doi:10.1006/jmbi.2001.5157 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 314, 789–796

JMB

Crystal Structure of ω Transcriptional Repressor Encoded by *Streptococcus pyogenes* Plasmid pSM19035 at 1.5 Å Resolution

Kazutaka Murayama[1], Peter Orth[1], Ana B. de la Hoz[2], Juan C. Alonso[2] and Wolfram Saenger[1]*

[1]*Institut für Kristallographie Freie Universität Berlin Takustraße 6, D-14195, Berlin Germany*

[2]*Departmento de Biotecnología Centro Nacional Biotecnología (CSIC), Campus Universidad Autónoma de Madrid Cantoblanco, 28049, Madrid Spain*

**Corresponding author*

The 71 amino acid residue ω protein encoded by the *Streptococcus pyogenes* non-conjugative plasmid pSM19035 is a transcriptional repressor that regulates expression of genes for copy number control and stable maintenance of plasmids. The crystal structure of ω protein has been determined by multiple isomorphous replacement, including anomalous scattering and refined to an *R*-factor of 21.1 % ($R_{free} = 23.2$ %) at 1.5 Å resolution. Two monomers related by a non-crystallographic 2-fold axis form a homodimer that occupies the asymmetric unit. Each polypeptide chain is folded into two α-helices and one β-strand forming an antiparallel β-ribbon in the homodimer. The N-terminal regions (1-23 and 1-22 in subunits I and II, respectively) are not defined in the electron density due to proteolysis of the N-terminal 20 amino acid residues during crystallisation and partial disorder. The ω protein belongs to the structural superfamily of MetJ/Arc repressors featuring a ribbon-helix-helix DNA-binding motif with the β-ribbon located in and recognizing the major groove of operator DNA; according to a modelled ω protein-DNA complex, residues Arg31 and Arg31′ on the β-ribbon are in positions to interact with a nucleobase, especially guanine.

Keywords: transcriptional repressor; ω protein; X-ray crystal structure; ribbon-helix-helix motif; MetJ/Arc superfamily

Introduction

The broad-host-range plasmid pSM19035 isolated from *Streptococcus pyogenes* is a member of the *inc18* family of plasmids and very stably maintained in Gram-positive bacteria with low G + C content in their DNA.[1] The pSM19035-encoded 71 amino acid residue ω protein, which occurs as homodimer in solution,[2] binds specifically to a DNA heptad repeat sequence located in the upstream promoter regions of *cop*S and δ genes and of the ω-ε-ζ operon.[3] Protein CopS corrects downward fluctuations in the number of plasmids, the ParA-like protein δ contributes actively in partitioning of plasmid copies,[3,4] and the ε and ζ proteins form an addiction or post-segregational killing system, with ζ acting as toxin and ε as antitoxin.[5] The ω protein directly represses transcription of the *cop*S, δ and ω genes, and indirectly controls expression of the ε and ζ genes. These functions of ω protein indicate that it is a global regulator actively engaged in copy number control and in the stable maintenance of plasmids of the *inc*18 family.[3]

Besides the ω protein, a similar ω2 protein has been isolated that features an identical amino acid sequence in positions 1 to 55 and an eight residue extension at its C terminus. The two proteins are encoded by the *inc18* family of plasmids, ω exclusively by the non-conjugative plasmid pSM19035 but both ω and ω2 by the conjugative plasmids pIP501 and pAMβ1 (A.B.d.l.H., unpublished results).

As shown by spectroscopic studies, the ω protein is largely α-helical and binds with high cooperativity to at least two consecutive copies of the 5′-$^{A}/_{T}$ATCAC$^{A}/_{T}$-3′ consensus repeat.[2,3] The

Abbreviations used: MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; PEG, polyethylene glycol; RHH, ribbon-helix-helix; r.m.s., root-mean-square.

E-mail address of the corresponding author: saenger@chemie.fu-berlin.de

0022-2836/01/040789–8 $35.00/0

heptads are arranged either in one contiguous block with seven direct and two inverted repeats $[(\rightarrow)_7(\leftarrow)_2]$ as in *P*δ (Figure 1), or in *n* copies of the block $[(\rightarrow)_2(\leftarrow)]_n$, with $n = 2$ or 3 and one additional direct or inverted repeat downstream of the block, as in *P*ω $[(\rightarrow)_2(\leftarrow)]_2(\leftarrow)$ and in *P*copS $[(\rightarrow)_2(\leftarrow)]_3(\rightarrow)$. The mechanism by which the ω protein controls transcription of the *P*copS, *P*δ and *P*ω promoters has not been established but it is clear that the ω protein does not inhibit binding of RNA polymerase to the promoter region.[3]

According to a BLAST search,[6] the ω protein shows no significant sequence homology to any other structurally known protein. To reveal details of the structure of the ω protein and to gain insight into its mechanism of action, we have crystallised ω protein and determined its structure at 1.5 Å resolution. To understand the interaction between double-helical DNA and ω protein, the putative complex between DNA and ω protein has been modelled.

Results and Discussion

Overall structure

The crystal structure of ω protein presented here includes two subunits, I and II, in the asymmetric unit, forming a homodimer (ω_2), Figure 2. The subunits are related by a local 2-fold rotation axis and can be superimposed with an r.m.s. deviation of 1.1 Å (mean value for the 48 common C^α atoms). According to the criteria programmed in WHAT_CHECK,[7] each subunit consists of one β-strand and two α-helices, A and B. The 20 N-terminal residues were lost by proteolysis during crystallisation (see Materials and Methods), and Lys21 to Asp23 (subunit I), and Lys21′ to Lys22′ (subunit II, the prime (′) referring to subunit II) could not be identified in the electron density map due to disorder. The polypeptide chains of ω protein were traced from residue Ile24 and Asp23′ to the C termini, respectively. A β-strand is formed by Lys28 (Lys28′) to Val32 (Val32′). Helix A begins with Ala34 (Ala34′) and ends at Asn47 (Lys46′). In both subunits, helix B starts at Val51 and terminates at Tyr66, and the C terminus is well defined except for the side-chains of residues 69-71. The two helices A and B are connected by a relatively flexible loop that includes two glycine residues, Gly48 and Gly49. In total, the secondary structure elements are 42 % α-helix, 7 % β-sheet and 51 % loop, calculated against full-length protein. This agrees with CD measurements of ω protein yielding 42 % α-helix, 13 % β-sheet and 45 % loop, respectively.[2] According to the CD spectra, the N-terminal region of ω protein is probably intrinsically flexible and does not fold into an α-helix.

Dimer structure and crystal packing

The local 2-fold rotation axis (*C*2) relating the two subunits of the ω protein is perpendicular to the crystallographic 6-fold screw axis of space group $P6_1$ (see Materials and Methods). The buried subunit-subunit interface was calculated to be 1800 Å^2. The dimer is stabilised by monomer-monomer contacts (<4 Å) between helices B and B′ involving a hydrogen bond Tyr66OH-Glu61′$O^{\varepsilon 2}$ and a salt-bridge Lys65N^{ζ}-Glu61′$O^{\varepsilon 1}$, and their *C*2-symmetry-related counterparts. Additional interactions occur among hydrophobic side-chains located at the dimer interface (Ile40, Val51, Val54, Met55, Ala58, Leu59, Tyr62, Ile63, Tyr66 and Leu67) and in β-strands (Val30 and Val32).

The twisted antiparallel β-ribbon is located on the surface of the homodimer and oriented perpendicular to the local *C*2-symmetry axis (Figure 2). Besides the normal main-chain hydrogen bonds, it is stabilised by hydrogen bonds and salt-bridges among peptide side-chains at both ends of the β-ribbon. At the N terminus of the β-strand of subunit I, two hydrogen bonds and five salt-bridges are formed: Arg33′N-Asp56$O^{\delta 1}$, Ala34′N-Gly26O, Arg33′N^{ε}-Asp35′$O^{\delta 1}$, Arg33′$N^{\eta 2}$-Asp35′$O^{\delta 2}$, Arg33′$N^{\eta 1}$-Glu60$O^{\varepsilon 1}$, Lys52N^{ζ}-Asp56$O^{\delta 2}$ and Lys52N^{ζ}-Asp27$O^{\delta 1}$ (Figure 3); corresponding symmetry-related interactions are found at the end of the N terminus of the β-strand of subunit II.

In the crystal, intermolecular contacts are formed mainly between the reference dimer and dimers that are symmetry-related by the 6-fold screw axis. Among surface-exposed amino acid residues and main-chain peptide atoms, seven hydrogen bonds and four salt-bridges are formed. They occur

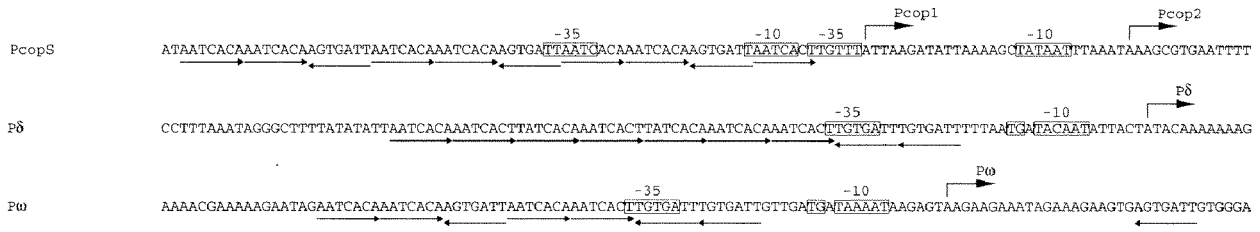

Figure 1. DNA recognition sequences of the ω regulator. The nucleotide sequences of the *P*copS (*P*cop1-*P*cop2), *P*δ and *P*ω promoters are shown for comparison. The transcription start sites of *P*cop1, *P*cop2, *P*δ and *P*ω are denoted by bent arrows. Conserved −35 and −10 regions are boxed. The heptamers and their relative orientations are indicated by arrows below the nucleotide sequences.

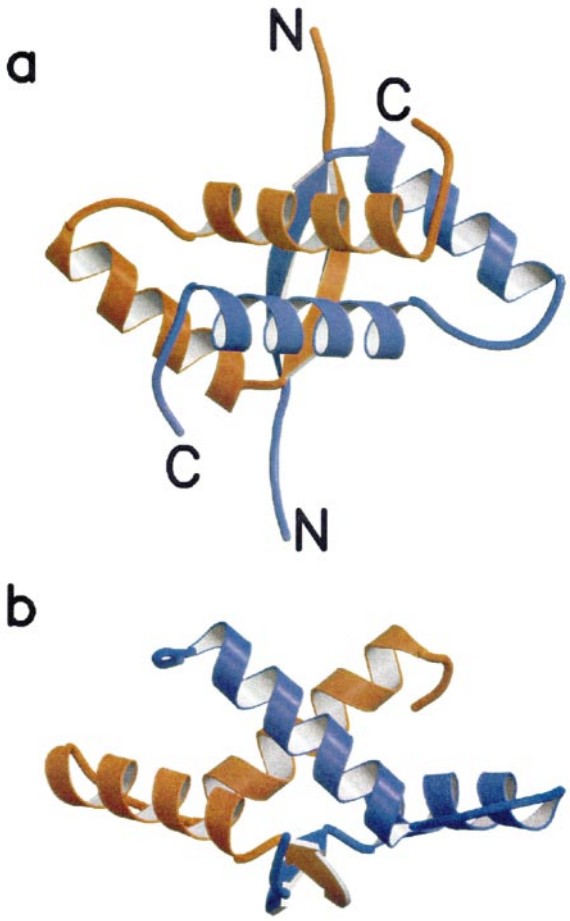

Figure 2. Ribbon diagram of the ω protein homodimer. Subunits I and II are drawn in red and blue, respectively. (a) Top view of the ω dimer. The local 2-fold axis is perpendicular to the paper plane. (b) Side view of the ω dimer, the local 2-fold axis is in the paper plane.

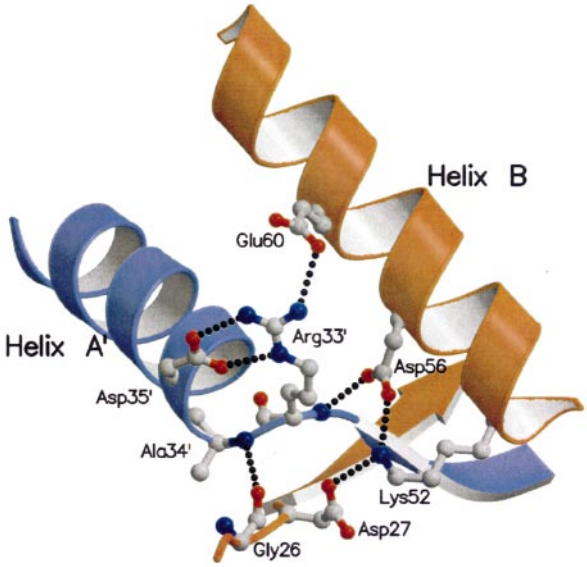

Figure 3. Network of hydrogen bonds and salt-bridges (dotted lines) at the N terminus of the β-strand of subunit I (red). Relevant amino acid residues are depicted as ball-and-stick models.

preferentially between the β-ribbon and helices B of a neighbour molecule. Additional contacts are formed between the C-terminal region of subunit I and the β-strand of subunit II of the symmetry-related dimer, including one hydrogen bond. These interactions are probably not relevant to cooperativity of binding to DNA heptads, as they involve the ß-ribbon that, according to the modelled complex (see below), is deeply inserted in the DNA major groove and adjacent dimers on the DNA are 23.8 Å apart.

The ω protein is a member of MetJ/Arc superfamily

A protein structure search using the DALI server[8] showed that ω protein belongs to the MetJ/Arc structural superfamily, although these proteins have sequence identity below 15% (Figure 4). In the last decade, four structures of repressors of the MetJ/Arc superfamily have been determined by X-ray crystallography: MetJ,[9,10] Arc,[11–14] CopG[15] and by NMR spectroscopy: Mnt,[16] Arc.[17–19] The common feature of all known members of the MetJ/Arc superfamily is a dimeric structure forming a ribbon-helix-helix (RHH) motif with a rigid hydrophobic core.[20–22]

The RHH structures are similar with the following r.m.s. deviations between ω protein dimer and other repressors: 2.2 Å (MetJ-ω, for common 85 C^{α} atoms), 2.9 Å (Arc-ω, for common 86 C^{α} atoms), 3.2 Å (Mnt-ω, for common 83 C^{α} atoms), and 2.7 Å (CopG-ω, for common 86 C^{α} atoms). The superimposed structures of MetJ and ω protein dimer are shown in Figure 5. Of the three secondary structure elements of the repressors, r.m.s. deviations of partial structures comprising β-strands and helices A between ω protein and the other repressors are large (2.2-3.6 Å), probably because these elements have individually different biological functions (base-pair recognition, DNA backbone contacts and tetramerisation on DNA). By contrast, r.m.s. deviations between helices B are small (1.0-1.5 Å), except for Mnt (2.8 Å), as many of the hydrophobic amino acid residues on helices B contribute to subunit dimerisation. The β-ribbon of each of the repressors comprises seven to nine amino acid residues per subunit in the case of MetJ, Arc, Mnt and CopG. By contrast, only five amino acid residues form this β-ribbon in ω protein in agreement with the shorter DNA-binding site (see below).

Of the above five repressors, MetJ and Mnt have an extra C-terminal region, which is attributed to their biological function (corepressor binding for

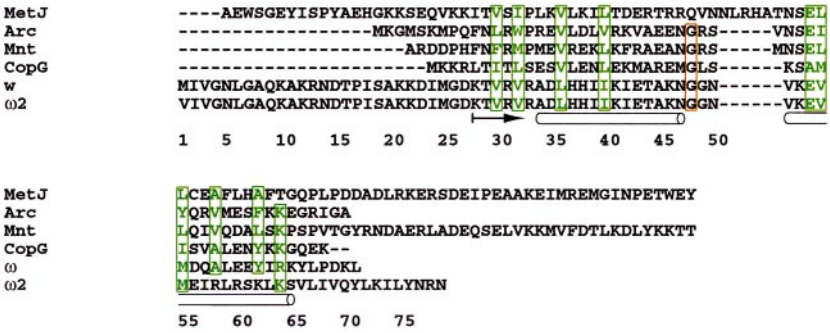

Figure 4. Alignment of amino acid sequences of the structurally known members of the MetJ/Arc superfamily. The conserved glycine residue in the turn connecting α-helices A and B is boxed in red, and type-conserved hydrophobic amino acids are boxed in green. The arrow and cylinders indicate β-strand and α-helices, respectively.

MetJ[9] and tetramerisation for Mnt[23]). Except for CopG, which has the shortest polypeptide chain (45 residues) of this superfamily, the N termini of these repressors are flexible. In the case of ω protein and MetJ, the flexible N-terminal region is 27 and 21 residues long, respectively. The N-terminal region of MetJ occupies different positions in free and in operator DNA-bound form.[10] Unfortunately, these parts are not seen in the crystal structure of the ω protein dimer.

As mentioned at the outset, the conjugative plasmids of the *inc18* family (pIP501, pAMβ1) encode for two ω-like proteins (ω and ω2), and the promoter region of many uncharacterised genes has the cognate binding site for ω protein (A.B.d.l.H., unpublished results). Except for the N-terminal amino acid residue, the amino acid sequence of ω2 protein is identical with that of ω protein until Met55 (Figure 4). The relevance of encoding two proteins with identical recognition site and different C terminus remains to be elucidated.

DNA-binding model

The X-ray crystal structures of repressor-DNA complexes of MetJ,[10] of Arc[11,13] and of CopG[15] have shown that in each case, two repressor dimers bind to double-stranded operator DNA. The distances between the binding centres of the two dimers are different: MetJ 8 bp apart, Arc 11 bp apart, Mnt and CopG each 9 bp apart. For the ω protein, it is assumed that the homodimer also binds to DNA with its β-ribbon motif and the binding centres of two dimers are 7 bp apart in agreement with the heptad repeat structure of operator DNA. The number of base-pairs between two binding centres determines the relative positioning of the protein dimers on the DNA strand,

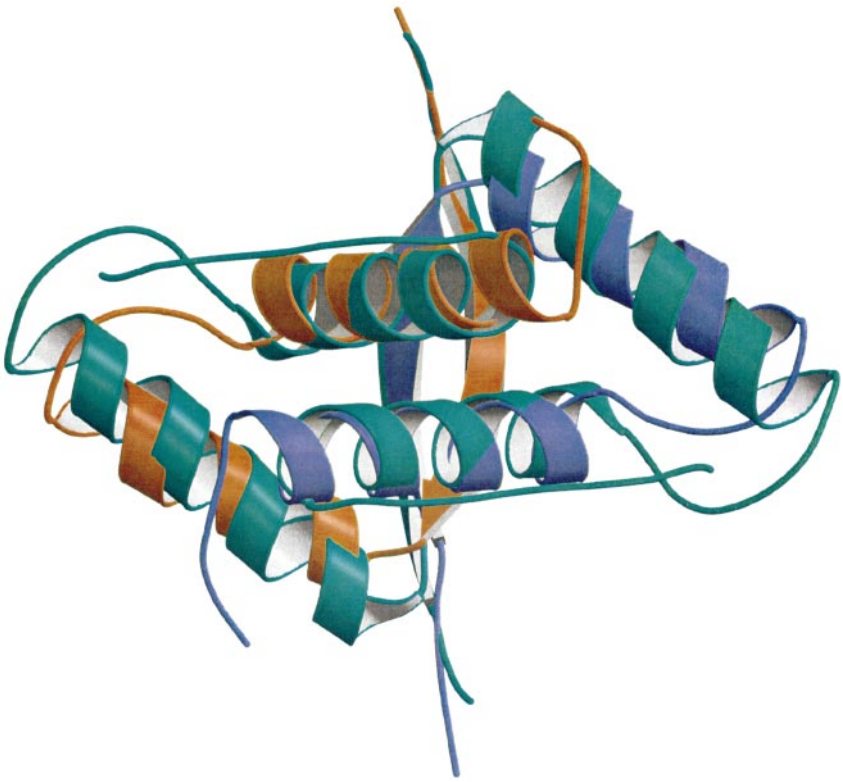

Figure 5. Superimposed structures of MetJ and ω protein, showing only the RHH motif: ω protein in red and blue, MetJ in green. The non-crystallographic 2-fold axis is perpendicular to the paper plane.

each base-pair counting for 3.4 Å separation and 36° rotation, yielding a total separation of 23.8 Å and rotation of 252° between ω_2 dimers bound to adjacent haptads.

To determine the geometry of the ω_2-DNA interaction, ω_2 was cocrystallized with two and three heptads long DNA (direct and inverted repeats). Since the obtained crystals had, however, only poor X-ray diffraction quality, no structure could be determined, and we resorted to model building studies utilising the atomic coordinates of the slightly bent *B*-DNA of the MetJ-DNA complex. In the modelled complex, the antiparallel β-ribbon of the ω protein dimer fits well to the DNA major groove (Figure 6). The β-ribbon contains two arginine residues (Arg31 and Arg31′) whose side-chains are directed toward the DNA major groove and would be able to interact with the two guanine bases in the heptad. This is consistent with the observation that minor groove binders (e.g. distamycin and spermine) do not affect binding of ω protein to its cognate binding site, whereas the major groove binder methyl green competes with binding of ω protein to this site (A.B.d.l.H. & J.C.A., unpublished results). In order to localise ω_2 protein on the DNA heptad, footprinting studies are in progress.

In the crystal structures of the complexes formed between MetJ, Arc and CopG and their operator DNA sequences, specific recognition occurs within the DNA major grooves. Although structures of repressor-DNA complexes of ω protein and of Mnt have not been determined, mutation studies for Mnt have suggested that of the five amino acid residues in the N-terminal region, the side-chains of Arg2 and Arg10 are important for operator binding.[24,25] In ω protein dimer, Arg31 and Arg31′ (corresponding to Arg10 in Mnt) are able to form contacts with base-pairs in the DNA major groove (Figure 6). Similar to other prokaryotic operators, the operators of MetJ, Arc and CopG have palindromic sequences, but Arc and CopG bind asymmetrically to the operator.[11,15] A similar asymmetry is also found for ω protein as the symmetrical dimer binds to a non-palindromic operator sequence (Figure 1). The complexes between operator DNA and MetJ, Arc and CopG show that amino acids in contact with the DNA backbone are located in the flexible N-terminal region and at the N terminus of helix B. For ω protein, the residues Lys21, Lys22, Lys21′, Lys22′ (N-terminal flexible part), Lys28 and Lys28′ (β-ribbon), Lys41, Lys41′ (C terminus of helix A) and Lys52, Lys52′ (N terminus of helix B) are potential phosphate-binding residues. If bound to adjacent heptads, the N termini of the two ω_2 dimers could possibly interact and contribute to the cooperativity of binding.

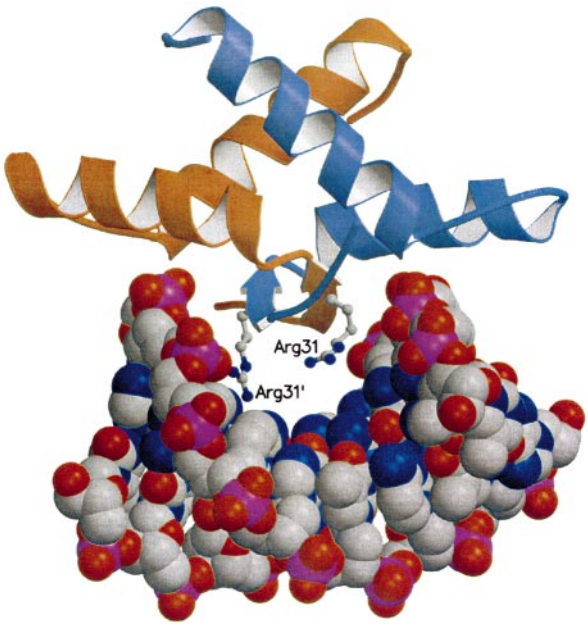

Figure 6. Modelled complex between the ω dimer and *B*-DNA. The β-ribbon binds to the major groove of DNA. The possible contact residues (Arg31 and Arg31′) are depicted as ball-and-stick models, and DNA is illustrated as a space-filling model.

Conclusions

The crystal structure described here shows that the ω protein homodimer is folded in the RHH motif characteristic of the MetJ/Arc structural superfamily. The two ω protein molecules in the dimer are related by a non-crystallographic 2-fold rotation axis and the dimer is stabilised by hydrophobic interactions and hydrogen bonds formed among helices B, B′ and the β-strands. The antiparallel β-ribbon as a potential DNA-binding motif lies at the surface of the dimer.

In the modelled ω protein-DNA complex, the β-ribbon is located in the major groove of the DNA binding site. Arg31 and Arg31′ can form hydrogen bonds to the guanine base, and Lys28, Lys28′, Lys41, Lys41′, Lys52 and Lys52′ are directed to the DNA backbone and may bind to phosphate groups. It is worthy of note that the ω repressor, which has a 2-fold symmetry, binds to a target DNA that does not show palindromic symmetry, in contrast to other repressor/operator systems belonging to the MetJ/Arc structural superfamily.

Previously we have shown that: (i) ω protein is a regulator that links plasmid copy number control and plasmid stability; (ii) ω protein is required to ensure active plasmid partitioning; (iii) ω protein is involved in postsegregational killing of plasmid-depleted cells; (iv) the conjugative plasmids of the *inc*18 plasmid encode for a second protein ω2 related to ω protein; and (v) the cognate DNA recognition sites for ω protein are present in all plasmids of the *inc*18 family.[3] It is likely, therefore, that ω and ω2 proteins are global regulators that recognise the same DNA-binding motif (see above), but unidentified factor(s) might allosterically enhance

or impair its (their) DNA-binding capacity and modulate transcription.

Materials and Methods

Crystallisation and data collection

ω protein was overexpressed and purified as described.[26] A single MALDI-TOF peak at a molecular mass of 7992 showed that the ω protein was pure and intact (the expected mass is 7989). The crystallisation conditions reported[26] have been modified: 2 μl of 25 mg/ml of ω protein in 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5% (v/v) glycerol were mixed with 2 μl of reservoir solution containing 30% (w/v) PEG 3350, 100 mM sodium acetate and 100 mM Tris-HCl (pH 8.5), and equilibrated at 18 °C against the reservoir solution using the hanging-drop vapour diffusion method. After three months, crystals grew as hexagonal columns, space group $P6_1$ with the unit cell dimensions given in Table 1. Two monomers are located in the asymmetric unit. According to MALDI-TOF of redissolved crystals, the molecular mass of 5839 indicates that the 20 N-terminal amino acid residues were cleaved off during crystallisation.

Heavy-atom derivatives were prepared by different soaking techniques. An osmium derivative was obtained by soaking crystals in 1 mM K_2OsCl_6 in harvesting buffer (35% PEG 3350, 100 mM Tris-HCl (pH 8.5), 100 mM sodium acetate) for 24 hours. Two halide ion derivatives were prepared by quick cryo-soaking[27] with 350 mM NaI and 500 mM NaBr respectively, in the harvesting buffer. Another iodine derivative was obtained by soaking in harvesting buffer containing 16 mM NaI/chloramin-T.[28]

X-ray diffraction data were collected under cryogenic conditions (100 K). Crystals were flash-frozen after incubation in the harvesting buffer. Initial X-ray data collection for native (data set-II) and three heavy-atom derivative crystals were conducted with in-house graphite monochromatised CuKα radiation (λ = 1.5418 Å) using a MAR Research Image Plate detector mounted on an Enraf-Nonius FR571 rotation anode X-ray generator operated at 40 kV, 70 mA. High-resolution native (data set I) and bromide derivative data were collected at the EMBL Outstation at DESY (Hamburg) on beamline X11 using a Mar Research CCD detector (Table 1). All X-ray diffraction data were integrated and scaled using the DENZO package.[29]

Multiple isomorphous replacement

The crystal structure of ω protein was determined by multiple isomorphous replacement including anomalous scattering (MIRAS). All computations for phasing were performed with the CCP4 program package[30] and SHELX.[31] The heavy-atom derivative data and the native data set-II were scaled using FHSCAL. The heavy-atom positions were determined by direct methods (SHELX) and anomalous and/or isomorphous difference Patterson maps. Minor heavy-atom sites of the four derivatives were determined by difference Fourier syntheses and refined using VECREF. Refinement of the heavy-atom temperature factors and occupancies and calculation of the native phases were done within MLPHARE.

Model building and solvent flattening

In the calculated electron density map, four α-helices and two β-strands were identified in the asymmetric unit. These secondary structure elements are related to each other by local 2-fold symmetry. Density modification using this symmetry improved the electron density map considerably. The polypeptide chain was traced

Table 1. Data collection and phasing statistics

	Native-I	Native-II	K_2OsCl_6	NaI/ch-T	NaI	NaBr
Space group	$P6_1$	$P6_1$	$P6_1$	$P6_1$	$P6_1$	$P6_1$
Cell dimensions (Å)						
$a = b$	46.11	45.74	45.90	45.60	45.78	46.31
c	88.29	88.08	88.56	87.92	88.18	88.23
Resolution range (Å)	29.6-1.5	29.5-2.6	23.7-2.6	29.4-2.6	29.5-2.6	29.6-2.2
X-ray source	DESY, X11	CuKα	CuKα	CuKα	CuKα	DESY, X11
Wavelength (Å)	0.9073	1.5418	1.5418	1.5418	1.5418	0.9073
No. reflections	69,483	8985	10,322	10,838	11,281	53,009
No. unique reflections	16,703	3259	3303	3246	3267	7243
Completeness (%)	97.6	99.5	99.8	99.7	99.6	99.3
R_{sym} (%)[a]	2.6	6.1	7.1	6.9	6.0	5.5
Average data multiplicity	4.2	2.8	3.1	3.3	3.5	7.3
Range of last shell (Å)	1.52-1.50	2.72-2.60	2.72-2.60	2.72-2.60	2.72-2.60	2.27-2.20
Completeness (%)	95.1	99.3	100	98.6	98.0	99.5
R_{sym} (%)[a]	38.4	12.2	20.7	20.2	12.2	31.1
Intensity ($I/\sigma(I)$), last shell	3.17	7.56	5.29	5.77	9.49	2.46
R_{deriv} (%)[b]			35.0	14.1	29.8	48.6
No. heavy-atom sites			2	4	2	3
Anomalous refinement			No	No	Yes	Yes
Phasing power[c], acent.			0.56	0.88	1.32	0.94
R_{cullis}[d], acent.			0.96	0.87	0.80	0.86

ch-T, chloramin-T.

[a] $R_{sym} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|/\Sigma_h\Sigma_i I_i(h)$.

[b] $R_{deriv} = \Sigma||F_{deriv}(h)| - |F_{nativ}(h)||/\Sigma_h|F_{nativ}(h)|$.

[c] Phasing power = r.m.s ($|F_H|/E$), where E is the lack of closure.

[d] $R_{cullis} = \Sigma_h||F_{PH} \pm F_P| - F_{H(calc)}|/\Sigma_h|F_{PH} \pm F_H|$.

Table 2. Refinement statistics

Resolution range used for refinement, Å	29.6-1.50
R_{factor}[a], free R_{factor} (%)	21.1, 23.2
No. of protein atoms	911
r.m.s. deviation	
Bonds (Å)	0.006
Angles (deg.)	1.054
r.m.s. difference in *B* over a bond	1.81
No. water molecules	125

[a] $R_{factor} = \Sigma_h||F_{obs}| - |F_{calc}||/\Sigma_h|F_{obs}|$.

on this electron density map as a polyalanine model using the program O.[32] Side-chains identified from this map were included in the initial structure model.

Structure refinement

The structure was refined using the CNS program package.[33] All refinement steps were monitored with the free *R*-factor based on 10% of the X-ray data. Following a simulated annealing protocol, the structure was refined using atom-positional and temperature factors refinement as well as manual model building. A total of 125 water molecules were added at peaks in the $F_o - F_c$ map with 3 r.m.s. deviation above mean density and within hydrogen bond distance to protein atoms or other water molecules. The N-terminal regions (amino acid residues Lys21-Asp23 in subunit I, Lys21′ and Lys22′ in subunit II) were not seen in the electron density map because of disorder. The loop regions connecting the α-helices A, B and A′, B′ as well as the C termini were less well ordered than the other parts of the structure as indicated by higher temperature factors (~60 Å^2). The stereochemical quality of the final model was assessed by WHAT_CHECK.[7] The refinement statistics are listed in Table 2. Figures were generated using programs MOLSCRIPT[34] and Raster 3D.[35]

Modelling of the complex between ω protein dimer and DNA

The complex between ω protein dimer and double-stranded DNA was modelled with LSQKAB[30] by fitting ω protein dimer to the major groove of the bent *B*-DNA of the MetJ-DNA complex (PDB code 1CMA).

Data Bank accession number

The atomic coodinates of the described crystal structure have been deposited with the RCSB Protein Data Bank under accession code RCSB005215 and with PDB under accession code 1IRQ.

Acknowledgements

This research was supported, in part, by grant BMC2000-0548 from the DGCYT and BIO4-CT98-0424 to J.C.A, and by grants Sa196/38-1 from the Deutsche Forschungsgemeinschaft and BIO4-CT98-0106 from EU to W.S. We are indebted to Dr Peter Franke for MALDI-TOF and to Wilhelm Weihofen and Bernhard Loll for help with the final manuscript. A.B.d.l.H. was the recipient of a Fellowship of the Gobierno Vasco.

References

1. Ceglowski, P., Boitsov, A., Chai, S. & Alonso, J. C. (1993). Analysis of the stabilization system of pSM19035-derived plasmid pBT233 in *Bacillus subtilis*. *Gene*, **136**, 1-12.
2. de la Hoz, A. B., Ayora, S., Missselwitz, R., Speck, Ch., Welfle, K., Murayama, K. *et al*. (2001). Analysis of the DNA binding domain of the ω_2 regulator protein from the broad-host range *Streptococcus pyogenes* plasmid pSM19035. *FEBS Letters*, **505**, 436-440.
3. de la Hoz A., B., Ayora, S., Sitkiewicz, I., Fernandez, S., Pankiewicz, R., Alonso, J. C. & Ceglowski, P. (2000). Plasmid copy number control and better-than-random segregation genes of pSM19035 share a common regulator. *Proc. Natl Acad. Sci. USA*, **97**, 728-733.
4. Brantl, S. & Wagner, E. G. H. (1997). Dual function of *cop*R gene production plasmid pIP501. *J. Bacteriol.* **179**, 7016-7024.
5. Gerdes, K. (2000). Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *J. Bacteriol.* **182**, 561-572.
6. Corpet, F., Gouzy, J. & Kahn, D. (1998). The ProDom database of protein domain families. *Nucl. Acids Res.* **26**, 323-326.
7. Hooft, R. W. W., Vriend, C., Sander, C. & Abola, E. E. (1996). Errors in protein structures. *Nature*, **381**, 272.
8. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138.
9. Rafferty, J. B., Somers, W. S., Saint-Girons, I. & Phillips, S. E. V. (1989). Three-dimensional crystal structures of *Escherichia coli* Met repressor with and without corepressor. *Nature*, **341**, 705-710.
10. Somers, W. S. & Phillips, S. E. V. (1992). Crystal structure of the Met repressor-operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. *Nature*, **359**, 387-393.
11. Raumann, B. E., Rould, M. A., Pabo, C. O. & Sauer, R. T. (1994). DNA recognition by β-sheets in the Arc repressor-operator crystal structure. *Nature*, **367**, 754-757.
12. Schildbach, J. F., Milla, M. E., Jefferey, P. D., Raumann, B. E. & Sauer, R. T. (1995). Crystal structure, folding and operator binding of the hyperstable Arc repressor mutant PL8. *Biochemistry*, **34**, 1405-1412.
13. Schildbach, J. F., Karzai, A. W., Raumann, B. E. & Sauer, R. T. (1999). Origins of DNA-binding specificity: role of protein contacts with the DNA backbone. *Proc. Natl Acad. Sci. USA*, **96**, 811-817.
14. Waldburger, C. D., Schildbach, J. F. & Sauer, R. T. (1995). Are buried salt-bridges important for protein stability and conformational specificity? *Nature Struct. Biol.* **2**, 122-128.
15. Gomis-Rüth, F. X., Solà, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R. *et al*. (1998). The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J.* **17**, 7404-7414.
16. Burgering, M. J. M., Boelens, R., Gilbert, D. E., Berg, J. N., Knight, K. L., Sauer, R. T. & Kaptein, R. (1994). Solution structure of dimeric Mnt repressor (1-76). *Biochemistry*, **33**, 15036-15045.
17. Berg, J. N., van Ophensden, J. H. J., Burgering, M. J. M., Boelens, R. & Kaptein, R. (1990). Structure

of Arc repressor in solution: evidence for a family of β-sheet DNA-binding proteins. *Nature,* **346**, 586-589.
18. Nooren, I. M. A., Kaptein, R., Sauer, R. T. & Boelens, R. (1999). The tetramerization domain of the Mnt repressor consists of two right-handed coiled coils. *Nature Struct. Biol.* **6**, 755-759.
19. Cordes, M. H. J., Walsh, N. P., McKnight, C. J. & Sauer, R. T. (1999). Evolution of a protein fold *in vitro*. *Science,* **284**, 325-327.
20. Phillips, S. E. V. (1994). The β-ribbon DNA recognition motif. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 671-701.
21. Raumann, B. E., Rould, M. A., Pabo, C. O. & Sauer, R. T. (1994). Major groove DNA recognition by β-sheets: the ribbon-helix-helix family of gene regulatory proteins. *Curr. Opin. Struct. Biol.* **4**, 36-43.
22. Suzuki, M. (1995). DNA recognition by a β-sheet. *Protein Eng.* **8**, 1-4.
23. Nooren, I. M. A., Rietveld, A. W. M., Melacini, G., Sauer, R. T., Kaptein, R. & Boelens, R. (1999). The solution structure and dynamics of an Arc repressor mutant reveal premelting conformational changes related to DNA binding. *Biochemistry,* **38**, 6035-6042.
24. Knight, K. L. & Sauer, R. T. (1989). Identification of functionally important residues in the DNA binding region of the Mnt repressor. *J. Biol. Chem.* **264**, 13706-13710.
25. Knight, K. L. & Sauer, R. T. (1992). Biochemical and genetic analysis of operator contacts made by residues within the β-sheet DNA binding motif of Mnt repressor. *EMBO J.* **11**, 215-223.
26. Murayama, K., de la Hoz, B. A., Alings, C., Lopez, G., Orth, P., Alonso, J. C. & Saenger, W. (1999). Crystallization and preliminary X-ray diffraction studies of *Streptococcus pyogenes* plasmid pSM19035-encoded ω transcriptional repressor. *Acta Crystallog. sect. D,* **55**, 2041-2042.
27. Dauter, Z., Dauter, M. & Rajashankav, K. R. (2000). Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallog. sect. D,* **56**, 232-237.
28. Bhwon, A. S., Mole, J. E., Hunter, F. & Bennet, J. C. (1980). High-sensitivity sequence determination of proteins quantitatively recovered from sodium dodecyl sulfate gels using an improved electrodialysis procedure. *Anal. Biochem.* **103**, 184-190.
29. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
30. Collaborative Computation Project No. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.
31. Sheldrick, G. M. (1990). Phase annealing in SHELX-90: direct method for larger structures. *Acta Crystallog. sect. A,* **46**, 467-473.
32. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110-119.
33. Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system (CNS): a new software system for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.
34. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
35. Merritt, E. A. & Bacon, D. J. (1997). Raster3D photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.

Edited by R. Huber

(Received 6 June 2001; received in revised form 1 October 2001; accepted 3 October 2001)