Nucleic Acids Research, 2001, Vol. 29, No. 14 **3059–3068**

Characterisation of the structure of ocr, the gene 0.3 protein of bacteriophage T7

C. Atanasiu, O. Byron[1], H. McMiken[2], S. S. Sturrock and D. T. F. Dryden[2,*]

Institute of Cell and Molecular Biology, The King's Buildings, University of Edinburgh, Edinburgh EH9 3JR, UK, [1]Institute of Biomedical and Life Sciences, Division of Infection and Immunity, Joseph Black Building, University of Glasgow, Glasgow G12 8QQ, UK and [2]Department of Chemistry, University of Edinburgh, The King's Builings, Edinburgh EH9 3JJ, UK

Received April 6, 2001; Revised and Accepted May 30, 2001

ABSTRACT

The product of gene 0.3 of bacteriophage T7, ocr, is a potent inhibitor of type I DNA restriction and modification enzymes. We have used biophysical methods to examine the mass, stability, shape and surface charge distribution of ocr. Ocr is a dimeric protein with hydrodynamic behaviour equivalent to a prolate ellipsoid of axial ratio 4.3 ± 0.7:1 and mass of 27 kDa. The protein is resistant to denaturation but removal of the C-terminal region reduces stability substantially. Six amino acids, N4, D25, N43, D62, S68 and W94, are all located on the surface of the protein and N4 and S68 are also located at the interface between the two 116 amino acid monomers. Negatively charged amino acid side chains surround W94 but these side chains are not part of the highly acidic C-terminus after W94. Ocr is able to displace a short DNA duplex from the binding site of a type I enzyme with a dissociation constant of the order of 100 pM or better. These results suggest that ocr is of a suitable size and shape to effectively block the DNA binding site of a type I enzyme and has a large negatively charged patch on its surface. This charge distribution may be complementary to the charge distribution within the DNA binding site of type I DNA restriction and modification enzymes.

INTRODUCTION

Upon injection of bacteriophage DNA from a viral particle into a bacterial cell, the viral DNA is often vulnerable to attack by bacterial restriction/modification (R/M) systems. In general, these R/M systems comprise a restriction endonuclease activity which cleaves viral DNA containing unmodified target sequences and a modification methyltransferase activity to modify target sequences on the bacterial DNA (reviewed in 1–3). Bacteriophage have developed numerous mechanisms to defend themselves from R/M systems including the synthesis of antirestriction proteins which block the action of the restriction endonuclease (2,4,5). One such protein is the product of gene 0.3 of bacteriophage T7.

This gene, often called ocr (overcome classical restriction), is the first to be transcribed and translated upon injection of T7 phage DNA into the *Escherichia coli* host (6,7). Once translated, the ocr protein targets the cells type I R/M systems and inactivates them by binding very tightly to the type I R/M enzyme (8,9). The blockage of the DNA binding site by the ocr protein prevents the type I R/M enzyme from binding to unmethylated phage DNA and allows the phage to propagate. Ocr operates against type I R/M enzymes that possess different DNA target specificities suggesting that ocr can operate in a non-sequence-specific manner against these R/M systems (7–11). Type I DNA R/M enzymes are large multifunctional molecular machines containing DNA methylation, DNA cleavage and ATPase-driven DNA translocation activities in one enzyme (reviewed in 3).

Ocr has been found to be a dimer comprising two 13.5 kDa subunits (8). Overall the protein is highly acidic containing 34 aspartic or glutamic acid residues and only six arginine or lysine residues (12). Many of the acidic residues, 14 out of 34, are located in the non-essential C-terminus (after amino acid W94) while none of the basic residues are in this region. This preponderance of negative charge throughout the protein led Dunn *et al.* (12) to suggest that ocr could act as a polyanion to compete with DNA for binding to the type I R/M enzymes. Of the 116 amino acids in each subunit, the first 94, including the single tryptophan at amino acid 94, are essential for activity. However, construction of a truncated protein containing the first 93 amino acids did not reveal whether the removal of W94 and subsequent amino acids, which destabilised the protein, led to the observed temperature sensitive activity or whether W94 actually played a functional role as well as or instead of a structural role.

We present an analysis of the stabilising effect of the C-terminus on ocr structure and the role of W94 in inhibition of type I R/M systems. In addition, the shape of ocr was determined using hydrodynamic methods and we have located amino acids potentially located at the interface of the two subunits. Fluorescence quenching shows that ocr has a large negatively charged patch, located around W94, which may be complementary to the charge distribution within the DNA binding site

*To whom correspondence should be addressed. Tel: +44 131 650 4735; Fax: +44 131 650 6453; Email: david.dryden@ed.ac.uk
Present address:
S. S. Sturrock, Edinburgh Biocomputing Systems Ltd, The Logan Building, Roslin BioCentre, Roslin, Midlothian EH25 9PS, UK

Table 1. Oligonucleotides used for mutagenesis

Mutation	Primer length	Strand	Sequence
N4C	44	Forward	5′-CAAGATGGCTATGTCTTGCATGACTTACAACAACGTTTTCGACC-3′
		Reverse	3′-GTTCTACCGATACAGAACGTACTGAATGTTGTTGCAAAAGCTGG-5′
D25C	41	Forward	5′-GCTGAAAGAAAACATCCGTTATTGTGACATCCGTGACACTG-3′
		Reverse	3′-CGACTTTCTTTTGTAGGCAATAACACTGTAGGCACTGTGAC-5′
N43C	44	Forward	5′-GCTATTCACATGGCTGCCGATTGTGCAGTTCCGCACTACTACGC-3′
		Reverse	3′-CGATAAGTGTACCGACGGCTAACACGTCAAGGCGTGATGATGCG-5′
D62C	39	Forward	5′-GGCAAGTGAGGGCATTTGCCTTGAGTTCGAAGACTCTGG-3′
		Reverse	3′-CCGTTCACTCCCGTAAACGGAACTCAAGCTTCTGAGACC-5′
S68C	39	Forward	5′-GACCTTGAGTTCGAAGACTGTGGTCTGATGCCTGACACC-3′
		Reverse	3′-CTGGAACTCAAGCTTCTGACACCAGACTACGGACTGTGG-5′
W94C	37	Forward	5′-GCAATTAACGATTGACCTCTGTGAAGACGCAGAAGAC-3′
		Reverse	3′-CGTTAATTGCTAACTGGAGACACTTCTGCGTCTTCTG-5′

of type I DNA R/M enzymes. Analysis of these data suggests that ocr is a rather elongated, non-spherical protein covered in negative charges. We also demonstrate that ocr can displace DNA from the methyltransferase core of a type I R/M enzyme and acts as a tight binding competitive inhibitor. A prolate ellipsoid modelled from our data is a suitable shape to effectively block the DNA binding site of a type I R/M enzyme, which typically covers ~30 bp of DNA (13–17).

MATERIALS AND METHODS

Two truncated forms of ocr were compared to the normal full-length protein; ocr109 containing the first 109 amino acids and ocr99 containing only the first 99 amino acids. For the purposes of this paper, the full-length ocr protein will be termed ocr or ocrfl. Variants of ocrfl containing amino acid substitutions will be referred to using the amino acid change, e.g. N4C for ocrfl containing cysteine at residue 4 instead of asparagine. The methyltransferase core of the type I R/M enzyme, *Eco*KI, will be referred to as M.*Eco*KI.

Ocr overexpression plasmids

Plasmids pAR2993, pAR3786 and pAR3790 were kind gifts from Dr A. Rosenberg and Prof. W. Studier (Brookhaven National Laboratory). Ocrfl is expressed from pAR2993 while pAR3786 and pAR3790 express ocr109 and ocr99 respectively. pAR2993 was constructed by ligating the *Alu*I fragment of the T7 genome (base pairs 837–1379) into the *Bam*HI site of pET1 using *Bam*HI linkers of sequence CCGGATCCGG. pAR3786 was constructed by cutting pAR2993 with *Rsa*I at T7 base pair 1258 and pAR3790 by cutting pAR2993 with *Hga*I at T7 base pair 1222. Appropriate palindromic linkers were added to supply a TAA termination codon and a *Bam*HI site for religation (A.Rosenberg, personal communication). The sequences of the regions containing the ocr gene or the truncations were determined on both strands by manual sequencing using the Sequenase 2.0 kit (Amersham).

Secondary structure prediction

In selecting amino acids within the ocr protein for alteration by site-directed mutagenesis, we were directed by predictions of protein secondary structure and solvent exposure produced by the PHD sequence alignment and structure prediction computer program (18–20). The reliability of the prediction is less than expected for this program as the amino acid sequence of ocr is unique in current public sequence databases. However, the protein was predicted to contain α-helices from amino acids 8–20, 31–43, 47–59 and 74–95, with β-strands between amino acids 24–26 and 63–67 (data not shown).

Site-directed mutagenesis

Site-specific mutations in the coding sequence of the ocr gene were created by PCR mutagenesis using the 'QuickChange' site-directed mutagenesis kit (Stratagene) following the manufacturer's instructions. The oligonucleotide primers used and the amino acid change expected are given in Table 1. The DNA sequences of the altered plasmids were checked using the Sequenase kit as for plasmid pAR2993.

In vivo restriction assay

Escherichia coli BL21 (DE3) pLysS cells were transformed with the plasmids expressing ocr or its derivatives. A conjugation between *E.coli* NK308 as a donor strain and the transformed *E.coli* BL21 cells as a recipient strain was then carried out. *Escherichia coli* NK308 contains the F'101 plasmid which is transferred to the *E.coli* BL21 cells during conjugation (21). F'101 expresses the *Eco*KI type I R/M system. The recipient strain, after transformation and conjugation, was plated on BBL agar plates and bacteriophage λ v.o (i.e. DNA susceptible to *Eco*KI cleavage) spotted on the plates at different dilutions. As a control, λ v.k (i.e. resistant to cleavage) was also tested. After incubation overnight at 37°C, the number of plaques was counted and the efficiency of plating calculated as a ratio between the titre of phage appearing on the restriction deficient and the restriction proficient strains.

Protein purification

The ocr proteins were produced by transformation of *E.coli* strain BL21 (DE3) pLysS with the plasmid pAR2993 or its derivatives. Purification using anion exchange chromatography, gel filtration chromatography and hydrophobic interaction chromatography yielded protein judged to be >99%

pure (22). The methyltransferase core of the type I R/M enzyme, *Eco*KI, was prepared as previously described (23).

Calculation of extinction coefficient

The extinction coefficients for pure ocr and its derivatives were calculated by comparing the absorbance of folded native protein and protein denatured in guanidium hyrochloride (GuCl) (24,25). A molar extinction coefficient of 32 095 M^{-1} cm^{-1} at 280 nm for folded ocr and 31 860 M^{-1} cm^{-1} for unfolded ocr were determined. Extinction coefficients for the mutant forms of the proteins were calculated from that of folded wild-type ocr accounting for changes in mass and the loss of tryptophan in the W94C variant.

Analytical ultracentrifugation

Sedimentation equilibrium and velocity analysis was performed on a Beckman Optima XL-A analytical ultracentrifuge equipped with scanning absorbance optics at 280 nm. Samples were suspended in 20 mM Tris–HCl, 0.3 M NH_4Cl, pH 8.0 [calculated density (ρ) 1.0040 g/ml at 20°C and 1.0056 g/ml at 4°C using SEDNTERP (26)]. Partial specific volumes (ν) for ocrfl and ocr99 were calculated to be 0.721 and 0.723 ml/g respectively, using the method of Cohn and Edsall (27) and the consensus partial specific volumes for amino acids (28). These values were assumed to remain constant upon oligomerisation.

Sedimentation equilibrium was attained for 1.5 mm solution columns and ascertained by ensuring overlay of scans acquired 3 h apart. Optical baselines, depleted of macromolecular solute, were obtained by overspeeding to 47 000 r.p.m. Nine loading concentrations of ocrfl from 2.0 to 13.1 μM and eight of ocr99 from 2.6 to 15.4 μM were analysed. Equilibrium distributions for ocr99 were obtained at rotor speeds of 25 000 and 32 000 r.p.m. at 20°C. Ocrfl was equilibrated at 4°C at 25 000 r.p.m. Equilibrium data were analysed by fitting with the equation describing the solute distribution in a molecular system comprising self-interacting species at thermodynamic equilibrium (29) using the program PCNONLIN (30), which permits the global fitting of up to 15 data sets with one self-association model.

Sedimentation velocity data were acquired at 40 000 r.p.m. by scanning the 10 mm solution column every 10 min. Data were analysed with the program SVEDBERG (26) and fitted with a modified form of the Fujita equation (31) for the concentration throughout the solution column. This analysis is valid for interacting macromolecular systems only if the kinetics of inter-conversion among species is slow compared to the experimental timescale. The sedimentation equilibrium experiments showed that ocr exists as a dimer, with no trace of monomer present so this analysis is appropriate.

Hydrodynamic modelling of spheres and ellipsoids of revolution

Standard theory describing the relation of hydrodynamic parameters to ellipsoids of revolution was applied to the measurements of sedimentation velocity (32–34). The relationship between sedimentation coefficient (s) and experimental frictional coefficient (f) is $f = M(1 - \nu\rho)/(N_A s)$ where N_A is Avogadro's number and M is the mass of the macromolecule. The frictional coefficient of an anhydrous sphere of radius r_o is $f_o = 6\pi\eta r_o$ where η is the solvent viscosity in Poise. The frictional ratio (f/f_o) represents the deviation of the macromolecule from an anhydrous sphere with $f/f_o = 1.0$. A typical protein hydration (δ) lies between 0.2 and 0.4 g water/g protein and the frictional ratio can be expressed as the product of the frictional ratio due to hydration $(f/f_o)_{hyd}$ and due to molecular asymmetry (shape) $(f/f_o)_{sh}$ as follows: $f/f_o = (f/f_o)_{sh} \cdot (f/f_o)_{hyd} = (f/f_o)_{sh} \cdot (1+\delta/\nu\rho)^{1/3}$. This allows the calculation of $(f/f_o)_{sh}$ at a range of assumed values of hydration. The ocr molecule was modelled as a prolate ellipsoid of revolution. Determining an axial ratio (p) from $(f/f_o)_{sh}$ is done using the following equation: $(f/f_o)_{sh} = [(p^2 - 1)^{1/2}]/\{p^{1/3} \cdot \ln[p + (p^2 - 1)^{1/2}]\}$ for a prolate ellipsoid of revolution.

Fluorescence spectroscopy

Steady state fluorescence spectroscopy measurements were performed on Perkin-Elmer LS50B or Edinburgh Instruments FS900 fluorimeters.

Chemical denaturation using GuCl was performed using the same procedures described for the *Eco*KI methyltransferase (23) and the data fitted to a two-state model after normalisation (25). The buffer was 20 mM Tris–HCl, 0.3 M NH_4Cl pH 8.0 at 20°C and the protein concentration was 0.7 μM. Denaturation produced a 20% decrease in fluorescence emission intensity and a concomitant decrease in fluorescence anisotropy.

Tryptophan fluorescence was quenched by CsCl or KI in 400 μl of 0.7 μM ocr protein in 20 mM Tris–HCl, pH 8.0 at 20°C. For some experiments, the ionic strength of the solutions was kept constant at 0.4 M by adding NH_4Cl to each solution, however, this had very little effect on the quenching so most experiments were performed by direct addition of a stock solution of the quencher to the protein in the Tris buffer. The addition of NH_4Cl had no effect upon the fluorescence emission. The solutions containing KI were supplemented with 2 mM $Na_2S_2O_3$ to prevent the formation of the I_3^- ion. Each sample was excited at 295 nm and emission measured at 350 nm with 10 and 2 mm pathlengths respectively. The Stern–Volmer equation was used to fit the fluorescence intensity (35).

Fluorescence lifetime measurements for the decay of the emission from W94 were performed using an Edinburgh Instruments 199T spectrofluorimeter. Excitation was at 295 nm, bandwidth 20 nm, with a hydrogen flashlamp operating at a repetition rate of 65 kHz in 0.5 atmospheres of H_2. Emission was collected through a 335 nm cutoff filter. The sample was in a 10 mm × 4 mm cuvette in 20 mM Tris–HCl, 0.3 M NH_4Cl pH 8.0 buffer at ~12 μM concentration. Excitation was through the 4 mm pathlength side of the cuvette. The lamp profile was collected at 295 nm using a dilute solution of Ludox colloidal silica suspension. The half-width of the lamp pulse was ~1.8 ns. Data were fitted using multiexponential fitting procedures incorporating lamp profile deconvolution (36).

The binding of DNA duplexes to the M.*Eco*KI methytransferase and their subsequent displacement by the addition of ocr was measured using 21-bp oligonucleotide duplexes containing the *Eco*KI target site sequence end-labelled with hexachlorofluorescein (16). The fluorescence anisotropy change caused by the addition of M.*Eco*KI was measured as described previously (16). DNA (1 μM) was titrated with M.*Eco*KI to a final concentration of 3.8 μM sufficient to bind all the DNA. Ocr was then added to a final concentration of 22 μM to displace DNA from M.*Eco*KI. Further experiments to examine the strength of binding of ocr to M.*Eco*KI used

1 nM DNA in 20 mM Tris–HCl, 6 mM $MgCl_2$, 7 mM β-mercaptoethanol, 100 µM *S*-adenosyl-methionine, pH 8.0 titrated with increasing concentrations of M.*Eco*KI as described (16). Ocr was included in the DNA solution at concentrations ranging from 2 to 5000 nM. The titration curves were fitted with DYNAFIT (37) using non-linear regression to the equilibrium model:

M.*Eco*KI + DNA ´ M.*Eco*KI–DNA Dissociation constant, $K_{d,DNA}$

M.*Eco*KI + ocr ´ M.*Eco*KI–ocr Dissociation constant, $K_{d,ocr}$

$K_{d,DNA}$ was held constant to that found in the absence of ocr.

Chemical modification of cysteine in mutant proteins

N-(1-pyrene)maleimide (Molecular Probes) was dissolved in dimethyl sulphoxide. An aliquot of 100 µl of 50 µM protein was incubated overnight at 4°C in the dark with a 50-fold molar excess of the label in 100 mM sodium phosphate, pH 7. The ocr protein containing the N4C mutation could only be labelled in the presence of 7 mM β-mercaptoethanol. The reaction was stopped by the addition of β-mercaptoethanol to 100 mM and unreacted label removed by gel filtration using a Sephadex G50 column (Pharmacia PD10) equilibrated in the same buffer. The degree of labelling was determined using an extinction coefficient of 40 000 M^{-1} cm^{-1} at 342 nm and 25 000 M^{-1} cm^{-1} at 280 nm (38). Fluorescence was excited at 342 nm with excitation and emission bandwidths of 5 and 10 nm and pathlength of 3 mm. The buffer was 0.1 M sodium phosphate at pH 7.

Analytical gel filtration chromatography

Analytical HPLC gel filtration chromatography was performed as described (39) using 50 µl of 20 µM ocr in 20 mM Tris–HCl, 20 mM 2-[*N*-morpholino]ethanesulphonic acid, 0.2 M NH_4Cl, 10 mM $MgCl_2$, 0.1 mM EDTA, 5 M GuCl, pH 6.5. β-Mercaptoethanol, if present, was at a concentration of 7 mM. Eluting proteins were detected by UV absorption at 280 nm and fluorescence emission at 350 nm with excitation at 280 nm.

Scanning calorimetry

Proteins were desalted into 20 mM Tris–HCl, 20 mM NH_4Cl, 6 mM $MgCl_2$, pH 8.0 and their concentration determined by UV absorption. The samples were degassed before loading 0.75 ml into a Microcal VP-DSC microcalorimeter. The temperature was raised from 20 to 100°C at a scan rate of 1°C/min. Subsequent cooling and reheating of the samples gave the same denaturation profiles indicating a reversible folding transition. The data were analysed using the Origin 5.0 program.

Mass spectroscopy

An aliquot of 10 µg of protein was desalted on an Applied Biosystems 130A Microbore Separation System with an aquapore RP-300 30 mm × 2.1 mm diameter microbore column eluted at 0.2 ml/min with a 7–70% v/v gradient of acetonitrile in 0.1% v/v trifluoroacetic acid. An aliquot of 50 µl of this desalted protein was mixed with 200 µl of 60% methanol in 1% v/v formic acid and infused by syringe at a flow rate of 3 µl/min into a Thermoquest LCQ mass spectrometer. The spectra collected in full scan mode were deconvoluted using BioExplore software. The mutant proteins all gave masses greater than expected by 76 Da corresponding to the attachment of β-mercaptoethanol to the cysteine residue.

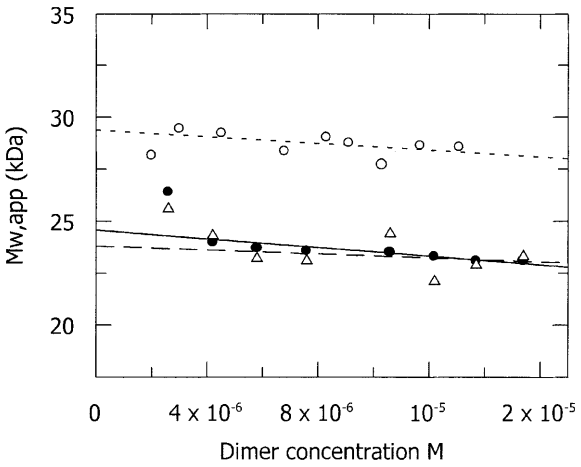

Figure 1. The apparent whole-cell weight average molecular mass of ocrfl (open circles) acquired at 25 000 r.p.m. and ocr99 at 25 000 r.p.m. (solid circles) and 32 000 r.p.m (open triangles), respectively, as a function of loading concentration in sedimentation equilibrium studies. Least-squares fits to the data yield values for the whole-cell weight average molecular mass at infinite dilution of $M^0_{w,app}$ = 29.4 ± 0.5 kDa for ocrfl at 25 000 r.p.m. (short dashes) and $M^0_{w,app}$ = 24.6 ± 0.4 kDa and 23.8 ± 0.7 kDa for ocr99 at 25 000 r.p.m. (solid line) and 32 000 r.p.m. (long dashes), respectively.

This modification can apparently survive the electrospray mass spectrometry (esms) sample preparation method. The β-mercaptoethanol could be removed by treatment with dithiothreitol, which was subsequently removed during the normal sample preparation procedure for the esms. The correct mass was then found for the N43C mutant and we assumed that a similar procedure would give the correct mass for the other proteins.

RESULTS

Hydrodynamic confirmation of protein dimerisation and determination of shape

From the amino acid sequence of full-length ocr, the monomer mass is known to be 13 678 Da. Thus, a dimer would have a mass of 27 356 Da. The extrapolation of sedimentation equilibrium data in Figure 1 yields a mass close to this value, 29.4 ± 0.5 kDa, indicating that ocr is a dimer in solution at the concentrations studied, in agreement with earlier experiments (8,40). The fits to the raw data obtained were good with little or no concentration dependence in the variance.

The monomer mass of ocr99 is expected to be 11 512 Da giving a dimer of 23 024 Da. The extrapolations in Figure 1 yield masses of 24.6 ± 0.4 kDa and 23.8 ± 0.7 kDa for the 25 000 and 32 000 r.p.m. data, respectively. These data indicate that ocr99 is a dimer in solution at the concentrations studied. Given that ocrfl and ocr99 are both dimeric, we assumed that ocr109 would also be a dimer.

Using the program SVEDBERG and extrapolating to infinite dilution, sedimentation velocity measurements gave values of the weight-average sedimentation coefficients ($s^0_{20,w}$) of 2.2 ± 0.1 S and 2.5 ± 0.1 S for ocr99 and ocrfl respectively. Given the mass of these two proteins and estimates of protein hydration, typically between 0.2 and 0.4 g water/g protein, one can

Table 2. Shape of ocrfl and ocr99 determined by sedimentation velocity

Protein	Hydration $g_{water}/g_{protein}$	Frictional ratio due to hydration[a] $(f/f_o)_{hyd}$	Frictional ratio due to asymmetry $(f/f_o)_{sh}$	Axial ratio a/b	Length of semi-axis of revolution a (nm)	Length of equatorial semi-axis b (nm)
Ocrfl	0.2	1.08	1.24 ± 0.05	4.9 ± 0.7	5.7	1.2
Ocrfl	0.3	1.12	1.20 ± 0.05	4.3 ± 0.7	5.2	1.2
Ocrfl	0.4	1.16	1.16 ± 0.05	3.7 ± 0.7	4.8	1.3
Ocr99	0.2	1.08	1.23 ± 0.09	4.6 ± 1.0	5.2	1.1
Ocr99	0.3	1.12	1.18 ± 0.09	4.0 ± 1.2	4.8	1.2
Ocr99	0.4	1.16	1.15 ± 0.09	3.5 ± 1.3	4.3	1.2

[a]The experimental frictional ratio for ocrfl is $f/f_o = 1.35 \pm 0.06$ with partial specific volume $\nu = 0.721$ ml/g. The expected frictional coefficient for a sphere of the same mass as ocrfl is $f_o = 0.3741 \times 10^{-7}$ g/s. The experimental frictional ratio for ocr99 is $f/f_o = 1.33 \pm 0.11$ with partial specific volume $\nu = 0.723$ ml/g. The expected frictional coefficient for a sphere of the same mass as ocr99 is $f_o = 0.3536 \times 10^{-7}$ g/s.

calculate approximate shapes for the proteins using ellipsoidal modelling. Both ocrfl and ocr99 are clearly not spherical and, if modelled as prolate ellipsoids of revolution, would be rather elongated (Table 2). These estimates of shape agree with previous determinations using light scattering (40). Modelling the shape of ocr as an oblate ellipsoid is also possible but produces a non-physical model (40).

Denaturation of ocr and the role of W94

The single tryptophan residue within each monomer of ocr provides a very convenient fluorescent signal for monitoring the effect of deleting the C-terminus on protein conformation. The fluorescence emission, when excited at 295 nm to produce only excitation of the tryptophan residues, showed an emission maximum of 350 nm (data not shown). This is close to that found for pure tryptophan in water, indicating that the tryptophan at position 94 in ocr is highly exposed to the aqueous solvent. Deletion of the C-terminus did not cause a change in the emission maximum of the protein but the fluorescence intensity at 350 nm increased by ~60% for both ocr109 and ocr99 relative to ocrfl (data not shown).

Addition of GuCl was used to denature ocrfl, ocr109 and ocr99 and the emission and anisotropy of W94 were used to follow the process. Unfolding was a co-operative process showing simultaneous decrease in fluorescence emission intensity and a decrease in fluorescence anisotropy as GuCl concentration increased. The fractional change in degree of unfolding was well fitted using a two-state model (Fig. 2). Ocr109 was nearly as stable as ocrfl but ocr99 was significantly destabilised by the loss of the C-terminus (Table 3). As more of the C-terminus was removed, the free energy of stability, the midpoint of the transition and the gradient of the transition all decreased. The small decrease in slope is expected as this parameter is proportional to the difference in accessible surface area between the native and the unfolded protein (41) and will decrease as protein size decreases. The process of denaturation was fully reversible as dilution of GuCl produced refolding curves identical to the unfolding curves (data not shown).

Denaturation of the three proteins was also carried out by heating in a differential scanning calorimeter. The thermograms were not accurately described with a two-state model and a more complex model incorporating an intermediate step where the dimer dissociates to folded monomers prior to denaturation of the monomers was employed (data not shown). The midpoints of thermal denaturation are shown in Table 4 and it is clear that while ocr109 is only 1.1°C less stable than

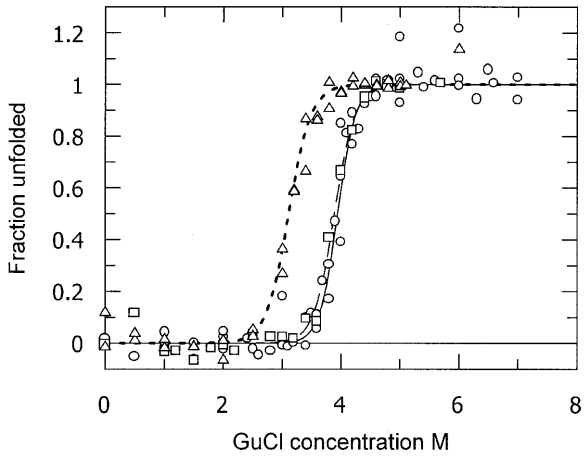

Figure 2. Fractional unfolding of ocrfl (circles), ocr109 (squares) and ocr99 (triangles) as a function of GuCl concentration determined from fluorescence emission intensity at 350 nm excited at 280 nm. Identical transitions were found with emission excited at 295 nm or fluorescence anisotropy. The fitted curves were determined by fitting the data to a two-state model of folding. Upon refolding, the fluorescence signals followed the same transitions.

Table 3. Chemical denaturation[a]

Protein	Free energy ΔG (kJ mol^{-1})	Gradient m (kJ mol^{-1} M^{-1})	Unfolding midpoint (M)
Ocrfl	64.2 ± 9.2	16.3 ± 2.3	3.94
Ocr109	55.2 ± 5.2	14.2 ± 1.3	3.89
Ocr99	40.0 ± 8.2	12.9 ± 2.0	3.11

[a]Excitation at 280 nm and the ratio of emission intensity at 350 nm to emission intensity at 380 nm was measured. Very similar results were obtained with excitation at 295 nm and with fluorescence anisotropy. The denaturation was fully reversible.

Table 4. Differential scanning calorimetry[a]

Protein	Concentration (μM)	Unfolding enthalpy (kJ mol^{-1})	Thermal denaturation midpoint, T_m (°C)	Change in T_m (°C)
Ocrfl	31.5	778 ± 1.0	72.2	0
Ocr109	13.5	699 ± 0.8	71.1	–1.1
Ocr99	22.6	523 ± 1.3	58.4	–13.8
N4C	17.0	728 ± 1.7	72.9	+0.7
D25C	29.1	837 ± 1.8	77.9	+5.7
N43C	41.3	799 ± 3.2	72.5	+0.3
D62C	18.0	770 ± 2.7	70.6	–1.6
S68C	21.7	782 ± 1.1	72.0	–0.2
W94C	27.9	728 ± 1.0	68.6	–3.5

[a]All denaturation transitions were reversible.

Table 5. Activity of ocr *in vivo*

Ocr protein produced[a]	F'101 plasmid present[b]	Efficiency of plating[c]
ocrfl	Yes	0.2600
ocr109	Yes	0.3200
ocr99	Yes	0.3000
N4C	Yes	0.2600
D25C	Yes	0.3000
N43C	Yes	0.3200
D62C	Yes	0.1600
S68C	Yes	0.3800
W94C	Yes	0.3200
None	Yes	0.0008
None	No	1.0000

[a]BL21 (DE3) pLysS cells are deficient in type I R/M systems. The plasmids express low levels of the indicated forms of ocr even in the absence of induction of protein overexpression.
[b]The F'101 plasmid is maintained in the cells and expresses the *Eco*KI type I R/M enzyme.
[c]The efficiency of plating of phage λ on a restricting strain relative to a non-restricting strain.

ocrfl, ocr99 is nearly 14°C less stable. The success of the two-state model in describing chemical denaturation but not thermal denaturation may indicate that tryptophan fluorescence is not sensitive to the presence of the dimer to monomer equilibrium detected by calorimetry. Alternatively, the difference may be attributable to the different magnitudes of the protein concentrations used in the two denaturation experiments.

Dunn *et al.* (12) were unable to distinguish, by successive truncation of the C-terminus, whether W94 was essential for activity or only stability. We used site-directed mutagenesis to make the amino acid substitution W94C within ocrfl. This mutant protein was still active *in vivo* as shown by its ability to reduce the effectiveness of restriction (Table 5). W94C could be purified in the same manner as wild-type ocrfl and it eluted

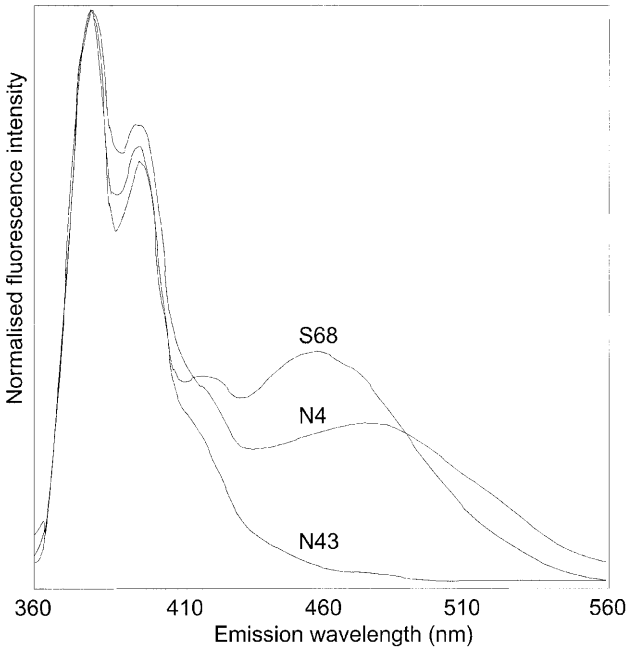

Figure 3. Normalised pyrene emission spectra for chemically modified ocr containing the N4C and S68C substitutions compared to one of the other mutant proteins, N43C. All pyrene-modified mutant proteins except N4C and S68C gave the same emission spectra. The presence of a broad emission band at long wavelengths for the N4C and S68C proteins indicates the formation of an excimer due to close proximity of the two pyrene moieties across the monomer–monomer interface in the ocr dimer.

at the same volume as wild-type ocrfl from the gel filtration column, indicating that it was still dimeric (data not shown). Differential scanning calorimetry (DSC) showed that the unfolding transition midpoint was 3.6°C lower than ocrfl, indicating that this non-conservative change affects stability rather than activity (Table 4).

Location of dimer interface and selected amino acids

In an effort to locate amino acids on the surface of ocrfl and the interface between the two monomers, we employed site-directed mutagenesis. Five amino acids, N4, D25, N43, D62 and S68, in addition to W94 described above, were chosen as targets for replacement by cysteine. Cysteine was chosen as the replacement amino acid as native ocr contains no cysteine. The single cysteine introduced into each subunit would therefore be a suitable target for chemical labelling.

It was possible that the single amino acid substitutions made within ocr could perturb the structure of the protein. Each protein was purified in the same manner as the native ocr and each formed a stable dimeric protein as shown by gel filtration chromatography (data not shown). DSC indicated minor changes in stability except for the D25C substitution, which was markedly stabilised by the decrease in size of side chain (Table 4). All of the mutant proteins showed activity *in vivo* (Table 5). Therefore, we conclude that they are structurally unaltered from wild-type ocrfl.

All six different substitutions of cysteine could be modified by pyrene maleimide to produce a stable protein. The level of

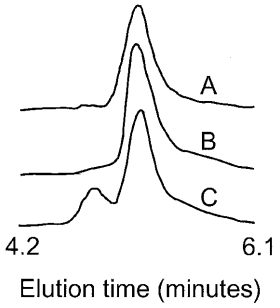

Figure 4. HPLC denaturing gel filtration for the N4C and S68C mutant proteins. In the absence (trace A) of β-mercaptoethanol, the S68C protein eluted in a single peak characteristic of the unfolded protein monomer. The protein containing the N4C substitution eluted as two peaks in the absence of β-mercaptoethanol (trace C), corresponding to unfolded monomer and unfolded dimer joined covalently by a disulphide bond, and predominantly as a single peak in the presence of the reducing agent (trace B). Elution was monitored at 280 nm.

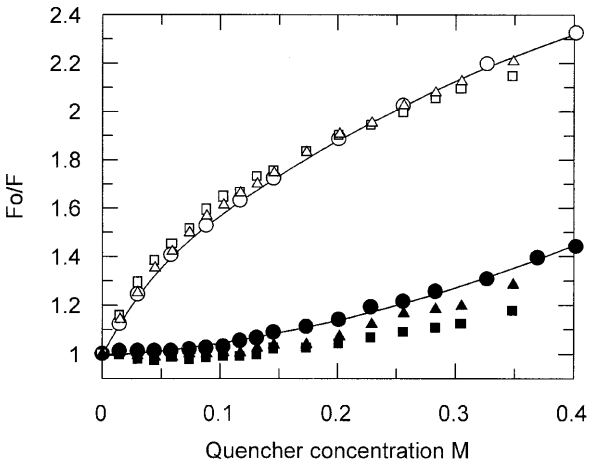

Figure 5. Stern–Volmer plot of the fluorescence quenching of ocrfl by Cs^+ (open circles) and I^- (closed circles). The Cs^+ data were fitted to a two-component quenching mechanism while the line through the I^- data is a simple polynomial to guide the eye. The other data are for ocr109 (squares) and ocr99 (triangles) with open symbols for Cs^+ quenching and closed symbols for I^- quenching.

modification was ~10–30%. The ability to label all six cysteine locations shows that these residues were at least partially exposed to the solvent. Five of the mutant proteins could be labelled by pyrene maleimide in the absence of β-mercaptoethanol, however, the N4C mutant protein could only be labelled when this reducing agent was also present. This suggested the presence of a disulphide bond across the dimer interface. A further indication that the N4 positions on each subunit lay in close proximity to each other came from the appearance of pyrene excimer fluorescence emission (Fig. 3). This could only arise when two pyrene molecules were within ~1 nm of each other and the conjugated ring systems stacked on top of each other (42). Confirmation of the formation of a disulphide bridge across the dimer interface in the N4C mutant protein was sought using esms and denaturing gel filtration. As well as detecting a species with the expected monomer mass, esms also showed a species of mass 27 326 Da, indicative of a dimer. Denaturing gel filtration chromatography showed a single species of low molecular weight, corresponding to a monomer, in buffers containing β-mercaptoethanol and an additional larger species, corresponding to a dimer, in the absence of the reducing agent. This additional species did not appear with ocrfl or any of the other mutant proteins (Fig. 4).

One other mutant protein produced pyrene excimer emission, namely S68C (Fig. 3). This protein could be labelled in the absence of reducing agent and produced only monomeric species as detected by esms and denaturing gel filtration (Fig. 4). Therefore, in the absence of pyrene labelling, the sulphydryl groups were too far apart for disulphide bond formation. The shift in the wavelength of maximum pyrene excimer emission relative to that from pyrene attached to N4C may be consistent with this greater separation across the dimer interface. This shift in the wavelength maximum suggested that the interaction between the two pyrene labels was different.

Ocr as a DNA mimic

Ocrfl contains a large excess of acidic amino acid side chains over basic side chains. This observation prompted the suggestion that the excess negative charge could be responsible for the tight binding of ocr to the DNA binding site of *Eco*KI (9,12). However, the removal of the highly negative C-terminus did not affect restriction activity (Table 5). It is possible that other highly charged regions exist on ocr and we decided to determine whether W94 was located in a highly charged region of the surface. Fluorescence quenching using the two charged quenchers Cs^+ and I^- was performed. It is known that Cs^+ is a poor dynamic quencher of tryptophan with only ~20% quenching efficiency whilst I^- is 100% efficient as a dynamic quencher. Their charge prevents both ions from penetrating into the protein matrix to quench buried fluorophores (35).

Normally, I^- quenches fluorescence better than Cs^+. In fact, the opposite was observed with I^- being virtually unable to cause any quenching and Cs^+ being an exceptionally good quencher of ocrfl (Fig. 5). We are not aware of any other quenching study in which Cs^+ is a better quencher than I^-. Cs^+ quenching produced a curved quenching profile requiring a two component fit. The quenching constants for these data were 0.56 ± 0.19 M^{-1} (46%) and 15.35 ± 2.23 M^{-1} (54%). The smaller of these quenching constants was similar to that found for free tryptophan (2.25 ± 0.02 M^{-1}; data not shown), indicating that the ion had easy access to a fraction of the conformational space available to the fluorophore. The very high quenching constant is far greater than physically possible by a normal collisional quenching mechanism. It is likely that this fraction of the quenching is due to the preferential attraction of Cs^+ to a negatively charged location near to the tryptophan residue.

The observed decrease in fluorescence emission intensity may result from dynamic collisional quenching during the lifetime of the fluorophore excited state or by static quenching of the ground state of the fluorophore. These two mechanisms were discriminated using time-resolved fluorescence spectroscopy. The fluorescence decay of the tryptophan emission from ocrfl could be fitted to a sum of two exponentials with lifetimes of 1.74 ± 0.13 ns (33.5%) and 4.30 ± 0.04 ns (66.5%) with a χ^2

value of 1.2 (data not shown). The two fluorescence lifetimes of W94 in ocrfl were both quenched by 0.2 M Cs^+ to values of 1.10 ± 0.13 ns (42.3%) and 2.62 ± 0.04 ns (57.7%), indicating that the environments giving rise to the different lifetimes were equally accessible to the quencher and that dynamic quenching is operative.

In contrast to Cs^+ quenching, I^- quenching was very poor (Fig. 5). At high concentrations of I^-, quenching was observed but so was some aggregation of protein, suggesting that the protein conformation was changing. Quenching by low concentrations of I^- was negligible with a quenching constant of <1 M^{-1}. This constant was far lower than the value of 12.23 ± 0.09 M^{-1} found for free tryptophan (data not shown). This indicated that a negatively charged region around W94 was repelling the negatively charged quencher.

The quenching behaviour of W94 in both ocr109 and ocr99 proteins with both quenchers was virtually identical to that in ocrfl (Fig. 5). This indicated that the negative charge repelling I^- was not a part of the highly acidic C-terminus found after W94.

Previous methods used to examine the interaction between ocr and type I R/M enzymes focused on inhibition of the various activities expressed by the enzyme such as DNA methylation, DNA cleavage and ATP hydrolysis (8,9). It was found that ocr could displace the type I R/M enzyme from DNA as long as ATP hydrolysis had not commenced. Inhibition was stoichiometric, implying tight binding of the ocr to the enzyme. Displacement of the DNA from the enzyme was shown by filter binding, gel filtration and electron microscopy (9). We employed an equilibrium method using a DNA duplex labelled with a fluorescent dye, hexachlororfluorescein (16). Displacement of the labelled duplex from M.*Eco*KI by binding of ocr led to a decrease in the fluorescence anisotropy of the dye. Figure 6A shows the increase in anisotropy of the DNA label as M.*Eco*KI was added to saturating levels, and the subsequent dissociation of the M.*Eco*KI–DNA complex as ocr was added. At the high concentrations used, binding of both M.*Eco*KI and ocr was tight and indicated that approximately one molecule of ocr dimer was sufficient to displace the DNA from the enzyme. The end point of the titration was not sufficiently clear to rule out other stoichiometric ratios and does not rule out other sites for binding ocr that do not affect DNA binding. Titration of 1 nM DNA with M.*Eco*KI in the presence of ocr concentrations ranging from 2 to 5000 nM gave binding curves shifted to progressively higher concentrations of M.*Eco*KI (Fig. 6B). Fitting the data with a single site binding equation for binding of M.*Eco*KI to DNA showed a decrease in apparent binding affinity as the ocr concentration was increased. The data were not well described by this model and there is no justification for a change in $K_{d,DNA}$ in the presence of ocr. Fixing $K_{d,DNA}$ to that found in the absence of ocr, namely 2.7 nM, and adding an equilibrium binding term for binding of ocr to M.*Eco*KI gave reasonable fits to all sets of data. The binding affinity $K_{d,ocr}$ for binding of ocr to M.*Eco*KI was estimated at ~100 pM but errors were very large, being ±100%. This binding affinity is far greater than that between M.*Eco*KI and DNA and lower than the concentration of DNA employed in the titration. These factors would account for the large error in the determination of $K_{d,ocr}$. Our instrument and the fluorescence anisotropy assay are not sensitive enough to better define $K_{d,ocr}$ but at least put an upper bound on its value.

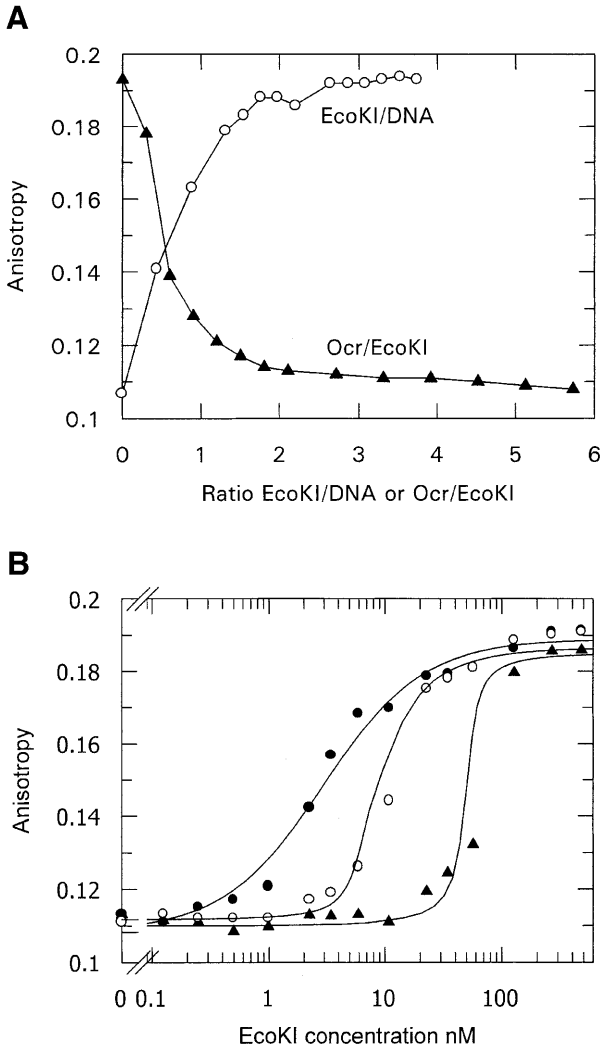

Figure 6. (**A**) Titration of a fluorescently labelled DNA oligonucleotide duplex with *Eco*KI methyltransferase (circles) increased the fluorescence emission anisotropy until all of the DNA was bound and a plateau was reached. Subsequent addition of ocr to the M.*Eco*KI–DNA complex (triangles) displaced the DNA from M.*Eco*KI and the anisotropy returned to the original value. The concentrations of all components were well above the expected dissociation constants for these interactions and the binding processes are essentially stoichiometric. It is apparent that approximately one molecule of the ocr dimer is sufficient to displace DNA from M.*Eco*KI. The initial DNA concentration was 1 µM and M.*Eco*KI was added to a final concentration of 3.8 µM. Ocrfl was then added to a final concentration of 22 µM. (**B**) Titration of 1 nM fluorescently labelled DNA with M.*Eco*KI in the presence of different concentrations of ocrfl. The anisotropy of the fluorophore increased as DNA was bound by M.*Eco*KI but the midpoint of the binding curve shifted to progressively higher concentrations as the concentration of ocrfl present increased. The concentrations of ocrfl were 0 nM (closed circles), 5 nM (open circles) and 50 nM (triangles). The fitted lines were determined using the competitive equilibrium model as described.

DISCUSSION

Our results can be summarised as follows. Ocr remains a dimeric protein even in the absence of the C-terminal 17 amino acids and it is definitely not spherical, in agreement with previous light scattering measurements (40). Denaturation

analysis shows that amino acids between 99 and 109 and the tryptophan side chain of W94 are important for protein stability. The earlier observation (12) that truncation of ocr to <94 amino acids leads to loss of activity can be attributed to a loss of stability rather than to the loss of W94 as the non-conservative substitution W94C does not affect activity. The pyrene excimer emission from both N4C and S68C mutant proteins indicates that both N4 and S68 are near to the interface between the two subunits of ocr but, since they can be chemically modified without denaturing the protein, they are not buried in the interface. The other four amino acid substitutions are all located upon the surface of ocr but distant from the interface. Fluorescence quenching indicates that W94 is located within a highly negatively charged patch upon the surface of ocr. The observation of similar quenching behaviour with ocr109 and ocr99 rules out the possibility that the negatively charged patch around W94 is comprised of amino acids 99–116. This is consistent with the observation that the C-terminus from amino acid 95 onwards is not required for activity *in vivo* (12). Lastly, we quantified earlier observations by Bandyopadhay *et al.* (9) that ocr displaces DNA from type I R/M enzymes and the smaller methyltransferase core of these enzymes with a binding affinity of at least 100 pM.

Type I R/M enzymes cover a length of ~30 bp of DNA depending upon the particular R/M enzyme (13–17). It has been suggested that ocr, with its ability to block the DNA binding sites of type I R/M enzymes, may resemble some aspect of DNA structure (9,12). This resemblance cannot extend to the fine details of DNA sequence recognition as the ocr protein acts against type I R/M enzymes with completely different DNA target sequences (8–11).

Our data show that the shape and size of ocr can be modelled as an elongated prolate ellipsoid in agreement with light scattering experiments (40). Preliminary X-ray crystallographic analysis of ocr also supports this elongated shape for the protein (M.Walkinshaw and D.Dryden, unpublished results). Intriguingly, this ellipsoidal model is similar in size and shape to the 25–30 bp DNA binding site of the methyltransferase core of a type I R/M enzyme. Therefore, it is feasible that ocr fits directly into the DNA binding site of the type I R/M enzyme in a manner similar to that found for the complex of uracil glycosylase inhibitor (UGI) with uracil glycosylase (43–45). UGI inhibits uracil glycosylase by mimicking electrostatic and structural interactions normally formed between uracil glycoslyase and DNA (43–45). It is also formally possible that ocr obscures the DNA binding site or causes some other more indirect conformational change rather than directly binding within the site. However, we believe it more likely that ocr binds directly within the DNA binding site of the type I R/M enzyme because its function in nature is so similar to that of UGI. UGI protects the DNA of phage PBS2, which contains uracil instead of thymine, from attack by the uracil glycosylase of *Bacillus subtilis* (46) just as ocr protects the DNA of phage T7, which contains unmethylated adenine instead of methylated adenine (47).

ACKNOWLEDGEMENTS

We thank Mr Gordon Campbell (Glasgow) for assistance with ultracentrifugation experiments, Dr Andy Cronshaw (Edinburgh) for running the esms samples and Drs Richard Virden and Jeremy Lakey (Newcastle University) for the loan of the 199T fluorimeter. The assays of *in vivo* activity were performed with the assistance of Professor Noreen Murray and Dr Svetlana Makovets (Edinburgh). Finally, we thank Professor William Studier and Dr Alan Rosenberg (Brookhaven National Laboratory) for generously supplying the ocr overexpression plasmids. H. McMiken contributed to this work as part of her undergraduate degree. This work was supported by the Leverhulme Trust (grant F/158/BC). S.S. thanks the Leverhulme Trust for their support, C.A. was supported by a Darwin Trust studentship and D.D. thanks the Royal Society for a University Research Fellowship.

REFERENCES

1. Wilson,G.G. and Murray,N.E. (1991) Restriction and modification systems. *Annu. Rev. Genet.*, **25**, 585–627.
2. Bickle,T.A. and Kruger,D.H. (1993) Biology of DNA restriction. *Microbiol. Rev.*, **57**, 434–450.
3. Murray,N.E. (2000) Type I restriction systems: sophisticated molecular machines. *Microbiol. Mol. Biol. Rev.*, **64**, 412–434.
4. Kruger,D.H. and Bickle,T.A. (1983) Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiol. Rev.*, **47**, 345–360.
5. Belogurov,A.A. and Delver,E.P. (1995) A motif conserved among the type I restriction–modification enzymes and antirestriction proteins: a possible basis for mechanism of action of plasmid-encoded antirestriction functions. *Nucleic Acids Res.*, **23**, 785–787.
6. Studier,F.W. (1975) Gene 0.3 of bacteriophage T7 acts to overcome the DNA restriction system of the host. *J. Mol. Biol.*, **94**, 283–295.
7. Moffatt,B.A. and Studier,F.W. (1988) Entry of bacteriophage T7 DNA into the cell and escape from host restriction. *J. Bacteriol.*, **170**, 2095–2105.
8. Mark,K.-K. and Studier,F.W. (1981) Purification of the gene 0.3 protein of bacteriophage T7, an inhibitor of the DNA restriction system of *Escherichia coli*. *J. Biol. Chem.* **256**, 2573–2578.
9. Bandyopadhay,P.K., Studier,F.W., Hamilton,D.L. and Yuan,R. (1985) Inhibition of the type I restriction-modification enzymes *Eco*B and *Eco*K by the gene 0.3 protein of bacteriophage T7. *J. Mol. Biol.* **182**, 567–578.
10. Kruger,D.H., Schroeder,C., Hansen,S. and Rosenthal,H.A. (1977) Active protection by bacteriophages T3 and T7 against *E.coli* B- and K-specific restriction of their DNA. *Mol. Gen. Genet.*, **153**, 99–106.
11. Kruger,D.H., Hansen,S. and Reuter,M. (1983) The ocr$^+$ gene function of bacteriophages T3 and T7 counteracts the *Salmonella typhimurium* DNA restriction systems SA and SB. *J. Virol.*, **45**, 1147–1149.
12. Dunn,J.J., Elzinga M., Mark,K.K. and Studier,F.W. (1981) Amino acid sequence of the gene 0.3 protein of bacteriophage T7 and nucleotide sequence of its messenger RNA. *J. Biol. Chem.* **256**, 2579–2585.
13. Taylor,I., Watts,D. and Kneale,G. (1993) Substrate recognition and selectivity in the type IC DNA modification methylase M.*Eco*R124I. *Nucleic Acids Res.*, **21**, 4929–4935.
14. Mernagh,D.R. and Kneale,G.G. (1996) High resolution footprinting of a type I methyltransferase reveals a large structural distortion within the DNA recognition site. *Nucleic Acids Res.*, **24**, 4853–4858.
15. Mernagh,D.R., Janscak,P., Firman,K. and Kneale,G.G. (1998) Protein–protein and protein–DNA interactions in the type I restriction endonuclease R.*Eco*R124I. *Biol. Chem.*, **379**, 497–503.
16. Powell,L.M., Connolly,B.A. and Dryden,D.T.F. (1998) The DNA binding characteristics of the trimeric *Eco*KI methyltransferase and its partially assembled dimeric form determined by fluorescence polarization and DNA footprint. *J. Mol. Biol.*, **283**, 947–961.
17. Powell,L.M., Dryden,D.T.F. and Murray,N.E. (1998) Sequence-specific DNA binding by *Eco*KI, a type IA DNA restriction enzyme. *J. Mol. Biol.*, **283**, 963–976.
18. Rost,B. and Sander,C. (1993) Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.*, **232**, 584–599.
19. Rost,B. and Sander,C. (1994) Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins*, **19**, 55–72.
20. Rost,B. and Sander,C. (1994) Conservation and prediction of solvent accessibility in protein families. *Proteins*, **20**, 216–226.
21. Makovets,S., Doronina,V.A. and Murray,N.E. (1999) Regulation of endonuclease activity by proteolysis prevents breakage of unmodified

bacterial chromosomes by type I restriction enzymes. *Proc. Natl Acad. Sci. USA*, **96**, 9757–9762.
22. Atanasiu,C. (2000) Characterisation of Ocr, the product of gene 0.3 from bacteriophage T7. PhD Thesis, University of Edinburgh, UK.
23. Dryden,D.T.F., Cooper,L.P. and Murray,N.E. (1993) Purification and characterization of the methyltransferase from the type I restriction and modification system of *Escherichia coli* K12. *J. Biol. Chem.*, **268**, 13228–13236.
24. Pace,C.N. and Schmidt,F.X. (1997) How to determine the molar absorbance coefficient of a protein. In Creighton,T.E. (ed.), *Protein Structure: A Practical Approach*. IRL Press, Oxford, UK, pp. 253–259.
25. Pace,C.N. and Scholtz,J.M. (1997) Measuring the conformational stability of a protein. In Creighton,T.E. (ed.), *Protein Structure: A Practical Approach*. IRL Press, Oxford, UK, pp. 299–348.
26. Philo,J.S. (1994) Measuring sedimentation, diffusion and molecular weights of small molecules by direct fitting of sedimentation velocity concentration profiles. In Schuster,T.M. and Laue,T.M. (eds), *Modern Analytical Ultracentrifugation*. Birkhauser, Boston, Basel, Berlin, pp. 156–170.
27. Cohn,E.J. and Edsall,J.T. (1943) *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*. Reinhold Publishing Corp., New York, NY.
28. Perkins,S.J. (1986) Protein volumes and hydration effects. *Eur. J. Biochem.*, **157**, 169–180.
29. Kim,H., Deonier,R.C. and Williams,J.W. (1977) The investigation of self association reactions by equilibrium ultracentrifugation. *Chem. Rev.*, **77**, 659–690.
30. Johnson,M.L., Correia,J.J., Yphantis,D.A. and Halvorson,H.R. (1981) Analysis of data from the analytical ultracentrifuge by nonlinear least squares techniques. *Biophys. J.*, **36**, 575–588.
31. Fujita,H. (1975) *Foundations of Ultracentrifugal Analysis*. Wiley, New York, NY.
32. Haschemeyer,R.H. and Haschemeyer,A.E.V. (1973) *Proteins: A Guide To Study By Physical And Chemical Methods*. J. Wiley and Sons, New York, NY.
33. Durchschlag,H., Zipper,P., Purr,G. and Jaenicke,R. (1996) Comparative studies of structural properties and conformational changes of proteins by analytical ultracentrifugation and other techniques. *Colloid Polymer Sci.*, **274**, 117–137.
34. Durchschlag,H. and Zipper,P. (1999) Calculation of structural parameters from hydrodynamic data. *Prog. Colloid Polymer Sci.*, **113**, 87–105.
35. Eftink,M.R. and Ghiron,C.A. (1981) Fluorescence quenching studies with proteins. *Anal. Biochem.*, **114**, 199–227.
36. Birch,D.J.S. and Imhof,R.E. (1985) Kinetic interpretation of fluorescence decays. *Anal. Instrum.*, **14**, 293–329.
37. Kuzmic,P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **237**, 260–273.
38. Wu,C.W., Yarbrough,L.R. and Wu,F.Y.H. (1976) *N*-(1-Pyrene)maleimide: a fluorescent cross-linking reagent. *Biochemistry*, **15**, 2863–2868.
39. Dryden,D.T.F., Cooper,L.P., Thorpe,P.H. and Byron,O. (1997) The *in vitro* assembly of the *Eco*KI type I DNA restriction/modification enzyme and its *in vivo* implications. *Biochemistry*, **36**, 1065–1076.
40. Blackstock,J.J., Egelhaaf,S.U., Atanasiu,C., Dryden,D.T.F. and Poon,W.C.K. (2001) Shape of ocr, the gene 0.3 protein of bacteriophage T7: modelling base on light scattering experiments. *Biochemistry*, in press.
41. Hamaguchi,K. (1992) *The Protein Molecule*. Springer Verlag, Berlin, Germany.
42. Lehrer,S.S. (1997) Intramolecular pyrene excimer fluorescence: a probe of proximity and protein conformational change. *Methods Enzymol.*, **278**, 286–295.
43. Mol,C.D., Arvai,A.S., Sanderson,R.J., Slupphaug,G., Kavli,B., Krokan,H.E., Mosbaugh,D.W. and Tainer,J.A. (1995) Crystal-structure of human uracil–DNA glycosylase in complex with a protein inhibitor—protein mimicry of DNA. *Cell*, **82**, 701–708.
44. Savva,R. and Pearl,L.H. (1995) Nucleotide mimicry in the crystal structure of the uracil–DNA glycosylase–uracil glycosylase inhibitor protein complex. *Nature Stuct. Biol.*, **2**, 752–757.
45. Putnam,C.D., Shroyer,M.J.N., Lundquist,A.J., Mol,C.D., Arvai,A.S., Mosbaugh,D.W. and Tainer,J.A. (1999) Protein mimicry of DNA from crystal structures of the uracil–DNA glycosylase inhibitor protein and its complex with *Escherichia coli* uracil–DNA glycosylase. *J. Mol. Biol.*, **287**, 331–346.
46. Beger,R.D., Balasubramamian,S., Bennett S.E., Mosbaugh,D.W. and Bolton,P.H. (1995) Tertiary structure of uracil–DNA glycosylase inhibitor protein. *J. Biol. Chem.* **270**, 16840–16847.
47. Eskin,B., Lautenberger,J.A. and Linn,S. (1973) Host-controlled modification and restriction of bacteriophage T7 by *Escherichia coli* B. *J. Virol.* **11**, 1020–1023.