NMR Spectroscopy Database and Searching System

Haw-Lih Su¹, Huei-Wen Winsome Teng²,³, Jeff Cheng-Lung Lee⁴, Mohammed Shkoor⁵, Mohammad Ibrahim Ahmad Ibrahim¹, Mohammed Hussain S A Alsafran¹,⁶ and Saeed H. Al-Meer¹,⁵

¹ Central Laboratories Unit, Qatar University, Doha, Qatar
² Department of Information Management and Finance, National Chiao Tung University, Hsinchu, Taiwan
³ Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan
⁴ Taiwan Police College, Taipei, Taiwan
⁵ Department of Chemistry and Earth Sciences, Qatar University, Doha, Qatar
⁶ Department of Biological and Environmental Sciences, Qatar University, Doha, Qatar

hsu@qu.edu.qa

Abstract
NMR spectroscopy is the most important analytical technology for organic compounds and plays the key role for the chemical characterizations and identifications in chemistry, pharmacy, materials science, environments, biology, and many related fields. In most cases, a NMR spectrum is compared with the known spectra to check if the sample contains a known chemical or is a new product. However, the current comparison process relayed on human beings and waste a lot of time and efforts. A new database and searching system is thus in need.

Typical Presentation Methods for a NMR Spectrum
A. Graphic Presentation: Good for human reading

B. Numerical Description: Good for digital searching

9-bromo-2,4-dimethyl-5-oxo-5H-chromeno[3,4-c]pyridine-1-carbonitrile

¹H NMR (CDCl₃, 600 MHz): δ 2.94 (s, 3H), 3.09 (s, 3H), 7.27 (d, 1H, J = 8.9 Hz), 7.75 (dd, 1H, J = 8.9 Hz, 2.0 Hz), 9.34 (d, 1H, J = 2.0 Hz). ¹³C{¹H} NMR (CDCl₃, 150 MHz): δ 25.2, 27.9, 100.0, 113.2, 116.5, 117.1, 117.8, 119.6, 128.4, 137.1, 142.6, 151.5, 157.6, 167.8, 168.0.

Current Databases and Literature
- Databases: Mostly in graphic; some in both. Only Searching by names and formula. Contains common chemicals only.
- Literature: Mostly in numerical; some has graphics in supporting information.
- Researchers need to search and compare by themselves. No convenient or quick searching methods.

Database Facilitated Screening and Chemical Identification Process

- ¹⁵N & ¹⁹F NMR Spectroscopy
- 2D correlation spectra, e.g. COSY, HMQC, HMBC, NOESY, DOSY, etc.
- GC , GC-MS, LC-MS, HPLC

- Unknown sample submissions
- ¹H and ¹³C NMR Spectra
- Database screening
- Percentage of matching for the preliminary judgement
- Additional analyses for confirmation
- Final judgement of the structure