UNDERSTANDING FLUORINE CHEMICAL SHIFTS AND THE MOLECULAR PROPERTIES OF ORGANIC MOLECULES

A Dissertation by
Chandana Kasireddy
Master of Science, Osmania University, 2010
Bachelor of Science, Osmania University, 2008

Submitted to the Department of Chemistry
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

May 2018

UNDERSTANDING FLUORINE CHEMICAL SHIFTS AND THE MOLECULAR PROPERTIES OF ORGANIC MOLECULES

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy with a major in Chemistry.

_____

Katie R. Mitchell-Koch, Committee Chair

_____

James G. Bann, Committee Member

_____

Douglas S. English, Committee Member

_____

Dennis H. Burns, Committee Member

_____

Anil Mahapatro, Committee Member

Accepted for the College of Liberal Arts and Sciences

_____

Ron Matson, Dean

Accepted for the Graduate School

_____

Dennis Livesay, Dean

DEDICATION

To my mother, who might have been the most happiest person seeing my progress.
You're the inspiration for each step of my success.

I also like to dedicate my work to my beloved family, friends and all my teachers.

ACKNOWLEDGMENTS

First and foremost, I would like to express my heartfelt gratitude to my advisor, Prof. Katie R. Mitchell-Koch for giving me the opportunity to work in her group and guiding me throughout the process. Thank you for giving me all the support, patient guidance and continuous encouragement. I am so grateful to you for showing confidence in me throughout the process, giving me the freedom and for always backing me up. It's been a privilege for me to work with you and without your guidance this dissertation would not have been possible. Thank you for all insightful discussions, ideas, feedback and for being the wonderful mentor, boss, teacher and a friend. Thank you for your support, kindness and for being a great role model in many ways.

My deepest appreciation goes to Prof. James G. Bann for his continuous support and constant motivation. I would like to pay special appreciation to all my dissertation committee members Prof. Douglas S. English, Prof. Dennis H. Burns and Prof. Anil Mahapatro for serving in my committee and for their continuous assistance and support.

My sincere thanks goes to department chair, Prof. David M. Eichhorn for supporting me and for all the advice. I would like to thank all the professors, faculty, technicians, and staff members who has provided me extensive personal and professional guidance throughout my stay here.

I am so grateful to all my present and past fellow lab mates for helping me all the time. My special thanks goes to Jayangika for her friendship and for being with me all these years in all situations.

Furthermore, I'm thankful to two important people, Mr. & Mrs. Jagadeeswar Reddy. I can never forget the goodness of your heart. Thank you for supporting me in reaching my dreams.

Lastly, I would like to thank my family for all their immense love and encouragement. For My dad Mr. Ram Reddy who raised me with a love and supported me in all my pursuits. Most importantly, I wish to thank my beloved and supportive husband Maruthi. I truly thank you for having faith in me and sticking by my side in all situations. You are the best I could ever ask for. Last but not the least my little sister Madhuri for being there for me all time .You people are amazing and I'm sure I might have never come this far without your support.

Finally, I want to thank each and every person in my life for their contribution in completing my dissertation.

ABSTRACT

Fluorinated molecules have emerged as important probes in drug design, and as strategies for enhancing protein stability and altering physicochemical properties of biological systems. Fluorine NMR plays a crucial role in studying protein structure, dynamics and protein-ligand interactions. Histidine is an important amino acid involved in many enzymatic reactions and protein functions. 2-fluorohistidine and 4-fluorohistidine are fluorinated analogues of histidine, offering valuable tools in biophysical studies with their altered *pKa* values compared to the canonical amino acid.

In this dissertation, the molecular properties of 2-fluorohistidine, 4-fluorohistidine, and analogues are studied with different electronic structure calculations. Among the two tautomeric states of histidine at neutral pH, it was found that 2-fluorohistidine prefers the τ-tautomeric state, whereas 4-fluorohistidine exclusively stays in the π-tautomeric state. While the ^{1}H signal shifts downfield in both isomers upon ring protonation, the ^{19}F signal of 2-fluorohistidine also shifts downfield but ^{19}F in 4-fluorohistidine shifts upfield. The reason for this unusual behavior is explained here. The effect of different environments (solvation effects, intramolecular hydrogen bonding effect, backbone effects) on tautomeric stabilities and fluorine chemical shifts are discussed. This work explains the role of histidine's C4-H bond as a potential hydrogen bond donor. The results indicate that the C4-H bonds in histidine and 2-fluorohistidine in the τ-tautomeric states act as strong hydrogen bond donors.

TABLE OF CONTENTS

Chapter **Page**

TABLE OF CONTENTS (continued)

Chapter **Page**

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES (continued)

Figure **Page**

LIST OF FIGURES (continued)

Figure **Page**

LIST OF TABLES

Table **Page**

LIST OF TABLES (continued)

Table **Page**

LIST OF ABBREVIATIONS

Abbreviation	Systematic name
AMP	Antimicrobial peptides
CPCM	Conductor polar continuum model
DFT	Density functional theory
DHFR	Dihydro foliate reductase
DNA	Deoxyribonucleic acid
ESP	Electrostatic potentia
FDA	Food and drug administration
GCN4	General control protein
GIAO	Gauge independent atomic orbital
HBF	Hydrogen bond formation
HF	Hartree-fock
HIV	Human immunodeficiency virus
MD	Molecular dynamics
MO	Molecular orbital
MP2	Second order møller-plesset perturbation theory
MRI	Molecular resonance imaging
NBO	Natural bonding orbital
NCS	Natural chemical shielding
NMR	Nuclear magnetic resonance

PET	Positron emission tomography
RNA	Ribonucleic acid
SCRF	Self consistent reaction field
XRD	X-Ray Diffraction
ZPE	Zero point energy
2-FHis	2-fluorohistidine
3-FY	3-fluorotyrosine
4-FHis	4-fluorohistidine
5-FU	5-fluorouracil
^{18}F FAHA	^{18}F-fluoroacetamido-hexanoic anilide
^{18}F-FDG	2-deoxy-2-^{18}F-fluoro-beta-D-glucose
^{18}F-FHBG	9-[4-^{18}F-fluoro-3-hydroxymethyl-butyl)guanine
^{18}F-FLT	3'-^{18}F-fluoro-3'-deoxy-thyamidine

CHAPTER 1

Introduction

1.1. Fluorine nucleus:

The fluorine nucleus is poised to be a powerful probe in nuclear magnetic resonance imaging and spectroscopy. Fluorine (^{19}F) is a half spin nucleas that is 100% naturally abundant. It has gyromagnetic ratio equal to 83% that of a proton's, giving rise to high sensitivity and strong signals. It is also considered to be isosteric to hydrogen, having a van der Waal radius of 14% larger than the proton makes ^{19}F compatible for incorporation of fluorinated tags into biomolecules.

^{19}F NMR shifts are influenced by many factors including lone pair electrons, electrostatic interactions, van der Waal forces, hydrogen-bonding interactions etc.[1] The wide range of chemical shifts (~800 ppm, for proton it is ~15 ppm) is one more advantageous property of fluorine NMR. Due to the lack of fluorine in naturally occurring biological systems, there will not be any overlap from background signals competing with the fluorolabelled residues, which makes spectra easy to analyze the signals. Also, compared to other nuclei such as carbon and nitrogen, the fluorine atom is highly sensitive: ~20, ~250 times higher, respectively, and almost as sensitive as hydrogen. All these features make fluorine an important probe for studying protein conformational changes, enzyme kinetics, protein folding, hydrophilic or hydrophobic interactions, and weak binding interactions.

1.2. Carbon-fluorine bond:

Fluorine is the ideal element to replace hydrogen due to being relatively it is isosteric to the proton. The C-F bond is the strongest bond, with a bond dissociation energy of ~105 kcal/mol, possessing high thermal stability. It is a highly polarized bond with a bond length of 1.35 Å. Due to the fact that the C-F bond is more polar than the C-H bond, fluorine replacement may lead to geometry distortions in hydrocarbons. Due to charge transfer effect from carbon to fluorine, as we go from methane to fluoromethane to difluoromethane to trifluoromethane, the C-F bond shortens from 1.39 Å to 1.32 Å with concomitant increase in bond dissociation energy. Also, going from methane to difluoromethane the H-C-H bond angle widens due to electron pull towards the fluorine from the sp^3 carbon, making carbon to attain a more like sp^2 character.[2] With its nearly ionic nature, the C-F bond possesses a high dipole moment. As the number of fluorines replacing hydrogen, the dipole moment increases. With its polarization and presence of lone pairs on fluorine atoms, the C-F bond also acts as a hydrogen bond acceptor. However, compared to typical hydrogen bond, these interactions are weak, with hydrogen bond distance of 2.5-3.0 Å.

Besides forming weak hydrogen bonds, the C-F bond acts as a good metal coordinator. Compared to hydrocarbon-metal complexes, the fluorocarbon-metal complexes are strong in nature with high thermal stability.[3] Preferably, it forms strong interactions with hard metal cations such as Na^+, Li^+, K^+.

1.3. Incorporation of fluorine probes into proteins:

Fluorine incorporation into proteins can be approached in many ways. In some cases, multiple fluorine insertions create a fluoro-stabilization effect in protein due to increased hydrophobicity, which in turn enhances the protein stability. Biosynthetic incorporation and site-specific incorporation are two popular methods being used to introduce fluorine probes. In many cases, fluorine insertion causes changes in ionization characteristics, leading to altered pKa values compared to natural amino acids. Table 1.1 shows the *pKa* natural amino acids and their fluorinated analogues.[4-6]

Table 1.1. *pKa* values of some amino acids and their fluorinated analogues. [(a) Lazar and Sheppard, **1968**, J Med Chem; b) Yeh et al, **1975**, J Am Chem Soc; c) Walborsky and Lang **1956**, J Am Chem Soc)]

Molecule	***pKa* Values**
Valine	2.29
4,4,4-Trifluorovaline	1.54[a]
Hexafluorovaline	1.21[a]
Methionine	2.16[a]
Trifluoromethionine	2.05[a]
Histidine	6[b]
2-fluorohistisine	1[b]
4-fluorohistidine	2.5[b]
Leucine	2.34[a]
Trifluoroleucine	2.05[c]
Hexafluoroleucine	1.81[a]

Biosynthetic incorporation is the standard approach to label via protein expression. Biosynthetic incorporation is performed by using the required auxotrophic regions of bacterial strains and by adding the fluorinated amino acids to the expression medium.[7] Even though it is a widely used method, there are some limitations. The main disadvantage of biosynthetic incorporation is reduced yield. Also, through this approach leads to modification of every amino acid of its kind in the whole protein, which further induces perturbations in protein structure and stability.

This difficulty can be overcome through site-specific incorporation, in which only the desired residue can be chemically modified.[8] In this method using tRNA/ aatRNA (aminoacyl) synthetase pairs, the required fluorinated amino acids are incorporated at codon-defined positions. Site-specific fluorolabeling is a cost effective and potentially successful method. Many successful attempts were made through this approach, some of which include incorporation of 4-fluorophenylalanine into β-lactamase[9] and trifluoromethyl phenylalanine (Tfm-Phe) into multiple sites of nitroreductase and histidinol dehydrogenase.[10] Wilkins et al. incorporated fluorotyrosine analogues, including 3-fluorotyrosine, 2-fluorotyrosine, and 2,6 difluorotyrosine into proteins expressed in *E.coli*.[11] Cheperone PapD was site specifically labeled with fluorophenyl alanine residue by Bann and Frieden to study protein folding and domain-domain interactions of papD.[12]

Many studies made clear the importance of fluorinated amino acids in proteins in enhancing protein stability, resistance to proteolysis, regulating protein interactions and constructing stable protein folds.[13-16] This could be from various factors including increased dipole-dipole interactions or enhanced hydrophobicity. Earlier, most studies

used fluorinated aromatic amino acids, but lately many aliphatic compounds have been inserted into proteins. Recently, incorporating aliphatic CF_3 reporters has attained wide applications in studying peptide-membrane interactions. Even though fluorine is isosteric with hydrogen, the CF_3 group's van der Waal radius is double that of a CH_3 group. This change definitely makes a difference in protein conformation. So, instead of replacing CH_3 groups with CF_3, it is ideal way to use CF_3 as a mimic for bulky groups such as ethyl groups.

The increased hydrophobic volume and surface area plays an important role in interacting and packing with other moieties. Marsh and co-workers reported that the enhanced stability of four-helix bundles of fluorinated analogues in the hydrophobic core is due to the increased hydrophobic interactions.[17]

1.4. Applications of ^{19}F NMR spectroscopy:

Fluorine NMR has become an important tool to understand and investigate protein structure dynamics and protein-ligand interactions. Fluorine chemical shifts are highly sensitive towards local environments compared to other nuclei such as proton, carbon and nitrogen. Relative to ^{15}N, ^{13}C 2D NMR, the one-dimensional ^{19}F NMR spectrum can be obtained with lower concentrations of compound and no need for expensive isotropic labeling. As much as an 8 ppm change in chemical shifts has been observed between native and denatured proteins.[1] Many factors, including short-range and electrostatic interactions affect the fluorine shieldings. Fluorinated amino acids are commonly used fluoroprobes to investigate protein aggregation, enzymatic pathways,

protein folding and protein membrane interactions. Fig. 1.1 shows examples of some fluorinated amino acids.

Fluorine NMR is widely used to monitor structural perturbations due to conformational changes in many protein-ligand interactions. Falke and Luck used fluorine NMR as a primary tool to scan for conformational changes in receptor within the *E.coli* chemosensory pathway upon sugar binding. By inserting different fluorinated amino acids, the experiments showed that sugar binding causes the global structural changes in the receptor.[18]

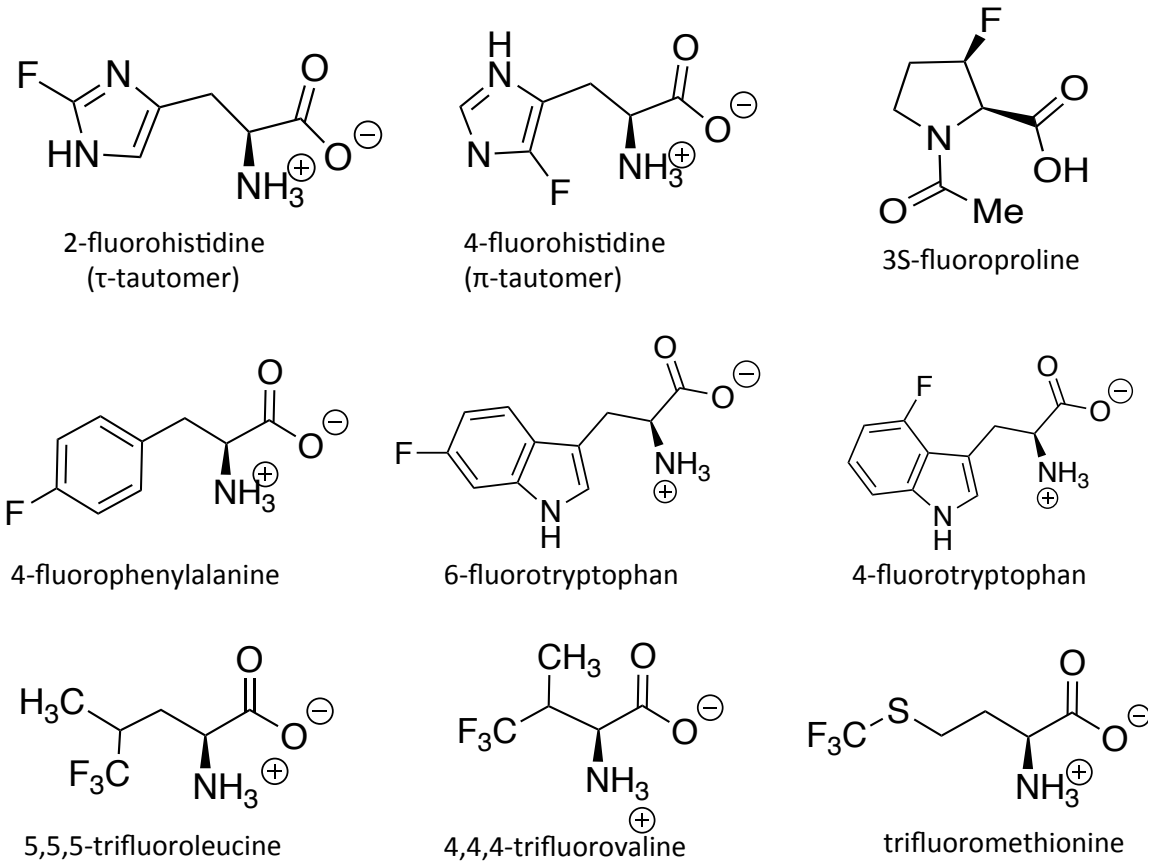


Figure 1.1. Examples of fluorinated amino acids.

Aliphatic amino acids are mainly involved in stabilizing folded proteins through hydrophobic interactions. Many aliphatic amino acids are inserted with CF_3 groups in place of CH_3. Fluorinated analogues of valine, isoleucine and leucine are frequently used fluoro probes in studying the effects of fluorination on protein structure and stability. Introduction of $5^3,5'^3$-F_6Leu in place of Leu and Ile residues in pexiganan showed enhanced stability towards trypsin and chymotrypsin digestion in the presence of a membrane environment.[19] Introducing 5^3-F_3Ile into transcription factor GCN4-p1's dimeric C-terminal subdomain increased the stability of the dimer. Amino acids containing trifluoromethyl groups have turned out to be good probes for increasing thermal stability as well as identifying conformational changes. In different experiments, labeling of basic leucine zippers with trifluorovaline (4,4,4-trifluorovaline) and trifluoroisoleucine (5,5,5-isoleucine) led to increased thermal stability of the proteins to different extents.[20] In other work with fluorinated aliphatic amino acids, studies showed that introducing 5-fluoroleucine in DHFR at thirteen sites showed wide range of fluorine chemical shift dispersion up to ~15 ppm.

1.5. Fluorine in medicinal chemistry:

It is an interesting fact that more than 25 % of pharmaceutical drugs contain at least one fluorine atom in them. Inserting a single fluorine might not lead to steric perturbations due to its (nearly) isosteric nature with hydrogen. However, due to the high electronegativity of fluorine, the lipophilicity and electrostatic interactions of pharmaceuticals can be greatly affected by fluorination. Variation in the basicity and acidity of fluorinated analogues compared to parent compounds leads to changes in binding affinity, bioavailability and pharmacokinetic properties. A number of fluorinated

compounds are approved by FDA as anticancer, antiviral, antifungal, anti-inflammatory and antidiabetic drugs. Below some of the fluorinated drugs and their importance are discussed.

Fludrocortisone and 5-fluorouracil (5-FU) were the first fluorine containing drugs synthesized in the 1950's.[21, 22] Fludrocortisone is an oral drug used in treatment of diseases like adrenal insufficiency, postural hypotension and adrenogenital syndrome by possessing high glucocorticoid activity. 5-FU is used in treating a variety of malignant tumors, including breast cancer, gastrointestinal cancer and neck cancer. It acts as an antimetabolite and anticancer drug by suppressing the thymidylate synthetase enzyme's activity, thus interfering with DNA and RNA synthesis. Fluoxetine, commonly known as Prozac, is an antidepressant drug used mainly to treat bulimia and obsessive-compulsive disorder. Studies showed that the presence of the trifluoro-methyl group on the aryl ring enhanced the potency up to 6 fold towards inhibiting 5-hydroxytryptamine uptake compared to non-fluorinated analogue. [23] Atorvastatin, also known as Lipitor, is one of the top selling drugs used in preventing heart attacks by treating high cholesterol level.[24]

Another widely used mono-fluorinated antibiotic used against both gram-negative and gram-positive bacteria is ciprofloxacin. It is used to treat respiratory tract infections and gastric infections. It acts by inhibiting the activity of bacterial enzymes tomoisomerase and DNA gyrase, causing cell-death. Fluorinated cationic antimicrobial peptides (AMPs) have been found to serve as better antimicrobial agents compared to non-fluorinated AMPs. Studies by Diaz et al. showed that the higher biological activity of fluorinated AMPs is due to their stronger interactions with the membrane interface.[25]

Fluorithromycin is a fluorinated analogue of the most widely used antibiotic, erythromycin. It was found that, compared to erythromycin, this molecule has higher bioavailability, better stability in acidic conditions and a prolonged serum half-life.[26 27] Indinavir is a protease inhibitor used in antiretroviral therapy to treat HIV.[28] Later experiments showed that the fluorinated analogue syn,anti-4 C17-*epi*-Indinavir posses 8-fold more potency compared to its parent compound.[29]

Over the past decade, many fluorine-containing drugs have been developed and released into the market. Some of these include cinacalcet, efavirenz, sitagliptin, amitiza, tamusulosin which are used to treat various diseases. The current statistics show that in the past ten years, there has been an almost 20% increase in the availability of fluorinated drugs on the market. This clearly indicates the continuing impact and importance of fluorinated compounds in the healthcare industry.

Positron emission tomography (PET) is a widely used method for the early detection of tumor cells and in molecular imaging of biochemical processes. ^{18}F is one of the commonly used fluorine isotopes in the radiolabeling of biomolecules.[30, 31] It is widely used in PET due to its short half life and positron emission. 2-deoxy-2-^{18}F-fluoro-beta-D glucose (^{18}F-FDG) is an analogue of glucose that is widely used for PET tracers in detecting tumor cells.[32] Several other compounds are used in PET for radiolabelling in electrophilic or nucleophilic reactions, some of which include 3'-^{18}F-fluoro-3'-deoxy-thymidine (^{18}F-FLT), 9-[4-^{18}F-fluoro-3-hydroxymethyl-butyl)guanine (^{18}F-FHBG), ^{18}F-fluoroacetamido-hexanoic anilide (^{18}F FAHA) and ^{18}F-fluorodopa.

^{19}F NMR is also largely used in molecular resonance imaging (MRI) and in gene therapy. ^{19}F MRI has attained more attention in capturing the dynamics of labeled cells. (fluorine MRS and MRI in biomedicine). Along with studying globular proteins, ^{19}F NMR is quite useful for studying disordered structures also. In separate experiments, 3-fluorotyrosine (3FY) and trifluoromethyl L-phenylalanine (tfmF) residues were incorporated into disordered alpha-synuclein protein (140 residues) at 39, 125, 133 and 136 positions using site specific labeling. From ^{19}F NMR, the NMR chemical shifts and peak broadening were evaluated to assess structural and dynamics details.[33] The results showed that the 221-residue side chain is more flexible/mobile compared to the one at the 39th position.

1.6. Histidine:

Histidine is a most versatile member of the naturally occurring amino acids. It is found to be present in more than 50% of enzyme active sites. The imidazole side chain of histidine is aromatic in nature and exists in different forms, depending on which nitrogen is protonated. At neutral pH, it exists in two forms; namely, τ-tautomer (N3-H) and π-tautomer (N1-H). Under acidic conditions, it exists in a protonated state in which both the nitrogens are protonated. Based on protonation states, histidine has the potential to act as Lewis acid, Lewis base, hydrogen bond acceptor and hydrogen bond donor. At neutral pH the imidazole ring possesses a *pKa* value of ~6, which is close to biological pH. These properties give histidine unique structure that participates in different biological functions.

1.6.1 Histidine tautomerization importance and characterization:

Based on the protonation state, histidine can either serve as hydrogen bond donor or acceptor. The previous studies showed that at neutral pH in aqueous solution, the τ-tautomer is more stable than the π-tautomer (existing in an 80:20 τ : π ratio). Several studies demonstrated the role of histidine tautomerization in proton transfer mechanisms.[34] Hass et al. reported that the histidine 61 residue in the copper-containing plastocyanine protein exists in different protonation forms. They reported the histidine's conformational exchange, tautomerization, and rotomerization through NMR measurements.[34] Shimahara et al. investigated the catalytic activity of human carbonic anhydrase II (hCAII), and it was revealed that the rate of proton transfer is highly efficient through a histidine tautomerization.[35] The significance of the histidine residue in ribonuclease A in enzymatic cleavage of RNA has been reported via many studies.[36, 37]

NMR studies serve as a good biophysical method for characterizing histidine tautomers. In subtilisin BPN', among the available six histidine residues, four of them are present in N3-H tautomer and the other two residues are dominant in the N1-H protonation state. The ^{13}C chemical shift indicates that if C4 carbon shifts are >122 ppm, the π-tautomer is present; if the shifts are <122 ppm, mostly the τ-tautomer is present.[38] NMR studies are used to determine the tautomeric states of histidine residues in phosphorylated and unphosphorylated forms of the protein III^{Glc}. The results showed that the His 75 residue remains dominantly in the τ-tautomeric state in both phosphorylated and unphosphorylated III^{Glc}. The His-90 residue primarily appears in the π-tauomeric state in unphosphorylated protein. However, upon phosphorylation, the N3 nitrogen also gets protonated, making histidine charged (cationic) residue.[39]

1.6.2. Histidine in metal coordination:

The significance of histidine in metal coordination in biological systems has been known for a long time. At neutral pH, the unprotonated nitrogen in both tautomeric states acts as a ligand in metal coordination. From many studies it is known that histidine in metalloproteins plays crucial roles in transport processes,[40] catalysis, electron transfer reactions,[41] protein aggregation[42] and oxidation reactions.[43] Studies have shown that, the histidine (metal free) tautomers are indicated by C4-C5 stretching bands via Raman spectra.[44, 45] The π tautomer stretching band is at higher frequencies (~1580 cm-1) compared to the more stable τ-tautomer (~1560 cm^{-1}). Recently, Miura *et al*. investigated metal binding properties of the histidine side chain using Raman spectra. Spectral analysis showed that upon binding to metal, the C-C stretch increased to higher wavenumber. However, upon metal coordination to N1-nitogen, in case of τ-tautomer, and to N3-nitrogen, in the case of π-tautomer, histidine exhibits increased C4-C5 stretching frequencies of ~1570 cm^{-1} and ~1600 cm^{-1} respectively.

Human carbonic anhydrase (CAII) is a zinc containing enzyme that catalyzes the reversible reaction of carbon dioxide hydration.[46] The zinc in the active site of CAII is coordinated by three His ligands, His94, His96 and His119, and one hydroxide ions forming a tetrahedral geometry (Fig. 1.2). Experiments substituting the histidine residues in CAII with aspargine and glutamine showed strongly diminished metal-protein binding affinity, which in turn alters the enzyme's activity.[47]

A. lesbiacum (Alyssum) is a terrestrial plant that is a hyper-accumulator and high tolerant of nickel. The studies showed that histidine plays a major role in nickel uptake by

metal chelation in plants, and this nitrogen containing ligand is involved in metal transport.[48] Histidine-87 is one of the important residues present in the binding pocket of the protein streptavidin. This surface exposed ligand was found to be crucial in protein's crystallization on a monolayer of iminodiacetate-Cu^{+2} lipid at the water-air interface.[49]

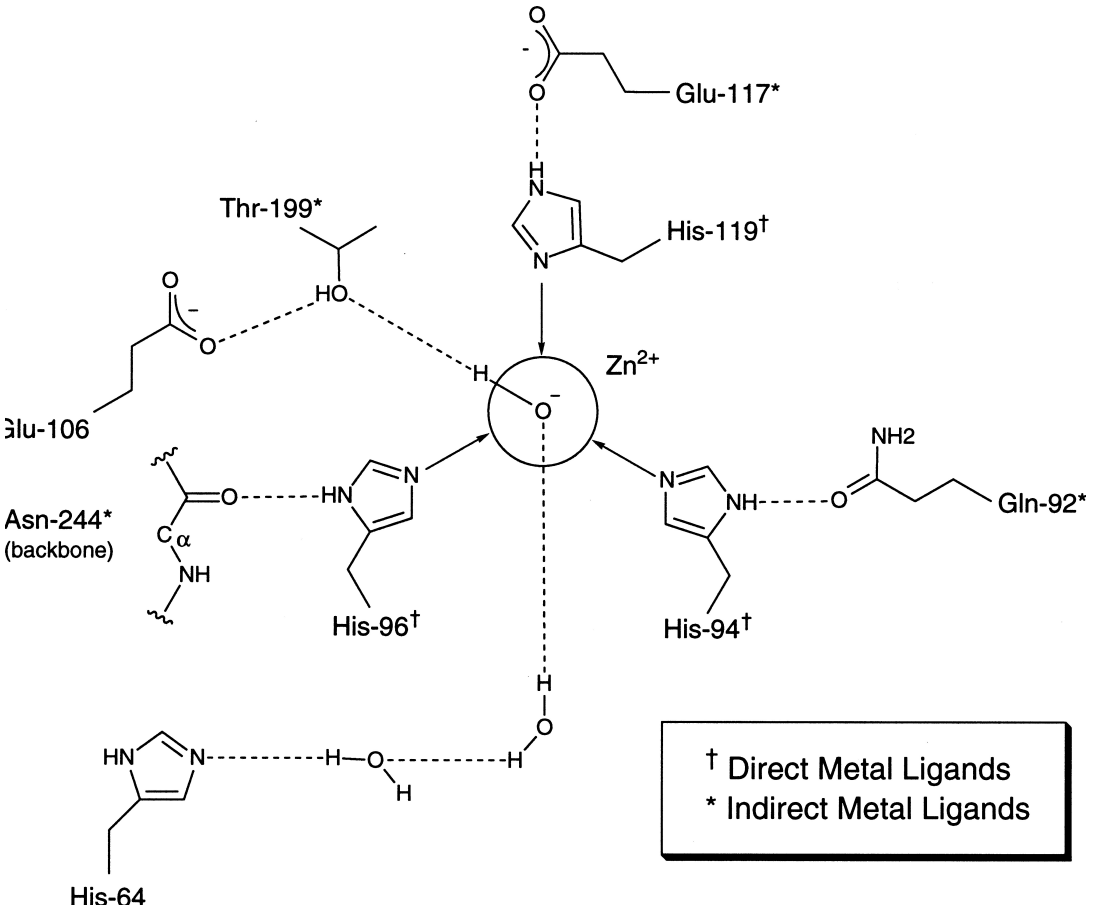


Figure 1.2. Zinc coordination with histidine ligands in the active site of human carbonic anhydrase II.[47]

1.7. 2-fluorohistidine and 4-fluorohistidine:

2-fluorohistidine and 4-fluorohistidine are fluorinated analogues of the canonical amino acid histidine. The two regioisomers have dramatically lower pH values respective to histidine: ~1 for 2-fluorohistidine and ~2.5 for 4-fluorohistidine. Most of the time, 4-fluoroimidazoles are stable under many chemical conditions, except in the presence of electron with drawing groups. In contrast, 2-fluoroimidazoles tend to undergo nucleophilic displacement due to fluorination at 2nd position. The photochemical Shiemann method is a widely known method used over the past ~35 years for 2-fluorohistidine and 4-fluorohistidine synthesis.[50] Fluorohistidine incorporation into proteins has advanced understanding of proteins structure and function. Several methods have been reported for incorporating flurohistidine isomers into proteins, both biochemically and synthetically. Wells and co-workers substituted catalytic residues, histidine -12 and histidine-119 in RNase A with unnatural amino acid 4-fluorohistidine.[51] Bann and co-workers incorporated fluorohistidine probes into papD chaperone using bacterial auxotrophs via biosynthetic incorporation.[50]

The second chapter of this dissertation summarizes the molecular and biophysical properties of the histidine analogues, 2-fluorohistidine and 4-fluorohistidine. In the third chapter, the reasons for chemical fluorine chemical shift behavior between 2-fluorohistidine and 4-fluorohistidine isomers upon ring protonation is explained. In the fourth chapter, the role of histidine's and 2-flurohistidine's C4-H bond in hydrogen bond interactions is studied. In the fifth chapter, various computational methods for calculating absolute and relative chemical shifts for fluorohistidine isomers and other considered fluorinated molecules are evaluated.

REFERENCES

[1] Lau, E. Y., and Gerig, J. T. (2000) Origins of fluorine NMR chemical shifts in fluorine-containing proteins? *J Am Chem Soc 122*, 4408-4417.

[2] O'Hagan, D. (2008) Understanding organofluorine chemistry. An introduction to the C-F bond. *Chem Soc Rev 37*, 308-319.

[3] Kiplinger, J. L., Richmond, T. G., and Osterberg, C. E. (1994) Activation of carbon fluorine bonds by metal-complexes. *Chem Rev 94*, 373-431.

[4] Yeh, H. J. C., Kirk, K. L., Cohen, L. A., and Cohen, J. S. (1975) F-19 and H-1 nuclear magnetic-resonance studies of ring-fluorinated imidazoles and histidines. *J Chem Soc Perk T 2*, 928-934.

[5] Lazar, J., and Sheppard, W. A. (1968) Fluorinated analogs of leucine, methionine, and valine. *J Med Chem 11*, 138-140.

[6] Walborsky, H. M., and Lang, J. H. (1956) Effects of the trifluoromethyl group .4. The Pks of W-trifluoromethyl amino acids. *J Am Chem Soc 78*, 4314-4316.

[7] Ngo, J. T., and Tirrell, D. A. (2011) Noncanonical amino acids in the interrogation of cellular protein synthesis. *Acc Chem Res 44*, 677-685.

[8] Link, A. J., Mock, M. L., and Tirrell, D. A. (2003) Non-canonical amino acids in protein engineering. *Curr Opin Biotech 14*, 603-609.

[9] Noren, C. J., Anthonycahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) A general-method for site-specific incorporation of unnatural amino-acids into proteins. *Science 244*, 182-188.

[10] Jackson, J. C., Hammill, J. T., and Mehl, R. A. (2007) Site-specific incorporation of a F-19-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J Am Chem Soc 129*, 1160-1166.

[11] Wilkins, B. J., Marionni, S., Young, D. D., Liu, J., Wang, Y., Di Salvo, M. L., Deiters, A., and Cropp, T. A. (2010) Site-specific incorporation of fluorotyrosines into proteins in Escherichia coli by photochemical disguise. *Biochemistry-Us 49*, 1557-1559.

[12] Bann, J. G., and Frieden, C. (2004) Folding and domain-domain interactions of the chaperone PapD measured by 19F NMR. *Biochemistry-Us 43*, 13775-13786.

[13] Marsh, E. N. G., and Suzuki, Y. (2014) Using F-19 NMR to probe biological interactions of proteins and peptides. *ACS Chem Biol 9*, 1242-1250.

[14] Biava, H., and Budisa, N. (2014) Evolution of fluorinated enzymes: An emerging trend for biocatalyst stabilization. *Eng Life Sci 14*, 340-351.

[15] Salwiczek, M., Nyakatura, E. K., Gerling, U. I. M., Ye, S. J., and Koksch, B. (2012) Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions. *Chem Soc Revs 41*, 2135-2171.

[16] Suzuki, Y., Buer, B. C., Al-Hashimi, H. M., and Marsh, E. N. G. (2011) Using fluorine nuclear magnetic resonance to probe changes in the structure and dynamics of membrane-active peptides interacting with lipid bilayers. *Biochemistry-US 50*, 5979-5987.

[17] Buer, B. C., Meagher, J. L., Stuckey, J. A., and Marsh, E. N. G. (2012) Structural basis for the enhanced stability of highly fluorinated proteins. *Proc Natl Acad Sci USA 109*, 4810-4815.

[18] Luck, L. A., and Falke, J. J. (1991) 19F NMR studies of the D-galactose chemosensory receptor. 1. Sugar binding yields a global structural change. *Biochemistry-US 30*, 4248-4256.

[19] Gottler, L. M., Lee, H. Y., Shelburne, C. E., Ramamoorthy, A., and Marsh, E. N. G. (2008) Using fluorous amino acids to modulate the biological activity of an antimicrobial peptide. *Chembiochem 9*, 370-373.

[20] Son, S., Tanrikulu, I. C., and Tirrell, D. A. (2006) Stabilization of bzip peptides through incorporation of fluorinated aliphatic residues. *Chembiochem 7*, 1251-1257.

[21] Fried, J., and Sabo, E. F. (1953) Synthesis of 17-alpha-hydroxycorticosterone and its 9-alpha-halo derivatives from 11-epi-17-alpha-hydroxycorticosterone. *J Am Chem Soc 75*, 2273-2274.

[22] Heidelberger, C., Chaudhuri, N. K., Danneberg, P., Mooren, D., Griesbach, L., Duschinsky, R., Schnitzer, R. J., Pleven, E., and Scheiner, J. (1957) Fluorinated pyrimidines, a new class of tumour-inhibitory compounds. *Nature 179*, 663-666.

[23] Wong, D. T., Bymaster, F. P., and Engleman, E. A. (1995) Prozac (fluoxetine, Lilly 110140), the first selective serotonin uptake inhibitor and an antidepressant drug: twenty years since its first publication. *Life Sci 57*, 411-441.

[24] Durazzo, A. E., Machado, F. S., Ikeoka, D. T., De Bernoche, C., Monachini, M. C., Puech-Leao, P., and Caramelli, B. (2004) Reduction in cardiovascular events after vascular surgery with atorvastatin: a randomized trial. *J Vasc Surg 39*, 967-975; discussion 975-966.

[25] Diaz, M. D., Palomino-Schatzlein, M., Corzana, F., Andreu, C., Carbajo, R. J., del Olmo, M., Canales-Mayordomo, A., Pineda-Lucena, A., Asensio, G., and Jimenez-Barbero, J. (2010) Antimicrobial peptides and their superior fluorinated analogues: Structure-activity relationships as revealed by NMR spectroscopy and MD Calculations. *Chembiochem 11*, 2424-2432.

[26] Mabe, S., Eller, J., and Champney, W. S. (2004) Structure-activity relationships for three macrolide antibiotics in Haemophilus influenzae. *Curr Microbiol 49*, 248-254.

[27] Fera, M. T., Giannone, M., Pallio, S., Tortora, A., Blandino, G., and Carbone, M. (2001) Antimicrobial activity and postantibiotic effect of flurithromycin against Helicobacter pylori strains. *Int J Antimicrob Agents 17*, 151-154.

[28] Vacca, J. P., Dorsey, B. D., Schleif, W. A., Levin, R. B., McDaniel, S. L., Darke, P. L., Zugay, J., Quintero, J. C., Blahy, O. M., Roth, E., and et al. (1994) L-735,524: an orally bioavailable human immunodeficiency virus type 1 protease inhibitor. *Proc Natl Acad Sci U S A 91*, 4096-4100.

[29] Myers, A. G., Barbay, J. K., and Zhong, B. Y. (2001) Asymmetric synthesis of chiral organofluorine compounds: Use of nonracemic fluoroiodoacetic acid as a practical electrophile and its application to the synthesis of monofluoro hydroxyethylene dipeptide isosteres within a novel series of HIV protease inhibitors. *J Am Chem Soc 123*, 7207-7219.

[30] Alauddin, M. M., Fissekis, J. D., and Conti, P. S. (2003) Synthesis of [F-18]-labeled adenosine analogues as potential PET imaging agents. *J Labelled Compd Rad 46*, 805-814.

[31] Alauddin, M. M. (2012) Positron emission tomography (PET) imaging with (18)F-based radiotracers. *Am J Nucl Med Mol Imaging 2*, 55-76.

[32] Pal, A., Balatoni, J. A., Mukhopadhyay, U., Ogawa, K., Gonzalez-Lepera, C., Shavrin, A., Volgin, A., Tong, W., Alauddin, M. M., and Gelovani, J. G. (2011) Radiosynthesis and initial in vitro evaluation of [18F]F-PEG6-IPQA--a novel PET radiotracer for imaging EGFR expression-activity in lung carcinomas. *Mol Imaging Biol 13*, 853-861.

[33] Li, C., Wang, G.-F., Wang, Y., Creager-Allen, R., Lutz, E. A., Scronce, H., Slade, K. M., Ruf, R. A., Mehl, R. A., and Pielak, G. J. (2009) Protein 19F NMR in Escherichia coli. *J Am Chem Soc 132*, 321-327.

[34] Hass, M. A., Hansen, D. F., Christensen, H. E., Led, J. J., and Kay, L. E. (2008) Characterization of conformational exchange of a histidine side chain: protonation, rotamerization, and tautomerization of His61 in plastocyanin from Anabaena variabilis. *J Am Chem Soc 130*, 8460-8470.

[35] Shimahara, H., Yoshida, T., Shibata, Y., Shimizu, M., Kyogoku, Y., Sakiyama, F., Nakazawa, T., Tate, S., Ohki, S. Y., Kato, T., Moriyama, H., Kishida, K., Tano, Y., Ohkubo, T., and Kobayashi, Y. (2007) Tautomerism of histidine 64 associated with proton transfer in catalysis of carbonic anhydrase. *J Biol Chem 282*, 9646-9656.

[36] Tanokura, M. (1983) H-1-NMR Study on the tautomerism of the imidazole ring of histidine-residues .1. microscopic pK values and molar ratios of tautomers in histidine-containing peptides. *Biochim Biophys Acta 742*, 576-585.

[37] Walters, D. E., and Allerhand, A. (1980) Tautomeric states of the histidine-residues of bovine pancreatic ribonuclease-A - application of C-13 nuclear magnetic-resonance spectroscopy. *J Biol Chem 255*, 6200-6204.

[38] Sudmeier, J. L., Bradshaw, E. M., Haddad, K. E., Day, R. M., Thalhauser, C. J., Bullock, P. A., and Bachovchin, W. W. (2003) Identification of histidine tautomers in proteins by 2D 1H/13C(delta2) one-bond correlated NMR. *J Am Chem Soc 125*, 8430-8431.

[39] Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from Escherichia coli, using two-dimensional heteronuclear NMR techniques. *Protein Sci 2*, 543-558.

[40] Zhou, L., Li, S. H., Su, Y. C., Yi, X. F., Zheng, A. M., and Deng, F. (2013) Interaction between histidine and Zn(II) metal ions over a wide pH as revealed by solid-state NMR spectroscopy and DFT calculations. *J Phys Chem B 117*, 8954-8965.

[41] Mayo, S. L., Ellis, W. R., Jr., Crutchley, R. J., and Gray, H. B. (1986) Long-range electron transfer in heme proteins. *Science 233*, 948-952.

[42] Liu, S. T., Howlett, G., and Barrow, C. J. (1999) Histidine-13 is a crucial residue in the zinc ion-induced aggregation of the A beta peptide of Alzheimer's disease. *Biochemistry-Us 38*, 9373-9378.

[43] Sundberg, R. J., and Martin, R. B. (1974) Interactions of histidine and other imidazole derivatives with transition-metal ions in chemical and biological-systems. *Chem Rev 74*, 471-517.

[44] Ashikawa, I., and Itoh, K. (1979) Raman-spectra of polypeptides containing L-histidine residues and tautomerism of imidazole side-chain. *Biopolymers 18*, 1859-1876.

[45] Miura, T., Satoh, T., Hori-i, A., and Takeuchi, H. (1998) Raman marker bands of metal coordination sites of histidine side chains in peptides and proteins. *J Raman Spectrosc 29*, 41-47.

[46] Ren, X. L., Tu, C. K., Laipis, P. J., and Silverman, D. N. (1995) Proton-transfer by histidine-67 in site-directed mutants of human carbonic-anhydrase-III. *Biochemistry-US 34*, 8492-8498.

[47] Lesburg, C. A., Huang, C. C., Christianson, D. W., and Fierke, C. A. (1997) Histidine -> carboxamide ligand substitutions in the zinc binding site of carbonic anhydrase II alter metal coordination geometry but retain catalytic activity. *Biochemistry-US 36*, 15780-15791.

[48] Kramer, U., CotterHowells, J. D., Charnock, J. M., Baker, A. J. M., and Smith, J. A. C. (1996) Free histidine as a metal chelator in plants that accumulate nickel. *Nature 379*, 635-638.

[49] Frey, W., Schief, W. R., Pack, D. W., Chen, C. T., Chilkoti, A., Stayton, P., Vogel, V., and Arnold, F. H. (1996) Two-dimensional protein crystallization via metal-ion coordination by naturally occurring surface histidines. *Proc Natl Acad Sci USA 93*, 4937-4941.

[50] Eichler, J. F., Cramer, J. C., Kirk, K. L., and Bann, J. G. (2005) Biosynthetic incorporation of fluorohistidine into proteins in E-coli: A new probe of macromolecular structure. *Chembiochem 6*, 2170-2173.

[51] Jackson, D. Y., Burnier, J., Quan, C., Stanley, M., Tom, J., and Wells, J. A. (1994) A designed peptide ligase for total synthesis of ribonuclease-A with unnatural catalytic residues. *Science 266*, 243-247.

CHAPTER 2

The Biophysical probes of 2-fluorohistidine and 4-fluorohistidine: Spectroscopic signatures and molecular properties

Note: Most of the contents in this chapter are published in Scientific Reports, 2017, 7, 42651 and Chemical Physics Letters, 2016, 666, 58.

2.1. INTRODUCTION

The difference in biophysical and molecular properties of the histidine analogues 2-fluorohistdine and 4-fluorohistidine are discussed in this chapter. In canonical histidine, the τ-tautomer is reported to be more stable compared to π-tautomer at neutral pH.[1] In several studies, NMR spectroscopy has been used to assign protonation states of histidine involved in protein function and structure. Numerous NMR experiments on proteins indicated the hidden states and rapid equilibration between tautomeric states involved in protein conformations. Traditionally, studies of histidine in protein function have used NMR spectroscopy to assign the protonation state.[1-11] The tautomeric states of histidine can be distinguished by ^{15}N NMR isotropic and anisotropic studies[12] and C4 and C5 ($C_{\delta 2}$ and C_{γ}) chemical shifts.[4] NMR experiments indicate rapid equilibration between tautomers,[13] and recent NMR studies of proteins have revealed hidden tautomeric states of histidine in transient protein conformations.[4] Work by Scheraga,[2] Oldfield and co-workers[7] showed the utility of calculated and experimental ^{13}C NMR shifts to evaluate tautomeric forms, and indicated that tautomeric form and chemical shifts of histidine are highly dependent on intramolecular hydrogen bonding. Recent studies showed the

advantage of ^{13}C NMR chemical shifts to evaluate tautomeric states and the dependence of these states on intramolecular hydrogen bonding.[4, 14]

The fluorinated analogues of histidine, which exist as two different regioisomers, have dramatically lower pK_a values than canonical histidine: ~1 (aqueous) for 2-fluorohistidine (2-FHis) and ~2 (aqueous) to >3 (proteins) for 4-fluorohistidine (4-FHis).[15] Our results showed that in the case of 2-fluorohistidine, the τ-tautomer is favored alike to canonical histidine, whereas the π-tautomer appeared to be significantly more stable in 4-fluorohistidine. Results from DFT calculations presented below enable for spectroscopic identification (via ^{19}F, ^{13}C chemical shifts) of the τ- and π-tautomers of 2- and 4-FHis in different environments (*i.e.,* with or without intramolecular hydrogen bonding, in different dielectric environments). The nomenclature used to describe atoms in the imidazole ring throughout this paper is illustrated in Fig. 2.1a, alongside Fig. 2.1b, which shows labels used to identify the atoms of the histidine side chain in other literature reports.

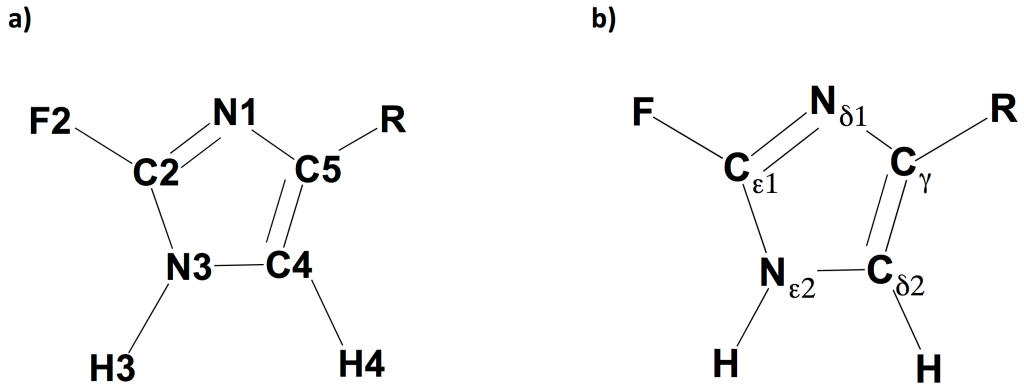


Figure 2.1. The imidazole ring of 2-fluorohistidine (τ-tautomer). The labels used for imidazole ring atoms in this work are shown in 1a, while 1b provides symbols used in other literature discussions of histidine.

2.2. METHODS

Energy minimization, frequencies, and NMR chemical shifts were calculated for Gly-FHis-Gly tripeptides and the following analogues, in order to simulate effects of peptide environment on fluoroimidazole ring properties: (2-fluorohistidine **(2-FHis)**, 4-fluorohistidine **(4-FHis)**, 2-fluoro-(5-methyl)-imidazole **(2F-MeIm)**, 4-fluoro-(5-methyl)-imidazole **(4F-MeIm)**, *N*-acetyl-2FHis-O-methyl ester **(NAc-2FHis-OMe)**, *N*-acetyl-4FHis-O-methyl ester **(NAc-4FHis-OMe),** 2-fluoro-(5-trifluoromethyl)-imidazole **(2F-CF_3Im)**, 4-fluoro-(5-trifluoromethyl)-imidazole **(4F-CF_3Im)**. All quantum chemical calculations were performed using Gaussian 09 package[16] with calculations set up in GaussView 5.[17] Multiple methods including B3LYP, BHandHLYP, HF are used for the calculations. The results with BHandHLYP[18] with 6-311++G(d,p) basis set are discussed in detail here. Initially, all the geometries were optimized to get a stable conformation, and all the positive frequencies verified that geometries are at local energy minima. These geometries were used for calculating absolute shielding constants, energies and electrostatic potential charge distribution (by ChelpG method)[19] at the same level of theory. Free energies were obtained from frequency calculations.

Self Consistent Reaction Field (SCRF) method is used for implicit solvation, with the conductor polar continuum model (CPCM),[20] to study the effect of dielectric environment/solvation on spectra and tautomeric stabilities. To estimate chemical shifts and molecular properties of buried residues (versus solvent-exposed side chains), aqueous (ε=78.3) calculations are compared to those in tetrahydrofuran (THF, ε=7.6) and acetonitrile (ε=35.7) to see effects of less polar environments.

NMR shielding constants were calculated using Gauge Independent Atomic Orbital (GIAO)[21 22] method, and are reported for ^{13}C and ^{19}F nuclei. The NMR chemical shifts (δ) reported for the compounds are calculated by the formula $\delta_{compound} = \sigma_{ref} - \sigma_{compound}$ from the shielding value (σ) of the compound and shielding of the reference compound: C_6F_6 for ^{19}F and tetramethylsilane (TMS) for ^{13}C. Shielding tensors for the reference compound are calculated with BHandHLYP/6-311++G(d,p) method after optimizing the molecule with the same method. Experimental ^{13}C NMR data for 2-fluorohistidine in water was collected on a Varian Mercury VX 300 MHz NMR equipped with a Varian 300 SW/BB probe.

2.3. RESULTS

The tautomeric stabilities of 2-fluorohistidine and 4-fluorohistidine incorporated into the Gly-FHis-Gly tripeptide were calculated using multiple methods. The tautomeric states of Gly-2FHis-Gly and Gly-4FHis-Gly are shown in Fig. 2.2. The energies are calculated with multiple methods including B3LYP, HF, BHandHLYP functionals. These energy differences were found to be similar across a survey of electronic structure methods, as shown in Table 2.1. Results with BHandHLYP density functional, 6-311++G(d,p) basis set, and CPCM implicit solvation are discussed in detail below. Yu *et al.* have recommended BHandHLYP functional for systematic amino acids from several density functional methods.[23]

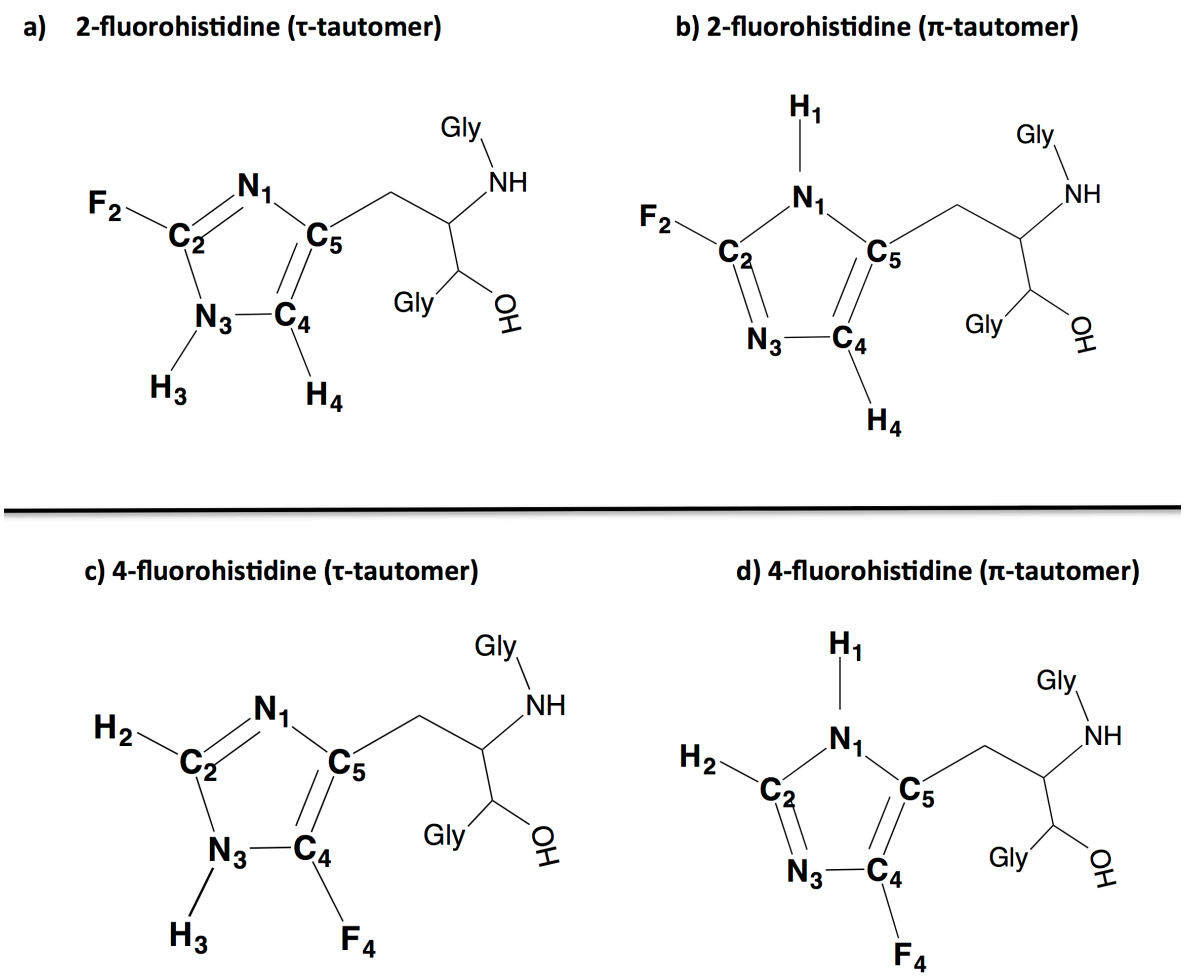


Figure 2.2. The two tautomeric forms of 2-fluorohistidine (a. τ-tautomer, b. π-tautomer) and 4-fluorohistidine (c. τ-tautomer, d. π-tautomer) in Gly-FHis-Gly tripeptide.

The calculations revealed that, for 2-fluorohistidine, the τ-tautomer (N3-H) is more stable than the π-tautomer (N1-H) by 1.70 kcal/mol, resulting in a Boltzmann ratio of 94:6 τ : π tautomer at 310 K. The energy differences between the tautomers of 4-fluorohistidine are substantial, with the π-tautomer being more favored than the τ-tautomer by an energy difference of 4.85 kcal/mol.

The crystal structure of 2-fluorohistidine indicates protonation at the τ-tautomer position.[24] Although XRD data cannot conclusively give proton locations, the N3-C2 bond length compared to the N1-C2 bond length indicates protonation at N3 (τ-tautomer).

In the case of 4-fluorohistidine, the stability of the π-tautomer can be rationalized by the position of electron-withdrawing fluorine in the imidazole ring: the nitrogen nearest to F4 is less basic (N3, unprotonated in the π tautomer), while the N1 position has higher electron density and is more basic.

Table 2.1. Relative stabilities of aqueous tripeptide tautomers, using different electronic structure methods. Energies are reported setting the lowest-energy tautomer to zero, using free energies obtained from frequency calculations.

2-fluorotripeptide free energies in kcal/mol		
Method	**τ-tautomer**	**π-tautomer**
B3LYP/6-31++G(d) (cpcm)	0	0.09
HF/6-311++G(d,p) (cpcm)	0	1.51
HF/6-31++G(d) (cpcm)	0	1.47
B3LYP/6-31++G(d) (smd)	0	0.73
BHandHLYP/6-31++G(d) (smd)	0	0.63
Average	0	0.90
4-fluorotripeptide free energies in kcal/mol		
Method	**τ-tautomer**	**π-tautomer**
B3LYP/6-31++G(d) (cpcm)	4.32	0
HF/6-311++G(d,p) (cpcm)	5.06	0
HF/6-31++G(d) (cpcm)	4.97	0
B3LYP/6-31++G(d) (smd)	4.58	0
BHandHLYP/6-31++G(d) (smd)	5.85	0
Average	5.00	0

In Fig. 2.3 the calculated ^{19}F NMR shifts are plotted with the experimental shifts. It illustrates how the experimental and calculated ^{19}F NMR spectra support the calculated tautomeric stabilities, by plotting experimental titration data for 2-FHis and 4-FHis[25] alongside the calculated ^{19}F chemical shifts of both tautomers and protonated forms of the fluorohistidine isomers. In experiments, 2-fluorohistidine shows a downfield shift of 4.49 ppm upon protonation (from 59.80 ppm at pH ~7 to 64.29 ppm at pH ~1). In the same way, calculations for zwitterionic 2-fluorohistidine show deshielding of 6.0 ppm upon protonation for the τ-tautomer, while the signal shifts downfield by 9.7 ppm when the π-tautomer is protonated. For 4-fluorohistidine, the calculated chemical shift upon protonation of the π-tautomer is 2 ppm upfield (*cf.* 3.74 ppm upfield experimentally). However the calculated chemical shift change of the τ-tautomer (calculated to be highly unstable) is 8.6 ppm *downfield*, opposite to the direction of the ^{19}F chemical shift observed in the experimental titration.

In summary, the simulated ^{19}F NMR chemical shifts upon acid titration for the stable tautomers are in good agreement with experimental results. However, the less stable tautomers show divergence from experimental results in the simulated acid titration, with 4-FHis giving shifts in the opposite direction and 2-FHis showing shifts in the same direction but with a bigger change in chemical shifts than what is observed.

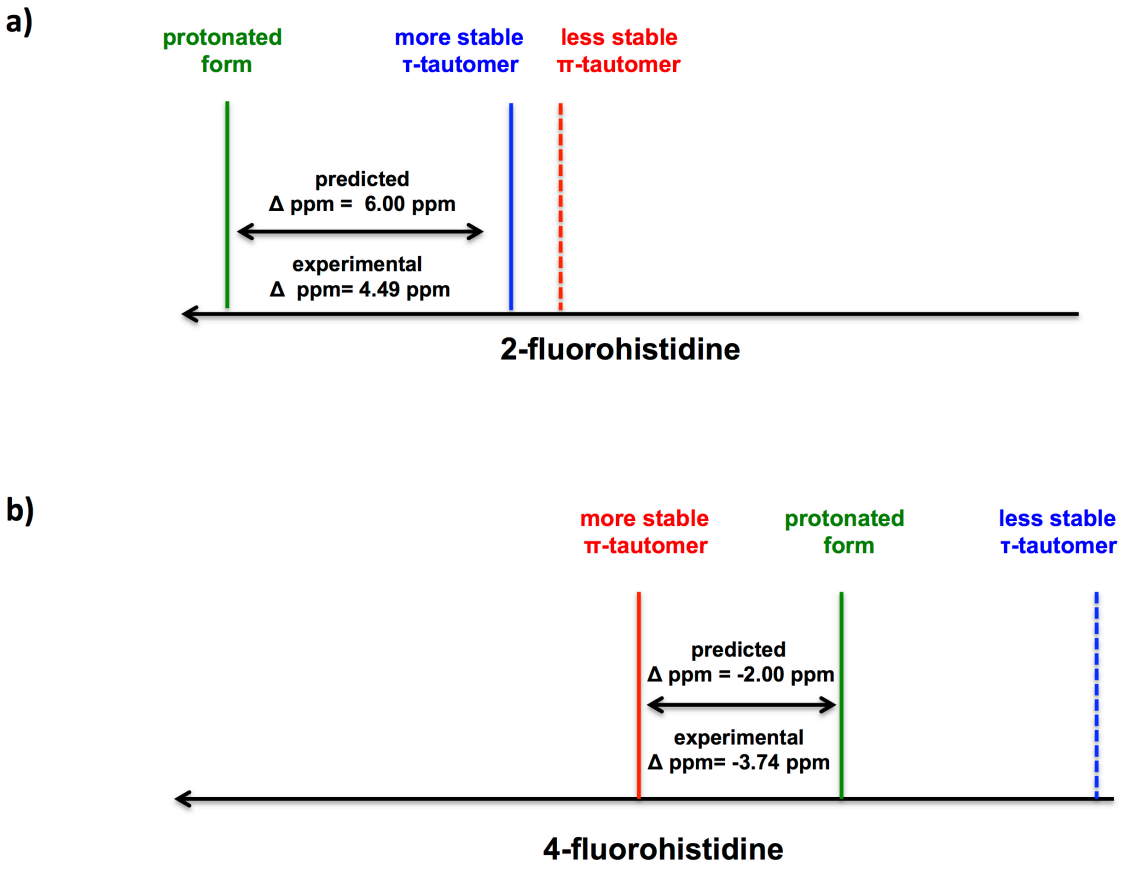


Figure 2.3. Calculated fluorine chemical shifts for the τ and π-tautomers and protonated forms of zwitterionic (a) 2-fluorohistidine and (b) 4-fluorohistidine.

2.3.1. What are the effects of environment on tautomeric stabilities and fluorine chemical shifts?

Calculations were carried out on the Gly-2FHis-Gly and Gly-4FHis-Gly tripeptides in different dielectric environments, using CPCM-modeled acetonitrile and tetrahydrofuran (THF), to see the effect on tautomeric stabilities and fluorine chemical shifts. The results are provided in Table 2.2. For 2-fluorohistidine, the tautomeric equilibrium slightly shifted towards π-tautomer as moving from aqueous to more hydrophobic environment. The relative stability of π-tautomer is observed to be ~1.35

kcal/mol higher than the τ-tautomer in acetonitrile and THF (compared to 1.70 kcal/mol in water). These relatively low energy differences in less polar environments, in combination with a hydrogen-bonding environment that favors stabilization of the π-tautomer over the τ-tautomer, are supported by experimental results. In the crystal structure of 2-fluorohistidine labeled anthrax protective antigen, the π-tautomer of 2-fluorohistidine is indicated to be more stable in some interior sites (*i.e.* 2-FHis616 in the hydrophobic core of domain 4, stabilized by a hydrogen bond to N3).[26] This is similar to histidine, which favors the π-tautomer in some protein environments. In less polar environments, the ^{19}F chemical shifts of both tautomers of 2-fluorohistidine becomes slightly more shielded, with changes on the order of 0.6 ppm.

Table 2.2. Effect of dielectric environment on ^{19}F chemical shifts (relative to C_6F_6) and relative tautomeric stabilities.

Molecule	Tautomer	In water (ε=78.4)		In acetonitrile (ε=35.7)		In tetrahydrofuran (ε=7.4)	
		^{19}F δ_{ppm}	ΔE	^{19}F δ_{ppm}	ΔE	^{19}F δ_{ppm}	ΔE
2F-tripeptide	**τ**	53.19 ppm	0.00 kcal/mol	52.83 ppm	0.00 kcal/mol	52.91 ppm	0.00 kcal/ mol
	π	52.22 ppm	1.70 kcal/mol	52.06 ppm	1.34 kcal/mol	51.32 ppm	1.36 kcal/mol
4F-tripeptide	**τ**	6.52 ppm	4.85 kcal/mol	6.48 ppm	4.74 kcal/mol	6.57 ppm	3.91 kcal/mol
	π	19.21 ppm	0.00 kcal/mol	19.31 ppm	0.00 kcal/mol	20.16 ppm	0.00 kcal/mol

For 4-fluorohistidine, the relatively hydrophobic environment of THF shifts the tautomeric equilibrium slightly towards the τ-tautomer with an energy difference of 3.91 kcal/mol higher than the π, acetonitrile shows a small change 4.75 kcal/mol, relative to 4.85 kcal/mol in water. These results indicate that in 4-fluorohistidine, the dielectric is seen to have very small influence on the ^{19}F chemical shift of the τ-tautomer (less than 0.1 ppm), while less polar environments deshield the ^{19}F nucleus of the 4-FHis π-tautomer from 19.21 ppm in water to 19.31 in acetonitrile and 20.16 ppm in THF.

2.3.2. Effects of intramolecular hydrogen bonding on fluorine chemical shifts and tautomeric stabilities.

"Capped" histidine (*N*-acetyl-FHis-O-methyl ester or NAc-FHis-OMe), which has no intramolecular hydrogen bonds with the imidazole ring, was compared to zwitterionic histidine, for which the lowest-energy conformation in implicitly-modeled solvent has hydrogen bonds between the imidazole ring (side chain) and amino or carboxylate moieties of the amino acid. Fig. 2.4 shows the lowest-energy optimized geometries of all tautomers of the capped and zwitterionic fluorohistidines. The relative energies and ^{19}F chemical shifts for these species are presented in Table 2.3.

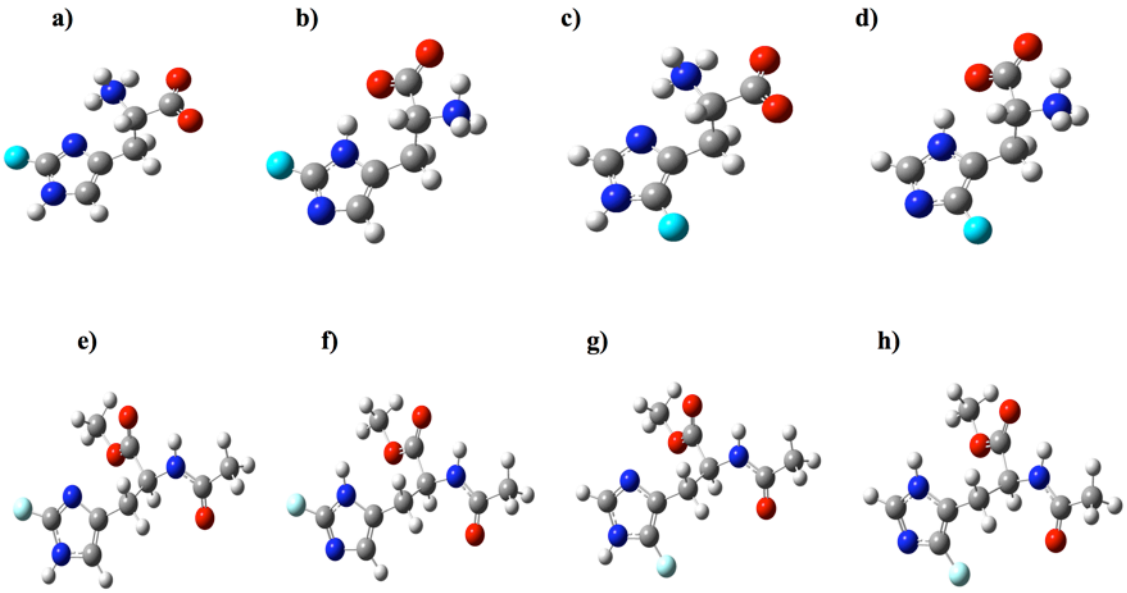


Figure 2.4. Lowest-energy optimized geometries for "capped" (NAc-FHis-OMe) and zwitterionic fluorohistidines: a) τ-tautomer of 2F-His b) π-tautomer of 2F-His c) a) τ-tautomer of 4F-His b) π-tautomer of 4F-His e) τ-tautomer of capped 2F-His f) π-tautomer of capped 2F-His g) τ-tautomer of capped 4F-His h) π-tautomer of capped 4F-His. Carbons are coloured grey, hydrogens white, nitrogens blue, oxygens red, and fluorines aqua.

Table 2.3. Effect of intramolecular hydrogen bonding on ^{19}F chemical shifts (in ppm, relative to C_6F_6) and relative tautomeric energies. Zwitterionic fluorohistidines have hydrogen bonds, while capped fluorohistidines do not.

Molecule	Tautomer	with intramolecular hydrogen bonding (zwitterion)		without intramolecular hydrogen bonding (capped)	
		^{19}F δ_{ppm}	ΔE	^{19}F δ_{ppm}	ΔE
2-fluorohistidine	**τ**	55.89 ppm	0.00 kcal/mol	55.50 ppm	0.00 kcal/mol
	π	52.28 ppm	1.70 kcal/mol	52.39 ppm	1.80 kcal/mol
4-fluorohistidine	**τ**	6.26 ppm	3.80 kcal/mol	6.83 ppm	4.69 kcal/mol
	π	16.89 ppm	0.00 kcal/mol	20.21 ppm	0.00 kcal/mol

The relative tautomeric stabilities of 2-fluorohistidine change by < 0.2 kcal/mol with the intramolecular hydrogen bond. Also, the calculations show that for 2-fluorohistidine, the effects of intramolecular hydrogen bonding on fluorine shifts are minor. In the case of τ-tautomer, the fluorine chemical shifts are affected more with hydrogen bonding showing a deshielding of ~0.4 ppm. However, in the case of the π-tautomer, the change is very little, with increased shielding of ~0.1 ppm with hydrogen bond.

In contrast, 4-fluorohistidine appears to be more sensitive to hydrogen bond effects: For the tautomeric stability, hydrogen bonding contributes 0.9 kcal/mol stabilization to the τ-tautomer of 4-fluorohistidine; yet the π-tautomer remains substantially more stable, by ~4 kcal/mol. Regarding the fluorine chemical shifts, of the stable π-tautomer has a difference of >3 ppm between forms, with hydrogen-bonded zwitterionic 4-fluorohistidine being more shielded at 16.89 ppm than the non-hydrogen-bonded capped 4-fluorohistidine at 20.21 ppm. This is another "reverse" effect in the fluorine chemical shift of 4-FHis, since hydrogen bonding would be expected to inductively shift electron density away from the fluorine.

2.3.3. Effect of C5 substitution on stability and fluorine chemical shifts

Finally, substitution at C5 of the fluoroimidazole was used to determine the sensitivity of tautomeric stabilities and fluorine chemical shifts to structure. C5 substitution can probe the sensitivity of fluorine chemical shifts to changes in primary and secondary structure, both of which give rise to changes in the wave function near the nucleus of interest. This is described by Oldfield and co-workers as a short-range contribution to chemical shifts, which propagates through the bonding framework.[27] In

this case, the tripeptide was compared with 2F- and 4F-(5-methyl)-imidazole and -(5-trifluoromethyl)-imidazole, providing information about electron-donating and –withdrawing effects, respectively. The relative tautomeric stabilities and fluorine chemical shifts for this series are presented in Table 2.4.

Table 2.4. Effects of C5 substitution on tautomeric stabilities and ^{19}F chemical shifts.

Molecule	-(5-methyl)-imidazole		-tripeptide		-(5-trifluoromethyl)-imidazole	
	^{19}F δ_{ppm}	ΔE	^{19}F δ_{ppm}	ΔE	^{19}F δ_{ppm}	ΔE
2F-τ	53.20 ppm	0.00 kcal/mol	53.19 ppm	0.00 kcal/mol	56.22 ppm	0.00 kcal/ mol
2F-π	52.41 ppm	0.02 kcal/mol	52.22 ppm	1.70 kcal/mol	58.96 ppm	3.50 kcal/mol
4F-τ	1.71 ppm	5.62 kcal/mol	6.52 ppm	4.85 kcal/mol	21.82 ppm	2.63 kcal/mol
4F-π	13.62 ppm	0.00 kcal/mol	19.21 ppm	0.00 kcal/mol	36.45 ppm	0.00 kcal/mol

As can be seen in Table 2.4, the relative stabilities of tautomers are dramatically affected by substitution at C5. In the case of 2-fluoro-(5-methyl)-imidazole, electron-donating substitution substantially stabilizes the π-tautomer, resulting in isoenergetic tautomers, while the electron-withdrawing CF_3 group strongly destabilizes the π-tautomer, relative to the τ-tautomer. The effect is opposite for 4-fluorohistidine: electron-donating substituents further stabilize the π-tautomer, while electron withdrawing groups drop the relative energy of the τ-tautomer. However, the energy differences between 4-

fluoroimidazole tautomers always remain such that the τ-tautomer remains <1% population at equillibrium.

The fluorine chemical shifts of 4-fluorohistidine are shown to depend dramatically on C5 substitution: electron donation causes an upfield shift (~5 ppm for both tautomers), while electron-withdrawing shifts the signal downfield (>15 ppm). However, in 2-fluorohistidine the ^{19}F NMR spectra are not as dramatically influenced by C5 substitution. The differences between tripeptide and (5-methyl)-imidazole are slight (<0.2 ppm), and CF_3, a rather large perturbation on electronic structure, shifts the signal 3 ppm and 6.7 ppm for the τ- and π-tautomers, respectively.

2.3.4. Tautomeric stabilities of 4-fluorohistidine in understanding proton transfer mechanism in RNA hydrolysis[28]

Ribonuclease A (RNase A) is a 13.7 kDa enzyme that has long served as a model for understanding enzymatic mechanisms[29-31] and more recently, the role of dynamics in enzyme catalysis.[32,33] One of the seminal contributions to understanding proton transfer in RNA hydrolysis was made by Jackson *et al*.[34] In their groundbreaking report, 4-fluorohistidine (4-FHis),[15, 35] was incorporated at the His12 and His119 sites during the total chemical synthesis of RNase A. These two histidines participate in the two steps of RNA hydrolysis within the active site, as shown in Fig. 2.5. In the labelling study with 4-fluorohistidine, it was found that catalysis was optimal at low pH values, consistent with the pK_a of 4-FHis, and supporting the notion that proton transfer in RNase A is not rate-limiting. However, some ambiguity in the histidines' roles, and the overall enzymatic mechanism, remain.[36, 37] Recent analysis by Harris and co-workers declares, 'effects of 4-

fluorohistidine substitution on catalysis are only ca. 10-fold, suggesting a relatively small Brønsted effect, or that a nonchemical step is partially rate-limiting'.[38]

The strong thermodynamic preference for the π-tautomer of 4-fluorohistidine looks to have significant impact for interpretation of the results of 4-fluorohistidine labeled RNase A. The tautomeric states of histidine in RNase A were assigned by multiple researchers using NMR spectroscopy,[9, 39] neutron diffraction,[40] and X-ray diffraction [41-46] (which of course does not definitively assign proton positions). Perhaps most relevant to understanding enzyme function is the tautomeric determination of the two RNase A active site histidines in aqueous solution (His12 and His119, shown in Fig. 2.5). Of these two residues replaced by 4-fluorohistidine in the work by Jackson *et al*.,[34] His12 in solution state is identified to be 100% in the π-tautomer, via ^{13}C NMR and determination of micro-pK_a constants. [9] In contrast, His119 in solution has been identified to exist in a τ:π tautomeric ratio of 85:15,[9, 39] which is close to the innate tautomeric stability of histidine in water (80:20) [30]. The ramifications of the strong favorability for the π-tautomer of 4-fluorohistidine are two fold in the ribonuclease A experiment. First, it can be seen from the results of Jackson *et al.* that catalysis by fluoro-labeled RNase A is efficient when both His12 and His119 are in the π-tautomeric form.[34] Secondly, it raises the question whether tautomerization contributes to the overall catalytic rate in wild-type RNase A, since the His119 tautomeric ratio is 85:15 τ:π, yet proton transfer presumably takes place via the π- tautomer when RNase A is labeled with 4-fluorohistidine. If his- tidine tautomerism is a mechanistic step in ribonuclease catalysis, then labeling with 4-fluorohistidine obviates tautomerization. Which tautomer of His119 participates in the native enzyme during catalysis appears to be an open question, and

should be considered when designing and interpreting future simulations and experiments. With these present results, the seminal work by Wells and co-workers can be re-evaluated for even more, valuable mechanistic information.

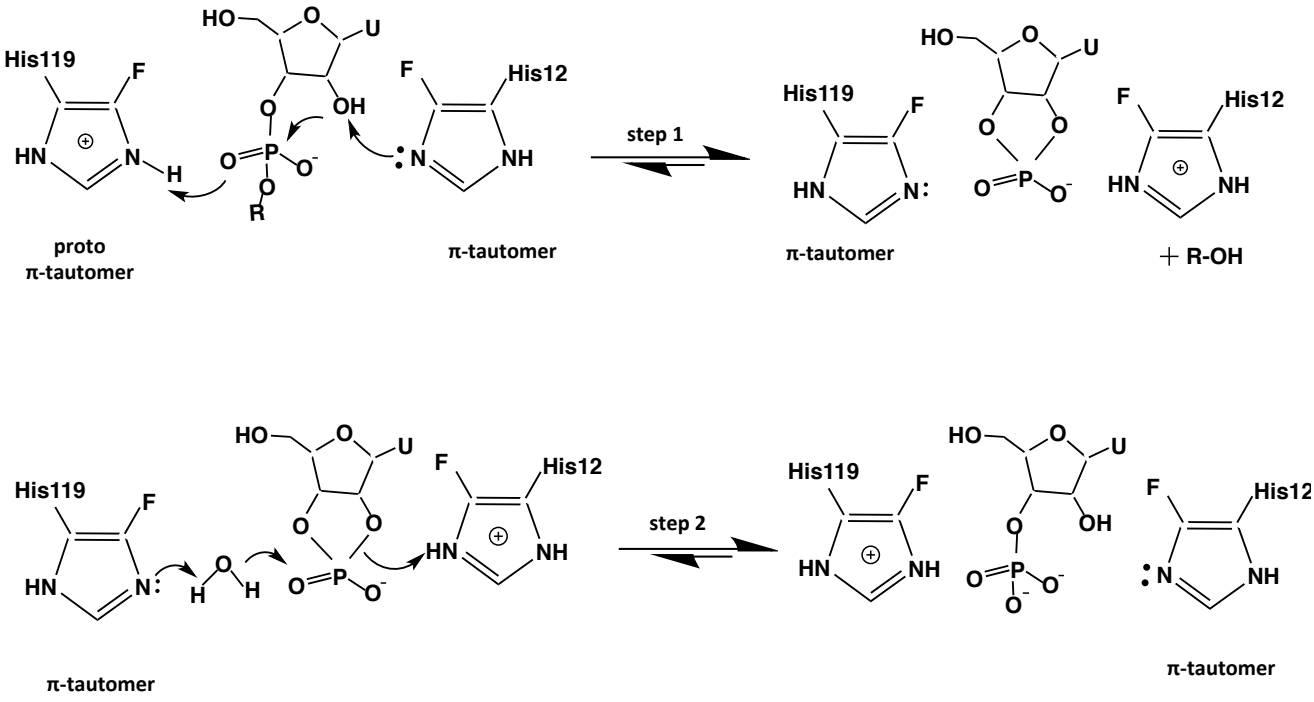


Figure 2.5. Role of 4-fluoro-His119 and -His12 in the two proton transfer steps of RNA hydrolysis.

2.3.5. The origin of pK_a differences between fluorohistidine isomers

^{19}F and ^{1}H NMR titrations by Yeh *et al*. showed that the aqueous fluorohistidine isomers have lower pK_a values of ~1 and ~2.5 for 2-fluorohistidine and 4-fluorohistidine, respectively, compared to canonical histidine (pK_a ~6)[25]. The differences in the relative

tautomeric stabilities between 2-fluorohistidine and 4-fluorohistidine give rise to an entropic factor in the differences in pK_a between two fluoroisomers. Neutral 2-FHis is entropically stabilized by its two thermally-accessible tautomeric states, whereas protonated imidazoliums and netural 4-fluorohistidine are not. Using the Boltzmann definition of entropy, the molar free energy of neutral fluorohistidine is reduced by RT ln 2 (drop in pK_a of 0.3) when the two tautomers are isoenergetic (degeneracy=2). In other cases, the closer the tautomers are in energy (and, hence, the closer their "microscopic" or individual pK_a values),[47, 48] the greater the reduction in observed pK_a value relative to microscopic pK_a. In summary, the pK_a of 2-fluorohistidine is reduced by tautomerism, whereas the same effect is not predicted in 4-fluorohistidine due higher energy gap.

Next, structural contributions to pK_a differences between the most stable tautomers of 2-FHis and 4-FHis were considered. The origins are not immediately obvious by considering resonance and inductive effects. To study this computationally, the gas phase energies of capped 2-fluorohistidine and 4-fluorohistidine, and their conjugate acids were calculated and are presented in Table 5. The gas phase energies ($E_{conjugate\ acid} - E_{neutral}$) of the two isomers were found to be virtually identical: within 0.2 kcal/mol of one another. Therefore, solvation appears to be a driving force for the differences in pK_a. Indeed, calculations confirm this, with protonation of aqueous 4-fluorohistidine being more favorable than protonation of 2-fluorohistidine by a difference of 1.4 kcal/mol. This is in very good agreement with the pK_a difference of ~1-2 reported for the isomers,[25] since a pK_a difference of 1 corresponds to a free energy difference of 1.37 kcal/mol at 300 K.

Table 2.5. Calculated pK_a differences and solvation energies (ZPE) between fluorohistidine isomers in kcal/mol.

Basicity	**capped 2FHis**		**capped 4FHis**	
ΔE $[E_{FHis^+(g)}-E_{FHis(g)}]$ **gas phase**	-59.99 kcal/mol		-60.17 kcal/mol	
ΔE $[E_{FHis^+(aq)}-E_{FHis(aq)}]$ **aqueous**	-28.79 kcal/mol		-30.23 kcal/mol	
solvation energy $[E_{(aq)}-E_{(g)}]$	**2FHis (τ)**	**2FHis (protonated)**	**4FHis (π)**	**4FHis (protonated)**
	-13.22 kcal/mol	-54.28 kcal/mol	-14.57 kcal/mol	-56.90 kcal/mol

The key to understanding how solvation energies affect pK_a lies in the unexpectedly favorable solvation energy of neutral 2-fluorohistidine, (given its smaller dipole of 5.2 Debye relative to the 4-FHis dipole of 7.1 Debye) which is very similar to ΔE_{solv} of 4-fluorohistidine (given in Table 2.5). Of course, the value of ΔE_{solv} depends on the charge distribution within the molecule. The charges for the imidazole side chains of protonated and neutral fluorohistidine tripeptides are presented in Fig. 2.6. The charges are color-coded, with green negative charges, red positive charges, and dark shading for near-neutral atoms.

The salient difference between 2-fluorohistidine (τ-tautomer) and all other species is the polarization of the C4-H4 bond, which has a carbon charge of -0.35 and hydrogen charge of +0.24. This charge difference is somewhat less, but on the order of, the N-H bond polarity seen in both fluorohistidine imidazole N-H bonds. The other charges, in the nitrogen with the lone pair, C-F bonds, and N-H bonds, are fairly similar among all species. Thus, the favourable solvation of the additional polar bond (C4-H4) in the imidazole ring (τ-tautomer only) of neutral 2-fluorohistidine appears to reduce the driving force in ΔE_{solv} for protonation, making 2-fluorohistidine more acidic than 4-fluorohistidine.

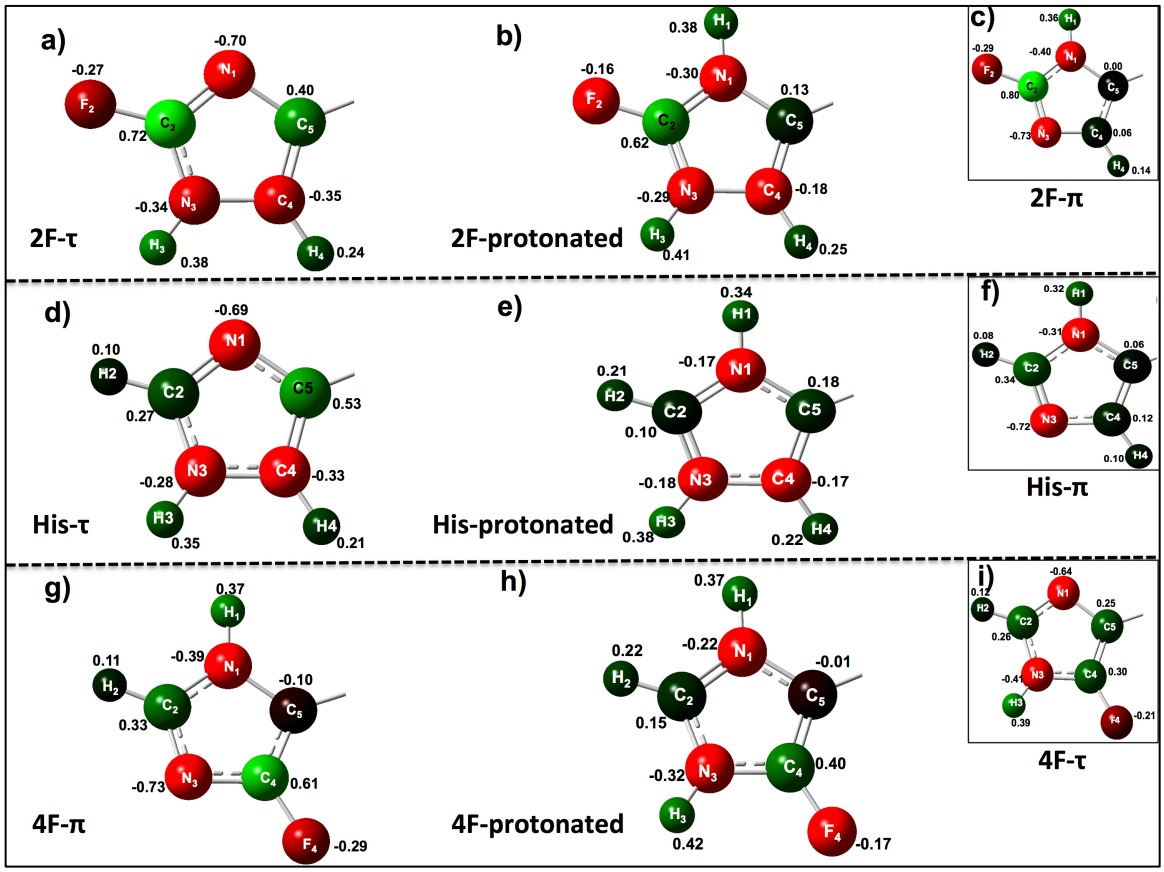


Figure 2.6. Calculated charges for imidazole side chains in Gly-FHis-Gly peptides for a) τ-tautomer of 2-fluorohistidine b) protonated form of 2-fluorohistidine c) π-tautomer of 2-fluorohistidine; d) τ-tautomer of histidine e) protonated form of histidine f) π-tautomer of histidine; g) π-tautomer of 4-fluorohistidine h) protonated form of 4-fluorohistidine i) τ-tautomer of 4-fluorohistidine.

The polarized C-H bond in the τ-tautomer of the 2-fluorohistidine side chain could be predicted to act as a hydrogen bond donor (see Ch.4), a distinctive chemical feature not seen in either tautomer of 4-fluorohistidine, nor the less stable π-tautomer of 2-fluorohistidine. Given that the tautomeric equilibrium is predicted to shift with protein environment (indicated in Tables 1, 2, and 3), this change in C-H bond polarity/hydrogen bonding ability that accompanies tautomerization of 2-fluorohistidine may have consequences that impact protein structure. However, it is interesting to note that

calculations with the same level of theory on canonical histidine indicate that the properties of 2-fluorohistidine track very well with those of histidine. The C4-H4 bond of histidine is also polar (charges of -0.275 for C and +0.201 for H) in the τ-tautomer, while the same bond in the π-tautomer is nonpolar (like 2-fluorohistidine). Hydrogen bonding between C2-H2 ($C_{\varepsilon 1}$-H) of the π-tautomer of histidines and amide carbonyls in serine proteases have been reported experimentally.[49] Reports of C4-H4 ($C_{\delta 2}$-H) hydrogen bonds are fewer.[50] Recently, Kubarych, Pecoraro, and co-workers reported on the polarity of the $C_{\delta 2}$-H bond (on a Zn-bound histidine residue) in a *de novo* peptide, pointing out that the QM results are substantiated by microwave spectroscopy of imidazole.[51, 52] It is seen below that the higher charge density of the C4 ($C_{\delta 2}$) carbon corresponds with a more shielded ^{13}C NMR shift for this carbon in the τ-tautomer imidazole ring in both histidine and 2-fluorohistidine.

2.3.6. Spectroscopic signals of the tautomers of 2-fluorohistidine

The results presented above indicate the π-tautomer of 4-fluorohistidine is so much more stable than the τ-tautomer that the τ-tautomer is not predicted to be observed. However, calculations indicate small relative energy differences between the τ- and π-tautomers of 2-fluorohistidine, and both tautomers of 2-fluorohistidine have been observed in the crystal structure of labelled anthrax protective antigen.[26] Therefore, the discussion below focuses primarily on distinguishing among the two tautomers of 2-fluorohistidine in solution. ^{13}C NMR chemical shifts for 2-fluorohistidine and 4-fluorohistidine in the Gly-FHis-Gly tripeptide are presented in Table 2.6.

Table 2.6. Carbon chemical shifts (in ppm, relative to TMS) for imidazole carbons in Gly-FHis-Gly tripeptide.

Gly-2FHis-Gly	(τ-tautomer)	(π-tautomer)	(protonated form)
C5(C_Y)	142.8 ppm	133.5 ppm	136.0 ppm
C4($C_{\delta 2}$)	118.3 ppm	129.7 ppm	122.0 ppm
C2 ($C_{\varepsilon 1}$)	157.4 ppm	156.9 ppm	153.0 ppm
Gly-4FHis-Gly	**(τ-tautomer)**	**(π-tautomer)**	**(protonated form)**
C5(C_Y)	121.2 ppm	112.1 ppm	118.4 ppm
C4($C_{\delta 2}$)	152.8 ppm	162.3 ppm	152.4 ppm
C2 ($C_{\varepsilon 1}$)	136.1 ppm	137.0 ppm	136.7 ppm

Previous work by Oldfield,[7] Scheraga, [2] and co-workers have shown that the tautomeric forms of canonical histidine depend strongly on the peptide environment (as seems to be the case for 2-fluorohistidine). Moreover, in their work, theoretical methods used to predict tautomeric fractions (relative stabilities) and the ^{13}C NMR chemical shifts of each tautomer have been shown to be reliable. Oldfield and co-workers used a large data set of histidine-containing peptides and provided method-dependent scaling factors for each of the carbons in the histidine imidazole ring.

As is seen in canonical histidine, the C2 chemical shift of 2-fluorohistidine and 4-fluorohistidine are invariant with tautomeric form, while the C4 and C5 chemical shifts distinguish tautomers.[2, 7] The C2 shift for both tautomers of 2-fluorohistidine is ~157 ppm: 157.4 ppm for the τ-tautomer and 156.9 ppm for the π-tautomer. The values of the

C4 and C5 chemical shifts exhibit a wide spread for the τ-tautomer, with chemical shifts at 118.3 (C4/Cδ2) and 142.8 (C5/C_γ) ppm. In contrast, the C4 and C5 chemical shifts are closer together in the π-tautomer, at 129.7 and 133.5 ppm, respectively. This range of chemical shifts between the 2-fluorohistidine tautomers (~11 ppm for C4 and ~10 ppm for C5) is on the order of the range observed for canonical histidine tautomers in a Gly-His-Gly peptide studied by Scheraga and co-workers.[2]

The range of chemical shifts for the carbons in the two tautomers of 2-fluorohistidine directly reflect the charge distribution on the carbons. The C4 carbon of the τ-tautomer is significantly more shielded (118.3 ppm) than the other carbons, reflecting its calculated negative charge (-0.4). As expected the C2 carbon bonded to fluorine is the most deshielded (~157 ppm for both tautomers), reflecting positive charges of +0.72 and +0.8 for the τ and π-tautomers, respectively. The polarization of C4 in the τ-tautomer coincides with polarization of the adjacent carbon, C5, which has a charge of +0.40 and a relatively deshielded shift at 142.8 ppm. In contrast, the charges on C4 and C5 of the less polar π-tautomer are nearly zero, and the chemical shifts are close together, at 129.7 and 133.5 ppm.

An experimental ^{13}C NMR spectrum was acquired for zwitterionic 2-fluorohistidine in water, and the imidazole ring shifts are as follows: 127.9 ppm, 113.6 ppm, 150.0 ppm for C5, C4 and C2 carbons respectively. The C5, C4, and C2 shifts for capped 2-fluorohistidine (close in structure to the aqueous zwitterion), provided in Table 2.7, are calculated to be 141.4 ppm, 118.9 ppm, 157.1 ppm respectively. Using a scaling factor approach in the same vein as Oldfield and co-workers, the recommended scaling factors for BHandHLYP/6-311++G(d,p) calculated ^{13}C chemical shifts of fluorohistidines

are as follows: 1.11, 1.05, 1.05 (C5, C4, C2). This is quite good agreement, indicating that the BHandHLYP functional may be a fairly reliable method for ^{13}C chemical shifts (as it is for ^{19}F).

Table 2.7. Carbon chemical shifts (in ppm, TMS reference) for imidazole carbons in aqueous capped 2-fluorohistidine and 4-fluorohistidine.

capped 2FHis	**(τ-tautomer)**	**(π-tautomer)**	**(protonated form)**
C5(C_γ)	141.4 ppm	133.1 ppm	137.0 ppm
C4($C_{\delta 2}$)	118.9 ppm	130.2 ppm	121.7 ppm
C2 ($C_{\varepsilon 1}$)	157.1 ppm	157.2 ppm	153.5 ppm
capped 4FHis	**(τ-tautomer)**	**(π-tautomer)**	**(protonated form)**
C5(C_γ)	120.0 ppm	112.1 ppm	119.8 ppm
C4($C_{\delta 2}$)	153.3 ppm	162.4 ppm	151.3 ppm
C2 ($C_{\varepsilon 1}$)	136.1 ppm	137.5 ppm	137.5 ppm

Although ^{13}C chemical shifts should not be needed to differentiate the tautomers of 4-fluorohistidine, given that the τ-tautomer is considerably unstable, it is worth remarking on the predicted spectra. The π-tautomer of 4-fluorohistidine exhibits a broader spread in carbon chemical shifts than the τ-tautomer. We have recently published the experimental ^{13}C spectrum of aqueous 4-fluorohistidine, and calculated values of the π-tautomer are in good agreement.[28] The C5 carbon has a relatively shielded nucleus at 101 ppm, and C4 (attached to fluorine) is the most deshielded at 152 ppm. The carbon chemical shifts again correlate with charge and polarity: the π-tautomer C-F

bond is more polar than the τ-tautomer, with a positive charge on the C4 carbon of +0.60 (-0.29 on F). In contrast, the most shielded carbon (C5) has a calculated charge of -0.10.

2.4. DISCUSSION

It has been observed that typically, biosynthetic incorporation of 2-fluorohistidine is more successful than biosynthetic incorporation of 4-fluorohistidine. [53-55] Although it has been argued that 4-FHis may not be a good substrate for the HisRS of *E.coli* for steric reasons,[30] the results presented above indicate that the stabilities *and* the molecular properties (charge distribution, C-H bond polarity) of the tautomers of 2-fluorohistidine are remarkably similar to canonical histidine, making its substitution into biomolecules more amenable than 4-fluorohistidine. Of course, the pK_a of 2-fluorohistidine is dramatically altered, providing a useful mechanistic probe for proton transfer steps in biological processes.

2.4.1. Studying tautomeric form in proteins. For 2-fluorohistidine, the τ-tautomer is more stable in most environments, as it is for native histidine. In the crystal structure of 2-fluorohistidine-labelled anthrax protective antigen, a mixture of τ- and π-tautomers were observed, with the identified π-tautomer (His616) buried in the hydrophobic core of domain 4. This is consistent with our findings that lower dielectric environments shift the tautomeric equilibrium of 2-fluorohistidine toward the π-tautomer. Furthermore, hydrogen bonding was shown in our results to shift the equilibrium toward the π-tautomer, which also corroborates with the hydrogen-bond stabilized His616 in π form. So it is anticipated that local hydrogen bonds and dielectric environment will dictate tautomeric form of 2-fluorohistidine in proteins, as is often the case for canonical

histidine in proteins. The introduction of the heavy atom (fluorine) in 2-fluorohistidine provides a distinctive marker in X-ray crystallography, allowing for histidine tautomeric form to be distinguished with low ambiguity in solved structures.

In contrast to 2-fluorohistidine, this work indicates that equilibrium tautomerism of 4-fluorohistidine is unfavorable in all environments. In all cases, the π-tautomer is favored by ~4 kcal/mol. This preference for the π-tautomer helps explain why 4-fluorohistidine could be successfully incorporated into the RNase A protein in work by Wells and co-workers.[54] Of the two active site histidines, His12 is indicated to be 100% π-tautomer,[39] while the tautomeric form of His119 during proton transfer is unknown (see our recent letter for further discussion).[28] Having a viable enzyme with 4-FHis substitution may indicate the π-tautomer is the active form during mechanistic steps. Overall, these results indicate that fluorohistidines may be poised for studying the role of histidine tautomerization or protonation state in biological mechanisms, as the τ-tautomer of 2-fluorohistidine is more stable, but tautomeric form can easily be identified by crystallography in labeled proteins and via ^{19}F and ^{13}C NMR. Meanwhile the π-tautomer is strongly favored in 4-fluorohistidine, and the effect of this tautomeric stability perturbation on protein function can be evaluated.

2.4.2. ^{19}F NMR spectra reflect protein environment. Fluorine chemical shifts are noted for their high sensitivity to local environment. Differences in 1º and 2º structure in proteins may give rise to conformational diversity and different inductive effects on the imidazole side chain, that, in combination with varying dielectric and hydrogen bonding environments, result in the range of ^{19}F signals observed for folded and unfolded proteins. For 4-fluorohistidine, differences in 1º and 2º structure in proteins may well be the

primary source of changes in fluorine chemical shifts in the folded protein environment. The π-tautomer of 4-fluorohistidine appears to be quite sensitive to hydrogen bonding and inductive effects in the imidazole ring, exhibiting changes upward of 4 ppm for each perturbation. However, dielectric environment has a relatively minor effect on 4-fluorohistidine chemical shift (up to 1 ppm). All environmental effects result in much smaller fluorine chemical shift changes for both tautomers of 2-fluorohistidine: on the scale of 0.1-0.5 ppm per environmental change. This is in line with the findings of Gerig and co-workers, who postulated that diversity in fluorine chemical shifts in denatured/folded proteins arises from a sum of changes in chemical shift due to factors such as van der Waals interactions, hydrogen bonds, electric fields, and local magnetic anisotropies. [56]

2.4.3. Probing proton transfer in proteins. The altered pK_a values of fluorohistidines can be used to determine whether histidines act as general acids and bases in enzymes.[54] The differences in pK_a between the two isomers, with 2-fluorohistidine being less basic than 4-fluorohistidine, can be attributed primarily to solvation effects. Calculated gas phase pK_a values are quite similar, but the aqueous pK_a of 4-fluorohistidine is higher than 2-fluorohistidine. This arises from favorable solvation of the polar charge distribution of 2-fluorohistidine (τ-tautomer) that reduces the driving force for protonation. The thermally-accessible tautomeric equilibrium of 2-fluorohistidine also lowers the pK_a value of this isomer somewhat (up to 0.3 pK_a units, when tautomers are isoenergetic). Since tautomerism is not favored by 4-fluorohistidine, the pK_a value for this isomer is not lowered by tautomeric equilibrium. In RNase A, the pK_a of 4-fluorohistidines in the

active site were indicated to be over 3.5 (*cf.* ~2 in aqueous solution), showing the influence of protein environment.[54]

2.5. CONCLUSIONS

Fluorohistidines can contribute to biophysical studies of proteins in numerous ways. They have the potential to be powerful probes for probing tautomeric state and the contributions of tautomeric form to protein stability and enzyme mechanism. In crystallography, the fluorine atoms' high electron density in the imidazole ring allows for definitive assignment of carbons and nitrogens, facilitating assignment of tautomeric form as well. The tautomeric stability of 2-fluorohistidine is predicted to shift with environment, with low dielectric environment shifting the tautomeric equilibrium toward the π-tautomer. While tautomeric form and tautomerism are acknowledged to be important factors in histidine's chemical versatility and role in protein structure-function, identification of tautomeric form (at equilibrium and in rate-determining steps) has proven difficult in large biomolecular systems. 4-fluorohistidine is poised to serve as a powerful probe of the role of tautomeric form in structure and catalytic steps. In addition, its altered pK_a provides a tool for determining how proton transfer affects catalysis, and at pH values that are less acidic than those required for 2-fluorohistidine labeled proteins.

Additionally, biosynthetic incorporation of either fluorohistidine isomer introduces a valuable spectroscopic label for ^{19}F NMR studies. Fluorine NMR spectra provide vastly simplified information relative to traditional protein NMR spectra, with the ability to report on site-specific dynamics[57] and ligand-binding events.[58] While the origin of changes in fluorine chemical shifts are not systematically understood, this work

provides some insight on how environment influences chemical shift. The ^{19}F signal of 2-fluorohistidine is shown to behave "normally"—in the way that would be predicted for ^{1}H chemical shifts. Deshielding is observed with hydrogen bonds, higher dielectric, and nearby electron withdrawing groups. In contrast, the ^{19}F chemical shifts of 4-fluorohistidine are "reverse", with environmental changes that normally deshield nuclei (such as hydrogen bonding, higher dielectric) resulting in shielding of the fluorine nucleus. Fluorohistidines are poised to serve as powerful biophysical probes of protein structure-dynamics-mechanisms. This work can serve as a guide for interpretation of mechanistic experiments and spectroscopic data with fluorohistidine-labeled biomolecules.

REFERENCES

[1] Reynolds, W. F., Peat, I. R., Freedman, M. H., and Lyerla, J. R. (1973) Determination of tautomeric form of imidazole ring of L-histidine in basic solution by C-13 magnetic-resonance spectroscopy. *J Am Chem Soc 95*, 328-331.

[2] Vila, J. A., Arnautova, Y. A., Vorobjev, Y., and Scheraga, H. A. (2011) Assessing the fractions of tautomeric forms of the imidazole ring of histidine in proteins as a function of pH. *Proc Natl Acad Sci U S A 108*, 5602-5607.

[3] Li, S. H., and Hong, M. (2011) Protonation, tautomerization, and rotameric structure of histidine: A comprehensive study by magic-angle-spinning solid-state NMR. *J Am Chem Soc 133*, 1534-1544.

[4] Hansen, A. L., and Kay, L. E. (2014) Measurement of histidine pK(a) values and tautomer populations in invisible protein states. *Proc Natl Acad Sci U S A 111*, 1705-1712.

[5] Wolff, N., Deniau, C., Letoffe, S., Simenel, C., Kumar, V., Stojiljkovic, I., Wandersman, C., Delepierre, M., and Lecroisey, A. (2002) Histidine pK(a) shifts and changes of tautomeric states induced by the binding of gallium-protoporphyrin IX in the hemophore HasA(SM). *Protein Sci 11*, 757-765.

[6] Bhattacharya, S., Sukits, S. F., MacLaughlin, K. L., and Lecomte, J. T. J. (1997) The tautomeric state of histidines in myoglobin. *Biophys J 73*, 3230-3240.

[7] Cheng, F., Sun, H. H., Zhang, Y., Mukkamala, D., and Oldfield, E. (2005) Solid state C-13 NMR, crystallographic, and quantum chemical investigation of chemical shifts and hydrogen bonding in histidine dipeptides. *J Am Chem Soc 127*, 12544-12554.

[8] Wasylishen, R. E., and Tomlinson, G. (1975) pH-dependence of C-13 chemical-shifts and C-13, H coupling-constants in imidazole and L-Histidine. *Biochem J 147*, 605-607.

[9] Tanokura, M. (1983) 1H-NMR study on the tautomerism of the imidazole ring of histidine residues. II. Microenvironments of histidine-12 and histidine-119 of bovine pancreatic ribonuclease A. *Biochim Biophys Acta 742*, 586-596.

[10] Shimba, N., Takahashi, H., Sakakura, M., Fujii, I., and Shimada, I. (1998) Determination of protonation and deprotonation forms and tautomeric states of histidine residues in large proteins using nitrogen-carbon J couplings in imidazole ring. *J Am Chem Soc 120*, 10988-10989.

[11] Wilbur, D. J., and Allerhand, A. (1977) Titration behavior and tautomeric states of individual histidine residues of myoglobins - application of natural abundance C-13 nuclear magnetic-resonance spectroscopy. *J Biol Chem 252*, 4968-4975.

[12] Munowitz, M., Bachovchin, W. W., Herzfeld, J., Dobson, C. M., and Griffin, R. G. (1982) Acid-base and tautomeric equilibria in the solid-state - N-15 NMR-spectroscopy of histidine and imidazole. *J Am Chem Soc 104*, 1192-1196.

[13] Blomberg, F., Maurer, W., and Ruterjans, H. (1977) Nuclear magnetic-resonance investigation of N-15-labeled histidine in aqueous-solution. *J Am Chem Soc 99*, 8149-8159.

[14] Chen, H., Viel, S., Ziarelli, F., and Peng, L. (2013) F-19 NMR: a valuable tool for studying biological events. *Chem Soc Rev 42*, 7971-7982.

[15] Yeh, H. J. C., Kirk, K. L., Cohen, L. A., and Cohen, J. S. (1975) F-19 and H-1 nuclear magnetic-resonance studies of ring-fluorinated imidazoles and histidines. *J Chem Soc-Perkin Trans 2*, 928-934.

[16] 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., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., *et al*. . (2009) Gaussian, Inc., Gaussian 09.

[17] Dennington, R., Keith, T. A., and Millam, J. M. (2009) Semichem Inc. Version 5.0. Shawnee Mission KS.

[18] Becke, A. D. (1993) A new mixing of Hartree-Fock and local density-functional theories. *J Chem Phys 98*, 1372-1377.

[19] Breneman, C. M., and Wiberg, K. B. (1990) Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J Comput Chem 11*, 361-373.

[20] Barone, V., and Cossi, M. (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J Phys Chem A 102*, 1995-2001.

[21] Ditchfield, R. (1974) Self consistent perturbation theory of diamagnetism .1. Gauge-invariant LCAO method for NMR chemical shifts. *Mol Phys 27*, 789-807.

[22] Wolinski, K., Hinton, J. F., and Pulay, P. (1990) Efficient implementation of the gauge-independent atomic orbital method for NMR chemical-shift calculations. *J Am Chem Soc 112*, 8251-8260.

[23] Yu, W. B., Liang, L., Lin, Z. J., Ling, S. L., Haranczyk, M., and Gutowski, M. (2009) Comparison of some representative density functional theory and wave function theory methods for the studies of amino acids. *J Comp Chem 30*, 589-600.

[24] Andra, K. K., Bullinger, J. C., Bann, J. G., and Eichhorn, D. M. (2010) 2-Fluoro-L-histidine. *Acta Crystallogr E 66*, O2713-U1695.

[25] Yeh, H. J. C., Kirk, K. L., Cohen, L. A., and Cohen, J. S. (1975) F-19 and H-1 nuclear magnetic-resonance studies of ring-fluorinated imidazoles and histidines. *J Chem Soc Perk T 2*, 928-934.

[26] Wimalasena, D. S., Janowiak, B. E., Lovell, S., Miyagi, M., Sun, J. J., Zhou, H. Y., Hajduch, J., Pooput, C., Kirk, K. L., Battaile, K. P., and Bann, J. G. (2010) Evidence that histidine protonation of receptor-bound anthrax protective antigen is a trigger for pore formation. *Biochemistry-US 49*, 6973-6983.

[27] Dedios, A. C., Pearson, J. G., and Oldfield, E. (1993) secondary and tertiary structural effects on protein NMR chemical shifts - an abinitio approach. *Science 260*, 1491-1496.

[28] Kasireddy, C., Ellis, J. M., Bann, J. G., and Mitchell-Koch, K. R. (2016) Tautomeric stabilities of 4-fluorohistidine shed new light on mechanistic experiments with labeled ribonuclease A. *Chem Phys Lett 666*, 58-61.

[29] Raines, R. T. (1998) Ribonuclease A. *Chem Rev 98*, 1045-1066.

[30] Quirk, D. J., and Raines, R. T. (1999) His ... Asp catalytic dyad of ribonuclease A: histidine pKa values in the wild-type, D121N, and D121A enzymes. *Biophys J 76*, 1571-1579.

[31] Harris, M. E., Piccirilli, J. A., and York, D. M. (2015) Enzyme transition states from theory and experiment. *Biochim Biophys Acta 1854*, 1727-1728.

[32] Watt, E. D., Shimada, H., Kovrigin, E. L., and Loria, J. P. (2007) The mechanism of rate-limiting motions in enzyme function. *Proc Natl Acad Sci U S A 104*, 11981-11986.

[33] Doucet, N., Watt, E. D., and Loria, J. P. (2009) The flexibility of a distant loop modulates active site motion and product release in ribonuclease A. *Biochemistry-US 48*, 7160-7168.

[34] Jackson, D. Y., Burnier, J., Quan, C., Stanley, M., Tom, J., and Wells, J. A. (1994) A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science 266*, 243-247.

[35] Kirk, K. L., and Cohen, L. A. (1973) Photochemistry of diazonium salts. I. Synthesis of 4-fluoroimidazoles, 4-fluorohistamine, and 4-fluorohistidine. *J Am Chem Soc 95*, 4619-4624.

[36] Kellerman, D. L., York, D. M., Piccirilli, J. A., and Harris, M. E. (2014) Altered (transition) states: mechanisms of solution and enzyme catalyzed RNA 2'-O-transphosphorylation. *Curr Opin Chem Biol 21*, 96-102.

[37] Lain, L., Lonnberg, H., and Lonnberg, T. (2015) Participation of an additional 4'-hydroxymethyl group in the cleavage and isomerization of ribonucleoside 3'-phosphodiesters. *Org Biomol Chem 13*, 4737-4742.

[38] Gu, H., Zhang, S., Wong, K. Y., Radak, B. K., Dissanayake, T., Kellerman, D. L., Dai, Q., Miyagi, M., Anderson, V. E., York, D. M., Piccirilli, J. A., and Harris, M. E. (2013) Experimental and computational analysis of the transition state for ribonuclease A-catalyzed RNA 2'-O-transphosphorylation. *Proc Natl Acad Sci U S A 110*, 13002-13007.

[39] Walters, D. E., and Allerhand, A. (1980) Tautomeric states of the histidine residues of bovine pancreatic ribonuclease A. Application of carbon 13 nuclear magnetic resonance spectroscopy. *J Biol Chem 255*, 6200-6204.

[40] Borah, B., Chen, C. W., Egan, W., Miller, M., Wlodawer, A., and Cohen, J. S. (1985) Nuclear magnetic resonance and neutron diffraction studies of the complex of ribonuclease A with uridine vanadate, a transition-state analogue. *Biochemistry-US 24*, 2058-2067.

[41] Santoro, J., Gonzalez, C., Bruix, M., Neira, J. L., Nieto, J. L., Herranz, J., and Rico, M. (1993) High-resolution three-dimensional structure of ribonuclease A in solution by nuclear magnetic resonance spectroscopy. *J Mol Biol 229*, 722-734.

[42] Schultz, L. W., Quirk, D. J., and Raines, R. T. (1998) His...Asp catalytic dyad of ribonuclease A: structure and function of the wild-type, D121N, and D121A enzymes. *Biochemistry-US 37*, 8886-8898.

[43] Gagne, D., French, R. L., Narayanan, C., Simonovic, M., Agarwal, P. K., and Doucet, N. (2015) Perturbation of the conformational dynamics of an active-site loop alters enzyme activity. *Structure 23*, 2256-2266.

[44] Tsirkone, V. G., Dossi, K., Drakou, C., Zographos, S. E., Kontou, M., and Leonidas, D. D. (2009) Inhibitor design for ribonuclease A: the binding of two 5'-phosphate uridine analogues. *Acta Crystallogr Sect F Struct Biol Cryst Commun 65*, 671-677.

[45] Boerema, D. J., Tereshko, V. A., and Kent, S. B. (2008) Total synthesis by modern chemical ligation methods and high resolution (1.1 A) X-ray structure of ribonuclease A. *Biopolymers 90*, 278-286.

[46] Larson, S. B., Day, J. S., Cudney, R., and McPherson, A. (2007) A new crystal form of bovine pancreatic RNase A in complex with 2'-deoxyguanosine-5'-monophosphate. *Acta Crystallogr Sect F Struct Biol Cryst Commun 63*, 728-733.

[47] Tanokura, M. (1983) H-1-NMR study on the tautomerism of the imidazole ring of histidine-residues .1. Microscopic Pk values and molar ratios of tautomers in histidine-containing peptides. *Biochim Biophys Acta 742*, 576-585.

[48] Khandogin, J., and Brooks, C. L. (2005) Constant pH molecular dynamics with proton tautomerism. *Biophys J 89*, 141-157.

[49] Ash, E. L., Sudmeier, J. L., Day, R. M., Vincent, M., Torchilin, E. V., Haddad, K. C., Bradshaw, E. M., Sanford, D. G., and Bachovchin, W. W. (2000) Unusual H-1 NMR chemical shifts support (His) C-epsilon 1-H .... O = C H-bond: Proposal for reaction driven ring flip mechanism in serine protease catalysis. *Proc Natl Acad Sci U S A 97*, 10371-10376.

[50] Schmiedekamp, A., and Nanda, V. (2009) Metal-activated histidine carbon donor hydrogen bonds contribute to metalloprotein folding and function. *J Inorg Biochem 103*, 1054-1060.

[51] Ross, M. R., White, A. M., Yu, F. T., King, J. T., Pecoraro, V. L., and Kubarych, K. J. (2015) Histidine orientation modulates the structure and dynamics of a de novo metalloenzyme active site. *J Am Chem Soc 137*, 10164-10176.

[52] Christen, D., Griffiths, J. H., and Sheridan, J. (1981) The microwave-spectrum of imidazole; complete structure and the electron distribution from nuclear quadrupole coupling tensors and dipole moment orientation. *Zeitschrift Fur naturforschung section a- J Phys Sci 36*, 1378-1385.

[53] Wimalasena, D. S., Cramer, J. C., Janowiak, B. E., Juris, S. J., Melnyk, R. A., Anderson, D. E., Kirk, K. L., Collier, R. J., and Bann, J. G. (2007) Effect of 2-fluorohistidine labeling of the anthrax protective antigen on stability, pore formation, and translocation. *Biochemistry-US 46*, 14928-14936.

[54] Jackson, D. Y., Burnier, J., Quan, C., Stanley, M., Tom, J., and Wells, J. A. (1994) A designed peptide ligase for total synthesis of ribonuclease-a with unnatural catalytic residues. *Science 266*, 243-247.

[55] Eichler, J. F., Cramer, J. C., Kirk, K. L., and Bann, J. G. (2005) Biosynthetic incorporation of fluorohistidine into proteins in E-coli: A new probe of macromolecular structure. *Chembiochem 6*, 2170-2173.

[56] Gerig, J. T. (1994) Fluorine NMR of proteins. *Prog Nuc Mag Reson Spectros 26*, 293-370.

[57] Hoeltzli, S. D., and Frieden, C. (1994) F-19 NMR-spectroscopy of [6-F-19]tryptophan-labeled Escherichia-coli dihydrofolate reductase equilibrium folding and ligand binding studies. *Biochemistry-US 33*, 5502-5509.

[58] Urick, A. K., Calle, L. P., Espinosa, J. F., Hu, H., and Pomerantz, W. C. (2016) Protein-observed fluorine NMR Is a complementary ligand discovery method to 1H CPMG ligand-observed NMR. *ACS Chem Biol 11*, 3154-3164.

CHAPTER 3

Demystifying fluorine chemical shifts: Electronic structure calculations address origins of seemingly anomalous ^{19}F-NMR spectra of fluorohistidine isomers and analogues

Note: Most of the contents in this chapter are published in Physical Chemistry Chemical Physics, 2015, 17, 30606 and Annual Reports on NMR Spectroscopy, 2018, 93, 281.

3.1. INTRODUCTION

Compared to ^{1}H and ^{13}C NMR spectroscopy, fluorine NMR gives a wide-range of shifts corresponding to even small changes in local environment.[1-3] While we have a robust framework for understanding and predicting NMR spectra for ^{1}H and ^{13}C nuclei, the origins of ^{19}F chemical shifts are not as well understood. Despite the sensitivity of ^{19}F-NMR and the breadth of literature on proteins labelled with ^{19}F-amino acids, there is still no unifying theoretical basis for predicting fluorine chemical shifts in proteins. When scientists wish to determine which fluorine chemical shift in a protein comes from a certain residue, it is usually necessary to make mutants at every fluorinated site, one-by-one eliminating fluorine NMR signals.[4-6] Such work is extremely time-intensive. This report is a step toward developing a quantitative predictive framework [7, 8] for ^{19}F chemical shifts, which will enable greater chemical insight. Such an advance would be a breakthrough, streamlining experiments and realizing the full potential of ^{19}F NMR spectroscopy.

Experimental work over four decades ago used substituent effects to build an understanding of fluorine electronic structure through ^{19}F NMR spectroscopy. Studies of the fluorine chemical shifts of substituted aromatic systems led scientists to postulate fluorine hyperconjugation,[9] and donation of fluorine *p* electrons to the π aromatic system (*p-π donation*).[10] Later experiments extended to substituted aliphatic systems, investigating polar and resonance effects on aliphatic fluorine chemical shifts.[1] This present work examines effects of fluorine electron delocalization on chemical shifts, through analysis of molecular orbitals calculated with DFT.

Previous theoretical efforts have also advanced our understanding of fluorine chemical shifts. This includes the assignment of fluorine signals from 5-fluorotryptophan residues in the solid-state NMR spectrum of the membrane-bound ion channel peptide gramicidin A by Sternberg *et al.*,[11] using *ab initio* calculations and semi-empirical bond polarization parameters for chemical shift calculations. Coupling calculated ^{19}F chemical shifts with Molecular Dynamics (MD) simulations, the authors were able to describe multiple conformations of tryptophan side chains in gramicidin A. Analysis by Lau and Gerig of the fluorine chemical shifts of dihydrofolate reductase (DHFR) labelled with 6-fluorotryptophan suggested that differences in fluorine chemical shifts are a sum of the following factors: hydrogen bonding, short-range interactions, electric fields and local magnetic anisotropies.[3] Additionally, Dalvit and Vulpetti classified different fluorine-containing functional groups based on fluorine electron density and their interactions with proteins.[12] Their experimental and theoretical results

showed that the most-shielded fluorine atoms are most likely to form interactions with hydrogen bond donors of a protein. On the other hand, more deshielded fluorine atoms interact with hydrophobic side chains and carbonyl carbons. These efforts in the past decade are significant steps toward understanding ^{19}F-NMR spectra. Still, a framework for the *a priori* assignment and prediction of fluorine chemical shifts remains to be developed.

Toward providing a better theoretical basis for predicting ^{19}F-chemical shifts, we have sought to explain the physical origins for a long-standing mystery regarding the ^{19}F chemical shifts of fluorohistidine isomers upon acid titration. Yeh *et al.* studied ^{19}F and ^{1}H NMR spectra for 2-fluoro- and 4-fluoro- histidine, -imidazole, and –(5-methyl)-imidazole in aqueous, basic, and acidic solutions.[13] As expected, Yeh *et al.* found that the ^{1}H chemical shifts for all fluoroisomers are deshielded upon protonation of the imidazole ring.[13] Likewise, 2-fluorohistidine and 2-fluoroimidazoles exhibited a downfield (higher frequency) shift in the ^{19}F signal upon acid titration. However, 4-flurohistidine and 4-fluroimidazoles exhibit an *upfield* (lower frequency) shift in the ^{19}F signal at low pH.

We have identified electronic structure methods that can replicate the experimental ^{19}F chemical shifts of 2- and 4-fluoro-histidine and –(5-methyl)-imidazole upon acid titration. Analysis provides a plausible explanation for the anomalous chemical shift changes observed. This system is a key place to start work on a rational framework for ^{19}F chemical shifts. Since the NMR spectra of 2 and 4-fluorohistidines are nearly identical to their imidazole analogues, the discussion here focuses on 2- & 4- (5-methyl)-imidazole, given that they are small

and easily amenable to computation. Understanding the physical origins of these specific chemical shift changes may provide clues that, in turn, will lead to a better understanding of ^{19}F chemical shift differences in proteins.

3.2. METHODS

All the input files were prepared in *GaussView 5*,[14] modifying the geometry provided in the 2F-histidine crystal structure[15] for desired protonation states and C5 substitution. All the electronic calculations were performed using *Gaussian 09* software.[16] Optimized geometries, energies, ^{19}F NMR shifts, electrostatic charge distributions, and Natural Chemical Shielding (NCS)[17] calculations were performed with multiple methods to ensure general conclusions. Calculations were performed with the BHandHLYP and B3LYP density functionals and the MP2 method, [18, 19] using 6-311++G(3df,2p), aug-cc-pVTZ, and 6-31+G* basis sets and water solvation with SMD[20] and CPCM[21] solvent models. Negative frequencies were not observed for any molecule, indicating that the geometries are at energy minima. All reported values are from calculations using BHandHLYP[18] hybrid density functional, 6-311++G(3df,2p) basis and SMD water solvation. Values from other methods are given in Table 3.1. Reported relative energies are free energies given from frequency calculations, setting the energy of the most stable tautomer to zero. NMR values are calculated with Gauge Independent Atomic Orbital (GIAO)[22] method, and isotropic chemical shifts are reported. Molecular orbitals are visualized using *Avogadro* software.[23] The standard reference for the ^{19}F NMR values reported in Table 2 is C_6F_6, with shifts reported

as σ_{ref} - σ_{calc}. The value of σ_{ref} comes from an optimized geometry of hexafluorobenzene with BHandHLYP/6-311++G(3df,2p) method, and SMD model water solvation.

3.3. RESULTS

3.3.1. Tautomer stabilities

At neutral pH, the imidazole ring exists in two tautomeric forms: τ (protonation at N3) or π (protonation at N1) (Fig. 3.1), depending on which nitrogen is protonated.[24] The tautomeric states of fluorohistidine/imidazoles have not been definitively determined by experiment, although the crystal structure suggests 2F-histidine is protonated at N3 (the τ-tautomer – Fig. 3.1a).[15] Thus, we carried out calculations on both tautomers of each fluoroisomer. The π-tautomer of 4F-histidine/(5-methyl)-imidazole was found to be more stable than the τ-tautomer (by ~15.0/25.8 kJ/mol using BHandHLPY/6-311++g(3df,2p) method). The two tautomers of 2F-(5-methyl)-imidazole were found to be nearly isoenergetic. However, we found that τ-tautomer of 2F-histidine is more stable than π-tautomer (by 5.3 kJ/mol), as it is in canonical histidine.[24] Our discussion from this point focuses on 2F- and 4F-(5-methyl)-imidazoles, since the simplified systems (relative to zwitterionic fluorohistidines) allow for more straightforward analysis, but shed light on experimental results for fluorohistidines.

Table 3.1. ^{19}F NMR Δppm values for acid titration of 2F-(5-methyl)-imidazole and 4F-(5-methyl)-imidazole, with different methods and solvation models and both π and τ tautomers. Here $\Delta_a = 2F(H^+) - 2F(\tau)$, $\Delta_b = 2F(H^+) - 2F(\pi)$, $\Delta_c = 4F(H^+) - 4F(\tau)$ and $\Delta_d = 4F(H^+) - 4F(\pi)$.

Method	Solvation model	$\Delta_a$ppm	$\Delta_b$ppm	$\Delta_c$ppm	$\Delta_d$ppm
MP2/6-31++g(d)	CPCM	3.82	4.08	9.22	-2.87
BHandHLYP/6-31+g(d)	CPCM	10.28	11.23	8.93	-2.22
BHandHLYP/aug-CC-pVTZ	CPCM	10.63	11.62	9.60	-2.18
BHandHLYP/6-31+g(d)	SMD	7.46	8.49	6.40	-4.15
BHandHLYP/aug-CC-pVTZ	SMD	7.77	8.77	6.94	-4.18
B3LYP/6-31+g(d)	SMD	7.52	8.22	8.21	-2.91
B3LYP/aug-CC-pVTZ	SMD	3.82	8.22	8.92	-2.67

The results reported herein are for chemical shifts of the putatively more stable tautomers of each fluoro-isomer: τ-tautomer (N3-H) for 2F-(5-methyl)-imidazole (Fig. 3.1a) and π-tautomer (N1-H) for 4F-(5-methyl)-imidazole, shown in Fig. 3.1c. Reassuringly, we found that the direction of the ^{19}F chemical shifts upon titration did not depend on the computational method used. The BHandHLYP/6-311++G(3df,2p) calculations with SMD water solvation gave the best quantitative agreement with experimental Δppm values, where Δppm is the difference between the chemical shifts of the fluoroimidazolium and the fluoroimidazole ($\Delta ppm = \delta_{Im+} - \delta_{Im}$, given as Δ_1 in Ref. 13).

Table 3.2. ^{19}F NMR shifts (experimental and calculated) and fluorine electrostatic potential (charge) for stable tautomers of 2-fluoro- and 4-fluoro-imidazole (2F-Im, 4F-Im) analogues 2- and 4-fluorohistidine (2F-His, 4F-His), 2- and 4-fluoro-(5-methyl)-imidazole (2F-MeIm, 4F-MeIm).

2F isomers, τ-tautomers	**2F-His**	**2F-MeIm**	**2F-Im**
ESP charge	-0.264	-0.278	-0.280
^{19}F NMR shifts (calculated)	54.96	53.62	52.50
^{19}F NMR shifts (experimental)	59.80	58.59	56.60
4F isomers, π-tautomers	**4F-His**	**4F-MeIm**	**4F-Im**
ESP charge	-0.290	-0.275	-0.308
^{19}F NMR shifts (calculated)	16.54	13.83	21.97
^{19}F NMR shifts (experimental)	20.25	16.91	23.43

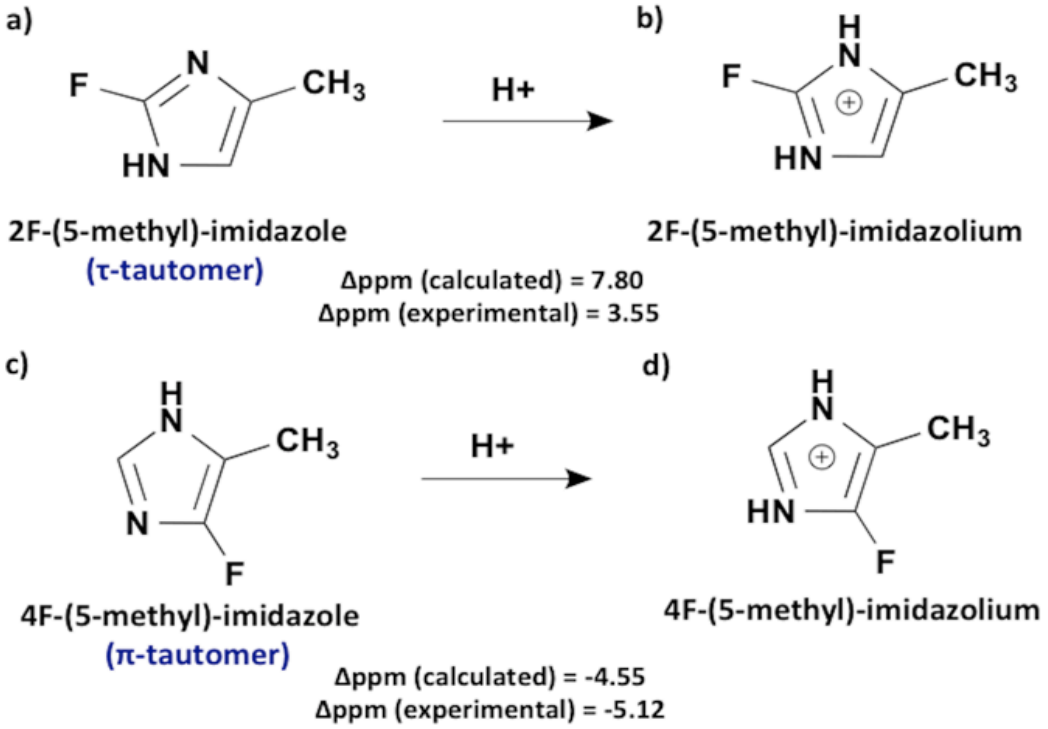
3.3.2. Partial charge distribution

Previous experimental work on the protonation of fluoropyridines, with results akin to 4F-(5-methyl)-imidazole, led researchers to postulate that anomalous values of Δppm arise from magnetic anisotropy.[25] Although we cannot account for the contribution of magnetic anisotropy in the method used here, the isotropic values in the chemical shielding tensor are able to reproduce the magnitude and direction of the fluorine chemical shifts of fluoroimidazoles/histidines upon

protonation.

For clues to the differences between 4F- vs 2F- ^{19}F chemical shifts, we first looked at the charges on the fluorine atom to get a simple picture of the electron

density available to shield the nucleus. In both the 2-fluoro and 4-fluoro cases, the charge on fluorine decreased upon protonation, as would be expected for systems gaining positive charge. For 2F-(5-methyl)-imidazole, the fluorine electrostatic charges calculated by the ChelpG procedure[26] were -0.278 for neutral and -0.153 for protonated form. For 4F-(5-methyl)-imidazole, the charges on fluorine were -0.275 (neutral) and -0.164 (protonated). Using straightforward shielding arguments by electron density, the deshielded ^{19}F chemical shifts for 2-fluorohistidine/imidazoles make sense, corresponding to reduced charge density. Meanwhile, the increase in shielding observed for 4F- imidazoliums is puzzling,



given the decrease in charge density that would seem to indicate less shielding.

Figure 3.1. Chemical structures and ^{19}F chemical shifts (Δppm= δIm+ - δIm) for (a) neutral N3-H/τ-tautomer and (b) protonated 2F-(5-methyl)- imidazole; (c) neutral N1-H/π-tautomer and (d) protonated 4F-(5-methyl)-imidazole.

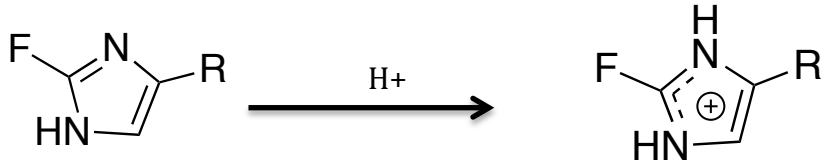
3.3.3. Natural Chemical Shielding Analysis

Next, we extended our analysis to Natural Chemical Shielding (NCS) analysis,[35] from which the contribution of diamagnetic and paramagnetic components towards total nuclear shielding can be computed. Natural chemical shielding analysis also provides a breakdown of contributions from each molecular orbital to the overall isotropic and anisotropic (off-diagonal) chemical shift tensors. This was used to assess changes in shielding/deshielding from each molecular orbital from protonation of the fluoroimidazole rings. Through natural chemical shielding analysis, the contribution of each Lewis and non-Lewis bonds and lone-pair orbital towards chemical shielding can be calculated. The diamagnetic shielding arises from localized electrons involved close to the respective atom, representing the ground state electron density. However, paramagnetic shielding arises from interactions with higher excited states and electron perturbations from ground state, due to the presence of the magnetic field. Most differences in shielding/deshielding of heavy nuclei arise from paramagnetic shielding. For 2-fluoro-(5-methylimidazole), the NMR calculations show that both τ and π tautomers exhibit downfield shifts of +3.55 and +8.8 ppm observed respectively upon acid titration. This is in line with the experimental change in chemical shift of + 7.8 ppm downfield. NCS analysis of the more stable τ-tautomer is presented in Table 3.3, and it shows that diamagnetic shielding of τ-tautomer (453.05 ppm) is higher than that of the protonated moiety (445.9 ppm). Again, this correlates with the ground state electronic structure and calculated charge on fluorine, which becomes less negative when the imidazole ring is protonated. When considering

paramagnetic deshielding, the protonated form is calculated to have slightly higher paramagnetic deshielding (-138.8 ppm), compared to τ-tautomer (-138.1 ppm). Hence, both diamagnetic and paramagnetic shielding parameters appear to be the reason for the observed downfield shift in ^{19}F NMR upon protonation of 2-fluorohistidine.

Table 3.3. Calculated ^{19}F shieldings for 2-fluoro-(5-methyl)-imidazole τ-tautomer and protonated form. (Note: Absolute (unreferenced) shieldings are given here.

2F-(5-methyl)-imidazole	Paramagnetic shielding (ppm)	Diamagnetic shielding (ppm)
τ-tautomer	-138.06	453.05
Protonated form	-138.78	445.92
Net change	0.72	7.56



2F-(5-methyl)-imidazole (τ-tautomer) **2F-(5-methyl)-imidazolium**

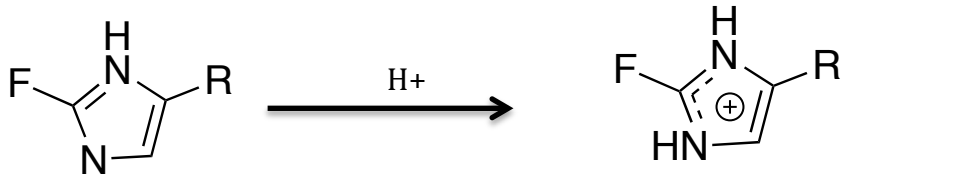
Experimental value: 7.80 ppm (deshielding)[13]
Net overall change (calculated): 3.55 ppm

In the case of the π-tautomer of 2-fluoro-(5-methyl)-imidazole, the same analysis holds. Data presented in Table 4 show that the calculated diamagnetic shielding is higher in the neutral species (454.3 ppm) compared to protonated form (445.92 ppm). Paramagnetic deshielding is marginally higher in the protonated form (-138.8 ppm), compared to the neutral structure (-138.4 ppm). Both of these

factors appear to contribute towards the downfield shift upon ring protonation of 2-fluoro-(5-methyl)-imidazole.

Table 3.4. Calculated ^{19}F shieldings for 2-fluoro-(5-methyl)-imidazole π-tautomer and protonated form.

2F-(5-methyl)-imidazole	Paramagnetic shielding (ppm)	Diamagnetic shielding (ppm)
π-tautomer	-138.35	454.32
protonated form	-138.78	445.92
Net change	0.40	8.40



2F-(5-methyl)-imidazole (π-tautomer) **2F-(5-methyl)-imidazolium**

Experimental value: NA
Net overall change (calculated): 8.8 ppm

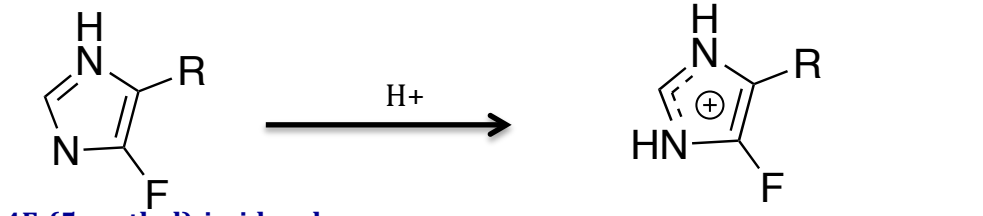
Similar analysis is performed on 4-fluoro-(5-methyl)-imidazole, considering both tautomers (τ and π) and protonated form. Here, the results are a little different compared to 2-fluoromethylimidazole. Among the two tautomers, the most stable π-tautomer exhibits an upfield shift of -5.12 ppm in ^{19}F NMR upon ring protonation, which is considered unusual chemical shift behavior. However, the less stable π-tautomer behaves normally with a downfield shift of 6.9 ppm at lower pH.

The shielding components of the ^{19}F nucleus in the π-tautomer of 4-fluoro-(5-methyl)-imidazole are presented in Table 3.5. The diamagnetic parameters show

that there is higher shielding in the neutral structure (460.5 ppm) respective to the protonated structure (454.0 ppm). The paramagnetic shielding is significantly reduced in protonated structure (-93.5 ppm) relative to the neutral π-tautomer (106.0 ppm). These results suggest illustrates that a significant decrease in paramagnetic deshielding is the primary cause for the upfield shift of ^{19}F that is experimentally observed upon protonation of 4-fluoro-(5-methyl)-imidazole.[25]

Table 3.5. Calculated ^{19}F shieldings for 4-fluoro-(5-methyl)-imidazole π-tautomer and protonated form.

4F-(5-methyl)-imidazole	Paramagnetic shielding (ppm)	Diamagnetic shielding (ppm)
π-tautomer	-105.95	460.54
Protonated form	-93.48	453.98
Net change	12.4	6.56



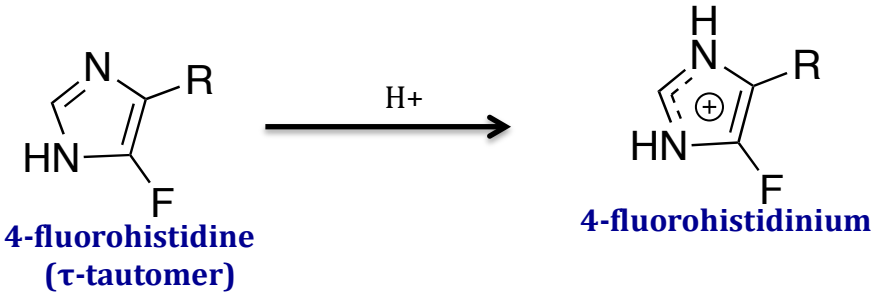
4F-(5-methyl)-imidazole (π-tautomer) → **4F-(5-methyl)-imidazolium**

Experimental value: -4.55 ppm (shielding)[13]
Net overall change (calculated): -5.12 ppm

In the case of the less stable τ-tautomer, calculations indicate that the downfield shift upon ring protonation arises from reduced diamagnetic shielding (460.2 ppm in neutral vs 453.9 ppm in protonated structure) and higher paramagnetic deshielding (-92.6 ppm vs -93.5 ppm), as shown in Table 3.6.

Table 3.6. Calculated ^{19}F shieldings for 4-fluoro-(5-methyl)-imidazole τ-tautomer and protonated form.

4F-(5-methyl)-imidazole	Paramagnetic shielding (ppm)	Diamagnetic shielding (ppm)
τ-tautomer	-92.61	460.27
Protonated form	-93.48	453.98
Net change	0.86	6.29



4-fluorohistidine (τ-tautomer) → **4-fluorohistidinium**

Experimental value: N/A
Net overall change (calculated): 6.99 ppm

Further, we extended our analysis to molecular orbital visualization. Out of 26 total molecular orbitals (MOs), 18 contribute to the ^{19}F isotropic chemical shifts of 2F-(5-methyl)-imidazole and 4F-(5-methyl)-imidazole. In mapping the corresponding molecular orbitals of 2F-(5-methyl)-imidazole and 2F-(5-methyl)-imidazolium to determine NCS contributions, we found that the molecular orbitals of the neutral and protonated species were very similar. There are differences in contributions to the fluorine chemical shift from every MO, but many of them appear to cancel each other out. One would expect that electrons highly localized around fluorine would have the greatest contribution to shielding of the ^{19}F

nucleus. The calculated molecular orbital 2 (MO 2, the second-lowest in energy), visualized in Fig. 3.2a and 3.2b using Avogadro software, corresponds to a lone pair orbital on fluorine. Differences between the Natural Chemical Shielding of this MO in 2F-(5-methyl)-imidazole and 2F-(5-methyl)-imidazolium results in a deshielded shift (4.3 ppm downfield/higher frequency). The overall difference in chemical shift was calculated to be 7.8 ppm (experimentally[13] it is 3.55 ppm), indicating that more than half of the change in the calculated chemical shift of 2F-(5-methyl)-imidazole upon acid titration can be attributed to changes in electron density in the fluorine lone pair molecular orbital. This corroborates with the calculated reduction in electrostatic charge on fluorine upon protonation, which leads to less diamagnetic shielding of the fluorine nucleus.

The shift in electron density away from the fluorine nucleus in 2F-(5-methyl)-imidazole can also be seen in changes in the dipole moment. The neutral species has a dipole moment of μ=5.2 Debye, whereas 2F-(5-methyl)-imidazolium shows a drastic change to 0.77 Debye, showing a reduction in the polarization toward the fluorine atom in the protonated species. However, the same pattern does not hold for 4F-(5-methyl)-imidazole. Neutral 4F-(5-methyl)-imidazole has a dipole moment of 7.1 Debye, and the dipole moment of 4F-(5-methyl)-imidazolium is only slightly reduced to 6.2 Debye. This may be seen as an indication that the electronic structure of the 4F-species is intrinsically different from the 2F-species. The authors of the pioneering experimental work on fluorinated imidazoles and histidines[31] suggested that differences in pK_a measured

for 2- vs 4- fluoro-imidazole rings may be due to electronic structure differences that arise in 4-fluoro species, beyond σ-inductive effects.

In analyzing Natural Chemical Shielding (NCS) data for 4F-(5-methyl)-imidazolium, it was found that there are drastic changes in the molecular orbitals of 4F-(5-methyl)-imidazole compared to 4F-(5-methyl)-imidazolium. So much so, that comparison of NCS data for the neutral and protonated species of 4F-(5-methyl)-imidazole is difficult. The 2nd-lowest energy molecular orbital (MO 2) of 4F-(5-methyl)-imidazolium is a lone pair orbital (delocalized somewhat toward the methyl group, see Fig. 2d) that contributes to shielding of the ^{19}F chemical shift. What is striking is that we observe *no lone pair electrons* on the fluorine for neutral 4F-(5-methyl)-imidazole. Fig. 2c shows the lowest energy molecular orbital (MO 5) that has any similarity to a lone pair orbital. As can be seen in the molecular orbital in Fig. 3.2c, while there is electron probability around the fluorine nucleus, the other lobe of the orbital encompasses the protonated nitrogen (N1) on the aromatic imidazole ring. The NCS data indicates that, relative to the shielding of MO 5 in the neutral species, MO 2 (lone pair) of 4F-(5-methyl)-imidazolium contributes more shielding (6.45 ppm) to the fluorine nucleus. Clearly, multiple orbitals contribute to chemical shielding, but these are the lowest-energy orbitals that have considerable electron density near the fluorine nucleus. The differences in chemical shielding provided by MO 2 (lone pair) of 4F-imidazolium and MO 5 (delocalized fluorine electrons) of neutral 4F-(5-methyl)-imidazole account for most of the change in chemical shift upon titration. Thus,

protonation appears to trigger fluorine electron localization into a lone pair in 4F-imidazolium species, giving rise to higher shielding at low pH.

3.3.4. When fluorine electron density is delocalized, do fluorine chemical shifts correlate with charge?

To address the question of whether 4-fluoroimidazoles have "predictable" chemical shifts when they undergo minor perturbations in electron density, we evaluated and compared the calculated charges and chemical shifts for 4F-(5-methyl)-imidazole (4F-MeIm), 4F-histidine (4F-His), 4F-imidazole (4F-Im). For 4-fluoro-substituted species, electron density (charge) is lowest for 4F-MeIm (-0.275), slightly increases for 4F-His, and is highest for 4F-Im (-0.31). Table 2 summarizes the fluorine electrostatic potential (charge) alongside experimental and calculated fluorine chemical shifts. As can be seen, the ^{19}F NMR shifts did not follow a predictable trend, with deshielded shifts corresponding to lower electron density. In fact, the chemical shifts are reversed: the most shielded ^{19}F chemical shift, for 4F-MeIm, has lowest charge density, while the most deshielded, 4F-Im, has highest charge density. However, the fluorine chemical shifts (though not the charges) do fit chemical intuition. That is to say, that the most shielded fluorine chemical shift corresponds to the imidazole with the most electron donating substituent at C5. Note that the computationally-calculated NMR chemical shifts for the 4F-imidazole series match the experimental ordering of ^{19}F chemical shifts. Thus, without the calculated fluorine charge density, the experimental results don't "raise any eyebrows".

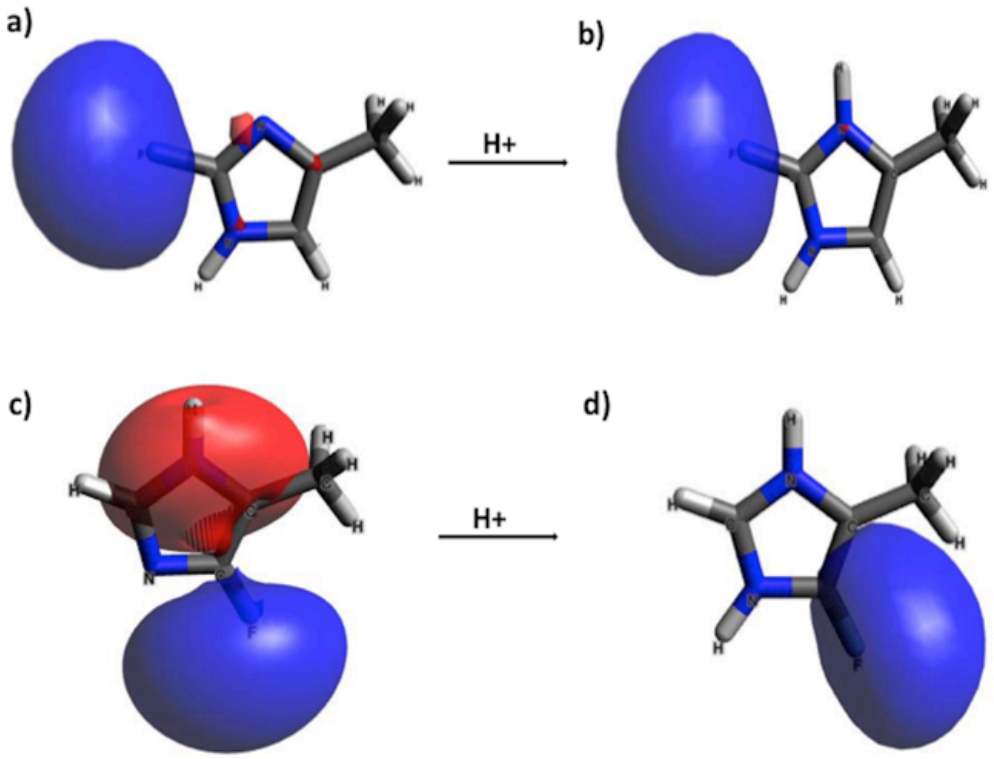


Figure 3.2. Lowest energy molecular orbitals containing fluorine electron density (MO 2) for **a)** 2F-(5-methyl)-imidazole **b)** 2F-(5-methyl)-imidazolium; **c)** lowest energy lone-pair like orbital (MO 5) for 4F-(5-methyl)-imidazole **d)** 4F-(5-methyl)-imidazolium (MO 2).

The peculiar behavior of 4F-imidazoles, in which chemical shift and electron density have reverse relationships relative to most NMR chemical shifts, is akin to that observed in aliphatic fluoride systems.[4] Adcock and Abeywickrema performed in-depth studies of relationships between ^{19}F substituent chemical shifts and fluorine electron density.[1] They studied substituent effects on a fluorinated bicyclo-octane, and found that most aliphatic fluoride chemical shifts become deshielded with increasing electron density. This is in contrast to studies on phenyl fluorides, which showed through substituent effects that fluorine chemical shifts

are deshielded as electron density decreases ("normal" behavior). So, Adcock and Abeywickrema concluded that aliphatic fluorides may have "reverse" chemical shift effects, while aromatic fluorides behave normally. They postulated that the "reverse" behavior of ^{19}F chemical shifts reflects the polarization of C-F σ bonds. To probe this in our data set, we looked at the charge separation between the carbon and fluorine electrostatic potential (ESP/charge). Indeed, the charge separation (polarity) of the C-F bond was greatest for the most deshielded shift, and least for the most shielded chemical shift. So, while charge of fluorine itself does not correlate with chemical shift in 4F-imidazole, which has all fluorine electron density delocalized, the extent of polarization does correlate in a reverse manner (polar bond = deshielding, less polar bond= shielding). This suggests that perhaps the reverse sigma effects seen for aliphatic systems also occur in aromatic systems in which there is full fluorine electron delocalization.

In contrast, the 2-fluoroimidazoles behaved normally, where more shielded ^{19}F shifts can be correlated with increased electron density. The electron density on fluorine is observed to increase from 2F-His to 2F-MeIm to 2F-Im. In accordance with electron density, shielding of the ^{19}F chemical shift is observed for the molecules in same order. However, there is no correlation between C-F charge separation and ^{19}F chemical shift for 2-fluoroimidazole and analogues.

3.4. DISCUSSION

Considering only 2F-(5-methyl)-imidazole, the fluorine chemical shifts fit "chemical intuition", with reduced fluorine electron density upon protonation leading to deshielding of the fluorine nucleus. Data from electronic structure

calculations supports chemical intuition as well, with less negative fluorine electrostatic charges and reduced dipole moment coinciding with deshielding of the fluorine nucleus. In contrast, 4F-(5-methyl)-imidazole does not show the same correlations among indicators of electron density and fluorine. (Fig. 3.3)

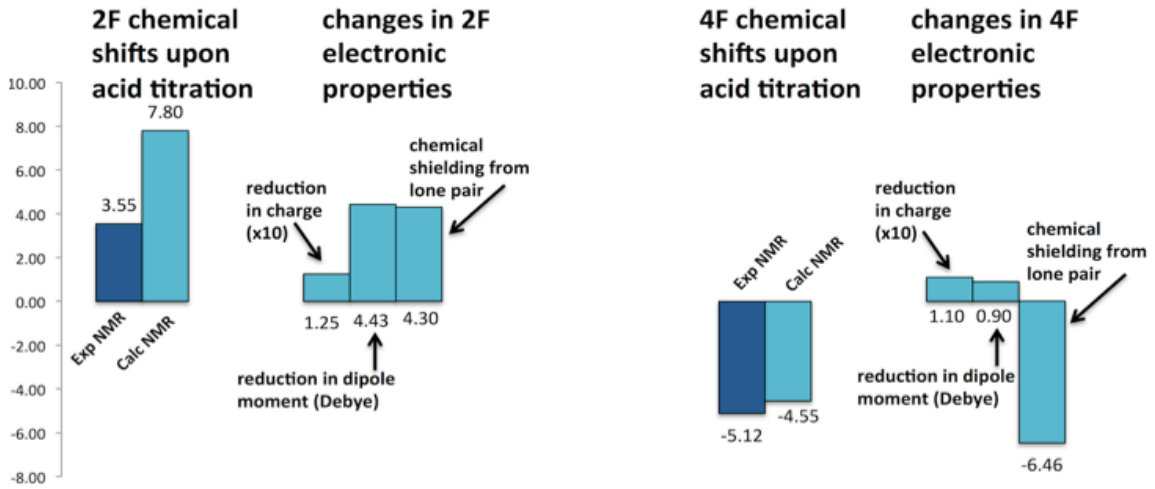


Figure 3.3. Summary of correlation between fluorine chemical shifts and electronic properties. Experimental and calculated ^{19}F chemical shifts are given, alongside reduction in charge and dipole moment upon acid titration, and natural chemical shielding (NCS) contributions from the lone pairs of 2F-(5-methyl)-imidazolium & 2F-(5-methyl)-imidazole, and lone pair (MO 2) of 4F-(5-methyl)-imidazolium & MO 5 of 4F-(5-methyl)-imidazole.

The electronic structure of the more stable π-tautomer of 4-fluoroimidazole is unique compared to the other fluorinated imidazoles we considered. In all the molecular orbitals of 4F-(5-methyl)-imidazole that contain fluorine electron density, there is conjugation within the ring and/or overlap with the adjacent methyl group. See Fig. 3.4 for visualization of all molecular orbitals of 2F- and 4F-(5-methyl)-imidazole and –imidazolium. The electronic structure of 4F-(5-methyl)-imidazole appears to have effects along the lines of the fluorine p-π

interaction proposed by Sheppard in 1965.[10] Using this information, we can conclude that protonation of 4F-(5-methyl)-imidazole changes the electronic structure such that the lone pair electrons on fluorine are restored (localized), providing greater shielding to the fluorine nucleus.

We postulated the anomalous electronic structure of the 4F-imidazole moiety may arise from the electron-donating nature of the methyl (methylene) group adjacent to the fluorine in 4F-(5-methyl)-imidazole (4F-histidine), giving an effect akin to hyperconjugation. To test this, we performed calculations that replaced the methyl group with an electron-withdrawing trifluoromethyl group. Upon protonation, this species still exhibited higher shielding (by 5.09 ppm calculated by BHandHLYP/6-311++G(3d,2p)), so we can conclude that the electronic structure does not appear to be dependent on the electron-donating or –withdrawing character of the adjacent substituent on the imidazole ring.

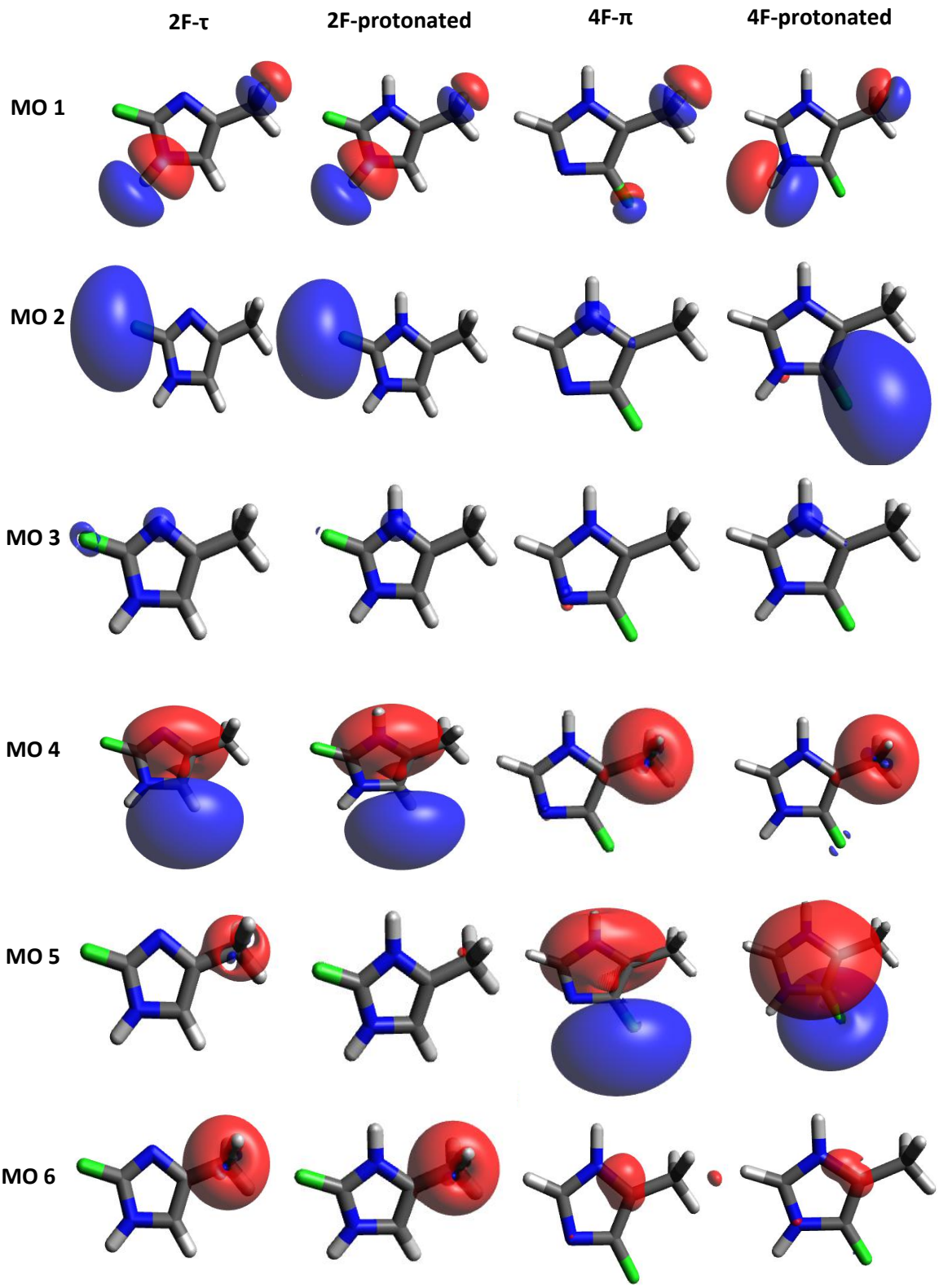


Figure 3.4. 2F- and 4F-(5-methyl)-imidazole and –imidazolium molecular orbitals (MOs) with BHandHLYP/6-311++G(3df,2p) method.

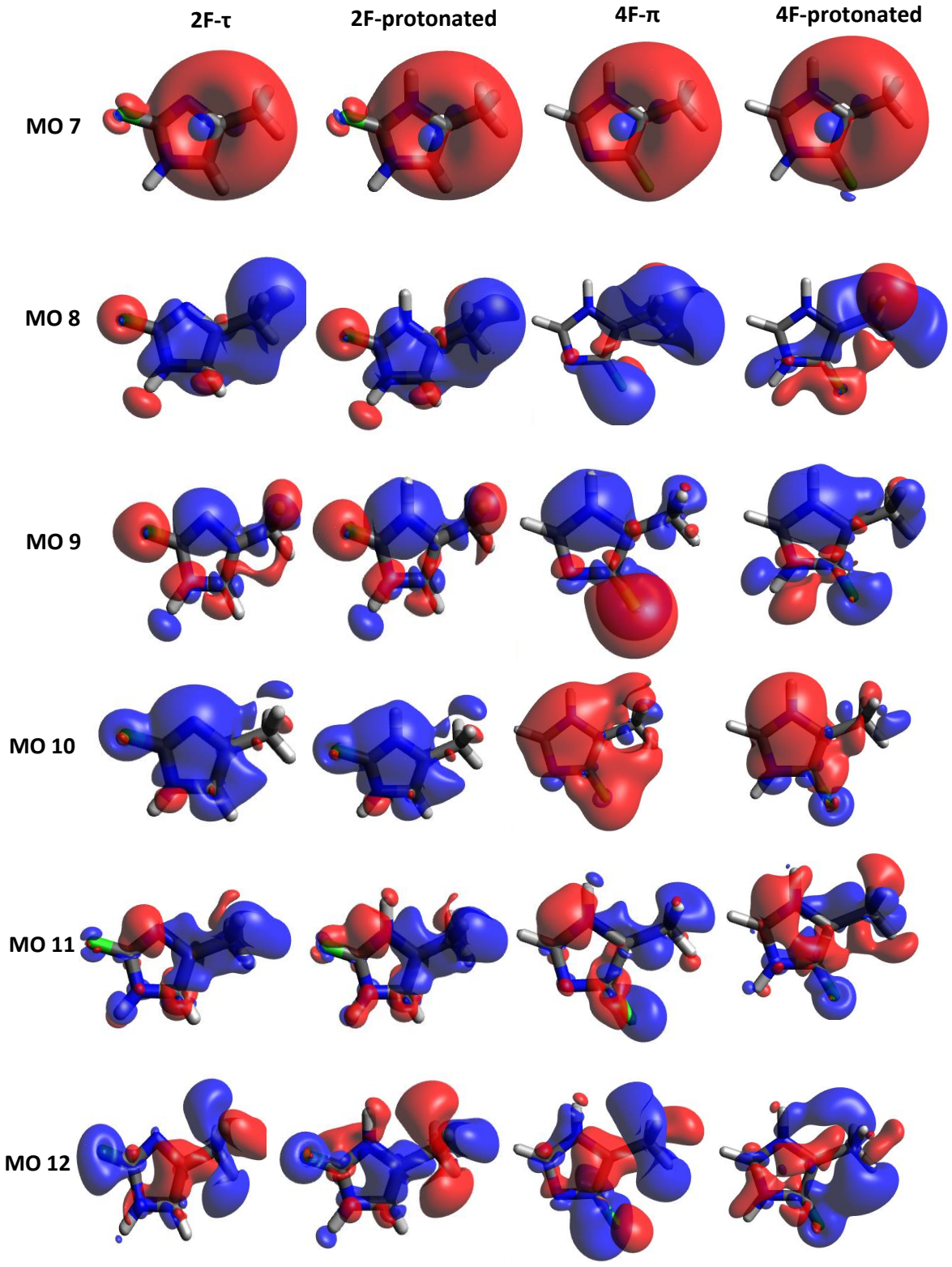


Figure 3.4. (continued)

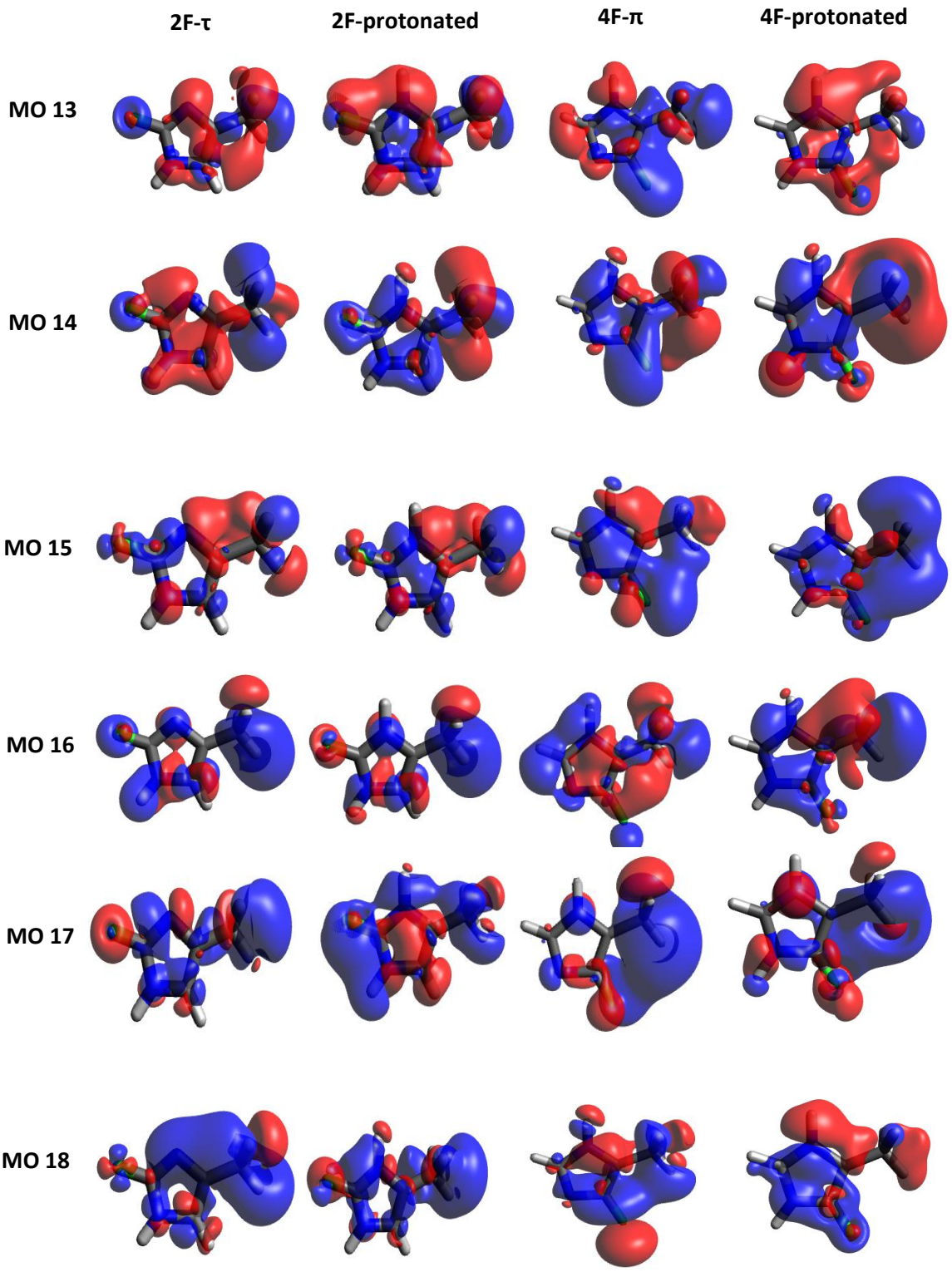


Figure 3.4. (continued)

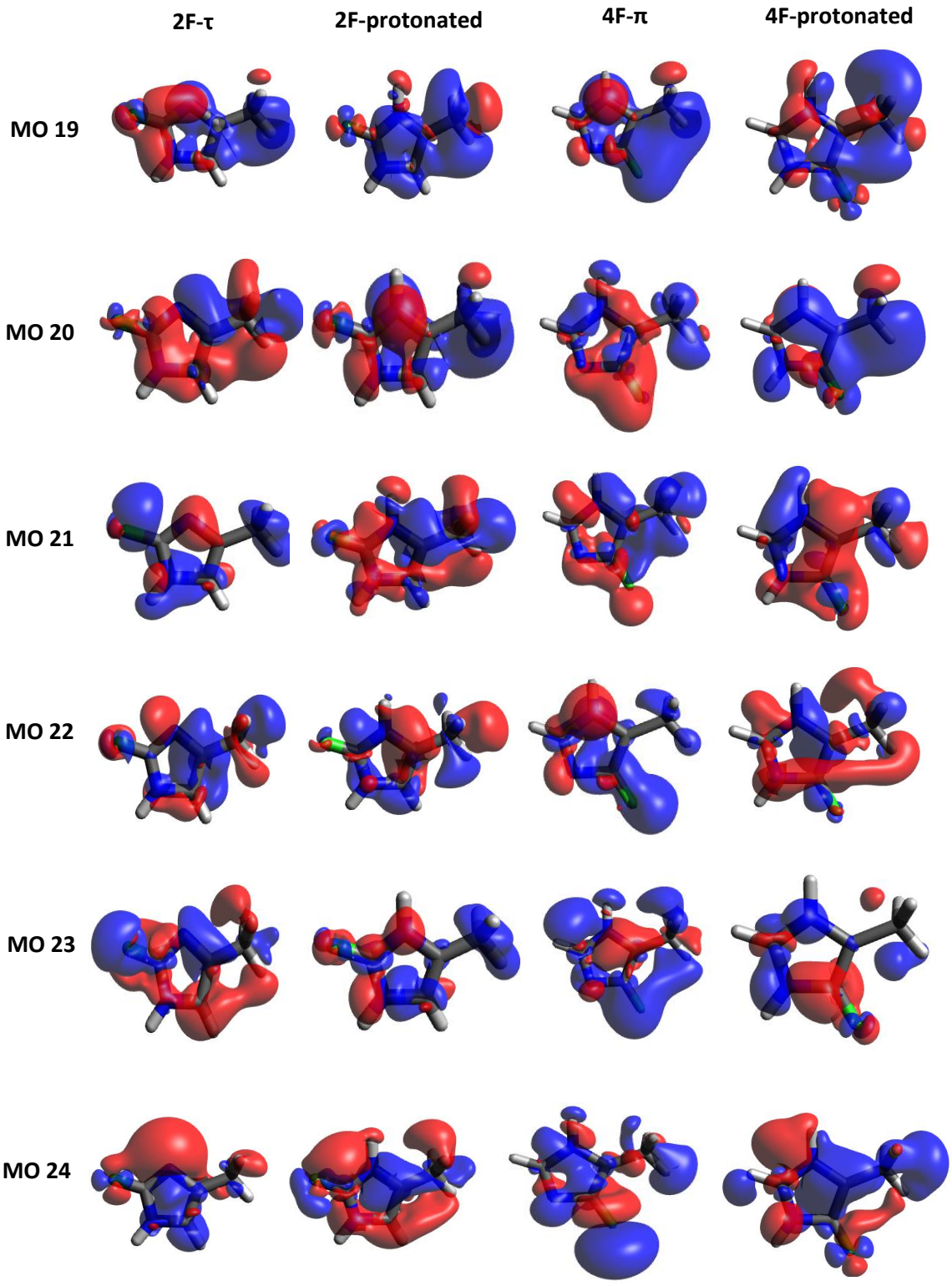


Figure 3.4. (continued)

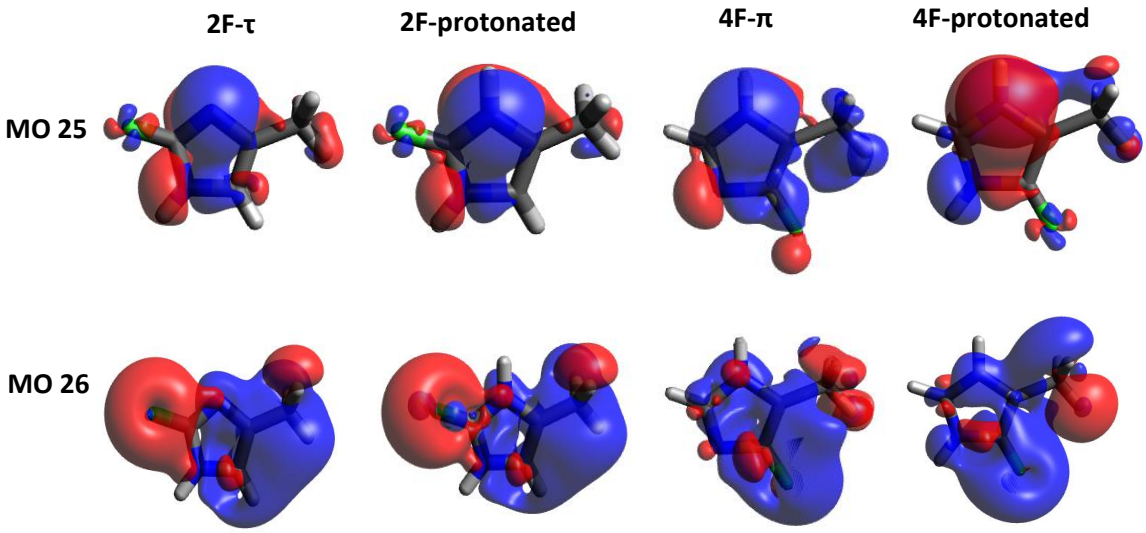


Figure 3.4. (continued)

The reverse behavior of ^{19}F NMR chemical shifts in substituted 4-fluoroimidazoles clearly relates these species with the behavior of aliphatic fluorides. Studies of the 4F-His/Im/MeIm series also indicated an inverse relationship among electron density and shielding of chemical shifts. Calculated charges of fluorine and C4 support Adcock and Abeywickrema's hypothesis that "reverse" or abnormal fluorine chemical shift behavior correlates with C-F bond polarity: deshielded shifts correspond to more polar C-F bonds (higher charge separation, provided in Table 3.7).

Table 3.7. ^{19}F NMR shifts (experimental and calculated) and fluorine electrostatic potential (charge) for stable tautomers of 2-fluoro- and 4-fluoro-imidazole analogues.

τ-tautomer	**2F-His**	**2F-MeIm**	**2F-Im**
Fluorine (F2) ESP charge	-0.264	-0.278	-0.280
Carbon (C2) ESP charge	0.718	0.737	0.712
Δq	0.982	1.015	0.992
π-tautomer	**4F-His**	**4F-MeIm**	**4F-Im**
Fluorine (F4) ESP charge	-0.290	-0.275	-0.308
Carbon (C4) ESP charge	0.597	0.509	0.651
Δq	0.887	0.784	0.959

It is interesting to note that the less stable τ-tautomer of 4F-(5-methyl)-imidazole has a lone pair orbital on fluorine and exhibits a deshielded chemical shift upon protonation (like 2F analogues do). Thus, for N1-protonated (π-tautomer), 4-fluoroimidazole moieties, the conjugated (delocalized) molecular orbitals of fluorine appear to be an inherent characteristic of the electronic structure. All of the systems analysed here support a hypothesis that when fluorines on an aromatic ring have no lone pair (*i.e.* completely delocalized electron density), the chemical shifts are abnormal or "reverse". All aromatic fluorines with a lone pair, in the systems studied here, have ^{19}F chemical shifts whose shielding/deshielding correlates with fluorine charge density. Further study with electronic structure methods is required to determine whether this is a general feature of ^{19}F chemical shifts.

3.5. Normal and reverse chemical shifts behavior between fluoropyridine isomers

Effects of protonation on fluorine chemical shifts for the regioisomers 2-fluoropyridine (2F-Pyr) and 3-fluoropyridine (3F-Pyr) was investigated four decades ago.[25] Practically, protonating the nitrogen atom in pyridine ring should lead to a downfield chemical shift of the fluorine nuclear resonance due to a net drain of electrons from the ring. The experimental results showed that 3-fluoropyridine shows a downfield shift in the ^{19}F NMR spectrum upon protonation; however, ring protonation of 2-fluoropyridine causes an upfield shift, which is an unusual finding. To understand if the

same theory holds as in fluorohistidine isomers the similar calculations are performed on fluropyridine regioisomers.

3.5.1. Computational details

The M06 density functional[27] with 6-31+g(d) basis set was considered for geometry optimization of fluoropyridine iso mers, followed by NMR calculations. The contribution of each atomic/ molecular orbital towards fluorine nuclear shielding and para/diamagnetic parameters were is computed using NCS analysis by providing the "NCS" keyword in addition to the NMR keyword in the input file. NMR calculations are calculated using GIAO method with implicit water solvation using the CPCM model. The geometries of 2-fluoropyridine and 3-fluoropyridine, and the protonated fluoropyridiniums, were prepared in GaussView 5[14] software followed by geometry optimization and subsequent NMR/NCS calculations in Gaussian09[16].

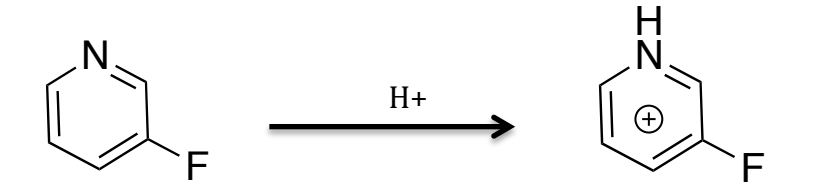
3.5.2. Results

The NMR calculations show that upon acid titration, the 3-fluoropyridine molecule shows a net downfield shift of 14.3 ppm, which is in good agreement with the experimental value of 12 ppm downfield. The fluorine's diamagnetic and paramagnetic shieldings in 3-fluoropyridine and 3-fluoropyridinium are given in Table 3.8. From the NCS analysis of the neutral and protonated forms of 3-fluoropyridine, it is seen that there is more diamagnetic shielding in the neutral molecule (478.6 ppm) compared to that of the protonated form (476.48 ppm), which reflects chemical intuition that the positive charge in the protonated pyridinium will withdraw electron density from the fluorine. The paramagnetic values reveal that, upon ring protonation, there is greater deshielding (− 177.9 ppm) of the fluorine atom compared to the neutral structure (− 166.3 ppm). Thus,

the normal chemical shift behavior upon protonation (net downfield ^{19}F NMR shift of 14.3 ppm) arises from both diamagnetic (lower ground-state electronic distribution) and paramagnetic effects (higher deshielding).

Table 3.8. Calculated ^{19}F shieldings for 3-fluoropyridine neutral and protonated form.

Molecule	Paramagnetic shielding (ppm)	Diamagnetic shielding (ppm)
3-fluoropyridine	-166.3	478.7
3-fluoropyridinium	-177.9	476.5
Net change	11.5	2.2



3-fluoropyridine **3-fluoropyridinium**

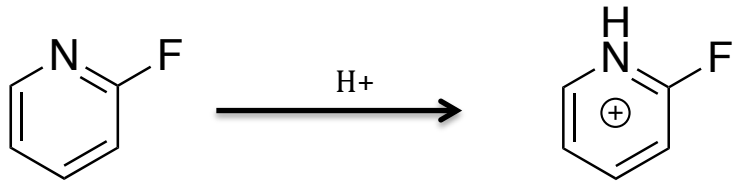
Experimental value[25]: 12 ppm (deshielding)
Net overall change (calculated): 14 ppm (deshielding)

Table 3.9 presents the calculated shieldings of ^{19}F in 2-fluoropyridine and its protonated form. Unlike 2-fluoropyridine, which could be considered to exhibit a normal fluorine chemical shift, protonation of 3-fluoropyridine leads to a more shielded ^{19}F nucleus (the extent of shielding is not quantified in the literature)[25]. Chemical shift calculations of the ^{19}F nucleus in 3-fluoropyridine and 3-fluoropyridinium predict an upfield shift of − 7.83 ppm upon acid titration. NCS analysis of diamagnetic shielding indicates the same trend as in 2-fluoropyridine, with a decrease in 3-fluoropyridine's diamagnetic shielding upon protonation (i.e. diamagnetic deshielding upon protonation). Again, the diamagnetic contribution trends with a decrease in ground-state charge

electron density or calculated charge. The observed shielding, then, is shown to arise from the paramagnetic contributions to chemical shift. It can be seen that in 3-fluoropyridine, paramagnetic deshielding is much higher in the neutral structure (− 220.2 ppm), compared to the protonated structure (− 208.6 ppm). These results illustrate that a decrease in paramagnetic deshielding is the primary cause for the upfield shift of ^{19}F that is experimentally observed in the protonated struc- ture.

Table 3.9. Calculated ^{19}F shieldings for 2-fluoropyridine neutral and protonated form.

Molecule	Paramagnetic shielding (ppm)	Diamagnetic shielding (ppm)
2-fluoropyridine	-220.2	476.1
2-fluoropyridinium	-208.7	472.0
Net change	-11.6	4.1



2-fluoropyridine **2-fluoropyridinium**

Experimental value not clearly quantified, but is shielding (negative)[25]
Net overall change (calculated): -7.8 ppm

It is interesting to note that the ^{13}C chemical shifts of pyridine (*not* fluoropyridine) upon protonation were found to depend on their position in the ring, with the carbons *ortho* to nitrogen being shielded (Fig. 5.5) and the other carbons in the ring deshielded. Meanwhile, the ^{15}N chemical shift in pyridine, when protonated, is significantly shielded. Duthaler and Roberts explain that the lone pairs of nitrogen in neutral pyridine are perpendicular to the aromatic π* orbitals, which leads to paramagnetic (de)shielding

through coupling of these orbitals in the presence of a magnetic field.[28] Thus, the shielding of ^{15}N upon protonation can be understood to arise because the nitrogen lone pair makes a significant paramagnetic deshielding contribution to neutral pyridine that is eliminated when it is protonated.

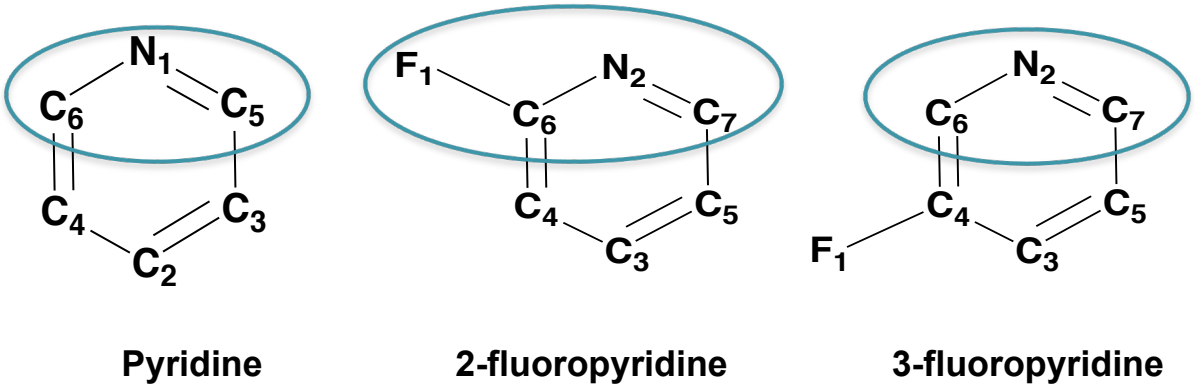


Figure 3.5. Structures of pyridine, 2-fluoropyridine and 3-fluoropyridine. Atoms showing reverse chemical shifts (upfield) upon ring protonation are circled.

The different results in shielding of fluorine at the 2nd position in pyridine after protonation (due to a reduction in paramagnetic deshielding) and deshielding of fluorine at the 3rd position in pyridine (due to an increase in paramagnetic deshielding) may arise from symmetry considerations in the coupling of fluorine lone-pair orbitals to orbitals in the aromatic ring. To understand this we also looked into para- and dia- magnetic contributions towards total ^{13}C and ^{15}N chemical shifts. The results show that the of nitrogen and the carbons ortho to nitrogen, in pyridine exhibit the same diamagnetic shielding in neutral and and protonated form. Also, the calculations display that for meta and para carbons (C4, C3, C2) there is more paramagnetic deshielding for protonated structure compared to the neutral one. However, for nitrogen and carbons ortho to

nitrogen (i.e. N1, C5, C6), there is higher paramagnetic (de)shielding in the neutral structure compared to the protonated. (Fig. 3.6) The calculations on 2F-Pyr and 3F-Pyr demonstrates the same behavior of increasing paramagnetic deshielding for nitrogens and ortho carbons in neutral structure respective to the protonated structure. These results explains that the decrease in paramagnetic deshielding is the predominant reason for the upfield shift of nitrogen and ortho carbons that is experimentally observed in the protonated structure. The different results in shielding of fluorine at the 2-position in pyridine after protonation (due to a reduction in paramagnetic deshielding) and deshielding of fluorine at the 3-position in pyridine (due to an increase in paramagnetic deshielding) may arise from symmetry considerations in the coupling of fluorine lone-pair orbitals to orbital s in the aromatic ring.

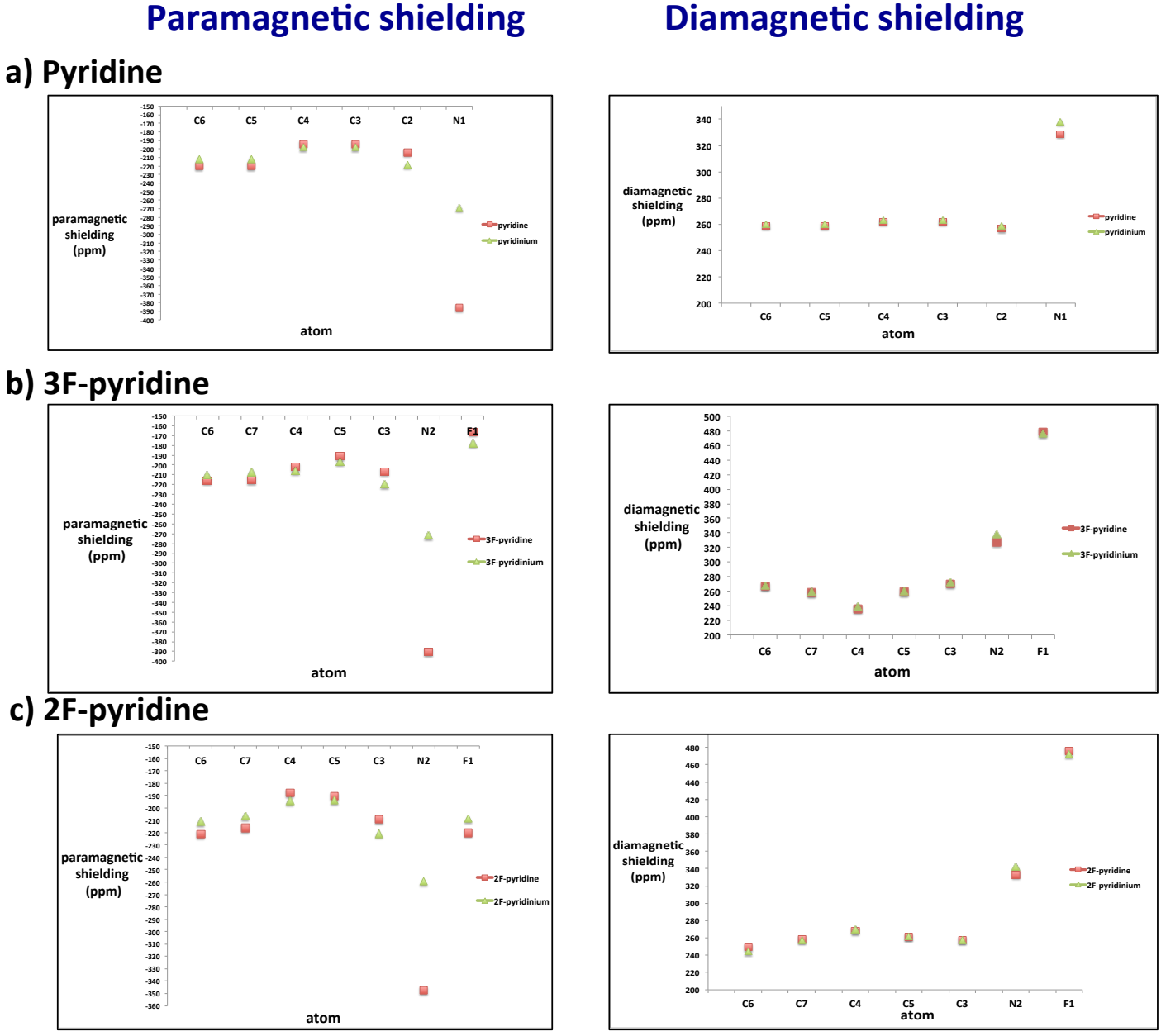


Figure 3.6. Paramagnetic and diamagnetic contribution towards total shielding for a) pyridine and pyridinium b) 3-fluoropyridine and 3-fluoropyridinium c) 2-fluoropyridine and 2-fluoropyridinium.

3.6. CONCLUSIONS

In summary, the electronic structure of 4F-(5-methyl)-imidazole (and histidine analogues) is unique, compared to the electronic structure of 2F-(5-methyl)-imidazole, 2F-(5-methyl)-imidazolium, and 4F-(5-methyl)-imidazolium analogues. Rather than a lone pair, all of the electrons and orbitals of fluorine in 4F-(5-methyl)-imidazole are conjugated with with the aromatic ring and adjacent methyl group. Since shielding of the fluorine nucleus upon protonation was calculated and observed experimentally for 4F-(5-methyl)-imidazole, 4F-imidazole, and 4F-histidine, and calculated for 4F-(5-trifluoromethyl)-imidazole, it seems that the delocalized electronic structure has little dependence on substitution at C5. Rather, delocalized fluorine electron density appears to be an inherent characteristic of the π-tautomer (N1-H) of 4-fluoro-imidazoles. For 2F-(5-methyl)-imidazole species, predictions of deshielding from shifts of electron density away from the fluorine nucleus upon addition of positive charge are substantiated in the less-negative fluorine electrostatic potential (charge), significantly smaller dipole moment, and deshielding contributions from the fluorine lone pair of MO 2 (calculated by NCS analysis). These reflect a reduction in diamagnetic (ground state) shielding. Fig. 3 shows how the ^{19}F-NMR chemical shifts of 2F-(5-methyl)-imidazole correlate with electronic structure properties, while only NCS data correlates well with the ^{19}F-NMR chemical shifts of 4F-(5-methyl)-imidazole. Within a series of C5-substituted 4F-imidazoles, correlations were found between increased C-F bond polarity and deshielding of ^{19}F chemical shifts. This abnormal or "reverse" relationship with bond polarity had been characterized previously

only for aliphatic systems.[1] Thus, when fluorine delocalization takes place in aromatic systems, fluorine chemical shift prediction may require more complex analysis, such as that provided by computational methods.

The concept of fluorine's complex electronic effects, such as hyperconjugation and lone pair back-donation, is not new,[27] but the importance in terms of ^{19}F-chemical shifts has been more difficult to nail down. The results here for 2F- and 4F-histidine/(5-methyl)-imidazole seem to suggest that when electron density is localized in a fluorine lone pair, changes in ^{19}F NMR chemical shifts might be understood and predicted with the same chemical knowledge and intuition as ^{1}H and ^{13}C chemical shifts. However, when electronic effects lead to conjugation of all fluorine orbitals, a framework for understanding fluorine chemical shifts is not so straightforward. For understanding ^{19}F chemical shifts in proteins, the next step will be to understand to what extent local environment induces fluorine lone pair localization/delocalization. With a database of proteins with known local environments and measured ^{19}F chemical shifts,[49] and continued computational efforts, we may not be too far off from a unifying protocol for *a priori* prediction of ^{19}F-NMR spectra in proteins.From the calculated paramagnetic and diamagnetic contributions towards total shielding it appears that in the case of 2-fluoropyridine, higher paramagnetic deshielding in the neutral structure (relative to the protonated state) appears to be the primary reason for the observed upfield shift upon ring protonation, whereas, in the case of 4-fluorohistidine, both factors, *i.e.*, increased diamagnetic shielding and lowered paramagnetic deshielding parameters, support the upfield shift, explaining the "reverse chemical shift"

behavior. All of these results and analysis demonstrate that NCS analysis is a valuable tool, helping to rationalize the origins of fluorine chemical shifts through analysis of diamagnetic and paramagnetic contributions to shielding.

REFERENCES

[1] Adcock, W., and Abeywickrema, A. N. (1982) Concerning the origin of substituent-induced F-19 chemical-shifts in aliphatic fluorides - C-13 and F-19 nuclear magnetic-resonance study of 1-fluoro-4-phenylbicyclo[2.2.2]octanes substituted in the arene ring. *J Org Chem 47*, 2945-2951.

[2] Hudson, E. P., Eppler, R. K., Beaudoin, J. M., Dordick, J. S., Reimer, J. A., and Clark, D. S. (2009) Active-site motions and polarity enhance catalytic turnover of hydrated subtilisin dissolved in organic solvents. *J Am Chem Soc 131*, 4294-4300.

[3] Lau, E. Y., and Gerig, J. T. (2000) Origins of fluorine NMR chemical shifts in fluorine-containing proteins? *J Am Chem Soc 122*, 4408-4417.

[4] Frieden, C., Hoeltzli, S. D., and Bann, J. G. (2004) The preparation of 19F-labeled proteins for NMR studies. *Methods Enzymol 380*, 400-415.

[5] Luck, L. A., and Falke, J. J. (1991) 19F NMR studies of the D-galactose chemosensory receptor. 1. Sugar binding yields a global structural change. *Biochemistry-US 30*, 4248-4256.

[6] Chen, H., Viel, S., Ziarelli, F., and Peng, L. (2013) F-19 NMR: a valuable tool for studying biological events. *Chem Soc Rev 42*, 7971-7982.

[7] Mulder, F. A., and Filatov, M. (2010) NMR chemical shift data and ab initio shielding calculations: emerging tools for protein structure determination. *Chem Soc Rev 39*, 578-590.

[8] Oldfield, E. (2002) Chemical shifts in amino acids, peptides, and proteins: From quantum chemistry to drug design. *Ann Rev Phys Chem 53*, 349-378.

[9] Holtz, D. (1971) Critical examination of fluorine hyperconjugation in aromatic systems. *Chem Rev 71*, 139-145.

[10] Sheppard, W. A. (1965) The Electronic Properties of Fluoroalkyl Groups. Fluorine p-π Interaction. *J Am Chem Soc 87*, 2410-2420.

[11] Sternberg, U., Klipfel, M., Grage, S. L., Witter, R., and Ulrich, A. S. (2009) Calculation of fluorine chemical shift tensors for the interpretation of oriented F-19-NMR spectra of gramicidin A in membranes. *Phys Chem Chem Phys 11*, 7048-7060.

[12] Dalvit, C., and Vulpetti, A. (2011) Fluorine-protein interactions and (1)(9)F NMR isotropic chemical shifts: An empirical correlation with implications for drug design. *ChemMedChem 6*, 104-114.

[13] Yeh, H. J. C., Kirk, K. L., Cohen, L. A., and Cohen, J. S. (1975) F-19 and H-1 nuclear magnetic-resonance studies of ring-fluorinated imidazoles and histidines. *J Chem Soc Perk T 2*, 928-934.

[14] Dennington, R., Keith, T. A., and Millam, J. M. (2009) Semichem Inc. Version 5.0. Shawnee Mission KS.

[15] Andra, K. K., Bullinger, J. C., Bann, J. G., and Eichhorn, D. M. (2010) 2-Fluoro-L-histidine. *Acta Crystallogr E 66*, O2713-U1695.

[16] 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., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G.; Sonnenberg, J. L., *et al.* . (2009) Gaussian, Inc., Gaussian 09. Wallingford, CT..

[17] Bohmann, J. A., Weinhold, F., and Farrar, T. C. (1997) Natural chemical shielding analysis of nuclear magnetic resonance shielding tensors from gauge-including atomic orbital calculations. *J Chem Phys 107*, 1173-1184.

[18] Becke, A. D. (1993) A new mixing of hartree-fock and local density-functional theories. *J Chem Phys 98*, 1372-1377.

[19] Kohn, W., and Sham, L. J. (1965) Self-consistent equations including exchange and correlation effects. *Phys Rev 140*, 1133-&.

[20] Marenich, A. V., Cramer, C. J., and Truhlar, D. G. (2009) 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 113*, 6378-6396.

[21] Barone, V., and Cossi, M. (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J Phys Chem A 102*, 1995-2001.

[22] Wolinski, K., Hinton, J. F., and Pulay, P. (1990) Efficient implementation of the gauge-independent atomic orbital method for NMR chemical-shift calculations. *J Am Chem Soc 112*, 8251-8260.

[23] Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., and Hutchison, G. R. (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminformatics 4*.

[24] Reynolds, W. F., Peat, I. R., Freedman, M. H., and Lyerla, J. R. (1973) Determination of tautomeric form of imidazole ring of L-histidine in basic solution by C-13 magnetic-resonance spectroscopy. *J Am Chem Soc 95*, 328-331.

[25] Giam, C. S., and Lyle, J. L. (1973) Medium effects on fluorine-19 magnetic-resonance spectra of fluoropyridines. *J Am Chem Soc 95*, 3235-3239.

[26] Breneman, C. M., and Wiberg, K. B. (1990) Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J Comput Chem 11*, 361-373.

[27] Zhao, Y., and Truhlar, D. G. (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor Chem Acc 120*, 215-241.

[28] Duthaler, R. O., and Roberts, J. D. (1978) Effects of solvent, protonation, and N-alkylation on N-15 chemical-shifts of pyridine and related compounds. *J Am Chem Soc 100*, 4969-4973.

CHAPTER 4

Characterization of CH-----O hydrogen bond in histidine/ fluorohistidine-water complexes

4.1. INTRODUCTION

The hydrogen bond has a crucial role in protein structure-functions affecting protein folding and enhancing the stability of folded proteins. Hydrogen bonds are involved in the stabilization of secondary structures, protein-ligand interactions and proton transfer mechanisms in many enzymatic reactions. Until recently, carbon has not been recognized as being involved in hydrogen bonding because of its low electronegativity compared to nitrogen or oxygen. But over the past few decades, studies have showed the role of the C-H bond as a potential hydrogen bond donor in protein structure and function. CH—O hydrogen bonding is observed in many protein crystal structures and nucleic acids.[1, 2] X-ray crystallographic studies illustrated that CH-O hydrogen bond interactions between oxygen of phosphate backbone and carbon atom of purines/pyrimidines tends to stabilize the RNA tertiary structure.[3] In a study it was revealed that the CH-O interaction in thyamine is a common and strong hydrogen bond, the same as that of a conventional NH—O hydrogen bond.[4] Even though the interactions where C is acting as a H-bond donor are weaker than the hydrogen bonds by more electronegative atoms such as nitrogen or oxygen, the studies showed that the presence of electronegative groups in close proximity helps activate C-H bonding ability by strengthening its acidity.[5 6 7] Examples of proteins and nucleic acids involved in CH---O interactions are depicted in Fig. 4.1.

Studies and protein crystal structures showed the role of imidazole ring carbon donor hydrogen bonds in protein stability interactions and in metal binding [8-10] (Fig. 4.2). Nanda and co-workers showed imidazole carbon in metal complexes form potent hydrogen bond donors.[11] In the copper containing metalloenzyme cytochrome *c* oxidase, the polar C-H bonds of imidazole are involved in the enzymes structural stability.[12]

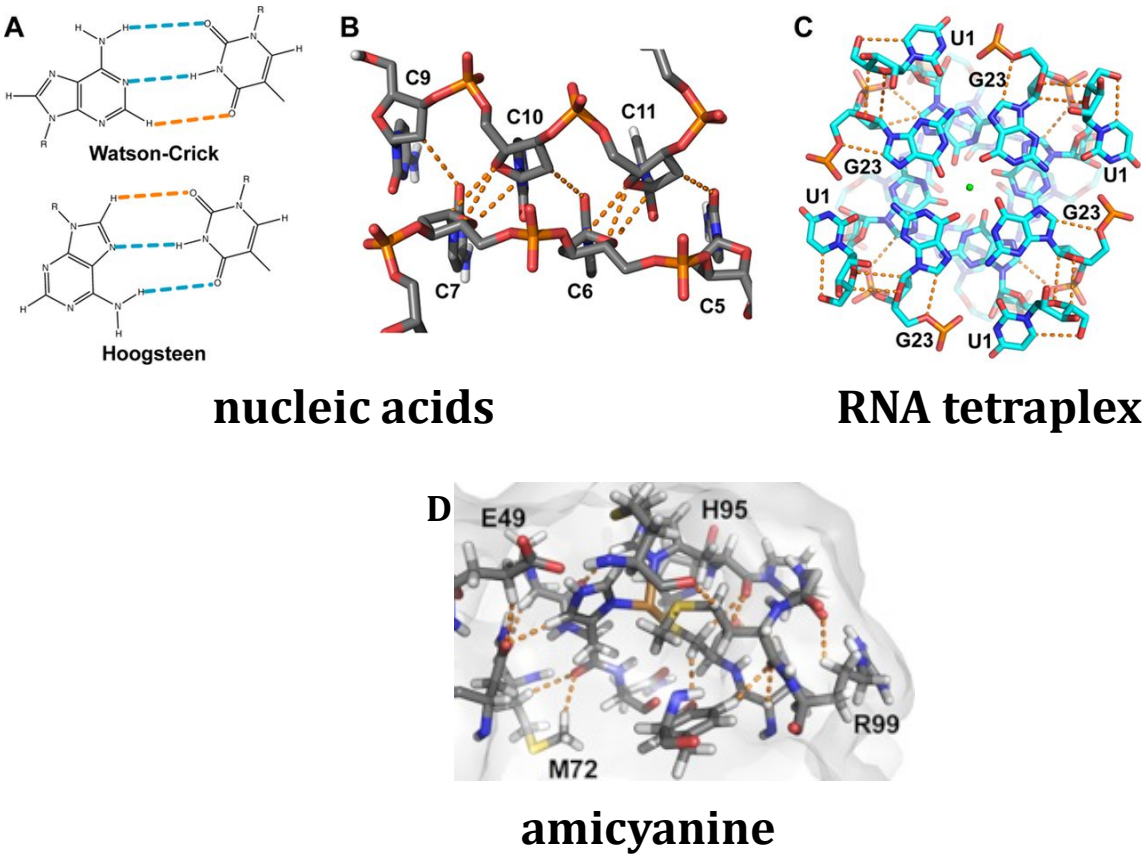


Figure 4.1. A) Hydrogen bonding patterns in Watson-Crick and Hoogstein adenine thyamidine basepairs. B) in DNA i-motif [13] C) in RNA tetraplex[14] D)extensive CH—O hydrogen bonding in copper coordination site in amicyanine.[15]

The most common method used to evaluate hydrogen bonding character is analysis of the donor acceptor (D-A) distance and bond angle (D-H—A) and the lengthening of the X-H bond distance (X=hydrogen bond donor atom such as N, O or C). Along with this, NMR chemical shifts are used to indicate hydrogen bonds and IR

spectroscopy is also used to measure the hydrogen bond strength. In CH—O interactions the hydrogen bond is characterized by the proton's downfield shift and a red shift in C-H IR stretching frequencies. The charge transfer phenomenon also is the fundamental cause for the change in C-H bond distances.[16, 17]

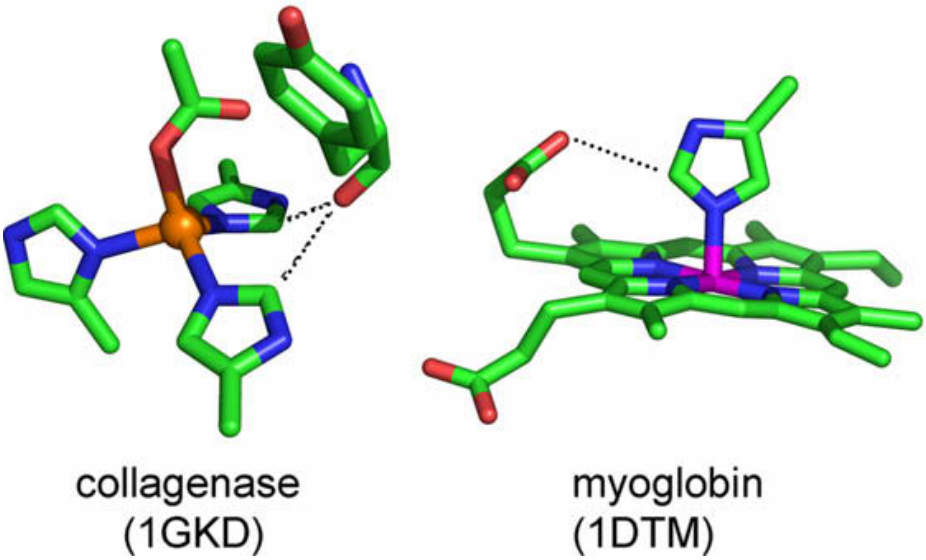


Figure 4.2. Examples of observed carbon donor hydrogen interactions involving metal-bound imidazoles. [11]

Previous studies showed the role of Cε (C2) carbon of histidine as a H-bond donor in hydrogen bonding interaction in metalloproteins and enzymes. Our electronic structure calculations indicate the Cδ-H bond is more polar in fluorohistidine and canonical histidine, which may reflect a propensity for hydrogen bonding. To check if there are any existing Cδ-H—O hydrogen bond interactions, we scrutinized the crystal structures of proteins which are solved through neutron diffractions. Through neutron diffraction, it is easy to differentiate the protonation states of imidazole rings, as this technique allows for well –defined position of all the atoms including protons.[18-20] The

screening shows that, in some proteins, the C4 carbon is involved in hydrogen bond interactions with other electronegative atoms. Examples of C2-H—O and C4-H---O interactions in protein crystal structures are shown in Fig. 4.3.[21-23]

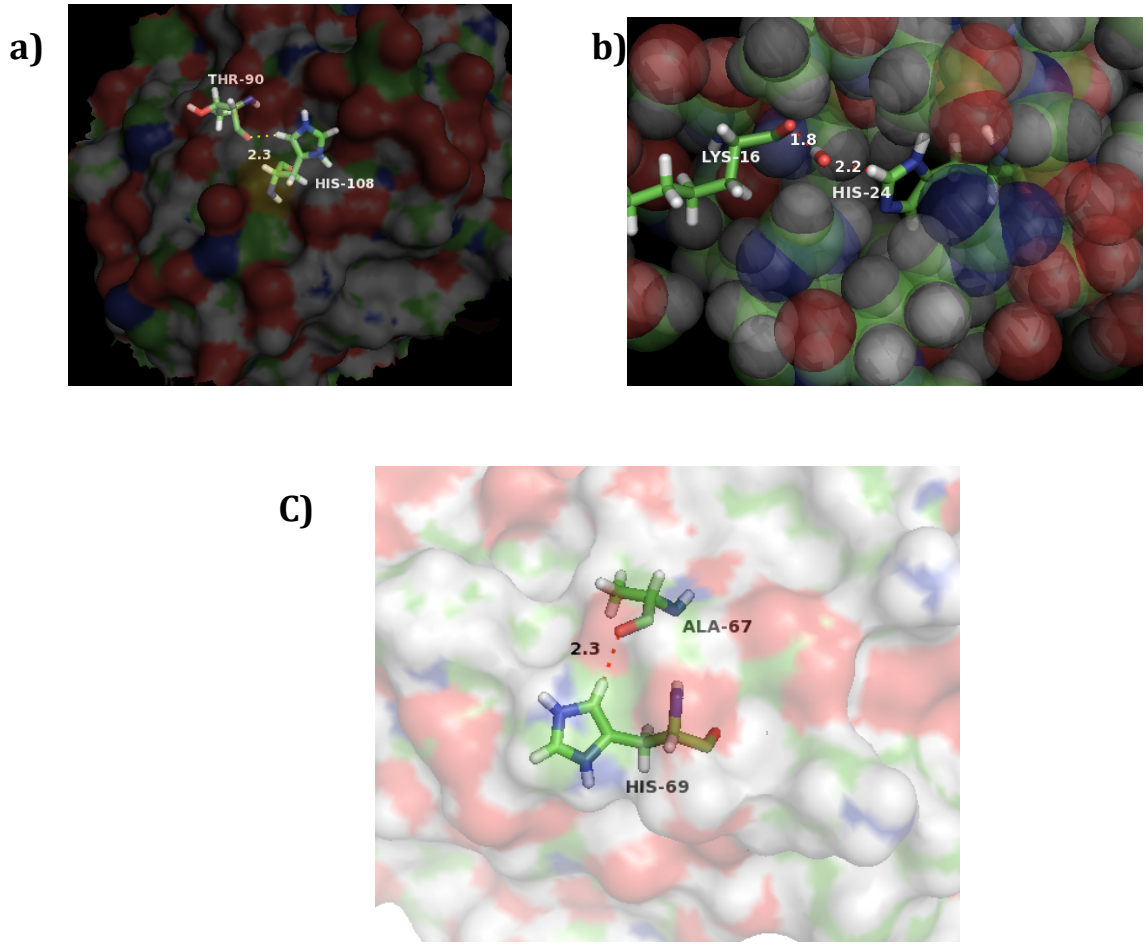


Figure 4.3. Examples of CεH---O (C2H--O) and Cδ-H---O (C4H---O) interactions in protein crystal structures. a) Photoactive yellow protein (PDB ID: 2ZOI)[21] b) HIV protease (PDB ID: 2ZYE)[22] c) Myoglobin (PDB ID: 1L2K).[23]

In this work we corroborated the presence of Cδ-H—O (C4-H) interaction between histidine and water complexes. The results indicate that histidine's imidazole ring in τ-tautomeric states acts as a strong hydrogen bond donor compared to imidazole ring in N1-H (π-tautomer) protonated state. Further the calculations demonstrate that the

presence of the electronegative atom fluorine at the 2nd position makes histidine a more potent hydrogen bond donor compared to canonical histidine.

4.2. COMPUTATIONAL DETAILS:

All the geometries are energy minimized using BHandHLYP and second order Moller-Plesset perturbation theory (MP2) functionals with 6-31+g(d) basis set. The geometries include histidine-τ tautomer (**His-τ**), histidine-π tautomer (**His-π**), complex1: histidine-τ/water complex (**His-τ/H2O**), complex2: histidine-π/water complex (**His-π/H2O**), 2-fluorohistidine-τ-tautomer (**2FHis-τ**), 2-fluorohistidine-π-tautomer (**2FHis-π**), complex3: 2-fluorohistidine-τ-water complex (**2FHis-τ/H2O**) and complex4: 2-flurohistidine-π-water complex (**2FHis-π/H2O**). Restricted (or partial) energy minimization is performed on complexes, in which only Cδ-Hδ atoms from histidine and H-O-H from water are allowed to move, leaving all other atoms rigid. The electrostatic potential distribution on each atom, free energies, C-H bond stretching frequencies, electron densities and natural bond orbital (NBO) calculations are performed using BHandHLYP and MP2 methods. The hydrogen bonding interaction energy, ΔE given in Eq. 4.1, is calculated by taking the difference between the energy of histidine-water complex (E_{AB}) and the summation of histidine energies (E_A and E_B respectively). The binding energies corrected for basis set superposition error by counterpoise method are also reported. All the calculations are performed in Gaussian 09 software. Gauss View5 software is used for molecular preparation and visualization.

$$\Delta E = E_{AB} - (E_A + E_B) \qquad (4.1)$$

4.3. RESULTS

4.3.1. Energies

The optimized geometries of histidine and 2-fluorohistidine tautomeric states with BHandHLYP/6-31+g(d) method are shown in Fig. 4.4. The calculated energies suggest that both the fluorinated and non-fluorinated analogues are stable in their N3-H (τ-tautomer) protonated conformation, compared to N1-H (π-tautomer) conformations. In the case of histidine, our calculations show that the τ-tautomer is stabilized by 1.5 kcal/mol with respect to π-tautomer. Experimentally the tautomeric ratio (τ : π) for canonical histidine is 80 : 20.[24] In the same way, in 2-fluorohistidine the τ-tautomer is observed to be more stable than π-tautomer with energy difference of 1.3 kcal/mol. Also, the calculated energies with MP2/6-31+g(d) are indicated in Table 4.1 within parentheses.

Fig. 4.5 shows the optimized conformations of histidine and water and 2-fluorohistidine and water in both the τ and π tautomeric states. The histidine-τ/H_2O complex is observed to be highly stable, with an energy difference by ~4 kcal/mol (table 4.1) with respect to the His-π/H2O conformation. Also, the hydrogen bond interaction with the fluorinated analogue is observed to enhance the τ-tautomer stability by an energy difference of ~4 kcal/mol with respect to the π-tautomer-water interaction. This data illustrates the stabilization of τ-tautomeric structures of histidine and 2-fluorohistidine by their hydrogen bond interactions with a water molecule.

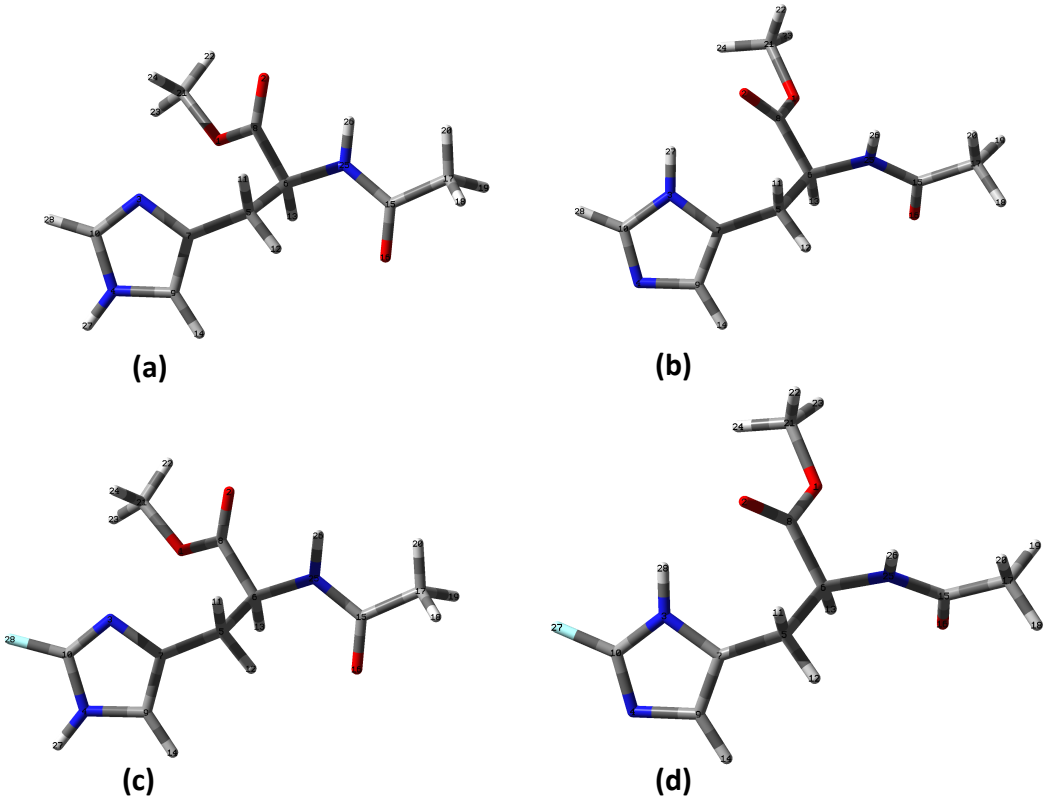


Figure 4.4. Optimized geometries of a) histidine-τ/H_2O (His-τ/H_2O), b) histidine-π (His-π/H_2O) c)2-fluorohistidine-τ/H_2O (2FHis-τ/H_2O), d) 2-fluorohistidine-π/H_2O (2FHis-π/H_2O)

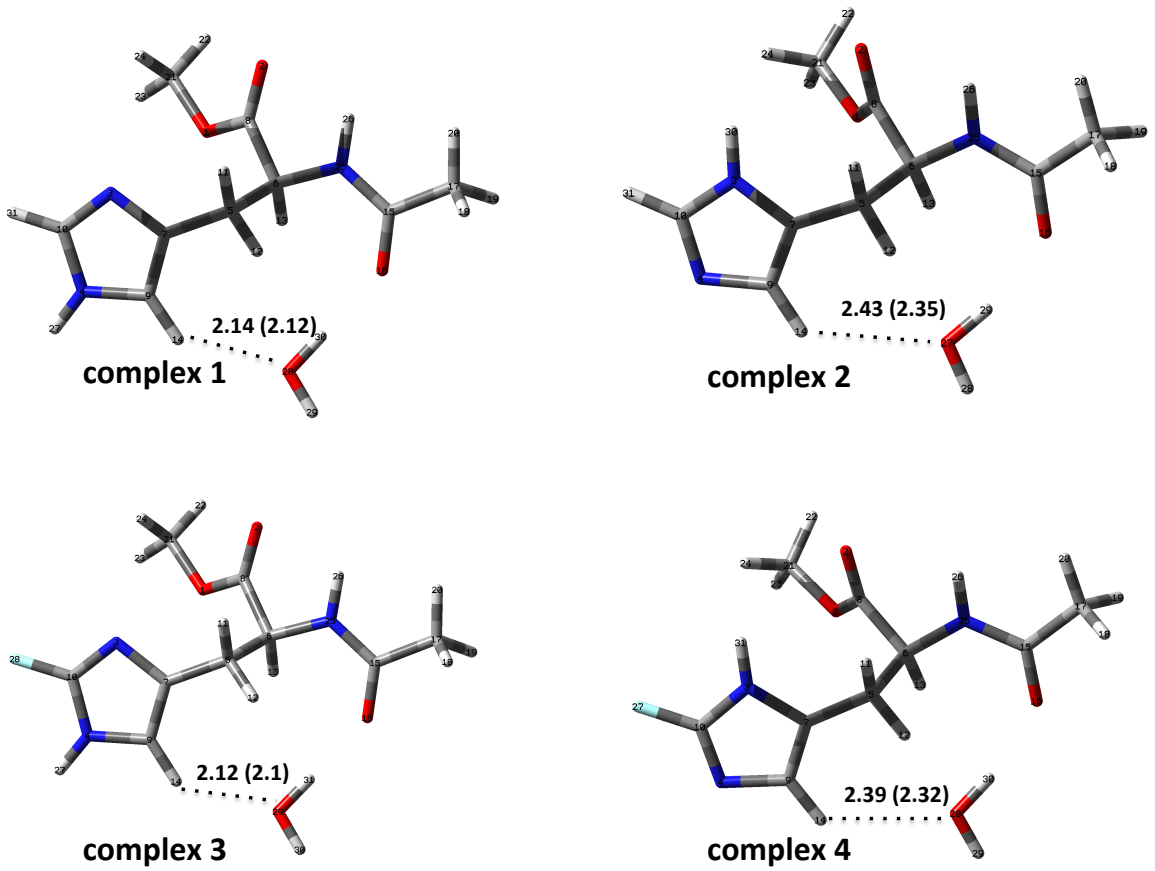


Figure 4.5. Optimized geometries of complex 1: histidine-τ/H_2O (His-τ/H_2O); complex 2: histidine-π/H_2O (His-π/H_2O); complex 3: 2-fluorohistidine-τ/H_2O (2FHis-τ/H_2O); complex 4: 2-fluorohistidine-π/H_2O (2FHis-π/H_2O). The dotted line indicates hydrogen bond and bond distance in Angstroms calculated with BHandHLYP/6-31+g(d) method. The values in paranthesis are bond distances calclated with MP2/6-31+g(d) method.

Table 4.1. Relative tautomeric energies for histidine and 2-fluorohistidine in their monomeric state and upon complexing with water.

Molecule	Tautomer	Monomer	Monomer+water complex
		ΔE	ΔE
histidine	**τ**	0.00 kcal/mol (0.00 kcal/mol)	0.00 kcal/mol (0.00 kcal/mol)
	π	1.52 kcal/mol (1.95 kcal/mol)	3.95 kcal/mol (3.70 kcal/mol)
2-fluorohistidine	**τ**	0.00 kcal/mol (0.00 kcal/mol)	0.00 kcal/mol (0.00 kcal/mol)
	π	1.28 kcal/mol (1.84 kcal/mol)	3.81 kcal/mol (3.52 kcal/mol)

4.3.2. Molecular properties

A lengthening in C-H bond distance is observed for all the complexes compared to their isolated/monomeric states (Table 4.2). Comparing the complexes, the elongation in H—O bond distance is observed up to 0.3 Å going from complex 2 to complex 1(His-π to His-τ) and from complex 4 to complex 3 (2FHis-π to 2FHis-τ). Further, it important to point out that by substituting the fluorine on the imidazole ring's τ-tautomer, the intermolecular bond distance (H---O) shortens to a greater extent (2.14 to 2.12 Å), indicating the strong interaction between water and 2FHis τ-tautomer. The increased hydrogen bond strength of τ-tautomer compared to π- tautomer in both fluorinated and nonfluorinated analogues is also corroborated by an increased C-H—O bond angle, with τ-tautomers having a more linear geometry.

Table 4.2. Calculated hydrogen bond distance (C-H, CH—O, Å) and bond angles (C-H—O, °) for isolated histidine molecules and complexes.

Compound	Parameters	BHandHLYP	MP2
His-τ	B_{C-H} C9-H14	1.071	1.082
His-τ/H20	B_{C-H} C9-H14	1.073	1.0826
(complex1)	$B_{H—O}$ H14-O28	2.140	2.12
	A_{C-H--O} C9-H14—O28	139.83	137.46
His-π	B_{C-H} C9-H14	1.073	1.083
His-π/H2O	B_{C-H} C9-H14	1.072	1.082
(complex2)	$B_{H—O}$ H14-O27	2.43	2.35
	A_{C-H--O} C9-H14—O27	125.2	124.86
2FHis-τ	B_{C-H} C9-H14	1.071	1.081
2FHis-τ/H2O	B_{C-H} C9-H14	1.073	1.082
(complex 3)	$B_{H—O}$ H14-O29	2.120	2.1
	A_{C-H--O} C9-H14—O29	138.95	136.55
2FHis-π	B_{C-H} C9-H14	1.073	1.082
2-FHis-π/H2O	B_{C-H} C9-H14	1.072	1.0813
(complex 4)	$B_{H—O}$ H14-O28	2.390	2.32
	A_{C-H--O} C9-H14—O28	125.09	124.7

4.3.3. Interaction energies

Interaction energies are calculated for all the complexes with and without basis set superposition error corrections. The complex with a more negative interaction energy indicates a more highly stabilized structure. In complex 1, the histidine's interaction with water is highly favored compared to complex 2. In the same way in the fluorinated

analogue also, the τ-tautomer's complexation with water is highly stabilized respective to the π-tautomer. In addition, it is found that adding one fluorine atom increases the binding stability energies up to ~0.5 kcal/mol. These results indicate the increased stability among complexes in the order of **complex 2 (His-π/H_2O) < complex 4 (2FHis-π/H_2O) < complex 1 (His-τ/H_2O) < complex 3 (2FHis-τ/H_2O)** with both BHandHLYP and MP2 methods. The interaction energies by considering counterpoise correction with BHand HLYP method are also reported in Table 4.3.

Table 4.3. Calculated hydrogen bond interaction energies with ($\Delta E_{HBF, CP}$) and without basis set superposition error (ΔE_{HBF}) in kcal/mol for all the complexes using BHandHLYP and MP2 methods with 6/31+g(d) basis set.

	Complex 1	**Complex 2**	**Complex 3**	**Complex 4**
ΔE_{HBF} (kcal/mol) (BHandHLYP)	-10.98	-8.54	-11.19	-8.66
ΔE_{HBF} (kcal/mol) (MP2)	-12.29	-10.54	-12.50	-10.82
$\Delta E_{HBF,CP}$ (kcal/mol) (BHandHLYP)	-11.31	-9.54	-11.54	-9.84

4.3.4. Partial charge distribution

Absolute charge (electrostatic potential) differences between hydrogen bond forming atoms are calculated are are listed in Table 4.4. The charge differences (Δq) are calculated using NBO and ChelpG methods with both MP2 and BHandHLYP functionals. The differences in Δq indicate the electrostatic attraction between interacting atoms; higher values indicate stronger coulombic interactions. The calculated NBO and ChelpG

charge difference between H and O atoms is observed to be more for complexes in τ-tautomeric states (complex 1 and complex 3) compared to those in π-tautomeric states (complexes 2 and 4).

The measured ESP charges via BHandHLYP method indicate that the τ-tautomer's C-H bond is highly polar compared to π-tautomer's, with the difference in Δq of 0.56*e* atomic units in ChelpG calculations. Further, upon substituting fluorine on the 2^{nd} carbon of the imidazole ring the C-H bond polarity is further increased to 0.65 a.u. This same trend can be observed via MP2 calculations.

Further, the electrostatic attraction between hydrogen bond forming atoms is measured by taking the absolute charge difference between respective atoms. As shown in Table 4.4, the ChelpG charge difference between hydrogen and oxygen illustrates that 2-fluorohistidine's C-H bond forms the strongest hydrogen bond with a charge difference of 1.32 a.u., higher than all other systems studied in here.

Table 4.4. Calculated NBO ChelpG charge difference between the atoms involved in hydrogen bonding interactions for complexes.

Method	System	NBO		ChelpG	
		C-H charge difference (Δq)	H-O charge difference (Δq)	C-H charge difference (Δq)	H-O charge difference (Δq)
	Complex 1	0.35	1.31	0.56	1.25
	Complex 2	0.33	1.29	0.00	1.07
BHandHLYP	Complex 3	0.36	1.31	0.65	1.32
	Complex 4	0.32	1.29	0.10	1.13
	Complex 1	0.34	1.32	0.58	1.29
	Complex 2	0.30	1.30	0.00	1.10
MP2	Complex 3	0.35	1.33	0.68	1.36
	Complex 4	0.29	1.31	0.13	1.17

4.3.5. Natural Bond Orbital (NBO) Analysis

The Table 4.5 shows the occupation number for donor oxygen lone pairs and acceptor C-H bonds involved in CH----O interactions. The second order perturbation orbital energies for each CH---O interaction are listed in kcal/mol. According to calculated methods, it is seen that the C-H occupation number is high in complex 1 and complex 3 compared to complex 2 and complex 4. Also it is important to note that, when going form non-fluorinated compound (complex 1) to fluorinated compound (complex 3), the C-H antibond occupation number slightly increases. As can be seen in Table 5 the interaction energy (E2) is ~3 kcal/mol higher in complexes 1 and 3 (where histidine is in τ-tautomeric states) compared the complexes 2 and 4 in which histidine is in π-tautomeric states.

Table 4.5. NBO analysis of complexes showing second order perturbation interaction energies (E2) and occupation number for oxygen lone pairs and C-H antibonds.

System	Method	Donor	Acceptor	Occupancy of donor	Occupancy of acceptor	E^2 kcal/mol
Complex 1	BHandHLYP	LP(2)O28	BD*(1)C9-H14	1.99234	0.01605	3.73
	MP2			1.99402	0.01301	3.85
Complex 2	BHandHLYP	LP(2)O27	BD*(1)C9-H14	1.99542	0.01464	0.56
	MP2			1.99975	0.0125	0.81
Complex 3	BHandHLYP	LP(2)O29	BD*(1)C9-H14	1.99199	0.01654	3.87
	MP2			1.9938	0.01333	3.96
Complex 4	BHandHLYP	LP(2)O28	BD*(1)C9-H14	1.99518	0.01456	0.66
	MP2			1.9962	0.01235	0.93

Various studies have shown that the appearance of a hydrogen bond critical point between donor and acceptor indicates formation of hydrogen bond.[25, 26] At the hydrogen bond critical point, the value of electron density is relatively low. The calculated electron density contours for all the complexes are shown in Fig. 4.6. One hydrogen bond critical point is observed in complexes 2 and 4 between O27-H29----O29 (or O28 on the backbone). In complex 1 and 3, an extra hydrogen bond critical point is observed involving C9-H14 as donor and O28 (or O29) as acceptor. By this analysis, a C4-H--O bond is only observed between atoms oxygen and C4-H (Cδ-H) when the imidazole ring is in τ-tautomeric form.

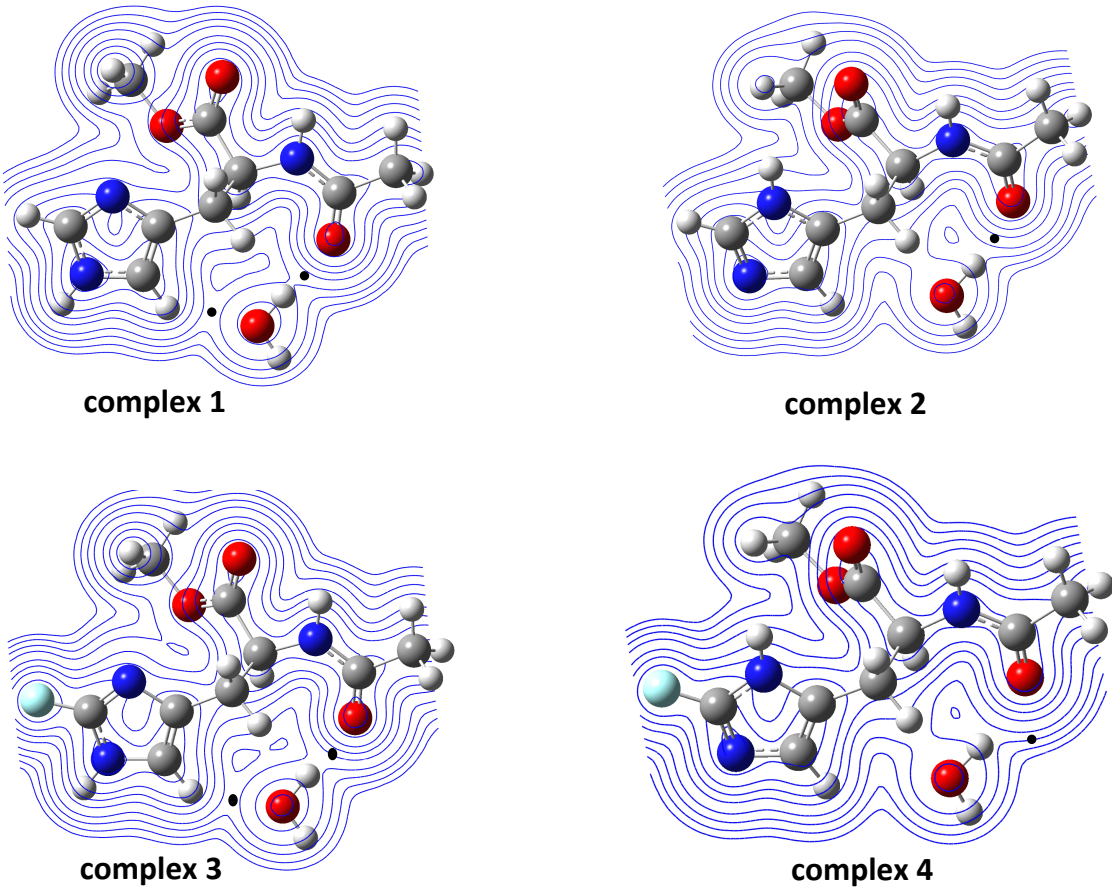


Figure 4.6. Electron density contour map for 1) complex 1 (His-τ/H2O) 2) complex 2 (His-π/H2O) 3) complex 3 (2FHis-τ/H2O) 4) complex 4 (2FHis-π/H2O).

4.3.6. IR stretching frequencies

C4-H bond stretching frequencies and IR intensities are calculated for His-τ, 2FHis-τ, alone and in complex with water. The interaction with the water molecule leads to lower frequencies and higher intensity of C-H stretching for both fluorinated and non-fluorinated histidine. The histidine's τ-tautomer interaction with water lead to lower the C-H bond stretching frequency from 3389 cm^{-1} to 3382 cm^{-1} with increased intensity (Table 4.6). Further, an enhanced red shift is observed (from 3395 to 3386 cm^{-1}) for fluorinated histidine when it is coupled to water. The higher values of red shift and

increased intensity in C-H stretching indicates strong interaction. Thus, the results show that 2FHis-τ is a potent hydrogen bond donor compared to its non-fluorinated analogue. This has implications for protein engineering work and fluorotagging of proteins with 2F-His for ^{19}F NMR spectroscopic studies.

Table 4.6. Calculated vibrational frequencies (cm^{-1}), IR intensity and red shifts of C4-H (Cδ-H) bond stretching upon complexation with water molecule.

System	C-H bond stretching frequencies		
	IR Frequency (cm^{-1})	Intensity	Red shift(cm^{-1})
His-τ	3389.3	0.3215	-
His-τ/H2O(complex1)	3382.57	27.3255	7
2FHis-τ	3395.3	1.2135	-
2FHis-τ/H2O(complex3)	3386.57	30.2261	9

4.4. DISCUSSION

In this study, we examined the role of Cδ-H bond of histidine as a potential hydrogen bond donor. From the results, Cδ-H appears to be a true hydrogen bond donor in the τ-tautomers of histidine and 2-fluorohistidine. The results indicate that when histidine and fluorohistidine are in N3-H protonation state (τ-tautomer) the molecules acts as a real hydrogen bond donor compared to when they are in N1-H protonation state. To our knowledge, this is the first identification of histidine Cδ-H---O hydrogen bonds in crystal structures. By surveying all the protein structures solved by neutron diffraction in

the RCSB (having resolution < 2.0 A°), it was found by inspection of crystallographic structures in the PyMOL visualization suite that 8 of the 13 structures had histidine C2-H (Cε-H) hydrogen bonds. (Out of 8 structures 4 histidines are protonated, 3 are in τ tautomeric state and 1 is in π tautomeric state). Meanwhile 5 histidine C4-H (Cδ-H) hydrogen bonds were identified, with 3 being in protonated state and π and τ tautomeric states have each. The C4-H bond is sterically hindered relative to the previously identified histidine C2-H hydrogen bond donor. However, this work suggests that the C4-H---O interactions of the τ-tautomer, when sterically feasible, is a strong interaction and researchers should be aware of it when analyzing protein structures to understand the stability and structure-function. Also, inserting the fluorine on imidazole carbon (C2) appears to enhance the hydrogen bond donor feature. Binding energies, interaction energies, and molecular properties suggest that the C-H—O bond interactions can be considered to have a potentially important role in protein folding and stabilization.

REFERENCES

[1] Derewenda, Z. S., Lee, L., and Derewenda, U. (1995) The occurrence of C-H...O hydrogen bonds in proteins. *J Mol Biol 252*, 248-262.

[2] Wahl, M. C., and Sundaralingam, M. (1997) C-H center dot center dot center dot O hydrogen bonding in biology. *Trends Biochem Sci 22*, 97-102.

[3] Brandl, M., Lindauer, K., Meyer, M., and Suhnel, J. (1999) C-H...O and C-H...N interactions in RNA structures. *Theor Chem Acc 101*, 103-113.

[4] Mandel-Gutfreund, Y., Margalit, H., Jernigan, R. L., and Zhurkin, V. B. (1998) A role for CH...O interactions in protein-DNA recognition. *J Mol Biol 277*, 1129-1140.

[5] Steiner, T., and Desiraju, G. R. (1998) Distinction between the weak hydrogen bond and the van der Waals interaction. *Chem Commun*, 891-892.

[6] Hobza, P., and Havlas, Z. (2002) Improper, blue-shifting hydrogen bond. *Theor Chem Acc 108*, 325-334.

[7] Hobza, P. (2002) N-H…F improper blue-shifting H-bond. *Int J Quantum Chem 90*, 1071-1074.

[8] Nanda, V., and Schmiedekamp, A. (2008) Are aromatic carbon donor hydrogen bonds linear in proteins? *Proteins 70*, 489-497.

[9] Harding, M. M. (1999) The geometry of metal-ligand interactions relevant to proteins. *Acta Crystallogr D 55*, 1432-1443.

[10] Horowitz, S., and Trievel, R. C. (2012) Carbon-Oxygen Hydrogen bonding in biological structure and function. *J Biol Chem 287*, 41576-41582.

[11] Schmiedekamp, A., and Nanda, V. (2009) Metal-activated histidine carbon donor hydrogen bonds contribute to metalloprotein folding and function. *J Inorg Biochem 103*, 1054-1060.

[12] Lockwood, C. W., Burlat, B., Cheesman, M. R., Kern, M., Simon, J., Clarke, T. A., Richardson, D. J., and Butt, J. N. (2015) Resolution of key roles for the distal pocket histidine in cytochrome C nitrite reductases. *J Am Chem Soc 137*, 3059-3068.

[13] Westler, W. M., Lin, I. J., Perczel, A., Weinhold, F., and Markley, J. L. (2011) Hyperfine-shifted 13C resonance assignments in an iron-sulfur protein with quantum chemical verification: aliphatic C-H...S 3-center-4-electron interactions. *J Am Chem Soc 133*, 1310-1316.

[14] Deng, J. P., Xiong, Y., and Sundaralingam, M. (2001) X-ray analysis of an RNA tetraplex (UGGGGU)(4) with divalent Sr2+ ions at subatomic resolution (0.61 angstrom). *Proc Natl Acad Sci USA 98*, 13665-13670.

[15] Sukumar, N., Mathews, F. S., Langan, P., and Davidson, V. L. (2010) A joint X-ray and neutron study on amicyanin reveals the role of protein dynamics in electron transfer. *Proc Natl Acad Sci USA 107*, 6817-6822.

[16] Wolfe, S., Pinto, B. M., Varma, V., and Leung, R. Y. N. (1990) The Perlin Effect - Bond Lengths, Bond Strengths, and the Origins of Stereoelectronic Effects Upon One-Bond C-H Coupling-Constants. *Can J Chem 68*, 1051-1062.

[17] Cieplak, A. S. (1981) Stereochemistry of Nucleophilic-Addition to Cyclohexanone - the Importance of 2-Electron Stabilizing Interactions. *J Am Chem Soc 103*, 4540-4552.

[18] Katz, A. K., Li, X., Carrell, H. L., Hanson, B. L., Langan, P., Coates, L., Schoenborn, B. P., Glusker, J. P., and Bunick, G. J. (2006) Locating active-site hydrogen atoms in D-xylose isomerase: time-of-flight neutron diffraction. *Proc Natl Acad Sci USA 103*, 8342-8347.

[19] Kossiakoff, A. A., and Spencer, S. A. (1981) Direct determination of the protonation states of aspartic acid-102 and histidine-57 in the tetrahedral intermediate of the serine proteases: neutron structure of trypsin. *Biochemistry-US 20*, 6462-6474.

[20] Blakeley, M. P., Mitschler, A., Hazemann, I., Meilleur, F., Myles, D. A., and Podjarny, A. (2006) Comparison of hydrogen determination with X-ray and neutron crystallography in a human aldose reductase-inhibitor complex. *Eur Biophys J 35*, 577-583.

[21] Yamaguchi, S., Kamikubo, H., Kurihara, K., Kuroki, R., Niimura, N., Shimizu, N., Yamazaki, Y., and Kataoka, M. (2009) Low-barrier hydrogen bond in photoactive yellow protein. *Proc Natl Acad Sci USA 106*, 440-444.

[22] Adachi, M., Ohhara, T., Kurihara, K., Tamada, T., Honjo, E., Okazaki, N., Arai, S., Shoyama, Y., Kimura, K., Matsumura, H., Sugiyama, S., Adachi, H., Takano, K., Mori, Y., Hidaka, K., Kimura, T., Hayashi, Y., Kiso, Y., and Kuroki, R. (2009) Structure of HIV-1 protease in complex with potent inhibitor KNI-272 determined by high-resolution X-ray and neutron crystallography. *Proc Natl Acad Sci USA 106*, 4641-4646.

[23] Ostermann, A., Tanaka, I., Engler, N., Niimura, N., and Parak, F. G. (2002) Hydrogen and deuterium in myoglobin as seen by a neutron structure determination at 1.5 angstrom resolution. *Biophys Chem 95*, 183-193.

[24] Reynolds, W. F., Peat, I. R., Freedman, M. H., and Lyerla, J. R. (1973) Determination of tautomeric form of imidazole ring of L-histidine in basic solution by C-13 magnetic-resonance spectroscopy. *J Am Chem Soc 95*, 328-331.

[25] Tang, T. H., and Cui, Y. P. (1996) A theoretical study of some X-H…π hydrogen-bonded complexes using the theory of atoms in molecules. *Can J Chem 74*, 1162-1170.

[26] Koch, U., and Popelier, P. L. A. (1995) Characterization of C-H-O Hydrogen-Bonds on the Basis of the Charge-Density. *J Phys Chem-US 99*, 9747-9754.

CHAPTER 5

Evaluating electronic structure methods for accurate calculation of ^{19}F chemical shifts in fluorohistidine analogues and other fluorinated amino acids.

Note: Most of the contents in this chapter are published in Journal of Computational Chemistry, 2017, 15, 2605.

5.1. INTRODUCTION

Fluorolabeling of biomolecules provides valuable information regarding protein structure, function, dynamics and ligand binding events, yet development of a robust computational protocol for reliable identification of fluorine chemical shifts is hampered by a lack of knowledge regarding the best electronic structure methods to accurately calculate fluorine shielding. Systematic studies of electronic structure methods for calculating fluorine chemical shifts are scarce. Isley *et al.* worked on a test set of fluorine-containing compounds before developing methodology for predicting fluorine chemical shifts in proteins, while commenting on the paucity of such studies.[1] Harding *et al.* examined methods for gas-phase fluorine chemical shifts in small molecules, focusing on high levels of theory.[2] The data set examined in this paper contains fluorinated amino acids and analogues for which fluorine chemical shifts have been determined experimentally. It surveys methods that are computationally tractable for these larger systems, in implicit solvent. In this work, computational methods for predicting absolute and *relative* fluorine chemical shifts are evaluated, studying fluorohistidine with known changes in fluorine shielding due to backbone, and protonation effects that may serve as proxies for changes in protein environment. For the ultimate goal of developing a

comprehensive and accurate method to predict fluorine chemical shifts in proteins, accurate determination of relative chemical shift (vs. absolute) is ostensibly the more necessary criterion. In addition, since amino acid chains represent many structures commonly seen in organic compounds, these results may serve as a guide to study fluorinated ligands and organic molecules.[3]

5.2. METHODS

The effects of protein environment can be propagated through the bonding framework,[4 5] or arise from differences in primary and secondary structure, different dielectric environments (interior *vs.* solvent-exposed), and site-specific differences in neighbors, close contacts, and hydrogen bonds. In this work, these effects are examined, respectively, by comparing "backbone effects" (changes in substituents on the amino acid side chain where C_α is bound); and changes in protonation state. Fig 5.1A. illustrates the systems studied within each of these categories. Experimental ^{19}F NMR chemical shifts were obtained from the literature for aqueous 2- and 4-fluorohistidine, 2- and 4-fluoro-(5-methyl)-imidazole, and 2- and 4-fluoroimidazole.[6]

The goal of this work is to identify efficient and accurate electronic structure calculations that can be carried out by typical computational resources. For that reason, Hartree-Fock, DFT, and MP2 methods (without computationally-expensive augmented correlation consistent basis sets) were considered. All the calculations were performed with the Gaussian 09 package [7] with Hartree-Fock (HF), MP2, and the density functional methods: BLYP, B3LYP,[8 9] BHandHLYP, M06, M06L, M062X,[10] PW91PW91,[11] PBE0,[12] PBEPBE,[11] ωB97X,[13] and ωB97XD.[14] While this set is not exhaustive, it does

contain both GGA, hybrid, and local functionals, commonly used methods, and those shown in other literature reports to provide reliable NMR chemical shifts. For instance, PBE1PBE (a.k.a. PBE0) has been shown to provide reliable values of ^{1}H, ^{13}C and ^{19}F chemical shifts in large and small organic compounds.[15] [16] Yu *et al.* showed BHandHLYP to provide reliable data for amino acids,[17] and our previous work indicated that BHandHLYP provided good prediction of ^{19}F chemical shifts in fluorohistidine isomers (much better than B3LYP).[18] This latter result hinted that a higher amount of Hartree-Fock exact exchange may improve accuracy for ^{19}F chemical shifts. Therefore, ωB97X, which has 100% long range Hartree-Fock exchange, was added to the set of methods, while ωB97XD allows for investigation of the extent to which dispersion affects calculated chemical shifts. When nuclear shielding is calculated with MP2 methods, values are provided in the output file for both the MP2 and SCF shielding. Both of these are considered in this work, indicated by "MP2" or "SCF" in tables and charts.

Two sizes of basis set were examined for each method, each containing both diffuse and polarization functions: 6-31+G* and 6-311++G(3df,2p). Solvation via the self consistent reaction field (SCRF) method with Conductor Polarized Continuum Model (CPCM)[19] and universal solvent model SMD[20] were used to perform the calculations, in order to see whether choice of implicit solvent model makes a difference in accuracy. The molecular structures were optimized with each level of method, and frequency calculations resulting in all positive frequencies indicated that the structures are at energy minima. NMR shielding tensors were calculated on optimized geometries, using the

GIAO (Gauge Independent Atomic Orbital) method.[21 22] For the chemical shift reference, $CFCl_3$ was used, and chemical shifts of the amino acids, $\delta_{amino\ acid}$, are provided as

$$\delta_{amino\ acid} = \sigma_{\mathrm{CFCl3}} - \sigma_{amino\ acid} + \delta_{ref.} \quad (1)$$

The absolute chemical shielding of $CFCl_3$ (σ_{CFCl3}) was calculated with each method (using optimized geometry at the same method), and the absolute chemical shielding of the amino acid, $\sigma_{amino\ acid}$, is provided in the GIAO calculations (at the same level of theory as the reference compound). In order to compare calculated values to experimental data, the molecules are referenced to hexafluorobenzene (C_6F_6) chemical shielding values were adjusted with δ_{ref}, which is the experimental difference between $CFCl_3$ and other reference compounds.[23] Error is evaluated throughout by taking the absolute value of the deviation of the calculated value from the experimental chemical shift, so that cancellation of positive and negative deviations from experimental values does not give rise to misleading conclusions regarding mean error. In the case of "relative" chemical shifts, in which changes in shielding of structural analogues or isomers with respect to a protonated structures are evaluated, the absolute deviation of the difference in chemical shifts (Δppm) is reported. All the geometries were prepared in GaussView 5,[24] using amino acid templates, modified by substitution of fluorine for hydrogen, and adjusted for specified stereochemistry. The different tautomeric states (τ and π) of 2-fluoro and 4-fluorohistidines and analogues were considered.[18 25 26] The fluorine chemical shift of the lowest energy tautomer is reported.

5.3. RESULTS

5.3.1. Absolute chemical shifts

Table 5.1 presents the average absolute deviation of fluorine chemical shifts (with respect to experimental values), determined for multiple methods (DFT, HF, MP2), with two basis sets (6-31+G* and 6-311++(3df,2p)) and two solvation models (SMD and CPCM). Also presented in Table 5.1 is the percentage of systems (12 total) in which the method calculates the chemical shift within 1 ppm of the experimental value. It can be seen in 5.1 that the top 14 methods are all combinations of basis sets and solvent models with three functionals: ωB97X, BHandHLYP, and ωB97XD and M06, which are the only methods to have average absolute errors of less than 5 ppm.

Table 5.1. Evaluation of electronic structure methods to predict chemical shifts (^{19}F) of fluorohistidine analogues in water: mean absolute error and percentage of systems for which calculated ^{19}F chemical shift is < 1.0 ppm.

Method ranked from lowest to highest error	Mean absolute error (ppm)	Percent systems error < 1.00 ppm
ωB97X/ 6-31+G* SMD,water	2.0	33.3
BHandHLYP/ 6-311++G(3df,2p) SMD,water	2.5	33.3
ωB97X/ 6-311++G(3df,2p) CPCM,water	2.8	8.3
BHandHLYP/ 6-31+G* SMD,water	3.0	25.0
ωB97X/ 6-311++G(3df,2p) SMD,water	3.3	16.7
ωB97XD/ 6-311++G(3df,2p) CPCM,water	3.3	25.0
M06/ 6-31+G* SMD,water	3.3	16.7
ωB97XD/ 6-31+G* CPCM,water	3.4	25.0
ωB97X/ 6-31+G* CPCM,water	3.7	8.3
M06/ 6-311++G(3df,2p) CPCM,water	4.2	16.7
ωB97XD/ 6-31+G* SMD,water	4.3	16.7
M062X/ 6-31+G* CPCM,water	4.6	0.0
M06/ 6-31+G* CPCM,water	4.7	0.0
BHandHLYP/ 6-311++G(3df,2p) CPCM,water	4.9	16.7
M062X/ 6-311++G(3df,2p) CPCM,water	5.1	0.0
BHandHLYP/ 6-31+G* CPCM,water	5.4	8.3
M06L/ 6-311++G(3df,2p) SMD,water	5.8	16.7
WB97XD/ 6-311++G(3df,2p) SMD,water	5.8	0.0
M06L/ 6-31+G* SMD,water	6.0	16.7
M06L/ 6-311++G(3df,2p) CPCM,water	6.2	8.3
PBE1PBE/ 6-311++G(3df,2p) CPCM,water	6.4	8.3
M06L/ 6-31+G* CPCM,water	6.4	0.0
PBE1PBE/ 6-31+G* CPCM,water	6.9	0.0
MP2/ 6-31+G* CPCM,water - MP2	7.1	8.3
M06/ 6-311++G(3df,2p) SMD,water	7.7	0.0
M062X/ 6-31+G* SMD,water	7.8	0.0
M062X/ 6-311++G(3df,2p) SMD,water	8.4	0.0
MP2/ 6-31+G* SMD,water - MP2	9.6	0.0

Table 5.1. (continued)

PBEPBE/ 6-31+G* SMD,water	10.2	0.0
PBEPBE/ 6-311++G(3df,2p) SMD,water	11.5	0.0
B3LYP/ 6-31+G* CPCM,water	15.8	0.0
B3LYP/ 6-311++G(3df,2p) CPCM,water	16.2	0.0
MP2/ 6-31+G* SMD,water- SCF	18.8	0.0
B3LYP/ 6-31+G* SMD,water	19.5	0.0
B3LYP/ 6-311++G(3df,2p) SMD,water	19.8	0.0
MP2/ 6-31+G* CPCM,water - SCF	21.5	0.0
HF/ 6-311++G(3df,2p) SMD,water	24.7	0.0
HF/ 6-31+G* SMD,water	25.6	0.0
HF/ 6-311++G(3df,2p) CPCM,water	27.0	0.0
BLYP/ 6-311++G(3df,2p) CPCM,water	27.6	8.3
HF/ 6-31+G* CPCM,water	27.8	0.0
BLYP/ 6-31+G* CPCM,water	28.7	0.0
PBEPBE/ 6-311++G(3df,2p) CPCM,water	30.5	0.0
PW91PW91/ 6-311++G(3df,2p) CPCM,water	30.8	0.0
PW91PW91/ 6-31+G* CPCM,water	32.6	0.0
PBEPBE/ 6-31+G* CPCM,water	32.6	0.0
PW91PW91/ 6-31+G* SMD,water	36.4	0.0
PW91PW91/ 6-311++G(3df,2p) SMD,water	37.5	0.0
BLYP/ 6-311++G(3df,2p) SMD,water	42.3	0.0
BLYP/ 6-31+G* SMD,water	43.0	0.0
PBE1PBE/ 6-311++G(3df,2p) SMD,water	54.7	0.0
PBE1PBE/ 6-31+G* SMD,water	57.0	0.0

5.3.2. Relative chemical shifts

In addition to determining the capability of computational methods to accurately calculate absolute values of ^{19}F chemical shifts in water, the ability to predict correct *relative* changes in fluorine shielding as a result of backbone effects and protonation effects was examined. The sets of comparative systems are illustrated in Fig 5.1A. The relative chemical shift is defined as $\Delta ppm = \delta_{amino\ acid} - \delta_{analogue}$. Here, the chemical shift of the reference amino acid structure is given by $\delta_{amino\ acid}$, while $\delta_{analogue}$ represents the chemical shift of the protonated structure or substituted side chain. Experimental values of Δppm are taken from published data, and evaluated error (absolute deviation) in calculated Δppm comes from reference to these experimental values.

Table 5.2. ranks the methods from lowest to highest average relative error, averaged over all fluorohistidine and their analogues. It also includes the percent time that the relative error is equal to or less than 0.5 ppm. The values of error in Δppm range from 0.02 for 4-fluorohistidine vs. 4-fluorohistidinium (M062X/6-311++g(3df,2p/SMD water) to 40.4 ppm for 4-fluorohistidine vs. 4-fluoroimidazole (BLYP/6-31+G*/CPCM water). The error in Δppm averaged across the 14 systems studied and all methods is 3.6 ppm. Taken altogether, this indicates a rather stringent requirement for accuracy and precision in computational methods that can reliably predict changes in shielding due to structural or environmental effects. Table 5.2. indicates that M062X/6-311++g(3df,2p) with SMD solvent model has the lowest error in calculating relative ^{19}F chemical shifts, predicting the relative error within 0.5 ppm for 29% of the systems.

Table 5.2. Ability of methods to calculate relative changes in ^{19}F chemical shifts among sets of fluorohistidine and analogues having structural changes (protonation, $C\alpha$ substitution).

Method ranked according to the lowest error in relative chemical shift values (Δppm)	Mean relative error (ppm)	Percent systems error < 0.5 ppm
MP2/ 6-31+G* CPCM,water -	1.83	14.29
M062X/ 6-31+G* SMD,water	2.04	7.14
M062X/ 6-311++G(3df,2p)	2.06	28.57
MP2/ 6-31+G* SMD,water - MP2	2.16	7.14
ωB97X/ 6-311++G(3df,2p)	2.26	14.29
ωB97X/ 6-31+G* SMD,water	2.28	0.00
ωB97XD/ 6-31+G* SMD,water	2.30	14.29
M06/ 6-311++G(3df,2p)	2.37	7.14
ωB97XD/ 6-311++G(3df,2p)	2.42	14.29
M06/ 6-31+G* SMD,water	2.47	0.00
BHandHLYP/ 6-31+G*	2.47	0.00
M06/ 6-31+G* CPCM,water	2.53	14.29
HF/ 6-311++G(3df,2p)	2.59	7.14
HF/ 6-31+G* SMD,water	2.59	14.29
B3LYP/ 6-311++G(3df,2p)	2.68	7.14
PBE1PBE/ 6-31+G* SMD,water	2.69	7.14
B3LYP/ 6-31+G* SMD,water	2.71	0.00
PBEPBE/ 6-311++G(3df,2p)	2.72	7.14
PBE1PBE/ 6-311++G(3df,2p)	2.74	0.00
PW91PW91/ 6-31+G*	2.97	0.00
PBEPBE/ 6-31+G* SMD,water	2.97	0.00
BHandHLYP/ 6-31+G*	3.07	7.14
M062X/ 6-31+G* CPCM,water	3.08	14.29
MP2/ 6-31+G* SMD,water- SCF	3.09	0.00
BHandHLYP/ 6-311++G(3df,2p)	3.09	14.29
ωB97X/ 6-31+G* CPCM,water	3.10	7.14
BLYP/ 6-31+G* SMD,water	3.11	0.00
HF/ 6-311++G(3df,2p)	3.11	14.29
HF/ 6-31+G* CPCM,water	3.15	14.29
ωB97X/ 6-311++G(3df,2p)	3.16	0.00
BLYP/ 6-311++G(3df,2p)	3.19	0.00
ωB97XD/ 6-31+G* CPCM,water	3.20	14.29
M062X/ 6-311++G(3df,2p)	3.22	7.14
PBE1PBE/ 6-311++G(3df,2p)	3.28	0.00
ωB97XD/ 6-311++G(3df,2p)	3.39	0.00
B3LYP/ 6-311++G(3df,2p)	3.47	0.00
PBEPBE/ 6-31+G* CPCM,water	3.48	0.00

Table 5.2. (continued)

PW91PW91/ 6-31+G*	3.50	0.00
M06/ 6-311++G(3df,2p)	3.53	0.00
PBEPBE/ 6-311++G(3df,2p)	3.54	7.14
M06L/ 6-311++G(3df,2p)	3.55	0.00
M06L/ 6-31+G* SMD,water	3.55	0.00
MP2/ 6-31+G* CPCM,water -	3.55	14.29
PW91PW91/ 6-311++G(3df,2p)	3.61	7.14
BHandHLYP/ 6-311++G(3df,2p)	3.66	0.00
PBE1PBE/ 6-31+G*	3.68	0.00
M06L/ 6-31+G* CPCM,water	3.98	7.14
M06L/ 6-311++G(3df,2p)	4.07	14.29
PW91PW91/ 6-311++G(3df,2p)	6.04	0.00
B3LYP/ 6-31+G* CPCM,water	7.94	7.14
BLYP/ 6-311++G(3df,2p)	15.66	0.00
BLYP/ 6-31+G* CPCM,water	16.03	0.00

5.3.3. Absolute and relative chemical shifts in other considered fluorinated molecules

Along with fluorohistidine and their analogues, our group extended the survey to 19 more fluorinated compounds (and their analogues) to generalize the best method in calculating absolute and relative fluorine chemical shifts. All the fluorinated systems considered for this survey have been shown in Fig. 5.1. In Fig. 5.2A, a pie graph indicates how many times each method provides the lowest error (in a set of 31 molecules) to calculate the absolute chemical shifts. It can be see that BHandHLYP/6-311G* method using CPCM solvent model and ωB97X/6-31+G* with SMD solvent model are nominally the best methods in predicting absolute ^{19}F NMR shift values, with each having the lowest error in 4 of the 31 systems studied (Table 5.3.). When considering only computational method and disregarding the specifics of basis set and solvation model, BHandHLYP, M06L, ωB97XD, and ωB97X functionals perform well for calculations of ^{19}F NMR chemical spectra. In addition to DFT methods, the MP2 method, which is a wave function theory (WFT) approach, performs well only in nonaromatic fluorinated systems in water, namely 3-fluoroproline, (2S, 3R)- and (2S, 3S)-4,4,4-trifluorovaline; (2S, 4R)- and (2S, 4S)-5,5,5-trifluoroleucine.

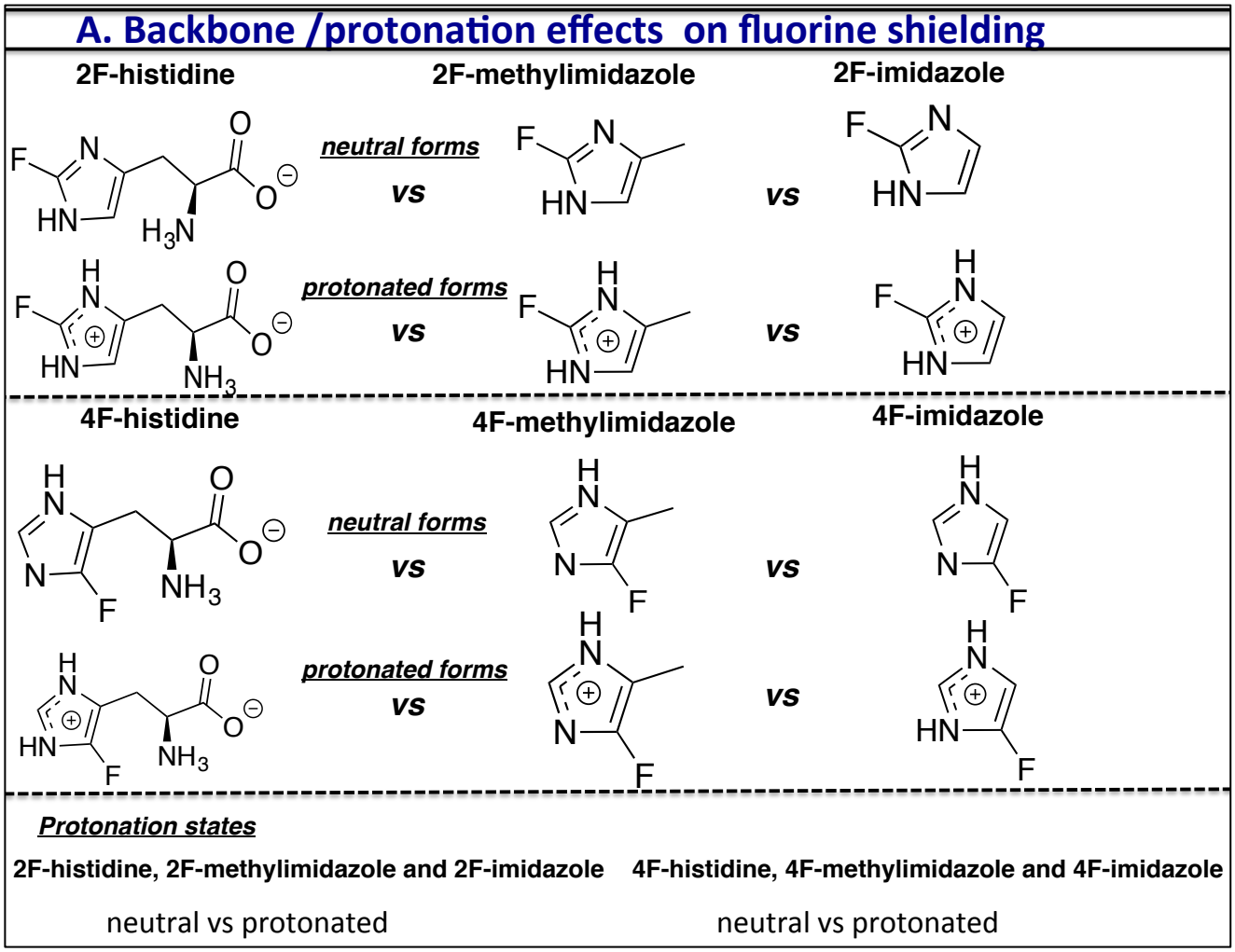


Figure 5.1. Fluorinated systems evaluated in this work. a) Backbone effects: fluorohistidine and analogues b) Backbone effects: fluorotryptophan and analogues, p-fluorophenylalanine and analogues. c) Isomeric effects: 3-fluoroproline, trifluoroleucine, and trifluorovaline. d) 3-fluorovaline and trifluoromethionine

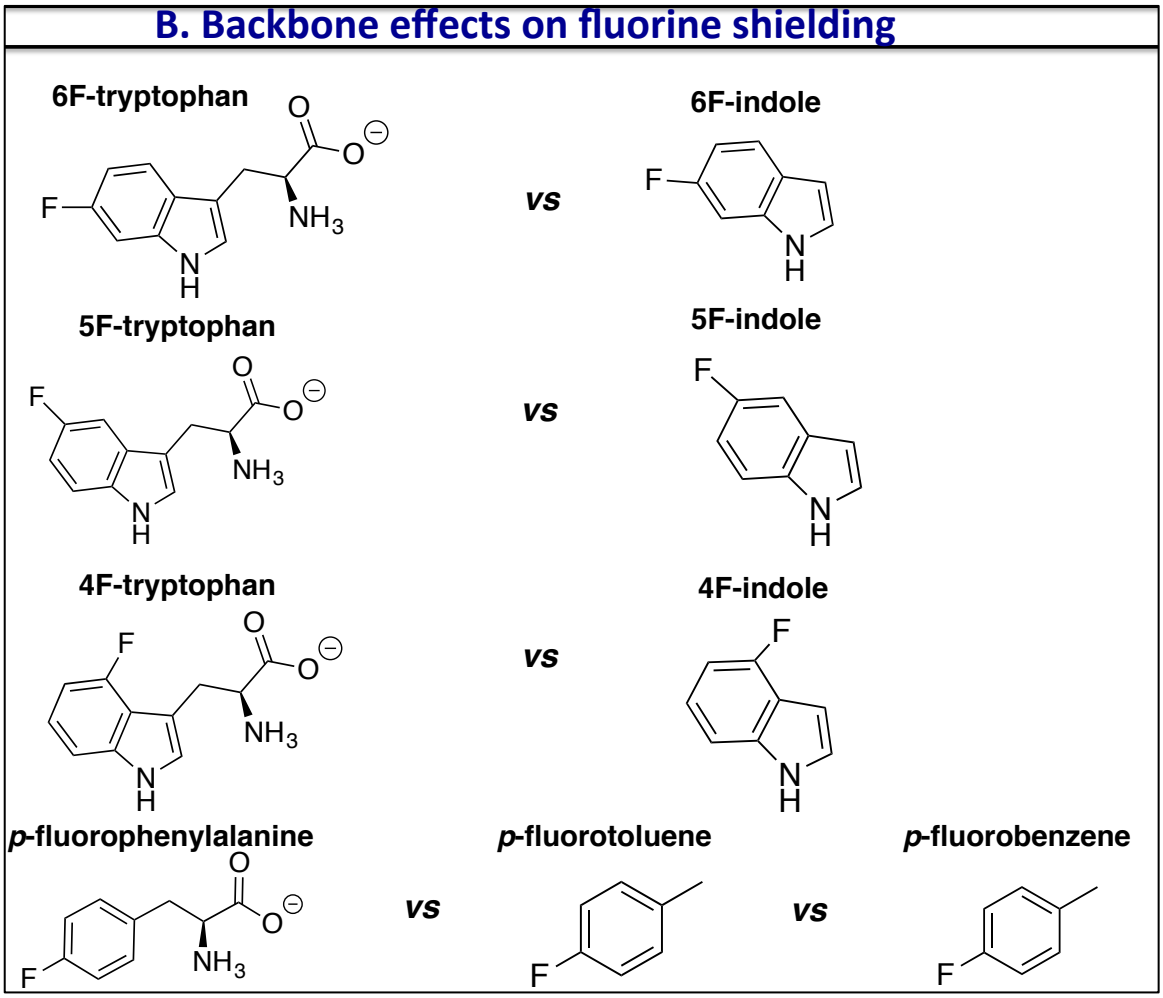


Figure 5.1. (continued)

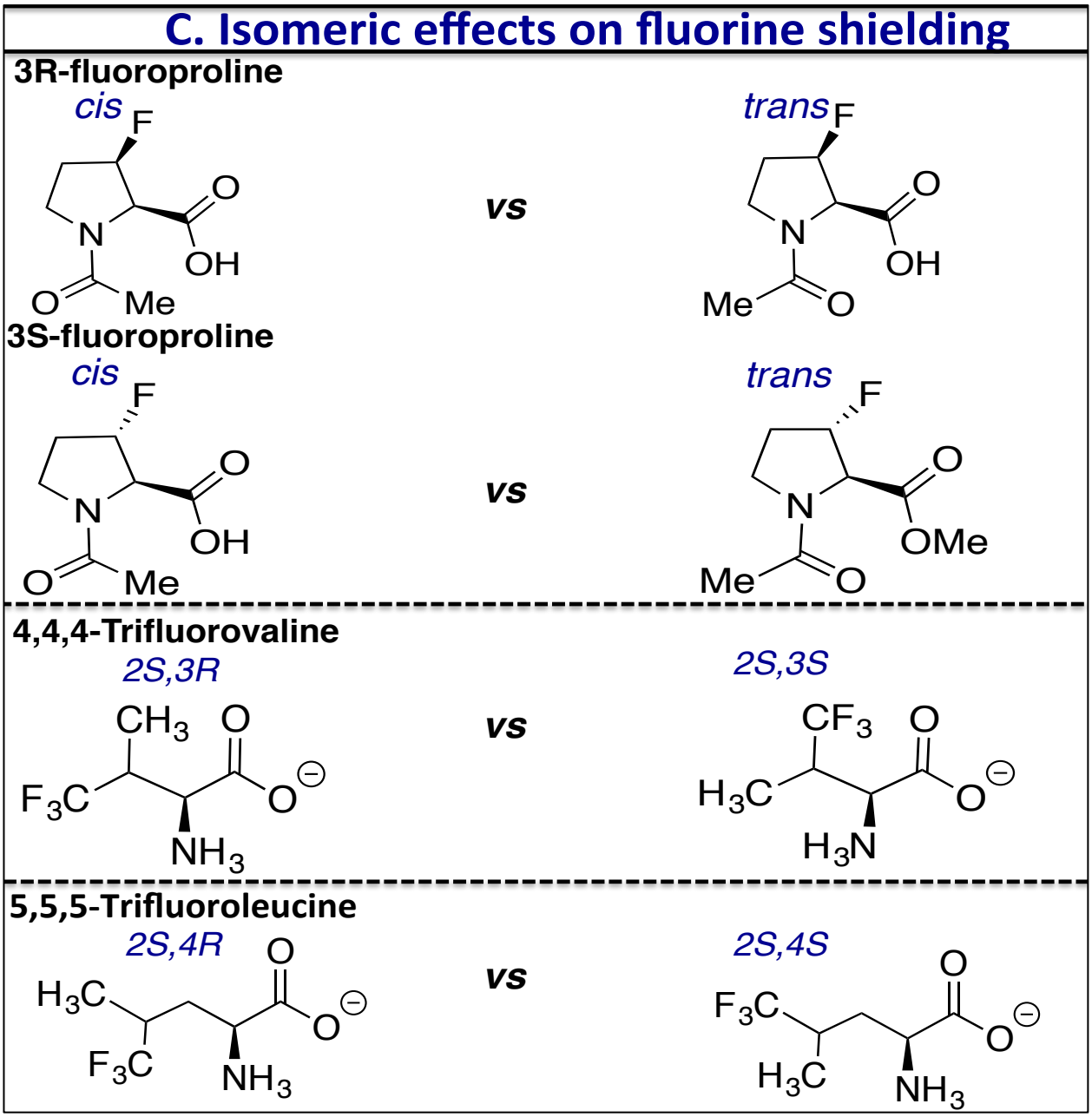


Figure 5.1. (continued)

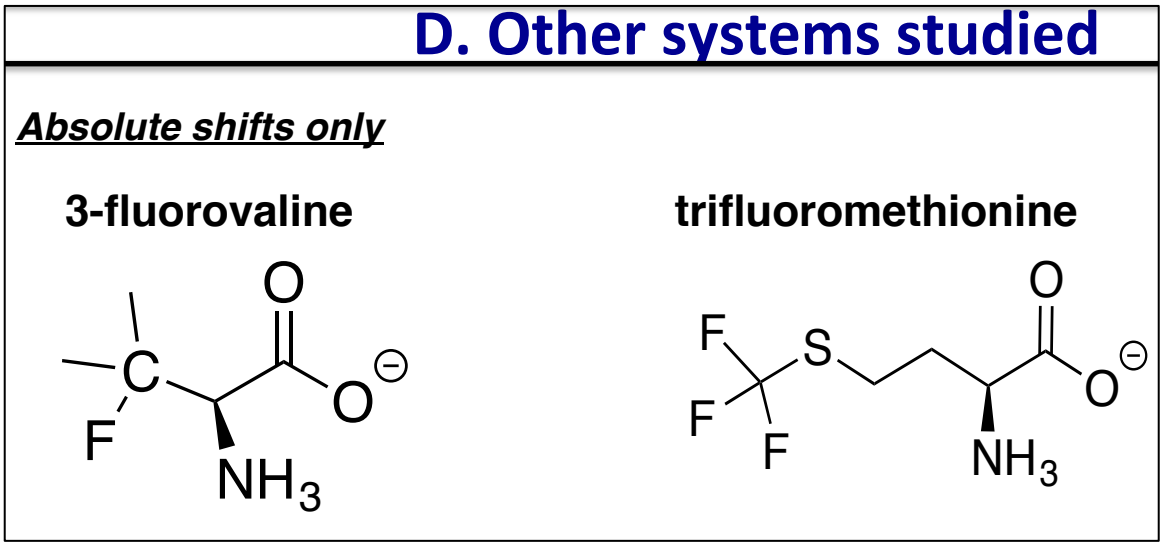


Figure 5.1. (continued)

Overall, Hartree-Fock (HF) methods, namely, HF/6–31 1 G* with either SMD or CPCM solvation, offer high accuracy and low computational cost. This finding should be advantageous for NMR chemical shifts prediction in large biomolecular systems. The relative chemical shift changes in fluorine shielding due to backbone effects, isomeric effects and protonation effects were studied for all the systems. Relative errors for all the considered fluorinated systems are reported in Table 5.4. Table 5.4 indicates that MP2/6–311G* with CPCM solvent model has the lowest error in calculating relative ^{19}F chemical shifts, predicting the relative error within 0.5 ppm for 22% of the systems (SMD model also has a low average error, but lower accuracy as well, predicting a ppm within 0.5 ppm for only 9% of the systems).

Table 5.3. Evaluation of electronic structure methods to predict chemical shifts (^{19}F) of fluorinated amino acids in water: mean absolute error and percentage of systems for which calculated ^{19}F chemical shift is < 1.0 ppm.

Method ranked from lowest to highest error	Mean absolute error (ppm)	Percent systems error < 1.00 ppm
ωB97X/ 6-31+G* SMD,water	2.68	29.03
BHandHLYP/ 6-311++G(3df,2p) SMD,water	3.01	35.48
BHandHLYP/ 6-31+G* SMD,water	3.09	16.13
ωB97XD/ 6-31+G* CPCM,water	3.18	12.90
BHandHLYP/ 6-31+G* CPCM,water	3.59	29.03
ωB97X/ 6-311++G(3df,2p) CPCM,water	3.85	19.35
ωB97X/ 6-31+G* CPCM,water	3.89	22.58
BHandHLYP/ 6-311++G(3df,2p) CPCM,water	3.92	16.13
ωB97XD/ 6-31+G* SMD,water	4.21	6.45
ωB97X/ 6-311++G(3df,2p) SMD,water	4.48	9.68
ωB97XD/ 6-311++G(3df,2p) CPCM,water	4.81	9.68
M06/ 6-31+G* SMD,water	5.47	9.68
MP2/ 6-31+G* CPCM,water - MP2	5.93	12.90
M062X/ 6-31+G* CPCM,water	6.19	3.23
M06/ 6-31+G* CPCM,water	6.23	3.23
M06L/ 6-311++G(3df,2p) SMD,water	6.46	25.81
M06L/ 6-311++G(3df,2p) CPCM,water	6.71	6.45
ωB97XD/ 6-311++G(3df,2p) SMD,water	7.11	0.00
MP2/ 6-31+G* SMD,water - MP2	7.39	12.90
M062X/ 6-311++G(3df,2p) CPCM,water	8.15	9.68
M062X/ 6-31+G* SMD,water	8.15	0.00
M06/ 6-311++G(3df,2p) CPCM,water	8.49	3.23
M06L/ 6-31+G* SMD,water	8.62	6.45

Table 5.3. (continued)

M06L/ 6-31+G* CPCM,water	8.95	0.00
PBE0/ 6-311++G(3df,2p) CPCM,water	10.39	3.23
M062X/ 6-311++G(3df,2p) SMD,water	10.48	0.00
M06/ 6-311++G(3df,2p) SMD,water	10.94	0.00
PBE0/ 6-31+G* CPCM,water	11.14	0.00
MP2/ 6-31+G* SMD,water- SCF	12.48	0.00
MP2/ 6-31+G* CPCM,water - SCF	14.06	0.00
B3LYP/ 6-31+G* CPCM,water	16.03	0.00
B3LYP/ 6-311++G(3df,2p) CPCM,water	18.25	0.00
B3LYP/ 6-31+G* SMD,water	18.29	0.00
HF/ 6-311++G(3df,2p) SMD,water	18.48	0.00
B3LYP/ 6-311++G(3df,2p) SMD,water	19.37	0.00
HF/ 6-311++G(3df,2p) CPCM,water	20.00	0.00
HF/ 6-31+G* SMD,water	20.22	0.00
HF/ 6-31+G* CPCM,water	21.64	0.00
PBEPBE/ 6-31+G* SMD,water	22.37	0.00
PBEPBE/ 6-311++G(3df,2p) SMD,water	22.54	0.00
PBE0/ 6-311++G(3df,2p) SMD,water	28.00	0.00
PBE0/ 6-31+G* SMD,water	28.51	0.00
PBEPBE/ 6-311++G(3df,2p) CPCM,water	28.87	0.00
PBEPBE/ 6-31+G* CPCM,water	29.99	0.00
BLYP/ 6-311++G(3df,2p) CPCM,water	31.72	3.23
PW91PW91/ 6-31+G* SMD,water	32.85	0.00
BLYP/ 6-31+G* CPCM,water	32.94	0.00
PW91PW91/ 6-311++G(3df,2p) SMD,water	33.11	0.00
PW91PW91/ 6-311++G(3df,2p) CPCM,water	35.50	0.00
PW91PW91/ 6-31+G* CPCM,water	36.70	0.00
BLYP/ 6-311++G(3df,2p) SMD,water	38.63	0.00
BLYP/ 6-31+G* SMD,water	39.03	0.00

Table 5. 4. Ability of methods to calculate relative changes in ^{19}F chemical shifts among sets of amino acids and analogues having structural changes (protonation, C_α substitution, isomers).

Methods ranked according to the lowest error in relative chemical shift values (Δppm)	Mean relative error (ppm)	Percent systems with error < 0.5 ppm
MP2/ 6-31+G* CPCM,water - MP2	1.84	21.74
MP2/ 6-31+G* SMD,water - MP2	1.88	8.70
M062X/ 6-31+G* SMD,water	2.13	13.04
ωB97X/ 6-31+G* SMD,water	2.19	8.70
ωB97X/ 6-311++G(3df,2p) SMD,water	2.30	13.04
M062X/ 6-311++G(3df,2p) SMD,water	2.31	21.74
BHandHLYP/ 6-31+G* SMD,water	2.32	4.35
ωB97XD/ 6-31+G* SMD,water	2.33	13.04
HF/ 6-31+G* SMD,water	2.33	17.39
ωB97XD/ 6-311++G(3df,2p) SMD,water	2.36	8.70
HF/ 6-311++G(3df,2p) SMD,water	2.39	4.35
M06/ 6-311++G(3df,2p) SMD,water	2.41	13.04
B3LYP/ 6-31+G* SMD,water	2.60	0.00
PBE0/ 6-31+G* SMD,water	2.62	8.70
PBE0/ 6-311++G(3df,2p) SMD,water	2.64	0.00
MP2/ 6-31+G* SMD,water- SCF	2.66	8.70
PBEPBE/ 6-311++G(3df,2p) SMD,water	2.67	8.70
BHandHLYP/ 6-311++G(3df,2p) SMD,water	2.69	8.70
ωB97X/ 6-31+G* CPCM,water	2.70	13.04
B3LYP/ 6-311++G(3df,2p) SMD,water	2.71	4.35
HF/ 6-31+G* CPCM,water	2.72	17.39
M062X/ 6-31+G* CPCM,water	2.75	17.39
BHandHLYP/ 6-31+G* CPCM,water	2.77	13.04
HF/ 6-311++G(3df,2p) CPCM,water	2.78	13.04
ωB97X/ 6-311++G(3df,2p) CPCM,water	2.78	4.35
PBE0/ 6-311++G(3df,2p) CPCM,water	2.83	8.70
MP2/ 6-31+G* CPCM,water - SCF	2.92	8.70
PBEPBE/ 6-31+G* SMD,water	2.92	4.35
M06L/ 6-311++G(3df,2p) SMD,water	2.94	13.04
M062X/ 6-311++G(3df,2p) CPCM,water	2.95	8.70

Table 5. 4. (continued)

BLYP/ 6-311++G(3df,2p) SMD,water	2.96	0.00
PW91PW91/ 6-31+G* SMD,water	3.00	4.35
BHandHLYP/ 6-311++G(3df,2p) CPCM,water	3.05	8.70
BLYP/ 6-31+G* SMD,water	3.08	0.00
PBEPBE/ 6-31+G* CPCM,water	3.08	4.35
PBE0/ 6-31+G* CPCM,water	3.09	0.00
PBEPBE/ 6-311++G(3df,2p) CPCM,water	3.12	13.04
M06/ 6-311++G(3df,2p) CPCM,water	3.16	0.00
PW91PW91/ 6-31+G* CPCM,water	3.24	4.35
PW91PW91/ 6-311++G(3df,2p) CPCM,water	3.26	8.70
ωB97XD/ 6-31+G* CPCM,water	3.33	17.39
ωB97XD/ 6-311++G(3df,2p) CPCM,water	3.52	0.00
M06/ 6-31+G* SMD,water	3.55	0.00
M06/ 6-31+G* CPCM,water	3.57	8.70
M06L/ 6-311++G(3df,2p) CPCM,water	3.79	13.04
M06L/ 6-31+G* SMD,water	4.24	4.35
PW91PW91/ 6-311++G(3df,2p) SMD,water	4.71	0.00
M06L/ 6-31+G* CPCM,water	4.98	4.35
B3LYP/ 6-31+G* CPCM,water	5.69	8.70
B3LYP/ 6-311++G(3df,2p) CPCM,water	6.08	4.35
BLYP/ 6-311++G(3df,2p) CPCM,water	10.50	8.70
BLYP/ 6-31+G* CPCM,water	13.73	0.00

The number of times each method appeared as the best method (with the most accurate calculation, represented by the least relative error) is shown in Fig 5.2B. According to Fig. 5.2B, when neglecting the details of basis set and solvation model, both MP2 and PBE0 methods have the highest number of most accurate ppm values. Overall, MP2/6–311G*/CPCM solvent performed best, with most accurate values calculated for 4 of the 23 comparative systems. Fig. 5.3 indicates how many times each method is within the top three methods for relative chemical shift calculations (Δppm values), based on lowest relative error. Fig. 5.3 also confirms that MP2 methods generally provide accurate

relative changes in ^{19}F chemical shifts of fluorinated systems in water. Simultaneously, M062X method also emerges many times among the best three methods. Other good performers include xB97XD, PBE0, HF, M06 and PBEPBE.

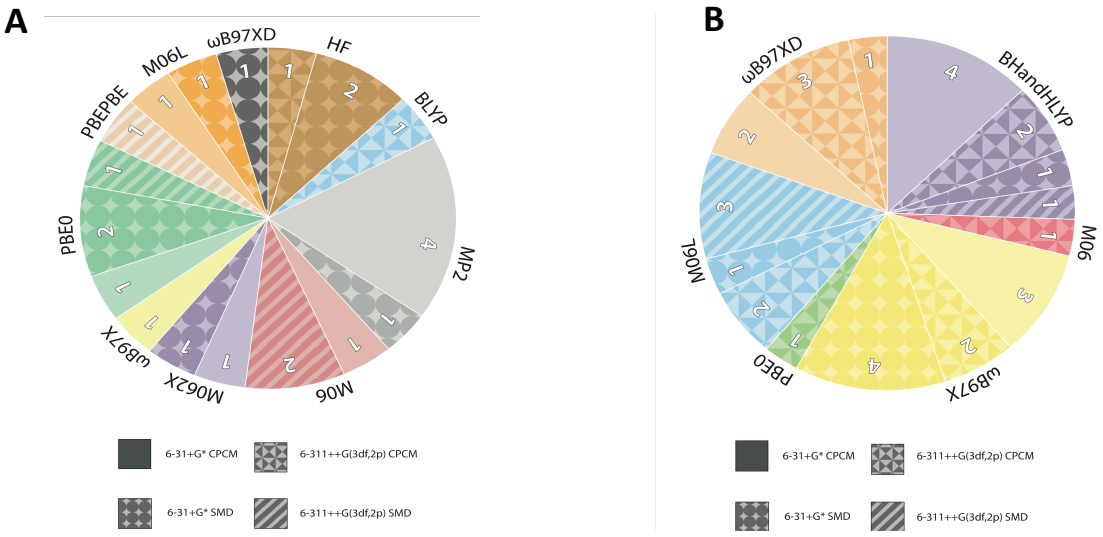


Figure 5.2. A) Number of molecules (out of 31) for which a method provides the least absolute error value in chemical shift calculations. B) Number of comparative systems (out of 23 total) for which a method provides the lowest error value in relative chemical shift (Δppm value).

When examining each fluorinated system solely considering lowest relative error, MP2, PBE0, M062X, HF, and M06L methods appear the most reliable in procuring correct relative ^{19}F chemical shifts for aromatic fluorinated compounds. The non-aromatic systems have a different behavior. In fluoroproline isomers, where cis-trans isomerization effects are considered, ωB97X and BHandHLYP methods also perform well in addition to MP2, M062X, and HF methods. But for the trifluorovaline systems, M062X method does not have good performance. When considering fluorinated systems independently, it is evident that SMD solvent model is better in most cases for calculating relative shifts. This is evident in data for fluorohistidine, fluoroimidazole, and fluoro-(5-methyl)-imidazole protonation effects, and fluorotryptophan versus fluoroindole backbone effects. It is also noticeable that the 6–31+G* basis set is the best performer, since it appears to correctly calculate changes in shielding in many comparative systems.

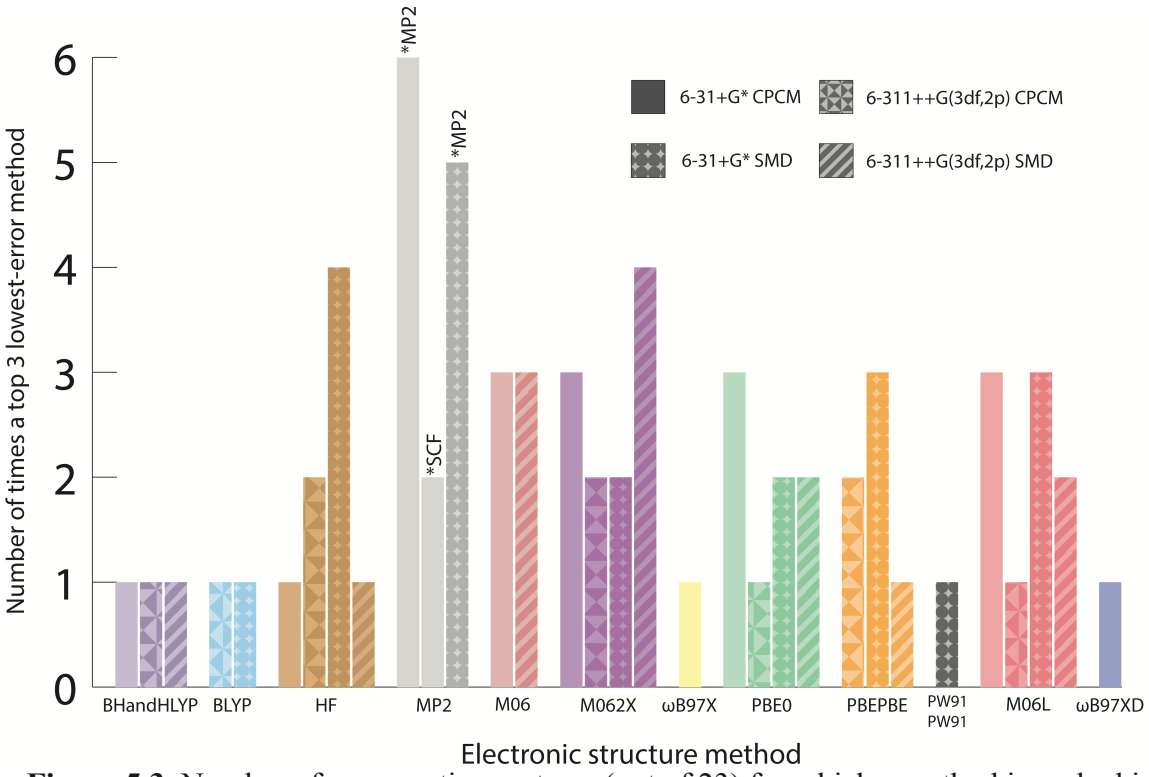


Figure 5.3. Number of comparative systems (out of 23) for which a method is ranked in the top three most accurate relative (Δppm) ^{19}F chemical shifts.

The best functionals to compute absolute fluorine chemical shifts for fluorohistidine and their analogues include BHandHLYP, ωB97X, M06, and ωB97XD. In general for all the fluorinated systems (31) considered above, along with BHandHLYP, ωB97X, ωB97XD functionals, the M06L functional also stands in top place in computing absolute chemical shifts. BHandHLYP has 50% HF exchange and 50% B88 GGA exchange. ωB97X and ωB97XD are range separated hybrid functionals that split the exact exchange contribution into short-range and long-range. In ωB97X, the exchange correlation function is composed of 16% short-ranged HF exchange, 100% long-ranged HF exchange and a value of 0.3 for ω. It should be noted that the smaller the ω value, the

longer the range of the short-ranged operator will be.[13] In ωB97XD, the long-range exchange has a longer length scale (value of 0.2 for ω); however, the amount of exact exchange in the short-range component is increased to 22%. Note the comparative performance of BHandHLYP over BLYP and B3LYP. As stated before, BHandHLYP has 50% HF exchange, whereas B3LYP has only 20% HF exchange and BLYP is not a hybrid functional (no HF exchange). The comparison across the series of functionals with Becke exchange and Lee-Yang-Parr GGA correlation may indicate that accurate shielding calculations of fluorine require higher amounts of HF exchange

The best performing functionals for relative changes in fluorine nuclear shielding (evaluated over structural analogues and isomers) are the Minnesota functionals M062X and M06. When considering just the fluorohistidine and their analogues M062X, ωB97X and MP2 stand in the top row. M062X is one of the recently developed Minnesota functionals with double the amount of nonlocal exchange. It is a global hybrid meta GGA with 54% HF exchange.[10] The M06 functional, however, has a lower amount of HF exchange (27%). M062x performed particularly well with fluorohistidine analogues and 4F-tryptophan versus 4F-indole, and performed poorly with trifluorovaline, whereas M06 did well calculating ^{19}F chemical shifts of trifluorovaline. Our dataset is admittedly biased toward fluorinated aromatic amino acids, which have been commoly used in protein fluorolabeling studies. It is possible that our dataset indicates that better results are obtained for fluorinated aromatic compounds when functionals with higher amounts of HF exchange are used. A summary of recommended methods to calculate absolute ^{19}F chemical shifts for each family of fluorinated compounds is depicted in Fig. 5.4.

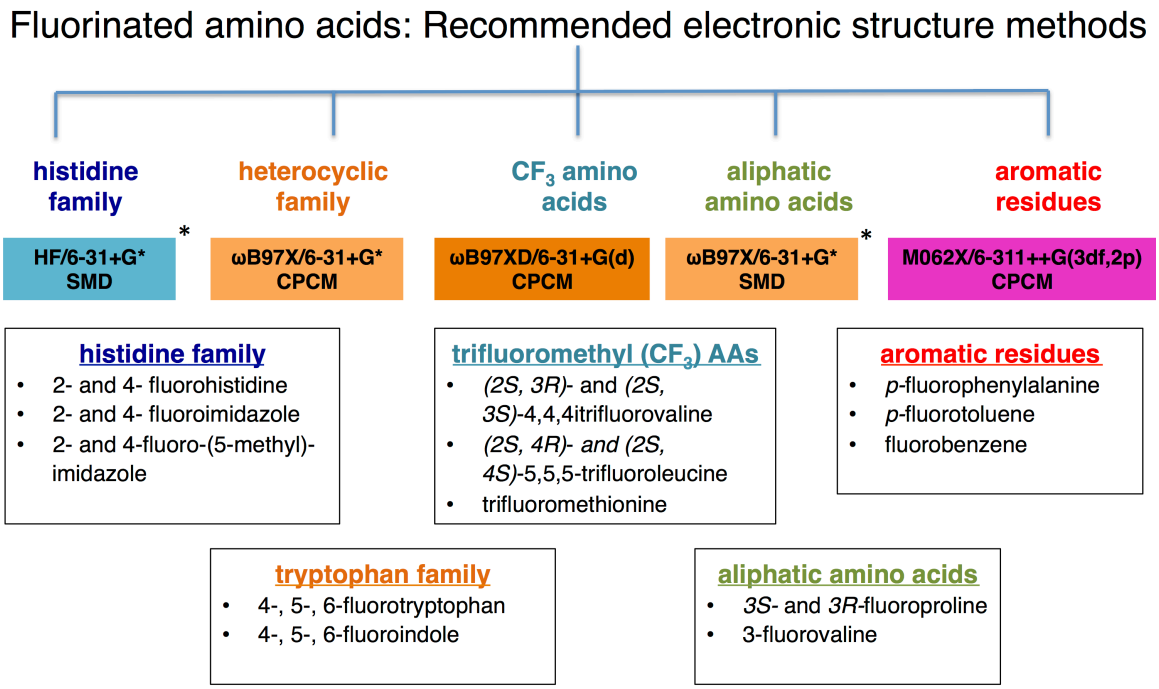


Figure 5.4. Flowchart showing recommended methods for different types of fluorinated families.

5.4. CONCLUSIONS

When considering future work on the fluorinated amino acids surveyed here, Fig. 5.4 provides general recommendations. In general, absolute chemical shifts of fluorine are most accurate using ωB97X and BHandHLYP functionals. Relative chemical shifts, that is, when examining effects of protein or structural changes on shielding, are most accurately calculated by MP2 calculations and the M062X functional. Best performance for relative changes in ^{19}F shielding is generally obtained with a small basis set (6–31+G*) and the SMD implicit model (in aqueous solutions).

REFERENCES

[1] Isley, W. C., Urick, A. K., Pomerantz, W. C. K., and Cramer, C. J. (2016) Prediction of F-19 NMR chemical shifts in labeled proteins: computational protocol and case study, *Mol Pharm 13*, 2376-2386.

[2] Harding, M. E., Lenhart, M., Auer, A. A., and Gauss, J. (2008) Quantitative prediction of gas-phase 19F nuclear magnetic shielding constants, *J Chem Phys 128*, 244111.

[3] Vulpetti, A., Hommel, U., Landrum, G., Lewis, R., and Dalvit, C. (2009) Design and NMR-based screening of lef, a library of chemical fragments with different local environment of fluorine, *J Am Chem Soc 131*, 12949-12959.

[4] Oldfield, E. (2002) Chemical shifts in amino acids, peptides, and proteins: From quantum chemistry to drug design, *Annu Rev Phys Chem 53*, 349-378.

[5] Dedios, A. C., Pearson, J. G., and Oldfield, E. (1993) Secondary and tertiary structural effects on protein NMR chemical-shifts - an abinitio approach, *Science 260*, 1491-1496.

[6] Yeh, H. J. C., Kirk, K. L., Cohen, L. A., and Cohen, J. S. (1975) F-19 and H-1 nuclear magnetic-resonance studies of ring-fluorinated imidazoles and histidines, *J Chem Soc Perk T 2*, 928-934.

[7] 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., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G.; Sonnenberg, J. L., *et al.* . (2009) Gaussian, Inc., Gaussian 09. Wallingford, CT.

[8] Becke, A. D. (1993) A new mixing of hartree-fock and local density-functional theories, *J Chem Phys 98*, 1372-1377.

[9] Stephens, P. J., Devlin, F. J., Chabalowski, C. F., and Frisch, M. J. (1994) Ab-initio calculation of vibrational absorption and circular-dichroism spectra using density-functional force-fields, *J Phys Chem-US 98*, 11623-11627.

[10] Zhao, Y., and Truhlar, D. G. (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, *Theor Chem Acc 120*, 215-241.

[11] Perdew, J. P., and Wang, Y. (1992) Accurate and simple analytic representation of the electron-gas correlation energy, *Phys Rev B Condens Matter 45*, 13244-13249.

[12] Adamo, C., and Barone, V. (1999) Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J Chem Phys 110*, 6158-6170.

[13] Chai, J. D., and Head-Gordon, M. (2008) Systematic optimization of long-range corrected hybrid density functionals, *J Chem Phys 128*, 084106.

[14] Chai, J. D., and Head-Gordon, M. (2008) Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections, *Phys Chem Chem Phys 10*, 6615-6620.

[15] Pierens, G. K. (2014) 1H and 13C NMR scaling factors for the calculation of chemical shifts in commonly used solvents using density functional theory, *J Comput Chem 35*, 1388-1394.

[16] Isley, W. C., 3rd, Urick, A. K., Pomerantz, W. C., and Cramer, C. J. (2016) Prediction of (19)F NMR Chemical Shifts in Labeled Proteins: Computational Protocol and Case Study, *Mol Pharm 13*, 2376-2386.

[17] Yu, W. B., Liang, L., Lin, Z. J., Ling, S. L., Haranczyk, M., and Gutowski, M. (2009) Comparison of some representative density functional theory and wave function theory methods for the studies of amino acids, *J Comput Chem 30*, 589-600.

[18] Kasireddy, C., Bann, J. G., and Mitchell-Koch, K. R. (2015) Demystifying fluorine chemical shifts: electronic structure calculations address origins of seemingly anomalous F-19-NMR spectra of fluorohistidine isomers and analogues, *Phys Chem Chem Phys 17*, 30606-30612.

[19] Cossi, M., and Barone, V. (2001) Time-dependent density functional theory for molecules in liquid solutions, *J Chem Phys 115*, 4708-4717.

[20] Marenich, A. V., Cramer, C. J., and Truhlar, D. G. (2009) 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 113*, 6378-6396.

[21] Ditchfield, R. (1974) Self-Consistent Perturbation-Theory of Diamagnetism .1. Gauge-Invariant Lcao Method for Nmr Chemical-Shifts, *Mol Phys 27*, 789-807.

[22] Wolinski, K., Hinton, J. F., and Pulay, P. (1990) Efficient implementation of the gauge-independent atomic orbital method for NMR chemical-shift calculations, *J Am Chem Soc 112*, 8251-8260.

[23] Reich, H. J. Fluorine NMR data, http://www.chem.wisc.edu/areas/reich/nmr/11-f-data.htm. [cited 18 October 2017].

[24] Dennington, R., Keith, T. A., and Millam, J. M. (2009) Semichem Inc. Version 5.0, Shawnee Mission KS.

[25] Kasireddy, C., Ellis, J. M., Bann, J. G., and Mitchell-Koch, K. R. (2016) Tautomeric stabilities of 4-fluorohistidine shed new light on mechanistic experiments with labeled ribonuclease A, *Chem Phys Lett 666*, 58-61.

[26] Kasireddy, C., Ellis, J. M., Bann, J. G., and Mitchell-Koch, K. R. (2017) The Biophysical Probes 2-fluorohistidine and 4-fluorohistidine: Spectroscopic Signatures and Molecular Properties, *Sci Rep 7*, 42651.