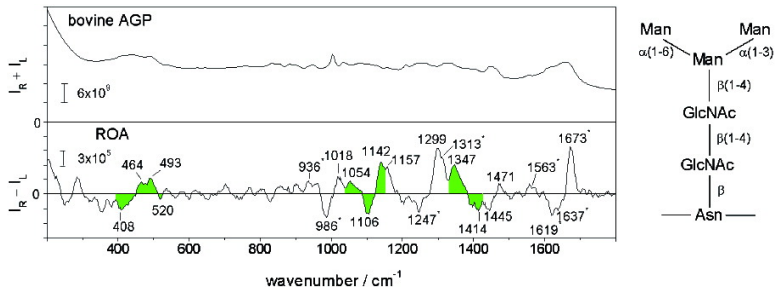
Communication

Polypeptide and Carbohydrate Structure of an Intact Glycoprotein from Raman Optical Activity

Fujiang Zhu, Neil W. Isaacs, Lutz Hecht, and Laurence D. Barron

J. Am. Chem. Soc., **2005**, 127 (17), 6142-6143• DOI: 10.1021/ja051048l • Publication Date (Web): 02 April 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 6 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 04/02/2005

Polypeptide and Carbohydrate Structure of an Intact Glycoprotein from Raman Optical Activity

Fujiang Zhu, Neil W. Isaacs, Lutz Hecht, and Laurence D. Barron*

Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, U.K.

Received February 18, 2005; E-mail: laurence@chem.gla.ac.uk

Glycoproteins play many fundamental roles in biochemistry and are of great interest to the pharmaceutical and biotechnology industries as drug targets, mediators in drug action, and therapeutics.[1] However, the structures of intact glycoproteins are notoriously difficult to characterize using the standard techniques of structural biology (X-ray crystallography and NMR supplemented with spectroscopic techniques, such as ultraviolet circular dichroism, infrared and Raman). In the few cases where X-ray crystal or solution NMR structures of intact or partially deglycosylated glycoprotein structures have been obtained, most of the oligosaccharide structure is poorly defined, in part, due to conformational heterogeneity, with only the first few residues attached to the protein adopting fixed conformations about their respective glycosidic links.[1,2] There is an urgent need for new techniques capable of determining the solution structures of the peptide and carbohydrate components in intact glycoproteins and how they modulate each other's stability and behavior.

Raman optical activity (ROA), which measures vibrational optical activity by means of a small difference in the intensity of vibrational Raman scattering from chiral molecules in right and left circularly polarized incident light (incident circular polarization, ICP, ROA) or, equivalently, as the intensity of a small circularly polarized component in the scattered light (scattered circular polarization, SCP, ROA),[3−6] is a powerful technique for the study of the structures of biomolecules in aqueous solution.[7] Proteins and carbohydrates are both excellent samples for ROA, giving rich and informative band structures over a wide range of the vibrational spectrum. ROA spectra of proteins provide information about the fold as well as the secondary structure of the polypeptide backbone[8,9] together with side chain conformation in some cases.[10] ROA spectra of monosaccharides contain information on sugar ring conformation, relative disposition of OH groups around the ring, the absolute configuration and axial or equatorial orientation of groups attached to the anomeric carbon, and the exocyclic CH_2OH conformation; those of di- and oligosaccharides contain information on the conformation of the C−O−C glycosidic links, and those of polysaccharides provide information about whether the structure is disordered or has extended order, such as helical.[7,11−13] Here, we show that ROA provides spectra of intact glycoproteins which contain information about the structures of both the polypeptide and carbohydrate components.

We focus on α_1-acid glycoprotein (AGP), a much-studied physiologically important glycoprotein whose exact biological function remains obscure.[14] To date, no X-ray crystal structure has been reported, but spectroscopic studies together with homology modeling of the polypeptide component have provided some insight into its structure.[15] The polypeptide is a single chain formed from 183 amino acids with two disulfide bridges. The carbohydrate content represents 45% of the molecular weight and contains five highly sialylated oligosaccharide chains linked via N-glycosidic bonds to asparagine residues. The polypeptide is thought to adopt a lipocalin-type fold based on an eight-stranded antiparallel β-sheet in the form of a barrel with strands linked by β-hairpins plus a long loop.[15]

The backscattered SCP ROA spectrum of bovine AGP, the structure of the N-linked pentasaccharide core common to all oligosaccharide chains, those of bovine β-lactoglobulin and *N*,*N*′-diacetylchitobiose, and their respective structures are displayed in Figure 1. These were measured on the Chiral*RAMAN* instrument recently introduced by BioTools, Inc., which is based on a design by W. Hug.[5,16] This instrument routinely extends protein ROA data acquisition to the low-wavenumber region ∼200−600 cm^{-1}, compared with a low-wavenumber limit of ∼600 cm^{-1} for proteins using the previous generation of instruments, thereby opening a new spectral window on proteins in general and glycoproteins in particular.

Those ROA bands of AGP, which we currently believe to originate mainly in carbohydrate and polypeptide structure, are highlighted in green and marked with an asterisk, respectively, in Figure 1a. The carbohydrate assignments are suggested by similar bands highlighted in green in the ROA spectrum of *N*,*N*′-diacetylchitobiose in Figure 1c. Band overlap in certain regions means there is some uncertainty in a few of these assignments. A number of the ROA bands of AGP assigned to the polypeptide are similar to those observed in β-lactoglobulin, which is consistent with the suggestion[15] that the AGP polypeptide also adopts a lipocalin-like fold. There are some differences, however, which may be due to structural perturbations, changes in hydration or hydrogen-bonding networks, etc., from carbohydrate−peptide interactions. The pentasaccharide core shown in Figure 1 is common to all N-linked oligosaccharides,[18] with the first two glycosidic links after the N-links to asparagine being of the β(1-4)-type, as in *N*,*N*′-diacetylchitobiose, so it is gratifying to observe a similar ROA band pattern in the range ∼400−550 cm^{-1}, not present in the ROA spectrum of β-lactoglobulin that may also be characteristic of this link in both *N*,*N*′-diacetylchitobiose and AGP [a similar pattern is observed in the β(1-4) disaccharide D-cellobiose, but not in the α-(1-4) disaccharide D-maltose, which exhibits opposite signs in the first two ROA bands in the ∼400−450 cm^{-1} region[12]]. This accords with the finding that the conformation of the di-*N*-acetylchitobiose core in N-linked glycoproteins is independent of the protein and is the same as that of the free sugar.[1] The negative ROA band at ∼1363 cm^{-1} in β-lactoglobulin is assigned to the β-turns associated with the antiparallel β-sheet in the protein.[8] An equivalent band in AGP may be obscured by the strong positive ∼1347 cm^{-1} carbohydrate band. We currently have no explanation for the nonappearance in the ROA spectrum of AGP of the strong negative ROA band of *N*,*N*′-diacetylchitobiose at ∼950 cm^{-1}, which is not present in the ROA spectrum of D-cellobiose[12] and which may therefore be associated with the *N*-acetyl groups.

Although we have discussed only a few assignments in this preliminary study and have offered only qualitative interpretations,

10.1021/ja051048l CCC: $30.25

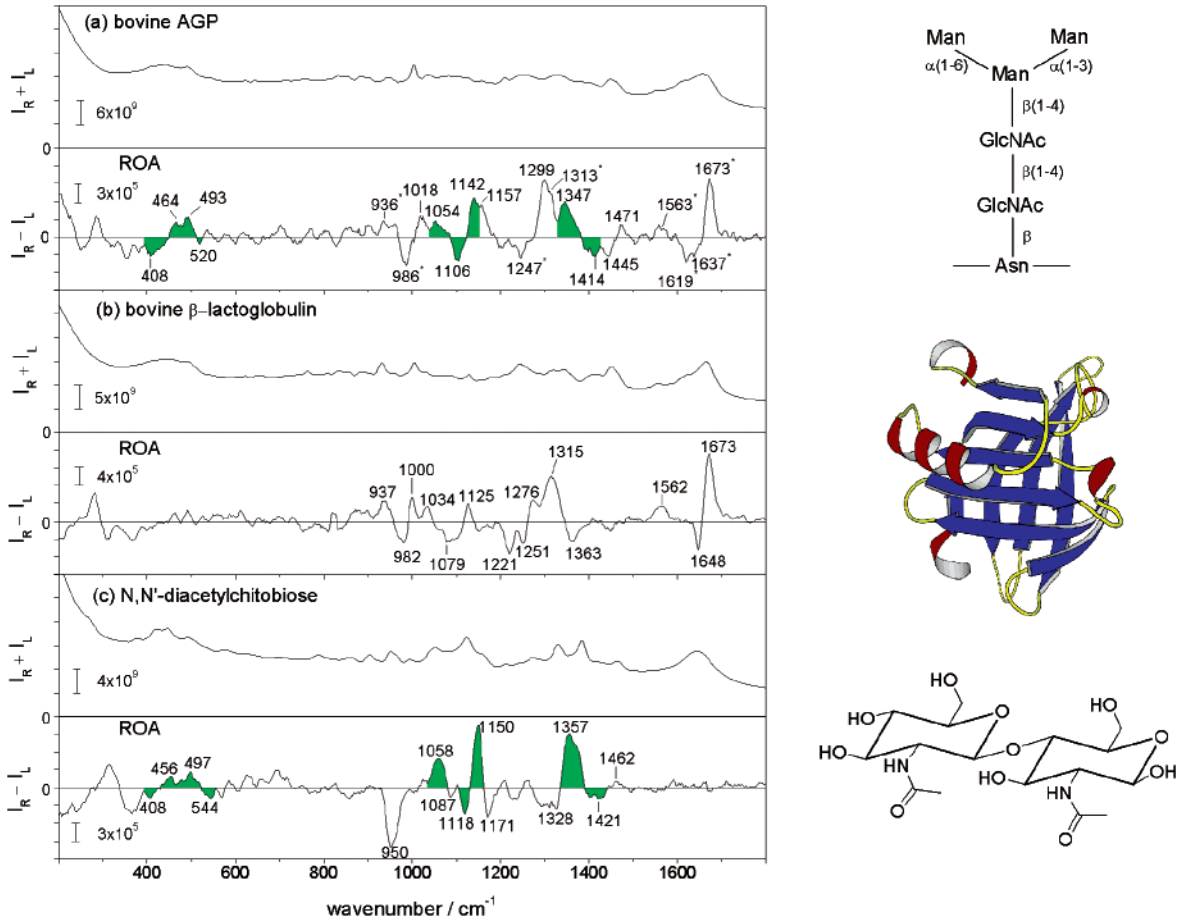


Figure 1. Backscattered SCP Raman ($I_R + I_L$) and ROA ($I_R - I_L$) spectra of (a) bovine α_1-acid glycoprotein (AGP) together with a diagram of the common pentasaccharide core; (b) bovine β-lactoglobulin together with its X-ray crystal structure (PDB code 1beb) drawn using MOLSCRIPT;[19] and (c) *N*,*N*′-diacetylchitobiose. All samples were purchased from the Sigma-Aldrich Company, Ltd. Experimental conditions: all spectra were measured in aqueous solution at concentrations of ∼50 mg/mL; frequency-doubled Nd:YAG laser wavelength at 532 nm; laser power of ∼500 mW at the sample; spectral resolution of ∼8 cm^{-1}; acquisition times of ∼4 h for the proteins and 2 h for the carbohydrate. Bands provisionally assigned to carbohydrate and polypeptide structure in (a) are highlighted in green and marked with an asterisk, respectively.

it is apparent from the plethora of ROA bands characteristic of both the peptide and carbohydrate components that ROA has the potential to provide structural information on glycoproteins well beyond the capabilities of other spectroscopic techniques. A detailed analysis of the band patterns may ultimately provide information on the more conformationally heterogeneous and functionally crucial peripheral oligosaccharide segments. For example, ROA has the sensitivity to distinguish, in principle, stacked glycosidic link conformations from extended conformations, both of which appear to be present in the oligomannose glycoforms of ribonuclease B.[1]

Acknowledgment. We thank the BBSRC for a research grant, and Dr. D. J. Procter for discussion.

References

(1) Dweck, R. A. *Chem. Rev.* **1996**, *96*, 683−720.
(2) Lapthorn, A. J.; Harris, D. C.; Littlejohn, A.; Lustbader, J. W.; Canfield, R. E.; Machin, K. J.; Morgan, F. J.; Isaacs, N. W. *Nature* **1994**, *369*, 455−461.
(3) Barron, L. D.; Bogaard, M. P.; Buckingham, A. D. *J. Am. Chem. Soc.* **1973**, *95*, 603−605.
(4) Barron, L. D. *Molecular Light Scattering and Optical Activity*, 2nd ed.; Cambridge University Press: Cambridge, 2004.
(5) Hug, W. In *Handbook of Vibrational Spectroscopy*; Chalmers, J. M., Griffiths, P. R., Eds.; Wiley: Chichester, 2002; Vol. 1, pp 745−758.
(6) Barron, L. D.; Hecht, L.; McColl, I. H.; Blanch, E. W. *Mol. Phys.* **2004**, *102*, 731−744.
(7) Barron, L. D.; Hecht, L.; Blanch, E. W.; Bell, A. F. *Prog. Biophys. Mol. Biol.* **2000**, *73*, 1−49.
(8) McColl, I. H.; Blanch, E. W.; Gill, A. C.; Rhie, A. G. O.; Ritchie, M. A.; Hecht, L.; Nielsen, K.; Barron, L. D. *J. Am. Chem. Soc.* **2003**, *125*, 10019−10026.
(9) McColl, I. H.; Blanch, E. W.; Hecht, L.; Barron, L. D. *J. Am. Chem. Soc.* **2004**, *126*, 8181−8188.
(10) Blanch, E. W.; Hecht, L.; Day, L. A.; Pederson, D. M.; Barron, L. D. *J. Am. Chem. Soc.* **2001**, *123*, 4863−4864.
(11) Wen, Z. Q.; Barron, L. D.; Hecht, L. *J. Am. Chem. Soc.* **1993**, *115*, 285−292.
(12) Bell, A. F.; Hecht, L.; Barron, L. D. *J. Am. Chem. Soc.* **1994**, *116*, 5155−5161.
(13) Bell, A. F.; Hecht, L.; Barron, L. D. *J. Raman Spectrosc.* **1995**, *26*, 1071−1074.
(14) Fournier, T.; Medjoubi, N. N.; Porquet, D. *Biochim. Biophys. Acta* **2000**, *1482*, 157−171.
(15) Kopecký, V., Jr.; Ettrich, R.; Hofbauerová, K.; Baumrauk, V. *Biochem. Biophys. Res. Commun.* **2003**, *300*, 41−46.
(16) Hug, W. *Appl. Spectrosc.* **2003**, *57*, 1−13.
(17) Brownlow, S.; Cabral, J. H. M.; Cooper, R.; Flower, D. R.; Yewdall, S. J.; Polikarpov, I.; North, A. C. T.; Sawyer, L. *Structure* **1997**, *5*, 481−495.
(18) Berg, M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 5th ed.; W. H. Freeman and Company: New York, 2002.
(19) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.

JA051048L