Articles

Structure-Based Redesign of Corepressor Specificity of the *Escherichia coli* Purine Repressor by Substitution of Residue 190[†,‡]

Fu Lu,[§,||] Maria A. Schumacher,[||,⊥] Dennis N. Arvidson,[||,⊥] Andreas Haldimann,[#] Barry L. Wanner,[#] Howard Zalkin,[*,§] and Richard G. Brennan[*,⊥]

Departments of Biochemistry and Biological Sciences, Purdue University, West Lafayette, Indiana 47907-1153, and Department of Biochemistry and Molecular Biology, Oregon Health Sciences University, Portland, Oregon 97201-3098

Received August 6, 1997; Revised Manuscript Received October 8, 1997

ABSTRACT: Guanine or hypoxanthine, physiological corepressors of the *Escherichia coli* purine repressor (PurR), promote formation of the ternary PurR−corepressor−operator DNA complex that functions to repress *pur* operon gene expression. Structure-based predictions on the importance of Arg190 in determining 6−oxopurine specificity and corepressor binding affinity were tested by mutagenesis, analysis of in vivo function, and in vitro corepressor binding measurements. Replacements of Arg190 with Ala or Gln resulted in functional repressors in which binding of guanine and hypoxanthine was retained but specificity was relaxed to permit binding of adenine. X-ray structures were determined for ternary complexes of mutant repressors with purines (adenine, guanine, hypoxanthine, and 6-methylpurine) and operator DNA. These structures indicate that R190A binds guanine, hypoxanthine, and adenine with nearly equal, albeit reduced, affinity in large part because of a newly made compensatory hydrogen bond between the rotated hydroxyl side chain of Ser124 and the exocyclic 6 positions of the purines. Through direct and water-mediated contacts, the R190Q protein binds adenine with a nearly 75-fold higher affinity than the wild type repressor while maintaining wild type affinity for guanine and hypoxanthine. The results establish at the atomic level the basis for the critical role of Arg190 in the recognition of the exocyclic 6 position of its purine corepressors and the successful redesign of corepressor specificity.

The *Escherichia coli* purine repressor (PurR) binds to a conserved 16 bp operator sequence and regulates transcription of multiple *pur* regulon genes, including nine operons for de novo purine nucleotide biosynthesis, the *purR* gene, and other genes with functions related to nucleotide metabolism (*1*). PurR is a 341-amino acid residue polypeptide that functions as a dimer. Each subunit is composed of an NH_2-terminal DNA binding domain (residues 1−60) and a larger COOH-terminal corepressor binding and dimerization domain (residues 61−341), termed the CBD. Binding of a corepressor, hypoxanthine or guanine, increases the operator DNA binding affinity. Consequently, PurR is the primary regulatory protein that coordinates expression of *pur* regulon genes to the availability of purines in *E. coli*. In addition, *purF*-encoded glutamine PRPP amidotransferase, the first enzyme of the pathway for de novo purine nucleotide synthesis, is subject to feedback regulation by purine nucleotides.

The X-ray structures have been determined for a PurR−hypoxanthine−*purF* operator complex (*2*) as well as the ligand-free PurR CBD dimer (*3*). The CBD consists of NH_2 and COOH subdomains which have very similar α/β topologies. Residues at the interface of the two subdomains and from two of the three crossover strands that connect the subdomains form the corepressor binding pocket. The structures of the PurR−hypoxanthine−*purF* operator complex and the corepressor-free CBD also identified key elements of the protein involved in DNA binding and have suggested the allosteric mechanism by which corepressor binding mediates specific binding to DNA (*3*). Upon binding corepressor, structural changes that include hinge-bending rotations of at least 17° between the NH_2 and COOH subdomains of each monomer result in the "closed" CBD conformation. These rotations were deduced to position the DNA-binding domains of each PurR monomer such that in the presence of operator DNA, the hinge region, residues 48−56, can undergo a coil to helix transition. Once formed, the hinge helix interacts specifically with the minor groove and causes the DNA to be kinked by greater than 50°. The 2-fold related helix−turn−helix motifs of PurR are then able to make a series of base-specific contacts in the DNA major

[†] Supported by U.S. Public Health Service Grants GM24658 (H.Z.) and GM49244 (R.G.B.) and a National Defense Science and Engineering Fellowship in Biosciences (M.A.S.). This is journal paper 15 568 from the Purdue University Agricultural Research Station.

[‡] Atomic models for all structures have been deposited with the Brookhaven Protein Data Bank. Their ID codes are 2PUA, 2PUB, 2PUC, 2PUD, 2PUE, 2PUF and 2PUG.

* Authors to whom correspondence should be sent. H.Z.: telephone (765)-494-1618; fax, (765)-494-7897; e-mail, zalkin@biochem.purdue.edu. R.G.B.: telephone, (503)-494-4427; fax: (503)-494-8393; e-mail, brennanr@ohsu.edu.

[§] Department of Biochemistry, Purdue University.

[||] F.L., M.A.S., and D.N.A. contributed equally to this work.

[⊥] Oregon Health Sciences University.

[#] Department of Biological Sciences, Purdue University.

S0006-2960(97)01942-9 CCC: $15.00
Published on Web 01/06/1998

Table 1: Strains and Plasmids

bacteria	description	construction, reference, or source
BL21(DE3)/pLysS	host for overexpression, Cm^R	Novagen
BW23473	Δ*lac-169 robA1 cre C510 hsdR514 endA recA1* Δ*uidA::pir*$^+$	Haldimann et al., 1996
BW23474	Δ*lac-169 robA1 cre C510 hsdR514 endA recA1* Δ*uidA::pir-116*	Haldimann et al., 1996
FL100	Δ*lac-169 purR220 purF-lacZ(kan)*$_{TX773}$	R220 with P1 on TX773
R220	Δ*lac-169 purR220*	Rolfes and Zalkin, 1988
TX883	Δ*lac-169 purR::Tn10 purF-lacZ(kan)*$_{TX773}$	He et al., 1990

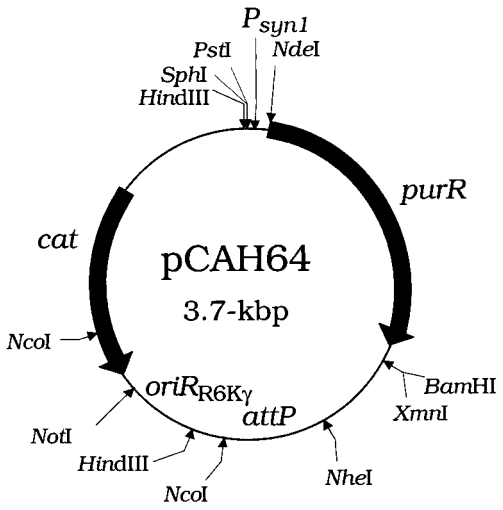
plasmid	description	construction, reference, or source
pCAH64	CmR, P_{syn1}-*purR*, *ori*$R_{R6K\gamma}$, *attP*	see text
pDNA100	1.1 kb *Nde*I/*Hind*III *pur*R+ in pET24(a)	D. N. Arvidson and R. G. Brennan
pET24(a)	T7 overexpression plasmid, Kan^R	Novagen
pHH7013	Amp^R, P_{syn1}	J. Hu
pINT-ts	Amp^R, *rep*101^{ts}, *cI*857, P_R-*int*	M. Koob
pPR1005−2	*purR*$^+$, *purO*$_2$, *Nde*I at Met^1, pUC119	Rolfes and Zalkin, 1990
pSK49Δ*uidA2*	Kan^R, *lacI*^{q}, P_{tac}-*uidA2*, *ori*$R_{R6K\gamma}$, *attP*	Haldimann et al., 1996

groove (*2*). Thus, PurR function requires a minimum of five steps: (i) corepressor binding, and (ii) a CBD structural transition that is (iii) transduced to the DNA binding domain, resulting in (iv) DNA binding and (v) inhibition of transcription.

Analysis of the corepressor binding pocket of the PurR implicated residue Arg190 as a key specificity determinant of PurR for 6-oxopurines (*2*). Specifically, the guanidinium group of this residue was found to donate two hydrogen bonds to the two lone pair acceptors of the O6 atom of hypoxanthine. Such hydrogen bond complementarity should preclude PurR binding to any purine that has hydrogen bond donor groups at this position, e.g., adenine. In addition to its proposed role in corepressor discrimination, the Arg190 guanidinium group is involved in a water-mediated hydrogen bond to the N1. Although not a corepressor-discriminating interaction, this water-mediated hydrogen bond likely contributes to the overall binding affinity.

In this report, we describe the structure-based design and analysis of two mutant PurR proteins in which residue Arg190 is substituted by alanine or glutamine. To study the function of mutations in vivo, we used conditionally replicative, integrative plasmids to incorporate *purR* into the chromosome in single copy. Next, we purified the mutant proteins to assess their corepressor and DNA binding activities. Finally, we describe seven crystal structures of mutant PurR−corepressor−*purF* operator complexes in which Arg190 has been replaced by either alanine (R190A) or glutamine (R190Q). Replacement of Arg190 by alanine or by glutamine expanded the corepressor specificity of PurR to include adenine as well as guanine and hypoxanthine and, in the case of the R190A mutant, 6-methylpurine. The structures of R190A bound to hypoxanthine, guanine, adenine, or 6-methylpurine reveal the structural and chemical features that allow these purines, which differ at their exocyclic 6 position, to bind the repressor with nearly equal, albeit lowered, affinity. The structures of R190Q bound to hypoxanthine, guanine, or adenine clarify the stereochemical details of the direct and water-mediated interactions which this protein utilizes to bind the natural corepressors with undiminished affinity and adenine with only a 6−16-fold lower affinity than hypoxanthine and guanine, respectively. Combined, the in vivo, in vitro, and structural data demonstrate that Arg190 is the key recognition determinant of PurR for 6-oxopurine corepressors.

A. Structure of pCAH64



B. Sequence of P_{syn1} region

```
PstI      BspEI SpeI          EcoRI      -35
CTGCAGCTGTCCGGAACTAGTGTCTTCAAGAATTCTAGGCTTTATACTTTAT
          -10          KpnI XbaI  rbs  NdeI
GCTTCCGGCTCGTACATTGTGTGGAGGTACCTCTAGACAGGAGACCATATG
```

FIGURE 1: Map of conditionally replicative integration vector. (**A**) The vector with *purR* cloned at *Nde*I and *Bam*HI sites is described in Materials and Methods. (B) Nucleotide sequence of the promoter region (overlined) and the ribosome binding site and translation start site (underlined).

MATERIALS AND METHODS

Strains and Plasmids. *E. coli* strains and plasmids are listed in Table 1. Plasmid pCAH64 (Figure 1A) is a choramphenicol-resistant, *ori*$R_{R6K\gamma}$, *attP* plasmid that was derived from pSK49Δu*idA2* (*4*). These plasmids are conditionally replicative because they have the γ replication origin of R6K (denoted *ori*$R_{R6K\gamma}$). Hence, they require the trans-acting Π protein, the *pir* gene product, to replicate as a plasmid. An *ori*$R_{R6K\gamma}$ plasmid replicates at an intermediate or high copy number, ca. 15 or 250 copies per cell, in a strain with *pir*$^+$ or *pir*-116, respectively, on the chromosome (*5*). All *pir*-dependent plasmids were propagated in the *pir*$^+$ host BW23473 or the *pir*-116 host BW23474 (*4*). These plasmids are integrative because they carry *attP* from phage λ. In a standard non-*pir* strain, an *ori*$R_{R6K\gamma}$, *attP* plasmid cannot replicate. Rather, it integrates into the chromosome

in single copy by site-specific recombination when λ integrase (int) is supplied in trans (*4*, *6*). In pCAH64, *purR* is transcribed from a weak synthetic promoter denoted P_{syn1} that originated from pHH7013 (from J. Hu). The promoter region, ribosome binding site, and translation start site are shown in Figure 1B. The *purR* wild type and mutant genes were subcloned into a precursor plasmid pCAH63 as *Nde*I/*Bam*HI fragments, thus yielding pCAH64. Similar plasmid constructs in which *purR* was expressed from P_{tac} instead of P_{syn1} were toxic, presumably due to high expression of *purR* (*7*).

For construction of strain FL100, the *purF-lacZ(kan)*$_{TX773}$ transcriptional fusion (from TX773) (*8*) was crossed into *purR* strain R220 by selecting Kan^R transductants (*9*). The resulting strain FL100 was transformed with pINT-ts (*6*) to provide integration function. pINT-ts is an ampicillin-resistant, low-copy plasmid with a temperature-sensitive replication origin; pINT-ts expresses *int* under control of the λcI-repressible P_r promoter and encodes the temperature-sensitive cI857 repressor.

Plasmid Integration. Plasmids were integrated into the chromosomal *attB* site in strain FL100 using pINT-ts (*4*). The *int* gene in pINT-ts is expressed at temperatures above 37 °C; the plasmid is also temperature-sensitive for replication (*5*). Thus, incubation of pINT-ts transformants at 43 °C simultaneously induces int synthesis and blocks plasmid replication. Accordingly, an ampicillin-resistant transformant of FL100 carrying pINT-ts was grown in Luria-Bertani (LB) broth with ampicillin (100 μg/mL) at 30 °C and transformed with a pCAH64*purR* plasmid. The mixture was incubated at 43 °C for 30 min and at 37 °C for 1 h, then spread onto LB agar containing chloramphenicol (6 μg/mL), and incubated at 37 °C. Several Cm^R colonies were purified on LB agar (without chloramphenicol), screened for ampicillin-sensitive ones, to verify loss of pINT-ts, and then tested by PCR for plasmid copy number. Once constructed, integrants were routinely grown in the absence of antibiotic selection. A pCAH64 plasmid integrant lacking *purR* was used as a negative control.

PCR Testing for Copy Number. Two primers specific for the vector *attP* (P2 and P3) and two specific for the chromosomal *attB* region (P1 and P4) were used to test integrants for those with a single copy of the integrated plasmid (Figure 2), as described (*4*). Primers P1 (GAAGGACGTTGATCGGGCGGGG), P2 (CAGTGACACAGGAACACTTAACG), P3 (CACGATAATATCCGGGTAGG), and P4 (GGCGCAATGCCATCTGGTATCAC) were obtained from IDT, Inc. (Coralville, IA). These primers were designed so that a non-integrant yielded a 231 bp PCR fragment with P1 and P4 (*attB* pair) and a single-copy integrant gave a 327 bp PCR fragment with P1 and P2 (*attL* pair) and a 592 bp PCR fragment with P3 and P4 (*attR* pair). Double or higher multiple integrants gave the 327 and 592 bp PCR fragments and an additional 698 bp PCR fragment with P2 and P3 (*attP* pair). These fragment sizes were verified when non-integrant, single-integrant, and double or multiple-integrant control strains were tested.

PCR was performed directly using bacterial colonies as described (*10*). Cells were picked with a glass capillary into 20 μL of lysis solution [1× PCR buffer (Promega, Madison, WI) containing 2.5 mM $MgCl_2$, 0.1 mg/mL nuclease-free BSA, 0.05 mg/mL Proteinase K, 20 mM dithiothreitol, and

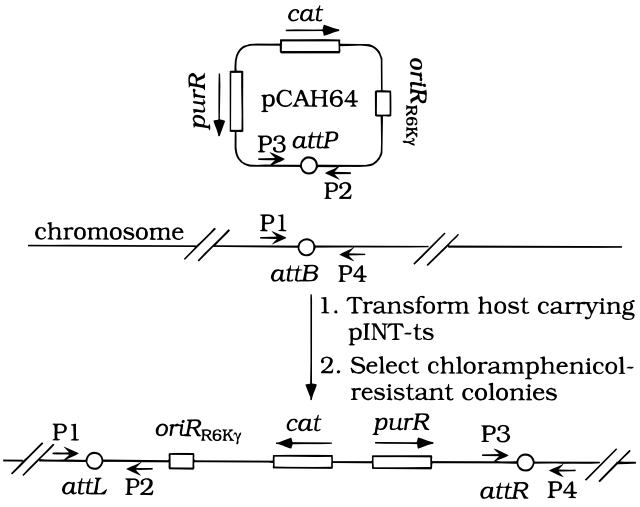


FIGURE 2: Scheme for integration of pCAH64 in *attB*. PCR primers P1−P4 are used to determine single-copy integration as described in the text.

1.8 μM sodium dodecyl sulfate]. The mixtures were incubated at 37 °C for 20 min, after which they were heated to 90 °C for 5 min. Five microliters of lysed cells, 50 pmol of each primer (P1−P4) and 2 units of Taq DNA polymerase (Promega) were combined with 1× PCR buffer, 2.5 mM $MgCl_2$, and each deoxynucleoside triphosphate at 200 μM in a final volume of 100 μL. PCR was carried out for 25 cycles (denaturing for 1 min at 94 °C, annealing for 1 min at 63 °C, and extending for 1 min at 72 °C).

Assay for Repression by purR. FL100 *purR* integrants were cultured overnight at 37 °C in minimal medium (*11*) containing 15 μg/mL chloramphenicol. Then, 100 μL of the fresh overnight culture was used to inoculate 5 mL of the same medium in a 1.5 cm × 12 cm tube. The cells were grown with shaking at 37 °C until reaching mid-log phase. The β-galactosidase activity was assayed in permeabilized cells according to the method of Miller (*12*).

Site-Directed Mutagenesis. Site-directed mutagenesis by the procedure of Kunkel et al. (*13*) was carried out as described previously (*14*). Single-stranded *purR* template DNA was prepared from plasmid pPR1005-2 (*11*).

Overexpression and Purification of PurR. Wild type and mutant *purR* genes in pPR1005-2 were transferred via flanking *Nde*I and *Hin*dIII sites to pET24(a) for overexpression. In this vector, *purR* is transcribed from a phage T7 promoter under the control of a *lac* operator. The pET24-(a) vector in strain BL21(DE3)pLysS (D. N. Arvidson and R. G. Brennan, unpublished) thus replaced the previous method for *purR* overexpression where pT7-7 and phage CE6 were employed to provide T7 RNA polymerase (*11*). These manipulations were necessary to negate purR toxicity from a multicopy plasmid. Plasmids were transformed into *E. coli* BL21(DE3)pLysS for IPTG-induced overexpression (*15*). Overexpressed protein was purified by two chromatographic steps, DEAE-Sepharose and heparin−agarose as described (*16*). For purified PurR, the protein concentration was determined spectrophotometrically at 280 nm using an A_{280} (1%) of 12.9 (*11*).

DNA Binding Site. PurR binds to a 16 bp palindromic site in 19 *E. coli* operons (*1*). Two binding sites were used in this study: the 16 bp *purF* operator and a 16 bp perfect palindromic sequence (*1*) designated *pur*. These sites which

have been used previously (*2*, *14*) were incorporated into oligonucleotides of different lengths. Flanking sequences in oligonucleotides of 24−30 bp have not been observed to influence PurR binding affinity.

Corepressor Binding. Equilibrium dialysis was used to determine binding of purine corepressors when the K_d for guanine and hypoxanthine was <20 μM and the K_d for adenine was <35 μM. The binding measurements were carried out as described previously (*16*). Binding data were fit to the Hill equation (*17*) using Ultrafit software (Biosoft, Cambridge, U.K.) (eq 1). *Y* is the fraction of protein bound;

$$Y = [L]^n/K_d^{\,n} + [L]^n \qquad (1)$$

K_d, the dissociation constant; *n*, the Hill coefficient; and L, free corepressor. Fits were made allowing both K_d and *n* to vary.

In some cases, K_d values were too high to be determined by equilibrium dialysis and gel retardation was used to estimate binding of corepressors to PurR (*11*). In this method, corepressor binding to PurR is coupled with binding of the holorepressor to operator DNA. The experiments were carried out under conditions where PurR binding to DNA is dependent upon corepressor. The method relies on the observation that the affinity of the PurR−corepressor complex for DNA is more than 200-fold greater than the affinity of PurR for corepressor (*11*, *14*, *16*). This relationship assures that once corepressor binds to PurR the complex will bind to DNA. A 30-mer DNA probe with the *pur* operator was labeled with ^{32}P as described (*14*). Binding was performed in corepressor-dependent buffer II [100 mM HEPES-KOH (pH 7.5), 250 mM potassium glutamate, 150 mM NaCl, 10 mM magnesium acetate, 1 mM EDTA, and 2% dimethyl sulfoxide]. The 20 μL incubation mixture contained 10 fmol of DNA probe, 15 nM PurR, and varied concentrations of corepressor and buffer II. The concentration of PurR refers to the 38 kDa subunit. The protein−DNA complex and free DNA were separated by electrophoresis, and the gel was dried and counted for radioactivity using a Packard Instant Imager. From these data, the fractional saturation of PurR was estimated at varied concentrations of corepressor. In previous experiments, similar K_d values were determined for binding of hypoxanthine and guanine to PurR by equilibrium dialysis (*16*) and estimated by gel retardation methods (*11*). In this work, a control experiment confirmed that the two methods gave similar K_d values for the PurR−hypoxanthine interaction (data not shown).

Finally, the binding constants for binding of 6-methylpurine to wild type PurR and the R190Q mutant were determined using a fluorescence anisotropy assay (*18*, *19*). In this assay, a 24 bp oligonucleotide containing the *pur* operator site was annealed to its complementary strand that is fluoresceinated on the 5′-end. The anisotropy measurements were carried out at room temperature in 1 mL of binding buffer II, which also contained 1 μg of poly(dI·dC), 50 nM PurR or R190Q, and 2 nM labeled duplex. The purine was titrated into this mixture. The anisotropy measurements were taken with a Panvera Beacon Fluorescence Polarization instrument. The data were plotted as anisotropy versus purine concentration, and the K_d was calculated using the equation (adapted from ref *19*):

$$A = \frac{\dfrac{A_{max}[P]_T}{K_c^{\,2} + {}^1/_2K_c[C] + [C]^2}[C]^2}{\dfrac{[P]_T}{K_c^{\,2} + {}^1/_2K_c[C] + [C]^2}[C]^2 + K_d} + A_o \qquad (2)$$

Using the experimental data for *A* (the measured anisotropy), $[P]_T$ (the total protein concentration), [C] (the corepressor or purine concentration), and K_d (the protein−DNA equilibrium dissociation binding constant), we fit the equation to determine K_c (the protein−purine equilibrium dissociation binding constant), A_{max} (the maximum anisotropy value at saturation), and A_o (the initial anisotropy value).

Measurement of ApoPurR−DNA Binding. The specific binding of wild type and mutant apoPurR to operator DNA was determined by gel retardation using corepressor-independent buffer I [10 mM HEPES-KOH (pH 8.0), 100 mM KCl, and 1 mM EDTA] (*11*). In this buffer, PurR assumes the conformation required for high-affinity binding to operator DNA without a requirement for corepressor. This assay was used to evaluate the DNA binding function of corepressor binding mutants.

Crystallization, Data Collection, and Structure Determination. Crystals of the mutant PurR−purine−*pur* operator complexes were grown by the hanging drop−vapor diffusion method, which was used previously to crystallize the PurR−hypoxanthine−*purF* operator complex (*20*). Briefly, protein at a concentration of 0.25−0.50 mM was saturated with corepressor and mixed 1:1 (v:v) with a 0.50 mM solution of the 16 bp *purF* operator or a *pur* operator site (*18*), both of which have an additional 5′-nucleoside overhang on both oligonucleotide strands, the identity of which is irrelevant to crystallization. The protein−corepressor−DNA solution was mixed 1:1 with the reservoir solution, which is 25% polyethylene glycol (PEG) 4000, 0.40 M ammonium sulfate, 50 mM cobalt hexammine chloride, and 0.10 M ammonium phosphate (pH 7.5). Crystals grow initially as thin two-dimensional plates and within a period of 2 weeks to several months melt and are replaced by large rhombohedral crystals (typically 0.6 mm × 0.2 mm × 0.3 mm). The mutant PurR−purine−*pur* operator crystals are isomorphous to the PurR−hypoxathine−*purF* operator crystals (Table 2). X-ray intensity data were collected at room temperature with an Area Detector Systems Corp. (ADCS) multiwire area detector (*21*) using a Rigaku RU200-H rotating anode operating at 40 kV and 150 mA. Selected crystallographic data for each mutant−purine−DNA complex are listed in Table 2.

Structure Determination and Refinement. All seven structures were solved by difference Fourier techniques using the wild type PurR−hypoxanthine−*purF* operator complex, minus the corepressor, the arginine side chain of residue 190 and solvent, as the starting model for each (*2*). There is one PurR monomer, purine molecule, and *pur* or *purF* operator half-site in the asymmetric unit. Refinement of each mutant complex structure began with several cycles of rigid body refinement and was followed by *XYZ* (positional) and *B* factor (thermal parameter) refinement as implemented in TNT (*22*). The R190Q−guanine−*purF* operator structure underwent only *XYZ* refinement due to the limited resolution of the data. Purines and water molecules were built into the structures in the final stages of refinement using either

Table 2: Summary of Selected Crystallographic Data

	R190−Hyp	R190A−Gua	R190A−Ade	R190A−6MP	R190Q−Hyp	R190Q−Gua	R190Q−Ade
Data Collection Statistics							
PDB code	2PUD	2PUC	2PUB	2PUA	2PUG	2PUF	2PUE
cell constants	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$
	$a = 175.93$ Å	$a = 175.93$ Å	$a = 175.81$ Å	$a = 175.65$ Å	$a = 175.98$ Å	$a = 175.92$ Å	$a = 175.81$ Å
	$b = 95.02$ Å	$b = 95.06$ Å	$b = 94.87$ Å	$b = 94.68$ Å	$b = 95.19$ Å	$b = 95.05$ Å	$b = 95.17$ Å
	$c = 81.78$ Å	$c = 81.53$ Å	$c = 81.75$ Å	$c = 81.83$ Å	$c = 81.22$ Å	$c = 81.59$ Å	$c = 81.31$ Å
unique reflections	20 826	20 405	18 609	14 911	18 672	13 317	17 561
total reflections	100 630	68 989	77 255	51 188	68 392	49 911	39 957
completeness	99	97	99	99	98	96	92
R_{symm} (%)[a]	6.8	6.0	5.3	9.0	7.6	9.8	6.8
$I/\sigma(I)$	6.6	7.6	7.2	5.7	8.5	3.6	6.5
Refinement Statistics							
resolution (Å)	2.6	2.6	2.7	2.9	2.7	3.0	2.7
R_{Factor} (%)[b]	18.0	17.3	17.1	16.6	17.1	22.2	17.5
number of atoms	3001	3002	3001	3005	3003	3006	3005
number of solvent molecules	61	55	60	60	64	35	54
Stereochemistry							
Rmsds[c]							
bond lengths (Å)	0.017	0.016	0.016	0.012	0.011	0.009	0.013
bond angles (deg)	2.49	2.45	1.67	1.76	1.06	0.085	1.19
Ramachandran Plot Analysis[d]							
region							
most favored (no./%)	254/85.5	260/87.5	253/85.2	262/87.9	261/87.9	251/84.5	249/83.8
additional allowed	39/13.1	33/11.0	39/13.1	33/11.1	32/10.8	42/14.1	44/14.8
generously allowed	2/0.7	2/0.7	3/1.0	1/0.3	2/0.7	3/1.0	2/0.7
disallowed	2/0.7	2/0.7	2/0.7	2/0.7	2/0.7	1/0.3	2/0.7

[a] $R_{symm} = \sum|I_o - \langle I\rangle|/I_o$, where I_o is the observed intensity and $\langle I\rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections. [b] $R_{Factor} = \sum||F_{obs}| - |F_{calc}||/|F_{obs}|$. [c] Rmsds are the root mean square deviations of the bond lengths and bond angles from their respective ideal values as implemented in TNT (*22*). [d] The Ramachandran analysis was generated with PROCHECK (*25*).

FRODO (*23*) or O (*24*). To ensure the unbiased placement of the purines and any side chains, which underwent significant movements, a series of $F_o - F_c$ omit maps were calculated for each structure. In these omit map calculations, the purine, attendant water molecules, and any rebuilt corepressor binding pocket residues were removed. The model was then subjected to 20−30 cycles of positional and thermal refinement before inspection of the omit density map. Refinement statistics and selected stereochemical parameters (*25*) are listed in Table 2.

RESULTS

In Vivo *Function of PurR Mutants.* On the basis of the critical interaction of Arg190 with the purine corepressor exocyclic 6 position, seen in the X-ray structure of holoPurR bound to operator DNA (*2*), we have mutagenized this residue to alanine or glutamine. Mutant genes were constructed and integrated into the phage λ attachment site of the *E. coli* chromosome in single copy in order to evaluate repressor function (Figure 2). The *purR* gene, in this construction (Figure 1A), is controlled by a weak synthetic promoter (Figure 1B) chosen to give expression comparable to that of the wild type gene at its normal site. Western blots of wild type cells and the integrants verified that all strains contained similar PurR levels (not shown). Data in Table 3 show repression of a *purF-lacZ* reporter by $PurR^+$ and mutant repressors. Repression of *purF* by $PurR^+$ and R190Q was 18- and 21-fold, respectively, values that are similar to the 17-fold repression by the wild type gene at its normal chromosomal position (*8, 26*). R190A retained approximately 60% of the function of $PurR^+$. It is necessary to emphasize that the repression assay depends upon corepressor-dependent binding of PurR to the *purF* operator but does not address which purines function as corepressor.

Table 3: In Vivo Function of PurR Mutants

PurR[a]	β-galactosidase[b] (Miller units)	fold repression[c]
$PurR^-$ (vector control)	37.4 ± 3.5	1
$PurR^+$	2.1 ± 0.3	17.8 ± 0.9
R190A	3.6 ± 0.6	10.4 ± 0.7
R190Q	1.8 ± 0.3	20.8 ± 3.1
S124A	2.4 ± 0.4	15.6 ± 2.7
S124A/R190A	27.7 ± 1.7	1.4 ± 0
S124D	23.2 ± 2.4	1.6 ± 0.3
S124D/R190A	24.2 ± 2.6	1.5 ± 0.1
S124N/R190A	34.0 ± 4.3	1.1 ± 0.1

[a] $PurR^+$, mutant genes, or the vector control was integrated into the *λattP* site of strain FL100. [b] Values are the average of three independent determinations ± the standard deviation. [c] Repression is calculated as the ratio of β-galactosidase from $PurR^-$/PurR (wild type or mutant).

PurR Purification. The wild type and R190A and R190Q proteins were overexpressed as described in Materials and Methods, and the repressor protein was purified in order to assess corepressor binding directly. For the wild type and the two mutants, PurR accounted for approximately 40% of the soluble protein on the basis of sodium dodecyl sulfate−polyacrylamide gel electrophoresis (not shown). The elution profiles of PurR from DEAE-Sepharose and heparin−agarose columns were identical for the wild type and mutants. In addition, binding constants similar to that of operator DNA were obtained for the wild type and mutant repressors when determined in corepressor-independent binding buffer I (*11*). The apparent K_d for wild type PurR was 11 nM, and the values for the mutants were between 7 and 16 nM (data not shown).

Binding of Purines to PurR. Three methods were used to determine binding affinities of purines to wild type and

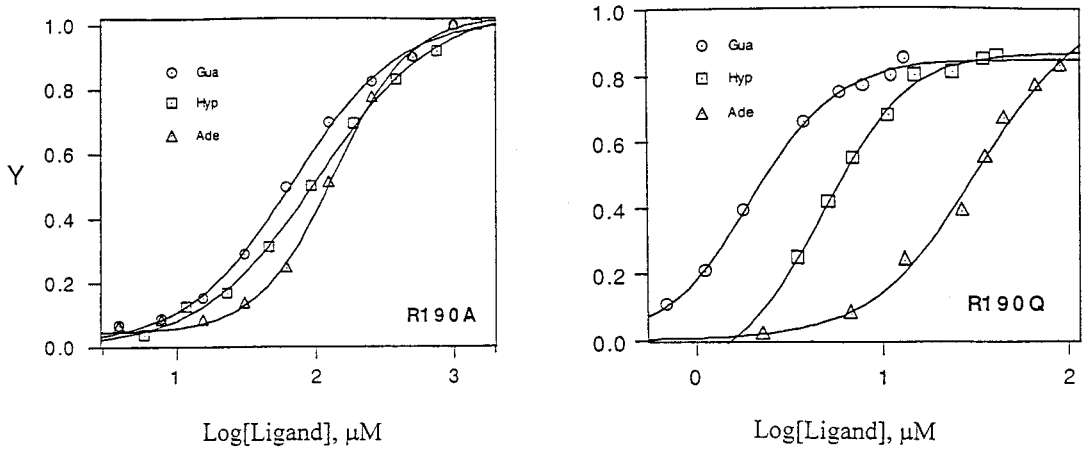


FIGURE 3: Binding of corepressors to R190 mutant proteins. (Left) R190A. Saturation curves were determined by gel retardation experiments. The fraction of bound DNA (*Y*) is plotted against the total ligand concentration. (Right) R190Q. Binding was determined by equilibrium dialysis. Repressor concentrations of 1, 10, and 25 μM were used to determine binding of Gua, Hyp, and Ade, respectively. The fraction of bound protein (*Y*) is plotted against the free ligand concentration.

Table 4: Affinity for Binding of Purines to PurR Mutants

	K_d (μM)[a]			
PurR	Gua	Hyp	Ade	6-MP[b]
wild type	**1.6** ± 0.5	**11** ± 1.6	2460 ± 835	>5000
R190A	69 ± 10	101 ± 5.5	134 ± 19	76 ± 8.5
R190Q	**2.0** ± 0.2	**5.5** ± 3.1	**33** ± 11	ND[c]

[a] K_d values (±standard deviation) in bold were determined by equilibrium dialysis and other values by the gel retardation method. [b] 6-MP is 6-methylpurine. The K_d was determined by a fluorescence anisotropy assay. [c] ND, not determined.

mutant repressors. Equilibrium dialysis was generally used to determine binding of hypoxanthine and guanine in cases where the K_d was less than 20 μM and for adenine when the K_d was less than 35 μM. Due to the limited solubility of PurR in dialysis buffer at 4 °C, in cases of low affinity, binding constants were estimated by a gel retardation assay (*11*). Binding constants are summarized in Table 4. For wild type PurR, K_d values of 1.6 and 11 μM were obtained for guanine and hypoxanthine, respectively, in close agreement with previous data (*16*). The apparent K_d value of wild type PurR for adenine was 2460 μM. Thus, the exocyclic atom at position 6 of the purine [a hydrogen bond acceptor for hypoxanthine (O6) and a hydrogen bond donor for adenine (N6)] is mainly responsible for the increased binding affinity of hypoxanthine compared to that of adenine by a factor of more than 220.

Binding of Purines to Arg190 Mutants. To evaluate the proposed key role of Arg190 in corepressor specificity, we determined binding affinities for binding of purines to the R190A and R190Q mutants. Binding curves are given in Figure 3 and K_d values summarized in Table 4. Data in Table 4 show that for the R190A repressor the binding affinity for guanine and hypoxanthine was decreased by 9−40-fold whereas the binding affinity for adenine increased by nearly 20-fold as compared to that of the wild type. These changes in binding constants effectively abolish the 220-fold specificity for hypoxanthine over adenine and 1500-fold specificity for guanine over adenine. In addition, it was found that the R190A repressor bound 6-methylpurine as well as guanine with an observed K_d of 76 μM. Specificity for corepressor was also changed in the R190Q repressor. In this mutant, wild type affinity for guanine and hypoxanthine was retained, while the affinity for adenine was increased over 75-fold. Thus, the specificity for hypoxanthine over adenine was reduced from 220- to 6-fold and for guanine over adenine from 1500- to 16-fold.

Crystal Structures of R190A Bound to the pur Operator and Corepressors Hypoxanthine, Guanine, Adenine, or 6-Methylpurine. To determine the structural basis of R190A binding to the physiological corepressors, hypoxanthine and guanine, and its increased affinity for adenine and 6-methylpurine, we determined four crystal structures: the R190A−hypoxanthine−*pur* operator complex, the R190A−guanine−*pur* operator complex, the R190A−adenine−*pur* operator complex, and the R190A−6−methylpurine-*pur* operator complex. Superimposition of the Cα carbons of these structures and the wild type PurR−hypoxanthine−*purF* operator structure confirms that they are identical with rmsds of 0.16 Å for the R190A−hypoxanthine−*pur* operator complex, 0.16 Å for the R190A−guanine−*pur* operator complex, 0.16 Å for the R190A−adenine−*pur* operator complex, and 0.22 Å for the R190A−6−methylpurine−*pur* operator complex, all of which are within error of the coordinates. As in the wild type PurR−DNA structure, the DNA is kinked about the central CpG base pair step. The DNA bending angle, which was calculated with CURVES (*27*) and excluded the first and last base pairs of the 16 bp operator because of their poor electron density, is 54° in the *pur* operator for R190A bound to hypoxanthine, guanine, and adenine and 52° for the *pur* operator in the R190A−6−methylpurine structure. These values indicate that the DNA structure is essentially identical to that of the PurR−hypoxanthine−*purF* operator structure (54° kink) (*2*, *18*). All other DNA properties, e.g., roll and twist angles and rise per nucleotide, were also nearly identical to those found in the wild type protein−corepressor−DNA structure (*2*, *18*).

To explain the expanded specificity and the equilibrium dissociation binding constants of R190A for these four purines, we predicted that all other protein−corepressor contacts would be maintained and water molecules would be recruited to the binding pocket to replace the missing Arg190 interactions with protein−water−corepressor contacts. Because of the potential of water to donate and accept hydrogen bonds, adenine would be expected to bind as well as hypoxanthine and guanine. Examination of the corepres-

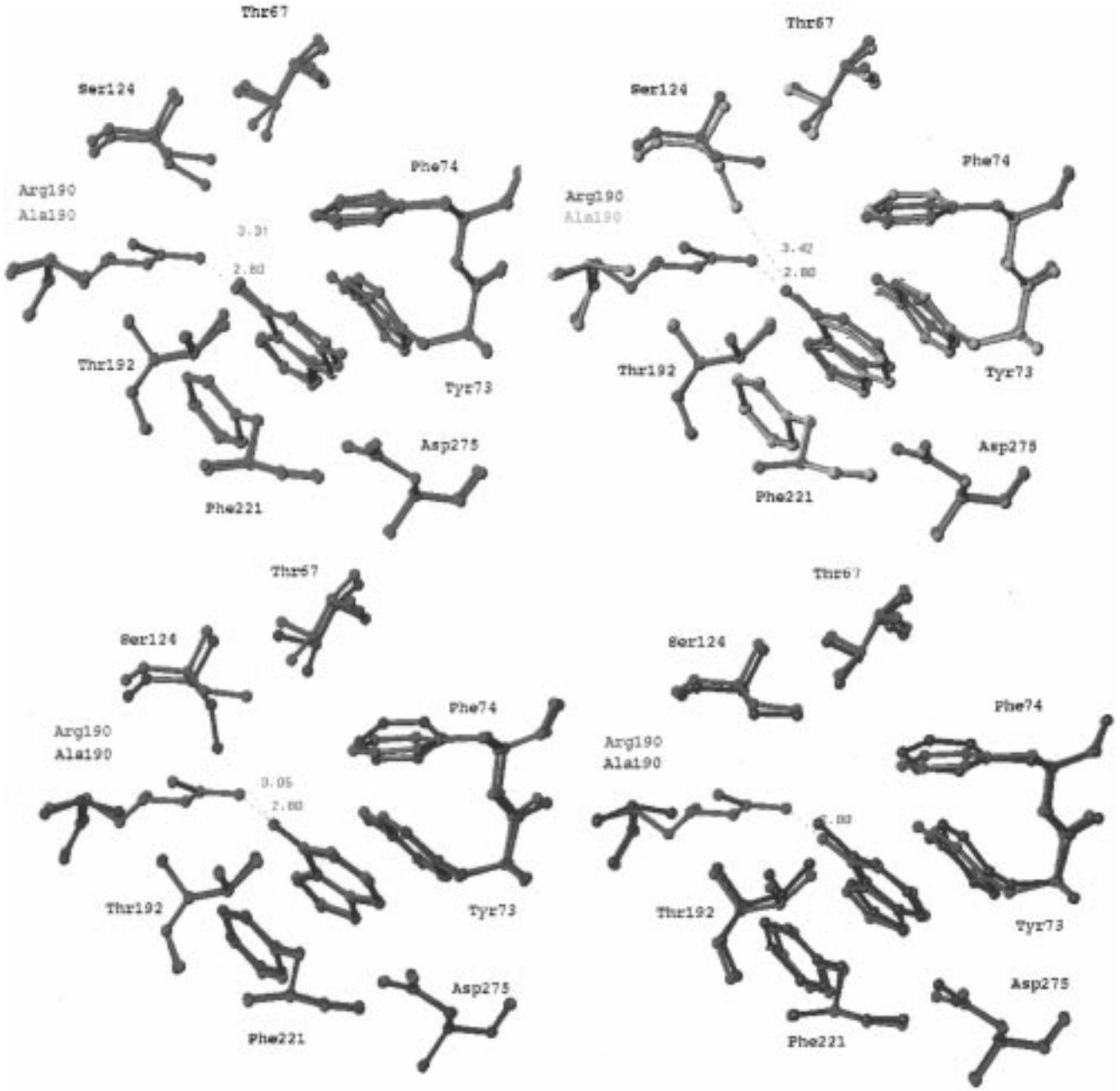


FIGURE 4: Views of the corepressor binding pocket of the R190A mutants after the superimposition of their Cα carbon atoms onto those of the wild type PurR protein bound to hypoxanthine: (A, top left) R190A−guanine (green) and PurR−hypoxanthine (red), (B, top right) R190A−hypoxanthine complex (yellow) and PurR−hypoxanthine (red), (C, bottom left) R190A−adenine (blue) and PurR−hypoxanthine (red), and (D, bottom right) R190A−6-methylpurine (gray) and PurR−hypoxanthine (red). Note the lack of rotation of the Ser124 hydroxyl group in the R190A−6-methylpurine structure. The hydrogen bond distance between the guanidinium NH2 of Arg190 and the O6 oxygen of hypoxanthine is given in each figure as are the hydrogen bonding distances between the hydroxyl OHγ of Ser124 and the O6 oxygens of hypoxanthine and guanine or the N6 of adenine. With the exception of that made by R190, all other wild type protein−purine contacts are found in each R190A mutant−purine complex. These include aromatic interactions between the purines and Tyr73, Phe74, and Phe221 van der Waals interaction to C8 and a hydrogen bond to N7 from the Thr192 side chain; and a hydrogen bond between N9 and the carboxylate group of Asp275. These figures were generated with the molecular graphics computer program O (*24*).

sor binding pockets of the structures of the R190A−corepressor−*pur* complexes reveals, as expected, that none of the remaining wild type protein−purine interactions are missing and that the purines do not take radically different binding positions (Figure 4A−D). Moreover, the large cavity created in the binding pocket by the loss of the Arg190 side chain is indeed filled with several well-ordered solvent molecules. However, none of these water molecules make direct hydrogen bonds to the exocyclic 6 atom. Rather, the side chain of Ser124, which normally makes a hydrogen bond to the side chain of Thr67 and stacks atop the Arg190 side chain, rotates into the binding pocket and donates a hydrogen bond to the O6 of guanine (OH−O6, 3.31 Å) (Figure 4A) and hypoxanthine (OH−O6, 3.42 Å) (Figure 4B) and accepts a hydrogen bond from adenine (HO−N6, 3.13 Å) (Figure 4C). When 6-methylpurine is bound, Ser124 maintains its hydrogen bond to the side chain of Thr67 but makes van der Waals contact to the exocyclic methyl group (Figure 4D).

The basis for the slightly increased affinity of R190A for guanine compared to those of the other three purines is

explained by both a specific protein−base−water mediated contact and protein−purine electrostatic complementarity. These contributions have been described recently for the wild type PurR−guanine−*purF* operator complex (*28*). In the R190A structure, the same interaction is found between a water molecule and the exocyclic hydrogen bond donor N2 atom of guanine (H_2O−N2, 3.33 Å; not shown). This hydrogen bond specifies a donor at this position on the purine ring as both hydrogen bond donors of this water molecule are taken: one by the carboxyl side chain of Glu222 (Oϵ1−HOH, 3.20 Å) and the other by the main chain carbonyl oxygen of Phe221 (CO−HOH, 3.04 Å). Additional specificity for guanine, in both the mutant and wild type proteins, arises from better electrostatic complementarity between the negative corepressor binding pocket and the positive exocyclic amino 2 group of guanine (*28*), a binding feature that is becoming more appreciated (*29*).

Because of the noted importance of the Ser124 side chain in the R190A mutant, Ser124 was replaced by an alanine or an aspartate to assess further its function in corepressor binding. Data in Table 3 indicate that a S124A mutant retained nearly 90% function but when that mutant was combined with an R190A mutation repression was virtually abolished (Table 3). In the latter double mutant, it was thought that an aspartate replacement of Ser124 might retain hydrogen bonding to the 6-amino group of adenine. However, the inactivity of these mutants suggests that adenine either was not bound or if bound was unable to activate PurR. Similar results were obtained for an R190A/S124N double mutant.

Crystal Structures of R190Q Bound to the purF Operator and Corepressors Hypoxanthine, Guanine, and Adenine. To elucidate the stereochemical and structural basis by which R190Q binds the natural corepressors, hypoxanthine and guanine, with unchanged affinity, and adenine with physiologically relevant affinity, we determined three crystal structures: the R190Q−hypoxanthine−*purF* operator complex, the R190Q−guanine−*purF* operator complex, and the R190Q−adenine−*purF* operator complex. Again as anticipated, the corepressor binding pockets of each of the R190Q−corepressor−*purF* operator complexes reveal that none of the remaining wild type protein−purine interactions are missing and that the bases do not take radically different binding positions (Figure 5A−D). Superimposition of the Cαs of each of these structures onto the corresponding atoms of the wild type PurR−hypoxanthine−*purF* operator structure confirms their structural identity and yields rmsds of 0.17 Å for the R190Q−hypoxanthine−*purF* operator complex, 0.27 Å for the R190Q−guanine−*purF* operator complex, and 0.18 Å for the R190Q−adenine−*purF* operator complex (all within error of the coordinates). Again as observed in the wild type PurR−hypoxanthine−*purF* operator and the R190A−corepressor(s)−*pur* operator structures, the DNA is kinked about the central CpG base pair step. The DNA bending angles, which were calculated with CURVES (*27*) and excluded the first and last base pairs of the 16 bp operator because of their poor electron density, are 50°, 50°, and 49° in the *purF* site for R190Q bound to hypoxanthine, adenine, and guanine, respectively, again indicating that the DNA structure is essentially identical to that seen in the PurR−hypoxanthine−*purF* operator structure (*2*, *18*). All other DNA properties, e.g., roll and twist angles and rise per nucleotide, were also nearly identical to those found in the wild type protein−corepressor−DNA structure.

To design a PurR with expanded corepressor specificity of high affinity, we predicted that a glutamine side chain, having both hydrogen bond donors and acceptors, would reorient in the binding pocket to allow the R190Q mutant to bind adenine as well as hypoxanthine and guanine. However, glutamine can replace only one of the two hydrogen bonds donated by the wild type arginine side chain to the purine O6, and we expected some loss of affinity for the natural corepressors. On the contrary, the in vitro corepressor binding data reveal that the R190Q mutant binds guanine and hypoxanthine with wild type affinity and, therefore, must compensate for the loss of this direct hydrogen bond. To do so, the R190Q mutant recruits a water molecule to the position occupied by the guanidinium NH2 of Arg190 in the wild type structure (compare panels A, C, and D of Figure 5 to panel B). From this position, the water molecule allows the R190Q mutant to form a protein−water−purine hydrogen bond that satisfies either the remaining lone pair electrons of the O6 atom of hypoxanthine and guanine or the remaining hydrogen atom of the N6 moiety of adenine. Moreover, in the R190Q−6−oxopurine structures, the glutamine carboxamide nitrogen engages in a hydrogen bond to the Oγ of Thr192 (Figure 5A, D). The result of this hydrogen bond network is a wild type affinity of R190Q for the physiological corepressors (Figure 5A, D) and by a 180° rotation about χ_3, which swaps the positions of the Oϵ and $NH_2\epsilon$ atoms (Figure 5C), its tight binding to adenine (Table 4). In fact, the R190Q mutant displays a more than 74-fold higher affinity for adenine than wild type PurR and thus has extended corepressor specificity as designed. However, as a consequence of the rotation about the χ_3 torsion angle, the side chain oxygen atoms of Gln190 and Thr192 face each other. To avoid their unfavorable electrostatic interaction, the side chains move away from each other slightly (Figure 5C). Although the resolution of the data does not allow us to quantify its movement, adenine is somewhat rotated clockwise in the binding pocket with respect to the other bound purines (Figure 6A) and guanine is pushed up (Figure 6B), presumably to make its best $NH_2\epsilon$−O6 hydrogen bond. Finally, there is a close approach of the Gln190 carboxamide oxygen and purine O6 oxygen atoms in the R190Q−guanine−*purF* operator complex that permits these atoms to make van der Waals contacts (Figure 5A).

DISCUSSION

In the X-ray structure of the PurR−hypoxanthine−*purF* operator complex, residue Arg190 was recognized to be a primary determinant of corepressor specificity by virtue of NHϵ and NH2 hydrogen bonds with the hypoxanthine O6. These interactions require the atom at the exocyclic 6 position of the purine to be a hydrogen bond acceptor and thereby effectively preclude adenine binding. In this work, we have replaced Arg190 with either alanine or glutamine in order to assess directly the function of this residue and the structural consequences of such substitutions. Measurements of repression and in vitro purine binding have confirmed the critical role of this residue in corepressor binding.

Replacement of Arg190 with alanine results in an 18-fold increase in affinity for adenine but diminishes the binding

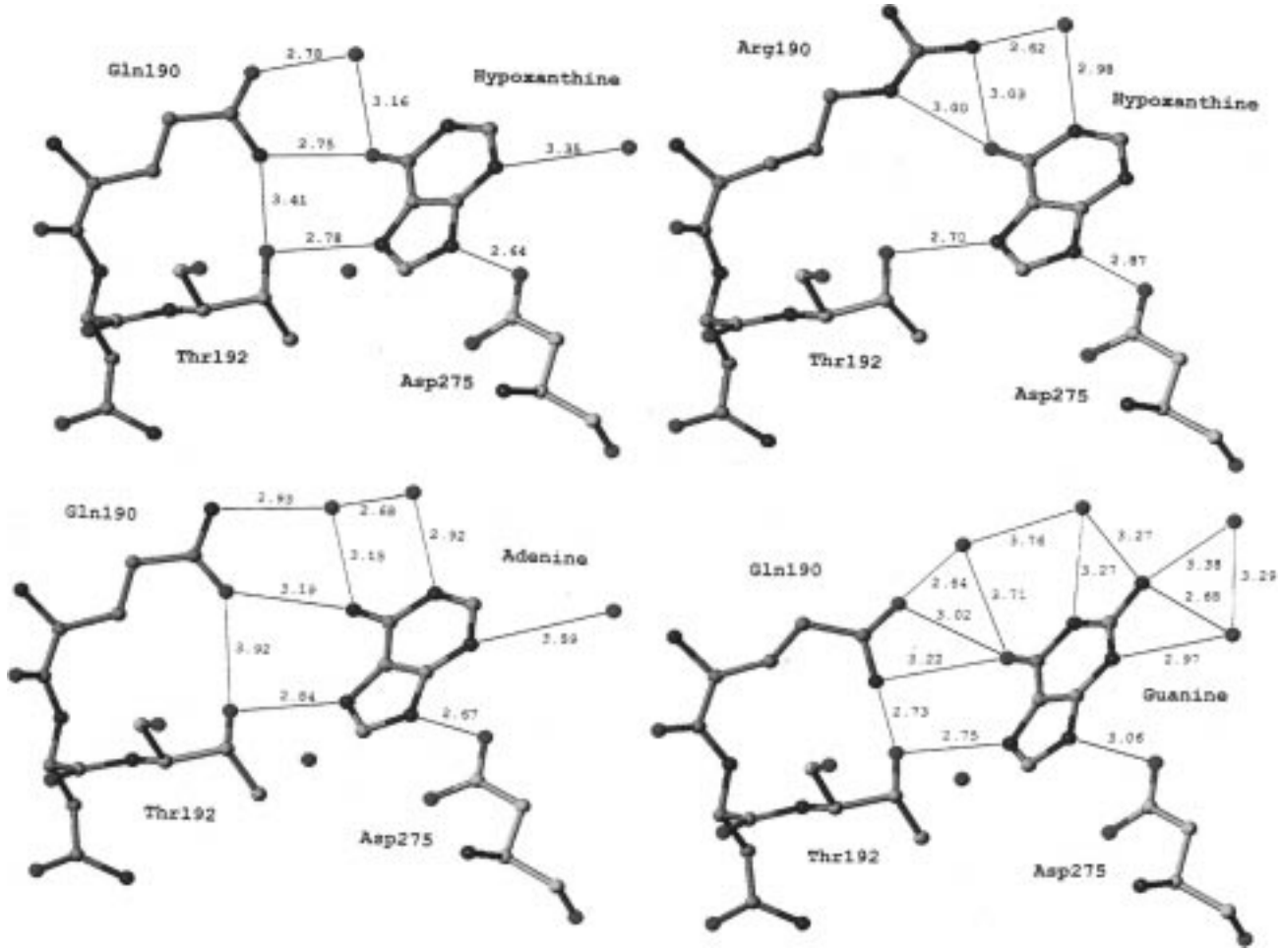


FIGURE 5: Closeup view of the corepressor binding pockets of the wild type PurR−hypoxanthine−*purF* operator and the complexes of the *purF* oligonucleotide with R190Q and hypoxanthine, adenine, and guanine: (A, top left) R190Q−hypoxanthine−*purF* operator complex, (B, top right) wild type PurR−hypoxanthine−*purF* operator complex, (C, bottom left) R190Q−adenine−*purF* operator complex, and (D, bottom right) R190Q−guanine−*purF* operator complex. All hydrogen bond donor−acceptor lengths are given as is the van der Waals contact distance between the Oϵ of Gln190 and the O6 of guanine (see also Figure 6A). In the R190Q−hypoxanthine−*purF* operator complex structure, weak electron density is observed near the N1 of hypoxanthine and is presumably due to the water molecule found in the other structures. However, it is not included here. The water molecule shown donating a hydrogen bond to the O6 of guanine was poorly refined, presumably due to the limited resolution of the data. As described in the figure legend of Figure 4, with the exception of that made by R190, all other wild type protein−purine contacts are found in each R190Q mutant−purine complex. These figures were generated with the molecular graphics computer program O (*24*).

affinity for guanine and hypoxanthine by 43- and 9-fold, respectively. The overall observed higher equilibrium dissociation constants of R190A for guanine and hypoxanthine arise in large part from the loss of three R190−purine interactions, two to the O6 atom and a water-mediated bond to the N1 atom of hypoxanthine or guanine (*2*, *28*). Surprisingly, unlike wild type PurR, the R190A mutant protein can also bind 6-methylpurine, a base without a hydrogen bond donor or acceptor at the exocyclic 6 position, as well as guanine. The crystal structures of R190A in complex with the *pur* operator site and each of these purines reveal the stereochemical bases of their binding. In the absence of the bulky arginine side chain, residue Ser124 breaks its wild type hydrogen bond with Thr67, moves toward the corepressor binding pocket, when occupied by hypoxanthine, guanine, or adenine, rotates its hydroxyl side chain toward their exocyclic 6 position, and engages in hydrogen bonds with each (Figure 4A−C). In this way, Ser124 acts as a replacement for the Arg190 side chain in this protein. However, the hydrogen bonds made between Ser124 and these purines are clearly not enough to restore wild type affinity (Table 4), and their strength, as estimated by the distance between the two interacting atoms, should not be considered the most important indicator of binding affinity (compare the hydrogen bond lengths and K_d values of R190A−adenine and R190A−hypoxanthine). Other factors, such as protein−purine electrostatic complementarity and purine desolvation, especially in the case of 6-methylpurine binding, must play a role (*28*). Another consequence of the replacement of arginine by alanine at position 190 is the increased hydrophobicity of the binding pocket, which aids indirectly the binding of 6-methylpurine. Although found in its wild type location, the Ser124 side chain again assists in this binding by providing van der Waals contacts to the 6-methyl group (Figure 4D).

The substitution of Ser124 by aspartate or asparagine in the context of the R190A replacement was an attempt to increase the affinity of this protein for adenine while decreasing its affinity for hypoxanthine and guanine. However, the double mutants failed completely to repress transcription in vivo (Table 3). The longer side chain of

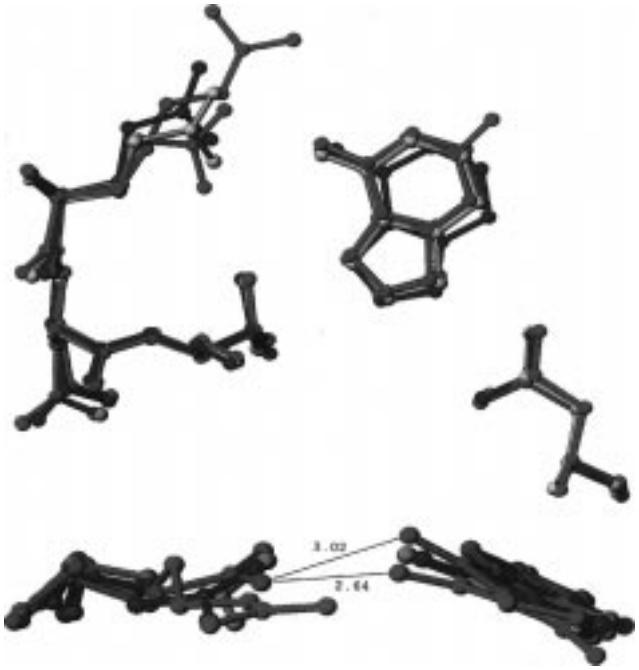


FIGURE 6: Views of the corepressor binding pockets of the R190Q mutant and wild type PurR−purine−DNA complexes after the superimposition of all their Cα carbons using O (*24*): wild type PurR−hypoxanthine, red; Gln190−hypoxanthine, yellow; Gln190−guanine, green; and Gln190−adenine, blue. For clarity, water molecules are not shown. (A) View nearly perpendicular to the plane of the purine ring. Note the relative rotation of adenine. (B) View rotated nearly 90° from that of panel A. Note the upward tilting of the guanine ring (green) to avoid unfavorable van der Waals contacts between the side chain Oϵ and exocyclic O6 atom that would occur if guanine bound to R190Q like hypoxanthine does to wild type PurR (red). This rotation increases the Oϵ−O6 distance from 2.64 to 3.02 Å. These figures were generated with the molecular graphics computer program O (*24*).

the aspartate or asparagine might possibly interfere sterically with corepressor binding.

The R190A mutation reveals the capacity of Ser124 to break its hydrogen bond with Thr67 and rotate into the corepressor binding pocket to hydrogen bond with purines, which contain hydrogen bond donors or acceptors at their exocyclic 6 position, and provides a striking example of the significance of plasticity and adaptability in proteins. Interestingly, Ser124 is one of only two Ramachandran outliers in PurR−corepressor−*pur* complexes (Table 2). This side chain is positioned perfectly to function in corepressor release as it is located above the "specificity loop" (residues 188−193) which undergoes a disorder to order transition upon corepressor binding (*3*). Thus, the strain associated with this side chain may play a role in corepressor binding dynamics. This is supported by the finding that this strain is nearly completely relieved in the unliganded form of PurR (*3*). The significance of such strain has been noted in other structures and is found to be intimately associated with the function of the protein (*30*, *31*). Striking examples include the histidine-containing phosphocarrier protein, HPr, where it has been postulated that the strain associated with Ala6 ultimately facilitates the phosphotransferase process from enzyme I to HPr (*30*) and β-lactamase where two Ramachandran outliers contribute in the positioning of key catalytic components (*32*). Consistent with the finding that ϕ and ψ outliers have significant functional importance, the other Ramachandran outlier of PurR, Asp275, makes a contact to the N9 atom of the physiologicial corepressors (*2*, *18*, *28*) and plays a role in maintaining the architecture of the corepressor binding pocket by making a salt bridge to the side chain of solvent-inaccessible Arg196.

The glutamine replacement of Arg190 was designed to permit this mutant protein to retain its capacity to bind tightly to guanine and hypoxanthine via a side chain $NH_2\epsilon$−O6 hydrogen bond while gaining the ability to bind adenine with high affinity by a side chain Oϵ−N6 hydrogen bond. Although the glutamine side chain can engage in only one direct hydrogen bond to the exocyclic 6 position, as compared to two provided by the arginine guanidinium group, the affinity of the R190Q protein for the natural corepressors is equal to that of the wild type PurR and for adenine is 75-fold higher (Table 4). As a result, in vivo function is indistinguishable from that of the wild type (Table 3). The crystal structures of the R190Q−purine−DNA complexes reveal the chemical basis for the high affinity of this mutant repressor for these purines (Figure 5A−D). In addition to the predicted direct hydrogen bond between the glutamine side chain and the purine O6 or N6 exocyclic groups, the remaining lone pair of the O6 of hypoxanthine or guanine or the remaining hydrogen atom of the N6 of adenine makes a second hydrogen bond to a water molecule that is recruited to the corepressor binding pocket (Figure 5A,C,D).

Clearly, water can play a critical role in multiple biological processes such as enzyme catalysis, thermostability, and protein−ligand binding (*28*, *33*, *34*), and recruitment of a water molecule to replace a missing side chain hydrogen bond donor or acceptor after its removal by site-directed mutagenesis has been reported (*35*, *36*). However, to our knowledge, this is the first structural example in which the recruited water molecule serves to maintain the wild type affinity for the natural ligands and to increase dramatically the affinity for a nonphysiologically relevant ligand. A priori one would think that the localization of these water molecules would be entropically unfavorable and, hence, result in a comparatively lower binding affinity of the mutant protein for the 6-oxopurines. On the other hand, the formation of the observed protein−water−purine hydrogen bonding network is enthalpically very favorable and should contribute significantly to the high affinity. These thermodynamic considerations parallel closely those of wild type PurR. In its corepressor-free form, the Arg190 side chain is on a loop that is completely disordered and, therefore, entropically favorable (*3*). Upon binding corepressor, the loop becomes locked down and the Arg190 side chain ordered. Again, the loss of side chain entropy is offset by the enthalpic gain of the three hydrogen bonds that Arg190 makes either directly or through water to two purine atoms (*2*, *28*).

Despite the high affinity of adenine for R190Q, it remains 6- and 16-fold lower than those of hypoxanthine and guanine, respectively (Table 4). This observation may be explained in part by considering the following series of purine−protein and protein−protein hydrogen bonds: (*1*) the carboxylate side chain of residue Asp275 can only accept a hydrogen bond from the protonated N9 nitrogen of these three purines, (*2*) the hydroxyl side chain of residue Thr192 can only donate a hydrogen bond to the deprotonated N7 of these three purines (Figure 5A,C,D), and (*3*) in the structures of R190Q bound to guanine or hypoxanthine, the observed hydrogen

bond between Gln190 and Thr192 requires that the glutamine side chain be the hydrogen bond donor and the Thr192 side chain the acceptor (Figure 5A,D). However, when adenine is bound and the Gln190 side chain is rotated 180° to satisfy the hydrogen bond acceptor requirements of N6, the hypoxanthine/guanine binding configuration would bring the lone pair electrons of the Thr192 and Gln190 side chains too close. To avoid such proximity, the Gln190 side chain moves away by at least 0.5 Å (Figure 5C). In turn, this movement is the likely cause of the slightly altered location of adenine with respect to the other purines (Figure 6A). Therefore, the absence of hydrogen bonding partners for Thr192 Oγ and Gln190 Oϵ provides a good chemical rationale for the lower affinity of R190Q for adenine than for the natural corepressors. Another contributor to the less favorable R190Q−adenine binding is the poorer electrostatic complementarity between the partial charges on its N1 ($\delta = -0.81$) and C2H ($\delta = 0.07$) atoms and the strongly negative electrostatic nature of the corepressor binding pocket. In support of this latter factor is our recent analysis of the crystal structure of the PurR−guanine−*purF* operator, which revealed that the better electrostatic complementarity between wild type PurR and the stronger partial positive charge of the guanine N1 ($\delta = 0.41$), N2H2 ($\delta = 0.43$), and N2H1 ($\delta = 0.43$) atoms, as compared to their N1 ($\delta = 0.41$) and C2H ($\delta = 0.010$) hypoxanthine counterparts, contribute in part to the higher affinity of guanine for PurR (*28*). A similar contribution to the increased affinity of R190Q for guanine compared to that for hypoxanthine is likely. This, of course, includes the reasonable assumption that the free energies of the corepressor-free conformations of wild type PurR and R190Q are very similar.

In conclusion, PurR mutant proteins R190Q and R190A are redesigned repressors with expanded corepressor specificity. Specifically, the R190A protein binds guanine, hypoxanthine, adenine, and 6-methylpurine with nearly equal affinity, albeit the affinity is reduced with respect to wild type PurR. Their binding affinities result in part from either a new compensatory hydrogen bond formed between the rotated hydroxyl side chain of Ser124 and the exocyclic 6 positions of guanine, hypoxanthine, and adenine or the combination of van der Waals contacts between the Ser124 side chain and the 6-methyl group of 6-methylpurine and the energetically more favorable desolvation of this purine. The R190Q protein, aided by direct and water-mediated contacts, is able to bind adenine with a nearly 75-fold higher affinity than wild type PurR while still maintaining the wild type affinity for guanine and hypoxanthine. By combining the R190Q mutation with additional structure-based substitutions of key residues in the corepressor-binding pocket, we should be able to increase further the affinity of PurR for adenine, e.g., E222Q or E222M, as well as to provide new corepressor specificities for PurR.

ACKNOWLEDGMENT

We thank Dr. Arthur Glasfeld for determining the equilibrium dissociation binding constants of 6-methylpurine for PurR and R190A.

REFERENCES

1. Zalkin, H., and Nygaard, P. (1996) Biosynthesis of purine nucleotides, in *Escherichia coli and Salmonella, cellular and molecular biology* Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S.; Riley, M., Schaechter, M., and Umbarger, H. E., Eds. pp 561−579, American Society for Microbiology, Washington, DC.
2. Schumacher, M. A., Choi, K.-Y., Zalkin, H., and Brennan, R. G. (1994) Crystal structure of LacI member, PurR bound to DNA: minor groove binding by α helices, *Science 266*, 763−770.
3. Schumacher, M. A., Choi, K.-Y., Lu, F., Zalkin, H., and Brennan, R. G. (1995) Mechanism of corepressor-mediated specific DNA binding by the purine repressor, *Cell 84*, 147−155.
4. Haldimann, A., Prahalad, M. K., Fisher, S. L., Kim, S.-H., Walsh, C. T., and Wanner, B. L. (1996) Altered recognition mutants of the response regulator PhoB: a new genetic strategy for studying protein-protein interactions, *Proc. Natl. Acad. Sci. U.S.A. 93*, 14361−14366.
5. Metcalf, W. W., and Wanner, B. L. (1993) Construction of new β-glucuronidase cassettes for making transcriptional fusions and their use with new methods for allele replacement, *Gene 129*, 17−25.
6. Hasan, N., Koob, M., and Szybalski, W. (1994) *Escherichia coli* genome targeting, I. Cre-*lox*-mediated in vitro generation of ori$^-$ plasmids and their in vivo chromosomal integration and retrieval, *Gene 150*, 51−56.
7. Rolfes, R. J., and Zalkin, H. (1988) *Escherichia coli* gene *purR* encoding a repressor protein for purine nucleotide synthesis. Cloning, nucleotide sequence, and interaction with the purF operator, *J. Biol. Chem. 263*, 19653−19661.
8. He, B., Shiau, A., Choi, K. Y., Zalkin, H., and Smith, J. M. (1990) Genes of the *Escherichia coli pur* regulon are negatively controlled by a repressor-operator interaction, *J. Bacteriol. 172*, 4555−4562.
9. Wanner, B. L. (1994) Gene expression in bacteria using Tn*phoA* and Tn*PhoA′* elements to make and switch *phoA* gene, *lacZ*(op) and *lacZ*(pr) fusions, *Methods in Molecular Genetics* (Adolph, K. W., Ed.) Vol. 3, Chapter 15, pp 291−310, Academic Press, Orlando, FL.
10. Atlas, R. M., and Bej, A. K. (1990) Detecting bacterial pathogens in environmental water samples by using PCR and gene probes, *PCR Protocols: A Guide to Methods and Applications.* Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., Eds. pp 399−406, Academic Press, Inc., San Diego, CA.
11. Rolfes, R. J., and Zalkin, H. 1990. Purification of the *Escherichia coli* purine regulon repressor and identification of corepressors, *J. Bacteriol. 172*, 5637−5642.
12. Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) Rapid and efficient site-specific mutagenesis without phenotype selection, *Methods Enzymol. 154*, 367−382.
14. Choi, K. Y., Lu, F., and Zalkin, H. (1994) Mutagenesis of amino acid residues required for binding of corepressors to the purine repressor, *J. Biol. Chem. 269*, 24066−24072.
15. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol. 185*, 60−89.
16. Choi, K. Y., and Zalkin, (1992) Structural characterization and corepressor binding of the *Escherichia coli* purine repressor, *J. Bacteriol. 174*, 6207−6214.
17. Liu, Y.-C., and Matthews, K. S. (1993) Dependence of *trp* repressor-operator affinity, stoichiometry, and apparent cooperativity on DNA sequence and size, *J. Biol. Chem. 268*, 23239−23249.
18. Glasfeld, A., Schumacher, M. A., Choi, K.-Y., Zalkin, H., and Brennan, R. G. (1996) A positively charged residue bound in the minor groove does not alter the bending of a DNA duplex, *J. Am. Chem. Soc. 118*, 13073−13074.
19. Lundblad, J. R., Laurance, M., and Goodman, R. H. (1996) Fluorescence Polarization Analysis of Protein-DNA and Protein-Protein Interactions, *Mol. Endocrinol. 10*, 607−612.

20. Schumacher, M. A., Choi, K.-Y., Zalkin, H., and Brennan, R. G. (1994) Crystallization and preliminary x-ray analysis of an *Escherichia coli* Purine Repressor-hypoxanthine-DNA complex, *J. Mol. Biol. 242*, 302−305.
21. Howard, A. J., Nielsen, C., and Xuong, N.-h. (1985) Software for a diffractometer with multiwire area detector, *Methods Enzymol. B115*; 452−472.
22. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) An efficient general-purpose least-squares refinement program for macromolecular structures, *Acta Crystallogr. A43*, 489−501.
23. Jones, T. A. (1985) Interactive computer program graphics: FRODO, *Methods Enzymol. B115*, 157−171.
24. Jones, T. A., Zou, J. Y.; Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
25. Laskowski, R. A., MacArthur, M. W., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
26. Rolfes, R., and Zalkin, H. (1988) Regulation of *Escherichia coli purF*. Mutations that define the promoter, operator, and purine repressor gene, *J. Biol. Chem. 263*, 19649−19652.
27. Lavory, R., and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids, *J. Biomol. Struct. Dyn. 6*, 63−91.
28. Schumacher, M. A., Glasfeld, A., Zalkin, H., and Brennan, R. G. (1997) The x-ray structure of the PurR-guanine-*purF* operator complex reveals the contributions of complementary electrostatic surfaces and a water-mediated hydrogen bond to corepressor specificity and binding affinity, *J. Biol. Chem. 272*, 22648−22653.
29. Honig, B., and Nicholls, A. (1995) Classical electrostatics in biology and chemistry, *Science 268*, 1144−1149.
30. Herzberg, O., and Moult, J. (1991) Analysis of the steric strain in the polypeptide backbone of protein molecules, *Proteins: Struct., Funct., Genet. 11*, 223−229.
31. Gunasekaran, K., Ramakrishnan, C., and Balaram, P. (1996) Disallowed Ramachandran conformations of amino acid residues in protein structures, *J. Mol. Biol. 264*, 191−198.
32. Jia, A., Vandonselaar, M., Quail, J. W., and Delbaere, L. T. J. (1993) Active-centre torsion-angle strain revealed in 1.6 Å resolution structure of histidine-containing phosphocarrier protein, *Nature 361*, 94−97.
33. Frey, M. (1993) Water structure of crystallized proteins: high-resolution studies, in, *Water and Biological Macromolecules* (Westhof, E., Ed.) pp 98−135, CRC Press, Boca Raton, FL.
34. Jeffrey, G. A., and Saenger, W. (1994) *Hydrogen Bonding in Biological Structures*, Springer-Verlag, Berlin.
35. Alber, T., Dao-pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., and Matthews, B. W. (1987) Contributions of hydrogen bonds to the thermodynamic stability of phage T4 lysozyme, *Nature 330*, 41−46.
36. Vermersch, P. S., Tesmer, J. J., Lemon, D. D., and Quiocho, F. A. (1990) A Pro to Gly mutation in the hinge of the arabinose-binding protein enhances binding and alters specificity. Sugar binding and crystallographic studies, *J. Biol. Chem. 265*, 16592−16603.

BI971942S