EDITORIAL

The indelible mark of computation on bio-NMR

Jeffrey C. Hoch · Martin Billeter

Computation played an enabling role in the birth of modern NMR spectroscopy, when Ernst and Anderson proposed using the discrete Fourier transform (DFT) to compute the spectrum from measurements of the free induction decay (FID). In retrospect their proposal for speeding up NMR was even more audacious than it appeared at the time, because they were unaware of the nearly contemporaneous publication of the Cooley-Tukey fast Fourier transform algorithm (personal communication from R. Ernst to J. H.). With the proposal by Jeener to parametrically sample indirect time dimensions by recording multiple FIDs, the role of computation for obtaining multidimensional spectra was even more firmly entrenched. Invention followed necessity as the ability to resolve individual resonances for complex biomacromolecules that was enabled by multidimensional experiments led to the development of software for computer visualization and analysis of complicated spectra. Systematic use of nuclear Overhauser effects for solving the three-dimensional structure of a protein relied on software for optimizing the fit to an underdetermined set of distance restraints. Richard Ernst summarized the impact of computation in bio-NMR most succinctly in the title of his opening lecture at a 1990 NATO Advanced Research Workshop: "Without Computers, No Biological NMR". In the proceedings volume, he would broaden this assertion to "No Modern NMR" (Ernst 1991).

Fast forward to today, where the depth and breadth of computer applications in bio-NMR is nothing short of staggering. Spectrum analysis has moved beyond the DFT to a host of non-Fourier methods that are capable of "super-resolution" from short data records, and are adept at handling non-uniformly sampled data. Quantum calculations of magnetic properties for large systems have advanced to useful accuracy. Simulations of spin system dynamics permit optimization of pulse shapes and sequences a priori. Structure computations can now be performed using a host of ambiguous experimental constraints including NOEs, chemical shifts, vicinal couplings, and residual dipolar couplings (RDCs). Meta-analyses made possible by public data repositories have facilitated novel insights, especially in the realm of structural correlates of chemical shifts. Model fits to relaxation data permit atomic-level descriptions of molecular dynamics, and reveal details of transient or low-population conformational states. Sophisticated statistical analyses are enabling NMR spectra of complex mixtures to reveal new insights into metabolism and identify disease biomarkers. Docking computations driven by NMR data are now a fixture of modern structure-aided drug discovery.

This special issue of the Journal of Biomolecular NMR presents a collection of papers that capture just a small window into the broad range of computational methods being applied in bio-NMR, but nevertheless give a vivid indication of the breadth of modern applications. Non-uniform sampling, coupled with non-Fourier methods of spectrum analysis, have proven a powerful combination for reducing the time required to perform multidimensional experiments. Somewhat less appreciated is their potential for improving sensitivity. Rovnyak and colleagues have

J. C. Hoch (✉)
Department of Molecular Biology and Biophysics, University of Connecticut Health Center, 263 Farmington Avenue, Farmington, CT 06030-3305, USA
e-mail: hoch@uchc.edu

M. Billeter
Department of Chemistry and Molecular Biology, University of Gothenburg, Box 462, 405 30 Göteborg, Sweden
e-mail: martin.billeter@chem.gu.se

devoted considerable attention to this prospect, and report their recent findings. Fitting dynamical and structural models to relaxation data has lead to novel insights across the full spectrum of biomolecular problems. Luchinat, Parigi, and Ravera use this approach to explore aspects of paramagnetism-induced relaxation to address the question of whether metal ions can be used to elicit dynamic nuclear polarization. Efficient matrix codes have enabled numerous applications of parallel factor analysis (and related methods of principle component analysis and multiway decomposition) in NMR and analytical chemistry. Nilsson and colleagues exploit the power of these methods for analyzing complex mixtures with diffusion-ordered spectroscopy. Nuclear magnetic relaxation rates and RDCs are both sensitive to molecular dynamics; Lee and colleagues describe a unified approach to characterizing protein dynamics that encompasses relaxation data and RDCs simultaneously. Finally, Vuister and colleagues provide a comprehensive overview of computational tools for validating protein structures based on NMR data.

It bears emphasizing that many of these applications are only now feasible because of remarkable advances in computing technology that have occurred in the past 30 years: desktop computers in 1984 were capable of about 100 K floating-point operations (flops) per second; modern desktops now achieve more than one teraflop. Looking forward, it is not difficult to contemplate that additional advances will similarly enable new applications of computers in bio-NMR that we can barely imagine today.

The advances in computational aspects of bio-NMR have been the work of many individuals, but we take special note of the recent passing of Flemming Martin Poulsen. Flemming held a very practical view of computers, recognizing that a utilitarian role was nonetheless truly enabling in the face of the data explosion associated with multidimensional NMR experiments on proteins. Among his contributions involving computation were methods for symmetry recognition to simplify multiplets in multidimensional spectra (Hengyi et al. 1990; Hengyi and Poulsen 1990, 1992; Hoch et al. 1987), and a comprehensive system for visualization and assignment of protein spectra (Kjaer et al. 1994). He was a co-organizer of a NATO workshop on computational aspects of bio-NMR in 1990, and a regular contributor to this journal. Among his contributions to JBNMR are highly-cited works on measurement of coupling constants (Madsen et al. 1993), determination of phi-angles (Ludvigsen and Poulsen 1992), and chemical shifts of intrinsically disordered or denatured proteins (Kjaergaard et al. 2011; Kjaergaard and Poulsen 2011). Although cancer claimed him prematurely, he leaves behind a vibrant Copenhagen legacy of computation in bio-NMR.

References

Ernst RR (1991) In: Hoch JC, Poulsen FM, Redfield C (eds) Computational aspects of the study of biological macromolecules by nuclear magnetic resonance spectroscopy. Plenum Press, Berlin, pp 1–25

Hengyi S, Poulsen FM (1990) Toward automated determination of buildup rates of nuclear overhauser effects in proteins, using symmetry projection operators. J Magn Reson 89:585–594

Hengyi S, Poulsen FM (1992) Group theoretical projection operators and symmetry-adapted functions in NMR spectral analysis. J Magn Reson 97:385–390

Hengyi S, Ludvigsen S, Poulsen FM (1990) Application of symmetry projection operators to measurements of coupling constants in 2D NMR spectra of proteins. J Magn Reson 90:346–354

Hoch JC, Hengyi S, Kjaer S, Poulsen FM (1987) Symmetry recognition applied to two-dimensional NMR data. Carlsberg Res Commun 52:111–122

Kjaer M, Andersen KV, Poulsen FM (1994) Automated and semiautomated analysis of homo- and heteronuclear multidimensional nuclear magnetic resonance spectra of proteins: the program Pronto. Methods Enzymol 239:288–307

Kjaergaard M, Poulsen FM (2011) Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. J Biomol NMR 50:157–165. doi:10.1007/s10858-011-9508-2

Kjaergaard M, Brander S, Poulsen FM (2011) Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. J Biomol NMR 49:139–149. doi:10.1007/s10858-011-9472-x

Ludvigsen S, Poulsen FM (1992) Positive theta-angles in proteins by nuclear magnetic resonance spectroscopy. J Biomol NMR 2:227–233

Madsen JC, Sorensen FM, Sorensen P, Poulsen FM (1993) Improved pulse sequences for measuring coupling constants in 13C, 15N-labeled proteins. J Biomol NMR 3:239–244