RICE UNIVERSITY

Polyelectrolyte and Hydrophobic Effects in Protein-DNA Interactions: PurR, LacI and Human p53 Proteins

by

Markos I. Moraitis

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

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

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

Susan I. Gibson, Assoc. Professor,
Biochemistry and Cell Biology

Seiichi P. T. Matsuda, Assoc. Professor,
Biochemistry and Cell Biology

Antonios G. Mikos, Professor,
Bioengineering

John S. Olson, Ralph & Dorothy Looney
Professor, Biochemistry and Cell Biology

Houston, TX
June, 2002
Abstract

Polyelectrolyte and Hydrophobic Effects in Protein-DNA Interactions: PurR, LacI and Human p53 proteins

by

Markos I. Moraitis

The polyelectrolyte and the hydrophobic effects influence affinity of protein-DNA interactions. Explorations of these phenomena form the focus of this thesis as applied to three biological systems: lactose (LacI) and purine (PurR) repressors from bacterial cells and p53 protein from human cells.

First, the presence of guanine or DNA enhanced DNA or guanine binding by PurR, respectively. Compared to the behavior of LacI, a homologous protein, PurR exhibits tighter control of the pur regulon, an effect that derives from both thermodynamic and kinetic differences. In vivo lactose utilization — controlled by LacI — is response to an environmental opportunity to which cells need to respond rapidly. In contrast, purine enzyme biosynthesis — controlled by PurR — is an energetically costly process so that cellular response must integrate signals of purine states over a period of time.

Second, the ionic and temperature dependence of DNA binding — reflecting the polyelectrolyte and hydrophobic effects, respectively — for the structurally similar
PurR and LacI proteins was examined. Both proteins exhibited a roughly \( \sim2 \)-fold increase in ion pairs formed in their high affinity form. Both proteins employ basic residues from their N-terminal helix-turn-helix DNA binding motif and core domain; however, PurR employs additional basic residues from its hinge helices and possibly interchain N-terminal to core ion pairs to account for the \( \sim15 \) ion pairs when PurR-guanine complexes DNA. The values of \( \Delta C_p \) determined by van't Hoff analysis, showed greater participation of local folding for PurR than LacI possibly due to more extensive hinge helix folding. These major differences between PurR and LacI, demonstrate that structurally similar proteins can display highly distinct behaviors.

Finally, ion concentration dependence of wild type, S392E, and a C-terminal deletion variant (\( \Delta33 \)) p53 proteins showed a remarkably low number of ion pairs formed compared to the number predicted from p53 core domain structural data. CD, temperature dependence, and water release studies of \( \Delta33 \) p53 are consistent with the notion of a structurally rigid core DNA binding domain. However, the DNA binding properties of \( \Delta33 \) p53 variant did not substantiate its characterization as “activated” form of p53 and disfavor the popular model of DNA binding regulation by the p53 C-terminal domain.
Acknowledgements

Keep Ithaka always in mind.
Arriving there is what you’ve destined for.
But don’t hurry the journey at all.
Better if it goes on for years
so you’re old by the time you reach the island,
wealthy with all you’ve gained on the way,
not expecting Ithaka to make you rich.
Ithaka gave you the marvelous journey.
Without her you wouldn’t have set out.
(C. Cavafy, Ithaca)

It was a full long journey indeed. The journey began with Drs. Constantine Demopoulos, Spyros Koinis and Andreas Tsatsas encouraging me to set out. After joining the Matthews lab, help and contribution from senior students and postdocs, Drs. Cathie Falcon, Jennifer Barry, Han Xu, Diane Wycuff, Likun Li and Liskin Swint-Kruije, was invaluable for me to begin my trip to Ithaca. For the p53 part of the journey, I was lucky to be with Emilia Mrowczinski and Dr. Nicole Nichols, to share the excitement and the frustrations of this difficult project. Then becoming senior graduate student, it was inspiring to see the enthusiasm of the newer members of the lab, Drs. Sarah Bondos and Xin-Xin Tang, Jamie Catanese, Miguellita Callabretta and Hongli Zhan.
However, I wouldn’t have reached Ithaca unless I had Dr. Kathy Matthews as my captain. Her guidance was crucial throughout my graduate career, with scientific ideas that culminated to the “hot” Biochemistry paper last year. Moreover, I am thankful to Dr. Matthews, that she helped me numerous times and most of them out of her big patient compassionate and generous heart. She helped me fight the inner deamons, the Cyclops and the angry Poseidon, inspiring me when I was down, magnifying the elation when experiments worked beautifully.

Outside the academics, Ms. Susan Norman was a special person for me that I love very much. Also, the Koemtzopoulos family have wrapped me with their love, so that I feel I have a second family here in Houston.

Above all, I dedicate this thesis to my parents and sister, who followed me along the journey with all their heart, at all times.
2 Materials and methods

2.1 Theoretical aspects ................................................. 44
  2.1.1 DNA binding ................................................. 44
  2.1.2 DNA binding dissociation kinetics .......................... 50
  2.1.3 Ion concentration dependence of DNA binding – Polyelectrolyte effect ................................................. 51
  2.1.4 Temperature and osmotic stress dependence of DNA binding – Hydrophobic effect ................................................. 54

2.2 Experimental aspects ................................................. 55
  2.2.1 Equilibrium and kinetic parameters for PurR and LacI ................................................. 55
  2.2.2 Ion concentration and temperature dependence of DNA binding in PurR and LacI proteins ................................................. 61
  2.2.3 Biochemical characterization of Δ33 C-terminal mutant of human p53 ................................................. 63

3 Thermodynamic parameters for ligand binding by PurR–Comparison to LacI ................................................. 69

3.1 Introduction ................................................. 69

3.2 Results ................................................. 70
  3.2.1 Guanine binding to PurR ................................................. 70
  3.2.2 DNA binding to PurR ................................................. 72
  3.2.3 Operator dissociation kinetics of PurR ................................................. 73
  3.2.4 Global equilibrium analysis of PurR ................................................. 75
  3.2.5 Thermodynamic and dissociation kinetic parameters of LacI DNA binding ................................................. 82

3.3 Discussion ................................................. 85
  3.3.1 Thermodynamic parameters for PurR DNA binding ................................................. 85
  3.3.2 Kinetic parameters for PurR DNA binding ................................................. 86
3.3.3 Comparison of PurR and LacI ........................................ 87

4 Ion concentration and temperature dependence of DNA binding:
  Comparison of PurR and LacI proteins .................................. 89
  4.1 Introduction ................................................................. 89
  4.2 Results .................................................................................. 90
    4.2.1 Ion concentration dependence of DNA binding of PurR .... 90
    4.2.2 Ion concentration dependence of DNA binding of LacI proteins 91
    4.2.3 Cation release by PurR and LacI proteins ......................... 95
    4.2.4 Temperature dependence of PurR and -11aa LacI proteins .... 97
  4.3 Discussion ............................................................................ 100
    4.3.1 Correlation of ion concentration dependence with structural
data for PurR ................................................................. 100
    4.3.2 Correlation of ion concentration dependence with structural
data for LacI ................................................................. 105
    4.3.3 Temperature dependence ................................................. 110

5 Biochemical characterization of Δ33 C-terminal deletion mutant of
human p53 .................................................................................. 115
  5.1 Introduction ........................................................................... 115
  5.2 Results ................................................................................. 117
    5.2.1 Secondary structure of Δ33 p53 ..................................... 117
    5.2.2 Protein assembly .............................................................. 120
    5.2.3 DNA binding of Δ33 p53 .................................................. 123
    5.2.4 Ion concentration dependence of DNA binding of Δ33 p53 .. 124
    5.2.5 Temperature dependence of DNA binding ....................... 124
    5.2.6 Water release ................................................................. 127
  5.3 Discussion ............................................................................. 132
5.3.1 Biochemical characterization of Δ33 p53 ................ 132
5.3.2 Implications for the activation model .................... 141

6 Conclusions ........................................ 145

7 Appendix ........................................... 151

7.1 Modelling of urea denaturation of tetrameric LacI ........ 151
7.1.1 Introduction .................................... 151
7.1.2 Models of urea denaturation of lactose repressor ...... 152
7.1.3 Application .................................... 153
7.2 Macroscopic and microscopic equilibrium constants ...... 155
# List of Figures

1.1 Modes of transcription regulation by LacI ........................................... 5
1.2 N-terminal headpieces of LacI complexed with DNA .............................. 8
1.3 LacI bound to DNA ........................................................................... 11
1.4 Core domain of LacI complexed with IPTG .......................................... 12
1.5 Assembly interfaces of LacI ............................................................... 13
1.6 Ion pairs involved in the allostery mechanism in LacI ............................ 16
1.7 Biology of purine repressor .............................................................. 19
1.8 Key steps in purine biosynthesis pathway .......................................... 20
1.9 Corepressor binding domain of purine repressor ................................. 22
1.10 Holorepressor PurR bound to purF and guanine ................................. 23
1.11 Summary of p53 biological function .................................................. 27
1.12 MDM2-p53 Interactions ................................................................... 31
1.13 Core domain of p53 ......................................................................... 33
1.14 Tetramerization domain of p53 .......................................................... 36
1.15 Activation model of p53 regulation ..................................................... 40

2.1 Thermodynamic cycle for guanine and DNA binding of PurR ............... 48
2.2 Polyelectrolyte effect ........................................................................ 51

3.1 Guanine binding affinities to PurR ..................................................... 71
3.2 Equilibrium DNA binding of PurR ..................................................... 72
3.3 PurR DNA dissociation kinetics ........................................ 74
3.4 Global data fitting ....................................................... 77
3.5 Species distribution ..................................................... 80
3.6 DNA binding affinity of LacI ......................................... 83
3.7 DNA binding dissociation kinetics of LacI ......................... 84

4.1 Representative DNA binding isotherms for purine repressor at various
[KCl] concentrations ...................................................... 92
4.2 Ion concentration dependence for DNA binding by PurR proteins .. 93
4.3 Ion concentration dependence for DNA binding by LacI proteins ... 94
4.4 van't Hoff plot of PurR and -11aa LacI DNA binding ............... 98
4.5 PurR helix-turn-helix contacts to DNA ............................ 101
4.6 PurR hinge helix contacts to DNA .................................. 102
4.7 PurR core domain ion pairs ........................................ 103
4.8 LacI helix-turn-helix contacts to DNA ............................. 107
4.9 LacI core domain contact to DNA .................................. 108

5.1 Thermal denaturation of Δ33 p53 as monitored by circular dichroism 118
5.2 Circular dichroism signal change as a function of temperature .... 119
5.3 Crosslinking analysis of Δ33 p53 ..................................... 121
5.4 Mobility shift of p53 proteins ........................................ 122
5.5 Ion concentration dependence of DNA binding of p53 proteins .... 125
5.6 van’t Hoff plots for specific and nonspecific binding of Δ33 p53 .... 128
5.7 Osmotic stress of DNA binding by p53 proteins ................... 130
5.8 Ion pairs between p53 central core and DNA ..................... 137

7.1 Tetrameric dissociative unfolding .................................... 154
## List of Tables

1.1 p53 modifications as a basis for the allosteric activation model .......................... 39
2.1 Conditions for globally analyzed data ........................................................................ 60
3.1 Thermodynamic parameters for guanine binding to PurR ........................................ 70
3.2 Thermodynamic parameters for operator binding to PurR ....................................... 73
3.3 DNA binding kinetics for PurR .................................................................................. 73
3.4 Thermodynamic parameters derived from global data fitting .................................. 76
3.5 Microscopic description of PurR thermodynamic cycle ............................................. 79
3.6 Summary of equilibrium and kinetic constants for LacI ............................................ 82
4.1 Number of ion pairs formed upon DNA binding ....................................................... 96
4.2 Thermodynamic parameters for DNA binding by van’t Hoff analysis ...................... 99
4.3 Protein-DNA electrostatic contacts for PurR and LacI proteins ............................... 106
5.1 DNA binding affinity of p53 proteins in 50 mM potassium phosphate buffer ........... 123
5.2 Number of ion pairs formed upon DNA binding of p53 proteins ......................... 127
5.3 Thermodynamic parameters of DNA binding of p53 proteins ................................. 129
5.4 Number of water molecules released for p53 proteins by osmotic stress .......... 130
5.5 p53-DNA electrostatic contacts ............................................................................... 136
Abbreviations

\( C_p \)  Heat Capacity (calK\(^{-1}\)mol\(^{-1}\))
\([D], [D_T]\)  Free and Total DNA Concentration
\( \text{IPTG} \)  isopropyl-1-thio-\( \beta \)-D-galactoside
\( k_a \)  Kinetic Association Constant
\( K_A \)  Equilibrium Association Constant
\( k_d \)  Kinetic Dissociation Constant
\( K_D \)  Equilibrium Dissociation Constant
\( \text{LacI} \)  Lactose repressor protein
\([P], [P_T]\)  Free and Total Protein Concentration
\( \text{PurR} \)  Purine repressor protein
\( R \)  Gas Constant, 8.31 J K\(^{-1}\)mol\(^{-1}\)
\( \text{ONPF} \)  \( \alpha \)-nitrophenyl-\( \beta \)-D-fucopyranoside
Chapter 1

Introduction

1.1 Thermodynamic regulation of protein-DNA interactions

Protein transcription factors provide cells with fine regulation of gene expression to ensure proper expression of genetic information at the right time and in the proper place. At the molecular level, this function is exerted primarily by protein-DNA and protein-protein interactions. The molecular mechanism of transcription control can be illuminated by understanding these interactions. DNA specificity — the ability to recognize a particular DNA sequence — depends on favorable ionic and hydrophobic interactions in the protein-DNA interface. After a transcription factor recognizes its site, transcription initiation depends on the appropriate tertiary assembly of the transcription machinery via protein-protein interactions. Transcription factors can both inhibit and facilitate the process of RNA synthesis. Protein-protein and protein-DNA complex formation are governed by thermodynamic principles. This thesis utilizes thermodynamic methods to examine protein-DNA binding, including the ionic and hydrophobic interactions that influence gene expression. These phenomena were studied in two different biological systems: lactose and purine repressor from the simple bacterium Escherichia coli and p53 tumor suppressor from the more
complex human cells.

1.1.1 Polyelectrolyte effect

Protein residues involved in electrostatic interactions with DNA are important for complex stability and protein-DNA recognition. For instance, for lactose repressor, mutations at R22 in the helix recognition motif eliminate binding to the native operator sequence but elicit binding to altered DNA sequences (Lehming et al., 1988). In purine repressor, mutations at K55, which contacts the minor groove of DNA, diminish DNA specificity (Glasfeld et al., 1999). Finally, mutations at R273 of p53 eliminate the transcription activation functions of p53 (Cho et al., 1994). In all three examples, favorable ionic interactions are formed between the positively charged protein side chains and the negatively charged phosphates of the DNA backbone. Such interactions are accompanied by the release of counterions from the protein and DNA surfaces. As the protein contacts DNA, counter anions and cations are released from the protein and DNA, respectively; a phenomenon termed the “polyelectrolyte effect” (reviewed in Record et al., 1991). This phenomenon results in an increase in entropy, which is a favorable drive for protein-DNA complex formation.

1.1.2 Hydrophobic effect

The term “hydrophobic effect” was first introduced to explain the driving force in protein folding (Privalov and Gill, 1988). As a protein folds, water molecules bound to the unfolded protein structure are released to the environment. This high entropically favored process overwhelms the entropic cost of adopting a minimal folded conformation among the vast number of possible conformations. Additionally, release of highly ordered clathrate-structured water molecules contacting exposed nonpolar surfaces in the unfolded protein results in a negative change in heat capacity ($\Delta C_p$) between the unfolded and folded state (Spolar et al., 1989). Most protein-DNA inter-
actions also display negative heat capacity changes, a phenomenon explained by the hydrophobic effect generated by simultaneous local protein folding and DNA binding (Spolar and Record, 1994).

1.1.3 Role of protein assembly for transcription factors

Protein transcription factor assembly provides crucial advantages for regulation of protein-DNA interactions (reviewed in Klemm et al., 1998). First, allostery — dependence of the affinity for a particular ligand upon binding of another ligand — is feasible by transmitting the allosteric signal through the oligomeric interface. Also, formation of protein complexes with multiple DNA binding sites allows another level of regulation, by involvement of distant cis DNA elements through DNA "looping" (reviewed in Matthews, 1992). Finally, oligomerization of transcription factors is an economical way for increasing the surface in the protein-DNA interface for formation of favorable protein-DNA interactions (Klemm et al., 1998).

1.2 Lactose repressor (LacI)

Sustenance of life requires cells to adapt depending on the availability of energy sources. One example of this principle is the regulation of the lactose operon in *E. coli* (reviewed in Matthews and Nichols, 1998). LacI is the key regulatory element for enzymes that enable bacteria to utilize lactose as an alternative carbon source when available (Miller and Reznikoff, 1970). LacI is the first historical example of the role of protein transcription factors in gene regulation, initially predicted by the seminal work of Jacob and Monod (1961).
1.2.1 Biological function of LacI

LacI was identified as a tetrameric protein of molecular weight of 150,000 kD (Gilbert and Müller-Hill, 1966; Riggs and Bourgeois, 1968) that represses expression of the three “structural” genes of the lac operon. The lac operon is composed of the following genes involved in lactose utilization by E. coli: β-galactosidase, lactose permease, and thiogalactoside transacetylase (Miller and Reznikoff, 1970). Upstream of these genes, there are two partially overlapping DNA sequences that regulate downstream gene expression: the promoter sequence (plac) and the operator sequence (O1) where RNA polymerase and LacI associate with DNA, respectively (Gilbert and Maxam, 1973; Majors, 1975). In the absence of lactose, the tetrameric LacI competes effectively with RNA polymerase for binding to its operator sequence (Schlax et al., 1995) and thus prevents RNA polymerase from initiating transcription (Figure 1.1). Indeed, LacI binds O1 with an affinity in the order of 10^{-13} M (Riggs et al., 1970b), and only ~10% of the repressor protein in a bacterial cell is free (Kao-Huang et al., 1977). The tight affinity for operator DNA is the underlying thermodynamic parameter connecting to biological effect, since the degree of repression depends strongly on the binding affinity for the operator (Oehler et al., 1994).

In addition to O1, LacI recognizes two additional operator sequences (O2 and O3) that are found down- and upstream of O1, respectively (Gilbert et al., 1974; Pfahl et al., 1979; Reznikoff et al., 1974). Although binding affinities of LacI to these “pseudo-operators” are 6 to 1000-fold weaker compared to O1 (Fried and Crothers, 1981; Winter and von Hippel, 1981), their presence is required for optimal repression in vivo (Eismann et al., 1987; Lehming et al., 1990; Mossing and Record, 1986; Oehler et al., 1994). DNA looping involves the “pseudo-operator” sequences in regulating lac operon expression (Eismann and Müller-Hill, 1990; Krämer et al., 1988). Formation of such DNA loops for LacI has been observed in vitro (Borowiec et al., 1987; Brenowitz
Figure 1.1: Modes of transcription regulation by LacI

In the absence of lactose (top panel), LacI (shown in green) binds at the operator sequence. As a result RNA polymerase cannot initiate transcription (negative regulation). The presence of allolactose (marked blue in bottom panel) relieves repression. In addition, in the absence of glucose as a carbon source, CAP protein binds to an upstream site and facilitates binding of RNA polymerase to the promoter sequence (positive regulation).
et al., 1991; Krämer et al., 1988; Whitson et al., 1987). The ability of LacI to form DNA loops relies on the oligomeric state of LacI; tetrameric lactose repressor contains two DNA binding sites (Culard and Maurizot, 1981; O'Gorman et al., 1980a) and exhibits higher repression efficiency than dimeric protein due to DNA looping (Oehler et al., 1994). In fact, dimeric LacI supplied with a basic C-terminal tail that can bind DNA non-specifically can form DNA loops and display more efficient repression (Kolkhof et al., 1995).

Binding of LacI to inducer molecules such as allolactose — the naturally occurring inducer (Jobe and Bourgeois, 1972) — or the commonly used gratuitous inducer, IPTG — isopropyl-1-thio-β-D-galactoside — enhances the dissociation rate of LacI from the DNA by at least 10³-fold (Barkley et al., 1975), thus diminishing the DNA binding affinity to non-specific levels (Barkley et al., 1975; Riggs and Bourgeois, 1968). As a result, RNA polymerase can initiate transcription of the lactose metabolic genes. This induction effect was ascribed to allosteric regulation of DNA binding by the binding of inducers. Moreover, binding of operator DNA prior to inducer reduces inducer affinity by 10-fold and makes the latter process cooperative (O'Gorman et al., 1980b). Employing this mechanism, the cell ensures the release of repression after a certain level of inducer is present.

In addition to this "negative" regulation mode, the lac operon genes are "positively" regulated by the function of catabolite activator protein or CAP (Miller and Reznikoff, 1970). The "positive" regulation is mediated by cAMP binding to CAP that increases the DNA binding affinity of the latter for promoter segments adjacent to O¹ which in turn facilitates transcription initiation by RNA polymerase (Fried and Crothers, 1984; Kolb et al., 1993). This effect occurs when glucose is absent; presence of glucose inhibits cAMP cyclase function and inducer uptake by the cells (Kolb et al., 1993). Cooperation between both modes of regulations, ensures that lactose utilization enzymes are produced in the presence of lactose but only when
the more common glucose is not readily available.

1.2.2 Functional domains of LacI

While LacI was purified and identified as the repressor agent for the lac operon over 35 years ago, only in the past few years have the X-ray structures of LacI been solved, and only very recently a high resolution structure resolved the side chains of the N-terminus. Nonetheless, earlier biochemical studies had identified two functional domains in lactose repressor: (a) a flexible N-terminal domain responsible for the DNA binding properties and (b) a larger proteolysis-resistant core domain that conferred the inducer binding and oligomerization properties to LacI.

N-terminus (amino acids 1-60)

Two experimental directions demonstrated that the first 60 N-terminal amino acids formed the DNA binding domain of LacI. Phenotypic studies showed that mutations in the N-terminus were more probable to eliminate DNA binding (Gordon et al., 1988; Markiewicz et al., 1994; Miller, 1979, 1984; Miller and Schmeissner, 1979). In the other direction, biochemical data with limited trypsin digestion located the DNA binding properties in this region, which contributes to both specific and non-specific DNA binding (Files and Weber, 1976; Geisler and Weber, 1977, 1978; Jovin et al., 1977; Kania and Brown, 1976; Ogata and Gilbert, 1978). DNA binding of this region was attributed to the helix-turn-helix motif (amino acids 1-50), which is common in DNA binding prokaryotic repressors (Brennan and Matthews, 1989).

The observed proteolytic susceptibility of the N-terminal domain was attributed to the presence of a hinge formed by amino acids 50-60 (Geisler and Weber, 1978). Indeed, the X-ray structure for the LacI-DNA complex demonstrated the presence of hinge helices that connect the N-terminal domain to the larger core domain (Lewis et al., 1996). NMR studies of free N-terminal headpieces alone and complexed
**Figure 1.2: N-terminal headpieces of LacI complexed with DNA**

The NMR structure of the N-terminal 1–62 headpiece of LacI complexed with $O^{\text{sym}}$ was derived from the PDB file 1CJG (Spronk *et al.*, 1999b). DNA is sketched in a ladder model and the two strands are colored in green and cyan. The two headpieces are sketched in a worm model based on the $C_\alpha$ atoms. The helix-turn-helix domains are colored in orange and mauve, whereas the hinge helices are colored in yellow.

with operator DNA, indicated that formation of the hinge helices is DNA-dependent, since they are not observed in the free form (Chuprina *et al.*, 1993; Spronk *et al.*, 1999a,b, 1996). These helices are involved in two types of interactions: (a) with the DNA minor groove by inserting leucine side chains to induce a bend in the DNA, and (b) with each other by hydrophobic association (Figure 1.2). Both of these interactions are important in the second step of a two-step DNA recognition by lactose repressor (Kalodimos *et al.*, 2002). In the first step, the flexible helix-turn-helix motif interacts with the major groove of the DNA (Slijper *et al.*, 1997, 1996). In the second step, the inter-
actions between the hinge helices and the minor groove of DNA determine the DNA selectivity (Falcon and Matthews, 2000, 2001; Kalodimos et al., 2001). The interactions between the hinge helices are important for transferring the allosteric response (Falcon and Matthews, 1999, 2001; Kalodimos et al., 2001).

Many factors affect the stability of LacI-DNA interaction. First, ionic strength affects adversely the DNA binding affinity of LacI, with a more pronounced effect on the nonspecific DNA binding affinity, indicating the greater ionic content of protein-nonspecific DNA interactions (Barkley et al., 1981; deHaseth et al., 1977b; Record et al., 1977; Whitson et al., 1986). LacI exhibits high DNA specificity with a $K_D$ of $\sim 10^{-11}$ M for a 40 bp sequence to $\sim 10^{-13}$ M for longer DNA sequences containing the operator (Barkley et al., 1975; Kao-Huang et al., 1977; Riggs et al., 1970a; von Hippel et al., 1974; Whitson et al., 1987; Whitson and Matthews, 1986). These values contrast to a $K_D$ of only $\sim 10^{-7}$ M for nonspecific sequences (Lin and Riggs, 1972; von Hippel et al., 1974). Symmetry in the DNA operator sequence affects DNA binding, since affinity of LacI for the synthetic symmetric operator $O_{sym}$ promoter (symmetric to promoter proximal half-site and missing the central basepair) is tighter than that to $O^1$ (Sadler et al., 1983), which is the primary operator in vivo. Additionally, the ability to form looped structures further stabilizes protein-DNA complex formation, primarily by reducing the dissociation rate of the complex (Eismann and Müller-Hill, 1990; Fried and Crothers, 1981; Hsieh et al., 1987; Whitson et al., 1987; Whitson and Matthews, 1986). Finally, the presence of inducer IPTG decreases DNA binding affinity of LacI to the level for nonspecific binding (Barkley et al., 1975; Butler et al., 1977; Revzin and von Hippel, 1977; Riggs and Bourgeois, 1968).

**Core Domain (amino acids 61–360)**

A major functional role of the core domain of LacI is to provide the essential elements for protein assembly into tetramers. The solution of the tetrameric X-ray
structures of the inducer-core domain and the DNA-LacI complexes confirmed that LacI assembled as a dimer of dimers (Friedman et al., 1995; Lewis et al., 1996). Each monomer of the dimeric subunit contributes an extensive dimer-dimer interface that positions two N-terminal DNA binding domains so that they each bind to a half-site of an operator sequence (Figure 1.3). Each dimeric subunit, also contributes to a dimer-dimer interface, coming together at the C-terminus to form a tetrameric coiled coil (Figure 1.4). Therefore, the intact LacI assembles into a V-shape tetramer that can bind two individual DNA operator sites. The DNA binding properties of various mutants have been employed elegantly, before the X-ray structures became available, to functionally distinguish the two assembly interfaces: the dimer-dimer and monomer-monomer interface (Figure 1.5).

The major energetic contribution to the dimer-dimer interface arises from the C-terminal region of the core domain. Amino acids 340–360 were identified as a leucine heptad repeat, an oligomerization motif found in eukaryotic transcription factors, such as fos and jun (Alberti et al., 1991; Chakerian et al., 1991). In LacI, this region forms a tetrameric antiparallel bundle which is stabilized mostly by the proper positioning of large apolar amino acids that form hydrophobic bonds among the helices (Alberti et al., 1993). In addition, formation of electrostatic interactions outside the bundle contributes to the stability of the helical bundle (Alberti et al., 1993). The assembly of the intact LacI tetramers is very stable; although estimations of the tetramer dissociation constants vary between $10^{-8}$ and $10^{-18}$, the value is sufficiently strong enough to be undetectable by conventional methods (Barry and Matthews, 1999b; Brenowitz et al., 1991; Levadoski et al., 1996; Royer et al., 1990, 1986). Mutation of crucial leucine residues at positions 342, 349 and 356, or deletion of this region, results in primarily dimeric species (Alberti et al., 1991; Chakerian and Matthews, 1991; Chen and Matthews, 1992; Chen et al., 1994).

The DNA binding properties of the -11aa LacI mutant reflect the moderate
Figure 1.3: LacI bound to DNA

Ribbons diagram is based on the X-ray crystal structure of the C-terminal deletion mutant 1-333 complexed with $O^{sym}$ and anti-inducer ONPF (Bell and Lewis, 2000). The protein lacks the tetramerization domain and therefore it is dimeric. One monomer is colored in red, whereas the other is colored according to its functional domains: helix-turn-helix, hinge helix and the core domains are colored in orange, yellow and gray, respectively. The anti-inducer sugar $\alpha$-nitrophenyl-$\beta$-D-fucopyranoside (ONPF) is shown in green.
Figure 1.4: Core domain of LacI complexed with IPTG

The tetrameric core domain is displayed across the dimer-dimer interface, and each dimeric unit is colored either green or blue. Complexed IPTG is displayed in orange. Ribbons diagram was generated from PDB file 1LBI (Friedman et al., 1995).
Figure 1.5: Assembly interfaces of LacI

Mutational analysis of LacI suggested that tetrameric LacI assembled as a dimer of dimers and formed two distinct oligomerization interfaces: the dimer-dimer and the monomer-monomer interface (perpendicular and parallel to the plane of print, respectively). For simplicity, only the core and the C-terminal tetramerization domains (spheres and cylinders, respectively) are displayed. The dimer-dimer interface is eliminated with mutations in the oligomerization domain (for instance deletion of the 11 C-terminal amino acids to form the short-axis dimer -11aa LacI). Specific mutations, such as Y282D, in the monomer-monomer interface destabilize both interfaces, resulting in monomeric protein. Stabilization in either interface (for instance, K84A/L or extention of the tetramerization domain) rescue the Y282D mutant, resulting in wild-type behavior or formation of a "long-axis" dimer. Detailed discussion about the mutational analysis can be found in text. The schematic was modified from Matthews and Nichols (1998).
stability of the monomer-monomer interface. Although the "apparent" DNA affinity of -11aa LacI is reduced relative to that of intact LacI, the intrinsic DNA affinity of -11aa LacI is only slightly reduced, suggesting that dimer dissociation into monomers accounts for the reduction in "apparent" DNA affinity (Chen and Matthews, 1994; Chen et al., 1994). In addition, introduction of the more stable dimeric GCN4 repeat (Chen and Matthews, 1994), or stabilization of the monomer-monomer interface by a second site mutation(K84A/-11aa LacI), returns the DNA binding affinity of the dimeric mutants to the levels of wild type (Gerke et al., 2000; Nichols and Matthews, 1997). On the other hand, destabilization of the monomer-monomer interface by mutations of residues, such as Y282 or L251A, that reside on the monomer-monomer interface, result in monomeric LacI (Daly and Matthews, 1986; Dong et al., 1999). In conjunction to the -11aa LacI, this observation suggests the requirement of dimer assembly for tetramers to form. Also, strengthening of the dimer-dimer interface by introduction of the LHR5 motif, resulted in "long axis" dimeric protein (in contrast to the "short axis" dimers of the monomer-monomer interface), characterized by the elongated shape of this protein that is incapable of binding DNA (Chen et al., 1994). The deleterious effect of Y282D can be reversed by combinatorial mutations at the monomer-monomer interface, such as K84A (Nichols and Matthews, 1997; Swint-Kruse et al., 2001).

The second role of the core domain is binding of sugar molecules. Most of the sugars recognized by LacI are galactoside derivatives whose hydroxyls at C3 and C6 confer most of the binding activity (Barkley et al., 1975; Chakerian et al., 1987). Most bind with at least millimolar affinity (Barkley et al., 1975; Butler et al., 1977). Two of those sugars will be used in this thesis, the inducer IPTG and the anti-inducer ONPF. Binding of IPTG diminishes DNA binding affinity of LacI (Barkley et al., 1975; Riggs and Bourgeois, 1968; Riggs et al., 1970a), whereas ONPF has a reverse effect, with slight increase of the LacI DNA binding affinity (Riggs et al., 1968, 1970b).
Inspired by the sequence homology of LacI to periplasmic binding proteins (Müller-Hill, 1983), Nichols et al. (1993) constructed a molecular model for the interpretation of the available biochemical data. Ultimately, the solution of the X-ray structure of the core domain complexed to inducer IPTG provided a detailed picture of the sugar binding interface (Friedman et al., 1995; Lewis et al., 1996). More specifically, the core domain has a cleft between the N-terminal and C-terminal subdomains, where the inducer molecule binds (Figures 1.3 and 1.4). IPTG can bind in two orientations by forming 3 to 4 hydrogen bonds (Lewis et al., 1996).

1.2.3 Allosteric regulation in LacI

Allosteric in LacI has been previously modelled according to Monod-Wyman-Changeux and Koshland models, where a relaxed (R) and a tense conformation (T) bind operator inducer and operator, respectively (Koshland et al., 1966; Monod et al., 1965; O’Gorman et al., 1980b). The availability of the X-ray structures for LacI proteins has provided insight on the allosteric mechanism (Bell and Lewis, 2000; Friedman et al., 1995; Lewis et al., 1996). Although there is not significant change in the area of interaction in the monomer-monomer interface between the IPTG and DNA-bound forms, different sets of interactions are present. The most striking difference lies in the content of the ionic interactions. In the IPTG-bound form, two salt bridges form across the monomer-monomer interface: K84-G100' and H74-R278' (Figure 1.6). However, formation of the latter ion pair is not essential for the allosteric transition (Barry and Matthews, 1999a). Most of the conformational change occurs within the N-subdomain of the core domain, whereas the C-subdomains remain fairly constant. In the inducer-free form, there is a 6° twist of the N-terminal subdomain, and the aforementioned salt bridges are disrupted. The allosteric effect on the N-terminal DNA binding region requires the hinge helices; in the IPTG-bound form, residues 62 within the dimer are moved apart by 3.5 Å (Lewis et al., 1996). As
**Figure 1.6: Ion pairs involved in the allosteric mechanism in LacI**

Residues that are altered by the allosteric process through formation of interchain electrostatic interactions are modelled as ball-and-stick. In the presence of inducer (top panel), LacI forms ion pairs K84-G100 (N-terminal core subdomain) and H74-R278 (C-terminal core subdomain). When bound to DNA, those ion pairs are not present (bottom panel).
a result, the hinge helices are not properly positioned in the IPTG-bound form to interact with each other and with the minor groove of DNA, preventing high affinity DNA association.

1.3 Purine repressor (PurR)

Propagation and translation of genetic information is based on assembly of the intracellular free nucleotides into DNA and mRNA. Biosynthesis of polynucleotides depends on the de novo synthesis of purines and pyrimidines (Nygaard, 1993; Zalkin and Nygaard, 1996). Expression of biosynthetic genes involved in this pathway is regulated by purine repressor (PurR). Although PurR has not been investigated in as much detail as its homologue, LacI, structural studies have provided significant molecular insight in the allosteric mechanism of gene expression in this system (Schumacher et al., 1995, 1994). Indeed, PurR became a model system that aided the solution of the X-ray structure of LacI (Lewis et al., 1996).

1.3.1 Biological function of PurR

Regulation of de novo purine biosynthesis

Cloning of E. coli PurR showed that the protein consisted of 341 amino acids and exhibited high sequence homology with members of the LacI protein family (~33 and 26% identity with GalR and LacI, respectively) (Rolfes and Zalkin, 1988a; Schumacher et al., 1993; Weickert and Adhya, 1992). Its first biologically characterized function was the regulation of purF, which is the enzyme for the first step in the de novo biosynthesis of purines (Rolfes and Zalkin, 1988b). In the presence of free purines, PurR inhibits expression of biosynthetic enzymes of the pur regulon, such as purF. When purines are freely available, the complex of purine repressor with guanine or hypoxanthine (Choi et al., 1994; Choi and Zalkin, 1992) represses transcription by
binding tightly to its DNA operator sequence (Figure 1.7), as a result of the increased affinity of purine repressor for its target sequence (Rolfes and Zalkin, 1990b). However, when intracellular purines are depleted, the purine aporepressor dissociates from its target DNA sequence as the DNA binding affinity of the aporepressor diminishes. As a consequence, RNA polymerase can transcribe purine biosynthetic enzymes, such as PurF.

Cellular regulation of the pur regulon appears to be more complex than the lactose operon. PurR was found to regulate expression of its own protein and most of the enzymes participating in de novo purine biosynthesis (Aiba and Mizobuchi, 1989; He et al., 1992; He and Zalkin, 1992, 1994; Kilstrup et al., 1989; Meng et al., 1990; Rolfes and Zalkin, 1990a; Tiedeman et al., 1990; Watanabe et al., 1989). This pathway starts with 5-phosphoribosyl-1-pyrophosphate and forms inosine monophosphate (Figure 1.8), which is the precursor molecule for the subsequent synthesis of AMP and GMP (Nygaard, 1993). Adding to the complexity, PurR regulates genes for pyrimidine biosynthesis (He et al., 1993; Kilstrup et al., 1989; Lu et al., 1995; Rappu et al., 1999; Shin et al., 1997; Steiert et al., 1990; Vial et al., 1993; Weng et al., 1995; Wilson and Turnbough, 1990).

1.3.2 Functional domains of PurR

PurR comprises of two separate domains, similarly to LacI: a 59 amino acid N-terminal domain that binds DNA and a larger 289 amino acid core domain responsible for corepressor binding.

N-terminal DNA binding domain (amino acids 1–59)

Earlier limited proteolysis experiments demonstrated the presence of a flexible N-terminal domain in PurR responsible for DNA binding (Choi and Zalkin, 1992). The N-terminal domain forms a separate globular domain that includes, similarly to
Figure 1.7: Biology of purine repressor

In the presence of hypoxanthine or guanine (blue circles), PurR (green) inhibits RNA polymerase from transcribing purine biosynthesis genes (top panel). When purines are not present, PurR dissociates from the operator DNA, allowing downstream gene transcription.
Figure 1.8: Key steps in purine biosynthesis pathway

Synthesis of GMP or AMP from ribose-5-phosphate involves a number of biosynthetic steps. Genes regulated by PurR are denoted in italics.

LacI, a helix-turn-helix DNA binding motif (Schumacher et al., 1993). This motif is the one that shows the highest homology among prokaryotic DNA binding repressor proteins (Brennan and Matthews, 1989; Weickert and Adhya, 1992). Its particular function in PurR is to recognize specific operator sequences, called the "PUR boxes", which share the following 12-bp sequence — GCAaCGtTTNC — as their consensus sequence (Mironov and Gelfand, 1999). Interestingly, in the absence of DNA, the NMR structure of this domain shows that it folds similarly to the structure found in the holorepressor X-ray structure (Nagadoi et al., 1995), with the exception of folding in the hinge region.

The N-terminal domain connects to the corepressor binding domain through a hinge helix consisting of amino acids 49–57 (Choi and Zalkin, 1994). The X-ray structure indicated van der Waals interactions of V50 and L54 to their corresponding partners in the hinge helix of the other monomer (Arvidson et al., 1998; Schumacher et al., 1994). Side chains of the hinge helix contact DNA bases or phosphates via van
der Waals (L54) or hydrogen bonding contacts (K55) in the minor groove in a manner
that results in kinking the DNA by 45° in a direction away from the protein. The
hinge helix was proposed to contribute DNA binding specificity, since mutations at
K55 diminish specific DNA binding (Glasfeld et al., 1999). Folding of the hinge helix
is DNA-dependent, since this region appears to be unordered in the NMR structure
of the isolated N-terminal DNA binding domain (Nagadoi et al., 1995).

Corepressor binding domain (amino acids 60–341)

High resolution X-ray structures of the apo- and holorepressor (Figures 1.9
and 1.10, respectively) provide molecular details of the organization of the core do-
main in the presence and absence of corepressor, respectively (Schumacher et al.,
1995, 1994, 1997). The core domain contains an N-terminal and a C-terminal sub-
domain, enclosing the corepressor binding site in between (Schumacher et al., 1995,
1994, 1997). Local conformational changes occur upon ligand binding: W147, which
occupies the binding pocket in the corepressor free form, moves to the exposed surface
where it forms a small hydrophobic patch with residues Y73 and F74, presumably
a first binding site for guiding the corepressor into the binding pocket (Schumacher
et al., 1995). Site-specificity for corepressor is conferred either directly or through sol-
vent interaction by key residues on the binding surface pocket: for instance, mutation
at R190 eliminates ligand specificity (Lu et al., 1998a). Moreover, the presence of a
water molecule in the binding pocket, interacting through hydrogen bonding with the
side chain of E222, can account for the differences between the binding affinities of
guanine and hypoxanthine to PurR (Schumacher et al., 1997).

1.3.3 Allosteric regulation of PurR

Effective DNA binding by PurR and subsequent inhibition of gene expres-
sion requires the presence of hypoxanthine or guanine as corepressors (Rolfes and
Figure 1.9: Corepressor binding domain of purine repressor

The X-ray structure of the core domain of PurR (amino acids 60–341) is derived from PDB file 1DBQ (Schumacher et al., 1995). The two monomers comprise the dimeric aporepressor. The N-terminal and C-terminal subdomain of each monomer, enclosing the purine binding cleft, are denoted with N and C, respectively. The structure is colored according to the secondary structure, with α-helices and β-sheets depicted in red and cyan, respectively.
Figure 1.10: Holorepressor PurR bound to \textit{purF} and guanine

The X-ray structure of the intact PurR bound to guanine and \textit{purF} operator DNA is derived from PDB file 1WET (Schumacher \textit{et al.}, 1997). DNA is modelled as ball and stick. One of the PurR monomers is colored in red. The other PurR monomer is colored according to its functional domains: Starting from the helix-turn-helix motif (orange) the protein crosses over to the opposite geometric side via the hinge-helix (yellow) with the core domain (cyan) binding to guanine (green space-filled).
Zalkin, 1990b). A global conformational change occurs in the N-terminal subdomain with corepressor binding; with ligand dissociation from the holorepressor, there is a rotational change of the N-terminal corepressor subdomain from 17 to 23° from the holorepressor to the aporepressor form (Schumacher et al., 1995), larger than the conformational change that occurs in LacI (Lewis et al., 1996). Although the DNA binding domain of the aporepressor structure is deleted, the transition moves the C-terminal end of the hinge helices about 3 Å further away, relative to the holorepressor form, presumably resulting in dissociation of the two hinge helices from the DNA (Schumacher et al., 1995). As the two hinge helices are pulled apart, they appear to unfold since this region is not structured in the NMR structure of the isolated monomeric N-terminal domain and is not detected in the aporepressor structure (Nagadoi et al., 1995).

1.4 p53 Tumor Suppressor Protein

Accurate transfer of genetic information requires molecular mechanisms that maintain the integrity of the genome. A major component of these mechanisms in higher eukaryotes involves the biological function of p53. p53 was named for the size of the protein that was detected to bind to the simian virus 40 T antigen and adenovirus E1B oncoproteins (Lane and Crawford, 1979; Linzer and Levine, 1979; Sarnow et al., 1982). Its similar classification as an oncogene was initially suggested by studies in the mid-80s (Eliyahu et al., 1985, 1984; Jenkins et al., 1984; Parada et al., 1984). Later studies attributed the oncogenic properties to the first isolated mutant form of p53 (Finlay et al., 1989; Hinds et al., 1989). Further findings, such as suppression of cell transformation in p53-deficient tumor cells, increased susceptibility of tumor formation in p53-deficient transgenic mice and Li-Fraumeni patients (germ line p53 mutations), and the occurrence of p53 mutations in ~50% of human cancers,
characterized p53 as a tumor suppressor (Baker et al., 1990; Donehower et al., 1992; Hollstein et al., 1991; Malkin et al., 1990). The intense interest in this molecule of the year for 1993 (Koshland, 1993) conferred its title of “cellular gatekeeper” or “guardian of the genome” based on its response to stress and DNA damage and induction of cell cycle arrest and apoptosis (Lane, 1992; Levine, 1997).

1.4.1 Biological function of p53

Response to genotoxic stress

UV, ionizing radiation, hypoxia and nucleotide depletion elicit an increase in cellular p53 protein levels, whereas in normal cells p53 is maintained at low levels (reviewed in Prives and Hall, 1999). Although such a response could arise at the transcriptional or translational level (Fu et al., 1996; Mosner et al., 1995; Sun et al., 1995), the rapid raise of p53 levels in the nucleus suggested that this response is mediated predominantly at the post-translational stage (Kastan et al., 1991; Midgley and Lane, 1997). In this way, cells would be able to respond promptly to genotoxic stresses (summarized in Figure 1.11).

The main mechanism of regulating the levels of p53 is the murine double minute protein, or MDM2 (Lane and Hall, 1997). This protein downregulates intracellular p53 levels. The importance of MDM2 in maintaining control of p53 levels was illustrated by the embryonic lethality of MDM2-deficient mice and their rescue by deleting p53 (Jones et al., 1995; Montes de Oca Luna et al., 1995). MDM2 was first characterized by its ability to bind p53 and inhibit the transactivation activity of p53 (Momand et al., 1992). Conversely, p53 activates MDM2 expression under stress conditions to form a negative autoregulatory loop (Mendrysa and Perry, 2000; Wu et al., 1993). MDM2 is a RING finger-dependent E3 ubiquitin ligase that ubiquitinates both p53 and itself (Fang et al., 2000; Honda et al., 1997; Honda and Yasuda,
2000). Moreover, MDM2 promotes p53 degradation through the proteasome pathway (Haupt et al., 1997; Kubbattat et al., 1997). Proteasome degradation occurs in the cytoplasm, and nucleolar p53 is exported to the cytoplasm either independently or with the assistance of MDM2 (Boyd et al., 2000; Freedman and Levine, 1998; Geyer et al., 2000; Roth et al., 1998; Stommel et al., 1999; Tao and Levine, 1999). Recent studies, however, have demonstrated that nuclear export is not required for ubiquitination and subsequent degradation of p53 (Xirodimas et al., 2001; Yu et al., 2000). To add to the complexity, MDMX, a homologous protein to MDM2, counteracts MDM2-dependent degradation, thus stabilizing intracellular p53 (Jackson and Berberich, 2000; Stad et al., 2001).

Cellular stabilization of p53 by down-regulation of MDM2-dependent degradation appears to be mediated by post-translational modifications of p53 and/or MDM2 (reviewed in Appella and Anderson, 2000). The major post-translational modifications of p53 are located in the N and C-terminal ends of the protein and involve mainly serine side chain phosphorylations and lysine residue acetylations. This mechanism allows p53 to integrate the genotoxic responses communicated by upstream kinase and acetylase/deacetylase enzymes.

Cell cycle arrest

Cell cycle arrest provides the necessary time for DNA to be repaired before the cell cycle continues (Kastan et al., 1991). DNA damage causes p53-dependent expression of candidate genes that are involved in G1/S arrest. Candidate genes include p21 and gadd45 (el Deiry et al., 1993; Kastan et al., 1992). p21 has a documented role in regulating G1/S progression by inhibiting phosphorylation of the retinoblastoma protein (RB) (Harper et al., 1993). For progression to S phase to occur, RB is phosphorylated and subsequently dissociates from E1B factor, which in turn activates a cascade of cycle progressing genes. RB also stabilizes p53 by abrogating
Figure 1.11: Summary of p53 biological function

The biological function in normal cells is indicated in normal type. MDM2 ensures that p53 is maintained at low levels by a ubiquitin-dependent proteolysis pathway. For simplicity, degradation is shown to occur in the cytoplasm, although nuclear export is not required and intranuclear degradation can occur. The biological function after genotoxic stress is indicated in bold type. p53 and MDM2 dissociate and p53 accumulates. p53 mediates pathways for cell cycle arrest (e.g. activation of p21) and/or apoptosis (e.g. activation of BAX). Moreover, p53 activates MDM2 expression, to ensure that levels of p53 decrease, after genotoxic stress subsides.
MDM2 binding (Hsieh et al., 1999; Yap et al., 1999). An alternative function of p21 is binding to the proliferating-cell nuclear antigen (PCNA) which inhibits elongation during DNA replication (Waga et al., 1994). A similar function involving PCNA has been proposed for gadd45 (Smith et al., 1994).

An increasing amount of data suggests that of p53 also induces cell cycle arrest at G2/M. p21 was proposed to be involved in this process also (Bunz et al., 1998). Additionally, DNA damage induces p53-dependent expression of 14-3-3σ, a homologue protein to the yeast rad24-rad25 (Hermeking et al., 1997). 14-3-3σ supposedly acts similar to the yeast rad24-rad25 to inhibit the function of cdc25, that phosphorylates cdc2 to proceed to mitosis.

**Programmed cell death – apoptosis**

Induction of p53-dependent apoptosis upon DNA damage is more difficult to explain because apoptosis appears to be mediated by various overlapping mechanisms (reviewed in Johnstone et al., 2002). Apoptosis can be induced by both p53-independent and p53-dependent pathways, whereas the latter can be p53-transcription dependent or independent (reviewed in Moll and Zaika, 2001). Transcription of genes that promote apoptosis involve an expanding number of genes, including BAX, the insulin growth factor IGF-BP3 (reviewed in May and May, 1999). p53-dependent apoptosis is thought to be conducted in two ways. First, apoptosis in conveyed by transcription of apoptotic genes such as BAX, probably with involvement of other factors (Thornborrow and Manfredi, 2001). In normal cells, BAX is inactivated by heteroligomerization with gene products from the anti-apoptotic gene family bcl-2 (reviewed in White, 1996). Recently, a p53-induced gene product was found (Noxa), that oligomerizes with bcl-2 to induce apoptosis (Oda et al., 2000). Secondly, p53 represses expression of the bcl-2 anti-apoptotic genes (Wu et al., 2001).

Decision between the apoptotic versus cell cycle arrest outcome is of paramount
significance regarding the clinical treatment of cancer (reviewed in Johnstone et al., 2002). Answering this question is not trivial, because of the multidimensional nature of p53 functionality. First, the biological outcome can be an inherent property of the p53 protein itself; for instance, p53 mutations have been identified that differentially affect cell cycle arrest and apoptotic functions (Almog et al., 2001; Aurelio et al., 1998; Flaman et al., 1998; Friedlander et al., 1996; Lassus et al., 1998). Secondly, p53 induces a variety of genes involved in cell cycle arrest and apoptosis, as discussed previously. Finally, the complexity of the protein networks that employ protein-protein interactions further complicates the analysis (Agarwal et al., 1998). What is evident is that the “gate-keeping” function of p53 (Levine, 1997) relies on a complex regulatory system of multiple interactions.

1.4.2 Functional domains of p53

To date, p53 has been identified in various species, both vertebrates (human, rodents, monkeys) and invertebrates (X. laevis, D. megalogaster, C. elegans) (Brodsky et al., 2000; Caron de Fromentel et al., 1992; Derry et al., 2001; Lamb and Crawford, 1986; Ollmann et al., 2000; Oren and Levine, 1983; Rigaudy and Eckhart, 1989; Soussi et al., 1988, 1987; Zakut-Houri et al., 1985). Comparison of the various p53 proteins revealed a protein whose functions can be roughly dissected in three domains: (a) the N-terminal transactivation domain, (b) the central core specific DNA binding domain and (c) the C-terminal domain that is comprised of the oligomerization domain and the C-terminal regulatory domain (Soussi et al., 1990). Moreover, p53 is characterized by the presence of five distinct regions that are conserved throughout evolution, with four of those regions located in the central core domain and containing most of the mutation hotspots found in human cancers (Soussi et al., 1990). The overall organization is retained in the p53 protein homologues p63 and p73 (reviewed in Arrowsmith, 1999) that were identified as earlier phylogenetic relatives of p53 (re-
viewed in Strano et al., 2001). The organization of p53 in roughly three domains, as indicated by the protein homology, was independently confirmed by biochemical studies (Pavletich et al., 1993; Wang et al., 1994, 1993, 1995).

**N-terminus transactivation domain (amino acids 1–90)**

The N-terminal domain of p53 contains two adjacent transactivation regions (amino acids 1–42 and 43–63) (Candau et al., 1997; Venot et al., 1999; Zhu et al., 1998) and a downstream hydrophobic proline-rich region (Matas et al., 2001; Müller-Tiemann et al., 1998; Sakamuro et al., 1997; Venot et al., 1998, 1999; Walker and Levine, 1996). Although this domain does not bind DNA per se, it is crucial for the DNA binding properties of p53, as this domain has a stabilizing effect on the protein-DNA complex (Hansen et al., 1996, 1998). This domain was found to be indispensable for *in vivo* transactivation properties of p53, and this activity has been pinpointed to residues 22 and 23 (Lin et al., 1994; Matas et al., 2001). This effect supposedly is conferred via protein-protein interactions; in fact, this domain is the binding site of transcription factors, such as the TATA box-binding protein (TBP), TFIID, and TBP-associated factors (reviewed in Ko and Prives, 1996). These protein-protein interactions are mostly hydrophobic in nature (Lin et al., 1994). The free N-terminal domain is mostly unstructured in solution, except for an amphipathic helix T18–L26 and two β-turns at M40–M48 and D48–W53, respectively (Lee et al., 2000).

This region also contains the binding site for MDM2 (Bottger et al., 1996; Lin et al., 1994; Picksley et al., 1994). Details of the MDM2-p53 interactions were provided by solution of the X-ray structure of p53 complexed to a 15 residue p53 peptide (Kussie et al., 1996). Hydrophobic residues F19, L22, W23 and L26 form an amphipathic helix that interacts with residues located in a hydrophobic pocket on MDM2 (see Figure 1.12). Stability of the p53-MDM2 complex is important for MDM2-dependent p53 degradation (Fuchs et al., 1998).
Figure 1.12: MDM2-p53 Interactions

Structure of the complex between MDM2 and the N-terminal region of p53 (amino acids 13–29) obtained from PDB file 1YCQ (Kussie et al., 1996). Both proteins are presented in ribbons, with MDM2 and p53 colored in blue and green, respectively.

Post-translational modifications of this domain affect the stability of the p53-MDM2 complex. Phosphorylation at S20 inhibits MDM2 binding to p53, protecting p53 from degradation (Craig et al., 1999; Unger et al., 1999). S20 is the target of checkpoint kinases Chk1 and Chk2, which are downstream of ATM and ATR kinases, connecting DNA damage and p53 stability (Chehab et al., 1999; Shieh et al., 2000, 1999). ATM and ATR also phosphorylate S15 in the presence of DNA damage (Banin et al., 1998; Canman et al., 1998; Nakagawa et al., 1999; Tibbetts et al., 1999).

Core specific-DNA binding domain (amino-acids 91–299)

The importance of the DNA binding properties of p53 as a carrier of its biological functions is demonstrated by the fact that most mutant p53 proteins from tumor cells display a loss in DNA binding (Bargonetti et al., 1991, 1993; Kern et al., 1991). The specific-DNA binding property of p53 was identified in the region com-
prised of amino acids 100–290 (Pavletich et al., 1993). Two independent approaches identified the half-site of the consensus sequence for the DNA binding target as 5'-PuPuPuC(a/t)(t/a)GPyPyPy-3' where Pu and Py represent purine and pyrimidine bases, respectively (el Deiry et al., 1992; Funk et al., 1992). The importance of the consensus sequence was also confirmed by site-directed mutagenesis of individual bases (Halazonetis et al., 1993). Although insertions within the half-site are not well-tolerated, insertions between two half-sites were tolerated only when the half-sites were on the same side of the DNA-helix (Tokino et al., 1994; Wang et al., 1995). The core domain has an inherent specific-DNA binding affinity ranging from 4 to 100 nM (Balagurumooorthy et al., 1995; Klein et al., 2001; Nagaich et al., 1997a; Nikolova et al., 1998). This range of affinities appears to be dependent on the flexibility properties of DNA, since the tighter affinities were associated with increased DNA bending (Nagaich et al., 1997a). Also, the topology of DNA is suggested to play a role, since p53 can bind better to supercoiled DNA or DNA that is presented as a stem-like conformation (Kim et al., 1997; Palecek et al., 1997).

Cho et al. (1994) solved the crystal structure of the core domain bound to a 21-bp DNA encompassing a p53 half-site (see Figure 1.13A). The overall core domain structure is characterized as a sandwich of two antiparallel β-sheets. This structure provides a scaffold for a loop-sheet-helix motif, which positions into the major groove of DNA. Another loop that inserts into the minor groove, is held in place by a zinc atom, which is liganded in a tetrahedral coordination. Although the asymmetric unit contained a complex of three p53 core domains, only one of them appeared to make specific contacts to the DNA, whereas the others were involved in non-specific DNA and protein-protein contacts to stabilize the crystal packing. Comparison of the three core domain monomers did not show any significant global structural differences, suggesting that DNA complex formation did not cause any significant conformational changes. Similar observations were observed in the structure of the solved X-ray
Figure 1.13: Core domain of p53

The X-ray structure of the core domain of human p53 (96–289) complexed with DNA (Cho et al., 1994) and free murine p53 (99–284) (Zhao et al., 2001) are displayed on top and bottom panel, respectively. The respective pdb files are 1TSR and 1HUB. The location of the Zn++ atom is illuminated. The L1 loop, which is major difference between the DNA-bound and free core domains (Zhao et al., 2001) is displayed in yellow on the free murine core domain structure.
structure of the murine core domain (Zhao et al., 2001). The major difference between the free-protein and the protein-DNA complex was position of the L1 loop relative to the H2 helix that makes contacts to the DNA (Figure 1.13B). In the protein-DNA complex, the L1 loop and the H2 helix are further way, allowing them to contact the major and minor groove of the DNA, respectively, whereas in the free protein, their position would sterically clash with the DNA.

Since most tumor-causing p53 mutations are located in the core domain, Cho et al. (1994) mapped the most common mutations on the solved structure. p53 hotspot mutations are classified into two groups: (a) mutations at residues such as R248 and R273 that contact DNA and (b) mutations at residues such as R175, R249, R282 and G245 that destabilize the structure of the DNA binding surface of p53. Overall, the most frequent mutations are located on the DNA binding surface or play a role in destabilizing it, whereas mutations further away appear to be more tolerated.

Denaturation studies confirmed in general the above predictions. Mutations at DNA contacts destabilized core domains to a mild to moderate degree, whereas mutations at sites that supported the structure displayed a more severe destabilization compared to wild type core, with the most pronounced effect on residues involving the zinc coordination residues (Bullock et al., 1997, 2000). In a different approach, protein conformations can be distinguished by their differential reactivity to specific monoclonal antibodies. Of particular interest are the antibodies pAb1620 (pAb246 for murine p53) and pAb240 that recognize “wild type” or “mutant” p53 phenotypes, respectively (Gannon et al., 1990; Kraiss et al., 1991; Milner and Medcalf, 1991). Reactivity for pAb1620 has been associated with intact DNA binding properties, although many mutant p53 proteins with eliminated DNA activity display pAb1620+ on phenotype, whereas reactivity for pAb240 in general corresponds with structural mutations (Rolley et al., 1995). Recent studies placed the antigenic epitopes for pAb1620 on an elongated surface containing residues 206–210, whereas
pAb240 mapped to the region 213–217 that is covered by the surface detected by pAb1620 (Wang et al., 2001). In this way, the two antibody reactivities are mutually exclusive, since pAb240 recognition requires the epitope of pAb1620 to unfold (Wang et al., 2001). An interesting temperature sensitive murine mutant, A138V, displayed a conversion from wild type to mutant conformation, as the temperature increased above 37°C (Ginsberg et al., 1991; Michalovitz et al., 1990). However, conformational analysis based on the monoclonal antibodies is not straightforward for every case; for instance, R175H displays a pAb1620 phenotype, but is incapable of binding DNA (Cohen et al., 1999).

**Oligomerization domain (amino acids 300–360)**

Although the core domain itself can bind DNA in a process where four monomers can bind cooperatively (Cho et al., 1994; Nagaich et al., 1997b), p53 has a separate region consisting of amino acids 300–360 that confers oligomeric properties of the full length proteins (Pavletich et al., 1993; Wang et al., 1993). p53 forms tetramers and multiples of tetramers, and this feature has been further attributed to the region of 323–355 (Wang et al., 1994). The DNA topology of the binding target also reflects the requirement for tetrameric formation (Wang et al., 1995). Oligomerization is required for effective DNA binding and wild type p53 activities *in vivo* (Chéne et al., 1997; Davison et al., 1998; Lomax et al., 1998; Rollenhagen and Chéne, 1998; Waterman et al., 1995, 1996). Effective DNA binding *in vivo* can be possibly mediated by formation of DNA loops (Stenger et al., 1994). Oligomerization properties of p53 have been implicated in the ‘gain of function’ phenomenon, where mutant p53 proteins exert their dominant effect over their wild type partners (reviewed in Roemer, 1999).

The structure of the oligomerization domain has been solved by both X-ray crystallographic and NMR studies (Clore et al., 1995; Jeffrey et al., 1995; Lee et al.,
Figure 1.14: Tetramerization domain of p53

Structure of the tetramerization domain obtained from PDB file 1C26 (Jeffrey et al., 1995). Dimerization occurs by formation of an antiparallel $\beta$-sheet and two antiparallel $\alpha$-helices by two monomers. Monomers belonging to a dimeric subunit are shown in red and orange or cyan and mauve. Two dimers associate into tetramers via hydrophobic interactions originating from the $\alpha$-helices.

1994). The structure employs a tetramer arranged as a dimer of dimers (see Figure 1.14). Each monomeric domain features a $\beta$-strand (amino acids 326–333), a tight turn (G334) and an $\alpha$-helix (amino acids 335–355). Two monomers associate into a dimer by forming an antiparallel $\beta$-sheet and two antiparallel $\alpha$-helices. Association into tetramers involves interaction between the $\alpha$-helices interacting mostly by hydrophobic contributions of L344 and L348 in each monomer (Jeffrey et al., 1995; McCoy et al., 1997; Stavridi et al., 1999). Initially, partially folded monomers dimerize transiently to produce structured dimers, which then associate into the final tetramers (Mateu et al., 1999). The tetramers formed are very stable in solution as displayed by thermal denaturation studies (Johnson et al., 1995; Mateu and Fersht, 1999) with the association stability in the $\mu$M range (Sakaguchi et al., 1997). Destabilization generally requires at least 2 to 3 simultaneous mutations, although mutations at crucial residues such as L344 are enhance to destabilization (Jeffrey et al., 1995; Mateu
et al., 1999). However, the low incidence of human cancers that contain mutations in this domain can attributed to both the relative stability of the oligomerization domain and the requirement of oligomerization for mutant p53 proteins to exert their oncogenic potential (Chéne and Bechter, 1999).

C-terminal regulatory domain (amino acids 360–393)

The final 30 amino acids at the C-terminal end form the C-terminal regulatory domain. This characterization originated from the effect of covalent and non-covalent modifications that appeared to increase DNA-binding by the central core (see Section 1.4.3). This domain affects biological function of p53 in vivo, since C-terminal deletion mutants display defects in their apoptotic but not cell cycle arrest responses (Lassus et al., 1998). p53 proteins with this domain altered have distinct functional properties, as found in alternatively spliced p53 in rodent cells (Almog et al., 1997, 2000; Miner and Kulesz-Martin, 1997; Rehberger et al., 1997; Wesierska-Gadek et al., 1999) and the p53 homologue proteins, p63 and p73, that contain different C-termini (reviewed in Strano et al., 2001).

Originally, this domain was proposed to confer the DNA binding properties of p53 (Foord et al., 1991). Later studies clarified the DNA binding properties of this domain and showed that it binds DNA in a non-specific manner (Wang et al., 1994). Moreover, this domain is able to bind single stranded DNA ends, an activity that contributes to p53’s ability to recognize gapped double stranded DNA, resembling damaged DNA (Bakalkin et al., 1995, 1994; Selivanova et al., 1996; Zotchev et al., 2000). This domain is also required for p53 to recognize DNA targets in chromatin-like structures (Espinosa and Emerson, 2001).

The C-terminal domain contains serine and lysine residues that undergo post-translational modifications (reviewed in Wolkowicz and Rotter, 1997). These modifications have been proposed to affect the DNA binding properties by both the central
core domain (see Section 1.4.3) and the C-terminal regulatory domain itself (Hoffmann et al., 1998). Phosphorylations at S315 and S392 have been demonstrated as downstream responses after DNA damage, induced by kinases p34 and CKII (Kapoor and Lozano, 1998; Keller et al., 2001; Lu et al., 1999b). In contrast, S376 seems to be constitutively phosphorylated (Waterman et al., 1998).

The cellular p53 stabilization occurring after DNA damage is proposed to be a result of both post-translational modifications of MDM2 that diminish its activity and acetylation induced at the p53 C-terminal regulatory domain of p53 (Grossman et al., 1998). Acetylation has been shown to occur in vivo at K373 and K382 by the CREB binding protein (CBP/p300) and at K320 by the CBP/p300 associated factor or PCAF (Liu et al., 1999; Sakaguchi et al., 1998). Acetylation has been proposed to aid in p53 transactivation by the recruitment of CBP/p300 (Avantaggiati et al., 1997) and deacetylation by histone deacetylases (HDACs) downregulates activity of p53 (Juan et al., 2000). Interestingly, MDM2 inhibits CBP/p300 induced p53 phosphorylation, implying participation of both N and C-terminal ends in regulating activity and cellular stability of p53 (Kobet et al., 2000).

1.4.3 Allosteric regulation of p53 – p53 activation

The notion that conformation was an important determinant of p53’s function was originally suggested by Milner and colleagues, based on antibody reactivities of wild type (pAb240+/pAb1620+) and mutant p53 proteins (pAb240-/pAb1620-) (Cook and Milner, 1990; Milner and Medcalf, 1990, 1991). The concept that the conformation of p53 was a crucial determinant of its function was further supported by interconversion of wild type and temperature sensitive p53 mutants (Hainaut et al., 1995; Medcalf et al., 1992; Milner and Medcalf, 1990).

Hupp and coworkers extended the concept of a conformational switch to formulate an activation model for p53 regulation. Unmodified p53 (for instance, when
expressed in bacterial cells) exhibited negligible or very low DNA binding affinity (Hupp et al., 1992). Deletion of the last 30 C-terminal amino acids resulted in a protein that bound DNA efficiently (Hupp et al., 1992). Further studies showed that a number of modifications, covalent or non-covalent, were required so that p53 exhibited DNA binding activity (Table 1.1). These modifications were located in or involved almost exclusively the last 30 amino acids of the protein. Therefore, an activa-

tion model was formulated to explain the DNA binding properties of p53 (Hupp and Lane, 1994, 1995; Jayaraman and Prives, 1995). According to this model, p53 in normal cells exists in a “latent” form, incapable of binding DNA (see Figure 1.15), because of the inhibitory effect of the native C-terminal regulatory domain. Modifications of the C-terminal regulatory domain relieve the inhibitory effects and induce a conformational change on p53 into an “activated” conformation capable of binding DNA. Since the activating processes include phosphorylation and acetylation modifications that occur in vivo, the activation model was widely accepted to describe p53's

Table 1.1: p53 modifications as a basis for the allosteric activation model

<table>
<thead>
<tr>
<th>Modification</th>
<th>Affected region</th>
</tr>
</thead>
<tbody>
<tr>
<td>phosphorylation by CKII²</td>
<td>S392</td>
</tr>
<tr>
<td>acetylation by p300/CBP³</td>
<td>K373, K382</td>
</tr>
<tr>
<td>deletion of last 30 amino acids⁴</td>
<td>363–393</td>
</tr>
<tr>
<td>binding by antibody pAb421⁵</td>
<td>371–380</td>
</tr>
<tr>
<td>single stranded DNA⁶</td>
<td>361–382</td>
</tr>
<tr>
<td>C-terminal peptides 369–382⁷</td>
<td>–</td>
</tr>
</tbody>
</table>

Figure 1.15: Activation model of p53 regulation

Regulation by the C-terminal regulatory region is the basis for the p53 activation model. For simplicity, p53 models contain only the DNA binding domain (green or red for DNA binding active or inactive, respectively) and the C-terminal regulatory domain (gray ellipse). Native C-terminal regulatory domain inhibits DNA binding by the central core (red). Covalent or noncovalent modifications of the C-terminal regulatory domain relieve the inhibitory effect, and consequently activate DNA binding by the central core (green) (Gu and Roeder, 1997, adapted).
biological functions.

Two mutant proteins that were employed for developing the allosteric model necessitate more detail. First, incorporation of a acidic amino acid at residue S389 in murine p53 (its human equivalent is S392) resulted in an activated protein (Hao et al., 1996). Phosphorylation of this residue is implicated in DNA damage (Kapoor and Lozano, 1998), it is phosphorylated in vitro by casein kinase II (CKII), and the phosphorylated protein exhibited activated properties (Hupp and Lane, 1995; Hupp et al., 1992). Secondly, the deletion of the C-terminal regulatory domain of p53 also increased its DNA binding properties dramatically and was the foundation of the activation model (Hupp et al., 1992).

Accumulating evidence, however, has called the proposed allosteric regulation model into question. According to the allosteric model, unmodified p53 was considered unstable, and activated p53 was stabilized by the allosteric effectors (Hall and Milner, 1995). In this case, cellular stability was misinterpreted as thermodynamic stability of the protein (Haupt et al., 1997; Kubbttat et al., 1997; Midgley and Lane, 1997). In fact, both wild type and S392E p53 proteins exhibited high thermal stability and retained structural elements up to 70 °C (Nichols and Matthews, 2001, 2002).

Anderson et al. (1997) questioned the conditions utilized for in vitro binding assays. More specifically, most DNA binding assays for p53 employed a large excess of non-specific DNA in the binding reactions. Anderson et al. (1997) showed that large non-specific double-stranded DNA inhibited specific DNA binding by p53. More recently, wild type p53 binding to chromatin-like DNA was not influenced by activating factors affecting the C-terminal regulatory domain (Espinosa and Emerson, 2001; Kaeser and Iggo, 2002). In fact, bacterially produced wild type p53 displayed a DNA binding affinity in the nM range when non-specific DNA was excluded from the binding assays (Nichols and Matthews, 2001). Also, structural data from NMR studies of unmodified p53 and protein with its C-terminal regulatory domain deleted
showed no significant structural changes (Ayed et al., 2001). Along this line, the S392E p53 mutant showed similar, if slightly weaker, DNA binding affinity compared to wild type (Nichols and Matthews, 2002).

1.5 Objectives

Complete and thorough thermodynamic description of protein-DNA interactions provides a level of understanding of the underlying mechanisms that govern the biological function of protein transcription factors. One of the proteins involved in this thesis, LacI, has been extensively employed the last 35 years for establishing a detailed thermodynamic description of protein-DNA interactions (reviewed in Matthews and Nichols, 1998). On the other hand, limited information is available for its homologue protein, PurR, that is involved in the de novo biosynthesis of purines. Therefore, complete thermodynamic description of ligand binding (purine or DNA) was undertaken in the equilibrium and kinetic studies presented in Chapter 3. To validate the consistency of the thermodynamic description, the data will be analyzed in two ways: (a) by dissecting the ligand binding processes into independent binding events, and (b) by simultaneous global analysis of all data. More importantly, comparison of the thermodynamic analyses of PurR and LacI will provide insight on how proteins with similar structures elicit different biological responses.

Protein-DNA interactions rely on both electrostatic and hydrophobic interactions. These are reflected by the “polyelectrolyte” and the “hydrophobic” effects, that are monitored by the ion concentration and temperature dependence of DNA binding (reviewed in Record et al., 1991; Spolar and Record, 1994). Therefore, it is meaningful to question whether the homology of the PurR and LacI proteins extends to the thermodynamic behavior of these proteins, as studied with regard to the “polyelectrolyte” and the “hydrophobic” effects in Chapter 4. Moreover, availability
of high resolution X-ray structures for the LacI and PurR proteins (Bell and Lewis, 2000; Schumacher et al., 1997) will allow insight at the atomic level, in an attempt to couple the macroscopic observations of the thermodynamic studies to the microscopic details of the molecular structures.

The studies in Chapters 3 and 4 revolve around proteins from ‘humble’ bacteria. In Chapter 5, the thermodynamic description will be extended to the p53 human suppressor protein, whose normal function is of paramount importance in the prevention of tumorigenesis (reviewed in Levine, 1997). Most biochemical studies in the early and mid-90s focused on a now questionable allosteric model of p53 regulation, with the C-terminal end of p53 negatively regulating its DNA binding affinity. However, no organized thermodynamic description existed before efforts in the Matthews lab to systematically characterize p53’s DNA binding properties. These efforts initially focused on wild-type and S392E p53. The latter is a mutant protein that had been considered be DNA-binding active according to the activation model (Nichols and Matthews, 2001, 2002). To further test the activation model for p53 regulation, thermodynamic characterization in the studies presented in this chapter will be focused on Δ33 p53, a mutant protein that lacks the C-terminal regulatory domain and therefore has been considered DNA-binding active (Hupp et al., 1992). Both the “polyelectrolyte” and “hydrophobic” effects displayed by Δ33 p53 will be compared to those of the previously characterized wild-type and S392E p53 proteins. These effects will also be analyzed with the aid of the only available solved structure of the core DNA binding domain by Cho et al. (1994).
Chapter 2

Materials and methods

2.1 Theoretical aspects

2.1.1 DNA binding

When a protein (P) binds DNA (D) to form a complex (P • D), the process can be modelled as a single binding site reaction: \( P + D \rightarrow P \cdot D \). The affinity for this reaction can be quantified with the expression of the equilibrium dissociation constant \( K_D \):

\[
K_D = \frac{[P][D]}{[PD]} \quad (2.1)
\]

Quantitative estimation of the \( K_D \) is derived from the binding isotherm, that is, the fraction of the total DNA (\( D_T \)) bound, \( R \). When non-stoichiometric conditions apply (\( D_T \ll K_D \)) then the free protein concentration is approximated as follows:

\[
[P] = \frac{[P_T]}{1 + \frac{[D]}{K_D}} \quad \text{or} \quad [P] \approx [P_T] \quad (2.2)
\]

\[
[2.3]
\]
Then the fraction of DNA bound, \( R \), can be estimated by the following equation:

\[
R = \frac{[PD]}{[D_T]} = \frac{[P_T]}{[P_T] + K_D}
\]  \hspace{1cm} (2.4)

Although this expression favors simplicity, its application is limited because:

- it does not contain any error factor regarding background or plateau, and

- it requires Y-data transformation to obtain the \( R \) data from the experimentally measured radioactivity signal, propagating experimental error in a non-controlled manner.

Therefore, for determining \( K_D \)s from DNA binding titrations, the following equation was used:

\[
Y_i = Y_0 + Y_m \frac{[P_T]}{[P_T] + K_D} = Y_0 + Y_r \frac{10^{\log [P_T]}}{10^{\log [P_T]} + 10^{\log K_D}}
\]  \hspace{1cm} (2.5)

where \( Y_i \) corresponds to the total signal from bound DNA at a given protein concentration, \( Y_0 \) is the background signal, \( Y_r \) is the range of the binding signal \( (Y_{max} - Y_0) \) and \( [P] \) is the total protein concentration, taking into account the oligomeric state of the examined proteins. The second part of equation (2.5) was used to analyze the data when presented in the traditional log-scale X axis of log\([P_T]\). Equation (2.5) implements equation (2.4) and addresses the original limitations. Indeed, the following transformation provides the value for the fractional saturation, \( R \):

\[
\frac{Y_i - Y_0}{Y_m} = R
\]  \hspace{1cm} (2.6)

A similar equation to Equation 2.4 expresses the Hill empirical model:

\[
Y_i = Y_0 + Y_r \frac{[P]^n}{[P]^n + K_D^n} = Y_0 + Y_m \frac{10^{\log [P_T]^n}}{10^{\log [P_T]^n} + 10^{\log K_D^n}}
\]  \hspace{1cm} (2.7)

This model assumes equivalent binding by \( n \) sites with an overall average \( K_D \). Of
note, this equation is empirical and does not derive from a physical model. Therefore, the values of $K_D$ derived by this model may not correspond to the affinity of a binding site, but represents the overall apparent $K_D$. Only when $n \approx 1$ will the apparent Hill dissociation constant represent the 'real' $K_D$ of a single binding site model.

The applicability of the nonstoichiometric condition relies on the fact that known total protein concentrations $[P_T]$ suffice for monitoring DNA binding experiments, rather than the free protein concentration $[P]$, which is rarely known. However, under stoichiometric conditions ($D_T >> K_D$), such as in DNA binding activity estimations, the free protein concentration is derived by solving the system of mass conservation equations for $P_T$ and $D_T$. The analytical solution for the free protein concentration $[P]$ is presented in its numerically stable form:

$$[P] = \frac{2P_T K_D}{K_D + D_T - P_T + \sqrt{(K_D + D_T - P_T)^2 + 4P_T K_D}}$$ (2.8)

This equation was applied for the estimation of the activity $A$ of the proteins used in the studies of this thesis, by substituting the free protein concentration, $[P]$, with the active free protein concentration, $[P]'$, as given by the following equation:

$$[P]' = A \times [P]$$ (2.9)

**Global equilibrium analysis for PurR**

Ligand (guanine) binding to PurR can be represented as a two-site binding model

$$P + G \rightleftharpoons K_1 P \cdot G \rightleftharpoons K_2 P \cdot G_2$$ (2.10)

where $P$ and $G$ denote the dimeric PurR and guanine, respectively. The affinities for the interactions for the first and second binding site can be represented by the
macroscopic equilibrium binding constants, \( K_1 \) and \( K_2 \), respectively. The mass conservation equations for dimeric protein and guanine can be expressed in terms of the binding constants, as follows:

\[
[P_T] = [P] + [PG] + [PG] = [P] + K_1[P][G] + K_1K_2[P][G]^2 \tag{2.11}
\]

\[
[G_T] = [G] + [PG] + 2[PG] = [G] + K_1[P][G] + 2K_1K_2[P][G]^2 \tag{2.12}
\]

Since, in most cases, the known experimental parameters are the total concentrations instead of free concentrations, these equations illustrate that there is no analytical solution that can express the free concentrations as a closed-form solution in terms of the total concentrations. In contrast, for a single binding site, the free concentrations can be expressed as a closed-form solution in terms of total concentrations, as discussed in Section 2.1.1. However, the above system of nonlinear equations can be transformed to a single nonlinear equation that can be solved numerically. Therefore, the fraction of free guanine and protein concentrations are defined as \( x \) and \( z \) respectively,

\[
x = \frac{[G]}{[G_T]} \tag{2.13}
\]

\[
z = \frac{[P]}{[P_T]} \tag{2.14}
\]

then the mass conservation nonlinear system can be transformed to the following equation

\[
(1 + 2K_2[G_T]x)K_1[P_T]zx + x - 1 = 0 \tag{2.15}
\]

where

\[
z = \frac{1}{1 + K_1[G_T]x (1 + K_2[G_T]x)} \tag{2.16}
\]

Binding of PurR to both guanine and DNA can be modelled by the following
thermodynamic cycle presented in Figure 2.1. Of note, description of the system requires a total of five thermodynamic parameters (depicted as the association constants). The binding equilibrium not indicated (dotted line) can be obtained as a combination of the depicted $K$, $M_1$, $M_2$, $N_1$ and $N_2$ parameters, since the system forms a complete cycle.

![Diagram](image)

**Figure 2.1: Thermodynamic cycle for guanine and DNA binding of PurR**

The dimeric purine holorepressor contains one DNA binding site and two guanine binding sites (one per monomer). Normal case letters represent the stoichiometries of various species, where $P$, $G$ and $D$ stand for dimeric PurR, guanine and DNA, respectively. In *italics* the macroscopic equilibrium association constants are indicated. Equilibrium indicated with dotted line can be derived from the other depicted parameters.

The mass conservation equations for dimeric protein, DNA and guanine can be expressed as follows:

$$[P_T] = [P] + [PG] + [PG_2] + [PDG] + [PDG_2]$$  \hspace{1cm} (2.17)

$$[D_T] = [D] + [PD] + [PDG] + [PDG_2]$$  \hspace{1cm} (2.18)

$$[G_T] = [G] + [PG] + 2[PG_2] + [PDG] + 2[PDG_2]$$  \hspace{1cm} (2.19)
Similarly, if the fractions of free guanine, DNA and protein concentrations are defined as $x$, $y$ and $z$, respectively, 

\[
\begin{align*}
x &= \frac{[G]}{[G_T]} \\
y &= \frac{[D]}{[D_T]} \\
z &= \frac{[P]}{[P_T]}
\end{align*}
\]

the 3x3 nonlinear system of the mass conservation equations can be transformed to a single nonlinear equation that can be numerically solved:

\[
(1 + z(M_1[P_T] + 2M_2[G_T]x + KN_1[P_T][D_T]y(1 + 2N_2[G_T]x)))x - 1 = 0
\]

where

\[
\begin{align*}
z &= \frac{2}{BB + \sqrt{BB^2 + 4AA}} \\
y &= \frac{1}{1 + K[P_T]Az} \\
AA &= K[P_T]A(1 + B) \\
BB &= 1 + B + KA([D_T] - [P_T]) \\
A &= 1 + N_1[G_T]x(1 + N_2[G_T]x) \quad \text{and} \\
B &= M_1[G_T]x(1 + M_2[G_T]x)
\end{align*}
\]

The thermodynamic parameters describing the global equilibrium cycle for PurR can be expressed via the microscopic binding constants. If $k$ and $m$ are defined as the microscopic association equilibrium constants for binding of DNA and guanine respectively, then the following sets of equations correlate the macroscopic to the
microscopic constants:

\[ K = k \]  \hspace{1cm} (2.24) \\
\[ M_1 = 2m \]  \hspace{1cm} (2.25) \\
\[ M_2 = \frac{sm}{2} \]  \hspace{1cm} (2.26) \\
\[ N_1 = 2s_1m \]  \hspace{1cm} (2.27) \\
\[ N_2 = \frac{s_2m}{2} \]  \hspace{1cm} (2.28)

where \( s, s_1 \) and \( s_2 \) stand for cooperativity constants of guanine binding to the second site, binding to the first and second site in the presence of DNA, respectively.

### 2.1.2 DNA binding dissociation kinetics

When a protein (\( P \)) binds DNA (\( D \)) with a 1 to 1 stoichiometry

\[ P + D \overset{k_d}{\underset{k_a}{\rightleftharpoons}} P \cdot D \]

the kinetics of the reactions are described by the following equation:

\[
\frac{d[P\cdot D]}{dt} = k_a[P][D] - k_d[P\cdot D] \tag{2.29}
\]

For the dissociation studies in this thesis, dissociation experiments were performed by monitoring the decrease of the concentration of complexed nonstoichiometric labelled DNA ([\( PD^* \)]) after excess non-specific DNA ([\( D_T \)]) was added. In this case, it can be assumed that the association term of for [\( PD^* \)] becomes negligible, because the high excess of unlabelled DNA competes effectively with any free labelled DNA for
the protein binding sites. Therefore, Equation 2.29 on the preceding page becomes:

\[
\frac{d[PD^*]}{dt} = -k_d[PD^*]
\]

(2.30)

\[
\frac{[PD^*]}{[PD^*_0]} = e^{-k_d t}
\]

(2.31)

where \([PD^*_0]\) is the labelled DNA-protein complex at the time when unlabelled excess of DNA is added.

### 2.1.3 Ion concentration dependence of DNA binding – Polyelectrolyte effect

Nucleic acids can be visualized as large polyanions because of the high negatively charged nature of their phosphate backbone (Figure 2.2). Consequently, nucleic acids attract a large number of cations to counteract the negative charge. This cation

![Figure 2.2: Polyelectrolyte effect](image)

Free DNA binds a number of counter-cations, due to the high negative charge. When basic protein sidechains bind to DNA, counter-cations are released. This process is entropically favorable.
accumulation is described by the term “polyelectrolyte” effect. Based on this perspective, protein-DNA interactions can be analyzed by a model where the positively charged protein side chains can serve as polycations that replace counter cations from the DNA surface.

Record et al. (1977, 1991) formulated a theoretical basis for studying the polyelectrolyte effect in protein-DNA interactions. Protein (P) DNA (D) complexation to yield a complex (P • D) can be represented as follows:

\[ P(aM^+, bH_2O) + D(cX^-, dH_2O) \rightarrow P \cdot D + aM^+ + cX^- + (b + d)H_2O \]

in which a, c, and (b+d) are the stoichiometric coefficients for release of cations, anions and water, respectively, the dependence of the binding affinity (\(K_A\)) as a function of salt concentration ([MX]) is provided by the following equation:

\[
\left( \frac{\partial \ln K_A}{\partial \ln [MX]} \right)_{PT} = -a - c + 2(b + d) \frac{[MX]}{[H_2O]} \tag{2.32}
\]

Based on experiments using divalent or polycations the dependence on the anion concentration was found to be negligible in magnitude (Record et al., 1991). This conclusion can be substantiated by the strong negative charge of DNA and exclusion of anions from the protein-DNA binding site. Therefore, the dependence of the binding affinity depends predominantly on the stoichiometry of the cation release:

\[
\frac{\partial - \log K_D}{\partial \log [MX]} = -a \tag{2.33}
\]

Furthermore, these experiments confirmed previous theoretical studies that predicted that DNA neutralization by counter cations accounted for 88% of the DNA charge. Taking this relationship into account, the ion concentration dependence of
the DNA binding affinity can be further simplified to (Record et al., 1991):

\[ \frac{\partial - \log K_D}{\partial \log [MX]} = -0.88 + Z \]  

(2.34)

where \( Z \) denotes the number of positive charges displaced from the negatively charged DNA backbone by the positively charged protein side-chains.

The above theoretical basis predicts that protein-DNA interactions are favored in low salt concentrations and that increase of the ionic strength results in decreased protein-DNA binding affinity. This prediction has been confirmed experimentally in most protein-DNA interactions (deHaseth et al., 1977a; Frank et al., 1997; Hamilton et al., 1998; Lohman et al., 1996; O’Brien et al., 1998). The simplification of Equation 2.32 relies on two assumptions. First, the dehydration (number of H\(_2\)O molecules released) upon complex formation is negligible relative to the cation release. This assumption holds true at low salt concentrations relative to the bulk concentration of water. For example, the apparent opposite behavior of the thermophilic TATA binding protein, with increasing DNA binding affinities at higher salt concentrations, is attributed to a large dehydration effect upon complex formation (O’Brien et al., 1998). The second assumption is that the anion release is negligible relative to the release of cations from the DNA backbone. This assumption holds true because of the highly negative charge nature of the DNA and has been confirmed by studies with divalent or polycations. Moreover, ion concentration dependence of DNA binding of LaCl indicated that the anion effect is negligible and that cation release is the governing factor in the ion concentration dependence of DNA binding affinity.
2.1.4 Temperature and osmotic stress dependence of DNA binding – Hydrophobic effect

Most protein-DNA interactions exhibit a large negative heat capacity change (Frank et al., 1997; Ha et al., 1989; Spolar and Record, 1994). For instance, earlier studies on LacI showed that protein-DNA complex formation is accompanied by a $\Delta C_p$ of -0.9 to -1.5 kcal/molK, depending on the operator sequence (Frank et al., 1997; Spolar and Record, 1994). As a result of the large heat capacity change, van’t Hoff plots (affinity vs 1/T) deviate from the typical linearity and are characterized by a highly nonlinear behavior (Ha et al., 1989). Ha et al. (1989) formulated a theoretical interpretation that quantifies the observed large negative heat capacity changes:

$$\ln K_A = \frac{\Delta C_p}{R} \left( \left( \frac{T_H}{T} \right) - \ln \left( \frac{T_S}{T} \right) - 1 \right)$$

(2.35)

where $T_H$ and $T_S$ represent the specific temperatures where enthalpic and entropic contributions to free energy are zero, respectively.

The large negative heat capacity change observed in protein-DNA interactions, has been attributed to the “hydrophobic” effect, a term originated to explain the negative heat capacity changes of protein folding transitions (Privalov and Gill, 1988). The “hydrophobic” effect is described by the release of clathrate-structured water molecules from solvent exposed protein apolar domains, upon protein folding (Spolar et al., 1989). Ha et al. (1989) extended this phenomenon to protein-DNA interactions and suggested that was a major driving force for protein-DNA complex stabilization. In fact, this hypothesis was employed and explained successfully the magnitude of observed heat capacity changes in examples of protein-DNA interactions, with the means of a model based on linkage of local or global protein folding to DNA binding equilibria (Spolar and Record, 1994). Although, the “hydrophobic” effect has been the major component of the negative heat capacity changes in many
protein-DNA interactions, there are a number of examples where there is discrepancy between the theoretical and observed heat capacity change; changes in internal vibrational modes of polar domains, protonation of charged groups, DNA base unstacking, or additional conformational changes, are other examples that can contribute to the negative heat capacity change, in addition to the release of water from the apolar surfaces (Keown et al., 1998; Kozlov and Lohman, 1999, 2000; Ladbury et al., 1994; Lundbäck et al., 1998; Morton and Ladbury, 1996; Oda et al., 1998; Privalov et al., 1999; Zou et al., 1998).

Release of water molecules from the protein-DNA interface may participate in the ion concentration dependence as shown in Equation 2.32. However, study of water release from ion concentration dependence is generally not practical because of the small or negligible contribution of the water release term, under the conditions employed for these studies. Alternatively, the extent of water release is measured by the osmolar dependence of DNA binding for protein-DNA interactions (Parsegian et al., 2000). The binding affinity depends on the osmolar concentration, \( m \), according to the following equation, modified from Garner and Rau (1995):

\[
\left( \frac{\partial (- \log K_D)}{\partial ([m])} \right)_{P,T} = \frac{\Delta n_w}{128}
\]  

(2.36)

where \( \Delta n_w \) denotes the number of water molecules released.

2.2 Experimental aspects

2.2.1 Equilibrium and kinetic parameters for PurR and LacI

Materials

PurR was kindly provided by Dr. Han Xu — a former graduate student in the lab — and the purification protocol is described in detail in Xu (2002). LacI was
purified according to the standard protocol (for instance, as in Barry and Matthews (1999b)).

A 32-bp double stranded DNA containing the purF site was used for DNA binding experiments of PurR. The two strands were as follows: 5'-GAATCCTACGC-AAACGTTTGCCTTTTCTG-3' and 5'-GACAGAAACGCAACGTTCGAG-GGAT-3'. Hybridization of the two strands leaves a 2 bp 5' overhang. For LacI DNA binding experiments, a 40-bp double stranded DNA containing the O1 site was employed. The DNA strand 5'-TGTTGTGGAATTGTGAGCCGGATAACAATT-TCACACAGG-3' and its complementary strand, were hybridized to yield the 40 bp double stranded DNA.

Hybridization of the complementary single stranded target DNA into double stranded DNA was conducted in the presence of 1XSequenase buffer, by heating the hybridization reaction to 70 °C, cooling slowly until room temperature, before transfer on ice. After hybridization, dsDNA was labeled with [γ-32P]-ATP by the T4 polynucleotide kinase method for 1.5 hrs at 37 °C, and the resulting labeled operator was purified by eluting with TE buffer in 400 μL from a Nick column.

**Equilibrium measurements**

All of the DNA binding experiments (equilibrium and kinetic measurements) for both proteins were carried in buffer E: 100 mM HEPES-KOH (pH 7.5), 250 mM potassium glutamate, 150 mM NaCl, 10 mM Mg(CH3COO)2, 1 mM EDTA, with the addition of 0.1 mg/ml bovine serum albumin (BSA) to reduce protein adsorbance to tubes and equipment.

Radiolabelled double-stranded purF and O1 were used at constant non-stoichiometric conditions (DNA concentration was below 1 x 10⁻¹¹ M) with protein titrating over a range of concentrations to obtain a binding isotherm. Equilibrium binding experiments in the presence of guanine were conducted under conditions to satu-
rate PurR corepressor binding sites. Guanine was diluted in buffer E containing BSA overnight, to ensure saturation of the stock solution. After filtration, guanine concentration was determined by measuring the absorbance at 246 and 275 and employing the extinction coefficients of 10700 and 8100 M\(^{-1}\)cm\(^{-1}\) (Windholz, 1983). For DNA binding experiments of PurR in the presence of guanine, reaction mixtures were assembled so that the final guanine concentration in the reaction mixture was 1 \(\times\) 10\(^{-5}\) M. After assembling, reactions were incubated for 1.5 hours in both presence or absence of guanine. Then, reactions were filtered through 0.45 \(\mu\)m nitrocellulose filters, preincubated in buffer E without BSA. After drying, each filter was placed in a scintillation vial containing 5 mL of scintillation liquid for subsequent scintillation counting.

Data were fit to the Hill empirical equation 2.5 with floating values for \(Y_0\), \(Y_r\), \(K_D\) and \(n\). Input data \([P]\) corresponds to dimeric PurR, since each dimer contains the DNA binding site. Data analysis was conducted with the appropriate fitting routine in IgorPro. For presentation purposes, experimental data were normalized and expressed as DNA fractional saturation, according to the fitted parameters.

For corepressor equilibrium studies a similar approach was employed. In the place of radiolabelled DNA, \(^{14}\)C\(\text{guanine}\) was used at 1 \(\times\) 10\(^{-8}\) M. This concentration was well below the estimated equilibrium constants and therefore constituted non-stoichiometric conditions. In the case of guanine binding in the presence of DNA, a constant amount of double stranded \(\text{purF}\) was added with a final concentration of 5 \(\times\) 10\(^{-5}\) M, so that the protein is saturated with DNA. Data analysis employed equation 2.7 with the exception that \([P]\) was expressed in monomeric subunits, since a monomer subunit of PurR contains a guanine binding site.

The equilibrium measurement of LacI employed the same protocol as PurR, with the exception of a 45 min incubation period for the reaction mixtures.
Dissociation kinetic measurements

Dissociation kinetics were performed under first-order reaction conditions (see Section 2.1.2). DNA binding dissociation rates for PurR were performed as follows: PurR (1 x 10^{-7} M dimers) was incubated for an hour with 1 x 10^{-11}M radiolabelled DNA. Then an aliquot was taken and filtered to represent the counts at time 0. Addition of unlabelled concentrated DNA (final concentration 5 x 10^{-7} M) — so that signal dilution due to addition of cold DNA is less than 2% — initiated the time measurement. After specific time points, aliquots were taken from the reaction and filtered on the nitrocellulose filter paper. For dissociation kinetics of PurR in the presence of guanine, guanine was included in the reaction at a concentration of 1 x 10^{-5} M to saturate the protein with guanine. Dissociation kinetics of LacI were performed similarly to the free PurR with the exception that incubation was 30 min.

Data analyses were conducted according to the following equation:

\[ Y(t) = Y_0 + Y_r e^{-kt} \]  

(2.37)

where \( Y(t) \), \( Y_0 \) and \( Y_r \) correspond to signal at a given time \( t \), the signal at equilibrium (background) and the signal range, respectively. This equation is derived from Equation 2.31 with the addition of the floating parameters \( Y_0 \) and \( Y_r \).

Global equilibrium analysis for PurR binding to guanine and DNA

PurR contains one DNA and two corepressor binding sites. The theoretical details on the approach for one binding site (PurR binding to DNA in the absence of DNA), two binding site (PurR binding to guanine) and a global thermodynamic cycle model (PurR binding to guanine and DNA simultaneously) were provided in the Theoretical section (see Figure 2.1). The equations describing the various models, as well as a numerical solver routine for numerical solution of the non-analytical
equations for the two binding and the global binding model, were introduced to the fitting function of NONLIN (Johnson and Frasier, 1985). The flexibility of NONLIN, compared to other commonly used commercial software, was essential for achieving the global data fitting.

The globally analyzed data were obtained by Han Xu, a former graduate student in the lab, because she obtained fluorescence titrations in addition to nitrocellulose filter data that were obtained in this thesis. A summary of the experimental conditions as well as the fitting functions are summarized in Table 2.1.
<table>
<thead>
<tr>
<th>[P&lt;sub&gt;T&lt;/sub&gt;]&lt;sup&gt;1&lt;/sup&gt;</th>
<th>[D&lt;sub&gt;T&lt;/sub&gt;]</th>
<th>[G&lt;sub&gt;T&lt;/sub&gt;]</th>
<th>Method&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Observed Signal&lt;sup&gt;3&lt;/sup&gt;</th>
<th>Model&lt;sup&gt;4&lt;/sup&gt;</th>
<th>Fitting Function&lt;sup&gt;5&lt;/sup&gt;</th>
</tr>
</thead>
</table>
| varied  | 2x10<sup>-12</sup> | –           | NC          | Fractional DNA saturation | 1:1      | \[
\frac{[PD]}{[D_T]} + \frac{[PDG]}{[PG] + 2[PG_2]} + [PDG_2]
\] |
| varied<sup>6</sup> | 2x10<sup>-12</sup> | 1x10<sup>-5</sup> | NC          | Fractional DNA saturation | global   | \[
\frac{[PD]}{[D_T]} + \frac{[PG]}{[PG] + 2[PG_2]}
\] |
| varied<sup>7</sup> | –           | 3x10<sup>-8</sup> | NC          | Fractional guanine saturation | 1:2      | \[
\frac{[G_T]}{[PG] + 2[PG_2]} + [G_T] + 2[PDG_2]
\] |
| 1x10<sup>-7</sup> | –           | varied       | FL          | Fractional protein saturation | 1:2      | \[
\frac{[PG]}{2[P_T]} + \frac{[PD]}{[PG] + 2[PG_2]} + [PDG] + 2[PDG_2]
\] |
| varied<sup>9</sup> | ≥ 10<sup>-6</sup> | 1x10<sup>-5</sup> | NC          | Fractional guanine saturation | global   | \[
\frac{[PG]}{2[P_T]} + \frac{G_T}{[PG] + 2[PG_2]} + [PDG] + 2[PDG_2]
\] |
| 1x10<sup>-7</sup> | 2x10<sup>-7</sup> | varied       | FL          | Fractional protein saturation | global   | \[
\frac{[PG]}{2[P_T]} + \frac{G_T}{[PG] + 2[PG_2]} + [PDG] + 2[PDG_2]
\] |

**Table 2.1: Conditions for globally analyzed data**

Data for the global analysis were obtained from Hax Xu. Initially, data were analyzed separately to obtain good starting parameter values for K<sub>M1,M2,N1</sub> and N<sub>2</sub> for the subsequent global fit. Protein concentration is in dimers. NC: Nitrocellulose filter binding, FL: Fluorescence signal change. Models are described in detail in Section 2.1.1: 1:1, single binding site; 1:2, double site binding site; global, global thermodynamic cycle. Observed signal is the Y dependent variable for the fitting function. Calculated fitting functions used in NONLIN. For the 1:1 model the closed form solution was used. For the other models, the fitting function was calculated according to the solutions as derived from the numerical solver. Titrations performed under saturating guanine concentrations and non-stoichiometric conditions for DNA. Therefore, PDG<sub>2</sub> is most likely the major contributing factor to the observed signal. Titrations performed under non-stoichiometric conditions for guanine. Therefore, PG is most likely the dominant factor to the observed signal, unless there is high positive cooperativity from the second binding site. The fitting model assumes that the fluorescence signal change is independent of the guanine binding site. Titrations performed under saturating DNA concentrations and non-stoichiometric conditions for guanine. Therefore, PDG is most likely the major contributing factor to the observed signal, unless the second guanine binding site exhibits high positive cooperativity.
2.2.2 Ion concentration and temperature dependence of DNA binding in PurR and LacI proteins

Materials

Purification protocols followed those described for wild type, -11 aa (Chen and Matthews, 1994), and H74W LacI proteins Barry and Matthews (1999a). PurR was provided by Han Xu — a former graduate student in the lab — purified according to Xu (2002). DNA binding activity for each of these proteins was determined by specific operator binding under stoichiometric conditions. The activity level measured was greater than 90% for all proteins examined.

The target DNA sequence used for PurR DNA binding experiments was the same 30 bp sequence described in Section 2.2.1. Similarly, the target sequence for LacI was the 40 bp sequence containing the O1 site. Hybridization and labelling procedures were performed with the protocol described in Section 2.2.1.

DNA binding conditions

The ion concentration dependence of DNA binding was monitored by measuring the $K_D$ under variable ionic conditions. Variable concentration of KCl was used in a binding buffer containing 10 mM Tris-HCl, pH 7.6; 5%(v/v) dimethylsulfoxide, and 0.1 mg/mL bovine serum albumin. Binding isotherms were obtained by titrating a constant amount of labeled dsDNA sequence under non-stoichiometric conditions. To ensure nonstoichiometric conditions the labeled DNA was kept below $10^{-12}$ M for high-affinity and $10^{-11}$ M for low-affinity modes, respectively. Under these conditions, both PurR and LacI proteins bind one DNA molecule at one DNA binding site, so we monitored one binding event as dimer. Similarly, for the tetrameric lactose proteins examined (H74W and wild type LacI), the conditions monitor the association of binding of one DNA molecule to a tetrameric protein. In contrast, for -11aa LacI,
DNA binding occurs with the protein being in the dimeric form.

When binding in the high-affinity DNA binding mode for PurR (presence of guanine), holorepressor binding to DNA was monitored using $4 \times 10^{-6}$ M guanine concentration to ensure that protein is saturated with guanine. Similarly, for the high affinity DNA binding mode of the LacI-ONPF complex, or for the low affinity DNA binding mode of LacI with IPTG inducer, $2 \times 10^{-3}$ M ligand was used. Reactions were incubated for 1.5 hrs and 45 min for PurR or LacI proteins, respectively. Subsequent filtering of the reaction mixtures on 0.45 μm nitrocellulose filters was followed by exposing the dry filters on the phosphoimager (Fuji). Visualization and quantitation were conducted by the program MacBas v2 (Fuji).

For temperature-dependence experiments, the binding buffer with 250 mM KCl was used. Reactions were incubated in a controlled temperature waterbath or on ice. Specific consideration (reactions were set up in separate Eppendorf tubes rather than the 96-well plate commonly used for these assays) was taken so that filtration was accomplished rapidly to minimize any temperature fluctuation that might occur due to heat transfer.

**Data analysis**

Since all of the DNA binding experiments were performed under nonstoichiometric conditions, DNA binding data (radioactive counts) were fit to Equation 2.5 to determine the $\log K_D$. Of note, the term $[P_T]$ was evaluated taking into account the oligomeric state of the examined proteins (dimeric for purine and -11aa LacI proteins; tetrameric for wild type and H74W LacI proteins). For illustration purposes, the DNA binding data were transformed to fractional saturation, $R$, and plotted against $\log[P_T]$. At least three independent experiments were conducted for each ion concentration or temperature condition to allow standard error analysis.

The ion concentration dependence of DNA binding and its correlation to the
number of positively charged protein side-chains that are involved in neutralizing negative charges on the DNA backbone were analyzed with the simplified Equation 2.34. On the other hand, the temperature dependence of DNA binding characterized with the thermodynamic parameters $\Delta C_p$, $T_H$ and $T_S$, was analyzed by fitting the van’t Hoff plot data to the following equation, modified from equation 2.35:

$$-\log K_D = \frac{\Delta C_p}{2.303R} \left( \left( \frac{T_H}{T} \right) - \ln \left( \frac{T_S}{T} \right) - 1 \right)$$

(2.38)

Implementation of the described equations for data analysis and curve fitting was done on the program IgorPro (Wavemetrics, CA).

### 2.2.3 Biochemical characterization of $\Delta 33$ C-terminal mutant of human p53

**Materials**

The $\Delta 33$ p53 mutant was constructed by introducing a stop codon corresponding to position 361 by PCR mutagenesis of the p53 gene, cloned into pRSET (ampicillin resistance). NdeI (5'-end) and XhoI (3'-end) were employed to clone the product into pET-28(b), which conferred kanamycin resistance. Changing antibiotic resistance is advantageous for selecting colonies. Wild type and S392E p53 proteins were kindly provided by Dr Nicole Nichols or purified according to the established protocols (Nichols and Matthews, 2001, 2002).

The expression and purification of $\Delta 33$ p53 was conducted according to the following protocol: BL21DE3 E. coli cells were transformed with $\Delta 33$ p53-containing pET-28(b) plasmid and plated on LB media containing 50 $\mu$g/mL kanamycin. Individual colonies resulting from overnight incubation were selected and inoculated in 12 x 1L flasks containing 2 x YT media with 50 $\mu$g/mL kanamycin at 37°C. When cell density reached $A_{600}$ of 0.6 to 0.8, cells were induced with 1 mM IPTG and grown
for another 3 hours at 30 °C. After harvesting, the cells were resuspended in lysis buffer (100 mM Tris-HCl, pH 7.6, 200 mM KCl, 1 mM DTT, 5% (v/v) glycerol, 5% (w/v) glucose, 1 mM PMSF and 1 mg/mL lysozyme), and frozen at -20 °C overnight. After thawing and addition of DNAase in the presence of 1 mM MgCl₂, the cell lysate was collected by centrifugation. The cell lysate was brought to 40% (w/v) ammonium sulfate saturation, and the precipitate was resuspended in HFH buffer (20%(w/v) ammonium sulfate, 150 mM potassium phosphate, pH 7.6, 5% glycerol and 5% glucose and 1 mM DTT) and incubated for 6 hrs. The supernatant was collected and applied on a 1 mL butyl hydrophobic column connected to a BIOCAD system. Elution was performed with a gradient of buffers HFH and HFL (20 mM potassium phosphate pH 7.6, 5% glycerol, and 5% glucose and 1 mM DTT), and Δ33 p53 eluted during the final fourth of this gradient. Fractions were collected and dialyzed three times for 2 hrs in HPL buffer (20 mM potassium phosphate, pH 7.6, 5%(v/v) glycerol and 5%(w/v) glucose and 1 mM DTT) and applied on a 5 mL heparin column. Δ33 p53 was eluted at the first half of a HPL-HPH (500 mM potassium phosphate, pH 7.6, 5%(v/v) glycerol, 1 mM DTT, and 5%(w/v) glucose) gradient. Fractions were examined by SDS-PAGE and monitored by silver stain and Western analysis with pAb-3 antibody (Oncogene, CA). Protein activity was in the range of 30 to 60% for various preparations.

For DNA binding experiments, two 20 bp DNA sequences were used. For specific DNA binding the two single stranded sequences 5'-AGACATGCTAGGACATGCT-3' and 5'-AGGCATGTCTAGGCGATGCT-3' were hybridized to form a 20 bp double stranded specific DNA target sequence that contains two repeats of the 10-bp consensus sequence. The following DNA strands 5'-TCCAGATGTAACCAACACGTTG-3' and 5'-CACGTTTGGTTACATCTGGA-3' were hybridized to form a 20 bp double stranded non-specific DNA target. This sequence was designed to contain similar base composition as the specific target, but with the least frequent nucleotides used
at each particular position (Nichols and Matthews, 2001). Hybridization and labelling protocol were as described in Section 2.2.1.

**Thermal denaturation as monitored by circular dichroism**

Circular dichroism spectra were collected on an AVIV model 62A DS circular dichroism spectrometer, and the software Star 3.0 Stationary was employed to visualize the spectra. To avoid interference by DTT, 150 μL of protein was diluted into 250 μL buffer containing 50 mM potassium phosphate, yielding 5 \times 10^{-6} \text{ M} tetrmeric protein. Individual circular dichroism spectra were collected in a 0.2 cm quartz cuvette with wavelengths ranging from 200 to 255 nM in a 0.5 nM wavelength step with 8 sec per point averaging. Scans were corrected for baseline, and data were expressed as the difference of the extinction coefficients of left and right polarized light:

\[
\Delta \epsilon = \epsilon(L) - \epsilon(R) = \frac{\Delta A}{c b}
\]

(2.39)

where \( c \) and \( b \) are the tetrmeric molar protein concentration and length of optical path, respectively.

Temperature denaturation was conducted by collecting circular dichroism data for every 10 °C ranging from 0 to 90 °C. Incubation at every temperature was 10 min before scans were collected. Throughout the experiments the cuvette was covered to be airtight to avoid buffer evaporation at the elevated temperatures. At the end of the experiment there was no visible indication of protein aggregation; however, data collected by decreasing the temperature did not yield the same circular dichroism spectra, suggesting that the temperature denaturation process was irreversible.
Protein assembly

For glutaraldehyde crosslinking experiments, 2 μL of protein was diluted into 8 μL of buffer containing 50 mM potassium phosphate, 5% (v/v) glycerol and 5% (w/v) glucose, to yield a protein concentration of 3 x 10^{-6} M tetramers. Final glutaraldehyde concentration was at 0.01% or 0.1% (v/v). After an incubation period of 10 or 30 min, samples were boiled in a 2 X SDS treatment buffer, and the protein bands were visualized by silver staining of a 5% SDS denaturing acrylamide gel.

DNA binding conditions

For DNA binding affinities, labelled specific and non-specific target DNA sequences were used at \( \sim 1 \times 10^{-11} \) M, to ensure non-stoichiometric conditions. Indeed, the strongest affinity as expressed in \( K_{ds} \) was at least 30-fold higher than then labelled DNA concentration. In every binding condition, at least three independent measurements were performed to allow for statistical error determination.

DNA binding in potassium phosphate buffer

A buffer containing 50 mM potassium phosphate pH 7.5, 1 mM DTT, 5% (v/v) glycerol, and 0.1 g/L BSA was used to determine the specific and non-specific DNA binding affinities for Δ33 p53. This buffer was also used for examining the temperature and osmolar dependence of DNA binding of p53 proteins (see below). Previously, analysis of binding affinities of wild type and S392E p53 proteins was conducted in a potassium phosphate based buffer (Nichols and Matthews, 2001, 2002).

Ion concentration dependence

Ion concentration dependence of DNA binding was monitored by using a Tris-KCl based buffer (10 mM Tris-HCl (pH 7.5), variable concentration of KCl, 5% (v/v) glycerol, and 0.5 mM DTT with the addition of 0.1 mg/mL BSA). The affinities of
DNA binding, as expressed by $K_D$, were measured within the 50 to 150 mM KCl range. Specific considerations and data analysis are explained in Section 2.2.2.

**Temperature dependence of DNA binding**

Temperature dependence of DNA binding was monitored by determining the affinities of Δ33 p53 protein at temperatures ranging between 0 and 37 °C. DNA binding experiments for this purpose were conducted in the 50 mM potassium phosphate buffer. Specific considerations and data analysis are explained in Section 2.2.2.

**Water release**

Osmolal dependence of DNA binding was used for estimating the number of water molecules released upon DNA binding. More specifically, DNA binding affinities were determined in a 50 mM potassium phosphate buffer with glycerol concentrations of 0, 5, 10 and 15% (v/v) as the mediator of osmolal concentration. Given (v/v) concentrations was converted to osmolality ($m$) with the following procedure: Data of standard aqueous glycerol solutions were plotted as osmolality versus concentration (v/v). Then the data can be approximated with the following binomial equation:

$$m = 0.055 + 0.089C + 0.0028C^2$$  \hspace{1cm} (2.40)

where $C$ denotes the (v/v) glycerol concentration. Using this function, the (v/v) glycerol concentrations (5, 10 and 15 %) were transformed to osmolal concentrations.

According to Equation 2.36, the slope of a plot of DNA binding affinity as a function of the osmolality depends linearly on the number of water molecules displaced. Therefore, this number can be derived by a linear fit of the dependence of the $-\log K_D$ on the osmolality ($m$). DNA binding affinity dependence on the osmolal concentration and estimation of the number of water molecules released were obtained.
by a linear fit of $-\log K_D$ as a function of the osmolal concentration according to Equation 2.36.

**Data Analysis**

All of the binding affinity analyses were conducted under non-stoichiometric binding conditions. At least three independent measurements were performed to evaluate appropriate statistical error estimation. DNA binding data were not fit to a 1:1 DNA binding model, but to the Hill empirical model (Equation 2.7) because the former model was not appropriate for analysis of all binding reactions. The protein concentration used for the data analysis was expressed in tetrameric p53 protein.
Chapter 3

Thermodynamic parameters for ligand binding by PurR—Comparison to LacI

3.1 Introduction

Understanding the thermodynamic framework for protein-DNA interactions is the underlying step to correlate those parameters with biological function of transcription factors \textit{in vivo}. This concept is applied to PurR, the key regulator of the \textit{de novo} purine biosynthetic genes. Previous studies have elucidated the fashion in which PurR exerts its function: corepressor (hypoxanthine or guanine) binds to the PurR aporepressor and increases its DNA binding affinity (Choi \textit{et al.}, 1994; Choi and Zalkin, 1992; Rolfes and Zalkin, 1990b). However, thermodynamic characterization of ligand binding (guanine and DNA) to PurR is far from complete.

Simultaneous binding of guanine and DNA can be described with a closed thermodynamic cycle (see Figure 2.1). In this chapter, the thermodynamic parameters for each binding event are determined. Ligand binding is also monitored from a kinetic perspective to further dissect the origins of the observed thermodynamic equilibrium parameters. For equilibrium measurements, two approaches of analysis are followed. In the first, the equilibrium constants for each ligand binding event are determined independently. In the second, the entire set of thermodynamic data is
analyzed according to a global thermodynamic cycle model.

Binding reactions were performed in a high concentration potassium glutamate buffer, which was proposed to resemble intracellular ionic conditions (Leirmo et al., 1987). This buffer system also confers corepressor sensitivity to PurR (Rolfes and Zalkin, 1990b). Given the homology of PurR to LacI, measurement of the equilibrium and kinetic parameters for LacI-O$^1$ complex under the same conditions allows comparison between these two structurally similar proteins.

3.2 Results

3.2.1 Guanine binding to PurR

The affinities of the apo- and the holo- PurR repressor for guanine were measured under non-stoichiometric conditions by using nitrocellulose filter binding. Representative titrations for the aporepressor (absence of guanine) and the holorepressor (presence of excess of DNA) are shown in Figure 3.1. Holorepressor PurR binds guanine tighter than aporepressor PurR, as indicated by the shift of the holorepressor binding isotherm to lower protein concentrations. A summary of the binding equilibrium constants ($K_D$) for the apo- and holorepressor PurR is presented in Table 3.1. The presence of DNA increases the corepressor affinity of PurR by 15-fold.

Table 3.1: Thermodynamic parameters for guanine binding to PurR

<table>
<thead>
<tr>
<th>Binding mode</th>
<th>$K_D$ M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence of DNA</td>
<td>2.8±1.8x10^{-6}</td>
</tr>
<tr>
<td>Presence of DNA</td>
<td>1.8±0.2x10^{-7}</td>
</tr>
</tbody>
</table>

$K_D$ values are reported in monomers
Figure 3.1: Guanine binding affinities to PurR

Guanine equilibrium binding of PurR by nitrocellulose filter retention was conducted in the presence (A) and the absence of DNA (B). Solid lines represent best fit of the data to Equation 2.7.
3.2.2 DNA binding to PurR

The DNA binding affinity of apo- and holo- PurR (absence and presence of guanine, respectively) was measured under non-stoichiometric conditions. A series of representative titrations for the aporepressor and the holoressor PurR is shown in Figure 3.2. Holoressor PurR binds operator purF tighter than aporepressor PurR, as indicated by the shift of the holoressor binding isotherm to lower protein concentrations. Quantitative results are shown in Table 3.2. The holoressor binds operator DNA with a \(~1000\)-fold higher affinity than free PurR does.

![Graph showing DNA fractional saturation vs log[PurR]](image)

**Figure 3.2: Equilibrium DNA binding of PurR**

Equilibrium DNA binding of PurR was conducted in the presence (■) and absence of guanine (●). Solid lines represent best fits to Equation 2.5.
Table 3.2: Thermodynamic parameters for operator binding to PurR

<table>
<thead>
<tr>
<th>Binding mode</th>
<th>$K_D$ M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence of guanine</td>
<td>$1.1\pm0.15\times10^{-7}$</td>
</tr>
<tr>
<td>Presence of guanine</td>
<td>$2.3\pm0.2\times10^{-10}$</td>
</tr>
</tbody>
</table>

$K_D$ values are reported in dimers

3.2.3 Operator dissociation kinetics of PurR

The DNA dissociation kinetics of apo- and holorepressor were measured by nitrocellulose filter binding. To ensure pseudo-first order conditions, an excess of unlabeled DNA was added, and the timepoint was set to 0 sec. Under the pseudo-first conditions employed for this study, the dissociation kinetics can be monitored over time as a zero-order complex dissociation, as described by Equation 2.31. Under these conditions, the DNA dissociation kinetics of the apo- and holorepressor are presented in Figure 3.3 and indicate that the protein-DNA complex is more kinetically labile in the absence of guanine. Aporepressor PurR-DNA complex releases DNA about 27-fold faster than the holorepressor complex (see Table 3.3).

Table 3.3: DNA binding kinetics for PurR

<table>
<thead>
<tr>
<th>Kinetic constant</th>
<th>Binding mode</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_d^\dagger$</td>
<td>Absence of guanine</td>
<td>$0.06\pm0.02$ s$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>Presence of guanine</td>
<td>$0.0022\pm0.0003$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_a^*$</td>
<td>Absence of guanine</td>
<td>$6\times10^5$ M$^{-1}$s$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>Presence of guanine</td>
<td>$1\times10^7$ M$^{-1}$s$^{-1}$</td>
</tr>
</tbody>
</table>

$\dagger$ Determined experimentally. * Calculated from $k_a=k_d/K_D$. 
Figure 3.3: PurR DNA dissociation kinetics

DNA dissociation kinetics of PurR in the absence (A) and presence of guanine (B). Dissociation rate in the absence of guanine is 27-fold faster than in the presence of guanine as indicated by the differences in the scaling of the time axis. Lines represent the best fit of the data to Equation 2.31.
3.2.4 Global equilibrium analysis of PurR

PurR binding to corepressor and DNA can be described by the thermodynamic loop illustrated in Figure 2.1. In this analysis, the macroscopic equilibrium constants \( K, M_1, M_2, N_1 \) and \( N_2 \) were fitted globally to titration data obtained by two different techniques, as summarized in Table 2.1. Estimation of these parameters by global fitting suffices to describe the system thermodynamically, since the remaining equilibria can be derived as a combination of the depicted thermodynamic parameters.

Initially, the binding data were analyzed independently to obtain good starting points for the global data fitting. Global data fitting converged to the parameters summarized in Table 3.4. Convergence to the estimated values was independent of a reasonable choice of initial values, indicating that convergence was global, rather than obtaining a local minimum. Comparison with the previously reported values by Xu (2002), as derived by fitting to a 1:1 or Hill empirical model are in good accordence, indicating the efficacy of the global fit. Of note, global fitting provided estimation for the affinity of guanine binding to the second binding site, both in the presence or absence of DNA (\( M_2 \) and \( N_2 \), respectively).

The goodness of the global fit was checked by comparing the simulated binding isotherms with the experimental data in Figure 3.4. In general, the binding isotherms correlated well with the experimental data, with the exception for the fluorescence guanine titration in the presence of saturating concentrations of DNA (Graph B of Figure 3.4). The origin of the discrepancy can be attributed to a variety of causes. First, global fitting takes into account all binding data and therefore precise estimation of the total protein, DNA, or guanine concentrations is paramount. Another source for the discrepancy could be the inherent assumption of the global fit model when constructing the fitting function for the fluorescent titrations; the change
Table 3.4: Thermodynamic parameters derived from global data fitting

<table>
<thead>
<tr>
<th>Eq. constants(^1)</th>
<th>Estim. constant(^2)</th>
<th>Confidence levels(^3)</th>
<th>Eq. constants(^4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(M^{-1})</td>
<td>(M^{-1})</td>
<td>(M^{-1})</td>
</tr>
<tr>
<td>(K)</td>
<td>4.7(\times)10(^7)</td>
<td>3.5 to 5.6(\times)10(^7)</td>
<td>2.5(\times)10(^7)</td>
</tr>
<tr>
<td>(M_1)</td>
<td>3.2(\times)10(^5)</td>
<td>2.3 to 3.5(\times)10(^5)</td>
<td>3(\times)10(^5)</td>
</tr>
<tr>
<td>(M_2)</td>
<td>8.0(\times)10(^4)</td>
<td>6.0 to 10.1(\times)10(^4)</td>
<td>ND</td>
</tr>
<tr>
<td>(N_1)</td>
<td>7.1(\times)10(^6)</td>
<td>6.0 to 8.3(\times)10(^6)</td>
<td>5(\times)10(^6)</td>
</tr>
<tr>
<td>(N_2)</td>
<td>5.0(\times)10(^5)</td>
<td>3.7 to 6.9(\times)10(^5)</td>
<td>ND</td>
</tr>
</tbody>
</table>

\(^1\) Equilibrium constants describing the PurR thermodynamic cycle (see Figure 2.1).  
\(^2\) Estimated association equilibrium constants from global data fitting of data from Xu (2002), which include fluorescence titrations for determining \(M_2\) and \(N_2\).  
\(^3\) Values presented are 67% or 1 standard deviation. Of note, because of the nonlinear behavior of the fit, opposite confidence levels might vary.  
\(^4\) Association equilibrium constants were taken from Xu (2002). Monomeric guanine equilibrium constants were converted to dimeric ones, as the global model presented in Figure 2.1 incorporates the dimeric form of the PurR for ligand binding reactions. ND: these binding constants cannot be determined, since the binding titrations provide no information about binding of the second binding site.
in the fluorescence signal upon binding of guanine to the binding sites is equivalent. The latter explanation is more plausible, since comparison of the thermodynamic parameters for guanine binding by analyzing the individual data also demonstrated some discrepancy.

Figure 3.4: Global data fitting

Lines represent best global data fit. (A) DNA binding in the presence and absence of guanine as derived from nitrocellulose filter binding. (B) Guanine binding in the presence and absence of DNA as derived from change of fluorescence signal. (C) Guanine binding in the presence and absence of DNA as derived from nitrocellulose filter binding.
Figure 3.4: Global data fitting (continued)
The PurR thermodynamic cycle can also be described by the thermodynamic parameters as expressed by the microscopic equilibrium constants (see Section 2.1.1 and Appendix 7.2). Since global data fitting estimated the affinities of the second guanine binding site, valuable information about the guanine binding site cooperativity can be derived. Such information is summarized in Table 3.5. Binding of the guanine to the aporepressor follows a site independent model, since binding of the first guanine did not have any effect on binding of the second guanine (s=1). However, binding of first guanine to the PurR-DNA complex exhibited a 22-fold enhanced intrinsic affinity compared to free protein. Interestingly, binding of the second guanine exhibits a 4-fold decreased value compared to the first binding site.

**Table 3.5: Microscopic description of PurR thermodynamic cycle**

<table>
<thead>
<tr>
<th>Constant</th>
<th>Description</th>
<th>Est. Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>DNA binding no guanine</td>
<td>$4.7 \times 10^7 \text{ M}^{-1}$</td>
</tr>
<tr>
<td>$m$</td>
<td>Guanine binding no DNA</td>
<td>$1.6 \times 10^5 \text{ M}^{-1}$</td>
</tr>
<tr>
<td>$s$</td>
<td>Guanine binding cooperativity for 1st site, no DNA</td>
<td>1</td>
</tr>
<tr>
<td>$s_1$</td>
<td>1st guanine binding site, ± DNA,</td>
<td>22</td>
</tr>
<tr>
<td>$s_2$</td>
<td>2nd guanine binding site, ± DNA</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Data are obtained from Table 3.4 by expressing the macroscopic equilibrium constants as microscopic parameters. See Section 2.1.1 and Appendix 7.2 for discussion.

Based on the derived thermodynamic parameters from global data fitting, the species distribution over concentration space can be simulated. The contributing species to the simulated DNA binding titration with $[D_T]$ at $1 \times 10^{-12}$ M (typical condition for non-stoichiometric experiments) over protein and guanine space are illustrated (Figure 3.5). At low guanine concentrations PD dominates the fraction of DNA bound. As guanine concentration increases, PDG becomes the most dominant
species, with an subsequent increase in DNA binding affinity as illustrated by the shift of the contour isotherms to lower protein concentrations. At higher guanine concentrations PDG is replaced by PDG$_2$ and a further increase in DNA binding affinity. The slight negative cooperativity of binding to the second site is illustrated by the fact that PDG levels contribute more than 50% of the combined signal. Finally, this simulation illustrates that binding titrations using 1 x 10$^{-5}$ M guanine reflect predominantly PDG$_2$ binding, validating the assumption that PurR is saturated with guanine under these conditions.

A. PD

![Graph](image)

**Figure 3.5: Species distribution**

DNA binding species distribution for PurR is simulated by using the thermodynamic parameters derived from Table 3.4. DNA binding titrations are simulated at 1 x 10$^{-12}$ M DNA with protein and guanine ranging from 10$^{-12}$ to 10$^{-4}$ dimers and 10$^{-9}$ to 10$^{-2}$ M, respectively. Blue, cyan and red curves at the bottom of the graphs represent binding isotherms at 10, 50 and 90% of DNA bound saturation, respectively. DNA binding species PD, PDG, PDG$_2$ (see Figure 2.1) are represented by graphs A, B and C, respectively.
Figure 3.5: Species distribution (continued)
3.2.5 Thermodynamic and dissociation kinetic parameters of LacI DNA binding

The affinity of LacI to $O^1$ was measured under non-stoichiometric conditions using the nitrocellulose filter binding method. The equilibrium titration is shown in Figure 3.6. Interestingly, under these conditions, the affinity of LacI for its operator is $\sim$40-fold weaker than the holorepressor PurR (presence of guanine), but $\sim$60-fold tighter than the aporepressor PurR. Therefore, LacI binds DNA with an intermediate affinity as compared to PurR. A summary of the results is presented in Table 3.6.

The dissociation kinetics of DNA from LacI-$O^1$ complex were measured by quenching an equilibrium complex of LacI with radiolabelled DNA by addition of a large excess of unlabelled DNA. Using the zero-order kinetic conditions, analysis of the data was conducted by using the Equation 2.31, as shown in Figure 3.7. The dissociation of the LacI-DNA complex is $\sim$20-fold faster than that of the holorepressor PurR, but only $\sim$1.5-fold slower than aporepressor. Therefore, the intermediate behavior of LacI versus the high and low affinity modes of PurR derives from effects on both association and dissociation kinetics.

**Table 3.6: Summary of equilibrium and kinetic constants for LacI**

<table>
<thead>
<tr>
<th>Constant</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_D^+$</td>
<td>$7.2 \pm 1.9 \times 10^{-9}$ M</td>
</tr>
<tr>
<td>$k_d^+$</td>
<td>$0.04 \pm 0.02$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_a^*$</td>
<td>$6 \times 10^6$ M$^{-1}$s$^{-1}$</td>
</tr>
</tbody>
</table>

$^+$ Determined experimentally. $^*$ Calculated from $k_a = k_d/K_D$. 
Figure 3.6: DNA binding affinity of LacI

DNA binding of LacI conducted under the same conditions as for PurR. The solid line represents the best fit to the data.
Figure 3.7: DNA binding dissociation kinetics of LacI

DNA dissociation kinetics of equilibrium LacI-DNA complex, monitored by quenching with excess of unlabelled DNA. The solid line represents the best fit to the data.
3.3 Discussion

3.3.1 Thermodynamic parameters for PurR DNA binding

In order to describe the function of biological systems, it is imperative to determine carefully the thermodynamic parameters that govern the behavior of their components. For the case of purine repressor, limited information for its thermodynamic properties was available. In the present study, the thermodynamics of the PurR protein-DNA and protein-ligand interactions were determined in detail in two ways: (i) by independent analysis of individual binding data obtained for DNA and guanine binding by nitrocellulose filter binding and (ii) by global analysis of binding data according to a thermodynamic cycle model. These analyses were consistent with each other.

Global data analysis provided additional information about guanine binding to the second binding site within a dimer. This information cannot be determined by analysis of the individual measurements. Most of the increase in the guanine affinity in the presence of DNA results from the increased intrinsic affinity from the first binding event. Binding of the second guanine molecule to the free PurR exhibits the same intrinsic affinity as binding of the first guanine to the free PurR. In contrast, binding of the second binding molecule to the PurR-DNA complex is accompanied by a relative decrease in the intrinsic guanine binding affinity of the second binding site relative to the first. This observation can be rationalized by taking into account the additional steric restraints imposed by the DNA bound to PurR, compared to the free PurR, where the N-terminal domains are free and flexible to accommodate any conformational changes in the core domain tertiary structure elicited by binding of the guanine molecules. Still, compared to the free PurR, the second guanine binding guanine event for PurR-DNA is tighter. The increased intrinsic affinities of both
guanine binding events work synergistically to increase the overall binding affinity of guanine to the PurR-DNA complex.

Comparison of the DNA binding affinities of PurR in the presence and absence of guanine provides insight on the thermodynamic mechanisms of PurR function. More specifically, the presence of guanine enables PurR to bind operator DNA with higher affinity, thereby shutting down transcription of genes involved in the biosynthesis of purines. Moreover, the protein-DNA complex exhibits higher binding affinity for guanine, increasing further the stability of the holorepressor-DNA complex and ensuring the effective repression of the purine biosynthetic genes. In this way, the DNA-bound protein can bind guanine at lower cellular levels than the free purine repressor. This arrangement enhances the ability of the protein to regulate repression efficiently. Only when the purine levels are depleted, the purine repressor dissociates from its operator sequence, allowing transcription of the purine biosynthetic genes to occur.

3.3.2 Kinetic parameters for PurR DNA binding

Determination of the kinetic parameters of PurR-DNA interactions can provide details on how transcription repression is mediated. Comparison of the kinetic parameters (Table 3.3) shows that the increased DNA binding affinity of the holorepressor — a crucial property for transcription repression — is achieved in two ways. First, the holorepressor binds the operator sequence more rapidly. Secondly, after the holorepressor-DNA complex has formed, its dissociation rate is slowed down. Both mechanisms contribute to holorepressor binding operator sequences for a larger fraction of time.
3.3.3 Comparison of PurR and LacI

Structural studies and sequence comparison have shown high homology between PurR and LacI. Determination of the binding and kinetic parameters for LacI under the same conditions as PurR allows direct comparison of their thermodynamic behavior. Interestingly, despite their sequence homology, LacI and PurR exhibit significant differences in their thermodynamic and kinetic behaviors. Under the conditions examined, LacI exhibits a 40-fold lower affinity for its operator sequence than PurR holorepressor. Therefore, PurR exhibits tighter repression of gene transcription.

This difference can be rationalized by considering the general biological roles of the two proteins. LacI regulates transcription of enzymes that participate in lactose uptake and utilization as an energy source. This catabolic process comes with an overall positive energy contribution to the cell. On the other hand, PurR regulates transcription of multiple genes that express purine biosynthetic enzymes. Moreover, the biosynthetic process is accompanied by high energy expenditure. Therefore, tighter regulation allows the cells to utilize other purine and pyrimidine salvage pathways before the cells commit to the energetically costly biosynthetic pathways.

Comparison of the kinetic parameters for both purine and lactose repressors indicates that the tighter regulatory control for the former is a synergy of both association and dissociation kinetic components. Indeed, LacI-DNA association occurs more slowly than PurR holorepressor-DNA complex formation. Once the protein-DNA complexes have formed, LacI dissociates more rapidly than PurR holorepressor. Thus, the LacI-DNA complex is more kinetically labile than the PurR holorepressor-DNA complex.

This difference in the kinetics of the high affinity protein-DNA complexes for lactose and PurR reflects the distinct physiological roles of the proteins. LacI exerts its function depending on the “opportunistic” presence of lactose in the environment.
As a result, its optimal function is a transient production of lactose metabolic enzymes that requires the kinetic lability of the LacI-DNA complex. A basal level of lactose uptake enzymes (permease) must be present to detect lactose in the environment. To this effect, ligand binding kinetics for lactose repressor are faster than PurR (Xu, 2002). Finally, the slower association of free LacI to its operator allows complete utilization of lactose before transcription repression occurs.

In contrast, intracellular purine levels are more stable, since they are required for cellular sustenance. Purine levels are regulated by both biosynthetic and salvage pathways. Under normal conditions, the kinetic stability of the PurR-operator complex suppresses purine enzyme biosynthesis, allowing cells to utilize salvage of existing purines. The slow holorepressor-DNA dissociation, and consequently slow onset of purine biosynthetic enzyme synthesis, occurs only after intracellular purine levels have been depleted for a sustained period of time. Further, increase of purine intracellular concentration results in the rapid association of holorepressor-DNA complex, ensuring prompt repression of purine biosynthesis to avoid potentially high lethal levels (Levine and Taylor, 1982).
Chapter 4

Ion concentration and temperature dependence of DNA binding: Comparison of PurR and LacI proteins

4.1 Introduction

LacI was the major biological system employed in formulating the theoretical bases of both the "polyelectrolyte" and the "hydrophobic" effects (Ha et al., 1989; Record et al., 1991). LacI binding to specific DNA sequences (high affinity mode) is accompanied by a release of 6 to 10 cations, depending on the length of the DNA target sequence (Frank et al., 1997; Lohman et al., 1980; Mossing and Record, 1985; Record et al., 1977; Whitson and Matthews, 1986; Winter and von Hippel, 1981). On the other hand, LacI binding to non-specific DNA was accompanied by release of 12 cations (deHaseth et al., 1977b; Revzin and von Hippel, 1977). The polyelectrolyte effect depended on the DNA length, as longer sequences increased the "polyelectrolyte" effect, presumably by contribution of both specific and non-specific binding events (Mossing and Record, 1985; Whitson and Matthews, 1986). However, a later study showed a decrease in cation release when specific sequence binding was gradually transformed to non-specific (Frank et al., 1997). In addition, this gradual
change in DNA binding specificity resulted in a concomitant reduction in the heat capacity change (Frank et al., 1997). Specific DNA binding of LacI was accompanied by significant heat capacity change (-0.9 to -1.5 kcal/molK) that was dependent on the affinity of the protein-DNA interaction (Ha et al., 1989; Spolar and Record, 1994). The negative heat capacity change for specific DNA binding to operator DNA was attributed to the hydrophobic effect of the hinge helix domain folding in the presence of DNA (Frank et al., 1997).

Studies of ion concentration and temperature dependence of DNA binding for PurR were utilized to examine the cation release and the heat capacity change for PurR and compare the behavior to that for its LacI homologue. As presented in Chapter 3 of this thesis, PurR bound its operator ~30-fold more tightly than LacI bound to $O^1$ in a high ionic strength buffer, indicating that although LacI and PurR are highly homologous, they exhibit different ion concentration dependence behavior. The eventual goal of this study was to correlate the macroscopic observations to microscopic data at the atomic level. This correlation is feasible because of the availability of high resolution X-ray structures for both LacI and PurR (Bell and Lewis, 2000; Schumacher et al., 1997).

4.2 Results

4.2.1 Ion concentration dependence of DNA binding of PurR

To investigate the ion concentration dependence of PurR binding to DNA, the equilibrium dissociation constants ($K_D$) over a range of KCl concentrations were determined by nitrocellulose filter binding. Measurements to determine cation release associated with complex formation required establishing the linear range for $K_D$ dependence on ion concentration. The linear range for PurR dependence was identified as 200 to 400 mM KCl, and this region was utilized for detailed measurements to
determine cation release. A series of representative titrations for high affinity (presence of guanine) and low affinity (absence of guanine) DNA binding is presented in Figure 4.1. Under all conditions, DNA concentration was maintained well below the estimated $K_D$ and titrated with increasing PurR concentrations until a saturating plateau was reached. Guanine concentration was set so that PurR was saturated. Therefore, only the formation of the holorepressor (PurR-guanine) complex with DNA was monitored. For both binding modes, increasing the KCl concentration resulted in decreased DNA binding affinity.

In this linear range, the ion concentration dependence is dramatic for DNA binding of the holorepressor, as indicated by the steep negative slope, with a differential in the DNA binding affinities of $\sim$5000-fold between 200 and 400 mM KCl (Figure 4.2). Over the same range, the effect on aporepressor binding is less pronounced, with a differential of $\sim$50-fold. As a result of the differences in the ion concentration behavior, affinities for both DNA binding modes of PurR converge at $\sim$400 mM KCl.

4.2.2 Ion concentration dependence of DNA binding of LacI proteins

To compare the ion concentration dependence results for PurR-DNA binding to its homologue LacI, the DNA binding affinities for a series of LacI proteins were measured. In almost all cases, the linear region for LacI and its variants was 100 mM to 300 mM KCl, which was lower than the region established for PurR DNA binding (see Section 4.2.1). The exception was the case of wild type LacI low affinity binding in the presence of inducer (IPTG), where a lower range of 50 to 150 mM KCl was necessary to achieve measurable $K_D$ values. The ion concentration dependence for DNA binding in the presence of the anti-inducer ONPF, which has an effect reverse to that of inducer binding (Barkley et al., 1975; Riggs et al., 1970b), was
Figure 4.1: Representative DNA binding isotherms for purine repressor at various [KCl] concentrations

Individual binding isotherms were obtained in the presence (top) and in the absence (bottom) of guanine. The binding isotherms correspond to 200 (△), 250 (▽), 314 (○) and 400 (☆) mM KCl. The solid lines correspond to the best fit according to Equation 2.5.
Figure 4.2: Ion concentration dependence for DNA binding by PurR proteins

The ion concentration dependence of PurR DNA binding in the presence (●) and absence (○) of guanine was determined in the linear range from 200 to 400 mM KCl. The solid and dashed lines correspond to weighted linear fits of the DNA binding data.

also measured. These experiments were necessary to correlate eventually the number of cations released with specific protein-DNA contacts in the recently solved high resolution X-ray structure of the dimeric LacI-ONPF:O$_{sym}$ complex (Bell and Lewis, 2000). To provide a wider range of proteins for interpreting the results of these measurements, the ion concentration dependence of LacI H74W, a mutant deficient in inducer response (Barry and Matthews, 1999a), and a LacI dimeric mutant (-11aa LacI) that exhibits thermodynamic linkage of protein assembly (monomer association to dimers) and DNA binding (Chen and Matthews, 1994) were also examined. For all of these LacI proteins, increased KCl concentration resulted in decreased DNA binding affinity as indicated by the negative slopes in Figure 4.3. Interestingly, the largest change in DNA binding affinities with increasing salt concentrations from 100 to 300 mM KCl was observed for H74W LacI-DNA binding and the smallest for H74W LacI binding in the presence of IPTG. In general, the ratio of high affinity to
Figure 4.3: Ion concentration dependence for DNA binding by LacI proteins

The solid and dashed lines represent best weighted linear fits to DNA binding data for LacI proteins. (A) LacI binding in the presence (●) and absence (○) of IPTG. The latter data were from Whitson and Matthews (1986). (B) H74W LacI binding in the presence (●) and absence (○) of IPTG. (C) LacI in the presence of ONPF (●) and -11aa LacI (○).
low affinity binding also decreased as the ionic strength increased. The significant
effect of inducer on operator DNA binding by wild type LacI is illustrated by the
differential between the two sets of data shown in Figure 4.3A. This differential is
diminished for H74W LacI (Figure 4.3B), as anticipated based on the resistance of
this mutant protein to induction. For H74W the affinities almost converge at $\sim$ 300
mM KCl, a pattern that mirrors PurR behavior except for the opposite effects of
effector ligand on DNA binding affinity. The dimeric deletion protein, -11aa LacI,
exhibited the lowest affinity and the least ion concentration dependence for DNA
binding of LacI proteins (Figure 4.3C). Ion concentration dependence of LacI-ONPF
binding to DNA follows a similar pattern to wild type LacI (Figure 4.3C).

4.2.3 Cation release by PurR and LacI proteins

The number of cations released is determined from the negative slope of the
ion concentration dependence and quantified by Equation 2.33. Moreover, the ion
pairs formed upon protein-DNA interactions can be calculated using the simplified
Equation 2.34. For the PurR and LacI proteins examined, a summary of ion pairs
formed upon protein-DNA complex formation is presented in Table 4.1.

When PurR holorepressor binds DNA, $\sim$15 ion pairs are formed, more than
double the value for low affinity PurR aporepressor DNA binding ($\sim$6). The 2.5-fold
difference in the slope of the ion concentration dependence is illustrated by the the
intersection of the two binding modes at $\sim$400 mM KCl. On the other hand, high
affinity binding modes of LacI, both in the absence and presence of the anti-inducer
ONPF, demonstrated the formation of a similar number of ion pairs as observed for
the aporepressor PurR ($\sim$6). The exception is H74W LacI which exhibited slightly
higher ion pair formation ($\sim$8). However, in no case of high affinity LacI binding
mode was ion pair formation as great as that shown for PurR holorepressor. The
ion pairs formed for the low DNA binding affinity mode of LacI (presence of IPTG)
Table 4.1: Number of ion pairs formed upon DNA binding

<table>
<thead>
<tr>
<th>Protein complex</th>
<th>Number of ion pairs formed</th>
<th>Affinity binding mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>PurR-guanine·purF</td>
<td>14.9±0.8</td>
<td>High</td>
</tr>
<tr>
<td>LacI·O&lt;sup&gt;1&lt;/sup&gt;</td>
<td>6.4±0.2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>High</td>
</tr>
<tr>
<td>LacI·ONPF·O&lt;sup&gt;1&lt;/sup&gt;</td>
<td>6.4±0.1</td>
<td>High</td>
</tr>
<tr>
<td>H74W LacI·O&lt;sup&gt;1&lt;/sup&gt;</td>
<td>8.1±0.9</td>
<td>High</td>
</tr>
<tr>
<td>-11aa LacI·O&lt;sup&gt;1&lt;/sup&gt;</td>
<td>4.3±0.4</td>
<td>High</td>
</tr>
<tr>
<td>PurR·purF</td>
<td>5.9±0.7</td>
<td>Low</td>
</tr>
<tr>
<td>LacI·IPTG·O&lt;sup&gt;1&lt;/sup&gt;</td>
<td>3.1±0.4</td>
<td>Low</td>
</tr>
<tr>
<td>H74W LacI·IPTG·O&lt;sup&gt;1&lt;/sup&gt;</td>
<td>4.5±0.2</td>
<td>Low</td>
</tr>
</tbody>
</table>

The number of ion pairs formed was calculated from the linear slopes of the ion concentration dependence plots in Figures 4.2 and 4.3, by using the simplified equation Equation (2.34) on page 53. <sup>a</sup>Data were obtained from Whitson and Matthews (1986) and are consistent with Winter and von Hippel (1981) and Frank et al. (1997).
dropped to half of the number observed for the high DNA binding affinity mode (3 to 4 versus 6 to 8 ion pairs). Therefore, the binding of LacI proteins mirrors PurR behavior, albeit with a lower overall level of cation release. High affinity binding involves 2-fold greater cation release compared to low affinity binding. Interestingly, the inverse relationship between effector ligand response for PurR and LacI does not alter this relationship.

4.2.4 Temperature dependence of PurR and -11aa LacI proteins

The temperature dependence of PurR DNA binding in both high and low affinity modes was examined by incubating the binding reactions at temperatures ranging from 1 to 37 °C in 250 mM KCl. The temperature dependence of PurR DNA binding was analyzed by using a van't Hoff plot (− log \( K_D \) vs T\(^{-1} \)) as shown in Figure 4.4. At all temperatures, PurR holo repressor binds DNA with higher affinity than PurR aporepressor, with a maximum difference in the middle of the temperature range. However, at the ends of the temperature range the differential of the DNA binding is decreased. Thus, the van’t Hoff curve for the DNA binding of the holo repressor PurR exhibits larger curvature than binding of PurR alone does.

A summary of the calculated thermodynamic parameters describing the temperature dependence of the binding affinities is shown in Table 4.2. PurR high affinity DNA binding (in the presence of corepressor) is accompanied by a large negative heat capacity change (-2.8±0.5 kcal/molK), with a much smaller \( \Delta C_p \) observed in the absence of corepressor (-0.8±0.3 kcal/molK). LacI high affinity binding to \( O^1 \) in the absence of inducer yields a \( \Delta C_p \) of -0.9±0.1 kcal/molK (Ha et al., 1989; Spolar and Record, 1994; Whitson and Matthews, 1986). Measurements for LacI binding to \( O^{sym} \) by others have yielded \( \Delta C_p \) values of -1.3±0.3 and -1.5±0.2 kcal/molK (Frank et al., 1997; Ha et al., 1989). Therefore, the PurR high affinity mode exhibits a more
Figure 4.4: van't Hoff plot of PurR and -11aa LacI DNA binding

van't Hoff plots of temperature dependence of DNA binding are presented for -11aa LacI (♦) and PurR in the presence of guanine (●) and in the absence of guanine (○). Temperature ranges from 0 to 37 °C. For simplicity the affinities (− log $K_D$) are plotted against 1000/T instead of 1/T. The solid and dashed curves represent the best nonlinear weighted fits to Equation 2.35.
Table 4.2: Thermodynamic parameters for DNA binding by van’t Hoff analysis

<table>
<thead>
<tr>
<th>Protein-DNA complex</th>
<th>$\Delta C_p$ (kcal mol$^{-1}$ K$^{-1}$)</th>
<th>$T_H$ (K)</th>
<th>$T_S$ (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PurR-guanine-purF</td>
<td>-2.8±0.5</td>
<td>294±1</td>
<td>300±1</td>
</tr>
<tr>
<td>PurR-purF</td>
<td>-0.8±0.3</td>
<td>303±5</td>
<td>314±10</td>
</tr>
<tr>
<td>-11aa LacI-O$^1$</td>
<td>-0.9±0.3</td>
<td>291±2</td>
<td>303±4</td>
</tr>
<tr>
<td>LacI-O$^1$</td>
<td>-0.9±0.1$^a$</td>
<td>294±1$^a$</td>
<td>312±2$^a$</td>
</tr>
<tr>
<td>LacI-O$^{sym}$</td>
<td>-1.3±0.3$^b$</td>
<td>293±1$^b$</td>
<td>305±3$^b$</td>
</tr>
<tr>
<td>LacI-O$^{sym}$</td>
<td>-1.5±0.2$^c$</td>
<td>290$^c$</td>
<td>294$^c$</td>
</tr>
</tbody>
</table>

The above thermodynamic parameters are derived from van’t Hoff analysis of the temperature dependence of DNA binding according to Equation 2.35. $^a$ From Ha et al. (1989); Spolar and Record (1994); Whitson and Matthews (1986). $^b$ From Ha et al. (1989); Spolar and Record (1994). $^c$ From Frank et al. (1997).
negative $\Delta C_p$ than LacI.

To investigate whether the large negative $\Delta C_p$ observed for PurR results from protein assembly equilibria, the temperature dependence for DNA binding of -11aa LacI was analyzed. This protein exhibits thermodynamic linkage of assembly and DNA binding equilibria. The -11aa LacI protein yielded a value for $\Delta C_p$ of -0.9±0.3 kcal/molK, comparable to the values measured for wild type LacI (Ha et al., 1989; Spolar and Record, 1994). Therefore, coupled assembly and DNA binding is not necessarily reflected in the magnitude of the heat capacity change.

4.3 Discussion

4.3.1 Correlation of ion concentration dependence with structural data for PurR

The high resolution X-ray structure for PurR-guanine-purF (Schumacher et al., 1997) allows interpretation of the cation release data at the atomic level. The type of the ionic contacts can be classified into three categories, depending on the origin of the interaction:

- Helix-turn-helix (K5, H20 and R26) as shown in Figure 4.5
- Hinge helix (R52 and K55) as shown in Figure 4.6
- Core (R115) as shown in Figure 4.7A

High affinity PurR-guanine-DNA binding was accompanied by release of $\sim$15 cations, while low affinity binding in the absence of guanine released $\sim$6 cations. The electrostatic contacts for DNA binding of the PurR low affinity binding mode presumably originate from the three helix-turn-helix domain interactions per monomer,
Figure 4.5: PurR helix-turn-helix contacts to DNA

Contacts of basic amino acid side chains of K5 (top), H20 (middle) and R26 (bottom) of PurR. Each PurR monomer is shown in green and mauve (chains A and D, respectively). The indicated residues from chain A are sketched in a ball-and-stick model, encapsulated by a transparent surface. DNA is sketched as a stick model, with its backbone colored in yellow and red for P and O atoms, respectively. The ion-pair formation is illuminated. Coordinates are from PDB file 1WET (Schumacher et al., 1997).
**Figure 4.6: PurR hinge helix contacts to DNA**

Contacts of basic amino acid side chains of R52 (top) and K55 (bottom) of PurR. Each PurR monomer is shown in green and mauve (chains A and D, respectively). The indicated residues from chain A are sketched in a ball-and-stick model, encapsulated by a transparent surface. DNA is sketched as a stick model, with its backbone colored in yellow and red for P and O atoms, respectively. The ion-pair formation is illuminated. Coordinates are from PDB file 1WET (Schumacher et al., 1997).
Figure 4.7: PurR core domain ion pairs

(A) Contact of basic amino acid side chain of R115 from the core domain of PurR to DNA. (B) Inter-subunit ion pair formation between H44 and E136 of amino acid chains A and D, respectively. Each PurR monomer is shown in green and mauve (chains A and D, respectively). Selected residues from chain A are sketched as a ball-and-stick model, encapsulated by a transparent surface. DNA is sketched in a stick model, with its backbone colored in yellow and red (for P and O atoms, respectively). The ion-pair formation is illuminated. Coordinates are from PDB file 1WET (Schumacher et al., 1997).
specifically the contacts of K5, H20, and R26, for a total of six per dimer, as observed experimentally. NMR data of the free N-terminus of PurR, demonstrated that this domain assumes a fold similar to that found for this region in the holorepressor structure, except for the hinge helix region which was unstructured (Nagadoi et al., 1995). The flexibility of the hinge helix region is implied by the inability to resolve the N-terminal domain in crystallographic studies (Schumacher et al., 1995) as well as its proteolytic susceptibility (Choi and Zalkin, 1992). The hinge helix plays a crucial role in DNA specificity, since mutation of K55 of PurR results in loss of the high affinity binding (Glasfeld et al., 1999). Therefore, the hinge helix contacts would not be anticipated to contribute to ion concentration dependence of DNA binding in the absence of guanine. Similarly, the separation of the core from the DNA binding domain due to hinge helix unfolding would preclude the ionic contact of the R115 side chain of the core domain.

Analysis of the X-ray structure of the PurR holorepressor complexed with DNA produces 6 contacts per monomer (K5, H20, R26, R52, K55 and R115) or a total of 12 ionic contacts for the dimeric PurR, which is the oligomeric form that binds DNA. However, the ion concentration dependence experiments showed 15 ionic contacts, leaving ~3 additional cations that must originate from other sources. Interestingly, the PurR-guanine-purF complex crystal structure revealed a potential inter-chain ionic pair formed by H44 to E136 of the opposite chain that may enhance protein-DNA binding by stabilizing the orientation of the N-terminal helix-turn-helix motif (Figure 4.7B). Structurally, these residues are part of an interface between the DNA binding domain of one monomer and the core domain of its partner within the dimer. This ionic interaction is likely to be present only in holorepressor bound to DNA, and ion concentration dependence differences between the low and high affinity forms of PurR may reflect changes in the N-terminal domain-core interface. However, formation of such an ion pair would have the same effect on the ion concentration
dependence as the release of two cations only if counterions are tightly bound at these sites. In that case, the slope of the ionic strength dependence would depend on the stoichiometries of both cation and anion release (see Equation 2.32). No structural evidence is in hand for such tightly bound ions to these amino acid side chains. A similar possibility arises from release of ions upon formation of inter-chain ion pairs at the monomer-monomer interface of the dimeric PurR protein at sites such as E70-R278. The same caveat relating to the requirement for tightly bound counterions would apply to this ion pair as well. Furthermore, contribution from such ion pairs requires monomer association to be linked to DNA binding equilibria. Analysis of the DNA binding curves for PurR does not immediately reveal such linkage, although more detailed studies would be required to eliminate this possibility entirely.

4.3.2 Correlation of ion concentration dependence with structural data for LacI

Lactose repressor binding to DNA has been the biological model for formulating the "polyelectrolyte" effect. In the past, structural data from NMR studies (Chuprina et al., 1993; Spronk et al., 1996) were utilized to identify positively charged side-chains of LacI that make contacts in the high affinity complex (Capp et al., 1996; Frank et al., 1997). Using NMR data, Frank et al. (1997) suggested that each of the monomeric N-terminal domains of LacI utilizes R22, H29, R51, and possibly K59 to contact phosphates on DNA. However, Capp et al. (1996) indicate that D8 is within 6 Å of R51, suggesting that R51 may be partially neutralized in the isolated N-terminus, consistent with release of ~6 cations from the DNA upon LacI complex formation. Although the X-ray structure of the LacI bound to its symmetric operator $O^{sym}$ was available, the positions of the N-terminal side-chains were unresolved (Lewis et al., 1996).

Recently, the X-ray structure of LacI dimer complexed to operator DNA in
the presence of the anti-inducer ONPF was solved, providing sufficient resolution of most of the N-terminal side-chains (Bell and Lewis, 2000). This structure of LacI in its high DNA binding affinity mode allows the identification of basic residues that make contact with the DNA backbone. In a fashion similar to purine repressor, the contacts can be classified depending on the origin of the residue:

- helix-turn-helix (R22 and H29) as shown in Figure 4.8

- core (R118) as shown in Figure 4.9

In addition, the backbone of K2 is within ~5 Å of the phosphate backbone, but the side chain is not resolved and may also contribute a partial charge to the interaction. A comparison of the various ion-pairs formed for both PurR and LacI is summarized in Table 4.3.

**Table 4.3: Protein-DNA electrostatic contacts for PurR and LacI proteins**

<table>
<thead>
<tr>
<th>Contact classification</th>
<th>Basic amino acid candidates (Distance in Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PurR·purF·guanine</td>
</tr>
<tr>
<td>Helix-turn-helix</td>
<td>K5 (5.0)</td>
</tr>
<tr>
<td></td>
<td>H20 (7.4)</td>
</tr>
<tr>
<td></td>
<td>R26 (6.3)</td>
</tr>
<tr>
<td>Hinge helix</td>
<td>R52 (4.3)</td>
</tr>
<tr>
<td></td>
<td>K55 (7.0)</td>
</tr>
<tr>
<td>Core domain</td>
<td>R115 (4.1)</td>
</tr>
<tr>
<td>Intersubunit</td>
<td>H44 (3.7)</td>
</tr>
</tbody>
</table>

Basic residues that are candidates for making ionic contacts to DNA for both PurR and LacI are classified by their origin. Cut-off distance for selecting the candidates was 7.5 Å.

A noted difference between the ionic contacts of PurR and LacI is the absence
Figure 4.8: LacI helix-turn-helix contacts to DNA

Contacts of basic amino acid side chains of R22 (top) and H29 (bottom) of LacI. Each LacI monomer is shown in mauve and green (chains A and B, respectively). The indicated residues from chain A are sketched as a ball-and-stick model, encapsulated by a transparent surface. DNA is sketched as a stick model, with its backbone colored in yellow and red for P and O atoms, respectively. The ion-pair formation is illuminated. Coordinates are from PDB file 1EFA (Bell and Lewis, 2000).
Figure 4.9: LacI core domain contact to DNA

Contact of basic amino acid side chains of R118 of LacI core domain to DNA is presented. Each LacI monomer is shown in mauve and green (amino acid chains A and B, respectively). R118 from chain A is sketched as a ball-and-stick model, encapsulated by a transparent surface. DNA is sketched as a stick model, with its backbone colored in yellow and red for P and O atoms, respectively. The ion-pair formation is illuminated. Coordinates are from PDB file 1EFA (Bell and Lewis, 2000).
of hinge-helix ionic contacts in LacI. For instance, the side chain of K59 of LacI is directed away from the DNA, although its backbone is within ~6 Å of the phosphate backbone. Also, the contact of R51, previously proposed by Frank et al. (1997), cannot occur in the intact protein, as the side chain of R51 is further than 7.5 Å from the phosphate backbone, presumably due to reorientation elicited by the presence of the hinge helix.

In general, analysis of either NMR or crystallographic structures for LacI independently is consistent with the ~6 cations released on specific DNA complex formation (whether anti-inducer is present or not), despite the difference regarding R51. The discrepancy between the two types of structures may reflect inherent flexibility in the specific contacts made by this protein under different conditions, particularly in the absence of the core domain. Interestingly, flexibility in binding has been invoked to explain other LacI DNA binding experiments. Wild type LacI binding to sequences that deviated stepwise from $O^1$ or $O^{sym}$ DNA released more cations than the parent DNA sequence (Frank et al., 1997; Mossing and Record, 1985). As DNA binding becomes less favorable, the protein-DNA interface apparently shifts to increase electrostatic interactions and thereby enhance affinity, a process termed adaptability (Mossing and Record, 1985).

To further explore this issue, H74W LacI, a mutant with significantly lower IPTG responsiveness was employed in this study (Barry and Matthews, 1999a). The goal was to determine whether the preference for the high affinity protein form, even in the presence of inducer, would be reflected in the ionic strength dependence. The results indicated ~8 ionic contacts were formed between H74W LacI and operator DNA, compared to ~6 for wild type LacI. In the presence of IPTG, H74W LacI made 4 to 5 contacts, a decrease of ~2-fold, but still an absolute value greater than that observed for wild type LacI. This increased cation release for both forms of H74W LacI may reflect “adaptation” of the protein-DNA interface to enhance binding affinity by
increasing the number of electrostatic interactions. This process could engage several potential side chains, including K2, R51, or K59. The elevated cation release for H74W in the presence of inducer might derive from a combination of adaptability and an incomplete transition to the low affinity conformation in the presence of inducer.

Interestingly, the 2-fold decrease in cation release for low affinity binding is consistent with the pattern observed for both wild type LacI and PurR proteins. In the high DNA binding affinity form, both proteins employ a number of ionic contacts, including an arginine residue from their core domains. In the low DNA binding affinity form, these proteins make fewer contacts that displace cations from the DNA backbone. These contacts presumably originate only from the helix-turn-helix domains of both proteins, because the N-terminal domains are not well aligned and the hinge helix is not present.

The -11aa LacI dimer provided the opportunity to compare behavior of a dimeric system that exhibits thermodynamic linkage of assembly and DNA binding (Chen and Matthews, 1994). The measurements for -11aa LacI indicate a decrease in the number of cations released relative to wild type LacI (∼4 versus ∼6, respectively). The observed decrease in cation release for high affinity -11aa LacI binding to DNA may arise from the significant release of water that would be expected to accompany assembly of the hydrophobic monomer-monomer interface and would oppose the contribution of ion release (see Equation 2.32). Given these data, coupling of assembly with DNA binding for PurR might be anticipated to decrease, rather than increase, the number of cations observed, and we conclude this to be an unlikely source for the additional cations for PurR.

4.3.3 Temperature dependence

To explain the unexpectedly large magnitude of $\Delta C_p$ observed in many cases of protein-DNA binding, linkage of protein domain folding with complex formation has
been suggested (Ha et al., 1989; Spolar and Record, 1994). An alternative hypothesis is that a large, negative $\Delta C_p$ is the signature for formation of a highly complementary specific interface, independent of the affinity of the interaction (Ladbury et al., 1994; Sturtevant, 1977). The formation of this specific interface restricts the degrees of freedom in the polar, hydrated surfaces involved (Ladbury et al., 1994; Sturtevant, 1977).

The behavior of PurR and LacI proteins can be interpreted in the context of both of these models. The folding of the hinge helices has been suggested as a key difference between the high affinity and low affinity forms of the PurR (Choi et al., 1994; Nagadoi et al., 1995; Schumacher et al., 1995, 1994) and LacI proteins (Bell and Lewis, 2000; Lewis et al., 1996; Spronk et al., 1999a, 1996). Frank et al. (1997) showed that folding of the two hinge helices (one per monomer) could account for $\sim 1.0$ kcal/mol K change in the heat capacity for LacI or PurR, consistent with the magnitude of the LacI dimer and tetramer $\Delta C_p$ values. In contrast, van't Hoff analysis of the temperature dependences of DNA binding showed that PurR-guanine-DNA complex formation was accompanied by a very large negative heat capacity change of $-2.8\pm0.5$ kcal/molK. In the absence of guanine, this value decreased to $-0.8\pm0.3$ kcal/molK. This result is consistent with the ion concentration dependence observations reported here that indicate that the hinge helix ionic contacts do not occur unless guanine is present. The larger $\Delta C_p$ for high affinity PurR binding compared to LacI may derive from a more extensive folding reaction for the PurR hinge region, consistent with the ability to detect (a) folding of the hinge helices for the isolated N-terminal domains of LacI bound to operator and (b) DNA binding by these domains for LacI but not for PurR (Nagadoi et al., 1995; Spronk et al., 1999a, 1996). Furthermore, the larger absolute magnitude of heat capacity change observed in the case of specific DNA binding by PurR-guanine versus PurR is consistent with the higher magnitude of the $\Delta C_p$ value observed for higher affinity LacI binding to $O_{sym}$.
and with the decrease in the absolute value of $\Delta C_p$ for LacI binding to progressively less optimal sequences derived from $O^{syn}$ operator (Frank et al., 1997).

Assuming linkage of protein folding to DNA binding, our results indicate that the local or global protein folding occurring upon protein-DNA complex formation is larger in the case of PurR holorepressor complex formation than for LacI or for PurR binding alone. The latter difference may originate in part from different degrees of hinge domain and/or N-terminal domain folding in the PurR-corepressor-DNA and the PurR-DNA complexes. In these experiments, any alterations due to PurR conformational changes between guanine-bound and free forms will not contribute to the observed effects, since the excess of corepressor in the binding reactions ensured PurR saturation throughout the titration range.

Although X-ray crystallographic structures for the holorepressor-DNA complex and the corepressor binding domain of PurR are available (Schumacher et al., 1995, 1994, 1997), interpretation of protein folding differences between the two complexes is not possible because no PurR-DNA crystal structure is available. The NMR solution structure of the PurR N-terminal domain superimposes on that of the holorepressor X-ray structure, except for the hinge helix region which was not folded (Nagadoi et al., 1995). An interesting possibility is that the contacts between the N-terminal helix-turn-helix and core domains bury additional apolar regions in the holorepressor complex.

Linkage of monomer-monomer assembly and DNA binding equilibria, as observed for LacI dimers (Chen and Matthews, 1994), is another potential source for extensive burial of hydrophobic residues. The large heat capacity change for PurR might derive from such coupled equilibria. In fact, the heat capacity changes estimated for PurR in our studies lie in the range of those observed for GalR and MntR that were interpreted in terms of coupling protein assembly with DNA binding equilibria (Spolar and Record, 1994). For comparison, we examined the temperature
dependence of DNA binding for -11aa LacI, a dimeric mutant protein for which coupling of monomer-monomer association to dimers and dimer DNA binding has been established (Chen and Matthews, 1994). The -11aa LacI exhibited a heat capacity change of -0.9 kcal/molK in its high affinity form, similar to the value observed for tetrameric wild type LacI-O$^1$ interactions (Spolar and Record, 1994). Although substantial effects presumably derive from removal of apolar residues at the subunit interface, these appear to be balanced by other influences to yield a similar $\Delta C_p$. We suggest that coupled assembly and DNA binding do not likely account for the magnitude of $\Delta C_p$ observed for PurR holorepressor binding to DNA, consistent with the cation release results for PurR.

A final element that may account for the variance between $\Delta C_p$ for LacI and PurR high affinity DNA binding is differential effects of complex formation on internal vibrational modes (Ladbury et al., 1994). In the complex, increased restriction of polar elements, including amino acid side chains, nucleotide bases/backbone, or water molecules, in the protein-DNA interface affects contributions to the heat capacity. This restriction factor has been proposed to account for heat capacity changes of almost the same magnitude as the hydrophobic effect (Keown et al., 1998; Ladbury et al., 1994; Lundbäck et al., 1998; Oda et al., 1998; Privalov et al., 1999; Zou et al., 1998). As the degrees of freedom of those polar elements decrease, the contribution of this effect to the absolute magnitude of heat capacity change is expected to increase relative to that of the hydrophobic effect. Indeed, the folding of the hinge helix and the formation of the N-terminal-core domain interface upon holorepressor-DNA complex formation would place additional restrictions on polar amino acid and DNA elements not present when DNA binds PurR alone. A similar argument is applicable to LacI. However, distinguishing factors between these proteins are the higher number of electrostatic interactions for PurR versus LacI and differences in the R factors for the crystal structures. The larger R factor reported for the LacI-ONPF-O$^{sym}$. 
crystal structure (R factor = 0.25) (Bell and Lewis, 2000) compared to that for PurR·guanine·purF (R factor = 0.17) (Schumacher et al., 1997), despite similar resolution (2.6 Å), might reflect higher contribution of restricting internal vibrational modes to the observed ΔC_p for PurR.
Chapter 5

Biochemical characterization of \( \Delta 33 \)

C-terminal deletion mutant of human p53

5.1 Introduction

DNA binding capacity of p53 is a determining factor for cellular life or death, since the most frequent mutations map to arginine residues that mediate this function by making crucial contacts to DNA (Cho et al., 1994). Despite the importance of DNA binding properties, only a limited amount of thermodynamic data exists and mainly from studies involving the core DNA binding domain (Balagurumoorthy et al., 1995; Klein et al., 2001; Nagaich et al., 1997a; Nikolova et al., 1998). Noting this need, efforts towards a thorough thermodynamic description of p53’s function were undertaken in the Matthews lab. Characterization of wild type p53 showed that in contrast to what was commonly accepted, bacterially expressed wild type p53 presented a thermostable structure and bound specifically to DNA with affinities in the nM range (Nichols and Matthews, 2001). Recently, S392E was shown to be more thermostable than wild type p53, with similar DNA affinity (Nichols and Matthews, 2002). The studies in this chapter focus on a mutant for which the 33 C-terminal amino acids are deleted — \( \Delta 33 \) p53 — in an effort to examine the effect of the native C-terminus on the structure and DNA binding properties of p53.
The previous chapter emphasized the description of the thermodynamic parameters of DNA binding for LacI and PurR in terms of the polyelectrolyte and the hydrophobic effect. In this chapter, similar analyses are applied to the DNA binding properties of human suppressor protein p53. By comparing the wild type, S392E and Δ33 p53 proteins, the effect of the C-terminus on thermodynamic behavior of these proteins is examined.

The focus of the current studies on the effect of the C-terminus (amino acids 360–393) originates from the proposed role of this region in regulating the DNA binding properties and consequently the biological function of p53. A now questionable allosteric model has been proposed in which the native C-terminus exerts inhibitory effect on the p53 DNA binding affinity (see Section 1.4.3). Modifications or even deletion of the C-terminal domain were postulated to activate p53 by relieving the inhibitory effect of the C-terminus (Gu and Roeder, 1997; Hupp and Lane, 1995; Hupp et al., 1992; Jayaraman and Prives, 1995; Selivanova et al., 1996, 1997). Qualitative studies showed that this protein is 'constitutively' active, presumably because of the absence of the negative regulatory C-terminal region (Hupp et al., 1992). Because Δ33 p53 would therefore not require any other effects of factors (for instance, phosphorylation or presence of pAb421) to activate its function, thermodynamic characterization of its DNA binding properties should reflect the putative activated state of p53. Moreover, the deletion of the C-terminal region eliminates any potential interference by p53 DNA binding in a non-specific manner, a property that has been attributed to this domain and questioned the activation model of DNA binding by p53 (Anderson et al., 1997; Nichols and Matthews, 2002). Characterization of the Δ33 p53 will provide information to test further the validity of the allosteric model.
5.2 Results

5.2.1 Secondary structure of Δ33 p53

The thermal stability of Δ33 p53 was monitored by circular dichroism spectroscopy as a function of temperature (Figure 5.1). In general, the shape of the circular dichroism spectra at relatively low temperatures resembles the circular dichroism behavior of both wild-type and S392E p53 proteins. However, as temperature increases there is a progressive change in the CD signal, most evident at the far UV circular dichroism wavelengths. The “trough” below 210 nm disappears as temperature increases. In contrast, the circular dichroism signal above 210 remains almost unaffected as temperature increases. Interestingly, even at the highest temperature (90 °C) employed in this study, Δ33 p53 retains most of its circular dichroism signal. Unfortunately, the thermal denaturation transition was irreversible, since renaturation of denatured protein did not produce identical circular dichroism signals.

Information about protein secondary structure character as derived from circular dichroism spectra could be obtained by applying spectra deconvolution methods (reviewed in Greenfield, 1996). Such approaches are cumbersome because (a) the reliability is variable since the deconvolution depends on the structures of the standard proteins used for sampling, and (b) special sample handling is required, to minimize interference from atmospheric oxygen in the wavelength range ≤200 nm. However, valuable information about the secondary structure content can be deciphered empirically by examining the signal at 208 or 218 nM that corresponds to the α-helical or β-sheet secondary structure, respectively. These changes of the circular dichroism signal as a function of temperature, are indicated more clearly in Figure 5.2 where the circular dichroism signals at 208 and 218 nm were monitored. Increasing temperature results in a loss of circular dichroism signal at 208, indicating that there is
Figure 5.1: Thermal denaturation of $\Delta 33$ p53 as monitored by circular dichroism

Circular dichroism spectra of $\Delta 33$ p53 with increasing temperature from 0 to 90 °C. The process is not reversible, as the decreasing temperature did not reproduce the original spectra. For demonstration purposes, spectra at every 20 °C are shown, with 0 °C corresponding to blue-color and 80 °C corresponding to red-color.
Figure 5.2: Circular dichroism signal change as a function of temperature

The normalized circular dichroism signals at 208 and 218 nm, corresponding to α-helical (●) and β-sheet (■) content of the secondary structure, was monitored as a function of increasing temperature. The process was not reversible. Signals were normalized to the values of the spectrum obtained at 0 °C.
a loss of α-helical content in the structure. In contrast, the signal at 218 nm, which corresponds to the β-sheet content of the structure, remains largely unaffected.

The thermal behavior of Δ33 p53 monitored by circular dichroism as a function of temperature can be compared to available data on wild-type and S392E p53. Wild-type p53 exhibited a significant loss of both α-helical and β-sheet secondary structure (≈50 and 20%, respectively) with increasing temperature, with a midpoint of the transition in the range of ≈70 °C (Nichols and Matthews, 2001). However, the S392E p53 mutant was more thermostable, since it retained its β-sheet structure and exhibited a modest loss (≈20%) of α-helical content (Nichols and Matthews, 2002). Therefore, Δ33 p53 thermal stability behavior resembles more clearly that of the S392E p53 protein.

5.2.2 Protein assembly

To examine the oligomeric state of Δ33 p53, glutaraldehyde crosslinking was employed. Δ33 p53 was incubated with varying amounts of glutaraldehyde over varying incubation time. The resulting crosslinked species examined on a denaturing acrylamide gel are shown in Figure 5.3. The distribution of the crosslinked species was dependent on the glutaraldehyde concentration and the incubation time. In general, higher oligomeric protein species appeared as the incubation time or glutaraldehyde concentration increased. Comparison to protein standards was employed to assign the molecular weight and identify the oligomeric state of the crosslinked species. Although, the apparent molecular weight of proteins as estimated from denaturing acrylamide gels depends both on size and shape of the molecule, the observed cross-linked species correspond approximately to monomeric, dimeric, and tetrameric species. Participation of the dimeric species appeared to be as a transient intermediate species. At higher glutaraldehyde concentrations, higher molecular weight bands appear that represent probably nonspecific intermolecular crosslinked species. The
**Figure 5.3: Crosslinking analysis of Δ33 p53**

Δ33 p53 (0.06 mg/mL) was incubated in the absence (lane 0) or presence of 0.01 % (lanes 1 and 2) or 0.1 % (v/v) (lanes 3 and 4) of glutaraldehyde. Incubation time varied from 5 min (lanes 1 and 3) to 30 min (lanes 2 and 4). Lane S represents protein standards with their molecular weight (in Kdalton) indicated. The resulting monomeric and cross-linked dimeric or tetrameric species are indicated as m, d and t on the right side.

crosslinking profile for Δ33 p53 was similar to that observed for wild-type and S392E p53 proteins (Nichols and Matthews, 2001, 2002).

Indirect information about protein assembly can also be derived from the electrophoretic behavior in mobility shift assays. Although traditionally used for examining protein-DNA interactions, this technique separates various species according to size, charge and shape. Wild type p53 formed a distinct complex that entered the native acrylamide gel in addition to some complex that stayed in the wells, probably corresponding to nonspecifically adherent high molecular weight complex. In contrast, Δ33 p53 protein formed only one complex that entered the gel and corresponded to possible tetrameric p53 binding to DNA. Therefore, the presence of the C-terminal region may mediate some of the adherent properties of p53 and perhaps promote higher order aggregates.

DNA binding examined with the mobility shift assay was performed under both stoichiometric and non-stoichiometric conditions. In theory, titrations under stoichiometric conditions will involve binding of all potential binding sites on the protein, whereas titrations under non-stoichiometric conditions monitor binding of
the site of highest affinity with a single binding event. Interestingly, a single complex was observed in the binding of Δ33 p53 suggesting a 4:1 protein monomer to DNA stoichiometry. This result can be extended to wild type p53, despite the presence of higher molecular weight complexes in the wells.

**Figure 5.4: Mobility shift of p53 proteins**

Mobility shift assays of Δ33 p53 (A) and wild type p53 proteins (B), performed under non-stoichiometric (ns) and stoichiometric (s) conditions (DNA concentration was kept constant at 5 x 10^{-12} and 2 x 10^{-8} M, respectively). Reaction mixtures in the 50 mM potassium phosphate binding buffer (with the exception that 10%(v/v) of glycerol was used, so that samples go to at the bottom of the wells during loading) were incubated for 45 min and loaded under high voltage on 4%(w/v) acrylamide:bisacrylamide 19:1 native gel.
5.2.3 DNA binding of \(\Delta 33\) p53

Since the biological functions of p53 rely on its DNA binding properties, the DNA binding affinities of \(\Delta 33\) p53 to specific and non-specific DNA were examined. In order to compare to previously characterized wild-type and S392E p53 proteins (Nichols and Matthews, 2001, 2002), the same buffer conditions (50 mM potassium phosphate buffer) were employed in this study. Interestingly, \(\Delta 33\) p53 showed a \(\sim5\)-fold increase in DNA binding affinity compared to wild-type p53 (summarized in Table 5.1). The DNA binding behavior of \(\Delta 33\) p53 protein is distinct from that

**Table 5.1: DNA binding affinity of p53 proteins in 50 mM potassium phosphate buffer**

<table>
<thead>
<tr>
<th>DNA binding mode</th>
<th>(K_{DS}) in M</th>
<th>DNA specificity(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta 33) p53 specific</td>
<td>3.1±1.5 \times 10^{-10}</td>
<td>70</td>
</tr>
<tr>
<td>(\Delta 33) p53 non-specific</td>
<td>2.0±1.0 \times 10^{-8}</td>
<td></td>
</tr>
<tr>
<td>wild-type p53 specific(^2)</td>
<td>1.6±0.1 \times 10^{-9}</td>
<td>30</td>
</tr>
<tr>
<td>wild-type p53 non-specific(^3)</td>
<td>5 \times 10^{-8}</td>
<td></td>
</tr>
<tr>
<td>S392E p53 specific(^3)</td>
<td>2.5±0.2 \times 10^{-9}</td>
<td>4</td>
</tr>
<tr>
<td>S392E p53 non-specific(^3)</td>
<td>1.1±0.1 \times 10^{-8}</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) DNA specificity was calculated as the ratio of the \(K_{DS}\) for non-specific over specific DNA binding. \(^2\) From Nichols and Matthews (2001). \(^3\) From Nichols and Matthews (2002).

exhibited by S392E p53, which showed a \(\sim2\)-fold weaker specific DNA binding affinity compared to wild type p53. The opposite behavior of \(\Delta 33\) and S392E p53 proteins extended to DNA binding specificity: S392E p53 showed minimal DNA specificity (\(\sim4\)-fold), whereas \(\Delta 33\) p53 displayed greater DNA specificity (\(\sim70\)-fold).
5.2.4 Ion concentration dependence of DNA binding of Δ33 p53

The polyelectrolyte effect on the p53-DNA interactions was utilized to determine the extent of cation release. Although the focus of this thesis is the Δ33 p53 protein, wild-type and S392E p53 were also employed to identify specific effects of the C-terminus. Using a variable KCl concentration of a Tris-KCl buffer system, the specific DNA binding affinities DNA binding of Δ33, wild-type and S392E p53 and non-specific DNA binding of Δ33 p53 were determined (Figure 5.5). Increase in the KCl concentration resulted in decreased binding affinities, a general phenomenon in protein-DNA interactions. Interestingly, the absence of the C-terminus did not alter dramatically the behavior of the p53 proteins, in terms of the effect of salt on DNA binding affinities.

The number of cations released and consequently the number of ion pairs formed between basic protein side chains and the DNA backbone, can be derived from the negative slopes of the ion concentration dependence, according to Equation 2.34. The results for the proteins examined are summarized in Table 5.2. In general, cation release from all proteins was generally very low, reflecting ion pair formation ranging from 0.5 to 3. Δ33 p53 showed a small but distinct difference in the ion concentration profile as compared to wild type and S392E p53 proteins. This finding suggests that the presence of the C-terminus of p53 may reduce the number of ionic interactions in the protein-DNA interface. Moreover, the S392E mutation mimicking phosphorylation did not reverse this behavior.

5.2.5 Temperature dependence of DNA binding

The extent of the “hydrophobic” effect on DNA binding of p53 was examined by monitoring the temperature dependence of DNA binding. Binding affinities of Δ33
Figure 5.5: Ion concentration dependence of DNA binding of p53 proteins

The solid lines represent best weighted linear fits to DNA binding of p53 proteins. (A) Δ33 p53 specific (●) and non-specific (■), respectively. (B) Wild type p53 specific DNA binding (○). (C) S392E p53 specific DNA binding (□).
Figure 5.5: (continued)
Table 5.2: Number of ion pairs formed upon DNA binding of p53 proteins

<table>
<thead>
<tr>
<th>Protein</th>
<th>DNA binding mode</th>
<th>No of ion pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ33p53</td>
<td>Specific</td>
<td>3.1±0.3</td>
</tr>
<tr>
<td>Δ33p53</td>
<td>Non-specific</td>
<td>1.8±0.5</td>
</tr>
<tr>
<td>Wild type p53</td>
<td>Specific</td>
<td>1.6±0.2</td>
</tr>
<tr>
<td>S392E</td>
<td>Specific</td>
<td>0.4±0.1</td>
</tr>
</tbody>
</table>

The number of ion pairs formed was calculated from the linear slopes of the ion concentration dependence plots in Figure 5.5, using the simplified Equation 2.34.

p53 for specific and nonspecific DNA were determined as a function of temperature ranging from 0 to 37 °C. Affinities were determined in 50 mM potassium phosphate buffer to allow comparison with previously studied wild-type and S392E p53 proteins (Nichols and Matthews, 2001, 2002). Analysis of the temperature dependence was examined by van’t Hoff plot. Both specific and nonspecific binding of Δ33 p53 exhibited nonlinear van’t Hoff behavior, with similar curvature (Figure 5.6). As a result, Δ33 p53 retained its DNA target selectivity throughout the temperature range examined. A summary of the thermodynamic parameters ΔC_p, T_H and T_S as derived from fitting to the temperature dependence data are summarized in Table 5.3. In general Δ33 p53 exhibited a modest negative heat capacity range in the range of -0.7 to -1.0 kcal/molK, slightly higher than previously reported for wild-type and S392E (Nichols and Matthews, 2001, 2002).

5.2.6 Water release

Osmolar stress can be used to detect the number of water molecules released upon protein-DNA complex formation (Garner and Rau, 1995). This technique was
Figure 5.6: van't Hoff plots for specific and nonspecific binding of Δ33 p53

Specific (•) and nonspecific (○) DNA binding of Δ33 p53 are shown. The curves represent best weighted nonlinear van't Hoff fit to the data, according to Equation 2.38.
Table 5.3: Thermodynamic parameters of DNA binding of p53 proteins

<table>
<thead>
<tr>
<th>Protein-DNA binding</th>
<th>$\Delta C_p$ (in kcal/molK)</th>
<th>$T_H$ (°C)</th>
<th>$T_S$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ33 specific</td>
<td>-0.7±0.2</td>
<td>283±3.2</td>
<td>300±3.6</td>
</tr>
<tr>
<td>Δ33 non specific</td>
<td>-1.0±0.2</td>
<td>282±1.6</td>
<td>293±1.0</td>
</tr>
<tr>
<td>wild-type specific$^1$</td>
<td>-0.5</td>
<td>288</td>
<td>313</td>
</tr>
<tr>
<td>S392E specific$^2$</td>
<td>-0.28</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

NR, not reported. $^1$ From Nichols and Matthews (2001). $^2$ From Nichols and Matthews (2002).

utilized to examine water release upon DNA binding for Δ33, S392E and wild-type p53 proteins. Increasing osmolar concentration of glycerol resulted in an slight increase of specific DNA binding affinities for all three p53 protein variants examined (Figure 5.7). From the linear best fits, the number of water molecules released was determined for all three proteins and is summarized in Table 5.4. The number of water molecules released upon DNA binding was quite low ranging in between $\sim$5 to 20 for all proteins examined.
Table 5.4: Number of water molecules released for p53 proteins by osmotic stress

<table>
<thead>
<tr>
<th>Protein</th>
<th>No of released water molecules$^{1,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ33p53</td>
<td>5±10</td>
</tr>
<tr>
<td>S392Ep53</td>
<td>10±5</td>
</tr>
<tr>
<td>Wild type p53</td>
<td>23±17</td>
</tr>
</tbody>
</table>

$^{1}$ Osmotic stress was conducted by increasing glycerol concentration. $^{2}$ The number of water molecules released upon binding of the examined proteins to specific DNA was determined by using Equation 2.36.

Figure 5.7: Osmotic stress of DNA binding by p53 proteins

Affinities are plotted against the osmolality (m) of increasing glycerol concentrations (0, 5, 10 and 15% (v/v), respectively). Solid lines represent best linear fits to DNA binding data of p53 proteins. (A) Δ33 p53 specific DNA binding (●). (B) Wild type p53 specific DNA binding (○). (C) S392E p53 specific DNA binding (■).
Figure 5.7: (continued)
5.3 Discussion

5.3.1 Biochemical characterization of Δ33 p53

In this chapter, the effect of the C-terminal regulatory domain of p53 on the biochemical properties of the protein (secondary structure and DNA binding) was examined by deleting this region to produce Δ33 p53 protein. Expression and purification of this protein from bacterial cells yielded a folded and functional tetrameric protein, as indicated by its secondary structure analysis, assembly and DNA binding characterization. In the past, similarly expressed wild-type and S392E p53 proteins have yielded functional full-length proteins, as shown by their biochemical and biophysical characterizations (Nichols and Matthews, 2001, 2002). Therefore, a comparison of the properties of Δ33 p53 to wild-type and S392E proteins was feasible.

Thermal denaturation of Δ33 p53 protein

Thermal denaturation of Δ33 p53 protein was monitored by observing the circular dichroism spectra as a function of temperature. Δ33 p53 displayed a thermal denaturation profile similar to that of the S392E p53 mutant, with a loss of intensity corresponding to α-helical structure but with almost unaltered β-sheet secondary structure (Nichols and Matthews, 2002). In contrast, wild-type p53 displayed a marked difference in its thermal denaturation profile, with loss in both α-helical and β-sheet secondary structure (Nichols and Matthews, 2001). Based on secondary structure prediction analysis, the striking difference between the thermal stability of the wild-type and S392E proteins was attributed to the core DNA binding domain, since it is the most dominant contributor of β-sheet secondary structure (Nichols and Matthews, 2002). The presence of the C-terminal regulatory domain most likely does not contribute to the α-helical and β-sheet structure, since structural studies
showed that this domain although adopting an α-helical structure upon binding to
the S100ββ protein, was unstructured in its free form (Rustandi et al., 2000). The
thermal denaturation profile of Δ33 p53 further elucidates the effect of the C-terminal
regulatory domain. Rather than providing stabilizing influence in its modified form,
the unmodified C-terminal regulatory domain, appears to destabilize the core DNA
binding domain. Furthermore, it provides further support to the hypothesis that the
structure of the core DNA binding domain when the C-terminal regulatory domain
is unmodified is fairly rigid (Nichols and Matthews, 2001).

**DNA binding**

DNA binding analysis of the Δ33 p53 protein was employed to study the
effect of the C-terminal regulatory domain on the DNA binding properties of p53.
Comparison of the DNA binding affinity of the Δ33 p53 to that of wild-type full-length
p53 proteins revealed a modest ~5-fold increase in specific DNA binding affinity in 50
mM potassium phosphate. This difference accounted in large part for the increased
DNA binding specificity of Δ33 p53 compared to wild-type p53. However, in a KCl-
based buffer utilized for examining ion concentration dependence, this difference was
diminished. These data suggest that the ionic component of the binding buffer, rather
than the intrinsic properties of the proteins, contribute to the differential behavior in
the presence or absence of the C-terminal regulatory domain. This differential effect
of phosphate was not examined further.

Despite the discrepancy between the DNA binding affinities for specific DNA
binding, Δ33 p53 exhibited DNA specificity in both buffer systems. However, differ-
ential binding for specific *versus* non-specific DNA sequences was not as high for p53
and its examined variants as observed for prokaryotic LacI and PurR, the other pro-
teins examined in this thesis. Although affinity for non-specific DNA was weaker, in
general, Δ33 p53 bound non-specific DNA in the submicromolar range. This finding
demonstrated that the core DNA binding domain can bind DNA in a non-specific manner, albeit with a lower affinity than observed in the presence of the C-terminal regulatory region. Previously, large non-specific double stranded DNA fragments have been shown to bind to the core domain and inhibit specific DNA binding (Anderson et al., 1997). Although the core domain preferentially binds to specific DNA, the low level of nonspecific binding could allow competition by non-specific DNA, particularly when the latter sequences exist in relatively high concentrations compared to specific sequences.

Low DNA binding specificity appears to be an intrinsic property of all of the p53 proteins examined. This phenomenon may arise from the fact that the DNA binding motif has a significant degree of variability (el Deiry et al., 1992). This notion implies that factors others than just DNA sequence participate in DNA recognition. For instance, DNA secondary structure can also play a crucial role, since p53 can bind better to the binding site when it is presented with a stem-like structure (Kim et al., 1997). Moreover, the DNA binding affinity of the central core depends on the flexibility of the target sequence (Nagaich et al., 1997a). In vivo, a set of p53 mutations distinguish the p21 and BAX promoters based on their degree of flexibility, therefore不同iating the cellular response (cell cycle arrest versus apoptosis, respectively) (Flaman et al., 1998).

**Protein assembly**

Two independent methods were employed to examine the protein assembly properties of Δ33 p53 protein: glutaraldehyde crosslinking and mobility shift. The former has been traditionally used for investigating assembly properties of p53 proteins. Δ33 p53 was shown to form tetramers, a finding that is consistent with previous studies. The mobility shift method revealed two additional pieces of information. First, only a single species, presumably tetramer, for Δ33 p53 and probably wild type
p53 proteins bound to DNA molecules. However, these data do not provide information on how many monomeric subunits interact with the binding site. Previously, modelling studies indicated that the binding site can accommodate four independent core domains with an accompanying degree of DNA bending (Nagaich et al., 1999, 1997b). However, a detailed view awaits a structure that includes the tetramerization domain and subsequently can demonstrate how the core domains are arranged in the tetrameric protein. Secondly, the presence of the C-terminal regulatory domain resulted in retention of wild type p53 protein-DNA complex species in the wells, indicating potential intermolecular interactions.

**Correlation of ion concentration dependence to ion pair formation**

Ion concentration dependence of DNA binding reflects the number of ions released upon DNA binding, a term described as the “polyelectrolyte” effect (reviewed in Record et al., 1991). Hence, the number of ion pairs formed upon DNA complexation can be inferred. In Chapter 4, results from the ion concentration dependence of DNA binding for PurR and LacI proteins were successfully correlated with the available structural data. This approach was extended to p53. The ion concentration dependence of all proteins examined (Δ33, S392E and wild type p53) showed a very low number of ion pairs when complexed to DNA, ranging from 0.5 to 3 pairs.

The only available relevant structure is the structure of the core DNA binding domain complexed to DNA (Cho et al., 1994). Although the overall spatial arrangement of the individual core domains with respect to the whole protein is not available yet, the rigid structure of this domain suggested by the minimal structural changes between the human p53 core domain-DNA complex (Cho et al., 1994) and the free murine core domain (Zhao et al., 2001) substantiates the hypothesis that at least locally the structure of this domain will be retained in the whole protein. Interestingly, the original structure contained three p53 core domains, with one of them not
involved in DNA complexation but stabilizing the whole complex via protein-protein interactions. Therefore, for investigating ion pair formation, only the core domain that makes most DNA contacts was examined (Chain B in PDB file 1TSR, see Figure 5.8). The structure revealed five basic residues that made ionic contacts with the DNA backbone as summarized in Table 5.5. Ion pair formation by R248 and R273 have been determined to be crucial for the protein-DNA formation as indicated by

**Table 5.5: p53-DNA electrostatic contacts**

<table>
<thead>
<tr>
<th>Residue</th>
<th>Distance in Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>K120\textsuperscript{a}</td>
<td>5.7</td>
</tr>
<tr>
<td>R248</td>
<td>4.5</td>
</tr>
<tr>
<td>R273\textsuperscript{b}</td>
<td>2.5</td>
</tr>
<tr>
<td>R280\textsuperscript{b}</td>
<td>4.6</td>
</tr>
<tr>
<td>R283</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Ion pairs formed between the core domain of p53 (Chain B) and a 21 bp double stranded DNA containing p53 DNA halfsite, as derived from PDB file 1TSR (Cho et al., 1994). \textsuperscript{a} Contact by K120 is evident in only one of the core subunits in the crystal structure. \textsuperscript{b} Both basic side chains of R273 and R280 are within 4.5 Å from the carboxy side chain of D281. Also, the side chain of R273 is within 6 Å of E285.

their high frequency in human p53 mutations (Cho et al., 1994). Ion pair formation for K120 was evident in only one core domain (Chain B) out of three that comprise the crystallographic unit, raising the question of whether this residue plays such a crucial role as R248 and R273. This hypothesis is also supported by the low incidence of human cancer mutations at this site, compared to R248 and R283. Also, in a set of p53 mutations that were able to distinguish p53-dependent transcription driven by the p21 and BAX promoters, K120 was dispensable for p21 transcription, in contrast to R248 and R273 whose presence is essential (Flaman et al., 1998). Therefore, contribution of K120 to cation release is questionable. Furthermore, the acidic side
Figure 5.8: Ion pairs between p53 central core and DNA

All of the ion pairs formed between the core domain of p53 and DNA as based on the B subunit of PDB file 1TSR (Cho et al., 1994) are indicated in the top panel. p53 residues K120, R248, R273, R280 and R283 (sketch in blue ball-and-stick model) contact the DNA backbone (sketch in ball-and-stick model). Acidic residues E285 and D281 (D281 is sketched in red ball-and-stick model, the view of E285 is obstructed) contact the basic residues R273 and R280 and neutralize their charges. The zinc atom is depicted in orange. The region of amino acids 270-287 is displayed in bottom panel. Blue colors correspond to the positively charged amino side chains of R273, R280 and R283. The negatively charged carboxy side chains of E285 and D281 are colored in red.
chain of D281 contacts both the basic side chains of R280 and R273, which in turn is also contacted E285. Thermodynamic data suggest that D281 is crucial for the stabilization of the helix H2 that provides most of the basic contacts to the DNA, by possibly stabilizing the DNA contact surface by the means of ionic interactions to R280 and R273 (Trulson and Millhauser, 1999). Therefore, R280 and R273 may not contribute significantly to cation release because of the charge neutralization imposed by D281 and E285. In aggregate, the monomeric core domain forms 2 to 3 ion pairs and subsequently releases 2 to 3 cations upon DNA binding.

Based on the observed ion pair formation as derived from the ion concentration dependence of DNA binding, a possible interpretation is that p53 employs one core domain and possibly a second in the case of Δ33 p53. However, the oligomeric status of p53 is tetrameric as derived from the crosslinking and mobility shift experiments. If all core domains contributed equally, then the number of cations released would be expected to be in the range of 8 to 12. Therefore, the observed cation release from the ion concentration dependence is significantly lower than expected for tetramers.

The apparent discrepancy in correlating the macroscopic observations to the structural data should be viewed in the light of the inherent limitations of both the structural model and the available data. The present analysis relied on the structure of the core domain, with no information on how the whole tetrameric protein is arranged when binding to DNA naturally. Also, the structure of the core domain itself presents intrinsic limitations: in the structure, the DNA is in linear B-form and cannot accommodate binding of four core domains; modelling studies indicate that such accommodation is feasible only with significant DNA bending, and such bending has been observed experimentally (Nagaich et al., 1997a, 1999, 1997b).

The observed cation release would correspond to a total of only one or possibly two subunits, if the assumption that the protein-DNA interactions presented
in the known structure are preserved locally in the whole protein given the rigid structure of the core domain. Interestingly, structural NMR and X-ray studies of the core domain imply the existence of a monomer-monomer interface that occurs via of protein-protein interactions involving helix H1, which also stabilizes the structure by coordinating the Zn atom (Klein et al., 2001; Zhao et al., 2001). In the free protein, these contacts preclude simultaneous binding of the helix H2 (amino acids 278–286) to the DNA by both monomeric subunits (Zhao et al., 2001).

The discrepancy of ion pair formation with the structural data could also be viewed in terms of the limitations of the applied model of analysis. The proteins examined indicated an interesting trend; the presence of the native C-terminus reduced the number of cations released (Δ33 versus wild type p53). Moreover, introduction of an additional negative charge in the penultimate residue of p53 further reduced the cation release (S392E versus wild type p53). These findings might be rationalized by cation uptake within the C-terminal regulatory domain. This hypothesis requires that DNA binding disrupts potential protein-protein interactions between the C-terminal regulatory region and the rest of the protein (intramolecular) and/or other domains of the oligomeric ensemble (intermolecular). Uptake of cations by the C-terminal domain, and particularly by the very C-terminal end of S392E p53 protein (390-DSD-COOH in wild type and 390-DED-COOH in S392E p53 protein), will reduce the overall number of cations released.

Interactions between the C-terminal regulatory domain and the core domain have been hypothesized to explain the data according to an allosteric model, involving the inhibitory function of the C-terminal domain. However, this allosteric model is under question by a number of recent studies, including data presented in this thesis (see Section 5.3.2). The proximity of the C-terminal regulatory domain and the N-terminus, has been indicated based on the fact that lysine mutations at the C-terminus interfere with MDM2 binding at the N-terminus (Almog et al., 2001; Nakamura et al.,
2000; Rodriguez et al., 2000). Moreover, modelling experiments based on peptides showed that the C-terminus could dock with the N-terminal region formed by amino acids 84–89 (Kim et al., 1999). Therefore, it is evident that solution of the X-ray structure of the overall p53 protein will be necessary to fully explain the behavior of this protein and its variants.

In summary, the number of ion pairs formed upon p53-DNA binding as derived from ion concentration dependence of DNA binding was found to be significantly lower than the expected number based on the available structure of the core p53-DNA complex. The present analysis has attempted to explain this discrepancy by taking into account the limitations of the structural model employed and a postulating possible mechanism of the protein-DNA interactions of p53. Clearly, elucidation of the whole protein structure will be essential to explaining the observed data.

**Hydrophobic effect**

To investigate the hydrophobic effect on the DNA binding of Δ33 p53, van't Hoff analysis of the temperature dependence of DNA binding was employed. In Chapter 4, a similar approach was used to analyze PurR and LacI temperature dependence data in terms of structural rearrangements upon DNA binding. The magnitude of heat capacity change for Δ33 p53 ranged between -0.7 and -1.0 kcal/molK for specific and non-specific DNA binding, respectively. The absolute values for Δ33 p53 were slightly larger than the values reported for S392E and wild type p53 (-0.3 to -0.5 kcal/molK); however, the low magnitude of these values indicated that there is minimal local folding upon DNA binding. This notion is further supported by the degree of water release upon DNA binding. Water release upon DNA binding was very low, with a number \( \leq 20 \). Therefore, both approaches indicated that local folding is a much less significant contributor to the binding process than for LacI and PurR proteins examined in Chapter 4. Moreover, the absence of the native C-terminal
domain did not alter the predominant feature for p53-DNA interactions: binding by the rigid core domain. Recent NMR studies have indicated that the absence of the C-terminus did not alter the core domain structure (Ayed et al., 2001). The rigidity of the core domain has also been postulated to explain the relative lack of temperature dependence for wild type and S392E p53 proteins (Nichols and Matthews, 2001, 2002).

5.3.2 Implications for the activation model

The characterization of Δ33 p53 protein was designed to evaluate the activation model for p53 regulation. This model implied that p53 exists in two conformations (Cook and Milner, 1990; Hupp and Lane, 1994; Jayaraman and Prives, 1995; Milner and Medcalf, 1991). Under physiological conditions p53 was proposed to exist in a “latent” conformation that displays little, if any, DNA binding affinity. The native C-terminal regulatory domain was proposed to impose an inhibitory effect on the binding by the central core. Covalent and non-covalent modifications of the C-terminal domain in vitro that are relevant as downstream effects of DNA damage pathways suggested that a second conformation of p53, with activated DNA binding properties, elicited the biological responses (Gu and Roeder, 1997; Hupp and Lane, 1995; Hupp et al., 1992, 1995; Jayaraman and Prives, 1995; Selivanova et al., 1996, 1997). According to this model, S392E p53 protein was perceived as a protein in the “active” conformation, since this mutant resembles the activating phosphorylation at S392E by CKII (Hao et al., 1996; Hupp and Lane, 1995; Hupp et al., 1992; Kapoor and Lozano, 1998). More importantly, a C-terminal deletion p53 mutant was used as the cornerstone of the activation theory, as this protein was shown in vitro to exhibit activated DNA properties, explained by the absence of the proposed inhibitory native C-terminal domain (Hupp et al., 1992). Therefore, selection of Δ33 p53 for systematic thermodynamic characterization would provide relevant information to evaluate
the proposed activation model. Expression and purification from bacterial cells that lack posttranslational modification machinery ensure that the protein is purified in an unmodified state. Also, the proposed constitutive activated phenotype of this protein simplifies the DNA binding studies, as no activation factors (covalent or non-covalent) are required.

The principal characteristic of activated p53 would be an increased DNA binding affinity as compared to native “latent” wild type protein. The DNA binding studies of Δ33 p53 found a small difference compared to wild type p53 protein insufficient to constitute an “activated” phenotype. However, the specificity for Δ33 p53 is somewhat higher than that of wild-type p53 protein. Although there was a small but distinct difference between the ion concentration profiles of Δ33 p53 and wild type p53, this variation does not support a fundamentally distinct mode of DNA binding. Indeed, both proteins appear to bind DNA specifically by making significantly fewer ion pairs than predicted from the structure of the compact predominantly β-sheet domain (Cho et al., 1994). Although the β-sheet structure, presumably corresponding to the core domain, was destabilized by the presence of the unmodified C-terminal regulatory domain, this observation is probably irrelevant to the activation model because it occurred at temperatures outside the physiological range. Therefore, the results for Δ33 p53 clearly disfavor the existence of an activation mechanism mediated by the C-terminal regulatory region.

The studies presented are concordant with recent results that also call the activation model into question. First, elucidation of the MDM2-dependent degradation pathway provided insight into the mechanisms for cellular stability of wild type p53, a concept often inappropriately extended to apply to thermodynamic stability (Haupt et al., 1997; Kubbuttat et al., 1997; Midgley and Lane, 1997). In vivo, C-terminal alternatively spliced p53 reduced the apoptotic capacity of wild type p53 (Almog et al., 1997, 2000). In addition, although the presence of the C-terminus was
required, modifications were not necessary for wild type p53 to bind efficiently to chromatin-like DNA (Espinosa and Emerson, 2001; Kaeser and Iggo, 2002). In fact, acetylation of p53 did not increase DNA binding but induced protein coactivators (Barlev et al., 2001). A recent NMR study showed no structural changes between dimeric p53 proteins in the presence or absence of the C-terminal regulatory region (Ayed et al., 2001). Further, alterations of p53 DNA binding activities, perceived as activation in in vitro studies, were the result of inhibition by non-specific double stranded DNA fragments that had been commonly used in large excesses as carrier DNA (Anderson et al., 1997). Notably, even in a large excess of non-specific DNA, p53 was capable to bind the naturally occurring mdm2 promoter, without the need of any activation (Kaku et al., 2001). Characterization of the DNA binding properties for wild-type and S392E p53, which would be considered p53 in active conformation, provided evidence against the activation model, since wild type p53 bound DNA with nM affinity, in the same range as that observed for S392E p53 protein (Nichols and Matthews, 2001, 2002). The results from the studies in this chapter further add to the evidence against the activation model, as C-terminal deletion, which provided pivotal evidence in favor of the activation model, does not generate any major differences from wild type p53 with regard to the overall DNA binding properties.

Challenging the activation model, however, raises the question of how p53 does elicit the responses in terms of downstream gene expression. Notably, uncontrolled p53 expression is deleterious for the cells, as demonstrated by the p53-dependent lethality observed for mdm2−/− mice (Jones et al., 1995; Montes de Oca Luna et al., 1995). Functional MDM2 ensures that p53 is maintained at appropriately low and presumably non-lethal levels (Haupt et al., 1997; Kubbattat et al., 1997). Accumulation of p53 due to reduction of its degradation when proper conditions dictate (for instance, DNA damage), can have a dual effect: (a) as protein levels increase p53 forms tetramers, which is the oligomerization form with the most efficient DNA
binding, and (b) tetramerization occludes a nuclear export sequence located in the tetramerization domain, and subsequently p53 is retained in the nucleus (Stommel et al., 1999). Further, Sakaguchi et al. (1997) found that phosphorylation at S392E, an event that elicited as part of the DNA damage response, stabilized p53 tetramers by a 10-fold factor. In addition to the "crude" accumulation response, the function of p53 can be further fine-tuned by protein-protein interactions between p53 and protein transcription activators. These interactions can have a dual effect: (a) inhibitory, such as MDM2 binding to the N-terminus and masking binding sites of basal transcription factors (Oliner et al., 1993), or (b) activating, such as the protein factor Sp1 that is required for BAX transcription (Thornborrow and Manfredi, 2001). Therefore, a wide range of external factors can contribute to the response of p53, without the need of an "activation" model.

To summarize, characterization of the DNA binding properties of the Δ33 p53 variant, was designed to examine the effect of the C-terminus. Hupp et al. (1992) showed that deletion of the C-terminus "activated" DNA binding of the native protein, suggesting an inhibitory effect by this domain. This concept has been the basis of the "activation" model for p53 regulation. In this chapter, the DNA binding properties of Δ33 p53 were carefully determined. In contrast to the earlier study by Hupp et al. (1992), the overall DNA binding properties of this variant did not substantiate "activated" behavior of Δ33 p53 compared to the native protein. This deviation can in fact be attributed to the presence of large excesses of non-specific DNA that had been used in the earlier studies (Anderson et al., 1997). The present study is not the first to disfavor the activation model of p53 regulation. However, since Δ33 p53 lacks the putative C-terminal regulatory domain crucial to the activation model, characterization of its DNA binding properties provides a straightforward test of the validity of the activation hypothesis.
Chapter 6

Conclusions

Biological responses to various environmental stimuli engage effector molecules or mechanisms that alter the protein-DNA interactions of transcription factors. To understand the thermodynamic basis of these interactions it is crucial to provide an ultimate description of these biological responses. To this end, a complete thermodynamic characterization of ligand interactions for PurR (guanine and DNA) and comparison to its homologue protein LacI was undertaken. Using a glutamate-based buffer that presumably resembles intracellular ionic conditions (Leirmo et al., 1987), the thermodynamic and kinetic parameters for guanine and DNA binding of PurR were determined. In the first approach, specifically selected conditions allowed determination of thermodynamic parameters of individual binding events.

The studies presented in Chapter 3 demonstrated increased DNA binding affinity by ~500-fold for PurR holorepressor (presence of guanine) compared aporepressor (absence of guanine). Guanine complexed to PurR reduced dissociation of DNA from the holorepressor complex. This finding suggests that the in vivo repression of purine biosynthetic genes by PurR when intracellular purines are abundant relies on the stabilization of the holorepressor PurR complex with the DNA operator sequences. Moreover, guanine binding to PurR is increased by ~15-fold in the presence of DNA, suggesting that if aporepressor-DNA complex does form, guanine
stabilizes the holorepressor-DNA complex, ensuring repression. In the second approach, the entire set of binding data was analyzed by global fitting to a complete thermodynamic cycle model. Although more complicated in implementation, global analysis showed the presence of a small degree of anticooperativity for the second guanine binding site in the presence of DNA, a finding explained by structural steric effects in the holorepressor-DNA complex. Nevertheless, both approaches were consistent. Equilibrium and kinetic parameters were also obtained for the homologous LacI repressor under the same conditions, so that comparison between the PurI and LacI became feasible. LacI bound DNA with a ~30 fold lower affinity than that for PurR, indicating tighter repression in the pur regulon. This difference is attributed primarily to the lower dissociation rate from DNA by PurR holorepressor.

The differences in the thermodynamic parameters between these two homologous repressor proteins are rationalized by the distinct biological functions for the two proteins. The lac operon encodes gene products that utilize lactose as an energy source, a process that provides energy to the cell. In contrast, the energetically costly biosynthetic processes controlled by PurR require the cells to commit energy. Therefore, a tighter repression is required of PurR. Relief of repression is faster for LacI than for PurR. Lactose utilization can be viewed as an opportunistic event, and therefore the cell is required to respond quickly. In contrast, purine biosynthesis is associated with cell homeostasis, a process that would be expected to integrate signals over a period of time before response. Therefore, the cells can distinguish their behavior to different environmental stimuli, in a manner dependent on the distinct intrinsic thermodynamic DNA binding properties of those two proteins, despite their homology.

Two phenomena contribute favorably to the intrinsic thermodynamic DNA binding properties of proteins interacting with DNA: the release of cations from the highly negatively charged DNA backbone displaced by basic protein side-chains (poly-
electrolyte effect) and the release of water molecules from exposed apolar protein surfaces (hydrophobic effect). These effects can be monitored by ion concentration and temperature dependence of DNA binding affinities, respectively. The studies undertaken in Chapter 4 examined these effects on the DNA binding of PurR and LacI proteins.

In their high affinity forms, PurR holorepressor binding to DNA involved \(\sim 15\) ion pairs, as compared to only \(\sim 6\) for LacI. However, in their low affinity forms both proteins displayed a similar \(\sim 2\)-fold reduction in the ion pairs formed upon protein-DNA complexation. The differences in their ion concentration behavior can be rationalized by examining candidate ion contacts in the high-resolution structures of PurR holorepressor and LacI complexed with DNA (Bell and Lewis, 2000; Schumacher et al., 1997). For the high affinity complex form, LacI engages ion pairs formed between the helix-turn-helix or core domain and the DNA backbone. In contrast, the holorepressor PurR engages additional ion pairs formed between the hinge helix and DNA backbone and possibly intersubunit ion pairs between the N-terminus and the core domain.

van’t Hoff analysis of the temperature dependence of DNA binding showed a negative heat capacity change of larger magnitude for PurR holorepressor than that for LacI. Although additional factors related to electrostatic interactions can be contributing to the hydrophobic effect, especially in the case of the PurR, the differences between PurR and LacI may relate to the more extensive folding of the hinge helices observed for the PurR. These distinct behaviors of the highly homologous LacI and PurR proteins illustrate an increasingly apparent phenomenon: significant plasticity in protein structure and function. The differences between these two repressor proteins demonstrates clearly that similar folds, even those with very high levels of sequence similarity, may undergo subtle structural shifts that nonetheless substantially alter thermodynamic behavior and allow them to perform distinct functions.
The studies undertaken in Chapter 5 examined the polyelectrolyte and the hydrophobic effect in p53 DNA binding. Ion concentration dependence experiments showed a minimal ionic dependence for Δ33 (a mutant missing the last 33 carboxy-terminal amino acids), S392E and wild-type p53 proteins. Binding of p53 proteins to their consensus DNA sequence indicated formation of 0.5 (wild-type and S392E) to 3 (Δ33 p53) ion pairs. Based on the available structure of the core DNA binding domain complexed with DNA, these numbers indicated that p53 engaged ion pairs contributed by one, or possibly two subunits, a finding that contrasts the tetrameric assembly of the protein as indicated by earlier studies and the cross-linking and mobility shift experiments undertaken in this thesis. These results can be rationalized by the notion that not all four monomers of the p53 tetramers contact the DNA, but some of them provide structural support by protein-protein interactions. Alternatively, since the overall structure of the p53 protein is unknown, the only available structure of the core DNA binding domain might not describe in detail the molecular structure of the whole protein. Notably, the structure proposed by Cho et al. (1994) contains DNA in B-form, whereas it has been demonstrated experimentally that p53-DNA binding of the core domain is accompanied by a bend in the DNA of at least 30°, depending on the DNA sequence (Nagaich et al., 1997a). Clearly, the structure of the entire protein is required for further analysis.

Despite the lack of an overall p53 structure, important information can be derived indirectly from the biochemical characterization of the Δ33 p53 protein. Thermal denaturation indicated circular dichroism behavior and enhanced thermostability similar to that for S392E p53 protein. These results suggest that the presence of the unmodified C-terminal regulatory region destabilizes the structure of the central domain. The central core of the Δ33 p53 retains its predominantly β-sheet structure, confirming that this domain forms a very rigid structure, a notion that is supported by structural data (Cho et al., 1994). Comparison of the highly homologous free
murine core domain to the human p53 core domain complexed with DNA revealed minimal structural change (Zhao et al., 2001). Therefore, the modest heat capacity change derived from van't Hoff analysis (-0.7 to -1.0 kcal/molK) is consistent with the lack of dramatic structural changes.

The study of Δ33 p53 was undertaken in part to test the allosteric model of regulation, as data were accumulating that called this model into question. This mutant protein was one of the fundamental pieces of evidence in favor of a conformational activation model, since Δ33 p53 was proposed to be constitutively active (Hupp et al., 1992). In the present study, the ~5-fold increase in specific DNA binding affinity in potassium phosphate buffer observed for Δ33 relative to wild-type p53 does not substantiate the allosteric model for "activation". Furthermore, under the conditions utilized for the ion concentration dependence experiments, Δ33 did not display significantly altered behavior to that found for wild type p53. Although, Δ33 displays somewhat different characteristics from S392E p53 protein (previously proposed "activated" mutant) with regard to specific and nonspecific DNA binding, may result with both proteins disfavor the activation model. In vivo the differential functionality of p53 may be a result from DNA-dependent (for instance, DNA conformation) (Kim et al., 1997) or protein-dependent factors (for instance, differential recruitment of transcription activation factors) (Thornborrow and Manfredi, 2001).

To summarize, the present thesis explored the thermodynamic description of protein-DNA interaction with an emphasis on the polyelectrolyte and hydrophobic effects in the homologous proteins LacI and PurR from E. coli and human p53. Both LacI and PurR — despite their distinct functions — share a common structural characteristic: both proteins employ varying degrees of protein folding occurring upon DNA binding, and this property is altered upon the allosteric binding of small effector molecules. In contrast, human p53 binds DNA with its relatively compact core domain and with significantly lower specificity compared to the prokaryotic LacI and
PurR. DNA binding affinities of the various p53 proteins alone are not sufficient to provide a complete understanding of the cellular responses from a thermodynamic standpoint. In contrast to the apparent more simple thermodynamic mechanism of prokaryotic gene regulation, eukaryotic cells employ additional complex mechanisms, mainly at the post-translational level (for instance, modifications and/or protein-protein interactions), to generate the required levels of specific gene regulation.
Chapter 7

Appendix

7.1 Modelling of urea denaturation of tetrameric LacI

7.1.1 Introduction

Protein stability of the oligomeric LacI can be evaluated by chemical denaturation. Increasing concentrations of the chemical denaturant results in structural changes that can be monitored by biophysical methods. More specifically, denaturation of lactose repressor causes a decrease in both the fluorescence signal and the helical content in the CD signal. In fact, both of those methods have been employed to derive valuable thermodynamic information for monomeric, dimeric and tetrameric lactose repressor mutant proteins.

In the previous studies, estimation of the thermodynamic stability depended on the estimation of the fraction of unfolded protein \( f_u \) as modelled by simple transitions. Modelling of the unfolding process was restricted to the use of closed-form solutions of the unfolded species. This section describes extension of the data analysis to tetrameric lactose repressor. The model for the unfolding transition is not described by a closed-form solution for the unfolded species. Application of this
model in unfolding data in was feasible due to the programming flexibility of the nonlinear regression analysis software NONLIN (Johnson and Frasier, 1985).

### 7.1.2 Models of urea denaturation of lactose repressor

Thermodynamic parameters for reversible protein unfolding by chemical denaturants can be described by linear dependence of the free energy of the unfolding ($\Delta G$) on the concentration of the denaturant ($[D]$) according to Pace et al. (1990):

\[
\Delta G = \Delta G^0 + m[D] \tag{7.1}
\]

where $\Delta G^0$ denotes the free energy of unfolding in the absence of denaturant and $m$ denotes a parameter that describes the denaturant concentration dependence of the free energy of unfolding. Therefore, the equilibrium constant for the unfolding reaction can be expressed as

\[
K = K^0 + e^{-\frac{m[D]}{RT}} \tag{7.2}
\]

where $K^0$ represents the unfolding equilibrium constant in the absence of denaturant.

A simple reversible protein unfolding reaction

\[
F \rightarrow U
\]

can be monitored by the fraction of unfolded protein ($f_u$) as derived from the protein mass conservation equation:

\[
f_u = \frac{K}{1 + K} \tag{7.3}
\]

In the case of the unfolding of dimeric lactose repressor

\[
P_2 \rightarrow 2U
\]
the fraction of unfolded protein \( f_u \) can be derived from the protein mass conservation equation, as presented by the following equation in its numerically stable form:

\[
    f_u = \frac{2K}{\sqrt{K^2 + 8KP_T} + K}
\]

(7.4)

where \( P_T \) represents the total protein concentration in monomer subunits. Of note, \( f_u \) is a function of protein concentration, and as a result chemical denaturation exhibits dependence on the protein concentration of the unfolding protein.

In both simple transition or dimeric unfolding models, data analysis is straightforward since \( f_u \) can be expressed as a closed-form solution from the protein mass conservation equations. However, unfolding of tetrameric lactose repressor according to the following non-cooperative model

\[
P_4 \rightarrow 4U
\]

cannot be analyzed directly, because there is no closed-form solution for \( f_u \). Data analysis for this model can be conducted by incorporating a numerical solver that solves for \( f_u \) according to the following equation, as derived from the protein mass conservation equation:

\[
    \frac{4f_u^4P_T^3}{K} + f_u - 1 = 0
\]

(7.5)

Programming flexibility of NONLIN allows incorporation of the numerical solver within the fitting algorithm.

### 7.1.3 Application

Jennifer Barry applied the unfolding model of tetrameric lactose repressor in her studies of lactose repressor stability. Urea denaturation data showed a protein concentration dependence that was modelled best by a tetrameric unfolding model
with participation of a tetrameric intermediate. Therefore, denaturation of tetrameric lactose repressor follows two steps: (i) an intramolecular unfolding of the less stable individual core monomeric subunits that are held together at the dimer-dimer interface, and (ii) unfolding of the more stable dimer-dimer interface to produce unfolded monomers.

**Figure 7.1: Tetrameric dissociative unfolding**

Protein unfolding of LacI was simulated as a function of urea and protein concentration in log units. The simulation was conducted according to the tetrameric dissociative model. Urea concentration range was from 0 to 6 M, whereas monomeric protein concentration range was $10^{-7}$ to $10^{-6}$ M (-7 and -5 in log units, respectively). Thermodynamic parameters of unfolding $m$ and $\Delta G^0$ for equation 7.2, were set at 9.2 and -49.1 kcal/mol, as described in Barry and Matthews (1999b). The blue, cyan and red lines represent contour lines for 10, 50 and 90% unfolding, respectively. Protein unfolding is concentration dependent as indicated by the shifting of the contour lines to higher urea concentration, as protein concentration increases.
7.2 Macroscopic and microscopic equilibrium constants

Binding of a ligand $L$ to a protein $PP$ that contains two independent ligand binding sites

$$PP + L \xrightleftharpoons[K_1]{K_2} PPL + L \xrightleftharpoons[LPPL]{LPPL}$$

can be described by the macroscopic equilibrium constants $K_1$ and $K_2$ for binding at the first and second site respectively. The protein mass conservation equation expressed as a function of the macroscopic equilibrium constants is expressed as:

$$P_T = [P] + K_1[P][L] + K_1 K_2[P][L]^2$$  \hspace{1cm} (7.6)

The macroscopic constants do not distinguish the individual binding events. Binding of ligand to the independent binding sites can be distinguished by the use of the microscopic binding equilibrium constants, which represent binding of each ligand binding site:

$$PP \xrightleftharpoons[k_1]{k_2} LPP \xrightleftharpoons[k_2]{k_1} PP + L \xrightleftharpoons[k_1 k_2]{k_1 k_2} LPPL$$

In this case, the protein mass conservation equation is expressed as:

$$P_T = [P] + (k_1 + k_2)[P][L] + k_1 k_2[P][L]^2$$  \hspace{1cm} (7.7)

Since Equations 7.6 and 7.7 are equivalent, then the coefficients of each equa-
tion should be equal, thus:

\[ k_1 + k_2 = K_1 \quad \text{and} \quad (7.8) \]
\[ k_1 \times k_2 = K_2 \quad \text{and} \quad (7.9) \]

If both ligand binding sites are equivalent \((k_1 = k_2 = k)\), then the macroscopic and the microscopic equilibrium binding constants are correlated according to the following equations:

\[ K_1 = 2k \quad \text{and} \quad (7.10) \]
\[ K_2 = \frac{k}{2} \quad \text{and} \quad (7.11) \]
Bibliography


Spolar, R. S. and Record, M. T., Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. Science 263, 777–784.


