Structure-based functional analysis of plasmid R6K’ s replication protein:
Key amino acids at the π/DNA interface

Selvi Kunnimalaiyaan, Sheryl A. Rakowski and Marcin Filutowicz*

Department of Bacteriology, University of Wisconsin
420 Henry Mall, Madison, Wisconsin 53706

Running title: Key amino acids for π binding and function

* Corresponding author. Mailing address: Department of Bacteriology, University of Wisconsin, 420 Henry Mall, Madison, Wisconsin 53706. Phone (608) 262-6947. Fax: (608) 262-9865. E-mail: msfiluto@facstaff.wisc.edu
**Abstract:** In previous work, we characterized the bases in an iteron of plasmid R6K that are important for the binding of $\pi$ protein monomers and dimers. Here, we investigate six amino acids of $\pi$, encoded by $pir$, hypothesized to be important for DNA contact: Ser71, Try74, Gly131, Gly211, Arg225 and Arg254.

**Background:** Rep proteins activate replication origins ($ori$s) by binding to tandem repeats in the DNA called iterons (3, 8, 13). When the sequences of iterons and rep genes are compared across families, striking similarities are observed (2, 10, 16-18). A breakthrough in Rep/iteron studies came when Komori *et al.* published the crystal structure of a RepE monomer (from F plasmid) in complex with its iteron DNA (10). The structure revealed that the Rep monomer contained two DNA-binding domains, an N-terminal WH1 (winged helix 1) and a C-terminal WH2, with both domains used for iteron binding. A sequence alignment of several Reps presented by this group and reprinted with permission (13) is a useful reference for identifying possible conserved amino acids in replication initiators including RepE and $\pi$ protein (described below). Based on the homologies observed in this class of proteins, Sharma *et al.* subsequently reported theoretical 3-dimensional models for several Reps (18). The accuracy of these models remained to be tested.

In a minimal R6K replicon, an $ori$ called $\gamma$ is activated by the Rep protein, $\pi$, which is encoded by the $pir$ gene (Fig. 1A) (4, 6, 7, 9, 15). $\pi$ exists in both monomeric and dimeric forms with monomers being activators of replication and dimers functioning both as replication inhibitors and autorepressors of $pir$. The latter function is mediated by $\pi$ dimers binding to inverted repeats in the promoter that share sequence similarity with the iterons. $\pi$ monomers show little to no affinity for the inverted repeats but bind to an iteron better than $\pi$ dimers (20). Insights into how $\pi$ achieves these similar but functionally distinct interactions were sought from analyses of Rep structure and the contacts that occur at the protein/DNA interface.

In previous work from this lab, we determined the importance of individual bases in the iteron for the binding of $\pi$ monomers and $\pi$ dimers (14). Beyond this, we knew of no other experimental data addressing the contact residues of $\pi$ protein and the DNA; a solved $\pi$/iteron co-crystal structure was not available at that time. In this study, we investigated the
protein/DNA interface again, this time with an interest in identifying amino acids of \( \pi \) that contact DNA. By analyzing the theoretical structural model of the \( \pi \) monomer/iteron complex, we identified approximately two dozen amino acids as candidates for DNA contact and narrowed the list down to six for further study (Fig. 1B). Four (Ser71, Gly131, Arg225 and Arg254) appeared to be positionally conserved within the RepE/iteron co-crystal structure (Ser75, Gly125, Arg206 and Arg233, respectively). Tyr74 was strongly suspected to contact a thymine residue in the half of the iteron bound only by \( \pi \) monomers (14). Lastly, Gly211 was of interest because it fell in an unstructured loop region positioned at the \( \pi \)/DNA interface; this region appears to be unique among modeled Rep proteins and its function, if any, is unknown. We hypothesized that these amino acids might be important for DNA binding and, consequently, the regulatory control of DNA replication. To test this, the six amino acids were changed (individually) via site-directed mutagenesis to alanines, and the resulting \( \pi \) variants were characterized in vivo and in vitro.

\( \pi \) variants display altered replication control in vivo: Each of the alanine substitution variants of \( \pi \) was tested for its ability to support the replication of a “suicide” \( \gamma \) ori plasmid, pFW25 (22), that does not encode its own \( \pi \). Plasmids expressing wt \( \pi \) (pJWW801) (12) and variant \( \pi \) proteins carried pir genes under the control of a T7 promoter. The plasmids were transformed into the \( E. coli \) strain, BL21(DE3) that carries a T7 RNA polymerase gene in the chromosome, which is IPTG-inducible. Western blots were done to confirm the expression of the \( \pi \) variants (Fig. 2 and data not shown). As expected, although IPTG induction did elevate \( \pi \) levels, detectable levels of the protein were also observed in the absence of IPTG due to "leaky" transcription. Plasmid-bearing BL21(DE3) strains that were grown in the absence of IPTG were made competent and used in transformation experiments with pFW25 DNA. We observed that the control strain producing wt \( \pi \) (from pJWW801) yielded colonies on selective medium (for pFW25) whether or not pir expression was induced by adding 0.1 mM IPTG to the plates. In contrast, when wt pir (on pJWW801) was replaced by an alanine substitution mutant, repeated attempts to obtain colonies transformed by pFW25 were unsuccessful for five of the six \( \pi \) variants tested; only \( \pi \)G211A produced colonies. As noted above, amino acid 211 falls in a unique unstructured loop region of the protein, the function of which cannot be inferred by homology modeling. These results are unlikely to be attributable to variation in the protein expression levels of our constructs for two reasons. First, repeated Western blots
failed to demonstrate reproducible differences in expression. Second, even if the variation observed had been reproducible, for each variant at least one level of $\pi$ expression (induced and/or uninduced) fell within the range of wt $\pi$ expression that allowed successful establishment of the $\gamma$ ori plasmid (Fig. 2).

In addition to being a replication initiator, $\pi$ can also inhibit replication (5). To test the inhibitor activity of the six $\pi$ variants, we performed another two-plasmid transformation assay, similar to the one described above except that the suicide plasmid was replaced with a complete $\gamma$ ori replicon that encodes its own wt $\pi$ (pFL129) (21). The $\pi$wt-producing pJWW801 plasmid is compatible with the pFL129 $\gamma$ ori replicon when low levels of $\pi$ protein are produced. However, the overproduction of wt $\pi$ inhibits the replication of $\gamma$ ori plasmids (e.g. pFL129), presumably due at least in part to increased competition for iteron binding by $\pi$ dimers (13).

Using this in vivo system, we compared the inhibitory activities of $\pi$ variants to the activity of wt $\pi$ (Fig. 3). The data show that when wt $\pi$ or $\pi$G211A was overproduced, the number of transformants obtained that exhibit the antibiotic resistances of both plasmids was reduced by approximately 50% compared to the uninduced control. In contrast, when any of the other variants was overproduced, the number of colonies was reduced by more than 90%. This observation suggests that these $\pi$ variants inhibit replication more strongly than wt $\pi$. The ability of a related assay to determine the relative strength of replication inhibition has been attributed to several factors (1). These factors include the fact that levels of chloramphenicol resistance, produced by the $\gamma$ ori construct, are plasmid copy number dependent, and the copy numbers of individual cells can vary in a population. Notably, in both in vivo assays, conditions that reduced the number of colonies (inhibited $\gamma$ ori plasmid replication) also led to reduced colony sizes (this work and ref. 1, unpublished observation). Taken together, the positive and negative functional assays indicate that five of the six amino acids of $\pi$ under study contribute to DNA replication control, consistent with predicted roles for these residues in iteron binding.

$\pi$ variants display altered binding patterns in vitro: To directly examine the DNA binding of the variants, DNA fragment containing a single iteron from pRK1 (11) was combined with purified, histidine-tagged $\pi$ variants over a range of protein concentrations. The preparation of
DNA fragment and protein purification were essentially as previously described (14, 23) except the latter used Ni Sepharose™ 6 Fast Flow resin (GE Healthcare) in the place of Ni-NTA agarose (Qiagen). Samples (15 µl final volume) for binding reactions were prepared in 1X binding buffer (2 mM Tris-HCl (pH 7.5), 0.6 mM MgCl₂, 0.1 mM EDTA and 10 mM potassium glutamate) with 500 pg labeled DNA, 65 ng poly(dI-dC) and 1 µl diluted protein. Proteins were diluted in TGE buffer (14) to concentrations of 25, 50, 100 and 200 ng/µl. Binding reactions were incubated for 15 min. at room temperature and samples were analyzed by EMSA (11). A representative gel is shown along with corresponding quantification data from lanes with 200 ng of each π variant (Fig. 4). The results demonstrate that alanine substitutions at positions 71, 74 and 131 produced binding deficits in π monomers but not π dimers. For Y74A the defect was particularly severe suggesting that this tyrosine is critical for monomer binding. Alanine substitutions at positions 225 and 254 affected the binding of both monomers and dimers. Interestingly, although the G211A substitution in the unstructured loop behaved the most like wt in the in vivo replication assays, it showed marked reductions in both monomer and dimer binding to the DNA. We also note that a reduction in dimer binding by 50% or more (G211A, R225A, R254A) did not weaken the inhibition of plasmid replication in the in vivo assay (Fig. 3). Importantly, however, work by Urh et al. demonstrated that DNA binding proficiency is not required for a π variant to inhibit replication of a γ ori plasmid (20). To assay the ability of a π variant to inhibit replication, the system must also include wt π (or other active initiator form of π) to support replication. The π variant does not need to have DNA binding activity to inhibit replication initiation by the wt protein; it needs only a functional dimerization domain. Presumably, by forming heterodimers with the wt π, the variant protein reduces the pool of monomeric initiator.

**Comments:** The data from the in vivo functional assays and the in vitro binding assays are consistent with roles for at least five of the six targeted amino acids in contacting the DNA. Mutations in the WH1 motif (S71A, Y74A and G131A) disturbed monomer binding only; mutations in the WH2 domain (R225A, R254A) affected the binding of both monomers and dimers.
At the time this work was being conducted, only two Rep structures had been solved from protein crystals; the structures of other Reps including π could only be hypothesized based on homologies in the Rep family. Recently, however, Swan et al., published the crystal structure of π protein in complex with the iteron DNA (19). Consistent with our results, their data also demonstrate the importance of amino acids Arg225 and Ser71 in DNA contact and replication and support DNA contact by Tyr74. In contrast, no mention is made of Gly131. Perhaps this is because their structure is derived from a quadruple substitution variant of π or perhaps the G131A substitution indirectly disrupts DNA binding. The Swan paper presents evidence for DNA contact by two immediately preceding amino acids, Asp129 and Glu130. Additionally, they suggest the importance of Arg254 but it is not clear on what basis Arg254 is identified as DNA-contacting. Here, we provide data consistent with a role for Arg254 in iteron binding. Taken together we believe the parallels between our findings bode well for the utility of homology-based structural models of Rep proteins for the dissection of structure/function relationships and for generating Rep protein variants with desired characteristics.

This work was supported by National Institutes of Health grant GM40314 to M.F.

References


Wu, F., J. Wu, J. Ehley, and M. Filutowicz. 1996. Preponderance of Fis-binding sites in the R6K γ origin and the curious effect of the penicillin resistance marker on


Figure Legends

Figure 1. Elements of a minimal R6K replicon. A) The roles of π protein in the regulation of γ ori. π protein encoded by the pir gene exists in two forms: monomers (shaded dark gray) and dimers (shaded light gray). Monomers activate replication (i), dimers inhibit replication (ii) and dimers also bind to the inverted repeat to repress pir transcription (iii). π binding sites are represented by black arrows. The seven iterons of γ ori, also called direct repeats (DRs), are indicated by tandem arrows, whereas a pair of shorter inverted arrows indicate the inverted repeats (IR) in the pir operator/promoter. There is an 8th iteron just upstream of the IR. The sequence of iteron number 1 is also shown and thymine 27 is marked in bold. B) Amino acids of π selected for mutational analysis. A model of a π monomer bound to a γ ori iteron (Kunnimalaiyaan et al. 2004) with the side chains of 6 amino acids targeted by site-directed mutagenesis is shown in space-filled form. The N-terminal and C-terminal domains of π are white and green respectively. DNA strands are represented in purple (top) and gray (bottom). Some of the iteron positions believed to be important for π monomer binding are shown as spheres on the DNA strand.

Figure 2. Western analysis. Cells containing pJWW801 (producing wt π) and derivative plasmids producing π variants were grown in the absence (-) or presence (+) of IPTG (0.1 mM). 0.1 ml of cells standardized to an OD600 of 1.2 were lysed and subjected to Western analysis using anti-π antibodies. A sample of purified wt π protein was loaded as control (C).

Figure 3. Testing the replication inhibitor function of π variants (in vivo). DNA from the γ ori plasmid pFL129 was added to competent cells containing pJWW801 (producing wt π) and derivative plasmids producing π variants (grown in the absence of IPTG). Transformation mixtures were plated on selective medium lacking IPTG (uninduced) or induced with 0.1mM IPTG to overexpress π variants (including wt π). The histogram shows the number of colonies obtained when the cells were uninduced (light shading) and induced (dark shading).
Figure 4. Binding of π variants to a single iteron. A) 3’ end-labeled DNA fragment containing single iteron was incubated with increasing amounts (25, 50, 100 or 200 ng) of π (wt or variants) and subjected to EMSA. The positions of free DNA (F), and nucleoprotein complexes containing π monomers (M) or π dimers (D) are indicated. B) The percentage of DNA bound by 200 ng of π (wt or variant) is shown in histogram.
A+T

pir

operator/promoter

(i) (ii)

AAACATGAGAGCTTAGTACGTT

TTTGTAACCTCTGAATCAGCA

1

44 27 23

22

A

B

Arg 254

Arg 225

Gly 131

Tyr 74

Ser 71

Gly 211

ACCEPTED on August 30, 2017 by guest http://jb.asm.org/ Downloaded from
% bound DNA

WT S71A Y74A G131A G211A R225A R254A

variants

DNA only