Progress in Nuclear Magnetic Resonance Spectroscopy 79 (2014) 48–68

Contents lists available at ScienceDirect

Progress in Nuclear Magnetic Resonance Spectroscopy

journal homepage: www.elsevier.com/locate/pnmrs

NMR of glycans: Shedding new light on old problems

Marcos D. Battistel, Hugo F. Azurmendi, Bingwu Yu, Darón I. Freedberg *

Laboratory of Bacterial Polysaccharides, Center for Biologics Evaluation and Research, Food and Drug Administration, 1401 Rockville Pike, Rockville, MD 20852-1448, United States

Edited by J.W. Emsley and J. Feeney

ARTICLE INFO

Article history:
Received 3 July 2013
Accepted 10 January 2014
Available online 14 February 2014

Keywords:
Glycan NMR
Hydroxyl groups
Hydrogen bonds
Isotopic labeling and 3D Structure
Residual dipolar coupling

ABSTRACT

The diversity in molecular arrangements and dynamics displayed by glycans renders traditional NMR strategies, employed for proteins and nucleic acids, insufficient. Because of the unique properties of glycans, structural studies often require the adoption of a different repertoire of tailor-made experiments and protocols. We present an account of recent developments in NMR techniques that will deepen our understanding of structure–function relations in glycans. We open with a survey and comparison of methods utilized to determine the structure of proteins, nucleic acids and carbohydrates. Next, we discuss the structural information obtained from traditional NMR techniques like chemical shifts, NOEs/ROEs, and coupling-constants, along with the limitations imposed by the unique intrinsic characteristics of glycan structure on these approaches: flexibility, range of conformers, signal overlap, and non-first-order scalar (strong) coupling. Novel experiments taking advantage of isotopic labeling are presented as an option for overcoming spectral overlap and raising sensitivity. Computational tools used to explore conformational averaging in conjunction with NMR parameters are described. In addition, recent developments in hydroxyl detection and hydrogen bond detection in protonated solvents, in contrast to traditional sample preparations in D_2O for carbohydrates, further increase the tools available for both structure information and chemical shift assignments. We also include previously unpublished data in this context. Accurate determination of couplings in carbohydrates has been historically challenging due to the common presence of strong-couplings. We present new strategies proposed for dealing with their influence on NMR signals. We close with a discussion of residual dipolar couplings (RDCs) and the advantages of using ^{13}C isotope labeling that allows gathering one-bond ^{13}C–^{13}C couplings with a recently improved constant-time COSY technique, in addition to the commonly measured ^{1}H–^{13}C RDCs.
Published by Elsevier B.V.

Contents

* Corresponding author.
E-mail address: daron.freedberg@fda.hhs.gov (D.I. Freedberg).

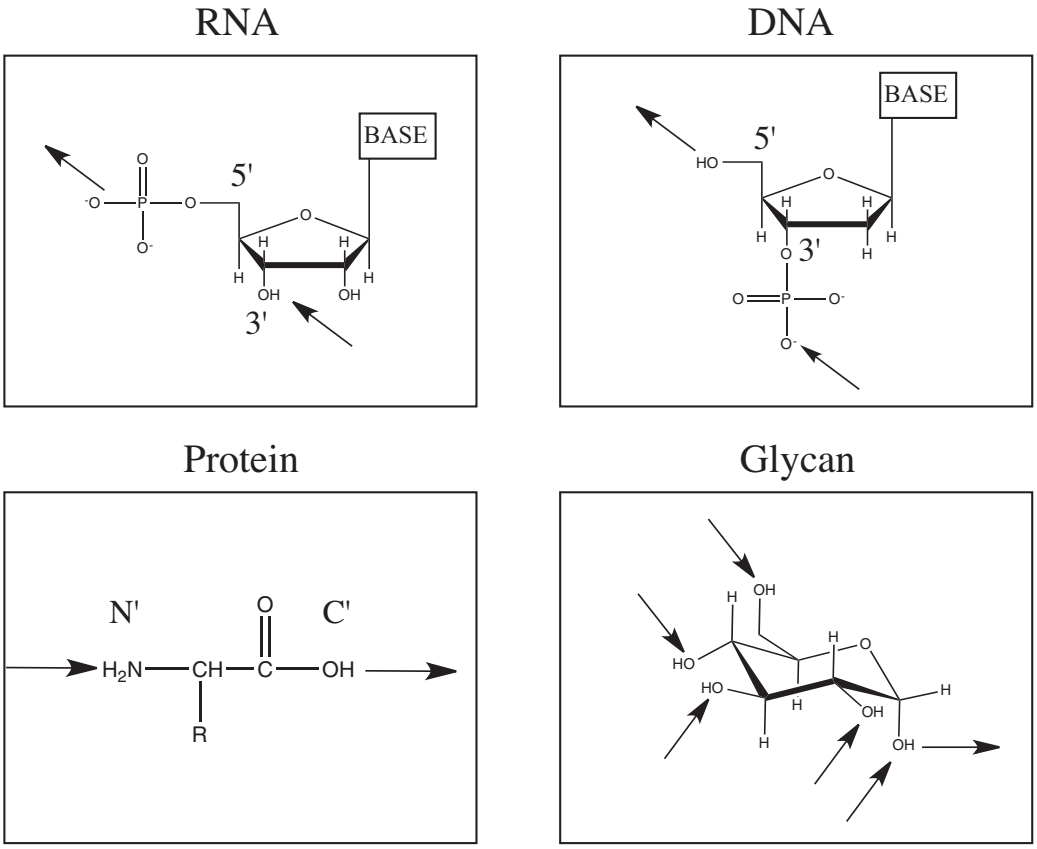
http://dx.doi.org/10.1016/j.pnmrs.2014.01.001
0079-6565/Published by Elsevier B.V.

1. Introduction

Carbohydrates are the most ubiquitous class of biomolecule in nature, yet we know little about their three-dimensional structures and more importantly, how their structures and flexibility translate into biological functions. In part, this is because the potential of carbohydrates to store chemical information, known as the "carbohydrate code" [1–3], is overwhelming. If each unique oligosaccharide sequence, and its corresponding three-dimensional structure, accounts for a different biological function, the potential for information storage in a glycan sequence is immense: for a linear carbohydrate hexamer with no repeating units, the number of possible combinations rises to $>1 \times 10^9$ sequences [4] compared to 6×10^6 and 6×10^5 for proteins and nucleic acids, respectively. Nucleic acids and proteins elongate unidirectionally, in a template-dependent manner (Scheme 1). Thus, their diversity and growth is restricted by the information contained in the nucleic acid template. In contrast, carbohydrate chain elongation proceeds in a template-independent fashion, in which a glycan unit, chemically reacting with an available hydroxyl group, forms a linkage to the growing chain through its anomeric position. Unlike proteins and nucleic acids, the glycan chain can be branched like glycogen (Scheme 1). Because the template constraint is not present, a leap in diversity occurs.

The carbohydrate code has usually been compared to a dictionary, in which each unit represents a word. In contrast, we propose that the carbohydrate code bears more of a resemblance to hieroglyphs (or glyphs), because each unit's meaning does not only depend on what precedes or follows (context, e.g. primary sequence), but also depends on its 3D shape; the code thus resembles more of an "idea" that can be built upon. This hieroglyph, is not only embedded in the chemical identity of its units but also in the α or β configuration at the anomeric carbon, which influences molecular shape; the possibility of multiple ring sizes (furanose or pyranose); the potential for chain elongation and branching at multiple chemically equivalent hydroxyl groups in different positions; and potential chemical modification of hydroxyl to groups such as amino, O-acetyl, N-acetyl, phosphate, sulfate or even to substitute a hydroxyl group by a hydrogen atom to give a deoxy sugar (Scheme 2). What remains unknown, however, is whether the increase in complexity reflects a functional importance.

Carbohydrates are often found covalently linked, or conjugated, to proteins and lipids. In this manner, the function of the carbohydrate-bearing biomolecule can be modified or modulated. The same carbohydrate fragment may appear in different cellular regions depending on the molecule to which it is conjugated. Because the molecules to which glycans bind can modulate glycan structure and function, their roles may depend on where they are



Scheme 1.