$ A. It is also a ring composed of 4 H_2O molecules, but the H_2O molecules lie much more in the plane of the ring than for $(H_2O)_4$ A, which has two hydrogen atoms positioned out of the plane of the ring at opposing corners of the cluster. This difference results in a small increase in aqueous free energy (1.0 kcal/mol relative to the lowest energy $(H_2O)_4$ B) and a more significant increase in gas phase free energy (2.3 kcal/mol relative to lowest energy $(H_2O)_4$ A). Finally, $(H_2O)_4$ D resembles two fused 3-member rings folded at approximately 83°, which results in a higher gas phase and aqueous relative free energy of 3.1 kcal/mol at the CCSD(T)/CBS(DT)-F12b level and 3.4 kcal/mol at the CCSD(T)/CBS(DT)-F12b/COSMO (G16) level.

Table A6.1.1. Gas Phase MP2/aug-cc-pV*n*Z ("a*n*"; *n* = D, T, Q) Total Energies in Hartrees

	MP2/aD			MP2/aT			MP2/aQ
Species	H_{298K} Corr.	G_{298K} Corr.	MP2 Elec. Energy	H_{298K} Corr.	G_{298K} Corr.	MP2 Elec. Energy	MP2 Elec. Energy
H^-	0.002361	-0.009999	-0.511958	0.002361	-0.010000	-0.515911	-0.517094
H	0.002360	-0.010654	-0.499334	0.002360	-0.010654	-0.499821	-0.499948
$(H_2O)_n$							
H_2O	0.025114	0.003676	-76.260910	0.025189	0.003766	-76.328992	-76.351828
$(H_2O)_2$	0.052858	0.019492	-152.530207	0.052999	0.019771	-152.666240	-152.711785
$(H_2O)_3$	0.081114	0.042853	-228.808800	0.081378	0.043400	-229.012934	-229.081177
$(H_2O)_4$ A	0.109246	0.065652	-305.089433	0.109570	0.066204	-305.361535	-305.452283
$(H_2O)_4$ B	0.108832	0.059054	-305.078385	0.109160	0.059610	-305.350416	-305.441289
$(H_2O)_4$ C	0.107887	0.065944	-305.085962	0.108205	0.066037	-305.358135	-305.448872
$(H_2O)_4$ D	0.109121	0.063038	-305.081700	0.109399	0.063614	-305.353484	-305.444277
$(H_2O)_8$	0.221816	0.154747	-610.192941	0.221816	0.154747	-610.735447	-610.916050
$(H_2O)_{12}$ A	0.331461	0.236106	-915.296950				
$(H_2O)_{12}$ B	0.333784	0.238374	-915.299276				
$(H_2O)_{16}$ A	0.334398	0.246280	-915.303969				
$(H_2O)_{16}$ B	0.446283	0.333019	-1220.417719				
$H•(H_2O)_n$							
$H(H_2O)$ A	0.027963	-0.003101	-76.760455	0.028042	-0.002880	-76.829046	-76.851964
$H(H_2O)$ B	0.027969	-0.003107	-76.760428	0.027088	-0.001978	-76.828975	-76.851940
$H(H_2O)_2$ A	0.055722	0.013744	-153.029813	0.055865	0.013939	-153.166382	-153.211999
$H(H_2O)_2$ B	0.055707	0.013204	-153.029828	0.055855	0.014567	-153.166424	-153.212230
$H(H_2O)_3$ A	0.083959	0.038266	-229.308502	0.084231	0.039068	-229.513130	-229.581470
$H(H_2O)_3$ B	0.083966	0.036725	-229.308362	0.084237	0.037743	-229.513014	-229.581333
$H(H_2O)_3$ C	0.083966	0.036531	-229.308370	0.084241	0.037762	-229.513021	-229.581340
$H(H_2O)_4$ A	0.112077	0.058318	-305.589163	0.112410	0.061305	-305.861787	-305.952610
$H(H_2O)_4$ B	0.112101	0.058190	-305.589014	0.112430	0.059866	-305.861631	-305.952722
$H(H_2O)_8$ A	0.224714	0.151656	-610.692772				
$H(H_2O)_8$ B	0.224716	0.152330	-610.692773				

$H^-(H_2O)_n$							
$H^-(H_2O)$	0.027405	0.003111	-76.801958	0.027492	0.003211	-76.874431	-76.898444
$H^-(H_2O)_4$ A	0.110684	0.062216	-305.659349	0.111030	0.062513	-305.936337	-306.028701
$H^-(H_2O)_4$ B	0.112128	0.067406	-305.664478	0.112554	0.068010	-305.941516	-306.033665
$H^-(H_2O)_4$ C	0.112743	0.067879	-305.661700	0.113042	0.067632	-305.938541	-306.030908
$H^-(H_2O)_4$ D	0.112221	0.063164	-305.661570	0.112681	0.064181	-305.938536	-306.030855
$H^-(H_2O)_8$ A	0.225293	0.156635	-610.797576				
$H^-(H_2O)_8$ B	0.224255	0.157494	-610.797462				
$H^-(H_2O)_8$ C	0.223871	0.152206	-610.792151				
$H^-(H_2O)_{12}$ A	0.338072	0.247287	-915.925369				
$H^-(H_2O)_{12}$ B	0.337367	0.250679	-915.928688				
$H^-(H_2O)_{12}$ C	0.337460	0.251096	-915.9283394				
$H^-(H_2O)_{12}$ D	0.337390	0.243077	-915.9121126				
$H^-(H_2O)_{16}$ A	0.451208	0.343902	-1221.053854				
$H^-(H_2O)_{16}$ B	0.451454	0.341159	-1221.054641				

Table A6.1.2. Total Electronic Energies Including Solvation at the MP2/aD Level in Hartrees

Species	COSMO (G16)	SMD (G16)	PCM (G16)	COSMO (G03) Elec Only
$(H_2O)_n$				
H_2O	-76.271643	-76.272329	-76.267525	-76.271621
$(H_2O)_2$	-152.548410	-152.549522	-152.541677	-152.548339
$(H_2O)_3$	-228.825940	-228.827286	-228.818569	-228.825734
$(H_2O)_4$ A	-305.109584	-305.110560	-305.100725	-305.109296
$(H_2O)_4$ B	-305.103142	-305.104794	-305.093229	-305.102803
$(H_2O)_4$ C	-305.108794	-305.109632	-305.099842	-305.108410
$(H_2O)_4$ D	-305.101872	-305.104017	-305.092818	-305.101218
$(H_2O)_8$	-610.233550	-610.235073	-610.220356	-610.232065
$(H_2O)_{12}$ A	-915.357657	-915.360029	-915.357657	-915.354435
$(H_2O)_{12}$ B	-915.335999	-915.340290	-915.317146	-915.332279
$(H_2O)_{16}$ A	-1220.473624	-1220.480191	-1220.448400	-1220.465989
$(H_2O)_{16}$ B	-1220.472343	-1220.475808	-1220.446141	-1220.468612
$H•(H_2O)_n$				
$H(H_2O)$ A	-76.771443	-76.770109	-76.767040	-76.770857
$H(H_2O)$ B	-76.771268	-76.769740	-76.767022	-76.771058
$H(H_2O)_2$ A	-153.048240	-153.047370	-153.041133	-153.047500
$H(H_2O)_2$ B	-153.048010	-153.046993	-153.041081	-153.047188
$H(H_2O)_3$ A	-229.325647	-229.324565	-229.318162	-229.325073
$H(H_2O)_3$ B	-229.325740	-229.325043	-229.318061	-229.324862
$H(H_2O)_3$ C	-229.325749	-229.325044	-229.318065	-229.324940
$H(H_2O)_4$ A	-305.609407	-305.607934	-305.600446	-305.608884
$H(H_2O)_4$ B	-305.609142	-305.608208	-305.599911	-305.608077
$H(H_2O)_8$ A	-610.733398	-610.732496	-610.720226	-610.731894
$H(H_2O)_8$ B	-610.733483	-610.732543	-610.720238	-610.731854
$H^-(H_2O)_n$				
$H^-(H_2O)$	-76.919790	-76.918174	-76.913117	-76.920025
$H^-(H_2O)_4$ A	-305.754314	-305.756353	-305.744676	-305.754314
$H^-(H_2O)_4$ B	-305.758988	-305.760548	-305.750353	-305.758988
$H^-(H_2O)_4$ C	-305.752734	-305.755338	-305.744423	-305.752734
$H^-(H_2O)_4$ D	-305.753811	-305.757039	-305.744576	-305.753059
$H^-(H_2O)_8$ A	-610.888979	-610.892285	-610.876088	-610.887271
$H^-(H_2O)_8$ B	-610.890120	-610.893386	-610.876278	-610.887721
$H^-(H_2O)_8$ C	-610.886303	-610.888719	-610.873334	-610.884709
$H^-(H_2O)_{12}$ A	-916.015393	-916.020876	-915.996261	-916.008551
$H^-(H_2O)_{12}$ B	-916.019017	-916.022799	-916.001967	-916.014608
$H^-(H_2O)_{12}$ C	-916.018228	-916.021915	-916.001967	did not converge
$H^-(H_2O)_{12}$ D	-916.004619	-916.008969	-915.987800	-916.001391
$H^-(H_2O)_{16}$ A	-1221.149463	-1221.152447	-1221.130880	-1221.145943
$H^-(H_2O)_{16}$ B	-1221.144490	-1221.148892	-1221.124090	-1221.138452

Table A6.1.3. Solvation Free Energies at the MP2/aD Level in kcal/mol

Species	COSMO (G03) Elec Only	COSMO (G03) Elec + Non-Elec [a]
$(H_2O)_n$		
H_2O	-6.7	-8.4
$(H_2O)_2$	-11.4	-13.7
$(H_2O)_3$	-10.6	-12.9
$(H_2O)_4$ A	-12.5	-14.8
$(H_2O)_4$ B	-15.3	-18.5
$(H_2O)_4$ C	-14.1	-18.0
$(H_2O)_4$ D	-12.2	-14.8
$(H_2O)_8$	-24.6	-29.6
$(H_2O)_{12}$ A	-20.7	-37.5
$(H_2O)_{12}$ B	-22.2	-24.3
$(H_2O)_{16}$ A	-31.7	-36.6
$(H_2O)_{16}$ B	-30.3	-52.0
$H•(H_2O)_n$		
$H(H_2O)$ A	-6.5	-8.2
$H(H_2O)$ B	-6.7	-8.4
$H(H_2O)_2$ A	-11.1	-13.4
$H(H_2O)_2$ B	-10.9	-13.2
$H(H_2O)_3$ A	-10.4	-12.6
$H(H_2O)_3$ B	-10.4	-12.6
$H(H_2O)_3$ C	-10.4	-12.6
$H(H_2O)_4$ A	-12.4	-14.6
$H(H_2O)_4$ B	-12.0	-14.3
$H(H_2O)_8$ A	-24.5	-29.7
$H(H_2O)_8$ B	-24.5	-29.6
$H^-(H_2O)_n$		
$H^-(H_2O)$	-74.1	-82.8
$H^-(H_2O)_4$ A	-59.3	-63.3
$H^-(H_2O)_4$ B	-59.0	-63.8
$H^-(H_2O)_4$ C	-56.8	-60.9
$H^-(H_2O)_4$ D	-57.4	-61.5
$H^-(H_2O)_8$ A	-56.3	-60.7
$H^-(H_2O)_8$ B	-56.6	-61.1
$H^-(H_2O)_8$ C	-58.1	-63.1
$H^-(H_2O)_{12}$ A	-52.2	-56.5
$H^-(H_2O)_{12}$ B	-53.9	-58.7
$H^-(H_2O)_{12}$ C	-53.9	-58.7
$H^-(H_2O)_{12}$ D	-56.0	-61.2
$H^-(H_2O)_{16}$ A	-57.8	-64.1
$H^-(H_2O)_{16}$ B	-52.6	-57.7

Table A6.1.4. Total Electronic Energies Including Solvation for H• and H^- at Various Levels of Theory in Hartrees

Energy [a]	H•	H^-
MP2/aD (gas)	-0.499334	-0.511958
CCSD/aD (gas)	-0.499334	-0.524029
QCISD/aD (gas)	-0.499334	-0.524029
MP2/aD/COSMO	-0.499480	-0.641485
MP2/aD/SMD	-0.497134	-0.638661
MP2/aD/PCM	-0.499353	-0.636056
CCSD/aD/COSMO	-0.499480	-0.650540
CCSD/aD/SMD	-0.497134	-0.647593
CCSD/aD/PCM	-0.499353	-0.645512
QCISD/aD/COSMO	-0.499480	-0.650540
QCISD/aD/SMD	-0.497134	-0.647593
QCISD/aD/PCM	-0.499353	-0.645511

Table A6.1.5. Solvation Free Energies for H• and H^- at Various Levels of Theory in kcal/mol

Energy [a]	H•	H^-
MP2/aD/COSMO	-0.1	-81.3
MP2/aD/SMD	1.4	-79.5
MP2/aD/PCM	0.0	-77.9
CCSD/aD/COSMO	-0.1	-79.4
CCSD/aD/SMD	1.4	-77.5
CCSD/aD/PCM	0.0	-76.2
QCISD/aD/COSMO	-0.1	-79.4
QCISD/aD/SMD	1.4	-77.5
QCISD/aD/PCM	0.0	-76.2

Table A6.1.6. Aqueous Electron Affinities (kcal/mol) and *E*° (V) for H•/H^- at 298 K

Method [a]	Energy	$H^-(H_2O)$/ $H•(H_2O)$	$H^-(H_2O)_4$/ $H•(H_2O)_4$	$H^-(H_2O)_8$/ $H•(H_2O)_8$		$H^-(H_2O)_{12}$/ $H•(H_2O)_8$		$H^-(H_2O)_{16}$/ $H•(H_2O)_8$	
			Surface	Embedded	Surface	Embedded	Surface	Embedded	Surface
MP2/aD	EA	101.0	100.4	105.2	104.1	100.0	104.2	99.2	97.8
	E°	0.10	0.07	0.28	0.23	0.06	0.24	0.02	-0.04
MP2/aT	EA	101.3	101.5						
	E°	0.11	0.12						
MP2/aQ with MP2/aT corr.	EA	101.3	101.1						
	E°	0.11	0.11						
G3MP2(B3)	EA	100.4	101.7						
	E°	0.07	0.13						
CCSD(T)/aD	EA	98.9	97.7	102.0	101.0	98.7	102.9	99.1	98.1
	E°	0.01	-0.04	0.14	0.10	0.00	0.18	0.02	-0.03
CCSD(T)/aT	EA	99.3	98.6	102.6	101.9				
	E°	0.03	0.00	0.17	0.14				
CCSD(T)/aD-F12b	EA	99.3	99.7	102.7	102.1	100.0	103.8	99.5	98.1
	E°	0.03	0.04	0.17	0.15	0.06	0.22	0.03	-0.03
CCSD(T)/aT-F12b	EA	99.3	98.6	102.6	102.1				
	E°	0.03	0.00	0.17	0.15				
CCSD(T)/ CBS(DT)-F12b	EA	99.3	98.7	102.5	102.1				
	E°	0.03	0.00	0.17	0.15				

[a] All methods except MP2/aD involve calculations using the electronic energy at the level specified in the table with the appropriate thermodynamic corrections calculated at the MP2/aD level. For clusters with 12 and 16 H_2O, the best available value for $\Delta G_{hyd}[H•]$ was used, 4.6 kcal/mol. All solution free energies are calculated at the MP2/aD/COSMO(G16) level.

Table A6.1.7. Anion…H-OH Bond Distances (Å) for F^- and H^- Clustered With 4 and 8 H_2O Molecules

Molecule	Structure	Anion…H-OH Bond Distances (Å)
$H^-(H_2O)_4$	A	1.5224 (x2), 1.7032 (x2)
$H^-(H_2O)_4$	B	1.4105, 1.5332, 1.7240
$H^-(H_2O)_4$	C	1.6991 (x4)
$H^-(H_2O)_4$	D	1.5823, 1.6226, 1.6447, 1.6757
$F^-(H_2O)_4$	A	1.6194 (x2), 1.7309 (x2)
$F^-(H_2O)_4$	B	1.5209, 1.5946, 1.7303
$F^-(H_2O)_4$	C	1.7136 (x4)
$F^-(H_2O)_4$	D	1.6349, 1.6852, 1.7092, 1.7124
$H^-(H_2O)_8$	A	1.4955 (x2), 1.5834 (x2)
$H^-(H_2O)_8$	B	1.5262 (x4)
$H^-(H_2O)_8$	C[a]	1.3068, 1.4142, 1.7336, 1.8273
$F^-(H_2O)_8$	A[b]	1.6338 (x4)
$F^-(H_2O)_8$	B[c]	1.6624 (x2), 1.6656 (x2)

[a] Unique conformation, "embedded." [b] On face of cube, like $H^-(H_2O)_8$ A and B. [c] Unique conformation.

Table A6.1.8. Natural Atomic Charges from Natural Bond Orbitals (NBO) for $H^-(H_2O)_n$ and $F^-(H_2O)_n$ for n = 4, 8 at the B3LYP/aug-cc-pVDZ Level

Species	X^-	O(H)	H(O)
$X^-(H_2O)_4$			
$H^-(H_2O)_4$ A	-0.69	-1.06, -1.06, -1.05, -1.05	0.47, 0.47, 0.48, 0.48 0.49, 0.49, 0.50, 0.50
$H^-(H_2O)_4$ B	-0.64	-1.05, -1.06, -1.07, -1.08	0.46, 0.47, 0.49, 0.49, 0.49, 0.49, 0.50, 0.51
$H^-(H_2O)_4$ C	-0.71	-1.06, -1.06, -1.06, -1.06	0.49, 0.49, 0.49, 0.49, 0.50, 0.50, 0.50, 0.50
$H^-(H_2O)_4$ D	-0.68	-1.04, -1.06, -1.06, -1.07	0.47, 0.48, 0.48, 0.48, 0.49, 0.50, 0.50, 0.50
$F^-(H_2O)_4$ A	-0.82	-1.04, -1.04, -1.04, -1.04	0.46, 0.46, 0.48, 0.48, 0.52, 0.52, 0.53, 0.53
$F^-(H_2O)_4$ B	-0.82	-1.05, -1.05, -1.05, -1.05	0.47, 0.48, 0.49, 0.50, 0.51, 0.52, 0.53, 0.53
$F^-(H_2O)_4$ C	-0.83	-1.05, -1.05, -1.05, -1.05	0.49, 0.49, 0.49, 0.49, 0.52, 0.52, 0.52, 0.52
$F^-(H_2O)_4$ D	-0.83	-1.04, -1.04, -1.05, -1.05	0.46, 0.48, 0.48, 0.48, 0.52, 0.52, 0.52, 0.53
$X^-(H_2O)_8$			
$H^-(H_2O)_8$ A	-0.62	-1.04, -1.04, -1.04, -1.04, -1.04, -1.04, -1.05, -1.05	0.48, 0.48, 0.49, 0.49, 0.49, 0.49, 0.50, 0.50, 0.50, 0.50, 0.51, 0.51, 0.51, 0.51, 0.51, 0.51
$H^-(H_2O)_8$ B	-0.62	-1.04, -1.04, -1.04, -1.04, -1.04, -1.04, -1.04, -1.04	0.48, 0.48, 0.48, 0.48, 0.49, 0.49, 0.49, 0.49, 0.50, 0.50, 0.50, 0.50, 0.51, 0.51, 0.51, 0.51
$H^-(H_2O)_8$ C	-0.58	-1.03, -1.03, -1.03, -1.03, -1.04, -1.04, -1.07, -1.09	0.45, 0.48, 0.48, 0.49, 0.49, 0.49, 0.50, 0.50, 0.50, 0.50, 0.50, 0.51, 0.51, 0.51, 0.51, 0.52
$F^-(H_2O)_8$ A	-0.80	-1.03, -1.03, -1.03, -1.03, -1.04, -1.04, -1.04, -1.04	0.48, 0.48, 0.48, 0.48, 0.51, 0.51, 0.51, 0.51, 0.51, 0.51, 0.51, 0.51, 0.52, 0.52, 0.52, 0.52
$F^-(H_2O)_8$ B	-0.81	-1.02, -1.02, -1.03, -1.03, -1.05, -1.05, -1.05, -1.05	0.47, 0.47, 0.48, 0.48, 0.51, 0.51, 0.51, 0.51, 0.52, 0.52, 0.52, 0.52, 0.53, 0.53, 0.53, 0.53

Table A6.1.9. Energy Values for reaction $OH^- + H_2 \rightarrow H_2O + H^-$ in kJ/mol

	OH^-	**H_2**	**H_2O**	**H^-**	**Source**
$\Delta H_{f,gas}$(0 K)	-139.091	0	-238.931	143.264	ATcT tables[a]
$\Delta H_{f,gas}$(298 K)	-139.058	0	-241.834	145.228	ATcT tables[a]
$H_{gas} \rightarrow G_{gas}$ (298 K)	4.897	0	13.244	-6.750	JANAF
ΔG_{gas}(298 K)	-134.16	0	-228.590	138.47	Sum
ΔG_{solv}(298 K)	-437.2[b]	17.8[c]	-35.15 ± 1.2[d]	-326.4 ± 8.4[e]	

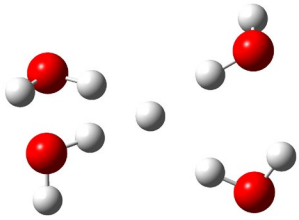
[a] The ATcT errors are ±0.027 kJ/mol. See text for references 81 and 82.
[b] From calculations. See text for reference 36.
[c] From Henry's Law constant. See reference 13.
[d] Hummer, G.; Pratt, L. R.; Garcia. A. E. Hydration Free Energy of Water. *J. Phys. Chem.* **1995**, 99, 14188 – 14194.
[e] This work.

Table A6.1.10. Reaction energetics for $OH^- + H_2 \rightarrow H_2O + H^-$

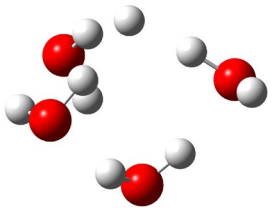
Energy	kJ/mol	kcal/mol
ΔH_{gas}(298 K)	42.45	10.1
ΔG_{gas}(298 K)	44.04	10.5
ΔG_{solv}(298 K)	57.8	13.8
ΔG_{aq}(298 K)	101.8	24.3

Table A6.1.11. Polarizabilities of H^- and F^- in $Bohr^3$

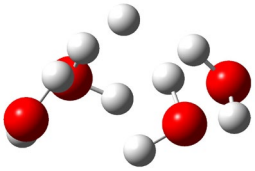
	H^-		F^-	
Basis Set	MP2	CCSD	MP2	CCSD(T)
daV5Z	108.18	157.10	16.29	16.27
daV6Z	110.93	163.66	16.33	16.30
taV5Z	128.54	227.90	16.56	16.66
taV6Z	128.92	229.78	16.54	16.61



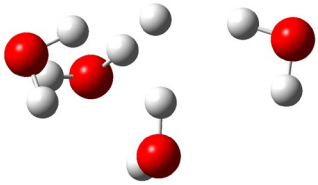
$H^-(H_2O)_4$ A
$\Delta H_{g,298K}$ = 2.3 kcal/mol
$\Delta G_{g,298K}$ = 1.1 kcal/mol
$\Delta G_{aq,298K}$ = 0.0 kcal/mol



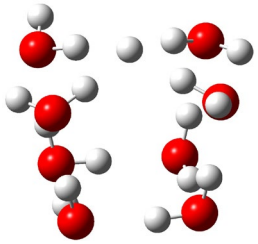
$H^-(H_2O)_4$ B
$\Delta H_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{g,298K}$ = 1.1 kcal/mol
$\Delta G_{aq,298K}$ = 0.3 kcal/mol



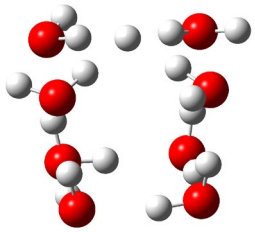
$H^-(H_2O)_4$ C
$\Delta H_{g,298K}$ = 1.7 kcal/mol
$\Delta G_{g,298K}$ = 2.7 kcal/mol
$\Delta G_{aq,298K}$ = 4.1 kcal/mol



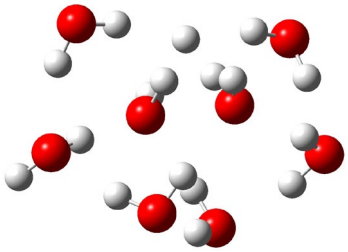
$H^-(H_2O)_4$ D
$\Delta H_{g,298K}$ = 1.6 kcal/mol
$\Delta G_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{aq,298K}$ = 0.7 kcal/mol



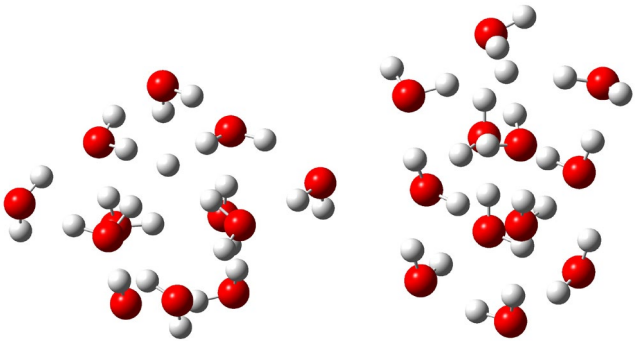
$H^-(H_2O)_8$ A
$\Delta H_{g,298K}$ = 0.4 kcal/mol
$\Delta G_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{aq,298K}$ = 0.6 kcal/mol



$H^-(H_2O)_8$ B
$\Delta H_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{g,298K}$ = 0.8 kcal/mol
$\Delta G_{aq,298K}$ = 0.6 kcal/mol

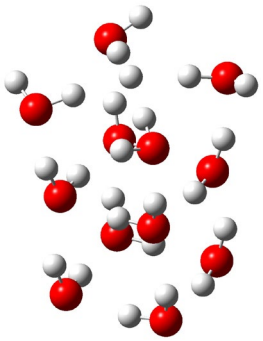


$H^-(H_2O)_8$ C
$\Delta H_{g,298K}$ = 3.4 kcal/mol
$\Delta G_{g,298K}$ = 1.1 kcal/mol
$\Delta G_{aq,298K}$ = 0.0 kcal/mol

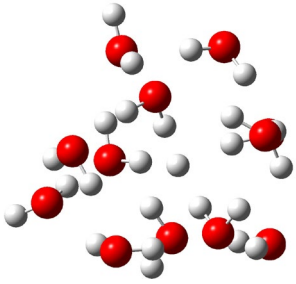


$H^-(H_2O)_{12}$ A
$\Delta H_{g,298K}$ = 2.5 kcal/mol
$\Delta G_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{aq,298K}$ = 2.7 kcal/mol

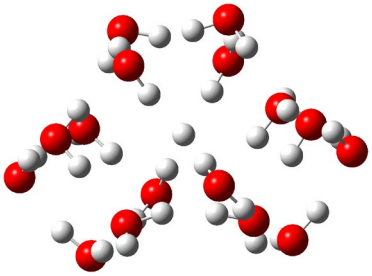
$H^-(H_2O)_{12}$ B
$\Delta H_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{aq,298K}$ = 0.0 kcal/mol



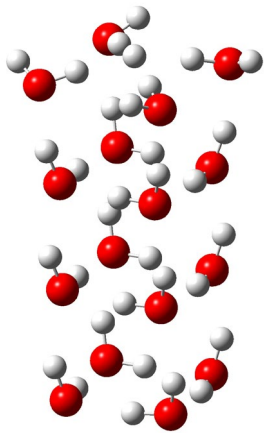
$H^-(H_2O)_{12}$ C
$\Delta H_{g,298K}$ = 0.3 kcal/mol
$\Delta G_{g,298K}$ = 0.5 kcal/mol
$\Delta G_{aq,298K}$ = 0.6 kcal/mol



$H^-(H_2O)_{12}$ D
$\Delta H_{g,298K}$ = 10.4 kcal/mol
$\Delta G_{g,298K}$ = 5.7 kcal/mol
$\Delta G_{aq,298K}$ = 4.3 kcal/mol



$H^-(H_2O)_{16}$ A
$\Delta H_{g,298K}$ = 0.3 kcal/mol
$\Delta G_{g,298K}$ = 2.2 kcal/mol
$\Delta G_{aq,298K}$ = 0.0 kcal/mol



$H^-(H_2O)_{16}$ B
$\Delta H_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{aq,298K}$ = 3.3 kcal/mol

Figure A6.1.1. Optimized $H^-(H_2O)_n$ clusters with n = 4, 8, 12, and 16. Gas phase and aqueous free energy values relative to the lowest energy isomer for a given species are calculated at the CCSD(T)/CBS(DT)-F12b level for n = 4 and 8 and at the MP2/aD level for n = 12 and 16. All solvation contributions are calculated at the MP2/aD/COSMO (G16) level.

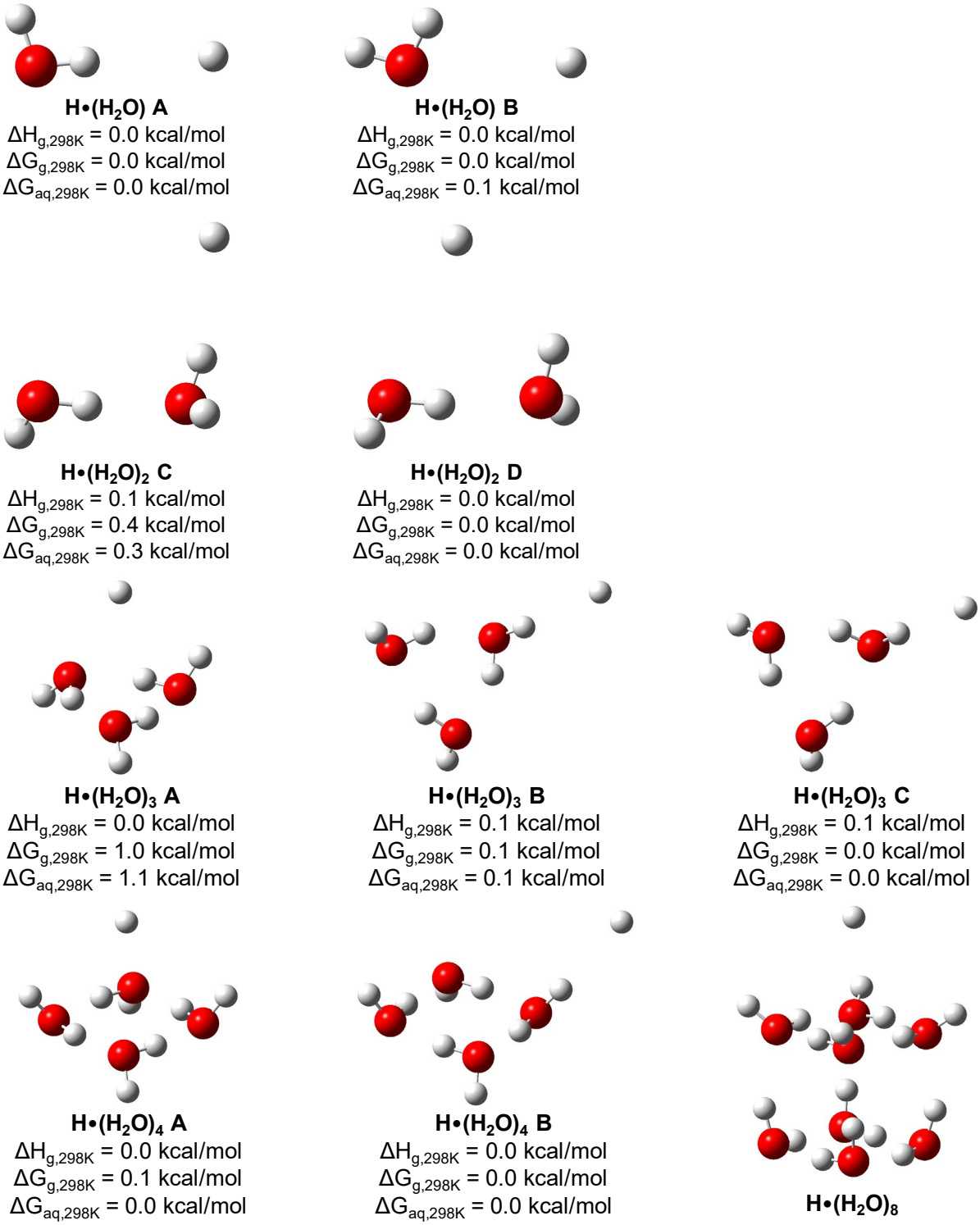
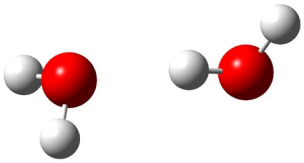
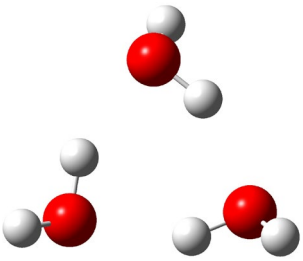


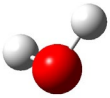
Figure A6.1.2. Optimized H•$(H_2O)_n$ clusters with n = 1 – 4 and 8. Gas phase and aqueous free energy values relative to the lowest energy isomer for a given species are calculated at the CCSD(T)/CBS(DT)-F12b level and solvation contributions are calculated at the MP2/aD/COSMO (G16) level.

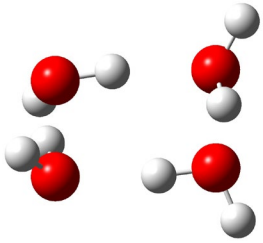


$(H_2O)_2$



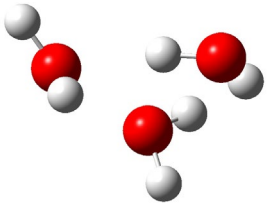
$(H_2O)_3$





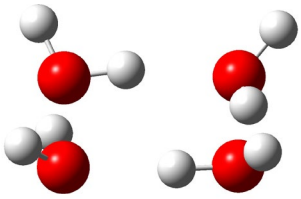
$(H_2O)_4$ A

$\Delta H_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{g,298K}$ = 0.0 kcal/mol
$\Delta G_{aq,298K}$ = 0.4 kcal/mol



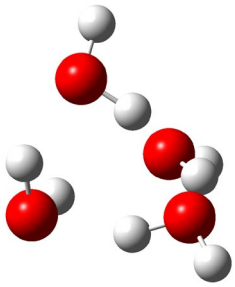
$(H_2O)_4$ B

$\Delta H_{g,298K}$ = 6.4 kcal/mol
$\Delta G_{g,298K}$ = 2.5 kcal/mol
$\Delta G_{aq,298K}$ = 0.0 kcal/mol



$(H_2O)_4$ C

$\Delta H_{g,298K}$ = 1.3 kcal/mol
$\Delta G_{g,298K}$ = 2.3 kcal/mol
$\Delta G_{aq,298K}$ = 1.0 kcal/mol



$(H_2O)_4$ D

$\Delta H_{g,298K}$ = 4.6 kcal/mol
$\Delta G_{g,298K}$ = 3.1 kcal/mol
$\Delta G_{aq,298K}$ = 3.4 kcal/mol

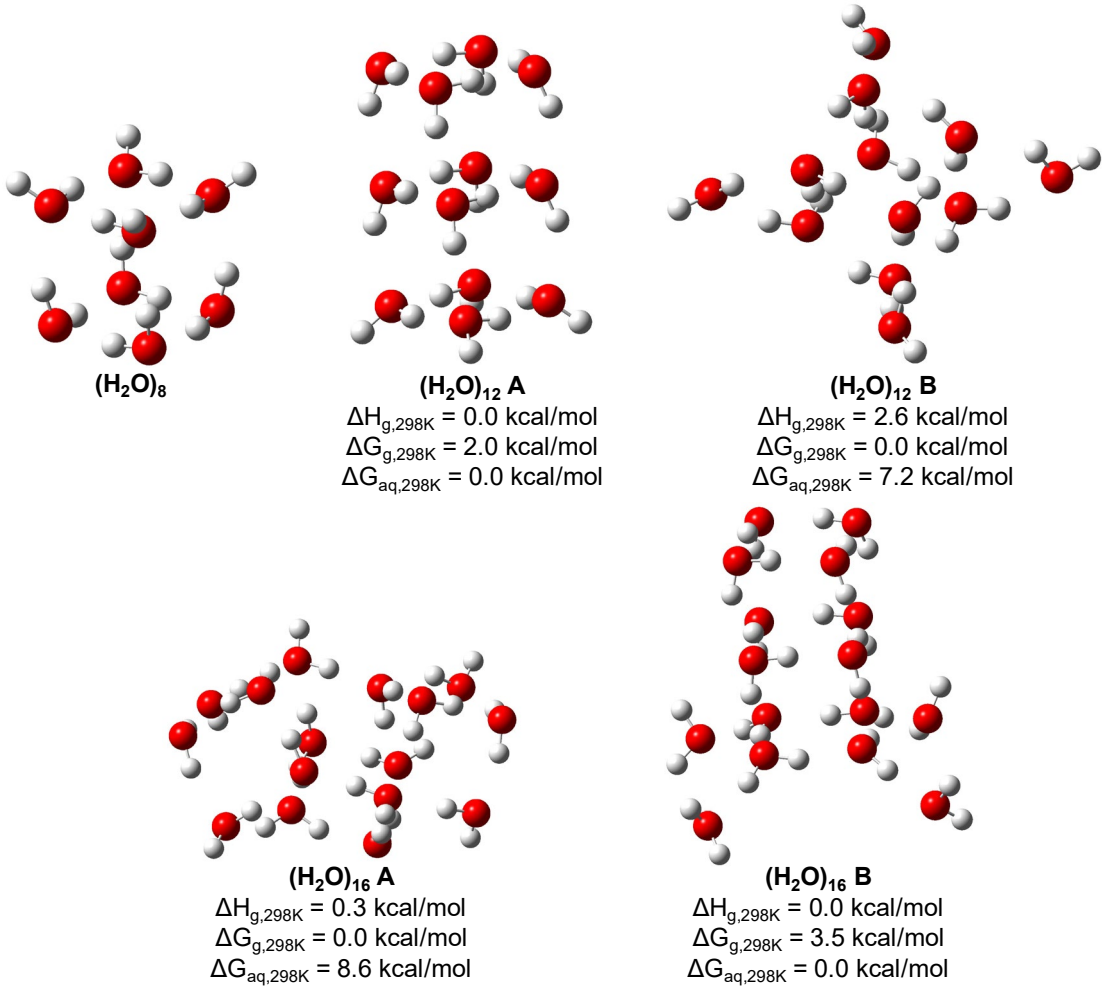


Figure A6.1.3. Optimized $(H_2O)_n$ clusters with n = 2 – 4, 8, 12, and 16. Gas phase enthalpy and free energy values are calculated at the CCSD(T)/CBS(DT)-F12b level for n = 2, 3, 4, and 8 and at the MP2/aD level for n = 12 and 16. All solvation contributions are calculated at the MP2/aD/COSMO (G16) level.

Appendix: Polarizabilities of H^-, Halides, Group I Cations, and the Rare Gases

Computational Methods

Even tempered expansion[1] was to obtain coefficients for modifying aug-cc-pVNZ[2,3,4,5,6,7,8,9] and aug-cc-pVNZ-PP[10] basis sets to be doubly and triply augmented with N = D (2), T (3), Q (4), 5, and 6. These basis sets will collectively be referred to as aVNZ(-PP). This is accomplished by taking the two outermost coefficients for each orbital type (s, p, d, etc.) to get a ratio that can be used to calculate the new coefficient for the next level of augmentation of each orbital type. This can be written as shown in equation (A6.2.1)

$$\alpha_i = \alpha\beta^{i-1} \tag{A6.2.1}$$

where αi represents the new diffuse function, α represents the most diffuse function, and β represents the ratio between the most diffuse function and the second-most diffuse function in the aVNZ(-PP) basis set. Doubly and triply augmented basis sets will be referred to as daVNZ(-PP) and taVNZ(-PP) respectively. This procedure was also carried out on the available all-electron relativistic aug-cc-pVNZ-DK (aVNZ-DK) and aug-cc-pwCVNZ-DK (awCVNZ-DK) basis sets[11,12,13] for each element studied to derive doubly augmented versions of these basis sets (daVNZ-DK and dawCVNZ-DK) with N up to 5.

Polarizability calculations were performed using these basis sets to determine polarizability values for hydride, the halides, the group I cations, and the rare gases as these are all isoelectronic species. Calculations were performed using both Gaussian 16[14] for the MP2[15,16] results listed and Molpro 2018[17,18] for the CCSD(T)[19,20,21,22,23,24,25] results listed.

Results

For hydride and the halide anions, Table A6.2.1 contains polarizabilities calculated using singly, doubly, and triply augmented (aVNZ, daVNZ, and taVNZ) basis sets. The singly

augmented basis sets do not perform well for calculating the polarizabilities of the anions, always underestimating the polarizability somewhat to, for hydride, severely. As shown, the polarizability values for fluoride are fairly consistent between MP2 and CCSD(T) calculations regardless of basis set at the triply augmented level, yielding a predicted polarizability of ca. 2.42 and 2.43 $Å^3$ at the MP2/taV6Z and CCSD(T)/taV6Z levels respectively. For hydride, the values predicted by MP2 and by CCSD(T) are very different, even with the triply augmented basis sets, as MP2 predicts just over half the magnitude of our best calculated polarizability, 33.67 $Å^3$ at the CCSD(T)/taV6Z level. For the other halides, Cl^-, Br^-, I^-, and At^-, the singly augmented basis sets are not in good agreement with our most sophisticated values. There is, however, good agreement between the various doubly or triply augmented basis sets and between MP2 and CCSD(T) results. Our best results predict that the polarizabilities for Cl^-, Br^-, I^-, and At^- are 5.30, 7.11, 10.05, and 12.33 $Å^3$ at the CCSD(T)/taV6Z level for Cl^- and the CCSD(T)/taV5Z-PP level for Br^-, I^-, and At^-. The av6Z-PP basis sets do not exist for bromine and for iodine and are, therefore, not represented in Table A6.2.1.

For hydride, the use of the all-electron aVNZ-DK basis sets makes virtually no difference to the calculated polarizabilities as there are only two electrons in the anion and they are always treated as valence. The same is also true for fluoride, as the number of electrons considered important to calculating the properties of the anion are the same in avNZ, aVNZ-DK, and awCVNZ-DK. All results from a DK basis set with the notable exception of hydride have excellent agreement between MP2 and CCSD(T) methods. Differences for the anions between the doubly augmented all-electron relativistic DK basis sets with and without weighted core, daVNZ-DK vs dawCVNZ-DK, are very nearly exact for H^-, F^-, Cl^-, and Br^- at the

CCSD(T)/daV5Z-DK and CCSD(T)/dawCV5Z-DK levels and small (less than 0.1 $Å^3$) for I^- and At^- at the CCSD(T)/daVQZ-DK and CCSD(T)/dawCVQZ-DK levels.

For the rare gases helium through radon, given in table A6.2.2, there is excellent agreement between the MP2 and CCSD(T) methods even with the smallest basis set employed, aVDZ. Results calculated at the CCSD(T)/aVTZ(-PP) level are in excellent agreement with those calculated at the doubly augmented level and with those calculated using the singly or doubly augmented all-electron basis sets with or without weighted core. In other words, for the rare gas atoms, the level of theory required to obtain excellent agreement with our most sophisticated calculation is only CCSD(T)/aVTZ(-PP) to be within a few hundredths of $Å^3$.

For the group I cations lithium through francium, given in Table A6.2.3, weighted core basis sets were required for accurate calculations due to the stripping of the valence electron for ionization. The polarizabilities of both Li^+ and Na^+ are fully converged at the HF/awCVDZ level, our absolute lowest level of theory used in this work for the group I cations. The HF/awCVDZ polarizability values for K^+, Rb^+, Cs^+, and Fr^+ are within 0.15 $Å^3$ of our best calculated value at the CCSD(T)/awCVQZ-DK level, and agreement improves to near perfection at the HF/awCVTZ-PP level of theory.

Table A6.2.1. Polarizabilities of X^- (X = H and Halogens) in $Å^3$ with Molpro 2018

	H^-		F^-		Cl^-		$Br^{-\,a}$		$I^{-\,a,b}$		$At^{-\,a,b}$	
Basis Set	MP2	CCSD	MP2	CCSD(T)	MP2	CCSD(T)	MP2	CCSD(T)	MP2	CCSD(T)	MP2	CCSD(T)
aVDZ	2.33	2.31	1.00	0.99	3.37	3.42	4.92	5.02	7.48	7.63	9.18	9.36
aVTZ	3.91	4.22	1.32	1.30	4.12	4.17	6.10	6.19	9.00	9.14	11.04	11.21
aVQZ	5.11	5.74	1.59	1.56	4.70	4.70	6.64	6.65	9.55	9.56	11.75	11.77
aV5Z	6.07	7.00	1.76	1.73	4.71	4.70	6.80	6.80	9.63	9.63	11.91	11.91
aV6Z	6.92	8.15	1.91	1.87	4.87	4.85						
daVDZ	14.63	20.70	2.15	2.20	5.30	5.43	7.12	7.32	10.11	10.39	12.25	12.55
daVTZ	14.81	20.79	2.30	2.32	5.41	5.46	7.31	7.41	10.40	10.54	12.61	12.77
daVQZ	15.46	22.03	2.39	2.40	5.44	5.39	7.26	7.23	10.28	10.23	12.44	12.43
daV5Z	16.03	23.28	2.41	2.41	5.41	5.35	7.23	7.18	10.22	10.16	12.37	12.31
daV6Z	16.44	24.25	2.42	2.42	5.40	5.35						
taVDZ	19.18	37.98	2.47	2.56	5.47	5.61	7.32	7.52	10.34	10.62	12.54	12.83
taVTZ	19.15	34.74	2.47	2.53	5.50	5.55	7.36	7.45	10.43	10.57	12.64	12.80
taVQZ	19.03	33.76	2.46	2.48	5.45	5.40	7.27	7.24	10.29	10.24	12.45	12.41
taV5Z	19.05	33.77	2.45	2.47	5.43	5.37	7.24	7.19	10.22	10.16	12.39	12.33
taV6Z	19.10	34.05	2.45	2.46	5.42	5.36						
aVDZ-DK	2.33	2.31	0.99	0.98	3.35	3.40	4.55	4.63				
aVTZ-DK	3.91	4.22	1.31	1.29	4.09	4.14	5.92	6.01	8.82	8.95	10.71	10.87
aVQZ-DK	5.11	5.74	1.58	1.55	4.67	4.67	6.37	6.39	9.45	9.46	11.46	11.49
aV5Z-DK	6.07	7.00	1.75	1.72	4.68	4.67	6.55	6.55				
daVDZ-DK	14.62	20.68	2.14	2.18	5.27	5.40	6.78	6.96				
daVTZ-DK	14.80	20.78	2.28	2.31	5.38	5.43	7.20	7.29	10.25	10.38	12.42	12.57
daVQZ-DK	15.45	22.01	2.38	2.38	5.41	5.36	7.14	6.97	10.15	10.10	12.28	12.24
daV5Z-DK	16.03	23.28	2.40	2.39	5.37	5.32	7.10	7.06				

taVDZ-DK	19.18	37.96	2.45	2.54	5.43	5.57	7.22	7.41				
taVTZ-DK	19.15	34.74	2.45	2.51	5.47	5.52	7.26	7.35	10.28	10.42	12.47	12.62
taVQZ-DK	19.03	33.75	2.44	2.47	5.42	5.37	7.18	7.15	10.16	10.11	12.31	12.26
taV5Z-DK	19.05	33.78	2.44	2.45	5.40	5.34	7.15	7.10				
awCVDZ-DK			0.99	0.98	3.36	3.42	4.60	4.68				
awCVTZ-DK			1.31	1.29	4.10	4.15	5.93	6.01	8.85	8.97	10.74	10.89
awCVQZ-DK			1.58	1.55	4.67	4.67	6.38	6.39	9.45	9.46	11.46	11.49
awCV5Z-DK			1.75	1.72	4.68	4.67	6.55	6.55				
dawCVDZ-DK			2.14	2.18	5.28	5.41	6.75	6.90				
dawCVTZ-DK			2.28	2.30	5.38	5.43	7.15	7.18	10.13	10.14	12.25	12.25
dawCVQZ-DK			2.38	2.38	5.41	5.36	7.12	7.08	10.10	10.04	12.20	12.15
dawCV5Z-DK			2.40	2.39	5.37	5.32	7.09	7.05				

[a] Pseudopotentials used for Br, I, and At. There is no aV6Z-PP basis set for Br, I, and At. [b] There is no aVDZ-DK or aV5Z-DK basis set for I and At.

Table A6.2.2. Polarizabilities of Rare Gases in Å^3 with Molpro 2018

	He		Ne		Ar		Kr [a]		Xe[a,b]		Rn [a,b]	
Basis Set	MP2	CCSD	MP2	CCSD(T)	MP2	CCSD(T)	MP2	CCSD(T)	MP2	CCSD(T)	MP2	CCSD(T)
aVDZ	0.20	0.20	0.29	0.29	1.45	1.46	2.28	2.31	3.75	3.79	4.67	4.71
aVTZ	0.20	0.20	0.36	0.36	1.60	1.61	2.49	2.50	4.05	4.07	5.00	5.03
aVQZ	0.20	0.20	0.39	0.38	1.65	1.64	2.51	2.51	4.07	4.06	5.01	5.01
aV5Z	0.20	0.20	0.40	0.39	1.64	1.64	2.51	2.51	4.07	4.07	5.01	5.01
aV6Z	0.20	0.20	0.40	0.39	1.65	1.64						
daVDZ	0.20	0.21	0.40	0.40	1.62	1.64	2.49	2.52	4.02	4.06	4.96	5.01
daVTZ	0.20	0.21	0.40	0.40	1.66	1.66	2.53	2.54	4.09	4.11	5.04	5.07
daVQZ	0.20	0.21	0.40	0.40	1.65	1.65	2.52	2.52	4.08	4.07	5.02	5.02
daV5Z	0.20	0.20	0.40	0.40	1.65	1.65	2.52	2.51	4.07	4.07	5.02	5.02
daV6Z	0.20	0.20	0.40	0.40	1.65	1.65						
aVDZ-DK	0.20	0.20	0.29	0.29	1.45	1.46	2.17	2.18				
aVTZ-DK	0.20	0.20	0.36	0.36	1.60	1.61	2.48	2.49	4.06	4.08	4.98	5.02
aVQZ-DK	0.20	0.20	0.39	0.39	1.65	1.64	2.50	2.50	4.07	4.07	5.01	5.01
aV5Z-DK	0.20	0.20	0.40	0.39	1.65	1.64	2.50	2.50				
daVDZ-DK	0.20	0.21	0.40	0.40	1.63	1.64	2.49	2.51				
daVTZ-DK	0.20	0.21	0.40	0.40	1.66	1.66	2.53	2.54	4.10	4.12	5.04	5.07
daVQZ-DK	0.20	0.21	0.40	0.40	1.66	1.65	2.52	2.52	4.08	4.07	5.02	5.02
daV5Z-DK	0.20	0.20	0.40	0.40	1.66	1.65	2.52	2.51				
awCVDZ-DK	0.20	0.20	0.29	0.29	1.46	1.47	2.18	2.20				
awCVTZ-DK	0.20	0.20	0.36	0.36	1.61	1.62	2.48	2.49	4.05	4.07	4.98	5.01
awCVQZ-DK	0.20	0.20	0.39	0.39	1.65	1.65	2.50	2.50	4.07	4.06	5.00	5.00

awCV5Z-DK	0.20	0.20	0.40	0.39	1.65	1.64	2.50	2.50				
dawCVDZ-DK	0.20	0.21	0.40	0.40	1.64	1.65	2.48	2.50				
dawCVTZ-DK	0.20	0.21	0.40	0.40	1.67	1.67	2.52	2.52	4.06	4.06	5.00	5.00
dawCVQZ-DK	0.20	0.21	0.40	0.40	1.66	1.65	2.51	2.51	4.06	4.06	5.01	5.01
dawCV5Z-DK	0.20	0.20	0.40	0.40	1.66	1.65	2.51	2.51				

[a] Pseudopotentials used for Kr, Xe, and Rn. There is no aV6Z-PP basis set for Kr, Xe, and Rn. [b] There is no aVDZ-DK or aV5Z-DK basis set for Xe and Rn.

Table A6.2.3. Polarizabilities of Alkali Metal Cations in $Å^3$ (CC = CCSD(T))

	Li^+			Na^+			$K^{+\,a}$			$Rb^{+\,a}$			$Cs^{+\,a}$			$Fr^{+\,a}$		
Basis Set	HF	MP2	CC	HF	MP 2	CC	HF	MP 2	CC	HF	MP 2	CC	HF	MP 2	CC	HF	MP 2	CC
awCVDZ	0.03	0.03	0.03	0.14	0.15	0.15	0.73	0.74	0.74	1.22	1.21	1.21	2.22	2.16	2.46	2.77	2.71	2.71
awCVTZ	0.03	0.03	0.03	0.14	0.15	0.15	0.80	0.81	0.81	1.34	1.37	1.37	2.35	2.37	2.38	2.89	2.92	2.93
awCVQZ	0.03	0.03	0.03	0.14	0.15	0.15	0.80	0.81	0.82	1.35	1.36	1.37	2.35	2.37	2.38	2.90	2.93	2.94
awCV5Z	0.03	0.03	0.03	0.14	0.15	0.15	0.80	0.81	0.81	1.35	1.36	1.36	2.35	2.37	2.37	2.90	2.93	2.92
awCVDZ -DK	0.03	0.03	0.03	0.12	0.12	0.12	0.69	0.69	0.69	1.23	1.24	1.24	2.24	2.25	2.31	2.75	2.77	2.78
awCVTZ-DK	0.03	0.03	0.03	0.11	0.12	0.12	0.80	0.82	0.82	1.34	1.36	1.37	2.33	2.34	2.35	2.90	2.93	2.94
awCVQZ -DK	0.03	0.03	0.03	0.14	0.15	0.15	0.80	0.82	0.82	1.34	1.35	1.36	2.33	2.34	2.35	2.90	2.93	2.94
awCV5Z-DK	0.03	0.03	0.03	0.14	0.15	0.15												

[a] Pseudopotential used for K, Rb, Cs, and Fr. awCV5Z-DK basis set does not exist for these atoms.

References

[1] Ulusoy, I. S.; Wilson, A. K. Slater and Gaussian Basis Functions and Computation of Molecular Integrals. In *Mathematical Physics in Theoretical Chemistry*; Blinder, S. M., House, J. E., Eds.; Elsevier: Cambridge, MA, 2019; pp 45.

[2] Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys*. **1989**, *90*, 1007 – 1023.

[3] Kendall, R. A.; Dunning, T. H. Jr.; Harrison, R. J., Electron Affinities of the First Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys*. **1992,** *96*, 6796 – 6806.

[4] van Mourik, T.; Wilson, A. K.; Dunning, Jr., T. H. Benchmark Calculations with Correlated Molecular Wavefunctions. XIII. Potential Energy Curved for He_2, Ne_2, and Ar_2 Using Correlation Consistent Basis Sets Through Augmented Sextuple Zeta. *Mol. Phys.* **1999**, *99*, 529 – 547.

[5] Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, Jr., T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. IX. The Atoms Gallium Through Krypton. *J. Chem. Phys.* **1999**, *110*, 7667 – 7676.

[6] Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning, Jr., T. H.; Wilson, A. K. Gaussian Basis Sets for Use in Correlated Molecular Calculations. VII. Valence and Core-Valence Basis Sets for Li, Na, Be, and Mg. *Theor. Chem. Acc.* **2011**, *128*, 69 – 82.

[7] Peterson, K. A. Systematically Convergent Basis Sets with Relativistic Pseudopotentials. I. Correlation Consistent Basis Sets for the Post-d Group. 13 – 15 Elements. *J. Chem. Phys.* **2003**, *119*, 11099 – 11112.

[8] Woon, D. E.; Dunning, Jr., T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum Through Argon. *J. Chem. Phys*. **1993**, *98*, 1358 – 1371.

[9] Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically Convergent Basis Sets with Relativistic Pseudopotentials. II. Small-Core Pseudopotentials and Correlation Consistent Basis Sets for the Post-d Group 16 – 18 Elements. *J. Chem. Phys*. **2003**, *119*, 11113 – 11123.

[10] Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically Convergent Basis Sets with Relativistic Pseudopotentials. II. Small-Core Pseudopotentials and Correlation Consistent Basis Sets for the Post-d Group 16 – 18 Elements. *J. Chem. Phys.* **2003**, *119*, 11113 – 11123.

[11] Bross, D. H.; Peterson, K. A. Correlation Consistent, Douglas-Kroll-Hess Relativistic Basis Sets for the 5*p* and 6*p* Elements. *Theor. Chem. Acc.* **2014**, *133*, 1434-1 – 1434-12.

[12] Hill, J. G.; Peterson, K. A. Gaussian Basis Sets for Use in Correlated Molecular Calculations. XI. Pseudopotential-Based and All-Electron Relativistic Basis Sets for Alkali Metal (K – Fr) and Alkaline Earth (Ca – Ra) Elements. *J. Chem. Phys.* **2017**, *147*, 244106-1 – 224106-12.

[13] Lim, I. S.; Schwerdtfeger, P.; Metz, B.; Stoll, H. All-Electron and Relativistic Pseudopotential Studies for the Group 1 Element Polarizabilities from K to Element 119. *J. Chem. Phys*. **2005**, *122*, 104103-1 – 104103-12.

[14] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision A.03, Gaussian, Inc., Wallingford CT, 2016.

[15] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 0618 – 0622.

[16] Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical Models Incorporating Electron Correlation. *Int. J. Quantum. Chem.* **1976**, *10*, 1 – 19.

[17] Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, D.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bennie, S. J.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Lee, S. J. R.; Liu, Y.; Lloyd, A. W.; Ma, Q.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Miller III, T. F.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Welborn, M. MOLPRO, version 2018.2, a package of *ab initio* programs, See http://www.molpro.net.

[18] Werner H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242 – 253.

[19] Purvis, G. D., III; Bartlett, R. J. A Full Coupled Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910 – 1918.

[20] Raghavachari, K.; Truck, G.W.; Pople, J.A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett*. **1989**, *157*, 479 – 483.

[21] Watts, J. D.; Gauss, J.; Bartlett, R. J. Coupled-Cluster Methods with Noniterative Triple Excitations for Restricted Open Shell Hartree-Fock and Other General Single Determinant Reference Functions. Energies and Analytical Gradients. *J. Chem. Phys.* **1993**, *98*, 8718 – 8733.

[22] Bartlett, R. J.; Musial, M. Coupled Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291 – 352.

[23] Knowles, P. J.; Hampel, C.; Werner, H.-J. Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **1993**, *99*, 5219 – 5227.

[24] Knowles, P. J.; Hampel, C.; Werner, H.-J. Erratum: Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **2000,** *112*, 3106 – 3107.

[25] Deegan, M. J. O.; Knowles, P. J. Perturbative Corrections to Account for Triple Excitations in Closed and Open Shell Coupled Cluster Theories. *Chem. Phys. Lett*. **1994**, *227*, 321 – 327.

CHAPTER 7

Absolute Hydration Free Energy of Small Acids and Anions

Introduction

The structural aspects of hydrated ions[1] play an important role in their thermodynamics. For example, based on experimental techniques such as diffraction, hydrated F^- has a well-ordered structure with a preference for embedded geometries. Going down the halogen column, surface geometries are preferred for the halide anions and the ordering of solvent H_2O molecules becomes less well defined. These preferences are consistent with calculations at the molecular orbital theory level of theory for the halide anions with 1 to 6 H_2O molecules with F^- preferring semi-surface structures even with a complete first solvation shell.[2,3] It has been shown using a combination of density functional theory (DFT) and corelated molecular orbital (MO) theory that the a fully solvated F^- requires on the order of 16 explicit H_2O molecules when embedded in a continuum solvent model for a converged free energy of solvation.[4] Full first hydration shells for F^-, Cl^-, and Br^- have been predicted to have 6 H_2O molecules and I^- is expected to have 9 H_2O in its first solvation shell, though the number of solvent molecules in the first shell can fluctuate for the larger halides, particularly for the weakly hydrated I^-.[1,5] Early tabulated values of hydration energies were commonly determined by assigning one ion a specific value and calculating other ions' properties with respect to them, such as the work of Rosseinsky.[6] This is often accomplished by calculating differences in single-ion free energies[7] or by taking advantage of thermodynamic cycles.[8] Kaya and de Farias define absolute ion hydration enthalpies from

hardness and volume-based thermodynamics.[9] The relationships between volume and hardness determined for group 1 cations and group 17 anions, respectively, are

$$V_m^{-1/3} = (9.645\eta^+ + 89.53)/48.2 \quad (7.1)$$

$$V_m^{-1/3} = (64.601\eta^- - 259.64)/214.7 \quad (7.2)$$

Additionally, polarizabilities of the free and hydrated halide cations have been reported at the DFT level.[10]

Other polyatomic anions of interest include hypohalite anions, which have been studied computationally to determine the structure and energetics of their hydrates,[11,12] such as ClO^-, BrO^-, IO^-. Additionally, hydride chalcogen anions OH^- and SH^-.[13,14,15] In this work, we are also interested in determining the properties of SeH^- and TeH^-.

Computational Methods

A hybrid supermolecule-continuum approach was used to estimate the absolute hydration free energy of X^- (X = F, Cl, Br, I, OH, SH, SeH, TeH, OCl, OBr, OI) and the corresponding acids in water. The absolute hydration free energy is the free energy of reaction (7.3), $\Delta G_{hyd}[X^-,n]$, converged to $n \rightarrow \infty$.

$$X^-(gas) + (H_2O)_n(aq) \rightarrow X^-(H_2O)_n(aq) \quad (7.3)$$

The corresponding value for the acid is $\Delta G_{hyd}[XH,n]$ for reaction (7.4).

$$XH(gas) + (H_2O)_n(aq) \rightarrow XH(H_2O)_n(aq) \quad (7.4)$$

In order to calculate this value, we must first determine the Gibbs free energy of the components of this reaction. Gas phase free energies are calculated directly and the aqueous free energies of $(H_2O)_n(aq)$, and $X^-(H_2O)_n(aq)$ and $XH(H_2O)_n(aq)$ are calculated as the sum of their respective gas phase free energy and the bulk solvent shift according to equations (7.5) to (7.7).

$$G[(H_2O)_n(aq)] = G[(H_2O)_n(gas)] + \Delta G_{sol}[(H_2O)_n] \quad (7.5)$$

$$G[X^-(H_2O)_n(aq)] = G[X^-(H_2O)_n(gas)] + \Delta G_{sol}[X^-(H_2O)_n] \quad (7.6)$$

$$G[XH(H_2O)_n(aq)] = G[XH(H_2O)_n(gas)] + \Delta G_{sol}[XH(H_2O)_n] \quad (7.7)$$

The bulk solvent shift from the gas phase cluster to the solution, $\Delta\Delta G_{sol}[X^-,n]$ and $\Delta\Delta G_{sol}[XH,n]$, and the hydration free energy of X^- in the gas phase cluster $X^-(H_2O)_n$ and the corresponding hydration free energy of XH in the gas phase cluster $XH(H_2O)_n$ are calculated according to equations (7.8) to (7.11), respectively.

$$\Delta\Delta G_{sol}[X^-,n] = \Delta G_{sol}[X^-(H_2O)_n] - \Delta G_{sol}[(H_2O)_n] \quad (7.8)$$

$$\Delta\Delta G_{sol}[XH,n] = \Delta G_{sol}[XH(H_2O)_n] - \Delta G_{sol}[(H_2O)_n] \quad (7.9)$$

$$\Delta G_{gas}[X^-,n] = G[X^-(H_2O)_n(gas)] - G[X^-(gas)] - G[(H_2O)_n(gas)] \quad (7.10)$$

$$\Delta G_{gas}[XH,n] = G[XH(H_2O)_n(gas)] - G[XH(gas)] - G[(H_2O)_n(gas)] \quad (7.11)$$

Then, we can evaluate the hydration free energy of X^- and XH using equations (7.12) and (7.13).

$$\Delta G_{hyd}[X^-,n] = \Delta G_{gas}[X^-,n] + \Delta\Delta G_{sol}[X^-,n] \quad (7.12)$$

$$\Delta G_{hyd}[XH,n] = \Delta G_{gas}[XH,n] + \Delta\Delta G_{sol}[XH,n] \quad (7.13)$$

At 298 K, $\Delta G_{hyd}[X^-,n]$ is converged to $\Delta G_{hyd}^{298}(X^-)$ and $\Delta G_{hyd}[XH,n]$ is converged to $\Delta G_{hyd}^{298}(XH)$ when $n \rightarrow \infty$. To predict $\Delta G_{hyd}[X^-,n]$ and $\Delta G_{hyd}[XH,n]$, $\Delta G_{gas}[X^-,n]$, $\Delta G_{gas}[XH,n]$, $\Delta\Delta G_{sol}[XH,n]$ and $\Delta\Delta G_{sol}[X^-,n]$ should be calculated at as high a level of theory as possible.

All geometry optimizations and frequency calculations were performed using the Gaussian 16[16] (G16) software package. Starting geometries for $OH^-(H_2O)_4$ and $F^-(H_2O)_4$ are taken from prior work[4,15] and retain their original labels. All other species were created by modifying these original structures by replacing OH^- with SH^-, SeH^-, TeH^-, OCl^-, OBr^-, and OI^- and by replacing F^- with Cl^-, Br^-, and I^-. The geometries were then optimized at the (DFT)level using the B3LYP[17,18] exchange-correlation functional with the aug-cc-pVDZ (aD) [19,20] basis set. For molecules containing Se, Br, Te, and I, the aug-cc-pVDZ-PP (aD-PP)[21] basis set with an

effective core potential was used for these atoms. The DFT geometries were then used as the initial geometries for further optimization at the correlated MO theory MP2[22,23] level using the aD basis set. Free energies of solvation in water at 298 K were calculated for MP2/aD optimized geometries from single point calculations using the self-consistent reaction field (SCRF) approach[24] with the COSMO[25,26] parameters as implemented in G16 at the MP2/aD(-PP) level.

Single point coupled cluster R/UCCSD(T)[27,28,29,30,31,32,33] or R/UCCSD(T)-F12b[34,35,36] calculations were done at the optimized MP2 geometries to improve the energetics. Thermodynamic values using the CCSD(T) energies include thermodynamic gas phase corrections at the MP2/aD level. The CCSD(T) single point values were calculated using aD and aug-cc-pVTZ (aT) for all species except for when aD(-PP) and aug-cc-pVTZ(-PP)[19,20] basis sets were necessary (Se, Br, Te, and I). The CCSD(T)-F12b calculations were performed using the aug-cc-pVDZ-F12 (aD-F12)[37,38] basis set for all species except those containing Se, Br, Te, and I, for which equivalent basis sets have not been reported. The CCSD(T) calculations were obtained using Molpro 2018.[39,40]

For the FPD calculations of the solvation free energies of all neutral and ionic species, geometries of the di- and triatomic molecules were additionally optimized at the MP2/aT level. Implicit aqueous solvation was also calculated at the MP2/aT(-PP) level using the self-consistent reaction field (SCRF) methods COSMO and SMD.[41] The radii developed by Klamt and co-workers[42,43] were used to define the cavity. The aqueous Gibbs free energy (free energy in aqueous solution), ΔG_{aq}, was calculated from eq. 7.14

$$\Delta G_{aq} = \Delta G_{gas} + \Delta G_{solv} \tag{7.14}$$

where ΔG_{gas} is the gas phase free energy and ΔG_{solv} is the aqueous solvation free energy, both of which are calculated at 298 K. A dielectric constant of 78.3553 corresponding to that of bulk water was used in the solvation calculations.

Beginning with the optimized bond length found at the MP2/aT(-PP) level, bond lengths, harmonic frequencies (ω_e) and anharmonic constants ($\omega_e\chi_e$) for the diatomics were obtained at the coupled cluster singles and doubles with perturbative triples (CCSD(T))[34,35,36] level using aug-cc-pwCV*n*Z-DK (awC*n*-DK, *n* = D, T, Q, 5) basis sets[44,45,46,47] with a 7-point Dunham expansion except for molecules with Te or I as discussed below.[48,49] For triatomic molecules, the MP2/aT optimized geometries were used for single point CCSD(T) calculations using the same basis sets. The scalar relativistic corrections (ΔE_{Rel}) at the second order Douglas-Kroll-Hess (DK)[50,51,52] level. All molecular species containing tellurium or iodine were calculated at the CCSD(T)/awCT-DK and CCSD(T)/awCQ-DK levels due to a lack of awCD-DK and awC5-DK basis sets for Te and I. This limitation requires us to report extrapolations to the complete basis set (CBS) using a variety of calculated value combinations. The extrapolations to the CBS limit were obtained using equation (7.15) with TQ and for those without Te or I with Q5.[53,54]

$$E_n = E_{CBS} + A(n + 1/2)^{-4} \quad (7.15)$$

For all molecules containing Te or I, we did the extrapolations using the pseudopotential and corresponding basis sets for T and Q and then corrected for any errors in the pseudopotential at the Q level using the DK results.

The above CCSD(T) calculations follow the Feller-Peterson-Dixon (FPD) approach.[55,56,57,58] The atomic spin-orbit corrections (ΔE_{SO}) for the heats of formation were taken from the experimental values (C = -0.08 kcal/mol, O = -0.22 kcal/mol, S = -0.56, Cl = -0.84, Se = -2.70, Br = -3.50, Te = -6.02, and I = -7.24).[59] The total atomization energies (TAE, $\Sigma D_{0,0K}$) for

the heats of formation for oxalic acid and the carbonyl radical are calculated using equation (7.16).

$$\Sigma D_{0,0K} = \Delta E_{CBS} + \Delta E_{ZPE} + \Delta E_{SO} \quad (7.16)$$

Heats of formation at 0 K were calculated from these TAEs and experimental heats of formation of the H, C, and O atoms at 0 K from the JANAF tables (51.63 kcal/mol for H, 169.98 kcal/mol for C, 58.98 kcal/mol for O, 65.66 kcal/mol for S, 112.53 kcal/mol for N, 28.59 kcal/mol for Cl, 57.90 kcal/mol for Se, 28.18 kcal/mol for Br, 50.70 kcal/mol for Te, and 25.61 kcal/mol for I).[60] Heats of formation at 298 K were then calculated by following the Curtiss et al. procedure.[61] Gas phase electron affinities were also calculated from the TAEs and are reported as the -ΔTAE of the gas phase electron capture reaction.

Results

Heats of Formation. Geometric parameters for the neutral triatomic species in this work at the MP2/aT and awCQ-DK level are given in Table A7.1.2 and show excellent agreement with experimental values. The optimized geometries of the diatomic anions, both at MP2/aT and from a 7-point fit at the CCSD(T)/awC5-DK level, are given in Table A7.1.3 and are also in excellent agreement with experimental values. The various calculated values used to determine the heats of formation given in Table 7.1 are given in Tables A7.1.4 and A7.1.5. Our calculated heats of formation at 298 K calculated at the CCSD(T)/CBS(Q5) level are in excellent agreement with the Active Thermochemical Tables (ATcT)[62] for HOCl, HOBr, HOI, H_2O, and their deprotonated anions; all species except for HOI and OI^- are within 0.5 kcal/mol of the experimental value. HOI and OI^- are within 1 kcal/mol of the ATcT values, which is also very good agreement. Other comparable experimental values for these species are also reported in Table 7.1.[63,64] Our calculated $\Delta H_{f,298K}(H_2S)$ and $\Delta H_{f,298K}(SH^-)$ are also in excellent agreement

with previously published values,[63,64,65,66] within 0.2 and 1 kcal/mol at the CCSD(T)/CBS(Q5) level respectively. The agreement between our calculated $\Delta H_{f,298K}(H_2Se)$ and $\Delta H_{f,298K}(H_2Te)$ and previously published values[67] is good with the former 1.2 kcal/mol too high and the latter 1.2 kcal/mol too low at the CCSD(T)/CBS(Q5) level.

Gas Phase Acidities. Gas phase acidity (reaction (7.17)) values were calculated at the highest

$$HA \rightarrow A^- + H^+ \qquad (7.17)$$

available level for each molecule (FPD derived CCSD(T)/CBS(Q5) for molecules containing F, Cl, O, and S and CCSD(T)/CBS(TQ) for molecules containing Br, I, Se, and Te) and are given in Table 7.2. The calculated values are all in excellent agreement with literature values, within better than 2 kcal/mol.[68,69,70,71,72,73] The agreement with the highly accurate ATcT values is excellent, within 0.5 kcal/mol. There are apparently larger differences with the values for the chalcogen hydrides but these are not as accurate experimental values with much larger error bars and our values are all within the error bars.

Solution Phase Acidity. As there are differing values for the hydration free energies of the anions in the literature, it can be difficult to compare them with our calculated values. This is due, in part, to the use of a variety of values for the hydration free energy of the proton. For example, Marcus uses a value of-254.5 kcal/mol for $\Delta G_{hyd}[H^+]$ for his calculated values of various ions' hydration free energies.[74] Stanbury uses either-254.3 or -261.0 kcal/mol.[12] In order to compare our calculated values to experiment, we will first compare to the value of $\Delta\Delta G_{aq}$ derived below from the gas phase acidity, the experimental pK_a,[67,69,75,] and the well-established hydration free energy of H^+ of -264.3 kcal/mol[76,77] as shown in Figure 7.1. The free energy of the aqueous deprotonation reaction (7.17) is given by reaction (7.18)

$$\Delta G_{rxn,aq} = G_{gas}(H^+) + G_{hyd}(H^+) + G_{gas}(A^-) + G_{hyd}(A^-) - G_{gas}(HA) - G_{hyd}(HA) \qquad (7.18)$$

where $\Delta G_{rxn,aq}$ is obtained from the pK_a using $\ln K = -\Delta G/RT$. Reaction (7.18) can be rewritten to combine various terms as

$$\Delta G_{rxn,aq} = \Delta G_{rxn,gas} + G_{hyd}(H^+) + \Delta\Delta G_{hyd} \quad (7.19)$$

where

$$\Delta G_{rxn,gas} = G_{gas}(H^+) + G_{gas}(A^-) - G_{gas}(HA) \quad (7.20)$$

and

$$\Delta\Delta G_{hyd} = \Delta G_{hyd}(A^-) - \Delta G_{hyd}(HA) \quad (7.21)$$

The gas phase values for use with Figure 7.1 are obtained at the CCSD(T)/CBS limit using the FPD method and are of high accuracy. In order to calculate $\Delta\Delta G_{hyd}$, values of the hydration energies are needed for HA and A^- and our approach and the appropriate values are described in subsequent sections.

Optimized Anionic $X^-(H_2O)_n$ Geometries. All monatomic anions with water tetramers were optimized with four starting geometries taken from previous work for fluoride.[4] The three $F^-(H_2O)_8$ geometries explored in this work were also taken from prior work.[4] All MP2/aD optimized Cl^- geometries with either four or eight water molecules, based on the previously published hydrated fluoride geometries, are shown in Figure 7.2 and MP2/aD optimized F^-, Br^-, and I^- geometries with four or eight water molecules are shown in Figures A7.1.1, A7.1.3, and A7.1.5 respectively. $Cl^-(H_2O)_4$ **A** has the chloride embedded between the four water molecules and the hydrogen bonds formed are tetrahedrally oriented around the anion. The **A** conformer of the other monatomic anion clusters are similar, though the distance between the pairs of water molecules and the anion increases with halide anion atomic radius. $Cl^-(H_2O)_4$ **B** is geometrically similar to $OH^-(H_2O)_4$ with the chloride interacting with three water molecules and a fourth water molecule beneath the water cluster interacting with two of those water molecules. $Cl^-(H_2O)_4$ **C**

has the chloride hydrogen bonded to a trimer of water molecules below and to a single water molecule above it. For the larger monatomic anions, the **B** and **C** geometries are more similar to one another, with the anion above the water trimer and the additional water to the side of this cluster, hydrogen bonding to both the anion and a water from the trimer. $Cl^-(H_2O)_4$ **D** and the **D** conformers for F^-, Br^-, and I^- have the anion on the face of a tetramer ring of water molecules. The geometry for $Cl^-(H_2O)_8$ **A** is shown in Figure 7.1 and consists of the chloride anion on the face of a cube of H_2O molecules, similar to $OH^-(H_2O)_8$. $Cl^-(H_2O)_8$ **B** has the chloride resting on a puckered rhombus of 8 water molecules. $Cl^-(H_2O)_8$ **C** optimized from a chloride between two tetramer rings of water molecules to the chloride embedded in the surface of a cube of water molecules similar to structure **A**. All relative energies for these isomeric clusters and others are given in Table 7.3.

For $Cl^-(H_2O)_4$, conformation **B** is lowest in free energy in the gas phase and conformation **A** is lowest in free energy in aqueous solution at the CCSD(T)-F12b/aD-F12 and CCSD(T)-F12b/aD-F12/COSMO levels. Conformation **C** is only 0.6 kcal/mol higher than the lowest energy conformer in both environments. Conformations **A** and **D** are *ca*. 1 kcal/mol higher in energy than conformation B in the gas phase. At the CCSD(T)/aD and CCSD(T)/aD/COSMO levels, $Cl^-(H_2O)_8$ **B** is the lowest in gas phase free energy and $Cl^-(H_2O)$ **C** is the lowest in aqueous gas phase free energy with structure **A** less than 2 kcal/mol higher than the most favorable energy in both phases.

For $Br^-(H_2O)_4$, conformation **C** is lowest in the gas phase with conformations B and **D** less than 0.5 kcal/mol higher in energy at the CCSD(T)/aT-PP level. In aqueous solution, $Br^-(H_2O)_4$ **A** is the lowest energy conformer with conformer **B** less than 0.5 kcal/mol higher in aqueous free energy at the CCSD(T)/aT-PP/COSMO level.

As with $Cl^-(H_2O)_8$, $Br^-(H_2O)_8$ **B** is lowest in gas phase free energy and $Br^-(H_2O)_8$ **C** is lowest in aqueous free energy at the CCSD(T)/aT-PP and CCSD(T)/aT-PP/COSMO levels, respectively. Both gas phase $Br^-(H_2O)_8$ **A** and **B** are within 1.5 kcal/mol of the lowest gas phase free energy cluster, **C**. $Br^-(H_2O)_8$ **A** is *ca*. 2 kcal/mol higher in aqueous free energy than the aqueous free energy cluster, **B**, and $Br^-(H_2O)_8$ **C** is 7.0 kcal/mol higher in aqueous free energy.

For $I^-(H_2O)_4$, conformers **C** and **D** have equivalent lowest gas phase free energy with conformer **B** less than 0.5 kcal/mol higher gas phase free energy at the CCSD(T)/aT-PP level. $I^-(H_2O)_4$ **A** and **C** have equivalent lowest aqueous free energy. As with $Cl^-(H_2O)_8$, $I^-(H_2O)_8$ **B** is lowest in gas phase free energy and $I^-(H_2O)_8$ **C** is lowest in aqueous free energy at the CCSD(T)/aT-PP and CCSD(T)/aT-PP/COSMO levels, respectively. Both gas phase $I^-(H_2O)_8$ **A** and **B** are within 2 to 3 kcal/mol of the lowest gas phase free energy cluster, **C**. $Br^-(H_2O)_8$ **A** is 2.6 kcal/mol higher in aqueous free energy than the aqueous free energy cluster, **B**, and $Br^-(H_2O)_8$ **C** is 5.4 kcal/mol higher in aqueous free energy.

Thus, the embedded **A** conformer for all monatomic anions with four water molecules is consistently lowest in aqueous free energy. For Cl^-, Br^-, and I^-, tetramer water cluster conformers **B** and **C**, which are geometrically very similar, are consistently the lowest gas phase free energy clusters. This is not true for gas phase $F^-(H_2O)_4$, for which conformer **A** is also preferred in the gas phase. For the clusters with an octamer of water molecules, conformer **B** is preferred in the gas phase and **C** is preferred in aqueous solution for Cl^-, Br^-, and I^-. $F^-(H_2O)_8$ is the only halide species that prefers conformation **A** in both the gas phase and in aqueous solution.

The starting geometries for $OH^-(H_2O)_4$ and $OH^-(H_2O)_8$ were taken from prior work.[15] The MP2/aD optimized $OH^-(H_2O)_4$ geometry was modified by changing the OH^- to SH^-, SeH^-, and TeH^- and optimized as described above. $OCl^-(H_2O)_4$ geometries were also created by modifying

the $OH^-(H_2O)_4$ MP2/aD optimized geometry with either the oxygen ($OCl^-(H_2O)_4$ **A**) or the chlorine ($OCl^-(H_2O)_4$ **B**) oriented toward the water cluster. The optimized geometries for all of these hydride chalcogen species with a water tetramer are shown in Figure 7.3. All of these structures maintained the same basic geometry for the water tetramer, with three water molecules hydrogen bonding to the chalcogen and a fourth water beneath hydrogen bonded to two of the water molecules. This is true even if it required fully rotating the anionic dimer to move the chalcogen closer to the water tetramer, as is the case for $OCl^-(H_2O)_4$ **B**. The hydrogen bonding between OH^- and the H_2O cluster is strong and results in a very directed hydrogen bond toward the center of the three interacting water molecules. For SH^-, SeH^-, and TeH^-, this hydrogen bonding is progressively weakened and results in a loss of directionality to the interaction, optimizing instead with the hydrogen of the anion leaning toward the interacting water molecules rather than directly away from the water cluster. For $OCl^-(H_2O)_4$, $OBr^-(H_2O)_4$, and $OI^-(H_2O)_4$ there is very little difference in the gas phase or aqueous free energies of the perfectly erect geometry (**A**) and the optimized geometry with the diatomic anion leaning toward the three interacting water molecules (**B**) by 2 to 3 kcal/mol in the gas phase and by less than 1 kcal/mol in aqueous solution.

These anions were all additionally optimized with 8 explicit H_2O molecules, shown in Figure 7.4. For the halide chalcogen anions, increasing size of the chalcogen pushes the water molecules in the top layer of the cluster away. $SH^-(H_2O)_8$ and $SeH^-(H_2O)_8$ remain perpendicular to the surface of H_2O molecules, but $TeH^-(H_2O)_8$ optimizes with the anion tilted to one side. The hypohalite anion clusters of geometry type **A** maintain a similar H_2O cluster shape to that of $OH^-(H_2O)_8$ with the halogen pointed away from the H_2O octet. These anionic clusters do not vary much as the side of the halide increases due to the halide being away from the cluster of H_2O

molecules. The hypohalite anions, OCl^-, OBr^-, and OI^-, with H_2O octamer were also optimized sandwiched between two water tetramers and are labeled as geometry **B**. The disruption of the strong hydrogen bonds between the octamer of water molecules is not compensated for by the improved solvation of both ends of the OX^- anions, however, therefore these geometries are more than 10 kcal/mol higher in both gas phase and aqueous free energy compared to geometry **A**.

Optimized Neutral $XH(H_2O)_n$ Geometries. Geometries for $HCl(H_2O)_4$ were created by adding a proton to the chloride of the $Cl^-(H_2O)_4$ geometries and optimizing as described above, shown in Figure 7.5 with isomer energies in Table 7.4. $HCl(H_2O)_4$ **A** was created by positioning HCl between two dimers of water and optimized to the HCl bound to only one H_2O molecule to the side of the cluster of water molecules. $HCl(H_2O)_4$ **B** and **B2** were created by placing the HCl above (with chloride near the dangling hydrogens) and below (with hydrogen near the plane of oxygen atoms) a tetramer ring of water molecules. $HCl(H_2O)_4$ **B** maintains its centered position on the ring of water molecules whereas $HCl(H_2O)_4$ **B2** shifts to one member of the water ring to form a hydrogen bond with a single water molecule. The large free energy difference between the **B** and **B2** conformers, particularly in the gas phase, is due to the fact that the oxygen of the water molecule (**B2**) is a much better hydrogen bond acceptor than the chlorine of HCl (**B**). $HCl(H_2O)_4$ **C**, created from $Cl^-(H_2O)_4$ **C**, optimizes to form a 5 membered pentagon from the four water molecules and the HCl. $HCl(H_2O)_4$ **B2** and $HCl(H_2O)_4$ **C** are the lowest energy structures in both the gas phase (equivalent energy) and in aqueous solution (**B2** is 0.5 kcal/mol higher than **C**) at the CCSD(T)-F12b/aD-F12 and CCSD(T)-F12b/aD-F12/COSMO levels. $HCl(H_2O)_4$ **D** has a much less ordered geometry similar in conformation and gas phase and aqueous free energy to $HCl(H_2O)_4$ **A**.

The five HCl structures were used as initial geometries to optimize comparable clusters for HF, HBr, and HI. The $HF(H_2O)_4$ **B2** geometry optimized to a geometry virtually identical to $HF(H_2O)_4$ **C**. The fluorine in the $HF(H_2O)_4$ clusters is much more electronegative than chlorine or the other halogens, resulting in stronger hydrogen bonds that incorporate the HF more fully into the cluster geometry. As with the tetramer water cluster with HCl, $HF(H_2O)_4$ **C** is the lowest in gas phase and aqueous free energy at 298 K followed by **B**, **D**, and **A**, respectively. The highest energy structure, $HF(H_2O)_4$ **A**, is very similar energetically to $HF(H_2O)_4$ **D** due to their similar structure, and is *ca*. 5 kcal/mol higher in energy than the lowest energy cluster at the CCSD(T)-F12b/aD-F12 and CCSD(T)-F12b/aD-F12/COSMO levels. The geometries for the five isomeric clusters of $HBr(H_2O)_4$ and $HI(H_2O)_4$ are shown in the Supporting Information.

For $HBr(H_2O)_4$ the cluster with lowest free energy at 298 K in the gas phase is $HBr(H_2O)_4$ **B2** and in aqueous solution is $HBr(H_2O)_4$ **C**, calculated at the CCSD(T)/aT-PP and CCSD(T)/aT-PP/COSMO levels. Each of these structures are within 1 kcal/mol of one another in the gas and aqueous phases. $HBr(H_2O)_4$ **B**, **D**, and **A** ascend in free energy in the gas and aqueous phases with the highest energy cluster, $HBr(H_2O)_4$ **A**, 3.2 kcal/mol higher than the lowest energy gas phase and aqueous solution structures respectively at the CCSD(T)/aT-PP and CCSD(T)/aT-PP/COSMO levels.

Similarly, for $HI(H_2O)_4$ the cluster with the lowest free energy at 298 K in the gas phase is $HI(H_2O)_4$ **B2** and in aqueous solution is $HI(H_2O)_4$ **C** at the CCSD(T)/aD-PP and CCSD(T)/aD-PP/COSMO levels. Notably, the gas phase free energy for $HI(H_2O)_4$ **B** is only 0.4 kcal/mol higher than that of $HI(H_2O)_4$ **B2**, indicating that there is a decreased preference for the orientation of the halogenic acid with respect to the water tetramer for HI. $HI(H_2O)_4$ **C** is an ion pair structure and it 4.6 kcal/mol higher in energy than the most stable hydrogen bonded

structure in the gas phase. Thus, four H_2O molecules are not enough for ion pair formation in the gas phase, consistent with observations on the reactivity of HI with NH_3 clusters where NH_3 is a stronger base.[78] In the presence of an implicit solvent field, $HI(H_2O)_4$ **C** is preferred in the aqueous phase by more than 5 kcal/mol. Thus, the difference in energy for the preferred gas phase and aqueous clusters is larger for $HI(H_2O)_4$ than for the other halogenic acids, due to the transfer of a proton in $HI(H_2O)_4$ **C**. The fact that an HI structure is not predicted to be stable in aqueous solution even in the presence of only 4 explicit H_2O molecules will impact our estimate of the pK_a of aqueous HI. The lowest energy aqueous cluster that maintains the H-I bond is $HI(H_2O)_4$ **B** or **B2**. The aqueous free energies for $HI(H_2O)_4$ **B** and **B2** are very similar energetically, differing by only 0.1 kcal/mol in favor of **B2** at the CCSD(T)/aT-PP/COSMO level. The free energy differences between the isomeric clusters of $HI(H_2O)_4$ are greater than that of the other halogenic acids, with $HI(H_2O)_4$ **A** and $HI(H_2O)_4$ **D** 3.6 kcal/mol higher than the lowest free energy gas phase cluster and *ca.* 8 kcal/mol higher than the lowest free energy aqueous cluster. Unlike $HCl(H_2O)_4$ **B** and $HBr(H_2O)_4$ **B**, $HF(H_2O)_4$ **B** and $HI(H_2O)_4$ **B** do not have the acid centered above the tetrameric ring of water molecules. In the case of $HF(H_2O)_4$ **B**, this is because the HF molecule rotated during optimization to orient the hydrogen toward the ring of water molecules rather than the fluorine, therefore actually resembling the **B2** structure for the other halogenic acids. This is not the case for $HI(H_2O)_4$ **B**, where optimization of the initially symmetric structure simply results in HI shifting to one side of the water tetramer.

Tetrameric clusters resembling the **C** geometries for the halogenic acids were optimized for H_2S, H_2Se, H_2Te, HOCl, HOBr, and HOI, shown in Figure 7.6. H_2S, H_2Se, and H_2Te optimize to very similar structure with the only major difference accounted for by the size of the chalcogen atom. HOCl and HOBr are also very similar to one another upon optimization.

However, HOI optimizes with a visibly different orientation of the halide atom to the hydrogen bonded ring. The angle between the halogen atom, its directly bonded oxygen atom, and the hydrogen atom opposite the acid in the five membered ring is 127.5° for HOCl and 125.4° for HOBr, but is much more acute for HOI at 99.3°, positioning the large iodine atom over the plane of the ring and slightly changing the orientation of the nearby water molecules.

Hydration of Anions. The previously calculated hydration free energy values for F^- and OH^- are -104.3 and -104.5 kcal/mol and were converged with about 16 explicit H_2O molecules.[4,15] Absolute hydration free energy ($\Delta G_{hyd}[X^-,4]$, Table 7.5) values for F^- and OH^- with 4 H_2O molecules are comparable at the CCSD(T)-F12b/aD-F12/COSMO level with $\Delta G_{hyd}[OH^-,4]$ *ca*. 5 kcal/mol higher in free energy than the most favorable value for $\Delta G_{hyd}[F^-,4]$ = -98.4 kcal/mol but mid-range compared to the four structures for $F^-(H_2O)_4$ considered (least favorable $\Delta G_{hyd}[F^-,4]$ = -90.9 kcal/mol). Absolute hydration free energy values are also calculated for $F^-(H_2O)_8$ and $OH^-(H_2O)_8$, giving -101.0 and -101.4 kcal/mol at the CCSD(T)/aD/COSMO level, respectively. These latter values are only slightly smaller than the previously reported converged values using larger clusters.

For the halide anions, $\Delta G_{hyd}[X^-]$ values decreases in magnitude with increasing anionic radius. The best free energy of hydration for $Cl^-(H_2O)_4$ is -67.8 kcal/mol for $Cl^-(H_2O)_4$ **C** at the CCSD(T)-F12b/aD-F12/COSMO level. Increasing the number of explicit water molecules to 8 yields a $\Delta G_{hyd}[Cl^-,8]$ = -73.2 kcal/mol for $Cl^-(H_2O)_8$ **B**, which is in good agreement with the experimental hydration free energy value of -72.7 kcal/mol.[79] Thus, the hydration free energy of chloride is fairly well converged with 8 explicit H_2O molecules. The hydration free energy for bromine with 4 and 8 H_2O molecules increases from -62.1 kcal/mol for $Br^-(H_2O)_4$ **A** at the CCSD(T)/aT-PP/COSMO level to -65.8 kcal/mol for $Br^-(H_2O)_8$ **B** at the CCSD(T)/aD-

PP/COSMO level. As with chloride, the bromide hydration free energy value with 8 H_2O molecules is reasonably well converged to the experimental value of -66.3 kcal/mol.[79] Iodine has a calculated hydration free energy of -56.0 kcal/mol for both $I^-(H_2O)_4$ **A** and **D** at the CCSD(T)/aT-PP/COSMO level and -59.0 kcal/mol for $I^-(H_2O)_8$ **B** at the CCSD(T)/aD-PP/COSMO level, all of which are in reasonable agreement with the experimental value of -57.4 kcal/mol.[79]

As observed for the halides, the $\Delta G_{hyd}[X^-,4]$ values decrease in magnitude for the chalcogen hydrides as the size of the anions increases. The hydration free energy of $SH^-(H_2O)_4$ is -62.8 kcal/mol at the CCSD(T)-F12b/aD-F12/COSMO level, which is only slightly improved when the number of explicit H_2O molecules is increased to 8 (-66.8 kcal/mol). As noted above, comparison with previously reported hydration energy values for these anions and for the hypohalite anions is difficult to determine due to a difference in the value of $\Delta G_{hyd}[H^+]$ used to derive the anionic hydration free energies. Again, these values are reasonably well converged with only 8 explicit H_2O molecules. The hydration free energies of $SeH^-(H_2O)_4$ (-58.4 kcal/mol) and $TeH^-(H_2O)_4$ (-51.7 kcal/mol) at the CCSD(T)/aT-PP/COSMO level, similar to those of the SH^- clusters, are only slightly improved when the number of explicit H_2O molecules are increased to 8 (-59.8 and -55.2 kcal/mol respectively at the CCSD(T)/aD-PP/COSMO level). $\Delta G_{hyd}[SH^-,8]$ is *ca*. 5 kcal/mol higher in energy than $\Delta G_{hyd}[Cl^-,8]$ at the CCSD(T)/aD/COSMO level. $\Delta G_{hyd}[SeH^-,8]$ is *ca*. 6 kcal/mol higher in energy than $\Delta G_{hyd}[Br^-,8]$ at the CCSD(T)/aD-PP/COSMO level. $\Delta G_{hyd}[TeH^-,8]$ is *ca*. 4 kcal/mol higher in energy than $\Delta G_{hyd}[I^-,8]$ at the CCSD(T)/aD-PP/COSMO level. For pairs other than F^- and OH^-, the chalcogen-based anion is less favorable in energy than even the least favorable geometry of the halide.

The values for $\Delta G_{hyd}[OCl^-,4]$ at the CCSD(T)-F12b/aD-F12/COSMO level with attachment to the water molecules through the oxygen atom are -72 to -71 kcal/mol and $\Delta G_{hyd}[OCl^-,8]$ is -83.3 kcal/mol at the CCSD(T)/aD/COSMO level. Attempts at optimizing the cluster attachment to the water tetramer through the chlorine resulted in rotation of the diatomic to preferentially bind through the oxygen. The optimized results for $OBr^-(H_2O)_4$ and $OI^-(H_2O)_4$ are strikingly similar to those of $OCl^-(H_2O)_4$, both in geometry and in hydration free energy. The hydration free energies for $OBr^-(H_2O)_4$ and $OI^-(H_2O)_4$ are -72 to -71 kcal/mol and -70 to -69 kcal/mol, respectively, at the CCSD(T)/aT-PP/COSMO level. These values drastically improve with an explicit H_2O octamer to -80.0 and -81.4 kcal/mol respectively at the CCSD(T)/aD/COSMO level.

Hydration of Acids. Absolute hydration free energies for the protonated (neutral) form of the anionic species presented in Table 7.5 are given in Table 7.6, all of which are modeled with a water tetramer. For the most favorable geometries of the halogenic acids with a water tetramer, the absolute hydration free energy increases as the size of the halide increases from HF (-4.3 kcal/mol

at CCSD(T)-F12b/aD-F12/COSMO), to HCl (1.2 kcal/mol at CCSD(T)-F12b/aD-F12/COSMO), to HBr (2.1 kcal/mol at CCSD(T)/aT/COSMO), to HI (-2.9 kcal/mol for ion pair cluster $HI^-(H_2O)_4$ **C**, and 4.1 or 4.2 kcal/mol for **B** or **B2** respectively at the CCSD(T)/aT/COSMO level.

There is a similar trend for the chalcogens starting from the water tetramer with H_2O (-6.0 kcal/mol at CCSD(T)-F12b/aD-F12/COSMO), to H_2S (4.9 kcal/mol at CCSD(T)-F12b/aD-F12/COSMO), to H_2Se (5.9 kcal/mol at CCSD(T)/aT/COSMO), to H_2Te (6.2 kcal/mol at CCSD(T)/aT/COSMO). To correctly calculate $\Delta G_{hyd}[H_2O,4]$, a correction for the change in standard state must be taken into account. The experimental Gibbs free energies of H_2O(gas) and

H_2O(liq) yield $\Delta G[H_2O(gas) \rightarrow H_2O(liq)]$ = -2.05 kcal/mol at 298 K.[63] This value is consistent with the typically used experimental hydration free energy of -6.32 kcal/mol when we account for the change in state. The entropy change of the expansion process to go from the density of liquid water to that of the ideal gas (1 atm) results in a change to the free energy of -4.3 kcal/mol.[80] For each cluster of H_2O molecules in an aqueous reaction, a correction of 4.3/n kcal/mol needs to be applied to the calculation of the free energy with a negative sign if it is a reactant and positive if it is a product.[80,81,82,83,84,85]

$$H_2O + (H_2O)_4 \rightarrow (H_2O)_5 \quad (7.22)$$

Thus, our calculated value for $\Delta G_{hyd}[H_2O,4]$ must have a correction term of (4.3/5) – (4.3) – (4.3/4) kcal/mol as reaction (7.22) is written. The correction terms for the pentamer (0.9 kcal/mol) and the tetramer (1.1 kcal/mol) of H_2O molecules are small and roughly cancel one another, so we can simply apply the correction term for the single molecule, -4.3 kcal/mol, to arrive at -6.0 kcal/mol at the CCSD(T)-F12b/aD-F12/COSMO level.

HOCl has an absolute free energy of hydration intermediate between HF and HCl at -1.3 kcal/mol at the CCSD(T)-F12b/aD-F12/COSMO level. Unlike the above trends, the free energy of hydration becomes more negative as the size of the halogen increases, yielding an absolute free energy of hydration of -2.0 kcal/mol for HOBr and of -2.7 kcal/mol for HOI, both at the CCSD(T)/aT-PP/COSMO level.

Calculation of $\Delta\Delta G_{aq}$ and pK_a. The above values can be used to calculate the value of $\Delta\Delta G_{aq}$. Given the difficulty in predicting converged solvation energies as a function of cluster size in an implicit self-consistent reaction field, there is good agreement for most neutral/anion pairs. For HX with X = F, Cl, and Br, the agreement between the calculated and experimental values is within 2 kcal/mol. We note that this is actually very good agreement given the very negative pK_a

for HCl and HBr. For HI, better agreement with experiment is found when using the ion pair structure for acid which suggests that the calculated solvation energy of undissociated HI is too positive, again given the limitation of measuring the acidity of HI in water.

For the chalcogen hydrides, the calculated value for H_2O, H_2S and H_2Se is in good agreement with the experimental $\Delta\Delta G_{aq}$ within *ca*. 1 kcal/mol. There is an error of *ca.* -4 kcal/mol for H_2Te suggesting that there is some combination of the calculation of the free energy of solvation of H_2Te being too positive and that of the anion being too negative.

The values of the HOX acids are in reasonable agreement for X = Cl and Br but again there is a larger disagreement with the value for one with the heaviest element in the group, HOI. The calculations seem to be overestimating the stability of OI^- in water, given that the solvation energies of HOI and OI^- have the same sign and that for HOI is not unreasonable. This may be an issue with the fact that not many solvation energies have been studied for the n = 5 period.

In addition to agreement with experimental $\Delta\Delta G_{aq}$, we additionally compare our calculated pK_a values with those experimentally determined and generally find good agreement. Our pKa value for HF is less than 1 pK_a unit higher than the literature pK_a value. The other hydrohalide acids, however, are 2 to 4 pK_a units lower than the experimental pK_a's, although this may be attributed to difficulty in determining these values experimentally as the pK_a's for these acids are very negative. Our calculated pK_a for H_2O, H_2S, and H_2Se are all within 1 pK_a unit too high compared to experimental values. H_2Te, on the other hand is calculated to be *ca.* 2 pK_a units lower than the literature value for this species. Finally, the calculated pK_a's of the hypohalite acids are systematically too low by 2.7, 0.7, and 4.3 pK_a units for HOCl, HOBr, and HI, respectively.

Conclusion

The heats of formation of a range of neutral acids and their corresponding anions after proton loss were predicted using the FPD method with the aug-cc-pwCVNZ-DK basis sets and are reported at the CCSD(T)/CBS(TQ) or CCSD(T)/CBS(Q5) level, where possible. These values are in very good agreement with experimental values with differences of 1.2 kcal/mol or less. Gas phase acidities were calculated from these FPD results as well and are in excellent agreement (within 1 kcal/mol) with values from the highly accurate ATcT values,[62] and within 2 kcal/mol of the other less accurate experimental values not in the ATcT. Our calculated aqueous acidities yield pKa values that are in good agreement with experiment, even though the aqueous free energy values of individual anions are not always in good agreement with literature values. Due to differing choices for the hydration free energy of the proton, previously published aqueous acidities can be significantly different from our calculated results. In order to compare with the literature values in another way, we calculated $\Delta\Delta G_{aq} = \Delta G_{aq}(A^-) - \Delta G_{aq}(HA)$ and found that, with only a few exceptions, the agreement with experimental $\Delta\Delta G_{aq}$ is within 2 kcal/mol. The ΔG_{hyd} values used to determine ΔG_{aq} were calculated via a supermolecule-continuum approach and were reported with 4 explicit H_2O molecules for the neutral acid species and with 4 and 8 H_2O molecules for the anions. The neutral clusters converge quickly to a reasonable value. Most of the anions with 8 explicit H_2O molecules are in good agreement with experiment, although additional explicit H_2O molecules may still be needed to be converged to within 1 kcal/mol of experiment.

Tables

Table 7.1. FPD and Literature Results for Small Triatomic Molecules and Diatomic Anions in kcal/mol

	$\sum D_{0,0K}$ (TQ)	$\sum D_{0,0K}$ (Q5)	$\Delta H_{f,0K}$ (TQ)	$\Delta H_{f,0K}$ (Q5)	$\Delta H_{f,298K}$ (TQ)	$\Delta H_{f,298K}$ (Q5)	$\Delta H_{f,298K}$ Expt
HOCl	156.5	156.6	-17.3	-17.4	-18.0	-18.2	-17.8,[a] -17.9 ± 0.3 [b] -18.36 ± 0.01 [c]
OCl^-	115.2	115.5	-27.6	-28.0	-27.6	-28.0	-28.2 [d] -28.28 ± 0.03 [c]
HOBr	150.7	150.8	-11.9	-12.1	-14.4	-14.6	-14.5 ± 0.3 [b] -14.84 ± 0.11 [c]
OBr^-	109.1	109.6	-21.9	-22.4	-23.7	-24.2	-24.4 [d] -24.77 ± 0.13 [c]
HOI	148.3		-12.1		-13.2		-16.6 ± 1.3 [b] -14.1 ± 0.7 [c]
OI^-	108.4		-23.9		-24.3		-25.6 [d] -24.9 ± 0.2 [c]
H_2O	219.3	219.2	-57.1	-57.0	-57.8	-57.7	-57.8 ± 0.01 [a] -57.80 ± 0.01 [c]
OH^-	143.5	143.7	-32.9	-33.0	-32.9	-33.0	33.24 ± 0.01 [c]
H_2S	173.3	173.3	-4.4	-4.4	-5.1	-5.1	-4.9,[a] -4.92 ± 0.12,[b] -5.33 ± 0.07 [e]
SH^-	136.9	136.8	-19.6	-19.5	-19.6	-19.5	-20.1 [a]
H_2Se	151.9	151.9	9.3	9.2	8.4	8.3	7.1 [f]
SeH^-	125.2	125.2	-15.7	-15.7	-15.9	-16.0	
H_2Te	130.3		23.7		22.6		23.8 [f]
TeH^-	113.8		-11.5		-11.9		

[a] Reference 63. [b] Reference 64. [c] Reference 62. [d] Reference 65. [e] Reference 66. [f] Reference67.

Table 7.2. Gas Phase and Aqueous Acidities at the Best Available Level at 298 K in kcal/mol

Species	$\Delta G_{gas,298K}$	$\Delta G_{gas,298K}$ Lit	$\Delta G_{aq,298K}$	$\Delta G_{aq,298K}$ Expt	$\Delta\Delta G$ (A^- - HA)	$\Delta\Delta G$ (A^- - HA) Expt	pK_a	pK_a Lit
HF	365.7	365.6 ± 0.02 [a]	5.2	4.3	-96.5	-97.1	3.8	3.17 [b]
HCl	328.5	328.0 ± 0.00 [a]	-10.2	-8.0	-74.1	-71.7	-7.5	-5.9 ± 0.4 [c]
HBr	318.1	318.2 ± 0.04 [a]	-14.1	-12.0	-67.9	-65.9	-10.4	-8.8 ± 0.8 [c]
HI	308.9	309.2 ± 0.01 [a]	-18.5	-13.0	-63.1, -56.1	-57.8	-13.6	-9.5 ± 1.0 [c]
H_2O	384.2	383.7 ± 0.02 [d]	20.2	19.1	-99.3	-100.3	14.8	13.995 [e]
H_2S	346.3	344.4 ± 3.1 [f]	10.3	9.6	-71.5	-70.5	7.6	7.05 [e]
H_2Se	336.0	335.2 ± 0.8 [g]	6.0	5.3	-65.7	-65.6	4.4	3.89 [e]
H_2Te	326.1	325.0 ± 2.1 [h]	0.4	3.5	-61.4	-57.2	0.3	2.6 [e]
HOCl	352.7	352.6 ± 0.04 [a]	6.4	10.1	-81.5	-78.2	4.7	7.40 [e]
HOBr	353.0	352.6 ± 0.24 [a]	10.8	11.7	-78.0	-76.6	7.9	8.55 [e]
HOI	351.5	351.8 ± 0.9 [a]	8.5	14.3	-78.7	-73.2	6.2	10.5 [e]

[a] Reference 68. [b] Reference 75. [c] Experimental gas phase acidity and calculated pK_a, based on experimental gas phase values and calculated ΔG_{aq}, from reference 69. [d] Reference 70. [e] Reference 67. [f] Reference 71. [g] Reference 72. [h] Reference 73.

Table 7.3. Relative Gas Phase and Aqueous Free Energies for Anionic Species with Multiple Conformations at 298 K in kcal/mol

Anionic Species	$\Delta G_{gas,298K}$ Relative	$\Delta G_{aq,298K}$ Relative
$F^-(H_2O)_4$ A [a]	0.0	0.0
$F^-(H_2O)_4$ B [a]	4.0	4.4
$F^-(H_2O)_4$ C [a]	0.8	2.4
$F^-(H_2O)_4$ D [a]	5.0	7.5
$F^-(H_2O)_8$ A [b]	0.0	0.0
$F^-(H_2O)_8$ B [b]	3.4	2.4
$F^-(H_2O)_8$ C [b]	4.8	7.2
$Cl^-(H_2O)_4$ A [a]	1.1	0.0
$Cl^-(H_2O)_4$ B [a]	0.0	2.0
$Cl^-(H_2O)_4$ C [a]	0.6	0.6
$Cl^-(H_2O)_4$ D [a]	1.0	4.0
$Cl^-(H_2O)_8$ A [b]	1.7	1.8
$Cl^-(H_2O)_8$ B [b]	2.4	0.0
$Cl^-(H_2O)_8$ C [b]	0.0	6.6
$Br^-(H_2O)_4$ A [c]	1.7	0.0
$Br^-(H_2O)_4$ B [c]	0.3	0.2
$Br^-(H_2O)_4$ C [c]	0.0	2.7
$Br^-(H_2O)_4$ D [c]	0.4	3.1
$Br^-(H_2O)_8$ A [d]	1.5	2.3
$Br^-(H_2O)_8$ B [d]	1.4	0.0
$Br^-(H_2O)_8$ C [d]	0.0	7.0
$I^-(H_2O)_4$ A [c]	1.9	0.0
$I^-(H_2O)_4$ B [c]	0.2	1.9
$I^-(H_2O)_4$ C [c]	0.0	0.0
$I^-(H_2O)_4$ D [c]	0.0	2.7
$I^-(H_2O)_8$ A [d]	2.6	2.6
$I^-(H_2O)_8$ B [d]	1.9	0.0
$I^-(H_2O)_8$ C [d]	0.0	5.4
$OCl^-(H_2O)_4$ A [a]	0.0	0.0
$OCl^-(H_2O)_4$ B [a]	1.8	0.9
$OCl^-(H_2O)_8$ A [b]	0.0	0.0
$OCl^-(H_2O)_8$ B [b]	12.2	13.5
$OBr^-(H_2O)_4$ A [c]	0.0	0.0
$OBr^-(H_2O)_4$ B [c]	1.9	0.7
$OBr^-(H_2O)_8$ A [d]	0.0	0.0
$OBr^-(H_2O)_8$ B [d]	11.4	10.8
$OI^-(H_2O)_4$ A [c]	0.0	0.0
$OI^-(H_2O)_4$ B [c]	3.2	0.9
$OI^-(H_2O)_8$ A [d]	0.0	0.0
$OI^-(H_2O)_8$ B [d]	13.5	13.9

[a] Calculated at the CCSD(T)-F12b/aD-F12 and CCSD(T)-F12b/aD-F12/COSMO level. [b] Calculated at the CCSD(T)/aD and CCSD(T)/aD/COSMO level. [c] Calculated at the CCSD(T)/aT-PP and CCSD(T)/aT-PP/COSMO level. [d] Calculated at the CCSD(T)/aD-PP and CCSD(T)/aD-PP/COSMO level.

Table 7.4. Relative Gas Phase and Aqueous Free Energies for Neutral Species with Multiple Conformations at 298 K in kcal/mol

Neutral Species	$\Delta G_{gas,298K}$ Relative	$\Delta G_{aq,298K}$ Relative
$HF(H_2O)_4$ A [a]	5.1	4.8
$HF(H_2O)_4$ B [a]	2.5	2.6
$HF(H_2O)_4$ C [a]	0.0	0.0
$HF(H_2O)_4$ D [a]	4.5	4.3
$HCl(H_2O)_4$ A [a]	2.8	3.0
$HCl(H_2O)_4$ B [a]	4.4	1.7
$HCl(H_2O)_4$ B2 [a]	0.0	0.5
$HCl(H_2O)_4$ C [a]	0.0	0.0
$HCl(H_2O)_4$ D [a]	2.6	2.9
$HBr(H_2O)_4$ A [b]	3.2	3.2
$HBr(H_2O)_4$ B [b]	2.9	1.7
$HBr(H_2O)_4$ B2 [b]	0.0	1.0
$HBr(H_2O)_4$ C [b]	0.6	0.0
$HBr(H_2O)_4$ D [b]	3.0	2.8
$HI(H_2O)_4$ A [b]	3.6	8.2
$HI(H_2O)_4$ B [b]	0.4	5.5
$HI(H_2O)_4$ B2 [b]	0.0	5.4
$HI(H_2O)_4$ C [b]	4.6	0.0
$HI(H_2O)_4$ D [b]	3.6	7.8

[a] Calculated at the CCSD(T)-F12b/aD-F12 and CCSD(T)-F12b/aD-F12/COSMO level. [b] Calculated at the CCSD(T)/aT-PP and CCSD(T)/aT-PP/COSMO level.

Table 7.5. Absolute Hydration Free Energy of Anions Calculated as the Converged Free Energy Change from X^-(gas) + $(H_2O)_n$(aq) → $X^-(H_2O)_n$(aq) at 298 K in kcal/mol

Species	MP2/aD	MP2/aD Lit. [a]	CCSD(T)/aD	CCSD(T)/aD Lit. [a]	CCSD(T)/aT	CCSD(T)-F12b/aD-F12	Best Est. Lit.
$F^-(H_2O)_4$ A	-97.1	*-94.3*	-98.8	*-96.0*	-99.2	-98.4	
$F^-(H_2O)_4$ B	-93.0	*-91.5*	-94.7	*-93.1*	-95.1	-93.9	
$F^-(H_2O)_4$ C	-94.6	*-90.8*	-96.4	*-92.7*	-96.8	-95.9	-104.3 ± 0.7 [a]
$F^-(H_2O)_4$ D	-89.8	*-88.3*	-94.8	*-90.3*	-91.9	-90.9	-82.5/-111.1 [b]
$F^-(H_2O)_8$ A	-99.3	*-99.8*	-101.0	*-101.8*	-101.4	-100.5	-103.0 [c]
$F^-(H_2O)_8$ B	-96.6	*-92.6*	-97.8			-98.1	
$F^-(H_2O)_8$ C	-91.2	*-84.0*	-92.9			-93.2	
$Cl^-(H_2O)_4$ A	-64.8		-65.5		-66.3	-65.8	
$Cl^-(H_2O)_4$ B	-63.2		-64.0		-64.4	-63.8	
$Cl^-(H_2O)_4$ C	-66.6		-67.0		-68.0	-67.8	-64.5/-81.3 [b]
$Cl^-(H_2O)_4$ D	-66.5		-67.0		-67.8	-67.2	-72.7 [c]
$Cl^-(H_2O)_8$ A	-69.3		-69.5			-71.2	-83.4/-77.2 [d]
$Cl^-(H_2O)_8$ B	-71.7		-71.3		-72.6	-73.2	
$Cl^-(H_2O)_8$ C	-63.9		-64.7			-65.9	
$Br^-(H_2O)_4$ A	-62.1		-61.6		-61.9		
$Br^-(H_2O)_4$ B	-62.6		-61.6		-62.1		
$Br^-(H_2O)_4$ C	-59.5		-59.0		-59.1		-59.8/-75.3 [b]
$Br^-(H_2O)_4$ D	-60.1		-59.6		-60.0		-66.3 [c]
$Br^-(H_2O)_8$ A	-64.7		-63.5				-76.0/-72.0 [d]
$Br^-(H_2O)_8$ B	-67.7		-65.8				
$Br^-(H_2O)_8$ C	-60.6		-58.8				
$I^-(H_2O)_4$ A	-55.5		-55.4		-56.0		
$I^-(H_2O)_4$ B	-55.9		-55.2		-56.0		
$I^-(H_2O)_4$ C	-53.2		-52.9		-53.3		-52.6/-65.7 [b]
$I^-(H_2O)_4$ D	-53.7		-53.4		-54.1		-57.4 [c]
$I^-(H_2O)_8$ A	-57.5		-56.4				-67.2/-63.2 [d]
$I^-(H_2O)_8$ B	-60.8		-59.0				

$I^-(H_2O)_8$ C	-55.1		-53.6				
$OH^-(H_2O)_4$	-92.6	*-90.2*	-93.9	*-91.7*	-94.0	-93.2	-104.5 [a]
$OH^-(H_2O)_8$	-100.2	*-99.9*	-101.4	*-101.4*			-82.5/-102.8 [b] -103.0 [c]
$SH^-(H_2O)_4$	-62.8		-63.1		-63.2	-62.8	-56.2/-70.5 [b]
$SH^-(H_2O)_8$	-67.1		-66.8				
$SeH^-(H_2O)_4$	-59.5		-58.8		-58.4		-57.4/-86.0 [b]
$SeH^-(H_2O)_8$	-62.0		-59.8				
$TeH^-(H_2O)_4$	-52.9		-52.2		-51.7		
$TeH^-(H_2O)_8$	-57.4		-55.2				
$OCl^-(H_2O)_4$ A	-74.4		-73.9		-73.0	-71.8	
$OCl^-(H_2O)_4$ B	-73.3		-73.1		-72.3	-70.9	-84.2/-90.3 [d]
$OCl^-(H_2O)_8$ A	-84.2		-83.3				
$OCl^-(H_2O)_8$ B	-70.1		-69.8				
$OBr^-(H_2O)_4$ A	-74.0		-72.8		-71.8		
$OBr^-(H_2O)_4$ B	-73.1		-72.1		-71.1		-81.9/-89.9 [d]
$OBr^-(H_2O)_8$ A	-81.5		-80.0				
$OBr^-(H_2O)_8$ B	-71.0		-69.1				
$OI^-(H_2O)_4$ A	-72.6		-71.2		-69.8		
$OI^-(H_2O)_4$ B	-71.7		-70.4		-68.9		-77.1/-88.2 [d]
$OI^-(H_2O)_8$ A	-83.5		-81.4				
$OI^-(H_2O)_8$ B	-69.9		-67.5				

[a] From References 4 and 15. [b] Theoretical/experimental values from Reference 74. Calculated values, based on ionic radius and calculated using $\Delta G_{hyd}[H^+]$ = -1056 kJ/mol (-254.5 kcal/mol). [c] Calculated values estimated based on cluster-pair approximations with error of ± 2 kcal/mol from Reference 79. [d] Theoretical/experimental values from reference 12. Values reported in this reference used values for $\Delta G_{hyd}[H^+]$ = -1092 kJ/mol (-261.0 kcal/mol) for theoretical results and $\Delta G_{hyd}[H^+]$ = -1064 kJ/mol (-254.3 kcal/mol) for experimental values.

Table 7.6. Absolute Hydration Free Energy of Acids Calculated as the Converged Free Energy Change from HX(gas) + $(H_2O)_n$(aq) → $HX(H_2O)_n$(aq) at 298 K in kcal/mol

Species	Bulk Solvent Shift (MP2/aD/COSMO)	MP2/aD	CCSD(T)/aD	CCSD(T)/aT	CCSD(T)-F12b/aD-F12
$HF(H_2O)_4$ A	-1.9	-0.1	-0.4	0.0	0.5
$HF(H_2O)_4$ B	-1.4	-2.6	-2.8	-2.4	-1.7
$HF(H_2O)_4$ C	-1.6	-4.9	-4.9	-4.9	-4.3
$HF(H_2O)_4$ D	-1.8	-0.6	-0.9	-0.5	0.0
$HCl(H_2O)_4$ A	-1.3	2.9	3.3	4.1	4.3
$HCl(H_2O)_4$ B	-4.3	2.4	2.3	2.6	2.9
$HCl(H_2O)_4$ B2	-1.0	0.3	0.7	1.4	1.8
$HCl(H_2O)_4$ C	-1.6	-0.5	0.5	0.9	1.2
$HCl(H_2O)_4$ D	-1.3	2.6	3.1	3.8	4.1
$HBr(H_2O)_4$ A	-1.1	3.6	4.6	5.3	
$HBr(H_2O)_4$ B	-2.3	3.0	3.4	3.8	
$HBr(H_2O)_4$ B2	-0.1	1.6	2.5	3.1	
$HBr(H_2O)_4$ C	-1.7	-0.1	1.8	2.1	
$HBr(H_2O)_4$ D	-1.2	3.2	4.3	4.9	
$HI(H_2O)_4$ A	-0.6	5.4	6.4	7.0	
$HI(H_2O)_4$ B	-0.1	3.1	3.7	4.1	
$HI(H_2O)_4$ B2	0.3	2.7	3.6	4.2	
$HI(H_2O)_4$ C	-9.7	-5.2	-1.8	-2.9	
$HI(H_2O)_4$ D	-0.9	4.7	6.0	6.6	
$(H_2O)_5$	-7.0	-6.7	-6.7	-6.4	-6.0
$H_2S(H_2O)_4$	-1.3	3.8	4.1	4.7	4.9
$H_2Se(H_2O)_4$	-1.2	4.2	5.3	5.9	
$H_2Te(H_2O)_4$	-1.5	4.8	5.8	6.2	
$HOCl(H_2O)_4$	-1.5	-2.5	-2.2	-1.8	-1.3
$HOBr(H_2O)_4$	-1.6	-3.1	-2.5	-2.0	
$HOI(H_2O)_4$	-2.1	-3.9	-3.1	-2.7	

Table 7.7. ΔG_{solv} for Single Molecules at 298 K in kcal/mol at MP2/aD/COSMO and MP2/aD/SMD

Species	ΔG_{solv} COSMO	ΔG_{solv} SMD	Our Best ΔG_{hyd}	Lit ΔG_{hyd}
HF	-5.4	-4.1	-4.9	
F^-	-87.9		-101.4	-104.3
HCl	-3.2	-1.6	0.9	
Cl^-	-71.6		-73.2	
HBr	-2.7	-0.6	2.1	
Br^-	-66.9		-65.8	
HI	-2.0	-2.1	-2.9 [a]	
I^-	-60.4		-59.0	
H_2O	-6.7	-7.1	-6.0	-5.7
OH^-	-86.2		-101.4	-104.5
H_2S	-3.0	-1.5	4.7	-0.5
SH^-	-67.4		-66.8	
H_2Se	-2.6	-2.1	5.9	
SeH^-	-63.9		-59.8	
H_2Te	-2.1	-1.5	6.2	
TeH^-	-58.0		-55.2	
HOCl	-4.7	-4.3	-1.8	
OCl^-	-71.5		-83.3	
HOBr	-4.8	-3.6	-2.0	
OBr^-	-69.3		-80.0	
HOI	-5.1	-7.5	-2.7	
OI^-	-66.3		-81.4	

[a] Neutral species dissociated to paired ions.

Figures

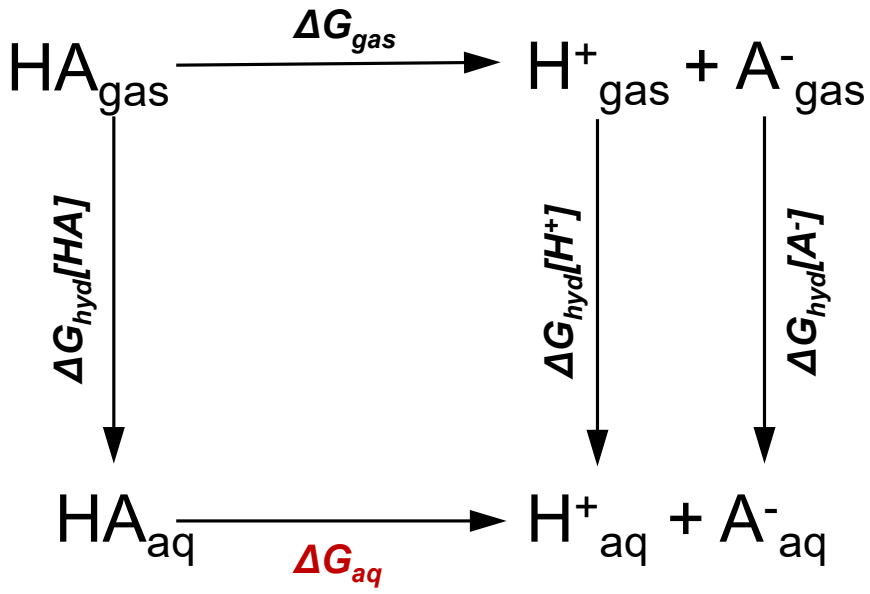


Figure 7.1. Thermodynamic cycle used for calculating the aqueous acidity of all species from gas phase acidity and hydration free energy of HA, H^+, and A^-.

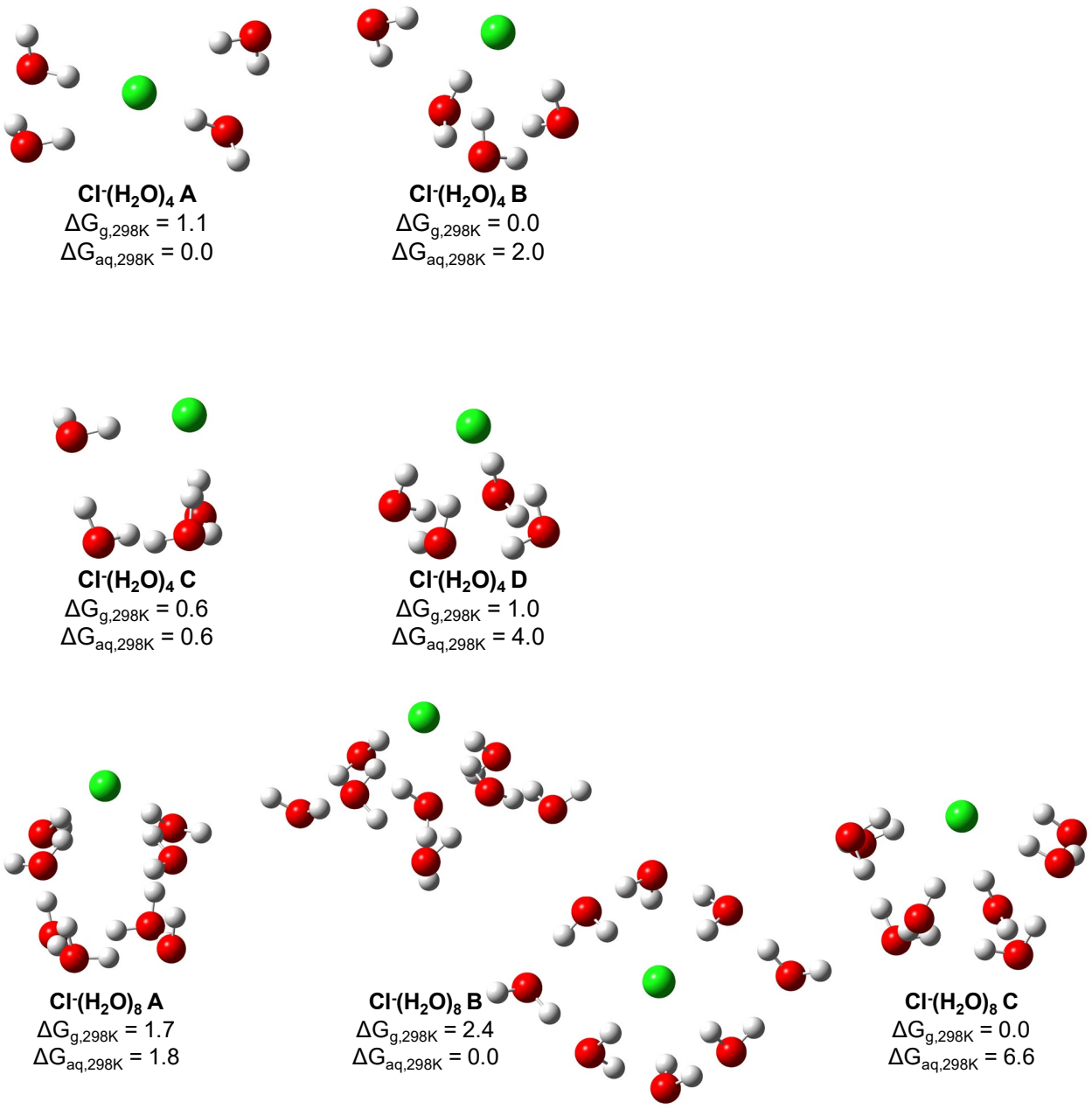


Figure 7.2. Conformers for $Cl^-(H_2O)_4$ and $Cl^-(H_2O)_8$ with relative gas phase and aqueous free energies at 298 K in kcal/mol. Energies for chloride with the water tetramer and octamer calculated at the CCSD(T)-F12b/aT-F12 level and with the water octamer calculated at the CCSD(T)/aD-F12b level. Solvation free energy calculated for all species at the MP2/aD/COSMO level.

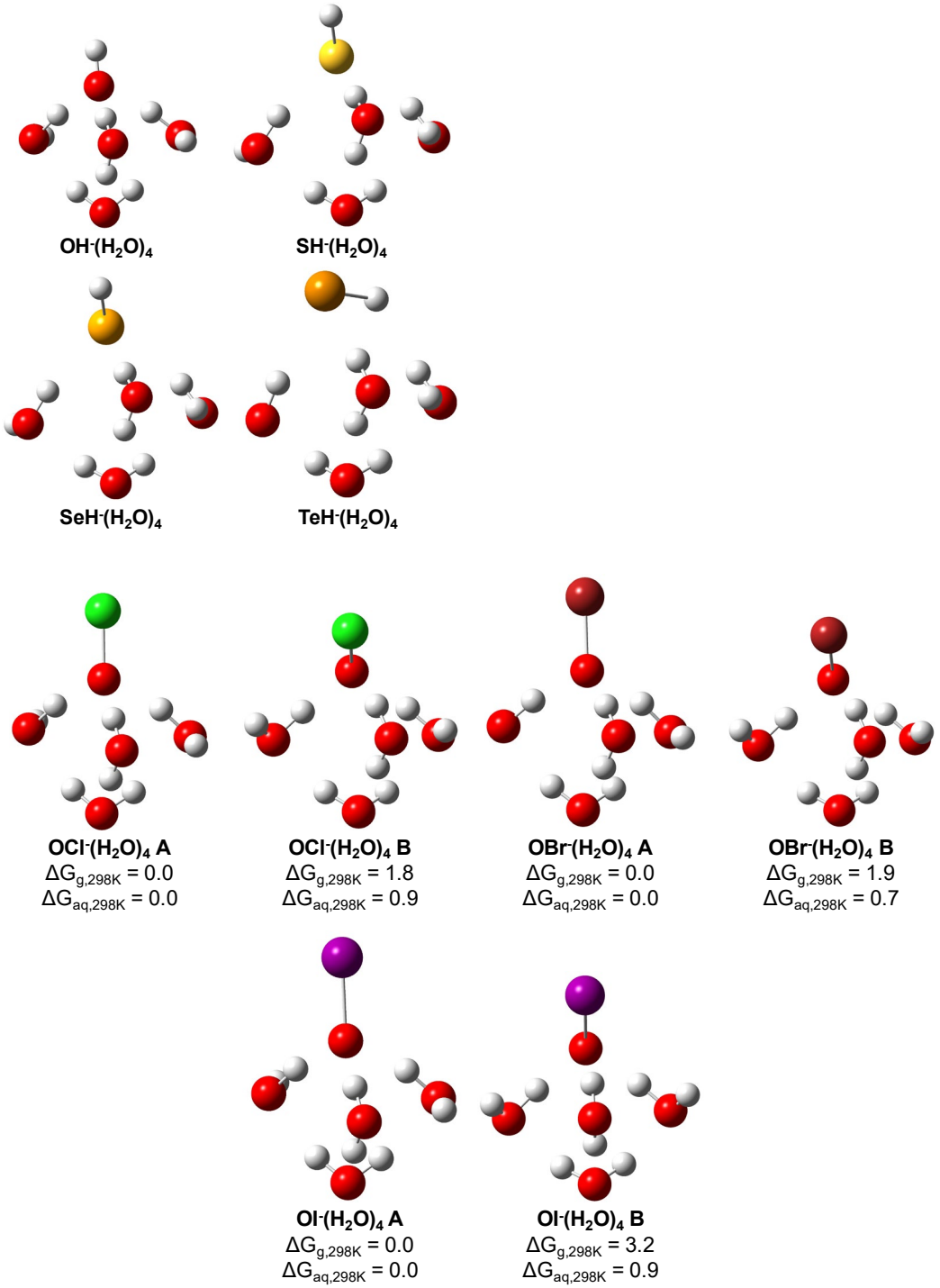


Figure 7.3. MP2/aD optimized geometries of diatomic anions with 4 H_2O molecules. Relative energies for $OCl^-(H_2O)_4$ **A** and $OCl^-(H_2O)_4$ **B** are calculated at the CCSD(T)-F12b/aug-cc-pVDZ-F12 level with and without solvation at the MP2/aug-cc-pVDZ/COSMO level at 298 K and are given in kcal/mol. $OBr^-(H_2O)_4$ **A**, $OBr^-(H_2O)_4$ **B**, $OI^-(H_2O)_4$ **A**, and $OI^-(H_2O)_4$ **B** are calculated at the CCSD(T)-F12b/aug-cc-pVTZ-PP level with and without solvation at the MP2/aug-cc-pVDZ/COSMO level at 298 K and are given in kcal/mol.

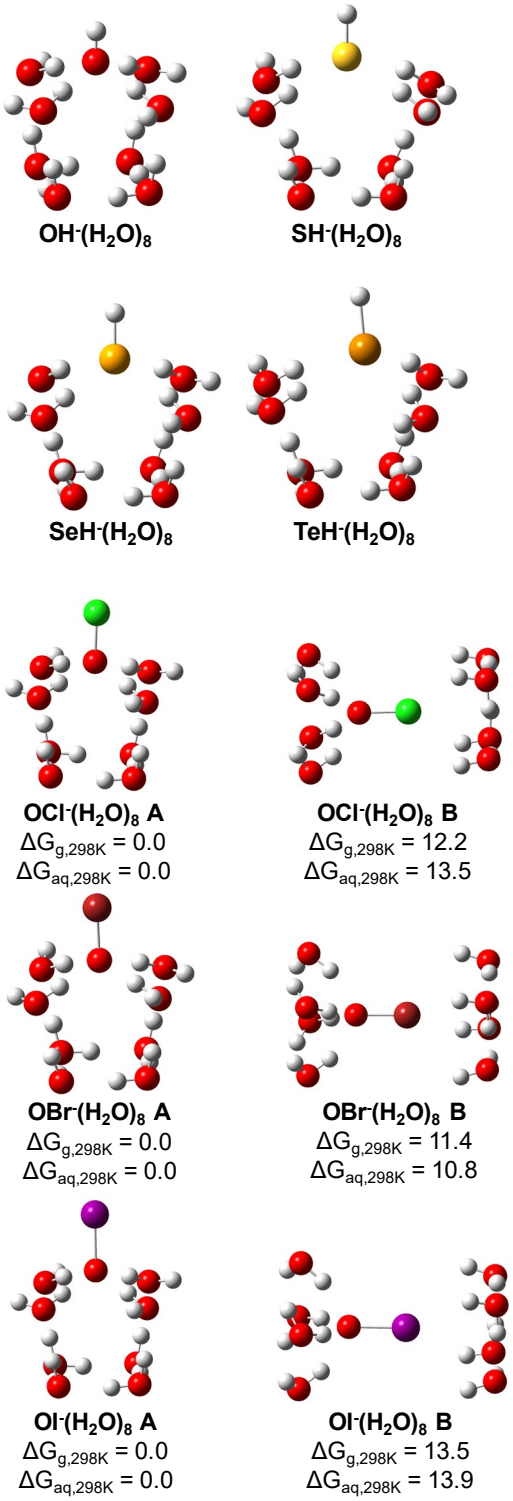


Figure 7.4. MP2/aD optimized geometries of diatomic anions with 8 H_2O molecules.

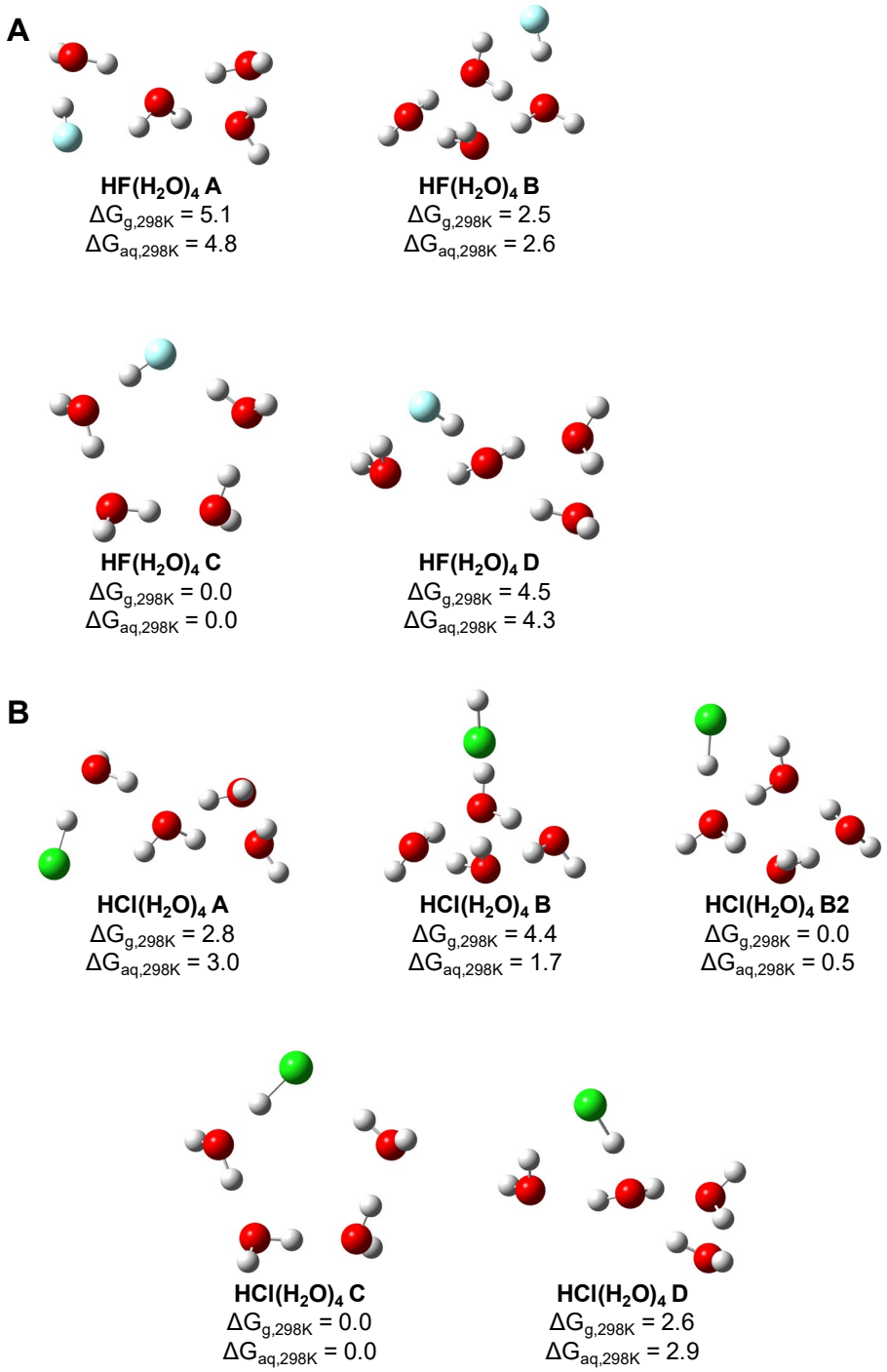


Figure 7.5. Conformers for (A) $HF(H_2O)_4$ and (B) $HCl(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol. Energies for HF and HCl with the water tetramer are calculated at the CCSD(T)-F12b/aD-F12 level. Solvation free energy is calculated for all species at the MP2/aD/COSMO level.

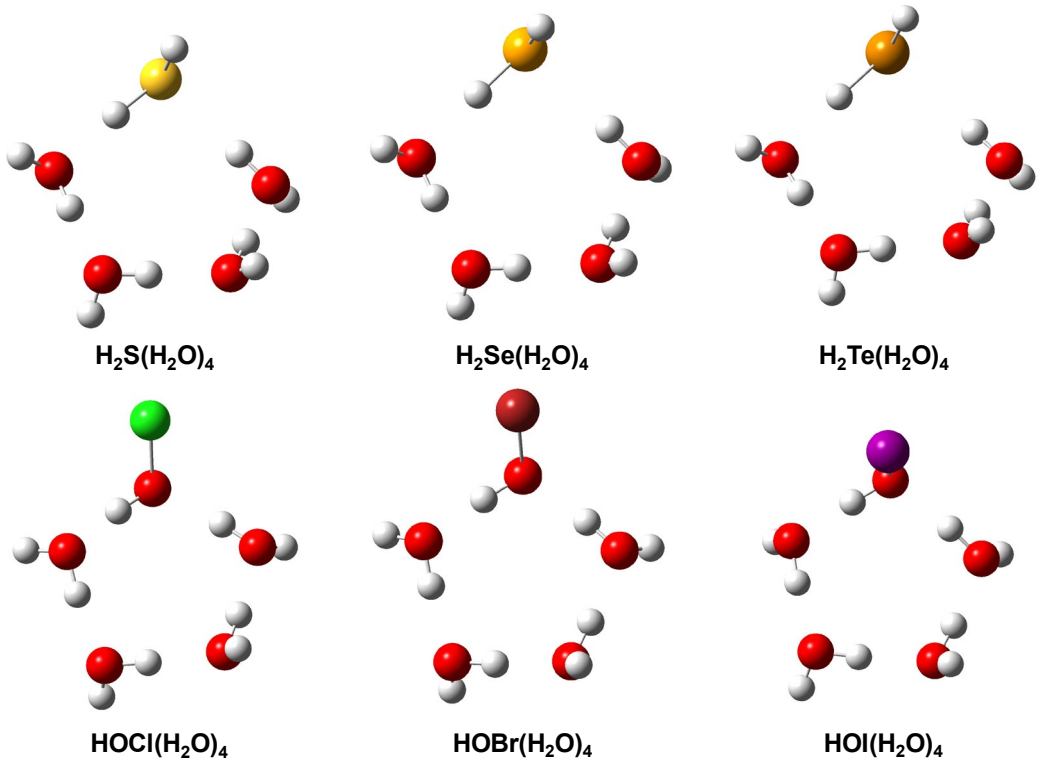


Figure 7.6. MP2/aD optimized clusters for H_2S, H_2Se, H_2Te, HOCl, HOBr, and HOI with 4 H_2O molecules.

References

[1] Ohtaki, H.; Radnai, T. Structure and Dynamics of Hydrated Ions. *Chem. Rev.* **1993**, *93*, 1157 – 1204.

[2] Kim, J.; Lee, H. M.; Suh, S. B.; Majumdar, D.; Kim, K. S. Comparative ab Initio Study of the Structures, Energetics and Spectra of $X^-(H_2O)_{n=1-4}$ [X = F, Cl, Br, I] Clusters. *J. Chem. Phys.* **2000**, *113*, 5259 – 5272.

[3] Lee, M. L.; Kim, D.; Kim, K. W. Structures, Spectra, and Electronic Properties of Halide-Water Pentamers and Hexamers, $X^-(H_2O)_{5,6}$ (X = F, Cl, Br, I): Ab Initio Study. *J. Chem. Phys.* **2002**, *116*, 5509 – 5520.

[4] Zhan, C.-G.; Dixon, D. A. Hydration of the Fluoride Anion: Structure and Absolute Hydration Free Energy from First-Principles Electronic Structure Calculations. *J. Phys. Chem. A* **2004**, *108*, 2020 – 2029.

[5] D'Angelo, P.; Migliorati, V.; Guidoni, L. Hydration Properties of the Bromide Aqua Ion: The Interplay of First Principle and Classical Molecular Dynamics, and X-ray Absorption Spectroscopy. *Inorg. Chem.* **2010**, *49*, 4224 – 4231.

[6] Rosseinsky, D. R. Electrode Potentials and Hydration Energies. Theories and Correlations. *Chem. Rev.* **1965**, *65*, 467 – 490.

[7] Klots, C. E. Solubility of Protons in Water. *J. Phys. Chem.* **1981**, *85*, 3585 – 3588.

[8] Kaya, S.; deFarias, R. F. Absolute Ion Hydration Enthalpies from Absolute Hardness and Some VBT Relationships. *Chem. Phys. Lett.* **2018**, *691*, 169 – 171.

[9] Riccardi, D.; Guo, H.-B.; Parks, J. M.; Gu, B.; Liang, L.; Smith, J. C. Cluster-Continuum Calculations of Hydration Free Energies of Anions and Group 12 Divalent Cations. *J. Chem. Theory Comput.* **2013**, *9*, 555 – 569.

[10] Serr, A.; Netz, R. R. Polarizabilities of Hydrated and Free Ions Derived From DFT Calculations. *Int. J. Quantum Chem.* **2006**, 106*, 2960 – 2974.*

[11] Gálvez, O.; Gómez, P. C. An ab Initio Study on the Structure and Energetics of the ClO Hydrates. *Chem. Phys. Lett.* **2007**, *448*, 16 – 23.

[12] Stanbury, D. M. Hydration Energies and Reactivity of the Hypohalite Anions. *Inorg. Chem.* **2018**, *57*, 1665 – 1669.

[13] Kritayakornupong, C.; Vchirawongkwin, V.; Rode, B. M. Determination of Structure and Dynamics of the Solvated Bisulfide (HS^-) Ion by an Initio QMCF Molecular Dynamics. *J. Phys. Chem. B* **2010**, *114*, 12883 – 12887.

[14] Parr, R. G.; Pearson, R. G. Absolute Hardness: Companion Parameter to Absolute Electronegativity. *J. Am. Chem. Soc.* **1983**, *105*, 7512 – 7516.

[15] Zhan, C.-G.; Dixon, D. A. First-Principles Determination of the Absolute Hydration Free Energy of the Hydroxide Ion. *J. Phys. Chem. A* **2002**, *106*, 9737 – 9744.

[16] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16, Revision A.03, Gaussian, Inc., Wallingford CT, 2016.

[17] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.

[18] Lee, C.; Yang, W.; Parr, R. G. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Physical Review B.* **1988**, *37*, 785 – 789.

[19] Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys*. **1989**, *90*, 1007 – 1023.

[20] Kendall, R. A.; Dunning, T. H. Jr.; Harrison, R. J., Electron Affinities of the First Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys*. **1992,** *96*, 6796 – 6806.

[21] Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically Convergent Basis Sets with Relativistic Pseudopotentials. II. Small-Core Pseudopotentials and Correlation Consistent Basis Sets for the Post-d Group 16 – 18 Elements. *J. Chem. Phys.* **2003**, *119*, 11113 – 11123.

[22] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 0618 – 0622.

[23] Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical Models Incorporating Electron Correlation. *Int. J. Quantum. Chem.* **1976**, *10*, 1 – 19.

[24] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Contiuum Solvation Models. *Chem. Rev.* **2005**, 105, 2999 – 3093.

[25] Klamt, A. *Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam, 2005.

[26] Klamt, A.; Schüürmann, G. J. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *Chem. Soc. Perkin Trans.* **1993**, 2, 799 – 805.

[27] Purvis, G. D., III; Bartlett, R. J. A Full Coupled Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910 – 1918.

[28] Raghavachari, K.; Truck, G.W.; Pople, J.A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett*. **1989**, *157*, 479 – 483.

[29] Watts, J. D.; Gauss, J.; Bartlett, R. J. Coupled-Cluster Methods with Noniterative Triple Excitations for Restricted Open Shell Hartree-Fock and Other General Single Determinant Reference Functions. Energies and Analytical Gradients. *J. Chem. Phys.* **1993**, *98*, 8718 – 8733.

[30] Bartlett, R. J.; Musial, M. Coupled Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291 – 352.

[31] Knowles, P. J.; Hampel, C.; Werner, H.-J. Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **1993**, *99*, 5219 – 5227.

[32] Knowles, P. J.; Hampel, C.; Werner, H.-J. Erratum: Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **2000,** *112*, 3106 – 3107.

[33] Deegan, M. J. O.; Knowles, P. J. Perturbative Corrections to Account for Triple Excitations in Closed and Open Shell Coupled Cluster Theories. *Chem. Phys. Lett*. **1994**, *227*, 321 – 327.

[34] Adler, T. B.; Knizia, G.; Werner, H.-J. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Phys. Chem.* **2007,** *127,* 221106-1 – 221106-4.

[35] Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009,** *130,* 054104-1 – 054104-20.

[36] Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-ζ Quality Coupled-Cluster Correlation Energies with Triple- ζ Basis Sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921 – 1930.

[37] Yousaf, K. E.; Peterson, K. A. Optimized Auxiliary Basis Sets for Explicitly Correlated Methods. *J. Chem. Phys.* **2008**, *129*, 184108-1 – 184108-7.

[38] Yousaf, K. E.; Peterson, K. A. Optimized Complementary Auxiliary Basis Sets for Explicitly Correlated Methods: aug-cc-pV*n*Z Orbital Basis Sets. *Chem. Phys. Lett.* **2009**, *476*, 303 – 307.

[39] Werner H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, T.; Korona, T.; Lindh, R.; et al. MOLPRO, version 2018.2, a package of *ab initio* programs, See http://www.molpro.net.

[40] Werner H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242 – 253.

[41] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions, *J. Phys. Chem. B.* **2009**, *113*, 6378 − 6396.

[42] Klamt, A. *Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*; Elsevier: Amsterdam.

[43] Klamt, A.; Schuurmann, G. Cosmo - a New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient, *J Chem Soc Perk T 2* **1993**, 799 − 805.

[44] de Jong, W. A.; Harrison, R. J.; Dixon, D. A. Parallel Douglas-Kroll Energy and Gradients in NWChem: Estimating Scalar Relativistic Effects Using Douglas-Kroll Contracted Basis Sets. *J. Chem. Phys*. **2001**, *114*, 48 – 53.

[45] Bross, D. H.; Peterson, K. A. Correlation Consistent, Douglas-Kroll-Hess Relativistic Basis Sets for the 5p and 6p Elements. *Theor. Chem. Acc*. **2014**, *133*, 1434-1 – 1434-12.

[46] Woon, D. E.; Dunning, T. H. Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-Valence Basis Sets for Boron Through Neon. *J. Chem. Phys.* **1995**, *103*, 4572 – 4585.

[47] Peterson, K. A.; Dunning, T. H. Accurate Correlation Consistent Basis Sets for Molecular Core-Valence Correlation Effects: The Second Row Atoms Al-Ar, and the First Row Atoms B-Ne Revisited. *J. Chem. Phys.* **2002**, *177*, 10548 – 10560.

[48] Dunham, J. L. The Energy Levels of a Rotating Vibrator. *Phys. Rev.* **1932**, *41*, 721 – 731.

[49] Dunham, J. L. The Wentzel-Brillouin-Kramers Method of Solving the Wave Equation. *Phys. Rev.* **1932**, *41*, 713 – 720.

[50] Douglas, M.; Kroll, N. M. Quantum Electrodynamical Corrections to the Fine Structure of Helium. *Ann. Phys.* **1974**, *82*, 89 – 155.

[51] Hess, B. A. Applicability of the No-Pair Equation with Free-Particle Projection Operators to Atomic and Molecular Structure Calculations. *Phys. Rev. A.* **1985**, *32*, 756 – 763.

[52] Hess, B. A. Relativistic Electronic-Structure Calculations Employing a Two-Component No-Pair Formalism with External-Field Projection Operators. *Phys. Rev. A.* **1986**, *33*, 3742 – 3748.

[53] Martin, J. M. L. Ab Initio Total Atomization Energies of Small Molecules - Towards the Basis Set Limit. *Chem. Phys. Lett.* **1996**, *259*, 669 – 678.

[54] Feller, D.; Peterson, K.A.; Hill, J. G. On the Effectiveness of CCSD(T) Complete Basis Set Extrapolations for Atomization Energies. *J. Chem. Phys.* **2011***, 135,* 044102-1 – 044102-18.

[55] Dixon, D. A.; Feller, D.; Peterson, K. A. A Practical Guide to Reliable First Principles Computational Thermochemistry Predictions Across the Periodic Table. In *Annual Reports in Computational Chemistry*, Ralph, A. W., Ed.; Elsevier: 2012; Vol. 8, pp 1 – 28.

[56] Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy: Current Strategies and Future Challenges. *Theor. Chem. Acc.* **2012**, *131*, 1 – 20.

[57] Feller, D.; Peterson, K. A.; Dixon, D. A. Further Benchmarks of a Composite, Convergent, Statistically Calibrated Coupled Cluster Based Approach for Thermochemical and Spectroscopic Studies. *Mol. Phys.* **2012**, *110*, 2381 – 2399.

[58] Feller, D.; Peterson, K. A.; Dixon, D. A. The Impact of Larger Basis Sets and Explicitly Correlated Coupled Cluster Theory on the Feller-Peterson-Dixon Composite Method. In *Annual Reports in Computational Chemistry*, Vol. 12, Ed. D. A. Dixon, Elsevier, Amsterdam, 2016, pp. 47 – 78.

[59] Moore, C. E. *Atomic Energy Levels as Derived from the Analysis of Optical Spectra, Volume 1, H to V;* U.S. National Bureau of Standards Circular 467, COM-72-50282; U.S. Department of Commerce, National Technical Information Service: Washington, DC, 1949.

[60] Chase, M. W., Jr. *NIST-JANAF Thermochemical Tables*, 4th ed. *J. Phys. Chem. Ref. Data.*, *Mono. 9*; American Institute of Physics: Woodbury, NY, 1998; Suppl 1.

[61] Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies of Formation. *J. Chem. Phys.* **1997**, *106*, 1063 – 1079.

[62] Ruscic, B.; Bross, D. H. Active Thermochemical Tables (ATcT) values based on ver. 1.122g of the Thermochemical Network (2019); available at ATcT.anl.gov.

[63] NIST Chemistry Webbook, NIST Standard Reference Database Number 69, Eds. P.J. Lindstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (https://webbook.nist.gov/), retrieved 4/15/2020.

[64] Sander, S. P.; Friedl, R. R.; Ravishankara, A. R.; Golden, D. M.; Kolb, C. E.; Kurylo, M. K.; Huie, R. E.; Orkin, V. L.; Molina, M. J.; Moortgat, G. K.; Finlayson-Pitts, B. J. *Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies: Evaluation Number 14*; JPL Publication 02-25; National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, 2003; http://jpldataeval.jpl.nasa.gov/pdf/JPL_02-25_rev02.pdf. Accessed March 7, 2020.

[65] Peterson, K. A.; Shepler, B. C.; Figgen, D.; Stoll, H. On the Spectroscopic and Thermochemical Properties of ClO, BrO, IO, and Their Anions. *J. Phys. Chem. A*, **2006**, *110*, 13877-13883.

[66] Grant, D. J.; Dixon, D. A.; Francisco, J. S.; Feller, D.; Peterson, K. A. Heats of Formation of the $H_{1,2}O_mS_n$ (m, n = 0-3) Molecules from Electronic Structure Calculations. *J. Phys. Chem. A* **2009**, *113*, 11343 – 11353.

[67] Lide, D. R. (E.) *CRC Handbook of Chemistry and Physics*, 86th Ed. CRC Press, Boca Raton, FL, USA, 2006.

[68] Blondel, C.; Delsart, C.; Goldfarb, F. Electron Spectrometry at the μeV level and the Electron Affinities of Si and F. *J. Phys. B: Atom. Mol. Opt. Phys.* **2001**, *34*, L281 – L288.

[69] Trummal, A.; Lipping, L.; Kaljurand, I.; Koppel, I. A.; Leito, I. Acidity of Strong Acids in Water and Dimethyl Sulfoxide. *J. Phys. Chem. A* **2016**, *120*, 3663 – 3669.

[70] Smith, J. R.; Kim, J. B.; Lineberger, W. C. High-Resolution Threshold Photodetachment Spectroscopy of OH^-. *Phys. Rev. A* **1997**, *55*, 2036 – 2043.

[71] Rempala, K.; Ervin, K. M. Collisional Activation of the Endoergic Hydrogen Atom Transfer Reaction $S^-(^2P) + H_2 \rightarrow SH^- + H$. *J. Chem. Phys.* **2000**, *112*, 4579 – 4590.

[72] Stoneman, R. C.; Larson, D. J. Photodetachment Spectroscopy of SeH^- in a Magnetic Field. *J. Phys. B* **1986**, *19*, L405 – L409.

[73] Gal, J.-F.; Maria, P.-C.; Decouzon, M. The Gas-Phase Acidity and Bond Dissociation Energies of Hydrogen Telluride. *Int. J. Mass Spectrom. Ion Processes* **1989**, *93*, 87 – 94.

[74] Marcus, Y. Thermodynamics of Solvation of Ions: Part 5. Gibbs Free Energy of Hydration at 298.16 K. *J. Chem. Soc. Faraday Trans.* **1991**, *87*, 2995 – 2999.

[75] Harris, D. C. *Quantitative Chemical Analysis*, 8th ed.; W. H. Freeman and Company: New York, NY, 2010, pp AP14.

[76] Zhan, C.-G.; Dixon, D. A. Absolute Hydration Free Energy of the Proton from First-Principles Electronic Structure Calculations. *J. Phys. Chem. A* **2001**, *105*, 11534 – 11540.

[77] Camaioni, D. M.; Schwerdtfeger, C. A. Comment on Accurate Experimental Values for the Free Energies of Hydration of H^+, OH^-, and H_3O^+. *J. Phys. Chem. A* **2005**, *109*, 10795 – 10797.

[78] Cheung, J. T.; Dixon, D. A.; Herschbach, D. R. Cluster Beam Chemistry: Adduct Complexes of Hydrogen Halides with Ammonia Clusters. *J. Phys. Chem.,* **1988**, *92*, 2536 – 2541

[79] Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle Jr., T. R. The Proton's Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data. *J. Phys. Chem. A* **1998**, *102*, 7787 – 7794.

[80] Ben-Naim, A.; Marcus, Y. J. Solvation Thermodynamics of Nonionic Solutes. *Chem. Phys.* **1984**, *81*, 2016 – 2027.

[81] Asthagiri, D.; Pratt, L. R.; Ashbaugh, H. S. Absolute Hydration Free Energies of Ions, Ion-Water Clusters, and Quasichemical Theory. *J. Chem. Phys*. **2003**, *119*, 2702 – 2708.

[82] Marcus, Y. A Simple Empirical Model Describing the Thermodynamics of Hydration of Ions of Widely Varying Charges, Sizes, and Shapes. *Biophys. Chem*. **1994**, *51*, 111 – 127.

[83] Hummer, G.; Pratt, L. R.; García, A. E. Hydration Free Energy of Water. *J. Phys. Chem*. **1996**, *100*, 1206 – 1215.

[84] Martin, R. L.; Hay, P. J.; Pratt, L. R. Hydrolysis of Ferric Ion in Water and Conformational Equilibrium. *J. Phys. Chem. A* **1998**, *102*, 3565 – 3573.

[85] Gutowski, K. E.; Dixon, D. A. Predicting the Energy of the Water Exchange Reaction and Free Energy of Solvation for the Uranyl Ion in Aqueous Solution. *J. Phys. Chem. A* **2006**, *110*, 8840 − 8856.

Full Reference 16 (Gaussian 16): Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision A.03, Gaussian, Inc., Wallingford CT, 2016.

Full Reference 39 (Molpro 2018): Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, D.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bennie, S. J.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Lee, S. J. R.; Liu, Y.; Lloyd, A. W.; Ma, Q.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Miller III, T. F.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.;Stone, A. J.; Tarroni, R.; Thorsteinsson, T.;Wang, M.; Welborn, M. MOLPRO, version 2018.2, a package of *ab initio* programs, See http://www.molpro.net.

Table A7.1.1. Bulk Solvent Shift ($\Delta\Delta G_{sol}[X^-,n]$) of Anions at 298 K in kcal/mol

Species	MP2/aD/COSMO (G16)	MP2/6-31++G** SVPE [a]
$F^-(H_2O)_4$ A	-47.5	-45.2
$F^-(H_2O)_4$ B	-47.0	-45.1
$F^-(H_2O)_4$ C	-45.8	-42.5
$F^-(H_2O)_4$ D	-45.0	-42.3
$F^-(H_2O)_8$ A	-33.4	-42.1
$F^-(H_2O)_8$ B	-34.4	-37.5
$Cl^-(H_2O)_4$ A	-41.7	
$Cl^-(H_2O)_4$ B	-40.7	
$Cl^-(H_2O)_4$ C	-44.8	
$Cl^-(H_2O)_4$ D	-43.7	
$Br^-(H_2O)_4$ A	-42.2	
$Br^-(H_2O)_4$ B	-43.8	
$Br^-(H_2O)_4$ C	-39.4	
$Br^-(H_2O)_4$ D	-39.9	
$I^-(H_2O)_4$ A	-40.1	
$I^-(H_2O)_4$ B	-42.1	
$I^-(H_2O)_4$ C	-37.5	
$I^-(H_2O)_4$ D	-38.4	
$OH^-(H_2O)_4$	-46.5	-42.1
$OH^-(H_2O)_8$	-32.6	-41.1
$SH^-(H_2O)_4$	-42.4	
$SH^-(H_2O)_8$	-30.0	
$SeH^-(H_2O)_4$	-41.3	
$SeH^-(H_2O)_8$	-31.4	
$TeH^-(H_2O)_4$	-38.8	
$TeH^-(H_2O)_8$	-29.5	
$OCl^-(H_2O)_4$ A	-41.6	
$OCl^-(H_2O)_4$ B	-42.4	
$OCl^-(H_2O)_8$ A	-31.3	
$OCl^-(H_2O)_8$ B	-29.9	
$OBr^-(H_2O)_4$ A	-41.0	
$OBr^-(H_2O)_4$ B	-42.2	
$OBr^-(H_2O)_8$ A	-29.7	
$OBr^-(H_2O)_8$ B	-30.3	
$OI^-(H_2O)_4$ A	-40.4	
$OI^-(H_2O)_4$ B	-42.7	
$OI^-(H_2O)_8$ A	-31.0	
$OI^-(H_2O)_8$ B	-30.7	

[a] From references 1 and 2.

Table A7.1.2. Geometry Parameters for CCSD(T)/awCQ-DK Optimized Triatomic Molecules

Species	Bond Length (Å)			Bond Angle (Degrees)		
	MP2/aT	**awCQ-DK**	**Lit [a]**	**MP2/aT**	**awCQ-DK**	**Lit [a]**
HOCl	0.969 O-H 1.690 O-Cl	0.964 O-H 1.691 O-Cl	0.975 O-H 1.690 O-Cl	102.71	102.94	102.5
HOBr	0.969 O-H 1.831 O-Br	0.964 O-H 1.828 O-Br	0.9613 O-H [b] 1.83415 O-Br [b]	102.81	102.97	102.33 [b]
HOI	0.969 O-H 1.996 O-I	0.963 O-H [c] 1.979 O-I [c]	0.967 O-H 1.9941 O-I	104.36	104.93 [c]	103.9
H_2O	0.961 O-H	0.9581 O-H	0.958 O-H [d]	104.12	104.42	104.48 [d]
H_2S	1.334 S-H	1.336 S-H	1.3356 S-H	92.18	92.23	92.12
H_2Se	1.459 Se-H	1.458 Se-H		90.74	90.94	
H_2Te	1.653 Te-H	1.648 Te-H [d]		90.38	90.32 [d]	

[a] Literature values are from reference 3 unless otherwise noted. [b] Reference 4. [c] Calculated with the aug-cc-pwCVTZ-DK basis set because aug-cc-pwCVDZ-DK does not exist for Te and I. [d] Reference 5.

Table A7.1.3. Polynomial Fit-Optimized Parameters for Diatomic Molecules at CCSD(T) Level with both aug-cc-pwCVQZ-DK [a] and aug-cc-pwCV5Z-DK Basis Sets

Species	Symmetry State	Bond Length (Å)			ω_e (cm^{-1})			$\omega_e x_e$ (cm^{-1})		
		awCQ-DK	awC5-DK	Lit	awCQ-DK	awC5-DK	Lit	awCQ-DK	awC5-DK	Lit
^2OCl	$X\ ^2\Pi_i$	1.570	1.567	1.56963	856.0	860.9	853.8	5.6	5.6	5.5
OCl^-		1.679	1.679		666.3	669.6		4.2	4.2	
^2OBr	$X1\ ^2\Pi_{3/2}$	1.720	1.717	1.7172	728.3	732.7	778.7	4.5	4.6	6.82
OBr^-		1.812	1.809		597.7	600.6		3.9	3.7	
^2OI	$X1\ ^2\Pi_{3/2}$	1.871		1.8676	674.2		681.47	4.3		4.29
OI^-		1.933			594.1			3.8		
$^3O_2$	$X\ ^3\Sigma_g^-$	1.206	1.205	1.20752	1601.7	1604.9	1580.193	11.0	11.0	11.981
$^2O_2^-$	$X\ ^2\Pi_{gi}$	1.347	1.345	1.35	1131.6	1136.1	1090	8.6	8.6	8.1
^2SH [b]	$X\ ^2\Pi_i$	1.340	1.340	1.3409	2700.3	2702.0	2711.6	48.1	48.3	59.9
SH^- [c]	$X\ ^1\Sigma^+$	1.343	1.343	1.3409	2650.6	2652.4	2711.6	51.9	52.1	59.9
^2SeH	$X1\ ^2\Pi_{3/2}$	1.463	1.463	1.475	2424.1	2425.2	2400	42.2	42.1	
SeH^-		1.467	1.467		2362.6	2363.9		45.3	45.0	
^2TeH		1.652			2146.3			36.2		
TeH^-		1.660			2077.6			38.8		
^2OH	$X\ ^2\Sigma^+$	0.970	0.969	0.96966	3741.9	3745.6	3737.76	83.1	83.4	84.881
OH^- [c]	$X\ ^1\Sigma^+$	0.964	0.964	0.970	3748.4	3753.1	3700	91.8	92.0	

[a] Basis sets are abbreviated in the table as awC*n*-DK. awCQ-DK is included because there is no awC5-DK basis set for iodine or tellurium. [b] Based on constants for SD.[6] [b] Ground state constants are indistinguishable from those of SH ($X\ ^2\Pi$).[7,8] [c] Literature value estimates are based on the analysis of photodetachment data.[9,10,11]

Table A7.1.4. FPD Data for Small Molecules (kcal/mol)

	CCSD(T)/ awCD-DK	CCSD(T)/ awCT-DK	CCSD(T)/ awCQ-DK	CCSD(T)/ awC5-DK	CBS (TQ)	CBS (Q5)	ZPE [a]	TC [a]
HOCl	150.01	160.40	163.85	164.89	165.84	165.99	-8.29	2.43
OCl^-	101.91	111.16	115.08	116.38	117.35	117.74	-1.13	2.13
HOBr	148.68	157.70	160.68	161.61	162.39	162.58	-8.02	2.46
OBr^-	100.05	108.33	111.81	113.03	113.82	114.30	-1.03	2.14
HOI		159.16	161.96		163.57		-7.81	2.48
OI^-		111.43	114.91		116.91		-1.00	2.14
H_2O	219.62	228.23	231.26	232.05	233.01	230.57	-13.44	2.37
OH^-	137.57	144.72	147.54	148.40	149.17	149.30	-5.42	2.07
H_2S	171.94	180.17	182.33	182.94	183.58	183.57	-9.71	2.38
SH^-	129.67	136.85	139.69	140.45	141.33	141.26	-3.89	2.07
H_2Se	154.81	160.83	162.36	162.82	163.24	163.31	-8.69	2.39
SeH^-	121.20	127.42	129.92	130.65	131.36	131.42	-3.47	2.07
H_2Te		141.91	143.18		143.92		-7.64	2.40
TeH^-		118.93	121.44		122.88		-3.05	2.07

[a] Calculated at the MP2/aug-cc-pVTZ level.

Table A7.1.5. Spin-Orbit Corrections and Respective $\Sigma D_{0,0K}$ Values in kcal/mol

	ΔE_{SO}	$\sum D_{0,0K}$ (TQ)	$\sum D_{0,0K}$ (Q5)	$\sum D_{0,0K}$ Literature Value
HOCl	-1.25 [a]	156.3	156.5	
^{2}OCl	-0.61 [b]	62.4	62.8	64.41,[b] 62.80 [d]
OCl^-	-0.62 [b]	116.1	116.5	
HOBr	-3.43 [a]	155.0	155.2	
OBr^-	-1.83 [b]	112.6	113.1	
HOI	0.47 [a]	156.2		
OI^-	-3.72 [b]	115.6		
H_2O	-0.22 [c]	219.3	216.9	
OH^-	-0.22 [c]	143.5	143.7	
H_2S	-0.56 [c]	173.3	173.3	
SH^-	-0.56 [c]	136.9	136.8	
H_2Se	-3.31 [a]	151.2	151.3	
SeH^-	-3.29 [a]	124.6	124.7	
H_2Te	-6.24 [a]	130.0		
TeH^-	-6.20 [a]	113.6		

[a] Calculated values for ΔE_{SO} are atomic SO values - molecular SO value. Atomic values for H (0.0 kcal/mol), O (-0.22 kcal/mol) are experimentally derived. Atomic values for Cl (-1.08 kcal/mol), Br (-3.91 kcal/mol), I (-7.73 kcal/mol) and for molecules containing Cl, Br, I were calculated with ADF. [b] Reference 12. [c] Reference 13. [d] Reference 14.

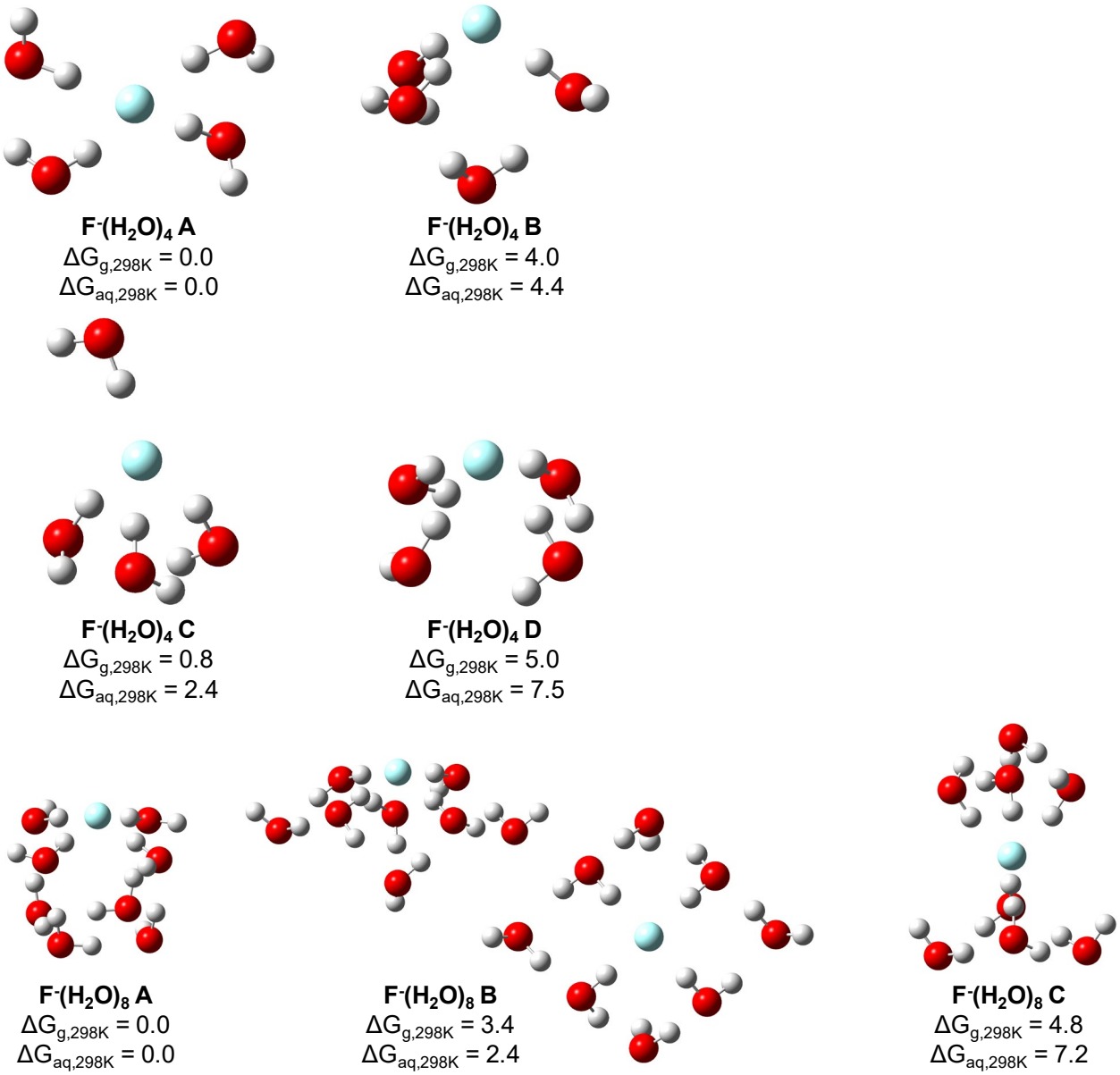


Figure A7.1.1. Conformers for $Cl^-(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol. Energies for chloride with the water tetramer are calculated at the CCSD(T)-F12b/aug-cc-pVDZ-F12 level and with the water octamer are calculated at the CCSD(T)/aug-cc-pVDZ. Solvation free energy is calculated for all species at the MP2/aug-cc-pVDZ/COSMO level.

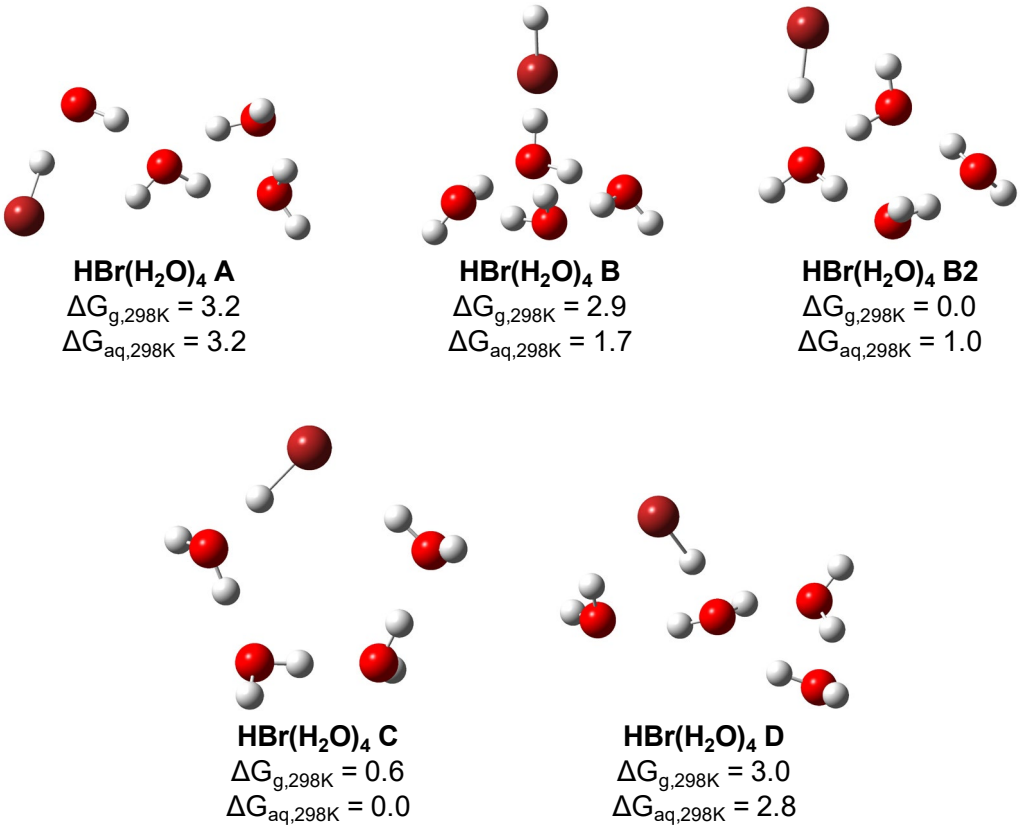


Figure A7.1.2. Conformers for $HBr(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol. Energies for HBr with the water tetramer are calculated at the CCSD(T)/aug-cc-pVTZ-PP level. Solvation free energy is calculated for all species at the MP2/aug-cc-pVDZ/COSMO level.

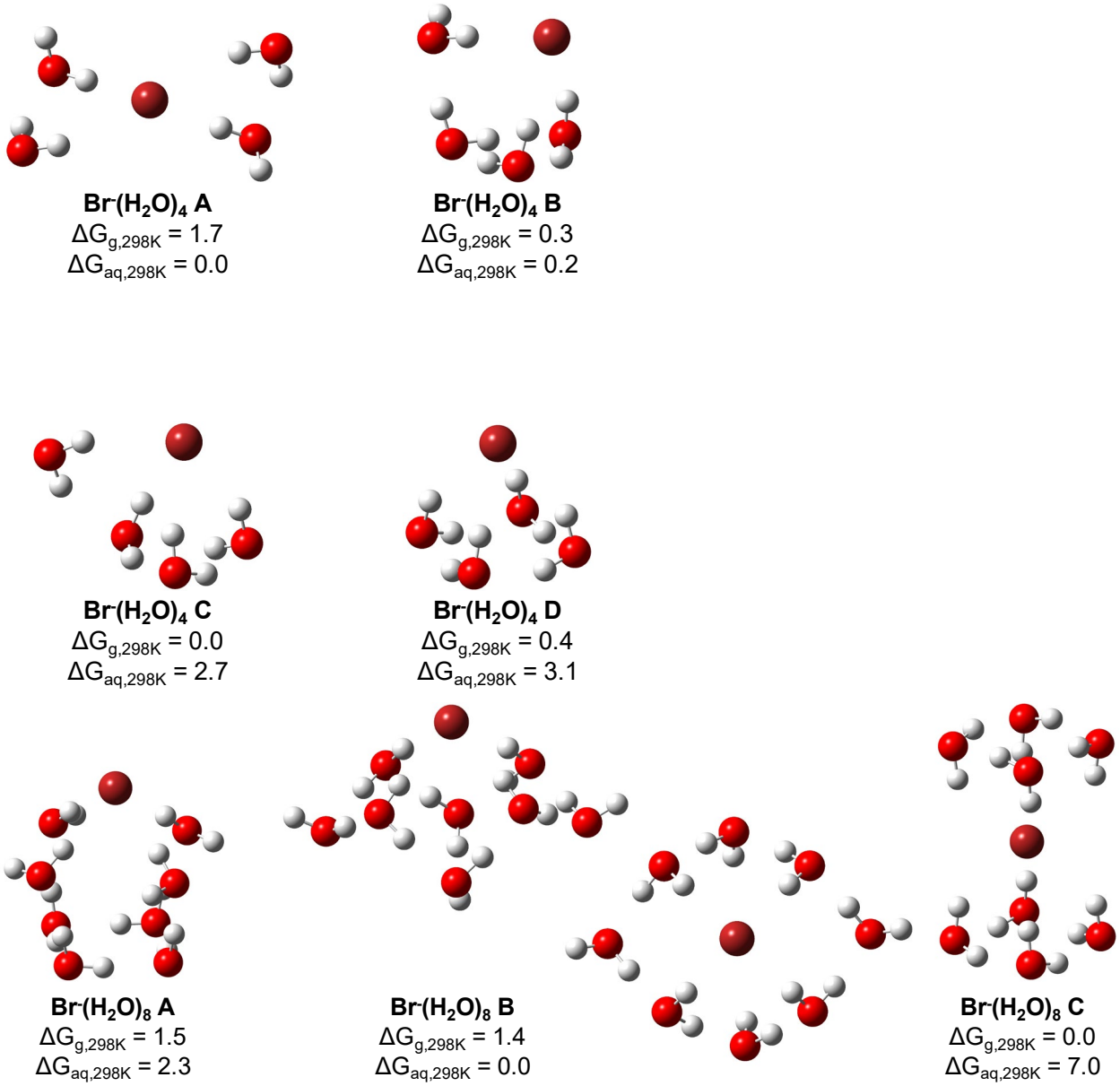


Figure A7.1.3. Conformers for $Br^-(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol. Energies for bromide with the water tetramer are calculated at the CCSD(T)/aug-cc-pVTZ-PP level and with the water octamer at the CCSD(T)/aug-cc-pVDZ-PP level. Solvation free energy is calculated for all species at the MP2/aug-cc-pVDZ/COSMO level.

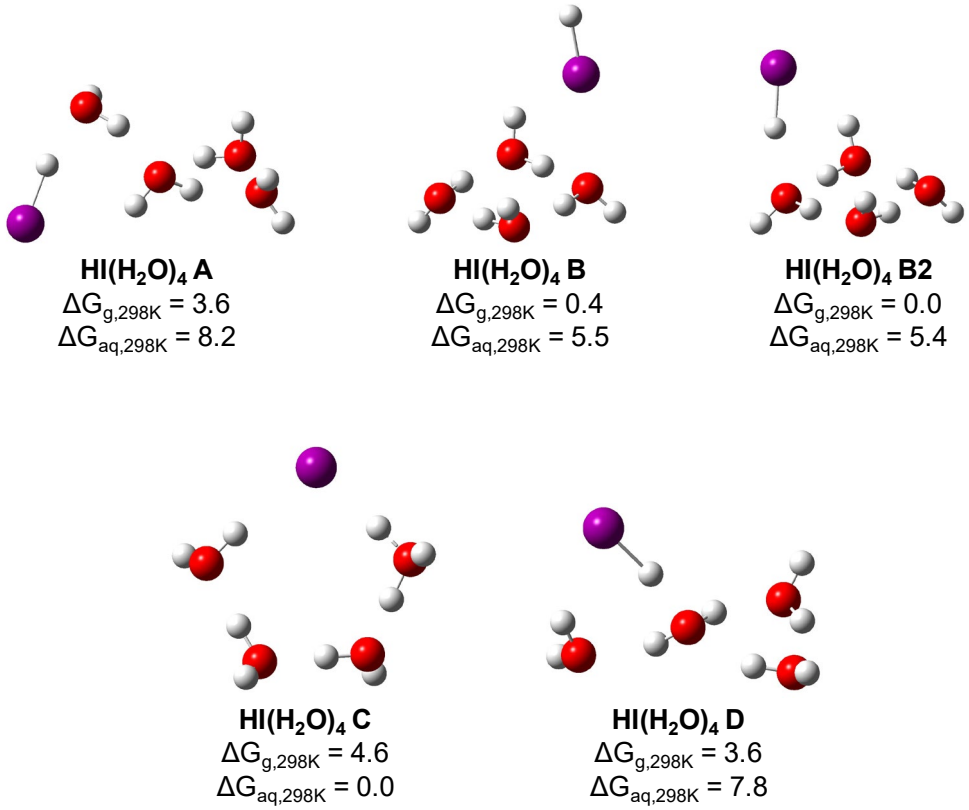


Figure A7.1.4. Conformers for $HI(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol. Energies for HI with the water tetramer are calculated at the CCSD(T)/aug-cc-pVDZ-PP level. Solvation free energy is calculated for all species at the MP2/aug-cc-pVDZ/COSMO level.

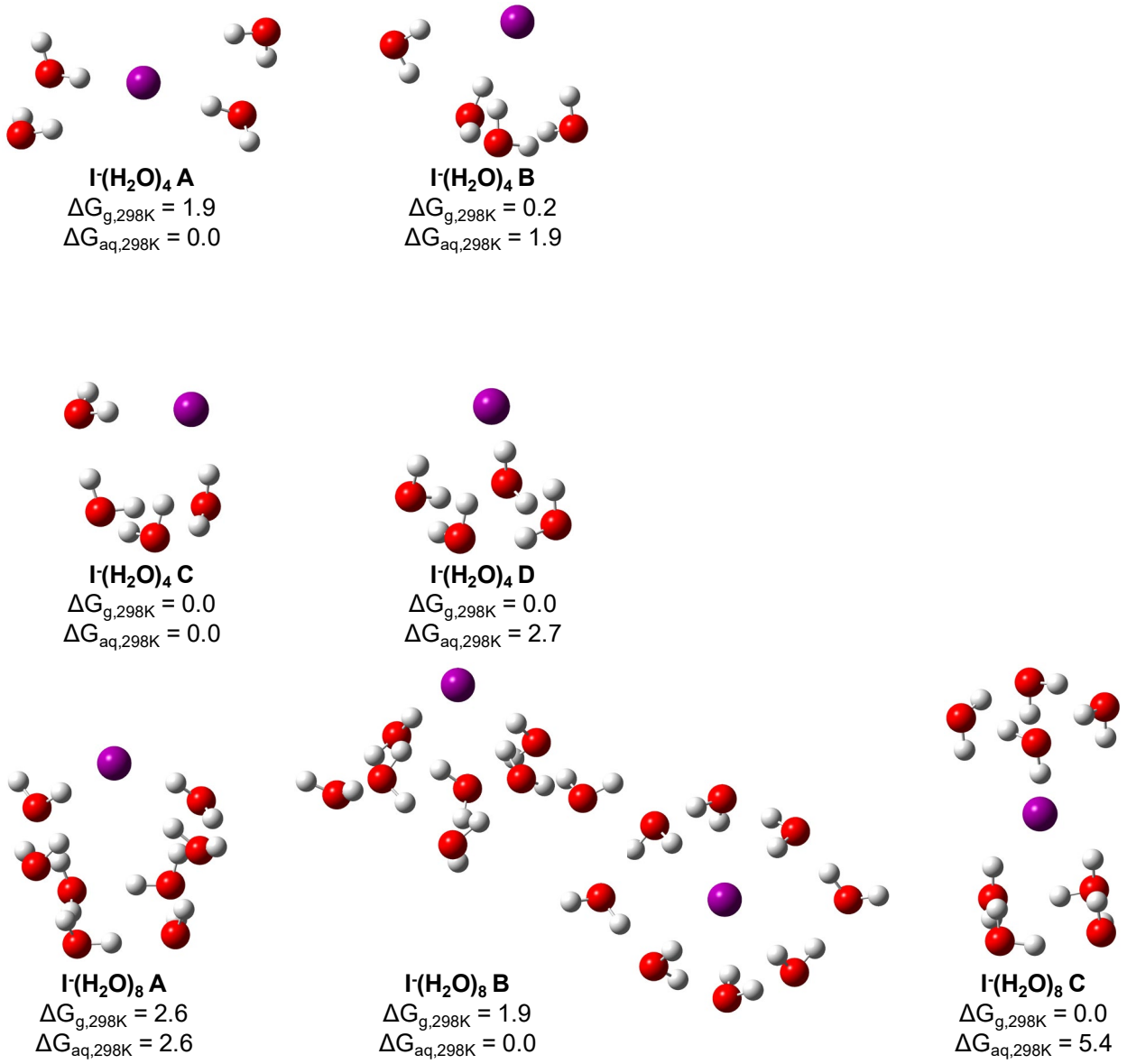


Figure A7.1.5. Conformers for $I^-(H_2O)_4$ with relative gas phase and aqueous free energies at 298 K in kcal/mol. Energies for iodide with the water tetramer are calculated at the CCSD(T)/aug-cc-pVTZ-PP level. Solvation free energy is calculated for all species at the MP2/aug-cc-pVDZ/COSMO level.

References

[1] Zhan, C.-G.; Dixon, D. A. Hydration of the Fluoride Anion: Structure and Absolute Hydration Free Energy from First-Principles Electronic Structure Calculations. *J. Phys. Chem. A* **2004**, *108*, 2020 – 2029.

[2] Zhan, C.-G.; Dixon, D. A. First-Principles Determination of the Absolute Hydration Free Energy of the Hydroxide Ion. *J. Phys. Chem. A* **2002**, *106*, 9737 – 9744.

[3] Lide, D. R. (E.) *CRC Handbook of Chemistry and Physics*, 86th Ed. CRC Press, Boca Raton, FL, USA, 2006.

[4] Koga, Y.; Takeo, H.; Kondo, S.; Sugie, M.; Matsumura, C.; McRae, G. A.; Cohen, E. A. The Rotational Spectra, Molecular Structure, Dipole Moment, and Hyperfine Constants of HOBr and DOBr. *J. Mol. Spectrosc*. **1989**, *138*, 467-481.

[5] NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, Ed. R. D. Johnson III, National Institute of Standards and Technology, Gaithersburg MD, 20899 (https://cccbdb.nist.gov/), retrieved 03/10/2020.

[6] Pathak, C. M.; Palmer, H. B. Chemical Excitation of the $A^2\Sigma^+$-$X^2\Pi$ System of SH and SH. *J. Mol. Spectrosc.* **1969**, *32*, 157 – 162.

[7] Cade, P. E. Hartree-Fock Wavefunctions, Potential Curves, and Molecular Properties for OH^- ($^1\Sigma^+$) and SH^- ($^1\Sigma^+$). *J. Chem. Phys.* **1967**, *47*, 2390 – 2406.

[8] Steiner, B. Photodetachment of Electrons from SH^-. *J. Chem. Phys.* **1968**, *49*, 5097 – 5104.

[9] Branscomb, L. M. Photodetachment Cross Section, Electron Affinity, and Structure of the Negative Hydroxyl Ion. *Phys. Rev.* **1966**, *148*, 11 – 18.

[10] Celotta, R. J.; Bennett, R. A.; Hall, J. L. Laser Photodetachment Determination of the Electron Affinities of OH, NH_2, NH, SO_2, and S_2. *J. Chem. Phys.* **1974**, *60*, 1740 – 1745.

[11] Hotop, H.; Patterson, T. A.; Lineberger, W. C. High Resolution Photodetachment Study of OH^- and OD^- in the Threshold Region 7000-6450 Å. *J. Chem. Phys.* **1974**, *60*, 1806 – 1812.

[12] Peterson, K. A.; Shepler, B. C.; Figgen, D.; Stoll, H. On the Spectroscopic and Thermochemical Properties of ClO, BrO, IO, and Their Anions. *J. Phys. Chem. A* **2006**, *110*, 13877 – 13883.

[13] Moore, C. E. *Atomic Energy Levels as Derived from the Analysis of Optical Spectra, Volume 1, H to V;* U.S. National Bureau of Standards Circular 467, COM-72-50282; U.S. Department of Commerce, National Technical Information Service: Washington, DC, 1949.

[14] Grant, D. J.; Garner, E. B. III; Matus, M. H.; Nguyen, M. T.; Peterson, K. A.; Francisco, J. S.; Dixon, D. A. Thermodynamic Properties of the XO_2, X_2O, XYO, X_2O_2, and XYO_2 (X, Y = Cl, Br, and I) Isomers. *J. Phys. Chem. A* **2010**, *114*, 4254 – 4265.

Appendix: Diatomic Electron Affinities

Computational Methods

All molecules were initially optimized at the density functional theory (DFT) level using the B3LYP[1,2] hybrid exchange-correlation functional with the aug-cc-pVTZ (aT) basis set[3,4] with pseudopotentials (PP) for Se, Br, Te, and I.[5,6] This geometry was then optimized at the higher-order MP2[7,8] level with the same aT(-PP) basis set. Frequencies were also calculated at each of these computational levels to verify that each optimized structure represents energetic minima. These calculations were performed with Gaussian 16.[9]

Beginning with the optimized bond length found at the MP2/aT(-PP) level, bond lengths, harmonic frequencies (ω_e) and anharmonic constants ($\omega_e\chi_e$) for the diatomics were obtained at the coupled cluster singles and doubles with perturbative triples (CCSD(T))[10,11,12] level using aug-cc-pwCV*n*Z-DK (awC*n*-DK, *n* = D, T, Q, 5) basis sets[13,14,15,16] with a 7-point Dunham expansion except for molecules with Te or I as discussed below.[17,18] For triatomic molecules, the MP2/aT optimized geometries were used for single point CCSD(T) calculations using the same basis sets. The scalar relativistic corrections (ΔE_{Rel}) at the second order Douglas-Kroll-Hess (DK)[19,20,21] level. All CCSD(T) calculations were performed using the Molpro 2015.1 software package.[22,23] All molecular species containing iodine were calculated at the CCSD(T)/awCT-DK and CCSD(T)/awCQ-DK levels due to a lack of awCD-DK and awC5-DK basis sets for Te and I. This limitation requires us to report extrapolations to the complete basis set (CBS) using a variety of calculated value combinations. The extrapolations to the CBS limit were obtained using equation (7.2.1) with TQ and for those without Te or I with Q5.[24,25]

$$E_n = E_{CBS} + A(n + 1/2)^{-4} \qquad (7.2.1)$$

For all molecules containing Te or I, we did the extrapolations using the pseudopotential and corresponding basis sets for T and Q and then corrected for any errors in the pseudopotential at the Q level using the DK results.

The above CCSD(T) calculations follow the Feller-Peterson-Dixon (FPD) approach.[26,27,28,29] The atomic spin-orbit corrections (ΔE_{SO}) for the heats of formation were taken from the experimental values (C = -0.08 kcal/mol, O = -0.22 kcal/mol, S = -0.56, Cl = -0.84, Se = -2.70, Br = -3.50, Te = -6.02, and I = -7.24).[30] For species containing chlorine, bromine, and iodine, alternate spin-orbit corrections were calculated with the ADF 2016 code[31,32] using the ZORA spin-orbit formalism[33,34,35,36,37] at the BLYP/TZ2P level or taken from previous second-order spin-orbit corrections reported in the literature.[38] The total atomization energies (TAE, $\Sigma D_{0,0K}$) for the heats of formation for oxalic acid and the carbonyl radical are calculated using equation (7.2.2).

$$\Sigma D_{0,0K} = \Delta E_{CBS} + \Delta E_{ZPE} + \Delta E_{SO} \quad (7.2.2)$$

Heats of formation at 0 K were calculated from these TAEs and experimental heats of formation of the H, C, and O atoms at 0 K from the JANAF tables (51.63 kcal/mol for H, 169.98 kcal/mol for C, 58.98 kcal/mol for O, 65.66 kcal/mol for S, 112.53 kcal/mol for N, 28.59 kcal/mol for Cl, 57.90 kcal/mol for Se, 28.18 kcal/mol for Br, 50.70 kcal/mol for Te, and 25.61 kcal/mol for I).[39] Heats of formation at 298 K were then calculated by following the Curtiss et al. procedure.[40] Gas phase electron affinities were also calculated from the TAEs and are reported as the -ΔTAE of the gas phase electron capture reaction.

Results

Optimized Geometries. Physical properties of all diatomic and species, including bond length, harmonic frequencies (ω_e) and anharmonic constants ($\omega_e\chi_e$) are reported in Table A7.2.1 at the

highest level of theory possible for each molecule. Because iodine- and tellurium-containing molecules cannot be reported at the awC5-DK level due to lack of basis set for these elements, all diatomic species are reported at the awCQ-DK level. Calculated values at the awCQ-DK and awC5-DK levels, where both are possible, are in excellent agreement with one another and with literature values.

Calculated values used for the FPD calculations to accurately estimate heats of formation for all diatomic and triatomic species are shown in Table A7.2.2 with the exception of spin-orbit corrections (ΔE_{SO}). Atomic experimental spin-orbit corrections were taken from the Moore tables[30] and are shown in Table 7.2.4 as well as the resulting TAE ($\sum D_{0,0K}$). Resulting enthalpy of formation values at 0 K and at 298 K calculated with the best available spin orbit correction are given in Table 7.2.5 and gas phase electron affinities for all diatomic molecules are given in Table 7.2.6.

Table A7.2.1. Polynomial Fit-Optimized Parameters for Diatomic Molecules at CCSD(T) Level with both aug-cc-pwCVQZ-DK [a] and aug-cc-pwCV5Z-DK Basis Sets [b]

Species	State	R (Å)			ω_e (cm^{-1})			$\omega_e x_e$ (cm^{-1})		
		awCQ-DK	awC5-DK	Lit	awCQ-DK	awC5-DK	Lit	awCQ-DK	awC5-DK	Lit
Cl_2	$^1\Sigma_g^+$	1.992	1.988	1.9879	558.0	562.2	559.72	2.6	2.6	2.675
$^2Cl_2^-$	$^2\Sigma$	2.576	2.570		261.3	262.8		1.7	1.8	
Br_2	$^1\Sigma_g^+$	2.282	2.278	2.28105	327.5	329.6	325.321	1.0	1.0	1.0774
$^2Br_2^-$	$^2\Sigma$	2.818	2.811		163.0	164.3		0.7	0.7	
I_2	$^1\Sigma_g^+$	2.565		2.6663	223.7		214.502	0.5		0.6147
$^2I_2^-$	$^2\Sigma$	3.180			115.7			0.4		
IBr	$^1\Sigma^+$	2.464		2.468989	275.7		268.640	0.7		0.814
$^2IBr^-$	$^2\Sigma$	2.984			140.6			0.4		
^2OCl	$^2\Pi_i$	1.570	1.567	1.56963	856.0	860.9	853.8	5.6	5.6	5.5
OCl^-	$^1\Sigma^+$	1.679	1.679		666.3	669.6		4.2	4.2	
^2OBr	$^2\Pi_{3/2}$	1.720	1.717	1.7172	728.3	732.7	778.7	4.5	4.6	6.82
OBr^-	$^1\Sigma^+$	1.812	1.809		597.7	600.6		3.9	3.7	
^2OI	$^2\Pi_{3/2}$	1.871		1.8676	674.2		681.47	4.3		4.29
OI^-	$^1\Sigma^+$	1.933			594.1			3.8		
$^3O_2$	$^3\Sigma_g^-$	1.206	1.205	1.20752	1601.7	1604.9	1580.193	11.0	11.0	11.981
$^2O_2^-$	$^2\Pi_{gi}$	1.347	1.345	1.35	1131.6	1136.1	1090	8.6	8.6	8.1
^2SH [c]	$^2\Pi_i$	1.340	1.340	1.3409	2700.3	2702.0	2711.6	48.1	48.3	59.9
SH^- [c]	$^1\Sigma^+$	1.343	1.343	1.3409	2650.6	2652.4		51.9	52.1	
^2SeH	$^2\Pi_{3/2}$	1.463	1.463	1.475	2424.1	2425.2	2400	42.2	42.1	
SeH^-	$^1\Sigma^+$	1.467	1.467		2362.6	2363.9		45.3	45.0	
^2TeH	$^2\Pi$	1.652			2146.3			36.2		
TeH^-	$^1\Sigma^+$	1.660			2077.6			38.8		
^2OH	$^2\Sigma^+$	0.970	0.969	0.96966	3741.9	3745.6	3737.76	83.1	83.4	84.881
OH^- [d]	$^1\Sigma^+$	0.964	0.964	0.970	3748.4	3753.1	3700	91.8	92.0	

[a] Basis sets are abbreviated in the table as awC*n*-DK. awCQ-DK is included because there is no awC5-DK basis set for iodine or tellurium. [b] Literature values obtained from NIST.[41] [c] Based on constants for SD.[42] Ground state constants are indistinguishable from those of SH ($X\ ^2\Pi$).[43,44] [d] Literature value estimates are based on the analysis of photodetachment data.[45,46,47]

Table A7.2.2. FPD Data for Small Molecules in kcal/mol

	CCSD(T)/ awCD -DK	CCSD(T)/ awCT -DK	CCSD(T)/ awCQ -DK	CCSD(T)/ awC5 -DK	CBS (TQ)	CBS (Q5)	CBS [a]	CBS + Corr [a]	ZPE [b]	TC [b]
Cl_2	46.56	54.97	57.65	58.69	59.19	59.78	60.00	59.07	-0.82	2.18
$^2Cl_2^-$	106.16	109.55	112.71	113.63	114.53	114.61			-0.38	2.36
Br_2	41.29	48.43	50.74	51.63	52.08	52.57	51.56	51.58	-0.49	2.31
$^2Br_2^-$	102.49	105.96	109.10	110.11	110.91	111.16			-0.24	2.45
I_2		44.02	46.63		48.14		47.97	47.66	-0.33	2.40
$^2I_2^-$		100.67	104.26		106.32				-0.17	2.51
IBr		47.56	49.84		51.16		51.39	50.72	-0.41	2.35
$^2IBr^-$		103.97	107.33		109.27				-0.21	2.48
^2OCl	52.23	60.38	62.83	63.72	64.24	64.66	64.75	63.41	-1.21	2.12
OCl^-	101.91	111.16	115.08	116.38	117.35	117.74			-1.13	2.13
^2OBr	48.57	54.83	56.58	57.30	57.59	58.06	58.34	56.59	-1.06	2.13
OBr^-	100.05	108.33	111.81	113.03	113.82	114.30			-1.03	2.14
^2OI		55.36	56.95		57.87		57.89	56.84	-0.96	2.15
OI^-		111.43	114.91		116.91				-1.00	2.14
$^3O_2$	107.28	114.92	118.24	119.14	120.15	120.10			-2.08	2.08
$^2O_2^-$	113.00	123.10	126.93	128.18	129.13	129.50			-1.57	2.09
^2SH	81.27	85.62	86.86	87.19	87.57	87.54			-3.95	2.07
SH^-	129.67	136.85	139.69	140.45	141.33	141.26			-3.89	2.07
^2SeH	73.27	76.64	77.51	77.78	78.02	78.06			-3.55	2.07
SeH^-	121.20	127.42	129.92	130.65	131.36	131.42			-3.47	2.07
^2TeH		67.93	68.68		69.11				-3.15	2.07
TeH^-		118.93	121.44		122.88				-3.05	2.07
^2OH	100.01	104.65	106.25	106.67	107.17	107.11			-5.43	2.07
OH^-	137.57	144.72	147.54	148.40	149.17	149.30			-5.42	2.07

[a] Values reported from previously published calculations as CBS extrapolated values with and without ZPE, SR, and CV corrections. SO Corrections are not included.[48] [b] Calculated at the MP2/aug-cc-pVTZ level.

Table A7.2.3. Spin-Orbit Corrections and Respective $\Sigma D_{0,0K}$ Values in kcal/mol

	ΔE_{SO}	$\sum D_{0,0K}$ (TQ)	$\sum D_{0,0K}$ (Q5)	$\sum D_{0,0K}$ Literature Value
Cl_2	0.00 [a]	58.4	59.0	57.3,[a] 57.39 [g]
$^2Cl_2^-$	0.00 [b]	114.2	114.2	
Br_2	-6.60 [a]	45.0	45.5	45.4,[a] 44.98 [g]
$^2Br_2^-$	-6.55 [b]	104.1	104.4	
I_2	-12.01 [c]	35.8		36.3,[a] 35.65 [g]
$^2I_2^-$	-12.02 [b]	94.1		
IBr	-8.93 [c]	41.8		41.79 [g]
$^2IBr^-$	-9.02 [b]	100.0		
^2OCl	-0.61 [e]	62.4	62.8	64.41,[e] 62.80 [g]
OCl^-	-0.62 [e]	116.1	116.5	
^2OBr	-2.05 [e]	54.5	54.9	57.99,[e] 54.54 [g]
OBr^-	-1.83 [e]	112.6	113.1	
^2OI	-3.35 [e]	53.6		57.24,[e] 53.49 [g]
OI^-	-3.72 [e]	115.6		
$^3O_2$	-0.44 [f]	117.6	117.6	
$^2O_2^-$	-0.44 [f]	127.1	127.5	
^2SH	-0.56 [f]	83.1	83.0	
SH^-	-0.56 [f]	136.9	136.8	
^2SeH	-0.77 [d]	73.7	73.7	
SeH^-	-3.29 [d]	124.6	124.7	
^2TeH	-1.62 [d]	64.3		
TeH^-	-6.20 [d]	113.6		
^2OH	-0.22 [f]	101.5	101.5	
OH^-	-0.22 [f]	143.5	143.7	

[a] Reference 38. [b] Values approximated by taking literature value and adding the difference between ADF-calculated values for related species. [c] Reference 49. [d] Calculated values for ΔE_{SO} are atomic SO values - molecular SO value. Atomic values for H (0.0 kcal/mol), O (-0.22 kcal/mol) are experimentally derived. Atomic values for Cl (-1.08 kcal/mol), Br (-3.91 kcal/mol), I (-7.73 kcal/mol) and for molecules containing Cl, Br, I were calculated with ADF. [e] Reference6. [f] Reference 30. [g] Reference 48.

Table A7.2.4. FPD Results for Small Molecules Using Best Spin Orbit Corrections with Comparable Literature Values in kcal/mol

	$\sum D_{0,0K}$ (TQ)	$\sum D_{0,0K}$ (Q5)	$\Delta H_{f,0K}$ (TQ)	$\Delta H_{f,0K}$ (Q5)	$\Delta H_{f,298K}$ (TQ)	$\Delta H_{f,298K}$ (Q5)	$\Delta H_{f,298K}$ Prev Pub Calc	$\Delta H_{f,298K}$ Expt
Cl_2	58.4	59.0	-1.2	-1.8	-1.2	-1.8	-0.2 [a]	0.0 [b,c]
$^2Cl_2^-$	114.2	114.2	-57.0	-57.0	-56.8	-56.9		-55.2 [d] -54.7±0.33 [e]
Br_2	45.0	45.5	11.4	10.9	7.8	7.3	7.9 [a]	7.4,[d] 7.39 ± 0.03 [c] 7.38±0.03 [e]
$^2Br_2^-$	104.1	104.4	-47.8	-48.0	-51.2	-51.4		-51.0±0.5 [e]
I_2	35.8		15.4		14.7		14.8 [f]	14.9,[d] 14.92 ± 0.02 [c] 14.918±0.00 [e]
$^2I_2^-$	94.1		-42.9		-43.6			-43.3 [d] -43.19±0.11 [e]
IBr	41.8		12.0		9.8		9.9 [f]	9.8,[d] 9.77 ± 0.02 [c] 9.74±0.02 [e]
$^2IBr^-$	100.0		-46.3		-48.3			-48.2 [d]
^2OCl	62.4	62.8	25.1	24.7	25.1	24.7	24.8 [a]	24.2,[d] 24.29 ± 0.03 [b,g] 24.31±0.01 [e]
OCl^-	116.1	116.5	-28.5	-28.9	-28.5	-28.9		-28.27±0.03 [e]
^2OBr	54.5	54.9	32.7	32.2	30.8	30.4	30.8 [f]	30.2 ± 0.4,[g] 29.5 ± 0.3 [h] 29.54±0.07 [e]
OBr^-	112.6	113.1	-25.5	-26.0	-27.3	-27.8		-24.77±0.13 [e]
^2OI	53.6		31.0		30.6		30.6 [f]	27.7 ± 1.2,[g] 31 ± 3 [i] 29.89±0.20 [e]
OI^-	115.6		-31.0		-31.5			-24.93±0.21 [e]
$^3O_2$	117.6	117.6	0.3	0.4	0.3	0.4		0.0 [g]
$^2O_2^-$	127.1	127.5	-9.2	-9.5	-9.2	-9.5		-10.3 [d] -10.21±0.04 [e]
^2SH	83.1	83.0	34.2	34.3	34.2	34.3	34.2 [j]	33.3,[d] 34.13 ± 0.68,[g] 33.56 ± 0.8 [l]
SH^-	136.9	136.8	-19.6	-19.5	-19.6	-19.5		
^2SeH	73.7	73.7	35.8	35.8	35.6	35.5		
SeH^-	124.6	124.7	-15.1	-15.1	-15.3	-15.4		
^2TeH	64.3		38.0		37.6			
TeH^-	113.6		-11.3		-11.7			

^{2}OH	101.5	101.5	9.1	9.2	9.1	9.2		9.3 [d] 8.96±0.006 [e]
OH^-	143.5	143.7	-32.9	-33.0	-32.9	-33.0		-33.24±0.006 [e]

[a] Reference 50. [b] Reference 51. [c] Reference 39. [d] Reference 41. [e] Reference 52. [f] Reference 48. [g] Reference 53. [h] Reference 54. [i] Reference 55. [j] Reference 56. [k] Reference 57. [l] Reference 58.

Table A7.2.5. Gas Phase Electron Affinities of Small Molecules Calculated at the CCSD(T) Level with aug-cc-pwCVnZ-DK or CBS Extrapolated TAE + ZPE With and Without Spin-Orbit Corrections in eV

Species	CBS (TQ)	CBS (Q5)	Experimental Gas Phase EA [a]
OCl	2.29	2.29	2.2775
OBr	2.37	2.37	2.3530
OI	2.41		2.3780
SH	2.33	2.33	2.31473
SeH	2.21	2.21	2.212525
TeH	2.14		2.112
CN	3.94	3.95	3.8620
O_2	0.41	0.43	0.4480
Cl_2	2.44	2.42	2.50 2.382 [b]
Br_2	2.57	2.56	2.42002 2.54±0.02 [b]
I_2	2.54		2.5240
IBr	2.53		2.5120
NO	0.12	0.12	0.024,[c] 0.026 [d]
OH	1.82	1.83	1.82767

[a] Reference 41 unless otherwise noted. [b] Reference 52. [c] Reference 59. [d] Reference 60.

References

[1] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.

[2] Lee, C.; Yang, W.; Parr, R. G. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Physical Review B.* **1988**, *37*, 785 – 789.

[3] Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys*. **1989**, *90*, 1007 – 1023.

[4] Kendall, R. A.; Dunning, T. H. Jr.; Harrison, R. J., Electron Affinities of the First Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys*. **1992,** *96*, 6796 – 6806.

[5] Peterson, K. A.; Yousaf, K. E. Molecular Core-Valence Correlation Effects Involving the Post-d Elements Ga-Rn: Benchmarks and New Pseudopotential-Based Correlation Consistent Basis Sets. *J. Chem. Phys.* **2010**, *133*, 174116-1 – 174116-8.

[6] Peterson, K. A.; Shepler, B. C.; Figgen, D.; Stoll, H. On the Spectroscopic and Thermochemical Properties of ClO, BrO, IO, and Their Anions. *J. Phys. Chem. A* **2006**, *110*, 13877 – 13883.

[7] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 0618 – 0622.

[8] Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical Models Incorporating Electron Correlation. *Int. J. Quantum. Chem.* **1976**, *10*, 1 – 19.

[9] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision A.03, Gaussian, Inc., Wallingford CT, 2016.

[10] Adler, T. B.; Knizia, G.; Werner, H.-J. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Phys. Chem.* **2007,** *127,* 221106-1 – 221106-4.

[11] Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009,** *130,* 054104-1 – 054104-20.

[12] Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-ζ Quality Coupled-Cluster Correlation Energies with Triple- ζ Basis Sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921 – 1930.

[13] de Jong, W. A.; Harrison, R. J.; Dixon, D. A. Parallel Douglas-Kroll Energy and Gradients in NWChem: Estimating Scalar Relativistic Effects Using Douglas-Kroll Contracted Basis Sets. *J. Chem. Phys*. **2001**, *114*, 48 – 53.

[14] Bross, D. H.; Peterson, K. A. Correlation Consistent, Douglas-Kroll-Hess Relativistic Basis Sets for the 5p and 6p Elements. *Theor. Chem. Acc*. **2014**, *133*, 1434-1 – 1434-12.

[15] Woon, D. E.; Dunning, T. H. Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-Valence Basis Sets for Boron Through Neon. *J. Chem. Phys.* **1995**, *103*, 4572 – 4585.

[16] Peterson, K. A.; Dunning, T. H. Accurate Correlation Consistent Basis Sets for Molecular Core-Valence Correlation Effects: The Second Row Atoms Al-Ar, and the First Row Atoms B-Ne Revisited. *J. Chem. Phys.* **2002**, *177*, 10548 – 10560.

[17] Dunham, J. L. The Energy Levels of a Rotating Vibrator. *Phys. Rev.* **1932**, *41*, 721 – 731.

[18] Dunham, J. L. The Wentzel-Brillouin-Kramers Method of Solving the Wave Equation. *Phys. Rev.* **1932**, *41*, 713 – 720.

[19] Douglas, M.; Kroll, N. M. Quantum Electrodynamical Corrections to the Fine Structure of Helium. *Ann. Phys.* **1974**, *82*, 89 – 155.

[20] Hess, B. A. Applicability of the No-Pair Equation with Free-Particle Projection Operators to Atomic and Molecular Structure Calculations. *Phys. Rev. A.* **1985**, *32*, 756 – 763.

[21] Hess, B. A. Relativistic Electronic-Structure Calculations Employing a Two-Component No-Pair Formalism with External-Field Projection Operators. *Phys. Rev. A.* **1986**, *33*, 3742 – 3748.

[22] Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, T.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.;Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. MOLPRO, version 2015.1, a package of *ab initio* programs, See http://www.molpro.net. Accessed March 1, 2016.

[23] Werner H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242 – 253.

[24] Martin, J. M. L. Ab Initio Total Atomization Energies of Small Molecules - Towards the Basis Set Limit. *Chem. Phys. Lett.* **1996**, *259*, 669 – 678.

[25] Feller, D.; Peterson, K.A.; Hill, J. G. On the Effectiveness of CCSD(T) Complete Basis Set Extrapolations for Atomization Energies. *J. Chem. Phys.* **2011***, 135,* 044102-1 – 044102-18.

[26] Dixon, D. A.; Feller, D.; Peterson, K. A. A Practical Guide to Reliable First Principles Computational Thermochemistry Predictions Across the Periodic Table. In *Annual Reports in Computational Chemistry*, Ralph, A. W., Ed.; Elsevier: 2012; Vol. 8, pp 1 – 28.

[27] Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy: Current Strategies and Future Challenges. *Theor. Chem. Acc.* **2012**, *131*, 1 – 20.

[28] Feller, D.; Peterson, K. A.; Dixon, D. A. Further Benchmarks of a Composite, Convergent, Statistically Calibrated Coupled Cluster Based Approach for Thermochemical and Spectroscopic Studies. *Mol. Phys.* **2012**, *110*, 2381 – 2399.

[29] Feller, D.; Peterson, K. A.; Dixon, D. A. The Impact of Larger Basis Sets and Explicitly Correlated Coupled Cluster Theory on the Feller-Peterson-Dixon Composite Method. In *Annual Reports in Computational Chemistry*, Vol. 12, Ed. D. A. Dixon, Elsevier, Amsterdam, 2016, pp. 47 – 78.

[30] Moore, C. E. *Atomic Energy Levels as Derived from the Analysis of Optical Spectra, Volume 1, H to V;* U.S. National Bureau of Standards Circular 467, COM-72-50282; U.S. Department of Commerce, National Technical Information Service: Washington, DC, 1949.

[31] *ADF Release 2016, ADF Users Guide*, SCM, Theoretical Chemistry, Vrije Universiteit: Amsterdam. https://www.scm.com/doc/ADF/downloads/ADF.pdf. Accessed January 5, 2018.

[32] te Velde, G.; Bickelhaupt, F. M.; van Gisbergen, S. J. A.; Fonseca Guerra, C.; Baerends, E. J.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931 – 967.

[33] van Lenthe, E.; Ehlers, A. E.; Baerends, E. J. Geometry Optimization in the Zero Order Regular Approximation for Relativistic Effects *J. Chem. Phys*. **1999**, *110*, 8943 – 8953.

[34] van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic Regular Two-Component Hamiltonians, *J. Chem. Phys* **1993**, *99*, 4597 – 4610.

[35] van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic Total Energy Using Regular Approximations, *J. Chem. Phys* **1994**, *101*, 9783 – 9792.

[36] van Lenthe, E.; Snijders, J. G.; Baerends, E. J. The Zero-Order Regular Approximation for Relativistic Effects: The Effect of Spin-Orbit Coupling in Closed Shell Molecules *J. Chem. Phys* **1996**, *105*, 6505 – 6516.

[37] van Lenthe, E.; van Leeuwen, R.; Baerends, E. J.; Snijders, J. G. Relativistic Regular Two-Component Hamiltonians *Int. J. Quantum Chem*. **1996**, *57*, 281 – 293.

[38] Feller, D.; Peterson, K. A.; de Jong, W. A.; Dixon, D. A. Performance of Coupled Cluster Theory in Thermochemical Calculations of Small Halogenated Compounds. *J. Chem. Phys.* **2003**, *118*, 3510 – 3522.

[39] Chase, M. W., Jr. *NIST-JANAF Thermochemical Tables*, 4th ed. *J. Phys. Chem. Ref. Data.*, *Mono. 9*; American Institute of Physics: Woodbury, NY, 1998; Suppl 1.

[40] Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies of Formation. *J. Chem. Phys.* **1997**, *106*, 1063 – 1079.

[41] NIST Chemistry Webbook, NIST Standard Reference Database Number 69, Eds. P.J. Lindstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (https://webbook.nist.gov/), retrieved 4/15/2020.

[42] Pathak, C. M.; Palmer, H. B. Chemical Excitation of the $A^2\Sigma^+$-$X^2\Pi$ System of SH and SH. *J. Mol. Spectrosc.* **1969**, *32*, 157 – 162.

[43] Cade, P. E. Hartree-Fock Wavefunctions, Potential Curves, and Molecular Properties for OH^- ($^1\Sigma^+$) and SH^- ($^1\Sigma^+$). *J. Chem. Phys.* **1967**, *47*, 2390 – 2406.

[44] Steiner, B. Photodetachment of Electrons from SH^-. *J. Chem. Phys.* **1968**, *49*, 5097 – 5104.

[45] Branscomb, L. M. Photodetachment Cross Section, Electron Affinity, and Structure of the Negative Hydroxyl Ion. *Phys. Rev.* **1966**, *148*, 11 – 18.

[46] Celotta, R. J.; Bennett, R. A.; Hall, J. L. Laser Photodetachment Determination of the Electron Affinities of OH, NH_2, NH, SO_2, and S_2. *J. Chem. Phys.* **1974**, *60*, 1740 – 1745.

[47] Hotop, H.; Patterson, T. A.; Lineberger, W. C. High Resolution Photodetachment Study of OH^- and OD^- in the Threshold Region 7000-6450 Å. *J. Chem. Phys.* **1974**, *60*, 1806 – 1812.

[48] Grant, D. J.; Garner, E. B. III; Matus, M. H.; Nguyen, M. T.; Peterson, K. A.; Francisco, J. S.; Dixon, D. A. Thermodynamic Properties of the XO_2, X_2O, XYO, X_2O_2, and XYO_2 (X, Y = Cl, Br, and I) Isomers. *J. Phys. Chem. A* **2010**, *114*, 4254 – 4265.

[49] Shepler, B. C.; Balabanov, N. B.; Peterson, K. A. Ab Initio Thermochemistry Involving Heavy Atoms: An Investigation of the Reactions Hg + IX (X = I, Br, Cl, O). *J. Phys. Chem. A* **2005**, *109*, 10363 – 10372.

[50] Matus, M. H.; Nguyen, M. T.; Peterson, K. A.; Francisco, J. S.; Dixon, D. A. $ClClO_2$ is the Most Stable Isomer of Cl_2O_2. Accurate Coupled Cluster Energetics and Electronic Spectra of Cl_2O_2 Isomers. *J. Phys. Chem. A*, **2008**, *112*, 9623 – 9627.

[51] Bloss, W. J.; Nickolaisen, S. L.; Salawitch, R. J.; Friedl, R. R.; Sander, S. P. Kinetics of the ClO Self-Reaction and 210 nm Absorption Cross Section of the ClO Dimer. *J. Phys. Chem. A* **2001**, *105*, 11226 – 11239.

[52] Ruscic, B.; Bross, D. H. Active Thermochemical Tables (ATcT) values based on ver. 1.122g of the Thermochemical Network (2019); available at ATcT.anl.gov.

[53] Sander, S. P.; Friedl, R. R.; Ravishankara, A. R.; Golden, D. M.; Kolb, C. E.; Kurylo, M. K.; Huie, R. E.; Orkin, V. L.; Molina, M. J.; Moortgat, G. K.; Finlayson-Pitts, B. J. *Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies: Evaluation Number 14*; JPL Publication 02-25; National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, 2003; http://jpldataeval.jpl.nasa.gov/pdf/JPL_02-25_rev02.pdf.

[54] Kim, H.; Dooley, K. S.; Johnson, E. R.; North, S. W. Photodissociation of the BrO Radical Using Velocity Map Ion Imaging: Excited State Dynamics and Accurate D_0^0 (BrO) Evaluation. *J. Chem. Phys.* **2006**, *124*, 134304-1 – 134304-8.

[55] Radlein, D. S. A. G.; Whitehead, J. C.; Grice, R. Bond Energy of the IO Radical from Molecular Beam Reactive Scattering Measurements. *Nature* **1975**, *253*, 37.

[56] Grant, D. J.; Dixon, D. A.; Francisco, J. S.; Feller, D.; Peterson, K. A. Heats of Formation of the $H_{1,2}O_mS_n$ (m, n = 0-3) Molecules from Electronic Structure Calculations. *J. Phys. Chem. A* **2009**, *113*, 11343 – 11353.

[57] Ruscic, B. Interim private communication of unpublished data based on Active Thermochemical Tables, ATcT ver. 1.25 and the Core (Argonne) Thermochemical Network ver. 1.0562006.

[58] Gurvich, L. V.; Veyts, I. V.; Alcock, C. B. *Thermodynamic Properties of Individual Substances*, 4th ed.; Hemisphere, New York, 1989.

[59] Siegel, M. W.; Celotta, R. J.; Hall, J. L.; Bennett, R. A. Molecular Photodetachment Spectrometry. I. The Electron Affinity of Nitric Oxide and the Molecular Constants of NO^-. *Phys. Rev. A* **1972**, *6*, 607 – 631.

[60] Arrington, C. A.; Dunning, T. H., Jr.; Woon, D. E. Electron Affinity of NO. *J. Phys. Chem. A* **2007**, *111*, 11185 – 11188.

CHAPTER 8

CONCLUSIONS

Accurate studies of the properties of anions in both the gas phase and in aqueous solution are critically important in a wide variety of chemical fields, from biochemistry to geochemistry to catalysis. Many of these properties, such as bond dissociation energies, electron affinities, gas phase and aqueous acidities, heats of formation, etc. are difficult, if not impossible, to determine experimentally for many species. In this dissertation, several representative cases are presented for the importance of computational chemistry in providing accurate results that complement and/or reinforce experimental results. Solvation properties have been explored for both explicit and implicit aqueous solvation; the inclusion of both of these forms of solvation is critical for accurately modeling many aqueous systems.

In chapter 2, the enthalpies of product formation for both direct peptide backbone bond dissociation and fragment rearrangement pathways have been predicted at the composite correlated molecule orbital theory G3(MP2) level for the dissociation of the $[M-H]^-$ precursor anions of glycine, alanine, the dipeptides, and the amide forms of these species. The lowest energy isomers for bond cleavage and rearrangement pathways were determined by extensive sampling of the available structures for each fragment using density functional theory. The G3(MP2) results are compared to the experimental low-energy collision-induced dissociation (CID) negative ion mode mass spectra. The y_1^- anion, $NH_2CHRCO_2^-$ was the most intense product ion in the CID spectra for the dipeptide species and their amides. Among

amide species, $N{\equiv}C{-}O^-$ was universally observed and the enthalpies for generation of this anion are less than ± 5 kcal/mol. Amino acids and their amides consistently lost neutral HN=CH−R (R = H or CH_3). Loss of the C-terminal carbon as CO_2 was only observed for diglycine and dialanine but loss of CO_2 or HN=C=O were not observed for the amino acids or any of the amide structures. The hydrogen at the C-terminus of the amino acid and dipeptide amide molecules facilitated the observable loss of H_2O in the experimental spectra. Loss of H_2O was not observed for non-amide species. Several energetically favorable fragmentation pathways were not detected in the low-energy CID mass spectra for these deprotonated anions. Selected transition states were calculated, and show that a number of the observed or low-energy pathways involve multistep mechanisms.

Gas-phase acidities (GA or ΔG_{acid}) of acidic di- and tripeptides are determined for the first time in chapter 3. The peptides studied are composed of inert alanine (A) residues and one X residue of either aspartic acid (D) or glutamic acid (E): AX, XA, AAX, AXA, and XAA. Experimental GAs were measured by the thermokinetic method of deprotonation ion/molecule reactions in a Fourier transform ion cyclotron resonance mass spectrometer. Calculated GAs were obtained by composite correlated molecular orbital theory at the G3(MP2) level for deprotonation of carboxylic acid groups both at the C-terminus and at the side chain. Excellent agreement was found between experimental and calculated GA values. There is a slight preference for peptides with D being more acidic than analogous peptides with E, which agrees with the GAs of the corresponding amino acids. Experiments showed that peptides are more acidic (lower numerical GA values) when the acidic residue is located at the C-terminus (i.e., AX or AAX). The lowest energy form of deprotonated AAE has a unique structure where the longer side chain of E allows the two carboxylates, which are in close proximity, to share the proton.

The tripeptides are less acidic (higher GA value) by 3−7 kcal/mol when the acidic residue is in the center. The tripeptides are more acidic (by 2−10 kcal/mol) than dipeptides containing the same acidic residue at the same location.

In chapter 4, we report the gas-phase and aqueous solution properties of neutral and anionic clusters of CO_2 with 3, 4, and 8 explicit H_2O molecules, calculated at the coupled cluster (CCSD(T)) level plus a self-consistent reaction field. Anionic clusters with the radical electron density localized on the carbon of the CO_2 molecule rather than localized on the H_2O molecules are more favorable energetically by 10−20 kcal/mol in the gas phase (ΔH_{gas}(298 K)) and 20−30 kcal/mol in aqueous solution (ΔG_{aq}(298 K)). The most favorable structures are those with the largest number of strong hydrogen bonds between the CO_2^- and the explicit H_2O molecules. Adiabatic electron affinities were calculated in the gas phase and in aqueous solution for the microsolvated anion. The adiabatic electron affinity of aqueous CO_2^- is predicted to be 2.35 ± 0.08 eV and is converged with as few as 3 explicit H_2O molecules plus a self-consistent reaction field. The calculated reduction potential for CO_2/CO_2^- is -1.93 ± 0.08 eV versus SHE (-4.28 eV), which is in good agreement with the experimental value of -1.90 eV. The EA of aqueous CO_2 is significantly greater than the aqueous solvation free energy of the electron. The vertical attachment energies to CO_2 and the vertical detachment energies from CO_2^- were calculated. The solvated CO_2^- anion is substantially bent to 135°, which requires 1.52 eV. The large energy required for bending in combination with the vertical detachment and attachment energies shows that substantial local solvent reorganization occurs on detachment or attachment of an electron to solvated CO_2. The formation of aqueous $C_2O_4^{2-}$ from CO_2^- was also explored, and dimerization is predicted to occur.

The alanine transaminase enzyme catalyzes the transfer of an amino group from alanine to α-ketoglutarate to produce pyruvate and glutamate. In chapter 5, isotope fractionation factors (IFFs) for the reaction $^{+}H_3NCH(CH_3)COO^{-} + {}^{-}OOCCH_2CH_2C(O)COO^{-} \leftrightarrow CH_3C(O)COO^{-} + {}^{+}H_3NCH(CH_2CH_2COO^{-})COO^{-}$ (zwitterionic neutral alanine + doubly deprotonated α-ketoglutarate ↔ pyruvate + zwitterionic glutamate anion) were calculated from the partition functions of explicitly and implicitly solvated molecules at 298 K. Calculations were done for alanine (non-charge separated, zwitterion, deprotonated), pyruvic acid (neutral, deprotonated), glutamic acid (non-charge separated, zwitterion, deprotonated, doubly deprotonated), and α-ketoglutaric acid (neutral, deprotonated, doubly deprotonated). The computational results, calculated from gas phase- and aqueous-optimized clusters with explicit H_2O molecules at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVDZ/COSMO levels, respectively, predict that substitution of ^{13}C at the C2 position of alanine and pyruvic acid and their various forms leads to the C2 position of pyruvic acid/pyruvate being enriched in $^{13}C/^{12}C$ ratio by 9‰. Simpler approaches that estimate the IFFs based solely on changes in the zero-point energies (ZPEs) are consistent with the higher-level model. ZPE-based IFFs calculated for simple analogues formaldehyde and methylamine (analogous to the C2 positions of pyruvate and alanine, respectively) predict a ^{13}C enrichment in formaldehyde of 7−8‰ at the MP2/aug-cc-pVDZ and aug-cc-pVTZ levels. A simple predictive model using canonical functional group frequencies and reduced masses for ^{13}C exchange between $R_2C{=}O$ and $R_2CH{-}NH_2$ predicted enrichment in $R_2C{=}O$ that is too large by a factor of two but is qualitatively accurate compared with the more sophisticated models. Our models are all in agreement with the expectation that pyruvate and formaldehyde will be preferentially enriched in ^{13}C because of the strength of their >C=O bond

relative to that of ≡C−NH_2 in alanine and methylamine. $^{13}C/^{12}C$ substitution is also modeled at the methyl and carboxylic acid sites of alanine and pyruvic acid, respectively.

In chapter 6, we employ the supermolecule-continuum approach with water clusters up to n = 16 H_2O molecules to predict the absolute hydration free energies at 298 K (ΔG_{hyd}) of both hydrogen (H•) and hydride (H^-) to be 4.6 and -78 ± 2 kcal/mol, respectively. These values are combined with a high accuracy prediction of the gas phase electron affinity ($\Delta G_{gas,298K}$ = -16.9 kcal/mol) to determine the aqueous electron affinity of H• of 99.5 ± 2 kcal/mol, which yields a reduction potential for H• of -0.03 ± 0.1 eV versus SHE (-4.28 eV). This value agrees within 0.2 V with most estimates obtained using a wide variety of approaches. These results can be used to improve the absolute hydricity scale in water which provides additional insights into how a putative hydride interacts with solvent, but do not change the ability to predict the relative reactivity of two species using relative hydricity scales. This work will enable us to develop an absolute hydricity scale for use in homogeneous and, potentially, heterogeneous catalyst design.

Chapter 7 describes a combination of Feller-Peterson-Dixon (FPD) high-accuracy gas phase calculations with supermolecule-continuum free energy of hydration values to calculate high accuracy gas phase and aqueous acidities for HX (X = F, Cl, Br, I), HOX, and H_2Y (Y = O, S, Se, Te). The heats of formation for both the acids and their corresponding anions and the gas phase acidities calculated using the FPD method are in very good agreement with experimental values with differences of 1.2 kcal/mol or less when compared with high accuracy experimental values from the ATcT tabulation. Using absolute hydration free energies determined with the supermolecule-continuum approach with 4 explicit H_2O molecules for the neutral acid species and with 4 and 8 H_2O molecules for the anions, aqueous acidities, $\Delta\Delta G_{aq}$, and pK_a values were calculated for all species. These values also show remarkable agreement with high accuracy

experimental values, typically within 1-2 kcal/mol or 1-2 pKa units, with a few exceptions, all in the n = 5 period.

References

[1] Simons, J.; Smith, W. D. Theory of Electron Affinities of Small Molecules. *J. Chem. Phys.* **1973**, *58*, 4899 – 4907.

[2] Pegg, D. J. Structure and Dynamics of Negative Ions. *Rep. Prog. Phys.* **2004**, *67*, 857 – 905.

[3] Rienstra-Kiracofe, J. C.; Tschumper, G. S.; Schaefer, H. F.; Nandi, S.; Ellison, G. B. Atomic and Molecular Electron Affinities: Photoelectron Experiments and Theoretical Computations. *Chem. Rev.* **2002**, *102*, 231 – 282.

[4] Lineberger, W. C. Atomic Electron Affinities. *IEEE T. Nucl. Sci.* **1976**, *NS-23*, 934 – 935.

[5] Simons, J. Theoretical Study of Negative Molecular Ions. *Annu. Rev. Phys. Chem.* **2011**, *62*, 107 – 128.

[6] Jensen, F. *Introduction to Computational Chemistry*, 2nd Ed; John Wiley & Sons, Ltd: West Sussex, England, 2007; pp 80 – 191.

[7] Jensen, J. H. *Molecular Modeling Basics*; CRC Press: Boca Raton, FL, 2010; pp 25 – 76.

[8] Dunning, Thomas H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys*. **1989**, *90*, 1007 – 1023.

[9] Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. Optimization of Gaussian-Type Basis Sets for Local Spin Density Functional Calculations Part I. Boron through Neon, Optimization Technique and Validation. *Can. J. Chem.* **1992**, *70*, 560 – 571.

[10] Sosa, C.; Andzelm, J.; Elkin, B. C.; Wimmer, E.; Dobbs, K. D.; Dixon, D.A. A Local Density Functional Study of the Structure and Vibrational Frequencies of Molecular Transition-Metal Compounds. *J. Phys. Chem.* **1992**, *96*, 6630 – 6636.

[11] Pople, J. A.; Seeger, R.; Krishnan, R. Variational Configuration Interaction Methods and Comparison with Perturbation Theory. *Int. J. Quantum Chem.* **1977**, *12*, 149 – 163.

[12] Raghavachari, K.; Schlegel, H. B.; Pople, J. A. Derivative Studies in Configuration-Interaction Theory. *J. Chem. Phys.* **1980**, *72*, 4654 – 4655.

[13] Raghavachari, K.; Pople, J. A. Calculation of One-Electron Properties Using Limited Configuration-Interaction Techniques. *Int. J. Quantum Chem.* **1981**, *20*, 1067 – 1071.

[14] Bartlett, R.J.; Musial, M. Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291 – 352.

[15] Bartlett, R. J. Many-Body Perturbation Theory and Coupled Cluster Theory for Electron Correlation in Molecules. *Ann. Rev. Phys. Chem.* **1981**, *32*, 359 – 401.

[16] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618 – 622.

[17] Freed, K. F. Many-Body Theories of the Electronic Structure of Atoms and Molecules. *Ann. Rev. Phys. Chem.* **1971**, *22*, 313 – 346.

[18] Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. Electron Correlation Theories and Their Application to the Study of Simple Reaction Potential Surfaces. *Int. J. Quantum Chem.* **1978**, *14*, 545 – 560.

[19] Purvis III, G. D.; Bartlett, R. J. A Full Coupled-Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910 – 1918.

[20] Raghavachari, K.; Trucks, G. W. Pople, J. A. Head-Gordon, M. A Fifth-order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett.* **1989**, *157*, 479 – 483.

[21] Watts, J. D.; Gauss, J.; Bartlett, R. J. Coupled-Cluster Methods with Noniterative Triple Excitations for Restricted Open-Shell Hartree-Fock and other General Single Determinant Reference Functions. Energies and Analytical Gradients. *J. Chem. Phys.* **1993**, *98*, 8718 – 8733.

[22] Knowles, P. J.; Hampel, C.; Werner, H.-J. Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **1993**, *99*, 5219 – 5227.

[23] Knowles, P. J.; Hampel, C.; Werner, H.-J. Erratum: Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys*. **2000,** *112*, 3106 – 3107.

[24] Deegan, M. J. O.; Knowles, P. J. Perturbative Corrections to Account for Triple Excitations in Closed and Open Shell Coupled Cluster Theories. *Chem. Phys. Lett*. **1994**, *227*, 321 – 327.

[25] Adler, T. B.; Knizia, G.; Werner, H.-J. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Chem. Phys.* **2007,** *127,* 221106-1 – 221106-4.

[26] Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009,** *130,* 054104-1 – 054104-20.

[27] Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-ζ Quality Coupled-Cluster Correlation Energies with Triple- ζ Basis Sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921 – 1930.

[28] Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864 – B870.

[29] Löwdin, P.-O. Twenty-Five Years of Sanibel Symposia: A Brief Historic and Scientific Survey. *Int. J. Quant. Chem.* **1986**, *S19*, 19.

[30] Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133 – A1138.

[31] Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200 – 1211.

[32] Perdew, J. P.; Wang, Y. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Phys. Rev. B* **1992**, *45*, 13244 – 13249.

[33] Perdew, J. P; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. Atoms, Molecules, Solids, and Surfaces: Application of the Generalized Gradient Approximation for Exchange and Correlation. *Phys. Rev. B.* **1992**, *46*, 6671 – 6687.

[34] Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, 38, 3098 – 3100.

[35] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865 – 3868.

[36] Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372 – 1377.

[37] Becke, A. D. Density-Functional Theormochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.

[38] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Physical Review B.* **1988**, *37*, 785 – 789.

[39] Cohen, A. J.; Mori-Sanchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289 – 320.

[40] Curtiss, L. A.; Refern, P. C.; Raghavachari, K. G*n* Theory. *WIREs Comput. Mol. Sci.* **2011**, *1*, 810 – 825.

[41] Bokatzian, S. S.; Stover, M. L.; Plummer, C. E.; Dixon, D. A.; Cassady, C. J. An Experimental and Computational Investigation into the Gas-Phase Acidities of Tyrosine and Phenylalanine: Three Structures for Deprotonated Tyrosine. *J. Phys. Chem. B* **2014**, *118*, 12630 – 12643.

[42] Li, Z.; Matus, M. H.; Velazquez, H. A.; Dixon, D. A.; Cassady, C. J. Gas-Phase Acidities of Aspartic Acid, Glutamic Acid, and their Amino Acid Amides. *Int. J. Mass Spectrom.* **2007**, *265*, 213 – 223.

[43] Stover, M. L.; Jackson, V.; Matus, M.; Adams, M.; Cassady, C. J.; Dixon, D. A. Fundamental Thermochemical Properties of Amino Acids: Gas-Phase and Aqueous Acidities and Gas-Phase Heats of Formation. *J. Phys. Chem. B* **2012**, *116*, 2905 – 2916.

[44] Stover, M. L.; Plummer, C. E.; Miller, S. R.; Cassady, C. J.; Dixon, D. A. Gas-Phase Acidities of Phosphorylated Amino Acids. *J. Phys. Chem. B* **2015**, *119*, 14604 – 14621.

[45] Plummer, C. E.; Stover, M. L.; Bokatzian, S. S.; Davis, J. T. M.; Dixon, D. A.; Cassady, C. J. An Experimental and Computational Study of the Gas-Phase Acidities of the Common Amino Acid Amides. *J. Phys. Chem. B* **2015**, *119*, 9661 – 9669.

[46] Bokatzian-Johnson, S. S.; Stover, M. L.; Dixon, D. A.; Cassady, C. J. Gas-Phase Deprotonation of the Peptide Backbone for Tripeptides and their Methyl Esters with Hydrogen and Methyl Side Chains. *J. Phys. Chem. B* **2012**, *116*, 14844 – 14858.

[47] Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V.; Pople, J. A. Gaussian-3 Theory using Reduced Møller-Plesset Order. *J. Chem. Phys.* **1999**, *110*, 4703 – 4709.

[48] Pople, J. A.; Head-Gordon, N.; Raghavachari, K. Quadratic Configuration Interaction. A general Technique for Determining Electron Correlation Energies. *J. Chem. Phys.* **1987**, *87*, 5968 – 5976.

[49] The G3MP2Large basis set, G3(MP2) total energies of the molecules, and a complete tabulation of deviations with experiment can be downloaded from the website http://chemistry.anl.gov/compmat/g3theory.htm.

[50] Grev, R. S.; Janssen, C. L.; Schaefer, H. F. Concerning Zeropoint Vibrational Energy Corrections to Electronic Energies. *J. Chem. Phys.* **1991**, *95*, 5128 – 5132.

[51] Baboul, G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-3 Theory Using Density Functional Geometries and Zero-Point Energies. *J. Chem. Phys.* **1999**, *110*, 7650 – 7657.

[52] Feller, D.; Dixon, D. A. Extended Benchmark Studies of Coupled Cluster Theory through Triple Excitations. *J. Chem. Phys.* **2001**, *115*, 3484 – 3496.

[53] Ruscic, B.; Wagner, A. F.; Harding, L. B.; Asher, R. L.; Feller, D.; Dixon, D. A.; Peterson, K. A.; Song, Y.; Qian, X.; Ng, C.; Liu, J.; Chen, W.; Schwenke, D. W. On the Enthalpy of Formation of Hydroxyl Radical and Gas-Phase Bond Dissociation Energies of Water and Hydroxyl. *J. Phys. Chem. A* **2002**, *106*, 2727 – 2747.

[54] Dixon, D. A.; Feller, D.; Peterson, K. A. Heats of Formation and Ionization Energies of NH_x, x=0-3. *J. Chem. Phys.* **2001**, *115*, 2576 – 2581.

[55] Pollack, L.; Windus, T. L.; de Jong, W. A.; Dixon, D. A. Thermodynamic Properties of the C_5, C_6, and C_8 *n*-Alkanes from Ab Initio Electronic Structure Theory. *J. Phys. Chem. A* **2005**, *109*, 6934 – 6938.

[56] Feller, D.; Peterson, K. A.; Dixon, D. A. A Survey of Factors Contributing to Accurate Theoretical Predictions of Atomization Energies and Molecular Structures. *J. Chem. Phys.* **2008**, *129*, 204015 – 204046.

[57] Dixon, D. A.; Feller, D.; Peterson, K. A. A Practical Guide to Reliable First Principles Computational Thermochemistry Predictions Across the Periodic Table. In *Annual Reports in Computational Chemistry*, Ralph, A. W., Ed.; Elsevier: 2012; Vol. 8, pp 1 – 28.

[58] Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy: Current Strategies and Future Challenges. *Theor. Chem. Acc.* **2012**, *131*, 1 – 20.

[59] Feller, D.; Peterson, K. A.; Dixon, D. A. Further Benchmarks of a Composite, Convergent, Statistically Calibrated Coupled Cluster Based Approach for Thermochemical and Spectroscopic Studies. *Mol. Phys.* **2012**, *110*, 2381 – 2399.

[60] Feller, D.; Peterson, K. A.; Dixon, D. A. The Impact of Larger Basis Sets and Explicitly Correlated Coupled Cluster Theory on the Feller-Peterson-Dixon Composite Method. In *Annual Reports in Computational Chemistry*, Vol. 12, Ed. D. A. Dixon, Elsevier, Amsterdam, 2016, pp 47 – 78.

[61] Douglas, M.; Kroll, N. M. Quantum Electrodynamical Corrections to the Fine Structure of Helium. *Ann. Phys.* **1974**, *82*, 89 – 155.

[62] Hess, B. A. Applicability of the No-Pair Equation with Free-Particle Projection Operators to Atomic and Molecular Structure Calculations. *Phys. Rev. A.* **1985**, *32*, 756 – 763.

[63] Hess, B. A. Relativistic Electronic-Structure Calculations Employing a Two-Component No-Pair Formalism with External-Field Projection Operators. *Phys. Rev. A.* **1986**, *33*, 3742 – 3748.

[64] De Jong, W. A.; Harrison, R. J.; Dixon, D. A. Parallel Douglas-Kroll Energy and Gradients in NWChem: Estimating Scalar Relativistic Effects Using Douglas-Kroll Contracted Basis Sets. *J. Chem. Phys*. **2001**, *114*, 48 – 53.

[65] Moore, C. E. *Atomic Energy Levels As Derived from the Analysis of Optical Spectra*; U.S. National Bureau of Standards Circular 467, COM-72-50282; U.S. Department of Commerce, National Technical Information Service: Washington, DC, 1949; Vol. *1*, H to V.

[66] Chase, M. W., Jr. *NIST-JANAF Thermochemical Tables*, 4th ed. *J. Phys. Chem. Ref. Data.*, *Mono. 9*; American Institute of Physics: Woodbury, NY, 1998; Suppl 1.

[67] Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies of Formation. *J. Chem. Phys.* **1997**, *106*, 1063 – 1079.

[68] McQuarrie, D. A. *Statistical Mechanics*. University Science Books: Sausalito, CA, **2000**.

[69] Omenn, G. S. The HUPO Human Proteome Project (HPP), a Global Health Research Collaboration. *Central Asian Journal of Global Health* **2012**, 1.

[70] Chial, H. DNA Sequencing Technologies Key to the Human Genome Project. *Nature Education* **2008**, *1*, 219.

[71] Lander E. S., Linton L. M., Birren B., et al. Initial Sequence Analysis of the Human Genome. *Nature* **2001**, *409*, 860 – 921.

[72] Venter J. C., Adams M. D., Myers E. W., et al. The Sequence of the Human Genome. *Science* **2001**, *291*, 1304 – 1351.

[73] Nilsson, T.; Mann, M.; Aebersold, R.; Yates, J. R.; Bairoch, A.; Bergeron, J. J. M. Mass Spectrometry in High-Throughput Proteomics: Ready for the Big Time. *Nat. Methods* **2010**, *7*, 681 – 685.

[74] Zeebe, R. E. Stable Boron Isotope Fractionation Between Dissolved $B(OH)_3$ and $B(OH)_4^-$. *Geochim. Cosmochim. Acta.* **2005**, *69*, 2753 – 2766.

[75] Drever, J. *The Geochemistry of Natural Waters.* Prentice Hall: New Jersey, **2002**, pp. 311 – 322.

[76] Conway, B. E. *Ionic Hydration in Chemistry and Biophysics*; Elsevier: New York, 1981.

[77] Marcus, Y. *Ion Solvation*; Wiley: New York, 1985.

[78] Weber, J. M.; Kelley, J. A.; Nielsen, S. B.; Ayotte, P.; Johnson, M. A. Isolating the Spectroscopic Signature of a Hydration Shell with the Use of Clusters: Superoxide Tetrahydrate. *Science* **2000**, *287*, 2461 – 2463.

[79] Pasquarello, A.; Petri, I.; Salmon, P. S.; Parisel, O.; Car, R.; Tóth, É.; Powell, D. H.; Fischer, H. E.; Helm, L.; Merbach, A. First Solvation Shell of the Cu(II) Aqua Ion: Evidence for Fivefold Coordination. *Science* **2001**, *291*, 856 – 859.

[80] Kropman, M. F.; Bakker, H. J. Dynamics of Water Molecules in Aqueous Solvation Shells. *Science* **2001**, *291*, 2118 – 2120.

[81] Robertson, W. H.; Diken, E. G.; Price, E. A.; Shin, J.-W.; Johnson, M. A. Spectroscopic Determination of the OH^- Solvation Shell in the $OH^-\cdot(H_2O)_n$ Clusters. *Science* **2003**, *299*, 1367 – 1372.

[82] Connelly, S. J.; Wiedner, E. S.; Appel, A. M. Predicting the Reactivity of Hydride Donors in Water: Thermodynamic Constants for Hydrogen. *Dalton Trans*. **2015**, *44*, 5933 – 5938

[83] Wiedner, E. S.; Chambers, M. B.; Pitman, C. L.; Bullock, R. M.; Miller, A. J. M.; Appel, A. M. Thermodynamic Hydricity of Transition Metal Hydrides. *Chem. Rev.* **2016**, *116*, 15, 8655 – 8692.

[84] Wiedner, E. S.; Linehan, J. C. Making a Splash in Homogeneous CO_2 Hydrogenation: Elucidating the Impact of Solvent on Catalytic Mechanisms. *Chem. Eur. J.* **2018**, *24*, 16964 – 16971.

[85] Tsay, C.; Livesay, B. N.; Ruelas, S.; Yang, J. Y. Solvation Effects on Transition Metal Hydricity. *J. Am. Chem. Soc.* **2015**, *137*, 14114 – 14121.

[86] Brereton, K. R.; Jadrich, C. N.; Stratakes, B. M.; Miller, A. J. M. Thermodynamic Hydricity Across Solvents: Subtle Electronic Effects and Striking Ligation Effects in Iridium Hydrides. *Organometallics* **2019**, *38*, 3104 – 3110.

[87] Pitman, C. L.; Brereton, K. R.; Miller, A. J. M. Aqueous Hydricity of Late Metal Catalysts as a Continuum Tuned by Ligands and the Medium. *J. Am. Chem. Soc*. **2016**, *138*, 2252 – 2260.

[88] Burgess, S. A.; Appel, A. M.; Linehan, J. C.; Wiedner, E. S. Changing the Mechanism for CO_2 Hydrogenation Using Solvent-Dependent Thermodynamics. *Angew. Chem. Int. Ed*. **2017**, *56*, 15002 – 15005.