Transcriptional repressor CopR: Amino acids involved in forming the dimeric interface

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · JUNE 2000

Impact Factor: 2.63 · DOI: 10.1002/(SICI)1097-0134(20000601)39:4<408::AID-PROT130>3.0.CO;2-0

PROTEINS: Structure, Function, and Genetics 39:408–416 (2000)

Transcriptional Repressor CopR: Amino Acids Involved in Forming the Dimeric Interface

Katrin Steinmetzer,[1] Alexander Hillisch,[2] Joachim Behlke,[3] and Sabine Brantl[1]*
[1]*Institut für Molekularbiologie, Friedrich-Schiller-Universität Jena, Jena, Germany*
[2]*Institut für Molekulare Biotechnologie, Jena, Germany*
[3]*Max-Delbrück-Centrum für Molekulare Medizin, Berlin-Buch, Germany*

ABSTRACT **Plasmid pIP501 encoded transcriptional repressor CopR is one of the two regulators of plasmid copy number. It acts as a transcriptional repressor at the essential *repR* promoter. Furthermore, CopR prevents convergent transcription from the *repR* and the antisense promoter, thereby indirectly increasing the amount of antisense-RNA, the second regulatory component. CopR binds as a dimer to a nearly palindromic operator with the consensus sequence 5′CGTG. Previously, a CopR structural model was built and used to identify amino acids involved in DNA binding. These data showed that CopR is a HTH protein belonging to the lambda repressor superfamily and allowed the identification of two amino acids involved in specific DNA recognition. Here, we describe site-directed mutagenesis in combination with EMSA, dimerization studies using sedimentation equilibrium, and CD measurements to verify the model predictions concerning amino acids involved in dimerization. With this approach, the dimeric interface could be located between amino acids I44 and L62. F5 located at the N-terminus is additionally required for proper folding, and could, therefore, not be unequivocally assigned to the dimeric interface. CD measurements at protein concentrations well below K_{Dimer} revealed that the monomer of CopR is folded. Proteins 2000;39:408–416.** © 2000 Wiley-Liss, Inc.

Key words: CopR; transcriptional repressor; dimeric interface; plasmid replication control; CD measurements; site-directed mutagenesis

INTRODUCTION

Protein-DNA recognition and protein dimerization have been studied in a variety of systems, both prokaryotic and eukaryotic. Whereas extensive studies were carried out to characterize transcriptional repressors controlling lysis/lysogeny of bacteriophages, e.g., λ (cI and Cro) or P22 (Arc, c2), and bacterial metabolism, e.g., Lac and Trp repressor, little information is available on transcriptional repressors involved in plasmid copy number control. The only two well-characterized transcriptional repressors are plasmid pLS1 encoded CopG,[1, 2] the crystal structure of which has been published recently,[3] and CopR[4] encoded by the streptococcal plasmid pIP501. The 10.4 kD CopR protein is one of the two regulators of plasmid copy number. It has a dual function: it acts as transcriptional repressor at the essential *repR* promoter pII by binding to a 44 bp inverted repeat upstream of and overlapping pII.[5] Furthermore, it prevents convergent transcription from pII and pIII (the antisense promoter), thereby indirectly increasing transcription initiation from pIII.[6] Two almost identical Cop proteins (95% sequence similarity) with the same functions are encoded by the related streptococcal plasmids pAMβ1 (CopF[7]) and pSM19035 (CopS[8]). CopR so far is the best-characterized Cop protein of this family. It binds asymmetrically at two consecutive sites (I and II) within the major groove of the DNA that share the sequence motif 5′CGTG. CopR can dimerize in solution.[9] It binds only as a dimer, and not as a monomer to its operator. The K_D values for CopR dimers and for the dimer-DNA complex were determined to be $1.44(\pm 0.5) \times 10^{-6}$ M and $4(\pm 1.3) \times 10^{-10}$ M, respectively.[10] The intracellular concentration of CopR in logarithmically grown *B. subtilis* cells was calculated to be $20–30 \times 10^{-6}$ M, indicating that CopR binds also as preformed dimer in vivo.[10] Previously, we constructed a structural model comprising the first 63 amino acids of CopR. This model was based on the fairly low (14%) sequence similarity to P22 c2 repressor. Furthermore, a model of the CopR-DNA complex based on experimental footprinting data, the CopR homology model and the crystal structure of the 434 c1 repressor-DNA complex was built.[11] Site-directed mutagenesis was used to verify amino acids predicted by the model to be involved in specific and nonspecific DNA contacts. This approach allowed to localize the predicted HTH motif between amino acids 18 and 37 and to determine two amino acids within the recognition helix that make specific contacts to the DNA, R29 and R34. The experimental data supported the model of CopR as a HTH protein belonging to the lambda repressor superfamily.[11]

Here, we use a similar approach to verify the model predictions concerning amino acids that form the dimeric interface. Site-directed mutagenesis was used to substitute amino acids that appear at the putative dimeric

Grant sponsor: Deutsche Forschungsgemeinschaft; Grant number: Br1552/4-1.

Alexander Hillisch's present address is EnTec GmbH, Adolf-Reichwein-Str. 20, 07745 Jena. E-mail: hillisch@imb-jena.de

*Correspondence to: Sabine Brantl, Institut für Molekularbiologie, Friedrich-Schiller-Universität Jena, Winzerlaer Str. 10, Jena, D-07745, Germany. E-mail: Sabine.Brantl@rz.uni-jena.de

Received 29 November 1999; Accepted 4 February 2000

interface individually or in combinations. Mutated proteins were purified and analyzed in analytical ultracentrifugation and crosslinking for their ability to dimerize, in EMSA for their ability to bind to the operator DNA, and in vivo for their ability to down-regulate pIP501 copy number. Furthermore, CD-measurements were performed to detect structural changes caused by the mutations. The results allow narrowing down the dimeric interface to a region between amino acids I44 and L62. Additionally, amino acids E2 and F5 could be part of the dimeric interface. The latter of them is, however, also involved in forming the hydrophobic core. Furthermore, based on CD-measurements with protein concentrations above and below K_{Dimer}, we confirm that dimerization of CopR occurs via a folded monomeric intermediate.

MATERIALS AND METHODS

DNA Preparation and Manipulation

Plasmid DNA was isolated from *B. subtilis* as reported previously.[4] DNA manipulations (restriction enzyme cleavage, ligation, etc.) were carried out at the conditions specified by the manufacturer or according to standard protocols.[12] Taq DNA polymerase was purchased from Perkin-Elmer. Sequencing reactions were performed according to Sanger et al.[13] using Sequenase from AmershamPharmaciaBiotech.

Construction of *E. Coli/B. Subtilis* Shuttle Vectors Containing Mutated *Copr* Genes

Plasmids pDim1 and pDim2 were constructed in the following way: Firstly, two PCR fragments were made on plasmid pCOP1B1 using oligonucleotide 486-34 (5′ TCT AGA GGA TCC GAA TTC GTT GTA AAA ATT GGG G) and oligonucleotides SB46 (5′ TGG GAT TAA ATC) or SB48 (5′ TGG CTC ATT TGG GAT TAA AGC CGC TGC TAG CGT TGA GTT AGT TA), respectively, or the reverse sequencing primer and either oligonucleotide SB45 (5′ATT GTA AAG TTA ACT AAC GCA GCG GCA GTA GTG GAT TTA ATC CCA) or oligonucleotide SB47 (5′TAA CTA ACT CAA CGC TAG CAG CGG CTT TAA TCC CAA ATG AGC CA). Secondly, a third PCR reaction was performed with the PCR fragments 1 and 2 for the corresponding mutation and oligonucleotides 486-30 and the reverse sequencing primer. The resulting 54-bp fragment was cleaved with EcoRI and inserted into the EcoRI site of plasmid pPR1E derived from pPR1[4] carrying an erythromycin instead of a phleomycin resistance gene. The mutations and the fragment orientation were confirmed by sequencing.

The same approach as for pDim1 and pDim2 was used for all other mutant *copR* genes (mutagenic primers 1 [pr1] and 2 [pr2] are shown in brackets) resulting in plasmids:
pDim3 (pr1: 5′GGA AAA TCA GAT CCA ACC GAC GCA ACA CTA GAA CAA ATT G)

(pr2: 5′CAATTTGTTCTAGTGTTGCGTCGGTTGGATC-TGATTTTCC)

pDim5 (pr1: 5′AAGTTAACTAACTCAACGGATAGTGT-GGATGATATCC-CAAATGAGCCAACA)

(pr2: 5′ TGTTGGCTCATTTGGGATATCATCCACACTA-TCCTTGA-GTTAGTTAACTT)

pDim6 (pr1: 5′AACTTGGAACTAGCAGTTAGAGAA-AGCTTA)

(pr2: 5′ TAAGCTTTCTCTAACTGCTAGTTCCAAGTT)

pDim7 (pr1: 5′ GATCCAACCGACGCAACAGCAGAA-CAAATTGTAAAG)

(pr2: 5′ CTTTACAATTTGTTCTGCTGTTGCGTCGGTT-GGATC)

Plasmid pDim4 was constructed using plasmid pDim2 as template and the mutagenic primers pr1 and pr2 designed for pDim3. Thus, pDim4 carries a combination of the mutations present in pDim2 and pDim3.

Construction of *E. Coli* Vectors for Overexpression of Mutated *Copr* Genes

In all cases, a single PCR step was used to amplify the mutated *copR* genes from the corresponding *E. coli/B. subtilis* shuttle vectors. All amplified fragments were inserted into expression vector pQE9 (Quiagen). In this way, all mutated *copR* genes contain 11 additional 5′ codons encoding: Met-Arg-Gly-Ser-His_6-Gly-Ser fused to the second codon of *copR* (Glu). In the case of pQED1 and pQED2, primers B618-30 and SB52 (5′ GAA TTC GGA TCC AAA AAA ATA AGC AAT GAT TTC G) were used, the PCR fragments were cleaved with BamHI and inserted into the unique BamHI site of pQE9. In the cases of pQED3, pQED4, pQED5, pQED6, and pQED7, PCR primers B618-30 (5′GAA TTC GGA TCC GAA CTA GCA TTT AGA GAA) and 951-30 (5′ GAA TTC CTG CAG TCA CAC GAA ATC ATT GCT) were used, the PCR fragments were cleaved with *BamHI* and *PstI* and inserted into pQE9 digested with the same pair of enzymes.

Determination of Copy Numbers in *B. subtilis*

The copy numbers of plasmids containing mutated *copR* genes in *B. subtilis* DB104 were determined as described before.[4]

Preparation of Labelled Wild-Type and Mutated CopR Targets

The preparation of CopR targets was performed as described before.[9] The following synthetic oligonucleotides were used: ks1, 5′-GGG GAA AAG CAA TGA TTT CGT GTG AAT AAT GCA CGA AAT CAT TGC TTA TTT TTT TAA GGG G and ks2, 5′-CCC CTT AAA AAA ATA AGC AAT GAT TTC GTG CAT TAT TCA CAC GAA ATC ATT GCT TTT CCC C.

Preparation of Crude Cell Extracts, Purification of CopR Proteins, and Determination of Protein Concentration

B. subtilis strain DB104 [pCOP7][4] expressing the wild-type *copR* gene was grown on TY with 1 μg/ml phleomycin. In the case of all other mutated *copR* genes, DB104 strains with the corresponding plasmids were grown on TY with 5 μg/ml erythromycin. Crude cell extracts from cultures grown for 6 h (late logarithmic phase) were prepared by sonication as described.[5] The supernatant obtained by centrifugation of the sonicated cells was stored at −20°C in 50% glycerol.

Overexpression and purification of the proteins from *E. coli* were performed as described before.[10]

The protein concentrations were determined by Bradford assay based on a calibration curve measured with His_6-CopR. An aliquot of the protein preparation used to obtain the calibration curve was subjected to amino acid hydrolysis to determine the concentration. For some of the proteins, the concentration was also determined spectrophotometrically in the presence of 6 M guanidine-HCl using the extinction coefficient of 1,280 M^{-1} cm^{-1} at 280 nm. Comparison of the values obtained by spectrophotometry with the Bradford-based values showed a fairly good agreement within an error range of about 10% for the proteins.

CopR-DNA Binding Reaction and EMSA

Binding reactions were performed in a final volume of 20 μl containing 1 nM labelled DNA and 20-7000 nM protein. After incubation at 30°C for 30 min in electrophoresis buffer (0.5 × TBE, 75 mM NaCl, and 0.025 g/l herring sperm DNA as competitor) aliquots of the reaction mixtures were separated on 8% native polyacrylamide gels run at 4°C for 3.5 h (4 V/cm) in 0.5× TBE and 75 mM NaCl. To ensure that the reaction reached equilibrium, aliquots of the mixtures were analyzed on polyacrylamide gels after several hours incubation, and no changes in the results were observed. Gels containing labelled DNA fragments were visualized and quantitated on a Fuji PhosphorImager.

Crosslinking

Protein-crosslinking with HPLC-purified proteins was performed in 50 mM phosphate buffer pH 7.9 containing 150 mM NaCl and 0.0014% glutardialdehyde. The final protein concentration in the reaction mixture was 9 μM**.** Protein, buffer, and glutardialdehyde were mixed on ice and incubated at 7°C for 1 h. Increasing reaction time up to 4 h or increasing glutardialdehyde concentration up to 0.035% did not increase the amount of crosslinked protein. Respective control experiments with α-lactalbumin (which exists exclusively as a monomer) at a glutardialdehyde concentration of 0.0014% and incubation at 30°C for 1 h yielded dimers, trimers and higher oligomers. In contrast, after incubation at 7°C, no such oligomers were observed (data not shown). Crosslinking of CopR at 30°C resulted in a higher amount of crosslinking products than at 7°C, but there was still a significant amount of dimers observable at the lower temperature, which was sufficient for our analyses. Based on these observations we were confident that the crosslinked dimers found under the conditions described above reflect only preformed dimers and not randomly crosslinked monomers. A glutardialdehyde-free control was always coincubated. After addition of loading dye containing 2% SDS, the samples were heated at 95°C for 5 min, cooled on ice and loaded onto a 17.5% SDS polyacrylamide gel. Protein bands were visualized by Coomassie staining till saturation and subsequent destaining. For quantification of the crosslinked products the gels were digitalized with a Stratagene EAGLE EYE™ and analyzed with TINA-Pcbas 2.0 software. Each experiment was carried out with 4–5 independent parallels.

Analytical Ultracentrifugation

The molecular mass of CopR was analyzed by means of an analytical ultracentrifuge XL-A (Beckman, Palo Alto, CA) using the sedimentation equilibrium technique. One hundred microliters of CopR samples dissolved in 0.5 × TBE buffer pH 8.0 containing 75 mM NaCl were filled in 6-channel cells and centrifuged for 2 h at 30,000 rpm (over speed) followed by about 20 h at 26,000 rpm (equilibrium speed). The radial absorbance distribution curves at sedimentation equilibrium were recorded at 225, 230, and 235 nm. The three different curves were fitted simultaneously as described previously.[10] The concentration dependent apparent molecular mass values (M_W) were analyzed according to a monomer-dimer equilibrium. An M_1 value of 11.900 Da for His_6-CopR was used.

Circular Dichroism Measurements and Guanidine-HCl Denaturation

Sample concentrations of 0.1 to 0.5 g/l in 50 mM phosphate buffer containing 150 mM NaCl and 50% glycerol were used. The CD spectra were measured at room temperature in the range from 192 to 260 nm with a JASCO model 710 spectropolarimeter at a scan speed of 50 nm/minute and 1 nm resolution. The path length of the cells used was 0.1 mm. The spectra were recorded as an average of 10 to 20 scans. The appropriate buffer baseline spectra were substracted from the protein spectra. To calculate the mean residue ellipticity the residue concentration used was obtained by multiplying the molar protein concentrations with the number of residues (103 aa for the His-tagged proteins CopR and the Dim-mutants).

Guanidine-HCl-induced unfolding of the proteins was monitored by CD changes at 228 nm. Protein samples were incubated at the various guanidine-HCl concentrations for at least 3 h at room temperature to reach equilibrium. Control measurements after 24 h incubation showed no changes in CD signals compared to samples incubated for 3 h. The values were obtained after an accumulation between 5 and 10 and corrected for the buffer signal. The apparent fraction of unfolded protein for each guanidine-HCl concentration was calculated from the CD-signal using following equation: $F_{app} = (Y_{obs}-Y_{nat})/(Y_{unf}-Y_{nat})$ where Y_{obs} is the measured CD-signal, and Y_{nat} and Y_{unf} were estimated from the extrapolated pre- and post-transitional baselines.

RESULTS

Three-Dimensional Model of CopR Predicts Amino Acids Involved in Dimerization

To aid defining amino acids of the CopR repressor involved in DNA-binding and in forming the dimeric interface, we calculated a 3-dimensional model of the first 63 amino acids of CopR based on the structural homology to c2 repressor of P22 and the 434 c1 repressor.[11] The predicted localization of the DNA binding motif and of amino acids involved in sequence specific and nonspecific recognition of the CopR operator could be confirmed experimentally.[11] Here, we use the model

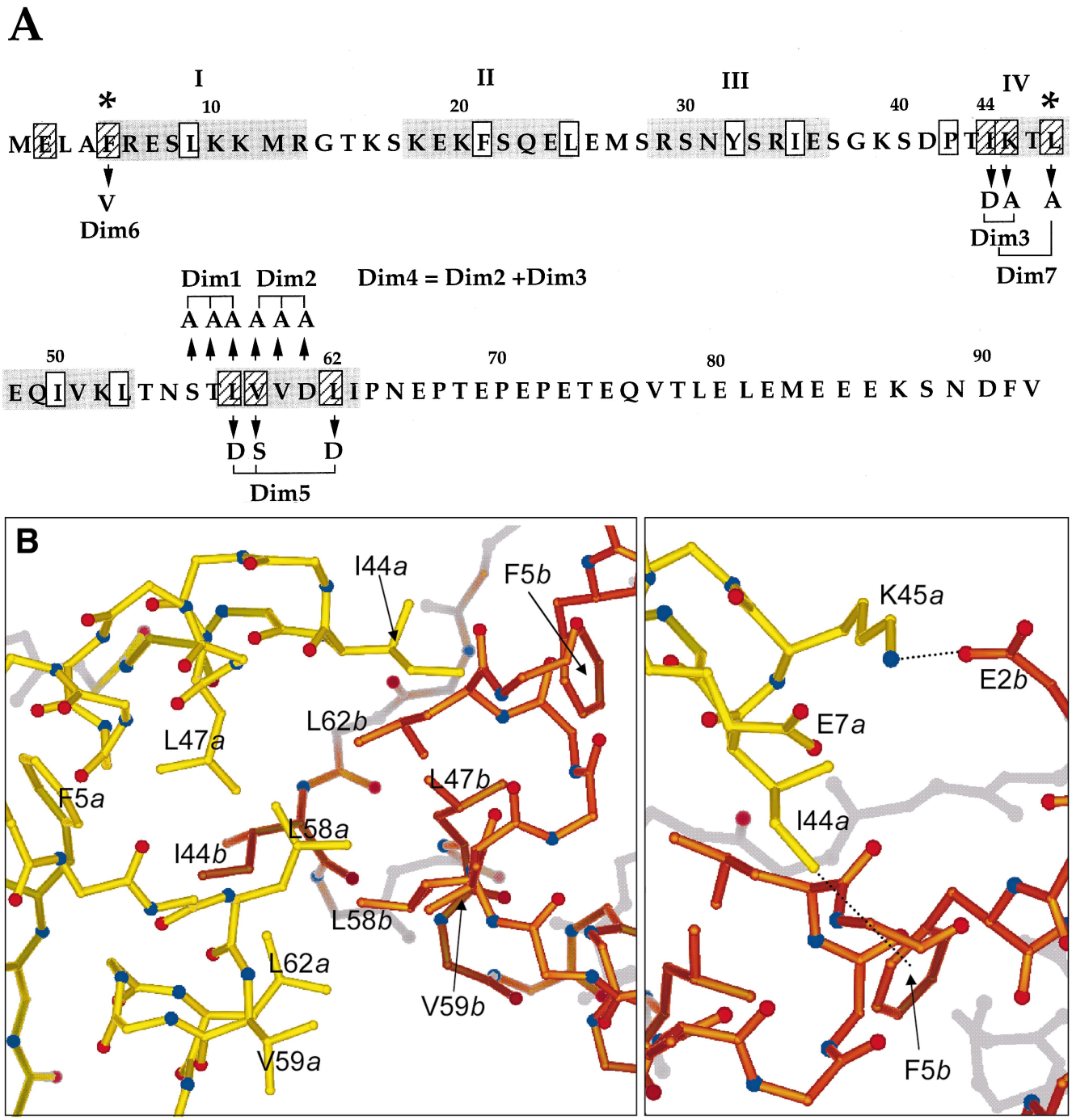

Fig. 1. Amino acids predicted to be located at the dimeric interface. **A:** Primary structure of CopR. The amino acids predicted to be located at the dimeric interface are shown as hatched boxes, amino acids forming the hydrophobic core are shown as white boxes. Asterisks label amino acids involved in dimerization and hydrophobic core formation. Mutated protein Dim4 combines the mutations of Dim2 and Dim3. **B:** Model of the dimeric interface. yellow: amino acids of monomer A; orange: amino acids of monomer B (designated *a* and *b*, respectively). For the sake of simplicity, only main chain atoms of the backbone of monomers A and B and, additionally, the side chains of amino acids predicted to be involved in forming the dimeric interface are shown. Dotted lane, intermolecular contacts.

to identify amino acids that contribute to the formation of the dimeric interface.

Following observations restrict the localization of amino acids involved in dimerization (1) the deletion of up to 27 C-terminal amino acids (CopR mutant CopRΔ27) does not influence significantly the ability of CopR to dimerize (data not shown) and (2) the amino acids between positions 18 and 37 form the DNA-binding helix-turn-helix motif.[11] Thus, the amino acids required for dimerization are most probably located at the N-

TABLE I. In Vivo Activity and Biochemical Properties of Wild-Type CopR and the Mutated Proteins[†]

Protein	In vivo activity (%)	K_{Dimer} (μM)	Relative amount of crosslinked dimer (%)	Activity of crude extracts from *B. subtilis* in EMSA (%)	[Guanidine-HCl] (M) 50% denatur.	% dimer in CD-meas. (protein conc. = 10 μM)
CopR	100	1.49 ± 0.49	16.2 ± 1.6	100	1.83	78
Dim1		(17)[a]	10.8 ± 0.9	15	2.0	40[b]
Dim2	inactive	(22)[a]	9.4 ± 0.8	n.d.	n.d.	n.d.
Dim3		10 ± 2.7	13.1 ± 1.9	16	1.45	50
Dim4	in vivo[c]	41 ± 3.9	3.9 ± 1.0	13	1.26	26
Dim5		70 ± 8.9	n.d.	13	1.36	18
Dim6		430 ± 178 mM	n.d.	0	1.76	0
Dim7		100 ± 17	n.d.	10	1.26	14

[†]n.d., not determined.
[a]Estimated from the crosslinking experiment.
[b]Based on K_D for protein-dimers estimated from crosslinking experiment.
[c]Inactivity in vivo observed for all Dim mutants results in a plasmid copy number of 50–100, comparable to the pIP501 derivative pPR1 lacking the *copR* gene.

terminal region of CopR and between residues 38 and 65. The C-terminal amino acids of CopRΔ27, P64 and N65, are not included in our model and, thus, nothing is known yet about their function.

As can be seen in Figure 1A, the amino acids predicted to be involved in formation of the dimeric interface are located mainly between positions 41 and 62, i.e., within the expected region between amino acid residues 38 and 65 (see above). Additionally, the N-terminal region, especially E2 and F5, seems to be important for intersubunit contacts. Both is consistent with the above-described observations. Most of the intersubunit contacts at the dimeric interface are based on hydrophobic interactions between the side chains of I44, L47, L58, V59, and L62 of both monomers. As an example, L58 and L62 form contacts to L58 and L62 of the other subunit as shown in Figure 1B. Furthermore, I44 of monomer A contacts L47, F5 and L62 of monomer B and vice versa. Apart from the hydrophobic interactions described above, the formation of two salt bridges between K45 of one monomer and E2 of the other monomer is predicted.

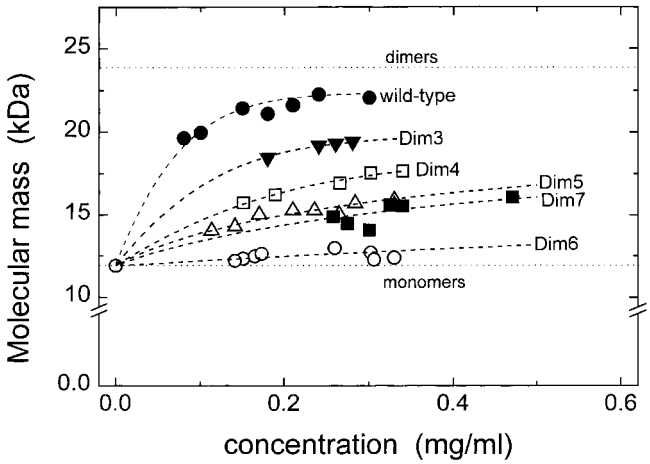

Fig. 2. Concentration-dependence of the average molecular mass of wild-type and mutated proteins determined by analytical ultracentrifugation. Equilibrium sedimentation measurements were performed at different protein concentrations in the presence of 75 mM NaCl. The plot shows a curve fit based on the assumption of a monomer-dimer equilibrium.

Design and Construction of Mutants to Verify the Model Predictions for Amino Acids Involved in Dimerization

To verify the model predictions, amino acids that appear at the putative dimeric interface were substituted in combinations or individually using site-directed mutagenesis by PCR (see Materials and Methods).

CopR mutant Dim3 carries two amino acid exchanges: the nonpolar I44 was replaced by the polar D and the polar K45 was replaced by the nonpolar A, the latter making the formation of the predicted salt bridge between K45 and E2 impossible. In Dim5 two nonpolar residues, L58 and L62 were replaced by D and the small nonpolar V59 was exchanged with the larger polar S. Thus, Dim5 contains mutations in three positions predicted to be involved directly in dimerization. Dim6 carries a single F5V substitution.[11] Dim7 represents a combination of Dim3 and the additional substitution of L47 by the smaller hydrophobic A. Both, Dim6 and Dim7 contain exchanges of amino acids also predicted to be involved in the hydrophobic core of the monomer subunit, namely F5V in Dim6 and L47V in Dim7. Dim4 is a combination of Dim3 and the three additional substitutions V59A, V60A, and D61A. Mutants Dim1 and Dim2 were constructed before the model was built, each of them containing three consecutive amino acid (S56, T57, L58 in Dim1 and V59, V60, D61 in Dim2) exchanges by alanine.

All Mutants Show a Decreased Ability to Dimerize

To characterize the ability of the mutated proteins to dimerize, glutardialdehyde crosslinking experiments were carried out (Table I). Compared to wild-type CopR, decreased amounts of crosslinked dimers were obtained for the examined dimerization mutants, indicating that the mutations affect dimerization either directly or indirectly. Furthermore, the equilibrium dissociation constants K_{Dimer} for protein dimers of Dim3, Dim4, Dim5, Dim6, and Dim7 were determined by analytical ultracentrifugation (Fig. 2 and Table I). Due to the low yield of mutated proteins Dim1 and Dim2 overexpressed in *E. coli*, no analytical ultracentrifugation was performed but K_{Dimer} was estimated from the

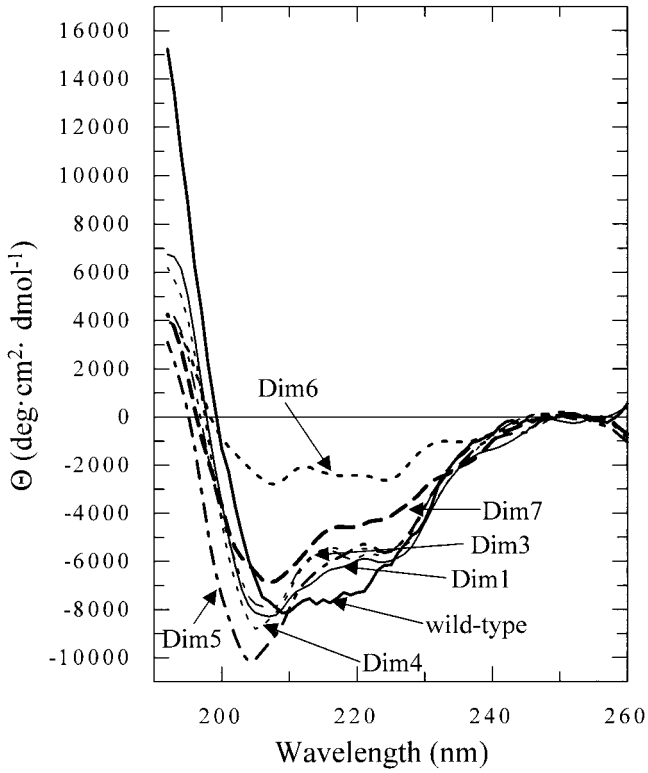

Fig. 3. Circular dichroism spectra of wild-type and mutated proteins. The measurements were performed at room temperature at 0.1 g/l protein concentration.

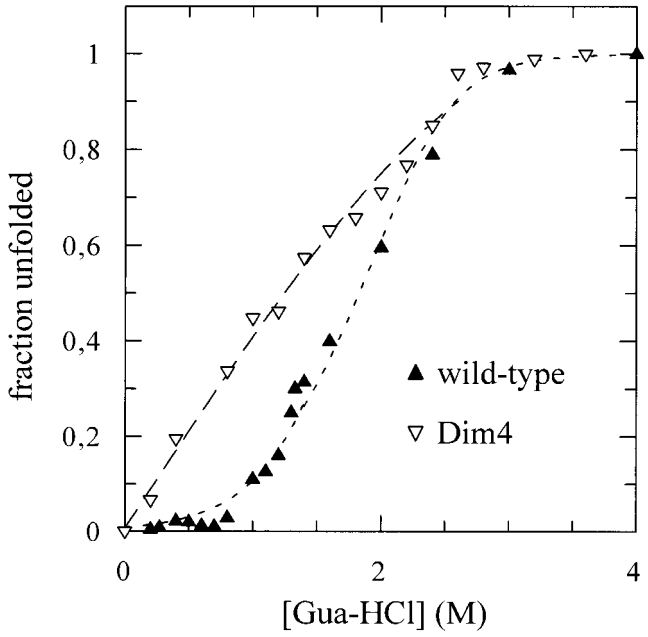

Fig. 4. Guanidine-HCl induced denaturation. Only the transition curves of wild-type protein and of mutant Dim4 are shown. The data represent the average of two measurements. The lines were drawn to guide the eye. Denaturation of the mutated proteins Dim1, Dim3, Dim5, and Dim7 yielded transition curves similar to Dim4.

crosslinking experiments, based on the observed good correlation between the results of the crosslinking and the analytical ultracentrifugation experiments.

As can be seen in Table I, dimerization of all mutants was impaired with K_{Dimer} values between 10 μM (Dim3) and 100 μM (Dim7) compared to that of the wild-type protein (1.44 μM). Dim6 (F5V) showed an extremely low tendency to dimerize with a K_{Dimer} of 430 mM.

Dimerization Defective Mutants Show Wild-Type Like Secondary Structures But Decreased Stability

Circular dichroism measurements were used to detect and evaluate structural alterations in the mutated proteins in comparison to the wild-type protein His_6-CopR. The far-UV CD spectra of wild-type and mutated proteins are displayed in Figure 3. The wild-type protein spectrum shows a positive band near 192 nm and two negative bands near 208 and 222 nm, indicating the presence of α-helical structures. Spectra with only minor deviations were obtained for the mutated proteins Dim1, Dim3, Dim4, and Dim5, indicating that the respective mutations did not cause significant structural alterations in the protein.

In contrast to the above-described spectra, the CD spectrum of Dim6 (F5V) looked completely different: The α-helical content was strongly decreased, indicating that the single mutation caused structural changes in the protein conformation. The CD spectrum of Dim7 also showed structural alterations that, however, were less pronounced than those of Dim6.

In order to analyze the stability of the proteins, guanidine-HCl denaturation experiments were performed. To monitor the unfolding process, changes in ellipticity at 228 nm were recorded for the different guanidine-HCl concentrations. For the wild-type protein, it was shown that the denaturation is fully reversible (data not shown). A monophasic cooperative transition from the completely folded to the unfolded state was observed for the wild-type protein, whereas mutated proteins Dim1, Dim3, Dim4, Dim5, and Dim7 showed clearly non-cooperative unfolding curves (shown only for Dim4 in Fig. 4) and for Dim3, Dim4, Dim5, and Dim7 a shift of the midpoint of denaturation towards lower guanidine-HCl concentrations as listed in Table I. This indicates reduced stability against guanidine-HCl induced denaturation. Although Dim1 also displays a non-cooperative denaturation curve, the midpoint of denaturation is shifted towards higher concentrations of denaturant.

In contrast, Dim6 revealing marked structural changes compared to the wild-type, differed from the above described in that at low guanidine-HCl concentrations the ellipticity at 228 nm decreases, which could be indicative for the occurrence of protein aggregation before the unfolding process proceeds at higher concentrations of denaturant [shown in Steinmetzer et al.[11]]. It has been shown that especially partially unfolded proteins have the propensity to aggregate.[14]

Circular Dichroism Can Be Used to Monitor Protein Dimerization

CD-spectra of wild-type protein and the mutated protein Dim3 were recorded at different protein concentrations,

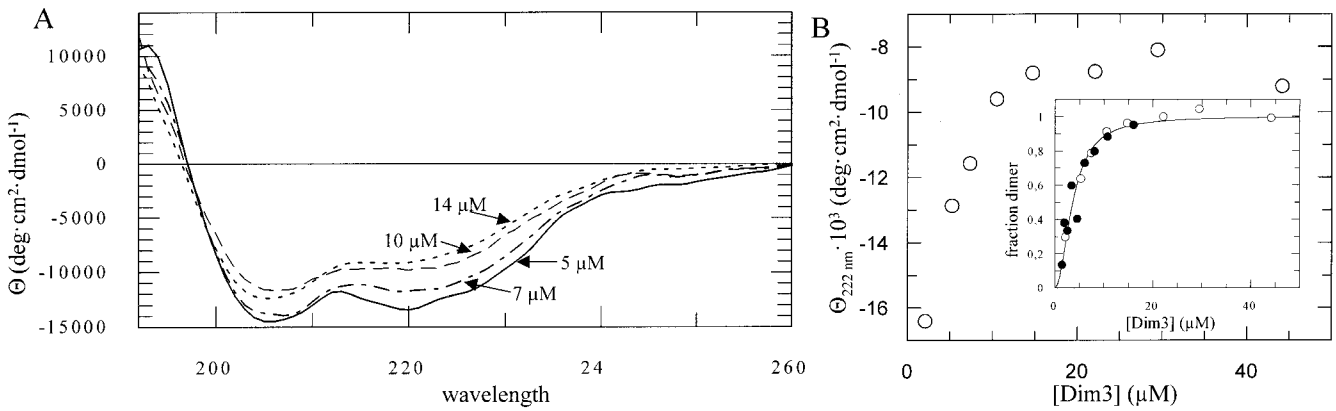

Fig. 5. **A:** Concentration-dependent changes in CD-spectra. The spectra of the mutated protein Dim3 were measured with protein dilutions made of a concentrated stock. For reasons of clarity, only four of the spectra measured are shown. The used concentrations are indicated. **B:** Concentration dependence of the mean molar ellipticity per residue Θ_{222nm} from the protein concentrations measured by circular dichroism. The values of Θ_{222nm} were derived from a series of normalized CD-spectra measured with the mutated protein Dim3. **Inset:** Data expressed as the concentration dependence of the fraction of dimers derived from two independent experiments (open and closed circles). The data were fitted according to a simple monomer-dimer equilibrium.

each derived by serial dilution of a concentrated stock. The respective normalized CD-spectra were found to be concentration dependent as shown in Figure 5A for the mutant Dim3. The spectra exhibit an isodichroic point near 198 nm indicating the presence of a simple two-state equilibrium. In Figure 5B the molar ellipticity at 222 nm is plotted vs. protein concentration as measured with the mutated protein Dim3. The observed changes in the CD-spectra presumably represent the dissociation of existing dimers upon dilution. An attempt to use circular dichroism measurements to monitor dimerization was made also by Franchini and Reid.[15] Based on the data presented in Figure 5B, the equilibrium dissociation constant for Dim3-dimers was determined and the value for $K_{Dimer} = 12.8 \pm 1.3$ μM derived is in good agreement with the respective value of 10 ± 2.7 μM obtained by analytical ultracentrifugation for the mutant. Using CD-measurements, the K_{Dimer} value for the wild-type protein was determined to be 0.8 ± 0.2 μM. This value is also in good correlation with the result of the analytical ultracentrifugation experiments. As expected, the mutated protein Dim6 that showed no tendency to dimerize in the micromolar concentration range, yielded concentration independent CD-spectra (data not shown).

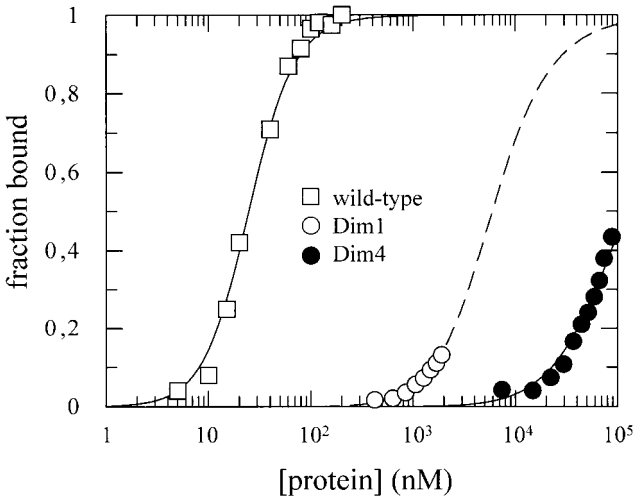

Fig. 6. DNA-binding curves. A constant amount of radioactively labelled DNA-fragment containing the wild-type operator of CopR was titrated with increasing concentrations of purified proteins and analyzed by EMSA. The binding-buffer, electrophoresis buffer, and polyacrylamide gel contained 75 mM NaCl. The curves were fitted based on a model of two coupled equilibria: protein dimerization and dimer-DNA complex formation.

Mutations Affecting Dimerization Influence DNA-Binding Activity In Vivo and In Vitro

It has been shown recently that CopR dimerization is a prerequisite for DNA-binding[10] and, therefore, it is expected that CopR mutants impaired in dimerization reveal an altered binding affinity for the CopR operator. The ability of the mutated proteins to bind the wild-type operator sequence was examined by EMSA using either *B. subtilis* crude cell extracts containing the respective protein or purified proteins. In the presence of 75 mM NaCl, only Dim1 and Dim4 were able to form stable complexes detectable by EMSA. The respective binding curves are shown in Figure 6. Based on the assumption of two coupled equilibria, protein-dimerization and DNA-binding, the data were used to calculate the equilibrium dissociation constants for the dimer-DNA complex $K_{complex}$. The obtained values for Dim1 and Dim4 were (0.9 ± 0.02) μM and (40 ± 1.5) μM, respectively. Under less stringent conditions without NaCl, EMSA with *B. subtilis* crude cell extracts containing wild-type CopR or the mutated proteins showed that DNA-binding of the dimerization mutants is reduced to 10–16% of the wild-type activity (see Table I). Consequently, the mutants exhibit the expected lower affinity for DNA-binding. As described previously, mutant Dim6 (F5V) is completely inactive in vitro.[11]

All dimerization mutants were inactive in vivo as assayed by copy number determination, indicating that the

observed residual activity in vitro was not sufficient to down-regulate plasmid copy number in the cell.

DISCUSSION

Amino Acids Between I44 and L62 Are Involved in the Formation of the Dimeric Interface

According to the model, amino acids E2, F5, I44, K45, L47, L58, V59, and L62 are located at the dimeric interface. Thereby, I44 of monomer A contacts F5, L47, and L62 of monomer B. E2 contacts K45 and L58 contacts L58 and L62 of the other subunit and vice versa. Furthermore, L58 of monomer A is located in between V59 and L62 of monomer B.

Here we present evidence that the region comprising amino acids I44 to L62 is directly involved in forming the dimeric interface of CopR. Analytical ultracentrifugation revealed that all mutated proteins with amino acid substitutions at positions predicted to be located at or near the dimeric interface showed a decreased ability to dimerize. Circular dichroism measurements of Dim1, Dim3, Dim4, and Dim5 showed only minor deviations from the wild-type spectrum, indicating that the respective mutations did not induce remarkable structural alterations in the proteins. Consequently, the impaired ability of these mutants to dimerize is caused directly by the introduced mutations and not due to structural alterations. Another argument in favour of this interpretation is that mutants impaired in specific DNA binding as R29Q and R34Q that also showed similarly slight changes in their CD spectra were not impaired in dimerization compared to wild-type CopR as demonstrated by analytical ultracentrifugation.[11] Apparently, subtle changes in the structure do not necessarily affect dimerization.

In contrast to the above discussed mutated proteins, Dim6 (F5V) and Dim7 (I44D, K45A, L47A) carry amino acid exchanges at positions predicted to be involved not only in dimerization but also in formation of the hydrophobic core of the protein, namely F5V and L47A. The phenyl ring of F5 is predicted to form an intersubunit contact to I44 but should also be involved in hydrophobic interactions to the side chains of L9, D61, P42, and L47 of its own subunit. Apart from the predicted contacts to the other monomer, L47 is also predicted to be involved in interactions with F5 and L58 of its own monomer. Thus, mutations in the positions F5 and L47 should not only lead to a decreased ability to dimerize but also to structural alterations in the protein fold and, indeed, both were observed. The observed drastic structural alterations caused by the substitution F5V are in accordance with the model, but due to the predicted involvement of F5 in several intrasubunit contacts our results do not unequivocally prove that F5 is directly involved in dimerization.

The combined substitutions of I44D and K45A (Dim3) caused a sevenfold increase of K_{Dimer}, indicating that at least one of the two amino acids is involved in dimerization. Since L47 is predicted to contact I44 on the other subunit, one would not expect a further decrease in the ability to dimerize by introducing the L47A mutation present in mutated protein Dim7. However, this substitution led to a further tenfold increased K_{Dimer} of Dim7 compared to Dim3. As mentioned above, L47 is also predicted to form hydrophobic contacts within the same subunit and thus the drastic effect upon dimerization caused by this additional mutation might also be the result of an improperly folded monomer. But, in contrast to Dim6, the structural alterations in the mutated protein Dim7 are much less pronounced.

CD Measurements Confirmed That Monomers of CopR Are Folded

Is the mutated protein Dim6 (F5V) almost unable to dimerize because it is misfolded or is it misfolded because it is almost unable to dimerize? The latter would implicate that the monomer of wild-type CopR might exist in an unfolded form. By dilution of mutated proteins to concentrations well below K_{Dimer}, CD spectra of predominantly monomeric proteins can be measured. The respective spectrum of Dim3, shown in Figure 5A, proves that significant amounts of secondary structure elements are present in the monomeric form. In contrast, the CD spectrum of Dim6 (F5V), a mutated protein that is present almost only in its monomeric form (analytical ultracentrifugation), reveals an unfolded protein. Therefore, it seems rather likely that the incorrect fold of the Dim6 monomer causes the significantly impaired ability to dimerize than the interruption of the predicted hydrophobic contact between F5 and I44 at the dimeric interface. The conclusion can be drawn that the dimerization of CopR occurs via a folded monomeric intermediate as it was also observed for the homodimeric Cro repressor of phage l,[16] whereas in the cases of P22 Arc repressor[17] and HPV-16E2-DNA binding domain[18] the monomer is unfolded.

Whether the dimerization of CopR is accompanied by secondary structural changes or involves additional interactions between secondary structure elements at the dimeric interface leading to the observed changes in the CD-signal upon dimerization, cannot be distinguished yet although the latter seems to be more likely.

The above discussed CD data showing that CopR monomers are folded are in good correspondence with the results of analytical ultracentrifugation where also only two CopR species—folded monomers and folded dimers—have been found.

Denaturation experiments with guanidine-HCl revealed that, in contrast to wild-type CopR, unfolding of the dimerization mutants occurs in a non-cooperative manner with midpoints of denaturation below that of the wild-type protein (Table I). This indicates that the dimers of the mutated proteins are less stable than the wild-type dimer although they do not differ significantly from the wild-type in their secondary structures as shown for Dim1, Dim3, Dim4, and Dim5.

The non-cooperative unfolding curves observed for the mutated proteins might be indicative for a molten-globule like state as it was observed for Arc repressor mutants with decreased ability to dimerize.[19] However, since we do not have data on the shape and compactness or on the exposure of hydrophobic side chains of the mutated proteins, further experiments are needed to answer this question.

All Dimerization Mutants Reveal a Significantly Decreased DNA Binding Affinity

All mutated Cop proteins were inactive in copy-number regulation in vivo. DNA binding in vitro as analyzed by EMSA under non-stringent conditions, i.e., in the absence of NaCl in both polyacrylamide gel and electrophoresis buffer, was significantly decreased for Dim1, 3, 4, 5, and 7 and completely abolished for Dim6 (Fig. 6). To be able to calculate the $K_{complex}$ from the binding curves, EMSA was performed also in the presence of 75 mM NaCl, i.e., under conditions used for the analytical ultracentrifugation experiments. Under these conditions, only the wild-type protein, Dim1 and Dim4, formed stable complexes that could be detected by EMSA. The values for $K_{complex}$ were 9.5×10^{-7} M and 4.0×10^{-5} M for Dim1 and Dim4, respectively, indicating a reduced affinity of the mutated proteins for the CopR operator when compared to the wild-type $K_{complex}$ of 4×10^{-10} M. Since dimerization of mutant Dim1 is less decreased than that of Dim4 when compared to the wild-type protein (Table I), it also binds the DNA with a higher affinity than Dim4. This is in good agreement with our previous data that CopR binds to the DNA only as a dimer and not as a monomer.[10]

Since the K_{Dimer}-values of all mutated proteins are below their estimated intracellular concentration in *B. subtilis* cells of 200 to 300 μM (analyzed by quantitative Western blotting, data not shown), there should be enough dimers in the cells to bind the CopR operator and to down-regulate the plasmid copy number. However, the existing dimers might not be able to form stable complexes with the DNA under cellular conditions with a physiological concentration of ≈150 mM NaCl. This could account for the inactivity of all Dim mutants in vivo. One possible explanation could be that mutations at the dimeric interface lead to unpredictable effects on the correct positioning of the two recognition helices of the CopR dimer towards the DNA that is necessary for high-affinity complex formation. Another hypothesis could be that the wild-type protein induces a conformational change in the DNA target upon binding to allow a tight fit. Indeed, using FRET (fluorescence resonance energy transfer) measurements and hydroxyl radical footprinting experiments we found recently that the operator DNA is slightly bent upon binding of CopR (Steinmetzer et al., unpublished data). Dimerization mutants might be unable to induce these structural changes in the DNA, and, consequently, their complexes with the DNA are less stable.

This defect could not be complemented by the concentration of these dimerization mutants well above their K_{Dimer}. Therefore, not only the ability to dimerize, but the ability to bind specifically to the DNA seems to be critical for in vivo function and, thus, regulation.

In summary, with the data presented here, we can narrow down the dimeric interface to a region between amino acids I44 and L62. Within this region, either amino acid I44 or K45 or both are clearly involved in dimerization. Furthermore, either L58, V59, or L62 or a combination of them are involved in dimerization. Due to their location in the hydrophobic core, the contribution of L47 to dimerization cannot be assessed unequivocally. The same holds true for F5 at the N-terminus of CopR, which is important for the structure of the monomer. Consequently, our experimental data represent further proof of our 3D model of CopR. However, to unequivocally prove the location of single amino acids at the dimeric interface, the tertiary structure of CopR has to be determined. Therefore, experiments are currently under way to obtain crystals of CopR.

ACKNOWLEDGMENTS

The authors thank E. Birch-Hirschfeld, Institute for Virology, Jena, for synthesizing the oligodeoxyribonucleotides. We are grateful to T. Steinmetzer, Institute for Biochemistry, Jena, for help with Figure 1B.

REFERENCES

1. del Solar GH, Perez-Martin 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 1990;265:12569–12575.
2. Acebo A, de Lacoba MG, Rivas G, Andreu JM, Espinosa M, del Solar G. 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. PROTEINS 1998;32:248–261.
3. Gomis-Rüth F X, Sola M, Acebo P, Parraga A, Guasch A, Eritja R, Gonzalez A, Espinosa M, del Solar G, Coll M. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. EMBO J 1998;17:7404–7415.
4. Brantl S, Behnke D. Copy number control of the streptococcal plasmid pIP501 occurs at three levels. Nucleic Acids Res 1992;20: 395–400.
5. Brantl S. The *copR* gene product of plasmid pIP501 acts as a transcriptional repressor at the essential *repR* promoter. Mol Microbiol 1994;14:473–483.
6. Brantl S, Wagner EGH. Dual function of the *copR* gene product of plasmid pIP501. J Bacteriol 1997;179:7016–7024.
7. Le Chatelier E Ehrlich SD, Janniere L. The pAMβ1 Cop repressor regulates plasmid copy number by controlling transcription of the *repE* gene. Mol Microbiol 1994;14:463–471.
8. Ceglowski P, Alonso JC. Gene organziation of the *Streptococcus pyogenes* plasmid pDB101: sequence analysis of the orf eta-*copS* region. Gene 1994;145:33–39.
9. Steinmetzer K, Brantl S. Plasmid pIP501 encoded transcriptional repressor CopR binds asymmetrically at two consecutive major grooves of the DNA. J Mol Biol 1997;269,684–693.
10. Steinmetzer K, Behlke J, Brantl S. Plasmid pIP501 encoded transcriptional repressor CopR binds to its target DNA as a dimer. J Mol Biol 1998;283:595–603.
11. Steinmetzer K, Hillisch A, Behlke J, Brantl S. Transcriptional repressor CopR: Structure model based localization of the DNA binding motif. Proteins 2000;38;393–406.
12. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning. A laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1989.
13. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 1977;74:5463–5467.
14. Fink AL. Protein aggregation: folding aggregates, inclusion bodies and amyloid. Fold Design 1998;3:R9–23.
15. Franchini PLA, Reid RE. A model for circular dichroism monitored dimerization and calcium binding in an EF-hand synthetic peptide. J Theor Biol 1999;199:199–211.
16. Jana R, Hazbun TR, Mollah AKMM, Mossing MC. A folded monomeric intermediate in the formation of lambda Cro dimer-DNA complexes. J Mol Biol 1997;273:402–416.
17. Bowie JU, Sauer RT. Equilibrium dissociation and unfolding of the Arc repressor dimer. Biochemistry 1989;28:7139–43.
18. Mok YK, de Prat-Gay G, Butler PJ, Bycroft M. Equilibrium dissociation and unfolding of the dimeric human papillomavirus strain-16 E2 DNA-binding domain. Protein Sci. 1996;5:310–319.
19. Milla ME, Sauer RT. Critical side-chain interactons at a subunit interface in the Arc repressor dimer. Biochemistry 1995;34:3344–3351.