

Chemical Shift Prediction of RNA Imino Group and the Application in Characterizing RNA Transient States

Yanjiao Wang¹, Ge Han¹, Tairan Yuwen², Hashim M. Al-Hashimi^{3,4}, Yi Xue¹

¹Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences, Beijing Advanced Innovation Center for Structural Biology, Tsinghua University, Beijing 100084, China; ²Departments of Chemistry, Molecular Genetics, and Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada; ³Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710, United States; ⁴Department of Chemistry, Duke University, Durham, North Carolina 27710, United States

¹⁵N and ¹H resonances of amide group in proteins or imino group in nucleic acids are particularly important for biomolecular NMR study, as these resonances are nicely resolved and exquisitely sensitive to conformational variations. However, NH group is also the most challenging target for chemical shift prediction, mainly due to the fact that this group is often involved in diverse hydrogen bonding, as well as dynamic and transient interactions with solvent and ions. In this work we focus on chemical shift prediction for imino group of RNAs where reliable predictors are unavailable heretofore. As imino resonances come only from guanine and uridine in base-paired regions, we hypothesize that a base pair triplet could be sufficient for determining all chemical shifts of the central base pair including the imino group of interest. In BMRB database the experimental data of imino chemical shift are deficient and sometimes are deteriorated by mis-assignment or other mistakes. To address this issue we prepared dozens of RNA hairpins containing all combinations of base pair triplet composed of GC, UA, and GU base pairs, and collected chemical shift data for imino group. In the end we established a lookup table which relates experimental imino chemical shifts to the corresponding base pair triplet, and we found that this table can be used to predict imino chemical shifts of RNAs with unprecedented accuracy (Fig. 1). More strikingly, by calculating ring-current shifts of the two flanking base pairs, we observed good correlation between imino chemical shifts of the central base pair predicted by the lookup table and the calculated ring-current shifts caused by bases of the two flanking base pairs, indicating that ring-current shift is the dominating factor for chemical shift variations among all base pair triplets (Fig. 2). This finding opens a new avenue for predicting NH and even CH chemical shift of more complicated motifs in RNAs, including those involving noncanonical base pairs, bulges and loops.

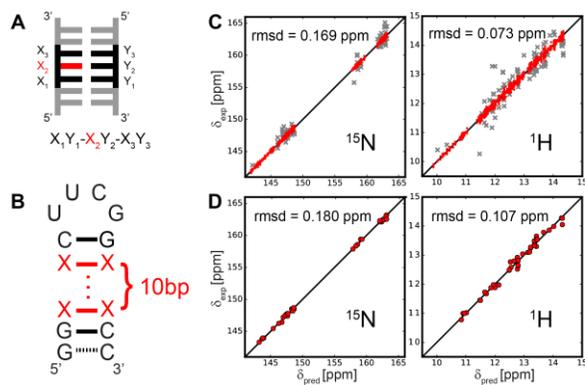


Figure 1. Base pair triplet (A); Hairpin RNAs that are used for collecting experimental imino chemical shift data (B); Correlation between predicted and experimental imino chemical shifts for RNAs in the training set (C) and the test set (D).

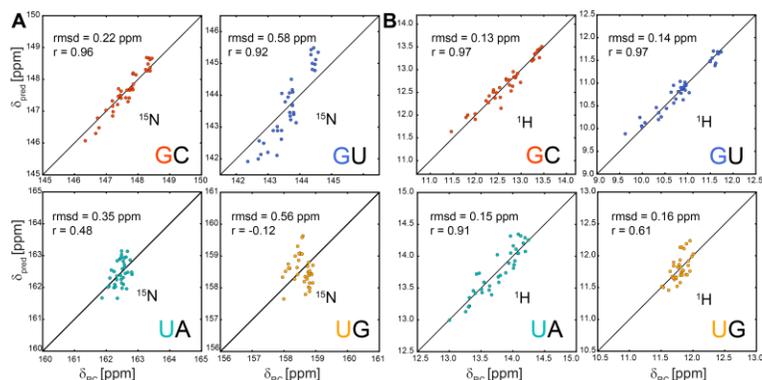


Figure 2. Correlation between chemical shifts obtained by ring-current calculation and those from our predictor for imino ¹⁵N (A) and imino ¹H (B) in various base pair types.

The new imino chemical shift predictor based on the lookup table has multiple potential applications, such as facilitating imino resonance assignment and RNA secondary structure determination. Here we demonstrated an application of this tool in secondary structure determination of RNA transient states by combining NMR relaxation dispersion (RD) experiments targeting at ¹⁵N and ¹H of imino group in RNAs. Previous RD studies for RNAs often make use of ¹³C probe, and more recently ¹⁵N and nonexchangeable ¹H probes. Nevertheless, RD detection of imino proton in RNAs has not been reported yet. In the current work we applied ¹⁵N CEST and ¹H CEST experiments to a previously well characterized transient state of P5abc, a subdomain of the Tetrahymena group I intron ribozyme. The imino chemical shifts of transient state derived from RD experiments are in excellent agreement with results from our chemical shift predictor. In addition, a previous speculated non-native G•G base pair in transient state has been confirmed by the new approach.