Shape of Ocr, the Gene 0.3 Protein of Bacteriophage T7: Modeling Based on Light Scattering Experiments†

J. J. Blackstock,‡ S. U. Egelhaaf,*,‡ C. Atanasiu,§ D. T. F. Dryden,§ and W. C. K. Poon‡

Department of Physics and Astronomy, The University of Edinburgh, Edinburgh EH9 3JZ, U.K., and Department of Chemistry, The University of Edinburgh, Edinburgh EH9 3JJ, U.K.

Received March 23, 2001

ABSTRACT: Ocr, the first protein expressed by bacteriophage T7, inhibits type I DNA restriction enzymes by preventing them from binding to DNA. This inhibition allows the phage to successfully infect the host. The shape of ocr is modeled on the basis of static and dynamic light scattering measurements. The static light scattering data confirm previous observations that ocr exists in solution as a dimer. The diffusion constant determined by dynamic light scattering indicates a nonspherical shape of the ocr dimer. Hydrodynamic models of ellipsoids are presented, and it is argued that ocr is best described by a prolate ellipsoid with dimensions of 10.4 nm by 2.6 nm. The size and shape predicted by this model are consistent with ocr acting as a mimic of the DNA structure bound by type I restriction enzymes.

The product of gene 0.3 of bacteriophage T7, also termed ocr[1] (overcome classical restriction), is the first protein to be produced by bacteriophage T7 upon infection of an *Escherichia coli* host cell (*1*, *2*). Once produced, ocr inhibits the type I restriction/modification (R/M) enzymes of the host cell, preventing them from destroying the invading DNA (*3*). This allows the remaining bacteriophage DNA to be transcribed.

The sequence of the 116 amino acid wild-type ocr monomer has been known for some time (*4*) and shows that ocr is a very acidic protein. This suggests that ocr acts as a polyanion that binds to the DNA binding site of type I enzymes and inactivates them. Later, it was observed that ocr competes with DNA for the DNA binding site of the type I enzymes *Eco*KI and *Eco*BI, causing these enzymes to dissociate from DNA (*5*).

As ocr operates against a number of type I R/M systems with different DNA sequences as targets (*1*, *5*−*8*), the mechanism by which ocr binds to these enzymes cannot be target sequence specific. In fact, the acidic nature of ocr combined with its competition for the DNA binding sites, suggests that ocr might mimic DNA structural and electrostatic features without being specific to a target sequence. Since ocr does not bind to the numerous other proteins which interact with DNA, it possibly resembles a distorted DNA conformation induced by the binding to type I R/M enzymes. Mimicry of distorted DNA has previously been demonstrated for the uracil glycosylase inhibitor protein (UGI) (*9*−*11*). This protein serves a function similar to that of ocr for the bactriophage PBS2 which infects *Bacillus subtilis*; the DNA of this phage contains uracil instead of thymine. UGI competes for the DNA binding site of *B. subtilis* uracil glycosylase (UNG) to preserve the phage DNA upon insertion into a host cell.

On the basis of sucrose gradient centrifugation and chromatography, Mark and Studier (*6*) demonstrated that ocr exists overwhelmingly (>99%) as a dimer in solution. We determine the shape of the ocr dimer in solution using static (SLS) and dynamic (DLS) light scattering. Light scattering techniques are used extensively to characterize biomolecules (*12*−*15*). Our experiments suggest that dimeric ocr is a prolate ellipsoid of revolution whose overall size and shape closely parallels those of the DNA target sites of type I R/M enzymes. This supports the possibility that ocr mimics the electrostatic and structural features of DNA bound by type I R/M enzymes.

MATERIALS AND METHODS

Protein Production. Ocr was produced by transformation of *E. coli* strain BL21 (DE3) pLysS with the plasmid pAR2993. Cells were grown at 37 °C in 10 × 500 mL of LB broth supplemented with 20 μg/mL ampicillin and 25 μg/mL chloramphenicol. Once the optical density reached a value of 0.6 at 650 nm, ocr expression was induced by the addition of isopropyl β-D-thiogalactoside to 0.4 mM in each flask. The cells were harvested by centrifugation after a further 2 h and resuspended in 20 mM Tris, 0.3 M NH_4Cl, 4% glycerol, 10 μM phenylmethanesulfonyl fluoride, and 10 μM benzamidine, pH 8.0.

The cells were broken by sonication in bursts for a total of 1 min/g of cell paste with the sample kept on ice. The cell debris was removed by centrifugation for 3 h at 30000*g*, and the supernatant was loaded onto a 20 × 1.6 cm diameter

† J.J.B. was partially supported by a Wellcome Trust studentship (Grant VS/00/EDI/004/CH/TH/LC). C.A. was supported by a Darwin Trust studentship and S.U.E. by Unilever Research. D.T.F.D. held a Royal Society University research fellowship and Leverhulme Trust Grant F/158/BC.

* Address correspondence to this author at the Department of Physics and Astronomy, The University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, U.K. Tel: +44-131-650 5291. Fax: +44-131-650 5902. E-mail: S.U.Egelhaaf@ed.ac.uk.

‡ Department of Physics and Astronomy.

§ Department of Chemistry.

[1] Abbreviations: ocr, overcome classical restriction; R/M, restriction/modification; UGI, uracil glycosylase inhibitor protein; UNG, uracil glycosylase; SLS, static light scattering; DLS, dynamic light scattering.

10.1021/bi010587+ CCC: $20.00
Published on Web 07/28/2001

DEAE-Sepharose fast flow ion-exchange column (Pharmacia) equilibrated in 20 mM Tris and 0.3 M NH_4Cl at pH 8.0 (buffer A) at 48 mL/h. After the column was washed in this buffer, a 500 mL gradient from 0.3 to 1 M NH_4Cl in buffer A was run at 24 mL/h. Fractions containing ocr, identified by sodium dodecyl sulfate−polyacrylamide gel electrophoresis, were applied to a HiLoad 16/60 Superdex 200 gel filtration column (Pharmacia) in buffer A. A slight contamination of nucleic acid was removed by a subsequent step using a 15 × 1.6 cm diameter phenyl-Sepharose column (Pharmacia) with a gradient running down from 2.4 to 0 M $(NH_4)_2SO_4$ in a buffer of 20 mM Tris, pH 8.0. The protein, estimated to be 99% pure, was stored at −20 °C in buffer A supplemented with glycerol to 50% v/v.

For the light scattering experiments, the protein was placed in buffer A which was previously filtered through several 0.2 μm filters. Glycerol was removed from the sample by several successive steps of dilution with 3 mL of buffer A, followed by concentration in a centrifugal ultrafiltration device (Vivaspin 10).

Special care was taken during the preparation of samples to avoid contamination with dust as large dust particles would dominate the scattering behavior. Once the glycerol was removed, the solution was passed through a microcentrifuge 0.2 μm filter and spun at ∼12000*g* for 45 min at 4 °C to force any remaining large dust particles to the bottom of the microcentrifuge tube. A microliter syringe was used to extract a sample from the top of the solution in the microcentrifuge tube, with care taken to not disturb the lower part. The sample (100 μL) was placed in a light scattering cell that had been cleaned with filtered buffer. Subsequently, the cell was sealed to prevent contamination.

Before each set of experiments, samples were spun in a centrifuge for 60 min at ∼7000*g*. Different concentrations were obtained by adding filtered buffer. Samples were stored and centrifuged at 12 °C.

UV Spectroscopy. The extinction coefficient for ocr was calculated by comparing the absorbance of folded native protein and protein denatured in GuCl (*16*). The molar extinction coefficient for folded ocr at 280 nm was found to be 31 095 M^{-1} cm^{-1}.

Light Scattering. Static (SLS) and dynamic (DLS) light scattering measurements were performed simultaneously using an ALV goniometer and correlator (ALV, Langen, Germany), equipped with an argon ion laser operating at a wavelength $\lambda_0 = 514.5$ nm and a power of 180 mW. The detection optics were modified to use an optical fiber to collect the scattered light (*17*). Experiments were performed at five different scattering angles within the range $50° \leq \theta \leq 70°$. Five individual 300 s runs per angle were taken. Measurements were performed at temperature $T = 19.0 \pm 0.3$ °C.

The average scattering intensity as a function of angle obtained by SLS was converted into absolute scattering intensities $\Delta R(Q)$ (i.e., the "excess Rayleigh ratio") using toluene as reference and subtracting the background

$$\Delta R(Q) = \frac{\langle I(Q)\rangle - \langle I_B(Q)\rangle}{\langle I_{tol}(Q)\rangle} R_{tol}(Q)\left(\frac{n}{n_{tol}}\right)^2 \qquad (1)$$

where $Q = (4\pi n/\lambda_0)\sin(\theta/2)$ is the magnitude of the scattering vector at scattering angle θ, $\langle I(Q)\rangle$, $\langle I_B(Q)\rangle$, and $\langle I_{tol}(Q)\rangle$ are the average scattering intensities of the solution, the background (buffer), and the reference solvent toluene, $R_{tol}(Q) = 39.6 \times 10^{-4}$ m^{-1} is the Rayleigh ratio of toluene (*18*), and $n = 1.33$ and $n_{tol} = 1.49$ are the indexes of refraction of the solution and toluene, respectively. The apparent molar mass M_{app} is calculated by

$$M_{app} = \lim_{Q\to 0}\langle I_n(Q)\rangle = \lim_{Q\to 0}\frac{\Delta R(Q)}{Kc} \qquad (2)$$

where $\langle I_n(Q)\rangle$ is the normalized average scattering intensity, c is the ocr concentration as determined by UV spectroscopy, $K = 4\pi^2 n^2 (dn/dc)^2/(N_A\lambda_0^4)$, and N_A is Avogadro's number. We took a typical value of the refractive index increment (dn/dc) for proteins, 0.19 mL/g, where variations of 0.01 are within the quoted errors (*19*).

The DLS experiments provide the intensity autocorrelation function $g^{(2)}(\tau)$ defined as (*13*)

$$g^{(2)}(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2} = 1 + f^2(\tau) \qquad (3)$$

where $\langle ...\rangle$ denotes averaging over start time t. At low enough concentrations particles diffuse independently, and the intermediate scattering function $f(\tau)$ decays exponentially (*13*, *20*):

$$f(\tau) = e^{-\tau/\tau_c} \qquad (4)$$

where the correlation time τ_c is related to the diffusion constant D by

$$D = \frac{1}{Q^2\tau_c} \qquad (5)$$

To obtain a diffusion constant D, the correlation functions were individually analyzed using second-order cumulant analysis (*21*) as well as CONTIN (*21*−*24*) and multiexponential fitting (DISCRETE) (*25*, *26*). If both CONTIN and DISCRETE yielded a second decay with an amplitude greater than 5% of the total intensity, this was taken as an indication of the presence of dust, and the measurement was disregarded. The remaining measurements were averaged for each scattering angle and sample. The results from the three methods are in very good agreement; the results presented here were obtained using CONTIN.

RESULTS

Static Light Scattering. The results of the SLS experiments are shown in Figure 1 for a sample with ocr concentration $c = 2.74$ mg/mL. No systematic dependence of the scattering intensity on Q is observed. This is expected for a protein which is small compared to the length scale L probed in a light scattering experiment, $L \sim 2\pi/Q > 300$ nm. Put differently, from the maximum dimensions expected for ocr, we expect Q dependence only for $Q \gtrsim 500$ μm^{-1}, which is well beyond the Q range of light scattering (Figure 1). We can thus obtain the apparent molar mass M_{app} by averaging the results obtained at all Q values. (We disregard the measurement at 15 μm^{-1}, which was contaminated by large dust particles as indicated by CONTIN.) This yields $M_{app} = 28 \pm 4$ kDa. A comparison with the molar mass of an ocr

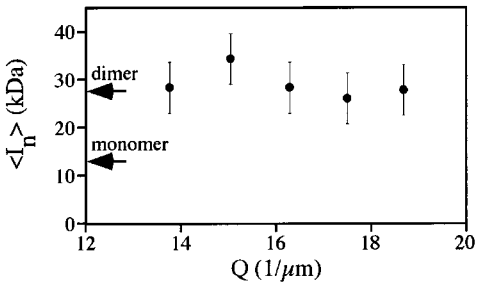

FIGURE 1: Results of the static light scattering experiments for an ocr concentration of 2.74 mg/mL. The normalized average scattering intensity $\langle I_n(Q)\rangle$ (defined in eq 2) is shown as a function of scattering vector Q. The expected values for ocr monomers and dimers are indicated by arrows.

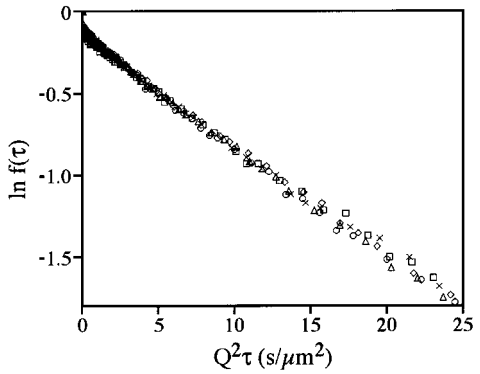

FIGURE 2: Natural logarithm of the intermediate scattering function $f(\tau)$ obtained from dynamic light scattering as a function of $Q^2\tau$ for one run at each of five scattering vectors. Q values for symbols: ◇, 13.8 μm^{-1}; □, 15.0 μm^{-1}; △, 16.3 μm^{-1}; ×, 17.5 μm^{-1}; ○, 18.7 μm^{-1}.

monomer, 13678 Da (*4*), indicates that ocr exists in solution as dimers. This is in agreement with earlier observations using sucrose gradient centrifugation and chromatography (*6*), which showed that more than 99% of ocr in solution is present in dimeric form.

To check that interaction effects were absent, we also performed measurements on a sample with a lower ocr concentration, $c = 1.37$ mg/mL. These returned the same M_{app} of 28 ± 8 kDa, with the larger error bar due to a worse signal-to-noise ratio. M_{app} can thus be taken as the molar mass M of the ocr dimer.

Dynamic Light Scattering. Using DLS, we determined the intermediate scattering function $f(\tau)$ as a function of $Q^2\tau$ for different scattering vectors and an ocr concentration of $c =$ 2.74 mg/mL. Figure 2 demonstrates that the measured $f(\tau)$ at different Q values fall on top of each other. The results are thus consistent with diffusion over the scattering vectors probed in our experiment (eq 5). Furthermore, the linearity of $f(\tau)$ over the time scale of interest indicates that the decay is dominated by one species.[2] The presence of particles of another size would introduce further decay times leading to a nonlinear behavior and, depending on their size, a variation of $f(\tau)$ with scattering vector. The second-order cumulant fit yields polydispersity indices of less than 0.12 and hence

[2] At very small $Q^2\tau$ a faster decay is observed, which is mainly due to the rotational diffusion of ocr. The rotational decay time is, however, too small to be accurately determined with our setup.

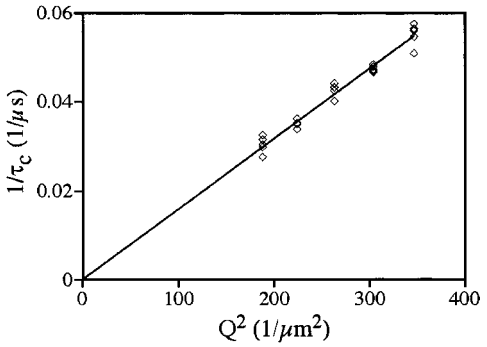

FIGURE 3: Inverse of the correlation time τ_c as a function of the square of the scattering vector, Q^2. The line is a linear fit to the data through the origin.

also supports the fact that only one species is present. From the correlation times τ_c an average diffusion constant can be calculated (eqs 4 and 5) and gives $D = 79.1 \pm 0.7$ $\mu m^2/s$. The values of $1/\tau_c$ obtained from CONTIN are plotted in Figure 3 as a function of Q^2. A linear fit, which passes through the origin, again shows the independence of D on Q (eq 5). The slope gives $D = 79.0$ $\mu m^2/s$, which is consistent with D determined from $f(\tau)$.

The diffusion constant was also determined for a lower ocr concentration, $c = 1.37$ mg/mL, which resulted in $D =$ 79.4 ± 1.6 $\mu m^2/s$. The diffusion constants determined for the two concentrations thus agree within experimental error. Again, this suggests that, as expected from the ocr concentration, ocr charge, and ionic strength of the solution, the values are not affected by interaction effects.

DISCUSSION

The diffusion constant D gives a hydrodynamic radius R_h from the Stokes−Einstein relation for solid spherical particles

$$R_h = \frac{k_B T}{6\pi\eta D} \tag{6}$$

where k_B is Boltzmann's constant and η is the viscosity of the solvent. The value obtained refers to the hydrated protein. To account for hydration, the hydrated partial specific volume of ocr ($\bar{\nu}_h$) is calculated from the specific volume of the protein, the specific volume of water ($\bar{\nu}_w = 1.0$ mL/g), and the level of hydration (δ_h) by

$$\bar{\nu}_h = \bar{\nu}_p + (\delta_h \times \bar{\nu}_w) \tag{7}$$

The value of $\bar{\nu}_p$ was obtained using the method of Cohn and Edsall (*27*) and the consensus partial specific volumes for the constituent amino acids (*28*). This resulted in $\bar{\nu}_p = 0.721$ mL/g. Hydration values for similar size proteins are typically between $0.2 \leq \delta_h \leq 0.4$ (*29*), where δ_h is in units of mass of water per mass of protein.

On the basis of eq 6 and our value of D, we obtain a hydrodynamic radius of $R_h = 2.60 \pm 0.03$ nm. Assuming a spherical shape of the ocr dimer and taking into account its mass and calculated $\bar{\nu}_p$, the dimer radius should be 2.0 nm. If we furthermore assume that the hydration water ($\delta_h =$ 0.3) forms a uniform shell surrounding the protein, this results in a hydrated radius $R_{sphere} = 2.2$ nm. This is significantly smaller than the observed R_h, which strongly suggests a nonspherical shape of the dimer. A spherical

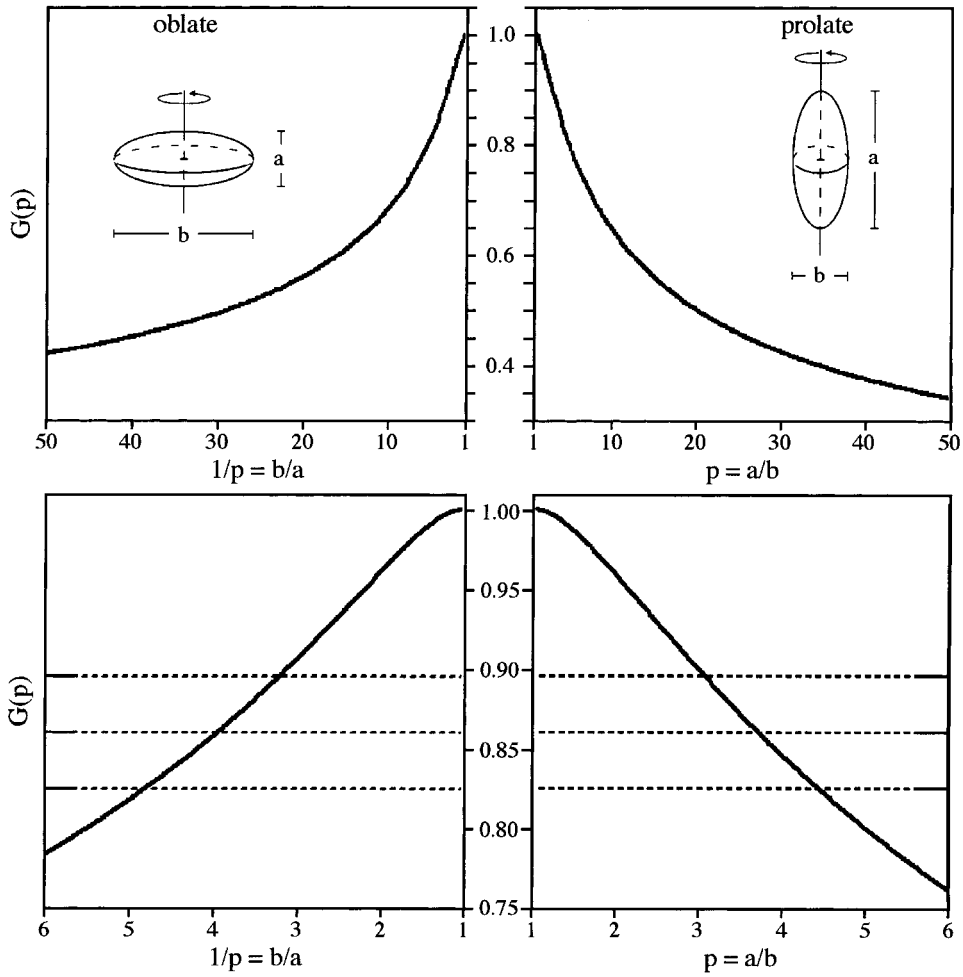

FIGURE 4: The function $G(p)$ as a function of $1/p$ (left) and p (right), which allows both oblate and prolate axial ratios p to be determined using the same definition of p. The bottom figures cover the range relevant to this study. The horizontal lines represent the experimentally determined average value and its uncertainty in the determination of the diffusion constant D and the hydration level δ_h; i.e., for the pairs (from top to bottom) $D = 79.8$ $\mu m^2/s$, $\delta_h = 0.4$; $D = 79.1$ $\mu m^2/s$, $\delta_h = 0.3$; $D = 78.4$ $\mu m^2/s$, $\delta_h = 0.2$.

shape, i.e., $R_{sphere} = R_h$, can only be enforced with unreasonable values for the level of hydration and protein specific volume, such as $\delta_h = 0.9$ and $\bar{\nu}_p = 0.721$ mL/g or $\delta_h = 0.6$ and $\bar{\nu}_p = 1.32$ mL/g. We therefore must consider ocr as nonspherical.

Nonspherical proteins have often been modeled as oblate or prolate ellipsoids of revolution (*29*−*32*). Their axial ratio $p = a/b$ is defined as the ratio of the end-to-end length of the axis of revolution (a) and the equatorial diameter of the ellipsoid (b). Thus for oblate proteins $p < 1$, while $p > 1$ for prolate proteins. Combined knowledge of the rotational and translational diffusion of nonspherical proteins can be used to determine directly their axial ratio (*13*). In the present case, however, the rotation is too fast to be accurately determined (Figure 2).[2] We thus use the center-of-mass translational diffusion constant D alone to infer the axial ratio p. They are related by (*33*, *34*)

$$D = H(M,\bar{\nu}_h)G(p) = \frac{k_B T}{6\pi\eta\left(\frac{3M\bar{\nu}_h}{4\pi N_A}\right)^{1/3}}G(p) \qquad (8)$$

where $H(M,\bar{\nu}_h)$ absorbs all the constants specific to the protein (except its shape) and $G(p)$ is a function of the axial ratio p only:

$$G(p) = \frac{p^{1/3}}{\sqrt{1-p^2}}\tan^{-1}\left(\frac{\sqrt{1-p^2}}{p}\right) \qquad \text{for } p < 1 \text{ (oblate)} \qquad (9A)$$

$$G(p) = \frac{p^{1/3}}{\sqrt{p^2-1}}\ln(p+\sqrt{p^2-1}) \qquad \text{for } p > 1 \text{ (prolate)} \qquad (9B)$$

Note that for a sphere $G(p=1) = 1$.

The function $G(p) = D/H(M,\bar{\nu}_h)$ is plotted in Figure 4. Note that, for greater clarity in the oblate case, the left-hand graphs show $G(p)$ as a function of $1/p$. As $G(p)$ is independent of protein-specific factors, such as mass or partial specific volume, Figure 4 shows the generic dependence of $G(p)$ for *any* protein on its axial ratio p for both prolate and oblate models.[3]

The lower graphs in Figure 4 cover the $G(p)$ values of interest to ocr. The horizontal lines represent the values of $G(p) = D/H(M,\bar{\nu}_h) = 0.860 \pm 0.036$ corresponding to the observed value of D with its associated error bar and $H(M,\bar{\nu}_h)$

[3] Dynamic light scattering measurements yield information similar to that of sedimentation experiments. An expression similar to eq 8 using the sedimentation coefficient (s) instead of D is given in the Appendix and can be used with the same $G(p)$.

Table 1: Dimensions of the Oblate and Prolate Ellipsoidal Models for both the Hydrated and Unhydrated Ocr Dimer Based on the Experimentally Determined Diffusion Constant D[a]

	oblate		prolate	
shape	hydrated	unhydrated	hydrated	unhydrated
axial ratio	0.25 ± 0.06	0.21 ± 0.05	3.7 ± 0.7	4.2 ± 1.0
long axis	7.1 ± 1 nm	6.7 ± 1 nm	10.8 ± 1 nm	10.4 ± 1 nm
short axis	1.8 ± 0.5 nm	1.4 ± 0.5 nm	3.0 ± 0.5 nm	2.6 ± 0.5 nm

[a] The values for the long and short axes are the end-to-end lengths.

for dimeric ocr with a hydration value $0.2 \leq \delta_h \leq 0.4$. The deviation of the experimentally determined $G(p)$ from unity, the value expected for a sphere, confirms that a spherical shape is inappropriate to describe the observed diffusion behavior.

The axial ratios for oblate and prolate ellipsoids obtained from numerical evaluations of eqs 8 and 9 are given in Table 1 and correspond to the hydrated ocr dimer. On the basis of these values we calculated the axial ratio for the ocr dimer without hydration layer. This was done assuming that the hydration shell has uniform thickness, which is given by the level of hydration. In addition, the end-to-end length of the long and short axes are given, which are calculated from the axial ratios and the volume.

Assuming an oblate shape of the ocr dimer, its long and short axis are 6.7 and 1.4 nm, respectively. This implies that the maximum thickness (the middle of the oblate) is only 1.4 nm. Typical elements of secondary structures, such as α helices and β strands, could thus only be accommodated at the very center of the ellipsoid. Moreover, a thin oblate ellipsoid has relatively low flexial rigidity, which makes it difficult to maintain a stable tertiary structure and thus an active conformation. These considerations render an oblate shape highly unlikely.

For a prolate shape, the long and short axes are found to be 10.4 and 2.6 nm, respectively. This size is strikingly similar to the size of DNA to which type I R/M enzymes bind. For example, *Eco*KI binds over a DNA sequence of 26 base pairs, which corresponds to a length of 9 nm and a double helix DNA diameter of 2 nm (*35*, *36*). The prolate model is therefore consistent with the observations that ocr competes for the DNA binding site of type I R/M systems (*5*). It is further supported by the fact that, like DNA, ocr is strongly acidic with 28 negative charges per monomer at neutral pH. It has been found that the type I R/M enzymes have approximately 2-fold rotational symmetry (*37*). The sequence recognition subunit of type I R/M enzymes contains two target recognition domains, each of which recognizes one part of the bipartite target sequence on the DNA. The 2-fold symmetry of dimeric ocr could be related to the 2-fold symmetry of the restriction enzyme and the bipartite DNA target sequence. With a prolate model of the ocr dimer, where the monomers face each other on the middle plane of the long axis, each monomer would be in the correct position to interact with one of the type I target recognition domains, the same as the two separately recognized parts of the DNA target sequence.

CONCLUSIONS

The results of the static (SLS) and dynamic (DLS) light scattering experiments support the existence of ocr as a dimer in solution as observed earlier using sucrose gradient centrifugation and chromatography (*6*). Assuming a spherical shape for the ocr dimer necessitates unreasonable values for the level of hydration and protein specific volume to account for the measured diffusion constant D. We therefore analyzed ellipsoidal models. An oblate shape is unlikely as the protein had to be unreasonably thin. Our data are, however, consistent with a prolate ellipsoid with long and short axes of 10.4 and 2.6 nm, respectively. This is in line with the trend that small proteins with a molar mass of less than about 100 kDa are prolate, while it is only the large proteins which tend to have an oblate shape. This has been noted by Durchschlag and Zipper on the basis of a wide variety of proteins studied by X-ray scattering and hydrodynamic methods (*30*).

The proposed geometry of the ocr dimer based on our data is very similar to a typical DNA to which type I R/M enzymes bind. Since its method of inducing binding of R/M systems must be nonspecific and ocr is a very acidic protein, it is conceivable that ocr dimers act as a mimic of the structural and electrostatic properties of the DNA structure recognized by type I R/M enzymes.

ACKNOWLEDGMENT

We thank A. Rosenberg and W. Studier (Brookhaven National Laboratory) for providing plasmid pAR2993. We also thank O. Byron (Glasgow University) for calculating the partial specific volume of ocr.

APPENDIX

Our calculation of $G(p)$ is based on equations presented by Durchschlag et al. (*29*, *30*) and Harding and Cölfen (31). We defined $H(M,\bar{\nu}_h)$ and $G(p)$ such that $G(p)$ is a generic function which is independent of the protein mass and specific volume and can be applied to both DLS and sedimentation. For DLS experiments, $H(M,\bar{\nu}_h)$ was already defined in the text. The appropriate constant for sedimentation measurements, $H_s(M,\bar{\nu}_h)$, is given by

$$s = H_s(M,\bar{\nu}_h)G(p) = \frac{\frac{M}{N_A}\left(1 - \frac{\bar{\nu}_p}{\bar{\nu}_w}\right)}{6\pi\eta\left(\frac{3M\bar{\nu}_h}{4\pi N_A}\right)^{1/3}}G(p) \qquad \text{(A1)}$$

where s is the sedimentation coefficient.

When the mass M of the protein is known and estimates for the level of hydration δ_h and protein partial specific volume $\bar{\nu}_p$ are available (and thus $\bar{\nu}_h$ can be calculated using eq 7), modeling DLS or sedimentation data using $G(p)$ can provide a valuable estimate of the protein shape. If adequate amounts of protein are available, measurements of further parameters, such as intrinsic viscosity or the second virial coefficient, can be obtained and modeling performed independent of δ_h. Such comprehensive modeling is discussed in detail by Harding and Cölfen (*31*) or Kumosinski and Pessen (*32*).

REFERENCES

1. Studier, F. W. (1975) *J. Mol. Biol. 94*, 283−295.

2. Kruger, D. H., Hansen, S., and Reuter, M. (1983) *J. Virol. 45*, 1147−1149.
3. Murray, N. E. (2000) *Microbiol. Mol. Biol. 64*, 412−434.
4. Dunn, J. J., Elzinga, M., Mark, K., and Studier, F. W. (1981) *J. Biol. Chem. 256*, 2579−2585.
5. Bandyopadhyay, P. K., Studier, F. W., Hamilton, D. L., and Yuan, R. (1985) *J. Mol. Biol. 182*, 567−578.
6. Mark, K., and Studier, F. W. (1981) *J. Biol. Chem. 256*, 2573−2578.
7. Spoerel, N., Herrlich, P., and Bickle, T. A. (1979) *Nature 278*, 30−34.
8. Kruger, D. H., and Bickle, T. A. (1983) *Microbiol. Rev. 47*, 345−360.
9. Mol, C. D., Arvai, A. A., Sanderson, R. J., Slupphaug, G., Kavli, B., Krokan, H. E., Mosbaugh, D. W., and Tainer, J. A. (1995) *Cell 82*, 701−708.
10. Savva, R., and Pearl, L. H. (1995) *Nat. Struct. Biol. 2*, 752−757.
11. Putnam, C. D., Shroyer, M. J. N., Lundquist, A. J., Mol, A. J., Arvai, C. D., Mosbaugh, D. W., and Tainer, J. A. (1999) *J. Mol. Biol. 287*, 331−346.
12. Harding, S. E. (1991) *Biochem. Soc. Trans. 19* (special edition on Laser Light Scattering in Biochemistry).
13. Berne, B., and Pecora, R. (1976) *Dynamic Light Scattering with Applications to Chemistry, Biology and Physics*, John Wiley & Sons Inc., London.
14. Cummins, H. Z. (1974) in *Photon Correlation and Light Beating Spectroscopy*, pp 285−330, Plenum Press, New York.
15. Dubin, S. B., Clark, N. A., and Benedek, G. B. (1971) *J. Chem. Phys. 54*, 5158−5164.
16. Pace, C. N., and Scholtz, J. M. (1997) in *Protein structure: a practical approach*, pp 229−348, IRL Press, Oxford.
17. Gisler, T., Rüger, H., Egelhaaf, S. U., Tschumi, J., Schurtenberger, P., and Rička, J. (1995) *Appl. Opt. 34*, 3546−3553.
18. Otto Glatter, personal communication (Graz, Austria).
19. Kuehner, D. E., Heyer, C., Ramsch, C., Fornefeld, U. M., Blanch, H. W., and Prausnitz, J. M. (1997) *Biophys. J. 73*, 3211−3224.
20. Pusey, P. N. (1974) in *Photon Correlation and Light Beating Spectroscopy*, pp 387−428, Plenum Press, New York.
21. Koppel, D. (1972) *J. Chem. Phys. 57*, 4814−4820.
22. Provencher, S. W. (1982) *Comput. Phys. Commun. 27*, 213−227.
23. Provencher, S. W. (1982) *Comput. Phys. Commun. 27*, 229−242.
24. Provencher, S. W., Hendrix, J., and De, L. (1978) *J. Chem. Phys. 69*, 4273−4276.
25. Provencher, S. W. (1976) *J. Chem. Phys. 64*, 2772.
26. Provencher, S. W. (1976) *Biophys. J. 16*, 27.
27. Cohn, E. J., and Edsall, J. T. (1943) in *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, Reinhold Publishing Corp., New York.
28. Perkins, S. J. (1986) *Eur. J. Biochem. 57*, 169−180.
29. Durchschlag, H., and Zipper, P. (1999) *Prog. Colloid Polym. Sci. 113*, 87−105.
30. Durchschlag, H., Zipper, P., Purr, G., and Jaenicke, R. (1996) *Colloid Polym. Sci. 274*, 117−137.
31. Harding, S. E., and Cölfen, H. (1995) *Anal. Biochem. 228*, 131−142.
32. Kumosinski, T. H., and Pessen, H. (1985) *Methods Enzymol. 117*, 154−182.
33. Perrin, F. (1936) *J. Phys. Radium, Ser. VII 7*, 1−11.
34. Perrin, F. (1934) *J. Phys. Radium, Ser. VII 5*, 497−511.
35. Powell, L. M., Connolly, B. A., and Dryden, D. F. T. (1998) *J. Mol. Biol. 283*, 947−961.
36. Taylor, I., Watts, D., and Kneale, G. (1993) *Nucleic Acids Res. 21*, 4929−4935.
37. Kneale, G. G. (1994) *J. Mol. Biol. 243*, 1−5.

BI010587+