On the Automated Determination of pKa by NMR in DMSO:Water Mixtures

Erik N. A. Sundén*, Staffan Karlsson*, Christoph Bauer, Per-Ola Norrby

Supporting Information

This document contains a description of the data processing and detailed example results for histidine that illustrate how we have obtained pKA values in an automated fashion.

Contents

peaKaOracle: Data Processing Method Overview

The algorithm for peaKaOracle can be described in several steps:

1. Reading in a series of NMR spectra. This requires the knowledge of the data structure. Magritek uses an openly readable format.
2. Apodization and zero filling of the data using nmrglue in python and conversion into ppm units
3. Phasing the signal using hard-coded solvent regions in ppm units:
   i. `water = [5.5,4.5]`
   ii. `dmso = [2.9, 2.5]`
4. Peak picking using scipy in python with the initial parameters of height=500 and prominence=500 with a distance of at least 0.2 ppm between peaks – this will heavily depend on the data set.
5. Grouping the peaks and sorting the peak groups over the series of data
6. Manual corrections to the automatically curated peak groups. This is usually only necessary for ill-behaved signals displaying poor S/N, resolution or a changing multiplet pattern.
7. Fitting the pKA values according to the equations in the main paper.

Example Result: Histidine

Figure S1 shows how we initially assign the peak groups over the course of the titration. The algorithm ignores the solvent region. There are minor peaks that are detected, e.g., the peak group labeled “4” at pH 7.12. Figure S2 shows the improvement that the curation yields. Four groups of peaks form the basis for the pKA fitting procedure. We obtain the final values from the curves displayed in Figure S3, where the four characteristic signals are traced, and their response is related to the closeness of the respective titratable group.

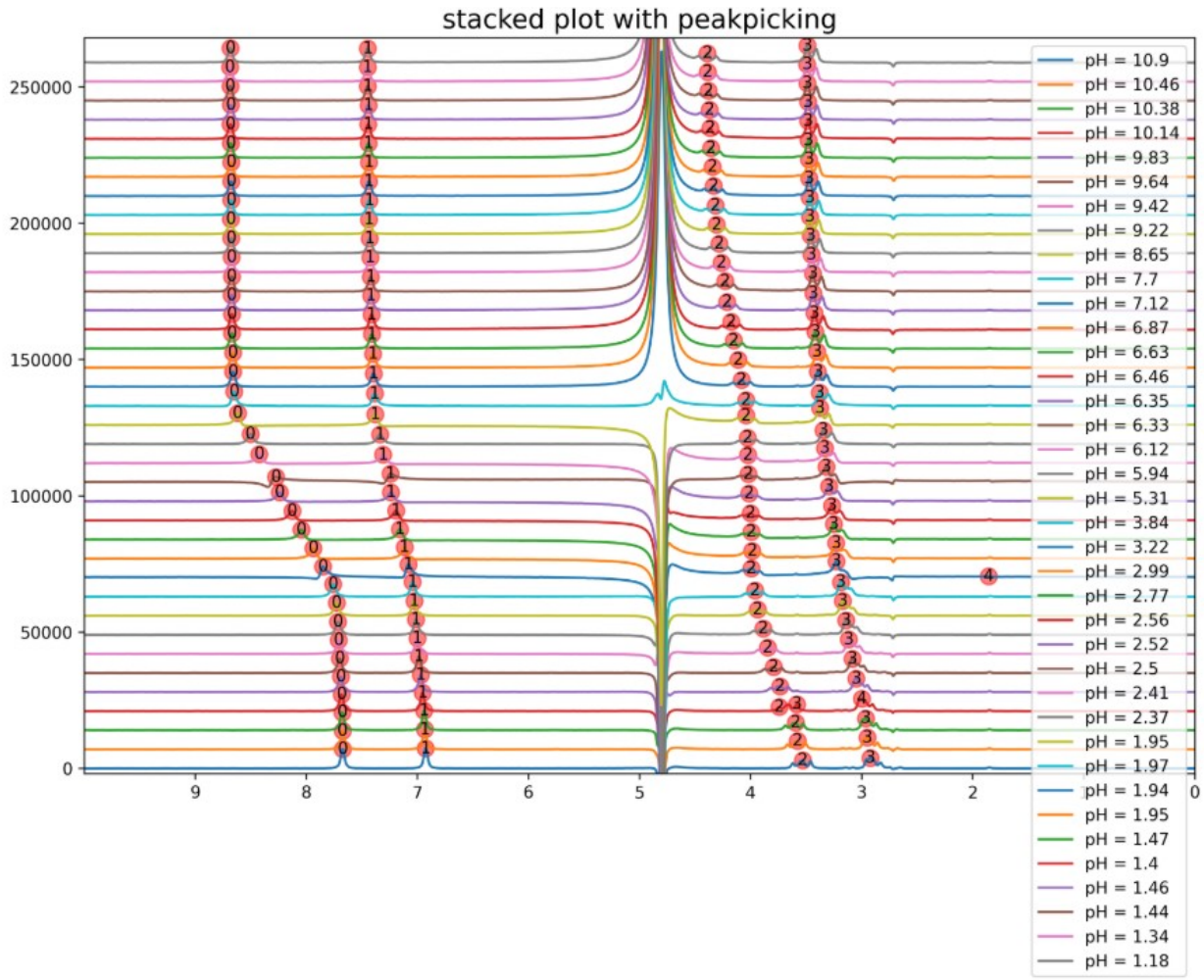

Figure S 1: Initial result of the peak picking and grouping for the histidine data set. The x-axis is in units of ppm. The lowest line represents the alkaline environment at pH 10.9.

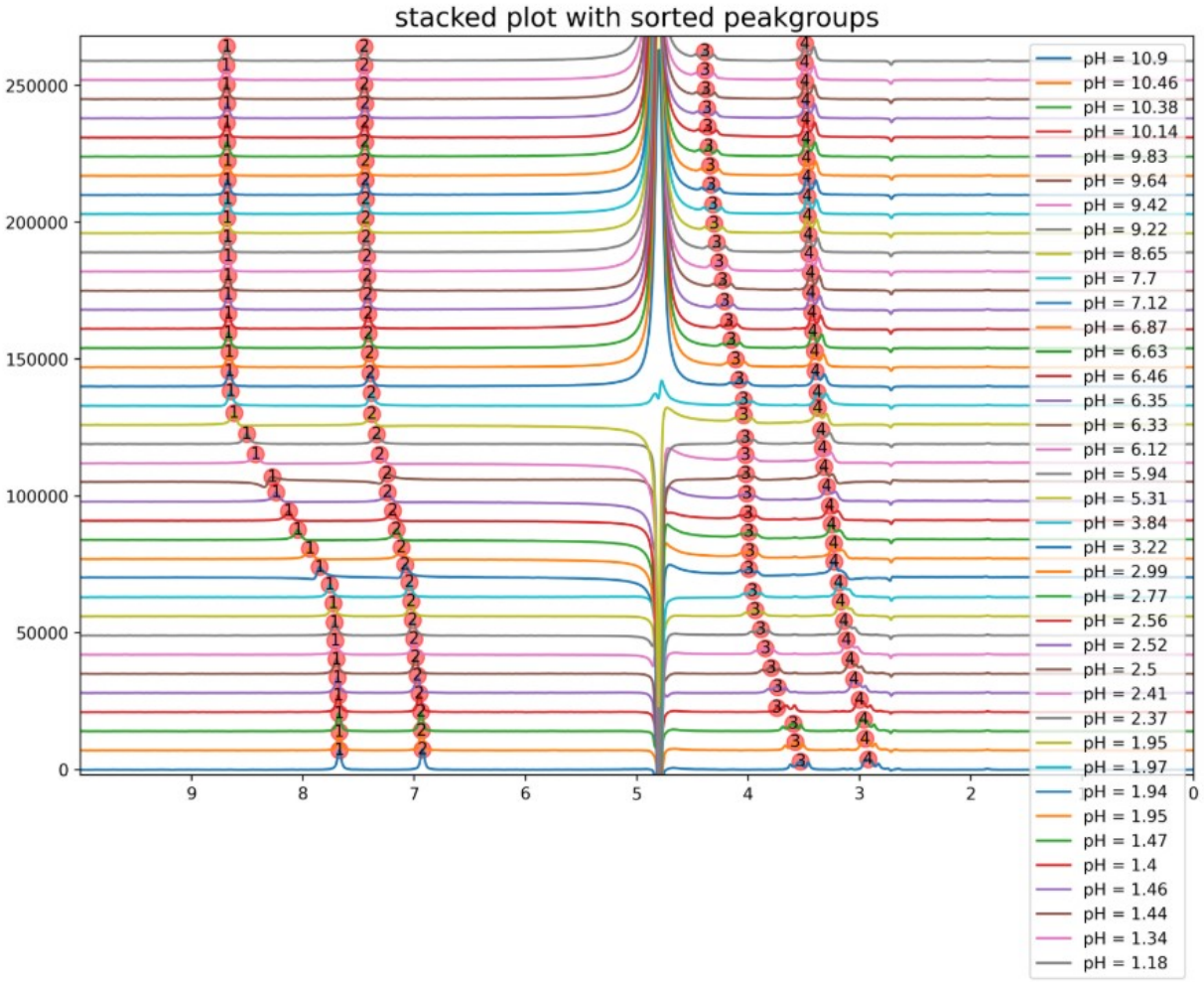

Figure S 2: Final result of the peak picking and grouping for the histidine data set. The x-axis is in units of ppm. The lowest line represents the alkaline environment at pH 10.9.

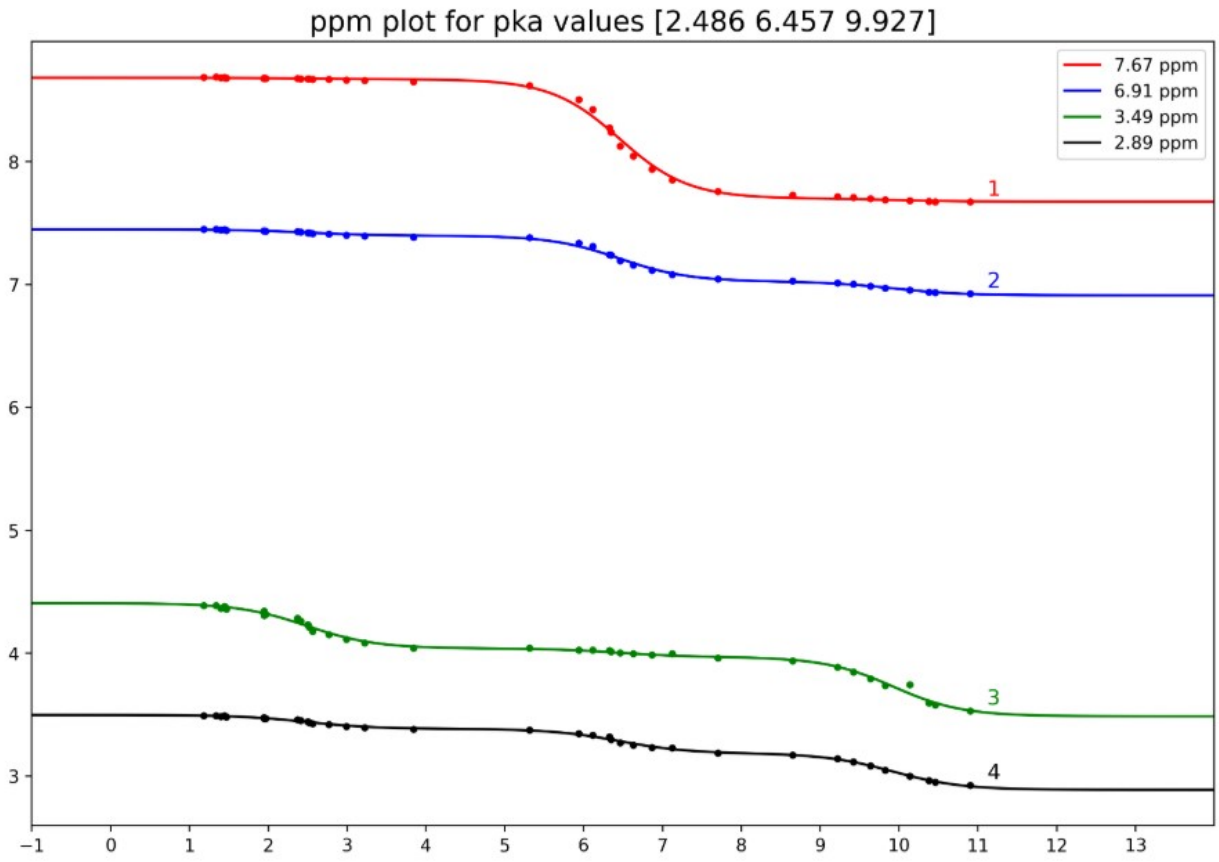

Figure S 3: Curves to fit pKA values for histidine. The x-axis represents the pH. The y-axis shows how the chemical shifts change in units of ppm.