Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1043 (2004) 57–64
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Evaluation of on-line high-performance size-exclusion chromatography, differential refractometry, and multi-angle laser light scattering analysis for the monitoring of the oligomeric state of human immunodeficiency virus vaccine protein antigen

John Barackman [a,*], Isaias Prado [b], Chulani Karunatilake [a], Kenji Furuya [a]

[a] *Chiron Analytical Development, Chiron Corporation, Emeryville, CA 94608, USA*
[b] *Chiron Formulation and Drug Delivery, Chiron Corporation, Emeryville, CA 94608, USA*

Available online 17 March 2004

Abstract

Chiron has developed a novel mutant form of the human immunodeficiency virus (HIV) envelop protein, o-gp140, that is currently entering Human Phase 1 clinical trials for testing as a prophylactic HIV vaccine. The o-gp140 protein is oligomeric and the quaternary structure is thought to play an important role in its activity as an antigen. As o-gp140 proceeds through the clinical trial process and prior to marketing approval, analytical methods that are able to demonstrate manufacturing consistency with respect to degree of oligomerization will need to be developed and validated. On-line high-performance size-exclusion chromatography, differential refractometry, and multi-angle laser light scattering analysis (HPSEC–RI–MALLS), a method commonly used to obtain the molar mass of macromolecules based on the Rayleigh–Gans–Debye approximation, was evaluated for this purpose. The results obtained demonstrated intra- and inter-day precisions to be 0.9 and 3.6% R.S.D., respectively. Accuracy was found to be equal to, or better than, 11% when comparing the known molar masses of test proteins to that of the molar masses determined by the method. Additionally, the method compared favorably to orthogonal native polyacrylamide gel electrophoresis and ultracentrifugation analyses. *R*-factor analysis was used to demonstrate that HPSEC–RI–MALLS is capable of discriminating compositional differences between o-gp140 test lots. Based on the data presented, HPSEC–RI–MALLS may be a suitable manufacturing control method.

Keywords: Human immunodeficiency virus; *R*-factor; Proteins

1. Introduction

The human acquired immunodeficiency syndrome (AIDS) and the human immunodeficiency virus (HIV) that causes AIDS is an epidemic that continues to grown in magnitude and scope reaching 1.5 million infected individuals in North America and Western Europe, and 42 million infected individuals globally by the end of 2002 [1]. Collaborations between government, academia, and the pharmaceutical industry have been assembled on a global scale to develop an effective vaccine against HIV, mostly utilizing new technologies in the pursuit of protective immunity [2–4]. One such new approach is to develop recombinant HIV envelop glycoproteins that better mimic native forms which, amongst other properties, are thought to exist as oligomers on the exterior surface of the virion particle [5,6]. Oligomeric forms of the HIV envelope glycoprotein have been found to bind neutralizing antibodies from primary isolates stronger than recombinant monomeric forms of the envelop glycoprotein [7]. In addition, deletions to the second hypervariable region (V2) have been shown to improve the ability of oligomeric HIV envelope glycoproteins to induce neutralizing antibodies still further [8]. Chiron has exploited both of these properties by developing novel V2 deletion mutants of the HIV envelop protein, o-gp140, that forms stable oligomers and that can be highly purified at clinically relevant scale [9]. The o-gp140 protein is currently entering Human Phase 1 clinical trials for testing as a prophylactic HIV vaccine. A recombinant protein, whether the intended end use is as a therapeutic or as a vaccine antigen, must have its basic physicochemical properties (i.e., primary, secondary, tertiary, and quaternary structures) extensively characterized prior to marketing approval [10]. In addition, analytical methods for demonstration of manufacturing consistency

* Corresponding author. Tel.: 1-510-923-8194; fax: 1-510-923-4116.
E-mail address: john_barackman@chiron.com (J. Barackman).

0021-9673/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2004.02.011

must also be validated prior to marketing approval; validation parameters of which precision, accuracy, and specificity are fundamental [11]. For o-gp140, the quaternary structure (i.e., the oligomeric state) is thought to play a key role in developing an effective vaccine, therefore, analytical methods that can accurately characterize the o-gp140 oligomers as well as provide a means to demonstrate manufacturing consistency of the degree of oligomerization are important. Analytical methods are required to both accurately assess the oligomeric state (i.e., the absolute size and relative distribution of oligomers), as well as demonstrate that a consistent oligomeric state is maintained from one product batch to another. To address this analytical need, high-performance size-exclusion chromatography (HPSEC) with differential refractometric (RI) and multi-angle laser light scattering (MALLS) detection was evaluated. This method, and variations of HPSEC–RI–MALLS, are used primarily to obtain the molecular mass of macromolecules, both mean and absolute [12], and have been used to characterized many different macromolecular systems, including polyethylene oxide star polymers [13], polysaccharides [14], surfactant micelle associations [15], receptor–ligand complex formations [16], and protein derivatization with poly(ethylene glycol) (PEGylation) [17]. However, HPSEC–RI–MALLS has been focused mainly on characterization; use of this technique for manufacturing control has not been fully exploited, particularly for protein pharmaceuticals, although some explorations of this use has been described [18]. It may be that simple and robust quantitative methods to render the HPSEC–RI–MALLS data suitable for utilization within a quality control environment are needed before the utility of this method for manufacturing consistency can be fully realized. Two statistical methods: differential analysis and *R*-factor analysis, adopted from peptide map "fingerprinting" techniques, are explored here [19–22]. These type of analyses allow differences between a test sample and a reference sample to be quantitatively measured, and therefore, "Conforms to Reference" type specifications be set [21]. In this paper, the results of studies to ascertain the accuracy and reproducibility of HPSEC–RI–MALLS are presented. In addition, the results of the HPSEC–RI–MALLS analysis of three lots of o-gp140 manufactured at different stages of the development cycle are presented. These samples, containing modest differences in purity and composition, were selected to act as "mock" manufacturing production runs in order to demonstrate that HPSEC–RI–MALLS, coupled with *R*-factor analysis, is able to address specificity, and therefore, is suitable as a method for assay of manufacturing consistency.

2. Experimental

2.1. Materials

Three lots of purified o-gp140 antigen formulated in 10 mM sodium citrate, 300 mM sodium chloride, pH 6.0 buffer (formulation buffer) were obtained from represented bench scale (LOT3), toxicology scale (LOT1), and phase 1 GMP scale (LOT2) purifications at Chiron (Emeryville, CA, USA). The antigen lots were produced in Chinese hamster ovary (CHO) cells and purified at bench scale (approximately 25 mg yield), toxicology scale (approximately 200 mg yield), and GMP scale (approximately 500 mg yield) by methodologies as described in [9] with the exception of CHO production levels and column dimensions appropriate for each purification scale, and with the exception that the GMP purification included additional steps for viral inactivation and removal. Final purity of each lot of o-gp140 was >90% by reducing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and each was verified to have the expected amino acid composition, amino acid sequence, monomer mass, CD4 binding, etc. by a panel of assays (data not shown). Monomeric HIV-1_{SF2} gp120 (gp120) was obtained from Chiron [23]. Bovine serum albumin (BSA), chicken egg ovalbumin (OVA), fetal calf serum fetuin (Fetuin), and bovine thyroglobulin (Thyroglobulin) were obtained from Sigma–Aldrich (St. Louis, MO, USA). All chemicals used were of reagent grade or equivalent.

2.2. HPSEC–RI–MALLS

Injections of test materials (40–100 μg typical) was separated by HPSEC using isocratic 50 mM sodium phosphate, 300 mM sodium chloride, pH 6.6 eluent at 0.5 ml/min flow rate on an Alliance 2695 HPLC system (Waters, Milford, MA, USA) equipped with a TSK-GEL G4000SWXL column (Tosoh Biosep LLC, Montgomeryville, PA, USA) heated to 35 °C. The detection system used consisted of three detectors connected on-line in the following order: a Waters 996 photodiode array detector, a Wyatt miniDAWN multi-angle laser light scattering detector (Wyatt Technology, Santa Barbara, CA, USA), and a Waters 2410 refractive index detector (Waters). The photodiode array detector was used to obtain 280 nm wavelength profiles. Data obtained from the HPSEC, RI, and MALLS portion of the setup was used to calculate molar mass as a function of elution volume using the Rayleigh–Gans–Debye approximation as previously described in detail [24,25]. Weight-average molar mass (M_w) for a given elution volume, differential mass fractions, and cumulative mass fractions for any given molar mass range were also calculated from the data as previously described [24]. These calculations were automated using ASTRA, a software program provided by Wyatt Technology. The RI detector output voltages were calibrated for specific refractive index increment (d*n*/d*c*) using sodium chloride standards, and the Rayleigh ratio of the MALLS detector was calibrated using pure toluene [25]. The MALLS detector was normalized for small differences in detector sensitivity, scattering volume, refractive index, and the volume offsets between detectors using BSA as recommended by Wyatt Technology. The accuracy of the d*n*/d*c* value used is inversely proportional to the accuracy of the molar mass

obtained by the Rayleigh–Gans–Debye approximation. Therefore, obtaining accurate d*n*/d*c* values are important, and, ideally, d*n*/d*c* values are empirically measured. However, prohibitively large amounts of o-gp140 was necessary to empirically determine its d*n*/d*c* value; therefore, it was decided to estimate the d*n*/d*c* value instead using the following calculation: $dn/dc = M_p \times dn/dc_p + M_c \times dn/dc_c$, where M_p and M_c are the mass fraction, and dn/dc_p and dn/dc_c are the refractive index increments for the protein and the carbohydrate portions, respectively. A dn/dc_p of 0.185 and a dn/dc_c of 0.135 was used throughout [26].[1] For consistency, the d*n*/d*c* of all other proteins examined were calculated in like manner. A molar mass of 1.14×10^5 g/mol for o-gp140 monomer, and a carbohydrate content of 36% (w/w) was calculated for o-gp140 matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) analysis (data not shown). A molar mass of 1.00×10^5 g/mol for gp120 based on MALDI-TOF [27] and a carbohydrate content of 46.2% based on the difference in mass between the amino acid composition and the mass determined by MALDI-TOF was used. Molar masses of 4.30×10^4, 4.63×10^4, 6.68×10^4, and 6.69×10^5 g/mol and carbohydrate contents of 3.2, 22.9, 0, and 8.7% by weight for OVA, fetuin, BSA, and thyroglobulin, respectively, were used for calculation and comparison purposes [28].

2.3. Precision and accuracy

Intra-day variability was estimated from the relative standard deviations (R.S.D.s) obtained from HPSEC–RI–MALLS analysis of three injections of LOT1 performed on a single day. Inter-day variability was estimated from R.S.D.s obtained from HPSEC–RI–MALLS analysis of one injection of LOT1 performed on each of six separate days. To determine accuracy, masses were obtained by HPSEC–RI–MALLS analysis after injection of OVA, fetuin, BSA, gp120, o-gp140 (LOT1), and thyroglobulin (40–100 μg typical) using the method as described above and compared to the known (expected) masses of each. The elution region ascribed to the monomer forms of each protein were utilized for this purpose. The percent difference for each protein was calculated as follows:

$$\text{difference (\%)} = \frac{(M_w \text{ found} - M_w \text{ expected})}{M_w \text{ expected}} \times 100.$$

2.4. Native PAGE

Analysis of o-gp140 by Blue Native PAGE was performed as described elsewhere [29] with some minor modifications. Lanes of NuPAGE 3–8% Tris–acetate gradient gels (Invitrogen, Carlsbad, CA, USA) were loaded with 5 μg LOT1 prepared in 100 mM Tris–HCl, 10% glycerol, 0.002% Coomassie R-250 dye (Bio-Rad, Hercules, CA, USA), pH 8.6 buffer. A lane containing 5 μg gp120 was included for comparison purposes. Gels were electrophoresed in a running buffer composed of 50 mM Tris–HCl, 50 mM tricine, pH 7.6 with the inclusion of 0.002% Coomassie R-250 dye in the cathode side for 90 min at 150 V constant. Gels were then destained as typical and scanned with a laser densitometer (Personal Densitometer SI, Amersham Parmacia Biotech, Piscataway, NJ, USA).

2.5. Ultracentrifugation

Samples of LOT1, LOT2, and LOT3 were diluted to 0.4 mg/ml with formulation buffer and analyzed by sedimentation velocity ultracentrifugation as described elsewhere [30] using a Beckman Optima XL-A analytical ultracentrifuge equipped with an AN-60Ti rotor at 35,000 rpm equilibrated to 20 °C. Samples were scanned at 280 nm wavelength approximately every 4.5 min for the duration of the analysis (approximately 5 h). Normalized sedimentation coefficient distribution plots were derived from the data obtained [30]. Centrifugation and distribution analyses were performed by Alliance Protein Labs. (Thousand Oaks, CA, USA).

2.6. Differential mass fraction and R-factor analysis

As noted above, ASTRA can provide cumulative mass fractions from HPSEC–RI–MALLS data. The cumulative mass fractions obtained for a given injection of o-gp140, calculated over the eluting species of interest, can be ported in the form of ASCII *x*, *y* data pairs, where *x* represents small molar mass increments (typically 1.00×10^3 g/mol) and *y* the percent fraction of particles within the increment. Species eluting between the excluded and included volume of each sample injection were typically used for these purposes, representing approximately 97% or greater of the total mass eluted. For this study, the cumulative mass fraction data from the HPSEC–RI–MALLS analysis of eight injections of LOT1 and one injection each of LOT2 and LOT3 were obtained. The differential mass fraction analysis was performed using Microsoft Excel after import of ASCII data. Molar mass increments were converted to their equivalent log base 10 units then grouped into 12 size 'bands' as follows: <5.0, ≥5.0 to <5.1, ≥5.1 to <5.2, ≥5.2 to <5.3, . . . , ≥5.9 to <6.0, and >6.0 log units. The cumulative mass fraction data was summed over the molar mass range specified by each size band. For LOT1, the data thus obtained from each of the eight injections was averaged and the standard error of the mean (S.E.M.) calculated at each size band. The data are presented directly comparing LOT1 to LOT2, and LOT1 to LOT3, and as 'differential mass fractions' after subtraction of LOT2 and LOT3 results from that of LOT1.

[1] The d*n*/d*c* of 0.185 is based on that of BSA and was obtained from [26]. The d*n*/d*c* of starch was used as an approximate d*n*/d*c* value for all protein oligosaccharides and is a value midway between that of dextran (0.137–0.147), heparin (0.129), and mucopolysaccharides (0.110). In addition, a comprehensive list of d*n*/d*c* values can be obtained from Wyatt Technology.

The data from LOT1, acting as a "pseudo-reference" were further compared to LOT2 and LOT3 by *R*-factor analysis, a statistical pattern matching algorithm [21,22] using the following equation:

$$R = \sqrt{\frac{\sum_1^n (\overline{\mathrm{MM}_{\mathrm{LOT1}i}} - \mathrm{MM}_{\mathrm{sample}_i})^2}{\sum_1^n (\overline{\mathrm{MM}_{\mathrm{LOT1}i}})^2}}$$

where $\overline{\mathrm{MM}}_{\mathrm{LOT1}_i}$ is the mean cumulative molar mass of LOT1 at each size band *i*, $\mathrm{MM}_{\mathrm{sample}_i}$ is the cumulative molar mass for either LOT2 or LOT3 at each size band *i*, and *n* is the total number of size bands analyzed. A value of zero for *R* indicates a perfect match between the sample and the reference. The sample becomes increasingly less like the reference as *R* increases in value.

3. Results and discussion

3.1. Overview of HPSEC–RI–MALLS analysis of o-gp140 samples

The elution profiles of o-gp140 obtained from the laser scatter detector at an angle of 90° (LS), the RI signal, and the absorbance at 280 nm obtained from injection of LOT1 is shown in Fig. 1. As indicated by Fig. 1, LS and RI signals do not overlay one another. This is characteristic of heterogeneous particulate mixtures since RI varies in proportion to $c \times \mathrm{d}n/\mathrm{d}c$, while the LS varies in proportion to $M \times c \times \mathrm{d}n/\mathrm{d}c^2$, where *c* is the eluent concentration and *M* the molar mass [25]. This heterogeneous nature of o-gp140 is clearly demonstrated in Fig. 2. As shown in Fig. 2, the molar masses obtained from the HPSEC–RI–MALLS analysis

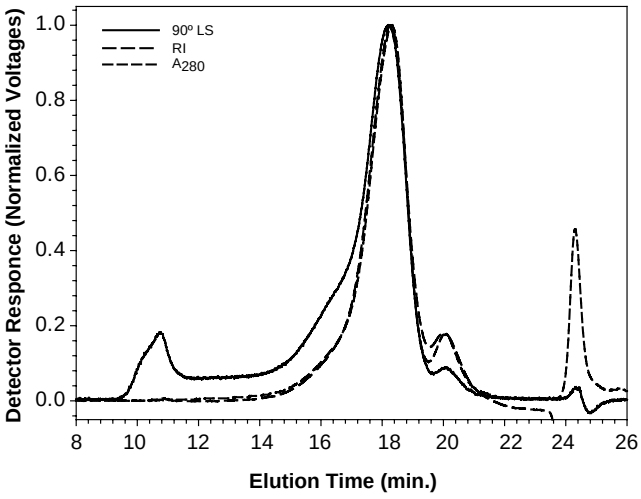

Fig. 1. Example HPSEC–RI–MALLS chromatogram of injection of o-gp140 (LOT1) showing the resulting detector voltages obtained from MALLS at 90° angle (90° LS), refractive index (RI), and UV-Vis at 280 nm wavelength (A_{280}). All voltages normalized to the maximum peak height, and retention times compensated for the minor volume offsets between detectors.

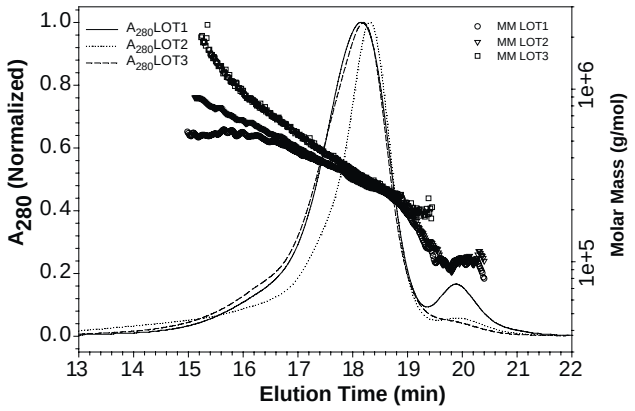

Fig. 2. Result HPSEC–RI–MALLS analysis of the o-gp140 samples. Shown are the molar masses (MM) obtained as a function of elution time from analysis of LOT1, LOT2, and LOT3 samples overlaid with each respective A_{280} elution profile.

appear as a negative sloping lines from relatively high molar masses at early elution volumes to relatively low molar masses at late elution volumes. The molar masses obtained across the major eluting species ranged from $\sim 1 \times 10^5$ g/mol up to $\sim 7 \times 10^5$ g/mol, corresponding to o-gp140 oligomers ranging from monomer to hexamer in size. Furthermore, the main elution peak as defined by the RI and A_{280} elution profiles centers around molar masses corresponding to dimeric and trimeric forms of o-gp140. Additionally, Fig. 2 suggests that the composition of LOT2 appears to be narrower in size distribution compared to that of LOT1 and LOT3. Note that changes in method flow rate, injection volume, or amount of o-gp140 injected (above limit of quantitation) does not alter either the elution profiles or molar masses obtained (data not shown).

3.2. Analysis of method precision

The elution profile obtained from inject of o-gp140 was divided arbitrary into three regions for the purpose of analyzing method precision: ALL, MAIN, and MONOMER as shown in Fig. 3. Weight-average molar masses were then obtained for each of these regions from injection of LOT1. The result of the analysis of intra-day variability is tabulated in Table 1, and the result of the analysis of inter-day variability is tabulated in Table 2. As indicated in Tables 1 and 2, the highest level of precision was obtained from the MAIN elution region (R.S.D. = 0.9% intra-day and R.S.D. = 3.6% inter-day variability), where the amount of material eluted was highest. The MONOMER and ALL elution regions demonstrated lower precision than this, particularly inter-day data (up to 16.1% R.S.D.). For the monomer peak, the relatively high inter-day variability (R.S.D. = 8.6%) is likely due to a combination of the low molar mass of the monomer, and the limited amount of material in the MONOMER peak. Uncertainty in molar mass measurements made by HPSEC–RI–MALLS is inversely dependant on par-

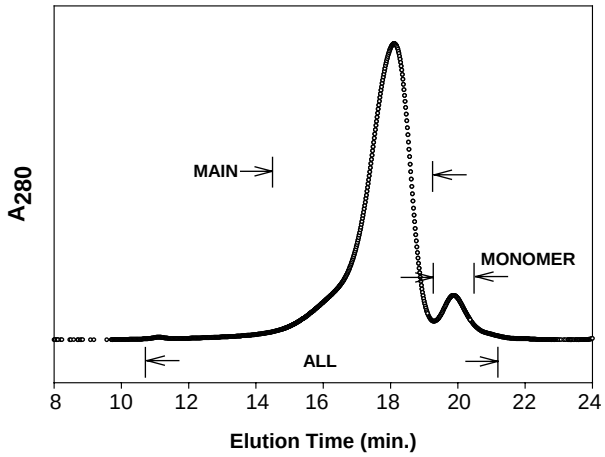

Fig. 3. Location of the o-gp140 chromatographic eluting regions labelled "ALL", "MAIN", and "MONOMER" as described in intra-, inter-day variability, and accuracy sections of text.

Table 1
Weight-average molar mass (M_w), average M_w, and R.S.D. by eluting region resulting from analysis of three injection of LOT1 performed on a single day HPSEC–RI–MALLS

Eluting region	Analysis	M_w (g/mol × 10^5)	Average M_w (g/mol × 10^5)	R.S.D. (%)
MONOMER	1	1.047		
	2	1.035		
	3	1.059	1.05	1.1
MAIN	1	3.081		
	2	3.128		
	3	3.078	3.10	0.9
ALL	1	3.405		
	2	3.474		
	3	3.465	3.45	1.1

Table 2
Weight-average molar mass (M_w), average M_w, and R.S.D. by eluting region resulting from analysis of one injection of LOT1 on each of 6 separate days by HPSEC–RI–MALLS

Eluting region	Analysis day	M_w g/mol × 10^5	Average M_w g/mol × 10^5	R.S.D. (%)
MONOMER	1	1.059		
	2	1.167		
	3	1.030		
	4	1.092		
	5	0.9172		
	6	0.9641	1.04	8.6
MAIN	1	3.078		
	2	3.132		
	3	3.103		
	4	3.382		
	5	3.112		
	6	3.201	3.17	3.6
ALL	1	3.465		
	2	3.548		
	3	3.033		
	4	4.374		
	5	4.634		
	6	3.578	3.77	16.1

Table 3
Comparison of M_w derived from analysis of various proteins by HPSEC–RI–MALLS to that of the expected molar mass to assess method accuracy

Sample	Found M_w g/mol × 10^5	Expected M_w g/mol × 10^5	Difference (%)[a]
OVA	0.3957	0.430	−8
Fetuin	0.5052	0.463	9
BSA	0.5925	0.668	−11
gp120	1.027	1.00	3
o-gp140 (LOT1 MONOMER)	1.050	1.14	−8
Thyroglobulin	6.611	6.69	−1

[a] Calculated as: (found − expected)/expected × 100.

ticle size, and directly dependent on eluent concentration [25]. At or near the exclusion volume, the minor amount of material eluted probably accounts for most of the uncertainty, although it is a possibility that trace variations in level of aggregates or column shedding could contribute to the variability seen. Although not explicitly determined for this study, injection of at least 40 μg o-gp140 was found to be required to be well above the limit of quantitation, at least for the majority of the eluting species (data not shown).

3.3. Analysis of method accuracy

The accuracy of the HPSEC–RI–MALLS was accessed based on comparison to other proteins that bracket o-gp140's mass range and carbohydrate content. The molar masses evaluated ranged from 4.35×10^4 g/mol (OVA) to 6.69×10^5 g/mol (thyroglobulin), and carbohydrate contents that ranged from 0% (BSA) to 46.2% (gp120) (w/w). As summarized in Table 3, the M_w obtained for OVA, fetuin, BSA, gp120, and thyroglobulin, as well as the o-gp140 monomer, were all within 11% of the expected molar masses. Although not investigated here, use of empirically determined d*n*/d*c* values, rather than estimated values, might reduce this uncertainty further. Orthogonal techniques were also investigated to compare, at least qualitatively, the oligomer composition resulting from HPSEC–RI–MALLS analysis of o-gp140. The assumption derived from HPSEC–RI–MALLS analysis is that o-gp140 is composed of a heterogeneous mixture of discreet particles of increasing mass. Native PAGE (Fig. 4) and ultracentrifugation (Fig. 5) were used to investigate this supposition. To note, HPSEC–RI–MALLS is capable of deriving molar mass independent of shape, whereas Native PAGE and ultracentrifugation cannot fully separate mass from shape. The assignment of an observed band (native PAGE) or peak (ultracentrifugation) to a given oligomer may be incorrect without either well characterized reference materials or additional information on shape. However, reasonable assignments can be made if some simplifications are assumed. For instance, if one assumes that all species resolved by ultracentrifugation are spherical (i.e., $f_o = 1$), then the molar mass (M) of a given specie can be related to its sedimentation coefficient using a combination of the

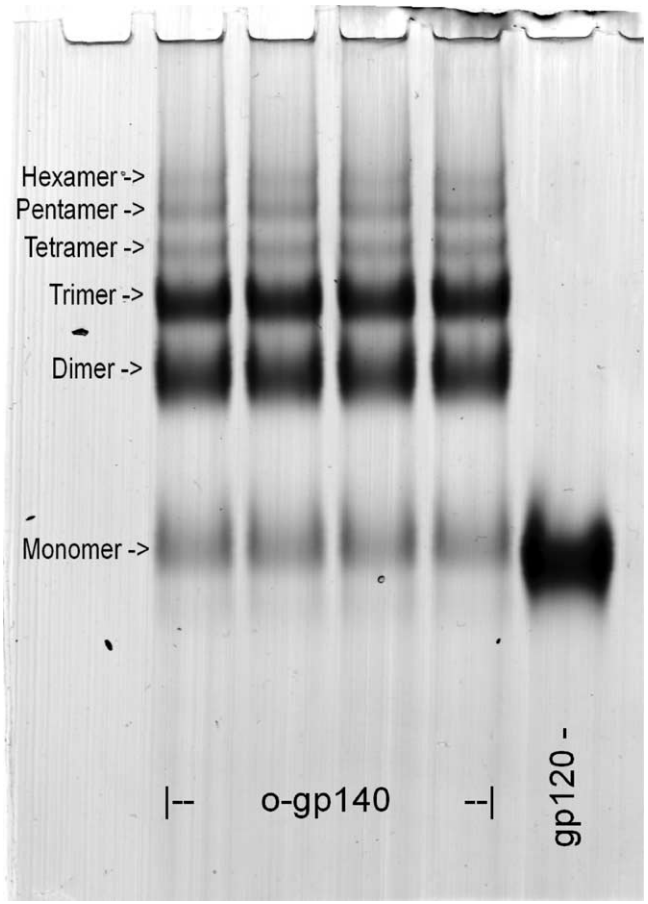

Fig. 4. Analysis of LOT1 by native PAGE. Shown is the laser densitometry image of a Coomassie stained native PAGE gel loaded with o-gp140 LOT1. A lane (at right) containing monomeric gp120 is included for comparison purposes.

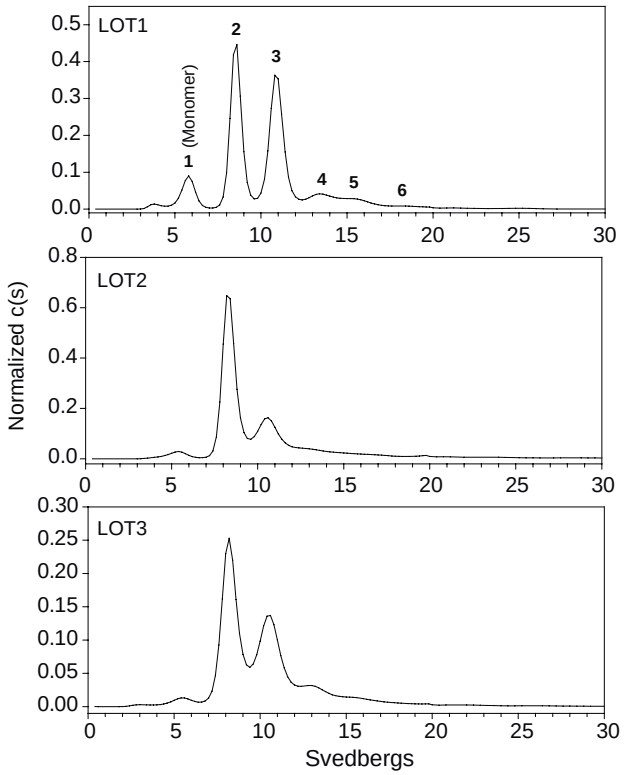

Fig. 5. Normalized sedimentation coefficient distributions obtained from LOT1, LOT2, and LOT3 o-gp140 samples by sedimentation velocity ultracentrifugation. Shown are the relative concentration of sample components as a function of sedimentation rate in units of Svedbergs. The peak marked "monomer" has a sedimentation coefficient appropriate for monomeric o-gp140 assuming a spherical particle with a partial specific volume, solvent density, etc. typical for globular proteins in water solutions. Larger components contained in the o-gp140 samples are indicated by peaks, numbered 2–6, with progressively larger sedimentation coefficient values. The exact nature of each peak resolved by sedimentation velocity ultracentrifugation is not known; mass, as well as shape, play a role in sedimentation rate.

Svedberg and Stokes equation as follows:

$$s = 0.012 \times \frac{M^{2/3}\,(1 - \bar{\nu}\rho)}{\sqrt[3]{\bar{\nu}}} \quad (1)$$

where $\bar{\nu}$ is the partial specific volume and ρ the solvent density [31]. If the molar mass of the monomer is known (as in the case of o-gp140), then the molar masses of each oligomer can be estimated using a derivation of Eq. (1) as:

$$M_2 = M_1 \times \sqrt{\left(\frac{s_2}{s_1}\right)^3} \quad (2)$$

where M_1, s_1 are the molar mass and sedimentation coefficient of the monomer, respectively; and M_2, s_2 are the molar mass and sedimentation coefficient, respectively, for the oligomer in question [32]. Using Svedberg values of 5.6, 8.3, 10.6, 13.1, 15.3, and 18.2, respectively (average of LOT1, LOT2, and LOT3) for peaks 1–6 shown in Fig. 5, and assuming a molar mass of 1.14×10^5 for peak 1 (the monomer), then the molar masses of 2.0×10^5, 2.8×10^5, 3.9×10^5, 4.9×10^5, and 6.3×10^5 can be derived for peaks 2–6, respectively. These masses can be roughly matched to the masses expected for the dimeric, trimeric, tetrameric, pentameric, and hexameric forms of o-gp140 while acknowledging that oligomers of o-gp140 are likely to deviate significantly from that of the assumed spherical form. The observed deviation from ideal may also be due to unexpected heterogeneity derived from differences in glycosylation or misfolding of the monomer units. A better fit to the observed peaks in Fig. 5 can be realized by determining the theoretical ratio of the monomer sedimentation coefficient to that of the oligomer sedimentation coefficient, and, assuming a spherical monomer, arrange the monomer into dumbbell (dimer), triangular (trimer), tetrahedral (tetramer), bipyramid (pentamer), and octahedron (hexamer) arrangements [32,33]. For native PAGE, it is simply assumed that the ladder pattern observed in Fig. 4 is due to regular increasing oligomer size, something that cannot be proven without further information. However, despite these uncertainties, it is clear that o-gp140 is composed of a heterogeneous mixture of oligomers. Furthermore, if one visually inspects the A_{280} plots verses the molar masses obtained from HPSEC–RI–MALLS (Fig. 2), one would

conclude qualitatively that the o-gp140 samples contain primarily dimeric and trimeric oligomers, with relatively minor amounts of monomeric, tetrameric, pentameric, and hexameric species. This conclusion is reasonable (although not proven) based on the assignments derived from native PAGE and ultracentrifugation.

3.4. Differential mass fraction and R-factor analysis

Cumulative mass fraction analysis is method that allows quantitation of the number of particles (expressed as a percent of the total detected mass) at given, discreet, size ranges. Differential mass fraction analysis allows easy comparison of weight fractions obtained from a sample to that of a reference. The *R*-factor analysis takes this comparison further by performing a transformation of the cumulative mass fraction information from a qualitative assessment to one of a quantitative comparison of total, aggregate difference between a sample and a reference. Thus, these types of analyses may provide the basis for "Conforms to Reference" type comparisons for proteins that are heterogeneous in oligomer content. To explore this concept, differential weight fraction and *R*-factor analysis was applied to the o-gp140 HPSEC–RI–MALLS data, the results of which are summarized in Table 4. As indicated in Table 4, the variance (expressed as S.E.M.) in percent content for any given size band for the eight injections of LOT1 analyzed (six of which were on different days) is less than or equal to 2.3%. However, it is clear from the data that differences exist between the o-gp140 lots analyzed (albeit as expected due to the differences in scale and purification methods). Examination of the differential mass fraction data (Δ) shown Table 4 makes these differences stand out. It is clear, for instance, that LOT2 contains 7.2% less particles in the log size range of 5.4–5.5 (2.51×10^5–3.16×10^5 g/mol) compared to LOT1; a conclusion that may be qualitatively drawn by examination of the A_{280} and molar mass plots (Fig. 2). It may be less apparent from examination of Table 4, however, that LOT2 is more different, relative to LOT1, than is LOT3. For this *R*-factor analysis was applied. The results indicated the *R*-factor for LOT2 is 0.31, while that of LOT3 is 0.24. Based on these *R*-values, LOT2 differs more than LOT3 compared to the pseudo-reference sample LOT1 as predicted.

Table 4
Percent cumulative mass fraction at each of 12 size bands resulting from analysis of o-gp140 samples by HPSEC–RI–MALLS

Size band, $\log_{10}$(molar mass)	LOT1, $\bar{x} \pm$ S.E.M.	LOT2		LOT3	
		%	Δ (%)	%	Δ (%)
<5.0	4.0 ± 1.0	2.3	−1.8	0.0	−4.0
5.0–5.1	6.3 ± 0.6	5.3	−1.0	0.0	−6.3
5.1–5.2	3.0 ± 0.5	7.3	4.3	0.0	−3.0
5.2–5.3	5.4 ± 0.5	9.7	4.4	3.7	−1.7
5.3–5.4	15.1 ± 0.5	13.1	−2.0	14.8	−0.3
5.4–5.5	27.6 ± 2.3	20.5	−7.2	30.6	2.9
5.5–5.6	19.8 ± 0.8	15.1	−4.7	22.0	2.2
5.6–5.7	8.4 ± 1.0	8.3	−0.1	12.0	3.6
5.7–5.8	5.4 ± 0.5	4.6	−0.8	5.7	0.3
5.8–5.9	1.6 ± 0.5	3.2	1.6	3.7	2.1
5.9–6.0	0.4 ± 0.2	2.1	1.7	1.9	1.6
>6.0	0.2 ± 0.1	5.1	4.9	2.4	2.2

Shown are the percent mean ($\bar{x}$) and standard error of the mean (S.E.M.) values for LOT1 from a total of eight measurements, as well as the percent and differential mass fraction relative to LOT1 (Δ) of one measurement each of LOT2 and LOT3. Total percent content values do not sum to 100% due to rounding errors and to the averaging used in the calculation of the results. Approximately 97% of all eluting material from the region labeled "ALL", Fig. 3, is represented.

4. Conclusions

Aggregation, even trace amounts, in a protein based therapeutic is typically considered an undesirable property associated with negative toxicological events. Purification development and careful formulation screening are often required to remove or minimize aggregation from proteins prone to this behavior. Analytical methods developed for manufacturing control often focus on demonstrating that the product is free from such aggregation phenomenon. Proteins developed for use as vaccine antigens, however, may demonstrate improved immunogenic properties when oligomeric. The o-gp140 antigen developed by Chiron is designed, in part, to be an oligomer for this reason [8,9], therefore, development of manufacturing methods for o-gp140 will strive to reliably and consistently produce the desired oligomeric state for the antigen. The oligomeric state of o-gp140 will need to be carefully characterized, and the analytical control methods to demonstrate manufacturing consistency will also need to be developed. This study has explored the use of HPSEC–RI–MALLS for these purposes. The data obtained here has shown that HPSEC–RI–MALLS methods can be developed that are both precise and accurate with respect to quantitation of mean size of o-gp140 oligomers, as well as the ability to discriminate lot to lot differences in oligomer composition. In addition, this demonstrated lot-to-lot specificity obtainable from HPSEC–RI–MALLS can readily be applied using statistical analyses in a manner suitable for use in manufacturing control.

Acknowledgements

We thank the following individuals for their contributions and support of this project: Susan Barnett, John Donnelly III, Indrish Srivastava, and Mark Wininger, Chiron Vaccines Research for guidance in the basic biology, purification, and biophysical properties of o-gp140, and for providing samples of research grade antigen for study. Dorota Brousmiche, Cynthia Cowgill, Barbara Gervase, and Xiaorong Wu, Chiron Purification Development for providing o-gp140 sample from toxilogic and GMP scale purifications and for general support of this project. Carol Coan, Viktorya Levi, Christo-

pher Yu, Chiron Analytical Development for continued help with interpretation of the data and providing analytical characterization of o-gp140 in support of this study. John Philo, Alliance Protein Laboratories for guidance in interpretation of the ultracentrifugation data. Finally, Michelle Chen, Wyatt Technology Corp., for guidance on MALLS analysis. This work was supported in part from a contract awarded by the NIH (N01-AI-05396).

References

[1] AIDS Epidemic Update, World Health Organization, Geneva, December 2002.
[2] Hope in the Modern Plague, Editorial, Nat. Med. 7 (2001) 551.
[3] G.K. Folkers, A.S. Fauci, Nat. Med. 4 (1998) 491.
[4] C.A. Heilman, D. Baltimore, Nat. Med. 4 (1998) 532.
[5] J.M. Binley, R.W. Sanders, B. Clas, N. Schuelke, A. Master, Y. Guo, F. Kajumo, D.J. Anselma, P.J. Maddon, W.C. Olson, J.P. Moore, J. Virol. 74 (2000) 627.
[6] C.P. Hill, D. Worthlake, D.P. Bancroft, A.M. Christensen, W.I. Sundquist, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 3099.
[7] T.R. Fouts, A. Trkola, M.S. Fung, J.P. Moore, AIDS Res. Hum. Retroviruses 14 (1998) 591.
[8] S.W. Barnett, S. Lu, I. Srivastava, S. Cherpelis, A. Gettie, J. Blanchard, S. Wang, I. Mboudjeka, L. Leung, Y. Lian, A. Fong, C. Buckner, A. Ly, S. Hilt, J. Ulmer, C.T. Wild, J.R. Mascola, L. Stamatatos, J. Virol. 75 (2001) 5526.
[9] I.K. Srivastava, L. Stamatatos, H. Legg, E. Kan, A. Fong, S.R. Coates, L. Leung, M. Wininger, J.J. Donnelly, J.B. Ulmer, S.W. Barnett, J. Virol. 76 (2002) 2835.
[10] Q6B Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products, Guidance for Industry, Center for Drug Evaluation and Research, US Department of Health and Human Services, 1999.
[11] Q2B Validation of Analytical Procedures: Methodology, Guidance for Industry, Center for Drug Evaluation and Research, US Department of Health and Human Services, 1996.
[12] H.H. Stuting, I.S. Krull, Anal. Chem. 62 (1990) 2107.
[13] D. Rintzler Yen, S. Raghavan, E.W. Merrill, Macromolecules 29 (1996) 8977.
[14] P. Roger, M.A.V. Axelos, P. Colonna, Macromolecules 33 (2000) 2446.
[15] D.P. Norwood, E. Minatti, W.F. Reed, Macromolecules 31 (1998) 2957.
[16] T. Horan, J. Wen, L. Narhi, V. Parker, A. Garcia, T. Arakawa, J. Philo, Biochemistry 35 (1996) 4886.
[17] B.S. Kendrick, B.A. Kerwin, B.S. Chang, J.S. Philo, Anal. Biochem. 299 (2001) 136.
[18] K. Ahrer, A. Buchacher, G. Iberer, D. Josic, A. Jungbauer, J. Chromatogr. A 1009 (2003) 89.
[19] G. Malmquist, R. Danielsson, J. Chromatogr. A 687 (1994) 71.
[20] G. Malmquist, J. Chromatogr. A 687 (1994) 89.
[21] M.W. Kretschmer, R. Mhatre, J. Woodward, A. Schwolow, A. MacLean, Presented at the WCBP Fifth Symposium on the Interface of Regulatory and Analytical Sciences for Biotechnology Health Products, Washington, DC, 2001, Abstracts.
[22] H.-J.P. Sievert, J. Chromatogr. 499 (1990) 221.
[23] C.J. Scandilla, J. Kilpatrick, W. Lidster, C. Parker, J.P. Moore, G.K. Moore, K.A. Mann, P. Brown, S. Coates, B. Chapman, F.R. Masiarz, K.S. Steimer, N.L. Haigwood, AIDS Res. Hum. Retroviruses 12 (1993) 1233.
[24] P.J. Wyatt, J. Colloid Interface Sci. 197 (1998) 9.
[25] P.J. Wyatt, Anal. Chim. Acta 272 (1993) 1.
[26] T. Arakawa, J. Wen, Anal. Biochem. 299 (2001) 158.
[27] C. Bochers, K.B. Tomer, Biochemistry 38 (1999) 11734.
[28] H.A. Sober (Ed.), CRC Handbook of Biochemistry: Selected Data for Molecular Biology, 2nd ed., The Chemical Rubber Co., 1970.
[29] N. Schülke, M.S. Vesanen, R.W. Sanders, P. Zhu, M. Lu, D.J. Anselma, A.R. Villa, P.W.H.I. Parren, J.M. Binley, K.H. Roux, P.J. Maddon, J.P. Moore, W.C. Olson, J. Virol. 76 (2002) 7760.
[30] J.S. Philo, Anal. Biochem. 279 (2000) 151.
[31] J. Lebowitz, M.S. Lewis, P. Schuck, Protein Sci. 11 (2002) 2067.
[32] J.S. Philo, Alliance Protein Laboratories, Thousand Oaks, CA, personal communication.
[33] J. Garcia de la Torre, V.A. Bloomfield, Q. Rev. Biophys. 14 (1981) 81.