Role of Residue 147 in the Gene Regulatory Function of the *Escherichia coli* Purine Repressor†

Joy L. Huffman,‡ Fu Lu,§ Howard Zalkin,§ and Richard G. Brennan*,‡

Department of Biochemistry and Molecular Biology, Oregon Health and Science University, Portland, Oregon, and Department of Biology, Purdue University, West Lafayette, Indiana

Received August 9, 2001; Revised Manuscript Received October 29, 2001

ABSTRACT: The crystal structures of corepressor-bound and free *Escherichia coli* purine repressor (PurR) have delineated the roles of several residues in corepressor binding and specificity and the intramolecular signal transduction (allosterism) of this LacI/GalR family member. From these structures, residue W147 was implicated as a key component of the allosteric response, but in many members of the LacI/GalR family, position 147 is occupied by an arginine. To understand the role of this tryptophan at position 147, three proteins, substituted by phenylalanine (W147F), alanine (W147A), or arginine (W147R), were constructed and characterized in vivo and in vitro, and their structures were determined. W147F displays a decreased affinity for corepressor and is a poor repressor in vivo. W147A and W147R, on the other hand, are super repressors and bind corepressor 13.6 and 7.9 times more tightly, respectively, than wild-type. Each mutant PurR−hypoxanthine−*purF* operator holo complex crystallizes isomorphously to wild-type. Whereas the apo corepressor binding domain (CBD) of W147F crystallizes under those conditions used for the wild-type protein, neither the apo CBD of W147R nor W147A crystallizes, although screened extensively for new crystal forms. Structures of the holo repressor mutants have been solved to resolutions between 2.5 and 2.9 Å, and the structure of the apo CBD of W147F has been solved to 2.4 Å resolution. These structures provide insight into the altered biochemical properties and physiological functions of these mutants, which appear to depend on the sometimes subtle preference for one conformation (apo vs holo) over the other.

The *Escherichia coli* purine repressor, PurR,[1] is a member of the LacI/GalR family of transcription repressors (reviewed in ref *1*). PurR is a 341 amino acyl residue, 38 kDa protein that is a functional homodimer. PurR consists of two domains, an N-terminal DNA binding domain (DBD) encompassing residues 1−60 and a C-terminal corepressor binding domain (CBD) comprised of residues 61−341 (*2*). The CBD is further divided into two subdomains, the N-CBD and C-CBD (*3*). PurR regulates the transcription of at least 18 genes that encode enzymes involved primarily in the de novo biosynthesis of purine nucleotides. A corepressor, hypoxanthine or guanine, must first bind to PurR to induce a conformational change, which enables PurR to bind its 16 base pair operator sites, thereby effecting transcription repression. Thus, the environmental signal of excess purine is transduced into a biological response, namely, repression of de novo purine nucleotide biosynthesis.

Several crystal structures of PurR−corepressor−DNA ternary complexes and the corepressor-free form of the CBD have been solved previous to this work (*2*−*5*). These crystal structures reveal residues involved in DNA or corepressor binding and indicate the route by which the intramolecular or allosteric signal of corepressor binding is passed from the corepressor binding domain to the DNA binding domain, which is over 40 Å from the corepressor binding pocket. Upon binding corepressor, structural changes that include rotations of at least 17° between the NH_2 and COOH subdomains of the CBD of each monomer result in its "closed" conformation (Figure 1A,B). These rotations restructure the interface of the N-subdomains of the CBD, which in turn positions 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 and DNA can subsequently be bound.

The structures of corepressor-bound and free PurR have revealed that residue W147 is one of the keys to intramolecular signal transduction, i.e., the allosteric response, yet the bound purine corepressor does not interact with W147 (*2*). The critical role for this residue is in the corepressor free form of the protein, in which W147 is found in the corepressor binding pocket, where the indole ring stacks with Y73 and F74 and acts as a structural surrogate for corepressor (Figure 1D). The W147 Nϵ also makes a hydrogen bond with the hydroxyl group of Y73, further stabilizing its position in the pocket. When PurR binds corepressor, W147 moves out of the pocket to stack with Y126 and interact with K151 in an "indole binding pocket" (Figure 1C). This translocation is over 9 Å (for W147 Cζ3) and also alters the

† R.G.B. is supported by NIH Grant GM49244. J.L.H. is supported by an NIH predoctoral training grant and a Tartar Trust Foundation grant.

* To whom correspondence should be addressed. E-mail: brennanr@ohsu.edu. Phone: (503) 494-4427. Fax: (503) 494-8393.

‡ Oregon Health and Science University.

§ Purdue University.

[1] Abbreviations: PurR, purine repressor; CBD, corepressor binding domain; hyp, hypoxanthine; *purF*, *purF* operator DNA; DBD, DNA binding domain; LacI, lactose repressor; GalR, galactose repressor; EDTA, ethylenediamine tetraacetic acid.

10.1021/bi0156660 CCC: $22.00
Published on Web 12/14/2001

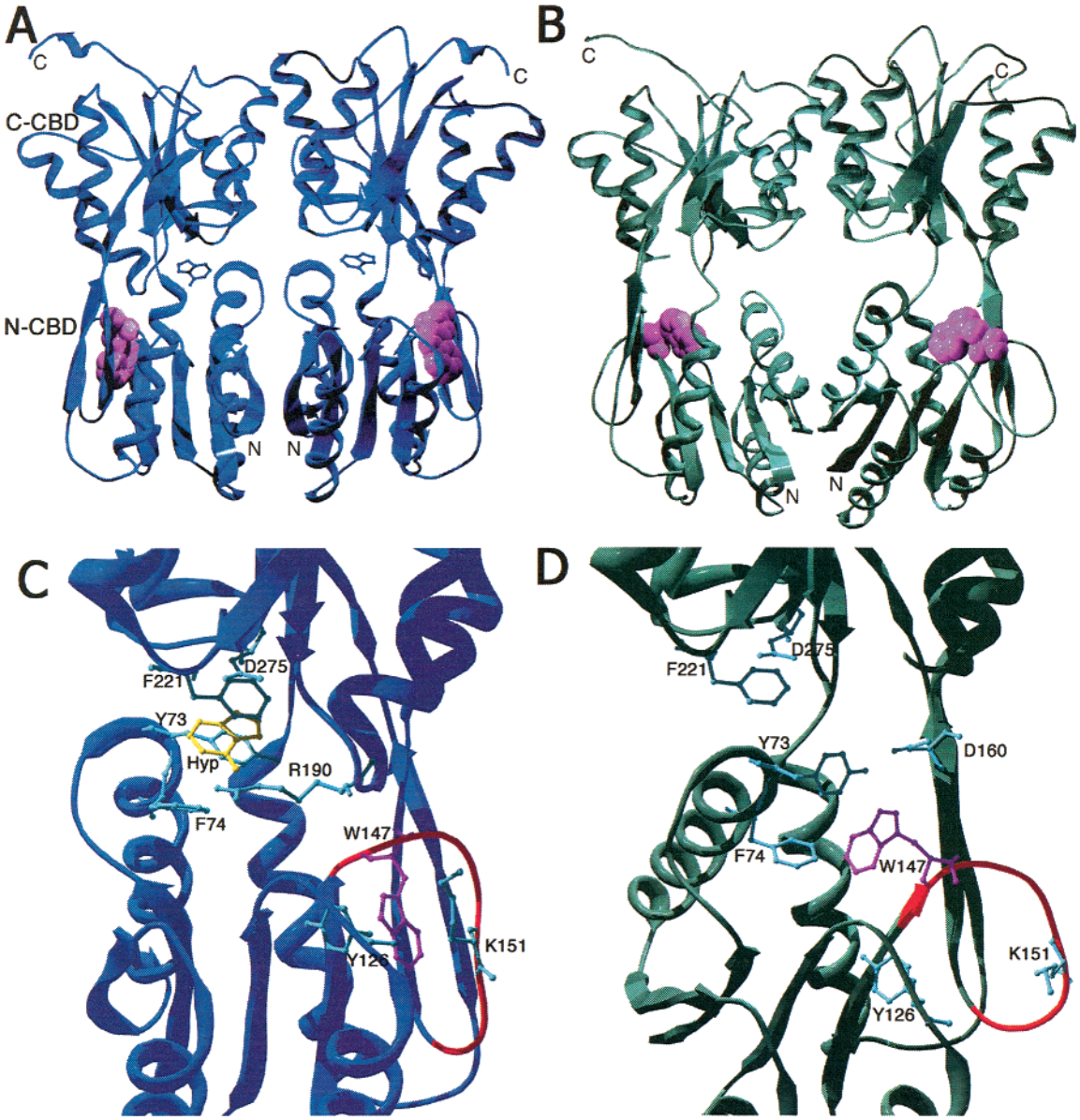

FIGURE 1: The structures of the corepressor binding domains of PurR in the presence or absence of corepressor. (A) In the presence of corepressor (hypoxanthine), i.e., the holo corepressor binding domain (CBD) structure. (B) In the absence of corepressor, i.e., the apo CBD structure. Holo complex structures are shown in blue, and apo CBD structures are shown in aquamarine. Hypoxanthine is displayed as blue balls and sticks, and residue Trp147 is shown as space-filled atoms (pink). Magnified views of the above corepressor binding pockets from the PurR−hypoxanthine−*purF* operator complex structure (C) and the corepressor free PurR CBD structure (D). In panels C and D, the allosteric switch loop is shown in red, Trp147 is shown in pink, other selected residues are shown in cyan as balls and sticks, and the hypoxanthine (in C only) is shown in yellow balls and sticks. The amino and carboxy termini are labeled as N and C, respectively.

position of residues 146−159 (compare Figure 1C,D). A subset of these, residues 146−154, are termed the "allosteric switch loop". The removal of W147 from the corepressor binding pocket to the indole binding pocket allows the change in the monomer−monomer interface of the N-subdomains.

To dissect the role of the indole side chain of PurR residue 147, three substitutions (to phenylalanine, alanine, or arginine) were constructed, their biological and biochemical effects were characterized, and their germane crystal structures were determined. Additionally, we modeled the CBD structures of W147A and W147R in their corepressor free forms because they could not be crystallized. The phenylalanine replacement was designed to test the effect of removing the ability of the side chain to hydrogen bond while retaining the hydrophobicity and ability to engage in the stacking interactions found with the wild-type tryptophan (*6*). The W147A mutant was made to examine the effect of removing the indole side chain altogether. The arginine replacement at position 147 was designed to conserve the side chain length and ability to hydrogen bond while removing the aromaticity of the wild-type tryptophan. Moreover, an arginine is often found at position 147 of LacI/GalR family members (*7*). The combination of the binding assays and crystal structures of these W147 mutants provides insight into the role of the indole side chain in allosteric regulation of PurR.

MATERIALS AND METHODS

Site-Directed Mutagenesis and in Vivo Function of PurR Mutants. Site-directed mutagenesis by the procedure of Kunkel et al. (*8*) was carried out as described previously (*9*). Single-stranded *purR* template DNA was prepared from

plasmid pPR1005-2 (*10*). Mutant genes were constructed and integrated into the phage λ attachment site of the *E. coli* chromosome in single copy to evaluate repressor function. The *purR* gene in this construction is controlled by a weak synthetic promoter chosen to give expression comparable to 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 (data not shown). We note that the repression assay depends on corepressor-dependent binding of PurR to the *purF* operator but does not address which purines function as corepressor.

Overexpression and Purification of PurR. Overexpressed PurR was purified as described (*11*). For purified wild-type PurR, the protein concentration was determined spectrophotometrically at 280 nm using A_{280} (1%) of 12.9 (*10*). This protein was then used to standardize a Bradford assay, as the spectroscopic properties of the mutant proteins were expected to be altered by removal of the tryptophan. An additional purification step was included for the W147A protein used in binding studies, as it had retained some purine base through the normal purification steps. W147A mutant protein was treated with 1 M guanidinium hydrochloride to allow the release of purine base, as described (*12*), followed by extensive dialysis to remove both base and chaotrope.

Isolation of the PurR Corepressor Binding Domain. The CBD of each mutant was prepared by utilizing a trypsin cleavage site between residues R52 and S53 in the hinge region of PurR. Trypsin digestion of PurR mutants was carried out as described previously (*11*). Modifications were made to the procedure for isolating the CBD from the DBD and trypsin. Instead of adding 5 mM phenylmethylsulfonyl fluoride (PMSF) and passing the solution over a sizing column, the trypsin was removed by gently shaking the reaction mixture with trypsin inhibitor cross-linked to agarose beads (Sigma) on ice for 15 min. The beads were pelleted by centrifugation, and the supernatant was removed. This procedure was repeated on the collected supernatant with fresh beads to ensure complete removal of trypsin. The 64 kDa CBD dimer was concentrated to ~10 mg/mL, and its buffer was exchanged to 100 mM NaCl, 5% glycerol, and 20 mM Tris, pH 7.6, using Millipore 50 kDa Molecular Weight cutoff filter devices. This step concentrated the protein for use in crystallization and effectively removed the 6 kDa DBD by filtration. Complete trypsin digestion and removal of the DBD were confirmed by SDS−PAGE (data not shown).

PurR:Corepressor Equilibrium Binding. Equilibrium dialysis was used to determine PurR binding constants for corepressor when the K_d for hypoxanthine was less than 20 μM. The binding measurements were carried out as described previously (*11*). Binding data were fit to the Hill equation using Ultrafit software (Biosoft, Cambridge, UK) (eq 1). Y is the fraction of protein bound; K_d, is the dissociation constant; n, is the Hill coefficient; and L, is free corepressor. Fits were made allowing both K_d and n to vary.

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

In the case where the K_d value was above 20 μM, gel retardation was used to estimate binding of hypoxanthine corepressor to PurR as described previously (*10*). Previous control experiments confirm that the two methods give similar K_d values for the PurR−hypoxanthine interaction (*5*).

PurR Equilibrium Binding to Operator DNA. Fluorescence polarization experiments were done using a PanVera Beacon Fluorescence Polarization system (PanVera Corporation). Samples were excited at 490 nm, and emission was measured at 530 nm. A 5′ fluoresceinated oligonucleotide corresponding to the *E. coli purF* operator (Oligos, Etc.) (5′-F-AAAGAAAACGTTTGCGTACCCCCTACGCAAACGTTTTCTTT-3′) was self-annealed by heating to 80 °C, followed by flash cooling, to form a stem−loop structure with the *purF* operator motif at the center of the base-paired stem. Binding was assayed in a 1 mL volume at 25 °C. The components of each binding experiment were 2 nM fluoresceinated DNA, 250 mM potassium glutamate, 150 mM NaCl, 10 mM magnesium acetate, 1 mM ethylenediamine tetraacetic acid (EDTA), 100 mM potassium HEPES, pH 7.5, and 1.0 μg/mL poly d(IC). Saturating amounts (200 μM) of hypoxanthine were included in the cuvette. The PurR protein was diluted serially between 100 and 0.3 μM and titrated into the cuvette. After each addition of protein, samples were incubated in the Beacon instrument at 25 °C for 30 s before a measurement was taken. A 30 s incubation allowed the binding equilibrium to be reached. The observed mP (milliPolarization) at each titration point represents the average of eight measurements integrated over 6 s.

The data of each binding isotherm were analyzed by least squares regression using SigmaPlot software (Jandel Corporation). The following equation was applied:

$$A = [(A_{bound} - A_{free})[P]/(K_d + [P])] + A_{free} \qquad (2)$$

where A is the polarization measured at a given total concentration of PurR protein [P], A_{free} is the initial polarization of the free DNA, and A_{bound} is the polarization of maximally bound DNA (*4*, *13*). Least squares regression analysis was used to determine A_{free}, A_{bound}, and K_d.

Crystallization and Data Collection. PurR mutants are stored at concentrations of 8−10 mg/mL in 200 mM potassium phosphate, pH 7.5, 0.1 mM EDTA, 0.2 mM dithiothreitol, and 5% glycerol. The 16 bp *purF* operator (Oligos, Etc.) has the sequence 5′-TACGCAAACGTTTTCTT-3′, containing a 5′ overhanging T when annealed to the complementary strand (*14*), and is stored in 10 mM sodium cacodylate, pH 6.5. Crystals of the PurR mutant−hypoxanthine−*purF* operator complex were grown using the hanging drop vapor diffusion method, as described previously (*14*). Briefly, 100 μM protein (8 mg/mL) solution saturated with purine base is mixed with *purF* operator so that a 1:1.5 molar ratio of protein:duplex DNA solution results. This solution is then mixed 1:1 with the crystallization solution, which is 25% PEG 4000, 0.4 M ammonium sulfate, 50 mM cobalt hexammine chloride, and 0.1 M ammonium phosphate, pH 7.5. The crystallization experiments are carried out at 23 °C. Crystals grow first as thin plates at the interface of the phase separation that occurs in the hanging drop. After several weeks to months, these plates yield large (0.6 × 0.3 × 0.2 mm) rhombohedrons. The resulting crystals are isomorphous to wild-type PurR−hypoxanthine−*purF* operator crystals.

Crystals of W147F-CBD were grown using hanging drop vapor diffusion. A solution of ~10 mg/mL protein was mixed 1:1 with a mother liquor of 20% PEG 600, 0.2 M $MgCl_2$,

and 0.1 M Tris HCl, pH 7.6. Crystals typically appear after 1 day but continue to grow for about one week. These crystals are long rods that have dimensions of up to 1.0 × 0.3 × 0.3 mm. W147F-CBD crystals are isomorphous to CBD crystals of the wild-type protein.

X-ray intensity data for the W147F and W147A−corepressor−*purF* operator complexes were collected at 22 °C on an ADSC area detector using a Rigaku RU-200 X-ray generator at 40 kV, 150 mA. The W147R−hypoxanthine−*purF* operator crystal was cryocooled by flash-freezing in liquid nitrogen after a 30 s incubation in a cryoprotectant mother liquor consisting of 25% PEG 4000, 0.4 M ammonium sulfate, 0.1 M ammonium phosphate, pH 7.5, and 35% ethylene glycol. X-ray intensity data were collected at −180 °C on beamline 7-1 at the Stanford Synchrotron Radiation Laboratory (SSRL). X-ray intensity data for the W147F-CBD crystals were also collected at SSRL beamline 7-1, at 25 °C.

Structure Determination and Refinement. Structures were solved using difference Fourier techniques utilizing the wild-type PurR−hypoxanthine−*purF* operator structure (1QPZ) as the initial model for each of the holo complexes, after removing the hypoxanthine, side chain at position W147, and all solvent. The asymmetric unit contains one PurR monomer, one purine base, and a *purF* operator half-site. The wild-type CBD structure (1DBQ) was used as the initial model for the W147F-CBD data.

Refinement of each complex was initiated with rigid body refinement, followed by XYZ (positional) refinement, and finally XYZ and B (thermal) refinement, as implemented in TNT (*15*). Hypoxanthine, the W147X side chains, and solvent molecules were built into the structures in the final stages of refinement using O (*16*). $F_o - F_c$ omit maps were calculated for each structure to ensure the unbiased placement of side chains, purines, and solvent molecules. The components in question were removed from the model, which was then refined until convergence, usually 20 cycles, before inspection of the omit density map. The stereochemical quality of each structure was ascertained using PROCHECK (*17*).

Atomic models of the W147A and W147R corepressor binding domains were built by simply mutating the tryptophan at position 147 in the wild-type structure to either an alanine or an arginine using the program O (*16*). To allow reasonable hydrogen bonding and stereochemistry, and to avoid steric clash, the R147 side chain was rotated, but the remainder of the protein was not altered in both mutant model structures. The W147R model was energy minimized using the simulated annealing energy minimization routine as implemented in CNS (*18*). No significant positional changes in the R147 side chain were observed.

RESULTS

In Vivo Function of PurR Mutants. To investigate the role of W147, which had been implicated in previous in vivo and structural studies (*3*, *9*, *12*) in the allosteric mechanism of PurR, we mutagenized this residue to phenylalanine (W147F), alanine (W147A), or arginine (W147R) and characterized their properties. The W147F mutation conserved the aromaticity of the side chain without allowing wild-type hydrogen bond formation with Y73. The W147A mutation was designed to reveal the effects of removing the entire indole moiety. The W147R mutation is conservative with respect to length and ability to hydrogen bond, but adds a positive charge. Interestingly, arginine is commonly found at position 147 in the LacI/GalR family (*7*, *19*).

Table 1: In Vivo Repression and in Vitro Equilibrium Binding of Wild-Type PurR and W147 Mutants as Determined by Fluorescence Polarization

	In vivo repression, fold	K_d (*purF*), nM	K_d (hyp), μM
wild-type	13.9[a]	2.5 ± 0.3	11.0 ± 1.6[a]
W147F	4.9[a]	2.2 ± 0.3	222 ± 15[a]
W147A	31.3[b]	0.7 ± 0.2	0.81 ± 0.04[b]
W147R	33.6[c]	2.0 ± 0.4	1.4 ± 0.1[c]

[a] These results were reported previously in ref *9*. [b] This result was reported in ref *12*. [c] This result was reported in ref *33*.

Repression of the *purF* operator−*lacZ* reporter by PurR$^+$, W147F, W147A, and W147R was evaluated [(*9*, *12*), Table 1]. W147F displays weaker repression (4.9 fold) than wild-type PurR (13.9 fold). W147A and W147R behave as "super repressors", displaying much higher in vivo repression (31.1 and 33.6 fold) than wild-type PurR.

W147F, A, and R Equilibrium Binding Studies. To determine the effects of these mutations on the ability of PurR to bind corepressor, we determined the in vitro equilibrium dissociation constants (K_d) of each of the W147 mutants and wild-type PurR for both *purF* operator DNA and one of its physiological corepressors, hypoxanthine (hyp). Previous studies have employed equilibrium dialysis and electrophoretic mobility shift assays for the determination of corepressor binding affinities [(*10*, *11*), Table 1]. We used fluorescence polarization to determine the affinities of each protein for *purF* operator. Our results for wild-type PurR agree well with these reported values [previous results: K_d(DNA) = 0.9−2.5 nM, K_d(hyp) = 7−11 μM, (*5*, *10*, *11*)]. Small variations can be attributed to the use of different methods. Binding data from the present experiments are listed in Table 1.

With the exception of W147A, the W147 mutant and wild-type affinities for *purF* operator are essentially identical (2.0 to 2.5 nM, Table 1). W147A bound *purF* operator with a slightly higher affinity (K_d = 0.7 nM) than wild-type PurR. W147F shows a significant decrease in affinity for corepressor (K_d = 222 μM), which is consistent with its weaker repressor phenotype. In contrast, the W147A and W147R mutants have increased affinities for corepressor (0.81 and 1.4 μM, respectively). Again, changes in in vitro corepressor affinity correlate well with in vivo repression activity and are not a result of a dramatically altered DNA affinity.

Global Conformations. Mutants W147F, W147A, and W147R were cocrystallized with the *purF* operator and hypoxanthine by standard methods (*14*), and each crystallized isomorphously to wild-type. We solved these structures by difference Fourier techniques. Selected crystallographic and refinement statistics are listed in Table 2.

As anticipated, the corepressor binding pocket of each of the mutant PurR−hypoxanthine−*purF* operator complex structures reveals that none of the wild-type protein−purine interactions are missing, nor does the hypoxanthine take a radically different binding position (Figure 2). Superimposition of the Cα carbons of the W147F structure onto the corresponding atoms of the wild-type PurR−hypoxanthine−

Table 2: Data Collection and Refinement Statistics for W147F Corepressor Binding Domain and Mutant PurR−Hypoxanthine−*PurF* Operator Complexes

	W147F-CBD	W147F−hyp−*purF*	W147A−hyp−*purF*	W147R−hyp−*purF*
PDB ID	1JHZ	1JFS	1JFT	1JH9
Data Collection:	SSRL[a]	ADSC[b]	ADSC	SSRL
resolution (Å)	2.40	2.90	2.50	2.55
highest resolution shell (Å)	2.45−2.40	3.12−2.90	2.69−2.50	2.61−2.55
unique reflections (#)	20934	13773	23194	22093
completeness (%)	92.2 (93.5)[c]	97.9 (97.7)	97.5 (95.4)	99.8 (100.0)
$I/\sigma(I)$ for data from				
11.5 to 2.0 Å	6.6 (1.9)			
10 to 2.9 Å		4.3 (1.4)		
11.5 to 2.5 Å			7.6 (1.2)	
∞ to 2.55 Å				4.9 (1.4)
Rsym (%)	7.9 (36.9)	7.5 (20.0)	6.4 (23.5)	11.2 (52.4)
temperature (°C)	22	22	22	−180
cell constants:				
a (Å)	38.0	176.1	175.9	175.4
b (Å)	126.1	95.0	95.3	92.6
c (Å)	61.8	81.5	81.4	80.6
α (°)	90	90	90	90
β (°)	99.7	90	90	90
γ (°)	90	90	90	90
Refinement:				
resolution (Å)	2.40	2.90	2.50	2.55
R factor (%)	17.8	15.7	19.1	21.9
total atoms (#)	4489	3040	3085	3121
water molecules (#)	151	32	89	107
rms deviations:				
bond length (Å)	0.009	0.016	0.017	0.016
bond angle (°)	1.018	1.757	1.745	1.655
B correlation (Å^2)	2.662	4.10	4.619	4.604

[a] SSRL = Stanford Synchrotron Radiation Laboratory. [b] ADSC = Area Detector Systems Corporation. [c] Numbers in parentheses refer to the highest resolution shells only.

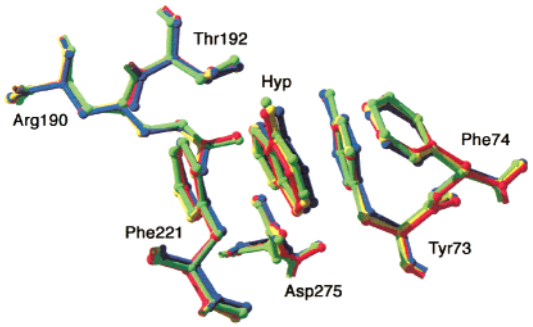

FIGURE 2: Composite overlay of corepressor binding pockets of each W147 mutant onto the wild-type PurR−hypoxanthine−*purF* operator structure. The wild-type PurR−hypoxanthine−*purF* operator structure is shown in yellow, W147F is in green, W147A is in red, and W147R is in blue. All labeled residues interact with corepressor.

purF operator structure (1QPZ) confirms their structural identity and yields a root mean-squared deviation (rmsd) of 0.24 Å (within the error of the coordinates). A similar superimposition of W147A onto the wild-type PurR structure results in an rmsd of 0.23 Å. Superimposition of the Cα atoms of residues 3−340 of the W147R structure onto the corresponding atoms of the wild-type holo complex structure confirmed their structural identity as well and yielded an rmsd of 0.43 Å. The slightly higher rmsd for this structure is likely attributable to small global changes that resulted from collection of the X-ray intensity data from a cryocooled crystal and comparing the resulting structure to the wild-type PurR structure, determined with room temperature X-ray intensity data.

As observed in the wild-type PurR−hypoxanthine−*purF* operator structure, the DNA is kinked about the central CpG base pair step in each mutant-hypoxanthine-*purF* operator structure. The global DNA bend angles (∼45°), which were calculated with CURVES (*20*) and excluded the first and last base pairs of the 16 base pair operator because of their poorer electron density, indicate that the DNA conformations in each structure are essentially identical to that seen in the PurR−hypoxanthine−*purF* operator structure (*2*). 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 (data not shown).

Only the W147F mutant crystallized in the apo conformation, i.e., the CBD form (*21*). These crystals were isomorphous to wild-type, and relevant crystallographic and refinement statistics for that structure determination are listed in Table 2. Extensive de novo crystallization attempts with the W147A-CBD and W147R−CBD proteins were unsuccessful after their failure to crystallize under our standard conditions.

Crystal Structures of PurR W147F−Hypoxanthine−purF Operator and W147F Corepressor Binding Domain. To elucidate the structural basis by which W147F binds the natural corepressor, hypoxanthine, with a 20-fold lower affinity than wild-type, we determined the crystal structures of the W147F−hypoxanthine−*purF* operator complex and the apo W147F corepressor binding domain (Figure 3A,B).

In the corepressor-bound or "closed" conformation, the F147 side chain stacks in an orientation similar to that of the W147 indole ring (Figure 3A). However, the loop that

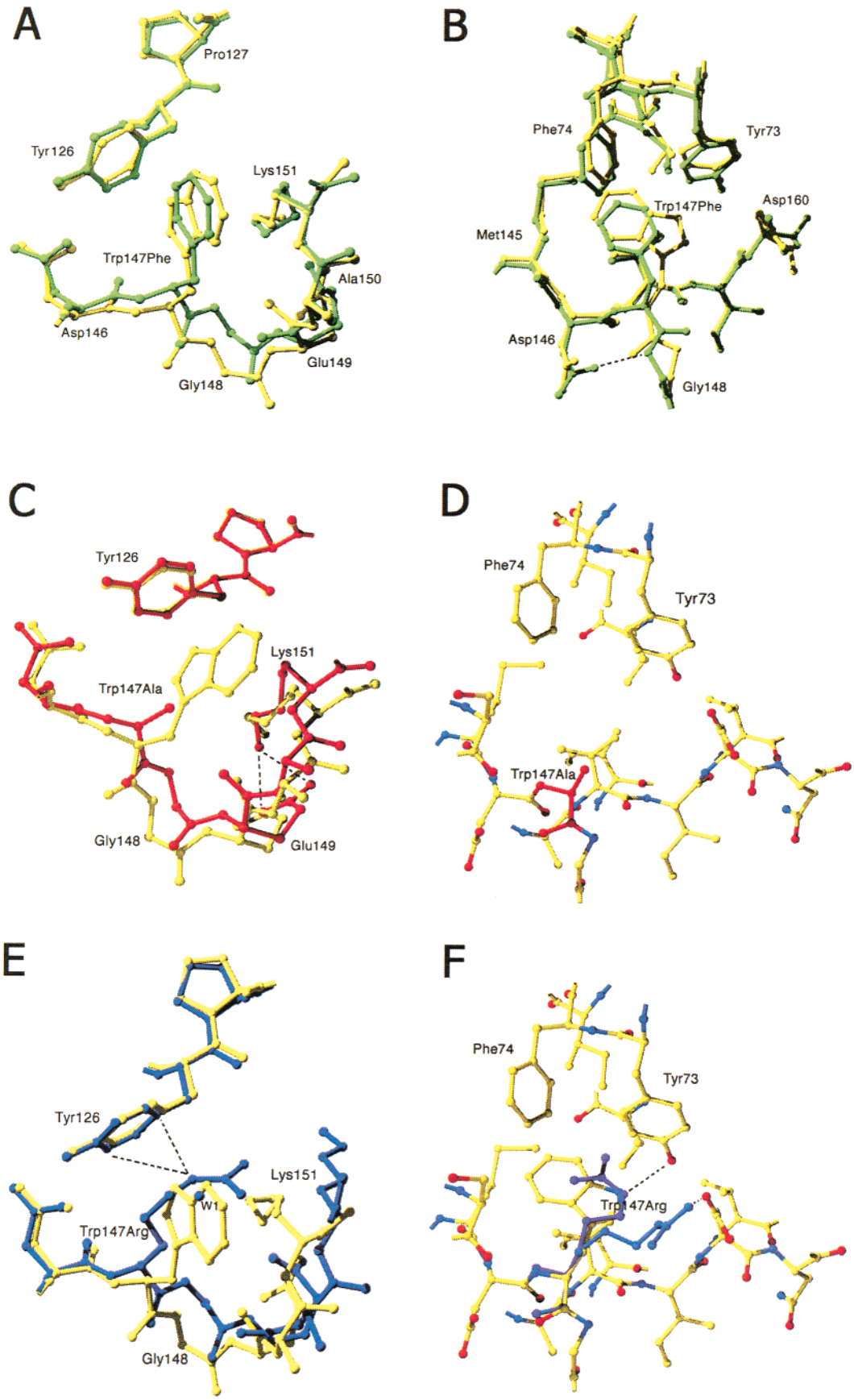

FIGURE 3: Overlays of the corepressor binding pockets of each of the W147 mutant structures onto the wild-type structure. (A) W147F−hypoxanthine−*purF* operator, in green, overlaid onto the wild-type structure, in yellow. (B) W147F-CBD is shown in green, overlaid onto wild-type CBD in yellow. The hydrogen bond between D146 Oδ1 and G148 N (2.7 Å) is shown as a dashed line. (C) W147A−hypoxanthine−*purF* operator, is in red, and wild-type is in yellow. Hydrogen bonds between K151 Nζ and E149 Oϵ1 (2.8 Å) and Oϵ2 (3.0 Å) are shown as dashed lines. (D) Model of the W147A-CBD corepressor binding pocket, with the alanine shown in red. (E) W147R−hypoxanthine−*purF* operator is in blue, and wild-type is in yellow. Cation−π interactions are shown as dashed lines, and W1 indicates a water molecule that is present in the W147R structure only. (F) Two modeled positions for the W147R−CBD. R147 is shown in light blue for one and in magenta for the other and potential hydrogen bonds are shown as dashed lines.

contains residues 146−149, and thus F147, moves toward Y126, and the Y126 phenol ring is rotated ∼25° relative to the wild-type structure. The rmsd between wild-type and W147F Cα atoms within this loop, which is extended to include residues 146−154, is 0.77 Å, three times that of the whole protein. More dramatic are the shifts of the F147 Cα (0.89 Å) and G148 Cα (1.83 Å) relative to their wild-type positions. These changes serve to strengthen the stacking interactions between F147 and Y126. Both the mutant F147 and wild-type W147 side chains make 26 contacts of < 4.0 Å to surrounding residues. In the W147F structure, residue K151 is in an orientation that allows dispersive contacts similar to those seen in the wild-type structure, largely between the phenyl ring and the Cβ and Cϵ carbons of K151. In the wild-type repressor, favorable van der Waals interactions are made between the indole ring and the Cβ and Cϵ methylene carbons of K151, while its Nζ points away from the W147 indole ring. K151 Nζ interacts with the Oϵ2 of E149 in both the W147F and wild-type structures (2.9 Å each), in the former better described as an electrostatic interaction and in the latter as a hydrogen bond. To maintain this interaction in the W147F protein, both Nζ and Oϵ take positions different from their wild-type locations.

The structure of the corepressor-free W147F-CBD is globally identical to wild-type, with a calculated rmsd of 0.47 Å for the Cα atoms of residues 61−341. However, several local changes are apparent that could contribute to an increased stabilization of this conformer (Figures 3B and 4). The phenylalanine side chain is located in the same general position as the wild-type tryptophan and stacks with Y73 and F74. To maintain this arrangement, the Y73 and F74 move closer to the F147. Additionally, the tyrosine at position 73 rotates approximately 25° to engage in better $\pi-\pi$ stacking interactions with the phenyl ring of F147 (*6*). Moreover, the electron density for the Y73 side chain is weak and suggests two possible orientations for the side chain, with one predominant, which is depicted in Figures 3B and 4. We have built this residue only into the strongest density because the electron density and resolution of the data are insufficient for refinement of high and low occupancy sites. The largest structural difference in the W147F-CBD is in the isomerization of the peptide linkage between residues 147 and 148. In this mutant, the protein backbone linkage between residues 147 and 148 is cis instead of trans (Figures 3B and 4). The cis conformation allows the formation of a good hydrogen bond between D146 Oδ1 and G148 N (2.7 Å). In contrast, the D146 Oδ1 of wild-type PurR makes a close contact with the carbonyl oxygen of W147 (2.9 Å), which is either stabilized by a shared proton or is an unfavorable interaction.

Crystal Structure of PurR W147A−Hypoxanthine−purF Operator and Model of W147A Corepressor Binding Domain. To understand more fully why the W147A PurR mutant binds hypoxanthine more tightly than wild-type PurR, and to provide a plausible reason for the "super repressor" phenotype of W147A containing bacteria, we determined the structure of this mutant in the holo complex, with hypoxanthine and *purF* operator. The alanine mutation is considerably more drastic than the phenylalanine replacement, and a larger change in local structure is observed (Figure 3C). The allosteric switch loop, residues 146−154, on which A147 resides, is shifted toward Y126 and more mobile than the wild-type loop. This is evidenced by the higher average thermal factors in this region, 80 Å^2 for W147A protein as compared to 66 Å^2 for wild-type PurR. The rmsd between the Cα carbons of residues 146−154 of wild-type and W147A is 0.89 Å. The Cα of A147 is shifted 0.78 Å, but as observed in the W147F mutant, the most dramatic Cα displacement, 1.69 Å, is observed for G148. This translocation into the indole binding pocket only partially fills the gap created by the removal of the large side chain of the native tryptophan. Whereas the side chain position of Y126 is nearly identical in W147A and wild-type, the side chain of residue K151 of the mutant moves toward Y126 also to help fill the indole ring binding pocket. The absence of the bulky tryptophan results in the rotation of the K151 Cϵ-Nζ bond in the W147A structure, which allows Nζ to hydrogen bond to the Oϵ1 and Oϵ2 atoms of residue E149 (2.8 and 3.0 Å, respectively). In the wild-type structure, only the K151 Nζ-E149 Oϵ1 hydrogen bond (2.9 Å) is observed.

Crystallization trials with the W147A-CBD were unsuccessful using standard conditions, and extensive de novo attempts failed. To visualize the corepressor binding pocket of apo W147A, a simple model was constructed by the removal of the indole side chain (Figure 3D). From this, it would appear that the W147A substitution would disfavor the apo conformation, as a large cavity in this pocket is created that renders the hydrophobic residues Y73 and F74 more solvent accessible, which is expected to be an energetically costly change. Cavity formation has been found to be thermodynamically unfavorable in other proteins, such as ROP, barnase, and T4 lysozyme (*22*−*24*), and might be a major contributor to shifting the conformational equilibrium of PurR toward the holo form. Moreover, using a cutoff distance of 4.0 Å, the wild-type W147 indole side chain makes 29 contacts to atoms of surrounding residues, including hydrogen bonds to the Y73 hydroxyl and D160 carboxylate side chain and numerous van der Waals interactions (Figure 1D). Using the same cutoff distances, no contacts were made to the Cβ of A147, supporting the hypothesis that the apo form of the W147A mutant is thermodynamically less favorable than the wild-type apo protein.

Crystal Structure of PurR−W147R−Hypoxanthine−purF Operator and Model of W147R Corepressor Binding Domain. To elucidate the structural basis of the higher affinity of W147R for hypoxanthine and the "super repressor" phenotype associated with *E. coli* containing this protein, we determined the structure of the W147R−hypoxanthine−*purF* operator complex.

A tryptophan to arginine mutation is relatively more conservative than an alanine replacement, as side chain size and the ability to hydrogen bond are maintained. Yet, very similar in vivo repression activity and in vitro corepressor affinity are observed for the W147R and W147A proteins (Table 1). In the W147R holo complex, significant rearrangement of the indole ring binding pocket is observed and is larger than that seen in the other W147 mutant proteins (compare Figure 3A,C,E). The entire allosteric switch loop, residues 146−154, is more flexible than the wild-type or W147F and is equally mobile as the same loop of the W147A protein, as quantitated by the higher average thermal factors for the Cα atoms in this region, 81 Å^2, as compared to 66 Å^2 for wild-type and 65 Å^2 for W147F. Residues 146−149 move closer to Y126 relative to their positions in the wild-

type structure and assume positions similar to those observed in the structure of the other super repressor, W147A (Figure 3C,E). The rmsd between the Cα atoms of W147R and wild-type for residues 146−149 is 1.14 Å, while that between W147R and W147A in this region is 0.64 Å. Thus, W147R and W147A are structurally more similar to each other than either is to wild-type PurR. On the other hand, W147R residues 150−154 shift away from the R147 and Y126 side chains, which is different from the indole pocket-filling movement seen in the W147A structure (Figure 3C,E). However, the Y126 side chain position is conserved absolutely among W147A, W147R, and wild-type PurR. In the W147R protein, the Y126 phenol ring engages the guanidinium group of the R147 side chain in extensive contacts (18 at a cutoff distance of less than 4.0 Å), including cation−π interactions (*25*), in which the two closest approaches are between R147 Nε and Y126 Cγ (3.0 Å) and Cδ2 (3.0 Å). The K151 side chain has moved farther out of the indole binding pocket, likely to minimize unfavorable charge−charge interactions between the positively charged guanidinium and amino groups. To complete the W147R indole binding pocket, a water molecule has been recruited to the location of the Cζ2 atom of the wild-type indole moiety.

All crystallization trials with the W147R-CBD were unsuccessful. Hence, to gain a rudimentary understanding of that conformation, we constructed a model of the apo corepressor binding pocket of the W147R mutant. When the R147 is modeled to lay flat in the plane occupied by W147, a hydrogen bond analogous to that found in wild-type can be made to the hydroxyl group of Y73 (Figure 3F). However, to allow this hydrogen bond to form, the arginine must push the Nζ1 and Nζ2 atoms of the guanidinium group deep into a hydrophobic region of the corepressor binding pocket, with no suitable hydrogen bonding partners or electrostatic components nearby. In contrast, the indole side chain of the wild-type PurR makes multiple hydrophobic contacts in this cleft to the side chains of I77, M122, M145, V158, and L297. An alternative, and seemingly better, model was constructed by allowing the R147 side chain to rotate and thereby allow its Nε to hydrogen bond with D160 (Figure 3F). However, this leaves the corepressor binding pocket open, and a large cavity is formed (similar to that of the modeled W147A-CBD), which is again likely to be thermodynamically unfavorable due to the exposure of a number of hydrophobic residues and the lack of any of the bracing hydrogen bonds that are seen in the wild-type CBD structure, e.g., between W147 and D160.

DISCUSSION

The crystal structures of the purine repressor in its corepressor bound and free conformations have shown that differences in intersubunit interactions between the N-terminal subdomain of the corepressor binding domain (CBD) and the DNA binding domain (DBD) are characterized by the movement of a "switch" loop, which encompasses residues 146−154. Upon corepressor binding, the removal of this loop from the purine binding pocket allows the transduction of the intramolecular signal from the CBD to the DBD and high affinity DNA binding (*2*, *3*). During this process, the C-terminal subdomain of the CBD moves very little, while the N-terminal subdomains of the CBD rotate ∼20°. Such rotation brings about a new N-terminal interface and juxtaposes the hinge regions of each PurR subunit, thus initiating a coil-to-helix transition of residues 48−56. A ligand-induced conformational change is observed also for the lactose repressor, LacI, the structure of which is remarkably similar to PurR. However, the structural changes in the N-terminal subdomain of LacI are more subtle, as it rotates only 6° and is translated ∼4 Å (*26*, *27*). Regardless, the mechanism of rotating the N-terminal subdomain of the CBD to alter the relative position of the hinge helices, and hence, DBD is conserved between the two repressors.

To explore the functional importance of the indole side chain at position 147, which is a key component of the switch loop, three substitutions, W147F, W147A, and W147R, were made and characterized in vivo and in vitro. Crystallographic studies were undertaken to reveal the stereochemical and structural parameters that contribute to the biological and biochemical functions of each mutant. In initial studies on corepressor binding to PurR, residue W147 was hypothesized to interact with the guanine or hypoxanthine corepressor directly because its mutation to a phenylalanine resulted in weak in vivo repression and lower affinity for hypoxanthine as corepressor (*9*, *28*). Subsequently, the crystal structures of PurR in its holo and apo forms revealed that W147 is involved, but surprisingly this residue makes no direct contacts to the corepressor (*2*, *3*). Rather, the indole ring of W147 acts as a structural, but not functional, surrogate of corepressor that stabilizes the "open", nonrepressing conformation of PurR by occupying part of the corepressor binding pocket. In contrast, when a corepressor is present, W147 translocates to an "indole ring binding pocket", which likely helps to maintain the "closed", repressing conformation by sequestering this weak corepressor competitor.

As previously reported, the W147F mutant is a weaker physiological repressor (*9*, Table 1). This correlates well with in vitro corepressor binding data, in which W147F binds corepressor 15-fold less tightly than wild-type PurR. However, in the presence of excess hypoxanthine, the DNA binding function of W147F is unimpaired (Table 1). The loss of corepressor affinity and the resulting weak in vivo repression of W147F can be attributed to a shift in its conformational equilibrium toward the corepressor-free (apo) conformation by either favoring this conformation or by disfavoring the corepressor-bound (holo) form.

The structure of the apo W147F protein shows that the phenyl ring engages in stacking interactions with the side chains of Y73 and F74. However, the hydrogen bond between the Nε of W147 and Y73 hydroxyl of wild-type PurR is lost and is the likely cause of the higher mobility of the Y73 side chain. Perhaps this gain in side chain entropy offsets at least somewhat the enthalpic loss of the Nε-HO hydrogen bond. Moreover, a new hydrogen bond is formed between D146 Oδ1 and G148 N, which is the result of the isomerization of the 147−148 peptide bond from trans to cis (Figure 4). A nonproline cis peptide linkage is normally energetically less favorable (*29*) but in the context of this mutant is likely stabilizing due to the removal of an unfavorable contact between the W147 carbonyl oxygen and D146 Oδ1 in wild-type PurR. In the corepressor-bound W147F structure, the side chain shifts from the corepressor binding pocket to the indole binding pocket where its phenyl moiety stacks against Y126 in a manner homologous to the

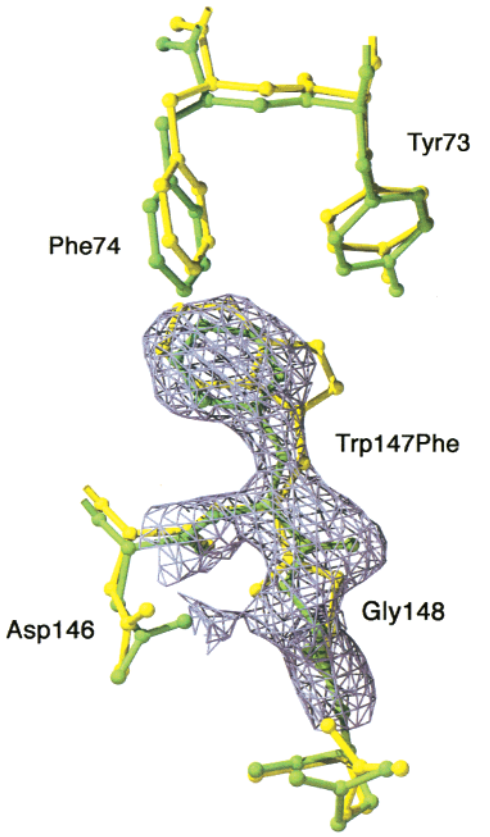

FIGURE 4: $2F_o-F_c$ omitted electron density map and close-up of the W147F (green) and wild-type (yellow) CBD superposition, which underscore the cis to trans peptide bond switch about F147−G148. The electron density is contoured at 1.5 σ and shown as lavender netting.

wild-type W147−Y126 interaction. However, a small cavity is created that is too small to fit a solvent molecule and hence is likely to be energetically unfavorable (*22*−*24*). Thus, the W147F protein is a poorer repressor than wild-type because of its slight preference for the corepressor-free conformation.

The W147A mutant phenotype is that of a super repressor (Table 1). In vitro binding data suggest that this phenotype can be attributed largely to the 14-fold increase in affinity of protein for corepressor. Our inability to grow crystals of the apo W147A CBD suggests that this conformation is disfavored relative to the wild-type PurR CBD. Modeling the W147A protein in the apo (corepressor-free) conformation provides obvious insight into its conformational preference. Unlike tryptophan, the alanine cannot form a hydrogen bond with Y73 or make stacking and van der Waals interactions with Y73, F74, or other residues that the indole ring of wild-type PurR contacts. Moreover, the tryptophan side chain shields interior hydrophobic residues, such as M145 and V158, from the bulk solvent. The exposure of these residues to bulk solvent that results from the W147A substitution, taken together with the energetic cost of creating a cavity and the removal of the indole ring, which can interfere sterically with corepressor access to its binding pocket, provide a plausible underlying foundation for the sharp increase in affinity for corepressor. Thus, W147A may be structurally more similar to the "closed" conformation even in the absence of corepressor, or shifted toward that conformation in the equilibrium of open and closed forms. This is supported by the higher affinity of W147A for *purF* operator than wild-type PurR (0.7 ± 0.2 nM and 2.5 ± 0.3 nM, respectively). Further, there is a small but significant effect on the structure of the W147A holo complex. The allosteric switch loop closes in on the cavity created by removal of the tryptophan, and we see the formation of a new hydrogen bond between the Nζ of K151 and the E149 Oϵ2, a bond that is sterically disfavored in wild-type PurR. We conclude that these interactions result in a shift in the conformational equilibrium of the protein away from the apo conformation. This shift toward the holo conformation is the most likely reason for the super repressor phenotype and high affinity corepressor and DNA binding properties of this mutant. However, our data do not exclude other possibilities, such as the generation of kinetically trapped structural intermediates.

As demonstrated by our in vivo repression experiments, W147R is also a super repressor (Table 1). Those studies are extended by our in vitro binding data, which demonstrate that W147R binds hypoxanthine with 8-fold higher affinity than wild-type, while cognate DNA binding affinity is unaffected. In the W147R−hypoxanthine−*purF* operator structure, we see that the R147 side chain moves closer to the Y126 side chain, and the entire side chain beyond Cβ makes energetically favorable contacts with Y126, including van der Waals interactions and a cation− π interaction (*30*) with the aromatic ring of Y126 (Figure 3E). The corepressor-free form of W147R, if it should resemble the wild-type structure, is likely to be unfavorable because the arginine side chain is unable to make a significant number of favorable interactions with the residues of the corepressor binding pocket or must take unfavorable side chain dihedral angles to engage in hydrogen bonding (*31*). Two models for R147 interactions in the corepressor binding pocket, in which we optimized hydrogen bonding and avoided steric clash (Figure 3F), revealed significant problems. Either a large cavity was formed, as in the W147A-CBD model, or a positive charge was buried in the midst of hydrophobic residues such as I77, M122, M145, V158, and L297. Both are likely to be energetically costly. The difficulty of the accommodation of R147 in the "open" conformation without significant restructuring of the apo corepressor binding pocket, along with the propitious interactions between R147 and Y126 in the indole binding pocket of the holo protein, i.e., the unfavorable interactions generated in the apo conformation and the favorable interactions of the holo conformation, readily account for the super repressor activity of the PurR−W147R mutant.

In conclusion, we have determined corepressor and cognate DNA binding affinities and supporting crystal structures of three allosteric PurR mutants, W147F, W147A, and W147R. Although this residue does not contact corepressor directly, mutations at position 147 have a significant effect on in vivo repression activity and in vitro corepressor binding. This residue and the allosteric switch loop to which it belongs play a key role in the allosteric response mechanism of PurR, not unlike the intrasteric autoregulatory sequences of many allosteric proteins (*32*). The crystal structures of each of these mutants give insight into the complexity of the interactions necessary for intramolecular signal transduction. Interestingly, LacI−inducer and LacI−DNA complexes do not reveal a similar "switch" in transiting between ligand (inducer) and DNA-bound conformations (*26*,

27). The movement of W147 and the entire "switch" loop is linked to the larger rotation of the N-terminal subdomain of the CBD (20° hinge bending), relative to that observed for LacI (6° hinge bending). Perhaps other LacI/GalR family members, which share a tryptophan at this position, such as the *E. coli* ribose repressor, will undergo similarly large rotations, while members that do not have a tryptophan at this position, e.g., the trehalose repressor, which has an arginine at position 147, will undergo smaller, but equally effective movements (*33*). Clearly, structural information from more family members in both their repressing and nonrepressing conformations will be required to determine whether the "switch" loop mechanism is unique to PurR.

REFERENCES

1. Zalkin, H., and Nygaard, P. (1996) in *Escherichia coli and Salmonella, Cellular and Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.) ASM Press, Washington, DC.
2. Schumacher, M. A., Choi, K. Y., Zalkin, H., and Brennan, R. G. (1994) *Science 266*, 763−770.
3. Schumacher, M. A., Choi, K. Y., Lu, F., Zalkin, H., and Brennan, R. G. (1995) *Cell 83*, 147−155.
4. Glasfeld, A., Schumacher, M. A., Choi, K.-Y., Zalkin, H., and Brennan, R. G. (1996) *J. Am. Chem. Soc. 118*, 13073−13074.
5. Lu, F., Schumacher, M. A., Arvidson, D. N., Haldimann, A., Wanner, B. L., Zalkin, H., and Brennan, R. G. (1998) *Biochemistry 37*, 971−982.
6. Burley, S. K., and Petsko, G. A. (1985) *Science 229*, 23−28.
7. Weickert, M. J., and Adhya, S. (1992) *J. Biol. Chem. 267*, 15869−15874.
8. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
9. Choi, K. Y., Lu, F., and Zalkin, H. (1994) *J. Biol. Chem. 269*, 24066−24072.
10. Rolfes, R. J., and Zalkin, H. (1990) *J. Bacteriol. 172*, 5637−5642.
11. Choi, K. Y., and Zalkin, H. (1992) *J. Bacteriol. 174*, 6207−6214.
12. Lu, F., Brennan, R. G., and Zalkin, H. (1998) *Biochemistry 37*, 15680−15690.
13. Lundblad, J. R., Laurance, M., and Goodman, R. H. (1996) *Mol. Endocrinol. 10*, 607−612.
14. Schumacher, M. A., Choi, K. Y., Zalkin, H., and Brennan, R. G. (1994) *J. Mol. Biol. 242*, 302−305.
15. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−501.
16. Jones, T. A. (1991) *Acta Crystallogr. A47*, 110−119.
17. Laskowski, R. A., MacArthur, M. W., and Thornton, J. M. (1993) *J. Appl. Crystallog. 26*, 283−291.
18. Brunger, A. T., et al. (1998) *Acta Crystallogr. D54*, 905−921.
19. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., and Miller, J. H. (1994) *J. Mol. Biol. 240*, 421−433.
20. Lavery, R., and Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.
21. Schumacher, M. A., Choi, K. Y., Zalkin, H., and Brennan, R. G. (1992) *J. Mol. Biol. 225*, 1131−1133.
22. Matthews, B. W. (1996) *FASEB J. 10*, 35−41.
23. Rashin, A. A., Rashin, B. H., Rashin, A., and Abagyan, R. (1997) *Protein Sci. 6*, 2143−2158.
24. Vlassi, M., Cesareni, G., and Kokkinidis, M. (1999) *J. Mol. Biol. 285*, 817−827.
25. Wintjens, R., Lievin, J., Rooman, M., and Buisine, E. (2000) *J. Mol. Biol. 302*, 395−410.
26. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) *Science 271*, 1247−1254.
27. Bell, C. E., and Lewis, M. (2000) *Nat. Struct. Biol. 7*, 209−214.
28. Schumacher, M. A., Macdonald, J. R., Bjorkman, J., Mowbray, S. L., and Brennan, R. G. (1993) *J. Biol. Chem. 268*, 12282−12288.
29. Jabs, A., Weiss, M., and Hilgenfeld, R. (1999) *J. Mol. Biol. 286*, 291−304.
30. Wouters, J. (1998) *Protein Sci. 7*, 2472−2475.
31. Carugo, O. and Argos, P. (1997) *Protein Eng. 10*, 777−787.
32. Kobe, B., and Kemp, B. E. (1999) *Nature 402*, 373−376.
33. Lu, F. (1999) Ph.D. Dissertation, Department of Biology, Purdue University, Purdue, IN.

BI0156660