INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600
RICE UNIVERSITY

CONSTRUCTION AND CHARACTERIZATION OF PURINE REPRESSOR MUTANTS WITH ALTERED OLIGOMERIZATION STATES

by

HAN XU

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
MASTER OF ARTS

APPROVED, THESIS COMMITTEE

Kathleen S. Matthews,
Stewart Memorial Professor
Biochemistry and Cell Biology

George Bennett, Professor
Biochemistry and Cell Biology

Micheal Gustin, Associate Professor
Biochemistry and Cell Biology

George Phillips Jr., Professor
Biochemistry and Cell Biology

Stephen Súbtelny, Professor
Ecology and Evolutionary Biology

Houston, Texas
April, 1997
ABSTRACT

Construction and Characterization of Purine Repressor Mutants with Altered Oligomerization States

by

Han Xu

X-Ray crystallographic structure and deletion mutagenesis of LacI have demonstrated that three-leucine-heptad repeats (LHR3), located at the C-terminus, are critical for dimer-dimer interaction to form tetramer. We are interested in exploring the potential for dimeric proteins, such as purine repressor, which share significant homology to lac repressor, to form higher oligomerization states when this assembly element is present. This information will expand our understanding of the general mechanism of protein oligomer formation.

To assess whether introduction of a LHR motif will elicit tetramer formation, two mutants of PurR, PL3 and PL3+, with different lengths of the LHR3 sequences of lac repressor, have been constructed. A monomer-dimer-tetramer equilibrium is observed for PL3, along with wild-type corepressor binding affinity and decreased operator binding affinity. PL3+ is a trimer, and has no detectable corepressor binding affinity and decreased operator binding affinity. These data indicate that construction of higher order oligomers may be complex.
ACKNOWLEDGMENTS

I would like to thank Dr. Kathleen S. Matthews for her valuable guidance, encouragement and understanding. I would also like to thank my committee members for their help and advice. Special appreciation goes to previous and present members in our lab for their consistent help and friendship. I would also like to thank all my friends and family for sharing smiles and tears in my life.
Table of Contents

Abstract .......................................................................................................................... i

Acknowledgments ......................................................................................................... ii

Table of Contents ......................................................................................................... iii

List of Abbreviations ..................................................................................................... iv

List of Figures ............................................................................................................... vi

List of Tables .............................................................................................................. vii

Introduction .................................................................................................................. 1

Purine and Pyrimidine Nucleotide Biosynthesis ......................................................... 1

Genetic Regulation of Nucleotide Biosynthesis .......................................................... 14

Structure and Properties of Purine Repressor ............................................................ 18

Objectives .................................................................................................................... 33

Materials and Methods ............................................................................................... 36

Plasmids and Strains ..................................................................................................... 36

Polymerase Chain Reaction ......................................................................................... 36

DNA Recombination .................................................................................................... 42

Expression and Purification of PurR, PL3, and PL3+ .................................................. 43

Oligomerization States of PurR, PL3, and PL3+ .......................................................... 44

DNA Binding Activity of PurR, PL3, and PL3+ ........................................................... 46

Corepressor Binding Activity of PurR, PL3, and PL3+ ............................................... 46
Results .................................................................................................................. 48

DNA Construction of PurR Mutants ................................................................. 48

Overexpression and Purification of PurR, PL3, and PL3+ ..................... 53

Oligomerization States of PurR, PL3, and PL3+ ..................................... 56

DNA Binding Activity of PurR, PL3, and PL3+ ....................................... 59

Corepressor Binding Activity of PurR, PL3, and PL3+ ......................... 70

Discussion ............................................................................................................. 76

References .......................................................................................................... 79
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>CBD</td>
<td>corepressor binding domain</td>
</tr>
<tr>
<td>DEAE</td>
<td>diethylaminoethyl</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetraacetate</td>
</tr>
<tr>
<td>$K_a$</td>
<td>equilibrium association constant</td>
</tr>
<tr>
<td>$K_d$</td>
<td>equilibrium dissociation constant</td>
</tr>
<tr>
<td>kD</td>
<td>kilo-Dalton</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>pfu</td>
<td>plaque forming unit</td>
</tr>
<tr>
<td>rpm</td>
<td>rotation per minute</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>SDS-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>w/v</td>
<td>weight/volume</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1.  Synthetic pathway of purine nucleotides ......................2

Figure 2.  Structural formulae of purine bases ..........................4

Figure 3.  The reactions of purine salvage pathways .................7

Figure 4.  Model of the effect of corepressor on PurR-DNA interaction .................................................................9

Figure 5.  Summary of the DNA sequences of “PurR Box” ........15

Figure 6.  Outline of gene regulation in E. coli controlled by PurR ............................................................19

Figure 7.  Amino acid sequence of purine repressor ..................22

Figure 8.  Domain organization of PurR .................................26

Figure 9.  X-Ray crystallographic structure of PurR .................28

Figure 10. The scheme for creating purine repressor mutants with different C-termini from lac repressor ..........49

Figure 11. The different C-terminal LHR motifs of PL3, PL3+, PL3d, PL3d+, and PL5 ..............................................51

Figure 12. Overexpression of PurR, PL3, and PL3+ .................54

Figure 13. The homogeneity of PurR, PL3, and PL3+ ...............57

Figure 14. Equilibrium sedimentation of PurR, PL3, and PL3+ ..60

Figure 15. Oligomerization state of PL3 detected on native polyacrylamide gel ..............................................64
Figure 16. Operator binding of PurR, PL3, and PL3+.................66
Figure 17. Guanine binding of PurR, PL3, and PL3+.................72
List of Tables

Table 1. List of purine biosynthetic enzymes..........................12
Table 2. Genotypes of *E. coli* strains involved in this work........37
Table 3. List of oligonucleotides involved in this work..............40
Table 4. Summary of operator binding activity for PurR, PL3, and PL3+.................................................................68
Table 5. Summary of guanine binding activity for PurR, PL3, and PL3+.............................................................................74
INTRODUCTION

Purine and Pyrimidine Nucleotide Biosynthesis

Nucleotides are very important to all biological organisms in several respects: they are substrates of multiple biochemical processes; they are monomeric units of nucleic acids; and they are the most important energy source for almost all energy-requiring reactions in vivo in their activated forms, i.e. nucleoside diphosphates and nucleoside triphosphates. However, nucleotide synthetic pathways are complicated, involving multi-step processes in the E. coli metabolic system. Purine and pyrimidine ribonucleotides are synthesized via different pathways. Purines are formed initially as ribonucleotides instead of free bases. In contrast, pyrimidines are synthesized as free bases first and then attached to the phosphorylated ribose ring to form nucleotides (Voet & Voet, 1995).

The process of purine, guanine or adenine, nucleotide generation can be divided into two major paths (Fig. 1; Zalkin & Dixon, 1992; Voet & Voet, 1995). The first process is a ten step pathway that converts ribose-5-phosphate to a hypoxanthine-containing molecule, inosine-5'-monophosphate (IMP) (Fig. 2). The second process is the conversion from IMP to either guanine or adenine nucleotide via reactions catalyzed by different sets of enzymes. Nucleoside diphosphates or triphosphates are generated by the
Figure 1. Synthetic pathway of purine nucleotides. Products of energy consuming reactions are underlined. The coding region of enzymes catalyzing the corresponding reaction is listed accordingly. Abbreviations are as defined next to the substrate.
Figure 2. Structural formulae of purine bases. The structure formulae of purine bases, hypoxanthine, adenine, and guanine, and pyrimidine bases, cytosine, uracil, and thymine are shown in this figure.
phosphorylation of nucleoside monophosphates. This pathway to purine nucleotides contains several energy-consuming reactions including phosphorylation, the addition of monocarbons, and the replacement of the carboxamide oxygen by an imino group.

In some circumstances, cells have excess free bases in vivo, either from the active turnover of nucleic acids or from accumulation of purines in the growth media. These free bases can be reconverted to the corresponding nucleotides through salvage pathways. Salvage pathways not only rapidly deplete the pool of free bases, which are toxic to cells (Levine & Taylor, 1982), but also provide a mechanism to save energy otherwise expended needlessly on de novo purine biosynthesis (Fig. 3).

The regulation of purine nucleotide synthesis occurs on two levels. One is the feedback inhibition or feedforward activation of enzyme activities. The accumulation of purine nucleoside mono-, di-, or triphosphates inhibits IMP synthesis. In contrast, the accumulation of PRPP activates the downstream synthetic step. The other level is the transcriptional regulation of those enzymes that are involved in the purine synthesis pathway (Fig. 4). Cells can decide which pathway, regular or salvage, to choose by sensing the availability of free bases and regulating the expression of those genes whose products are involved in the synthetic pathways. The expression of most of these genes is found to be highly regulated so that nucleotides will be generated de novo only
Figure 3. The reactions of purine salvage pathway. When excess purine bases are available, they react with PRPP directly to form AMP, IMP, and GMP. Abbreviations are the same as defined previously. PPi: pyrophosphate.
Adenine + PRPP $\rightarrow$ AMP + PPI

Hypoxanthine + PRPP $\rightarrow$ IMP + PPI

Guanine + PRPP $\rightarrow$ GMP + PPI
Figure 4. Model of the effect of corepressor on PurR-DNA interaction. PurR requires the binding of corepressor, guanine or hypoxanthine, to change conformation from low affinity aporepressor to high affinity holorepressor. The downstream expression is repressed upon the association of PurR.
when required to assure energetic efficiency.

Fourteen genes involved in *de novo* purine synthesis are found dispersed throughout chromosome as either monocistronic or small polycistronic operons (Table. 1). Among these genes, *purD* and *purH*, *purM* and *purN*, and *guaA* and *guaB* are organized into small polycistronic operons, *purHD*, *purMN*, and *guaBA*, respectively (Bachmann, 1983; Flannigan et al., 1990). When sequenced, genes *purE* and *purK* were found to be encoded by the same coding sequence, and therefore named *purEK*. Gene *purF* is associated with two other genes, *cvpA*, which is involved in colicin A production, and *dedF*, which is an undetermined gene so far, to form a small polycistronic operon, *cvpApurFdedF* (Makaroff & Zalkin, 1985; Tso et al., 1982; Fath et al., 1989; Nonet et al., 1989; He & Zalkin, 1994). Genes *purA*, *purB*, *purC*, and *purL* are monocistronic operons (Bachmann, 1983; Tiedeman et al., 1990).

The biosynthesis of UMP, a precursor of all pyrimidine nucleotides, is also a multiple enzymatic process (Neuhard & Nygaard, 1987). There are six enzymes required for *de novo* UMP biosynthesis, and these proteins are encoded by six genes and small operons, designated as *carAB*, *pyrBI*, *pyrC*, *pyrD*, *pyrE* and *pyrF* (Neuhard & Nygaard, 1987). The expression of these genes is highly regulated through multiple mechanisms to assure energetic efficiency.
Table 1. List of purine biosynthetic enzymes. Enzymes involved in purine biosynthesis pathway, their corresponding coding regions, and their positions on *E. coli* chromosomal DNA are listed in this table.
<table>
<thead>
<tr>
<th>Enzymes</th>
<th>E. coli genes</th>
<th>Map position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glutamine PRPP-amidotransferase (EC 2.4.2.14)</td>
<td>purF</td>
<td>50.0</td>
</tr>
<tr>
<td>GAR synthetase (EC 6.3.4.13)</td>
<td>purD</td>
<td>90.3</td>
</tr>
<tr>
<td>GAR transformylase (EC 2.1.2.2)</td>
<td>purN</td>
<td>53.5</td>
</tr>
<tr>
<td>FGAM synthetase (EC 6.3.5.3)</td>
<td>purL</td>
<td>55.2</td>
</tr>
<tr>
<td>AIR synthetase (EC 6.3.3.1)</td>
<td>purM</td>
<td>53.5</td>
</tr>
<tr>
<td>AIR carboxylase (EC4.1.1.21)</td>
<td>purEK</td>
<td>12.2</td>
</tr>
<tr>
<td>SAICAR synthetase (EC 6.3.2.6)</td>
<td>purC</td>
<td>53.3</td>
</tr>
<tr>
<td>Adenylosuccinate lyase (EC 4.3.2.2)</td>
<td>purB</td>
<td>25.2</td>
</tr>
<tr>
<td>AICAR transformylase (EC 6.3.2.6)</td>
<td>purH</td>
<td>90.3</td>
</tr>
<tr>
<td>IMP cyclohydrolase (EC 3.5.4.10)</td>
<td>purH</td>
<td>90.3</td>
</tr>
<tr>
<td>Adenylosuccinate synthetase (EC 6.3.4.4)</td>
<td>purA</td>
<td>95.0</td>
</tr>
<tr>
<td>Adenylosuccinate lyase (EC4.3.2.2)</td>
<td>purB</td>
<td>25.2</td>
</tr>
<tr>
<td>IMP dehydrogenase (EC 1.1.1.205)</td>
<td>guaB</td>
<td>54.0</td>
</tr>
<tr>
<td>GMP synthetase (EC 6.3.4.1)</td>
<td>guaA</td>
<td>54.0</td>
</tr>
</tbody>
</table>

*From Zalkin and Dixon, 1992*
Genetic Regulation of Nucleotide Biosynthesis:

The addition of exogenous purine, such as adenine, causes the repression of all the genes involved in the pathway from PRPP to IMP (Meng et al., 1990; He et al., 1992; Makaroff & Zalkin, 1985; He & Zalkin, 1992; He et al., 1990). The expression of LacZ fused to different pur operator sequences is repressed to different levels: purF, purHD, purEK and purC are tightly regulated by 17-fold, 12-fold, 15-fold and 11-fold, respectively; purL and purMN are regulated to a lesser degree: 6.7-fold and 4.6-fold, respectively. Enzymes involved in the synthesis of AMP and GMP from IMP, purA, purB, and guaBA, are negatively regulated much more weakly, at about 2.5-fold (He & Zalkin, 1994; He et al., 1992; Meng et al., 1990).

The expression of all the operons involved in purine nucleotide biosynthesis is regulated by a negative regulator, purine repressor (PurR) (He et al., 1990; Meng et al., 1990; He & Zalkin, 1994; He et al., 1992; Rolfes & Zalkin, 1988; Makaroff & Zalkin, 1985). In vitro experiments, such as DNAse I footprinting and gel retardation, have shown that purine repressor has high affinity for a consensus sequence, ACGCAAACGGTTTNCGT, although the affinity varies with different sequences (Fig. 5; He et al., 1990; He & Zalkin, 1994). Sequences with high homology to this consensus sequence have been located at the 5' regulatory region of operons of cupA purF dedF (Makaroff & Zalkin, 1985), purC, purMN, purHD, purL, purEK,
Figure 5. Summary of the DNA sequences of "PurR Box". Operator sites that are recognized by PurR are listed. For individual operators, highly conserved nucleotides are in bold. For the consensus sequence (Consen), N = any nucleotide; lower case letters = less conserved nucleotides; upper case letters = highly conserved sequences. ♦ = center of symmetry
and guaAB, between positions -46 and 10 relative to the start codon of transcription (Zalkin & Dixon, 1992; Meng et al., 1990; He et al., 1990). Two consensus sequences have been located at the 5' regulatory region of purA (Meng et al., 1990; He & Zalkin, 1994). Both operator sites are required for efficient repression by PurR. In addition, one consensus sequence has been found far downstream of the promoter and inside the coding region of purB (He et al., 1992; He & Zalkin, 1992).

Although pyrimidine nucleotide biosynthesis is an independent pathway from purine nucleotide synthesis, the regulation of pyrimidine nucleotide production also resides on two different levels. One is the regulation of enzyme activity, and the other one is transcriptional regulation. The expression of all the enzymes involved in pyrimidine biosynthesis, including pyrC and pyrD, is primarily regulated by the in vivo pool level of pyrimidine nucleotides (Neuhard & Nygaard, 1987; Schwartz & Neuhard, 1975). However, the addition of exogenous free purine also causes the repression of pyrC and pyrD (Wilson & Turnbough, 1990; Choi & Zalkin, 1990). Consensus operator sequences are also located in the regulatory regions of these two genes (Fig. 5; Wilson & Turnbough, 1990; Choi & Zalkin, 1990; Vial et al., 1993). In addition, purine repressor regulates the expression of several other genes, such as codA (Kilstrup et al., 1989), required for pyrimidine salvage pathway, glyA (Steiert et al., 1990) whose gene product,
hydroxymethyl transferase, provides one carbon units for purine nucleotide synthesis, glycine inducible gcv system (Plamann et al., 1983), which encodes enzymes responsible for glycine cleavage to NH₃ and CO₂ (Wilson et al., 1993), glnB, which encodes a protein that functions in conversion from glutamate to glutamine, prsA, which encodes PRPP synthetase that provides PRPP for nucleotides and amino acid synthesis, as well as speA, which is involved in the synthesis of polyamines (He et al., 1993). Similar DNA sequences have also been found at the regulatory region of these genes, although with less homology than observed for purine biosynthetic enzymes (Fig. 5). Two consensus sequences have been located at the regulatory region of purR, indicating that the transcription of purine repressor is autoregulated (Meng et al., 1990; Rolfes & Zalkin, 1988; Rolfes & Zalkin, 1990b). One of the operator sites is downstream from the promoter in the 5' mRNA untranslated region, and the other one is at the beginning of the coding region of purR. All these results demonstrated that purine repressor is a master element which regulates several divergent biosynthetic pathways (Fig. 6).

**Structure and properties of purine repressor:**

The coding region for purine repressor has been cloned and sequenced (Rolfes & Zalkin, 1988). PurR has been overexpressed and purified to almost homogeneity (Rolfes & Zalkin, 1990a). PurR is a homodimer with 341 amino
Figure 6. Outline of gene regulation in *E. coli* controlled by purine repressor. Genes regulated by PurR are indicated by dashed arrows. Corepressors, guanine and hypoxanthine, are underlined. Abbreviations are the same as defined previously. C<sub>1</sub>-FH<sub>4</sub> is 5,10-methylenetetrahydrofolate.

This figure is adapted from Zalkin and Dixon, 1992.
acids (Fig. 7). It is a member of the LacI family of DNA binding proteins (Schumacher et al., 1993). There are more than 21 members in the LacI family, including CytR (Valentin-Hansen et al., 1986), GalR (von Wilcken-Bergmann & Müller-Hill, 1982), RibR (Mauzy & Hermodson, 1992), MalR (Reidl et al., 1989), RafR (Aslanidis & Schmitt, 1990), evolved β-galactosidase repressor (Stokes & Hall, 1985), FruR (Jahreis et al., 1991), B. subtilis amylase repressor protein (Henkin et al., 1991), among others. Except for tetrameric protein lac repressor (Alberti et al., 1991; Chakerian et al., 1991; Alberti et al., 1993; Chang et al., 1993), fructose repressor (Cortay et al., 1994), and raffinose repressor (Jaenicke et al., 1990; Aslanidis & Schmitt, 1990), the remaining members of this family are dimers (Gralla, 1992; Rolfes & Zalkin, 1988; Weickert & Adhya, 1992). Members of this family contain two separate functional domains: a helix-turn-helix (HTH) DNA binding domain (Brennan & Matthews, 1989) and a ligand/dimerization domain (Gralla, 1992; Rolfes & Zalkin, 1988; Weickert & Adhya, 1992). LacI forms a tetramer via a four-helical-bundle formed by four identical 3-leucine-heptad-repeats located at its C-terminus (Alberti et al., 1993; Alberti et al., 1991; Chakerian et al., 1991). PurR also contains a N-terminal DNA binding domain and a corepressor binding domain (CBD), which can be separated by proteolytic cleavage at Arg57 of PurR (Choi & Zalkin, 1994).
Figure 7. Amino acid sequence of purine repressor.
<table>
<thead>
<tr>
<th></th>
<th>Met</th>
<th>Ala</th>
<th>Thr</th>
<th>Ile</th>
<th>Lys</th>
<th>Asp</th>
<th>Val</th>
<th>Ala</th>
<th>Lys</th>
<th>Arg</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>210</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>290</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>340</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
LacI family members display a high degree of sequence homology throughout their entire sequence, especially in the DNA binding domain (Weickert & Adhya, 1992). PurR and LacI share almost 50% identity within the DNA binding domain (Schumacher et al., 1993). The region critical for dimerization also displays significant sequence identity among all LacI members. PurR, LacI and CytR share about 60% identity within this region (Schumacher et al., 1993).

Structural homology has been found between the LacI family of DNA binding proteins and another group of E. coli binding proteins, the periplasmic binding proteins (PBP), despite the low level of sequence homology (Müller-Hill, 1983; Vartak et al., 1991; Mauzy & Hermodson, 1992; Weickert & Adhya, 1992). PBPs contain a short N-terminal signal sequence and a large metabolite binding domain (Spurlino et al., 1991). The signal sequence targets these proteins to the inner cell membrane and is cleaved thereafter. The metabolite binding domain is divided into two subdomains (designated as N-subdomain and C-subdomain) connected by several crossovers (Vyas et al., 1991; Mowbray & Cole, 1992). The homology between PBP and LacI family proteins has been found within the ligand binding domains of LacI family proteins and the metabolite binding domains of PBP. The CBD of purine repressor is strikingly homologous to the metabolite binding domain of the group of periplasmic binding proteins,
especially to ribose binding protein (RBP) (Bowie et al., 1991). The amino acid sequence alignment of purine repressor with RBP and other periplasmic binding proteins led to the prediction of the tertiary structure map for the CBD of purine repressor (Schumacher et al., 1993). The X-ray crystallographic structure of PurR bound to one of its corepressors, hypoxanthine, and a cognate DNA, a 16 bp purF operator, as well as the corepressor free-CBD of PurR have been determined (Schumacher et al., 1994a, b; Schumacher et al., 1992). The crystal structure confirms that PurR has a PBP like corepressor binding domain and a helix-turn-helix DNA binding domain (Fig. 8, Fig. 9).

**DNA binding domain:** The DNA binding domain can be divided into two important functional groups, the HTH motif (residues 4-23) and the adjacent short helix (residues 30-43) connected by a short loop (residues 24-29). PurR recognizes pseudo-palindromic, 16-base pair, operator sites. The binding affinity to purF operator was determined, and the dissociation constant (K_d) was found to be 3.4 nM (Rolfes & Zalkin, 1990a). Amino acid side chains of both DNA binding domains make extensive contacts to the bases and phosphate backbone of purF operator. The HTH motif contributes the primary component of specific recognition and contacts the DNA in the major groove. The function of the following turn and the third helix is to provide structural support for the interaction between the HTH and operator DNA, as well as
Figure 8.  Domain organization of PurR
Figure 9. X-Ray crystallographic structure of PurR. Helix-turn-helix DNA binding domains are in blue. Hinge helices are in green. Corepressor binding domains are in yellow. This structure is from Schumacher et al. (1994) is constructed from the coordinates in PDB file number 1BDI.
making some contacts with the operator DNA itself (Schumacher et al., 1994a; Nagadoi et al., 1995). A helix was found in the region that connects the HTH and core binding domains in the X-ray crystallographic structure (Choi & Zalkin, 1994; Schumacher et al., 1994a). This "hinge helix" (residues 49-57) also makes contact with the cognate DNA, but in the minor groove rather than major groove (Schumacher et al., 1994a). The hinge helices interact with each other within the dimer through a series of van der Waals interactions between side chains of Val50, Val50', Leu54, and Leu54' to form double helices. The insertion of the hinge helix into the minor groove causes an opening of the minor groove by the side chains of Leu54 and Leu 54', which elicits a "kink" of 45° at the central basepair of the operator DNA. The consequence of this kink is a DNA bend toward the major groove and away from the protein (Schumacher et al., 1994a).

Corepressor binding and dimerization domain: Hypoxanthine and guanine are the corepressors of PurR based on in vivo and in vitro experiments (Rolfes & Zalkin, 1990a; Meng & Nygaard, 1990). These compounds bind to purine repressor with affinities of \( \sim 10^{-5} \) M and \( \sim 10^{-6} \) M, respectively, measured by equilibrium dialysis and intrinsic fluorescence signal change (Choi & Zalkin, 1992).

The CBD of both holo- and apo-purine repressor consists of two subdomains, the N-subdomain and the C-subdomain, linked by three
noncontiguous crossovers (Schumacher et al., 1994a; Schumacher et al., 1995). The corepressor binding site is located in the cleft between these two subdomains. Amino acids responsible for sugar binding are conserved among PBP family members, and include Asp146, Arg196 and Asp275. These residues of PurR are critical for corepressor binding (Choi et al., 1994). Trp147 is also required for corepressor binding (Choi et al., 1994). In aporepressor, Trp147 is found in the corepressor binding pocket, replacing the corepressor structurally without eliciting the conformational shift required for specific DNA binding. Trp147 moves out of the binding cleft upon corepressor binding (Schumacher et al., 1994a; Schumacher et al., 1995). The residues on the surface formed by the two subdomains form a high affinity ligand binding site utilizing polar and non polar interactions. The side chains on the surface of this pocket are also responsible for the specificity of ligand binding. In particular, the hydrogen bond between ε-NH$_2$ of Arg190 and O$_6$ acceptor of hypoxanthine allows PurR to distinguish the exocyclic atom of a purine ring from the NH$_2$ donor group of adenine (Schumacher et al., 1994a).

The CBD is the primary site of interaction between two monomers so that it plays the key role in dimerization (Schumacher et al., 1994a). Subunit interactions between the two N-subdomains are generated by residues 68 through 115. The contact between the two C-subdomains is scattered among residues 223-229, 249-267, 278-285, and 328-328. Interestingly, three cross-
subunit interactions between the DNA binding domain of one subunit and the CBD of the other subunit are generated by Gln113, Ala49, Arg115, Ser46. In addition, two salt bridges connect the N-subdomain of one subunit to the C-subdomain of the other subunit.

**Effect of corepressor dissociation on DNA binding affinity:** Most LacI family proteins have higher DNA binding affinities in the conformation without their effectors. In contrast, purine repressor requires either guanine or hypoxanthine as a corepressor to enhance the DNA binding activity.

The solved crystal structures of hypoxanthine-bound purine repressor and corepressor-free CBD demonstrated that there is a global conformational change of the CBD upon dissociation of corepressors (Schumacher *et al.*, 1995). With dissociation of the corepressors, the C-subdomains remain essentially in the same position relative to each other, while N-subdomains rotate about 17° to 23°. This rotation appears to derive from hinge bending motions that affect the amino acids found in the crossover segments between the CBD subdomains (Schumacher *et al.*, 1995). This rotation ultimately results in an open conformation of aporepressor relative to the closed conformation of holorepressor. In the open conformation of CBD, the DNA binding domain and the hinge helix appear to be rearranged with respect to one another (Schumacher *et al.*, 1994a, Schumacher *et al.*, 1995; Nagadoi *et al.*, 1995). NMR studies on the DNA binding domain along with the hinge
helix of PurR demonstrated that amino acids in the hinge helix region form no secondary structure in the absence of operator DNA (Nagadoi et al., 1995).

These studies provide the basis for a hypothesis for the mechanism of DNA binding regulated by corepressor. The regulation of DNA binding via association and dissociation of corepressors includes two steps in the signal transduction. The first step is from CBD to hinge helix. The correct position of the hinge helix from each monomer is induced by the proper conformational change upon corepressor binding, and the conformational change of CBD caused by dissociation of corepressors destabilizes the hinge helices and causes them to unravel. The second step is from DNA binding domain to hinge helix. The specific DNA binding initiated by the DNA binding domain activates the local folding of the hinge helices. Dissociation of corepressor causes hinge helices to lose contact with each other, and, in turn, to lose contact with the minor groove of DNA and unfold. Dissociation and unfolding change the orientation of the HTH so that it cannot occupy the major groove in a correct orientation, resulting in loss of contacts in the major groove. In this manner, the DNA binding activity is abolished after the dissociation of corepressors (Schumacher et al., 1995).

Objectives

Protein association is a common feature in genetic regulation. Almost all transcriptional and translational apparati in both prokaryotic and
eukaryotic cells involve multiple proteins associating with each other. In addition to the interactions among proteins with different functions, in many circumstances, protein self-association is required for maintaining or promoting proper function. Therefore, understanding the mechanism of protein-protein interactions is very important.

The LacI family of DNA binding proteins includes dimeric and tetrameric proteins and provides an especially useful system for understanding the interaction between identical dimers or monomers. The homology among these family members and the different oligomerization states provide the opportunity for the alteration of oligomerization state by insertion, deletion or domain swapping mutagenesis. X-Ray crystallographic structural studies have demonstrated that LacI is a dimer of dimers (Lewis et al., 1996; Friedman et al., 1995). Two dimers are associated by the four-helical-bundle, formed by four identical three-leucine-heptad-repeats (LHR3) at the C-terminus of LacI. Previous mutagenesis studies have successfully constructed dimeric LacI mutants by deleting different lengths of the leucine-heptad-repeats (Li & Matthews, 1995; Chen & Matthews, 1992). Thus, the four-helical-bundle is designated as tetramerization domain. A single amino acid change at residue 282, from tyrosine to any amino acid other than phenylalanine and leucine, completely disrupts the monomer-monomer interface and creates a monomer (Chakerian & Matthews, 1991). However, elongation of the tetramerization domain from LHR3 to five-leucine-heptad-repeats (LHR5) not only stabilizes
the tetramer of LacI, but also produces a dimer when coupled with monomeric mutant Y282D (Chen et al., 1994). These results promote the interest in whether the tetramerization domain of LacI is sufficient for other homologous proteins within LacI family, such as PurR, to form tetramer.
MATERIALS AND METHODS

Plasmids and Strains

Plasmid pPR1005-2, kindly provided by Dr. Howard Zalkin at Purdue University, is a derivative of pUC119, and contains the purR coding region. Plasmid pT7-7, a derivative of a pET vector (Rolfes & Zalkin, 1990a), is for the expression of purR and its mutants. Plasmid pPR1006 contains two PurR operator binding sites, which can be digested out of the plasmid singly or in combination by restriction enzymes. So, this plasmid is used as the source of PurR operator with different numbers of binding sites. Plasmid pJC1 was created by Dr. Jie Chen (Chen & Matthews, 1992), and it is a derivative of pEMBL9+ with the coding region of wild-type lacI. A variety of mutants of lac repressor were generated using pJC1 as a backbone.

_E. coli_ 71/18 has been used to grow plasmid DNA. _E. coli_ DH5α has been used for the cloning of PCR fragments. _E. coli_ BL21 has been used for the expression of PurR and its mutants. The genotypes of all strains involved in this research are listed in Table 2.

Polymerase Chain Reaction

The coding region of LHR3 (Arg340-Gln360 from _lac_ repressor), LHR3+ (Pro332-Gln360 from _lac_ repressor), LHR3d (Arg340-Gln360 from the L349A mutant of _lac_ repressor), LHR3d+ (Pro332-Gln360 for the L349A
Table 2. Genotypes of *E. coli* strains involved in this work.
<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>71/18</td>
<td>$F' \ lac^F \Delta(lacZ)M15proA^+B^+/(\Delta(lac-proAB)\ thi\ supE)$</td>
</tr>
<tr>
<td>BL21</td>
<td>$F^-\ ompT\ hsdS_\beta(r_B^+m_B^-)\ gal\ dcm$</td>
</tr>
<tr>
<td>ED8739</td>
<td>$F^-\ metB\ supE\ supF\ hsdS(r_{12}^+m_{12}^-)$</td>
</tr>
<tr>
<td>DH5α</td>
<td>$endA1\ hsdR17(r_k^+m_k^+)^{supE44}\ thi-1\ recA1\ gyrA\ (Nal^r)\ \Delta(lacZYA-argF)U169\ deoR(\Phi80dlac\Delta(lacZ)\ M15)$</td>
</tr>
</tbody>
</table>
mutant of lac repressor), LHR5 (Arg340-Glu373 from the "long axis" dimeric mutant of lac repressor), and purine repressor were amplified by PCR using GeneAmp PCR kit from Perkin Elmer, and the reaction was carried out on the GeneAmp PCR system 9600. NcoI sites were generated at the N-terminus of LHR3, LHR3+, LHR3d, and LHR5, and at the C-terminus of the purine repressor. Wild-type lacI gene in pJC1 (Chen & Matthews, 1992) was used as the template for LHR3 and LHR3+. L349A mutant of lacI in pJC1 was used as the template for LHR3d. "Long-axis" dimeric mutant of lacI in pJC1 was used as the template for LHR5. Wild-type purR gene in pPR1005-2 was used as the template to amplify purR. The templates were purified as double-stranded plasmids by phenol/chloroform extraction from the crude cell extract, followed by precipitation with ethanol.

Oligonucleotides employed for PCR are listed in Table 3. LHR3, LHR3d, LHR3+ and LHR3d+ were amplified out of pJC1. PurR was amplified out of pPR1005-2. All the oligonucleotides were synthesized on a Biosearch 8600 DNA Synthesizer and purified by Econo-Pac 10DG Columns from Bio-Rad. The reaction mixture, including 1 ng of template, 1.2 μg of each oligonucleotide, 2.5 μM of each dNTP, 2 mM of magnesium chloride, 10 μl of 10x reaction buffer (100 mM Tris-HCl, pH 8.3; 500 mM KCl; 15 mM MgCl₂; 0.01% (w/v) gelatin) and 2.5 units of Taq DNA polymerase from Perkin Elmer in 100 μl as a total volume, but without dNTPs and Taq DNA polymerase, was heated to 94°C for five minutes to improve denaturation
Table 3. List of oligos involved in this work. Oligonucleotides used in polymerase chain reactions for the amplification of coding regions are listed in this table. The sequences of the NcoI site are capitalized.
<table>
<thead>
<tr>
<th>Coding region</th>
<th>N-terminal oligo</th>
<th>C-terminal oligo</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHR3</td>
<td>5’cgcaaacgcCCATGGuccccgcgcgg3’</td>
<td>5’gacgttgtaaaactgcagccagtgaatcc3’</td>
</tr>
<tr>
<td>LHR3d</td>
<td>5’cgcaaacgcCCATGGuccccgcgcgg3’</td>
<td>5’gacgttgtaaaactgcagccagtgaatcc3’</td>
</tr>
<tr>
<td>LHR5</td>
<td>5’cgcaaacgcCCATGGuccccgcgcgg3’</td>
<td>5’gacgttgtaaaactgcagccagtgaatcc3’</td>
</tr>
<tr>
<td>LHR3+</td>
<td>5’ccaCCATGGuccccaatacgcaaacgcgg3’</td>
<td>5’gacgttgtaaaactgcagccagtgaatcc3’</td>
</tr>
<tr>
<td>LHR3d+</td>
<td>5’ccaCCATGGuccccaatacgcaaacgcgg3’</td>
<td>5’gacgttgtaaaactgcagccagtgaatcc3’</td>
</tr>
<tr>
<td>PurR</td>
<td>5’gcagggagctggagttagggtctgg3’</td>
<td>5’gcaacggCCATGGuacgaacgatagtgcgcgg3’</td>
</tr>
</tbody>
</table>
of the double-stranded plasmid templates. Twenty-five three-stepped cycles were carried out, as indicated in the following description, after the addition of dNTPs and Taq DNA polymerase. For every cycle, the templates were denatured at 94°C for one minute; annealing was carried out at 55°C for one minute; and chain elongation was carried out at 72°C for two minutes. The final PCR products were purified by Quickspin Column PCR Purification Kit form Qiagen and were checked by electrophoresis on 0.9-1.6% agarose gel depending on the size of the PCR products.

DNA Recombination

PCR fragments of LHR3, LHR3d, LHR3+, LHR3d+, LHR5, and purR were cloned into the pCRII vector from Invitrogen. Vector (50 ng), equal molar amount of insertion fragments, 3 Weiss units DNA ligase, and 1 μL of 10x reaction buffer (300 mM Tris-HCl, pH 7.8, 100 mM MgCl₂, 100 mM DTT, and 5 mM ATP) were mixed together in 10 μL. After incubating at 16°C for 4-16 hours, the mixture was heated to 65°C for 10 minutes and was transformed into E. coli 71/18 to screen for recombinants. Since it bears proper unique restriction enzyme sites, plasmid pKK332-2 was chosen to serve for the recombination of purR and each of the five different C-termini of lac repressors. The recombination was carried under similar conditions with the following alterations: 1. All six fragments of purR, LHR3, LHR3d, LHR3+, LHR3d+, and LHR5 were digested out of pCRII with proper
restriction enzymes. 2. After enzyme digestion, the desired fragments were separated from other fragments by electrophoresis on 0.9-1.6% agarose gel depending on the size of the fragments. 3. The desired bands were purified from agarose gels using DEAE membranes (Ausubel et al., 1995). The same procedure was used to clone the recombinant fragments of PL3 (purR+LHR3), PL3+ (purR+LHR3+), PL3d (purR+LHR3d), PL3d+ (purR+LHR3d+), and PL5 (purR+LHR5) into the expression vector, pT7-7.

**Expression and Purification of PurR, PL3, and PL3+**

Induction by T7 RNA polymerase was used to express wild-type PurR, PL3, and PL3+. pT7-7 containing PurR, PL3, or PL3+ was grown in *E. coli* BL21 in 4 L LB media with 0.5% maltose and 50 µg/ml ampicillin, and absorbance at 600 nm was measured periodically. Glucose was added at an absorbance of 0.3 to final concentration 0.4%, and magnesium sulfate and λ phage CE6 were added at an absorbance of 0.5 to final concentration 0.01 M and 2-4 x 10⁹ pfu, respectively. The overexpression of PurR, PL3, and PL3+ was checked by the SDS-PAGE profile of the whole cell extracts. Cells were harvested after 3 more hours growth by centrifugation and resuspended into 40 mL buffer A (10 mM Hapes, 5% glycerol, 0.2 mM DTT and 0.1 mM EDTA) (Rolfes & Zalkin, 1990a). The cells were lysed by sonication using a Branson Sonifier 450 sonicator. The supernatant from the subsequent centrifugation was diluted into buffer A at a total protein concentration of
10 mg/ml. This extract was loaded on to a 60 mL DEAE-Sephadex column which was equilibrated by buffer A. A 200 ml (total volume) gradient of KCl from 0 - 200 mM in buffer A was applied to the column after a thorough wash by buffer A. The eluate was collected by a 2122 HeliRac fraction collector. The protein concentration of each fraction was monitored by absorbance at 280 nm with the 2238 UVICORD S II. PurR, PL3 and PL3+ were all eluted at about 170 mM KCl. The fractions containing PurR, PL3, or PL3+ were examined by SDS-PAGE. The proper fractions were pooled and diluted 3-fold before they were loaded on an 80 ml phosphocellulose column, which was equilibrated by 30 mM KCl in buffer A. A 200 ml (total volume) gradient of KCl from 30 mM - 400 mM in buffer A was applied to the column after a thorough wash by 30 mM KCl. All three proteins were eluted at 300 mM - 350 mM KCl. The eluate was monitored and collected under the same conditions described previously. The quantity and purity of all three proteins were detected by SDS-PAGE.

**Oligomerization States of PurR, PL3 and PL3+**

The oligomerization states of PurR, PL3, and PL3+, and the equilibrium parameters for conversion between different states were determined by equilibrium sedimentation. Purified PurR, PL3, and PL3+ were diluted to 0.1, 0.2, 0.5 mg/ml into buffer II (100 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 250 mM potassium glutamate, 10 mM magnesium acetate
and 1 mM EDTA) (Rolfes & Zalkin, 1990a). These samples were centrifuged at 15,000 rpm for 16 hours at 20°C using an Optima XL-A Analytical Ultracentrifuge. The data were fit into a multiple data set analysis model using the program package with the analytical ultracentrifuge from Beckman. The fitting equation used by this model is:

\[
A = \exp[\ln(A_{O_1}) \cdot H \cdot M \cdot (X^2 - X_0^2) - B \cdot M \cdot (A_r - A_0)] \\
+ \exp[N_2 \cdot \ln(A_{O_1}) + \ln(KA_2) + N_2 \cdot H \cdot M \cdot (X^2 - X_0^2)] \\
+ \exp[N_3 \cdot \ln(A_{O_1}) + \ln(KA_3) + N_3 \cdot H \cdot M \cdot (X^2 - X_0^2)] \\
+ \exp[N_4 \cdot \ln(A_{O_1}) + \ln(KA_4) + N_4 \cdot H \cdot M \cdot (X^2 - X_0^2)] + E
\]

where

- \(A_r\) is the absorbance at radius \(X\)
- \(AO_1\) is the absorbance of the monomer at reference radius \(X_0\)
- \(A_0\) is the total absorbance of all species at the reference radius \(X_0\)
- \(B\) is the second virial coefficient
- \(E\) is the baseline offset
- \(H\) is the constant \((1-pV)\sigma^2/(2RT)\)
- \(KA_2\) is the association constant for the monomer-nmer equilibrium of species 2
- \(KA_3\) is the association constant for the monomer-nmer equilibrium of species 3
$K_{A_4}$ is the association constant for the monomer-nmer equilibrium of species 4

$M$ is the monomer molecular weight

$N_2$ is the stoichiometry for species 2

$N_3$ is the stoichiometry for species 3

$N_4$ is the stoichiometry for species 4

$XO$ is the reference radius

**DNA Binding Activity**

The DNA binding activities of PurR and PL3 were determined by gel retardation. The NdeI-HindIII fragment of pPR1006, which contains one operator binding site of purine repressor, was purified by electrophoresis as described above. The operator was labeled with $[\gamma-^{32}P]$ ATP by T4 DNA kinase. Operator DNA (6.25 µg) was incubated with 16 µl of $[\gamma-^{32}P]$ ATP (10 mCi/ml, $2.2 \times 10^{-4}$ M), 1 µl of T4 DNA kinase (Promega) and 2.5 µl of 10x T4 DNA kinase buffer (700 mM Tris-HCl, pH 7.6, 100 mM MgCl$_2$ and 50 mM DTT) in 25 µl of reaction mixture at 37°C for half an hour. Then T4 DNA kinase was inactivated by heating the mixture up to 65°C for 10 minutes, and the labeled operator was purified using a Nickel Column from Pharmacia. The protein-operator binding assay was carried out in buffer system II as defined previously. The DNA concentration employed was $10^{-12}$ M, and the protein concentration was varied from $5 \times 10^{-11}$ to $5 \times 10^{-7}$ M. The reaction
mixture was incubated at room temperature for two hours due to the slow binding of guanine to purine repressor. Free DNA and DNA-protein complex were resolved on a 5% polyacrylamide gel.

**Corepressor Binding Activity**

Corepressor binding activities of PurR and PL3 were determined by nitrocellulose filter binding. The reaction was carried out in buffer system II. \[^{14}\text{C}\]-Guanine was employed at concentration of $1 \times 10^{-7}$ M, and the protein concentration was varied from $1 \times 10^{-7}$ M to $5 \times 10^{-5}$ M. The reaction mixture was incubated at room temperature for two hours (Choi & Zalkin, 1992) before it was filtered through nitrocellulose membrane (Schleicher and Schuell) using a 96-well filtering apparatus. The radioactive density of each point was scanned in a Fuji Phospho-imager.
RESULTS

DNA Construction of PurR Mutants

As described in Materials and Methods, PurR was amplified by PCR, and an NcoI site was created at the C-terminus of the gene, changing the stop codon TAA to serine TCC. The coding regions corresponding to the C-termini designated as LHR3, LHR3+, LHR3d, LHR3d+, and LHR5 were also amplified by PCR with the creation of NcoI sites at positions shown in Figure 10. All PCR products were cloned into the pCRII vector using the PCR Cloning Kit from Invitrogen. In order to construct PL3, PL3+, PL3d, PL3d+, and PL5, the fragment of purR coding region with NcoI site at the C-terminus from pCRII was first cloned into vector pKK332-2 between EcoRI and NcoI sites, replacing the ABP coding region in that vector to create pKK-purR. Thereafter, the fragments of LHR3, LHR3+, LHR3d, LHR3d+, and LHR5 from pCRII were cloned into pKK-purR between NcoI and HindIII sites, so that all five different C-termini of lac repressor were placed at the C-terminus of purine repressor in the proper reading frame. The differences among five C-termini of lac repressor are shown in Figure 11. In order to overexpress these five recombinant proteins, the coding regions of PL3, PL3+, PL3d, PL3d+, and PL5 were cloned into an expression vector pT7-7 between NdeI and HindIII sites. However, only the overexpression of PL3 and PL3+ were attempted due to the negative DNA binding results for both
Figure 10. The scheme for creating purine repressor mutants with different C-termini from *lac* repressor. The border of the coding region for purine repressor and different C-termini of lactose repressor is the position of the NcoI site. The altered amino acids due to the creation of NcoI site are indicated under the border line.
purine repressor

PL5

residues 1-341 of purine repressor  
-ser-met-ala-  
residues 340-373  
of the end-to-end  
dimer of lac repressor

PL3

residues 1-341 of purine repressor  
-ser-met-ala-  
residues 340-360  
of lac repressor

PL3+

residues 1-341 of purine repressor  
-ser-met-ala-  
residues 332-360  
of lac repressor

PL3d

residues 1-341 of purine repressor  
-ser-met-ala-  
residues 340-360  
of the L349 mutant  
of lac repressor

PL3d+

residues 1-341 of purine repressor  
-ser-met-ala-  
residues 332-360  
of the L349A mutant  
of lac repressor
Figure 11. The different C-terminal leucine-heptad-repeat motifs of PL3, PL3+, PL3d, PL3d+, and PL5. The leucines in different types of leucine-heptad-repeats are in bold.
these proteins.

**Overexpression and Purification of PurR, PL3 and PL3+**

Since purine repressor at high levels is toxic to host cells (Zalkin & Doxin, 1992), induction by T7 RNA polymerase was employed to overexpress PurR, PL3, and PL3+ as described in *Materials and Methods*. The plasmids containing the structural gene of the recombinant proteins driven by the T7 promoter were transformed into the cell line BL21 which lacks endogenous T7 RNA polymerase so that there is no overexpression of the lethal product during normal cell growth. λ phage CE6, which encodes the T7 RNA polymerase, was infected into BL21 cells at mid log-phase of growth. After 3 hours growth following the phage induction, the target proteins, PurR, PL3, and PL3+, were overexpressed and accumulated in the cell extract at about 30-40% of the total protein amount (Fig. 12). About 20 g cells were obtained from a 4 L culture of each protein. The proteins were purified using DEAE cellulose and phosphocellulose column chromatography, as described in *Materials and Methods*. However, when they were eluted from the phosphocellulose column, more PL3 and PL3+ were present in the flowthrough than PurR (data not shown). About 40 mg of PurR were eluted from the column, whereas, only ~15 mg PL3 and ~10 mg of PL3+ were eluted from the column. This result could be due to the presence of monomer in the equilibrium among different oligomerization states of PL3 and PL3+. 
Figure 12. Overexpression of PurR, PL3, and PL3+. The overexpression of PurR, PL3, and PL3+ was detected by electrophoresis on 10% SDS-PAGE. Cell extract from 75 μL of cell culture was loaded in each lane. Lane 1: uninduced PL3+; Lane 2: induced PL3+; Lane 3: uninduced PL3; Lane 4: induced PL3; Lane 5: uninduced PurR; Lane 6: induced PurR.
which will be discussed in the following section. The homogeneity of both proteins is better than 85% as demonstrated by SDS-PAGE (Fig. 13).

**Oligomerization States of PurR, PL3 and PL3+**

Purine repressor was previously determined to be a dimer by cross-linking (Choi & Zalkin, 1992), but gel filtration indicated a species with a molecular weight between monomer and dimer (Rolfes & Zalkin, 1990a), apparently due to the interaction between PurR and the gel matrix, *i.e.* Sephadex (Fu Lu, personal communication). X-Ray crystallographic structural studies demonstrated that PurR is a homodimer with two identical monomers (Schumacher *et al.*, 1994a). The data obtained from equilibrium sedimentation experiments with different concentrations of PurR were not fit well by a simple exponential equation, which indicated that there was more than one species in the equilibrium system. The data were fit by a multiple exponential equation as described in *Materials and Methods*. The best fit yield indicates an equilibrium between monomer and dimer with the $K_a = 2.2 \times 10^{40}$ M, which means that dimer is the major species in the equilibrium at protein concentrations employed for assay (Fig. 14).

The oligomerization state of PL3 and PL3+ was also determined by equilibrium sedimentation. The data from different concentrations were fit by a multiple exponential curve only, and the fitting curve showed that monomer, dimer, and tetramer were all present in the equilibrium system of PL3 with
Figure 13. The homogeneity of PurR, PL3, and PL3+. The homogeneity of all three proteins after the purification by DEAE and phosphocellulose columns was checked by electrophoresis on 10% SDS-PAGE. Lane 1: PL3+; Lane 2: PL3; Lane 3: PurR.
$K_{a2} = 3.37 \text{ M}$, and $K_{a4} = 9335 \text{ M}$ (Fig. 14). These association constants indicate that the addition of LHR3 may have disrupted the monomer-monomer interface of wild-type purine repressor so that it is easier for monomer and tetramer to form. The presence of monomer was confirmed by non-denaturing polyacrylamide gel electrophoresis at protein concentration 0.1 mg/ml. Only one species with higher mobility than purine repressor was observed for PL3 (Fig. 15). The ultracentrifugation data from different concentrations of PL3+ were fit by both associating and non-associating models. The molecular weight determined from a non-associating model yielded the best fit consistent with the known monomer molecular weight. The species molecular weight was 102 kDa, which is close to that of a trimer of PurR.

**DNA Binding Activity of PurR, PL3, and PL3+**

The operator binding activity of PL3 and PL3+ was studied by gel retardation. In order to compare PL3, PL3+ and PurR, gel retardation of all three proteins was carried out in buffer system II. From Figure 16 and Table 4, it is obvious that PL3 and PL3+ have lower DNA binding affinity than wild-type purine repressor under the buffer conditions employed. Equilibrium dissociation constants ($K_d$) and Hill coefficients of all three proteins are summarized in Table 4. Operator binding affinity was determined by fitting the data to the equation, $Y = [\text{protein}]^n/(K_d^n + [\text{protein}]^n)$. 
Figure 14. Equilibrium sedimentation of PurR, PL3, and PL3+. The molecular weights of all three proteins were determined by equilibrium sedimentation in buffer system II. Samples were centrifuged at 15,000 rpm for 16 hours at 20°C. A, Data for PurR was fit well in a nondissociating model with a molecular weight of 67 kD; B, Data for PL3 was fit well in a monomer-dimer-tetramer associating model (II), but not nondissociating dimer (I) or trimer (III); C, Data for PL3+ was fit well in a non dissociating model with a molecular weight of 105 kD (III), but not dimer (I) or monomer-dimer-tetramer associating model (II).
Figure 15. Oligomerization state of PL3 detected on native acrylamide gel.

The oligomerization state of PL3 was detected on a 10% nondenaturing acrylamide gel. *Lac* repressor and wild-type PurR were used as standards. The concentration of all three proteins was 0.1 mg/ml, and the loading volume was 10 μL. Lane 1: *lac* repressor; Lane 2: PurR; Lane 3: PL3.
Figure 16. Operator binding of PurR, PL3, and PL3+. Operator binding affinities of all three proteins were determined by gel retardation in buffer system II. Guanine concentration was $1 \times 10^{-5}$ M, and the operator concentration was $1 \times 10^{-12}$ M. The percentage of the operator-repressor complex is plotted against the logarithm of the protein concentration for each protein. ◆: PurR; ○: PL3; ■: PL3+.
Table 4. Summary of operator binding activity for PurR, PL3, and PL3+.

Apparent $K_d$s and Hill coefficients deduced from the operator binding assays for all three proteins are listed in this table. $K_d$: equilibrium dissociation constant; n: Hill coefficient.
<table>
<thead>
<tr>
<th>Protein</th>
<th>$K_d$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>PurR</td>
<td>$2.0 \pm 0.2 \times 10^{-9}$</td>
<td>$1.0 \pm 0.1$</td>
</tr>
<tr>
<td>PL3</td>
<td>$1.1 \pm 0.2 \times 10^{-7}$</td>
<td>$2.0 \pm 0.4$</td>
</tr>
<tr>
<td>PL3+</td>
<td>$2.5 \pm 0.9 \times 10^{-7}$</td>
<td>$1.0 \pm 0.3$</td>
</tr>
</tbody>
</table>
As deduced from equilibrium sedimentation, monomer, which does not bind DNA, is the major species of PL3 in the protein concentration range in which DNA binding activities were determined. So the protein concentration of PL3 does not generate the species, dimer or tetramer, which are capable of binding to the operator. Since the percentage of each species has not been determined, the apparent DNA binding affinity obtained from gel retardation cannot reflect the intrinsic operator binding affinity of PL3. The sedimentation data from PL3+ were fit by a nonassociating model with a molecular weight corresponding to a trimer of PurR. DNA binding affinity may have been abolished because of the misarrangement of three DNA binding domains.

Corepressor Binding Activity of PurR, PL3, and PL3+

Wild-type purine repressor requires corepressor, either guanine or hypoxanthine, to bind to the operator at physiological conditions. The \( K_d \) for binding of guanine to wild-type purine repressor was determined by equilibrium dialysis to be \( 1.5 \times 10^{-6} \) M (Choi & Zalkin, 1992). In our experiments, guanine binding activity of PurR, PL3, and PL3+ was determined by nitrocellulose filter binding. The data (Fig. 17) showed a similar \( K_d \) of PL3 binding to guanine as found for wild-type purine repressor. However, the binding affinity of PL3+ to guanine was significantly diminished.
(Fig. 17). No binding activity was observed within the protein concentration range employed in this experiment. The $K_d$s and Hill coefficients for all three proteins are listed in Table 5. Guanine binding affinity was determined by fitting the data to the equation, $Y = [\text{protein}]^n / (K_d^n + [\text{protein}]^n)$. The trimerization of PurR may have caused a dramatic rearrangement of all three monomers. In this way, the corepressor binding sites may have been blocked so that the corepressor binding activity is abolished.
Figure 17. Guanine binding of PurR, PL3, and PL3+. Guanine binding affinity of all three proteins was determined by filter binding in buffer system II. Guanine concentration was $1 \times 10^{-7}$ M. The percentage of the retention of radioactivity on each filter is plotted against the logarithm of the protein concentration for each protein. ◆: PurR; ●: PL3; ■: PL3+. 
Table 5. Summary of guanine binding activity for PurR, PL3, and PL3+. $K_d$s and Hill coefficients deduced from guanine binding assays for all three proteins are listed in this table. $K_d$: equilibrium dissociation constant; n: Hill coefficient; N/D: not detectable.
<table>
<thead>
<tr>
<th>Protein</th>
<th>$K_d$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>PurR</td>
<td>$5.7 \pm 1.6 \times 10^6$</td>
<td>$1.0 \pm 0.1$</td>
</tr>
<tr>
<td>PL3</td>
<td>$4.9 \pm 0.4 \times 10^6$</td>
<td>$1.7 \pm 0.2$</td>
</tr>
<tr>
<td>PL3+</td>
<td>N/D</td>
<td>N/D</td>
</tr>
</tbody>
</table>
DISCUSSION

The importance of protein association in genetic regulation elicits the necessity for understanding the mechanism of protein-protein interactions. Previous studies have successfully generated dimeric and monomeric mutants of tetrameric DNA binding protein, LacI, and demonstrated that the four-helical-bundle formed by the C-terminal amino acids is responsible for dimer-dimer interactions (Chakerian et al., 1991; Alberti et al., 1991; Chen and Matthews, 1992; Alberti et al., 1993). However, the potential effect of the sequence of the region that forms this four-helical-bundle on other homologous proteins, such as PurR remains unclear. We postulated that the four-helical-bundle could cause dimeric PurR to form tetramer, if it is properly positioned.

In order to investigate this possibility, leucine-heptad-repeats with different lengths for the flanking sequence at the N-terminus, i.e. LHR3 and LHR3+, were added at the C-terminus of PurR. Different lengths for the flanking region were used in order to provide a proper distance between the CBD of PurR and the tetramerization domain, so that both domains can fold properly and form a tetramer. LHR3d and LHR3d+, which bears a single amino acid mutation L349A (Chakerian et al., 1991; Chakerian & Matthews, 1992), caused the dissociation of tetramer to dimer in the LacI protein. These sequences were also added at the C-terminus of PurR to create PL3d and
PL3d+, respectively. These two constructs were used as controls in order to exclude functional changes caused by the addition of the peptide sequences, and to separate such effects from the addition of a properly folded tetramerization domain. The elongated leucine-heptad-repeat (LHR5) sequence (Chen et al., 1994) was also added to PurR to create PL5, in order to create a more stable tetramer.

A monomer-dimer-tetramer association is determined for PL3 by equilibrium sedimentation, and PL3+ is found to be trimeric by the same method. These results indicate that the addition of LHR3 or LHR3+ has disrupted the monomer-monomer interface in PurR. The X-ray crystallographic structures of PurR and LacI have shown that the C-terminus of PurR does not align with LacI in their tertiary structures (Schumacher et al., 1994a; Lewis et al., 1996; Friedman et al., 1996) The last seven amino acids of PurR (residue 335-341) form a helix-like structure, and these helices point away from each other in the dimer (Schumacher et al., 1994a). LHR3 sequences at the C-terminus of PL3 are very likely to interact with each other to form helices or a helical bundle. This tendency may destabilize the PurR dimer because of the orientation of the C-termini away from each other so that dimer and tetramer formation may be competitive. Thus, a monomer-dimer-tetramer equilibrium is observed in equilibrium sedimentation studies. Of note is the fact that the operator binding affinity for this protein is significantly decreased. Although LHR3+ provides more flexibility between
the CBD and the assembly motif, the tendency still exists for LHR3+ to form 
helices or a helical bundle and for the C-terminus from each monomer of 
PurR to position these helices in altered orientation. This conflict appears to 
result in formation of a trimer, in which the relative position of each 
monomer is different from the native dimer. Although the arrangement of 
three monomers is not clear, the corepressor binding pocket appears to be 
blocked, and the DNA binding domain is impaired in its recognition capacity. 
Thus, the corepressor and operator binding affinities of PL3 are abolished.

The effects of LHR3 and LHR3+ on the oligomerization states of PL3 
and PL3+, respectively, and, in turn, the binding properties of these two 
proteins, are strikingly different. The addition of LHR3 destabilized PurR 
and caused a monomer-dimer-tetramer equilibrium. On the other hand, the 
addition of LHR3+, with 8 extra amino acids in the region connecting CBD 
and the assembly motif, completely disrupted the monomer-monomer 
interface, and, in turn, caused a dramatic conformational rearrangement. 
These results indicate a profound effect on protein conformation and function 
caused by relatively subtle difference in amino acid sequences. From these 
results, we conclude that design of oligomerization motifs, even using 
segments from closely related proteins, is complex, and products of such 
experiments must be analyzed carefully.
REFERENCES


