The BMRB archive of Protein, Nucleic Acid and Metabolite NMR Data

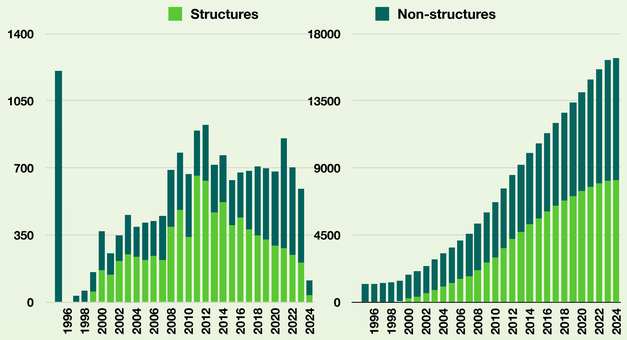
Kumaran Baskaran, Jonathan R. Wedell, Hongyang Yao, Dimitri Maziuk, Hamid R. Eghbalnia, Michael M. Gryk, and Jeffrey C. Hoch

Department of Molecular Biology and Biophysics, UConn Health,
263 Farmington Avenue, Farmington, CT 06030, USA

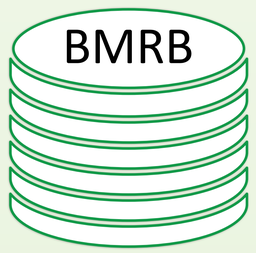
The Biological Magnetic Resonance Data Bank[1] (BMRB: https://bmrb.io) serves the biomolecular NMR community by supporting a curated archive of primary and derived data and metadata linked to scientific investigations under the "FAIR Principles" (Findable, Accessible, Interoperable, and Reusable)[2]. BMRB is a member and a core archive of the Worldwide Protein Data Bank (wwPDB: https://www.wwpdb.org), which collects the coordinate data along with assigned chemical shifts and restraints from NMR structural studies through the OneDep (https://deposit.wwpdb.org) deposition system. NMR data from other studies are collected by the BMRBdep (https://deposit.bmrb.io) deposition system. BMRB's goal is to empower scientists in their analysis of the structure, dynamics, and chemistry of biological systems and to support further developments in the field of biomolecular NMR spectroscopy.

BMRB growth statistics

The bar graph shows the growth of the BMRB archive. As of March 2024, BMRB holds 16332 entries with 8180 entries having corresponding coordinate data in the PDB archive and 8152 BMRB only entries.



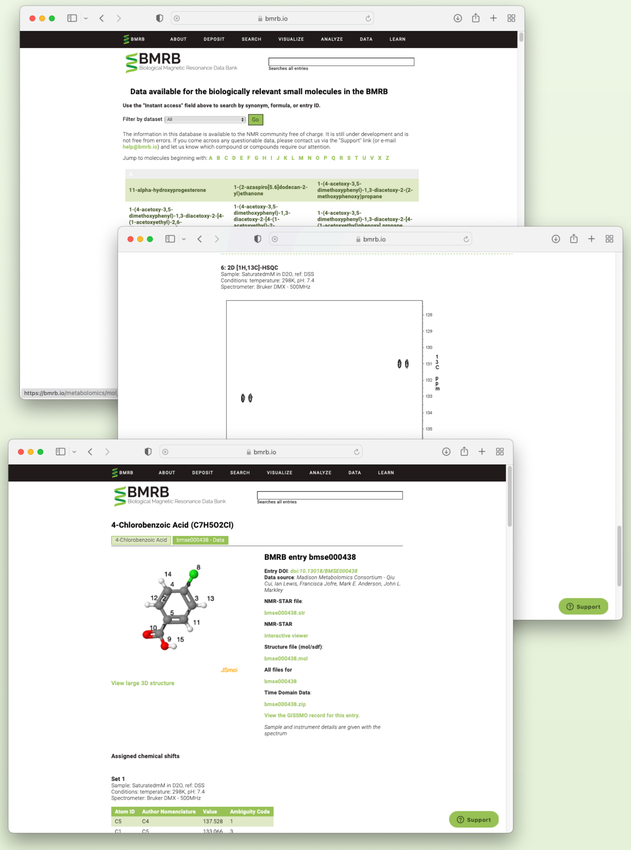
Data access



- **F**ind data using the instant search or advance search
- **A**ccess through web browser or BMRB API
- **I**nteroperate with the help of PyNMRSTAR
- **R**euse/reinterpret with new software tools

Small molecules library

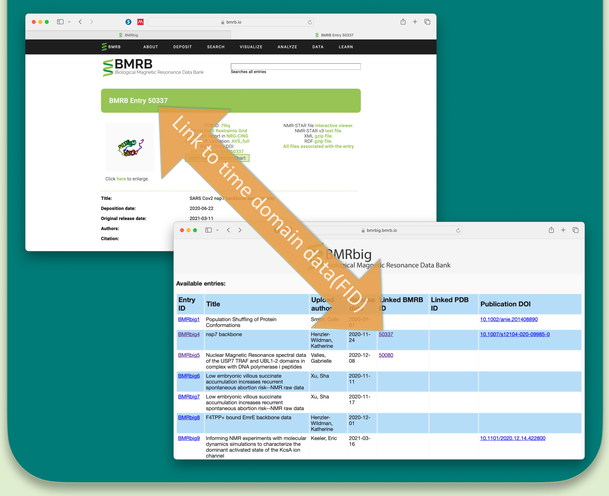
BMRB maintains a library of carefully curated NMR spectroscopic data of over 1000 biologically important small molecules.



New

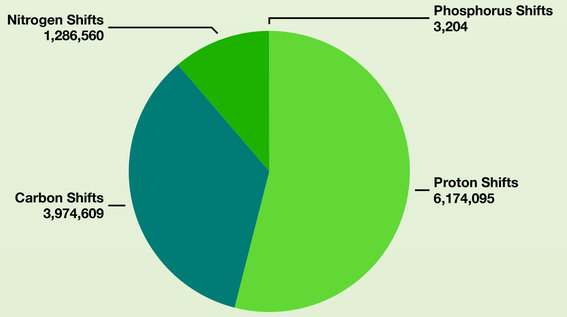
Time Domain Data

BMRB invites depositors of existing entries in BMRB to upload the time domain data that support their BMRB entries. A simple web interface will be available on the main web page for depositors to upload the time domain data to BMRbig and link them to the corresponding BMRB entry. A community effort to increase the amount of time domain data in BMRB holds the promise of further enabling applications of machine learning in structural biology.

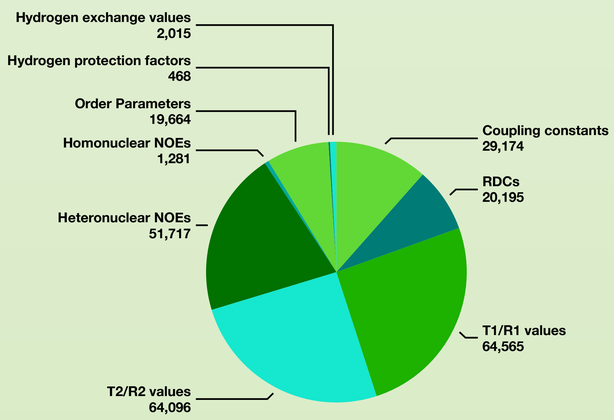


BMRB data content (chemical shifts)

As of March 2024, BMRB holds over 11.4 million chemical shifts from 15210 protein, 572 DNA and 540 RNA entries. In addition to the chemical shift data, other NMR data are also available at the BMRB



BMRB data content (other NMR data)



NMR-STAR data model

NMR-STAR[3] is the official data format of BMRB and the deposition and archival format of NMR spectroscopic data at the wwPDB. It is a Self-defining Text Archive and Retrieval (STAR) format with controlled vocabularies(tags) defined by the NMR-STAR dictionary. The NMR-STAR data model supports many kinds of NMR data, metadata and derived data.

PDF version of the poster

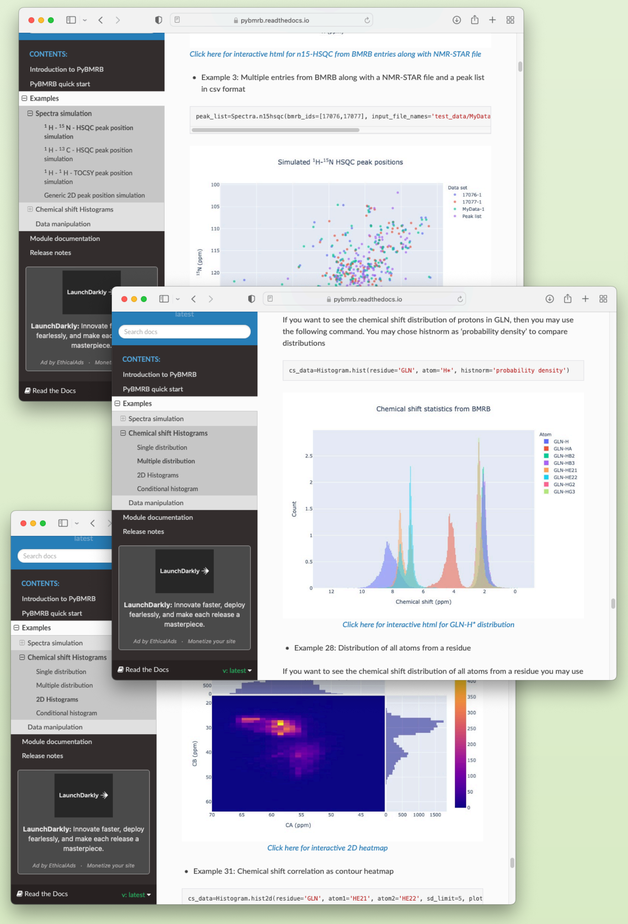
BMRB Software resouce

BMRB GitHub: https://github.com/bmrb-io

- BMRB - API : provides machine to machine access to BMRB data base
- PyNMRSTAR : Python NMR-STAR parser
- PyBMRB: BMRB data visualization tool in Python
- RBMRB: BMRB data visualization tool in R

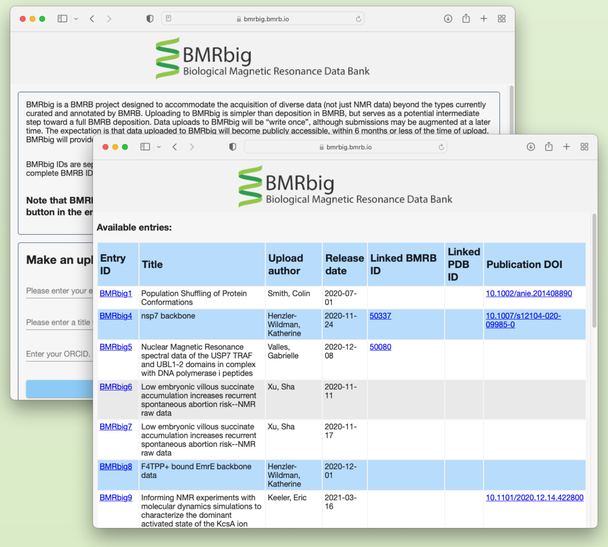
BMRB data visualization

Chemical shift histograms and simulated HSQCs can be generated using PyBMRB or RBMRB.



BMRBig

BMRbig (https://bmrbig.bmrb.io/) is a BMRB project designed to accommodate the acquisition of diverse data (not just NMR data) beyond the types currently curated and annotated by BMRB. Uploading to BMRbig is simpler than deposition in BMRB, but serves as a potential intermediate step toward a full BMRB deposition. Data uploads to BMRbig will be "write once", although submissions may be augmented at a later time.



1. *Hoch, J. C.; et al. Nucleic Acids Research 2023, 51*
2. *Wilkinson, M. D.; et al. Scientific data 2016, 3, 160018.*
3. *Ulrich, E. L.; et al. J. Biomol. NMR 2019, 73 (1-2), 5.*

BMRB is supported by NIGMS through grant R24GM150793