HHS Public Access

Author manuscript

Biophys Chem. Author manuscript; available in PMC 2021 December 01.

Published in final edited form as:

Biophys Chem. 2020 December ; 267: 106476. doi:10.1016/j.bpc.2020.106476.

Using quantum chemistry to estimate chemical shifts in biomolecules

David A. Case[1,*]

[1]Dept. of Chemistry & Chemical Biology, Piscataway, NJ 08854

Abstract

An automated fragmentation quantum mechanics/molecular mechanics approach (AFNMR) has shown promising results in chemical shift calculations for biomolecules. Sample results for ubiquitin, and an RNA hairpin and helix are presented, and used to recent directions in quantum calculations. Trends in chemical shift are stable with regards to change in density functional or basis sets, and the use of the small “pcSseg-0” basis, which was optimized for chemical shift prediction,[1] opens the way to more extensive conformational averaging, which can often be necessary, even for fairly well-defined structures.

Keywords

chemical shift NMR conformational averaging

1. Introduction

Nuclear magnetic resonance (NMR) spectroscopy plays an important role in the determination of three dimensional structural ensembles of macromolecules. In principle, but not so much in practice, chemical shifts could provide important information here: they provide sensitive local probes at almost every location, and (in the common case of fast exchange) are averaged in a simple additive way over the molecule’s conformational ensemble. As an electronic structure property, they report on a large number environmental factors, such as the conformation of neighboring residues, hydrogen bonding, ring-current effects, and long-range electrostatics [2–4]. A robust understanding of the connection between structure and chemical shifts could help in characterizing protein and nucleic acid structure and dynamics.

[*]**Corresponding author.** david.case@rutgers.edu.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

One obvious, and well-studied, approach seeks to train predictors using the large number of observed shifts collected in the BioMagResBank,[5] associating these values in some way with the structures that serve as input to the calculations.[6–12] One limitation here is that the (conformationally-averaged) structures that give rise to observed shifts are almost never known at the level of detail and precision that one might like. This can be addressed to some extent by using quantum chemical calculations in place of, or in addition to, observed shifts as input to the training procedure.[13, 14] This eliminates uncertainty about the connection between the structure and the computed chemical shift, but at the cost of losing the effects of conformational averaging that is encoded in observed shifts. A third approach, analyzed here, tries to estimate shifts directly from assumed ensembles of input structures, using established quantum chemical procedures to compute the shifts. This clearly requires procedures that are sufficiently accurate, affordable and automated to allow practical calculations to be carried out.

NMR chemical shifts describe the electronic response of a molecule to magnetic fields arising from the spectrometer and from nuclear magnetic moments. Methods and computer codes to carry out such calculations at various levels of theory are well-established.[15] Density functional theory provides a computationally straightforward way to incorporate some important effects beyond the Hartree-Fock approximation, in a manner efficient enough to be applied to molecules or fragments containing one to two hundred atoms. Since proteins and nucleic acids are too big for routine application of quantum chemical methods, fragmentation approaches that divide the system into (generally overlapping) fragments are attractive, particularly for local properties like chemical shifts. This paper gives an short overview of some recent results using one such method (“AFNMR”), setting the stage for what may be a new generation of studies that use quantum chemistry directly to interpret biomolecular chemical shifts.

2. The automated fragmentation (AFNMR) method

The AFNMR model [16–18] takes the entire protein and solvent effect into consideration, while limiting the size of the quantum region, as described in Fig. 1. Overlapping fragments are used to break the larger problem into pieces, and the electrostatic effects of the solvent and the remaining parts of the biomolecule are computed using a Poisson-Boltzmann procedure implemented in the MEAD package.[19, 20] This uses a three-dielectric model, with $\varepsilon = 1$ for the quantum region, $\varepsilon = 4$ for the remainder of the biomolecule, and $\varepsilon = 78$ for the solvent region. Fitting these to point charges surrounding the molecule allows for easy incorporation into most quantum chemical programs. The AFNMR code creates input files for Gaussian,[21] ORCA,[22] QChem[23] and deMon;[24] once the quantum calculations are complete, it parses their outputs, assembling results for the entire molecule in a convenient format. Further details are given elsewhere.[17] This implementation is not unique, and similar results have been reported by others.[25–29] The automatic fragmentation method used here was first made available for general use in 2015,[17] and has been applied to a fair number of problems in proteins and nucleic acids.[30–38] The code is available at https://github.com/dacase/afnmr.

3. Some illustrative results

In the limited space available here, I will discuss some results from three systems where NMR solution structures have been generated from conventional restraints in combination with a large number of residual dipolar couplings. The first is an RNA hairpin closed by a UUCG tetraloop (PDB ID 2koc);[39–41] the second is a helix from the ψ^{746} element in E. coli (PDB ID 2gbh);[42–44] and the third the monomeric form of ubiquitin (PDB ID 1d3z).[45, 46] Results are chosen to illustrate particular issues that arise in the calculations, rather than for how informative they are with respect to the structures themselves. Unless otherwise noted, all calculations used the OLYP functional[47] and the psSseg-*n* basis sets.[1].

3.1. A UUCG RNA hairpin

This 14-nucleotide UUCG hairpin (PDB ID 2koc[39] has long served as a paradigm of a small RNA structure. In addition to a large number of NMR restraints, the chemical shift list is quite complete, and includes shifts for protons bonded to nitrogen and oxygen as well as to carbon. We reported some results for this system in the original AFNMR paper;[17] since then, we have generally moved to using the pcSseg-*n* basis sets optimized for chemical shielding calculations.[1] The smallest of these sets (called pcSseg-0) is at the double-zeta level with no polarization functions; the next larger (pcSseg-1) is at the [2s,1p] level for hydrogen, and the [3s,3p,1d] level for first-row atoms. Fig. 2 shows that *trends* in shifts, which are of primary importance in biomolecular studies, are nearly the same with the two basis sets. The main exception is for protons bonded to nitrogen or oxygen (HO2', H41/H42, H61/H62, H1, H3), which have smaller shifts in the smaller basis set. Even here, trends for a given type of proton are nearly the same. Calculations for single structures of 2koc (taken from the deposited NMR set) at the pcSseg-2 level are nearly identical to those at the pcSseg-1 level (results not shown, but similar calculations on ubiquitin are discussed below.)

The pcSseg-0 basis is about 5 times faster than pcSseg-1 using the deMon program, with the average fragment requiring 14 minutes on a single core of an Intel E5-2695 CPU. Since the fragment calculations are independent, all 14 fragments for a given conformation can be computed at the same time on a single CPU. This allows one to analyze 100 RNA conformations per day on a single CPU, allowing for much more extensive sampling than has been the case in the past. The smaller basis is also less prone to convergence failures, whereas about 1% of fragments analyzed at the pcSseg-1 level failed to converge, requiring hand intervention to tweak SCF convergence parameters.

Fig.3 compares computed and experimental shifts at the pcSseg-1 level. Even with the larger basis, results for protons bonded to nitrogen and oxygen are not well-predicted. It is known that the continuum electrostatic model for solvent does not faithfully capture the effects of hydrogen bonds to solvent molecules,[3, 48], and explicit solvent molecules are probably required for these protons. The use of explicit water molecules introduces a new set of challenges for conformational modeling; one approach is discussed in Section 3.3, below.

3.2. Ubiquitin

Like the UUCG hairpin, the ubiquitin monomer has long served as a paradigm of a folded globular protein that can be characterized by NMR.[45, 46, 49, 50] Here I give some information about how results depend upon the method used, using an early set of NMR structures (PDB ID 1d3z[45]) as a reference point. Fig. 4 shows the trends seen for pcSseg-0 calculations; here only protons bonded to carbon are considered. Both overall trends, and those for particular types of atom, are reasonably well reproduced. Of perhaps greater interest are results reported in Table 1, illustrating the dependence on method for estimates of backbone 15N shifts, which generally has a large shift dispersion in folded proteins. Empirical methods like sparta+[9] and shiftx+,[10] trained on observed shifts of globular proteins, have a smaller mean error than do the AFNMR results, but this is partially achieved by using slopes near 0.9 that somewhat limit the range of the predictions. The quantum results shown here are generally consistent with other studies: using a more expensive basis set or a hybrid functional such as B3LYP[51] gives remarkable similar results when trends across the entire protein are considered.

3.3. RNA helix 35-ψ^{746}

The 24-nt stem-loop corresponding to the Helix 35ψ from the 23S ribosomal RNA (pdb ID 2gbh[42]) has a 16nt helical step that is characterized by 17 residual-dipolar couplings per nucleotide, surpassing any other structure reported to date. Its structure is also informed by an extensive set of distance and torsional angle restraints. Importantly, it has recently been shown that unconstrained molecular dynamics simulations recapitulate NMR data to a high degree,[44] providing a new avenue for studying the effects of conformational averaging on chemical shifts.

The need for, and effects of, conformational averaging can be especially noticeable when individual conformations are drawn from a molecular dynamics simulation. Fig. 5 illustrates this for ^{13}C shifts for this 16-nucleotide helix. The left panel shows typical results for a single conformation; each type of carbon (see, *e.g.* C1’ or C4’) tends to have nearly the same experimental shift at each position, but computed shifts are spread out over about 20 ppm due to instantaneous differences in local conformations. By the time one has averaged 10 (data not shown) or 20 conformations (right panel), the computed shifts have converged to the same average at each nucleotide in good agreement with the observed shifts. Future work will look at the more flexible loop region not analyzed here.

4. Conclusions

The connections between chemical shift and structure in proteins and nucleic acids is a very broad subject. The direct application of quantum chemical methods avoids the need to connect structures to observed shifts in the training of a predictive algorithm, and should be able to reflect in an unbiased (if still approximate) way, the dependence of shift on structure, whether one is dealing with very local conformational changes or more global ones. Dealing with conformational variability is likely to be a key component of getting more accurate results, and also opens the possibility of using shifts to help generate (or to cross-validate) ensembles generated from other sorts of data. The results outlined here show some progress

toward an automated and affordable scheme, but it may be that fewer or smaller fragments are required, or that alternative methods for creating and analyzing them can be used. The methods described here can also be used to analyze the anisotropy of shielding tensors,[30, 34] which introduces an additional set of challenges and opportunities.

Acknowledgments

I thank Jason Swails for developing the templates used for the plots presented here; Ad Bax for pointing me to the observed shifts for helix 35 ψ ; Alex Grishaev and Christina Bergonzo for providing snapshots from their molecular dynamics simulations; and Honglue Shi and Hashim Al-Hashimi for many useful discussions. This work was supported by NIH grant U54 GM103297.

References

[1]. Jensen F, Segmented Contracted Basis Sets Optimized for Nuclear Magnetic Shielding, J. Chem. Theory Comput 11 (2015) 132–138. [PubMed: 26574211]

[2]. Buckingham A, Schaefer T, Schneider W, Solvent effects in nuclear magnetic resonance spectra, J. Chem. Phys 32 (1960) 1227–1233.

[3]. Sitkoff D, Case D, Density functional calculations of proton chemical shifts in model peptides, J. Am. Chem. Soc 119 (1997) 12262–12273.

[4]. Sitkoff D, Case D, Theories of chemical shift anisotropies in proteins and nucleic acids, Prog. NMR Spectr 32 (1998) 165–190.

[5]. Ulrich E, Akutsu H, Doreleijers J, Harano Y, Ioannidis Y, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte C, Tolmie D, Kent Wenger R, Yao H, Markley J, BioMagResBank, Nucl. Acids Res 36 (2008) D402–408. [PubMed: 17984079]

[6]. Ösapay K, Case D, A new analysis of proton chemical shifts in proteins, J. Am. Chem. Soc 113 (1991) 9436–9444.

[7]. Meiler J, PROSHIFT: Protein chemical shift prediction using artificial neural networks, J. Biomol. NMR 26 (2003) 25–37. [PubMed: 12766400]

[8]. Cavalli A, Salvatella X, Dobson C, Vendruscolo M, Protein structure determination from NMR chemical shifts, Proc. Natl. Acad. Sci. USA 104 (2007) 9615–9620. [PubMed: 17535901]

[9]. Shen Y, Bax A, SPARTA plus : a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network, J. Biomol. NMR 48 (2010) 13–22. [PubMed: 20628786]

[10]. Han B, Liu Y, Ginzinger S, Wishart D, SHIFTX2: significantly improved protein chemical shift prediction, J. Biomol. NMR 50 (2011) 43–57. [PubMed: 21448735]

[11]. Li D, Brüschweiler R, PPM_One: a static protein structure based chemical shift predictor, J. Biomol. NMR 62 (2015) 403–409. [PubMed: 26091586]

[12]. Marchant J, Summers M, Johnson B, Assigning NMR spectra of RNA, peptides and small organic molecules using molecular network visualization software, J Biomol NMR 73 (2019) 525–529. [PubMed: 31325088]

[13]. Xu X, Case D, Automated prediction of ^{15}N, $^{13}C\alpha$, $^{13}C\beta$ and ^{13}C' chemical shifts in proteins using a density functional database, J. Biomol. NMR 21 (2001) 321–333. [PubMed: 11824752]

[14]. Paruzzo F, Hofstetter A, Musil F, De S, Ceriotti M, Emsley L, Chemical shifts in molecular solids by machine learning, Nature Commun. 9 (2018) 4501. [PubMed: 30374021]

[15]. Helgaker T, Jaszunski M, Ruud K, Ab initio methods for the calculation of NMR shielding and indirect spin-spin coupling constants, Chem. Rev 99 (1999) 293–352. [PubMed: 11848983]

[16]. He X, Fusti-Molnar L, Merz K Jr., Accurate Benchmark Calculations on the Gas-Phase Basicities of Small Molecules, J. Phys. Chem. A 113 (2009) 10096–10103. [PubMed: 19694482]

[17]. Swails J, Zhu T, He X, Case D, AFNMR: Automated fragment quantum mechanical calculation of NMR chemical shifts for biomolecules, J. Biomol. NMR (2015) 125–139. [PubMed: 26232926]

[18]. Zhu T, He X, Zhang J, Fragment density functional theory calculation of NMR chemical shifts for proteins with implicit solvation, Phys. Chem. Chem. Phys 14 (2012) 7837–7845. [PubMed: 22314755]

[19]. Bashford D, Karplus M, pKa's of ionizable groups in proteins: Atomic detail from a continuum electrostatic model, Biochemistry 29 (1990) 10219–10225. [PubMed: 2271649]

[20]. Chen J, Noodleman L, Case D, Bashford D, Incorporating solvation effects into density functional electronic structure calculations, J. Phys. Chem 98 (1994) 11059–11068.

[21]. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA Jr., Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ, Gaussian16 Revision C.01, Gaussian Inc Wallingford CT (2016).

[22]. Neese F, Software update: the ORCA program system, version 4.0, WIREs Comput. Mol. Sci 8 (2018) e1327.

[23]. Shao Y, Gan Z, Epifanovsky E, Gilbert ATB, Wormit M, Kussmann J, Lange AW, Behn A, Deng J, Feng X, Ghosh D, Horn MGPR, Jacobson LD, Kaliman I, Khaliullin RZ, KÃ°s T, Landau A, Liu J, Proynov EI, Rhee YM, Richard RM, Rohrdanz MA, Steele RP, Sundstrom EJ, III HLW, Zimmerman PM, Zuev D, Albrecht B, Alguire E, Austin B, Beran GJO, Bernard YA, Berquist E, Brandhorst K, Bravaya KB, Brown ST, Casanova D, Chang C-M, Chen Y, Chien SH, Closser KD, Crittenden DL, Diedenhofen M, Jr RAD., Dop H, Dutoi AD, Edgar RG, Fatehi S, Fusti-Molnar L, Ghysels A, Golubeva-Zadorozhnaya A, Gomes J, Hanson-Heine MWD, Harbach PHP, Hauser AW, Hohenstein EG, Holden ZC, Jagau T-C, Ji H, Kaduk B, Khistyaev K, Kim J, Kim J, King RA, Klunzinger P, Kosenkov D, Kowalczyk T, Krauter CM, Lao KU, Laurent A, Lawler KV, Levchenko SV, Lin CY, Liu F, Livshits E, Lochan RC, Luenser A, Manohar P, Manzer SF, Mao S-P, Mardirossian N, Marenich AV, Maurer SA, Mayhall NJ, Oana CM, Olivares-Amaya R, O'Neill DP, Parkhill JA, Perrine TM, Peverati R, Pieniazek PA, Prociuk A, Rehn DR, Rosta E, Russ NJ, Sergueev N, Sharada SM, Sharmaa S, Small DW, Sodt A, Stein T, StÃ1/4ck D, Su Y-C, Thom AJW, Tsuchimochi T, Vogt L, Vydrov O, Wang T, Watson MA, Wenzel J, White A, Williams CF, Vanovschi V, Yeganeh S, Yost SR, You Z-Q, Zhang IY, Zhang X, Zhou Y, Brooks BR, Chan GKL, Chipman DM, Cramer CJ, III WAG, Gordon MS, Hehre WJ, Klamt A, III HFS, Schmidt MW, Sherrill CD, Truhlar DG, Warshel A, Xua X, Aspuru-Guzik A, Baer R, Bell AT, Besley NA, Chai J-D, Dreuw A, Dunietz BD, Furlani TR, Gwaltney SR, Hsu C-P, Jung Y, Kong J, Lambrecht DS, Liang W, Ochsenfeld C, Rassolov VA, Slipchenko LV, Subotnik JE, Voorhis TV, Herbert JM, Krylov AI, Gill PMW, , Head-Gordon M, Advances in molecular quantum chemistry contained in the Q-Chem 4 program package, Mol. Phys 113 (2015) 184–215.

[24]. Koster A, Geudtner G, Calaminici P, Casida M, Dominguez V, Flores-Moreno R, Gamboa G, Goursot A, Heine T, Ipatov A, Janetzko F, del Campo J, Reveles J, Vela A, Zuniga-Gutierrez B, Salahub D, deMon2k, Version 3, The deMon developers, 2011.

[25]. Frank A, Onila I, Möller H, Exner T, Toward the quantum chemical calculation of nuclear magnetic resonance chemical shifts of proteins, Proteins 79 (2011) 2189–2202. [PubMed: 21557322]

[26]. Frank A, Möller H, Exner T, Toward the Quantum Chemical Calculation of NMR Chemical Shifts of Proteins. 2. Level of Theory, Basis Set, and Solvents Model Dependence, J. Chem. Theory Comput 8 (2012) 1480–1492. [PubMed: 26596758]

[27]. Exner T, Frank A, Onila I, Moeller H, Toward the Quantum Chemical Calculation of NMR Chemical Shifts of Proteins. 3. Conformational Sampling and Explicit Solvents Model, J. Chem. Theory Comput 8 (2012) 4818–4827. [PubMed: 26605634]

[28]. Kobayashi R, Amos RD, Reid DM, Collins MA, Application of the Systematic Molecular Fragmentation by Annihilation Method to ab Initio NMR Chemical Shift Calculations, J. Phys. Chem. A 122 (2018) 9135–9141. [PubMed: 30398349]

[29]. Unzueta P, Beran G, Polarizable continuum models provide an effective electrostatic embedding model for fragment-based chemical shift prediction in challenging systems, J. Computat. Chem (in press).

[30]. Tang S, Case D, Calculation of chemical shift anisotropy in proteins, J. Biomol. NMR 51 (2011) 303–312. [PubMed: 21866436]

[31]. Fu I, Case D, Baum J, Dynamic Water-Mediated Hydrogen Bonding in a Collagen Model Peptide, Biochemistry 54 (2015) 6029–6037. [PubMed: 26339765]

[32]. Zhang H, Hou G, Lu M, Ahn J, Byeon I, Langmead C, Perilla J, Hung I, Gorkov P, Gan Z, Brey W, Case D, Schulten K, Gronenborn A, Polenova T, HIV-1 Capsid Function Is Regulated by Dynamics: Quantitative Atomic-Resolution Insights by Integrating Magic-Angle-Spinning NMR, QM/MM, and MD, J. Am. Chem. Soc 138 (2016) 14066–14075. [PubMed: 27701859]

[33]. Jin X, Zhu T, Zhang J, He X, A systematic study on RNA NMR chemical shift calculation based on the automated fragmentation QM/MM approach, RSC Adv 6 (2016) 108590–108602.

[34]. Fritz M, Quinn C, Wang M, Hou G, Lu X, Koharudin L, Struppe J, Case D, Polenova T, Gronenborn A, Accurate determination of backbone chemical shift tensors in microcrystalline proteins by integrated MAS NMR and QM/MM, Phys. Chem. Chem. Phys 20 (2018) 9543–9553. [PubMed: 29577158]

[35]. Shi H, Clay M, Rangadurai A, Sathyamoorthy B, Case D, Al-Hashimi H, Atomic Structures of Excited State A-T Hoogsteen Base Pairs in Duplex DNA by Combining NMR Relaxation Dispersion, Mutagenesis, and Chemical Shift Calculations, J. Biomol. NMR 70 (2018) 229–244. [PubMed: 29675775]

[36]. Jin X, Zhu T, Zhang J, He X, Automated fragmentation QM/MM calculation of NMR chemical shifts for protein-ligand complexes, Front. Chem 6 (2018) 150. [PubMed: 29868556]

[37]. Zhou H, Sathyamoorthy B, Stelling A, Xu Y, Xue Y, Pigli Y, Case D, Rice P, Al-Hashimi H, Characterizing Watson-Crick versus Hoogsteen Base Pairing in a DNA-Protein Complex Using Nuclear Magnetic Resonance and Site-Specifically 13C-and 15N-Labeled DNA, Biochemistry 58 (2019) 1963–1974. [PubMed: 30950607]

[38]. Nam H, Becette O, LeBlanc R, Oh D, Case D, Dayie T, Deleterious effects of carbon-carbon dipolar coupling on RNA NMR dynamics, J. Biomol. NMR 74 (2020) 321–331. [PubMed: 32363430]

[39]. Nozinovic S, Fuertig B, Jonker H, Richter C, Schwalbe H, High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA., Nucl. Acids Res 38 (2010) 683–694. [PubMed: 19906714]

[40]. Giambasu G, York D, Case D, Structural fidelity and NMR relaxation analysis in a prototype RNA hairpin, RNA 21 (2015) 963–974. [PubMed: 25805858]

[41]. Borkar A, Vallurupalli P, Camilloni C, Kay L, Vendruscolo M, Simultaneous NMR characterisation of multiple minima in the free energy landscape of an RNA UUCG tetraloop, Phys. Chem. Chem. Phys 19 (2017) 2797–2804. [PubMed: 28067358]

[42]. O'Neil-Cabello E, Bryce D, Nikonowicz E, Bax A, Measurement of five dipolar couplings from a single 3D NMR multiplet applied to the study of RNA dynamics, J. Am. Chem. Soc 126 (2004) 66–67. [PubMed: 14709062]

[43]. O'Neil-Cabello E, Development of NMR Methods for Measuring Residual Dipolar Couplings and ^{3}1P Chemical Shift Anisotropy in RNA Including Structural Applications, using Ribosomal Helix-35 ψ746 as a Model Molecule, Ph.D. thesis, Rice University, 2004.

[44]. Bergonzo C, Grishaev A, Maximizing accuracy of RNA structure in refinement against residual dipolar couplings, J. Biomol. NMR 73 (2019) 117–139. [PubMed: 31049778]

[45]. Cornilescu G, Marquardt J, Ottiger M, Bax A, Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase, J. Am. Chem. Soc 120 (1998) 6836–6837.

[46]. Maltsev A, Grishaev A, Roche J, Zasloff M, Bax A, Improved Cross Validation of a Static Ubiquitin Structure Derived from High Precision Residual Dipolar Couplings Measured in a Drug-Based Liquid Crystalline Phase, J. Am. Chem. Soc 136 (2014) 3752–3755. [PubMed: 24568736]

[47]. Hoe W, Cohen A, Handy N, Assessment of a new local exchange functional OPTX, Chem. Phys. Lett 341 (2001) 319–328.

[48]. Zhu T, Zhang J, He X, Automated Fragmentation QM/MM Calculation of Amide Proton Chemical Shifts in Proteins with Explicit Solvent Model, J. Chem. Theory Comput 9 (2013) 2104–2114. [PubMed: 26583557]

[49]. Schneider D, Dellwo M, Wand A, Fast internal main-chain dynamics of human ubiquitin, Biochemistry 31 (1992) 3645–3652. [PubMed: 1314645]

[50]. Lange O, Lakomek N-A, Farès C, Schröder G, Walter K, Becker S, Meiler J, Grubmüller H, Griesinger C, de Groot B, Recognition dynamics up to microseconds revealed from a RDC-derived ubiquitin ensemble in solution, Science 320 (2008) 1471–1475. [PubMed: 18556554]

[51]. Becke A, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys 98 (1993) 5648–5652.

Highlights

- An automated fragmentation quantum mechanics/molecular mechanics approach (AFNMR) has shown promising results in chemical shift calculations for biomolecules. Sample results for ubiquitin, and an RNA hairpin and helix are presented, and used to recent directions in quantum calculations. Trends in chemical shift are stable with regards to change in density functional or basis sets, and the use of the small “pcSseg-0” basis, which was optimized for chemical shift prediction, opens the way to more extensive conformational averaging, which can often be necessary, even for fairly well-defined structures.

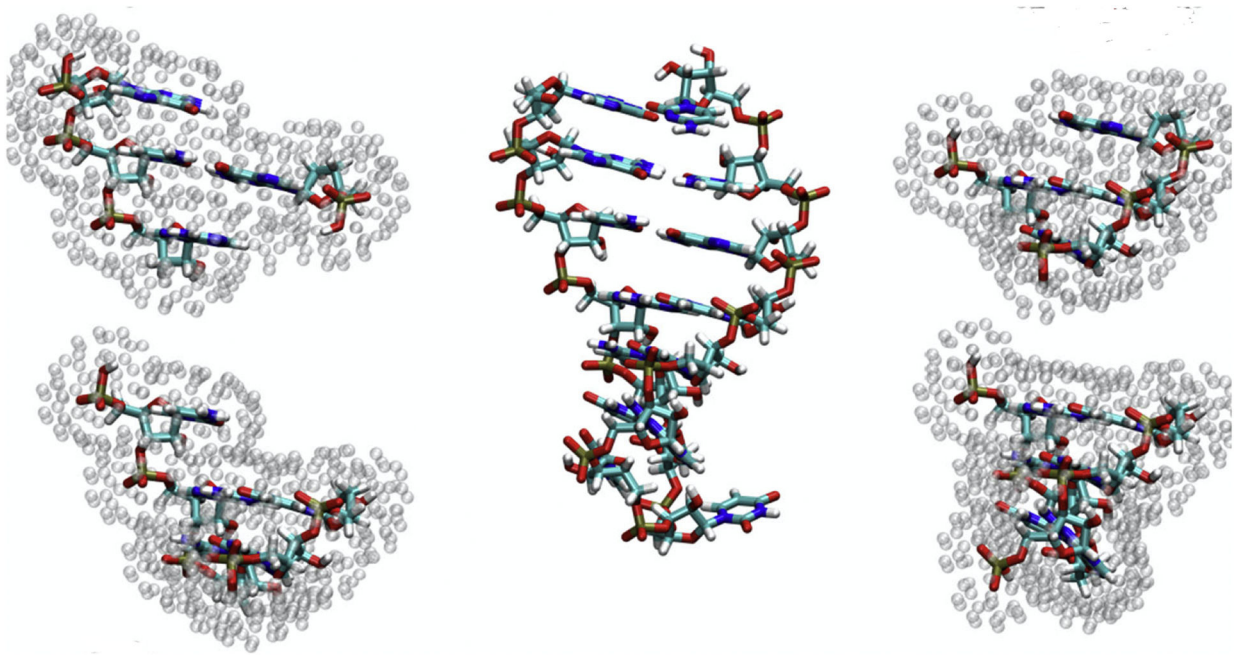

Figure 1:

AFNMR model for an RNA hairpin. The solute (RNA) molecule is illustrated in the center. A local fragment is constructed centered on each nucleotide (or amino acid for proteins); four such fragments are illustrated at on the left and right. The atoms shown for each fragment form the quantum region, and point charges surrounding each fragment are used to represent the electrostatic effects arising from the remaining RNA atoms (which result from Amber charges embedded in the dielectric of $\varepsilon = 4$), and effects of solvent (represented by a dielectric $\varepsilon = 78$ and a Debye-Huckel ionic strength taken as 0.1M by default).

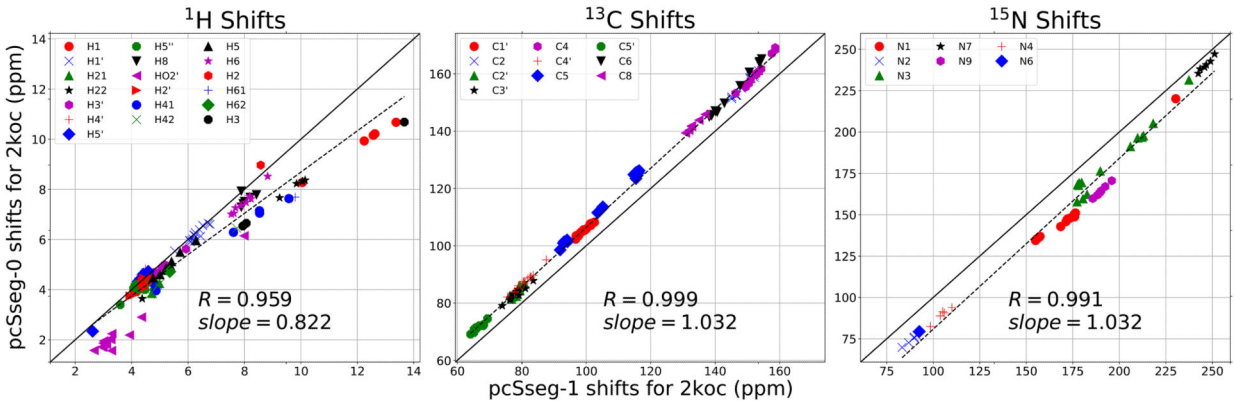

Figure 2:

Average shifts in the 10 NMR structures from pdb ID 2koc. Calculations carried out with the deMon 5.0 code, using the OLYP functional and either the pcSseg-0 or pcSseg-1 basis sets.

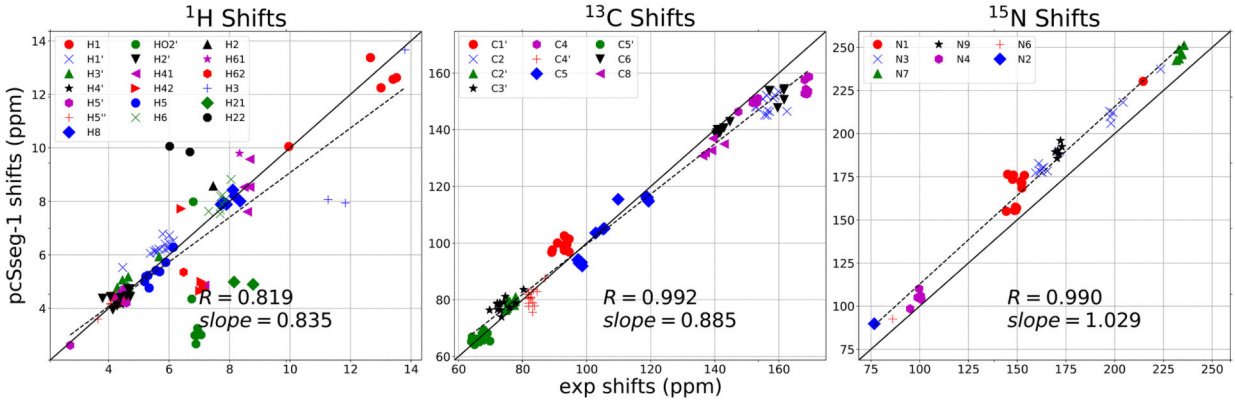

Figure 3:

Average shifts in the 10 NMR structures from pdb ID 2koc. Calculations carried out at the OLYP/pcSseg-1 level; experimental shifts from BMRB entry *bmr5705*.

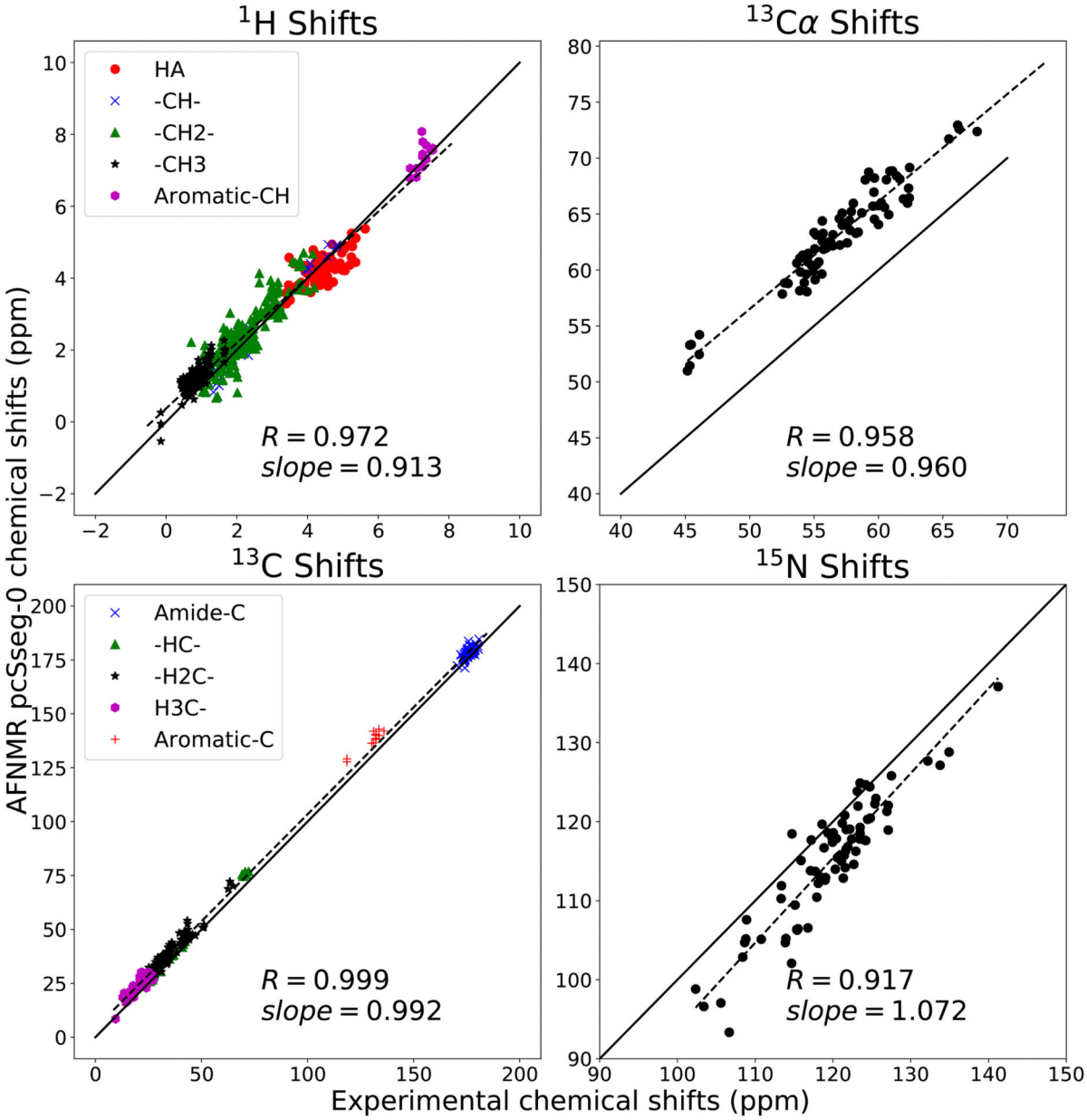

Figure 4:

Calculated vs. observed shifts for ubiquitin. Calculated values are the average shift for the ten models in PDB ID 1d3z, at the OLYP/pcSseg-0 level. Observed shifts are from BMRB entry *bmr17769*. Solid line is $y = x$; dashed line is a least-squares best-fit line,whose slope is shown. R is the Pearson correlation coefficient.

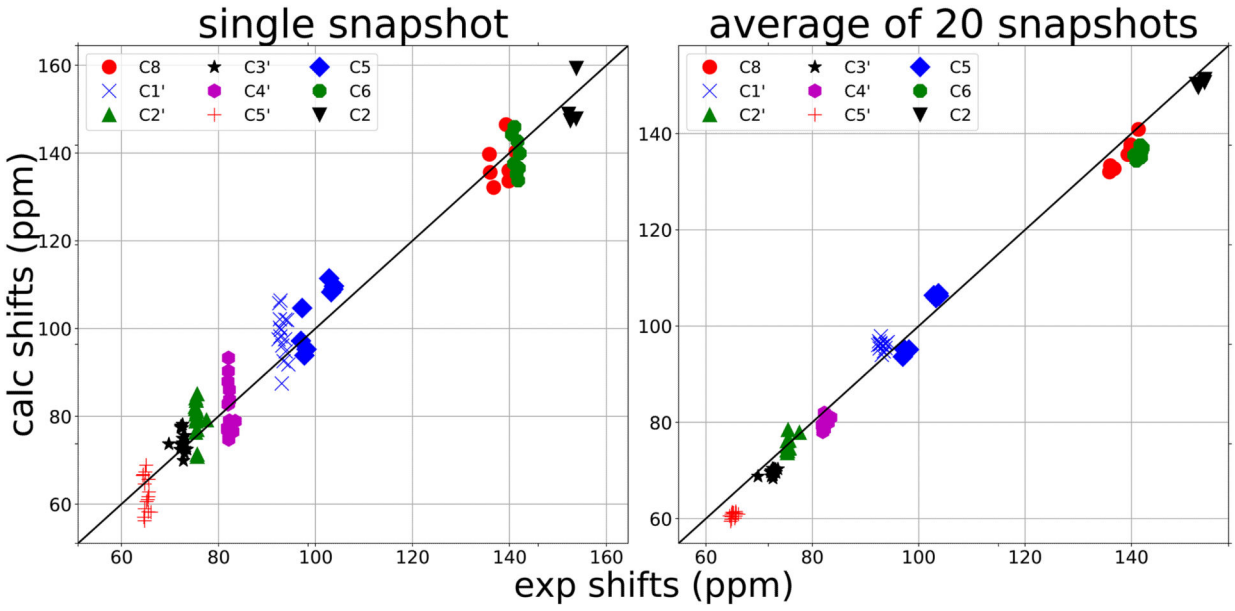

Figure 5:

Effects of conformational averaging on the ^{13}C shifts in 2gbh. Snapshots (including solvent waters and ions) are taken from Ref. [44]. Calculated shifts at the OLYP/pcSseg-0 level are shown for a single snapshot (*left*) and for the average over 20 snapshots spaced 0.5 nsec apart (*right*). Experimental shifts from Ref. [43].

Table 1:

Performance of various predictors for the 15N shifts in ubiquitin, using the first model from pdb ID 1d3z as the input structure. Comparisons are made to the observed shifts from BMRB entry *bmr17769*. R is the Pearson correlation coefficient, slope is the slope of the best fit line (shown as dashed line in Fig. 4), and RMSE is the root-mean-square difference between calculated and observed shifts.

Method	R	slope	RMSE
sparta+[9]	0.969	0.919	1.61
shiftx+[10]	0.953	0.892	1.98
OLYP/pcSseg-O	0.908	1.069	3.41
OLYP/pcSseg-1	0.916	0.982	3.33
OLYP/pcSseg-2	0.915	1.067	3.25
B3LYP/pcSseg-1	0.897	1.156	4.05