PROTEINS: Structure, Function, and Genetics 32:248–261 (1998)

Structural Features of the Plasmid pMV158-Encoded Transcriptional Repressor CopG, a Protein Sharing Similarities With Both Helix-Turn-Helix and β-Sheet DNA Binding Proteins

Paloma Acebo, Mario García de Lacoba, Germán Rivas, José M. Andreu, Manuel Espinosa, and Gloria del Solar*

Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

ABSTRACT **The small transcriptional repressor CopG protein (45 amino acids) encoded by the streptococcal plasmid pMV158 was purified to near homogeneity. Gel filtration chromatography and analytical ultracentrifugation showed that the native protein is a spherical dimer of identical subunits. Circular dichroism measurements of CopG indicated a consensus average content of more than 50% α-helix and 10–35% β-strand and turns, which is compatible with the predicted secondary structure of the protein. CopG exhibited a prolonged intracellular half-life, but deletions in regions other than the C-terminal affected the global structure of the protein, severely reducing the half-lives of the CopG variants. This indicates that CopG has a compact structure, perhaps constituted by a single domain. Molecular modeling of CopG showed a good fitting between the helix-turn-helix motifs of well-known repressor proteins and a bihelical unit of CopG. However, modeling of CopG with ribbon-helix-helix class of DNA binding proteins also exhibited an excellent fit. Eleven out of the 12 replicons belonging to the pMV158 plasmid family could also encode Cop proteins, which share features with both helix-turn-helix and β-sheet DNA binding proteins. Proteins 32:248–261, 1998.** © **1998 Wiley-Liss, Inc.**

Key words: copy number control; plasmids; transcriptional repressors; homology modeling; protein comparison

INTRODUCTION

The most relevant feature of bacterial plasmids is that replication of their DNA is autonomous and subjected to autogenous control. Control of replication by plasmid-encoded elements defines and maintains a mean plasmid copy number, which is constant for a given host under fixed growth conditions. In many cases, the stage subjected to such a control is the synthesis of the plasmid-encoded initiator of replication (Rep) protein, whose concentration determines the rate of initiation of the replicative process.[45] In these cases, control of replication is generally achieved by pairing between a leader region in the *rep* mRNA and a *trans*-acting antisense RNA, which results in either premature termination of the synthesis of the *rep* mRNA, or inhibition of *rep* translation.[47] This strategy may be combined with repression of the *rep* promoter.[17] If so, the *rep* promoter can be repressed either by the same Rep protein involved in initiation,[31] or by a different plasmid-encoded transcriptional repressor protein (generically termed Cop), which regulates the levels of *rep* mRNA.[6,16] Although these two situations are mechanistically similar, plasmid-encoded Cop proteins offer the possibility of studying a protein which has, in principle, only one biological function.

Genes encoding Cop proteins have been reported and characterized in a variety of theta-replicating plasmids isolated from Gram-negative[27] and Gram-positive bacteria.[6] In addition, putative *cop* genes have been found in a family of rolling circle-replicating plasmids,[18] although physical evidence of Cop repressor proteins has only been provided in plasmids pMV158[15] and pE194.[10] Plasmid pMV158 encodes CopG, a 45 amino acid-long protein. Gene *copG* is co-transcribed with the *repB* gene from the single promoter P_{cr}, which is included within the target of CopG[16] (see Fig. 1). Computer-aided predictions indicate that CopG has a simple secondary structure, presenting one β-strand and two α-helices separated by a small turn.[15] Based on similarities with well-characterized transcriptional repressors,[32,33] the region including the two putative α-helices of CopG was postulated to constitute a helix-turn-helix (HTH) motif involved in the binding to its cognate DNA.[15,16] However, no function was as-

Abbreviations: HTH, helix-turn-helix; RC, random coil.

Grant sponsor: CICYT; Grant number: BIO97–0347; Grant sponsor: DGCyT; Grant numbers: PB95–0116 and PB95-0120; Grant sponsor: the CSIC Program on Protein Structure and Function.

*Correspondence to: G. del Solar, Centro de Investigaciones Biológicas, Velázquez, 144, E-28006 Madrid, Spain.

Received 9 January 1998; Accepted 12 March 1998

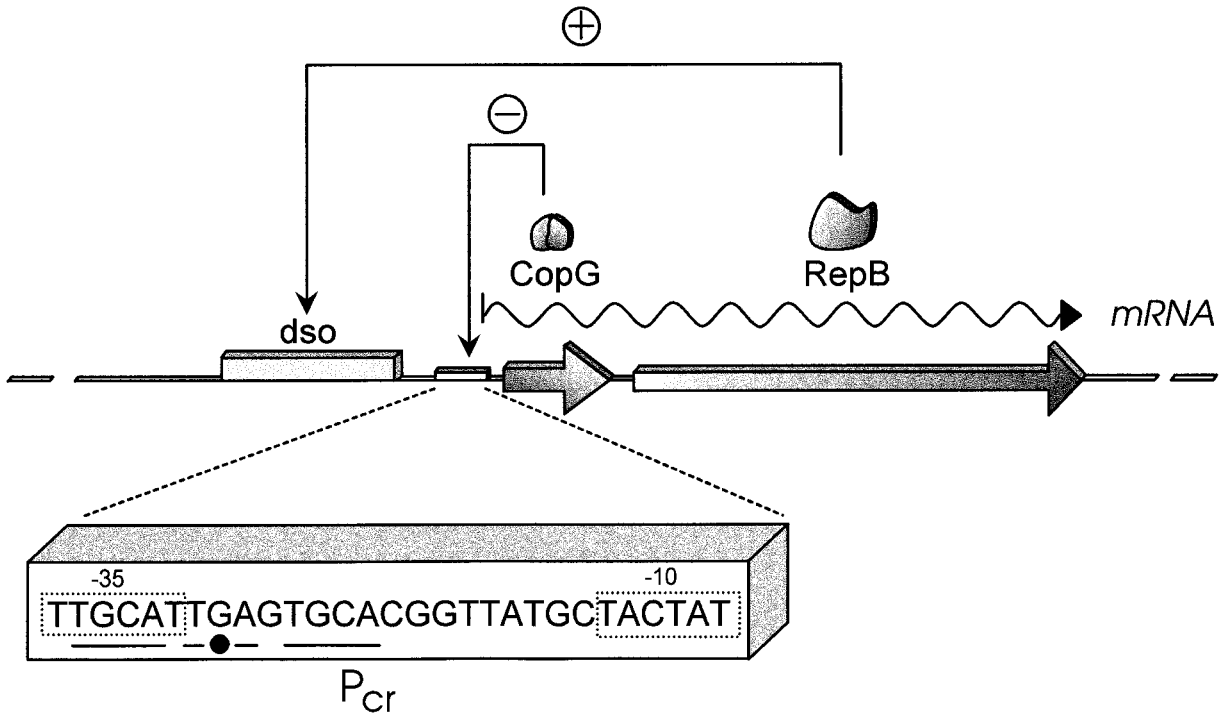


Fig. 1. Scheme of the CopG-dependent regulation of the replicative process in plasmid pMV158. The positive effector (+) is the RepB protein, which initiates replication at the *dso*. Synthesis of RepB is negatively controlled (-) at the transcriptional level by the CopG protein. Genes *copG* and *repB* are represented by arrows. The region encompassing the -35 and -10 boxes of the P_{cr} promoter, which directs synthesis of the *cop-rep* mRNA (wavy arrow), is shown amplified below. Overlapping with the -35 box, the 13-bp symmetric element is represented, with palindromic bases underlined and the center of symmetry indicated by a point.

signed to the β-strand region of the protein, in spite of the secondary structure predicted for CopG being similar to that found in the bacteriophage P22 Arc protein dimer.[7] Arc repressor is a small protein (53 amino acids/monomer) that belongs to the ribbon-helix-helix class of DNA-binding proteins. This group of prokaryotic regulatory proteins uses an antiparallel β-sheet, constituted by a β-strand from each of two identical monomers, to make specific base contacts in the major groove of its target DNA.[33] CopG has been chemically synthesized, and the chemical protein has been shown to have an activity indistinguishable from the biologically purified protein, indicating that chemical synthesis did not affect the proper folding of the protein.[19]

In the present work, we performed gel filtration, analytical ultracentrifugation, and circular dichroism measurements with purified CopG to define the biophysical parameters of the protein. Several mutations within the *copG* gene have been constructed. Overexpression of these mutant genes showed that deletions affecting regions other than the C-terminal end of the protein strongly reduced the intracellular stability of the respective CopG variants. Based on similarities with both the HTH and the β-sheet DNA binding motifs of well-characterized regulatory proteins,[2,3,11,33] we performed molecular modeling of putative structural motifs of CopG. In addition to pMV158, ten replicons belonging to this plasmid family may encode Cop proteins which share features with CopG from pMV158.

MATERIALS AND METHODS

Biological Material: Bacterial Strains and Plasmids

Escherichia coli BL21(DE3) ($r_B^-m_B^-$, *gal, ompT, int::*P_{lacUV5}-T7 gene 1 *imm21 nin5*; a gift of F.W. Studier) was used as the host for the in vivo expression of genes. This strain contains a single copy of the gene for bacteriophage T7 RNA polymerase inserted in the chromosome, under the control of the inducible *lac*UV5 promoter.[44] The expression plasmids were based on the commercially available phagemid pALTER-1, and they carried the pMV158-*copG* (wild-type or mutants) and -*repB* genes, and the pC194-*cat* gene, under the control of the ϕ10 promoter of phage T7. These plasmids were used for overexpression of *wt* and mutant *copG* alleles, in order to purify the gene products and to measure their intracellular stability. Construction of plasmids harboring *copG* or *copG*7 genes has been described previously,[20] whereas the plasmid containing *copG*8 was constructed here, following the same fragment-swapping procedure as for the plasmids above. Cultures of the various host/recombinant plasmid strains were treated with IPTG and with

rifampicin to selectively express genes under the ϕ10 promoter.[15,20]

Construction of Deleted *copG* Genes by Site-Directed Mutagenesis

The "Altered Sites" kit (Promega, Madison, WI), designed for in vitro mutagenesis, was used. To construct the different deletion mutants, the following oligonucleotides (corresponding to the coding strand of pMV158, and spanning between the plasmid coordinates indicated in parentheses[23]) were employed: i) 46-mer *Δrc1* (638–692): 5′-ATTTTGAGAGGTGACGCATG/TTGACGATAACATTAAGTGAATCGGT-3′ ii) 44-mer *Δβ* (646–698): 5′-AGGTGACGCATGAAAAAAAGATTG/TTAAGTGAATCGGTACTTGA-3′ iii) 45-mer *Δα1* (681–734): 5′-AAGTGAATCGGTACTTGAAAATCTT/GCAAGAGAGATGGGGTTATC-3′ iv) 45-mer *Δrc2* (752–805): 5′-CTGTTGCCTTGGAAAATTAC/CAAGAAAAATAAAAAAAGCCGTGCT-3′.

The slash indicates the junction between the ends of the 9 nucleotide-long deletion. Each oligonucleotide was annealed to single-stranded DNA prepared from the expression plasmid. This single-stranded DNA contained the noncoding strand of the *wt copG* gene. The mutagenesis procedure was carried out as specified by the supplier. Expression of the truncated *copG* genes, constructed in this way, would also be under the control of the ϕ10 promoter. The different mutants obtained were characterized by determination of the entire nucleotide sequence of their *copG* genes. Mutants *Δrc1*, *Δβ*, *Δα1*, and *Δrc2*, lacked, respectively, codons 2–4, 6–8, 18–20, and 40–42 of the *copG* gene.

Purification of CopG Proteins

Wild-type CopG, and variant CopG7 and CopG8 proteins were overproduced in *E. coli* BL21(DE3) cells harboring the corresponding expression plasmid. Proteins were purified following a new large-scale purification method, based on that previously reported.[17] When labeled CopG proteins were to be used for gel filtration chromatography, a 5-ml sample of crude protein extract was mixed with 500 µl (12,000 cpm μl^{-1}) of extract labeled with [^{35}S]Met as described previously[15] before starting the purification steps. Due to the presence of a single Tyr residue in the amino acid sequence of CopG, determination of the protein concentration using an extinction coefficient of 1,490 M^{-1} cm^{-1} at 280 nm was not very accurate with small amounts of the protein. For this reason, CopG concentration was determined by quantitative amino acid analysis. In addition, amino acid sequence determination of purified CopG and chemical synthesis of the protein were performed following published procedures.[15,19]

Gel Filtration Chromatography

A 38-ml (1 × 48 cm) Agarose (Bio Rad A-0.5 m, 200–400 mesh) column, equilibrated in buffer A (20 mM Tris/HCl pH 7.5, 500 mM NaCl, 1 mM EDTA, 5% ethylene glycol), was used for the chromatographic analysis of the molecular size of native CopG protein. Samples consisting of 200 µl of protein solution in buffer A supplemented with 5% glycerol were applied under the eluent. All chromatographic runs were performed at 4°C, with a flow rate of 4.5 ml h^{-1}, and the column was washed with five volumes of buffer A before a new protein sample was loaded. The column was calibrated by loading solutions (3–5 mg ml^{-1}) of standard proteins of known Stokes radius, as the behavior of proteins during gel filtration correlates with the molecular size rather than with the molecular weight.[13,42] The standards used were: bovine serum albumin (35.5 Å), carbonic anhydrase (20.1 Å), cytochrome c (17.4 Å), and ribonuclease A from bovine pancreas (16.4 Å). Elution positions of the standard proteins was monitored with an on-line UV monitor (Isco) at 280 nm. The test sample consisted of a solution of pure [^{35}S]Met-labeled, CopG protein (12 µM, specific activity 10,000 cpm $nmol^{-1}$). In this case, the elution profile was monitored by collecting 0.5-ml fractions from the eluate and measuring the radioactivity in each fraction. The K_{av} parameter was calculated for each protein as:[42]

$$K_{av} = (V_e - V_0)/(V_t - V_0)$$

where V_e is the elution volume of the center of the protein peak; V_0 is the void volume (determined by elution of blue dextran); and V_t is the total volume of the gel bed.

The frictional ratio (f/f_0) was calculated from the equation: $f/f_0 = R_s/(3\nu Mr/4\pi N)^{1/3}$, where R_s = Stokes radius, ν = partial specific volume, N = Avogadro's number, M_r = molecular mass[28,42]. The molecular mass value of the native CopG protein was obtained from equilibrium ultracentrifugation ($M^c_{w,a}$).

MALDI-TOF Mass Spectrometric Analysis of CopG

Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF, Beckman, Palo Alto, CA) mass spectrometry of purified CopG protein was performed on a Brucker Biflex instrument (Bruker-Franzen Analytik, Bremen, Germany) using insulin as standard. The spectra (average of 100 shots) were recorded in the linear mode at 19.5 kV.

Sedimentation Equilibrium

The experiments were performed in a Beckman Optima XL-A analytical ultracentrifuge equipped with absorbance optics and using an An60Ti rotor. Standard cells (12 mm optical pathlength) with double sector centerpieces of charcoal-filled Epon were used. Protein CopG (80 µl, loading concentra-

TABLE I. General Properties of CopG

Number of residues: 45	
Amino acid sequence	
Edman's degradation:	
MKKRLTITLSESVLENLEKMAREMGL-SKSAMISVAL(E)NY(K)KGQ(EK) (Residues in parenthesis were not unambiguously determined)	
Chemical synthesis:	
MKKRLTITLSESVLENLEKMAREMGL-SKSAMISVALENYKKGQEK	
Molecular weight determined from:	
DNA sequence	5,118.7
MALDI-TOF	5,085.7
SDS-Urea-PAGE	5,000
Analytical ultracentrifugation	10,400
Partial specific volume: 0.752 ml/g	
Stokes radius: 16 Å (partition chromatography)	
Frictional ratio (f/f_o): 1.1	

tion range of 12–200 µM) equilibrated in buffer B (20 mM Tris/HCl, pH 8.0, and 75 mM KCl) was centrifuged at 25,000 rpm and 20°C until sedimentation equilibrium was reached. Then absorbance scans were taken at 230 and 275 nm. In several experiments a second equilibrium speed was also performed (30,000 rpm). Baseline offsets were determined by high-speed centrifugation (42,000 rpm). Whole cell apparent weight-average molecular weights ($M^c_{w,a}$) were determined with the programs XLAEQ and EQASSOC (supplied by Beckman[30]), using a partial specific volume for CopG of 0.752 ml g^{-1}. This value (see Table I) was determined from the amino acid composition of the protein.[24,25]

Sedimentation Velocity

Sedimentation velocity experiments were performed as described above, except that protein samples of 300 µl were centrifuged at 60,000 rpm. The experimental data were analyzed with the program SVEDVERG,[37] which uses the Faxen's approximation of the Lamm equation to simultaneously fit multiple datasets in up to four different sedimenting species. The sedimentation coefficients obtained were corrected to standard conditions to get the corresponding $S_{20,w}$ coefficients.[46] The translational frictional coefficient (f) of CopG was determined from the molecular mass and sedimentation coefficient of the protein.[46,48] The frictional coefficient of the equivalent hydrated sphere (f_0) was estimated using a hydration of 0.3 g H_20 per g protein.[36] From these parameters, the translational frictional ratio (f/f_0) was calculated, which allows a gross estimation of the hydrodynamic shape of CopG.[48]

Circular Dichroism and Fluorescence Spectroscopy

CD spectra of 20 to 200 µM CopG (0.9 to 9 mM mean residue) were acquired with a Jasco J-720 dichrograph employing 0.1 to 0.01 cm cells at 25°C, as described previously.[14] The CD of CopG was analyzed employing three different methods, since the estimation of the percentage of secondary structure depends on the method and the choice of standards. The least squares method of Yang et al.[50] was employed to fit the experimental CD from 190 to 240 nm with a linear combination of four standard spectra of α-helix, β-sheet, turn, and random coil components extracted from a dataset of 15 proteins of known 3D structure. The Convex Constraint Analysis (CCA[34]) uses no 3D standards and was employed to extract singular components from a dataset including the 195 to 240 nm CD spectra of 25 proteins plus CopG. Alternatively, the problem CD spectrum was analyzed as a linear combination of a four reference curve set obtained by the CCA of the 25 reference spectra, employing the Lincomb program.[35] Fluorescence spectra of 20 µM CopG were obtained with a Shimadzu RF540 (Shimadzu Corp., Kyoto, Japan) spectrofluorometer with excitation and emission slits of 2 and 5 nm, respectively, employing a 0.5 × 0.5 cm cell at 25°C.

Half-Life Measurements

The intracellular stability of wild-type and variants CopG proteins was analyzed in pulse-chase experiments. *E. coli* BL21(DE3) cells containing the different expression plasmids (each corresponding to one *copG* allele) were grown at 37°C in M9 minimal medium until an OD_{650} of 0.4 was reached. The synthesis of T7 RNA polymerase was then induced by addition of 1 mM IPTG. Incubation continued for 10 more min and rifampicin was added to 0.2 mg ml^{-1} to inhibit the host RNA polymerase. After a further 10 min incubation, a 2-ml aliquot was taken and mixed with 10 µl (100 µCi) [^{35}S]Met (specific activity >1,000 Ci $mmol^{-1}$). Labeling of de novo-synthesized proteins was allowed for 1 min at 37°C and chased by addition of 2 µl of 33.5 mM cold Met (a 700 times excess). Incubation continued and, at different times, 200 µl samples were withdrawn and frozen in liquid nitrogen until all were ready to be processed together. Cells were then collected by centrifugation and disrupted as described.[23] The proteins were separated by Tricine-SDS-polyacrylamide gel electrophoresis on 16% polyacrylamide gels.[41] [^{35}S]-labeled protein bands were detected by autoradiography on Kodak X-Omat films (Eastman Kodak, Rochester, NY). The counts in the bands were directly quantified by means of the storage phosphor technology, with the aid of a Phosphorimager Image Quant equipment (Molecular Dynamics, Sunnyvale, CA).

Model-Building Procedure

The models of the 3D structure of CopG were built from the amino acid sequence of the protein[15,19] by using knowledge-based protein modeling methods. Regions of structural similarity among the reference groups of well-characterized proteins containing ei-

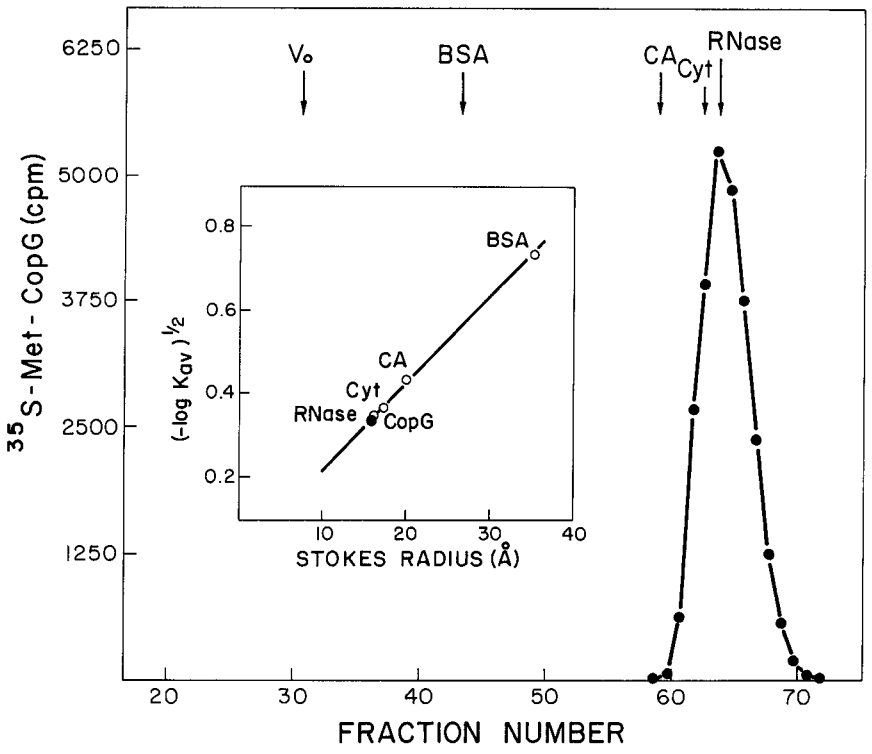


Fig. 2. Gel filtration chromatography and estimation of the Stokes radius of CopG. Ordinates show total cpm due to [^{35}S]Met-labeled CopG protein in each collected fraction. Arrows indicate the elution position of the standard proteins. Insert, the gel filtration data were plotted according to Siegel and Monty.[42] Standard proteins (○), CopG protein (●). BSA, bovine serum albumin; CA, carbonic anhydrase; Cyt, cytochrome c; RNase, ribonuclease A from bovine pancreas.

ther HTH (λ-cI, λ-Cro and Lac proteins), or β-sheet (Arc and MetJ repressors) DNA binding motifs were found through an iterative procedure.

In the case of the putative HTH model of CopG, the β-strand region could not be reliably modeled with these constraints. The CopG-HTH model was built from the high-resolution crystal structures of the λ-Cro repressor[2] (2.2 Å) and the λ-cI repressor[3] (1.8 Å), and from the NMR spectroscopy-determined structure of the Lac repressor.[11] Their coordinates were taken from the Brookhaven Protein Data Bank (PDB codes 2CRO, 1LMB, and 1LCD, respectively). Computations were performed on a DIGITAL MicroVAX 3400 and a MicroVAX 3100 workstation by using the Chem-X software package (Chemical Design, Oxon, UK). The energy calculations were carried out using the CHARMM[8] forcefield.

To build the CopG model based on the ribbon-helix-helix class of proteins, the recognition of an overall fold of CopG was carried out from the resulting threading predictions of THREADER 2.[29] The model of the amino acid sequence of CopG was built from the high-resolution crystal structure of the MetJ repressor[38] (1.8 Å) and from the NMR spectroscopy-determined structure of the Arc repressor.[7] Their coordinates were taken from the Brookhaven Protein Data Bank (PDB codes 1CMB and 1ARR, respectively). Computations were performed on an SGI Power Challenge R10000 by using the BIOSYM software package, Release 95.0 (Molecular Simulations, San Diego, CA).

RESULTS AND DISCUSSION

Physical Properties of CopG

CopG is one of the pMV158-encoded elements that negatively regulates synthesis of the initiator of replication protein, RepB (Fig. 1). Genes *copG* and *repB* are transcriptionally coupled, being RepB the positive effector of the circuit. Both genes are transcribed from promoter P_{cr}, five out of the six bases of its -35 region belonging to a 13-bp symmetric element, which is a part of the CopG target (Fig. 1). DNase I and hydroxyl radical footprints performed on linear DNA have shown that the regions protected by purified CopG protein span about 50 bp.[16]

The primary structure of CopG protein was verified by complete protein sequencing and the protein was chemically synthesized[15,19] (Table I). The Met1 residue is present in the native protein, not being formylated or subjected to processing. In addition, we determined other physical properties of CopG (Table I). Gel electrophoresis analyses, under denaturing conditions, showed that purified CopG migrated as a single band (M_r of about 5,000), which agreed with the data derived from DNA sequence (5118.7 Da). MALDI-TOF mass spectrometric analysis of purified CopG gave a single peak with the value of 5114.7 Da.

Analytical gel filtration of purified CopG (12 μM loading concentration) showed that the native protein eluted as a nearly symmetric single peak, indicating that most of the protein exists as a single

macromolecular component (Fig. 2). Values of K_{av} were calculated from the elution position of each protein (see Materials and Methods) and plotted as $(-\log K_{av})^{1/2}$ against Stokes radii (Fig. 2, insert). CopG eluted at about the same position as RNase A (Fig. 2). The Stokes radius of CopG was calculated, from a simple linear regression, to be 16 Å.

Analytical ultracentrifugation (sedimentation equilibrium and velocity) was performed at the CopG concentration of 12 μM. Sedimentation equilibrium (Fig. 3A) showed that the average molecular mass ($M^c_{w,a}$) was 10,400 ± 400, a value that fits well with the theoretical mass of a CopG dimer (10,237). Similar average molecular masses were determined when CopG concentrations between 25 μM ($M^c_{w,a}$ = 10,800 ± 400) and 780 μM ($M^c_{w,a}$ = 10,750 ± 250) were used. Sedimentation velocity profiles of CopG protein revealed an apparent single boundary. Using SVEDBERG[37] the data were well fitted to a single sedimenting species, with an $s_{20,w}$ value of 1.30 ± 0.10 S (Fig. 3B). There was no improvement in the best fit parameters values if more sedimenting species were considered, a good indication of sample homogeneity.

The Stokes radius of CopG deduced from the relative method of gel filtration (16 Å) together with the absolute average molecular weight calculated from sedimentation equilibrium (10,400), allowed us to calculate that the (f/f_0) frictional ratio of CopG was 1.10. The frictional ratio was also estimated from the absolute values obtained in the analytical ultracentrifuge. This f/f_0 value was 1.08. Therefore, the hydrodynamic behavior of the CopG protein dimer slightly deviates from the one corresponding to a rigid spherical particle.[48] We conclude that the native configuration of CopG is a nearly spherical dimer of identical subunits.

Circular Dichroism of CopG Protein

The product of gene *copG* is predicted to have a relatively simple secondary structure using several programs,[15,22] consisting of one β-strand, two α-helices separated by a turn, and disordered N- and C-terminal ends (Fig. 4A). The experimental CD spectrum of CopG (Fig. 4B) contains a very marked α-helical contribution, evidenced by the maximum at 194 nm, and minima at 209 nm and 222 nm. The spectrum was independent of protein concentration in the range of 20–200 μM CopG, and very similar to 20 μM chemically synthesized CopG (not shown). Deconvolution of the CD spectrum with three different methods (Fig. 4C) gave a consensus average secondary structure content of 50–70% α-helix, 10–35% β-strand and turn, and 15–20% disordered, which is compatible with the predicted secondary structure. In addition, preliminary fluorescence spectroscopy characterization of CopG has indicated that the emission spectrum of the single tyrosine residue is similar to the model compound N-acetyltyrosin-

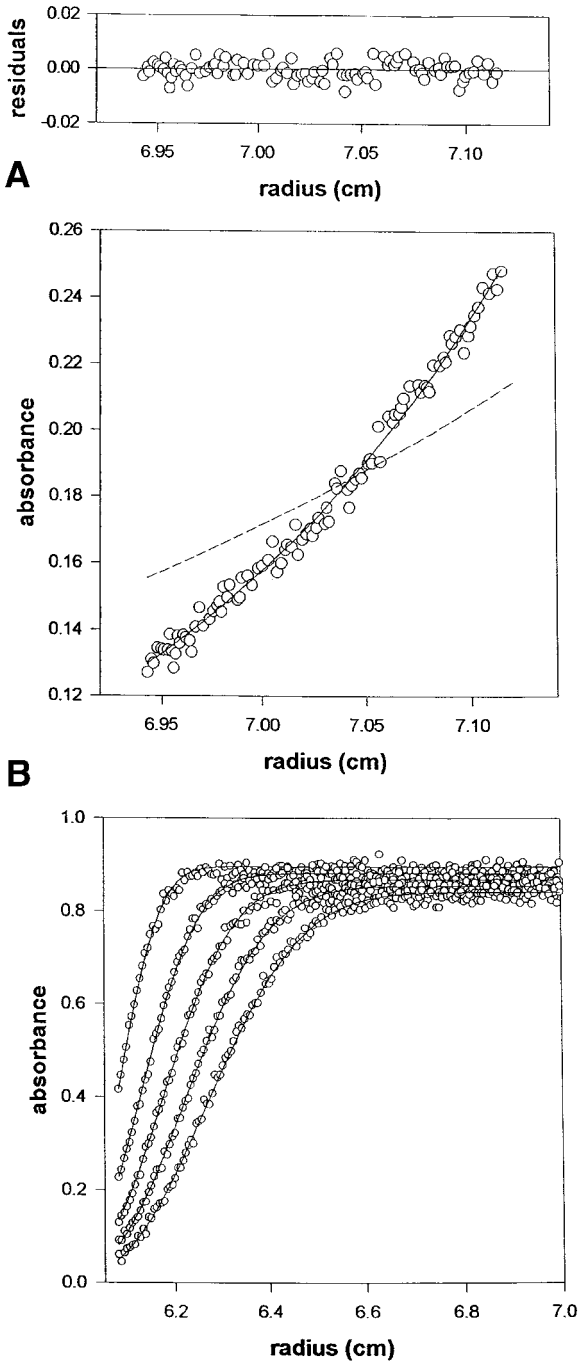
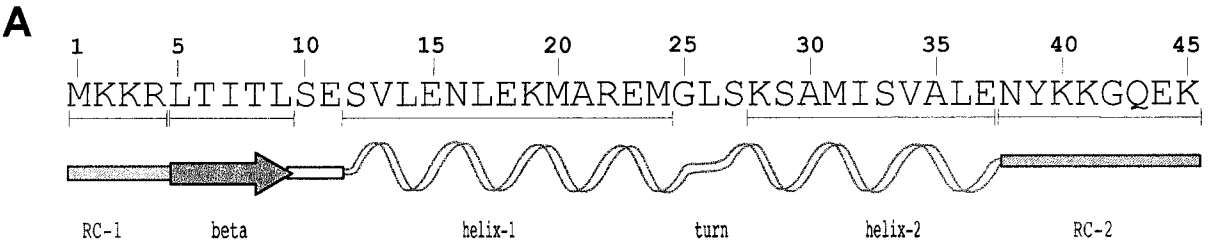
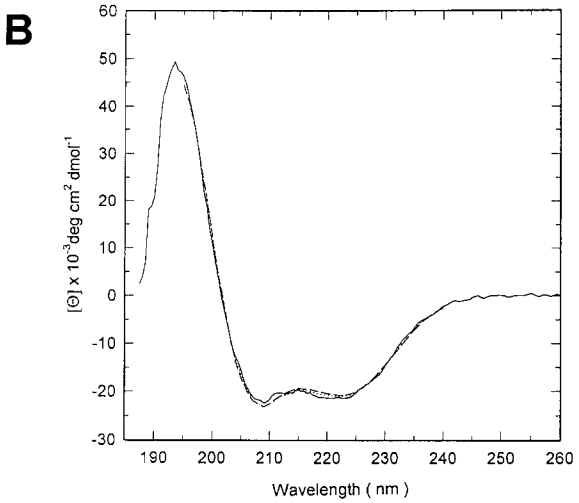


Fig. 3. Analytical ultracentrifugation of CopG. **A:** Sedimentation equilibrium profile of 12 μM CopG taken at 25,000 rpm, 20°C, and at a wavelength of 230 nm. The symbols represent the experimental data. The solid line shows the best fit weight average molecular weight (10,400); the dotted line indicates the theoretical gradient of CopG monomer (5,119). **B:** Sedimentation velocity analysis of 12 μM CopG at 60,000 rpm and 20°C. The symbols represent the measured absorbances at 218 nm, and the lines are the best fits to a model of a single sedimenting species (1.30 S).





C

	α-helix	β-p + turn	β-a	Turn	Unordered
Yang	70	-	-	10	20
CCA	67	1.5	19.5	-	12
Lincomb	47	36	1	-	16
Average	50-70		10-35		15-20

Fig. 4. Secondary structure of CopG. **A:** Prediction of secondary structure of CopG derived from the PROTEAN Computer program (DNASTAR). A scheme of the different regions of secondary structure is shown below the amino acid sequence of CopG (residue number indicated above). RC-1 and RC-2 indicate the disordered regions located at the N- and C-terminal ends, respectively. Beta indicates a β-strand region. **B:** Circular dichroism spectrum of CopG protein (200 μM) at 25°C. Continuous line: experimental spectrum, unsmoothed; dashed line, model spectrum fitted with the program Lincomb; dotted line, spectrum fitted with the program CCA (see Materials and Methods). **C:** Results of the deconvolution of the CD spectrum of CopG.

amide in buffer, suggesting that Tyr39, located at the C-terminal end of the predicted helix-2 (Fig. 4A), is exposed to the solvent (not shown).

Mutations in *copG* Gene and Features of the Products

An important parameter to characterize the structure and function of a protein is its intracellular stability (i.e., its half-life). In this sense, measurement of half-lives of the products of mutated genes is an approach to characterize the contribution of the different regions to the structural properties of the protein. In the case of the N-terminal domain of the lambda repressor, mutations have been described which severely affect intracellular stability of the otherwise long-lived wild-type protein.[39] We have found and characterized two spontaneous point mutations, termed *copG7* (formerly *cop7*[17]), and *copG8*.[1] The *copG7* mutation results in a protein with one amino acid change in the putative α-helix2 (Ala30 → Glu), whereas the *copG8* mutation leads to the change Gly25 → Glu within the turn between the two putative α-helices (Table II). Both mutations increased fivefold the number of copies of plasmids harboring the pMV158 replicon, indicating synthesis of a defective repressor (not shown).

To know which regions within the predicted secondary structure of CopG were important for the global structure of the protein, we designed several mutations in the *copG* gene by introducing 9-bp deletions. The mutations would then result in CopG variants lacking three amino acids in several positions (Table II). In addition, the constructions used for site-directed mutagenesis contained the *copG* gene under the control of phage T7 ϕ10 promoter. Then, overexpression of *copG* genes (wild-type or mutants) would allow us to compare the synthesis and to measure the half-lives of CopG and of CopG variant proteins. Four deletion mutants were constructed (Table II) and they removed the *copG* coding sequences for the following residues: Lys2, Lys3, and Arg4 (*copGΔrc1*); Thr6, Ile7, and Thr8 (*copGΔβ*); Glu18, Lys19, and Met20 (*copGΔα1*), and Lys40, Lys41, and Gly42 (*copGΔrc2*). The products should affect the RC1, the β-strand, the α-helix1, and the RC2 putative regions of CopG, respectively (Fig. 4A). To measure the

TABLE II. Features of CopG and Variant Proteins

Protein	Amino acid sequence	Half life (min)	Binding to target DNA
	RC1 Beta H-1 H-2 RC2		
CopG	MKKRLTITLSESVLENLEKMAREMGLSKSAMISVALENYKKGQEK	>180	+++
CopG7	MKKRLTITLSESVLENLEKMAREMGLSKSEMISVALENYKKGQEK	>180	+
CopG8	MKKRLTITLSESVLENLEKMAREMELSKSAMISVALENYKKGQEK	>180	+
CopGΔRC1	M---LTITLSESVLENLEKMAREMGLSKSAMISVALENYKKGQEK	8	ND
CopGΔβ	MKKRL---LSESVLENLEKMAREMGLSKSAMISVALENYKKGQEK	0.5	ND
CopGΔα1	MKKRLTITLSESVLENL---AREMGLSKSAMISVALENYKKGQEK	0.3	ND
CopGΔRC2	MKKRLTITLSESVLENLEKMAREMGLSKSAMISVALENY---QEK	150	ND

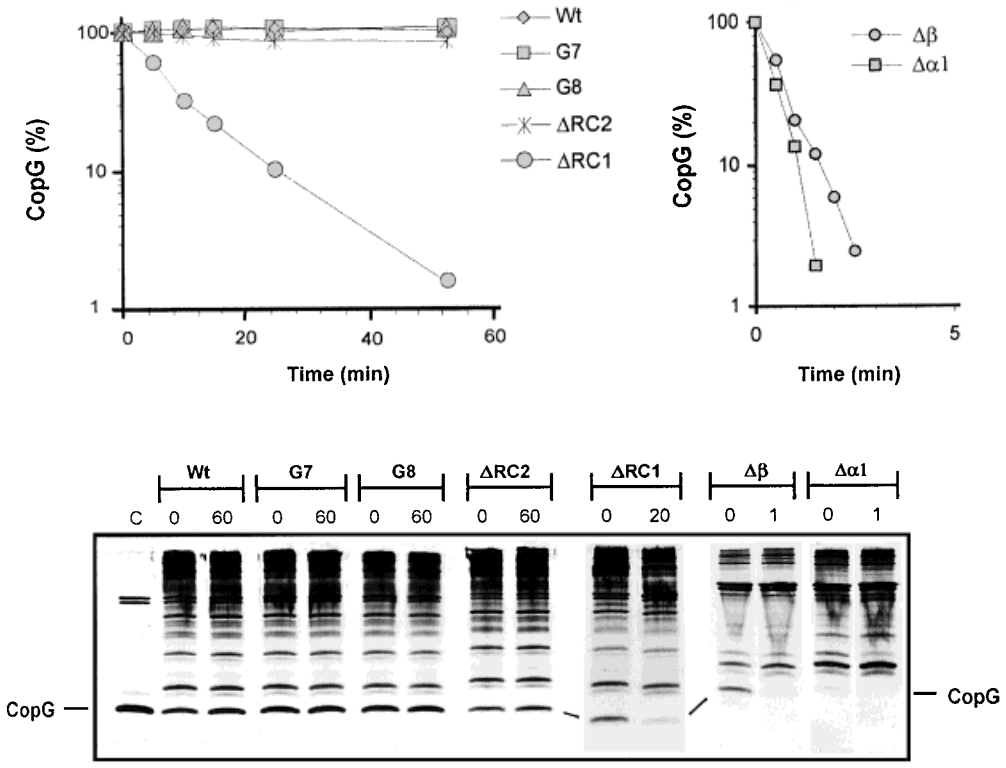


Fig. 5. Half-life of wild-type and variant CopG proteins in *E. coli* cells. Pulse-chase experiments were performed for each pair *E. coli* BL21(DE3)/expression plasmid containing the indicated *copG* allele: Wt, *copG*; G7, *copG7*; G8, *copG8*; ΔRC1, *copGΔrc1*; Δα1, *copG Δα1*; Δβ, *copG*Δβ; ΔRC2, *copGΔrc2*. The amount of radioactivity in the corresponding CopG proteins at different times after chasing was determined and plotted as percentage of radioactivity relative to that in time zero (100%). Plots for highly or moderately stable proteins are shown on the left, whereas those for very unstable proteins are represented on the right. The autoradiogram at bottom shows the decay of the different *copG* products, each from the indicated *copG* allele, at a representative time (1, 20, or 60 min) after chasing, as compared with time zero. As controls (C), in vitro synthesized proteins[20] were employed. The position of the products of different *copG* alleles is indicated.

half-lives of the various *copG* products, *E. coli* cultures harboring plasmids were grown, treated with IPTG and rifampicin, and pulse-chased with labeled [^{35}S]-Met. Proteins were separated on SDS-gels (Fig. 5). Quantification of the labeled products (Fig. 5, Table II) showed that proteins CopG, CopG7, and CopG8 were quite stable, with half-lives of at least 180 min. Band shift assays, performed with CopG7 and CopG8 protein variants indicated that they were impaired in their ability to recognize target DNA, as compared to the wild-type CopG (not shown). Protein CopGΔRC2 was slightly less stable (150 min of half life). Proteins CopGΔRC1 and CopGΔβb were very unstable, with half-lives of 8 and 0.5 min, respectively. Finally, CopGΔα1 was the most unstable protein, with a half-life of about 20 sec. The instabil-

ity of these altered proteins was most likely due to degradation by host-encoded proteases, although the bacterial host employed to overproduce the proteins was defective in the Lon and the OmpT proteases (this latter enzyme cleaving between Lys-Lys, Lys-Arg or Arg-Arg residues[49]). We conclude that the amino acid substitutions in CopG7 and CopG8, and the deletion in the C-terminal region of the protein, can be tolerated without gross changes in the global conformation of the protein, but that the other deletions severely affect the structure of the proteins. We take these results as indicative that protein CopG has a compact structure, not having domains connected through hinges like the Lac repressor protein.[26] The N-terminal region, presumably structured as a β-strand, and the bihelical region of CopG would need each other for stabilization of the global structure of the protein. Since β-strands are stable only when incorporated into a β-sheet, CopG dimerization could be required in order to maintain the appropriate conformation of the global native protein. In the case of the Arc repressor, folding and dimerization are the same process.[9] Appropriate folding of CopG could be achieved by its interaction with host-encoded chaperones. According to the algorithm of Rüdiger et al.[40] for DnaK binding sites, we calculated that the CopG region spanning from Lys2 to Leu14 has an energy value of -10.83. Eighty-two percent of the peptides analyzed by Rüdiger et al.[40] having an energy value ≤-5 were experimentally verified as DnaK binders. Most of these sites were found in β-sheet elements. Thus, the potential β-strand region of CopG has a high probability of containing a DnaK binding region.

CopG7 and CopG8 variant proteins were purified following the same procedure as for the wild-type CopG. Both variant proteins eluted, in gel filtration chromatography, at the position expected for a dimer (not shown). CD spectra of these two proteins indicated that their overall structure was similar to that of the wild-type CopG (not shown). These findings suggest that the amino acid changes in both CopG variants do not result in gross changes in the protein structure, which agrees with the prediction of their secondary structures being very similar to the wild-type CopG (not shown) and with the unchanged intracellular stability. The impaired DNA binding ability of CopG7 and CopG8 does not necessarily imply that amino acids in the putative α-helix2 or turn of the bihelical unit of CopG are involved in specific contacts with the target DNA, since these changes may also alter the optimal conformation of the protein. In addition, protein–protein contacts, which should be relevant for the highly cooperative binding to target DNA observed for CopG, could be affected in both variant proteins. In this sense, it is worth noting that: 1) the dimer–dimer interface in the Arc-operator complex is formed by the sidechains of several amino acid residues located in the turn and in the helix2 of the bihelical unit of the protein; and 2) mutations affecting these amino acids impair DNA binding and cooperativity.[43]

Molecular Modeling of CopG Motifs

The predicted secondary structure of CopG indicates the existence of two relevant putative regions in the protein, namely, the bihelical unit and the N-terminal β-strand. Let us consider which could be the role of these motifs. The region structured as two α-helices would expand from residues Ser12 to Met24 (α-helix1) and Lys28 to Glu37 (α-helix2). The general structure of the region suggested to us that it could constitute an HTH motif which may be involved in DNA recognition.[15] There is a large hydrophobic moment which corresponds almost entirely to α-helix1. A helical wheel projection along the axis of this α-helix1 (Fig. 6A) shows that, indeed, the distribution of hydrophilic and hydrophobic residues would define two subdomains within this region. Thus, the hydrophilic face of CopG would be exposed to the aqueous medium, whereas the hydrophobic face would be confronted with α-helix2, the putative "reading head" of the HTH motif.

Alignment of the primary structure of the putative HTH motif of CopG with the HTH DNA binding motifs of several well-characterized transcriptional repressor proteins[33] (Fig. 7A) showed that the Gly25 of CopG would correspond to the highly conserved Gly residue in the turn (relative position 9 in the HTH motif), whereas the Ala21 residue would be placed at relative position 5, most frequently occupied by a small residue (Ala or Gly). In addition, the putative CopG-HTH motif fulfills other requirements predicted to be important for the stabilization of the arrangement of the two helices in the HTH unit, and for the package of this unit against the rest of the protein[32]: 1) no Pro residues (which would break an α-helix) are found within the putative α-helices (residues in relative positions 1–8 and 12–20 in Fig. 7A); 2) residues in the relative positions 4, 8, 10, and 15 (corresponding to CopG residues Met20, Met24, Leu26, and Met31, respectively) are hydrophobic; and 3) residues Lys19 and Glu23 (corresponding to the relative positions 3 and 7, respectively) are hydrophilic, as is to be expected for external residues with a sidechain fully exposed to the solvent. However, computation of the CopG putative HTH motif (between residues Leu17 and Leu36) by employment of the systematic matrix developed by Dodd and Egan[21] gave a score of 2.0 SD, below the 2.5 SD significance threshold chosen by the authors to assign a protein as likely bearing an HTH DNA binding motif. The score obtained for CopG is close to the 2.2 SD value reported for TrpR, an HTH protein containing atypical residues that allow it to respond to the presence or absence of its co-repressor. Construction of a molecular model of

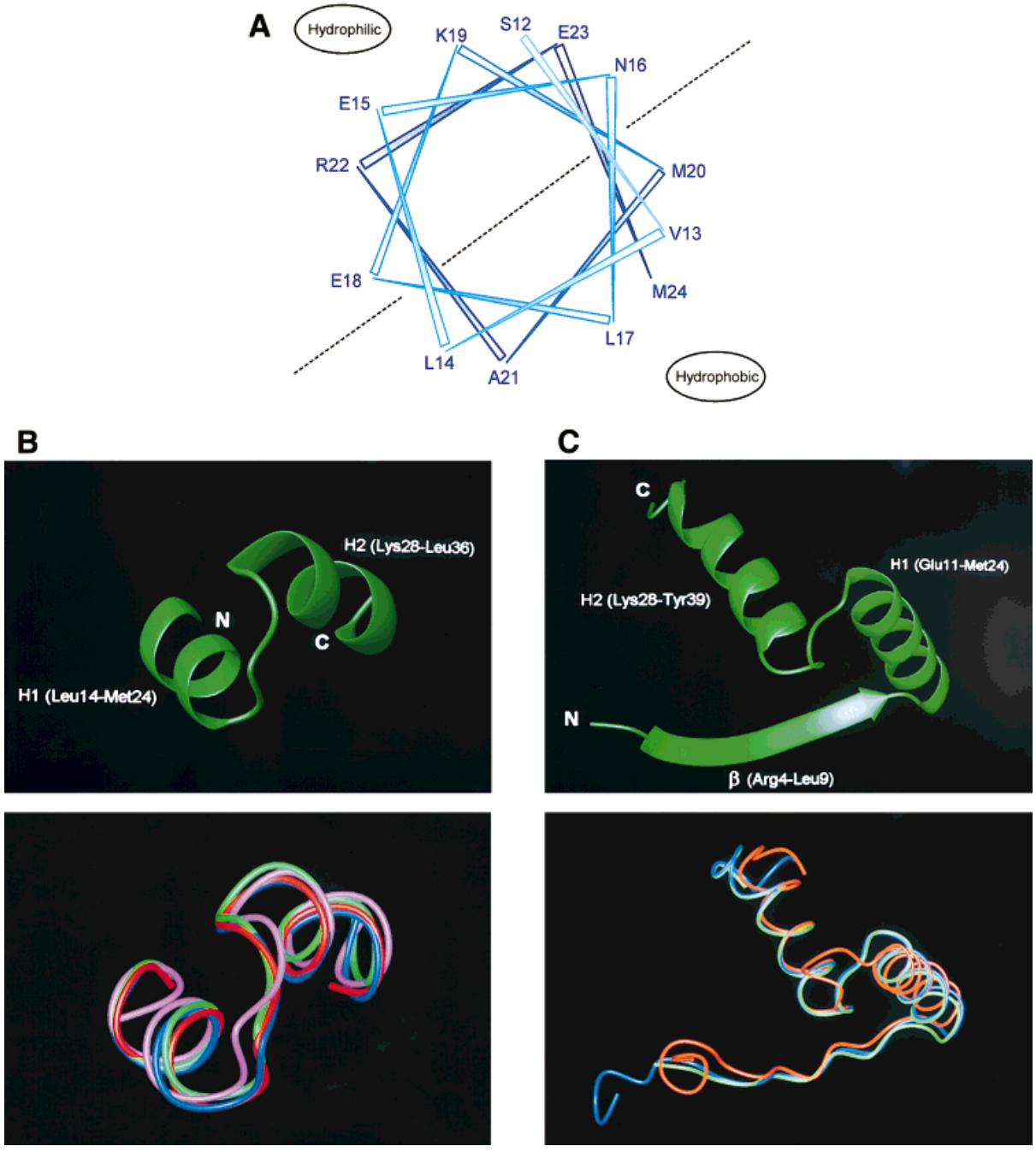


Fig. 6. Putative tertiary structures of CopG. **A:** Helical wheel projection along the axis of the putative α-helix1 of CopG (from Ser12 to Met24). Amino acids residues are plotted every 100° consecutive around the wheel. An imaginary plane (represented by the dashed line in the projection) would divide this helix into a hydrophilic and a hydrophobic moiety, based on the nature of the amino acid residues. Molecular modeling of CopG (**B,C**) show the drawings of the polypeptide backbone for the homology-modeled structures. **B:** A ribbon drawing of the modeled HTH putative CopG motif (upper part), and comparison of this motif (green) with λ-cI (red), λ-Cro (purple), and LacR (blue) repressor proteins (lower part). **C:** A ribbon drawing of the β-strand-helix-helix modeled structure of CopG (upper part), and comparison of this model (green) with Arc (blue) and MetJ (red). Note that both models represent one subunit of the proteins. In the ribbon drawings, residues spanning putatives helices (H1 and H2) and β-strand of CopG are indicated. N and C indicate the N- and C-terminal ends, respectively.

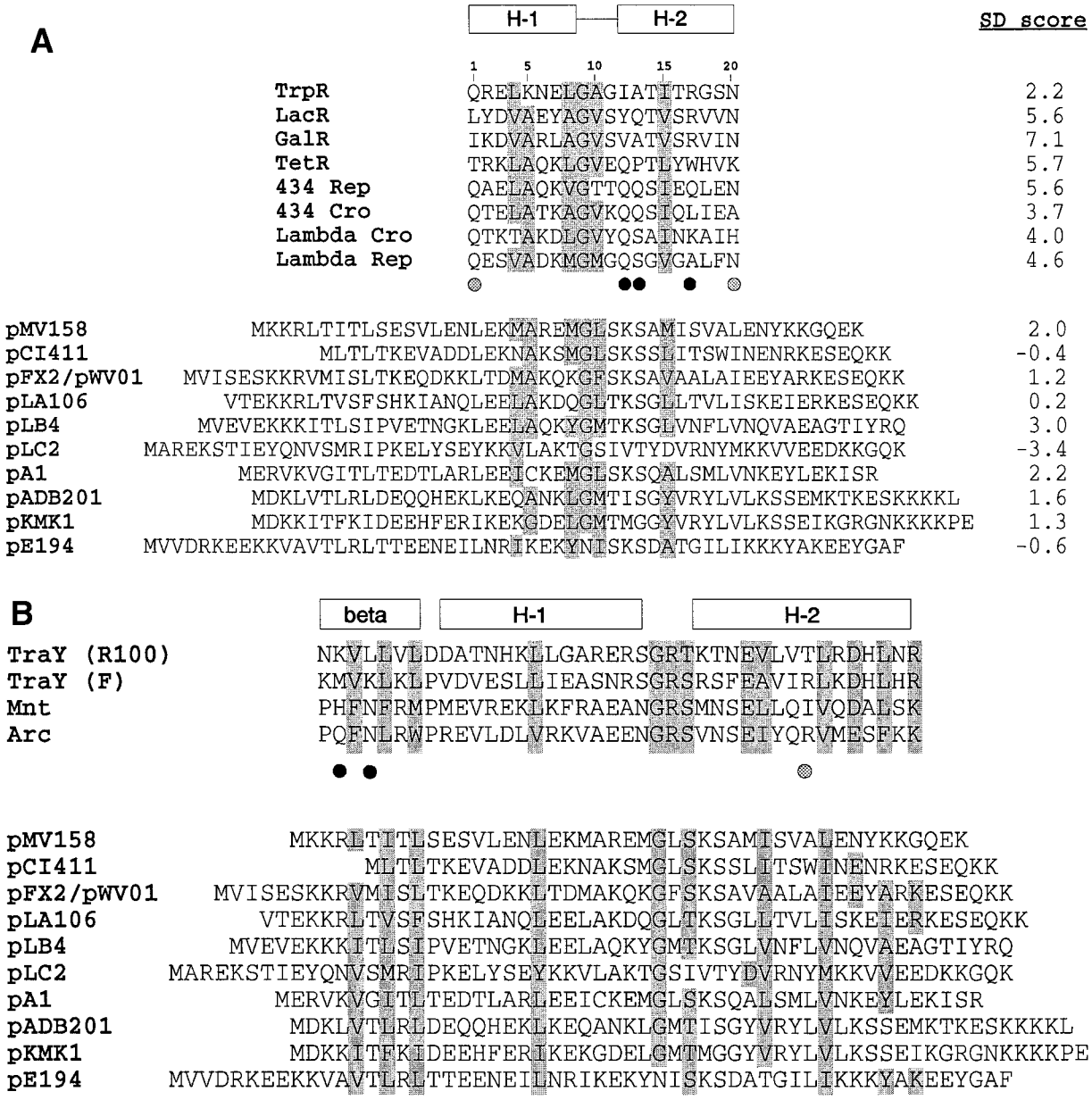


Fig. 7. Alignment of Cop proteins of the pMV158 plasmid family with relevant regions of some reference HTH (**A**) or ribbon-helix-helix (**B**) DNA binding proteins. Positions with generally conserved features in either class of reference proteins are shaded: hydrophobic (A, C, V, I, L, M, F, Y, W), small (G, A), acidic (D, E), or basic (R, K). Positions which interact with the backbone (○) or with bases in the major groove (●) of the DNA in the co-crystal resolved structures of the reference proteins are indicated.[33] In A, the relative positions of residues in the HTH motif are depicted. On the right, the SD score values of the reference proteins[21] and of the Cop proteins (calculated in this work) are shown.

the HTH motif of CopG, based on replacement of λ-Cro residues by those of CopG (not shown) gave satisfactory results for the α-helix1, but not for α-helix2. However, when the molecular model of the putative HTH motif of CopG was based on the construction of a conserved framework from the structures of λ-cI, λ-Cro, and LacR repressor proteins, followed by refinement, we found a good fit between the HTH motifs of these three proteins and the putative motif of CopG (Fig. 6B).

A puzzling observation is that all proteins contacting their cognate DNA by HTH motifs have a third α-helix which is thought to stabilize the protein,[33] whereas CopG does not seem to have this third α-helix. This stabilizing role could be played by the essential N-terminal region of the protein, for which a β-strand from each subunit of a CopG dimer could be postulated. Among prokaryotic repressor proteins with solved structures, the P22 bacteriophage Arc dimer has a structure that fits better with the

predictions of the secondary structure of CopG. Arc belongs to a class of regulatory proteins that use an antiparallel β-sheet to bind its cognate DNA. Arc forms dimers in solution and each subunit contains a β-strand (which is involved in dimerization) and two α-helices.[7] In the Arc dimer, the α-helical regions are packed against the β-sheet, and against each other to stabilize the dimer. Then, the predicted β-strand of CopG (spanning at least from Leu5 to Leu9) may be involved in dimerization and/or DNA contacts. Within the Arc class of proteins, limited sequence similarities among the few known members have made the detection of new members difficult, although searches based upon hydrophobicity patterns identified TraY proteins as belonging to this protein class.[5] Knowledge of the permitted amino acid substitutions was obtained by cassette mutagenesis, which led to Arc variants able to fold into a stable structure.[4] Comparison of the amino acid sequence in the predicted β-strand region of CopG with those of Mnt and Arc indicates obvious conservation without any apparent violation (Fig. 7B). When CopG was modeled on the structures of Arc and MetJ proteins, an excellent fit was observed (Fig. 6C). Consequently, at our present stage of knowledge we cannot discard that CopG belongs to the Arc class of proteins that use an antiparallel β-sheet for DNA binding. Be that as it may, both kinds of regulatory proteins with which CopG compares bind their DNA target through the major groove. In this sense, the pattern of hypersensitive sites to DNase I and the footprints found in CopG-DNA complexes suggested to us that CopG could bind DNA through the major groove (not shown). The resolution of the 3D structure of CopG (alone and complexed with its target DNA) can elucidate which is the DNA binding motif of the protein. These structural analyses have been hampered because of the low yield of purified CopG. Three strategies are being pursued to overcome this problem: 1) use of a new expression vector; 2) a recently developed scaling-up procedure, and 3) longer expression times based on the observed high intracellular stability of CopG.

cop Genes in Plasmids of the pMV158 Family

The pMV158 plasmid family is so far composed of 12 replicons isolated from different bacterial species. All these plasmids show a similar genetic structure at their replication and control regions.[18] Experimental data are available only for pMV158 and pE194, whereas information on Cop proteins from the other plasmids relies on sequence-derived similarities. With the exception of plasmid pHPK255, which seems to lack a *cop* gene, the rest of the replicons may have a *cop-rep* promoter with features similar to those of the pMV158 P_{cr}, so that their *cop* and *rep* genes may be co-transcribed. In the case of pCI411, for which no *cop* gene was reported,[12] we found a possible GTG initiation codon for a Cop protein. These Cop proteins would have a number of residues ranging from 42 (pCI411-Cop) to 56 (pE194-Cop and pLC2-Cop). We aligned the HTH region of various well-characterized repressor proteins with the Cop proteins of plasmids of the pMV158 family (Fig. 7A). No gaps were allowed and alignment was started from the highly conserved Gly residue of the turn. All Cop proteins have this Gly, except Cop-pE194, which would have an Asn residue at this position, as in the phage ϕ105 repressor.[21] The alignment indicates that all Cop proteins could have a putative HTH motif lacking Pro residues in the predicted α-helices. Sequence positions that are generally hydrophobic or small were conserved in several, but not in all, Cop proteins. Plasmid pLC2-Cop exhibited the lowest fit, since only two out of the six most conserved positions were maintained. On the other hand, a low SD score, below the significance threshold (2.5 SD[21]), is obtained for all these Cop proteins, with the exception of the pLB4-Cop, for which a score of 3.0 SD was calculated (Fig. 7A). Proteins scoring from 2.50 SD to 2.99 SD were expected to contain an HTH DNA binding motif with a frequency of about 25%.[21]

Due to the analogies between CopG and the less numerous proteins of the Arc group of β-sheet DNA binding proteins, an alignment of these proteins with Cop proteins of the pMV158 family was performed (Fig. 7B). We found amino acid sequences compatible with the existence of a β-strand in the Cop proteins of the pMV158 family, perhaps with the exception of pCI411-Cop. Using this alignment, it was evident that some other residues located at conserved positions in the ribbon-helix-helix proteins were maintained in most of the Cop proteins (Fig. 7B).

In conclusion, CopG protein of plasmid pMV158 would represent the prototype of a family of DNA binding proteins which seem to share general features with both HTH and β-strand proteins. Since there are no similarities at the amino acid level among the regulatory proteins belonging to either group, assignment of a given protein to an HTH or a β-sheet motif relies on features conserved in specific positions. However, this assignment may prove to be difficult, especially when the protein shares common features with both motifs. Cop proteins of the pMV158 plasmid family, while showing conserved positions with both DNA binding motifs, exhibit a bihelical region with a low SD score (Fig. 7A), as is the case for Arc and Mnt ribbon-helix-helix proteins (-1.9 SD and -1.3 SD, respectively). Since most of the conserved positions in either group of DNA binding motifs fulfill stereochemical requirements in the protein global structure, Cop proteins of the pMV158 family may represent an atypical subgroup of HTH proteins lacking a third α-helix and having, instead, a β-strand. Alternatively, they may belong to the ribbon-helix-helix class of proteins. At present, only the solution of the CopG-DNA tertiary structure will

provide a direct answer to which domains of the protein are involved in binding to DNA.

ACKNOWLEDGMENTS

We thank M.T. Alda for assistance in purifying CopG protein, J. Evangelio for analysis of the CD spectra, B. Rüdiger for providing the program to analyze DnaK binding sites, and A. Prieto for the MALDITOF analysis. We thank A. Hurtado, M. Fontenla, and V. Muñoz for the artwork.

REFERENCES

1. Acebo, P., Alda, M.T., Espinosa, M., del Solar, G. Isolation and characterization of pLS1 plasmid mutants with increased copy numbers. FEMS Microbiol. Lett. 140:85–91, 1996.
2. Anderson, W.F., Ohlendorf, D.H., Takeda, Y., Matthews, B.W. Structure of the Cro repressor from bacteriophage λ and its interaction with DNA. Nature 290:754–758, 1981.
3. Beamer, L.J., Pabo, C.O. Refined 1.8 Å crystal structure of the λ repressor-operator complex. J. Mol. Biol. 227:177–196, 1992.
4. Bowie, J.U., Sauer, R.T. Identifying determinants of folding and activity for a protein of unknown structure. Proc. Natl. Acad. Sci. USA 86:2152–2156, 1989.
5. Bowie, J.U., Sauer, R.T. TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. J. Mol. Biol. 211:5–6, 1990.
6. Brantl, S. The *copR* gene product of plasmid pIP501 acts as a transcriptional repressor at the essential *repR* promoter. Mol. Microbiol. 14:473–483, 1994.
7. Breg, J.N., van Ophesden, J.H.J., Burgering, M.J., Boelens, R., Kaptein, R. Structure of Arc repressor in solution: Evidence for a family of β-sheet DNA-binding proteins. Nature 346:586–589, 1990.
8. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., et al. A program for macromolecular energy, minimization, and dynamics calculation. J. Comp. Chem. 4:187–217, 1983.
9. Brown, B., Sauer, R.T. Assembly of the Arc repressor-operator complex: Cooperative interactions between DNA-bound dimers. Biochemistry 32:1354–1363, 1993.
10. Byeon, W.-H., Weisblum, B. Replication genes of plasmid pE194-*cop* and *repF*: Transcripts and encoded proteins. J. Bacteriol. 172:5892–5900, 1990.
11. Chuprina, V.P., Rullmann, J.A.C., Lamerichs, R.M.J.N., et al. Structure of the complex of *lac* repressor headpiece and an 11 bp half-operator determined by nuclear magnetic resonance spectroscopy and restrained molecular dynamics. J. Mol. Biol. 234:446–462, 1993.
12. Coffey, A., Harrington, A., Kearney, K., Daly, C., Fitzgerald, G. Nucleotide sequence and structural organization of the small broad-host-range plasmid pCI411 from *Leuconostoc lactis*. Microbiology 140:2263–2269, 1994.
13. Cölfen, H., Harding, S.E., Boulter, J.M., Watts, A. Hydrodynamic examination of the dimeric cytoplasmic domain of the human erythrocyte anion transporter, band 3. Biophys. J. 71:1611–1615, 1996.
14. de Pereda, J.M., Leynadier, D., Evangelio, J.A., Chacón, P., Andreu, J.M. Tubulin secondary structure analysis, limited proteolysis sites and homology to FtsZ. Biochemistry 35:14203–14215, 1996.
15. del Solar, G., de la Campa, A.G., Pérez-Martín, J., Choli, T., Espinosa, M. Purification and characterization of RepA, a protein involved in the copy number control of plasmid pLS1. Nucleic Acids Res. 17:2405–2420, 1989.
16. del Solar, G., Pérez-Martín, J., Espinosa, M. Plasmid pLS1-encoded RepA protein regulates transcription from *repAB* promoter by binding to a DNA sequence containing a 13-base pair symmetric element. J. Biol. Chem. 265: 12569–12575, 1990.
17. del Solar, G., Espinosa, M. The copy number of plasmid pLS1 is regulated by two trans-acting plasmid products: The antisense RNA II and the repressor protein, RepA. Mol. Microbiol. 6:83–94, 1992.
18. del Solar, G., Moscoso, M., Espinosa, M. Rolling circle-replicating plasmids from Gram-positive and Gram-negative bacteria: A wall falls. Mol. Microbiol. 8:789–796, 1993.
19. del Solar, G., Albericio, F., Eritja, R., Espinosa, M. Chemical synthesis of a fully active transcriptional repressor protein. Proc. Natl. Acad. Sci USA 91:5178–5182, 1994.
20. del Solar, G., Acebo, P., Espinosa, M. Replication control of plasmid pLS1: The antisense RNA II and the compact *rnaII* region are involved in translational regulation of the initiator RepB synthesis. Mol. Microbiol. 23:95–108, 1997.
21. Dodd, I.B., Egan, J.B. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Res. 18:5019–5026, 1990.
22. Gribskov, M., Burgess, R.R., Devereux, J. PEPLOT, a protein secondary structure analysis program for the UW-GCG sequence analysis software package. Nucleic Acids Res. 14:327–334, 1986.
23. Lacks, S.A., López, P., Greenberg, B., Espinosa, M. Identification and analysis of genes for tetracycline resistance and replication functions in the broad-host-range plasmid pLS1. J. Mol. Biol. 192:753–765, 1986.
24. Laue, T.M., Shah, B.D., Ridgeway, T.M., Pelletier, S.L. Computer-aided interpretation of analytical sedimentation data for proteins. In "Analytical Ultracentrifugation in Biochemistry and Polymer Sciences." Harding, S.E., Rowe, A., Horton, J.C. (eds.). Cambridge: Royal Society of Chemistry, 1992:90–125.
25. Lee, J.C., Timasheff, S.N. The calculation of partial specific volumes of proteins in 6M guanidine hydrochloride. Methods Enzymol. 61:49–57, 1979.
26. Lewis, M., Chang, G., Horton, N.C., et al. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. Science 271:1247–1254, 1996.
27. Light, J., Molin, S. The sites of action of the two copy number control functions of plasmid R1. Mol. Gen. Genet. 187:486–493, 1982.
28. Maeser, S., Scholz, P., Otto, S., Scherzinger, E. Gene F of plasmid RSF1010 codes for a low-molecular-weight repressor protein that autoregulates expression of the *repAC* operon. Nucleic Acids Res. 18:6215–6222, 1990.
29. Miller, R.T., Jones, D.T., Thornton, J.M. Protein fold recognition by sequence threading—Tools and assessment techniques. FASEB J. 10:171–178, 1996.
30. Minton, A.P. Conservation of signal: A new algorithm for the elimination of the reference concentration as an independently variable parameter in the analysis of sedimentation equilibrium. In "Modern Analytical Ultracentrifugation." Schuster, T.M., Laue, T.M. (eds.). Boston: Birckhouser, 1994:81–93.
31. Müller, A.K., Rojo, F., Alonso, J.C. The level of the pUB110 replication initiator protein is autoregulated, which provides an additional control for plasmid copy number. Nucleic Acids Res. 23:1894–1900, 1995.
32. Ohlendorf, D.H., Anderson, W.F., Matthews, B.W. Many gene-regulatory proteins appear to have a similar α-helical fold that binds DNA and evolved from a common precursor. J. Mol. Evol. 19:109–114, 1983.
33. Pabo, C.O., Sauer, R.T. Transcription factors: Structural families and principles of DNA recognition. Annu. Rev. Biochem. 61:1053–1095, 1992.
34. Perczel, A., Hollósi, M., Tusnády, G., Fasman, G. D. Convex constraint analysis: A natural deconvolution of circular dichroism curves of proteins. Protein Eng. 4:669–679, 1991.
35. Perczel, A., Park, K., Fasman, G.D. Deconvolution of the circular dichroism spectra of proteins: The circular dichroism spectra of the antiparallel β-sheet in proteins. Proteins 13:57–69, 1992.
36. Pessen, H., Kumosinsky, T.F. Measurement of protein hydration by various techniques. Methods Enzymol. 117: 219–255, 1985.
37. Philo, J. An improved function for fitting sedimentation

velocity data for low-molecular weight solutes. Biophys. J. 72:435–444, 1997.
38. Rafferty, J.B., Somers, W.S., Saint-Girons, I., Phillips, S.E.V. Three dimensional crystal structures of *Escherichia coli* Met repressor with and without corepressor. Nature 341:705–710, 1989.
39. Reidhaar-Olson, J.F., Parsell, D.A., Sauer, R.T. An essential proline in λ repressor is required for resistance to intracellular proteolysis. Biochemistry 29:7563–7571, 1990.
40. Rüdiger, S., Germeroth, L., Schneider-Mergener, J., Bukau, B. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. EMBO J. 16:1501–1507, 1997.
41. Schagger, H., von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166:368–379, 1987.
42. Siegel, L.M., Monty, K.J. Determination of molecular weights and frictional ratios of proteins in impure systems by the use of gel filtration and density gradient centrifugation. Applications to crude preparations of sulfite and hydroxylamine reductases. Biochem. Biophys. Acta 112: 346–362, 1966.
43. Smith, T.L., Sauer, R.T. P22 Arc repressor: Role of cooperativity in repression and binding to operators with altered half-site spacing. J. Mol. Biol. 249:729–742, 1995.
44. Studier, F.W., Rosenberg, A.H., Dunn, J.J., Dubendorff, J.W. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. 185:60–89, 1990.
45. Thomas, C.M. Recent studies on the control of plasmid replication. Biochem. Biophys. Acta 949:253–263, 1988.
46. van Holde, K.E. "Physical Biochemistry." Englewood Cliffs, NJ: Prentice-Hall, 1985.
47. Wagner, E.G.H., Simons, R.W. Antisense RNA control in bacteria, phages, and plasmids. Annu. Rev. Microbiol. 48:713–742, 1994.
48. Waxman, E., Laws, W.R., Laue, T.M., Nemerson, Y., Ross, J.B.A. Human factor VIIa and its complex with soluble tissue factor: Evaluation of asymmetry and conformational dynamics by ultracentrifugation and fluorescence anisotropy decay methods. Biochemistry 32:3005–3012, 1993.
49. White, C.B., Chen, Q., Kenyon, G.L., Babbitt, P.C. A novel activity of OmpT. Proteolysis under extreme denaturing conditions. J. Biol. Chem. 270:12990–12994, 1995.
50. Yang, J.T., Chuen-Shang, C.W., Martinez, H.M. Calculation of protein conformation from circular dichroism. Methods Enzymol. 130:208–269, 1986.