Characterization of Human Cytomegalovirus Protease Dimerization by Analytical Centrifugation

James L. Cole*

Department of Antiviral Research, Merck Research Laboratories, West Point, Pennsylvania 19486

Received July 15, 1996; Revised Manuscript Received September 23, 1996[⊗]

ABSTRACT: Human cytomegalovirus, a member of the herpesvirus family, encodes a maturational protease required for processing of the assembly protein and virus replication. The protease is synthesized as a precursor protein that undergoes autoproteolytic cleavage to yield a mature, 28-kDa enzyme. It has recently been demonstrated that mature human cytomegalovirus protease is capable of dimerization and that the dimer is the active species [Darke, P. L., Cole, J. L., Waxman, L., Hall, D. L., Sardana, M. K., & Kuo, L. C. (1996) *J. Biol. Chem. 271*, 7445−7449; Margosiak, S. A., Vanderpool, D. L., Sisson, W., Pinko, C., & Kan, C.-C. (1996) *Biochemistry 35*, 5300−5307]. Here, analytical equilibrium and velocity sedimentation measurements were used to define the thermodynamics of protease dimerization. Protease dimerization is well described by a homogeneous, reversible mass-action equilibrium. The apparent molecular weight of the protease decreases at higher protein concentrations, and good global fits to sedimentation equilibrium data require a positive value of the second virial coefficient, indicating that the protein exhibits thermodynamic nonideality. The magnitude of the nonideality is higher than expected on the basis of excluded volume and electrostatic effects and is not very sensitive to salt concentration, as would be expected for electrostatic effects. The dimer dissociation constants are in agreement with the values we previously determined by activity measurements and hydrodynamic techniques. Dimerization is enhanced by addition of glycerol or NaCl. The temperature dependence of the dimerization constant indicates that both $\Delta H°$ and $\Delta S°$ are negative, which is commonly observed in protein self-association reactions.

Human cytomegalovirus (hCMV),[1] a member of the herpesvirus family, encodes a maturational protease required for processing of the assembly protein and virus replication (Baum et al., 1993; Welch et al., 1991). The hCMV protease is synthesized as a precursor protein that undergoes autoproteolytic cleavage to yield a mature, 28-kDa enzyme (Baum et al., 1993; Welch et al., 1991, 1993). The known substrates of the hCMV protease are the precursor protease itself and the viral assembly protein (Welch et al., 1991) that participates in the assembly of the viral capsid. Hydrodynamic measurements, such as sedimentation velocity and size-exclusion chromatography (Darke et al., 1996) and dynamic light scattering (Margosiak et al., 1996), indicate that the mature hCMV protease dimerizes. Enzyme assays of monomeric and dimeric species eluted from a size-exclusion column demonstrate that the dimer is the active species (Darke et al., 1996). The specific activity of the hCMV protease is dependent on protein concentration and these data are also consistent with a model in which only the dimeric species is active (Darke et al., 1996; Margosiak et al., 1996). The activity of the hCMV protease is affected by the presence of a variety of solutes, such as glycerol (Burck et al., 1994; Darke et al., 1996; Pinko et al., 1995), DMSO, sodium citrate, and sodium chloride (Margosiak et al., 1996; Pinko et al., 1995). Activation of hCMV by glycerol is due, in part, to stabilization of the dimer form (Darke et al., 1996). DMSO and citrate also enhance oligomerization of the hCMV protease; however, species larger than dimer are observed (Margosiak et al., 1996).

The protease is an attractive target for novel therapeutics directed against hCMV; a potential mechanism of enzymatic inhibition is disruption of dimerization. For example, a subnanomolar dissociative inhibitor of the heterodimeric herpes simplex virus ribonucleotide reductase has been reported that has *in vivo* antiviral activity (Liuzzi et al., 1994). In order to fully characterize the hCMV protease monomer ↔ dimer equilibrium we have performed sedimentation equilibrium, sedimentation velocity, and static light scattering measurements as a function of glycerol concentration, salt concentration, and temperature. A preliminary account of this work has been presented (Darke et al., 1996).

MATERIALS AND METHODS

The mature form of the cytomegalovirus protease containing the mutations V141G and V207G to abolish autoproteolysis was expressed and purified as previously described (Darke et al., 1996; Sardana et al., 1994). Enzyme was equilibrated in buffers containing 10 mM HEPES, 0.1 mM EDTA, and 1 mM DTT, pH 7.5, with variable concentrations of glycerol or NaCl using Bio-Rad Bio-Spin 6 spin columns. Protein concentration was measured spectrophotometrically. The molar extinction coefficient at 280 nm was determined by amino acid analysis to be 2.92×10^4 M^{-1} cm^{-1}. The partial specific volume of hCMV protease, $\bar{\nu}$, was calculated

* Address correspondence to this author at WP 26-344, Merck Research Laboratories, West Point, Pennsylvania 19486.

[⊗] Abstract published in *Advance ACS Abstracts*, November 15, 1996.

[1] Abbreviations: DMSO, dimethyl sulfoxide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HEPES, *N*-(2-hydroxyethyl)-piperazine-N′-(2-ethanesulfonic acid); hCMV, human cytomegalovirus; MES, 2-(*N*-morpholino)ethanesulfonic acid; TAPSO, 3-[*N*-tris(hydroxymethyl)methlyamino]-2-hydroxypropanesulfonic acid.

S0006-2960(96)01719-9 CCC: $12.00

to be 0.727 at 25 °C using the method of Cohn and Edsall (1943) and adjusted for temperature according to Durschlag (1986) such that $\bar{\nu} = 0.725$ at 20 °C. The solvent densities, ρ, were measured using an Anton Paar DMA 48 density meter. For equilibrium centrifugation experiments, 110-μL samples were loaded into six-channel (1.2-cm path) and 30-μL samples were loaded into two-channel (0.3-cm path) charcoal-Epon centrifuge cells under argon. Equilibrium analytical centrifugation was performed using a Beckman XL-A centrifuge and an An-60Ti rotor at speeds of 22 000 and 28 000 rpm and a temperature of 20 °C unless otherwise indicated. Scans were recorded using 0.001-cm point spacing and 10 readings were averaged at 230 or 280 nm. Equilibrium was judged to be achieved by the absence of systematic deviations in a plot of the difference between successive scans taken 4 h apart. At 20 °C, in the presence of 20% glycerol, equilibrium was generally achieved within 20 h; however, longer spin times were required at lower temperatures. It was verified that the absorbance readings were linear through 0.5 OD at 230 nm and 1.8 OD at 280 nm. Data at 230 nm never exceeded 0.4 OD, and the data obtained at 280 nm were truncated at 1.2 OD prior to analysis.

Equilibrium analytical centrifugation data were fit globally to several models using the program NONLIN (Johnson et al., 1981). The two models that were considered are an ideal monomer ↔ dimer equilibrium

$$C(r) = \delta + C_{10}\exp(\sigma_1\xi) + {C_{10}}^2\exp(\ln K_2 + 2\,\sigma_1\xi) \quad (1)$$

and a nonideal monomer ↔ dimer equilibrium

$$C(r) = \delta + C_{10}\exp\{\sigma_1\xi - 2BM_1[C(r) - \delta]\} + {C_{10}}^2\exp\{\ln K_2 + 2[\sigma_1\xi - 2BM_1[C(r) - \delta]]\} \quad (2)$$

where $C(r)$ is the total protein concentration (milligrams per milliliter) at a given radius r, δ is the baseline offset, C_{10} is the monomer concentration at the reference distance r_0, and σ_1 is the reduced molecular weight of the monomer

$$\sigma_1 = \frac{M_1(1 - \bar{\nu}\rho)\omega^2}{RT} \quad (3)$$

where M_1 is molecular weight of the monomer and ω is the angular velocity of the rotor.

$$\xi = (r^2 - {r_0}^2)/2,$$

K_2 is the dimerization constant in units of (milligrams per milliliter)$^{-1}$, and B is the second virial coefficient. Because the protein concentration units are expressed in milligram per milliliter, the product of the second virial coefficient and the monomer molecular weight, BM_1, is in units of (milligrams per milliliter)$^{-1}$. For computational reasons the natural logarithm of K_2 is the actual fitted quantity.

Sedimentation velocity experiments were performed using 400-μL samples loaded into two-channel (1.2-cm path) carbon-Epon centerpieces in an An-60Ti titanium rotor. The rotor was equilibrated at 20 °C for 1 h prior to each experiment. The rotor speed was 45 000 rpm. Scans were recorded at 230 or 280 nm at 15-min intervals with a point spacing of 0.003 cm in continuous mode. The sedimentation coefficients were obtained by fitting the data using the program SVEDBERG (Philo, 1994) and were corrected for buffer density and viscosity. Similar values of the sedimentation coefficients were obtained when the data were fit using time-derivative, radial distance derivative, and second-moment methods. For the data obtained in the absence of glycerol, K_d was obtained by fitting the velocity data to an expression appropriate for a weight-averaged sedimentation coefficient in a rapid monomer ↔ dimer equilibrium

$$s_w = \frac{s_m[M] + 2s_D[D]}{[M] + 2[D]} \quad (4)$$

where s_W, s_M, and s_D are the weight-average, monomer, and dimer sedimentation coefficients, respectively, and [M] and [D] are the molar concentrations of monomer and dimer, respectively. For a simple monomer ↔ dimer equilibrium it can be shown that the concentration of monomer is given by

$$[M] = \frac{-1 + (1 + 8K_2[E]t)^{1/2}}{4K_2} \quad (5)$$

where $[E]_t$ is the total enzyme molar concentration in monomer equivalents.

Table 1: Apparent Molecular Weights of hCMV Protease

% glycerol	method	[hCMV protease] (mg/mL)	$M_{w,app}$ (kg/mol)
0	sedimentation equilibrium[a]	0.02	29.6
	sedimentation velocity[b]	0.05	30.1
	sedimentation equilibrium	1.5	34.4
	static light scattering[c]	2.0	33.9
20[d]	sedimentation equilibrium	0.008	39.5
	sedimentation velocity	0.02	35.3
	sedimentation equilibrium	1.5	42.1
	static light scattering	2.6	44.0

[a] Results obtained from a fit of data collected at 22 000 and 26 000 rpm to a single ideal species model using NONLIN (Johnson & Correia, 1981). [b] Molecular weight derived from sedimentation coefficients and diffusion constants obtained by fitting sedimentation velocity profiles using SVEDBERG (Philo, 1994). Data were recorded at 45 000 rpm at 15-min intervals. [c] Static light scattering data were recorded with a Wyatt Dawn F detector downstream from a Waters size-exclusion column (Ultrahydrogel). $M_{w,app}$ was obtained from Debye plots using ASTRA (Wyatt Instruments). [d] The value of $\bar{\nu}$ used in sedimentation calculations was corrected for preferential hydration induced by glycerol (Gekko & Timasheff, 1981) as described in the Results section.

RESULTS

Table 1 shows the apparent molecular weights ($M_{w,app}$) of the hCMV protease obtained by sedimentation and light-scattering methods as a function of protein and glycerol concentrations. There is reasonably good agreement among the various techniques for the values of $M_{w,app}$ obtained at a given protein and glycerol concentration. In the absence of glycerol, the apparent molecular weights are close to that expected for the monomer of hCMV protease ($M_r = 27\,960$) at protein concentrations of < 0.1 mg/mL. At higher enzyme concentrations of 1.5−2 mg/mL, the molecular weights increase slightly to around 33 000 but do not indicate complete dimerization even at the highest concentrations examined. Addition of 20% glycerol results in a strong enhancement of dimerization, such that the apparent molecular weights are appreciably higher than the monomer, even

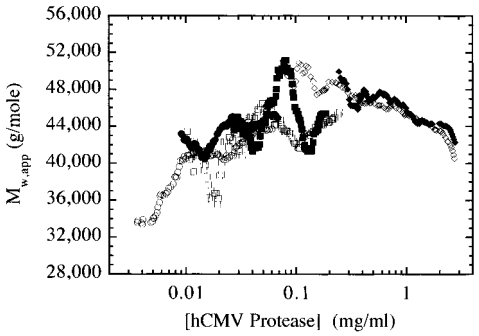

FIGURE 1: Plot of $M_{w,app}$ vs enzyme concentration. Data obtained in buffer containing 20% glycerol at protein loading concentrations of 0.05, 0.125, and 1.5 mg/mL and rotor speeds of 20 000 and 26 000 rpm. $M_{w,app}$ was obtained from the absorption vs radial distance scans using Beckman XL-A data analysis software version 2.0 with a 40-point smoothing.

at the lowest protein concentrations examined. However, at high enzyme concentrations of 1.5−2 mg/mL (53−93 μM), $M_{w,app} \sim 43\,000$, which is considerably less than expected for a dimer ($M_r = 55\,910$). Note that under similar conditions we previously found that $K_d \sim 0.6$ μM (Darke et al., 1996). There was no evidence for any species larger than dimer in sedimentation equilibrium experiments performed at speeds as low as 18 000 rpm (data not shown).

We have considered three possible explanations for this low value of $M_{w,app}$ at concentrations well above K_d: (1) an error in the calculation of $\bar{\nu}$, (2) heterogeneity in the enzyme preparation, such that a fraction of the protein is incompetent to dimerize, or (3) thermodynamic nonideality. In the case of sedimentation methods, $M_{w,app}$ is influenced by the calculated value of $\bar{\nu}$. The good agreement between the predicted monomer molecular weight and $M_{w,app}$ in the absence of glycerol indicates that $\bar{\nu}$ is likely correct. However, glycerol is known to preferentially hydrate proteins, such that the value of $\bar{\nu}$ increases with increasing glycerol concentration (Gekko & Timasheff, 1981). The effect of glycerol on $\bar{\nu}$ is similar for four proteins examined by Gekko and Timasheff, and we calculate an average correction factor for $\bar{\nu}$ of $\Delta\bar{\nu}/\Delta$(% volume glycerol) = (3.33 ± 0.38) × 10^{-4} (average of four proteins) from their data. This factor was used in the molecular weight calculations in Table 1, and in any case, it results in less than a 5% change in the calculated values of $M_{w,app}$ at 20% glycerol. In contrast to sedimentation methods, the measurement of $M_{w,app}$ by static light scattering is not dependent on the value of $\bar{\nu}$. Thus, the good agreement between the light-scattering and sedimentation results further suggests that $\bar{\nu}$ is not in serious error even in the presence of 20% glycerol.

Additional equilibrium sedimentation studies of hCMV were used to discriminate between sample heterogeneity and nonideality and to define the association thermodynamics. In sedimentation equilibrium of a homogeneous ideal or nonideal associating system the value of $M_{w,app}$ obtained at each point in a cell should be a function of protein concentration but should be independent of protein loading concentration or rotor speed. Figure 1 shows a plot of $M_{w,app}$ vs protein concentration in 20% glycerol obtained from cells loaded at three different concentrations and centrifuged at two rotor speeds. There is no systematic dependence of $M_{w,app}$ on loading concentration or rotor speed, suggesting that the system is in a homogeneous mass-action equilibrium. The value of $M_{w,app}$ increases with protein concentrations up to 0.1−0.2 mg/mL, as expected for an associating system, but decreases at higher enzyme concentrations. This decrease at high protein concentration is characteristic of positive nonideality (Tanford, 1961). Qualitatively similar behavior is observed in the absence of glycerol or in the presence of NaCl (data not shown).

Sedimentation equilibrium data sets were collected using cells loaded over a wide concentration range of 0.008−1.5 mg/mL and centrifuged at rotor speeds of 22 000 and 28 000 rpm. The complete data sets were analyzed using the global nonlinear least-squares algorithm NONLIN (Johnson et al., 1981). Having defined that hCMV protease undergoes a reversible dimerization, two different models were considered: ideal monomer ↔ dimer and nonideal monomer ↔ dimer. In all of the fits described, the value of the monomer molecular weight is fixed at the value derived from the amino acid sequence ($M_r = 27\,960$), and the values of ln K_2 and *BM* are treated as global fitting parameters. Figure 2 shows an overlay of the data obtained in the presence of 20% glycerol and the best fit to the ideal monomer ↔ dimer model, the insets show the residuals, and Table 2 contains the best-fit parameters. In 20% glycerol the data fit poorly to the ideal model: the value of the RMS deviation is about 2-fold above the noise level and systematic variations are observed in the residuals. Figure 3 shows that a much better fit is obtained using the nonideal model where the value of *BM* is not constrained at 0: the RMS value drops to 0.00465 and the residuals are randomly distributed about zero. The value of ln K_2 corresponds to a K_d for dimerization of 5.7 μM.

Relatively poor fits are also obtained when the data obtained in the absence of glycerol are fit to the ideal model (Table 2), but substantial improvement occurs when *BM* is allowed to vary. Figure 4 shows that the best nonideal fit in the absence of glycerol is also characterized by little systematic deviation in the residuals. The value of ln K_2 obtained from the fit corresponds to a K_d of 59 μM. The value of *BM* of 35 mL/g is quite close to that obtained in the presence of glycerol (25.4 mL/g). Thus, glycerol serves to enhance dimerization without substantially affecting the second virial coefficient. As would be expected, for hCMV protease dissolved at an intermediate concentration of glycerol of 10%, the value of K_d is intermediate between that obtained at 0 and 20% glycerol, with a similar value of *BM* (Table 2).

For the measurements obtained in 10% glycerol, the data were further analyzed to probe for any systematic dependence of the fitting parameters on rotor speed or protein loading concentration. With both ln K_2 and *BM* as floating parameters, the value of ln K_2 increases slightly with increasing rotor speed from 22 000 to 28 000 rpm but then decreases at 34 000 rpm. Note that for a heterogeneous system the value of ln K_2 is expected to decrease at higher rotor speeds. In contrast, the value of *BM* shows a weak systematic increase with rotor speed. This increase may reflect the limitations in the simple model used to describe the nonideal association (see Discussion). For samples analyzed at only one or two protein loading concentrations it is not possible to obtain unique fits with both ln K_2 and *BM* as floating parameters; therefore, these data are analyzed by fixing either ln K_2 or *BM* at the value determined in the global fit and allowing the other parameter to vary. With the value of ln K_2 fixed, the value of *BM* is similar to the value obtained

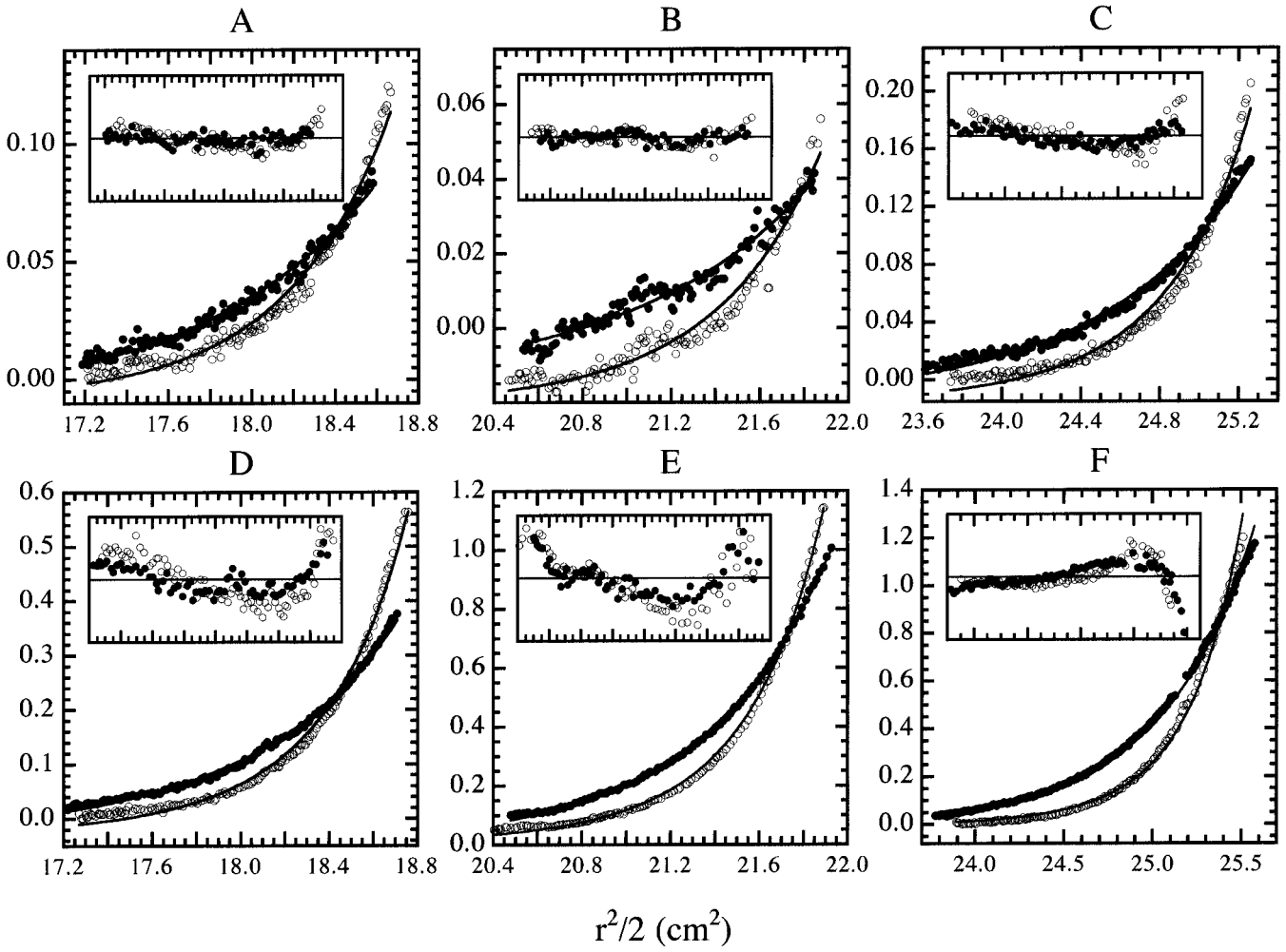

FIGURE 2: Analysis of sedimentation equilibrium of hCMV in 20% glycerol at 20 °C: ideal monomer ↔ dimer fit. The data were obtained at rotor speeds of (●) 20 000 and (○) 26 000 rpm. The protein loading concentrations are (A) 0.008, (B) 0.02, (C) 0.05, (D) 0.125, (E) 0.3, and (F) 1.5 mg/mL. The data in panel A were obtained at 230 nm and the other channels were obtained at 280 nm. The path length for panels A−E was 1.2 cm and for panel F it was 0.3 cm. The solid line is the global fit of all 12 channels using the ideal monomer ↔ dimer model. The insets show the residuals, where for clarity only every other data point is shown. The scale for the residuals is −0.03 to 0.03 OD, except for panel F where the scale is −0.08 to 0.08 OD. The results of the fit are given in Table 2.

with the global fit for the data at 1.5 mg/mL and is slightly higher for the data at 0.125 and 0.3 mg/mL. Conversely, with the value of *BM* fixed, the values of ln K_2 obtained at either concentration range are quite close to the global fit value.

It should be noted that the hCMV protease used in these studies is purified from *Escherichia coli* inclusion bodies solubilized using a buffer containing 7 M urea and followed by a refolding step (Darke et al., 1996). The enzyme may also be purified from the soluble fraction.[2] Dimerization of this hCMV protease preparation was examined by sedimentation equilibrium in buffers containing 20% glycerol. The values of ln K_2 and *BM* are similar regardless of purification method.

The effect of salt was also investigated on the sedimentation equilibrium of hCMV protease. As was observed above, in 0.1 and 0.5 M NaCl the fit quality is improved by inclusion of a floating *BM* parameter (Table 3). The best fit value of BM is essentially unchanged in 0.1 M and slightly lower at 0.5 M NaCl than in the absence of salt, but the differences do not indicate a strong systematic effect of NaCl. There are more significant changes in the value of K_d. At 0.1 M salt the K_d decreases about two-fold below the value in the absence of salt, whereas at 0.5 M the value of K_d decreases to close to the value obtained in 20% glycerol.

Figure 5 shows a van't Hoff plot of the temperature dependence of hCMV protease dimerization constant over a range of 5−30 °C obtained by sedimentation equilibrium in 10% glycerol. The values of ln K_2 were obtained from fits to the nonideal monomer ↔ dimer model. The value of BM increases weakly with increasing temperature, varying from 33.3 mL/g at 5 °C to 54.0 mL/g at 30 °C. The temperature dependence of ln K_2 is well described by a linear fit, indicating that $\Delta H°$ is approximately constant over this range. The fitted value of the enthalpy change is $\Delta H° = -11.3 \pm 1.3$ kcal/mol and at 20 °C $\Delta G° = -5.9$ kcal/mol, giving $\Delta S° = -18$ cal/(K mol).

The hCMV monomer ↔ dimer equilibrium was also characterized by sedimentation velocity measurements in the presence of 20% glycerol and in the absence of glycerol. In the presence of 20% glycerol a single boundary is observed at both low (0.08 mg/mL) and high (3 mg/mL) protein concentrations (Figure 6). Figure 7A shows a plot of the

[2] The initiating methionine residue is completely processed in the insoluble fraction, whereas only about 50% processing occurs in the soluble fraction (L. Waxman, unpublished results). The kinetic properties of these two forms are indistinguishable, and the difference in molecular weight is too small to affect the sedimentation results.

Table 2: Analysis of Sedimentation Equilibrium Data for hCMV Protease: Effect of Glycerol

glycerol concentration[a]	condition	model	ln K_2[b]	K_d (μM)	BM[c] (mL/g)	RMS[d] ($\times 10^{-3}$)
0%		ideal	−1.305 (−1.653, −0.945)			9.32
		nonideal	0.186 (−0.086, 0.457)	59	35.4 (31.1, 39.6)	5.90
10%		ideal	−0.684 (−1.08, −0.276)			13.2
		nonideal	1.277 (0.961, 1.604)	20	39.1 (35.5, 42.6)	7.14
	22 000 rpm only	nonideal	1.231 (0.869, 1.605)	21	36.3 (31.7, 40.9)	6.67
	28 000 rpm only	nonideal	1.352 (1.004, 1.713)	19	42.5 (38.4, 46.7)	7.35
	34 000 rpm only	nonideal	0.949 (0.577, 1.331)	28	56.6 (50.9, 62.2)	8.55
	1.5 mg/mL only	nonideal	1.487 (1.000, 2.029)	16	39.1 (fixed)	12.1
	1.5 mg/mL only	nonideal	1.277 (fixed)	20 (fixed)	38.6 (33.9, 43.4)	12.3
	0.125, 0.3 mg/mL only	nonideal	1.169 (0.907, 1.445)	22	39.1 (fixed)	7.29
	0.125, 0.3 mg/mL only	nonideal	1.277 (fixed)	20 (fixed)	56.5 (38.3, 74.7)	6.96
20%		ideal	0.8776 (0.504, 1.276)			9.82[e]
		nonideal	2.534 (2.221, 2.872)	5.7	25.4 (23.0, 27.9)	4.65[e]
	soluble fraction[f]	ideal	0.956 (0.562, 1.376)			11.6
	soluble fraction[f]	nonideal	2.443 (2.102, 2.808)	6.2	33.9 (30.1, 37.5)	6.05
	100 mM HEPES, 30 °C[g]	ideal	0.514 (0.139, 0.915)			11.0
	100 mM HEPES, 30 °C[g]	nonideal	1.943 (1.615, 2.284)	10	28.7 (25.0, 32.4)	6.03

[a] All data were collected at 20 °C in buffer containing 10 mM HEPES, 0.1 mM EDTA, and 1 mM DTT, pH 7.5, plus additives indicated. Unless otherwise indicated, data at six loading concentrations from 0.008 to 1.5 mg/mL and two rotor speeds (22 000 and 28 000 rpm) were included in fit. [b] Natural logarithm of K_2, where the units of K_2 are $(mg/mL)^{-1}$. The values in parentheses refer to the 95% confidence intervals. The molar dissociation constant is given by $K_d = 2/(27\,956 \times K_2)$. [c] The product of the second virial coefficient and the monomer molecular weight. The values in parentheses refer to the 95% confidence intervals. [d] Root mean square deviation of the fit in absorbance units. [e] Data obtained at 20 000 and 26 000 rpm. [f] CMV protease purified from the soluble fraction of the *E. coli* cell lysate rather than refolding of urea-solubilized inclusion bodies. [g] The channel loaded at 0.008 mg/mL was omitted because the high UV absorbance of this buffer precludes measurements at 230 nm.

corrected sedimentation coefficients vs protein concentration. $s_{20,w}$ increases modestly from 3.3 S at the lowest protein concentration to 3.6 S at higher protein concentrations. The sedimentation traces obtained at 3 mg/mL were globally fit to obtain s and D, and the molecular weight was obtained from the Svedberg equation, giving a value of $M_r = 49\,100$ (48 600, 49 600) (95% confidence intervals). These data are consistent with a predominance of the dimeric species in 20% glycerol over the concentration range of 0.08−3 mg/mL.

In the absence of glycerol, a single moving boundary is detected at both low (0.05 mg/mL) and high (3 mg/mL) protein concentrations (Figure 8); similarly, a single boundary is also present at intermediate concentrations (data not shown). Analysis of the traces obtained at 0.08 mg/mL for s and D returns a value of $M_r = 30\,100$ (29 600, 30 600). These data confirm that hCMV protease is monomeric at low concentration in the absence of glycerol. The sedimentation coefficients obtained in the absence of glycerol increase with protein concentration (Figure 7A), and at the higher protein concentrations the data obtained in 0 and 20% glycerol converge to a similar value of $s_{20,w} \sim 3.6$ S. These results are consistent with a rapid equilibrium between the monomer and dimer species on the time scale of the sedimentation experiments such that the apparent sedimentation coefficient is a weight-average quantity. Thus, the data obtained in the absence of glycerol can be analyzed to obtain a value of K_d. Figure 7B shows a fit of the 0% glycerol data to eq 4; the best-fit parameters are $s_M = 2.38$, $s_D = 3.63$, and $K_2 = 6.0 \times 10^4\ M^{-1}$, which corresponds to $K_d = 17\ \mu M$. The frictional coefficient of an oligomer of identical subunits may be predicted from the monomer frictional coefficient using a theory developed by Kirkwood (Van Holde, 1975). For a monomer ↔ dimer system, the value of s_D is expected be $1.5 s_M$. From the nonlinear fit to sedimentation velocity data in 0% glycerol $s_D/s_M = 1.53$, which agrees well with theory.

DISCUSSION

For typical globular proteins of moderate size, nonideality is not often observed for dilute solutions of <5 mg/mL near neutral pH (Tanford, 1961). Hence, we were surprised to observe this effect in our equilibrium sedimentation studies of hCMV protease. The evidence for nonideality is the following: (1) At protein concentrations well above K_d in the presence of 20% glycerol, $M_{w,app}$ obtained by both sedimentation equilibrium and static light scattering is well below the expected value for the hCMV dimer; (2) $M_{w,app}$ increases with protein concentration in the range below 0.1−0.2 mg/mL, which is due to dimerization, but then decreases at higher protein concentration; and (3) good global nonlinear least-squares fits to the sedimentation equilibrium data obtained under all conditions that were examined requires a positive second virial coefficient. The apparent nonideality is not due to detectable heterogeneity in the hCMV protease preparation. The protein is ≥95% homogeneous by denaturing electrophoresis, electrospray mass spectrometry gives a single species within 10 mass units of the expected mass, and ≥95% of the hCMV protease is capable of dimerizing in 20% glycerol at high protein concentrations (Darke et al., 1996). The good overlay of sedimentation equilibrium data sets in Figure 1 and the absence of systematic variations of the fitted values of K_2 (Table 2) obtained at various protein loading concentrations and rotor speeds are characteristic of a homogeneous and reversible mass-action equilibrium system. In sedimentation velocity measurements, only a single moving boundary is observed at each protein concentration irrespective of the presence of glycerol, which is also consistent with a reversible monomer ↔ dimer equilibrium system. Thus, by both hydrodynamic and equilibrium methods the hCMV protease appears to be homogeneous. We have also been careful to avoid optical artifacts by maintaining sample absorption within the linear range

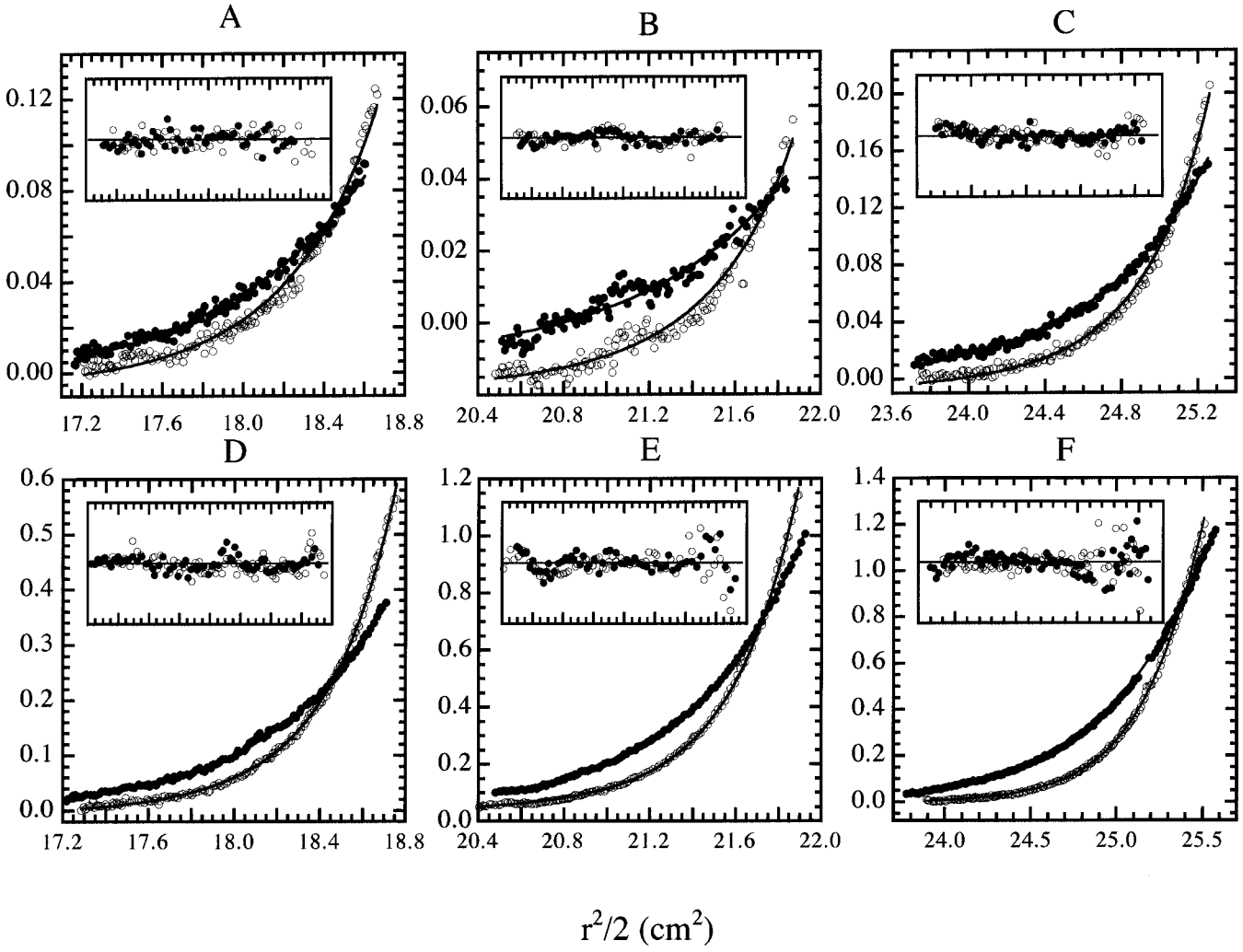

FIGURE 3: Analysis of sedimentation equilibrium of hCMV in 20% glycerol: nonideal monomer ↔ dimer fit. Refer to the legend for Figure 2 for details. The scale for the residuals is −0.03 to 0.03 OD. The results of the fit are given in Table 2.

for the XL-A detection system and avoiding generating steep absorption gradients which may contribute to Wiener skewing (Svensson, 1954).

The resolution of a nonideal association model is considered a difficult problem to fit since the effects of association and nonideality on molecular weight tend to cancel each other out and thus are highly correlated (Johnson et al., 1981). We have constrained these parameters by using data obtained over a broad concentration range and at two rotor speeds. In the fits described above the cross correlation coefficients are 0.71 (0% glycerol), 0.62 (10% glycerol), and 0.56 (20% glycerol), which are not very large. Johnson et al. (1981) have estimated that the lower limit for *BM* which can be resolved simultaneously with K_2 is $\sim$ 0.4 mL/g, which is considerably below the values reported here.

The chief sources of nonideality in proteins that have been previously described are excluded volume and electrostatic effects; generally, the latter is the dominant contribution (Tanford, 1961). For a rigid sphere the contribution of excluded volume is independent of molecular weight and corresponds to $BM = 4\bar{\nu} \sim 3$ mL/g (Tanford, 1961). However, for long, rod-shaped particles $BM = (L/d)\bar{\nu}$, where L is the length and d is the diameter, and thus higher values of the second virial coefficient are observed for asymmetric, fibrous proteins. For example, *BM* for myosin is 75 mL/g (Tanford, 1961) and for spectrin, $BM = 50$ mL/g (Ralston, 1992). The overall shape and hydration of proteins is often parameterized by calculation of the ratio of the experimentally determined frictional coefficient to the frictional coefficient expected for an unhydrated sphere, f/f_{min}. For typical globular proteins $f/f_{min} = 1.1-1.3$, for fibrous proteins $f/f_{min} = 2.3-6.8$, and for myosin the value is 3.5 (Cantor & Schimmel, 1980; Tanford, 1961). We calculated the value of f/f_{min} for the monomeric hCMV protease in the absence of glycerol by extrapolating the values of $s_{20,w}$ and $D_{20,w}$ to zero protein concentration, giving values of $s^0_{20,w} = 2.4$ S and $D^0_{20,w} = 7.3 \times 10^{-7}$ cm²/s. From these parameters, $f/f_{min} = 1.4-1.5$ for the hCMV monomer. Thus, the frictional ratio for hCMV is somewhat higher than is usually found for globular proteins but well below that expected for a highly asymmetric, fibrous protein. Thus, the hydrodynamic shape of hCMV is more similar to globular proteins, and the slightly higher value of f/f_{min} is due to either a slightly elongated shape or a high degree of hydration. However, given the value of the frictional ratio it seems unlikely that excluded volume effects can fully account for the observed nonideality of hCMV protease.

Electrostatic effects may also contribute to nonideality. Roark and Yphantis (1971) have presented an expression to predict the contribution of the Donnan effect to the second

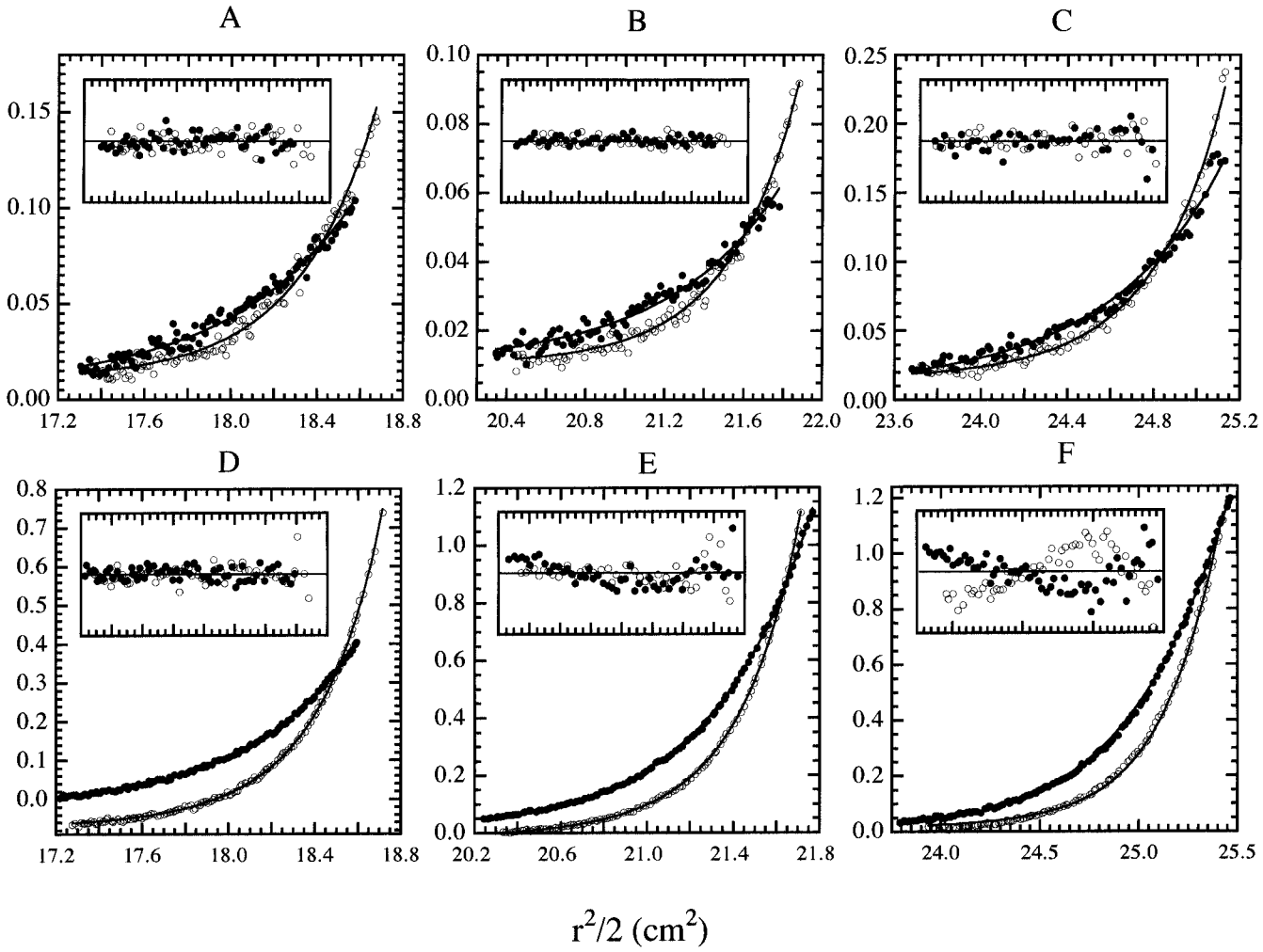

FIGURE 4: Analysis of sedimentation equilibrium of hCMV in 0% glycerol: nonideal monomer ↔ dimer fit. The data were obtained at rotor speeds of (●) 22 000 and (○) 28 000 rpm. Refer to the legend for Figure 2 for details. The scale for the residuals is −0.03 to 0.03 OD. The results of the fit are given in Table 2.

Table 3. Analysis of Sedimentation Equilibrium Data for hCMV Protease: Effect of NaCl

conditions[a]	model	Ln K_2[b]	K_d (μM)	BM[c] (mL/g)	RMS[d] ($\times 10^{-3}$)
no additive	ideal	−1.305 (−1.653, −0.945)			9.32
	nonideal	0.186 (−0.086, 0.457)	59	35.4 (31.1 39.6)	5.90
0.1 M NaCl[e]	ideal	−0.988 (−1.391, −0.573)			9.70
	nonideal	0.783 (0.443, 1.122)	33	35.8 (30.3, 41.1)	6.15
0.5 M NaCl	ideal	1.375 (1.017, 1.757)			10.3
	nonideal	2.213 (1.841, 2.627)	7.8	18.7 (14.5, 22.8)	7.93

[a] All data were collected at 20 °C in buffer containing 10 mM HEPES, 0.1 mM EDTA, and 1 mM DTT, pH 7.5, plus additives indicated. Unless otherwise indicated, data at six loading concentrations from 0.008 to 1.5 mg/mL and two rotor speeds (22 000 and 28 000 rpm) were included in fit. [b] Natural logarithm of K_2, where the units of K_2 are $(mg/mL)^{-1}$. The values in parentheses refer to the 95% confidence intervals. The molar dissociation constant is given by $K_d = 2/(27\,956 \times K_2)$. [c] The product of the second virial coefficient and the monomer molecular weight. The values in parentheses refer to the 95% confidence intervals. [d] Root mean square deviation of the fit in absorbance units. [e] Data obtained at 22 000 rpm only.

virial coefficient values determined in sedimentation equilibrium in the presence of a monovalent salt:

$$B = \frac{Z^2}{4M^2m} \tag{6}$$

where Z is the effective charge on the protein, M is the molecular weight of the protein, and m is the molality of the salt. For hCMV protease the isoelectric point is calculated from the amino acid sequence to be 5.7 and at pH 7.5 the net charge is calculated to be approximately −6. The actual net charge is likely to be less. Since the dominant contribution to the ionic strength in the buffers containing ≥0.1 NaCl is due to the monovalent salt, in 0.5 M NaCl the Donnan contribution to BM is estimated to be <0.6 mL/g and in 0.1 M NaCl BM is estimated at <3 mL/g. These values are far below those determined experimentally at the corresponding NaCl concentrations (Table 2). Note that the Donnan contribution to nonideality should be inversely proportional to the monovalent salt concentration, whereas we see a fairly weak dependence on ionic strength.

The McMillan−Mayer theory (Wills & Winzor, 1992) may be used to estimate the contribution of charge−charge interactions to B, assuming a symmetrically distributed net

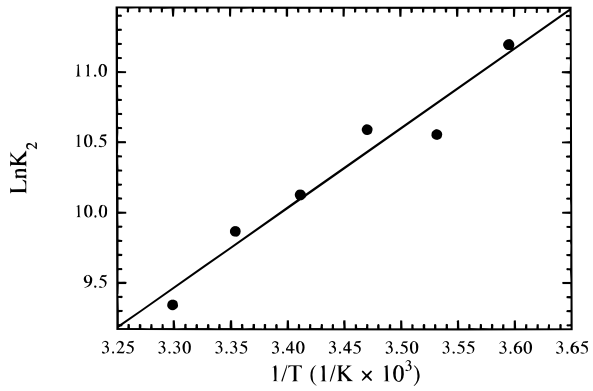

FIGURE 5: Temperature dependence of hCMV dimerization in 10% glycerol. Sedimentation equilibrium data was obtained as described in Materials and Methods, except that a single rotor speed of 25 000 rpm was employed and the temperature was varied from 5 to 30 °C. The data were fit to the nonideal monomer ↔ dimer model to obtain ln K_2 and BM. The solvent density was measured at each temperature and the calculated value of $\bar{\nu}$ was adjusted for temperature (Durchschlag, 1986).

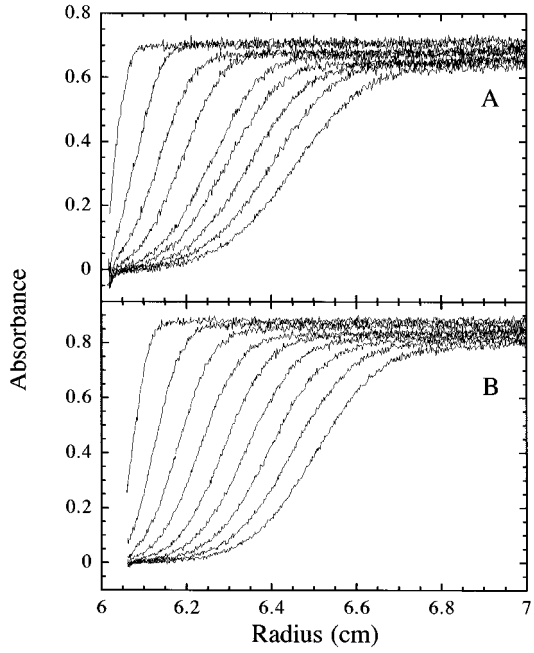

FIGURE 6: Sedimentation velocity of hCMV protease in 20% glycerol: A 0.08 mg/mL recorded at 230 nm in a 1.2-cm path length cell and (B) 3 mg/mL recorded at 280 nm in a 0.3-cm pathlength cell. The scans were taken at 20 °C, 45 000 rpm, using a point spacing of 0.002 cm. The scans were recorded at 15-min intervals up to 220 min; however, only every third scan is shown. The data plotted correspond to $\omega^2 t = 2.8 \times 10^{10}$, 8.8×10^{10}, 1.5×10^{11}, 8.8×10^{10}, 2.1×10^{11}, 2.7×10^{11}, 3.3×10^{11}, 3.9×10^{11}, 4.5×10^{11}, and 5.1×10^{11} s^{-1}.

charge on a spherical molecule:

$$B = \frac{Z^2(1 + 2\kappa R)}{4I(1 + \kappa R)^2} \qquad (7)$$

where κ is the Debye−Huckel inverse screening length, R is the radius of the macromolecule, and I is the ionic strength. In aqueous solutions near room temperature (Cantor & Schimmel, 1980), $\kappa = 0.33(I^{1/2})$ Å^{-1}. For the monomeric hCMV protease we approximate R as the hydrated radius of 28 Å derived from sedimentation velocity data. Under these circumstances, we estimate the value of BM due to charge−

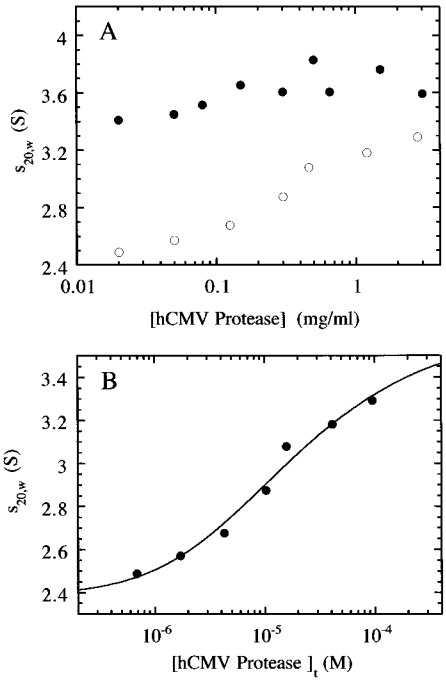

FIGURE 7: (A) Corrected sedimentation coefficients of hCMV protease as a function of protein concentration in either 0% (○) or 20% (●) glycerol. Sedimentation coefficients at each protein concentration were obtained by fitting 8 radials absorbance scans using the program SVEDBERG. The sedimentation coefficients were corrected using measured solvent densities and tabulated solvent viscosities and protein partial specific volumes. (B) Fit of corrected sedimentation coefficients obtained in 0% glycerol to eq 5. The best-fit parameters are $s_M = 2.38$, $s_D = 3.63$, and $K_2 = 6.0 \times 10^4$ $M^{-1,}$ which corresponds to $K_d = 17$ μM.

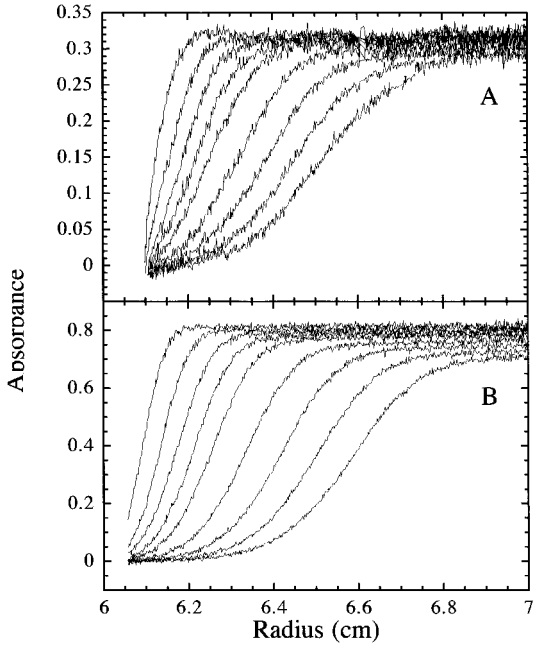

FIGURE 8: Sedimentation velocity of hCMV protease in 0% glycerol: (A) 0.05 mg/mL recorded at 230 nm in a 1.2-cm path length cell and (B) 2.78 mg/mL recorded at 280 nm in a 0.3-cm pathlength cell. The scans were taken at 20 °C, 45 000 rpm, using a point spacing of 0.002 cm. The scans were recorded at 15-min intervals up to 220 min. The data plotted correspond to $\omega^2 t = 2.8 \times 10^{10}$, 4.8×10^{10}, 6.8×10^{10}, 8.8×10^{10}, 1.1×10^{11}, 1.5×10^{11}, 1.9×10^{11}, 2.3×10^{11}, and 2.7×10^{11} s^{-1}.

charge interactions to be <0.2 mL/g at 0.5 M NaCl and <1.4 at 0.1 M NaCl. Like the Donnan effect, the predicted

contribution of charge−charge interactions to B is far less than is observed experimentally, and is even more strongly dependent on ionic strength.

In summary, we find that hCMV protease behaves as a nonideal associating protein in a variety of solvents, and the magnitude of nonideality is more than predicted by conventional theory for excluded volume effects or electrostatic interactions. Further structural and thermodynamic studies will be necessary to clarify the origin of this unusual behavior. Owing to recent advances in instrumentation and data analysis methods there has been a resurgence in the use of sedimentation methods to characterize protein association. As more proteins are characterized by sedimentation equilibrium, additional examples of unexpected nonideal behavior are likely to be found.

Solution nonideality may be biologically significant in the context of the solvent conditions in a cellular compartment or, as in the present case, within an assembling viral capsid. In the herpesviruses, capsid formation proceeds via an intermediate (B-capsid) that lacks DNA but contains assembly protein, which is not present in mature capsids. (Harper, 1994). By analogy to bacteriophage scaffolding proteins (Casjens & King, 1975), it is believed that the herpes assembly proteins aid in capsid formation. Interestingly, herpes simplex virus assembly protein associates to form a toroidal structure similar to that found within the center of B-capsids *in vivo* (Newcomb & Brown, 1991). The herpesvirus assembly proteins are cleaved at a site near the C-terminus by the action of the herpesvirus protease. The protease also cleaves itself to release an N-terminal catalytic domain; only the processed forms of the protease and assembly proteins are found associated with B-capsids (Harper, 1994). It is believed that these processing reactions occur within the capsid (Robertson et al., 1996). Molecular crowding effects will tend to increase the rates and extents of macromolecular associations (Zimmerman, 1993). Thus, within the immature capsid, dimerization of the hCMV protease and its interaction with assembly protein may be dramatically enhanced. In addition, Herzfeld has shown that in crowded mixed solutions, entropically driven spatial self-organization of macromolecules can be explained in terms of excluded volume effects (Herzfeld, 1996), and such concepts have been applied to polymerization of sickle cell hemoglobin within red cells and formation of cytoskeletal filaments. Similarly, assembly of herpesvirus capsids may be enhanced by molecular crowding, and these effects may be enhanced by the markedly nonideal behavior we have observed for the mature hCMV protease.

The dimer dissociation constants for hCMV protease determined in the present study in the absence of glycerol (10 mM HEPES, pH 7.5, at 20 °C) are K_d = 59 μM by sedimentation equilibrium and K_d = 17 μM by sedimentation velocity. The differences between these values are slight and may reflect the fact that nonideality was not taken into account in the analysis of the hydrodynamic data (see below). In a recent report, Margosiak et al. (1996) determined a dissociation constant of K_d = 8 μM in the absence of glycerol (25 mM MOPS and 2% DMSO, pH 7.2, at 25 °C) by activity assays. There are a number of differences between their conditions and ours. Their buffers include 2% DMSO, which is known to enhance dimerization. Also, without accompanying data concerning the rates at which hCMV achieves dimerization equilibrium under their experimental conditions, it is not clear that the K_d values determined by activity measurements truly reflect equilibrium dissociation constants.

The present study also confirms previous reports that glycerol and other solutes enhance dimerization of hCMV protease. Using sedimentation equilibrium, we find that K_d decreases about 3-fold (to 20 μM) in 10% glycerol and about 10-fold (to 5.7 μM) in 20% glycerol. In our previous study, activity assays give values of K_d = 6.6 μM in 10% glycerol and 0.55 μM in 20% glycerol (both in 52 mM TAPSO, 52 mM MES, 100 mM diethanolamine, and 0.05% bovine serum albumin, pH 7.5, at 30 °C), and fitting of size-exclusion chromatography data gives K_d = 0.55 μM in 20% glycerol (100 mM HEPES, pH 7.5, at 30 °C) (Darke et al., 1996). Thus, a similar enhancement of dimerization by glycerol is observed by either activity assays or sedimentation equilibrium. Under the same conditions of the size-exclusion chromatography experiment in 20% glycerol, sedimentation equilibrium gives K_d = 10.2 μM (Table 2). Again, this difference may reflect the influence of thermodynamic nonideality. In the NONLIN algorithm, and in most analyses of associating systems, it is assumed there is a single value of the second virial coefficient common to all species present (Adams & Fujita, 1963); this leads to the convenient result that equilibrium constants are simple ratios of concentrations and are not affected by nonideality. Note, however, that some workers have questioned the validity of this assumption (Jacobsen & Winzor, 1992). Thus, if the value of B for the dimer is different from the monomer this would influence the apparent value of K_d determined by various methods. In any case, an 18-fold difference in K_d corresponds to only about 1.6 kcal/mol difference in free energy. However, the K_d values that we have obtained in buffers containing 20% glycerol are more than 2 orders of magnitude higher than the value of K_d = 1.9 nM in 25% glycerol (25 mM MOPS and 2% dimethyl sulfoxide, pH 7.2, at 25 °C) reported by Margosiak et al. (1996). This discrepancy appears to be much greater than can be accounted for by differences in experimental conditions. The much lower apparent K_d reported by their activity measurements may be due to the effects of substrate-induced dimerization.

We have also found that NaCl enhances dimerization; in 0.5 NaCl K_d decreases about 8-fold relative to the data obtained in the absence of added salt. Enzyme kinetic and size-exclusion methods (Darke et al., 1996) indicate that NaCl decreases K_d (U. Schmidt, unpublished observations). Margosiak et al. (1996) reported that NaCl weakly enhances the activity of hCMV protease.

There is precedence for the enhancement of protein oligomerization by glycerol (Bonafe et al., 1991; Pedrosa & Ferreira, 1994), and in these cases it is believed that the enhancement of oligomerization induced by glycerol is due, in part, to preferential hydration of the protein, i.e., exclusion of glycerol from the protein surface region (Gekko & Timasheff, 1981). This situation is thermodynamically unfavorable and can be partially relieved by reduction of the exposed surface area of the protein by folding or oligomerization. Anions high in the Hofmeister series are also known to induce preferential hydration. Indeed, citrate ion, which is high in the Hofmeister series, is known to activate both the closely related HSV-1 protease (Hall & Darke, 1995) and hCMV protease (Margosiak et al., 1996; U. Schmidt, unpublished observations). In the former case,

citrate activation is accompanied by an enhanced dimerization (U. Schmidt, unpublished observations), whereas in the latter case light-scatter measurements suggest formation of large, heterogeneous aggregates (Margosiak et al., 1996). Using sedimentation equilibrium, we have not observed any hCMV aggregates higher than dimer in the presence of glycerol, citrate, or NaCl (J. L. Cole, unpublished results). In contrast to glycerol and citrate, NaCl does not induce preferential hydration yet still has the ability to enhance dimerization of hCMV protease. Further studies are required to define the mechanism of solute effects on the hCMV protease.

The temperature-dependence studies of dimerization reveal that hCMV protease association is characterized by a negative enthalpy change and a negative entropy change, as is observed in many protein association reactions (Ross & Subramanian, 1981). Thus, the dimerization is not driven by the hydrophobic effect but by hydrogen bonding and/or van der Waals interactions.

ACKNOWLEDGMENT

I thank Paul Darke, Dawn Hall, and Lloyd Waxman for supplying purified hCMV protease, Mo Sardana for amino acid analysis, Paul Darke, Tom Laue, David Olsen, and Lloyd Waxman for valuable discussions, and Lawrence Kuo for guidance and encouragement.

REFERENCES

Adams, E. T., Jr., & Fujita, H. (1963) in *Ultracentrifugal Analysis in Theory and Experiment*. (Williams, J. W., Ed.) Academic Press, New York.

Baum, E. Z., Bebernitz, G. A., Hulmes, J. D., Mizithras, Y. P., Jones, T. R., & Gluzman, Y. (1993) *J. Virol. 67*, 497−506.

Bonafe, C. F. S., Villas-Boas, M., Suarez, M. C., & Silva, J. L. (1991) *J. Biol. Chem. 266*, 13210−13216.

Burck, P. J., Berg, D. H., Luk, T. P., Sassmannhausen, L. M., Wakulchik, M., Smith, D. P., Hsiung, H. M., Becker, G. W., Gibson, W., & Villarreal, E. C. (1994) *J. Virol. 68*, 2937−2946.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Vol. III, W. H. Freeman and Co., New York.

Casjens, S., & King, J. (1975) *Annu. Rev. Biochem. 44*, 555−611.

Cohn, E. J., & Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, p 157, Rheinhold, New York.

Darke, P. L., Cole, J. L., Waxman, L., Hall, D. L., Sardana, M. K., & Kuo, L. C. (1996) *J. Biol. Chem. 271*, 7445−7449.

Durchschlag, H. (1986) in *Thermodynamic Data for Biochemistry and Biotechnology* (Hinz, H. J., Eds.), pp 45−128, Springer-Verlag, New York.

Gekko, K., & Timasheff, S. N. (1981) *Biochemistry 20*, 4667−4676.

Hall, D. L., & Darke, P. L. (1995) *J. Biol. Chem. 270*, 22697−22700.

Harper, D. R. (1994) *Rev. Med. Virol. 4*, 119−128.

Herzfeld, J. (1996) *Acc. Chem. Res 29*, 31−37.

Jacobsen, M. P., & Winzor, D. J. (1992) *Biophys. Chem. 45*, 119−132.

Johnson, M. L., Correia, J. J., Yphantis, D. A., & Halvorson, H. R. (1981) *Biophys. J. 36*, 575−588.

Liuzzi, M., Dezlel, R., Moss, N., Beaulieu, P., Bonneau, A.-M., Bousquet, C., Chafouleas, J. G., Garneau, M., Jaramillo, J., Krogsrud, R. L., Lagace, L., McCollum, R. S., Nawoot, S., & Guindon, Y. (1994) *Nature 372*, 695−698.

Margosiak, S. A., Vanderpool, D. L., Sisson, W., Pinko, C., & Kan, C.-C. (1996) *Biochemistry 35*, 5300−5307.

Newcomb, W. W., & Brown, J. C. (1991) *J. Virol. 65*, 613−620.

Pedrosa, C., & Ferreira, S. T. (1994) *Biochemistry 33*, 4046−4055.

Philo, J. S. (1994) in *Modern Analytical Ultracentrifugation* (Shuster, T. M., and Laue, T. M., Eds.), pp 156−170, Birkhauser, Boston, MA.

Pinko, C., Margosiak, S. A., Vanderpool, D., Gutowski, J. C., Condon, B., & Kan, C.-C. (1995) *J. Biol. Chem. 270*, 23634−23640.

Ralston, G. B. (1992) *Biophys. Chem. 44*, 175−186.

Roark, D. E., & Yphantis, D. A. (1971) *Biochemistry 10*, 3241−3249.

Robertson, B. J, McCann, P. J., III, Matusick-Kumar, L., Newcomb, W. W., Brown, J. C., Colonno, R. J., & Gao, M. (1996) *J. Virol. 70*, 4317−4328.

Ross, P. D., & Subramanian, S. (1981) *Biochemistry 20*, 3096−3102.

Sardana, V. V., Wolfgang, J. A., Veloski, C. A., Long, W. J., LeGrow, K., Wolanski, B., Emini, E. A., & LaFemina, R. L. (1994) *J. Biol. Chem. 269*, 14337−14340.

Svensson, H. (1954) *Opt. Acta 1*, 25−32.

Tanford, C. (1960) *Physical Chemistry of Macromolecules* John Wiley and Sons, New York.

Van Holde, K. E. (1975) *The Proteins* (Neurath, H., & Hill, R. L., Eds.) pp 225−291, Academic Press, New York.

Welch, A. R., Woods, A. S., McNally, L. M., Cotter, R. J., & Gibson, W. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10792−10796.

Welch, A. R., McNally, L., Hall, M., & Gibson, W. (1993) *J. Virol. 67*, 7360−7372.

Wills, P. R., & Winzor, D. J. (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., & Horton, J. C., Eds.) pp 311−330, Royal Society of Chemistry, Cambridge, England.

Zimmerman, S. B. (1993) *Biochim. Biophys. Acta 1216*, 175−185.

BI961719F