

I. AMHB: (ANTI)AROMATICITY-MODULATED HYDROGEN BONDING
II. EVALUATION OF IMPLICIT SOLVATION MODELS FOR PREDICTING HYDROGEN BOND
FREE ENERGIES

By

Tayeb Kakeshpour

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry—Doctor of Philosophy

2019

ABSTRACT

I. AMHB: (ANTI)AROMATICITY-MODULATED HYDROGEN BONDING II. EVALUATION OF IMPLICIT SOLVATION MODELS FOR PREDICTING HYDROGEN BOND FREE ENERGIES

By

Tayeb Kakeshpour

My doctoral research under Professor James E. Jackson focused on hydrogen bonding (H-bonding) using physical organic chemistry tools. In the first chapter, I present how I used quantum chemical simulations, synthetic organic chemistry, NMR spectroscopy, and X-ray crystallography to provide robust theoretical and experimental evidence for an interplay between (anti)aromaticity and H-bond strength of heterocycles, a concept that we dubbed (Anti)aromaticity-Modulated Hydrogen Bonding (AMHB). In the second chapter, I used accurately measured hydrogen bond energies for a range of substrates and solvents to evaluate the performance of implicit solvation models in combination with density functional methods for predicting solution phase hydrogen bond energies. This benchmark study provides useful guidelines for *a priori* modeling of hydrogen bonding-based designs. Coordinates of the optimized geometries and crystal structures are provided as supplementary materials.

Copyright by
TAYEB KAKESHPOUR
2019

ACKNOWLEDGEMENTS

My doctoral study was a wonderful experience which would not have been possible without the love and support of many people to whom my gratitude cannot be fully conveyed in words. Beyond giving me the freedom to experience and explore various aspects of my interests from programming to synthetic chemistry himself, Dr. Jackson and his family have been a second family to me during the seven years I have been far from my own family. In addition, I want to thank our collaborators Dr. Judy I. Wu, Dr. Xuefei Huang, Dr. Babak Borhan, Dr. Jetze J. Tepe, Dr. Daniel C. Whitehead, and Dr. Carl L. Lira and my committee members Dr. Benjamin L. Levine and Dr. John L. McCracken who have been supportive of my research during these years. My research was made possible by the technical support and patience of many MSU staff members in explaining various tools to me. Among those I also want to express my gratitude to Dr. Daniel Holmes, Dr. Richard J. Staples, Dr. Chun-Min Chang, Paul Reed, and Robert Rasico. I also would like to express my appreciation towards the Jackson group members and the numerous high school and undergraduate students with whom I have worked with. Furthermore, I want to thank Hadi Gholami for both teaching me organic synthesis and being a supportive friend during my time at MSU. Lastly, I want to thank my parents, my sister, and my brother for their endless love and support.

TABLE OF CONTENTS

LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
LIST OF SCHEMES.....	x
KEY TO SYMBOLS AND ABBREVIATIONS.....	xi
CHAPTER I: AMHB: (ANTI)AROMATICITY-MODULATED HYDROGEN BONDING.....	1
I.1. Introduction.....	1
I.2. A broad in silico survey: generality and robustness of AMHB.....	6
I.2.1. Designing examples to explore the generality of AMHB.....	8
I.2.2. Nucleus-independent chemical shift.....	10
I.2.3. Energetic calculations.....	14
I.2.4. Aromaticity gain strengthens H-bonds.....	15
I.2.5. Aromaticity loss weakens H-bonds.....	16
I.2.6. Antiaromaticity increase weakens H-bonds.....	17
I.2.7. Antiaromaticity relief strengthens H-bonds.....	18
I.3. Tuning association energies via AMHB.....	20
I.4. Remote AMHB effects in fused rings.....	23
I.5. Experimental measurement of the AMHB effect.....	25
I.5.1. Energetic, geometric, and magnetic evidence for the interplay of aromaticity and H-bonding.....	35
I.5.2. Charge polarization vs. AMHB: Chemical shifts.....	37
I.5.3. Hybridization vs. AMHB: Energetics.....	38
I.5.4. Dipole-dipole interactions vs. AMHB: Energetics.....	39
I.6. Resonance-assisted H-bonding (RAHB) and AMHB.....	41
I.7. AMHB in redox systems.....	43
I.8. Conclusions.....	49
I.9. Computational and experimental details.....	49
I.9.1. Computational details.....	49
I.9.2. Details of NMR measurements.....	51
I.9.3. Synthesis of compounds 1, 1', 2 and 2'.....	52
CHAPTER II: EVALUATION OF IMPLICIT SOLVATION MODELS FOR PREDICTING H-BOND FREE ENERGIES.....	58
II.1. Introduction.....	58
II.2. Experimental reference H-bond energies.....	60
II.3. CCSD(T)/CBS: a reference for gas-phase H-bond energies.....	62

II.4. The performance of DFT methods vs CCSD(T)/CBS in the gas-phase.....	63
II.5. Comparison of theory and experiment.....	68
II.6. Conclusions.....	80
II.7. Experimental H-bond measurements.....	80
APPENDICES.....	82
APPENDIX A: Chemical shifts used for obtaining associations.....	83
APPENDIX B: Plots of fits for obtaining thermodynamic values from NMR spectra.....	98
APPENDIX C: Stacked ¹ H NMR spectra used for obtaining dimerization energies.....	105
APPENDIX D: NMR and IR spectra of compounds 1, 1', 2 and 2'.....	141
APPENDIX E: Wavefunction-based energies.....	162
APPENDIX F: Calculated NICS values.....	179
APPENDIX G: The Python code for insertion of NICS probes.....	191
BIBLIOGRAPHY.....	203

LIST OF TABLES

Table I-1. Thermodynamic and chemical shift values (at 298.15 K and vs tetramethyl silane) obtained from NMR spectroscopy.....	33
Table I-2. Calculated dipole-dipole interaction energies (E_{d-d}). This table was adapted from <i>Angewandte Chemie International Edition</i> , 2017 , <i>56</i> , 9842. Copyright 2017 John Wiley and sons.	40
Table II-1. Measured H-bond thermodynamic values and extrapolated monomer and dimer chemical shifts.....	61
Table II-2. Calculated H-bond electronic energies at CCSD(T)/CBS and a few less correlated wavefunction-based approximations.....	63
Table II-3. Mean absolute errors (MAE) of calculated electronic energies of H-bonds using various levels of theory vs CCSD(T)/CBS. The energies are not counterpoise corrected.....	65
Table II-4. Extracted parameters from NMR measurements at 298.15 K.....	81

LIST OF FIGURES

- Figure I-1.** The structures of cycloheptatrienylium and cyclopentadienylium ions. This figure was recreated from Doering's paper mentioned above.¹⁹ 4
- Figure I-2.** Intramolecular H-bond interaction in Stipitatic Acid. This figure was recreated from Dewar's paper.²⁰ 5
- Figure I-3.** Archetype examples of AMHB. H-bond electronic energies were computed at the PBE0/aDZ level of theory. 7
- Figure I-4.** The schematic presentation of the Python code written to perform the molecular rotation needed to prepare inputs for NICS(1)_{zz} calculations. 12
- Figure I-5.** AMHB strengthens H-bonded dimers of **1** and **2** in contrast to their respective reference compounds **1'** and **2'**. Nonbonding electron pairs not involved in π -delocalization are not shown. Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries..... 16
- Figure I-6.** AMHB weakens H-bonded dimers of **3** and **4** in contrast to their respective reference compounds **3'** and **4'**. Nonbonding electron pairs not involved in π -delocalization are not shown. Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries..... 17
- Figure I-7.** AMHB weakens H-bonded dimers of **5** and **6** in contrast to their respective reference compounds **5'** and **6'**. Nonbonding electron pairs not involved in π -delocalization are not shown. Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries..... 18
- Figure I-8.** AMHB strengthens H-bonded dimers of **7** and **8** in contrast to their respective reference compounds **7'** and **8'**. Nonbonding electron pairs not involved in π -delocalization are not shown. Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries..... 19
- Figure I-9.** Comparison of X-ray crystal structure H-bond donor/acceptor distances for (A) **2a** vs. **2a'** derivatives and (B) **3a** vs. **3a'** derivatives. R groups may be H, alkyl, phenyl, 4-chlorophenyl, 4-trifluoromethylphenyl, and in a few cases, 4-methoxy and 3,4-dimethoxy phenyl, with the aryl group twisted out of the plane of the five-membered rings. 20
- Figure I-10.** The dimerization energies of compound **9** and its hydrogenated analogues. The resonance forms that are enhanced upon H-bonding are shown. Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries. 21

Figure I-11. Comparison of X-ray crystal structure H-bond donor-acceptor distances for 9_{dimer} and 12_{dimer} derivatives.....	22
Figure I-12. (A) How to explain the AMHB effect in fused rings with extended π -systems. (B) Two cases where the enhanced aromaticity of one ring disrupts or enhances that of the second one calculated at DF-MP2/aQ5Z//MP2/aDZ using homodimers.	24
Figure I-13. Tautomerization energies for studied compounds.....	28
Figure I-14. Stacked NMR spectra of the NH region of compound 1' over a range of concentrations at four temperatures.	31
Figure I-15. The plot of N-H chemical shift of compound 1' by varying concentration at four temperatures. The markers are experimental data for the corresponding temperature and the grey lines are the best fits obtained from equation I-11...	32
Figure I-16. The two cases of aromaticity-modulated H-bonding. Enthalpies were measured in benzene via NMR spectroscopy. Aromatic rings are shown in red.....	36
Figure I-17. H-bonded heteroatom distances from X-ray structures of 1 , 1_{t-butyl} and 2 (left) vs their reference compounds 1' , 1'_{t-butyl} and 2' (right). Oxygen, nitrogen, carbon and hydrogen atoms are red, blue, gray and white respectively.....	37
Figure I-18. Gas-phase dimerization electronic energies (kcal/mol). Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.	38
Figure I-19. Two examples of cases of RAHB.....	41
Figure I-20. Two cases for evaluation of the magnitude of the RAHB effect. Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.	42
Figure I-21. Reduction energy modulations via AMHB. Energy (in kcal/mol) and NICS(1) _{zz} (in ppm) values are respectively calculated at CCSD(T)/CBS and mPW91PW91 using MP2/aDZ geometries.	45
Figure I-22. H-bonding modulation of flavin upon H-bonding.	48
Figure II-1. Sorted by sum of mean absolute errors of various levels of theory vs gas-phase CCSD(T)/CBS electronic energies and experimental H-bond free energies.	71
Figure II-2. (A), (B), and (C) Individual signed errors for the three best-performing and (D) the most economical level of theory.....	79

LIST OF SCHEMES

- Scheme I-1** . Timeline for the development of H-bonding, aromaticity and AMHB concepts. 6
- Scheme I-2**. Generating four cases for exploring the scope of AMHB. X can be NH, O, or S..... 9
- Scheme I-3**. (A) Structures of compounds used for experimental evaluation of the AMHB concept, and (B) and (C) two examples showing the significance of their H-bond formation in two protein binding pockets.....27
- Scheme II-1**. (A) Complications from H-bonding polymerization in alcohols. (B) H-bonding species studied here.....59

KEY TO SYMBOLS AND ABBREVIATIONS

Å	Angstrom; 10^{-10} meter
°	degrees
°C	degree centigrade
K	Kelvin
(anti)aromaticity	Antiaromaticity and aromaticity
H-bond	Hydrogen Bond
AMHB	(Anti)aromaticity-Modulated Hydrogen Bonding
RAHB	Resonance-Assisted Hydrogen Bonding
NICS	Nucleus-Independent Chemical Shift
NICS(1) _{zz}	Dissected Nucleus-Independent Chemical Shift
NMR	Nuclear Magnetic Resonance
ppm	Parts per million
M	Molar; moles per liter
mM	Millimolar; 10^{-3} molar
µM	Micromolar; 10^{-6} molar
SCF	Self-consistent field

DFT	Density Functional Theory
PBE0	The PBE1PBE DFT method
HF	Hartree-Fock method
MP2	2nd order Møller–Plesset perturbation method
CCSD(T)	Coupled Cluster with singles, doubles and perturbative triples
DF-HF	Density-fitted HF method
DF-MP2	Density-fitted MP2 method
aDZ	Dunning aug-cc-pVDZ basis set
aTZ	Dunning aug-cc-pVTZ basis set
aQZ	Dunning aug-cc-pVQZ basis set
a5Z	Dunning aug-cc-pV5Z basis set
aQ5Z	Dunning Basis set extrapolation using cardinal numbers 4 and 5.
kcal/mol	Kilocalories per mol
a.u.	Hartree per particle; 1 a.u. = 627.509 kcal/mol

CHAPTER I: AMHB: (ANTI)AROMATICITY-MODULATED HYDROGEN BONDING

I.1. Introduction

This chapter explores and generalizes an interplay between (anti)aromaticity and hydrogen bonding (H-bonding) of heterocycles which leads to modulations of their H-bond strengths. The concepts of aromaticity and H-bonding, which were formally developed during the first half of the 19th century, have been around for many decades.^{1,2} They have been used by chemists to explain chemical phenomena and to develop new chemical designs. However, their connection has remained underappreciated during these decades.

The credit for the first mention of the H-bonding concept in the literature can be given to Moore and Winmill³ for their 1912 work on properties of aqueous solutions of ammonium salts. About two decades before the term H-bond was coined by Linus Pauling in his famous book titled “The Nature of the Chemical Bond”, they used the concept of H-bonding to rationalize the stronger basicity of tetramethyl ammonium hydroxide than that of trimethyl ammonium hydroxide.⁴ A definition of H-bonds, very similar to what we teach in general chemistry courses using Lewis structures, appeared in the literature in 1920 in a paper by Latimer and Rodebush working with G. N. Lewis on the structure and properties of water:⁵

*"[A] free pair of electrons on one water molecule might be able to exert sufficient force on a hydrogen held by a pair of electrons on another water molecule to bind the two molecules together... Such an explanation amounts to saying that the hydrogen nucleus held between 2 octets constitutes a weak 'bond'."*⁵

Interestingly, they called it a “bond” in this description. Despite these and a few more appearances of the H-bonding concept in the literature,^{6,7} the term “hydrogen bond” was

only dubbed by Pauling in his first edition of his book "The Nature of the Chemical Bond" in 1939:⁸

"[U]nder certain conditions an atom of hydrogen is attracted by rather strong forces to two atoms, instead of only one, so that it may be considered to be acting as a bond between them. This is called the hydrogen bond" (taken from page 445 of the third edition published in 1959).⁸

Although this definition looks like the classical definition of H-bonding presented earlier, he goes on to put it in firm quantum chemical terms a few pages later on page 450:

"[A] hydrogen atom, with only one stable orbital, cannot form more than one pure covalent bond and that the attraction of two atoms observed in hydrogen-bond formation must be due largely to ionic forces."⁸

In 1957, a symposium was held in Ljubljana, Yugoslavia on the topic of H-bonding which was attended by the most prominent names in chemical bonding and theory such as Pauling, Pople, and Pimentel. In this meeting Coulson presented his paper which was simply titled "The Hydrogen Bond".⁹ In this landmark paper, which was a critical review of the current status of theory in understanding H-bonding, he went beyond the work of Lennard-Jones and Pople¹⁰ by dividing the H-bond interaction energy into four components: electrostatic, covalent, repulsive and dispersion. About 20 years later, Kitaura and Morokuma developed a SCF-based ab initio theory¹¹⁻¹³ that decomposed the H-bond energy into five components: electrostatic (ES), polarization (PL), exchange repulsion (EX), charge transfer (CT), and coupling (MIX) components which set the foundation for the many energy

decomposition methods which continue to be further developed and improved even up to the present date.¹⁴

Finally, in their 1960 book simply called "The Hydrogen Bond", Pimentel and McClellan provided a broad and practical definition for H-bonding that does not require the atoms (other than the hydrogen atom) involved in H-bonding to be limited to the three most electronegative atoms, nitrogen, oxygen and fluorine.¹⁵

*"A H Bond exists between a functional group A-H and an atom or a group of atoms B in the same or a different molecule when (a) there is evidence of bond formation (association or chelation), (b) there is evidence that this new bond linking A-H and B specifically involves the hydrogen atom already bonded to A."*¹⁵

In about the same period, due to the explosion of knowledge in bonding theory and quantum chemistry, the concept of aromaticity was being developed in parallel. In contrast to H-bonding, the "aromaticity" term was proposed in the literature by August Wilhelm Hofmann in 1857, many decades before it was understood in terms of chemical bonding theory.¹⁶ This early naming was mainly based on the distinct odors of benzene and toluene rather than on bonding notions or spectroscopic behavior of related compounds, although the substitution vs addition behavior was seen as a hallmark of aromatic compounds even then. The word "aromatic" originates from the Greek word aroma and means having a pleasant and distinctive smell. However, it was a decade later in 1866 when the cyclohexatriene structure of benzene was proposed by August Kekulé.¹⁷ The term aromatic sextet was introduced for the first time in 1925 by Sir Robert Robinson who used it to explain

the resistance of the six electron system to disruption and the puzzling inertness of its compounds to addition reactions known for olefins.¹⁸ This stability was formulated in molecular orbital theory terms by Hückel in 1931; his concepts, however, remained unrecognized for two decades. Although the more subtle aromaticity of tropolone was explained by Dewar in 1945, the famous Hückel $4n + 2$ rule was first articulated formally and clearly by Doering in a 1951 article in the *Journal of the American Chemical Society* on the subject of the synthesis and properties of tropolone:¹⁹

“The general molecular orbital theory ascribing peculiar stability to cyclic molecular orbitals containing $2 + 4n$ electrons has, among others, the corollary that the cycloheptatrienylium ion (I) should be more stable than the cyclopentadienylium ion (II) whereas the stability of the anions, $C_7H_7^-$ and $C_5H_5^-$, should be reversed.”¹⁹

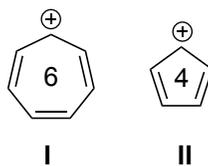


Figure I-1. The structures of cycloheptatrienylium and cyclopentadienylium ions. This figure was recreated from Doering's paper mentioned above.¹⁹

Interestingly, even before the $4n + 2$ rule was recognized by the community in a widespread manner, Dewar in 1945 noted the interplay between aromaticity and H-bonding in a study on Stipitatic Acid, without explicitly invoking aromaticity:²⁰

“When the appropriate triketocycloheptene carboxylic acid is written in the dienol forms (I) and (II), the possibility of resonance between them by hydrogen bond chelation becomes

evident; such a compound would be expected to show abnormal stability and lack of ketonic function."²⁰

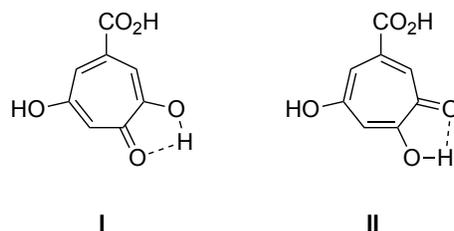
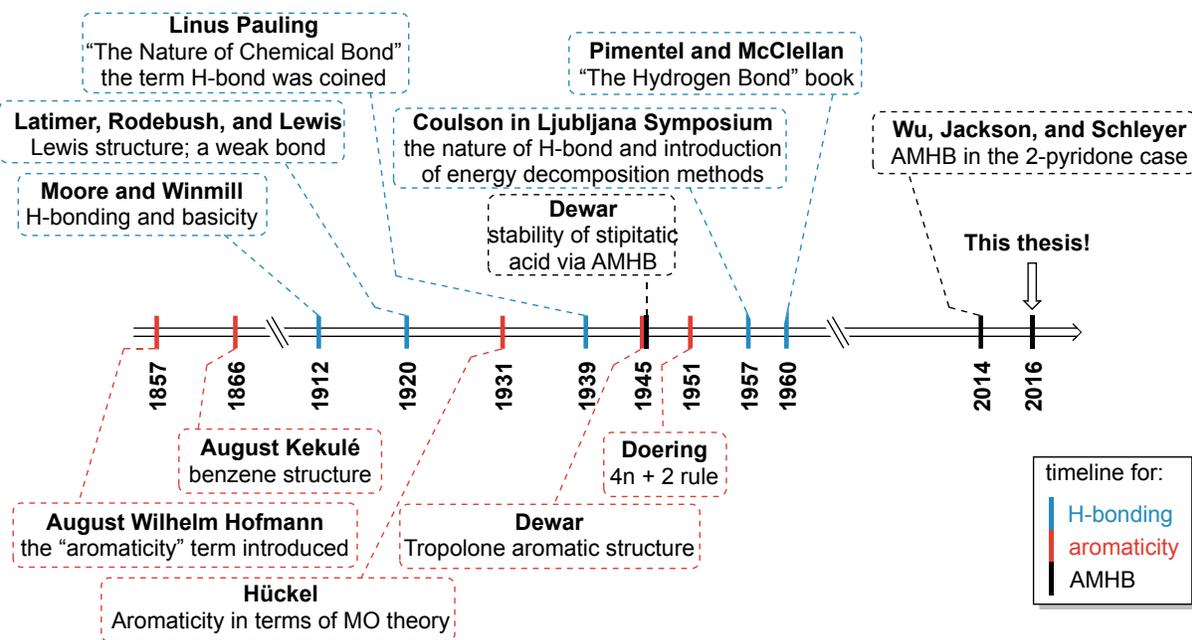


Figure I-2. Intramolecular H-bond interaction in Stipitic Acid. This figure was recreated from Dewar's paper.²⁰

Fast forward to 2014, the first study to directly point out the reciprocal interplay between aromaticity and H-bonding was published in a collaboration between Dr. Judy I. Wu, a postdoctoral fellow in the laboratory of Paul von Ragué Schleyer and now an assistant professor at University of Houston, and Dr. James E. Jackson, my Ph.D. advisor. The rest of this chapter presents my work which comprises of two main parts. The first part uses quantum chemical simulations to examine the generality of the interplay between H-bonding and both aromaticity and antiaromaticity of heterocycles for a broad range of examples. It uses reference compounds to quantify the magnitude of the energetic effects and provides magnetic evidence for changes of ring currents upon H-bonding. The second part, which is experimental, provides geometrical, energetic, and magnetic evidence for the effect. A Schematic timeline for the above history is presented in **Scheme I-1** below. The materials of this chapter were, in part, published and were adapted with permission from the publishers: (1) *The Journal of the American Chemical Society*, **2016**, *138*, 3427. Copyright 2016 American Chemical society. (2) *Angewandte Chemie International Edition*, **2017**, *56*, 9842. Copyright 2017 John Wiley and sons.



Scheme I-1 . Timeline for the development of H-bonding, aromaticity and AMHB concepts.

I.2. A broad in silico survey: generality and robustness of AMHB

Since the C-N bond is not a complete double bond, 2-pyridone (case **a** in **Figure I-3**) is only partially aromatic. But when it dimerizes, a resonance form is enhanced in which the aromaticity of the ring, and thus the stability of the system (aromatic rings are shown in red) increases. Therefore, its H-bond should be stronger than in case **e**, which includes the same H-bonding moieties, but lacks a cyclic π -system. However, in case **c**, H-bonding enforces some antiaromaticity on the four-membered $4\pi e^-$ ring, since the C=N π -bond character of the amidine group is decreased upon H-bond formation. As a result, its H-bonding is weaker than that of case **e**, in which no antiaromaticity is involved. Conversely, in cases **b** and **d**, H-bond formation respectively disrupts (favorable) aromaticity and relieves (unfavorable) antiaromaticity effects leading to weaker and stronger H-bond energies than in case **f**, respectively. The measured homo-dimerization enthalpies of **a** and **b** by Inuzuka and co-workers^{21,22} show that **a** binds 9.6 kcal/mol stronger than **b** (15.6 vs 6.0 kcal/mol).

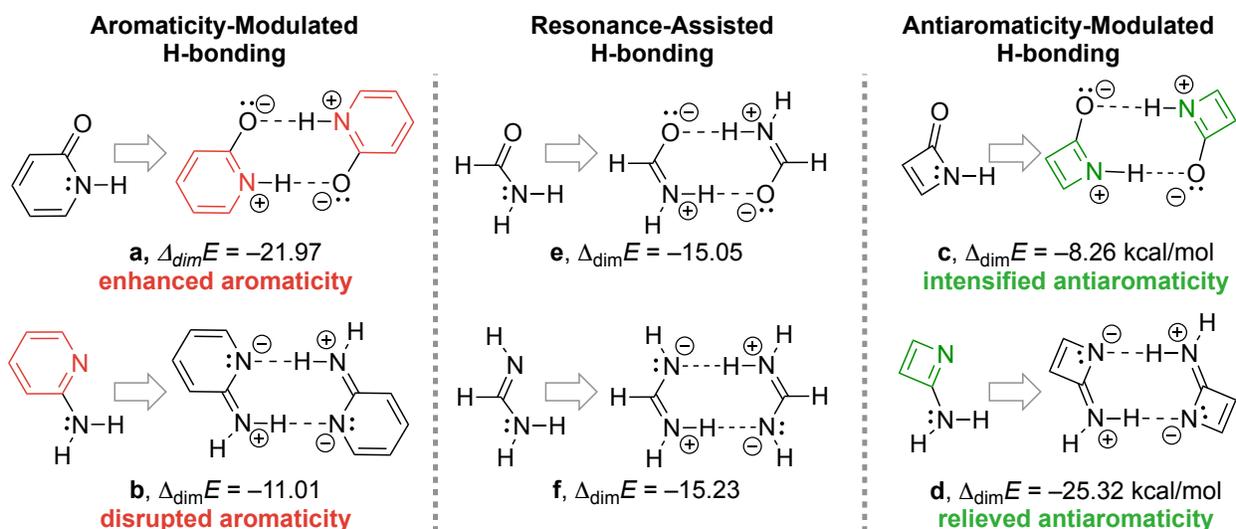


Figure I-3. Archetype examples of AMHB. H-bond electronic energies were computed at the PBE0/aDZ level of theory.

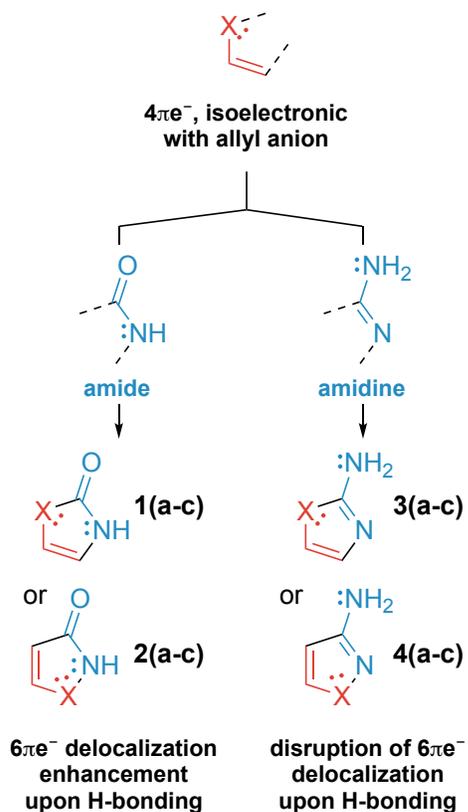
While the above points all make sense intuitively, here, the questions are: how strong is the AMHB effect? And, these H-bonds are weakened and strengthened with respect to what? In other words, what are the appropriate reference compounds to compare these (anti)aromatic systems with to quantify the magnitude of the AMHB effect on H-bond strengths. These reference compounds should provide the same H-bonding moieties as the (anti)aromatic heterocycles under study but lack their conjugation to (anti)aromaticity. Fortunately, these structures can be easily constructed by hydrogenation of the double bonds in the rings. The described reference molecules for an (anti)aromatic heterocycle can be obtained simply by disconnecting the (anti)aromatic π -circuit via hydrogenation of a ring double bond. For a six-membered ring, e.g. 2-pyridone, two isomers are possible (four if hydrogenation of both double bonds and 1,4 hydrogenation are also considered). In addition, the flexibility of the hydrogenated ring might lead to multiple conformers, further complicating the study by increasing the number of structures that need to be calculated. Four-membered rings introduce strain and pyramidalization as additional issues. For these

reasons, five-membered rings, which can be designed to be aromatic (e.g. thiophene) or antiaromatic (e.g. cyclopentadienone) were chosen as the templates for this study.

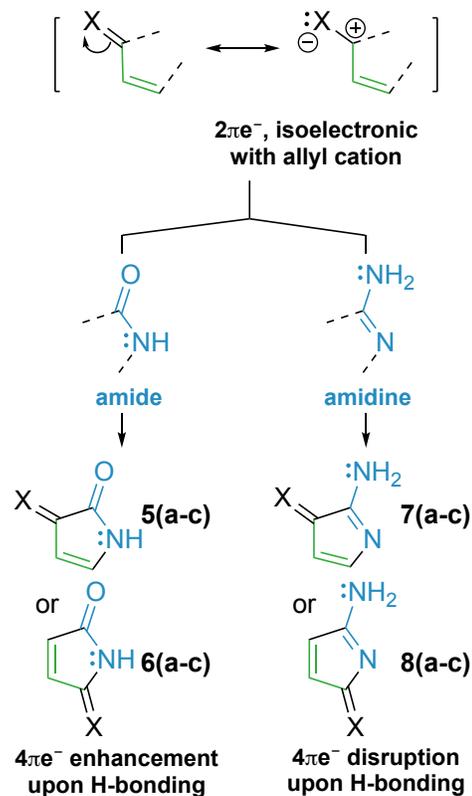
I.2.1. Designing examples to explore the generality of AMHB

This section concentrates on designing a broad set of heterocycles capable of coupling (anti)aromaticity and H-bonding to investigate the generality of the AMHB concept (**Scheme I-2**). For this purpose, amide and amidine groups were used as H-bond moieties. These groups make two-point H-bonds that can polarize them to change their C=N π -bond orders. Considering the usual resonance structures, the amide group increases the π -bond character between C and N upon H-bonding. Conversely, in the amidine case, H-bonding decreases the π -bond character of the C=N bond. These changes in π -bond character can be coupled to both aromaticity and antiaromaticity depending on how many π -electrons are paired with them in a cycle. If amide or amidine groups are coupled to groups isoelectronic to the allyl cation, they form partially or fully antiaromatic rings respectively (total of four π -electrons in the ring). On the other hand, if these H-bonding groups form a cycle with groups that are isoelectronic with allyl anion, they form partially or fully aromatic rings.

(A) Generating cyclic $6\pi e^-$ cases for aromaticity-modulated H-bonding



(B) Generating cyclic $4\pi e^-$ cases for antiaromaticity-modulated H-bonding



Scheme I-2. Generating four cases for exploring the scope of AMHB. X can be NH, O, or S.

The hypothesis is that these changes of π -bond character of the H-bond moieties could couple to the stabilizing and destabilizing effects of aromaticity or antiaromaticity. Hence, these perturbations should respond back and control the strength of H-bond. In this scenario, there are four possible cases upon H-bonding. The H-bond formation can

1. enhance aromaticity, and as a result thus be strengthened (lowered in energy upon association).
2. disrupt an aromatic ring and be weakened
3. relieve some antiaromaticity from a ring and be strengthened
4. enforce more antiaromaticity on the cycle and thus be weakened

Obtaining reference compounds in the abovementioned manner seems trivial for the AMHB case, but the challenge of designing reference compounds has led to controversial papers on the resonance-assisted hydrogen bonding concept (RAHB).^{23,24} With a wide range of examples for all the four possible cases of AMHB in hand, we put the hypothesis to test for its generality, but before that, the method used to follow the changes in magnetic properties of heterocycles, i.e. the degree of (anti)aromaticity, is discussed in the next section.

I.2.2. Nucleus-independent chemical shift

There is a number of descriptors for aromaticity that are based on geometric, energetic and magnetic consequences of aromatic stabilization. Each one has its own shortcomings and advantages.²⁵ Among these methods, nucleus-independent chemical shift (NICS) is simple to implement, and complements the energetic calculations by providing magnetic evidence for changes in “ring current” and thus (anti)aromaticity of the studied heterocycles. Thus, this method was chosen to follow the changes of aromaticity or antiaromaticity of the π -conjugated heterocycles upon H-bonding. The NICS values are simply obtained by calculating the magnetic shielding of a ghost atom in shielding or deshielding cones of aromatic or antiaromatic rings, respectively. For convenience of interpretation in analogy with ¹H NMR chemical shifts, the negative sign of the calculated shielding is usually reported, as recommended in the original paper.²⁶ Since the so-called ghost atom is just a measurement location, with neither a positively charged nucleus nor any basis set or electrons, it does not perturb the probed molecule. Thus, the calculated wave function or electron density of the probed molecule is not affected during the NICS simulations. This is a major advantage over using experimental probes such as alkali metals (Li) or protons complexed to the systems, as the probe can be selectively placed in locations

close to the molecule's site of interest to enhance the sensitivity of the method without interfering with the quantum mechanical calculations.

The dissected version of NICS, which minimizes the effect of sigma electrons on the calculated magnetic shielding, was used in this study. The NICS probe in the original work of Schleyer *et al.*²⁶ was placed at the center of the studied rings and the negative sign of the overall calculated shielding was used for reporting its value. However, the probe was later moved to 1 Å above the ring center and the negative sign of only the zz tensor of the magnetic shielding was used for reporting its value. Such a treatment not only decreases the effect of sigma electrons but also increases the sensitivity of the calculated shielding to π -electron ring currents by moving away from the former and closer to the latter. The notation of this dissected NICS is as NICS(1)_{zz}, in which (1) stands for the 1 Å distance and zz refers to the used tensor of computed shielding.

Calculation of NICS(1)_{zz} values generally requires finding the plane of a ring and reorientation of the molecule in the Cartesian coordinate system. As mentioned above, the NICS(1)_{zz} probe is, by definition, placed at one angstrom above (or below) the plane of the studied (anti)aromatic ring. Since, unlike benzene, many of the rings studied in this work are not completely flat, calculation of a least-squares mean plane was needed to be able to place the NICS(1)_{zz} probe in a well-defined manner. Then, for the extracted zz tensor to be meaningful, the molecule needed to be rotated with the mean plane perpendicular to the z-axis of the Cartesian coordinate system. **Figure I-4** shows the rotation required to prepare a NICS(1)_{zz} input file, along with a schematic presentation of a Python²⁷⁻²⁹ code written to automate this tedious and repetitive work.

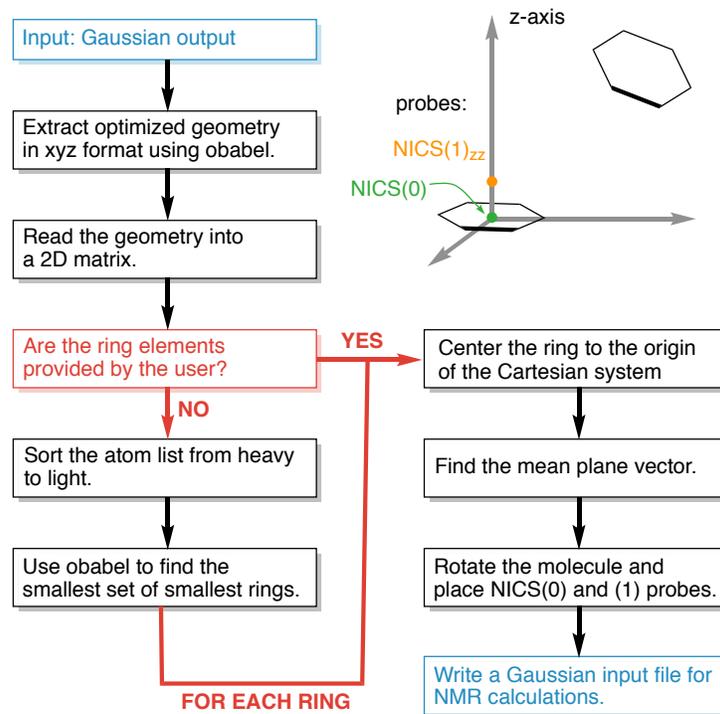


Figure I-4. The schematic presentation of the Python code written to perform the molecular rotation needed to prepare inputs for NICS(1)_{zz} calculations.

The code first uses open Babel³⁰ to extract the coordinates of an optimized geometry from a Gaussian output file and reads it into a 2-D matrix. Then if the labels of the ring atoms are provided by the user, the code proceeds to find the mean plane of the specified ring, rotate the molecule, and place the NICS(1)_{zz} probes before saving the Gaussian input file for NMR calculations. The centering (non-mass-weighted) of the ring is only for convenience and is not required; however, the NICS(1)_{zz} probe is then convenient to place 1 Å above this ring center. Placement of the NICS(1)_{zz} probe at the non-mass-weighted center of the ring is arbitrary. If the ring of interest is not specified by the user, the code again uses open Babel to find the smallest set of smallest rings (SSRs) in the structure and generates multiple NICS input files if more than one ring is found. The only reason the elements are sorted from heavy to light before searching for the SSRs is that the open Babel module removes the hydrogen

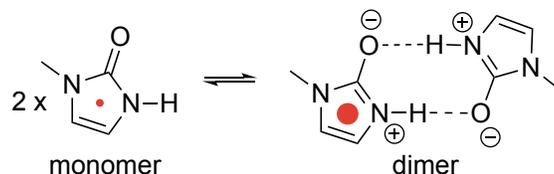
atoms before the SSRs search and as a result the indexing of the found ring atoms in the reported array may not match the original indexing of the structure. Considering that most non-flat rings may not be symmetrical on both faces, NICS(1)_{zz} probes were in all cases placed on both “top” and “bottom” of the rings. The resulting NICS(1)_{zz} values were averaged to eliminate the arbitrariness of choosing one face over the other one.

To follow the NICS(1)_{zz} changes upon dimerization of aromatic and antiaromatic heterocycles, Δ NICS(1)_{zz} was obtained as:

$$\Delta\text{NICS}(1)_{zz} = \text{NICS}(1)_{zz}^{\text{dimer}} - \text{NICS}(1)_{zz}^{\text{monomer}}$$

in which NICS(1)_{zz}^{dimer} and NICS(1)_{zz}^{monomer} refer to the calculated NICS(1)_{zz} values for dimer and monomer, respectively. Since the negative sign of the calculated zz tensor of shielding is taken as the NICS(1)_{zz} value, the reported values can be interpreted in a manner similar to ¹H NMR chemical shifts. For example, the calculated NICS(1)_{zz} value for benzene is -10.6 ppm³¹ as a proton in the shielding cone of an aromatic ring has a chemical shift of -4.0 ppm,³² both of which are large negative numbers. Presumably, the reason the experimental shift is 6.6 ppm more downfield than the NICS(1)_{zz} value is: (1) there is no electron around the NICS(1)_{zz} probe to shield it from the anisotropic magnetic field of benzene and (2) the NICS(1)_{zz} probe is closer to the center of the benzene ring than any protons could experimentally be located; hence it experiences a larger magnetic shielding. In contrast, the calculated NICS(1)_{zz} value for cyclobutadiene is +12.7 ppm, since the top of antiaromatic rings, where the NICS(1)_{zz} probe is placed, is in the deshielding cone. For interpretation of the reported NICS(1)_{zz} values in studying the cases of AMHB, the following short statement is helpful: *A negative Δ NICS(1)_{zz} value implies enhanced aromaticity or decreased*

antiaromaticity, while a positive value means decreased aromaticity or increased antiaromaticity.



The $\text{NICS}(1)_{zz}$ calculations were performed at the mPWPW91/6-311++G(3df,3pd) level of theory. Since these calculations are essentially NMR calculations, the choice of method was guided by benchmark study of ^1H NMR calculations by Lodewyk *et al.*³³ These authors have shown that the root mean square deviation (RMSD) of chemical shifts calculated with mPWPW91/6-311+G(2d,p) is only 0.16 ppm from a wide range of experimental ^1H NMR values. Here, the larger basis set 6-311++G(3df,3pd) was used.

I.2.3. Energetic calculations

Although in our original paper³⁴ the energetic effects of AMHB were calculated at the PBE0/6-311++G(3df,3pd) level of theory, all the energy calculations in this section are performed at the wavefunction-based CCSD(T)/CBS level of theory, the so-called “gold standard” method for non-covalent interactions.³⁵ The geometries, however, were optimized at the MP2/aDZ level of theory, which is the highest wavefunction level of theory accessible for the studied systems considering the resources available. Here, CCSD(T)/CBS refers to a composite method that approximates the complete basis set by summing three components:

$$E_{\text{approx.}}^{\text{CCSD(T)/CBS}} = E_{\text{elec.}}^{\text{DF-HF}} + E_{\text{extrapol.}}^{\text{DF-MP2}} + \delta_{\text{MP2}}^{\text{CCSD(T)}}$$

in which $E_{\text{approx.}}^{\text{CCSD(T)/CBS}}$, $E_{\text{elec.}}^{\text{DF-HF}}$, $E_{\text{extrapol.}}^{\text{DF-MP2}}$, and $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ are approximate CCSD(T)/CBS energy, density-fitted Hartree-Fock (DF-HF) energy, two-point extrapolated density-fitted MP2 (DF-MP2) correlation energy, and CCSD(T) correlation energy, all at the highest basis set

possible, respectively. The CCSD(T) correlation energy is defined as the energy difference between MP2 and CCSD(T) methods at the aDZ basis set. That is why the $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ term has an “MP2” subscript. Considering the size of our systems and computational resources available at MSU, we were able to use a5Z, aQZ to a5Z extrapolated, and aDZ basis sets for $E_{\text{elec}}^{\text{DF-HF}}$, $E_{\text{extrapol.}}^{\text{DF-MP2}}$, and $\delta_{\text{MP2}}^{\text{CCSD(T)}}$ calculations, respectively. Burns *et al.*³⁶ have shown that the extrapolated DF-MP2 and CCSD(T) correlation energies obtained using the above basis sets for a variety of non-covalent interaction energies of small systems are each within about 0.1 kcal/mol of the corresponding correlation energies calculated at the highest computationally possible basis sets. Thus, the calculated energies here should be within 0.2 kcal/mol of directly extrapolated CCSD(T) energies using Dunning basis sets with the highest possible cardinal numbers (two-point extrapolation).

I.2.4. Aromaticity gain strengthens H-bonds

Compounds **1(a-c)** and **2(a-c)**, are six cases of heterocycles that are partially aromatic (**Figure I-5**). They are called partially aromatic because the amide C-N has only a partial π -bond character. The partial aromaticity is also supported by their calculated small and negative NICS(1)_{zz} values. Their aromaticity can then be enhanced upon H-bonding, which leads to stronger H-bonds. Indeed, in all the cases in **Figure I-5**, the H-bonds in the aromatic heterocycles are stronger than their respective reference compounds. The enhancement of aromatic character upon H-bonding is supported by negative values of the calculated $\Delta\text{NICS}(1)_{zz}$. Thus, energetic and magnetic calculations suggest that enhancement of aromaticity upon H-bonding can strengthen H-bonds to a significant degree.

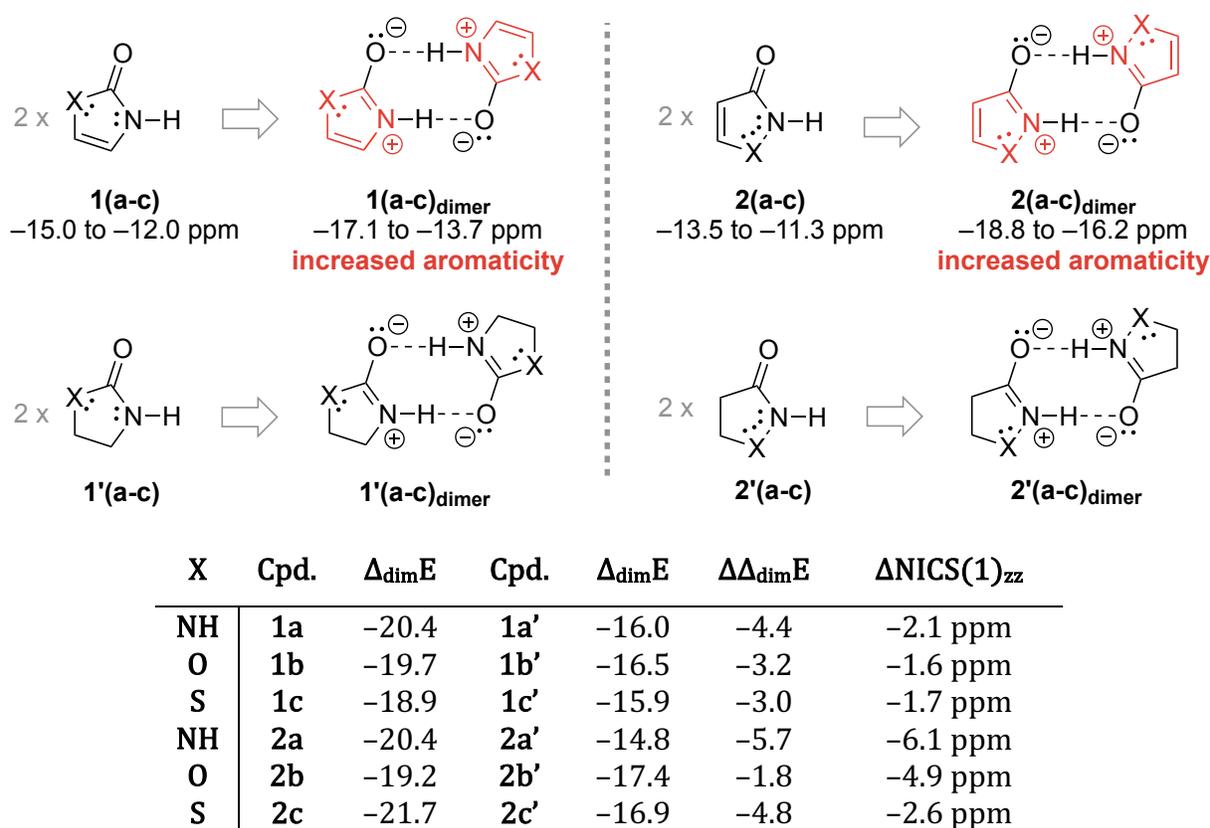
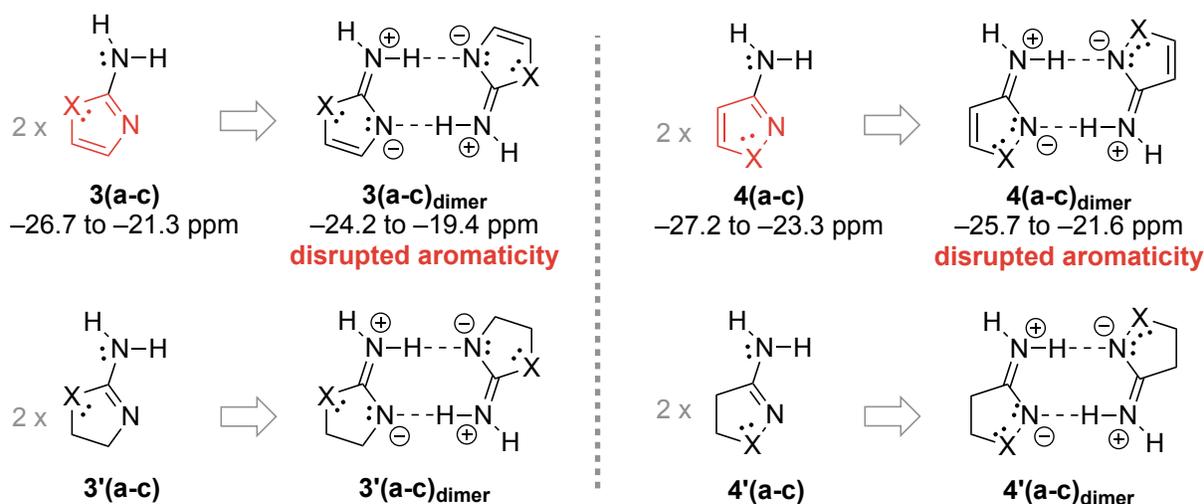


Figure I-5. AMHB strengthens H-bonded dimers of **1** and **2** in contrast to their respective reference compounds **1'** and **2'**. Nonbonding electron pairs not involved in π -delocalization are not shown.

Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.

I.2.5. Aromaticity loss weakens H-bonds

Supported by their calculated large and negative $\text{NICS}(1)_{zz}$ values, all the main compounds **3(a-c)** and **4(a-c)** in **Figure I-6** are strongly aromatic. Their aromaticity is disrupted upon H-bonding and hence their H-bonds are weaker than their respective reference compounds with the same H-bonding moieties but lacking aromaticity. The positive values of $\Delta\text{NICS}(1)_{zz}$ for all the six compounds upon H-bonding point to the disrupted aromatic rings. Therefore, in contrast to the first case, aromatization of heterocycles weakens H-bonds of potential practical interest. These findings may have practical applications as the 2-aminoimidazole class of compounds covers a wide range of experimental and approved drugs.



X	Cpd.	Δ_{dimE}	Cpd.	Δ_{dimE}	$\Delta\Delta_{\text{dimE}}$	$\Delta\text{NICS}(1)_{zz}$
NH	3a	-13.6	3a'	-14.5	+0.9	+2.5 ppm
O	3b	-14.3	3b'	-15.1	+0.8	+1.9 ppm
S	3c	-12.9	3c'	-14.8	+1.8	+2.4 ppm
NH	4a	-8.9	4a'	-12.2	+3.3	+1.5 ppm
O	4b	-11.7	4b'	-13.5	+1.8	+1.6 ppm
S	4c	-10.7	4c'	-13.3	+2.6	+2.0 ppm

Figure I-6. AMHB weakens H-bonded dimers of **3** and **4** in contrast to their respective reference compounds **3'** and **4'**. Nonbonding electron pairs not involved in π -delocalization are not shown.

Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.

I.2.6. Antiaromaticity increase weakens H-bonds

All the six compounds **5(a-c)** and **6(a-c)** have four π -electrons in their cycles, and hence they are formally antiaromatic (**Figure I-7**). However, only compounds **5(a-c)** show positive NICS values. The negative values in **6(a-c)** are probably due to isolation of the endocyclic double bond from the nitrogen non-bonding electron pairs by the two electron withdrawing groups adjacent to it. The changes of $\text{NICS}(1)_{zz}$ upon H-bonding, however, suggest enhanced paratropicity, i.e. antiaromaticity, in the all the dimers. Energetic comparisons also confirm weaker H-bonds in the π -cyclic systems compared to their

respective reference compounds. These results, although small in some cases, are in line with the AMHB hypothesis.

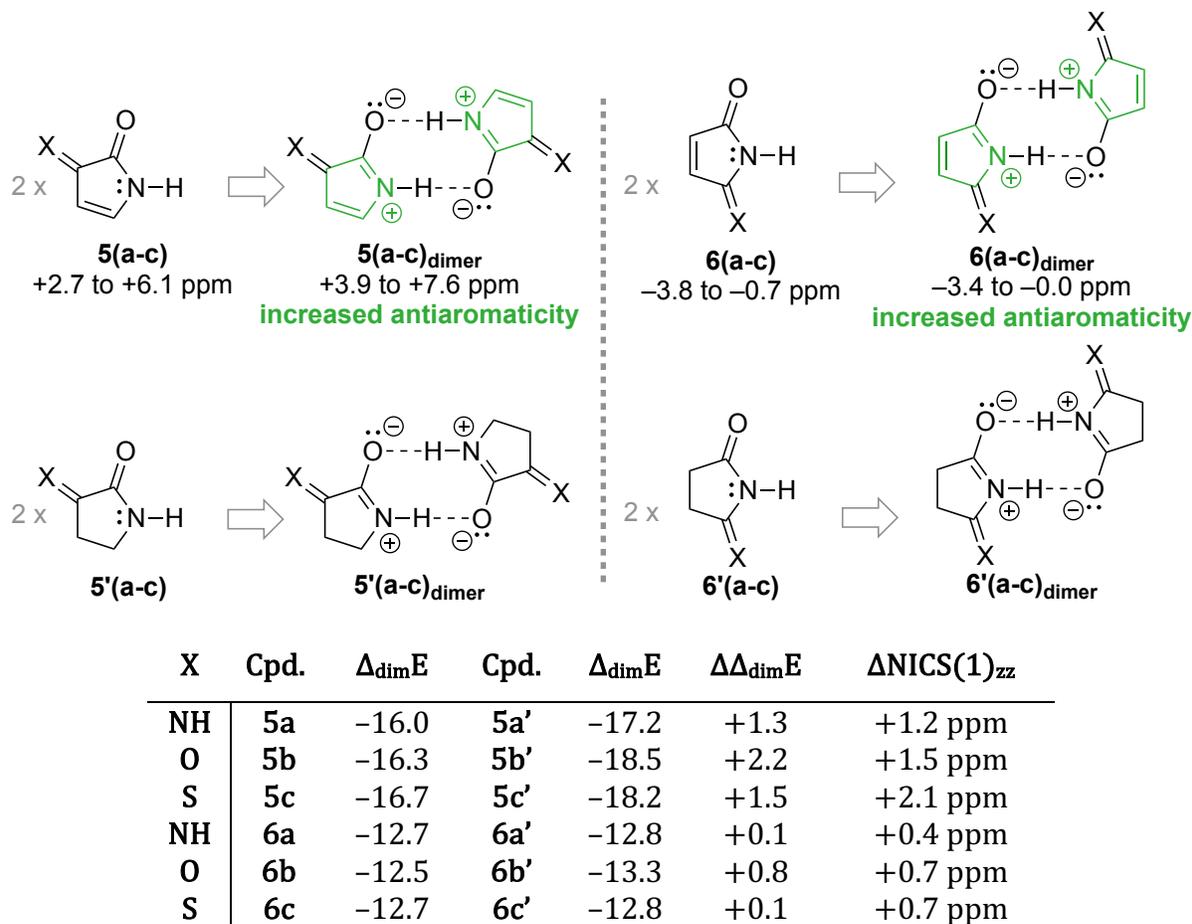


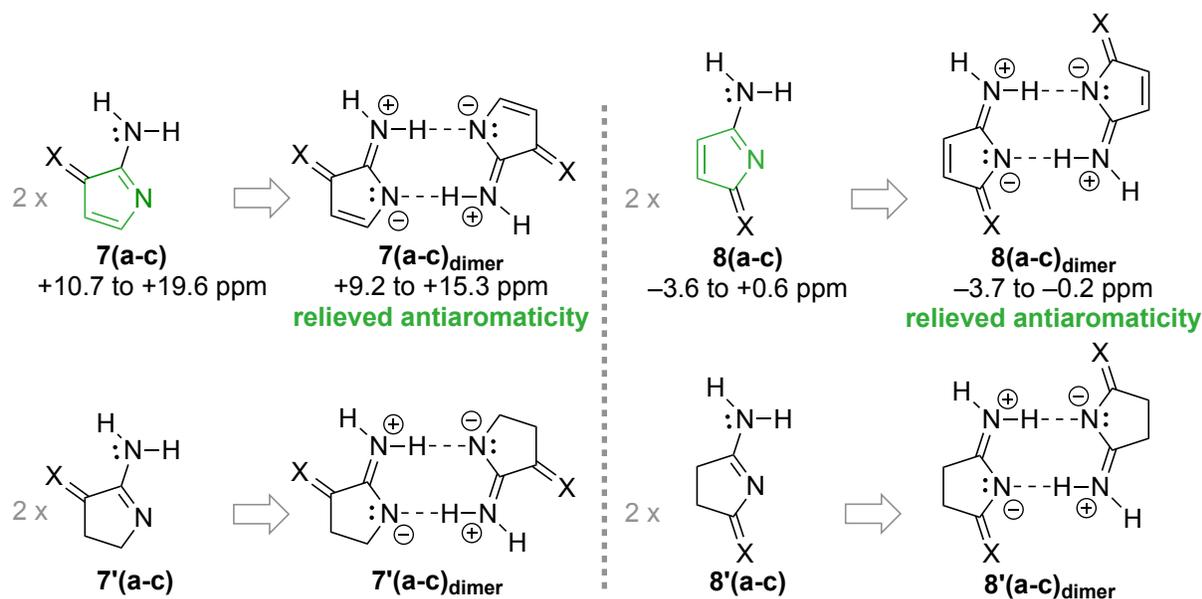
Figure I-7. AMHB weakens H-bonded dimers of **5** and **6** in contrast to their respective reference compounds **5'** and **6'**. Nonbonding electron pairs not involved in π -delocalization are not shown.

Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.

I.2.7. Antiaromaticity relief strengthens H-bonds

The last case of AMHB is the situation where relieving destabilization from antiaromatic π -system rings results in stronger H-bonds than the reference compounds. Compounds **7(a-c)** and **8(a-c)** in **Figure I-8** are all formally antiaromatic, and hence their NICS(1)_{zz} values are expected to be large and positive. This is true for the former group. However, the latter have small and negative values. Similar to the cases of **6(a-c)**, compounds

8(a-c) escape some antiaromaticity, by delocalizing the endocyclic C=N π -bonds outside the rings, isolating it from coupling to the other rings' π -bonds. Despite the small effect in **8(a-c)**, all the antiaromatic compounds relieve some antiaromaticity upon H-bonding, as suggested by their positive $\Delta\text{NICS}(1)_{zz}$ values. Accordingly, all the antiaromatic compounds form H-bonds stronger than those of their reference compounds.



X	Cpd.	Δ_{dimE}	Cpd.	Δ_{dimE}	$\Delta\Delta_{\text{dimE}}$	$\Delta\text{NICS}(1)_{zz}$
NH	7a	-16.9	7a'	-12.7	-4.2	-1.5 ppm
O	7b	-17.8	7b'	-13.0	-4.8	-2.5 ppm
S	7c	-18.3	7c'	-12.3	-6.0	-6.6 ppm
NH	8a	-16.8	8a'	-15.8	-1.0	-0.1 ppm
O	8b	-22.9	8b'	-21.1	-1.8	-0.7 ppm
S	8c	-23.7	8c'	-23.3	-0.4	-0.6 ppm

Figure I-8. AMHB strengthens H-bonded dimers of **7** and **8** in contrast to their respective reference compounds **7'** and **8'**. Nonbonding electron pairs not involved in π -delocalization are not shown.

Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.

A survey of the online Cambridge X-ray Structural Database (WebCSD),³⁷ in which H-bond donor-acceptor distances were compared, found that the average distance for **2a** in five crystals of its derivatives³⁸⁻⁴² is 0.116 Å shorter (2.707(6) vs 2.823(10) Å, **Figure I-9-A**) than

for **2a'** in nine crystal structures.⁴³⁻⁵⁰ In contrast, the average N----N distance in **3a** derivatives in three crystal structures⁵¹⁻⁵³ is 0.105 Å longer (3.013(6) vs. 2.908(6), **Figure I-9-B**) than **3a'** derivatives in four crystal structures.⁵⁴⁻⁵⁶ These observations offer strong experimental geometric support for the notion of aromaticity assisted H-bonding and aromaticity disrupted H-bonding in the cases of **2a** and **3a**, respectively.

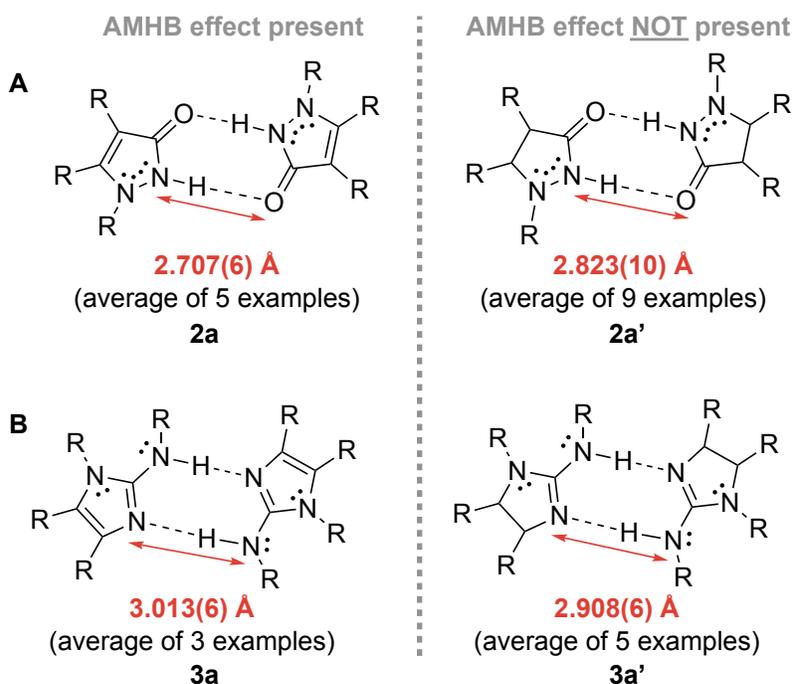


Figure I-9. Comparison of X-ray crystal structure H-bond donor/acceptor distances for (A) **2a** vs. **2a'** derivatives and (B) **3a** vs. **3a'** derivatives. R groups may be H, alkyl, phenyl, 4-chlorophenyl, 4-trifluoromethylphenyl, and in a few cases, 4-methoxy and 3,4-dimethoxy phenyl, with the aryl group twisted out of the plane of the five-membered rings.

I.3. Tuning association energies via AMHB

1H-imidazo[1,2-a]imidazole, compound **9** in **Figure I-10**, is a framework in which dimerization energy could in principle be predictably tuned via AMHB considerations from favorable to unfavorable through hydrogenation of different C=C π -bonds. In **9_{dimer}**, H-bonding enhances a resonance form in which an aromatic Clar sextet is formed as an

imidazolium ring while an imidazole sextet is disrupted. These described changes in ring currents are clearly supported by probing $\Delta\text{NICS}(1)_{zz}$ for both rings A and B. The $\Delta\text{NICS}(1)_{zz}$ is +2.5 ppm for the imidazole ring, whereas it is -0.8 ppm for the non-sextet ring (see **Figure I-10** under each ring for the $\text{NICS}(1)_{zz}$ values).

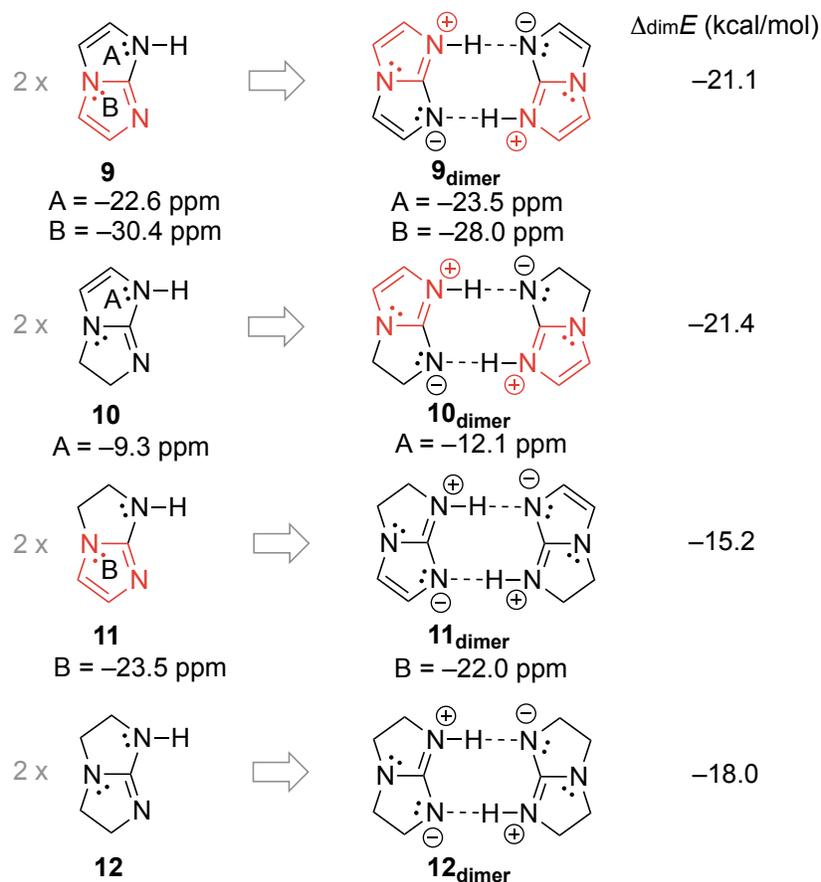


Figure I-10. The dimerization energies of compound **9** and its hydrogenated analogues. The resonance forms that are enhanced upon H-bonding are shown. Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.

Comparison of dimerization energies between compounds **9** and **12**, the related bicyclic with no cyclic π conjugation, reveals that the added $10\pi e^-$ aromatic system and its further delocalization upon H-bonding leads to a 3.1 kcal/mol stronger H-bond in **9** than in **12**. In **10_{dimer}**, an even higher dimerization energy is expected, since no imidazole sextet is

disrupted as the imidazolium sextet is enhanced. This difference, however, is quite small. On the other hand, dimerization of **11** should disrupt the imidazole sextet, weakening the H-bonding; this effect is observed as a 2.8 kcal/mol decrease in dimerization energy compared to **12**_{dimer}. In this case, the $\Delta\text{NICS}(1)_{zz} = +2.4$ ppm supports the idea that the diatropicity of the ring current is decreased upon dimerization. Such considerations for H-bonded system design would be of interest, since the H-bonding strength for the same guanidine moiety could be electronically tuned with a slight change in sterics, i.e. the hydrogenation of double bonds. Eventhough **9**, **11**, and **12** are known,^{49,51} compound **10** is not likely to be thermodynamically stable in a chemical environment; its computed electronic energy, without thermal corrections, is 9.3 kcal/mol above that of its isomer **11**.

Published crystal structures of derivatives of **9**_{dimer} and **12**_{dimer} also show shorter N---N distances in the former than the latter, in agreement with the calculated dimerization energies (**Figure I-11**). The H-bond length for **11**_{dimer} is 3.014(1), longer than those in both **9**_{dimer} and **12**_{dimer} (**Figure I-17**).

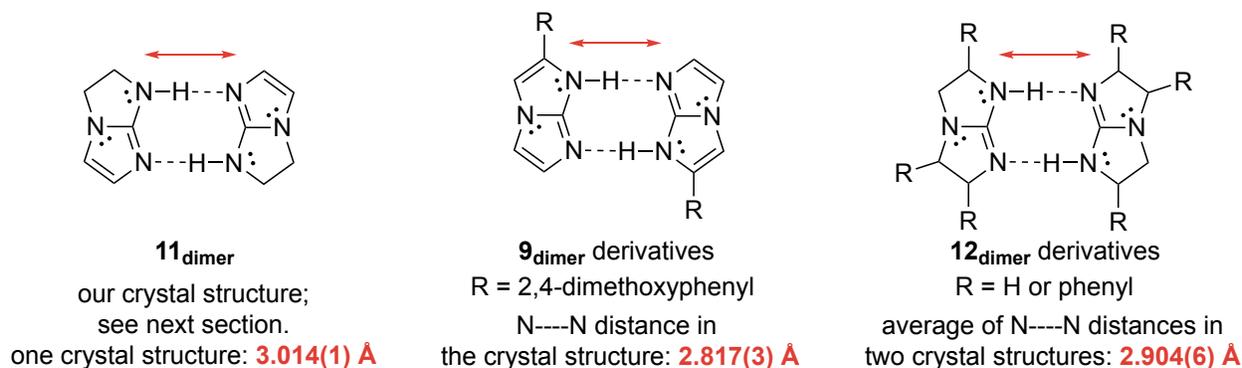
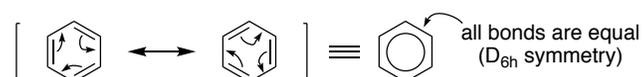


Figure I-11. Comparison of X-ray crystal structure H-bond donor-acceptor distances for **9**_{dimer} and **12**_{dimer} derivatives.

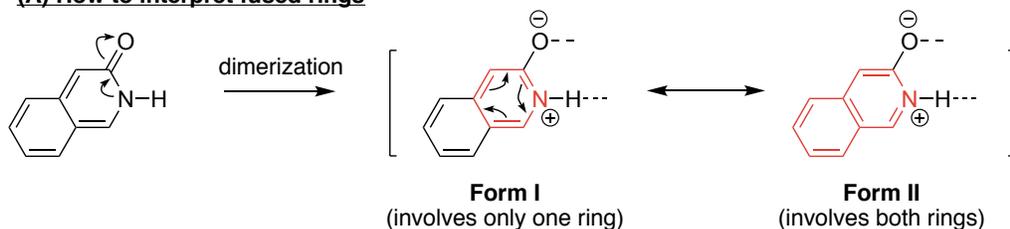
I.4. Remote AMHB effects in fused rings

The π -electron delocalization of cyclic rings due to AMHB leads to alteration in the π -bond character inside the rings. An extreme case is that of benzene in which the two resonance forms lead to complete equalization of π -bonds around the ring leading to the highly symmetrical D_{6h} point group.



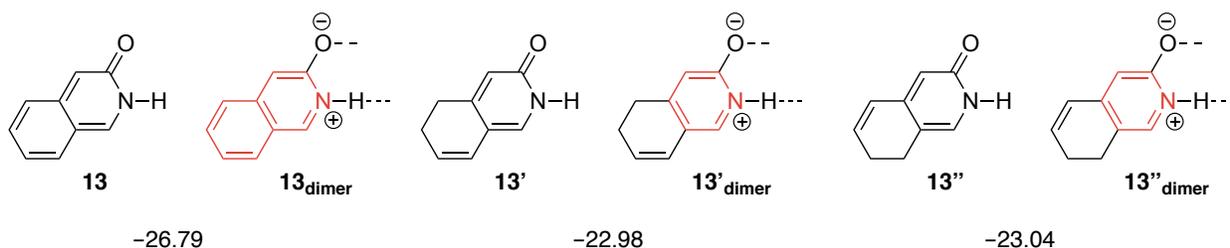
Hence, these π -bond order changes due to AMHB could, in principle, couple to (anti)aromaticity of a second ring fused to them. In previous sections, the resonance form shown in AMHB cases was only the one involving the amide or amidine moiety since that was enough to explain the changes of (anti)aromaticity of the rings upon H-bonding. However, in the fused rings studied in this section, the local functional group polarizations shown by the abovementioned resonance forms cannot be used to explain the effect of AMHB on the second ring. (Form I in **Figure I-12-A**) However, if the π -bonds of the ring in Form I are rotated, as shown using arrows in **Figure I-12-A**, resonance Form II is obtained, which can better describe the polarization effects of the first ring on the second one. In this case, the enhanced aromaticity of the first ring enhances the π -bond order of the bond shared between the two fused rings and hence the aromaticity of the second ring.

(A) How to interpret fused rings



(B) Two cases of fused rings

Case A: The aromaticity of both rings enhances upon H-bonding.



Case B (two possible examples): The aromaticity of one ring is disrupted while the other one is enhanced.

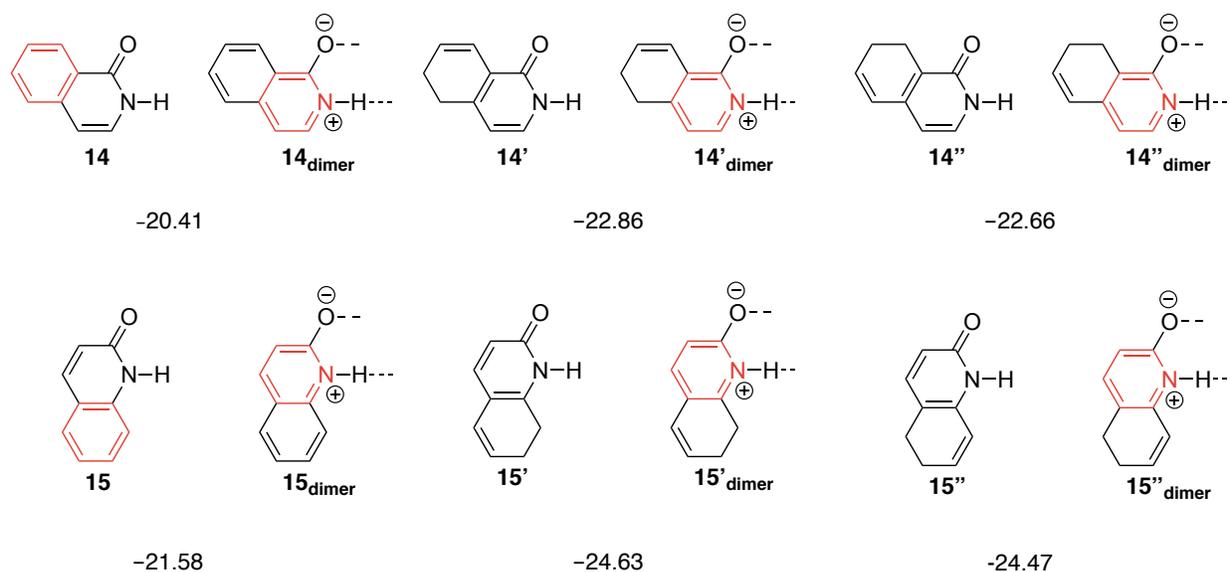


Figure I-12. (A) How to explain the AMHB effect in fused rings with extended π -systems. (B) Two cases where the enhanced aromaticity of one ring disrupts or enhances that of the second one calculated at DF-MP2/aQ5Z//MP2/aDZ using homodimers.

Two cases in which the enhanced aromaticity of the first ring upon H-bond formation can enhance (case A) or disrupt (case B) the aromaticity of the second one are shown in **Figure I-12-B**. In compound **13**, both rings are partially aromatic as suggested by their

calculated NICS(1)_{zz} values (see numbers inside respective rings). Upon H-bonding, the aromaticity of both rings is supposed to increase based on the resonance form shown. This hypothesis is supported by the upfield shifts of the NICS(1)_{zz} values of both rings. This leads to a ca. 4 kcal/mol stronger H-bond in **13** when compared to **13'** or **13''** in which the aromaticity of the second ring is breached by hydrogenation of any of the double bonds of the second ring.

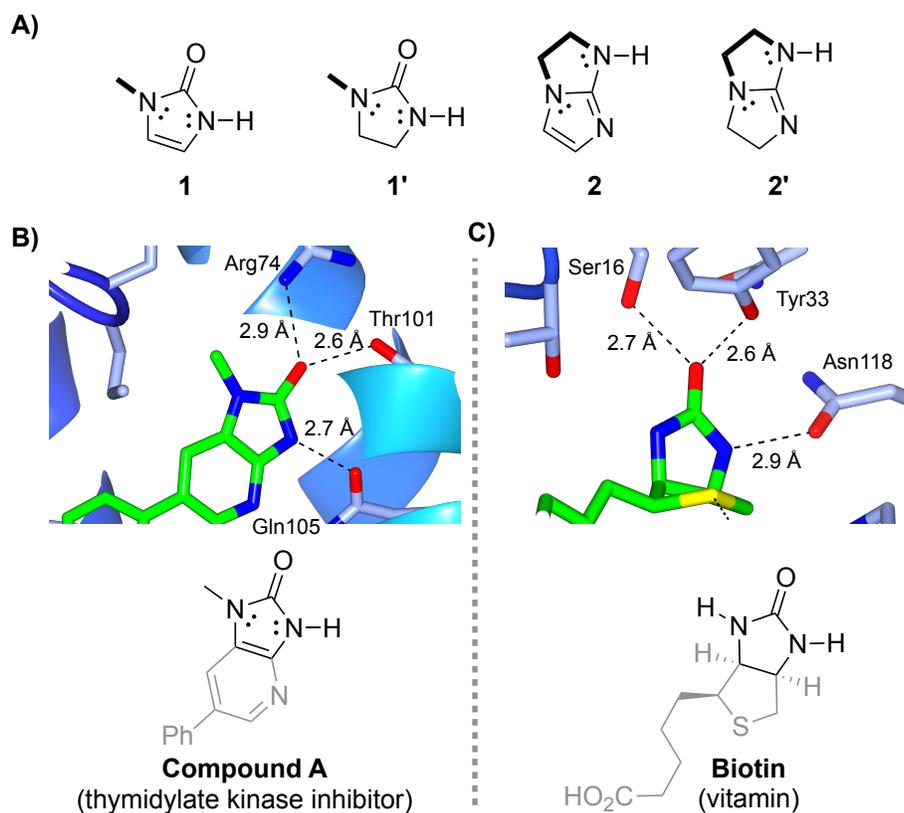
Conversely, in **14**, the enhancement of aromaticity of the first ring upon H-bonding leads to the disruption of that of the second ring. Thus, the H-bond energy of **14** is ca. 2 kcal/mol weaker than those of **14'** or **14''** in which the aromaticity of the second ring is eliminated via hydrogenation of any of the rings and cannot be disrupted. Notably, the H-bond energies of **13'**, **13''**, **14'** and **14''** are only marginally different meaning the location of the second ring and π -patterns that are not cyclic do not significantly interfere with AMHB. Lastly, for the sake of completeness of the discussion, compound **15** and its reference compounds, **15'** and **15''**, were explored. Similar to **14**, the aromaticity of the second ring of **15** is disrupted upon H-bond formation, and hence its H-bond energy is ca. 3 kcal/mol weaker than either of its reference compounds **15'** or **15''**.

I.5. Experimental measurement of the AMHB effect

In the previous section, quantum chemical analyses of H-bonding heterocycles, using the nucleus-independent chemical shift (NICS) method as an index of aromatic or antiaromatic π -delocalization, we found a systematic relationship between their (anti)aromaticity and H-bond association energies. Specifically, "H-bonding interactions that enhance aromaticity or relieve antiaromaticity are fortified, whereas those that intensify antiaromaticity or disrupt aromaticity are weakened, relative to analogues lacking full π -

circuits.”³⁴ As H-bonds play key roles in the interactions of biologically important heterocycles such as DNA bases, enzyme cofactors, and drugs, such insights into the factors affecting H-bond strength are of broad interest and use.⁵⁷⁻⁶⁰

The interplay between aromaticity and H-bonding can be probed by experimentally comparing heterocyclic dimers with H-bonds that strengthen or weaken aromatic delocalization. Therefore, compounds **1** and **2** (**Scheme I-3-A**) were chosen for these two cases, respectively. To quantify the effect, their H-bond energies are compared with those of compounds **1'** and **2'**, which have the same H-bonding motifs but lack π -conjugated circuits. Apart from their simplicity for this fundamental study, these heterocyclic frameworks can be found in many biologically active compounds. As an example for **1**, compound A (**Scheme I-3-B**) accepts two H-bonds from Arg74 and Thr101, while donating one to Gln105 in the binding pocket of pseudomonas aeruginosa thymidylate kinase.⁶¹ A well-known instance involving ring **1'** is biotin, which reacts with avidin to form one of the tightest complexes known by accepting two H-bonds from residues Ser16 and Tyr33, and donating one to Asn118 (**Scheme I-3-C**).⁶²



Scheme I-3. (A) Structures of compounds used for experimental evaluation of the AMHB concept, and (B) and (C) two examples showing the significance of their H-bond formation in two protein binding pockets.

Measurements of association energetics in classic H-bonded systems such as alcohols or water are complicated by formation of a range of aggregates. To avoid this complexity, **1**, **1'**, **2**, and **2'** were designed for simple pairwise homodimerization by blocking any extra H-bond donor moieties with methyl (in **1** and **1'**) or ethanato (in **2** and **2'**) groups. Another possible complication is tautomerization, as in 2-pyridone vs. 2-hydroxypyridine; such tautomeric equilibria are strongly affected by H-bonding media.^{63,64} However, G4(MP2)⁶⁵ calculations find compounds **1**, **1'**, and **2** to be well below their tautomers in energy, both in the gas-phase, and in implicit benzene solvent (**Figure I-13**).

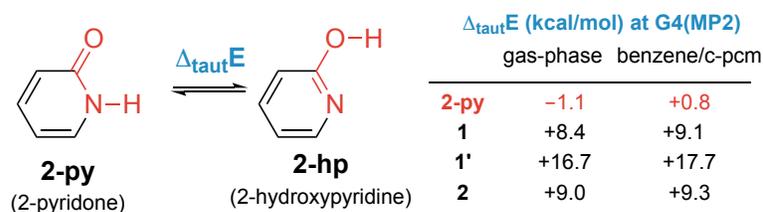


Figure I-13. Tautomerization energies for studied compounds.

For compounds **1**, **1'**, **2**, and **2'** in C_6D_6 , the homodimerization equilibria were established faster than the NMR time scale, so observed chemical shifts appeared as weighted averages of monomer and dimer forms. The molar equilibrium constant K_T (at temperature T) for the equilibrium between monomer M and its H-bonded dimer D can be written as follows:

$$K_T = \frac{[D]}{[M]^2} \quad \text{equation I-1}$$

where $[M]$ and $[D]$ denote the molarities of M and D , respectively. In a solution with formal concentration of C_0 , i.e. $C_0 = [M] + 2[D]$, the $[D]$ could be written as a function of $[M]$

$$[D] = \frac{C_0 - [M]}{2} \quad \text{equation I-2}$$

substitution of equation I-2 in I-1 leads to a quadratic equation, after rearrangement:

$$2K_T[M]^2 + [M] - C_0 = 0 \quad \text{equation I-3}$$

solving via the quadratic formula yields one positive and one negative answer for $[M]$. However, since negative molarities are not meaningful, the only real answer for the above equation would be:

$$[M] = \frac{\sqrt{1+8K_T C_0} - 1}{4K_T} \quad \text{equation I-4}$$

Since the dynamics of H-bonding equilibria are usually faster than the NMR time scale, the proton involved in H-bonding is observed as a weighted average δ_{AVG} of its monomer and dimer forms:

$$\delta_{AVG} = X_M(\delta_M - \delta_D) + \delta_D \quad \text{equation I-5}$$

where X_M is the mole fraction of the nucleus of interest in the monomer form, and δ_M and δ_D are the chemical shifts of that nucleus in the monomer and dimer forms at temperature T , respectively. For a solution with C_0 concentration, X_M can be written as:

$$X_M = \frac{[M]}{C_0} \quad \text{equation I-6}$$

Substitution of $[M]$ from equation I-4 in equation I-6, and then X_M from equation I-6 in I-5 results in equation I-7:

$$\delta_{AVG} = \frac{\sqrt{1+8K_T C_0} - 1}{4K_T C_0} (\delta_M - \delta_D) + \delta_D \quad \text{equation I-7}$$

in which, K_T , δ_M , and δ_D can be replaced by

$$K_T = e^{-\frac{\Delta_{dim}H - T\Delta_{dim}S}{RT}} \quad \text{equation I-8}$$

$$\delta_M = \delta_{M0} + d\delta_M/dT(T - T_0) \quad \text{equation I-9}$$

$$\delta_D = \delta_{D0} + d\delta_D/dT(T - T_0) \quad \text{equation I-10}$$

which results in

$$\delta_{AVG} = \frac{\sqrt{1+8e^{-\frac{\Delta_{dim}H - T\Delta_{dim}S}{RT}} C_0} - 1}{4e^{-\frac{\Delta_{dim}H - T\Delta_{dim}S}{RT}} C_0} ([\delta_{M0} + d\delta_M/dT(T - T_0)] - [\delta_{D0} + d\delta_D/dT(T - T_0)]) + [\delta_{D0} + d\delta_D/dT(T - T_0)] \quad \text{equation I-11}$$

in which, $\Delta_{dim}H$, $\Delta_{dim}S$, δ_{M0} , $d\delta_M/dT$, δ_{D0} , and $d\delta_D/dT$ are respectively, the enthalpy of dimerization (kcal/mol), the entropy of dimerization (kcal/molK), the monomer chemical shift at the lowest measurement temperature (ppm), temperature dependence of monomer chemical shift (ppm/K), dimer chemical shift at the lowest measurement temperature (ppm), and the temperature dependence of the dimer chemical shift (ppm/K). The constant R is in units of kcal/molK. Substitution of equations I-8, I-9 and I-10 into equation I-7 implies

that the $\ln K_7$ is constrained to be a linear function of $1/T$ based on the van 't Hoff equation and the chemical shifts of monomer and dimer are assumed to be a linear function of temperature.

The averaged chemical shifts, which are the only experimental observable, were then measured for series of samples over a range of concentrations at four temperatures to obtain the spectra presented in **Figure I-14**. Then, by varying the six independent variables, $\Delta_{\text{dim}}H$, $\Delta_{\text{dim}}S$, δ_{M0} , $d\delta_M/dT$, δ_{D0} , and $d\delta_D/dT$ the best fit of equation I-11 to the experimental data was obtained.^{66,67} **Figure I-15** is a visual presentation of the result of the fit for compound **1'**. The markers show the experimental chemical shifts at various concentrations at four temperatures, and the grey lines are the plot of equation I-11 using the optimized values for the abovementioned six parameters. With the best values of these six parameters in hand, the thermodynamic and NMR parameters at 298.15 K were calculated. These quantities, along with Gibbs free energy $\Delta_{\text{dim}}G$ and equilibrium constant K values at 298.15 K, are summarized in **Table I-1** for compounds **1**, **1'**, **2**, and **2'**, **BA**, **IM** in four solvents. The near perfect fit of experimental data to equation I-11 supports the absence of aggregates larger than dimers in the temperature and concentration ranges chosen for this study. All measurements in benzene which are used in this chapter were performed in triplicate to calculate the standard deviations of enthalpies shown in **Figure I-16**. While about 90% data completeness was obtained and needed for species **1** and **2'** for **1'** and **2**, data covering about 70% (due to solubility limitations) of the monomer-dimer chemical shift range was enough to obtain the uncertainties (standard deviation) seen in **Figure I-16**. Data coverage index is defined as the ratio of the range of chemical shift measured experimentally divided by the difference between extrapolated dimer and monomer chemical shifts multiplied by 100%.

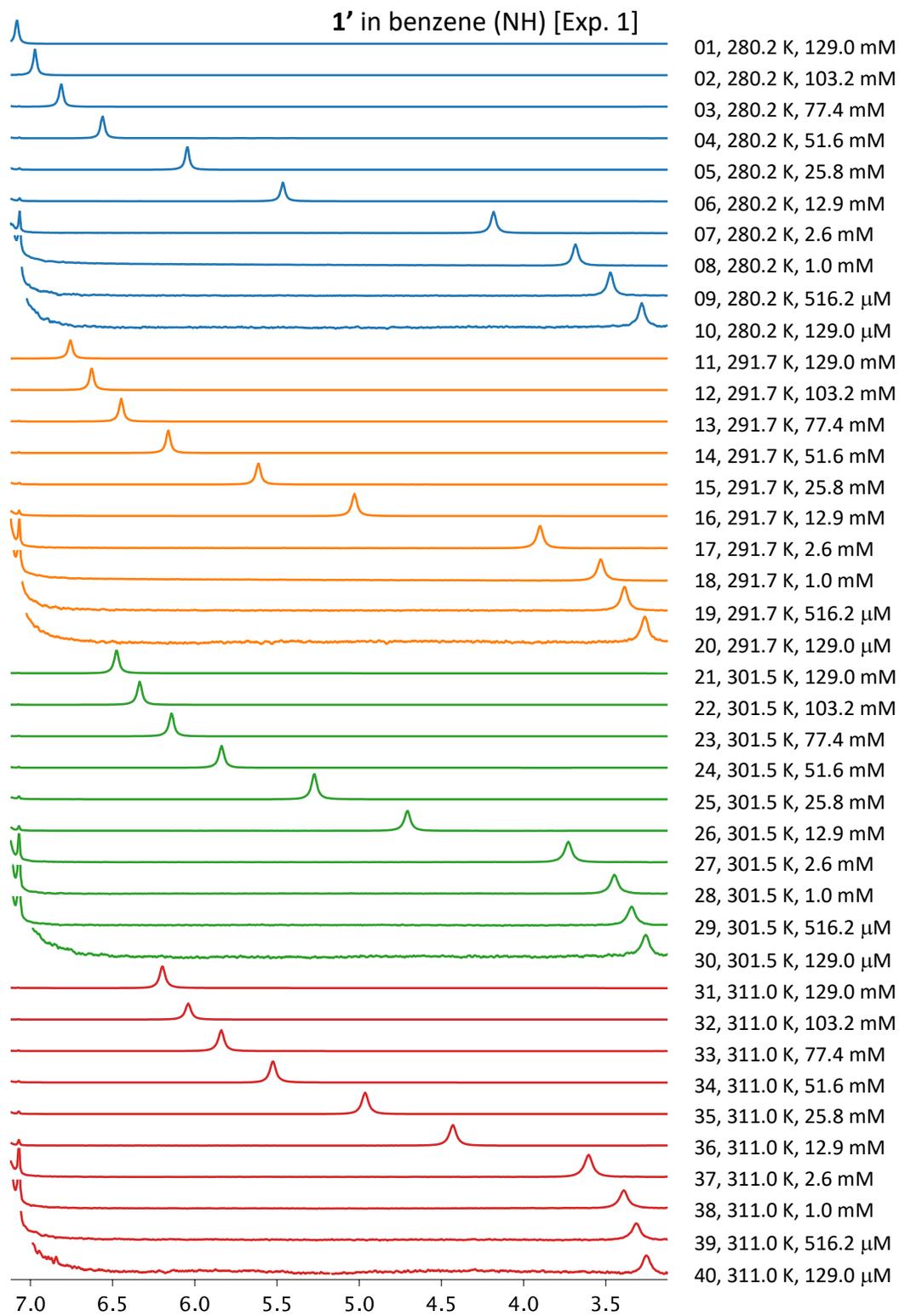


Figure I-14. Stacked NMR spectra of the NH region of compound 1' over a range of concentrations at four temperatures.

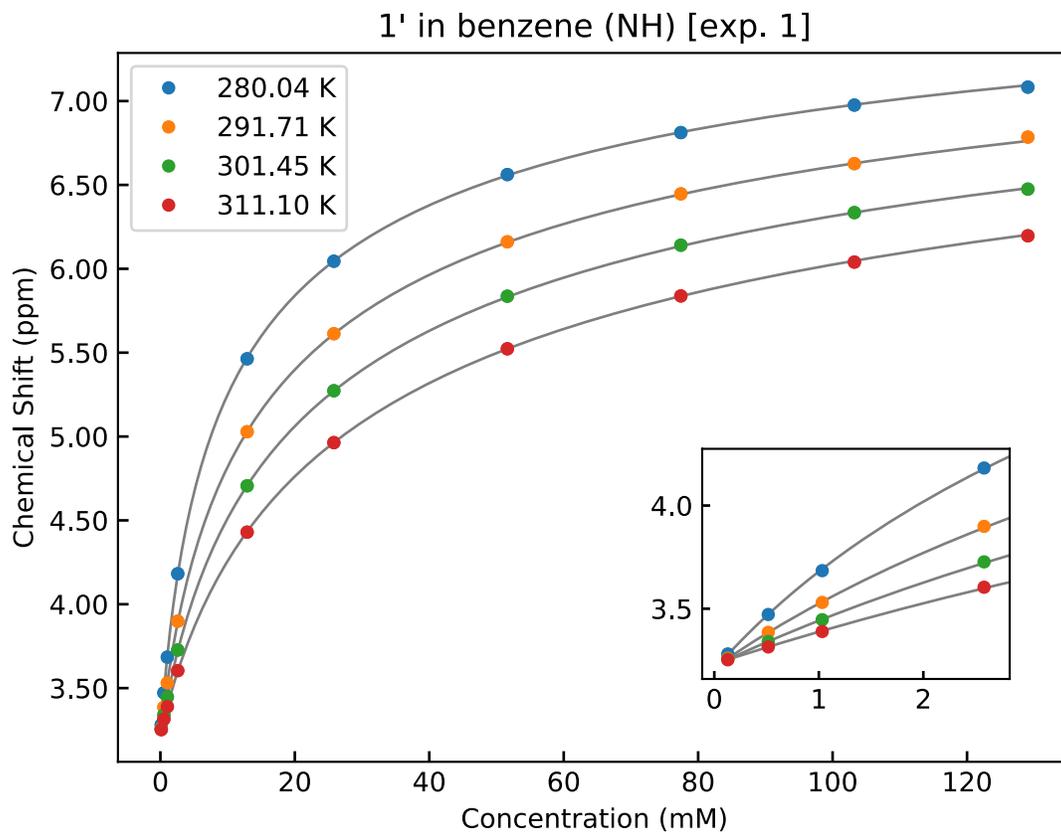


Figure I-15. The plot of N-H chemical shift of compound 1' by varying concentration at four temperatures. The markers are experimental data for the corresponding temperature and the grey lines are the best fits obtained from equation I-11.

Table I-1. Thermodynamic and chemical shift values (at 298.15 K and vs tetramethyl silane) obtained from NMR spectroscopy.

exp.	ΔH kcal/mol	ΔS cal/molK	ΔG kcal/mol	Keq 1/M	δ_D ppm	δ_M ppm	$d\delta_D/dT$ ppb/K	$d\delta_M/dT$ ppb/K	data coverage (%)
1 in C ₆ D ₆ (NH) [exp. 1]	-9.05	-15.17	-4.52	2073.19	12.5257	5.6646	-7.1	1.5	91
1 in C ₆ D ₆ (NH) [exp. 2]	-9.03	-15.18	-4.50	1995.22	12.5398	5.6849	-6.5	1.3	89
1 in C ₆ D ₆ (NH) [exp. 3]	-8.75	-14.19	-4.52	2065.10	12.5287	5.6952	-7.3	-0.2	92
1' in C ₆ D ₆ (NH) [exp. 1]	-6.55	-15.32	-1.98	28.11	8.0744	3.2207	-9.4	0.9	76
1' in C ₆ D ₆ (NH) [exp. 2]	-6.52	-15.23	-1.97	27.96	8.0710	3.2178	-9.5	0.7	77
1' in C ₆ D ₆ (NH) [exp. 3]	-6.64	-15.65	-1.98	28.18	8.0649	3.2232	-9.4	0.8	77
2 in C ₆ D ₆ (NH) [exp. 1]	-6.56	-14.98	-2.09	34.28	8.1583	2.8941	-10.7	0.7	65
2 in C ₆ D ₆ (NH) [exp. 2]	-6.54	-14.91	-2.09	34.23	8.1617	2.8950	-10.7	0.7	65
2 in C ₆ D ₆ (NH) [exp. 3]	-6.62	-15.21	-2.08	33.66	8.1856	2.9005	-9.3	0.6	65
2' in C ₆ D ₆ (NH) [exp. 1]	-8.85	-18.01	-3.48	354.32	9.5827	3.1315	-8.6	0.2	86
2' in C ₆ D ₆ (NH) [exp. 2]	-9.14	-18.76	-3.55	400.65	9.5174	3.1039	-9.0	5.5	83
2' in C ₆ D ₆ (NH) [exp. 3]	-8.89	-18.14	-3.48	355.53	9.5924	3.0896	-7.8	2.4	83

Because the N–H resonances show by far the largest changes in chemical shift between dimer and monomer forms, they were the primary data used in computing association constants. Between the monomer and dimer chemical shift values δ_M and δ_D of the studied compounds' N–H sites, δ_D is uniformly several ppm downfield from δ_M due to H-bond induced deshielding.

C_6D_6 was the solvent selected for the present study because: 1) It has weak interactions with solute due to its low polarity. 2) It dissolves all compounds studied. 3) It also has the minimum number of overlapping peaks in the chemical shift range where the studied N–H peaks appear. 4) its freezing and boiling temperatures allowed measurement in the temperature range available on MSU's 900 MHz NMR instrument, i.e. ca. 5 to 40 °C. Deuterated methanol was used for temperature calibration of the cryogenic probe.⁶⁸ To rigorously exclude water (which could hinder analyses by forming H-bonded complexes with the substrates)⁶⁹ C_6D_6 was dried over molecular sieves, the substrates were sublimed before measurement, and samples were prepared in a dry box. In all experiments, no or negligible peaks were observed at 0.4 ppm, the chemical shift associated with residual water impurity in C_6D_6 .⁷⁰

For the dimerization energy differences between **1** and **2** and their reference compounds to be meaningful, accurate equilibrium data was needed. Though 1H NMR spectroscopy was a natural tool to use, MSU's 500 and 600 MHz spectrometers did not allow measurements of sufficiently low concentrations to yield reliable association energies for compound **1**. The high sensitivity of the 900 MHz NMR instrument equipped with a cryogenic probe (Bruker AVANCE) was the key to this problem, enabling useful spectra to be collected at concentrations as low as 10 μM . There, even compound **1**, the strongest H-bonded dimer

studied, was substantially dissociated. Measured enthalpies of dimerization are presented in **Figure I-16**. The chemical shift values for monomers and dimers, and associated temperature coefficients were optimized as parameters to fit Equation I-11 to the experimental data. They are corrected to 298.15 K for the sake of comparison. The aromatic hydrogens of compounds **1** and **2** were assigned by NOE. The assignment of compound **1** was also confirmed by HMBC. The observed methine CH chemical shift changes upon dimerization serve as probes for the enhanced resonance forms shown in **Figure I-16**. The details of association measurements using NMR and peak assignments along with associated spectra are presented at the end of this chapter.

To complement the NMR data, single crystals of compounds **1** and **2** (grown and characterized by X-ray crystallography) and known crystal structures of compounds **1_{t-butyl}**, **1'**, **1'_{t-butyl}**, and **2'** were examined to assess the geometric effects of AMHB.^{56,71-73} Unlike **1'**, compound **1** and the tert-butyl derivatives, **1_{t-butyl}** and **1'_{t-butyl}**, form simple dimers in their crystals.

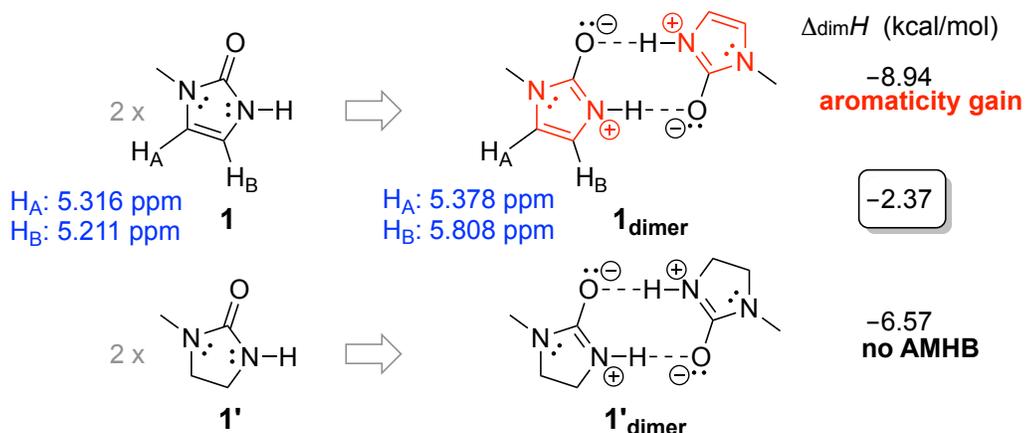
I.5.1. Energetic, geometric, and magnetic evidence for the interplay of aromaticity and H-bonding

Experimental energetic, geometric, and magnetic evidence document the effects of aromaticity gain and its effect by H-bond strengthening. Upon dimerization of **1**, cyclic six π -electron delocalization is enhanced in the five-membered rings of **1_{dimer}** (**Figure I-16-A**, top, red rings). The measured dimerization enthalpy is 2.37(15) kcal/mol more negative for **1** than it is for **1'** ($\Delta_{dim}H = -8.94$ vs. -6.57 kcal/mol), the analogue with no π -conjugated ring and thus no possible AMHB. The N \cdots O distance shortening (by 0.061(4) and 0.092(2) Å) in the structures of **1** vs. **1'** and **1_{t-butyl}** vs. **1'_{t-butyl}** (see top and middle in **Figure I-17**) also support

the stronger H-bonding in **1** and **1_{t-butyl}**. Downfield shifted ¹H NMR signals for H_A and H_B in **1_{dimer}** (by 0.062(1) ppm and 0.597(1), relative to those of **1**; see blue values, **Figure I-16-A**).

These document the magnetic effects of AMHB.

A) H-bond strengthening due to aromaticity enhancement



B) H-bond weakening due to aromaticity disruption

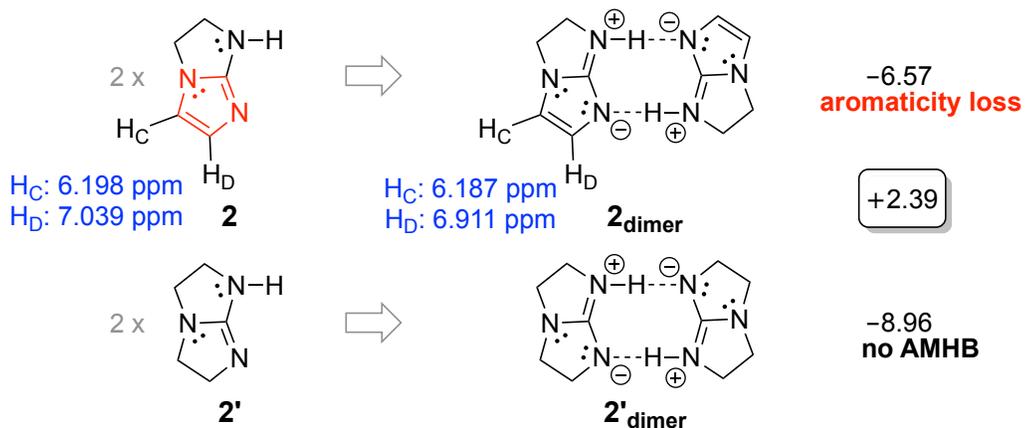


Figure I-16. The two cases of aromaticity-modulated H-bonding. Enthalpies were measured in benzene via NMR spectroscopy. Aromatic rings are shown in red.

Dimerization of **2** decreases cyclic six π-electron aromaticity in the lower ring of **2_{dimer}** (**Figure I-16-B**), resulting in a 2.39(16) kcal/mol *weaker* dimerization enthalpy for **2** vs. **2'** (Δ_{dim}H = -6.57 vs. -8.96 kcal/mol, respectively), a 0.108(2) Å *longer* intermolecular N⋯N distance in **2_{dimer}** than in **2'_{dimer}** (see **Figure I-17**), and *upfield* shifts (by 0.011(1) and 0.128(1)

ppm) in the ^1H NMR signals for H_E and H_F in $\mathbf{2}_{\text{dimer}}$ relative to those in $\mathbf{2}$ (see the blue values in Figure I-16-B).^[10]

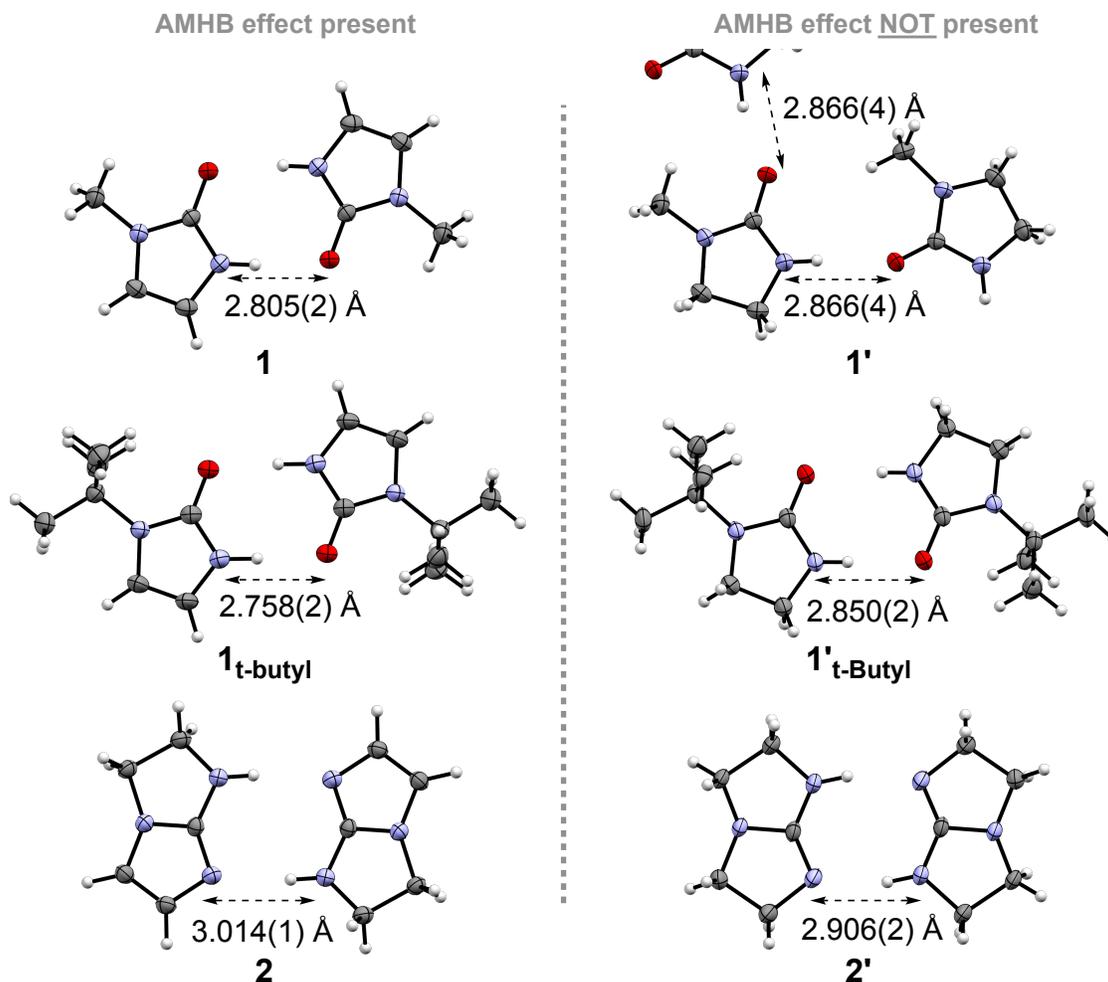


Figure I-17. H-bonded heteroatom distances from X-ray structures of **1**, **1_{t-butyl}** and **2** (left) vs their reference compounds **1'**, **1'_{t-butyl}** and **2'** (right). Oxygen, nitrogen, carbon and hydrogen atoms are red, blue, gray and white respectively.

I.5.2. Charge polarization vs. AMHB: Chemical shifts

The respective downfield and upfield shifts of the methine protons of compounds **1** and **2** upon dimerization qualitatively agree with expected AMHB trends. But these changes reflect not only π -delocalization effects; H-bond-induced polarization (expected to work in the same direction as aromaticity enhancement) must also be considered. For example, the

two methylene groups of compound **1'** shift by +0.502(4) and +0.126(19) ppm upon H-bonding. Thus, the methine peak shifts in **1** and **2** reflect both π -aromaticity perturbations and overall polarization effects (**Figure I-16**, resonance structures) for **1**_{dimer}, **1'**_{dimer}, and **2**_{dimer}; in **2'**_{dimer}, fast degenerate tautomerization averages resonances, complicating this analysis. Nonetheless, all the proton peaks of these compounds shift downfield upon dimerization, *except* those of the methine groups of compound **2**; it is this exception that highlights the role of AMHB.

I.5.3. Hybridization vs. AMHB: Energetics

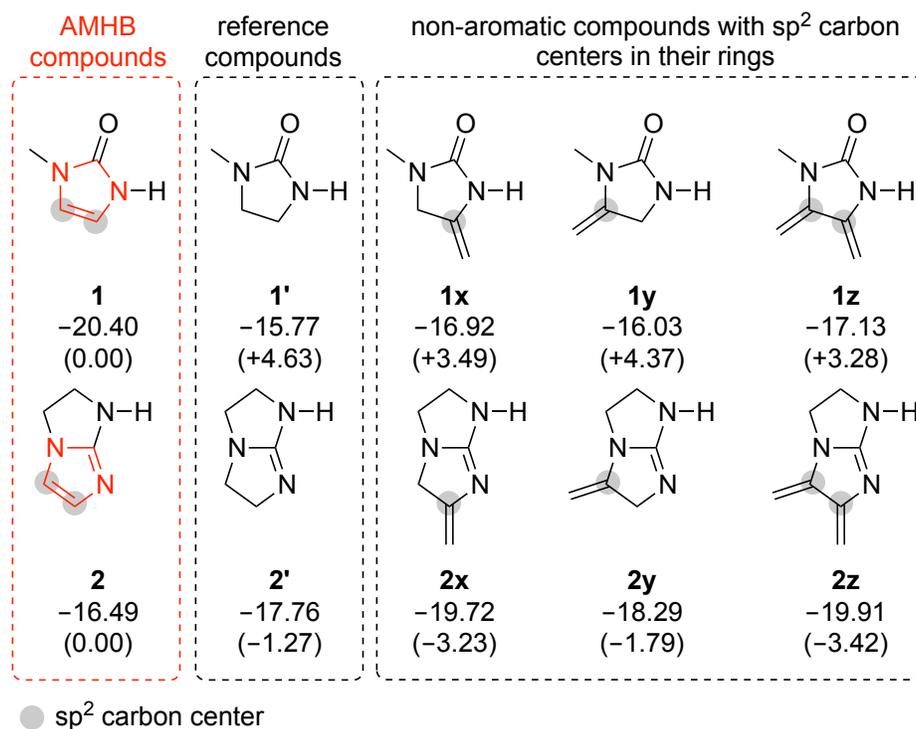


Figure I-18. Gas-phase dimerization electronic energies (kcal/mol). Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.

Comparisons of the computed association energies of **1** and **2** to a series of reference compounds without a cyclic π -conjugated circuit, **1'** and **1x-z** and **2'** and **2x-z** (**Figure I-18**), show that ring carbon hybridizations of the reference compounds have little effect on AMHB.

As shown in **Figure I-18**, the computed $\Delta_{\text{dim}}E$ for **1** (two sp^2 carbons) is consistently more negative than for **1'** (zero sp^2 carbons), **1x** and **1y** (one sp^2 carbon each), and **1z** (two sp^2 carbons), indicating aromaticity gain in **1_{dimer}**, while that of **2** (two sp^2 carbons) is *less* negative compared to **2'** (zero sp^2 carbons), **2x** and **2y** (one sp^2 carbon each), and **2z** (two sp^2 carbons). These comparisons support the notion that aromaticity modulates H-bond strengths in a manner distinct from simple rehybridization effects.

I.5.4. Dipole-dipole interactions vs. AMHB: Energetics

Dipole moment (μ) is another property that changes upon going from compounds **1** and **2** to their corresponding reference compounds besides aromaticity. Thus, to understand the magnitude of the dipole effects on interaction energies vs AMHB, the μ vectors of monomers of **1**, **1'**, **2**, and **2'** were calculated using their geometries found in the optimized C_i or C_{2h} symmetry dimers without changing the Cartesian coordinates of the extracted monomers, at the MP2/aDZ level of theory using wB97X-D/aDZ geometries. This level of theory has been shown to predict dipole moment of a wide range of compounds with a mean absolute error of 0.08 D vs experimental values.⁷⁴ Dipole-dipole distances (r) and angles (θ) were then obtained by center of symmetry operation on the calculated dipole moment vectors. They were then used in the following formula⁷⁵ for antiparallel dipoles to calculate the dipole-dipole interaction energies ($E_{\text{d-d}}$) that are listed in **Table I-2**.

$$E_{\text{d-d}} = \frac{-2\mu(3\cos^2\theta - 1)}{4\pi\epsilon_0 r^3}$$

in which $\epsilon_0 = 8.85 \times 10^{-12}$ F/m is the vacuum permittivity.

Notably, for all four species, the μ values are about 5 D and the angles (θ) relating the antiparallel dipoles to the inter-dipole axes in the dimers are close to 54.74° (the magic

angle) making all the E_{d-d} values small. For **1** and **1'**, these E_{d-d} values are close to zero and essentially negligible, whereas those of **2** and **2'** are slightly attractive. In both cases, the dipole effects are opposite to and substantially smaller than those of AMHB. Thus, dipole-dipole interactions do not account for the dimerization energy differences seen upon hydrogenation.

Table I-2. Calculated dipole-dipole interaction energies (E_{d-d}). This table was adapted from *Angewandte Chemie International Edition*, **2017**, *56*, 9842. Copyright 2017 John Wiley and sons.

Cpd.	d (Å)	θ (°)	μ (D)	E_{d-d} (kcal/mol)
1	5.313	54.52	4.3458	0.02
1'	5.424	55.27	4.7805	-0.05
2	5.171	58.83	4.6188	-0.44
2'	5.299	56.63	3.8340	-0.13

In sum, experimental and theoretical evidence clearly point to significant geometric, energetic, and magnetic effects of AMHB. As illustrated herein, perturbations of aromaticity can modulate the H-bonding energies of heterocycles in the solution phase by as much as one-third. Quantum chemical calculations find that although hybridization and dipole effects are present, the effects due to changes in aromatic character are dominant. The reported NMR data for H-bonding interactions in the solvent medium may also serve as reference values for the evaluation and improvement of theoretical solvation models, enabling more realistic descriptions of molecules and their reactions in solution, which is the subject of the next chapter.

I.6. Resonance-assisted H-bonding (RAHB) and AMHB

Resonance-assisted hydrogen bonding (RAHB) is a concept explaining shorter H-bond distances seen in crystal structures of compounds in which the H-bond donor and acceptor sites are conjugated.⁷⁶ This concept rationalizes the strong and short H-bonds in carboxylic acids, amides, DNA base pairs, etc. as a result of an increased contribution from the zwitterionic resonance form that strengthens H-bonds (**Figure I-19**).

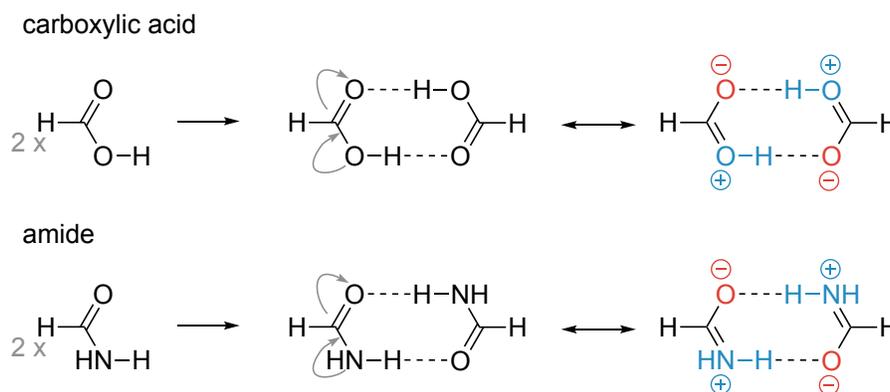


Figure I-19. Two examples of cases of RAHB.

The concepts of RAHB and AMHB are closely related since they both correlate the polarizations of π -orbitals to those involved in H-bonding in the σ -orbitals of H-bond donors and acceptors. Although RAHB is supported by broad surveys of the crystal database by Gilli and co-workers,⁷⁶⁻⁷⁸ it has been a controversial subject mainly, in my opinion, due to the ill-defined reference compounds chosen for the critical studies and ignorance of the effects of (anti)aromaticity in some cases.^{23,79,80} In this section, reference compounds are designed to quantitatively evaluate the magnitude of the RAHB effect.

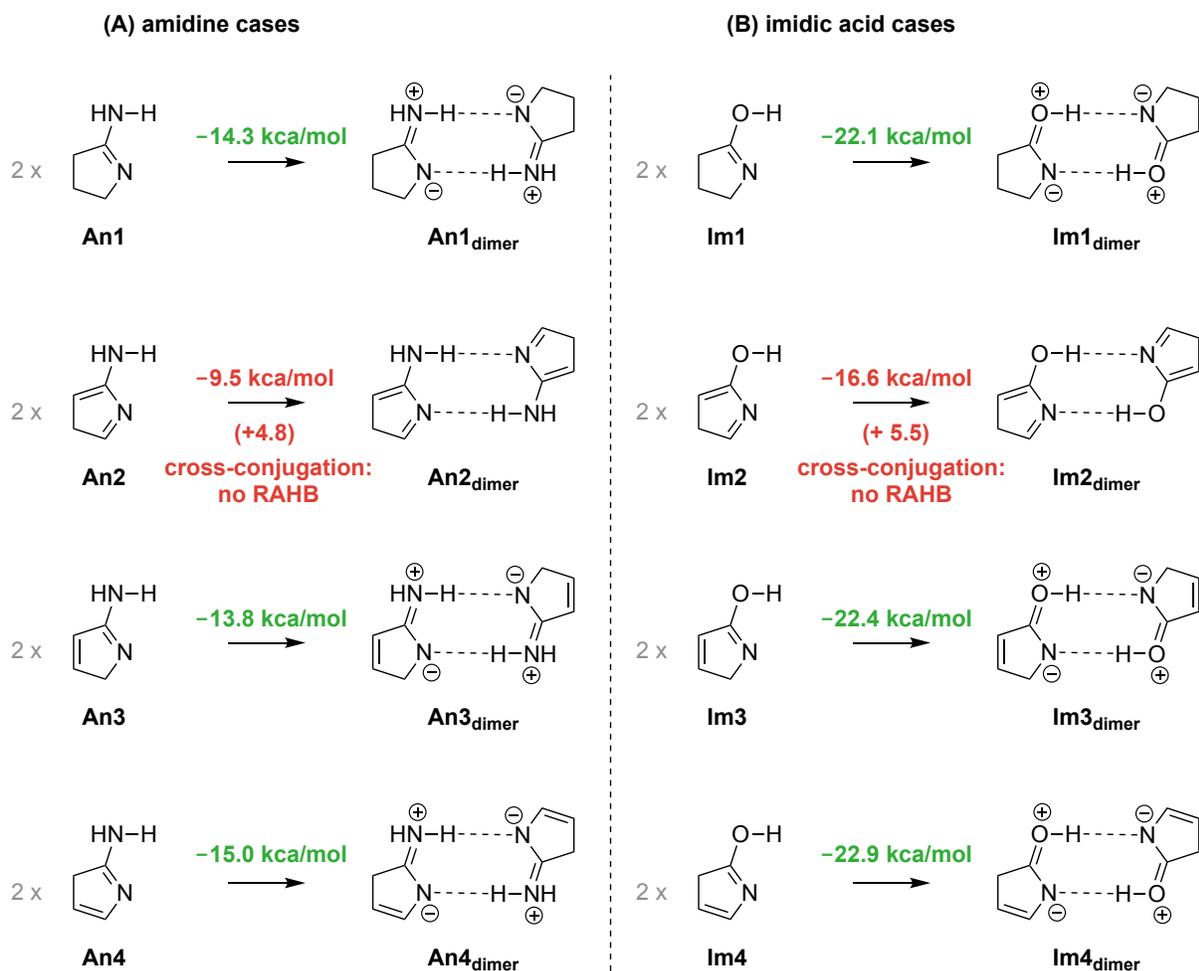


Figure I-20. Two cases for evaluation of the magnitude of the RAHB effect. Energies were calculated at CCSD(T)/CBS using MP2/aDZ geometries.

Upon H-bonding of **An1** (Figure I-20), a resonance form is enhanced in which the charge separation as shown in **An1_{dimer}** favors H-bonding, and thus its H-bond energy should be stronger than that of **An2** in which such a resonance form is not possible due to cross-conjugation. Calculations show that this difference is significant (4.8 kcal/mol). They also show that hybridization changes of the ring carbons is not the main cause of this large difference since **An3** and **An4** have H-bond energies that are barely different from that of **An1**. Similarly, while calculated H-bond energies of **Im1**, **Im3**, and **Im4** are within 1 kcal/mol of each other, they are significantly stronger than that of **Im2**. Specifically, **Im1** is more tightly

bound than **Im2** by 5.5 kcal/mol. **An2** and **Im2** have the same hybridization as **An1** and **Im2** on the atom centers involved in H-bonding respectively, so considering them as reference compounds, the RAHB effect is about 5 kcal/mol for both cases (**Figure I-20**). Therefore, these calculations provide strong evidence for a large RAHB effect at least in these two cases.

I.7. AMHB in redox systems

So far, it has been demonstrated how the H-bond strength of heterocycles can be modulated by their (anti)aromaticity. In this section, the energetic effects of H-bond formation on (anti)aromaticity of fused-ring heterocycles is probed via calculation of changes of reduction energies upon H-bonding. **Figure I-21** (three pages long) shows the computational results. Importantly, π -electron counts vary by ± 2 potentially switching between aromatic and antiaromatic values. Given the opposite effects of AMHB on these two classes, a particularly pronounced modulation by AMHB of redox energy changes may be envisioned. The blue numbers are the H-bond electronic energy of the starting materials and the numbers under the arrows are the reduction electronic energies in the absence of formic acid. Compound **A** has a total of $8\pi e^-$. Based on calculated NICS(1)_{zz} values (shown outside the rings), while ring I is partially antiaromatic, ring II is weakly aromatic. Upon H-bonding, the antiaromaticity of ring I is increased, and aromaticity of ring II is disrupted, suggesting destabilization of the π -system upon H-bonding. On the other hand, the hydrogenated version of **A**, **A_{H2}**, has two aromatic rings both of which increase in aromaticity upon H-bonding to formic acid. As a result of these (anti)aromaticity changes, the H-bond energy of **A_{H2}** is 3.84 kcal/mol stronger than that of **A**. This means the reduction of **A** to **A_{H2}** is favored by 3.84 kcal/mol in the presence of H-bonding to formic acid. This would translate into a difference in 2-electron reduction potential of 83 mV. Similar analysis can be used for

explaining the observed effect in the rest of compounds **A-H** in which effects as large as 6.37 kcal/mol are observed. In all these cases, the enhancement of aromaticity on both rings in the reduced product, and the enhancement of antiaromaticity on one or both rings in the starting upon H-bonding, favors the reduction in the presence of formic acid. Notably, in cases **D**, **E**, and **G**, in which the presence of two electron withdrawing groups (EWG) of carbonyls on ring II localizes its electrons, the ring I behaves like an isolated 2-pyridone and enhances in aromaticity upon H-bonding resulting in a smaller effect. This is supported by the upfield shifts of the calculated $\text{NICS}(1)_{zz}$ values on ring I in all these cases.

I, **J** and **M** behave simply since the 2-pyridone ring is effectively isolated due to the presence of two carbonyl groups on ring II. As a consequence, H-bonding leads to enhancement of aromaticity on ring I and relief of antiaromaticity on ring II which leads to a stronger H-bonding than their reduced forms in which the aromaticity of both rings is enhanced upon H-bonding. This leads to weakly disfavoring of the reduction reaction in the presence of formic acid (the left column, small positive values). However, in the rest of the cases **I-N**, in which the two rings are strongly conjugated, the fused ring behaves as an antiaromatic $8\pi e^-$ system increasing in antiaromaticity upon H-bonding, which favors reduction in the presence of formic acid. In all the **I-N** cases, the rings do not act individually, and the aromaticity of both rings is enhanced upon H-bonding. This is presumably due to electron donating OH groups pushing electrons towards the electron deficient 2-pyridone rings. The very small or no effects are also supported by the calculated small $\Delta\text{NICS}(1)_{zz}$ values. Cases **O-U** can be analyzed in a similar fashion.

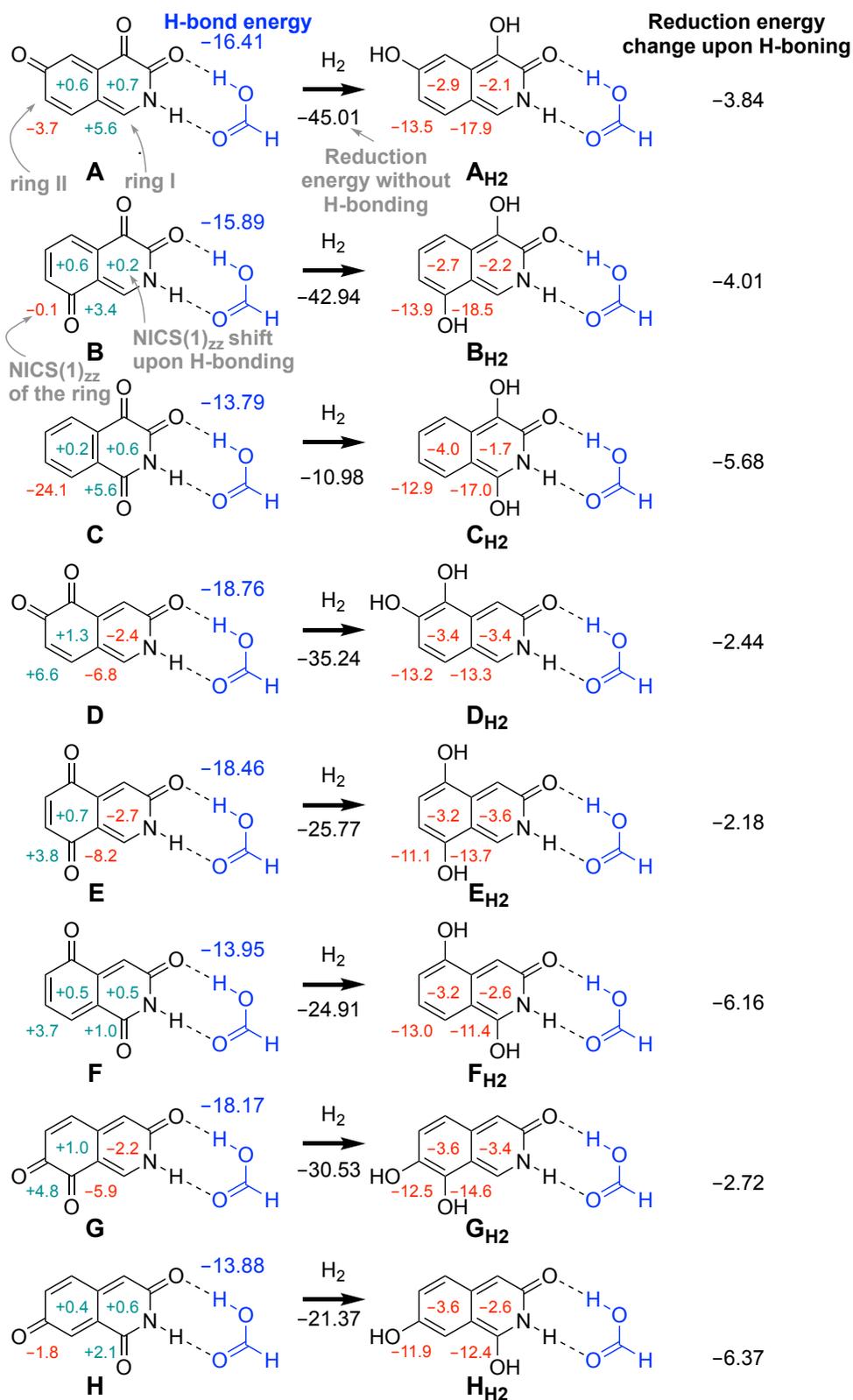


Figure I-21. Reduction energy modulations via AMHB. Energy (in kcal/mol) and NICS(1)_{zz} (in ppm) values are respectively calculated at CCSD(T)/CBS and mPW91PW91 using MP2/aDZ geometries.

Figure I-21 (cont'd)

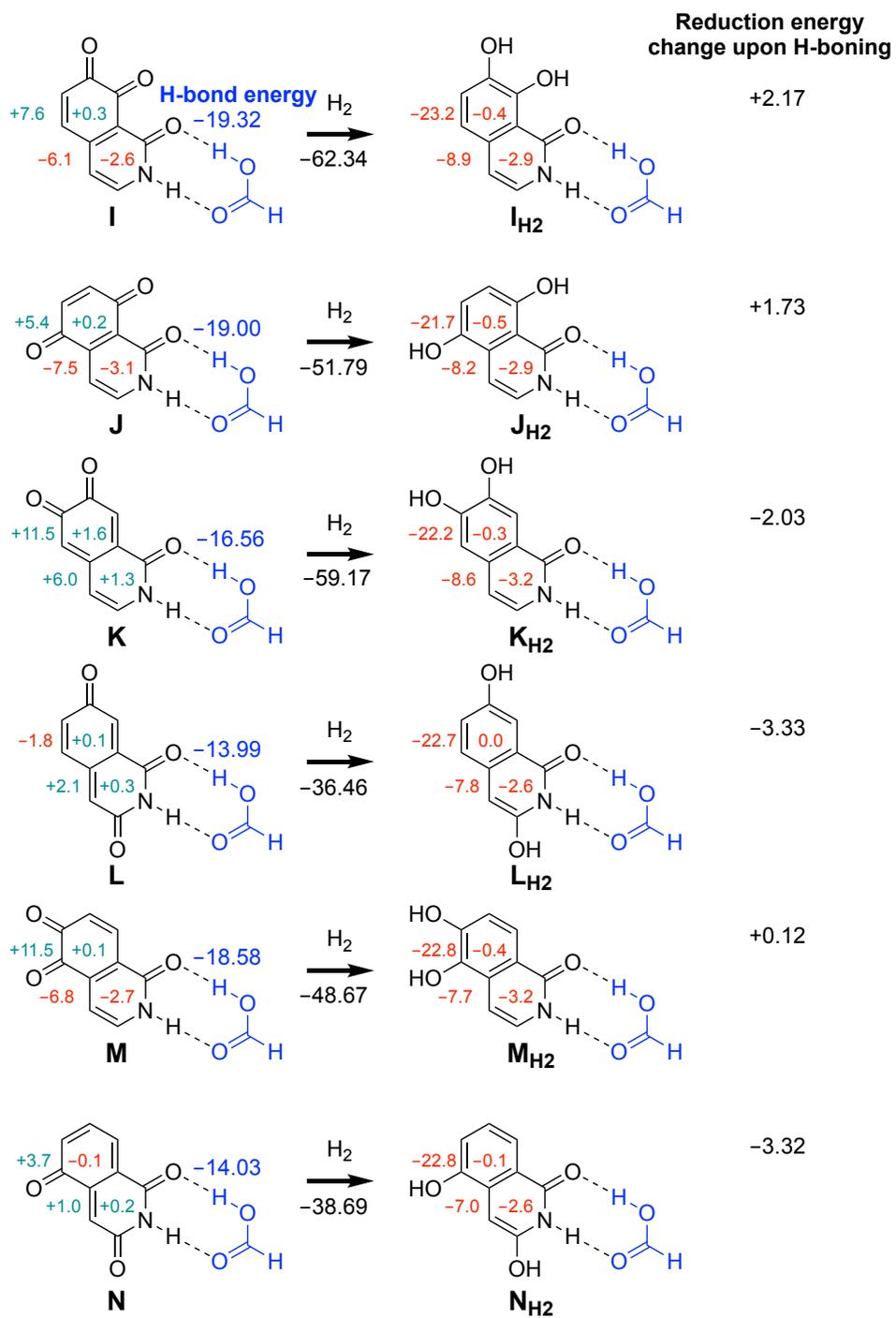
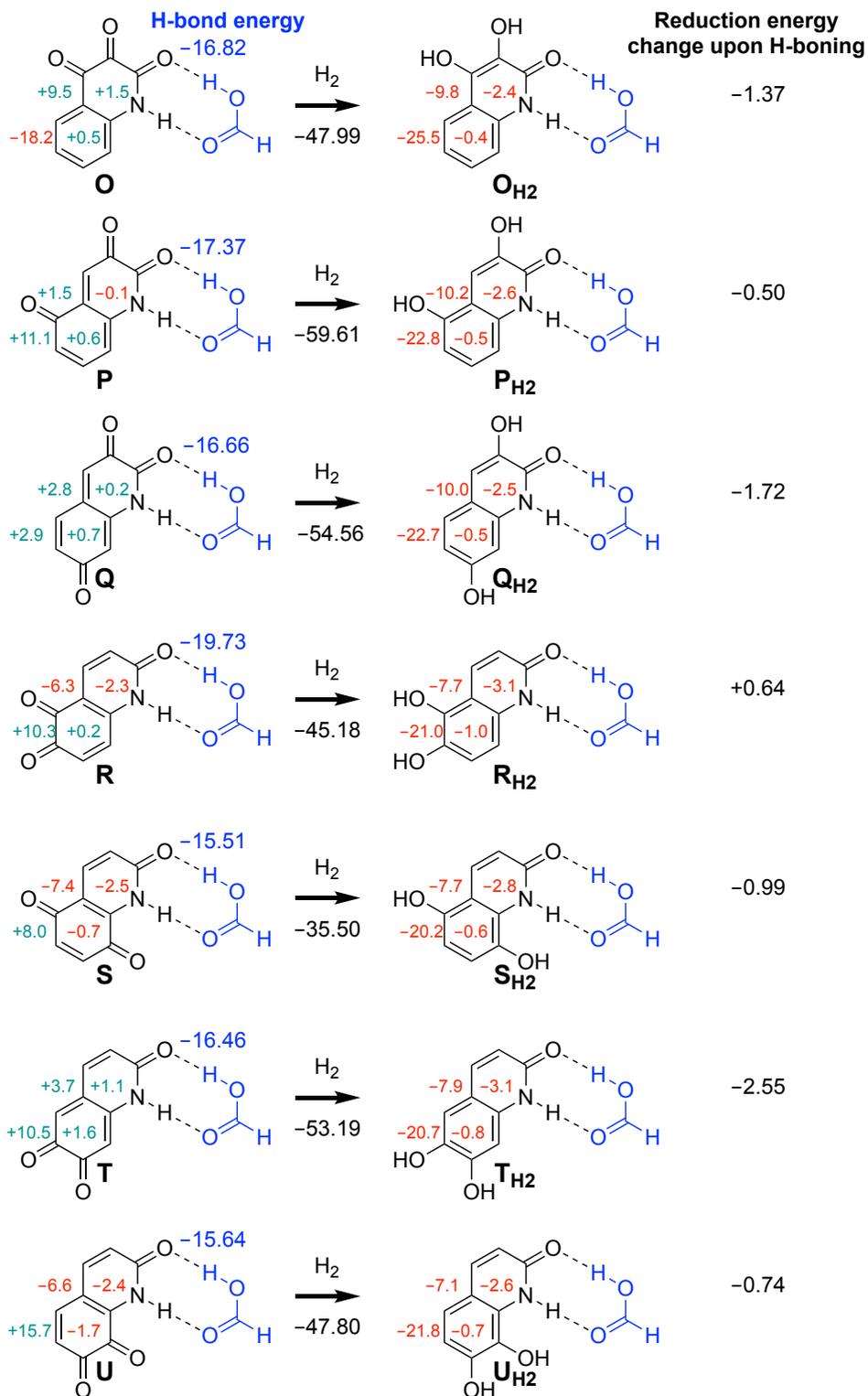


Figure I-21 (cont'd)



These calculations clearly show that while AMHB can affect the redox chemistry of fused rings, the interpretation of the results may be less straightforward, especially in cases where the comprising rings' π -systems are not strongly coupled and act as individual ones.

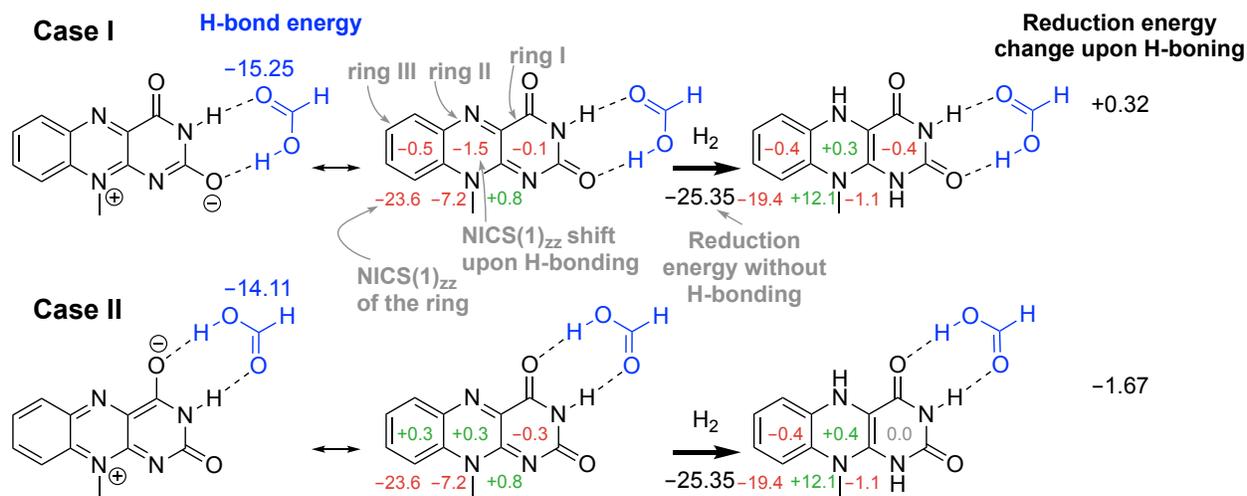


Figure I-22. H-bonding modulation of flavin upon H-bonding.

One biologically relevant case for which the redox chemistry could be affected by H-bond formation via AMHB is flavin. Flavin can H-bond in two modes which are shown as cases I and II (**Figure I-22**). In case I, the aromaticity of rings II and III is increased upon H-bonding. This is consistent with the calculated negative $\Delta\text{NICS}(1)_{zz}$ values for these rings (see numbers inside the corresponding rings). A resonance form consistent with these magnetic changes is shown in **Figure I-22**. In this resonance form, the sum of rings II and III are similar to naphthalene, a bicyclic $10\pi e^-$ aromatic compound. This stabilization causes the reduction reaction to be favored towards the starting flavin, i.e. disfavoring the reduction by 0.32 kcal/mol. Conversely, in case II, H-bond formation favors a resonance form in flavin in which the rings II and III are not aromatic. This is consistent with the calculated $\Delta\text{NICS}(1)_{zz}$ values (see numbers inside rings). As a result, the starting material's π -system is destabilized leading to a more favored reduction reaction by 1.67 kcal/mol. Thus, the active

site binding mode found in the protein pockets of various flavin binding enzymes can potentially adjust the redox potential over this nontrivial range.

I.8. Conclusions

In conclusion, the generality of AMHB was demonstrated using a broad computational survey of heterocycles via calculations of energetic and magnetic effects of AMHB. Then, experimental measurements provided energetic, geometric and magnetic evidence for AMHB. These studies hinged on the clear definition of reference compounds in which the H-bonding sites were the same as (anti)aromatic heterocycles subjected to AMHB, except that their cyclic π -systems were breached via hydrogenation of rings' π -bonds. The concept of AMHB was then used to explain remote H-bond energy modulations of bicyclic aromatic heterocycles. Calculations of the reduction of a wide range of heterocycles show that the energy modulations of cyclic π -systems upon H-bonding can be used to explain energy variations among redox active heterocycles such as flavin.

I.9. Computational and experimental details

I.9.1. Computational details

All the geometries in this chapter were optimized at the MP2⁸¹⁻⁸⁵/aDZ⁸⁶⁻⁹⁰ level of theory, with the core electrons frozen, unless otherwise mentioned. Frequency calculations confirmed the stationary nature of the minima and the default geometry convergence criteria in Gaussian16⁹¹ after frequency calculations were used to confirm complete optimization. These default values for Maximum Force, RMS Force, Maximum Displacement, and RMS Displacement were 4.5×10^{-4} , 3.0×10^{-4} , 1.8×10^{-3} , and 1.2×10^{-3} respectively. In cases where not all the four criteria were converged, if the Maximum Force and RMS Force were smaller than 4.5×10^{-4} and 3.0×10^{-4} the geometries were accepted as optimized.

This exception was only used in cases where even optimizations with calculation of Hessian at each point did not resolve the issue and Gaussian16 itself accepted those geometries. In some cases, some dimers could form in two possible symmetries, C_i and C_2 . In those cases, the dimers were optimized in both symmetries and the lowest energy dimer was used for calculation of H-bond energies, after the CCSD(T)⁹²⁻⁹⁶ calculations described below.

Using the optimized geometries, the wavefunction-based scheme recommended by Burns *et al.*³⁶ was used to obtain reference H-bonding interaction energies in the gas-phase.

$$E_{approx.}^{CCSD(T)} = E_{elec.}^{DF-HF/aQZ} + E_{extrapol.}^{DF-MP2/aQ5Z} + \delta_{MP2/aDZ}^{CCSD(T)/aDZ}$$

The scheme is based on density-fitted⁹⁷⁻¹⁰⁵ Hartree-Fock (DF-HF)¹⁰⁶ electronic energies calculated at DF-HF/a5Z, $E_{elec.}^{HF/aQZ}$, DF-MP2 correlation energies extrapolated from aQZ to a5Z (abbreviated as DF-MP2/aQ5Z), $E_{extrapol.}^{MP2/aQ5Z}$, and CCSD(T)/aDZ additional correlation energies, $\delta_{MP2/aDZ}^{CCSD(T)/aDZ}$, added on top of those of DF-MP2. The sum of these three components gives an approximation of the CCSD(T) energies extrapolated to the complete basis limit (CBS), denoted as $E_{approx.}^{CCSD(T)}$ in the above equation. The extrapolation scheme used for the DF-MP2 calculations is that of Halkier *et al.*¹⁰⁷ in the form below

$$E_n^{corr.} = E_\infty^{corr.} + An^{-3}$$

in which n is the cardinal number of the Dunning basis set, $E_n^{corr.}$ is the calculated correlation energy using that basis set, A is a constant, and $E_\infty^{corr.}$ is the extrapolated correlation energy at cardinal number of infinity. By using two basis sets with consecutive cardinal numbers of n and $n+1$, the equation can be solved for $E_\infty^{corr.}$.

$$E_\infty^{corr.} = \frac{(n+1)^3 E_{n+1}^{corr.} - n^3 E_n^{corr.}}{(n+1)^3 - n^3}$$

In our case where cardinal numbers $n = 4$ and $n + 1 = 5$ are used, the equation simplifies as follow

$$E_{\text{extrapol.}}^{\text{DF-MP2/aQ5Z}} = E_{\infty}^{\text{corr.}} = \frac{125 \times E_{\text{a5Z}}^{\text{corr.}} - 64 \times E_{\text{aQZ}}^{\text{corr.}}}{61}$$

After calculating the $E_{\text{approx.}}^{\text{CCSD(T)}}$ for each monomer and dimer, the electronic dimerization energies, $\Delta_{\text{dim}}E$, were calculated as

$$\Delta_{\text{dim}}E = E_{\text{approx.}}^{\text{CCSD(T)}}(\text{dimer}) - 2 E_{\text{approx.}}^{\text{CCSD(T)}}(\text{monomer})$$

in which the denotations are obvious.

I.9.2. Details of NMR measurements

Compounds **1**, **1'**, **2**, and **2'** were sublimed and the sublimation apparatus was opened inside the nitrogen filled the dry box. The anhydrous C_6D_6 containing tetramethylsilane was prepared as follows: Tetramethylsilane (1 μL , NMR grade, Across Organics) and 3 Å molecular sieves (30 % v/v, Sigma-Aldrich) were added to a 10-gram bottle of C_6D_6 (99.5 % deuterated, Cambridge Isotope Laboratories) and the bottle was sealed for two days prior to the experiments. The C_6D_6 bottle was also opened inside the dry box, where all the samples were made.

Thick-walled volumetric flasks (Corning Pyrex, 5- and 10 mL) were dried in an oven at ca. 140 °C for about 2 days followed by 2 hours at 300 °C. This drying procedure did not change the calibration of the flasks (vide infra). NMR tubes were dried at ca. 140 °C for 36-48 hours. All the above items were transferred to the vacuum chamber of the dry box, while still hot, and were kept under 100 mTorr for overnight. Syringes (Hamilton, 50- and 500 μL) were dried at room temperature for overnight and were also kept under 100 mTorr for overnight. Samples were made by sequential dilution of a concentrated solution of the

analyte. After the solutions were made, the NMR tubes were capped with regular polyethylene caps, and the tops were wrapped with parafilm.

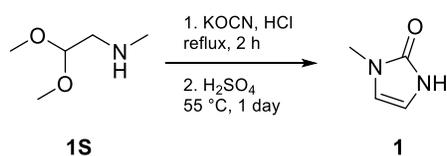
All the volumetric glassware and syringes were calibrated, and their errors were determined by weighing water on a METTLER AE200 balance. A statistical sample population of 10 was used for all the standard deviation determinations. The balance showed a repeatability of ± 0.1 mg in our lab environment, which makes it suitable for calibration of the volumetric flasks and the 500- μ L syringe but not the 50- μ L one. Standard deviations of 7.5 and 15.7 μ L was obtained for 5- and 10-mL flasks respectively, which are in agreement with the ± 20 μ L tolerance provided in the supplier catalog. This leads to an uncertainty less than 0.5 % in volumetric measurements carried out with these flasks. The 500 μ L syringe however showed a standard deviation of 0.8 μ L for measuring 100 μ L, the smallest volume it was used for, which is again in good agreement with the 1% error provided in the supplier catalog. The 50 μ L syringe gave standard deviation of 0.1 μ L, which is in the range of repeatability of the balance. Since we were not able to estimate the error of the 50 μ L syringe, we assumed the standard deviation of 0.1 μ L which is provided by the supplier. Considering all the above standard deviations, errors of roughly up to 1% and 2% should be expected for the mM and μ M concentration ranges, respectively.

I.9.3. Synthesis of compounds 1, 1', 2 and 2'

^1H NMR spectra were recorded on an Agilent DDR2 500 MHz spectrometer. Chemical shifts are reported in ppm from tetramethylsilane with the residual non-deuterated solvent resonance as the internal standard (CHCl_3 : δ 7.26 ppm). Data are reported as follows: chemical shift (multiplicity, coupling constant, number of hydrogens). Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. ^{13}C

NMR spectra were also recorded on the above Agilent DDR2 500 MHz spectrometer with complete proton decoupling. Chemical shifts are reported in ppm from tetramethylsilane with the solvent resonance as the internal standard (CDCl₃: δ 77.16 ppm). Infrared (IR) spectra were recorded on a JASCO FT/IR 6600 type A (ATR mode) spectrophotometer. Peaks are reported in cm⁻¹.

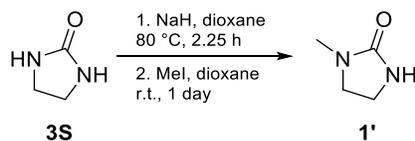
All synthetic procedures were carried out using reagent grade solvents in air unless otherwise mentioned. Reagents and chemicals were used as arrived, without further purification. Diethylenetriamine, sodium hydride (60% in mineral oil), 2,2-dimethoxy-N-methylethan-1-amine, 2-imidazolidinethione, imidazolidin-2-one, and 2,2-dimethoxyethan-1-amine were purchased from Sigma-Aldrich. CS₂ was purchased from Matheson Coleman and Bell. Methyl iodide was purchased from Alfa Aesar.



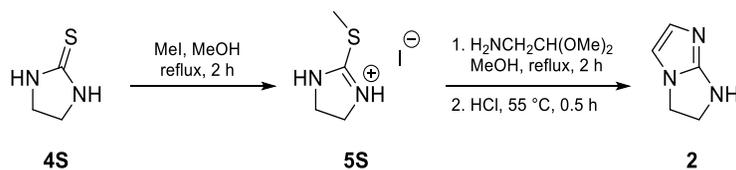
1-methyl-1,3-dihydro-2H-imidazol-2-one **1** (method A). This procedure was adapted from the literature.¹⁰⁸ A solution of KOCN (1.20 g, 14.79 mmol) in water (2.6 mL) was added to a solution of **1S** (1.20 g, 10.07 mmol) in HCl (4.0 mL, 3 M). The resulting solution was refluxed for 2 hours. Then, it was cooled down, saturated with sodium chloride, and extracted with dichloromethane (4 x 5 mL). Evaporation of all volatiles to dryness yielded a white solid (1.10 g). The solid was then dissolved in H₂SO₄ (1.1 mL, 0.25 N) and stirred at 57 °C for 2.2 hours. After that the solution was cooled down, saturated with NaCl (0.5 g), and extracted with dichloromethane (5 mL). After drying with magnesium sulfate, evaporation of all volatiles and subsequent sublimation (80 °C, 20 mTorr), compound **1** was obtained as

a white solid (70 mg, 7% isolated yield). Characterization data were in accordance with literature.¹⁰⁹ ¹H NMR (500 MHz, CDCl₃) δ 11.12 (s, 1H), 6.27 (t, J = 2.6 Hz, 1H), 6.16 – 6.11 (m, 1H), 3.24 (s, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 155.26, 112.51, 108.41, 30.00. IR (neat): 3107, 3030, 2994, 2843, 2809, 1655, 1469, 1402, 1280, 909, 785, 671, 531 cm⁻¹. Single crystals suitable for X-ray diffraction were grown by slow evaporation of a solution in benzene.

1-methyl-1,3-dihydro-2H-imidazol-2-one **1** (method B). This method is an improved version of method A. A solution of KOCN (1.20 g, 14.79 mmol) in water (2.6 mL) was added to a solution of **1S** (1.20 g, 10.07 mmol) in HCl (4.0 mL, 3 M). The resulting solution was refluxed for 30 minutes. Then, it was cooled down, saturated with sodium chloride (1.1 g), and extracted with dichloromethane (3 x 10 mL). After the organic phase was dried with MgSO₄ and evaporated to dryness, a white solid (1.6 g) was obtained. The solid was then dissolved in dichloromethane (5 mL), trifluoroacetic acid (5 mL) was added, and the resulting solution was refluxed for 20 hours. After that, all the volatiles were evaporated under a stream of nitrogen, the residue was taken up in chloroform (30 mL), stirred with potassium carbonate (5 g) for 30 minutes, and the solids were filtered off. The neutralization with potassium carbonate was repeated, and all the volatiles were evaporated to obtain a yellowish white solid. After sublimation, a white solid (610 mg, 62%) was obtained. ¹H NMR (500 MHz, CDCl₃) δ 11.08 (s, 1H), 6.28 (t, J = 2.6 Hz, 1H), 6.14 (dd, J = 2.8, 2.2 Hz, 1H), 3.25 (s, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 155.21, 112.56, 108.44, 77.41, 77.16, 76.91, 30.03.



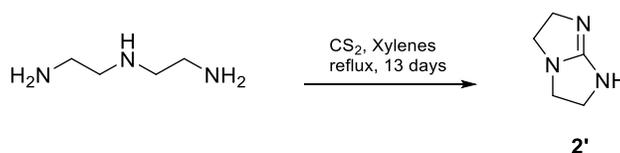
1-methylimidazolidin-2-one 1'. This procedure was adapted from the literature.¹¹⁰ A mixture of 2-imidazolidinone **3S** (1.52 g, 96%, 16.95 mmol), sodium hydride (0.84 g, 60 % in mineral oil, 21.00 mmol), and 1,4-dioxane (20 mL, kept over 3 Å sieves for overnight) was stirred at 80 °C for 2.25 hours under an inert atmosphere of nitrogen. After that, the mixture was cooled down in an ice-water bath, and methyl iodide (2.0 mL, 31.80 mmol) was added. After stirring for one more day at ambient temperature, the solids were filtered off, and the filter cake was washed with more 1,4-dioxane (20 mL). Evaporation of all volatiles to dryness yielded crude **1'** which was purified via column chromatography using silica as the stationary phase and 10 % v/v methanol in dichloromethane as the eluent. Compound **1'** was isolated as a white solid (0.47 g, 28% isolated yield). Characterization data were in accordance with literature.¹¹⁰ ¹H NMR (500 MHz, CDCl₃) δ 6.02 (s, 1H), 3.24 (s, 4H), 2.61 (s, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 163.38, 47.10, 37.62, 30.18. The pure compound was sublimed 60 °C under 50 mTorr and opened in the dry box.



2,3-dihydro-1H-imidazo[1,2-a]imidazole 2. This procedure was adapted from the literature.¹¹¹ Methyl iodide (14.9 g, 105.0 mmol) was added to a solution of **4S** (9.82 g, 96.1 mmol) in methanol (20 mL), and the resulting mixture refluxed for 2 hours. After that, the solution was cooled down to room temperature, and ether (20 mL) was added. The precipitated white solid was filtrated and washed with more ether (50 mL). Recrystallization from methanol (15 mL) yielded **5S** as a white solid (15.5 g, 66%). Characterization data were in accordance with literature.¹¹¹ ¹H NMR (500 MHz, DMSO-*d*₆) δ 9.91 (s, 2H), 3.87 (s, 4H),

2.62 (s, 3H). ^{13}C NMR (126 MHz, $\text{DMSO-}d_6$) δ 170.45, 45.37, 14.01.

A solution of **5S** (1.50 g, 6.2 mmol) and $\text{H}_2\text{NCH}_2\text{CH}(\text{OMe})_2$ (1.50 g, 14.3 mmol) in methanol (2.0 mL) was refluxed for 12 hours. After that, all the volatiles were evaporated and the resulting oil was further dried by heating at 110 °C under high vacuum for one hour. The residual gel was then dissolved in concentrated HCl (1.9 mL) and stirred at 55 °C for half an hour. After that, it was cooled down in an ice-water bath, and concentrated ammonium hydroxide (5.0 mL) and sodium chloride (4.0 g) were added, respectively. Insoluble solid was filtered off and the aqueous phase was extracted with dichloromethane (3 x 10 mL), and subsequently dried with magnesium sulfate. Evaporation of all volatiles to dryness yielded crude **2** which was purified via column chromatography using silica as the stationary phase and dichloromethane: triethylamine: methanol (7:2:1) as the eluent. After sublimation (80 °C, 50 mTorr), compound **2** was achieved as a white solid (0.21 g, 31% isolated yield). ^1H NMR (500 MHz, CDCl_3) δ 6.64 (d, $J = 1.6$ Hz, 1H), 6.54 (d, $J = 1.6$ Hz, 1H), 4.84 (s, 1H), 4.01 – 3.88 (m, 4H). ^{13}C NMR (126 MHz, CDCl_3) δ 158.40, 129.33, 110.36, 77.16, 76.90, 76.65, 47.94, 43.27. Single crystals were grown by slow evaporation of a solution in benzene. IR (neat): 3147, 3108, 2973, 2899, 1574, 1514, 1480, 1325, 1286, 1264, 1203, 1161, 1096, 1061, 1032, 939, 876, 848, 829, 738, 685, 620 cm^{-1} .



2,3,5,6-tetrahydro-1H-imidazo[1,2-a]imidazole 2'. This procedure was adapted from the literature.¹¹² CS_2 (12.0 mL, 0.20 mol) was added to a stirring solution of diethylenetriamine (20.6 g, 0.20 mol) in xylenes (300 mL), and the resulting mixture was

refluxed under an inert atmosphere of nitrogen for 13 days. After that, about two-third of the solvent was evaporated by blowing nitrogen on top of the condenser. The remaining solution was then decanted into an Erlenmeyer flask while still hot and cooled down slowly overnight to precipitate a brownish white solid (15.3 g). Sublimation (60 °C, 50 mTorr) of a portion of this solid (1.00 g) yielded compound **2'** as a yellowish white solid (0.50 g). Characterization data were in accordance with literature.¹¹² ¹H NMR (500 MHz, CDCl₃) δ 6.76 (s, 1H), 3.76 (t, J = 7.0 Hz, 4H), 3.01 (t, J = 7.1 Hz, 4H). ¹³C NMR (126 MHz, CDCl₃) δ 171.52, 52.70, 49.53.

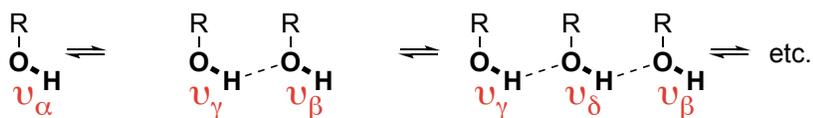
CHAPTER II: EVALUATION OF IMPLICIT SOLVATION MODELS FOR PREDICTING H-BOND FREE ENERGIES

II.1. Introduction

Accurate methods to predict the strengths of H-bonds are of value in the design and analysis of molecular systems across various fields of chemistry. The development of such tools relies on systematic datasets of accurately measured solution phase H-bond energies. However, many simple compounds of interest, such as alcohols, aggregate via multi-species equilibria in solution, placing H-bonds in a complex mixture of different settings. To probe H-bonding, no common spectroscopic tool is ideal: fast exchange averages NMR peaks for H-bonding species, while IR spectra show broad overlapping bands with widely varying frequency-dependent extinction coefficients. The above issues have hindered quantitative molecular-level determination of association energetics for most H-bonds among simple compounds such as water, alcohols, or amines in solution (**Scheme II-1-A**).

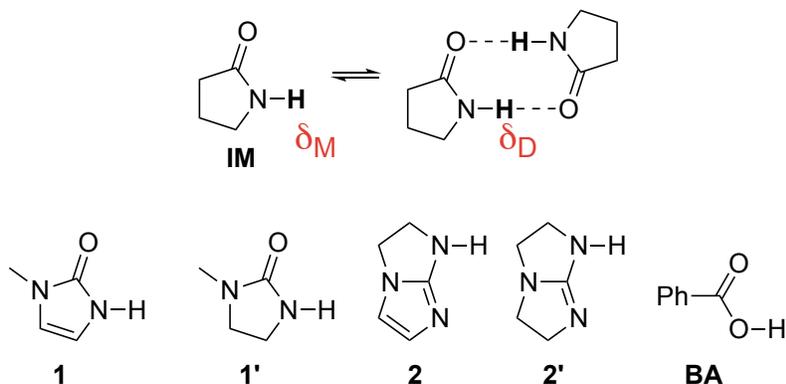
A number of H-bonded small-molecule systems have been studied in the gas-phase using specialized spectroscopic methods.¹¹³ Modern high-level quantum chemical protocols can compute energetics for such gas-phase systems at levels of accuracy that rival experiment.^{114,115} Thus, high-level calculations generally serve well as reference methods for gas-phase systems (*vide infra*).^{116,117} Yet the condensed phase is the context of broadest interest. Recent years have seen the development of sophisticated and general continuum methods to simulate solvent environments. These methods are mostly optimized to compute free energies of solvation.¹¹⁸ In this chapter, H-bonded association energies as another metric to assess the performance of such tools are explored.

A) Complicated hydrogen bonding equilibria of alcohols



At least five overlapping peaks in infrared spectra or averaged peaks in NMR spectra.

B) Simple dimerization equilibria studied here



Scheme II-1. (A) Complications from H-bonding polymerization in alcohols. (B) H-bonding species studied here.

One way to get around the measurement challenges noted above is to design systems that establish simple “ $2 \times \text{monomer} \rightleftharpoons \text{dimer}$ ” equilibria in solution (**Scheme II-1-B**). Systems chosen for this study only form dimers, mostly via two-point H-bond homodimerization of amide or carboxylic acid functional groups (six cases). Though tautomerization to the 2-hydroxypyridine structure complicates the analysis of 2-pyridone, for the amides examined here, the potential tautomeric forms are strongly disfavored, as confirmed by quantum chemical simulations. In rigorously dry solvents, the systems studied herein show simple monomer-dimer equilibria that accurately fit an analytical expression for the observed NMR data, enabling extraction of thermodynamic and extrapolated spectroscopic parameters. While NMR instruments with common field values and probes are limited in sensitivity for measurements of strongly bound systems, MSU's 900 MHz NMR

spectrometer, equipped with a cryogenic probe, enabled accurate quantification of H-bonds as strong as 4.5 kcal/mol.

The solvents chosen for this study were benzene, dichloromethane, chloroform, and acetonitrile covering a range of dielectric constants from 2.3 to 37.5. These are among the most common organic solvents, and their polarity ranges from strongly polar to hydrophobic, roughly the range seen in the binding sites of various proteins. H-bonding solvents and those with exchangeable protons were not considered here. The computational survey includes a wide range of common popular and newer density functional methods, with solvation effects incorporated using the implicit solvation models IEF-PCM C-PCM, and SMD models.¹¹⁸⁻¹²¹ Gas-phase benchmark studies have shown that some DFT methods can compete in accuracy (≤ 0.5 -1.0 kcal/mol) with highly correlated wavefunction-based levels of theory.¹²² The use of continuum solvation models entails manageably small numbers of modest-sized calculations, obviating the need for explicit quantum chemical calculations on the dynamically variable clusters of solvent molecules that form the solvation shell of a given solute. Therefore, a survey of combinations of DFT methods and continuum solvation models constitutes the main part of this study.

II.2. Experimental reference H-bond energies

As mentioned earlier, NMR spectroscopy was used to measure solution phase H-bond energies. Samples were prepared at concentrations chosen to access a range of monomer:dimer ratios, and measured at four temperatures in the range of 5-35 °C. From these measurements, enthalpy, entropy and free energies of dimerization were extracted. The details of procedures and data fitting are similar to those presented in chapter I of this

dissertation. The N-H or O-H proton chemical shifts were used as the primary peaks for extracting thermodynamic data for **IM** and **BA** respectively.

Table II-1. Measured H-bond thermodynamic values and extrapolated monomer and dimer chemical shifts.

Compound or solvent	Δ_{dimG} (kcal/mol)	Δ_{dimH} (kcal/mol)	Δ_{dimS} (cal/molK)	δ_{M} (ppm)	δ_{D} (ppm)	K_{eq} (1/M)
Benzene						
1	-4.52	-8.94	-14.85	5.682	12.531	2044.50
1'	-1.98	-6.57	-15.40	3.221	8.070	28.08
2	-2.09	-6.57	-15.03	2.897	8.169	34.06
2'	-3.50	-8.96	-18.30	3.108	9.564	370.17
IM	-2.17	-7.30	-17.22	4.165	9.079	38.93
BA	-3.78	-9.18	-18.1	5.864	13.981	593.29
Chloroform						
1	-2.82	-5.84	-10.11	6.956	11.637	117.52
1'	-0.40	-3.39	-10.02	4.137	7.185	1.98
2	-0.06	-1.99	-6.48	3.870	6.991	1.10
2'	-0.87	-4.41	-11.88	4.001	8.735	4.31
IM	-0.62	-3.63	-10.08	5.329	8.573	2.85
BA	-3.48	-8.16	-15.68	6.794	13.383	355.86
Dichloromethane						
1	-2.70	-6.59	-13.04	6.948	11.599	95.76
1'	-0.19	-4.01	-12.8	4.095	7.248	1.39
2	-0.33	-3.33	-10.07	3.851	7.228	1.74
2'	-1.54	-5.48	-13.2	4.021	8.496	13.46
IM	-0.42	-4.44	-13.48	5.287	8.5044	2.03
BA	-3.14	-8.18	-16.91	6.977	13.390	200.28
Acetonitrile						
1	-0.98	-4.43	-11.57	7.797	11.413	5.2
1'	0.65	-3.03	-12.34	4.618	7.175	0.33
2	0.45	-2.71	-10.59	4.305	7.196	0.47
2'	-0.84	-5.87	-16.88	4.516	8.824	4.12
IM	0.62	-3.09	-12.41	5.827	8.445	0.35
BA^a	1.44	-3.81	-17.6	9.391	18.542	0.09

^aLocking down the dimer chemical shift to 13.40 ppm and its temperature dependence to 6.1 ppb/K gave: $\Delta_{\text{dimG}} = +0.84$ kcal/mol, $\Delta_{\text{dimH}} = -1.75$ kcal/mol, $\Delta_{\text{dimS}} = -8.71$ cal/mol/K, $\delta_{\text{M}} = 9.38$ ppm, and $d\delta_{\text{M}}/dt = -5.8$ ppb/K

Table II-1 presents the thermodynamic data extracted for all compounds studied. As expected, going from non-polar to polar solvents, the dimerization energies weakened for all compounds. The dimerization energies in CDCl_3 and CD_2Cl_2 are very similar. Interestingly, the AMHB effect^{34,123,124} manifests itself in dimerization energies across all solvents, even though its effect is smaller in the more polar CD_3CN solvent. The free energies were used as reference points as they are reported in this table.

II.3. CCSD(T)/CBS: a reference for gas-phase H-bond energies

The wavefunction-based scheme recommended by the Sherrill group³⁶ was used to obtain reference H-bonding interaction energies in the gas phase. The geometries used for these calculations were fully optimized at MP2/aDZ using the Gaussian16 software. As described in detail in the previous chapter, the CCSD(T)/CBS energy is the sum of the DF-HF/a5Z total electronic energy, DF-MP2 correlation energy extrapolated from aQZ to a5Z basis set, and CCSD(T) correlation energy calculated using the aDZ basis set. DF-MP2 correlation energy is the difference between DF-MP2 and DF-HF total electronic energies calculated using the same basis set. Likewise, the CCSD(T) correlation energy is the difference between the CCSD(T) and MP2 (not DF-MP2) energies calculated at the same basis set. Both CCSD(T) and DF-MP2 correlation energies are negative. **Table II-2** shows the DF-HF,/a5Z extrapolated DF-MP2/aQ5Z, and CCSD(T)/aDZ interaction energies as well as the estimated CCSD(T)/CBS values for all the dimers studied here in various symmetries.

While DF-HF interaction energies were off by a few kcal/mol, extrapolated DF-MP2/aQ5Z and CCSD(T)/aDZ were close to the CCSD(T)/CBS reference methods, with the former performing slightly better. The poorer performance of CCSD(T)/aDZ is likely due to

slow convergence of the correlation energy and use of the smallest aDZ Dunning basis set.¹⁰⁷

The CCSD(T)/CBS energies were used as reference values in the following discussion.

Table II-2. Calculated H-bond electronic energies at CCSD(T)/CBS and a few less correlated wavefunction-based approximations.

Cpd. (sym.)	DF-HF/a5Z	DF-MP2/aQ5Z	CCSD(T)/aDZ	CCSD(T)/CBS
1 (C2h)	-12.36	-21.09	-22.26	-20.40
1' (C2)	-9.55	-15.65	-17.23	-15.77
1' (Ci)	-9.47	-15.59	-17.13	-15.69
2 (C2)	-8.23	-17.14	-18.57	-16.49
2 (Ci)	-8.64	-16.95	-18.29	-16.39
2' (C2)	-9.98	-18.66	-19.51	-17.76
2' (Ci)	-9.82	-18.62	-19.34	-17.66
BA (C2h)	-12.53	-17.43	-18.88	-17.79
IM (C2)	-9.53	-16.30	-17.76	-16.31
IM (Ci)	-9.52	-16.31	-17.78	-16.32

II.4. The performance of DFT methods vs CCSD(T)/CBS in the gas-phase

Deviations for calculated H-bond electronic energies vs the CCSD(T)/CBS composite level of theory for **1**, **1'**, **2**, **2'**, **BA** and **IM** as well as their mean absolute errors (MAE, the right column) are presented in **Table II-3**. For reference, the CCSD(T)/CBS energies of each species is shown on top of their corresponding columns. The deviations (signed errors) for each species at various DFT methods and basis sets (DFT/base), denoted as D_x^{DFT} , were calculated as

$$D_x^{\text{DFT/base}} = E_x^{\text{DFT/base}} - E_x^{\text{CCSD(T)/CBS}}$$

in which, $E_x^{\text{DFT/base}}$ and $E_x^{\text{CCSD(T)/CBS}}$ are total electronic energies of dimerization for species X (X = **1**, **1'**, **2**, **2'**, **BA**, or **IM**) at DFT/base and CCSD(T)/CBS levels of theory, respectively. The values of $D_x^{\text{DFT/base}}$ are presented in columns 2-7 of **Table II-3**. The MAE values for each level of theory over the 6 complexes were then calculated as

$$MAE = \frac{|E_1^{\text{DFT/base}}| + |E_{1'}^{\text{DFT/base}}| + |E_2^{\text{DFT/base}}| + |E_{2'}^{\text{DFT/base}}| + |E_{\text{BA}}^{\text{DFT/base}}| + |E_{\text{IM}}^{\text{DFT/base}}|}{6}$$

Consistent with recommendations from the Thanthiriwatte *et al.* study of formic acid, formamide and formamidine homo- and hetero-dimers,¹²² methods such as ω B97X-D,¹²⁵ PBE,^{126,127} or M06-2X¹²⁸ produce H-bond energies with MAE values of less than 0.5 kcal/mol is some combinations with a large basis sets such as 6-311++G(3df,3pd), aDZ or aTZ. To shorten denotation of large Gaussian basis sets, their '6-311++' term is replaced by 'L'. This simplifies these basis sets as: 6-311++G(3df,3pd) = LG(3df,3pd), 6-311++G(2df,2p) = LG(2df,2p), and 6-311++G(d,p) = LG(d,p).

It is notable that PBE0 (also known as PBE1PBE),¹²⁹ M06-2X, and PBE (not to be confused with PBE0) can achieve high accuracy when used in combination with the moderate 6-31+G(d,p) basis set saving hugely in computational resources. Using this relatively small basis set, methods such as MN15¹³⁰ and ω B97X-D also perform acceptably. However, LC- ω HPBE¹³¹ and TPSS¹³² produce errors of 1-2 kcal/mol, and APFD¹³³ and B3LYP methods are not recommended to be used with this basis set. The use of HF and S-VWN¹³⁴⁻¹³⁶ is to be avoided with any basis set. While the calculated energies of the former method are under-bound by up to 6 kcal/mol, the latter method predicts H-bond energies that are over-bound by up to 19 kcal/mol. In addition to these two methods, the use of Dunning's double zeta basis set without augmentation, DZ, should be avoided. This basis set is the smallest one in this study. On the other hand, the augmented version of this basis set, aDZ, appears to construct the best performing levels of theory when used in conjunction with appropriate DFT methods. It is notable that as opposed to wavefunction-based methods, for density functional methods, the increase of basis set size does not necessarily lead to more

accurate H-bond energies. For example, the PBE method used with the aDZ basis set leads to more accurate results than the significantly larger basis set aTZ. On a second-row element such as carbon, there are 14, 15, 19, 22, 23, 30, 34, 39, and 46 basis functions for DZ, 6-31G(d), 6-31+G(d,p), LG(d,p), aDZ, TZ, LG(2df,2p), LG(3df,3pd), and aTZ basis sets, respectively.

Table II-3. Mean absolute errors (MAE) of calculated electronic energies of H-bonds using various levels of theory vs CCSD(T)/CBS. The energies are not counterpoise corrected.

	1	1'	2	2'	BA	IM	MAE
CCSD(T)/CBS values:	-20.40	-15.77	-16.49	-17.76	-17.79	-16.32	kcal/mol
ω B97X-D/LG(3df,3pd)	0.40	-0.04	-0.64	-0.85	-0.73	0.03	0.45
ω B97X-D/aTZ	0.47	0.09	-0.61	-0.88	-0.55	0.12	0.46
PBE/aDZ	0.59	0.49	-0.16	-0.79	-0.73	0.02	0.46
M06-2X/aDZ	0.34	0.51	0.80	0.41	-0.27	0.46	0.47
PBE0/aDZ	0.56	0.67	0.68	-0.01	-0.50	0.42	0.47
M11/LG(d,p)	0.53	0.22	0.54	0.21	-0.93	0.43	0.48
MN15/aDZ	0.50	0.14	0.80	0.23	-0.96	0.26	0.48
ω B97X-D/LG(2df,2p)	0.33	-0.14	-0.84	-1.15	-0.63	-0.07	0.53
PBE0/6-31+G(d,p)	0.65	0.58	0.49	-0.02	-0.90	0.64	0.55
M06-2X/6-31+G(d,p)	0.47	0.43	0.65	0.43	-0.70	0.76	0.57
PBE/6-31+G(d,p)	0.88	0.62	-0.08	-0.46	-1.04	0.45	0.59
PBE0/TZ	0.80	0.68	0.40	-0.28	-0.93	0.45	0.59
M11/6-31+G(d,p)	-0.16	0.05	0.97	0.60	-1.60	0.26	0.61
M06-2X/TZ	0.58	0.51	0.81	0.46	-0.88	0.54	0.63
M05-2X/aDZ	0.53	0.55	1.30	0.88	-0.32	0.38	0.66
M06-2X/LG(d,p)	1.00	0.73	0.65	0.42	-0.21	0.98	0.67
M05-2X/LG(d,p)	0.91	0.53	0.90	0.61	-0.44	0.67	0.68
PBE0/LG(d,p)	1.34	1.00	0.46	-0.09	-0.22	0.98	0.68
M05-2X/TZ	0.53	0.37	1.11	0.77	-1.08	0.26	0.68
M11/aDZ	-0.68	-0.41	0.26	-0.65	-1.66	-0.53	0.70
PBE/TZ	0.71	0.36	-0.56	-1.07	-1.41	-0.11	0.70
M11/TZ	0.73	0.48	1.11	0.63	-0.85	0.50	0.71
PBE/LG(d,p)	1.58	1.02	-0.13	-0.54	-0.21	0.81	0.72
M05-2X/6-31+G(d,p)	0.64	0.46	1.17	0.97	-0.55	0.69	0.75
PBE/LG(2df,2p)	1.72	1.27	0.52	0.11	-0.17	0.98	0.80
ω B97X-D/LG(d,p)	0.14	-0.43	-1.53	-1.81	-0.79	-0.28	0.83
PBE/aTZ	1.65	1.33	0.61	0.22	-0.22	0.97	0.83
M11/LG(2df,2p)	0.91	0.61	1.30	0.88	-0.68	0.73	0.85
PBE/LG(3df,3pd)	1.68	1.27	0.64	0.34	-0.39	0.95	0.88
ω B97X-D/6-31+G(d,p)	-0.35	-0.63	-1.29	-1.52	-1.23	-0.41	0.90
MN15/TZ	1.38	0.81	1.15	0.82	-0.67	0.88	0.95
MN15/6-31+G(d,p)	1.17	0.69	1.22	1.03	-0.69	1.00	0.97
PBE0/LG(2df,2p)	1.57	1.32	1.18	0.60	-0.07	1.22	0.99

Table II-3 (cont'd)

M06-2X/aTZ	1.04	0.99	1.51	1.25	-0.28	1.09	1.03
MN15/LG(d,p)	1.69	0.95	1.15	0.88	-0.31	1.20	1.03
M05-2X/aTZ	0.89	0.82	1.75	1.44	-0.56	0.78	1.04
LC- ω HPBE/6-31G(d)	-0.53	-0.74	-0.65	-1.26	-2.23	-0.93	1.06
M11/LG(3df,3pd)	1.16	0.85	1.70	1.39	-0.42	0.94	1.08
PBE0/aTZ	1.57	1.45	1.32	0.80	-0.07	1.31	1.09
ω B97X-D/aDZ	-0.67	-0.78	-1.37	-1.83	-1.05	-0.85	1.09
M06-2X/LG(2df,2p)	1.38	1.16	1.48	1.16	0.10	1.33	1.10
PBE0/LG(3df,3pd)	1.54	1.36	1.32	0.87	-0.29	1.24	1.10
ω B97X-D/TZ	-0.32	-0.71	-1.59	-1.98	-1.37	-0.74	1.12
M05-2X/LG(2df,2p)	1.36	1.03	1.85	1.52	0.00	1.05	1.14
TPSS/TZ	2.04	1.72	1.08	0.68	-0.15	1.25	1.15
B3LYP/6-31G(d)	-0.58	-0.85	-0.80	-1.24	-2.24	-1.27	1.16
MN15/aTZ	1.59	1.11	1.63	1.25	-0.15	1.27	1.17
M06-2X/LG(3df,3pd)	1.31	1.21	1.62	1.46	-0.27	1.34	1.20
M11/aTZ	1.29	1.12	1.99	1.54	-0.19	1.16	1.21
TPSS/aDZ	1.77	1.77	1.38	0.93	0.33	1.24	1.24
M05-2X/LG(3df,3pd)	1.30	1.10	1.98	1.82	-0.42	1.10	1.29
TPSS/6-31+G(d,p)	1.91	1.73	1.39	1.12	-0.10	1.53	1.30
MN15/LG(2df,2p)	1.96	1.31	1.88	1.60	0.02	1.51	1.38
TPSS/6-31G(d)	-0.60	-0.80	-1.24	-1.71	-2.67	-1.51	1.42
MN15/LG(3df,3pd)	1.97	1.31	1.97	1.75	-0.17	1.51	1.45
LC- ω HPBE/TZ	2.19	1.90	1.88	1.36	0.37	1.82	1.59
B3LYP/TZ	2.29	1.81	1.79	1.51	0.59	1.61	1.60
LC- ω HPBE/aDZ	1.86	1.82	2.06	1.50	0.72	1.72	1.62
TPSS/LG(d,p)	2.61	2.16	1.35	1.07	0.71	1.89	1.63
LC- ω HPBE/6-31+G(d,p)	2.03	1.85	2.01	1.68	0.53	2.06	1.70
MN15/6-31G(d)	-1.14	-1.54	-1.18	-1.65	-3.24	-1.65	1.73
B3LYP/DZ	-1.10	-1.12	-1.57	-2.01	-3.22	-1.42	1.74
B3LYP/aDZ	2.10	1.89	2.07	1.66	1.12	1.66	1.75
M05-2X/6-31G(d)	-1.54	-1.62	-1.19	-1.65	-2.80	-1.88	1.78
LC- ω HPBE/DZ	-1.32	-1.16	-1.48	-2.20	-3.55	-1.30	1.84
M06-2X/6-31G(d)	-1.62	-1.54	-1.57	-2.02	-2.87	-1.69	1.88
B3LYP/6-31+G(d,p)	2.30	1.98	2.22	2.10	0.80	2.03	1.90
LC- ω HPBE/LG(d,p)	2.68	2.18	1.85	1.45	1.04	2.29	1.92
TPSS/LG(2df,2p)	2.82	2.42	1.95	1.65	0.80	2.08	1.95
TPSS/LG(3df,3pd)	2.76	2.43	2.05	1.87	0.55	2.06	1.95
APFD/aTZ	-1.32	-1.37	-2.02	-2.58	-3.12	-1.40	1.97
APFD/LG(3df,3pd)	-1.35	-1.46	-2.04	-2.52	-3.35	-1.47	2.03
APFD/LG(2df,2p)	-1.31	-1.48	-2.16	-2.78	-3.13	-1.47	2.05
TPSS/aTZ	2.91	2.62	2.14	1.88	0.87	2.21	2.11
PBE0/6-31G(d)	-1.51	-1.60	-1.85	-2.59	-3.17	-2.00	2.12
M05-2X/DZ	-1.80	-1.57	-1.62	-2.00	-3.94	-1.82	2.13
M06-2X/DZ	-1.87	-1.48	-1.96	-2.36	-3.61	-1.60	2.15
TPSS/DZ	-1.30	-1.17	-2.13	-2.67	-3.88	-1.81	2.16
B3LYP/LG(d,p)	2.94	2.32	2.17	2.02	1.47	2.32	2.21
LC- ω HPBE/LG(2df,2p)	2.95	2.56	2.69	2.27	1.27	2.58	2.39

Table II-3 (cont'd)

MN15/DZ	-1.85	-1.87	-1.86	-2.37	-4.46	-1.91	2.39
APFD/LG(d,p)	-1.47	-1.74	-2.87	-3.46	-3.21	-1.65	2.40
LC- ω HPBE/LG(3df,3pd)	2.98	2.65	2.88	2.59	1.10	2.65	2.48
M11/6-31G(d)	-2.43	-2.35	-1.57	-2.19	-3.89	-2.58	2.50
LC- ω HPBE/aTZ	3.03	2.78	2.88	2.50	1.36	2.77	2.55
B3LYP/LG(2df,2p)	3.19	2.62	2.77	2.61	1.60	2.56	2.56
B3LYP/LG(3df,3pd)	3.19	2.65	2.88	2.82	1.42	2.56	2.59
B3LYP/aTZ	3.16	2.69	2.86	2.73	1.60	2.58	2.60
APFD/6-31+G(d,p)	-2.16	-2.15	-2.77	-3.37	-3.93	-1.97	2.73
APFD/aDZ	-2.36	-2.16	-2.70	-3.45	-3.54	-2.31	2.75
PBE0/DZ	-2.21	-1.96	-2.66	-3.49	-4.24	-2.31	2.81
HF/DZ	3.28	2.63	3.12	3.07	1.96	2.84	2.81
APFD/TZ	-2.10	-2.17	-3.00	-3.69	-3.99	-2.27	2.87
HF/6-31G(d)	3.31	2.60	3.43	3.26	1.88	2.83	2.89
PBE/6-31G(d)	-2.00	-2.27	-3.12	-3.71	-4.02	-3.01	3.02
ω B97X-D/6-31G(d)	-2.33	-2.64	-3.52	-3.97	-3.37	-2.83	3.11
M11/DZ	-3.11	-2.54	-2.43	-3.25	-5.31	-2.77	3.24
ω B97X-D/DZ	-3.13	-3.11	-4.50	-5.00	-4.48	-3.26	3.91
PBE/DZ	-2.83	-2.77	-4.16	-4.84	-5.55	-3.40	3.92
HF/6-31+G(d,p)	4.60	3.72	4.71	4.83	3.43	4.21	4.25
HF/LG(d,p)	5.29	4.27	5.01	5.24	4.00	4.64	4.74
HF/aDZ	5.15	4.45	5.53	5.84	4.21	4.70	4.98
HF/TZ	5.51	4.57	5.33	5.45	4.23	4.82	4.98
APFD/6-31G(d)	-4.27	-4.34	-5.18	-5.89	-6.05	-4.61	5.06
HF/LG(2df,2p)	5.81	4.74	5.69	5.90	4.45	5.06	5.27
HF/LG(3df,3pd)	5.89	4.90	5.89	6.24	4.37	5.24	5.42
HF/aTZ	5.98	5.02	5.91	6.20	4.62	5.33	5.51
APFD/DZ	-5.14	-4.81	-6.14	-6.96	-7.36	-5.04	5.91
S-VWN/aTZ	-9.10	-8.54	-9.18	-10.64	-13.73	-8.88	10.01
S-VWN/LG(3df,3pd)	-9.19	-8.66	-9.23	-10.63	-13.93	-8.99	10.10
S-VWN/LG(2df,2p)	-9.15	-8.65	-9.43	-10.97	-13.90	-8.93	10.17
S-VWN/LG(d,p)	-9.13	-8.75	-10.18	-11.73	-13.67	-8.93	10.40
S-VWN/6-31+G(d,p)	-10.04	-9.17	-9.94	-11.55	-14.99	-9.33	10.84
S-VWN/TZ	-10.03	-9.48	-10.33	-11.98	-14.86	-9.93	11.10
S-VWN/aDZ	-10.79	-9.92	-10.65	-12.66	-14.57	-10.43	11.50
S-VWN/6-31G(d)	-11.95	-11.47	-12.66	-14.27	-16.31	-12.25	13.15
S-VWN/DZ	-14.07	-12.86	-14.40	-16.38	-19.49	-13.58	15.13

To summarize this part, the combined choice of basis set and DFT method is critical for achieving high accuracy for gas phase H-bond energies. The results obtained here are consistent with those of the previous study by Sherrill's group,¹²² but inclusion of smaller basis sets in this work enables recommendations suitable for larger systems without

compromise in the accuracy. Specifically, the use of PBE0, M06-2X and PBE functionals is recommended in combination with the 6-31+G(d,p) basis set when quick calculations of small molecules are needed or the system of study is too large for use of more complete basis sets.

II.5. Comparison of theory and experiment

With CCSD(T)/CBS energies as reference for gas-phase “electronic energies” and experimental dimerization “free energies” as reference for solution phase, the performance of various combinations of DFT methods (and HF) and basis sets can be evaluated both in the gas- and in the solution phases. For the latter, implicit solvation models SMD, C-PCM, and IEF-PCM were used to simulate the four studied solvents. The geometries for both gas-phase and solution phase structures were fully optimized at the specified level of theory in the corresponding medium. All the DFT calculations were performed using Gaussian 16. Due to the large number of calculations (about 30,000 geometry optimizations) the convergence of the geometries, maximum force, RMS force, maximum displacement, and RMS displacement were only checked before frequencies. The default Gaussian16 values were used for these four criteria of geometry convergence. Since the experimental free energies used as reference were at room temperature (298.15 K), no special correction was made to the calculated thermal corrections to free energies and standard solvent parameters as they were implemented in Gaussian16 were used. The harmonic frequencies for each fully optimized structure were calculated in the solution phase and used for calculation of thermal correction to free energy. The following formula was used to calculate the dimerization free energies:

$$\Delta_{\text{dim}}G = G_{\text{dimer}} - 2 \times G_{\text{monomer}} - 1.89 \text{ kcal/mol}$$

in which G_{dimer} and $G_{monomer}$ are dimer and monomer free energies which were directly extracted from the thermochemistry section of the Gaussian output files by searching for the “Sum of electronic and thermal Free Energies” string. The 1.89 kcal/mol is the correction for free energy of dimerization that originates from the contraction from gas-phase standard conditions (1 atm) to the solution-phase standard conditions (1 M) at 298.15 K. This value was obtained using the following formula

$$\Delta G(1 \text{ atm} \rightarrow 1 \text{ M}) = nRT \ln\left(\frac{V_{1\text{atm}}^{\text{gas}}}{V_{1\text{M}}^{\text{sol}}}\right)$$

in which $n = 1 \text{ mol}$, $R = 1.987 \times 10^{-3} \text{ kcal/molK}$, $T = 298.15 \text{ K}$, $V_{1\text{M}}^{\text{sol}} = 1 \text{ L}$, and $V_{1\text{atm}}^{\text{gas}}$ for an idea gas at these conditions was calculated using $PV = nRT$ to be 24.465 L. Substituting all these values in the above equation, $\Delta G(1\text{atm} \rightarrow 1\text{M}) = +1.89 \text{ kcal/mol}$ was obtained, which is also reported in the original SMD parametrization paper.¹¹⁸ The harmonic frequencies for each fully optimized structure were calculated in the solution phase and used for calculation of thermal correction to free energy.

Monomers of **1** and **BA** were optimized in C_s symmetry. However, in some cases, the symmetric monomer of **1** led to an imaginary frequency. In those cases, the symmetry was broken. Dimers of **1** and **BA** were optimized in C_{2h} symmetry. For dimers of **1**, again, this symmetry resulted in imaginary frequencies in some cases, so C_i and C_2 symmetry geometries were also optimized for the dimers of **1**. For the rest of the molecules (**1'**, **2**, **2'** and **IM**), the monomers were optimized in C_1 symmetry and the dimers were optimized in both C_2 and C_i symmetries. For extraction of data, a home-made python code was used. If the C_s symmetry monomer of **1** had an imaginary frequency, then the C_1 one was used. Similarly, for the dimer of **1** if the geometry with C_{2h} symmetry had any imaginary frequencies then the C_2 and C_i were compared and the one with lowest energy and no imaginary frequency was used. For **BA**, the C_s monomers and C_{2h} dimers

did not have any imaginary frequency at any level of theory. For the other cases with C₂ and C_i dimer, the symmetry with the lowest energy and no imaginary frequency was used.

The criteria for the best-performing levels of theory, though somewhat arbitrary, was the sum of gas-phase electronic energy MAEs vs CCSD(T)/CBS reference, and the solution phase free energies vs experimental values. **Figure II-1**, which is extended to a few pages, shows the plots comparing the performance of the studied methods. The Y-axis shows various levels of theory and the X-axis is the sum of gas-phase and solution phase MAEs, by which the Y-axis is sorted. Looking at the end of the plot, it is obvious that choosing HF and S-VWN methods in combination with any basis set leads to large errors in both the gas- and solution-phase. It is also clear that while using the DZ basis set (the smallest one in this study) in combination with any method leads to substantial errors, its augmented version, the aDZ basis set, is among the best performing ones when used with DFT methods such as PBE0, MN15, M06-2X, M05-2X or PBE. The colors of the bars show the solvation model used as guided in the legend, and the MAE for the corresponding levels of theory both in the gas-phase and solution phase are labeled on each bar. Nominally, PBE0/aDZ/C-PCM is the best performing method among all. Its MAE values for gas- and solution phases are 0.47 and 0.60 kcal/mol respectively. Changing the solvation model from C-PCM to IEF-PCM leads to a slight increase of 0.06 kcal/mol in its MAE. However, using the SMD solvation model with this level of theory leads to solution phase MAE of 1.47 kcal/mol. This comparison points out to the importance of the choice of implicit solvation model in addition to the DFT method and basis set.

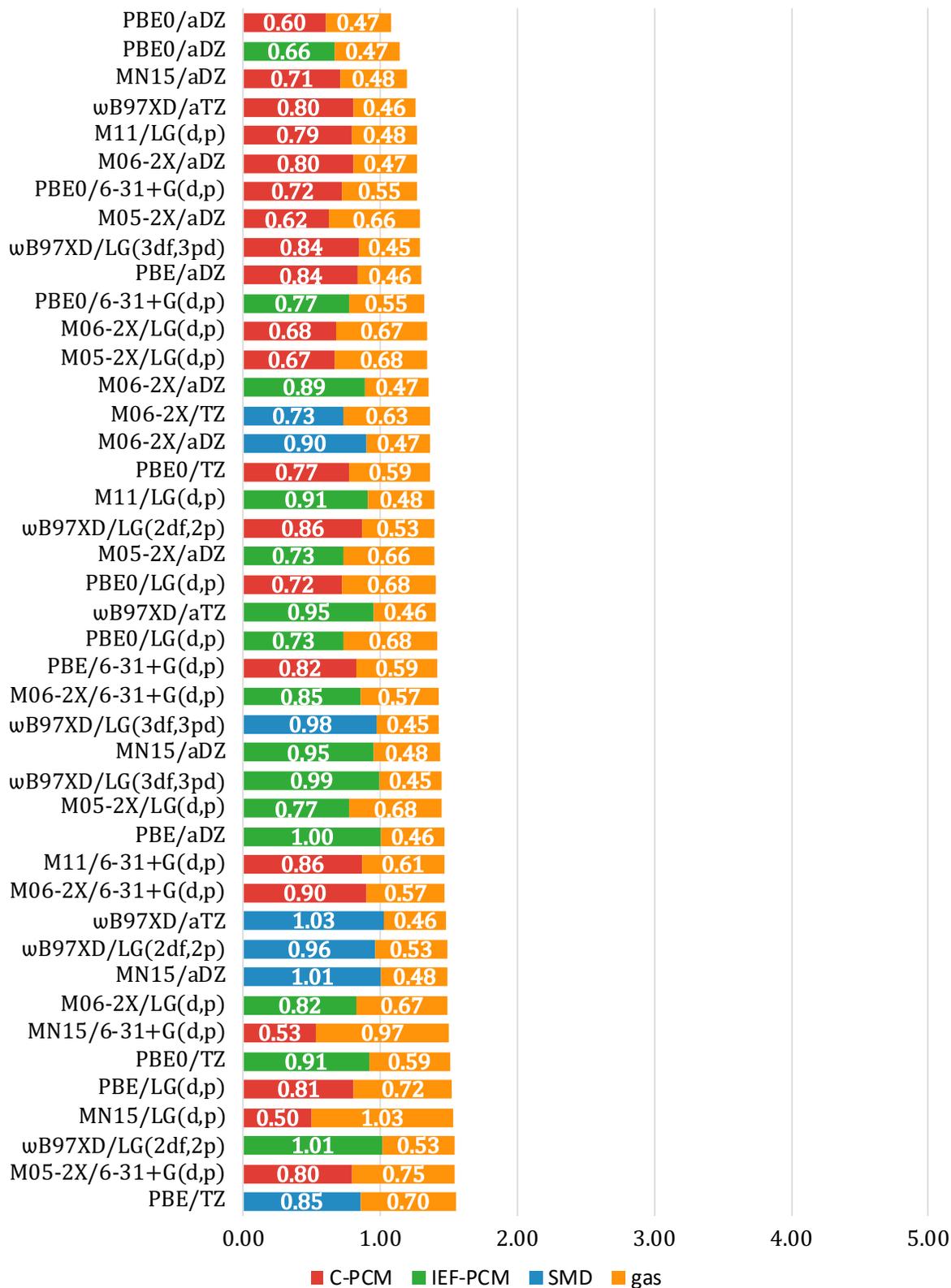


Figure II-1. Sorted by sum of mean absolute errors of various levels of theory vs gas-phase CCSD(T)/CBS electronic energies and experimental H-bond free energies.

Figure II-1 (cont'd)

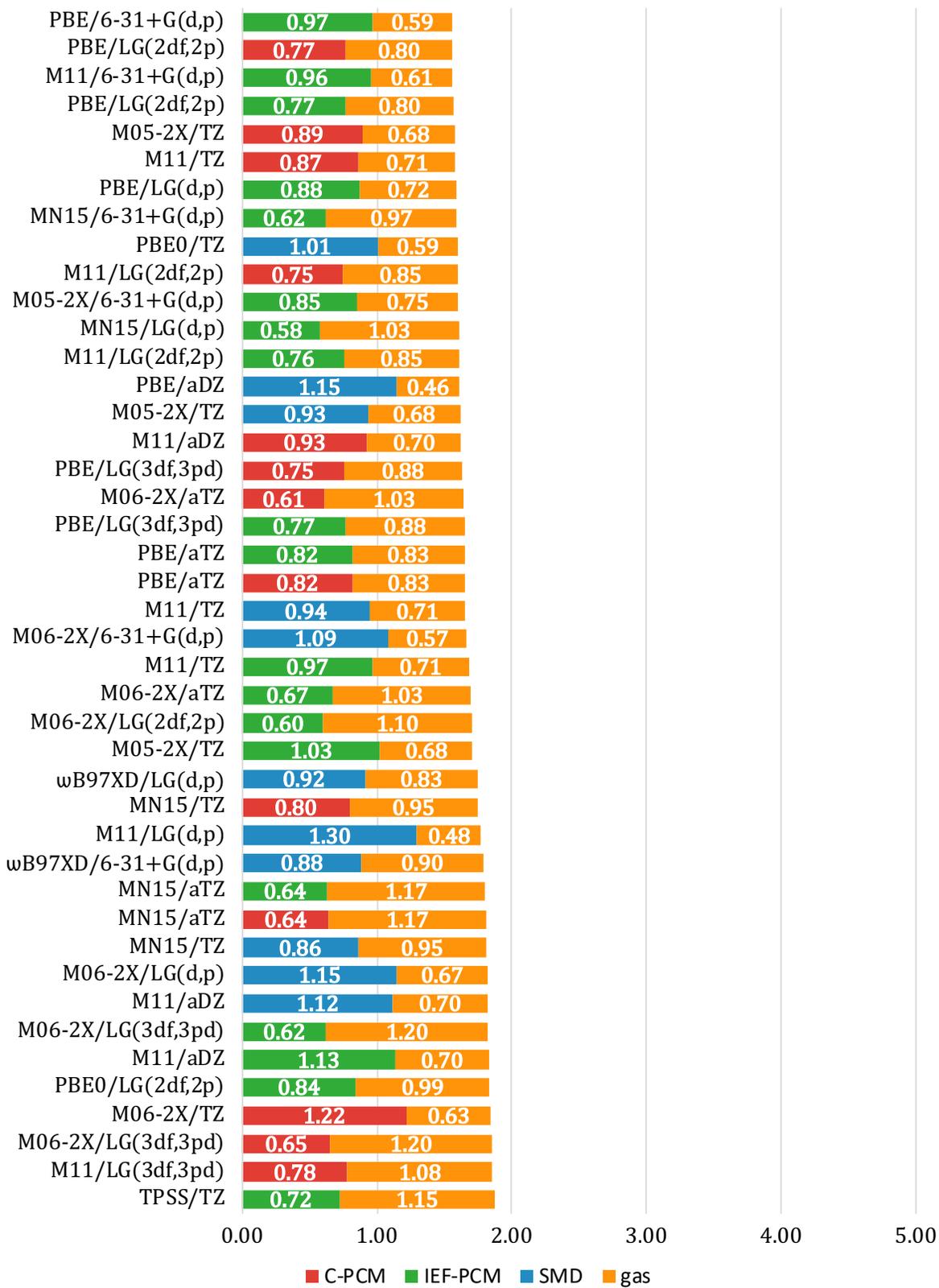


Figure II-1 (cont'd)

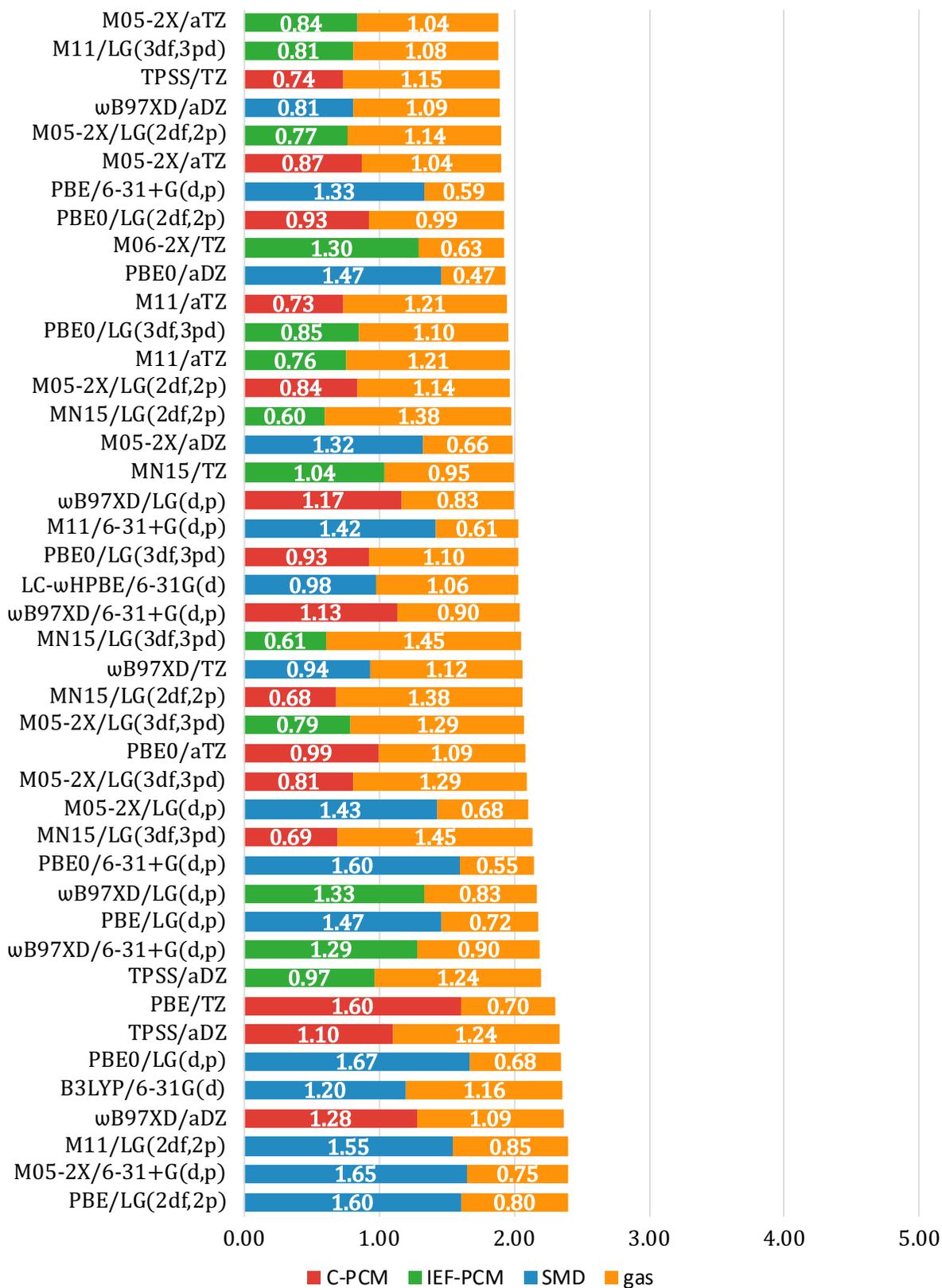


Figure II-1 (cont'd)

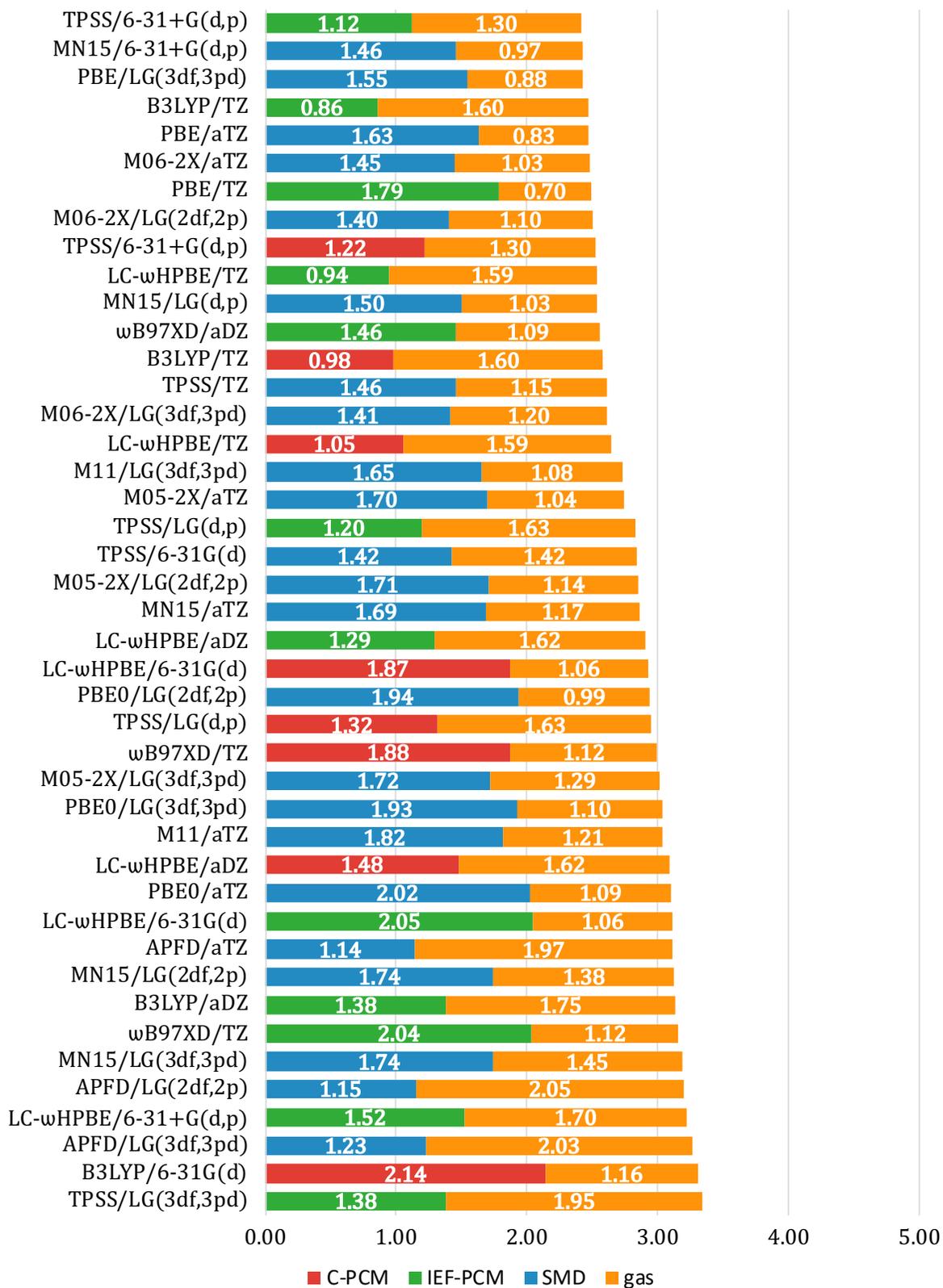


Figure II-1 (cont'd)

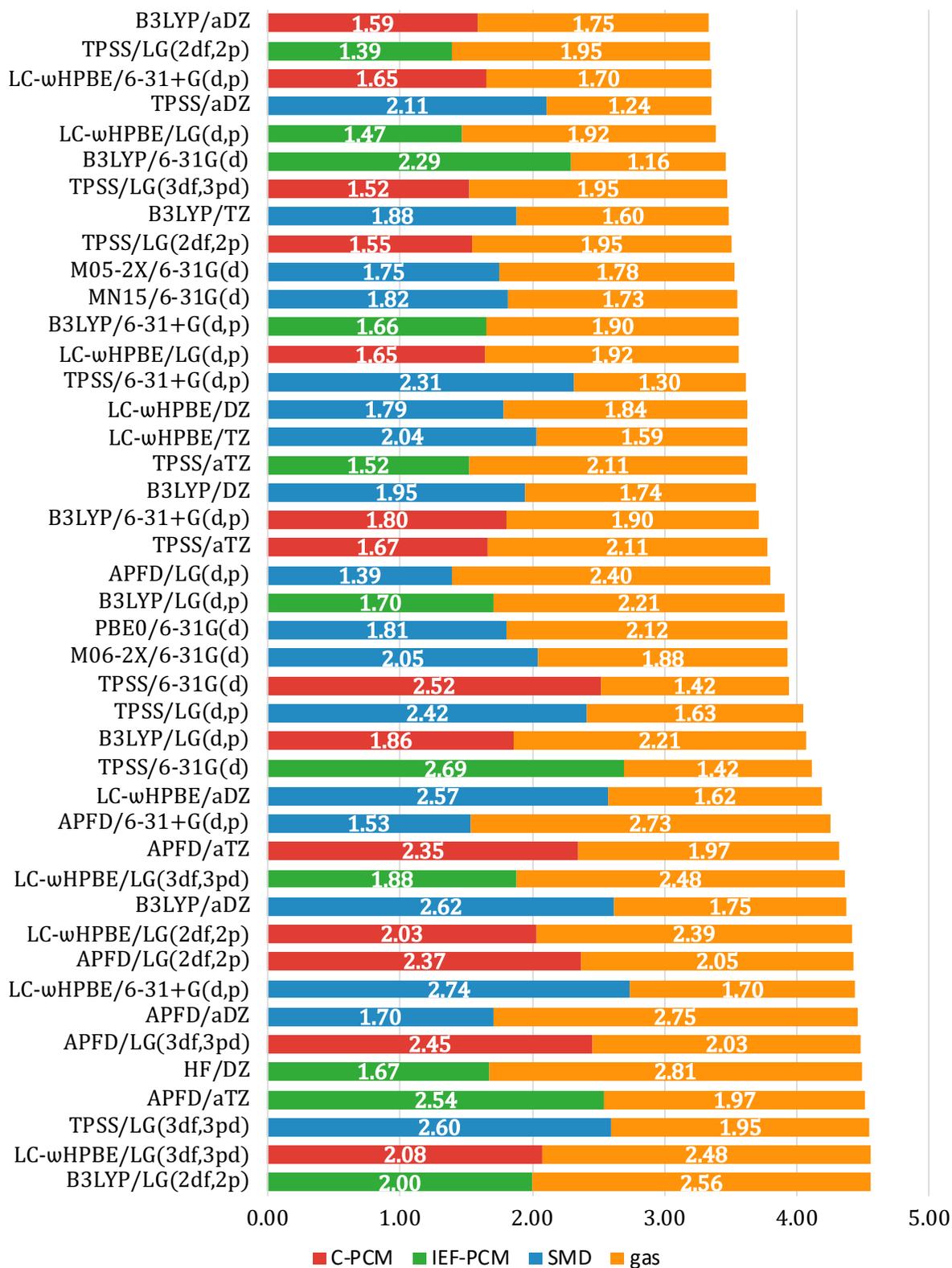


Figure II-1 (cont'd)

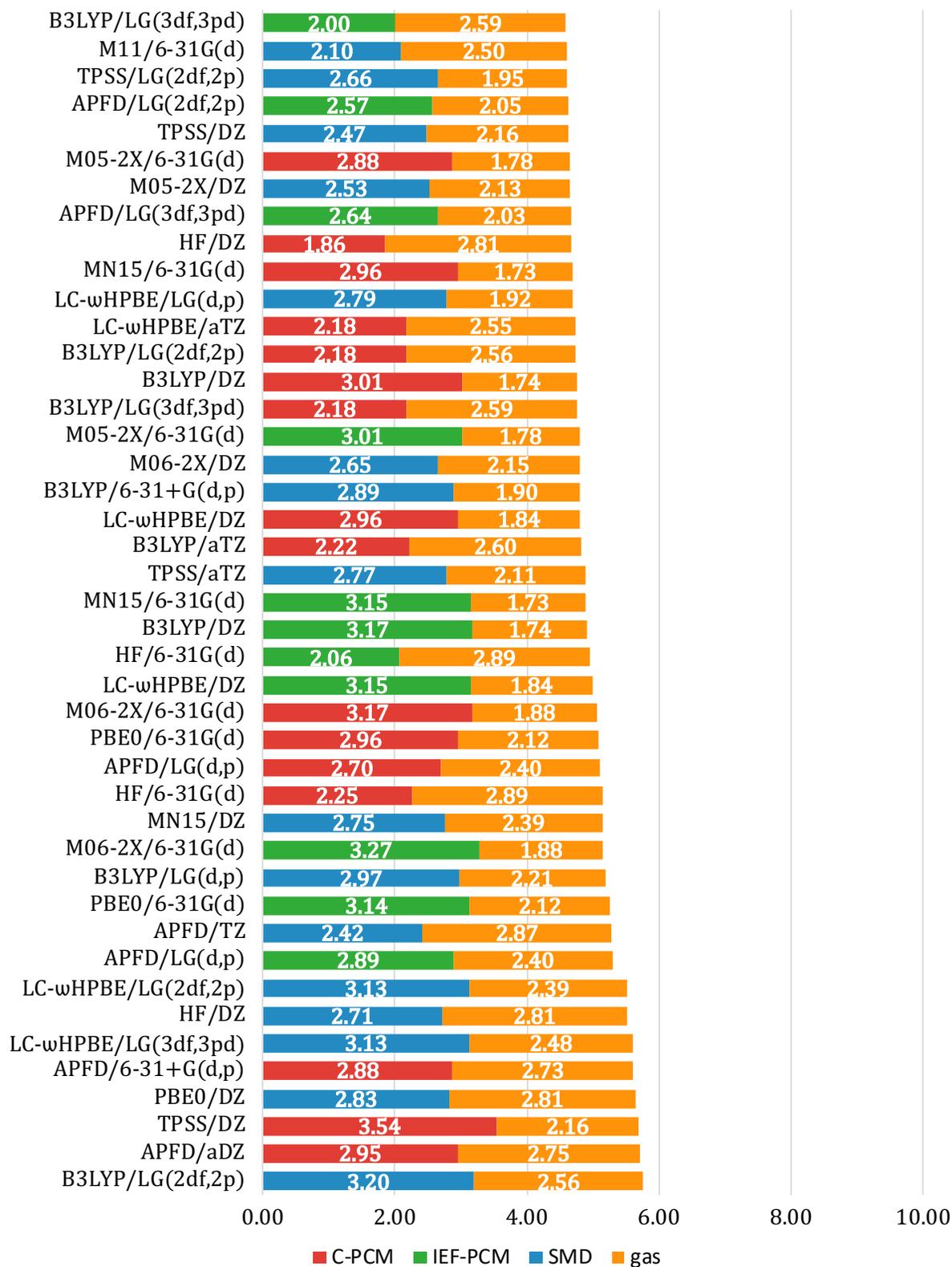


Figure II-1 (cont'd)

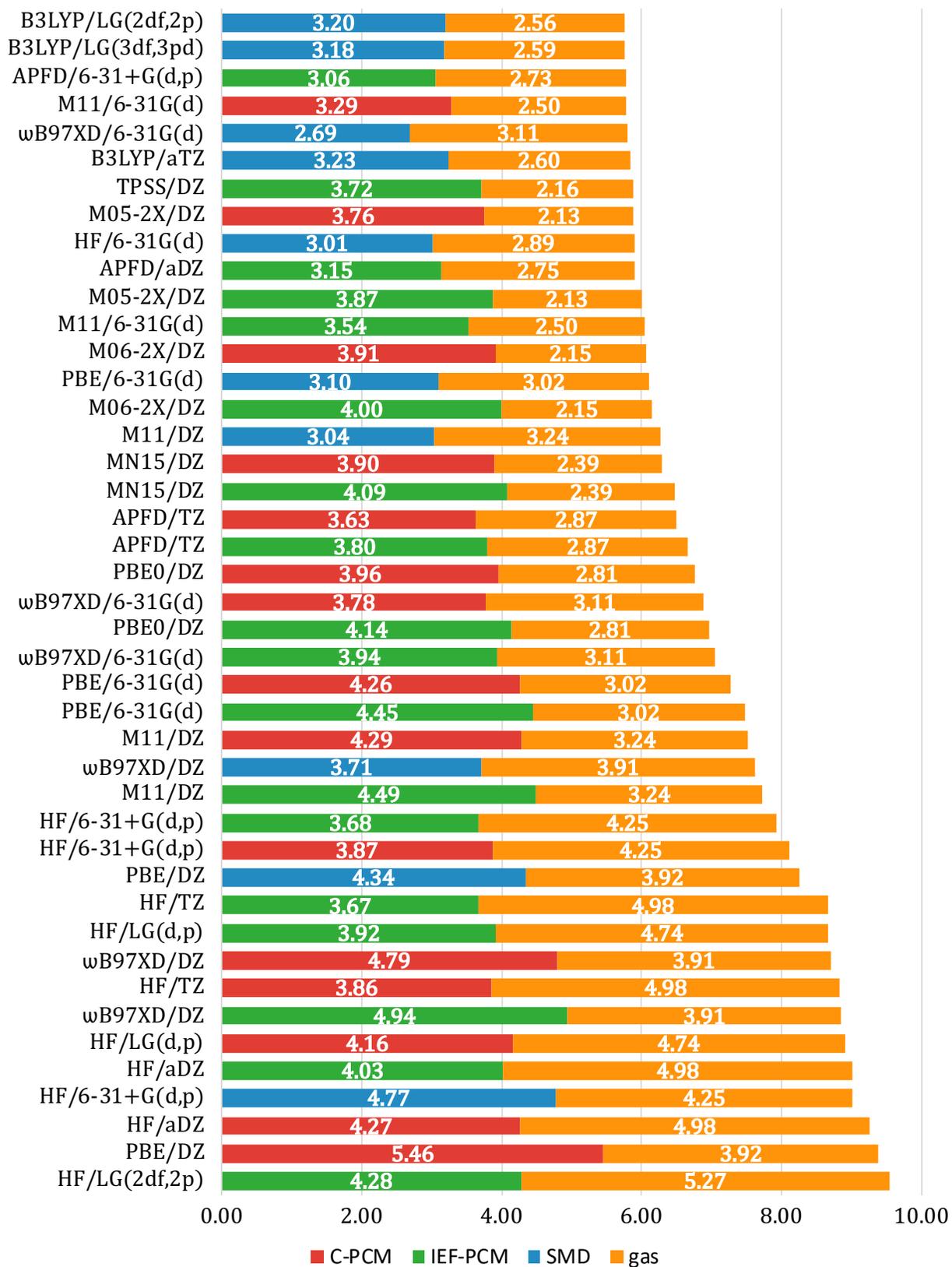
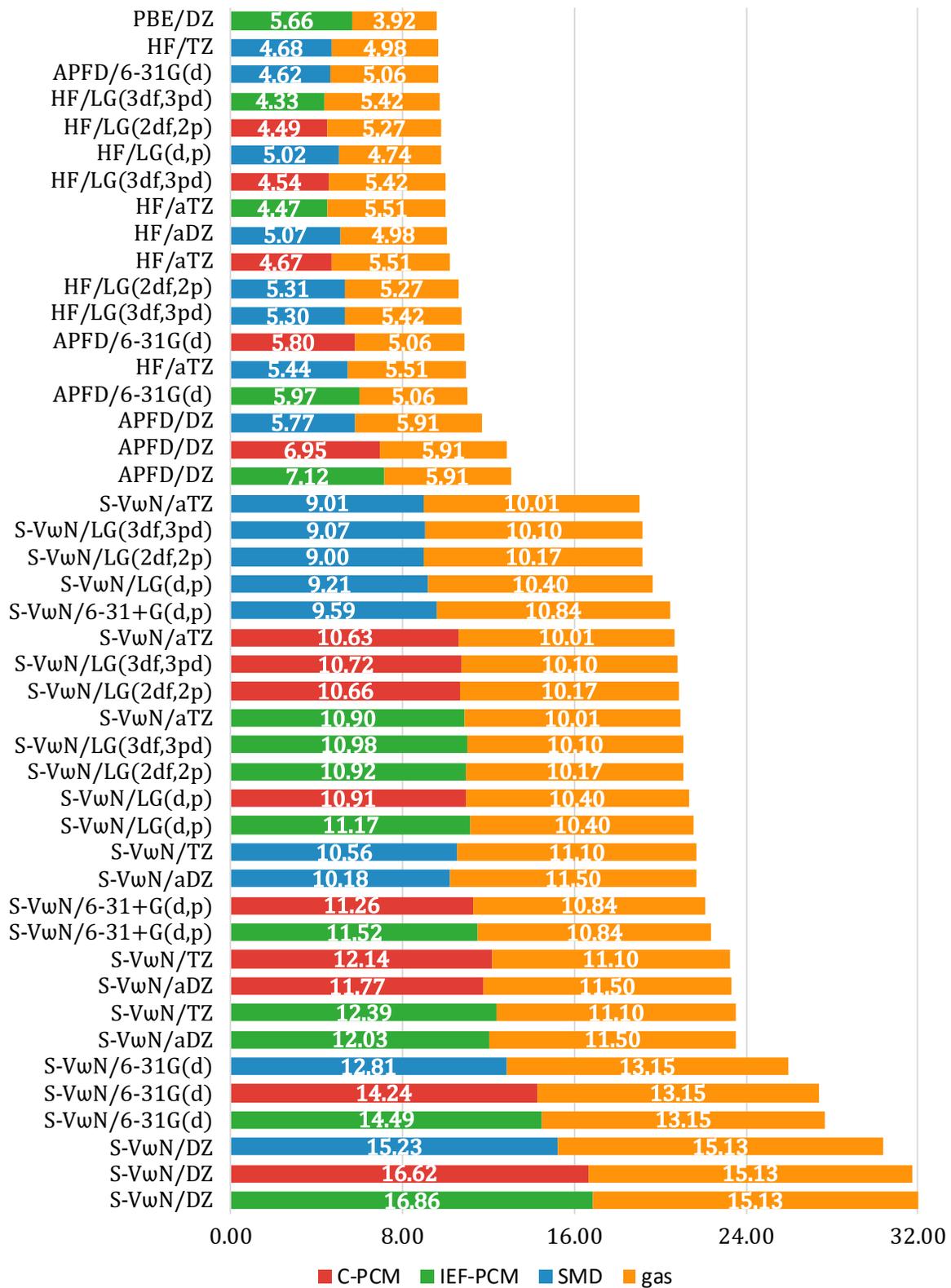


Figure II-1 (cont'd)



Interestingly, as opposed to wave function methods, using substantially larger basis sets than aDZ such as aTZ, LG(3df,3pd), or LG(2df,2p) does not lead to systematic improvement of accuracy. For example, replacing the aDZ basis set with the substantially larger aTZ basis set for the best-performing level of theory (PBE0/aDZ/C-PCM), in fact, increases the gas and solution phase MAE values to 1.09 and 0.99 kcal/mol, respectively. PBE0/6-31+G(d,p)/C-PCM seems to be a good balance between accuracy and computational resource demand. Its solution and gas-phase MAE values are only about 0.1 kcal/mol larger than those of the best-performing level of theory, which uses the aDZ basis set with all the same other parameters.

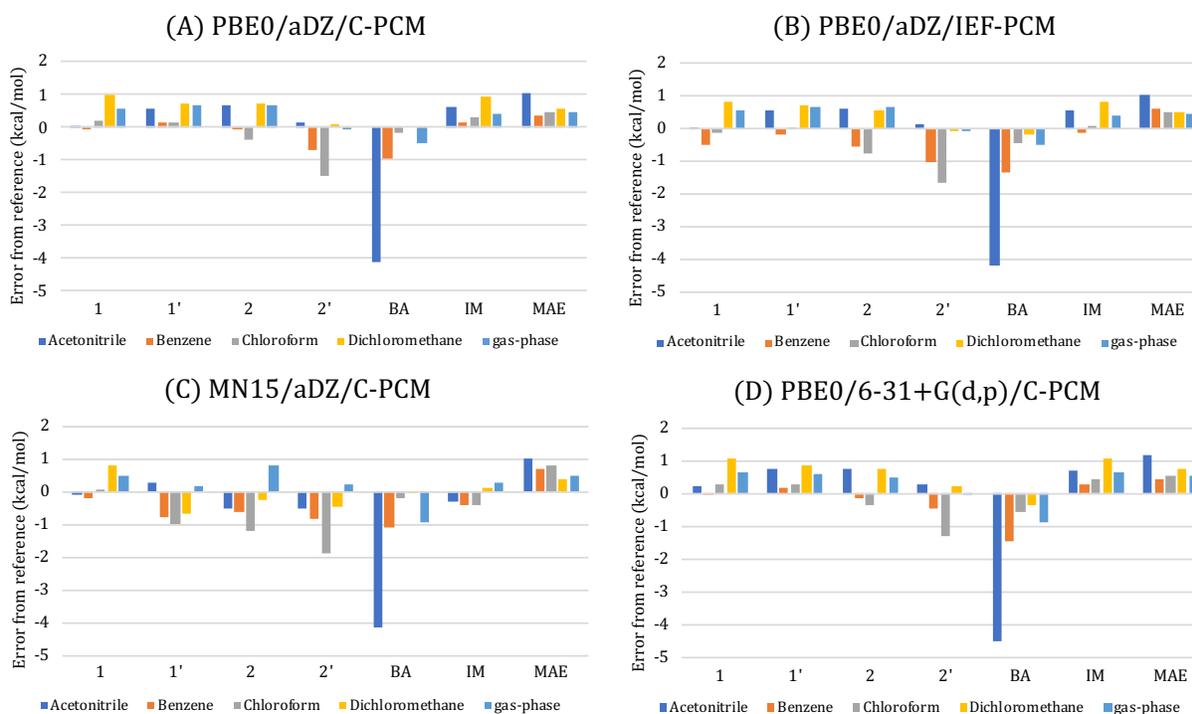


Figure II-2. (A), (B), and (C) Individual signed errors for the three best-performing and (D) the most economical level of theory.

The individual signed errors for the three best-performing levels of theory as well as PBE0/6-31+G(d,p)/C-PCM are shown in **Figure II-2**. While most errors are small, those of

BA in acetonitrile appear to be all negative and about 4 kcal/mol. This is presumably due to slight strong polarization of the carboxylate group in acetonitrile that can be challenging to model via implicit solvation models. Finding the best-performing method based on MAE values calculated by eliminating the **BA** in acetonitrile data did not change the conclusion from **Figure II-1** significantly. In this condition, the order changed very slightly and MN15/aDZ/C-PCM appeared as the best-performing level of theory.

II.6. Conclusions

Gas-phase CCSD(T)/CBS electronic energies and experimental solution phase free energies were used as reference to assess the performance of DFT method in the gas-phase as well as in the solution phase when used in combination with implicit solvation models. The benchmark study shows that in addition to the solvation model, the choice of both method and basis set is critical in choosing a level of theory for simulation of H-bond free energies in solution; using a long-range corrected DFT method or a large basis set does not guarantee a more accurate result. PBE0/aDZ/C-PCM appeared to be the best-performing level of theory with gas- and solution phase MAE values of 0.47 and 0.60 kcal/mol. The benchmark study also recommends the use of PBE0/6-31+G(d,p)/C-PCM levels of theory when calculation of larger systems is needed. The MAE values for both gas- and solution phases are only about 0.1 kcal/mol different from those of PBE0/aDZ/C-PCM.

II.7. Experimental H-bond measurements

The H-bond energies for species studied in this chapter were measured as described in the previous chapter. The NMR and thermodynamic values corrected for 298.15 K are shown in **Table II-4**. Data coverage is defined as the ratio of the measured shift range to the difference of the extrapolated monomer and dimer shifts multiplied by 100%.

Table II-4. Extracted parameters from NMR measurements at 298.15 K.

exp.	ΔH kcal/mol	ΔS cal/molK	ΔG kcal/mol	K_{eq} 1/M	δ_D ppm	δ_M ppm	$d\delta_D/dT$ ppb/K	$d\delta_M/dT$ ppb/K	data coverage (%)
1 in CD ₃ CN (NH)	-4.43	-11.57	-0.98	5.20	11.4131	7.7974	-5.9	-4.2	57
1 in CDCl ₃ (NH)	-5.84	-10.11	-2.82	117.52	11.6371	6.9555	-4.2	-3.1	82
1 in CD ₂ Cl ₂ (NH)	-6.59	-13.04	-2.70	95.76	11.5988	6.9484	-4.1	-2.3	86
1' in CD ₃ CN (NH)	-3.03	-12.34	0.65	0.33	7.1753	4.6176	-3.1	-3.2	32
1' in CDCl ₃ (NH) [exp. 1]	-3.44	-10.14	-0.42	2.04	7.1034	4.1408	-8.2	-1.9	34
1' in CDCl ₃ (NH) [exp. 2]	-3.29	-9.65	-0.42	2.02	7.1876	4.1396	-8.3	-1.9	54
1' in CDCl ₃ (NH) [exp. 3]	-3.39	-10.02	-0.40	1.98	7.1848	4.1367	-8.2	-1.9	50
1' in CD ₂ Cl ₂ (NH)	-4.01	-12.80	-0.19	1.39	7.2477	4.0950	-3.5	-1.6	49
2 in CD ₃ CN (NH)	-2.71	-10.59	0.45	0.47	7.1961	4.3054	-9.1	-2.4	19
2 in CDCl ₃ (NH) [exp. 1]	-1.94	-6.32	-0.06	1.10	6.9976	3.8711	-8.9	-1.8	38
2 in CDCl ₃ (NH) [exp. 2]	-2.04	-6.63	-0.06	1.10	6.9838	3.8690	-8.2	-1.7	44
2 in CD ₂ Cl ₂ (NH)	-3.33	-10.07	-0.33	1.74	7.2284	3.8512	-6.1	-1.5	50
2' in CD ₃ CN (NH)	-5.87	-16.88	-0.84	4.12	8.8239	4.5160	2.2	-3.5	45
2' in CDCl ₃ (NH)	-4.41	-11.88	-0.87	4.31	8.7345	4.0005	-6.2	-2.0	63
2' in CD ₂ Cl ₂ (NH)	-5.48	-13.20	-1.54	13.46	8.4964	4.0211	-6.4	-0.6	74
BA in CD ₃ CN (OH)	-3.81	-17.60	1.44	0.09	18.5424	9.3908	92.0	-5.9	13
BA in C ₆ D ₆ (OH)	-9.12	-17.91	-3.78	588.95	13.9893	5.8759	-7.4	-1.3	72
BA in CDCl ₃ (OH)	-8.16	-15.68	-3.48	355.86	13.3832	6.7941	-6.3	-5.8	81
BA in CD ₂ Cl ₂ (OH)	-8.18	-16.91	-3.14	200.28	13.3895	6.9767	-6.2	-4.8	88
Im in CD ₃ CN (NH)	-3.09	-12.41	0.62	0.35	8.4454	5.8274	-3.2	-3.5	39
Im in C ₆ D ₆ (NH)	-7.30	-17.22	-2.17	38.93	9.0788	4.1653	-10.2	1.6	85
Im in CDCl ₃ (NH)	-3.63	-10.08	-0.62	2.85	8.5734	5.3294	-5.4	-2.6	67
Im in CD ₂ Cl ₂ (NH)	-4.44	-13.48	-0.42	2.03	8.5044	5.2873	-4.2	-2.1	63

APPENDICES

APPENDIX A:

Chemical shifts used for obtaining associations

1 in acetonitrile (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	277.05 K	286.82 K	296.61 K	305.95 K
1.85E+02	---	9.8130	9.4222	9.4222
1.39E+02	9.8531	9.6453	9.2476	9.2476
9.25E+01	9.6093	9.3953	8.9981	8.9981
4.63E+01	9.1797	8.9690	8.6047	8.6047
2.31E+01	8.7710	8.5865	8.2827	8.2827
5.78E+00	8.2061	8.0958	7.9319	7.9319
2.31E+00	8.0239	7.9593	7.8333	7.8333
1.16E+00	7.9610	7.9002	7.7961	7.7961
5.78E-01	7.9268	7.8743	7.7866	7.7866

1 in benzene (NH) [exp. 1]				
	Chemical Shifts (ppm)			
conc. (mM)	280.15 K	291.50 K	301.25 K	310.50 K
3.19E+01	12.2766	12.0769	11.6730	11.6730
2.55E+01	12.2399	12.0268	11.5960	11.5960
1.92E+01	12.1876	11.9500	11.4880	11.4880
1.28E+01	12.0878	11.8229	11.2803	11.2803
6.39E+00	11.8710	11.5410	10.8730	10.8730
3.19E+00	11.5729	11.1580	10.3371	10.3371
6.39E-01	10.4560	9.7980	8.6950	8.6950
2.55E-01	9.5320	8.7800	7.6200	7.6200
1.28E-01	8.7320	7.9590	---	---
3.19E-02	---	6.6210	6.1145	6.1145
1.28E-02	---	6.1000	5.8660	5.8660

1 in benzene (NH) [exp. 2]				
	Chemical Shifts (ppm)			
conc. (mM)	280.15 K	291.40 K	301.04 K	311.00 K
2.70E+01	12.2458	12.0348	11.6085	11.6085
2.16E+01	12.2010	11.9794	11.5210	11.5210
1.62E+01	12.1420	11.8950	11.3970	11.3970
1.08E+01	12.0360	11.7560	11.1880	11.1880
5.40E+00	11.8080	11.4570	10.7480	10.7480
2.70E+00	11.4700	11.0451	10.1820	10.1820
5.40E-01	10.2870	9.6060	8.4360	8.4360
2.16E-01	9.3220	8.5520	7.3900	7.3900
1.08E-01	8.4800	7.7390	6.7810	6.7810
2.70E-02	overlap	6.5100	6.0550	6.0550
1.08E-02	6.3100	6.0570	---	---

1 in benzene (NH) [exp. 3]				
	Chemical Shifts (ppm)			
conc. (mM)	280.04 K	291.82 K	301.76 K	310.50 K
2.65E+01	12.2440	12.0315	11.6110	11.6110
2.12E+01	12.2013	11.9750	11.5240	11.5240
1.59E+01	12.1377	11.8950	11.3980	11.3980
1.06E+01	12.0316	11.7510	11.1890	11.1890
5.30E+00	11.8000	11.4470	10.7530	10.7530
2.65E+00	11.4730	11.0350	10.1790	10.1790
5.30E-01	10.2910	9.6061	8.4580	8.4580
2.12E-01	9.3040	8.5384	---	---
1.06E-01	8.5030	---	---	---
2.65E-02	overlap	6.5200	6.0500	6.0500
1.06E-02	6.3310	6.0700	5.8400	5.8400

1 in chloroform (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	277.05 K	286.70 K	296.53 K	304.01 K
1.86E+02	11.2470	11.1235	10.8766	10.8766
1.48E+02	11.1994	11.0680	10.8059	10.8059
1.11E+02	11.1281	10.9846	10.6992	10.6992
7.42E+01	11.0272	10.8681	10.5522	10.5522
3.71E+01	10.7778	10.5815	10.1962	10.1962
1.86E+01	10.4018	10.1582	9.6946	9.6946
4.64E+00	9.4754	9.1616	8.6237	8.6237
1.86E+00	8.7547	8.4427	7.9593	7.9593
9.28E-01	8.2227	7.9558	7.5736	7.5736
4.64E-01	7.7812	7.5686	---	---
2.32E-01	7.4721	---	---	---

1 in dichloromethane (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	276.10 K	283.91 K	291.74 K	299.58 K
1.50E+02	11.1476	11.0288	10.7711	10.7711
1.20E+02	11.0960	10.9692	10.6936	10.6936
9.02E+01	11.0176	10.8794	10.5782	10.5782
6.01E+01	10.8905	10.7331	10.3932	10.3932
3.01E+01	10.6028	10.4078	9.9940	9.9940
1.50E+01	10.2325	9.9983	9.5185	9.5185
3.76E+00	9.2297	8.9392	8.4095	8.4095
1.50E+00	8.5059	8.2326	7.7874	7.7874
7.52E-01	8.0098	7.7841	7.4479	7.4479
1.88E-01	7.3557	7.2456	7.1007	7.1007

1' in acetonitrile (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	277.11 K	286.73 K	296.48 K	305.76 K
5.79E+02	5.4236	5.3192	5.1343	5.1343
4.63E+02	5.3352	5.2351	5.0591	5.0591
3.47E+02	5.2271	5.1337	4.9714	4.9714
2.32E+02	5.0986	5.0140	4.8713	4.8713
1.16E+02	4.9239	4.8601	4.7468	4.7468
5.79E+01	4.8166	4.7654	4.6747	4.6747
1.45E+01	4.7207	4.6827	4.6153	4.6153
5.79E+00	4.6987	4.6635	4.6012	4.6012
2.89E+00	4.6938	4.6588	4.5984	4.5984
1.45E+00	4.6903	4.6603	4.5952	4.5952

1' in benzene (NH) [exp. 1]				
	Chemical Shifts (ppm)			
conc. (mM)	280.04 K	291.71 K	301.45 K	311.10 K
1.29E+02	7.0830	6.7850	6.1967	6.1967
1.03E+02	6.9760	6.6269	6.0400	6.0400
7.74E+01	6.8120	6.4468	5.8380	5.8380
5.16E+01	6.5618	6.1609	5.5231	5.5231
2.58E+01	6.0450	5.6130	4.9635	4.9635
1.29E+01	5.4634	5.0284	4.4298	4.4298
2.58E+00	4.1819	3.8990	3.6042	3.6042
1.03E+00	3.6845	3.5305	3.3897	3.3897
5.16E-01	3.4720	3.3857	3.3150	3.3150
1.29E-01	3.2810	3.2600	3.2526	3.2526

1' in benzene (NH) [exp. 2]				
	Chemical Shifts (ppm)			
conc. (mM)	280.26 K	291.71 K	301.45 K	311.10 K
1.35E+02	7.0982	6.7740	6.2155	6.2155
1.08E+02	6.9873	6.6477	6.0693	6.0693
8.08E+01	6.8080	6.4710	5.8608	5.8608
5.39E+01	6.5852	6.1890	5.5507	5.5507
2.69E+01	6.0790	5.6489	4.9972	4.9972
1.35E+01	5.4894	5.0515	4.4492	4.4492
2.69E+00	4.1833	3.9005	3.6049	3.6049
1.08E+00	3.6960	3.5386	3.3962	3.3962
5.39E-01	3.4780	3.3895	3.3157	3.3157
1.35E-01	3.2800	3.2600	3.2465	3.2465
5.39E-02	3.2440	3.2380	3.2333	3.2333

1' in benzene (NH) [exp. 3]				
	Chemical Shifts (ppm)			
conc. (mM)	280.26 K	291.61 K	301.45 K	310.90 K
1.36E+02	7.1068	6.7830	6.2286	6.2286
1.09E+02	7.0032	6.6597	6.0802	6.0802
8.14E+01	6.8429	6.4800	5.8724	5.8724
5.43E+01	6.5896	6.1982	5.5614	5.5614
2.71E+01	6.0811	5.6506	4.9967	4.9967
1.36E+01	5.5076	5.0702	4.4660	4.4660
2.71E+00	4.2260	3.9362	3.6270	3.6270
1.09E+00	3.7061	3.5458	3.3986	3.3986
5.43E-01	3.4876	3.3967	3.3203	3.3203
1.36E-01	3.2865	3.2683	3.2543	3.2543
5.43E-02	3.2420	3.2380	3.2426	3.2426

1' in chloroform (NH) [exp. 1]				
	Chemical Shifts (ppm)			
conc. (mM)	277.10 K	286.68 K	297.58 K	304.04 K
1.12E+02	5.1836	5.0437	4.8218	4.8218
8.97E+01	5.0725	4.9392	4.7306	4.7306
6.72E+01	4.9348	4.8113	4.6231	4.6231
4.48E+01	4.7599	4.6533	4.4935	4.4935
2.24E+01	4.5299	4.4509	4.3365	4.3365
1.12E+01	4.3714	4.3171	4.2388	4.2388
2.24E+00	4.2229	4.1964	4.1535	4.1535
8.97E-01	4.1974	4.1756	4.1389	4.1389
4.48E-01	4.1900	4.1696	4.1364	4.1364
1.12E-01	4.1836	4.1635	4.1305	4.1305

1' in chloroform (NH) [exp. 2]				
	Chemical Shifts (ppm)			
conc. (mM)	277.21 K	286.63 K	296.41 K	303.84 K
3.81E+02	5.8488	5.6985	5.4432	5.4432
3.05E+02	5.7329	5.5825	5.3286	5.3286
2.29E+02	5.5777	5.4268	5.1798	5.1798
1.52E+02	5.3727	5.2268	4.9930	4.9930
7.62E+01	---	---	---	---
3.81E+01	4.6902	4.5929	4.4498	4.4498
9.53E+00	4.3460	4.2975	4.2250	4.2250
3.81E+00	4.2516	4.2200	4.1693	4.1693
1.91E+00	4.2176	4.1925	4.1513	4.1513
9.53E-01	4.1994	4.1773	4.1411	4.1411
4.76E-01	4.1908	4.1700	4.1339	4.1339

1' in chloroform (NH) [exp. 3]				
	Chemical Shifts (ppm)			
conc. (mM)	283.01 K	292.66 K	302.37 K	311.81 K
3.88E+02	5.7553	5.6065	5.3302	5.3302
3.10E+02	5.6391	5.4899	5.2171	5.2171
2.33E+02	5.4881	5.3409	5.0765	5.0765
1.55E+02	5.2737	5.1324	4.8871	4.8871
7.76E+01	4.9273	4.8076	4.6060	4.6060
3.88E+01	4.6271	4.5372	4.3920	4.3920
3.10E+01	4.5605	4.4778	4.3469	4.3469
1.55E+01	4.3933	4.3348	4.2428	4.2428
7.76E+00	4.2898	4.2502	4.1833	4.1833

1' in dichloromethane (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	276.21 K	283.92 K	291.76 K	299.57 K
3.98E+02	5.6930	5.5655	5.3303	5.3303
3.18E+02	5.5785	5.4500	5.2135	5.2135
2.39E+02	5.4257	5.2950	5.0636	5.0636
1.59E+02	5.2080	5.0786	4.8650	4.8650
7.95E+01	4.8551	4.7497	4.5764	4.5764
3.98E+01	4.5709	4.4933	4.3685	4.3685
3.18E+01	4.5066	4.4365	4.3232	4.3232
3.98E+00	4.1930	4.1693	4.1271	4.1271

2 in acetonitrile (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	283.04 K	292.56 K	302.25 K	311.55 K
2.01E+02	4.8420	4.7551	4.6088	4.6088
1.61E+02	4.7666	4.6868	4.5536	4.5536
1.21E+02	4.6819	4.6113	4.4934	4.4934
8.05E+01	4.5859	4.5268	4.4280	4.4280
4.02E+01	4.4735	4.4297	4.3557	4.3557
2.01E+01	4.4109	4.3763	4.3145	4.3145
5.03E+00	4.3608	4.3344	4.2838	4.2838
2.01E+00	---	---	x4.2751	x4.2751
1.01E+00	4.3449	4.3205	4.2749	4.2749
5.03E-01	4.3449	4.3198	4.2738	4.2738

2 in benzene (NH) [exp1]				
	Chemical Shifts (ppm)			
conc. (mM)	280.26 K	291.61 K	301.45 K	310.90 K
4.20E+01	6.5024	6.0749	5.3822	5.3822
3.36E+01	6.3369	5.8917	5.1903	5.1903
2.52E+01	6.1036	5.6418	4.9343	4.9343
1.68E+01	5.7616	5.2844	4.5880	4.5880
8.39E+00	5.1277	4.6587	4.0420	4.0420
4.20E+00	4.4731	4.0676	3.5950	3.5950
8.39E-01	3.3952	3.2259	3.0765	3.0765
3.36E-01	3.1130	3.0357	2.9767	2.9767
1.68E-01	3.0052	2.9680	2.9402	2.9402
8.39E-02	2.9510	2.9330	2.9225	2.9225

2 in benzene (NH) [exp2]				
	Chemical Shifts (ppm)			
conc. (mM)	280.26 K	291.71 K	301.45 K	311.10 K
3.95E+01	6.4566	6.0227	5.3286	5.3286
3.16E+01	6.2920	5.8415	5.1374	5.1374
2.37E+01	6.0545	5.5870	4.8794	4.8794
1.58E+01	5.7044	5.2214	4.5299	4.5299
7.90E+00	5.0625	4.5927	3.9879	3.9879
3.95E+00	4.4254	4.0255	3.5649	3.5649
7.90E-01	3.3686	3.2067	3.0645	3.0645
3.16E-01	3.1001	3.0295	2.9721	2.9721
1.58E-01	2.9965	2.9622	2.9403	2.9403
7.90E-02	2.9408	2.9250	2.9212	2.9212
5.53E-02	2.9290	2.9210	2.9170	2.9170

2 in benzene (NH) [exp3]

Chemical Shifts (ppm)

conc. (mM)	280.04 K	291.71 K	301.45 K	310.90 K
4.03E+01	6.4741	6.0421	5.3488	5.3488
3.23E+01	6.3088	5.8605	5.1583	5.1583
2.42E+01	6.0860	5.6128	4.9046	4.9046
1.61E+01	5.7247	5.2240	4.5512	4.5512
8.06E+00	5.0762	4.6098	4.0017	4.0017
4.03E+00	4.4455	4.0419	3.5764	3.5764
8.06E-01	3.4040	3.2331	3.0830	3.0830
3.23E-01	3.1070	3.0330	2.9760	2.9760
1.61E-01	3.0020	2.9658	2.9403	2.9403
8.06E-02	2.9510	2.9351	2.9249	2.9249

2 in chloroform (NH) [exp. 1]

Chemical Shifts (ppm)

conc. (mM)	276.87 K	286.73 K	296.39 K	303.61 K
3.32E+02	5.1195	5.0149	4.8416	4.8416
2.65E+02	5.0051	4.9024	4.7366	4.7366
1.99E+02	4.8473	4.7503	4.5956	4.5956
1.33E+02	4.6443	4.5589	4.4238	4.4238
6.63E+01	4.3676	4.3016	4.2001	4.2001
3.32E+01	4.1647	4.1180	4.0441	4.0441
8.29E+00	---	---	3.9142	3.9142
3.32E+00	3.9384	3.9170	3.8816	3.8816
1.66E+00	3.9234	3.9039	3.8716	3.8716
8.29E-01	3.9178	3.8994	3.8673	3.8673
4.15E-01	3.9129	3.8957	3.8639	3.8639

2 in chloroform (NH) [exp. 2]

Chemical Shifts (ppm)

conc. (mM)	276.19 K	285.86 K	295.71 K	303.43 K
4.65E+02	5.3126	5.2033	5.0123	5.0123
3.72E+02	5.1951	5.0861	4.9058	4.9058
2.79E+02	5.0350	4.9291	4.7588	4.7588
1.86E+02	4.8152	4.7174	4.5635	4.5635
9.29E+01	4.5006	4.4226	4.3022	4.3022
4.65E+01	4.2510	4.1950	4.1066	4.1066
1.16E+01	4.0073	overlap	3.9308	3.9308
4.65E+00	3.9496	3.9266	3.8895	3.8895
5.81E-01	3.9135	3.8953	3.8653	3.8653

2 in dichloromethane (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	276.12 K	283.95 K	291.73 K	299.55 K
3.71E+02	5.6215	5.4986	5.2634	5.2634
2.97E+02	5.5002	5.3763	5.1426	5.1426
2.22E+02	5.3335	5.2051	4.9772	4.9772
1.48E+02	5.0869	4.9706	4.7560	4.7560
7.41E+01	4.7062	4.6048	4.4290	4.4290
3.71E+01	4.3886	4.3135	4.1869	4.1869
1.85E+01	4.1754	4.1228	4.0368	4.0368
3.71E+00	3.9555	3.9323	3.8926	3.8926
2' in acetonitrile (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	279.15 K	288.82 K	298.69 K	308.43 K
8.32E+01	6.4010	6.1270	5.6357	5.6357
6.66E+01	6.2366	5.9750	5.5001	5.5001
4.99E+01	6.0430	5.7686	5.3267	5.3267
3.33E+01	5.7634	5.5048	5.1144	5.1144
1.66E+01	5.3463	5.1429	4.8379	4.8379
8.32E+00	5.0276	4.8786	4.6631	4.6631
4.16E+00	4.8353	4.7350	4.5867	4.5867
1.66E+00	4.6928	4.6296	4.5287	4.5287
2' in benzene (NH) [exp1]				
	Chemical Shifts (ppm)			
conc. (mM)	280.15 K	291.61 K	301.25 K	310.10 K
5.04E+01	9.0822	8.7875	8.2084	8.2084
4.03E+01	9.0121	8.6949	8.0770	8.0770
3.02E+01	8.9008	8.5582	7.8848	7.8848
2.02E+01	8.7330	8.3392	7.5847	7.5847
1.01E+01	8.3582	7.8736	6.9739	6.9739
5.04E+00	7.8855	7.3036	6.2807	6.2807
1.01E+00	6.4173	5.6852	4.6569	4.6569
4.03E-01	5.2648	4.6103	3.8693	3.8693
2.02E-01	4.5750	4.0619	3.5616	3.5616
1.01E-01	4.0367	3.6920	3.3850	3.3850
7.05E-02	3.8270	3.5590	---	---

2' in benzene (NH) [exp2]

Chemical Shifts (ppm)

conc. (mM)	280.04 K	291.71 K	301.45 K	310.90 K
5.07E+01	9.0866	8.7936	8.2189	8.2189
4.06E+01	9.0145	8.7008	8.0846	8.0846
3.04E+01	8.9130	8.5671	7.9000	7.9000
2.03E+01	8.7375	8.3473	7.5959	7.5959
1.01E+01	8.3639	7.8818	6.9840	6.9840
5.07E+00	7.8983	7.3203	6.2999	6.2999
1.01E+00	6.5648	5.8407	4.7956	4.7956
4.06E-01	5.3180	4.6610	3.9113	3.9113
2.03E-01	4.6400	4.1360	3.6298	3.6298
1.01E-01	4.0470	3.6914	---	---

2' in benzene (NH)

Chemical Shifts (ppm)

conc. (mM)	280.15 K	291.40 K	301.45 K	310.70 K
4.72E+01	9.0570	8.7527	8.1574	8.1574
3.78E+01	8.9830	8.6600	8.0291	8.0291
2.83E+01	8.8740	8.5239	7.8390	7.8390
1.89E+01	8.6990	8.2980	7.5306	7.5306
9.45E+00	8.3120	7.8221	6.9103	6.9103
4.72E+00	7.8250	7.2320	6.2010	6.2010
9.45E-01	6.2510	5.5248	4.5310	4.5310
3.78E-01	5.2100	4.5590	3.8480	3.8480
1.89E-01	4.4840	3.9900	3.5220	3.5220

2' in chloroform (NH)

Chemical Shifts (ppm)

conc. (mM)	277.02 K	286.76 K	296.54 K	303.81 K
3.18E+02	7.1011	6.8722	6.4690	6.4690
2.55E+02	6.9650	6.7192	6.3107	6.3107
1.91E+02	6.7584	6.4961	6.0763	6.0763
1.27E+02	6.4250	6.1708	5.7497	5.7497
6.37E+01	5.8560	5.6043	5.2129	5.2129
3.18E+01	5.3090	5.0895	4.7665	4.7665
7.96E+00	4.5068	4.3920	4.2339	4.2339
3.18E+00	4.2536	4.1872	4.0953	4.0953
1.59E+00	4.1612	4.1143	4.0504	4.0504

2' in dichloromethane (NH)

Chemical Shifts (ppm)

conc. (mM)	276.18 K	283.91 K	291.77 K	299.54 K
2.37E+02	7.5115	7.3374	7.0093	7.0093
1.89E+02	7.4179	7.2287	6.8837	6.8837
1.42E+02	7.2860	7.0609	6.6876	6.6876
9.46E+01	7.0681	6.8216	6.4044	6.4044
4.73E+01	6.5801	6.3452	5.8983	5.8983
2.37E+01	6.0052	5.7682	5.3494	5.3494
5.92E+00	4.9763	4.8064	4.5430	4.5430
2.37E+00	4.5061	4.3999	4.2276	4.2276
1.18E+00	4.2949	4.2308	4.1452	4.1452
2.96E-01	4.1305	4.1105	4.0850	4.0850

BA in acetonitrile (OH)

Chemical Shifts (ppm)

conc. (mM)	276.10 K	282.98 K	294.68 K	305.90 K
4.57E+02	10.2738	10.1966	9.9416	9.9416
3.66E+02	10.1437	10.0727	9.8350	9.8350
2.74E+02	10.0075	9.9413	9.7230	9.7230
1.83E+02	9.8610	9.8017	9.6056	9.6056
9.15E+01	9.6968	9.6463	9.4762	9.4762
4.57E+01	9.6115	9.5662	9.4115	9.4115
2.29E+01	9.5599	9.5178	9.3715	9.3715
9.15E+00	9.5384	9.4983	9.3560	9.3560
4.57E+00	9.5328	9.4937	9.3548	9.3548

BA in benzene (OH)

Chemical Shifts (ppm)

conc. (mM)	279.04 K	288.62 K	298.46 K	307.55 K
8.60E+00	12.7324	12.2956	11.2723	11.2723
7.74E+00	12.6601	12.2064	11.1517	11.1517
6.88E+00	12.5878	12.1171	11.0310	11.0310
5.16E+00	12.3801	11.8676	10.7000	10.7000
3.44E+00	12.0366	11.4592	10.1858	10.1858
1.72E+00	11.3449	10.6590	9.2491	9.2491
8.60E-01	10.5080	9.7409	8.3182	8.3182
2.15E-01	8.5840	7.8614	6.8460	6.8460

BA in chloroform (OH)				
	Chemical Shifts (ppm)			
conc. (mM)	279.00 K	288.59 K	298.56 K	307.67 K
2.10E+02	13.1634	13.0204	12.6729	12.6729
1.68E+02	13.1277	12.9770	12.6076	12.6076
1.26E+02	13.0772	12.9114	12.5113	12.5113
8.40E+01	12.9998	12.8038	12.3508	12.3508
4.20E+01	12.7934	12.5555	11.9890	11.9890
2.10E+01	12.5305	12.2325	11.5332	11.5332
5.25E+00	11.6784	11.2173	10.2156	10.2156
2.10E+00	10.8811	10.3195	9.2088	9.2088
1.05E+00	10.1701	9.5635	8.4867	8.4867
5.25E-01	9.3687	8.7740	7.8436	7.8436

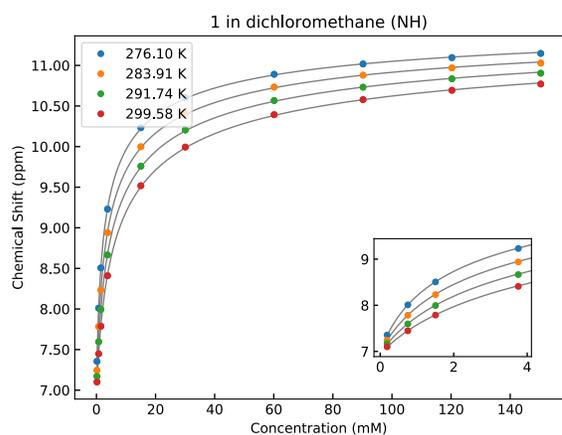
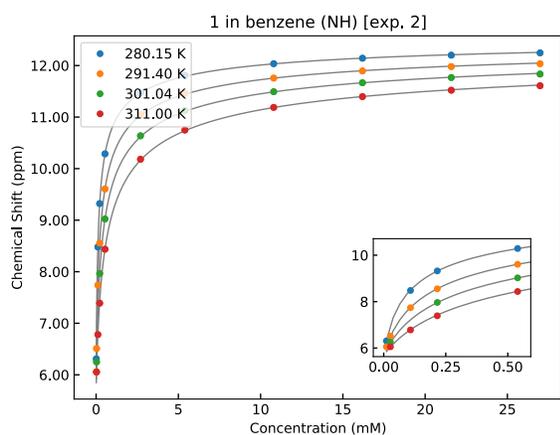
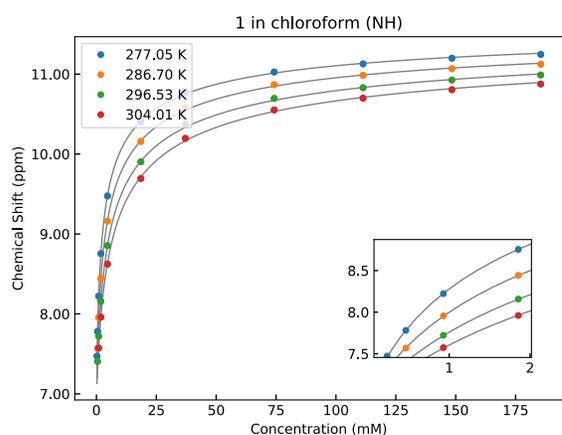
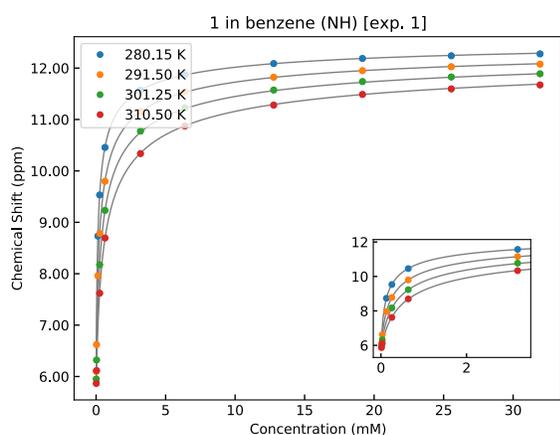
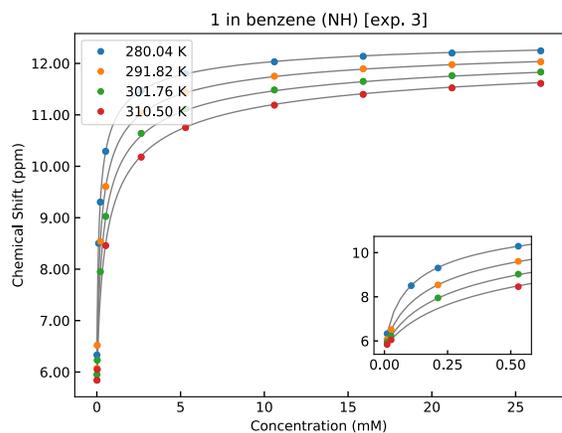
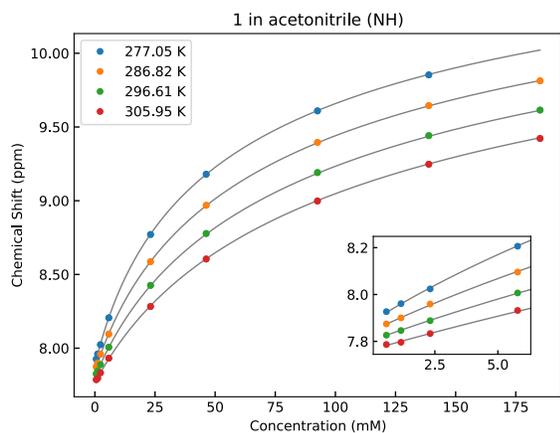
BA in dichloromethane (OH)				
	Chemical Shifts (ppm)			
conc. (mM)	276.16 K	283.95 K	291.84 K	299.62 K
1.65E+02	13.0834	12.9418	12.6126	12.6126
1.32E+02	13.0355	12.8836	12.5337	12.5337
9.92E+01	12.9620	12.7961	12.4106	12.4106
6.62E+01	12.8454	12.6548	12.2179	12.2179
3.31E+01	12.5851	12.3455	11.8002	11.8002
1.65E+01	12.2330	11.9309	11.2623	11.2623
4.14E+00	11.1942	10.7602	9.8737	9.8737
1.65E+00	10.2922	9.8035	8.9106	8.9106
8.27E-01	9.5494	9.0690	8.2771	8.2771
2.07E-01	8.1981	7.8679	7.4199	7.4199

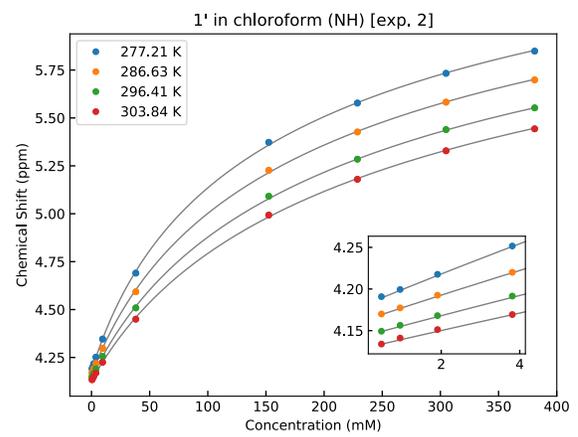
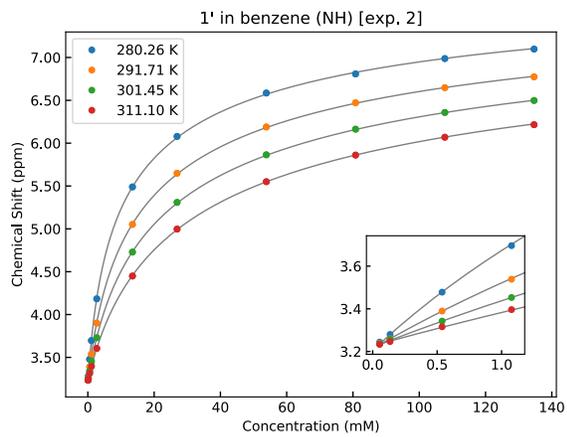
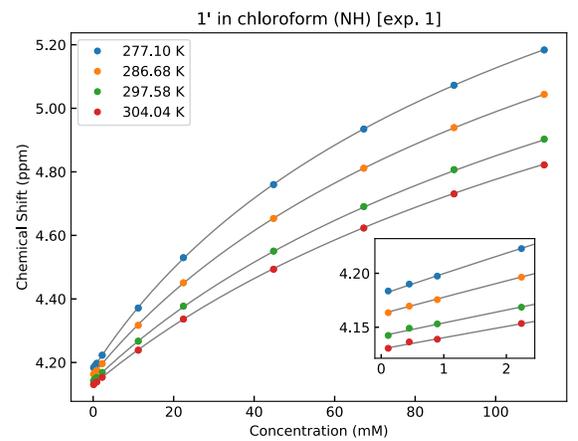
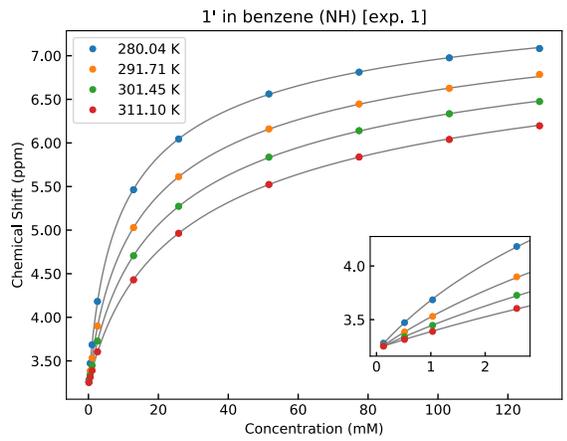
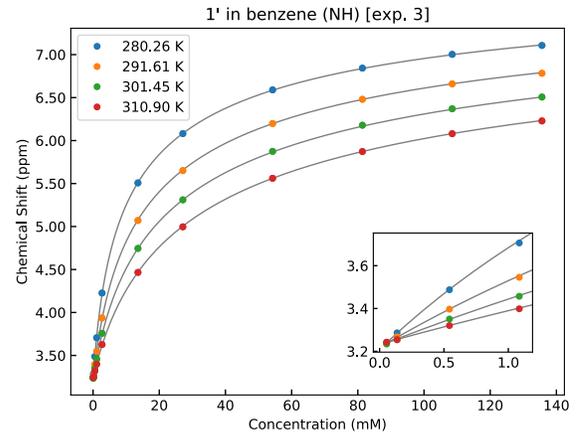
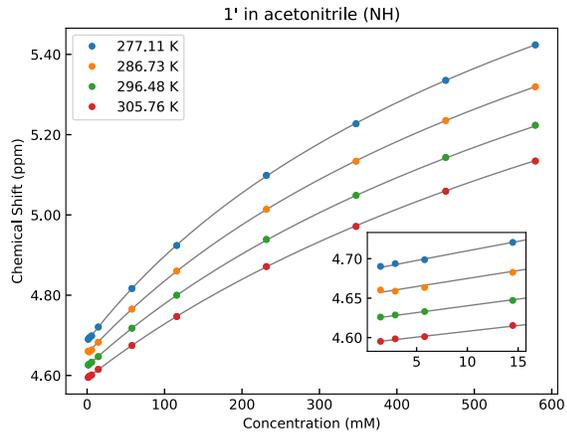
IM in acetonitrile (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	276.22 K	285.87 K	295.72 K	303.43 K
8.25E+02	6.8452	6.7291	6.5353	6.5353
6.60E+02	6.7499	6.6354	6.4481	6.4481
4.95E+02	6.6269	6.5181	6.3446	6.3446
3.30E+02	6.4713	6.3685	6.2125	6.2125
1.65E+02	6.2529	6.1699	6.0390	6.0390
8.25E+01	6.0992	6.0346	5.9327	5.9327
2.06E+01	5.9590	5.9147	5.8423	5.8423
4.13E+00	5.9169	5.8817	5.8182	5.8182

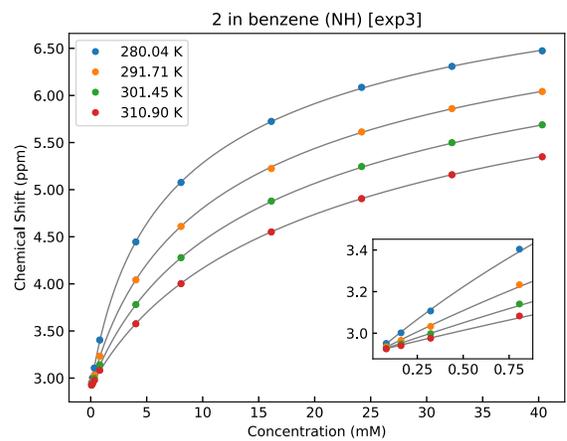
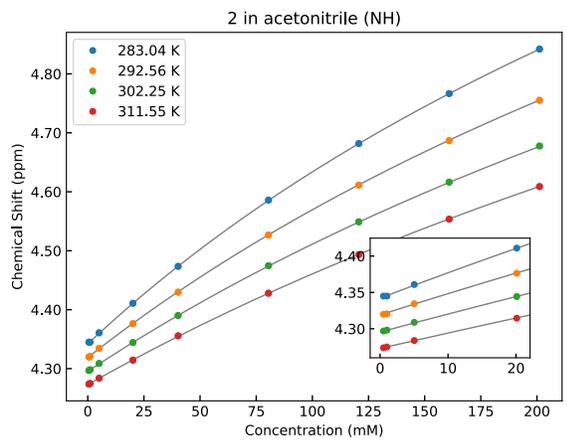
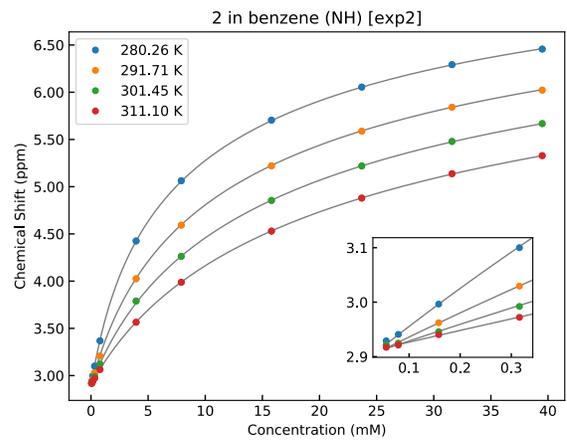
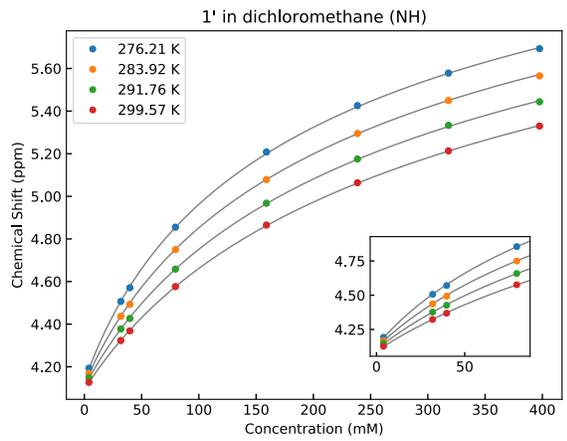
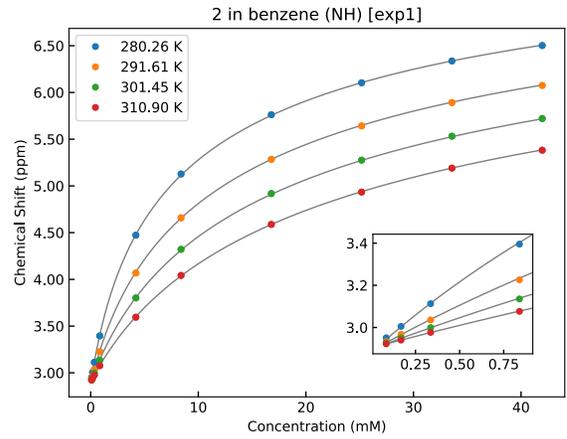
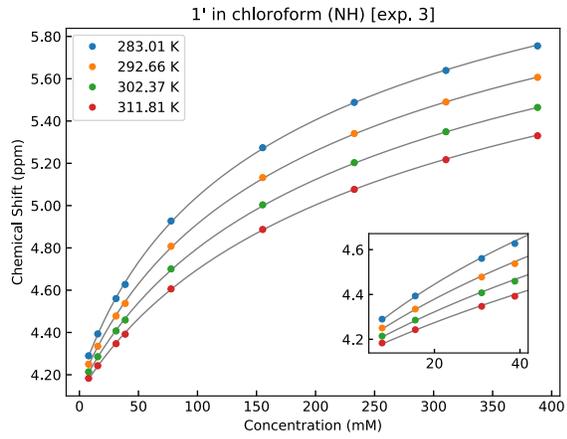
IM in benzene (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	276.10 K	285.94 K	295.59 K	305.28 K
2.49E+02	8.6470	8.4149	7.9250	7.9250
1.99E+02	8.5701	8.3177	7.7951	7.7951
1.49E+02	8.4584	8.1841	7.6183	7.6183
9.96E+01	8.2814	7.9793	7.3581	7.3581
4.98E+01	7.8951	7.5373	6.8410	6.8410
2.49E+01	7.4717	7.0798	6.3385	6.3385
6.23E+00	6.3519	5.9333	5.2705	5.2705
2.49E+00	5.5454	5.2019	4.7396	4.7396
1.25E+00	5.0364	4.7820	4.4787	4.4787
3.11E-01	4.4226	4.3283	4.2433	4.2433
IM in chloroform (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	276.09 K	283.93 K	291.82 K	299.57 K
6.19E+02	7.5596	7.4532	7.2186	7.2186
4.95E+02	7.4703	7.3515	7.1106	7.1106
3.71E+02	7.3246	7.2066	6.9583	6.9583
2.47E+02	7.1215	6.9899	6.7383	6.7383
1.24E+02	6.7336	6.6023	6.3593	6.3593
6.19E+01	6.3411	6.2189	6.0108	6.0108
1.55E+01	5.7484	5.6778	5.5551	5.5551
6.19E+00	5.5554	5.5094	5.4297	5.4297
3.09E+00	5.4786	5.4433	5.3820	5.3820
7.73E-01	5.4133	5.3902	5.3414	5.3414
IM in dichloromethane (NH)				
	Chemical Shifts (ppm)			
conc. (mM)	276.18 K	283.92 K	291.82 K	299.53 K
6.66E+02	7.4135	7.2821	7.0254	7.0254
5.33E+02	7.3168	7.1788	6.9193	6.9193
4.00E+02	7.1703	7.0359	6.7629	6.7629
2.66E+02	6.9551	6.8067	6.5347	6.5347
1.33E+02	6.5765	6.4303	6.1762	6.1762
6.66E+01	6.1944	6.0618	5.8450	5.8450
1.67E+01	5.6536	5.5820	5.4706	5.4706
6.66E+00	5.4835	5.4385	5.3655	5.3655
8.33E-01	---	---	---	---

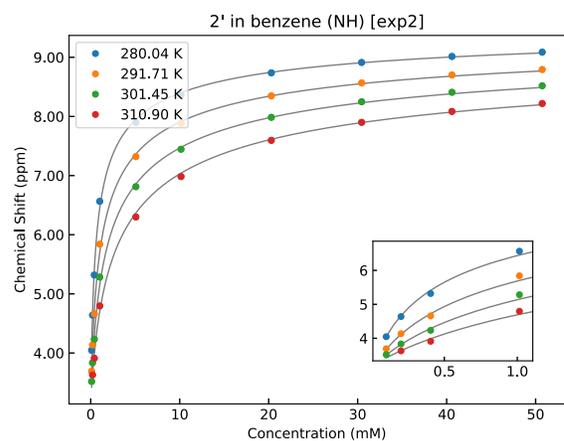
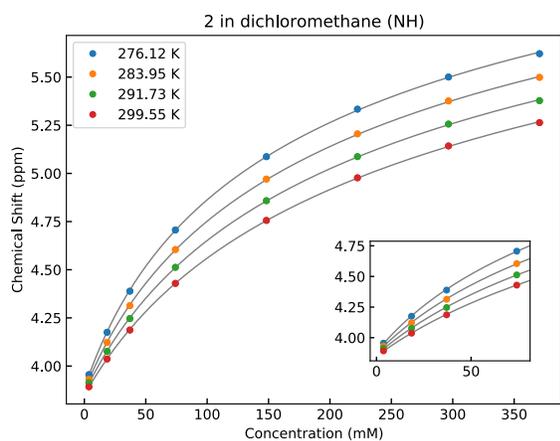
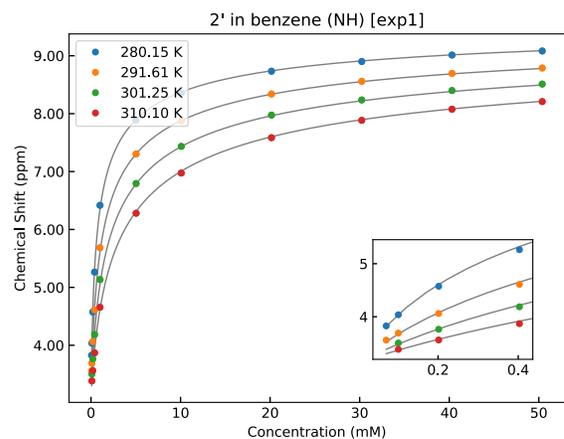
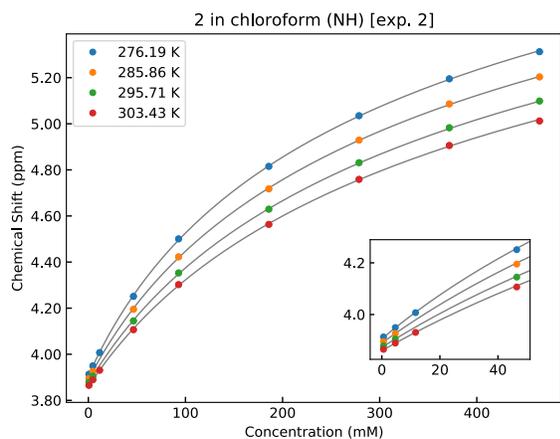
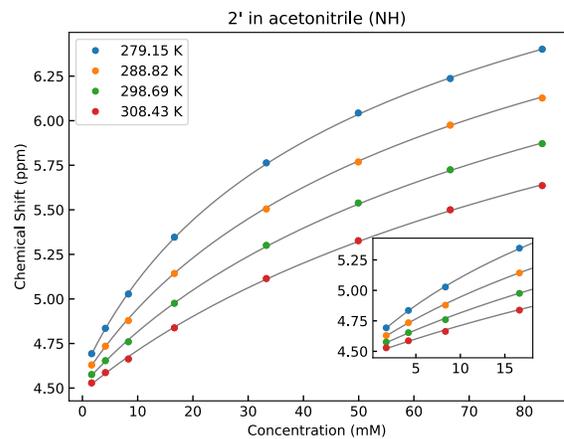
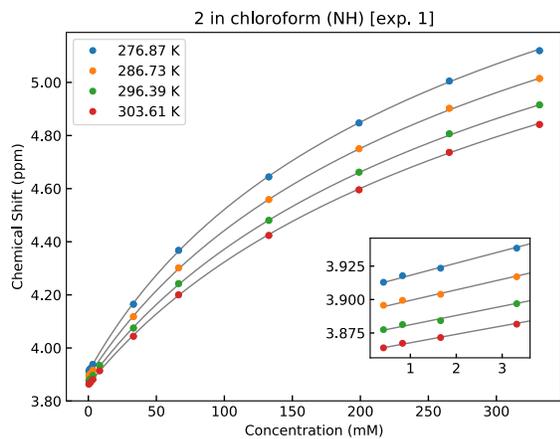
APPENDIX B:

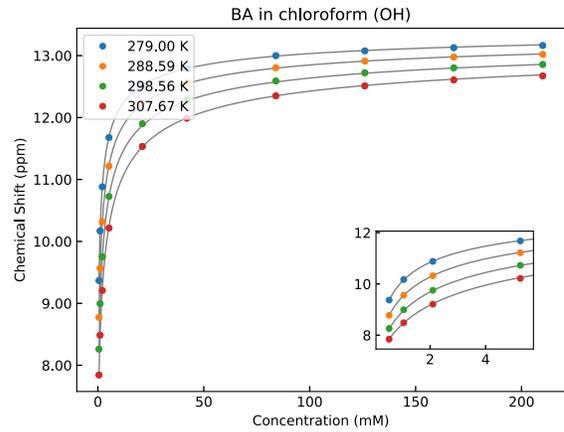
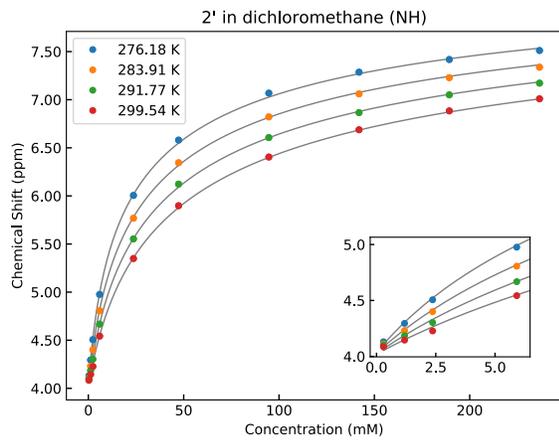
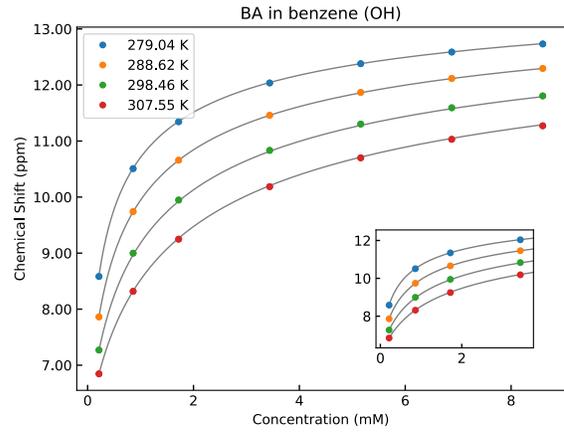
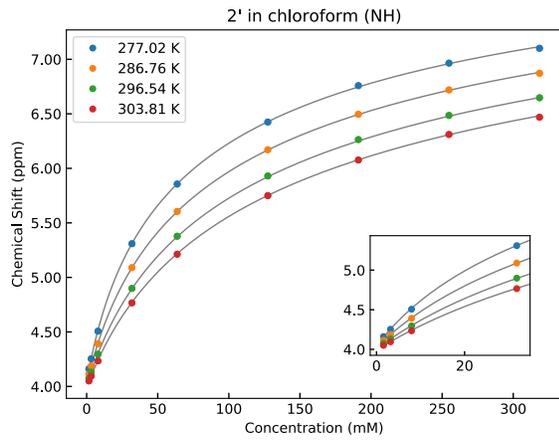
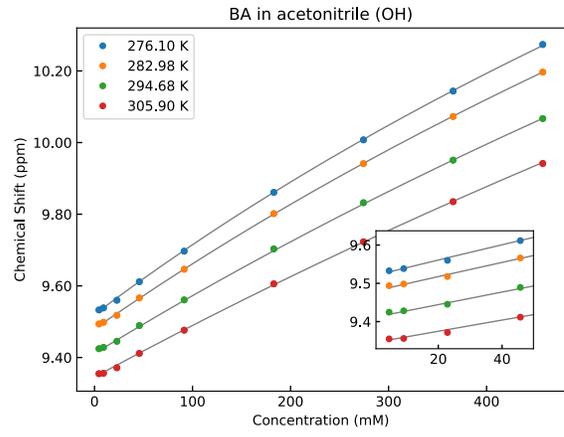
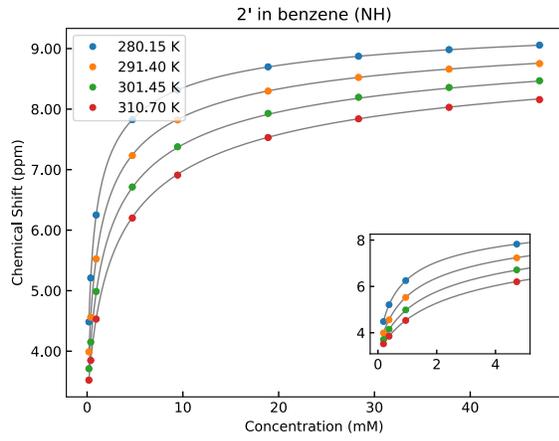
Plots of fits for obtaining thermodynamic values from NMR spectra

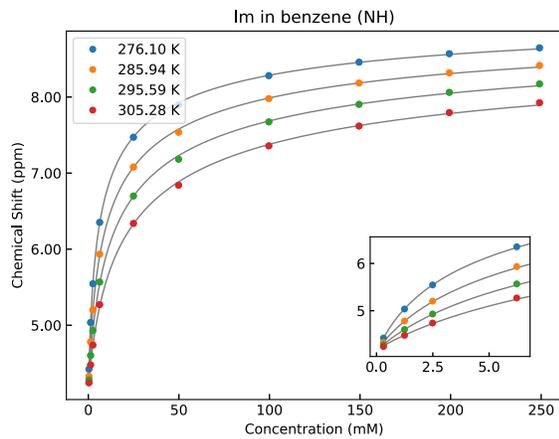
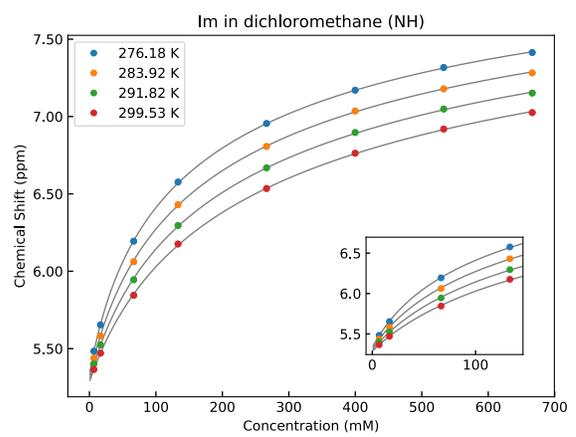
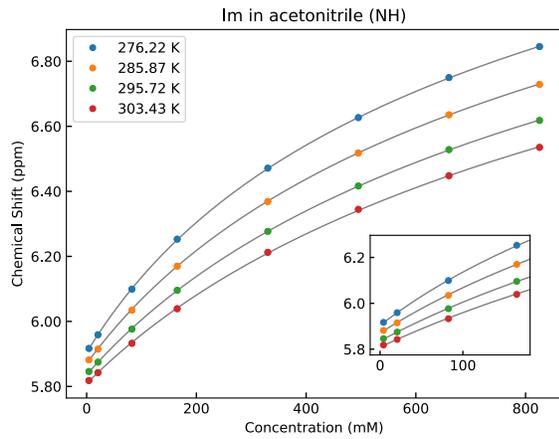
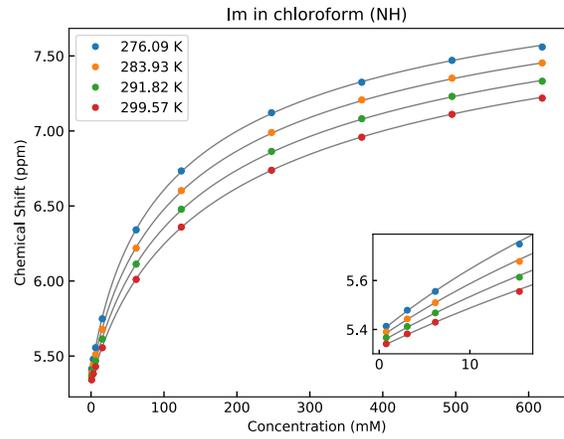
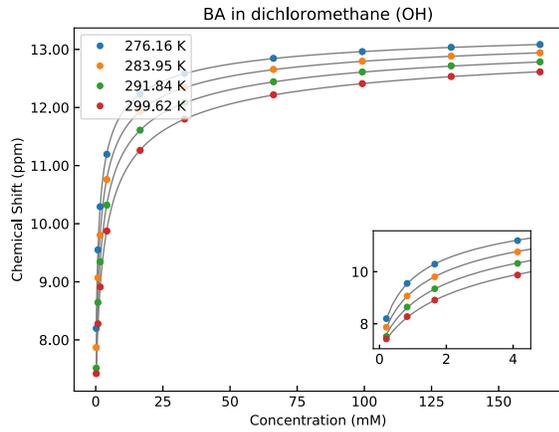






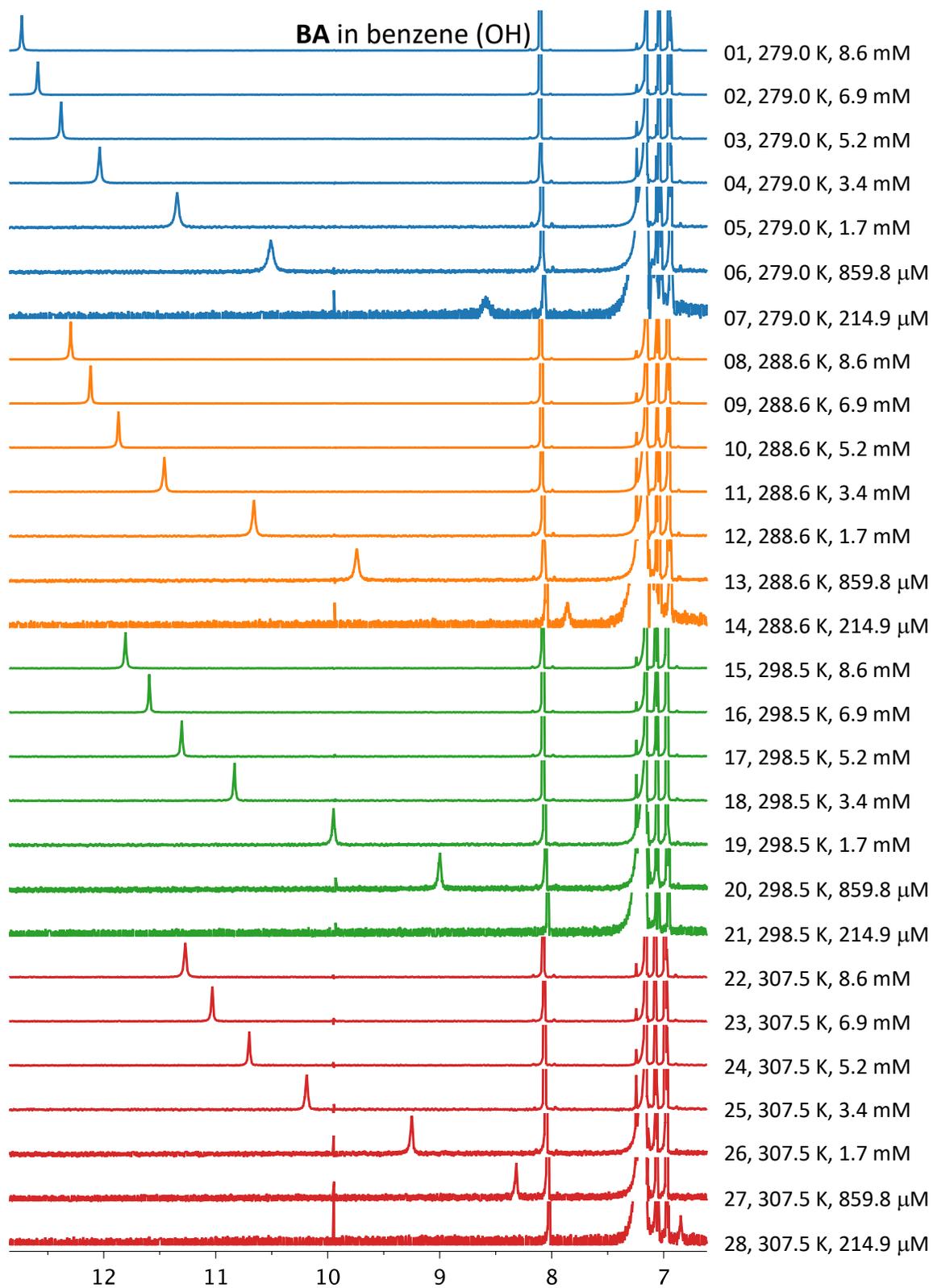


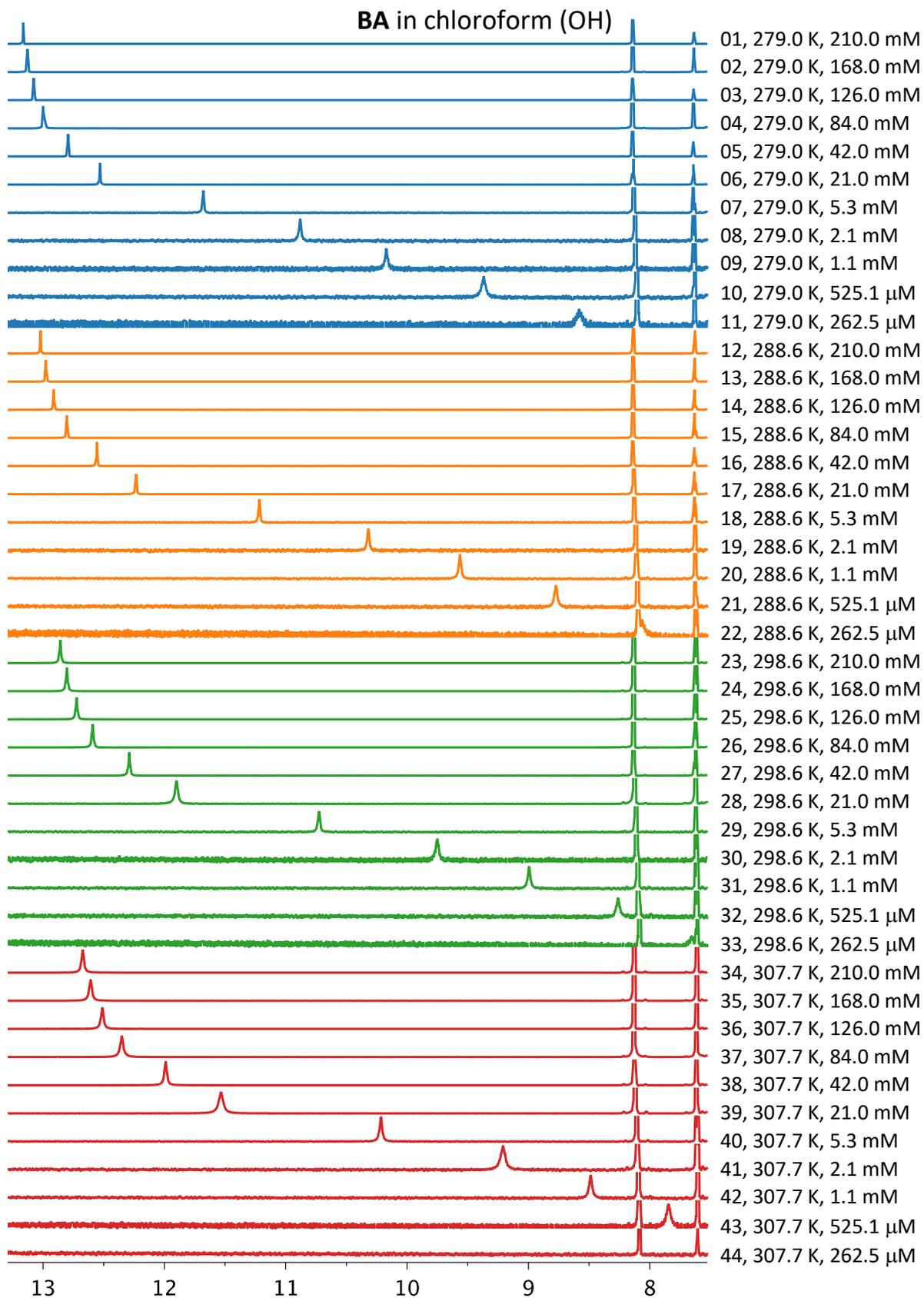




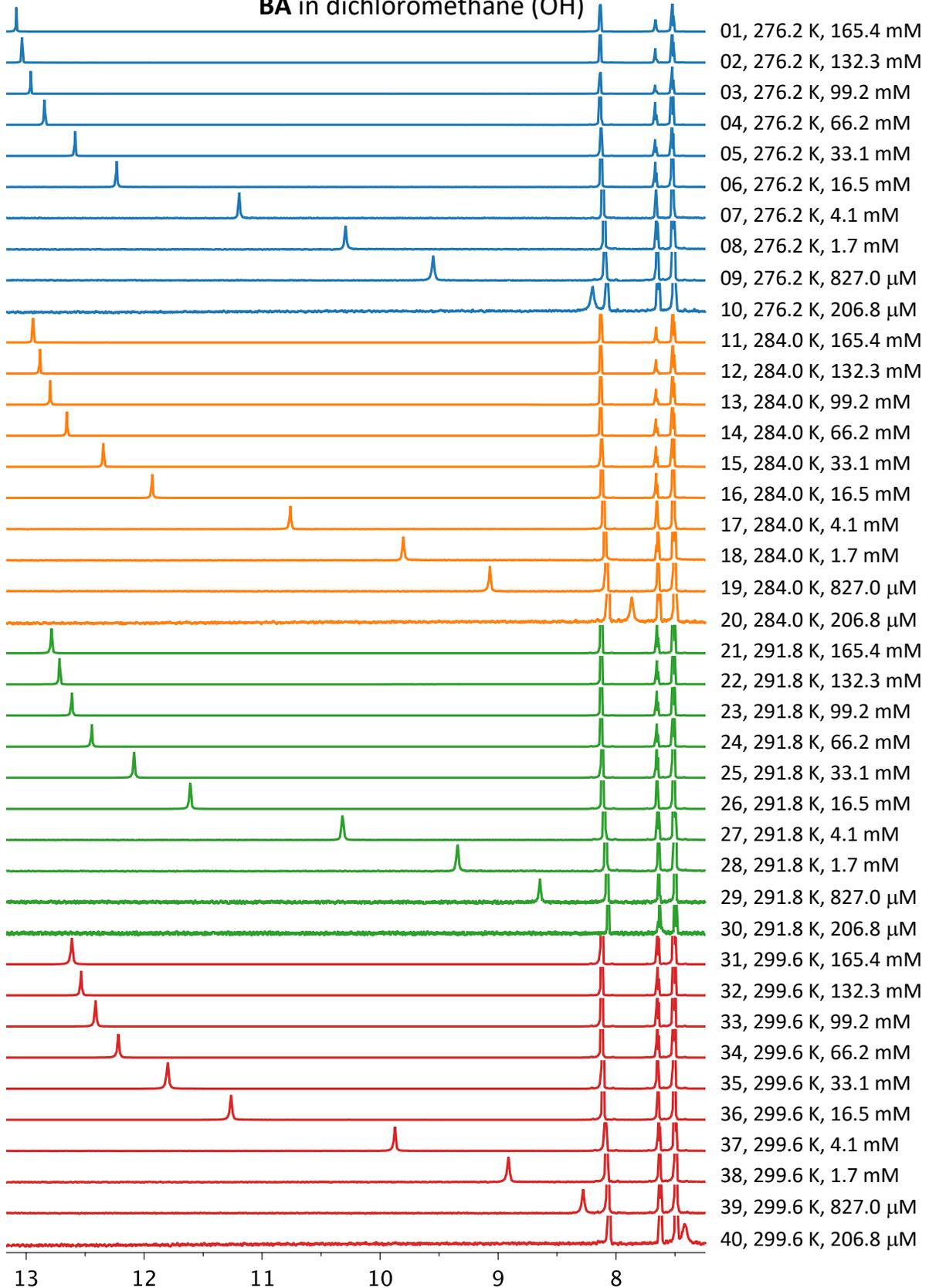
APPENDIX C:

Stacked ^1H NMR spectra used for obtaining dimerization energies

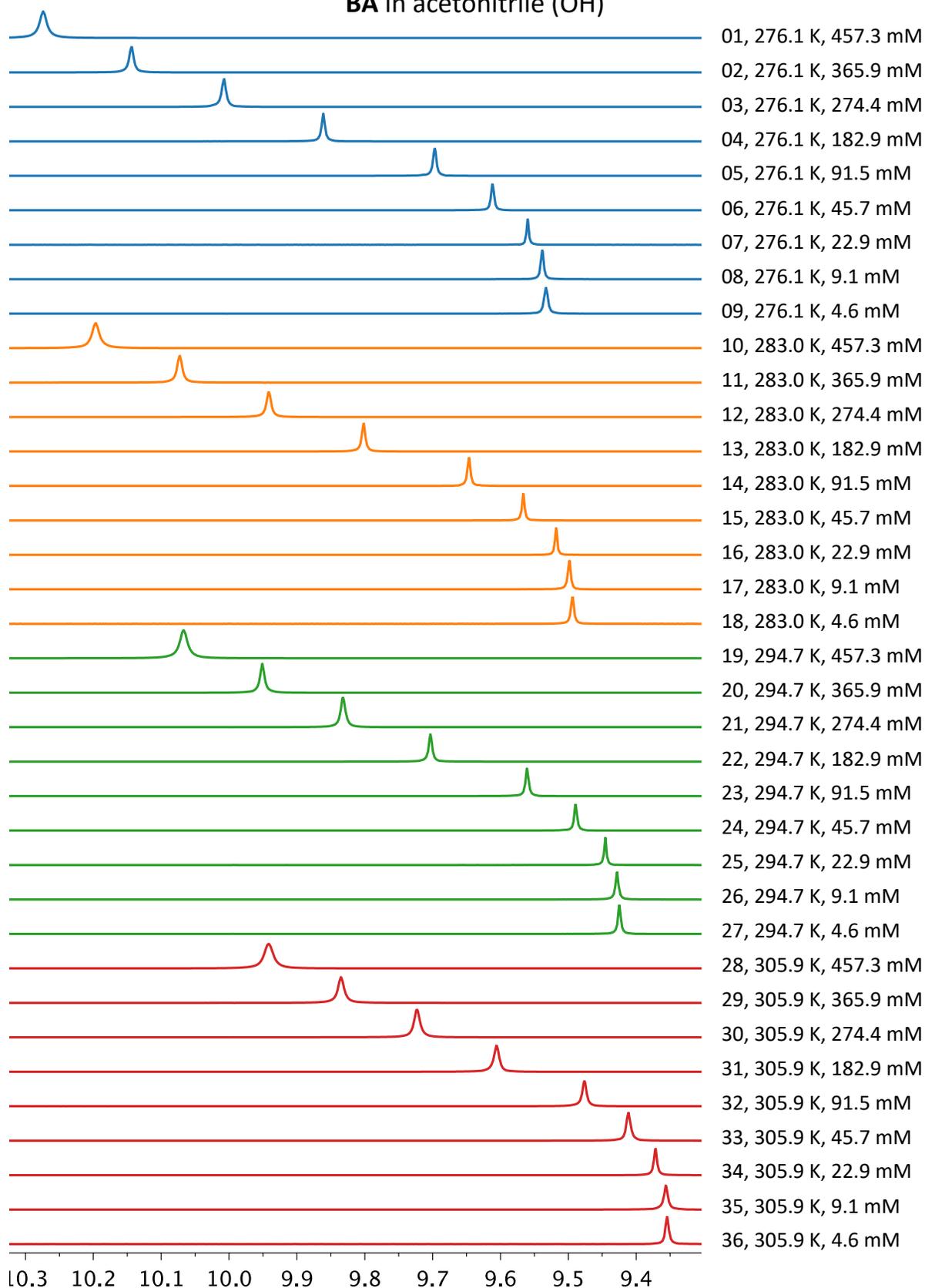


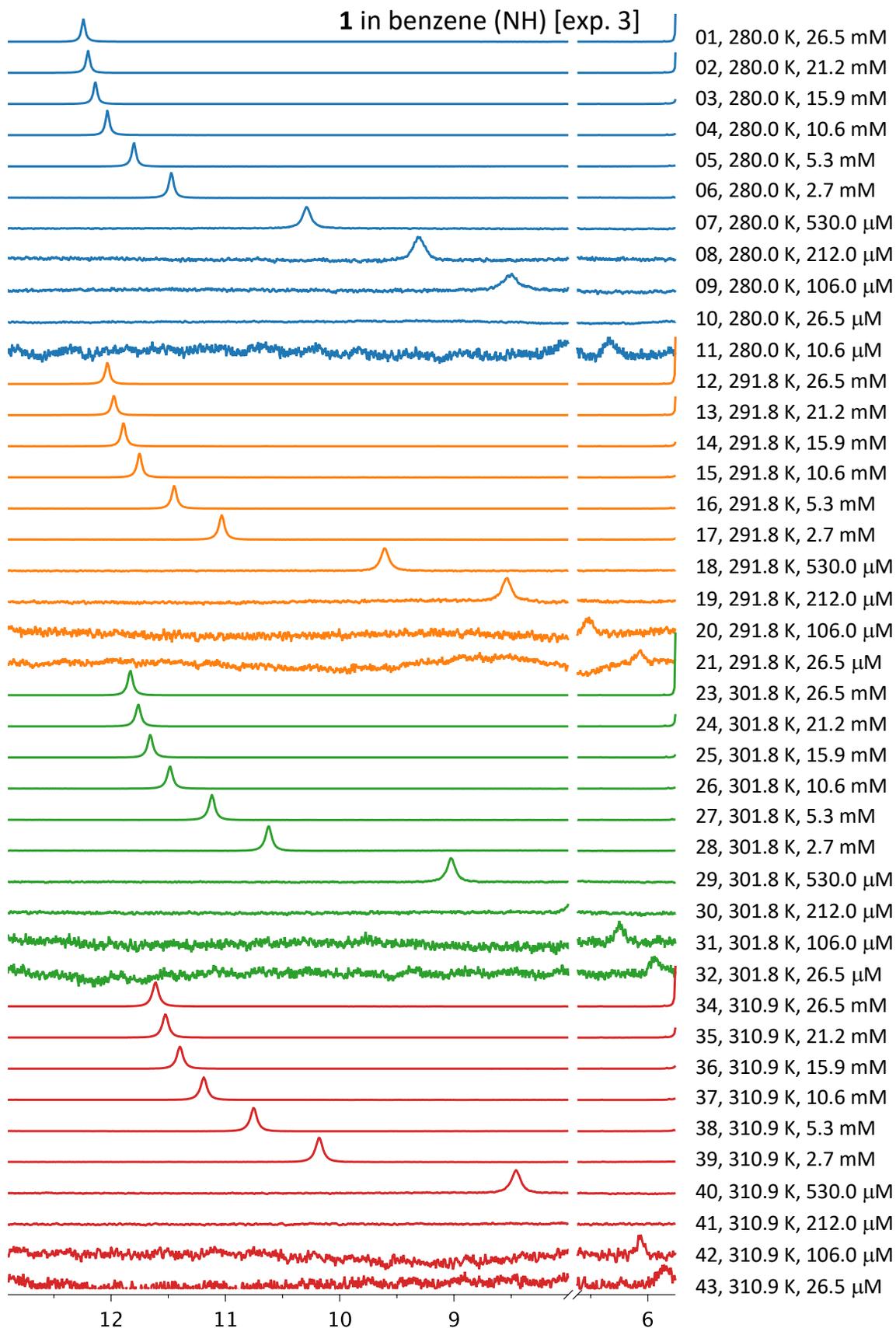


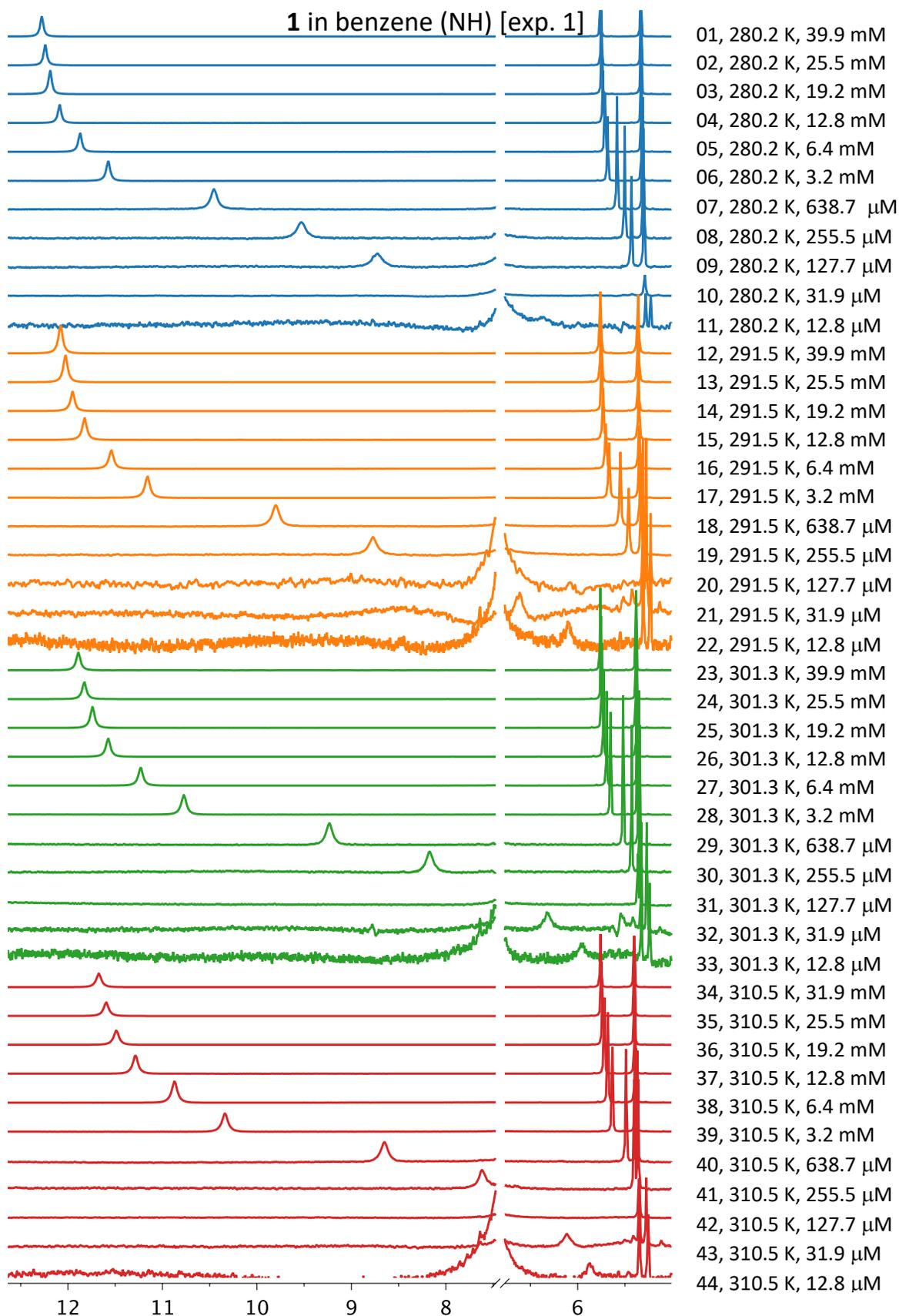
BA in dichloromethane (OH)

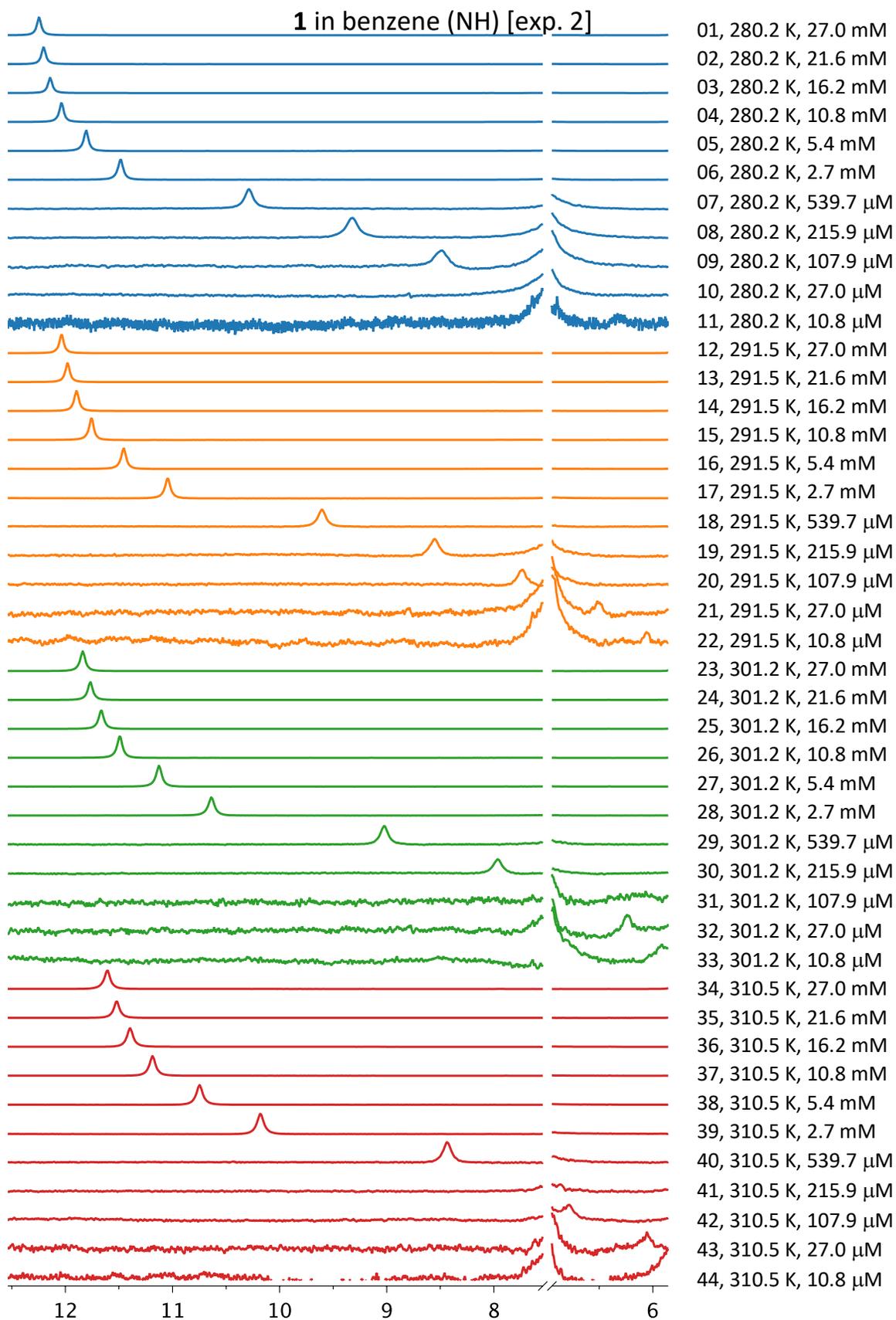


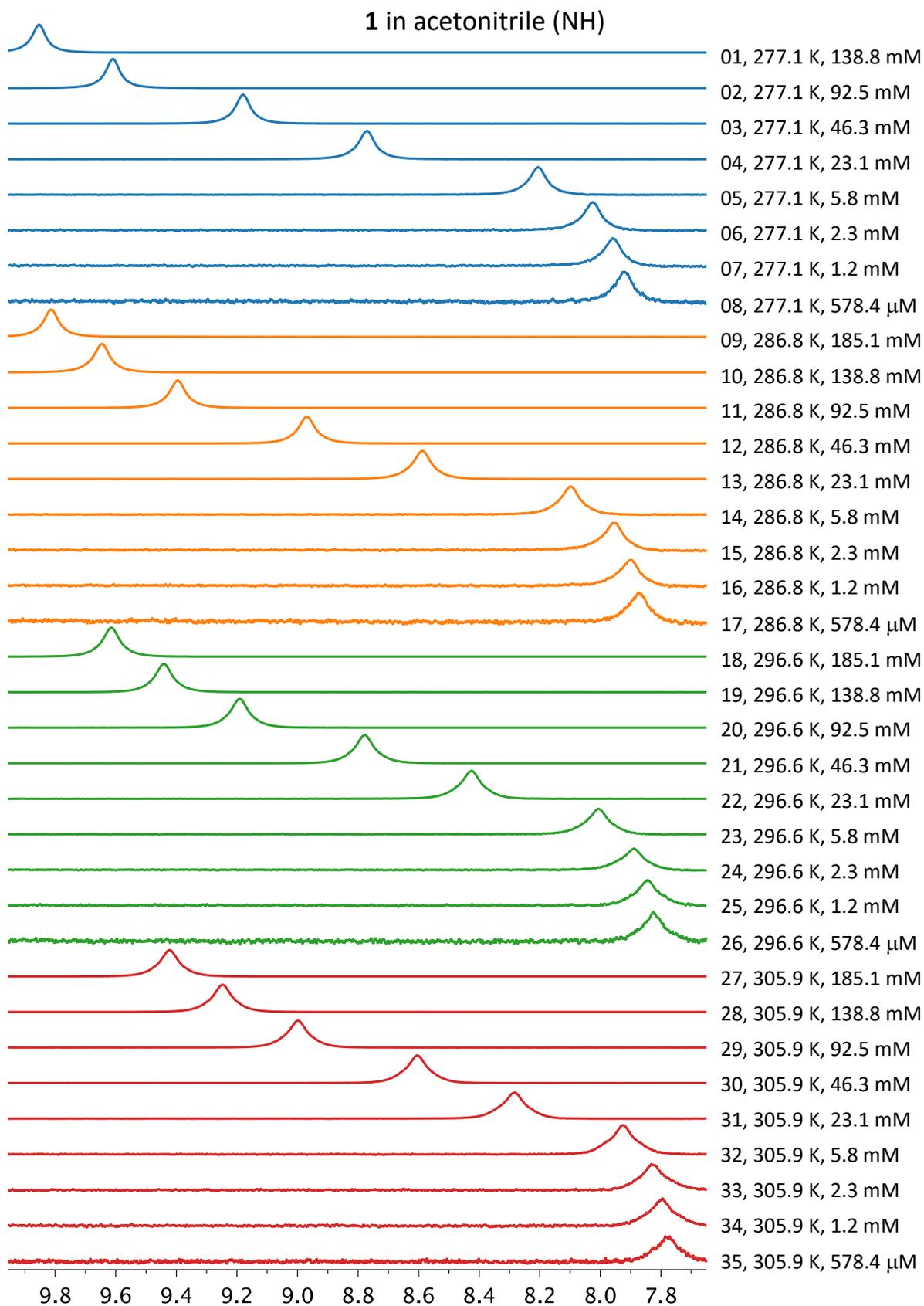
BA in acetonitrile (OH)



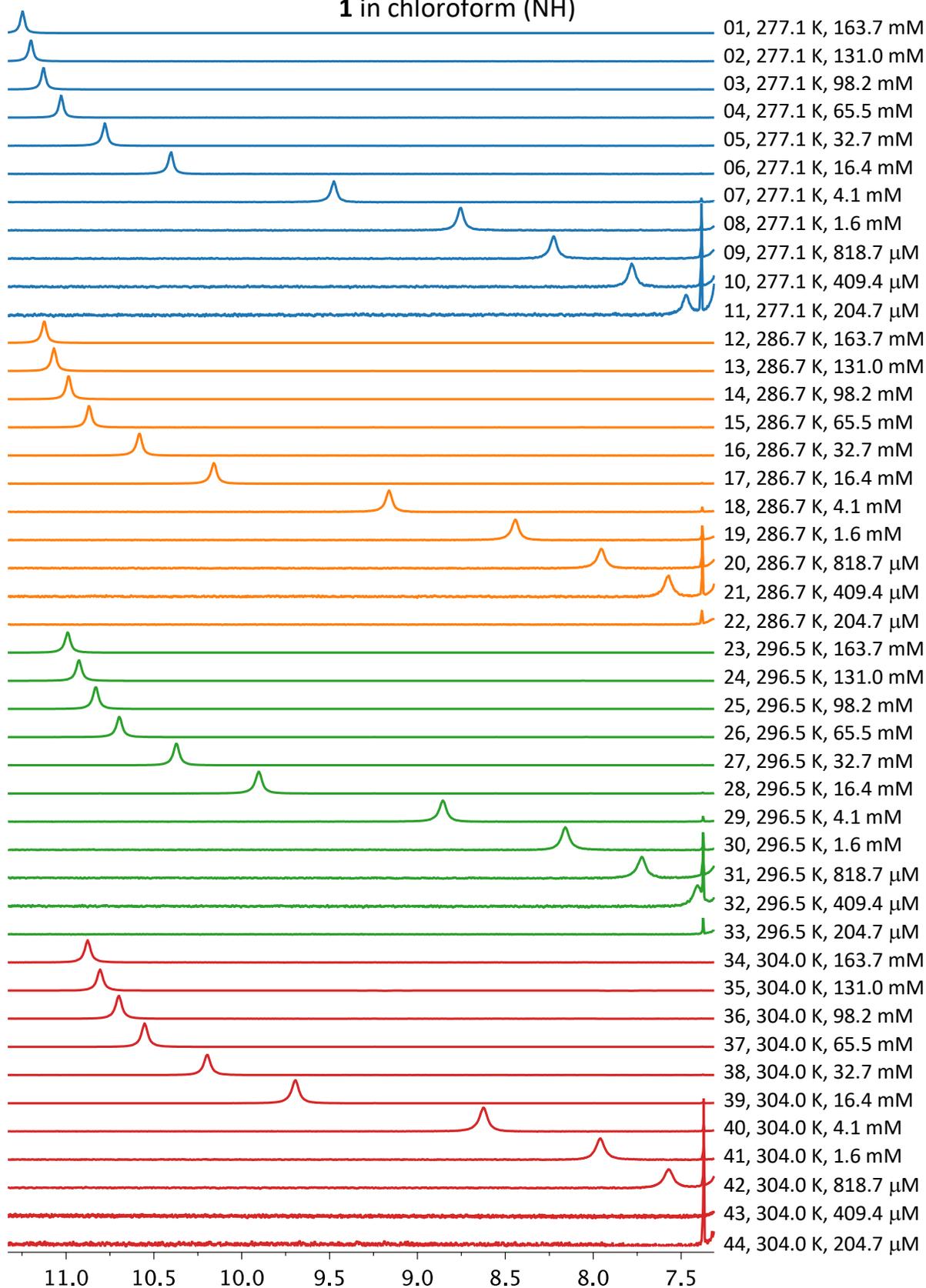




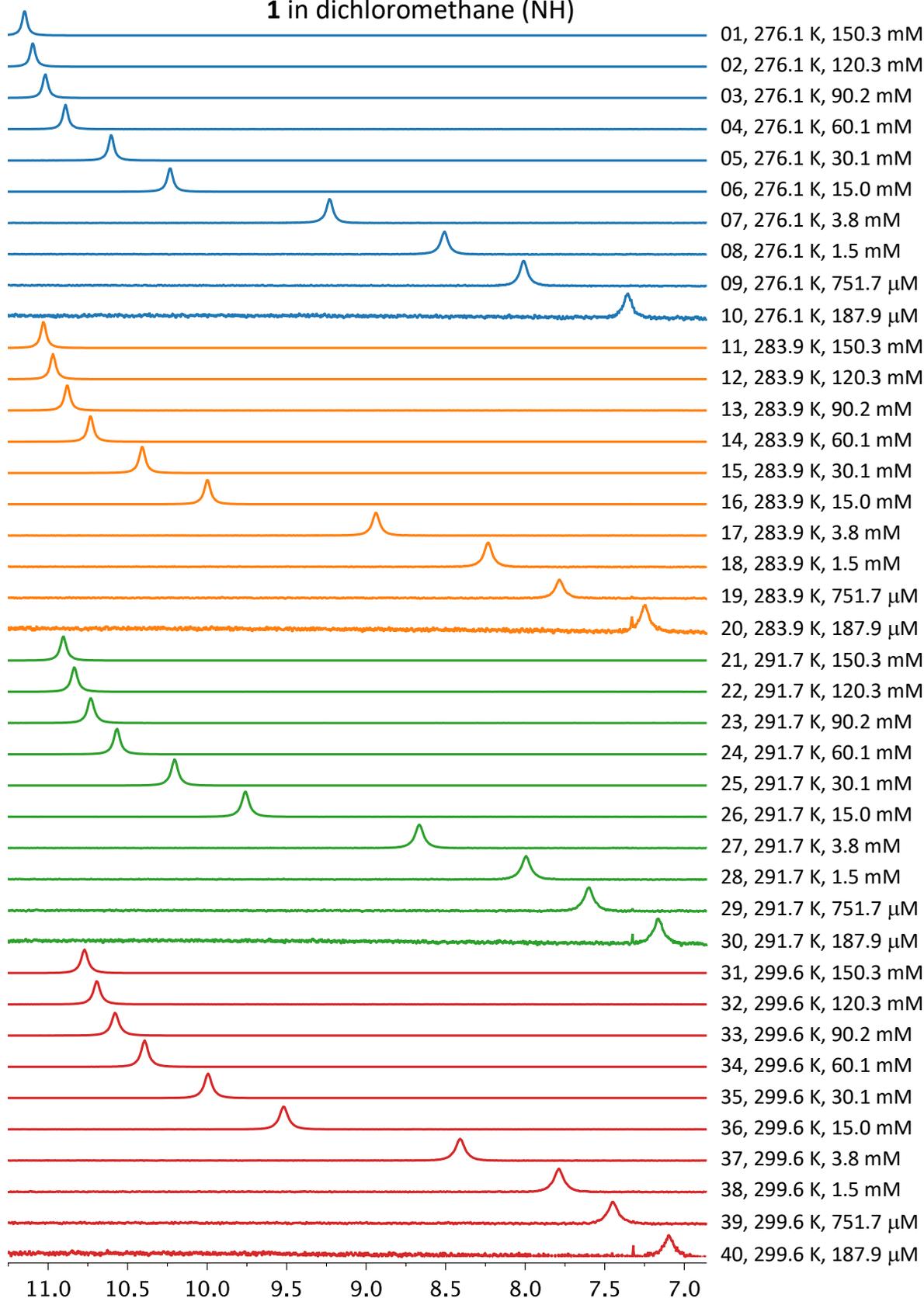


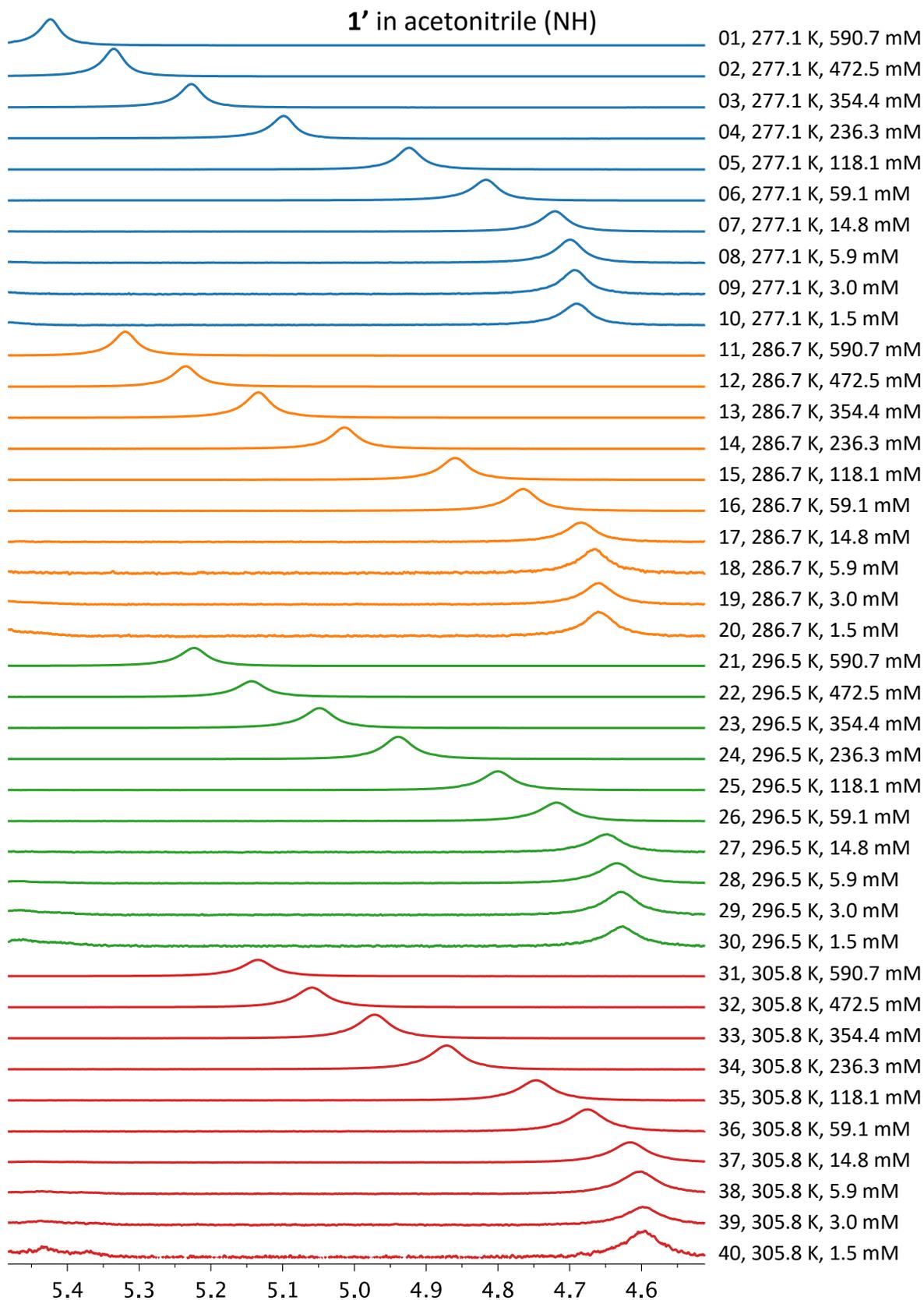


1 in chloroform (NH)

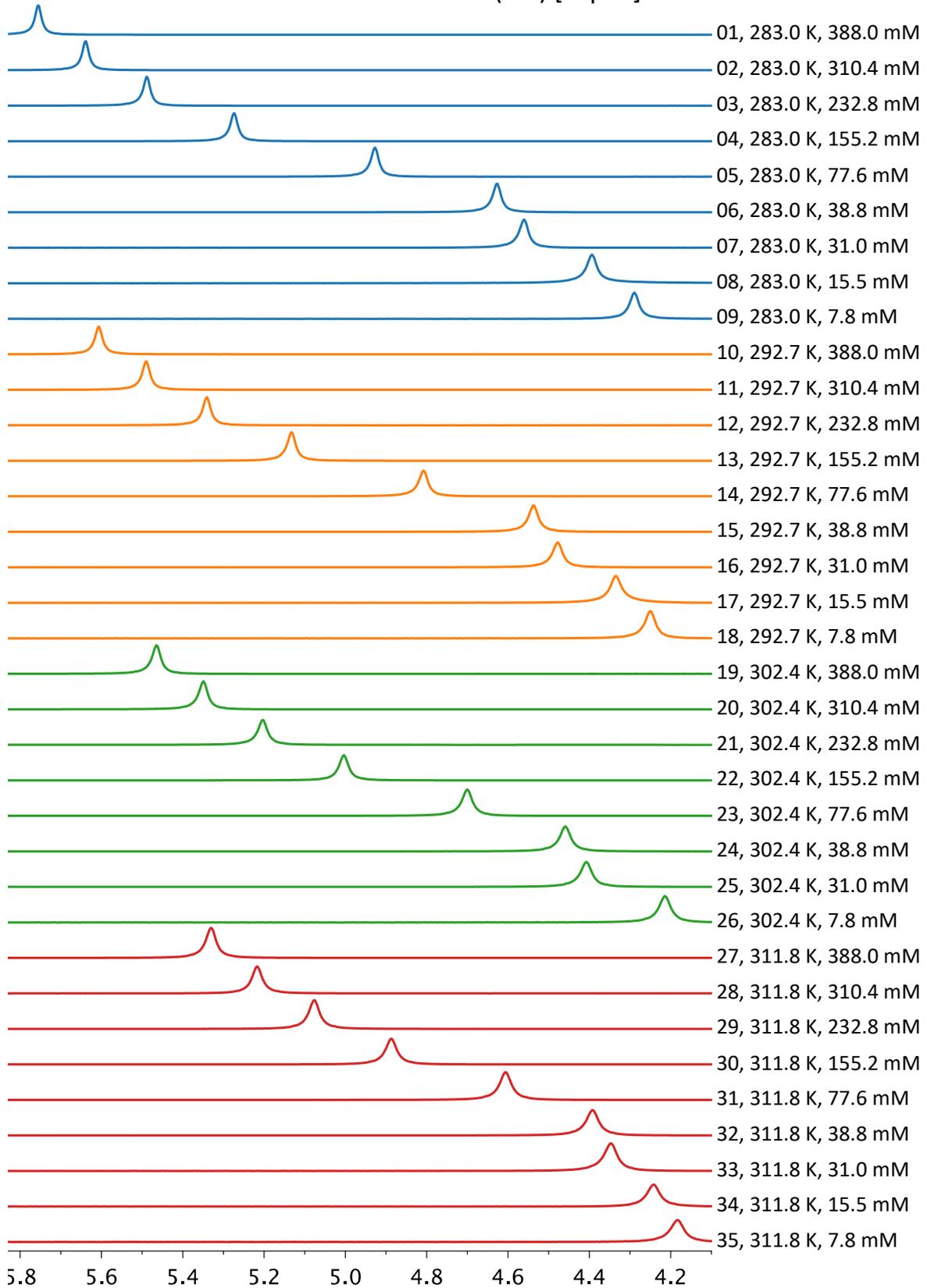


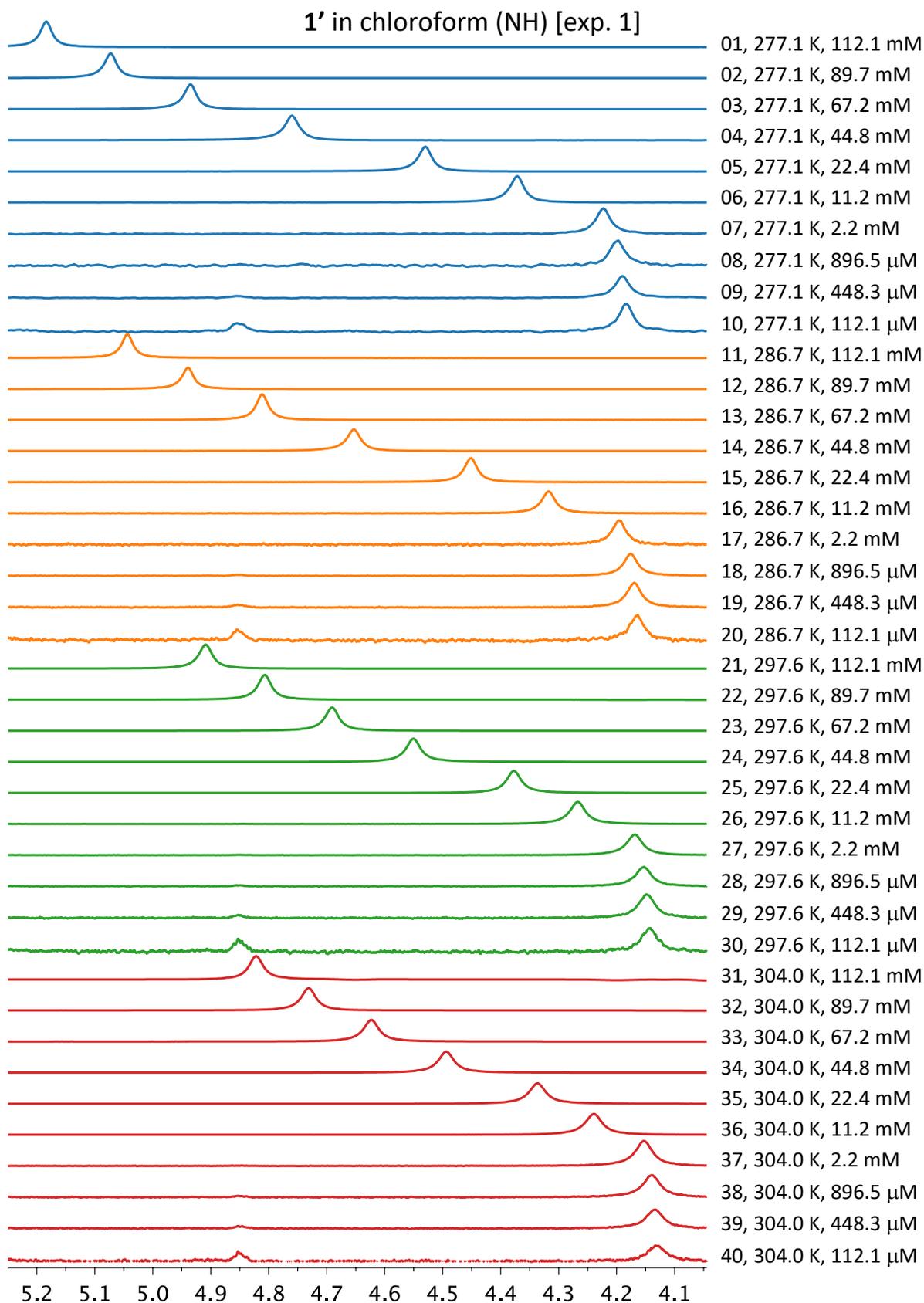
1 in dichloromethane (NH)

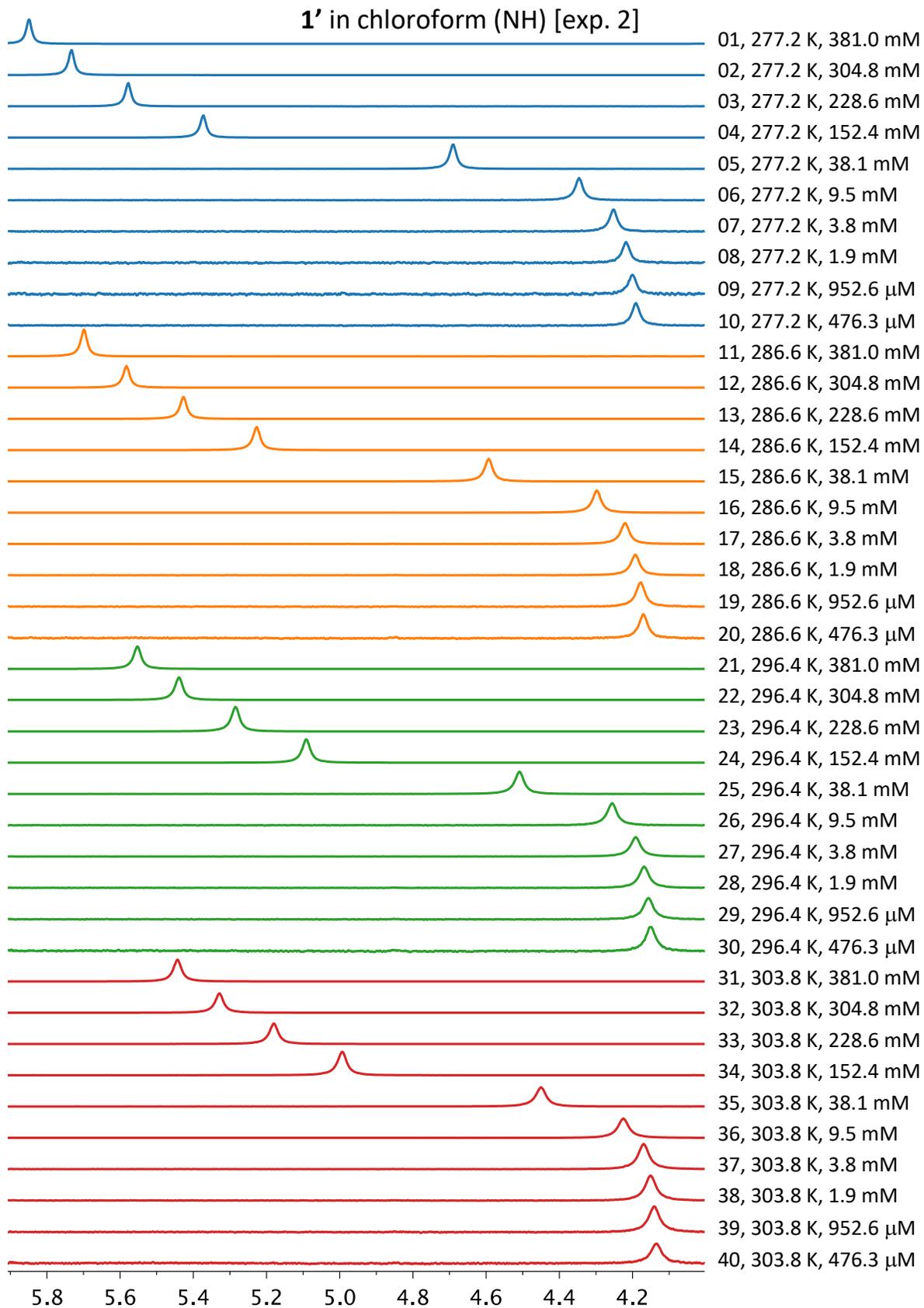




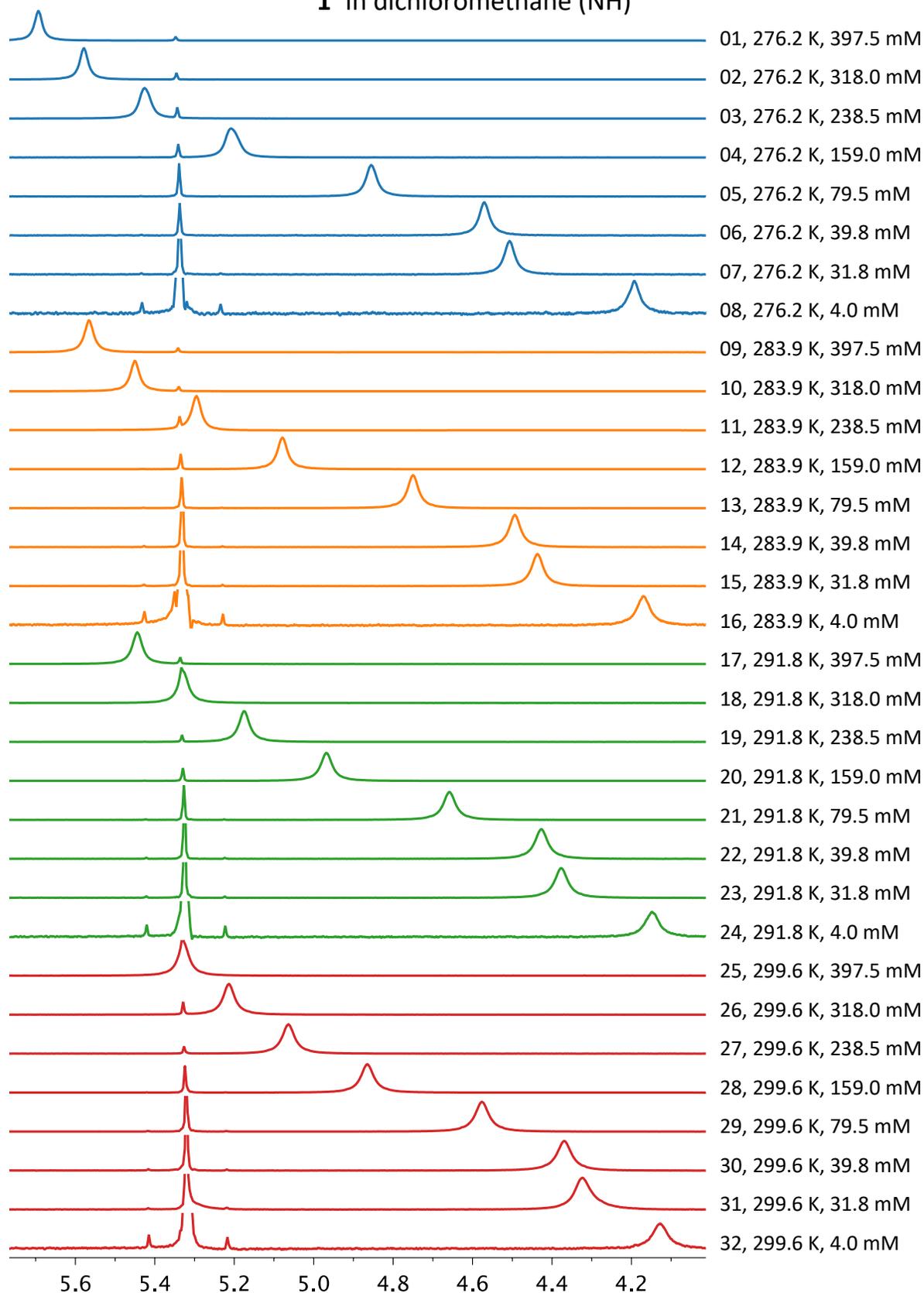
1' in chloroform (NH) [exp. 3]

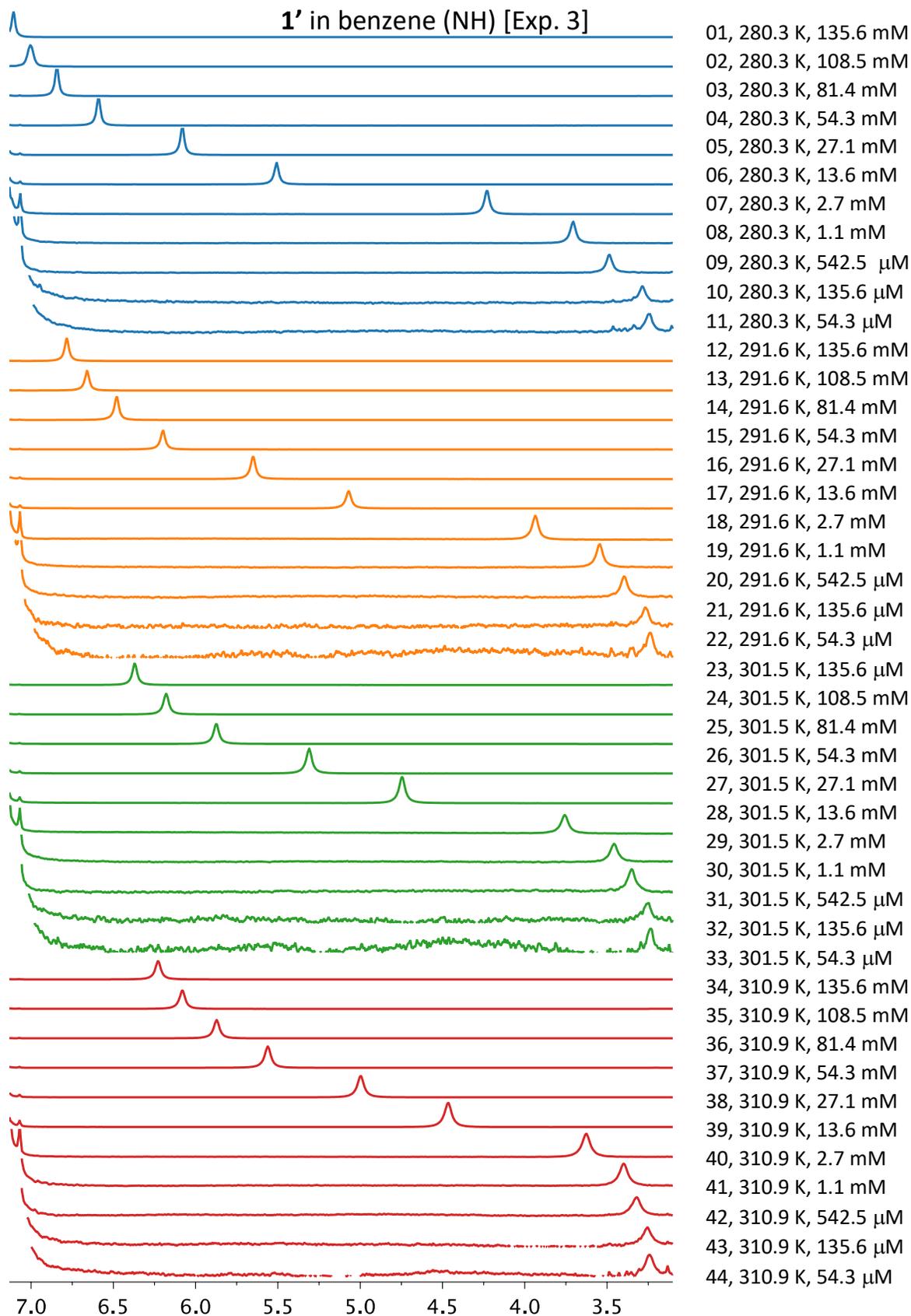


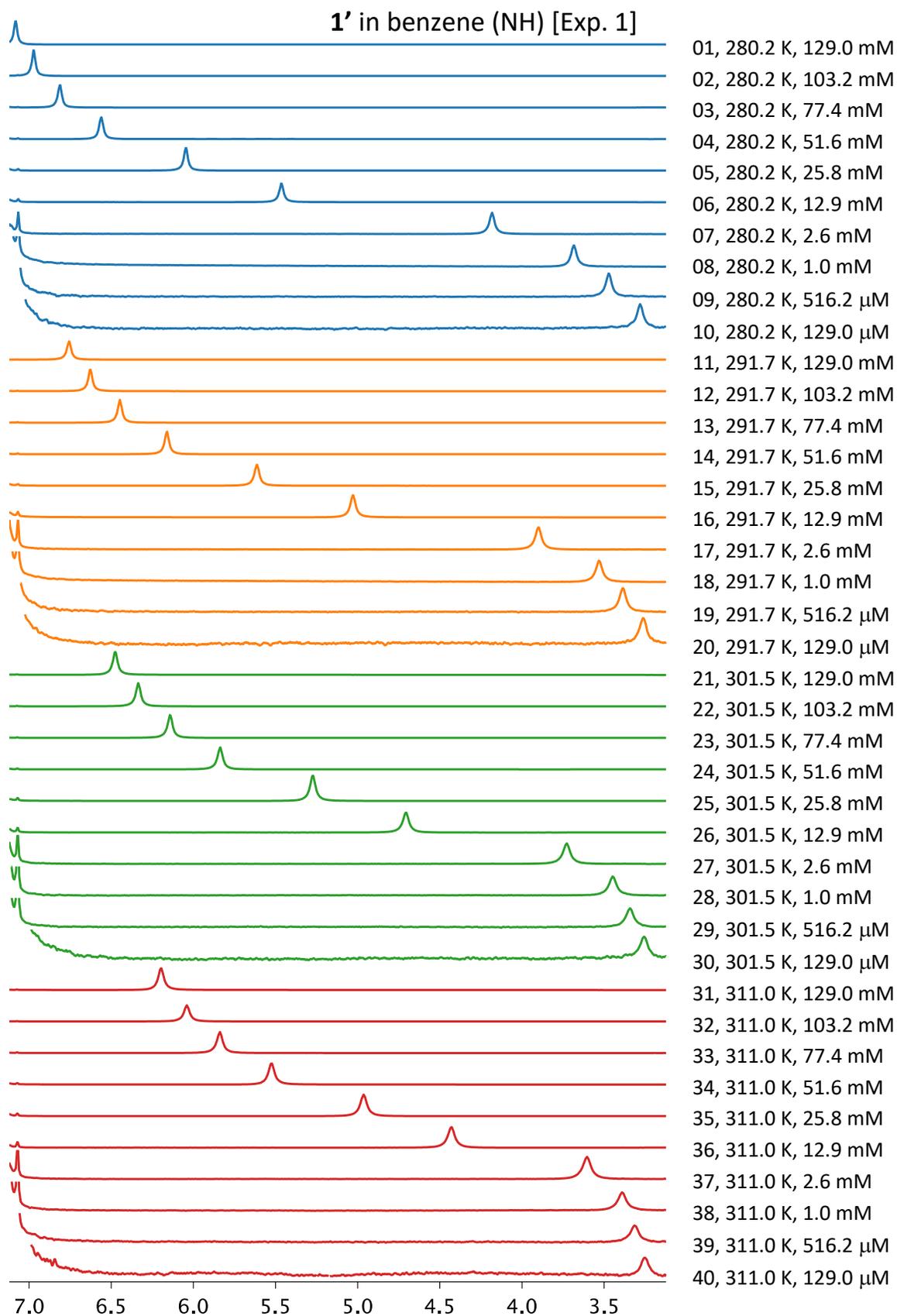


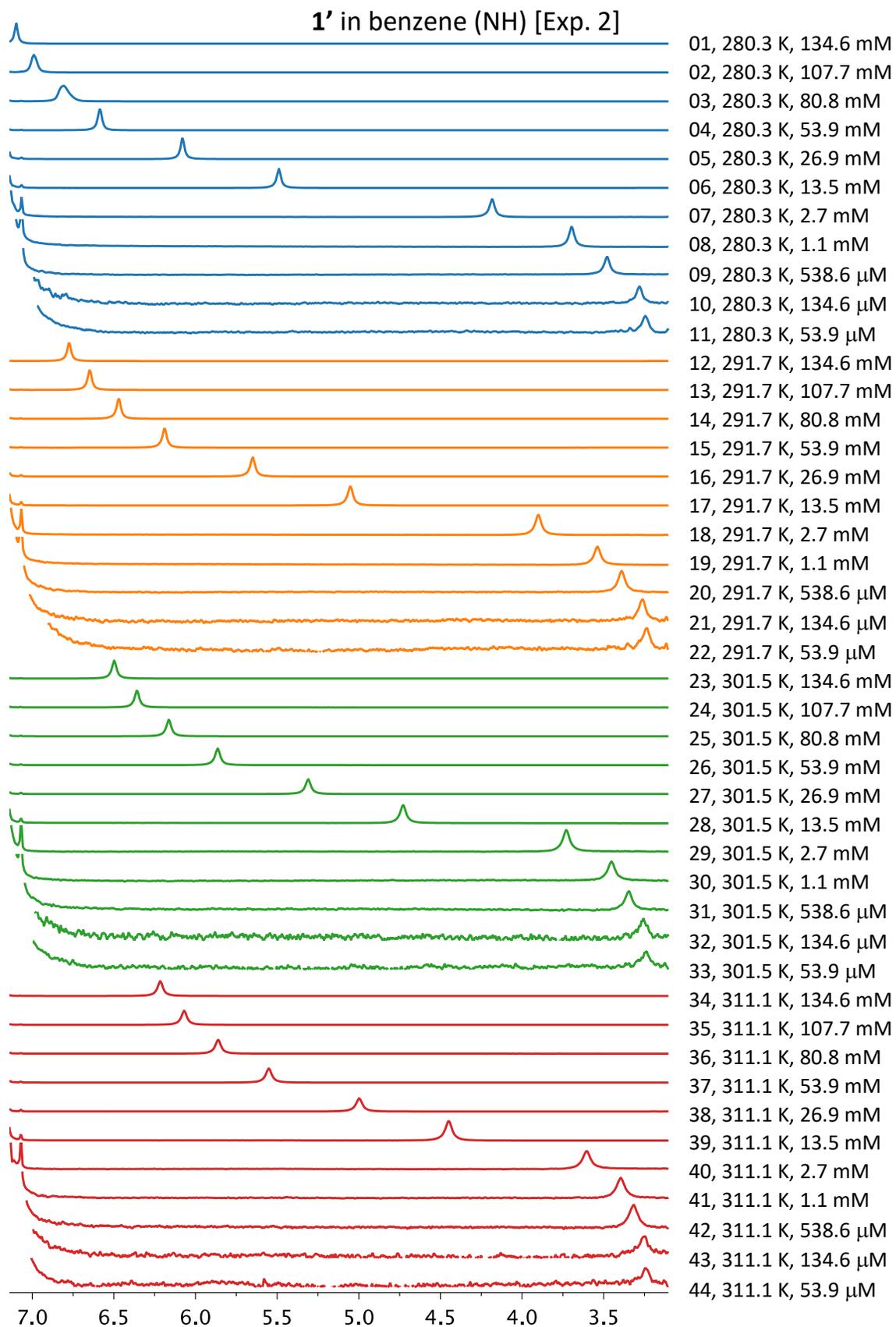


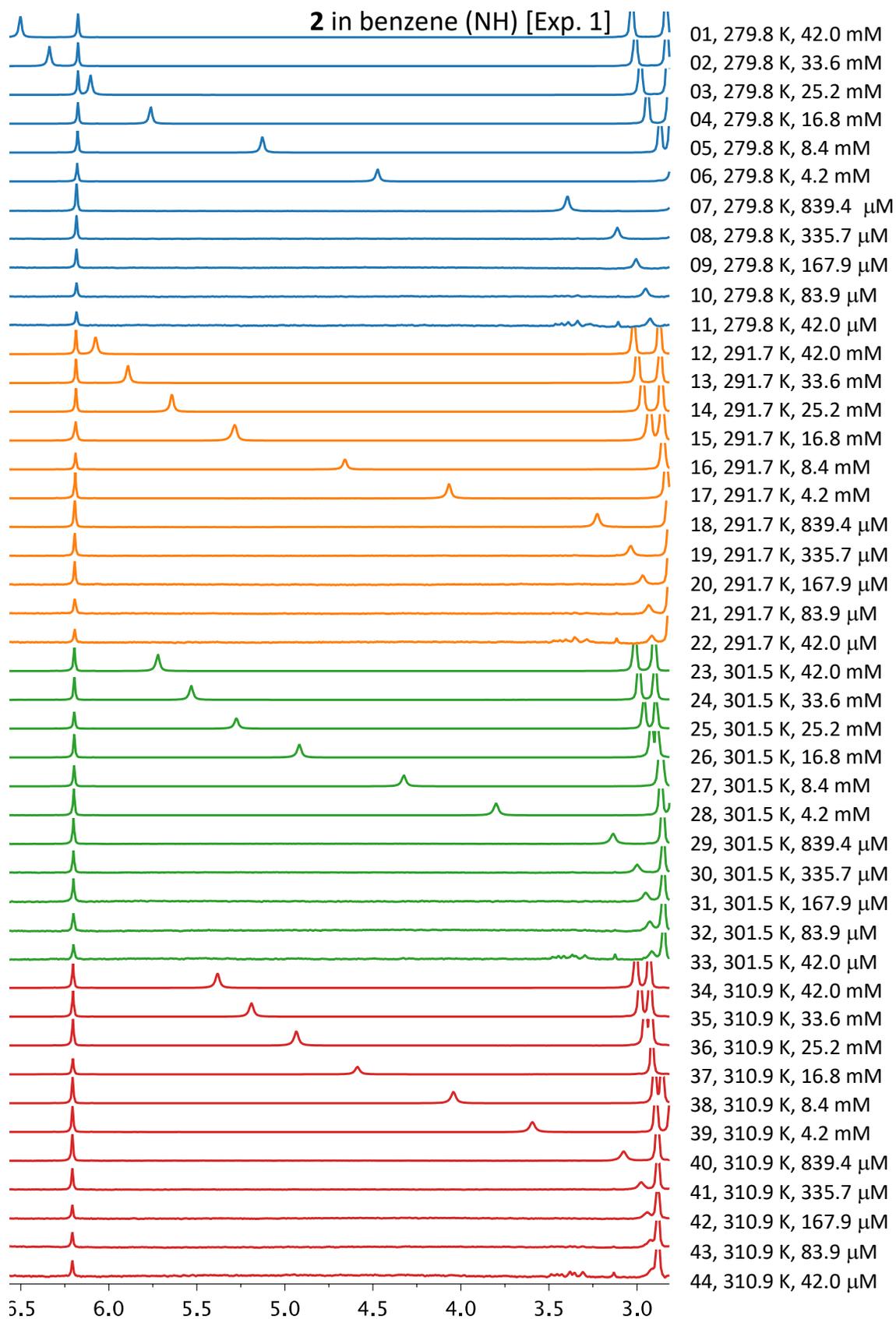
1' in dichloromethane (NH)

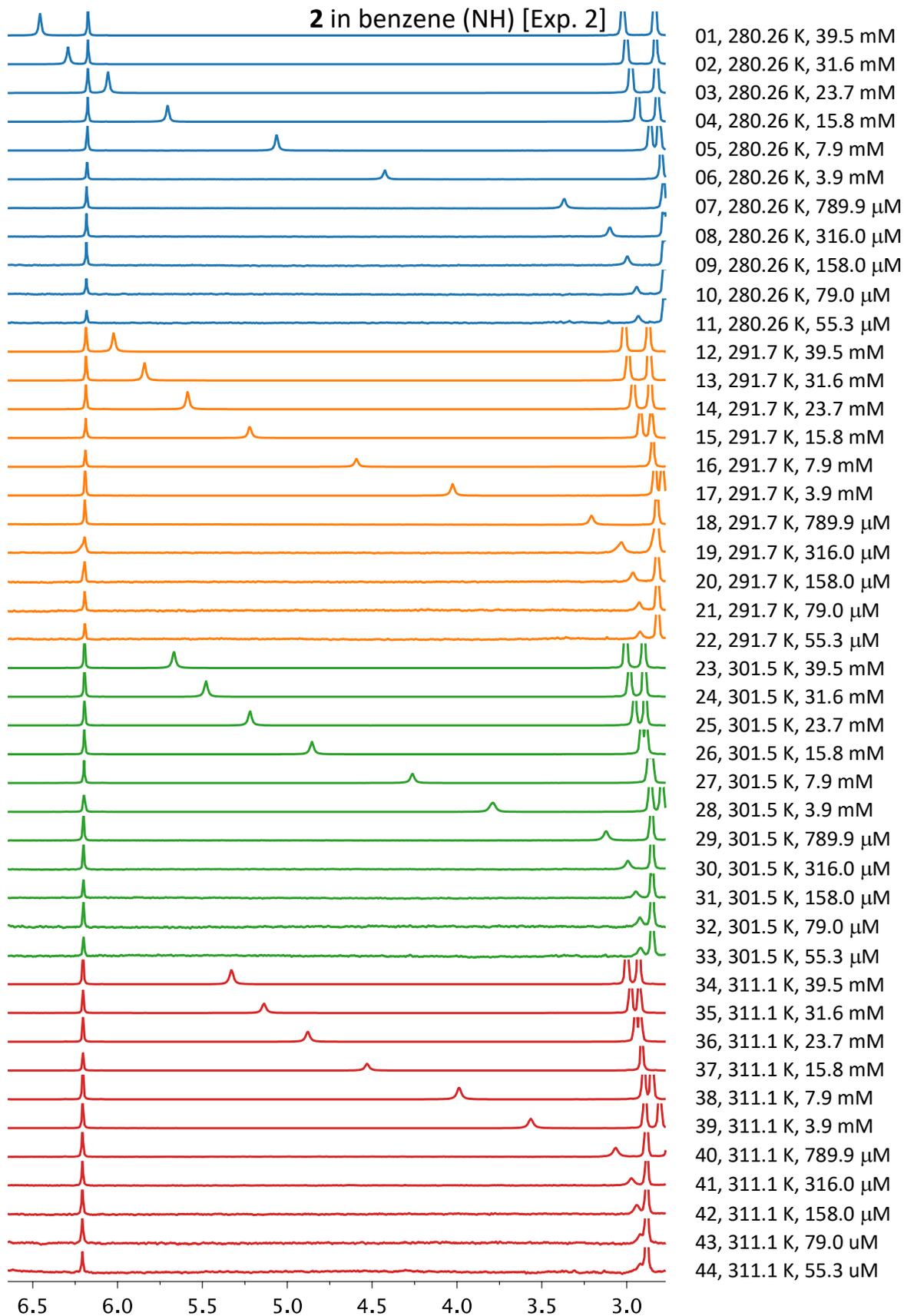


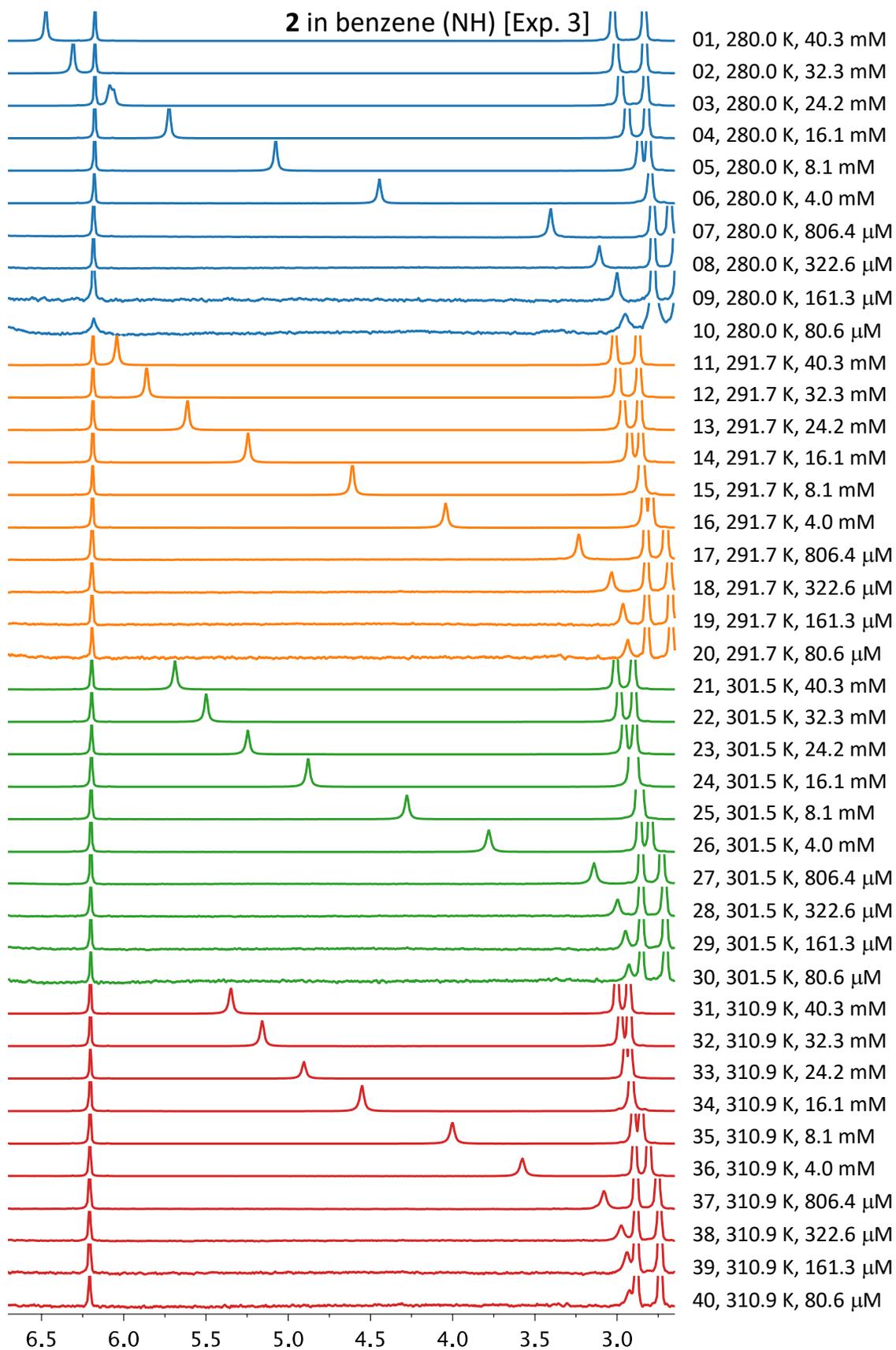




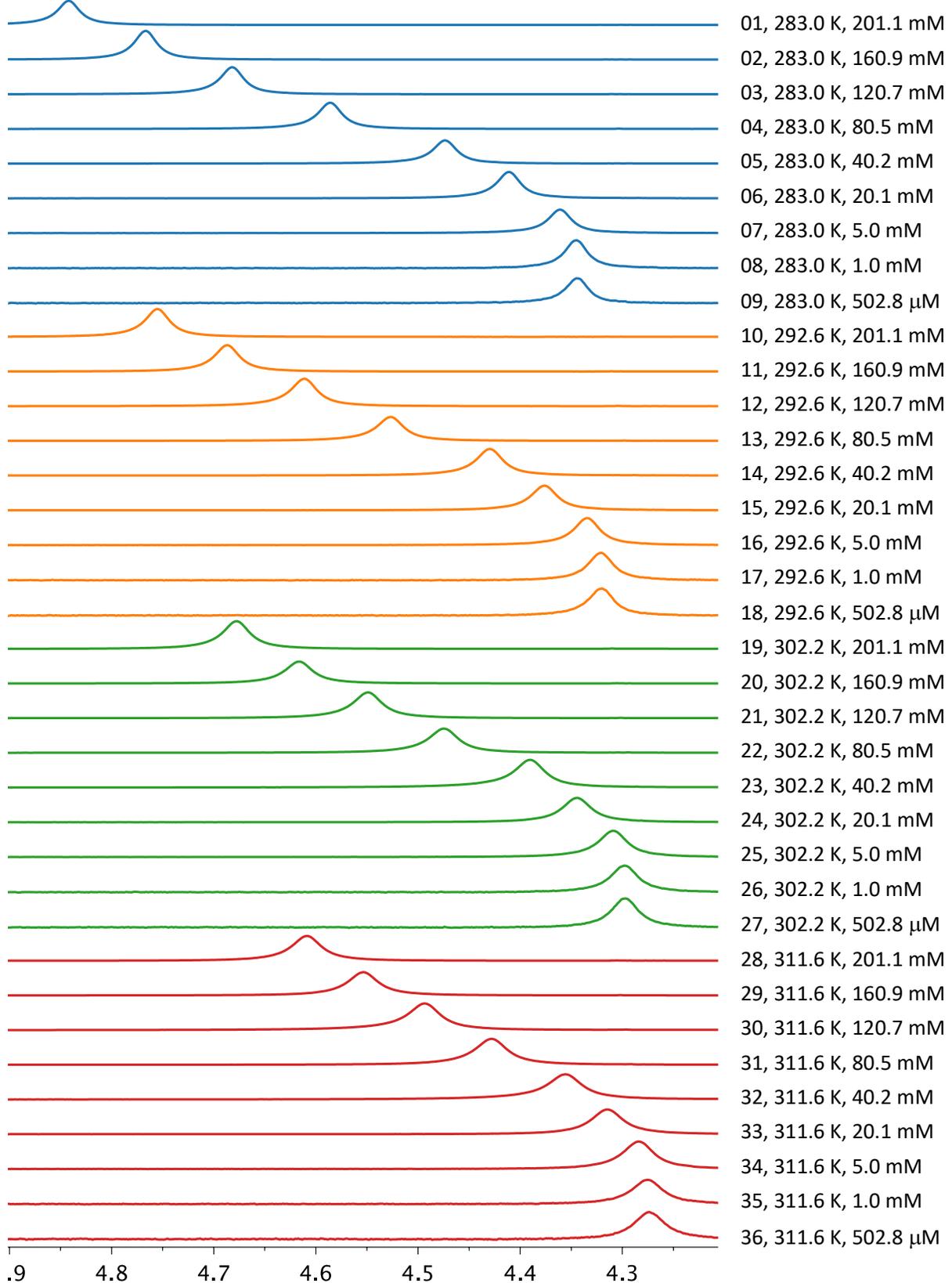


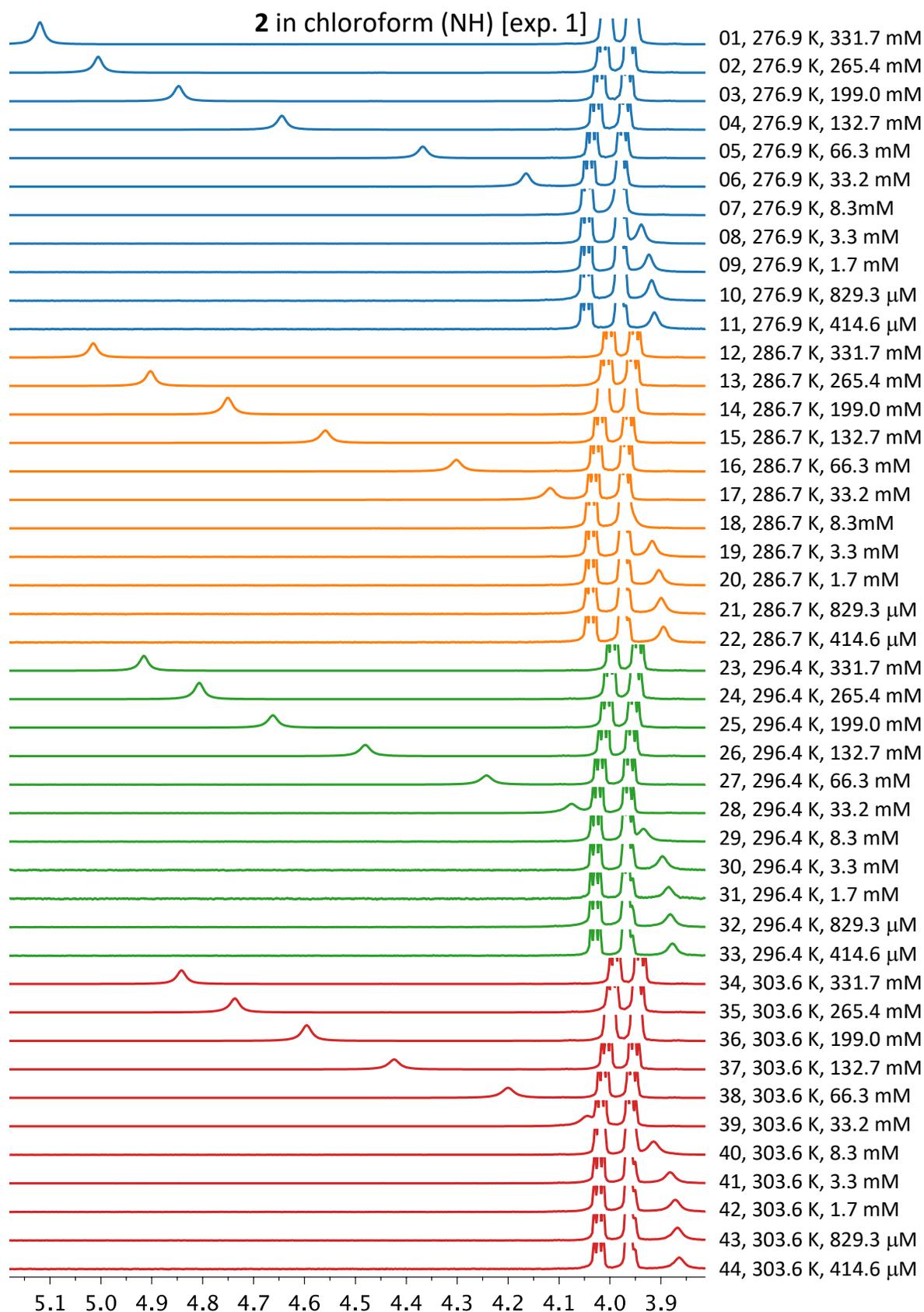


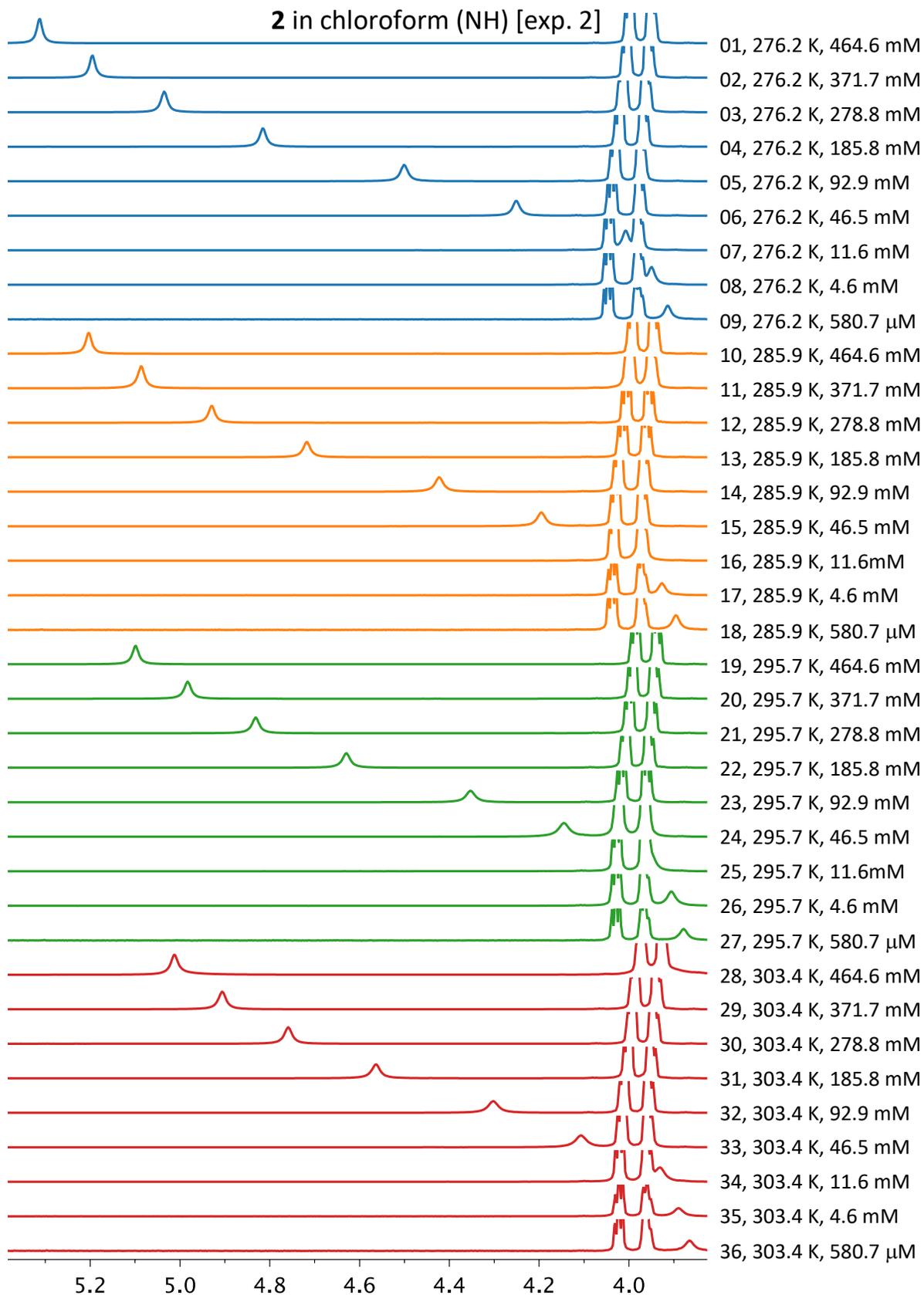


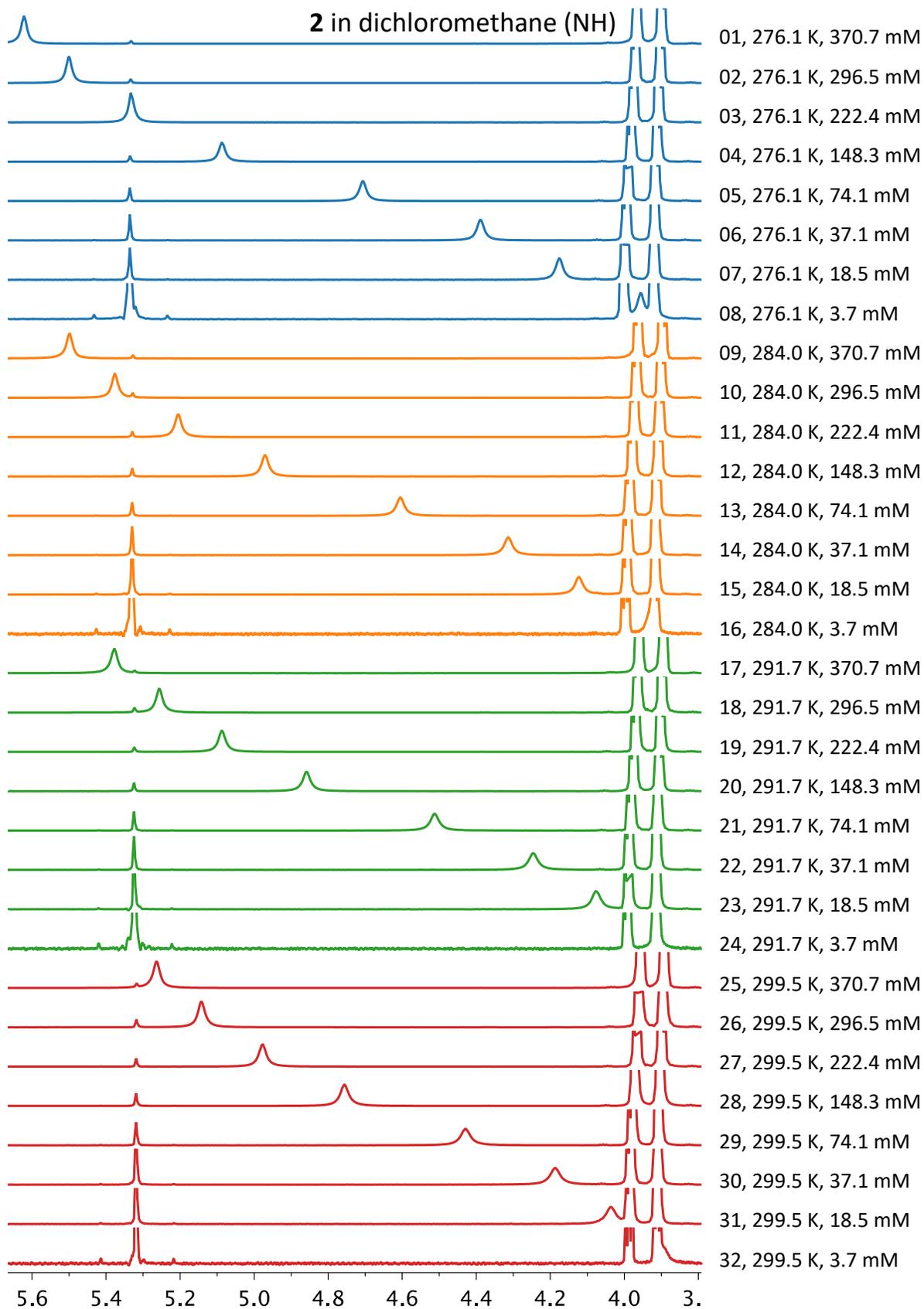


2 in acetonitrile (NH)

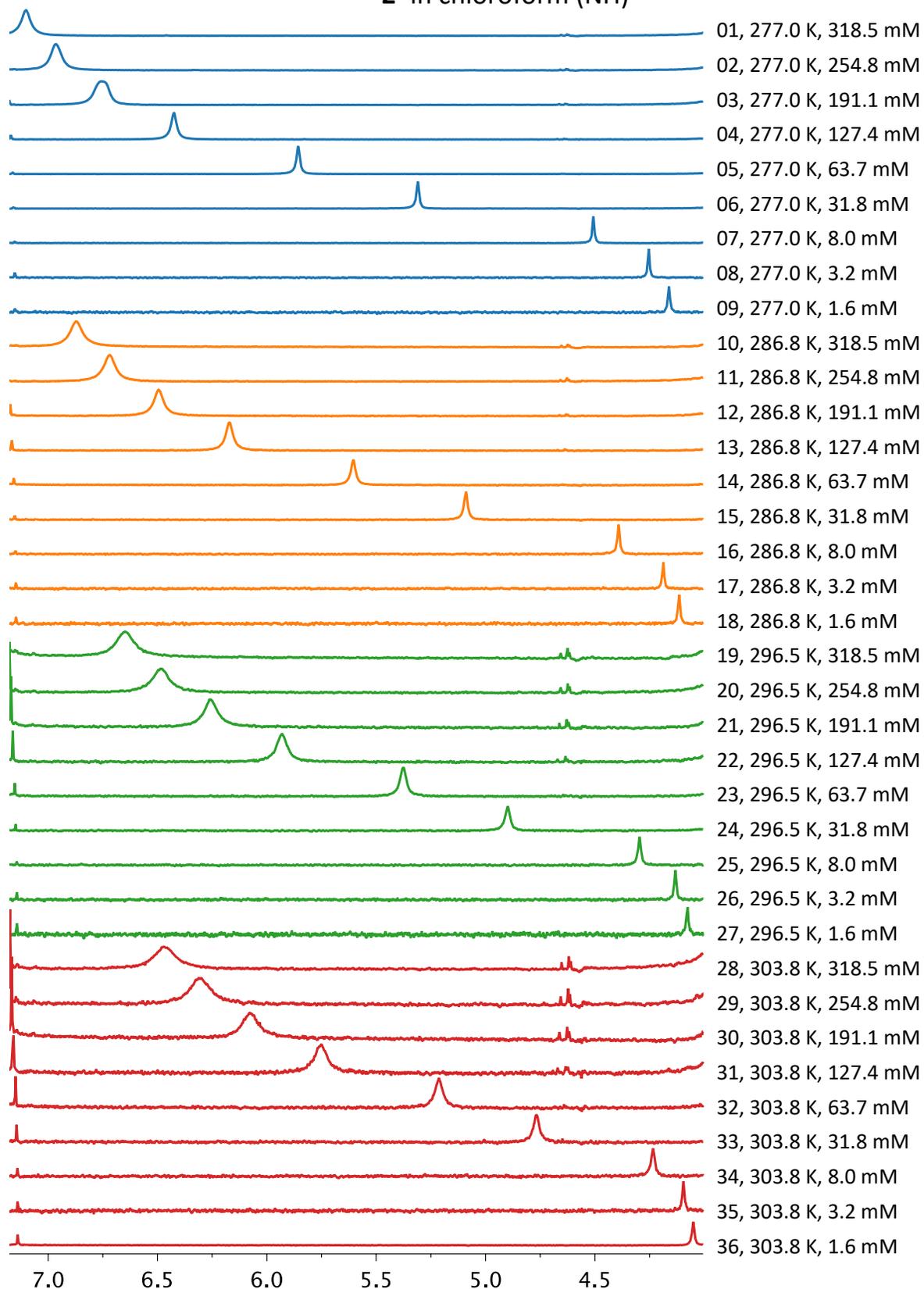


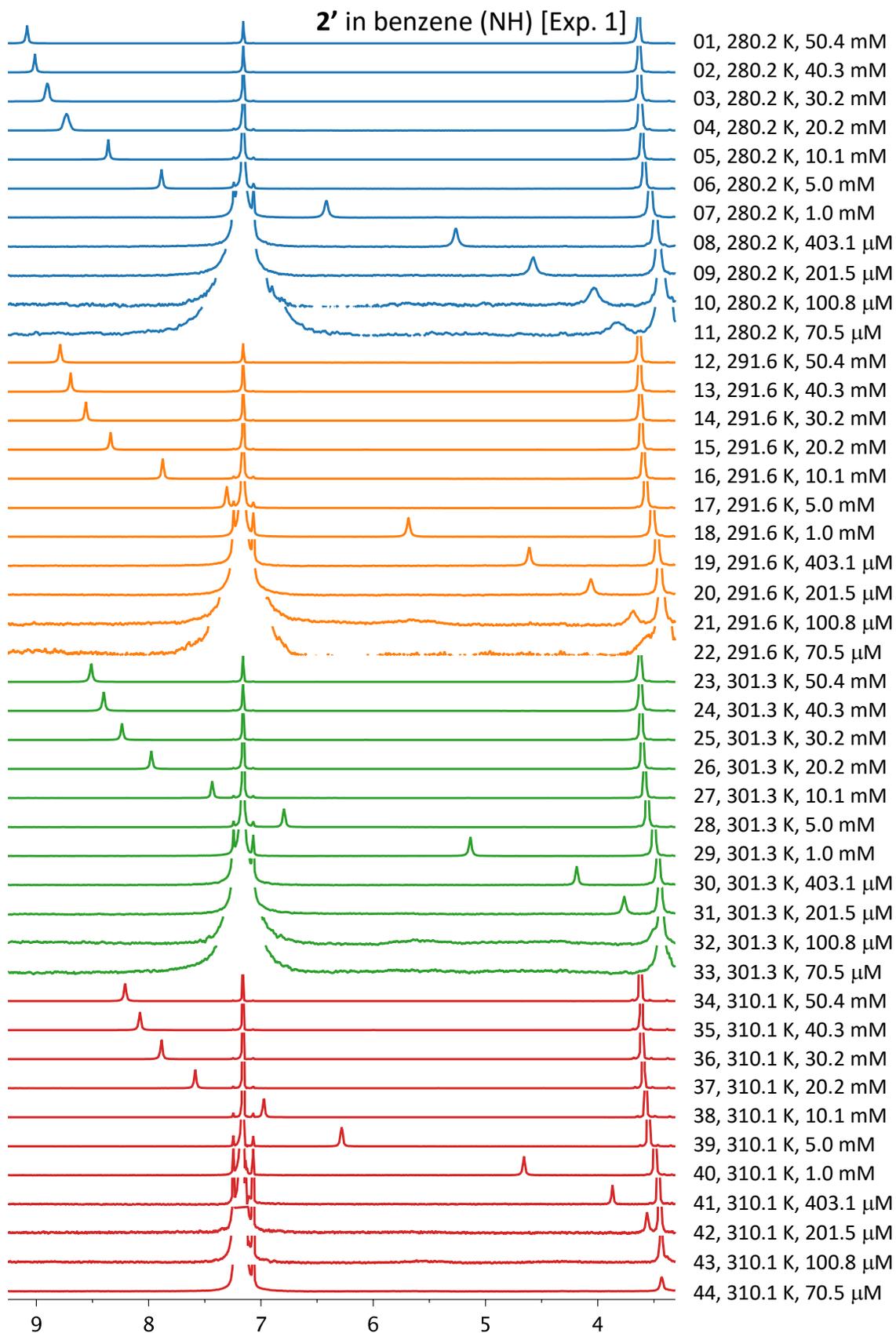


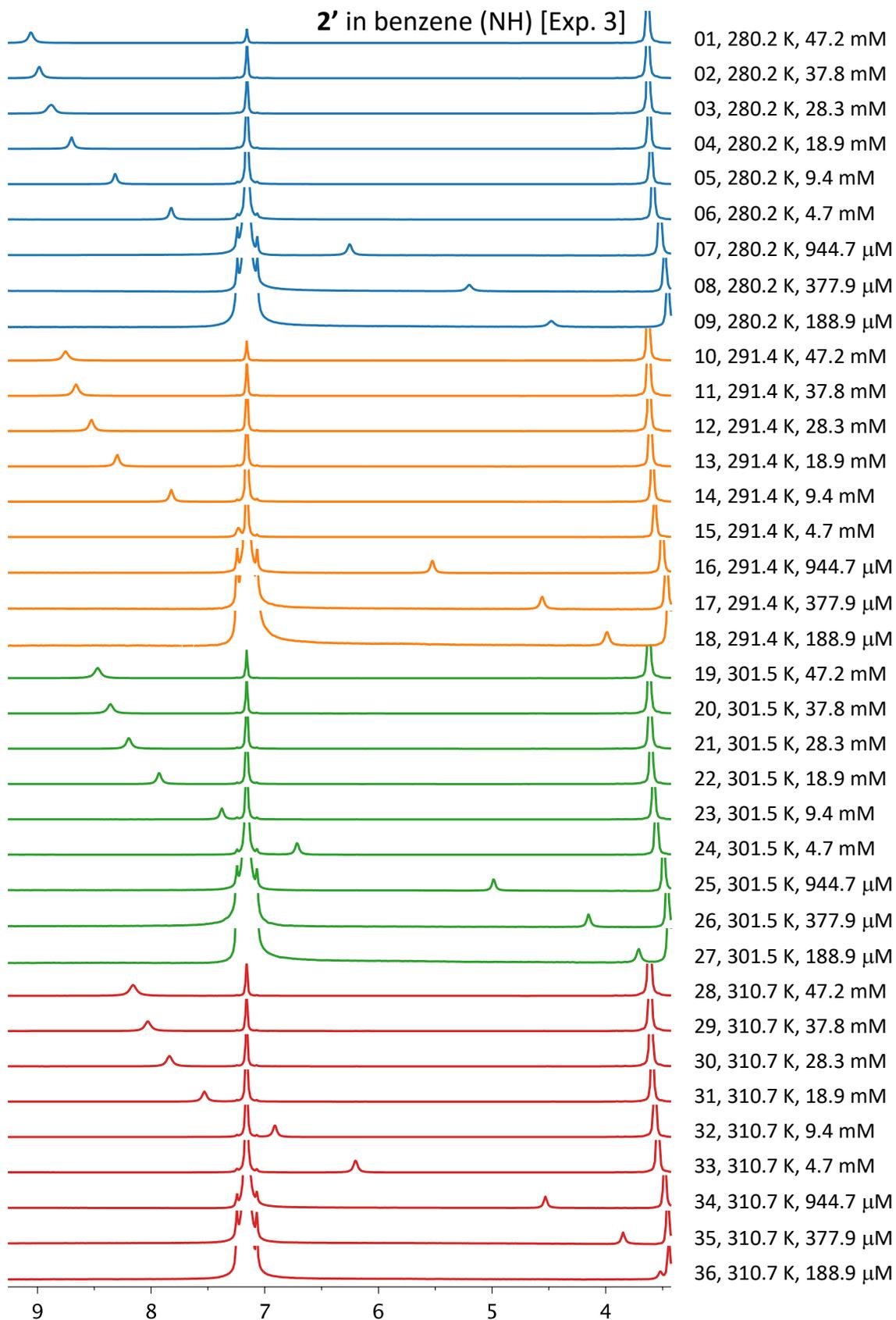




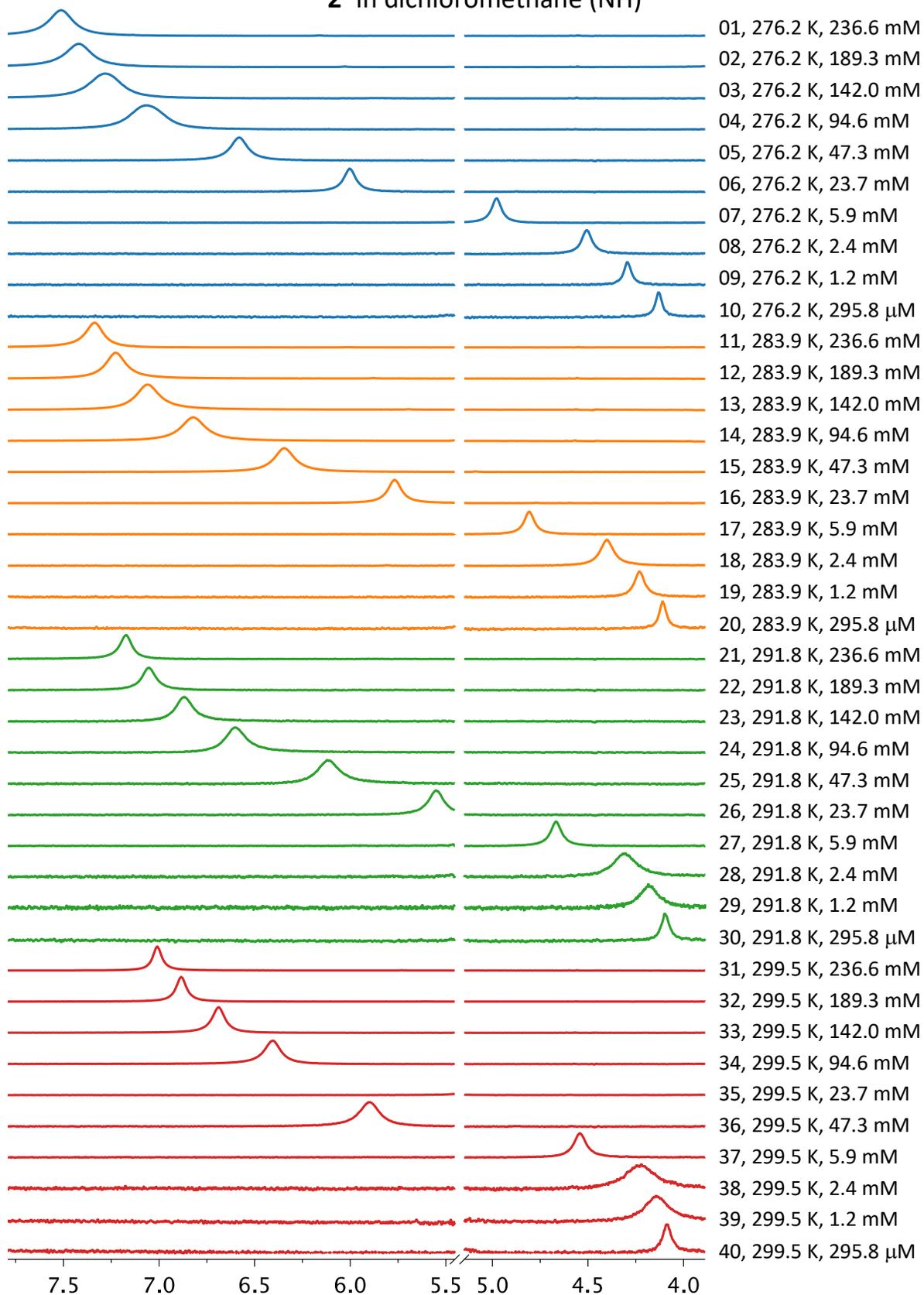
2' in chloroform (NH)



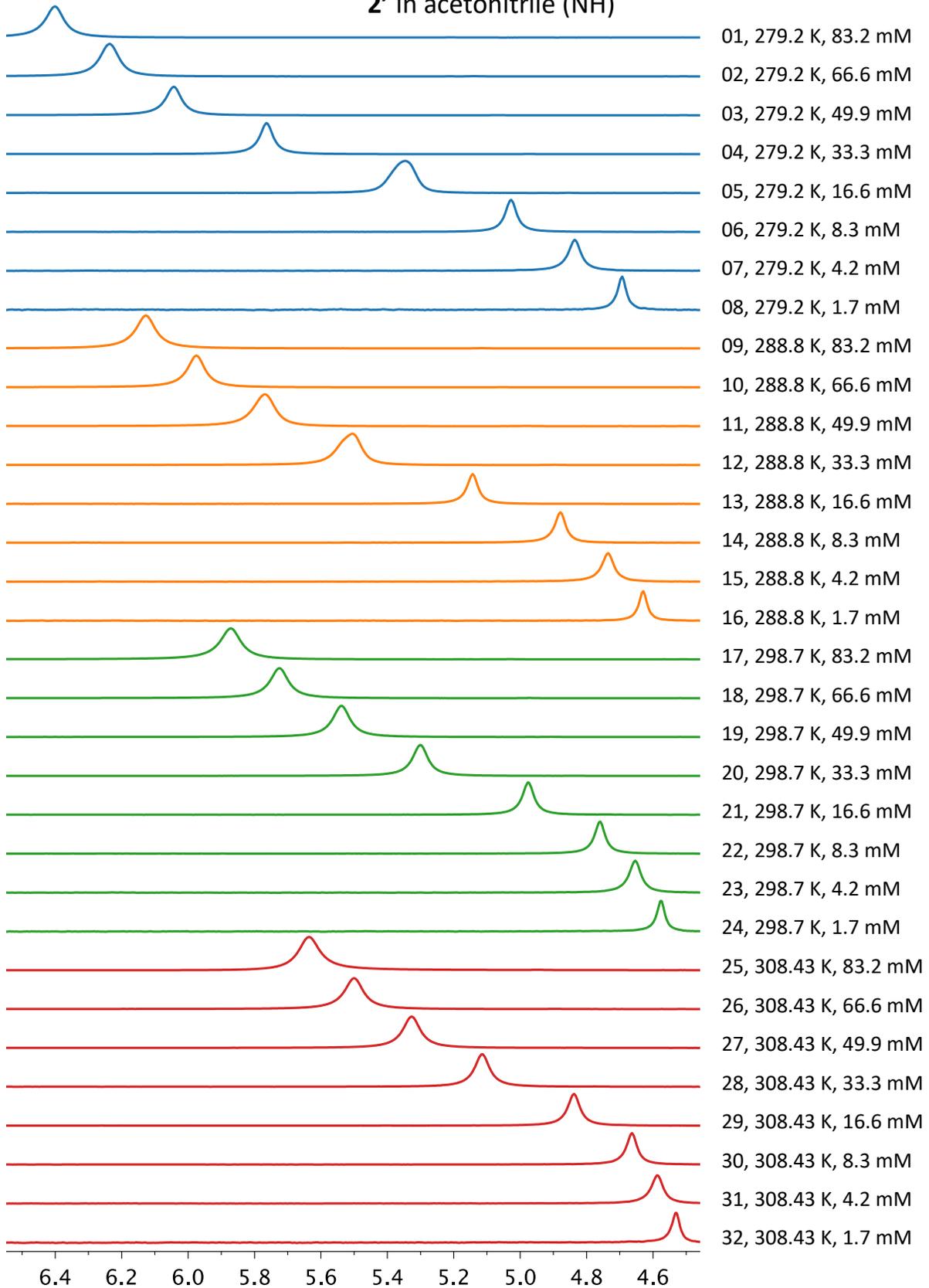


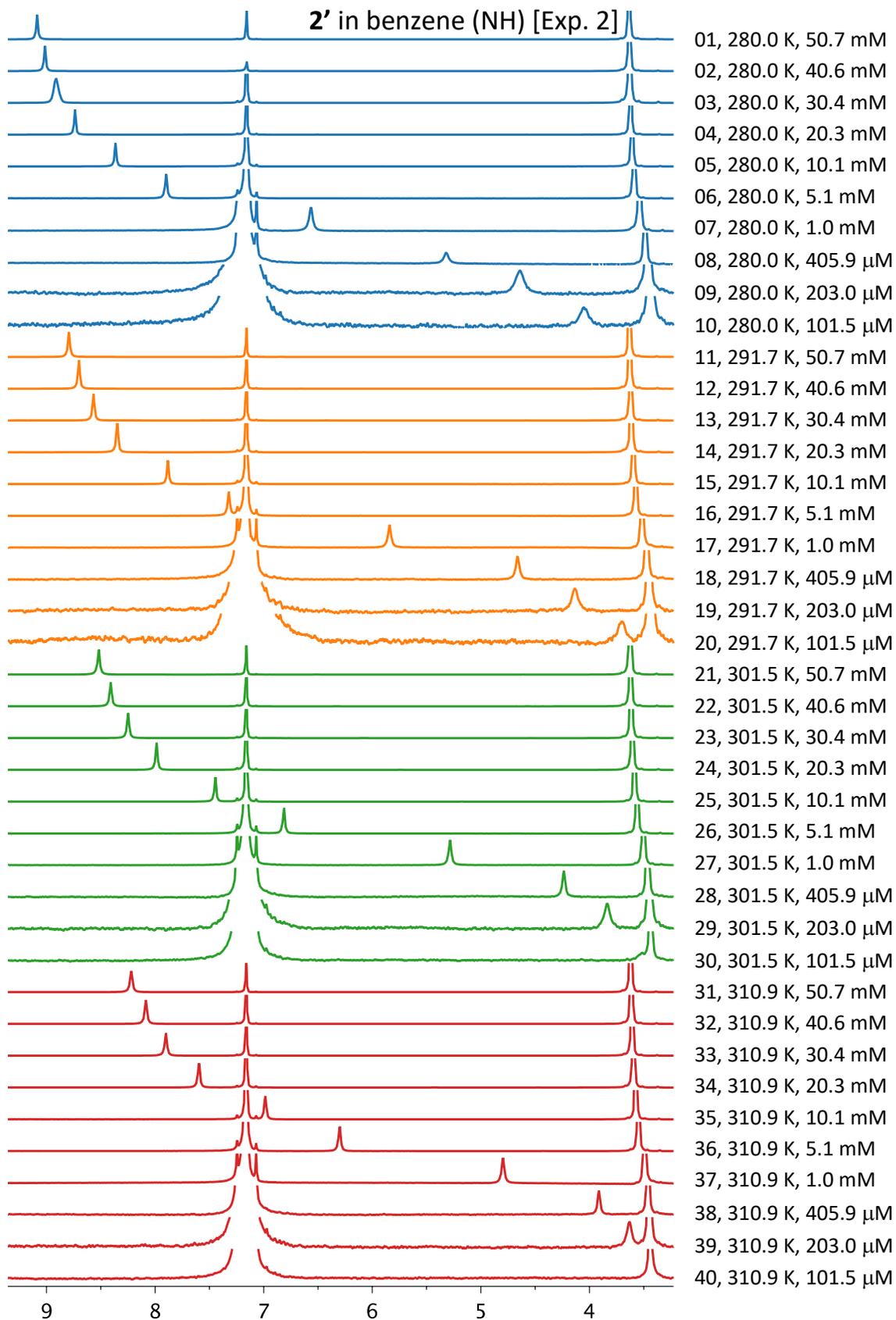


2' in dichloromethane (NH)

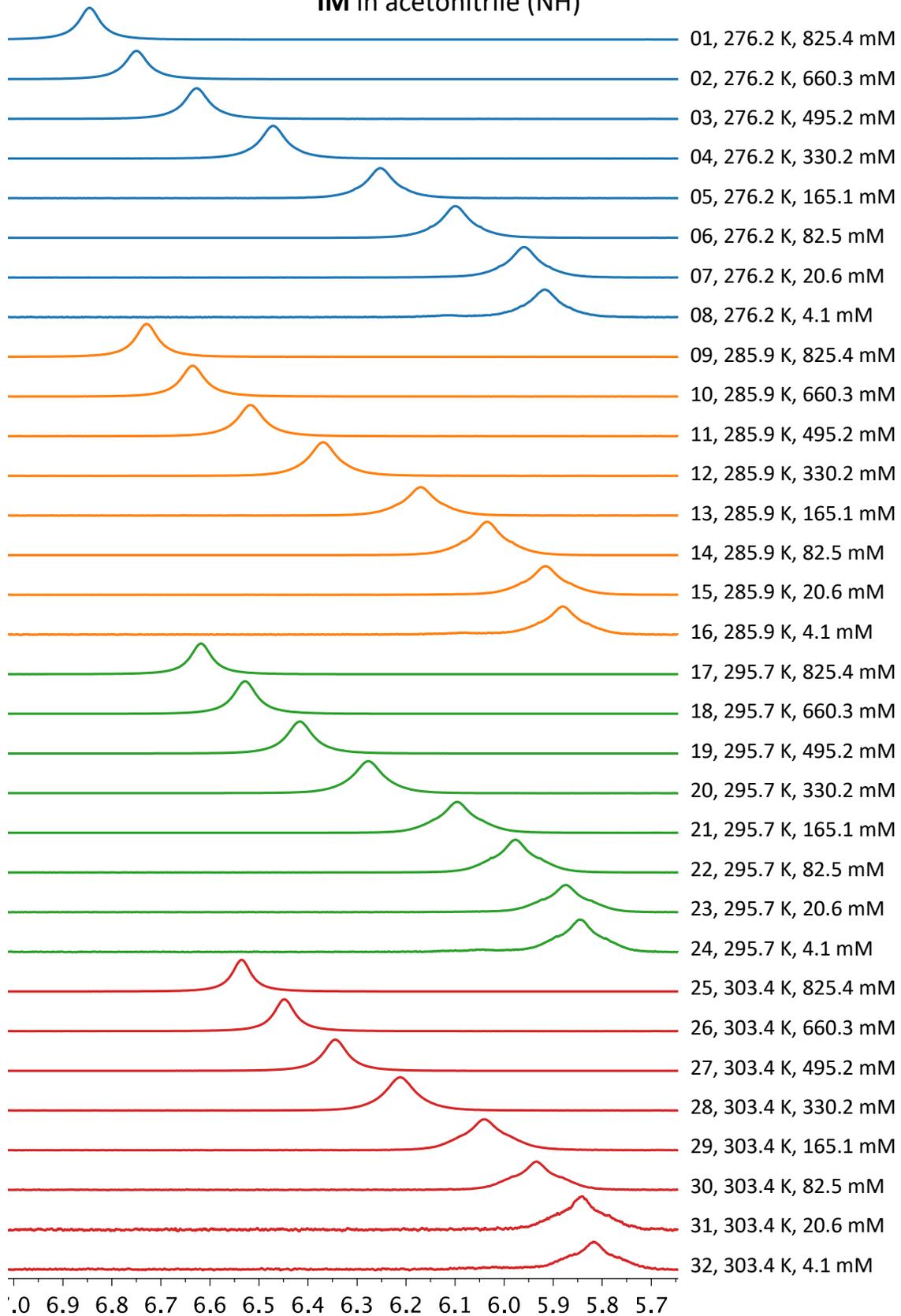


2' in acetonitrile (NH)

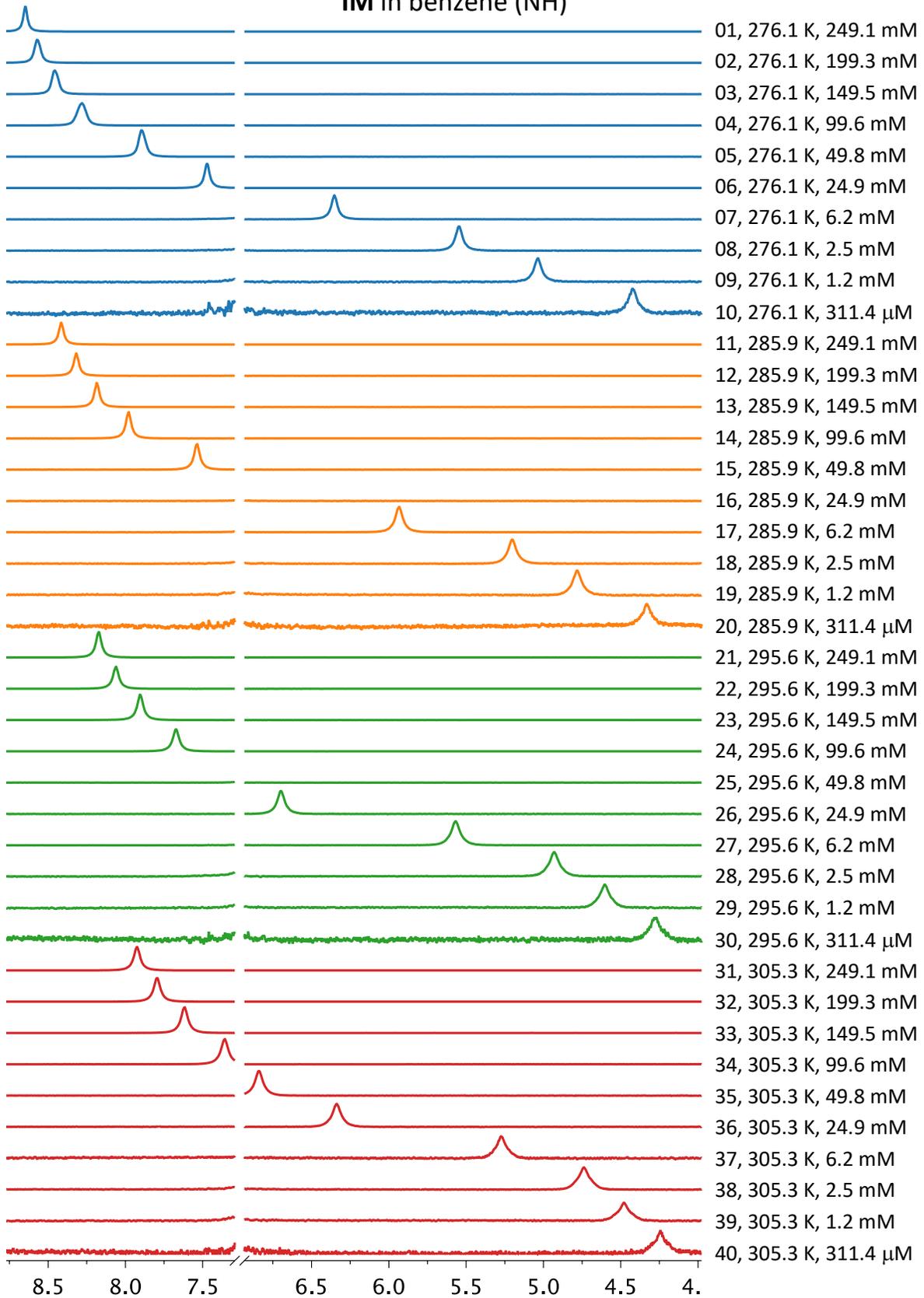




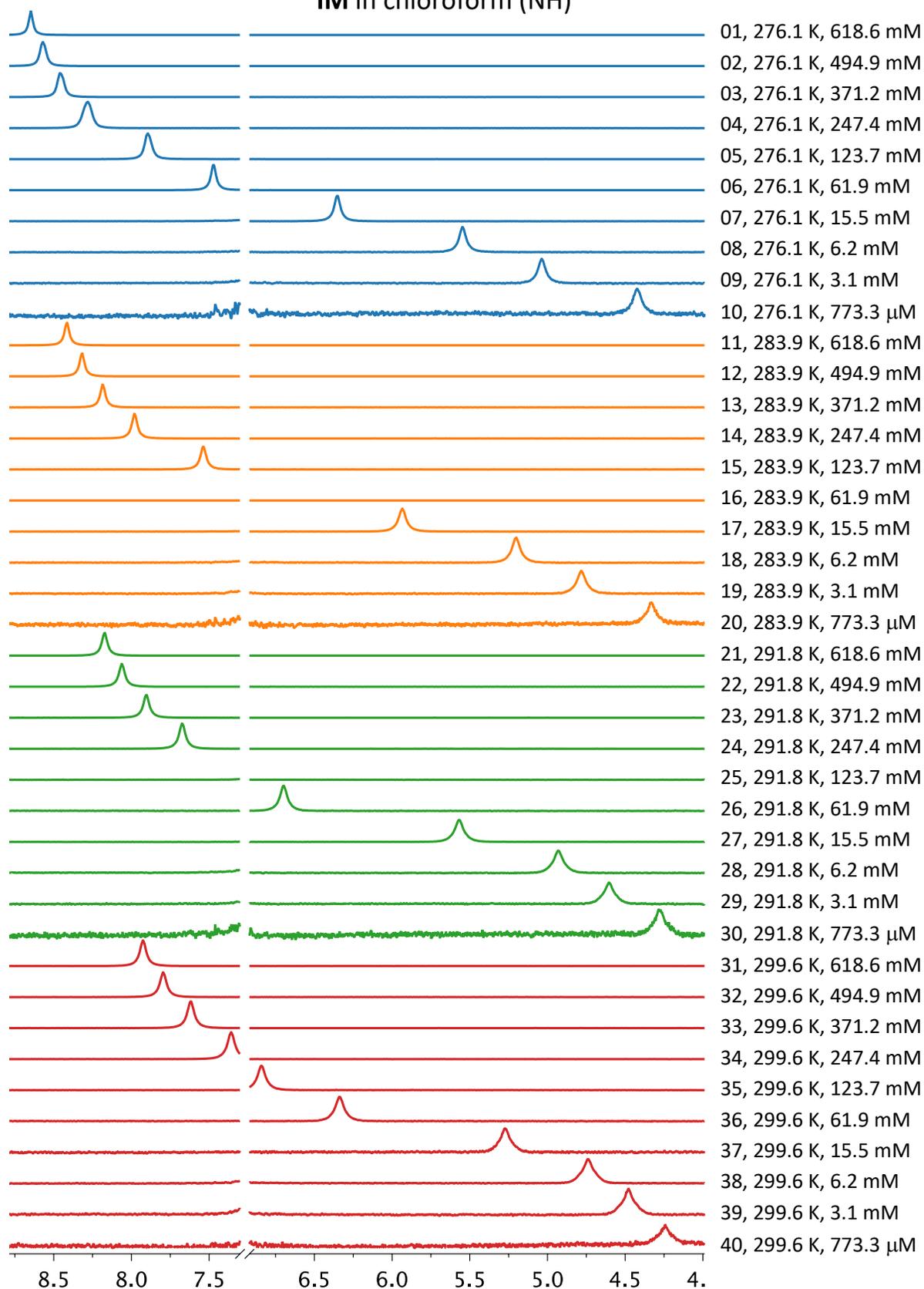
IM in acetonitrile (NH)



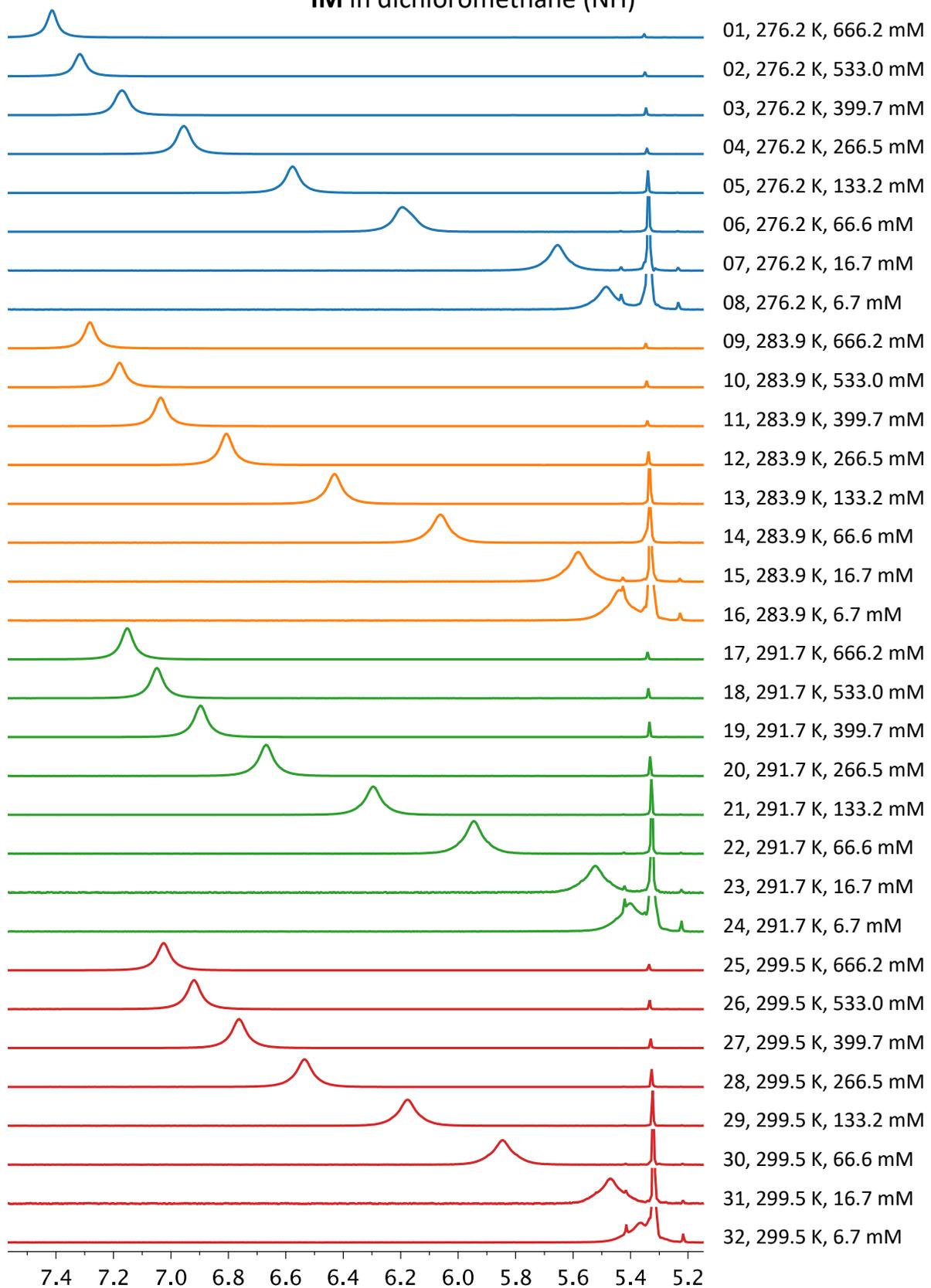
IM in benzene (NH)



IM in chloroform (NH)



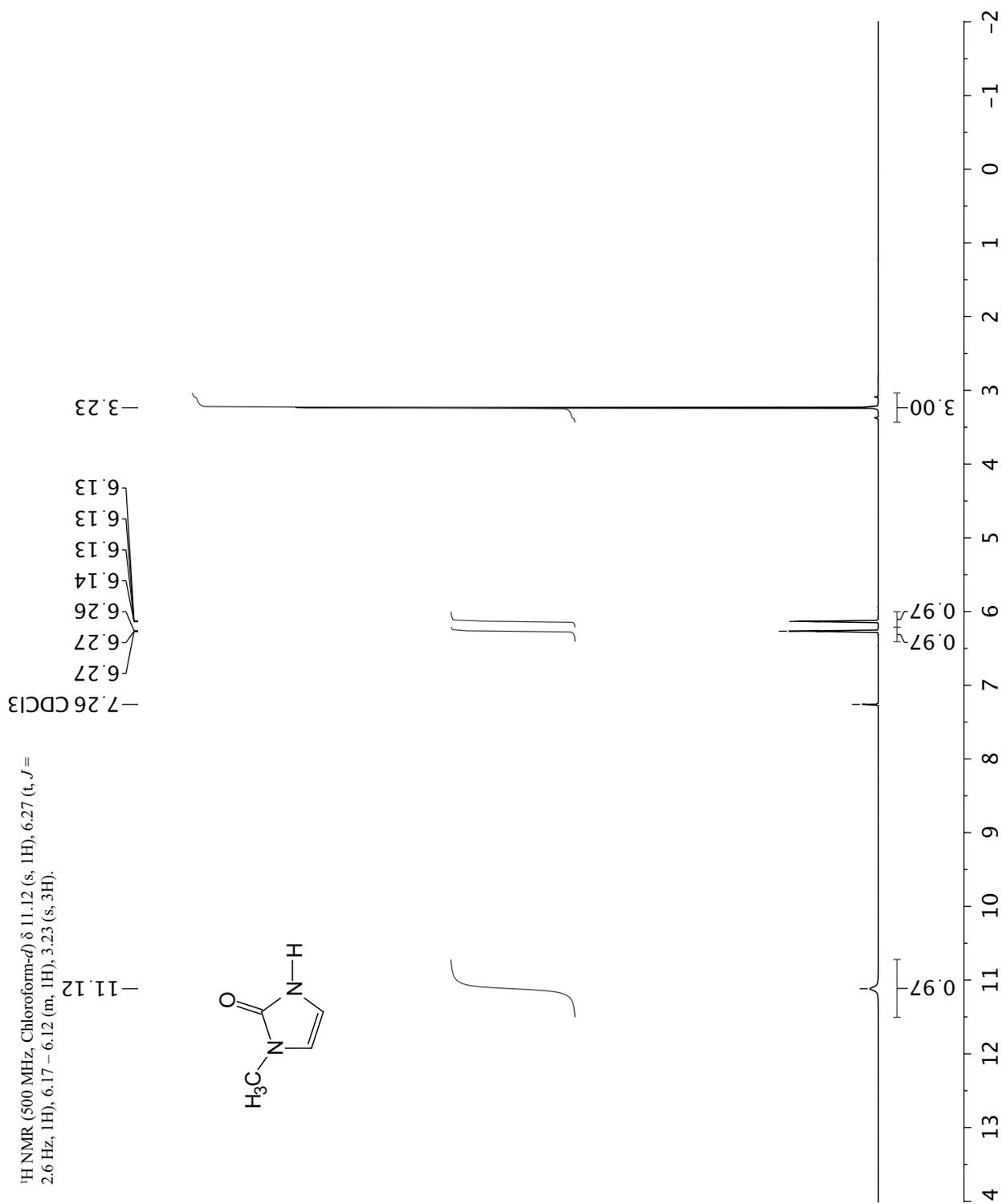
IM in dichloromethane (NH)



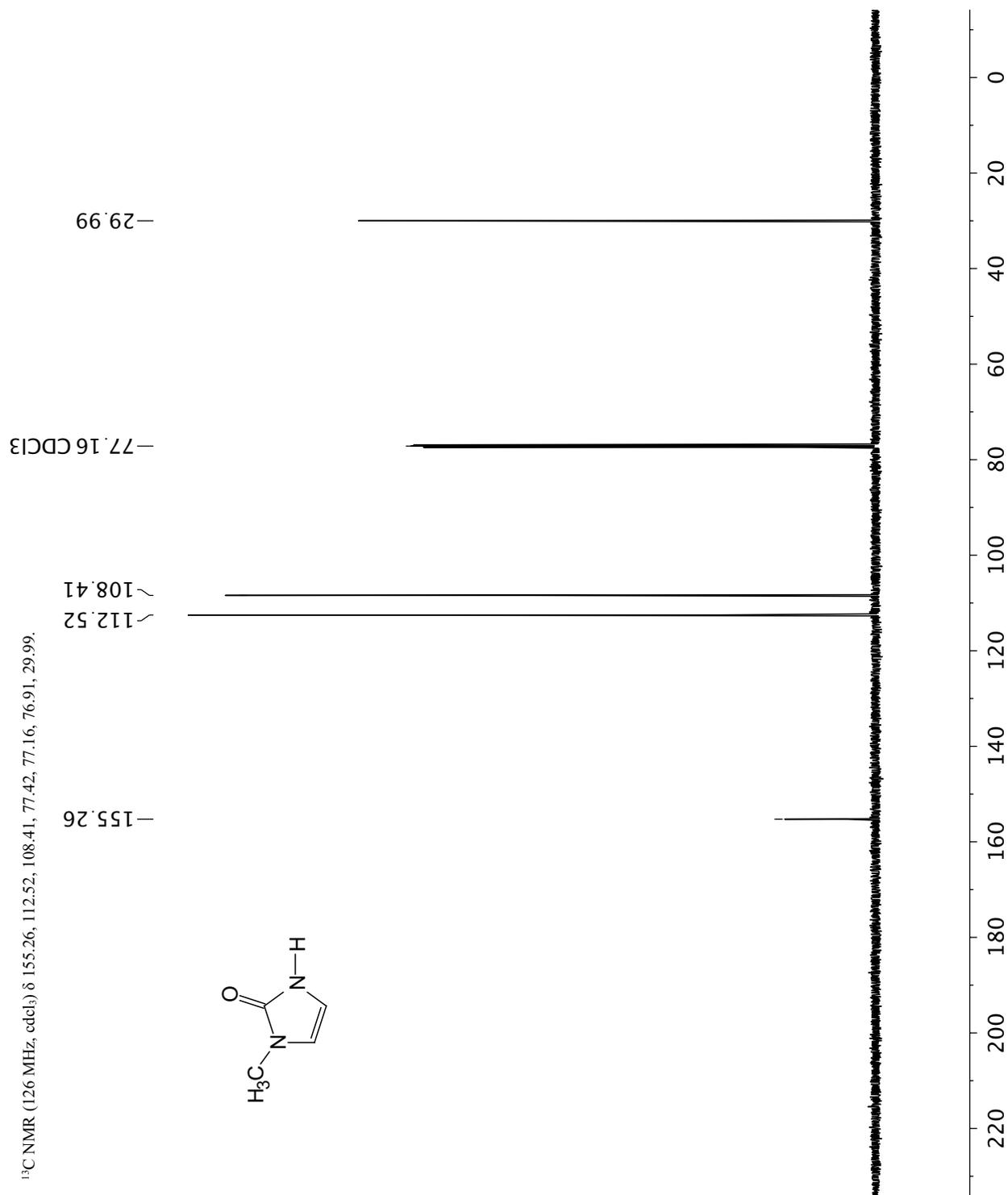
APPENDIX D:

NMR and IR spectra of compounds 1, 1', 2 and 2'

^1H NMR spectrum of **1** in CDCl_3

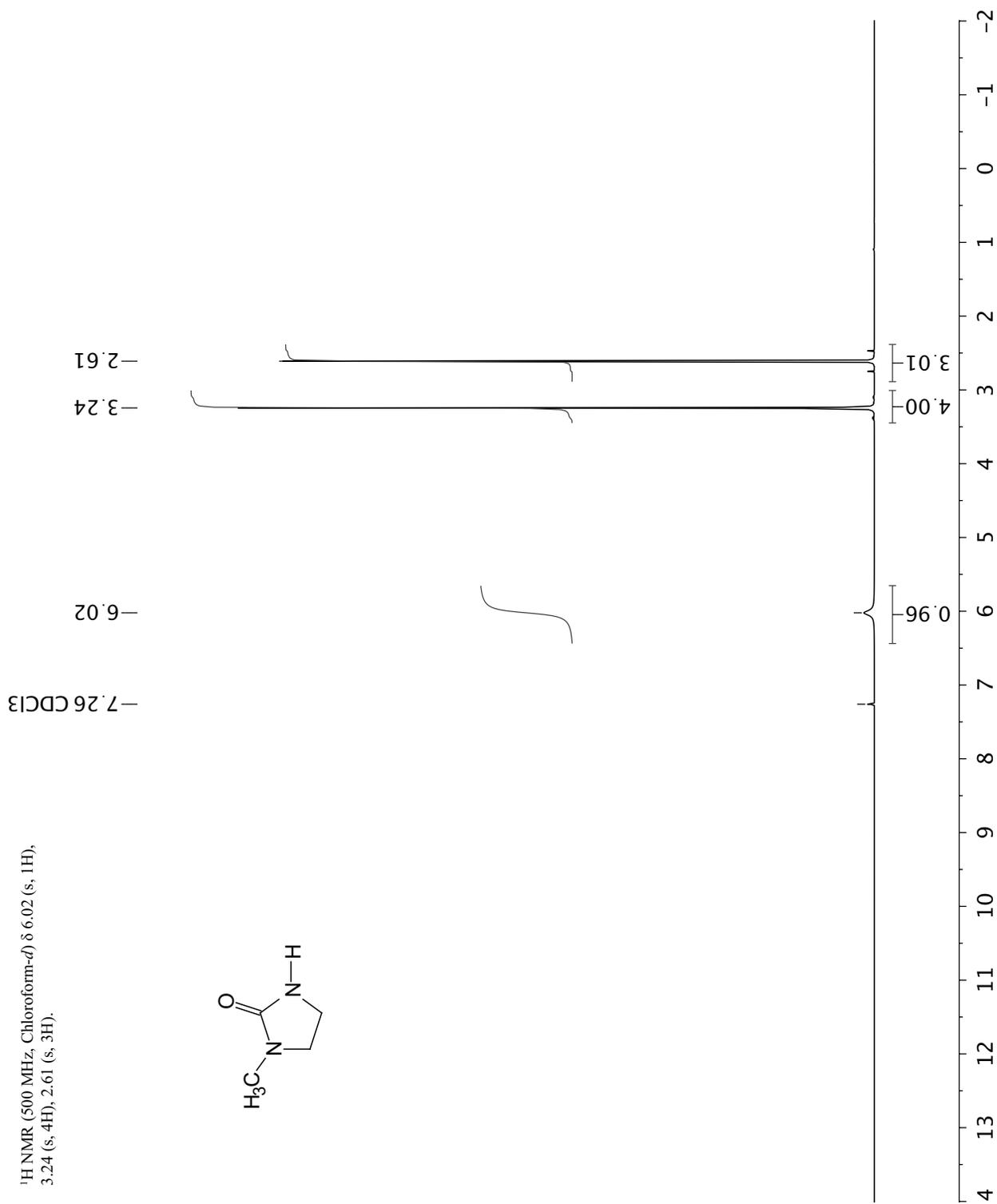
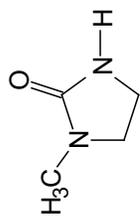


^{13}C NMR spectrum of **1** in CDCl_3

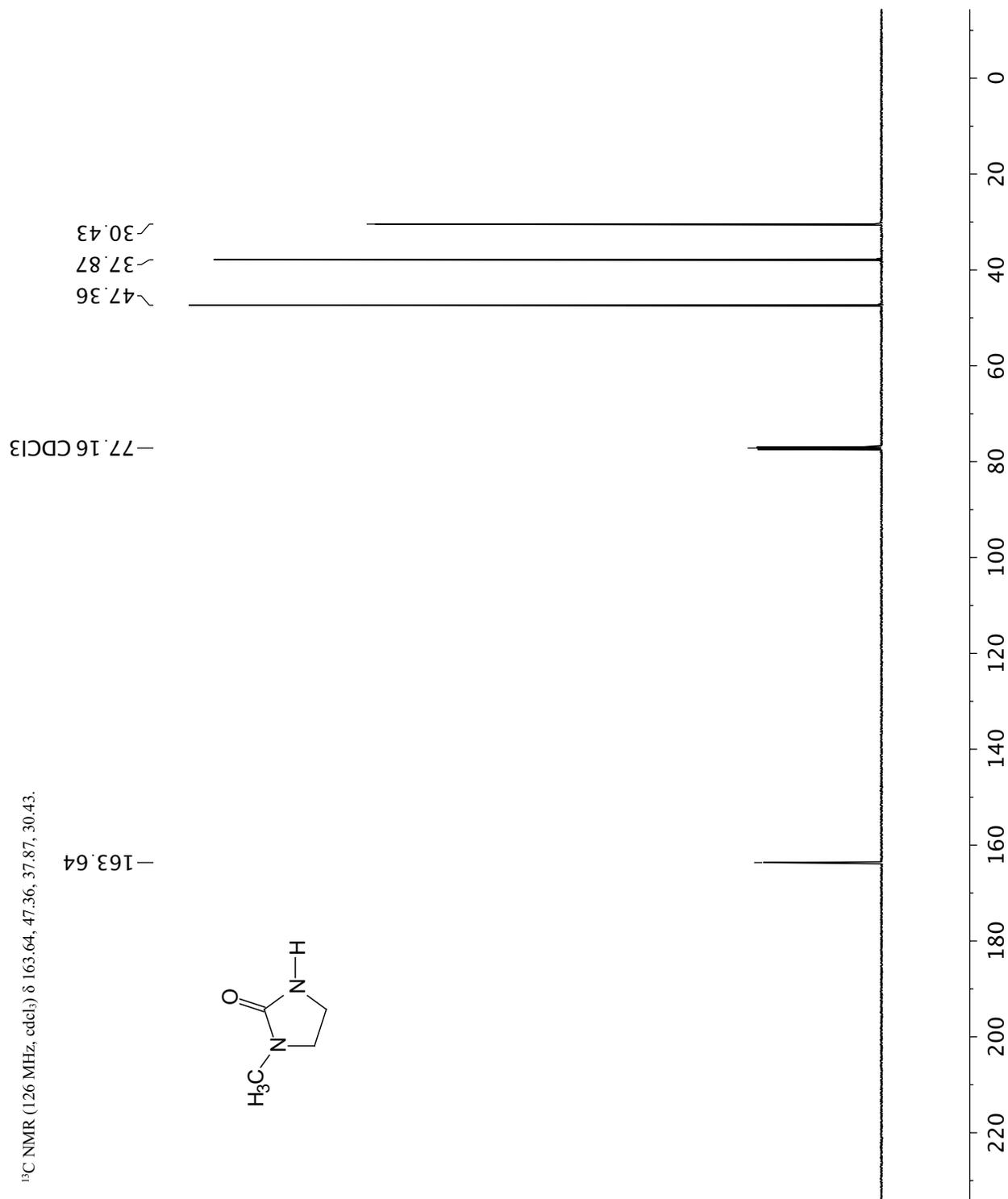


^1H NMR spectrum of **1'** in CDCl_3

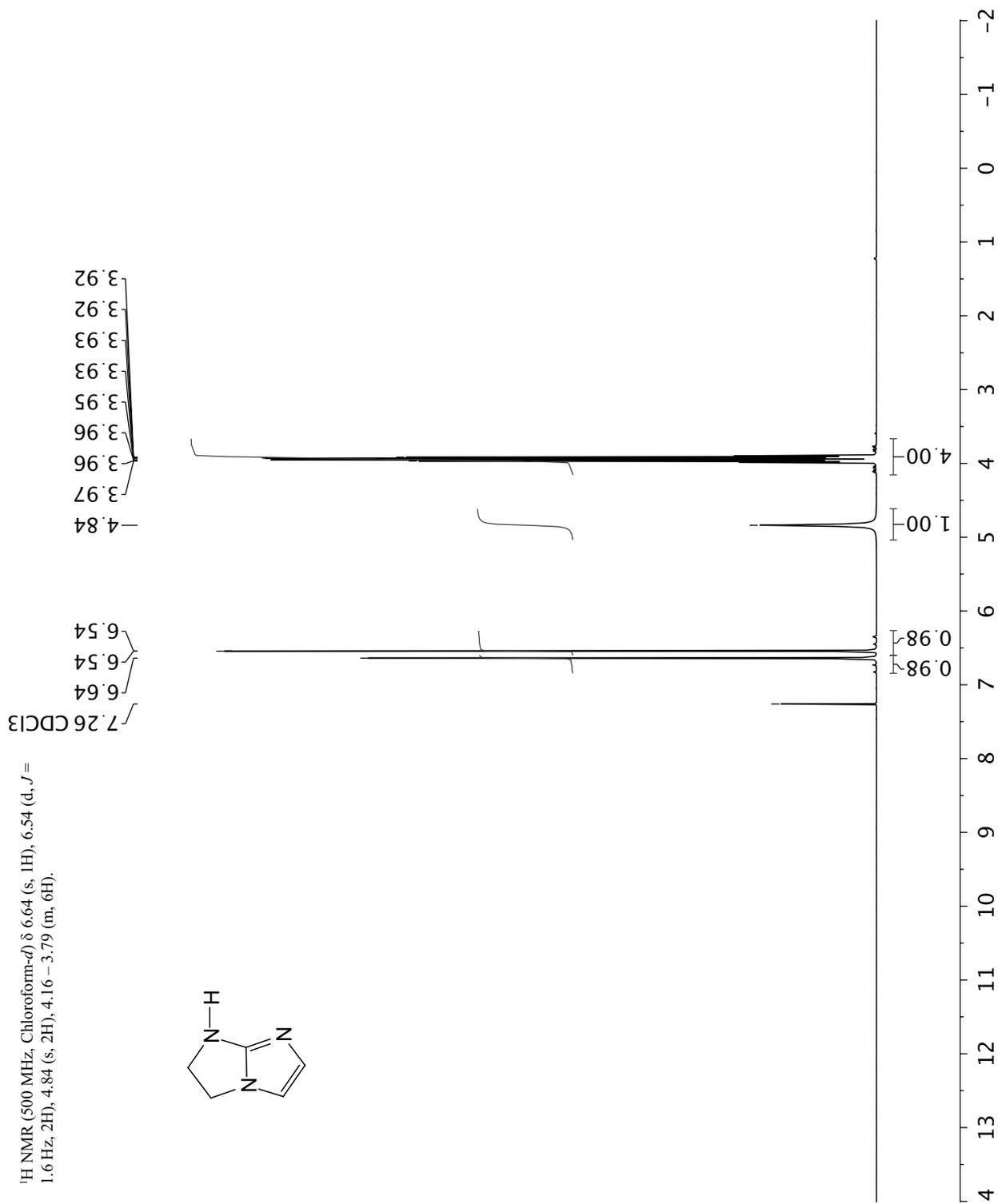
^1H NMR (500 MHz, Chloroform-*d*) δ 6.02 (s, 1H),
3.24 (s, 4H), 2.61 (s, 3H).



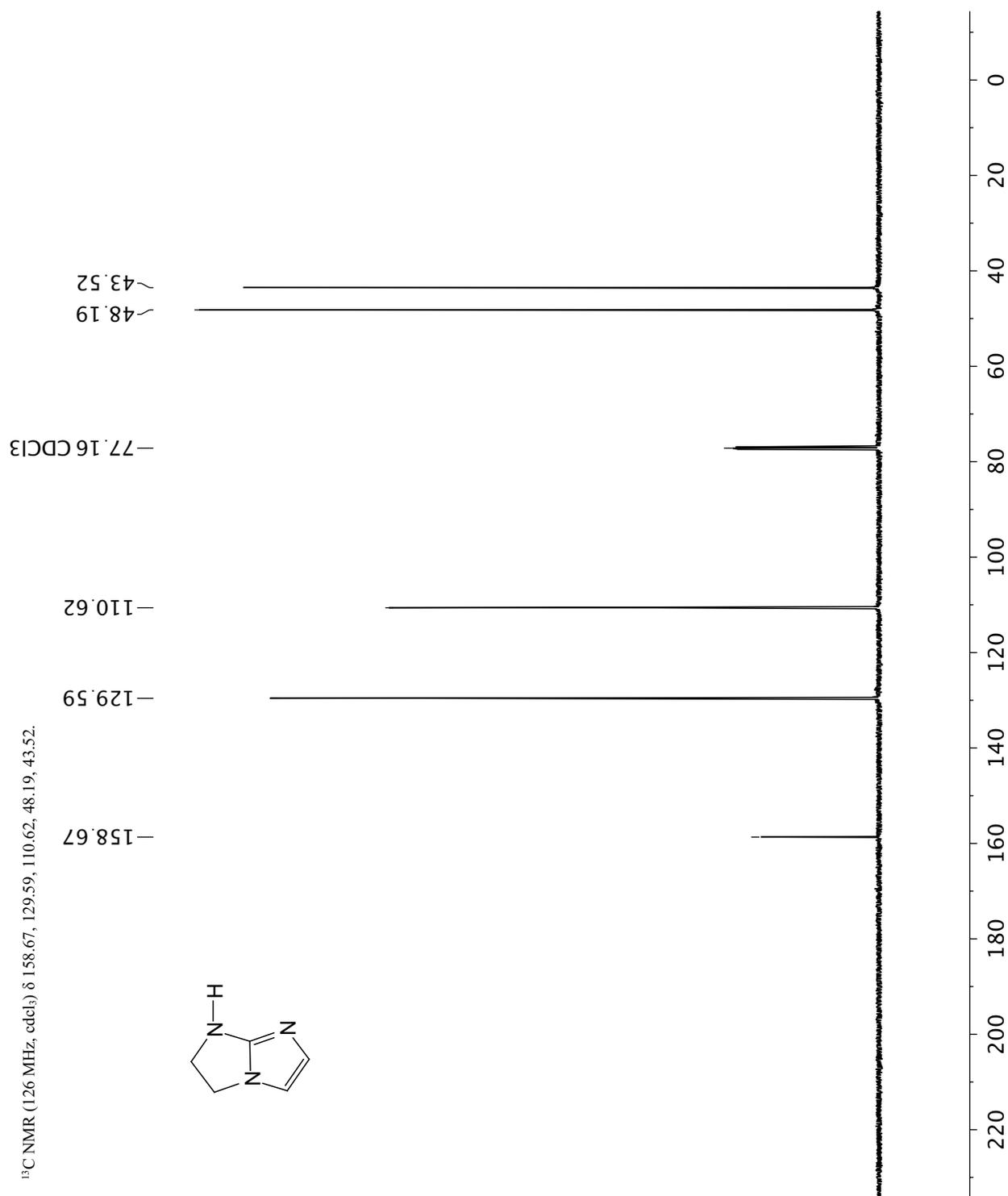
^{13}C NMR spectrum of **1'** in CDCl_3



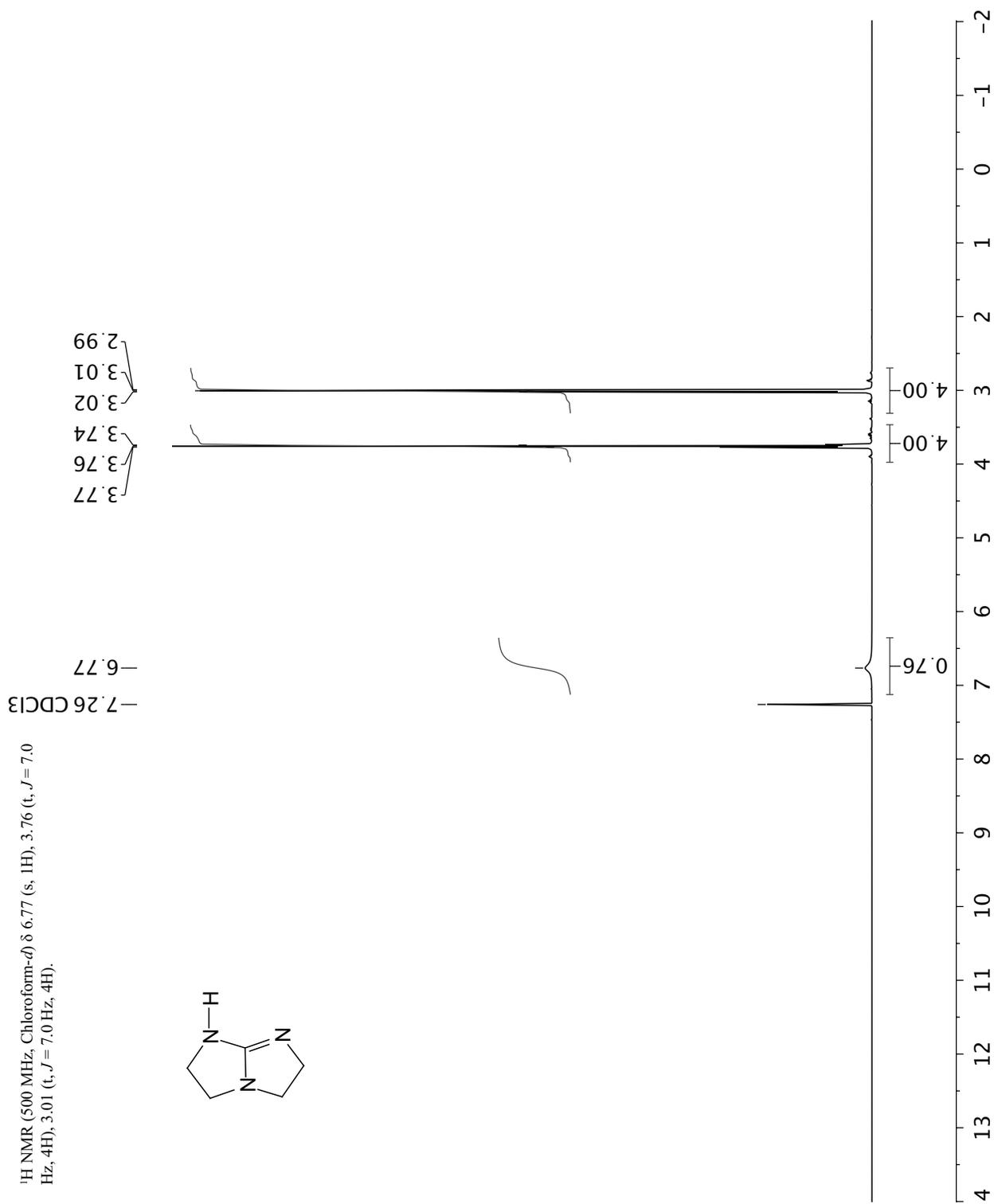
^1H NMR spectrum of **2** in CDCl_3



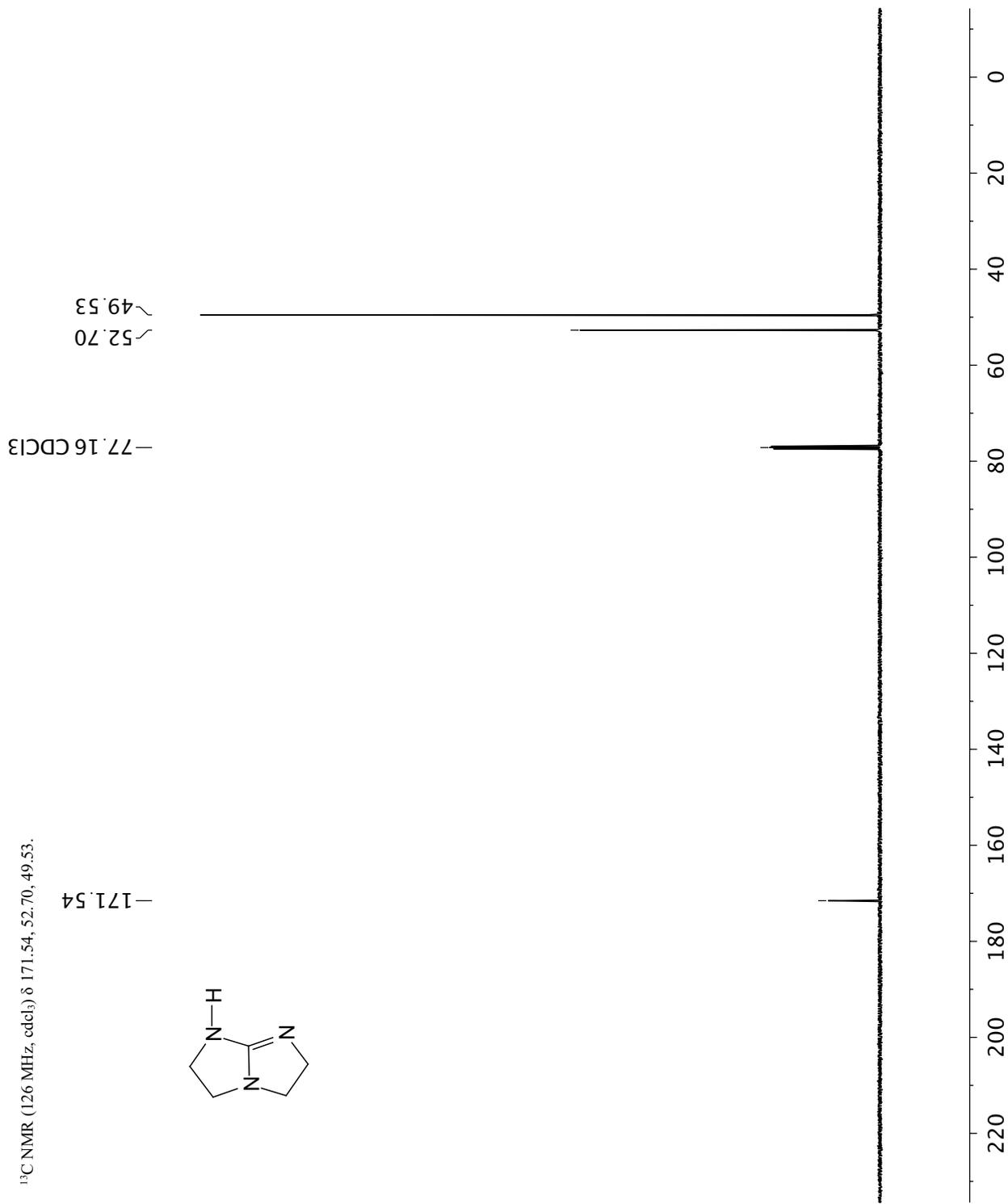
^{13}C NMR spectrum of **2** in CDCl_3



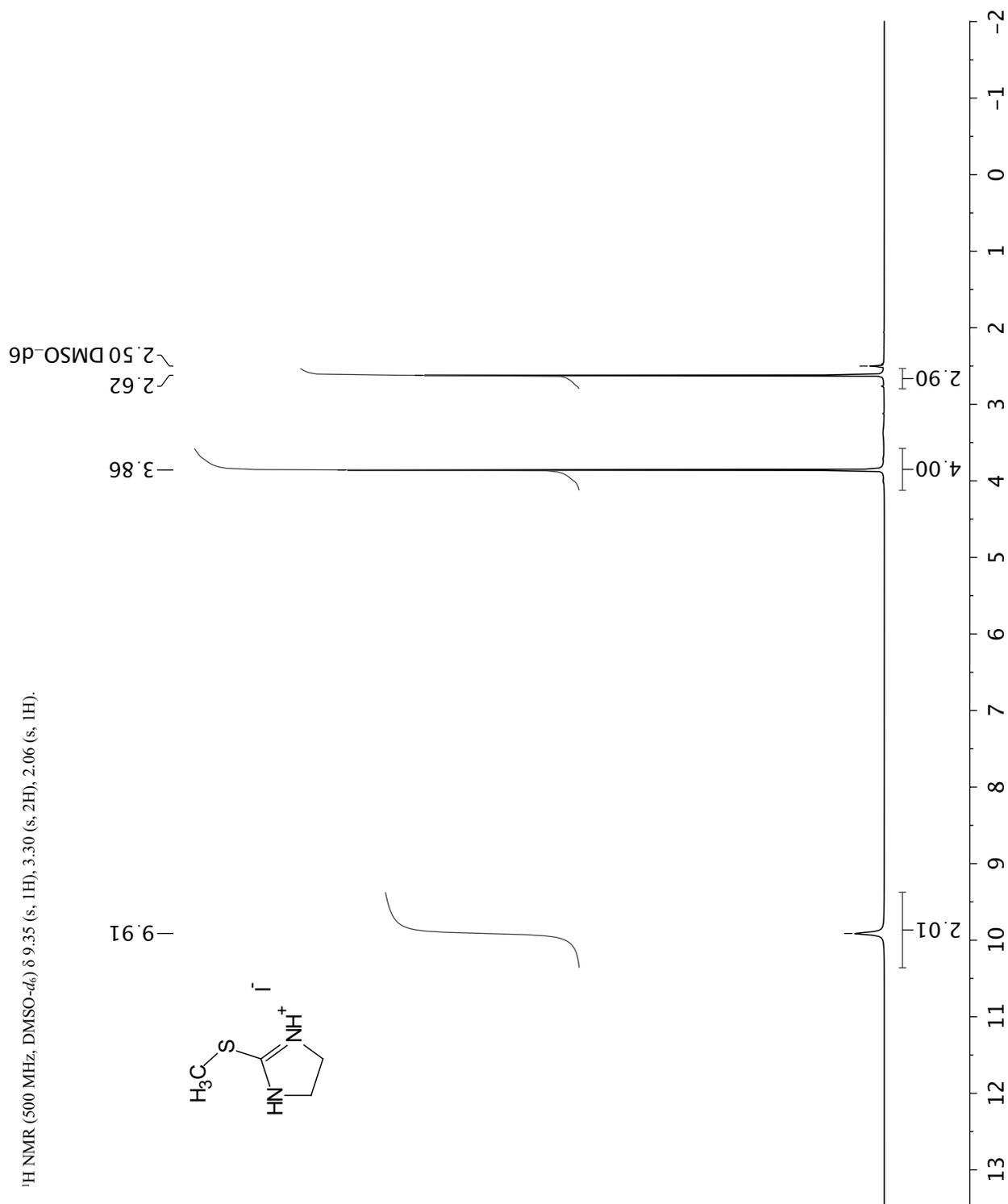
^1H NMR spectrum of **2'** in CDCl_3



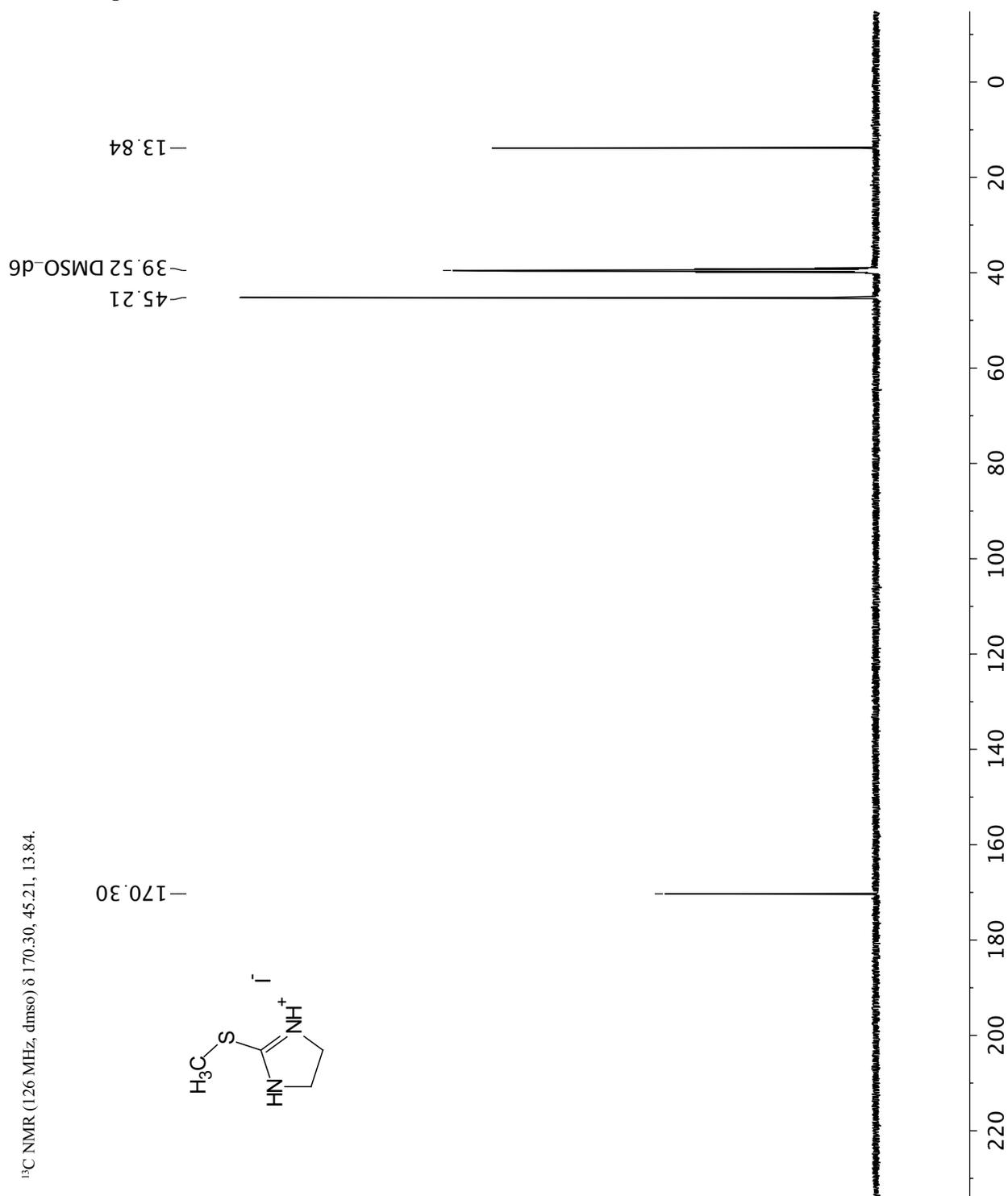
^{13}C NMR spectrum of **2'** in CDCl_3



^1H NMR spectrum of **5S** in $\text{DMSO-}d_6$

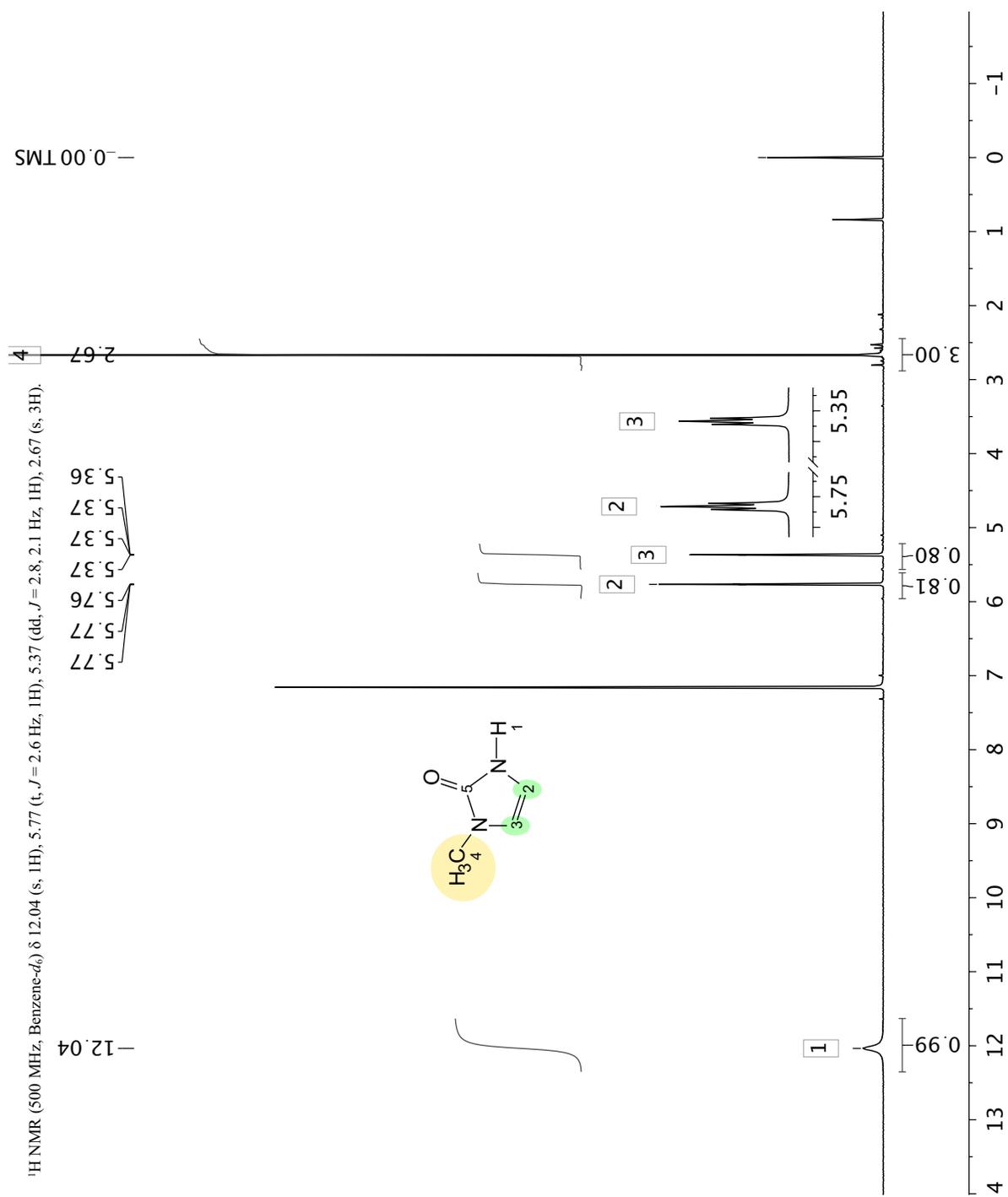


^{13}C NMR spectrum of **5S** in $\text{DMSO-}d_6$

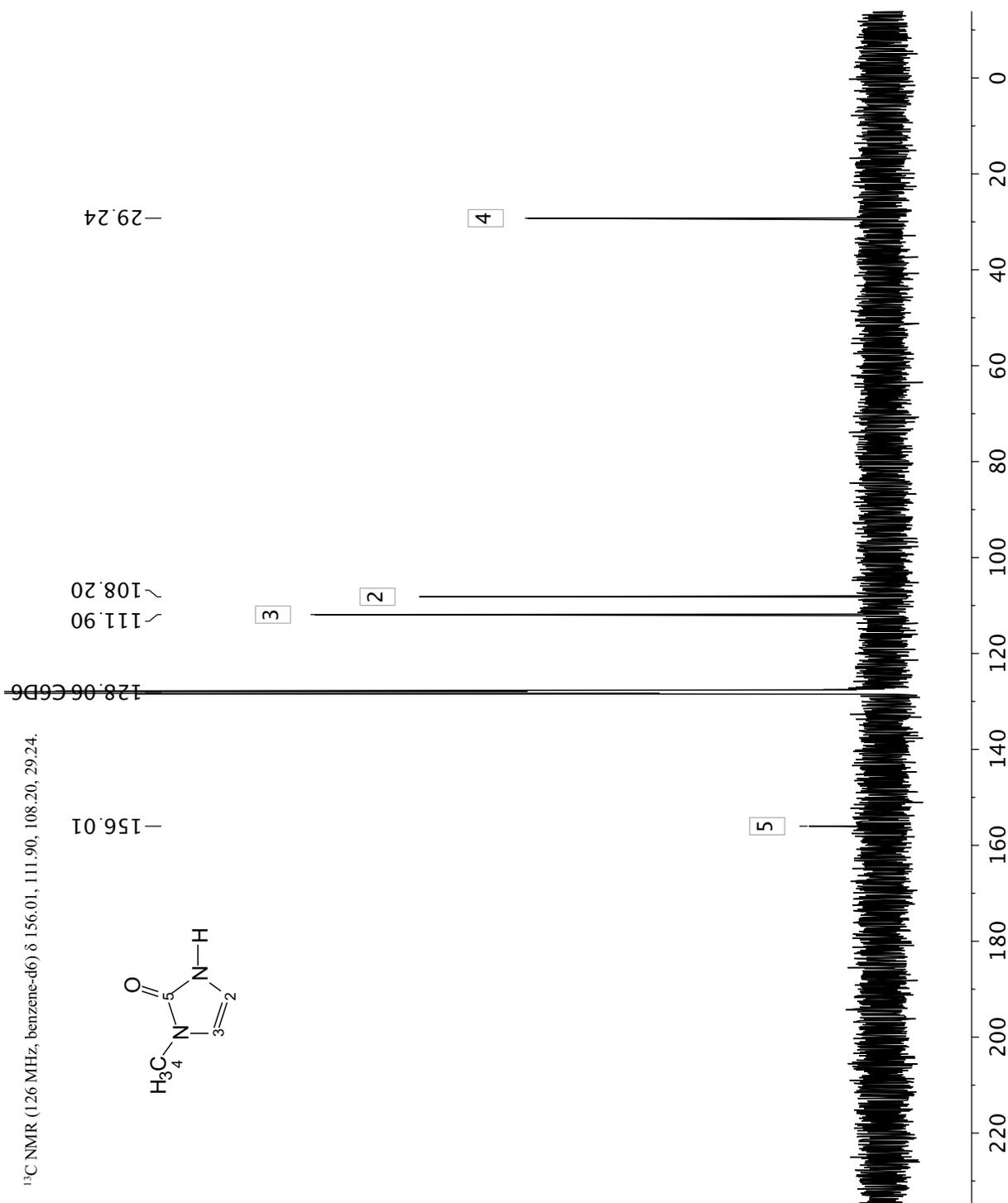


1D and 2D spectra used for the assignment of aromatic hydrogens of compounds 1 and 2

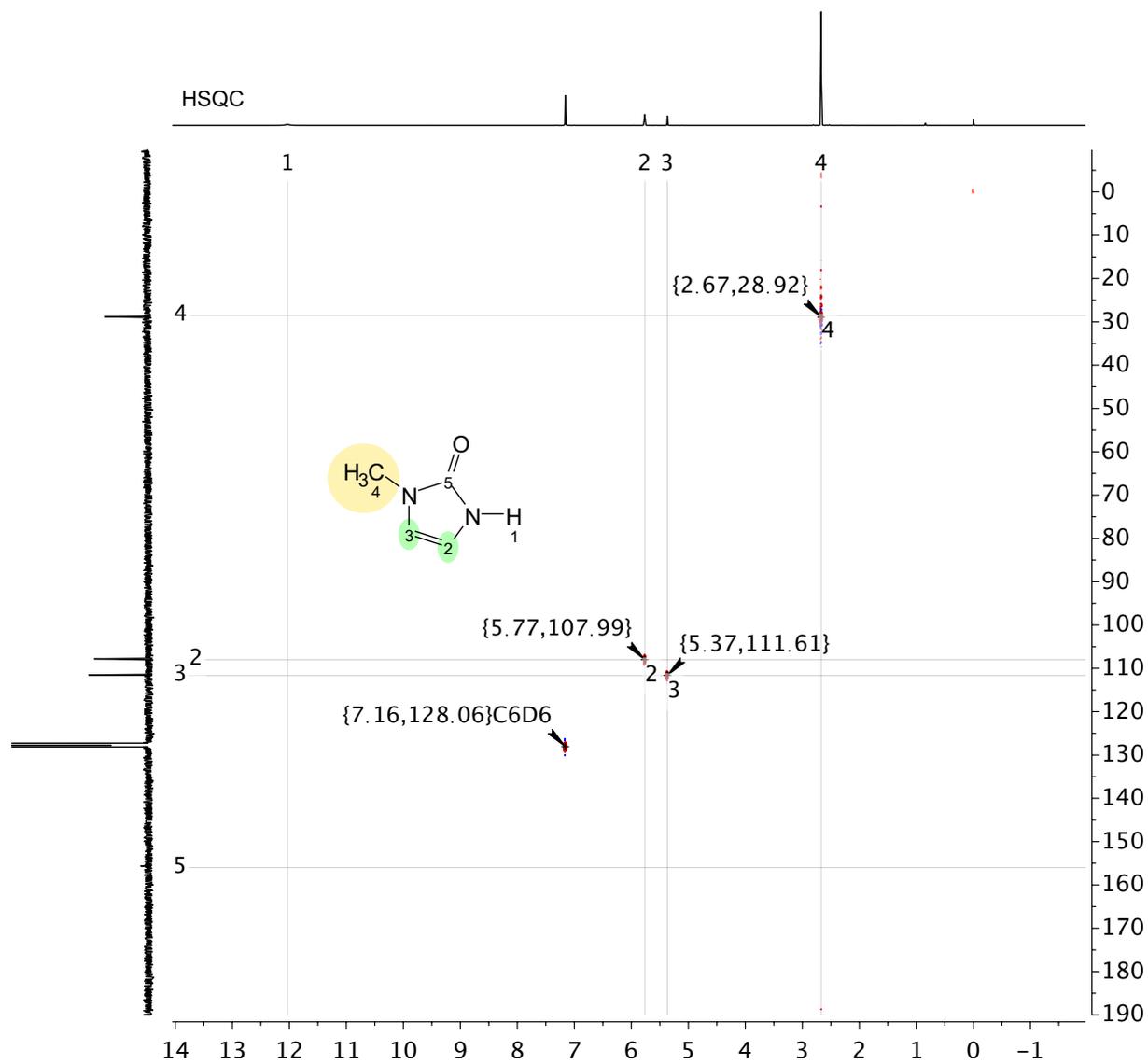
¹H NMR spectrum of 1 in C₆D₆



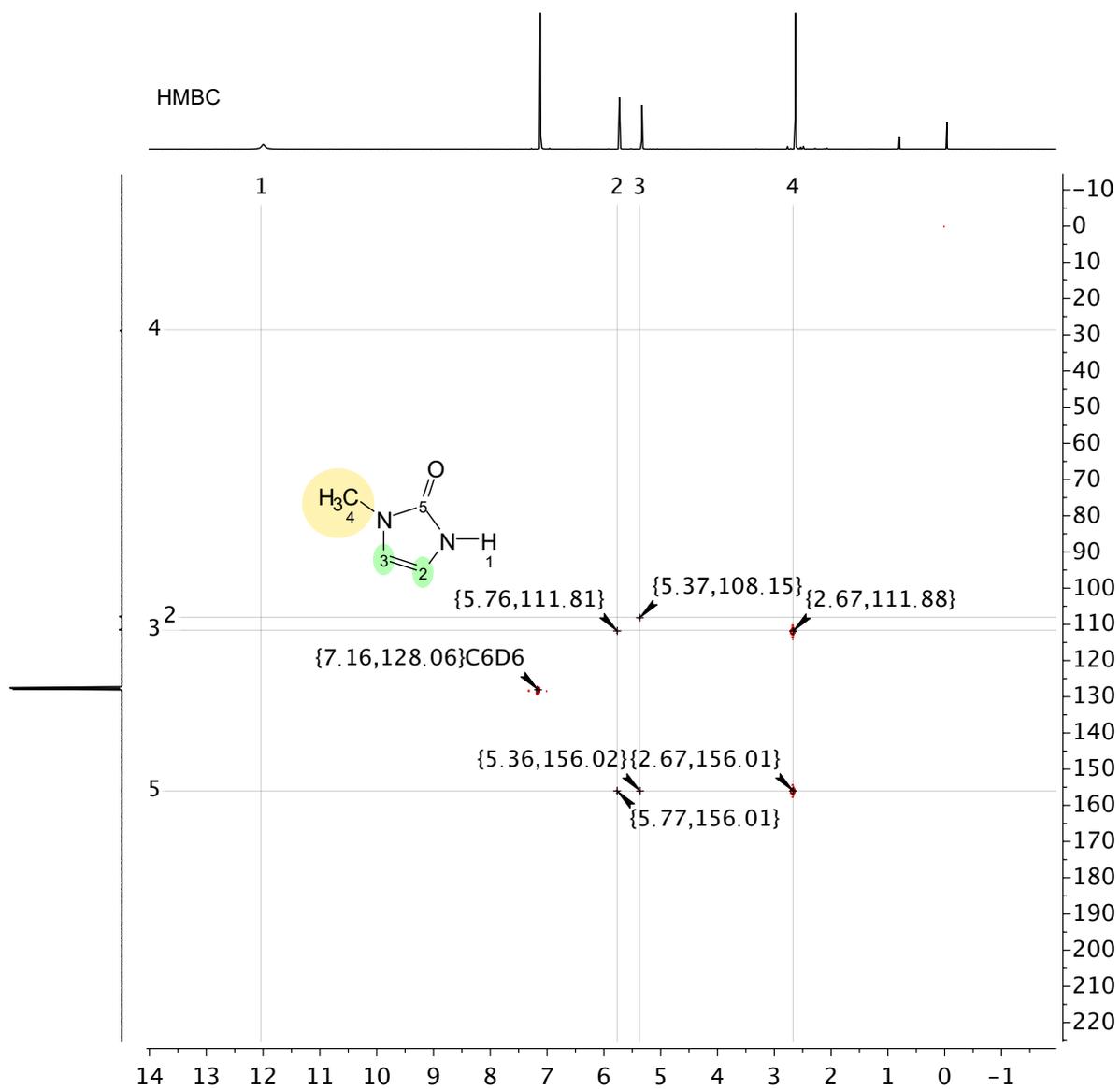
^{13}C NMR spectrum of **1** in C_6D_6



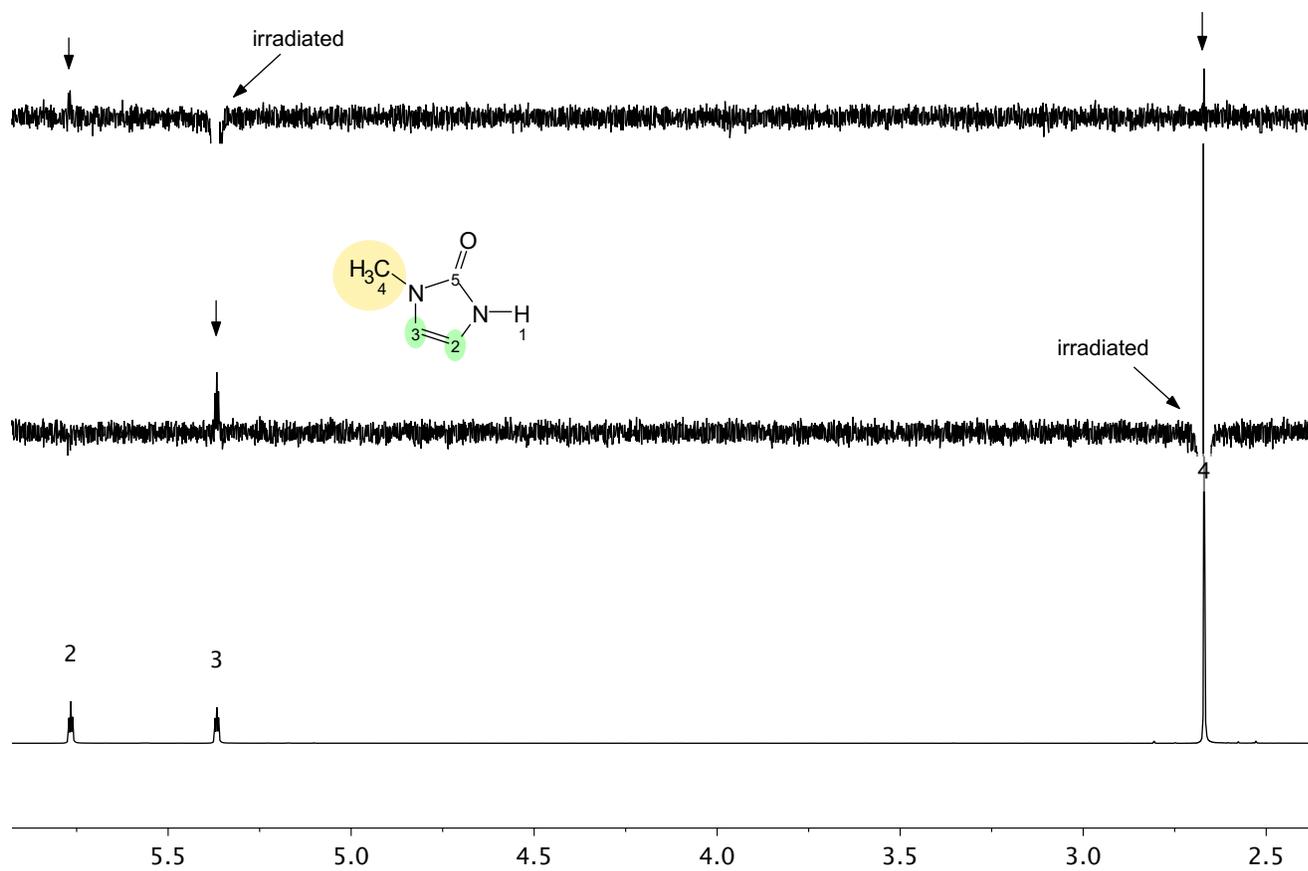
HSQC spectrum of **1** in C₆D₆



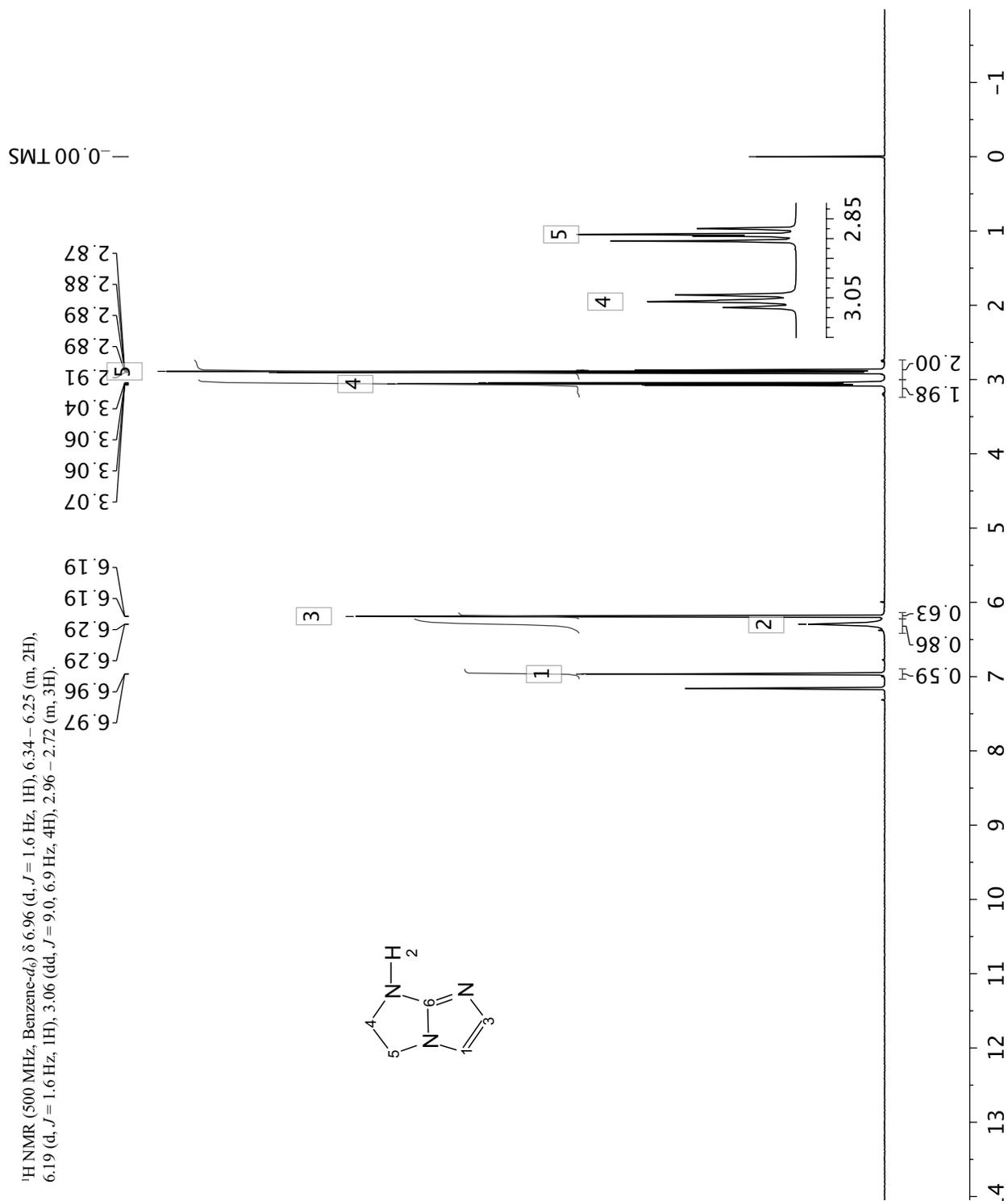
HMBC spectrum of **1** in C₆D₆



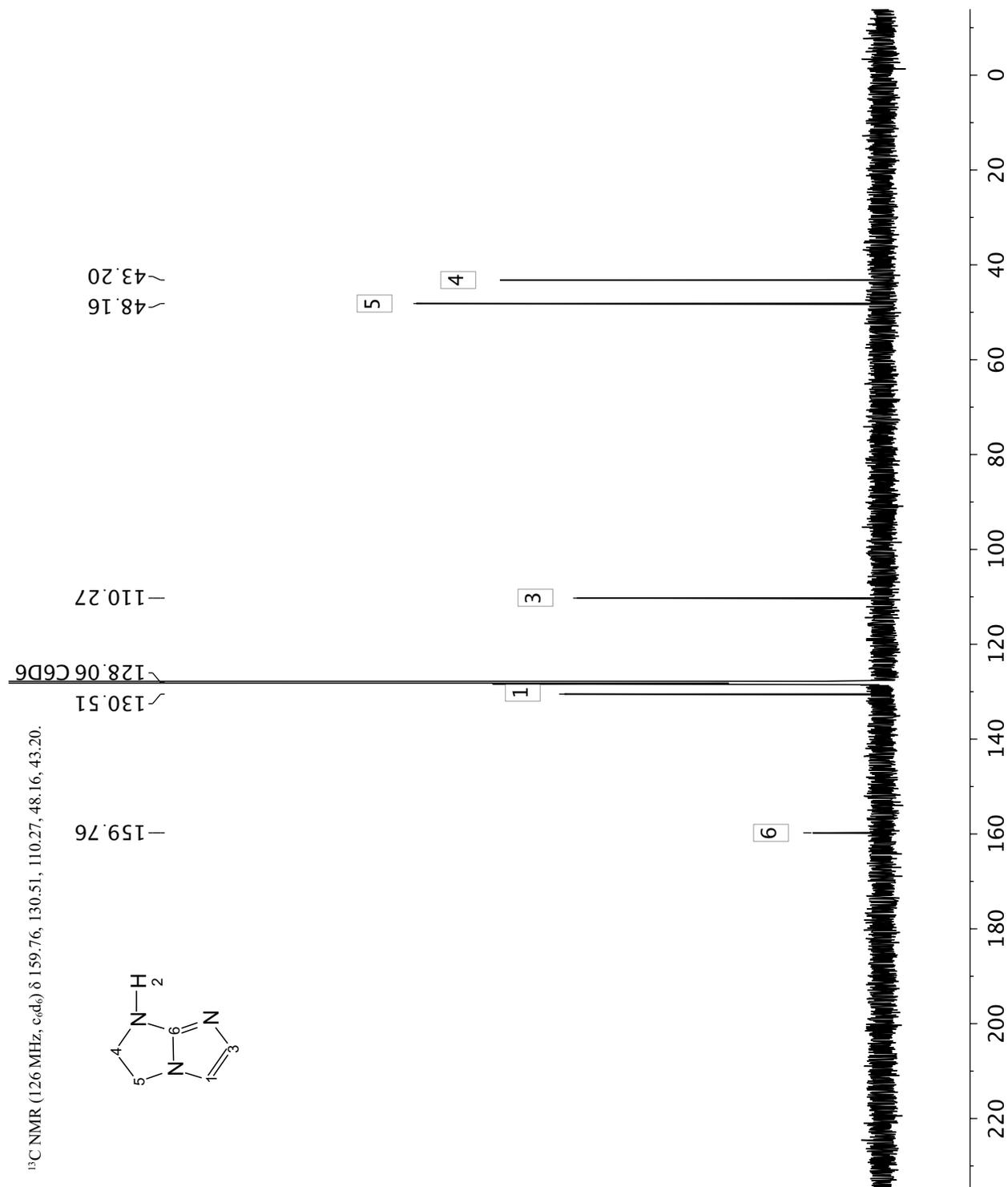
Stacked NOE spectra of **1** in C₆D₆



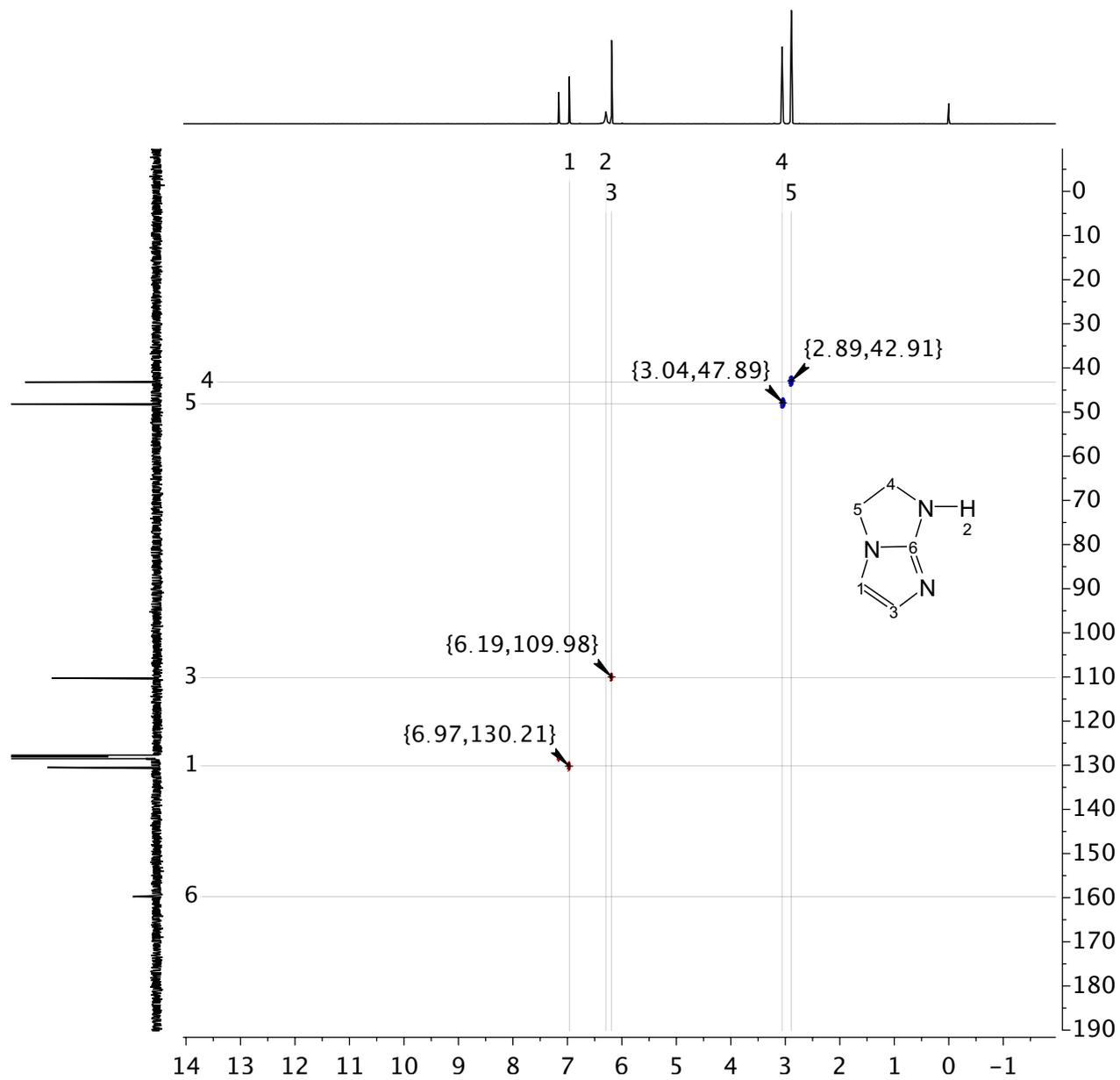
^1H NMR spectrum of **2** in C_6D_6



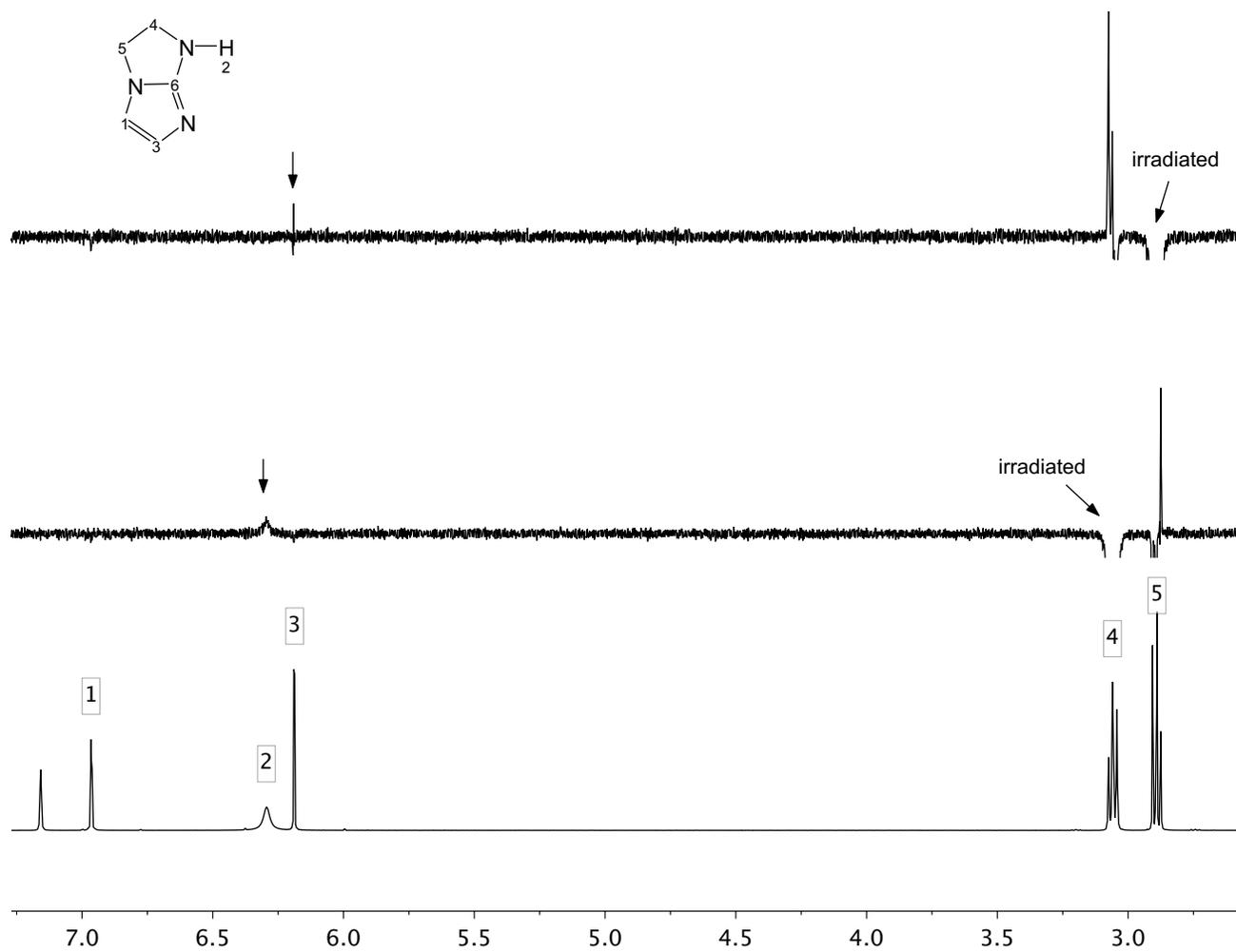
^{13}C NMR spectrum of **2** in C_6D_6



HSQC spectrum of **2** in C₆D₆

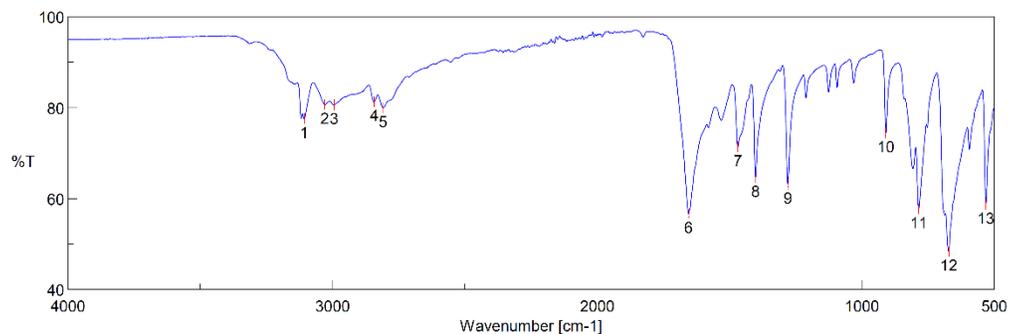


Stacked NOE spectra of **2** in C₆D₆



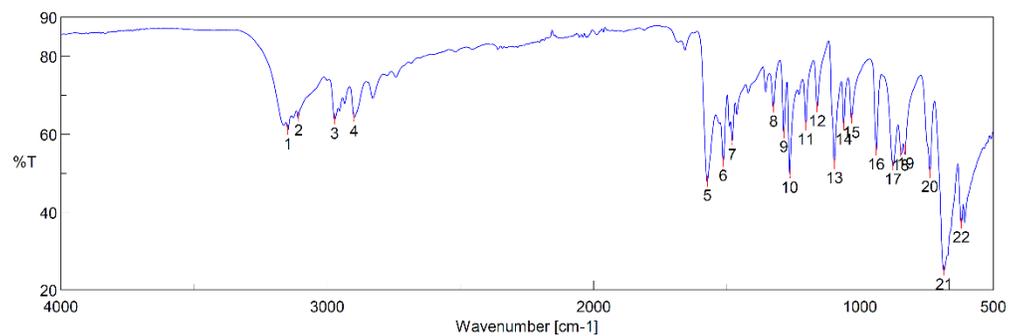
IR Spectra of compounds 1 and 2

Compound 1:



No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity
1	3107	77.5074	2	3030	80.631	3	2994	80.6081
4	2843	81.2464	5	2809	79.9288	6	1655	56.6362
7	1469	71.3278	8	1402	64.5854	9	1280	63.0954
10	909	74.4361	11	785	57.6802	12	671	48.3399
13	531	58.646						

Compound 2:



No.	Position	Intensity	No.	Position	Intensity	No.	Position	Intensity
1	3147	61.0129	2	3108	64.5119	3	2973	63.7132
4	2899	64.3007	5	1574	47.7998	6	1514	52.8813
7	1480	58.2944	8	1325	66.8319	9	1286	60.3116
10	1264	49.8722	11	1203	62.2041	12	1161	66.8608
13	1096	52.3351	14	1061	62.1982	15	1032	63.9797
16	939	55.8193	17	876	51.9343	18	848	55.5555
19	829	56.0611	20	738	50.0195	21	685	24.9754
22	620	37.0757						

APPENDIX E:

Wavefunction-based energies

Wavefunction-based energies for all species are presented in the appendix. DF-HF/aQZ values are the total electronic energies (density fitted) at the specified level; DF-MP2/aQZ values are the two-point extrapolated “MP2 correlation energies” (density fitted) at the specified level; CCSD(T)/aDZ values are the calculated total electronic energy (not density-fitted) at the specified level, CCSD(T)/CBS values are approximated CCSD(T) energies extrapolated to the complete basis set limit. The symmetries of species are shown in parentheses, if applicable. The table is divided into sections based on which figure the named structures belong in. For the redox systems, the hydrogenated species are specified using “-H2” at the end of their names, and the formic acid bound compounds are specified by “:FA”.

The following code was used to extract the CCSD(T)/CBS energies from a directory that had Gaussian16 geometry output files in the current directory (./), DF-MP2/aQ5Z Molpro^{137,138} output files in the “./DF-MP2/done” subdirectory, and CCSD(T)/aDZ Molpro output files in the “aDZ_CCSD_T/done/” subdirectory:

```
#!/bin/sh
echo "File DF-HF/a5Z DF-MP2/aQ5Z CCSD(T)/aDZ CCSD(T)/CBS/(kcal/mol)" >
CCSDT_CBS.csv
for fname in *.log; do
    name=${fname%.*}
    data=$(tail -3 aDZ_CCSD_T/done/$name.out 2> /dev/null |sed '2,3d')
    CCSDT=$(echo $data | awk '{print $1}')
    MP2=$(echo $data | awk '{print $2}')
    HF=$(echo $data | awk '{print $3}')
    MP2_aQ5Z=$(grep "SETTING DF_MP2_CORR_AQ5Z_M" DF-MP2/done/$name.out
2> /dev/null | awk '{print $4}')
    HF_a5Z=$(tail -3 DF-MP2/done/$name.out 2> /dev/null |sed '2,3d' | awk '{print
$2}') )
    if [ -n "$CCSDT" ] && [ -n "$MP2_aQ5Z" ]; then
        cbs=$(echo "627.509*($HF_a5Z + $MP2_aQ5Z + $CCSDT - $MP2)" | bc)
        fi
        [ -z "$CCSDT" ] && CCSDT="---" && cbs="---"
        [ -z "$MP2" ] && MP2="---" && cbs="---"
        [ -z "$HF" ] && HF="---" && cbs="---"
        [ -z "$MP2_aQ5Z" ] && MP2_aQ5Z="---" && cbs="---"
        [ -z "$HF_a5Z" ] && HF_a5Z="---" && cbs="---"
        echo "$name $HF_a5Z $MP2_aQ5Z $CCSDT $cbs" >> CCSDT_CBS.csv
done

cat CCSDT_CBS.csv | column -t
sed -i 's| |,|g' CCSDT_CBS.csv
```

Example CCSD(T)/aDZ input file for Molpro (formic acid):

```
geometry={  
5  
FA.log Energy: -118470.4941797  
O   1.17367   0.10891  -0.00000  
O  -1.04407  -0.44342   0.00000  
C   0.00000   0.42556  -0.00000  
H  -0.38064   1.46061  -0.00000  
H  -0.65615  -1.33794   0.00000  
}
```

```
memory,2400,m
```

```
basis=aVDZ  
hf  
HF_ADZ_M=energy  
mp2  
MP2_ADZ_M=energy  
ccsd(t)  
CCSD_T_ADZ_M=energy
```

```
DELTA = CCSD_T_ADZ_M - MP2_ADZ_M
```

```
show,DELTA
```

Example DF-MP2/aQ5Z input file for Molpro (formic acid):

```
geometry={
```

```
5
```

```
FA.log Energy: -118470.4941797
```

```
O    1.17367    0.10891   -0.00000
```

```
O   -1.04407   -0.44342    0.00000
```

```
C    0.00000    0.42556   -0.00000
```

```
H   -0.38064    1.46061   -0.00000
```

```
H   -0.65615   -1.33794    0.00000
```

```
}
```

```
memory,2400,m
```

```
BASIS=aVDZ
```

```
DF-HF,DF_BASIS=aVDZ/JKFIT
```

```
DF_HF_ADZ_M=energy
```

```
DF-MP2,DF_BASIS=aVDZ/MP2FIT
```

```
DF_MP2_ADZ_M=energy
```

```
DF_MP2_CORR_ADZ_M=(DF_MP2_ADZ_M - DF_HF_ADZ_M)
```

```
BASIS=aVTZ
```

```
DF-HF,DF_BASIS=aVTZ/JKFIT
```

```
DF_HF_ATZ_M=energy
```

```
DF-MP2,DF_BASIS=aVTZ/MP2FIT
```

```
DF_MP2_ATZ_M=energy
```

```
DF_MP2_CORR_ATZ_M=(DF_MP2_ATZ_M - DF_HF_ATZ_M)
```

```
BASIS=aVQZ
```

```
DF-HF,DF_BASIS=aVQZ/JKFIT
```

```
DF_HF_AQZ_M=energy
```

```
DF-MP2,DF_BASIS=aVQZ/MP2FIT
```

```
DF_MP2_AQZ_M=energy
```

```
DF_MP2_CORR_AQZ_M=(DF_MP2_AQZ_M - DF_HF_AQZ_M)
```

```
GTHRESH,THROVL=1.00D-9
```

```
BASIS=aV5Z
```

```
DF-HF,DF_BASIS=aV5Z/JKFIT
```

```
DF_HF_A5Z_M=energy
```

```
DF-MP2,DF_BASIS=aV5Z/MP2FIT
```

```
DF_MP2_A5Z_M=energy
```

```
DF_MP2_CORR_A5Z_M=(DF_MP2_A5Z_M - DF_HF_A5Z_M)
```

```
DF_MP2_CORR_ADTZ_M = ((27 * DF_MP2_CORR_aTZ_M) - (8 * DF_MP2_CORR_aDZ_M))/19
```

```
DF_MP2_ADTZ_M = DF_MP2_CORR_ADTZ_M + DF_HF_ATZ_M
```

$DF_MP2_CORR_ATQZ_M = ((64 * DF_MP2_CORR_aQZ_M) - (27 * DF_MP2_CORR_aTZ_M)) / 37$

$DF_MP2_ATQZ_M = DF_MP2_CORR_ATQZ_M + DF_HF_AQZ_M$

$DF_MP2_CORR_AQ5Z_M = ((125 * DF_MP2_CORR_a5Z_M) - (64 * DF_MP2_CORR_aQZ_M)) / 61$

$DF_MP2_AQ5Z_M = DF_MP2_CORR_AQ5Z_M + DF_HF_A5Z_M$

show,DF_MP2_ADTZ_M

show,DF_MP2_ATQZ_M

show,DF_MP2_AQ5Z_M

Energies of the structures for the main AMHB case in Figures I-5 to I-8				
cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
1a (C2v)	-299.832448	-1.24936507	-300.7481821	-188971.2323
1a dimer (C2h)	-599.6851583	-2.51207132	-601.5315286	-377962.8869
1a' (C2)	-301.0167305	-1.27288822	-301.9596451	-189736.4059
1a' dimer (C2)	-602.049544	-2.55504901	-603.9468021	-379488.8103
1a' dimer (Ci)	-602.049418	-2.55509392	-603.9466734	-379488.7423
1b (Cs)	-319.6603015	-1.27932864	-320.5794488	-201430.1431
1b dimer (C2h)	-639.3419929	-2.56911508	-641.1925899	-402879.9692
1b' dimer (C2)	-641.7216002	-2.6159716	-643.6234896	-404415.6176
1b' dimer (Ci)	-641.7215321	-2.61600044	-643.6234333	-404415.5801
1b'	-320.8518087	-1.30394142	-321.7976807	-202199.5802
1c (Cs)	-642.3187447	-1.219865	-643.2039551	-403869.2189
1c dimer (C2h)	-1284.656011	-2.45252837	-1286.440211	-807757.3607
1c' dimer (C2)	-1287.021655	-2.50067745	-1288.86303	-809285.8658
1c' dimer (Ci)	-1287.021617	-2.50073281	-1288.863079	-809285.8771
1c'	-643.503011	-1.24538408	-644.4177855	-404634.9709
2a dimer (C2)	-599.575739	-2.52632162	-601.4386773	-377903.6912
2a dimer (Ci)	-599.5764303	-2.52563735	-601.4390471	-377903.8741
2a	-299.7848202	-1.24702766	-300.7041933	-188942.3328
2a' dimer (C2)	-601.9366708	-2.56505541	-603.852725	-379427.371
2a' dimer (Ci)	-601.9367001	-2.56506823	-603.8527681	-379427.3966
2a'	-300.9623507	-1.27654466	-301.9135506	-189706.3192
2b dimer (C2h)	-639.1978489	-2.58084518	-641.0672456	-402799.5931
2b	-319.5923714	-1.27957856	-320.5182185	-201390.1982
2b' [1]	-320.7701895	-1.30901944	-321.7276326	-202153.975
2b' dimer (C2) [1]	-641.5532279	-2.62932188	-643.4792946	-404322.6879
2b' dimer (C2)	-641.5532219	-2.6293234	-643.4792965	-404322.6863
2b' dimer (Ci) [1]	-641.5530149	-2.62962658	-643.4793782	-404322.7243
2b' dimer (Ci)	-641.5530181	-2.62962489	-643.479378	-404322.7251
2b'	-320.7683902	-1.30859578	-321.7253257	-202152.6483
2c	-642.2995114	-1.21686903	-643.1806455	-403856.2939
2c dimer (C2h)	-1284.616293	-2.45386143	-1286.398341	-807734.2958
2c' dimer (C2)	-1286.975364	-2.49948661	-1288.815633	-809258.2046
2c' dimer (Ci)	-1286.975311	-2.49952168	-1288.81566	-809258.1963
2c'	-643.4801478	-1.24357489	-644.3934189	-404620.6555
3a dimer (C2)	-559.959781	-2.45504016	-561.8064213	-353005.0346
3a dimer (Ci)	-559.9605655	-2.45449653	-561.8065949	-353005.2882
3a	-279.9737399	-1.22289028	-280.8917333	-176495.8547
3a' dimer (C2)	-562.3026734	-2.5006276	-564.2056673	-354521.3281
3a' dimer (Ci)	-562.302701	-2.50076544	-564.2058235	-354521.4192
3a'	-281.1447868	-1.2447543	-282.090536	-177253.4595
3b dimer (C2)	-599.6268368	-2.50883664	-601.4740125	-377927.2579
3b dimer (Ci)	-599.6268812	-2.50880481	-601.4741065	-377927.2641
3b	-299.8066536	-1.24954215	-300.7247067	-188956.4894
3b' [1]	-300.9857816	-1.27417056	-301.9317204	-189719.1392
3b' dimer (C2) [1]	-601.9853038	-2.56003645	-603.8896909	-379453.3735
3b' dimer (C2)	-601.9853205	-2.5600193	-603.889695	-379453.3758
3b' dimer (Ci) [1]	-601.9853519	-2.55973707	-603.8893022	-379453.2201
3b' dimer (Ci)	-601.9854326	-2.55998908	-603.8898898	-379453.4695
3b'	-300.9856599	-1.27456875	-301.9320807	-189719.2758
3c dimer (C2)	-1244.941148	-2.39553118	-1246.722872	-782805.8504
3c dimer (Ci)	-1244.941163	-2.395501	-1246.722949	-782805.8357
3c	-622.4651953	-1.19255927	-623.3499945	-391396.4522
3c' [1]	-623.6390123	-1.21738416	-624.5548889	-392156.4622
3c' dimer (C2) [1]	-1247.289869	-2.44788776	-1249.135677	-784327.6011
3c' dimer (C2)	-1247.289867	-2.44788891	-1249.135677	-784327.6004

Energies of the structures for the main AMHB cases in Figures I-5 to I-8 (cont'd)

cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
3c' dimer (Ci) [1]	-1247.29007	-2.44775494	-1249.135892	-784327.7067
3c' dimer (Ci)	-1247.290065	-2.44775867	-1249.135892	-784327.7057
3c'	-623.6397267	-1.21776693	-624.5561494	-392157.1343
4a dimer (C2)	-559.9030376	-2.46676874	-561.7600546	-352975.5168
4a dimer (Ci)	-559.9038405	-2.46646775	-561.76045	-352975.9007
4a	-279.948716	-1.22934588	-280.872362	-176483.5199
4a' dimer (C2)	-562.2186284	-2.51166597	-564.139689	-354478.0762
4a' dimer (Ci)	-562.2182279	-2.51163535	-564.1391964	-354477.7881
4a'	-281.1043282	-1.25057197	-282.059204	-177232.9371
4b dimer (C2)	-599.5213348	-2.52271558	-601.3896215	-377871.6539
4b dimer (Ci)	-599.5214208	-2.52273166	-601.3898117	-377871.7164
4b	-299.754534	-1.25844087	-300.684893	-188930.0232
4b' [1]	-300.9185656	-1.28272791	-301.8794936	-189684.5496
4b' dimer (C2) [1]	-601.8507968	-2.57435002	-603.7819249	-379382.5823
4b' dimer (C2)	-601.850794	-2.57435339	-603.7819236	-379382.5818
4b' dimer (Ci) [1]	-601.8506779	-2.57444455	-603.7819824	-379382.5585
4b' dimer (Ci)	-601.8506692	-2.57445237	-603.7819823	-379382.5573
4b'	-300.918578	-1.28272152	-301.8794902	-189684.5512
4c dimer (C2)	-1244.931001	-2.39971615	-1246.709161	-782801.6239
4c dimer (Ci)	-1244.930994	-2.3997711	-1246.709315	-782801.6528
4c	-622.4608337	-1.19600709	-623.3449405	-391395.46
4c' [1]	-623.6274378	-1.21648372	-624.5428171	-392149.9044
4c' dimer (C2) [1]	-1247.266842	-2.44392539	-1249.108635	-784313.1005
4c' dimer (C2)	-1247.266874	-2.44428202	-1249.109325	-784313.39
4c' dimer (Ci) [1]	-1247.266746	-2.44390521	-1249.108692	-784313.068
4c' dimer (Ci)	-1247.266903	-2.44428194	-1249.109392	-784313.4254
4c'	-623.6275261	-1.21679088	-624.5432165	-392150.1108
5a (Cs)	-337.6812385	-1.4065704	-338.7268169	-212829.414
5a dimer (C2h)	-675.3788149	-2.82188242	-677.4814207	-425674.793
5a' dimer (C2)_C1	-677.760215	-2.88561946	-679.9255397	-427219.5603
5a' dimer (C2)	-677.7601984	-2.88562483	-679.9255345	-427219.5541
5a' dimer (Ci)_C1	-677.760201	-2.88562793	-679.9255407	-427219.5575
5a' dimer (Ci)	-677.7602019	-2.88562725	-679.9255406	-427219.5575
5a'	-338.8719849	-1.43723477	-339.9478799	-213601.1692
5b (Cs)	-357.5255426	-1.43668033	-358.570739	-225296.2295
5b dimer (C2h)	-715.0690572	-2.88077114	-717.1697118	-450608.7812
5b' dimer (C2)	-717.4474094	-2.94974257	-719.6155108	-452154.5346
5b' dimer (Ci)	-717.4474113	-2.94974081	-719.6155118	-452154.5347
5b'	-358.7146109	-1.46924096	-359.7918445	-226068.0226
5c (Cs)	-680.1705273	-1.3760872	-681.184147	-427726.8077
5c dimer (C2h)	-1360.358914	-2.76022996	-1362.397339	-855470.2684
5c' dimer (C2)	-1362.727137	-2.82701809	-1364.831827	-857008.5261
5c' dimer (Ci)	-1362.727139	-2.82701879	-1364.831829	-857008.5283
5c'	-681.3548878	-1.40768469	-682.4000485	-428495.1663
6a (Cs)	-337.7015359	-1.40253589	-338.7419187	-212839.2526
6a dimer (C2h)	-675.4133547	-2.81490363	-677.5067141	-425691.2425
6a'	-338.8961842	-1.43404447	-339.9666354	-213613.5978
6a' dimer (C2)	-677.8029968	-2.87758514	-679.9562559	-427240.0088
6a' dimer (Ci)	-677.8029888	-2.87759078	-679.9562548	-427240.006
6b (C2v)	-357.5599612	-1.4317008	-358.5995488	-225314.6375
6b dimer (C2h)	-715.131351	-2.87152289	-717.2214201	-450641.7763
6b' (C2v)	-358.7582231	-1.46420513	-359.8283774	-226091.643
6b' dimer (C2h)	-717.5288431	-2.93697234	-719.6804382	-452196.621
6c (Cs)	-680.1977379	-1.37011576	-681.2053354	-427740.2312

Energies of the structures for the main AMHB cases in Figures I-5 to I-8 (cont'd)

cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
6c dimer (C2h)	-1360.405725	-2.75012508	-1362.433726	-855493.2113
6c' (Cs)	-681.3929678	-1.40065894	-682.429896	-428514.4292
6c' dimer (C2h)	-1362.796663	-2.81080759	-1364.88298	-857041.6882
7a (Cs)	-317.8245251	-1.37306597	-318.8661613	-200352.3597
7a dimer (C2h)	-635.6624271	-2.7610224	-637.7619436	-400721.6628
7a' dimer (C2)	-638.0386673	-2.8298886	-640.2050049	-402265.9496
7a' dimer (Ci)	-638.0386666	-2.8298885	-640.2050056	-402265.9495
7a'	-319.0145976	-1.4084834	-320.0909168	-201126.3818
7b (Cs)	-337.6735434	-1.40266605	-338.7146345	-212822.1596
7b dimer (C2h)	-675.3611736	-2.82091003	-677.4600272	-425661.9909
7b' dimer (C2h)	-677.7423679	-2.89238294	-679.909839	-427210.5157
7b'	-338.8663862	-1.43945113	-339.9430023	-213598.4854
7c (Cs)	-660.3178984	-1.34195136	-661.327294	-415252.3459
7c dimer (C2h)	-1320.649741	-2.70111973	-1322.687075	-830523.1416
7c' dimer (C2)	-1323.01857	-2.77234806	-1325.124987	-832063.7405
7c' dimer (Ci)	-1323.018574	-2.77234444	-1325.124989	-832063.7435
7c'	-661.5050004	-1.3791661	-662.5507178	-416025.2904
8a	-317.8378639	-1.37291703	-318.8784942	-200360.2197
8a dimer (C2h)	-635.6916364	-2.75831081	-637.7868921	-400737.6705
8a' dimer (C2)	-638.0825183	-2.81918323	-640.2364409	-402286.3048
8a' dimer (Ci)	-638.0825154	-2.81918393	-640.2364376	-402286.3054
8a'	-319.0341643	-1.40307833	-320.1038207	-201134.9573
8b (Cs)	-337.6904596	-1.40143194	-338.7294508	-212831.6775
8b dimer (C2h)	-675.4048676	-2.81677851	-677.4983396	-425686.3694
8b' (Cs)	-338.8935895	-1.4329658	-339.9621784	-213611.0581
8b' dimer (C2h)	-677.8094579	-2.87903367	-679.9615314	-427243.7079
8c (Cs)	-660.3279831	-1.34046815	-661.3353167	-415257.3891
8c dimer (C2h)	-1320.680992	-2.69560652	-1322.712404	-830538.9061
8c' (Cs)	-661.5271963	-1.36894468	-662.562158	-416032.9355
8c' dimer (C2h)	-1323.0799	-2.75200196	-1325.166016	-832089.8919

Energies of the structure for the AMHB case study in Figure I-10				
cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
9 (Cs)	-355.7027398	-1.55668884	-356.874087	-224234.7747
9 dimer (C2h)	-711.4256407	-3.131799	-713.7876858	-448492.0543
10	-356.8679428	-1.576723	-358.0678758	-224988.9224
10 dimer (C2)	-713.754775	-3.17312756	-716.1733532	-449999.4834
10 dimer (Ci)	-713.7549317	-3.1729724	-716.1734777	-449999.5378
11	-356.8834637	-1.57942871	-358.0832706	-224998.2001
11 dimer (C2)	-713.7800417	-3.17304979	-716.1961342	-450012.8939
11 dimer (Ci)	-713.780695	-3.17209429	-716.1956881	-450012.7883
12	-358.0573836	-1.60268975	-359.2861315	-225758.2854
12 dimer (C2)	-716.1306709	-3.21920925	-718.6033596	-451534.3314
12 dimer (Ci)	-716.1304132	-3.21939837	-718.6030909	-451534.2311

Energies of the structures for the fused rings in Figure 12				
	DF-HF/a5Z	DF-MP2/aQ5Z	CCSD(T)/aDZ	CCSD(T)/CBS
	(a.u)	(a.u)	(a.u)	(kcal/mol)
13 dimer	-948.8070042	-4.13180154	---	---
13	-474.3895175	-2.05817079	---	---
13' dimer	-951.1401735	-4.18095861	---	---
13'	-475.559067	-2.08331416	---	---
13'' dimer	-951.137489	-4.18151208	---	---
13''	-475.5579203	-2.08327427	---	---
14 dimer	-948.8563775	-4.11812359	---	---
14	-474.4188472	-2.05245742	---	---
14' dimer	-951.139054	-4.18516881	---	---
14'	-475.5584618	-2.08561762	---	---
14'' dimer	-951.1377559	-4.18586352	---	---
14''	-475.5582475	-2.08555362	---	---
15 dimer	-948.8555022	-4.11858376	---	---
15	-474.4186015	-2.05117752	---	---
15' dimer	-951.138329	-4.18492582	---	---
15'	-475.5578982	-2.08409325	---	---
15'' dimer	-951.1346375	-4.18657193	---	---
15''	-475.5560914	-2.08480694	---	---

Energies of the structures for Hybridization vs AMHB in Figure I-18

cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
1 dimer (C2h)	-677.77540313	-2.90080507	-679.93541184	-427231.87750
1	-338.87785110	-1.44343803	-339.94997571	-213605.73788
1' dimer (C2)	-680.13827916	-2.94347318	-682.34992915	-428757.14985
1' dimer (Ci)	-680.13815120	-2.94351287	-682.34976998	-428757.07518
1'	-340.06152125	-1.46688351	-341.16123427	-214370.69011
1x dimer (C2)	-755.87689744	-3.26154981	-758.33400267	-476496.05399
1x dimer (Ci)	-755.87690335	-3.26151781	-758.33396021	-476496.03434
1x	-377.93036012	-1.62548236	-379.15196639	-238239.56944
1y dimer (C2)	-755.88063589	-3.26343939	-758.33776576	-476498.92407
1y dimer (Ci)	-755.88063541	-3.26343938	-758.33765113	-476498.90171
1y	-377.93220565	-1.62720212	-379.15501415	-238241.44610
1z dimer (C2h)	-831.61806225	-3.58981478	-834.32905164	-524241.70092
1z (Cs)	-415.80089362	-1.78964045	-417.14917160	-262112.28792
2 dimer (C2)	-713.78004172	-3.17304979	-716.19613421	-450012.89386
2 dimer (Ci)	-713.78069503	-3.17209429	-716.19568807	-450012.78826
2	-356.88346369	-1.57942871	-358.08327056	-224998.20007
2' dimer (C2)	-716.13067085	-3.21920925	-718.60335960	-451534.33144
2' dimer (Ci)	-716.13041319	-3.21939837	-718.60309089	-451534.23110
2'	-358.05738357	-1.60268975	-359.28613151	-225758.28540
2x dimer (C2)	-791.87521525	-3.53497689	-794.59160233	-499275.53573
2x dimer (Ci)	-791.87522087	-3.53494563	-794.59151194	-499275.50652
2x	-395.92829942	-1.76032625	-397.27844990	-249627.90569
2y dimer (C2)	-791.86991180	-3.53542570	-794.58582645	-499272.48856
2y dimer (Ci)	-791.86975027	-3.53560816	-794.58568330	-499272.45192
2y	-395.92656687	-1.76082872	-397.27686794	-249627.10029
2z dimer (C2)	-867.61261553	-3.85951196	-870.57993548	-547016.98642
2z dimer (Ci)	-867.61248704	-3.85960817	-870.57985739	-547016.94083
2z	-433.79700540	-1.92240565	-435.27242091	-273498.53764

Energies of the structures for the RAHB cases in Figure I-20

cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
An1	-265.13813394	-1.20508534	-266.06902938	-167188.12341
An1 dimer (C2)	-530.28744479	-2.42254765	-532.16174110	-334390.00889
An1 dimer (Ci)	-530.28724165	-2.42255752	-532.16147312	-334389.85457
An1 [1]	-265.13796719	-1.20487476	-266.06862419	-167187.90691
An1 dimer (C2) [1]	-530.28744081	-2.42254782	-532.16174310	-334390.00774
An1 dimer (Ci) [1]	-530.28751061	-2.42251430	-532.16190189	-334390.07838
An2	-263.94049754	-1.17368485	-264.84090254	-166411.68355
An2 dimer (C2)	-527.88630703	-2.35826355	-529.69917488	-332832.86635
An2 dimer (Ci)	-527.88644225	-2.35796503	-529.69899968	-332832.79280
An3	-263.95130335	-1.17349423	-264.85012300	-166417.76188
An3 dimer (C2)	-527.91503085	-2.35740303	-529.72374648	-332849.28243
An3 dimer (Ci)	-527.91502350	-2.35749936	-529.72396678	-332849.35855
An4	-263.94924588	-1.17262218	-264.84758672	-166416.25909
An4 dimer (C2)	-527.91229461	-2.35686637	-529.72097080	-332847.55804
An4 dimer (Ci)	-527.91229214	-2.35686758	-529.72097107	-332847.55742
Im1	-284.98023806	-1.23682218	-285.91408322	-179655.59797
Im1 dimer (C2)	-569.98017201	-2.49166072	-571.86558879	-359333.34022
Im1 dimer (Ci)	-569.98017501	-2.49165854	-571.86559512	-359333.33998
Im2	-283.77951520	-1.20623025	-284.68369942	-178877.77958
Im2 dimer (C2h)	-567.57232537	-2.42733072	-569.39618998	-357772.20058
Im3	-283.79316520	-1.20560447	-284.69512538	-178885.20317
Im3 dimer (C2h)	-567.60838937	-2.42644714	-569.42798867	-357792.83229
Im4	-283.78962941	-1.20504254	-284.69168748	-178883.06402
Im4 dimer (C2h)	-567.60051784	-2.42751212	-569.42198472	-357789.05941

Energies of the structures for the redox systems in Figure I-21

cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
1A:FA	-811.86618309	-3.28304168	-814.26210801	-511611.54702
1A-H2:FA	-813.06764830	-3.33770007	-815.50028113	-512398.52756
1A-H2	-624.18805068	-2.60710724	-626.11021733	-393398.67040
1A	-622.99163528	-2.55494417	-624.87824469	-392615.53098
1B:FA	-811.86766202	-3.28384051	-814.26391382	-511612.68232
1B-H2:FA	-813.06516629	-3.33902443	-815.49910362	-512397.76923
1B-H2	-624.18606312	-2.60853819	-626.10959557	-393398.26877
1B	-622.99411115	-2.55558731	-624.88087676	-392617.19196
1C:FA	-811.92203832	-3.28040312	-814.30864401	-511641.15668
1C-H2:FA	-813.06351232	-3.33762954	-815.49598075	-512395.94976
1C-H2	-624.18589954	-2.60621643	-626.10710339	-393396.87422
1C	-623.05162532	-2.55232067	-624.92890475	-392647.76340
1D:FA	-811.87551412	-3.28502902	-814.27160523	-511617.39859
1D-H2:FA	-813.06076823	-3.33566660	-815.49218901	-512393.21265
1D-H2	-624.18001703	-2.60458791	-626.10080148	-393392.40490
1D	-622.99820985	-2.55497765	-624.88400848	-392619.03366
1E:FA	-811.89315024	-3.28007016	-814.28492174	-511625.90779
1E-H2:FA	-813.06008206	-3.33348587	-815.49047977	-512391.98880
1E-H2	-624.18033938	-2.60227867	-626.09994794	-393391.74666
1E	-623.01626462	-2.55018132	-624.89782829	-392627.84807
1F:FA	-811.89282739	-3.27629674	-814.28312268	-511624.99970
1F-H2:FA	-813.06467092	-3.33362457	-815.49351729	-512394.20987
1F-H2	-624.18626634	-2.60202681	-626.10376118	-393394.48897
1F	-623.02223148	-2.54813807	-624.90315164	-392631.44283
1G:FA	-811.88548652	-3.27931548	-814.27722553	-511621.04509
1G-H2:FA	-813.05993994	-3.33451070	-815.49088389	-512392.42573
1G-H2	-624.17981540	-2.60321291	-626.09998608	-393391.93207
1G	-623.00869374	-2.55007572	-624.89057830	-392623.27121
1H:FA	-811.89790024	-3.27411838	-814.28617043	-511626.95300
1H-H2:FA	-813.06333421	-3.33220373	-815.49127114	-512392.83123
1H-H2	-624.18491853	-2.60022297	-626.10131839	-393392.97078
1H	-623.02725482	-2.54620031	-624.90630076	-392633.46452
1I:FA	-811.87068517	-3.28307436	-814.26630154	-511613.98509
1I-H2:FA	-813.09400923	-3.33270698	-815.52179125	-512412.29442
1I-H2	-624.21948707	-2.60281132	-626.13648043	-393415.53727
1I	-622.99210295	-2.55368405	-624.87779564	-392615.06127
1J:FA	-811.88174834	-3.28331868	-814.27685073	-511620.67692
1J-H2:FA	-813.08915021	-3.33166597	-815.51674897	-512408.87356
1J-H2	-624.21437470	-2.60184465	-626.13127286	-393411.99623
1J	-623.00387368	-2.55357474	-624.88882354	-392622.07044
1K:FA	-811.86034320	-3.27874303	-814.25572811	-511607.42453
1K-H2:FA	-813.09039383	-3.32700361	-815.51317712	-512406.75213
1K-H2	-624.21392406	-2.59666223	-626.12580692	-393408.56133
1K	-622.98590120	-2.55039104	-624.87156346	-392611.26240
1L:FA	-811.89783450	-3.27441712	-814.28643577	-511627.05677
1L-H2:FA	-813.09013261	-3.32466252	-815.51038706	-512404.98108
1L-H2	-624.21606270	-2.59405919	-626.12487768	-393408.05868
1L	-623.02725187	-2.54620023	-624.90630240	-392633.46365
1M:FA	-811.87690457	-3.28617995	-814.27424685	-511618.97242
1M-H2:FA	-813.08781999	-3.32789096	-815.51165497	-512405.66000
1M-H2	-624.21149550	-2.59758572	-626.12445732	-393407.59044
1M	-623.00008500	-2.55592516	-624.88684077	-392620.78606
1N:FA	-811.89284947	-3.27640819	-814.28334778	-511625.07486
1N-H2:FA	-813.08996155	-3.32531799	-815.51091723	-512405.21865
1N-H2	-624.21576762	-2.59479615	-626.12537678	-393408.26767

Energies of the structures for the redox systems in Figure I-21 (cont'd)

cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
1N	-623.02223239	-2.54813555	-624.90315343	-392631.44294
1O:FA	-811.88827587	-3.28371053	-814.28081609	-511623.30500
1O-H2:FA	-813.09570462	-3.32832648	-815.51934826	-512410.79753
1O-H2	-624.22002574	-2.59777511	-626.13230898	-393412.99903
1O	-623.01417072	-2.55416321	-624.89600415	-392626.87940
1P:FA	-811.86178466	-3.28370848	-814.26004763	-511610.15931
1P-H2:FA	-813.09077455	-3.32931687	-815.51564804	-512408.40958
1P-H2	-624.21582799	-2.59852817	-626.12912688	-393410.93394
1P	-622.98783130	-2.55293773	-624.87440177	-392613.18684
1Q:FA	-811.87054758	-3.28068758	-814.26565370	-511613.73517
1Q-H2:FA	-813.09107332	-3.32853973	-815.51512634	-512408.14754
1Q-H2	-624.21546135	-2.59748494	-626.12778271	-393410.15815
1Q	-622.99708141	-2.55083601	-624.88113826	-392617.46745
1R:FA	-811.88369610	-3.28106808	-814.27737791	-511621.06669
1R-H2:FA	-813.08534147	-3.32659026	-815.50851314	-512403.73378
1R-H2	-624.20875810	-2.59532507	-626.12017735	-393405.03694
1R	-623.00485806	-2.55114581	-624.88798694	-392621.72765
1S:FA	-811.89170662	-3.28166611	-814.28503066	-511625.94561
1S-H2:FA	-813.08030134	-3.32666885	-815.50378180	-512400.57022
1S-H2	-624.20776139	-2.59557055	-626.11952881	-393404.46038
1S	-623.01971890	-2.55175444	-624.90238768	-392630.83012
1T:FA	-811.86585862	-3.27819422	-814.26000001	-511610.13005
1T-H2:FA	-813.08610715	-3.32657652	-815.50878471	-512404.00158
1T-H2	-624.20975820	-2.59515021	-626.12057318	-393405.38739
1T	-622.99235288	-2.54888686	-624.87580397	-392614.06118
1U:FA	-811.87062985	-3.28744193	-814.26909956	-511615.70658
1U-H2:FA	-813.08315765	-3.32702062	-815.50638431	-512402.38812
1U-H2	-624.21103043	-2.59561218	-626.12227730	-393406.39467
1U	-622.99856791	-2.55734416	-624.88622578	-392620.45735
FA (formic acid)	-188.85623790	-0.72047604	-189.35617046	-118979.60464
H2	-1.13333857	-0.03429902	-1.16486912	-738.13271

Energies of the structures for the flavin case in Figure I-22				
cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
Flavin	-789.08189262	-3.32531923	-791.55931102	-497346.39130
Flavin-H2	-790.25727546	-3.35901174	-792.76088999	-498109.86912
Flavin:FA	-977.95469469	-4.05350101	-980.94140714	-616341.25163
Flavin-H2:FA	-979.12866832	-4.08797460	-982.14241829	-617104.41145
Flavin:FA [2]	-977.95225562	-4.05390936	-980.93961334	-616340.11124
Flavin-H2:FA [2]	-979.13008002	-4.08801201	-982.14383970	-617105.26172
FA (formic acid)	-188.85624878	-0.72046662	-189.35617236	-118979.60668
H2	-1.13333686	-0.03429923	-1.16486966	-738.13210

Gas-phase references for DFT method evaluation in Table II-2				
cpd.	DF-HF/a5Z (a.u)	DF-MP2/aQ5Z (a.u)	CCSD(T)/aDZ (a.u)	CCSD(T)/CBS (kcal/mol)
1 (Cs)	-338.87785831	-1.44343752	-339.94997234	-213605.74002
1 (C1)	-338.87785283	-1.44343571	-339.94997767	-213605.73870
1 dimer (C2h)	-677.77541637	-2.90078920	-679.93541288	-427231.87661
1 dimer (C2)	-677.77541379	-2.90079875	-679.93541066	-427231.87938
1 dimer (Ci)	-677.77540841	-2.90080308	-679.93541027	-427231.87847
1' (C1)	-340.06152715	-1.46687997	-341.16123370	-214370.69122
1' dimer (C2)	-680.13828087	-2.94347183	-682.34992928	-428757.15015
1' dimer (Ci)	-680.13815180	-2.94351255	-682.34976991	-428757.07531
2 (C1)	-356.88346293	-1.57942919	-358.08327050	-224998.19985
2 dimer (C2)	-713.78004220	-3.17304927	-716.19613438	-450012.89394
2 dimer (Ci)	-713.78069548	-3.17209433	-716.19568763	-450012.78829
2' (C1)	-358.05738394	-1.60268953	-359.28613145	-225758.28546
2' dimer (C2)	-716.13067195	-3.21920831	-718.60335976	-451534.33164
2' dimer (Ci)	-716.13041342	-3.21939844	-718.60309077	-451534.23121
BA (benzoic acid) (Cs)	-418.49976816	-1.75484204	-419.80271227	-263773.69211
BA dimer (C2h)	-837.01950532	-3.51749366	-839.63550465	-527565.17058
IM (C1)	-285.00267390	-1.23336713	-285.93159430	-179666.75116
IM dimer (C2)	-570.02053756	-2.47751571	-571.89149807	-359349.81330
IM dimer (Ci)	-570.02052101	-2.47754720	-571.89152938	-359349.82211

APPENDIX F:

Calculated NICS values

NICS(0) and NICS(1)_{zz} values for all species are presented in this appendix. To avoid arbitrary choice of the face of the ring for placing the NICS(1)_{zz} probe in cases where the ring was not completely flat, NICS(1)_{zz} was averaged after calculations of its values on both faces at no extra computational cost. The “NICS(1)_{zz}” on each face is denoted as “NICS(1)_{zz} [1]” and “NICS(1)_{zz} [2]”, and their average is denoted as “NICS(1)_{zz} [avg]”. In fused ring systems, the ring involved in H-bonding was considered as “ring I”, and the other ring was named “ring II”. For the flavin case, with three rings, the farthest ring from the H-bonding ring was named “ring III”. The symmetries of species are shown in parentheses, if applicable. The table is divided into sections based on which figure the named structures belong. For the redox systems, the hydrogenated species are specified using “-H2” at the end of their names and the Formic acid bound compounds are specified by “:FA”.

The code used for extracting NICS values from Gaussian16 output files in the current directory:

```
#!/bin/sh
echo "File NICS(0) NICS(1)zz[1] NICS(1)zz[2] NICS(1)zz[avg]" > nics.csv

for gout in *.log; do

nics0_raw=$(grep "Bq Isotropic =" $gout -A 4 | awk 'NR==1' | awk '{print $8}')
nics1a_raw=$(grep "Bq Isotropic =" $gout -A 4 | awk 'NR==9' | awk '{print $6}')
nics1b_raw=$(grep "Bq Isotropic =" $gout -A 4 | awk 'NR==14' | awk '{print $6}')

nics0=$(echo " $nics0_raw * -1" | bc)
nics1a=$(echo " $nics1a_raw * -1" | bc)
nics1b=$(echo " $nics1b_raw * -1" | bc)
nics1=$(echo "scale=4; $nics1a/2 + $nics1b/2" | bc)

echo "${gout%. *} $nics0 $nics1a $nics1b $nics1" >> nics.csv

done

cat nics.csv | column -t
sed -i 's| |,|g' nics.csv
```

Example NICS input file for Gaussian16:

```
# nmr=giao gen nosymm mpwpw91
```

This is a comment line.

```
0 1
O      2.46555400 -0.00000000 -0.00000000
N      0.36074400  1.09014000 -0.00000000
N      0.36074400 -1.09014000  0.00000000
C     -0.97593600  0.68527000 -0.00000000
C     -0.97593600 -0.68527000 -0.00000000
C      1.23038400 -0.00000000 -0.00000000
H      0.70669400  2.04064000 -0.00000000
H      0.70669400 -2.04064000  0.00000000
H     -1.80551600  1.38528000 -0.00000000
H     -1.80551600 -1.38528000 -0.00000000
bq      0.00000000  0.00000000  0.00000000
bq      0.00000000  0.00000000  1.00000000
bq      0.00000000  0.00000000 -1.00000000
```

```
O H N C 0
6-311++G(3df,3pd)
****
```

NICS values for the four cases of AMHB in Figures I-5 to I-8				
cpd.	NICS(0)	NICS(1)zz [1]	NICS(1)zz [2]	NICS(1)zz [avg]
1a (C2v)	-25.3221	-15.0420	-15.0420	-15.0420
1a dimer (C2h)	-23.2633	-17.0960	-17.0960	-17.0960
1b (Cs)	-27.4987	-12.5100	-12.5100	-12.5100
1b dimer (C2h)	-25.9574	-14.0729	-14.0729	-14.0729
1c (Cs)	-29.1508	-12.0275	-12.0275	-12.0275
1c dimer (C2h)	-27.5345	-13.7494	-13.7494	-13.7494
2a	-23.6321	-12.5416	-12.7829	-12.6623
2a dimer (C2)	-18.4351	-19.1302	-18.6532	-18.8917
2a dimer (Ci)	-18.5625	-19.0828	-18.4517	-18.7673
2b	-24.6175	-11.7602	-10.8795	-11.3199
2b dimer (C2h)	-19.8805	-16.1784	-16.1784	-16.1784
2c	-19.1221	-13.5330	-13.5282	-13.5306
2c dimer (C2h)	-16.9614	-16.1620	-16.1620	-16.1620
3a	-9.6613	-26.5917	-26.8984	-26.7451
3a dimer (C2)	-12.7358	-24.6245	-24.5039	-24.5642
3a dimer (Ci)	-13.2597	-24.1810	-24.2526	-24.2168
3b	-15.2836	-21.4652	-21.1983	-21.3318
3b dimer (C2)	-17.9751	-19.4091	-19.3623	-19.3857
3b dimer (Ci)	-17.9912	-19.3526	-19.4225	-19.3876
3c	-16.0074	-22.0304	-21.5200	-21.7752
3c dimer (C2)	-19.0985	-19.3305	-19.3756	-19.3531
3c dimer (Ci)	-19.0968	-19.3157	-19.4351	-19.3754
4a	-8.3504	-27.2163	-27.2015	-27.2089
4a dimer (C2)	-9.5270	-26.0313	-25.7398	-25.8856
4a dimer (Ci)	-9.9176	-25.6014	-25.7921	-25.6968
4b	-12.2128	-23.2381	-23.2875	-23.2628
4b dimer (C2)	-13.5093	-21.7088	-21.5691	-21.6390
4b dimer (Ci)	-13.5296	-21.5541	-21.7410	-21.6476
4c	-10.2972	-24.7611	-24.5850	-24.6731
4c dimer (C2)	-11.1600	-22.7620	-22.6592	-22.7106
4c dimer (Ci)	-11.1649	-22.6302	-22.8117	-22.7210
5a (Cs)	-28.3943	2.6965	2.6965	2.6965
5a dimer (C2h)	-29.8807	3.8509	3.8509	3.8509
5b (Cs)	-24.5804	6.1100	6.1100	6.1100
5b dimer (C2h)	-27.2502	7.6389	7.6389	7.6389
5c (Cs)	-24.0058	4.1206	4.1206	4.1206
5c dimer (C2h)	-27.5684	6.2059	6.2059	6.2059
6a (Cs)	-28.7793	-3.7599	-3.7599	-3.7599
6a dimer (C2h)	-28.7834	-3.4012	-3.4012	-3.4012
6b (C2v)	-30.8265	-0.6794	-0.6794	-0.6794
6b dimer (C2h)	-31.2348	0.0372	0.0372	0.0372
6c (Cs)	-27.3301	-2.9479	-2.9479	-2.9479
6c dimer (C2h)	-28.1249	-2.2927	-2.2927	-2.2927
7a (Cs)	-34.1613	10.7246	10.7246	10.7246
7a dimer (C2h)	-33.8665	9.2360	9.2360	9.2360
7b (Cs)	-39.0621	15.5995	15.5995	15.5995
7b dimer (C2h)	-37.9094	13.1353	13.1353	13.1353
7c (Cs)	-41.3565	19.6025	19.6025	19.6025
7c dimer (C2h)	-38.4509	15.3109	15.3109	15.3109
8a	-25.1258	-3.6774	-3.5824	-3.6299
8a dimer (C2h)	-25.9216	-3.7261	-3.7261	-3.7261
8b (Cs)	-29.7228	0.5930	0.5930	0.5930
8b dimer (C2h)	-30.0552	-0.1696	-0.1696	-0.1696
8c (Cs)	-29.1740	-1.2013	-1.2013	-1.2013
8c dimer (C2h)	-27.7666	-1.9457	-1.9457	-1.9457

NICS values for the case study in Figure I-10				
cpd.	NICS(0)	NICS(1)zz [1]	NICS(1)zz [2]	NICS(1)zz [avg]
9 (Cs) [ring I]	-16.4712	-22.6335	-22.6335	-22.6335
9 dimer (C2h) [ring I]	-15.4878	-23.4836	-23.4836	-23.4836
9 (Cs) [ring II]	-7.7871	-30.3510	-30.3510	-30.3510
9 dimer (C2h) [ring II]	-10.2526	-27.9754	-27.9754	-27.9754
10	-29.6438	-8.9333	-9.5831	-9.2582
10 dimer (C2)	-27.0103	-11.5620	-12.6744	-12.1182
10 dimer (Ci)	-27.0832	-11.5082	-12.6336	-12.0709
11	-13.8733	-23.8597	-23.0941	-23.4769
11 dimer (C2)	-15.6099	-21.8891	-22.0145	-21.9518
11 dimer (Ci)	-16.0699	-21.4433	-21.5664	-21.5049

NICS values for the fused rings in Figure I-12				
cpd.	NICS(0)	NICS(1)zz [1]	NICS(1)zz [2]	NICS(1)zz [avg]
13 [ring I]	-9.2153	-14.9718	-14.9718	-14.9718
13 dimer [ring I]	-6.7532	-18.3134	-18.3134	-18.3134
13 [ring II]	-5.3612	-14.0180	-14.0180	-14.0180
13 dimer [ring II]	-2.6724	-18.0441	-18.0441	-18.0440
13'	-13.5527	-8.1439	-8.7720	-8.4579
13' dimer	-11.8505	-11.3145	-10.6261	-10.9702
13''	-11.5286	-8.3448	-8.3394	-8.3421
13'' dimer	-9.6177	-10.9856	-10.9710	-10.9783
14 [ring I]	-12.8723	-8.3881	-8.3881	-8.3880
14 dimer [ring I]	-11.1953	-11.1508	-11.1508	-11.1508
14 [ring II]	-5.9773	-26.2968	-26.2968	-26.2968
14 dimer [ring II]	-5.6686	-26.2369	-26.2369	-26.2368
14'	-13.4711	-8.8383	-9.0845	-8.9613
14' dimer	-11.9102	-11.5195	-11.2612	-11.3903
14''	-12.6342	-9.3119	-9.6620	-9.4869
14'' dimer	-10.7572	-11.7407	-12.1514	-11.9460
15 [ring I]	-14.7782	-8.1397	-8.1397	-8.1396
15 dimer [ring I]	-12.5067	-10.9006	-10.9006	-10.9006
15 [ring II]	-3.9241	-24.7713	-24.7713	-24.7712
15 dimer [ring II]	-4.3776	-25.1814	-25.1814	-25.1814
15'	-12.8530	-8.5305	-9.1794	-8.8549
15' dimer	-10.5383	-11.1043	-11.6962	-11.4002
15''	-11.0463	-8.7719	-8.8522	-8.8120
15'' dimer	-9.1417	-11.4900	-11.2467	-11.3683

NICS values for the redox systems in Figure I-20				
cpd.	NICS(0)	NICS(1)zz [1]	NICS(1)zz [2]	NICS(1)zz [avg]
1A [ring I]	-22.0703	5.5648	5.5648	5.5648
1A [ring II]	-16.1816	-3.7078	-3.7078	-3.7078
1A:FA [ring I]	-23.7034	6.3103	6.3103	6.3102
1A:FA [ring II]	-16.6617	-3.0683	-3.0683	-3.0682
1A-H2 [ring I]	-8.8965	-17.8766	-17.8766	-17.8766
1A-H2 [ring II]	-7.9184	-13.4616	-13.4616	-13.4616
1A-H2:FA [ring I]	-7.1452	-19.9900	-19.9900	-19.9900
1A-H2:FA [ring II]	-6.1670	-16.3935	-16.3935	-16.3934
1B [ring I]	-20.6435	3.3945	3.3945	3.3944
1B [ring II]	-17.9179	-0.0867	-0.0867	-0.0866
1B:FA [ring I]	-21.7974	3.6018	3.6018	3.6018
1B:FA [ring II]	-18.5429	0.4900	0.4900	0.4900
1B-H2 [ring I]	-7.4796	-18.4821	-18.4821	-18.4820
1B-H2 [ring II]	-7.3826	-13.8950	-13.8950	-13.8950
1B-H2:FA [ring I]	-5.7661	-20.6838	-20.6838	-20.6838
1B-H2:FA [ring II]	-5.8210	-16.6205	-16.6205	-16.6204
1C [ring I]	-25.5448	5.5904	5.5904	5.5904
1C [ring II]	-3.1417	-24.1280	-24.1280	-24.1280
1C:FA [ring I]	-27.1939	6.1756	6.1756	6.1756
1C:FA [ring II]	-3.0388	-23.9200	-23.9200	-23.9200
1C-H2 [ring I]	-9.6202	-16.9714	-16.9714	-16.9714
1C-H2 [ring II]	-6.4896	-12.8757	-12.8757	-12.8756
1C-H2:FA [ring I]	-8.0501	-18.6321	-18.6321	-18.6320
1C-H2:FA [ring II]	-3.9654	-16.8764	-16.8764	-16.8764
1D [ring I]	-13.1462	-6.7539	-6.7539	-6.7538
1D [ring II]	-16.8412	6.6499	6.6499	6.6498
1D:FA [ring I]	-11.1770	-9.1046	-9.1046	-9.1046
1D:FA [ring II]	-17.7483	7.9208	7.9208	7.9208
1D-H2 [ring I]	-11.1856	-13.2577	-13.2577	-13.2576
1D-H2 [ring II]	-6.7412	-13.1671	-13.1671	-13.1670
1D-H2:FA [ring I]	-8.7869	-16.6993	-16.6993	-16.6992
1D-H2:FA [ring II]	-4.8956	-16.5935	-16.5935	-16.5934
1E [ring I]	-13.3780	-8.2297	-8.2297	-8.2296
1E [ring II]	-27.0482	3.7772	3.7772	3.7772
1E:FA [ring I]	-11.3804	-10.9178	-10.9178	-10.9178
1E:FA [ring II]	-27.8508	4.4628	4.4628	4.4628
1E-H2 [ring I]	-9.9742	-13.7492	-13.7492	-13.7492
1E-H2 [ring II]	-9.3637	-11.1163	-11.1163	-11.1162
1E-H2:FA [ring I]	-7.5019	-17.3543	-17.3543	-17.3542
1E-H2:FA [ring II]	-7.6035	-14.3069	-14.3069	-14.3068
1F [ring I]	-18.8506	1.0033	1.0033	1.0032
1F [ring II]	-20.9527	3.7096	3.7096	3.7096
1F:FA [ring I]	-19.1998	1.4858	1.4858	1.4858
1F:FA [ring II]	-21.4346	4.2236	4.2236	4.2236
1F-H2 [ring I]	-12.7409	-11.4262	-11.4262	-11.4262
1F-H2 [ring II]	-8.8304	-13.0186	-13.0186	-13.0186
1F-H2:FA [ring I]	-10.8469	-14.0190	-14.0190	-14.0190
1F-H2:FA [ring II]	-6.8437	-16.1937	-16.1937	-16.1936
1G [ring I]	-15.3350	-5.9111	-5.9111	-5.9110
1G [ring II]	-17.8488	4.7897	4.7897	4.7896
1G:FA [ring I]	-13.6633	-8.1188	-8.1188	-8.1188

NICS values for the redox systems in Figure I-20 (cont'd)				
cpd.	NICS(0)	NICS(1)zz [1]	NICS(1)zz [2]	NICS(1)zz [avg]
1G:FA [ring II]	-18.2541	5.8155	5.8155	5.8154
1G-H2 [ring I]	-8.4441	-14.6484	-14.6484	-14.6484
1G-H2 [ring II]	-7.7247	-12.5234	-12.5234	-12.5234
1G-H2:FA [ring I]	-5.9641	-18.0909	-18.0910	-18.0909
1G-H2:FA [ring II]	-5.5475	-16.0817	-16.0817	-16.0816
1H [ring I]	-19.9265	2.0932	2.0932	2.0932
1H [ring II]	-19.6795	-1.8054	-1.8054	-1.8054
1H:FA [ring I]	-20.3538	2.6493	2.6493	2.6492
1H:FA [ring II]	-19.8561	-1.4298	-1.4298	-1.4298
1H-H2 [ring I]	-11.0673	-12.3981	-12.3981	-12.3980
1H-H2 [ring II]	-9.6685	-11.9355	-11.9355	-11.9354
1H-H2:FA [ring I]	-8.9993	-15.0172	-15.0172	-15.0172
1H-H2:FA [ring II]	-7.2129	-15.5246	-15.5246	-15.5246
1I [ring I]	-12.2112	-6.1004	-6.1004	-6.1004
1I [ring II]	-17.8875	7.5817	7.5817	7.5816
1I:FA [ring I]	-10.9809	-8.6723	-8.6723	-8.6722
1I:FA [ring II]	-18.7629	7.8634	7.8634	7.8634
1I-H2 [ring I]	-12.8923	-8.8707	-8.8707	-8.8706
1I-H2 [ring II]	-1.6646	-23.1691	-23.1691	-23.1690
1I-H2:FA [ring I]	-10.6191	-11.7324	-11.7324	-11.7324
1I-H2:FA [ring II]	-1.5661	-23.6720	-23.6720	-23.6720
1J [ring I]	-11.0223	-7.5254	-7.5254	-7.5254
1J [ring II]	-28.9323	5.3570	5.3570	5.3570
1J:FA [ring I]	-9.2311	-10.5925	-10.5925	-10.5924
1J:FA [ring II]	-29.0645	5.5231	5.5231	5.5230
1J-H2 [ring I]	-13.9942	-8.1776	-8.1776	-8.1776
1J-H2 [ring II]	-2.5190	-21.7100	-21.7100	-21.7100
1J-H2:FA [ring I]	-11.8233	-11.0977	-11.0977	-11.0976
1J-H2:FA [ring II]	-2.4726	-22.1888	-22.1888	-22.1888
1K [ring I]	-21.1834	5.9895	5.9895	5.9894
1K [ring II]	-20.3840	11.5024	11.5024	11.5024
1K:FA [ring I]	-22.0148	7.3009	7.3009	7.3008
1K:FA [ring II]	-21.1735	13.1282	13.1282	13.1282
1K-H2 [ring I]	-13.4269	-8.6351	-8.6351	-8.6350
1K-H2 [ring II]	-4.2612	-22.1850	-22.1850	-22.1850
1K-H2:FA [ring I]	-11.1109	-11.8135	-11.8135	-11.8134
1K-H2:FA [ring II]	-4.0717	-22.4913	-22.4913	-22.4912
1L [ring I]	-19.9269	2.0931	2.0931	2.0930
1L [ring II]	-19.6796	-1.8044	-1.8044	-1.8044
1L:FA [ring I]	-20.7374	2.4038	2.4038	2.4038
1L:FA [ring II]	-19.7758	-1.6755	-1.6755	-1.6754
1L-H2 [ring I]	-13.3535	-7.7658	-7.7658	-7.7658
1L-H2 [ring II]	-1.6156	-22.7282	-22.7282	-22.7282
1L-H2:FA [ring I]	-11.4425	-10.3484	-10.3484	-10.3484
1L-H2:FA [ring II]	-1.7522	-22.7716	-22.7716	-22.7716
1M [ring I]	-12.0636	-6.8052	-6.8052	-6.8052
1M [ring II]	-20.0076	11.4940	11.4940	11.4940
1M:FA [ring I]	-10.1110	-9.4558	-9.4558	-9.4558
1M:FA [ring II]	-19.7339	11.5565	11.5565	11.5564
1M-H2 [ring I]	-14.8548	-7.6612	-7.6612	-7.6612
1M-H2 [ring II]	-1.4256	-22.7646	-22.7646	-22.7646

NICS values for the redox systems in Figure I-20 (cont'd)				
cpd.	NICS(0)	NICS(1)zz [1]	NICS(1)zz [2]	NICS(1)zz [avg]
1M-H2:FA [ring I]	-12.7185	-10.8524	-10.8524	-10.8524
1M-H2:FA [ring II]	-1.0153	-23.1836	-23.1836	-23.1836
1N [ring I]	-18.8508	1.0036	1.0036	1.0036
1N [ring II]	-20.9520	3.7096	3.7096	3.7096
1N:FA [ring I]	-19.6608	1.1813	1.1813	1.1812
1N:FA [ring II]	-21.2115	3.6285	3.6285	3.6284
1N-H2 [ring I]	-14.6864	-7.0398	-7.0398	-7.0398
1N-H2 [ring II]	-1.1963	-22.7900	-22.7900	-22.7900
1N-H2:FA [ring I]	-12.9266	-9.6538	-9.6538	-9.6538
1N-H2:FA [ring II]	-1.3853	-22.8921	-22.8921	-22.8920
1O [ring I]	-17.3325	9.5119	9.5119	9.5118
1O [ring II]	-4.0831	-18.1754	-18.1754	-18.1754
1O:FA [ring I]	-19.3306	10.9806	10.9806	10.9806
1O:FA [ring II]	-4.5099	-17.7130	-17.7130	-17.7130
1O-H2 [ring I]	-15.1882	-9.8411	-9.8411	-9.8410
1O-H2 [ring II]	-4.2141	-25.5487	-25.5487	-25.5486
1O-H2:FA [ring I]	-12.9059	-12.2358	-12.2358	-12.2358
1O-H2:FA [ring II]	-4.7560	-25.9800	-25.9800	-25.9800
1P [ring I]	-22.2825	1.5260	1.5260	1.5260
1P [ring II]	-24.2215	11.1064	11.1064	11.1064
1P:FA [ring I]	-23.3315	1.4525	1.4525	1.4524
1P:FA [ring II]	-24.9905	11.7146	11.7146	11.7146
1P-H2 [ring I]	-15.6243	-10.2321	-10.2321	-10.2320
1P-H2 [ring II]	-0.6553	-22.7917	-22.7917	-22.7916
1P-H2:FA [ring I]	-13.4305	-12.7961	-12.7961	-12.7960
1P-H2:FA [ring II]	-1.0890	-23.3105	-23.3105	-23.3104
1Q [ring I]	-23.0745	2.8150	2.8150	2.8150
1Q [ring II]	-23.1363	2.8795	2.8795	2.8794
1Q:FA [ring I]	-24.2501	3.0149	3.0149	3.0148
1Q:FA [ring II]	-23.7070	3.5343	3.5343	3.5342
1Q-H2 [ring I]	-15.9113	-9.9538	-9.9538	-9.9538
1Q-H2 [ring II]	-1.6881	-22.7084	-22.7084	-22.7084
1Q-H2:FA [ring I]	-13.7595	-12.5012	-12.5012	-12.5012
1Q-H2:FA [ring II]	-1.6402	-23.2339	-23.2339	-23.2338
1R [ring I]	-15.9078	-6.3335	-6.3335	-6.3334
1R [ring II]	-20.1806	10.3407	10.3407	10.3406
1R:FA [ring I]	-14.0969	-8.6455	-8.6455	-8.6454
1R:FA [ring II]	-19.9524	10.5117	10.5117	10.5116
1R-H2 [ring I]	-14.2084	-7.7352	-7.7352	-7.7352
1R-H2 [ring II]	-2.4202	-21.0146	-21.0146	-21.0146
1R-H2:FA [ring I]	-11.7689	-10.8282	-10.8282	-10.8282
1R-H2:FA [ring II]	-1.6284	-21.9963	-21.9963	-21.9962
1S [ring I]	-15.4667	-7.3744	-7.3744	-7.3744
1S [ring II]	-30.7135	7.9965	7.9965	7.9964
1S:FA [ring I]	-13.3240	-9.8601	-9.8601	-9.8600
1S:FA [ring II]	-30.1211	7.2869	7.2869	7.2868
1S-H2 [ring I]	-15.0052	-7.7062	-7.7062	-7.7062
1S-H2 [ring II]	-3.4372	-20.1983	-20.1983	-20.1982
1S-H2:FA [ring I]	-12.5932	-10.4790	-10.4790	-10.4790
1S-H2:FA [ring II]	-3.0360	-20.8043	-20.8043	-20.8042
1T [ring I]	-23.2912	3.7265	3.7265	3.7264

NICS values for the redox systems in Figure I-20 (cont'd)				
cpd.	NICS(0)	NICS(1)zz [1]	NICS(1)zz [2]	NICS(1)zz [avg]
1T [ring II]	-21.3442	10.5248	10.5248	10.5248
1T:FA [ring I]	-23.5063	4.7972	4.7972	4.7972
1T:FA [ring II]	-21.9598	12.0934	12.0934	12.0934
1T-H2 [ring I]	-14.1931	-7.8732	-7.8732	-7.8732
1T-H2 [ring II]	-5.2989	-20.6548	-20.6548	-20.6548
1T-H2:FA [ring I]	-11.7794	-10.9700	-10.9700	-10.9700
1T-H2:FA [ring II]	-5.0252	-21.4612	-21.4612	-21.4612
1U [ring I]	-14.6400	-6.5957	-6.5957	-6.5956
1U [ring II]	-22.6971	15.6840	15.6840	15.6840
1U:FA [ring I]	-12.7615	-8.9704	-8.9704	-8.9704
1U:FA [ring II]	-21.6600	13.9676	13.9676	13.9676
1U-H2 [ring I]	-16.0243	-7.0817	-7.0817	-7.0816
1U-H2 [ring II]	-2.4025	-21.7625	-21.7625	-21.7624
1U-H2:FA [ring I]	-13.5768	-9.7055	-9.7055	-9.7054
1U-H2:FA [ring II]	-2.1362	-22.4626	-22.4626	-22.4626

NICS values for the flavin case in Figure I-22				
cpd.	NICS(0)	NICS(1)zz [1]	NICS(1)zz [2]	NICS(1)zz [avg]
Flavin [ring 1]	-18.2123	-7.2093	-7.2093	-7.2092
Flavin [ring 2]	-21.7554	0.7828	0.7828	0.7828
Flavin [ring 3]	-3.3670	-23.5947	-23.5947	-23.5946
Flavin:FA [ring 1]	-16.8633	-8.7288	-8.7288	-8.7288
Flavin:FA [ring 2]	-21.0450	0.7086	0.7086	0.7086
Flavin:FA [ring 3]	-3.9667	-24.0879	-24.0879	-24.0878
Flavin:FA [2] [ring 1]	-18.3505	-6.8818	-6.8818	-6.8818
Flavin:FA [2] [ring 2]	-22.8244	0.4949	0.4949	0.4948
Flavin:FA [2] [ring 3]	-2.8266	-23.2572	-23.2572	-23.2572
Flavin-H2 [ring 1]	-34.5131	12.5943	11.5510	12.0726
Flavin-H2 [ring 2]	-18.6058	-1.8251	-0.3350	-1.0800
Flavin-H2 [ring 3]	-2.8642	-20.3206	-18.4402	-19.3804
Flavin-H2:FA [ring 1]	-34.5565	11.4364	13.2445	12.3404
Flavin-H2:FA [ring 2]	-18.6950	-0.6893	-2.2535	-1.4713
Flavin-H2:FA [ring 3]	-2.9742	-18.9620	-20.6499	-19.8059
Flavin-H2:FA [2] [ring 1]	-34.7277	13.3280	11.6260	12.4770
Flavin-H2:FA [2] [ring 2]	-18.6201	-1.8217	-0.3801	-1.1008
Flavin-H2:FA [2] [ring 3]	-2.8129	-20.5851	-18.9837	-19.7843

APPENDIX G:

The Python code for insertion of NICS probes

```

import numpy, math, argparse, ast, sys
import pybel, rdkit
from rdkit import Chem
from pprint import pprint
from scipy.optimize import curve_fit
from subprocess import Popen, PIPE, call

```

```
'''
```

How to run this script:

make sure there is Gaussian input file names "3c.com" exists in the directory and type:

```
python nics.py -I 3c.com -R "1 2 3 4 5 6"
```

```
'''
```

```

# parse arguments
parser = argparse.ArgumentParser()
parser.add_argument("-I", "--input", help="input file name: it can be either Gaussian16
input (*.com) or output (.log)")
parser.add_argument("-R", "--rings", help="ring element indices starting from 1. Also,
makes sure single quotes are included. e.g: '1 3 6 7 9'")
parser.add_argument("-M", "--method", help="method: e.g mpwpw91")
parser.add_argument("-B", "--basis", help="basis set: e.g 6-311+G(2d,p)")
parser.add_argument("-C", "--charge", help="charge : e.g -1")
parser.add_argument("-S", "--spin", help="basis set: e.g 2")
args = parser.parse_args()

input = args.input
rings = args.rings
method = args.method
basis = args.basis
charge = args.charge
spin = args.spin

# set default value for level of theory if not specified.
if method == None:
    method = 'mpwpw91'

if basis == None:
    basis = '6-311++G(3df,3pd)'
if charge == None:
    charge = '0'
if spin == None:
    spin = '1'

```

```

gaussian_card = '# nmr=giao gen nosymm ' + method
charge_and_multiplicity = charge + ' ' + spin

# dictionary of atoms
atomlist = {'H': 1, 'He': 2, \
            'Li': 3, 'Be': 4, 'B': 5, 'C': 6, 'N': 7, 'O': 8, 'F': 9, 'Ne': 10, \
            'Na': 11, 'Mg': 12, 'Al': 13, 'Si': 14, 'P': 15, 'S': 16, 'Cl': 17, 'Ar': 18, \
            'K': 19, 'Ca': 20, 'Ga': 31, 'Ge': 32, 'As': 33, 'Se': 34, 'Br': 35, 'Kr': 36, \
            'Rb': 37, 'Sr': 38, 'In': 49, 'Sn': 50, 'Sb': 51, 'Te': 52, 'I': 53, 'Xe': 54, \
            'Cs': 55, 'Ba': 56, 'Tl': 81, 'Pb': 82, 'Bi': 83, 'Po': 84, 'At': 85, 'Rn': 86}

#####
##### DEFINING FUNCTIONS #####
#####
# functions for getting geometries:
def getGeometryGeneral(input):
    """
    this function tries to distinguish between a gaussian input and output and extract
    geometries from them.
    """
    if input.endswith('.log'):
        result = get_lines(input, 'Standard orientation', 'Rotational constants
(GHZ)', 5, -1, '1 3 4 5')
    elif input.endswith('.com'):
        result = getGINgeometry(input)
    elif input.endswith('.mol2'):
        result = None
    elif input.endswith('.xyz'):
        result = None
    elif input.endswith('.mol'):
        result = None
    elif input.endswith('.sdf'):
        result = None
    elif input.endswith('.pdb'):
        result = None
    elif input.endswith('.xml'):
        result = None
    return result

def getGINgeometry(input):
    """
    This function extracts the xyz coordinates from a xyz format Gaussian16 input file.
    it might work with Gaussian09 and Gaussian03 input files, too. At this point, the
    extension of the file does matter.
    """
    # open the file read it in and close it.

```

```

with open(input) as f:
    gin = f.readlines()
# split each line into list to make a nested list.
lines = [i.split() for i in gin]
# convert the numbers in the list to floats if possible.
final = []
for i in range(len(lines)):
    final0 = []
    for j in range(len(lines[i])):
        b = lines[i][j]
        try:
            a = float(b)
        except ValueError:
            # print("error")
            a = b
        # print(type(a))
        final0.append(a)
    final.append(final0)
# check the lines tha have atomic coordinates and add them to the final list.
lines = final
geometry = []
for line in lines:
    if len(line) == 4 and type(line[1]) == float and type(line[2]) == float and
type(line[3]) == float:
        geometry.append(line)
return geometry

```

```

def
get_lines(g16,start_phrase,end_phrase,lines_after_start,lines_before_end,columns_matrix):
'''

```

This function gets the line between two phrases and extracts the requested columns between them.

Also, lines before and after the end and start phrase can be eliminated.

start_phrase is where the point of interest begins.

lines_after_atart is the number of lines to be eliminated after the start phrase.

end phrase is the point where marks the point after which we are not interested in.

lines_before_end is the number of lines that need to be eliminated before the end phrase

columns_matrix is a string of integers separated by space which show the columns of interest starting from 0.

```
'''
```

```
# Open the file and read it as a list by each line
```

```
with open(g16) as f:
```

```
    gout = f.readlines()
```

```
# Find the lines that have the geometries in between
```

```
start_phrases = []
```

```

for i, elem in enumerate(gout):
    if start_phrase in elem:
        start_phrases.append(i)
end_phrases = []
for i, elem in enumerate(gout):
    if end_phrase in elem:
        end_phrases.append(i)
# Get the last geometry in the file.
start_line = start_phrases[-1] + lines_after_start
end_line = end_phrases[-2] + lines_before_end
# Sometime the frequencies cause confusion so correct for that.
if (start_line > end_line):
    end_line = end_phrases[-1] + lines_before_end
# Split each atom's properties and delete useless columns from the 2D list.
lines = gout[start_line:end_line]
lines = [i.split() for i in lines]
columns_matrix0 = columns_matrix.split()
columns = [int(i) for i in columns_matrix0]
rows = list(range(0,(len(lines))))
final = []
for row in rows:
    final0 = []
    for column in columns:
        b = lines[row][column]
        a = ast.literal_eval(b)
        final0.append(a)
    final.append(final0)
#pprint(final)
return final

# functions for finding mean plane:
def SortByAtomicNumber(unsorted):
    """
    This function sort a list of atoms by their atomic numbers.
    """
    for i, atom in enumerate(unsorted):
        a = unsorted[i][0]
        b = atomlist.get(a)
        atom.insert(0,b)
    sorted1 = sorted(unsorted, key=lambda x: x[0], reverse=True)

    for i, atom in enumerate(unsorted):
        atom.pop(0)
    # pprint2D(sorted1,False)
    return sorted1

```

```

def getAxis(geometry,column):
    """
    this function extracts a column of interest from a 2D nested list.
    """
    if column == 'X' or column == 'x':
        column = 1
    elif column == 'Y' or column == 'y':
        column = 2
    elif column == 'Z' or column == 'z':
        column = 3
    axis = []
    for i, row in enumerate(geometry):
        element = geometry[i][column]
        axis.append(element)
        # print(axis)
    return axis

def getRings(geometry):
    """
    gets the ring elements matrix using rdkit. both open Babel and rdkit need to be
    installed.
    """
    # write an xyz file of the original coordinates
    orig_stdout = sys.stdout
    number = len(geometry)
    sys.stdout = open('temp.xyz','wt')
    print(number)
    print("This is a comment line.")
    print2D(geometry,False)
    sys.stdout.close()
    sys.stdout=orig_stdout
    # convert the xyz file into a mol file to be opened by rdkit
    call(['obabel', 'temp.xyz', '-O', 'temp.mol'], stdout=PIPE, stderr=PIPE)
    # Popen(['obabel', '1.xyz', '-O', '1.mol'], stdout=PIPE, stderr=PIPE)
    m = Chem.MolFromMolFile('temp.mol')
    ssr = Chem.GetSymmSSSR(m)
    call(['rm', 'temp.xyz', 'temp.mol'])
    return ssr

def getRingAxis(centered_geometry,ring_elements,column):
    """
    This function extracts partial columns of interest from a 2D nested list of geometry.
    """
    if column == 'X' or column == 'x':
        column = 1
    elif column == 'Y' or column == 'y':

```

```

        column = 2
    elif column == 'Z' or column == 'z':
        column = 3
    ring_elements = ring_elements.split()
    ring_elements = [int(i) for i in ring_elements]
    ring_elements = [i - 1 for i in ring_elements]
    # print(ring_elements)
    axis = getAxis(centered_geometry,column)

    ring_axis = []

    for row in ring_elements:
        element = centered_geometry[row][column]
        ring_axis.append(element)
    return ring_axis

def centerRing(geometry,ring_elements,x_column,y_column,z_column):
    ring_elements = ring_elements.split()
    ring_elements = [int(i) for i in ring_elements]
    ring_elements = [i - 1 for i in ring_elements]
    # print(ring_elements)
    x = getAxis(geometry,x_column)
    y = getAxis(geometry,y_column)
    z = getAxis(geometry,z_column)
    x_ring = []
    y_ring = []
    z_ring = []
    for row in ring_elements:
        element = geometry[row][1]
        x_ring.append(element)
    for row in ring_elements:
        element = geometry[row][2]
        y_ring.append(element)
    for row in ring_elements:
        element = geometry[row][3]
        z_ring.append(element)

    x_center = sum(x_ring)/len(x_ring)
    y_center = sum(y_ring)/len(y_ring)
    z_center = sum(z_ring)/len(z_ring)
    x_centered = [round(element - x_center,10) for element in x]
    y_centered = [round(element - y_center,10) for element in y]
    z_centered = [round(element - z_center,10) for element in z]
    centered_geometry = []
    for i,element in enumerate(x_centered):
        atom = []

```

```

        e = geometry[i][0]
        atom.append(e)
        x = x_centered[i]
        atom.append(x)
        y = y_centered[i]
        atom.append(y)
        z = z_centered[i]
        atom.append(z)
        centered_geometry.append(atom)
    return centered_geometry

```

```
def findMeanPlane(geometry,ring_elements):
```

```

'''
    Given a 2D nested list of xyz coordinates (geometry) and a string containing the
    indices of the ring elements
    in th form of 'int int int int ...' (ring_elements, a string separated by spaces) finds the
    a, b, and c of a plane.

```

```

    The geometry does not need to be centered
'''

```

```

# center the geometry and get the ring atom coordinates
centered_geometry = centerRing(geometry,ring_elements,'x','y','z')

```

```

x_ring = getRingAxis(centered_geometry,ring_elements,'x')

```

```

y_ring = getRingAxis(centered_geometry,ring_elements,'y')

```

```

z_ring = getRingAxis(centered_geometry,ring_elements,'z')

```

```

# find the best fit

```

```

def function(big_x,a,b):
'''

```

```

    this is a function for ring plane.
'''

```

```

'''

```

```

    x,y = big_x

```

```

    return (a*x) + (b*y)
'''

```

```

p0 = 1, 1

```

```

best_vals, covar = curve_fit(function, (x_ring,y_ring), z_ring, p0)

```

```

[a0,b0,c0] = [best_vals[0], best_vals[1], -1]

```

```

# convert the mean plane vector to a unit vector

```

```

length = (a0**2 + b0**2 + c0**2)**(0.5)

```

```

[a,b,c] = [-a0/length, -b0/length, -c0/length]

```

```

return (a,b,c)

```

```
def insertNICS(geometry,ring_elements):
```

```

'''
    This function takes a 2D nested list of atomic coordinates, center them to the
    ring_elements which is

```

```

    a string of integers separated by spaces, finds the mean plane, and rotates the
    molecule so the vector of

```

the mean plane is in the Z direction. It also add both NICS(1)zz and NICS(0) probes
'''

```
atoms = getAxis(geometry,0) # get atoms column
```

'''

This section was added later to fix the cases that have their rings already completely flat and in the XY plane.

What it does, it only rotates the molecule to a random angle around the Y axis. In case problems raise, the value of the random variable can be changed.

'''

```
RANDOM_PHI = 1.25
xx0 = getAxis(geometry,'x')
yy0 = getAxis(geometry,'y')
zz0 = getAxis(geometry,'z')
[xx01,yy01,zz01] = PhiRot(xx0,yy0,zz0,RANDOM_PHI) # rotate the molecule
around Y by A RANDOM phi
randomly_rotated_geometry = [atoms,xx01,yy01,zz01] #
randomly_rotated_geometry
randomly_rotated_geometry = [list(i) for i in zip(*randomly_rotated_geometry)]
```

```
# End of the patch up for random rotation of the system. -Tayeb on April 5, 2019
```

```
centered_geometry =
centerRing(randomly_rotated_geometry,ring_elements,'x','y','z') # center the ring
```

```
[a,b,c] = findMeanPlane(centered_geometry,ring_elements)
[rho,tetha,phi] = car2sph(a,b,c) # convert the mean plane from cartesian to
spherical coordinates
```

```
xx = getAxis(centered_geometry,'x')
yy = getAxis(centered_geometry,'y')
zz = getAxis(centered_geometry,'z')
centered = [atoms,xx,yy,zz]
#print2D(centered,True)
##### ROTATE AROUND Z BY -THETA #####
[xx1,yy1,zz1] = TethaRot(xx,yy,zz,tetha) # rotate the molecule around Z by -tetha
[aa1,bb1,cc1] = TethaRot(a,b,c,tetha) # rotate the plane vector around Z by -tetha
[rho1,tetha1,phi1] = car2sph(aa1,bb1,cc1) # convert again the plane vector into
spherical coordinates
# print('rotated by tetha around z rho => tetha phi = ', rho1, tetha1, phi1)
# tetha_around_z = [atoms,xx1,yy1,zz1]
# print2D(tetha_around_z, True)
##### ROTATE AROUND Y BY -PHI #####
[xx2,yy2,zz2] = PhiRot(xx1,yy1,zz1,-phi1) # rotate the molecule around Y by phi
```

```

[aa2,bb2,cc2] = PhiRot(aa1,bb1,cc1,-phi1) # rotate the plane vector around Y by phi
# print('rotated by tetha around y by phi => rho tetha phi = ', rho1, tetha1, phi1)
# phi_around_y = [atoms,xx2,yy2,zz2]
# print2D(phi_around_y, True)
##### ROTATE AROUND Y BY PHI #####
[xx3,yy3,zz3] = PhiRot(xx1,yy1,zz1,phi1) # rotate the molecule around Y by phi
[aa3,bb3,cc3] = PhiRot(aa1,bb1,cc1,phi1) # rotate the plane vector around Y by phi
# print('rotated by tetha around y by -phi => rho tetha phi = ', rho1, tetha1, phi1)
# minus_phi_around_y = [atoms,xx3,yy3,zz3]
# print2D(minus_phi_around_y, True)
# find which courdinates to use
if abs(cc2) > abs(cc3):
    [xx_final,yy_final,zz_final] = [xx2,yy2,zz2]
else:
    [xx_final,yy_final,zz_final] = [xx3,yy3,zz3]
xx_final = [ '%.8f' % elem for elem in xx_final ]
yy_final = [ '%.8f' % elem for elem in yy_final ]
zz_final = [ '%.8f' % elem for elem in zz_final ]
xx_final = numpy.array(xx_final).tolist()
yy_final = numpy.array(yy_final).tolist()
zz_final = numpy.array(zz_final).tolist()
# format the columns and get them ready for printing.
xx_final = [i.rjust(14,' ') for i in xx_final]
yy_final = [i.rjust(14,' ') for i in yy_final]
zz_final = [i.rjust(14,' ') for i in zz_final]
# add the nics probes and transpose the matrix
nics_bq = [['bq',' 0.00000000',' 0.00000000',' 0.00000000'],\
['bq',' 0.00000000',' 0.00000000',' 1.00000000'],\
['bq',' 0.00000000',' 0.00000000',' -1.00000000']]
nics = [atoms,xx_final,yy_final,zz_final]
nics = [list(i) for i in zip(*nics)]
nics.append(nics_bq[0])
nics.append(nics_bq[1])
nics.append(nics_bq[2])
return nics

```

functions for manipulating the coordinates:

```

def car2sph(x,y,z):
    """
    This function converts Cartesian coordinates of a XYZ point to spherical coordinates.
    """
    rho = (x**2+y**2+z**2)**(0.5)
    tetha = -math.atan(y/x)
    phi = -math.atan(((x**2+y**2)**0.5)/z)
    return (rho, tetha, phi)

```

```

def TethaRot(x,y,z,tetha):
    """
    This function rotates coordinates of a single point or a set of points around the Z-
    AXIS by theta.
    """
    x0 = numpy.array(x)
    y0 = numpy.array(y)
    z0 = numpy.array(z)
    x1 = (x0*math.cos(tetha))-(y0*math.sin(tetha))
    y1 = (x0*math.sin(tetha))+(y0*math.cos(tetha))
    z1 = z0
    return (x1,y1,z1)

def PhiRot(x,y,z,phi):
    """
    This function rotates coordinates of a single point or a set of points around the Y-
    AXIS by phi.
    """
    x0 = numpy.array(x)
    y0 = numpy.array(y)
    z0 = numpy.array(z)
    x1 = (x0*math.cos(phi))+(z0*math.sin(phi))
    y1 = y0
    z1 = (z0*math.cos(phi))-(x0*math.sin(phi))
    return (x1,y1,z1)

# printing tool:
def print2D(list_name,transpose):
    """ This function prints a 2D list in a nice way with columns and everything lined up
    the list name is the name of the 2D list that you want to print. the transpose part is
    for whether
    you want to transpose the matrix before printing. if so type "True", otherwise type
    "False".
    """
    if transpose == True:
        list_name = [list(i) for i in zip(*list_name)]
    mx = max((len(str(ele)) for sub in list_name for ele in sub))
    for row in list_name:
        print(" ".join(["{:<{mx}}".format(ele,mx=mx) for ele in row]))

#####
##### END OF FUNCTIONS #####
#####

# import geometry from a Gaussian input or output.
unsorted_geometry = getGeometryGeneral(input)

```

```

# check if ring atoms are requested manually.
if rings == None:
    # if ring atoms are not specified, sort the atoms and find the rings.
    geometry = SortByAtomicNumber(unsorted_geometry)
    ssr = getRings(geometry)
else:
    # otherwise, do not sort the geometry just convert the ring indices that were input
    to an appropriate nested list.
    geometry = unsorted_geometry
    ssr = rings.split()
    ssr = [int(x) for x in ssr]
    ssr = [x - 1 for x in ssr]
    ssr = [ssr]

# extract the atom names for the basis set.
gaussian_atoms = " ".join(set(getAxis(unsorted_geometry,0))) + ' 0'

# generate the input files for nics(1)zz
for i in range(0,len(ssr)):
    # define the name of the files to be generated.
    file_name = str(input)
    file_name = file_name[0:-4]
    file_name = file_name + '_ring' + str(i + 1) + '.com'
    # print(file_name)
    ring_elements = list(ssr[i])
    ring_elements = [i + 1 for i in ring_elements]
    ring_elements = " ".join(map(str, ring_elements))
    # insert the NICS probe and write the file
    nics = insertNICS(geometry,ring_elements)
    orig_stdout = sys.stdout
    sys.stdout = open(file_name,'wt')
    print(gaussian_card)
    print()
    print("This is a comment line.")
    print()
    print(charge_and_multiplicity)
    print2D(nics,False)
    print()
    print(gaussian_atoms)
    print(basis)
    print('****')
    print()
    sys.stdout.close()
    sys.stdout=orig_stdout
    print(ring_elements)
    print2D(nics, False)

```

BIBLIOGRAPHY

BIBLIOGRAPHY

- (1) Smith, D. A. A Brief History of the Hydrogen Bond. In *Modeling the Hydrogen Bond*; ACS Symposium Series; American Chemical Society, 1994; Vol. 569, pp 1–5.
- (2) Rocke, A. J. It Began with a Daydream: The 150th Anniversary of the Kekulé Benzene Structure. *Angew. Chem. Int. Ed.* **2015**, *54* (1), 46–50.
- (3) Pauling, L. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*, 3. ed., 17. print.; Cornell Univ. Press: Ithaca, New York, 2000.
- (4) Moore, T. S.; Winmill, T. F. CLXXVII.—The State of Amines in Aqueous Solution. *J. Chem. Soc. Trans.* **1912**, *101* (0), 1635–1676.
- (5) Latimer, W. M.; Rodebush, W. H. Polarity and Ionization from the Standpoint of the Lewis Theory of Valence. *J. Am. Chem. Soc.* **1920**, *42* (7), 1419–1433.
- (6) Oddo, G.; Puxeddu, E. Sui 5-Azoeugenoli e La Costituzione Dei Cosidetti o-Ossiazocompositi. *Gazzetta Chim. Ital.* *36*, 1–48.
- (7) Pfeiffer, P.; Fischer, Ph.; Kuntner, J.; Monti, P.; Pros, Z. Zur Theorie der Farblacke, II. *Justus Liebigs Ann. Chem.* **1913**, *398* (2), 137–196.
- (8) Pauling, L. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals; an Introduction to Modern Structural Chemistry (3rd Ed.)*, 3rd ed.; Cornell University Press: Ithaca (NY), 1960.
- (9) Coulson, C. A. The Hydrogen Bond: A Review of the Present Interpretations and an Introduction to the Theoretical Papers Presented at the Congress. In *Hydrogen Bonding*; Hadži, D., Ed.; Pergamon, 1959; pp 339–360.
- (10) Lennard-Jones, J. E.; Pople, J. A. Molecular Association in Liquids I. Molecular Association Due to Lone-Pair Electrons. *Proc. R. Soc. Lond. Ser. Math. Phys. Sci.* **1951**, *205* (1081), 155–162.
- (11) Morokuma, K. Molecular Orbital Studies of Hydrogen Bonds. III. C=O···H–O Hydrogen Bond in H₂CO···H₂O and H₂CO···2H₂O. *J. Chem. Phys.* **1971**, *55* (3), 1236–1244.
- (12) Kitaura, K.; Morokuma, K. A New Energy Decomposition Scheme for Molecular Interactions within the Hartree-Fock Approximation. *Int. J. Quantum Chem.* **1976**, *10* (2), 325–340.

- (13) Morokuma, K. Why Do Molecules Interact? The Origin of Electron Donor-Acceptor Complexes, Hydrogen Bonding and Proton Affinity. *Acc. Chem. Res.* **1977**, *10* (8), 294–300.
- (14) Phipps, M. J. S.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. Energy Decomposition Analysis Approaches and Their Evaluation on Prototypical Protein–Drug Interaction Patterns. *Chem. Soc. Rev.* **2015**, *44* (10), 3177–3211.
- (15) McClellan, A. L.; Pimentel, G. C. *The Hydrogen Bond*; W.H. Freeman & Co Ltd, 1960.
- (16) Hoffmann, A. W. I. On Insolinic Acid. *Proc. R. Soc. Lond.* **1857**, *8*, 1–3.
- (17) Kekulé, Aug. Untersuchungen über aromatische Verbindungen Ueber die Constitution der aromatischen Verbindungen. I. Ueber die Constitution der aromatischen Verbindungen. *Ann. Chem. Pharm.* **1866**, *137* (2), 129–196.
- (18) Armit, J. W.; Robinson, R. CCXI.—Polynuclear Heterocyclic Aromatic Types. Part II. Some Anhydronium Bases. *J Chem Soc Trans* **1925**, *127* (0), 1604–1618.
- (19) von E. Doering, W.; Detert, F. L. Cycloheptatrienylium Oxide. *J. Am. Chem. Soc.* **1951**, *73* (2), 876–877.
- (20) Dewar, M. J. S. Structure of Stipitatic Acid. *Nature* **1945**, *155* (3924), 50–51.
- (21) Inuzuka, K.; Fujimoto, A. Theoretical Considerations of the Dimer-Formation Energy and Electronic Structures of the Monomer and Dimer of the 2-Pyridone by the Molecular Orbital Method. *Bull. Chem. Soc. Jpn.* **1982**, *55* (8), 2537–2540.
- (22) Inuzuka, K.; Fujimoto, A. Hydrogen Bond Energies of 2-Aminopyridine Dimer and 2-Aminopyridine-2-Pyridone Complex Formation. *Spectrochim. Acta Part Mol. Spectrosc.* **1984**, *40* (7), 623–627.
- (23) Sanz, P.; Mó, O.; Yáñez, M.; Elguero, J. Resonance-Assisted Hydrogen Bonds: A Critical Examination. Structure and Stability of the Enols of β -Diketones and β -Enaminones. *J. Phys. Chem. A* **2007**, *111* (18), 3585–3591.
- (24) Sanz, P.; Mó, O.; Yáñez, M.; Elguero, J. Non-Resonance-Assisted Hydrogen Bonding in Hydroxymethylene and Aminomethylene Cyclobutanones and Cyclobutenones and Their Nitrogen Counterparts. *ChemPhysChem* **2007**, *8* (13), 1950–1958.
- (25) Chen, Z.; Wannere, C. S.; Corminboeuf, C.; Puchta, R.; Schleyer, P. von R. Nucleus-Independent Chemical Shifts (NICS) as an Aromaticity Criterion. *Chem. Rev.* **2005**, *105*, 3842–3888.

- (26) Schleyer, P. von R.; Maerker, C.; Dransfeld, A.; Jiao, H.; van Eikema Hommes, N. J. R. Nucleus-Independent Chemical Shifts: A Simple and Efficient Aromaticity Probe. *J. Am. Chem. Soc.* **1996**, *118* (26), 6317–6318.
- (27) Rossum, G. van. *Python Tutorial*; CS-R9526; Centrum voor Wiskunde en Informatica (CWI): Amsterdam, 1995.
- (28) Jones, E.; Oliphant, T.; Peterson, P.; others. *SciPy: Open Source Scientific Tools for Python*; 2001.
- (29) Oliphant, T. E. *A Guide to NumPy*; Trelgol Publishing USA, 2006; Vol. 1.
- (30) O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminformatics* **2011**, *3* (1), 33.
- (31) Schleyer, P. v. R.; Manoharan, M.; Wang, Z.-X.; Kiran, B.; Jiao, H.; Puchta, R.; van Eikema Hommes, N. J. R. Dissected Nucleus-Independent Chemical Shift Analysis of π -Aromaticity and Antiaromaticity. *Org. Lett.* **2001**, *3* (16), 2465–2468.
- (32) Pascal, R. A.; Grossman, R. B.; Van Engen, D. Synthesis of In-[34,10][7]Metacyclophane: Projection of an Aliphatic Hydrogen toward the Center of an Aromatic Ring. *J. Am. Chem. Soc.* **1987**, *109* (22), 6878–6880.
- (33) Lodewyk, M. W.; Siebert, M. R.; Tantillo, D. J. Computational Prediction of ^1H and ^{13}C Chemical Shifts: A Useful Tool for Natural Product, Mechanistic, and Synthetic Organic Chemistry. *Chem. Rev.* **2012**, *112* (3), 1839–1862.
- (34) Kakeshpour, T.; Wu, J. I.; Jackson, J. E. AMHB: (Anti)Aromaticity-Modulated Hydrogen Bonding. *J. Am. Chem. Soc.* **2016**, *138* (10), 3427–3432.
- (35) Řezáč, J.; Hobza, P. Describing Noncovalent Interactions beyond the Common Approximations: How Accurate Is the “Gold Standard,” CCSD(T) at the Complete Basis Set Limit? *J. Chem. Theory Comput.* **2013**, *9* (5), 2151–2155.
- (36) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Comparing Counterpoise-Corrected, Uncorrected, and Averaged Binding Energies for Benchmarking Noncovalent Interactions. *J. Chem. Theory Comput.* **2014**, *10* (1), 49–57.
- (37) Thomas, I. R.; Bruno, I. J.; Cole, J. C.; Macrae, C. F.; Pidcock, E.; Wood, P. A. WebCSD : The Online Portal to the Cambridge Structural Database. *J. Appl. Crystallogr.* **2010**, *43* (2), 362–366.
- (38) De Camp, W. H.; Stewart, J. M. The Crystal and Molecular Structure of 3-Methyl-3-Pyrazolin-5-One. *Acta Crystallogr. B* **1971**, *27* (6), 1227–1232.

- (39) Shahani, T.; Fun, H.-K.; Ragavan, R. V.; Vijayakumar, V.; Sarveswari, S. Tert-Butyl 3-Oxo-2,3,4,5,6,7-Hexa-Hydro-1H-Pyrazolo[4,3-c]Pyridine-5-Carboxyl-Ate. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2009**, *66* (Pt 1), o142–o143.
- (40) Loh, W.-S.; Fun, H.-K.; Ragavan, R. V.; Vijayakumar, V.; Venkatesh, M. 5-Ethyl-4-Phenyl-1H-Pyrazol-3(2H)-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2011**, *67* (Pt 2), o403–o404.
- (41) Loh, W.-S.; Fun, H.-K.; Ragavan, R. V.; Vijayakumar, V.; Sarveswari, S. 4-Methyl-5-Phenyl-1H-Pyrazol-3(2H)-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2010**, *67* (Pt 1), o151–o152.
- (42) Shahani, T.; Fun, H.-K.; Ragavan, R. V.; Vijayakumar, V.; Sarveswari, S. 5-Ethyl-4-Methyl-1H-Pyrazol-3(2H)-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2010**, *66* (Pt 6), o1357–o1358.
- (43) Sbit, M.; Dupont, L.; Dideberg, O.; Snoek, J. P.; Delarge, J. Structure Du Phenyl-1 Pyrazolidinone-3 (Phénidone). *Acta Crystallogr. C* **1987**, *43* (4), 718–720.
- (44) Shi, H.; Zhu, H.-J.; Yin, P.-W.; Wang, J.-T.; Shi, X.-L. 5-(4-Methoxyphenyl)-1-Phenylpyrazolidin-3-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2005**, *61* (7), o2246–o2247.
- (45) Dai, B.-J.; Liu, Y.-Y.; Xu, Q.-B.; Hu, J.; Zhu, H.-J. 1-(4-Fluoro-Phen-Yl)-5-(4-Methoxy-Phen-Yl)Pyrazolidin-3-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2009**, *65* (Pt 4), o675–o675.
- (46) Jia, H.-S.; Li, Y.-F.; Liu, Y.-Y.; Liu, S.; Zhu, H.-J. 1-(4-Isopropyl-Phen-Yl)-5-(4-Methoxy-Phen-Yl)Pyrazolidin-3-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2008**, *64* (Pt 5), o855–o855.
- (47) Sun, Y.-F.; Jia, H.-S.; Liu, S.; Luo, Z.-H.; Zhu, H.-J. 5-(3,4-Dimethoxyphenyl)-1-(4-Methylphenyl)Pyrazolidin-3-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2007**, *63* (8), o3657–o3657.
- (48) Shintani, R.; Soh, Y.-T.; Hayashi, T. Rhodium-Catalyzed Asymmetric Arylation of Azomethine Imines. *Org. Lett.* **2010**, *12* (18), 4106–4109.
- (49) Yan, S.-J.; Huang, C.; Li, Y.-M.; Yan, Y.-Y.; Lin, J. 1-(2,6-Difluoro-Benzo-Yl)-3-(2,3,5-Trichloro-Phen-Yl)Urea. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2008**, *64* (Pt 11), o2102–o2102.
- (50) Shi, H.; Zhu, H.-J.; Wang, J.-T. 5-(4-Chlorophenyl)-1-Phenylpyrazolidin-3-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2006**, *62* (1), o233–o235.

- (51) Rybakov, V. B.; Babaev, E. V.; Belykh, E. N. 2-Amino-5-(4-Chlorophenyl)-1-Methylimidazole. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2002**, *58* (2), o126–o128.
- (52) Guo, X.; Chen, W.; Chen, B.; Huang, W.; Qi, W.; Zhang, G.; Yu, Y. One-Pot Three-Component Strategy for Functionalized 2-Aminoimidazoles via Ring Opening of α -Nitro Epoxides. *Org. Lett.* **2015**, 1157–1159.
- (53) Ermolat'ev, D. S.; Alifanov, V. L.; Rybakov, V. B.; Babaev, E. V.; Van Der Eycken, E. V. A Concise Microwave-Assisted Synthesis of 2-Aminoimidazole Marine Sponge Alkaloids of the Isonaamines Series. *Synthesis* **2008**, No. 13, 2083–2088.
- (54) Lambert, P.; Evrard, G.; Durant, F. 2,3,5,6-Tetrahydro-5-Phenyl-1H-Imidazo[1,2-a]Imidazole (Imafen). *Acta Crystallogr. B* **1978**, *34* (8), 2654–2655.
- (55) Isobe, T.; Fukuda, K.; Yamaguchi, K.; Seki, H.; Tokunaga, T.; Ishikawa, T. Modified Guanidines as Potential Chiral Superbases. 3. Preparation of 1,4,6-Triazabicyclooctene Systems and 1,4-Disubstituted 2-Iminoimidazolidines by the 2-Chloro-1,3-Dimethylimidazolinium Chloride-Induced Cyclization of Guanidines with a Hydroxyethyl Su. *J. Org. Chem.* **2000**, *65* (23), 7779–7785.
- (56) Cotton, F. A.; Murillo, C. A.; Wang, X.; Wilkinson, C. C. Homologues of the Easily Ionized Compound Mo₂(Hpp)₄ Containing Smaller Bicyclic Guanidines. *Inorg. Chem.* **2006**, *45* (14), 5493–5500.
- (57) Malerich, J. P.; Hagihara, K.; Rawal, V. H. Chiral Squaramide Derivatives Are Excellent Hydrogen Bond Donor Catalysts. *J. Am. Chem. Soc.* **2008**, *130* (44), 14416–14417.
- (58) McGuirk, C. M.; Katz, M. J.; Stern, C. L.; Sarjeant, A. A.; Hupp, J. T.; Farha, O. K.; Mirkin, C. A. Turning On Catalysis: Incorporation of a Hydrogen-Bond-Donating Squaramide Moiety into a Zr Metal–Organic Framework. *J. Am. Chem. Soc.* **2015**, *137* (2), 919–925.
- (59) Saez Talens, V.; Englebienne, P.; Trinh, T. T.; Noteborn, W. E. M.; Voets, I. K.; Kieltyka, R. E. Aromatic Gain in a Supramolecular Polymer. *Angew. Chem. Int. Ed.* **2015**, *54* (36), 10502–10506.
- (60) Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. The Nature of the Hydrogen Bond in DNA Base Pairs: The Role of Charge Transfer and Resonance Assistance. *Chem. – Eur. J.* **1999**, *5* (12), 3581–3594.
- (61) Choi, J. Y.; Plummer, M. S.; Starr, J.; Desbonnet, C. R.; Soutter, H.; Chang, J.; Miller, J. R.; Dillman, K.; Miller, A. A.; Roush, W. R. Structure Guided Development of Novel Thymidine Mimetics Targeting *Pseudomonas Aeruginosa* Thymidylate Kinase: From Hit to Lead Generation. *J. Med. Chem.* **2012**, *55* (2), 852–870.

- (62) Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. Three-Dimensional Structures of Avidin and the Avidin-Biotin Complex. *Proc. Natl. Acad. Sci.* **1993**, *90* (11), 5076–5080.
- (63) Szyk, Ł.; Guo, J.; Yang, M.; Dreyer, J.; Tolstoy, P. M.; Nibbering, E. T. J.; Czarnik-Matusiewicz, B.; Elsaesser, T.; Limbach, H.-H. The Hydrogen-Bonded 2-Pyridone Dimer Model System. 1. Combined NMR and FT-IR Spectroscopy Study. *J. Phys. Chem. A* **2010**, *114* (29), 7749–7760.
- (64) Alkorta, I.; Elguero, J. Influence of Intermolecular Hydrogen Bonds on the Tautomerism of Pyridine Derivatives. *J. Org. Chem.* **2002**, *67* (5), 1515–1519.
- (65) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 Theory Using Reduced Order Perturbation Theory. *J. Chem. Phys.* **2007**, *127* (12), 124105.
- (66) Meschede, L.; Gerritzen, D.; Limbach, H.-H. Dynamic NMR Study of the Interference between Cyclic Proton Exchange, Selfassociation and Hindered Rotation of Diphenylformamidine in Tetrahydrofuran. *Berichte Bunsenges. Für Phys. Chem.* **1988**, *92* (4), 469–485.
- (67) Tan, H. K. S. Determination of the NMR Monomer Shift and Dimerization Constant in a Self-Associating System by Direct Application of the Least-Squares Method. *J. Chem. Soc. Faraday Trans.* **1994**, *90* (23), 3521.
- (68) Findeisen, M.; Brand, T.; Berger, S. ¹H-NMR Thermometer Suitable for Cryoprobes. *Magn. Reson. Chem.* **2007**, *45* (2), 175–178.
- (69) Muller, N.; Hughes, O. R. Nuclear Magnetic Resonance Dilution Shifts for Carboxylic Acids in Rigorously Dried Solvents. II. Benzoic Acid in Benzene. *J. Phys. Chem.* **1966**, *70* (12), 3975–3982.
- (70) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **1997**, *62* (21), 7512–7515.
- (71) Rong, Y.; Al-Harbi, A.; Krieger, B.; Parkin, G. Structural Characterization of 2-Imidazolones: Comparison with Their Heavier Chalcogen Counterparts. *Inorg. Chem.* **2013**, *52* (12), 7172–7182.
- (72) Caudle, M. T.; Tassone, E.; Groy, T. L. Anhydrous 1-Methylimidazolidin-2-One. *Acta Crystallogr. Sect. E Struct. Rep. Online* **2005**, *61* (10), o3269–o3270.
- (73) Al-Harbi, A.; Sattler, W.; Sattler, A.; Parkin, G. Synthesis and Structural Characterization of Tris(2-Oxo-1-Tert-Butylimidazolyl) and Tris(2-Oxo-1-Methylbenzimidazolyl)Hydroborato Complexes: A New Class of Tripodal Oxygen Donor Ligand. *Chem. Commun.* **2011**, *47* (11), 3123.

- (74) Hickey, A. L.; Rowley, C. N. Benchmarking Quantum Chemical Methods for the Calculation of Molecular Dipole Moments and Polarizabilities. *J. Phys. Chem. A* **2014**, *118*(20), 3678–3687.
- (75) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science: Sausalito, CA, 2006.
- (76) Gilli, G.; Bellucci, F.; Ferretti, V.; Bertolasi, V. Evidence for Resonance-Assisted Hydrogen Bonding from Crystal-Structure Correlations on the Enol Form of the .Beta.-Diketone Fragment. *J. Am. Chem. Soc.* **1989**, *111* (3), 1023–1028.
- (77) Gilli, G.; Gilli, P. Towards an Unified Hydrogen-Bond Theory. *J. Mol. Struct.* **2000**, *552* (1–3), 1–15.
- (78) Gilli, P.; Bertolasi, V.; Ferretti, V.; Gilli, G. Evidence for Resonance-Assisted Hydrogen Bonding. 4. Covalent Nature of the Strong Homonuclear Hydrogen Bond. Study of the O-H--O System by Crystal Structure Correlation Methods. *J. Am. Chem. Soc.* **1994**, *116* (3), 909–915.
- (79) Alkorta, I.; Elguero, J.; Mó, O.; Yáñez *, M.; Del Bene, J. E. Do Coupling Constants and Chemical Shifts Provide Evidence for the Existence of Resonance-Assisted Hydrogen Bonds? *Mol. Phys.* **2004**, *102* (23–24), 2563–2574.
- (80) Góra, R. W.; Maj, M.; Grabowski, S. J. Resonance-Assisted Hydrogen Bonds Revisited. Resonance Stabilization vs. Charge Delocalization. *Phys. Chem. Chem. Phys.* **2013**, *15* (7), 2514.
- (81) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A Direct MP2 Gradient Method. *Chem. Phys. Lett.* **1990**, *166* (3), 275–280.
- (82) Head-Gordon, M.; Head-Gordon, T. Analytic MP2 Frequencies without Fifth-Order Storage. Theory and Application to Bifurcated Hydrogen Bonds in the Water Hexamer. *Chem. Phys. Lett.* **1994**, *220* (1–2), 122–128.
- (83) Sæbø, S.; Almlöf, J. Avoiding the Integral Storage Bottleneck in LCAO Calculations of Electron Correlation. *Chem. Phys. Lett.* **1989**, *154* (1), 83–89.
- (84) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153* (6), 503–506.
- (85) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Semi-Direct Algorithms for the MP2 Energy and Gradient. *Chem. Phys. Lett.* **1990**, *166* (3), 281–289.
- (86) Peterson, K. A.; Woon, D. E.; Dunning, T. H. Benchmark Calculations with Correlated Molecular Wave Functions. IV. The Classical Barrier Height of the $\text{H}+\text{H}_2 \rightarrow \text{H}_2 + \text{H}$ Reaction. *J. Chem. Phys.* **1994**, *100* (10), 7410–7415.

- (87) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96* (9), 6796–6806.
- (88) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90* (2), 1007–1023.
- (89) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993**, *98* (2), 1358–1371.
- (90) Wilson, A. K.; van Mourik, T.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. VI. Sextuple Zeta Correlation Consistent Basis Sets for Boron through Neon. *J. Mol. Struct. THEOCHEM* **1996**, *388*, 339–349.
- (91) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. *Gaussian 16 Revision A.03*; 2016.
- (92) Purvis, G. D.; Bartlett, R. J. A Full Coupled-cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76* (4), 1910–1918.
- (93) *Advances in Chemical Physics: LeFebvre/Advances*; LeFebvre, R., Moser, C., Eds.; Advances in Chemical Physics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1969.
- (94) Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F. An Efficient Reformulation of the Closed-shell Coupled Cluster Single and Double Excitation (CCSD) Equations. *J. Chem. Phys.* **1988**, *89* (12), 7382–7387.
- (95) Scuseria, G. E.; Schaefer, H. F. Is Coupled Cluster Singles and Doubles (CCSD) More Computationally Intensive than Quadratic Configuration Interaction (QCISD)? *J. Chem. Phys.* **1989**, *90* (7), 3700–3703.
- (96) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. Quadratic Configuration Interaction. A General Technique for Determining Electron Correlation Energies. *J. Chem. Phys.* **1987**, *87* (10), 5968–5975.
- (97) Jung, Y.; Sodt, A.; Gill, P. M. W.; Head-Gordon, M. Auxiliary Basis Expansions for Large-Scale Electronic Structure Calculations. *Proc. Natl. Acad. Sci.* **2005**, *102* (19), 6692–6697.
- (98) Whitten, J. L. Coulombic Potential Energy Integrals and Approximations. *J. Chem. Phys.* **1973**, *58* (10), 4496–4501.

- (99) Stuebing, E. W.; Weare, J. H.; Parr, R. G. Discontinuous Approximate Molecular Electronic Wave-Functions. *Int. J. Quantum Chem.* **1977**, *11* (1), 81–102.
- (100) Werner, H.-J.; Manby, F. R.; Knowles, P. J. Fast Linear Scaling Second-Order Møller-Plesset Perturbation Theory (MP2) Using Local and Density Fitting Approximations. *J. Chem. Phys.* **2003**, *118* (18), 8149–8160.
- (101) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral Approximations for LCAO-SCF Calculations. *Chem. Phys. Lett.* **1993**, *213* (5–6), 514–518.
- (102) Gill, P. M. W.; Johnson, B. G.; Pople, J. A.; Taylor, S. W. Modeling the Potential of a Charge Distribution. *J. Chem. Phys.* **1992**, *96* (9), 7178–7179.
- (103) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On Some Approximations in Applications of $X\alpha$ Theory. *J. Chem. Phys.* **1979**, *71* (8), 3396–3402.
- (104) Skylaris, C.-K.; Gagliardi, L.; Handy, N. C.; Ioannou, A. G.; Spencer, S.; Willetts, A. On the Resolution of Identity Coulomb Energy Approximation in Density Functional Theory. *J. Mol. Struct. THEOCHEM* **2000**, *501–502*, 229–239.
- (105) Baerends, E. J.; Ellis, D. E.; Ros, P. Self-Consistent Molecular Hartree—Fock—Slater Calculations I. The Computational Procedure. *Chem. Phys.* **1973**, *2* (1), 41–51.
- (106) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23* (2), 69–89.
- (107) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-Set Convergence in Correlated Calculations on Ne, N₂, and H₂O. *Chem. Phys. Lett.* **1998**, *286* (3–4), 243–252.
- (108) Han, S.; Zard, S. Z. Stabilization of Radicals by Imides. A Modular Stereoselective Approach to Protected Functional 1,2-Diamines. *Org. Lett.* **2014**, *16* (20), 5386–5389.
- (109) Wong, O.; Tsuzuki, N.; Richardson, M.; Rytting, H.; Konishi, R.; Higuchi, T. An Improved Synthesis of 1-Alkyl-4-Imidazolin-2-Ones. *HETEROCYCLES* **1987**, *26* (12), 3153.
- (110) Powers, T. S.; Wulff, W. D.; Quinn, J.; Shi, Y.; Jiang, W.; Hsung, R.; Parisi, M.; Rahm, A.; Wu Jiang, X.; Yap, G. P. A.; et al. The Preparation of Imidazolidinone and Oxazolidinone Chelated Carbene Complexes. *J. Organomet. Chem.* **2001**, *617–618*, 182–208.
- (111) Asier Gómez-SanJuan; José Manuel Botija; Almudena Méndez; Nuria Sotomayor; Esther Lete. C-N Bond Forming Reactions in the Synthesis of Substituted 2-Aminoimidazole Derivatives. *Arkivoc* **2013**, *2014* (2), 44.

- (112) Kieseewetter, M. K.; Scholten, M. D.; Kirn, N.; Weber, R. L.; Hedrick, J. L.; Waymouth, R. M. Cyclic Guanidine Organic Catalysts: What Is Magic About Triazabicyclodecene? *J. Org. Chem.* **2009**, *74* (24), 9490–9496.
- (113) Ottiger, P.; Pfaffen, C.; Leist, R.; Leutwyler, S.; Bachorz, R. A.; Klopper, W. Strong N–H··· π Hydrogen Bonding in Amide–Benzene Interactions. *J. Phys. Chem. B* **2009**, *113* (9), 2937–2943.
- (114) Kalescky, R.; Kraka, E.; Cremer, D. Accurate Determination of the Binding Energy of the Formic Acid Dimer: The Importance of Geometry Relaxation. *J. Chem. Phys.* **2014**, *140* (8), 084315.
- (115) Kollipost, F.; Larsen, R. W.; Domanskaya, A. V.; Nörenberg, M.; Suhm, M. A. Communication: The Highest Frequency Hydrogen Bond Vibration and an Experimental Value for the Dissociation Energy of Formic Acid Dimer. *J. Chem. Phys.* **2012**, *136* (15), 151101.
- (116) Frey, J. A.; Leutwyler, S. An Ab Initio Benchmark Study of Hydrogen Bonded Formamide Dimers. *J. Phys. Chem. A* **2006**, *110* (45), 12512–12518.
- (117) Assaf, K. I.; Florea, M.; Antony, J.; Henriksen, N. M.; Yin, J.; Hansen, A.; Qu, Z.; Sure, R.; Klapstein, D.; Gilson, M. K.; et al. HYDROPHOBE Challenge: A Joint Experimental and Computational Study on the Host–Guest Binding of Hydrocarbons to Cucurbiturils, Allowing Explicit Evaluation of Guest Hydration Free-Energy Contributions. *J. Phys. Chem. B* **2017**, *121* (49), 11144–11162.
- (118) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113* (18), 6378–6396.
- (119) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102* (11), 1995–2001.
- (120) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24* (6), 669–681.
- (121) Scalmani, G.; Frisch, M. J. Continuous Surface Charge Polarizable Continuum Models of Solvation. I. General Formalism. *J. Chem. Phys.* **2010**, *132* (11), 114110.
- (122) Thanthiriwatte, K. S.; Hohenstein, E. G.; Burns, L. A.; Sherrill, C. D. Assessment of the Performance of DFT and DFT-D Methods for Describing Distance Dependence of Hydrogen-Bonded Interactions. *J. Chem. Theory Comput.* **2011**, *7* (1), 88–96.

- (123) Wu, J. I.; Jackson, J. E.; Schleyer, P. von R. Reciprocal Hydrogen Bonding–Aromaticity Relationships. *J. Am. Chem. Soc.* **2014**, *136* (39), 13526–13529.
- (124) Kakeshpour, T.; Bailey, J. P.; Jenner, M. R.; Howell, D. E.; Staples, R. J.; Holmes, D.; Wu, J. I.; Jackson, J. E. High-Field NMR Spectroscopy Reveals Aromaticity-Modulated Hydrogen Bonding in Heterocycles. *Angew. Chem. Int. Ed.* **2017**, *56* (33), 9842–9846.
- (125) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10* (44), 6615.
- (126) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.
- (127) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]. *Phys. Rev. Lett.* **1997**, *78* (7), 1396–1396.
- (128) Zhao, Y.; Truhlar, D. G. 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.* **2008**, *120* (1–3), 215–241.
- (129) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170.
- (130) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham Global-Hybrid Exchange–Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7* (8), 5032–5051.
- (131) Henderson, T. M.; Izmaylov, A. F.; Scalmani, G.; Scuseria, G. E. Can Short-Range Hybrids Describe Long-Range-Dependent Properties? *J. Chem. Phys.* **2009**, *131* (4), 044108.
- (132) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.
- (133) Austin, A.; Petersson, G. A.; Frisch, M. J.; Dobek, F. J.; Scalmani, G.; Throssell, K. A Density Functional with Spherical Atom Dispersion Terms. *J. Chem. Theory Comput.* **2012**, *8* (12), 4989–5007.

- (134) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871.
- (135) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), A1133–A1138.
- (136) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200–1211.
- (137) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *WIREs Comput Mol Sci* **2012**, *2*, 242–253.
- (138) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, D.; Korona, T.; Lindh, R.; et al. *MOLPRO, Version 2018.2, a Package of Ab Initio Programs*; molpro, 2018.