### **Supplementary Material**

to

# *pyr* **RNA binding to the** *Bacillus caldolyticus* **PyrR attenuation protein. Characterization and regulation by uridine and guanosine nucleotides**

by

Casper Møller Jørgensen, Christopher J. Fields, Preethi Chander, Desmond Watt, John W. Burgner, II, Janet L. Smith and Robert L. Switzer

# **Doc. S1. Supplementary Results and Discussion**

### **Structural requirements of effectors for affecting PyrR binding to BcBL2**

To learn more about how PyrR differentiates between purine and pyrimidine nucleotides, we tested the ability of purine and pyrimidine nucleotide structural variants to activate or inhibit binding of BcBL2 to PyrR (Table S1). All of the pyrimidine nucleotides tested were capable of activating RNA binding to PyrR, which suggests that pyrimidine nucleotides with bulky additions to the pyrimidine ring (OMP and TMP) can bind to PyrR. This is in agreement with the crystal structure [1], in which this part of the pyrimidine ring does not make direct contacts with PyrR.

The structural requirements for the effects of purine nucleotides are more restrictive than for pyrimidine nucleotides. Substitution of the 2-exocyclic amino group in GMP (R5 in Table S1) with an oxo-group in XMP essentially abolished the effect of the nucleotide on RNA binding. Removal of this amino group (IMP) permitted inhibition of binding to BcBL2, but the apparent dissociation constant for binding of RNA was 10-fold lower than with GMP. AMP did not affect binding of BcBL2 to PyrR. The data suggest that both the exocyclic oxo and amino groups of GMP make functionally important contacts to PyrR.

Both 2'-GMP and 3'-GMP increased the apparent dissociation constant for BcBL2 binding by two to three-fold over the no effector level, in contrast to the 40-fold effect of 5'-GMP. These observations, together with their low intracellular

concentrations, suggest that these nucleotides do not play a significant role in regulation of *pyr* gene expression *in vivo*. 2', 3'-cGMP did not affect binding of PyrR to BcBL2, but 3', 5'-cGMP increased the apparent dissociation constant by almost eight-fold. The physiological relevance of these findings is doubtful, but the data indicate that the 2'hydroxyl group of ribose in GMP is important for binding of the nucleotide to PyrR or modulating its effect on RNA binding.

#### **Binding of BcBL2 structural variants to PyrR**

We wanted to determine whether the binding of RNA to *B. caldolyticus* PyrR exhibits the high sequence specificity expected from previous genetic and biochemical studies with *B. subtilis* PyrR. To examine this question, we used the filter binding assay to examine binding of *B. caldolyticus* PyrR to three *B. caldolyticus* PyrR binding loops containing single base substitutions (Fig. 1B) that were observed in previous gel shift studies with *B. subtilis* PyrR to have very different apparent  $K_d$  values [2].

The G726A sequence variant of BcBL2 has a changed base pair in the upper stem predicted not to affect the structure of the binding loop and does not alter a conserved nucleotide. *B. caldolyticus* PyrR bound to the G726A structural variant with the same apparent  $K_d$  as the wild type binding loop, both in the presence and absence of nucleotides (Table S2). Gel shift assay of the corresponding structural variant in *B. subtilis* (G728A) also showed tight binding to PyrR in the presence of UMP [2].

The G at position 723 in BcBL2 is part of a sequence of nucleotides in the terminal loop that is conserved in PyrR binding loops from many bacteria; replacement of this nucleotide is predicted to disrupt binding to PyrR. Filter binding studies of the *B. caldolyticus* G723A BcBL2 replacement determined that the apparent K<sub>d</sub> in the presence of UMP was increased by more than 1300-fold relative to the  $K_d$  of wild type BcBL2 (Table 6). Impaired binding of PyrR to the G723A structural variant was even more pronounced without nucleotide effectors or in the presence of 0.5 mM GMP. Gel mobility shift assays with *B. subtilis* PyrR also indicated that replacement of this G with

an A (G725A in BsBL2) disrupted PyrR binding in the presence of UMP; the variant RNA bound with an apparent  $K_d$  that was more than 800-fold larger than the  $K_d$  for the wild type BsBL2 binding loop [2].

The A724C replacement in BcBL2 (BcBL2-A724C) is expected to reduce binding by PyrR, since it, like BcBL2-G723A, changes a conserved residue in the terminal loop. In previous gel mobility shift assays with this structural variant (A726C in *B. subtilis*) and *B. subtilis* PyrR, a very high apparent  $K_d$  was observed (20  $\mu$ M in the presence of UMP [2]). However, we did not observe the same large effect on binding of the BcBL2- A724C RNA variant in the filter binding studies with *B. caldolyticus* PyrR (Table 6). In the presence of UMP, the apparent  $K_d$  was close to that of wild type BcBL2. With no effector and with GMP, on the other hand, decreased affinity of PyrR for the A724C structural variant was observed, indicating that the A724 nucleotide is indeed involved in PyrR binding to BcBL2.

The binding of the A724C structural variant of BcBL2 was investigated further using the gel shift method and *B. caldolyticus* PyrR. Strikingly, no binding of the BcBL2 A724C RNA to PyrR (up to 10 µM PyrR) was detected under the conditions used for the studies in Table 4 and Fig. 3. This was in sharp contrast to the relatively tight binding of the same RNA observed by filter binding. However, when the  $Mg^{2+}$  concentration in the electrophoresis gel was increased to 10 mM, binding of this RNA variant was detected ( $K_d$  of 3500 nM in the absence of effector, bimodal binding with  $K_{d1} = 1.2$  nM and  $K_{d2}$ >20,000 nM in the presence of 0.5 mM UMP). These values are much larger than those obtained by filter binding (in which the concentration of  $Mg^{2+}$  is also 10 mM). These observations suggest that the A724C variant RNA differs from the wild type BcBL2 in its interaction with  $Mq^{2+}$ . They also point out an unusual instance in which the filter binding and gel shift assays yielded highly divergent results even though the same RNA and PyrR were used under as close to identical conditions as possible.

# **Effects of Mg2+, pH and temperature on binding of PyrR to BcBL2**

The standard buffer used in filter binding assays of RNA binding contained 10 mM  $Mq^{2+}$ and 50 mM K<sup>+</sup>; varying the concentration of Mg<sup>2+</sup> in the range from 1 mM to 50 mM or the K<sup>+</sup> concentration between 10 mM and 300 mM did not affect binding of PyrR to BcBL2 or BcBL3 (data not shown). However, when  $Mq^{2+}$  was omitted from the binding buffer, BcBL2 RNA bound to PyrR with an affinity about 15-fold lower than when  $Mq^{2+}$ was present. Inclusion of 10 mM EDTA in the binding reaction without added  $\text{Mq}^{2+}$ further reduced the apparent affinity of PyrR for BcBL2 by about 100 fold. This documents a requirement of the system for divalent cations for RNA binding and suggests that traces of such cations may contaminate the reaction components.

The effects of changing the pH of the binding buffer in the range of pH 5.5 to 8.5 on the binding of BcBL2 to PyrR are shown in Fig. S1A. With no effector present, PyrR bound 50-fold less tightly to BcBL2 at pH 5.5 than at the standard pH of 7.5 used in this study. When UMP was present in the reaction mix, the same change in pH resulted in less than a three-fold change in the apparent  $K<sub>d</sub>$ . The addition of GMP, on the other hand, had an opposite and very large effect on binding of PyrR to BcBL2, i.e., a more than 300-fold higher apparent  $K_d$  was observed at pH 8.5 than at pH 6.5. These results indicate that when the pH is changed, a residue in PyrR that is important for binding of RNA or nucleotides undergoes altered ionization. This residue could be a histidine, since the pKa value for the imidazole side chain of N-acetyl-histidine ethyl ester is 7. No ionizable group in the GMP binding site in the crystal structure accounts for the observed pH dependence. Although none of the histidine residues in PyrR lie closer to a purine nitrogen of GMP than 8 Å [1], His82 and His21 are candidates for mediating the observed pH effect. His82 resides on a flexible loop adjacent to the nucleotide binding site, and may be proximal to the guanine base in unseen conformations of the enzyme. His21 lies in the putative RNA binding face of PyrR [1, 3], and the equivalent residue in *B. subtilis* PyrR, His22, was strongly implicated in RNA binding by the mutagenesis studies of Savacool and Switzer [4]. In *B. caldolyticus* PyrR, His21 is in the interface between dimers that forms the closed tetramer, a position in which its ionization state may influence tetramer-dimer equilibration. The intracellular pH in *B. subtilis* cells has been determined to lie in the range from 7.6 to 8.3 in growth media with pH buffered at

values from 6 to 8 [5] and is likely similar in *B. caldolyticus*. Thus, the marked differences in GMP and UMP effects on RNA binding occur at a physiologically relevant pH.

## **Analytical ultracentrifugation of PyrR and PyrR-RNA complexes**

Several sedimentation velocity studies were performed on both native and Histagged PyrR as a function of protein and salt concentrations. Since  $Mg^{2+}$  is required for RNA binding, 10 mM  $Mg^{2+}$  was also included in all cases. Figure S2 shows the results of sedimentation velocity studies on both the native (Panels A, B, and C) and Histagged (Panels D, E, and F) PyrR. The data in both cases were analyzed with the program SEDFIT [6]. Panels A and D show boundary movement as a function of time with dotted lines, the solid lines show fits to the data based on the c(s) distributions as a function of  $s_{20,w}$  shown in Panels C and F. Panels B and E show the residuals between observed and fit data. The distributions of PyrR as a function of  $s_{20,w}$  in Panels C and F indicate that nearly all of the PyrR, regardless of the source, sediments with s values near 5 S (see Figure S2 legend for further details). The calculated masses of the 5 S particles are 83 and 100 kDa for the native and His-tagged PyrR, respectively. A sedimentation velocity coefficient of 5.2 S was calculated from the Protein Data Base file for the native PyrR x-ray structure (1XZN) using the bead modeling program HYDROPRO [7], which is in good agreement with the observed  $s_{20,w}$  values. The calculated mass values are somewhat greater than the sequence mass for a tetramer (Table S3), which is likely the result of rapid, weak formation of higher order oligomers (octomers, etc.). Further characterization of these larger oligomers was not attempted. The values for  $s_{20,w}$  given in legend of Figure S2 were further corrected for protein concentration using SEDINTERP; the frictional ratio  $(f/f_0)$  was estimated from these values and the corresponding sequence masses of the tetramer (Table S3). The value of this ratio was calculated using the Teller algorithm from the program SEDNTERP v1.09 (which was obtained from http://www.rasmb.bbri.org/rasmb/ms\_dos/) and assuming 0.4 g of water per 1.0 g of protein [8]; it is about 1.28 for the oligomeric form

of both native and His-tagged PyrR, which suggests that the addition of the His-tag does not significantly alter the shape of the tetramer. Thus, we suggest that using the Histagged PyrR for the long-term titration studies shown in Figures 5 and S5 to avoid possible ribonuclease contamination is reasonable.

As noted in the Discussion, crystallographic studies on this PyrR suggest that the potential binding site for RNA is sterically hindered in the tetramer. Titration studies in the analytical ultracentrifuge (Figures 4 and 5) indicate that the PyrR-RNA complex likely consists of a dimer of PyrR and a single molecule of RNA. To investigate the tetramer-dimer equilibrium further, we undertook sedimentation equilibrium experiments with both the native and His-tagged PyrR at monomer concentrations ranging from 0.25  $\mu$ M to 25  $\mu$ M, which represents the lower end of the useful range for detecting an equilibrium distribution of PyrR in the ultracentrifuge. As part of this experiment, we also collected "approach to equilibrium" data [9] at the slowest speed used (8000 and 10,000 rpm for the native and His-tagged PyrR, respectively). Figure S3 shows the approach to equilibrium data for three different concentrations of native PyrR; each pair of plots, reading horizontally, shows the interference and absorbance data as well as (below) the calculated residuals at the PyrR concentrations shown in the legend. The first 10 absorbance and interference scans were used such that the sample was close to but not at equilibrium. (Including an excess of scans at equilibrium or very near equilibrium conditions seems to prevent an accurate estimation of particle mass, according to studies with other proteins; Burgner, J. W. unpublished observations). The data at each concentration were fitted to a single species model using SEDPHAT 4.1. Masses of 61, 71, and 82 kDa were estimated from low to high PyrR concentrations. In our hands, this method gives estimates of the mass for simple systems within 1 to 5% of the sequence mass, but the estimate of the sedimentation coefficient is much more variable, as should be expected from the lack of sharpness in the boundaries. The decrease in the weight average mass, which is calculated here, with decreasing protein concentration is consistent with a dissociation constant for the tetramer that is less than  $10^{-7}$  M.

Figure S4 shows the absorbance distribution as a function of rotational speed for native (Panel A) and His-tagged (Panel B) at equilibrium. The equilibrium data were collected at ~60 h for the slowest speed (see Figure S4 legend) and after an additional 40 h at each of the remaining speeds. These absorbance and interference (not shown) data were collected (see Figure S3, additional details) and fitted to a single species model using mass conservation. The calculated mass the native PyrR is 78.3 kDa for both the absorbance data alone and a combined set of both. The calculated mass for the His-tagged PyrR is 90.7 and 91.6 kDa for the absorbance and combined data, respectively. These values compare well with the sequence masses of the native (79.7 kDa) and His-tagged (89.2 kDa) PyrR. As indicated in the Results section, these data were also fit using a dimer-tetramer formation model using SEDPHAT and mass conservation between speeds. The estimated value for the  $K_d$  for a tetramer to dimer disassociation model is about  $3 \times 10^{-8}$  M for both PyrR forms. The interference data were not included in these calculations because blank data were unfortunately not collected; thus, corrections for changes in the baseline as a function of speed could not be made with sufficient accuracy. This is not the case for either the equilibrium absorbance data or all of the approach to equilibrium data, where time and radial independent noise can be removed algebraically [9]. Values for the global reduced  $\chi^2$ were calculated for each fit and appear in the legend of Figure S4. It would appear from these values that the dimer-tetramer model does not provide a significantly better fit to the data than a tetramer-only model. However, the decrease in observed masses with decreased PyrR concentration observed in Figure S3 is suggestive of dissociation of the tetramer. Thus, assuming the weight average molecular mass describes only a dimertetramer equilibrium, the observed 60 kDa at the lowest native PyrR concentration would predict a Kd of about 2  $\times$  10<sup>-8</sup> M, which is in reasonable agreement with that calculated from the equilibrium data.

Figure S5 shows the sedimentation velocity data from the titration of His-tagged PyrR with RNA as well as the c(s) distribution plots for interference, A260 and A278 data and the residual plots from these fits. The distribution plots obtained for each

titration point as well as for the reactants are qualitatively the same. While there are sufficient data to determine, by solving the appropriate simultaneous equations, the ratio of RNA to PyrR in the new ~4.7S peak found, the signal to noise prevented an accurate assessment of this ratio by multiwavelength analysis. However, the plot of unreacted RNA against the ratio of PyrR added to total RNA predicts a 2 to 1 ratio of PyrR to RNA (see Figure 5, main text).

## **References**

1. Chander P, Halbig, KM, Miller JK, Fields CJ, Bonner HKS, Grabner GK, Switzer RL & Smith, JL (2005) Structure of the nucleotide complex of PyrR, the *pyr* attenuation protein from *Bacillus caldolyticus*, suggests dual regulation by pyrimidine and purine nucleotides. *J Bacteriol* **187**, 1773-1782.

2. Bonner ER, D'Elia JN, Billips BK & Switzer RL (2001) Molecular recognition of *pyr*  mRNA by the *Bacillus subtilis* attenuation regulatory protein PyrR, *Nucleic Acids Res*  **29**, 4851-4865.

3. Tomchick DR, Turner RJ, Switzer RL & Smith JL (1998) Adaptation of an enzyme to regulatory function: structure of *Bacillus subtilis* PyrR, a bifunctional *pyr* RNA-binding attenuation protein and uracil phosphoribosyltransferase. *Structure* **6**, 337-350.

4. Savacool HK & Switzer RL (2002) Characterization of the interaction of *Bacillus subtilis* PyrR with *pyr* mRNA by site-directed mutagenesis of the protein. *J Bacteriol*  **184**, 2521-2528.

5. Breeuwer P, Drocourt J-L, Rombouts FM & Abee T (1996) A novel method for continuous determination of the intracellular pH in bacteria with the internally conjugated fluoresecent probe 5 (and 6-)-carboxyfluorescein succinimidyl ester. *Appl Environ Microbiol* **62**, 178-183.

6. Schuck P (2000) Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys J* **78**, 1606-1619.

7. Garcia de la Torre J, Navarro S, Lopez Martinez MC, Diaz FG & Lopez Cascales JJ

(1994) HYDRO: a computer software for the prediction of hydrodynamic properties of macromolecules. *Biophys J* **67**, 530-531.

8. Laue TM, Shah BD, Ridgeway TM & Pelletier SL (1992) Computer-aided interpretation of analytical sedimentation data for proteins. in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding SE, Rowe AJ & Horton JC, eds) pp. 90-125, The Royal Society of Chemistry, Cambridge, UK. 9. Schuck P & Demeler B (1999) Direct sedimentation analysis of interference-optical

data in analytical ultracentrifugation. *Biophys J* **76**, 2288-2296.

**Table S1. Structural variants of purine and pyrimidine nucleotides that affect binding of PyrR to BcBL2 as determined by filter binding**





The effectors were present at 0.5 mM. The data are averages of at least three independent determinations and include standard deviations of the mean value.





NE, no effector; n.b., no detectable binding. UMP and GMP were present at 0.5 mM.

The data are averages of three independent determinations and include standard deviations of the mean value.

Sample Analyzed	0.1 <sub>M</sub> NaCl <sup>a</sup>	Initial Concentration	$S_{20,w}$	$f/f_0^b$	Calculated Mass <sup>c</sup>
		μM	S		kDa
Native PyrR	$+$	50	4.83	1.27	83.4
Native PyrR	$+$	5	4.78	1.31	101
Native PyrR		33	5.09	1.22	83
Native PyrR		$\overline{2}$	5.04	1.24	84
His-tagged PyrR	$\ddot{}$	28	5.28	1.28	99
His-tagged PyrR	$+$	$\overline{2}$	5.05	1.37	94
His-tagged PyrR		27	5.4	1.27	95
His-tagged PyrR		5.6	5.2	1.33	98

**Table S3**. **Sedimentation velocity analysis of native and His-tagged PyrR**

 $^a$  All samples included 10 mM Tris-Cl, pH 7.5 and 10 mM MgCl<sub>2</sub>.

 $b$  f/f<sub>0</sub> calculated using the vbar method in Sednterp with 0.4 g water per g of protein.

 $\degree$  Sequence masses for the tetramer are 79,754 Da and 89,152 Da for native and Histagged PyrR, respectively.



# **Table S4. Sedimentation velocity constants from the titration of BcBL2 with PyrR**

<sup>a</sup>Ratio = [PyrR]/[BcBL2] where [PyrR] is the subunit concentration and BcBL2 is the 36nt synthetic RNA corresponding to nt nos.702 through 737 of BcBL2 shown in Fig. 1.  $^{\text{b}}$ Concentration, A<sub>260</sub>, is the integrated area under the peak for the given s value in A<sub>260</sub> units (loading concentration).

<sup>c</sup>The loading concentration of BcBL2 was 0.3  $\mu$ M and of PyrR was 1.2  $\mu$ M subunit.

<sup>d</sup>The values in parentheses are for a non-interacting 2 species model.

### **Supplementary Material: Figure Legends**

**Figure S1.** Effects of the pH of the binding buffer on reactions incubated at 0°C (**A**) and of temperature on reactions at pH 7.5 (**B**) on binding of BcBL2 to PyrR as determined by filter binding. Open circles: no effector; closed circles: 500 µM UMP; closed triangles: 500 µM GMP.

**Figure S2.** Sedimentation velocity studies on native and His-tagged PyrR at moderate ionic strength. Panels A, B, and C show the results of a sedimentation velocity run with native PyrR in 0.1 M NaCl at 50,000 rpm and 20°C. Panel A shows changes in absorbance as a function of radial position (dots) at six min intervals and the results (solid lines) from a c(s) distribution analysis using the program SEDPHAT 4.1. Panel B shows the calculated residuals against radial position for each scan in Panel A, and Panel C shows the  $c(s)$  distribution against s corrected to  $s_{20,w}$ . The weight average sedimentation coefficient from the integration of the principal peak is 4.83 S. Panels D, E, and F show the results from an identical sedimentation velocity experiment and analysis using His-tagged PyrR. The weight average  $s_{20,w}$  calculated from these results is 5.34 S. Additional details for both experiments are found in the Data Supplement text and Table S3.

**Figure S3**. An approach to equilibrium study with the native PyrR at low ionic strength. The initial concentrations of PyrR studied were: Panels A and B, 0.25 µM; Panels C and D, 1.25  $\mu$ M; Panels E and F, 25  $\mu$ M. The data were collected at 8000 rpm. Rayleigh interference optics were used for Panels A, C, and E; absorbance optics at 232 nm

were used for Panels B and D; and absorbance at 278 nm was used for Panel F. The calculated values for  $s_{20,w}$  and observed mass based on a global single species fit to both data sets with the program SEDPHAT 4.1 are 5.7 S and 61 kDa, respectively, for a native PyrR concentration of 0.25  $\mu$ M. The calculated values for  $s_{20,w}$  and mass for PyrR at an initial concentration of 1.25  $\mu$ M are 3.7 S and 71 kDa, respectively, and a loading concentration of 25  $\mu$ M PyrR, the calculated values for  $s_{20,w}$  and mass are 5.1 S and 82 kDa. The solid lines in each panel are the fitted results from the single species model (see Data Supplement text); the calculated residuals against radial position for each scan are shown in the boxes below each panel.

**Figure S4.** Equilibrium sedimentation studies of native and His-tagged PyrR at low ionic strength and concentrations shown. The absorbance data in the upper panel (native PyrR) and the lower panel (His-tagged PyrR) were obtained from a continuation of the experiments shown in Figure S2 in that the samples were allowed to come to equilibrium at 8 k, 12 k, 1k 17 k, and 20 k rpm and absorbance scans across the cell were collected at  $A_{232}$ ,  $A_{232}$ , and  $A_{278}$ , respectively, from left to right. The solid lines were calculated based on a single discrete species model, with mass conservation, using SEDPHAT 4.1 and a mass of 78.3 kDa. The data in the lower panel were obtained using the same procedures as in the upper panel, except that the samples were allowed to come to equilibrium at 12 k, 15 k, 17 k, and 20 k rpm. Absorbance scans were collected at the same wavelengths indicated in the upper panel. The solid lines were calculated based on a single discrete species model, with mass

conservation, using SEDPHAT4.1 and a mass of 91.6 kDa. Calculated residuals against radial position for each scan are shown in the boxes below each panel.

**Figure S5.** Sedimentation velocity studies of the interaction between RNA and Histagged PyrR. The concentrations of PyrR and RNA in each sample are shown at the left of each row of panels; additional details appear in the legend of Figure 4. The three sets of scans shown in each row of panels were collected using Rayleigh interference,  $A_{260}$  and  $A_{278}$ , from left to right, respectively. The data were fitted using the  $c(s)$ distribution algorithm of SEDPHAT as described in the Methods section; c(s) distributions were calculated form interference data (red lines), absorbance at 260 nm (green lines), absorbance at 278 nm (blue lines).

**Figure S1**



**Figure S2**











**Figure S4**



