The anti-toxin ParD of plasmid RK2 consists of two structurally distinct moieties and belongs to the ribbon-helix-helix family of DNA-binding proteins

Monika OBERER*, Klaus ZANGGER†, Stefan PRYTULLA‡ and Walter KELLER*1

*Institut für Chemie, Arbeitsgruppe Strukturbiologie, Karl-Franzens-Universität Graz, Heinrichstrasse 28, A-8010 Graz, Austria, †Institut für Chemie, Bereich Organische Chemie, Karl-Franzens-Universität Graz, Heinrichstrasse 28, A-8010 Graz, Austria, and ‡m-phasys GmbH, Vor dem Kreuzberg 17, D-72070 Tübingen, Germany

INTRODUCTION
Plasmid-addiction systems are genetic entities of low-copy-number plasmids (for reviews, see [1–6]) that selectively kill plasmid-free cells and thus mediate stable inheritance of the plasmid in bacteria. They consist of an operon coding for two small proteins, acting as a toxin and an anti-toxin. The comparatively less stable anti-toxin counteracts toxicity by forming a toxin–antidote complex and exhibits negative autoregulatory activity through DNA binding to its promoter region. Plasmid-protected cells cannot produce the antidote, leaving the stable toxin active through DNA binding to its promoter region. Plasmid-toxin–antidote complex and exhibits negative autoregulatory function by binding to its own promoter P

The uniformly 15N- and 13C-labelled ParD protein was expressed and purified using a slightly modified version of the procedure described in [8], in order to avoid proteolytic degradation. In

Abbreviations used: NOE, nuclear Overhauser enhancement; HSQC, heteronuclear single quantum coherence; CSI, chemical-shift index; TFE, 2,2,2-trifluoroethanol; MALDI–TOF MS, matrix-assisted laser-desorption ionization–time-of-flight MS; 2D, two-dimensional; 3D, three-dimensional.

1 To whom correspondence should be addressed (e-mail walter.keller@uni-graz.at).
brief, the induced cells were harvested, lysed by sonication and after centrifugation the soluble fraction was heated to 85–90 °C for 3–5 min. After slowly cooling to room temperature and centrifugation the degraded impurities remained in the insoluble pellet, whereas refolded ParD remained in the soluble supernatant and was purified according to the described protocol [8], comprising a heparin column chromatography step followed by anion-exchange and size-exclusion chromatography. Uniformly labelled samples were produced in minimal medium (6.8 g/l Na₂HPO₄/3.0 g/l KH₂PO₄/0.5 g/l NaCl) containing 1.5 g/l (¹³¹⁵NH₄)₂SO₄ and 2.0 g/l [¹³]Cglucose as the sole nitrogen and carbon sources, respectively. The medium was complemented with 1 μg/l biotin, 1 μg/l thiamin and 1 ml of 1000 x microsols [1 M MgCl₂/150 mM CaCl₂/20 mM FeCl₂/50 mM H₃BO₃/150 μM CoCl₂/800 μM CuCl₂/1 mM MnCl₂/1.5 mM ZnCl₂/15 μM (NH₄)₂-MoO₃/2₄₋₄H₂O).

The molecular mass of the protein was determined by matrix-assisted laser-desorption ionization–time-of-flight MS (MALDI–TOF MS) on a Shimadzu Kompakt MALDI II mass spectrometer. ParD (0.4 mg/ml) was applied to a reversed-phase HPLC column and eluted with an acetonitrile gradient and 0.1 % trifluoroacetic acid. α-Cyano-4-hydroxy-cinnamic acid was used as the matrix.

The NMR samples contained about 0.6 mM ParD protein based on the molecular mass of monomeric ParD in 90 % buffer (20 mM phosphate, pH 6.0/50 mM KCl) and 10 % ²H₂O.

**NMR spectroscopy**

All NMR measurements were performed at 30 °C on a Varian Unity INOVA 600 MHz spectrometer equipped with a triple-resonance z-gradient probe. The assignment of ¹H, ¹³N and ¹³C resonances was based on the following experiments: two-dimensional (2D) ¹H-¹³N-heteronuclear single quantum coherence (HSQC), 2D TOCSY, 3D HNCA, 3D CBCA(CO)NH, 3D HNCO, 3D HN(CA)CO, ¹³C-edited TOCSY-HSQC and 3D HCCH-TOCSY (for a review, see [21]). The correctness of the NMR backbone assignments was confirmed by sequential nuclear Overhauser enhancements (NOEs) in the ¹³C-edited NOESY-HSQC spectra. The ¹H chemical shifts were calibrated relative to 2.2-dimethyl-2-silapentane-5-sulphonate, sodium salt; ¹³N and ¹³C shifts were indirectly referenced [21]. The data were processed using NMRPipe software [22], and visualized and analysed with the XEASY program suite [23] as well as NMRView [24]. ¹H-¹H and ¹H-¹³C coupling constants were calculated from the integrated peaks of an HNHA experiment. ¹³N longitudinal (T₁) and transverse (T₂) relaxation times were obtained by measuring peak heights and using the two-parameter fit \( T(\tau) = I₀ \exp(-\tau/T) \). The ¹H-¹³N-heteronuclear NOEs were obtained from the ratio of peak intensity for ¹H-saturated and -unsaturated spectra (saturation time, 2.5 s). Slowly exchanging NH protons were detected by recording two ¹H-¹³C-HSQC spectra after dissolution of lyophilized ¹³C-labelled ParD protein (20 mM phosphate, pH 6.0/50 mM KCl) in ²H₂O after 45 and 120 min.

**Structure prediction and sequence alignments**

A multiple sequence alignment of ParD, other known members of the ribbon-helix-helix family of proteins and the antidote PsaA [25] was generated using CLUSTAL W (version 1.8, June 1999) [26] with the following parameters: weight matrix, blosum; gap-opening penalty, 15; gap-extension penalty, 0.05; hydrophilic gaps and residue-specific gap penalties, on; hydrophilic residues, GPSDQERK. In addition, secondary-structure predictions and database threading were performed using the programs PSIPred [27] and GenTHREADER [28]. GenTHREADER yielded an alignment of the N-terminal part of ParD (Met¹–Lys⁵⁰) with CopG protein (PDB code, 2cpg [29]). This alignment, however, contained a 3-amino-acid gap inserted between helices A and B. Closing this gap resulted in an intact short turn and conserved hydrophobic positions in helix B, in agreement with the multiple alignment.

**Conformational changes under hydrophobic conditions**

The influence of 2,2,2-trifluoroethanol (TFE) was monitored by titrating TFE into the ¹⁵N-labelled sample to reach final concentrations of 1, 3, 5, 10 and 15 %. TFE. The normalized shift differences were calculated for each residue that could be unambiguously identified using the equation:

\[
\text{norm} \Delta \delta = \Delta \delta \cdot ||H\| + 0.2 \cdot \Delta \delta ||¹⁵N||
\]

where \( \Delta \delta ||H|| \) and \( \Delta \delta ||¹⁵N|| \) are the absolute values of the chemical-shift differences of amide protons and nitrogen atoms, respectively. Unlabelled ParD protein at a final concentration of 23 μM, in 20 mM potassium phosphate buffer, pH 7.5/20 mM KCl, incubated with increasing amounts of TFE (0–80 %), was used for CD measurements carried out on a Jasco J-715 spectropolarimeter. Five averages were taken for each spectrum with the following parameters: step resolution, 0.2 nm; speed, 50 nm/min; response, 1 s; bandwidth, 1.0 nm. The secondary-structure contents were estimated using the secondary-structure estimation program provided by Jasco.

**RESULTS**

**Protein purification**

Expression and purification of labelled and unlabelled ParD were performed essentially as described in [8] except that a heat-denaturation and refolding step was introduced prior to the first chromatography step. It has been shown previously by gel-shift analysis [17], CD spectroscopy and differential scanning calorimetry [8] that ParD regains its DNA-binding activity and its native fold after heat treatment. The existence of the dimeric form of ParD was verified using glutaraldehyde cross-linking and gel-permeation chromatography (results not shown). The susceptibility to proteolytic degradation formed a serious obstacle to NMR spectroscopic investigations because data had to be collected with several preparations of ParD and the integrity of the protein had to be checked frequently by recording ¹⁵N-HSQC spectra. The initial experiments for the assignment of the backbone were carried out with ParD purified according to the established protocol. All subsequently performed NMR spectra were obtained with heat-treated ParD, which exhibited stability and intactness even over extended measurement times. The agreement of the chemical shifts between both sets of experiments indicated clearly that the native fold of ParD was recovered after heat treatment.

The integrity and uniform labelling of ParD protein were confirmed by MALDI–TOF MS (Figure 1). The experimentally obtained molecular mass of 9467 Da corresponds to a fully ¹⁵N-labelled protein devoid of Met¹. The theoretical molecular mass was 9470 Da.

**Resonance assignment and structural features**

For ParD protein, 79 out of 80 possible backbone amide resonances (83 residues minus Pro⁵⁷, the missing Met¹ and N-terminal Ser¹) in the ¹⁵N-HSQC spectrum were assigned un-
ParD is a ribbon-helix-helix DNA-binding protein

Figure 1 MALDI–TOF spectra of unlabelled and $^{13}$C- and $^{15}$N-labelled ParD

The mass spectra of purified ParD show the main peaks at $m/z = 9126$ and 9467 Da, corresponding to the integer unlabelled and labelled proteins, respectively. %Int, relative peak intensity.

Figure 2 2D $^1$H-$^1$N-HSQC spectrum of ParD

The spectrum of 0.6 mM ParD in 90% buffer (20 mM phosphate, pH 6.0/50 mM KCl) and 10% $^2$H$_2$O is shown with $^1$H-$^1$N resonance assignments.

ambiguously (Figure 2). The dimeric form is symmetrical, because only one set of signals was observed. The N-terminal residue Arg could not be found in the $^{15}$N-edited spectra but was observed in $^{13}$C-edited spectra. For the other nuclei the percentage of assignment was 98% for C$\alpha$, 97% for C$\beta$ and C$\gamma$. The side-chain assignment was also essentially complete where the missing resonances could be attributed mainly to the less well-ordered C-terminal part of the protein. A table of the chemical shifts assigned has been deposited in the BioMagRes data bank (http://www.bmrb.wisc.edu) under the accession number 4792.

The secondary-structure assignment of ParD is based on the consensus chemical-shift index (CSI) method, including $^{13}$Ca, $^{13}$C$\beta$, $^{13}$C$\gamma$ and Hz shifts [30]. These data show very clearly the formation of a $\beta$-strand in the N-terminal region (Leu–Thr) followed by $\alpha$-helical stretches that are interrupted by short loop regions (Figure 3). The first three $\alpha$-helical regions were well defined in the CSI data, reaching from Gln to Ala, Lys to Leu and Ala to Leu, respectively, which make 35% of the $\alpha$-helical regions and 7% of the $\beta$-sheet content. Consensus chemical shifts and NOE data (Figures 3B and 3C) indicate that the C-terminal region is less well ordered or undergoing rapid conformational change on the NMR timescale. However, $^{13}$Ca chemical shifts show a tendency for the formation of further $\alpha$-helices beyond the well-ordered N-terminal domain.

Unambiguous $^3J_{Hz,HN}$ coupling constants characteristic of $\beta$-structures were derived for the residues Leu, Thr and Met; small coupling constants characteristic of $\alpha$-helices were found clustered in the three $\alpha$-helical regions of ParD (Figure 3B). These data are in good agreement with the predicted secondary structure (Figure 3A) reporting an N-terminal $\beta$-ribbon (Arg–Ile) followed by $\alpha$-helical stretches (Asp to Gln, Ile to Leu, Ala to Arg and Val to Leu), which correspond to 57% of the $\alpha$-helices and 5% of the extended structure. Protein-fold-recognition database searches [28] using only residues Met–Lys found a match with CopG protein, a known member of the ribbon-helix-helix family.

The $\beta$-ribbon was formed by a typical alternating hydrophobic–hydrophilic pattern of amino acids (Leu–Thr–Ile–Asp–...).
of the respective proteins with ParD are listed. Conserved residues are depicted in bold, conserved hydrophobic residues are shaded in light grey, other residues are important for forming the hydrophobic core of this folding motif. The turn between the first and second helices starts nicely into this pattern of conserved hydrophobic residues, as nicely in the context of conserved hydrophobic residues, as demonstrated in Figure 4. From the known 3D structures of Arc, Mnt and CopG [20, 31–33], it can be seen clearly that these residues are important for forming the hydrophobic core of this folding motif. The turn between the first and second helices starts in all cases with a glycine residue displaying the typical GXT pattern (Figure 4). The N-terminal part of PasA (Met 1–Trp 46) fits nicely into this pattern of conserved hydrophobic residues, as demonstrated in Figure 4. After a gap of 5 amino acids corresponding to Gln 57–Thr 58 in ParD, the C-terminal sequence aligns well with ParD. The sequence identity/homology between PasA and ParD were 27/50 % for the N-terminal domain (Arg 3–Trp 48) and 26/47 % for the whole sequence, respectively. These values differ slightly from the ones published previously [25] due to minor deviations in alignment. Identities and homologies between ParD and members of the ribbon-helix-helix family were calculated and showed surprisingly low sequence identities but high homologies (Figure 4).

Figure 3 Summary of the structural information about ParD obtained from NMR

(A) The amino acid sequence of ParD and its predicted (Pred) secondary structure [27]. The arrow represents the β-strand region and bars represent α-helical regions. (B) Observed short-range NOEs typical of α-helices and β-structures are depicted, with line thickness representing the strength of NOEs; NH–NH resonances that were indistinguishable due to overlapping signals are tagged with *. The amide signals present 45 min after dissolving the lyophilized protein in 2H2O are represented by #. NH–NH coupling constants typical for extended and helical structural elements are represented by □ and ■, respectively. (C) The extracted data from chemical-shift analysis with the consensus CSI displayed in the bottom row.

Figure 4 Alignment of ParD (Met 1–Trp 46) with the ribbon-helix-helix proteins CopG, Arc and Mnt and the closest antidote relative PasA

The secondary-structure diagram was derived from the 3D structure of CopG protein. Conserved residues are depicted in bold, conserved hydrophobic residues are shaded in light grey, other conserved residues are shaded in dark grey. The highly conserved turn connecting helices A and B is shown in italics. Sequence identities (Id) and homologies (Ho) of the pairwise alignments of the respective proteins with ParD are listed.

Met-Thr*). Further hydrophobic residues at conserved positions were Leu 16, Ala 17, Ala 20, Ala 21 and Leu 25. These corresponded to Leu 17, Met 18, Ala 20, Leu 22 and Leu 26 in CopG, Val 22, Val 23, Ala 24, Ile 37 and Val 41 in Arc and Leu 19, Val 22, Ala 25, Leu 28 and Val 39 in Mnt (Figure 4). From the known 3D structures of Arc, Mnt and CopG [20, 31–33], it can be seen clearly that these residues are important for forming the hydrophobic core of this folding motif. The turn between the first and second helices starts in all cases with a glycine residue displaying the typical GXT/S pattern (Figure 4). The N-terminal part of PasA (Met 1–Trp 46) fits nicely into this pattern of conserved hydrophobic residues, as demonstrated in Figure 4. After a gap of 5 amino acids corresponding to Gln 57–Thr 58 in ParD, the C-terminal sequence aligns well with ParD. The sequence identity/homology between PasA and ParD were 27/50 % for the N-terminal domain (Arg 3–Trp 48) and 26/47 % for the whole sequence, respectively. These values differ slightly from the ones published previously [25] due to minor deviations in alignment. Identities and homologies between ParD and members of the ribbon-helix-helix family were calculated and showed surprisingly low sequence identities but high homologies (Figure 4).

Amide hydrogen-exchange data indicated that the most protected area ranged from Leu 22 to Phe 36, which resides within the cluster of slowly exchanging amides covering most of helices A and B. This region, together with the hydrophobic residues of the β-ribbon, forms the central core and the dimerization interface of the protein.

To characterize the dynamic properties of ParD, relaxation measurements were performed with 15N-labelled protein. Heteronuclear NOEs sensitive to rapid internal motions were calculated from the ratio of the peak heights in the spectra acquired with and without 1H saturation. The missing resonances (Ser 15, Leu 18, Met 22, Asn 25, Arg 26, Asn 28, Asp 29 and Glu 32 for T1, Asp 31, Ala 32, Gln 37 and Leu 32 for T2, plus Ile 37, Leu 31 and Ile 32 for the heteronuclear NOE) could not be obtained reliably owing to spectral overlap (Figure 5). The negative NOE values for the C-terminal domain (Asp 30–Ala 31) indicate the presence of significant flexibility in this region. This is in excellent agreement with the 1H-T1 and 15N-T2 relaxation times, which clearly show two structurally different moieties ranging from Leu 4 to Ile 37 for the well-structured N-terminus and from Asn 39 to Ala 83 for the remaining flexible C-terminal region. Cα chemical-shift data
ParD is a ribbon-helix-helix DNA-binding protein

Figure 5 Relaxation data for the $^{15}$N nuclei of ParD
$^{15}$N-$T_1$, $^{15}$N-$T_2$ and $[^1H]^{15}$N-heteronuclear NOE data are plotted as a function of residue number. show a down-field shift from the random-coil values for residues in the C-terminus for residues 58–64, 66, 68–72, 74–81 and 83, indicating a tendency for the formation of helical structures. The deviations are, however, smaller than in the N-terminal region, which again supports the presence of a predominantly unstructured C-terminal domain. As the C-terminal fragment of ParD appears to be responsible for ParE binding [17], we expect this domain to become ordered upon anti-toxin–toxin complex formation.

TFE-induced secondary-structure changes

TFE was added to a ParD solution and changes in the secondary structure were monitored by CD spectroscopy. Upon an increase in the TFE concentration (from 0 to 5, 10, 20, 30, 40, 50, 60, 70 and 80 %) the $\alpha$-helical content rose gradually (from 37 to 44, 50, 53, 55, 62, 65, 68, 70 and 72 %) with an isodichroic point at 200.4 nm (Figure 6). The same titration was investigated using NMR spectroscopy. With increasing TFE concentrations the amide proton line widths were broadened in the acquired 2D $[^1H]-^{15}$N-HSQC spectra and led to high spectral overlap, thus preventing a more detailed study of the dynamic behaviour (results not shown). As can be seen from the combined chemical-shift changes (Figure 7), addition of TFE influenced mainly the residues starting from Gly$^{20}$, resulting in major chemical-shift differences. This indicates a transition from the more flexible and random-coil C-terminal moiety to an $\alpha$-helical domain consistent with the CD experiment.

DISCUSSION

The secondary-structure assignment based on the CSI and sequential NOEs clearly shows the existence of a short N-terminal $\beta$-sheet followed by three $\alpha$-helices. The overall secondary-structure content, as estimated from the consensus CSI, matched those data determined by CD spectroscopy, namely 37 % $\alpha$-helical regions and 12 % $\beta$-structure [8]. The small discrepancy in $\beta$-content arises from the difficulty in getting accurate estimates of the $\beta$-structure by CD spectroscopy due to the less-pronounced contributions of $\beta$-structures compared with the $\alpha$-helical signal.

Figure 6 Effect of the hydrophobic agent TFE
The far-UV CD spectra of ParD (23 $\mu$M) in 20 mM potassium phosphate buffer, pH 7.5/20 mM KCl, incubated with increasing amounts of TFE (0%, thin line; 20%, ○; 80%, thick line), were monitored.

Figure 7 Titration of ParD with TFE
The chemical-shift differences in the resonances observed in $[^1H]-^{15}$N-HSQC spectra between ParD in phosphate buffer and ParD in buffer plus 15 % TFE are plotted against residue number. Missing bars are either due to spectral overlap or ambiguous shifts after titration.
Second-order-structure prediction yields a significantly higher α-helical content (57%), which results from an additional helix (Figure 3A). Experimental 1H NMR chemical-shift data, however, only indicate a propensity for α-helical regions in the C-terminus.

We tried to induce this ‘missing’ helicity by the addition of TFE, which is known to stabilize helical structures but also to disrupt native tertiary structure by modifying non-local hydrophobic interactions in polypeptides and proteins [34–36]. The structural changes of ParD, which are induced by TFE addition as observed in the near-UV CD spectrum, indicate the formation of additional helical structure. As monitored by NMR spectroscopy, C-terminal residues appear to be involved mainly in these changes. However, due to extensive line-broadening and resulting overlaps the NMR data were only interpretable up to a TFE concentration of 15%. Therefore a more detailed assignment of the TFE-induced changes to particular regions of the polypeptide chain could not be carried out.

ParD protein can be aligned with CopG and Arc, which belong to a family of small prokaryotic repressors displaying a homodimeric ribbon-helix-helix fold in their 3D structures. This distribution of secondary-structural elements is corroborated by our experimental data derived from NMR solution studies (Figure 3).

It has been hypothesized previously that ParD might belong to this fold family [37], despite lack of experimental evidence and the low sequence homology between ParD and the mentioned repressor proteins. On the basis of the alignment of secondary-structure elements and the pattern of hydrophobic residues it is now evident that ParD belongs to the ribbon-helix-helix fold family and that the mode of DNA binding is analogous to that seen in the crystal structures of the DNA complexes of Arc [31], MetJ [38] and CopG [20]. Therefore we expect the short intermolecular β-ribbon to insert into the major groove of the DNA at the specific binding site.

Second-order-structure predictions for the antidotes Phd, CcdA, PemI [39], RelB and PasA (results not shown) suggest the presence of a β-ribbon followed by α-helices. In analogy, antidote proteins of these functional homologous addiction systems most probably belong to this fold family as well, even though the systems are only related weakly at the protein-sequence level. Homology has been reported for ccdA and pemI [40] as well as for mazE (chpA1) [41], and within the large family of relBE systems [14]. Amongst all the characterized antidote proteins, the closest related to ParD is PasA from the Gram-negative acidophilic bacterium Thiothrix ferroxidans [25]. The complete amino acid sequence of PasA (74 residues) can be aligned with that of ParD and, consequently, the ribbon-helix-helix motif is proposed for the C-terminal domain (Figure 4).

The relaxation measurements together with the chemical-shift data indicate clearly that the ParD protein is divided into two domains: a well-ordered N-terminal domain containing the βαβ secondary-structure elements and a very flexible C-terminal domain.

These data are in excellent agreement with the two biological roles of ParD, namely its repressor and antidote functions, and with biochemical data showing that the DNA-binding function resides in the N-terminal part of the protein while the toxin-binding activity is suggested to reside in the C-terminal part [17]. Similar to the conformational changes reported of the CcdA/CcdB and PhD/Doc modules upon complex formation [42,43], we expect the C-terminal part of ParD to acquire a higher degree of order upon binding to the ParE protein.

The availability of an increasing number of complete genomes has revealed the presence of addiction systems on many chromosomes of Gram-positive and Gram-negative bacteria as well as the Archaea. Although their function has not been elucidated they are bound to convey some evolutionary advantage to the host. The low sequence homology between antidotes of the diverse addiction systems may reflect the necessity to develop a high specificity (for DNA and target binding) in order to be useful to the host organism.

We thank Don Helinski for expression strains, Robert Konrat and Heinz Sterk for fruitful discussions, and Edith Kramer for help with preparation of the manuscript. This work was supported by the Fonds zur Förderung der Wissenschaftlichen Forschung (grant no. P12559-GEN). K.Z. thanks the Austrian Academy of Sciences for an APART Fellowship.

REFERENCES

ParD is a ribbon-helix-helix DNA-binding protein


Received 9 July 2001/10 October 2001; accepted 25 October 2001