Characterization of Biotherapeutics by Chemometric and Machine Learning Analysis of NMR Spectra

Frank Delaglio

Protein therapeutics are vitally important clinically and commercially, with monoclonal antibody (mAb) therapeutic sales alone accounting for $115 billion in revenue for 2018. In order for these therapeutics to be safe and efficacious, their protein components must maintain their three-dimensional fold and not aggregate. Analytical methods to characterize this higher order structure (HOS) can help establish comparability between drug products and provide impact throughout the life cycle of a therapeutic, from development to manufacturing. Heteronuclear Nuclear Magnetic Resonance (NMR) is a practical multi-attribute method that can be applied directly to intact biologics and their formulations to meet this important measurement need.

We have previously demonstrated practical approaches to generate two-dimensional $^{1}H/^{13}C$ NMR spectra at natural isotopic abundance for molecules as large as intact monoclonal antibodies, and the robustness and precision of these methods has recently been confirmed by an international multi-laboratory study with participation from 25 academic, pharmaceutical, governmental, and regulatory organizations.

Using measurements on the IgG1k NIST reference mAb (NISTmAb), we demonstrate that small variations of structure and interaction can be revealed and classified by direct computational analysis of the shapes of such spectra, as an alternative to time-consuming and subjective interactive analysis and assignment of spectral features, paving the way for NMR characterization of protein therapeutics via chemometrics and machine learning that are both objective and automated.