Molecular Weight of Escherichia coli β-Galactosidase in Concentrated Solutions of Guanidine Hydrochloride

ROBERT P. ERICKSON*
Laboratory of Chemical Biology, National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, Bethesda, Md. 20014, U.S.A.

(Received 5 June 1970)

The molecular weight of Escherichia coli β-galactosidase was determined in 6M- and 8M-guanidine hydrochloride by meniscus-depletion sedimentation equilibrium, sedimentation velocity and viscosity. Sedimentation equilibrium revealed heterogeneity with the smallest component having a molecular weight of about 50,000. At lower speeds, the apparent weight-average molecular weight is about 80,000. By use of a calculation based on an empirical correlation for proteins that are random coils in 6M-guanidine hydrochloride, sedimentation velocity gave a molecular weight of 91,000, and the intrinsic viscosity indicated a viscosity-average molecular weight of 84,000. Heating in 6M-guanidine hydrochloride lowered the viscosity of β-galactosidase in a variable manner.

The molecular weight of β-galactosidase (β-D-galactoside galactohydrolase, EC 3.2.1.23) purified from Escherichia coli has previously been determined in 6M-guanidine hydrochloride and found to be in general accord with the 'monomer' molecular weight of 135,000 (Craven, 1967; Ullmann, Goldberg, Perrin & Monod, 1968a). The latter study utilized low-speed sedimentation equilibria, which would poorly distinguish multiple components, and an empirical correlation of apparent molecular weights in guanidine hydrochloride [lim_{x->0} M_{app} (1-γp)] to actual molecular weights determined for several well-studied proteins, thus neglecting variable buoyancies. This result has been used to argue that the 'monomer' is one polypeptide chain (Ullmann, Jacob & Monod, 1968b) despite the fact that it is easily dissociable into smaller subunits by detergent, dilute acid or dilute base (Wallenfels, Sund & Weber, 1963; Steers, Craven & Anfinsen, 1965a). We find the molecular weights of β-galactosidase determined in concentrated solutions of guanidine hydrochloride to be significantly lower than is compatible with a single chain of molecular weight 135,000. A preliminary report of these results has been published (Erickson & Steers, 1969a). Evidence for blocked N-terminal amino acids in β-galactosidase as pyrrolidonecarboxylic acid has been reported elsewhere (Erickson & Steers, 1969b).

* Present address: Department of Pediatrics, University of California Medical Center, San Francisco, Calif. 94122, U.S.A.

MATERIALS AND METHODS

Purification of β-galactosidase. The β-galactosidase used in this study was prepared from the regulator-constitutive strain of E. coli K12 3300 by the procedure of Marchesi, Steers & Shifrin (1969). Briefly, the bacterial lysate was fractionated with (NH4)2SO4 and the 20-40% saturation fraction was subjected to gel filtration on Sepharose 4-B (Pharmacia Fine Chemicals, Uppsala, Sweden) in buffer B (0.1M-NaCl-0.01M-MgCl2-0.05M-2-mercaptoethanol-0.05M-tris-HCl buffer, pH 7.4). The major peak of activity was further purified by ion-exchange chromatography on DEAE-cellulose (Whatman DE23; Reeve Angel, Clifton, N.J., U.S.A.) with a combination pH- NaCl gradient. Sometimes a Sephadex G-200 (Pharmacia) purification step was interposed between the first gel-filtration and the ion-exchange steps. Purity was checked by polyacrylamide disc gel electrophoresis. Amidoschwarz (Allied Chemical Corp., New York, N.Y., U.S.A.) staining revealed no contaminating bands when 100-200 μg quantities of the purified enzyme were subjected to electrophoresis on 5% polyacrylamide gels at 3 mA/tube for 3 h. The purified protein was stored under 50% (w/v) (NH4)2SO4 at 4°C.

Preparation of solutions of β-galactosidase in concentrated aqueous guanidine hydrochloride. Spectral-quality guanidine hydrochloride was obtained from Aldrich Chemical Co., Milwaukee, Wis., U.S.A., or from Mann Research Laboratories, New York, N.Y., U.S.A. A β-galactosidase solution in buffer B was diluted to the desired concentration with (or directly dialysed against) 6M- or 8M-guanidine hydrochloride-0.1M-2-mercaptoethanol-0.1M-tris-HCl buffer, pH 7.5, and dialysed against three 100ml portions of the buffer for 3 days. The resultant solution (and control samples of dialysate) were filtered with pressure through 0.7-1 μm millipore filters (Millipore Filter Corp., Bedford,
Mass, U.S.A.) Concentrations in the usual buffers were calculated by using an extinction coefficient at 280 nm of 2.09 mg/ml for a 1 cm light-path (Craven, Steers & Anfinsen, 1965). By difference spectroscopy in the Cary model 15 recording spectrometer, an extinction coefficient at 280 nm of 2.207 mg/ml (1 cm light-path) was found for β-galactosidase in 6M- or 8M-guanidine and this value was used to calculate the protein concentrations of guanidine solutions.

**Sedimentation-equilibrium measurements.** High-speed sedimentation-equilibrium experiments were performed as described by Yphantis (1964) at 20°C, double-sector duraluminum-filled epoxy or six-sector carbon-filled epoxy centrepieces and sapphire windows being used. (Base fluid FC43 Beckman, Palo Alto, Calif., U.S.A.) was used in the two- but not the six-sector experiments. Optical blanks were obtained by a preliminary run with water in the cells. One window assembly was then removed (without dismantling the windows or cell centrepiece), water was carefully removed and the cell was dried over P2O5. The cell was then ready for the experimental run. The concentrations across the cell were recorded with the interference optical system after 24 h at each increasing speed. Equilibrium had been achieved, as exposures at 18 h were occasionally taken and were not measurably different from those at 24 h. The Spinco model E instrument equipped with electronic speed control was used and photographic plates were measured with a Nikon 6C comparator. Densities of concentrated aqueous solutions of guanidine hydrochloride were taken from Kawahara & Tanford (1966). The effective specific volume, ϕ, was taken as 0.75 ml/g, 0.01 ml/g less than the previously determined partial specific volume of 0.76 (Wallenfels, Sund & Weber, 1963) following Tanford (Tanford, Kawahara & Lapanje, 1967). A computer program in BASIC fitted linear least-squares slopes for plots of ln c versus r², determined true slopes for the second component by subtracting the extrapolated first slope from the observed second slope when polydispersity was indicated by breaks in the ln c versus r² plots, calculated molecular weights, and determined amounts of each component by integrating areas under the curve of c versus r².

**Sedimentation-velocity determinations.** The duraluminum-filled epoxy, capillary artificial-boundary, double-sector cell was used with sapphire windows. The determinations were carried out at 56 000 rev./min and 20°C by using schlieren optics in the analytical ultra-centrifuge equipped with electronic speed control. Measurements of boundary movement were made with the Nikon 6C comparator.

**Viscosity measurements.** Viscosities were determined at 20.0 ± 0.01°C in an Ostwald viscometer of the dilution type. Protein concentrations were checked on the final dilution of the series. Some determinations were made with a four-chamber Ubbelohde viscometer that allowed variation in shear forces.

**RESULTS**

Sedimentation-equilibrium experiments by the meniscus-depletion method gave weight-average values for the molecular weight of β-galactosidase of 70 000–80 000 in concentrated guanidine hydro-chloride at moderate speeds (Table 1). At higher speeds heterogeneity became apparent (Fig. 1). Here three segments of the curve are distinguishable; the three curves are distinguishable on multiple experiments. The first component ends at one fringe, so that the value calculated is of questionable accuracy. However, it is given in Table 2 and used in Fig. 3. If the first slope is considered to be erroneous and the second slope is taken to be the initial slope, a molecular weight of 70 000 is calculated for this slope. That the 70 000 value represents the weight average of several components and the first slope truly represents the initial component becomes clearer when a solution of 1 mg/ml is run under the same conditions (Fig. 2). The initial slope is now much clearer and extends for two fringes although we have obscured.
Table 2. Sedimentation-equilibrium-determined molecular weights of β-galactosidase in 6M-guanidine hydrochloride

Results are given ±s.e.m. for the plot of ln c versus r²; if only one value is listed the plot is linear; if two values are listed, the second molecular weight is that of the second component as determined by subtraction of the extrapolated first slope from the observed second slope.

<table>
<thead>
<tr>
<th>Molecular weight</th>
<th>Heated at 100°C for 20 min</th>
<th>0.25mg/ml</th>
<th>1mg/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed (rev./min)</td>
<td>40000</td>
<td>0</td>
<td>42000±2950</td>
</tr>
<tr>
<td></td>
<td>36000</td>
<td>52000±49000</td>
<td>35000±1400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50000±2000</td>
<td>117000±8850</td>
</tr>
<tr>
<td></td>
<td></td>
<td>167000±27500</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Sedimentation equilibrium of β-galactosidase (1mg/ml) in 6M-guanidine hydrochloride after 24h at 36000rev./min at 20°C.

the third slope. A molecular weight of 35000 is indicated by the first slope, and subtraction of the extrapolated first slope from the second observed slope shows that the component contributing to the second slope has a molecular weight of about 117000 (Table 2; non-heated, 1mg/ml). Some of the values were determined after heating the solution of β-galactosidase in concentrated guanidine hydrochloride and an identically treated portion of the final dialysate (for the blank) at 100°C for 20 min. This was because of a change in viscosity noted after such treatment (see below). No significant changes in sedimentation-equilibrium values were noted and the values are included with the other determinations.

Although only two protein concentrations were used in these experiments, variable amounts of protein were sedimented to the bottom of the ultracentrifuge cell as a function of the speed so that a range of concentrations was achieved in the solutions. Since the volume of solvent remained the same, the concentration of the component of lowest molecular weight is measured by the determination of the quantity in solution from integrating the area under the curve of c versus r² with boundaries of the meniscus and base. The plot of these concentrations against the molecular weight shows little dependence of molecular weight on concentration (Fig. 3), although there is a fair amount of scatter. (In general, sedimentation-equilibrium determinations in concentrated solutions of guanidine hydrochloride are less precise than those in usual solvents because of variations in density with exposure of solvent to the atmosphere, leakage due to ‘creep’ of the solutions between windows and centrepieces, etc.) However, the concentration determined by this integration is for the whole cell rather than the concentration at the point where the values generating the molecular-weight value are obtained. As such, the plot is preferable to one of initial protein concentrations for estimating concentration effects but not as good as point weight-average molecular weights would be. The latter, however, cannot be obtained over any significant concentration range for the multicomponent curves herein obtained.

The results shown in Tables 1 and 2 and Fig. 3 leave little doubt that the smallest component of β-galactosidase in concentrated solutions of guanidine hydrochloride has a molecular weight of about 50000. The molecular weight of the second component is much more variable. Of the four values in Tables 1 and 2, the range is from 116000 to 212000 with a mean of about 160000. This is most compatible with an undissociated-monomer weight of 135000. Since a third component is observed beyond this in some experiments (Fig. 1), it is highly unlikely that there is any component larger than undissociated monomer in these solutions, it is possible that the second component represents a fragment of β-galactosidase remaining when a 50000 dalton piece is dissociated from the 135000 dalton monomer. On this interpretation, the second component should have a molecular weight of about 90000 and the third component would be undissociated monomer of molecular weight 135000. The large discrepancy between a theoretical
Results are from heated or non-heated enzyme, 6M- or 8M-guanidine hydrochloride. The inclusion of several components are being 'averaged' in the peak, amplification of any error in partial specific volume (\(\phi\)) by the high solvent-density term (\(\rho\)), and the amplification of error in the experimental value by the exponential relationship of the number of residues, \(n\), to it. The latter factor also effects the relationship of molecular weight to viscosity in concentrated guanidine hydrochloride solutions.

The intrinsic viscosity measured in 6M-guanidine hydrochloride was high (Fig. 5). On multiple determinations, the intrinsic viscosity of \(\beta\)-galactosidase in 6M-guanidine hydrochloride was 60±3 ml/g. The inclusion of 10mM-EDTA or 10mM-magnesium chloride in the solvent or reduction and carboxymethylation (Steers, Craven & Anfinsen, 1965b) of the protein were without effect on the viscosity. Again, the number, \(n\), of amino acid residues per chain in a random-coil peptide in guanidine hydrochloride may be calculated from Tanford's empirical equation (Tanford et al. 1967):

\[
\frac{s^0}{1 - \phi' \rho} = 0.286 n^{0.473}
\]

where \(\rho\) is the density of the solvent and \(\phi'\) the effective specific volume. From this formula, the number of residues calculated is 790. Taking a residue-average molecular weight of 114.9 (Craven, Steers & Anfinsen, 1965), this corresponds to a molecular weight of 91000. The deficiencies inherent in this approach include the decreased accuracy for the apparent sedimentation coefficient when several components are being 'averaged' in the peak, amplification of any error in partial specific volume (\(\phi\)) by the high solvent-density term (\(\rho\)), and the amplification of error in the experimental value by the exponential relationship of the number of residues, \(n\), to it. The latter factor also effects the relationship of molecular weight to viscosity in concentrated guanidine hydrochloride solutions.

second-component molecular weight of 90 000 and the observed value is only partly explainable by the error amplification that occurs during the subtraction of the extrapolated first slope from the observed second slope.

Only a single peak was visible when \(\beta\)-galactosidase was run in sedimentation-velocity experiments in 6M-guanidine hydrochloride. For these very low sedimentation values, we would not expect the several components to be resolved as individual peaks. The values of \(s^0\), the apparent sedimentation coefficient, obtained at 20°C are plotted against concentration in Fig. 4. An \(s^0\) of 0.95 results when the plot is linearly extrapolated to zero protein concentration. It is possible that we are ignoring an upward deflexion at low concentrations of \(\beta\)-galactosidase but nonlinear plots were not noted by Tanford et al. (1967). (They do emphasize the poor precision of the results for \(s\) values in concentrated solutions of guanidine hydrochloride.) The number, \(n\), of amino acid residues per chain in a random-coil peptide in guanidine hydrochloride may be calculated from a second empirical formula (Tanford et al. 1967):

\[
[\eta] = 0.761 n^{0.66}
\]
where \( [\eta] \) is the intrinsic viscosity. From this formula, the number of residues calculated is 740 corresponding to a molecular weight of 84,000. This is a viscosity-average molecular weight.

When \( \beta \)-galactosidase is heated for 20 min at 100°C in 6M-guanidine hydrochloride–0.1 M-2-mercaptoethanol–0.1 M-tris-HCl buffer, pH 7.5, there is a marked change in the intrinsic viscosity (Fig. 5). A similar volume of solvent is heated in the same way for the blank. This result is reproducible but the degree of decreased viscosity is variable. The fall in viscosity is not found if the concentration of \( \beta \)-galactosidase is greater than 8–10 mg/ml when it is heated. The change in viscosity is not due to loss of the volatile 2-mercaptoethanol. The addition of more mercaptoethanol to the protein solution and the blank after heating did not change the results and reduced carboxymethylated \( \beta \)-galactosidase also showed a similar fall in intrinsic viscosity after being heated in 6M-guanidine hydrochloride. Such a change in viscosity with heating in 6M-guanidine hydrochloride solutions is not found with bovine serum albumin. Determinations in a four-chamber Ubbelohde viscometer at different shear forces showed a trend for the \( \beta \)-galactosidase heated in 6M-guanidine hydrochloride to be less viscous at lower shear forces, but this was not observed with the non-heated solutions. This represented a total variation of 5% and is unexplained, if it is indeed significant. The change in viscosity with heating in the presence of concentrated guanidine hydrochloride seems reversible, as the initial intrinsic viscosity gradually returns while the \( \beta \)-galactosidase remains in the guanidine hydrochloride solution (Fig. 5). In 6M-guanidine hydrochloride the recovery takes about 10 days.

**DISCUSSION**

The measurements of the molecular weight of \( \beta \)-galactosidase in solution in weak acid, weak base and detergent might be artificially low because of charge effects; careful measurements at low concentrations are needed. This source of error seems unlikely to apply to the sedimentation-equilibrium results in concentrated solutions of guanidine hydrochloride and the concentration plot of Fig. 3 argues against such an effect. The sedimentation-velocity method determines the molecular weight of a component if the molecle follows random-coil behaviour in concentrated guanidine hydrochloride. The value of 91,000 determined might represent the average for several components. Thus the values of 80,000–90,000 daltons found as averages by meniscus-depletion sedimentation-equilibrium at lower speeds, by sedimentation velocity and by viscosity determinations might represent a mixture of about 35% 50,000 dalton material, about 35% 90,000 dalton material and about 30% 135,000 dalton material.

The integrated areas under the \( c \) versus \( r^2 \) curves always show that the order of quantities of components is first > second > third. This is not unexpected, as each heavier component in turn would lose a greater proportion by sedimentation to the bottom of the cell. The presence of undissociated monomer is suggested by the change in viscosity on heating in concentrated guanidine hydrochloride. The viscosity determination should result in a viscosity-average molecular weight if there are several components in the solution of 6M-guanidine hydrochloride. The method could give too low a result if the protein(s) is incompletely uncoiled. However, the decrease in viscosity on heating in guanidine hydrochloride suggests that dissociation is incomplete. This average molecular weight of 84,000 is close to that determined at lower speeds during the sedimentation-equilibrium studies.

These results are not in agreement with those of Ullmann et al. (1968a) or the single determination by Goldberg & Edelstein (1969). However, Goldberg & Edelstein (1969) detected approximately equinolar concentrations of 43,000 dalton and 109,000 dalton components in complemented \( \beta \)-galactosidase. Their result for complemented \( \beta \)-galactosidase is very similar to these results with native \( \beta \)-galactosidase. The discrepancy is unlikely to be due to a difference in strains of \( E. \) coli (3300 is a Pasteur Institute strain) but there could be strain differences in quantity of interpeptide-chain bonds. It is unlikely that our preparations of \( \beta \)-galactosidase have been modified by proteolytic enzymes. The molecular weight of these preparations remains 135,000 in 8M-urea. This should be compared with trypsin-treated \( \beta \)-galactosidase which retains full activity and has a sedimentation coefficient of 16S from velocity measurements, but is dissociated into four small components in 8M-urea (Givol, Craven, Steers & Anfinsen, 1966). The point has previously been made that \( \beta \)-galactosidase is dissociable into components smaller than the monomer by treatments known to break ester bonds (Craven, 1967). Such results suggest that the \( \beta \)-galactosidase monomer is actually a protomer of several polypeptide chains held together by covalent bonds other than the peptide ones. If so, these ‘ester’ bonds must be at least partially dissociable by guanidine hydrochloride or be present in variable amounts, e.g. a proportion of \( \beta \)-galactosidase protomer would have fewer interpeptide-chain bonds.

Difference spectra and fluorescence-emission spectra show that the optically active amino acids, tyrosine and tryptophan, are fully exposed to the solvent in 4M-guanidine hydrochloride after 3h (R. P. Erickson, unpublished work). In 6- and 8M-guanidine hydrochloride the spectra are stable
almost immediately; there is a slight increase in quenching of the absorption spectra as the concentration of guanidine increases from 4 to 8M. These results suggest that only a small part of the $\beta$-galactosidase protomer, if any, is resistant to denaturation by concentrated solutions of guanidine.

Although all of the proteins studied for random-coil behaviour in 6M-guanidine hydrochloride by Tanford's group were denatured at 25°C, denaturation was not complete until 12h for bovine serum albumin (Tanford, Kawahara & Lapanne, 1967). Harrington's group found that myosin had to be heated at 55°C for unspecified times to be completely dissociated in 6M-guanidine hydrochloride (Woods, Himmerfarb & Harrington, 1969). Normally iodinated thyroglobulin is not completely devoid of ordered structure by the criteria of viscosity, optical rotatory dispersion and fluorescence polarization when reduced with 15mM-2-mercaptoethanol in 5.8M-guanidine hydrochloride at neutral pH (Edelhoch & Steiner, 1966). Further, synthetic pol-L-leucine shows helical stability in 7.2M-guanidine hydrochloride at temperatures up to 95°C (Auer & Doty, 1966). Avidin retains a normal spectrum in 6.4M-guanidine hydrochloride if complexed with biotin; biotin will reverse the denaturation spectra only at 3M-guanidine hydrochloride (Green, 1963). Thus, $\beta$-galactosidase would not be unique if it partially resists denaturation in concentrated guanidine hydrochloride.

Presumptive genetic evidence for multiple products of the z gene, the structural gene for $\beta$-galactosidase, was provided by Ullmann et al. (1965) and Ullmann, Jacob & Monod (1967) with studies of complementation between different z-gene deletions, a complementation mediated by separable polypeptides. Recently this group studied $\beta$-galactosidase purified from strains bearing operator-distal point mutations, s908, u386, sexduced with F-lac bearing an operator-proximal deletion, B9 (Ullmann et al., 1968b). They found that the $\beta$-galactosidase purified from these complementary diploids differed from the wild-type enzyme in heat stability and for the quantity of complementing peptides extractable. A proportion of the 'complemented' enzyme was always similar to the wild-type. The authors concluded that the complemented $\beta$-galactosidase was produced by association between peptides corresponding to different fragments of the wild-type chain with redundancy of some amino acids, i.e. an unused ‘tail’ must project from the folded monomer (Goldberg, 1969). The wild-type monomer was considered to be a single polypeptide chain. However, the presence of a heterogeneous population of complemented $\beta$-galactosidase molecules, some wild-type and some altered, argues for an alternative explanation. Association of non-mutant polypeptide chains would result in wild-type enzyme when four such protomers formed the tetramer, and a tetramer containing even one protomer created by association of a point-mutant chain and wild-type chains, complemented to native configuration by the other protomers, would show altered characteristics.

A genetic argument against a polycistronic z gene, i.e. in favour of a polypeptide of molecular weight 135,000, was the polarity gradient extending across the z gene for transacetylation production (Newton, 1967). However, Martin (1966) found a polarity gradient across the C gene of the histidine operon in Salmonella, a gene known to determine two polypeptide chains, and finer analysis reveals several peaks in the z gene polarity gradient (Zipser, Zabell, Rothman, Grodzicker & Noritski, 1970). Also, recent evidence for reinitiation sites in the z gene (Grodzicker & Zipser, 1968; Newton, 1969) argues that it is polycistronic. These sites may correspond to initiation sites of chains with the blocked N-terminal pyrrolidinocarboxylic acid (Erickson & Steers, 1969b).

I thank Dr D. Robard for preparing the computer program, Dr Harvey Pollard for help with the Xephytes cells and Dr P. Charlwood for helpful comments on the manuscript. I am also grateful for the encouragement and support of Dr E. Steers, jun., and Dr C. B. Anfinsen.

REFERENCES

MOLECULAR WEIGHT OF β-GALACTOSIDASE IN GUANIDINE