Discriminating Native from Non-Native RNA Structure Using Unassigned Chemical Shift Data

Toward Rapid RNA Structure Elucidation

William Fuh and Aaron T. Frank

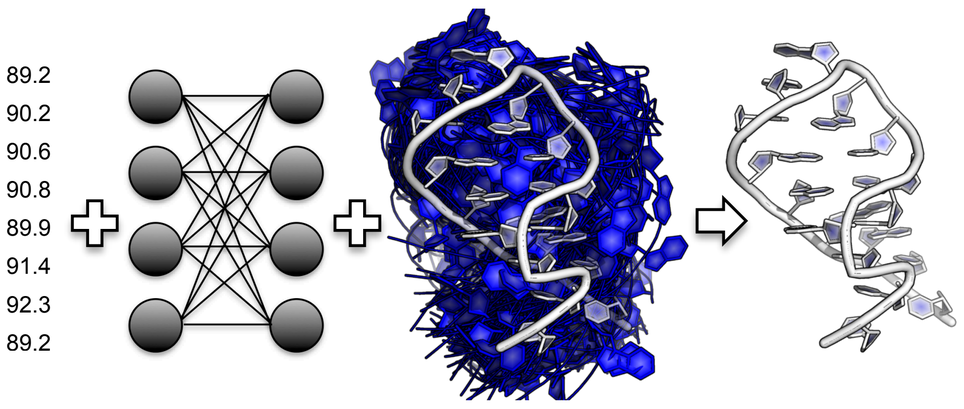
Department of Chemistry and Department of Biophysics, University of Michigan, Ann Arbor, Michigan, 48109

Abstract

In this work, we demonstrate that unassigned NMR chemical shifts, that is, an “anonymized” list of observed chemical shift peak values, can be used to identify native-like structures of RNAs. We achieve this by casting the problem of “assigning” unassigned chemical shift peaks to specific sites in an RNA as a linear assignment problem, and then solving it using the fast Kuhn–Munkres bipartite matching algorithm – also commonly referred to as the Hungarian algorithm. Using our assignment method, which we refer to as SCAHA (Structure-Based Chemical-Shift Assignment via the Hungarian Algorithm), we found that given an accurate structure of an RNA, and chemical shift data free of referencing errors, unassigned chemical shift peaks can be assigned to specific sites in the RNA with an accuracy of about ~0.2 and ~1.0 ppm for non-exchangeable ^{1}H and ^{1}C nuclei, respectively. By comparing SCAHA-assigned chemical shifts to computed chemical shifts, we demonstrate that we can identify, with high sensitivity, native-like structures in conformational pools that contain both native- like and non-native structures. Our results suggest that hybrid methods that combine state-of-the-art structure prediction methods with accurate structured-based assignment methods, like SCAHA, will soon enable RNA structure to be rapidly elucidated from unassigned NMR spectra.

Introduction

- Recent work has demonstrated that assigned chemical shift data can, in several cases, be used to predict the structure of non-canonical motifs of RNA with near atomic accuracy.
- However, acquiring assigned chemical shift data, especially for medium-sized to large RNAs, can be both time consuming and expensive.
- As such, there is immense interest in developing methods that enable multi-resolution structural information to be extracted from ”unassigned” chemical shift data.
- Can unassigned chemical shift be really discriminate between native and non-native structures of an RNA?



- We developed novel approach, SCAHA (**S**tructure-Based **C**hemical Shift **A**ssignment via the **H**ungarian **A**lgorithm) and used it to test the hypotheses that:

(1) native-like structures exhibit the lower assignment errors than non-native decoys and

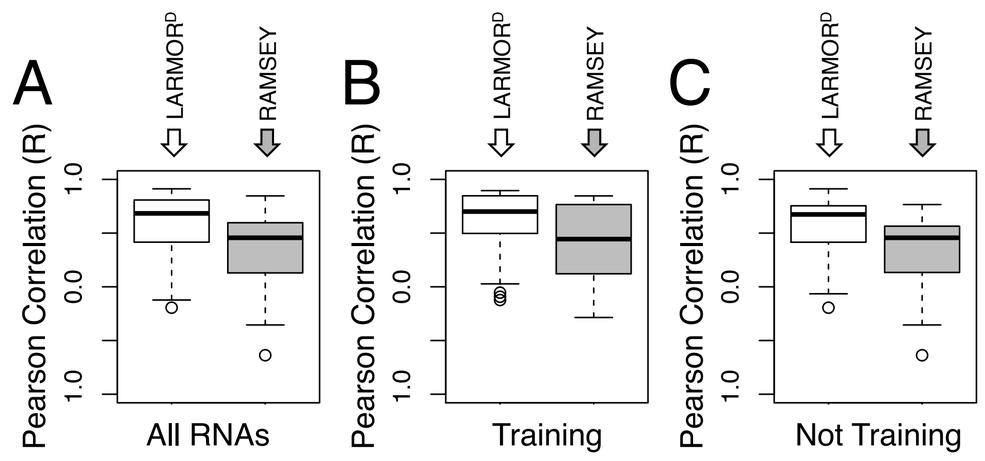
(2) native-like structures exhibit the lower errors between the optimally “assigned” chemical shifts and computed chemical shifts than non-native decoys.

Results

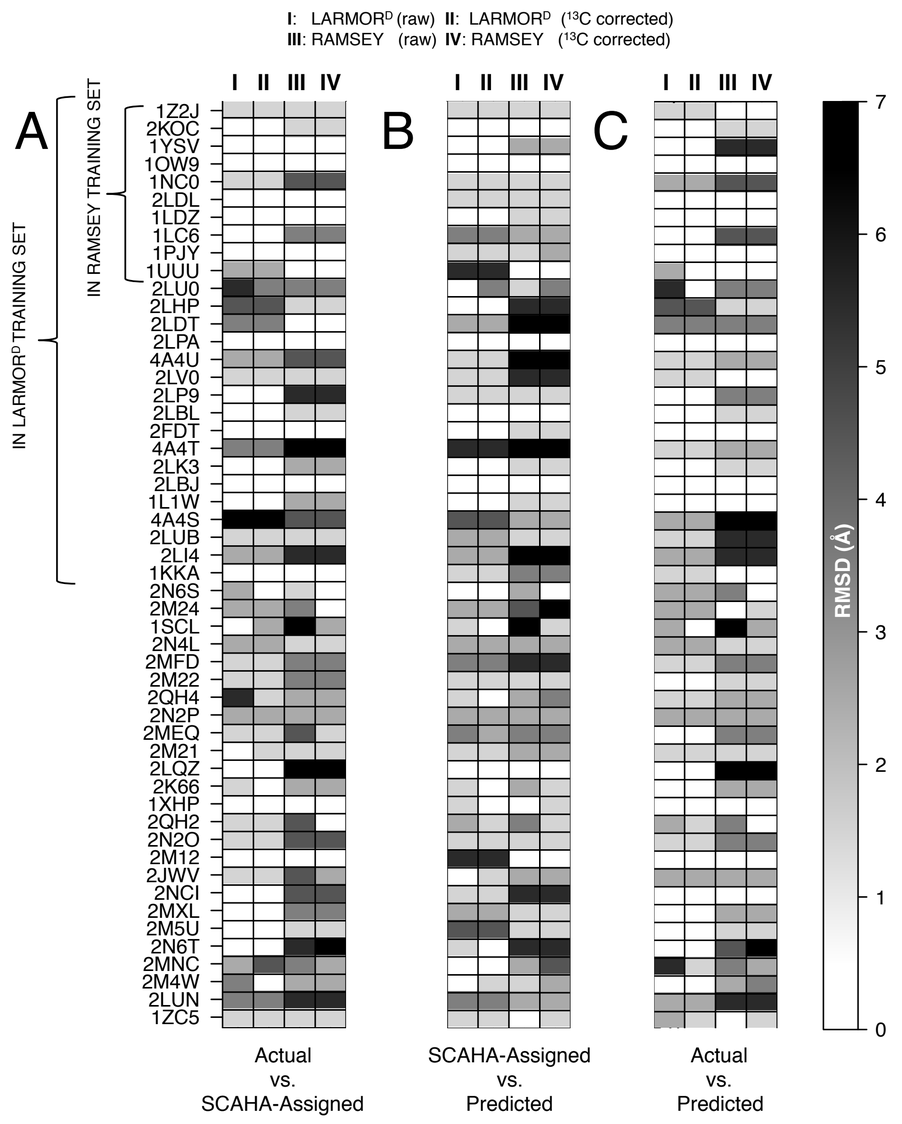
SCAHA can assign chemical shifts to with ~0.20 and~ 1.0 ppm for ^{1}H and ^{13}C Non-exchangeable Nuclei

	$^{1}\mathbf{H}$		$^{13}\mathbf{C}_{\text{corrected}}$	
dataset	$\Delta\delta$ (ppm)	IQR (ppm)	$\Delta\delta$ (ppm)	IQR (ppm)
all	0.15/0.16	0.06/0.04	0.84/0.77	0.76/0.20
training	0.14/0.16	0.03/0.01	0.71/0.74	0.29/0.66
not training	0.17/0.16	0.08/0.06	1.13/0.78	0.73/0.13

SCAHA assignment errors are positively correlated with RMSD



Low Error Structures tend to Native-Like



Take Away

The structures that exhibited the lowest assignment errors tended to be native-like, as were the structures that exhibited the lowest errors between the optimally “assigned” chemical shifts and computed chemical shifts.