Virology 407 (2010) 281–288

Contents lists available at ScienceDirect

Virology

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

A simple and general method for determining the protein and nucleic acid content of viruses by UV absorbance

J. Zachary Porterfield, Adam Zlotnick *

Department of Biochemistry & Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, USA
Department of Molecular & Cellular Biochemistry and Biology, Indiana University, Bloomington, Indiana 47405, USA
Department of Biology, Indiana University, Bloomington, Indiana 47405, USA

ARTICLE INFO

Article history:
Received 21 July 2010
Returned to author for revision
10 August 2010
Accepted 14 August 2010
Available online 17 September 2010

Keywords:
Capsid assembly
Virus assembly
Electrostatics
Hepatitis B virus
Brome Mosaic Virus
Absorbance spectroscopy
RNA content
DNA content

ABSTRACT

UV spectra of viruses are complicated by overlapping protein and RNA absorbance and light scattering. We describe and validate methodology for estimating RNA and protein concentration from such spectra. Importantly, we found that encapsidation did not substantially affect RNA absorbance. Combining absorbance data with a known T number, we confirmed that brome mosaic virus packages about 3100 nucleotides/capsid, consistent with its genome. E. coli-expressed hepatitis B virus (HBV) packages host RNA based on capsid charge and volume. We examined HBV capsid protein (Cp183, +15 charge) and a less basic mutant (Cp183-EEE, +12 charge) that mimics a phosphorylated state. Cp183-EEE packaged ~3450 nucleotides per T=4 capsid and Cp183 packaged ~4800 nucleotides, correlating to the size of HBV's RNA pre-genome and mature DNA genome, respectively. The RNA:protein charge ratio (about 1.4 phosphates per positive charge) was consistent with that of other ssRNA viruses. This spectroscopic method is generalizable to any virus-like particle.

Introduction

There are numerous examples of virus-like particles assembled in vitro and in heterologous systems where the nucleic acid content is of interest but has not been rigorously determined (Annamalai and Rao, 2005; Grieger and Samulski, 2005; Johnson et al., 2002; Krishna et al., 2003; Ludwig and Wagner, 2007; Ma and Vogt, 2002; Mukherjee et al., 2007; Pattenden et al., 2005; Satheshkumar et al., 2005; Stockley et al., 2007). This sampling of systems ranges from bacteriophages to retroviruses; they package RNA and DNA. In some cases, packaging is sequence specific, in other cases it is promiscuous. The nucleic acid may function as structural scaffolding (Tihova et al., 2004) or as an allosteric effector (Stockley et al., 2007). The nucleic acid content may be of particular interest as genetic information in virus-based transfection systems, especially in the context of gene therapy (Grieger and Samulski, 2005). One of the most accessible methods for defining the nucleic content of a sample is UV absorbance.

Accurate interpretation of the absorbance spectra of viruses, and other large nucleo-protein complexes, are complicated by several factors: the UV absorption of protein and nucleic acid overlap, light scattering systematically distorts the spectra, and there is the possibility that structural change of the nucleic acid may alter its hyperchromicity. The overlap in the absorption of proteins and nucleic acids can be resolved (Gonen and Rytwo, 2009; Kalb and Bernlohr, 1977; Mukherjee et al., 2010). Methods have also been described to evaluate the effect of light scattering on absorbance spectra (Cox et al., 2002; Leach and Scheraga, 1960). An extinction coefficient for a large nucleoprotein complex can be experimentally determined that explicitly accounts for all of these features, as has been done for adenovirus (Sweeney and Hennessey, 2002). However, an ad hoc extinction coefficient will not always yield a meaningful measurement of concentration and has no predictive power for differing conditions and/or mutations. When the stoichiometry of chromophores is unknown and/or the light scattering contribution to absorbance is variable, a self-consistent method for determining the concentrations of protein and nucleic acid is required.

In this paper we describe and test a simple method for analyzing absorbance spectra of viruses, correcting for light scattering and determining accurate concentrations of both protein and nucleic acid. To demonstrate the utility of this method we examine two viruses. Brome Mosaic Virus (BMV) packages an average of about 3100 nucleotides/capsid in vivo and virus-like particles assembled in vitro (Fox et al., 1998). In vivo, Hepatitis B Virus (HBV) in a phosphorylated state packages the ~3200 nucleotide RNA pre-genome (Seeger et al., 2007). We determine

* Corresponding author. Department of Molecular and Cellular Biochemistry, Indiana University, 212 S. Hawthorne Dr., Simon Hall 220, Bloomington, IN 47405-7003, USA. Fax: +1 812 856 5710.
E-mail address: azlotnic@indiana.edu (A. Zlotnick).

0042-6822/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2010.08.015

the RNA content of two versions of E. coli expressed HBV capsids: Cp183, a full-length 183 residue capsid (core) protein, and Cp183-EEE, which incorporates three Ser to Glu mutations on each RNA-binding domain to mimic the phosphorylated state and effectively decrease the net charge of each RNA-binding domain from +15 to +12, demonstrating the relationship between protein charge and packaged nucleic acid.

Results

Nucleic acid and protein spectra can be differentiated qualitatively and quantitatively (Fig. 1A). Stacked purines and pyrimidines absorb light with an absorption maximum at 260 nm and a 260 nm/280 nm ratio of ~2.0 for ssRNA (Glasel, 1995). Protein absorbance is dominated by tryptophans, tyrosines, and disulfide bonds. Protein absorbance has a peak near 280 nm and a characteristic shoulder at 290 nm. The 260 nm/280 nm ratio for protein is ~0.6 (Glasel, 1995; Goldfarb et al., 1951). Here, we consider the absorbance spectrum to be a sum of protein and nucleic acid components, determined algebraically from the UV absorption at 260 nm and 280 nm after removing the background light scattering. (A spreadsheet application of the following equations is included in Supporting Information.)

Light scattering corrections

Light scattering increases the 260 nm/280 nm absorbance ratio for large complexes, e.g. for viruses. Thus, for accurate absorbance-based nucleoprotein concentration calculations, the contribution of light scattering to the apparent absorption must be estimated and subtracted from the spectra. According to the Rayleigh approximation, light scattering of spherical solutes is proportional to λ^{-4}; the amount of scattering at a given wavelength is proportional to the molecular mass of the solute. When the diameter of the scatterer is much less than the wavelength of the incident light, the amount of light scattering is well predicted by the Rayleigh approximation (Cox et al., 2002; Young, 1982). Fig. 1B shows absorbance spectra from several pure HBV capsid protein samples with varying amounts of light scattering. The variation in light scattering may be attributed to the amount of capsid and aggregation.

We find that a nucleo-protein absorbance spectrum can be corrected for light scattering using a two point approximation. Measured absorbance can be partitioned into light that is actually absorbed ($A_{corrected,\lambda}$) and the apparent absorption that is due to light scattering at a given wavelength ($A_{LS,\lambda}$). $A_{LS,\lambda}$ for a solution of particles that are small compared to the length of scattered light can be expressed as a function of λ.

$$A_{LS,\lambda} = c_1(\lambda)^{-4} + c_2 \quad (1)$$

Where c_1 is a coefficient for wavelength-dependent light scattering and c_2 accounts for baseline offset. Though a more rigorous fit can be accomplished, for simplicity, we estimated these constants from the apparent absorbance at 340 nm and 360 nm. At these wavelengths, there is no significant absorbance by either protein or nucleic acid so that absorbance is essentially A_{LS}. The light scattering at these wavelengths can therefore be expressed as

$$A_{340} = c_1(340nm)^{-4} + c_2 \quad (2)$$

$$A_{360} = c_1(360nm)^{-4} + c_2 \quad (3)$$

c1 and c2 can be solved from this set of equations

$$c_1 = \frac{A_{340} - A_{360}}{(340nm)^{-4} - (360nm)^{-4}} \quad (4)$$

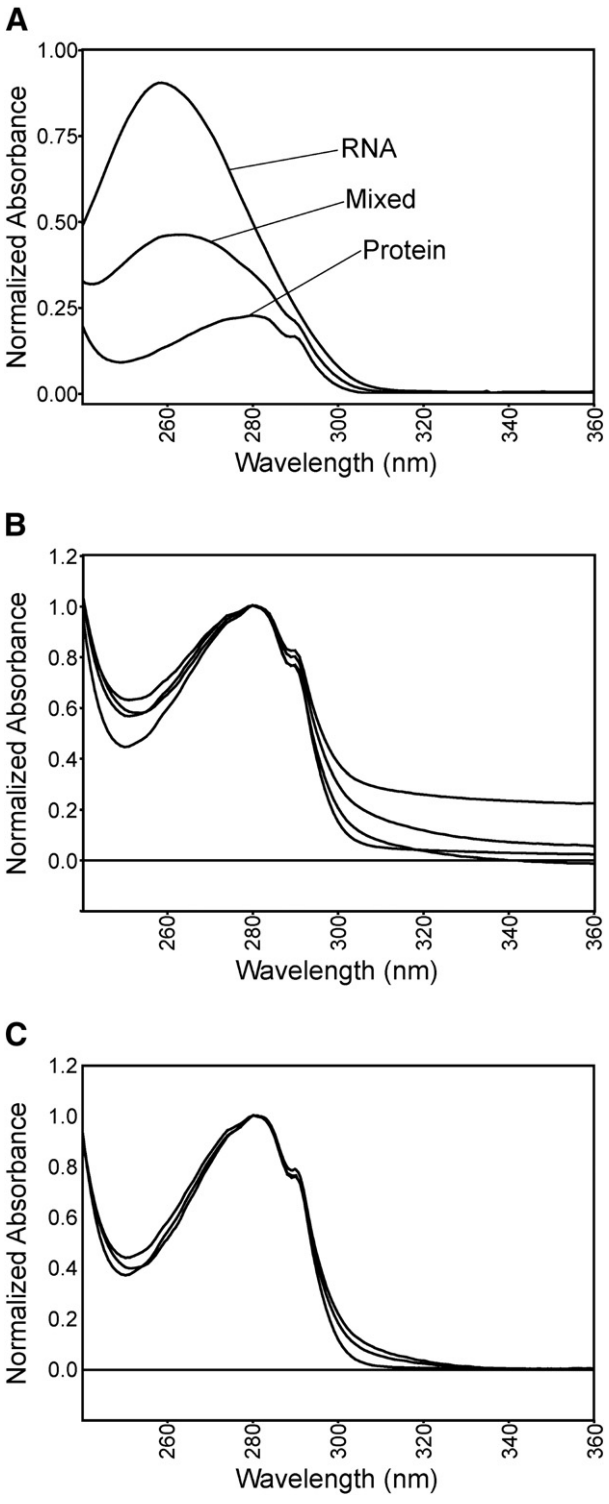

Fig. 1. Standard, experimental, and corrected absorbance spectra. (A) Typical UV absorbance spectra for pure protein, pure RNA, and a mixture of protein and RNA. These spectra can be differentiated qualitatively by shape and quantitatively by 260/280 nm ratio. Quantitative evaluation of UV absorbance spectra requires that the contribution of light scattering to the apparent absorbance be accounted for at each wavelength. (B) Spectra from pure protein samples show varying severity of light scattering and baseline offsets. The degree of light scattering was estimated using the apparent absorbance at 340 nm and 360 nm, where protein and RNA absorbance is minimal, assuming a λ^{-4} relationship between light scattering and wavelength. (C) The spectra from panel B were corrected for light scattering and baseline offset. The corrected spectra have much more consistent 260 nm/280 nm ratios. The spectra in panels B and C were normalized to a 280 nm absorbance of one for comparison to highlight the improvement of the 260 nm/280 nm ratio.